International PhD Program in Biomolecular Sciences

Department of Cellular, Computational and Integrative Biology – CIBIO

34th Cycle

"Histidine ligated Iron-Sulfur Proteins and Peptides"

Tutor/Advisor

Sheref S. Mansy

Ph.D. Thesis of

Luca Valer

Academic Year 2021-2022

Declaration of Authorship

I, Luca Valer, declare that this thesis titled “Histidine ligated Iron-Sulfur Proteins and Peptides” is my own work and the use of all material from other sources has been properly and fully acknowledged.

Abstract

Iron-sulfur clusters play a fundamental role in biology and are believed to be ancient cofactors that could have played a role in early protometabolic systems. Thus far, redox active, prebiotically plausible iron-sulfur clusters have always been obtained through cysteine coordination to the iron ions. However, extant iron-sulfur proteins can be found to exploit other modes of binding, including ligation by histidine residues, as seen with [2Fe-2S] Rieske and MitoNEET proteins. In this thesis, we investigated the ability of cysteine- and histidine-containing proteins and peptides to coordinate mononuclear [1Fe-0S] centers and a [2Fe-2S] clusters. The iron-sulfur peptides were characterized by UV-vis, circular dichroism, and paramagnetic NMR spectroscopies and cyclic voltammetry. Small (≤ 6 amino acids) peptides can coordinate [2Fe-2S] clusters through a combination of cysteine and histidine residues with similar reduction potentials as their corresponding proteins. Such complexes may have been important for early cell-like systems.

Thesis Organization

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.

Chapter 2

Chapter 2 is focused on the expression, purification and subsequent spectrophotometric characterization of several iron-sulfur protein. Supporting information for Chapter 2 are in Appendix 1

Chapter 3

Chapter 3 is focused on the synthesis and characterization of iron-sulfur cluster using small peptides containing histidine and cysteine. Supporting information for Chapter 3 are in Appendix 2

Chapter 4

Chapter 4 summarize the results of this work and discuss the future perspectives of this project.

The data presented in this thesis have been published in:

Luca Valer, Daniele Rossetto, Taylor Parkkila, Lorenzo Sebastianelli, Graziano Guella, Amber Hendricks, James Cowan, Lingzi Sang, Sheref S. Mansy. Histidine ligated iron-sulfur peptides. In *ChemBioChem* (June 2022) (https://doi.org/10.1002/cbic.202200202)

List of Publications

- **Luca Valer**, Daniele Rossetto, Taylor Parkkila, Lorenzo Sebastianelli, Graziano Guella, Amber Hendricks, James Cowan, Lingzi Sang, Sheref S. Mansy. Histidine ligated iron-sulfur peptides. In *ChemBioChem* (June 2022)

- Daniele Rossetto, **Luca Valer**, Noel Yeh, Graziano Guella, Yayoi Hongo & Sheref S. Mansy. Prebiotic Environments with Mg^{2+} and Thiophilic Metal Ions Increase the Thermal Stability of Cysteine and Non-cysteine Peptides. In *ACS Earth and Space Chemistry* (April 2022)

- **Luca Valer**, Daniele Rossetto, Simone Scintilla, Yin Juan Hu, Anju Tomar, Serge Nader, Isaiah O. Betinol & Sheref S. Mansy. Methods to identify and characterize iron-sulfur oligopeptides in water. In *Canadian Journal of Chemistry* (January 2022)

- Alessandro Groaz, Silvia Galvan, **Luca Valer**, Daniele Rossetto, Filippo Benedetti, Graziano Guella, Ömer Duhan Toparlak & Sheref S. Mansy. Cell-Free Synthesis of Dopamine and Serotonin in Two Steps with Purified Enzymes. In *Advanced Biosystems* (October 2020)

- Noel Yeh, **Luca Valer** & Sheref S. Mansy. Toward long-lasting artificial cells that better mimic natural living cells. In *Emerging Topics in Life Sciences* (September 2019)

- Claudia Bonfio, **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. In *Nature Chemistry* (July 2017)

Acknowledgements

This thesis represents the end of a almost four year long journey. During this time I met a lot of wonderful people that helped me to reach this goal. Now I would like to take some time to thank them for the support they showed me.

First, I would like to thank Prof. Sheref S. Mansy for giving me this wonderful opportunity to do my PhD in his laboratory. His support was fundamental; his way of thinking, both in science and in life, deeply inspired me.

Secondly I would like to thank all the Mansy Lab members, past and present. I'll start by thanking especially Claudia Bonfio for introducing me to this wonderful team and for guiding my first steps in the origin of life field when I was a bachelor student. I would also like to thanks all the other member of the lab Rajanya Banarjee, Noël Yeh-Martin, Shibaji Basak, Simone Scintilla, Duhan Toparlak, Dario Cecchi, Giuliano Berloffa, Alessandro Groaz, Filippo Benedetti, Dongman Xi, Yin Juan Hu, Taylor Parkkila. In particular I'd like to thanks Daniele Rossetto, Anju, Alexandre Baccouche Serge Nader and Lorenzo Sebastianelli, for sharing with me the best (and the worst) of Canada.

I would like to acknowledge our collaborators as well. I thank Graziano Guella from the University of Trento, James Cowan from the Ohio State University and Lingzi Sang from the University of Alberta.

I would also like to thank my friends. They have always supported me even though what I do is basically incomprehensible for them.

And finally, I cannot describe how much grateful I am to my family, my mom, my dad, my brother for never ending support and love.

Table of Content

List of Abbreviations

CD Circular Dichroism

CV Cyclic Voltammetry

K_D Dissociation constant

LUCA Last Universal Common Ancestor

NMR Nuclear Magnetic Resonance

RR Resonance Raman

SHE Standard Hydrogen Electrode

SPPS Solid Phase Peptide Synthesis

UV Ultraviolet

UV-vis. Ultraviolet-visible

Chapter 1 - Introduction

1.1 The early Earth

The Earth formed around 4.56 billion years ago (Ga).[1] The period of time from its formation to 4 Ga is referred to as the Hadean. According to the standard model, the Earth grew by agglomeration of planetesimals that formed in the solar nebula.[2] Over a period of 100 million years, the young Earth formed a partially solid surface as temperature decreased. During that time, volatile species started to be released due to the volcanic activity of the newly formed crust. The most abundant gases were water (H_2O), carbon dioxide (CO_2) and molecular nitrogen (N_2). Other species present were hydro hydrogen sulfide (SH_2), methane (CH_4) and ammonia (NH_3). Molecular hydrogen (H_2) was also likely present, although it can reach easily escape velocity and thus its concentration in the early atmosphere would have dropped over time.[3] With the temperature decreasing below the water condensation point, the water present in the atmosphere began to condense, forming the first clouds and falling as rains. The water would eventually accumulate into oceans and other smaller bodies of water. The studies of sedimentary rocks shows that the oceans were rich in carbon and in many other elements, including Na, Cl, K, B, F, P, Mg, Ca, Si, Mo, Co, Ni, and Fe. However, over geological time scales, composition of both the atmosphere and the ocean went through drastic changes (Figure 1.1).[4]

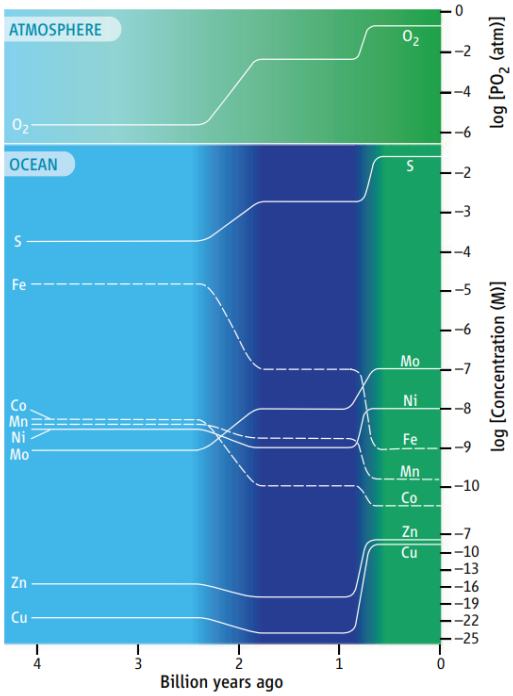

Figure 1.1 - Variations in elemental concentrations through time. Color gradients indicate a transition from sulfur-poor, anoxic oceans before 2.4 billion years ago (light blue region) to hydrogen sulfide-rich oceans between 1.8 and 0.8 billion years ago (blue region) to the complete oxygenation of the ocean (green region). Dashed lines are for elements with decreasing concentrations (adapted from Anbar, 2008).

It is believed that life appeared on Earth not long after the oceans formed. Life must have emerged from a process or a series of processes and not in a sudden event.[5] Lifemust have started from simple abiotic processes, involving one or more sources of energy and matter, in particular carbon (C), hydrogen (H), nitrogen (N), oxygen (O), Phosphorus (P) and sulfur (S). The earliest undisputed fossil evidence places life on Earth prior to 3.5 Ga, showing bacterial-like cellular entities encased in sedimentary rocks.[6] The isotope ratio between $^{12}C/^{13}C$ is often used as a biosignature due to the isotopic fractionation that occurs when carbon is fixed. Such signatures may have been found in zircons dating back to 4.1 Ga.[7] Estimates suggested that the last universal common ancestor (LUCA) existed in the Archean, a few hundred million years before the earliest fossil evidence of life (Figure 1.2).[8]

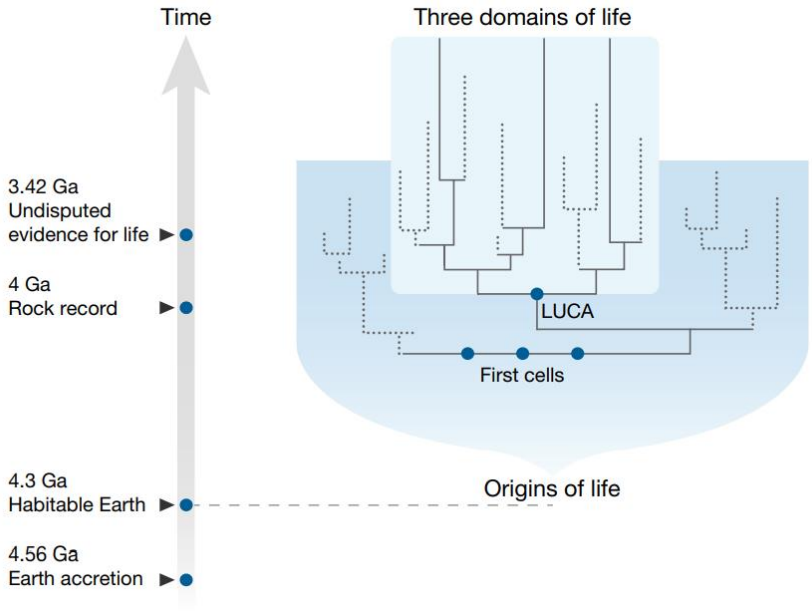

Figure 1.2 – Schematic representation of the evolution of early life (adapted from Javaux, 2019)

1.2 The beginning of origin of life research

The cell was first discovered by Robert Hooke in 1665. At that time, most of the people believed that living creatures possessed a magical property called the "vital spark" that made them different from inanimate objects. It took two centuries for biologists to realize that all organisms were made of cells, and this realization led to the development what is now known as the cell theory.

In 1828 Friedrich Wöhler, reporting the chemical synthesis of urea from ammonium cyanate, demonstrated for the first time that biological molecules could be made from simpler, "non-living" molecules.[9] Only in 1859, with his famous experiments, Louis Pasteur disproved the spontaneous generation theory.[10] In the same year, Charles Darwin published The Origin of Species and he consciously avoided discussing the origin of life. However in some correspondence with friends and colleagues he conceived a "warm little pond" as

the place where life started.[11] In the 1920s, this idea sparked the interest of Aleksandr Oparin and John Haldane. They both independently proposed what's now called the Oparin-Haldane hypothesis.[12,13] In the Oparin-Haldane hypothesis, life formed in a "primordial soup" of simple chemicals that could be combined together to form modern macromolecules. However, no experimental data was provided to confirm their theory.

In a 1952 paper Harold C. Urey, following the ideas of Oparin and Haldane, defined three key requirement for the origin of life: the spontaneous formation of the chemical compounds which form the physical bodies of living organisms, the evolution of the complex chemical reactions which are the dynamic basis of life, and a source of free energy which can maintain the chemical reactions and synthesize the chemical compounds.[14] The first part of this three pieces puzzle was addressed by Stanley L. Miller, a graduate student in Urey's laboratory. Miller used an apparatus to reproduce the primordial Earth's atmosphere-ocean system to test the first question. A 500 mL flask half-full of water simulated the ocean, and it was connected to a bigger flask filled with methane, ammonia, and hydrogen, that acted as a primordial atmosphere. The water in the smaller flask was heated to induce evaporation, and the water vapor was allowed to enter the larger flask. Inside the larger flask, a continuous electrical spark was fired between two electrodes, to simulate lightning. Then the simulated atmosphere was cooled again so that the water vapors were condensed back into the "ocean" (Figure 1.3). The system was left cycling for more than a week and then the "ocean" water was analyzed. Surprisingly, the results showed the accumulation of organic molecules, such as simple amino acids, in the water.[15]

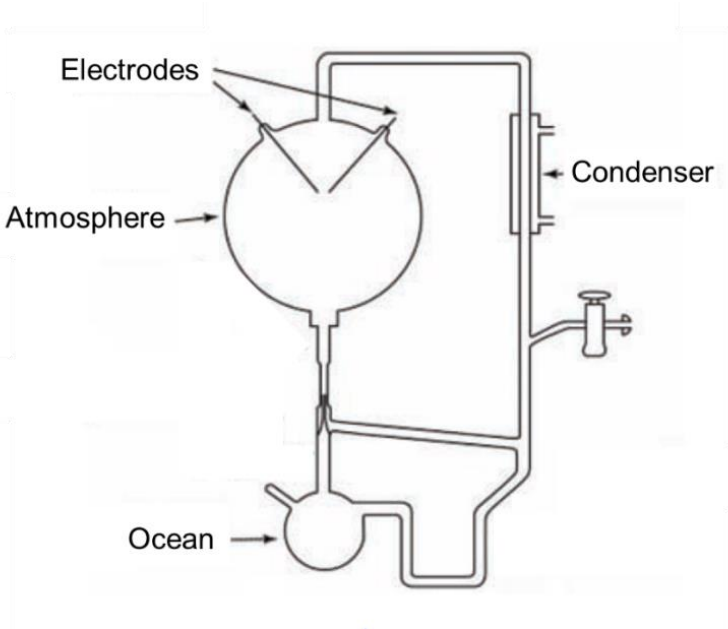

Figure 1.3 – The spark discharge apparatus used by Miller in its experiments (adapted from Johnson et al., 2008)

Later studies demonstrated that the primordial atmosphere might have had a mixture of gases different from the one used in this experiment.[16] However, after Miller's death, Jeffrey L. Bada reanalyzed many of Miller's old samples. In 2011 he published the results of previously unpublished experiment, where Miller added to the mixture of gases also H_2S. Strikingly, the molar ratio of the synthesized amino acids was very similar to those found in the Murchison meteorite carbonaceous chondrites.[17] In any case, whether the experiments reproduced accurately the conditions of the early Earth or not, this experiment has been considered the experiment that kick-started the field of the modern research on the origins of life.

1.3 Bottom-up and top-down approaches

To tackle the origin of life problem one can decide to go from simple chemistry and geology onward or to go from the complexity of extant biology backward. The two different approaches take the name of bottom-up and top-down, respectively. Both the approaches are undoubtedly important to understand how life emerged on the primitive Earth.

Planetary science and geology indicate that the primordial Earth possessed a wide variety of environments, such as a shallow pond, as Darwin imagined, or a deep-sea hydrothermal vent. Each environment can differ widely from one another and yet, there is insufficient constraints from geochemistry alone to settle on one particular scenario. After Miller's publication, the synthetic quest continued with efforts focused on the synthesis of other important bio-molecules. The synthesis of amino acids is achieved usually involving the Strecker or the Bücherer−Bergs synthesis. Sugar are produced through chemistry usually based on the formose reaction, and nucleobase exploiting the chemistry of nitriles and their derivatives. Scientists, like John Sutherland, proposed a the type of chemistry based heavily on hydrogen cyanide, that could abiotically produce life's building blocks from a set of interconnected smaller systems (Figure 1.4).[18]

Another aspect of the bottom-up approach focuses on how compartmentalization was achieved. Generally, scientists have used fatty acids vesicles as models of protocells. While fatty acids are considered prebiotically plausible, they are not stable across a wide range of pH and metal ion concentrations. Recently Krishnamurthy and his team showed that using the prebiotically plausible phosphorylation agent diamidophosphate (DAP) is possible to synthesize cyclicphospholipids from a mixture of glycerol and nonanoic acid.[19,20] The cyclicphospholipids produced vesicles that are stable across several units of pH and were not disrupted by high concentration of metal ions, such as Na^+.[21]

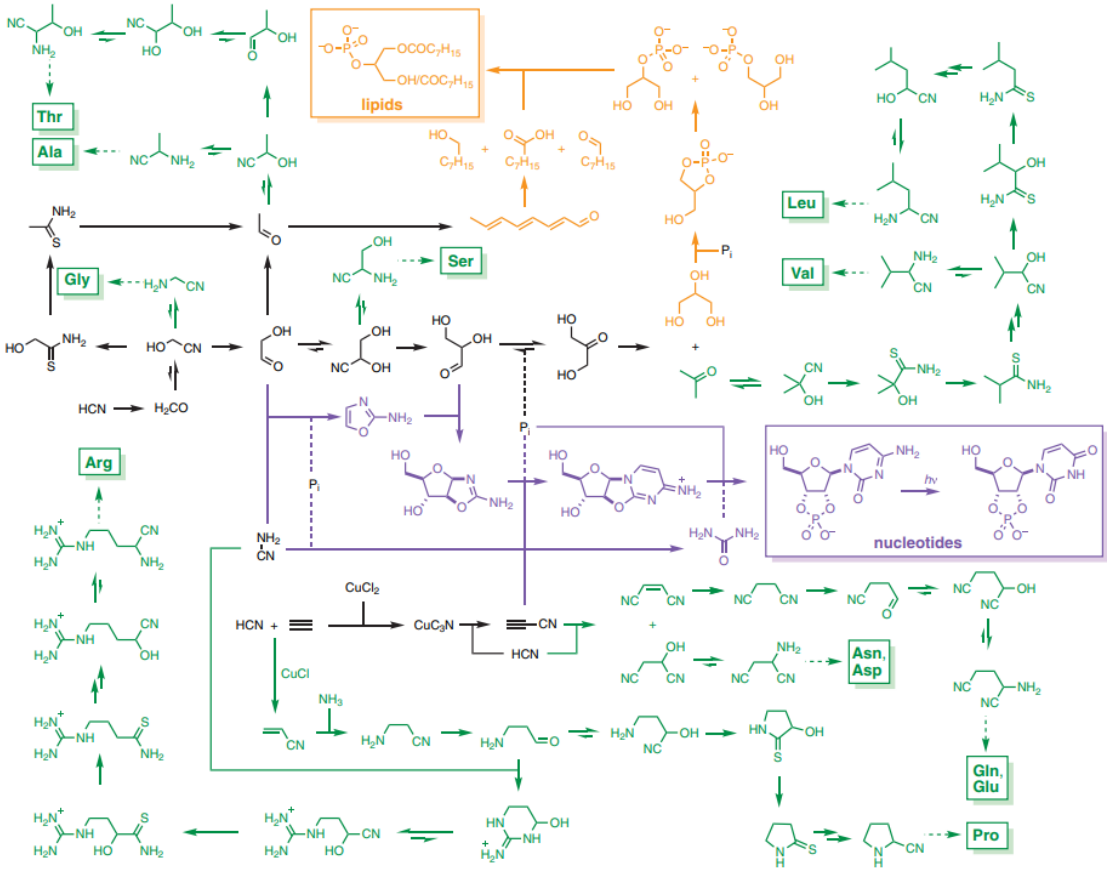

Figure 1.4 - Cyanosulfidic protometabolic reaction network leading from HCN and derivatives thereof to pyrimidine ribonucleotides, and amino acid and lipid precursors. Pathways, intermediates and products specific to a particular product class are shown in color (pyrimidine ribonucleotides, purple; amino acids, green; lipids, orange) with products boxed, shared starting materials, pathways and intermediates are shown in black. (adapted from Wu et al., 2021)

One could say that the top-down approach was first used by Darwin after his discovery of evolution; however, this approach did not produce many results until scientists started to understand how genetics worked and started to compare genomes of different organisms. Their goal was to understand what are the minimal components (genes) on which all life forms relied upon, speculating that this minimal genome might be similar to the genetic content possessed by LUCA. Prokaryotes, are the simplest forms of cellular life, but even among them their genetic content may hugely vary (Figure 1.5).[22] Genome sequences show that parasitic bacteria have 500–1200 genes, free-living bacteria have 1500–7500 genes, and archaea have 1500–2700 genes. However, the study of phylogeny can help only partially to understand how chemistry produced the first cell. Even the experimental works conducted by scientists like Craig Venter reported the creation of a synthetic genome having the fewest genes of any free living organism (473 genes) can take us that far.[23] Some argue that the gap between the origin of life and the biochemical nature of LUCA is so vast that whatever

constituted the chemistry leading to the earliest life could have been rewritten multiple times and changed beyond recognition to then become LUCA.[24]

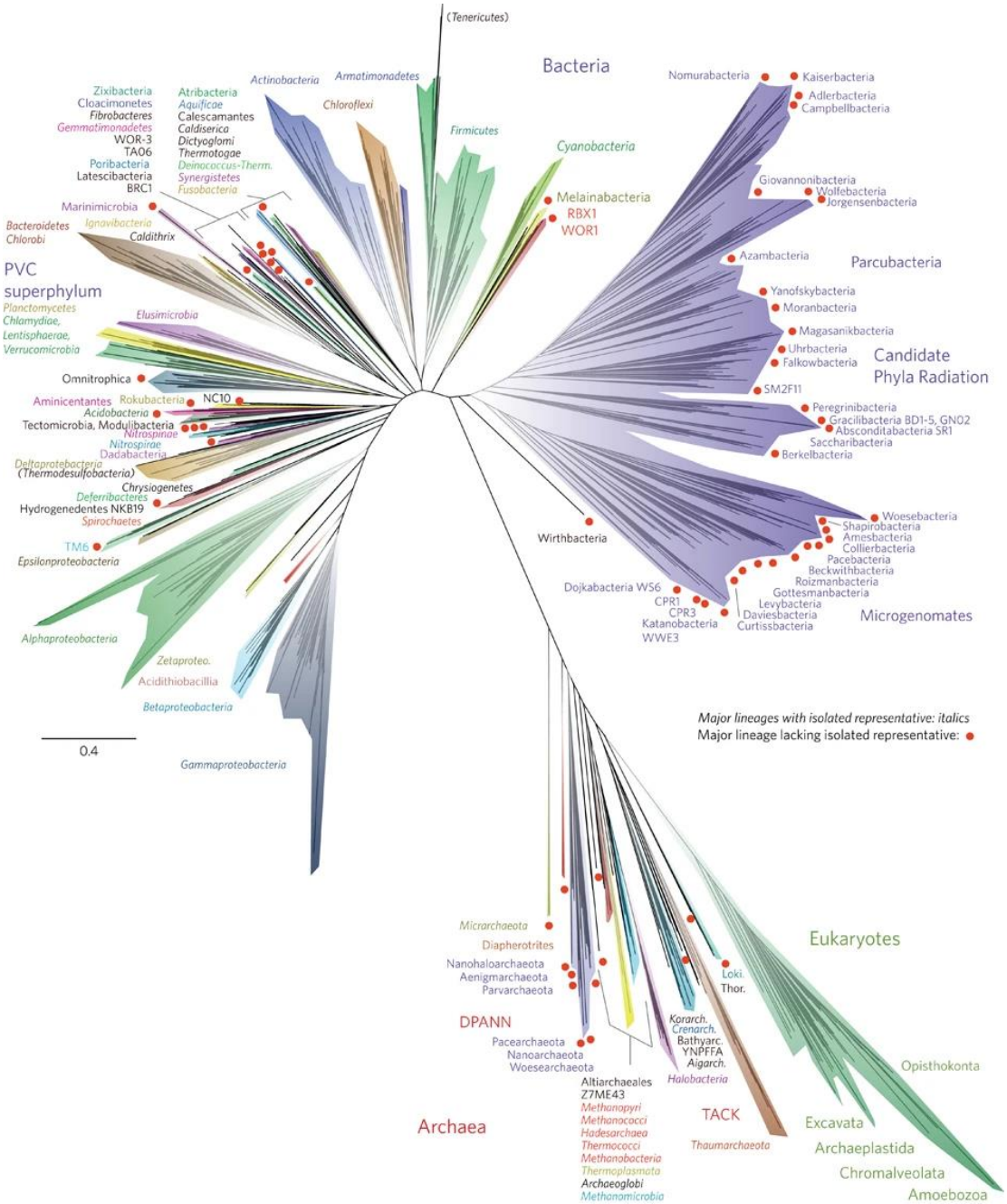

Figure 1.5 – A current view of the tree of life, encompassing the total diversity represented by sequenced genomes. (adapted from Hug et al., 2016)

1.4 RNA World

In the same year of Miller's publication, Watson and Crick announced the structure of DNA. In the decades that followed that discovery, biologist unveiled the central role and functions of protein and nucleic acids. Origin of life researcher started to hypothesize plausible ways from which such complexes and intertwined mechanisms came to be. The ability of RNA to fold suggests the

ability to catalyze chemical reactions. In 1967, Woese, Crick, and Orgel were the first to suggest that RNA could act as a catalyst.[25,26] Only in 1982 was this hypothesis shown to be true when Cech and colleagues demonstrated that RNA molecules do not only carry the genetic information but they can have catalytic functions and can participate in cellular reactions.[27]

The term “RNA World” was first used by Walter Gilbert in 1986 to describe the scenario where molecules of RNA could replicate and at the same time catalyze the indispensable reactions that could have led to the origin of life.[28] Many found the idea compelling: a world where RNA molecules, catalyzing the formation of indispensable cellular scaffolds, from which somehow cells emerged. However, other scientists were not completely convinced and pointed out the several problems with this scenario. The lack of templates enabling the polymerization of RNA in the prebiotic complex mixture, RNA’s extreme instability at moderate to high temperatures, and its susceptibility to degradation caused by base-catalyzed hydrolysis are examples. Another big criticism to the hypothesis is that RNA catalyzed reactions are not known in central metabolism in extant biology, so it is still unclear how this kind of reaction could have been carried out within a strict RNA model.

From the first formulation of the RNA world hypothesis some challenges have been surpassed. The prebiotic synthesis of nucleotides initially appeared problematic suffering from limited regio- and stereoselectivity and competing side reactions. However in recent years, progress has been made, for example advances towards a prebiotically plausible synthesis of nucleotides.[29] In 2015 Sutherland and co-workers showed that the synthesis of nucleotides could be pursued following an alternative pathway to the sequential sugar-nucleobases-phosphate scheme.[30] Alternative candidates to RNA for the primordial replicator have been proposed in the past years, they were obtained by changes in the backbone structure, such as with amide bonds in peptide nucleic acids (PNAs),[31] or composition, such as by substituting ribose with threose in threose nucleic acids (TNAs)[32] and in glycol nucleic acids (GNAs).[33]

1.5 Iron-sulfur world

In the late 1980s, a hypothesis, in contrast to the RNA world, was postulated by Günter Wächtershäuser. He proposed the idea of a “Metabolism-first World”, in which a primitive metabolism predated genetics. His idea was that early life may have been formed on the surface of iron sulfide minerals. Metal-composed minerals could have been early plausible alternatives to the metalloproteins that exist today.[34] In particular, he proposed that simple prebiotic molecules such as CO_2 and formaldehyde could have been reduced to produce more complex organic molecules in the presence of iron sulfide systems. The ability of mineral surfaces to adsorb organic molecules and act as a template to catalyze specific chemical reactions is widely recognized. Moreover, the adsorption of chemicals

in the interlayers of minerals could have had a protecting function against degradation induced by several external factors.[35]

In the following years, the geologist Michael Russell and the biologist William Martion expanded Wächtershäuse's theory to propose the idea of an iron-sulfur world. Iron sulfides in hydrothermal vents could have acted on the primitive Earth as catalysts. In those conditions, iron sulfides could potentially drive dehydration, isomerization, cleavage reactions as well as oxidation and reductions of organic compounds.[36–38] According to this theory, the transition metal centers in the mineral could have catalyzed autotrophic carbon fixation in order to produce organic compounds from inorganic gases and even synthetic reactions leading to primordial small peptides (figure 1.6).[34,39] Harold Morowitz and colleagues built upon the idea of an autotrophic chemical network to provide a theory for the origin of metabolic biochemistry. At the core of the metabolic network, Morowitz placed the (reductive) citric acid cycle, which represents the central driving force for all biosynthetic pathways.[40]

Even though the iron-sulfur world theory is intriguing, it has still unsolved problems. First of all, there is little experimental evidence for the processes that Russell and Martin describe. Another problem is that this theory does not provide a clear way for passage from mineral-based catalysis to the modern enzymes. However, some aspects of this theory must certainly be true, such as the fundamental role of metal ions. In fact, many prebiotically relevant reactions involve the use of metal redox couples or the reducing power of hydrogen sulfide, while other metals, such as magnesium, are crucial in some prebiotic polymerization reactions.

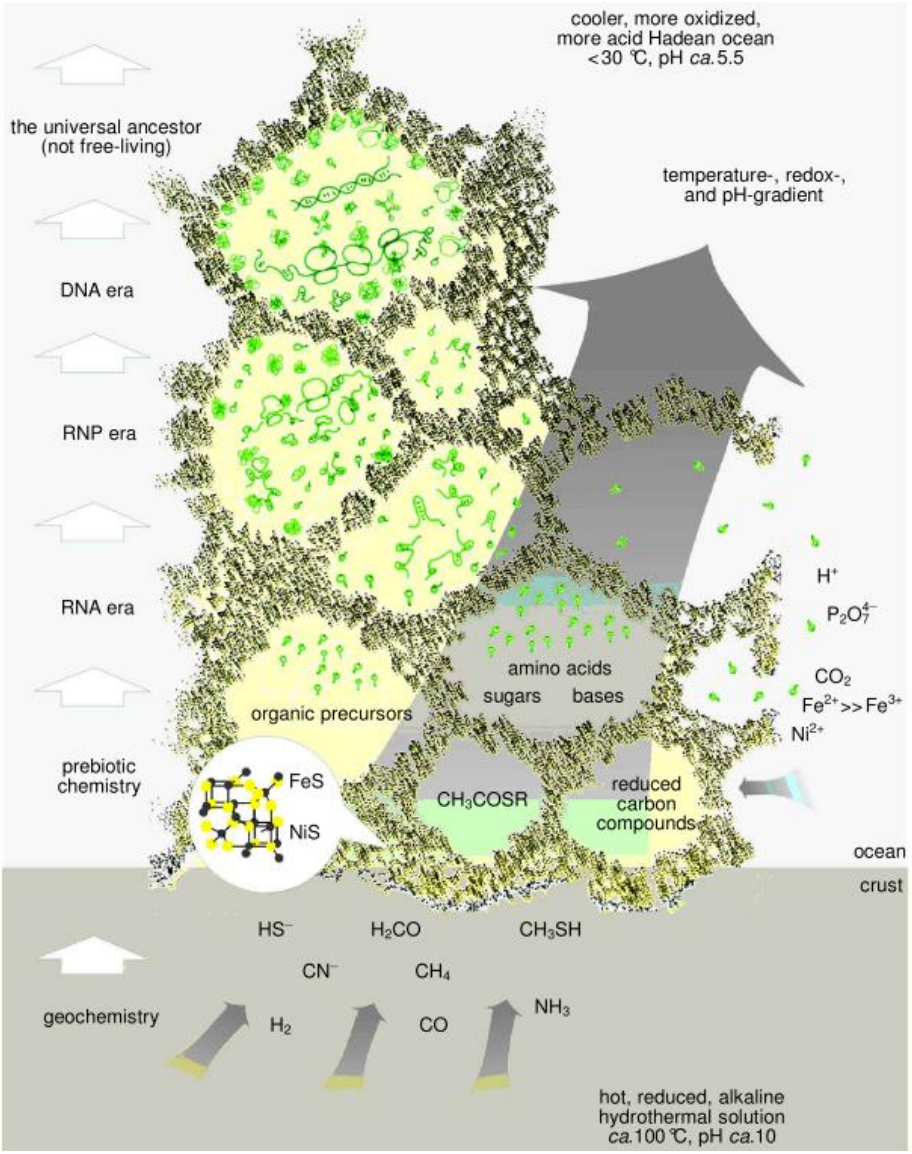

Figure 1.6 - A model for the origin of life at a redox, pH and temperature gradient at a submarine hydrothermal vent, as proposed by Russell and Martin. (adapted from William Martin and Michael J. Russell, 2003)

1.6 Many worlds

The RNA world and the iron-sulfur world theories have often been portrayed in opposition to each other.[41] However, other plausible prebiotic worlds came to light, each of them focusing on a preferred class of biomolecules. Scientists proposed the idea of a peptide world[42–44], or a lipid world[45,46] Other have theorized of coenzymes based world[47] or even a virus world[48] has been hypothesised. Each prebiotic world produces invaluable insight on its own class of biomolecules. But in the light of this, this way of thinking has the intrinsic limitation of considering its type of biomolecules in a privileged manner. This privileged position imposes a sequential hierarchy on the steps of the origin of life process. Each subsequent world predates the previous one until a living cell

originates. Still, the question of which class of biomolecules initiated the origin of life is still an open question.[49]

There is a strong argument to be made for the emergence of essential biomolecules to have been at least to some extent contemporaneous or in a interdependent way. In fact, cells are not mere collections of their chemical components, but highly dynamic, complex systems with multiple interlocked processes involving those components. For this reason, the emergence of life cannot be distilled down to biomolecular synthesis only. If common requirements between the origin of each biomolecule are found, they may provide necessary constraints for the origin of life problem. Ultimately, this might also help bridging different prebiotic world and produce a big advance in the origin of life field.

1.7 The importance of metals

Life as we know it needs metals. Metalloproteins and metalloenzymes are present in every living organism and metals play crucial roles in many life processes. Among the many functions that they play inside a cell, metals are also involved in signal transduction, usually Ca, Na, and Mg. They are exploited to reduce toxic reactive oxygen species, like Mn, Fe and Cu, and the last two are also able to perform electron transfer reaction, making them especially versatile in drive redox reaction.

Modern cells employ metal ions, like Zn, K, Mg and Mn, to stabilize the folding of proteins and enzymes as well as DNA and ribozymes. Additional vital roles provided by metalloproteins include oxygen transport (as hemoglobin), electron transfer vectors (as cytochrome) and many important enzymes necessary for the fixation of various elements in molecules, as the classes of enzymes dismutases, hydrogenases, nitrogenases, and catalases. In addition, even transcription factors and ribozymes have been classified as metalloproteins, as they require metal ions in their active sites to perform their catalytic function. These properties are strongly bound to the nature of the metal ions. Each metal ion has a preference in coordinating some ligands more than other.

Protein consensus sequences that coordinate a certain metal are found to be very well conserved among all life kingdoms and living organisms, underlining the presence of a specific integration of a specific metal on a specific protein moiety. And it is very attractive that the protein moiety of an enzyme could not have a single preference for a metal ion: DNA polymerase, an enzyme involved in the replication of DNA, usually bind Mg^{2+}, but can coordinate Mn^{2+} as well. However, its substitution affects activity of the enzyme.

An important consideration is that metals diffused in the prebiotic Earth were constituted in the same way as metals involved in the extant life. Redox-active metals such as Fe, Ni, Mn or Co, could have mainly promoted redox reactions also before the emergence of life. Interestingly, bare metal ions already possess

catalytic activity: Calvin showed in 1957 that iron ions can decompose hydrogen peroxide. However, metal ions could improve their catalytic activity when bound to polymer scaffolds: iron's ability to decompose hydrogen peroxide increases 1000-fold when bound to porphyrin, and an additional 10^7-fold improvement was observed when the porphyrin-bound iron was included in the protein scaffold of the enzyme catalase. More recently it has been shown that metals, like Zn^{2+}, Cr^{3+} and Fe^0, can drive non-enzymatically CO_2 fixation and several steps of the reverse citric acid cycle.[50,51]

The continued dependence on metals ions has led to the speculation that cellular chemistry observed today is a reflection of the types of reactions that were needed by the first protocellular systems that emerged on our planet. Iron is the most abundant transition metal of our universe, our planet, and of extant biology. Therefore, prebiotic chemistry on the early Earth must have been influenced by the iron that was present. Early enzymes may have more frequently exploited iron ions than genomic analyses indicate, as some modern-day enzymes that make use of other divalent cations may have coordinated Fe^{2+} in the past.

1.8 Iron-Sulfur proteins in modern cells

An important class of metalloproteins is the iron-sulfur proteins. The most common structures of iron-sulfur cluster, coordinated by such proteins, are mononuclear [1Fe-0S], [2Fe-2S] and [4Fe-4S] clusters (Figure 1.7). Polynuclear iron-sulfur clusters are additionally coordinated by acid labile, bridging sulfides.[52] Although protein ligation is typically provided by cysteine residues, there are several examples of non-cysteinyl ligands. For example, mononuclear, [2Fe-2S], and [4Fe-4S] clusters can be ligated by the sidechain of an Asp, Glu, or His.[53] In particular, [2Fe-2S] clusters in which one iron ion is coordinated by two cysteines and the other iron center is coordinated by two histidines take the name of Rieske cluster, from the name of the scientist that discovered them.[54]

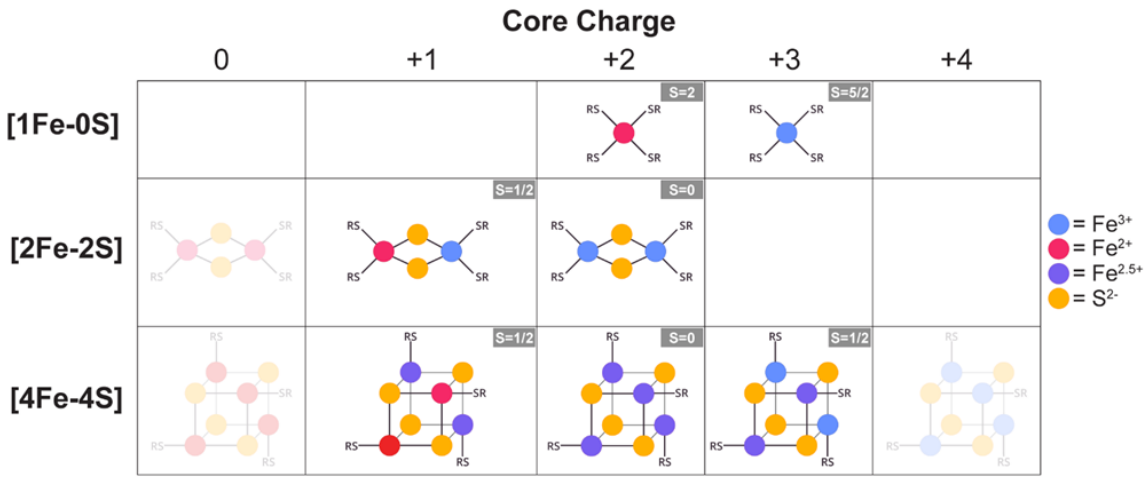

Figure 1.7 - Different types of iron-sulfur clusters commonly found in biology. [2Fe-2S]0, [4Fe-4S]0, and [4Fe-4S]4+ clusters do not typically form. (adapted from Valer et al., 2022)

One of the main feature of iron-sulfur clusters is their ability to switch between different oxidation states (Figure 1.6). [1Fe-0S] and [2Fe-2S] clusters are typically found in biology in two different oxidation states, i.e. $[1Fe\text{-}0S]^{3+/2+}$ and $[2Fe\text{-}2S]^{2+/+}$.[55] [4Fe-4S] proteins can be divided into two categories, low and high potential. Low potential iron-sulfur proteins contain a $[4Fe\text{-}4S]^{2+/+}$ cluster, and high potential iron-sulfur proteins possess a $[4F\text{-}4S]^{3+/2+}$ cluster.[56] Reduction potentials are strongly influenced by their coordination sphere within the polypeptide. This ability is one of the main reasons why these proteins are generally involved in mediating electron transfer reactions in metabolic processes, such as in respiration or in photosynthetic pathways. However, it is possible to find some unusual clusters such as the iron-sulfur cluster in aconitase. Aconitase is a [4Fe-4S] protein that under some conditions can lose an iron, thereby making a [3Fe-4S] cluster that is very unstable. The [4Fe-4S] cluster of aconitase does not take part in redox processes, instead the iron-sulfur cluster serves as a Lewis acid catalyst to convert citrate to isocitrate.[57] Iron-sulfur cluster are also involved in cellular processes, such as nitrogen fixation, protein and DNA synthesis. Their ubiquity is one of the reasons why these complexes are believed to be evolutionary ancient. Researches have also pointed out the structural similarities between some iron-sulfur clusters, like [4Fe-4S] clusters and the crystalline structure of minerals, like greigite.[58]

1.9 Prebiotic Iron-Sulfur Clusters

This synthetic approach to study such complexes started in early 1970s.[59] However, the synthesis were carried out in organic solvents, like methanol or ethanol. While these synthetic analogues are near to identical of the protein-bound sites and enabled to gain important details about their nuclearity, geometry and electronic structure, these systems did not raise prebiotic interest, as the majority of the reactions on the early Earth are thought to have taken place in the aqueous phase.

