



**University of  
Nottingham**

UK | CHINA | MALAYSIA

**A Minimal Peptide Scaffold for  
[4Fe-4S]-Cluster Interrogation  
and Catalysis**

**Andrea Bombana**

**School of Chemistry**

**A thesis submitted to the University of Nottingham for  
the degree of Doctor of Philosophy**

## Declaration

I hereby declare that, except where specific reference is made to other sources, the work contained within this thesis is my own original research. This thesis has been composed by me and has not been submitted, in whole or part, for any other degree, diploma or other qualification. I confirm that appropriate credit has been given within the thesis to the work of collaborators. The following thesis contains results reported in the following publication:

Bombana, A., Shanmugam, M., Collison, D., Kibler, A. J., Newton, G. N., Jager, C. M., Croft, A. K., Morra, S., Mitchell, N. J.; Application of a Synthetic Ferredoxin-Inspired [4Fe4S]-Peptide Maquette as the Redox Partner for an [FeFe]-Hydrogenase; *ChemBioChem*; **2023**; DOI: 10.1002/cbic.202300250

Andrea Bombana

## ACKNOWLEDGEMENTS

Despite being the only one responsible for the content of this thesis, this work could only be achieved with the aid of talented people that helped me throughout my journey.

Above all, I would like to thank my supervisor Dr. Nicholas Mitchell. Your passion and encouragement have always pushed me to give my best, please always keep up this mentality. A special thank goes to my second supervisors Prof. Anna Croft and Dr. Christof Jager for their precious help during the past three years.

I would like to thank all former and current members of the Mitchell group and everyone I shared the lab and the office with. A special mention goes to Fran, Justyna, Salwa, Jo, Lola, and Rhys. I sincerely enjoyed my time with you all and I wish you a successful career, wherever it might bring you.

Thanks to Dr. Luisa Ciano for letting me use her glove box and for sharing her valuable and vast knowledge about EPR. Mi mancheranno le chiacchiere in laboratorio. Thanks to Mr. Mark Guyler for his technical support with the use of the glove box.

Thanks to Dr. Muralidharan Shanmugam and everyone at the EPR facility in Manchester who contributed to the analysis of my samples.

Grazie alla mia famiglia per il sostegno economico e morale fornitomi durante questi anni di studio. Il mio successo accademico è frutto di un'opportunità che mi è stata donata grazie al duro lavoro di mia mamma e di mio papà. Vi sarò per sempre grato dell'educazione che mi avete impartito e degli insegnamenti che mi hanno reso ciò che sono oggi.

Grazie a te Ale per essere sempre al mio fianco e avere la forza di spronarmi ogni giorno ad essere una persona migliore. La tua presenza nella mia vita e la gratitudine e ammirazione che ho nei tuoi confronti hanno un significato che va al di là di ogni parola che io possa cercare per ringraziarti.

## ABBREVIATION LIST

Fe-S	Iron-sulfur
HiPIPs	High-potential iron-sulfur proteins
DNA	Deoxyribonucleic acid
SAM	S-adenosyl-L-methionine
5'-dAdo•	5'Deoxyadenosyl radical
EPR	Electron paramagnetic resonance
ENDOR	Electron nuclear double resonance
PFL-AE	Pyruvate formate-lyase activating enzyme
FNR	Fumarate nitrate reduction
FAD	Flavin adenine dinucleotide
ATP	Adenosine triphosphate
RiPPs	Ribosomally synthesised and post-translationally modified peptides
RNA	Ribonucleic acid
ISCA	Iron-sulfur cluster assembly
βME	2-Mercaptoethanol
DTT	Dithiothreitol
DMSO	Dimethyl sulfoxide
Ac	Acetylated
UV-Vis	Ultraviolet-visible
LMCT	Ligand-to-metal charge transfer
MO	Molecular orbitals
MLCT	Metal-to-ligand charge transfer
DT	Sodium dithionite
NADH	Nicotinamide adenine dinucleotide
SPPS	Solid phase peptide synthesis

Fmoc	Fluorenylmethoxycarbonyl
HEPES	(4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
KCl	Potassium chloride
SEC	Size Exclusion Chromatography
Da	Dalton
EPSRC	Engineer and physical sciences research council
NRF	National research facility
MoFe	Molybdenum-iron
SCE	Saturated calomel electrode
NADPH	Nicotinamide adenine dinucleotide phosphate
LUMO	Lowest unoccupied molecular orbital
HOMO	Highest occupied molecular orbital
IUPAC	International union of pure and applied chemistry
$E_{p,c}$	Cathodic peak
$E_{p,a}$	Anodic peak
$I_{p,a}$	Anodic current
$I_{p,c}$	Cathodic current
TOF	Turnover frequency
Sub	Substrate
SET	Single electron transfer
PC	Photocatalyst
PEG	Polyethylene glycol
HAT	Hydrogen atom transfer
TCEP	Tris(2-carboxyethyl)phosphine
TEMPO	(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl
HRMS	High resolution mass spectrometry
NMR	Nuclear magnetic resonance

ROS	Reactive oxygen species
ONOOH	Peroxynitrous acid
HOCl	Hypochlorous acid
HOSCN	Hypothiocyanous acid
HOBr	Hypobromous acid
HPLC	High performance liquid chromatography
PCA	Principal component analysis
$\Delta_r H^\circ$	Reaction enthalpy
NaOCl	Sodium hypochlorite
ESI	Electron spray ionisation
HPC	High performance computing
OS	Operative system
DMF	<i>N,N</i> -Dimethylformamide
CH <sub>2</sub> Cl <sub>2</sub>	Dichloromethane
DIC	<i>N,N</i> -Diisopropylcarbodiimide
CTC	2-Chlorotriptyl resin
iPr <sub>2</sub> Net	<i>N,N</i> -Diisopropylethylamine
MeOH	Methanol
A	Absorbance
Df	Dilution factor
V	Volume
TFA	Trifluoroacetic acid
TIPS	Triisopropylsilane
TMSOTf	Trifluoromethanesulfonate
RP	Reverse phase
mCPBA	<i>meta</i> -chloroperoxybenzoic acid

## AMINO ACID ABBREVIATION LIST

<b>Amino acid</b>	<b>Three letter symbols</b>	<b>One letter symbol</b>
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Selenocysteine	Sec	U
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

## Table of Contents

ABSTRACT .....	12
1 INTRODUCTION.....	14
1.1 ORIGINS OF IRON-SULFUR CLUSTERS .....	14
1.2 RADICAL SAM ENZYMES .....	15
1.2.1 Mechanism of Radical SAM Enzymes.....	17
1.3 BACTERIAL FERREDOXIN.....	19
1.4 INTRODUCTION TO MAQUETTE CHEMISTRY.....	21
1.5 [4Fe-4S] CLUSTERS IN CATALYSIS.....	22
1.5.1 Radical SAM in Catalysis.....	22
1.6 IN VIVO AND IN VITRO FORMATION OF [4Fe-4S] CLUSTERS.....	26
1.6.1 In Vitro Formation of [4Fe-4S] Cluster Maquettes .....	28
1.6.2 Bacterial Ferredoxin Maquettes.....	29
1.7 SPECTROSCOPICAL TECHNIQUES FOR THE CHARACTERISATION OF [4Fe-4S] CLUSTERS .....	31
1.7.1 Ultraviolet-Visible Spectroscopy .....	31
1.7.2 Electron Paramagnetic Resonance Spectroscopy.....	33
1.8 AIMS AND OBJECTIVES .....	34
2 SYNTHESIS AND CHARACTERISATION OF [4Fe-4S] CLUSTER MAQUETTES .....	37
2.1 AIMS AND OBJECTIVES .....	37
2.2 RESULTS AND DISCUSSION.....	37



2.2.1 Initial Attempts on Radical SAM-type Maquettes.....	37
2.2.2 $\beta$ ME and Single Amino Acid Screening.....	43
2.2.3 Bacterial Ferredoxin-Type Maquettes.....	48
2.2.4 Radical SAM-Type Maquettes.....	71
2.2.5 Reactivity of [4Fe-4S] Cluster Maquettes to Molecular Oxygen.....	75
2.2.6 Exchange of Sulfur by Selenium in [4Fe-4S] Cluster Maquettes ...	81
2.3 CONCLUSION AND FUTURE WORK.....	87
3 REDOX PROPERTIES OF [4Fe-4S] CLUSTER MAQUETTES.....	89
3.1 INTRODUCTION.....	89
3.2 AIMS AND OBJECTIVES.....	89
3.3 ELECTROCHEMICAL PRINCIPLES OF [4Fe-4S] CLUSTERS.....	90
3.4 RESULTS AND DISCUSSION.....	95
3.4.1 Buffer and Controls.....	95
3.4.2 Cyclic Voltammetry of [4Fe-4S] Cluster Maquettes in absence of $\beta$ ME .....	99
3.5 CONCLUSION AND FUTURE WORK.....	112
4 INTEGRATION OF [4FE-4S] CLUSTER MAQUETTES IN AN ELECTRON TRANSPORT PATHWAY.....	114
4.1 INTRODUCTION TO HYDROGENASE.....	114
4.2 AIMS AND OBJECTIVES.....	114
4.3 RESULTS AND DISCUSSION.....	116
4.3.1 Experimental Set-Up.....	116

4.3.2 Hydrogenase Experiments on Purified [4Fe-4S] Cluster Maquette Samples .....	118
4.3.3 Hydrogenase Experiments on [4Fe-4S] Cluster Maquettes in Absence of $\beta$ ME.....	125
4.4 CONCLUSION AND FUTURE WORK.....	133
5 VISIBLE-LIGHT-MEDIATED CYSTEINE DESULFURISATION AND $\beta$ -ARYLATION IN PEPTIDES .....	135
5.1 INTRODUCTION .....	135
5.2 AIMS AND OBJECTIVES .....	135
5.3 METHODOLOGIES FOR THE MODIFICATION OF CYSTEINE .....	136
5.3.1 Introduction to Bioconjugation Strategies with Cysteine .....	136
5.3.2 Michael Addition Strategy .....	136
5.3.3 Photocatalytic Modification of Cysteine .....	138
5.4 RESULTS AND DISCUSSION.....	147
5.4.1 Introduction.....	147
5.4.2 Boronic Acid Screening .....	148
5.4.3 Proposed Mechanism .....	154
5.5 CONCLUSION AND FUTURE WORK.....	155
6 REDUCTION MONITORING OF DISULPHIDE BONDS IN PEPTIDES .....	157
6.1 INTRODUCTION .....	157
6.2 AIMS AND OBJECTIVES .....	160
6.3 RESULTS AND DISCUSSION.....	162
6.3.1 Preliminary Computational Calculations .....	162

6.3.2 Chemical Reduction of Disulphide Bonds.....	175
6.4 CONCLUSION AND FUTURE WORK.....	200
7 THESIS CONCLUSIONS AND FUTURE WORK.....	203
8 EXPERIMENTAL.....	206
8.1 GENERAL METHODS.....	206
8.2 GENERAL PROTOCOLS.....	210
General Protocols for Chapter Two.....	210
General Protocols for Chapter Three.....	214
General Protocols for Chapter Four.....	214
General Protocols for Chapter Five.....	216
General Protocols for Chapter Six.....	218
General Protocol for Appendix 2.....	219
8.3 PEPTIDE SYNTHESIS.....	220
8.4 COMPOUNDS SYNTHESIS.....	238
8.5 SYNTHESIS OF MODIFIED PEPTIDES.....	244
REFERENCES.....	251
APPENDIX 1.....	266
APPENDIX 2.....	285

## ABSTRACT

The S-adenosyl-L-methionine (SAM) dependant superfamily of enzymes is capable of initiating a variety of biologically essential radical reactions. Likewise, bacterial-type ferredoxins mediate electron transfer to facilitate a broad range of biological redox processes. Both families of enzymes host a cubic  $[4\text{Fe-4S}]^{2+/+}$  motif that initiates and mediates the reactivity of these proteins.

The synthesis of short peptide maquettes that mimic the conserved cysteine-rich motif responsible for coordinating the  $[4\text{Fe-4S}]$  cluster represents an exciting and easy-to-synthesise tool with which to explore the utility of these systems. In Chapters 2 and 3 of this thesis we describe the preparation of a series of cluster maquettes to gather information regarding the influence of the local electronic and steric environment on the formation and redox potential of the cluster, as well as exploring the stability of the clusters to aerobic conditions. Despite being routinely formed under strict anaerobic conditions, the use of photoactivated NADH is herein described for the successful reduction of oxidised  $[4\text{Fe-4S}]$  cluster when reconstituted in presence of atmospheric oxygen. Different peptides mimicking both the radical SAM and bacterial-type ferredoxins families are investigated in this thesis with the aim of producing an engineered maquette. In Chapter 4 we demonstrate the integration of synthetic, bacterial ferredoxin-type maquette into a  $\text{H}_2$ -powered electron transport chain to showcase the potential utility of these systems within synthetic biology.

Site-selective methods for peptide and protein modification allow for the late-stage installation of various moieties into peptides and enable protein engineering via bioconjugation. In Chapter 5 of this thesis, we describe a visible-light-mediated desulfurative  $\text{C}(\text{sp}^3)\text{-C}(\text{sp}^2)$  bond forming method that enables the site-selective installation of aromatic analogues into small peptides.

Disulphide bonds are well known for playing a key role in defining and stabilising proteins structures. More recently, data has revealed that disulphide bonds are crucially involved in oxygen sensing by interacting with oxidising agents. In Chapter 6 of this thesis, we investigate the effect that different amino acids in close proximity to a cysteine residue involved in the disulphide bond can have on the reduction of the

disulphide in small peptide models. Computational simulations were conducted to integrate the experimental data.

# 1 INTRODUCTION

## 1.1 ORIGINS OF IRON-SULFUR CLUSTERS

Iron-sulfur (Fe-S) clusters are classed among the oldest cofactors on earth, with their presence in numerous metabolic reactions. These systems normally take part in redox reactions and are composed of iron and inorganic sulfur atoms. Despite the absence of clear evidence regarding the origins of Fe-S clusters, these are commonly accepted to have preceded oxygenic photosynthesis; the metabolic pathway that enabled complex life formation on earth.<sup>1</sup> When oxygen became more abundant in the atmosphere, microbes adopted different strategies to prevent Fe-S clusters from oxidative degradation. Among these adaptations we can find restriction of the clusters to anoxic environment, reduction of oxygen and other oxidative species, and mechanisms to repair Fe-S clusters.<sup>2</sup> All these actions to protect the integrity of Fe-S clusters showcase the importance these cofactors played in supporting life. Fe-S proteins were firstly discovered in the 1960 due to a characteristic electrochemical feature that emerged after exposure of these proteins to reducing agents. This peculiarity called “g value” had never been observed for any other metalloprotein.<sup>3</sup> Following this discovery, hundreds of thousands of proteins were identified that belonged to this family. The ease with which this family of metalloproteins was explored is due to the enormous abundance of these proteins, their exclusive spectroscopic features, and their highly charged nature, which makes their purification and analysis uncomplicated.<sup>4</sup> Metalloproteins are generally classified in different sub-families based on the number of iron and sulfur atoms in their metal centres and on their electrochemical properties. The major groups are: rubredoxins, characterised by a [1Fe-4S] cluster; ferredoxins, generally with low-potential and [2Fe-2S], [3Fe-4S], and [4Fe-4S] clusters; Rieske proteins, characterised by high-potential and [2Fe-2S] clusters; and high-potential iron-sulfur proteins (HiPIPs), with high-potential [4Fe-4S] clusters.<sup>4</sup> Various Fe-S proteins carry multiple binding sites and can therefore host more than one cluster.

The most common combinations of iron and sulfur in clusters encountered in nature are represented in Figure 1. These include rhombic [2Fe-2S], and cuboidal [3Fe-4S] and [4Fe-4S].

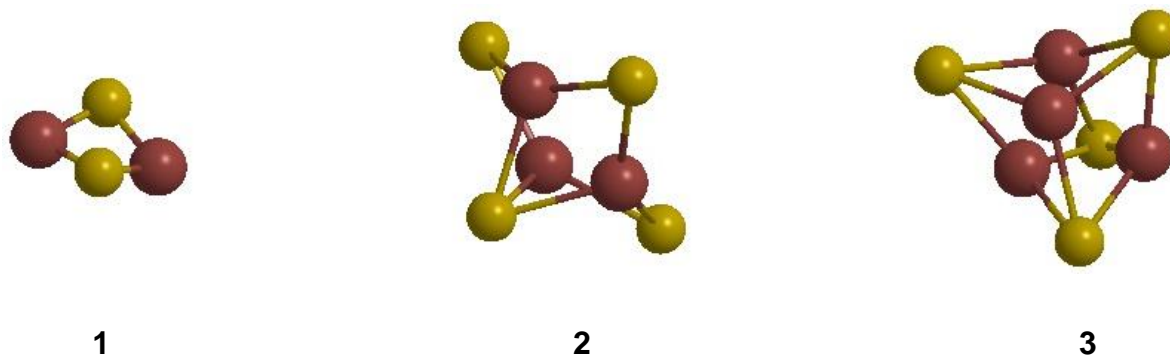


Figure 1 – Most common Fe-S clusters found in nature: 1 rhombic [2Fe-2S], 2 cuboidal [3Fe-4S], and 3 cuboidal [4Fe-4S].

Fe-S clusters are incorporated by proteins via coordination between the iron ions and the side chains of the amino acids that compose the protein sequence. In the early stages, it is thought that prebiotically formed peptides stabilised the clusters and that as life became more complex, the short peptides evolved into proteins.<sup>5</sup> The thiolate group of Cys is the most common residue found in the coordination of Fe-S clusters in proteins, despite other ligands have also been identified.<sup>6</sup> During the course of evolution, various family of enzymes have developed to incorporate Fe-S clusters for a wide range of essential-to-life tasks including catalysis, electron transfer, nitrogen fixation, gene expression, cofactor biosynthesis, and DNA repair.<sup>7</sup> In this document, we focused on two family of enzymes: radical SAM and bacterial ferredoxin.

## 1.2 RADICAL SAM ENZYMES

The radical SAM superfamily of enzyme was first classified in 2001 and subsequent bioinformatic studies have revealed it to be one of the biggest protein superfamilies in nature, with over 400,000 annotated sequences to date stretching over all kingdoms of life.<sup>8–10</sup> This class of enzymes adopt a [4Fe-4S]-cluster and a SAM molecule to initiate a broad range of radical reactions, most of which occur via generation of a 5'-deoxyadenosyl radical (5'-dAdo•) intermediate, as displayed in Scheme 1.

The 5'-dAdo• radical was successfully characterised by Yang *et al.* via photoinitiated electron transfer from the reduced form of the cluster [4Fe-4S]<sup>1+</sup> to the coordinated SAM.<sup>11</sup> Isotopically labelled SAM was used in combination with electron paramagnetic resonance (EPR) and electron nuclear double resonance (ENDOR) spectroscopy to

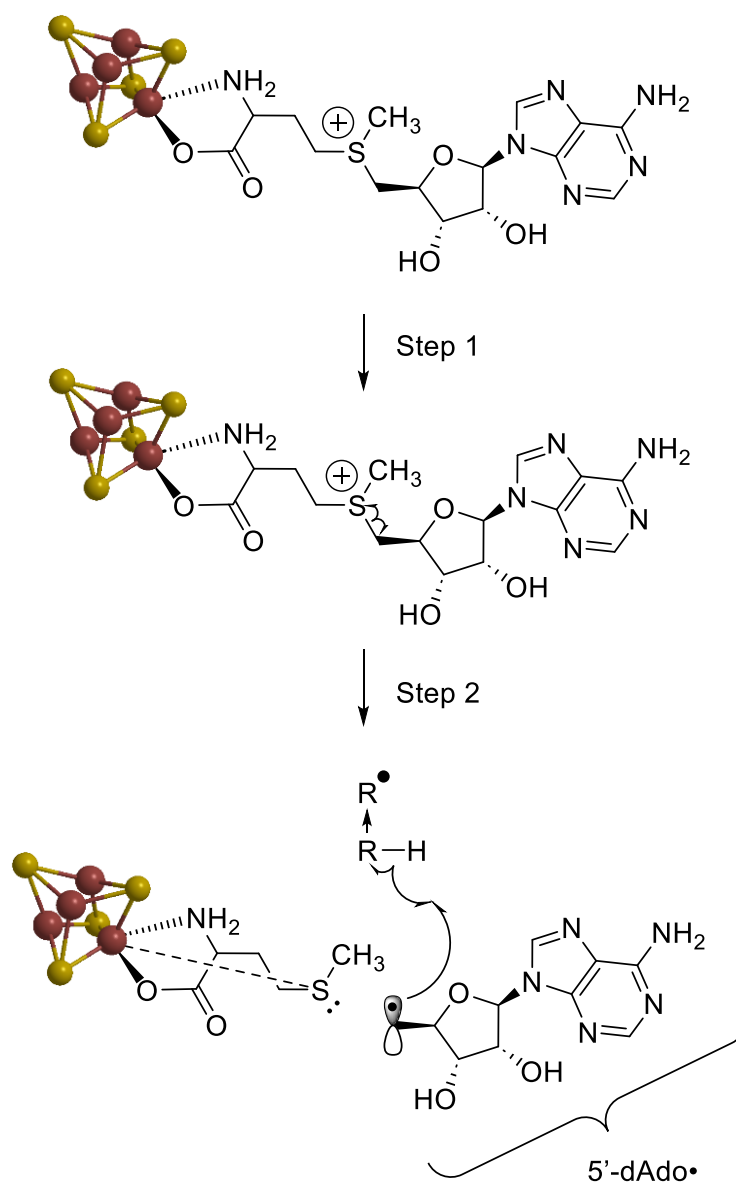
identify the radical species. To obtain further details, SAM was labelled with  $^{15}\text{N}$  at the amino group and  $^{17}\text{O}$  and  $^{13}\text{C}$  at the carboxylate group. Labelled SAM was added to the reduced cluster in presence of pyruvate formate-lyase activating enzyme (PFL-AE), an enzyme that belongs to the radical SAM family. ENDOR studies registered signals that were attributed to the interaction between the amino acid moiety of SAM to the unique iron ion of the reduced cluster.

Even though 2001 marks the identification of the radical SAM protein superfamily, the discovery of iron metalloenzymes employing SAM to initiate radical reactions precedes this date by more than a decade.<sup>12</sup> For instance, previous studies on the activation of PFL proved that it required the generation of a stable protein radical, and was initiated by the presence of SAM, iron, and a molecule from the cell extract now confirmed to be the PFL-AE.<sup>13,14</sup> The radical on PFL was then revealed to be located on a specific Gly residue and has been one of the first stable protein radicals to be characterised.<sup>15</sup> Eventually, PFL-AE was proved to contain a catalytically crucial Fe-S cluster and to implement SAM as a necessary component for the PFL activation.<sup>16,17</sup>

The key motif of most radical SAM enzymes ( $\geq 90\%$ ) is a conserved **CX<sub>3</sub>CX $\phi$ C** (where  $\phi$  = Tyr, Phe, His, or Trp), in which the three Cys residues are crucial for the recruitment of the [4Fe-4S]-cluster, whereas the fourth unique iron site is complexed to the amino and carboxylate groups of the Met component of SAM, as depicted in Scheme 1.<sup>18,19</sup>

There also exist variations to this pathway with a different number of residues in between the three Cys, showing how diverse the primary structure used to generate this three-dimensional unit can be.<sup>20</sup>



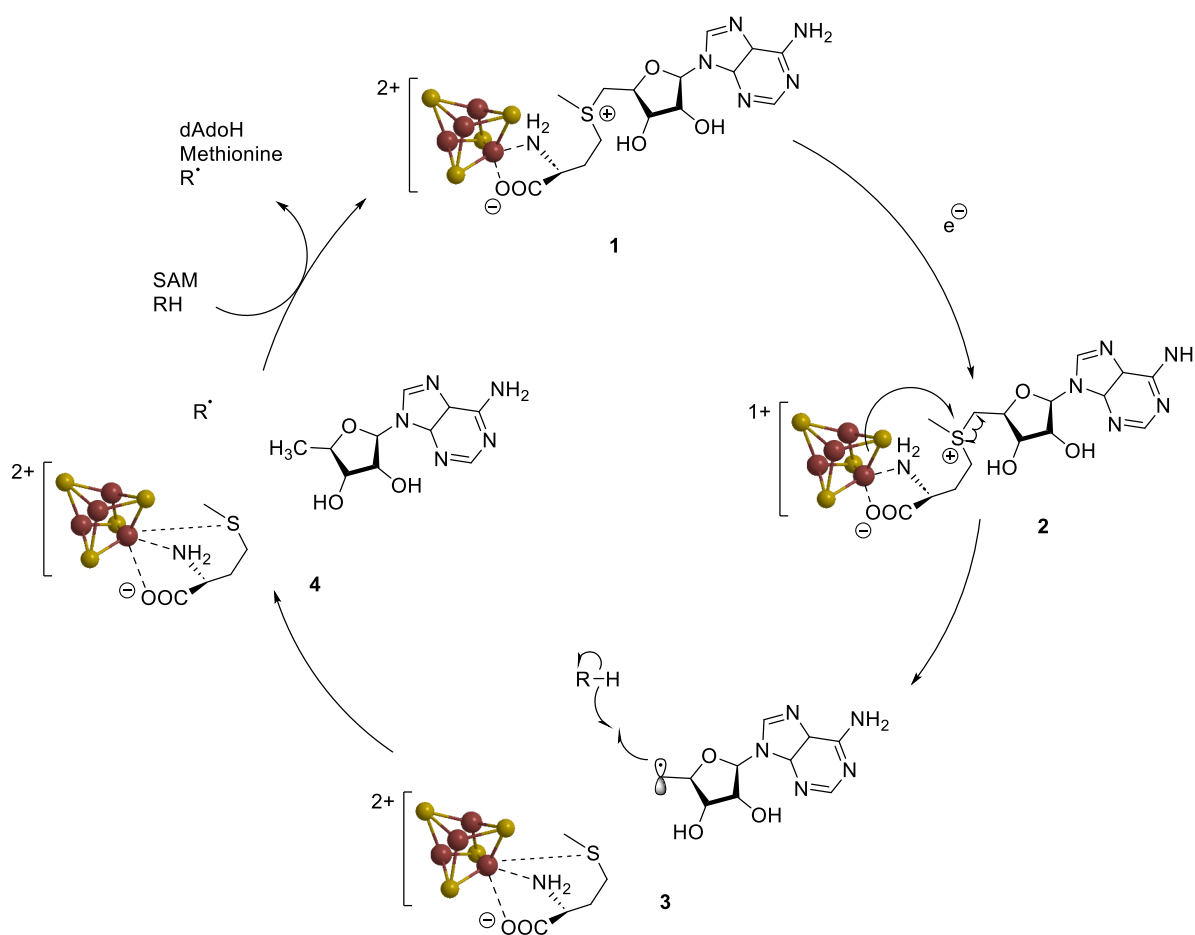


*Scheme 1 – Formation of the 5'-dAdo• radical via homolytic cleavage (Step 1) and hydrogen atom abstraction (Step 2).*

### 1.2.1 Mechanism of Radical SAM Enzymes

The structure of the [4Fe-4S]-cluster centre and its activation in terms of reactivity are thought to be analogous throughout all the members of the group.<sup>21</sup> Until recently, the well accepted mechanism for the radical SAM enzymes involved the coordination of the unique, non-Cys coordinated iron of the [4Fe-4S]-cluster to a SAM molecule through the amino and carboxylic acid groups of the Met component of SAM (Scheme 2, 1).<sup>22</sup> An electron transfer step from the reduced and catalytically-active state of the

radical SAM enzymes [4Fe-4S]<sup>+</sup>-cluster (Scheme 2, **2**) to SAM triggers the inner sphere reductive cleavage of the sulfonium bond S-C5' of SAM. The electron in this step is provided by redox active species present in living organisms, such as NADH. The intermediate 5'-dAdo• radical immediately abstracts a H-atom from a substrate molecule (Scheme 2, **3**) and generates the primary substrate radical (R•) and dAdoH (Scheme 2, **4**).<sup>23</sup> Overall, from a SAM molecule and a general substrate R-H, this mechanism yields a radical that can undergo organic reactions, as well as a Met molecule coordinated to the oxidised [4Fe-4S]<sup>2+</sup> cluster and a 5-deoxyadenosine molecule.<sup>24</sup> Examples include: the formation of α-, β-, and γ-thioether linkages;<sup>25–28</sup> carbon-carbon bond formation;<sup>29–33</sup> methylation;<sup>34,35</sup> epimerisation;<sup>36,37</sup> and tyramine excision.<sup>38</sup>



*Scheme 2 – The mechanism for a radical SAM enzyme begins with the formation of the 5'-dAdo• radical by one electron reductive cleavage of a [4Fe-4S] cluster-bound SAM, which then withdraws a proton from the substrate (R) to generate the active radical (R•).*

More recently, the organometallic intermediate  $\Omega$  (Figure 2) has been successfully characterised using EPR and ENDOR spectroscopy using both SAM and [4Fe-4S]

cluster isotopically labelled, the latter with  $^{57}\text{Fe}$ . In this intermediate, the unique iron ion of the cluster is directly bound to the adenosyl moiety of SAM through a Fe-C bond.<sup>39</sup> The mechanism of formation of this organometallic intermediate has not been uncovered yet. However, the most plausible theory involves the reductive cleavage of SAM to generate the 5'-dAdo• radical, which then undergoes oxidative addition to the oxidised  $[\text{4Fe-4S}]^{2+}$  cluster.<sup>40</sup> The presence of  $\Omega$  in various enzymes proves that this species is a pivotal intermediate for the mechanism of radical initiation, even if clear evidence in support of this theory is yet to be discovered. It is plausible that  $\Omega$  plays a role in re-storing the 5'-dAdo• radical after its formation from homolytic cleavage of the Fe-C5' bond (Scheme 1).<sup>40</sup>

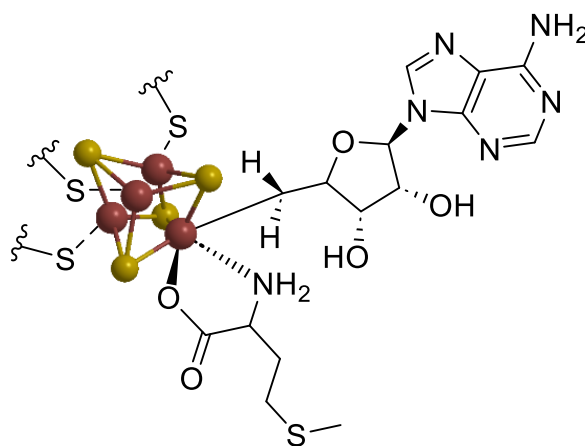
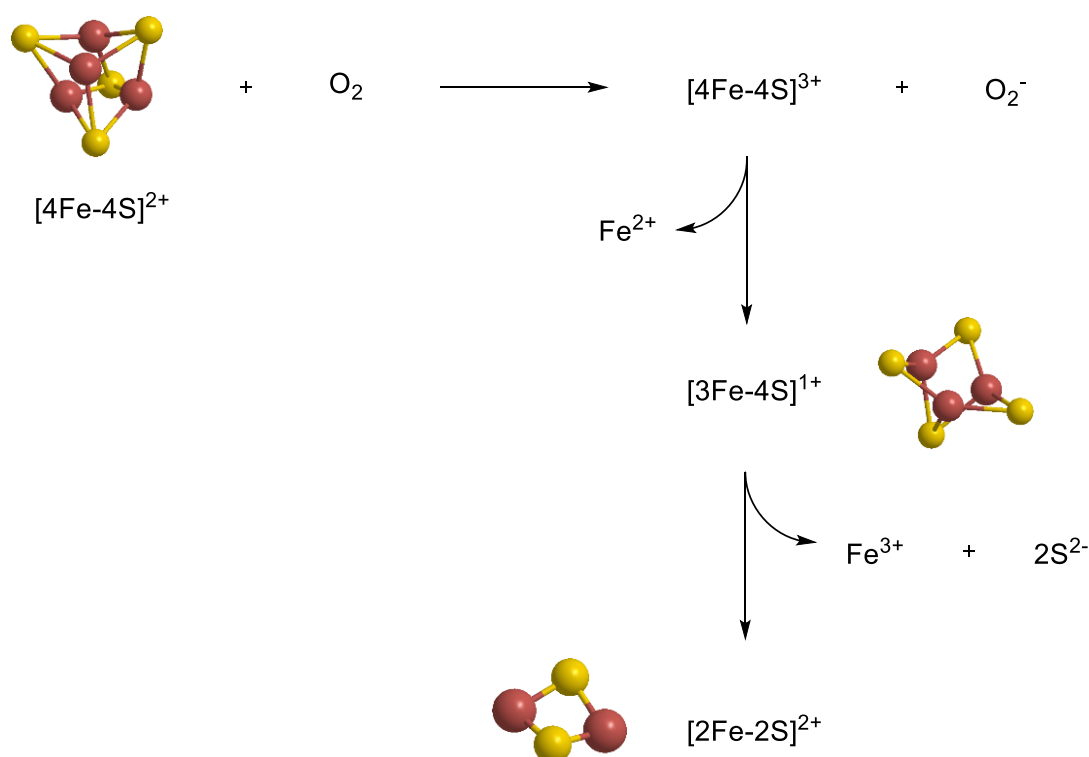


Figure 2 – Structure of the organometallic intermediate  $\Omega$ .

### 1.3 BACTERIAL FERREDOXIN

Radical SAM enzymes have raised interest due to the numerous reactions they catalyse and the intricate mechanism through which they achieve their purpose. Conversely, bacterial ferredoxins have so far mainly been investigated to gather more information about the structural requirements of the protein sequences and the redox properties of the clusters they incorporate. The bacterial ferredoxin family is involved in biological pathways such as photosynthesis and respiration, where its main function is to act as an electron carrier exploiting the redox abilities of the incorporated Fe-S clusters.<sup>41</sup> A class of bacterial ferredoxin, called hydrogenases, employs this electron-mediating catalyst activity for the biological production or use of hydrogen gas.<sup>42</sup> For

this reason, bacterial ferredoxins have developed a broad variety of cluster nuclearities and coordination geometry, being therefore able to accommodate all the different types of Fe-S clusters and using non-cysteiny amino acids (*i.e.*, His, Arg, Lys, and Ser) to coordinate them.<sup>43</sup> An interesting characteristic of bacterial ferredoxins is that not only they dispose of different cluster nuclearities, but they are also able to convert one cluster to another. This ability is the result of an intricate mechanism that allows bacterial ferredoxins to transport electrons across different species to finally achieve the required redox process. An example is given by the fumarate and nitrate reduction (FNR) regulator, as depicted in Scheme 3.<sup>44</sup>



*Scheme 3 – Mechanism for the reaction between a [4Fe-4S]<sup>2+</sup> cluster and O<sub>2</sub>.*

In the first step the [4Fe-4S]<sup>2+</sup> cluster reacts with O<sub>2</sub> to give a superoxide anion (O<sub>2</sub><sup>-</sup>) and the unstable superoxidised species [4Fe-4S]<sup>3+</sup> cluster. The latter species immediately releases an Fe(II) atom to give a [3Fe-4S]<sup>1+</sup> cluster. Subsequently, an Fe(III) and two sulfide ions S<sup>2-</sup> are lost from the [3Fe-4S]<sup>1+</sup> cluster to generate the final product of this mechanism: a [2Fe-2S]<sup>2+</sup> cluster. Confirmation of this mechanism were given by the EPR observation of a transient species with characteristics attributable to the [3Fe-4S]<sup>1+</sup> cluster. The rate of formation and loss of this latter species matched with a two-steps mechanism, where the [3Fe-4S]<sup>1+</sup> cluster, after being formed, is

converted into a  $[2\text{Fe}-2\text{S}]^{2+}$  cluster. Moreover, the presence of the superoxide anion produced in the first step has been demonstrated by reduction of oxidised cytochrome c, further confirming this mechanism of action.<sup>45</sup>

The crucial structural difference between radical SAM and bacterial ferredoxin families is the number of Cys residues used to coordinate the cluster. Whereas the former adopts SAM to coordinate the unique iron ion of the Fe-S centre, the latter is provided with four Cys residues that fully coordinate the  $[4\text{Fe}-4\text{S}]$  cluster. The consistent motif of the bacterial ferredoxin family is  $\mathbf{C}_1\text{XXX}\mathbf{C}_2\text{XX}\mathbf{C}_3\text{XX}\mathbf{C}_4$ , even though variations have been described where  $\mathbf{C}_1$  is located in several amino acids distance from the core motif. Ferredoxins generally use a cofactor such as flavin adenine dinucleotide (FAD) or adenosine triphosphate (ATP) to achieve low reduction potentials and initiate the single electron transfer process.<sup>46</sup>

## 1.4 INTRODUCTION TO MAQUETTE CHEMISTRY

A maquette is a small peptide with 3 to 16 residues that is either designed or inspired from a biological protein. In maquette chemistry, this short peptide-based synthetic models are designed to mimic the macromolecular system of a protein. These miniaturised proteins contain the minimum set of constituents required for an accurate replica of defined structures in order to be able to reproduce the desired functions of the enzyme they mimic. They are generally implemented in this context to replicate the binding site of Fe-S cluster enzymes to investigate specific properties of these systems. Questions addressed with the aid of synthetic maquettes include structural composition of the amino acid sequence, chemical reactions and catalytic activity occurring around the metal centre, and electronic properties of the peptide-cluster complex.<sup>47</sup> In case of radical SAM enzymes, given a parent motif  $\mathbf{Y}\mathbf{C}\mathbf{X}_3\mathbf{C}\mathbf{X}_2\mathbf{C}\mathbf{Y}$ , the synthetic peptide could act as a scaffold for the incorporation of a  $[4\text{Fe}-4\text{S}]$ -cluster. Furthermore, short peptides can extend the capability of the corresponding enzymes due to the ability of covering a more extended conformational space being easier to synthesise and isolate. For this reason, maquette chemistry can be implemented not only to explore native enzymes, thus replicating their functions, but both the

intervening (X) and flanking (Y) amino acids can be exchanged to investigate novel combinations that can possibly lead the sequence to show specific characteristics.<sup>48</sup>

Because the main characteristic of bacterial ferredoxin arises from its redox properties, most peptide models replicating this class of enzyme have been designed to untangle the reduction potential properties of the Fe-S clusters incorporated by these systems.<sup>49</sup> The redox properties of this class of enzyme are crucial for understanding how they catalyse electron transport processes and has therefore been the main focus in most research groups.

In radical SAM enzymes, Galambas *et al.*, recently proved that maquette chemistry can also be extended to the radical SAM enzymes motif, and appropriate peptides can successfully coordinate the redox active [4Fe-4S]-cluster.<sup>20</sup> In their study they investigated both bio-inspired and nascent radical SAM maquettes consisting of 7- to 9-mer peptides, obtaining a high degree of reconstitution, from 80 to 100%. For this reason, maquette chemistry could possibly be implemented to investigate the stability, reactivity, and utility of the vast radical SAM family of enzymes by using miniaturised proteins as scaffolds for the broad radical reactions to take place.

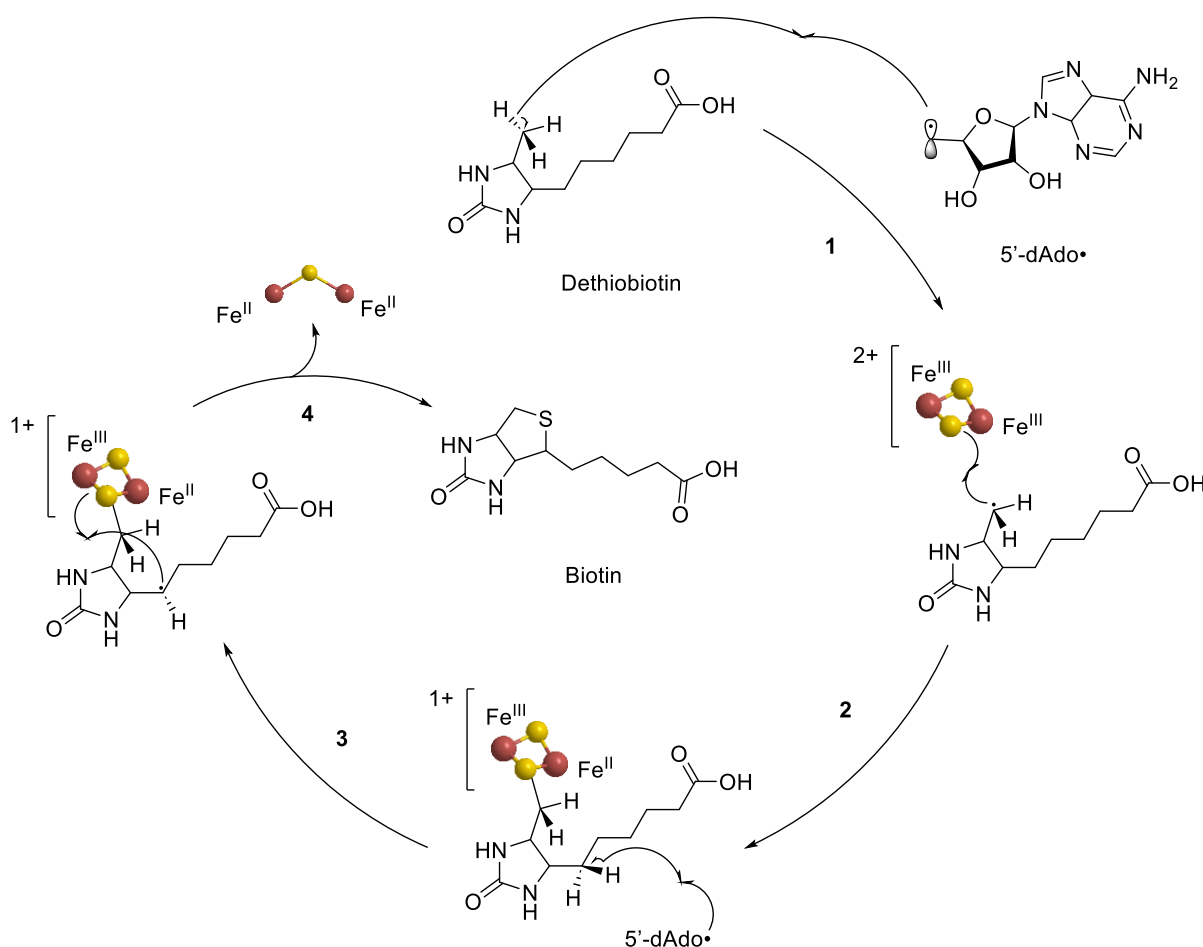
## 1.5 [4Fe-4S] CLUSTERS IN CATALYSIS

### 1.5.1 Radical SAM in Catalysis

Over the past decades, an increasing number of organic reactions in nature have been proved to be catalysed by radical SAM enzymes. Due to their radical-based reactivity, these enzymes can be implemented in a wide range of reactions. Among the first discovered radical SAM enzymes there was biotin synthase, which catalyses the thiophene ring formation during the last step of biotin synthesis (Scheme 4).<sup>50</sup> This transformation converts dethiobiotin to biotin by inserting a sulfur atom between the non-activated carbons C6 and C9 of dethiobiotin.<sup>51</sup>

The 5'-dAdo• (see Scheme 2) radical performs H-atom abstraction to generate a dethiobiotinyl C9 carbon radical (Scheme 4, 1). This radical is immediately sequestered by a secondary [2Fe-2S]<sup>2+</sup> cluster presented in the enzyme, to form a C-

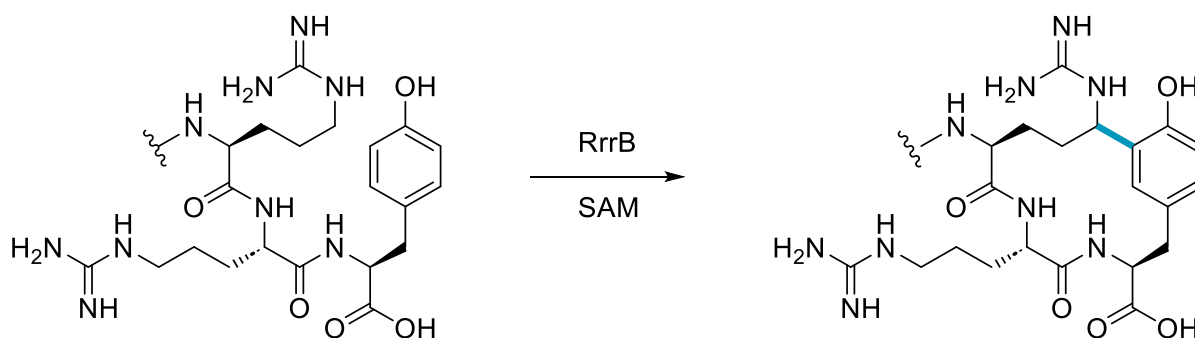
S bond and consequently reduce an iron atom from  $\text{Fe}^{3+}$  to  $\text{Fe}^{2+}$  (Scheme 4, **2**). At this stage, a second radical is generated from the  $5'$ -dAdo $\bullet$  intermediate by abstracting a H-atom on the C6 carbon of dethiobiotin (Scheme 4, **3**). The radical attacks the sulfur of the  $[2\text{Fe}-2\text{S}]$  cluster to yield the thiophene ring of biotin and an unstable  $2\text{Fe}-\text{S}$  cluster that possibly dissociates (Scheme 4, **4**).<sup>50</sup> This example highlights the utility of Fe-S clusters as sulfur donors in organic reactions.



Scheme 4 – Proposed mechanism for the synthesis of biotin from dethiobiotin, catalysed by the radical SAM enzyme biotin synthase.

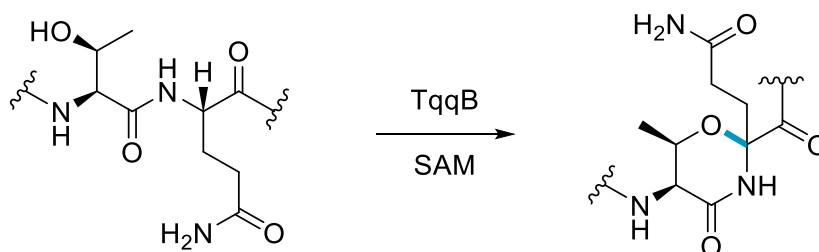
More recently, Caruso *et al.* identified a wide range of ribosomally synthesised and post-translationally modified peptides (RiPPs) gene clusters that are responsible for producing one or more radical SAM enzymes.<sup>52</sup> These modified peptides encode metalloenzymes for the installation of various alterations, some yet to be discovered. RiPPs consist of peptides with ribosomal origin that can undergo enzymatic post-translational modification, which are chemical transformations that occur after the

peptide sequence has been synthesised (translated) following the transcription of DNA to RNA in the nucleus of the cells.<sup>53</sup> These systems are capable of catalysing chemical modifications that plays crucial roles in challenging biosynthetic pathways and mimicking their functions is remarkably difficult. For instance, the radical SAM enzyme RrrB links the inactivated  $\delta$ -carbon of an Arg side chain to the *ortho* position of a Tyr-phenol via the installation of a macrocyclic C-C bond (Scheme 5). This post-translational crosslink gives rise to an unprecedented 14-membered macrocycle. Macrocycles are well known for facilitating the transport of ions across the hydrophobic membrane of cells. Furthermore, macrocycles are becoming more popular in drug discovery due to their improved pharmacokinetics despite the relatively low molecular weight, and for showing good drug-like properties for challenging protein targets.<sup>54</sup>



Scheme 5 – Reaction carried out by the radical SAM enzyme RrrB with the newly installed bond in light blue.

Clark and co-workers recently characterised TqqB, a radical SAM enzyme able to link the oxygen of a Thr side chain to the  $\alpha$ -carbon of a Gln residue, as highlighted in Scheme 6.<sup>26</sup> This reaction constitutes the first aliphatic ether bond formation performed by a radical SAM enzyme and provides a heterocyclisation strategy for peptides.



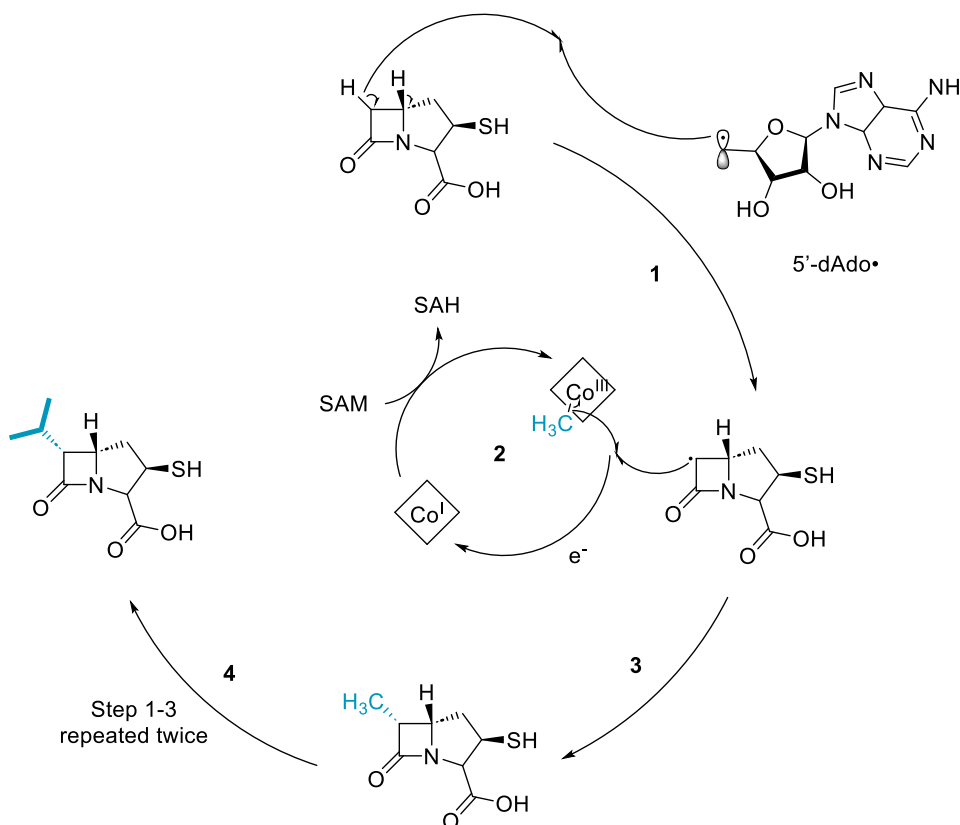
Scheme 6 – Reaction carried out by the radical SAM enzyme TqqB with the newly installed bond in light blue.



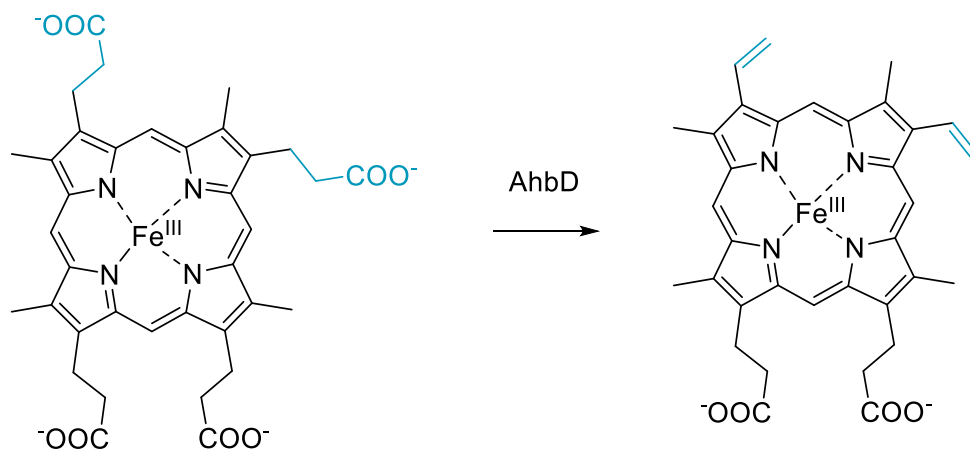
Sinner *et al.*, investigated the activity of TokK, a cobalamin-dependent radical SAM methylase that has been confirmed to catalyse three methylation reactions in a sequential mechanism to form an isopropyl group attached to the C6 carbon of a pantetheine-containing carbapenem substrate (Scheme 7).<sup>55</sup> What is particularly interesting about this radical SAM enzyme family is that they use SAM in two different ways. One equivalent undergoes reductive cleavage via the [4Fe-4S]-cluster to generate the 5'dAdo• radical (Scheme 7, **1**), whereas the second is used as a methyl donor to obtain the active methyl-cobalamin cofactor (Scheme 7, **2**). The radical obtained in the first step accepts a methyl radical from the cobalamin complex obtained in the second step to yield the methylated product (Scheme 7, **3**). This methylation is repeated three times in total to achieve the isopropyl moiety (Scheme 7, **4**).

Kühner *et al.* studied the radical SAM heme synthase AhbD, which is capable of catalysing the oxidative decarboxylation of two propionate side chains of iron-coproporphyrin III into the corresponding vinyl groups of heme (Scheme 8).<sup>56</sup> Interestingly, AhbD contains two redox active [4Fe-4S]-clusters of which one is needed for the canonical reductive cleavage of SAM, whereas the second one is thought to play a crucial role in electron transfer from the reaction intermediate to a so far unknown electron acceptor during the final step.

The role of heme in most living organisms is crucial for catalysing enzyme activity. Being radical SAM enzymes and more in general [4Fe-4S] clusters often found in cooperation with this cofactor, showcases the importance of these classes of enzymes that incorporate an Fe-S cluster for the correct functioning of life.



Scheme 7 – Reaction carried out by the radical SAM enzyme TokK with the newly installed isopropyl moiety in light blue.



Scheme 8 – Reaction carried out by the radical SAM enzyme AhbD with the decarboxylated moiety in light blue.

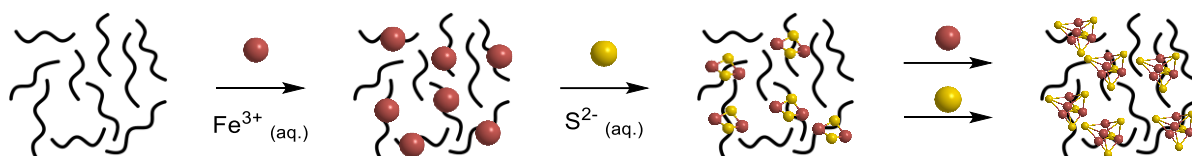
## 1.6 IN VIVO AND IN VITRO FORMATION OF [4Fe-4S] CLUSTERS

The mechanism for the formation of [4Fe-4S] clusters *in vivo* has not been completely unveiled yet. Studies on the assembly of [4Fe-4S] clusters in mitochondria both in

human and yeast cells proved that the process is achieved by the action of specific proteins named iron-sulfur cluster assembly (ISCA).<sup>57</sup> Firstly, a [2Fe-2S] cluster is *de novo* assembled by seven different proteins starting from elemental iron and sulfur obtained from Cys desulfurase.<sup>58</sup> Secondly, the [2Fe-2S] cluster is transferred to a different protein that inserts the cluster into mitochondrial [2Fe-2S]-requiring proteins. Finally, two [2Fe-2S]<sup>2+</sup> clusters are assembled into a [4Fe-4S]<sup>2+</sup> cluster.<sup>59</sup> It is plausible that the same mechanism, governed by different proteins, takes place in the cytosol and nucleus of human cells; however, further studies will be required to confirm this hypothesis.

The biophysical and chemical principles that govern the formation of synthetic peptide cluster-maquettes *in vitro* are, similarly to *in vivo*, not clear. Using molecular dynamic simulations, Hanscam *et al.* investigated the secondary structure of aqueous peptide models in the first stage of maquettes formation.<sup>60</sup> The mechanism they propose is described in Scheme 9. Firstly, the dissolved short peptides are solvated in the buffer solution. As the iron solution is added, the ferric ions (Fe<sup>3+</sup>) are bound to the peptide at their deprotonated Cys residues and as the sulphide ions (S<sup>2-</sup>) are added, they spontaneously assemble into [2Fe-2S]<sup>2+</sup>-clusters. Eventually, upon addition of the final aliquots of both FeCl<sub>3</sub> and Na<sub>2</sub>S solutions two more ferric and sulphide ions are incorporated into the cluster, resulting in the formation of peptide maquettes each coordinating an [4Fe-4S]<sup>1+</sup>-cluster. This final step is thought to be more complex than a reductive coupling of two [2Fe-2S]<sup>2+</sup> clusters, however more studies would be required for a definite answer.

