UNIVERSITY OF TRENTO - ITALY

International PhD Program in Biomolecular Sciences

Centre for Integrative Biology (CIBIO)

XXIXth Cycle

Prebiotic Synthesis of Redox-Active Iron-Sulfur Clusters

Author:

Claudia BONFIO

PhD Supervisor:

Prof. Sheref S. MANSY

Armenise-Harvard Laboratory of Synthetic and Reconstructive Biology

Centre for Integrative Biology (CIBIO)

Academic Year 2015 – 2016

In memory of my friend Maddalena.

Declaration of Authorship

I, Claudia BONFIO, declare that this thesis titled "Prebiotic synthesis of redox-active iron-sulfur clusters" is my own work and the use of all material from other sources has been properly and fully acknowledged.

I confirm that:

- o This work was done while in candidature for a research degree at this University;
- o Where I have consulted the published work of others, this is always clearly attributed;
- o Where I have quoted from the work of others, the source is always given;
- o I have properly and fully acknowledged all main sources of help;
- o Where the work was done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Abstract

Iron-sulfur clusters are indispensable to extant metabolism and are thought to have had an ancient role in mediating the chemical reactions that led to life. However, there has been no clear proposal for how these inorganic clusters came to occupy such an important position in biology. In this thesis I describe my efforts in delineating a plausible path from short, prebiotically plausible peptides to longer sequences with similar features to modern day iron-sulfur proteins.

Small organic thiolates and short cysteine-containing peptides can give rise to [2Fe-2S] and [4Fe-4S] clusters in aqueous solution when irradiated with UV light in the presence of iron ions. Additionally, duplications of tripeptides coordinated iron-sulfur clusters give sequences which are better able to stabilize iron-sulfur clusters, resembling motifs with cysteinyl ligand spacing highly similar to contemporary ferredoxins. Moreover, the studied iron-sulfur clusters are redox active and are able to mimic extant metabolic pathways, such as the first step of the electron transport chain, within protocells favouring the formation of a proton gradient which could be exploited for central biosynthetic processes.

Thesis Organization and Contribution

This PhD Thesis summarizes my four years of work on the importance of iron-sulfur clusters in the research on the origin of life. Since iron-sulfur proteins may have evolved from simple iron-sulfur peptides, my first project was focused on the synthesis and the characterization of iron-sulfur clusters coordinated by peptides of different length (Chapter 2). I focused then on the identification of a prebiotically plausible synthesis of iron-sulfur clusters under conditions compatible with the synthesis of amino acid, nucleotide, sugar and lipid precursors (Chapter 3). My last project was focused on determining the catalytic activity of prebiotic iron-sulfur clusters on relevant metabolic reactions (Chapter 4). Even though all the Chapters are interconnected and contribute to the aim of understanding the role of iron-sulfur clusters on the early Earth, each project could be viewed as a separate part of the full story, so that I have chosen to dedicate a separate Chapter to each different topic.

Chapter 1

Chapter 1 is a general introduction on the origin of life and the research studies aimed at understanding how live emerged on the early Earth. For the sake of keeping the introduction as general as possible, the specific questions of this Thesis are not introduced here. Claudia BONFIO wrote this Chapter.

Chapter 2

Chapter 2 is focused on the synthesis and characterization of iron-sulfur clusters coordinated by L-glutathione and its polymers in aqueous solution. All the wet-lab experiments concerning L-glutathione iron-sulfur clusters have been performed in collaboration with Dr. Simone SCINTILLA (University of Trento). The synthesis and characterization studies of iron-sulfur clusters coordinated by L-glutathione polymers have been accomplished by Claudia BONFIO. NMR studies were performed by Dr. Simone SCINTILLA. Structure calculations were completed by Dr. Luca BELMONTE and spectral decomposition and statistical analyses by Dr. Michele FORLIN. All the experiments were conceived by Dr. Simone SCINTILLA, Claudia BONFIO and Prof. Sheref MANSY. Claudia BONFIO wrote this Chapter. Chapter 2 has previously been published with minor changes as the following publication: Scintilla, S., Bonfio, C., *et al.* Duplications of an iron-sulfur

tripeptide leads to the formation of a protoferredoxin. *Chem. Commun.* 52, 13456-13459 (2016).

Chapter 3

Chapter 3 is focused on the UV light-driven prebiotic synthesis of iron-sulfur clusters. This Chapter is the result of collaboration with Prof. David J. EVANS (Dept. of Chemistry, University of Hull, Hull, UK), Prof. Jack W. SZOSTAK (Howard Hughes Medical Institute, Boston, MA, US), Prof. Dimitar D. SASSELOV (Dept. of Astronomy, Harvard University, Cambridge, MA, US), and Prof. John D. SUTHERLAND (MRC - Laboratory of Molecular Biology, Cambridge, UK). All the experiments have been performed by Claudia BONFIO, except for the Mössbauer analyses. All the experiments were conceived by Claudia BONFIO and Prof. Sheref MANSY. Claudia BONFIO wrote this Chapter. Supporting materials for Chapter 3 appear in Appendix 1. Chapter 3 has previously been published with minor changes as the following publication: <u>Bonfio, C.</u>, *et al.* UV light-driven prebiotic synthesis of iron-sulfur clusters. *Nat. Chem.* (2017). doi:10.1038/nchem.2817.

Chapter 4

Chapter 4 is focused on the evaluation of the catalytic activity of iron-sulfur clusters. This Chapter is the result of a collaboration with Dr. Maddalena CORSINI (Dept. of Chemistry, University of Siena, Siena, IT), and Prof. Graziano GUELLA (Dept. of Physics, University of Trento, Trento, IT). All the experiments have been performed by Claudia BONFIO. All the experiments were conceived by Claudia BONFIO and Prof. Sheref MANSY. Claudia BONFIO wrote this Chapter. Supporting materials for Chapter 4 appear in Appendix 2.

List of Publications

- <u>Claudia Bonfio</u>, Elisa Godino, Maddalena Corsini, Fabrizia Fabrizi De Biani, Graziano Guella, John D. Sutherland, Sheref S. Mansy - **Prebiotic redox-active iron-sulfur peptides generate a proton gradient across model protocells**. (Under review).

- Sambuddha Sen, <u>Claudia Bonfio</u>, Sheref S. Mansy, James A. Cowan - **Investigation of glutathione-derived electrostatic and hydrogen bonding interactions and their role in defining Grx5 [2Fe-2S] cluster stability, optical spectra and transfer chemistry**. (Under review).

- <u>Claudia Bonfio</u>, Sheref S. Mansy - **The chemical roots of iron-sulfur dependent metabolism**. *Biochemistry* (2017), DOI: 10.1021/acs.biochem.7b00842 (invited article).

- <u>Claudia Bonfio</u>, Luca Valer, Simone Scintilla, Sachin Shah, David J. Evans, Lin Jin, Jack W. Szostak, Dimitar D. Sasselov, John D. Sutherland, Sheref S. Mansy - **UV Light-driven prebiotic synthesis of iron-sulfur clusters**. *Nature Chemistry* (2017), DOI: 10.1038/nchem.2817.

- Simone Scintilla, <u>Claudia Bonfio</u>, Luca Belmonte, Michele Forlin, Daniele Rossetto, Jingwei Li, James A. Cowan, Angela Galliani, Fabio Arnesano, Michael Assfalg, Sheref S. Mansy - **Duplications of an Iron-Sulfur Tripeptide Leads to the Formation of a Protoferredoxin**. *Chem. Comm.* (2016), 52, 13456-13459, DOI: 10.1039/C6CC07912A.

- Luca Belmonte, Daniele Rossetto, Michele Forlin, Simone Scintilla, <u>Claudia Bonfio</u>, Sheref S. Mansy - **Cysteine containing dipeptides show a metal specificity that matches the composition of seawater**. *Phys. Chem. Chem. Phys.* (2016), 18, 20104-8, DOI: 10.1039/C6CP00608F.

Acknowledgements

"I have just been lucky."

"No, Claudia, you have made your fortune."

On my way back home after my interview four years ago, I would have never imagined what this moment would have meant to me. I would have never been able to achieve this goal without the continuous support and guidance (and patience) of my advisor Sheref Mansy. Not only did he allow me to enter this new field of research on the origin of life, but he made me love it. And now, I cannot wait to start my new adventure in Cambridge (UK) in John Sutherland's group.

Sheref also gave me the great opportunity to spend some time in some of the best places in the world for scientific research. First, I felt in love with Cambridge (UK). Working on my project in John Sutherland's lab has been an overwhelming experience. John is a great person and a great scientist, always positive and passionate about research. A year later, I felt also in love with Boston (MA, US). Becoming part of Dimitar Sasselov and Jack Szostak's groups has allowed me to expand my interests towards a more interdisciplinary view of my research. With a bit of astrophysics and a bit of biology, I came back to Italy with an even greater passion for science.

I loved attending seminars and meeting people at international conferences working in my field. That's also how I met my thesis reviewers. Without them, and without many comments and suggestions that came from many other open-minded and generous colleagues, this work would have never seen the light.

My Ph.D. experience has not always been easy. At one point, I even considered abandoning it and looking for something else. Luckily, my family and my friends have always been ready to help me smile and move forward.

Many thanks to Betty and Anna for the continuous support for all of my weird requests; to Cristina, Allegra and Charlie, for the great time we spent together in Italy and abroad; to Elisa, for the energy she brought in the lab during my last year; to my bench-mate Silvia, for all the time spent together drinking Emma's cocktails and hot chocolate; to Francesca and Rosa, for our many rescheduled coffees; to Luisa, for the lunch she accepted to have with a stranger on the very first day of our Ph.D. journey; to my Trento lab-mates, in particular

Table of Contents

List of Abbreviations

Ac Acetyl

ATP Adenosine Triphosphate

E$_{cp}$ Cathodic Peak Potential

E° Standard Midpoint Potential

E°' Formal Midpoint Potential

FAD Flavin Adenine Dinucleotide

FMN Flavin Mononucleotide

ESI Electrospray Ionization

HPLC High Performance Liquid Chromatography

K$_a$ Acid dissociation constant

K$_D$ Dissociation constant

LUCA Last Universal Common Ancestor

MS Mass Spectroscopy

NAD Nicotinamide Adenine Dinucleotide

NADP Nicotinamide Adenine Dinucleotide Phosphate

NHE Normal Hydrogen Electrode

NMR Nuclear Magnetic Resonance

POPC 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine

SPPS Solid Phase Peptide Synthesis

UV Ultraviolet

UV-vis. Ultraviolet-visible

Chapter 1 - Introduction

The research on the origin of life

The fundamental question of how life emerged on Earth has always fascinated scientists and philosophers. Extant life mainly exploits six chemical elements (C, H, O, N, P, and S) to form biological macromolecules.[1] However, simple geochemical models and inferences from ancient sediments would suggest that the elemental composition of Earth's oceans and atmosphere underwent significant changes on geological time scales (Figure 1.1).[2]
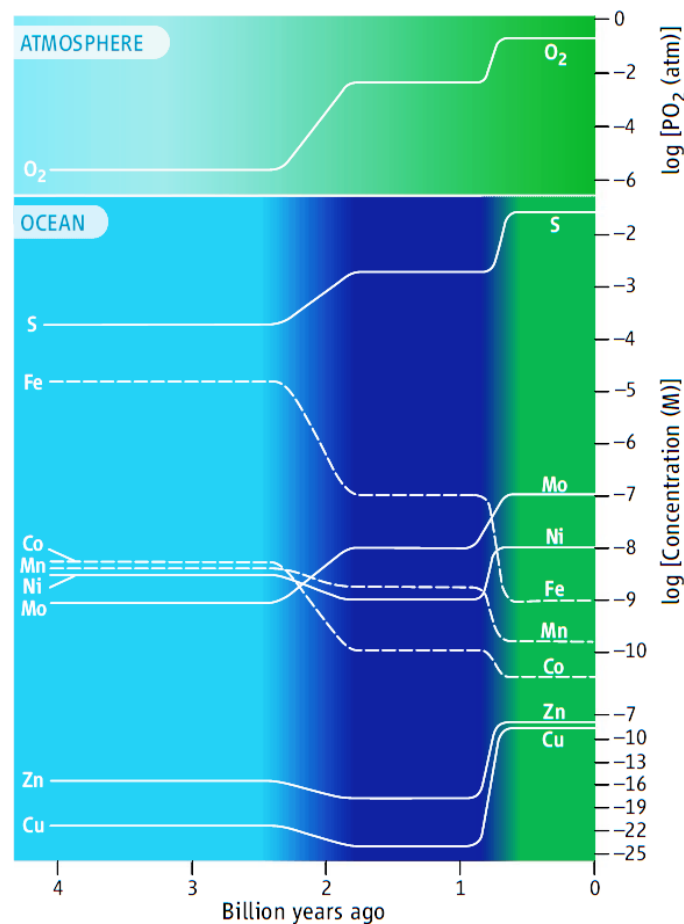
Figure 1.1 | Variations in elemental concentrations through time, based on geochemical models and extrapolations from ancient sediments. Colour gradients indicate a transition from sulfur-poor, anoxic oceans before 2.4 billion years ago (light blue region) to hydrogen sulfide-rich oceans between 1.8 and 0.8 billion years ago (blue region) to the complete oxygenation of the ocean (green region). Dashed lines are for elements with decreasing concentrations (adapted from Anbar A. D.[2]).

Alterations in the environmental availability of essential elements could have imposed selection pressures which would have changed the composition of biological macromolecules and shaped the emergence of life. However, revealing the historical flow of events that could explain how life arose may not be a realistic goal due to the lack of "chemical fossils"[3] and defined details of the physicochemical conditions present on primordial Earth. Thus, a systematic scientific approach emerged in the past years aimed at discovering plausible solutions for the prebiotic origins of various classes of biological building blocks and their respective biopolymers.[4] This challenging point of view represents the central open question for research on the origins of life: is it possible to demonstrate that life could emerge from the self-assembly of small molecules and networks of chemical transformations?[3] In order to find an appropriate answer, research on the origins of life requires interdisciplinary support on topics including the astrophysical context, the development of prebiotic chemistry, the assembly of the first cells, and the advent of Darwinian evolution. That is, the main challenge on the path to understand how life arose on Earth is to imagine an *ensemble* of molecules that is simple enough to form by self-assembly, yet sufficiently complex to display the essential properties of a living organism.[5]

There is little consensus on the definition of life. Life forms could be considered as self-replicating molecular systems able to progressively evolve.[6] Extant life is based on genetic polymers able to store information (nucleic acids), genome-encoded catalysts that facilitate self-replication (proteins), and compartmentalization systems capable of protecting their content and enabling evolution (membranes). A convenient way of representing how life emerged has been proposed by Robert Pascal, Addy Pross and John Sutherland[7] in the context of a graph with time or system complexity as the abscissa and degree of "aliveness" as the ordinate, where the starting equilibrium state represents the inanimate and the ending state represents the animate (Figure 1.2). The intermediate states would require a disequilibrium route which would correspond to innovation steps, such as RNA replication or vesicle division, and optimization steps, such as mutation or selection processes.[4] Thus, the main challenge for a chemical approach to the research on the origins of life would be to understand and reproduce all the steps leading to the transition from prebiotic chemistry to protobiology.[3] However, aliveness is admittedly an arbitrary and ambiguous term, and attempts to define life may appear irrelevant to scientific efforts to understand the origin of life.[8]
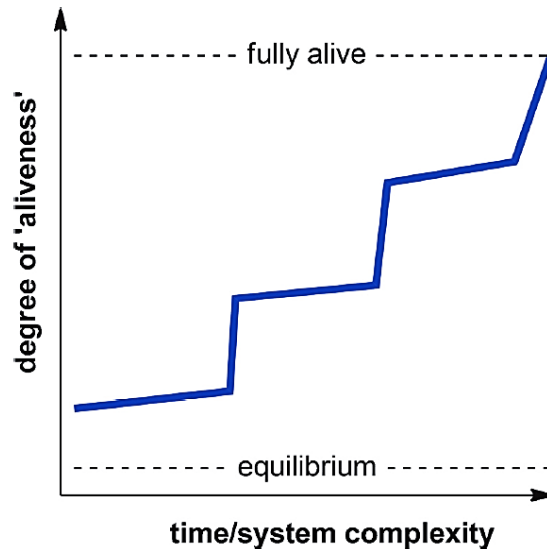
Figure 1.2 | Evolution from the inanimate to the animate state (adapted from Sutherland J. D.[4]).

1.1 First studies on the origins of life

Cells were first discovered in the 17th century. However, biologists realized that all living things are made of cells only in the 19th century.[9] In that period, vitalism was the most commonly accepted credence, sustaining that living things were endowed with a magical property that could make animate organisms different from inanimate objects. However, in 1828 Friedrich Wöhler reported the synthesis of urea from ammonium cyanate, demonstrating for the first time that biological molecules could be made from simpler, "non-living" molecules.[10]

Aleksandr Oparin and John Haldane were among the first scientists who got deeply interested in studying how life emerged on Earth. They independently recovered Charles Darwin idea from 1871 concerning a "warm little pond" as the potential prebiotic environment in which a mixture of inorganic salts and simple organic molecules formed more complex compounds through electricity-, heat- or light-driven reactions. In the Oparin-Haldane hypothesis, life formed in a "primordial soup" of simple chemicals that could be combined together to form modern macromolecules.[11,12] However, they did not provide experimental data to confirm their theory.

In 1952 Stanley Miller ran his well-known experiment, simulating the early Earth conditions by observing the effect of electrical discharges on a mixture of water, hydrogen, ammonia and methane.[13] The results he published in collaboration with his PhD advisor Harold Urey showed the formation of traces of amino-acids.[14] Even though later studies evidenced that the primordial atmosphere had a different mixture of gases, this experiment

3

has been considered the first fundamental experiment in the field of the modern research on the origins of life.[15]

After the discoveries of the roles of nucleic acids and proteins in cells and their interconnected activities, any explanation for the origins of life was required to show how the central dogma of molecular biology came into existence. Since RNA was structurally similar to DNA and could fold like proteins, potentially performing enzymatic activity, in 1968 Leslie Orgel suggested RNA as the primordial highly versatile living molecule at the basis of life.[16] He suggested that the crucial aspect of life, and thus the first that emerged, was the ability of a living system to reproduce itself. The discovery in 1982 of ribozymes, RNA sequences with catalytic activity, by Thomas Cech and colleagues seemed to confirm Orgel's hypothesis.[17] Few years later Walter Gilbert coined the expression "RNA world" to describe how RNA molecules led to the emergence of life.[18] Alternative candidates to RNA for the primordial replicator were proposed in the past years, obtained by changes in the backbone structure, such as with amide bonds in peptide nucleic acids[19] (PNAs), or composition, such as by substituting ribose with threose in threose nucleic acids[20] (TNAs) and in glycol nucleic acids[21] (GNAs).

However, throughout the years many researchers pointed out that other features of life seemed essential, such as metabolic pathways or compartments. An alternative view for the emergence of life was proposed in the late 1980s by Günter Wächertshäuser.[22] He envisioned the first metabolic cycle as a pull of chemical reactions taking place in a hot aqueous high-pressure environment enriched by volcanic gases and minerals. The discovery of less extreme alkaline "hydrothermal vents" in the deep oceans, made of porous rocks densely populated by bacteria and rich in iron-sulfur based minerals such as pyrite pushed the geologist Michael Russell[23] and later the biologist William Martin[24] to modify Wächertshäuser theory and propose a new hypothesis, the "iron-sulfur world", for describing the origins of life.

In the last decades, a third approach has come out, related to the importance of membranes. David Deamer and Pier Luigi Luisi independently suggested that, in order to have either a working metabolism or a self-replicating system, a way to keep all the essentials of life together, such as simple fatty acid vesicles, must have been formed first.[25–27] On the basis of the intrinsic chemical properties of lipids and their capacity to undergo spontaneous self-organization into vesicles, Daniel Segrè proposed a "lipid world" as an early evolutionary step in the emergence of life on Earth.[28] Recently, more unified approaches in the research on the origins of life are emerging, showing that RNA replication

and compartmentalization could happen simultaneously[29] or that the prebiotic conditions for the synthesis of amino acids and nucleotides could be the same.[5]

1.2 The "RNA world"

How could modern organism have arisen from a "prebiotic soup", since modern DNA depends on enzymes for replication and modern proteins depend on DNA for their genomic information? The postulation of an RNA world seems able to solve this famous chicken-and-egg problem (Figure 1.3). The RNA world theory is based on the idea that there was a period in the early history of life on Earth when RNA, or something chemically similar, was able to store genetic information and to carry out most of the replicative and metabolic processes needed for chemistry to be converted into biology.[6]
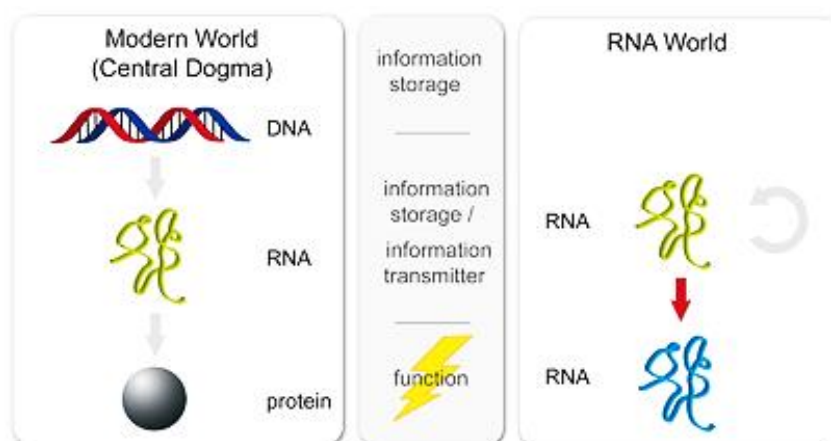

Figure 1.3 | Schematic representation of the "RNA world" concepts: on early Earth RNA would have been able to behave both as a protein and DNA (adapted from www.chemistryworld.com).

As previously mentioned, in the 1960s Orgel and other researchers recognized that RNA was a good candidate to explain the origins of life on Earth, since a system based on RNA would be a good prebiotic ancestor to the more complex triad DNA-RNA-proteins at the basis of modern molecular biology.[16] Discoveries of naturally occurring ribozyme catalysts,[17] with the demonstration that ribosomal RNA catalyses peptide bond formation[30,31] allowed the RNA world theory to become widely accepted and established. However, few arguments about the plausibility of the RNA world theory still remain unanswered. Could RNA have formed on the early Earth? Could RNA self-replicate and perform all the functions of life by itself?

The prebiotic synthesis of nucleotides from a prebiotically plausible mixture of compounds initially appeared challenging suffering from limited regio- and stereoselectivity

and detrimental side reactions. In recent years, progress has been made, for example advances towards a prebiotically plausible synthesis of nucleotides,[32] a ribozyme-mediated conversion of the 5'-hydroxy groups of RNA into 5'-triphosphates,[33] and the ribozyme-mediated polymerization of nucleoside triphosphates into RNA.[34] Indeed, in 2009 Sutherland and co-workers realized that the synthesis of nucleotides could be pursued following an alternative pathway to the sequential sugar-nucleobases-phosphate scheme. Few prebiotically plausible molecules could react together under fairly mild conditions forming activated nucleotides in the presence of metal ions and ultraviolet (UV) light and inorganic phosphate, that could behave as catalyst, reactant and buffer at the same time.[35] Moreover, being inspired by Orgel's pioneering work,[36] Jack Szostak showed that RNA activated monomers could assemble into polymers by using a template to orient and guide all the pieces together within lipid vesicles (Figure 1.4).[29] Regardless of the question of when cells originated and started to play a fundamental role in enhancing life metabolic and evolutionary potential, the coexistence of RNA-based reactions and lipid vesicles represented a great step forward towards the understanding of how life emerged.[37]
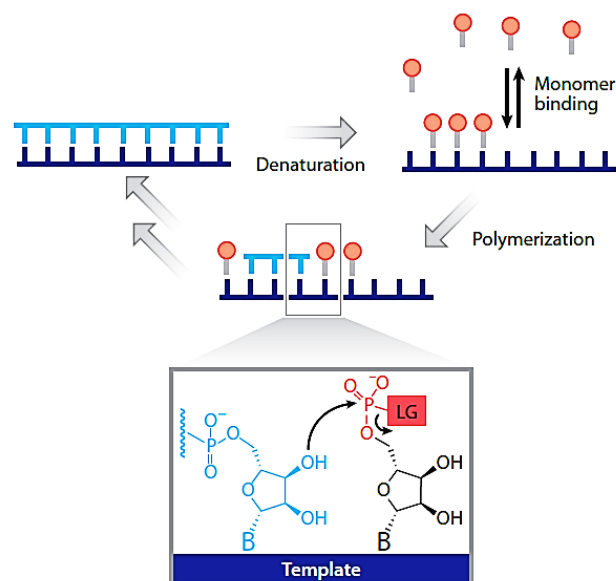

Figure 1.4 | In non-enzymatic nucleic acid self-replication, activated nucleotides or short oligonucleotides bind to a complementary base in a single-stranded template (B stands for nitrogenous Base, LG stands for Leaving Group) (adapted from Blain J. C. *et al.*[38]).

The ability of RNA to behave as an enzyme was explored even more. In 2001 David Bartel made an RNA enzyme called R18 that could add new nucleotides to a strand of RNA, based on an existing template, not just adding random nucleotides, but correctly copying a sequence.[39] Ten years later Philipp Holliger's group showed that a modified R18 could copy sequences up to 95 nucleotides long, which is 48% of its own length.[40] An alternative

approach was published in 2009 by Gerald Joyce, who created an RNA enzyme which could replicate itself indirectly. This enzyme joins together two short pieces of RNA to create a second enzyme, able to join together other two RNA pieces to recreate the original enzyme.[41] However, regarding the RNA world there is one problem that scientists have not focused on and solved yet. The first organism must have had some form of metabolism and no data have been shown so far on a RNA-based prebiotic metabolism.

1.3 The "metabolism-first" theory and the "iron-sulfur world"

Over the last century, chemists have shown many successful pathways for the prebiotic synthesis of amino acids, nucleotides and lipids. However, there is strikingly little similarity between this chemistry and the metabolic pathways of cells, in terms of substrates, catalysts, and synthetic pathways.[42] The alikeness of the metabolic networks of all living organisms highlight the great universality within the diversity of extant life. Perhaps modern metabolic pathways contain information about the primordial metabolic chemistry that existed on the early Earth. In the late 20th century, Harold Morowitz and colleagues built upon the idea of an autotrophic chemical network to provide a theory for the origin of metabolic biochemistry.[43] At the core of the metabolic network, Morowitz placed the (reductive) citric acid cycle, which represents the central driving force for all biosynthetic pathways. Indeed, sugars and amino acids can be synthesized by phosphoenol pyruvate and keto acids, respectively, and nucleic acids are derived therefrom. The crucial point of the whole metabolic pathway would be the autocatalysis of each individual synthetic step.[44] Moreover, the network would represent a sink for carbon-rich compounds, driving the synthesis of more complex molecules starting from carbon dioxide. According to the view of Morowitz, the origin of metabolism would have been, therefore, a prebiotic non-enzymatic reductive citric acid cycle eventually involving catalytic surfaces such as metal sulfides to drive biochemical reactions.[45] Indeed, modern metabolism is dependent on enzymes with metal ions in their active sites, so it seems plausible that metal ions and simple metal-complexes took part in some way to prebiotic metabolic-like reactions.

The importance of metals, in particular iron sulfides, was at the basis of Günter Wächtershäuser's theory, which suggests the involvement of iron pyrite in the emergence of life. In 1988 he proposed that simple prebiotic molecules such as carbon dioxide and formaldehyde could have been reduced to more complex organic molecules due to the low redox potential of the iron(II)/iron(III) sulfide systems. Since iron sulfide minerals were

thought to be quite abundant in hydrothermal vents on early Earth, the reactions happening on their surfaces could have represented the connection between geochemistry and biochemistry.[22] The involvement of minerals in the emergence of life is based on few concepts. Since the hydrolytic power of water would have easily destroyed any synthesized organic molecule and the "prebiotic soup" would have been rather diluted, it seems more plausible that high concentrations of chemicals were present in locations isolated from the ocean, such as small ponds, moist soil and hydrothermal sites, which would be subject to hydration-dehydration cycles and temperature gradients. In such environments, organic molecules would have been into close contact with mineral surfaces. The ability of mineral surfaces to adsorb organic molecules and act as a template to catalyse specific chemical reactions is widely recognized. Moreover, the adsorption of chemicals in the interlayers of minerals could have had a protecting role against degradation induced by light.[46]

Wächtershäuser based his theory on the formation of a mineral surface, that is pyrite (FeS$_2$) derived from iron monosulfide (FeS) and hydrogen sulfide (H$_2$S).

$$\text{FeS} + \text{H}_2\text{S} \rightarrow \text{FeS}_2 + 2\text{e}^- + 2\text{H}^+$$

In an acidic geochemical solution, pyrite is positively charged and can bind phosphates, thiolates and carboxylates, molecules which are essential in modern cellular metabolism. The first organism was then a chemoautotroph, able to use pyrite as a source of electrons and energy. Carbon dioxide would have been reduced resulting in organic molecules such as carbohydrates, sugars and amino acid precursors, while some metals, such as those related with modern protein functions, could have evolved into more complex system until modern biology appeared.[47,48]

A few years later, the geologist Michael Russell modified and amplified Wächtershäuser hypothesis of a primordial "iron-sulfur world". Since iron sulfides are dominant in hydrothermal vents, they could have acted on the primitive Earth both as catalysts and reagents. They could potentially drive dehydration, isomerization, cleavage reactions as well as oxidation and reductions and organic syntheses. Moreover, he realized that the chemical and temperature gradients present in these sites could have easily allowed the development of ancient forms of life.[49–51] Expanding and emphasizing Russell's ideas, William Martin's studies on the Last Universal Common Ancestor (LUCA), a hypothetical primordial organism[52,53] from which life on Earth is thought to have descended approximately 3.5 billion years ago, revealed that many ancient enzymes were iron-based enzymes (Figure 1.5).[24,54,55] Indeed, from an evolutionary point of view, proteins containing

iron-sulfur systems, called iron-sulfur clusters, such as ferredoxins and rubredoxins, are considered among the first enzymes appeared on early Earth and a plausible precursor for these proteins could have been the iron(II)/iron(III) sulfide system.[56,57]
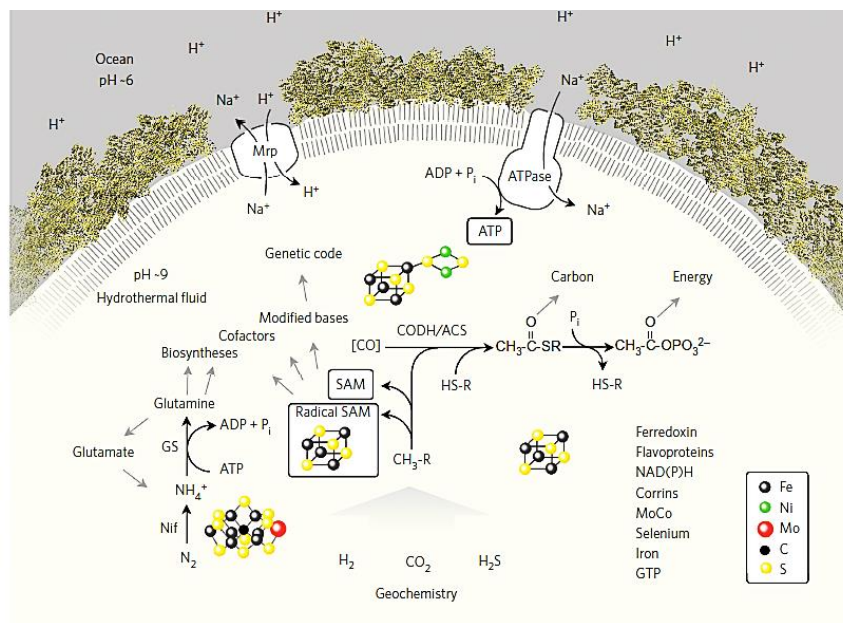

Figure 1.5 | Summary of the main interactions of LUCA with its environment, a vent-like geochemical setting, as inferred from genome data (adapted from Weiss M. C. *et al.*[55]).

Even though the iron-sulfur world theory is intriguing, it has mainly two unsolved problems. First of all, there is no, or little, experimental evidence for the processes that Russell and Martin describe. Many research groups are trying to demonstrate the plausibility of this theory, but still no valid data have been shown. The location of hydrothermal vents in the deep sea is the second problem, since it seems unlikely that long-chain organic molecules and biopolymers could have formed and been stable in water without undergoing hydrolytic degradation. However, some aspects of this theory must certainly be true, such as the fundamental role of metal ions. Indeed, Sutherland's chemical reactions involve the use of the redox couple copper(I)/copper(II) cyanide and the reducing power of hydrogen sulfide, and magnesium ions are crucial in Szostak's RNA polymerization.

1.4 Bottom up and top down approaches

In theory, the studies on the origins of life can go from geochemistry forward or from biology backward. Focusing on the bottom-up approach, astrophysics and planetary science suggest that a wide range of conditions and environments could have existed on the early Earth. Thus, a huge variety of chemistry could be plausible in submarine vents, or meteorites

impact craters, or in warm drying lagoons. However, this wide amount of possibilities means that there may not be sufficient constraints to choose one particular scenario and to explore the chemistry of each environment in a systematic way in order to discover how biomolecules could have favoured the rise from simple precursors. Regarding the top-down approach, in the 19th century Darwin proposed his Doctrine of Common Descent: "[P]robably all of the organic beings which have ever lived on this Earth have descended from some one primordial form…".[7] The origin of this communal life is presumed to have occurred on the early Earth at some point. In April 2016, scientists presented an updated version of the "tree of life", a kind of family tree for every living species. Almost all of the branches are bacterial and the shape of the tree suggests that a bacterium was the common ancestor of all life.[58]

However, phylogeny can help only partially to understand how chemistry evolved into biology. The conceptual and experimental studies on the reduction of cells to the minimal cell have led to systems based on minimal genetic, metabolic, catalytic and compartment-forming subsystems. However, such simplified assembly would require too evolved and sophisticated independent prebiotic biochemistry processes, which would interfere with each other.[4]

1.5 The holistic approach

Either considering the divergence between the bottom-up and the top-down approaches or between the metabolism-first and the RNA-first theories, the problem of the origins of life seems difficult to be solved by pointing to one scenario as the correct one. It was previously suggested in this Chapter that the origin and evolution of life as we know it should be seen as a stepwise dynamic process in terms of time and system complexity.[7] Theories suggesting a purely metabolic or genomic basis for the origin of life, and not considering the importance of cellular membranes, comes across several theoretical and experimental problems, as previously described.[59] Therefore, it seems more logical to consider that a wide variety of molecules were available on the early Earth and that different physicochemical processes and reaction mechanisms would have led through selection steps to compartmentalized primordial self-replicating systems sustained by metabolic networks.

A more realistic and holistic approach has begun to take hold over the last years, aimed at understanding how much different and compatible are the various chemistries, geological and astrophysical environments, and prebiotically plausible scenarios leading to

each subsystem. This new approach, also referred to as systems chemistry,[59] suggest to focus not on single reactions in isolation, but on the whole series of contemporaneous processes which shared common reactants, products and conditions to allow the emergence of life (Figure 1.6).[37,59]
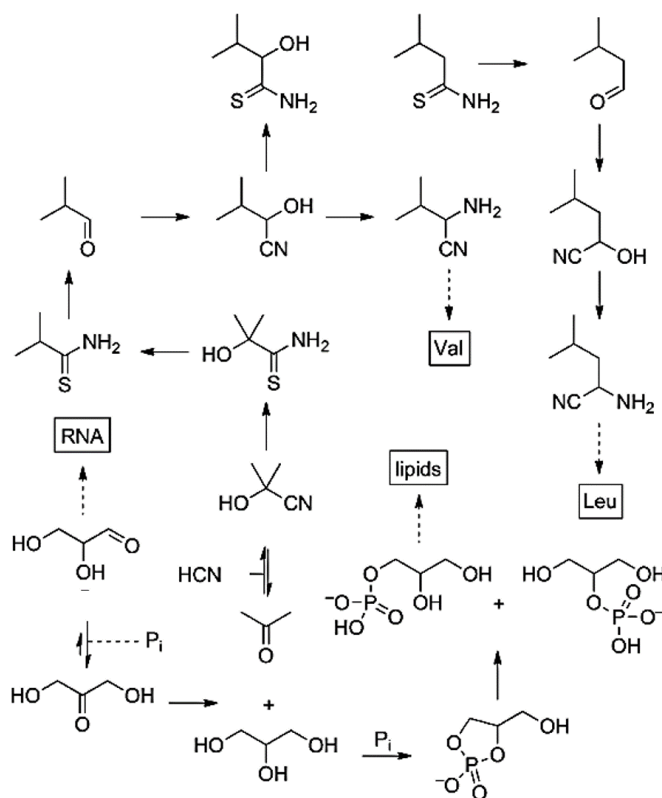

Figure 1.6 | First signs of a linkage between all subsystems through cyanosulfidic chemistry (adapted from Sutherland J. D.[4]).

According to the holistic approach, Sutherland and co-workers envisioned an iterative process in which chemical results would constrain geochemical scenarios and the proposed scenario would back-condition the chemistry, leading to refinements in both the chemistry and the scenario.[4] Indeed, finding reasonable geological environments for the origins of life is a critical aspect of understanding the whole pathway. Research on the origins of life is no longer restricted to the idea of life emerging from diluted oceanic soups or at deep sea hydrothermal vents. More recent studies on surface settings which could allow organic molecules to accumulate over time, so that prebiotic chemistry could be driven by UV light coming from the young Sun in very concentrated solutions in geothermally active areas, represent a great advance.[60]

1.6 The importance of metals

Modern life is strongly dependent on metal ions, which are able to catalyse metabolic reactions, to influence protein and RNA folding, to transduce signals, and to form concentration gradients used to drive unfavourable reactions. Therefore, it seems plausible that modern day metabolism has been partially built on prebiotic reactions accelerated by metal ions. Moreover, from an evolutionary point of view, metal ions coordinated by peptides on the early Earth would have represented a reasonable connection point between free metal ions and modern proteins, where the reactivity of the metal centre would have been modified in a sequence-dependent manner.[61]

The catalytic activity of free metal ions has been widely studied. For example, free divalent cations can catalyse the hydrolysis of ATP[62] and ferric ions in an aqueous environment can decompose hydrogen peroxide.[63] In 2014 Markus Ralser and co-workers showed that ferrous ions alone at elevated temperatures could catalyse the majority of the reactions of the glycolysis and the pentose phosphate pathways. Even though magnesium and zinc dependent enzymes are exploited for most of the steps of modern glycolysis, free iron(II) ions have been shown to be fundamental to recover RNA catalytic activity in the absence of magnesium ions.[64] Moreover, bioinformatic analyses of protein-metal binding motifs encoded in genomes revealed that, relative to prokaryotes, which are considered to be more ancient organisms, eukaryotes require more zinc, and less iron, manganese and cobalt.[2]

Metal ions such as iron, molybdenum and nickel can form catalytically active complex, inorganic clusters resembling modern-day protein active sites.[65] Indeed, the two additional pathways exploited by modern organisms to boost energy yields from the catabolism of sugars, proteins and lipids, are driven by more complex proteins. The citric acid cycle and the electron transport chain are mediated by iron-sulfur and heme-binding proteins, even though it is unclear whether this incongruence reflects differences in the environment from which the proteins emerged or in the evolutionary timeline.[61]

Moreover, inspection across the tree of life of proteins, which catalyse electron transfer revealed that the most abundant transition metal is iron, present in ~70% of oxidoreductases either as a free ion, bound to sulfide in iron-sulfur clusters, or as porphyrin-coordinated heme cofactor (Figure 1.7).[1]
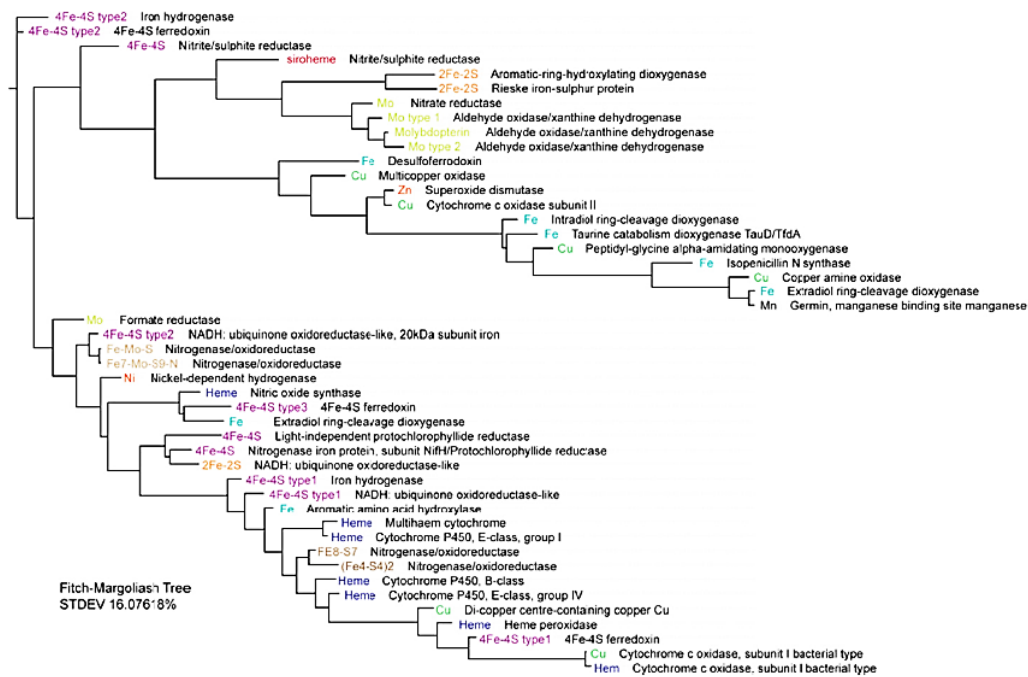
Figure 1.7 | A topologically derived phylogenetic correlation of oxidoreductases extrapolated from the structural analyses of the metal binding sites (adapted from Falkowski P. G.[1]).

Hence, the similarities between free metal ions and modern metal-based enzymes could represent the scientific starting point to experimentally explore the metal ion binding ability and catalytic activity of model prebiotic peptides to gain insight into how life as we know it could have emerged.

1.7 The importance of membranes

The idea that life emerged on the early Earth as self-assembled structures of organic molecules was first proposed by Oparin.[11] According to the "compartment first" theorists, the origin of life required the presence of at least three different components on the early Earth: a source of free energy, a plausible solvent, and organic molecules capable of self-assembly, such as lipids.[26] In particular, self-assembling compounds must have provided monomers for polymer synthesis and the formation of defined boundaries. Indeed, a prebiotic functional species must have had definite margins that separate the system itself from the surrounding environment.[25]

A fundamental question regarding self-assembling compounds concerns the nature of the lipid-like molecules able to self-assemble into vesicular structures which could be found on the early Earth (Figure 1.8). Even though phospholipids, the main component of extant cellular membranes, could have been synthesized later on during prebiotic chemical evolution, a wide variety of single-chain amphiphiles that are believed to have been readily

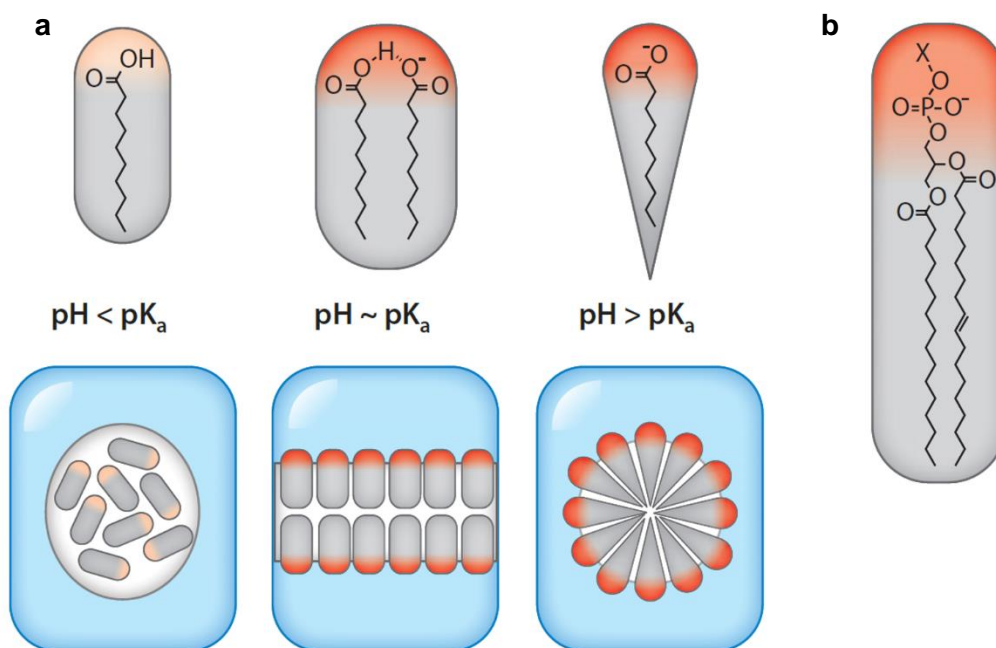available on the early Earth, such as fatty acids or fatty alcohols, would have been able to form cell-like membranes.[66]

Figure 1.8 | Common components of protocell compartments. a, Fatty acids spontaneously form aggregates above a certain concentration. At a pH value lower than their acid dissociation constant (pK_a), the aggregates are amorphous precipitates or oils, while above their pK_a, the formation of small micelles can be observed. Near their pK_a, the partially ionized fatty acids form bilayer vesicles. b, Phospholipids are composed of two aliphatic chains and form more stable membranes under a wider range of conditions (adapted from Blain J. C. and Szostak J. W.[38]).

Indeed, primordial membranes would have required a certain extent of permeability to facilitate the free diffusion of polar and ionic solutes between the prebiotic cell-like structure (or protocell) and the environment. Since membrane permeability is inversely proportional to the degree of unsaturation and the length of the aliphatic chains, membranes formed of short-chain fatty acids or fatty alcohols would have been advantageous for prebiotic systems.[38] Moreover, complete replication cycles of growth and division of fatty acid membranes have been shown under prebiotically plausible conditions.[67] However, the stability of short-chain amphiphiles membranes is tremendously affected by environmental conditions such as divalent cations, pH, salinity, and ionic strength.[26] For example, fatty acids can form membranes when the pH of the solution is approximately the pK_a of the carboxylate headgroup, thus implying a short pH range of availability of such membranes in aqueous conditions. Moreover, fatty acids protocells are unstable in the presence of divalent cations such as Mg^{2+}.[38] That is, more complex membranes would be required for more

advanced protocells. In order to prevent the salt-induced disruption of fatty acid vesicles, or to generate chemical gradients across the membranes, phospholipid-based membranes must have become essential.[68]

It is probable that early membranes have undergone a significant evolutionary process as the first living organisms evolved and developed new abilities.[69] Indeed, understanding how life emerged and evolved seems to be closely related to understanding how cellular membranes emerged and evolved. The lack of stabilizing structures and protein systems responsible for the exchange of molecules across the membrane in early protocells seems to suggest highly permeable membranes, as in the case of fatty vesicles. As simple transport systems across the membrane emerged, more robust cell boundaries would have been needed, such as short-chain phospholipid and fatty acids mixtures. Finally, through progressive steps of evolution and optimization, homogenous phospholipid membranes would have appeared.[25,26]

1.8 Bibliography

1.   Falkowski, P. G. From Light to Life. *Orig. Life Evol. Biosph.* (2015). doi:10.1007/s11084-015-9441-6

2.   Anbar, A. D. Elements and Evolution. *Science (80-.).* **322,** 1481–1483 (2008).

3.   Krishnamurthy, R. Giving Rise to Life: Transition from Prebiotic Chemistry to Protobiology. *Acc. Chem. Res.* **50,** 455–459 (2017).

4.   Sutherland, J. D. The Origin of Life — Out of the Blue. *Angew. Chemie* **55,** 104–121 (2016).

5.   Szostak, J. W., Bartel, D. & Luisi, P. L. Synthesizing life. *Nature* **409,** 387–390 (2001).

6.   Müller, U. F. & Tor, Y. Citric Acid and the RNA World. *Angew. Chemie* **53,** 5245–5247 (2014).

7.   Pascal, R., Pross, A. & Sutherland, J. D. Towards an evolutionary theory of the origin of life based on kinetics and thermodynamics. *Open Biol.* **3,** 1–9 (2013).

8.   Szostak, J. W. Attempts to define life do not help to understand the origin of life. *J. Biomol. Struct. Dyn.* **29,** 599–600 (2012).

9.   Gest, H. The discovery of microorganisms by Robert Hooke and Antoni van Leeuwenhoek, Fellows of The Royal Society. *Notes Rec. R. Soc.* **58,** 187–201 (2004).

10. Kinne-Saffran, E. & Kinne, R. K. H. Vitalism and synthesis of urea. *Am. J. Nephrol.* **19,** 290–294 (1999).

11. Oparin, A. I. *The Origin of Life.* (The Macmillan Company, 1938). doi:10.2113/gselements.12.6.413

12. Haldane, J. B. S. The origin of life. *Ration. Annu.* **148,** 3–10 (1929).

13. Miller, S. L. A production of amino acids under possible primitive earth conditions. *Science (80-.).* **117,** 528–529 (1953).

14. Miller, S. L. & Urey, H. C. Organic compound synthesis on the primitive earth. *Science* **130,** 245–51 (1959).

15. Parker, E. T. *et al.* Primordial synthesis of amines and amino acids in a 1958 Miller H2S-rich spark discharge experiment. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 5526–5531 (2011).

16. Orgel, L. E. Evolution of the genetic apparatus. *J. Mol. Biol.* **38,** 381–393 (1968).

17. Kruger, K. *et al.* Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* **31,** 147–157 (1982).

18. Gilbert, W. The RNA world. *Nature* **319,** 618 (1986).

19. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **254,** 1497–1500 (1991).

20. Schoning, K.-U. *et al.* Chemical Etiology of Nucleic Acid Structure: The alpha - Threofuranosyl-(3'rightarrow 2') Oligonucleotide System. *Science (80-.).* **290,** 1347–1351 (2000).

21. Zhang, L., Peritz, A. & Meggers, E. A simple glycol nucleic acid. *J. Am. Chem. Soc.* **127,** 4174–4175 (2005).

22. Wächtershäuser, G. Pyrite Formation, the First Energy Source for Life: a Hypothesis. *Syst. Appl. Microbiol.* **10,** 207–210 (1988).

23. Russell, M. J., Daniel, R. M. & Hall, A. J. On the emergence of life via catalytic iron-sulphide membranes. *Terra Nov.* **5,** 343–347 (1993).

24. Martin, W. & Russell, M. J. On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos. Trans. R. Soc. B Biol. Sci.* **358,** 59–85 (2003).

25. Morowitz, H. J., Heinz, B. & Deamer, D. W. The chemical logic of a minimum protocell. *Orig. Life Evol. Biosph.* **18,** 281–287 (1988).

26. Monnard, P. A. & Deamer, D. W. Membrane self-assembly processes: Steps toward

the first cellular life. *Anat. Rec.* **268,** 196–207 (2002).

27. Walde, P., Wick, R., Fresta, M., Mangone, A. & Luisi, L. P. Autopoietic Self-Reproduction of Fatty Acid Vesicles. *J. Am. Chem. Soc.* **116,** 11649–11654 (1994).

28. Segré, D., Ben-Eli, D., Deamer, D. W. & Lancet, D. The lipid world. *Orig. Life Evol. Biosph.* **31,** 119–145 (2001).

29. Mansy, S. S. *et al.* Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **454,** 122–5 (2008).

30. Cech, T. R. The ribosome is a ribozyme. *Science (80-.).* **289,** 878–9 (2000).

31. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The Structural Basis of Ribosome Activity in Peptide Bond Synthesis. *Science (80-.).* **289,** 920–930 (2000).

32. 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.* **7,** 301–307 (2015).

33. Moretti, J. E. & Muller, U. F. A ribozyme that triphosphorylates RNA 5'-hydroxyl groups. *Nucleic Acids Res.* **42,** 4767–4778 (2014).

34. Attwater, J., Wochner, A. & Holliger, P. In-ice evolution of RNA polymerase ribozyme activity. *Nat. Chem.* **5,** 1011–8 (2013).

35. Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459,** 239–242 (2009).

36. Tohidi, M., Zielinski, W. S., Chen, C. H. B. & Orgel, L. E. Oligomerization of 3'-amino-3'-deoxyguanosine-5'-phosphorimidazolidate on a d(CpCpCpCpC) template. *J. Mol. Evol.* **25,** 97–99 (1987).

37. Higgs, P. G. & Lehman, N. The RNA World: molecular cooperation at the origins of life. *Nat. Rev. Genet.* (2014). doi:10.1038/nrg3841

38. Blain, J. C. & Szostak, J. W. Progress Toward Synthetic Cells. *Annu. Rev. Biochem.* **83,** 615–640 (2014).

39. Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E. & Bartel, D. P. RNA-Catalyzed RNA Polymerization: Accurate and General RNA-Templated Primer Extension. *Science (80-.).* **292,** 1319–1325 (2001).

40. Wochner, A., Attwater, J., Coulson, A. & Holliger, P. Ribozyme-Catalyzed Transcriptionof an Active Ribozyme. *Science (80-.).* **61911,** 2009–2012 (2011).

41. Lincoln, T. A. & Joyce, G. F. Self-sustained replication of an RNA enzyme. *Science (80-.).* **323,** 1229–32 (2009).

42. Sojo, V., Herschy, B., Whicher, A., Camprubí, E. & Lane, N. The Origin of Life in

Alkaline Hydrothermal Vents. *Astrobiology* **16,** 181–197 (2016).

43. Morowitz, H. J., Kostelnik, J. D., Yang, J. & Cody, G. D. The origin of intermediary metabolism. **2000,** (2000).

44. Morowitz, H. & Smith, E. Energy flow and the organization of life. *Complexity* **13,** 51–59 (2007).

45. Smith, E. & Morowitz, H. J. Universality in intermediary metabolism. *Proc. Natl. Acad. Sci.* **101,** 13168–13173 (2004).