In recent year, iron-sulfur peptides have been synthesized also in aqueous solution. For example, the Cowan group synthesized a [2Fe-2S] cluster coordinated to the tripeptide L-glutathione (γECG) in water.[60] Later works of our group showed that iron-sulfur clusters could be stabilized in aqueous solution by a wide range of cysteine-containing di- and tripeptides. Also, synthesis can be achieved prebiotically exploiting UV light irradiation. Indeed, UV light is able to photoxidize Fe^{2+} to Fe^{3+} and at the same time it causes the photolysis of the sulfur of the from the thiols to produce SH^{-}, which are subsequently assembled into [2Fe-2S] and [4Fe-4S] clusters (Figure 1.8A).[61] Later we have shown that such metallopeptides are redox active and, as their biological counterparts, they are able to accept electrons from donors like NADH, then they are able to donate them to suitable acceptors (Figure 1.8B).[62]

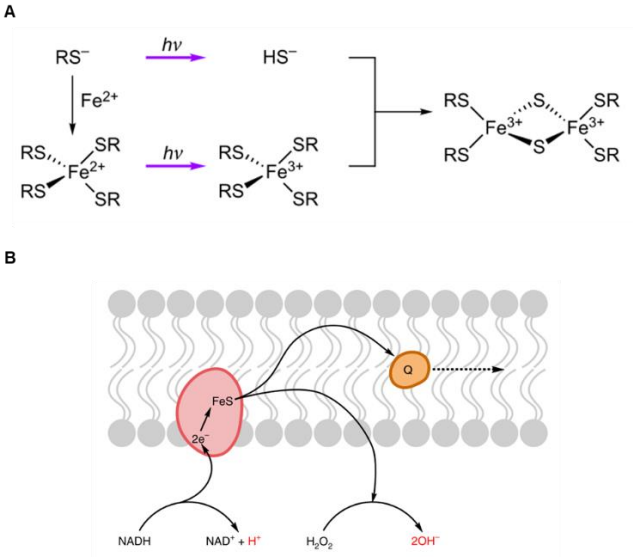

Figure 1.8 – **A.** Proposed mechanism of photooxidation and photolysis for the synthesis of a [2Fe–2S] cluster from a thiolate solution in the presence of ferrous ions. (Adapted from Bonfio et al., 2016) **B**. Schematic showing the path of electron transfer. NADH donates two electrons to two $[1Fe\text{-}0S]^{3+}$ peptides. The reduced metallopeptides can either transfer these electrons to a terminal electron acceptor (for example, H_2O_2), thereby making OH^- or to an intermediate electron carrier (for example, ubiquinone). (Adapted from Bonfio et al.,2017)

Interestingly, duplications of the glutathione sequence into hexa- and dodeca-peptides increase the stability of the cluster in water (Figure 1.9A).[63] This might suggest a plausible way for the evolution of these metalloenzymes in a way similar to the one proposed by Eck and Dayhoff. In a 1966 paper, they were the first to consider the possibility that modern metalloproteins could have arisen from repeated sequences of short peptides. The authors noted that a 55-amino acid redox-active protein from the bacterium *Clostridium pasteurianum* likely resulted from a duplication event (Figure 1.9B).[64]

The synthesis of longer iron-sulfur peptides frequently exploits motifs found in nature. For example, the $CX_3CX_2CX_2C$ motif, derived from the *Peptococcus aerogenes* ferredoxin I, have been exploited from the Dutton group to produce a [4Fe-4S] cluster maquette, however, the need for β-mercaptoethanol may suggests that one of the cysteines do not coordinate the cluster.[65] In particular, this motif possesses a CX_2C sequence that is extremely common in biology and likely is central for cluster loading and stability for [1Fe-0S], [2Fe-2S], and [4Fe-4S] clusters.[53] Frequently in iron-sulfur proteins, three of the coordinating cysteine residues are close in primary sequence (2-3 amino acids of spacing) with the forth in a more distant position. Alternatively, two different two cysteine-containing motifs can be found in which all four cysteines together bind a single [1Fe-0S] or [2Fe-2S] cluster.[66] Other motifs that generally have been used to synthesize iron-sulfur peptides include $CX_2CX_2CX_2C$ for ferredoxin-like sequences. Radical SAM-inspired motifs that contain only three cysteines, such as CX_2CX_2C and CX_3CX_2C, have been also used to synthesize iron-sulfur clusters.[67]

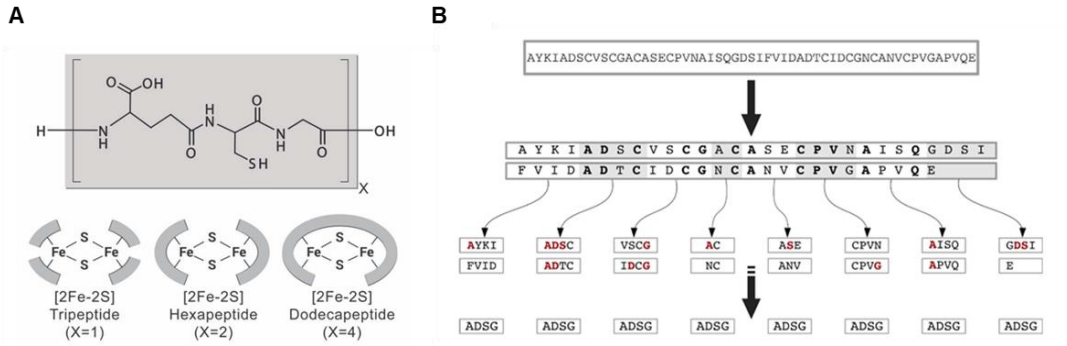

Figure 1.9 – **A.** Glutathione (top, grey block) can stabilize the formation of a [2Fe-2S] cluster (X=1, bottom left). Two and four duplications of the tripeptide lead to longer [2Fe-2S] polymers (X=2 and 4, bottom centre and bottom right, respectively). **B.** Schematic representation of Eck and Dayhoff scenario: extant metalloprotein can be broken down into repeating peptide units. Each box represents a peptide. The full-length sequence of the iron–sulfur cluster protein ferredoxin from Clostridium pasteurianum is shown atthe top. Identical residues at equivalent positions are shown in bold. Positions in red are equivalent to a hypothetical progenitor tetrapeptide [Aspartic Acid (Asp) – Glutamic Acid (Glu) – Serine (Ser) – Glycine (Gly)]. (adapted from Metal Catalysts and the Origin of Life).

In the case of scaffolds that do not provide complete four thiolate coordination to the cluster, the same peptide can bind different types of iron-sulfur clusters at the same time. For example, it is not unusual for peptides, that contains only one or two cysteine, to stabilize the formation of a complex mixture of [1Fe-0S], [2Fe-2S], and [4Fe-4S] clusters in aqueous solution, with the relative ratios reflecting the amount of hydrosulfide present.[68] Due to a large variety of iron-sulfur clusters, it is necessary to have proper tools to characterize these clusters in a precise and efficient way.[69]

1.10 How to study Iron-Sulfur cluster coordinated by small ligands

Iron-sulfur peptides can be challenging to study. In addition to being often unstable in vitro, iron–sulfur clusters can be found in a wide variety of forms with varied characteristics, which makes it difficult to easily discern what is in solution. This difficulty is mainly caused by the dynamics of iron–sulfur peptides, which frequently coordinate multiple types of clusters simultaneously. In fact, such complexes are generally studied combining multiples techniques.

UV–vis spectroscopy is the easiest and most accessible technique to characterize iron–sulfur proteins and peptides. Generally, iron–sulfur clusters absorb in the 300–800 nm range, reflecting the LMCT (ligand-to-metal charge transfer) transitions. The wavelength of the absorption maxima can provide some insight on presence of specific iron-sulfur species (figure 1.10A). For example, $[1Fe\text{-}0S]^{3+}$ clusters display a single broad band at 490 nm,[70] $[2Fe\text{-}2S]^{2+}$ have two absorption maxima at 420 and 450 nm[71] and $[4Fe\text{-}4S]^{2+}$ shows a broad band

between 400 and 420 nm.[72] However other techniques are generally required to fully characterize a mixture where multiple iron-sulfur species are present together (Figure 1.10B).

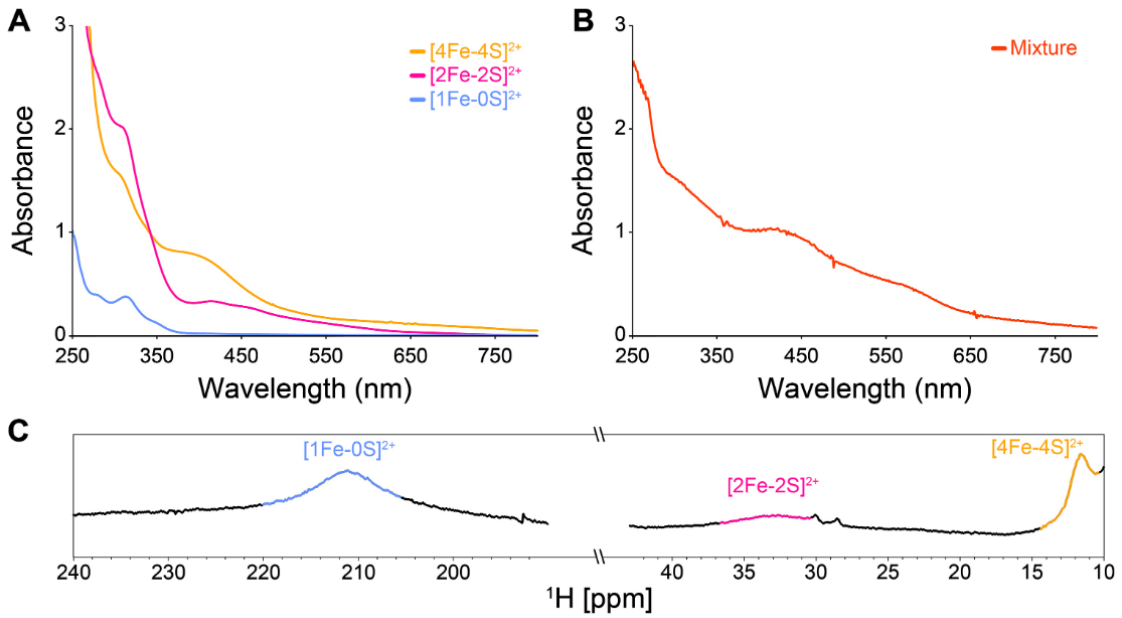

Figure 1.10 - Spectra of aqueous mixtures of iron–sulfur clusters stabilized by glutathione (γECG). Spectra are of $[1Fe\text{-}0S]^{2+}$, $[2Fe\text{-}2S]^{2+}$, and $[4Fe\text{-}4S]^{2+}$ clusters (**A**). UV–vis spectrum of 40 mM glutathione, 0.5 mM Na_2S, and 0.5 mM $FeCl_3$ at pH 8.6, which contained a mixture of types of iron–sulfur clusters (**B**). Paramagnetic 1H NMR spectrum of a similar sample revealed the presence of $[1Fe\text{-}0S]^{2+}$, $[2Fe\text{-}2S]^{2+}$, and $[4Fe\text{-}4S]^{2+}$ clusters at 210 ppm (blue), 33 ppm (red), and 11.8 ppm (yellow), respectively (**C**). (Adapted from Valer et al., 2021)

The environment surrounding an iron–sulfur cluster within a protein is asymmetric, which can be exploited by circular dichroism (CD) spectroscopy to gain some insight into the type of cluster present. CD spectra to study iron-sulfur cluster are collected in the near-UV and visible regions.[73] The $[1Fe\text{-}0S]^{3+}$ of rubredoxin also has a positive band near 450 nm but a negative band near 550 nm.[74] CD spectra of $[2Fe\text{-}2S]^{2+}$ and $[4Fe\text{-}4S]^{2+}$clusters tend to both show two clear positive bands near 450 nm and 550 nm.[75] However, given that CD spectra are greatly impacted by geometry and symmetry, spectra can look different in wavelength and intensity for the same cluster type coordinated to different proteins.

Resonance Raman (RR) spectroscopy is a powerful tool for the identification and characterization of metal sites and the elucidation of structure–function relationships of metalloproteins and metallopeptides. RR spectroscopy is particularly insightful when investigating iron–sulfur clusters that are diamagnetic and thus EPR silent, such as $[1Fe\text{-}1S]^{3+}$, $[2Fe\text{-}2S]^{2+}$ and $[4Fe\text{-}4S]^{2+}$ clusters.[76] Incident light excites the Fe ← S charge transfer electronic transitions (LMCT).

For this reason, the resulting spectra are sensitive to the configuration, symmetry, nature of ligands, and ultimately the type of iron–sulfur cluster present.[77]

Unlike many of the techniques used to interrogate iron–sulfur clusters, paramagnetic nuclear magnetic resonance (NMR) can typically observe iron–sulfur clusters at all of their oxidation states. Unpaired electrons possess a magnetic moment ~600 fold greater than the magnetic moment of a proton, which significantly influences the surrounding nuclei, with broadening observed with a dependence from the distance of the metal center. Depending on the rate of electronic relaxation of the paramagnetic center, nearby nuclei may either be quenched or shifted, with the latter giving rise to what is known as hyperfine shifted resonances. In the case of iron–sulfur clusters, the presence of the paramagnetic core considerably influences the resonances of surrounding nuclei, particularly the protons of ligating cysteines (e.g. Cys H_α and Cys H_β) (Figure 1.10C).

The presence of unpaired electrons within an iron–sulfur cluster makes EPR spectroscopy perfectly suitable for the interrogation of the cluster. Just as for NMR spectroscopy, electron paramagnetic resonance (EPR) spectroscopy exploits the presence of a magnetic moment (electronic, in this case) immersed in an external magnetic field, that interacts with an electromagnetic frequency. However, unlike paramagnetic NMR spectroscopy, only iron–sulfur clusters with an S integer (S = 1, 2, … n) or half-integer are EPR active. Anti-ferromagnetic spin coupling in $[2Fe\text{-}2S]^{2+}$ and $[4Fe\text{-}4S]^{2+}$ make these clusters EPR silent (S = 0).

Mössbauer spectroscopy is extremely useful in characterizing iron–sulfur clusters. This technique can identify every type of iron–sulfur cluster and every iron center present within the sample in a way that reveals their spin state, oxidation state, and electronic structure and gives insight into the coordination sphere. Similar to EPR and paramagnetic NMR spectroscopies, the information obtained from Mössbauer spectroscopy is due in large part to hyperfine interactions, i.e., interactions between the nucleus and nearby electrons.[78] Thus, this technique can probe very accurately the electronic environment around the iron nuclei. While Mössbauer can be considered one of the best technique to study iron-sulfur complexes it has its limitations. Limitations of the technique include the need for ^{57}Fe labelled samples to absorb the γ-rays emitted from the ^{57}Co source, and the need to either work with frozen samples at low temperatures or with precipitated complexes at higher temperature.[79]

The reduction potential of iron ions are strongly affected by solvent accessibility, the charges of neighboring residues, and the ligating residues. Therefore, reduction potentials can be correlated to some extent with the type of iron–sulfur cluster (Figure 1.11), but corroborating evidence is typically needed. Iron-sulfur clusters' redox potential decrease as the number of iron in the complex increase ([1Fe-0S]>[2Fe-2S]>[4Fe-4S]).

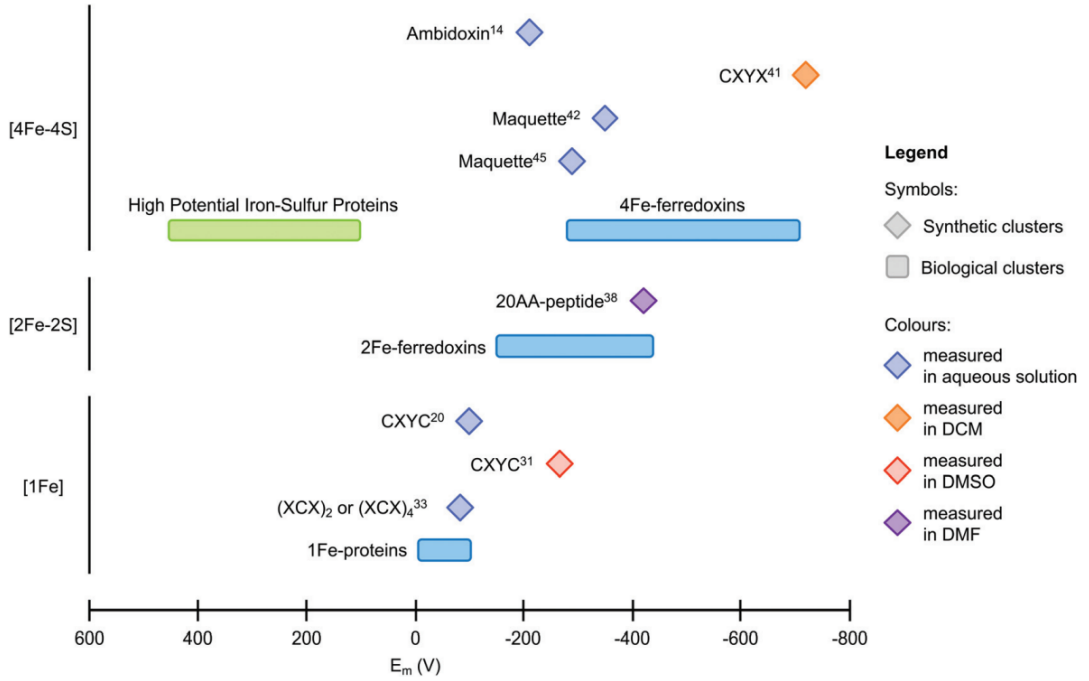

Figure 1.11 - Comparison between redox midpoint potentials of different iron–sulfur systems. (adapted from Bonfio, 2021)

1.11 References

1. Wetherill, G. W. Formation of the Earth. *Annu. Rev. Earth Planet. Sci.* **18**, 205–256 (1990).
2. Oschmann, W., Grasshof, M. & Gudo, M. The Early Evolution of the Planet Earth and the Origin of Life. *Constr. Morphol. Evol.* (1991) doi:10.1007/978-3-642-76156-0.
3. Kasting, J. F. . Earth's Early Atmosphere. *Adv. Sci.* **259**, 920–926 (2008).
4. Anbar, A. D. Elements and Evolution. *Science (80-.).* **322**, 1481–1483 (2008).
5. Sutherland, J. D. The Origin of Life - Out of the Blue. *Angew. Chemie - Int. Ed.* **55**, 104–121 (2016).
6. Ohtomo, Y., Kakegawa, T., Ishida, A., Nagase, T. & Rosing, M. T. Evidence for biogenic graphite in early archaean isua metasedimentary rocks. *Nat. Geosci.* **7**, 25–28 (2014).
7. Bell, E. A., Boehnke, P., Harrison, T. M. & Mao, W. L. Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 14518–14521 (2015).
8. Javaux, E. J. Challenges in evidencing the earliest traces of life. *Nature* **572**, 451–460 (2019).
9. Hopkins, F. G. The centenary of Wöhler's synthesis of urea (1828–1928). *Biochem. J.* **22**, 1341–1348 (1928).

10. Levine, R. & Evers, C. The Slow Death of Spontaneous Generation. *North Carolina State Univ.* 7–8 (2000).

11. Peretó, J., Bada, J. L. & Lazcano, A. Charles darwin and the origin of life. *Orig. Life Evol. Biosph.* **39**, 395–406 (2009).

12. Oparin, A. I. The origin of life. *Macmillan Co.* 1–234 (1938) doi:10.4324/9781315133638.

13. Haldane, J. B. S. The Origin of Life. *Ration. Annu.* **148**, 3–10 (1929).

14. Urey, H. C. On the Early Chemical History of the Earth and the Origin of Life. *Proc. Natl. Acad. Sci.* **38**, 351–363 (1952).

15. Miller, S. L. the Production of Amino Acids Under Possibly Primitive Earth Conditions. *Science (80-.).* **117**, 528–529 (1953).

16. Johnson, A. P. *et al.* The Miller volcanic spark discharge experiment. *Science (80-.).* **322**, 404 (2008).

17. Parker, E. T. *et al.* Primordial synthesis of amines and amino acids in a 1958 Miller H 2S-rich spark discharge experiment. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 5526–5531 (2011).

18. Wu, L. F. & Sutherland, J. D. Provisioning the origin and early evolution of life. *Emerg. Top. Life Sci.* **5**, 459–468 (2021).

19. Gibard, C., Bhowmik, S., Karki, M., Kim, E. K. & Krishnamurthy, R. Phosphorylation, oligomerization and self-assembly in water under potential prebiotic conditions. *Nat. Chem.* **10**, 212–217 (2018).

20. Gibard, C. *et al.* Geochemical Sources and Availability of Amidophosphates on the Early Earth. *Angew. Chemie - Int. Ed.* **58**, 8151–8155 (2019).

21. Toparlak, D., Karki, M., Egas Ortuno, V., Krishnamurthy, R. & Mansy, S. S. Cyclophospholipids Increase Protocellular Stability to Metal Ions. *Small* **16**, 1–8 (2020).

22. Hug, L. A. *et al.* A new view of the tree of life. *Nat. Microbiol.* **1**, 1–6 (2016).

23. Fraser, C. M. *et al.* The minimal gene complement of Mycoplasma genitalium. *Science (80-.).* **270**, 397–403 (1995).

24. Morange, M. Some considerations on the nature of LUCA, and the nature of life. *Res. Microbiol.* **162**, 5–9 (2011).

25. Crick, F. H. C. The origin of the genetic code. *J. Mol. Biol.* **38**, 367–379 (1968).

26. Orgel, L. E. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* **39**, 99–123 (2004).

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

28. Gilbert, W. Origin of life: The RNA world. *Nature* vol. 319 618 (1986).

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

30. 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).

31. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science (80-.).* **254**, 1497–1500 (1991).

32. Schoning, K.-U. *et al.* Chemical etiology of nucleic acid structure. *Science (80-.).* **290**, 1347–1352 (2000).

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

34. 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).

35. Maurel, M. C. & Leclerc, F. From foundation stones to life: Concepts and results. *Elements* **12**, 407–412 (2016).

36. 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).

37. Mielke, R. E. *et al.* Iron-sulfide-bearing chimneys as potential catalytic energy traps at life's emergence. *Astrobiology* **11**, 933–950 (2011).

38. Weiss, M. C. *et al.* The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **1**, 1–8 (2016).

39. 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).

40. Morowitz, H. J., Kostelnik, J. D., Yang, J. & Cody, G. D. The origin of intermediary metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7704–7708 (2000).

41. Preiner, M. *et al.* The future of origin of life research: Bridging decades-old divisions. *Life* **10**, 1–25 (2020).

42. Ikehara, K. Evolutionary steps in the emergence of life deduced from the bottom-up approach and GADV hypothesis (Top-down approach). *Life* **6**, 1–15 (2016).

43. Plankensteiner, K., Reiner, H. & Rode, B. Prebiotic Chemistry: The Amino Acid and Peptide World. *Curr. Org. Chem.* **9**, 1107–1114 (2005).

44. Caetano-Anollés, G., Wang, M., Caetano-Anollés, D. & Mittenthal, J. E. The origin, evolution and structure of the protein world. *Biochem. J.* **417**, 621–637 (2009).

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

46. Tessera, M. Origin of evolution versus origin of life: A shift of paradigm. *Int. J. Mol. Sci.* **12**, 3445–3458 (2011).

47. Sharov, A. A. Coenzyme world model of the origin of life. *BioSystems* **144**, 8–17 (2016).

48. Koonin, E. V., Senkevich, T. G. & Dolja, V. V. The ancient virus world and evolution of cells. *Biol. Direct* **1**, 1–27 (2006).

49. Lanier, K. A. & Williams, L. D. The Origin of Life: Models and Data. *J. Mol. Evol.* **84**, 85–92 (2017).

50. Muchowska, K. B. *et al.* Metals promote sequences of the reverse Krebs cycle. *Nat. Ecol. Evol.* **1**, 1716–1721 (2017).

51. Varma, S. J., Muchowska, K. B., Chatelain, P. & Moran, J. Native iron reduces CO2 to intermediates and end-products of the acetyl-CoA pathway. *Nat. Ecol. Evol.* **2**, 1019–1024 (2018).

52. Johnson, M. K. & Smith, A. D. Iron-Sulfur Proteins. in *Encyclopedia of Inorganic and Bioinorganic Chemistry* (John Wiley & Sons, Ltd, 2011). doi:10.1002/9781119951438.eibc0109.

53. Belmonte, L. & Mansy, S. S. Patterns of Ligands Coordinated to Metallocofactors Extracted from the Protein Data Bank. *J. Chem. Inf. Model.* **57**, 3162–3171 (2017).

54. Link, T. A. The Structures of Rieske and Rieske-Type Proteins. in *Advances in Inorganic Chemistry* vol. 47 83–157 (1999).

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

56. Meyer, J. Iron–sulfur protein folds, iron–sulfur chemistry, and evolution. *JBIC J. Biol. Inorg. Chem.* **13**, 157–170 (2008).

57. Jones, K., Gomes, C. M., Huber, H., Teixeira, M. & Wittung-Stafshede, P. Formation of a linear [3Fe-4S] cluster in a seven-iron ferredoxin triggered by polypeptide unfolding. *J. Biol. Inorg. Chem.* **7**, 357–362 (2002).

58. Russell, M. J. & Martin, W. The rocky roots of the acetyl-CoA pathway. *Trends Biochem. Sci.* **29**, 358–363 (2004).

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

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

61. Bonfio, C. *et al.* UV-light-driven prebiotic synthesis of iron–sulfur clusters. *Nat. Chem.* **9**, 1229–1234 (2017).

62. Bonfio, C. *et al.* Prebiotic iron–sulfur peptide catalysts generate a pH gradient across model membranes of late protocells. *Nat. Catal.* **1**, 616–623 (2018).

63. Scintilla, S. *et al.* Duplications of an iron–sulphur tripeptide leads to the formation of a protoferredoxin. *Chem. Commun.* **52**, 13456–13459 (2016).

64. 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).

65. 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).

66. Kim, J. D. *et al.* Minimal Heterochiral de Novo Designed 4Fe–4S Binding Peptide Capable of Robust Electron Transfer. *J. Am. Chem. Soc.* **140**, 11210–11213 (2018).

67. Galambas, A. *et al.* Radical S-adenosylmethionine maquette chemistry: Cx3Cx2C peptide coordinated redox active [4Fe–4S] clusters. *JBIC J. Biol. Inorg. Chem.* **24**, 793–807 (2019).

68. Jordan, S. F. *et al.* Spontaneous assembly of redox-active iron-sulfur clusters at low concentrations of cysteine. *Nat. Commun.* **12**, 1–14 (2021).

69. Valer, L. *et al.* Methods to identify and characterize iron-sulfur oligopeptides in water. *Can. J. Chem.* (2022) doi:10.1139/cjc-2021-0237.

70. Sugiura, Y., Ishizu, K. & Kimura, T. A rubredoxin-like mononuclear FeS4 derivative of adrenal iron-sulfur protein (adrenodoxin). *Biochem. Biophys. Res. Commun.* **60**, 334–340 (1974).

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

72. Devanathan, T., Akagi, J. M., Hersh, R. T. & Himes, R. H. Ferredoxin from Two Thermophilic Clostridia. *J. Biol. Chem.* **244**, 2846–2853 (1969).

73. Stephens, P. J. *et al.* Circular Dichroism and Magnetic Circular Dichroism of Iron-Sulfur Proteins. *Biochemistry* **17**, 4770–4778 (1978).

74. Dos Santos, P. C. *Fe-S Proteins*. vol. 2353 (Springer US, 2021).

75. Webert, H. *et al.* Functional reconstitution of mitochondrial Fe/S cluster synthesis on Isu1 reveals the involvement of ferredoxin. *Nat. Commun.* **5**, (2014).

76. Todorovic, S. & Teixeira, M. Resonance Raman spectroscopy of Fe–S proteins and their redox properties. *J. Biol. Inorg. Chem.* **23**, 647–661 (2018).

77. Spiro, T. G. & Czernuszewicz, R. S. Resonance Raman Spectroscopy of Metalloproteins. *Methods Enzymol.* **246**, 416–460 (1995).

78. Vrajmasu, V. V., Bominaar, E. L., Meyer, J. & Münck, E. Mössbauer Study of Reduced Rubredoxin As Purified and in Whole Cells. Structural Correlation Analysis of Spin Hamiltonian Parameters. *Inorg. Chem.* **41**, 6358–6371 (2002).

79. Clémancey, M., Geneviève Blondin, Latour, J.-M. & Garcia-Serres, R. *Metalloproteins*. *Methods in Molecular Biology* vol. 1122 (Humana Press, 2014).

Chapter 2 – Histidine ligated Iron-Sulfur clusters in proteins

2.1 Introduction

Although iron-sulfur clusters are primarily coordinated by four cysteinyl ligands,[1] there are several cases where the cluster is not completely ligated by cysteines. For example, mononuclear [1Fe-0S], [2Fe-2S], and [4Fe-4S] clusters can be ligated by the sidechain of an aspartic acid, glutamic acid, or hisdine.[2] In the early 1960s, Beinert and Sands reported an Electron Paramagnetic Resonance (EPR) spectroscopic signal around g = 1.94 in fragments of the mitochondrial respiratory chain.[3] A few years later, in 1964, Rieske and co-workers discovered that the origin of that signal was due to the presence of a particular [2Fe-2S] cluster with unique spectroscopic properties.[4] After this discovery many scientist started to report similar findings. For example, an unknown iron-sulfur cluster was detected in complex III of the mitochondria membranes or another iron-sulfur cluster discovered in the bf_6 complex of photosynthetic electron transport chain and in membranes of archaebacteria.[5] The proteins that contain such clusters took later the name of Rieske protein, from the name of the first researcher that discovered them.[6] In biology, Rieske iron-sulfur clusters are strictly related to cytochrome bc complexes (bc_1 in mitochondria, bf_6 in thylakoids) located in respiratory and photosynthetic electron-transport chain.[7] These complexes are involved in the transduction of energy by oxidizing quinol and so coupling electron transfer with the formation of a proton gradient across the membranes. This electrochemical gradient subsequently drives the synthesis of ATP.

The distinctive feature of Rieske clusters consists of the coordination environment of the [2Fe-2S] cluster core: one iron atom is coordinated by the thiols of two cysteines, and the other iron atom is coordinated by two δ nitrogens of the imidazole rings of two histidine residues. This was confirmed by sequencing, three-dimensional structure, and magnetic spectroscopy, both NMR than EPR. The unusual coordination environment on the active site leads to a distorted tetrahedral iron center, different from the Ferredoxin-like [2Fe-2S] clusters.[8,9]

Rieske and full-cysteine coordinating [2Fe-2S] clusters have different coordination spheres and hence different folds, caused by the chemical asymmetry of the two distinct set of Fe_S-S and Fe_N-S bond lengths.[8] Nevertheless, the topology of the Rieske cluster binding subdomain shows a stunningly relationship with that of rubredoxin and zinc-ribbon motif typically found in RNA polymerases or TFII transcription factors.[10–12]

In addition, Rieske and Rieske-type proteins have two conserved cysteines that form a disulfide bridge, believed to increase the stability of the prosthetic group

and, above all, the redox potential of the cluster.[13–15] Proteins lacking this functional group have been observed in lower levels of ubiquinol oxidation.[16,17]

Other unique features of the Rieske clusters regard the electrochemical behavior. Indeed, reduction potentials of iron-sulfur clusters are strongly influenced by their ligands. The substitution of a negatively charged cysteine ligand with a neutral histidine typically increases the reduction potential of the cluster. Another unique feature of the Rieske proteins regards a pH-dependent redox potential. This is due to the protonation of nitrogen ε of the two coordinating histidines of Fe, which is strictly related to cluster reduction.[18,19]

To better understand the kind of coordination provided by histidine, we have selected three different proteins and designed several mutations (Table 2.1). The first protein we have chosen is a rubredoxin from *Clostridium pasteurianum*. This is a small protein (6 kDa) contains a motif of four cysteine residues with the spacing $CX_2CX_{29}CX_2C$ that holds a mononuclear [1Fe-0S] center. For this reason rubredoxin is considered the simples iron-sulfur protein.[20] This rubredoxin is presumed to be an electron transport protein, cycling between ferric and ferrous forms, as this role has been demonstrated in both the sulfate-reducing bacterium *Desulfovibrio gigas*[21] and the aerobe *Pseudomonas oleovorans*.[22] The mutants of this [1Fe-0S] protein were obtained by mutating one or two cysteine into histidine (C42H, C9H C42H).

Human ferredoxin and the soluble respiratory-Type Rieske protein from *Thermus thermophilus* were chosen as examples of [2Fe-2S] clusters.[23] Human mitochondrial ferredoxin 1 (FDX1) is present in the matrix of human mitochondria. This ferredoxin is a small iron-sulfur protein (12 kDa) that can accept and carry a single electron and functions as an electron transfer protein in the mitochondrial cytochrome P450 systems. The single [2Fe-2S] cluster is coordinated by a $CX_5CX_2CX_{36}C$. The Rieske protein from *Thermus thermophilus* is the bigger of the three chosen weighting around 20 kDa. This protein acts inside the cell as a component of the membrane-bound bc_1 respiratory complex and contains a [2Fe-2S] cluster coordinated by two cysteine and two histidine through the motif $CXHX_{16}CX_2H$. The mutants of the Rieske protein were obtained, as in the case of the rubredoxin, by the substitution of one or two cysteine residues with histidine residues. Since Rieske proteins contain a disulfide bond between C93 and C109, it was possible that the recombinantly expressed constructs in which ligating histidines were substituted led to the recruitment of C93 and/or C109 for the stabilization of the cluster. To confirm that the [2Fe-2S] cluster was coordinated by the original positions of the ligands, the cysteines involved in the formation of the disulfide bond were substituted with Ala. Both the rubredoxins and the Rieske proteins were tagged with a maltose binding protein (MBP) (42 kDa) to facilitate the purification. After the expression and purification UV-vis and CD spectroscopy was carried out to verify the successful formation of the iron-sulfur cluster. Based on those sequences of the purified protein later we designed several peptides. The

collected spectrographic date were employed as references for the synthetic iron-sulfur clusters coordinated by peptides discussed in Chapter 3.

Name	Protein sequence
WT Rubredoxin ROI	T**C**TV**C**G...V**C**PL**C**G
C42H Rubredoxin ROI	T**C**TV**C**G...V**C**PL**H**G
C9H C42H Rubredoxin ROI	T**C**TV**H**G...V**C**PL**H**G
WT Ferredoxin ROI	A**C**EGTLA**C**ST**C**H...G**C**Q
WT Rieske ROI	V**C**T**H**LG<u>C</u>IVSQWVADEEAAL**C**P<u>C</u>**H**G
H90C Rieske ROI	V**C**T**C**LG<u>C</u>IVSQWVADEEAAL**C**P<u>C</u>**H**G
H90C H107C ROI	V**C**T**C**LG<u>C</u>IVSQWVADEEAAL**C**P<u>C</u>**C**G
C93A C109A ROI	V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G
H90C C93A C109A ROI	V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G
H90C C93A H107C C109A ROI	V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G

Table 2.1 - ROI indicates the region of interest of the protein. Positions in bold indicate coordinating residues (red cysteine, blue histidine). Underlined residues indicate the position of the disulfide bridge (green cysteine). The "..." of rubredoxin indicates 27 amino acids, and 34 amino acids for ferredoxin. (The full gene sequences for the proteins are reported in Appendix 1)

2.2 Results

2.2.1 Spectroscopic characterization of Rubredoxin and mutants

The rubredoxin and its mutants were recombinantly expressed (Figure S1.1A) and were purified through affinity chromatography with amylose resin. The brown color of recombinant rubredoxin preparations suggested that the iron center had been assembled and incorporated into the protein *in vivo*, without the need for an *in vitro* reconstitution of the cluster. This was indeed confirmed by optical spectroscopy. The UV-vis spectra of oxidized rubredoxins display absorption peaks at 380 nm, 490 nm, and a shoulder at 580 nm.[24] The purified wild type rubredoxin showed absorption maxima consistent with previously reported spectra for this type of proteins (Figure 2.1).[25,26]

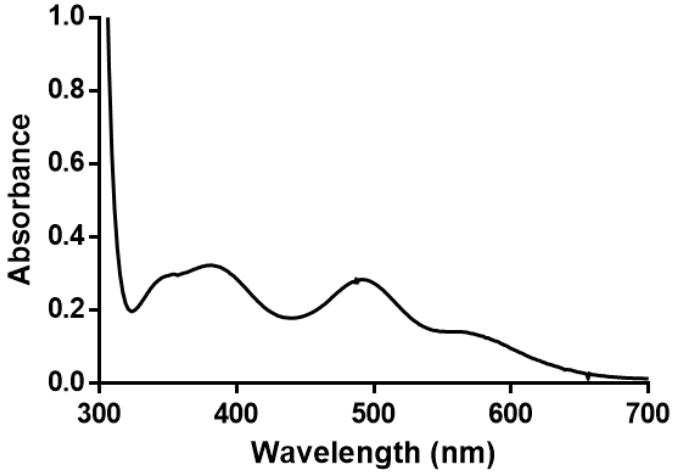

Figure 2.1 – UV-vis spectra absorption of Wild Type rubredoxin

The circular dichroism (CD) spectra were also collected. The CD spectrum of the wild type rubredoxin showed a positive band near 450 nm but a negative band near 500 nm (Figure 2.2). The spectrum was in accordance with previously reported ones, showing a large Cotton effect.[27] This is due to the high asymmetric arrangement around the iron ion.

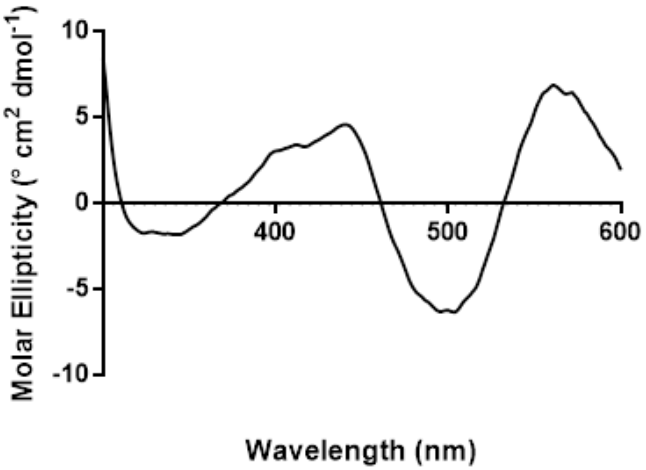

Figure 2.2 – Circular dichroism spectra of oxidized wild type rubredoxin.

However, the purified fraction of both C42H rubredoxin and C9H C42H rubredoxin did not show the intense red/brown color showed by the wild type. After the protein was concentrated, the UV-vis adsorption spectra were acquired. The spectra were greatly diminished in intensity. In particular, the mutant C42H did not seem to show any feature in the spectra (Figure 2.3A). The mutant C9H C42H spectrum showed a shoulder around 410 nm (Figure 2.3B).

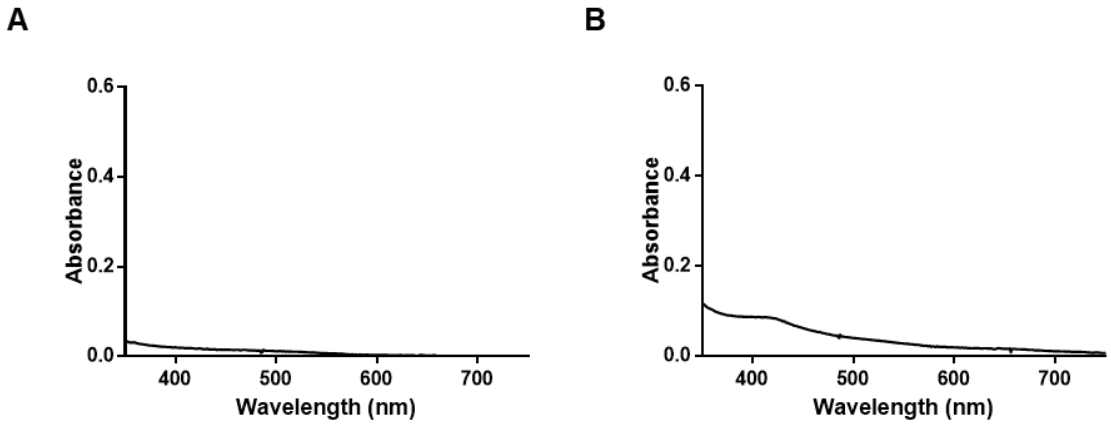

Figure 2.3 – UV-vis spectra absorption of Rubredoxin mutants. A. C42H rubredoxin B. C9H C42H rubredoxin

CD spectroscopy was also performed on the two his-containing mutants. The CD spectra of both the C42H and C9H C42H mutants were featureless (Figure 2.4), which was not surprising considering the low intensity of absorption of the UV-vis spectra. Also, if water molecules were present in the coordination sphere of the metal ion, they could alter the asymmetry of the environment reducing the Cotton effect.