In case of radical SAM maquettes, the unique Fe site of the [4Fe-4S]-cluster that would be bound to the amino acid moiety of SAM, is instead coordinated by 2-mercaptoethanol (βME) acting as a fourth ligand in addition to the three Cys of the synthetic peptide. The use of βME is also crucial to ensure all the Cys residues in the peptide sequence are kept in the correct oxidation state. When in oxidising conditions, the thiols (-SH) of Cys residues can react between each other to form disulphide bonds (S-S). The formation of disulphide bonds is detrimental for the correct synthesis of Fe-S clusters as the lack of thiols able to coordinate the iron of the cluster could lead to lower cluster yields. For this reason, reducing agents such as βME and dithiothreitol (DTT) are often found as additives in Fe-S -cluster maquette reconstitution protocols.



Scheme 9 – Proposed mechanism for the formation of [4Fe-4S] clusters *in vitro*.

### 1.6.1 In Vitro Formation of [4Fe-4S] Cluster Maquettes

Early studies in the field of Fe-S cluster reconstitution demonstrated that these clusters can be easily formed by simply mixing iron and sulfur in a reductive environment.<sup>61</sup> However, the reconstitution of these clusters in peptides and proteins had proved more challenging, showcasing the sensitivity of these systems to specific environments and the importance of robust protocols for their synthesis. The reason for that is explained by the fact that the insertion of a Fe-S cluster into a protein or a peptide is not a spontaneous process, as the formation of the cluster itself, but rather a catalysed one.<sup>62</sup> The [2Fe-2S] cluster is considered to be the elementary building block for other, more complex clusters; and as described in Scheme 9, [4Fe-4S] clusters are currently thought to be formed by a sort of reductive coupling between two [2Fe-2S] clusters.<sup>63</sup>

The past and current literature has focused on getting information about what are the minimal requirements for binding a Fe-S cluster using small peptides that mimic the more elaborated structures of Fe-S proteins. There are two main approaches for the synthesis of [4Fe-4S] cluster peptide maquettes. In one method the cluster is transferred to the peptide from a cluster donor, whereas in the second the cluster is built in solution and directly coordinated by the peptide. An example of the first approach involves the use of  $[4\text{Fe-4S}(\text{S-t-Bu})_4]^{2-}$  as the cluster donor in presence of a base and in a solvent such as dimethyl sulfoxide (DMSO).<sup>64</sup> The cluster is transferred to the peptide by a ligand exchange mechanism with the tert-butyl group acting as a leaving group. Because this technique has not been thoroughly investigated yet, its application is limited. For example, there is no precedence of the use of this protocol in aqueous media, and the concentration or nature of the coordinating ligands and the neighbouring amino acid on the peptide chain has not been fully explored yet. The second approach is the most studied and the principle is described in Scheme 9. In this case the Fe-S cluster is generated *in situ* and directly coordinated by the thiols of

the Cys residues that compose the peptide. Because this coordination happens between the thiols of the peptide and the Fe(III) in solution, peptides are protected on the C- and N-termini to avoid S,O- and S,N-coordination and ensure S-only coordination.<sup>65</sup>

The amino acids surrounding the Cys residues has also been proved to play a crucial role in the stability of the cluster maquette. Overall, it was showed that hydrophobic amino acids favour the electrostatic interaction between the ionic moiety of the Fe-S cluster and the peptide environment itself.<sup>66</sup> The synthesis of peptide maquettes begins with the equilibration of the peptide in an aqueous buffer in presence of a reducing agent (generally  $\beta$ ME or DTT) in concentration that can vary from 2 to 10% in volume of the final solution. The use of a reducing agent is required to ensure the thiols of the Cys residues are in the reduced state and not oxidised into disulphides.<sup>66</sup> The most adopted cluster sources are FeCl<sub>3</sub> and Na<sub>2</sub>S. These reagents are normally used with a six-fold excess with respect to the peptide.<sup>67</sup> The presence of an excess in solution of these ions favours the formation of the cluster, and any residual Fe<sup>3+</sup> and S<sup>2-</sup> ions form an FeS salt that is insoluble in the aqueous media and easily removed from the cluster maquette. Despite the initial and final concentration of the ions is not crucial, provided they are present in excess to the peptide, the order of addition seems to play a crucial role for the reasons depicted in Scheme 9. Therefore, the iron source is generally slowly added to the peptide before the addition of the sulfur source.<sup>20,66–</sup>

68

### 1.6.2 Bacterial Ferredoxin Maquettes

Gibney *et al.* investigated the use of peptide cluster maquettes to better understand how each amino acid composing the peptide sequence contributes to the assembly of [4Fe-4S] clusters.<sup>66</sup> The maquettes they studied were inspired by ferredoxins containing multinuclear clusters and they implemented both natural proteins and inorganic model complexes to develop the cluster maquettes. In their studies they noted that replacing one or two Cys residues out of four led to lower yields of cluster formation. The cluster could still be formed with  $\beta$ ME, hydroxide ion, or another peptide as the third and fourth ligand; in a similar fashion to what is observed for radical SAM

maquettes. Replacement of three Cys residues or use of L-Cys alone led to almost complete loss of cluster reconstitution. The intervening amino acids were also found to play an important role. A loss of over 10% in yield was registered when all the non-Cys amino acids were exchanged with Gly residues. Finally, the electrochemical studies they carried out showed that the cluster resides in an environment with a high degree of solvent exposure, a common trait among [4Fe-4S] clusters with low potential.

In a different study, Mulholland *et al.*, investigated the effect of replacing up to three of the four Cys residues in a sixteen amino acids ferredoxin maquette.<sup>68</sup> Starting from the peptide sequence Ac-KL**C**<sub>1</sub>EGGC**C**<sub>2</sub>IA**C**<sub>3</sub>GAC**C**<sub>4</sub>GGW-NH<sub>2</sub>; all the four Cys residues were sequentially replaced by an Ala residue. It was observed that replacement of either **C**<sub>1</sub>, **C**<sub>2</sub>, or **C**<sub>3</sub> led to an overall loss of 44 to 48%. Replacement of **C**<sub>4</sub> led instead to a 90% loss in cluster reconstitution, suggesting that the Cys residue proximal to the C-terminus plays a more crucial role in the cluster formation and stability. Interestingly, replacement of **C**<sub>3</sub> with either an Ala or a Leu residue led to an identical loss of 48% in cluster reconstitution, meaning that the presence of a hydrophobic residue in this position does not greatly affect the cluster formation. Replacement of two Cys residues with Ala residues in the **C**<sub>2</sub>-**C**<sub>3</sub> and **C**<sub>1</sub>-**C**<sub>4</sub> positions led to a decrease of respectively 71% and over 95 % in cluster reconstitution, once again demonstrating how the fourth position is pivotal for the cluster formation and stability. While exogenous solvent was thought to coordinate the cluster when only one Cys residue was replaced, gel permeation chromatography demonstrated that two peptides were required to fulfil the ligand requirements of the two-Cys containing peptide maquettes. This is an interesting finding as it proves that despite βME being in larger concentration with respect to the peptide, the cluster only coordinates two molecules of the latter rather than two molecules of the former.

Further to this, the importance of non-liganding amino acids for [4Fe-4S] cluster reconstitution was investigated by Mulholland *et al.*<sup>69</sup> The original, 16-mer peptide sequence they developed (Ac-KL**C**<sub>1</sub>EGGC**C**<sub>2</sub>IA**C**<sub>3</sub>GAC**C**<sub>4</sub>GGW-NH<sub>2</sub>) was shortened to a 7-mer peptide sequence (Ac-**C**<sub>2</sub>IA**C**<sub>3</sub>GAC**C**<sub>4</sub>-NH<sub>2</sub>) by removing all the flanking amino acids and one Cys residue. This peptide sequence successfully incorporated a [4Fe-4S] cluster, despite affording a lower yield (29% compared to the 16-mer peptide).

Replacement of all the non-Cys amino acids with either Gly or Ala resulted in low cluster-reconstitution, with respectively 6 and 7% yield compared to the original sequence. Starting from the poly- Gly sequence Ac-C<sub>2</sub>GGC<sub>3</sub>GGC<sub>4</sub>-NH<sub>2</sub>; they observed that replacement of one of the Gly residues with an Ile residue either in the second, third, or in the sixth position restored the ability of cluster coordination to respectively 28, 15, and 27%. On the other hand, replacement of the Gly residue with an Ile residue in the fifth position resulted in complete loss of cluster affinity. This position proved to be a critical site for cluster-binding and is generally occupied by a Gly residue in natural ferredoxins. Interestingly, the introduction of an Ile residue in the eighth position of a poly-Gly 16-mer sequence led to an increase in cluster affinity from 11% (for Ac-GGC<sub>1</sub>GGGC<sub>2</sub>GGG<sub>3</sub>GGC<sub>4</sub>GGW-NH<sub>2</sub>) to 61% (for Ac-GGC<sub>1</sub>GGGC<sub>2</sub>I<sub>8</sub>GC<sub>3</sub>GGC<sub>4</sub>GGW-NH<sub>2</sub>).

## 1.7 SPECTROSCOPICAL TECHNIQUES FOR THE CHARACTERISATION OF [4Fe-4S] CLUSTERS

### 1.7.1 Ultraviolet-Visible Spectroscopy

Ultraviolet-Visible (UV-Vis) spectroscopy serves as a simple and rapid technique in Fe-S species investigation. For proteins and peptides containing Fe-S centres the characteristic absorption occurs in the range of 380-420 nm, due to the ligand-to-metal charge transfer (LMCT).<sup>70</sup> In the context of Fe-S-cluster maquette, the ligand is the thiol group of the Cys residues on the peptide, and the metal is the Fe<sup>2+</sup>/ Fe<sup>3+</sup> ion constituting the cluster, as displayed in Figure 3.

The LMCT complexes results from electrons being transferred from molecular orbitals (MO) with a ligand-like character, to MO with metal-like character. LMCT complexes are generally opposed to metal-to-ligand charge transfer (MLCT) complexes, where the opposite is true, and therefore electrons move from metal-like MO to ligand-like MO.<sup>71</sup> In Fe-S-cluster maquettes the ligand, the thiolate moiety, is a good nucleophile and therefore has a high energy lone pair that can easily transfer its electrons across the metal orbitals.<sup>72</sup> When a LMCT transition occurs, electrons are excited from the

MOs of the  $\pi$ -donor ligand to the MOs of the metal, which are therefore mostly metal in character.

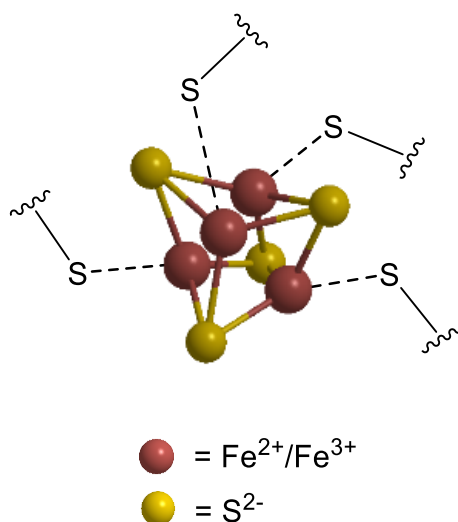


Figure 3 – [4Fe-4S] Cluster coordinated by Cys residues.

Unfortunately, UV-Vis spectroscopy is not certain proof of [4Fe-4S]-cluster presence. The interaction that results in an absorbance event can be produced by any molecule acting as a ligand for the metal centre. When ferric (Fe<sup>3+</sup>) ions are mixed with thiol-containing molecules in aqueous solution, different species can possibly interfere with the formation of Fe-S cluster coordinated by the Cys-rich motif of the synthetic peptide. For example,  $\beta$ ME has been proved to spontaneously form [2Fe-2S]-clusters when in sufficiently high concentration.<sup>5</sup>  $\beta$ ME is an additive routinely used in Fe-S cluster incorporation procedures to ensure the Cys residues of the synthetic peptides are kept in the reduced form. [2Fe-2S]-clusters have a similar absorption curve to the one of [4Fe-4S]-clusters, causing uncertainty in addressing which species is presented in solution. Luckily, [2Fe-2S] clusters are generally less abundant than the corresponding [4Fe-4S]-clusters.  $\beta$ me could also form rubredoxin-like complexes that can compete with the formation of [4Fe-4S]-clusters. These systems are small, iron-containing molecules that could be generated by the thiol of  $\beta$ ME coordinating Fe<sup>3+</sup> ions in solution.<sup>73</sup> These species are similar to the desired Fe-S cluster maquette, but lack the inorganic sulphide, as depicted in Figure 4.



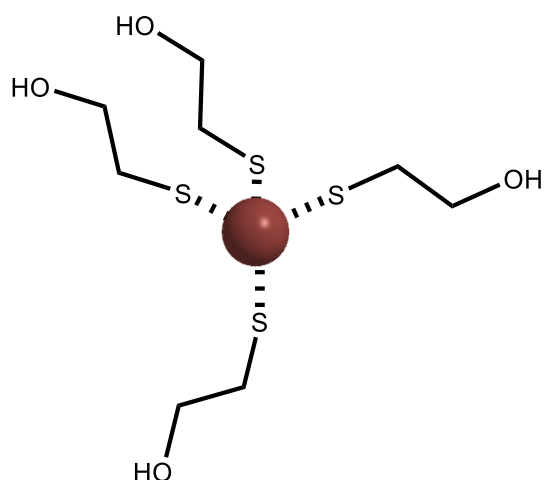


Figure 4 – Species potentially formed in presence of  $Fe^{3+}$  and  $\beta ME$ .

### 1.7.2 Electron Paramagnetic Resonance Spectroscopy

The characterisation of [4Fe-4S]-cluster maquettes is normally achieved by combining UV-Vis spectroscopy with EPR spectroscopy. The latter represents a powerful tool for in-depth characterisation of Fe-S cluster maquettes but requires more effort for sample preparation and extensive experience for the spectra interpretation. EPR spectroscopy requires the sample to have unpaired electrons.<sup>74</sup> Both iron and sulfur have wide redox chemistry that ranges from Fe (-II) to Fe (VI), and from S (-II) to S (VI). However, due to the limitations imposed by water as the solvent and a synthetic peptide as the ligand, iron has a limited oxidation state range that spans from Fe(I) to Fe(IV). When reconstituted, [4Fe-4S]-clusters are in their oxidised  $[4Fe-4S]^{2+}$  state, which is diamagnetic and thus silent for EPR spectroscopy.<sup>75</sup> The diamagnetic state is caused by the presence of two ferrous (2+) and two ferric (3+) atoms in the cluster. These two sets of pairs are magnetically coupled and therefore have a total spin  $S = 0$  due to the paired electrons. The Fe-S centre needs to be reduced to the paramagnetic  $[4Fe-4S]^{1+}$  ( $S = \frac{1}{2}$ ) state to be observed by EPR spectroscopy. This condition is generally achieved by reacting the oxidised  $[4Fe-4S]^{2+}$  cluster with a reducing agent, such as sodium dithionite (DT,  $Na_2S_2O_4$ ) and then allowing an incubation period before flash-freezing the samples for analysis. In this paramagnetic state, the cluster is theoretically composed of three ferrous (2+) atoms and one ferric (3+) atom. However, the unpaired electron is formally delocalised over two atoms of iron, giving rise to an equal-valence pair ( $2Fe^{2+}$ ) and a mixed-valence pair ( $2Fe^{2.5+}$ ).<sup>76</sup>

An important parameter in EPR spectroscopy is the  $g$ -value. This value determines the resonant field position, which corresponds to the magnetic field at which the absorption peak is observed for the species analysed. The  $g$ -value is also influenced by the micro-environment surrounding the unpaired electron and is therefore unique for a given species (i.e., an [4Fe-4S]-cluster coordinated by a synthetic peptide). A free electron in the vacuum has a  $g$ -value  $g_e = 2.0023$  and different species with an unpaired electron can differ slightly or greatly from this value based on the extent of the interaction with their environment. Normally, transition metals (i.e., iron for Fe-S clusters) have a discrete spin-orbit coupling, which means that the corresponding  $g$ -value can substantially differ from that of a free electron. The spin-orbit coupling is defined as the motion of a particle's spin and the motion of its orbit.<sup>77</sup> For this reason, the  $g$ -value for an unpaired electron in a d complex will depend on how much the orbital angular momentum contributes to the overall magnetic moment of the molecule, and whether this is parallel or opposed to the spin angular momentum.<sup>78</sup> Normally, the  $g$ -value for a species is determined either by accurate measurement of the magnetic field and frequency of the microwave power, or by referencing an unknown sample to one with a known  $g$ -value. The  $g$ -value also provides key information to define the symmetry of the absorption peak. The majority of [4Fe-4S]-cluster complexes presents an axial peak where  $g_x = g_y < g_z$ . In this situation, one axis ( $g_z$ ) is parallel to the magnetic field, whereas the other two ( $g_x$  and  $g_y$ ) are perpendicular.

## 1.8 AIMS AND OBJECTIVES

The aim of this thesis is to expand the comprehension of Cys-containing peptides both in biologically relevant processes and in synthetic chemistry. The chapters of this thesis describe the importance of Cys as a molecular handle that from prebiotic life played a crucial role in coordinating Fe-S clusters in proteins; and how Cys is used in cells to cope with oxidative damage. The exploration of a photosynthetic methodology of Cys residues in peptides will also be explored to further emphasise how this amino acid can be used to install various modifications.

The aim of Chapters 2, 3, and 4 of this thesis is to provide more details concerning the use of radical SAM and bacterial ferredoxin peptide maquettes in electron transfer

processes. Understanding and controlling oxygen-sensitive peptide maquettes could be of great importance in exploring new ways to access challenging chemical functionalisation reactions. Pharmaceuticals are often small molecules the synthesis of which could involve trivial steps; or computational chemistry could point towards targets that medicinal chemists might find difficult to synthesise. The ability to overcome these and other challenges could lay behind using nature to catalyse chemical transformations. The family of radical SAM enzymes presents a myriad of members capable of catalysing several organic reactions often impossible to replicate at the current stages of synthetic chemistry. Bacterial ferredoxin proteins are well known for catalysing the transfer of electrons in several steps critical for bacterial functions. The application of these models to biotechnology could have an enormous impact on the future of drug discovery. The high reactivity of the radical generated by radical SAM enzymes and the high sensitivity to oxygen characteristic of both classes of enzymes has always presented challenges in the development of these systems in industrial processes. Our aim is to better understand what the limits of these classes of enzymes are by synthesising and interrogating small peptides that mimic their functions. An interesting but so far not deeply investigated idea is to implement short peptides [4Fe-4S]-cluster maquettes in catalysis. These catalysts could be used to perform organic transformations that would otherwise be extremely challenging to achieve. Enzymes are known for being extremely selective and capable of executing a modification on a substrate among several similar compounds. Peptides can be quickly and easily synthesised and provide an easier architecture to modify with respect to proteins. They can also be readily functionalised by changing their flanking (Y) or intervening (X) amino acids in a sequence with this  $YCX_3CX_2CY$  (radical SAM) or this  $YCX_3CX_2CX_2CY$  (bacterial ferredoxin) pattern. In Chapters 2, 3, and 4 we will address questions like: how sensitive to oxygen are these systems? What is the redox potential and how tuneable is it? Can key feature of these peptide maquettes be changed without affecting the stability or to improve their characteristic? How do these systems perform when implemented into an electron transport process?

Chapter 5 of this thesis describes the modification of Cys in peptides using a photosynthetic method. Post-synthetic modifications are crucial for selectively installing modifications in peptides and proteins, and Cys is often targeted with this respect due to its unique reactivity. The importance of this chalcogen containing amino

in proteins is well described in the literature, and the post-synthetic modification reported in this thesis adds a new transformation that can be achieved using Cys as the starting point.

Cys residues under the form of disulphide bonds are crucial in cells for oxygen sensing and this process is further investigated in Chapter 6. The importance of Cys in proteins spans over various aspects, but in this thesis we focused on exploring its use as a anti oxidative tool in small peptides. Due to its range of oxidation states, the sulfur atom of Cys is adopted in cells to store oxygen atoms in radicals that could otherwise cause damage and affect the correct functioning of the cell. The focus in Chapter 6 is on understanding how proximal amino acids to the Cys residues can affect the stability of the disulphide bond when this is exposed to oxidative conditions.

## 2 SYNTHESIS AND CHARACTERISATION OF [4Fe-4S] CLUSTER MAQUETTES

### 2.1 AIMS AND OBJECTIVES

In this chapter we describe the synthesis and characterisation of a small library of [4Fe-4S]-cluster maquettes. The intervening (X) and flanking (Y) amino acid of the canonical key Cys-rich motif  $Y\mathbf{CX}_n\mathbf{CX}_n\mathbf{CY}$  will be changed to investigate how these modifications affect the overall ability of the synthetic peptide to incorporate the cluster. The synthetic peptides implemented for the cluster reconstitution span across two different family of enzymes: radical SAM and bacterial ferredoxins. The synthesis of these systems was performed in strict anaerobic conditions to preserve the cluster from oxidation, and their characterisation involved the use of both UV-Vis and EPR spectroscopy. The role of  $\beta$ ME as a routinely used reducing agent in cluster-maquette reconstitution procedure has been extensively investigated. The samples have been treated with nicotinamide adenine dinucleotide (NADH) followed by photoactivation to explore if a different source of electrons could be successful in yielding the reduced [4Fe-4S]<sup>1+</sup> cluster when the reconstitution procedure is performed in aerobic conditions. Finally, the replacement of sulphide ( $S^{2-}$ ) with selenide ( $Se^{2-}$ ) has been studied to further explore the possible modifications achievable with these cluster maquettes. Selenium has been tested as an alternative to sulfur in the construction of the cluster, but also in the form of Sec for the synthesis of a selenopeptide.

### 2.2 RESULTS AND DISCUSSION

#### 2.2.1 Initial Attempts on Radical SAM-type Maquettes

The peptides described in this thesis were all synthesised using Solid Phase Peptide Synthesis (SPPS) and the fluorenylmethoxycarbonyl (Fmoc) strategy. This methodology allows for the peptide sequence to be assembled one amino acid at a time, with the amine moiety being protected with an Fmoc protecting group. The reaction by-products can be easily removed via washes of the solid support. More

detailed procedures can be found in the Experimental section, at 8.1 GENERAL METHODS.

The synthetic peptide Ac-**CGGGCGGCY**-NH<sub>2</sub> (maquette1, **MQ1**, Figure 5) was chosen as the first candidate for [4Fe-4S]-cluster incorporation as it proved successful in the work published in 2019 by Galambas *et al.*<sup>20</sup>

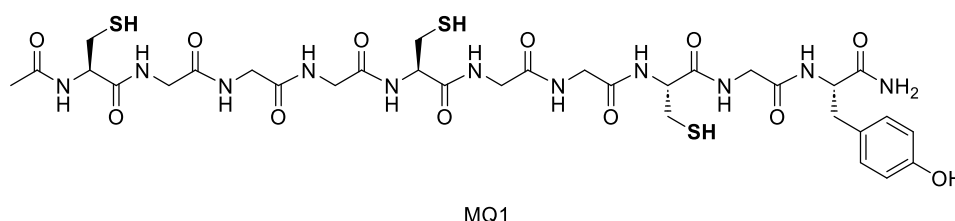


Figure 5 – Peptide sequence of **MQ1**.

This synthetic peptide represents the simplest motif with three Cys residues and no flanking amino acids on either side of the sequence. With only three Cys residues this peptide mimics the radical SAM motif **CX<sub>3</sub>CX<sub>2</sub>C** and requires βME to act as a fourth ligand in the coordination of the cluster, as showed in Figure 6.

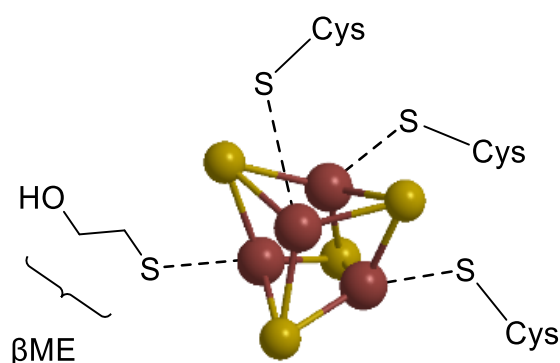


Figure 6 – [4Fe-4S] Cluster coordinated by three Cys residues and βME.

The first technique we implemented to investigate successful formation of the cluster maquette was UV-Vis and the results agreed with the reported literature, as showed in Figure 7.

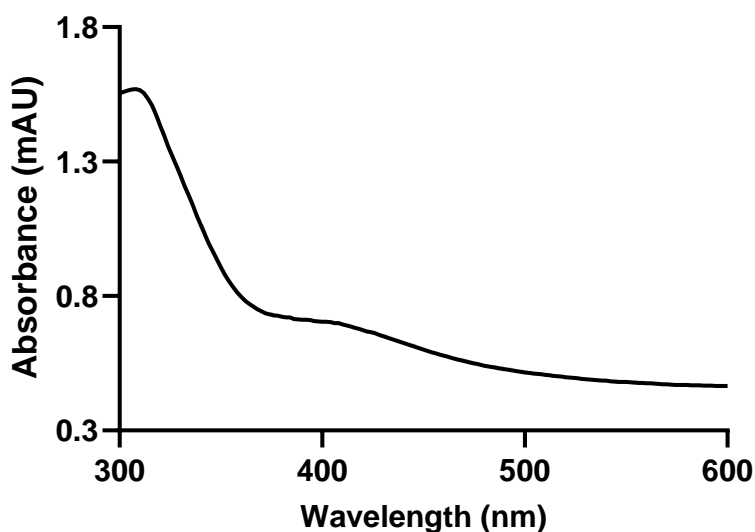


Figure 7 – UV-Vis trace for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ1**. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 100  $\mu$ M. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

The first peak registered at 300 nm is typical of peptides containing Tyr, Trp or Phe, which can absorb in the aromatic region due to the chromophores presented in their side chains.<sup>79</sup> The characteristic absorbance peak for Fe-S clusters was observed at approximately 400-420 nm, suggesting the successful formation of the [4Fe-4S] cluster maquette.

We then proceeded to analyse the sample by EPR spectroscopy. The analysis was recorded on an anaerobically reconstituted maquette reduced with DT to ensure the cluster would be in the EPR-visible reduced state [4Fe-4S]<sup>1+</sup>. The reconstitution procedure under anaerobic conditions is described in 8.2 *GENERAL PROTOCOLS*. Briefly, degassed HEPES 50 mM and KCl 10 mM buffer at pH = 8.0 was used to dissolve the peptide. To this solution were progressively added the reducing agent and fourth ligand  $\beta$ ME, and the two solutions of FeCl<sub>3</sub> and Na<sub>2</sub>S for the cluster formation. Finally, the sample was reduced using DT, transferred into an EPR tube and flash frozen for the analysis. Before recording the spectrum for the reduced cluster-maquette we performed the analysis of the buffer and buffer with DT solutions. All EPR samples were supplemented with glycerol to a final concentration of 10% v/v.

EPR analysis was carried out at the EPSRC National Research Facility (NRF) for EPR Spectroscopy in Manchester under the supervision of Dr. Muralidharan Shanmugam

and Prof. David Collison. Dr. Muralidharan Shanmugam ran all the samples and analysed the results for us.

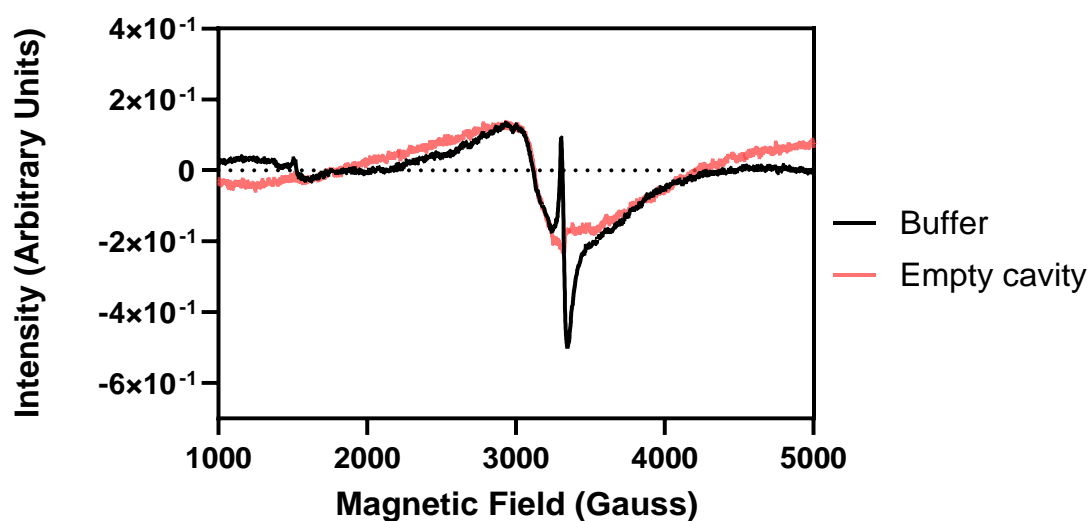


Figure 8 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed buffer (black trace) and empty cavity (red trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

Figure 8 shows the trace for the buffer solution (black) compared to the empty cavity (red) of the instrument. The sharp peak observed at approximately 3300 G was possibly due to the presence of a radical species. The superoxide radical is sometimes observed in buffer analysed by EPR and can be produced by one-electron reduction of oxygen.<sup>80</sup> Despite the thorough procedure followed for degassing the buffer (see 8.2 GENERAL PROTOCOLS), it is possible that residual oxygen was still available in the sample. Unfortunately, the presence of a radical in the buffer could have negative effects on the analysis of the cluster maquette sample. The second analysis we performed was on a sample containing a solution of buffer and DT, as presented in Figure 9.



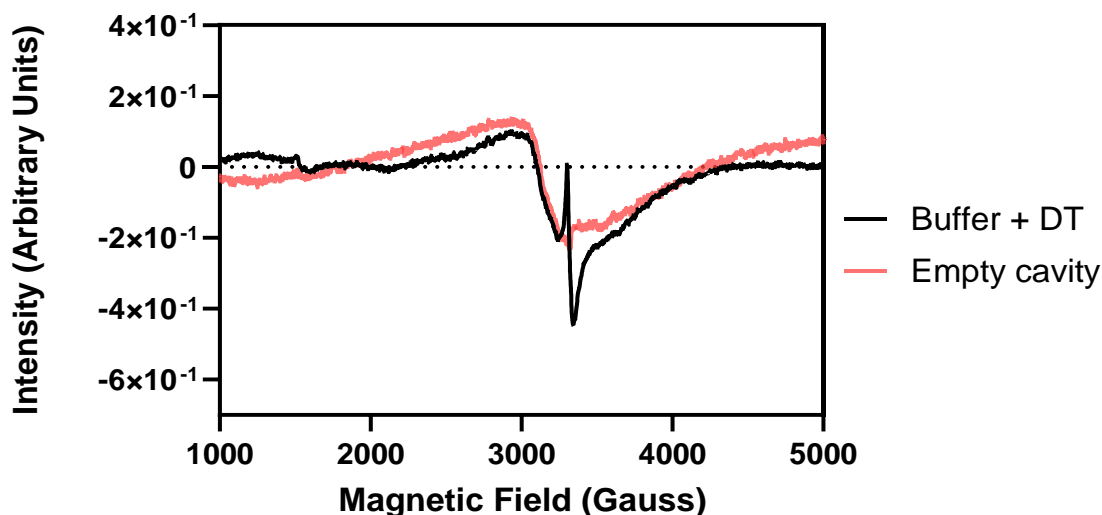


Figure 9 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed buffer + DT (black trace) and empty cavity (red trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

The trace for the solution DT in buffer was similar to that of pure buffer, and the radical signal was still present in this sample. We then proceeded by analysing the cluster maquette samples anaerobically reconstituted both before and after reduction with DT. The results are displayed in Figure 10 and are compared to the signal for the empty cavity.

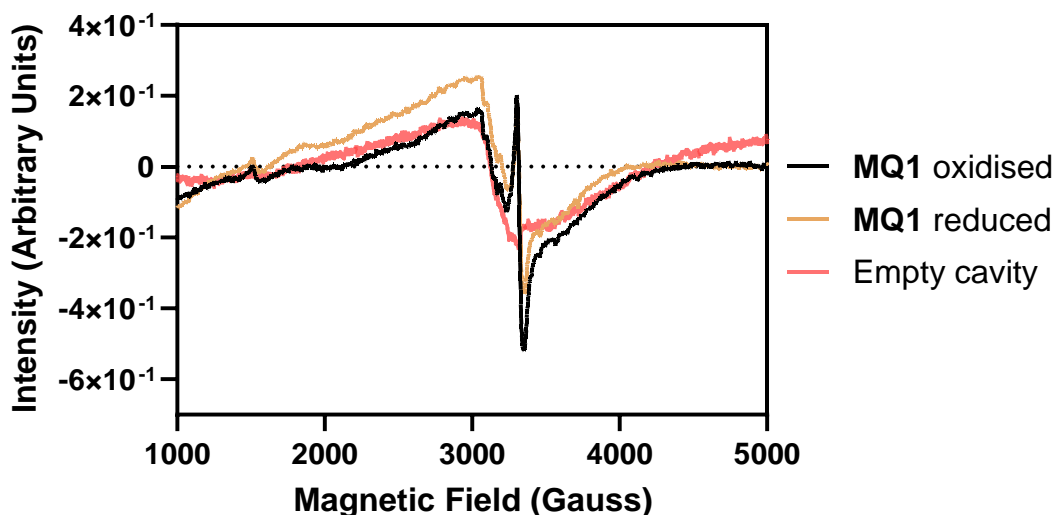


Figure 10 – X-band cw-EPR spectra at 20 K a frozen solution of degassed (anaerobically reconstituted in presence of  $\beta$ ME) MQ1 oxidised (black trace), dithionite-reduced (gold trace), and empty cavity (red trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

The spectrum registered for the reduced sample was comparable to the one recorded for the oxidised sample, meaning that DT did not produce any effect on the redox centre of the Fe-S cluster presented in the sample. Alternatively, the reason why the two spectra were comparable is because no cluster was present in the sample that could be reduced and observed by EPR spectroscopy. These results were not in agreement with previous literature, where Galambas *et al.* were able to successfully reconstitute an [4Fe-4S]-cluster using an identical synthetic peptide.<sup>20</sup> Despite several attempts at different concentrations of peptide, FeCl<sub>3</sub> and Na<sub>2</sub>S, and βME, all our efforts could not produce positive results. The reasons behind the negative results could be explained by different salts of the synthetic peptide after purification (i.e. formate salt or trifluoroacetate salt) that could have affected the reconstitution procedure. The synthesis of these cluster maquettes requires extensive care and experience to ensure every step is followed thoroughly, and perhaps this could have led to unsuccessful results in the case of **MQ1**.

Due to the inconclusive results obtained by EPR spectroscopy we questioned whether the cluster maquette had been successfully synthesised in the first place. UV-Vis spectroscopy could have possibly given a false positive result due to the interaction between βME and excess Fe<sup>3+</sup> in solution. This complex did not show the characteristic EPR signal due to the absence of a [4Fe4S] cluster in the sample. We therefore decided to investigate whether the presence of βME could indeed lead to the formation of an iron complex and compete with the cluster maquette in the determination of its absorbance. Size Exclusion Chromatography (SEC) is a purification technique that implements gel filtration to separate small to medium molecules based on their size. In principle, molecules larger than the largest pores in the column matrix are excluded from the matrix and are therefore eluted first. On the contrary, molecules smaller than the largest pore in the matrix penetrate the pores to varying extent and are therefore eluted after the larger molecules. The columns with the lowest exclusion limit that could be found on the market at the time of these experiments were the PD MidiTrap G-10™, produced and commercialised by Cytiva. The exclusion limit of these columns is 700 Da (or 700 g/mol), which means that molecules with a lower size than 700 Da would spend more time in the column and be eluted more slowly than molecules with a size greater than 700 Da. However, the mass of **MQ1** is 816.23 Da, which is not too far from the exclusion limit and could therefore

lead to impurities presented in the sample. Moreover, the nature of the complexes that formed during the synthesis could not be determined and it was therefore impossible to propose a mass for these impurities.

### 2.2.2 $\beta$ ME and Single Amino Acid Screening

The unsatisfactory results obtained with EPR spectroscopy regarding the maquette incorporated using peptide Ac-**CGGGCGGCY**-NH<sub>2</sub> (**MQ1**), led us to reconsider our reconstitution procedure, and various control experiments were run and analysed by UV-Vis spectroscopy. The first point of interest addressed in these experiments was the reducing agent  $\beta$ ME. Due to the higher concentration with respect to the peptide used in the reaction mixture, and its ability to coordinate Fe<sup>3+</sup> ions due to the thiol group, this agent was thought to be interfering with the cluster formation to the point that all the iron ions were being consumed before reacting with the Cys residues of the peptide. As a result, the cluster could not be formed and coordinated by the Cys residues on the peptide sequence. A screening with the synthetic peptide Ac-**CGGGCGGCY**-NH<sub>2</sub> (**MQ1**) and different concentration of  $\beta$ ME was performed with the aim of investigating how different concentration of this reagent could affect the coordination between the Cys residues and the iron ions. Five experiments were performed including four different concentrations of  $\beta$ ME and a negative control without the reducing agent. The UV-Vis traces are showed in Figure 11.

Four concentrations of  $\beta$ ME were chosen: 0.18 mM, which is equimolar to the peptide concentration, 1 mM, 100 mM, and 280 mM, which is equivalent to 2% v/v of the final solution and was adopted in previous literature for the cluster reconstitution procedure.<sup>20</sup>

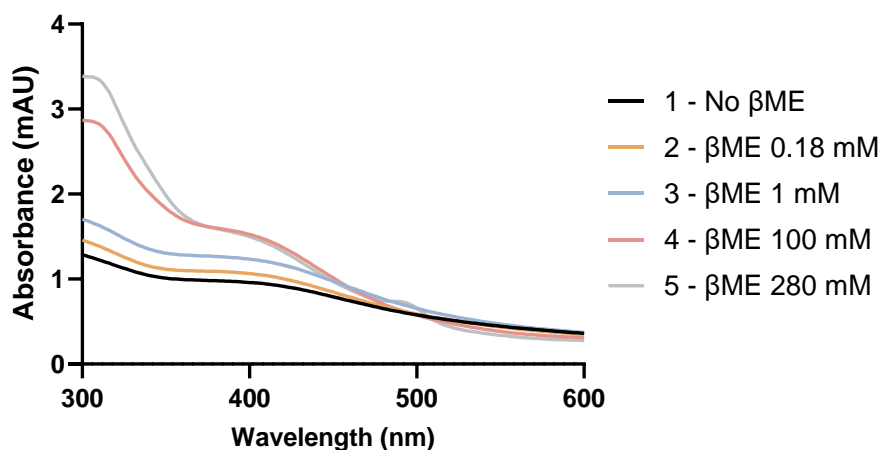


Figure 11 – UV-Vis traces for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ1**. Screening of  $\beta$ ME concentration at: 0 (**1**), 0.18 mM (**2**), 1 mM (**3**), 100 mM (**4**), and 280 mM (**5**). Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 0.18 mM. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

The results displayed in Figure 11 present a clear peak at approximately 400 nm. Moreover, the intensity of the peak increases as the amount of  $\beta$ ME in the sample is increased. This outcome supported the hypothesis by which  $\beta$ ME participated in a competitive reaction with the ferric ions and formed a different rubredoxin-type complex. The concentration of the latter was directly correlated to the concentration of  $\beta$ ME. Entries 1, 2, and 3 showed a green solution accompanied by a dark precipitate, suggesting that  $\beta$ ME was not concentrated enough to coordinate all the  $\text{Fe}^{3+}$ , which therefore precipitated with the  $\text{HS}^-$  ions in solution to form  $\text{FeS(s)}$  insoluble particles. The solution of entry 4 had small quantities of precipitate, whereas 5 did not present precipitate at all, suggesting that the amount of thiol in solution was high enough to coordinate all the ferric ions. The peak observed in absence of  $\beta$ ME, despite low in intensity, could possibly arise from coordination of a Fe-S cluster by the peptide sequence. Unfortunately, no EPR spectroscopy was performed on this sample to confirm this hypothesis. Table 1 provides more details on the concentrations of the various reagents presented in solution.

The concentrations of  $\text{FeCl}_3$  and  $\text{Na}_2\text{S}$  were kept consistent at 800  $\mu\text{M}$  for all the five entries, as well as the concentration of peptide at 180  $\mu\text{M}$ . The only variable in this set of experiments was the concentration of  $\beta$ ME.

To confirm that the characteristic key motif of radical SAM enzymes (**CX<sub>3</sub>CX<sub>2</sub>C**) played a role in [4Fe-4S]-cluster incorporation, a new set of experiments was subjected to UV-Vis spectroscopy analysis. These cluster incorporation experiments were run using two different synthetic peptides: Ac-GGGGGGGG-NH<sub>2</sub> (**MQ2**) and H<sub>2</sub>N-CGGGCGGC-OH (**MQ3**) (Figure 12).

Table 1 –  $\beta$ ME concentration screening. All reported values refer to the concentration of each component in the reconstitution mixture. The synthetic peptide employed in these experiments was **MQ1**. All solutions presented a green/brown transparent colour.

Entry	[Peptide] (mM)	[ $\beta$ Me] (mM)	[FeCl <sub>3</sub> ] ( $\mu$ M)	[Na <sub>2</sub> S] ( $\mu$ M)	Physical appearance
1	0.18	-	800	800	Black precipitate
2	0.18	0.18	800	800	Black precipitate
3	0.18	1	800	800	Black precipitate
4	0.18	100	800	800	Traces of black precipitate
5	0.18	280	800	800	No precipitate

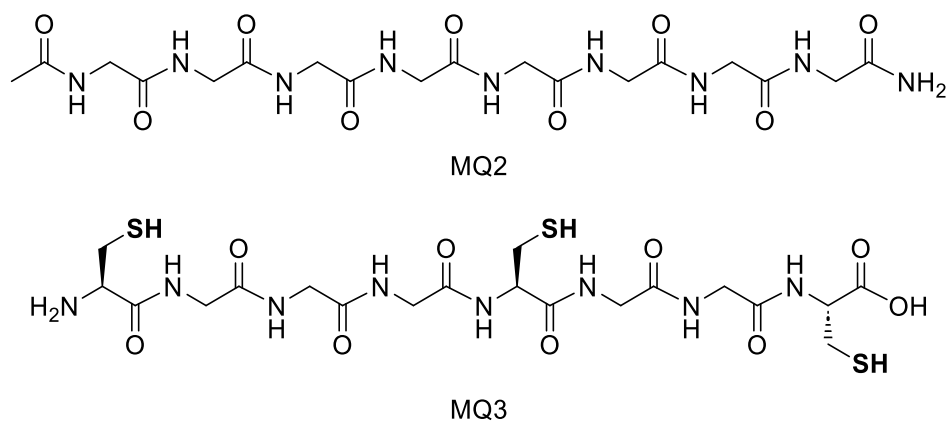


Figure 12 – Peptide sequence of **MQ2** and **MQ3**.

Both synthetic peptides were used in reconstitution procedure with and without  $\beta$ ME. **MQ2** did not contain any Cys residue, and for this reason was not expected to show any trace of cluster incorporation. This experiment was considered as a negative control to understand if any peptide sequence, even if not bearing a thiol on the side chain, could incorporate a Fe-S cluster. The second peptide sequence (**MQ3**) was meant to show traces of cluster incorporation as it contained the characteristic motif of radical SAM enzymes and differed from **MQ1** for the absence of a Tyr residue and the presence of the free carboxylic acid at the C-terminus and the free amine at the N-terminus. We argued that such a difference in electronics at both termini of the peptide sequence could have an impact on the cluster formation and stabilisation. The latter peptide was proved to successfully incorporate a [4Fe4S]-cluster in previous literature.<sup>73</sup> The UV-Vis traces for the four different reconstitution experiments are showed in Figure 13.

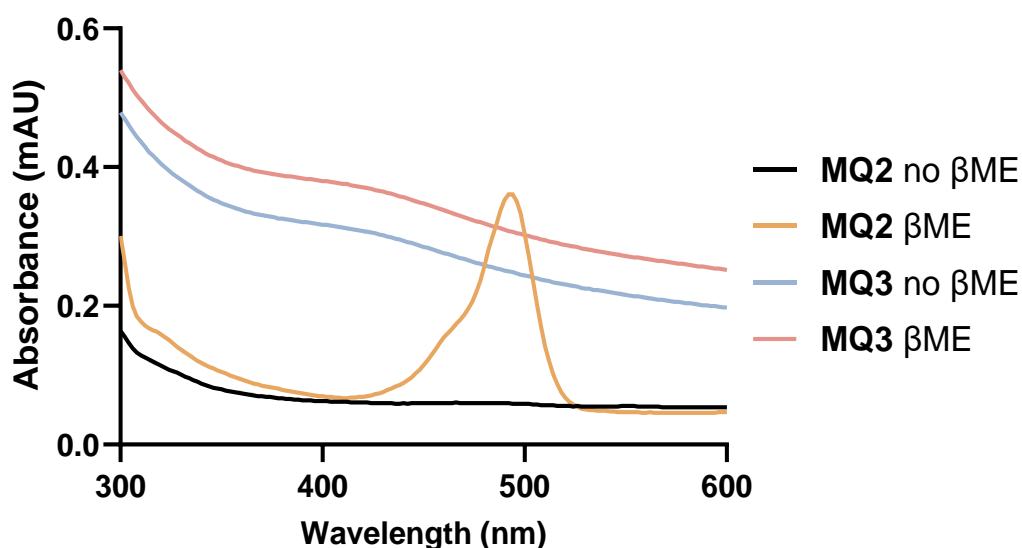


Figure 13 – UV-Vis traces for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ2** and **MQ3**. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 80  $\mu$ M. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

The experiments run using **MQ3** both with (red trace) and without (light-blue trace)  $\beta$ ME showed a curve at around 420 nm, which could be explained by the presence of an Fe-S species.<sup>4</sup> The resemblance between the two trends could be reasoned by the presence of the same species, with the concentration, and thus the intensity, being enhanced in presence of  $\beta$ ME (**MQ3**, red trace). The bend observed for the

experiments without  $\beta$ ME (**MQ3**, light-blue trace) could also arise from an interaction between the thiol group of a Cys and a single  $\text{Fe}^{3+}$  ion. This interaction would absorb at the characteristic wavelength of [4Fe-4S]-clusters, despite not being necessarily due to a Fe-S cluster species. The spectra registered for the experiment run using **MQ2** in absence of  $\beta$ ME (**MQ2**, black trace) showed a flat line, which means that no Fe-S or Fe---thiol species were produced in those specific conditions. This result agreed with the expectations as no thiol, either from Cys or  $\beta$ ME, was available to incorporate a cluster. Interestingly, the spectrum for the reaction in presence of  $\beta$ ME (**MQ2**, gold trace), showed a sharp peak at approximately 490 nm. This wavelength region is at too high frequency to be correlated to any Fe-S cluster species but has been found characteristic of rubredoxins species.<sup>81</sup> This iron-thiolate complex was only visible in presence of the poly-Gly peptide, whereas it was not observed in analogous conditions with different synthetic peptides such as **MQ1** and **MQ2**. A plausible explanation for this outcome is that the absence of Cys residues on the peptide made  $\beta$ ME the only thiol-containing species in solution that could coordinate with the  $\text{Fe}^{3+}$  ions and form a rubredoxin-type system. Since no competitive reactions could occur due to the inactivity of the synthetic peptide, this complex was the only species in solution and thus showed an intense signal due to the high concentration. Despite these experiments providing insightful information regarding the nature of the synthetic peptides required for binding the ferric ions in solution, the results obtained could not be considered conclusive. The characteristic absorbance at 420 nm arises from the interaction between the sulfur atom of either the Cys residue or  $\beta$ ME and the  $\text{Fe}^{3+}$  ion. Therefore, when experiments performed with synthetic peptides bearing a Cys residue in absence of  $\beta$ ME showed an absorbance increase in the 400-420 nm range, this could derive from the sole interaction between the Cys residue and a ferric ion. Both complexes showed in Figure 14 bear the characteristic Fe-thiol interaction that absorb in the 400-420 nm region.

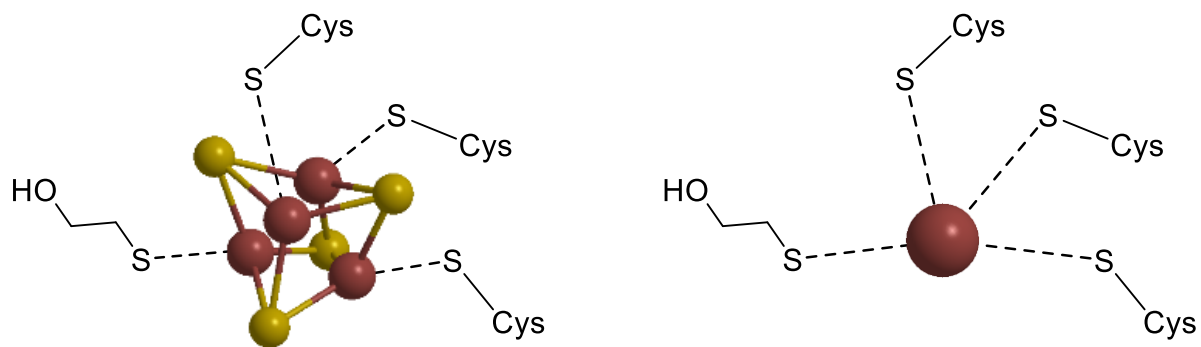


Figure 14 – Comparison between a [4Fe-4S] cluster (left) and a Fe<sup>3+</sup> ion (right) coordinated by three Cys residues and  $\beta$ ME.

For this reason, we concluded that UV-Vis spectroscopy alone could never be a suitable analytical technique for proving the presence of any Fe-S cluster species when the crude sample was not eluted through SEC columns.

### 2.2.3 Bacterial Ferredoxin-Type Maquettes

Since the short sequence of the synthetic peptide **MQ1** did not show any proof of Fe-S cluster assembly, efforts were focused on a different sequence: Ac-KL**CEGGCIACGACGGW**-NH<sub>2</sub> (**MQ4**) (Figure 15). This synthetic peptide was proved to successfully incorporate a [4Fe-4S]-cluster, as confirmed by Mulholland *et al.*<sup>68</sup> The 16-mer peptide was derived from a natural Fe-S protein that belongs to the bacterial ferredoxin family and presents the classic **CX<sub>3</sub>CX<sub>2</sub>CX<sub>2</sub>C** motif, with the fourth Cys ligand located four positions toward the amino terminus. The extra Cys was thought to be a crucial difference with respect to the 9-mer peptide of **MQ1** in incorporating and stabilising the [4Fe-4S]-cluster.

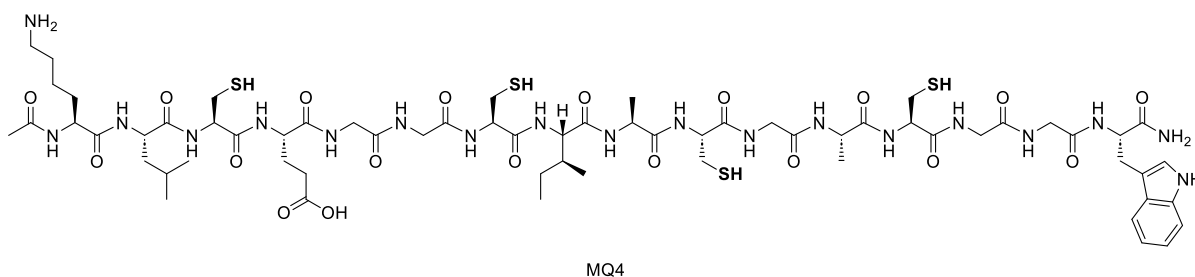


Figure 15 – Peptide sequence of **MQ4**.



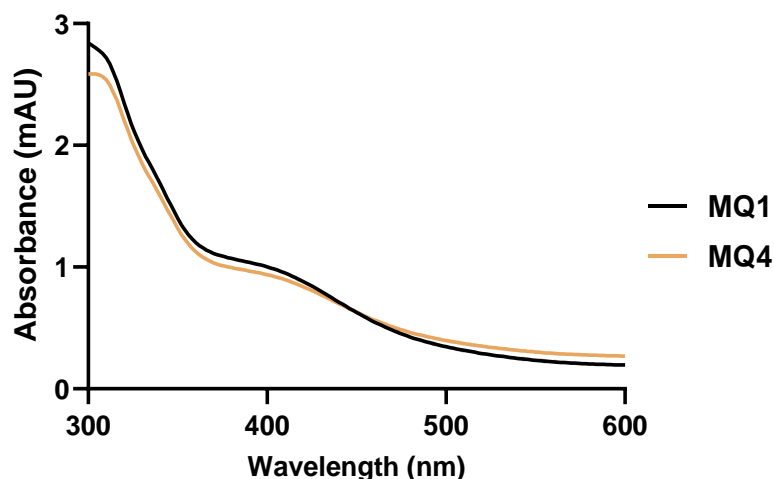


Figure 16 – UV-Vis traces for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ1** (black trace) and **MQ4** (gold trace). Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 80  $\mu$ M. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

Firstly, we proceeded by reconstituting the cluster maquette using synthetic peptide **MQ4** and analysed it by UV-Vis spectroscopy. The reconstitution procedure was slightly adapted from the previous literature and is described in 8.2 *GENERAL PROTOCOLS*. Overall, the synthetic procedure of this cluster maquette was similar to that of **MQ1** and involved the addition of FeCl<sub>3</sub> and Na<sub>2</sub>S to a solution of peptide and 10% v/v  $\beta$ ME in buffer under strict anaerobic conditions. The outcome is showed in Figure 16 and produced a similar trend to the one obtained using **MQ1**.

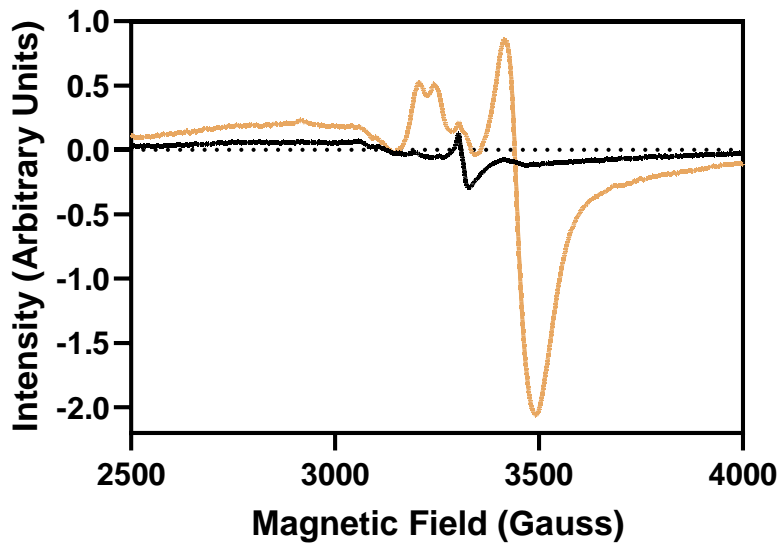
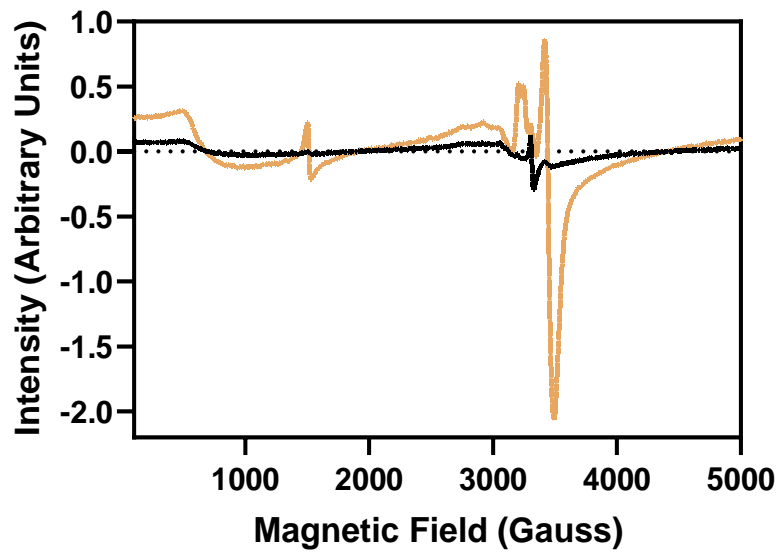
As previously discussed, this outcome could not be considered conclusive and therefore EPR spectroscopy was required to confirm the successful incorporation of a [4Fe-4S] cluster into the peptide maquette.