46. Maurel, M. & Leclerc, F. From Foundation Stones to Life: Concepts and Results. *Elements* **12,** 407–412 (2016).

47. Blöchl, E., Keller, M., Wachtershäuser, G. & Stetter, K. O. Reactions depending on iron sulfide and linking geochemistry with biochemistry. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 8117–8120 (1992).

48. Wächtershäuser, G. On the chemistry and evolution of the pioneer organism. *Chem. Biodivers.* **4,** 584–602 (2007).

49. Cairns-smith, A. G., Hall, A. J. & Russell, M. J. Mineral Theories of the Origin of Life and An Iron Sulfide Example. *Orig. Life Evol. Biosph.* **22,** 161–180 (1992).

50. Nitschke, W. & Russell, M. J. Hydrothermal Focusing of Chemical and Chemiosmotic Energy, Supported by Delivery of Catalytic Fe, Ni, Mo/W, Co, S and Se, Forced Life to Emerge. *J. Mol. Evol.* **69,** 481–496 (2009).

51. Mielke, R. E. *et al.* Iron-Sulfide-Bearing Chimneys as Potential Catalytic Energy Traps at Life's Emergence. *Astrobiology* **11,** 933–950 (2011).

52. Darwin, C. *On the origin of species. On the origin of species* (D. Appleton and Company, 1859). doi:10.4324/9780203509104

53. Theobald, D. L. A formal test of the theory of universal common ancestry. *Nature* **465,** 219–222 (2010).

54. Lane, N., Allen, J. F. & Martin, W. How did LUCA make a living? Chemiosmosis in the origin of life. *BioEssays* **32,** 271–280 (2010).

55. Weiss, M. C. *et al.* The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **1,** 16116 (2016).

56. Baltscheffsky, H., Blomberg, C., Liljenström, H., Lindahl, B. I. & Arhem, P. On the origin and evolution of life: an introduction. *J. Theor. Biol.* **187,** 453–9 (1997).

57. Ycas, M. On certain homologies between proteins. *J.Mol.Evol.* **7,** 215–244 (1976).

58. Hug, L. A. *et al.* A new view of the tree of life. *Nat. Microbiol.* **1,** 16048 (2016).

59. Ruiz-mirazo, K., Briones, C. & Escosura, D. Prebiotic Systems Chemistry : New

Perspectives for the Origins of Life. (2014).

60. Gross, M. How life can arise from chemistry. *Curr. Biol.* **26,** R1247–R1249 (2016).

61. Belmonte, L. & Mansy, S. S. Metal Catalysts and the Origin of Life. *Elements* **12,** 413–418 (2016).

62. Lowenstein, J. M. Non-enzymic formation of hydroxamates catalysed by manganous ions. *Biochim. Biophys. Acta* **28,** 206–7 (1958).

63. Calvin, M. Evolution of enzymes and the photosynthetic apparatus. *Science* **130,** 1170–4 (1959).

64. Athavale, S. S. *et al.* RNA Folding and Catalysis Mediated by Iron (II). *PLoS One* **7,** e38024 (2012).

65. McMillan, R. S., Renaud, J., Reynolds, J. G. & Holm, R. H. Biologically related iron-sulfur clusters as reaction centers. Reduction of acetylene to ethylene in systems based on [Fe4S4(SR)4]3-. *J. Inorg. Biochem.* **11,** 213–227 (1979).

66. Deamer, D. W. Boundary structures are formed by organic components of the Murchison carbonaceous chondrite. *Nature* **317,** 792–794 (1985).

67. Budin, I., Debnath, A. & Szostak, J. W. Concentration-driven growth of model protocell membranes. *J. Am. Chem. Soc.* **134,** 20812–20819 (2012).

68. Chen, I. A. & Walde, P. From Self-Assembled Vesicles to Protocells. *Cold Spring Harb. Perspect. Biol.* **2,** 1–13 (2010).

69. Deamer, D. & Weber, A. L. Bioenergetics and Life ' s Origins. 1–16 (2010).

Chapter 2 - Duplication of a model prebiotic iron-sulfur tripeptide leads to the formation of a protoferredoxin

2.1 Introduction

Combinatorial shuffling and selection events starting from a basic set of simple protein patterns could suggest how modern proteins emerged and evolved. On early Earth short peptide sequences gave rise by repetition, recombination and fusion events to the first folded domains, which can be phylogenetically traced back to LUCA, around 3.5 billion years ago. In a hypothetic RNA world, an ancestral set of peptides may have cooperated with RNA to expand the functionalities of RNA, since peptides are able to bind nucleic acids and coordinate small molecules, metal ions, iron-sulfur clusters and nucleotide-derived cofactors.[1]

The combined analysis of sequence, structure and shared metabolism on the wide variety of proteins suggest that the activity of all the most ancient enzymes is dependent on metal ion coordination.[2] Moreover, the primary and secondary structures of domains binding metal ions are often highly conserved. It has been proposed that the repetition of short sequences of amino acids mimicking the structural features and the amino acid sequences of conserved motifs could have given rise to modern metalloproteins.[3,4] About 50 years ago Richard Eck and Margaret Dayhoff hypothesized that a 55-amino acid redox-active protein found in the obligate anaerobic bacterium *C. pasteurianum* was evolved from duplication events. It was suggested that the protein could be broken down into progressively smaller repetitive units, all the way down to a peptide constituted of 4 amino acids, Ala-Asp-Ser-Gly. Although the full-length protein coordinates two Fe_4S_4 clusters, the tetrapeptide was not assumed to bind a metal ion. Instead, it was proposed that later mutations resulted in properly positioned cysteines for iron-sulfur cluster coordination.[3] According to this hypothesis, metal ions would have been coordinated first by small molecules or amino acids, then by short peptides. Modern metalloproteins would have been built from metal-coordinating peptide sequences conferring some kind of selective advantage, resulting in longer and more active domains. Therefore, it may be possible to recognize within the primary structure of modern metalloproteins these ancient metal-coordinating peptide motifs.[5]

Common features across modern organisms suggest the existence of molecular relics from earlier phases of evolution. The dependence of all living systems on metal-based chemistry for fundamental processes suggests the importance of metal ions on the emergence and evolution of life as we know it. There are compelling reasons why metals would have been exploited in prebiotic chemistry and in the earliest phases of protocellular evolution. Metals are common catalysts and are capable of funnelling model prebiotic reactions towards the synthesis of nucleotides, amino acids, and lipids[6] and facilitate the emergence of metabolic-like pathways.[7] Hadean Earth was rich in metal ions, particularly in iron,[8] suggesting that nascent oligonucleotides or peptides could have exploited the available metals for folding.[5] In fact, *in vitro* attempts to mimic evolutionary processes that lead to active peptide and nucleic acid folds invariably give sequences that depend on metal coordination for activity.[3] Therefore, metal ions and metal-based moieties, in particular iron-sulfur clusters, could be viewed as inorganic, molecular relics. Iron-sulfur clusters are needed for cellular life to survive. Moreover, iron and iron-sulfur proteins may have played an even larger role in the past, as some contemporary non-iron enzymes may have evolutionarily ancient progenitors that were iron or iron-sulfur dependent, including NAD(P)-proteins and Mg^{2+}-dependent RNAs.[9,10] The biological iron-sulfur cluster synthesis machinery is evolutionarily ancient with at least one component hypothesized to be needed to support minimal cellular life.[11] Iron-sulfur proteins are believed to be among the most ancient proteins emerged[9] and are fundamental to several physiological processes, including central metabolism and protein and DNA synthesis.[12] Iron-sulfur cluster stability is dependent on several factors, including the nature of the coordinating residue, the rigidity of the protein backbone, and the protective effect of the backbone on the cluster from hydrolysis.[13] The latter two factors are governed by the fold of the protein, suggesting that long peptide chains are needed to achieve stable iron-sulfur cluster coordination in aqueous solution. These considerations would indicate that iron-sulfur clusters did not emerge on early Earth until protein synthesis machinery evolved or that protocellular life exploited iron-sulfur minerals to catalyse the biochemical processes which now are based on iron-sulfur proteins. Moreover, it is unclear if the dependence on iron-sulfur cores can be traced all the way back to the origins of cellular life or was a later adaptation that came to play a crucial role near the advent of the last universal common ancestor.

In the iron-sulfur proteins and their synthetic analogues, each iron atom is generally covalently linked to four sulfur atoms (Figure 2.1). A cluster may contain from one to eight Fe atoms and a given protein may contain several Fe-S clusters. The three most common

iron–sulfur clusters are mononuclear, Fe_2S_2, and Fe_4S_4 clusters. In all these clusters, each tetrahedral iron is frequently coordinated by cysteine side-chains and, in the case of the most common Fe_2S_2 and Fe_4S_4 clusters, the iron is additionally coordinated by inorganic sulfides. Several iron-sulfur cluster motifs are easily identifiable in modern protein sequences. One class of Fe_4S_4 proteins coordinate their cluster, in part, through a Cys-Xxx-Xxx-Cys-Xxx-Xxx-Cys motif, where Xxx is any amino acid. These proteins additionally contain a fourth cysteine ligand further away in the primary sequence that is necessary to coordinate the cluster. Similar motifs can be found for specific classes of mononuclear iron and Fe_2S_2 proteins, although iron-sulfur clusters can be even coordinated to quite different sequences.
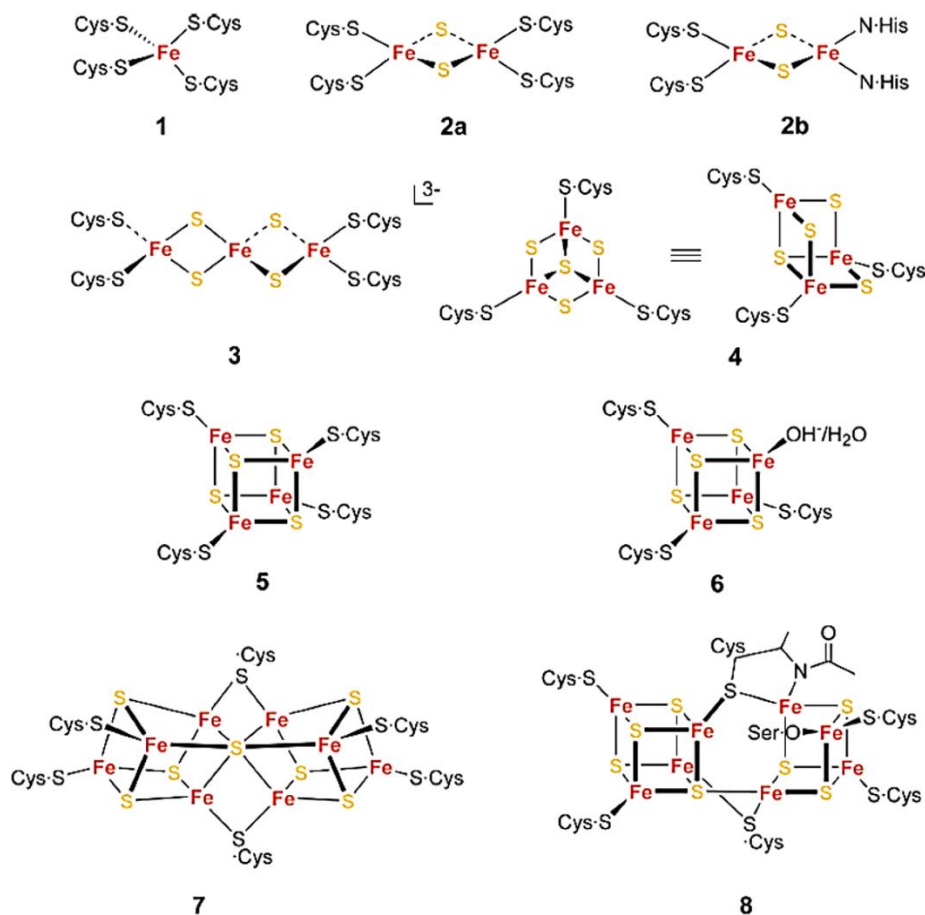
Figure 2.1 | Schematic representations of crystallographically demonstrated iron-sulfur protein cores.

In terms of phylogeny, the most studied iron-sulfur proteins are the ferredoxins, which are present in organisms from each of the three primary lines of evolutionary descent. Phylogenetic comparison identifies ferredoxins (and ATPases) as components of an early stage of cell metabolism.[14] There is possibly an earlier, simpler, and potentially related fold

in desulforedoxin, where a metal (usually iron) is bound to two cysteine residues.[15] The peptides that gave rise to iron-sulfur proteins may have impacted the evolution of other protein folds as well, because early metal complexes with peptides were probably quite unrestrained in coordinating metal ions. It is quite common in modern proteins to exhibit a different metal (Fe^{2+}, Ni^{2+}, and Zn^{2+}) coordination with the same residue. A recent attempt to deduce prebiotic metal-coordinating peptide sequences was focused on the search for short sequences that primarily consisted of prebiotic amino acids (glycine, alanine, aspartic acid, and valine) and that coordinated at least one metal ion. Two different metal ion-binding motifs of Asp-Xxx-Asp-Xxx-Asp and Asp-Xxx-Xxx-Xxx-Xxx-Asp-Xxx-Asp were found, associated in modern biology with the binding of Mg^{2+}, Mn^{2+}, Zn^{2+}, and Ni^{2+}.[16]

An alternative bottom-up way to gain insight into when iron-sulfur proteins could have emerged is to determine the minimum unit needed to coordinate and stabilize iron-sulfur clusters in aqueous solution. Several prebiotically plausible amino acid polymerization reactions have been described.[17] The ability of short peptides to stably coordinate iron-sulfur clusters, for example, would support the presence of peptide-coordinating iron-sulfur clusters on prebiotic Earth if the needed components were available. Modern day iron-sulfur proteins can be as small as about 50 amino acids, but synthetic iron-sulfur peptides can be smaller. In the last decades, many efforts have been made to study the spontaneous assembly of Fe-S cluster coordinate by non-genetically encoded sterically hindered thiolates mainly in organic solvents, such as methanol, tetrahydrofuran or acetonitrile.[13] A 16 amino acid [4Fe-4S] peptide was built by incorporating a sequence found in *Peptococcus aerogenes* ferredoxin.[18] More recently, efforts to elucidate the mechanism of iron-sulfur cluster biosynthesis and transport demonstrated that the tripeptide L-glutathione (Figure 2.2) coordinates a $[2Fe-2S]^{2+}$ cluster in aqueous solution.[19] The reaction was performed by mixing aqueous ferric and sulfide ions in the presence of an excess of the coordinating tripeptide.
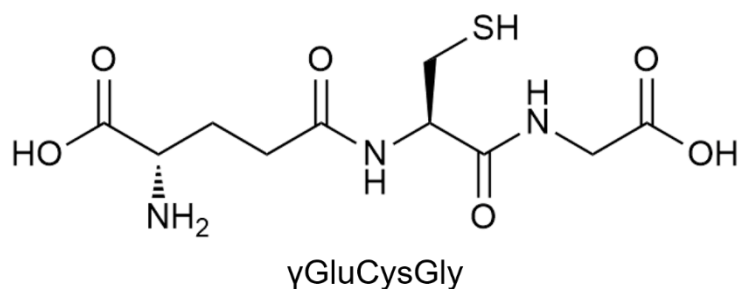
Figure 2.2 | Chemical structure of reduced L-glutathione (γGluCysGly).

Thiol-containing organic molecules are widely accepted in terms of prebiotic availability. For example, the Strecker precursor of cysteine, beta-mercaptoacetaldehyde, is used for the prebiotic synthesis of nucleotides.[20] Additionally, data collected from the Miller experiment performed in the presence of H_2S showed the presence of degradation products of cysteine,[21] suggesting that cysteine was originally formed together with several other sulfur-containing organic compounds.[22] The fact that a short Cys-containing peptide can coordinate an iron-sulfur cluster in water suggested to us the possibility that short iron-sulfur peptides could have existed on prebiotic Earth. Therefore, the first step in this direction would be further probe the stability of L-glutathione [2Fe-2S] cluster to model prebiotic conditions and determine if contemporary iron-sulfur proteins could have originated from short peptide sequences able to undergo redox cycles. Secondly, to determine if such a short peptide could give rise to a protoferredoxin upon polymerization, based on quantum mechanics and molecular dynamics simulations used to gain insight into the potential structure, one, two, and three replication events were pursued by synthesizing peptides of six - (γGluCysGly)$_2$ -, nine- (γGluCysGly)$_3$ -, and twelve - (γGluCysGly)$_4$ - amino acids in length (Figure 2.3).
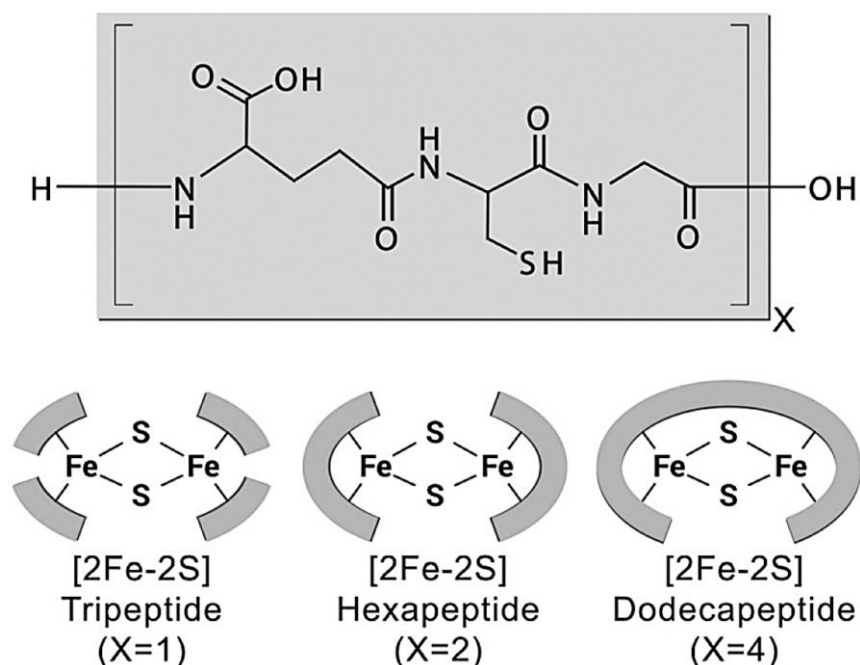
Figure 2.3 | Schematic structures of the [2Fe-2S] clusters coordinated by L-glutathione polymers. The [2Fe-2S] cluster coordinated by the nonapeptide is not shown for simplicity.

2.2 Results

To gain insight into the environmental conditions compatible with iron-sulfur cluster coordination to L-glutathione, a series of variables were screened, including L-glutathione, iron and sulfide concentration and pH on [2Fe-2S] stability. However, since several species other than L-glutathione [2Fe-2S] cluster can be formed under the explored synthetic prebiotic conditions and absorb in the visible region, reference UV-vis. absorbance spectra were first acquired so that spectral decomposition could be used to follow the distribution of each species under different conditions. Four reference spectra were collected, including oxidized mononuclear iron complex coordinated to L-glutathione, oxidized L-glutathione in the presence of $FeCl_3$, oxidized L-glutathione [2Fe-2S] cluster synthesized following previously reported procedures[19], and reduced L-glutathione [2Fe-2S] cluster (Figure 2.4). Subsequent UV-vis. absorbance spectra were then decomposed to a weighted sum of these reference spectra.
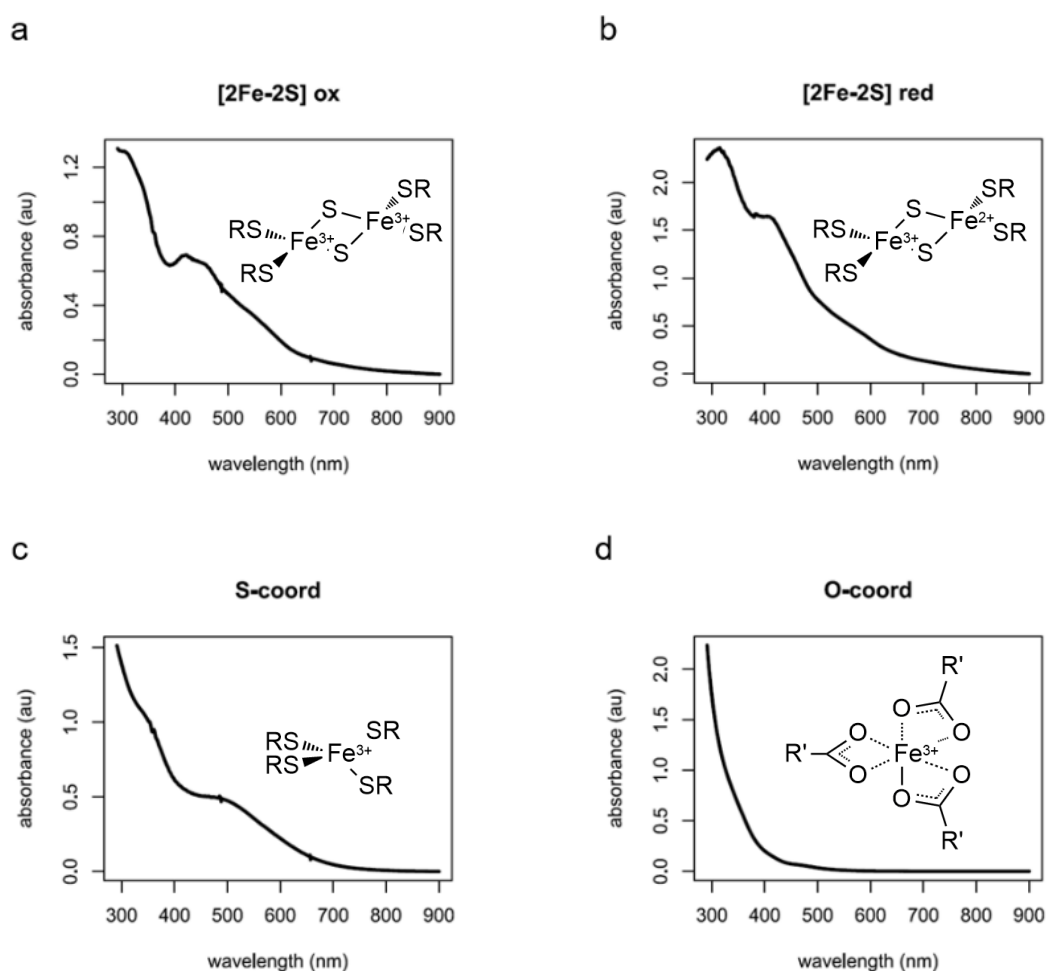
Figure 2.4 | UV-Visible absorption spectra of the four reference species used for UV-Visible spectral decomposition. a, oxidized form of L-glutathione [2Fe-2S] cluster; b, reduced form of L-glutathione [2Fe-2S]

cluster; c, oxidized mononuclear L-glutathione-iron complex (*S*-coordination to iron); d, oxidized L-glutathione iron complex (*O*-coordination to iron). Simplified schematic structures of the iron-sulfur systems are reported for each studied species.

Previous studies on iron-sulfur cluster in organic solvents with different polarity have shown that the synthesis of [2Fe-2S] clusters can be easily obtained in inert atmosphere by the addition of a stoichiometric amount of each component in slightly basic conditions, in order to deprotonate the coordinating thiol group.[13] However, hydrolytically stable iron-sulfur clusters coordinated by small thiols have not been frequently studied in aqueous environment. The pioneering work of Hisashi Tanaka on sulfur-containing amino acids and peptides able to coordinate [2Fe-2S] clusters in water was poorly characterized and documented, and no subsequent work was pursued.[23] Few years later, Hiroshi Matsubara and Akira Nakamura reported data on the stability of a reduced mononuclear iron complex coordinated by cysteine-containing peptides in aqueous Triton X-100 micelle solution.[24,25] In all the cases, the authors pointed out the need to use an excess of the binding ligand in order to prevent iron-sulfur clusters from hydrolysis.

A second obstacle is represented by the narrow pH range required for the synthesis of iron-sulfur clusters in aqueous solutions. Iron ions tend to precipitate at neutral and basic pH as mixed oxides and hydroxides,[26] in the same pH range in which sulfide ions can be available in solution for the assembly of iron-sulfur clusters (pK_a H_2S = ~7.00)[27] and the thiol groups are mostly deprotonated and prone to coordinate iron ions (pK_a values are usually around 8-9, depending on the neighbouring groups).[27] Moreover, it is widely reported in the literature that redox-active metal ions, such as Fe and Cu, are able to efficiently oxidize thiols in aqueous solutions.[28,29] Indeed, when a stoichiometric amount of each component was mixed under anaerobic conditions, *i.e.* 2:1:1 L-glutathione:$FeCl_3$:Na_2S in order to have four L-glutathione molecules available to coordinate one [2Fe-2S] cluster, evidence of a [2Fe-2S] cluster was not observed. Diamagnetic ^{1}H NMR spectroscopy showed that, instead of [2Fe-2S] cluster formation, L-glutathione became oxidized, by donating electrons to the ferric ions (Figure 2.5).
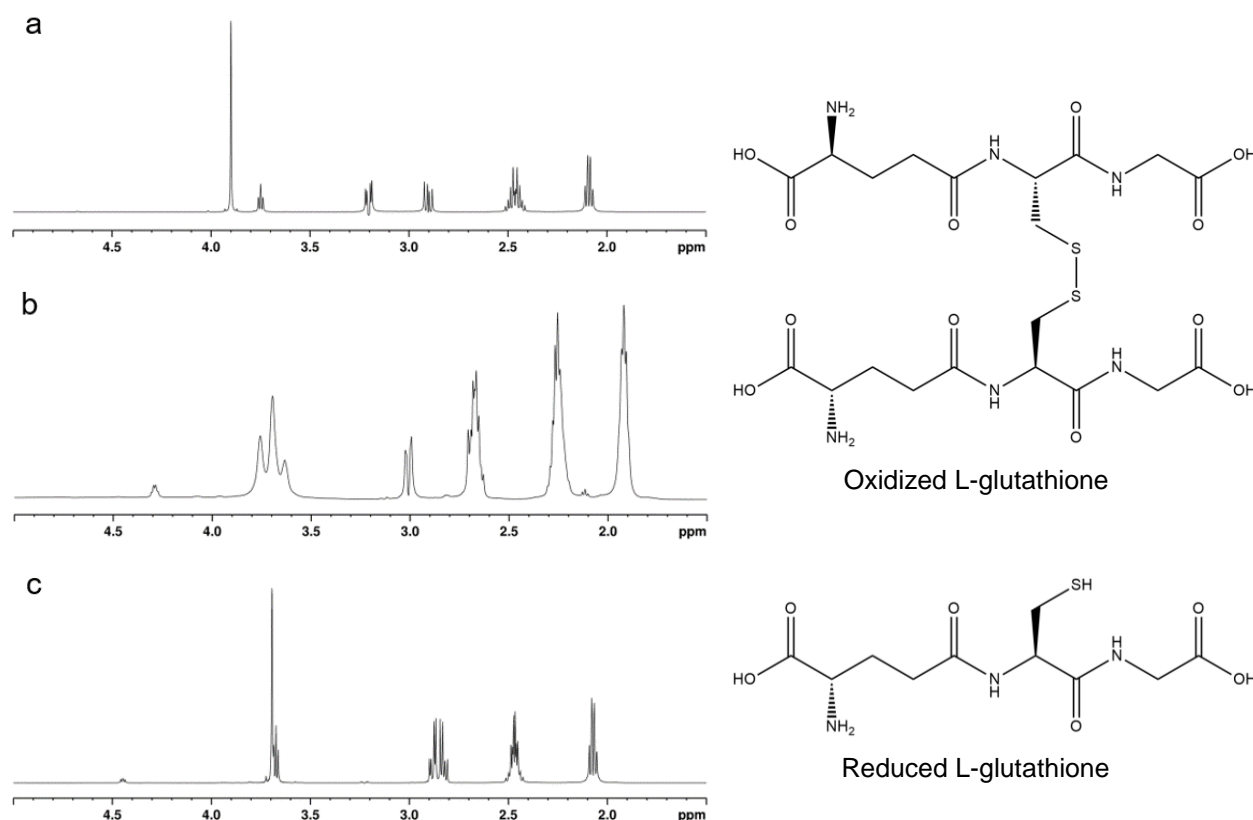
Figure 2.5 | ^{1}H NMR spectra of a, oxidized L-glutathione/Fe^{2+} (20 mM oxidized L-glutathione, ratio 2:1, pH 8.6) b, mixture of L-glutathione/S^{2-}/Fe^{3+} (40 mM L-glutathione, ratio 2:1:1, pH 8.6) and c, L-glutathione (40 mM L-glutathione, pH 8.6), in D_2O. On the right, chemical structures of the oxidized and the reduced forms of L-glutathione.

Further studies on the best concentration range of L-glutathione were then performed. Spectral decomposition of L-glutathione titrations where iron and sulfide were kept constant showed the appearance of the [2Fe-2S] cluster at 20 mM L-glutathione (40:1 L-glutathione:$FeCl_3$) with 40 mM L-glutathione leading to maximal quantities of [2Fe-2S] (Figure 2.6). L-glutathione [2Fe-2S] cluster showed UV-vis. spectral features similar to that of human ferredoxin (Figure 2.6), and mass spectrometry data were consistent with four L-glutathione molecules coordinated to one [2Fe-2S] cluster, as previously reported.[30] Next, the range of sulfide and iron concentrations amenable to L-glutathione [2Fe-2S] cluster formation was evaluated. When keeping $FeCl_3$ (0.5 mM) and L-glutathione (40 mM) constant the addition of approximately 0.2 mM Na_2S gave the highest amount of [2Fe-2S] cluster (Figure 2.6). Increasing concentrations of Na_2S produced water-insoluble inorganic iron sulfide species which compete with [2Fe-2S] cluster formation. No oxidized L-glutathione [2Fe-2S] cluster was observed at 0.5 mM Na_2S or higher. A similar titration with constant L-glutathione (40 mM) and Na_2S (0.18 mM) revealed an optimum $FeCl_3$ concentration of 0.5

mM (Figure 2.6). Higher concentrations increased the fraction of competing products, including oxidized mononuclear iron complex and inorganic iron sulfide species. In the presence of excess L-glutathione, the highest amount of [2Fe-2S] cluster was observed upon the addition of Na_2S first, and then $FeCl_3$. Use of Fe^{2+} in place of Fe^{3+} failed to give a [2Fe-2S], which was consistent with previous Mössbauer studies showing the presence of diferric metal centers.[19] However, when the solution containing $FeCl_2$, Na_2S, L-glutathione was exposed to oxygen, the sample gave rise to UV-vis. spectral bands consistent with a [2Fe-2S] cluster.

L-glutathione [2Fe-2S] cluster is stable to alkaline conditions. To further define the range of conditions compatible with [2Fe-2S] cluster synthesis, the influence of pH on [2Fe-2S] cluster yield was evaluated. L-glutathione contains several potential metal ligands, each exhibiting different pK_a values, including two carboxylates (2.1, 3.5), a thiol (8.6), and an amino-group (9.1). Additionally, as previously mentioned, inorganic sulfide and ferric ions stability and solubility are pH dependent. Sulfide ions equilibrate between S^{2-}, HS^-, and H_2S based on the pH of the solution, where H_2S is a volatile component which can escape as a gas and thus deplete the pool of sulfide necessary for [2Fe-2S] formation. Insoluble metal hydroxides form under alkaline conditions, thus decreasing the availability of iron ions for cluster formation at high pH values. A pH titration revealed that L-glutathione [2Fe-2S] cluster persisted between pH 7.5 and pH 10, with maximal L-glutathione [2Fe-2S] cluster observed near pH 8.5 (Figure 2.6). The optimal pH for cluster formation ensured the presence of deprotonated cysteine residues (as said, pK_a 8.6) and decreased quantities of sulfur species that can form inorganic, water-insoluble iron sulfide species. The lack of L-glutathione [2Fe-2S] cluster under acidic conditions was consistent with the presence of protonated cysteine side-chains, the lability of the bridging inorganic sulfides, and the volatility of H_2S ($pK_a = 7.0$). On the other hand, a too alkaline pH (over pH 10) would induce the predominant precipitation of oxo-hydroxo iron species and iron sulfide species rather than the formation of the [2Fe-2S] cluster.
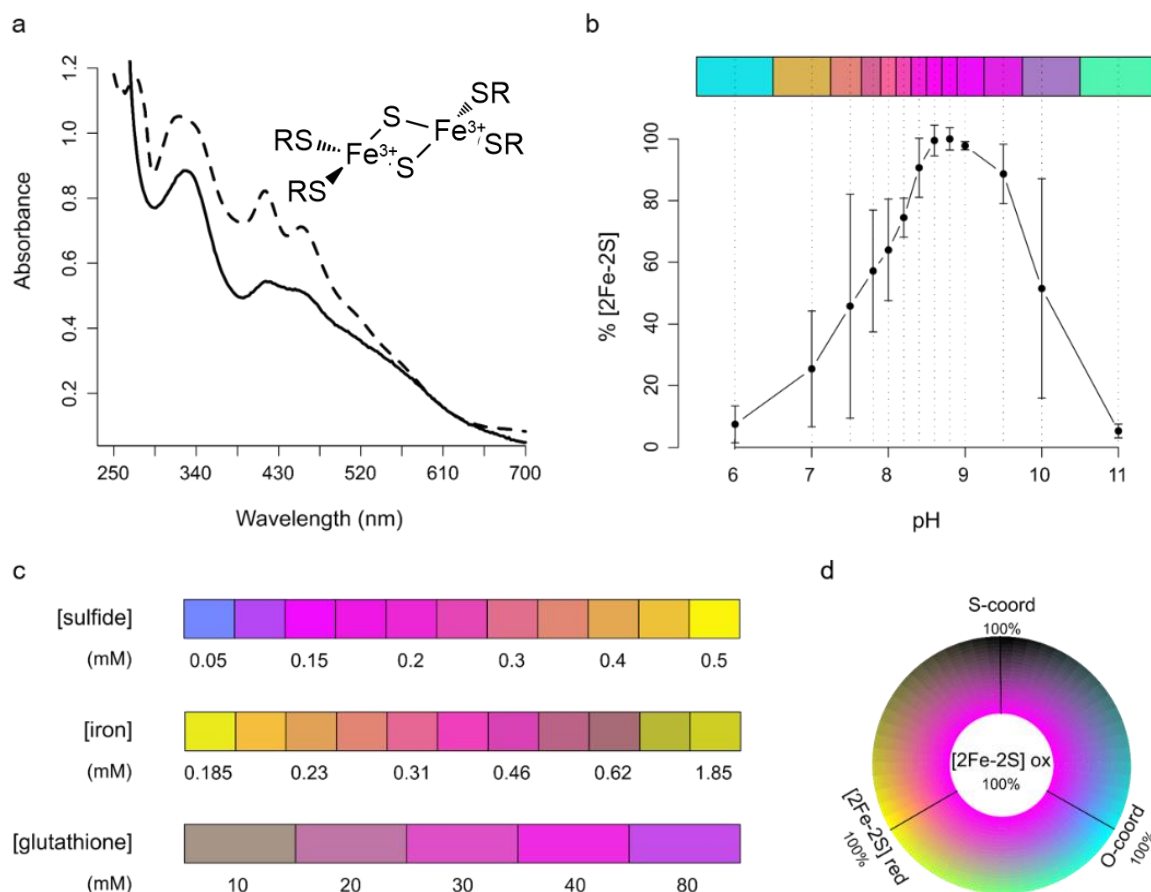
Figure 2.6 | L-glutathione stabilized [2Fe-2S] cluster. a, UV-vis. spectra of human ferredoxin (dashed line) and L-glutathione [2Fe-2S] cluster (solid line). The conditions for L-glutathione [2Fe-2S] cluster were 40 mM L-glutathione, 0.18 mM sodium sulfide, and 0.5 mM ferric chloride at pH 8.6. A simplified schematic structure of the [2Fe-2S] cluster is reported. b, Screening profiles for L-glutathione, sodium sulfide, and ferric chloride concentrations obtained by spectral decomposition. The data show that the optimal values were 40 mM L-glutathione, 0.18 mM sodium sulfide, and 0.5 mM ferric chloride. Data shown are averages of measurements on three independently run reactions. c, The influence of pH on the formation of L-glutathione [2Fe-2S] cluster obtained by spectral decomposition. The resulting sample composition at each pH is shown in the coloured bar above the plot. Data shown are averages of measurements on three independently run reactions with error bars signifying standard deviation. d, Colour wheel legend explaining the meaning of the colours used in panels B, C, and D are composed of four colours indicating the contribution from oxidized L-glutathione [2Fe-2S] cluster (magenta), oxidized mononuclear L-glutathione-iron complex (black), reduced L-glutathione [2Fe-2S] cluster (yellow), and oxidized L-glutathione in the presence of ferric chloride (cyan).

Paramagnetic NMR studies performed on L-glutathione [2Fe-2S] cluster resulted similar to those previously observed for [2Fe-2S] ferredoxin.[31] The paramagnetic NMR spectra of L-glutathione [2Fe-2S] cluster were consistent with ligand exchange rates which are fast on the NMR time scale. To confirm whether a dynamic equilibrium existed whereby at any given time different molecules were coordinated to a single [2Fe-2S] cluster, a chromatography assay was employed using L-glutathione-conjugated Sepharose resin

(GSTrap column). This resin was chosen to increase the likelihood of the cluster surviving the column since the resin itself would be capable of ligating the iron centres. When L-glutathione [2Fe-2S] cluster was loaded on the column with 40 mM L-glutathione in the running buffer, two peaks eluted that absorbed at 405 nm (Figure 2.7). The first, larger peak exhibited an UV-vis. spectrum consistent with L-glutathione [2Fe-2S] cluster, whereas the UV-vis. spectrum of the second, smaller peak was dominated by oxidized mononuclear iron complex coordinated to L-glutathione. When the experiment was repeated without L-glutathione in the running buffer, only one clear peak eluted with an UV-vis. spectrum consistent with a hexa-aquo Fe^{3+} complex and not a [2Fe-2S] cluster. The data support the existence of a dynamic equilibrium that requires excess L-glutathione to stabilize [2Fe-2S] clusters.
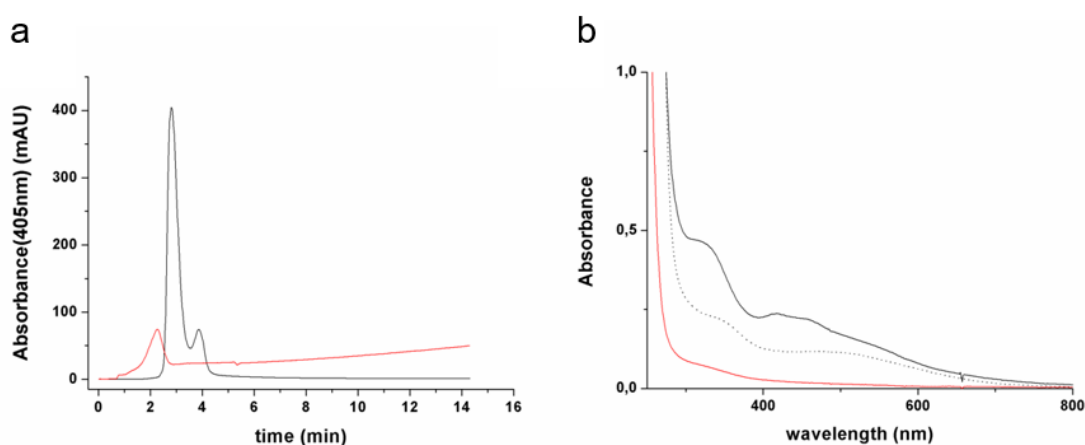
Figure 2.7 | a, Comparison of chromatograms of L-glutathione [2Fe-2S] cluster eluted from a L-glutathione column (GSTrap) by using 40 mM L-glutathione, pH 8.6 (black line) and water (red line) as eluent. b, UV-vis. spectra of the fractions eluted. When the mixture is eluted with 40 mM L-glutathione pH 8.6 (a, black line), two discrete peaks are detected, corresponding to the elution of L-glutathione [2Fe-2S] cluster (b, black solid line) and oxidized mononuclear L-glutathione-iron complex (b, black dashed line), respectively. None of the two species is isolated if the running buffer lacks L-glutathione (b, red line).

Since the paramagnetic NMR data suggested the presence of reduced [2Fe-2S] cluster, the redox-activity of the [2Fe-2S] coordinated to L-glutathione was tested. The oxidized cluster, *i.e.* [2Fe-2S]$^{2+}$, is EPR silent because of antiferromagnetic coupling, whereas the reduced, mixed valence state of ferredoxin-like centres showed a distinctive g=1.94 signal.[32] The addition of the reductant dithionite to a freshly prepared sample of oxidized [2Fe-2S] cluster showed the presence of a reduced [2Fe-2S]$^{1+}$ cluster coordinated to L-glutathione. To determine whether the reduced state was stable and capable of returning to the oxidized state, a reduced sample was run through a Sephadex G-10 gel

filtration column to trap within the resin the reductant and any degraded cluster or free ion that may have formed. The eluent was then oxidized with hydrogen peroxide. The resulting UV-vis. absorption spectrum showed the presence of $[2Fe-2S]^{2+}$ L-glutathione, confirming that L-glutathione [2Fe-2S] cluster can go through at least one complete round of reduction-oxidation without being degraded (Figure 2.8).
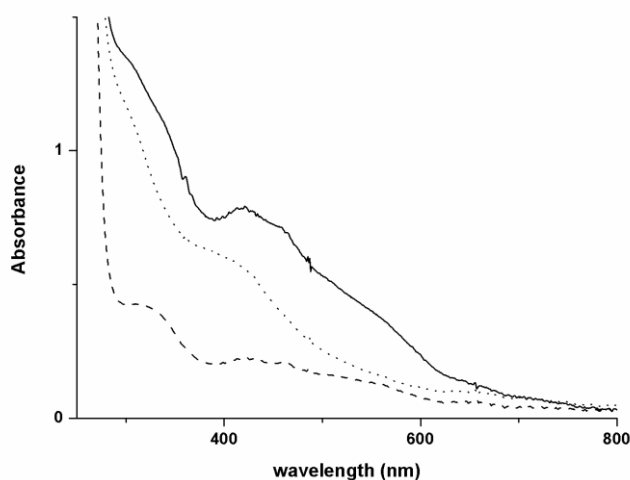

Figure 2.8 | UV-vis absorption spectra of L-glutathione [2Fe-2S] cluster before (full line) and after the reduction with sodium dithionite (dotted line). After the reduction, the mixture containing L-glutathione [2Fe-2S] cluster is purified on Sephadex G-10 resin. Upon the addition of hydrogen peroxide to the eluate, it is possible to observe again the bands at 420 and 450 nm distinctive of oxidized [2Fe-2S] cluster (dashed line).

Several rounds of reduction and oxidation with minimal degradation of the iron-sulfur cluster could be achieved by adding 0.5 equivalents of either reductant or oxidant with respect to the cluster concentration (Figure 2.9). The data were consistent with the ability of a tripeptide to mimic the cluster coordination and the redox activity of modern day ferredoxin proteins.
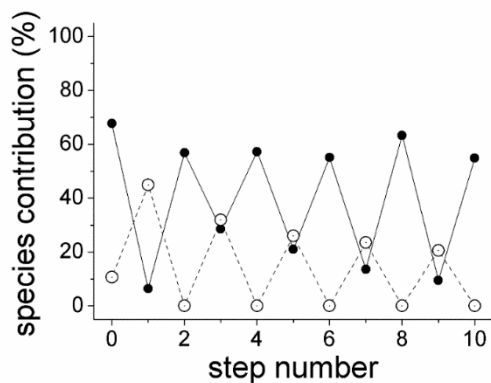
Figure 2.9 | L-glutathione $[2Fe–2S]^{2+}$ (filled circles) and $[2Fe–2S]^{1+}$ (open circles) clusters during reduction and oxidation steps determined by UV-vis. spectral decomposition.

The structure of L-glutathione suggests a path to a protoferredoxin. Molecular dynamic studies were performed on L-glutathione [2Fe-2S] cluster in order to gain insight into the molecular interactions present among each peptide coordinating the metal core. The results showed that the structure is stabilized by intra- and inter-molecular backbone electrostatic contacts. One particularly important interaction was between the α-amino group of glutamate and the α-carboxylate of glycine of an adjacent L-glutathione, as previously predicted.[30,33] This interaction suggested that the incorporation of a peptide bond at this position would not significantly disrupt the overall structure of the complex. To gain insight into whether duplication events leading to polymers of L-glutathione would result in [2Fe-2S] peptides of greater stability, peptides of six, nine, and twelve amino acids in length corresponding to polymers consisting of two, three, and four L-glutathione units were synthesized and evaluated for iron-sulfur cluster binding and stability. Two out of three peptides were capable of coordinating a [2Fe-2S] cluster with significantly greater stability than the tripeptide L-glutathione (Figure 2.10). Whereas the lifetime of L-glutathione [2Fe-2S] cluster was approximately 7 h, the hexapeptide and the dodecapeptide clusters persisted for 22 h and 35 h, respectively.
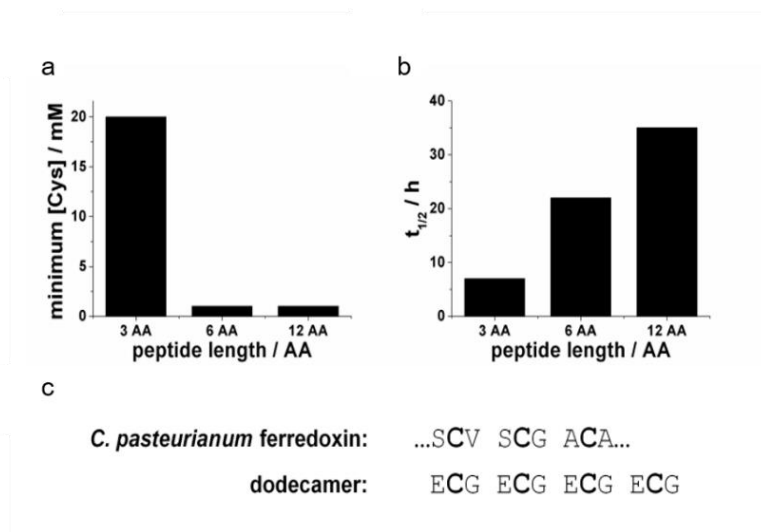
Figure 2.10 | Stability of peptide [2Fe-2S] cluster. a, The minimum concentration of cysteinyl ligand needed to stabilize a [2Fe-2S] for L-glutathione and peptides containing two and four L-glutathione units. b, The t1/2 of the peptide coordinated [2Fe-2S] increased as the length of the peptide increased. c, The dodecamer contained a spacing of cysteines similar to that of the first three cysteine ligands of *C. pasteurianum* ferredoxin. L-glutathione, hexapeptide, and dodecapeptide are abbreviated as 3 AA, 6 AA, and 12 AA, respectively.

However, while spectral decomposition showed that peptide [2Fe-2S] cluster was the predominant species for the hexa- and dodecapeptides, the nonapeptide slowly converted

to mononuclear iron complex. The stoichiometry between ligands and the [2Fe-2S] cluster was fully satisfied by either two hexapeptides or one dodecapeptide in a manner that did not disrupt the overall structure of the complex. Conversely, a single nonapeptide would not provide enough cysteine ligands for cluster coordination, and two nonapeptides would distort the structure of the complex. Nevertheless, duplication events leading to a protoferredoxin would not have required passage through a nonapeptide state. A duplication of a tripeptide followed by a duplication of the resulting hexapeptide would give the dodecapeptide without an intermediate, less stable nonapeptide transition.

2.3 Discussion

Similarities between geochemistry and biochemistry may seem to suggest a progression from abiotic to biotic systems that transitioned through an intermediate, hybrid, life-like structure dependent upon minerals for survival. The possibility is attractive, because minerals can catalyse RNA synthesis,[34] vesicle formation,[35] and mediate metabolic-like transformations.[36] However, minerals could have participated in the synthesis of starting material without functioning as an integral cellular component. A nascent cellular system could have instead exploited short iron-sulfur clusters coordinated by peptides. Model prebiotic reactions indicate that ribonucleotides, amino acids, and lipids can emerge from common, cyanosulfidic-based systems chemistry.[6] Peptide-coordinating iron-sulfur clusters could have then emerged along with a genetic system, whereas no analogous path exists for minerals. Additionally, [2Fe-2S] L-glutathione exhibits a redox activity similar to [2Fe-2S] ferredoxins. As previously mentioned, Eck and Dayhoff noted that the amino acid sequence of ferredoxin appears to consist of a repeating four amino acid sequence originating from Ala-Asp-Ser-Gly. That is, ferredoxin came into existence through a process of duplication events starting from a tetrapeptide that did not coordinate an iron-sulfur cluster but did provide some type of selective advantage to early life. At a later point, after repeated duplications, mutations resulted in properly positioned cysteines that then gave rise to iron-sulfur coordination. This theory fits well with the belief that short peptides are incapable of coordinating iron-sulfur clusters. Interestingly, the L-glutathione sequence γECG bears some resemblance to this ADSG ferredoxin motif. Glutamate and aspartate both contain carboxylate side-chains, cysteine is isosteric with serine, and both sequences end with glycine. Additionally, the spacing of six of the eight cysteines found in ferredoxin is the same as found in polymers made of repeating L-glutathione units (Cys-(Xxx)$_2$-Cys). Since repeating units of γECG results in polymers with increased [2Fe-2S] cluster stability, it now

seems possible that ferredoxin could have emerged from an analogous short peptide sequence that coordinated an iron-sulfur cluster from the beginning.

There are still some obvious difficulties in understanding how the first ferredoxin formed. The analysis of Eck and Dayhoff was on *Clostridium pasteurianum* ferredoxin, which forms a dimeric structure containing two [4Fe-4S] clusters. Prior studies in organic solvent indicated that [4Fe-4S] clusters are easier to form than [2Fe-2S] clusters,[13] and [4Fe-4S] ferredoxins are thought to be more ancient.[37] The smallest [4Fe-4S] peptide synthesized to date is 16 amino acids long.[4] Conversely, a much simpler tripeptide was found to stabilize a [2Fe-2S] cluster in aqueous solution under prebiotically reasonable conditions. To better understand which types of iron-sulfur peptides were possible, a more systematic investigation of peptide sequence space is needed. The structural model of [2Fe-2S] L-glutathione is consistent with stabilization that is largely mediated by backbone hydrogen bonding. This suggests that other tripeptide sequences will be capable of coordinating an iron-sulfur cluster. Comparing the stabilities and potential catalytic activities of a series of iron-sulfur peptides may help to reveal the role played by iron-sulfur clusters in the origins of life.

2.4 Materials and Methods

Materials. All reagents were bought from VWR Chemicals and Sigma Aldrich and used with no further purification. All the chromatographic columns were purchased from GE Healthcare. All the synthetic procedures were performed with a Schlenk line and Schlenk glassware under controlled N_2 flow. MilliQ water was distilled under N_2 atmosphere to remove oxygen from the solvent. The samples were preserved under N_2 atmosphere and transferred to anaerobically sealed Hellma quartz cuvettes, or NMR tubes capped with rubber septa for UV-vis. and NMR spectroscopy, respectively.

UV-visible spectroscopy. An Agilent 8453 UV-vis. diode array spectrophotometer was used to collect UV-vis. absorption spectra of freshly prepared solutions (integration time = 0.5 s, interval = 1 nm).

Spectral decomposition and [2Fe-2S] half-life estimation. R statistical computing software was used to perform all the analyses. All the recorded UV-vis. spectra (y) were fit using the least squares method to the linear combination of four reference spectra (s_1, s_2, s_3, s_4):

$$y = p_1 \cdot s_1 + p_2 \cdot s_2 + p_3 \cdot s_3 + p_4 \cdot s_4$$

Parameters were constrained to be null or positive and normalized ($p_1 + p_2 + p_3 + p_4 = 1$). The estimated values for each of the four parameters represented an approximation of the contribution of the corresponding reference spectrum to the overall UV-vis. spectrum. The half-life time of the [2Fe-2S] cluster was estimated by fitting the [2Fe-2S] cluster contribution in both its oxidation states over time. A decreasing curve after an initial maximum ended into a plateau. Based on the features of the profile, the fitting of the curve was obtained from one of the following equations:

$$[2Fe-2S](t) = \frac{A_1 - A_2}{1 + e^{(t-t_0)/dt}} + A_2$$

or

$$[2Fe-2S](t) = [2Fe-2S](0) + Ae^{-t/\tau}$$

The half-life time ($t_{1/2}$) could be estimated as:

$$t_{1/2} = t_0$$

or

$$t_{1/2} = \tau \ln(2)$$

Peptide [2Fe-2S] cluster synthesis. An aqueous solution of each peptide (40 mM, 1 mL) was prepared in a Schlenk round bottom flask under anaerobic conditions and the pH was adjusted to 8.6. Unless otherwise reported, sodium sulfide ($Na_2S \cdot 9H_2O$, final concentration 0.185 mM) and, subsequently, ferric chloride ($FeCl_3 \cdot 6H_2O$, final concentration 0.5 mM) were added to synthesize the [2Fe-2S] cluster. The same conditions were used for the pH screening using different solution pH (pH range between 6 and 11). 0.185 mM sodium sulfide and 0.5 mM ferric chloride were used for L-glutathione concentration screening (L-glutathione concentration range explored between 10 and 80 mM). 40 mM L-glutathione and 0.5 mM ferric chloride were used for sulfide screening (sulfide concentration range explored between 0.05 and 0.5 mM). 40 mM L-glutathione and 0.185 mM sodium sulfide were used for iron screening (iron concentration range explored between 0.05 and 2.0 mM). The spectrum of reduced L-glutathione [2Fe-2S] cluster was recorded after the [2Fe-2S] cluster was incubated at room temperature for over 180 min.