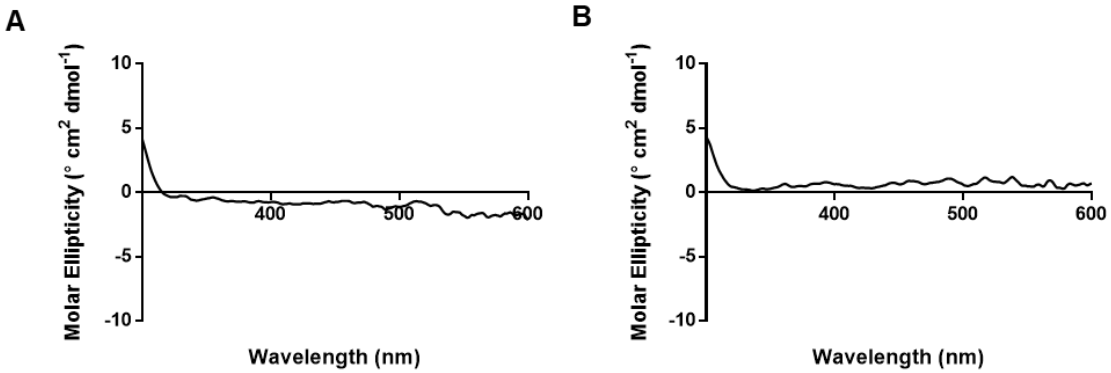

Figure 2.4 – Circular dichroism of spectra of oxidized rubredoxin mutants. . A. C42H rubredoxin B. C9H C42H rubredoxin.

2.2.2 Spectroscopic characterization of Ferredoxin

Human ferredoxin (12 kDa) was successfully expressed (Figure S1.1B) and was isolated by two successive chromatographic techniques, an ion exchange chromatography performed on a DEAE column followed by a size exclusion. After the size exclusion the fractions containing the proteins were combined and concentrated. And subsequently, the UV-vis spectrum of the oxidized protein was acquired to confirm the presence of the iron-sulfur cluster.

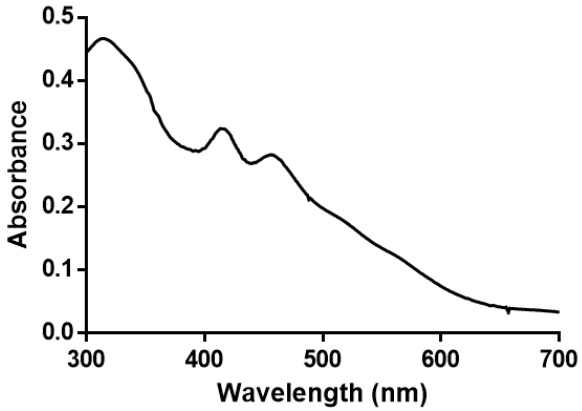

Figure 2.5 – UV-vis absorption spectra of oxidized wild type human ferredoxin (20 mM Tris-HCl, pH 7.6)

The UV-vis spectrum showed two characteristic peak of a $[2Fe\text{-}2S]^{2+}$ cluster at 420 and 450 nm (Figure 2.5). Such peaks are one of the most characteristic spectroscopic features of [2Fe-2S] cluster,[24,28] confirming that also in this case the cluster was successfully assembled *in vivo* by *E. coli*.

The CD spectrum of the oxidized ferredoxin was also collected. As in the case of the wild type rubredoxin, the ferredoxin spectra appears well structured with a huge positive contribution around 450 nm (Figure 2.6). The CD spectra features were similar to the previously reported ones.[29] As in the rubredoxin case, the CD spectra of ferredoxin indicated that the environment around the cluster is high asymmetric.

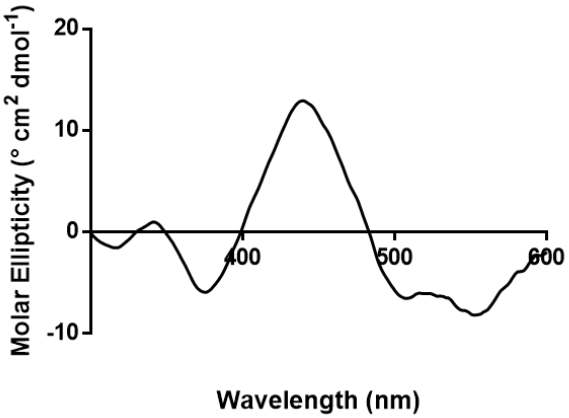

Figure 2.6 – Circular dichroism spectra of oxidized wild type human ferredoxin (20 mM Tris-HCl, pH 7.6)

2.2.3 Spectroscopic characterization of Rieske Protein and mutants

As for the previous iron-sulfur proteins, the soluble Rieske protein from *Thermus Thermophilus*, was successfully expressed (Figure S1.2) and

subsequently purified using an amylose resin column. The UV-vis spectra of both the wild type and the C93A C109A mutant were consistent with past reports other oxidized Rieske cluster, with peaks at 450 nm and 580 nm (Figure 2.7A).[30,31] The removal of the cysteines involved in the formation of the disulfide bridge did not seem to impair the ability of the protein to successfully coordinate a $[2Fe\text{-}2S]^{2+}$ cluster (Figure 2.7B). After reduction with dithionite, the spectra shows in both cases two peaks at 420 nm and 520 nm.[32]

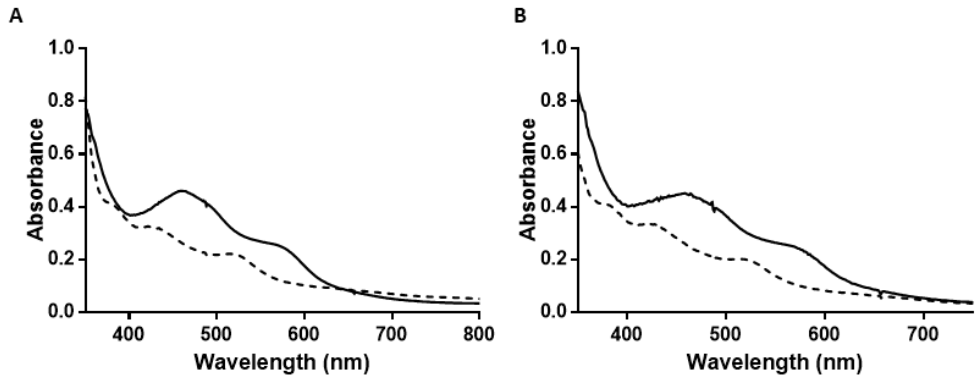

Figure 2.7 - UV-vis absorption spectra of WT *Thermus thermophilus* Rieske. Spectra of oxidized (solid line) and dithionite reduced (dashed line) species are shown. A. Wild Type Rieske. B. C93A C109A Rieske (no disulfide bridge).

We also performed CD spectroscopy on these proteins (Figure 2.8). The main feature of the spectra was a negative contribution around 390 nm, which corresponds with previously reported spectra.[18] The spectra of both the WT and its mutants without the disulfide bridge were almost identical confirming that even without the bridge the geometry of the environment around the cluster is unaltered.

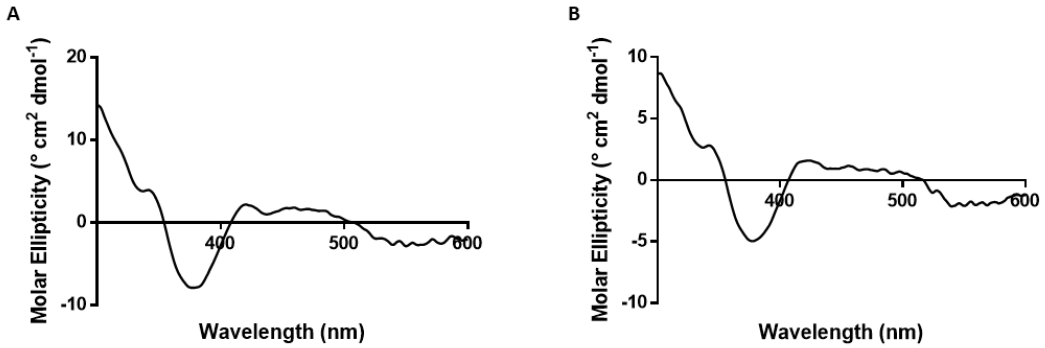

Figure 2.8 – CD spectra of WT *Thermus thermophilus* Rieske. Spectra of oxidized (solid line) and dithionite reduced (dashed line) species are shown. A. Wild Type Rieske. B. C93A C109A Rieske (no disulfide bridge).

The oxidized Rieske mutants containing 1 His, 3 Cys ligands, both with and without the disulfide bridge, had a prominent peak at 470 nm (Figure 2.9). This type of coordination is typical of a class of mitochondrial protein called

MitoNEET.[33] The UV-vis spectra of the two purified MitoNEET-like mutants are similar to previously reported $[2Fe\text{-}2S]^{2+}$ MitoNEET absorbance spectra. However the mutant without the disulfide bridge displayed a lower intensity. Also, upon reduction with dithionite the spectra were consistent with previously reported ones. It is worth noting that the UV-vis absorption spectra of MitoNEET [2Fe-2S] cluster were quite similar to the one of Rieske protein with the main difference displayed by the disappearance of the prominent shoulder at 580 nm.[34]

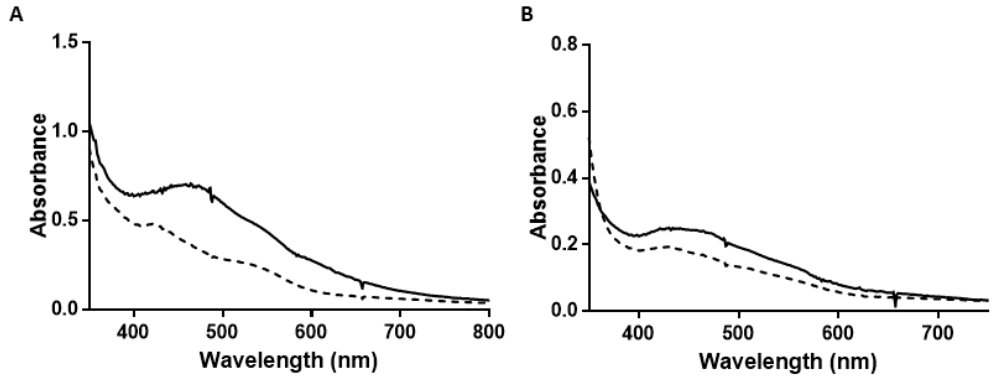

Figure 2.9 - UV-vis absorption spectra mutant *Thermus thermophilus* Rieske. Spectra of oxidized (solid line) and dithionite reduced (dashed line) species are shown. A. H107C Rieske B. C93A H107C C109A Rieske (no disulfide bridge).

The CD spectroscopy of these two Rieske mutants showed less feature compared to the WT (Figure 2.10). In particular the MitoNEET-like mutants with the disulfide bridge showed around 470 nm and 580 nm two negative peak, which are displayed also by other $[2Fe\text{-}2S]^{2+}$ MitoNEET protein.[35] On the other hand the three cysteine one histidine mutant without the disulfide bridge showed quite different pattern especially in the region below 400 nm. This indicates that the geometry of the environment around the cluster is slightly different between the two.

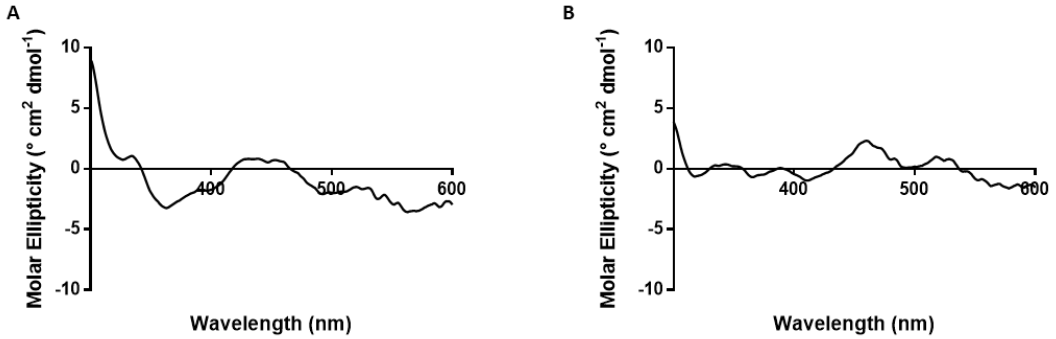

Figure 2.10 – CD spectra mutant Thermus thermophilus Rieske. Spectra of oxidized (solid line) and dithionite reduced (dashed line) species are shown. A. H107C Rieske B. C93A H107C C109A Rieske (no disulfide bridge).

Finally, the last mutants contained no ligating histidines (H90C H107C and H90C C93A H107C C109A). The UV-vis spectrum of the mutant with the disulfide bridge showed clearly two peaks at 420 nm and 450 nm (Figure 2.11), which are the main features of a full-cysteine coordinated $[2Fe\text{-}2S]^{2+}$ cluster.[24,36] The mutant with the disulfide bridge showed the same two peaks; however, they appeared less pronounced, similar to the one displayed by $[2Fe\text{-}2S]^{2+}$ cluster coordinates by peptides, possibly meaning that this mutant is less effective in stabilizing the cluster. Dithionite reduction decrease the intensity of the spectra leaving as the main feature the peak around 420 nm.

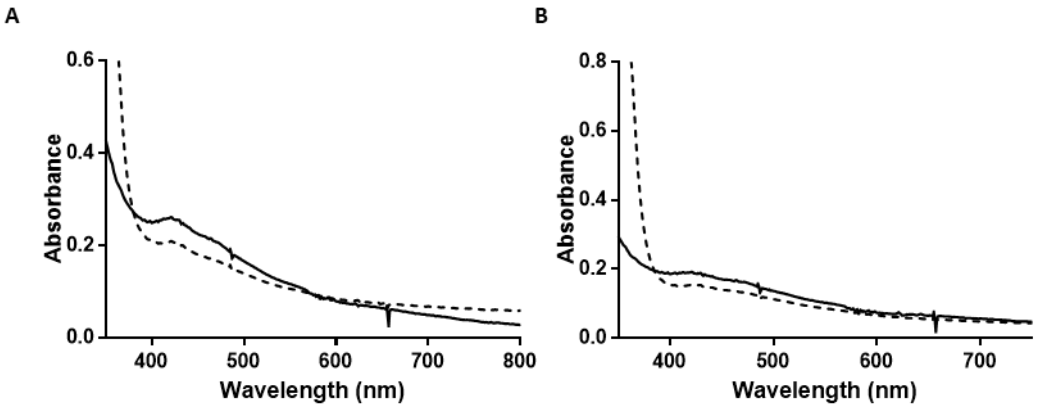

Figure 2.11 - UV-vis absorption spectra of mutant *Thermus thermophilus* Rieske. Spectra of oxidized (solid line) and dithionite reduced (dashed line) species are shown. A. H90C H107C Rieske B. H90C C93A H107C C109A Rieske (no disulfide bridge).

CD was performed on these mutants; however, the collected spectra did not show any prominent features and looks quite similar between the version with the disulfide bridge and without (Figure 2.12). The lack of any peak could mean that the environment around the cluster might not be structured or might have increased symmetry compared to the wild type.

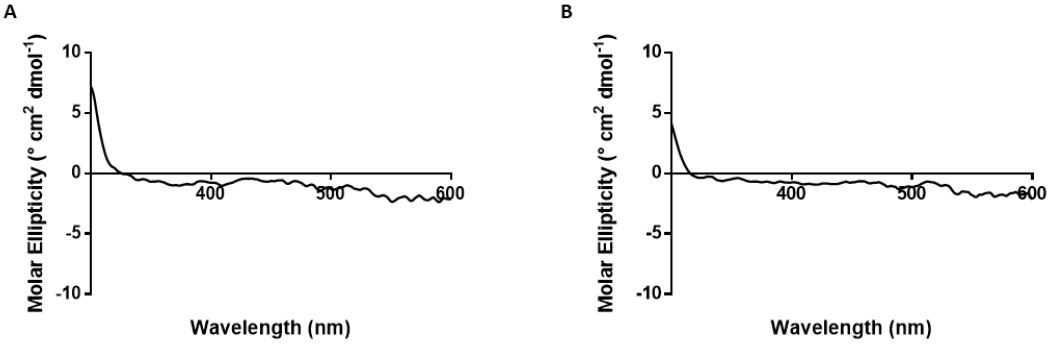

Figure 2.12 - CD mutant Thermus thermophilus Rieske. Spectra of oxidized (solid line) and dithionite reduced (dashed line) species are shown. A. H90C H107C Rieske B. H90C C93A H107C C109A Rieske (no disulfide bridge).

2.3 Discussion

UV–vis spectroscopy is the easiest and most accessible technique to characterize iron–sulfur proteins. Generally, iron–sulfur clusters absorb in the 300–800 nm range, reflecting the ligand-to-metal charge transfer (LMCT) transitions. the bands appearing in this region correspond to Fe(3d) ← μ-S-Cys(n) electronic transition and in some cases can be quite diagnostic of the type of clusters present in solution. CD is more difficult to interpret, given that spectra are greatly impacted by geometry and symmetry. Spectra can look different both in wavelength and intensity for the same cluster type coordinated to different proteins.

For such [1Fe-0S] complexes, the UV–vis spectra of oxidized $[1Fe\text{-}0S]^{3+}$ rubredoxin show several peaks between 350 and 380 nm, with a maximum absorbance at 490 nm and a broad band centered near 560 nm. This is consistent with the spectra obtained for the purified rubredoxin. However, the two mutants (C42H and C9H, C42H) showed spectra without many features and of diminished intensity, this might be due to an incomplete coordination. This is possible when the coordination is achieved exploiting the oxygen atom of one or more water molecules instead of the nitrogen from the histidine residues. Thus, these results cannot confirm the nitrogen coordination of a [1Fe-0S] center.

All-cysteine coordinated $[2Fe\text{-}2S]^{2+}$ clusters give the most distinctive UV–vis spectra with bands at approximately 420 and 450 nm. The UV-vis spectra collected for the human ferredoxin shows clearly the two characteristic peaks of such cluster, confirming the successful coordination of a [2Fe-2S] cluster. Also, the CD shows the positive peak at 450 nm typical of such cluster.

Both the Wild Type Rieske protein show bands at 460 and 490 nm, which are in accordance with previously reported spectra. Also, the mutant constructs in which one of the histidine or both of the histidine were substituted with cysteine possessed spectra similar to MitoNEET and ferredoxin, respectively. These results confirmed that we successfully mutated a Rieske protein to incorporate a MitoNEET- and ferredoxin-like clusters. The mutations did not disrupt the biosynthetic route for the cluster formation. In fact, all the iron-sulfur cluster discussed here were assembled *in vivo* by *E. coli*, without the need for a further reconstitution *in vitro*.

Interestingly, with all the three types of coordination (Rieske-, MitoNEET- and ferredoxin-like) the UV-vis spectra seems unaltered by the removal of the bridge. This confirms previous studies, showing that the [2Fe-2S] cluster can be assembled even if the disulfide bridge that connects the two loops of the protein surrounding the cluster is removed. However, while the cluster stability is not impaired by this mutation, it has been shown that in such mutants significant structural changes occurred in in the Rieske protein environment around the [2Fe-2S] cluster, impairing its activity. The CD spectra, in the case of three and

four cysteine coordination, was less structured meaning that those mutations can alter severely the geometry of the cluster environment.

2.4 Material and methods

Materials

All the reagents were bought from Carlo Erba, VWR Chemicals and Sigma Aldrich and no further purification were performed. The pH of every solution was monitored with a Mettler Seven2Go pH meter. pH meter calibration was considered successful only if calibration percentage was at least 96%. pH meter calibration was performed every time the tool was used.

Constructs

All the constructs were synthesized by Genescript. Plasmids encoding rubredoxin and Rieske iron-sulfur proteins and their respective mutants were synthesized by Genscript and inserted into pMal-c4x (from NEB), generating a Maltose Binding Protein (MBP) tag. The plasmid encoding ferredoxin was synthesized by Genscript and inserted in pET-21b. Human ferredoxin was not tagged.

Protein Expression

Recombinant expression of the genes in pMal-c4x was with *Escherichia coli* NEBExpress. Recombinant expression of ferredoxin was with *E. coli* BL21(DE3) pLysS. For both, transformed cells were grown overnight at 37 °C in 10 mL Luria-Bertani (LB) supplemented with ampicillin (50 mg/mL). 1 mL of this starter culture was used to inoculate 1 LB, 50 mg/mL ampicillin in a 5 L flask and grown at 37 °C with shaking to an OD_{600} of 0.6. Expression was induced with 0.4 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated at room temperature for an additional 15 h. Cells were harvested by centrifugation at 4000 x g for 15 min. The cell pellet was resuspended in 40 mL of sonication buffer (20 mM Tris HCl, 200 mM NaCl, pH 7.6,) and lysed by sonication (20 cycles of 10 s pulses followed by 50 s incubation on ice). Cellular debris was removed by centrifugation at 15000 x g for 15 min. For the purification of the MBP tagged proteins, the lysate was loaded on an amylose column (NEB) and washed with 20 column volumes of amylose column buffer (20 mM Tris HCl, 200 mM NaCl, 0.1 mM EDTA, pH 7.6). The protein was eluted with 10 mM maltose in column buffer. The colored fractions were collected and concentrated using a Amicon Ultra centrifugal filter units (MWCO 10 kDa). Proteins were dialyzed against 20 mM Tris HCl, pH 7.6 overnight at 4 °C. Ferredoxin was purified in two-steps, as previously reported.29 The first step of the purification exploited an ion exchange column with DEAE resin (Bio-Rad) and washed with 20 column volumes of DEAE column buffer (50 mM Tris-HCl, pH 7.4). The proteins were eluted with a 90 mL gradient of 0-500 mM NaCl in

DEAE column buffer. The protein eluted at approximately 350-400 mM NaCl. The colored fractions were collected and concentrated with Amicon Ultra centrifugal filter units (MWCO 10 kDa). The second step of the purification relied on size-exclusion chromatography with a HiLoad 16/600 Superdex 75 pg column and a FPLC (Äkta pure, GE Healthcare). Subsequently, the protein-containing fractions were collected and concentrated with Amicon Ultra centrifugal filter units (MWCO 10 kDa).

UV-Vis absorption and Circular Dichroism spectroscopy

The UV-vis spectra of the samples were recorded inside a glovebox with a Genesys 150 UV-Vis spectrophotometer (ThermoFisher) with an integration time of 0.5 s and an interval of 1 nm. Circular Dichroism (CD) spectra were collected with a JASCO J-815 CD spectrometer, using a scan rate of 200 nm/min (spectral window 200-600 nm). For each CD spectra a minimum of 20 scans were collected. Spectra manager II (JASCO) software was used to analyze data.

2.5 References

1. Johnson, D. C., Dean, D. R., Smith, A. D. & Johnson, M. K. Structure, function, and formation of biological iron-Sulfur clusters. *Annu. Rev. Biochem.* **74**, 247–281 (2005).

2. Belmonte, L. & Mansy, S. S. Patterns of Ligands Coordinated to Metallocofactors Extracted from the Protein Data Bank. *J. Chem. Inf. Model.* **57**, 3162–3171 (2017).

3. Sands, R. H. & Beinert, H. Studies on mitochondria and submitochondrial particles by paramagnetic resonance (EPR) spectroscopy. *Biochem. Biophys. re* **3**, 47–52 (1960).

4. Rieske, J. S., MacLennan, D. H. & Coleman, R. Isolation and properties of an iron-protein from the (reduced coenzyme Q)-cytochrome C reductase complex of the respiratory chain. *Biochem. Biophys. Res. Commun.* **15**, 338–344 (1964).

5. Schneider, D. & Schmidt, C. L. Multiple Rieske proteins in prokaryotes: Where and why? *Biochim. Biophys. Acta - Bioenerg.* **1710**, 1–12 (2005).

6. Colbert, C. L., Couture, M. M. J., Eltis, L. D. & Bolin, J. T. A cluster exposed: Structure of the rieske ferredoxin from biphenyl dioxygenase and the redox properties of Rieske Fe-S proteins. *Structure* **8**, 1267–1278 (2000).

7. Molik, S., Karnauchov, I., Weidlich, C., Herrmann, R. G. & Klösgen, R. B. The Rieske Fe/S protein of the cytochrome b6/f complex in chloroplasts: Missing link in the evolution of protein transport pathways in chloroplasts? *J. Biol. Chem.* **276**, 42761–42766 (2001).

8. Ali, M. E. *et al.* The iron-sulfur core in Rieske proteins is not symmetric. *J. Biol. Inorg. Chem.* **19**, 1287–1293 (2014).

9. Iwata, S., Saynovits, M., Link, T. A. & Michel, H. Structure of a water soluble fragment of the 'Rieske' iron-sulfur protein of the bovine heart mitochondrial cytochrome bc1 complex determined by MAD phasing at 1.5 Å resolution. *Structure* **4**, 567–579 (1996).

10. Kusakabe, T., Hine, A. V., Hyberts, S. G. & Richardson, C. C. The Cys4 zinc finger of bacteriophage T7 primase in sequence-specific single-stranded DNA recognition. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 4295–4300 (1999).

11. Todone, F., Weinzierl, R. O. J., Brick, P. & Onesti, S. Crystal structure of RPB5, a universal eukaryotic RNA polymerase subunit and transcription factor interaction target. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6306–6310 (2000).

12. Adman, E. T., Sieker, L. C. & Jensen, L. H. Structure of rubredoxin from Desulfovibrio vulgaris at 1·5 Å resolution. *J. Mol. Biol.* **217**, 337–351 (1991).

13. Verhagen, M. F. J. M., Link, T. A. & Hagen, W. R. Electrochemical study of the redox properties of [2Fe-2S] ferredoxins Evidence for superreduction of the Rieske [2Fe-2S] cluster. *FEBS Lett.* **361**, 75–78 (1995).

14. Zu, Y. *et al.* Reduction Potentials of Rieske Clusters: Importance of the Coupling between Oxidation State and Histidine Protonation State. *Biochemistry* **42**, 12400–12408 (2003).

15. Brown, E. N. *et al.* Determining Rieske cluster reduction potentials. *J. Biol. Inorg. Chem.* **13**, 1301–1313 (2008).

16. Botelho, H. M. *et al.* Role of a novel disulfide bridge within the all-beta fold of soluble Rieske proteins. *J. Biol. Inorg. Chem.* **15**, 271–281 (2010).

17. Merbitz-Zahradnik, T., Zwicker, K., Nett, J. H., Link, T. A. & Trumpower, B. L. Elimination of the Disulfide Bridge in the Rieske Iron-Sulfur Protein Allows Assembly of the [2Fe-2S] Cluster into the Rieske Protein but Damages the Ubiquinol Oxidation Site in the Cytochrome bc1 Complex. *Biochemistry* **42**, 13637–13645 (2003).

18. Konkle, M. E. *et al.* Effects of pH on the Rieske Protein from Thermus thermophilus : A Spectroscopic and Structural Analysis ,. *Biochemistry* **48**, 9848–9857 (2009).

19. Lin, I. J. *et al.* Rieske protein from Thermus thermophilus: 15N NMR titration study demonstrates the role of iron-ligated histidines in the pH dependence of the reduction potential. *J. Am. Chem. Soc.* **128**, 10672–10673 (2006).

20. Xiao, Z. *et al.* The rubredoxin from Clostridium pasteurianum: Mutation of the iron cysteinyl ligands to serine. Crystal and molecular structures of oxidized and dithionite-treated forms of the Cys42Ser mutant. *J. Am. Chem. Soc.* **120**, 4135–4150 (1998).

21. Santos, H. *et al.* Aerobic Metabolism of Carbon Reserves by the "Obligate Anaerobe" Desulfovibrio gigas. *Biochemical and Biophysical Research Communications* vol. 195 551–557 (1993).

22. Eggink, G., Engel, H., Vriend, G., Terpstra, P. & Witholt, B. Rubredoxin reductase of Pseudomonas oleovorans. Structural relationship to other flavoprotein oxidoreductases based on one NAD and two FAD fingerprints. *J. Mol. Biol.* **212**,

135–142 (1990).

23. Fee, J. A. *et al.* Purification and characterization of the Rieske iron-sulfur protein from Thermus thermophilus. Evidence for a [2Fe-2S] cluster having non-cysteine ligands. *J. Biol. Chem.* **259**, 124–133 (1984).

24. Valer, L. *et al.* Methods to identify and characterize iron-sulfur oligopeptides in water. *Can. J. Chem.* (2022) doi:10.1139/cjc-2021-0237.

25. Moura, I., Bruschi, M., Le Gall, J., Moura, J. J. G. & Xavier, A. V. Isolation and characterization of desulforedoxin, a new type of non-heme iron protein from Desulfovibrio gigas. *Biochem. Biophys. Res. Commun.* **75**, 1037–1044 (1977).

26. Maelia, L. E., Millar, M. & Koch, S. A. General synthesis of iron(III) tetrathiolate complexes. Structural and spectroscopic models for the [Fe(Cys-S)4] center in oxidized rubredoxin. *Inorg. Chem.* **31**, 4594–4600 (1992).

27. Dos Santos, P. C. *Fe-S Proteins*. vol. 2353 (Springer US, 2021).

28. Coghlan, V. M. & Vickery, L. E. Expression of human ferredoxin and assembly of the [2Fe-2S] center in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 835–839 (1989).

29. Webert, H. *et al.* Functional reconstitution of mitochondrial Fe/S cluster synthesis on Isu1 reveals the involvement of ferredoxin. *Nat. Commun.* **5**, (2014).

30. Link, T. A. The Structures of Rieske and Rieske-Type Proteins. in *Advances in Inorganic Chemistry* vol. 47 83–157 (1999).

31. Emptage, M. H., Kent, T. A., Kennedy, M. C., Beinert, H. & Münck, E. Mössbauer and EPR studies of activated aconitase: development of a localized valence state at a subsite of the [4Fe-4S] cluster on binding of citrate. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 4674–4678 (1983).

32. Boxhammer, S., Glaser, S., Kühl, A., Wagner, A. K. & Schmidt, C. L. Characterization of the recombinant Rieske [2Fe–2S] proteins HcaC and YeaW from E. coli. *BioMetals* **21**, 459–467 (2008).

33. Lin, J., Zhou, T., Ye, K. & Wang, J. Crystal structure of human mitoNEET reveals distinct groups of iron-sulfur proteins. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14640–14645 (2007).

34. Wiley, S. E. *et al.* The outer mitochondrial membrane protein mitoNEET contains a novel redox-active 2Fe-2S cluster. *J. Biol. Chem.* **282**, 23745–23749 (2007).

35. Conlan, A. R. *et al.* Mutation of the His ligand in mitoNEET stabilizes the 2Fe-2S cluster despite conformational heterogeneity in the ligand environment. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 516–523 (2011).

36. Coghlan, V. M. & Vickery, L. E. Expression of human ferredoxin and assembly of the [2Fe-2S] center in Escherichia coli. *Proc. Natl. Acad. Sci.* **86**, 835–839 (1989).

Chapter 3 – Histidine ligated Iron-Sulfur Peptides

3. 1 Introduction

The synthetic approach to the study of iron-sulfur cluster started after the first successful synthesis of iron-sulfur clusters in vitro by Holm and colleagues in the 1970s,[1] using thiols containing organic molecules as scaffold. Originally, these reactions happened in organic solvents, such as methanol, tetrahydrofuran (THF) or acetonitrile. In 2008 Meyer and his colleagues reported the synthesis of a synthetic analogue of a Rieske cluster, a [2Fe-2S] with one iron coordinated by two thiols and the other iron coordinated by two imidazoles.[2] However, such complexes, as the ones studied previously by Holm, were obtained in organic solvent (THF) and the ligands consisted of non-protein scaffolds. Thus, these findings did not spark any prebiotic interest. The main obstacle was the condition in which the iron-sulfur cluster were obtained. In fact, most of the reactions that occurred on the early Earth are thought to have taken place in water. Recently, while investigating the cellular synthesis and transport of iron−sulfur clusters, the Cowan group revealed the ability of the tripeptide L-glutathione (γECG) to coordinate a [2Fe-2S] cluster in aqueous solution.[3] Soon after, our group showed that iron-sulfur clusters can be synthesized with simple, prebiotically plausible ligands (cysteine-containing di- and tripeptides) under UV light.[4] Exposure to UV irradiation leads to the photolysis of cysteine and other thiol-containing compounds, thereby releasing bisulfide (SH^-) ions into solution, as well as ferric ions (Fe^{3+}) through photooxidation of ferrous ions (Fe^{2+}). Notably, this synthetic route is similar to the modern biosynthetic pathway in terms of donation of sulfide during the synthesis of iron-sulfur clusters. Our group also demonstrated that the duplication of the tripeptide γECG sequence to produce longer peptides (up to 12 amino acids) increased the cluster's stability leading to the formation of a protoferredoxin;[5] this suggest that ancient iron-sulfur proteins may have emerged from simple and small peptides, in a modified Eck-Dayhoff scenario.[6] Later our group has shown that such simple prebiotic iron-sulfur cluster are able to mediate redox processes. Indeed, they can accept electrons from donors, such as NADH, and subsequently donate them to a suitable acceptor, such as H_2O_2.[7] This reaction if confined inside a model protocell can generate a pH gradient across the membrane. This is remarkable because the prebiotic synthesis of iron-sulfur cluster resemble the biological one, and these complexes can have an activity that mimics the one of evolved and complex proteins.

Modern day iron-sulfur proteins exploit primarily cysteine residues to coordinate their cluster. Indeed, all of the above mentioned studies used cysteine containing protein to synthesized the clusters. However, in biology there are several cases where the cluster is not completely ligated by cysteines. Beside the coordination provided by the sulfur atoms of cysteines, iron-sulfur proteins can exploit the

oxygen of the carboxylic group of aspartic and glutamic acids or the nitrogen of histidines.[8] None of these previous efforts have demonstrated the ligation of an iron-sulfur cluster by imidazole or an imidazole sidechain with a non-protein scaffold in water.

To investigate if nitrogen coordination to an iron-sulfur cluster can be achieved in aqueous solution, fourteen different peptides (Table 3.1) have been designed based on the sequence of the rubredoxin and Rieske protein discussed in Chapter 2. The four ligands to the [2Fe-2S] cluster of Rieske are within a 25 amino acid region of the protein that contains CXC/H and CXXC/H motifs. Two CXC/H pentapeptides (GCTHG and GCTCG) and one CXXH hexapeptide (GCPAHG) were tested. Three 25 amino acid peptides that encompassed the entire [2Fe-2S] binding site were also synthesized with the disulfide positions substituted with Ala (linear peptides). Stapled versions of the 25-residue peptide were attempted and failed due to the distance and flexibility between the Asp and Lys residues. The smaller 16-residue version was designed by removing the residues involved in the formation of an intervening anti-parallel β-sheet[9] in between the CXH and CXXH motifs. The staple was a lactam bridge formed by using an orthogonal protecting group strategy during synthesis followed by crosslinking with N,N-diisopropylethylamine (DIPEA), which led to the covalent joining of the side chains of aspartic acid and lysine. Rubredoxin contains two CXXC motifs separated by 29 amino acids. Four hexapeptides based on these motifs were synthesized, including GCTVCG, GCTVHG, GCPLCG, GCPLHG. A GCG peptide was also prepared for comparison (Table 3.1).

The designed peptides, N-acetyl Cys methyl ester (N-Ac-Cys-OMe) and N-acetyl His methyl ester (N-Ac-His-OMe) were incubated with Fe^{2+}, Co^{2+}, Zn^{2+}, Fe^{3+} and Fe^{3+} and sulfide (S^{2-}) to be evaluated. Nevertheless, these metallocofactors can be challenging to characterize when coordinated to multiple, small oligopeptides (≤ 6 amino acids), because the resulting solutions are typically in equilibrium between multiple types of iron–sulfur clusters each absorbing in the same region.[10] We employed several spectroscopic techniques, including UV-vis, CD and paramagnetic NMR spectroscopy, as well as cyclic voltammetry (CV) to confirm if they are able to coordinate iron-sulfur clusters.

In particular, CV is quite useful not only to understand the electrochemical properties of an iron-sulfur clusters but also to distinguish between different types of coordination. Biology makes use of different types of coordination to alter the redox properties of the iron ions. One notable effect of ligation by nitrogen atoms of histidine residues is a shift in reduction potential of the cluster. [2Fe-2S] clusters solely coordinated by cysteine, as found in ferredoxins, display reduction potentials from -150 to -400 mV (vs SHE),[11] while histidine coordination shifts the reduction potential to more positive values. The [2Fe-2S] clusters of MitoNEET have potentials from -100 mV to +50 mV.[12] The reduction potentials of [2Fe-2S] Rieske clusters vary from -150 to +400 mV.[13] The protonation state of the coordinated histidine impacts the reduction potential of the iron-sulfur cluster, and

so alterations of the pK_a of the histidine can be used to fine-tune potentials.[14] Such a strategy was used to adjust the reduction potential of MitoNEET across a 700 mV range.[15]

Name	Peptides sequence
Rieske Protein sequence	V**C**T**H**LG<u>C</u>IVSQWVADEEAAL**C**P<u>C</u>**H**G
CXH peptide	G**C**T**H**G
CXC peptide	G**C**T**C**G
CX_2C peptide	G**C**PA**H**G
CHCH linear	V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G
CCCH linear	V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G
CCCC linear	V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G
CHCH stapled	V**C**T**H**LG<u>D</u>IVSL**C**P<u>K</u>**H**G
CCCH stapled	V**C**T**C**LG<u>D</u>IVSL**C**P<u>K</u>**H**G
CCCC stapled	V**C**T**C**LG<u>D</u>IVSL**C**P<u>K</u>**C**G
Rubredoxin sequence	T**C**TV**C**G…V**C**PL**C**G
CX_2C peptide	G**C**TV**C**G
CX_2H peptide	G**C**TV**H**G
CX_2C peptide	G**C**PL**C**G
CX_2H peptide	G**C**PL**H**G

Table 3.1 - Designed peptides used in this study. Positions in bold indicate coordinating residues. Underlined residues indicate the position of the original disulfide bond, which was substituted with a lactam bridge for the stapled peptides.

3.2 Results

3.2.1 Absorption Spectra of Mononuclear Co^{2+} complexes

Metal binding was first evaluated with Co^{2+}, as the UV-visible absorption spectra of Co^{2+} complexes are easily interpretable in terms of the number of Cys ligands.[16] In fact the wavelength of the last peak in the UV-vis spectra is directly correlated to the number of cysteine involved in the coordination: four cysteines coordination show a peak at 750 nm, three cysteines coordination at around 700 nm and two cysteine coordination at 670-680 nm.

All of the small peptides (≤ six residues) with two cysteine and no histidine showed an absorbance peak at 750 nm (Figure 3.1 red curve), consistent with coordination by four thiolates.[16] To achieve such coordination, two peptides would have had to bind a single Co^{2+}. When one of the cysteine was substituted with a histidine, the absorbance peak shifted to 660 nm (Figure 3.1 blue curve), as would be expected for a complex with two cysteinyl and two histidinyl ligands.[16]

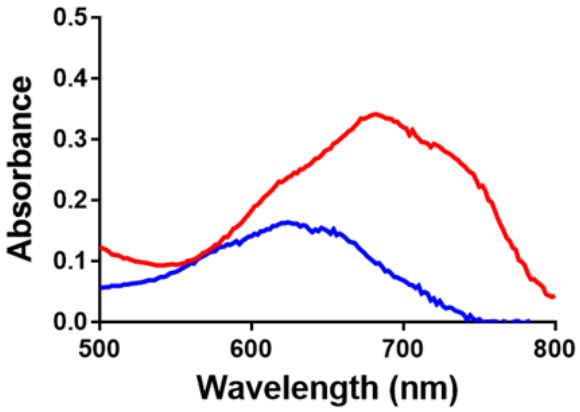

Figure 3.1 - Comparison between all Cys and mixed Cys - His coordination for Co^{2+}. UV-vis spectra and Co^{2+} coordinated to GCTVCG (red) and GCTVHG (blue). Conditions were 1.25 mM peptide, 1 mM metal ion, 20 mM Gly-Gly, pH 8.8.