Figure 17 illustrates the EPR measurements of the anaerobically reconstituted **MQ4** both before (black trace) and after (gold trace) reduction with DT. These results proved the presence of a one-electron [4Fe-4S] cluster that was produced when the oxidised maquette was treated with the reducing agent DT.

This result was encouraging because they reflected the nature of [4Fe-4S]-clusters to become EPR-visible when reduced from [4Fe-4S]<sup>2+</sup> to [4Fe-4S]<sup>1+</sup> in presence of DT.<sup>82</sup> This means that the incorporation performed in presence of **MQ4** led to the formation of a [4Fe-4S]-cluster that was not visible in the oxidised (black trace) spectra because in its paramagnetic (+2) state, but when it was reduced to the diamagnetic (+1) state,

it became detectable. The broad EPR signal observed at the low magnetic field between 0 and 1000 G was thought to be arising from a high-spin ( $S = 2$ ) ferrous centre possibly presented in the  $\text{FeCl}_3$  batch used for the cluster reconstitution. The sharp signal registered at approximately 1500 G was due to high-spin ferric ion that could be present in solution as the iron source for the cluster is used in excess during the reconstitution procedure.<sup>83</sup> The spectrum for the reduced sample displayed the characteristic rhombic shape for  $[\text{4Fe-4S}]$  clusters slightly altered by the presence of low intensity peaks. These signals were thought to arise from traces of  $\text{Mn}^{2+}$  ion possibly contaminating the  $\text{FeCl}_3$  salt. To confirm this hypothesis, a sample containing  $\text{MnCl}_2$  in buffer was analysed in the same conditions, the results are displayed in Figure 18.

As showed in Figure 18, the manganese sextet appears in a similar region of magnetic field where the characteristic  $[\text{4Fe-4S}]^{1+}$  cluster is recorded. This outcome corroborated the idea by which the sample was potentially contaminated by traces of  $\text{Mn}^{2+}$ . The sextet produced by  $\text{Mn}^{2+}$  is characteristic of its  $I = 5/2$  spin number, which result in six peaks due to the Pascal's triangle.<sup>84</sup> The spectrum recorded for manganese also matched with literature examples and we therefore concluded that the signal was associated with manganese in solution.



— MQ4 oxidised    — MQ4 reduced

Figure 17 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in presence of  $\beta$ ME) MQ4 oxidised (black trace) and dithionite-reduced (gold trace). Bottom panel is expansion of the area between 2500 and 4000 G. EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

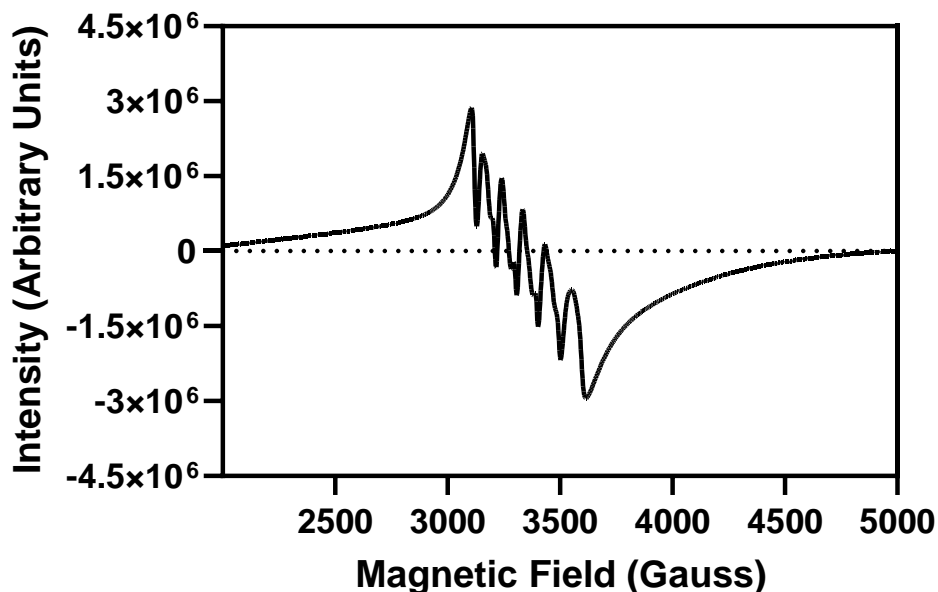


Figure 18 – X-band cw-EPR spectra at 20 K of a frozen solution of a sample containing  $\text{MnCl}_2$  in buffer. EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

Dr. Muralidharan Shanmugam produced simulations of the EPR spectrum for the sample reconstituted using **MQ4**, which are displayed in Figure 19.

The EPR spectrum was simulated (red dotted trace) by considering two,  $S = \frac{1}{2}$  species (species A, pink trace; and species B, blue trace). The spin-Hamiltonian parameters used to model the experimental (black trace) spectrum are stated in Table 2 and are in good agreement with the reported values for the reduced  $[\text{4Fe-4S}]^{1+}$  cluster coordinated with the peptide sequence Ac-KL**C**EGG**C**I**A**C**G**A**C**GGW-NH<sub>2</sub>.<sup>20</sup>

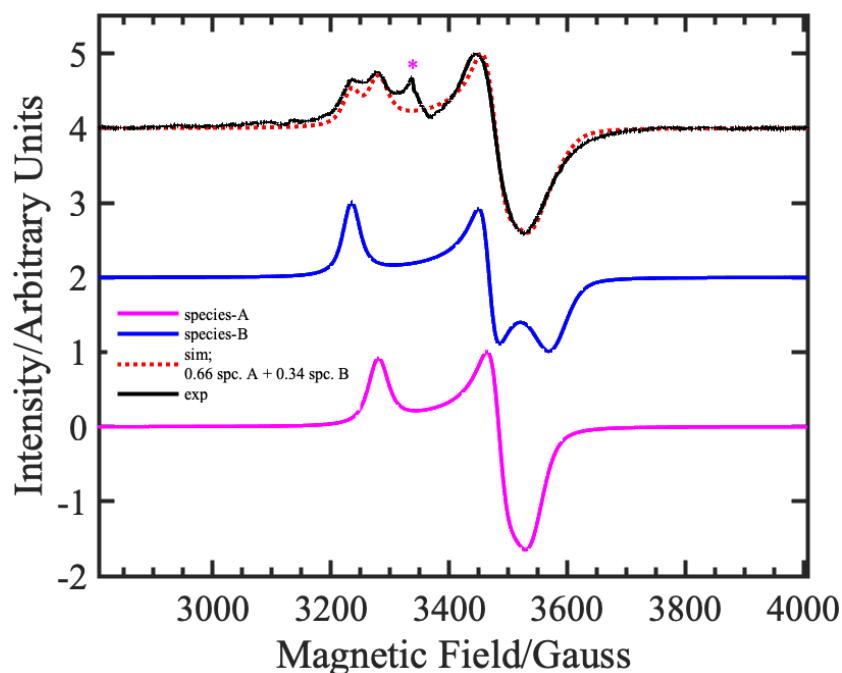


Figure 19 – X-band cw-EPR spectra of a frozen solution of anaerobically reconstituted, dithionite reduced, **MQ4** cluster measured at 20 K (black trace) and its simulations (red dotted trace with simulation parameters in Table 2). Magenta asterisk denotes the EPR signal arises from the buffer control.

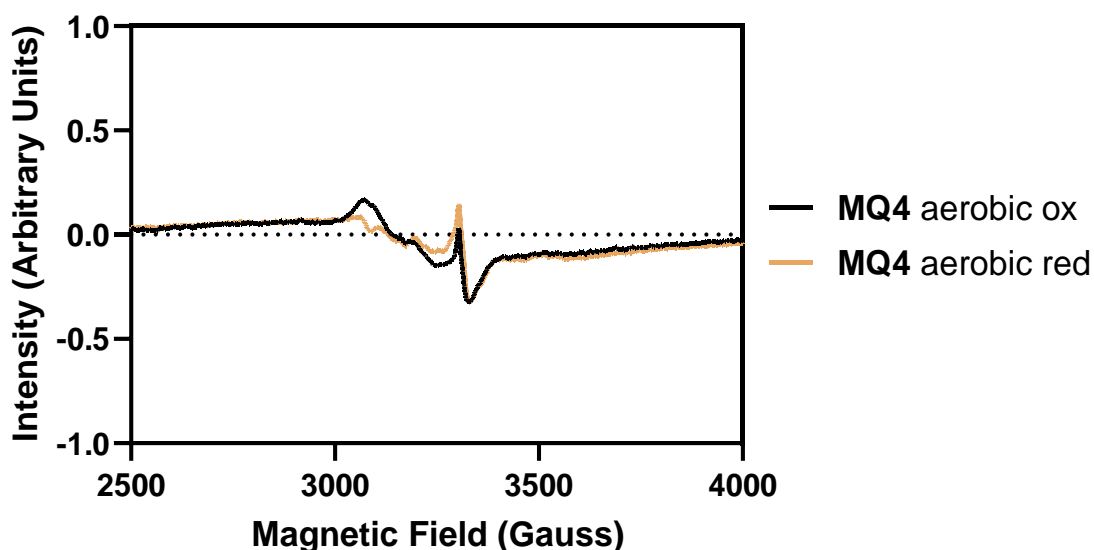
Table 2 – Simulated EPR spectroscopies properties for [4Fe-4S]<sub>1+</sub> cluster maquettes.

Entry	g <sub>1</sub>	g <sub>2</sub>	g <sub>3</sub>
Species A	1.895	1.924	2.044
Species B	1.876	1.933	2.073
Ac-KLCEGGCIACGACGGW-NH <sub>2</sub>	2.047	1.923	1.878

Due to the high sensitivity of [4Fe-4S] cluster maquettes to atmospheric oxygen, all samples were prepared and handled in strict anaerobic conditions, as detailed in 8.2 *GENERAL PROTOCOLS*. To ensure that these samples were not at any point in contact with oxygen, control experiments were prepared in aerobic conditions and analysed by EPR. The samples prepared in aerobic conditions were synthesised following the same protocols as for the anaerobic reconstitution of cluster maquettes, but the buffer and all the solutions involved were not degassed and the reaction was

performed in a standard fume hood. Figure 20 presents the EPR results obtained for **MQ4** oxidised (black trace) and reduced with DT (gold trace) when reconstituted in aerobic conditions.

These results clearly show how the oxidised and reduced samples presented the same spectrum, meaning that no cluster was observed for either of the two. This batch was prepared the same day of the anaerobically reconstituted sample, using the same peptide,  $\text{FeCl}_3$ ,  $\text{Na}_2\text{S}$ , buffer salts, and DT. As previously described in Figure 10, the sharp signal detected in both the oxidised and reduced sample can be attributed to the buffer solution. The importance of these control experiments was crucial to ensure that every time a sample was prepared and analysed by EPR spectroscopy the results could be trusted. The presence of a radical signal in these control experiments run in aerobic conditions, could potentially indicate a fault in the analysis. Even if not mentioned, EPR spectroscopy was recorded for all the experiments described in this thesis both on anaerobically and aerobically reconstituted samples.



*Figure 20 – X-band cw-EPR spectra at 20 K of a frozen solution of aerobically reconstituted in presence of  $\beta$ ME **MQ4** oxidised (black trace) and dithionite-reduced (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.*

Further experiments were performed to investigate how the EPR signals for the anaerobically reconstituted **MQ4** changed when the analysis was obtained at 10, 15, and 20 K. As displayed in Figure 21, the spectra recorded at 15 and 20 K looked

similar, with minor differences in the intensity of the signal. On the contrary, the spectrum obtained at 10 K showed a distorted shape for the rhombic signal.

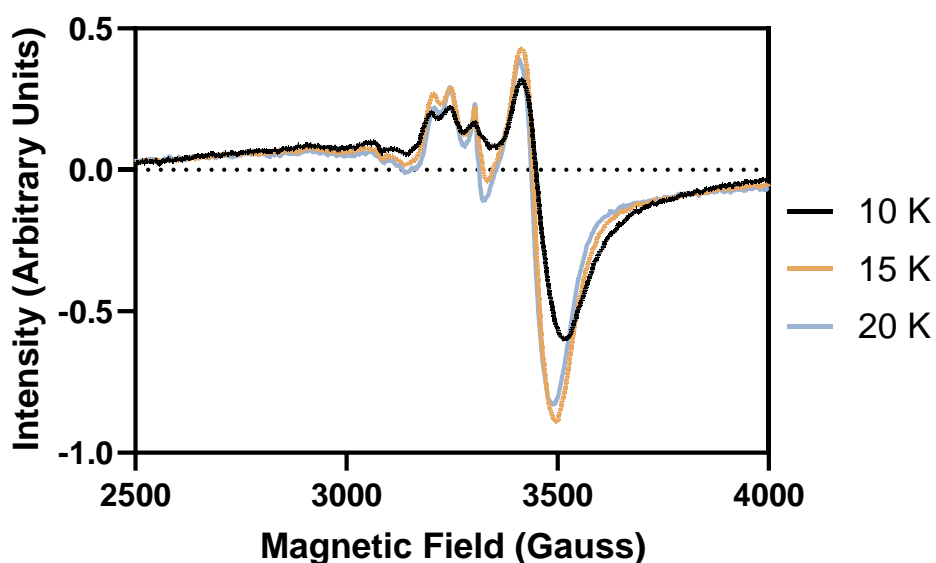


Figure 21 – Temperature dependent behaviour of anaerobically reconstituted in presence of  $\beta$ ME) dithionite-reduced **MQ4**: 10 K (black trace), 15 K (gold trace), and 20 K (light-blue trace). EPR spectra were recorded at 5 MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

In theory, the lowest possible temperature should afford the highest signal intensities due to the energy difference between the two  $S = -\frac{1}{2}$  and  $S = +\frac{1}{2}$  energy levels that are involved in the EPR transition. In order to produce a signal, the electronic population of these two levels should be as high as possible, and this is achieved at lower temperatures. However, the experimental results we obtained seemed to contradict this theory. Figure 22 shows a representation of the electron distribution in the two energy levels,  $S = -\frac{1}{2}$  and  $S = +\frac{1}{2}$ , with a slight excess in the lower level. During an EPR experiment, the electrons in the lower level absorb quanta of energy and are excited to the higher level of energy. When this happens, the population distribution is reversed, and more electrons occupy the  $S = +\frac{1}{2}$  level. However, an absorption signal is only observed in the case where the lower energy level has a higher occupancy of electrons. Due to the spin-lattice relaxation property of the sample, some electrons lose energy and move to the lower energy level, thus allowing the cycle to be repeated. When the electrons relaxation is limited due to the too low temperature, the electron populations in the two levels reaches an equilibrium more quickly and therefore the EPR signal is lost.

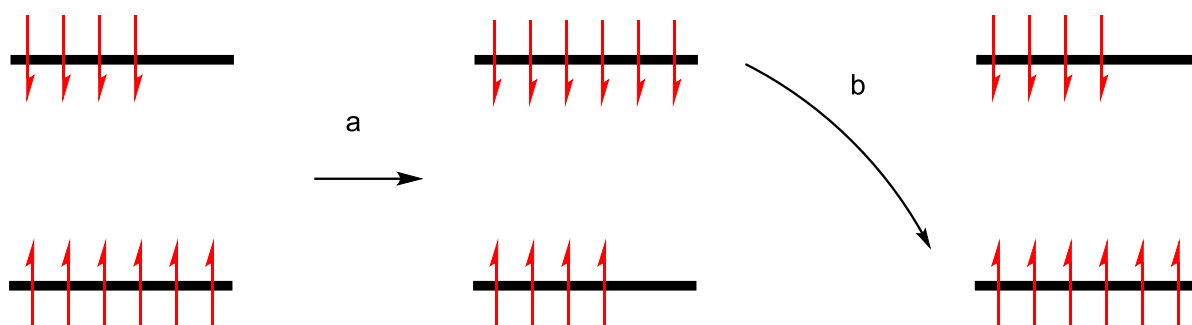


Figure 22 – Electrons distribution in the two  $S = -\frac{1}{2}$  and  $S = +\frac{1}{2}$  energy levels. a) absorption. b) relaxation.

Efforts were then focused to simplify the synthetic peptide **MQ4** used for [4Fe-4S]-cluster incorporation and develop an engineered sequence. The idea was to maintain the characteristic ferredoxin-type motif (**CX<sub>3</sub>CX<sub>2</sub>CX<sub>2</sub>C**) that was successful in the incorporation of a [4Fe-4S]-cluster. All the flanking amino acids both on the C- and N-terminus on the native sequence of **MQ4** were removed, and the internal residues were all switched to Gly residues as showed in Figure 23.

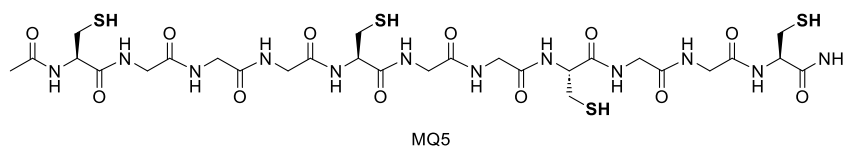


Figure 23 - Peptide sequence of **MQ5**.

The newly engineered synthetic peptide Ac-**CGGGCGGCGGC**-NH<sub>2</sub> (**MQ5**) was subjected to the same incorporation procedure that yield to [4Fe-4S]-cluster with the native sequence **MQ4**. UV-Vis traces of the [4Fe-4S]-cluster maquette incorporated using both synthetic peptides are displayed in Figure 24.

The maquette synthesised using **MQ5** produced similar results to those obtained with **MQ4**. The trends for oxidised maquettes were identical and characteristic of a ligand-to-metal interaction with a curve at around 400-420 nm. Even though reduced **MQ4** showed a flatten trend, reduced **MQ5** displayed a curve in proximity of the Fe-S species absorption region. This absorption could be explained by a portion of the cluster population not being fully reduced by DT.



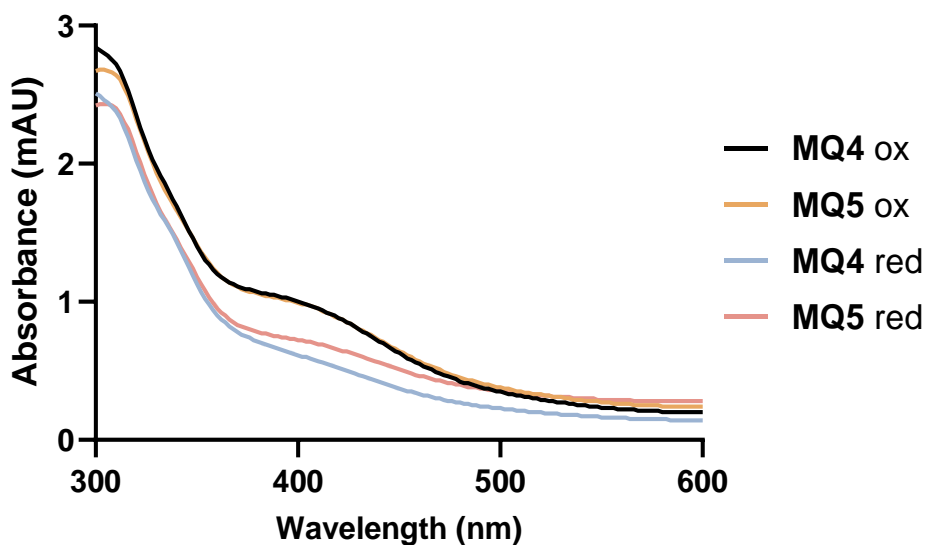


Figure 24 – UV-Vis trace for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ4** and **MQ5**. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 80  $\mu$ M. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

The EPR spectra of the anaerobically reconstituted **MQ5** before and after reduction are displayed in Figure 25. Only a tiny peak was observed in the reduced sample at 3500 G, proving the presence of a small population of [4Fe-4S]-cluster available in the sample after reduction with DT. When the reduced samples of **MQ4** and **MQ5** were compared to each other, a great difference in cluster population was observed (Figure 26).

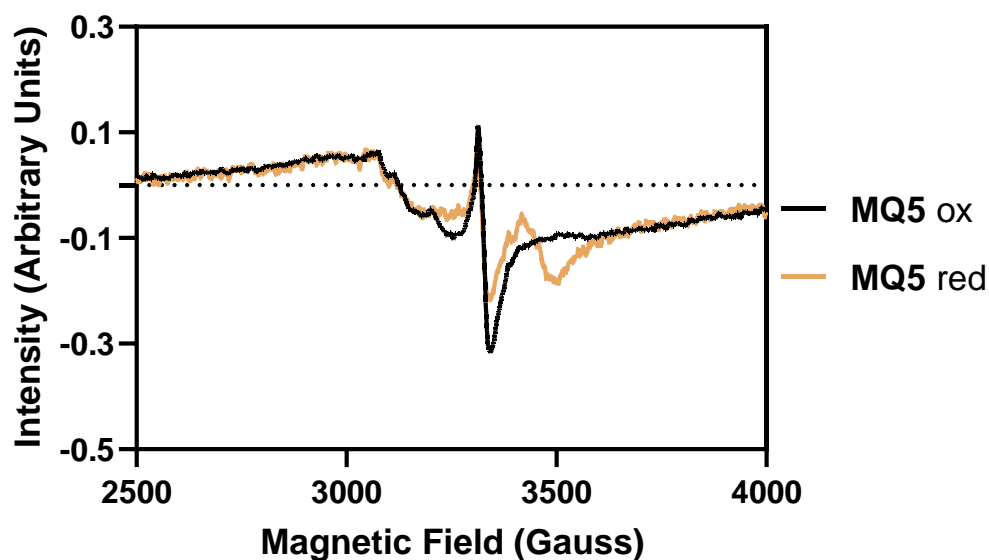


Figure 25 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in presence of  $\beta$ ME) **MQ5** oxidised (black trace) and dithionite-reduced (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

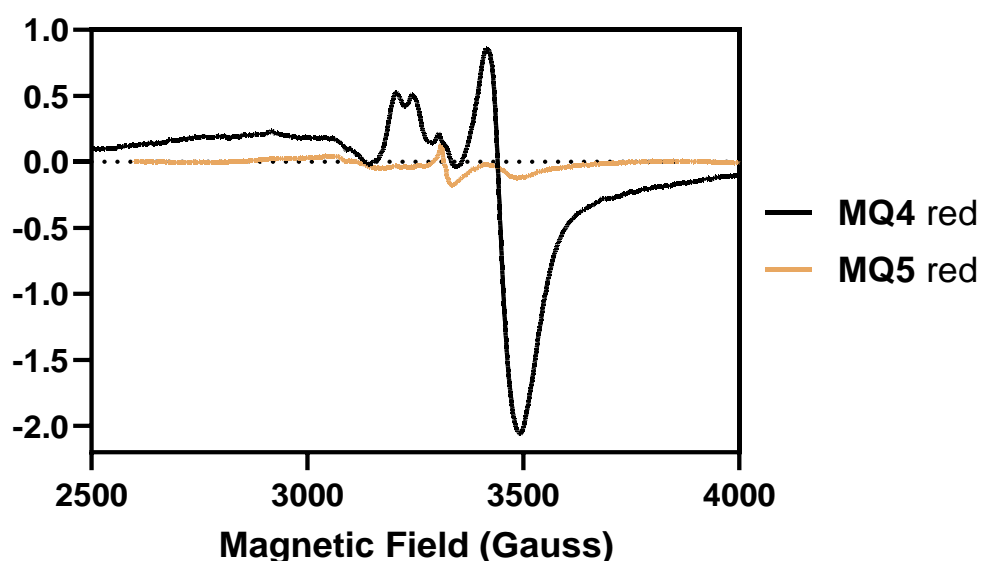


Figure 26 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in presence of  $\beta$ ME) dithionite-reduced: **MQ4** (black trace) and **MQ5** (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

Because both **MQ4** and **MQ5** were prepared following the same reconstitution procedure and concentration of reagents, the outcome of these results could possibly be explained by the different synthetic peptide used to incorporate the [4Fe-4S]-clusters. It was hypothesised that the longer sequence used for **MQ4** could partially fold around the cluster and exert a stability effect that protected the metal cubane from

its decomposition. The same effect is not guaranteed by the shorter sequences used for **MQ5** and therefore a much smaller concentration of [4Fe-4S] cluster could be produced.

We proceeded by exploring a cluster-maquette reconstitution protocol that lacked the reducing agent  $\beta$ ME. In the literature, there is no precedent of reconstituting Fe-S clusters with peptides in absence of reducing agents, and we were therefore the first to attempt such a methodology. Because the motif for the bacterial ferredoxin maquettes contains four Cys residues, the synthetic peptide should alone be capable of coordinating the Fe-S cluster without the aid of an extra ligand. The lack of reducing agent would allow for an easier purification that does not require separation of the thiol from the cluster maquette. Moreover, the UV-Vis trace would be a true representation of the LMCT produced by the [4Fe-4S]-peptide system only, without raising doubts about the presence of other species formed after the coordination between  $\beta$ ME and excess  $\text{Fe}^{3+}$  in solution. We used the same synthetic peptide of **MQ5** in a reconstitution protocol that omitted the reducing agent and compared the UV-Vis spectra (Figure 27).

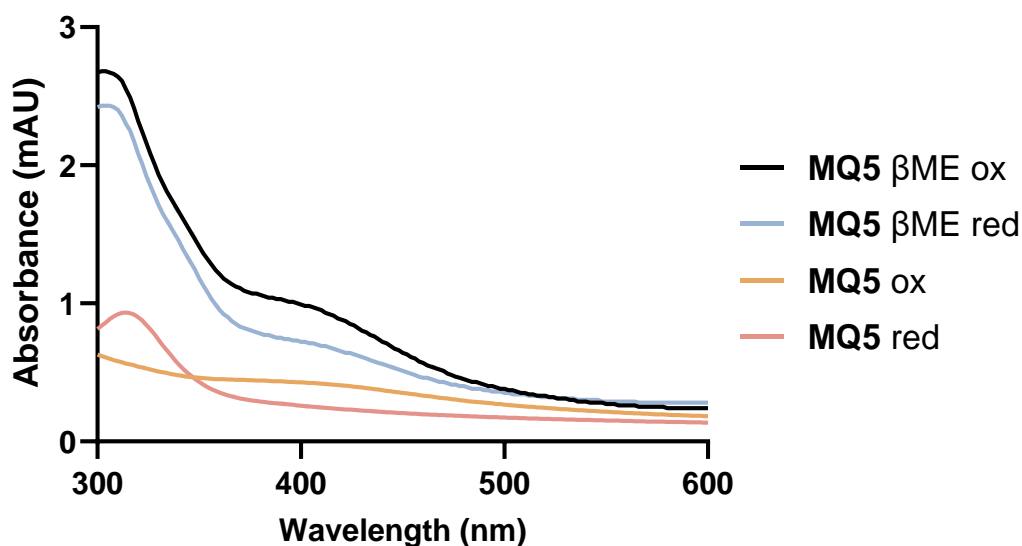


Figure 27 – UV-Vis trace for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ5** with and without  $\beta$ ME. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 80  $\mu\text{M}$ . The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

The difference in absorbance is clear between the two oxidised samples (black trace and gold trace) reconstituted respectively with and without the reducing agent. The

oxidised sample reconstituted in absence of  $\beta$ ME (gold trace) seems to show a flat line in proximity of the region where the LMCT absorbs, but this is solely due to the different scale. The sample reconstituted without reducing agent (gold trace) provides a true reflection of the Fe-S cluster presented in solution, whereas the sample containing  $\beta$ ME is enlarged by the LMCT effect exerted by the free thiol. The absorbance observed for the sample reconstituted with  $\beta$ ME and subsequently reduced with DT (light-blue trace) is larger than that of the oxidised sample reconstituted without  $\beta$ ME (gold trace), proving the degree of inflation detected when the thiol is in use. The reduced sample reconstituted without  $\beta$ ME (red trace) displays a new peak at approximately 310-320 nm, characteristic for excess DT in solution.<sup>85</sup> This peak confirms that the entire population of Fe-S cluster has been successfully reduced in the sample. However, this peak is not observed for the reduced sample of the  $\beta$ ME-reconstituted sample (light-blue trace), meaning that not all the cluster or the thiol-iron complex has been reduced. These results confirmed that, when  $\beta$ ME is used in the reconstitution process of the cluster-maquette, UV-Vis analysis could potentially be misleading. When the oxidised and reduced sample reconstituted using **MQ5** without  $\beta$ ME were zoomed in, the characteristic absorbance for the ligand-to-metal charge could be observed in the oxidised sample (Figure 28).

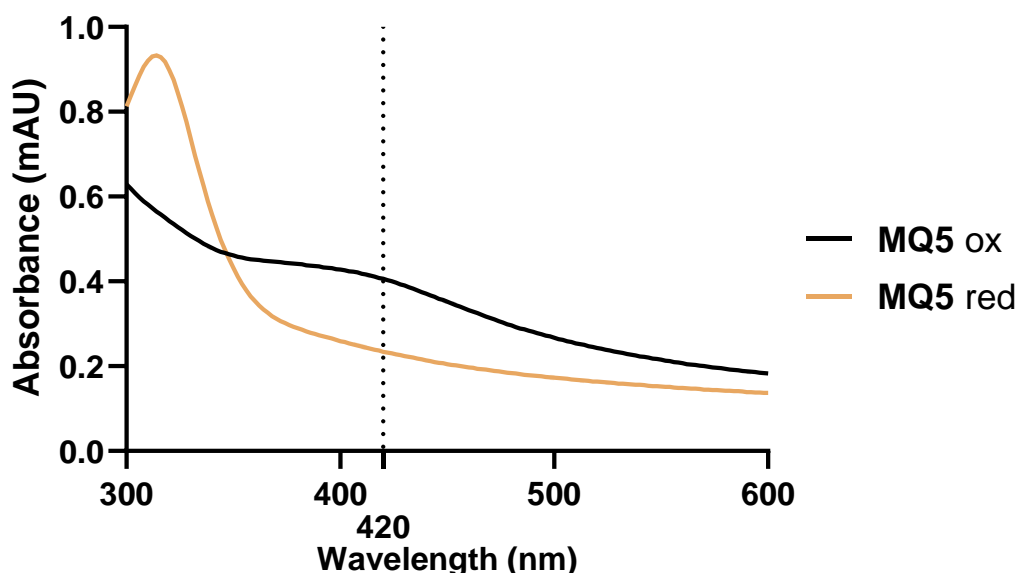


Figure 28 – UV-Vis trace for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ5**. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 100  $\mu$ M. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

To confirm that the LMCT absorbance observed was characteristic for the interaction between the peptide and the cluster, we synthesised a control peptide where all the Cys residues have been replaced by Gly residues. The resulting sequence, inspired by **MQ1**, was Ac-KLGEGGGIAGGAGGGW-NH<sub>2</sub> (**MQ6**) (Figure 29).

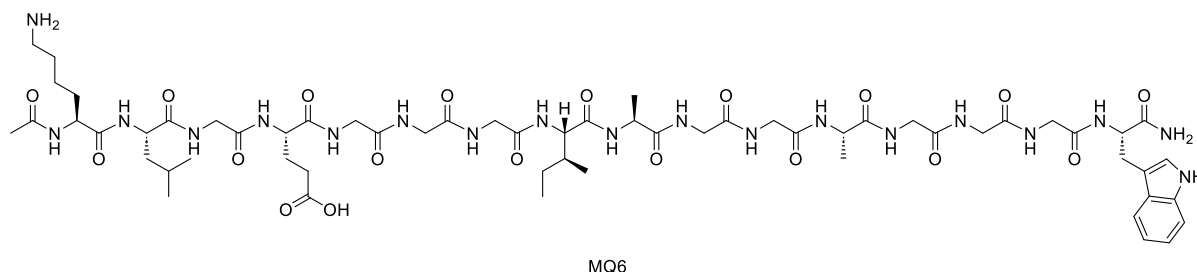


Figure 29 – Peptide sequence of **MQ6**.

Figure 30 shows the comparison of the UV-Vis traces for **MQ4** and **MQ6**. As expected, the sequence lacking the Cys residue showed a flat curve (gold trace) in proximity of the 400-420 nm range, meaning that no LMCT interaction was observed. On the contrary, the synthetic peptide with the four Cys residues displayed the characteristic peak at 400 nm.

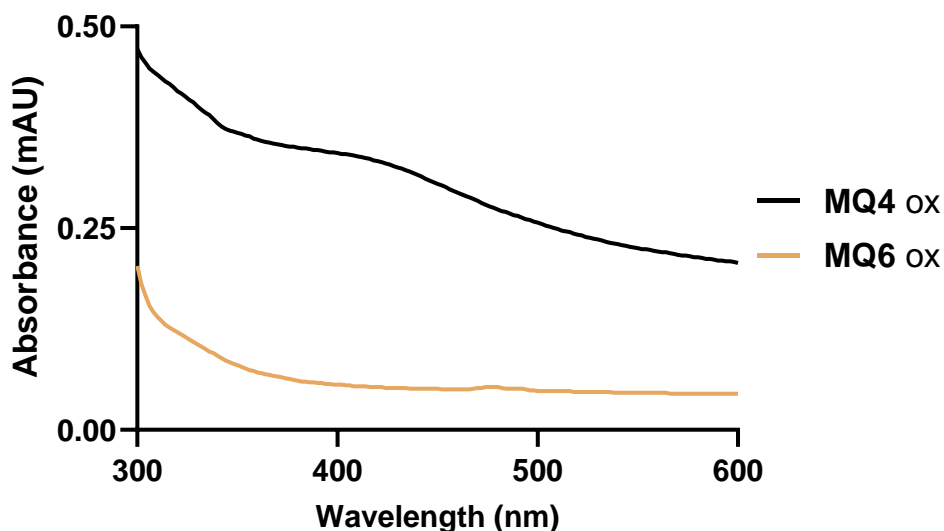


Figure 30 – UV-Vis trace for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ6**. The spectrum was recorded at peptide concentration of 100  $\mu$ M. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

This result confirmed that when **MQ4** was used to reconstitute the [4Fe-4S] cluster maquette in absence of  $\beta$ ME, the Cys residues were key for the successful incorporation and stabilisation of the cluster.

To prove that the reconstitution procedure could be successful even in absence of the reducing agent  $\beta$ ME, EPR analysis was required. The spectrum for the oxidised and reduced samples recorded for **MQ5** without  $\beta$ ME are reported in Figure 31.

EPR spectroscopy confirmed the presence of a population of [4Fe-4S]-cluster being reduced by the addition of DT to the oxidised sample. These results proved that the use of  $\beta$ ME is not required for the successful incorporation of a cluster when a maquette bearing four Cys residues is used. However,  $\beta$ ME was thought to exert a stabilisation effect by ensuring all the Cys residues remain reduced in solution before incorporation of the Fe-S cluster. To study this aspect, we compared the EPR results recorded for the two **MQ5** samples reconstituted both with and without the reducing agent (Figure 32).

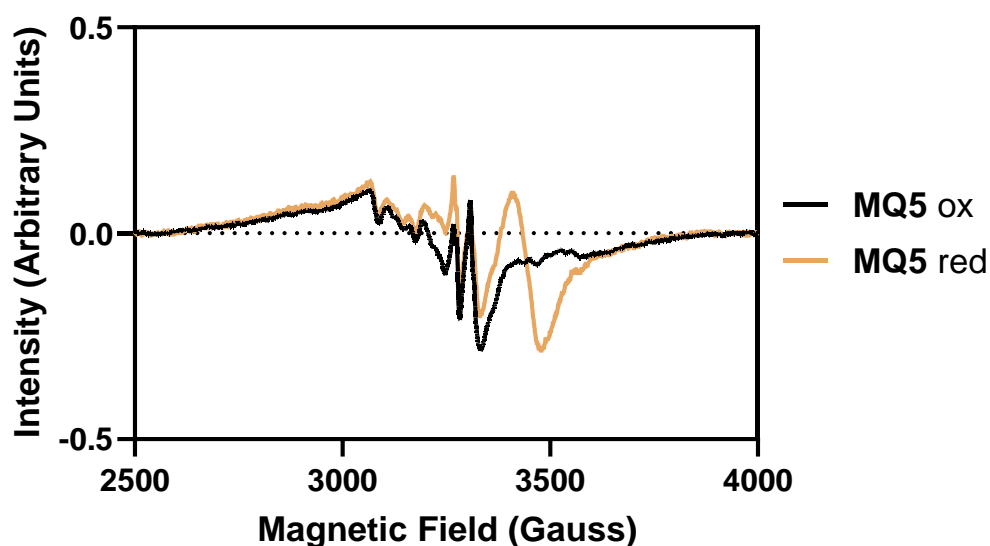


Figure 31 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in absence of  $\beta$ ME) **MQ5** oxidised (black trace) and dithionite-reduced (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

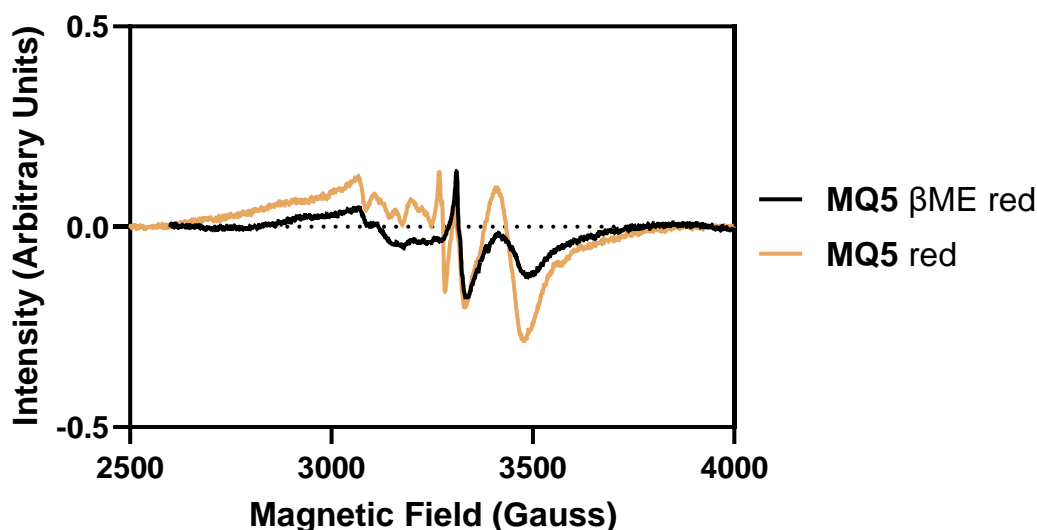


Figure 32 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted) **MQ5** dithionite-reduced and reconstituted in presence (black trace) and in absence (gold trace) of  $\beta$ ME. EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

Interestingly, the cluster population produced in presence of  $\beta$ ME was lower than that produced in absence of the reducing agent. It is reasonable to think that an excess of  $\beta$ ME in solution could coordinate a great fraction of the iron ions available to the synthetic peptide. As a result of this competing reaction, the amount of peptide could be larger than the amount of  $\text{Fe}^{3+}$  ions available to compose the cluster. Moreover,  $\beta$ ME is a much smaller molecule than the synthetic peptide and can easily and more quickly secure the iron ions to form a complex.

Convinced that the peptide sequence could be further engineered to increase the concentration of Fe-S cluster, we added two flanking amino acids on both the C- and N-terminus of the peptide. The aim was to increase the length of the synthetic peptide to improve the folding activity, but also to add a barrier on both sides of the peptide and not leave the two Cys residues exposed at each ends. The sequence we synthesised was Ac-G**CGGGCGGCGGCG**-NH<sub>2</sub> (**MQ7**) and is showed in Figure 33. The synthesis of this peptide, as well as all the others presented in this thesis, was achieved using the Fmoc SPPS strategy and detailed protocols are described in 8.2 GENERAL PROTOCOLS.

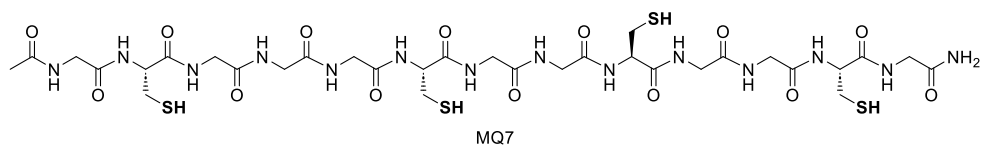


Figure 33 – Peptide sequence of **MQ7**.

The incorporation of the Fe-S cluster using **MQ7** was revised and  $\beta$ ME was no longer included in the protocol. Contrarily to **MQ1** that was engineered by taking inspiration from the radical SAM class of enzymes and exhibited three Cys residues, the synthetic peptide of **MQ6** alone should be able to fully coordinate the four iron ions of the cluster without the assistance of an extra ligand. In the synthesis of cluster maquettes  $\beta$ ME is generally used as a reducing agent to maintain the Cys residues reduced and avoid formation of disulphide bonds that would prevent the ligand-to-metal interaction with a detrimental effect on the cluster coordination by the peptide. For this reason,  $\beta$ ME is frequently found as an additive in the reconstitution buffer for cluster maquettes or, more in general, to ensure the stability of peptides bearing Cys residues.<sup>86</sup> Figure 34 shows the UV-Vis spectra recorded for the Fe-S cluster maquette sample reconstituted using **MQ7** without  $\beta$ ME.

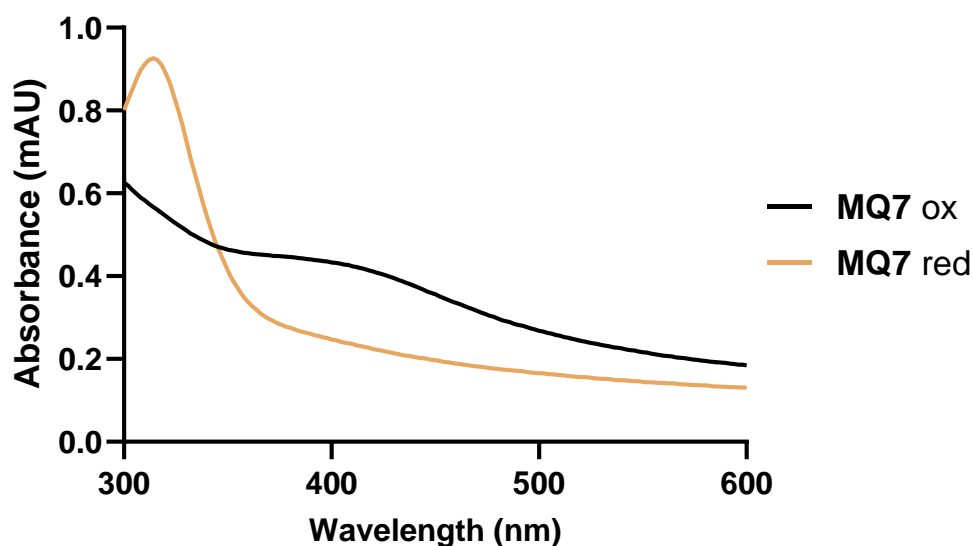


Figure 34 – UV-Vis trace for  $[4\text{Fe-4S}]$ -cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ7**. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 80  $\mu\text{M}$ . The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

The oxidised sample (black trace) showed the characteristic curve at approximately 420 nm, as expected for the LMCT of a Fe-S cluster. Due to the absence of  $\beta$ ME in



the sample, the signal recorded was a true reflection of the concentration of Fe-S cluster in solution. Under these circumstances, UV-Vis spectroscopy could be used to determine whether the reconstitution procedure had been successful or not and could act as an alternative to EPR spectroscopy. The sample treated with DT (gold trace) showed conversion from the oxidised  $[4\text{Fe-4S}]^{2+}$  cluster to the reduced  $[4\text{Fe-4S}]^{1+}$  species. A peak appeared at approximately 310-320 nm that is characteristic of excess DT in solution.<sup>85</sup> The UV-Vis traces for sample reconstituted with **MQ7** looked similar to those obtained for **MQ5**; and when the four traces were plotted together, they nearly superimposed with each other (Figure 35).

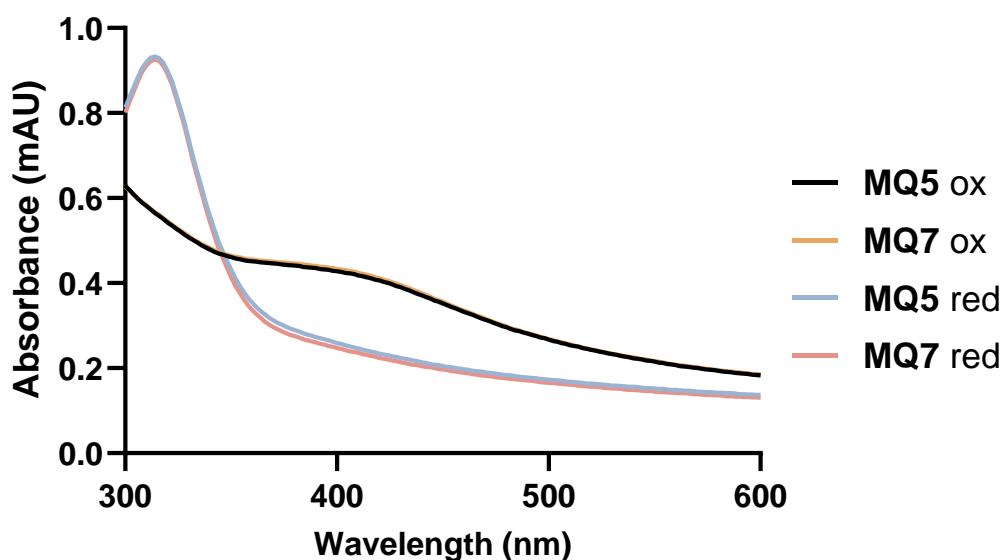


Figure 35 – UV-Vis trace for  $[4\text{Fe-4S}]$ -cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ5** and **MQ7**. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of  $80 \mu\text{M}$ . The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

Again, when the spectra for the oxidised samples obtained using **MQ4** (Ac-KLCEGGCIACGACGGW-NH<sub>2</sub>) and **MQ7** (Ac-GCGGGCGGCGGCG-NH<sub>2</sub>) were compared, the difference in intensity of the peaks was more evident (Figure 36).

Such a difference in absorbance could be explained by two reasons. Firstly, the presence of  $\beta\text{ME}$  in **MQ4** could interact with the excess of  $\text{Fe}^{3+}$  in solution and produce the LMCT with the characteristic peak in absorption between 400 and 420 nm. Secondly, the absence of reducing agent in **MQ7** could lead to partial oxidation of the Cys residues, limiting the formation of  $[4\text{Fe-4S}]$  clusters. Disulphide bond formation in small peptides and proteins has been shown to be favoured in basic conditions, at pH

greater than 8.5.<sup>87</sup> The buffer used for the cluster-maquette reconstitution procedure has a slightly basic pH of 8.0, but lower than the optimal pH for the disulphide bond formation. For this reason, we speculated that the free thiol on the Cys residues should be favoured, providing adequate conditions for the cluster incorporation. The EPR analysis for the sample produced using **MQ7** proved the presence of a [4Fe-4S] cluster in the sample treated with DT, as showed in Figure 37.

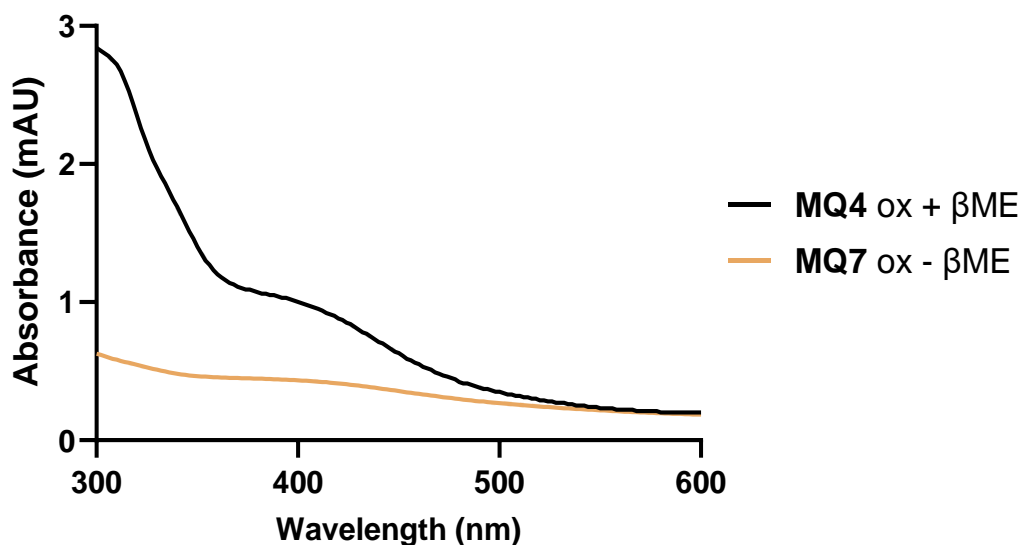


Figure 36 – UV-Vis trace for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ4** and **MQ7**. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 80  $\mu$ M. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

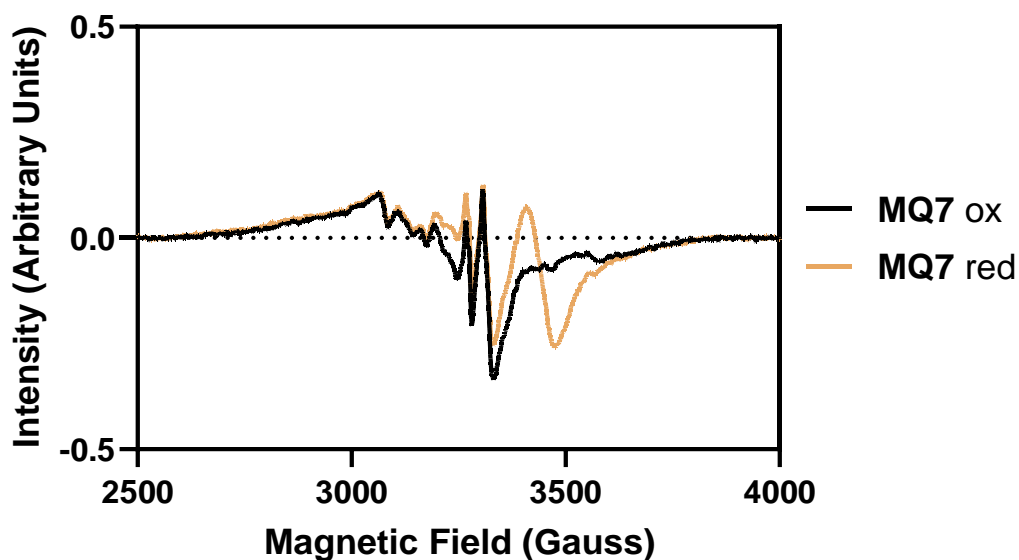


Figure 37 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in absence of  $\beta$ ME) **MQ7** oxidised (black trace) and dithionite-reduced (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

The sample treated with DT showed the presence of a small peak, confirming that the [4Fe-4S]-cluster maquette was successfully synthesised. The reconstitution procedure in absence of  $\beta$ ME showed reproducible results that could be extended to different synthetic peptides. For this reason, we decided to synthesise a final maquette with the bacterial ferredoxin motif. This synthetic peptide was obtained by adding a Tyr residue to the C-terminus of the previous maquette, yielding the peptide sequence Ac-GCGGGCGGCGGY-NH<sub>2</sub> (**MQ8**) (Figure 38). The presence of a Tyr residue acts as a UV-Vis spectroscopy handle by absorbing at approximately 280 nm due to the conjugated system of its side chain, therefore facilitating the purification of the peptide.

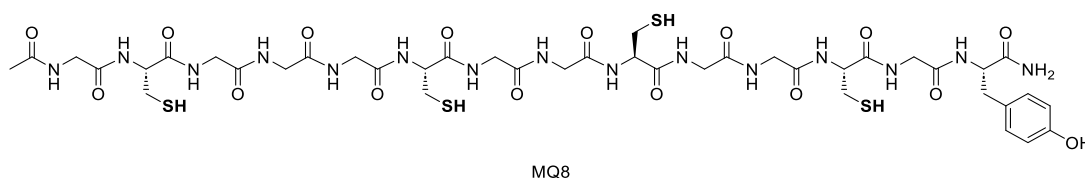


Figure 38 – Peptide sequence of **MQ8**.

Due to the similarity of this peptide sequence to the other previously investigated (**MQ5** and **MQ8**), similar results were expected for both the UV-Vis and EPR analysis. The UV-Vis spectroscopy results for the oxidised and reduced samples are displayed in Figure 39.

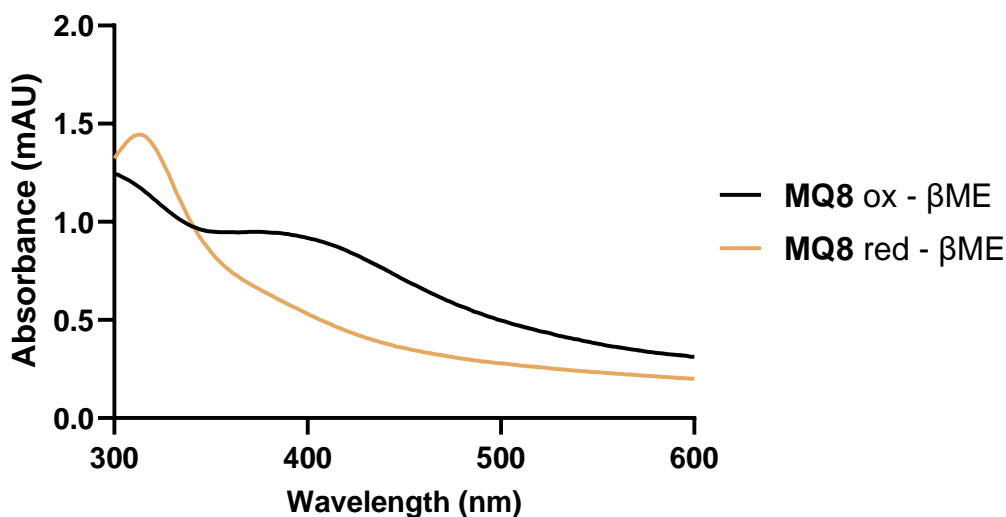


Figure 39 – UV-Vis trace for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ8** in absence of βME. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 80 μM. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

The reconstitution procedure was performed in absence of βME, and therefore the characteristic peak observed at 400 nm reflected the presence of a Fe-S cluster. The sample treated with DT showed a flattened trace in proximity of the region where the LMCT absorbs, proving that the cluster was reduced from the 2+ to the 1+ oxidation state. A peak at 310 nm was also detected in the reduced sample, confirming that DT was in excess and successfully reduced the totality of the cluster species in solution. The EPR analysis proved the presence of a reduced [4Fe-4S]<sup>1+</sup> cluster in the sample exposed to DT, as displayed in Figure 40.