NMR spectroscopy. A 600 MHz Bruker Avance III spectrometer equipped with a triple resonance TCI cryogenic probe was used to acquire NMR spectra at 25 °C. One-dimensional ^{1}H-NMR spectra were recorded with an excitation sculpting water suppression pulse sequence that utilizes water selective 180° pulses. Two-dimensional ^{1}H homonuclear chemical shift correlation spectroscopy (COSY) and [1]H-^{13}C heteronuclear single quantum correlation (HSQC) experiments were performed for resonance assignment.

Fast protein liquid chromatography. An ÄKTA purifier P-900 system with a Frac-920 fraction collector was used to perform the analysis. L-glutathione [2Fe-2S] cluster was injected (500 µL) onto a L-glutathione-conjugated agarose-based column (GSTrap HP, 1 mL bed volume, GE Healthcare). The flow rate was set to 0.5 mL/min. A 40 mM aqueous solution of L-glutathione, pH 8.6 or MilliQ water, pH 8.6 were filtered and used as running buffers. The elution was monitored at 405 nm.

UV-vis. monitored redox experiments. In the reduction step, to a solution containing oxidized L-glutathione [2Fe-2S] cluster (1 mL, ~93 uM) was added under inert N_2 atmosphere an aqueous solution of the reducing agent, sodium dithionite (2 uL, final concentration 50 uM), and UV-vis. absorption spectrum of the reduced L-glutathione [2Fe-2S] cluster was recorded. Next, an aqueous solution of the oxidizing agent, hydrogen peroxide (2 uL, final concentration 50 uM) was added and UV-vis. absorption spectrum of the oxidized L-glutathione [2Fe-2S] cluster was recorded. For the redox cycle studies, each reduction and oxidation step was consecutively repeated and UV-vis. absorption spectra collected after every addition, until distinctive UV-vis. spectral features of L-glutathione [2Fe-2S] cluster were not anymore detectable.

Gel filtration chromatography. A solution of reduced L-glutathione [2Fe-2S] cluster, obtained by the addition of sodium dithionite as previously described, was loaded under inert N_2 atmosphere onto a Sephadex G-10 column, using a filtered 40 mM aqueous solution of L-glutathione, pH 8.6 as running buffer. The collected fractions were transferred to an anaerobically sealed cuvette and UV-vis. absorption spectra recorded. The reduced L-glutathione [2Fe-2S] cluster was subsequently oxidized by addition of hydrogen peroxide as previously described, and UV-vis. absorption spectra recorded.

Determination of $t_{1/2}$ for iron-sulfur clusters. [2Fe-2S] clusters coordinated by L-glutathione, hexapeptide and dodecapeptide were freshly synthesized as previously described, keeping fixed the amount of sodium sulfide and ferric chloride (final concentrations 0.1 mM and 0.25 mM, respectively) and using different concentrations of ligand (cysteinyl ligand final concentration of 20, 10, 5, 2.5 and 1 mM, *i.e.* L-glutathione concentrations tested 20, 10, 5, 2.5, 1 Mm - 1 cysteinyl residue per molecule; hexapeptide concentrations tested 10, 5, 2.5, 1.25, 0.5 mM - 2 cysteinyl residues per molecule; dodecapeptide concentrations tested 5, 2.5, 1.25, 0.625, 0.25 mM - 4 cysteinyl residues per molecule). Each solution was time-course monitored by UV-vis. spectrophotometry and decomposed as previously described.

Solid phase peptide synthesis. Peptide synthesis was performed as previously described.[38] *N,N*-dimethyl formamide (DMF) and dichloromethane (DCM) were used as solvents for washing and coupling steps. Preloaded fluorenylmethyloxycarbonyl-glycyl Wang resin (Fmoc-Gly Wang) was chosen to obtain a free carboxylic group upon cleavage. Fmoc-glycine (Fmoc-Gly-OH), trityl-protected Fmoc-L-cysteine (Fmoc-L-Cys(Trt)-OH), and tert-butyl-protected Fmoc-L-glutamic acid (Fmoc-L-Glu-OtBu) were chosen as building blocks. The peptide chain was elongated by sequential Fmoc deprotection of the last amino acid bound to the resin and Fmoc-Xxx-OH (Xxx = Gly, Cys, and Glu) coupling. For Fmoc deprotection the peptide-bound resin was incubated with 20% (v/v) solution of piperidine in DMF. For coupling steps, an excess (Fmoc-Xxx-OH:peptide-bound resin, 3:1) of the Fmoc-amino acid derivative was used. Apart from Fmoc-L-Cys(Trt)-OH, Fmoc-amino acid derivatives were mixed with hydroxyl-benzotriazole (HOBt), *O*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU), and *N,N*-diisopropylethylamine (DIPEA) before adding to the resin and stirring for 1 hour. Fmoc-L-Cys(Trt)-OH was activated with a HOBt/*N,N'*-diisopropylcarbodiimide (DIC) mixture. After each coupling step, the peptide-bound resin was washed with DMF and DCM to remove the coupling reagents leftover. Once the elongation was complete, the polymers were deprotected and cleaved from the resin with trifluoroacetic acid (TFA) adding 1,2-ethanedithiol (EDT), water, and triisopropyl silane (TIS) as scavengers (volume ratio 92.5:2.5:2.5:2.5). The solution was subsequently precipitated with a cold solution of diethyl ether/petroleum ether (70:30% (v/v)) and dried under N_2 atmosphere.

2.5 Bibliography

1. Alva, V., Söding, J. & Lupas, A. N. A vocabulary of ancient peptides at the origin of folded proteins. *Elife* **4,** 1–19 (2015).

2. Goldman, A. D., Baross, J. A. & Samudrala, R. The enzymatic and metabolic capabilities of early life. *PLoS One* **7,** 1–7 (2012).

3. Eck, R. V. & Dayhoff, M. O. Evolution of the Structure of Ferredoxin Based on Living Relics of Primitive Amino Acid Sequences. *Science (80-.).* **152,** 363–366 (1966).

4. Mulholland, S. E., Gibney, B. R., Rabanal, F. & Dutton, P. L. Determination of nonligand amino acids critical to [4Fe-4S](2+/+) assembly in ferredoxin maquettes. *Biochemistry* **38,** 10442–10448 (1999).

5. Belmonte, L. & Mansy, S. S. Metal Catalysts and the Origin of Life. *Elements* **12,** 413–418 (2016).

6.   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.* **7,** 301–307 (2015).

7.   Keller, M. A., Turchyn, A. V & Ralser, M. Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean. *Mol. Syst. Biol.* **725,** 1–12 (2014).

8.   Anbar, A. D. Elements and Evolution. *Science (80-.).* **322,** 1481–1483 (2008).

9.   Beinert, H. Iron-sulfur proteins: ancient structures, still full of surprises. *J. Biol. Inorg. Chem.* **5,** 2–15 (2000).

10.   Daniel, R. M. & Danson, M. J. Did primitive microorganisms use nonhem iron proteins in place of NAD/P? *J. Mol. Evol.* **40,** 559–563 (1995).

11.   Gil, R., Silva, F. J., Pereto, J. & Moya, A. Determination of the Core of a Minimal Bacterial Gene Set. *Microbiol. Mol. Biol. Rev.* **68,** 518–537 (2004).

12.   Lill, R. Function and biogenesis of iron-sulphur proteins. *Nature* **460,** 831–838 (2009).

13.   Venkateswara Rao, P. & Holm, R. H. Synthetic Analogues of the Active Sites of Iron-Sulfur Proteins. *Chem. Rev.* **104,** 527–559 (2004).

14.   Koch, A. L. & Schmidt, T. M. The First Cellular Bioenergetic Process: Primitive Generation of a Proton-Motive Force. *J. Mol. Evol.* **33,** 297–304 (1991).

15.   Falkowski, P. G. From Light to Life. *Orig. Life Evol. Biosph.* (2015). doi:10.1007/s11084-015-9441-6

16.   van der Gulik, P., Massar, S., Gilis, D., Buhrman, H. & Rooman, M. The first peptides: the evolutionary transition between prebiotic amino acids and early proteins. *J. Theor. Biol.* **261,** 531–9 (2009).

17.   Danger, G., Plasson, R. & Pascal, R. Pathways for the formation and evolution of peptides in prebiotic environments. *Chem. Soc. Rev.* **41,** 5416–5429 (2012).

18.   Mulholland, S. E., Gibney, B. R., Rabanal, F. & Dutton, P. L. Characterization of the fundamental protein ligand requirements of [4Fe- 4S](2+/+) clusters with sixteen amino acid maquettes. *J. Am. Chem. Soc.* **120,** 10296–10302 (1998).

19.   Qi, W. *et al.* Glutathione Complexed Fe–S Centers. *J. Am. Chem. Soc.* **134,** 10745–10748 (2012).

20.   Powner, M. W., Zheng, S. L. & Szostak, J. W. Multicomponent assembly of proposed DNA precursors in water. *J. Am. Chem. Soc.* **134,** 13889–13895 (2012).

21.   Parker, E. T. *et al.* Primordial synthesis of amines and amino acids in a 1958 Miller

H2S-rich spark discharge experiment. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 5526–5531 (2011).

22. Parker, E. T. *et al.* 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.* **41,** 201–212 (2011).

23. Sugiura, Y. & Tanaka, H. Iron-sulfide chelates of some sulfur-containing peptides as model complex of non-heme iron proteins. *Biochem. Biophys. Res. Commun.* **46,** 335–340 (1972).

24. Sun, W., Ueyama, N. & Nakamura, A. Stabilization of hydrolytically labile iron(II)–cysteine peptide thiolate complexes in aqueous triton X-100 micelle solution: Spectroscopic properties mimicking of reduced rubredoxin. *Biopolymers* **46,** 1–10 (1998).

25. Nakata, M. *et al.* Chemical simulation of rubredoxin by iron(II)/tetrapeptide complexes in aqueous Triton X-100 micelle solution. *Biochim. Biophys. Acta* **788,** 306–312 (1984).

26. Greenwood, N. N. & Earnshaw, A. *Chemistry of the Elements.* (Butterworth-Heinemann, 1997).

27. Lide, D. R. CRC Handbook of Chemistry and Physics. 3485 (2003). doi:978-1466571143

28. Neville, R. G. The Oxidation of cysteine by Iron and Hydrogen Peroxide. *J. Am. Chem. Soc.* **79,** 2456–2457 (1957).

29. Pecci, L., Montefoschi, G., Musci, G. & Cavallini, D. Novel findings on the copper catalysed oxidation of cysteine. *Amino Acids* **13,** 355–367 (1997).

30. Qi, W. *et al.* Glutathione-complexed iron-sulfur clusters. Reaction intermediates and evidence for a template effect promoting assembly and stability. *Chem. Commun.* **49,** 6313 (2013).

31. Machonkin, T. E., Westler, W. M. & Markley, J. L. Paramagnetic NMR spectroscopy and density functional calculations in the analysis of the geometric and electronic structures of iron-sulfur proteins. *Inorg. Chem.* **44,** 779–797 (2005).

32. Beinert, H., Holm, R. H. & Münck, E. Iron-sulfur clusters: nature's modular, multipurpose structures. *Science* **277,** 653–659 (1997).

33. Li, J., Pearson, S. A., Fenk, K. D. & Cowan, J. A. Glutathione-coordinated [2Fe-2S] cluster is stabilized by intramolecular salt bridges. *J. Biol. Inorg. Chem.* **20,** 1221–

1227 (2015).

34. Ferris, J. P. Mineral Catalysis and Prebiotic Synthesis: Montmorillonite-Catalyzed Formation of RNA. *Elements* **1,** 145–150 (2005).

35. Hanczyc, M. M., Mansy, S. S. & Szostak, J. W. Mineral surface directed membrane assembly. *Orig. Life Evol. Biosph.* **37,** 67–82 (2007).

36. Cody, G. D. *et al.* Primordial carbonylated iron-sulfur compounds and the synthesis of pyruvate. *Science (80-.).* **289,** 1337–1340 (2000).

37. Fitch, W. M. & Bruschi, M. The evolution of prokaryotic ferredoxins--with a general method correcting for unobserved substitutions in less branched lineages. *Mol. Biol. Evol.* **4,** 381–394 (1987).

38. Jensen, Knud J.; Tofteng Shelton, Pernille; Pedersen, S. L. Peptide Synthesis and Applications. *Methods Mol. Biol.* **1047,** 252 (2013).

Chapter 3 - UV light-driven prebiotic synthesis of iron-sulfur clusters

3.1 Introduction

The conversion of free L-cysteine to free L-alanine is a fundamental step for the formation of biological Fe-S clusters. Cysteine desulfurase is a pyridoxal phosphate-dependent homodimeric enzyme that catalyses the beta-elimination of thiols from free L-cysteine to yield L-alanine and elemental sulfur or hydrogen sulfide. Sulfur removal is then followed by the formation of a persulfide[1] on the active site cysteine.[2] Cysteine desulfurase represents a broad class of proteins that use L-cysteine as source of sulfur for the biosynthesis of a variety of sulfur-containing prosthetic groups, such as iron-sulfur clusters, as well as biotin, thiamine and thionucleotides.[3] On the other hand, the chemical removal of elemental sulfur or hydrogen sulfide from simple thiols has been investigated[4–6] in the past years, exploiting ultraviolet (UV) radiation inducing photolytic radical reactions.[7] However, an efficient process for sulfur mobilization and related cofactors synthesis adaptable to plausible prebiotic conditions has not been proposed yet.

Since Sunlight can be considered the Earth's largest energy source[8] and continues to be exploited in biology today for the synthesis of molecules such as vitamin D3[9] and for biological processes such as photosynthesis to convert light into chemical energy, photochemical reactions must be considered in the chemical evolution of biomolecules that eventually led to life.[8] About 4.5 billion years ago the UV environment would have been characterized by a broadband input, and wavelengths below 204 nm and 168 nm would have been shielded out by atmospheric CO_2 and water, respectively.[10] Moreover, the absence of protective biogenic oxygen in the primordial terrestrial atmosphere exposed the early Earth surface to much more intense UV flux than the present day.[10] UV light-driven chemical reactions have been shown to generate the prebiotic molecules needed for life.[11–13] John Sutherland suggested UV radiation as a fundamental component in the prebiotic synthesis of amino acid, sugar, nucleotide, and lipid precursors, and thus all the cellular subsystems could have risen simultaneously through common photochemistry.[12] It is thus reasonable to expect that UV light could have driven many other non-spontaneous chemical pathways on the early Earth.

The prebiotic chemical synthesis of iron-sulfur clusters in aqueous solution is typically accomplished by mixing Fe^{3+}, S^{2-}, and suitable thiolate ligands under anaerobic

conditions.[14–16] However, the high volatility of hydrogen sulfide,[17,18] mainly present in volcanic environments, and iron ions, mostly available in the reduced form,[19] challenge the understanding of how iron-sulfur cluster synthesis could have been driven on the early Earth. Therefore, studies investigating the effect of ultraviolet light as the driving force for the synthesis of iron-sulfur peptides could be fundamental to probe the compatibility of iron-sulfur clusters with prebiotically plausible environmental conditions.

It was previously shown in Chapter 2 that the tripeptide L-glutathione is able to coordinate a redox active [2Fe-2S] cluster in aqueous solution and that oligomerization of this tripeptide sequence to the dodecapeptide leads to a protoferredoxin with increased cluster stability. However, it is not clear if L-glutathione is the shortest peptide unit that could coordinate an iron-sulfur cluster, nor if the coordination is sequence-specific and sequence-dependent. Even though a few other short Cys-containing peptides and thiol-containing organic molecules have been studied as coordinating ligands for iron-sulfur clusters,[14,20–22] no effort has been put so far into understanding the chemical features required to efficiently coordinate and stabilize an iron-sulfur cluster. A full screening of different cysteine-containing peptide sequences and branched, bulky or aliphatic organic thiols has been undertaken to allow prebiotic chemists and molecular biologists to have a deeper knowledge in iron-sulfur clusters primordial and extant biosynthesis.

Considering the holistic view as the most reliable approach to the studies on the origin of life, a deep understanding of the compatibility among the metal systems studied in this Thesis and the other relevant components of cellular life would be needed to test the plausibility of such metal models on the early Earth. In particular, since iron-sulfur proteins are prevalently embedded within lipid bilayers, the compatibility of primitive cells, or protocells, and iron-sulfur clusters should be subject to further investigations. Models of the first primordial cells are used to study the origin and early evolution of life. The protocell membrane is a compartment boundary which provides a selective barrier between the cell and its environment, allowing the influx of nutrients and the efflux of waste while stably encapsulating the macromolecular contents of the protocell.[23] Fatty acids spontaneously self-assemble into vesicles that can serve as protocell membranes. A pool of fatty acids likely existed on the early Earth as they have been found in meteorites[24] and can be abiotically synthesized in a prebiotic environment by Fischer-Tropsch synthesis in hydrothermal vents.[25] Fatty acid vesicles have many physical properties such as

permeability and stability[26,27] that are appropriate for a protocell with minimal (or no) evolved biochemical machinery. To a certain extent, one can tune the permeability, stability, and dynamics of fatty acid vesicles by changing the acyl-chain length and saturation state of the fatty acid components or by adding, for example, fatty alcohols, fatty esters, and polycyclic aromatic hydrocarbons.[27] In general, fatty acid membranes are much more permeable to small charged molecules than are phospholipid membranes. However, the incompatibility with metal ions and the limitation in the range of pH which can be exploited represent the major drawbacks. That is, for more advanced but less prebiotically plausible artificial cell models, more complex membranes, such as phospholipid-based vesicles, are required to broaden the working pH range and the preservation of the encapsulated molecules within the vesicle.[26]

One of the main focus in the prebiotic research on compartmentalization is to understand and reproduce prebiotically plausible reactions within membranes. Encapsulation would be required for two reasons: first, it generates a boundary between self and non-self, needed to exclude parasitic molecules and to allow Darwinian evolution, and second, it allows primordial organisms to retain the molecules necessary for their metabolic processes.[28] Recent studies have shown that non-enzymatic RNA polymerization is feasible within vesicles,[29] and that peptide-driven catalytic activity is enhanced by compartmentalization.[30] However, the efficiency of prebiotic reactions performed in bulk solutions, especially if driven by UV light as in the case of Sutherland's prebiotic pathways[12], could be strongly influenced by the encapsulation efficiency of each component, which is generally low, and the penetration of UV radiation through membrane walls. Indeed, UV light may be partially shielded, reflected or absorbed by the lipid layer, thus resulting in a reduced effect on the encapsulated molecules. Moreover, in the case of metal-based reactions, an additional issue related to the induced degradative effect of metal ions on the stability of fatty acid vesicles need to be addressed.[31] Even though the effect of 360 nm UV radiation on fatty acid vesicles has been recently studied,[32] not much effort has been put into testing the effect of short-wave UV light on the stability of prebiotic membranes in order to evaluate the compatibility of lipid layers with the UV-driven prebiotic synthesis of biomolecules.

On the other hand, the vesicle degradation induced by metal ions has been shown to be strongly reduced by the chelate effect of specific molecules behaving as coordinating ligands.[31] Indeed, Mg^{2+}-dependent non-enzymatic RNA polymerization has been shown to be efficient within fatty acid vesicles in the presence of citrate as chelating moiety for the divalent cation.[28,31] However, the light-driven prebiotic reactions suggested so far for the

synthesis of amino acid, lipid and nucleotide precursors exploit a more energetic radiation (~250 nm)[12] and the most abundant metal ion on the prebiotic Earth was presumably Fe^{2+}.[19] That is, studying whether iron-sulfur cluster synthesis could be performed within vesicles exploiting UV light could be a fundamental step in the understanding of the plausibility of iron-sulfur systems on the prebiotic Earth.

3.2 Results

The experiments previously described (see Chapter 2) were based on L-glutathione as a model prebiotic tripeptide to coordinate and stabilize [2Fe-2S] clusters. However, since L-glutathione may not have been present on the primordial Earth, a library of peptides was investigated to identify the sequences able to coordinate and stabilize iron-sulfur clusters in aqueous solution. Therefore, several tripeptides were synthesized either keeping fixed the CysGly moiety and changing the *N*-terminus amino acid, and keeping fixed the GlyCys moiety and changing the *C*-terminus amino acid. Then palindrome sequences with the cysteinyl residue in the middle and the same residue both at the *N*- and the *C*-termini were tested, together with few peptides where the *N*- and the *C*-termini were occupied by different amino acids. Finally, the cysteinyl residue was moved at the *C*- or at the *N*-termini (both in the free or in the protected form) by synthesizing peptides composed by 2 or 3 residues. The residues other than the cysteine were selected in order to include in the investigation at least one component of each amino acid class (polar, aliphatic, aromatic, or charged residues).

That is, 35 different cysteine-containing tripeptides were synthesized (Table A1.1), then mixed with ferric and sulfide ions and tested for their ability to coordinate [2Fe-2S] clusters exploiting UV-vis. spectrophotometry analyses (Table 3.1). Such peptide screen was designed to probe the influence of vicinal groups on the coordination of a [2Fe-2S] cluster by the cysteinyl residue.

Table 3.1. Absorbance maxima for synthesized [2Fe-2S] clusters with different tripeptides. The pH value of each tripeptide solution was adjusted to the pK_a value of the thiol/thiolate couple.

Ligand		Absorbance maximum (nm)	Absorbance maximum (nm)
GCX	GCA	413	459
	GCE	415	456
	GCK	422	447
	GCQ	430	455
	GCS	420	454
	GCT	419	453
	GCV	418	457
XCG	ACG	421	457
	DCG	415	451
	βDCG	417	459
	ECG	425	455
	γECG	423	457
	FCG	422	456
	KCG	424	451
	NCG	423	451
	QCG	418	454
	SCG	420	447
	TCG	419	447
	VCG	422	456
	WCG	-	-
XCX	ACA	417	449
	ECE	423	459
	GCG	421	455
	KCK	425	451
	SCS	418	447
	TCT	421	455
	VCV	418	457
XCY	ACT	425	453
	SCA	422	452
Protected and unprotected CXG	CGG	-	-
	AcCGG	413	456
	CγEG	-	-
	AcCγEG	415	462
XGC	GGC	415	455
	γEGC	416	453
Negative control	GMG	-	-

Since cysteine appears fundamental for the coordination of ferric ions in iron-sulfur clusters, first the cysteine residue was removed and substituted with another sulfur-

47

containing amino acid, *i.e.* methionine. As expected, no [2Fe-2S] cluster was observed (Figure A1.1). The second test was focused on determining the importance of the cysteine position for cluster coordination. The L-glutathione sequence (γECG) was scrambled maintaining the glutamic acid γ-peptide bond. The data showed that moving Cys at the *C*-terminus, *i.e.* γEGC, but not at the *N*-terminus, *i.e.* CγEG, was enough to obtain the formation of a [2Fe-2S] cluster. Since the presence of a free amino group close to the metal core induces a strong decrease in the pK_a of the thiol (ΔpK_a ~1.5), thus avoiding the formation of a [2Fe-2S] cluster, the amino terminus of CγEG was acetylated. Upon protection of the *N*-terminus, the ability to coordinate a [2Fe-2S] cluster was recovered (Figure A1.2). The following step was to investigate the importance of the γ-peptide bond. The introduction of a generic α-peptide bond in ECG did not inhibit cluster formation and stabilization (Figure A1.3). Moreover, substituting the glutamate with an aspartate, thus shortening the side-chain with either an α- or β-peptide bond, *i.e.* DCG and βDCG, had no effect on cluster coordination, still leading to the formation of a [2Fe-2S] cluster (Figure A1.3). Finally, a broad variety of cysteine containing tripeptides more diverged from the sequence of L-glutathione were screened. The presence of hydrophobic (valine, alanine, phenylalanine), polar (asparagine, glutamine, serine, threonine), and charged (aspartic acid, glutamic acid, lysine) residues did not interfere with the ability of the adjacent cysteine residue to coordinate the [2Fe-2S] cluster. WCG was the only cysteine containing sequence that failed to coordinate a [2Fe-2S] cluster. Besides the fact that tryptophan is unlikely to have been a prebiotic amino acid,[33] the size of the tryptophan side-chain likely hindered cluster coordination. Finally, [2Fe-2S] cluster coordination was recovered on the peptide CGG by simply acetylating the free amino group at the *N*-terminus as previously reported with the other *N*-terminal cysteine containing peptides (CγEG and acetylated CγEG). Since the synthesis of all the [2Fe-2S] clusters depended on the presence of deprotonated thiol groups in cysteine side-chains, the pH value of each anaerobic peptide solution was adjusted to the pK_a of the thiol/thiolate couple, which changed depending on the peptide sequence. Since the first pK_a of H_2S is 7.05[34] and H_2S is highly volatile, thus not easily available to react, a slightly alkaline pH is needed to trap in solution HS^-. While amino-terminal cysteines exhibited pKa values lower than 7, the pK_a values were between 7.3 and 8.1 for carboxy-terminal or central cysteines (Figure A1.4 and Table A1.2). Therefore, it appears clear that that most sequences are competent for the coordination of a [2Fe-2S] cluster, so that there is no sequence-dependence in the ability of a cysteine-containing tripeptide to bind an iron-sulfur cluster.

Having shown that a simple tripeptide is able to bind an iron-sulfur cluster, the next step was to understand if a tripeptide is the shortest unit able to coordinate an iron-sulfur cluster. Therefore, the ability of smaller ligands to bind a [2Fe-2S] cluster in aqueous solution (Table A1.3) was investigated. When sulfide and ferric ions were added to the dipeptides GC and acetylated CG solutions, the typical spectroscopic features of the [2Fe-2S] clusters were recorded. Accordingly, *N*-acetylated L-cysteine and *N*-acetylated L-cysteine methyl ester but not L-cysteine methyl ester nor unprotected free cysteine were similarly able of binding an iron-sulfur cluster (Figure A1.5). Following previous studies on iron-sulfur cluster synthesis,[14,35] non-genetically encoded thiolate ligands were then tested, including cysteamine, 3-mercaptopropionic acid, 2-mercaptoethanol, 2-mercaptopyrimidine, dithiothreitol, 1,2-ethanedithiol, and dihydrolipoic acid. Of these ligands, only 3-mercaptopropionic acid, 2-mercaptoethanol, and dithiothreitol were shown to coordinate an iron-sulfur cluster (Table 3.2).

Table 3.2. Absorbance maxima for [2Fe-2S] clusters coordinated to Cys-containing dipeptides, protected and unprotected L-cysteine and small molecule thiols.

Ligand		Absorbance maximum (nm)	Absorbance maximum (nm)
Dipeptides	CG	-	-
	AcCG	418	452
	GC	419	453
Protected and unprotected Cys	L-cysteine	-	-
	L-cysteine methyl ester	-	-
	N-acetyl L-cysteine	420	454
	N-acetyl L-cysteine methyl ester	420	451
Small thiolates	DL-dithiothreitol	419	457
	2-mercaptoethanol	420	455
	3-mercaptopropionic acid	419	451

Indeed, in both cysteamine and the tautomeric form of 2-mercaptopyridine a -NH_2 group is located in close proximity to the binding thiolate, while 1,2-ethanedithiol and dihydrolipoic acid exhibit two coordinating thiolates not sufficiently spaced by respectively two and three methylene groups, as previously reported[36–38]. However, although three small thiolate ligands were found capable of coordinating a [2Fe-2S] cluster, the concentration of the ligand required for cluster synthesis was higher than that for the studied cysteine-containing tripeptides. In particular, 240 mM 2-mercaptoethanol was needed to stabilize the iron-sulfur cluster in the presence of a fixed amount of sulfide and ferric ions, while only 40 mM L-

glutathione was required under the same conditions. It is possible that longer molecules are able to better protect the iron-sulfur cluster from hydrolytic attack, as suggested from previously discussed data on L-glutathione polymers (see Chapter 2). Taken together, simpler ligands than L-glutathione were identified to be able to coordinate a [2Fe-2S] cluster in aqueous solution, suggesting that iron-sulfur clusters could be easily synthesized in the presence of a wide variety of thiolate ligands, enhancing their prebiotic plausibility on the early Earth.

In order to understand if and how UV light coming from the young Sun would have affected the synthesis of iron-sulfur clusters on the prebiotic Earth, a solution containing L-glutathione [2Fe-2S] cluster was irradiated with 254 nm UV light (Figure 3.1).
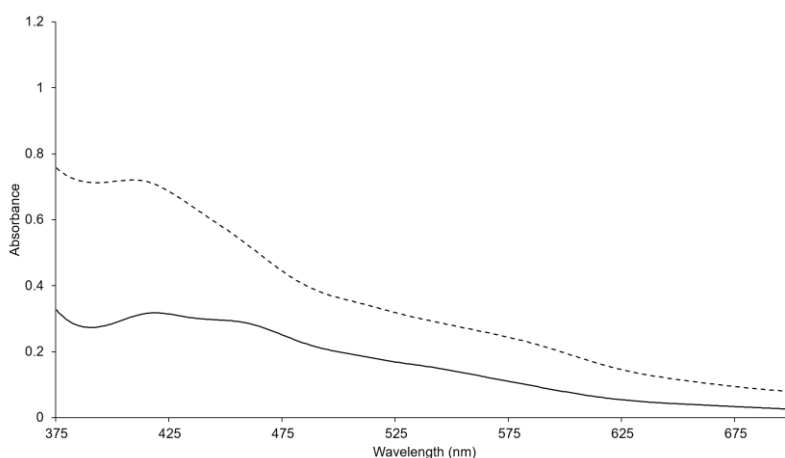
Figure 3.1 | UV-vis. absorption spectra of L-glutathione [2Fe-2S]$^{2+}$ cluster solution exposed to UV light. The [2Fe-2S] cluster (solid line) is converted to a higher nuclearity iron-sulfur cluster (dashed line) by UV irradiation.

UV radiation induced an evident change in the UV-vis. spectrum of the [2Fe-2S] cluster without any apparent precipitation or pH alteration suggesting a possible conversion of the starting material into an iron-sulfur cluster with different nuclearity.

The facile conversion between Fe-S structures have been previously observed both in proteins and in synthetic analogues. The exposition of a [4Fe-4S] cluster to oxygen or ferricyanide gives rise to a [2Fe-2S] cluster, usually through a [3Fe-4S] intermediate,[39–41] while reducing a [2Fe-2S] cluster with dithionite induce the formation of a [4Fe-4S] cluster.[42,43] On the other hand, Kim and Lippard have independently conducted detailed studies on iron-nitrosyl species, which can be converted into [2Fe-2S] clusters releasing nitrosyl groups by the addition of thiolates or by UV light treatment in the presence of elemental sulfur.[44–46] However, these conversion pathways have been conducted exploiting organic solvents and non-genetically encoded organic thiolates, far apart from the proposed

prebiotic conditions and the plausible coordinating moieties needed for the synthesis of primordial biological frameworks. Moreover, no studies have been conducted on UV light-driven conversion of mononuclear iron-thiolate complexes into higher nuclearity iron-sulfur clusters in prebiotically plausible conditions without further addition of elemental sulfur. Therefore, based on the preliminary results shown on UV irradiated [2Fe-2S] cluster, an anaerobic solution containing only iron ions and a thiolate ligand, *i.e.* the model tripeptide L-glutathione, was studied upon UV light treatment by UV-vis. and Mössbauer spectroscopies. Irradiation of a solution containing only Fe^{3+} and L-glutathione to 254 nm UV light led to the synthesis of polynuclear iron-sulfur clusters (Figure 3.2).
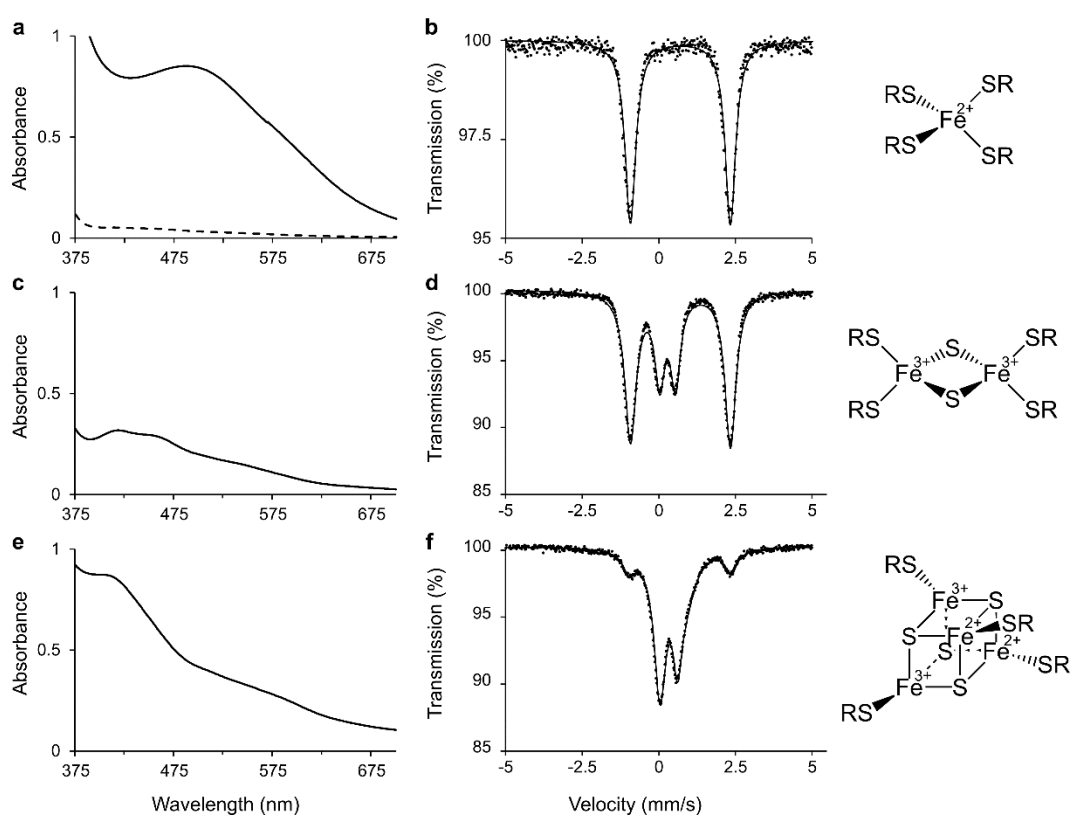
Figure 3.2 | UV light-driven prebiotic synthesis of iron-sulfur clusters coordinated by thiolates. The mixture of Fe^{3+} ions and L-glutathione results in an initial UV-visible spectrum consistent with an oxidized mononuclear Fe^{3+}-L-glutathione complex (a, solid line), which is rapidly (over 180 s) reduced by free thiolates to a Fe^{2+}-L-glutathione complex (a, dashed line, and b). Exposure to 254 nm UV light results in the formation of a [2Fe-2S] cluster (c and d) and then a [4Fe-4S] cluster (e and f). a, c, and e: UV-vis. absorption spectra; b, d, and f: Mössbauer spectra. d shows a mixture of mononuclear Fe^{2+} and $[2Fe-2S]^{2+}$. f shows a mixture of mononuclear Fe^{2+}, $[2Fe-2S]^{2+}$, and $[4Fe-4S]^{2+}$.

Since the standard protocols for the *in vitro* synthesis of iron-sulfur proteins in aqueous solution utilize Fe^{3+} salts, ferric ions were used instead of ferrous ions. Prior to UV

51

light irradiation, the solution of Fe^{3+}-L-glutathione complex was violet showing a UV-vis. absorption spectrum similar to that of the mononuclear iron protein rubredoxin (Figure 3.2a, solid line). Since the oxidized mononuclear iron-L-glutathione complex was not stable, it underwent rapid reduction due to the presence of free thiolates (Figure 3.2a, dashed line). Mössbauer spectroscopy evidenced an isomeric shift (IS) and quadrupole splitting (QS) of 0.68 mm/s and 3.26 mm/s, respectively, consistent with the appearance of a reduced mononuclear Fe^{2+} complex (Figure 3.2b and Table A1.4).[47]

Exposure to 254 nm UV light for 30 s turned the solution red and gave a UV-vis. absorption spectrum consistent with that of a [2Fe-2S] ferredoxin (Figure 3.2c). The presence of a [2Fe-2S] cluster was confirmed by Mössbauer studies. 21% of the Mössbauer spectrum, expressed as molar contribution, fit a [2Fe-2S]$^{2+}$ cluster with an IS of 0.27 mm/s and QS of 0.52 mm/s, whereas the remaining signals were due to the presence of the reduced iron-L-glutathione complex (Figure 3.2d). Upon further (up to 180 s) UV exposure of the mixture, the solution turned brown and exhibited a UV-vis. absorption spectrum similar to [4Fe-4S] ferredoxins (Figure 3.2e). Mössbauer spectra were fit to 10mol% [4Fe-4S]$^{2+}$ (IS=0.48 mm/s; QS=1.06 mm/s), 64mol% [2Fe-2S]$^{2+}$ (IS=0.31 mm/s; QS=0.57 mm/s), and 26mol% Fe^{2+} (IS=0.68 mm/s; QS=3.32 mm/s) (Figure 3.2f). Since the extinction coefficient of [2Fe-2S] clusters is generally lower than that of [4Fe-4S] clusters[42], the absorption signal of the [4Fe-4S] cluster dominated the UV-vis. spectrum. In the absence of UV radiation, no conversion was detected within 24 h (Figure 3.3).
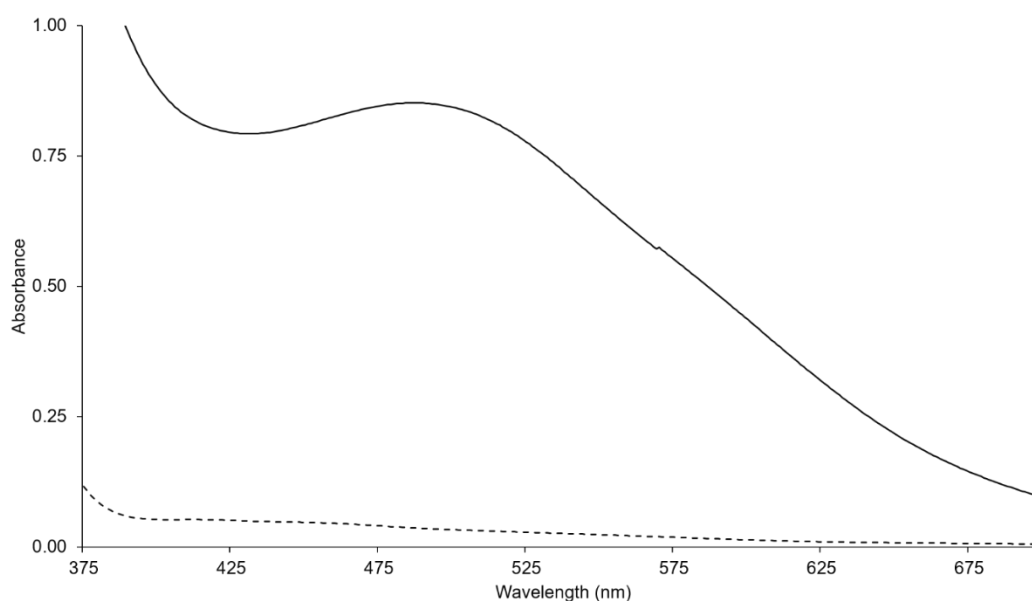
Figure 3.3 | UV-vis. absorption spectra of L-glutathione solution with Fe^{3+} not exposed to UV light. In the absence of irradiation, the Fe^{3+} mononuclear complex (solid line) undergoes reduction to the Fe^{2+} mononuclear complex (dashed line) without conversion to higher nuclearity clusters up to 24 h.

Since preliminary experiments were run with L-glutathione as model peptide, solutions of 2-mercaptoethanol, 3-mercaptopropionic acid, *N*-acetyl cysteine, *N*-acetyl cysteine methyl ester, and some peptides such as ACG, KCG, GCK, TCT, and ACT were investigated in the presence of ferric ions to test the ability to lead to cluster formation upon exposure to UV light. All the studied thiolates irradiated with 254 nm UV light yielded the formation of [2Fe-2S] clusters as confirmed by UV-vis. spectroscopy (Figure A1.6-A1.14).

High nuclearity iron-sulfur clusters contain inorganic sulfide within the metal core. However, no inorganic sulfide ions were added to the solution before exposure to UV light, the reaction was probed by ^{1}H NMR spectroscopy to determine the source of the sulfide (Table A1.5-A1.6). When a solution of L-glutathione was exposed to UV light, the side-chain methylene resonance of cysteine at 2.9 ppm decreased over time and a new doublet related to the methyl group of an alanine at 1.3 ppm concomitantly appeared and increased (yield ~60%, Figure 3.4a). The conversion of the cysteine of L-glutathione (γECG, $[M+H]^+$=308.2 m/z) to an alanine (γEAG, $[M+H]^+$=276.2 m/z) was confirmed also by mass spectrometry (Figure A1.15-A1.16). Exposure to UV light also induced the formation of the oxidized disulfide form of the thiolates (*i.e.* cystines), which were more slowly converted to alanine upon further irradiation (Figure A1.17). However, if hypophosphite[48] as a prebiotically plausible reductant was added to the solution, L-glutathione was completely converted by UV light to the corresponding alanine containing tripeptide (Figure 3.4b).

The conversion of cysteine into alanine was analogously obtained in the presence (Figure A1.18) of iron ions, showing that the conversion was not inhibited by iron ions. Therefore, exposure to UV light would lead to release of sulfide ions in solution upon radical-induced cleavage of the cysteine C-S bond. Thus, light-induced desulfurization increased sulfide concentration and induced conversion from mononuclear, to [2Fe-2S], and [4Fe-4S] clusters, since the type of cluster formed is dependent upon the sulfide to iron ion ratio.[49] Finally, NMR analyses showed that such photolytic reactions happen for all the studied thiolates, independently from their physical or chemical properties (Figure A1.19-A.23).
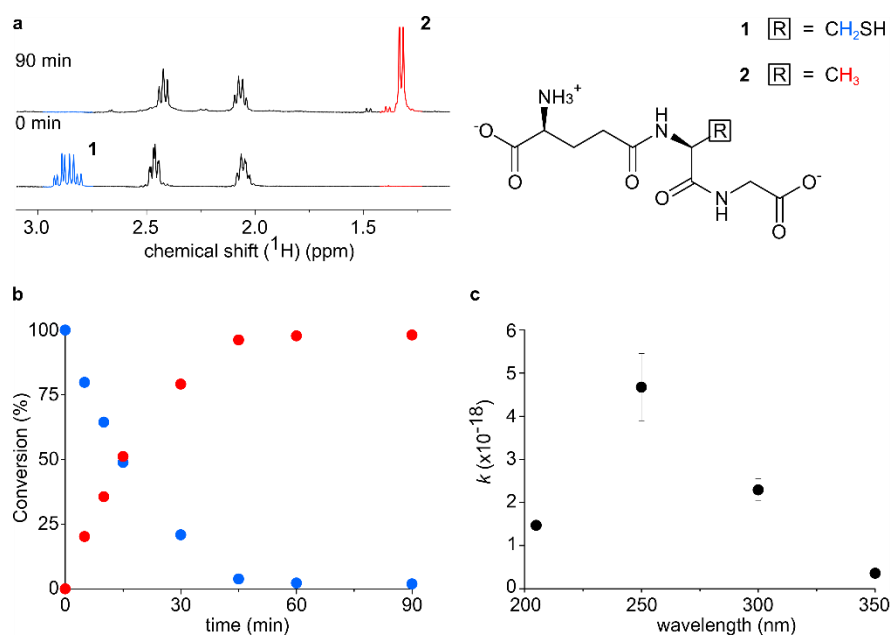
Figure 3.4 | UV light desulfurization of the cysteine of L-glutathione to an alanine. a, Conversion of the thiolate 1 (blue) to the corresponding desulfurized species 2 (red) shown by ^{1}H NMR, consistent with the release of sulfide ions. b, Time course of conversion as obtained by the integration of NMR peaks for the irradiated solution of L-glutathione in the presence of hypophosphite. Blue, L-glutathione; red, alanyl-analogue. c, Dependence of the conversion rate constant on the irradiation wavelength. The data are congruent with the electronic transitions of thiolates. The rate constant (expressed as s-1) was normalized by the photon flux (expressed as s). Data represent n ≥ 3 replicates (mean and SEM).

To gain better insight into whether the observed photochemical reaction could have driven the formation of iron-sulfur clusters on the primordial Earth, the effect of light more similar to that of the young Sun on the desulfurization reaction was assessed. On the primordial Earth the UV environment would be characterized by broadband input, and wavelengths below 204 nm and 168 nm would be shielded out by atmospheric CO_2 and water, respectively.[10] thus oxidized mononuclear Fe^{3+}-L-glutathione complex was subjected to exposure to 205 nm, 250 nm, 300 nm, and 350 nm UV light. To perform these experiments on the wavelength dependence of photolysis a lower intensity lamp (≤0.3 mW) than that exploited above (ca. 35 W) was used, to better mimic early Earth conditions. Exposure to 250 nm light yielded desulfurization at a rate 3-fold, 2-fold, and 13-fold higher than at 205 nm, 300 nm, and 350 nm, respectively (Figure 3.4c). The faster desulfurization at 250 nm is consistent with the electronic transition in the UV region of thiolate bonds, which have an absorbance maximum at ca. 250 nm.

After the identification of the source of the inorganic sulfide, the next question was how iron-sulfur clusters that contain Fe^{3+} centres could have been synthesized from a solution consisting of Fe^{2+}-L-glutathione complex, since in the presence of excess thiolate

ligands mononuclear ferric complexes undergo spontaneous rapid reduction to the corresponding ferrous complexes. Without UV irradiation a [2Fe-2S] cluster cannot be synthesized from Fe^{2+}, S^{2-}, and L-glutathione unless the solution is subsequently exposed to atmospheric oxygen to oxidize the iron ions (Figure 3.5a).
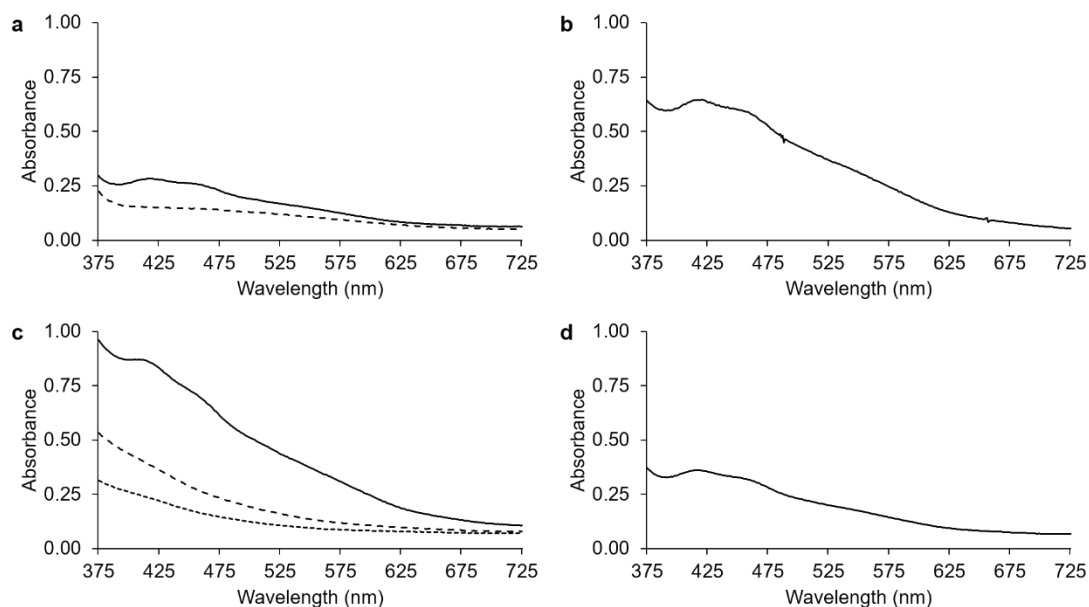
Figure 3.5 | a, Starting from a solution of Fe^{2+}, S^{2-}, and L-glutathione (dashed line), a [2Fe-2S] cluster could not be synthesized unless subsequently exposed to air to oxidize the iron ions (solid line). b, When S^{2-} is premixed with L-glutathione, the addition of Fe^{3+} leads to the formation of a [2Fe-2S] cluster. c, If Fe^{3+} is incubated with L-glutathione prior to the addition of S^{2-}, then the [2Fe-2S] cluster does not form. Sulfide was added after 60 s from premixing of L-glutathione and iron ions (solid line), after 60 min (dashed line) and after 6 h (dotted line). d, The [2Fe-2S] cluster can form by irradiation of a solution of L-glutathione and Fe^{2+} ions.

If S^{2-} ions were mixed with L-glutathione prior to the addition of Fe^{3+} ions, then the [2Fe-2S] cluster could be formed, because the rate of cluster synthesis was faster than the reduction of Fe^{3+} ions (Figure 3.5b). However, if Fe^{3+} ions were added to L-glutathione prior to the addition of S^{2-}, then [2Fe-2S] clusters did not form, because of the reduction of ferric ions prior to the availability of sulfide ions needed for cluster synthesis (Figure 3.5c). That is, UV light irradiation seemed to overcome the reductive effects of iron ions. A solution of reduced mononuclear complex was converted to L-glutathione [2Fe-2S] cluster upon exposure to UV light (Figure 3.5d). Confirming previous reports regarding the photooxidation of ferrous ions in water,[50] the increased formation of ferric thiocyanate and decreased formation of ferrous phenanthroline complexes were monitored in the absence and in the presence of thiolates (Figure 3.6).
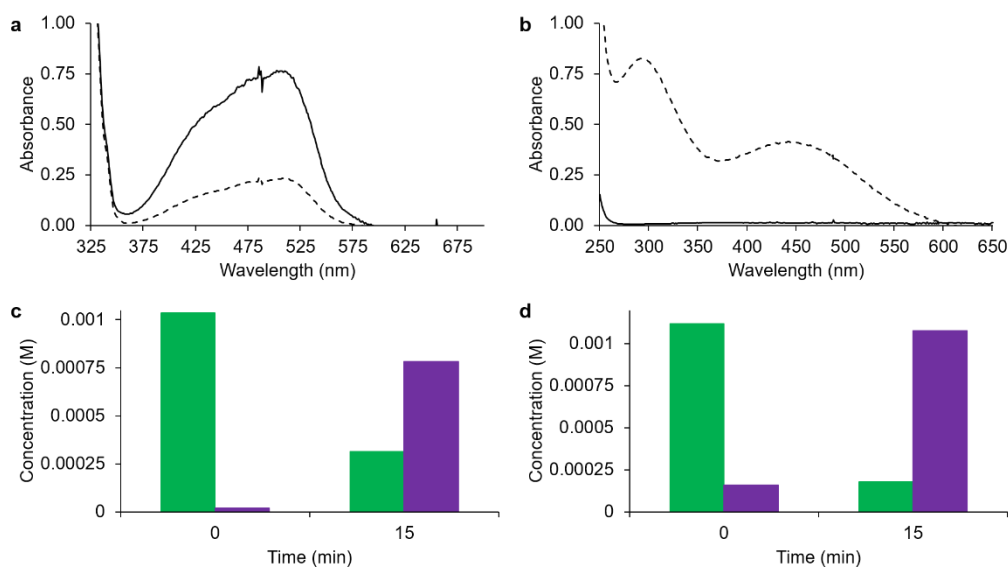
Figure 3.6 | a, UV-vis absorption spectra of a solution of Fe^{2+} ions in water before (solid line) and after (dotted line) exposure to UV light (254 nm) for up to 15 min in the presence of 1,10-phenanthroline, an indicator of Fe^{2+} ions. b, UV-vis spectra of a solution of Fe^{2+} ions in water before (solid line) and after (dotted line) exposure to UV light (254 nm) in the presence of thiocyanate, an indicator of Fe^{3+} ions, for 15 min. c, Change in concentration of Fe^{2+} (green bars) and Fe^{3+} (violet bars) ions upon exposure to UV light in aqueous solution. d, Change in concentration of Fe^{2+} (green bars) and Fe^{3+} (violet bars) ions upon exposure to UV light in the presence of L-glutathione.

The identified photolysis and photooxidation reactions allowed to outline the step-by-step mechanism for the synthesis of prebiotic iron-sulfur clusters and appear to be analogous to the modern biological pathways for the synthesis of several molecules needed to support extant life (Figure 3.7). Indeed, in modern biology a highly conserved cysteine desulfurase present in all three Kingdoms of life[51] (IscS and homologues NifS and SufS) and thought to extend back to at least LUCA,[52] catalyses the conversion of free cysteine to alanine, releasing the sulfide ions necessary for the synthesis of sulfur containing vitamins, thiolated tRNA molecules and iron-sulfur clusters.[1] Since UV light is one of the most abundant sources of energy, similar sulfide releasing mechanisms could have occurred on the surface of the prebiotic Earth or in the first layers of primordial oceans before desulfurase enzymes appeared. It is plausible to expect that such photochemistry pathways took part in the formation of biological molecules in the past[8], since UV light continues to be exploited for the synthesis of modern biological molecules, such as vitamin D_3. Analogously, the photooxidation pathways may also prelude modern day steps in iron-sulfur cluster biosynthesis. Indeed, the *isc* operon that encodes the evolutionarily conserved protein machinery needed for the synthesis of iron-sulfur clusters contains a necessary redox active ferredoxin.
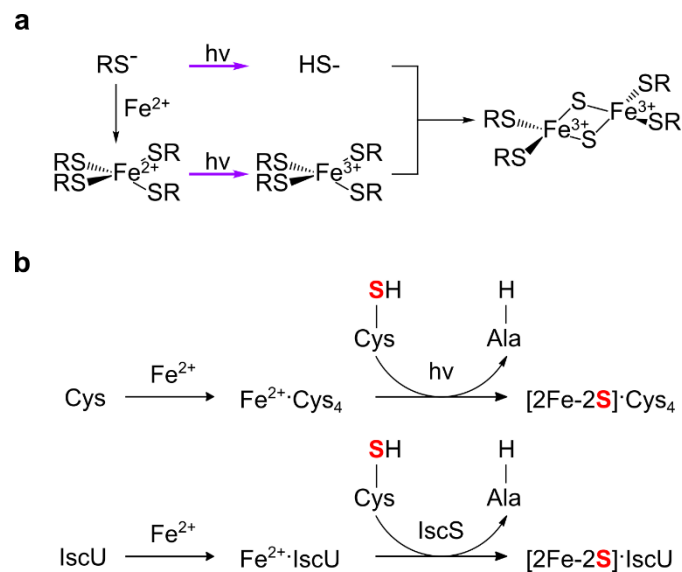
Figure 3.7 | Prebiotic pathways for the synthesis of iron-sulfur clusters. a, Proposed mechanism of photooxidation and photolysis for the synthesis of a [2Fe-2S] cluster from a thiolate solution in the presence of ferrous ions. b, A comparison between model prebiotic (top) and extant, biosynthetic (bottom) pathways for the donation of sulfide during the synthesis of an iron-sulfur cluster.