These results showed that Co^{2+} is coordinated by the expected number of cysteine for each of these peptide. Short peptides (≤ 6 amino acid) containing two cysteine produced a spectrum with a peak at 750 nm indicating the presence of an all-cysteine tetrahedral coordination around the metal, while peptides containing one cysteine and one histidine displayed a peak at 670 nm hinting at a tetrahedral coordination where only two cysteine are involved and the two other position are occupied by histidine side chains. The longer linear and stapled Rieske peptides with four potential ligands to the metal ion produced UV-vis spectra congruent with the number of cysteine residues. CHCH, CCCH, and CCCC Rieske peptides, both linear (Figure 3.2) and stapled had absorbance maxima at 660 nm, 710 nm, and 750 nm, respectively, when bound to Co^{2+}.

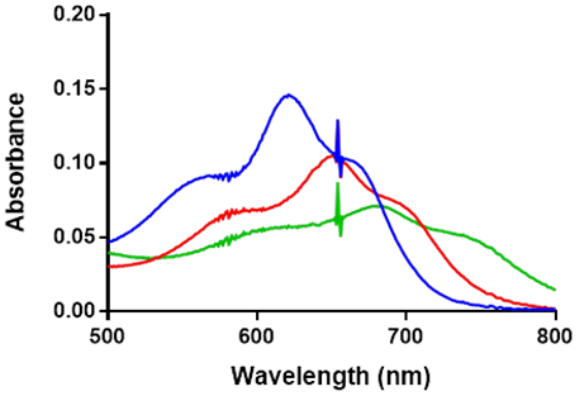

Figure 3.2 - UV-vis spectra of Co^{2+} coordinated by peptides with an increasing number of Cys. Spectra are of Co^{2+} peptides V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G (blue, CHCH), V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G (red, CCCH), and V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G (green, CCCC). Ligands are in bold. Solution conditions were 2 mM peptide, 0.5 mM $CoCl_2$, pH 8.8.

3.2.2 Peptide Affinities for Co^{2+}, Zn^{2+}, and Fe^{2+}

We next evaluated the binding affinity for Co^{2+}, Zn^{2+}, and Fe^{2+} (Figure 3.3 – S2.1-6, Table S2.1). Since Zn^{2+} complexes do not absorb in the UV-Vis region, the affinity for Zn^{2+} was determined by substitution of bound Co^{2+}, in a competition titration experiment.

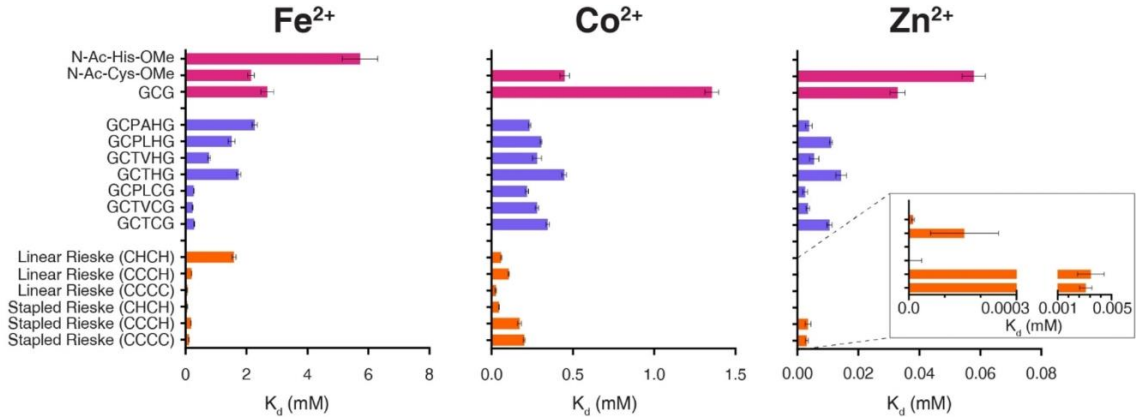

Figure 3.3 - K_d of peptides for Fe^{2+}, Co^{2+}, and Zn^{2+}. Data obtained represent mean and SD of distinct samples, n=3. Peptides with a single ligand (Cys or His) are in red, peptides with two ligands are in blue, and peptides with four ligands are in orange. The K_d of N-Ac-His-OMe is not reported for Co^{2+} and Zn^{2+}.

Binding affinities for N-Ac-Cys-OMe, N-Ac-His-OMe, and GCG, i.e. molecules with only one potential ligand, were always less (higher K_d) than peptides that contained more than one potential ligand for all three metal ions.

Molecules with two ligands possessed, on average, 2.7-fold greater affinity for the tested metal ions, and peptides with four potential ligands bound the metal ions 2.8-fold more tightly than peptides with two ligands. Affinities followed the Irving-Williams series, with peptides binding Zn^{2+} more strongly than Co^{2+} followed by Fe^{2+}. There was a stronger sequence dependence, confirmed by calculations of coefficients of variation, of affinities for Fe^{2+} followed by Zn^{2+} and then Co^{2+}. For Fe^{2+}, the preference for cysteine over histidine was most evident for the two ligand containing peptides (Figure 3.3 blue bars, table S2.2). Of these, peptides that contained two cysteine residues bound Fe^{2+}, on average, 5.9-fold greater than peptides with one cysteine and one histidine. The Linear Rieske (CCCC) peptide bound all the tested metal ions more strongly than the other Linear Rieske peptides. Conversely, of the Stapled Rieske peptides, the best binder was (CHCH), i.e. the peptide that contained two cysteine and two histidine. The lactam bridge may have helped pre-organize the ligating residues for binding and thus decrease the entropic cost associated with coordination.

3.2.3 Absorption Spectra of $[1Fe\text{-}0S]^{2+}$ cluster

Generally, absorption spectra with Fe^{2+} do not have any feature above 400 nm.[10] The only peaks present in an all cysteines coordinated $[1Fe\text{-}0S]^{2+}$ fall near 310 nm and possibly also a shoulder at 350 nm is present. The incubation of the peptides with Fe^{2+} produced either a clear peak at 315 nm when coordination by four cysteine was possible or a shoulder in the 300-310 nm when histidine was present (Figure 3.4). The peak at 315 nm is consistent with previously reported $[1Fe\text{-}0S]^{2+}$ UV-vis spectra.[17,18]

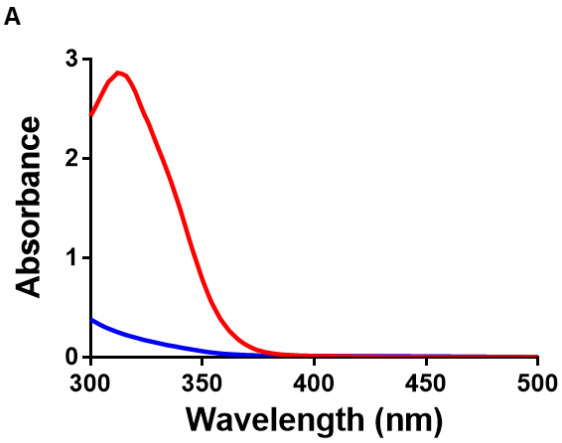

Figure 3.4 - Comparison between all Cys and mixed Cys - His coordination for $[1Fe\text{-}0S]^{2+}$. UV-vis spectra of Fe^{2+} coordinated to GCTVCG (red) and GCTVHG (blue). Conditions were 1.25 mM peptide, 1 mM metal ion, 20 mM Gly-Gly, pH 8.8.

There were only two exceptions: N-Ac-Cys-OMe and N-Ac-His-OMe. The first one gave a featureless UV-vis spectrum above 250 nm, while the second one showed a broad shoulder below 300 nm (Figure 3.5). The UV-vis spectra of peptidyl complexes with Fe^{2+}, due to the lack of characteristic features, were less clearly interpretable than spectra with Co^{2+}. However, there was a completely different spectral behavior between the only-cysteine containing peptides and the mixed ligands one.

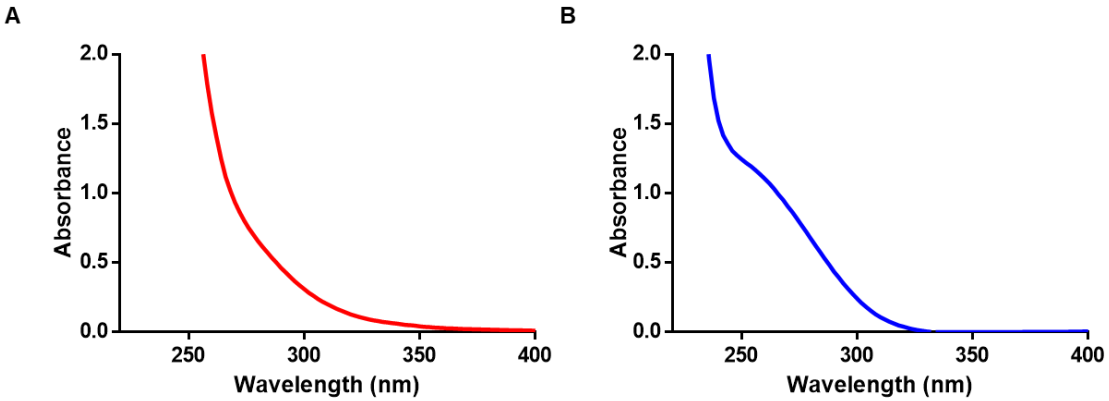

Figure 3.5 - UV-vis spectra of $[1Fe\text{-}0S]^{2+}$ centers coordinated to N-Ac-Cys-OMe (**A**) and N-Ac-His-OMe (**B**). Conditions were 2.5 mM amino acid, 10 mM $FeCl_3$, 20 mM Gly-Gly pH 8.8.

3.2.4 Absorption Spectra of $[1Fe\text{-}0S]^{3+}$ cluster

The peptides were incubated with Fe^{3+} to evaluate their ability to coordinate a $[1Fe\text{-}0S]^{3+}$ cluster. The ratio between the ligands (cysteine + histidine) and the metal was kept 16:1. Generally, $[1Fe\text{-}0S]^{3+}$ peptides produce UV-vis spectra that do not look like the spectra of oxidized rubredoxin.[19,20] However, they still display bands above 400 nm. Both the short and long all-cysteine containing peptides showed a broad band at 490 nm (Figure 3.6A, S2.7A, S2.8A), while mixed ligands peptides had two bands at 450 and 570 nm (Figure 3.6B, S2.7B, S2.8B-C).

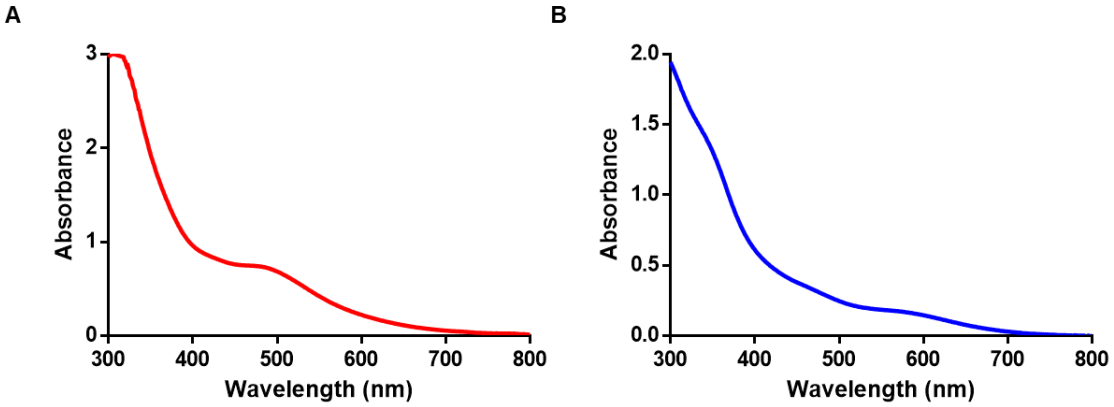

Figure 3.6 - UV-vis spectra of $[1Fe\text{-}0S]^{3+}$ centers coordinated to GCTVCG (**A**) and GCTVHG (**B**). Conditions were 4 mM peptide, 0.5 mM $FeCl_3$, pH 8.8.

The addition of Fe^{3+} to the 25 residue Rieske-like peptides and the shorter Stapled peptides led to broad, featureless absorption in the UV-Vis region, with absorption increasing with increasing number of Cysteine (Figure 3.7). The stapled peptide absorbed more strongly than the 25 residue peptides upon the addition of Fe^{2+}. In general, the same pattern of peaks observed with the short peptides was displayed by the spectra of the $[1Fe\text{-}0S]^{3+}$ cluster coordinated by both the linear and stapled Rieske peptides. The only exception is the spectra produced by the complex coordinated by the stapled CCCH peptide that showed a band more similar to the one of an all-cysteine ligated $[1Fe\text{-}0S]^{3+}$, possibly indicating that more peptides come together to fully coordinate the iron ion.

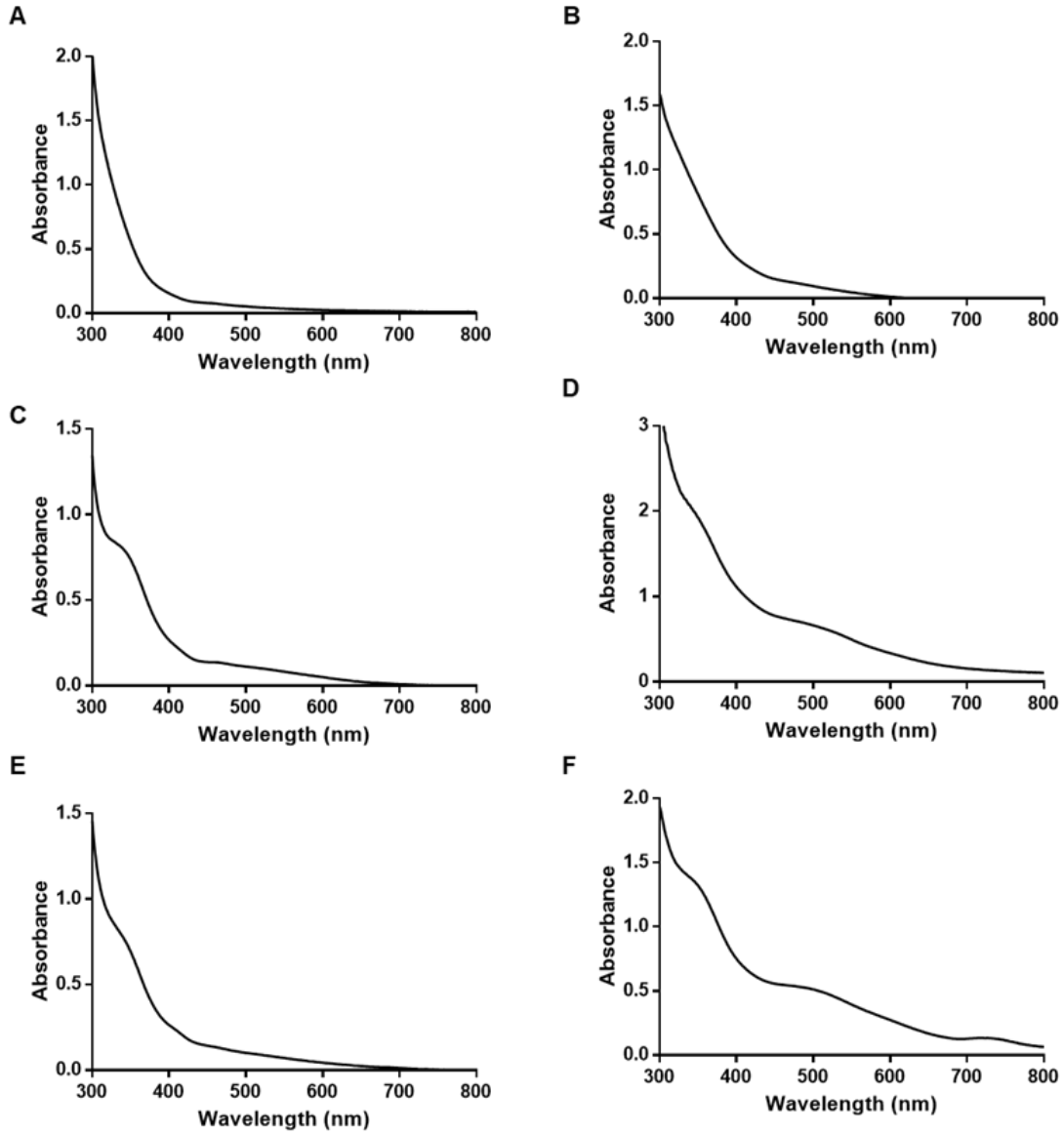

Figure 3.7 - UV-vis spectra of [1Fe-0S] centers coordinated to Linear and Stapled Rieske peptides. The peptides were Linear (CHCH): V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G (**A**), Stapled (CHCH): V**C**T**H**LGDIVSL**C**PK**H**G (**B**), Linear (CCCH): V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G (**C**), Stapled (CCCH): V**C**T**C**LGDIVSL**C**PK**H**G (**D**), Linear (CCCC): V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G (**E**), Stapled (CCCC): V**C**T**C**LGDIVSL**C**PK**C**G (**F**). Solutions contained 2 mM peptide, 0.5 mM $FeCl_3$, pH 8.8. Bold positions indicate ligands, and underlined positions indicate residues joined by a lactam bridge.

3.2.5 Stability of $[1Fe\text{-}0S]^{3+}$ cluster

Due to the excess amount of thiols presents, $[1Fe\text{-}0S]^{3+}$ peptides complex in water tends to get constantly reduced to $[1Fe\text{-}0S]^{2+}$.[4] The process can be monitored through UV-vis spectroscopy. In fact, upon reduction the UV-vis spectra feature above 400 nm starts to gradually lose intensity until they disappear.

The all-cysteine peptides showed from the beginning a gradually decrease of the intensity of the peak at 490 nm (Figure 3.8A). The absorbance values over time were best fit by a two-phase decay, that produced two half-life values, one slow and one fast (Table 3.2). This might indicate that two reduction processes are happening simultaneously, most probably one is due to the intramolecular oxidation of thiols present on the same peptides and the other due to the intermolecular oxidation among thiols on different peptides.

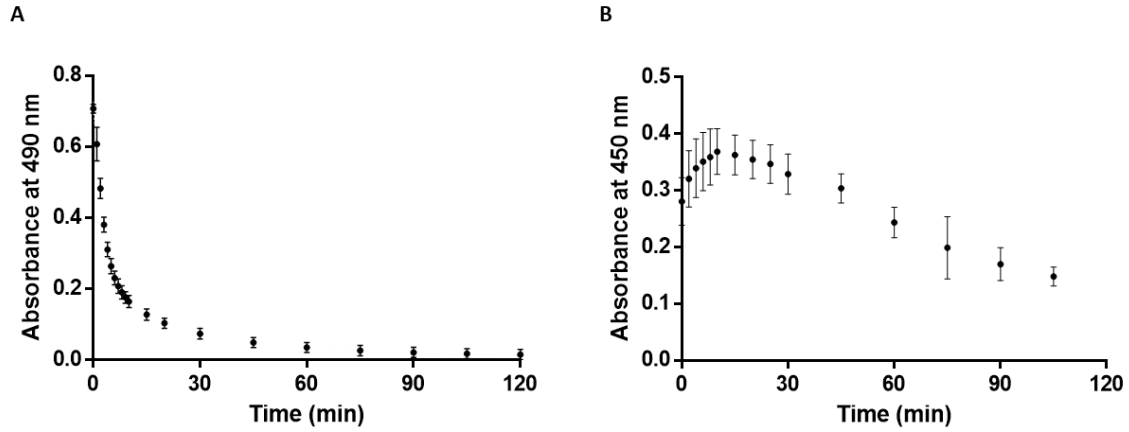

Figure 3.8 – Stability over time of [1Fe-0S] center coordinated by GCTVGC (**A**) and GCTVHG (**B**). Conditions were 4 mM peptide, 0.5 mM $FeCl_3$, pH 8.8. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

Peptide sequence	**$[1Fe\text{-}0S]^{3+}$ fast half-life (min)**	**$[1Fe\text{-}0S]^{3+}$ slow half-life (min)**
GCTCG	1.077	12.94
GCTVCG	2.334	23.43
GCPLCG	1.896	12.35

Table 3.2 – Half-life of $[1Fe\text{-}0S]^{3+}$ clusters coordinated by CXC and CX_2C peptides. Conditions were 4 mM peptide, 0.5 mM $FeCl_3$, pH 8.8.

The UV-vis spectra of $[1Fe\text{-}0S]^{3+}$ cluster coordinated by short histidine-containing peptides displayed a different behavior comparted to the all-cysteine ones. For the peptides GCTHG and GCTVHG, we observed a sort of initial maturation phase for the first minutes of the reaction, with an increase in the absorption intensity of the region above 400 nm, followed by a slow decrease (Figure 3.8B – S2.9). Also, the model peptide GCG showed the same behavior of the histidine and cysteine containing peptides. However, the peptides GCPAHG and GCPLHG did not display the initial phase (Table 3.3) (Figure S2.10).

Peptide sequence	"Maturation" time (min)
GCTHG	≈ 14
GCTVHG	≈ 9
GCPAHG	no maturation
GCPLHG	no maturation
GCG	≈ 15

Table 3.3 – "Maturation" time of CXH, CX_2H and GCG.

3.2.6 Paramagnetic NMR of Mononuclear [1Fe-0S] cluster

To confirm the coordination of iron, paramagnetic ^1H NMR spectra were acquired. $[1Fe\text{-}0S]^{3+}$ rubredoxin contains a single ferric ion with S = 5/2. The fast relaxation occurring in the presence of this high spin metal ion strongly influences both the hyperfine shift of the resonances and their linewidths. Cysteine H_α nuclei are typically easily observed within the 10–20 ppm region.[21] However, due to the presence of excess thiols, $[1Fe\text{-}0S]^{3+}$ is rapidly reduced to $[1Fe\text{-}0S]^{2+}$. The paramagnetism of the Fe^{2+} center gives rise to hyperfine shifted resonances of Cysteine H_α in the 10-20 ppm region and Cysteine H_β within the 100-300 ppm range. Indeed, the paramagnetic 1H NMR spectra of the cysteine-containing peptides incubated with Fe^{3+} showed the presence of such paramagnetically shifted resonances of cysteine H_α and H_β (Figure 3.9), confirming both the coordination and the reduction of Fe^{3+} to Fe^{2+}.

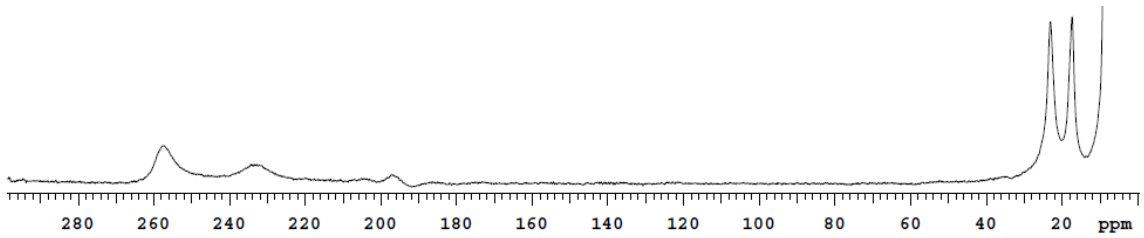

Figure 3.9 - Paramagnetic NMR spectra of [1Fe-0S] clusters coordinated by GCTVCG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8.

Significant paramagnetically shifted resonances were not detected for peptides containing histidine (Figure S2.7). Unfortunately, the Linear and Stapled Rieske peptides were not sufficiently soluble for the concentrations needed for paramagnetic NMR spectroscopy (Figure S2.12-17).

3.2.7 Absorption and Circular Dichroism Spectra of $[2Fe\text{-}2S]^{2+}$ cluster

To determine whether the synthesized peptides could coordinate a [2Fe-2S] cluster, each peptide was anaerobically incubated with Fe^{3+} and HS^- and the UV-vis spectra was collected. The UV-vis spectra of cysteine-containing small (≤ 6

amino acids) peptides that lacked histidine residues showed bands at 420 nm and 450 nm indicative of a ferredoxin-like [2Fe-2S] cluster (Figure 3.10A, S2.16A, S2.17A). Near-UV-visible circular dichroism (CD) spectra showed a band at 470 nm that was similar to the 450 nm peak seen with ferredoxin, but with lower molar ellipticity (Figure 3.10B).

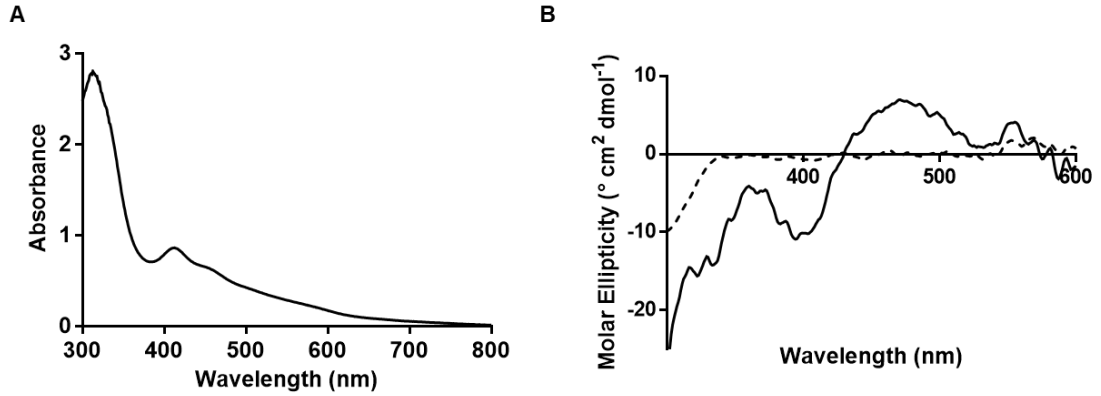

Figure 3.10 - UV-vis (**A**) and CD (**B**) spectra of $[2Fe\text{-}2S]^{2+}$ cluster coordinated by GCTVCG. In panel B the CD spectra of the apo peptide is represented with a dashed line. Solution conditions were 4 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8.

Conversely, UV-Vis absorption spectra of small peptides that possessed both Cys and His residues did not resemble the spectra of ferredoxin or the Rieske protein (Figure 3.11A). The spectra display a shoulder around 410 nm and a band at 590 nm. Although the CD spectra lack any feature (Figure 3.11B, S2.16B, S2.17B-C), suggesting that there is not a unique coordination geometry.

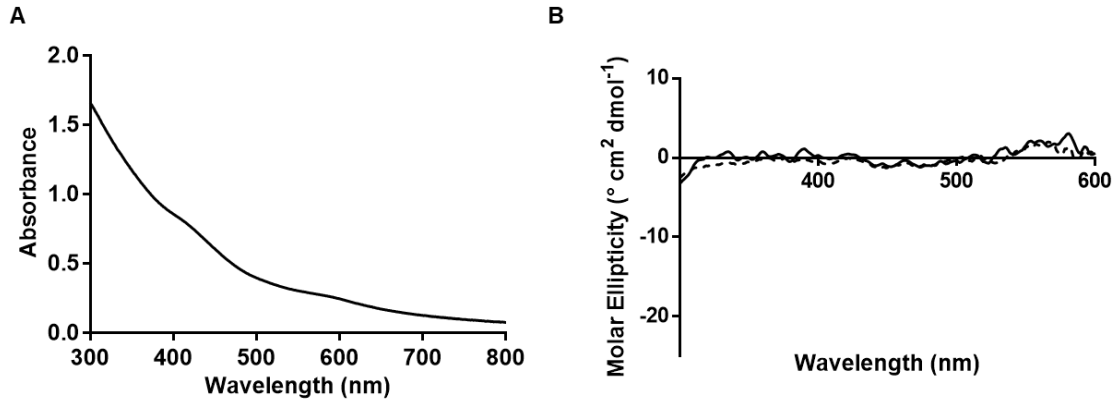

Figure 3.11 - UV-vis (**A**) and CD (**B**) spectra of $[2Fe\text{-}2S]^{2+}$ cluster coordinated by GCTVHG. In panel B the CD spectra of the apo peptide is represented with a dashed line. Solution conditions were 4 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8.

In presence of Fe^{3+} and HS^-, UV-Vis spectra were indicative of the presence of a ferredoxin-like $[2Fe\text{-}2S]^{2+}$ cluster for linear peptides CCCH and CCCC and not for the more Rieske-like sequence found in the linear CHCH peptide (Figure 3.12A-C-E). The CD spectra of the cluster stabilized by CHCH and CCCC linear

peptides did not display any discernible peak, while the CCCH peptide produced CD spectrum with a broad positive peak around 450 nm, similarly to ferredoxin-like $[2Fe\text{-}2S]^{2+}$ clusters (Figure 3.12B-D-F).

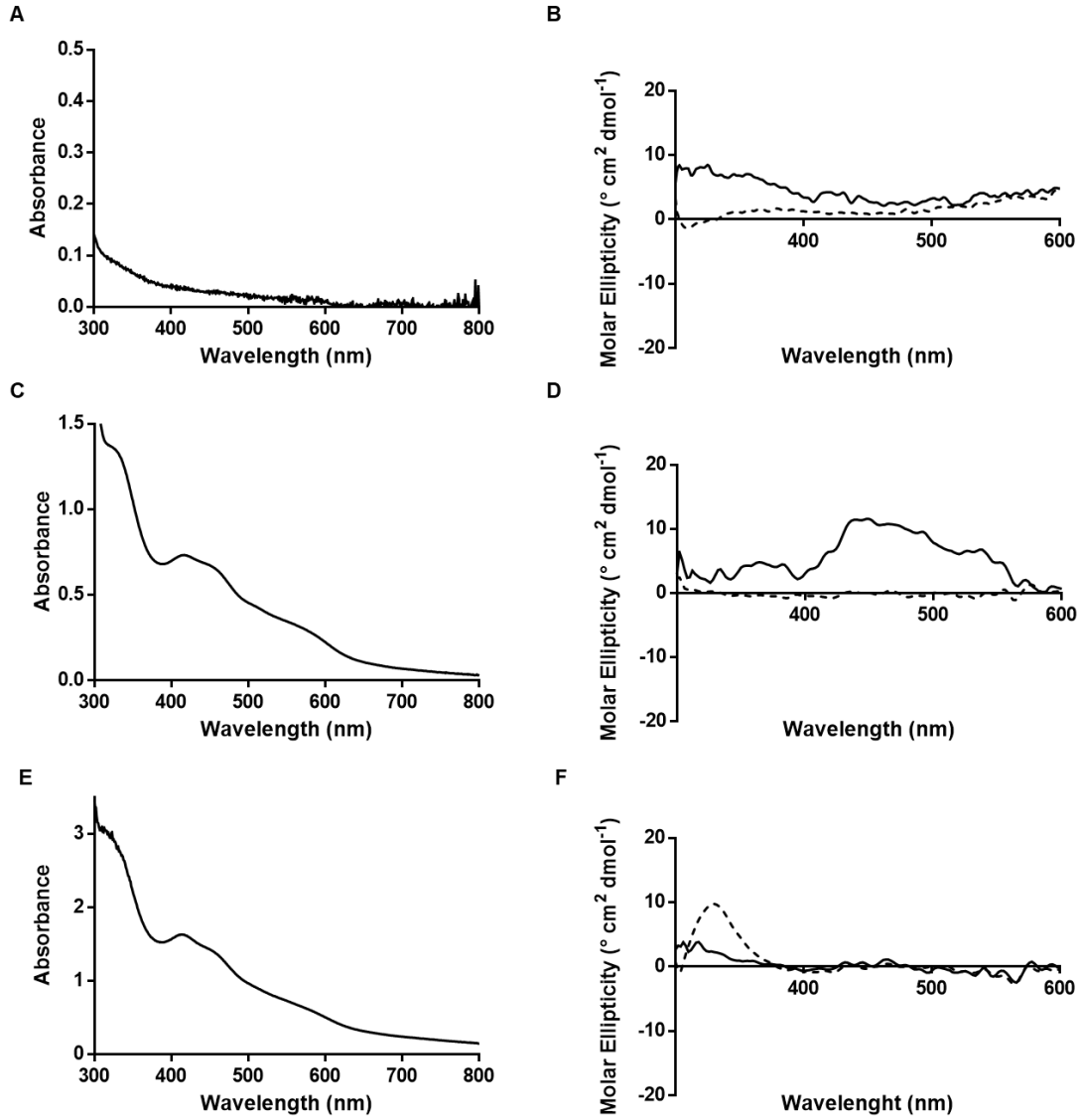

Figure 3.12 – UV-vis and CD spectra of $[2Fe\text{-}2S]^{2+}$ clusters coordinated by linear Rieske peptides. UV-vis spectra of $[2Fe\text{-}2S]^{2+}$ coordinated by linear CHCH (**A**), CCCH (**C**) and CCCC (**E**) linear peptide. UV-vis spectra of $[2Fe\text{-}2S]^{2+}$ coordinated by linear CHCH (**B**), CCCH (**D**) and CCCC (**F**) linear peptide. Dashed line represent the CD spectra of the apo peptides. Solution conditions were 2 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8.

UV-Vis spectra of the stapled peptide lacked well defined peaks, but as seen with the linear peptides, the bands became more prominent as the number of Cys increased (Figure S2.20). The data suggested that the linear peptides better stabilized the formation of a $[2Fe\text{-}2S]^{3+}$ cluster than the stapled peptides.

3.2.8 Stability of $[2Fe\text{-}2S]^{2+}$ cluster

Next, we evaluated the stability of $[2Fe\text{-}2S]^{2+}$ cluster stabilized by peptides. We have monitored the UV-vis spectra of the clusters over 10 h. Contrary to the $[1Fe\text{-}0S]^{3+}$ peptides, $[2Fe\text{-}2S]^{2+}$ clusters were stable for a longer period of time. For the clusters coordinated by cysteine containing peptides (Figure 3.13), the absorbance was continuously increasing over the first hour. After the intensity of the peaks was slowly decreasing, however the ratio between the absorbance at 450 and 420 nm was stable across the 10 h. This ratio is generally used to determine if the cluster is getting reduced, in fact the main difference between the UV-vis spectra of $[2Fe\text{-}2S]^{2+}$ and $[2Fe\text{-}2S]^{1+}$ is the lack of the peak at 450 nm. The stability of this ratio is an indication that the cluster is not getting significantly reduced over time.

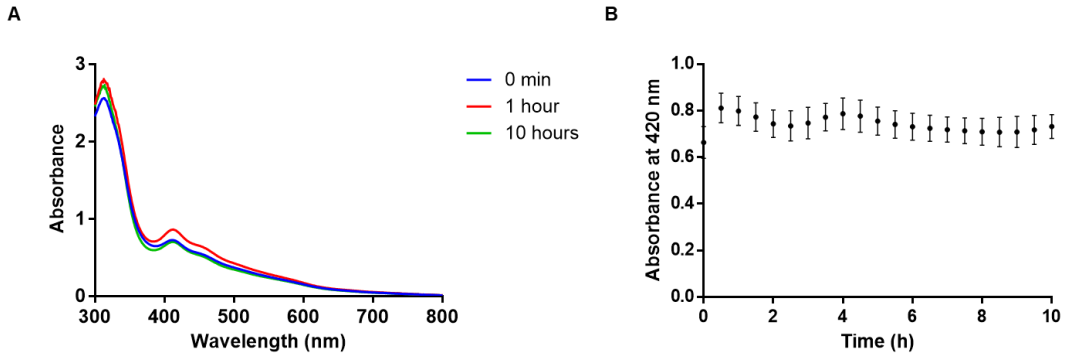

Figure 3.13 – Stability over time of $[2Fe\text{-}2S]^{2+}$ cluster coordinated by GCTVCG. **A**. UV-vis spectra of GCTVCG $[2Fe\text{-}2S]^{2+}$ cluster at 0, 1 and 10 hours. Conditions were 4 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8. **B**. Absorbance at 420 nm over time. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

The histidine-containing peptides in presence of Fe^{3+} and SH^- produced spectra with absorbance that is quite stable over time (Figure 3.14). As showed by the cluster coordinated by cysteine-containing peptides, the $[2Fe\text{-}2S]^{2+}$ coordinated by CXH and CX_2H show that the peak at 410 nm increased its intensity over the first hour of incubation. Over the subsequent hours the peak intensity is decreased.

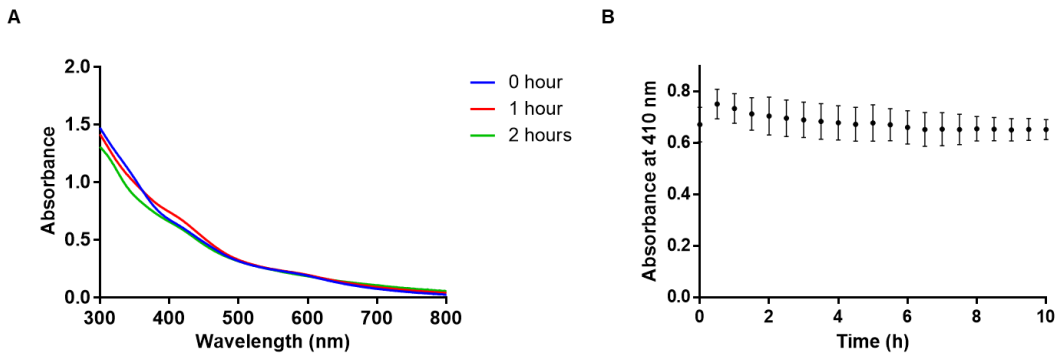

Figure 3.14 - Stability over time of $[2Fe\text{-}2S]^{2+}$ cluster coordinated by GCTVHG. **A**. UV-vis spectra of GCTVHG $[2Fe\text{-}2S]^{2+}$ cluster at 0, 1 and 10 hours. **B**. Absorbance at 420 nm over time. Conditions were 4 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

3.2.9 Paramagnetic NMR of $[2Fe\text{-}2S]^{2+}$ cluster

Paramagnetic ^1H NMR spectroscopy was performed to confirm the presence of the $[2Fe\text{-}2S]^{2+}$ cluster. The NMR spectra produced by the cluster coordinated by all-cysteine short peptides showed a broad resonance at 34 ppm (Figure 3.15). This peak is produced by the H_β resonances of ligating cysteines in [2Fe-2S] cluster. The NMR spectrum also revealed the continued presence of the mononuclear center that was not observable by UV-vis absorption. In fact only a [1Fe-0S] cluster displays resonance falling within the 100–300 ppm range.[10] Such finding was not surprising, since many peptides exist in an equilibrium between the coordination of a mononuclear center and a [2Fe-2S] cluster when in an environment containing iron and sulfide ions.[5,10,22]

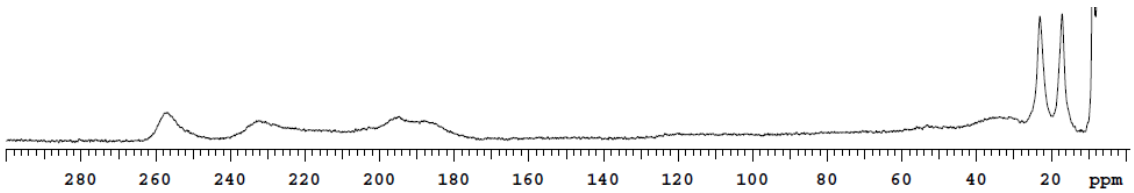

Figure 3.15 - Paramagnetic NMR spectra of a mixture of $[1Fe\text{-}0S]^{2+}$ and $[2Fe\text{-}2S]^{2+}$ clusters coordinated by GCTVCG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, 2 mM Na_2S, pH 8.8.

In the histidine-containing peptides' spectra paramagnetic ^1H NMR spectra showed sharp peaks near 12 ppm and 35 ppm in addition to a shoulder at 25 ppm (Figure 3.16A). The peak around 35 ppm was consistent with previously reported spectra of Rieske proteins.[13,23]

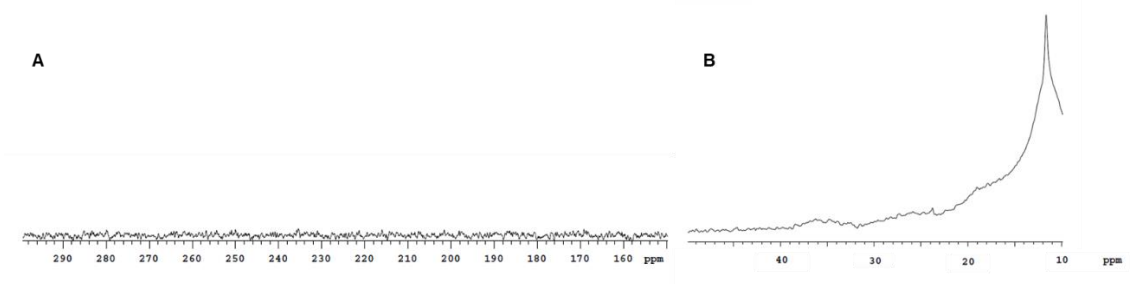

Figure 3.16 - Paramagnetic NMR spectra of $[2Fe\text{-}2S]^{2+}$ clusters coordinated by GCTVHG. Downfield (**A**) and upfield (**B**) region. Conditions were 16 mM peptide, 2 mM $FeCl_3$, 2 mM Na_2S, pH 8.8.