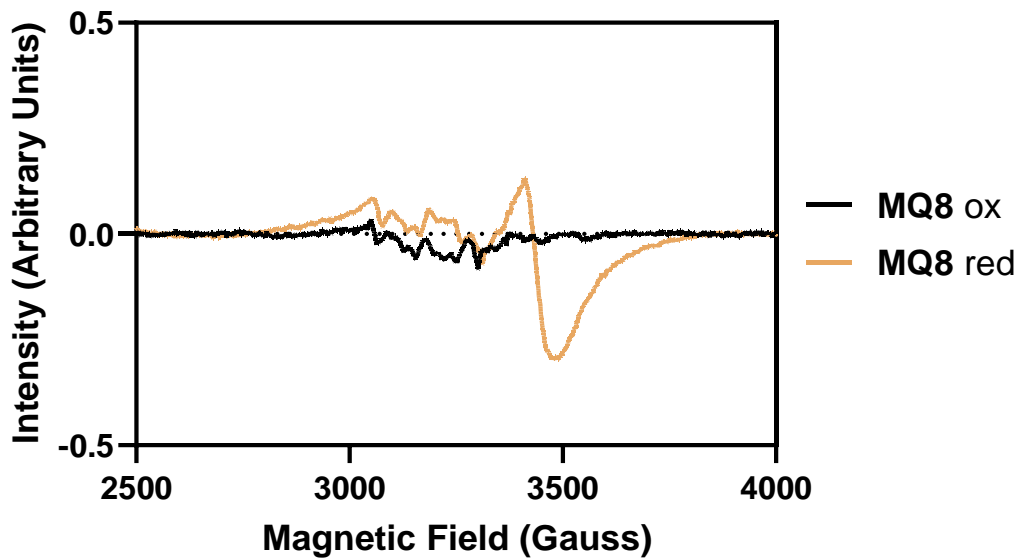


Figure 40 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in absence of  $\beta$ ME) **MQ8** oxidised (black trace) and dithionite-reduced (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

As for the other two engineered synthetic peptides **MQ5**, and **MQ7**; **MQ8** too showed low intensity for the cluster population. Figure 41 compares the EPR results obtained for the reduced samples **MQ5**, **MQ7**, and **MQ8**.

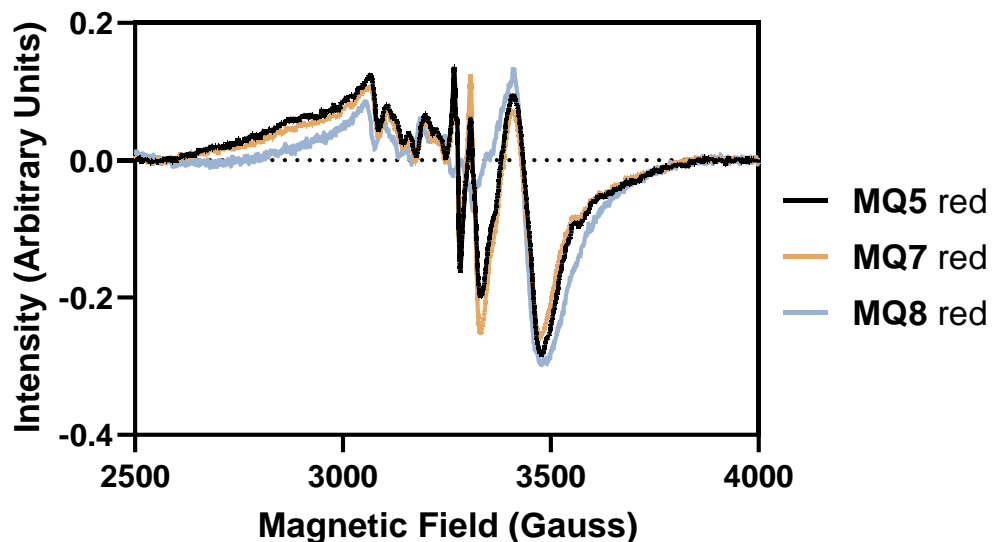


Figure 41 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in absence of  $\beta$ ME) dithionite-reduced: **MQ4** (black trace), **MQ7** (gold trace), and **MQ8** (light-blue trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

The results obtained for the three different Fe-S cluster maquettes reconstituted in absence of  $\beta$ ME were all comparable to each other. These results proved that synthetic peptides that only differed for the flanking amino acids on both the C- and N-terminus did not provide great effect on the reconstitution capabilities of the maquette. Figure 42 compares the EPR results obtained for **MQ8** to those obtained when the cluster maquette was reconstituted using **MQ4** in absence of  $\beta$ ME.

This outcome demonstrated that, in absence of  $\beta$ ME, the engineered **MQ8** was more successful in coordinating a [4Fe-4S] cluster with respect to the native-inspired **MQ4**. To conclude, we were able to produce an engineered maquette that proved successful in reconstituting an [4Fe-4S]-cluster and showed positive results when analysed using both UV-Vis and EPR spectroscopy. Furthermore, we developed a new reconstitution protocol for the synthesis of Fe-S cluster maquettes that did not involve the use of  $\beta$ ME as a reducing agent. We envision that this synthetic procedure could be of great help when using cluster maquettes as catalyst in organic reactions or in electron transport chain, as it removes a particularly reactive chemical from the reaction mixture and does not require the sample to be eluted through SEC columns for purification.

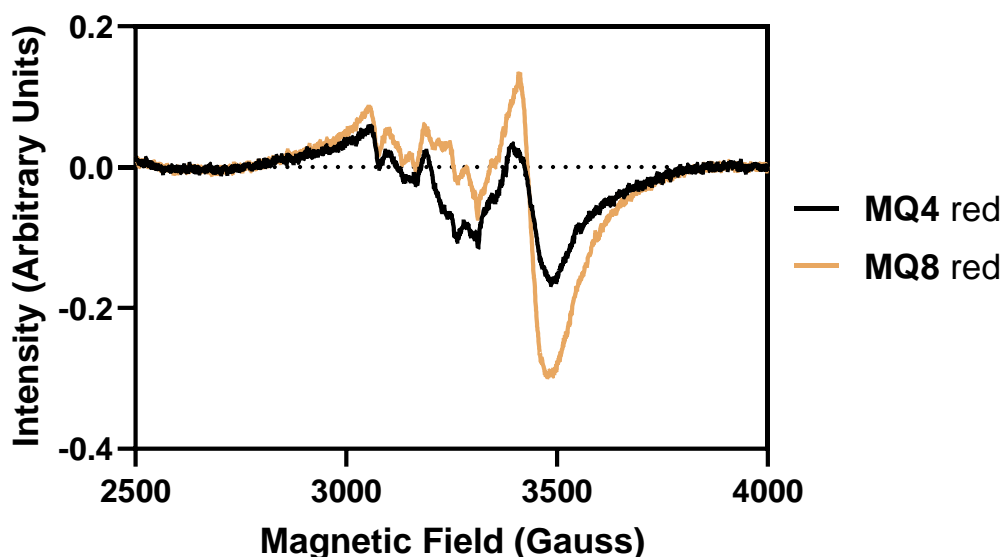


Figure 42 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in absence of  $\beta$ ME) dithionite-reduced: **MQ4** (black trace) and **MQ7** (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.



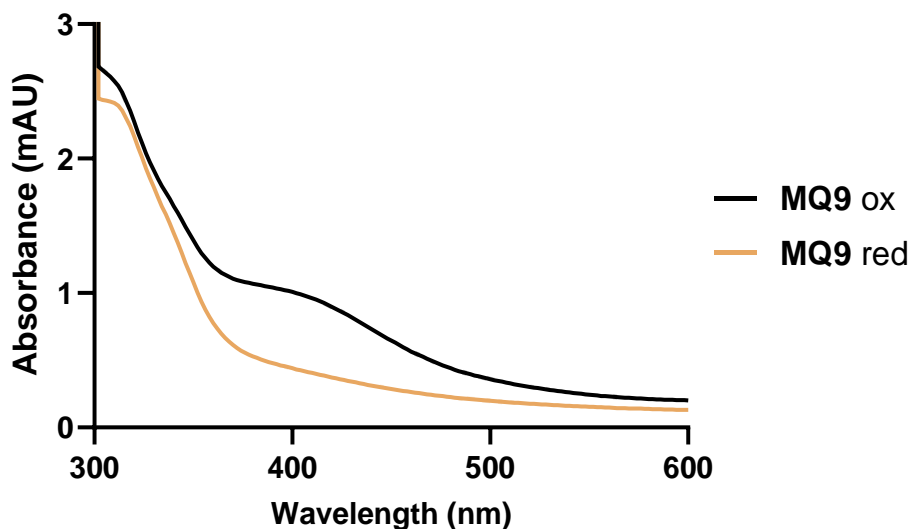


Figure 44 – UV-Vis trace for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ9**. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 80  $\mu\text{M}$ . The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

The characteristic peak at 420 nm was observed for the oxidised sample and as expected, treatment with DT flattened the curve in the reduced sample. However, due to the presence of the free thiol acting as an extra ligand, UV-Vis analysis could not be considered satisfactory to confirm the presence of a [4Fe-4S]-cluster. For this reason, EPR analysis was carried out and the results are presented in Figure 45.

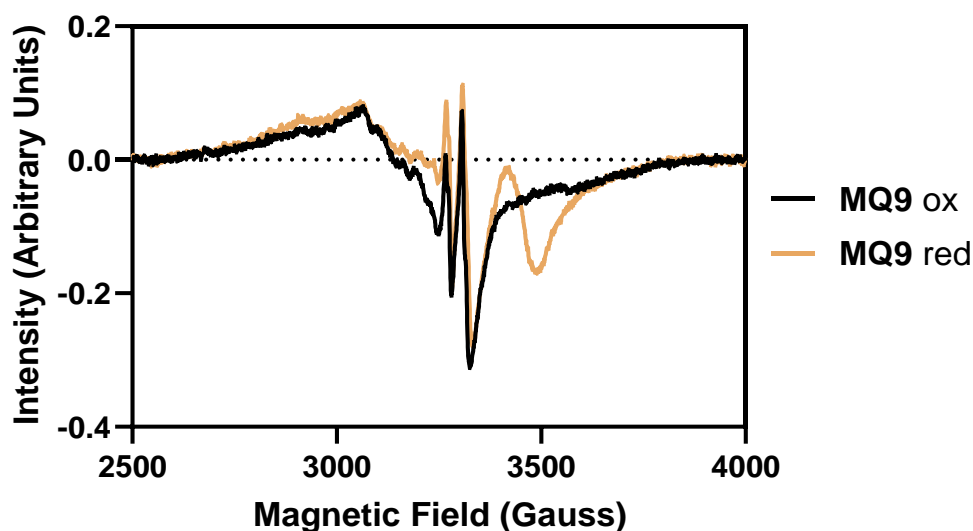


Figure 45 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in absence of  $\beta\text{ME}$ ) **MQ9** oxidised (black trace) and dithionite-reduced (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.



EPR spectroscopy confirmed the presence of a reduced  $[4\text{Fe}4\text{S}]^{1+}$  cluster when the sample was treated with DT (gold trace). However, the population was at low concentration and therefore the signal was weak. It is plausible that the amount of DT used was not enough to successfully reduce both the maquette and the  $\beta\text{ME}$ -iron ions complexes produced in presence of an excess of these two reagents. As explored for the bacterial ferredoxin maquettes, we investigated the effect of the amino acid residues on the capabilities of the maquette to coordinate the cluster. For this reason, we changed all non-Cys residues to Gly, apart from the Trp residue on the C-terminus, and we left all the flanking amino acids. The sequence of the engineered synthetic peptide  $\text{Ac-GGCGGGCGGCGGGGGW-NH}_2$  (**MQ10**) is displayed in Figure 46.

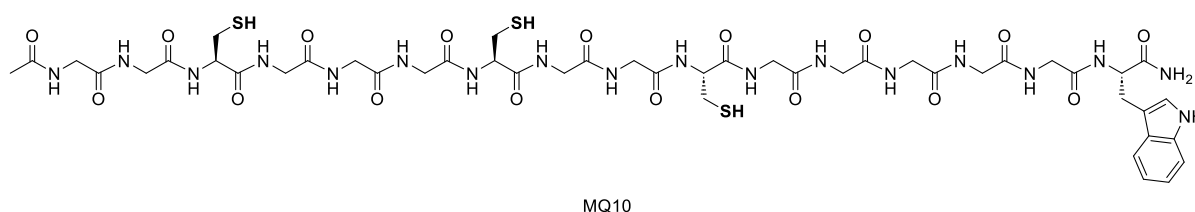


Figure 46 – Peptide sequence of **MQ10**.

UV-Vis analysis was consistent with the expectations and showed an absorption peak in the range between 400 and 420 nm for the oxidised sample (Figure 47, black trace). To reiterate, this absorption could not be considered satisfactory for the presence of a  $[4\text{Fe}-4\text{S}]$  cluster, due to the use of  $\beta\text{ME}$  in the reconstitution process.

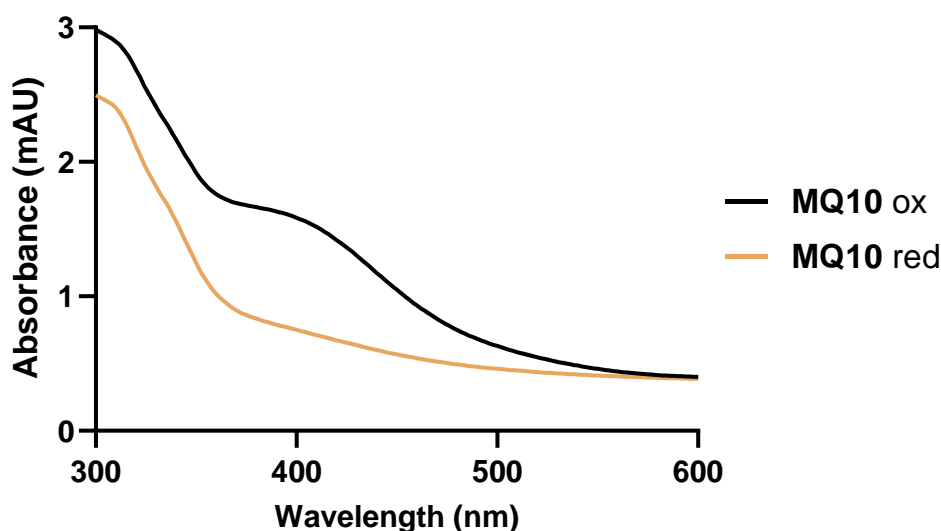


Figure 47 – UV-Vis trace for  $[4\text{Fe}-4\text{S}]$ -cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ10**. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 80  $\mu\text{M}$ . The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

EPR spectroscopy confirmed the presence of a reduced Fe-S cluster in the sample treated with DT, as presented in Figure 48.

Unfortunately, the cluster population obtained with **MQ10** was extremely low and the signal was barely distinguished from the baseline. When the EPR analysis obtained for the reduced sample of both **MQ8** and **MQ9** were compared, the difference in signal intensity and shape was substantial (Figure 49).

These results confirmed that, as well as for the bacterial ferredoxin motif, the amino acid residues play a role in the cluster maquette reconstitution. Interestingly, the engineered poly-Gly sequence **MQ8** (Ac-GCGGGCGGCGGCGY-NH<sub>2</sub>) produced better results compared to the native sequence **MQ4** (Ac-KLCEGGCIACGACGGW-NH<sub>2</sub>) when the cluster was reconstituted in absence of  $\beta$ ME. In this case, the poly-Gly sequence **MQ10** (Ac-GGCGGGCGGCGGGGGW-NH<sub>2</sub>) yielded a smaller population of [4Fe-4S] cluster when compared to the native-inspired sequence **MQ9** (Ac-KLCEGGCIACGAGGGW-NH<sub>2</sub>). Overall, these results proved that a maquette mimicking the radical SAM family of enzyme could be produced provided the **CX<sub>3</sub>CX<sub>2</sub>C** motif is retained.

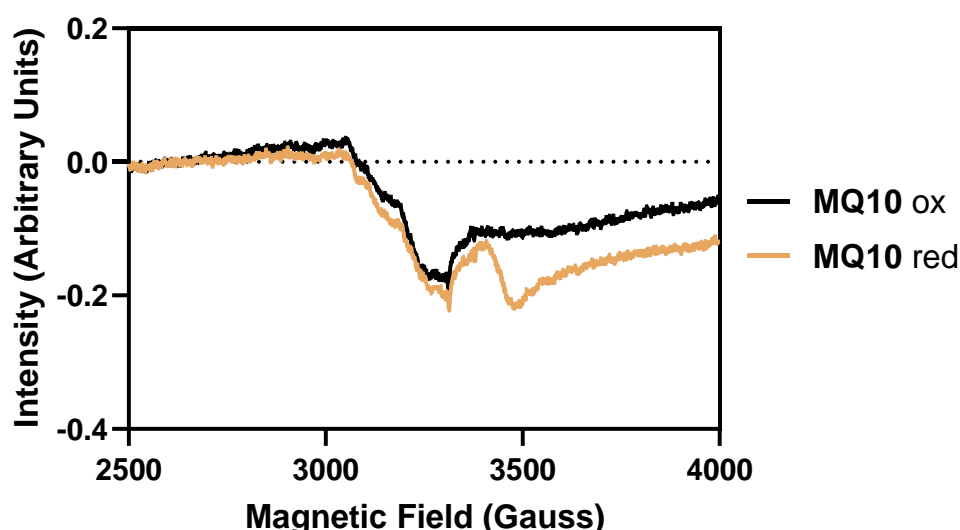


Figure 48 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in absence of  $\beta$ ME) **MQ10** oxidised (black trace) and dithionite-reduced (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

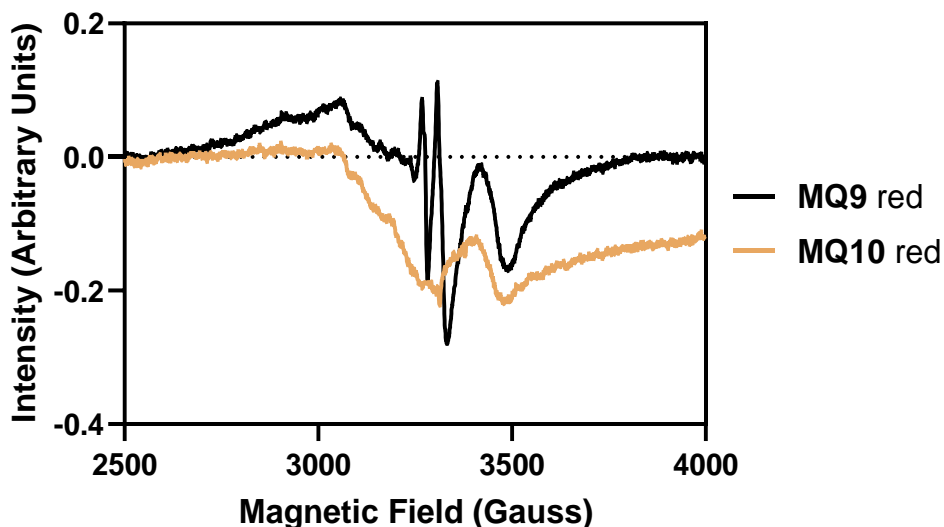
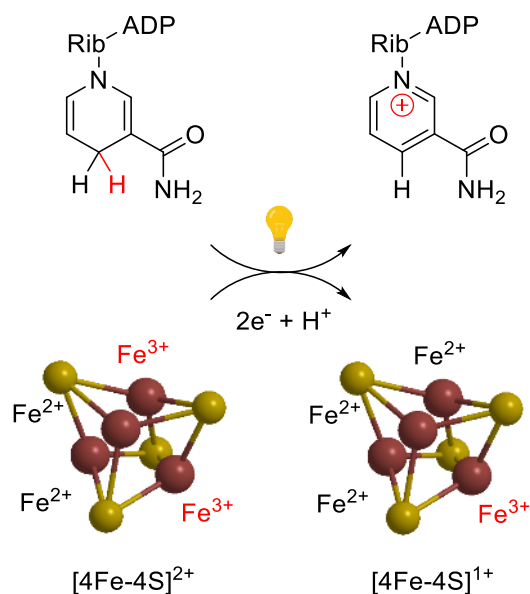


Figure 49 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in absence of  $\beta$ ME) dithionite-reduced: **MQ9** (black trace) and **MQ10** (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

### 2.2.5 Reactivity of [4Fe-4S] Cluster Maquettes to Molecular Oxygen

To investigate the reactivity of [4Fe-4S]-cluster maquettes towards molecular oxygen, we analysed by EPR spectroscopy samples that were prepared in aerobic conditions. As previously described, all samples analysed by EPR spectroscopy were prepared in both anaerobic and aerobic conditions, the latter as a control. Dr. Muralidharan Shanmugam once attempted to use a different reducing agent (NADH) with respect to DT to investigate whether this could be powerful enough to induce the reduction from the 2+ to the 1+ state in a [4Fe-4S] cluster. The samples were synthesised following the same reconstitution protocol used in previous experiments (see 8.2 *GENERAL PROTOCOLS*), but this time outside a glove box and using non-degassed solvents. Previously, photoactivated NADH was used as an electron donor to reduce a [2Fe-2S]<sup>2+</sup> cluster, showcasing the property of this coenzyme as a powerful reducing agent.<sup>88</sup> The photoactivation of NADH at 365 nm triggers the release of solvated electrons and radical species that can reduce redox centres in fluid or frozen samples.<sup>89</sup> We postulated that the [4Fe-4S] cluster could be formed in aerobic conditions and that the use of photoactivated NADH could access the reduced, 1+, state of the cluster to convert it into an EPR-visible species. This hypothesis assumes that the cluster is successfully formed and coordinated by the peptide in an oxygen-

rich environment. Scheme 10 displays the reaction involved in these experiments, where photoactivated NADH is oxidised to reduce the cluster from the 2+ to the 1+ oxidation state.



*Scheme 10 – Schematic photoactivated NADH reduction of a [4Fe-4S] from the 2+ to the 1+ oxidation states.*

The first experiment was performed using **MQ4** to incorporate the Fe-S cluster in presence of  $\beta$ ME. In this case, we hypothesized that reconstitution of the cluster-maquette in an oxygen-rich atmosphere required the reducing agent to ensure the thiols of the Cys were in the reduced oxidation state. Once the cluster maquette was reconstituted, the sample was transferred into an EPR tube and flash frozen in liquid nitrogen for analysis. Figure 50 displays the EPR analysis of the cluster maquette sample to which NADH was added, before (black trace) and after (gold trace) photoactivation of NADH.

From the EPR results it is clear how before activation of NADH at 365 nm (black trace), the only signal observed was the one produced by the buffer solution. After photoactivation of NADH (gold trace), the characteristic rhombic signal for the [4Fe-4S]-cluster became visible. It is plausible to think that the oxidised [4Fe-4S]<sup>2+</sup> cluster was successfully synthesised in aerobic conditions and using non-degassed solvents. Interestingly, when the aerobically reconstituted sample was treated with DT, no reduced [4Fe-4S]<sup>1+</sup> cluster could be detected, as presented in Figure 51. The oxidised (black trace) and DT-reduced (gold trace) samples displayed a similar trend and the only signal observed was the buffer background signal.

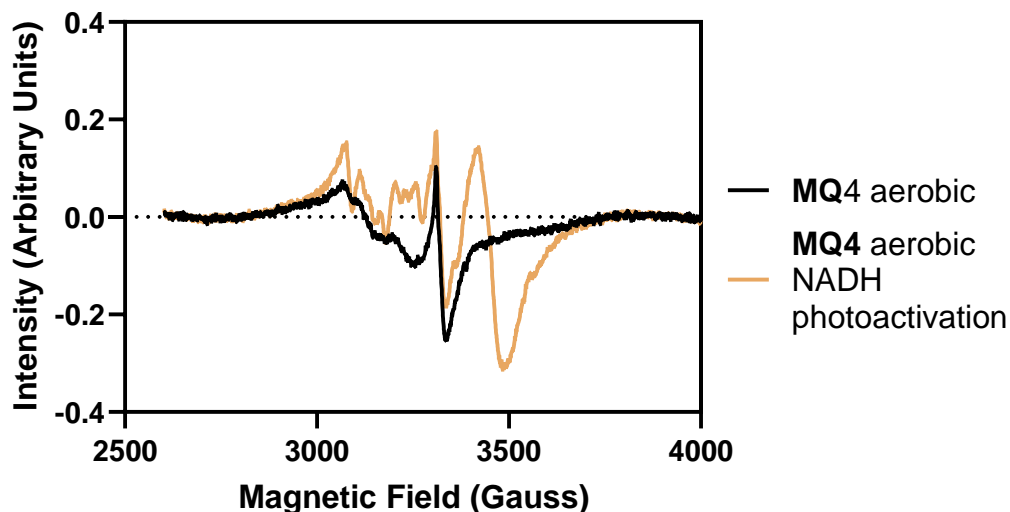


Figure 50 – X-band cw-EPR spectra at 20 K of a frozen solution aerobically reconstituted in presence of  $\beta$ ME of **MQ4** pre (black trace) and post (gold trace) photoactivation of NADH at 365 nm. EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

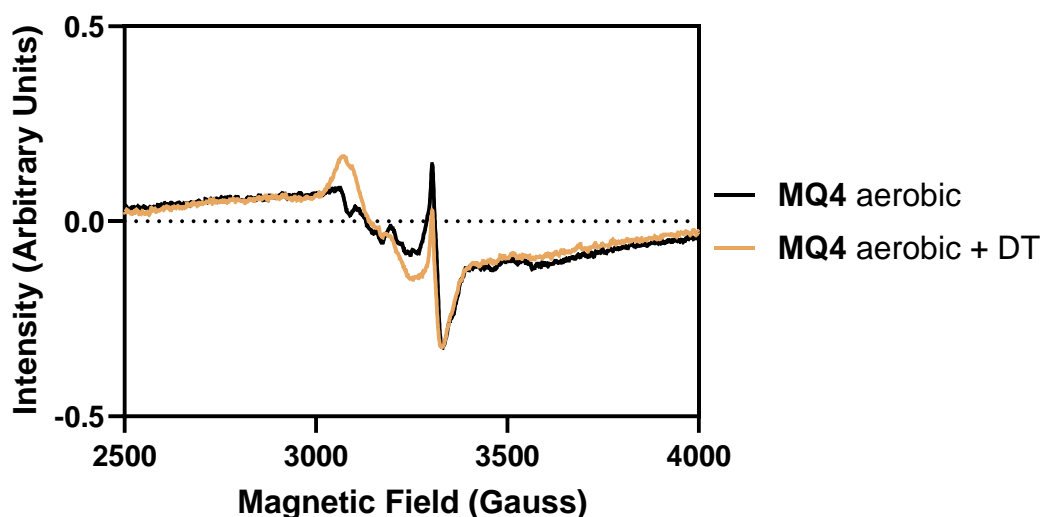


Figure 51 – X-band cw-EPR spectra at 20 K of a frozen solution aerobically reconstituted in presence of  $\beta$ ME of **MQ4** oxidised (black trace) and dithionite-reduced (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

Interestingly, despite DT being successful in reducing the Fe-S cluster in anaerobically reconstituted samples, this was not the case when the sample was synthesised in aerobic conditions. When the EPR results for the NADH-photoactivated sample was compared to that of the anaerobically reconstituted sample in presence of  $\beta$ ME and treated with DT, it was clear that the population of reduced cluster obtained by photoactivation of NADH was in lower concentration (Figure 52).

It is arguable that, despite  $\beta$ ME being used as an additive for the aerobically reconstituted sample, the amount of oxygen was enough to partially oxidise the Cys residues of the synthetic peptide, preventing them from coordinating the cluster. To explore the reproducibility of these findings, we performed a similar experiment but on the radical SAM-inspired maquettes **MQ9** (Ac-KL**C**EGG**C**IACGAGGGW-NH<sub>2</sub>) and **MQ10** (Ac-GG**C**GGG**C**GG**C**GGGGW-NH<sub>2</sub>). Again, the samples were reconstituted in aerobic conditions with  $\beta$ ME and then flash frozen prior exposure to DT. The EPR spectra of both **MQ9** and **MQ10** in presence of NADH before and after photoactivation are displayed in Figure 53.

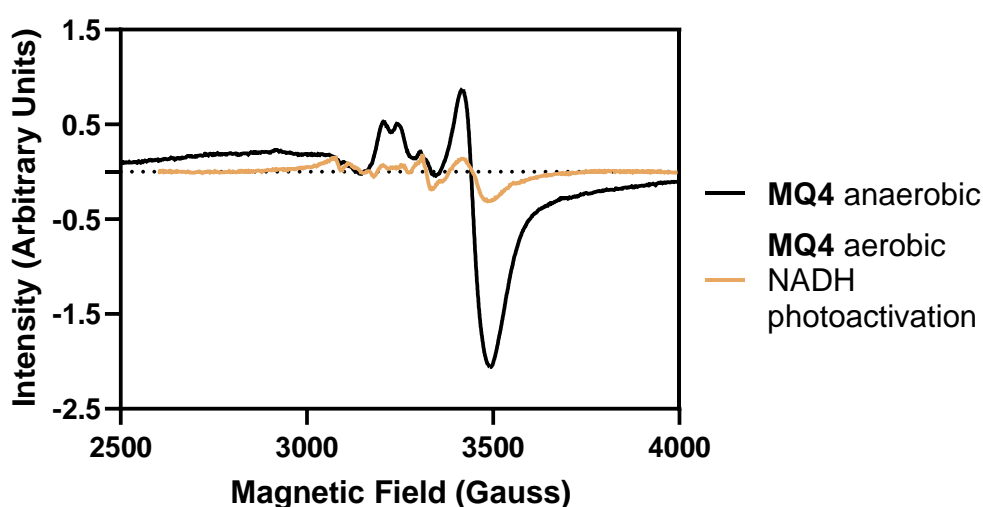


Figure 52 – X-band cw-EPR spectra at 20 K of a frozen solution reconstituted in presence of  $\beta$ ME of **MQ4** anaerobically reconstituted (black trace) and aerobically reconstituted after NADH photoactivation (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

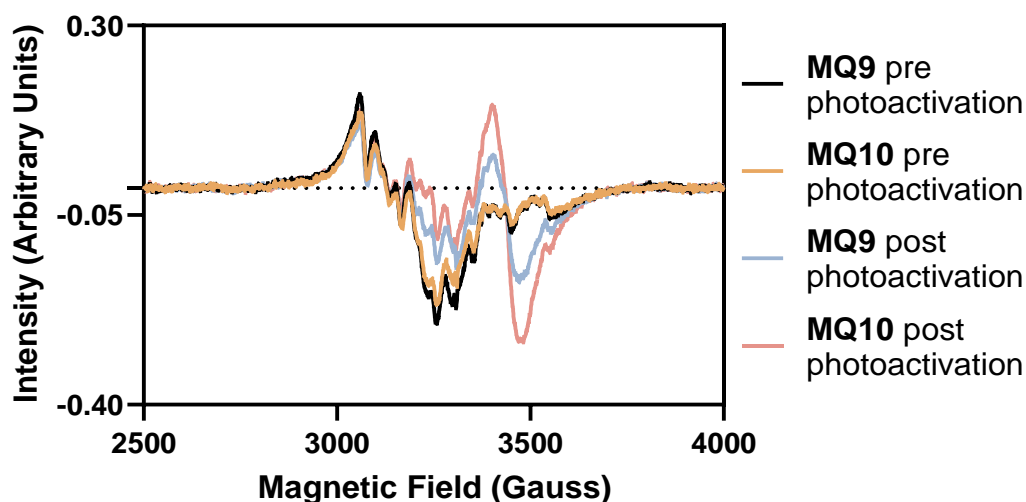


Figure 53 – X-band cw-EPR spectra at 20 K of a frozen solution aerobically reconstituted in presence of  $\beta$ ME of **MQ9** and **MQ10** pre (respectively black and gold trace) and post (respectively light-blue and red trace) NADH photoactivation. EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

As previously observed for the bacterial ferredoxin **MQ4** (Ac-KLCEGGCIACGACGGW-NH<sub>2</sub>), the EPR signals before photoactivation of NADH only showed the background signal of the buffer (black and gold traces). After photoactivation at 365 nm, the characteristic rhombic signal for the reduced [4Fe-4S]<sup>1+</sup> cluster became visible for both **MQ9** and **MQ10**, respectively light-blue and red traces. Interestingly, the analysis for the poly-Gly **MQ10** showed a more intense signal, implying a larger population of cluster present in the sample. This result contrasts with that presented in Figure 49, where the cluster population coordinated by **MQ9** was proved to be larger than that coordinated by **MQ10** when the samples were reconstituted in anaerobic conditions and reduced using DT. It is plausible that, because the analysis was carried on two different batches of samples, one aerobically and one anaerobically reconstituted, the cluster reconstitution was not consistent between the two samples and therefore produced a slight difference in the intensity of the EPR signals.

Inspired by our findings regarding the avoidable use of  $\beta$ ME as a reducing agent for the reconstitution of cluster maquettes, we tested the NADH photoactivation of aerobically reconstituted samples in absence of the thiol. In this experiment we were interested in investigating whether the presence of  $\beta$ ME was again required for the successful incorporation of the cluster in the synthetic peptide, or if it could be superfluous. The cluster was reconstituted using both **MQ4** and **MQ8** (Ac-

GCGGGC<sub>2</sub>GGC<sub>2</sub>GGC<sub>2</sub>GY-NH<sub>2</sub>) in aerobic conditions. The EPR results are displayed in Figure 54 and displays the traces before and after photoactivation of NADH.

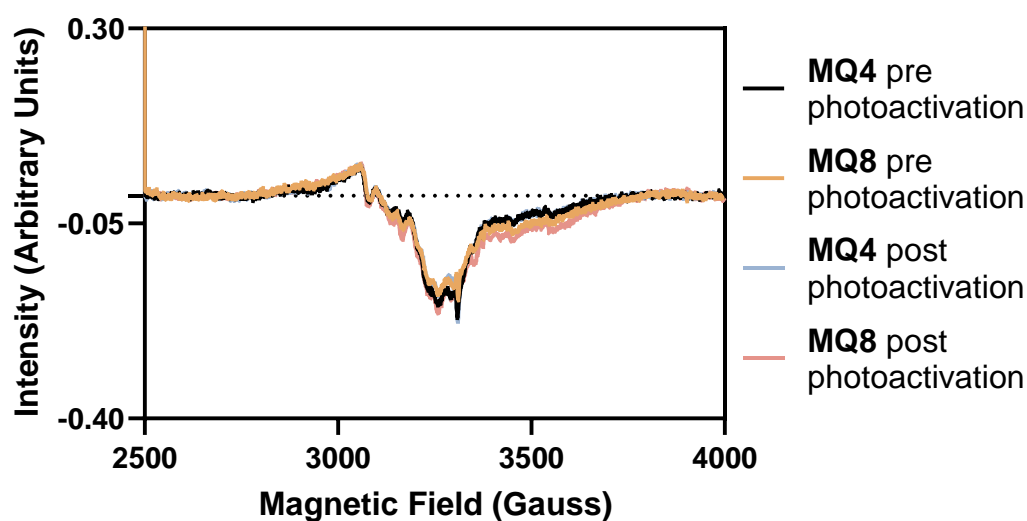


Figure 54 – X-band cw-EPR spectra at 20 K a frozen solution aerobically reconstituted in absence of  $\beta$ ME) of **MQ4** and **MQ8** pre (respectively black and gold trace) and post (respectively light-blue and red trace) NADH photoactivation. EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

From these results we deduced that no reduced cluster species could be detected after photoactivation of NADH at 365 nm. The signals registered before and after activation were identical and no traces of a  $[4\text{Fe-4S}]^{1+}$  could be detected by EPR. We therefore concluded that  $\beta$ ME is required for the successful incorporation of a  $[4\text{Fe-4S}]$  cluster using both bacterial ferredoxin and radical SAM inspired maquettes when the reconstitution is attempted in aerobic conditions. These results were, to the best of our knowledge, never presented before, and could lead to new applications of these Fe-S cluster maquettes in less strict anaerobic conditions. With these experiments we concluded that Fe-S clusters can be successfully reconstituted in aerobic conditions using the peptide maquettes discussed. These findings were never discovered or described in the literature before, to the best of our knowledge. To achieve the reduced  $[4\text{Fe-4S}]^{1+}$  cluster, the more powerful reducing activity of photoactivated NADH was required, as DT did not yield the expected results. The presence of 10%  $\beta$ ME proved crucial for the successful reconstitution of the cluster in aerobic conditions. We believe that under aerobic conditions the formation of disulphide bonds is boosted, with a detrimental effect on the presence of reduced thiols that can coordinate the iron ions of the cluster. For this reason, the presence of a reducing agent is required to ensure

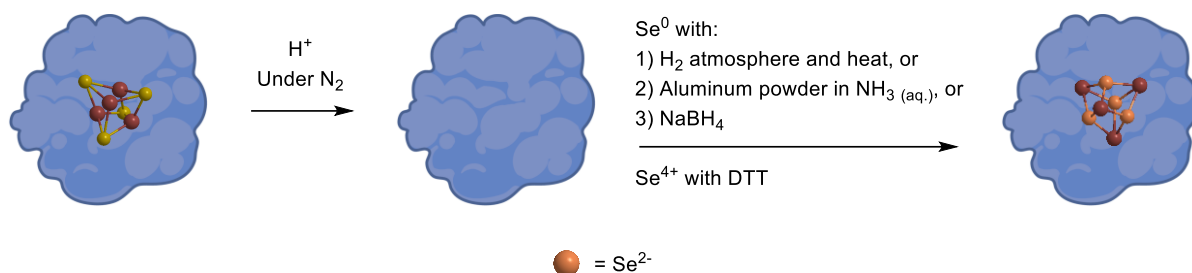


all the Cys residues in solution are kept in the reduced state. Overall, these results showcase the stability of Fe-S cluster maquettes to oxidizing conditions and provide an example of how these systems evolved to survive in ambient atmosphere.

## 2.2.6 Exchange of Sulfur by Selenium in [4Fe-4S] Cluster Maquettes

### 2.2.6.1 *Synthesis and Characterisation of a Selenopeptide [4Fe-4S] Cluster Maquette*

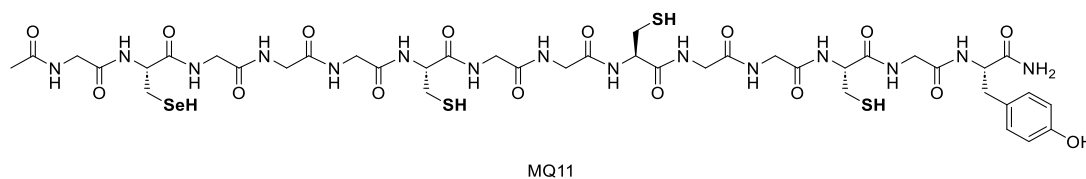
Here we describe the use of selenium as an alternative to sulfur in the reconstitution of cluster maquettes. The abundance of selenium in the Earth's crust is much lower with respect to that of sulfur. The estimated abundance for sulfur is 0.035 %, compared to  $5 \times 10^{-6}$  % for selenium; 7000 times lower.<sup>90</sup> The difference in abundance is perhaps the reason why life, and in particular bacteria governed by Fe-S clusters, adapted to the use of sulfur as a mediator of electrons rather than selenium. Selenium is indeed crucial for the constitution of selenoproteins, i.e. proteins where Cys is replaced by Sec, that governs crucial roles in our body such as DNA synthesis, reproduction, hormone metabolism, and anti-oxidative protection from free radicals and oxidative agents.<sup>91</sup> Another reason why selenium has not been deeply studied in iron cluster formation, is its increased reactivity with respect to sulfur, caused by the greater polarity and therefore lower strength of the C-, N-, and O-Se bonds compared to those formed with sulfur.<sup>92</sup> In the literature, examples can be found where Cys or Met have been replaced by Sec in proteins and small peptides, mainly for the augmented properties of selenium in acting as a probe.<sup>93</sup> Selenium has an atomic mass number higher than that of sulfur (79 and 32 respectively), which produces a considerable change in the vibrational properties of the modified proteins and peptides, making the Se-containing molecules more suitable for a range of spectroscopic techniques including EPR, Raman and X-ray absorption spectroscopy.<sup>93</sup> In this case, the conversion from Cys or Met to Sec is achieved by either chemical synthesis or chemical modification of the protein. In other examples, selenium is used instead of sulfur for the reconstitution of Fe-Se clusters. This process involves treatment of the cluster-protein species under inert conditions with an acid (either hydrochloric or trichloroacetic acid) to allow the inorganic sulfur of the cluster to escape as H<sub>2</sub>S.<sup>92</sup> The resulting apoprotein is then treated with either reduced elemental selenium (Se<sup>0</sup>) or selenite (Se<sup>+4</sup>) to coordinate the Fe-Se cluster, as illustrated in Scheme 11.



*Scheme 11 – Exchange of  $\text{S}^{2-}$  by  $\text{Se}^{2-}$  in  $[\text{4Fe-4S}]$ -clusters.*

More recently, Buscagan *et al.*, observed the exchange of sulfur by selenium in a  $[\text{4Fe-4S}]$  nitrogenase protein when this was treated with potassium selenocyanate.<sup>94</sup> This Fe nitrogenase protein governs the ATP mediated electron transfer to the nitrogenase molybdenum-iron (MoFe) protein during nitrogen fixation. Nitrogenases are enzyme complexes produced by bacteria to convert atmospheric nitrogen to ammonia, which is a more useful source of life.<sup>95</sup> Most nitrogenase complexes are composed of a MoFe protein that is thought to contain the active site where  $\text{N}_2$  is converted into  $\text{NH}_3$ , and a Fe-S protein that catalyse the transformation by transporting electrons. Because of the chalcogenide exchange observed, the Fe-S protein is now considered to provide an active site where small molecules reacting with the nitrogenase could be directly bound. These finding could lead to study new molecules that successfully produce ammonia from atmospheric  $\text{N}_2$ , with several applications in everyday life. Our efforts in this chapter were aimed at understanding how selenium could be used as an alternative for sulfur in  $[\text{4Fe-4S}]$ -cluster maquettes.

We firstly investigated the synthesis of a peptide where a Cys residue was exchanged with Sec. The synthesis of the Sec building block is described in 8.4 COMPOUNDS SYNTHESIS. The sequence was inspired from the poly-Gly **MQ7**, where the Cys closer to the *N*-terminus was replaced with Sec to achieve Ac-GUGGGCGGCGGY- $\text{NH}_2$  (**MQ11**) (Figure 55).



*Figure 55 – Peptide sequence of **MQ11**.*

This synthetic peptide only had one Cys residue replaced with Sec, but we speculated that it could provide insightful information in whether a [4Fe-4S]-cluster could be incorporated by this sequence. The reconstitution of the cluster was performed using DTT, and the detailed procedure is described in *8.2 GENERAL PROTOCOLS*. Briefly, DTT was used as a reducing agent to ensure both the thiol and the selenol groups would remain in the reduced state, similarly to what is achieved in presence of  $\beta$ ME. Figure 56 presents the UV-Vis spectroscopy analysis obtained when the oxidised and reduced samples were analysed.

These results clearly show the presence of a peak at approximately 420 nm for the oxidised peak, which is turned into a flat line when the sample is exposed to the reducing agent DT. The absorbance observed for the reduced sample from 240 to approximately 370 nm saturated the spectrophotometer, possibly because of the excess DT in solution. However, this did not affect the results of the analysis as the area around 400 and 420 nm was clear and showed the expected signal. The faint absorption peaks detected at approximately 300 nm in the oxidised sample could instead be assigned to the Tyr conjugated system of the synthetic peptide. Despite the mechanism of action for DTT being different to that of  $\beta$ ME (Scheme 12), both present a thiol moiety that, as previously discussed, could coordinate the excess  $\text{Fe}^{3+}$  in solution to produce a misleading signal.<sup>96</sup>

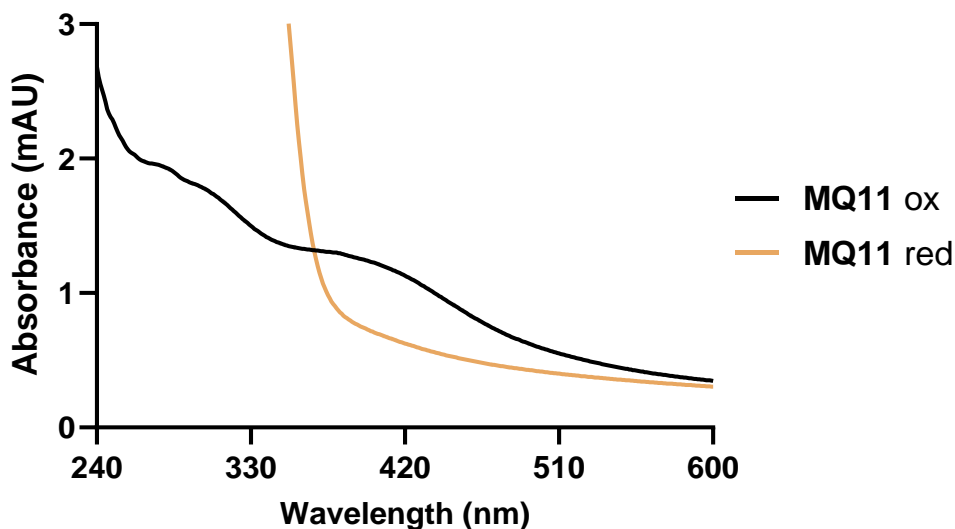
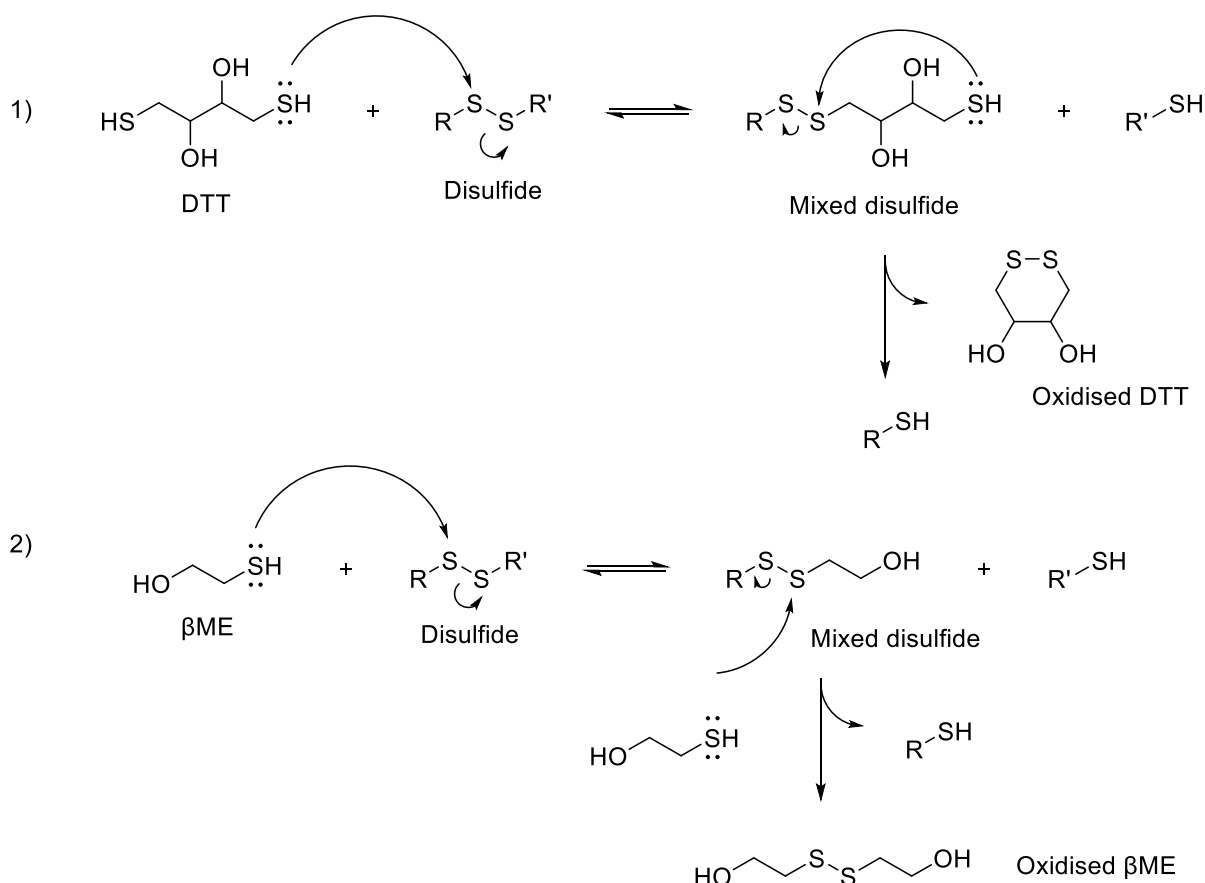


Figure 56 – UV-Vis trace for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ11**. Abs. at 420 nm represents the LMCT band for the cluster. The spectrum was recorded at peptide concentration of 80  $\mu$ M. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.



Scheme 12 – Mechanism for the reduction of disulphide bonds with 1) DTT and 2)  $\beta$ ME.

For this reason, we proceeded by running EPR spectroscopy analysis (Figure 57), to confirm the successful incorporation of a cluster using the selenopeptide **MQ11**.

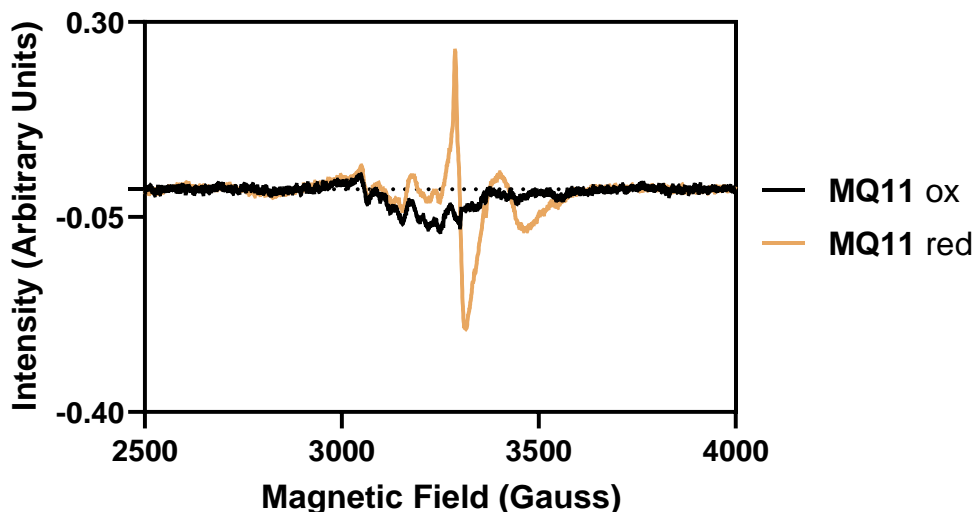


Figure 57 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in presence of DTT) **MQ11** oxidised (black trace) and dithionite-reduced (gold trace). EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

Despite low in intensity, a signal was observed in the reduced sample (gold trace) with a peak at approximately 3500 G, characteristic of a reduced  $[4\text{Fe-4S}]^{1+}$  cluster. The sample prior treatment with DT (black trace) only presented the background noise for the buffer solution. These results confirmed that a  $[4\text{Fe-4S}]$  cluster could be successfully reconstituted even if one of the Cys residues along the peptide sequence is replaced by a Sec residue. Unfortunately, due to lack of time, further exploration of Sec-containing peptides could not be investigated. However, these results pose the basis for future studies involving the use of selenopeptides for the incorporation of  $[4\text{Fe-4S}]$ -clusters and the examination of this new class of cluster maquettes.

#### 2.2.6.2 Attempt on the Reconstitution of a $[4\text{Fe-4Se}]$ Cluster Maquette

Finally, we investigated the use of selenium as a replacement for sulfur in the formation of a cluster to be coordinated by a Cys-containing peptide. The idea was to use sodium selenite ( $\text{Na}_2\text{SeO}_3$ ) instead of sodium sulphide ( $\text{Na}_2\text{S}$ ) for the construction of a  $[4\text{Fe-4Se}]$  cluster, in which the Fe ions would be coordinated by the Cys residue of the peptide. For doing so, we selected the engineered poly-Gly **MQ8** that was inspired by the bacterial ferredoxin family and therefore bears four Cys residues capable of coordinating all iron ions of the cluster. The reconstitution procedure was slightly

adjusted from previous literature, and a detailed method can be found in 8.2 *GENERAL PROTOCOLS*.<sup>92</sup> In brief, the peptide was equilibrated in HEPES and KCl buffer solution at pH = 8.0 with 50-fold excess of DTT. FeCl<sub>3</sub> dissolved in buffer was added followed by the source of Se<sup>2-</sup>; both were used in 32-fold excess with respect to the synthetic peptide. The selenium source was separately prepared by dissolving and mixing for five minutes a solution of Na<sub>2</sub>SeO<sub>3</sub> in buffer containing DTT. The resulting mixture was left stirring for one hour before being analysed by UV-vis and EPR spectroscopy.

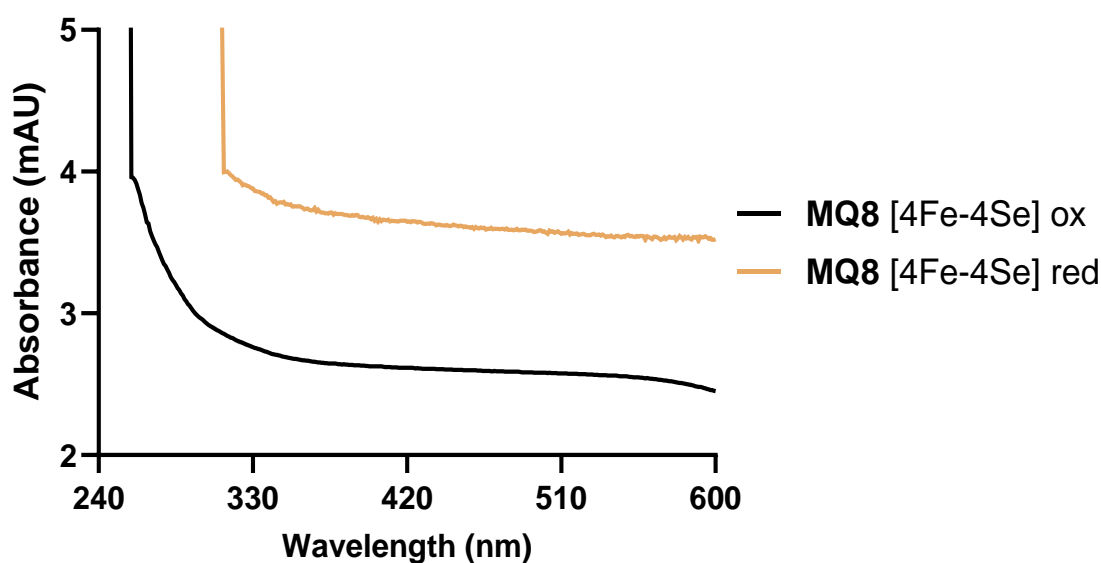


Figure 58 – UV-Vis trace for [4Fe-4S]-cluster maquette reconstituted in anaerobic conditions using synthetic peptide **MQ8**. The spectrum was recorded at peptide concentration of 80  $\mu$ M. The peptide was obtained as a TFA salt after purification and the buffer is 50mM HEPES, 10mM KCl at pH = 8.0.

Unfortunately, no LMCT interaction could be observed in the UV-Vis analysis (Figure 58). The oxidised sample (black trace) presented a flat line in proximity of the 420 nm region, meaning that no interaction occurred between the Cys residues of the synthetic peptide and the iron ions of the Fe-Se cluster. This result suggested that the [4Fe-4Se]-cluster species was not successfully obtained. The sample treated with DT (gold trace) showed an incremented absorbance, and a flat line across the range between 330 and 600 nm. Overall, the absorbance observed for both the oxidised and reduced samples was higher than usual, possibly suggesting the presence of particulate in solution that increased the absorbance due to scattering of the radiation.<sup>97</sup> This hypothesis was supported by the values registered in the range between 240 and 330 nm, where the values of absorbance saturated the detector and therefore masked the

sample trend. The results described in Figure 58 cannot be trusted due to the high absorbance values obtained and the high value of the baseline. Despite the negative results obtained with UV-Vis spectroscopy, the sample was subjected to EPR analysis to gather more information about its nature (Figure 63).