Having shown that the synthesis of iron-sulfur clusters can be driven by UV light in bulk solution, the next step was focused on the compatibility of the studied synthetic pathway with membrane models, since iron-sulfur proteins are usually membrane-bound proteins. Fatty acid vesicles are often used to build model protocells, because fatty acids can be abiotically synthesized and can form vesicles that acquire nutrients, grow, divide, and host nucleic acid non-enzymatic polymerization.[53,54] However, fatty acid vesicles are highly sensitive to the presence of divalent cations.[55] It was previously shown that the stability of vesicles composed of oleate, myristoleate, or decanoate to Mg^{2+} could be strongly enhanced by the presence of citrate ions as chelating agent.[31] Since thiolates can bind iron ions, the stability of oleic acid and decanoic acid vesicles to Fe^{2+} was tested in the absence and in the presence of L-glutathione as model thiolate. Oleic acid and 4:1:1 decanoic acid:decanol:monocaprin vesicles (Table A1.7) containing fluorescein labelled dextran were incubated for 60 min with 5 mM Fe^{2+} ions either in the presence or in the absence of 200 mM L-glutathione. The stability of the fatty acid vesicles was then assessed by quantifying through size-exclusion chromatography and fluorescence spectroscopy the leakage of the fluorescent molecule out of the vesicles. In the absence of L-glutathione, the vesicles were destroyed and white aggregates started to form upon addition of iron ions, with a complete leakage of the fluorescein-dextran adduct. Conversely, the vesicles were completely stable

in the presence of the mononuclear Fe^{2+}-L-glutathione (Figure A1.24-A1.25). The next question was focused on understanding whether the UV light-driven synthesis of iron-sulfur clusters was compatible with fatty acid vesicles (Figure 3.8).
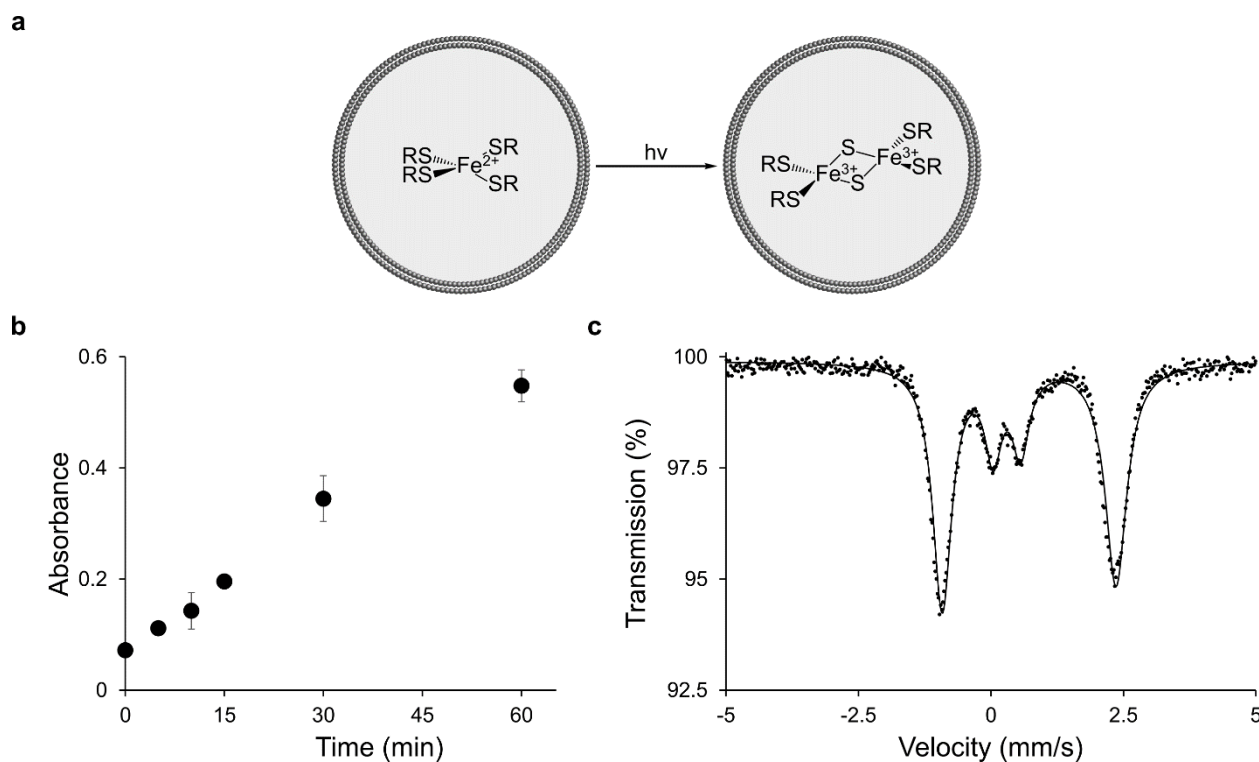
Figure 3.8 | [2Fe-2S] cluster formation in situ within fatty acid vesicles. a, Schematic representation of light-driven, [2Fe-2S] cluster synthesis within fatty acid vesicles. b, A methylene blue assay detected the release of sulfide within oleic acid vesicles induced by UV light, demonstrating the compatibility of the photolysis of entrapped thiolates with fatty acid vesicles. Colorimetric data represent mean and SEM, n ≥ 3 replicates. c, Mössbauer spectrum of the $[2Fe-2S]^{2+}$ cluster, minor component, formed within oleic acid vesicles by UV light.

First, the UV light stability of mononuclear iron-L-glutathione complexes encapsulated within oleic acid and 4:1:1 decanoic acid:decanol:monocaprin vesicles was tested by exposing aliquots of purified vesicles to 254 nm light for 0 or 60 min. Then the samples were fractionated by size-exclusion chromatography and subjected to a colorimetric assay by chelating iron ions with tiron. No iron ions were released from either type of vesicles, with or without UV irradiation (Figure A1.26-1.27). To determine whether sulfide could be generated from cysteine residues inside of the vesicles, purified aliquots of vesicles were subjected to UV light exposure and the release of sulfide ions was assessed by a fluorimetric assay checking the reductive conversion of 7-azido-4-methyl coumarin into 7-amino-4-methyl coumarin by S^{2-} (Figure A1.28). The final step was to test whether UV light could lead to the formation of [2Fe-2S] clusters within fatty acid vesicles. Therefore,

purified vesicles containing mononuclear iron-L-glutathione complex were exposed to 254 nm UV light and the formation of iron-sulfur clusters observed with a methylene blue assay by adding N,N-dimethyl-p-phenylenediamine (DMPD) in hydrochloric acid to each aliquot. Since DMPD reacts with H_2S in the presence of ferric ions to form methylene blue, the appearance of methylene blue spectral bands upon the addition of DMPD indicated the presence of inorganic sulfide and indirectly the presence of an iron-sulfur cluster (Figure 3.8 and A1.29). Mössbauer spectroscopy confirmed the formation of [2Fe-2S] L-glutathione from mononuclear iron-L-glutathione complex (Table A1.8). The results show that cysteine-containing peptides can not only prevent model protocells from iron ions-induced disruption but also allow UV light driven iron-sulfur cluster synthesis within model protocells.

3.3 Discussion

Our data support the ability of UV light coming from the young Sun to influence two critically important processes in iron-sulfur cluster assembly. First, UV light induces the release of the sulfide ions needed for the formation of polynuclear iron-sulfur clusters from thiol containing ligands. Second, UV light allows the photooxidation of ferrous ions, generating the ferric ions needed for cluster synthesis. Both of these UV light mediated steps are analogous to protein mediated biosynthetic processes for the synthesis of iron-sulfur clusters. Moreover, our synthetic model of prebiotic iron-sulfur cluster shows little dependence on the coordinating peptide-sequence and is fully performant within model protocells, suggesting that cellular iron-sulfur cluster synthesis could have emerged at a very early stage in the evolution of life.

A broad range of prebiotically plausible conditions and peptide sequences are compatible with the formation of iron-sulfur clusters that are soluble in aqueous solution. Environments rich in suitable thiolate ligands, Fe^{3+}, and S^{2-}, including a wide variety of cysteine containing peptides, would have allowed the synthesis of [2Fe-2S] and [4Fe-4S] clusters on the early Earth. If only ferrous ions were available, then the needed Fe^{3+} would have been anaerobically converted through photooxidation upon exposure to UV light from the young Sun. Analogously, if volatile hydrogen sulfide was not present in solution, then photolysis of thiol containing molecules would release the necessary sulfide ions for cluster synthesis. Moreover, the photolytic generation of H_2S could influence the synthesis of amino acid, nucleotide, and lipid precursors.[12] The results also revealed a driving force for high nuclearity iron-sulfur cluster emergence in the presence of sulfide: in the absence of sulfide, mononuclear iron coordination predominates, while in the presence of available sulfide ions

the formation of [2Fe-2S] and [4Fe-4S] clusters is preferred. Therefore, modern metabolic pathways that solely depend on enzymes containing mononuclear metal centres, such as glycolysis,[56] may have not needed more complex metal clusters or may have emerged from environments lacking accessible sulfide, while metabolic reactions that exploit polynuclear iron-sulfur clusters may reflect a high environmental presence of sulfide ions.[56] This hypothesis fits well with the possibility of ancestral iron-peptides complexes evolving towards the exploitation of other metal ions when the availability of soluble ferrous ions decreased due to the great oxidation event.[19,57]

Iron-sulfur clusters coordinated by cysteine-containing peptides are stable to the pH and salt conditions needed for the formation of fatty acid vesicles. Moreover, the coordinating peptides prevent vesicles from the degradative effects of thiophilic, soft metal ions. Since the stability of fatty acid vesicles is not compromised by UV light, the photolysis and photooxidation steps needed for the formation of iron-sulfur clusters are not only compatible with model protocells but also extremely similar to modern biosynthetic mechanisms. However, how such iron-sulfur peptides clusters could have impacted the fitness of a protocell is still unclear. Nevertheless, the ease with which iron-sulfur peptides form under prebiotically plausible conditions, the compatibility of iron-sulfur peptide synthesis with model protocells, the prevalence of iron-sulfur dependent metabolism in all known living systems, and the similarity between abiotic iron-sulfur cluster synthesis with extant cellular pathways, suggest that the early Earth surface conditions were disposed for iron-sulfur peptides to have impacted protometabolic processes within self-replicating protocells.

3.4 Materials and Methods

Materials and methods. Unless otherwise specified, all reactions and manipulations were carried out at room temperature using vinyl anaerobic chambers that provide a strict anaerobic atmosphere of 0-5 parts per million (ppm) using a palladium catalyst and a hydrogen gas mix of 5%. The products were maintained under inert atmosphere and transferred to NMR tubes sealed with rubber septa, anaerobic sealed Hellma quartz cuvettes, or sealed glass vials for NMR, UV-Vis spectroscopy, and Mössbauer spectroscopy, or vesicle experiments, respectively. Reagents and solvents were from Sigma-Aldrich, including ^{57}Fe, and VWR International and were used without further purification, unless otherwise stated. Peptides ACG, ECG, FCG, KCG, QCG, WCG, and

GMG were bought from ProteoGenix SAS. The *N*-acetylated compounds Ac-Ala-OH and Ac-Ala-OMe were prepared according to published procedures.

A Mettler Toledo InLab Flex-Micro pH meter was used to monitor pH. H_2O was degassed by the freeze-pump-thaw method. Unless otherwise stated, all irradiations were carried out in a RPR-200 Rayonet reactor using 8 lamps with a principal emission at 254 nm and with the cooling fan turned on resulting in an internal reactor temperature of ca. 35 °C. Kinetic irradiation studies at variable wavelengths were carried out with a Photon Technology International UV tunable lamp with a 75 W Xenon lamp as the power source. Tunable wavelength selection was enabled by adjusting the angle between the diffraction grating and the slit. The bandwidth of radiation used for all trials was set to 10 nm. Kinetic values obtained from these studies were normalized by the incident photon flux, as the flux varies from one wavelength to another. ^{1}H-NMR spectra were acquired using a Bruker Ultrashield 400 Plus operating at 400.1 MHz 1H frequency equipped with a cryogenic inverse probe. The sample temperature was 298 K. Samples consisting of H2O/D2O mixtures were analysed using HOD suppression to collect the ^{1}H-NMR data. Chemical shifts (δ) are shown in ppm. The yields of conversion were determined by relative integrations of the signals in the ^{1}H-NMR spectrum or by spiking with internal standards, unless otherwise stated. UV-Vis absorption spectra of freshly prepared solutions were recorded with an Agilent 8453 UV-Vis diode array spectrophotometer with an integration time of 0.5 s and an interval of 1 nm. Mössbauer spectra were recorded at 80 K in zero magnetic field in the solid state on an ES-Technology MS-105 Mössbauer spectrometer with a 60 MBq 57Co source in a rhodium matrix at ambient temperature. Spectra were referenced against a 25 µm iron foil at 298 K and spectrum parameters obtained by fitting with Lorentzian lines. Samples were prepared by grinding with boron nitride under an inert atmosphere. Mass spectra for L-glutathione samples were acquired on an Agilent 1200 LC-MS system that employs a 6130 Quadrupole spectrometer. The solvent system used for liquid chromatography (LC) was 0.2 % formic acid in H_2O as buffer A, and 0.2 % formic acid in acetonitrile (MeCN) as buffer B. Samples were injected into Phenomenex Jupiter C18 column (150 x 200 mm, 5 µm) and subsequently into the mass spectrometer using a fully automated system. Spectra were acquired in the positive mode and analysed using the MS Chemstation software (Agilent Technologies). The deconvolution program provided by the software was used to obtain the mass spectra. Absorbance and fluorescence analyses of vesicle experiments were performed with a SpectraMax i3x Multi-Mode Detection Platform (Molecular Devices).

Solid-phase peptide synthesis (SPPS). All the peptides were synthesized according to previously published SPPS procedures by using Fmoc-protected L-amino acids.[58] *N,N*-dimethyl formamide (DMF) was used as the solvent and preloaded fluorenylmethyloxycarbonyl-glycyl-Wang resin (Fmoc-Gly-Wang) was used as the starting polymeric support. Trityl-protected Fmoc-cysteine (Fmoc-Cys(Trt)-OH) and tert-butyl (OtBu) side chain-protected Fmoc-α-amino acids were used as building blocks. In general, the peptide chain was elongated by sequential Fmoc deprotection of the residue anchored to the resin and Fmoc-AA-OH coupling. Acetic anhydride in DMF was used to achieve *N*-acetylation. Fmoc deprotection was obtained by washing the mixture with 20% (v/v) solution of piperidine in DMF. For each coupling, an excess (Fmoc-AA-OH: anchored AA, 4:1) of the Fmoc-α-amino acid derivative was added to the resin. Apart from Fmoc-Cys(Trt)-OH, Fmoc-α-amino acid derivatives were activated with a mixture of hydroxyl-benzotriazole (HOBt), *N,N,N',N'*-tetramethyl-O-(benzotriazol-1-yl)uronium tetrafluoborate (TBTU), and *N,N*-diisopropylethyl amine (DIPEA). Fmoc-Cys(Trt)OH was activated with a *N,N'*-diisopropylcarbodiimide (DIC)/HOBt mixture. At the end of the coupling, the polymers were cleaved from the resin and deprotected by treatment with a solution of trifluoroacetic acid (TFA):H_2O:triisopropyl silane (TIS):1,2-ethanedithiol (EDT) (volume ratio 97:1:1:1) for 2 h. The volume was successively reduced under nitrogen atmosphere to avoid cysteinyl thiol oxidation, and the product was precipitated with a cold solution of diethyl ether/petroleum ether (30:70 (v/v)) followed by washing cycles with diethyl ether or extracted 3 times with 20% acetic acid/chloroform and finally dried under inert atmosphere.

Sulfide-release studies. An aqueous solution containing each thiol (5 mL, 0.24 M) was prepared and the pH was adjusted to the pK_a of the thiol under anaerobic conditions. Subsequently, the solution was transferred to an NMR tube, capped with a rubber septum, and irradiated at 254 nm.[59] Hypophosphite was used only in the set of experiments described in Fig. 2a-2b. That is, hypophosphite was not used for the synthesis of iron-sulfur clusters unless specifically indicated.

Photo-oxidation colorimetric assay. An aqueous solution of Fe^{2+} (1 mL, 0.001 M) and an aqueous L-glutathione-Fe^{2+} solution (1 mL, L-glutathione 0.24 M, Fe^{2+} 0.001 M) were irradiated (254 nm) for 5 min. An excess of 1,10-phenantroline (0.005 M), which reacts with Fe^{2+} ions, or potassium thiocyanate (0.005 M) solution, which reacts with Fe^{3+} ions, was added to each solution before transfer to anaerobic sealed Hellma quartz cuvettes.

[2Fe-2S] cluster synthesis. Ferric chloride hexahydrate (25 uL, 0.1 M) was added to an aqueous solution containing each thiol (5 mL, 0.24 M) and sodium sulfide nonahydrate (10

uL, 0.1 M) after pH adjustment under anaerobic conditions. For UV-driven experiments, ferric chloride hexahydrate (25 uL, 0.1 M) was added to an aqueous solution containing each thiol (5 mL, 0.24 M) after pH adjustment under anaerobic conditions. Subsequently, the solution was transferred to a NMR tube capped with a rubber septum or an anaerobic sealed Hellma quartz cuvette and irradiated with UV light. For Mössbauer experiments, degassed 2-propanol was added in a 40-fold excess to solutions containing L-glutathione mononuclear Fe^{2+} complex, a [2Fe-2S] cluster, and a [4Fe-4S] cluster to induce precipitation. The solution was centrifuged for 20 min at 5000 g and the solvent was discarded. The obtained precipitate was dried under N_2 atmosphere.

Vesicle formation. Under anaerobic conditions, 100 mM oleic acid and 100 mM decanoic acid:decanol:glycerol monodecanoate 4:1:1 vesicles were prepared as previously described[60] by direct dispersion of the amphiphiles in either buffer (0.2 M L-glutathione, pH 8.6) or in solution containing the Fe^{2+} mononuclear complex (0.2 M L-glutathione, 5 mM Fe^{2+}, pH 8.6) with 0.5 equivalents of NaOH. Samples were briefly vortexed, tumbled at room temperature for 2 h, then extruded with 11 passages through a 400 nm pore membrane (Avanti Polar Lipids) using a Mini-Extruder (Avanti Polar Lipids). Vesicles containing Fe^{2+} mononuclear complex were then loaded for purification onto a Sepharose 4B (6 mL bed volume) size exclusion column to remove unencapsulated material, using 0.2 M Tris-HCl pH 8.5 as the running buffer supplemented with 0.25 mM of oleic acid or 25 mM 4:1:1 decanoic acid:decanol:glycerol monodecanoate. Elution was monitored at 420 nm and the fractions containing vesicles were collected.

Vesicle stability assay. 100 mM oleic acid and 100 mM 4:1:1 decanoic acid:decanol:glycerol monodecanoate vesicles containing 0.1 mM FITC-dextran (MW 4000) were prepared and purified by size exclusion chromatography with a Sepharose 4B column. 500 µL of the vesicle solution was then incubated for 60 min with 500 µL of 200 mM Tris-HCl, pH 8.5, 500 µL of 200 mM L-glutathione, pH 8.6, and 500 µL of a L-glutathione-Fe^{2+} mononuclear complex (5 mM Fe^{2+}, 200 mM GSH) or an aqueous solution containing 5 mM Fe^{2+}. All the samples were loaded onto a Sepharose 4B (6 mL bed volume) size exclusion column to assess vesicle stability, using 0.2 M Tris-HCl, pH 8.5 as the running buffer supplemented with 0.25 mM oleic acid or 25 mM 4:1:1 decanoic acid:decanol:glycerol monodecanoate. The elution profile was monitored by fluorescence (excitation 490 nm, emission 525 nm).

Vesicle sulfide-release assay. A fluorescence assay was performed to detect the release of sulfide from L-glutathione encapsulated within oleate vesicles. Freshly prepared 100 mM

7-azido-4-methyl coumarin in DMSO was added to each fraction (200 uL) of purified oleate vesicles solution containing 200 mM L-glutathione that was previously exposed to UV light (254 nm). The formation of 7-amino-4-methyl coumarin upon reaction with sulfide ions was detected by fluorescence (excitation wavelength: 345 nm, emission wavelength: 440 nm).

Colorimetric assays for iron-sulfur cluster formation within fatty acids vesicles. A tiron assay was performed to detect fractions containing iron. Freshly prepared 100 mM tiron in 0.2 M Tris-HCl, 0.25 mM oleic acid, pH 8.6 was added to each fraction (2.5 mM) that was previously incubated with ethanol (25% v/v) for 5 min. The formation of the tiron complex was detected by absorbance at 484 nm. Alternatively, a methylene blue assay was employed to detect fractions containing sulfide. *N,N*-dimethyl-p-phenylenediamine (DMPD) (10 mg) was dissolved in 10 mL 35% (v/v) HCl to make a DMPD stock solution. DMPD was added to each fraction (2.5 mM) in the presence of Fe^{3+} (1.25 mM) to increase the sensitivity of the assay and left to react for 1 h. The samples were then centrifuged and the absorbance at 668 nm recorded.

3.5 Bibliography

1. Mihara, H. & Esaki, N. Bacterial cysteine desulfurases: Their function and mechanisms. *Appl. Microbiol. Biotechnol.* **60,** 12–23 (2003).

2. Schwartz, C. J., Djaman, O., Imlay, J. a & Kiley, P. J. The cysteine desulfurase, IscS, has a major role in in vivo Fe-S cluster formation in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 9009–9014 (2000).

3. Kim, J. H., Frederick, R. O., Reinen, N. M., Troupis, A. T. & Markley, J. L. [2Fe-2S]-Ferredoxin Binds Directly to Cysteine Desulfurase and Supplies an Electron for Iron–Sulfur Cluster Assembly but Is Displaced by the Scaffold Protein or Bacterial Frataxin. *J. Am. Chem. Soc.* **135,** 8117–8120 (2013).

4. Hoffmann, F. W., Ess, R. J., Simmons, T. C. & Hanzel, R. S. The desulfurization of mercaptans with trialkyl phosphites. *J. Am. Chem. Soc.* **78,** 6414 (1956).

5. González, A. & Valencia, G. Photochemical desulfurization of L-cysteine derivatives. *Tetrahedron Asymmetry* **9,** 2761–2764 (1998).

6. Walling, C. & Rabinowitz, R. The reaction of thiyl radicals with trialkyl phosphites. *J. Am. Chem. Soc.* **79,** 5326 (1957).

7. Knight, A. R. *The chemistry of the thiol group. John Wiley & Sons* (1974). doi:10.1021/cr60109a003

8. Rapf, R. J. & Vaida, V. Sunlight as an energetic driver in the synthesis of molecules

necessary for life. *Phys. Chem. Chem. Phys.* **18,** 20067–20084 (2016).

9. Shane, B. *Encyclopedia of Biological Chemistry. Encyclopedia of Biological Chemistry* **2,** (Elsevier Inc., 2013).

10. Ranjan, S. & Sasselov, D. D. Influence of the UV Environment on the Synthesis of Prebiotic Molecules. *Astrobiology* **16,** 68–88 (2016).

11. Ritson, D. & Sutherland, J. D. Prebiotic synthesis of simple sugars by photoredox systems chemistry. *Nat. Chem.* **4,** 895–9 (2012).

12. 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.* **7,** 301–307 (2015).

13. Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459,** 239–242 (2009).

14. Sugiura, Y. & Tanaka, H. Iron-sulfide chelates of some sulfur-containing peptides as model complex of non-heme iron proteins. *Biochem. Biophys. Res. Commun.* **46,** 335–340 (1972).

15. Balasubramaniam, a. & Coucouvanis, D. Peptide cluster analogue of two iron ferredoxin. *Inorganica Chim. Acta* **78,** 289 (1983).

16. Qi, W. *et al.* Glutathione Complexed Fe–S Centers. *J. Am. Chem. Soc.* **134,** 10745–10748 (2012).

17. Morse, J., Millero, F., Cornwell, J. & Rickard, D. The chemistry of the hydrogen sulfide and iron sulfide systems in natural waters. *Earth-Science Rev.* **24,** 1–42 (1987).

18. Kabil, O. & Banerjee, R. Redox Biochemistry of Hydrogen Sulfide. *J. Biol. Chem.* **285,** 21903–21907 (2010).

19. Anbar, A. D. Elements and Evolution. *Science (80-.).* **322,** 1481–1483 (2008).

20. Mulholland, S. E., Gibney, B. R., Rabanal, F. & Dutton, P. L. Determination of nonligand amino acids critical to [4Fe-4S](2+/+) assembly in ferredoxin maquettes. *Biochemistry* **38,** 10442–10448 (1999).

21. Ueno, S., Ueyama, N., Nakamura, A. & Tukihara, T. Synthesis of tetrapeptide 2Fe-2S complexes of Cys-X-Y-Cys segments by a ligand-exchange reaction. Peptide models of iron ferredoxin characterized by electrochemistry and spectroscopy. *Inorg. Chem.* **25,** 1000–1005 (1986).

22. Sun, W., Ueyama, N. & Nakamura, A. Stabilization of hydrolytically labile iron(II)–cysteine peptide thiolate complexes in aqueous triton X-100 micelle solution:

Spectroscopic properties mimicking of reduced rubredoxin. *Biopolymers* **46,** 1–10 (1998).

23. Szostak, J. W., Bartel, D. & Luisi, P. L. Synthesizing life. *Nature* **409,** 387–390 (2001).

24. Lawless, J. G. & Yuen, G. U. Quantification of monocarboxylic acids in the Murchison carbonaceous meteorite. *Nature* **282,** 396–398 (1979).

25. McCollom, T. M., Ritter, G. & Simoneit, B. R. T. Lipid Synthesis Under Hydrothermal Conditions by Fischer- Tropsch-Type Reactions. *Orig. Life Evol. Biosph.* **29,** 153–166 (1999).

26. Blain, J. C. & Szostak, J. W. Progress Toward Synthetic Cells. *Annu. Rev. Biochem.* **83,** 615–640 (2014).

27. Budin, I. & Szostak, J. W. Physical effects underlying the transition from primitive to modern cell membranes. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 5249–54 (2011).

28. Müller, U. F. & Tor, Y. Citric Acid and the RNA World. *Angew. Chemie* **53,** 5245–5247 (2014).

29. Mansy, S. S. *et al.* Template-directed synthesis of a genetic polymer in a model protocell. *Nature* **454,** 122–5 (2008).

30. Adamala, K. & Szostak, J. W. Competition between model protocells driven by an encapsulated catalyst. *Nat. Chem.* **5,** 495–501 (2013).

31. Adamala, K. & Szostak, J. W. Nonenzymatic template-directed RNA synthesis inside model protocells. *Science* **342,** 1098–100 (2013).

32. Zhu, T. F., Adamala, K., Zhang, N. & Szostak, J. W. Photochemically driven redox chemistry induces protocell membrane pearling and division. *Proc. Natl. Acad. Sci.* **109,** 9828–9832 (2012).

33. Kolb, V. *Astrobiology: an evolutionary approach.* (CRC/F&T, 2014).

34. Lide, D. R. CRC Handbook of Chemistry and Physics. 3485 (2003). doi:978-1466571143

35. Yang, C. S. & Huennekens, F. M. Iron-Mercaptoethanol-Inorganic sulfide complex. Possible model for chromophore of nonheme iron proteins. *Biochemistry* **9,** 2127–2133 (1970).

36. Sugiura, Y., Kunishima, M. & Tanaka, H. Spectral similarity of some dithiol-iron(III)-sulfide complexes to ferredoxin. *Biochem. Biophys. Res. Commun.* **48,** 1400–1403 (1972).

37. Sugiura, Y., Kunishima, M. & Tanaka, H. Some factors controlling incorporation of

labile sulfur into thiol-iron(III) complexes. *Biochem. Biophys. Res. Commun.* **49,** 1518–1524 (1972).

38. Bonomi, F., Werth, M. T. & Kurtz, D. M. Assembly of [FenSn(SR)4]2- (n = 2, 4) in aqueous media from iron salts, thiols, and sulfur, sulfide, or thiosulfate plus rhodanese. *Inorg. Chem.* **24,** 4331–4335 (1985).

39. Conover, R. C. *et al.* Spectroscopic characterization of the novel iron-sulfur cluster in Pyrococcus furiosus ferredoxin. *J. Biol. Chem.* **265,** 8533–8541 (1990).

40. Crack, J. C. Detection of Sulfide Release from the Oxygen-sensing [4Fe-4S] Cluster of FNR. *J. Biol. Chem.* **281,** 18909–18913 (2006).

41. Fu, W., Morgan, T. V., Mortenson, L. E. & Johnson, M. K. Resonance Raman studies of the [4Fe-4S] to [2Fe-2S] cluster conversion in the iron protein of nitrogenase. *FEBS Lett.* **284,** 165–168 (1991).

42. Duin, E. C. *et al.* [2Fe-2S] to [4Fe-4S] cluster conversion in Escherichia coli biotin synthase. *Biochemistry* **36,** 11811–11820 (1997).

43. Zhang, B. *et al.* Reversible cycling between cysteine persulfide-ligated [2Fe-2S] and cysteine-ligated [4Fe-4S] clusters in the FNR regulatory protein. *Proc. Natl. Acad. Sci.* **109,** 15734–15739 (2012).

44. Fitzpatrick, J. & Kim, E. New Synthetic Routes to Iron–Sulfur Clusters: Deciphering the Repair Chemistry of [2Fe–2S] Clusters from Mononitrosyl Iron Complexes. *Inorg. Chem.* **54,** 10559–10567 (2015).

45. Tran, C. T., Williard, P. G. & Kim, E. Nitric oxide reactivity of [2Fe-2S] clusters leading to H2S generation. *J. Am. Chem. Soc.* **136,** 11874–7 (2014).

46. Harrop, T. C., Song, D. & Lippard, S. J. Reactivity pathways for nitric oxide and nitrosonium with iron complexes in biologically relevant sulfur coordination spheres. *J. Inorg. Biochem.* **101,** 1730–1738 (2007).

47. Pandelia, M. E., Lanz, N. D., Booker, S. J. & Krebs, C. Mössbauer spectroscopy of Fe/S proteins. *Biochim. Biophys. Acta* **1853,** 1395–1405 (2015).

48. Pasek, M. A., Kee, T. P., Bryant, D. E., Pavlov, A. A. & Lunine, J. I. Production of potentially prebiotic condensed phosphates by phosphorus redox chemistry. *Angew. Chem. Int. Ed. Engl.* **47,** 7918–20 (2008).

49. Venkateswara Rao, P. & Holm, R. H. Synthetic Analogues of the Active Sites of Iron-Sulfur Proteins. *Chem. Rev.* **104,** 527–559 (2004).

50. Braterman, P. S., Cairns-Smith, A. G., Sloper, R. W., Truscott, T. G. & Craw, M. Photo-oxidation of iron(II) in water between pH 7.5 and 4.0. *J. Chem. Soc. Dalt.*

Trans. **7,** 1441–1445 (1984).

51. Lill, R. Function and biogenesis of iron-sulphur proteins. *Nature* **460,** 831–838 (2009).

52. Gil, R., Silva, F. J., Pereto, J. & Moya, A. Determination of the Core of a Minimal Bacterial Gene Set. *Microbiol. Mol. Biol. Rev.* **68,** 518–537 (2004).

53. Mansy, S. S. & Szostak, J. W. Reconstructing the emergence of cellular life through the synthesis of model protocells. *Cold Spring Harb. Symp. Quant. Biol.* **74,** 47–54 (2009).

54. Stano, P. & Luisi, P. L. Achievements and open questions in the self-reproduction of vesicles and synthetic minimal cells. *Chem. Commun.* **46,** 3639 (2010).

55. Monnard, P., Apel, C. L., Kanavarioti, A. & Deamer, D. W. Influence of ionic inorganic solutes on self-assembly and polymerization processes related to early forms of life: implications for a prebiotic aqueous medium. *Astrobiology* **2,** 139–152 (2002).

56. Belmonte, L. & Mansy, S. S. Metal Catalysts and the Origin of Life. *Elements* **12,** 413–418 (2016).

57. Hsiao, C. *et al.* RNA with iron(II) as a cofactor catalyses electron transfer. *Nat. Chem.* **5,** 525–528 (2013).

58. Verlander, M. S., Fuller, W. D. & Goodman, M. Rapid, quantitative peptide synthesis using mixed anhydrides. US004351762 (1982).

59. Yamashita, T. Process for chemical desulfurization of coal. US005350431A (1994).

60. Zhu, T. F. & Szostak, J. W. Uniform-sized, multi-drug carrying, and photosensitive liposomes for advanced drug delivery. US 20110105995A1 (2011).

Chapter 4 - Towards iron-sulfur enzymes and protometabolic pathways

4.1 Introduction

Life as we know it has two key features: information transfer and energy transduction.[1] In the last decades, research on the origin of life has primarily focused on RNA, since membrane-encapsulated[2] or mineral-adsorbed[3] RNA can serve both as an information reservoir and as a catalyst for functions that can be subjected to Darwinian evolution.[4] Being inspired by Peter Mitchell's studies on energy transduction,[5,6] in 1985 Arthur Koch hypothesized that a proton gradient was required to drive endergonic reactions.[7] If electrons, but not protons or charged species can pass through the phospholipid membrane, a reduction and an oxidation on different sides of the cellular membrane can easily allow for energy storage in the form of charge displacement and production of a proton motive force across the membrane.[8]

The fundamental enzymatic features of present-day living systems is probably the result of the progressive development of ancestral catalytic properties present in simple ions or molecules found in the prebiotic environment.[9] Therefore, it is highly plausible that metal ions available on the early Earth and widely found in extant proteins were coordinated by prebiotic organic and inorganic ligands and acted as primordial catalysts with low activity and broad specificity. In fact, free metal ions such as iron, molybdenum, and zinc, or relatively simple artificial metal-bound complexes have been shown to have catalytic activity more or less similar to that of present day enzymes.[9–11] That is, modern enzymes containing iron, molybdenum, and zinc could have evolved from protoenzymes catalysing the transfer of electrons, the activation or incorporation of small molecules, and hydrolytic or transferring reactions, respectively, without constrained specificity and with weak activity.[12] One group of modern enzymes which is widely dispersed in present day biological systems is that centred around the element iron. In 1959 Melvin Calvin published a research study focused on the evolution of modern enzymes. He recognized that the most important features of extant enzymes derive from the evolutionary development of the intrinsic features of simple molecules and metal ions present in the primordial environment.[9] By studying biological systems centred around the element iron, he found that the catalytic ability of decomposing hydrogen peroxide into oxygen and water progressively increases from 10^{-5} for the aqueous ferric ion to 10^{-2} for the heme, to 10^5 for the complete catalase enzyme.

Extant iron-based proteins exhibit many functions, mainly involving oxidation or reduction reactions, which explains why iron enzymes, and in particular iron-sulfur enzymes, are mostly involved in biological electron transfer reactions.[9] Indeed, the ubiquity of iron-sulfur proteins in all modern-day electron-transfer chains could suggest an early and central role for non-protein iron-sulfur complexes in energy transduction.[8] As previously reported in this thesis, from an evolutionary point of view, iron-sulfur proteins, such as rubredoxins and ferredoxins, are thought to be among the most ancient proteins found in biology and are critical to a wide variety of the chemistry that supports life, including photosynthesis, nitrogen fixation, protein and DNA synthesis, and the electron transfer chain.[13] In 1976 Ycas proposed that, since no enzymatic activity is required to coordinate iron with the protein sequence of ferredoxin, ferredoxin is likely to be more ancient than other cofactor-coordinating proteins that require enzymatic assembly.[14] Moreover, homology studies on ferredoxin and rubredoxin sequences suggest that rubredoxin could be older.[15] Additionally, bioinformatics and structural analyses indicate that rubredoxins are evolutionary related to flavodoxins.[15] Indeed, there is a remarkable structural similarity between the $Fe(Cys)_4$ unit found in rubredoxin-like proteins and the $Zn(Cys)_4$ unit found in DNA-binding zinc fingers.[16,17] Taken together, a path from a simple iron-binding protein (rubredoxin binds a single iron ion) to nucleotide-binding proteins (flavodoxins bind the nucleotide containing molecules flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD)) begins to take shape. Thus, it is plausible that many biochemical processes developed mainly by gene duplication and subsequent specialization from one primordial catalyst. That is, reproducing these active iron-based systems could help to reveal their ancestral role as catalysts on the early Earth.

The work presented in the previous Chapters shows one potential path to the emergence of modern day iron-sulfur proteins. The next goal is then to exploit the model prebiotic iron-sulfur peptides to mimic extant metabolic reactions catalysed by iron-sulfur proteins in modern biology in order to determine the chemical roots of protometabolism. The preliminary data shown in Chapter 2 reports that [2Fe-2S] clusters are redox-active toward simple reductants and oxidants, such as sodium dithionite and hydrogen peroxide. Since iron-sulfur proteins are widely involved in shuttling electrons in the advanced machinery of mitochondrial respiration, the next step aims at reproducing a prebiotically plausible, more simplified electron transport chain. Prebiotic redox-active iron-sulfur centres could then mediate electron transfer involving biologically relevant electron acceptors, such as

ubiquinone (CoQ10), and electron donors, such as nicotinamide adenine dinucleotide (NADH) and nicotinamide adenine dinucleotide phosphate (NADPH) (Figure 4.1).
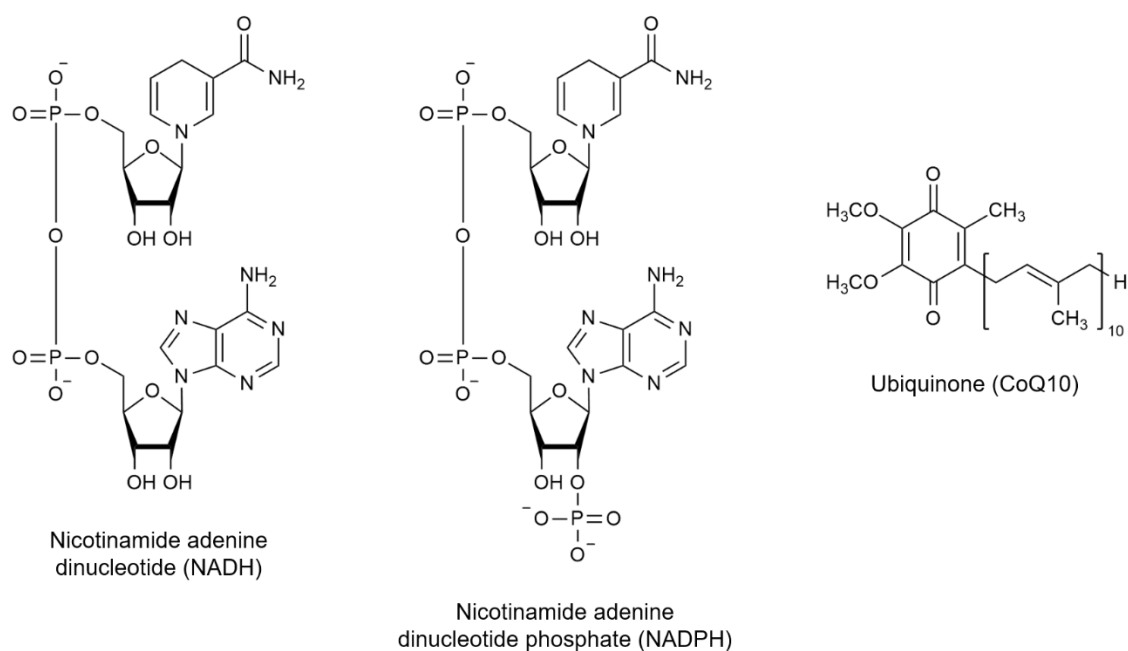
Figure 4.1 | Chemical structures of studied electron donors and acceptors.

Such reactions would prebiotically mimic what is found in extant organisms, since typically iron and iron-sulfur proteins form intra- and inter-protein complex systems containing FMN and FAD to catalyse electron transfer reactions involving NADH and NADPH and to form a proton gradient to power the synthesis of ATP.

4.2 Results

Aqueous solutions of iron ions were previously shown to catalyse several redox reactions, such as the decomposition of hydrogen peroxide and the oxidation of free thiols.[9,18,19] In fact, as shown in Chapter 3, in slightly alkaline pH, thiol-containing molecules could easily coordinate ferric ions forming a tetrahedral mononuclear complex which undergoes rapid reduction due to the presence in solution of extra thiolates prone to be oxidized. The expected results were then confirmed by high performance liquid chromatography (HPLC) to quantify the amount of converted oxidized thiolates and UV-vis. spectrophotometry in the presence of resazurin as a redox indicator (E° = -0.042 V vs Normal Hydrogen Electrode (NHE)) (Figure A2.1-A2.2). Indeed, once the free thiolates were oxidized by the mononuclear Fe^{3+} complex, resazurin was reduced to resorufin. Therefore, the identification of iron complexes able to take part into more biologically-relevant metabolic

pathways such as those involving the NADH/NAD+ species could have a fundamental impact on prebiotic chemistry studies.

The midpoint potential of Fe(III)/Fe(II) is strongly positive (E° = +0.771 V vs NHE, Figure A2.3) and thus free ferric ions would be able to catalyse the oxidation of NADH (E° = -0.320 V vs NHE).

$$NADH \rightarrow 2e^- + H^+ + NAD_+$$
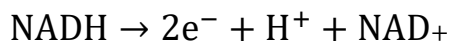
In fact, the results obtained by HPLC showed that Fe^{3+} ions could lead to the oxidation of NADH to NAD+ (Figure 4.2). However, the acidic Fe(III) solution led also to the degradation of the starting material, while an alkaline Fe(III) solution caused the precipitation of insoluble oxo-hydroxo iron complexes.
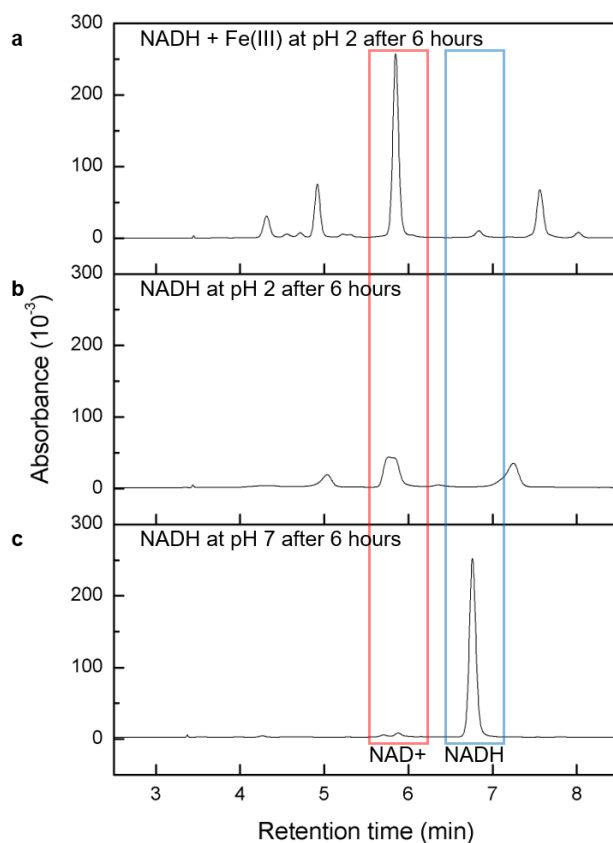
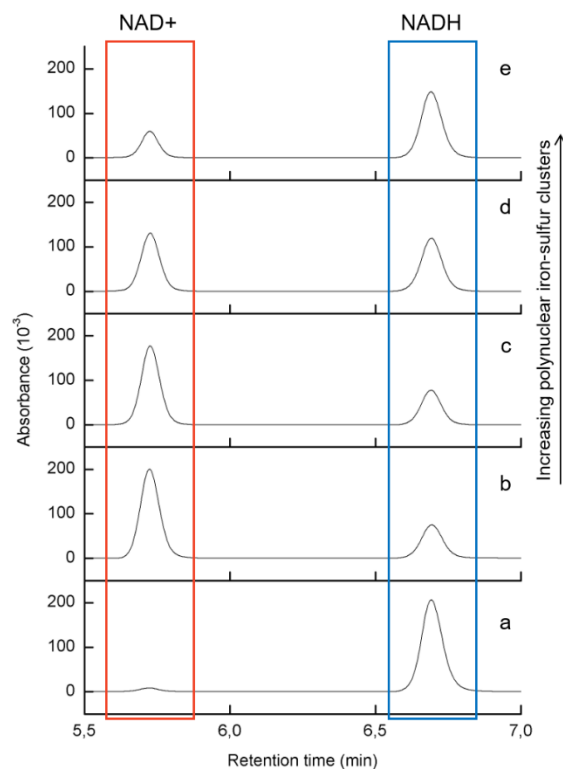**Figure 4.2 |** Effect of pH and Fe(III) ions on NADH studied by HPLC analysis (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm). a, Solution of NADH (0.5 mM) and Fe^{3+} ions (0.5 mM) in acidic pH after 6 hours from mixing. b, Solution of NADH (0.5 mM) in acidic pH after 6 hours. c, Solution of NADH (0.5 mM) in neutral pH after 6 hours. The peak centred at 6.69 min is due to NADH, the peak centred at 5.72 min is due to NAD+.

In order to be biologically relevant, the reaction should ideally occur at neutral pH without inducing the degradation of the substrate. Therefore, several *O*-coordinating or *S*-coordinating ligands such as L-serine, citric acid, 2-mercaptoethanol, L-cysteine, *N*-acetyl L-cysteine methyl ester and L-glutathione were tested for their ability to form complexes with iron ions and catalyse the oxidation of NADH (Scheme 4.1 and Table A2.1).
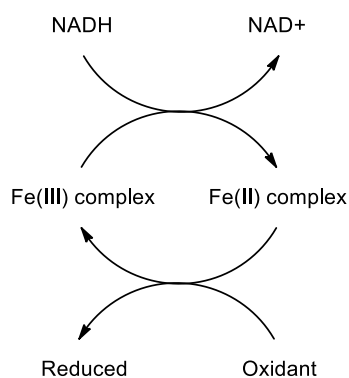
Scheme 4.1 | Schematic representation of the redox processes studied. Upon NADH oxidation the Fe(III) complex is converted to the reduced form. In order to close the redox cycle, an oxidant must then be available to oxidize back the Fe(II) complex.

For example, the carboxylate groups of L-serine and citric acid are able to bind ferric ions at pH 3.5 and 6.5, respectively. 2-mercaptoethanol, *N*-acetyl L-cysteine methyl ester and L-glutathione were previously shown to accommodate ferric ions in a tetrahedral geometry through *S*-coordination in slightly alkaline aqueous solutions. Moreover, L-cysteine can coordinate Fe(III) ions with the α-carboxyl, the α-amino and the side-chain thiol groups at pH 8.4. Since the coordination of iron ions would affect the tendency of the Fe(III) centre to be reduced, cyclic voltammetry studies were performed on each ferric complex to determine the reduction potential (Table A2.2). The performed electrochemical studies showed that all the Fe(III) species would be able to induce the oxidation of NADH, with the lowest value for the 2-mercaptoethanol Fe^{3+} complex (-0.189 V vs NHE) and the highest value for the L-serine Fe^{3+} complex (+0.085 V vs NHE) (Figure A2.4-A2.7). All but *N*-acetyl cysteine methyl ester mononuclear complex exhibited an electrochemically irreversible profile, which could be due to a different coordination of iron ions between the two oxidation states or to a different interface interaction between the electrode and the two redox species. However, the HPLC analyses following the oxidation of NADH to NAD+ demonstrated that *O*-coordination inhibits the reducing power of ferric ions on NADH (Figure 4.3). Moreover, no NADH oxidation was observed for the mixed coordination of L-cysteine on Fe(III) ions (Figure A2.8).
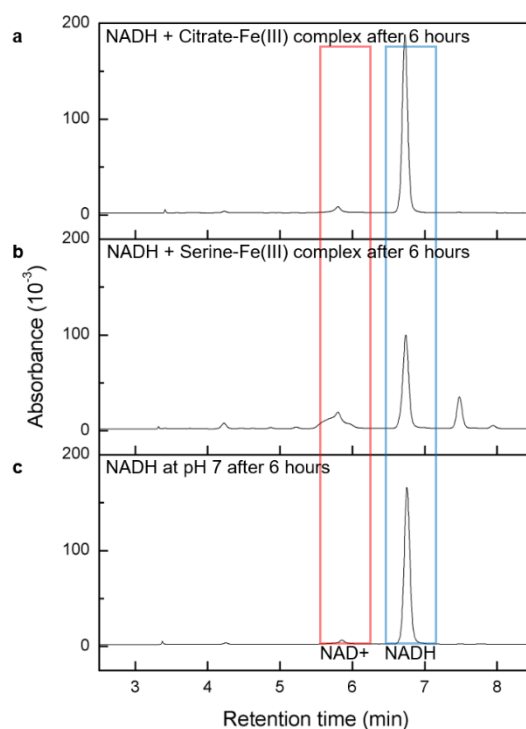
Figure 4.3 | Effect of Fe(III) ions *O*-coordinated by L-serine and citrate on NADH studied by HPLC analysis (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm). a, Solution of NADH (0.5 mM) and citrate-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM citric acid, pH 6.5) after 6 hours from mixing. b, Solution of NADH (0.5 mM) and L-serine-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM L-serine, pH 3.5) after 6 hours from mixing. The incubation of NADH and the L-serine ferric complex in acidic conditions induced the degradation of the substrate. c, Solution of NADH (0.5 mM) in neutral pH after 6 hours. The peak centred at 6.69 min is due to NADH, the peak centred at 5.72 min is due to NAD+.

Conversely, thiol-Fe(III) complexes were shown to efficiently influence the NADH oxidation (Figure 4.4). To understand the importance of the coordinating group on the activity of the ferric ions, L-glutathione was studied as a ligand at different pH values to explore the selective binding of iron ions through the *O*- (pH = 3) or *S*-coordination (pH = 8.6) (Figure A2.9-A2.10). Cyclic voltammetry data confirmed that the coordination and thus the midpoint potential of ferric ions are strongly dependent on the pH value. In acidic conditions, only carboxylates are available to coordinate iron ions and the cathodic peak potential was 0.371 V vs NHE. HPLC studies showed that *O*-coordinated L-glutathione iron complex do not exhibit any effect on NADH oxidation (Figure A2.11). Slightly alkaline conditions induce the *S*-coordination of the cysteine residues on iron ions and thus a shift of the reduction peak to -0.080 V vs NHE. As in the case of other thiol ferric complexes, the L-glutathione mononuclear complex is able to efficiently oxidize NADH (Figure 4.4).

Electrochemical data on 2-mercaptoethanol and *N*-acetyl L-cysteine methyl ester ferric complexes resulted in reduction peaks at about -0.189 V and -0.120 V, respectively. Unsurprisingly, the *N*-acetyl cysteine methyl ester mononuclear complex was the only complex exhibiting an electrochemically reversible profile, probably because thiolates are the only available coordinating group for both ferric and ferrous ions. The overall results on iron complexes with *O*- or *S*-ligands suggest that the coordinating groups may play a central role in the interaction with the substrate. Moreover, the tetrahedral *S*-coordination of iron-thiolate complex mimics the geometry occurring in extant iron-sulfur proteins involved in catalytic NADH oxidation processes occurring in several modern biological processes.

Focusing on iron-thiolate mononuclear complexes, several experiments were performed aimed at understanding whether the structure and the length of the ligand moiety could influence the oxidation of NADH (Figure 4.4 and A2.12). Together with the previously studied coordinating thiolates, *N*-acetyl L-cysteine and 3-mercaptopropionic acid iron complexes were also tested in slightly basic conditions, according to each thiol pK_a value. Moreover, the activity of ferric ions coordinated by L-glutathione polymers (see Chapter 2) was investigated, and compared with rubredoxin, a small redox-active protein containing one iron atom in the active site.
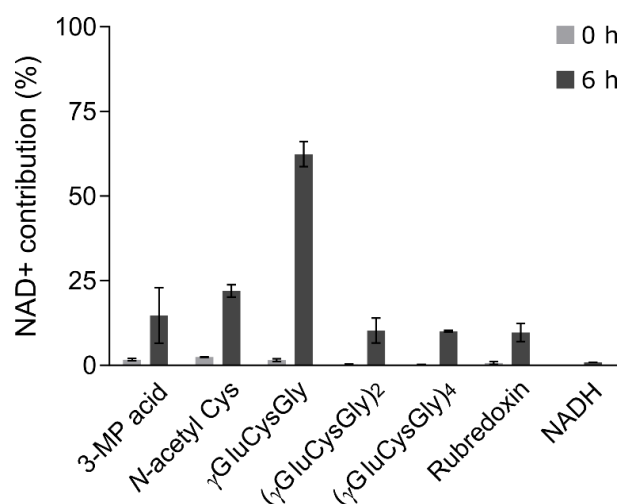
Figure 4.4 | NADH (0.5 mM) conversion efficiency for different mononuclear complexes (0.5 mM) based on HPLC data (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm). The investigated ligands (40 mM) were 3-mercaptopropionic acid (3-MP acid), *N*-acetyl cysteine (*N*-acetyl Cys), L-glutathione (γGluCysGly), L-glutathione polymers ((γGluCysGly)$_2$ and (γGluCysGly)$_4$), rubredoxin. The spontaneous oxidation of NADH is reported as a control. Baseline correction and integration analysis performed with Origin Pro 8. Presented data represent mean and SEM, n ≥ 2 replicates. Dark grey bars represent NAD+ concentration at time 0 h calculated from absorbance values, light grey bars represent NAD+ concentration at time 6 h calculated from absorbance values, both corrected by the extinction coefficients of NADH and NAD+, respectively.