Given that both ferredoxin-like and Rieske-like $[2Fe\text{-}2S]^{2+}$ show their characteristic signal around 35 ppm (Figure 3.15 – 3.16), the NMR data collected with the short peptides were not sufficient to distinguish between four cysteine coordination and the one with two cysteine and histidine.

3.2.10 Cyclic voltammetry on [1Fe-0] and [2Fe-2S] clusters

Reduction potentials of iron-sulfur clusters are strongly influenced by their ligands. The substitution of a negatively charged Cys ligand with a neutral His typically increases the reduction potential of the cluster. Therefore, to confirm whether the histidine residues were ligating the iron ions, cyclic voltammetry was carried-out on the prepared complexes (Figure 3.17, Table S2.3).

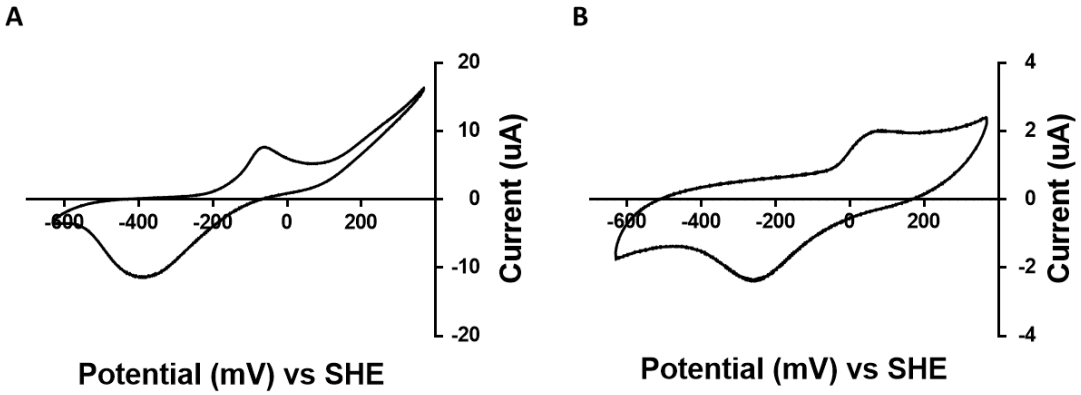

Figure 3.17 - Cyclic voltammetry of [2Fe-2S] peptides. Cyclic voltammetry of [2Fe-2S] GCTVCG (A) and [2Fe-2S] GCTVHG (B). Conditions were 16 mM peptide, 2 mM $FeCl_3$, 2 mM Na_2S, pH 8.8, scan rate 100 mV/s.

First, we evaluated the reduction potential of the mononuclear cluster (Figure S2.21-29). All of the mononuclear complexes had reduction potentials in the range expected for rubredoxin-like proteins, and other reported mononuclear centers coordinated by small thiol-containing molecules.[11] The most negative reduction potential measured was for N-Ac-Cys-OMe (-104 ± 21 mV; all reduction

potentials are vs. SHE). The most positive reduction potential was that of the pentapeptide GCTHG (+16 ± 1 mV) (Figure 3.18). [1Fe-0S] complexes with N-Ac-Cys-OMe and GCG had lower reduction potentials than peptides containing two potential ligands to the iron ion. Although the average reduction potentials were greater for histidine-containing peptides, the values were within error. Therefore, it was not possible to confirm ligation by histidine for the mononuclear complexes.

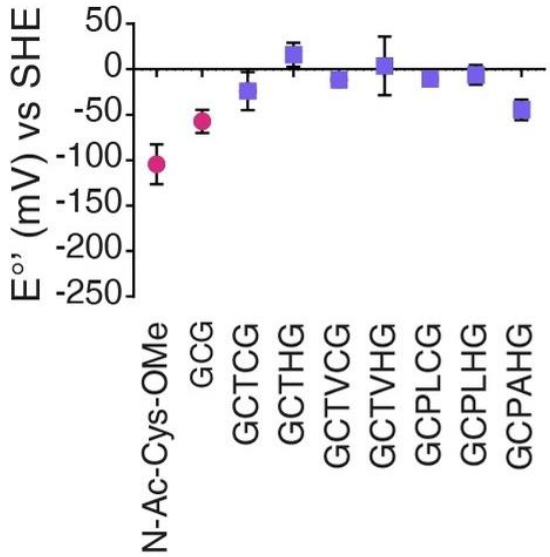

Figure 3.18 - Measured formal redox potentials of [1Fe-0S] peptides (vs SHE). Data obtained represent mean and SD of distinct samples, n = 3 replicates. Condition were 16 mM ligating residues, 1 mM $FeCl_3$, 1 mM Na_2S, pH 8.8. Red circle, 1 ligand; blue square, 2 ligands.

Conversely, cyclic voltammetry of [2Fe-2S] clusters coordinated to peptides showed two different types of behavior (Figure S2.30-36). The all cysteine containing peptides had reduction potentials in the range of -210 mV to -231 mV, with the exception of GCG (-117 ± 12 mV). The data were similar to [2Fe-2S] ferredoxin with four ligating Cys. The peptides that contained both His and Cys gave more positive reduction potentials between -63 mV and -83 mV (Figure 3.19), as would be expected for ligation by histidine and similar to Rieske proteins. We were not able to determine the reduction potentials of all the peptides, because several did not produce distinguishable reduction and oxidation peaks. These peptides included GCPLCG, GCTCG, and most of the Linear and Stapled Riske peptides. The only exceptions were the Stapled Rieske peptides with two cysteine and two histidine (CHCH) and with three cysteine and one histidine (CCCH), which had reduction potentials of -63 mV and -49 mV, respectively, consistent with nitrogen coordination.

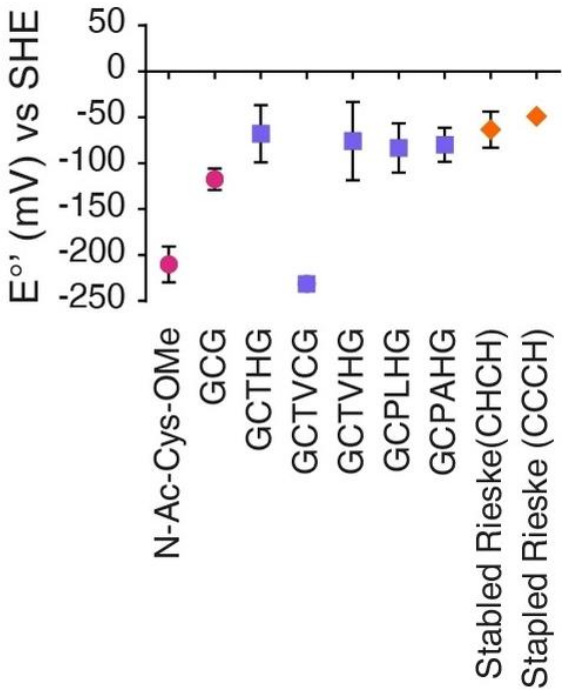

Figure 3.19 - Measured formal redox potentials of [2Fe-2S] peptides. Data obtained represent mean and SD of distinct samples, n = 3 replicates. Condition were 16 mM ligating residues, 1 mM $FeCl_3$, 1 mM Na_2S, pH 8.8. Red circle, 1 ligand; blue square, 2 ligands; orange diamond, stapled CHCH and CCCH peptides.

The reduction potential of Rieske proteins is pH dependent.[24] Lower pH leads to increased protonation of the ligating His and thus higher reduction potential. Therefore, if the investigated peptides bound the iron-sulfur cluster, in part, through His residues, then the reduction potential would be expected to be influenced by pH. Two [2Fe-2S] peptides, one lacking His (GCTVCG) and one containing His (GCTVHG) were subjected to cyclic voltammetry between pH 7.8 and 9.8. As expected, the reduction potential of the [2Fe-2S] cluster coordinated by GCTVHG decreased as the pH increased with a 92 mV/pH dependence (Figure 3.20, Table S2.4). Rieske proteins show a similar dependence (120 mV/pH).[25] Conversely, the reduction potential of the [2Fe-2S] cluster coordinated to GCTVCG was consistently near -230 mV between pH 8.3 and 9.8. Data are not reported at pH 7.8, because [2Fe-2S] GCTVCG was not stable at this pH. The data were consistent with the histidine residues coordinating the iron-sulfur cluster.

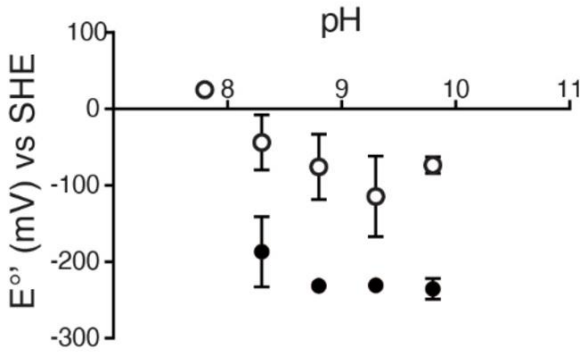

Figure 3.20 - Formal reduction potentials of [2Fe-2S] GCTVCG (filled circles) and [2Fe-2S] GCTVHG (open cicles) at different pH. Measurements were on 8 mM peptide, 1 mM $FeCl_3$, 1 mM Na_2S, pH 8.8 with a scan rate of 100 mV/s. At pH 7.8 the cluster coordinated by GCTVCG was not formed. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

3.3 Discussion

In the past, prebiotic iron-sulfur cluster research focused their interest on all-cysteine coordinated iron-sulfur cluster. Even though in Miller's experiments with atmosphere containing H_2S, the presence cysteamine, which is the degradation product of cysteine, was observed,[26] only recently a prebiotic plausible synthetic route for the formation of such amino acid has been reported.[27] On the other hand, although histidine is often argued to be prebiotically implausible because of phylogenetic analyses, phylogeny tells us little about the chemistry that transpired on the prebiotic Earth,[28] and histidine can be synthesized from erythrose and formamidine.[29] Additionally, the side-chain of histidine, i.e. imidazole, is frequently invoked as a leaving group in the nonenzymatic polymerization of nucleic acids.[30] It is also possible to envision a scenario were the coordination and redox capabilities of different iron-sulfur peptides might have guided a prebiotic selection of the sequence most suited for specifics environments and tasks.

Here, we show that in the absence of evolved scaffolds, [2Fe-2S] clusters can form coordinated to peptides containing cysteine and histidine ligands. Ligation by histidine alters the redox property of the cluster, shifting the reduction potential by 150 mV. In this work, we showed that if small peptides with cysteine and histidines existed on the prebiotic Earth, then such peptides may have been capable of coordinating a [2Fe-2S] cluster with reduction potentials distinct from iron-sulfur peptides that relied on complete cysteinyl coordination. The possibility of having metallopeptides with different redox behavior might indicate that more simplistic ways of assembling an electron transport chain with metallopeptides could be possible. Indeed, differently ligated cluster might have also engaged in more complex or multi step redox processes in a manner similar to how in the modern electron transport chain different iron-sulfur proteins.

3.4 Material and methods

Materials

All reagents were bought from VWR Chemicals and Sigma Aldrich and used with no further purification. All the synthetic procedures were performed with a Schlenk line and Schlenk glassware under controlled N2 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.

Peptide Synthesis

Peptides were synthesized via solid phase peptide synthesis, as previously described.11 N,N-dimethyl formamide (DMF) was used as the solvent and 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 the building blocks. The peptide chain was elongated by sequential Fmoc deprotection of the residue anchored to the resin and Fmoc-AA-OH coupling. 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. with a N,N'-diisopropylcarbodiimide (DIC)/HOBt mixture. For the cyclized peptides, Fmoc protected Asp (Allyl) and Lys (Alloc) were used. After the synthesis of the full peptide, before the last Fmoc deprotection, Asp and the Lys were deprotected by treatment with 0.5 equiv $Pd(PPh_3)_4$, 10 equiv of $PhSiH_3$ for 2 h, in the dark under constant flux of N_2 and repeated twice. After deprotection, the resin was washes five times with dichloromethane (DCM), sodium diethyldithiocarbamate in DMF (0.1% w/w) and DMF. Cyclization was achieved with a 24 h treatment hydroxyl-benzotriazole (HOBt), N,N-diisopropylethyl amine (DIPEA) and (7-Azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyAOP) in DMF. At the end of the synthesis, 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) (97:1:1:1, v/v) for 2 h. The product was precipitated with cold diethyl ether/petroleum ether (30:70, v/v) followed by washing cycles with diethyl ether. The peptides were finally solubilized in 20% acetic acid (v/v), flash frozen in liquid nitrogen and lyophilized overnight at -84 °C (Labconco FreeZone Freeze Dryer). The stapled peptides were further purified by HPLC (Agilent 1100 Series) with a reverse phase Zorbax SB-C12 column (Agilent Technologies) using H_2O and acetonitrile (both

containing 0.1 % TFA) as the mobile phase. The gradient was from 20% to 100% acetonitrile over 15 min. After purification, fractions containing peptide were lyophilized overnight at -84 °C.

Synthesis of the clusters

All steps were carried-out under anoxic conditions inside a Genesis 2P glovebox system (Vacuum Atmospheres Company) with O2 < 1 ppm. The peptides were dissolved in ultrapure degassed H2O (Synergy UV water purification system, Merck, Darmstadt, Germany) and the pH was adjusted to pH 8.8 using NaOH (5 M), if not specified otherwise. For the formation of the mononuclear complex, FeCl3 was added from a 0.33 M stock using to 1:8 Fe3+:(Cys+His) concentration. [2Fe-2S] clusters were synthesized by the addition of FeCl3 (0.33 M stock) and Na2S (0.33 M stock) to 1:1:8 Fe3+:S2-:(Cys+His) concentration.

UV-Vis absorption and Circular Dichroism spectroscopy

The UV-vis spectra of the samples were recorded inside a glovebox with a Genesys 150 UV-Vis spectrophotometer (ThermoFisher) with an integration time of 0.5 s and an interval of 1 nm. Circular Dichroism (CD) spectra were collected with a JASCO J-815 CD spectrometer, using a scan rate of 200 nm/min (spectral window 200-600 nm). 20 scans were collected. Spectra manager II (JASCO) software was used to analyze data.

Affinity Measurements

The peptide solutions were prepared by dissolving the peptides in 20 mM Gly-Gly, pH 8.8 inside a glovebox to a total concentration of cysteine plus histidine of 2.5 mM. After sequential addition of the chosen metal (Fe^{2+}, Co^{2+}), UV-vis spectra were collected and fit with GraphPad Prism v. 6.00 (GraphPad Software, La Jolla Californiaa USA) to the following equation: y=BMax*xh/(Kdh + xh), where BMax was the absorbance reached at saturation and h was the Hill slope. For Zn^{2+}, a competition titration against Co^{2+} was used. The Co^{2+} complex was formed at the K_d prior to titration with Zn^{2+}, and the absorption fit to the revised Cheng-Prusoff equation.

Paramagnetic NMR spectroscopy

Samples were prepared in a glovebox, as described above, and 700 µL placed in 5 mm NMR tubes (Sigma-Aldrich) with 10% D_2O. Spectra were recorded at room temperature using a VNMRS four-channel 600 MHz Varian spectrometer (5000 scans, 0.4 s recycle delay, 0.25 s acquisition time).

Cyclic Voltammetry

Experiments were conducted under ambient conditions, using an electrochemical workstation CHI660A (CH Instruments). Samples were prepared inside a glovebox and then transferred into a custom-made electrochemical cell. All solutions were purged with argon for 10 min before analysis, to ensure that no O2 was trapped during the transfer from the glovebox to the cell. A standard three

electrode cell composed of a glassy carbon working electrode (BASi, 3 mm diameter), an Ag/AgCl reference electrode (1 M KCl) (calibration curve Figure S37), and a platinum wire counter electrode was used for the electrochemical measurements. A scan rate from 50 mVs to 500 mVs was used. No supporting electrolyte was used to avoid interference with stability of the cluster; however, a small amount of NaCl was present from pH adjustment with NaOH and HCl. The working electrode was cleaned with a 0.05 µm aluminum oxide slurry on a polishing cloth and thoroughly rinsed with milliQ H_2O before each experiment. The reference electrode was sonicated in milliQ H_2O for 5 min before each experiment.

3.5 References

1. Venkateswara Rao, P. & Holm, R. H. Synthetic Analogues of the Active Sites of Iron-Sulfur Proteins. *Chem. Rev.* **104**, 527–559 (2004).
2. Ballmann, J. *et al.* A Synthetic Analogue of Rieske-Type [2Fe-2S] Clusters. *Angew. Chemie Int. Ed.* **47**, 9537–9541 (2008).
3. Qi, W. *et al.* Glutathione Complexed Fe–S Centers. *J. Am. Chem. Soc.* **134**, 10745–10748 (2012).
4. Bonfio, C. *et al.* UV-light-driven prebiotic synthesis of iron-sulfur clusters. *Nat. Chem.* **9**, 1229–1234 (2017).
5. Scintilla, S. *et al.* Duplications of an iron–sulphur tripeptide leads to the formation of a protoferredoxin. *Chem. Commun.* **52**, 13456–13459 (2016).
6. 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).
7. Bonfio, C. *et al.* Prebiotic iron–sulfur peptide catalysts generate a pH gradient across model membranes of late protocells. *Nat. Catal.* **1**, 616–623 (2018).
8. Belmonte, L. & Mansy, S. S. Patterns of Ligands Coordinated to Metallocofactors Extracted from the Protein Data Bank. *J. Chem. Inf. Model.* **57**, 3162–3171 (2017).
9. Hunsicker-Wang, L. M. *et al.* High-resolution structure of the soluble, respiratory-type Rieske protein from Thermus thermophilus: Analysis and comparison. *Biochemistry* **42**, 7303–7317 (2003).
10. Valer, L. *et al.* Methods to identify and characterize iron-sulfur oligopeptides in water. *Can. J. Chem.* (2022) doi:10.1139/cjc-2021-0237.
11. Bonfio, C. The curious case of peptide-coordinated iron-sulfur clusters: prebiotic and biomimetic insights. *Dalt. Trans.* (2021) doi:10.1039/x0xx00000x.
12. Bak, D. W. & Elliott, S. J. Alternative FeS cluster ligands: tuning redox potentials and chemistry. *Curr. Opin. Chem. Biol.* **19**, 50–58 (2014).
13. Brown, E. N. *et al.* Determining Rieske cluster reduction potentials. *J. Biol. Inorg. Chem.* **13**, 1301–1313 (2008).
14. Klingen, A. R. & Ullmann, G. M. Negatively charged residues and hydrogen bonds tune the ligand histidine pKa values of rieske iron-sulfur proteins. *Biochemistry* **43**, 12383–12389 (2004).

15. Zuris, J. A. *et al.* Engineering the redox potential over a wide range within a new class of FeS proteins. *J. Am. Chem. Soc.* **132**, 13120–13122 (2010).

16. Sivo, V. *et al.* Co(II) Coordination in Prokaryotic Zinc Finger Domains as Revealed by UV-Vis Spectroscopy. *Bioinorg. Chem. Appl.* **2017**, (2017).

17. Ueyama, N., Nakata, M., Fuji, M. A., Terakawa, T. & Nakamura, A. Analogues of Reduced Rubredoxin: Positive Shifts of Redox Potentials of Cysteine-Containing Peptide Iron(II) Complexes. *Inorg. Chem.* **24**, 2190–2196 (1985).

18. Sun, W. Y., Ueyama, N. & Nakamura, A. Reduced rubredoxin models containing Z-Cys-Pro-Leu-Cys-Gly-NH-C6H4-p-X (X = MeO, H, F, CN): electronic influence by a distant para substituent through NH---S hydrogen bonds. *Inorg. Chem.* **30**, 4026–4031 (1991).

19. Sugiura, Y., Ishizu, K. & Kimura, T. A rubredoxin-like mononuclear FeS4 derivative of adrenal iron-sulfur protein (adrenodoxin). *Biochem. Biophys. Res. Commun.* **60**, 334–340 (1974).

20. Hamed, M. Y., Silver, J. & Wilsonx, M. T. Studies of the reactions of ferric iron with glutathione and some related thiols. *Inorganica Chim. Acta* **78**, 1–11 (1983).

21. Werth, M. T. & Donald M. Kurtz, J. 1H NMR Spectra of Rubredoxins: New Resonances Assignable to a-CH and B-CH2, Hydrogens of Cysteinate Ligands to Iron(II). *Am. Chem. Soc.* **109**, 273–275 (1987).

22. Bonfio, C. *et al.* UV-light-driven prebiotic synthesis of iron–sulfur clusters. *Nat. Chem.* **9**, 1229–1234 (2017).

23. Holz, R. C., Small, F. J. & Ensign, S. A. Proton nuclear magnetic resonance investigation of the [2Fe-2S]1- - containing 'Rieske-type' protein from Xanthobacter strain Py2. *Biochemistry* **36**, 14690–14696 (1997).

24. Link, T. A. The Structures of Rieske and Rieske-Type Proteins. in *Advances in Inorganic Chemistry* vol. 47 83–157 (1999).

25. Link, T. A., Hagen, W. R., Pierik, A. J., Assmann, C. & von Jagow, G. Determination of the redox properties of the Rieske [2Fe-2S] cluster of bovine heart bc1 complex by direct electrochemistry of a water-soluble fragment,. *Eur. J. Biochem.* **208**, 685–691 (1992).

26. Parker, E. T. *et al.* Primordial synthesis of amines and amino acids in a 1958 Miller H 2S-rich spark discharge experiment. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 5526–5531 (2011).

27. Foden, C. S. *et al.* Prebiotic synthesis of cysteine peptides that catalyze peptide ligation in neutral water. *Science (80-.).* **370**, 865–869 (2020).

28. Nader, S., Sebastianelli, L. & Mansy, S. S. Protometabolism as out-of-equilibrium chemistry. *Philos Trans R Soc A* (2022) doi:10.1098/rsta.2020.0423.

29. Shen, C., Mills, T. & Oró, J. Prebiotic synthesis of histidine. *J. Mol. Evol.* **31**, 175–179 (1990).

30. Walton, T., Zhang, W., Li, L., Tam, C. P. & Szostak, J. W. The mechanism of nonenzymatic template copying with imidazole-activated nucleotides. *Angew. Chemie Int. Ed.* **58**, 10812–10819 (2019).

Chapter 4 - Final remarks

4.1 Conclusions

Life as we know it is completely dependent on metal catalyst. It is possible that of the many potential reaction networks that could have emerged, modern day metabolism reflects, at least in part, those reactions that were accelerated by metal ions. It has also been showed that Fe^{2+} ions promote an abiotic chemical pathway that resembles the Krebs cycle.[1] However, such systems do not offer an explanation on how they could have evolved over time. Calvin noted that the ability of free Fe^{3+} to catalyse the degradation of hydrogen peroxide to water and oxygen was increased 1,000-fold when the iron was coordinated to porphyrin.[2] This example illustrates how the intrinsic activity of a metal center can be augmented by the coordination. If peptides capable of coordinating the available metal ions were present on the early Earth then the reactivity of the metal center would have been modified in a sequence-dependent manner.[3] It is still not clear how metabolic-like reactions could have been harnessed by protocellular structures. It is, therefore, important to experimentally explore and evaluate the metal ion binding ability of amino acid residues other than cysteine inside model prebiotic peptides to gain insight into what was and was not possible.

The data previously published showed that iron-sulfur cluster could have been synthesized on the early Earth starting from three simple and readily available ingredients: thiol containing molecules, ferric ions and UV-light.[4] Sunlight represents one of the Earth's largest and earliest energy sources and has been implicated in the prebiotic synthesis of amino acids, sugars, nucleotides, lipid precursors.[5] Iron-sulfur proteins are known to be excellent catalysts and such functions may have played a crucial role in shaping the chemistry that led to life. 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 catalyze one-electron transfer reactions exploiting the redox couple Fe^{2+}/Fe^{3+}. While there is extensive studies on the activity of iron-sulfur proteins, unfortunately research, beside redox potential measurements, has not been conducted extensively on the reactivity of iron-sulfur cluster coordinated by small ligands. However an example of the activity they might have performed on the early Earth has been shown: small prebiotically plausible iron-sulfur clusters coordinated by a tripeptide are able to accept and donate electrons, and are able to establish a pH gradient if the reaction is confined a protocell.[6]

This work attempted to shed light, by comparing spectroscopic data or iron-sulfur protein and peptides, on how short peptides containing both cysteine and histidine could have coordinated the Rieske-like iron-sulfur clusters long before the emergence of proteins and enzymes currently known to contain such prosthetic groups. The data obtained throughout this project suggest that a

Rieske-like iron-sulfur cluster can be indeed coordinated by short and prebiotically plausible peptides. Moreover the [2Fe-2S] cluster coordinated by cysteine and histidine display a shift in the redox potential compared to the one of all-cysteine coordinated cluster. Such differences over long periods of time might have guided the early evolution from short peptides to more complex protein scaffolds, selecting the motifs more suited for specific redox processes.

4.2 Future perspective

Prebiotically plausible iron-sulfur peptides can be studied to reproduce reactions, catalyzed by iron-sulfur proteins, in a prebiotically plausible way in order to determine the chemical roots of protometabolism. Iron-sulfur cluster are known to mediate many redox processes inside the cells, and during this work while we have evaluated the formal redox potential of non-biological Rieske-like iron-sulfur cluster, we have not explored yet the catalytic activity of such metallopeptides on biologically relevant molecules such as NADH or ubiquinol. After exploring the kind of substrate and reactions which might have utilized prebiotic iron-sulfur cluster. An interesting progression of the project would be to study the possibility to create an electron transfer based on a system composed by Rieske peptides and other iron-sulfur complexes, as the one described by Bonfio.[6] Metallopeptides could be anchored to vesicle membrane through hydrophobic residues to investigate the possible enhancement of electron transfer in an attempt to reconstruct the first example of plausible prebiotic metabolic pathway.

Another possible step forward in the study of prebiotic iron-sulfur cluster is a further expansion of the coordinating residues. In this work we have probed the ability of histidine to coordinate metal centers inside proteins and peptides however in the future it could be worth exploring how other possible plausible residues, such as aspartic and glutamic acid, might influence the chemicals properties of prebiotic iron-sulfur peptides. Or even the removal of one of the coordinating residues may produce an incomplete coordination. A vacant position on one iron ion might leave the space for possible substrate as in some modern enzymes, such as aconitase.[7]

4.3 References

1. Muchowska, K. B., Varma, S. J. & Moran, J. Synthesis and breakdown of universal metabolic precursors promoted by iron. *Nature* **569**, 104–107 (2019).
2. Calvin, M. Evolution of Enzymes and the Photosynthetic Apparatus: Primitive photochemistry and porphyrin catalysis after separate genesis join in modern photosynthesis. *Science (80-.).* **130**, 1170–1174 (1959).
3. Belmonte, L. & Mansy, S. S. Metal Catalysts and the Origin of Life. *Elements* **12**, 413–418 (2016).

4. Bonfio, C. *et al.* UV-light-driven prebiotic synthesis of iron-sulfur clusters. *Nat. Chem.* **9**, 1229–1234 (2017).

5. 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).

6. Bonfio, C. *et al.* Prebiotic iron–sulfur peptide catalysts generate a pH gradient across model membranes of late protocells. *Nat. Catal.* **1**, 616–623 (2018).

7. Meyer, J. Iron–sulfur protein folds, iron–sulfur chemistry, and evolution. *JBIC J. Biol. Inorg. Chem.* **13**, 157–170 (2008).

Appendix 1

Gene sequences

In this section are reported the gene sequence for the protein expressed and purified in Chapter 2.

Wild Type Rubredoxin sequence (cloned inside a pMAL-c4x vector)

ATGAAAAAATACACCTGCACCGTCTGTGGCTACATCTACAACCCGGAAGATGGCGACCCGGATAATGGCGTCAACCC
GGGCACCGATTTTAAAGATATTCCGGACGACTGGGTCTGTCCGCTGTGCGGCGTCGGTAAAGATCAATTTGAAGAAG
TTGAAGAA

C42H Rubredoxin sequence (cloned inside a pMAL-c4x vector)

ATGAAAAAATACACCTGCACCGTCTGTGGCTACATCTACAACCCGGAAGATGGCGACCCGGATAATGGCGTCAACCC
GGGCACCGATTTTAAAGATATTCCGGACGACTGGGTCTGTCCGCTGCATGGCGTCGGTAAAGATCAATTTGAAGAAG
TTGAAGAA

C9H C42H Rubredoxin sequence (cloned inside a pMAL-c4x vector)

ATGAAAAAATACACCTGCACCGTCCATGGCTACATCTACAACCCGGAAGATGGCGACCCGGATAATGGCGTCAACCC
GGGCACCGATTTTAAAGATATTCCGGACGACTGGGTCTGTCCGCTGCATGGCGTCGGTAAAGATCAATTTGAAGAAG
TTGAAGAA

Wild Type Ferredoxin sequence (cloned inside a pMAL-c4x vector)

ATGAGCTCCTCGGAAGACAAAATCACGGTCCATTTTATCAACCGCGACGGCGAAACGCTGACGACGAAAGGCAAAGT
GGGCGACTCGCTGCTGGATGTGGTTGTCGAAAACAATCTGGATATTGACGGCTTTGGTGCGTGCGAAGGCACCCTG
GCCTGCAGTACGTGTCATCTGATTTTCGAAGATCACATCTATGAAAAACTGGATGCAATTACCGACGAAGAAAACGAT
ATGCTGGACCTGGCGTACGGCCTGACGGATCGTAGCCGCCTGGGTTGCCAGATCTGTCTGACCAAATCTATGGACA
ATATGACGGTTCGTGTCCCGGAAACGGTCGCTGACGCCCGCCAGAGCATTGATGTCGGTAAAACCTCGTAA

Wild Type Rieske Protein sequence (cloned inside a pMAL-c4x vector)

ATGACCCCGGAGAAGGAACCGCTGAAACCGGGCGACATCCTGGTTTATGCGCAGGGTGGCGGTGAGCCGAAGCCG
ATTCGTCTGGAGGAACTGAAACCGGGCGACCCGTTTGTGCTGGCGTATCCGATGGACCCGAAGACCAAAGTGGTTA
AGAGCGGTGAAGCGAAAAACACCCTGCTGGTTGCGCGTTTTGATCCGGAGGAACTGGCGCCGGAAGTGGCGCAGC
ATGCGGCGGAAGGCGTGGTTGCGTACAGCGCGGTTTGCACCCACCTGGGTTGCATCGTGAGCCAATGGGTTGCGG
ATGAGGAAGCGGCGCTGTGCCCGTGCCATGGCGGTGTGTATGATCTGCGTCACGGCGCGCAGGTTATTGCGGGTC
CGCCGCCGCGTCCGGTGCCGCAACTGCCGGTGCGTGTTGAGGACGGCGTGCTGGTTGCGGCGGGTGAATTCCTGG
GTCCGGTGGGTGTTCAAGCGAGCGCGGGTGCGTACACCTGGCGTGTGTAA

H90C Rieske Protein sequence (cloned inside a pMAL-c4x vector)

ATGACCCCGGAGAAGGAACCGCTGAAACCGGGCGACATCCTGGTTTATGCGCAGGGTGGCGGTGAGCCGAAGCCG
ATTCGTCTGGAGGAACTGAAACCGGGCGACCCGTTTGTGCTGGCGTATCCGATGGACCCGAAGACCAAAGTGGTTA
AGAGCGGTGAAGCGAAAAACACCCTGCTGGTTGCGCGTTTTGATCCGGAGGAACTGGCGCCGGAAGTGGCGCAGC
ATGCGGCGGAAGGCGTGGTTGCGTACAGCGCGGTTTGCACCTGCCTGGGTTGCTGCGTGAGCCAATGGGTTGCGG
ATGAGGAAGCGGCGCTGTGCCCGTGCCATGGCGGTGTGTATGATCTGCGTCACGGCGCGCAGGTTATTGCGGGTC
CGCCGCCGCGTCCGGTGCCGCAACTGCCGGTGCGTGTTGAGGACGGCGTGCTGGTTGCGGCGGGTGAATTCCTGG
GTCCGGTGGGTGTTCAAGCGAGCGCGGGTGCGTACACCTGGCGTGTGTAA

H90C H107C Rieske Protein sequence (cloned inside a pMAL-c4x vector)

ATGACCCCGGAGAAGGAACCGCTGAAACCGGGCGACATCCTGGTTTATGCGCAGGGTGGCGGTGAGCCGAAGCCG
ATTCGTCTGGAGGAACTGAAACCGGGCGACCCGTTTGTGCTGGCGTATCCGATGGACCCGAAGACCAAAGTGGTTA
AGAGCGGTGAAGCGAAAAACACCCTGCTGGTTGCGCGTTTTGATCCGGAGGAACTGGCGCCGGAAGTGGCGCAGC
ATGCGGCGGAAGGCGTGGTTGCGTACAGCGCGGTTTGCACCTGCCTGGGTTGCTGCGTGAGCCAATGGGTTGCGG
ATGAGGAAGCGGCGCTGTGCCCGTGCTGCGGCGGTGTGTATGATCTGCGTCACGGCGCGCAGGTTATTGCGGGTC
CGCCGCCGCGTCCGGTGCCGCAACTGCCGGTGCGTGTTGAGGACGGCGTGCTGGTTGCGGCGGGTGAATTCCTGG
GTCCGGTGGGTGTTCAAGCGAGCGCGGGTGCGTACACCTGGCGTGTGTAA

C93A C109A Rieske Protein sequence (cloned inside a pMAL-c4x vector)

ATGACCCCGGAGAAGGAACCGCTGAAACCGGGCGACATCCTGGTTTATGCGCAGGGTGGCGGTGAGCCGAAGCCG
ATTCGTCTGGAGGAACTGAAACCGGGCGACCCGTTTGTGCTGGCGTATCCGATGGACCCGAAGACCAAAGTGGTTA
AGAGCGGTGAAGCGAAAAACACCCTGCTGGTTGCGCGTTTTGATCCGGAGGAACTGGCGCCGGAAGTGGCGCAGC

ATGCGGCGGAAGGCGTGGTTGCGTACAGCGCGGTTTGCACCCACCTGGGTTGCGCGGTGAGCCAATGGGTTGCGG
ATGAGGAAGCGGCGCTGTGCCCGGCGCATGGCGGTGTGTATGATCTGCGTCACGGCGCGCAGGTTATTGCGGGTC
CGCCGCCGCGTCCGGTGCCGCAACTGCCGGTGCGTGTTGAGGACGGCGTGCTGGTTGCGGCGGGTGAATTCCTGG
GTCCGGTGGGTGTTCAAGCGAGCGCGGGTGCGTACACCTGGCGTGTGTAA

H90C C93A C109A Rieske Protein sequence (cloned inside a pMAL-c4x vector)

ATGACCCCGGAGAAGGAACCGCTGAAACCGGGCGACATCCTGGTTTATGCGCAGGGTGGCGGTGAGCCGAAGCCG
ATTCGTCTGGAGGAACTGAAACCGGGCGACCCGTTTGTGCTGGCGTATCCGATGGACCCGAAGACCAAAGTGGTTA
AGAGCGGTGAAGCGAAAAACACCCTGCTGGTTGCGCGTTTTGATCCGGAGGAACTGGCGCCGGAAGTGGCGCAGC
ATGCGGCGGAAGGCGTGGTTGCGTACAGCGCGGTTTGCACCTGCCTGGGTTGCGCGGTGAGCCAATGGGTTGCGG
ATGAGGAAGCGGCGCTGTGCCCGGCGCATGGCGGTGTGTATGATCTGCGTCACGGCGCGCAGGTTATTGCGGGTC
CGCCGCCGCGTCCGGTGCCGCAACTGCCGGTGCGTGTTGAGGACGGCGTGCTGGTTGCGGCGGGTGAATTCCTGG
GTCCGGTGGGTGTTCAAGCGAGCGCGGGTGCGTACACCTGGCGTGTGTAA

H90C C93A H107C C109A Rieske Protein sequence (cloned inside a pMAL-c4x vector)

ATGACCCCGGAGAAGGAACCGCTGAAACCGGGCGACATCCTGGTTTATGCGCAGGGTGGCGGTGAGCCGAAGCCG
ATTCGTCTGGAGGAACTGAAACCGGGCGACCCGTTTGTGCTGGCGTATCCGATGGACCCGAAGACCAAAGTGGTTA
AGAGCGGTGAAGCGAAAAACACCCTGCTGGTTGCGCGTTTTGATCCGGAGGAACTGGCGCCGGAAGTGGCGCAGC
ATGCGGCGGAAGGCGTGGTTGCGTACAGCGCGGTTTGCACCTGCCTGGGTTGCGCGGTGAGCCAATGGGTTGCGG
ATGAGGAAGCGGCGCTGTGCCCGGCGTGCGGCGGTGTGTATGATCTGCGTCACGGCGCGCAGGTTATTGCGGGTC
CGCCGCCGCGTCCGGTGCCGCAACTGCCGGTGCGTGTTGAGGACGGCGTGCTGGTTGCGGCGGGTGAATTCCTGG
GTCCGGTGGGTGTTCAAGCGAGCGCGGGTGCGTACACCTGGCGTGTGTAA

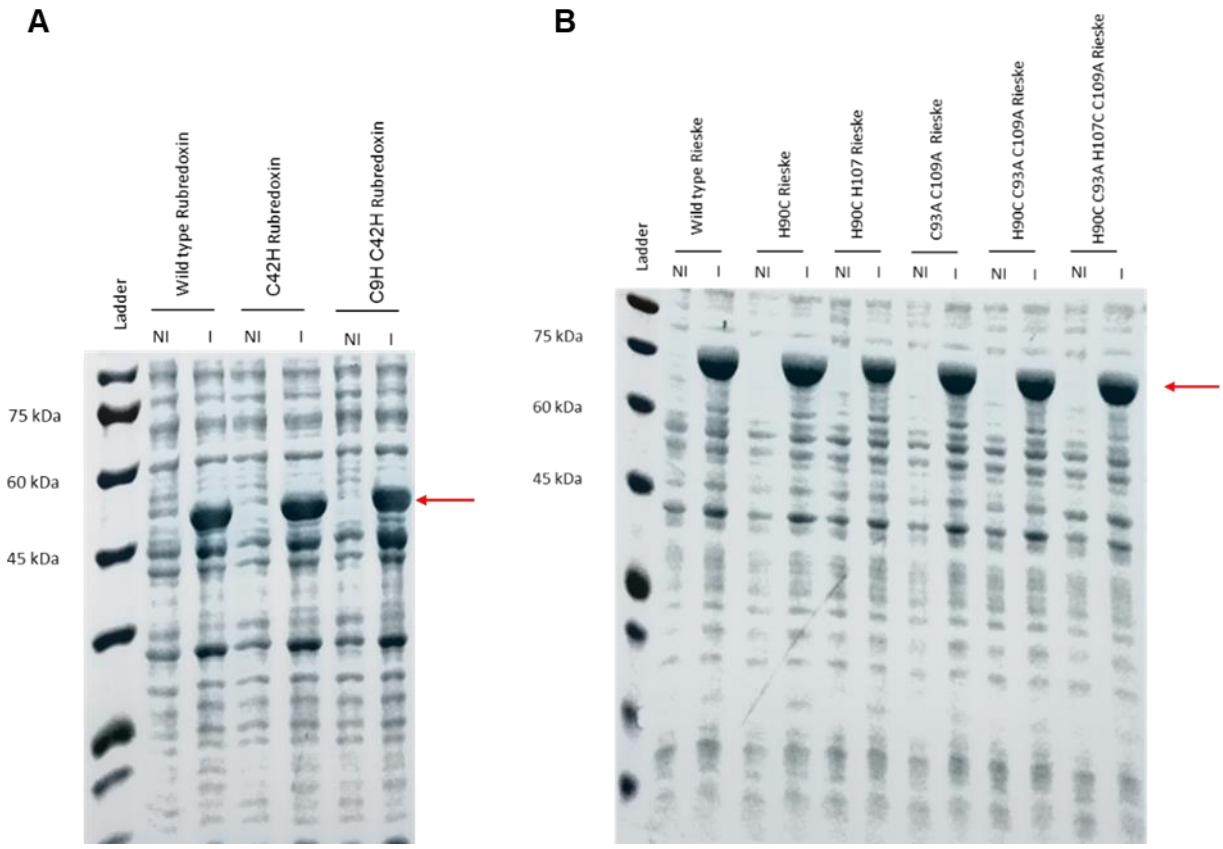

Figure S1.1 – SDS-page to check the expression of the iron-sulfur proteins cloned in pMAL-c4x vectors. **A**. Expression of WT rubredoxin and C42H, C9H C42H mutants (rubredoxin 6.1 kDa + Maltose Binding Protein 42.5 kDa). **B**. Expression of WT Rieske proteins and mutants (Rieske (17.5 kDa + Maltose Binding Protein 42.5 kDa). NI = non-induced, I = induced with 0.4 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG).