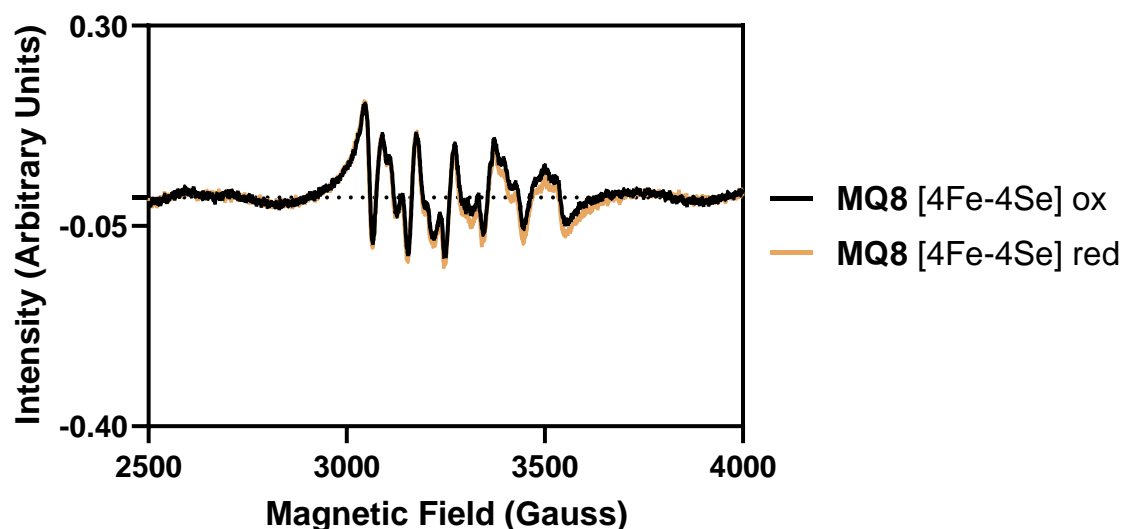


Figure 59 – X-band cw-EPR spectra at 20 K of a frozen solution of degassed (anaerobically reconstituted in presence of DTT) **MQ8** oxidised (black trace) and dithionite-reduced (gold trace).  $Se^{2-}$  was used instead of  $S^{2-}$  for the cluster constitution. EPR spectra were recorded at 5 G MA, with microwave power of 20 dB (~2.2 mW). EPR samples contained glycerol 10% v/v.

Unfortunately, EPR spectroscopy confirmed the absence of a  $[4Fe-4Se]^{1+}$  species in the sample exposed to DT (gold trace). The spectrum recorded assumed the shape of a multiplet and was possibly due to manganese presented in the iron source, as previously discussed.

## 2.3 CONCLUSION AND FUTURE WORK

To conclude, both the radical SAM and bacterial ferredoxin families of enzyme have been investigated to engineer synthetic enzyme sequences and successfully incorporate a reduced  $[4Fe-4S]$  cluster. The cluster has been characterised using both UV-vis and EPR spectroscopy. Modification of standard reconstitution procedure for cluster maquettes has allowed the successful incorporation of a cluster using a bacterial ferredoxin-inspired peptide in absence of the reducing agent  $\beta$ ME. This novel procedure, despite yielding lower concentration of cluster maquette, could be of use

in applications where reactive additives are not tolerated. Moreover, the absence of  $\beta$ ME consents the use of UV-Vis spectroscopy as a reliable tool for the characterisation of cluster maquettes.

Attempts aimed at modifying the radical SAM-inspired peptide sequence by changing both intervening and flanking amino acids did not produce improved results compared to the original peptide sequence, therefore further exploration in this area should be undertaken. Moreover, the use of SAM in place of  $\beta$ ME for the reconstitution of the cluster could be explored. The use of SAM would allow for the enzyme moiety to be mimicked, possibly allowing for the formation of the active radical to catalyse chemical modifications. These modifications could initially target the side chain of amino acid residues on the peptide sequence, but also other functionalities linked to the peptide.

Formation and reactivity of the cluster maquette was then explored in presence of oxygen. Aerobically reconstituted samples were treated with NADH and photoactivated at 365 nm to show the presence of a reduced  $[4\text{Fe-4S}]^{1+}$  cluster by EPR spectroscopy. Identical samples treated with DT as the reducing agent did not provide the same result, showcasing the ability of photoactivated NADH as a powerful reducing agent. Unfortunately, observation of the reduced cluster was only possible when the sample was reconstituted in presence of  $\beta$ ME. These findings allow for the cluster maquette to be reconstituted in non-strict anaerobic conditions, broadening the potential applications. However, a deeper investigation is required to understand the reasons behind the mandatory presence of  $\beta$ ME in the reaction mixture to allow for the successful formation of reduced cluster.

Finally, one of the Cys residues was replaced by the amino acid Sec in a bacterial ferredoxin-inspired maquette. The peptide sequence obtained was implemented in the reconstitution of a  $[4\text{Fe-4S}]$ -cluster; and EPR spectroscopy confirmed the presence of a reduced Fe-S species. Future studies should be aimed at substituting all the four Cys residues with Sec to explore the effect on cluster reconstitution and EPR signal. Unfortunately, the use of selenium as an alternative to sulfur for the cluster reconstitution did not produce successful results. Efforts should be aimed at optimising this synthetic methodology. Furthermore, a  $[4\text{Fe-4Se}]$  cluster could be incorporated by a Sec-containing peptide, to investigate a novel species with potentially different spectroscopic characteristics.



## 3 REDOX PROPERTIES OF [4Fe-4S] CLUSTER MAQUETTES

### 3.1 INTRODUCTION

Bacterial ferredoxin proteins can be classified into two different families based on their redox potential values. The first one is the HiPIPs family, which consist of small proteins with a reduction potential that ranges between -0.14 and 0.16 V vs saturated calomel electrode (SCE).<sup>82</sup> Ferredoxin proteins are mostly present in photosynthetic bacteria where, thanks to their high redox potentials, they are able to serve as an electron donor and acceptor for tetraheme cytochrome.<sup>98</sup> The high redox potential of HiPIPs is thought to derive from the conformation in which the [4Fe-4S] cluster is located. This is indeed coordinated by the four Cys residues of the protein surrounded by hydrophobic amino acids. This hydrophobic pouch allows for the cluster to be inaccessible to the solvent and results in a higher redox potential for the interchange between [4Fe-4S]<sup>+2</sup> and [4Fe-4S]<sup>+1/+3</sup>.<sup>7</sup> HiPIPs are opposed to low potential bacterial ferredoxins, which present a redox potential that spans from -0.49 to -0.89 mV vs SCE. This class of ferredoxins is normally found in anaerobic bacteria where they conjugate substrate oxidation with the reduction of FAD, nicotinamide adenine dinucleotide phosphate (NADP<sup>+</sup>) and other coenzymes for the correct function of living organisms.<sup>7</sup>

### 3.2 AIMS AND OBJECTIVES

In this chapter we describe the analysis carried out to determine the redox potentials of three different [4Fe-4S]-cluster maquettes (Figure 60). This small variety of maquettes includes two peptide sequences bearing four Cys residues (**MQ4**: Ac-KLCEGGCIACGACGGW-NH<sub>2</sub>, and **MQ8**: Ac-GCGGGCGGCGGY-NH<sub>2</sub>); and one peptide sequence that presents three Cys residues and one Sec residue (**MQ11**: Ac-GUGGGCGGCGGY-NH<sub>2</sub>). The [4Fe-4S]-cluster was incorporated into three different synthetic peptides to investigate how the environment provided by the different amino acids can alter the redox activity of the cluster. Different amino acids in close proximity to the Cys residues involved in the coordination of the cluster could indeed provide both steric and electronic effects that are expected to influence the redox potentials of the maquette. The aim of these experiments was to determine the

accurate redox potential at which the cluster maquettes are oxidised and reduced. The information would then be used to find a biological pathway in which to use the cluster maquette as either a reducing or oxidising agent. Ultimately, this would prove the utility of cluster maquettes as electron transport agents in bio-catalysis. The aim of this chapter is not just to determine the redox potentials of these cluster maquettes to gather more information about the ability of these miniaturised system to model proteins. We intent to probe the tuning of the redox potentials of these clusters to build synthetic and semi-synthetic electron transfer systems to be employed in synthetic biology and biotechnology.

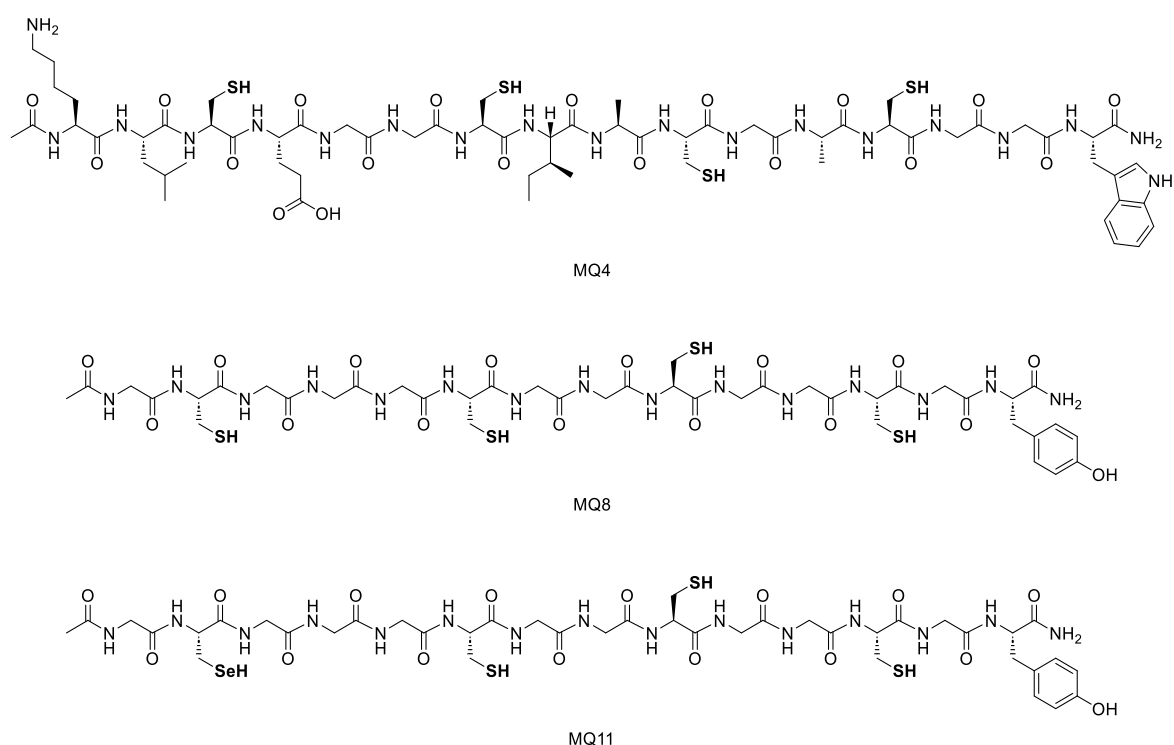


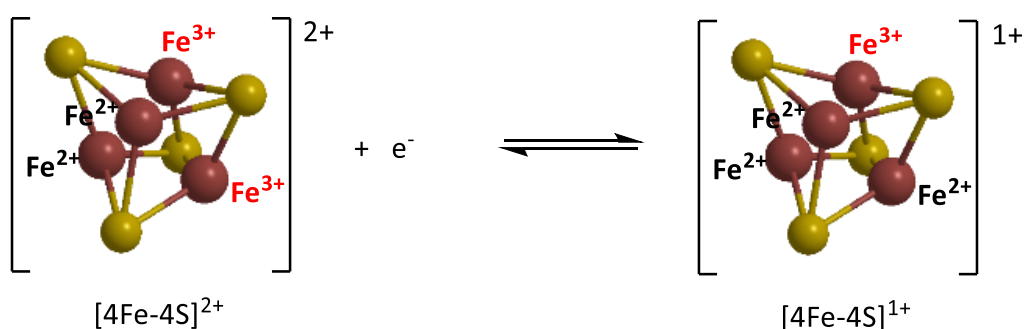
Figure 60 – Peptide sequence of **MQ4** (top), **MQ8** (middle), and **MQ11** (bottom) used for the cyclic voltammetry experiments.

### 3.3 ELECTROCHEMICAL PRINCIPLES OF [4Fe-4S] CLUSTERS

Electrochemical analysis is implemented to gather information regarding the reduction and oxidation of redox active species, in this case the Fe-S cluster coordinated by the synthetic peptide. The redox ability of the cluster maquette can be used to catalyse a suitable biological pathway that requires either a reducing or an oxidising agent. Furthermore, determination of the accurate potential at which the cluster maquette is

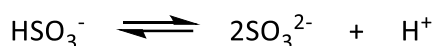
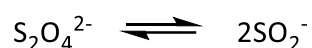
reduced and oxidised is valuable information to investigate the tunability of this biocatalyst. Ideally, the redox potential could be shifted to higher or lower potentials by changing the peptide sequence, while keeping the consistent Cys-rich motif required for the coordination of the cluster.

The main characteristic of Fe-S clusters that made them so attractive for enzymes when they firstly started to spread in prebiotic life was their ability to exchange electrons. This attribute quickly became essential for enzymes that are required to catalyse biochemical pathways involving electron transfer reactions. Electrochemistry is a key tool when it comes to measuring the flow of electrons in such chemical systems. In [4Fe4S]-cluster maquettes, the electron exchange takes place as a redox process involving the metal complex formed by the interaction between the four iron atoms and the four sulphide ions. When reconstituted, [4Fe-4S] clusters are in the oxidised +2 state with two iron ions in the +3 oxidation state and two iron ions in the +2 oxidation state. When the cluster is reduced, only one iron ion is in the +3 oxidation state, whereas three iron ions are now in the +2 oxidation state (Scheme 13).



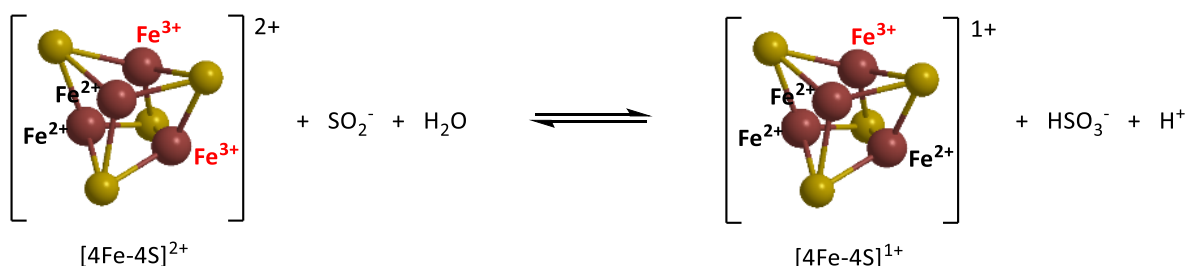
*Scheme 13 – Reduction of a [4Fe-4S] cluster. Iron atoms in red and sulfur atoms in yellow. The oxidation state for the iron atoms is showed in red (+3) and black (+2).*

When using a reducing agent to reduce the [4Fe-4S]-cluster in a chemical system, the source of electrons is a chemical species, such as DT. DT is a popular reducing agent in biochemistry and has a potential of approximately -0.9 V when measured at pH = 7.0 and compared to SCE.<sup>82</sup> The reduction carried out by DT is thought to happen via formation of the radical species  $\text{SO}_2^-$  has suggested from Equation 1.



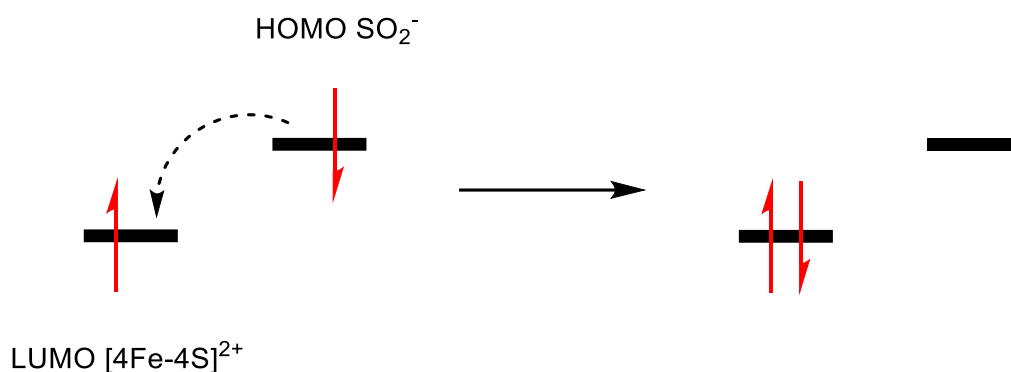
Equation 1 – Equations describing the dissociation of the dithionite anion ( $\text{S}_2\text{O}_4^{2-}$ ).

When in water or in a buffer at close-to-neutral pH, DT dissolves into  $2\text{Na}^+$  and the dithionite dianion  $\text{S}_2\text{O}_4^{2-}$ . The latter undergoes monomerisation to give the radical species  $\text{SO}_2^-$ , which is the species thought to act as the actual reducing agent when DT is used in a chemical reduction.<sup>99</sup> The sulfur dioxide radical is subjected to a one-electron oxidation to give the bisulphite  $\text{HSO}_3^-$ . This one-electron oxidation step fits with the reduction of the Fe-S cluster of the maquette in the chemical reduction process showed in Scheme 14. The sulfur atom in this step is oxidised from the +3 oxidation state in  $\text{SO}_2^-$  to the +4 oxidation state in  $\text{HSO}_3^-$ . The final step involves the ionisation of the bisulphite ion to  $\text{SO}_3^{2-}$  at a pH of approximately 6.9.<sup>82</sup>



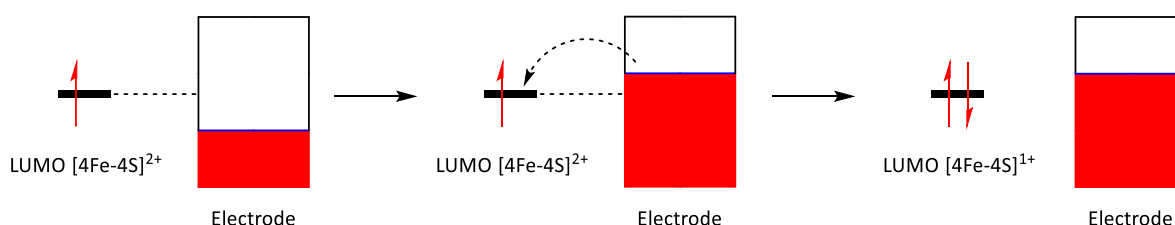
Scheme 14 – Chemical reduction of a  $[4\text{Fe-4S}]$  cluster. Iron atoms in red and sulfur atoms in yellow.

In this redox reaction, one electron is transferred from DT to the Fe-S cluster. The reason why this happens is because the lowest unoccupied molecular orbital (LUMO) of the oxidised cluster is at a lower energy than the electron in the highest occupied molecular orbital (HOMO) of DT (Scheme 15).<sup>100</sup> The driving force of this process is therefore the energy associated with the electron that is being transferred from one species to another, and more in particular in the energy difference between the respective HOMO and LUMO.



*Scheme 15 – Electron transfer from the HOMO of  $\text{SO}_2^-$  to the LUMO of the oxidised  $[\text{4Fe-4S}]^{2+}$  cluster.*

When an electrode and a potentiostat are used instead of a reducing agent in an electrochemical reaction, the electrons are transferred from the electrical conductor to the cluster. The energy of the electrons on the surface of the electrode can be increased or decreased by modulating the voltage applied to the electrode from the potentiostat. This means that the energy associated to the electrons is not fixed, in contrast to the energy of the electron in the HOMO during the electrical reduction process. Once the energy of the electrons on the electrode is higher than the energy level of the LUMO of the oxidised cluster, one electron is transferred to the HOMO, resulting in the cluster being reduced (Scheme 16).



*Scheme 16 – Electron transfer from the electrode to the LUMO of the oxidised  $[\text{4Fe-4S}]^{2+}$  cluster.*

While a chemical reduction of the oxidised cluster can be performed by simply adding DT, or another reducing agent, to a solution of the cluster; the electrochemical version of this redox process requires an electrochemical cell to be set up. Figure 61 shows the electrochemical cell setup used to collect the data in our experiments.

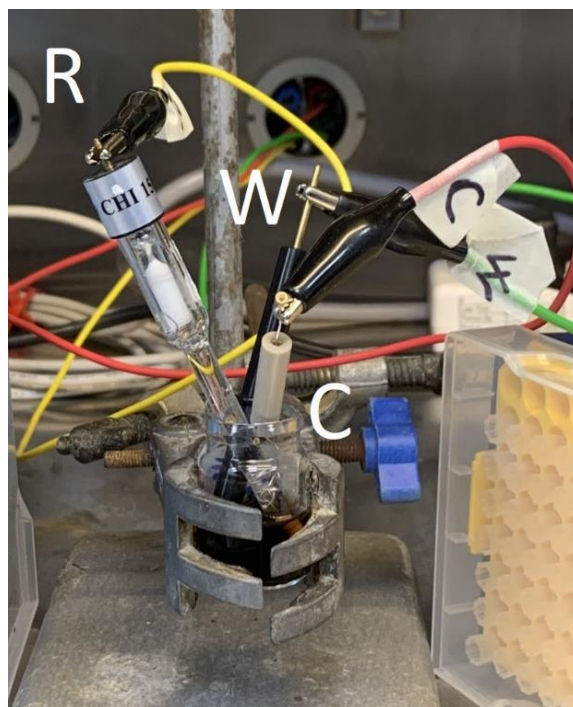
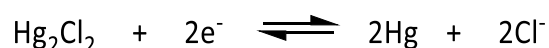


Figure 61 – Electrochemical cell setup. The working electrode (**W**) is a glassy carbon electrode, the counter electrode (**C**) is a platinum wire, and the reference electrode (**R**) is a SCE.

The electrochemical cell always consists of a working electrode, which is the component on the surface of which the electrochemical event happens. The working electrode is directly connected to the potentiostat so that the potential applied to the electrical conductor can easily be adjusted. The counter electrode is required to close the circuit and therefore accept the current flowing from the working electrode and through the solution where the redox active analyte is dissolved. Because the current is registered as the electrons flow from the working to the counter electrode, when the redox active species absorbs or donate electrons these can be measured as a change in the flow and therefore produce a voltammogram. Finally, the reference electrode is used as a reference against which the potential of another electrode can be measured. The reference electrode has a stable potential that is measured based on a known redox reaction, and for our experiments we used SCE. Its reaction is reported in Equation 2 and involves elemental mercury and Hg(I) chloride.



Equation 2 – Equation describing the redox reaction in a SCE reference electrode.

The equilibrium in this reaction is between Hg(I) chloride and elemental mercury. Because mercury has an oxidation state of +1 in the salt form, it requires one electron to be reduced to the elemental state, where it has an oxidation state of 0. The chloride

salt of Hg(I) is also known as “calomel”, hence the name for the reference electrode based on this reaction.<sup>101</sup>

Before being analysed, all samples were supplemented with a 100 mM concentration of NaCl as a supporting electrolyte. Supporting electrolytes are necessary to increase the solution conductivity and reduce the resistance. Once the adequate potential is reached for electrons migration from the electrode, the supporting electrolyte migrate to balance the charge and close the circuit. The supporting electrolyte keeps the ionic strength of the solution to a high level and ensure that the electric field remains homogeneous and does not get perturbed by the reduction or oxidation of the analytes.<sup>98</sup>

## 3.4 RESULTS AND DISCUSSION

### 3.4.1 Buffer and Controls

All analysis were performed following the international union of pure and applied chemistry (IUPAC) convention and moving from high potentials to low potentials. The reason why we decided to start at high potential and move towards lower values was due to the oxidation state of the reconstituted cluster maquette. When synthesised, cluster maquettes are in their oxidised (+2) state, which can be converted to the reduced (+1) state.

The procedure we followed for the synthesis of [4Fe-4s]-cluster maquettes involved the addition of FeCl<sub>3</sub> and Na<sub>2</sub>S to a buffer solution where the synthetic peptide was dissolved. In order to identify the characteristic reduction and oxidation peaks for the cluster maquette, the Fe<sup>3+</sup> and S<sup>2-</sup> sources, as well as the buffer solution, were firstly analysed. As for any spectroscopic measurements, it is important to determine whether the buffer or any other reagent used in the reaction mixture contributes to the overall absorption of the sample to accurately assign each peak. Figure 62 shows the measurement obtained for the 50 mM HEPES and 10 mM KCl buffer solution at pH = 8.0.

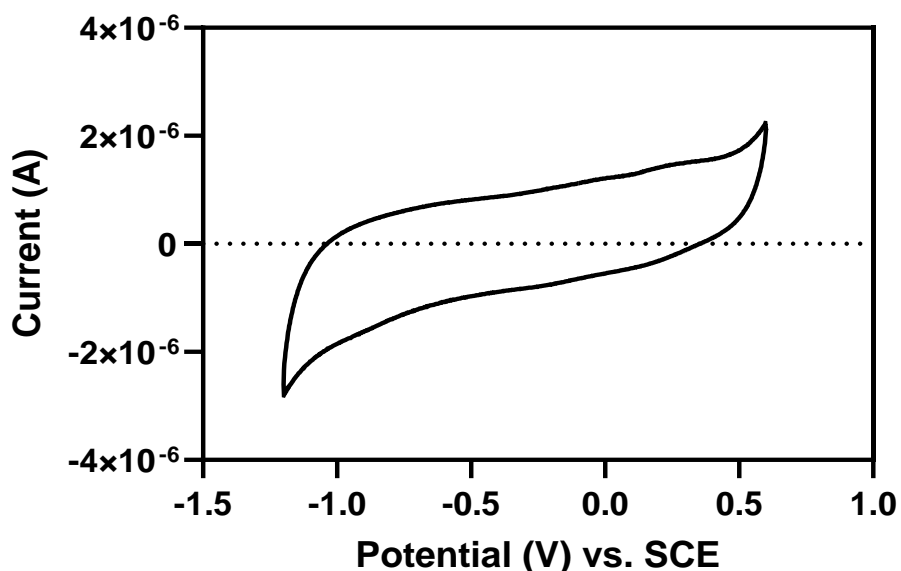


Figure 62 – Cyclic voltammogram for the buffer solution 50 mM HEPES 10 mM KCl at pH = 8.0.

As expected, no peaks were detected for the buffer solution, meaning that no redox active species were identified. Measuring the redox profile of the buffer solution not only informed us that no redox active species were present in solution, but it also gave us an idea of what the potential window of this buffer is in the specific conditions that we used to measure it. In this case, the redox potential window for this buffer was determined between -1.2 and 0.6 V. Every solvent has a specific potential window in which no redox activity is expected to happen and in which the solvent can therefore be considered stable. The solvent window can be affected by different variables, including the supporting electrolyte used and the nature of the working electrode. Water, for example, normally has a narrow electroactive window, ranging from  $\sim -0.7$  V to  $\sim 0.7$  V when at pH = 7.0 and determined using a SCE as the reference electrode.<sup>102</sup> What happens to the solvent at lower and higher potentials than its electroactive window, is that it begins to be reduced and oxidised. Because the solvent is in much higher concentration with respect to any other redox active species dissolved into it, the reduction and oxidation events of the solvent completely mask any other activity and therefore make it impossible to observe the redox profile of the analytes.



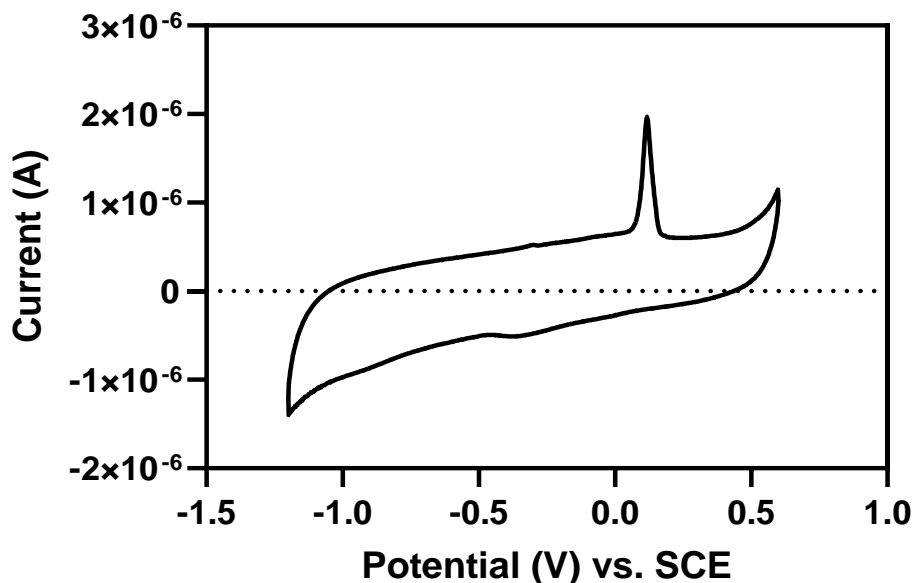


Figure 63 – Cyclic voltammogram for  $\text{FeCl}_3$  in 50 mM HEPES 10 mM KCl at  $\text{pH} = 8.0$ .

The sample containing  $\text{FeCl}_3$  displayed a tiny reduction peak at  $-0.45$  V, and an oxidation peak at  $0.12$  V, as shown in Figure 63. The most common oxidation states for iron are +2 and +3. Because iron is in the +3 oxidation state in  $\text{FeCl}_3$ , the cyclic voltammetry measurement was expected to only show a reduction peak for its reduction to  $\text{Fe}^{2+}$ . However, only a small reduction peak was detected for this sample at  $-0.45$  V, whereas a much more intense peak was registered for an oxidation event at  $0.12$  V. This could be referred to the oxidation of  $\text{Fe}^{3+}$  to  $\text{Fe}^{4+}$ . It is possible that the range in which the analysis was performed, was not wide enough to cover the full redox cycle for  $\text{FeCl}_3$ , meaning that only its oxidation from +3 to +4 was observed. This scenario could happen because the portion of the sample that is affected by the measurement is only that in the area surrounding the electrode, and fresh  $\text{Fe}^{+3}$  would be available at every new cycle to be reduced. To investigate the nature of this peaks and to understand whether they were linked between each other, we decided to shorten the redox potential from  $-1.2$  to  $0.6$  V to  $0$  to  $0.6$  V, as showed in Figure 64.

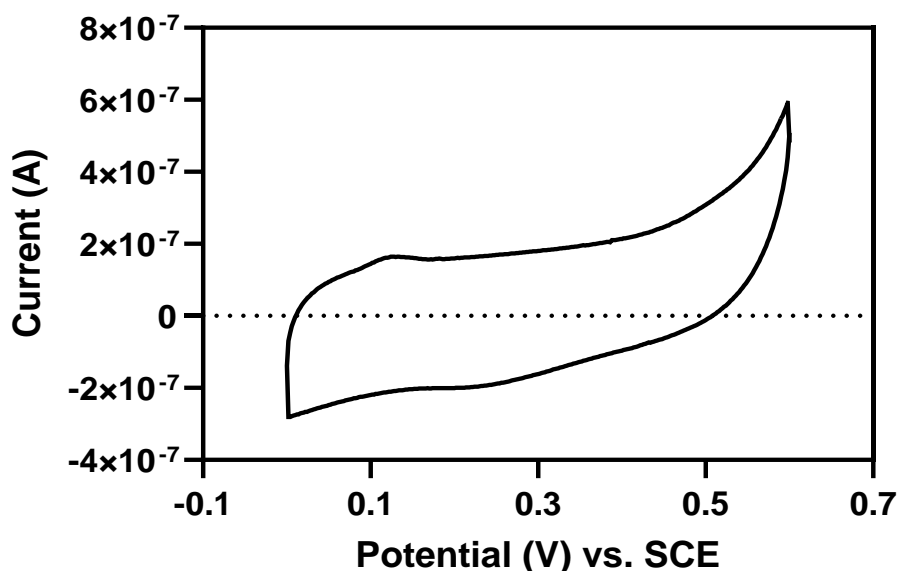


Figure 64 – Cyclic voltammogram for  $\text{FeCl}_3$  in 50 mM HEPES 10 mM KCl at  $\text{pH} = 8.0$ .

In this way we were able to avoid the reduction event at  $-0.45$  V, which caused the oxidation peak observed at  $0.12$  V to disappear too. We therefore concluded that the two redox events were related to each other. It is possible that some impurities presented in the  $\text{FeCl}_3$  batch were responsible for the redox activity that we observed, but that the potential window provided by the buffer we used was not wide enough to cover any of the reduction or oxidation related to the  $\text{Fe}^{3+}$  ions.

The cyclic voltammetry profile for  $\text{Na}_2\text{S}$  showed a reduction and an oxidation peak respectively at  $-0.774$  V and  $-0.068$  V, as displayed in Figure 65, and was comparable to that of previous literature.<sup>103</sup>

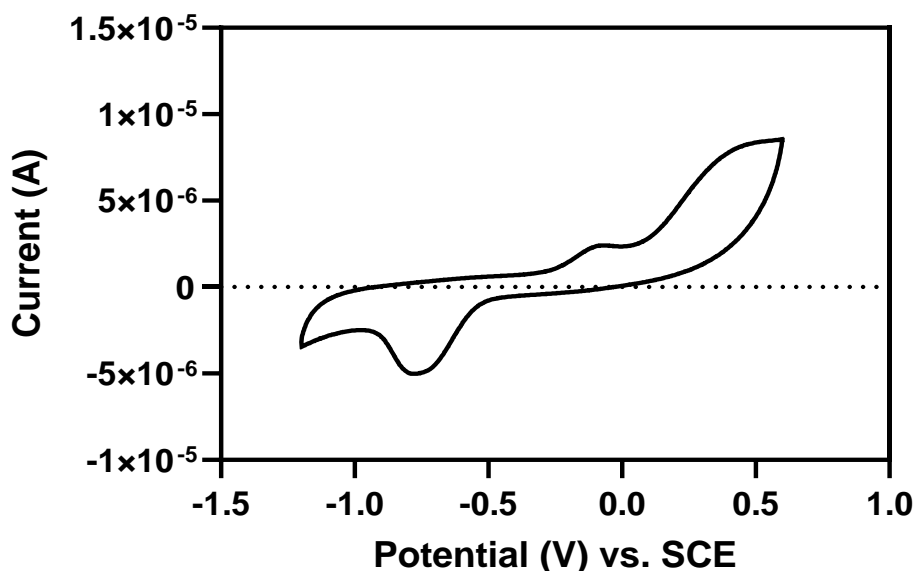


Figure 65 – Cyclic voltammogram for  $\text{Na}_2\text{S}$  in 50 mM HEPES 10 mM KCl at  $\text{pH} = 8.0$ .

The reduction peak was not expected to be observed because sulphides, in which sulfur has an oxidation state of -2, normally represents the most reduced state for this element. It is plausible that due to the absence of reducing agents in solution (i.e.,  $\beta\text{ME}$  or DTT), a portion of the  $\text{S}^{2-}$  oxidised to the -1 oxidation state and was therefore reduced during the cyclic voltammetry measurement. On the positive potential, the oxidation peak was expected to be registered and it refers to the oxidation of  $\text{S}^{-2}$  to  $\text{S}^{-1}$ . Towards the end of the potential range, close to 0.5 V, it is possible to notice that the current value increases possibly to detect a second oxidation peak. However, this is not fully identified. Sulfur has other oxidation states apart from -1 and -2, however, because only one electron is associated with each peak, we can assume that the other oxidation states are not available in the potential range in which the measurement was performed. Furthermore, 0.5 V is also very close to the limit of the potential range given by the buffer, which means that the peak could also be related to oxidation of the solvent.

### 3.4.2 Cyclic Voltammetry of [4Fe-4S] Cluster Maquettes in absence of $\beta\text{ME}$

All the experiments presented here were run on samples reconstituted under anaerobic conditions with peptides synthesised using the Fmoc-SPPS strategy, as detailed in 8.2 GENERAL PROTOCOLS. Due to difficulties encountered in the

purification of the samples via SEC columns, these experiments were performed in absence of the reducing agent  $\beta$ ME, unless otherwise mentioned. The decision to reconstitute the samples in absence of the thiol was taken to ensure that no other chemicals, apart from those necessarily required for the cluster formation, would interfere with the analysis. The first maquette to be analysed was the one obtained using synthetic peptide Ac-KLCEGGCIACGACGGW-NH<sub>2</sub> (**MQ4**) in the coordination of the [4Fe-4S] cluster. The cyclic voltammety spectrum for this cluster-maquette is showed in Figure 66.

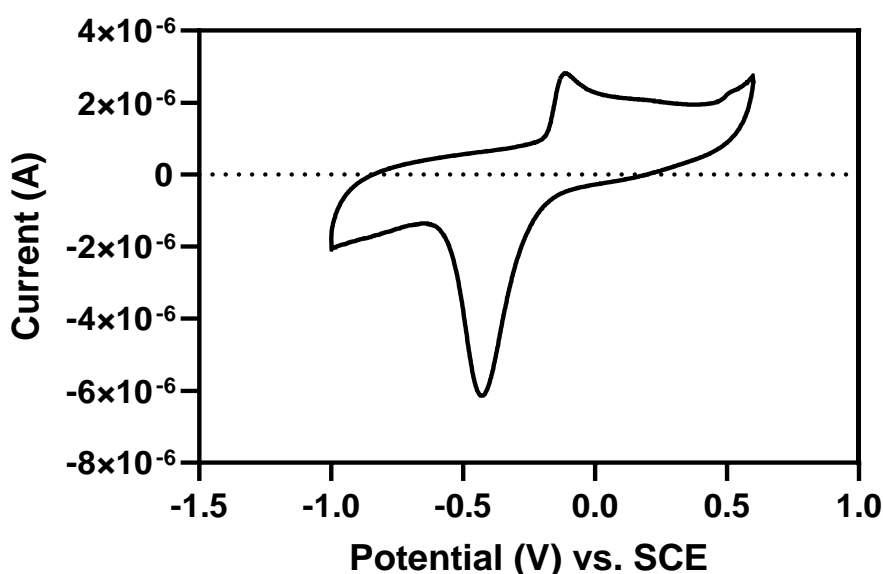


Figure 66 – Cyclic voltammogram for the [4Fe-4S] cluster coordinated by synthetic peptide **MQ4** in 50 mM HEPES 10 mM KCl at pH = 8.0. Sample reconstituted in absence of  $\beta$ ME.

The main features of this cyclic voltammogram consist in a single cathodic peak ( $E_{p,c}$ ) at -0.429 V vs SCE (-188 mV vs. standard hydrogen electrode (SHE)), and a single anodic peak ( $E_{p,a}$ ) at -0.112 V vs SCE. As described in Figure 67, the  $E_{p,c}$  is defined as the peak potential reached by the cathodic current ( $i_{p,c}$ ).<sup>104</sup> The  $E_{p,c}$  is achieved when all the sample at the surface of the electrode has successfully been reduced, and therefore corresponds to the reduction peak. On the contrary, when the potential is scanned from negative to positive values, the anodic current ( $i_{p,a}$ ) is produced. The peak reached by  $i_{p,a}$  is called  $E_{p,a}$ , and corresponds to the sample at the surface of the electrode being completely oxidised. In this case, the peak is called the oxidation peak.<sup>105</sup>

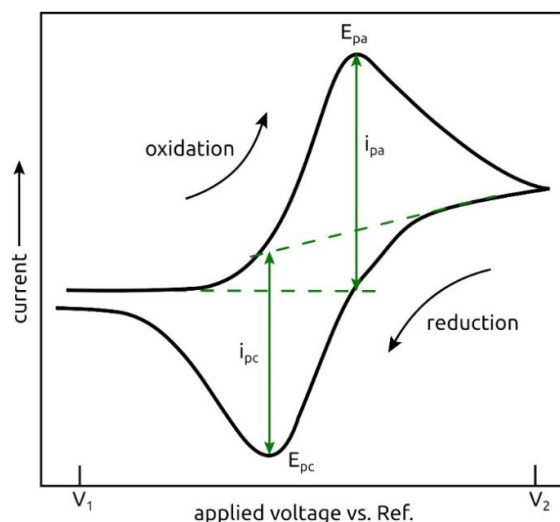


Figure 67 – Voltammogram of a single oxidation-reduction process.<sup>104</sup>

The redox potential was consistent to that observed in previous literature and identified the maquette in a range closer to that of HiPIPs. The peptide sequence used to incorporate this cluster was representative of a ferredoxin maquette that belongs to the *Peptococcus aerogens* bacteria.<sup>68</sup> The original ferredoxin belongs to the low potential family of enzymes, which contrasts with our experimental results. The reason why the reduction peak of this maquette does not fit the range of the protein the sequence is representative of, could be explained by the fact that taking a section of the sequence from the *Peptococcus Aerogens* ferredoxin could have influenced its redox potential. Despite the original ferredoxin is part of the low potential family, decreasing its length might have caused the redox potential to increase. It is worth mentioning that apart from Glu, all the other amino acid residues that constitutes the synthetic peptide are hydrophobic, and if the peptide sequence is long enough to fold around the cluster that it coordinates, this could help keeping the solvent away and increase the redox potential. To further prove that the reduction and oxidation events observed were characteristic of the anaerobically reconstituted sample, **MQ4** was used to reconstitute a sample in aerobic conditions. The cyclic voltammogram recorded for the aerobically assembled maquette is displayed in Figure 68.

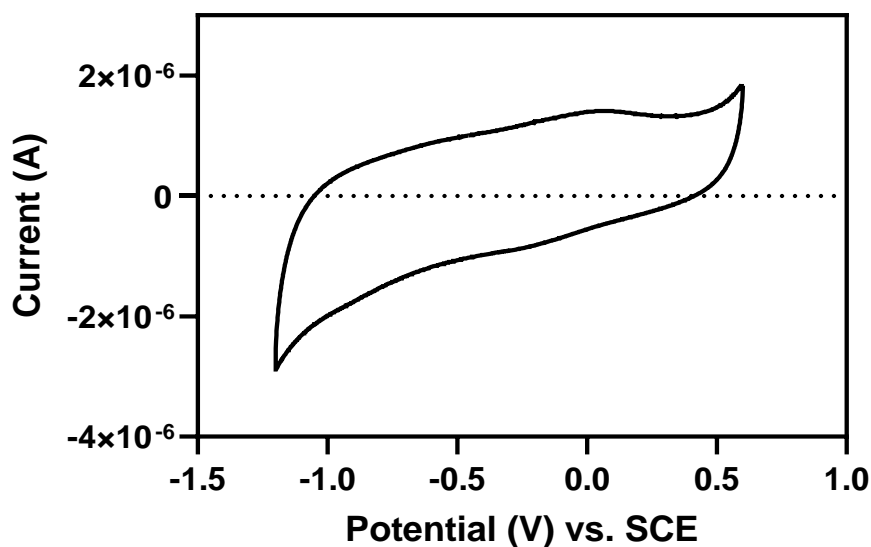


Figure 68 – Cyclic voltammogram for the aerobically reconstituted [4Fe-4S] cluster coordinated by synthetic peptide **MQ4** in 50 mM HEPES 10 mM KCl at pH = 8.0. Sample reconstituted in absence of  $\beta$ ME.

As expected, the voltammogram did not show the characteristic reduction and oxidation peaks observed for the maquette sample synthesised in anaerobic conditions.

The second cluster maquette that was analysed by cyclic voltammetry was the one incorporated using synthetic peptide Ac-GCGGGCGGCGGCGY-NH<sub>2</sub> (**MQ8**). This sequence is two amino acids (12.5 %) shorter with respect to the previous and is almost entirely composed by Gly residues. This peptide is shorter than the previous one and slightly less polar, with an isoelectric point of pH = 0 compared to pH = 5.8 for **MQ4**. For these reasons, the redox potential fingerprint was not expected to be dissimilar from that of the previous maquette. As showed in Figure 69, the shape of the spectra is almost identical, and so are the redox potentials for both the  $E_{p,c}$  and  $E_{p,a}$  peaks, with respectively -0.471 V vs SCE (-230 mV vs. SHE) and -0.108 V vs SCE.

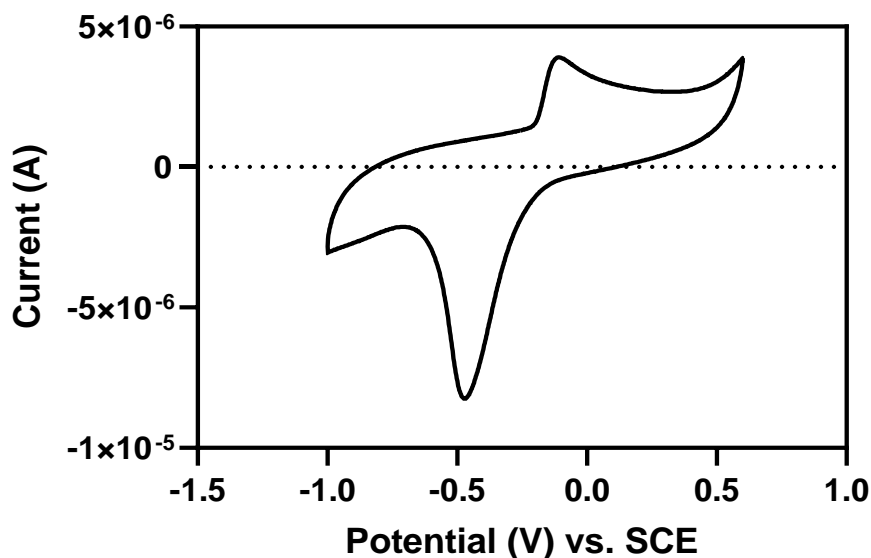


Figure 69 – Cyclic voltammogram for the [4Fe-4S] cluster coordinated by synthetic peptide **MQ8**. Sample reconstituted in absence of  $\beta$ ME.

The reduction peak came with a 42 mV difference with respect to that of the previous maquette, whereas a change of only 4 mV was observed for the oxidation peak. Overall, the shift happened towards lower redox potentials, and it is plausible that it was caused by the different environment and length of the new maquette. Again, to prove that these redox potential values were characteristic of the synthetic [4Fe-4S] cluster maquette, we performed cyclic voltammetry analysis on a sample that was reconstituted in aerobic conditions.

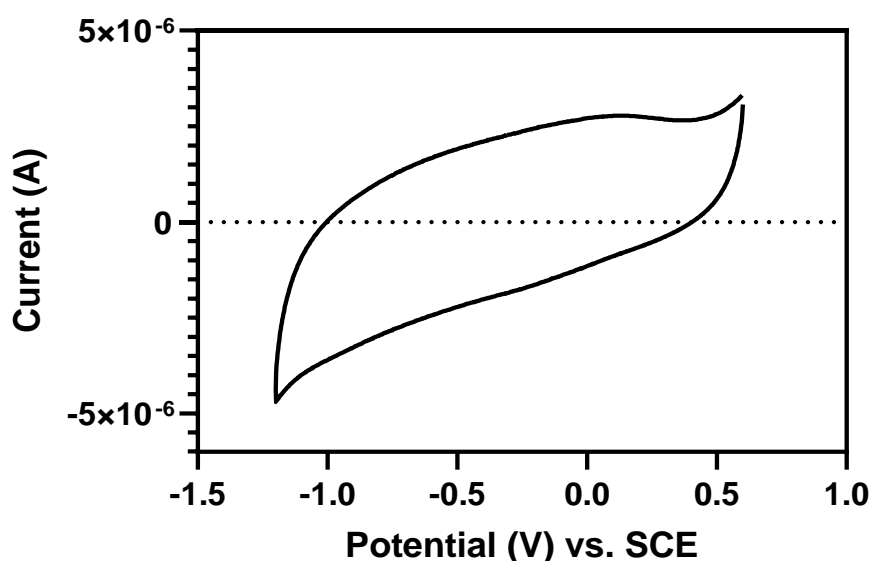


Figure 70 – Cyclic voltammogram for the aerobically reconstituted [4Fe-4S] cluster coordinated by synthetic peptide **MQ8**. Sample reconstituted in absence of  $\beta$ ME.

The results presented in Figure 70 proved that no reduction and oxidation events were observed for the sample reconstituted in presence of atmospheric oxygen. This outcome confirmed that the cyclic voltammogram showed in Figure 69 presents the reduction and oxidation potentials of the cluster maquette synthesised using **MQ8**. A common feature of the cyclic voltammograms showed in Figure 66 and Figure 69 respectively for **MQ4** and **MQ8** is that  $i_{p,a}$  is significantly lower than  $i_{p,c}$ . The peaks were also offset from the reduction process. The degree of irreversibility in our samples is unusual, however, quasi-reversible processes have previously been noted in synthetic  $[4Fe-4S]^{2+/+}$  systems.<sup>106</sup> A quasi-reversible process is obtained when  $E_{p,c}$  and  $E_{p,a}$  (Figure 71 (b)) are far from each other and the mid-point redox potential ( $E_{1/2}$ ) is not close to the two peaks.<sup>107</sup> On the contrary, a reversible system (Figure 71 (a)) is characterised by an  $E_{p,c}$  and  $E_{p,a}$  close to each other, with  $E_{1/2}$  being at or close to the two peaks. In an irreversible reaction (Figure 71 (c)), the potential required to fully reduce or oxidise the species is higher, therefore resulting in a broader peak.

Due to the non-classical shape of the voltammograms obtained for **MQ4** (Figure 66) and **MQ8** (Figure 69),  $E_{1/2}$  could not be accurately determined. However, based on the measured  $E_{p,c}$ , the  $E_{1/2}$  are expected to be at more positive potentials than those typically observed for ferredoxin proteins (-300 to -700 mV vs. SHE) or the previously reported non-acetylated version of **MQ4** ( $H_2N$ -KLCEGGCIGCGACGGW-NH<sub>2</sub>, -350 mV vs. SHE).<sup>66,108</sup> Interestingly, our data were closer to that registered for a 16-mer peptide similar to **MQ4**, but with a single Ala to Gly substitution (Ac-KLCEGGCIGCGACGGW-NH<sub>2</sub>), which registered an  $E_{p,c}$  of -289 mV vs. SHE.<sup>109</sup>

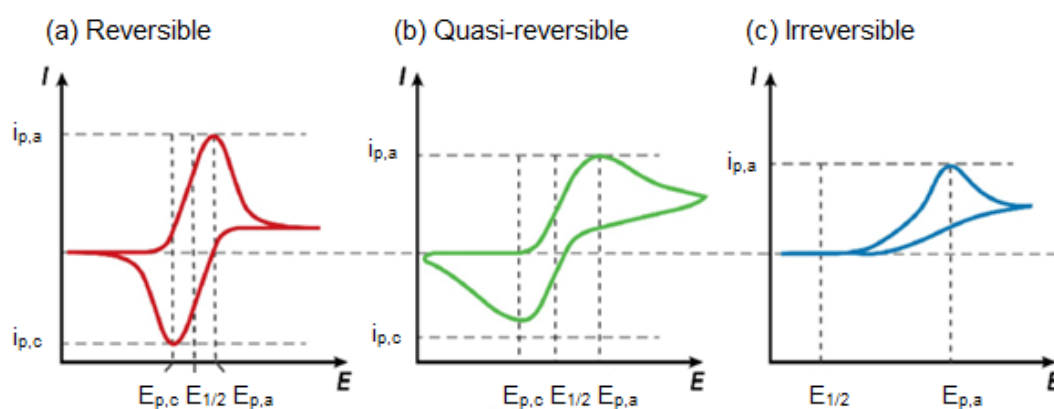


Figure 71 – Cyclic voltammograms for a reversible (a), quasi-reversible (b), and irreversible (c) systems in a typical redox reaction.<sup>107</sup>



The final cluster-maquette to be analysed was the peptide sequence Ac-GUGGGCGGCGGY-NH<sub>2</sub> (**MQ11**), employed for the incorporation of a [4Fe-4S] cluster. In this peptide sequence, the Cys residue on the N-terminus was replaced by a Sec residue. Proteins containing selenium are normally oxidoreductases involved in the defence of the organism against oxidants. In selenoproteins, the selenium residue is close to a Cys residue to form a ring that can be used for redox regulation of intracellular signalling.<sup>110</sup> Despite sulfur and selenium having similar electronegativity, respectively 2.58 and 2.55, Sec is a more reactive nucleophile due to the lower pK<sub>a</sub> value of the selenol moiety (5.43) with respect to the thiol moiety (8.22).<sup>111</sup> This means that at physiological pH, and in the buffer solution in which our experiments were performed, Sec would predominantly exist as a selenolate anion (R-Se<sup>-</sup>), whereas the equilibrium for the Cys residue would be towards the neutrally charged thiol.<sup>111</sup> Previous literature about the redox potentials of Cys and Sec unprotected amino acids reported respectively -0.233 V and -0.488 V with NADPH as a reference.<sup>112</sup> The difference of 250 mV clearly states how Sec is a significantly more powerful reducing agent than Cys. This study did not investigate the redox potential of Cys and Sec in a peptide sequence, and therefore does not give any suggestion on how proximal amino acid residues or the overall environment could affect the reducing properties of the two moieties. Studies on the redox potentials of Sec-containing peptides and their relative Cys-containing homologs showed how the shift in redox potential can be as little as 20-25 mV.<sup>110</sup> The factors to account that influence the redox potentials are multiple and include the solvation of the peptide sequence, the pK<sub>a</sub> of the amino acid residues, the weak forces, the entropy, and the nucleophilicity.<sup>113</sup> Figure 72 shows the cyclic voltammogram for the cluster maquette reconstituted using synthetic peptide **MQ11**.

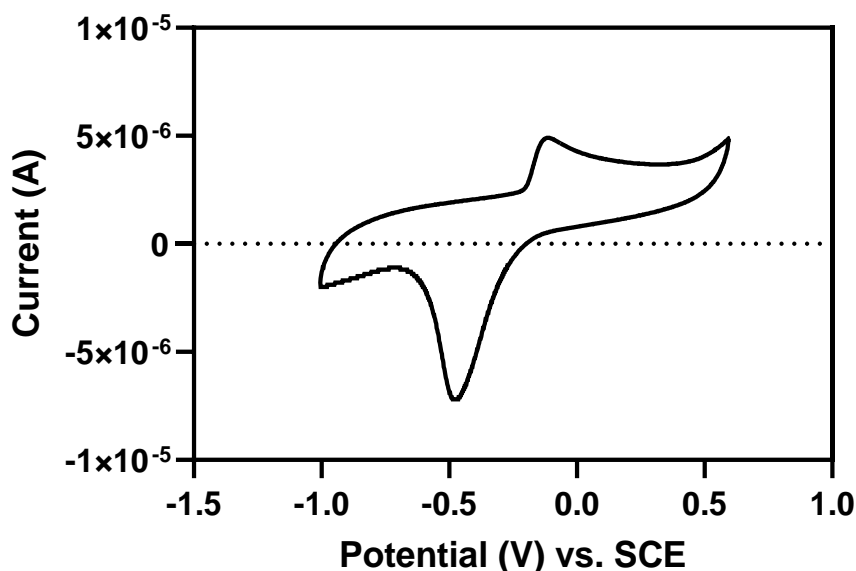


Figure 72 – Cyclic voltammogram for the [4Fe-4S] cluster coordinated by synthetic peptide **MQ11** in 50 mM HEPES 10 mM KCl at pH = 8.0. Sample reconstituted in presence of DTT.

Again, the redox profile was similar to those obtained for **MQ4** and **MQ8**. The  $E_{p,c}$  was registered at -0.476 V vs SCE (-235 mV vs. SHE), which is 47 mV lower than that of bacterial ferredoxin **MQ4**, and 5 mV lower to that of maquette **MQ8**, reconstituted using the poly-Gly peptide. The  $E_{p,a}$  was observed at -0.113 V vs SCE. This potential was similar to both **MQ4** and **MQ8** and was respectively only 1 mV and 5 mV lower. Overall, exchanging one Cys residue for one Sec residue did not play a significant role in affecting the redox potential of the maquette. Interestingly, despite a substantive difference in the electrochemical properties of the single amino acid Cys and Sec, their difference is not perpetuated when they are part of a longer peptide.