The data showed that a higher conversion of NADH to NAD+ can be detected for 3-mercaptopropionic acid, *N*-acetyl L-cysteine and L-glutathione mononuclear complexes with respect to 2-mercaptoethanol and *N*-acetyl cysteine methyl ester analogues (Figure A2.12). The common chemical feature for the most active iron complexes is the presence of carboxylate groups in the ligand moiety, which increase the negative charge of the mononuclear complex. In fact, the thiolate-ferric complexes range from an overall charge of -1 for 2-mercaptoethanol and *N*-acetyl cysteine methyl ester coordination to -5 for 3-mercaptopropionic acid, *N*-acetyl L-cysteine and L-glutathione coordination. The results seem to suggest that ionic interactions take part to the oxidation process, helping iron-thiolate complexes and NADH molecules to approach and react. From the comparison of equally-charged mononuclear complexes, the length of the ligand chain seems to have an important role as well. According to previous data on the stability of iron-sulfur systems (see Chapter 2 and 3), the coordination to L-glutathione highly enhances the activity of ferric ions towards NADH conversion in comparison with shorter, equally charged ferric species. As previously hypothesized, L-glutathione may be able to protect the Fe(III) complex from hydrolysis more proficiently than 3-mercaptopropionic acid and *N*-acetyl L-cysteine, thus allowing the redox activity to occur to a higher extent. However, the experiments performed on the mononuclear complexes involving L-glutathione polymers and rubredoxin showed a ~10% conversion after 6 h. The observed trend could suggest that too short ligands may not be able to prevent the degradation of the mononuclear complex, while too long ligands may shield the metal active site impeding the interaction with the substrate.

Since the preliminary studies of the thiolate ferric complexes gave promising results in terms of redox activity towards the NADH oxidation reaction, the [2Fe-2S] and [4Fe-4S] clusters reported in Chapter 3 were also tested (Figure 4.5). As previously shown, the amount of inorganic sulfide ions can drive the synthesis of different iron-sulfur species. No inorganic sulfide is required for the tetrahedral coordination of ferric ions by thiolates, while low and high sulfide quantities favour the synthesis of [2Fe-2S] clusters and [4Fe-4S] clusters, respectively. To test the efficiency of different iron-sulfur systems on NADH oxidation, iron-thiolate complex solutions with increasing amount of sulfide ions were studied in the presence of NADH, by going from a 100% solution of mononuclear complex towards binuclear and tetranuclear Fe-S clusters.

Figure 4.5 | Effect of different iron-sulfur systems on NADH studied by HPLC analysis (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm). a, Solution of NADH (0.1 mM). b, Solution of Fe(III) (0.5 mM) with L-glutathione (40 mM, pH 8.6) and NADH (0.1 mM) after 6 hours. c, Solution of sulfide (0.185 mM) and Fe(III) (0.5 mM) with L-glutathione (40 mM, pH 8.6) and NADH (0.1 mM) after 6 hours. d, Solution of sulfide (0.5 mM) and Fe(III) (0.5 mM) with L-glutathione (40 mM, pH 8.6) and NADH (0.1 mM) after 6 hours. e, Solution of sulfide (2.5 mM) and Fe(III) (0.5 mM) with L-glutathione (40 mM, pH 8.6) and NADH (0.1 mM) after 6 hours. The peak centred at 6.69 min is due to NADH, the peak centred at 5.72 min is due to NAD+.

The experiments showed that prebiotic iron-sulfur systems could induce the conversion of NADH into NAD+ in a manner that is dependent on the concentration of sulfide. The mononuclear Fe(III) complex resulted the only prebiotic iron-sulfur system taking part to the oxidation of NADH. By increasing the concentration of sulfide, and thus decreasing the concentration of the mononuclear complex, the conversion efficiency drastically reduced. In extant biology, proteins coordinating iron ions through cysteine residues such as rubredoxins are often involved in the oxidation of NADH and NADPH.[20] Moreover, the midpoint potential value of the mononuclear iron-protein is usually between +0.1 and -0.2 V, while [2Fe-2S] proteins span between -0.2 and -0.4 V and for [4Fe-4S] proteins values are even more negative.[21] Since the midpoint potential of the NAD+/NADH

species is -0.32 V, only rubredoxins would be electrochemically allowed to catalyse the oxidation of NADH.

Cyclic voltammetry and dye photoreduction studies[22] confirmed the obtained results on each iron-sulfur species previously tested (Figure A2.13-A2.19). All the studied iron-sulfur systems are subjected to an electrochemical irreversible reduction. The cathodic peak potentials (E_{cp}) resulted at -0.092 V vs NHE for the reduction of the mononuclear complex, at -0.490 V vs NHE for the [2Fe-2S] cluster and at -0.571 V vs NHE for the [4Fe-4S] cluster, with L-glutathione as the model ligand (Figure A2.13). As previously mentioned, the electrochemical irreversibility of the iron-sulfur species could be due to a chemically reversible or irreversible changes in the coordination sphere of the iron ions in the ferric and the ferrous states. Moreover, the ferric mononuclear complex is strongly adsorbed onto the electrode, as seen from the sharp reduction peak which an increase in intensity observed during sequential measurements.

Since it was not possible to extrapolate the midpoint potential of the studied species, due to their electrochemical irreversibility, redox titrations of each iron-sulfur system were performed in the presence of a redox dye with sodium dithionite as reducing agent. The recorded UV-vis. spectra were decomposed and analysed to find the midpoint potential. The chemical properties required to select the appropriate redox dyes are the midpoint potential, which must be close to the expected midpoint potential of the analysed species, and the absorbance spectrum, which must not overlap the absorbance spectrum of the analysed species. Two series of experiments were then performed. In the first titration studies, benzyl viologen (E° = -0.359 V vs NHE) was added to different solutions of L-glutathione coordinated mononuclear complex, [2Fe-2S] cluster and [4Fe-4S] cluster, and then sodium dithionite was titrated (Figure A2.14-A2.16). Since the midpoint potential for the mononuclear complex was supposed to be much higher than the midpoint potential of benzyl viologen, the reducing agent would have first reacted with the thiolate ferric species, and then with the dye. Conversely, the addition of sodium dithionite to a solution containing benzyl viologen and the [4Fe-4S] cluster would have immediately reduced the dye, without affecting the iron-sulfur species which was supposed to have a lower midpoint potential than the dye. As expected, the data showed that benzyl viologen represents an appropriate redox dye only for the [2Fe-2S] cluster, having a too high or too low midpoint potential for the [4Fe-4S] species and the mononuclear iron complex, respectively. The second set of experiments was performed testing a redox dye with an appropriate midpoint potential for each species (Figure A2.17-A2.19). Based on literature data and previous electrochemical studies, 2,3,5-

triphenyl tetrazolium (E° = -0.07 V vs NHE) was selected for the mononuclear complex and methyl viologen (E° = -0.446 V vs NHE) for the [4Fe-4S] cluster. Upon decomposition of the recorded UV-vis. spectra and analysis of the obtained data, the midpoint potential for the mononuclear complex, the [2Fe-2S] and the [4Fe-4S] clusters resulted -0.025 V vs NHE, -0.311 V vs NHE, -0.381 V vs NHE (Figure 4.6). That is, the ferric mononuclear complex would be able to oxidize NADH, while the more negative midpoint potential of the [2Fe-2S] and the [4Fe-4S] clusters would not allow the reaction to occur.
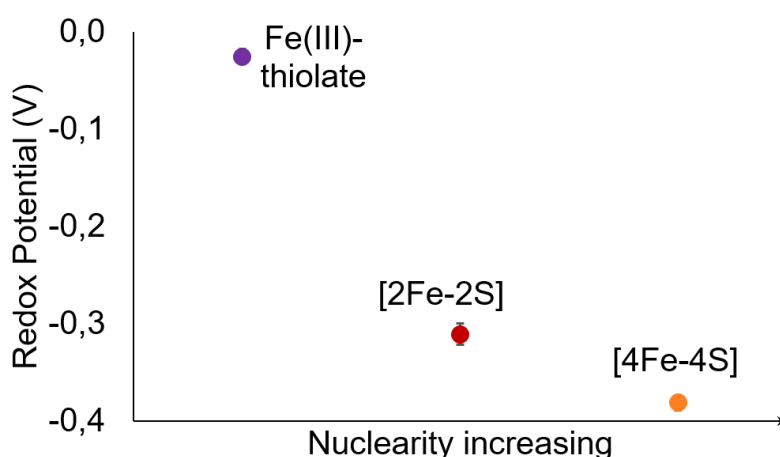
Figure 4.6 | Midpoint potential values of studied iron-sulfur species with different nuclearities as obtained with the dye photoreduction procedure. Presented data represent mean and SEM, n ≥ 2 replicates.

Since the L-glutathione-ferric complex showed the highest efficiency in terms of influencing the NADH oxidation process, such mononuclear complex was further characterized. Dissociation constant (K_D) measurements for both Fe(II) and Fe(III) complexes were performed (Figure A2.20), as well as UV-vis. studies on the stability of the system over time in the absence or in the presence of NADH (Figure A2.21). The obtained K_D values for Fe^{3+} and Fe^{2+} mononuclear complex were 0.33 mM and 0.74 mM, respectively. The high K_D values could thus explain why a wide excess of ligand is required in order to form the mononuclear iron complex. Moreover, the ferrous complex obtained by rapid reduction of Fe(III) as reported in Chapter 3 gave stable complexes for up to 18 h.

In order to get information on the NADH oxidation kinetics and compare the mononuclear complex with enzymes involved in similar processes, the reaction was studied with a fixed concentration of the mononuclear complex (0.5 mM) by varying the concentration of the substrate (NADH concentrations studied: 0.05 mM, 0.1 mM, 0.25 mM, 0.5 mM, 1 mM, 2.5 mM). The obtained data were graphically plotted as reaction rate vs substrate concentration profiles (Figure 4.7).
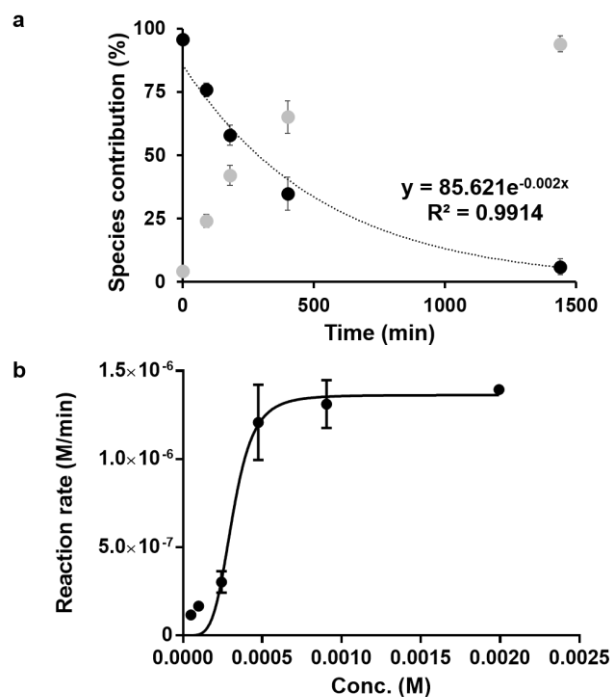
Figure 4.7 | Reaction kinetics for NADH oxidation conversion in the presence of L-glutathione-ferric mononuclear complex. a, First order kinetics for the equimolar reaction between the mononuclear complex (0.5 mM) and NADH (0.5 mM). NADH and NAD+ concentrations are represented as black and grey dots, respectively. b, Reaction rate vs substrate concentration graph with fixed concentration of mononuclear complex (0.5 mM) and variable NADH concentration. Presented data represent mean and SEM, n ≥ 2 replicates.

The NADH oxidation kinetics observed in equimolar concentrations of substrate and L-glutathione-iron complex resulted in a first-order reaction, similar to that observed for the kinetics enzymes. Interestingly, the graph resulting from the data with increasing NADH concentrations did not match a typical Michaelis-Menten kinetic profile, but resulted in a sigmoidal curve, which often indicates allosteric cooperative binding either of two enzymatic units or of the substrate to the active site. Additionally, similar results were obtained with NADPH as the substrate, even though the conversion efficiency to NADP+ was lower than for the NADH/NAD+ species (Figure A2.22). The obtained kinetic data showed similarities between the mononuclear ferric complex and modern enzymes, strengthening the hypothesis that simple metallocomplexes coordinated by peptides could have played a fundamental role as protoenzymes on primordial Earth. Moreover, the sigmoidal profile obtained scanning a wide range of substrate concentrations suggests that molecular interactions are essential for the redox reaction to occur.

In order to further probe and understand whether ionic interactions are essential for the NADH conversion process as suggested before, two sets of experiments were

performed, either exploring the ionic interactions occurring between NADH molecules and the active site or among thiolate-ferric species themselves. Indeed, the oxidation of NADH is a redox process involving the transfer of two electrons, so that two ferric mononuclear complexes would be required to oxidize one molecule of NADH. Since NADH is oxidized in a two-electron concerted process while each iron-sulfur species can accept only one electron per cluster, iron-sulfur enzymes usually exploit complementary biological cofactors such as FMN or FAD to transfer the electrons from NADH to the iron-sulfur core.[23] Indeed, FMN and FAD are able to accept two electrons and then donate one electron at a time to an iron-sulfur acceptor (Scheme 4.2), passing through a stable radical intermediate.
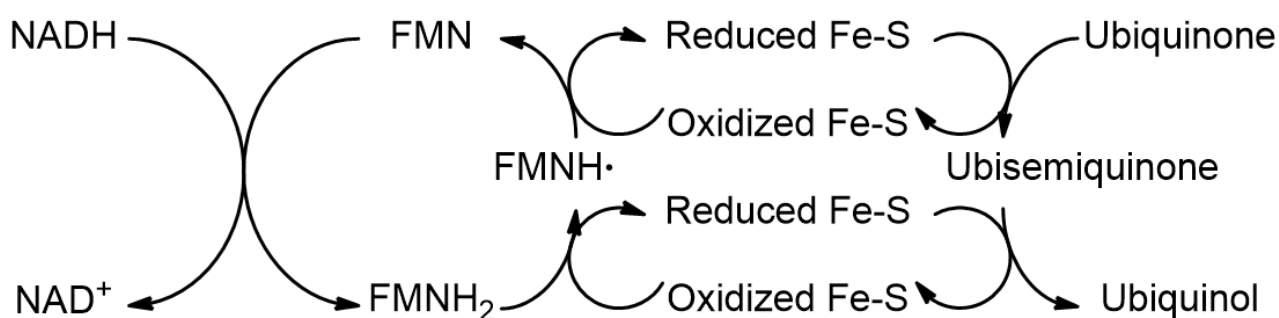
Scheme 4.2 | Schematic representation of the redox processes occurring in the Complex I of the electron transport chain.

To understand the importance of the ionic interactions between NADH and mononuclear complexes, the oxidation reaction was performed in the presence of different concentrations of magnesium or barium ions, keeping fixed the concentration of the ferric complex and the substrate (Figure 4.8). Magnesium ions were chosen since small divalent cations are known to strongly interact with biological phosphorylated molecules such as NADH, NADPH or adenosine triphosphate (ATP), while barium ions present a lower binding affinity towards nucleotides.[24] As expected, magnesium ions decreased the ionic interactions between the NADH moiety and the mononuclear complex, thus reducing the rate constant of the oxidation process, while barium ions did not negatively influence the reaction kinetics. To further verify the importance of NADH-related ionic interactions, the concentration of NADH was varied as well, equimolar to Mg^{2+}, keeping fixed the amount of ferric complex. The resulting profile was then compared with the previous reaction rate vs substrate concentration plot obtained in the absence of magnesium ions (Figure 4.8). In the presence of magnesium ions, the reaction profile followed a Michaelis-Menten kinetics with a reduced efficiency than in the absence of the divalent cation, suggesting that the NADH

molecules would be less inclined to interact with the mononuclear complex and thus be oxidized.
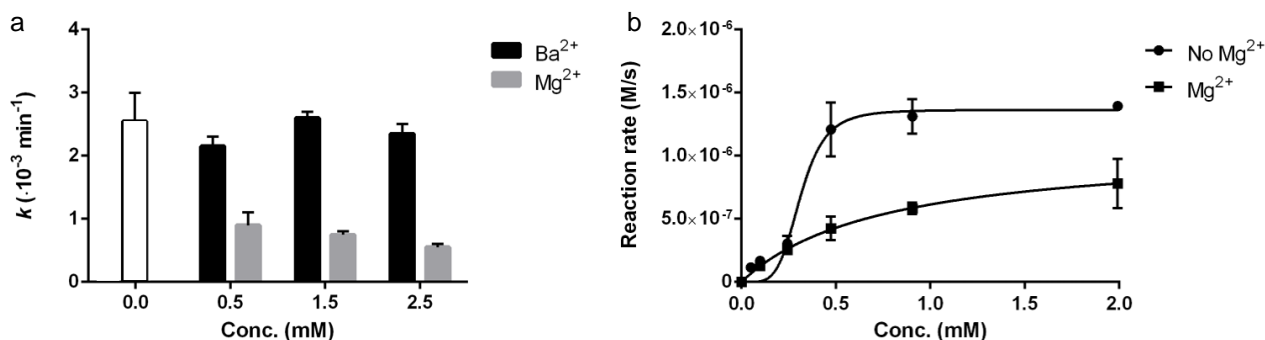
Figure 4.8 | Effect of magnesium and barium ions on the NADH oxidation kinetics. a, Effect induced by different concentration of Mg^{2+} (grey bars) or Ba^{2+} (black bars) ions on the rate constant of the NADH oxidation reaction catalysed by the L-glutathione mononuclear complex (0.5 mM Fe3+, 40 mM reduced L-glutathione). The control experiment in the absence of metal ions is reported as an empty bar. Presented data represent mean and SEM, n ≥ 2 replicates. b, Reaction rate vs substrate concentration graph with fixed concentration of mononuclear complex (0.5 mM Fe^{3+}, 40 mM reduced L-glutathione) and variable NADH concentration. Black dots represent reactions in the presence of Mg^{2+} ions (equimolar to NADH), grey dots represent reactions without Mg^{2+} ions. Presented data represent mean and SEM, n ≥ 2 replicates.

An alternative way to study the ionic interactions occurring between molecules in solution would be to change the peptide sequence of the coordinating ligand involved in iron binding. The commonly used procedures for solid phase peptide synthesis provide peptides in the cationic form with trifluoroacetate, acetate or chloride as counterions, which means that for solutions of chemically synthesized peptides the salt concentration and thus the ionic strength of the solution is higher than with commercial, harvested from biological sources, zwitterionic L-glutathione. In fact, when the NADH oxidation reaction is run in the presence of an excess of sodium chloride, the efficiency of the conversion process is strongly reduced, further evidencing the importance of specific ionic interactions for the reaction to occur (Figure A2.23). Therefore, since the removal of the trifluoroacetate counterion for such short peptides was problematic, the comparison between different ligands was done by comparing peptides synthesized with the same procedure.

Seven Cys-containing tripeptides were then synthesized and studied as coordinating ligands: GluCysGlu, GluCysGly, GlyCysGly, *N*-acetyl GlyCysGly amide, LysCysGly, GlyCysLys and LysCysLys. The results confirmed that the efficiency of the conversion is dependent on the molecular interactions occurring between the substrate and the ferric complexes. Since the NADH molecules is negatively charged (2-), the presence of positively

charged complexes (as in the case of the most efficient LysCysGly ferric complex) would allow ionic interactions to keep in close proximity the substrate moiety and the required mononuclear complexes. In the case of negatively charged complexes, hydrogen-bonding interactions would occur as well, involving the carbonyl, carboxyl, amino and amide groups present both in NADH and the peptide moiety (Figure A2.24 and Table A2.3). The profile of the *N*-acetyl GlyCysGly amide ferric complex exhibits a Michaelis-Menten kinetics for different NADH concentrations, because of the absence of any molecular interaction among spatially-close mononuclear complexes (Figure 4.9). Moreover, an increased conversion efficiency was recovered by exploiting the GlyCysGly and GluCysGly as coordinating moieties, since molecular interactions occurring among ionized groups were restored.
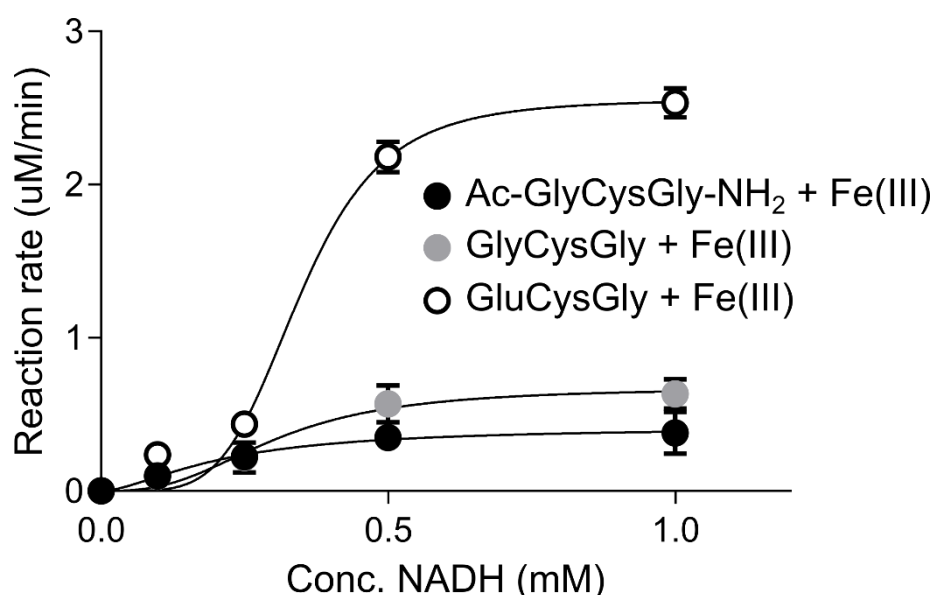
Figure 4.9 | Dependence of NADH oxidation kinetics on the peptide sequence coordinating the mononuclear complex. Reaction rate vs substrate concentration graph with fixed concentration of mononuclear complex (0.5 mM Fe^{3+}, 40 mM peptide) and variable NADH concentration. Black dots represent reactions with *N*-Acetyl Gly-Cys-Gly-NH_2 ferric complex, grey dots represent reactions with Gly-Cys-Gly ferric complex, and empty dots represent reactions with Glu-Cys-Gly ferric complex.

It was previously shown in Chapter 2 that the L-glutathione [2Fe-2S] cluster is redox active and can be subjected to several cycles of reduction and oxidation, whereas an analogous ability for the L-glutathione mononuclear ferric complex was not known yet. Moreover, the yield of NADH conversion was higher than 50% when an equimolar mixture of substrate and L-glutathione ferric complex was analysed, implying that the mononuclear complex was capable of undergoing several reduction and oxidation cycles. Therefore, L-glutathione-Fe^{3+} mononuclear complex was repetitively exposed to sodium dithionite as

reducing agent and hydrogen peroxide as oxidizing agent and UV-vis. spectra of the resulting solutions were recorded. The analysed data assessed that the mononuclear complex could be cyclically reduced and oxidized with minimal degradation (Figure 4.10).
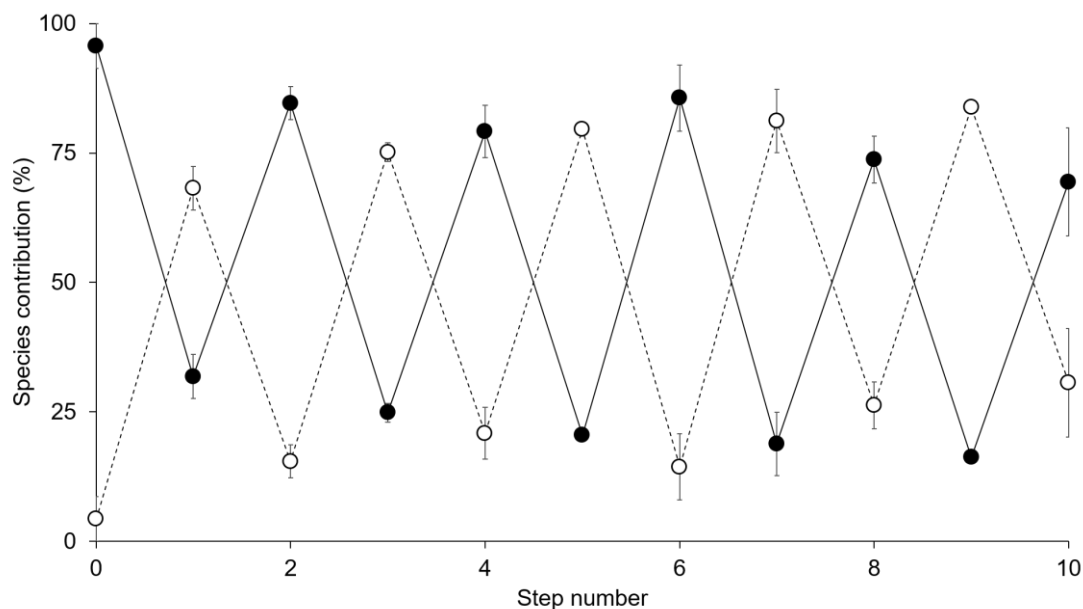
Figure 4.10 | Redox cycles of L-glutathione-Fe^{3+} mononuclear complex. Ferric complex (open circles) and ferrous complex (filled circles) coordinated to L-glutathione during reduction (sodium dithionite addition) and oxidation (hydrogen peroxide addition) determined by UV spectral decomposition. Presented data represent mean and SEM, $n \geq 2$ replicates.

The fact that more than 50% of NADH was oxidized to NAD+ after 6 h suggests that the ferrous complex formed by NADH conversion is constantly oxidized back by an oxidant present in solution to perform at least one more redox cycle. Since the oxidized form of L-glutathione is spontaneously converted from reduced L-glutathione in the presence of Fe(III), the opposite reaction (oxidation of ferrous to ferric complex) would not be spontaneous in the presence of oxidized L-glutathione. Another possibility was the ability of oxidized L-glutathione to oxidize NADH, so that even once all the ferric complex was over, the conversion of NADH would have been catalysed by oxidized L-glutathione. Thus, a mixture of oxidized L-glutathione and NADH was analysed by HPLC, showing that no reaction occurred in the absence of the mononuclear complex (Figure A2.25-A2.26). It was previously shown that, when NADH was not present, the mononuclear complex (0.5 mM) was able to considerably oxidize the free L-glutathione molecules present in excess (40 mM) in solution, up to 50% in 90 min (Figure A2.1). Surprisingly, when NADH is present in solution (0.5 mM), only NADH was oxidized, even though L-glutathione was in excess, showing a

strong specificity of the mononuclear complex towards the nucleotide instead of the peptide as substrate for the redox process (Figure A2.27). Therefore, since oxidized L-glutathione could not take part in the redox cycles, an additional component had to be considered for the oxidation of the ferrous complex. All the previous experiments were performed under anaerobic conditions by preparing the samples with a Schlenk line and Schlenk glassware. However, many manual steps have to be accomplished for the preparation of each sample, so that small infiltrations of oxygen have to be expected ($[O_2] \sim 2\text{-}10\cdot10^2$ ppm). To understand the role of molecular oxygen in the reaction, the NADH oxidation experiment was repeated either under aerobic conditions ($[O_2] \sim 2.1\cdot10^5$ ppm) or under strictly anaerobic conditions in a glove box ($[O_2] < 20$ ppm). In the first case, the massive presence of oxygen in solution allowed the fast degradation of the mononuclear complex and the precipitation of oxo-hydroxo iron species, with negligible conversion of NADH to NAD+ in 6 h. Conversely, in the strict absence of oxygen, the NADH oxidation occurred only to a low extent (Figure A2.28). Supposing that the presence of small quantities of oxidant are needed for the reduced mononuclear complex to be oxidized and continue to oxidize NADH, after 6 h an excess of hydrogen peroxide (H_2O_2) was added to the solution still, while maintaining strictly anaerobic conditions. HPLC data showed that the activity was recovered and NADH was converted comparably to previous results obtained in the presence of a small concentration of O_2. Moreover, hydrogen peroxide seems to be a more prebiotically plausible oxidant.[25] Thus, a redox cycle with NADH as reductant and H_2O_2 as oxidant would be more plausible (Figure A2.29).

Since the mononuclear complex resulted in an enzyme-like activity involving biologically relevant molecules, further experiments were performed to understand whether the NADH oxidation reaction was compatible with protocells. Fatty acid and phospholipid vesicles were shown to be stable in the presence of iron ions once coordinated by L-glutathione molecules (see Chapter 3, Table A2.4 and Figure A2.30). Therefore, NADH and the mononuclear complex were encapsulated within vesicles made with fatty acids (oleic acid) or phospholipids (POPC). After incubation, the vesicles were purified and analysed at specific time points by HPLC. The results showed that NADH oxidation occurs also within vesicles with a similar efficiency to the unencapsulated reaction (Figure 4.11 and A2.31).
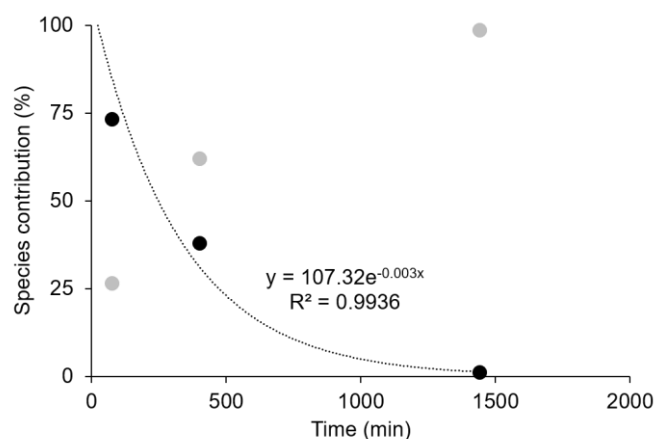
Figure 4.11 | NADH oxidation reaction occurring within phospholipid vesicles. First order kinetics was observed for the equimolar reaction between the mononuclear complex (0.5 mM) and NADH (0.5 mM) within phospholipid vesicles (20 mM). NADH and NAD+ concentrations are represented as black and grey dots, respectively.

In extant biology, a wide variety of electron acceptors are available to receive electrons from iron-sulfur enzymes, depending on the chosen metabolic pathway. In the electron transport chain, the final electron acceptor in the first step of the pathway is ubiquinone. Therefore, preliminary studies were performed exploiting the mononuclear complex as a "shuttle of electrons" from the electron donor NADH to the electron acceptor ubiquinone. The lipophilic ubiquinone was thus embedded in the phospholipid bilayer and NADH and the ferric complex were encapsulated. After extrusion (400 nm pore membrane) and purification under anaerobic conditions, HPLC analyses were performed to detect the conversion of ubiquinone (oxidized form) to ubiquinol (reduced form). The data showed that the reduction of ubiquinone to ubiquinol (up to ~60%) was detected, concomitant to the oxidation of NADH to NAD+ (Figure A2.32).

Since iron-sulfur proteins are usually embedded in or anchored to the membrane, the possibility to localize the mononuclear complex on the lipid layer was then tested. Arginine-rich peptides are known to interact with the polar headgroup of fatty acids and phospholipids, allowing for the localization of RNA to protocells' membrane.[26] Therefore, an arginine-rich peptide was designed containing an iron-binding motif (GCG), a short spacer (GAG) and a arginine tail (R$_8$). To better visualize fatty acid and phospholipid vesicles by confocal microscopy, the peptide was fluorescently-labelled with fluorescein (FITC). A similar fluorescently-labelled peptide without the arginine motif was used as a negative control. Oleic acid and POPC giant vesicles[27] were prepared and deposited on the glass slide, subsequently a freshly-prepared solution of peptide-mononuclear complex was added. The

collected images showed that only arginine-rich peptides are able to electrostatically interact with membranes and thus localize mononuclear complex to the membrane (Figure 4.12).
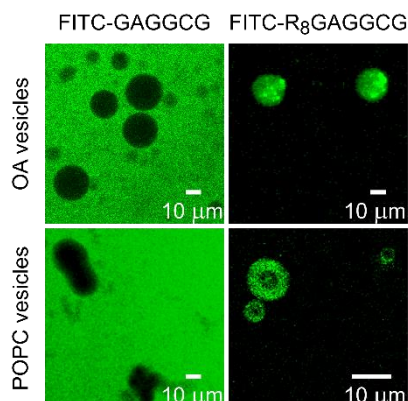
Figure 4.12 | Localization of the mononuclear complex on the membrane. The experiment was performed with OA (top) and POPC (bottom) vesicles. Left, The mononuclear complex is coordinated by the fluorescently-labelled GAGGCG peptide, therefore no localization to the membrane could occur (green background and dark vesicles). Right, The mononuclear complex coordinated by the fluorescently-labelled R_8GAGGCG peptide is localized to the membrane (green vesicles and dark background). Leica TCS SP5: confocal microscope, upright microscope (Leica DM6000 CS), lamp Ar (488 nm) 40x dry objective used for OA vesicles, 63x oil objective used only for POPC vesicles with R_8GAGGCG-mononuclear complex. Software: Leica Application Suite Advanced Fluorescence (LAS AF).

Therefore, further experiments were performed to evaluate to which extent the efficiency of the oxidation of NADH is affected by the localization of the mononuclear complex to the membrane. The mononuclear complex coordinated by the arginine-rich peptide in the absence of vesicles resulted inefficient for the conversion of NADH to NAD+, as previously reported for long peptides. However, in the presence of model protocells exploiting ubiquinone as the only electron acceptor, the ferric complex anchored to the membrane exhibited an increased NADH conversion efficiency up to 20% with respect to the non-anchored L-glutathione ferric complex.

In modern organisms the electron transport chain is related to the formation of a proton gradient between the two sides of the membrane. The proton motive force thus generated can be then exploited for the synthesis of ATP. The possibility to create a proton gradient and a chemical driving force for the synthesis of important biomolecules without complex enzymatic machineries could be an attractive perspective for prebiotic chemistry studies. Therefore, the oxidation of NADH driven by ferric complexes was followed in the presence of a pH indicator dye (thymol blue, $pK_a = 8.9$, from yellow to blue in the 8.0-9.2 range). While the oxidation of NADH involves the release of one H^+ ion per each NADH

molecule converted, the reduction of hydrogen peroxide and molecular oxygen induces the release of 2 OH⁻ ions per time.

$$H_2O_2 \rightarrow 2OH^- + 2e^-$$

$$\frac{1}{2}O_2 + H_2O \rightarrow 2OH^- + 2e^-$$
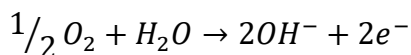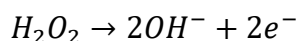
According to the redox reactions involved, a pH increase due to the excess of released OH⁻ would then be expected. The first experiments were performed in solution under strictly anaerobic conditions by using H_2O_2 to oxidize the ferrous complexes or under less strictly anaerobic conditions by exploiting small quantities of O_2. The results showed a shift of the solution colour from yellow to blue, resulting in an increase of 0.1 in the absorbance of the dye at 590 nm and thus in a pH change towards more alkaline values (~0.5 pH unit increase) (Figure A2.33-A2.34). Therefore, the reaction was performed within POPC vesicles (Figure 4.13). A variation in the absorbance at 590 nm similar to what observed in solution was recorded by UV-vis. spectroscopy within vesicles. Moreover, no leakage of thymol blue was observed over time, suggesting that the detected pH increase occurred inside the model protocells.
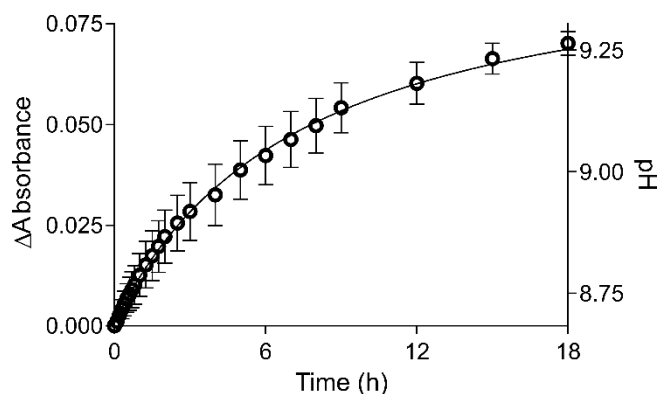
Figure 4.13 | 590 nm absorbance variation of the pH indicator dye within phospholipid vesicles. The experiment was set up as reported elsewhere with an excess of NADH (2.5 mM) and the mononuclear ferric complex (0.5 mM). Vesicles were extruded (400 nm) and purified, and no leakage of the pH dye was detected after 24 h, suggesting that the pH increase occurred inside the vesicles.

4.3 Discussion

The mononuclear complex formed by iron ions and simple organic thiols or Cys-containing peptides was able to influence the conversion of NADH and NADPH to NAD+ and NADP+ in a way that mimics modern-day rubredoxins, which are involved in several multi-complex enzymes, such as fatty acid and alkane hydroxylases. Moreover, the

mononuclear complex could be recycled in the presence of small amounts of oxidants, so that the reaction could be performed within protocells using a biologically relevant lipophilic electron acceptor embedded in the membrane and generating a proton gradient. Finally, the mononuclear complex could be anchored to protocell lipid layers exploiting electrostatic interactions between an arginine-rich motif and the negatively charged lipid headgroups, thus enhancing the electron transfer from NADH to ubiquinone or simpler prebiotic analogues. Therefore, the obtained results may allow to understand to which extent the mononuclear complex can mimic the electron transfer chain in a prebiotically plausible and biologically relevant way, being considered the first example of a primordial metabolic process that could be active on the early Earth. Moreover, due to the functional importance that flavin prosthetic groups exert together with iron-sulfur proteins in the electron transport chain, further analyses on the studied system can be performed in the presence of FAD or FMN. An interesting progression of the study would be to explore a broader metal peptide pool involving also His-binding peptides together with Cys-binding peptides, as analogues of the Rieske proteins, on different redox reactions. Such experiments could suggest the importance of peptide sequences and molecular interactions in the efficiency of redox reactions and could improve the understanding of how modern proteins evolved from simple metal peptides.

An additional study could then test how prebiotic chemistry may have exploited proton gradients formed across membranes. As previously mentioned, in modern biology proton gradients are coupled with electron transfer reactions and are used to drive physiologically essential reactions such as the synthesis of ATP. It is thus plausible that a prebiotic version of such simple energy storage mechanism could spontaneously emerge from solutions containing lipid membranes, iron-sulfur peptides, and redox active nucleotides, and be used to drive protometabolic networks. The demonstration of the prebiotic plausibility of such a system would greatly advance the understanding of possible paths towards protometabolic processes and give insight into what could be possible on other planets.

4.4 Materials and Methods

Materials. All reagents were bought from VWR Chemicals and Sigma Aldrich and used with no further purification. All the chromatographic columns were purchased from GE Healthcare. All the synthetic procedures were performed with a Schlenk line and Schlenk glassware under controlled N_2 flow. MilliQ water was distilled under N_2 atmosphere to remove oxygen from the solvent. The samples were preserved under N_2 atmosphere and

transferred to anaerobically sealed Hellma quartz cuvettes or anaerobically sealed glass vials for UV-vis. spectroscopy and HPLC studies, respectively.

UV-visible spectroscopy. An Agilent 8453 UV-vis. diode array spectrophotometer was used to collect UV-vis. absorption spectra of freshly prepared solutions (integration time = 0.5 s, interval = 1 nm).

Thiol-Fe^{3+} mononuclear complex synthesis. An aqueous solution of each thiol (40 mM, 1 mL) was prepared in a Schlenk round bottom flask under anaerobic conditions and the pH was adjusted to a value equal to the pK_a. Unless otherwise reported, ferric chloride ($FeCl_3 \cdot 6H_2O$, final concentration 0.5 mM) was added to synthesize the mononuclear complex.

NADH oxidation experiments. An aqueous solution of each thiol (40 mM, 1 mL) was prepared in a Schlenk round bottom flask under anaerobic conditions and the pH was adjusted to a value equal to the pK_a. Unless otherwise reported, NADH (final concentration 0.5 mM) and ferric chloride ($FeCl_3 \cdot 6H_2O$, final concentration 0.5 mM) were added. For kinetics studies, NADH was added to each sample to a final concentration of 0.05 mM, 0.1 mM, 0.25 mM, 0.5 mM, 1 mM and 2.5 mM and the data analyses performed correcting for the volume increase. For Mg^{2+} and Ba^{2+} studies, divalent cations were added prior to NADH addition in order to maximize the eventual interaction between the metal ions and the nucleotide molecules.

Proton gradient detection experiments. An aqueous solution of L-glutathione (40 mM, 1 mL) was prepared in a Schlenk round bottom flask under anaerobic conditions or in the glove box and the pH was adjusted to a value equal to the pK_a. NADH (final concentration 2.5 mM), ferric chloride ($FeCl_3 \cdot 6H_2O$, final concentration 0.5 mM) and thymol blue (final concentration 6.25 uM) were added. Absorbance spectra were recorded up to 24 hours and the contribution of the mononuclear complex at 590 nm was removed for analysis purposes. For the glove box experiments, hydrogen peroxide was added after 2 hours from the initial mixing of the reactants. For the experiments performed with POPC vesicles, vesicles were prepared, extruded and purified as described elsewhere. After 24 h, a leakage test was run to evaluate the amount of pH dye leaked out. No leakage was detected.

Cyclic voltammetry. Cyclic voltammetry experiments were performed in a three-electrode cell having a glassy-carbon or a 4,4'-bipyridyl-modified gold working electrode (1.5 mm in diameter), a platinum-wire counter electrode and an Ag/AgCl reference electrode. An Amel Model 566 analogue function generator and an Amel Model 552 potentiostat were used as polarizing units. The Amel potentiostat was connected to an Amel Model 558 integrator.

Under the present experimental conditions, the one electron oxidation of ferrocene occurs at $E°' = +0.435$ V vs Ag/AgCl at +20°C.

Redox titration experiments. Solutions containing 0.5 mM mononuclear complex, [2Fe-2S] cluster, and [4Fe4S] cluster were prepared as previously described. The selected redox dye was then added (final concentration 0.05 mM). A stock solution of sodium dithionite (10 mM) was then prepared and a UV-vis. spectrum was recorded after each aliquot (10 uL). The UV-vis. spectra were then decomposed using a|e UV-vis. IR spectral software to quantify the amount of oxidized and reduced iron-sulfur species, and oxidized and reduced dye. The spanned midpoint potential range was then calculated through the Nernst equation by using the redox dye concentration values.

High performance liquid chromatography. The HPLC measurements were performed with a Shimadzu High Performance LC system (CBM-20 A, binary pump LC-20AB, Italy). The analytes were chromatographically resolved working in reverse-phase (Kinetex C18 column, 100 Å pore size, 4.6 mm ID, 2.6 μm particle size, 25 cm length, Phenomenex, Italy). For NADH/NAD+ detection, the mobile phase was composed of solvent A, 20 mM pH 5.1 ammonium acetate, and solvent B, acetonitrile. The gradient elution program started with 0% B and reached 15% B in 15 min, at the steady flow rate of 1.0mL/min. For reduced L-glutathione/oxidized L-glutathione, the mobile phase was composed of solvent A, water + 0.1% trifluoroacetic acid, and solvent B, acetonitrile + 0.1 trifluoroacetic acid. The gradient elution program started with 3% B and reached 18% B in 15 min, at the steady flow rate of 1.0 mL/min. For ubiquinol/ubiquinone, the mobile phase was composed of solvent A, hexane:isopropanol:methanol 2:1:1, and solvent B, isopropanol:methanol 1:9. The gradient elution program started with 90% B and reached 100% B in 10 min, at the steady flow rate of 1.0mL/min. Each experiment was performed twice. HPLC data were analysed using the Origin Pro software. For each peak, the absorbance values were divided by the extinction coefficient of each species. The resulting peaks were then integrated to get quantitative information.

Solid-phase peptide synthesis (SPPS). All the peptides were synthesized according to previously published SPPS procedures by using Fmoc-protected L-amino acids. *N,N*-dimethyl formamide (DMF) was used as the solvent and preloaded fluorenylmethyloxycarbonyl-glycyl-Wang resin (Fmoc-Gly-Wang) was used as the starting polymeric support. Trityl-protected Fmoc-cysteine (Fmoc-Cys(Trt)-OH) and tert-butyl (OtBu) side chain-protected Fmoc-α-amino acids were used as building blocks. In general, the peptide chain was elongated by sequential Fmoc deprotection of the residue anchored to

the resin and Fmoc-AA-OH coupling. Acetic anhydride in DMF was used to achieve *N*-acetylation. Fmoc deprotection was obtained by washing the mixture with 20% (v/v) solution of piperidine in DMF. For each coupling, an excess (Fmoc-AA-OH: anchored AA, 4:1) of the Fmoc-α-amino acid derivative was added to the resin. Apart from Fmoc-Cys(Trt)-OH, Fmoc-α-amino acid derivatives were activated with a mixture of hydroxyl-benzotriazole (HOBt), *N*,*N*,*N'*,*N'*-tetramethyl-*O*-(benzotriazol-1-yl)uronium tetrafluoborate (TBTU), and *N*,*N*-diisopropylethyl amine (DIPEA). Fmoc-Cys(Trt)OH was activated with a *N*,*N'*-diisopropylcarbodiimide (DIC)/HOBt mixture. At the end of the coupling, the polymers were cleaved from the resin and deprotected by treatment with a solution of trifluoroacetic acid (TFA):H_2O:triisopropyl silane (TIS):1,2-ethanedithiol (EDT) (volume ratio 97:1:1:1) for 2 h. The volume was successively reduced under nitrogen atmosphere to avoid cysteinyl thiol oxidation, and the product was precipitated with a cold solution of diethyl ether/petroleum ether (30:70 (v/v)) followed by washing cycles with diethyl ether or extracted 3 times with 20% acetic acid/chloroform and finally dried under inert atmosphere.

Competent cells. *E. coli* BL21(DE3) pLysS cells were grown overnight at 37°C in Luria Broth (LB) medium (Sigma Aldrich) and bacteria were then diluted 1:100 into 50 mL of fresh medium and incubated at 37°C until an OD_{600} of 0.5 was reached. The culture was then centrifuged at 4°C at 5000 xg for 15 minutes. The supernatant was discarded and the cell pellet was resuspended in 15 ml of Transformation Buffer (10 mM Tris-HCl pH 7.0, 50 mM $CaCl_2$) and stored in ice for 15 min. Cells were then centrifuged at 4°C at 5000 xg for 15 min, the supernatant was discarded and the pellet was resuspended in 1 ml of Transformation Buffer. Cells were aliquoted with 10-15% v/v glycerol and stored at -80 °C.

Transformation. An aliquot of *E. coli* BL21(DE3) pLysS competent cells (200 µL) was thawed in ice and mixed with 100 ng of plasmid encoding the *C. pasteurianum* rubredoxin gene under the control of an IPTG inducible T7 promoter. Cells were transformed in ice for 30 min, heat shocked at 42°C for 2 min and cooled in ice for 1 min. Cells were then incubated at 37°C in the termoshaker with the addition of 500 µL of LB for 1 hour. Finally, bacteria were plated on LB agar (Sigma Aldrich) plates supplemented with antibiotic (50 µg/ml ampicillin) and incubated at 37°C overnight.

Rubredoxin expression. A colony of the previously transformed *E. coli* strain was inoculated in 50 mL of LB medium at 37°C overnight. 10 mL of the inoculum was transferred into two different flasks and mixed with 1 L of Terrific Broth medium. The culture was grown until OD_{600} of 0.6 was reached and then induced with IPTG (1 mM). A supplement of iron was added at the moment of the induction (0.5 mM $FeCl_3$).

Rubredoxin purification. Induced cells were centrifugated at 4°C at 18000 xg for 40 min, the supernatant was discarded and the pellet was resuspended in 50 mL of 50 mM Tris-HCl pH 7.5 and sonicated at 4°C keeping the tube on ice for short pulses (5-10 s) followed by 50 s pauses until the solution became clear. After centrifugation at 18000 xg at 4°C for 20 min, the protein present in the supernatant fraction was purified by ion exchange chromatography, performed with a DEAE resin column, eluting with 100-300 mM NaCl gradient in 50 mM Tris-HCl buffer pH 7.4. A second run of purification was performed with the same column eluting with 50-300 mM NaCl gradient in 50 mM Tris-HCl buffer pH 7.4.

Vesicle formation. Under anaerobic conditions, 20 mM POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine) vesicles were prepared as previously described by thin-film rehydration using a solution containing the Fe^{3+} mononuclear complex (0.2 M L-glutathione, 5 mM Fe^{3+}, pH 8.6) in the absence or in the presence of NADH (5 mM). Samples were briefly vortexed, tumbled at room temperature for 30 min, then extruded with 11 passages through a 400 nm pore membrane (Avanti Polar Lipids) using a Mini-Extruder (Avanti Polar Lipids). Vesicles samples (200 uL) were then loaded for purification onto a Sepharose 4B (6 mL bed volume) size exclusion column to remove unencapsulated material, using 0.2 M Tris-HCl pH 8.5 as the running buffer. Elution was monitored at 420 nm and the fractions containing vesicles were collected.

Vesicle stability assay. 20 mM POPC containing 0.1 mM fluorescein (MW 4000) were prepared and purified by size exclusion chromatography with a Sepharose 4B column. 500 µL of the vesicle solution was then incubated for 0 min, 180 min, 360 min with 500 µL of a L-glutathione-Fe^{3+} mononuclear complex (5 mM Fe^{3+}, 200 mM GSH) in the presence of NADH (5 mM). All the samples were loaded onto a Sepharose 4B (6 mL bed volume) size exclusion column to assess vesicle stability, using 0.2 M Tris-HCl, pH 8.5 as the running buffer. The elution profile was monitored by fluorescence (excitation 490 nm, emission 525 nm).

Vesicle localization of the ferric complex. Oleic acid and phospholipid vesicles localization experiments were prepared as previously reported. Lipid films were hydrated with a solution of 100 mM TRIS and 200 mM sucrose pH 8.5 and giant vesicles were allowed to form at 37°C overnight. A solution of each fluorescently-labelled and not fluorescently-labelled peptide (1:20 ratio, 1 mM peptide content) was used to avoid saturation of the fluorescence signal. 100 mM TRIS and 200 mM glucose pH 8.5 was used to solubilize the peptides and keep an alkaline pH. Ferric ions (12.5 uM) were then added to form the mononuclear complex. For confocal microscopy analyses, 10 uL droplets were deposited

on untreated glass slides with a final concentration of 2 mM vesicles, 25 μM peptide and 0.3125 uM mononuclear complex. For visualization purposes a Leica TCS SP5: confocal microscope with upright microscope (Leica DM6000 CS) was used, equipped with a Ar lamp (488 nm), and 40x dry objective (63x oil objective used only for $R_8GAGGCG$ with POPC vesicles). Software: Leica Application Suite Advanced Fluorescence (LAS AF).

4.5 Bibliography

1.   Deamer, D. & Weber, A. L. Bioenergetics and Life ' s Origins. 1–16 (2010).

2.   Adamala, K. & Szostak, J. W. Nonenzymatic template-directed RNA synthesis inside model protocells. *Science* **342,** 1098–100 (2013).

3.   Ferris, J. P. Mineral Catalysis and Prebiotic Synthesis: Montmorillonite-Catalyzed Formation of RNA. *Elements* **1,** 145–150 (2005).

4.   Scott, W. G., Szöke, A., Blaustein, J., O'Rourke, S. M. & Robertson, M. P. RNA catalysis, thermodynamics and the origin of life. *Life* **4,** 131–41 (2014).

5.   Mitchell, P. Proton translocation mechanisms and energy transduction by adenosine triphosphatases: an answer to criticisms. *FEBS Lett.* **50,** 95–97 (1975).

6.   Mitchell, P. Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. *Biol. Rev.* **41,** 445–502 (1966).

7.   Koch, A. L. Primeval cells: Possible energy-generating and cell-division mechanisms. *J. Mol. Evol.* **21,** 270–277 (1985).

8.   Koch, A. L. & Schmidt, T. M. The First Cellular Bioenergetic Process: Primitive Generation of a Proton-Motive Force. *J. Mol. Evol.* **33,** 297–304 (1991).

9.   Calvin, M. Evolution of enzymes and the photosynthetic apparatus. *Science* **130,** 1170–4 (1959).

10.   Shelton, V. M. & Morrow, J. R. Catalytic Transesterification and Hydrolysis of Rna by Zinc(II) Complexes. *Inorg. Chem.* **30,** 4295–4299 (1991).

11.   Schrauzer, G. N., Kiefer, G. W., Doemeny, P. A. & Kisch, H. Chemical evolution of a nitrogenase model. VI. Reduction of cyanide ion, azide ion, nitrous oxide, molecular nitrogen, and other substrates by molybdocysteine catalysts in the presence of nucleoside phosphates. *J. Am. Chem. Soc.* **95,** 5582–5592 (1973).

12.   Egami, F. Origin and Early Evolution of Transition Element Enzymes. *J. Biochem.* **77,** 1165–1169 (1975).

13.   Lippard, S. J. Iron sulfur coordination compounds and proteins. *Acc. Chem. Res.* **6,** 282–288 (1973).

14. Ycas, M. On certain homologies between proteins. *J.Mol.Evol.* **7,** 215–244 (1976).

15. Baltscheffsky, H. A new hypothesis for the evolution of biological electron transport. *Orig. Life* **5,** 387–395 (1974).

16. Miseta, A. & Csutora, P. Relationship Between the Occurrence of Cysteine in Proteins and the Complexity of Organisms. 1232–1239 (1989).

17. Belmonte, L. & Mansy, S. S. Metal Catalysts and the Origin of Life. *Elements* **12,** 413–418 (2016).

18. Neville, R. G. The Oxidation of cysteine by Iron and Hydrogen Peroxide. *J. Am. Chem. Soc.* **79,** 2456–2457 (1957).