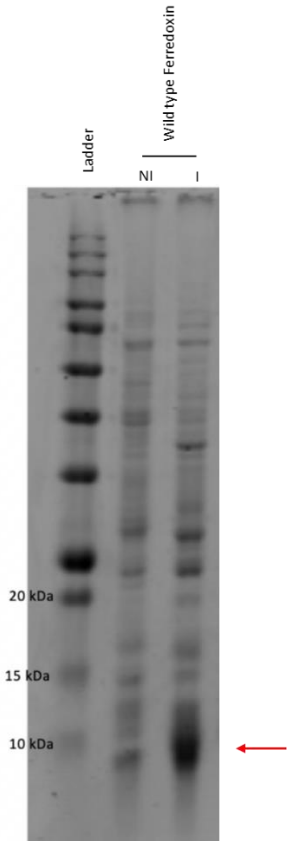

Figure S1.2 – SDS-page to check the expression of human ferredoxin (12 kDa)

Appendix 2

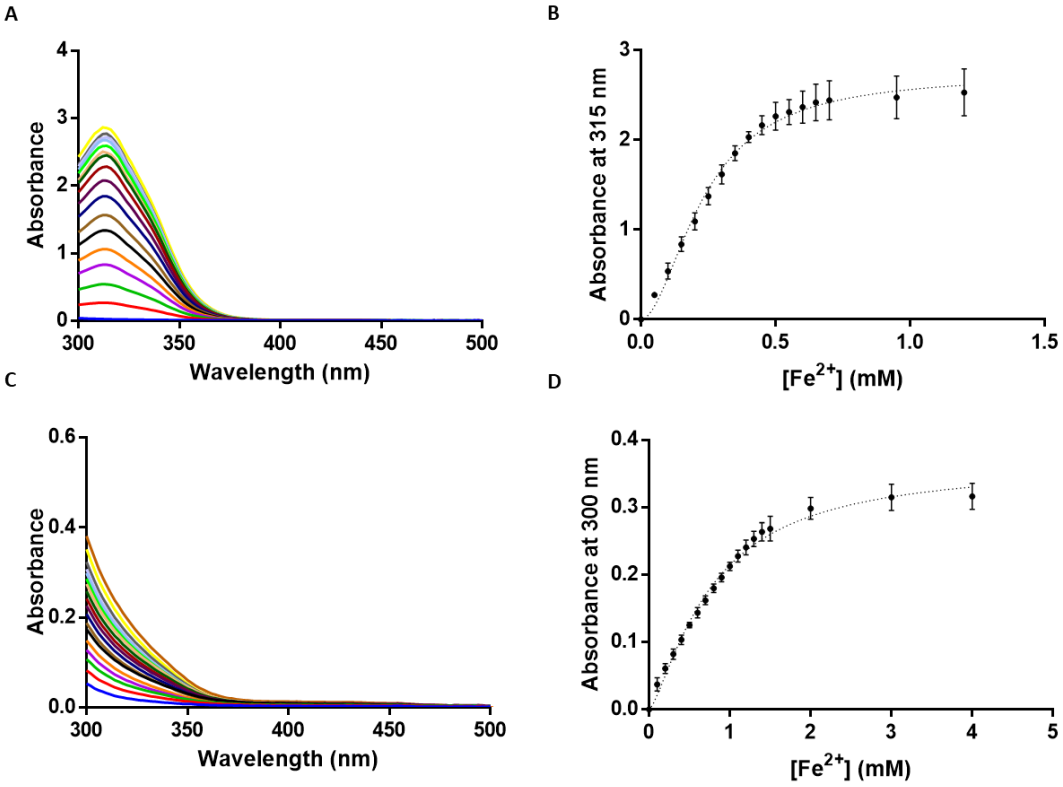

Figure S2.1 - Fe^{2+} affinity for GCTVCG and GCTVHG. **A**. Fe^{2+} titration of GCTVCG. **B**. Fitting of the Fe^{2+} titration of CGTVCG. **C**. Fe^{2+} titration of GCTVHG. **D**. Fitting of the Fe^{2+} titration of CGTVHG. Condition of the titrations were peptides 1.25 mM, Gly-Gly 20 mM, pH 8.8. Data represent mean and SD of distinct samples, n = 3 replicates.

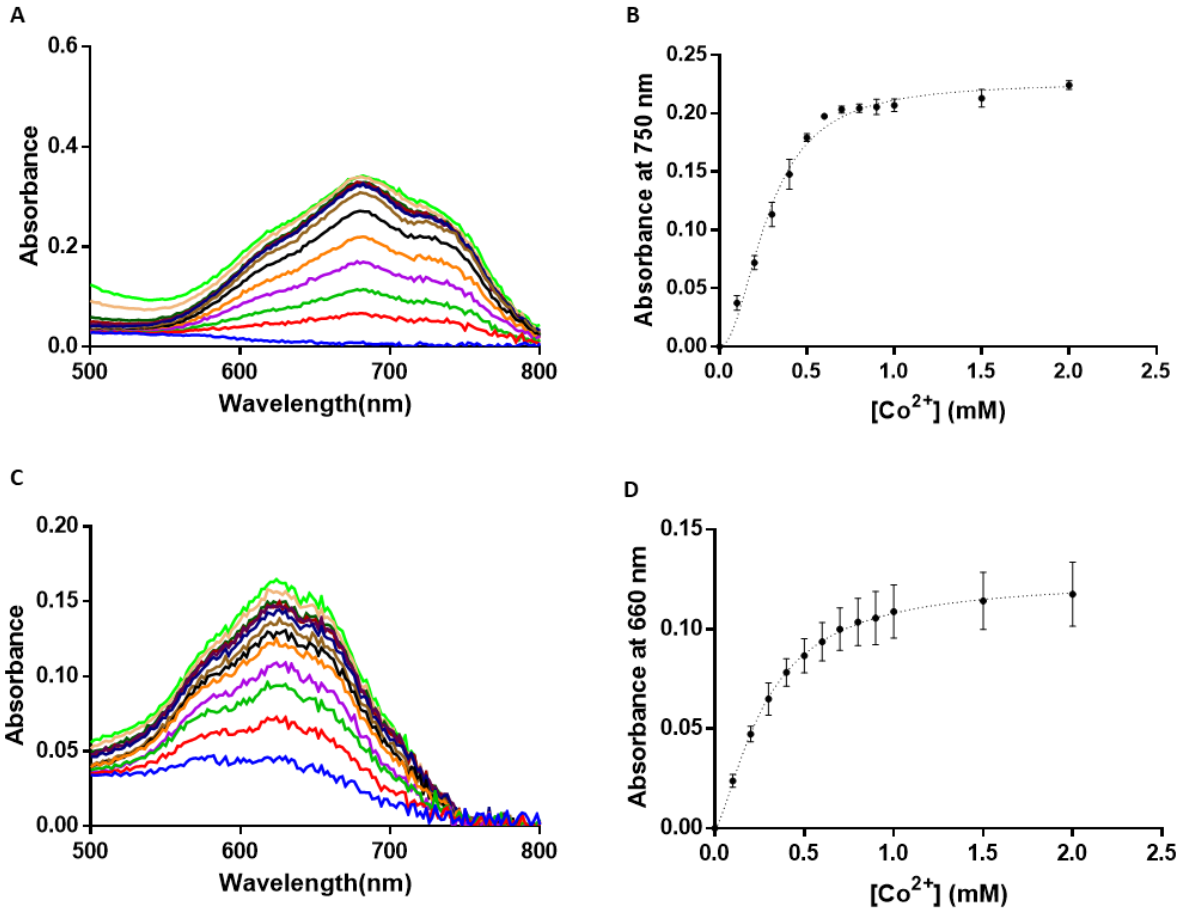

Figure S2.2 - Co^{2+} affinity for GCTVCG and GCTVHG. **A**. Co^{2+} titration of GCTVCG. **B**. Fitting of the Co^{2+} titration of CGTVCG. **C**. Co^{2+} titration of GCTVHG. **D**. Fitting of the Co^{2+} titration of CGTVHG. Condition of the titrations were peptides 1.25 mM, Gly-Gly 20 mM, pH 8.8. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

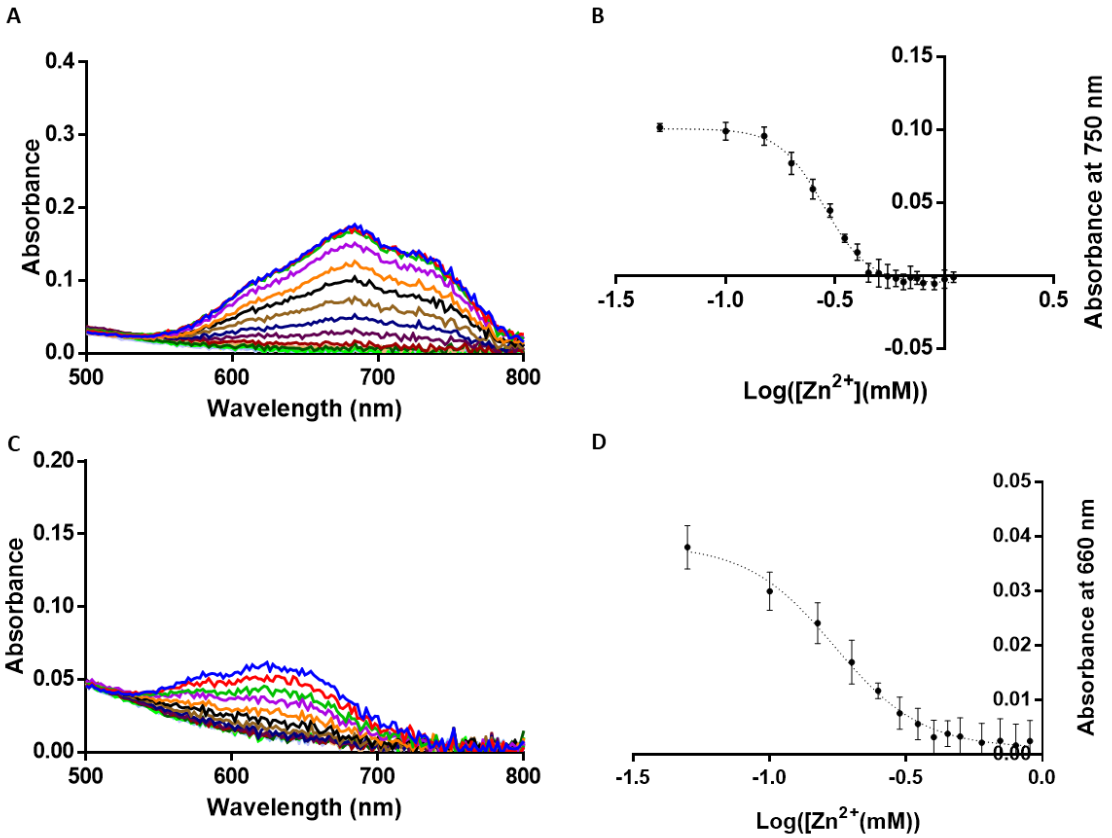

Figure S2.3 - Zn^{2+} affinity for GCTVCG and GCTVHG. **A**. Zn^{2+} titration of GCTVCG. **B**. Fitting of the Zn^{2+} titration of CGTVCG. **C**. Zn^{2+} titration of GCTVHG. **D**. Fitting of the Zn^{2+} titration of CGTVHG. Condition of the titrations were peptides 1.25 mM, Gly-Gly 20 mM, pH 8.8. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

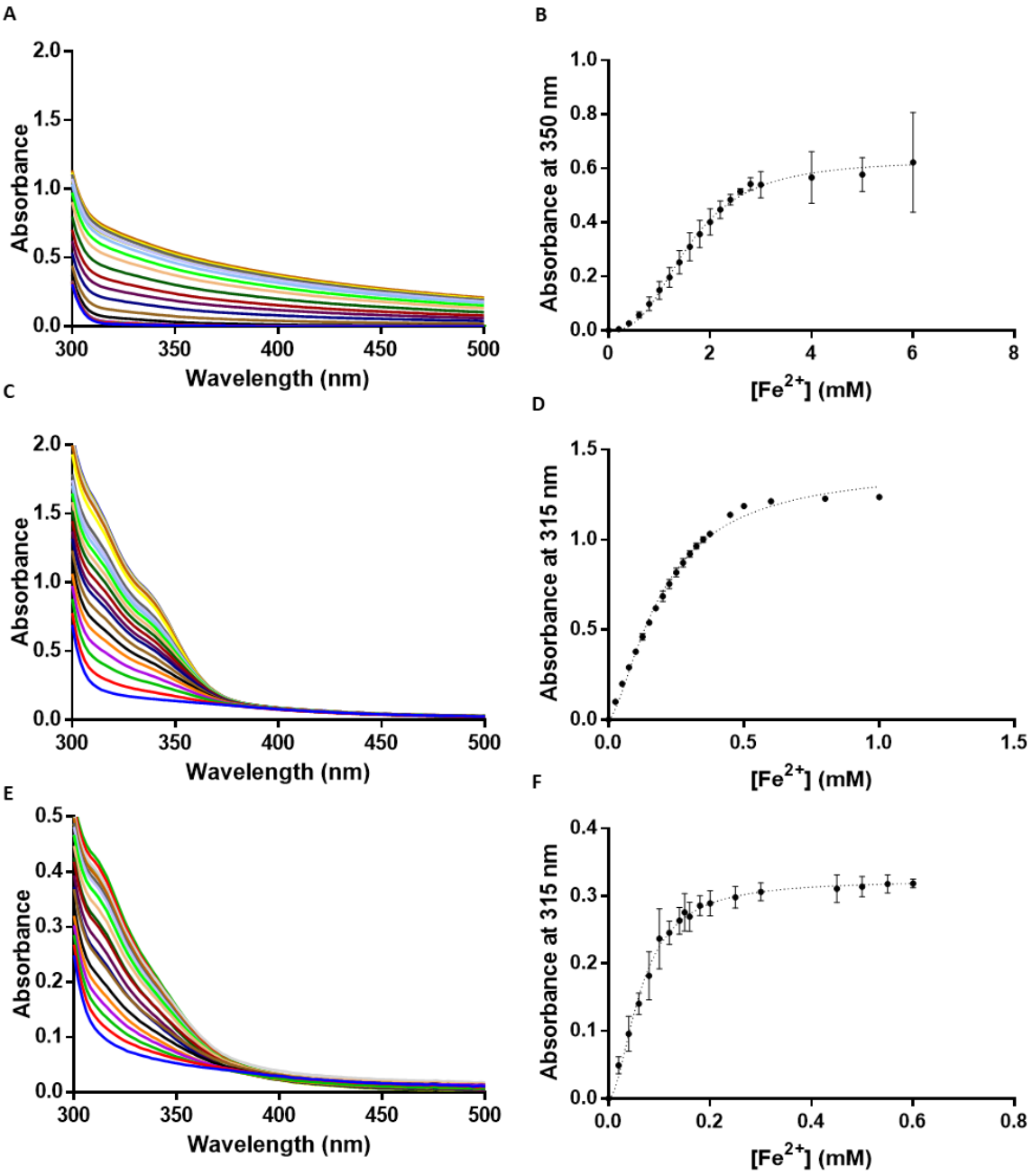

Figure S2.4 - Fe^{2+} affinity for the linear Rieske peptides VCTHLGAIVSQWVADEEAALCPAHG, VCTCLGAIVSQWVADEEAALCPAHG, and VCTCLGAIVSQWVADEEAALCPACG. A. Fe^{2+} titration of VCTHLGAIVSQWVADEEAALCPAHG. B. Fitting of the Fe^{2+} titration of VCTHLGAIVSQWVADEEAALCPAHG. C. Fe^{2+} titration of VCTCLGAIVSQWVADEEAALCPAHG. D. Fitting of the Fe^{2+} titration of VCTCLGAIVSQWVADEEAALCPAHG. E. Fe^{2+} titration of VCTCLGAIVSQWVADEEAALCPACG. F. Fitting of the Fe^{2+} titration of VCTCLGAIVSQWVADEEAALCPACG. Condition of the titrations were peptides 1.25 mM, Gly-Gly 20 mM, pH 8.8. Data represent mean and SD of distinct samples, n = 3 replicates.

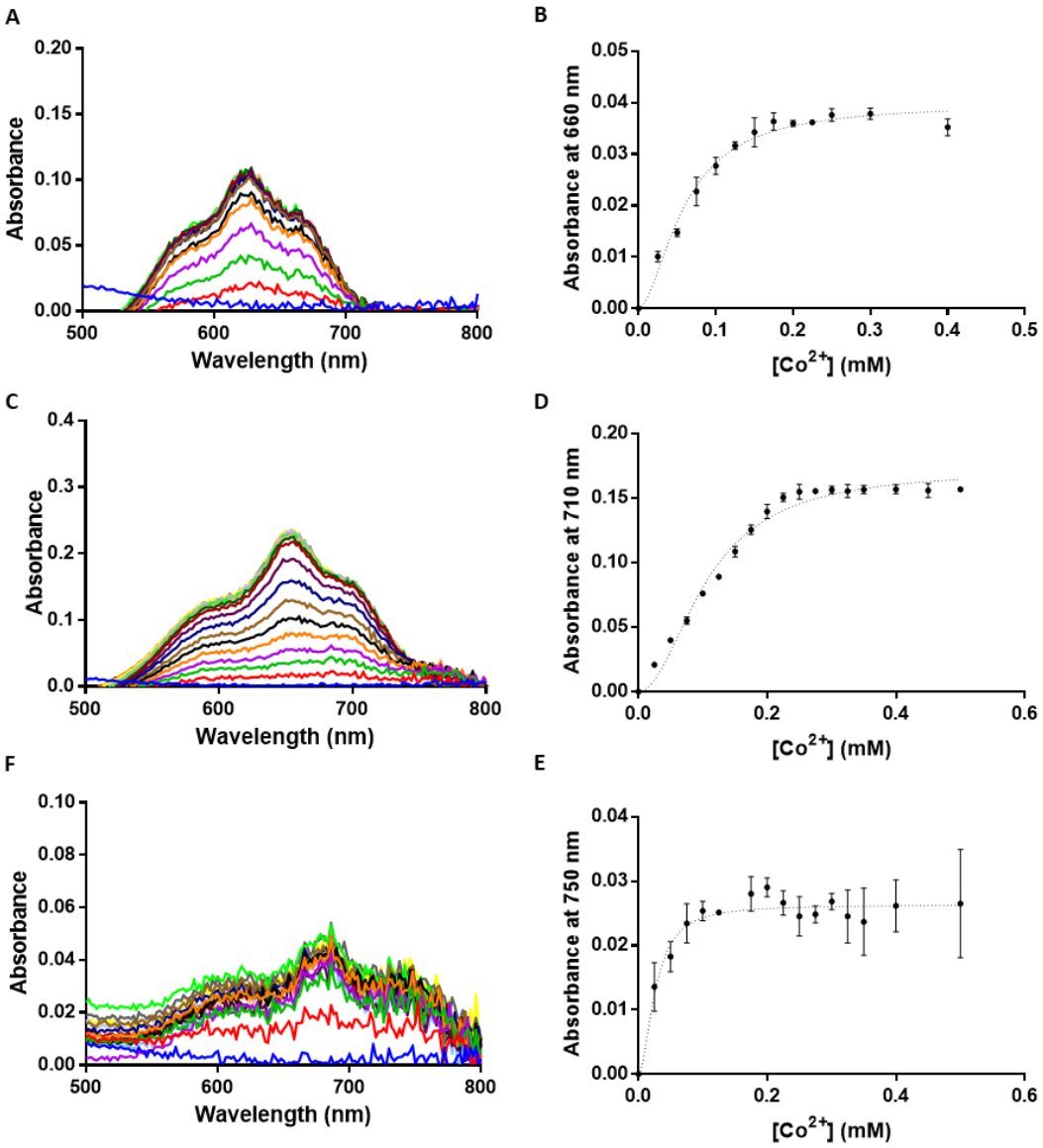

Figure S2.5 - Co^{2+} affinity for small linear Rieske peptides: V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G, V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G, and V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G. **A**. Co^{2+} titration of V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G. **B**. Fitting of the Co^{2+} titration of V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G. **C**. Co^{2+} titration of V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G. **D**. Fitting of the Co^{2+} titration of V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G. **E**. Co^{2+} titration of V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G. **F**. Fitting of the Co^{2+} titration of V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G. Condition of the titrations were peptides 0.75 mM, Gly-Gly 20 mM, pH 8.8. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

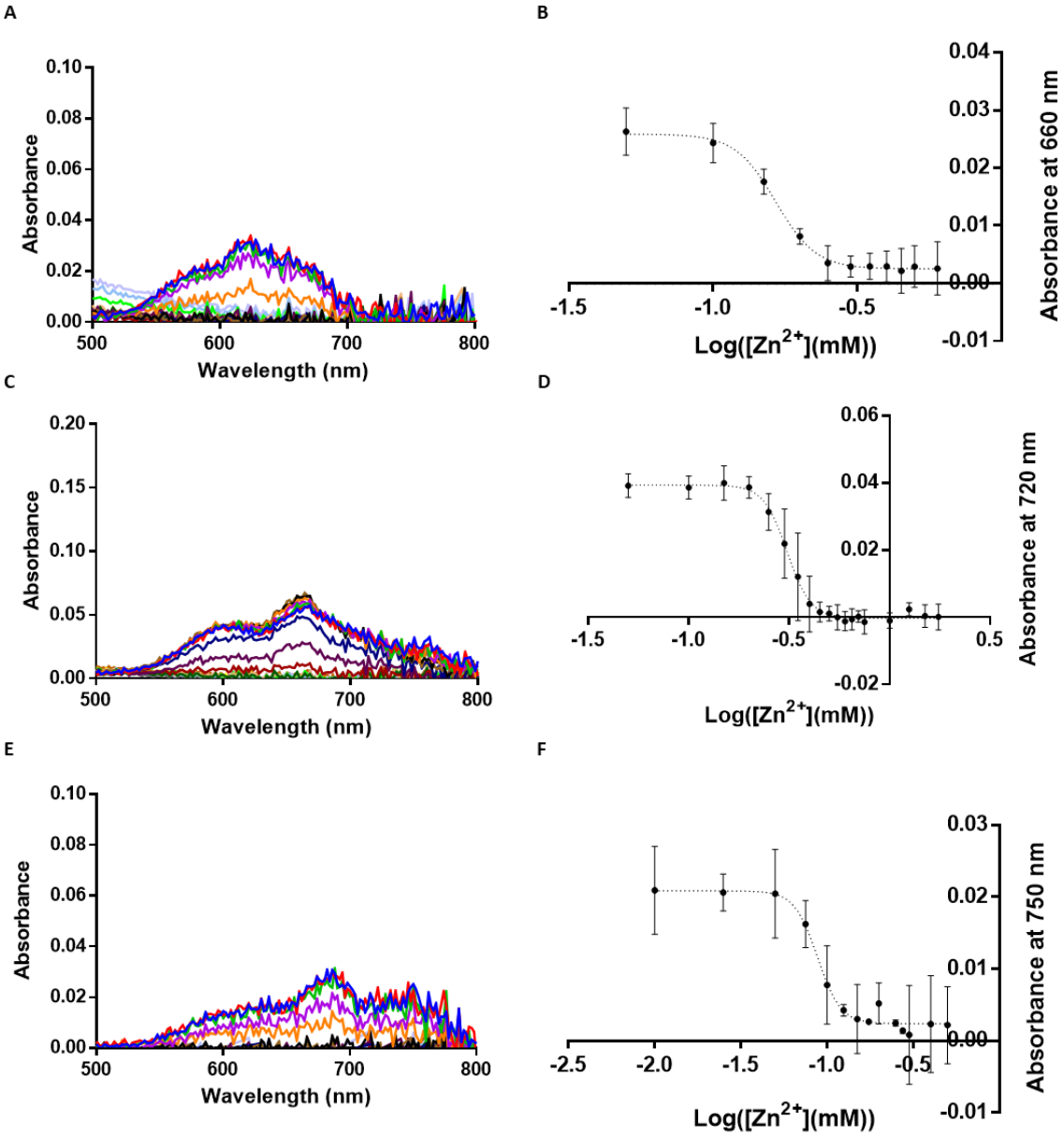

Figure S2.6 - Zn^{2+} affinity for linear Rieske peptides: V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G, V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G, and V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G. **A**. Zn^{2+} titration of V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G. **B**. Fitting of the Zn^{2+} titration of V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G. **C**. Zn^{2+} titration of V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G. **D**. Fitting of the Zn^{2+} titration of V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G. **E**. Zn^{2+} titration of V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G. **F**. Fitting of the Zn^{2+} titration of V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G. Condition of the titrations were peptides 0.75 mM, Gly-Gly 20 mM, pH 8.8. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

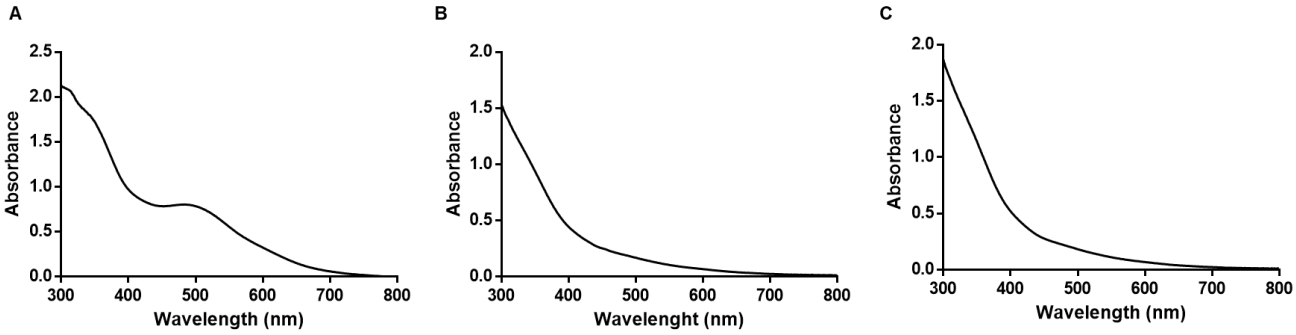

Figure S2.7 - UV-vis spectra of $[1Fe\text{-}0S]^{3+}$ centers coordinated to the hexapeptides GCPLCG (**A**), GCPLHG (**B**), and GCPAHG (**C**). Conditions were 4 mM peptide, 0.5 mM $FeCl_3$, pH 8.8.

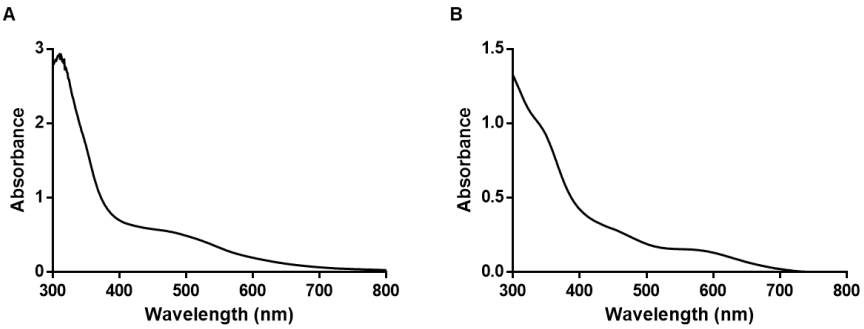

Figure S2.8 - UV-vis spectra of $[1Fe\text{-}0S]^{3+}$ centers coordinated to the pentapeptides GCTCG (**A**) and GCTHG (**B**). Solutions contained 4 mM peptide, 0.5 mM $FeCl_3$, pH 8.8.

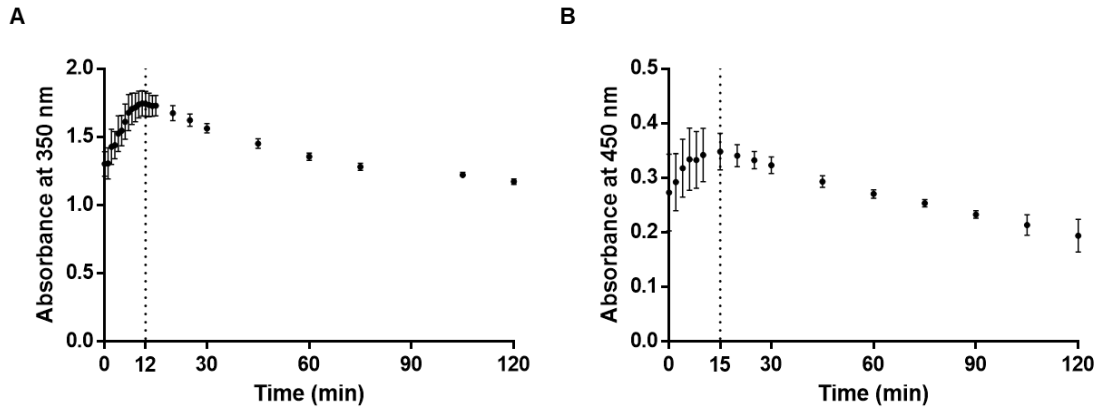

Figure S2.9 - Stability over time of $[1Fe\text{-}0S]^{3+}$ center coordinated by GCPAHC (A) and GCPLHG (B). Conditions were 4 mM peptide, 0.5 mM FeCl3, pH 8.8. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

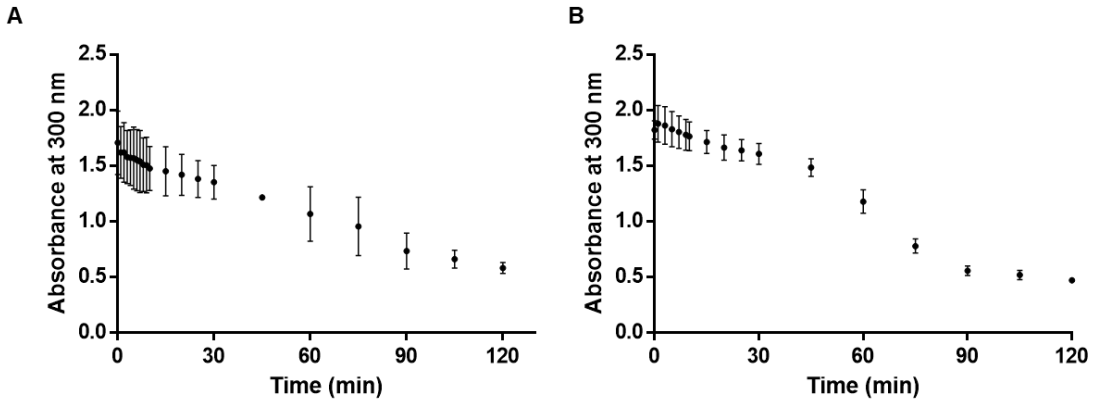

Figure S2.10 - Stability over time of $[1Fe\text{-}0S]^{3+}$ center coordinated by GCPAHC (A) and GCPLHG (B). Conditions were 4 mM peptide, 0.5 mM FeCl3, pH 8.8. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

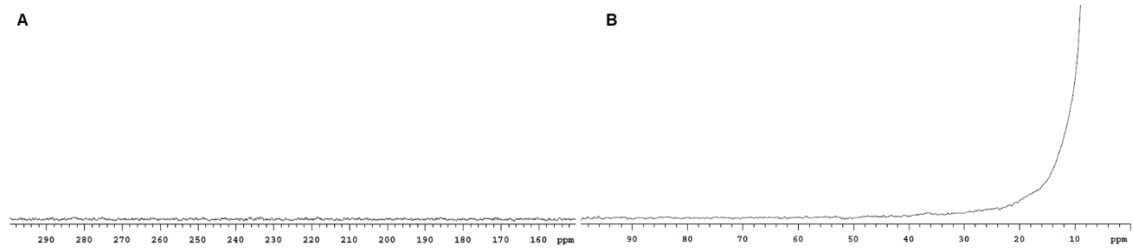

Figure S2.11 - Paramagnetic NMR spectra of $[1Fe\text{-}0S]^{3+}$ clusters coordinated by GCTVHG. Upfield (**A**) and downfield (**B**) region. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8.

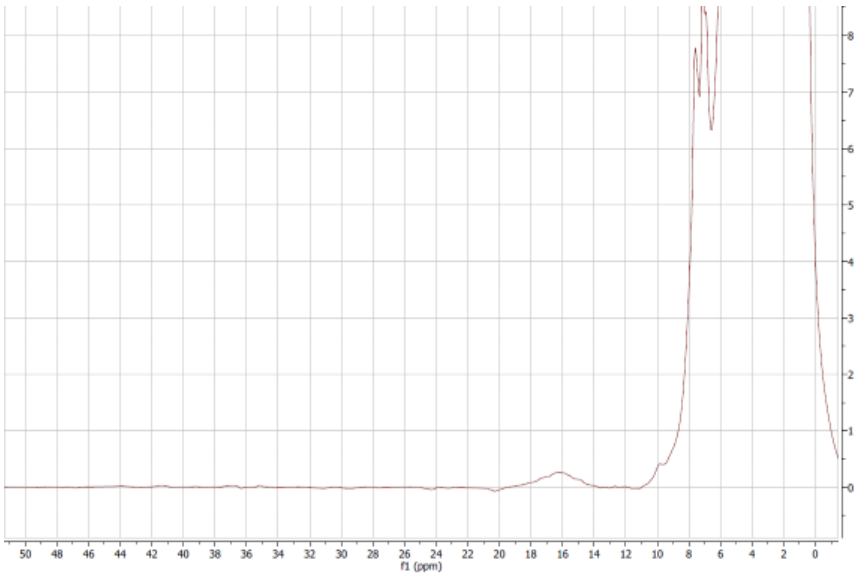

Figure S2.12 - Paramagnetic ^1H NMR spectrum of [2Fe-2S] V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G (8 mM Linear Rieske (CHCH), 2 mM $FeCl_3$, 2 mM Na_2S pH 8.8).

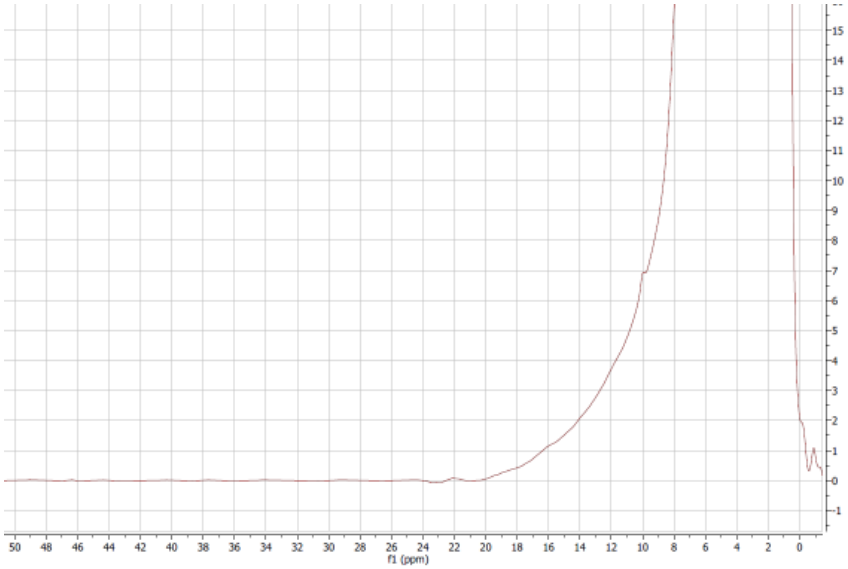

Figure S2.13 - Paramagnetic ^{1}H NMR spectrum of [2Fe-2S] V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G (8 mM Linear Rieske (CCCH), 2 mM $FeCl_3$, 2 mM Na_2S pH 8.8).

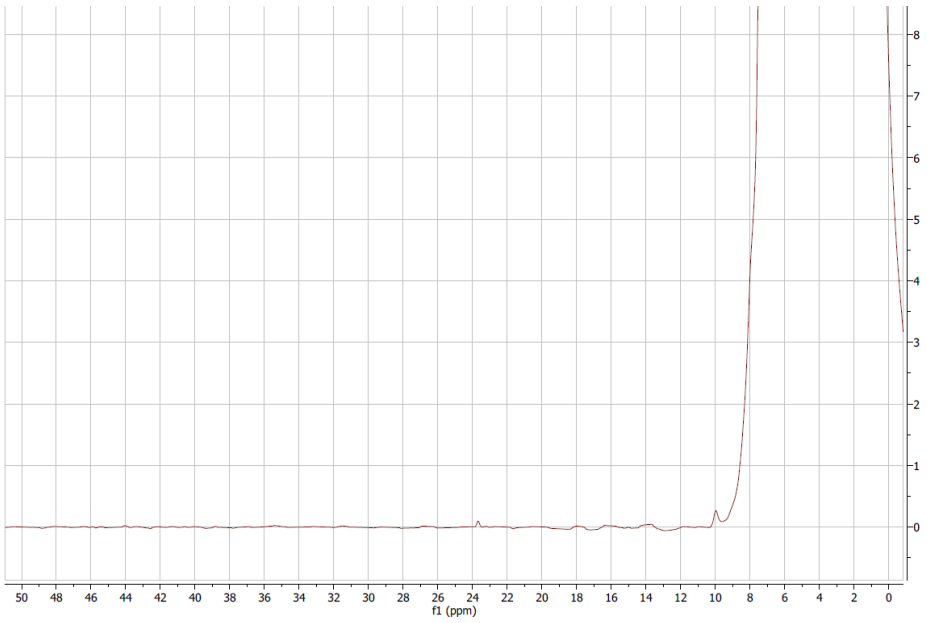

Figure S2.14 - Paramagnetic ^{1}H NMR spectrum of [2Fe-2S] V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G (8 mM Linear Rieske (CCCC), 2 mM $FeCl_3$, 2 mM Na_2S pH 8.8).

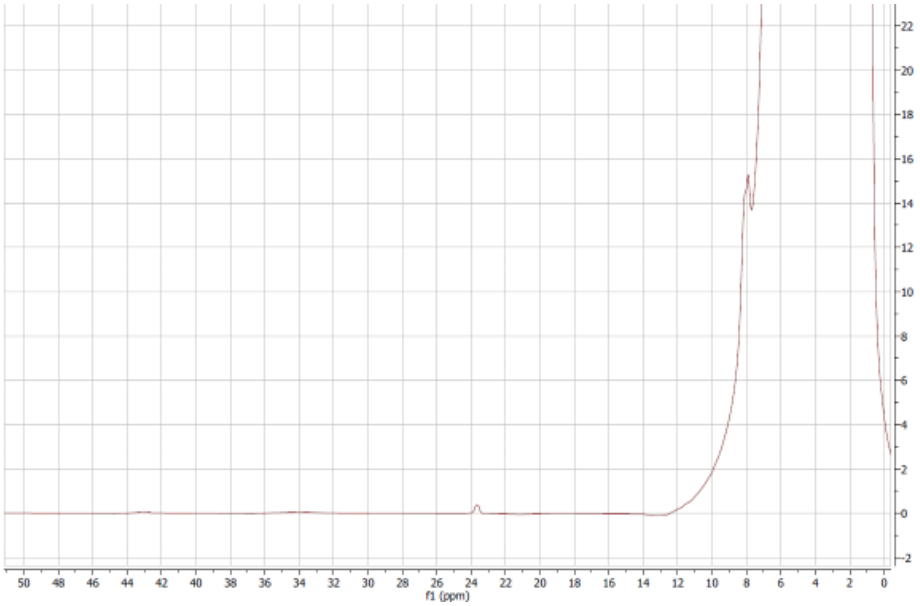

Figure S2.15 - Paramagnetic ^{1}H NMR spectrum of $[2Fe\text{-}2S]^{2+}$ V**C**T**H**LGDIVSL**C**PK**H**G (2 mM Stapled Rieske (CHCH), 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8).

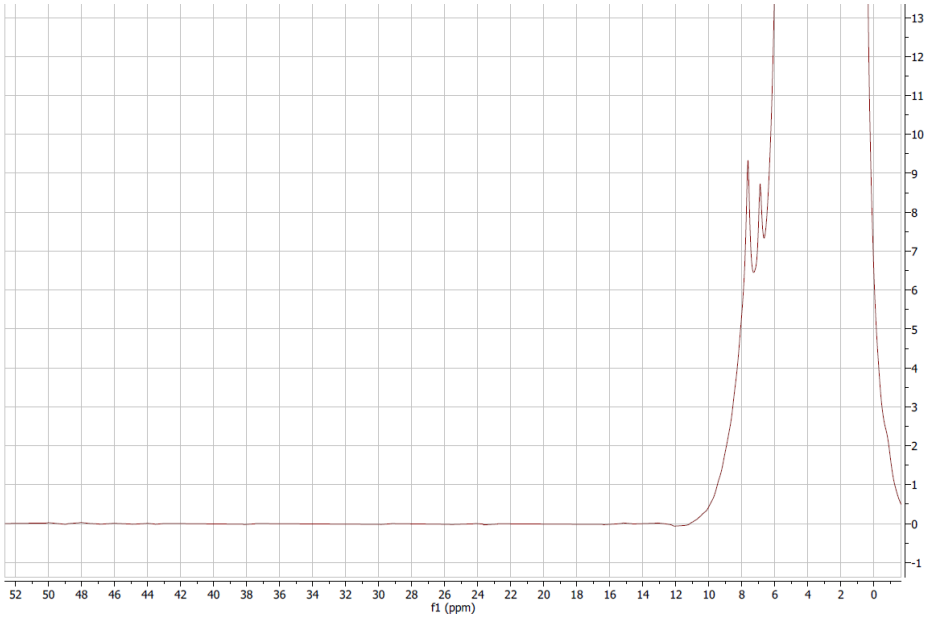

Figure S2.16 - Paramagnetic ^{1}H NMR spectrum of $[2Fe\text{-}2S]^{2+}$ V**C**T**C**LGDIVSL**C**PK**H**G (2 mM Stapled Rieske (CCCH), 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8).