Figure 73 shows how the cyclic voltammograms for the three different maquettes compared with each other. It is clear that the shape and relative position of both the oxidation and reduction peaks all have similarities. Minor differences were observed for the reduction and oxidation peaks, but the ranges in which the two redox events were registered were consistent throughout the three cluster maquettes. The reduction peaks for the three maquettes were determined between -0.476 and -0.429 V vs. SCE, whereas the oxidation peaks were all registered between -0.108 and -0.113 V. the shape of the voltammograms is somewhat unusual, but overall consistent with a reversible system, despite quasi-reversible processes have previously been reported for [4Fe-4S]<sup>2+/+</sup> systems.<sup>106</sup> Due to the uncommon shape of the voltammograms, the

mid-point redox potentials could not be determined. However, given the oxidation and reduction potentials, the mid-point values are likely to be higher than those typically seen for ferredoxin proteins (-300 to -700 mV vs. SHE) and the previously reported non-acetylated version of **MQ4** (-350 mV vs SHE).<sup>66,108</sup> Our results are closer to those registered for a 16-mer ferredoxin-like maquette with one single Ala to Gly substitution relative to **MQ4**, which presented a reduction potential of -289 mV vs. SHE.<sup>68</sup>

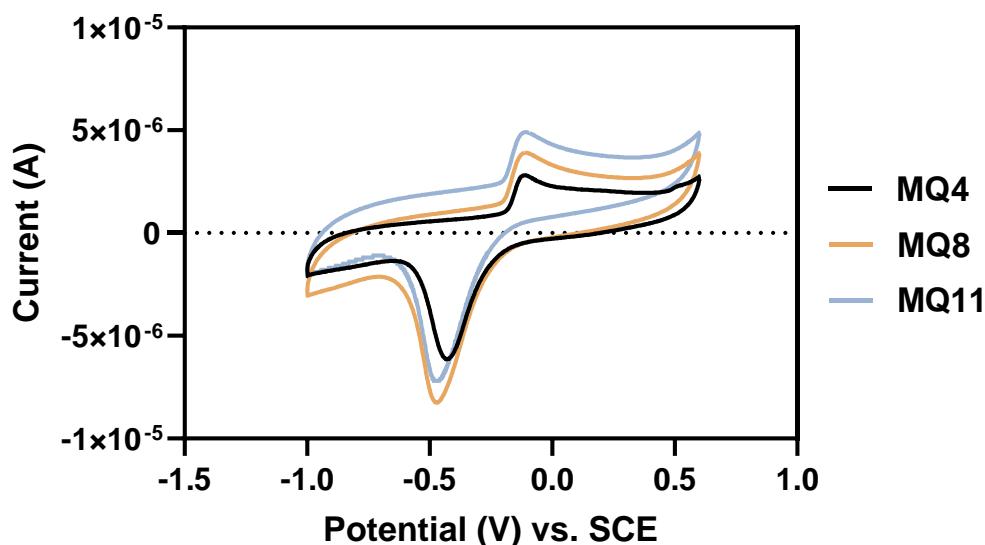


Figure 73 – Comparison of the cyclic voltammograms for the [4Fe-4S] cluster coordinated by synthetic peptides **MQ4** (black trace), **MQ7** (gold trace), and **MQ10** (light-blue trace) in 50 mM HEPES 10 mM KCl at pH = 8.0.

### 3.4.3 Cyclic Voltammetry of [4Fe-4S] Cluster Maquettes in Presence of $\beta$ ME

Despite the fact that no reducing agent was used in the synthesis of the maquettes presented in the previous discussion, and the incorporation of the cluster was afforded by simply mixing the selected peptide with  $\text{FeCl}_3$  and  $\text{Na}_2\text{S}$ ; an experiment was performed using  $\beta$ ME in the reconstitution process. The cluster maquette previously discussed were also not eluted through SEC columns due to issues related to the size of the peptides being too close to the size limit of the pores. Briefly, the peptide was dissolved in the buffer solution containing 5% v/v of  $\beta$ ME. Once the peptide was dissolved and equilibrated in solution, the  $\text{Fe}^{3+}$  and  $\text{S}^{2-}$  ions sources were added.  $\beta$ ME is frequently used in the synthesis of Fe-S cluster maquettes as it ensures the Cys residue on the peptide or protein are fully reduced and in the optimum redox state to

coordinate the iron ions of the cluster. Firstly, a sample of pure  $\beta$ ME in buffer was subjected to the potentiostat voltage to determine its cyclic voltammogram. Figure 74 shows the reduction profile for  $\beta$ ME in 50 mM HEPES 10 mM KCl buffer at pH = 8.0. Due to the powerful reducing activity of  $\beta$ ME, no reduction event was expected to be observed and the voltammogram was consistent with this anticipation. At approximately 0.2 V, the thiol begins to be oxidised and results in an increase in current.

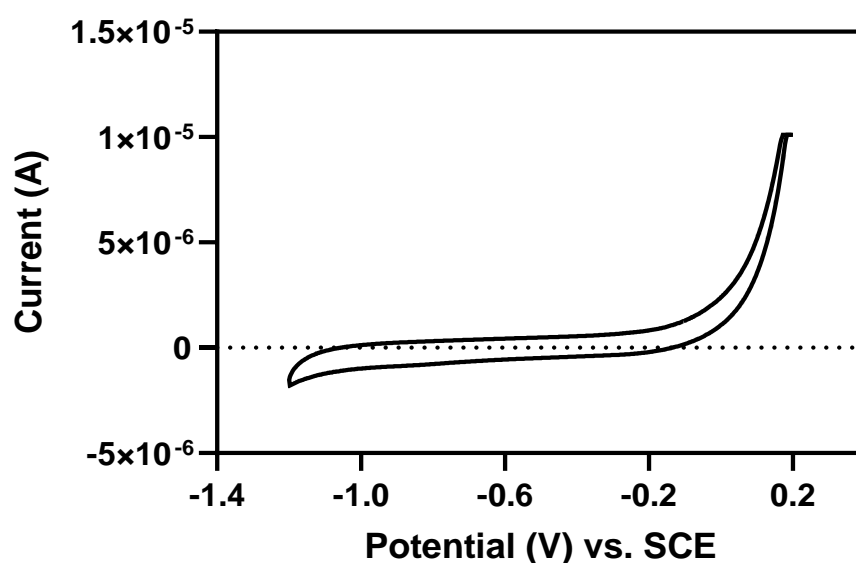


Figure 74 – Cyclic voltammogram for  $\beta$ ME in 50 mM HEPES 10 mM KCl at pH = 8.0.

Figure 75 shows the cyclic voltammogram for the maquette reconstituted using peptide **MQ8** in presence of  $\beta$ ME (black trace). Compared to the cyclic voltammogram of the maquette reconstituted without the reducing agent (gold trace), the notable difference relies in the suppression of both the reduction and oxidation peaks. This effect is clearer on the oxidation peak, as this assumed a flattened shape. The reduction peak was also affected by the suppression, as the current value did not reach the same value when the thiol was presented, however, the shape of the peak remained consistent. The oxidation peak also seemed to be shifted to higher potential, but this effect could also be caused by the different shape of the peak. We could argue that the addition of thiol to such a higher concentration with respect to the peptide, could influence the binding between the cluster and the Cys residues. It is plausible that the thiol of  $\beta$ ME is exchanging the thiol of the Cys residue, causing the reduction and oxidation peaks to be gradually suppressed due to a competitive reaction exerted by  $\beta$ ME.

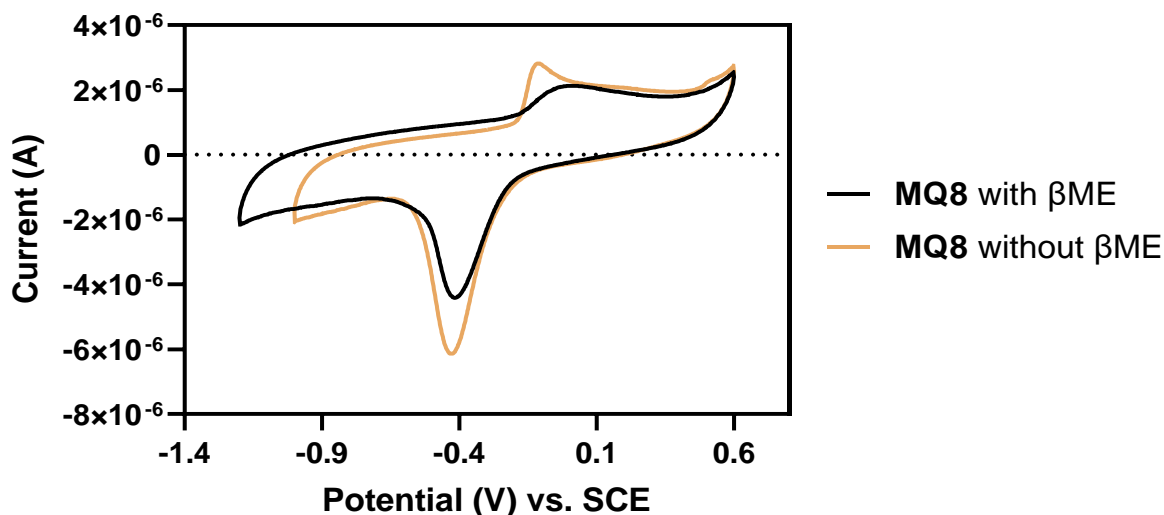


Figure 75 – Comparison of the cyclic voltammograms for **MQ8** in 50 mM HEPES 10 mM KCl at pH = 8.0 reconstituted in presence (black trace) and in absence (gold trace) of  $\beta$ ME.

To further investigate this result, twenty more equivalents of  $\beta$ ME with respect to the peptide concentration were added to the solution.

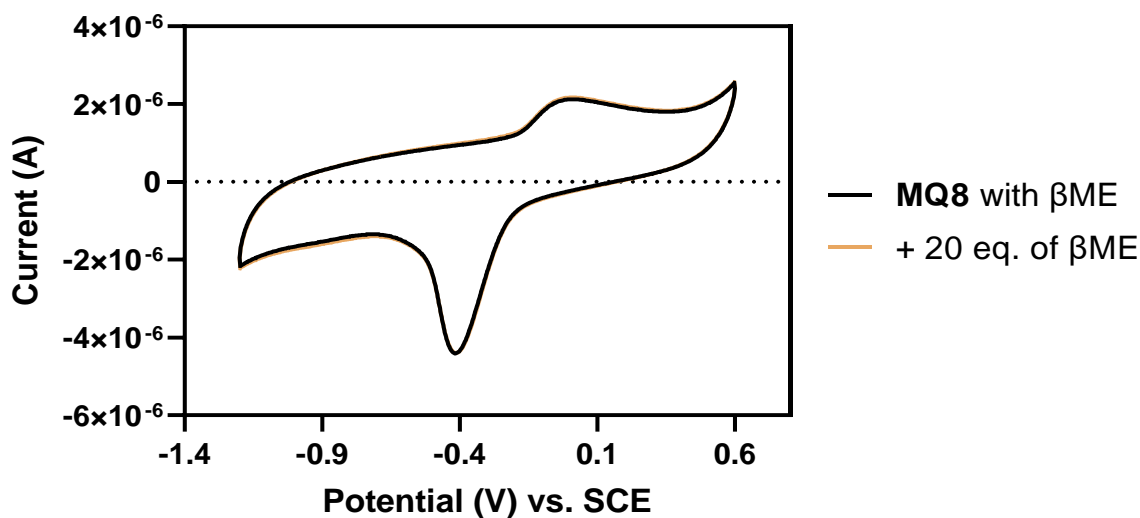


Figure 76 – Comparison of the cyclic voltammograms for **MQ8** in 50 mM HEPES 10 mM KCl at pH = 8.0 reconstituted in presence of  $\beta$ ME (black trace) and after addition of further 20 equivalents of  $\beta$ ME after reconstitution (gold trace).

Figure 76 shows the voltammograms registered before (black trace) and after (gold trace) the addition of the extra equivalents of thiol, however no difference was observed, and the two spectra are identical to each other.

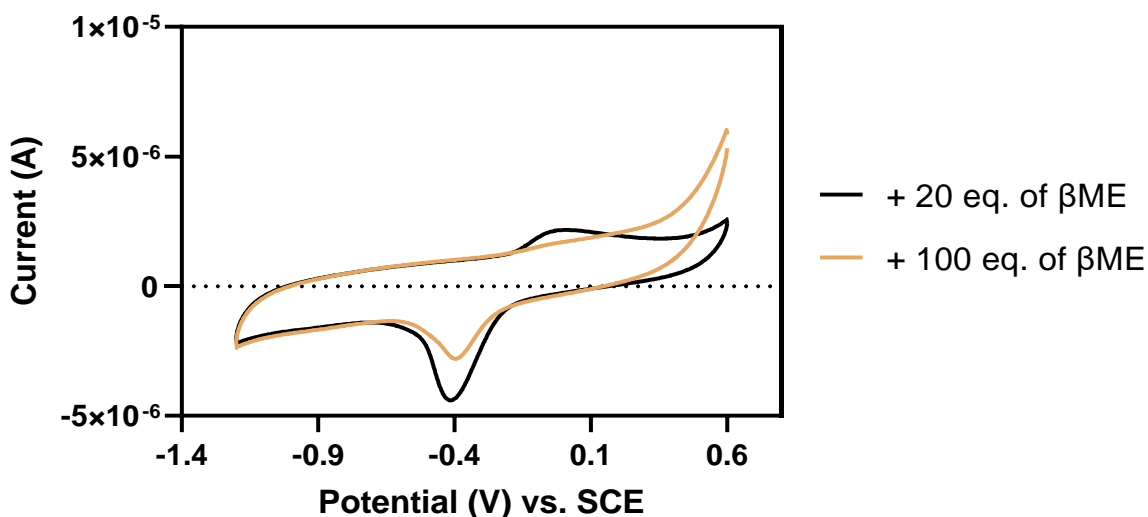


Figure 77 – Comparison of the cyclic voltammograms for **MQ8** in 50 mM HEPES 10 mM KCl at pH = 8.0 after addition of further 20 (black trace) and 100 (gold trace) equivalents of  $\beta$ ME after reconstitution.

The addition of further one hundred equivalents of thiol with respect to the peptide caused the suppression effect on the oxidation peak to be even greater, as displayed by the gold trace in Figure 77. The decrease in current had also increased on the reduction peak, despite this being still visible. Two hundred more equivalents of  $\beta$ ME were added and the resulting cyclic voltammogram is showed in Figure 78. Interestingly, the results were different from what could be anticipated. Based on what happened after the addition of twenty and one hundred equivalents, it was expected that both the oxidation and reduction peaks would be suppressed until no more visible. However, the reduction peak increased in intensity, and the thiol oxidation became more visible. This result implied that the cluster was still present, and the addition of thiol increased its response towards the reduction. It is also plausible that by this point all the peptide was completely exchange by  $\beta$ ME in coordinating the  $\text{Fe}^{3+}$  ions and therefore the reduction peak was purely determined by the redox activity of the interaction between the thiol and the iron ions. Another explanation to describe the reason why the reduction peak was still visible, but the oxidation peak was not, is that  $\beta$ ME could be interacting with the reduced cluster to form a complex that then prevents the cluster to be oxidised.

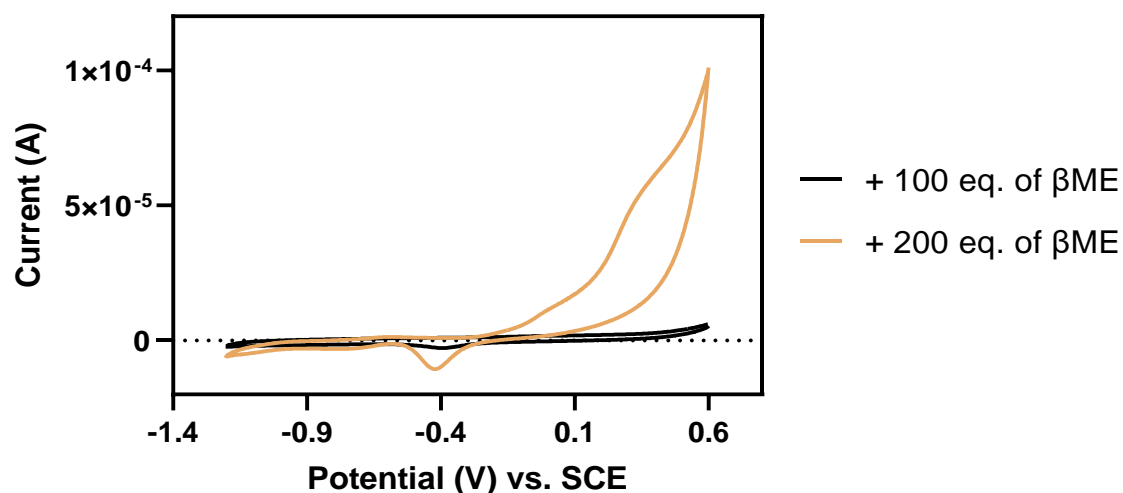


Figure 78 – Comparison of the cyclic voltammograms for **MQ8** in 50 mM HEPES 10 mM KCl at pH = 8.0 after addition of further 100 (black trace) and 200 (gold trace) equivalents of  $\beta$ ME after reconstitution.

One final addition of  $\beta$ ME was performed to bring the final concentration to five hundred equivalents with respect to the peptide concentration, as showed in Figure 79.

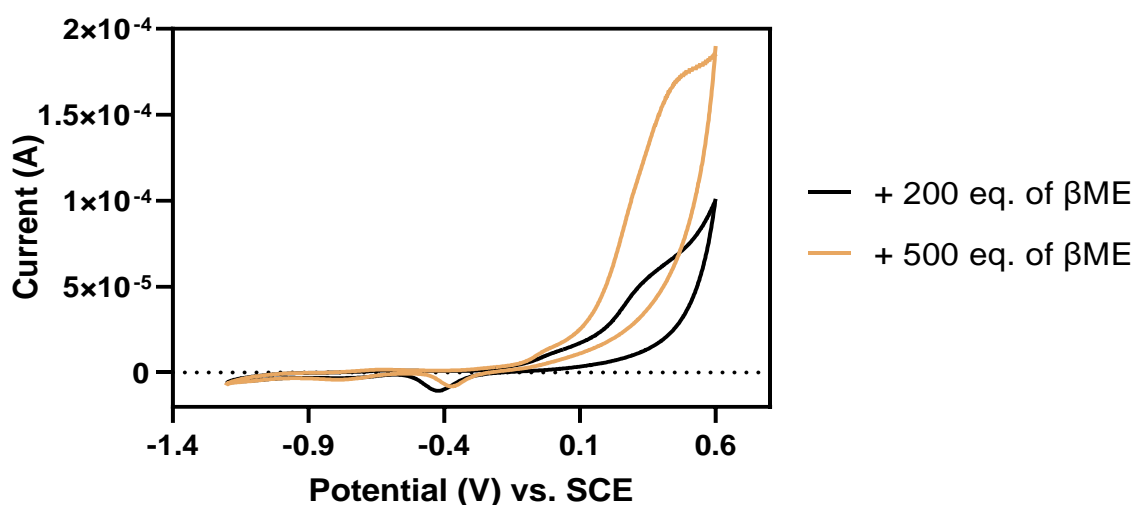


Figure 79 – Comparison of the cyclic voltammograms for **MQ8** in 50 mM HEPES 10 mM KCl at pH = 8.0 after addition of further 200 (black trace) and 500 (gold trace) equivalents of  $\beta$ ME after reconstitution.

The pattern seemed to be similar to the previous addition of reducing agent, however the reduction peak was shifted towards higher potential and the oxidation peak increased greatly and became broader. At this point, the volume of  $\beta$ ME was approximately 9% of the solution, and it was therefore acting as a co-solvent.

### 3.5 CONCLUSION AND FUTURE WORK

To conclude, in this chapter we investigated the redox properties of [4Fe-4S] cluster maquettes by cyclic voltammetry measurements. The reduction and oxidation potentials were successfully observed for the three different maquettes synthesised in absence of the reducing agent  $\beta$ ME. As the presented data was collected from samples that were not eluted through SEC columns, cyclic voltammetry was carried out on control samples of peptides,  $\text{Na}_2\text{S}$ , and  $\text{FeCl}_3$ . The reagents used for maquette formation were also combined under aerobic conditions, under which the cluster formation was not observed. Only when the reagents were combined under anaerobic conditions the values discussed were observed. Therefore, all these control analyses confirmed that the redox events detected were due to the [4Fe-4S]-cluster being sequentially reduced and oxidised. Because the maquettes were not eluted through SEC columns, we cannot rule out alternative unidentified molecular species accounting for the lower than expected  $E_{p,c}$  and  $E_{p,a}$  values.

**MQ8** showed a reduction peak 42 mV vs. SCE towards lower redox potentials when compared to **MQ4**. This drop can be considered as a considerable shift and gives us confidence for the possibility of tuning the redox potentials of these systems. These values are overall in agreement with ferredoxin proteins and with a previously reported acetylated version of **MQ4**. The shape of the obtained voltammograms is quasi-reversible, which is not unusual, but literature examples generally display reversible patterns. **MQ11** presented a reduction profile that was shifted by 45 mV vs. SCE towards lower potentials when compared to **MQ4**. Interestingly, **MQ8** and **MQ11** showed similar values, meaning that the single amino acid change (Cys for **MQ8** and Sec for **MQ11**) does not play a more pivotal role than the overall peptide sequence. The effect of  $\beta$ ME on the redox potentials were also investigated when **MQ8** was reconstituted in presence of the reducing agent. Further additions of the thiol showed suppression of the oxidation peak and a shift towards firstly lower and then higher redox potentials based on the concentration of  $\beta$ ME added to the sample.

In future work, different peptide sequences could be synthesised to incorporate an [4Fe-4S] cluster and generate a library to better model the change of redox potential based on the amino acids used to build the synthetic peptide. Different amino acids bearing various charges and polarity in the side chain could be implemented in the



peptide sequence to modulate the redox potential of the maquette. Furthermore, different reducing agent could be implemented to observe how these affect the reduction and oxidation of the cluster maquette. More analysis would be required to fully understand what causes the suppression of the oxidation peak in presence of  $\beta$ ME, and the shifts on the reduction event.

## 4 INTEGRATION OF [4FE-4S] CLUSTER MAQUETTES IN AN ELECTRON TRANSPORT PATHWAY

### 4.1 INTRODUCTION TO HYDROGENASE

Biological reactions mainly rely on intramolecular electron-transfer pathways to exchange electrons and fulfil their scope. These processes couple external electron donors and acceptors, and frequently implement redox cofactor catalysts such as Fe-S clusters to speed up the process. Exploiting the fundamental mechanism that governs the energetic and kinetic properties of intermolecular electron transfer could have tangible application in the control of catalysis. Among the most studied biological processes that take place in nature, hydrogen oxidation and proton reduction represent one of the most interesting and profitable cycle. Exploiting the potential of this process could lead to new ways of converting hydrogen into a green source of energy.<sup>114</sup> Hydrogenases are enzymes that catalyse the reversible oxidation of molecular hydrogen ( $H_2$ ) to two protons and two electrons ( $H^+ + 2e^-$ ).<sup>115</sup> This reversible oxidation of  $H_2$  catalysed by hydrogenases is achieved by a specialised Fe-S cluster named H-cluster. The direction in which this redox process goes depends on the redox coupling partner, which is required by the enzyme to complete the cycle.<sup>116</sup> The [FeFe]-hydrogenase used in our experiments is CaHydA from *Clostridium acetobutylicum*, this has a natural redox partner that is a bacterial-type 2[4Fe-4S] ferredoxin expected to bind the *N*-terminus of the enzyme via the so-called F-domain. The turnover frequency (TOF) is described as the number of chemical conversions of a substrate molecule that a single active site of an enzyme is able to execute.<sup>117</sup> The TOF is measured in  $k_{cat}$ , which corresponds to (moles of product/second)/(moles of enzyme) or as  $s^{-1}$ . The TOF for the physiological process between the CaHydA hydrogenase and its ferredoxin redox partner has been calculated to be  $901 \pm 225 s^{-1}$ .<sup>118</sup>

### 4.2 AIMS AND OBJECTIVES

The metabolic process used by microbes to produce energy involves  $H_2$  as the main source of energy. Enzymes such as the hydrogenase family that contains Fe-S

clusters are widely adopted by microbes in this metabolic pathway. In this process, the enzyme is reduced to favour the oxidation of molecular hydrogen to two atoms of  $H^+$ . To complete the cycle, a coupling partner is required for the hydrogenase to be oxidised. Nature has developed a partnership between different Fe-S proteins to fulfil this redox step, meaning that a second enzyme oxidises the hydrogenase.<sup>119</sup> Due to the low reduction potential of the oxidised (2+) and reduced (1+) [4Fe-4S]-clusters, these systems fall in the redox window of hydrogen metabolism and could potentially be implemented as components for  $H_2$ -technology in tandem with hydrogenase proteins. In Chapter 4 we describe the implementation of a [4Fe-4S]-cluster maquette in the mimicked metabolic pathway that couples the reduction of a hydrogenase enzyme with the oxidation of molecular hydrogen. We sought to substitute the natural coupling partner for the  $H_2$ -oxidation process catalysed by the CaHydA hydrogenase and the bacterial-type ferredoxin, by substituting the latter for our cluster maquette. In particular, we describe the use of an [4Fe-4S]-cluster maquette that acts as the oxidising agent for the hydrogenase, which then oxidises molecular hydrogen to  $H^+$  (Figure 80). **MQ8** was selected to act as the catalytic partner in a biologically relevant electron transport process. This maquette showed the highest conversion in [4Fe-4S]-cluster incorporation by EPR spectroscopy, as described in Chapter 2 of this thesis.

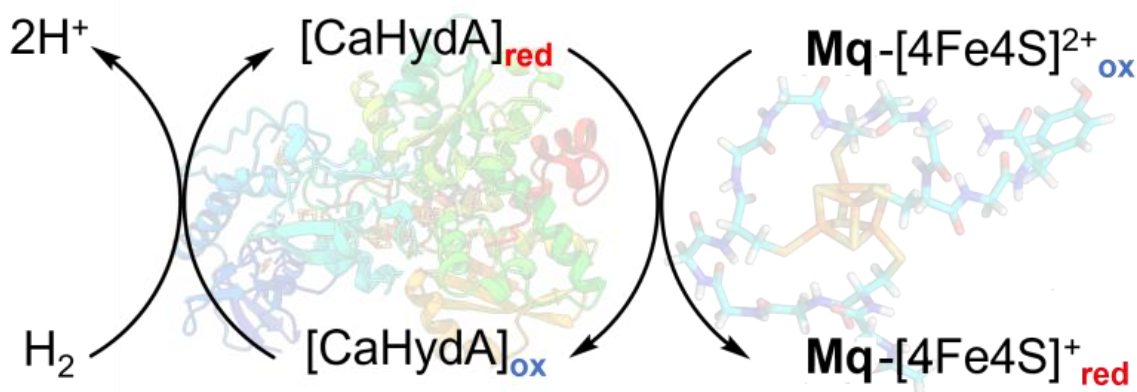


Figure 80 – Integration of the peptide-cluster-maquette **MQ8** (Ac-GCGGGCGGCGCGY-NH<sub>2</sub>) as the terminal electron acceptor within a hydrogenase catalysed  $H_2$ -oxidation pathway.

## 4.3 RESULTS AND DISCUSSION

### 4.3.1 Experimental Set-Up

Assembly of the cell in which the experiments were performed consisted in a 1 mL quartz cuvette containing the cluster maquette in 50 mM HEPES and 10 mM KCl buffer at pH = 8.0 (Figure 81). The cuvette was sealed under anaerobic conditions to ensure the cluster never came in contact to atmospheric oxygen.



*Figure 81 – Quartz cuvette containing the [4Fe-4S] cluster maquette sample sealed under anaerobic conditions.*

Once the cuvette was removed from the glove box under anaerobic conditions, molecular hydrogen was bubbled through the solution for a period of five minutes. This step ensured that the solution and the head space of the cuvette would be completely saturated with hydrogen. The concentration of hydrogen in solution could not be accurately determined, however, literature shows that the solubility of hydrogen gas in water at 25 °C at a pressure of one atmosphere is approximately 1.55 milligrams per kilogram of water.<sup>120</sup> Given a volume of 700 microliters of sample, the mass of hydrogen dissolved in solution can be estimated at approximately 1.09 nanograms (or 0.77 M). The volume of hydrogenase added to conduct the experiments was of 2  $\mu$ L and the final concentration of enzyme in the sample was 0.2 mg/mL, or 7.6 nM.

The first challenge in proving the utility of the cluster maquette in acting as a coupling partner in the metabolic pathway that ultimately leads to the oxidation of molecular hydrogen was finding a suitable analytical technique. Different options were available to determine the appearance or consumption of one of the partners. The first option we evaluated was to determine the concentration of hydrogen in the head space.

Hydrogen is being consumed during the cycle to produce  $H^+$ , and therefore determining the consumption of the former could provide us with a measure of how efficient the [4Fe-4S]-cluster maquette was as a coupling partner. An issue with this option is that the initial concentration of hydrogen in solution could not be determined, being this in large excess with respect to both the maquette and the hydrogenase. Moreover, taking in consideration the equilibrium between the hydrogen dissolved in solution, and that present in the head space of the cuvette would have led to complicated calculations. A second option we evaluated was to determine the concentration of enzyme by using UV-Vis spectroscopy. The first measurement to be registered would be the absorbance for the oxidised cluster presented in the enzyme before oxidation of the molecular hydrogen dissolved in solution. The second analysis would be the absorbance of the cluster after the enzyme oxidises the gas. We previously described how the interaction between the cluster and the amino acid ligands results in a visible absorbance that can be observed with a spectrophotometer. The oxidised, 2+, cluster absorbs at approximately 420 nm, whereas the reduced, 1+, cluster does not absorb UV radiation. In this case, a drop in absorbance would be expected given the oxidised cluster is progressively reduced as it oxidises the molecular hydrogen dissolved in the buffer. The issue with this procedure is given by the low concentration of enzyme presented in solution with respect to the cluster maquette. Because they both absorb in the same region of wavelength, the absorption of the maquette would mask the absorption produced by the enzyme. Given a concentration of 0.6 mg/mL of peptide and assuming 100% conversion for the Fe-S cluster reconstitution, the concentration of maquette in solution would be 0.6 mg/mL. The concentration of enzyme in the stock solution was 0.2 mg/mL, and 2  $\mu$ L were used in the experiments. This led to an enzyme concentration of 0.00057 mg/mL, which is approximately 1000-fold lower than the maquette concentration. The final option available was to monitor the concentration of [4Fe-4S]-cluster maquette throughout the experiments. Contrary to the enzyme, the cluster maquette was present in a concentration high enough to be analysed. Its absorbance was expected at 420 nm due to the interaction between the iron atoms of the cluster and the thiol moiety of the peptide. We therefore chose to determine the concentration of the cluster maquette in solution prior to the addition of molecular hydrogen and enzyme. The observed value would have then been compared to the absorbance showed by the cluster following its reduction in presence of the enzyme.

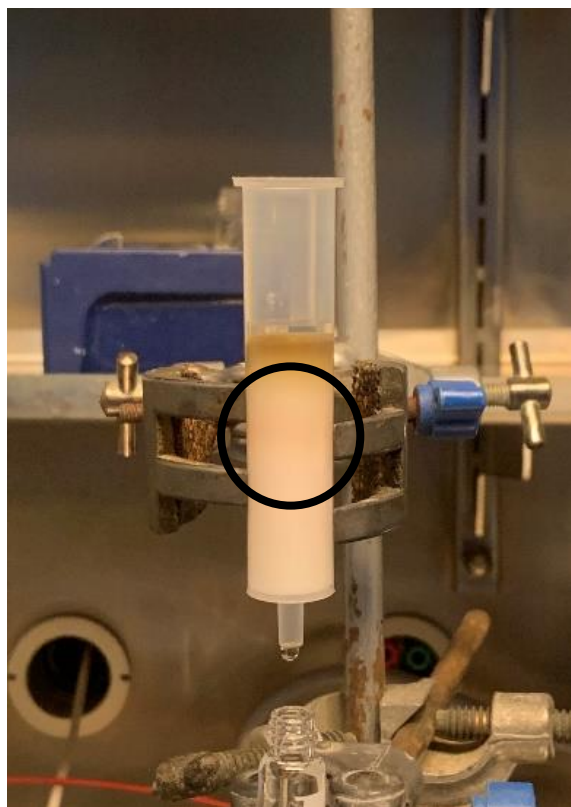
#### 4.3.2 Hydrogenase Experiments on Purified [4Fe-4S] Cluster Maquette Samples

We have previously shown how the presence of free thiol  $\beta$ ME can coordinate with excess  $\text{Fe}^{2+}$  ions in solution to yield a coordinated compound that, due to its similar nature to the cluster maquette, absorbs at an identical wavelength. The use of this thiol in the reconstitution of Fe-S cluster maquettes therefore masks the observation of the latter when UV-Vis spectroscopy is the analytical tool implemented for their observation. The first option we evaluated was to reconstitute the Fe-S cluster maquette using the standard procedure that involves  $\beta$ ME to ensure all Cys residues are reduced, and we then proceeded to purify the cluster using SEC columns.

This chromatographic technique allows for the separation of molecules in solution based on their size and molecular weight. Briefly, when the solution is loaded on top of the column, the small molecules are trapped in the pores of the adsorbent stationary phase. The molecules small enough to be held by the depression on the surface of the beads, or that enter the channels that cross the beads are therefore slowed in passing through the stationary phase. On the contrary, bigger molecules are allowed to freely elute through the column and can be collected in the first fractions. The SEC columns used in these experiments had an exclusion limit of 700 g/mol. This means that every molecule with a molecular mass lower than 700 g/mol would be trapped by the stationary phase, whereas molecules with a higher molecular weight would be allowed to quickly elute.  $\beta$ ME has a molecular weight of 78.13 g/mol, and therefore should be slowed down by the stationary phase. On the other side, the cluster maquette has a mass of approximately 1500 g/mol, however, this molecular weight is still close to the limit of the SEC column since the cluster is not a globular protein.

An issue encountered when firstly using these SEC columns was regarded to the logistics involved when moving the columns into the glove box. SEC columns are pre-packed with a 20% ethanol in water storage solution that has not been degassed, meaning that it might contain dissolved oxygen. All solutions that are used in the glove box in this text have previously been degassed using the freeze-pump-thaw method described in *8.2 GENERAL PROTOCOLS*. However, it was logistically difficult to degas the SEC columns using a similar procedure, and no other protocols were found online. The best option seemed to be to transport the columns inside the glove box as they were purchased and equilibrate them with degassed buffer once in anaerobic

conditions. The columns were therefore moved inside the glove box and equilibrated with 20 mL of buffer before the sample was loaded (see 8.2 *GENERAL PROTOCOLS* for a detailed description of how the sample was loaded on SEC columns). When the sample containing the Fe-S cluster maquette reconstituted in presence of **MQ4** (Ac-KLCEGGCIACGACGGW-NH<sub>2</sub>) and  $\beta$ ME was loaded on the column, a red band quickly appeared while eluting the stationary phase (Figure 82).



*Figure 82 – SEC column used for the purification of the [4Fe-4S] cluster maquette sample. The red band is circled in black.*

This band was thought to be the cluster maquette eluting through the pores. A dark green-brown layer stacked on top of the stationary phase and was hypothesised to be FeS solid particles formed due to excess of FeCl<sub>3</sub> and Na<sub>2</sub>S used in the reconstitution mixture.  $\beta$ ME was thought to be slowly eluting through the column at a much lower pace with respect to the cluster maquette and was therefore expected to be separated from the Fe-S cluster maquette. The red band eluting from the SEC column was collected in three different fractions, which were analysed by UV-Vis spectroscopy (Figure 83).

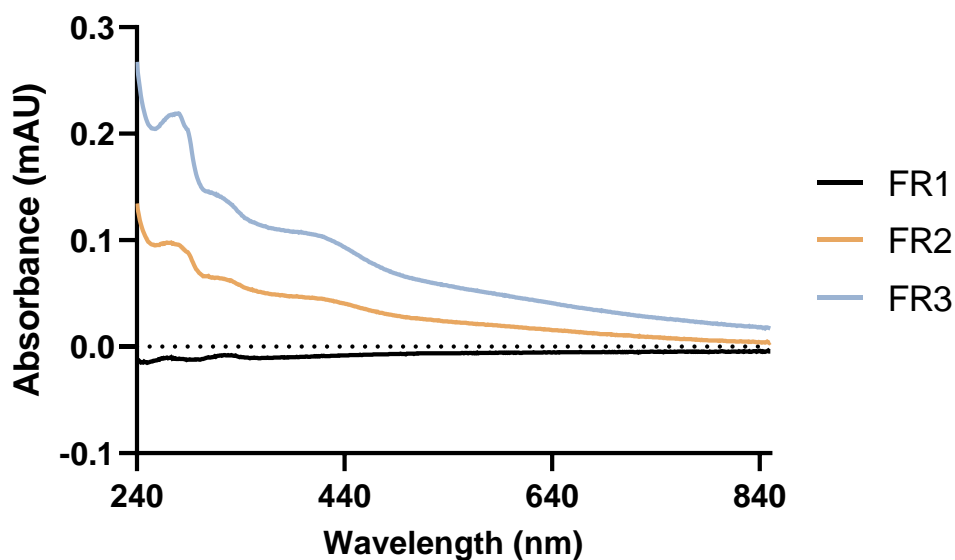


Figure 83 – Comparison of UV-Vis trace for the three fractions collected after purification by SEC columns. Samples were analysed directly, and concentrations could not be quantified.

The spectrum for FR1 proved the absence of any species absorbing at approximately 420 nm, the characteristic wavelength at which the LMCT interaction is observed for Fe-S cluster maquettes. FR2 showed an absorbance at 420 nm, but the absorbance registered was low and the peak not sufficiently defined from the baseline. In addition to the peak at 420 nm, a peak at approximately 280 nm was observed that is characteristic for the amino acid Trp.<sup>79</sup> This observation raised confidence towards the successful purification of the cluster maquette from the salts and the thiol used in the reconstitution mixture. The UV-Vis spectrum for FR3 showed a similar result to that of FR2, but with a stronger absorbance. This fraction contained most of the red band that eluted from the SEC column and had a volume of approximately 2mL. Because the volume of sample loaded was of 2 mL, it is possible that the additional volume of buffer required to completely elute the cluster maquette through the column diluted the sample by some degree. FR3 was selected for the experiments with molecular hydrogen and the hydrogenases due to the stronger signal registered by its spectrum. Initially, the sample was sparged with molecular hydrogen for five minutes and a spectrum was recorded to monitor any change in absorbance. As showed in Figure 84, the absorbance dropped slightly after the addition of hydrogen to the solution. Despite being only a modest change in absorbance, this could imply that the presence of hydrogen affected the cluster stability to a certain extent.



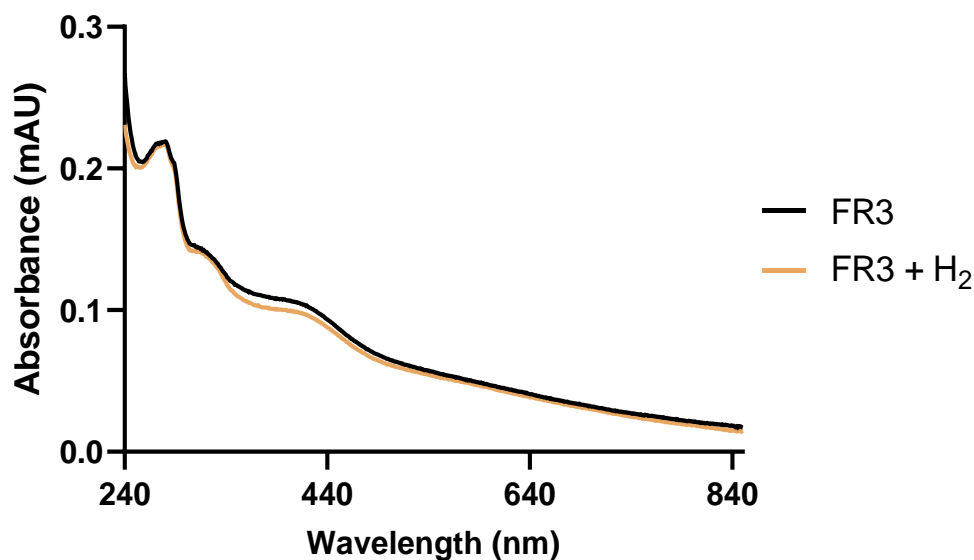


Figure 84 – Comparison of UV-Vis trace for FR3 before (black trace) and after (gold trace) sparging with H<sub>2</sub>.

Once the enzyme was added to the solution this was left reacting for approximately ten minutes and then analysed by UV-Vis spectroscopy. The hydrogenase enzyme that we used, CaHydA from *Clostridium acetobutylicum*, incorporates two [4Fe-4S]-clusters that would also absorb at approximately 420 nm. The recorded spectrum is showed in Figure 85 along with the trace before the addition of enzyme.

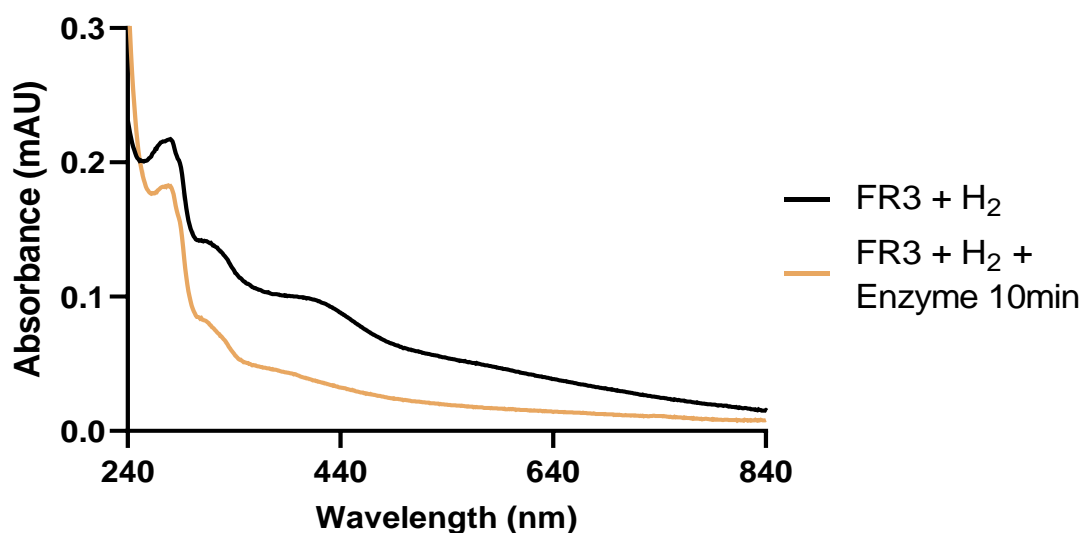


Figure 85 – Comparison of UV-Vis trace for FR3 after sparging with H<sub>2</sub> before (black trace) and after (gold trace) addition of enzyme and equilibration for 10 minutes.

The different absorbance registered before and after addition of the enzyme gave clear evidence that the oxidised cluster-maquette was progressively reduced in presence of

the enzyme. The reaction mixture was analysed again after five minutes to confirm that the absorbance did not drop further, and the spectrum is displayed in Figure 86.

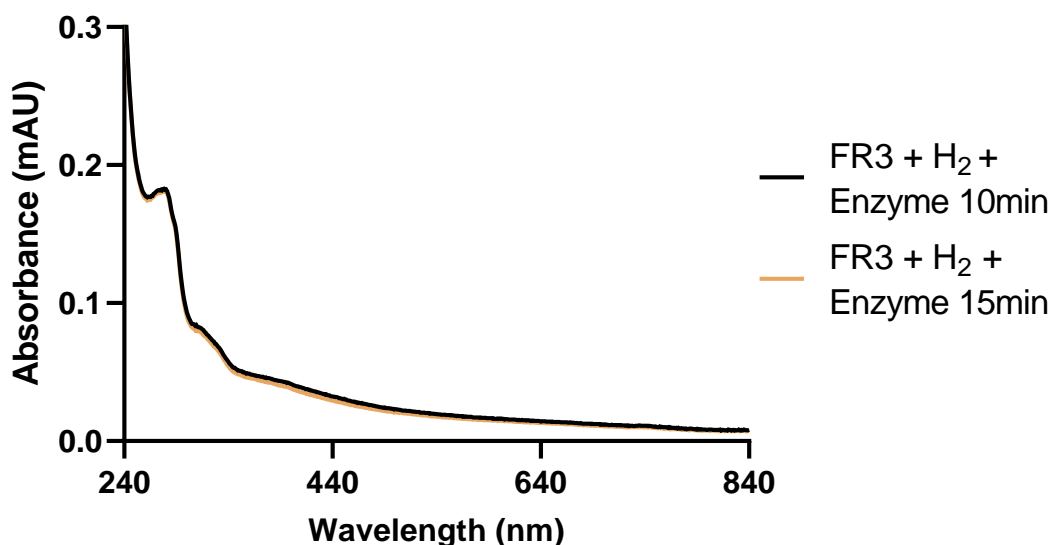


Figure 86 – Comparison of UV-Vis trace for FR3 after enzyme addition and equilibration for 10 (black trace) and 15 (gold trace) minutes.

The two spectra looked very similar and therefore it was assumed that an equilibrium was reached, and the absorbance would not have decreased any further. The experiment with the hydrogenase was repeated to record a time-point analysis and investigate if the absorbance would sharply or gradually decrease over a period of time. The result would give an idea of how quickly the [4Fe-4S]-cluster is reduced in presence of the hydrogenase and molecular hydrogen. Figure 87 shows the absorbance trend for the reduction from [4Fe-4S]<sup>2+</sup> to [4Fe-4S]<sup>1+</sup> registered at 420 nm. From this experiment we concluded that the reduction of the cluster happens gradually over ten minutes and eventually reaches an equilibrium between the two species.

To investigate if the cluster-maquette could be re-oxidised after being reduced in presence of molecular hydrogen and the hydrogenase, we decided to expose the reaction mixture to oxygen. To do so, we used a syringe to pump 5 mL of ambient atmosphere in the quartz cuvette containing the reaction mixture. A spectrum was recorded after this experiment and is showed in Figure 88.

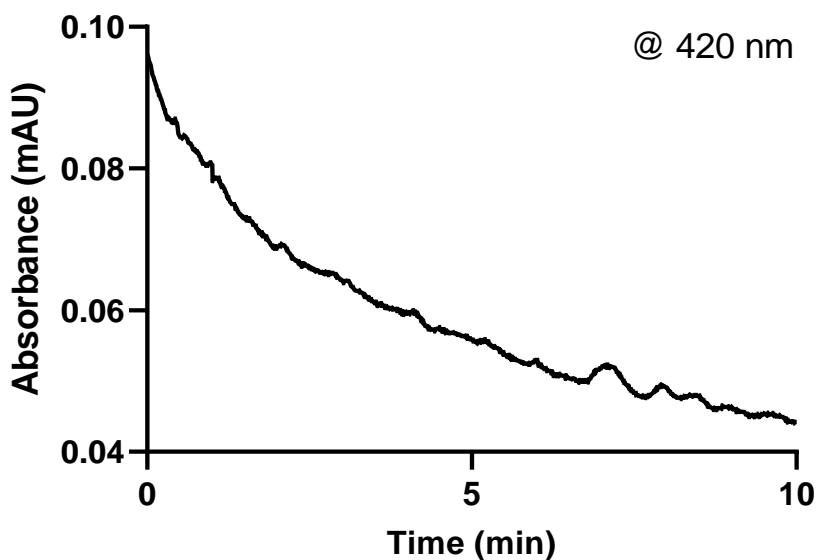


Figure 87 – Absorption trend registered from time zero to 10 minutes at 420 nm for sample FR3 sparged with  $H_2$  and treated with enzyme.

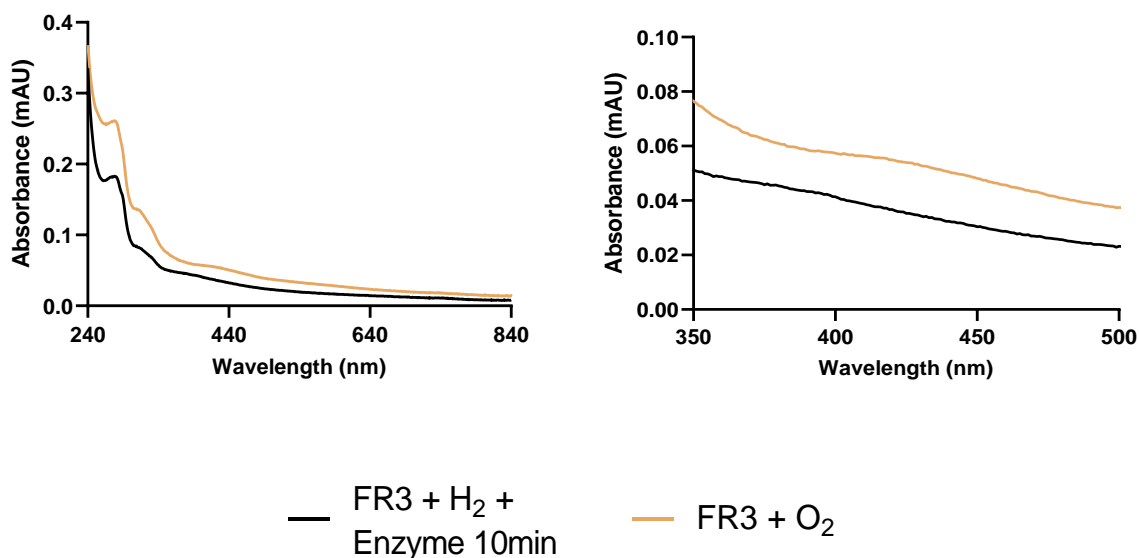


Figure 88 – Comparison of UV-Vis trace for FR3 sparged with  $H_2$  after addition of enzyme and equilibration for 10 minutes (black trace) and after exposure to atmospheric oxygen (gold trace). Left panel is expansion of the area where the peak was observed.

The presence of molecular oxygen had a small positive effect on the absorbance of the reaction mixture. A slight peak appeared at approximately 420 nm, suggesting that the cluster was still available after its reduction, and it could be re-oxidised in presence of an oxidant (i.e., the oxygen available in ambient atmosphere). We then proceeded to add an extra aliquot of enzyme to the reaction mixture to explore whether the re-oxidised cluster could also be re-reduced. Figure 89 shows the spectrum for the reaction mixture after the addition of enzyme (gold trace), proving that the cluster can

be reduced again after being exposed to oxygen, but only to a certain extent. In this case, the reduction in absorbance from the pre-reduction (black trace) and post-reduction (gold trace) proves that the oxidised cluster can act as a catalytic oxidising agent in the electron transport process and be reduced.

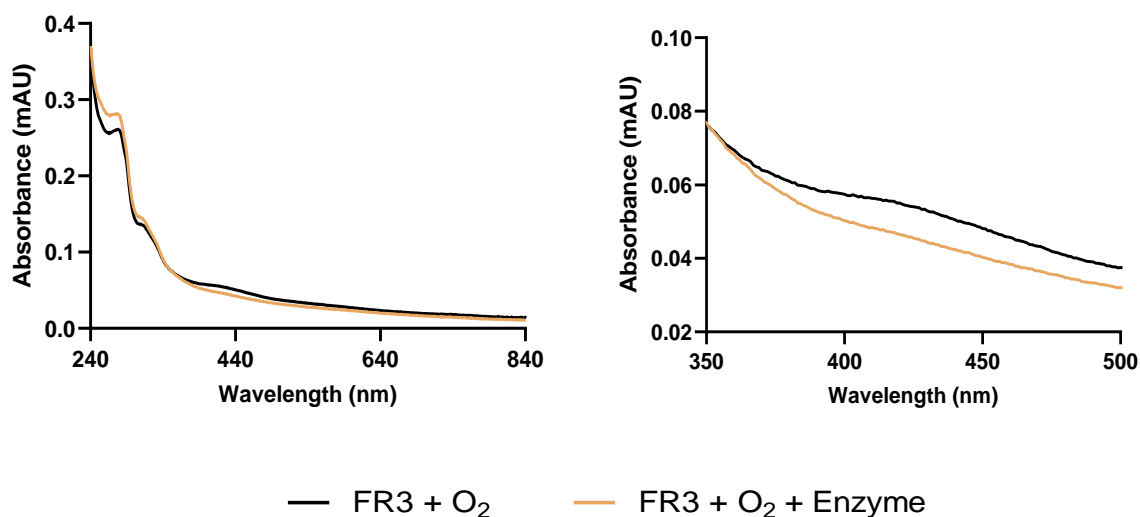


Figure 89 – Comparison of UV-Vis trace for FR3 sparged with H<sub>2</sub> after addition of enzyme and exposure to atmospheric oxygen (black trace), and after addition of further enzyme (gold trace).

When compared to the spectrum registered after the first reduction (Figure 90, black trace), it is possible to notice that the second reduction (Figure 90, gold trace) did not reach the same level of absorbance.

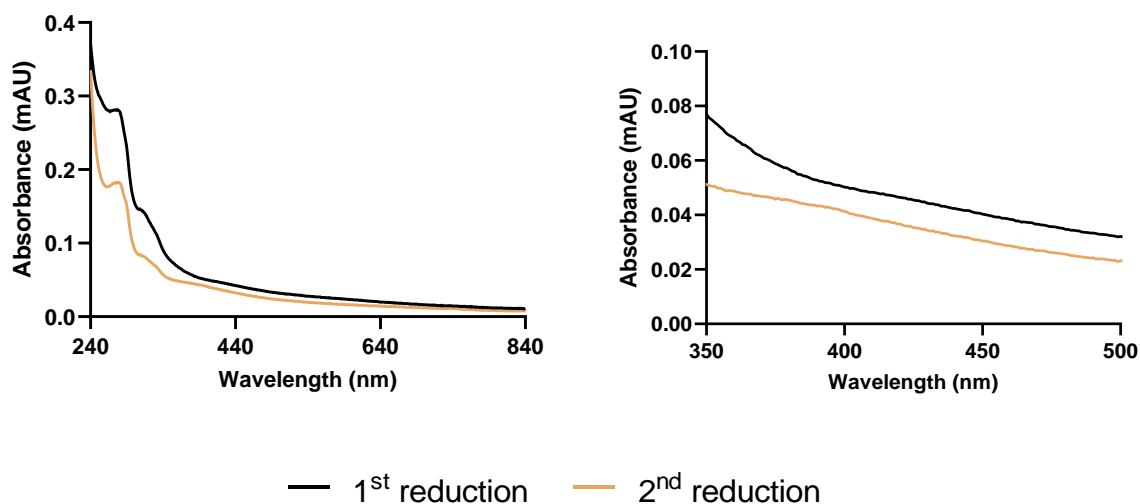


Figure 90 – Comparison of UV-Vis trace for treatment with enzyme during the first (black trace) and second (gold trace) reduction cycles. Left panel is expansion of the area where the peak was observed.

The ability of the maquette to retain its interaction with the cluster during a redox cycle would be of great interest. If the cluster is stable in these conditions and can be cycled between the oxidised (2+) and reduced (1+) state, it could be used as a catalyst in iterative processes.