19. Bohning, J. J. & Weiss, K. The Kinetics of the Oxidation of 3-Mercaptopropionic Acid with Potassium Ferricyanide. *J. Am. Chem. Soc.* **82,** 4724–4728 (1960).

20. Van Beilen, J. B. *et al.* Rubredoxins involved in alkane oxidation. *J. Bacteriol.* **184,** 1722–1732 (2002).

21. Bak, D. W. & Elliott, S. J. Alternative FeS cluster ligands: tuning redox potentials and chemistry. *Curr. Opin. Chem. Biol.* **19,** 50–58 (2014).

22. Chang, F. C. & Swenson, R. P. Regulation of oxidation-reduction potentials through redox-linked ionization in the Y98H mutant of the Desulfovibrio vulgaris [Hildenborough] flavodoxin: direct proton nuclear magnetic resonance spectroscopic evidence for the redox-dependent shift in the . *Biochemistry* **36,** 9013–21 (1997).

23. Verkhovskaya, M. & Bloch, D. A. Energy-converting respiratory Complex I: On the way to the molecular mechanism of the proton pump. *Int. J. Biochem. Cell Biol.* **45,** 491–511 (2013).

24. Nanninga, L. The association constant of the complexes of adenosine triphosphate with magnesium, calcium, strontium, and barium ions. *Biochim. Biophys. Acta* **16,** 330–338 (1961).

25. Ball, R. & Brindley, J. Hydrogen peroxide thermochemical oscillator as driver for primordial RNA replication. *J. R. Soc. Interface* 1–8 (2014).

26. Kamat, N. P., Tobé, S., Hill, I. T. & Szostak, J. W. Electrostatic Localization of RNA to Protocell Membranes by Cationic Hydrophobic Peptides. *Angew. Chemie* **127,** 11901–11905 (2015).

27. Zhu, T. F., Budin, I. & Szostak, J. W. *Preparation of fatty acid or phospholipid vesicles by thin-film rehydration. Methods in enzymology* **533,** (Elsevier Inc., 2013).

Chapter 5 - Final Remarks

5.1 Conclusions

The formation of water-insoluble iron-sulfur precipitates achieved with iron(II) ions in sulfur-rich environments restrained the presence of iron-sulfur polymers to extreme deep-sea hydrothermal settings. While such iron-sulfur minerals could have helped catalyse metabolic-like reactions at the emergence of life, a path from iron-sulfur minerals to biologically active iron-sulfur proteins is not apparent. In the last decades, many efforts have been made to study the spontaneous assembly of Fe-S clusters coordinated by non-genetically encoded sterically hindered thiolates mainly in organic solvents, such as methanol, tetrahydrofuran or acetonitrile. However, hydrolytically stable iron-sulfur clusters coordinated by small thiols have not been frequently studied or well characterized in anaerobic aqueous environment. In this thesis, I expanded previously published data on iron-sulfur cluster synthesis, focusing on how a nascent primordial cellular system could have exploited short iron-sulfur clusters coordinated by peptides. Under prebiotically reasonable conditions the simple tripeptide L-glutathione was shown to coordinate and stabilize a [2Fe-2S] cluster upon adding sulfide and ferric ions in aqueous solution. Moreover, the binding of polypeptides containing repeated sequences of L-glutathione resulted in more stable iron-sulfur clusters and ferredoxin-like coordination motifs.

The low accessibility of sulfide ions, mainly present in volcanic environments, and iron ions, mostly available in the reduced form, challenged the understanding of how iron-sulfur cluster synthesis could have been driven on the early Earth. Sunlight can be considered the Earth's largest energy source, so that photochemical reactions must be considered fundamental in the chemical evolution of biomolecules that eventually led to life. UV radiation has been suggested as a fundamental component in the prebiotic synthesis of amino acid, sugar, nucleotide, and lipid precursors, and thus all the cellular subsystems could have arisen simultaneously through common photochemistry. In this thesis, I described two critical steps in the iron-sulfur cluster assembly mechanism to be mediated by UV light coming from the young Sun. Indeed, the photolysis of thiol-containing molecules induced by UV light allows the release of the sulfide ions needed for the assembly of polynuclear iron-sulfur clusters. Moreover, ferrous ions can be photooxidised by UV light, generating the ferric ions necessary for iron-sulfur cluster synthesis. The proposed synthetic procedure of prebiotic iron-sulfur cluster shows little dependence on the amino acid

composition of the coordinating peptide. Such findings broaden the pool of sequences capable of stabilizing iron-sulfur clusters and the absence of important requirement concerning the amino acid availability consolidate the prebiotic plausibility of iron-sulfur clusters. Moreover, the compatibility of the described chemical pathway for the synthesis of iron-sulfur clusters within model fatty acid protocells further suggests that biological iron-sulfur cluster assembly could have occurred at a very early phase in the evolution of life. Finally, the chemical process found appeared strikingly similar to the way iron-sulfur clusters are synthesized in modern living cells. Indeed, the UV light-driven assembly of iron-sulfur clusters is analogous to the protein mediated biosynthetic processes for the synthesis of iron-sulfur clusters. At some point during the evolution timeline, an enzyme started to catalyse the cleavage of the carbon-sulfur bond converting cysteine molecules into alanine residues and releasing sulfide implemented in the formation of iron-sulfur clusters, perhaps replacing the ancestral role of UV-light along the same pathway. The observed strong parallelism could suggest that a deep knowledge on how the molecules needed for life were prebiotically synthesized could help us to better observe modern biochemical behaviours. Moreover, the study of extant biomolecules may strongly enhance our understanding of how their primordial ancestors emerged on early Earth.

Iron-sulfur proteins can perform a wide variety of functions in living cells. Iron-sulfur clusters are mainly known as redox-active biological cofactors, able to catalyse one-electron transfer reactions exploiting the redox couple Fe(II)/Fe(III). Prebiotic iron-sulfur peptides could be then studied to reproduce extant metabolic reactions catalysed by iron-sulfur proteins in a prebiotically plausible way in order to determine the chemical roots of protometabolism. In this thesis, I showed that ferric ions coordinated by small organic thiols or cysteine-containing tripeptides can induce the oxidation of biologically relevant electron donors such as NADH and NADPH within fatty acid and phospholipid vesicles. The electron transfer is enhanced by molecular interactions present between mononuclear complex molecules and the substrate and is dependent on the ionic strength of the solution. The studied redox system was encapsulated within model protocells and implemented with electron acceptors such as ubiquinone, hydrogen peroxide and molecular oxygen, to determine the ability of iron-sulfur species to undergo consecutive redox cycles. Moreover, a proton gradient induced by hydrogen peroxide and molecular oxygen reduction was detected, suggesting that iron-sulfur systems could serve as useful models and plausible prebiotic analogues of the prebiological steps leading to the actual biological ion pumps.

5.2 Future perspectives

The future step of this work would be to further expand the pool of reactions which could be catalysed by peptide-coordinated iron-sulfur clusters. In addition to their electron transfer function, iron-sulfur proteins act as catalytic centres and sensors of iron and oxygen. Iron-sulfur clusters may be inserted or removed from proteins, can influence protein structure by preferential side chain ligation, and can be interconverted. Moreover, structurally advanced iron-sulfur clusters can take part in activation and conversion reactions, by substituting labile ligands and accommodating substrate molecules. Taken together, iron-sulfur proteins are known to be fundamental to several physiological processes, including photosynthesis, central metabolism, nitrogen fixation and protein and DNA synthesis. Therefore, water-soluble iron-sulfur clusters coordinated by peptides embedded within protocells could be tested for a wide variety of biologically relevant reactions which could have played a role in the origin and the evolution of life-like systems.

An interesting alternative would focus on iron-sulfur proteins as the starting point for a top-down approach-based study. Indeed, the amino acid removal, insertion and modification of extant iron-sulfur enzymes could lead to new clues in the evolution of iron-sulfur clusters on the early Earth. It is known that iron-sulfur clusters can be easily interconverted and protein motifs accommodating iron-sulfur moieties are also found in Zn^{2+}-based enzymes. Therefore, the identification of the chemical and biological features which enhance the stabilization of a specific metal species would allow researchers in the field of origins of life to understand how and when specialized iron-sulfur proteins started to play a fundamental role in biology. In conclusion, studies on how iron-sulfur proteins evolved may lead to a better understanding of how their primordial ancestors emerged on early Earth.

Appendix 1

Supplementary Materials for Chapter 3

Table A1.1 | Chemical structures of studied Cys-containing tripeptide sequences.

Peptide seq.	Peptide structure	Peptide seq.	Peptide structure
ACA		GCQ	
ACG		GCS	
ACT		GCT	
CGG		GCV	
AcCGG		GGC	
CγEG		GMG	
AcCγEG		KCG	
DCG		KCK	
βDCG		NCG	
ECE		QCG	
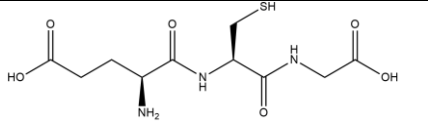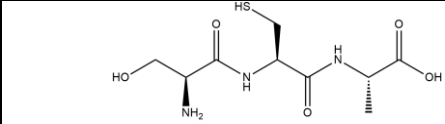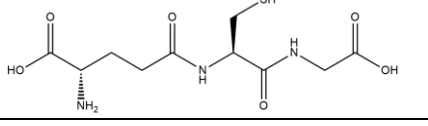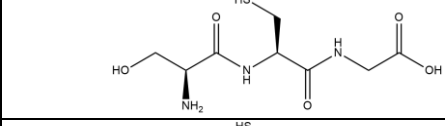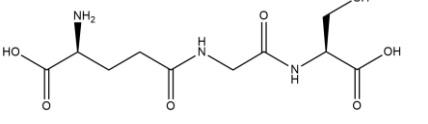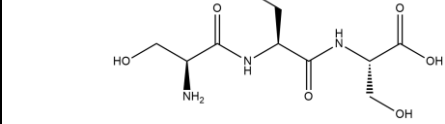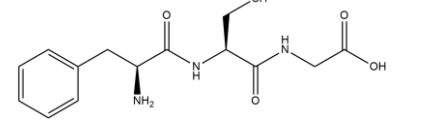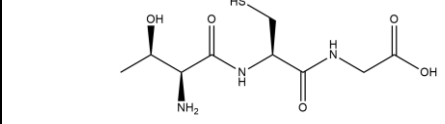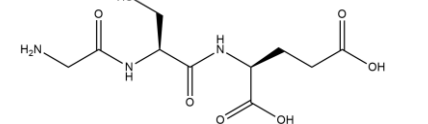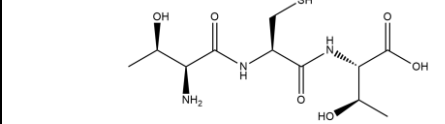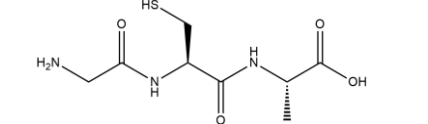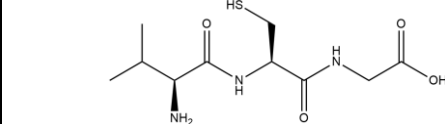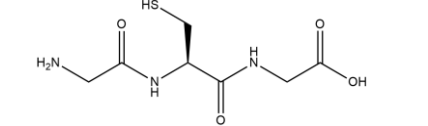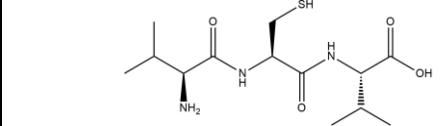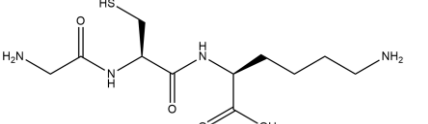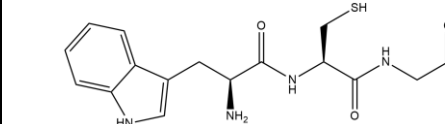
ECG		**SCA**	
γECG		**SCG**	
γEGC		**SCS**	
FCG		**TCG**	
GCA		**TCT**	
GCE		**VCG**	
GCG		**VCV**	
GCK		**WCG**	

Table A1.2 | Inflection points and pK$_a$ values for model Cys-containing tripeptides.

Peptide	pH - first inflection point	pH - second inflection point	Thiol pK$_a$
GGC	5.02	10.88	8.01
GCG	6.25	11.06	8.04
CGG	5.52	10.61	6.91
AcCGG	6.12	10.47	8.06
ECG	5.83	11.03	7.74
SCG	5.20	10.99	7.28
VCG	5.32	10.70	7.68
VCV	5.74	10.44	7.61

Table A1.3 | Chemical structures of studied Cys-containing dipeptides and thiol-containing molecules.
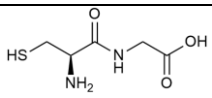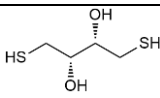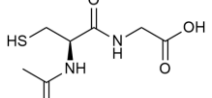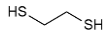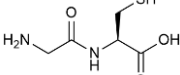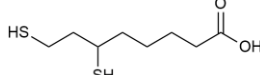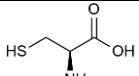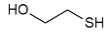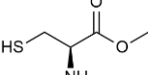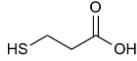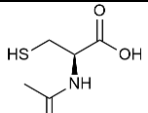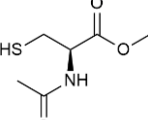
Ligand	Structure	Ligand	Structure
CG		DL-dithiothreitol	
AcCG		1,2-ethanedithiol	
GC		dihydrolipoic acid	
L-cysteine		2-mercaptoethanol	
L-cysteine methyl ester		3-mercaptopropionic acid	
N-acetyl L-cysteine		2-mercaptopyrimidine	
N-acetyl L-cysteine methyl ester		cysteamine	

Table A1.4 | Mössbauer data obtained on Fe^{3+} mononuclear complex, [2Fe-2S], and [4Fe-4S] clusters with L-glutathione (mms^{-1}, errors $\leq \pm 0.01$ mms^{-1} unless shown otherwise in parenthesis, at 80 K).

Irradiation time	IS*	QS*	h.w.h.m.*	Molar contribution (%)	Identity	Geometry
0 s	0.68	3.26	0.22	100	Fe^{2+}	4-coord., T_d
30 s	0.68	3.26	0.23	79	Fe^{2+}	4-coord., T_d
	0.27	0.52	0.21	21	Fe^{3+}	[2Fe-2S], T_d
180 s	0.68	3.32	0.23	26	Fe^{2+}	4-coord., T_d
	0.31	0.57	0.22	64	Fe^{3+}	[2Fe-2S], T_d
	0.48	1.06(2)	0.26(2)	10	$Fe^{2+/3+}$	[4Fe-4S], T_d

*IS, isomeric shift; QS, quadrupole splitting; h.w.h.m., half of the line width at half maximum.

Table A1.5 | Chemical shift (ppm) data collected from ^{1}H-NMR spectra of thiol containing solutions exposed to UV irradiation at 254 nm.

	Reduced thiol (ppm)		Oxidized thiol (ppm)		Desulfurized analogue (ppm)	
	R-CH$_2$-C*H*$_2$-SH	R-C*H*$_2$-CH$_2$-SH	(R-CH$_2$-C*H*$_2$-S)$_2$	(R-C*H*$_2$-CH$_2$-S)$_2$	R-CH$_2$-C*H*$_3$	R-C*H*$_2$-CH$_3$
2-Mercapto ethanol	δ 2.57 (t, J = 7.0 Hz, 2H)	δ 3.55 (t, J = 7.0 Hz, 2H)	δ 2.83 (t, J = 6.0 Hz, 2H)	δ 3.81 (t, J = 6.0 Hz, 2H)	δ 1.11 (t, J = 7.1 Hz, 2H)	δ 3.59 (q, J = 7.1 Hz, 3H)
3-Mercapto propionic acid	δ 2.62 (t, J = 7.6 Hz, 2H)	δ 2.37 (t, J = 7.6 Hz, 2H)	δ 2.53 (t, J = 7.1 Hz, 2H)	δ 2.86 (t, J = 7.1 Hz, 2H)	δ 0.99 (t, J = 7.7 Hz, 2H)	δ 2.11 (q, J = 7.7 Hz, 3H)
Cysteamine	δ 2.69 (t, J = 6.7 Hz, 2H)	δ 3.04 (t, J = 6.7 Hz, 2H)	δ 2.91 (t, J = 6.5 Hz, 2H)	δ 3.25 (t, J = 6.5 Hz, 2H)	δ 1.20 (t, J = 7.3 Hz, 2H)	δ 2.98 (q, J = 7.3 Hz, 3H)

Table A1.6 | Chemical shift (ppm) data collected from ^{1}H-NMR spectra of protected and unprotected L-cysteine and L-glutathione solutions exposed to UV irradiation at 254 nm.

	Reduced thiol (ppm)		Oxidized thiol (ppm)		Desulfurized analogue (ppm)	
	R-CH-CH_AH_B-SH	R-CH_X-CH$_2$-SH	(R-CH-CH_AH_B-S)$_2$	(R-CH_X-CH$_2$-S)$_2$	R-CH-CH_3	R-CH-CH$_3$
L-cysteine	δ 2.93 (d, J = 5.3 Hz, 2H)	δ 3.77 (t, J = 5.3 Hz, 1H)	δ$_A$ 3.27 (dd, 14.7, 4.0 Hz, 1H) δ$_B$ 3.06 (dd, 14.7, 8.0 Hz, 1H)	δ$_X$ 3.94 (dd, 8.0, 4.0 Hz, 1H)	δ 1.40 (d, J = 7.2 Hz, 3H)	δ 3.70 (q, J = 7.2 Hz, 1H)
L-cysteine methyl ester	δ$_A$ 2.89 (dd, 13.9, 5.3 Hz, 1H) δ$_B$ 2.81 (dd, 13.9, 4.9 Hz, 1H)	δ$_X$ 3.80 (dd, 5.3, 4.9 Hz, 1H)	δ$_A$ 3.07 (dd, 14.1, 5.3 Hz, 1H) δ$_B$ 3.00 (dd, 14.1, 6.4 Hz, 1H)	δ$_X$ 3.86 (dd, 6.4, 5.3 Hz, 1H)	δ 1.34 (d, J = 7.2 Hz, 3H)	δ 3.83 - 3.77 (m, 1H)
***N*-acetyl L-cysteine**	δ$_A$ 2.87 (dd, 13.7, 4.2 Hz, 1H) δ$_B$ 2.78 (dd, 13.7, 7.3 Hz, 1H)	δ$_X$ 4.28 - 4.15 (m, 1H)	δ$_A$ 3.18 (dd, 14.0, 4.1 Hz, 1H) δ$_B$ 2.88 (dd, 14.0, 4.7 Hz, 1H)	δ$_X$ 4.47 - 4.37 (m, 1H)	δ 1.25 (d, J = 7.3 Hz, 3H)	δ 4.12 - 4.00 (m, 1H)
***N*-acetyl L-cysteine methyl ester**	δ 2.89 (d, J = 5.7 Hz, 2H)	δ 4.53 (t, J = 5.7 Hz, 1H)	δ$_A$ 3.22 (dd, 14.4, 4.7 Hz, 1H) δ$_B$ 2.98 (dd, 14.4, 8.4 Hz, 1H)	n.d.	δ 1.33 (d, J = 7.3 Hz, 3H)	δ 4.36 - 4.25 (m, 1H)
L-glutathione	δ$_A$ 2.90 (dd, 14.0, 5.1 Hz, 1H) δ$_B$ 2.83 (dd, 14.0, 7.4 Hz, 1H)	δ$_X$ 4.43 (dd, 7.4, 5.1 Hz, 1H)	δ$_A$ 3.24 (dd, 14.2, 4.5 Hz, 1H) δ$_B$ 2.90 (dd, 14.2, 8.4 Hz, 1H)	n.d.	δ 1.33 (d, J = 7.2 Hz, 3H)	δ 4.29 (q, J = 7.2 Hz, 1H)

Table A1.7 | Chemical structures of studied lipid systems.
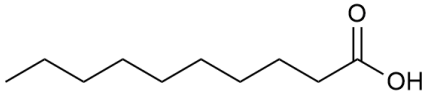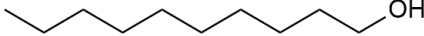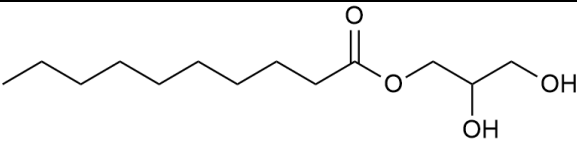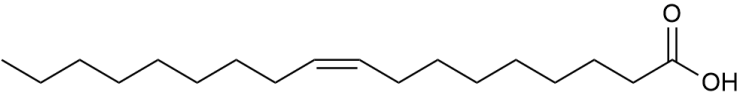
Lipid name	Lipid structure
Decanoic acid	
Decanol	
Glycerol monodecanoate	
cis-9-octadecenoic acid (oleic acid)	

Table A1.8 | Mössbauer data obtained on irradiated and non-irradiated oleate vesicles containing Fe^{3+} -L-glutathione (mms^{-1}, errors ≤ ± 0.01 mms^{-1} unless shown otherwise in parenthesis, at 80 K).

Irradiation time	IS*	QS*	h.w.h.m.*	Molar contribution (%)	Identity	Geometry
0 s	0.72	3.50(3)	0.16(2)	23	Fe^{2+}	4-coord., distorted T_d
	0.70	3.11(3)	0.20	37	Fe^{2+}	4-coord., T_d
	0.38	0.70	0.29	40	Fe^{3+}	4-coord., distorted T_d
180 s	0.72	3.49(2)	0.18	32	Fe^{2+}	4-coord., distorted T_d
	0.69	3.13(2)	0.21	54	Fe^{2+}	4-coord., T_d
	0.28	0.53	0.21	14	Fe^{3+}	[2Fe-2S], T_d

*IS, isomeric shift; QS, quadrupole splitting; h.w.h.m., half of the line width at half maximum.
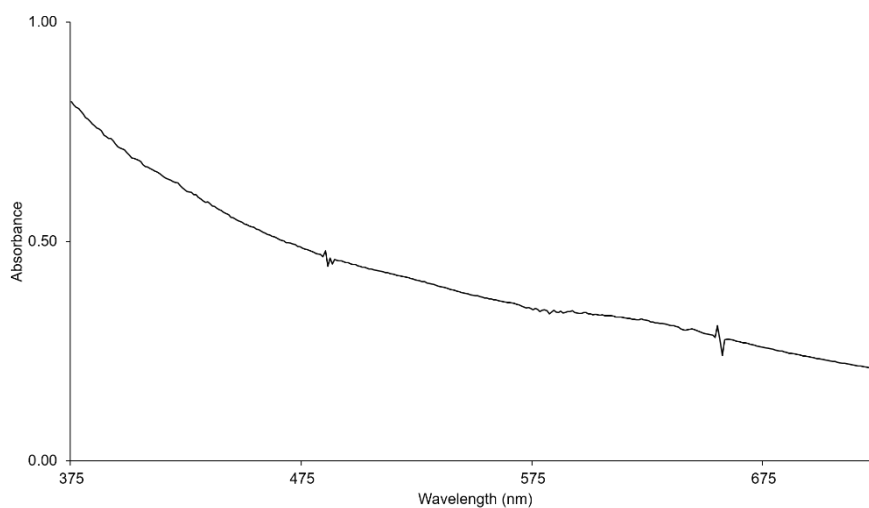
Figure A1.1 | UV-vis spectrum of a solution containing 40 mM GMG, 0.5 mM iron ions, 0.2 mM sulfide at pH 8. The [2Fe-2S] cluster is not detectable.
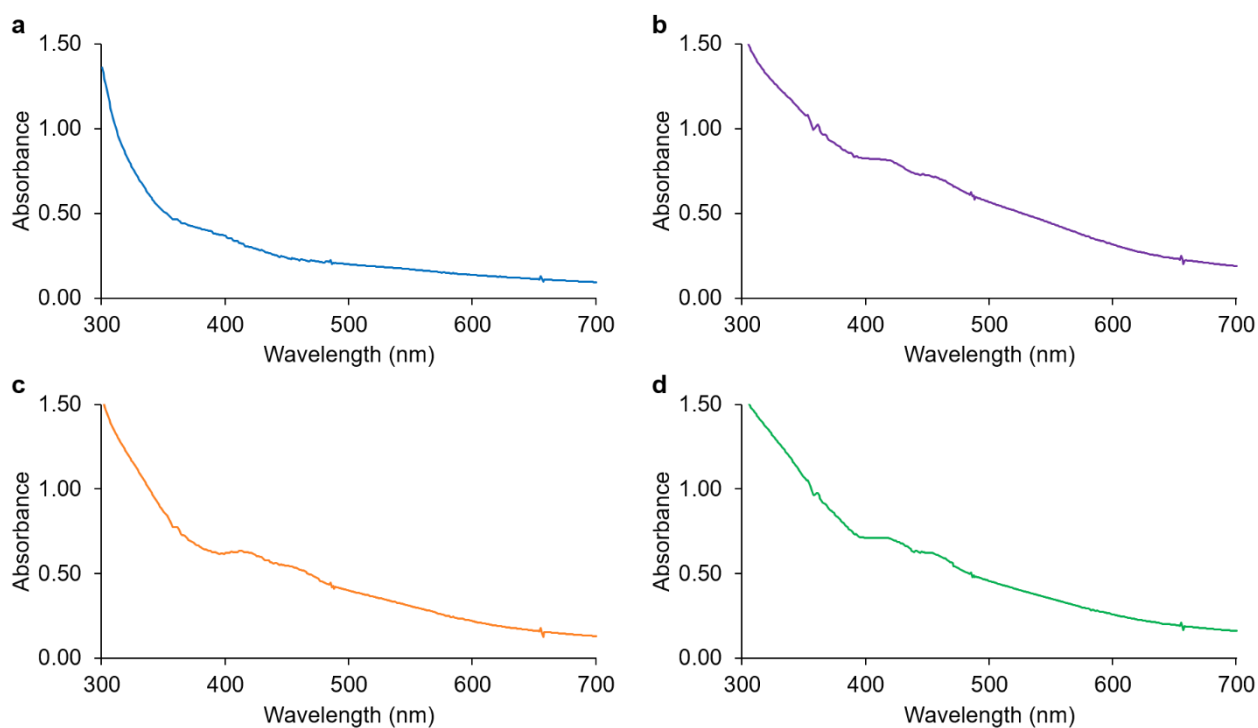
Figure A1.2 | a, UV-vis spectrum of the CγEG peptide in the presence of Fe^{3+} and S^{2-} ions. The [2Fe-2S] cluster is not detectable. b, UV-vis spectrum of [2Fe-2S] cluster coordinated by γECG. c, UV-vis spectrum of [2Fe-2S] cluster coordinated to γEGC. d, UV-vis spectrum of [2Fe-2S] cluster coordinated to AcCγEG.
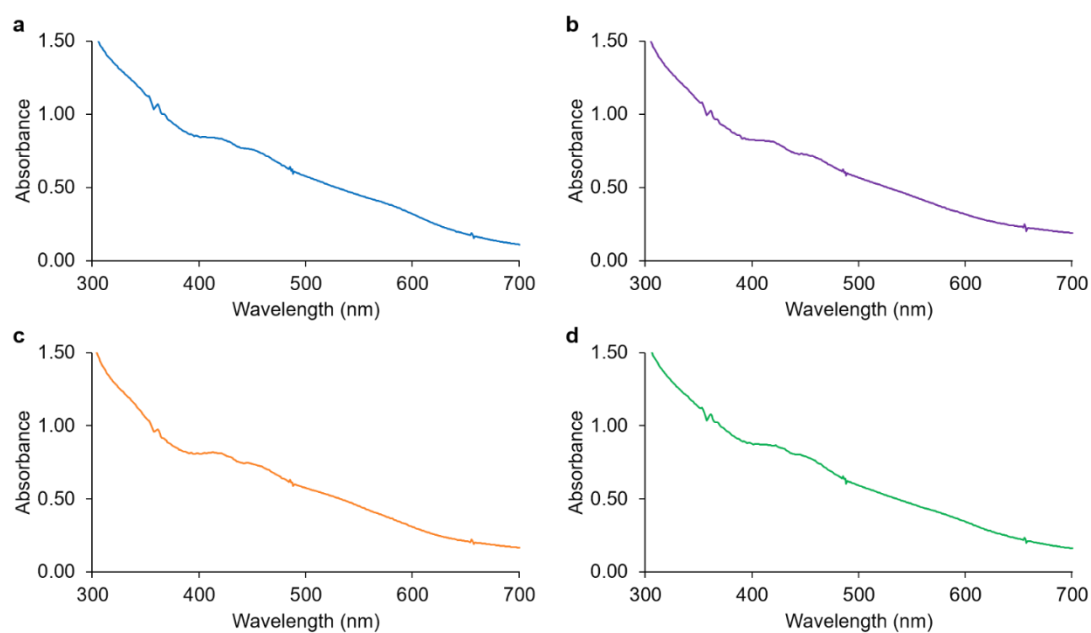
Figure A1.3 | a, UV-vis spectrum of [2Fe-2S] cluster coordinated to DCG. b, UV-vis spectrum of [2Fe-2S] cluster coordinated to γECG. c, UV-vis spectrum of [2Fe-2S] cluster coordinated to βDCG. d, UV-vis spectrum of [2Fe-2S] cluster coordinated to ECG.
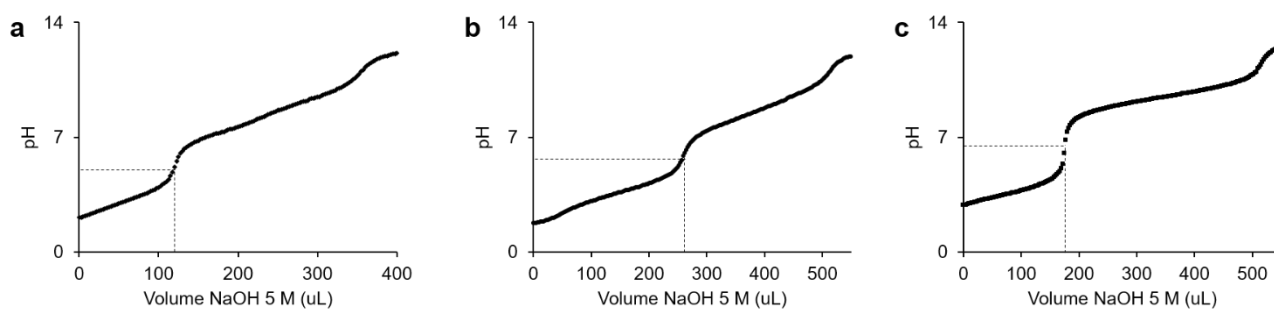
Figure A1.4 | a, pH titration of SCG. b, pH titration of ECG. c, pH titration of γECG. The first inflection point is highlighted for each titration. Acetylation of the amino-terminus increased the pK_a of the thiol to 8.1. For amino-terminal cysteine peptides, the pK_a was lower (ca. 6.9) due to the proximity of the α-amino group to the side-chain thiol. For these cases, the pH was adjusted to 6.9 or 7.3, in order to shift the equilibrium of H_2S dissociation. However, *N*-terminal Cys peptides failed to coordinate a [2Fe-2S] cluster in both the cases.
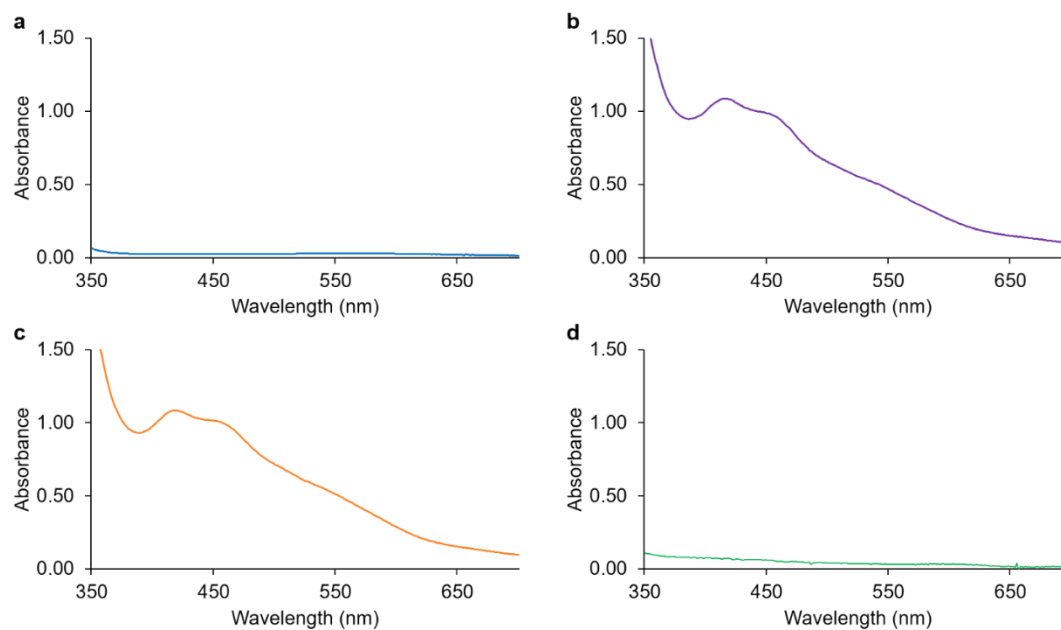
Figure A1.5 | a, UV-vis spectrum of L-cysteine methyl ester in the presence of Fe^{3+} and S^{2-} ions. The [2Fe-2S] cluster is not detectable. b, UV-vis spectrum of [2Fe-2S] cluster coordinated by *N*-acetyl L-cysteine methyl ester. c, UV-vis spectrum of [2Fe-2S] cluster coordinated by *N*-acetyl L-cysteine. d, UV-vis spectrum of L-cysteine in the presence of Fe^{3+} and S^{2-} ions. The [2Fe-2S] cluster is not detectable.
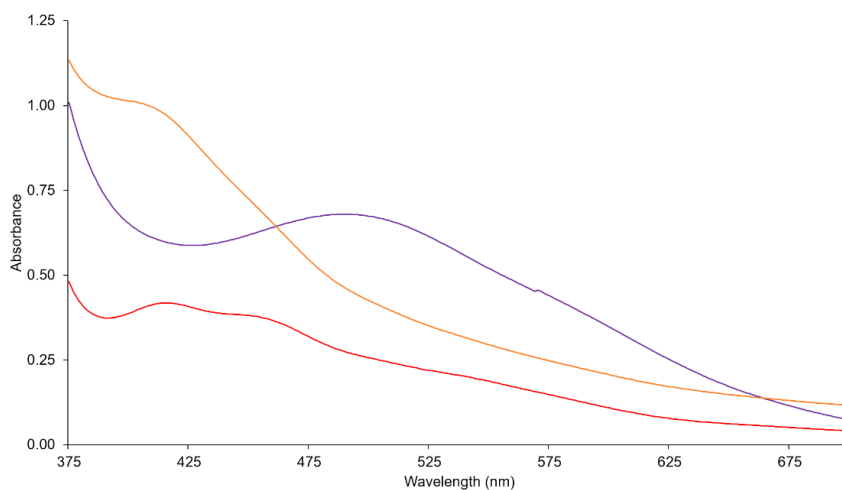
Figure A1.6 | UV-vis spectra of a solution of 2-mercaptoethanol and Fe^{3+} exposed to UV light (254 nm). The Fe^{3+}-mononuclear complex (violet line) is converted within 30 s to the corresponding [2Fe-2S] cluster (red line). Upon further irradiation (180 s) the formation of the [4Fe-4S] cluster can be detected (orange line).
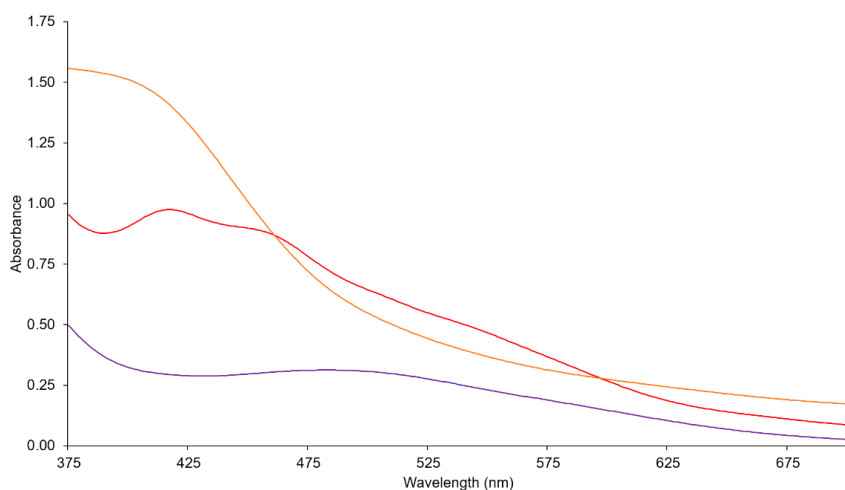
Figure A1.7 | UV-vis spectra of a solution of 3-mercaptopropionic acid and Fe^{3+} exposed to UV light (254 nm). The Fe^{3+}-mononuclear complex (violet line) is converted within 30 s of exposure to UV light to the corresponding [2Fe-2S] cluster (red line). Upon further irradiation (180 s) the formation of the [4Fe-4S] cluster can be detected (orange line).
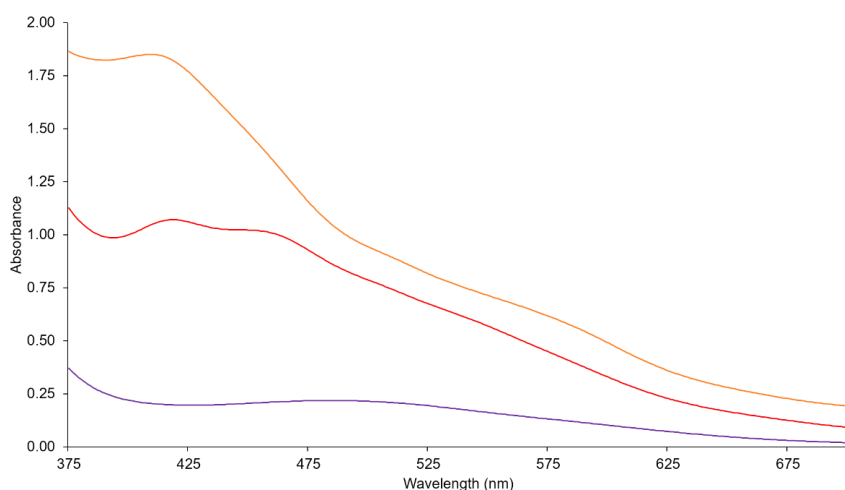
Figure A1.8 | UV-vis spectra of a solution of *N*-acetyl L-cysteine and Fe^{3+} exposed to UV light (254 nm). The Fe^{3+}-mononuclear complex (violet line) is converted within 30 s of exposure to UV light to the corresponding [2Fe-2S] cluster (red line). Upon further irradiation (180 s) the formation of the [4Fe-4S] cluster can be detected (orange line).
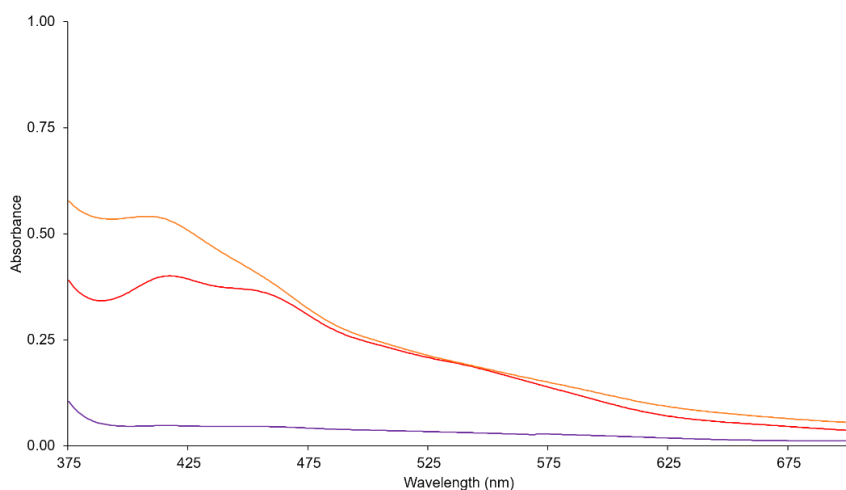
Figure A1.9 | UV-vis spectra of a solution of *N*-acetyl L-cysteine methyl ester and Fe^{3+} exposed to UV light (254 nm). The Fe^{2+}-mononuclear complex (violet line) is converted within 30 s of exposure to UV light to the corresponding [2Fe-2S] cluster (red line). Upon further irradiation (180 s) the formation of the [4Fe-4S] cluster can be detected (orange line).
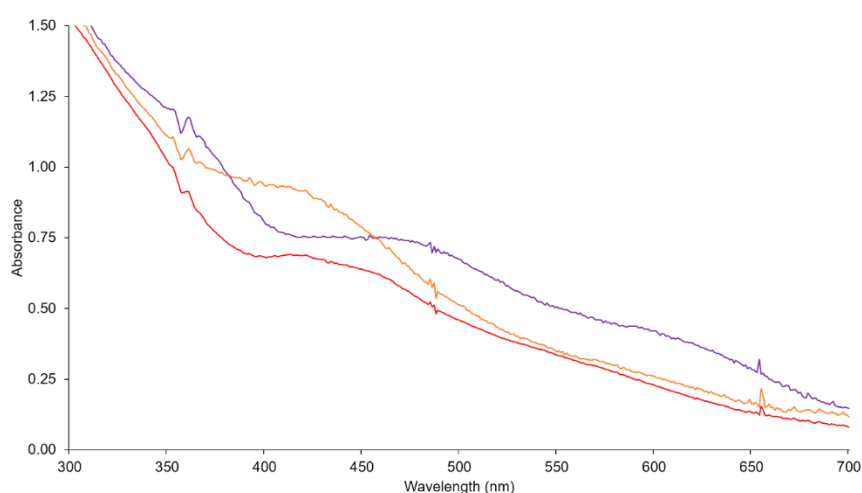
Figure A1.10 | UV-vis spectra of a solution of ACG and Fe^{3+} exposed to UV light (254 nm). The Fe^{2+}-mononuclear complex (violet line) is converted within 30 s of exposure to UV light to the corresponding [2Fe-2S] cluster (red line). Upon further irradiation (180 s) the formation of the [4Fe-4S] cluster can be detected (orange line).