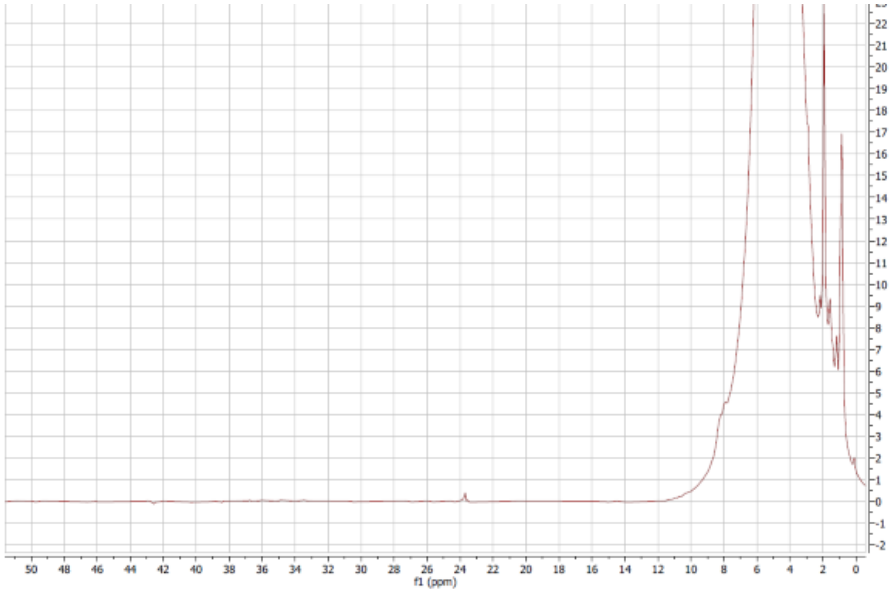

Figure S2.17 - Paramagnetic ^{1}H NMR spectrum of $[2Fe\text{-}2S]^{2+}$ V**C**T**C**LGDIVSL**C**PK**C**G (2 mM Stapled Rieske (CCCC), 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8).

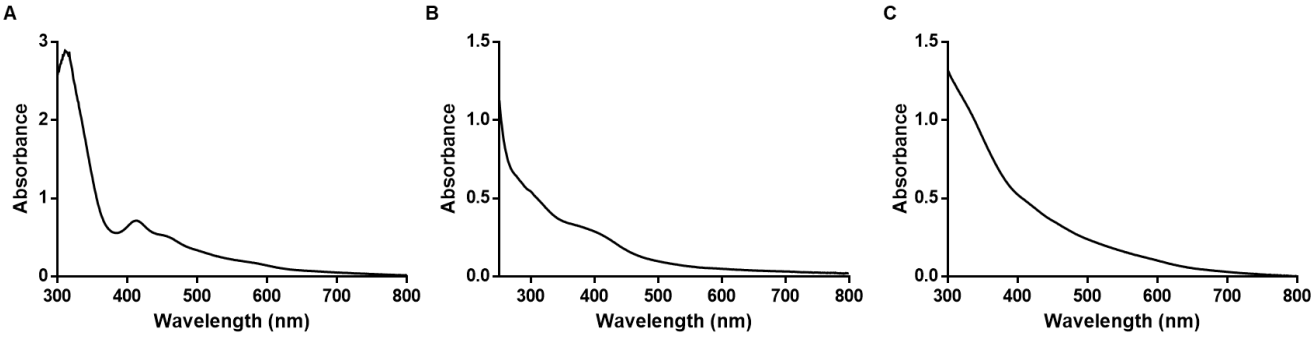

Figure 2.18 - UV- vis spectra of iron-sulfur hexapeptides. **A.** $[2Fe\text{-}2S]^{3+}$ GCPLCG (4 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8). **B**. $[2Fe\text{-}2S]^{3+}$ GCPLHG (4 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8). **C**. $[2Fe\text{-}2S]^{3+}$ GCPAHG (4 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8).

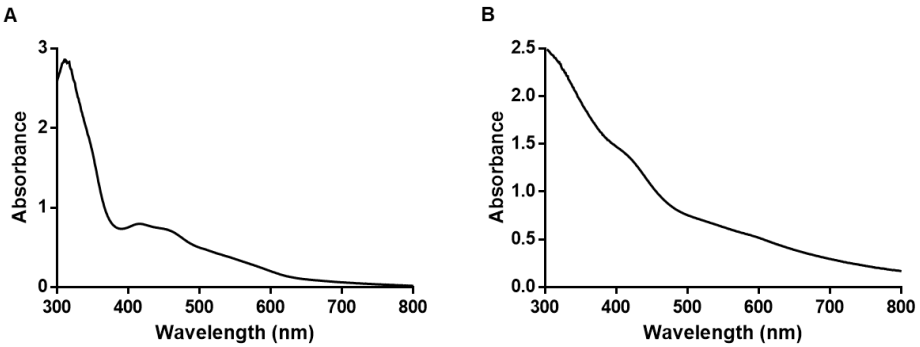

Figure S2.19 - UV-vis spectra of the $[2Fe\text{-}2S]^{3+}$ pentapeptides GCTCG (**A)** and GCTHG (**B**). Solutions contained 4 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8.

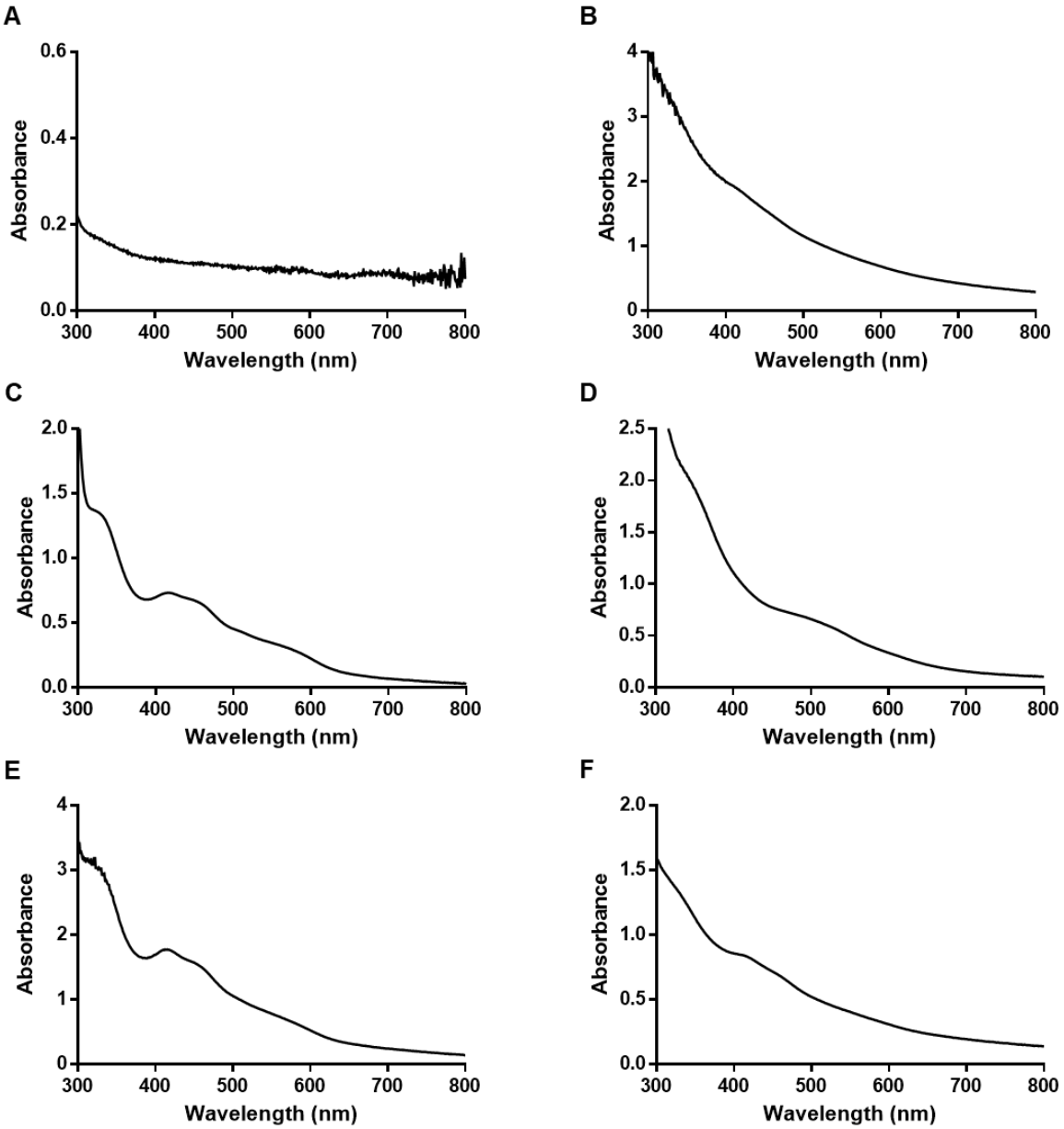

Figure S2.20 - UV-vis absorption spectra of $[2Fe\text{-}2S]^{2+}$ Linear and Stapled Rieske peptides. The peptides were Linear (CHCH): V**C**T**H**LGAIVSQWVADEEAAL**C**PA**H**G (**A**), Stapled (CHCH): V**C**T**H**LGDIVSL**C**PK**H**G (**B**) Linear (CCCH) V**C**T**C**LGAIVSQWVADEEAAL**C**PA**H**G (**C**), Stapled (CCCH): V**C**T**C**LGDIVSL**C**PK**H**G (**D**), Linear (CCCC): V**C**T**C**LGAIVSQWVADEEAAL**C**PA**C**G (**E**), Stapled (CCCC): V**C**T**C**LGDIVSL**C**PK**C**G (**F**). Conditions were 2 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8.

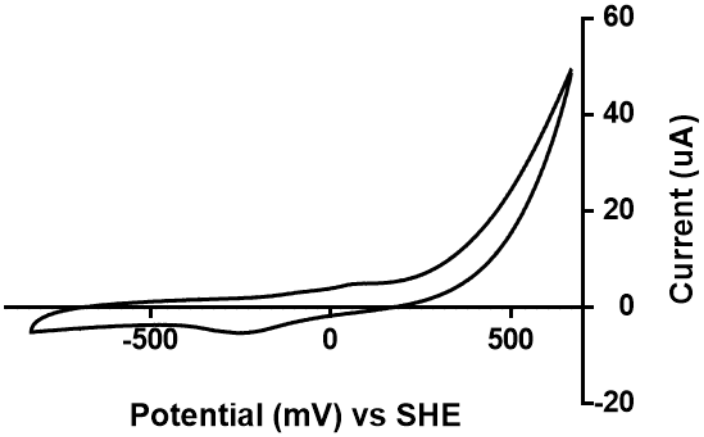

Figure S2.21 - Cyclic voltammetry of [1Fe-0S] NAc-Cys-OMe. Conditions were 32 mM peptide, 2 mM $FeCl_3$, pH 8.8, scan rate 100 mV/s.

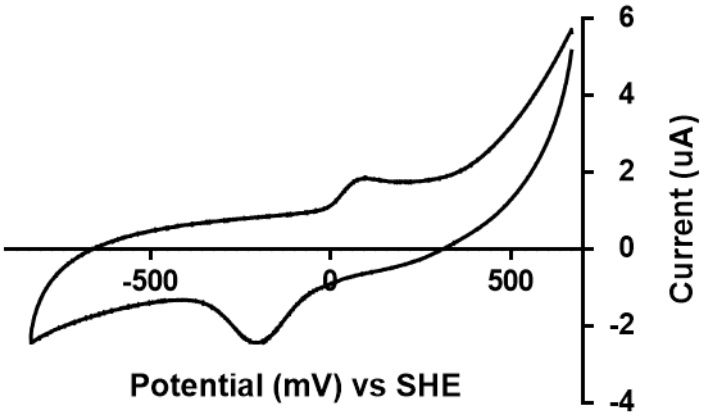

Figure S2.22 - Cyclic voltammetry of [1Fe-0S] GCG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8, scan rate 100 mV/s.

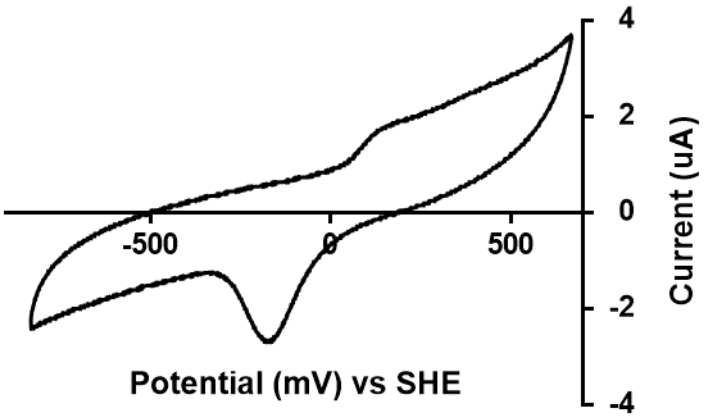

Figure S2.23 - Cyclic voltammetry of [1Fe-0S] GCTVCG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8, scan rate 100 mV/s.

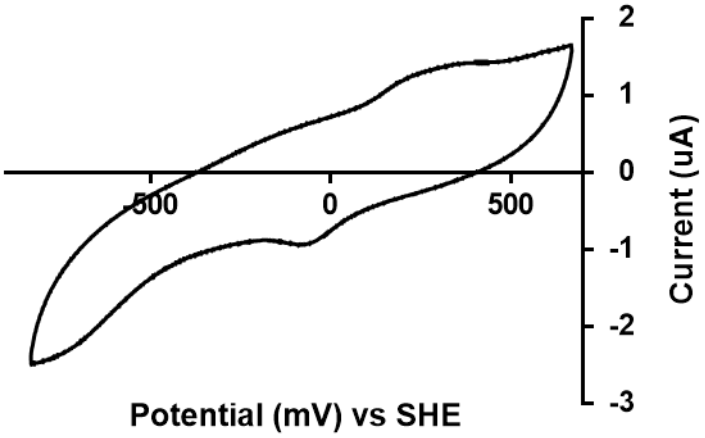

Figure S2.24 - Cyclic voltammetry of [1Fe-0S] GCTVHG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8, scan rate 100 mV/s.

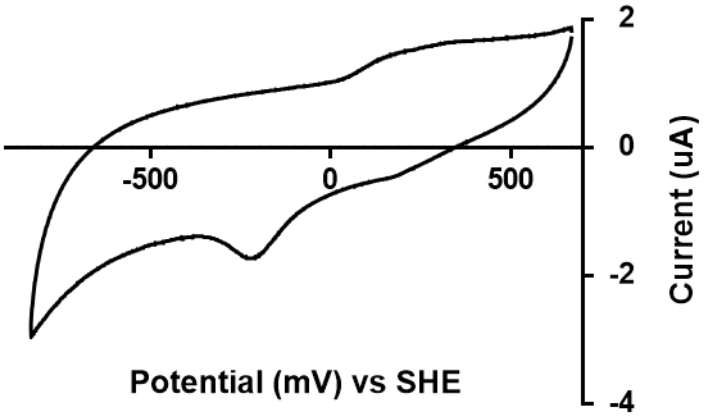

Figure S2.25 - Cyclic voltammetry of [1Fe-0S] GCTCG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8, scan rate 100 mV/s.

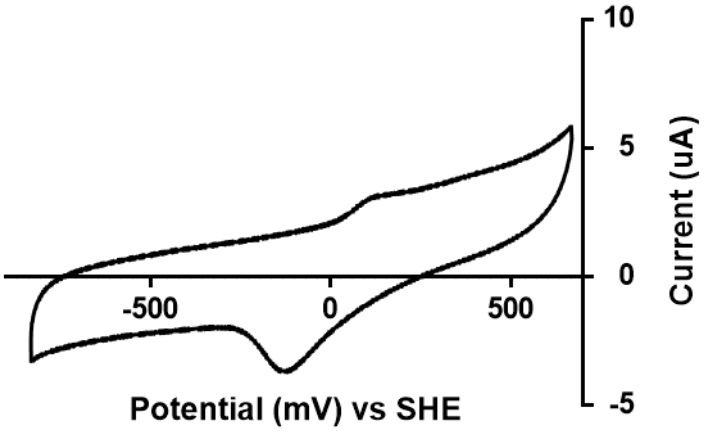

Figure S2.26 - Cyclic voltammetry of [1Fe-0S] GCTHG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8, scan rate 100 mV/s.

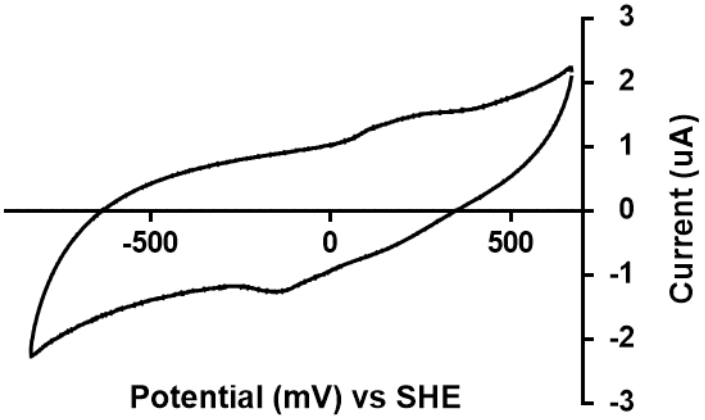

Figure S2.27 - Cyclic voltammetry of [1Fe-0S] GCPLCG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8, scan rate 100 mV/s.

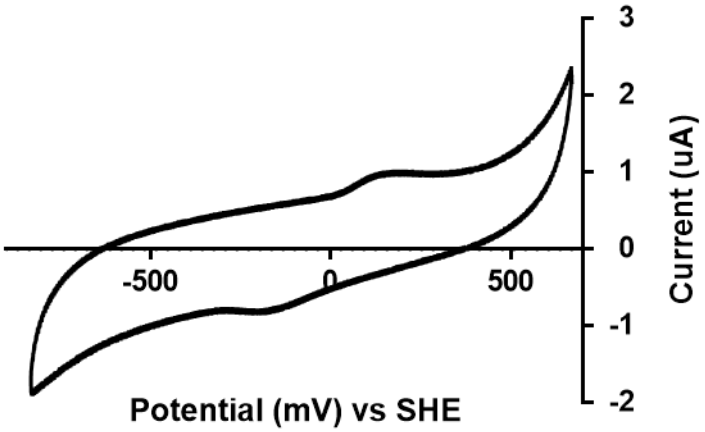

Figure S2.28 - Cyclic voltammetry of [1Fe-0S] GCPLHG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8, scan rate 100 mV/s.

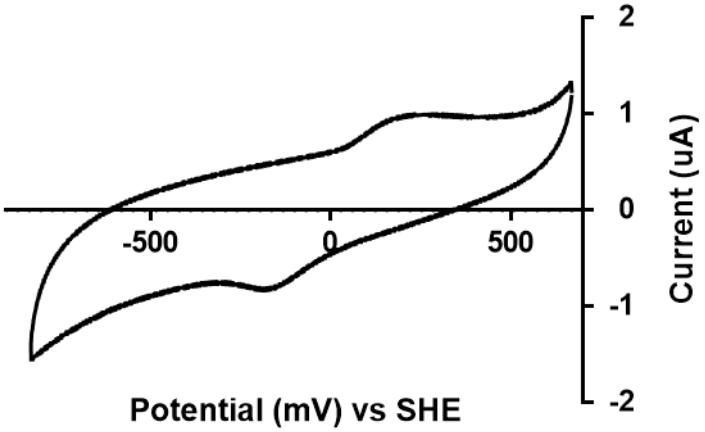

Figure S2.29 - Cyclic voltammetry of [1Fe-0S] GCPAHG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8, scan rate 100 mV/s.

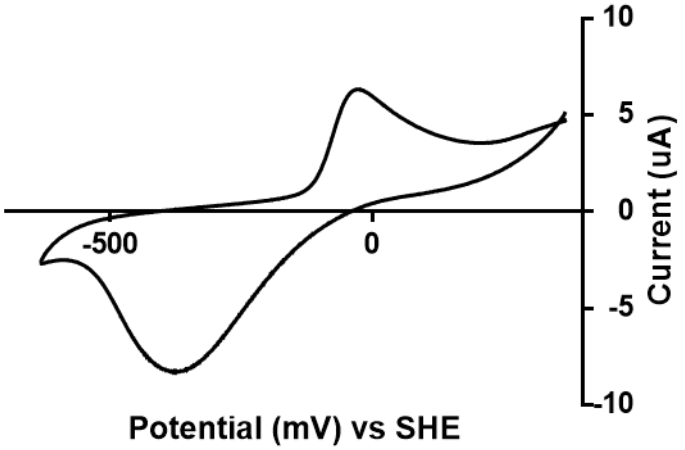

Figure S2.30 - Cyclic voltammetry of [2Fe-2S] NAc-Cys-OMe. Conditions were 32 mM peptide, 2 mM $FeCl_3$, 2 mM Na_2S, pH 8.8, scan rate 100 mV/s.

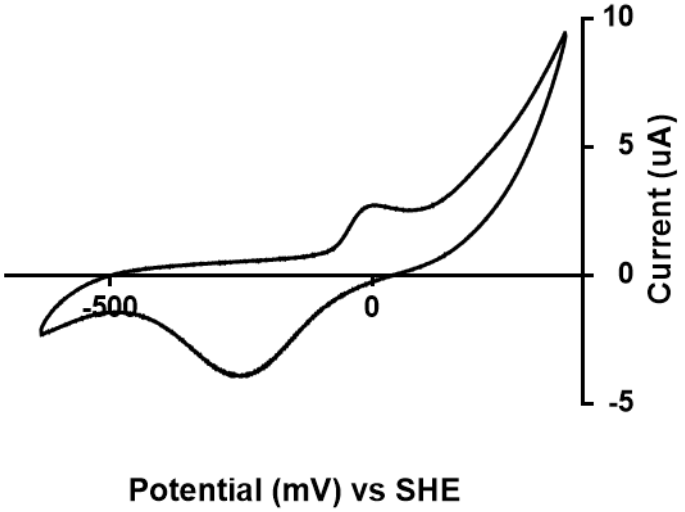

Figure S2.31- Cyclic voltammetry of [2Fe-2S] GCG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, 2 mM Na_2S, pH 8.8, scan rate 100 mV/s.

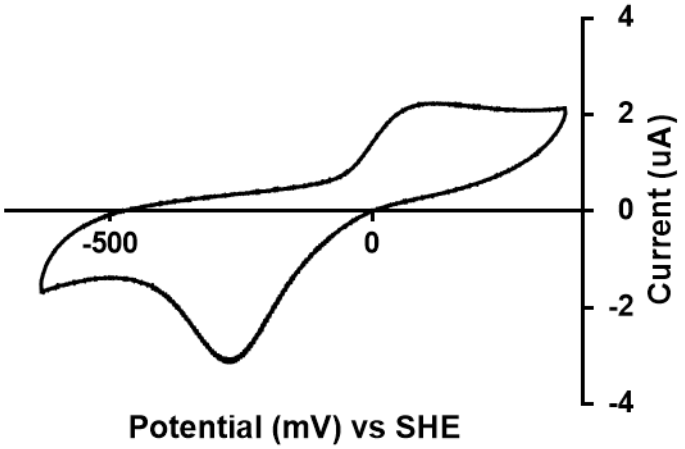

Figure S2.32 - Cyclic voltammetry of [2Fe-2S] GCTHG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, pH 8.8, 2 mM Na_2S, scan rate 100 mV/s.

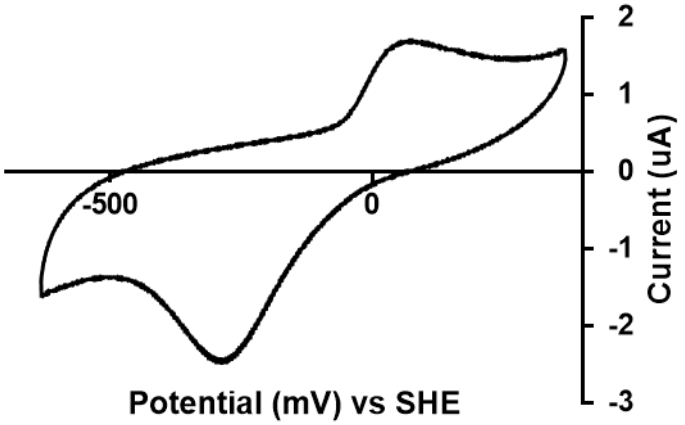

Figure S2.33 - Cyclic voltammetry of [2Fe-2S] GCPLHG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, 2 mM Na_2S, pH 8.8, scan rate 100 mV/s.

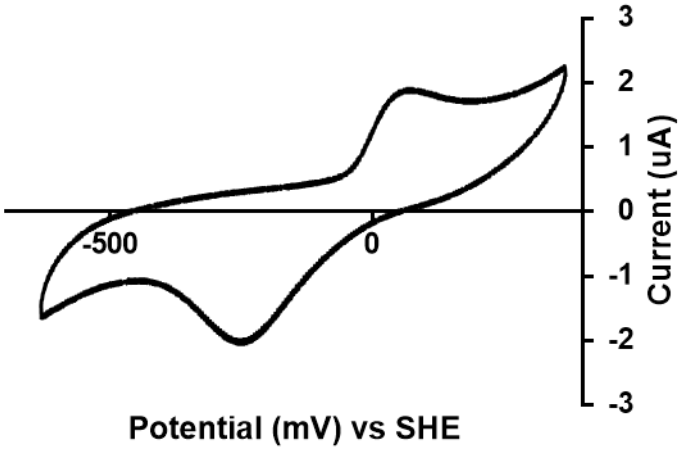

Figure S2.34 - Cyclic voltammetry of [2Fe-2S] GCPAHG. Conditions were 16 mM peptide, 2 mM $FeCl_3$, 2 mM Na_2S, pH 8.8, scan rate 100 mV/s.

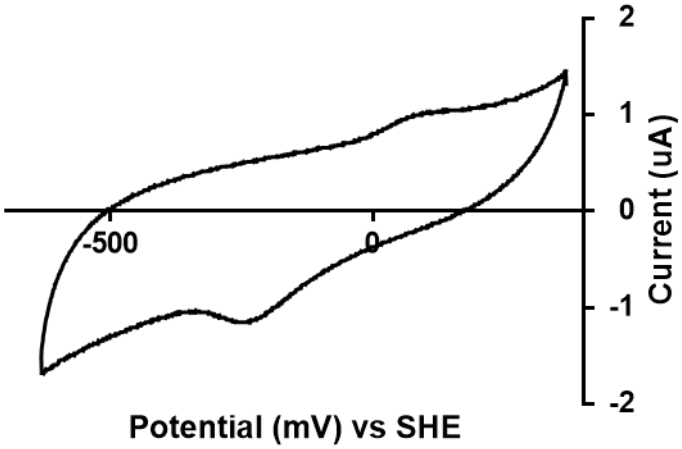

Figure S2.35 - Cyclic voltammetry of [2Fe-2S] stapled Rieske peptides (CHCH). Conditions were 8 mM peptide, 2 mM $FeCl_3$, 2 mM Na_2S, pH 8.8, scan rate 100 mV/s.

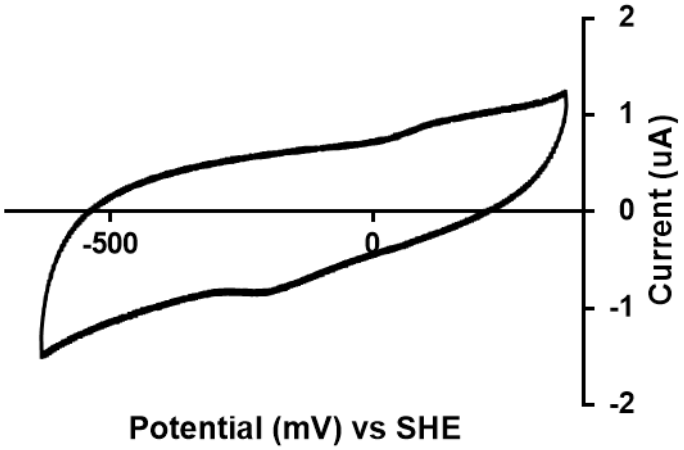

Figure S2.36 - Cyclic voltammetry of [2Fe-2S] stapled Rieske peptides (CCCH). Conditions were 8 mM peptide, 2 mM $FeCl_3$, 2 mM Na_2S, pH 8.8, scan rate 100 mV/s.

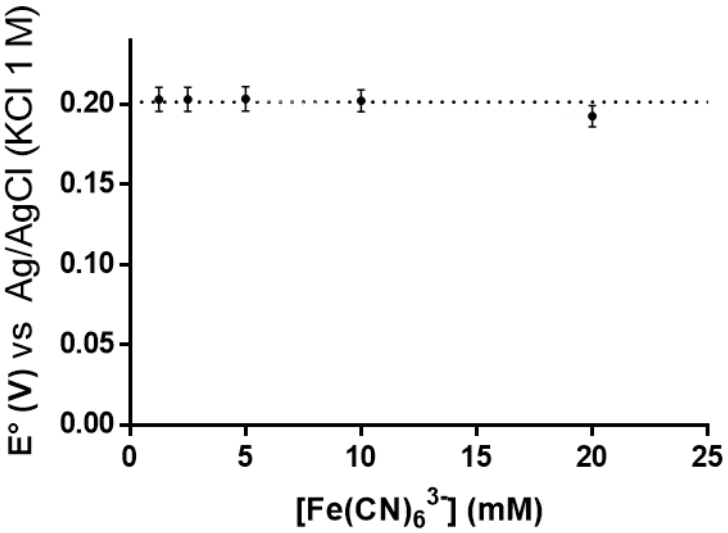

Figure S2.37 - Calibration curve of our electrochemical system. Reported in the graph are the measured potential of solution with different concentration of $K_3Fe(CN)_6$ vs Ag/AgCl (1 M KCl) (201.3 mV). Reported standard potential of $K_3Fe(CN)_6$ vs. SHE is -370.4 mV.[1]

Table S2.1 - K_d Co^{2+}, Fe^{2+}, and Zn^{2}+. Data are mean and SD of distinct samples, n = 3 replicates. The Hill coefficient (h) is for the binding of Co^{2+} and Fe^{2+} and not for Zn^{2+}, since the K_d of Zn^{2+} was determined by displacement of bound Co^{2+}.

	K_d Co^{2+} (mM)	h	K_d Fe^{2+} (mM)	h	K_d Zn^{2+} (µM)
N-Ac-Cys-OMe	0.449± 0.029	1.10± 0.082	2.15± 0.106	1.28± 0.066	57.8± 3.7
N-Ac-His-OMe	-	-	5.72± 0.576	1.31± 0.101	-
GCG	1.35± 0.04	2.05± 0.092	2.68± 0.211	1.23± 0.064	32.9± 2.5
GCPLHG	0.305± 0.005	1.66± 0.051	1.51± 0.118	1.32± 0.083	11.1± 0.4
GCPLCG	0.217± 0.009	2.46± 0.238	0.273± 0.009	2.03± 0.125	2.6± 0.8
GCPAHG	0.235± 0.006	1.49± 0.061	2.27± 0.080	1.25± 0.028	3.8± 1.1
GCTVHG	0.279± 0.027	1.46± 0.196	0.773± 0.044	1.36± 0.075	5.5± 1.5
GCTVCG	0.279± 0.010	2.06± 0.128	0.234± 0.011	1.86± 0.145	3.5± 0.6
GCTHG	0.446± 0.014	1.22± 0.033	1.74± 0.072	1.16± 0.033	14.3± 1.8
GCTCG	0.345± 0.011	2.03± 0.119	0.288± 0.008	1.72± 0.062	10.5± 0.7
Linear Rieske (CHCH)	0.059± 0,003	1.74± 0.167	1.59± 0.067	2.68± 0.244	0.012± 0.004
Linear Rieske (CCCH)	0.106± 0.003	1.99± 0.132	0.196± 0.006	1.51± 0.049	0.155± 0.095
Linear Rieske (CCCC)	0.026± 0.004	1.86± 0.616	0.064± 0.003	1.85± 0.17	$1.32x10^{-4}$± $6.75x10^{-5}$
Stapled Rieske (CHCH)	0.046± 0.001	1.81± 0.098	0.057± 0.017	0.839± 0.205	$1.37x10^{-3}$± $3.45x10^{-4}$
Stapled Rieske (CCCH)	0.172± 0,011	1.44± 0.110	0.175± 0.004	1.86± 0.078	3.5± 0.9
Stapled Rieske (CCCC)	0.201± 0.005	2.02± 0.094	0.113± 0.003	2.07± 0.124	3.1± 0.45

Table S2.2 - Coefficient of variation (CoV) of the K_d values. CoV were calculated as the ratio between the standard deviation σ and the mean µ.

N° ligands	Fe^{2+}	Co^{2+}	Zn^{2+}
1	0.55	0.71	0.39
2	0.82	0.25	0.63
4	1.64	0.69	1.49

Table S2.3 - Redox potentials of [1Fe-0S] and [2Fe-2S] clusters coordinated to peptides. Dashes indicate complexes that did not give reduction or oxidation peaks. Data represent mean and SD of distinct samples, n = 3 replicates.

Peptide Sequence	**Ligand:Fe^{3+}=16:1** E° vs SHE (mV)	**Ligand:Fe^{3+}:S^{2-}=16:1:1** E° vs SHE (V)
N-Acetyl-L-cysteine methyl ester	-104±21	-210±19
GCG	-57±12	-117±11
GCPLHG	-6.1±11	-83±26
GCPLCG	-10±7.8	-
GCPAHG	-44±11	-79±18
GCTVHG	3.7±32	-75±42
GCTVCG	-11±7.7	-231±6.5
GCTHG	16±13	-68±31
GCTCG	-24±20	-
Stapled Rieske (CHCH)	-	-63±20
Stapled Rieske (CCCH)	-	-49±7.5
Stapled Rieske (CCCC)	-	-

Table S2.4 - pH dependence of the redox potential of [2Fe-2S] cluster coordinated to GCTVCG and GCTVHG. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

Peptide Sequence	pH	E° (mV) vs. SHE
GCTVCG	7.8	-
GCTVCG	8.3	-187±45
GCTVCG	8.8	-231±6
GCTVCG	9.3	-230±8
GCTVCG	9.8	-235±13
GCTVHG	7.8	25±5
GCTVHG	8.3	-43±35
GCTVHG	8.8	-75±42
GCTVHG	9.3	-114±52
GCTVHG	9.8	-73±11

References

1. Rock, P. A. The standard oxidation potential of the ferrocyanide-ferricyanide electrode at 25° and the entropy of ferrocyanide ion. *J. Phys. Chem.* **70**, 576–580 (1966).

ChemBioChem

Research Article
doi.org/10.1002/cbic.202200202

Chemistry Europe
European Chemical Societies Publishing

www.chembiochem.org

Histidine Ligated Iron-Sulfur Peptides

Luca Valer,[a, b] Daniele Rossetto,[a, b] Taylor Parkkila,[b] Lorenzo Sebastianelli,[b] Graziano Guella,[c] Amber L. Hendricks,[d] James A. Cowan,[d] Lingzi Sang,[b] and Sheref S. Mansy*[a, b]

Iron-sulfur clusters are thought to be ancient cofactors that could have played a role in early protometabolic systems. Thus far, redox active, prebiotically plausible iron-sulfur clusters have always contained cysteine ligands to the cluster. However, extant iron-sulfur proteins can be found to exploit other modes of binding, including ligation by histidine residues, as seen with [2Fe-2S] Rieske and MitoNEET proteins. Here, we investigated the ability of cysteine- and histidine-containing peptides to coordinate a mononuclear Fe^{2+} center and a [2Fe-2S] cluster and compare their properties with purified iron-sulfur proteins. The iron-sulfur peptides were characterized by UV-vis, circular dichroism, and paramagnetic NMR spectroscopies and cyclic voltammetry. Small ($\leq$6 amino acids) peptides can coordinate [2Fe-2S] clusters through a combination of cysteine and histidine residues with similar reduction potentials as their corresponding proteins. Such complexes may have been important for early cell-like systems.

Introduction

Although iron-sulfur clusters are primarily coordinated by four cysteinyl ligands,[1] there are several cases where the cluster is not completely ligated by cysteines. For example, mononuclear, [2Fe-2S], and [4Fe-4S] clusters can be ligated by the sidechain of an Asp, Glu, or His.[2] Of these, ligation by His is rarer. Two examples of His ligated iron-sulfur clusters are found in Rieske and MitoNEET, which coordinate a [2Fe-2S] cluster with two Cys and two His, and with three Cys and one His, respectively. One notable effect of ligation by His is a shift in reduction potential of the cluster. [2Fe-2S] clusters solely coordinated by Cys, as found in ferredoxins, display reduction potentials from −150 to −400 mV (vs SHE),[3] while His coordination shifts the reduction potential to more positive values. The [2Fe-2S] clusters of MitoNEET have potentials from −100 mV to +50 mV.[4] The reduction potentials of [2Fe-2S] Rieske clusters vary from −150 to +400 mV.[5] The protonation state of the coordinated His impacts the reduction potential of the iron-sulfur cluster, and so alterations of the pK_a of the His can be used to fine-tune potentials.[6] Such a strategy was used to adjust the reduction potential of MitoNEET across a 700 mV range.[7]

The synthesis of non-proteinaceous iron-sulfur clusters has a long history. Holm and colleagues expertly deciphered the rules governing the assembly of all the major types of iron-sulfur clusters found in biology.[8] Although these reactions were typically in organic solvent, their characteristics were similar to their protein bound counterparts in aqueous solution. More recently, iron-sulfur clusters coordinated by small peptides in aqueous solution have been investigated.[9] The most well characterized of these are the ferredoxin maquettes from the Dutton group, where 15–20 amino acid peptides, based on natural protein sequences, were synthesized and found to coordinate a $[4Fe\text{-}4S]^{2+/+}$ cluster.[10] More recent efforts have focused on the synthesis of model prebiotic iron-sulfur peptides.[11] However, none of these previous efforts have demonstrated the ligation of an iron-sulfur cluster by imidazole or an imidazole sidechain with a non-protein scaffold in water. To determine if such clusters could be built in aqueous solution, we synthesized peptides between 5 and 25 amino acids in length containing Cys and His residues and compared their ability to coordinate iron-sulfur clusters with recombinant iron-sulfur proteins. Human ferredoxin and the soluble respiratory-Type Rieske protein from *Thermus thermophilus* were chosen as examples of [2Fe-2S] clusters.[12] *Clostridium pasteurianum* rubredoxin was the representative of a mononuclear center. Mutants of each were generated for comparison to the synthesized peptides. Mononuclear Fe^{2+}-peptides were spectroscopically different from rubredoxin but possessed similar reduction potentials. While the data were inconclusive for His ligation of mononuclear centers, UV-Visible absorption and NMR spectra in addition to reduction potentials were consistent with the ligation of [2Fe-2S] clusters by His-containing peptides.

[a] *L. Valer, D. Rossetto, S. S. Mansy*
D-CIBIO, University of Trento
via Sommarive 9, 38123 Trento 28123 (Italy)
[b] *L. Valer, D. Rossetto, T. Parkkila, L. Sebastianelli, L. Sang, S. S. Mansy*
Department of Chemistry, University of Alberta
11227 Saskatchewan Drive
Edmonton T6G 2G2, Alberta (Canada)
E-mail: sheref.mansy@ualberta.ca
[c] *G. Guella*
Department of Physics, University of Trento
Via Sommarive 14, Trento 38123 (Italy)
[d] *A. L. Hendricks, J. A. Cowan*
Department of Chemistry and Biochemistry
The Ohio State University
100 West 18th Ave, Columbus, OH 43210 (USA)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202200202

Results and Discussion

Reference Iron-sulfur proteins

Recombinantly expressed and purified Rieske proteins were prepared to generate reference UV-Vis absorption spectra. Their spectra were consistent with past reports,[12–14] with peaks at 450 nm and 580 nm (Figure 1C). Mutant constructs (Table 1) in which one of the His or both of the His were substituted with Cys possessed UV-vis spectra similar to MitoNEET and ferredoxin, respectively. H90 C Rieske (1 His, 3 Cys ligands) had a prominent peak at 470 nm (Figure 1D), and H90 C H107 C Rieske (4 Cys ligands) displayed absorption maxima at 420 nm and 450 nm (Figure 1E). Since Rieske proteins contain a disulfide bond between C93 and C109, it was possible that the recombinantly expressed constructs in which ligating His were substituted led to the recruitment of C93 and/or C109 for the stabilization of the cluster. To confirm that the [2Fe-2S] cluster was coordinated by the original positions of the ligands, the Cys involved in the formation of the disulfide bond were substituted with Ala. C93A H107C C109 A and H90C C93A H107C C109A Rieske showed absorption and visible CD spectra similar to constructs that contained C93 and C109 (Figures S1–S2). Continued coordination of the [2Fe-2S] cluster in the absence of a disulfide bond was consistent with past reports.[15] Recombinantly expressed rubredoxin had absorption maxima at 380 nm, 490 nm, and a shoulder at 580 nm (Figure 1A).