Unfortunately, when the quartz cuvette containing the reaction mixture was open to dispose of the content an unpleasant odour of thiol could be detected. This was presumed to be arising from residual  $\beta$ ME present in the sample. Despite the purification step performed using the SEC columns, we concluded that a certain amount of thiol co-eluted with the cluster-maquette. Even a small quantity of thiol in solution could be responsible for a strong smell to be detected, however, its presence could be altering the nature of the cluster maquette interaction affecting the result of the analysis. We have previously described how  $\beta$ ME can coordinate with iron ions in solution to produce a similar result to that observed for the peptide interaction with the cluster. If the SEC purification successfully removed all iron particles in solution, the thiol that co-eluted in the final sample would not be able to coordinate with any  $\text{Fe}^{2+}$  ions to produce the misleading interaction that could be observed in the UV-Vis spectrum. However, the presence of thiol in solution with the purified maquette could be detrimental for its stability.  $\beta$ ME could indeed act as a scavenger and progressively substitute the peptide in the interaction with the cluster, gradually decreasing its concentration.

#### 4.3.3 Hydrogenase Experiments on [4Fe-4S] Cluster Maquettes in Absence of $\beta$ ME

To avoid contamination of the samples and false negative or false positive results, we decided to reconstitute the cluster maquette in the absence of  $\beta$ ME, and therefore exclude this uncontrollable variable from the experiments. For these experiments we coordinated the Fe-S cluster using synthetic peptide **MQ7** (Ac-GCGGGCGGCGGY-NH<sub>2</sub>), as this was the candidate that showed the best EPR results among the bespoke maquettes that we developed. The first spectrum was recorded on the reconstituted sample before sparging hydrogen in the reaction mixture (Figure 91).

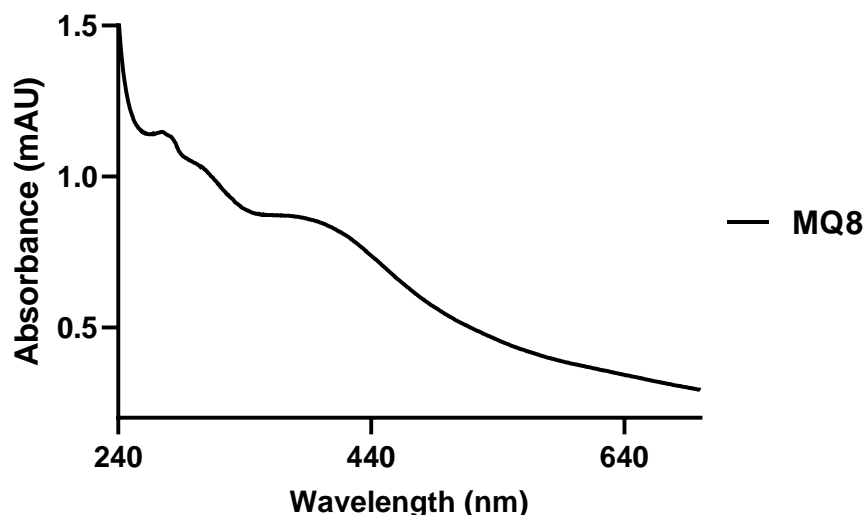


Figure 91 – UV-Vis absorption for **MQ8**. The spectrum was recorded at peptide concentration of 100  $\mu\text{M}$ .

When comparing this spectrum to the one recorded for the cluster maquette purified by SEC, the absorbance values are notably higher possibly because the sample was not diluted by extra buffer in the elution process. Another difference is given by the smaller peak observed at 280 nm, and this is because in this case the peak is characteristic of Tyr instead of Trp. This amino acid showed a weaker peak possibly because of the lower grade of complexity of the surrounding amino acid, which might have decreased the degree of conjugation. The extinction coefficient for Tyr and Trp are respectively  $\epsilon = 1280 \text{ M}^{-1}\text{cm}^{-1}$  and  $\epsilon = 5690 \text{ M}^{-1}\text{cm}^{-1}$ , therefore resulting in different absorption values.<sup>121</sup> Moreover, Tyr normally gives stronger peaks at higher pHs, but always in the range between 230 and 300 nm, as observed for the spectra displayed in Figure 91.<sup>122</sup> The sample was then sparged with molecular hydrogen for five minutes and the reaction mixture analysed (Figure 92). As previously observed for the sample purified by SEC columns, no major changes in the UV-Vis trace were recorded and the sample was assumed to be stable and in equilibrium.

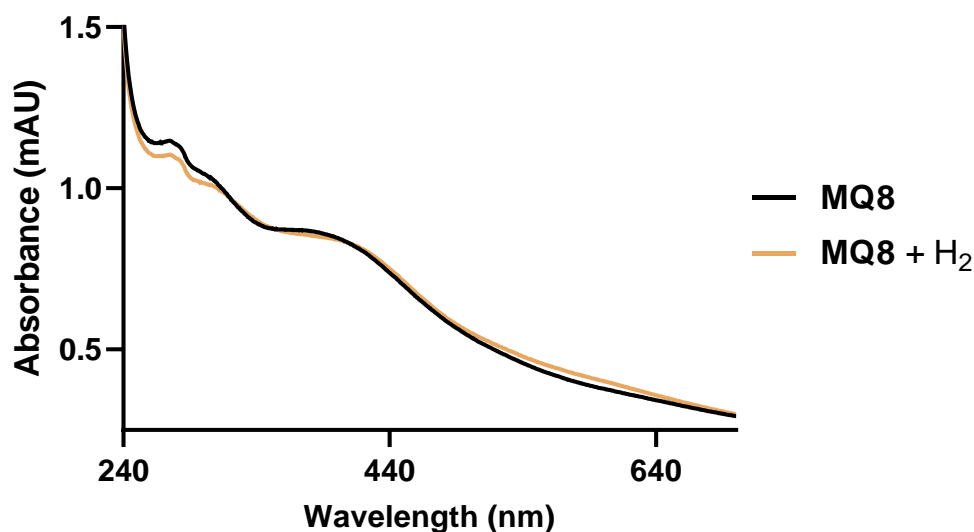


Figure 92 – Comparison of UV-Vis trace for **MQ8** before (black trace) and after (gold trace) sparging with  $H_2$ . The spectrum was recorded at peptide concentration of  $100 \mu M$ .

The enzyme was added to the reaction mixture and left reacting for 20 minutes. The UV-Vis spectrum recorded after this time is displayed in Figure 93 and shows a clear drop in absorbance in the characteristic area where the oxidised cluster maquette absorbs. The decrease in absorbance does not lead to a flat spectrum, meaning that the cluster population is not fully reduced in these conditions. Instead, it is plausible that an equilibrium was reached between the oxidised and reduced cluster.

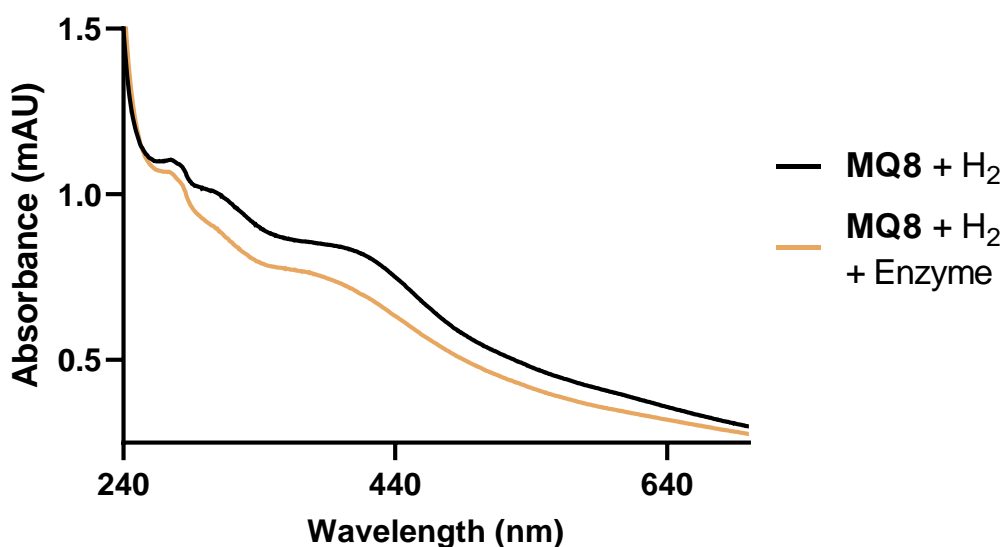


Figure 93 – Comparison of UV-Vis trace for **MQ8** sparged with  $H_2$  before (black trace) and after (gold trace) addition of enzyme. The spectrum was recorded at peptide concentration of  $100 \mu M$ .

The time-points determined for the reduction of the cluster maquette is displayed in Figure 94. The peak maximum was observed at 420 nm, and this wavelength was therefore selected for the time-points observation.

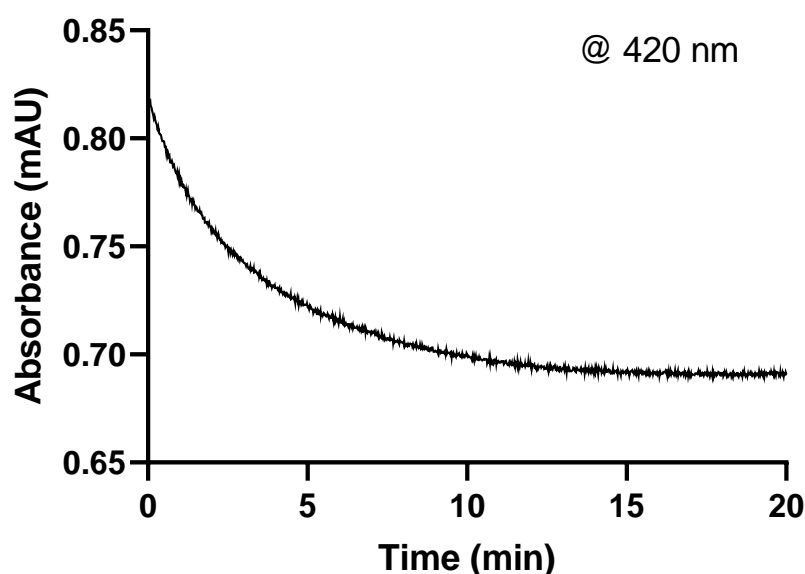


Figure 94 – Absorption trend registered from time zero to 20 minutes at 420 nm for **MQ8** sparged with  $H_2$  and treated with enzyme.

From the graph we concluded that the equilibrium is reached after approximately 15 minutes from the addition of enzyme, and then the trend comes to a plateau. To determine what the trend would look like in case the cluster maquette would be fully reduced, we added DT to the reaction mixture. DT is a reducing agent commonly implemented in biological experiments that has a reduction potential of approximately -0.420 V vs SCE.<sup>123</sup> Figure 95 displays the spectra recorded before and after addition of DT to the reaction mixture after the hydrogenase experiment.

The spectrum registered after the addition of DT (gold trace) showed a reduced value of absorbance around 420 nm, which is expected for a completely reduced cluster. The peak at approximately 315 nm is characteristic for DT accumulation.<sup>85</sup> This observation proves that the Fe-S cluster is fully reduced as the excess dithionite absorbed in the UV-Vis spectrum. The overall reduction profile observed in the area around 420 nm is showed in Figure 96.



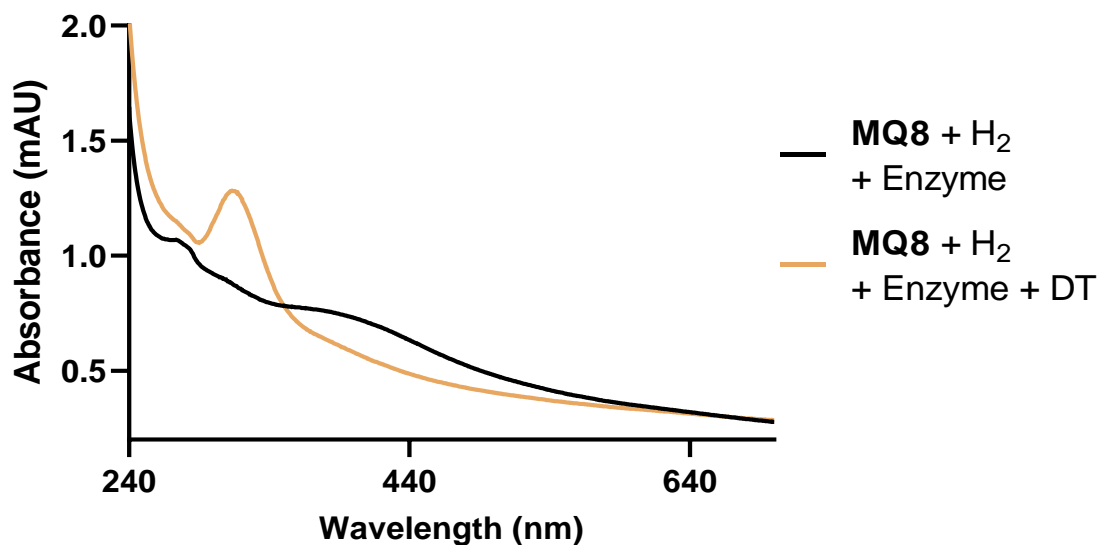


Figure 95 – Comparison of UV-Vis trace for **MQ8** sparged with  $H_2$  after addition of enzyme before (black trace), and after addition of DT (gold trace). The spectrum was recorded at peptide concentration of  $100 \mu M$ .

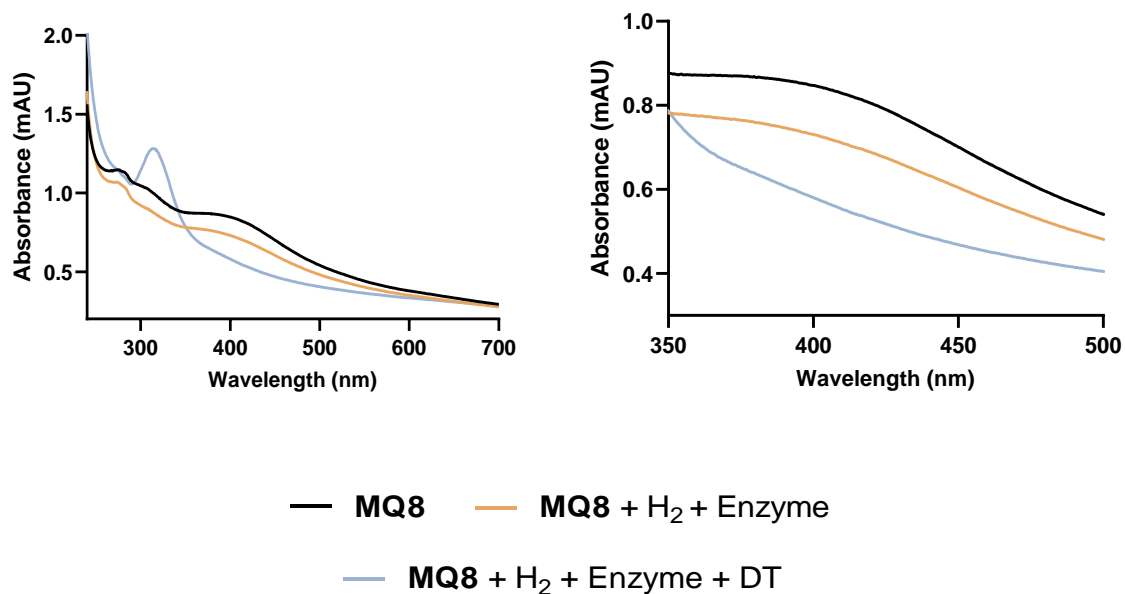


Figure 96 – Comparison of UV-Vis trace for **MQ8** (black trace), after  $H_2$  sparging and enzyme addition (gold trace), and after reduction with DT (light-blue trace). Left panel is expansion of the area where the peak was observed.

The absorbance before and after addition of enzyme was respectively 0.84 and 0.73 mAU, and the drop was therefore of 0.11 mAU after the addition of hydrogenase. The reduction with DT decreased the absorbance to 0.58 mAU. Overall, the absorbance decreased by 42% after the addition of enzyme, when considering the value observed

after the addition of DT as the maximum reachable. Despite not being a great percentage, this number is close to 50%, which could suggest an equilibrium between the oxidised and reduced cluster being achieved.

To further explore the activity of the cluster maquette in presence of the hydrogenase, we performed a series of experiments where the cluster was exposed to different concentrations of enzyme. This investigation was begun by adding a smaller volume of hydrogenase solution to the sample containing the cluster maquette sparged with hydrogen. In the previously reported experiments, we used 2  $\mu\text{L}$  of a 0.2 mg/mL, or 3.05  $\mu\text{M}$ , solution of enzyme. Figure 97 shows the time points taken at 420 nm for the standard experiment and for an experiment where we used 0.2  $\mu\text{L}$  of a 0.2 mg/mL solution of hydrogenase.

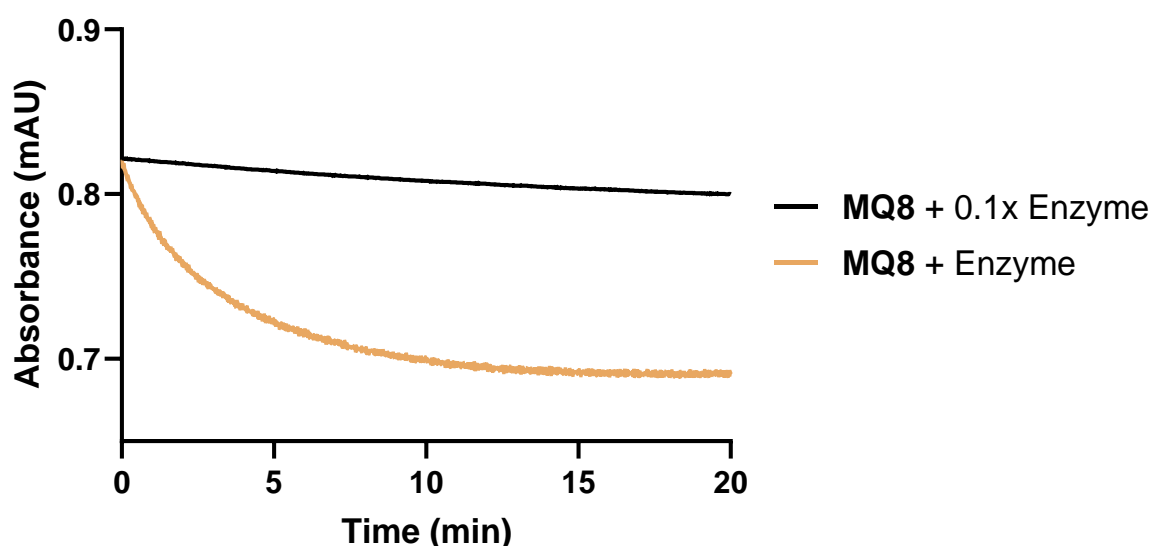


Figure 97 – Comparison of the absorption trend registered from time zero to 20 minutes at 420 nm for **MQ8** sparged with  $\text{H}_2$  and treated with 0.1-fold (black trace) and 1-fold (gold trace) of enzyme.

The cluster maquette treated with 0.1-fold the concentration of enzyme seemed not to be reduced over the course of twenty minutes. The absorbance decreased from 0.82 mAU to 0.80 mAU, for an overall drop of 0.02 mAU. It is plausible that the concentration of enzyme was too low to initiate the cycle and most of the cluster population remained in the oxidised state. On the other hand, the standard sample registered a drop of 0.13 mAU, from 0.82 mAU to 0.69 mAU. Because the experiment where we decreased the concentration of hydrogenase did not produce any interesting results due to poor conversion to the reduced cluster, we moved on to explore how

the system behaved under incrementing concentration of enzyme. We started by adding twice the amount of enzyme, and therefore 4  $\mu\text{L}$  of a 3.05  $\mu\text{M}$  stock solution of enzyme. The gold trace showed in Figure 98 represents the standard concentration of enzyme (7.6 nM), whereas the light-blue trace represents the twice as much concentrated (15.3 nM). Despite the start and end point being similar between one another, the rate at which the absorbance decreased was much faster for the sample treated with twice the amount of enzyme. After five minutes the absorbance registered for the standard sample was 0.72 mAU, whereas the absorbance for the twice as concentrated sample was 0.69 mAU. The plateau for the first experiment was reached after approximately 13.5 minutes, whereas it was reached in 6.5 minutes in the second case.

A similar trend was observed when the cluster was exposed to a concentration three times higher than the standard (22.9 nM), the trace of which experiment is coloured in red. The volume of enzyme stock solution used for this experiment was 6  $\mu\text{L}$ . Again, the absorbance decreased quickly during the first three minutes, before reaching a plateau. The cluster maquette exposed to three times the concentration of enzyme reached a plateau in absorbance after approximately 3.5 minutes. In this case, the absorbance registered after ten minutes from the beginning of the time-points analysis was 0.69 mAU and was already close to the lowest value registered for this sample. The final experiments we performed involved the addition of ten-times the concentration of enzyme with respect to the original experiment. In this case we exposed the cluster maquette to 20  $\mu\text{L}$  of a 0.2 mg/mL stock solution for a final concentration of enzyme of 76.3 nM. The drop of absorbance registered for this experiment (light grey trace) happened much faster than for the previous ones. A consequence of this is the time at which the plateau was reached, which was less than one minute after the enzyme was added to the cluster maquette sparged with hydrogen.

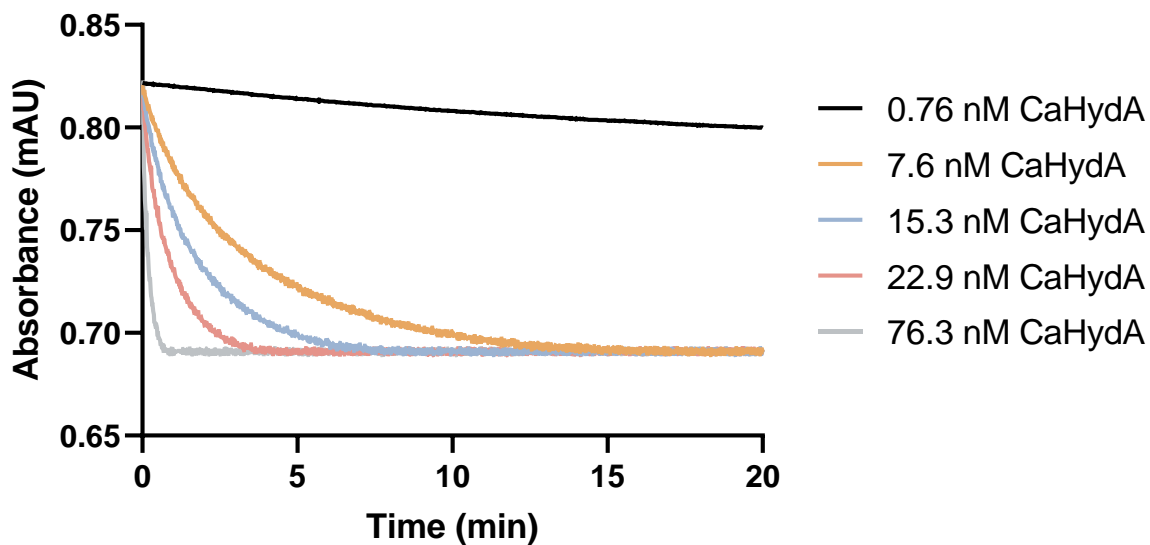


Figure 98 – Comparison of the absorption trend registered from time zero to 20 minutes at 420 nm for MQ8 sparged with H<sub>2</sub> and treated with various concentrations of enzyme.

Interestingly, apart from the experiment run using 0.1-times the concentration of enzyme, all the other entries reached the same value of absorbance over the course of the time-points analysis. It is plausible that the first entry would have also reached the same value provided the assay was run for an adequate period of time. It was hypothesised that increasing concentrations of enzyme would have led to higher reduction profiles for the cluster maquette and therefore lower values of absorbance. The fact that this hypothesis did not manifest could be explained by the reduction of the cluster maquette being governed by the overall environment in which the experiments were performed. Factors including pH of the buffer, concentration of hydrogen dissolved in solution, and concentration of the cluster maquette were all consistent within all these experiments and probably had a stronger influence on the reduction of the cluster. Finally, we explored the correlation between the enzyme concentration used in these experiments and the reaction rate for the reduction of the cluster. The reaction rate describes how quickly the cluster maquette was reduced, or how quickly the absorbance decreased, and is expressed in mAU per seconds. The resulting plot is displayed in Figure 99.

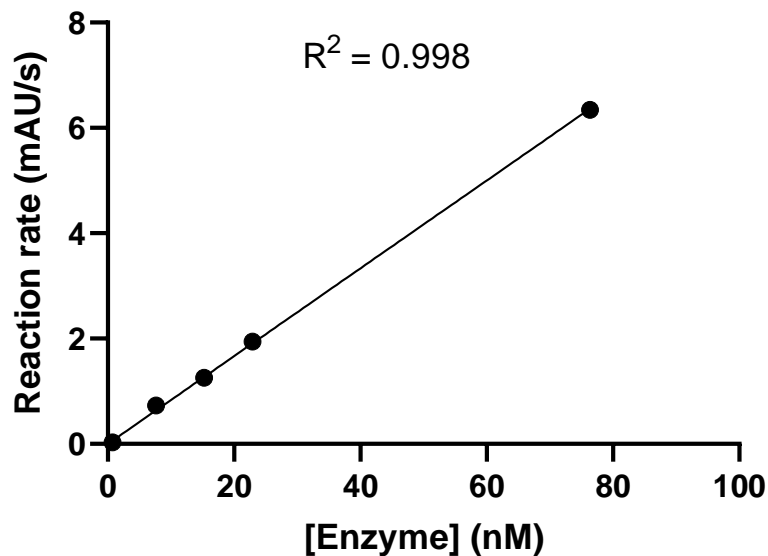


Figure 99 – Plotting of the reaction rate vs the concentration of enzyme.

The data points followed a linear trend meaning that the rate of reduction of the cluster maquette increases as the concentration of enzyme in solution is increased. The pace for the cluster maquette reduction is therefore adjustable by changing the amount of enzyme that we add to the reaction mixture. This result confirms that the synthetic maquette can directly interact with the enzyme resulting in productive electron transfer. The TOF for our electron transfer process was determined at a value of  $79 \pm 5.8 \text{ s}^{-1}$ , assuming a cluster concentration equal to that of the peptide concentration of the sample. Based on the comparison of the EPR data for the reconstitution of **MQ4** with  $\beta$ ME and **MQ8** without  $\beta$ ME, the extent of **MQ8** cluster reconstitution appears to be approximately 40% relative to **MQ4**. When this correction is taken into consideration, the TOF is closer to  $31 \pm 5.8 \text{ s}^{-1}$ , with respect to  $901 \pm 225 \text{ s}^{-1}$  for the original redox partners. For this reason, we can conclude that the TOF our semi-synthetic redox system is at least one order of magnitude lower than the physiological redox partnership that take advantage of two [4Fe-4S] clusters.<sup>118</sup>

#### 4.4 CONCLUSION AND FUTURE WORK

In conclusion, we described the interactions between a synthetic cluster maquette and a hydrogenase in a hydrogen-rich environment. The challenges encountered during the purification of the cluster maquette using SEC columns were overcome by

changing the reconstitution procedure and avoiding the use of  $\beta$ ME. The absence of the reducing agent did not affect the formation of the cluster maquette and revealed the robustness of these systems. We proceeded by describing the utilisation of a synthetic cluster maquette in a hydrogen-powered biological pathway. The electron transport chain that we explored links the reduction of a cluster maquette to the oxidation of a hydrogenase, which is reduced by the oxidation of molecular hydrogen to hydrogen ions. We argue that explorations of this kind could find industrial applications in the market that grows around hydrogen and its use as a more environmentally friendly source of energy.

In the future, further exploration in the synthetic peptide sequences could lead to new cluster maquettes able to achieve the reduction of a higher portion of the cluster population. This could have a positive impact on the efficiency of the electron transport chain and on the amount of hydrogen oxidised. The use of different reducing agent to improve the cluster formation in the reconstitution step could also be explored. However, purification of the synthetic maquette from the latter would be vital to ensure the electron transport chain is not affected by any residuals. The use of a maquette reconstituted using a Sec-containing peptide would also be of interest, especially if it produces different trends in the rate of reaction.

## 5 VISIBLE-LIGHT-MEDIATED CYSTEINE DESULFURISATION AND $\beta$ -ARYLATION IN PEPTIDES

### 5.1 INTRODUCTION

Bioconjugation refers to the modification of an amino acid within a peptide or a protein to form a stable covalent link with a chemical moiety. This technique aims at modifying the affinity, stability, and physicochemical properties of the biomolecule of interest to enable its application in various fields. Examples of functionalities in which modified biomolecules can be implemented are imaging of specific biomarkers, delivering of drugs to targeted cells, revealing of enzyme functions, and determination of proteins biodistribution.<sup>124</sup> It is vital for the bioconjugation strategy to maintain the structure of bioactive molecules intact. In their native cellular environment, peptides and proteins are often folded into secondary and tertiary structures. These intricate and delicate arrangements can be easily disrupted by fluctuations in the pH and temperature, and by the activity of reagents. Breakdown of the structure of the biomolecule eventually lead to complete loss of its function.<sup>125</sup> It is therefore mandatory for bioconjugation methodologies to implement mild and biocompatible reaction conditions that involve aqueous buffered solvents, ambient temperature, atmospheric pressure, and physiological pH. In addition, an ideal bioconjugation strategy should be chemo- and regioselective to ensure only the targeted residue of the peptide is modified in the desired conformation. Finally, it is key that bioconjugation strategies lead to stable conjugates, proceed with fast reaction kinetics, and avoid formation of toxic by-products.<sup>126</sup> For this reason, many strategies have been developed in past and recent years to improve and expand the knowledge in this area.<sup>127</sup>

### 5.2 AIMS AND OBJECTIVES

Inspired by the present literature that aims at Cys modification in peptides, we noticed that there were few methodologies that targeted the direct arylation of polypeptide scaffolds. We therefore developed a strategy that allows for the arylation of Cys in  $\beta$ -position via desulfurisation. This protocol allows for the formation of a stable C(sp<sup>3</sup>)-C(sp<sup>2</sup>) bond via visible-light-mediated desulfurisation, it is selective for the residue

Cys, and it is carried out in an aqueous buffer/organic solvent mixture. The utility of this methodology lies in extending the scope for Cys modification in peptides. Biologically relevant molecules (i.e., fluorescent tags, drugs carrier, biomarkers etc.) could be appended to the targeted residue before desulfurisation to suppress its reactivity. This strategy also allows for the insertion of non-canonical or modified amino acids. Finally, the C-C bond formed with this strategy is more stable in physiological conditions than the thioether moiety obtained with non-desulfurisation methods, allowing for a more stable modification.

## 5.3 METHODOLOGIES FOR THE MODIFICATION OF CYSTEINE

### 5.3.1 Introduction to Bioconjugation Strategies with Cysteine

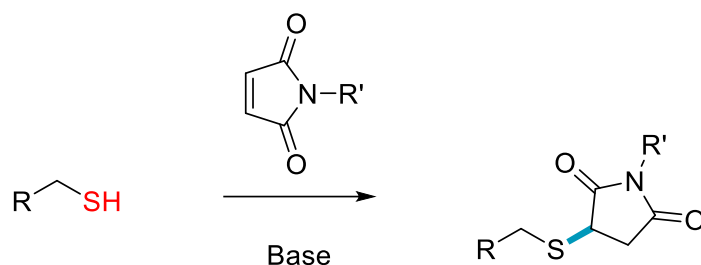
Most bioconjugation techniques involving peptides rely on the nucleophilicity or acid-base nature of the amino acid side chains, such as those of Cys and Lys, to introduce a chemical modification. Particularly, Cys is one of the most targeted residues among the endogenous amino acids in peptides and proteins. The relatively low natural abundance of Cys in the proteome (1-2%), means that modification on this residue allows for higher selectivity and less chances of isoform mixtures. Moreover, Cys displays a unique reactivity with respect to other amino acid side chains, given by the relatively high nucleophilicity of the thiolate moiety ( $pK_a = 8.2$ ).<sup>128</sup> Together, these characteristics of Cys and its thiol moiety, render it suitable for both ionic and radical reactions, with a constant interest for milder and more selective strategies for its continuous growth in bioconjugation chemistry. A wide range of reactions is described in the literature and include nucleophilic substitution<sup>128–130</sup> and addition<sup>131–137</sup>, thiol-ene chemistry<sup>138</sup>, and metal-free<sup>139</sup> or transition-metal-catalysed arylation<sup>140</sup>.

### 5.3.2 Michael Addition Strategy

Michael addition using maleimides and other electrophiles, is one of the most investigated strategies for modifying the Cys thiolate into a thiosuccinimide bond.<sup>128</sup> A broad range of methodologies have been proposed and some maleimide-thiol

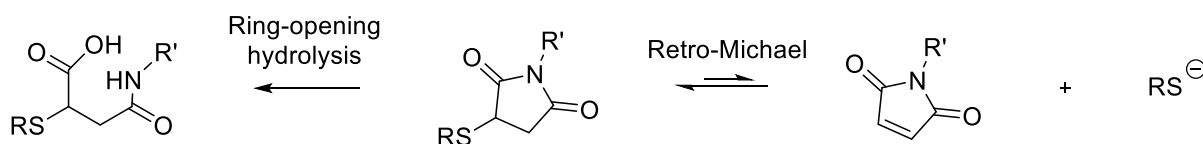


conjugation products are also currently present on the market (i.e. Cimzia, an antibody-drug conjugate used for the treatment of Crohn's disease).<sup>141</sup> Scheme 17 represents the general reaction involving a thiol and a maleimide reagent to afford the thiosuccinimide product.



Scheme 17 – Cys-mediated modification with maleimide reagents.  $R'$  = metalloenzymes, fluorescent tags, radio labels, polymers, drugs, lipids, and albumin-binding moieties.

Succinimide thioether bond formation between maleimides and Cys offers many advantages including the high selectivity, the solubility of all reagents in aqueous solvents, and the lack of by-products formation. Furthermore, maleimide reagents can be functionalised on the nitrogen atom to install a broad range of conjugation partners, as highlighted in Scheme 17, starting by appending an alkene and functionalising it.<sup>128</sup> Problems with the sensitivity of the succinimide thioether moiety were highlighted especially in reducing environments and near-physiological conditions, where retro-Michael addition and hydrolysis of the bond could affect this bioconjugation strategy, as showed in Scheme 18.<sup>128</sup>



Scheme 18 – Ring-opening and retro-Michael of maleimide-thiol conjugates.

In particular, the retro-Michael pathway results in free maleimide in solution that was observed to form conjugates with albumin.<sup>142</sup> Conversely, ring opening of the maleimide moiety could be triggered by hydrolysis. In this case, however, the product formed is stable and does not release any free drug. These two downsides of the Michael addition pathway could also be limited by positioning specific substituents on the ring that can affect the pKa of the protons next to the carbonyl groups.

### 5.3.3 Photocatalytic Modification of Cysteine

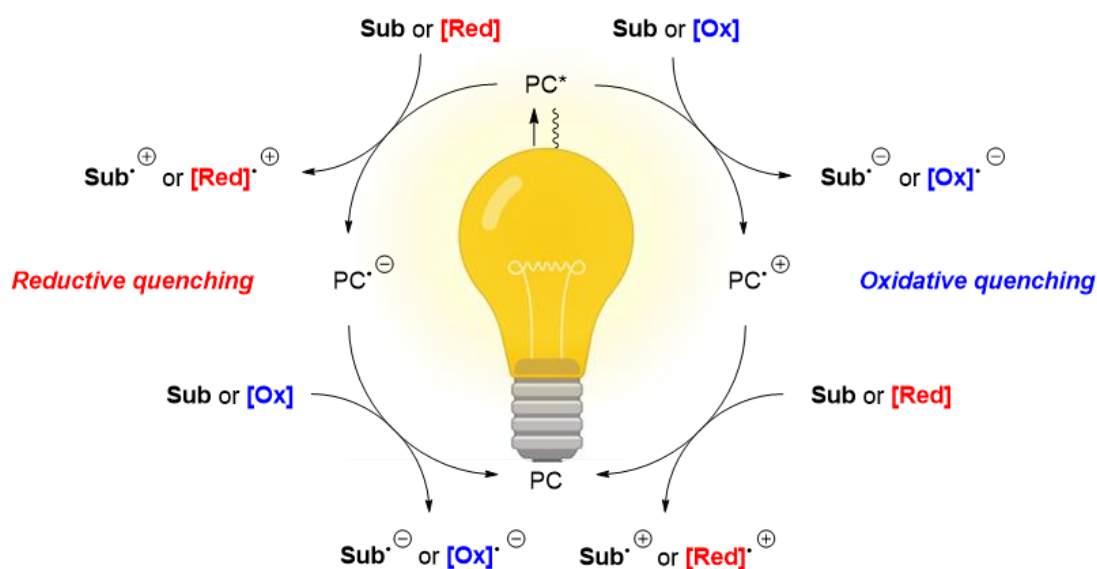
#### 5.3.3.1 Introduction

During the past two decades, photocatalysis has been widely investigated to achieve chemical transformations in organic synthesis. This, along with the growing interest in peptides as drug candidates, is leading to a constant rise in photocatalytic strategies for the incorporation of chemical moieties in these bioactive molecules.<sup>143</sup> One of the advantages of photocatalysis lies in the use of milder conditions to activate the organic substrate. More traditional radical processes often require high energy reagents or initiator to trigger the radical formation, which can potentially lead to unwanted reactivity on other sites of the bioactive molecule. In addition, most organic substrates are not capable of absorbing light in the visible range and possess energy barriers too high to be overcome at room temperature, which makes photocatalytic reactions highly selective and with little chances for side reactions.<sup>144</sup> Another advantage of photochemistry is given by the application of visible-light irradiation as a green and traceless reagent that helps preserving the fragile architecture of the bioactive substrates.<sup>145</sup> Finally, the strong dependence of photocatalytic reactions on the photon flux, makes their reaction kinetic easily controllable by simply turning off the light radiation.<sup>146</sup> The ease with which photochemical reactions can be controlled and quenched makes them the ideal strategy for bioconjugation methodologies, and drastically simplify the purification step by circumventing the need for quenchers in the reaction mixture.

#### 5.3.3.2 Principles of Photocatalytic Modifications

In photochemistry, the organic substrate (Sub) (i.e., a Cys residue in a peptide), is activated, and subsequently modified, by the ability of a photocatalyst (PC) to be excited by visible light and convert the absorbed energy into an electrochemical potential. The PC in its excited state is able to engage in a single electron transfer (SET) the substrate by acting as either an electron donor or acceptor, depending on the respective redox potentials. This step forms a radical on the organic substrate and activates it towards the desired reactivity.<sup>144</sup> Most catalytic cycles begin with the visible-light-mediated excitation of a PC, which can be either an organic or a transition-

metal-based molecule (Scheme 19). Different PCs possess different absorption maxima, and it is therefore crucial to select the most appropriate source of visible light to optimise the PC response.<sup>147</sup> The excited state photocatalyst ( $PC^*$ ) is quenched by following a SET pathway that results in either reduction or oxidation of the substrate. In oxidative quenching,  $PC^*$  transfers an electron to the substrate (or an oxidant), generating the oxidised photocatalyst  $PC^{+\bullet}$ . Conversely, in reductive quenching the excited PC is quenched by accepting an electron from the substrate (or a reductant), and the reduced photocatalyst  $PC^{\ominus\bullet}$  is formed. At this point a radical is formed on the substrate (i.e., the substrate is activated), and it can undergo reactions such as atom abstraction, nucleophilic or electrophilic attack, and bond cleavage.<sup>126</sup>

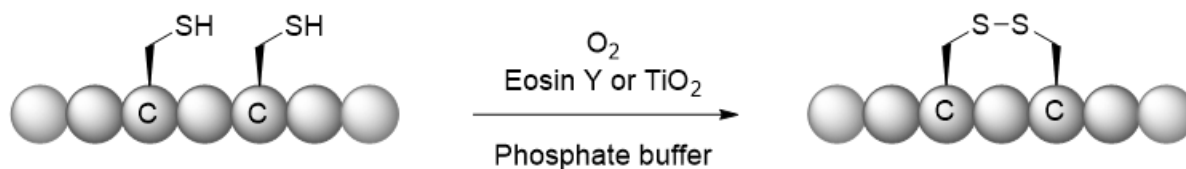


Scheme 19 – Reductive and oxidative quenching cycle of a PC.<sup>126</sup>

### 5.3.3.3 Photocatalytic Oxidation of Cysteine

Photochemistry has been widely implemented to catalyse reactions involving the chemical modification of Cys that previously involved too harsh conditions to be used for bioconjugation strategies. An example in this context is the formation of disulphide bonds; a crucial tool for integrity, folding, and oxygen sensing in peptides and proteins. Traditionally, disulphide bonds formation in peptides required strong oxidising agents (i.e., iodine or potassium ferricyanide) and/or harsh conditions (i.e. acidic pH) to be successful.<sup>148</sup> Due to these strong conditions, orthogonal protection of Cys residues is frequently needed to limit the reactivity of other functionalities on the peptide chain,

therefore adding extra steps to the strategy. Greener conditions were developed as well and included the use of DMSO or atmospheric oxygen to catalyse the oxidation of intramolecular Cys residues in peptides.<sup>149,150</sup> Limitations of these two strategies included the complication of removing DMSO upon completion of the oxidation, and the slow reaction rates when using molecular oxygen to catalyse disulphide bonds formation. Noel and coworkers developed a strategy that circumvented these limitations by adopting milder conditions (visible light, ambient temperature, and neutral-pH buffer solutions), and by accelerating the reaction rates via the use of either Eosin Y or TiO<sub>2</sub> as PCs (Scheme 20).<sup>151,152</sup> Examples of biologically relevant peptides presenting a disulphide bond obtained with this strategy were the neuropeptide and peptide hormone oxytocin, the thiuram disulphide therapeutic disulfiram, and the reduced form of the lipid-binding peptide y1fatc.<sup>151,152</sup>



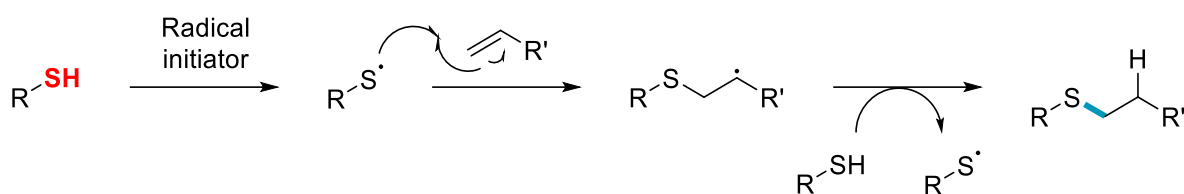
*Scheme 20 – Disulphide formation in Cys-containing peptides catalysed by Eosin Y or heterogeneous TiO<sub>2</sub> PC.*

This strategy showed selectivity towards Cys when performed on fully deprotected peptides, and great tolerance with sensitive amino acid residues (Glu, Arg, Tyr, Trp, His). The method was performed in flow to achieve faster reaction rates and improved irradiation of the reaction mixture. As a limitation to this method, due to the high sensitivity towards the gas-to-liquid mass transfer effect between the solution and atmospheric oxygen, the reaction cannot be easily scaled up.

#### 5.3.3.4 Photocatalytic C-S Bond Formation with Cysteine

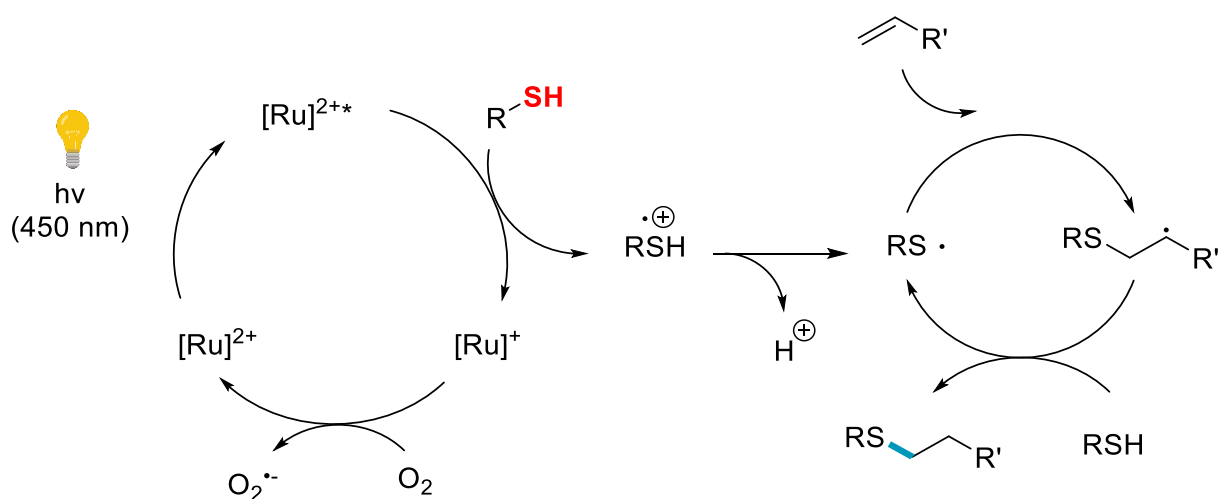
Another widely investigated bioconjugation tool for Cys-containing peptides regards alkylation strategies. Among this broad family of reactions, the reactivity between thiols and alkenes has been deeply studied as a click approach with numerous advantages including having rapid reaction rates, being regioselective, allowing for a broad scope with numerous commercially available alkenes, and being insensitive to oxygen and water.<sup>138</sup> The radical thiol-ene reaction represents one of the most common strategy

to install a thioether moiety and results from the anti-Markovnikov addition of a thiol on an olefin (Scheme 21).<sup>153</sup>



Scheme 21 – Proposed mechanism for the photocatalytic thiol-ene reaction of Cys residues.

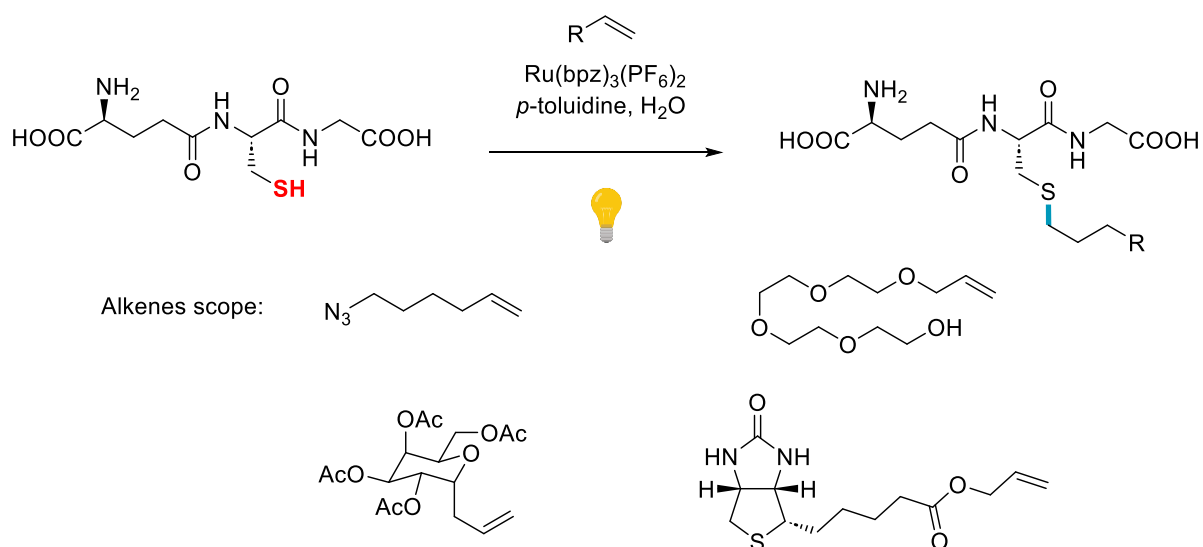
The proposed mechanism for the radical thiol-ene reaction catalysed by transition metal PC is showed in Scheme 22. Photoexcitation from a visible-light source promotes a MLCT that afford an oxidised excited state of the ruthenium catalyst. This species rapidly undergoes reductive quenching with the thiol to generate a thiyl radical cation and the reduced catalyst. The radical cation is deprotonated to afford the thiyl radical, which subsequently adds across the olefin in an anti-Markovnikov orientation. The alkyl radical abstracts a hydrogen atom from another thiol to afford the hydrothiolated product and another equivalent of thiyl radical. Alongside this photocatalytic cycle, the reduced catalyst is most likely re-oxidised to its ground state by a molecule of atmospheric oxygen.<sup>154</sup>



Scheme 22 – Mechanism of the thiol-ene photocatalytic reaction catalysed by ruthenium.

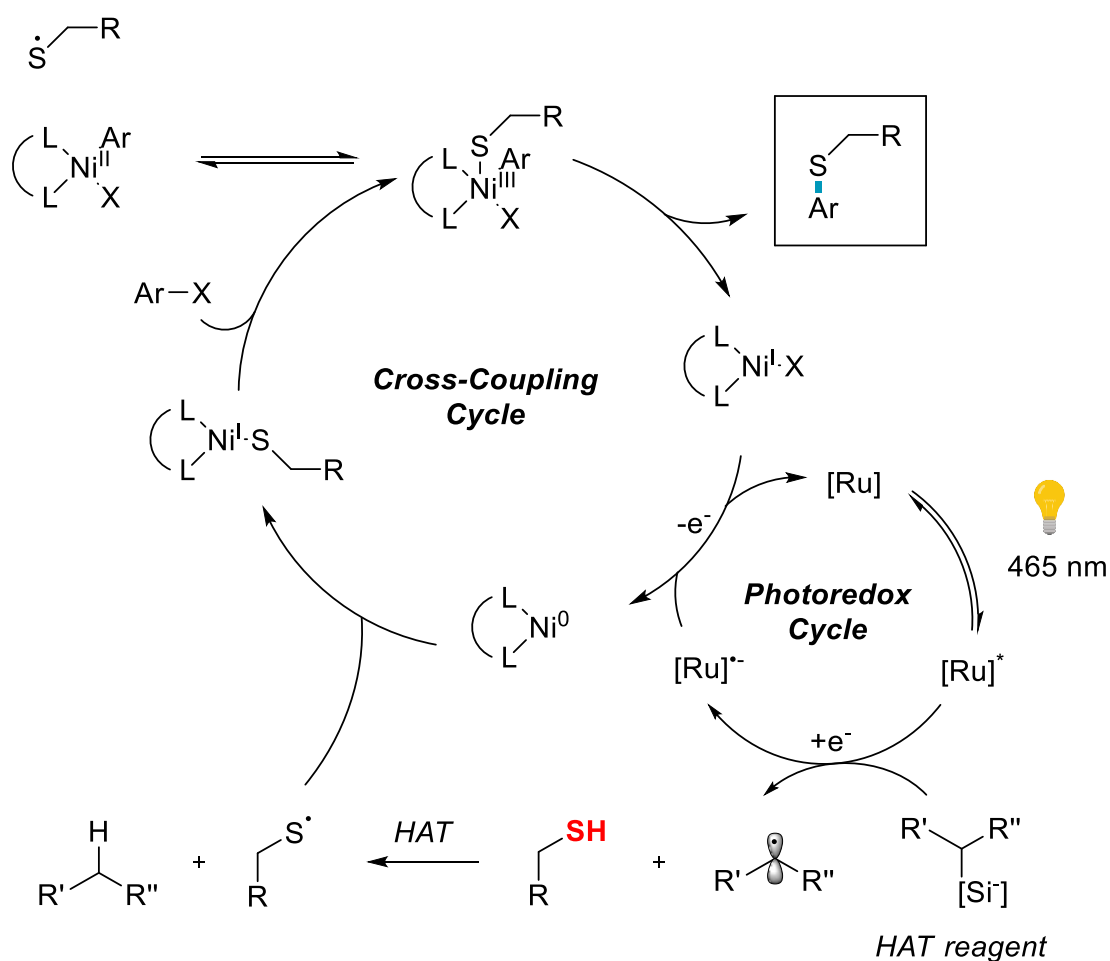
Numerous photocatalytic approaches for the C-S bond formation using the thiol-ene strategy have been developed over the past years. Tyson *et al.* in 2014 proposed a visible-light-mediated hydrothiolation of olefins that proceeds in water using a ruthenium PC.<sup>153</sup> The suitability of this method for bioconjugation was tested by reacting a series of biologically relevant olefins partner with glutathione, a tripeptide

formed by a Cys coupled onto a Gly residue and the carboxyl group of glutamate side chain that is an endogenous antioxidant found in most living cells.<sup>155</sup> As showed in Scheme 23, the examined coupling partners included alkene modified polyethylene glycol (PEG) oligomers, biotin, protected sugars, and azides. A limitation of this methodology is that it requires p-toluidine as an essential redox mediator to circumvent the slow reaction rate of the direct photooxidation of the thiol with the ruthenium catalyst. The thiyl radical cation step is therefore omitted, and the thiyl radical is directly generated by a more rapid electron transfer event via a p-toluidine radical cation.<sup>153</sup>



*Scheme 23 – Visible-light-mediated photocatalytic coupling of glutathione with different alkenes in aqueous media.*

Another useful strategy to label Cys residues is given by arylation of its thiol moiety. In this context, Vara et al. developed a transition-metal catalysed Cys arylation procedure and tested its suitability for bioconjugation purposes on glutathione.<sup>156</sup> As showed in Scheme 24, the photochemical cycle is initiated by photoactivation of the ruthenium PC, which, in presence of a hydrogen atom transfer (HAT) silicate reagent, generates the required thiyl radical. Precisely, homolytic cleavage of the C-Si bond upon oxidation of the silicate leads to a carbon-centred radical that abstracts a hydrogen atom from the Cys thiol to generate the thiyl radical. This radical is trapped by the Ni(0) species and enters the cross-coupling cycle. After oxidative addition of the aryl bromide, the reductive elimination steps yield the desired thioarylated product with the newly installed C(sp<sup>3</sup>)-S bond.



Scheme 24 – Proposed mechanism for the Ni/photocatalytic cycle and thioarylation reaction.

Aryl bromides were used in this procedure as coupling partners in the oxidative addition step, allowing for a wide scope of arylated products. It was found that various aryl halides greatly affected the final yield of this reaction, as the scope ranged from 14% (with a para-alcohol) to 83% (with 4-bromobenzyl boronic acid). A variety of functional groups decorating the aryl bromide coupling partners were tolerated (carboxylate groups, alcohol, ether, sulphonamide, amide, and nitrile). This thioarylation method was successfully applied to a 9-mer, fully deprotected peptide containing an internal Cys and an array of amino acids with a reactive sidechain (Trp, His, Glu, and Tyr), even if lower peptide concentrations were required for the full conversion. A limitation of this strategy was found to be the use of toxic and non-green DMF as solvent, and water was tolerated only up to 2% by volume.<sup>157</sup> A different photocatalytic approach to Cys arylation in peptides was proposed by Noel *et al.* in 2017.<sup>158</sup> In this work, aryl radicals were generated from benzene diazonium salts in presence of the PC Eosin Y. The thiol moiety of the Cys residue was then used to trap the electrophilic aryl radical and yield the thioarylated product. In this case, the radical

is not formed on the thiol group of Cys (i.e., as a thiyl radical), but it is generated on the coupling partner. The diazonium salts are generated *in situ* in presence of *tert*-butyl nitrate, *p*-toluenesulfonic acid, and an aryl amine. A wide range of amines containing electron deficient and electron donating groups were tolerated. This procedure was performed on a model peptide in biologically-like conditions (aqueous phosphate buffer at pH = 8.0) and afforded the arylated peptide after an irradiation time of 30 minutes. Interestingly, the presence of other nucleophilic sites on the peptide (Lys and Ser residues side chain) did not affect the modification, proving the selectivity of this strategy. Limiting this procedure could be the fact that due to the highly reactive conditions required for the formation of diazonium salts *in situ*, pre-made diazonium salts were required to preserve the delicate nature of the peptide.