111
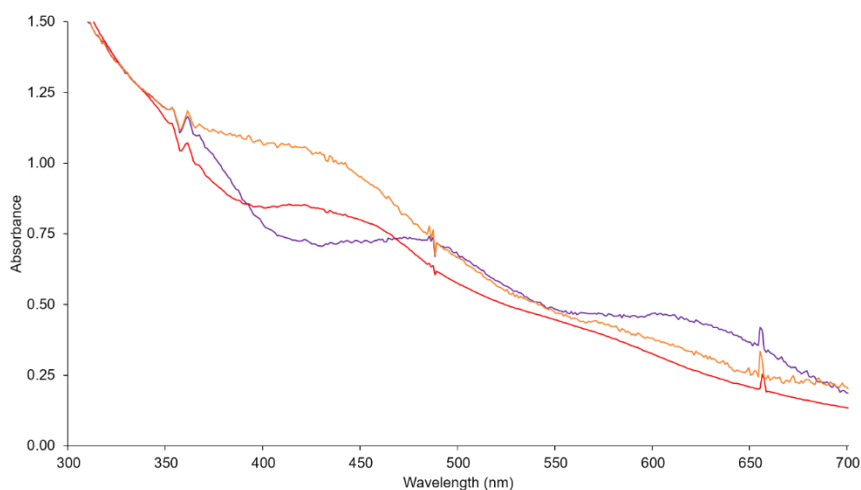
Figure A1.11 | UV-vis spectra of a solution of KCG and Fe^{3+} exposed to UV light (254 nm). The Fe^{2+}-mononuclear complex (violet line) is converted within 30 s of exposure to UV light to the corresponding [2Fe-2S] cluster (red line). Upon further irradiation (180 s) the formation of the [4Fe-4S] cluster can be detected (orange line).
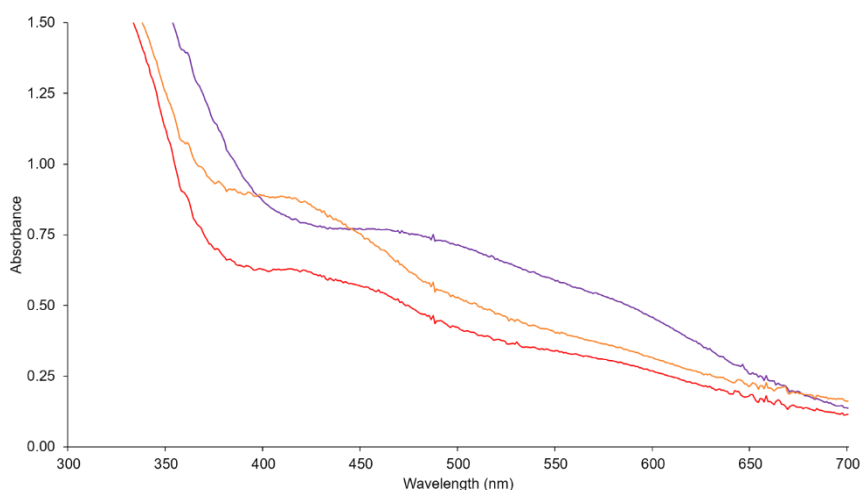
Figure A1.12 | UV-vis spectra of a solution of GCK and Fe^{3+} exposed to UV light (254 nm). The Fe^{2+}-mononuclear complex (violet line) is converted within 30 s of exposure to UV light to the corresponding [2Fe-2S] cluster (red line). Upon further irradiation (180 s) the formation of the [4Fe-4S] cluster can be detected (orange line).
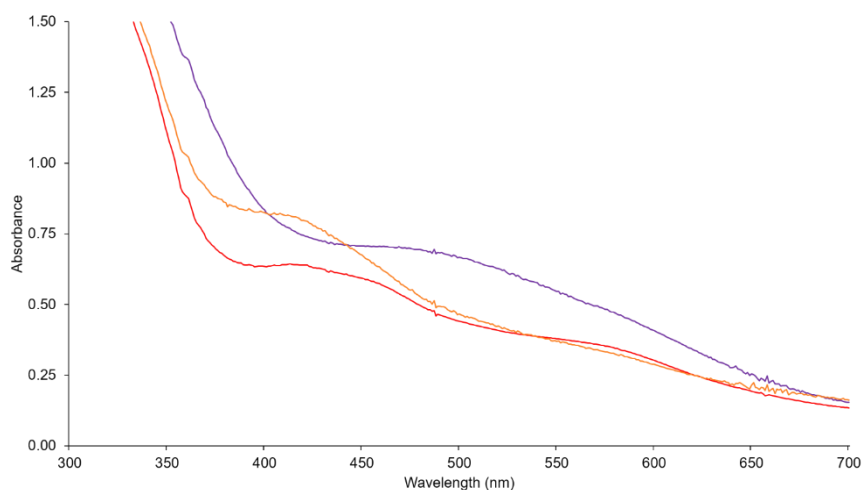
Figure A1.13 | UV-vis spectra of a solution of TCT and Fe^{3+} exposed to UV light (254 nm). The Fe^{2+}-mononuclear complex (violet line) is converted within 30 s of exposure to UV light to the corresponding [2Fe-2S] cluster (red line). Upon further irradiation (180 s) the formation of the [4Fe-4S] cluster can be detected (orange line).
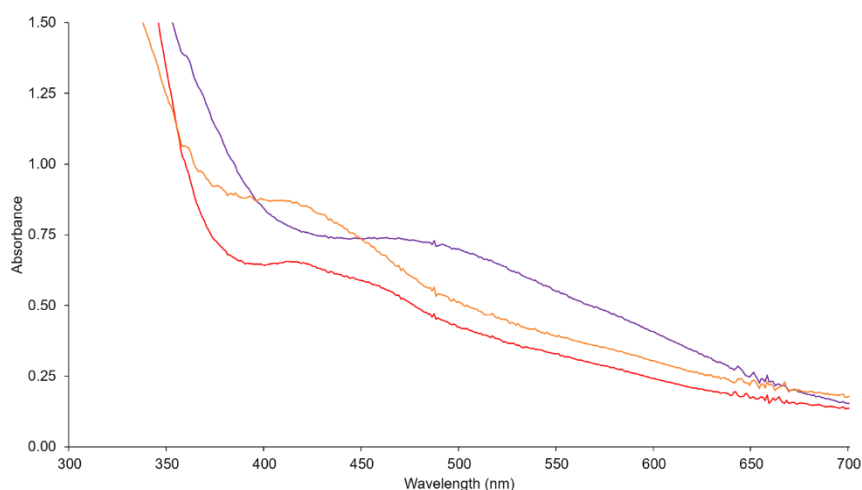
Figure A1.14 | UV-vis spectra of a solution of ACT and Fe^{3+} exposed to UV light (254 nm). The Fe^{2+}-mononuclear complex (violet line) is converted within 30 s of exposure to UV light to the corresponding [2Fe-2S] cluster (red line). Upon further irradiation (180 s) the formation of the [4Fe-4S] cluster can be detected (orange line).
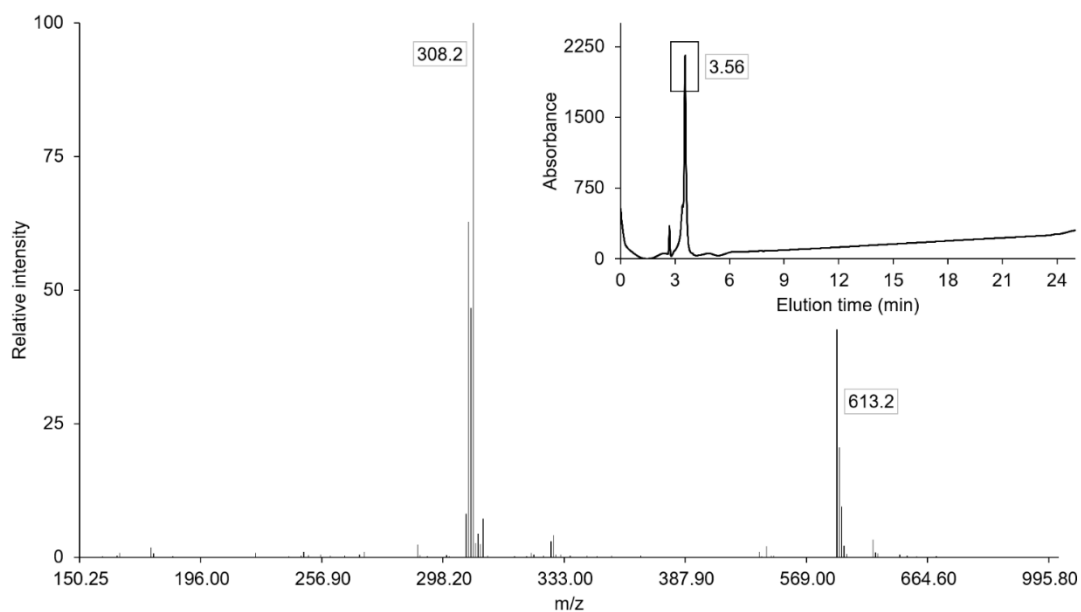
Figure A1.15 | LC-MS analysis of L-glutathione solution before irradiation. One peak was detected in the chromatogram (inset, elution time: 3.56 min, detector: 215 nm) which corresponds to L-glutathione, as shown by the mass spectrum.
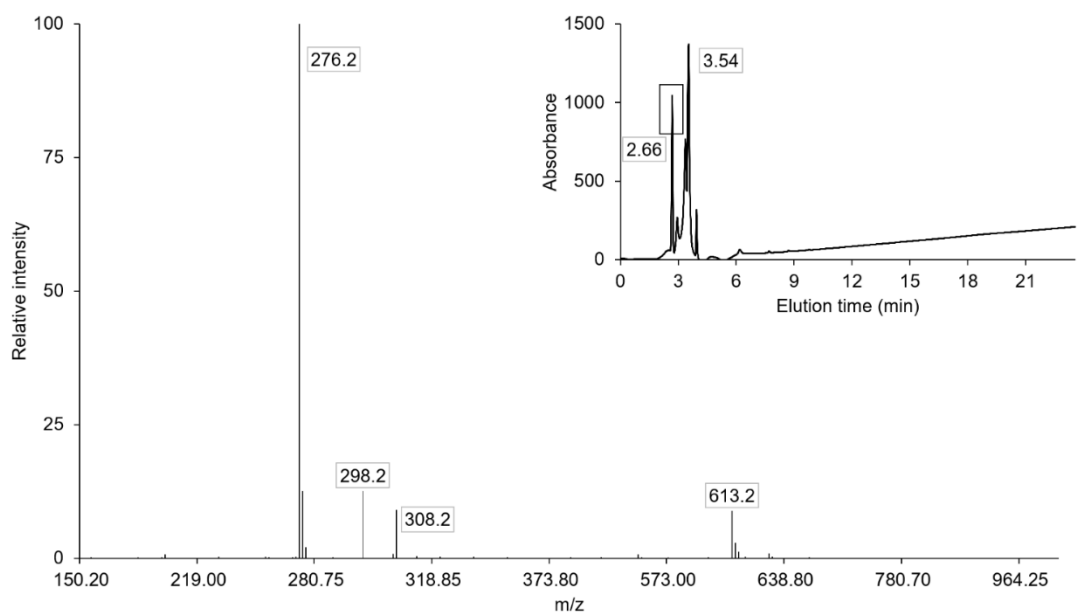
Figure A1.16 | LC-MS analysis of L-glutathione solution after 15 min irradiation. Two peaks were detected in the chromatogram (inset, elution time: 2.66 and 3.54 min, detector: 215 nm), which correspond to the alanyl analogue of L-glutathione, as shown by the mass spectrum, and L-glutathione.
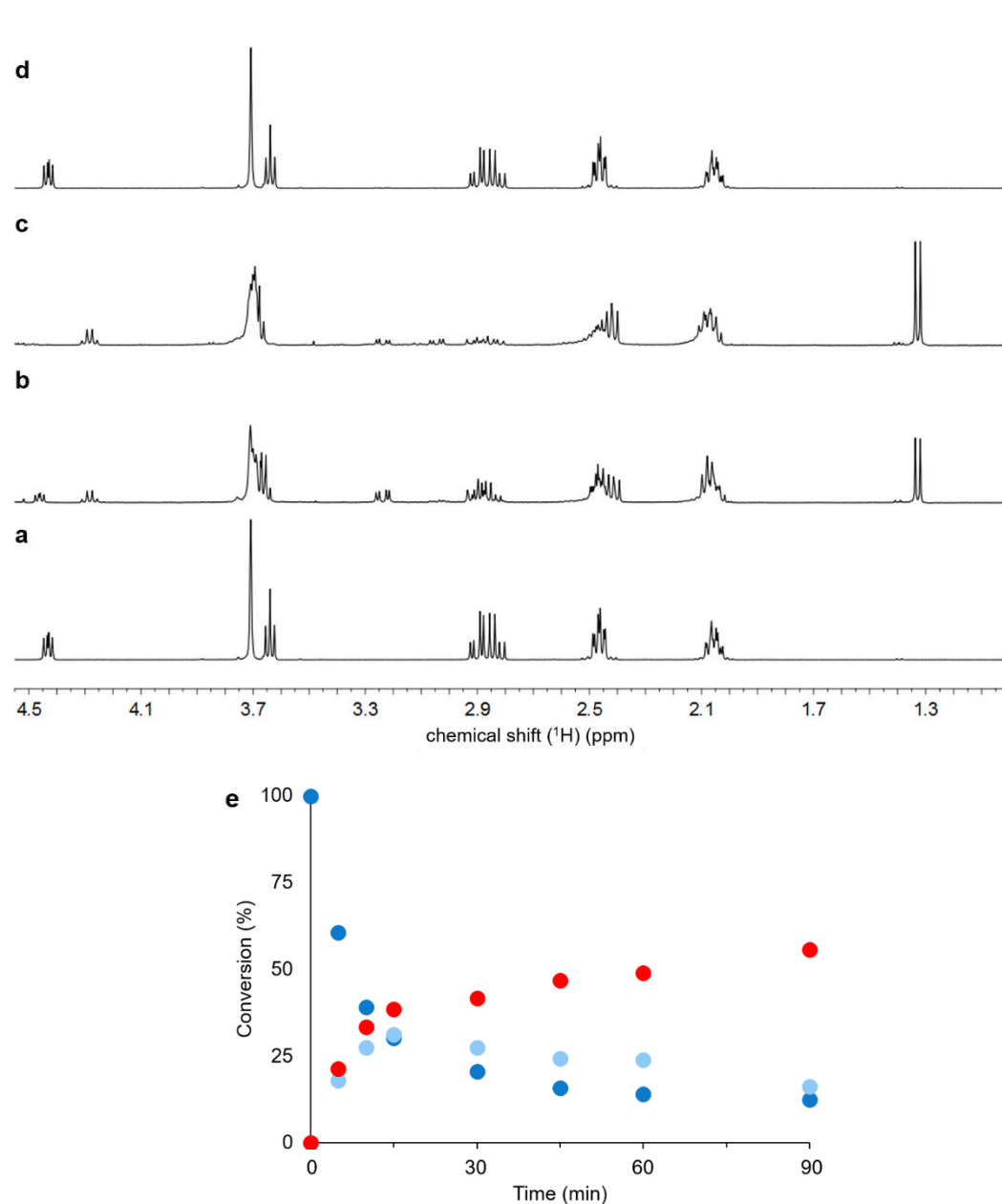
Figure A1.17 | a, ^{1}H-NMR spectrum of a L-glutathione solution before exposure to UV light in the absence of hypophosphite. b, ^{1}H-NMR spectrum of a L-glutathione solution after 15 min exposure to UV light in the absence of hypophosphite. The alanyl species is already detectable at 1.33 ppm together with the disulfide species of L-glutathione at 2.94 and 3.23 ppm. c, The conversion to the alanyl-analogue is almost complete after 90 min of exposure to UV light. d, ^{1}H-NMR spectrum of a L-glutathione solution without exposure to UV light after 90 min. e, Time course of conversion as obtained by integration of NMR peaks for the irradiated solution of L-glutathione in the absence of hypophosphite. Blue, reduced L-glutathione; light blue, oxidized L-glutathione; red, alanyl-analogue.
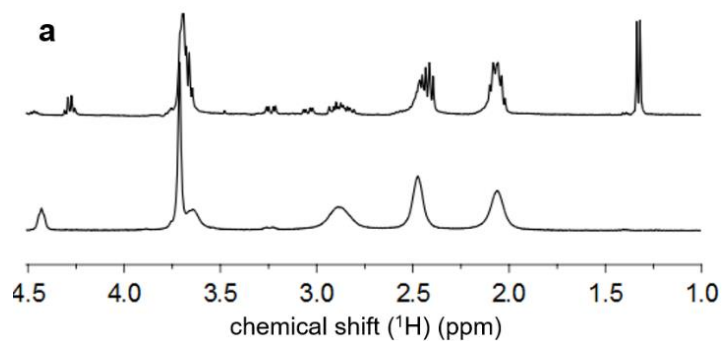
Figure A1.18 | ^{1}H-NMR spectra of an irradiated solution of L-glutathione in the presence of ferric ions. The bottom spectrum is of the starting solution. The top spectrum is after 15 min of irradiation, which shows a new peak at 1.33 ppm indicating that photolysis occurs also in the presence of iron ions. Significant line broadening due to the paramagnetism of the iron ions and ligand exchange were observed.
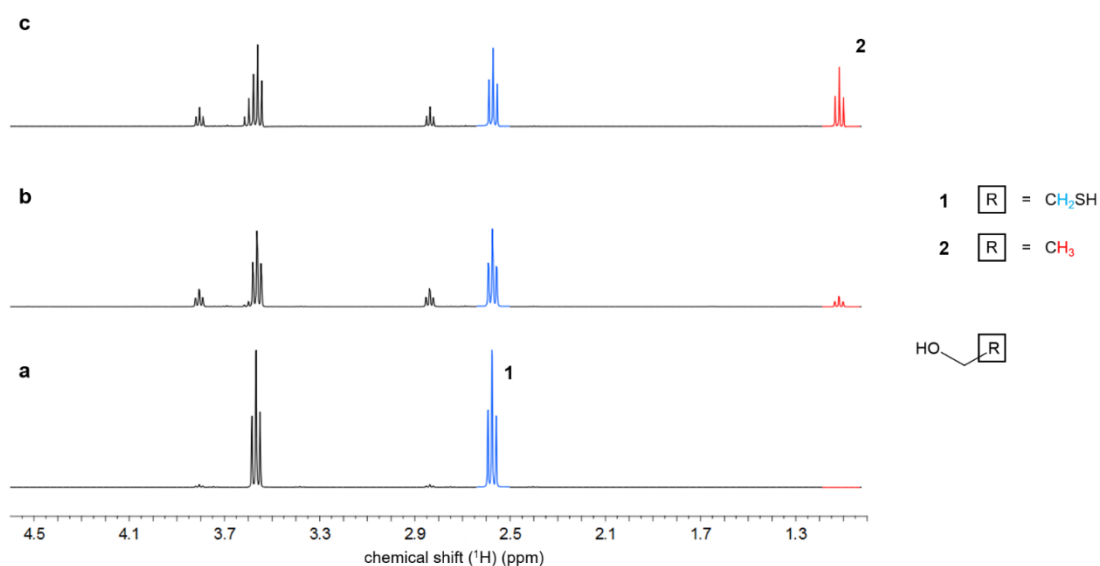
Figure A1.19 | a, ^{1}H-NMR of 2-mercaptoethanol solution before UV light (254 nm) irradiation. b, ^{1}H-NMR of 2-mercaptoethanol solution after 15 min UV light irradiation. c, ^{1}H-NMR of irradiated 2-mercaptoethanol solution after spiking with ethanol.
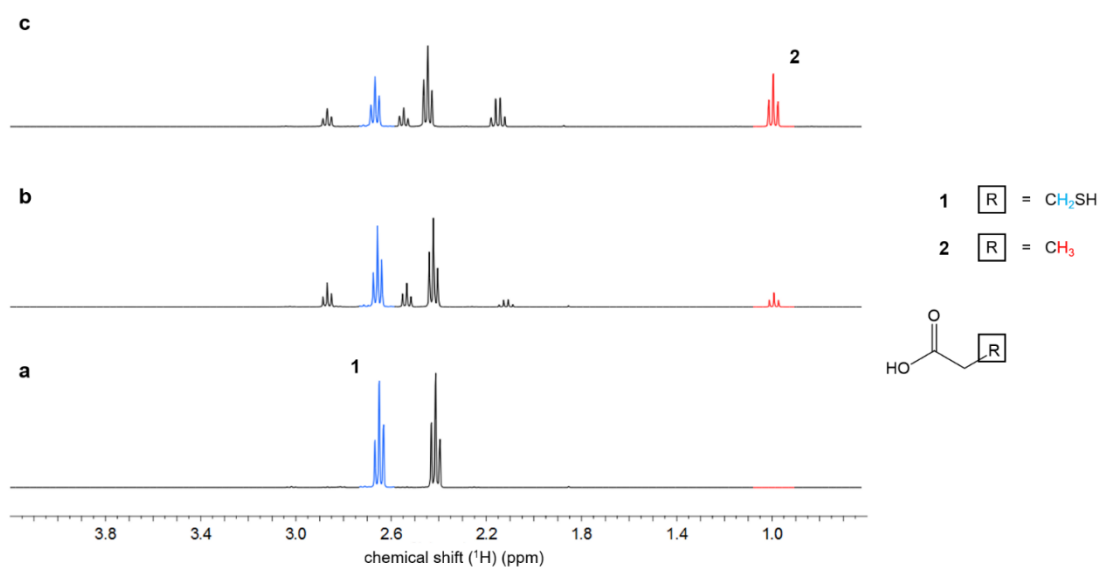
Figure A1.20 | a, ^{1}H-NMR of 3-mercaptopropionic acid solution before UV light (254 nm) irradiation. b, ^{1}H-NMR of a 3-mercaptopropionic acid solution after 15 min UV light irradiation. c, ^{1}H-NMR of an irradiated 3-mercaptopropionic acid solution after spiking with propionic acid.
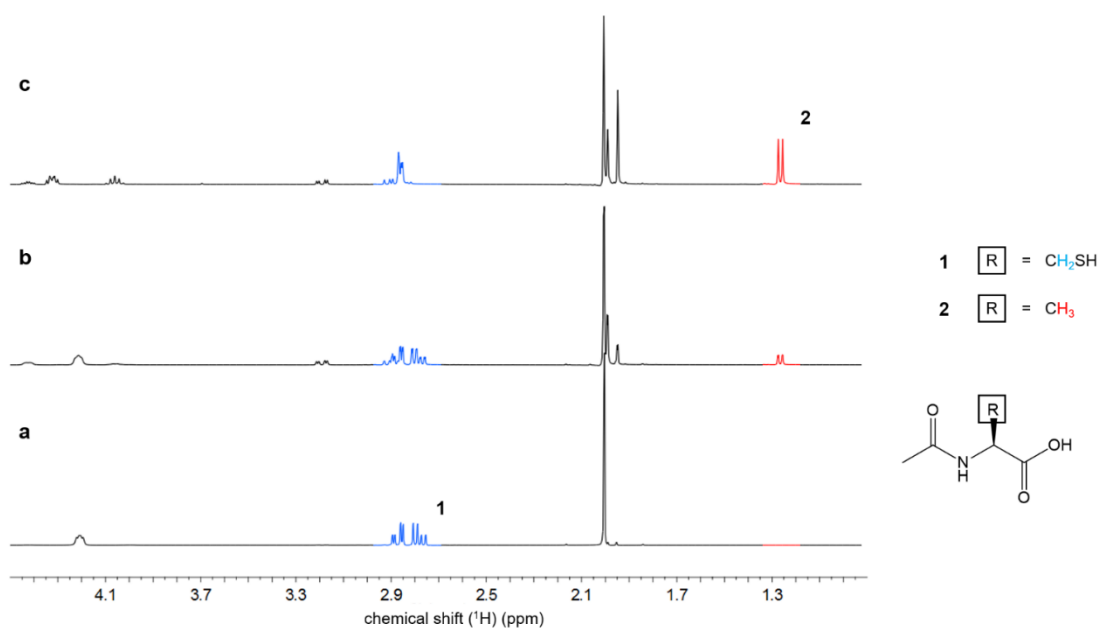
Figure A1.21 | a, ^{1}H-NMR of a *N*-acetyl L-cysteine solution before UV light (254 nm) irradiation. b, ^{1}H-NMR of *N*-acetyl L-cysteine solution after 15 min UV light irradiation. c, ^{1}H-NMR of an irradiated *N*-acetyl L-cysteine solution after spiking with *N*-acetyl alanine.
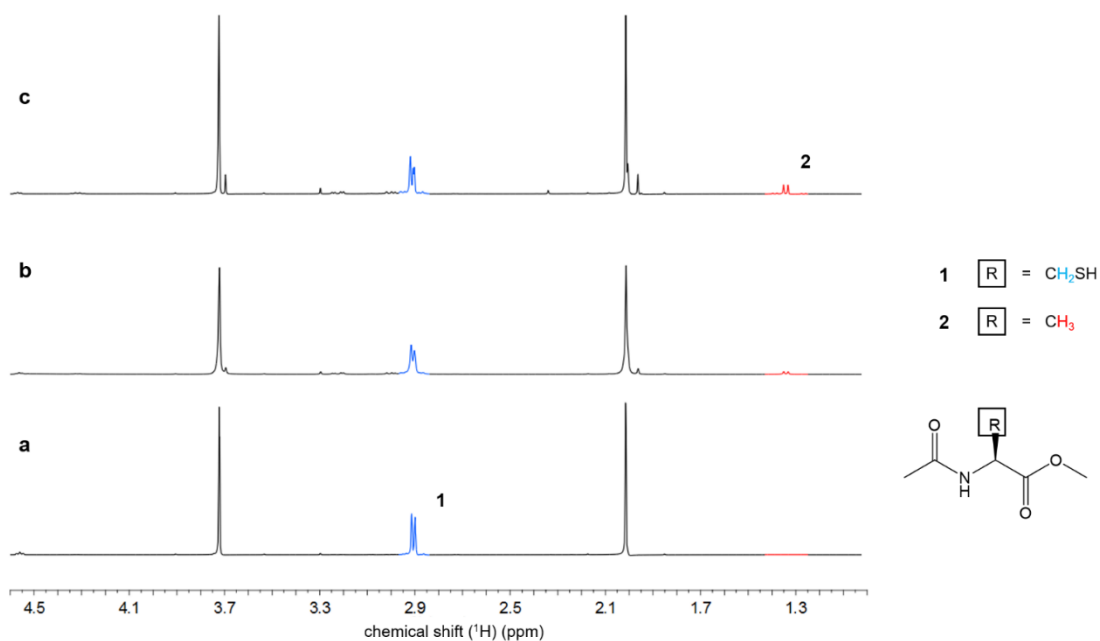
Figure A1.22 | a, ¹H-NMR of *N*-acetyl L-cysteine methyl ester solution before UV light (254 nm) irradiation. b, ¹H-NMR of *N*-acetyl L-cysteine methyl ester solution after 15 min UV light irradiation. c, ¹H-NMR of irradiated *N*-acetyl L-cysteine methyl ester solution after spiking with *N*-acetyl alanine methyl ester.
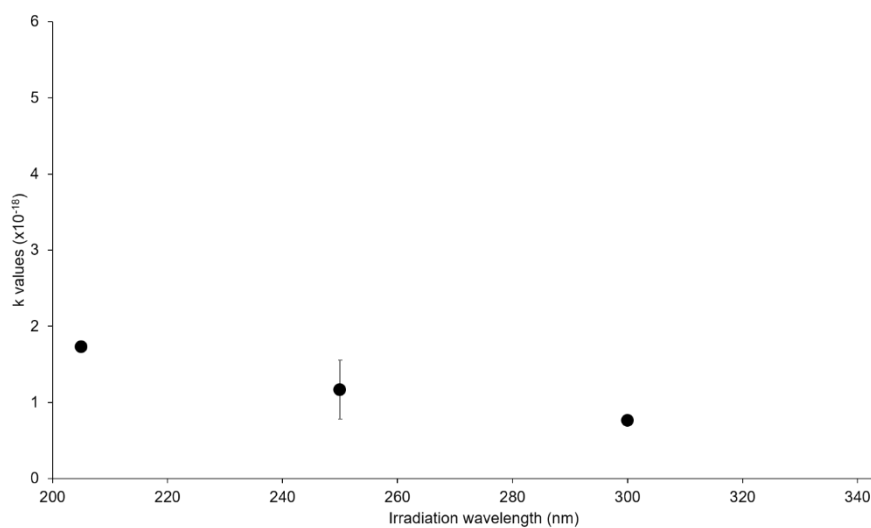
Figure A1.23 | Dependence of the conversion rate constant on the irradiation wavelength for 2-mercaptoethanol. The rate constant (expressed as s^{-1}) was normalized by the photon flux (expressed as s) and each wavelength measurement had a bandwidth of 10 nm. Data represent $n \geq 3$ (mean and SEM).
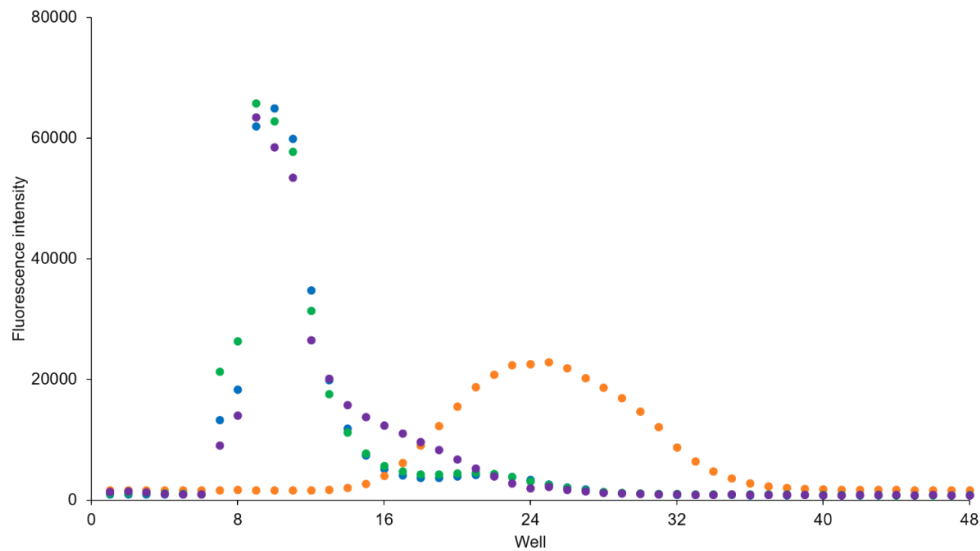
Figure A1.24 | Oleate vesicle stability in the presence of Fe^{2+} and Fe^{2+}- L-glutathione. Oleate vesicles containing encapsulated FITC-dextran were incubated with either 200 mM Tris-HCl, pH 8.5 (blue dots), 200 mM L-glutathione (green dots), 5 mM Fe^{2+} (orange dots), or 5 mM Fe^{2+} plus 200 mM L-glutathione (violet dots) for 60 min. The integrity of the vesicles was then assessed by quantifying unencapsulated, leaked FITC-dextran by size exclusion chromatography with Sepharose 4B followed by fluorescence spectroscopy.
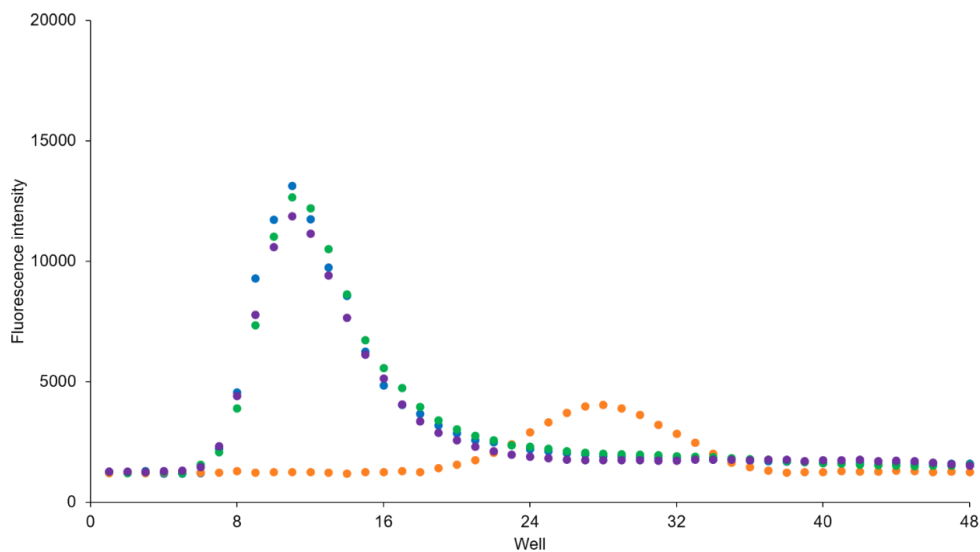
Figure A1.25 | 4:1:1 decanoic acid:decanol:monocaprin vesicle stability to Fe^{2+} and Fe^{2+}-L-glutathione. 4:1:1 decanoic acid:decanol:monocaprin vesicles containing encapsulated FITC-dextran were incubated with either 200 mM Tris-HCl, pH 8.5 (blue dots), 200 mM L-glutathione (green dots), 5 mM Fe^{2+} (orange dots), or 5 mM Fe^{2+} plus 5 mM L-glutathione (violet dots) for 60 min. The integrity of the vesicles was then assessed by quantifying unencapsulated, leaked FITC-dextran by size exclusion chromatography with Sepharose 4B followed by fluorescence spectroscopy.
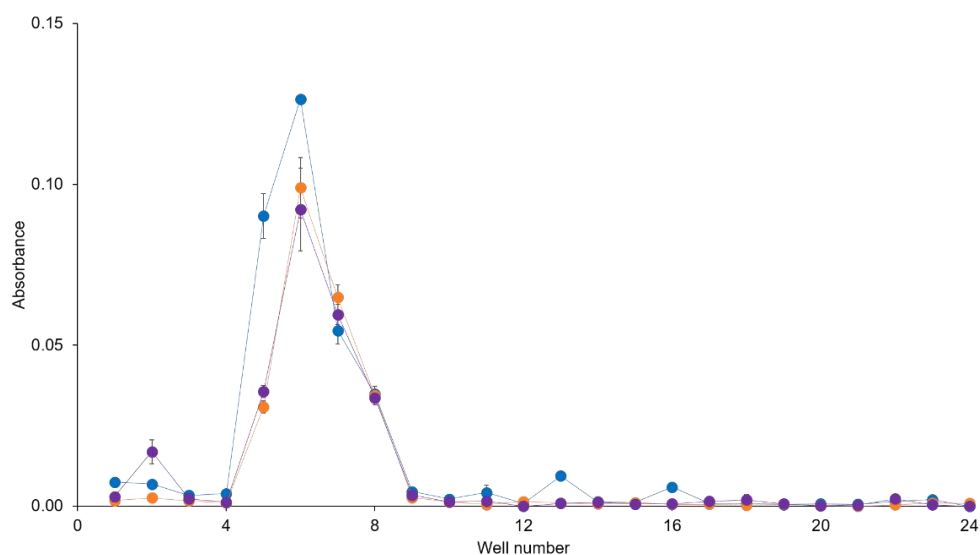
Figure A1.26 | Oleate vesicle stability assay to UV light. Oleate vesicles containing encapsulated Fe^{2+}-L-glutathione was irradiated at 254 nm for 60 min (orange dots). Control reactions included an aliquot of the same sample prior to exposure to UV light (blue dots) and vesicles kept at room temperature for 60 min (violet dots) without exposure to UV. Absorbance of the tiron-iron complex was exploited to detect leaked iron ions, thus indicating structurally compromised vesicles. No leaked iron ions were detected.
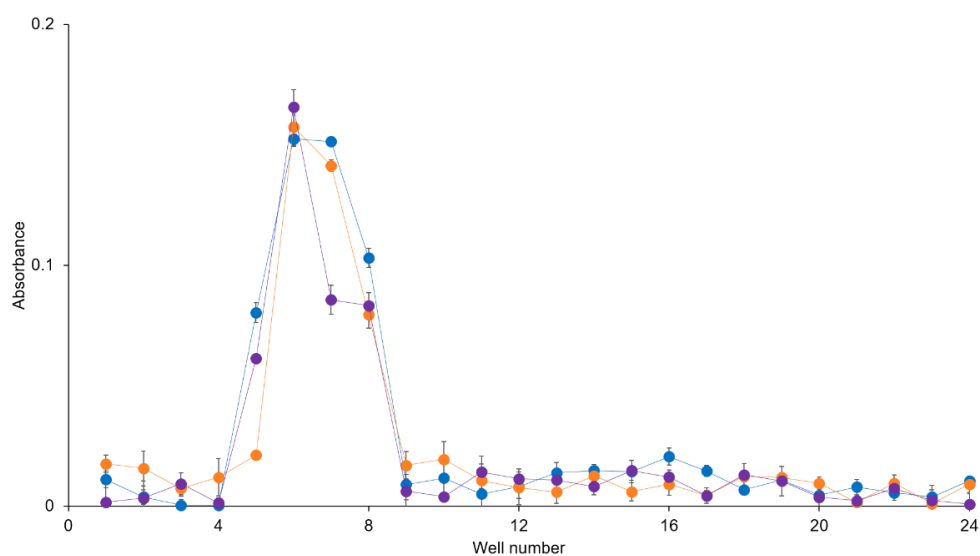
Figure A1.27 | 4:1:1 decanoic acid:decanol:monocaprin vesicle stability to UV light. 4:1:1 decanoic acid:decanol:monocaprin vesicles containing encapsulated Fe^{2+}-L-glutathione was irradiated at 254 nm for 60 min (orange dots). Control reactions included an aliquot of the same sample prior to exposure to UV light (blue dots) and vesicles kept at room temperature for 60 min (violet dots) without exposure to UV. Absorbance of the tiron-iron complex was exploited to detect leaked iron ions, thus indicating structurally compromised vesicles. No leaked iron ions were detected.
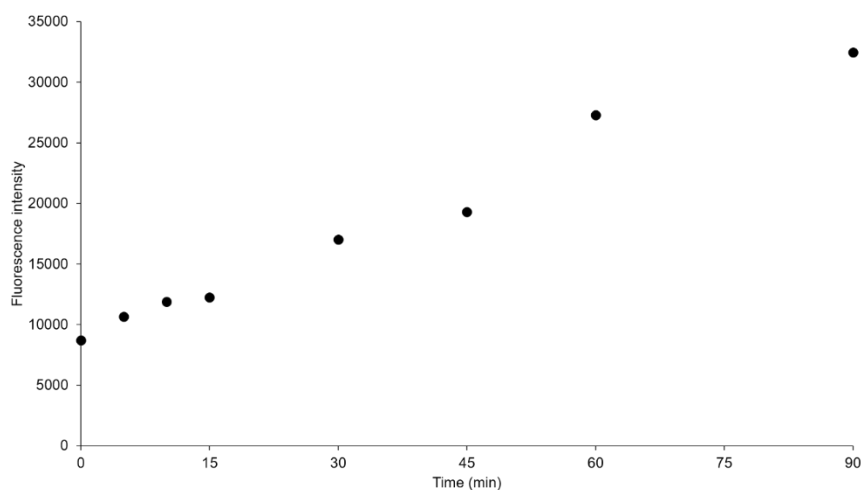
Figure A1.28 | 7-amino-4-methyl coumarin assay for sulfide production within oleate vesicles induced by UV-light. Aliquots of oleate vesicles were irradiated for 5, 10, 15, 30, 45, 60 and 90 min. 7-azido-4-methyl coumarin was added to each aliquot and left to react for 1 h. The formation of 7-amino-4-methyl coumarin upon reaction with sulfide ions was detected by fluorescence (excitation wavelength: 345 nm, emission wavelength: 440 nm).
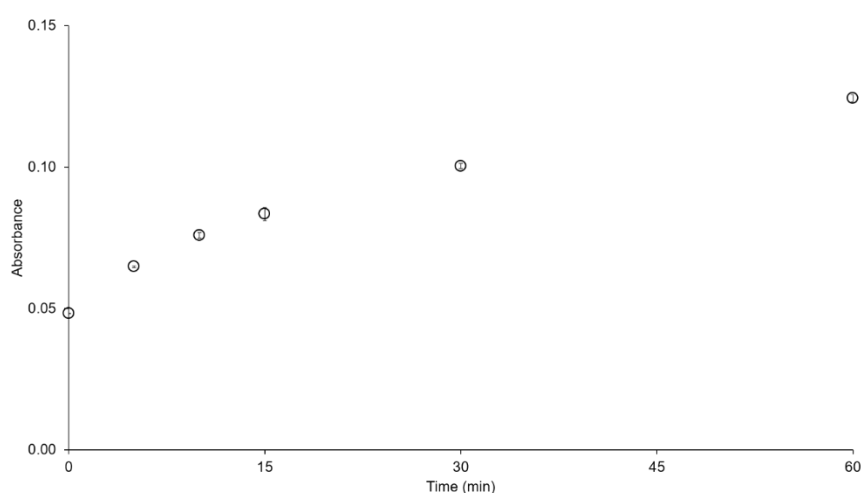
Figure A1.29 | Methylene blue assay for sulfide production within 4:1:1 decanoic acid:decanol:monocaprin vesicles induced by UV-light. Aliquots of 4:1:1 decanoic acid:decanol:monocaprin vesicles were irradiated for 5, 10, 15, 30, 45 and 60 min. DMPD solution (2.5 mM) was added to each aliquot upon the addition of Fe^{3+} ions (1.25 mM) and left to react for 1 h. The absorbance at 668 nm was then recorded. Colorimetric data represent mean and SEM in n ≥ 3 replicates.

Appendix 2

Supplementary Materials for Chapter 4

Table A2.1 | Chemical structures of studied *O*-coordinating and *S*-coordinating ligands.
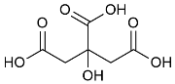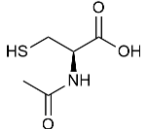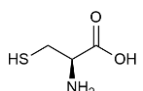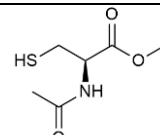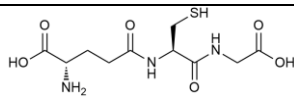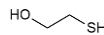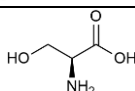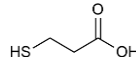
Ligand	Structure	Ligand	Structure
Citric acid		**N-acetyl L-cysteine**	
L-cysteine		**N-acetyl L-cysteine methyl ester**	
L-glutathione		**2-mercaptoethanol**	
L-serine		**3-mercaptopropionic acid**	

Table A2.2 | Formal electrode potential and cathodic peak potential values (expressed in V, vs NHE) for studied Fe^{3+} mononuclear complexes and Fe-S clusters.

Ligand	pH	Fe-S species	Methodology	E°' (V)	E$_{cp}$ (V)
2-mercaptoethanol	9.6	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, GC electrode)	-	-0.189
L-serine	3.5	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, GC electrode)	-	+0.085
L-cysteine	8.4	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, GC electrode)	-	n.d.
N-acetyl L-cysteine methyl ester	8.4	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, GC electrode)	-0.044	-0.114
N-acetyl L-cysteine methyl ester	8.4	[2Fe-2S] cluster	Cyclic voltammetry (50 mV/s, modified Au electrode)	-	-0.595
N-acetyl L-cysteine methyl ester	8.4	[4Fe-4S] cluster	Cyclic voltammetry (50 mV/s, modified Au electrode)	-	-0.635*
L-glutathione	3.5	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, GC electrode)	-	+0.371

L-glutathione	8.6	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, GC electrode)	-	-0.080
L-glutathione	8.6	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, modified Au electrode)	-	-0.092
L-glutathione	8.6	Fe^{3+} mononuclear complex	Spectroelectrochemistry (2 mV/s, Pt electrode)	-	-0.126
L-glutathione	8.6	Fe^{3+} mononuclear complex	Dye photoreduction method	-0.025 ± 0.008[#]	-
L-glutathione	8.6	[2Fe-2S] cluster	Cyclic voltammetry (50 mV/s, modified Au electrode)	-	-0.490
L-glutathione	8.6	[2Fe-2S] cluster	Dye photoreduction method	-0.311 ± 0.011[#]	-
L-glutathione	8.6	[4Fe-4S] cluster	Cyclic voltammetry (50 mV/s, modified Au electrode)	-	-0.571
L-glutathione	8.6	[4Fe-4S] cluster	Dye photoreduction method	-0.381 ± 0.007[#]	-
(GlyCysGly)₂	8.0	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, GC electrode)	-	-0.089
(GlyCysGly)₄	8.0	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, GC electrode)	-	-0.035
(GluCysGly)₂	8.6	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, GC electrode)	-	-0.109
(GluCysGly)₄	8.6	Fe^{3+} mononuclear complex	Cyclic voltammetry (50 mV/s, GC electrode)	-	-0.047

* The cathodic peak of the [4Fe-4S] cluster and the cathodic peak of sulfide ions reduction are partially overlapped.

[#] Presented data obtained represent mean and SEM, n ≥ 2 replicates.

Table A2.3 | Overall charge values for model Cys-containing tripeptides and related Fe^{3+} mononuclear complexes.
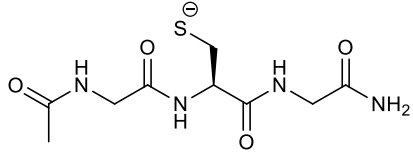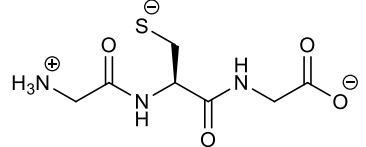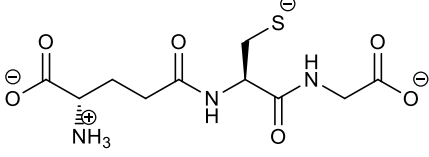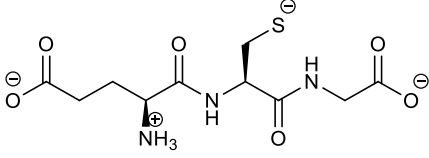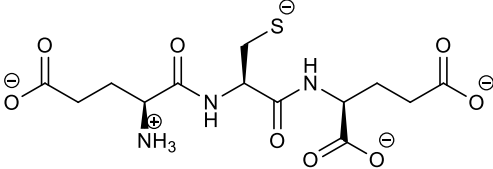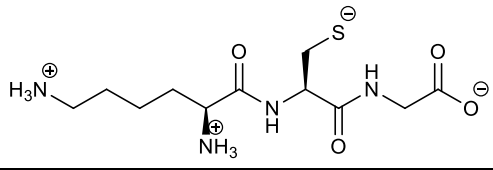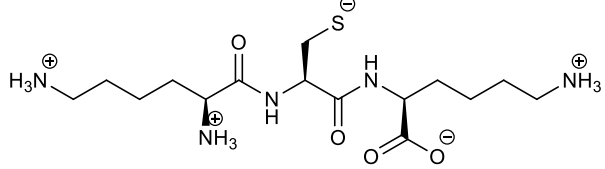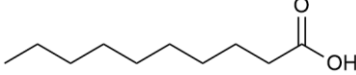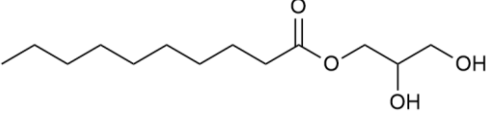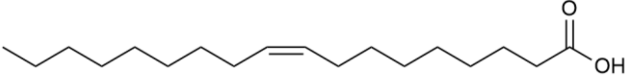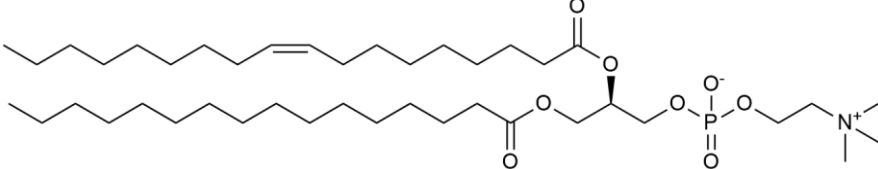
Peptide sequence	Peptide structure (pH = $pK_{a(thiol)}$)	Peptide charge	$[Fe^{3+}(peptide)_4]$ charge
N-acetyl GlyCysGly amide		-1	-1
GlyCysGly		-1	-1
γGluCysGly		-2	-5
GluCysGly		-2	-5
GluCysGlu		-3	-9
LysCysGly		0	+3
LysCysLys		+1	+7

Table A2.4 | Chemical structures of studied lipid systems.

Lipid name	Lipid structure
Decanoic acid	
Decanol	
Glycerol monodecanoate (monocaprin)	
cis-9-octadecenoic acid (oleic acid)	
1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC)	
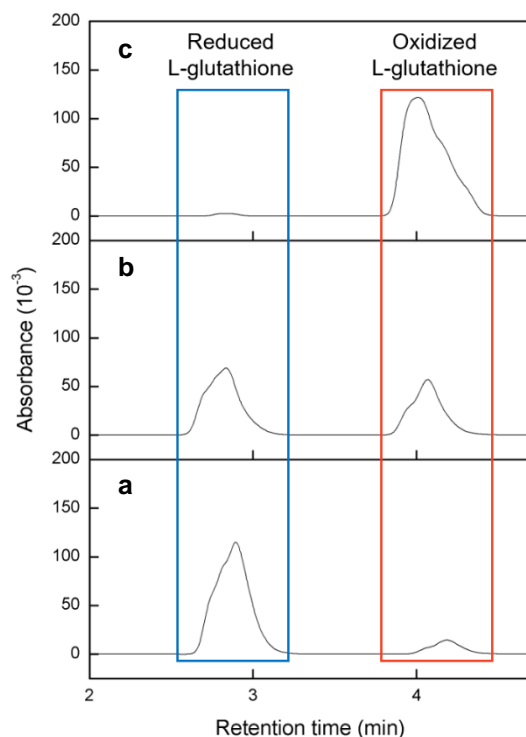
Figure A2.1 | Effect of mononuclear Fe(III) complex on free thiolates studied by HPLC analysis (gradient run: 3-18% eluent B, eluent A: water + 0.1% TFA, eluent B: acetonitrile + 0.1% TFA, flow rate: 1 mL/min, detection wavelength: 215 nm). Solution of mononuclear complex (0.5 mM Fe^{3+}, 40 mM reduced L-glutathione) at 0 min (a), 90 min (b), 360 min (c), showing the complete conversion of reduced L-glutathione to oxidized L-glutathione after 6 hours. The peak centred at 2.88 min is due to reduced L-glutathione, the peak centred at 4.24 min is due to oxidized L-glutathione.
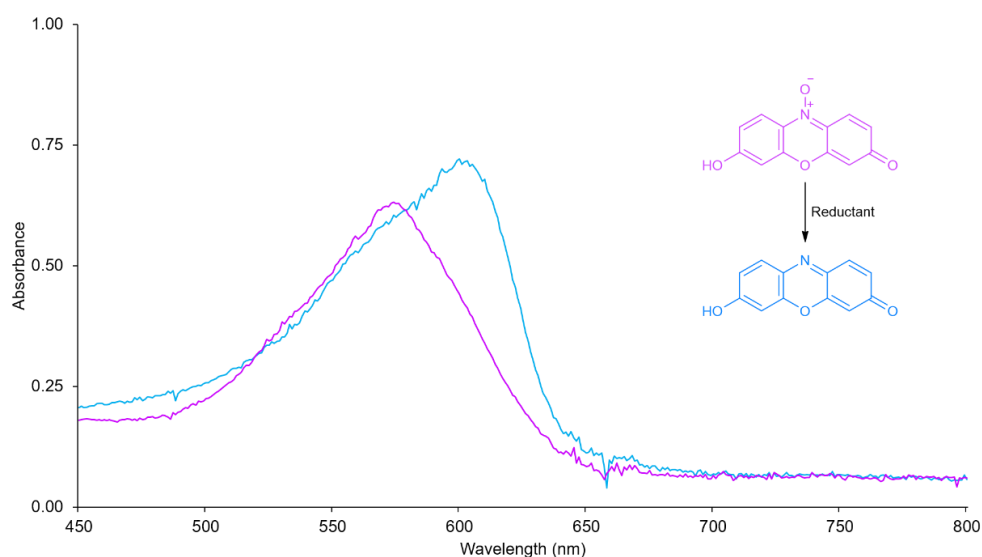
Figure A2.2 | Redox activity performed by the mononuclear ferric complex on free thiolates in the presence of a redox indicator (resazurin). UV-vis. spectra of a solution of the mononuclear ferric complex (0.5 mM Fe^{3+}, 40 mM reduced L-glutathione) after 30 min and 180 min from the addition of the redox indicator resazurin (0.1 mM). The ferrous mononuclear complex formed by oxidation of free thiolates is converted back to the oxidized species while resazurin (λ_{max} 605 nm) is reduced to resorufin (λ_{max} 575 nm).

127
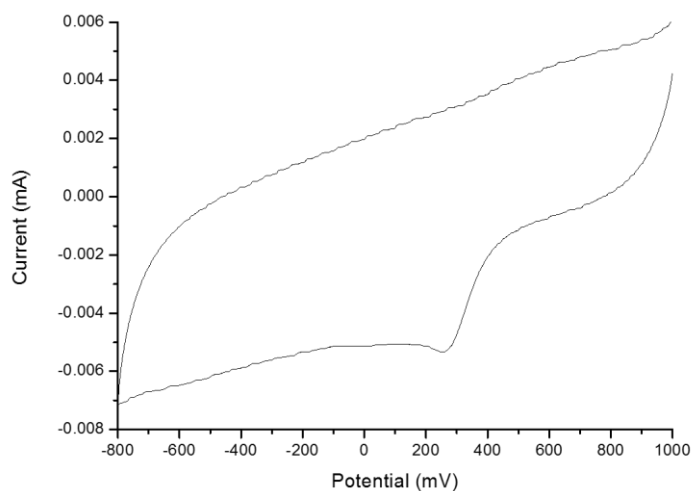
Figure A2.3 | Cyclic voltammetry data on aqueous iron ions. The experiments have been performed by using Ag/AgCl as reference electrode, platinum as counter electrode and glassy-carbon electrode on an acidic solution of Fe(III) (0.5 mM) in water (scan rate: 50 mV/s, E_{cp} = 0.263 V vs Ag/AgCl).
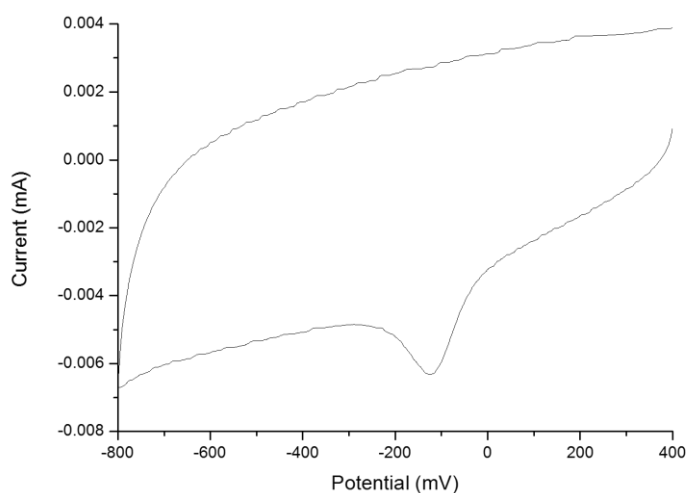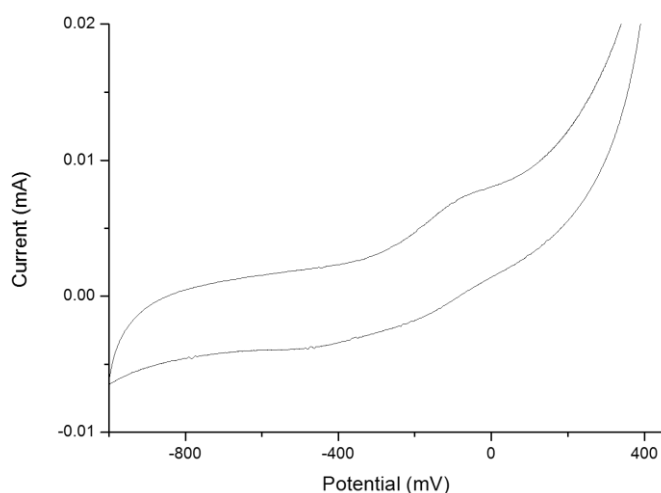

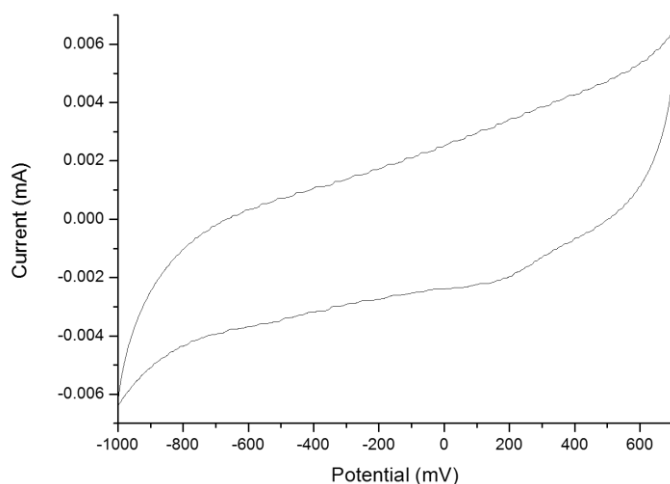**Figure A2.4 |** Cyclic voltammetry data on aqueous iron ions in the presence of L-serine. The experiments have been performed by using Ag/AgCl as reference electrode, platinum as counter electrode and glassy-carbon electrode on an acidic solution of L-serine (40 mM) and Fe(III) (0.5 mM) in water (scan rate: 50 mV/s, E_{cp} = -0.120 V vs Ag/AgCl).
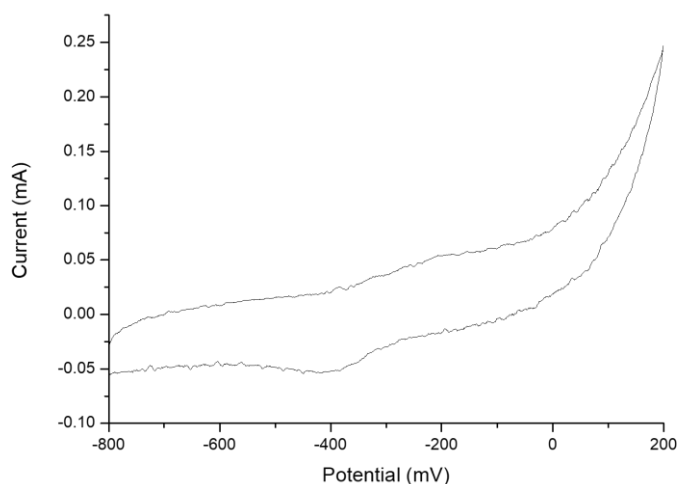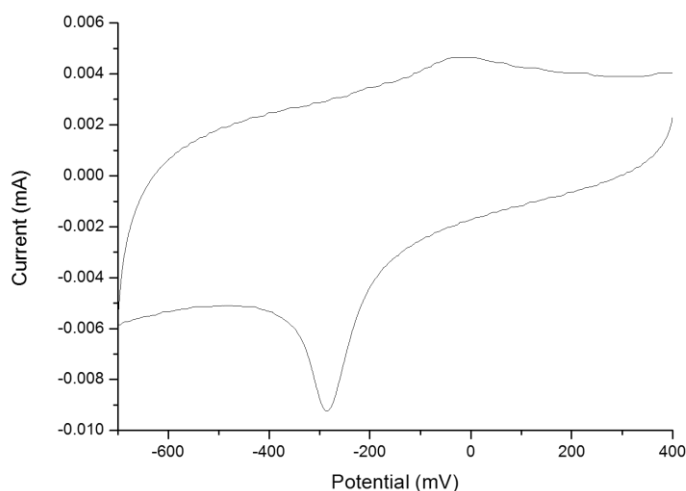
Figure A2.5 | Cyclic voltammetry data on aqueous iron ions in the presence of 2-mercaptoethanol. The experiments have been performed by using Ag/AgCl as reference electrode, platinum as counter electrode and glassy-carbon electrode on a solution of 2-mercaptoethanol (40 mM, pH 9.2) and Fe(III) (0.5 mM) in water (scan rate: 50 mV/s, E_{cp} = -0.394 V vs Ag/AgCl).
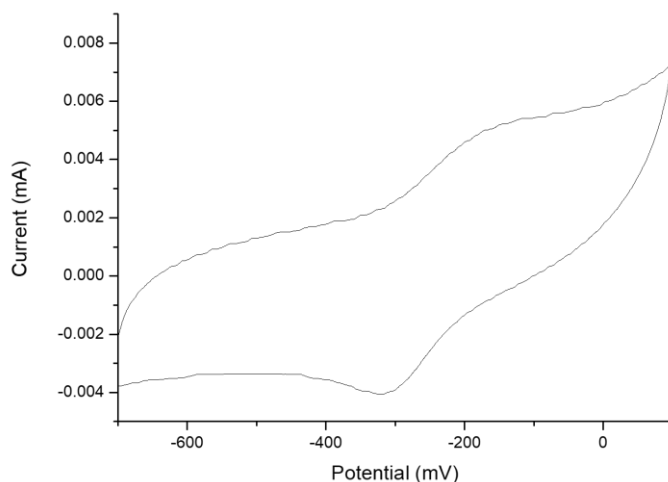

Figure A2.6 | Cyclic voltammetry data on aqueous iron ions in the presence of *N*-acetyl L-cysteine methyl ester. The experiments have been performed by using Ag/AgCl as reference electrode, platinum as counter electrode and glassy-carbon electrode on a solution of *N*-acetyl L-cysteine methyl ester (40 mM, pH 8.4) and Fe(III) (0.5 mM) in water (scan rate: 50 mV/s, $E^{\circ'}$ = -0.249 V vs Ag/AgCl).

Figure A2.7 | Cyclic voltammetry data on aqueous iron ions in the presence of L-cysteine. The experiments have been performed by using Ag/AgCl as reference electrode, platinum as counter electrode and glassy-carbon electrode on a solution of L-cysteine (40 mM, pH 8.4) and Fe(III) (0.5 mM) in water (scan rate: 50 mV/s, E_{ap} = 0.123 V vs Ag/AgCl).
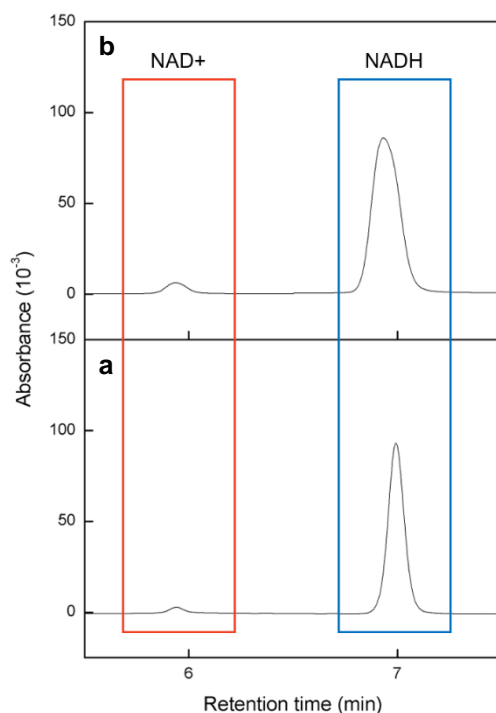
Figure A2.8 | Effect of Fe(III) ions coordination by L-cysteine on NADH studied by HPLC analysis (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm). a, Solution of NADH (0.5 mM) and L-cysteine-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM L-cysteine, pH 8.4) immediately after mixing. b, Solution of NADH (0.5 mM) and L-cysteine-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM L-cysteine, pH 8.4) after 6 hours from mixing. The peak centred at 6.99 min is due to NADH, the peak centred at 5.92 min is due to NAD+.

Figure A2.9 | Cyclic voltammetry data on aqueous iron ions in the presence of reduced L-glutathione. The experiments have been performed by using Ag/AgCl as reference electrode, platinum as counter electrode and glassy-carbon electrode on a solution of reduced L-glutathione (40 mM, pH 3.5) and Fe(III) (0.5 mM) in water (scan rate: 50 mV/s, E_{ap} = 0.166 V vs Ag/AgCl).