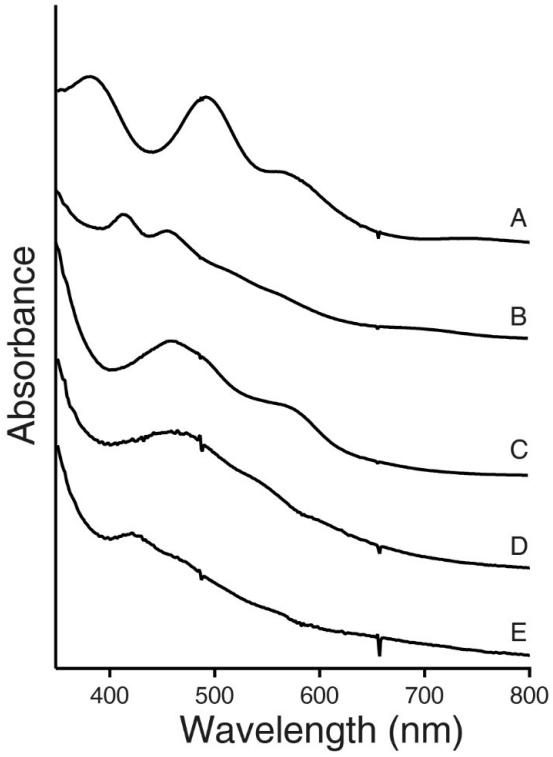

Figure 1. UV-vis spectra of oxidized iron-sulfur proteins. A. Wild Type (WT) rubredoxin (4 Cys ligands). B. WT human ferredoxin (4 Cys ligands). C. WT Rieske (2 His, 2 Cys ligands). D. H90C Rieske (1 His, 3 Cys ligands). E. H90C H107C Rieske (4 Cys ligands).

Table 1. Iron-sulfur proteins and peptides used in this study.

Name[a]	Protein/peptide sequence[b]
WT Rieske ROI	VCTHLGCIVSQWVADEEAALCPCHG
H90C Rieske ROI	VCTCLGCIVSQWVADEEAALCPCHG
H90C H107C ROI	VCTCLGCIVSQWVADEEAALCPCCG
C93A C109A ROI	VCTHLGAIVSQWVADEEAALCPAHG
H90C C93A C109A ROI	VCTCLGAIVSQWVADEEAALCPAHG
H90C C93A H107C C109A ROI	VCTCLGAIVSQWVADEEAALCPACG
CXH peptide	GCTHG
CXC peptide	GCTCG
CXXH peptide	GCPAHG
CHCH linear	VCTHLGAIVSQWVADEEAALCPAHG
CCCH linear	VCTCLGAIVSQWVADEEAALCPAHG
CCCC linear	VCTCLGAIVSQWVADEEAALCPACG
CHCH stapled	VCTHLGDIVSLCPKHG
CCCH stapled	VCTCLGDIVSLCPKHG
CCCC stapled	VCTCLGDIVSLCPKCG
WT Rubredoxin ROI	TCTVCG...VCPLCG
C42H Rubredoxin ROI	TCTVCG...VCPLHG
C9H C42H ROI	TCTVHG...VCPLHG
CXXC peptide	GCTVCG
CXXH peptide	GCTVHG
CXXC peptide	GCPLCG
CXXH peptide	GCPLHG
WT Ferredoxin ROI	ACEGTLACSTCH...GCQ

[a] ROI indicates the region of interest of the protein. [b] Positions in bold indicate coordinating residues. Underlined residues indicate the position of the original disulfide bond, which was substituted with a lactam bridge for the stapled peptides. The "…" of rubredoxin indicates 27 amino acids and 34 amino acids for ferredoxin.

However, the spectra of both C42H rubredoxin and C9H C42H rubredoxin were greatly diminished in intensity (Figure S3).

To investigate the ability of small peptides containing both Cys and His to coordinate metal ions similarly to proteins, fourteen different peptides and N-acetyl Cys methyl ester (N–Ac-Cys-OMe) and N-acetyl His methyl ester (N–Ac-His-OMe) were incubated with Fe^{2+}, Co^{2+}, and Zn^{2+} and spectroscopically evaluated. Cys-containing peptides, but not peptides that contained both Cys and His, were previously demonstrated to coordinate transition metals.[16,17] Here, we focused on the binding of Fe^{2+} because iron ions were easily accessible before

the great oxidation event,[18] are frequently found in model prebiotic pathways,[19–22] and are exploited throughout extant metabolism. Zn^{2+} was investigated because this metal ion can bind to the same motifs as iron-sulfur clusters,[2] and Zn^{2+} has been hypothesized to be important in early evolution.[23] Co^{2+} is a useful probe that facilitates studies with the spectroscopically silent Zn^{2+} through metal substitution.[24] The majority of the exploited peptide sequences were derived directly from sequences found in iron-sulfur proteins (Table 1). For example, the four ligands to the [2Fe-2S] cluster of Rieske are within a 25 amino acid region of the protein that contains CXH and CXXH motifs. Two CXC/H-containing pentapeptides (GCTHG and GCTCG) and one CXXH-containing hexapeptide (GCPAHG) were tested. Three 25 amino acid peptides that encompassed the entire [2Fe-2S] binding site of the Rieske protein from *T. thermophilus* were also synthesized with the Cys disulfide positions substituted with Ala. Longer peptides with four Cys were previously found to stabilize the formation of [2Fe-2S] clusters.[25] These 25 amino acid peptides either contained two (CHCH), three (CCCH), or four (CCCC) Cys residues (Scheme S1). Finally, three smaller (16 residues), stapled versions of the peptide with four potential ligands to the cluster were synthesized (Scheme S2). Staples were used to attempt to spatially organize the ligands for the binding of cluster. Stapled versions of the 25-residue peptide failed, likely due to the distance and flexibility within the peptide between the side-chains that would form the staple.[26] The smaller 16-residue version was designed by removing the residues involved in the formation of an intervening anti-parallel β-sheet[27] in between the CXH and CXXH motifs. The staple was a lactam bridge formed by using an orthogonal protecting group strategy during synthesis, which led to the covalent joining of the side chains of Asp and Lys. Rubredoxin contains two CXXC motifs separated by 27 amino acids. Four hexapeptides based on these motifs were synthesized, including GCTVCG, GCTVHG, GCPLCG, GCPLHG (Table 1). A GCG peptide was prepared and served, along with N–Ac-Cys-OMe and N–Ac-His-OMe, as a representative molecule containing a single potential ligand to the metal center.

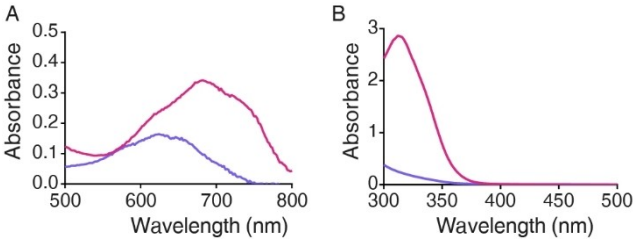

Figure 2. Comparison between all Cys and mixed Cys - His coordination for Co^{2+} and Fe^{2+}. UV-vis spectra of Co^{2+} (A) and Fe^{2+} (B) coordinated to GCTVCG (red) and GCTVHG (blue). Conditions were 1.25 mM peptide, 1 mM metal ion, 20 mM Gly-Gly, pH 8.8.

Absorption spectra of mononuclear Co^{2+} and Fe^{2+} centers

Metal binding was first evaluated with Co^{2+}, as the UV-visible absorption spectra of Co^{2+} complexes are easily interpretable in terms of the number of Cys ligands.[24] All of the small peptides (≤ six residues) with two Cys and no His showed an absorbance peak at 750 nm, consistent with coordination by four thiolates (Figure 2A).[24] To achieve such coordination, two peptides would have had to bind a single Co^{2+}. When one of the Cys was substituted with a His, the absorbance peak shifted to 660 nm, as would be expected for a complex with two Cys and two His ligands (Figure 2B).[24] The longer linear and stapled Rieske peptides with four potential ligands to the metal ion gave UV-vis spectra congruent with the number of Cys residues. CHCH, CCCH, and CCCC Rieske peptides had absorbance maxima at 660 nm, 710 nm, and 750 nm, respectively, when bound to Co^{2+} (Figure S4).

The UV-vis spectra of peptidyl complexes with Fe^{2+} were less clearly interpretable than spectra with Co^{2+} and were unlike that of rubredoxin. Absorption spectra with Fe^{2+} either showed a clear peak at 315 nm when coordination by four Cys was possible or a shoulder in the 300–310 nm when His was present (Figures 2B-S5-S6-S7). The only exceptions were N–Ac-Cys-OMe, which gave a featureless UV-vis spectrum above 250 nm, and N–Ac-His-OMe, which showed a broad shoulder below 300 nm (Figure S8). To confirm the coordination of Fe^{2+}, paramagnetic ^{1}H NMR spectra were acquired. The paramagnetism of the Fe^{2+} center gives rise to hyperfine shifted resonances of Cys Hα in the 10–20 ppm region and Cys Hβ within the 100–300 ppm range.[28] The paramagnetic ^{1}H NMR spectra of the Cys-containing peptides incubated with Fe^{2+} showed the presence of such paramagnetically shifted resonances of Cys Hα and Cys Hβ (Figure S9A-9B). Paramagnetically shifted resonances were not detected for peptides containing His (Figure S10A). The addition of Fe^{2+} to the 25 residue Rieske-like peptides and the shorter Stapled peptides led to broad, featureless absorption in the UV-Vis region, with absorption increasing with increasing number of Cys (Figure S11). The stapled peptide absorbed more strongly than the 25 residue peptides upon the addition of Fe^{2+}.

Peptide affinities for Co^{2+}, Zn^{2+}, and Fe^{2+}

We next evaluated the binding affinity for Co^{2+}, Zn^{2+}, and Fe^{2+}. Since Zn^{2+} complexes do not absorb in the UV-Vis region, the affinity for Zn^{2+} was determined by substitution of bound Co^{2+}. Binding affinities for N–Ac-Cys-OMe, N–Ac-His-OMe, and GCG i.e. molecules with only one potential ligand, were always less than peptides that contained more than one potential ligand for all three metal ions. Molecules with two ligands possessed, on average, 2.7-fold greater affinity for the tested metal ions, and peptides with four potential ligands bound the metal ions 2.8-fold more tightly than peptides with two ligands (Figures S12–S13–S14). Affinities followed the Irving-Williams series, with peptides binding Zn^{2+} more strongly than Co^{2+} followed by Fe^{2+}. There was a stronger sequence dependence, confirmed by calculations of coefficients of variation, of affinities for Fe^{2+}

followed by Zn^{2+} and then Co^{2+} (Figure 3 and Tables S1–S2). For Fe^{2+}, the preference for Cys ligands over His ligands was most evident for the two ligand containing peptides (Figure 3, blue bars). Of these, peptides that contained two Cys residues bound Fe^{2+}, on average, 5.9-fold greater than peptides with one Cys and one His. The Linear Rieske (CCCC) peptide bound all the tested metal ions more strongly than the other Linear Rieske peptides (Figures S15–S16–S17). Conversely, of the Stapled Rieske peptides, the best binder was (CHCH), i.e. the peptide that contained two Cys and two His ligands. The lactam bridge may have helped pre-organize the Cys and His for binding and thus decreased the entropic cost associated with coordination.

Absorption spectra of [2Fe-2S] peptides

To determine whether the synthesized peptides could coordinate a [2Fe-2S] cluster, each peptide was anaerobically incubated with Fe^{3+} and HS^-. The ratio of ligating residues (L) to Fe^{3+} and HS^- (i.e. L:Fe^{3+}:HS^-) was 16:1:1. The UV-vis spectra of Cys containing small ($\leq$ 6 amino acids) peptides that lacked His residues showed bands at 420 nm and 450 nm indicative of a ferredoxin-like [2Fe-2S] cluster (Figure 4A and S18A-S19 A).[29] Near-UV-visible circular dichroism (CD) spectra showed a band at 470 nm that was similar to the 450 nm peak seen with ferredoxin, but with lower molar ellipticity (Figure S20A). Paramagnetic ^{1}H NMR spectra confirmed the presence of the [2Fe-2S]$^{2+}$ cluster with a broad resonance at 30 ppm (Figure S9D).[28] The NMR spectrum also revealed the continued presence of the mononuclear center that was not observable by UV-vis absorption (Figure S9C). Such a result was not surprising, since peptides exist in an equilibrium between the coordination of a mononuclear center and a [2Fe-2S] cluster when in an environment containing iron and sulfide ions.[28] Conversely, UV-Vis absorption spectra of small ($\leq$ 6 amino acids) peptides that possessed both Cys and His residues did not resemble the spectra of ferredoxin or the Rieske protein (Figure 4B-S18BC–S19B). Although there was absorption in the visible region, no clear peaks were discernable, and their CD spectra were featureless (Figure S14B). Paramagnetic ^{1}H NMR spectra showed sharp peaks near 10 ppm and 30 ppm in addition to a shoulder at 13 ppm (Figure S7B). The peak at 30 ppm was consistent with previously reported spectra of Rieske proteins;[5,30] however, the data were not sufficient to distinguish between four peptides, each providing a single Cys to a single cluster or two peptides that ligated the [2Fe-2S] with two Cys and two His in total. In the presence of Fe^{3+} and HS^-, UV-Vis spectra were indicative of the presence of a ferredoxin-like [2Fe-2S]$^{2+}$ cluster for Linear peptides CCCH and CCCC and not for the more

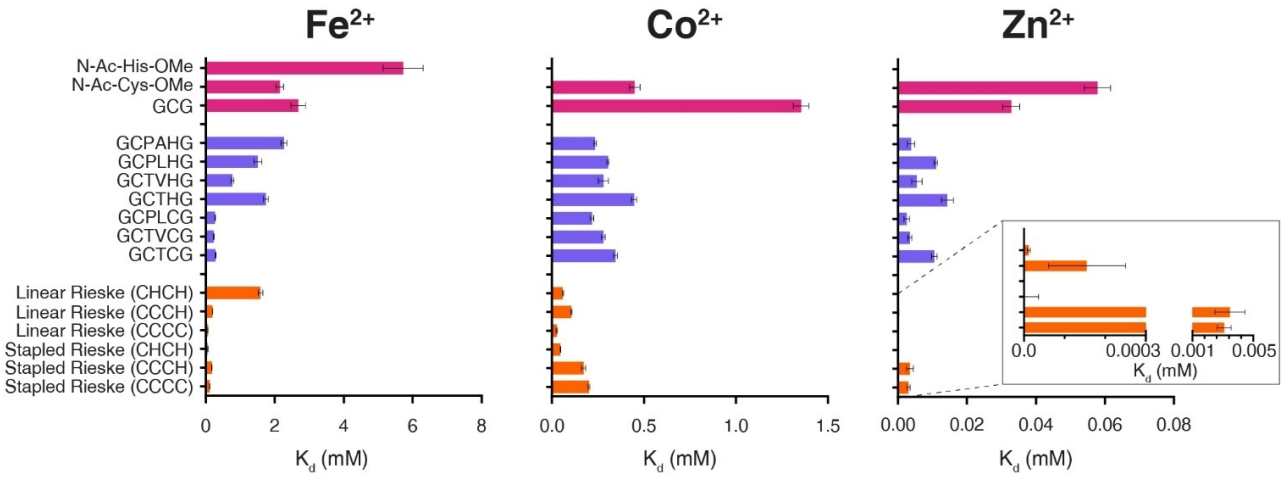

Figure 3. K_d of peptides for Fe^{2+}, Co^{2+}, and Zn^{2+}. Data obtained represent mean and SD of distinct samples, n = 3. Peptides with a single ligand (Cys or His) are in red, peptides with two ligands are in blue, and peptides with four ligands are in orange. The K_d of N–Ac-His-OMe is not reported for Co^{2+} and Zn^{2+}.

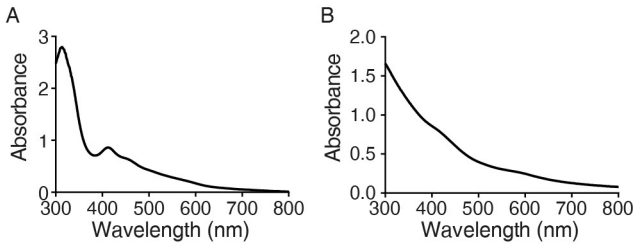

Figure 4. UV-vis spectra [2Fe-2S]$^{2+}$ cluster coordinated by small peptides. A. [2Fe-2S]$^{2+}$ coordinated by GCTVCG. B. [2Fe-2S]$^{2+}$ coordinated by GCTVHG. Solution conditions were 4 mM peptide, 0.5 mM $FeCl_3$, 0.5 mM Na_2S, pH 8.8.

Rieske-like sequence found in the Linear CHCH peptide (Figure S21A-S21C–S21E). UV-Vis spectra of the stapled peptide lacked well defined peaks, but as seen with the Linear peptides, the bands became more prominent as the number of Cys increased (Figure S21B-S21D–S21F). The data suggested that the Linear peptides better stabilized the formation of a [2Fe-2S]$^{3+}$ cluster than the Stapled peptides (Figures S22–S23–S24–S25–S26–S27).

Cyclic voltammetry confirms His-coordination to [2Fe-2S] clusters

Reduction potentials of iron-sulfur clusters are strongly influenced by their ligands. The substitution of a negatively charged Cys ligand with a neutral His typically increases the reduction potential of the cluster. Therefore, to confirm whether the His residues were ligating the iron ions, cyclic voltammetry was carried-out on the prepared complexes. All of the mononuclear complexes had reduction potentials in the range expected for rubredoxin-like proteins,[31] and other reported mononuclear centers coordinated by small thiol-containing molecules (Figures S28–S29–S30–S31–S32–S33–S34–S35–S36, Table S3).[22] The most negative reduction potential measured was for N–Ac-Cys-OMe (-104 ± 21 mV; all reduction potentials are vs. SHE). The most positive reduction potential was that of the pentapeptide GCTHG ($+16 \pm 1$ mV) (Figure 5A). Complexes of Fe^{2+} with N–Ac-Cys-OMe and GCG had lower reduction potentials than peptides containing two potential ligands to the iron ion. Although the average reduction potentials were greater for His-containing peptides, the values were within error. Therefore, it was not possible to confirm ligation by His for the mononuclear complexes.

Conversely, cyclic voltammetry of [2Fe-2S] clusters coordinated to peptides showed two different types of behavior (Figures S37-S38-S39-S40-S41-S42-S43-S44, Table S4). The all Cys containing peptides had reduction potentials in the range of -210 mV to -231 mV, with the exception of GCG (-117 ± 12 mV). The data were similar to [2Fe-2S] ferredoxin with four ligating Cys. The peptides that contained both His and Cys gave more positive reduction potentials between -63 mV and -83 mV (Figure 5B), as would be expected for ligation by His, and similar to Rieske proteins.[4] We were not able to determine the reduction potentials of all the peptides, because several did not produce distinguishable reduction and oxidation peaks. These peptides included GCPLCG, GCTCG, and most of the Linear and Stapled Riske peptides. The only exceptions were the Stapled Rieske peptides with two Cys and two His (CHCH) and with three Cys and 1 His (CCCH), which had reduction potentials of -63 mV and -49 mV, respectively, consistent with coordination by His.

The reduction potential of Rieske proteins is pH dependent.[32] Lower pH leads to increased protonation of the ligating His and thus higher reduction potential. Therefore, if the investigated peptides bound the iron-sulfur cluster, in part, through His residues, then the reduction potential would be expected to be influenced by pH. Two [2Fe-2S] peptides, one lacking His (GCTVCG) and one containing His (GCTVHG) were subjected to cyclic voltammetry between pH 7.8 and 9.8. As expected, the reduction potential of the [2Fe-2S] cluster coordinated by GCTVHG decreased as the pH increased with a 92 mV/pH dependence (Figure 6C). Rieske proteins show a similar dependence (120 mV/pH).[33] Conversely, the reduction potential of the [2Fe-2S] cluster coordinated to GCTVCG was consistently near -230 mV between pH 8.3 and 9.8 (Table S4). Data are not reported at pH 7.8, because [2Fe-2S] GCTVCG was not stable at this pH. The data were consistent with the His residues coordinating the iron-sulfur cluster.

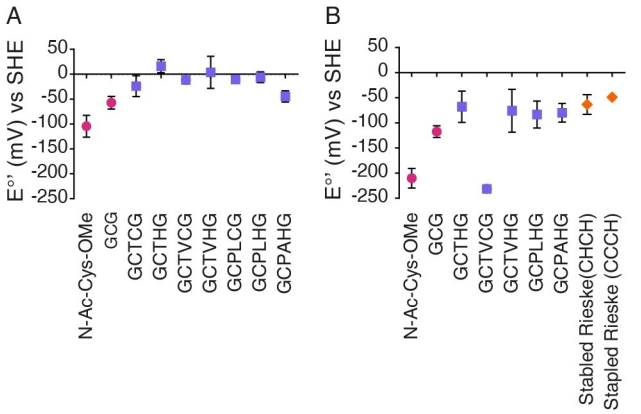

Figure 5. Measured formal reduction potentials of iron-sulfur peptides. A. Measured potentials of mononuclear centers stabilized by small peptides (vs. SHE). B. Measured potential of [2Fe-2S] peptides. Data obtained represent mean and SD of distinct samples, n = 3 replicates. Conditions were 16 mM ligating residues, 1 mM $FeCl_3$, 1 mM Na_2S, pH 8.8. Red circle, 1 ligand; blue square, 2 ligands; orange diamond, Stapled CHCH and CCCH peptides.

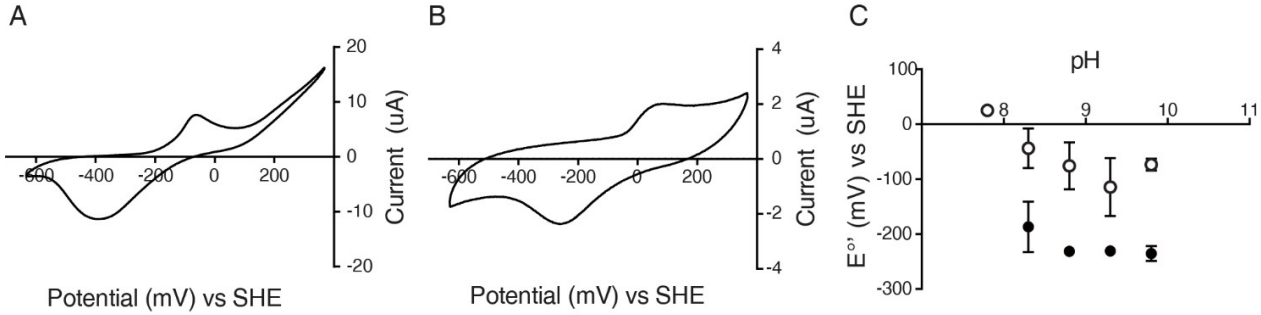

Figure 6. Cyclic voltammetry of [2Fe-2S] peptides. Cyclic voltammetry of [2Fe-2S] GCTVCG (A) and [2Fe-2S] GCTVHG (B). Conditions were 16 mM peptide, 2 mM $FeCl_3$, 2 mM Na_2S, pH 8.8, scan rate 100 mV/s. C. Formal reduction potentials of [2Fe-2S] GCTVCG (filled circles) and [2Fe-2S] GCTVHG (open circles) at different pH. Measurements were on 8 mM peptide, 1 mM $FeCl_3$, 1 mM Na_2S, pH 8.8 with a scan rate of 100 mV/s. At pH 7.8 the cluster coordinated by GCTVCG was not formed. Data obtained represent mean and SD of distinct samples, n = 3 replicates.

Conclusions

Due to their central role in metabolism and the ease with which these metallocofactors form, there has been much speculation as to how iron-sulfur clusters could have impacted the origins of life.[34] Thus far, peptidyl mimics of prebiotic iron-sulfur clusters have focused on thiolate coordination; however, extant electron transfer pathways additionally make use of Fe^{2+} coordinated to His and porphyrins with steps mediated by organic molecules and other metal ions.[35] Although His is often argued to be prebiotically implausible because of phylogenetic analyses, phylogeny tells us little about the breadth of chemistry that transpired on the prebiotic Earth.[36] Nevertheless, the prebiotic presence of His remains unclear. Oro and coworkers did show that His can be synthesized from erythrose and formamidine.[37,38] However, the reaction is inefficient and relies on high concentrations of formamidine, which rapidly hydrolyzes.[39] Similar ambiguities existed for Cys in the past. Cys was frequently referred to as prebiotically implausible, and an analysis of the 1958 H_2S-rich spark discharge experiment showed the presence of the break-down products of Cys rather than Cys itself.[40,41] However, more recent efforts have uncovered prebiotically plausible routes to Cys.[42] Our data do not speak to the prebiotic plausibility of peptides that contain these amino acid residues but rather to their characteristics.

To determine if simpler analogues of the electron transport chain could have predated the complex pathways observed today, data are needed regarding the formation of different metallocomplexes and their electron transfer characteristics. Contemporary proteins can tune the reduction potential of their metallocofactor through subtle interactions arising from the local environment created within a complex tertiary fold. Here, we show that in the absence of such evolved scaffolds, [2Fe-2S] clusters can form coordinated to peptides containing Cys and His ligands with reduction potentials distinct from iron-sulfur peptides that rely on complete cysteinyl coordination. Since ligation by His shifts the reduction potential by 150 mV, iron-sulfur peptide-dependent electron transport chains much simpler than found in biology today may have been possible.

Experimental Section

Protein expression and purification: Plasmids encoding rubredoxin and Rieske iron-sulfur proteins and their respective mutants were synthesized by Genscript and inserted into pMal-c4x (NEB), generating a Maltose Binding Protein (MBP) tag. The plasmid encoding ferredoxin was synthesized by Genscript and inserted in pET-21b. Human ferredoxin was not tagged. Recombinant expression of the genes in pMal-c4x was with *Escherichia coli* NEBExpress. Recombinant expression of ferredoxin was with *E. coli* BL21(DE3) pLysS. For both, transformed cells were grown overnight at 37 °C in 10 mL Luria-Bertani (LB) supplemented with ampicillin (50 mg/mL). 1 mL of this starter culture was used to inoculate 1 L LB with 50 mg/mL ampicillin in a 5 L flask and grown at 37 °C with shaking to an OD^{600} of 0.6. Expression was induced with 0.4 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated at room temperature for an additional 15 h. Cells were harvested by centrifugation at 4000 x g for 15 min. The cell pellet was resuspended in 40 mL of sonication buffer (20 mM Tris HCl, 200 mM NaCl, pH 7.6,) and lysed by sonication (20 cycles of 10 s pulses followed by 50 s incubation on ice). Cellular debris was removed by centrifugation at 15000 x g for 15 min. For the purification of the MBP-tagged proteins, the lysate was loaded on an amylose column (NEB) and washed with 20 column volumes of amylose column buffer (20 mM Tris HCl, 200 mM NaCl, 0.1 mM EDTA, pH 7.6). The protein was eluted with 10 mM maltose in column buffer. The colored fractions were collected and concentrated using Amicon Ultra centrifugal filter units (MWCO 10 kDa). Proteins were dialyzed against 20 mM Tris HCl, pH 7.6 overnight at 4 °C. Ferredoxin was purified in two-steps, as previously reported.[29] The first step of the purification exploited an ion exchange column with DEAE resin (Bio-Rad) and washed with 20 column volumes of DEAE column buffer (50 mM Tris-HCl, pH 7.4). The proteins were eluted with a 90 mL gradient of 0–500 mM NaCl in DEAE column buffer. The protein eluted at approximately 350–400 mM NaCl. The colored fractions were collected and concentrated with Amicon Ultra centrifugal filter units (MWCO 10 kDa). The second step of the purification relied on size-exclusion chromatography with a HiLoad 16/600 Superdex 75 pg column and a FPLC (Äkta pure, GE Healthcare). Subsequently, the protein-containing fractions were collected and concentrated with Amicon Ultra centrifugal filter units (MWCO 10 kDa).

Peptide synthesis: Peptides were synthesized via solid phase peptide synthesis, as previously described.[11] *N*,*N*-Dimethyl formamide (DMF) was used as the solvent and 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 the building blocks. The peptide chain was elongated by sequential Fmoc deprotection of the residue anchored to the resin and Fmoc-AA-OH coupling. 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. For the cyclized peptides, Fmoc protected Asp (Allyl) and Lys (Alloc) were used. After the synthesis of the full peptide, before the last Fmoc deprotection, Asp and the Lys were deprotected by treatment with 0.5 equiv. $Pd(PPh_3)_4$, 10 equiv. of $PhSiH_3$ for 2 h, in the dark under constant flux of N_2 and repeated twice. After deprotection, the resin was washed five times with dichloromethane (DCM), sodium diethyldithiocarbamate in DMF (0.1% w/w) and DMF. Cyclization was achieved with a 24 h treatment hydroxyl-benzotriazole (HOBt), *N*,*N*-diisopropylethyl amine (DIPEA), and (7-azabenzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyAOP) in DMF. At the end of the synthesis, 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) (97:1:1:1, v/v) for 2 h. The product was precipitated with cold diethyl ether/petroleum ether (30:70, v/v) followed by washing cycles with diethyl ether. The peptides were finally solubilized in 20% acetic acid (v/v), flash frozen in liquid nitrogen, and lyophilized overnight at −84 °C (Labconco FreeZone Freeze Dryer). The stapled peptides were further purified by HPLC (Agilent 1100 Series) with a reverse phase Zorbax SB–C12 column (Agilent Technologies) using H_2O and acetonitrile (both containing 0.1% TFA) as the mobile phase. The gradient was from 20% to 100% acetonitrile over 15 min. After purification, fractions containing peptide were lyophilized overnight at −84 °C.

Synthesis of the clusters: All steps were carried-out under anoxic conditions inside a Genesis 2P glovebox system (Vacuum Atmos-

pheres Company) with $O_2 < 1$ ppm. The peptides were dissolved in ultrapure degassed H_2O (Synergy UV water purification system, Merck, Darmstadt, Germany) and the pH was adjusted to pH 8.8 using NaOH (5 M), if not specified otherwise. For the formation of the mononuclear complex, $FeCl_3$ was added from a 0.33 M stock to 1:8 Fe^{3+} : (Cys + His) concentration. [2Fe-2S] clusters were synthesized by the addition of $FeCl_3$ (0.33 M stock) and Na_2S (0.33 M stock) to 1:1:16 Fe^{3+}: S^{2-} : (Cys + His) concentration.

UV-Vis absorption and Circular Dichroism spectroscopy: The UV-vis spectra of the samples were recorded inside a glovebox with a Genesys 150 UV-Vis spectrophotometer (ThermoFisher) with an integration time of 0.5 s and an interval of 1 nm. Circular Dichroism (CD) spectra were collected with a JASCO J-815 CD spectrometer, using a scan rate of 200 nm/min (spectral window 200–600 nm). 20 scans were collected. Spectra manager II (JASCO) software was used to analyze data.

Affinity measurements: The peptide solutions were prepared by dissolving the peptides in 20 mM Gly-Gly, pH 8.8 inside a glovebox to a total concentration of cysteine plus histidine of 2.5 mM. After sequential addition of the chosen metal (Fe^{2+}, Co^{2+}), UV-vis spectra were collected and fit with GraphPad Prism v. 6.00 (GraphPad Software, La Jolla California USA) to the following equation: $y = B_{Max}*x^h / (K_d{}^h + x^h)$, where B_{Max} was the absorbance reached at saturation and h was the Hill slope. For Zn^{2+}, a competition titration against Co^{2+} was used. The Co^{2+} complex was formed at the K_d prior to titration with Zn^{2+}, and the absorption fit to the revised Cheng-Prusoff equation.[43]

Paramagnetic NMR spectroscopy: Samples were prepared in a glovebox, as described above, and 700 μL placed in 5 mm NMR tubes (Sigma-Aldrich) with 10% D_2O. Spectra were recorded at room temperature using a V NMRS four-channel 600 MHz Varian spectrometer (5000 scans, 0.4 s recycle delay, 0.25 s acquisition time).

Cyclic voltammetry: Experiments were conducted under ambient conditions using an electrochemical workstation CHI660 A (CH Instruments). Samples were prepared inside a glovebox and then transferred into a custom-made electrochemical cell. All solutions were purged with argon for 10 min before analysis to ensure that no O_2 was trapped during the transfer from the glovebox to the cell. A standard three electrode cell composed of a glassy carbon working electrode (BASi, 3 mm diameter), an Ag/AgCl reference electrode (1 M KCl) (calibration curve Figure S28), and a platinum wire counter electrode was used for the electrochemical measurements. A scan rate from 50 mVs to 500 mVs was used. No supporting electrolyte was used to avoid interference with stability of the cluster; however, a small amount of NaCl was present from pH adjustment with NaOH and HCl. The working electrode was cleaned with a 0.05 μm aluminum oxide slurry on a polishing cloth and thoroughly rinsed with MilliQ H_2O (Synergy UV water purification system, Merck, Darmstadt, Germany) before each experiment. The reference electrode was sonicated in MilliQ H_2O for 5 min before each experiment.

Acknowledgements

We thank the Simons Foundation (290358FY18 and 290358FY19) and the Natural Sciences and Engineering Research Council of Canada (NSERC) (RGPIN-2020-04375) for funding and A. Dallapè and S. Scintilla for preliminary experiments.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are openly available in Zenodo at https://doi.org/10.5281/zenodo.6425646, reference number 1.

Keywords: bioinorganic chemistry · iron-sulfur clusters · metallopeptides · prebiotic chemistry · Rieske

[1] D. C. Johnson, D. R. Dean, A. D. Smith, M. K. Johnson, *Annu. Rev. Biochem.* **2005**, *74*, 247–281.
[2] L. Belmonte, S. S. Mansy, *J. Chem. Inf. Model.* **2017**, *57*, 3162–3171.
[3] C. Bonfio, *Dalton Trans.* **2021**, *50*, 801–807.
[4] D. W. Bak, S. J. Elliott, *Curr. Opin. Chem. Biol.* **2014**, *19*, 50–58.
[5] E. N. Brown, R. Friemann, A. Karlsson, J. V. Parales, M. M. J. Couture, L. D. Eltis, S. Ramaswamy, *J. Biol. Inorg. Chem.* **2008**, *13*, 1301–1313.
[6] A. R. Klingen, G. M. Ullmann, *Biochemistry* **2004**, *43*, 12383–12389.
[7] J. A. Zuris, D. A. Halim, A. R. Conlan, E. C. Abresch, R. Nechushtai, M. L. Paddock, P. A. Jennings, *J. Am. Chem. Soc.* **2010**, *132*, 13120–13122.
[8] P. Venkateswara Rao, R. H. Holm, *Chem. Rev.* **2004**, *104*, 527–559.
[9] W. Qi, J. Li, C. Y. Chain, G. A. Pasquevich, A. F. Pasquevich, J. A. Cowan, *J. Am. Chem. Soc.* **2012**, *134*, 10745–10748.
[10] B. R. Gibney, S. E. Mulholland, F. Rabanal, P. L. Dutton, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 15041–15046.
[11] C. Bonfio, L. Valer, S. Scintilla, S. Shah, D. J. Evans, L. Jin, J. W. Szostak, D. D. Sasselov, J. D. Sutherland, S. S. Mansy, *Nat. Chem.* **2017**, *9*, 1229–1234.
[12] J. A. Fee, K. L. Findling, T. Yoshida, R. Hille, G. E. Tarr, D. O. Hearshen, W. R. Dunham, E. P. Day, T. A. Kent, E. Münck, *J. Biol. Chem.* **1984**, *259*, 124–133.
[13] B. Xia, H. Cheng, V. Bandarian, G. H. Reed, J. L. Markley, *Biochemistry* **1996**, *35*, 9488–9495.
[14] J. Lin, T. Zhou, K. Ye, J. Wang, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 14640–14645.
[15] T. Merbitz-Zahradnik, K. Zwicker, J. H. Nett, T. A. Link, B. L. Trumpower, *Biochemistry* **2003**, *42*, 13637–13645.
[16] M. Lukács, D. Csilla Pálinkás, G. Szunyog, K. Várnagy, *ChemistryOpen* **2021**, *10*, 451–463.
[17] C. D. Douglas, A. V. Dias, D. B. Zamble, *Dalton Trans.* **2012**, *41*, 7782–7791.
[18] D. Rossetto, S. S. Mansy, *Front. Cell Dev. Biol.* **2022**, *10*, 2020–2023.
[19] J. Xu, D. J. Ritson, S. Ranjan, Z. R. Todd, D. D. Sasselov, J. D. Sutherland, *Chem. Commun.* **2018**, *54*, 5566–5569.
[20] M. A. Keller, A. V. Turchyn, M. Ralser, *Mol. Syst. Biol.* **2014**, *10*, 1–12.
[21] K. B. Muchowska, S. J. Varma, J. Moran, *Nature* **2019**, *569*, 104–107.
[22] C. Bonfio, E. Godino, M. Corsini, F. Fabrizi de Biani, G. Guella, S. S. Mansy, *Nat. Catal.* **2018**, *1*, 616–623.
[23] V. Alva, J. Söding, A. N. Lupas, *eLife* **2015**, *4*, 1–19.
[24] V. Sivo, G. D'abrosca, L. Russo, R. Iacovino, P. V. Pedone, R. Fattorusso, C. Isernia, G. Malgieri, *Bioinorg. Chem. Appl.* **2017**, *2017*, 1527247.
[25] S. Scintilla, C. Bonfio, L. Belmonte, M. Forlin, D. Rossetto, J. Li, J. A. Cowan, A. Galliani, F. Arnesano, M. Assfalg, S. S. Mansy, *Chem. Commun.* **2016**, *52*, 13456–13459.
[26] Y. S. Tan, D. P. Lane, C. S. Verma, *Drug Discovery Today* **2016**, *21*, 1642–1653.
[27] L. M. Hunsicker-Wang, A. Heine, Y. Chen, E. P. Luna, T. Todaro, Y. M. Zhang, P. A. Williams, D. E. McRee, J. Hirst, C. D. Stout, J. A. Fee, *Biochemistry* **2003**, *42*, 7303–7317.
[28] L. Valer, D. Rossetto, S. Scintilla, Y. J. Hu, A. Tomar, S. Nader, I. O. Betinol, S. Mansy, *Can. J. Chem.* **2022**, https://doi.org/10.1139/cjc-2021–0237.
[29] V. M. Coghlan, L. E. Vickery, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 835–839.
[30] R. C. Holz, F. J. Small, S. A. Ensign, *Biochemistry* **1997**, *36*, 14690–14696.
[31] Y. Sugiura, K. Ishizu, T. Kimura, *Biochem. Biophys. Res. Commun.* **1974**, *60*, 334–340.

[32] T. A. Link, in *Adv. Inorg. Chem.* **1999**, pp. 83–157.
[33] T. A. Link, W. R. Hagen, A. J. Pierik, C. Assmann, G. von Jagow, *Eur. J. Biochem.* **1992**, *208*, 685–691.
[34] Y. Bromberg, A. A. Aptekmann, Y. Mahlich, L. Cook, S. Senn, M. Miller, V. Nanda, D. U. Ferreiro, P. G. Falkowski, *Sci. Adv.* **2022**, *8*, 1–14.
[35] S. Dagar, S. Sarkar, S. Rajamani, *ChemBioChem* **2022**, *23*, e202200013.
[36] S. Nader, L. Sebastianelli, S. S. Mansy, *Philos. Trans. R. Soc. London Ser. A* **2022**, https://doi.org/10.1098/rsta.2020.0423.
[37] C. Shen, T. Mills, J. Oró, *J. Mol. Evol.* **1990**, *31*, 175–179.
[38] C. Shen, L. Yang, S. L. Miller, J. Oro, *Orig. Life* **1987**, *17*, 295–305.
[39] J. L. Bada, J. H. Chalmers, H. J. Cleaves, *Phys. Chem. Chem. Phys.* **2016**, *18*, 20085–20090.
[40] E. T. Parker, H. J. Cleaves, M. P. Callahan, J. P. Dworkin, D. P. Glavin, A. Lazcano, J. L. Bada, *Origins Life Evol. Biospheres* **2011**, *41*, 201–212.
[41] E. T. Parker, H. J. Cleaves, J. P. Dworkin, D. P. Glavin, M. Callahan, A. Aubrey, A. Lazcano, J. L. Bada, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 5526–5531.
[42] C. S. Foden, S. Islam, C. Fernández-García, L. Maugeri, T. D. Sheppard, M. W. Powner, *Science* **2020**, *370*, 865–869.
[43] H. C. Cheng, *J. Pharmacol. Toxicol. Methods* **2001**, *46*, 61–71.

Manuscript received: April 8, 2022
Revised manuscript received: June 8, 2022
Accepted manuscript online: June 8, 2022
Version of record online: June 23, 2022