Figure A2.10 | Cyclic voltammetry data on aqueous iron ions in the presence of reduced L-glutathione. The experiments have been performed by using Ag/AgCl as reference electrode, platinum as counter electrode and glassy-carbon electrode on a solution of reduced L-glutathione (40 mM, pH 8.6) and Fe(III) (0.5 mM) in water (scan rate: 50 mV/s, E_{ap} = -0.285 V vs Ag/AgCl).
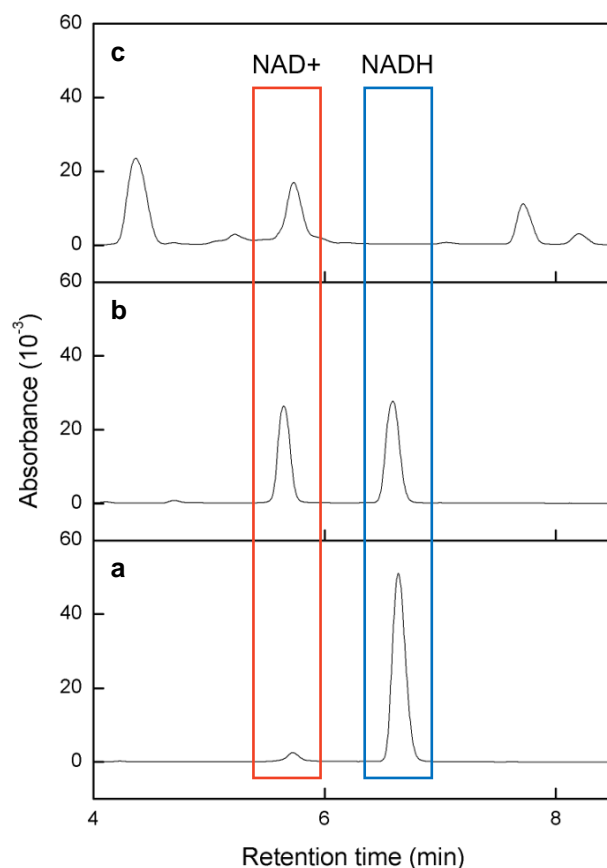
Figure A2.11 | Effect of Fe(III) ions *O*- and *S*-coordination by reduced L-glutathione on NADH studied by HPLC analysis (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm). a, Solution of NADH (0.1 mM) in neutral pH after 6 hours. b, Solution of NADH (0.5 mM) and L-glutathione-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM L-glutathione, pH 8.6) after 6 hours from mixing. c, Solution of NADH (0.5 mM) and L-glutathione-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM L-glutathione, pH 3.5) after 6 hours from mixing. The incubation of NADH and the L-glutathione ferric complex in acidic conditions induced the degradation of the substrate. The peak centred at 6.69 min is due to NADH, the peak centred at 5.72 min is due to NAD+.
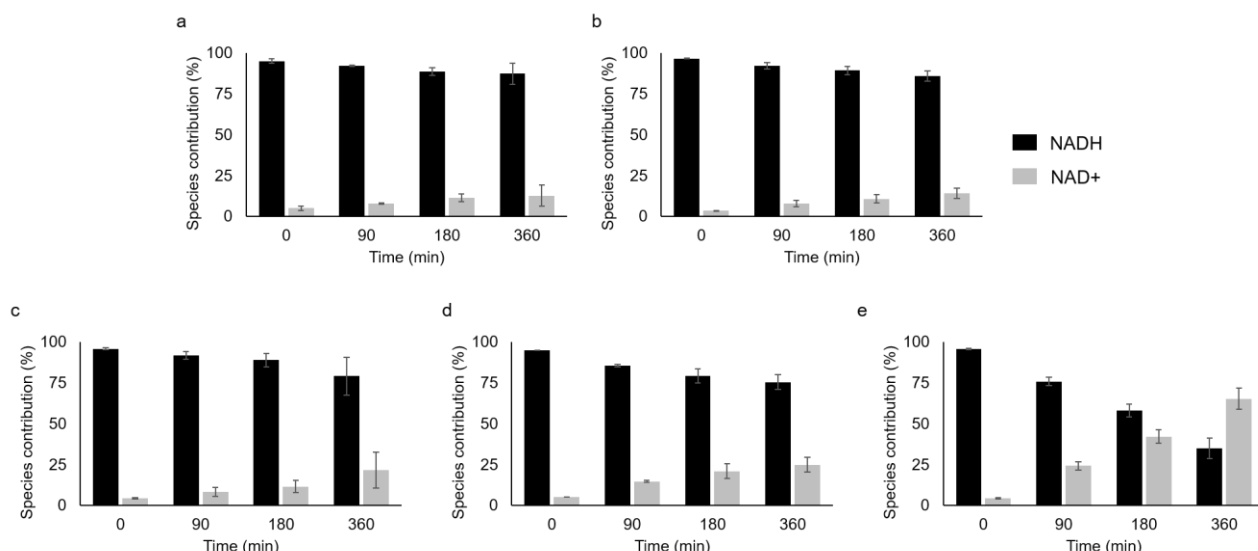
Figure A2.12 | Conversion efficiency for different mononuclear complexes based on HPLC data (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm). a, Solution of Fe(III) (0.5 mM) with 2-mercaptoethanol (40 mM, pH 9.2) and NADH (0.5 mM). b, Solution of Fe(III) (0.5 mM) with *N*-acetyl cysteine methyl ester (40 mM, pH 8.4) and NADH (0.5 mM). c, Solution of Fe(III) (0.5 mM) with 3-mercaptopropionic acid (40 mM, pH 10.3) and NADH (0.5 mM). d, Solution of Fe(III) (0.5 mM) with *N*-acetyl cysteine (40 mM, pH 9.6) and NADH (0.5 mM). e, Solution of Fe(III) (0.5 mM) with L-glutathione (40 mM, pH 8.6) and NADH (0.5 mM). Baseline correction and integration analysis performed with Origin Pro 8. Presented data represent mean and SEM, n ≥ 2 replicates. Black bars represent NADH concentration calculated from absorbance values, grey bars represent NAD+ concentration calculated from absorbance values, both corrected by the extinction coefficients of NADH and NAD+, respectively.
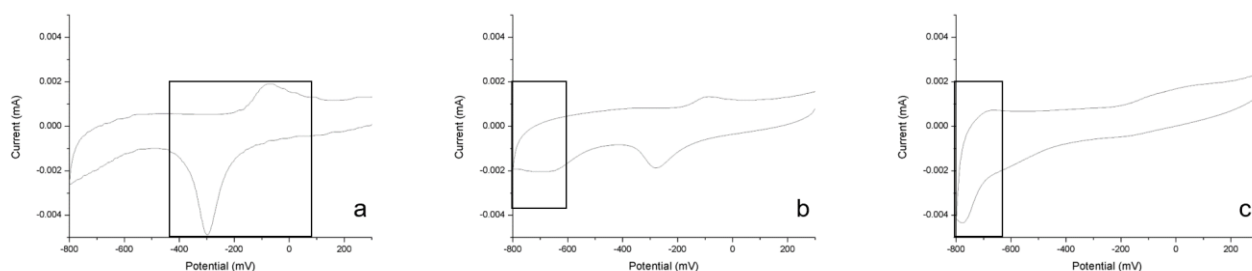
Figure A2.13 | Cyclic voltammetry data on different iron-sulfur systems. The experiments have been performed by using Ag/AgCl as reference electrode, platinum as counter electrode and a 4,4'-bipyridyl-modified gold electrode. A black rectangle identifies the major signal, due to the predominant species in each voltammogram. a, Solution of Fe(III) (0.5 mM) with L-glutathione (40 mM, pH 8.6) (scan rate: 100 mV/s, E_{cp} = -0.298 V vs Ag/AgCl). b, Solution of sulfide (0.5 mM) and Fe(III) (0.5 mM) with L-glutathione (40 mM, pH 8.6) (scan rate: 50 mV/s, E_{cp} = -0.695 V vs Ag/AgCl,). c, Solution of sulfide (5 mM) and Fe(III) (0.5 mM) with L-glutathione (40 mM, pH 8.6) (scan rate: 50 mV/s, E_{cp} = -0.776 V vs Ag/AgCl).
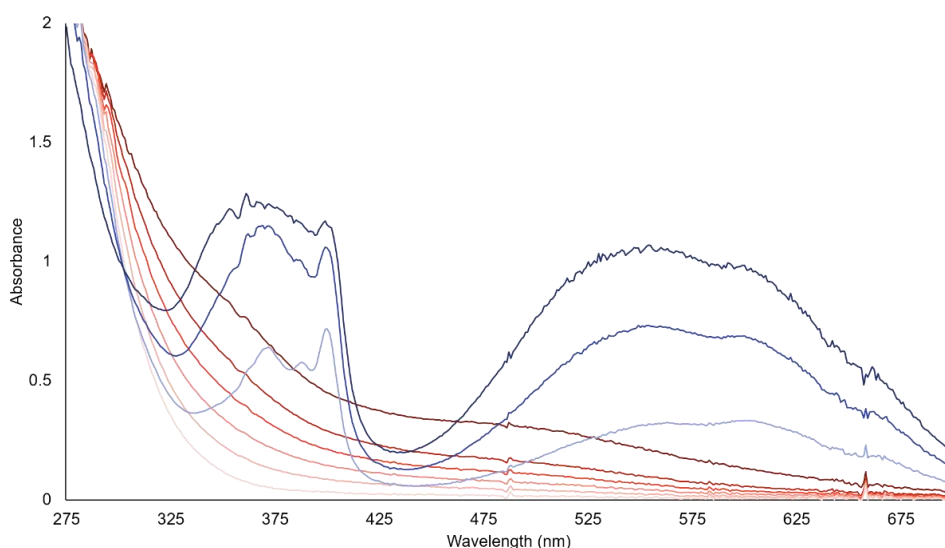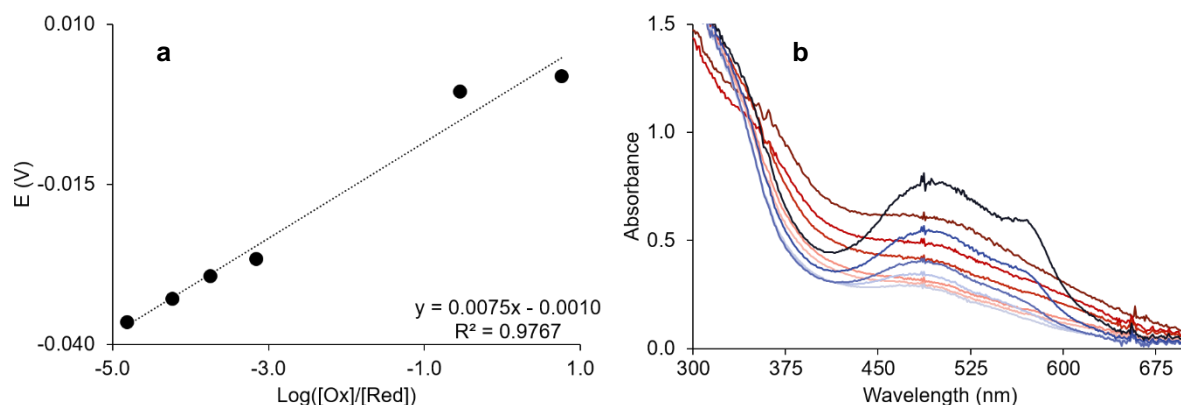
Figure A2.14 | Redox titration studies on mononuclear L-glutathione-ferric complex in the presence of a redox indicator (benzyl viologen). UV-vis. spectra of a solution of the mononuclear ferric complex (0.5 mM Fe^{3+}, 40 mM reduced L-glutathione) titrated with the reducing agent sodium dithionite (0.1 umoles addition) in the presence of benzyl viologen (0.05 mM). Since the redox potential of benzyl viologen is lower than the redox potential of the iron complex, the ferric mononuclear complex (λ_{max} ~500 nm) is reduced by sodium dithionite, without influencing the redox state of the dye. In the presence of an excess of the reducing agent benzyl viologen undergoes reduction (λ_{max} 560 nm).
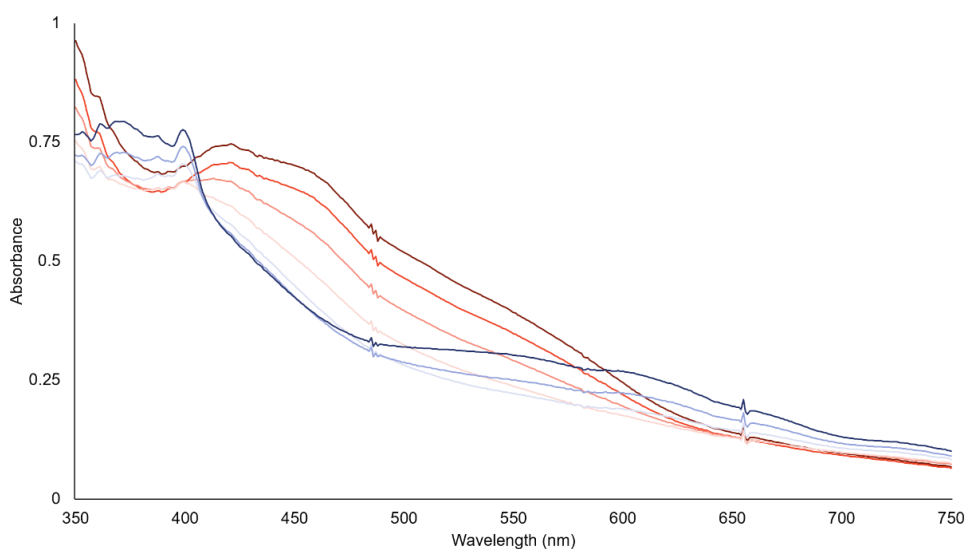
Figure A2.15 | Redox titration studies on L-glutathione [2Fe-2S] cluster in the presence of a redox indicator (benzyl viologen). UV-vis. spectra of a solution of the mononuclear ferric complex (0.5 mM Fe^{3+}, 0.2 mM S^{2-}, 40 mM reduced L-glutathione) titrated with the reducing agent sodium dithionite (0.1 umoles addition) in the presence of benzyl viologen (0.05 mM). The L-glutathione [2Fe-2S] cluster (λ_{max} ~420 nm and ~450 nm) is reduced by sodium dithionite. Since the redox potential of benzyl viologen is close to the redox potential of the [2Fe-2S] cluster, benzyl viologen undergoes simultaneously reduction (λ_{max} 560 nm).
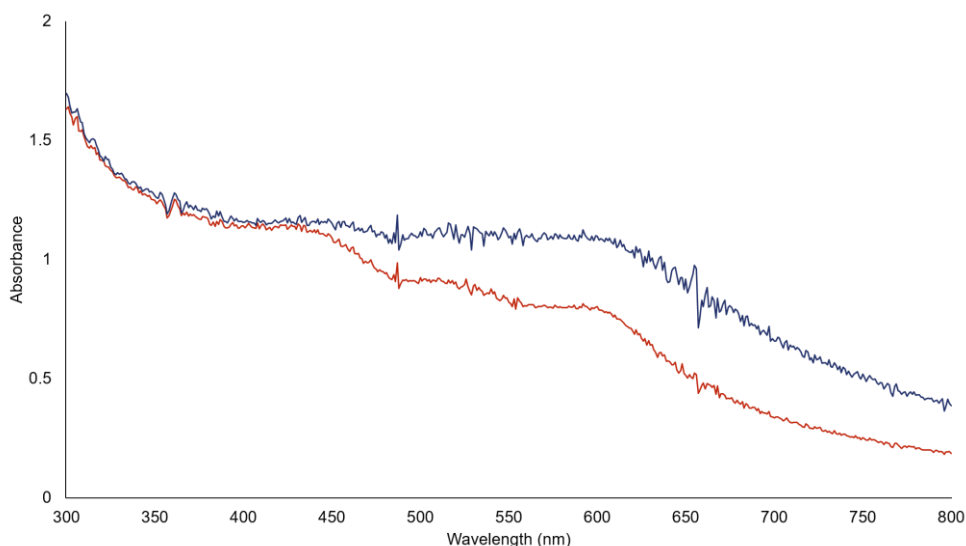
Figure A2.16 | Redox titration studies on L-glutathione [4Fe-4S] cluster in the presence of a redox indicator (benzyl viologen). UV-vis. spectra of a solution of the mononuclear ferric complex (0.5 mM Fe^{3+}, 5 mM S^{2-}, 40 mM reduced L-glutathione) titrated with the reducing agent sodium dithionite (0.1 umoles addition) in the presence of benzyl viologen (0.05 mM). Since the redox potential of benzyl viologen is higher than the redox potential of the [4Fe-4S] cluster, the oxidized dye is reduced by sodium dithionite, without influencing the redox state of L-glutathione [4Fe-4S] cluster (λ_{max} ~415 nm).

Figure A2.17 | Example of redox titration studies on mononuclear L-glutathione-ferric complex in the presence of a redox indicator (2,3,5-triphenyl tetrazolium). The redox potential was determined by the equilibration of the iron system with the redox dye during reduction with sodium dithionite under strict anaerobic conditions. a, Determination of the redox potential for the couple Fe^{3+}/Fe^{2+} mononuclear complex. For the interpolation of the data, the system potentials were calculated by using the Nernst equation from the ratio of the concentrations of the oxidized and reduced forms of the redox dye. b, UV-vis. spectra of a solution of the mononuclear ferric complex (0.5 mM Fe^{3+}, 40 mM reduced L-glutathione) titrated with the reducing agent sodium dithionite (0.1 umoles addition) in the presence of 2,3,5-triphenyl tetrazolium (0.05 mM, λ_{max} 490 nm). Red curves represent the absorbance spectra of the oxidized species mixture (iron-sulfur system and dye), blue curves represent the absorbance spectra of the reduced species mixture. The study has been performed as n ≥ 2 replicates.
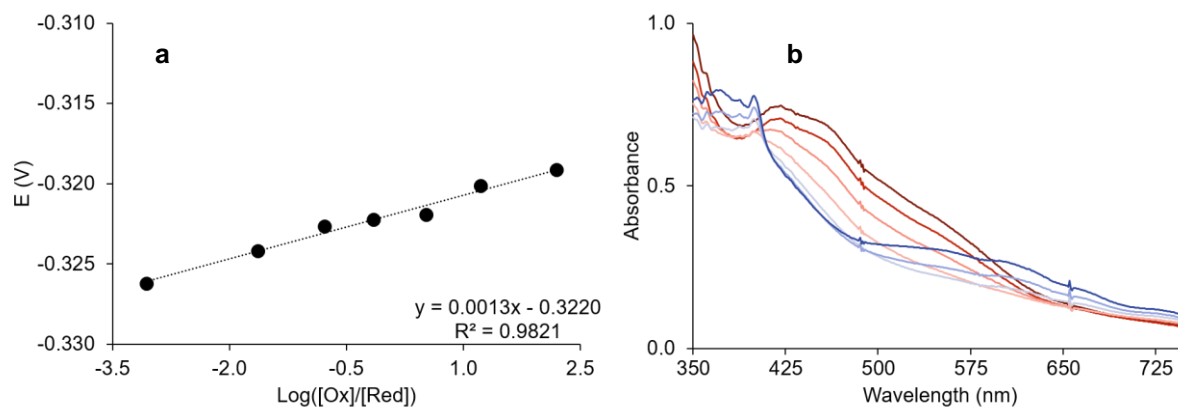
135
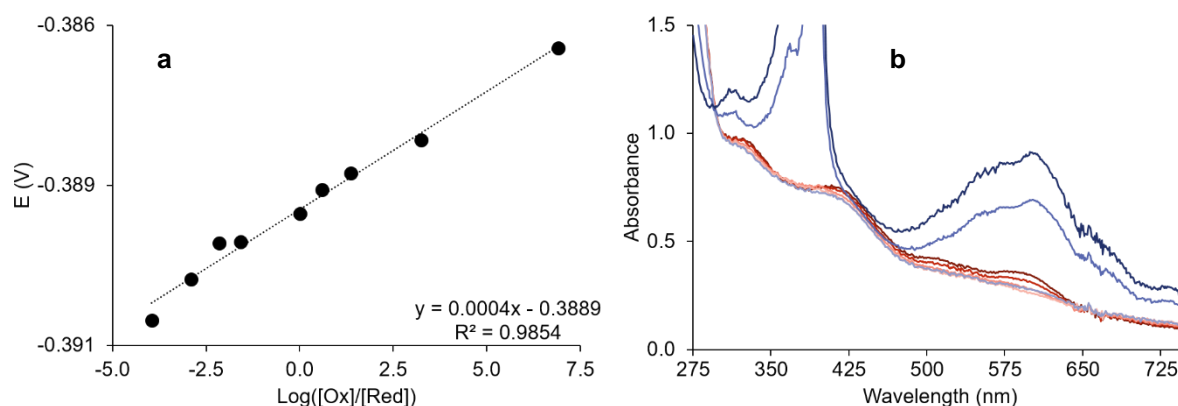
Figure A2.18 | Example of redox titration studies on mononuclear L-glutathione [2Fe-2S] cluster in the presence of a redox indicator (benzyl viologen). The redox potential was determined by the equilibration of the iron-sulfur system with the redox dye during reduction with sodium dithionite under strict anaerobic conditions. a, Determination of the redox potential for the couple $[2Fe-2S]^{2+/1+}$ cluster. For the interpolation of the data, the system potentials were calculated by using the Nernst equation from the ratio of the concentrations of the oxidized and reduced forms of the redox dye. b, UV-vis. spectra of a solution of the $[2Fe-2S]^{2+}$ cluster (0.5 mM Fe^{3+}, 0.2 mM S^{2-}, 40 mM reduced L-glutathione) titrated with the reducing agent sodium dithionite (0.1 umoles addition) in the presence of benzyl viologen (0.05 mM, λ_{max} 560 nm). Red curves represent the absorbance spectra of the oxidized species mixture (iron-sulfur system and dye), blue curves represent the absorbance spectra of the reduced species mixture. The study has been performed as n ≥ 2 replicates.
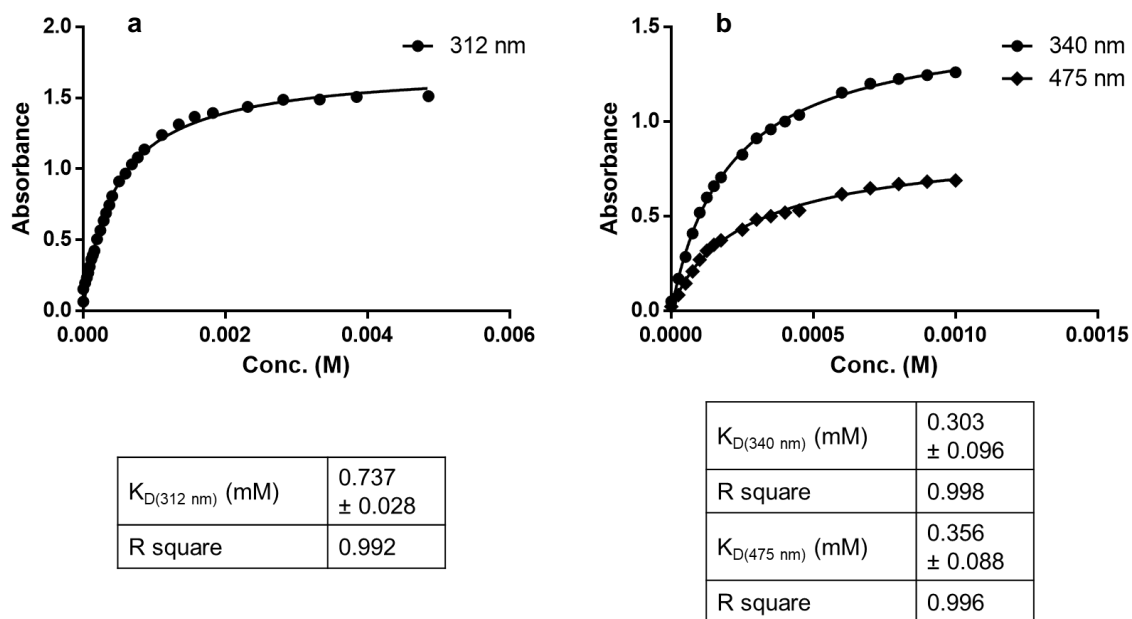

Figure A2.19 | Example of redox titration studies on mononuclear L-glutathione [4Fe-4S] cluster in the presence of a redox indicator (methyl viologen). The redox potential was determined by the equilibration of the iron-sulfur system with the redox dye during reduction with sodium dithionite under strict anaerobic conditions. a, determination of the redox potential for the couple $[4Fe-4S]^{2+/1+}$ cluster. For the interpolation of the data, the system potentials were calculated by using the Nernst equation from the ratio of the concentrations of the oxidized and reduced forms of the redox dye. b, UV-vis. spectra of a solution of the $[4Fe-4S]^{2+}$ cluster (0.5 mM Fe^{3+}, 5 mM S^{2-}, 40 mM reduced L-glutathione) titrated with the reducing agent sodium dithionite (0.1 umoles addition) in the presence of methyl viologen (0.05 mM, λ_{max} 600 nm). Red curves represent the absorbance spectra of the oxidized species mixture (iron-sulfur system and dye), blue curves represent the absorbance spectra of the reduced species mixture. The study has been performed as n ≥ 2 replicates.

$K_{D(312\,nm)}$ (mM)	0.737 ± 0.028
R square	0.992

$K_{D(340\,nm)}$ (mM)	0.303 ± 0.096
R square	0.998
$K_{D(475\,nm)}$ (mM)	0.356 ± 0.088
R square	0.996

Figure A2.20 | Example of dissociation constant (K_D) studies based on Fe(III)/(II) ions titration on reduced L-glutathione solution. Aqueous ferrous (a) or ferric (b) ions aliquots were added to a reduced L-glutathione (40 mM, pH 8.6) solution and formation of the mononuclear complex followed by UV-vis. spectroscopy (λ_{max} 312 nm for Fe^{2+} complex, and 340 or 475 nm for Fe^{3+} complex). Presented results represent mean and SEM, n ≥ 2 replicates.
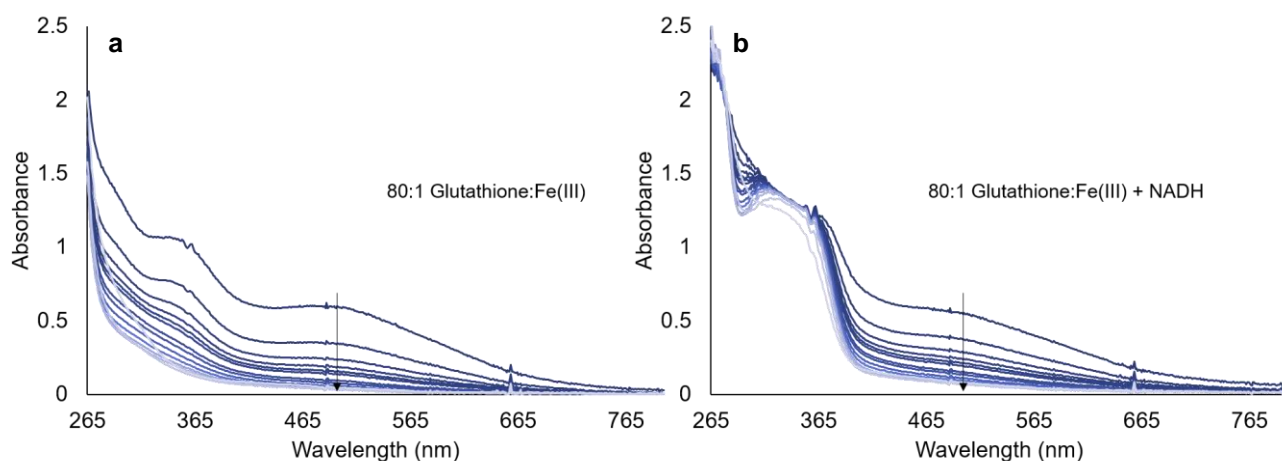
Figure A2.21 | Stability studies for reduced L-glutathione ferric mononuclear complex. UV-vis. spectra of a solution of ferric mononuclear complex (0.5 mM Fe^{3+}, 40 mM reduced L-glutathione) over time (up to 18 hours) in the absence (a) or in the presence (b) of NADH (0.5 mM). The study has been performed as n ≥ 2 replicates.
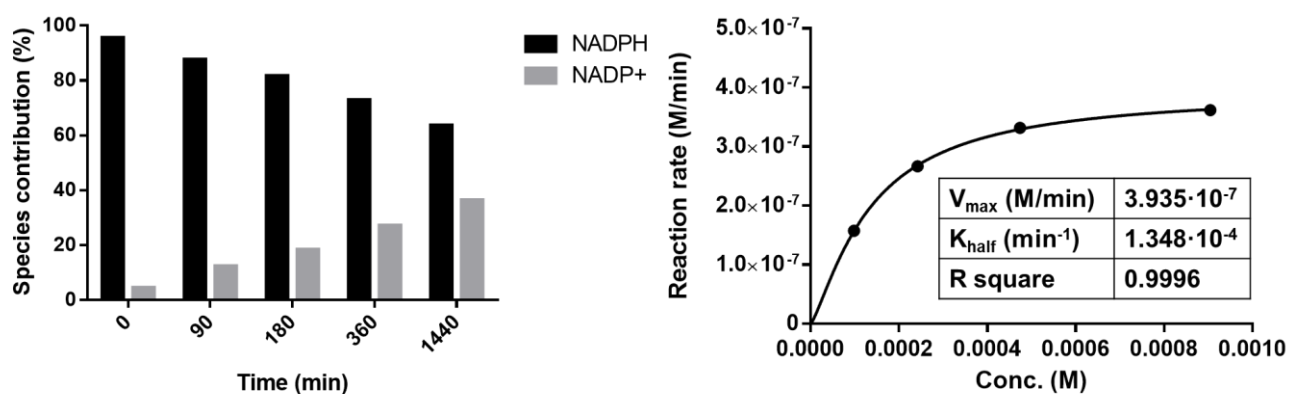
Figure A2.22 | Reaction kinetics for NADPH oxidation conversion in the presence of L-glutathione-ferric mononuclear complex. a, Conversion studies for the equimolar reaction between the mononuclear complex (0.5 mM) and NADPH (0.5 mM). NADPH and NADP+ concentrations are represented as black and grey bars, respectively. b, Reaction rate vs substrate concentration graph with fixed concentration of mononuclear complex (0.5 mM Fe^{3+}, 40 mM reduced L-glutathione) and variable NADPH concentration.
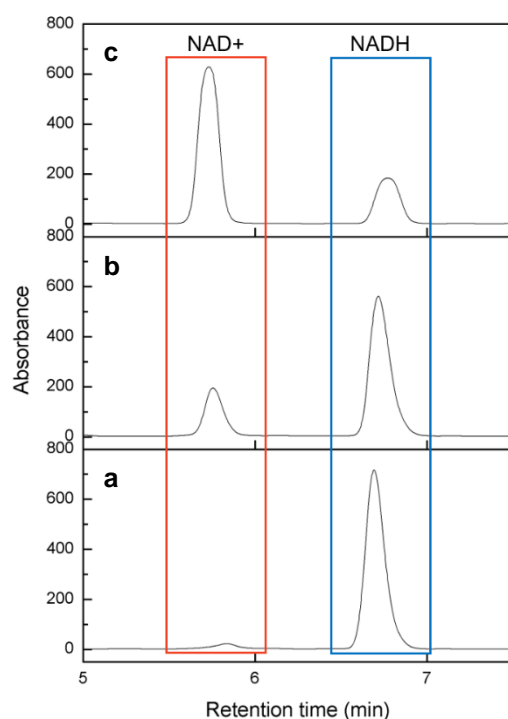
Figure A2.23 | Effect of NaCl on the NADH oxidation reaction induced by reduced L-glutathione ferric mononuclear complex studied by HPLC analysis (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm). a, Solution of NADH (0.1 mM) in neutral pH after 6 hours. b, Solution of NADH (0.5 mM) and L-glutathione-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM L-glutathione, pH 8.6) after 6 hours from mixing. c, Solution of NADH (0.5 mM) and L-glutathione-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM L-glutathione, pH 8.6) in the presence of NaCl (40 mM) after 6 hours from mixing. The peak centred at 6.69 min is due to NADH, the peak centred at 5.72 min is due to NAD+.
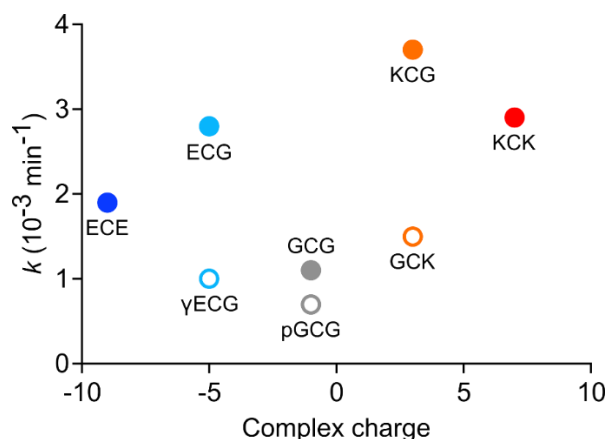
Figure A2.24 | Relationship between the overall ferric complex charge with a different peptide coordination and the value of the rate constant, thus the efficiency of the reaction. The best result was obtained with the positively charged mononuclear complex with LysCysGly. The other charged peptides LysCysLys, GluCysGly and GluCysGlu made the ferric complex still efficient towards NADH oxidation. GlyCysGly in the unprotected and protected forms made the mononuclear complex less functional. Even though γGluCysGly and GlyCysLys have the same charge of GluCysGly and LysCysGly respectively, the related mononuclear complexes appear less efficient, thus implying that not only the presence of charged groups, but also the peptide structure is fundamental to obtain a functional ferric complex. Data are obtained from HPLC analyses.
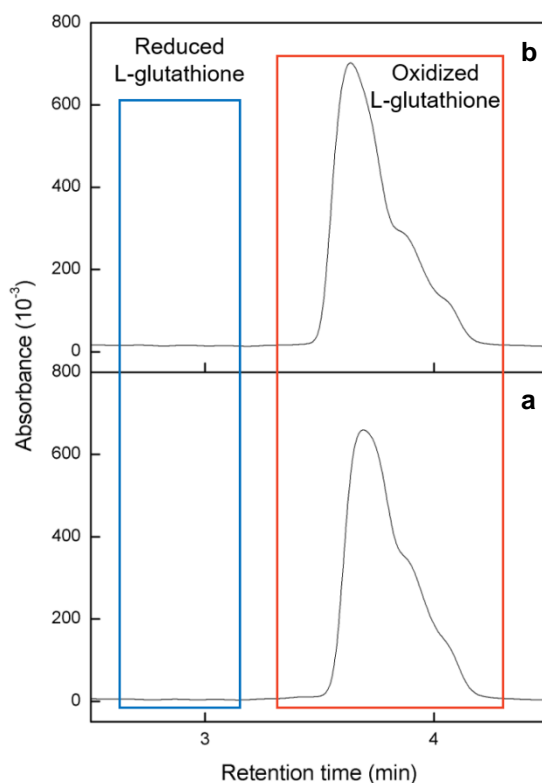
Figure A2.25 | Effect of oxidized L-glutathione on NADH studied by HPLC analysis (gradient run: 3-18% eluent B, eluent A: water + 0.1% TFA, eluent B: acetonitrile + 0.1% TFA, flow rate: 1 mL/min, detection wavelength: 215 nm). Solution of 40 mM oxidized L-glutathione and 0.5 mM NADH (pH 8.6) at 0 min (a) and 360 min (b), showing that no conversion of oxidized L-glutathione to reduced L-glutathione occurs. The peak centred at 4.24 min is due to oxidized L-glutathione.
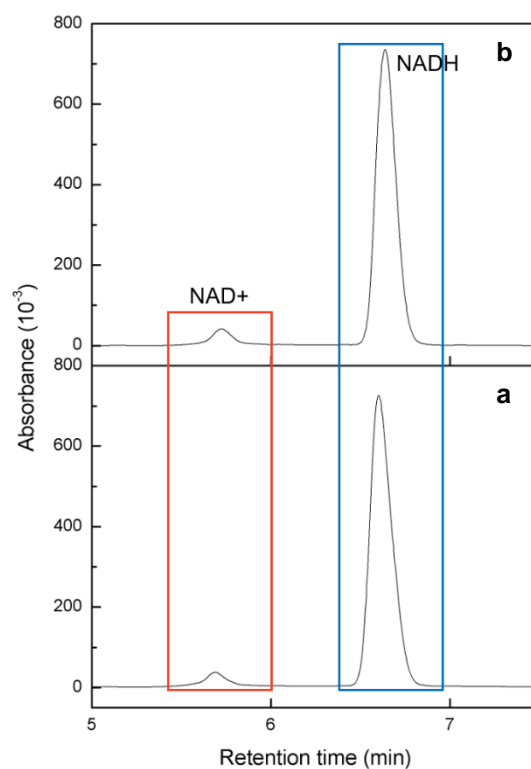
Figure A2.26 | Effect of oxidized L-glutathione on NADH studied by HPLC analysis (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm). Solution of 40 mM oxidized L-glutathione and 0.5 mM NADH (pH 8.6) at 0 min (a) and 360 min (b), showing that no conversion of NADH to NAD+ occurs. The peak centred at 6.69 min is due to NADH, the peak centred at 5.72 min is due to NAD+.
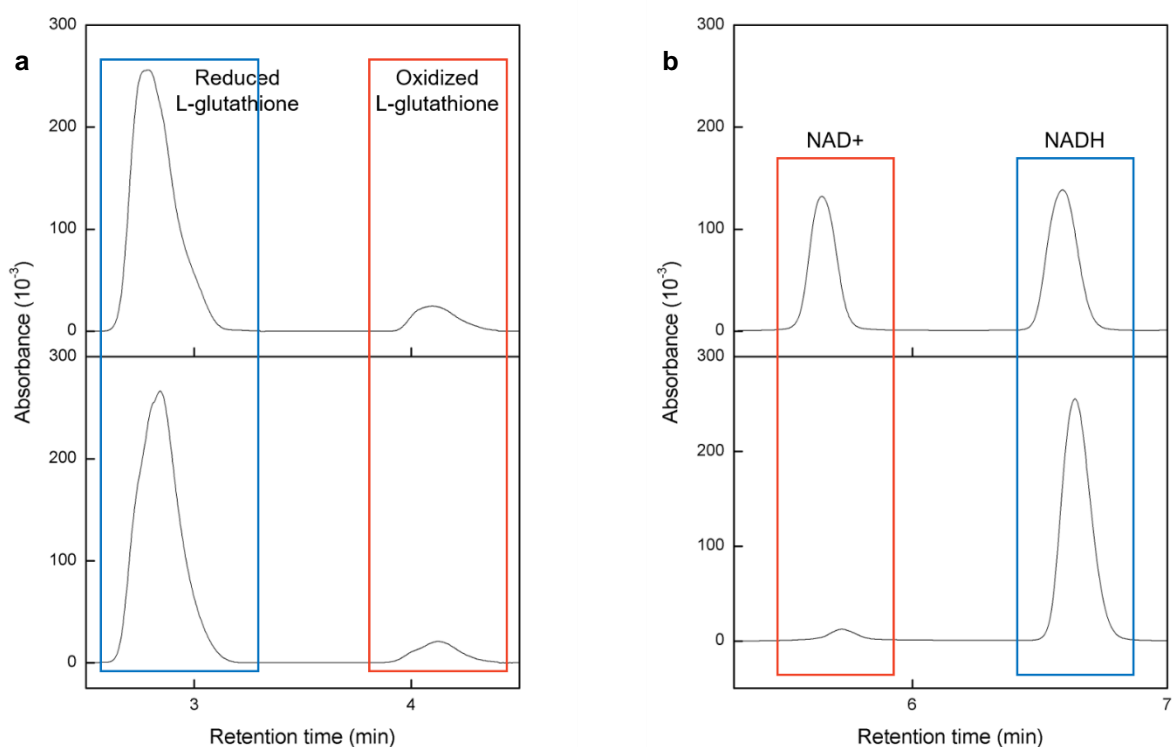
Figure A2.27 | Effect of NADH oxidation induced by the mononuclear Fe(III) complex on free thiolates studied by HPLC analysis. a, Solution of mononuclear complex (0.5 mM Fe^{3+}, 40 mM reduced L-glutathione) with 0.5 mM NADH at 0 min (bottom), 180 min (top), showing that no conversion of reduced L-glutathione to oxidized L-glutathione occurs after 3 hours. The peak centred at 2.88 min is due to reduced L-glutathione, the peak centred at 4.24 min is due to oxidized L-glutathione (gradient run: 3-18% eluent B, eluent A: water + 0.1% TFA, eluent B: acetonitrile + 0.1% TFA, flow rate: 1 mL/min, detection wavelength: 215 nm). b, Solution of mononuclear complex (0.5 mM Fe^{3+}, 40 mM reduced L-glutathione) with 0.5 mM NADH at 0 min (bottom), 180 min (top), showing that NADH conversion occurs after 3 hours up to 50%. The peak centred at 6.69 min is due to NADH, the peak centred at 5.72 min is due to NAD+ (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm).
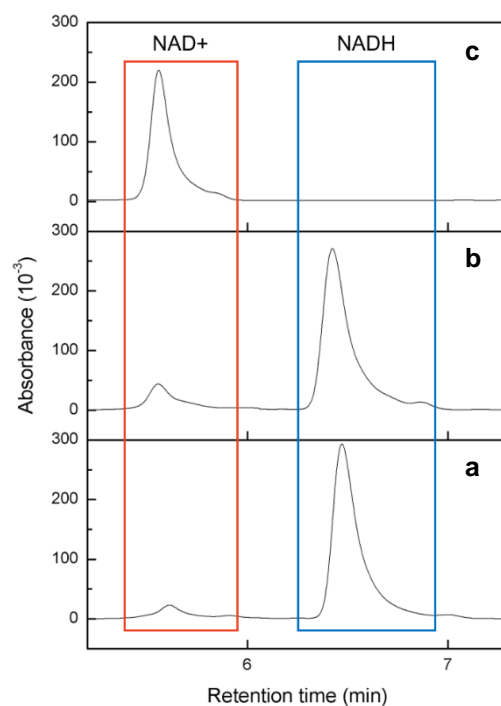
Figure A2.28 | NADH oxidation reaction induced by reduced L-glutathione ferric mononuclear complex performed in glove box ($[O_2]$ <20 ppm) studied by HPLC analysis (gradient run: 0-15% eluent B, eluent A: ammonium acetate 20 mM, pH 5.1, eluent B: acetonitrile, flow rate: 1 mL/min, detection wavelength: 260 nm). a, Solution of NADH (0.5 mM) and L-glutathione-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM L-glutathione, pH 8.6) after mixing. b, Solution of NADH (0.5 mM) and L-glutathione-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM L-glutathione, pH 8.6) after 6 hours from mixing. c, Effect of the addition of hydrogen peroxide (0.5 mM) on a solution of NADH (0.5 mM) and L-glutathione-Fe^{3+} ions (0.5 mM Fe^{3+}, 40 mM L-glutathione, pH 8.6) after 6 hours from mixing. The peak centred at 6.69 min is due to NADH, the peak centred at 5.72 min is due to NAD+. Hydrogen peroxide does not show any effect on NADH alone.
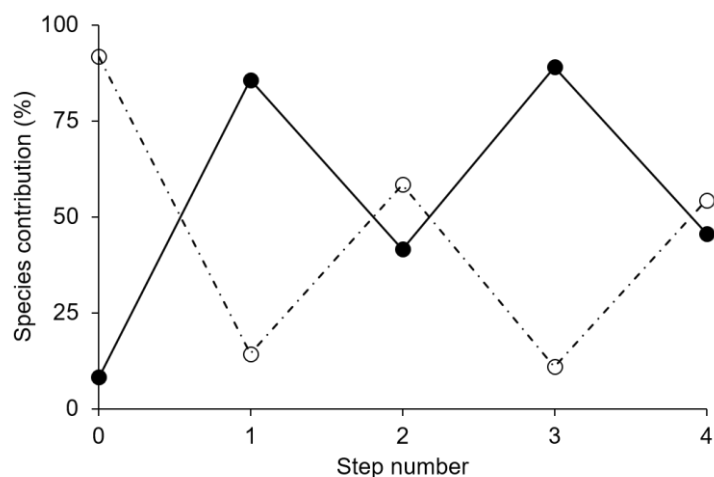
Figure A2.29 | Redox cycles of L-glutathione-Fe^{3+} mononuclear complex in the presence of NADH and H_2O_2. Ferric complex (open circles) and ferrous complex (filled circles) coordinated to L-glutathione during reduction (NADH addition) and oxidation (hydrogen peroxide addition) determined by UV spectral decomposition.
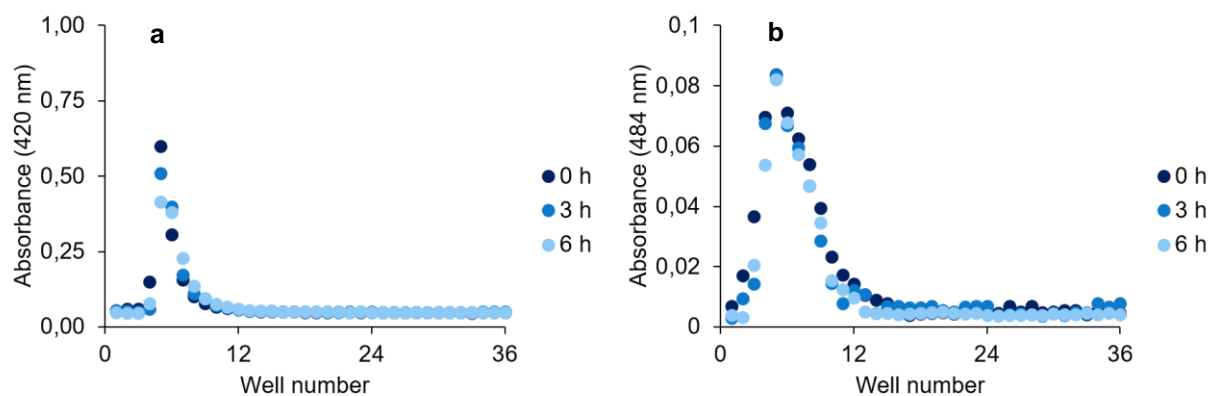

Figure A2.30 | Phospholipid vesicles stability assay. POPC vesicles containing encapsulated Fe^{3+}-glutathione in the presence of NADH at 0 min (dark blue dots), 180 min (blue dots), 360 min (light blue dots). a, The absorbance of POPC vesicles (420 nm) containing Fe^{3+}-L-glutathione in the presence of NADH was exploited to indicate structurally compromised vesicles. No decrease of the vesicle peak area indicating destruction of vesicles was detected. b, The absorbance of the tiron-iron complex was exploited to detect leaked iron ions from POPC vesicles containing Fe^{3+}-glutathione in the presence of NADH, to indicate structurally compromised vesicles. No leaked iron ions were detected.
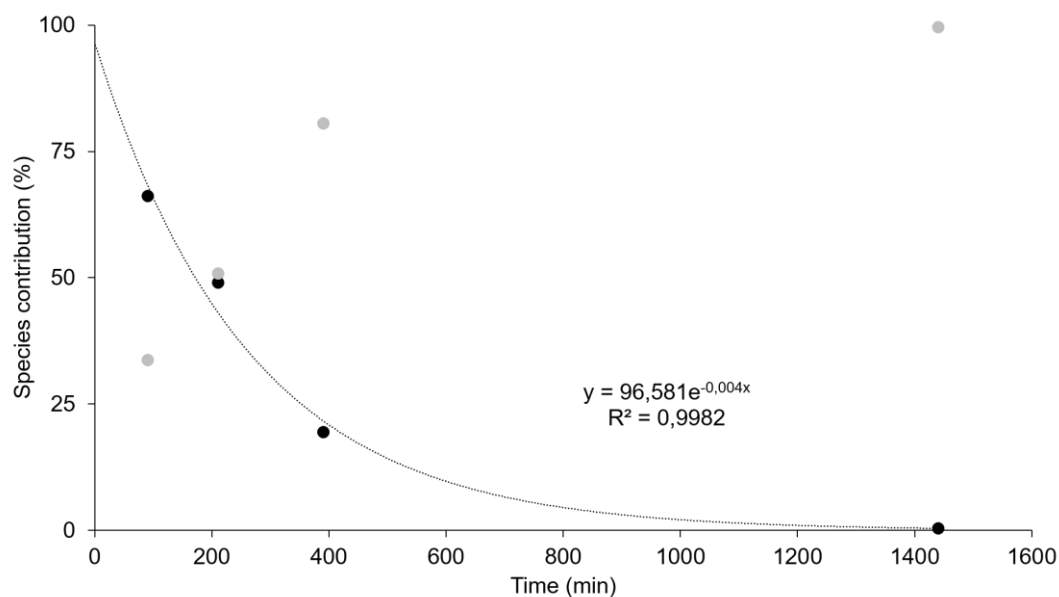
Figure A2.31 | NADH oxidation reaction occurring within oleic acid vesicles. First order kinetics was observed for the equimolar reaction between the mononuclear complex (0.5 mM) and NADH (0.5 mM) within oleic acid vesicles (20 mM). NADH and NAD+ concentrations are represented as black and grey dots, respectively.
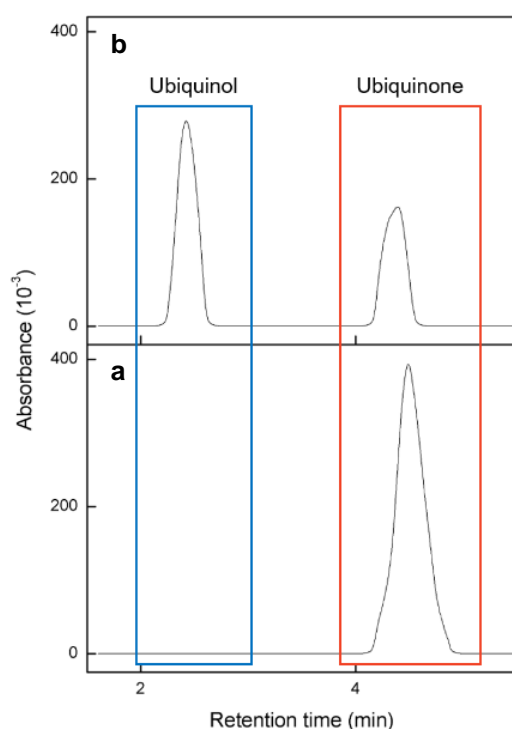

Figure A2.32 | Reduction of ubiquinone reaction performed within phospholipid vesicles studied by HPLC analysis (gradient run: 90-100% eluent B, eluent A: hexane:isopropanol:methanol 2:1:1, eluent B: isopropanol:methanol 1:9, flow rate: 1 mL/min, detection wavelength: 275 nm). Solution of 40 mM oxidized L-glutathione and 0.5 mM NADH (pH 8.6) encapsulated within 20 mM POPC vesicles enriched with 0.5 mM ubiquinone at 0 h (a) and 24 h (b), showing that a partial conversion of ubiquinone to ubiquinol occurs. The peak centred at 4.45 min is due to ubiquinone, the peak centred at 2.42 min is due to ubiquinol.
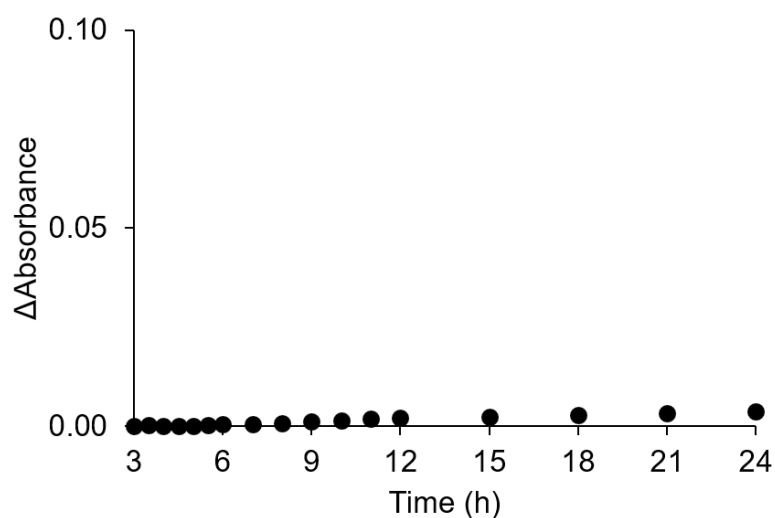
Figure A2.33 | 590 nm absorbance variation of the pH indicator dye in the absence of oxidizing agents for the mononuclear ferrous complex. The absorbance values are plotted from 3 to 24 hours in order to remove the initial contribute given by the mononuclear ferric complex at 590 nm. The experiment was set up as reported elsewhere with an excess of NADH (2.5 mM) and the mononuclear ferric complex (0.5 mM).
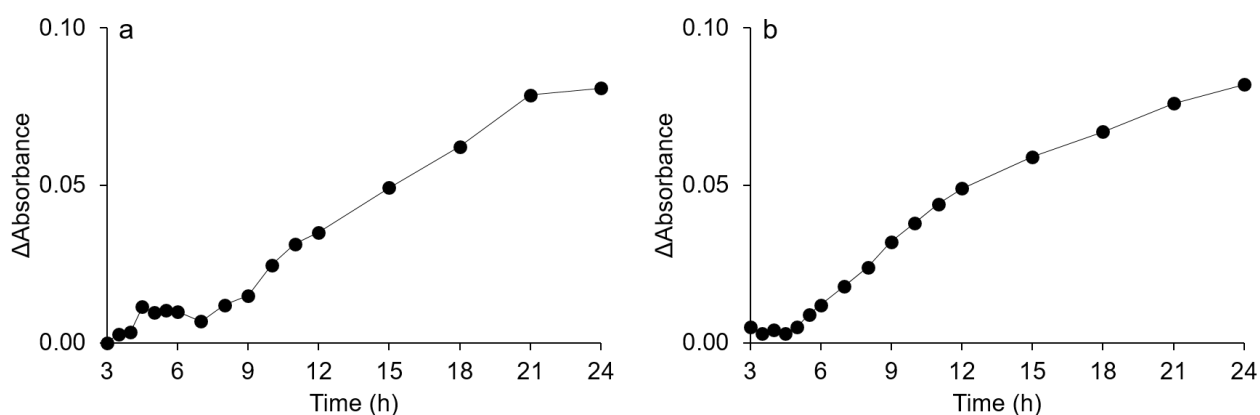

Figure A2.34 | 590 nm absorbance variation of the pH indicator dye induced by the oxidation of the mononuclear ferrous complex driven by O_2 or H_2O_2. The absorbance values are plotted from 3 to 24 hours in order to remove the initial contribute given by the mononuclear ferric complex at 590 nm. a, Absorbance variation of thymol blue (590 nm) induced by the release of OH^- from an excess of hydrogen peroxide. The experiment was set up as reported elsewhere with an excess of NADH (2.5 mM) and the mononuclear ferric complex (0.5 mM). b, Absorbance variation of thymol blue (590 nm) induced by the release of OH^- from molecular oxygen. The experiment was set up as reported elsewhere with an excess of NADH (2.5 mM) and the mononuclear ferric complex (0.5 mM).