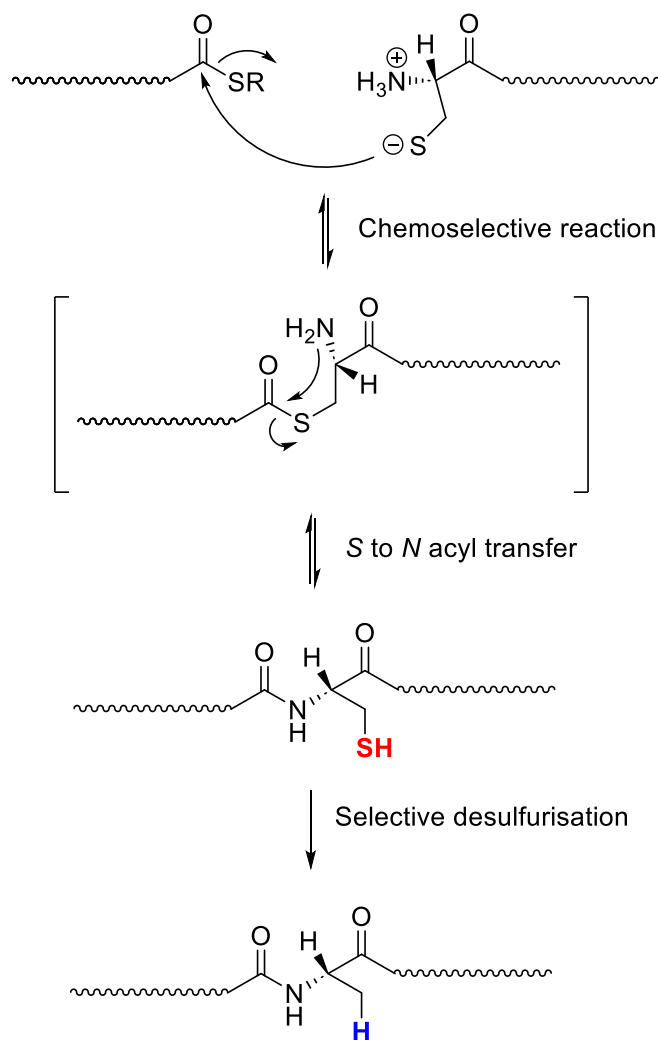
Another example of Cys arylation was proposed by Beard *et al.* in 2019, when they developed a method for the ligand-directed labelling of a Cys residue on a protein surface.<sup>159</sup> The biomolecule of interest was an anti-apoptotic BCL-2 protein and current anticancer target, human MCL-1. The strategy they adopted was to attach the ruthenium PC to the *N*-terminus of a 19-mer peptide known to successfully bind the target protein. The mechanism of reaction was not proposed, but in similar reaction pathways, the excited state of the ruthenium PC loses an electron towards a sacrificial oxidant (ammonium persulfate or atmospheric oxygen), and catalyses the reaction between the amino acid residues (i.e. Cys or Tyr) and an electron-rich dimethylaniline that acts as a radical trapping agent.<sup>160</sup> In this example, Bon and co-workers were able to functionalise the aniline with a fluorescent and a biotinylated label. This strategy was proved to be working smoothly even in a mixture of structurally related proteins, highlighting the selectivity for the Cys residue on the targeted MCL-1 protein. Finally, it was noted that complete conversion could not be achieved and irradiation for a time longer than 1 minute led to increased concentration of oxidised protein species, possibly resulting from oxidation of the protein by the PC.

#### 5.3.3.5 Cysteine Desulfurisation in Peptides

A different tool in bioconjugation involves the chemical ligation between two peptides bearing one an *N*-terminal unprotected Cys residue, and the other a *C*-terminal



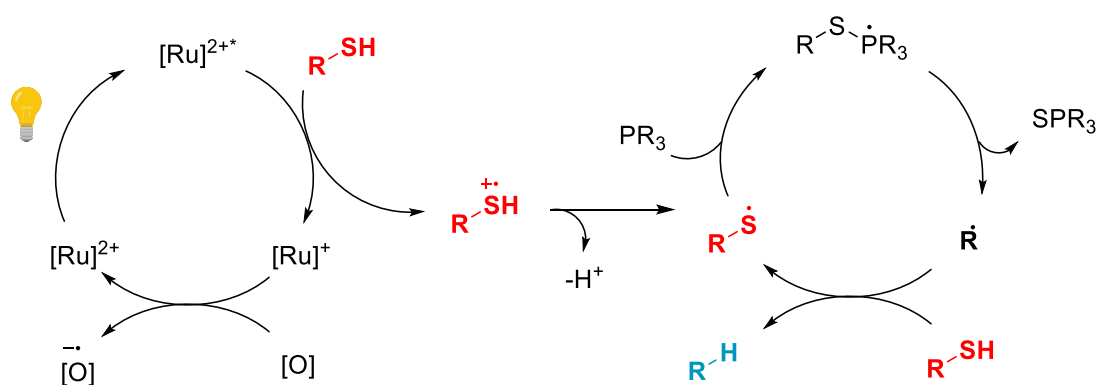
thioester.<sup>161</sup> This is one of the most widely used condensation technique for peptide segments to synthesise longer peptides or proteins, and involves a transthioesterification and S-to-N acyl transfer steps (Scheme 25).



*Scheme 25 – Native chemical ligation and desulfurisation reaction.*

The free thiol of a Cys residue is a key component of this ligation technique. However, as described earlier, Cys has a low abundance and most proteins either do not contain this residue or, if it is present, it might not be located in a region suitable for a ligation site. For this reason, desulfurisation methods are of great importance after the ligation step, to convert the no longer desired Cys residue into Ala. Ala is more abundant in the proteosome, and therefore gives a greater choice of ligation function. Precedent desulfurisation methods employed palladium or Raney nickel under a stream of H<sub>2</sub>. However, such harsh conditions were generally not compatible with specific functional groups and led to low recovery of the desired product.<sup>162</sup> An alternative protocol was developed in 2007 and involved the use of a phosphine and a radical initiator to

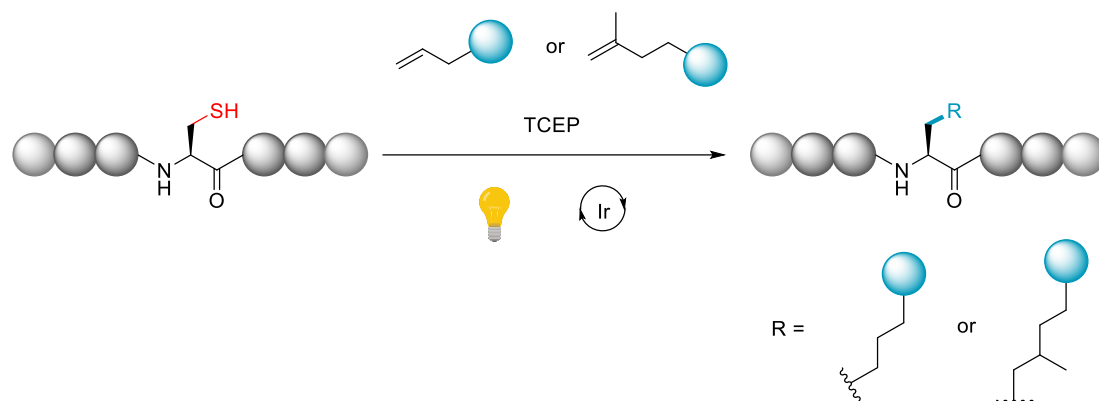
desulfurise a range of peptides containing a Cys residue. The issues with this method involved the high temperature (between 37°C and 65°C) required for the radical initiator to be activated, and an excess of tris(2-carboxyethyl)phosphine (TCEP) to drive the reaction to completion.<sup>163</sup> Gao *et al.*, developed a visible-light-induced desulfurisation method for cysteinyl peptides that is catalysed by a ruthenium PC and triphenylphosphine-3,3,3-trisulfonic acid.<sup>162</sup> This strategy proceeds in aqueous or aqueous/organic solvent mixtures at room temperature to convert Cys residues into Ala. These milder conditions were suitable to different functional groups and were successfully implemented in the ligation of poly- and glycopeptides. The reaction mechanism was proposed based on previous studies on phosphine radical traps (Scheme 26). Firstly, the ruthenium PC is excited by accepting a photon from the visible light. The activated ruthenium catalyst intermediate is then reduced by the thiol moiety to produce the reduced ruthenium catalyst and the thiyl radical cation. As previously seen for the thiol-ene reaction mechanism, the thiyl radical cation is deprotonated to yield the thiyl radical. At this point, the thiyl radical is attacked by the phosphine to form a phosphoranyl radical, which undergoes  $\beta$ -scission to produce an alanyl radical. Finally, the alanyl radical is protonated by a second thiol group to yield the desulfurised product and a further equivalent of thiyl radical. Concerted with this catalytic cycle, the reduced ruthenium catalyst is re-oxidised to its starting ground state by either molecular oxygen or other oxidative species.<sup>162</sup>



Scheme 26 – Proposed mechanism for the Ru-catalysed desulfurisation of Cys residues.

In our research group, two methodologies for the desulfurative C-C bond formation on Cys residues were recently investigated.<sup>164,165</sup> Both works described the protocol for a light-mediated reaction that enabled the site-selective modification of peptides and proteins via desulfurative C(sp<sup>3</sup>)-C(sp<sup>3</sup>) bond formation (Scheme 27).<sup>165</sup> The mechanism proceeds via formation and interception of an alanyl radical using the

water-soluble phosphine TCEP to favour the desulfurisation process via a phosphoranyl radical.<sup>163,166</sup> Various radical traps were investigated that allowed the installation of various moieties, including biologically relevant biotin.



Scheme 27 – Cys-selective peptide/protein modification via desulfurative C(sp<sup>3</sup>)-C(sp<sup>3</sup>) bond formation.

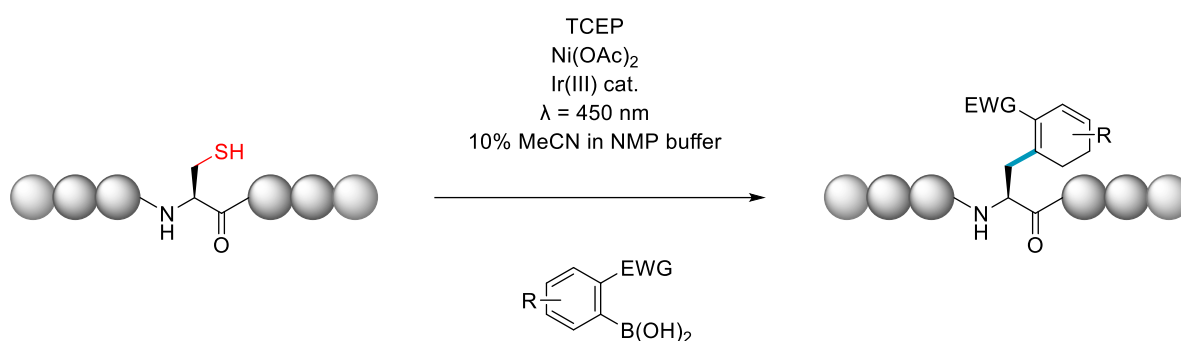
These bioconjugation methodologies were successfully applied for the modification of small proteins, respectively ubiquitin<sup>165</sup> and a histone H4 protein.<sup>164</sup>

## 5.4 RESULTS AND DISCUSSION

### 5.4.1 Introduction

In this chapter, the selective visible-light-mediated arylation via desulfurisation for Cys-containing small peptides is described (Scheme 28). In this strategy, a new C(sp<sup>3</sup>)-C(sp<sup>2</sup>) bond is formed upon desulfurisation of the Cys thiol with the aim of expanding the chemical diversity in the proteome and provide a new tool in post-synthetic modification of peptides. Previous studies in the group involved the trapping of desulfurisation with persistent radical (2,2,6,6-tetramethylpiperidin-1-yl)oxyl (TEMPO) that allowed for the post-synthetic installation of various moieties appended to the radical traps.<sup>165,166</sup> The coupling partners chosen for this  $\beta$ -arylation strategy were phenyl boronic acids. The feasibility of using aryl boronic acids in a nickel-catalysed arylation reaction of Cys was previously showcased by Hanaya *et al.* in 2019.<sup>167</sup> In their work, this Cys modification tool was successfully tested on various peptides (6- to 15-mer) and on a protein (bovine serum albumin) in an *N*-methylmorpholine buffer (10 mM, pH = 7.5). Boronic acids are known to have good stability and solubility in

aqueous buffered solvents and are therefore considered to be suitable reagents for bioconjugation techniques.



Scheme 28 – General conditions for the desulfurisation and  $\beta$ -arylation reaction on Cys residues.

The initial conditions we screened are described in Scheme 28 and consisted of 5 mol % of Ir(II) PC and 5 equivalents of Ni(II) salt as co-catalyst, 25 equivalents of TCEP, and 5 equivalents of boronic acid as the coupling partner. Presence of the phosphine in the reaction mixture was required to achieve the desulfurisation of the Cys residue, as previously described by Gao *et al.*<sup>162</sup> The Ni(II) salt was proved to be necessary for catalysing the arylation step in previous work.<sup>156,167,168</sup> The solvent we selected for this method was a mixture of 0.01 M NMM buffer with 10% acetonitrile. The percentage of acetonitrile was required to fully solubilise the Ir(II) PC in the reaction mixture.

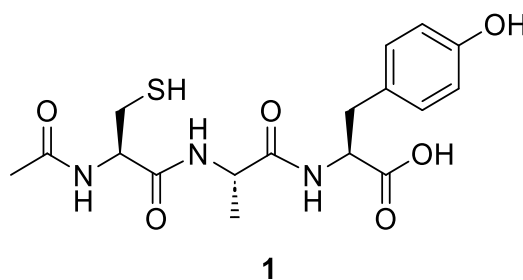


Figure 100 – Peptide sequence of 1.

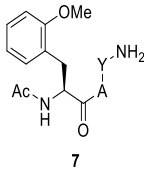
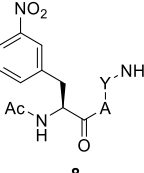
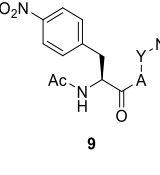
#### 5.4.2 Boronic Acid Screening

We began by screening a range of boronic acids bearing different functional groups in *ortho*-position. The substrate selected for this reaction was the model peptide Ac-CAY-NH<sub>2</sub> (**1**) (Figure 100). This small peptide provides a Cys residues required for the transformation, as well as a Tyr residue to introduce complexity to the molecule that can test the robustness and selectivity of the modification. Moreover, the Tyr residue

makes the purification of the peptide easier due to the absorption in the aromatic region.

Table 3 – Boronic acid screening and corresponding yields.

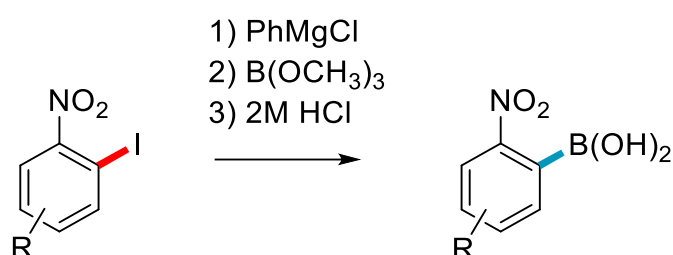
Entry	Boronic Acid	Product	Isolated yield (%)
1	2-Nitrophenyl	 2	32
2	2-Cyanophenyl	 3	25
3	2-Trifluoromethyl	 4	23
4	2-Carbamoyloxyphenyl	 5	16
5	2-Carboxyphenyl	 6	-

6	2-Methoxyphenyl	 7	-
7	3-Nitrophenyl	 8	-
8	4-Nitrophenyl	 9	-

The results presented in Table 3 showed that the nitro group in *ortho*-position to the boronic acid gave the best isolated yield (Entry 1, 32%). Other electron withdrawing groups in *ortho*-position also showed good yields with respectively 25% for the cyano group (Entry 2) and 23% for the trifluoromethyl group (Entry 3). The slightly activating amide group in 2-position to the boronic acid gave very poor conversion to the arylated-desulfurised product, with only 16% isolated yield (Entry 4). The carboxamide group has limited electron withdrawing power, which could explain the lower conversion towards the arylated product. Finally, the carboxylic acid group in *ortho*-position was unreactive, despite its inductive effect to deactivate the aromatic ring (Entry 5). These results were partially in agreement with previous reports, where the nitro group in 2-position proved to be the best electron deficient substituent.<sup>167</sup> The presence of the electron donating and activating methoxy group in *ortho*-position (Entry 6) hampered the reactivity of the substrate in this arylation reaction, confirming that electron withdrawing, deactivating groups were required for this reactivity to take place. Previous literature in a similar methodology reasoned the higher reactivity displayed by the nitro group with the need of an electron deficient  $\pi$  system as well as a  $\sigma$ -electron-withdrawing effect.<sup>167</sup> To further confirm that the electron withdrawing

group was required in *ortho*-position to the boronic acid, the nitro group was moved around the aromatic ring, in *meta* (Entry 7) and *para* (Entry 8) positions. Again, these two regioisomers were not reactive, proving that the deactivating substituent was required to be placed in 2-position to the boronic acid.

To further extend the scope of this methodology we synthesised a range of boronic acids starting from the corresponding aryl iodide, following a previous literature.<sup>169</sup> Briefly, the aryl iodide was reacted with phenyl magnesium chloride to achieve the I-Mg exchange, which was stable at temperatures below -40 °C. Further reaction with the electrophile trimethyl borate afforded the corresponding boronic acid (Scheme 29).

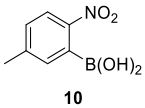
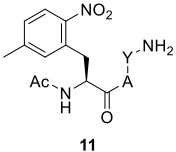
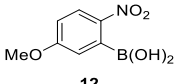
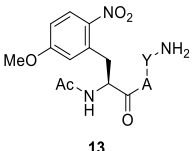
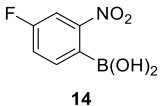
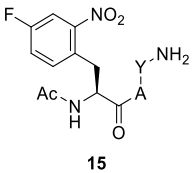


Scheme 29 – Synthesis of functionalised *ortho*-nitrophenylboronic acids.

The three functionalities that we selected were: methyl (**10**), methoxy (**12**), and fluorine (**14**). The methyl and methoxy groups were appended in *meta* position to the boronic acid, whereas the fluorine substituent was attached in *para* position. All these three modified boronic acids were successfully installed onto the model peptide Ac-CAY-NH<sub>2</sub> (Table 4), widening the scope of this proposed desulfurisation and arylation technique for small peptides.

Yields consistent with the other boronic acid explored previously were observed for the 5-methyl boronic acid (Entry 1, Table 4). Interestingly, the 4-fluoro-2-nitrophenyl boronic acid (Entry 3, Table 4) afforded the modified peptide **15** in higher yields than those observed for the 2-nitrophenyl boronic acid (Entry 1, Table 3). Finally, the 5-methoxy boronic acid (Entry 2, Table 4) gave poor yields for the desulfurisation and  $\beta$ -arylation modification, affording only 8% of the modified peptide **13**.

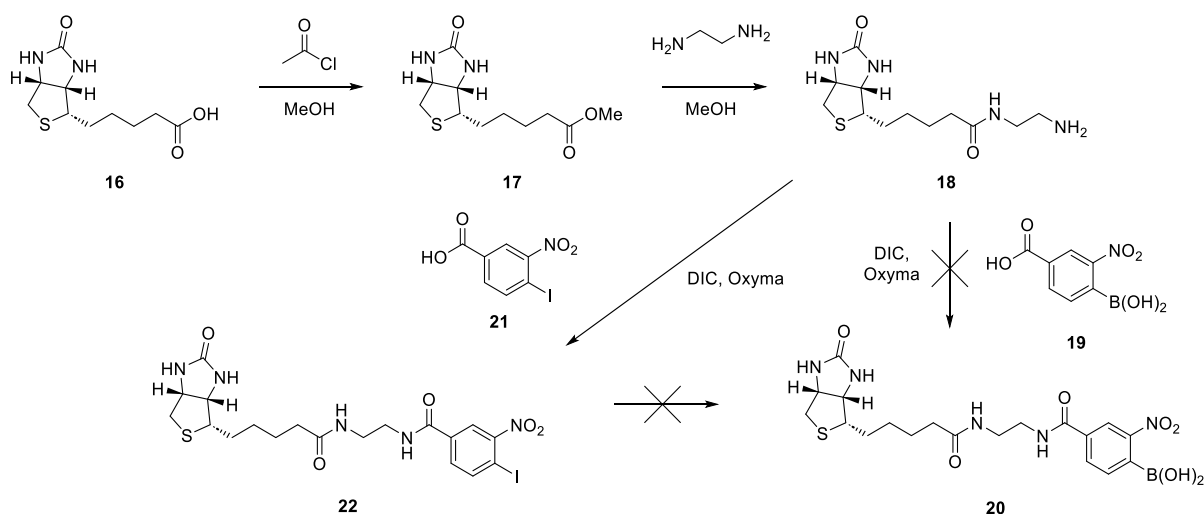
Table 4 – Boronic acid scope.

Entry	Boronic Acid	Product	Isolated yield (%)
1	 <p>10</p>	 <p>11</p>	25
2	 <p>12</p>	 <p>13</p>	8
3	 <p>14</p>	 <p>15</p>	33

In an effort to further extend the scope of this methodology towards more biologically relevant molecules, we opted for the synthesis of **19**. This molecule had the potential to extend the utility of this arylation technique to biologically relevant experiments, including protein modification. Biotin is indeed small in size and therefore unlike to disturb the natural activity of the biomolecule. Moreover, biotin presents an exceptionally strong binding affinity towards avidin and streptavidin, which makes it an ideal tool for isolating biotinylated molecules.<sup>170</sup> In these experiments, biotin is attached to the peptide or protein of interest, often to the amino group of the sidechain in a Lys residue, and then bond to streptavidin beads. The high affinity between biotin and streptavidin ensure that this interaction is fast, and with high specificity. Once the peptide or protein is subjected to the enzyme assay, the magnetic beads to which avidin or streptavidin is bound are easily separated from the rest of the solution, simplifying the purification and isolation of the biological molecule from other reagents. The high affinity with which biotin binds to avidin and streptavidin is of great interest because such a high affinity interaction is difficult to replicate via H-bonding and electrostatics in synthetic systems. These interactions arise between the valeric acid



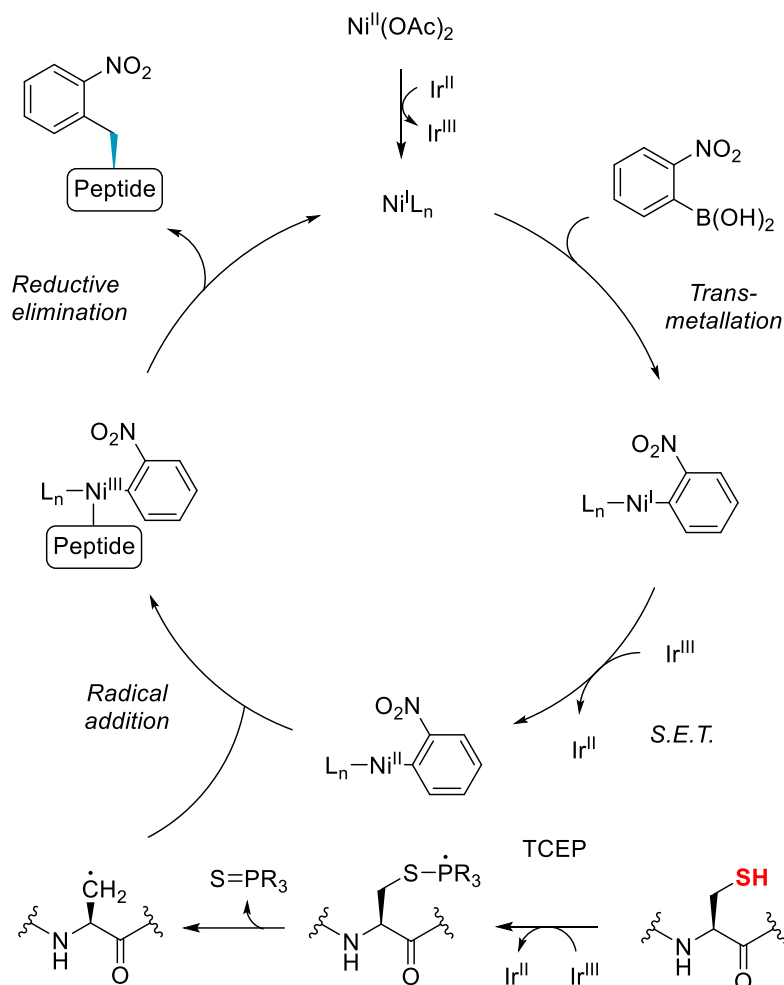
moiety of biotin and the lipophilic region of the protein binding pocket.<sup>171</sup> Issues were encountered during the synthesis of this analogue that prevented us from testing the feasibility of this reactivity in our proposed arylation and desulfurisation strategy (Scheme 30). Initially, Biotin (**15**) was reacted with *in situ* formed hydrogen chloride to convert the carboxylic acid into a methyl ester (**16**), to install the methoxy group as a good leaving group. This intermediate was reacted with ethylenediamine to form an amide bond (**17**), as previously described in a method for the condensation of  $\beta$ -ketoesters with amines.<sup>172</sup> In the next step, the primary amine of **17** was expected to react with the carboxylic acid of **18** and afford the expected product **19**, but this reaction was not successful and both high resolution mass spectrometry (HRMS) and nuclear magnetic resonance (NMR) only showed the unreacted starting materials. The reactivity was assumed to be hampered by the presence of the boronic acid on the aromatic ring of **18**, therefore **17** was instead reacted with the aryl iodide **21**, with the intent of installing the boronic acid moiety in a second instance. The modified biotin analogue **17** was successfully attached to the carboxylic acid of the aryl iodide **21**. However, when this was reacted under the conditions previously described in Scheme 29 to afford the corresponding boronic acid (**20**), the expected product was not obtained. Instead, HRMS showed two peaks that corresponded to the mass of the mono- and bis-boronic acid. Interestingly, the I-Mg exchange step did not occur, and the aryl iodide was apparently subjected to mono- and di-borination. Due to the lack of time, efforts could not be put into further expanding this or other analogues to widen the scope of this methodology.



Scheme 30 – Reaction scheme for the attempted synthesis of biotin ortho-nitrophenylboronic acid (**20**).

### 5.4.3 Proposed Mechanism

The proposed mechanism for this reaction is described in Scheme 31, and its formulation was based on prior studies of radical-mediated Ni(II)-catalysed reactions.<sup>162,173,174</sup> The Ni(II) salt is initially reduced to Ni(I) by the Ir(II) PC, which is oxidised to Ir(III). TCEP acts as a ligand for the reduced Ni species, which further binds to the boronic acid in the trans-metalation step. At this stage, a single electron transfer step regenerates the Ir(II) PC by oxidising the Ni(I) to Ni(II). Concurrently to this cycle, the Ir(III) PC forms a thiol radical cation on the Cys residue. This intermediate, upon deprotonation, forms a thiyl radical that reacts with TCEP to form a phosphoranyl radical. Rearrangement of this intermediate yields an alanyl radical and a sulfur phosphide, which thanks to the high dissociation enthalpy for the S=P bond is the driving force of this step. The alanyl radical is trapped by the Ni(II) species through the radical addition step. Finally, reductive elimination from the oxidised Ni(III) intermediate yield the final product and restore the Ni(I)-TCEP species.



Scheme 31 – Proposed mechanism for the Ir- and Ni-catalysed desulfurisation and  $\beta$ -arylation of Cys residues in small peptides.

## 5.5 CONCLUSION AND FUTURE WORK

In conclusion, a novel method for the desulfurative  $\beta$ -arylation of Cys residues in small peptides has been developed. This method uses TCEP as a desulfurative agent to generate a radical in  $\beta$ -position to the Cys residue. Low equivalents of the corresponding boronic acid are coupled via the activity of an Ir(II) PC and a Ni(II) co-catalyst. The reaction is performed in aqueous buffer with 10% of organic solvent at neutral pH and is completed in one hour. Initially, a study was performed to identify the most effective electron-withdrawing group that was required to be placed in *ortho*-position to the boronic acid. The nitro group was selected as the best candidate. Despite moderate yield, the installation of a range of non-canonical aromatic amino

acids was described, proving the feasibility of this strategy for the post-synthetic modification of small peptides.

In the future, the installation of different moieties could be attempted to further expand the utility of this method particularly towards more biologically relevant substrates. In an attempt to increase the moderate yields obtained so far, the Cys residue could be moved to the *C*- or *N*-terminus to test if this has an effect on the reaction. This strategy could be optimised and carried out on a protein to showcase its utility in a substrate with a higher level of complexity. Computational modelling studies could be implemented to further understand the mechanism of action behind this desulfurisation and  $\beta$ -arylation reaction. This information could be used to explore different ligands and substrates to further extend the utility of this bioconjugation strategy.

## 6 REDUCTION MONITORING OF DISULPHIDE BONDS IN PEPTIDES

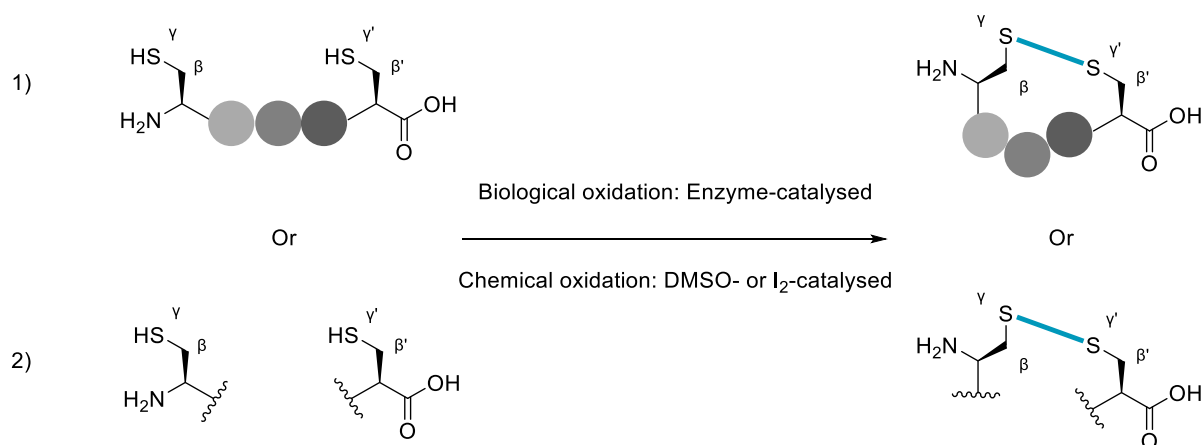
### 6.1 INTRODUCTION

Disulphide bridges are key elements known for playing a crucial role in structural stabilisation and folding of a wide variety of peptides and proteins, including enzymes, immunoglobulins, hormones, and toxins.<sup>175</sup> Of the 90 000 proteins contained in the Protein data bank, 21% feature at least one disulphide bond and the percentage increases when only structural and extracellular matrix proteins are considered.<sup>176</sup> Formation of a disulphide bond involves a two-electron oxidation process that links two side chain sulfur atoms of spatially proximal Cys residues, yielding an oxidised cystine (S-S) from two sulfhydryl groups (S-H). The extent of architectural stabilisation provided by the presence of disulphide bridges can be measure by the entropy of the system. The entropy measures the thermal energy of the system that is not available for producing useful work. Because work can only be obtained from the motion of ordered molecules, the entropy of a system also defines the molecular disorder or randomness.<sup>177</sup> Low entropy means a regular and highly ordered system, whereas high entropy results from a disordered and less stable system. The presence of a covalent crosslink between two sections of a polypeptide chain significantly decreases the conformational freedom of the strand. Restriction of the conformational space leads to a decrease in the entropy of the system, resulting in stabilisation of the unfolded state of the protein.<sup>178</sup>

Disulphide bonds are also key tools in defending biochemical systems against oxidative damage, thanks to their anti-oxidative properties. Most living organisms, including humans, use  $O_2$  as the terminal electron acceptor in energy production processes because of its great electrophilicity.<sup>179</sup> However, if the reduction of oxygen is not complete and water is not formed, reduced forms of molecular oxygen can be produced that still contain electrophilic oxygen and are therefore reactive. These oxidants are called reactive oxygen species (ROS) and they can oxidise vital biochemical compounds such as lipids, fatty acids, proteins, and the nitrogenous bases of DNA, overall inducing oxidative damage to the system.<sup>180</sup> For instance, the superoxide radical ( $O_2^{\cdot-}$ ) is produced by NADPH oxidases and can be partitioned into  $H_2O_2$  by the activity of superoxide dismutase.<sup>181</sup> Both  $O_2^{\cdot-}$  and  $H_2O_2$  are by-products

of the oxygen metabolism and can induce oxidative damage if not regulated. Nitric oxide synthases produce the oxidative agent nitric oxide (NO·), which is a signalling molecule involved in a variety of pathophysiological processes such as inflammation, cell apoptosis, smooth muscle relaxation and neurotransmission.<sup>182,183</sup> Reaction of  $O_2^-$  with NO· can generate the strong oxidant peroxynitrous acid (ONOOH) that was found to contribute in the pathogenesis of inflammatory processes, neurodegenerative disorders, sepsis, and others.<sup>184</sup> Another example of oxidative damage is given by the generation of myeloperoxidase by polymorphonuclear neutrophils, macrophages and monocytes. This leukocyte-derived enzyme can react with  $H_2O_2$  to catalyse the formation of powerful two-electron oxidants such as hypochlorous acid (HOCl), hypothiocyanous acid (HOSCN), and hypobromous acid (HOBr).<sup>185</sup>

Disulphide bonds results from the oxidation reaction of the sulfhydryl (SH) group of two Cys residues, moving from an oxidation state of -2 at the sulfur atom to an oxidation state of -1 at each of the two linked sulfur atoms in the cystine moiety (Scheme 32). Disulphide bonds can be formed either between two Cys residues of the same peptide sequence (intra-molecular, Scheme 32, 1), or between Cys residues belonging to two different sequences (inter-molecular, Scheme 32, 2).



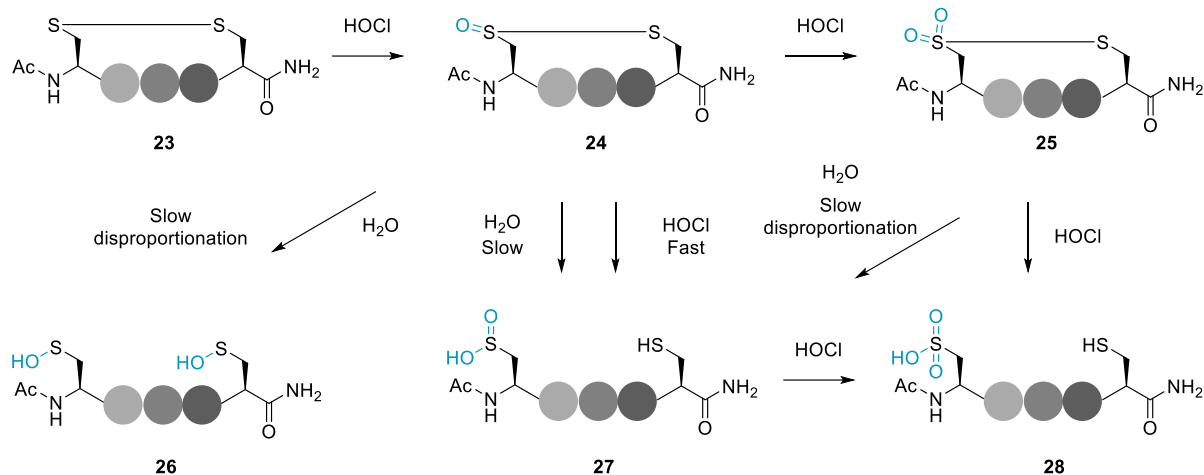
*Scheme 32 – Biological and chemical oxidation of Cys residues for disulphide bond formation.*

This interaction acts as a molecular “staple” that secures the 3D structure of the chain by imposing distance and angle constraints between the β-carbon and γ-sulfur atoms of the two joined Cys residues.<sup>186</sup> In cellular environment the oxidation of sulfhydryl groups to form disulphide bridges is catalysed by enzymes such as protein disulphide isomerases or thioredoxin in a process called oxidative protein folding.<sup>187</sup> Cleavage of disulphide bridges in proteins could potentially result in collapsing of the native

conformational and biological functions. On the other side, failure in disulphide bonds formation may result in protein aggregation and consequent cellular stress and cell death.<sup>188</sup>

Recently, attention has been focused on the artificial introduction of disulphide bridges in peptides and proteins to obtain enhanced stability and pharmacokinetic profiles in the corresponding analogues. An example in this context is given by insulin, a small peptide hormone that is crucial in controlling the glucose homeostasis and that already contains three disulphide bonds. Vinther *et al.*, showed that the addition of a fourth disulphide bond led to increased affinity for the insulin receptor, augmented glucodynamic potency, and an increase of 34.6 °C in the melting point.<sup>189</sup>

More recently, disulphide bonds were found to rapidly react with available lone pairs of electrons provided by biological oxidants. These findings proved that disulphides are redox active species that can catalyse enzymatic reactions and coordinate protein function by directly modify the enzyme activity through a conformational change. Disruption of disulphide bridges in biologically active peptides and proteins is often associated with a loss in their function and activity.<sup>176</sup> For this reason, understanding the causes that control the oxidation of disulphides could help in regulating the stability and pharmacokinetic properties of peptides and proteins containing linked Cys residues. When the disulphide bond is cleaved in presence of HOCl, for example, the product does not correspond to the reduced Cys residues, and this is due to the oxidative conditions induced by the oxidative agent. Karimi *et al.* identified a series of oxidation species produced by disulphide bond cleavage in presence of HOCl, as detailed in Scheme 33.<sup>190</sup> Thiosulfinate **24** is the first intermediate isolated after treatment of the cyclised peptide **23** with HOCl. Following further oxidation or slow hydrolysis, **24** is converted into the thiosulfonate **25**. Subsequent oxidation of the disulphide-S-monoxide and disulphide-S-dioxide yields a number of oxidised products where the disulphide bond was cleaved. These species included the bis-sulfenic acid **26**, the sulfinic acid **27**, and the sulfonic acid **28**. In Scheme 33 only oxidation happening at the *N*-terminal residue is reported, however, evidence was detected for a similar process occurring at the *C*-terminal Cys residue as well.



*Scheme 33 – Oxidation species produced from an intra-molecular disulphide bond after treatment with aqueous HOCl.*

## 6.2 AIMS AND OBJECTIVES

In Chapter 6 we investigated the reactivity of disulphide bonds toward the oxidative agent HOCl in three different 6-mer, *N*-acetylated peptide sequences. The peptide sequences were synthesised with a consistent spacing between the two Cys residues (Ac-**CAGACY**-NH<sub>2</sub>, Ac-**CEGACY**-NH<sub>2</sub>, and Ac-**CKGACY**-NH<sub>2</sub>). An intra-molecular disulphide bridge was synthetically introduced between the proximal Cys residues, and monitoring of the S-S bond reduction was explored using high performance liquid chromatography (HPLC) analysis. The synthesis of the cyclised peptides is described in General Protocols for Chapter Six. The three selected sequences were non-native, 6-mer peptides, all differing for the residue in the second position from the *N*-terminus. These three peptides were chosen as a starting point for the analysis. They provided small sequences with two Cys residues required for the formation of an intramolecular disulphide bond and an aromatic residue (Tyr) to assist purification. The different residues we examined were Gly, Glu, and Lys, which conferred respectively no charge, a negative charge, and a positive charge next to one of the Cys residues involved in the disulphide bond, when the peptide is dissolved at physiological pH. Figure 101 shows the structures for the three reduced peptides (Gly **29a**, Lys **30a**, and Glu **31a**) and the respective cyclised peptides (Gly **29b**, Lys **30b**, and Glu **31b**).



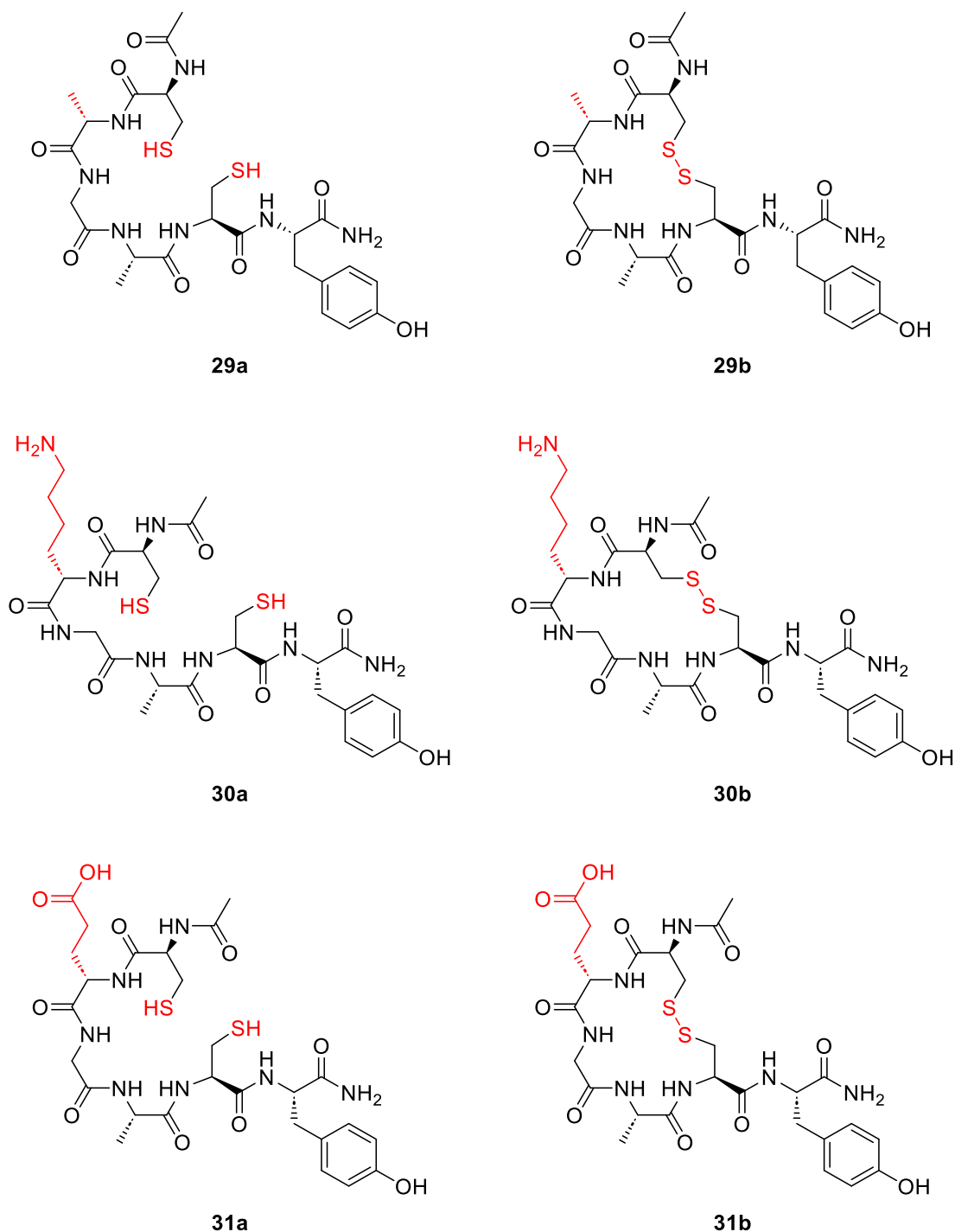


Figure 101 – Synthetic peptides utilised for the reduction monitoring of disulphide bonds.

The reduction of the intramolecular disulphide bridge installed in the synthetic peptide was performed with HOCl, a reactive species containing a nucleophilic oxygen that was reported as an oxidative agent in biochemical processes. Following optimisation of the reaction conditions, monitoring of the reduction profile was conducted by analytical HPLC and characterisation of the products was performed with LC-MS. The scope of this effort was to produce preliminary data that, alongside computational

experiments, could provide an insight in whether different residues, local steric, and electrostatic environment placed next to one of the two Cys residues involved in the disulphide bond could play a role in stabilising the cyclic peptide and slow the reduction of the disulphide bond and the activity of HOCl.

## 6.3 RESULTS AND DISCUSSION

### 6.3.1 Preliminary Computational Calculations

#### 6.3.1.1 Principal Component Analysis

Molecular dynamic simulations were produced for all the three synthetic peptides bearing a disulphide bond: Ac-CAGACY-NH<sub>2</sub> (**29b**), Ac-CEGACY-NH<sub>2</sub> (**30b**), and Ac-CKGACY-NH<sub>2</sub> (**31b**). From now on, both reduced (**29a**, **30a**, and **31a**) and oxidised (**29b**, **30b**, and **31b**) peptides will be abbreviated to the amino acid sequence only and the *N*- and *C*- termini modifications, respectively acetylated (Ac) and amide (NH<sub>2</sub>), will be omitted.

First of all, principal component analysis (PCA) was performed to explore the free energy landscape of the different cyclised peptides and determine their relative cluster centres. PCA is a statistical technique that reduces the dimensionality of a dataset and increases the interpretability while minimising the loss of information. This result is obtained by generating new uncorrelated variables, the principal components, that maximise the statistical information.<sup>191</sup> Figure 102 displays the results for the free energy landscape map of synthetic peptide CAGACY (**a**), CEGACY (**b**), and CKGACY (**c**). The energy landscape provides a mapping of all the possible states of the system and describes all the conformation and the spatial positions of the interacting functional groups.

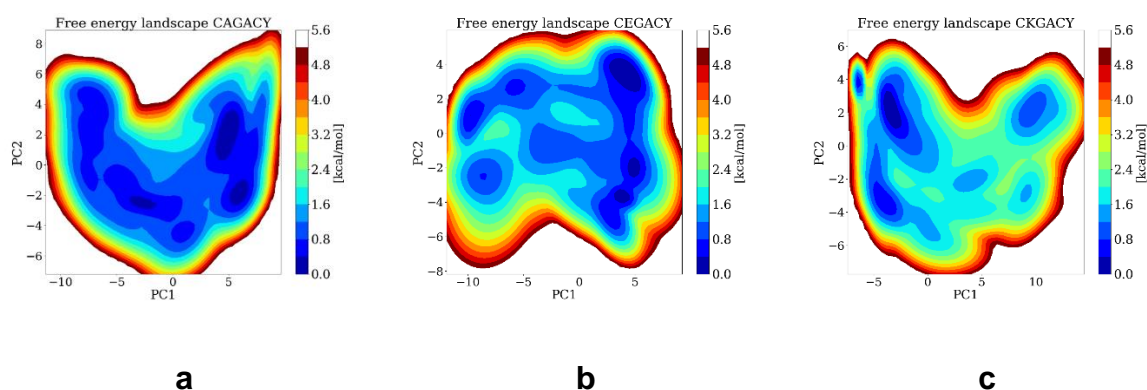


Figure 102 – Estimated free energy landscape of peptide CAGACY (a), CEGACY (b), and CKGACY (c) along the distribution of PC1 and PC2 constructed using 1498500 snapshots from 299.7 ns of MD simulation (1498503 snapshots for CAGACY). PCA was carried out by calculating root-mean-square deviation (RMSD) of all backbone atoms including C $\beta$  and S atoms of Cys from the reference average structure.

As detailed by the legend on the right-hand side of each graph, low-energy areas are characterised by dark-blue colours, whereas high-energy areas are characterised by red colours. Each of these regions define an ensemble of energetically stable structures.

The free energy landscape maps for the three peptide sequences were analysed using the clustering method InfleCS. This tool was implemented to extract well-defined core states from the energy landscape, and the results are showed in Figure 103.

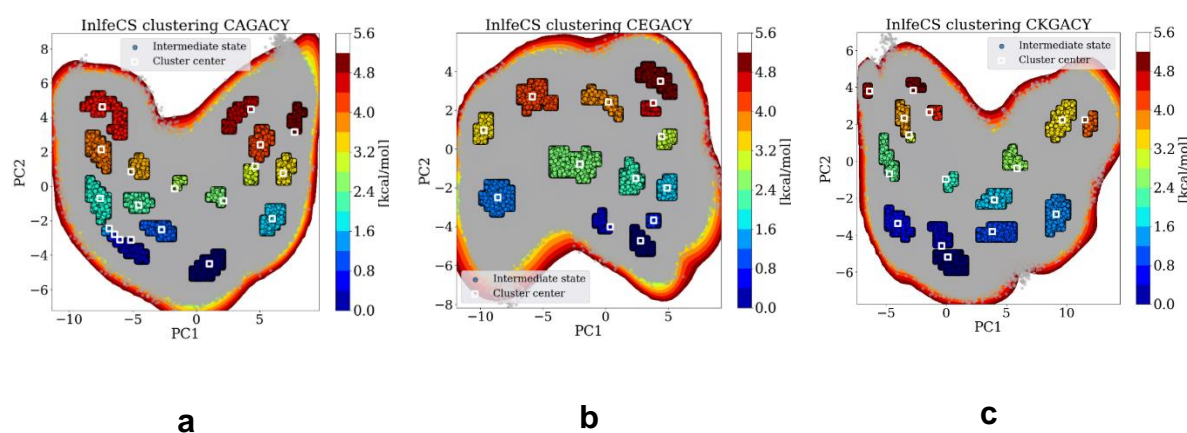


Figure 103 – Results from clustering using the InfleCS method. The final density landscape consists of 19 Gaussian components for CAGACY (a), 13 for CEGACY (b), and 16 for CKGACY (c). The identified core states based on the estimated density on a 100 x 100 grid, are coloured by cluster labels. Cluster centres and transition points are shown as white squares and grey dots, respectively.

The state populations obtained from Inflex from each peptide sequence were plotted against their probability to gather an idea of the most energetically favoured clusters. The results are presented in Figure 104.

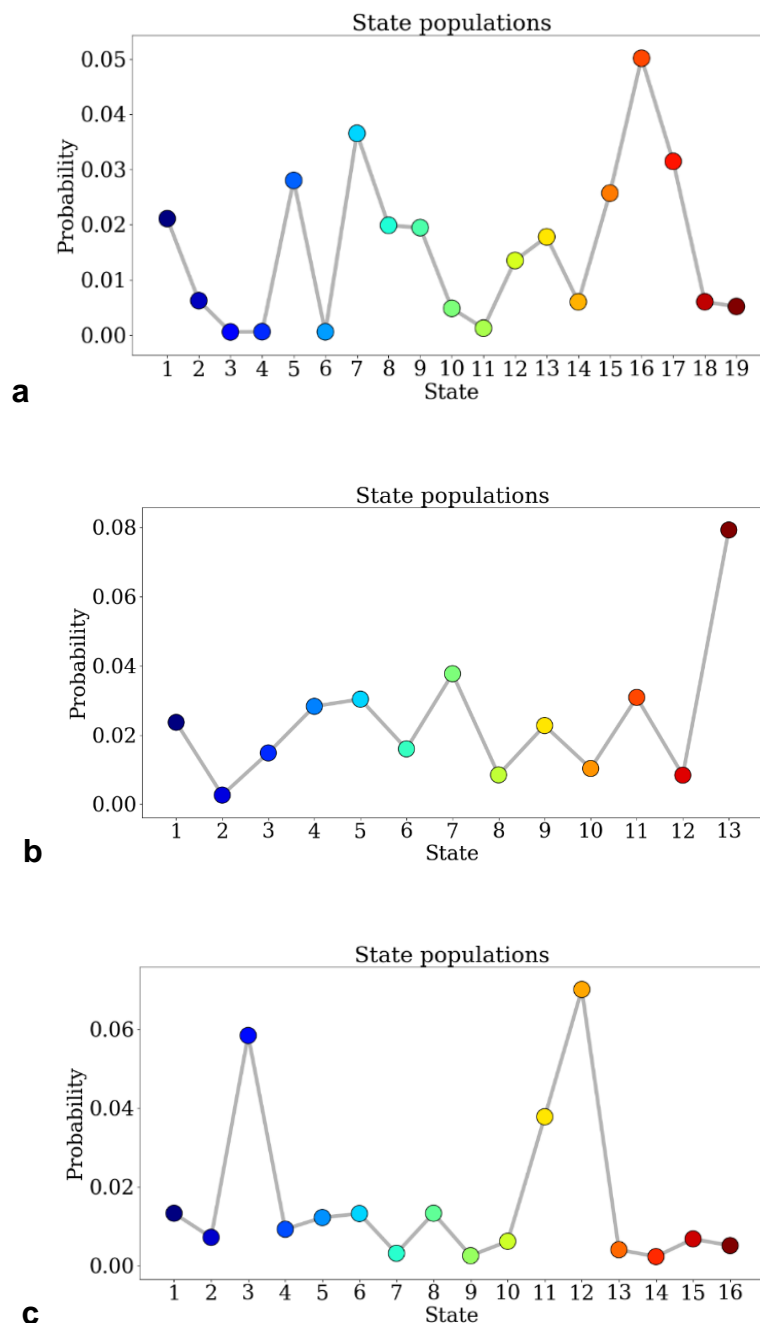


Figure 104 – Normalised state populations for synthetic peptide CAGACY (a), CEGACY (b), and CKGACY (c). The coloured dots in the plot refer to each clustering presented in Figure 103.

Four population states were identified for synthetic peptide CAGACY (5, 7, 16, and 17). The structures representing each cluster centre are shown in Figure 105.

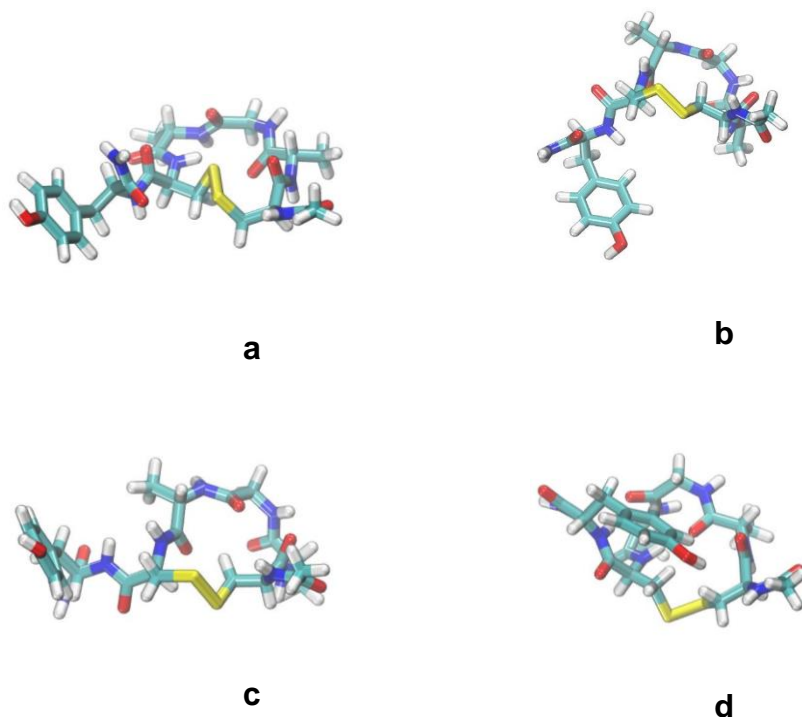


Figure 105 – Structures representing the cluster centres extracted from dominant clusters 5 (a), 7 (b), 16 (c), and 17 (c) for synthetic peptide CAGACY. The structures are visualised with VMD.

InfleCS clustering for synthetic peptide CEGACY produced one population state (13) that emerged from the others. The structure related to the centre of this cluster is showed in Figure 106.

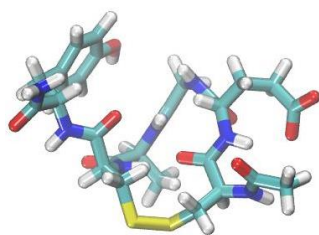


Figure 106 – Structure representing the cluster centre extracted from dominant cluster 13 for synthetic peptide CEGACY. The structure is visualised with VMD.

Synthetic peptide CKGACY was characterised by three population states at high probability, and the structures related to each cluster centre, respectively 3 (a), 11 (b), and 12 (c), are showed in Figure 107.

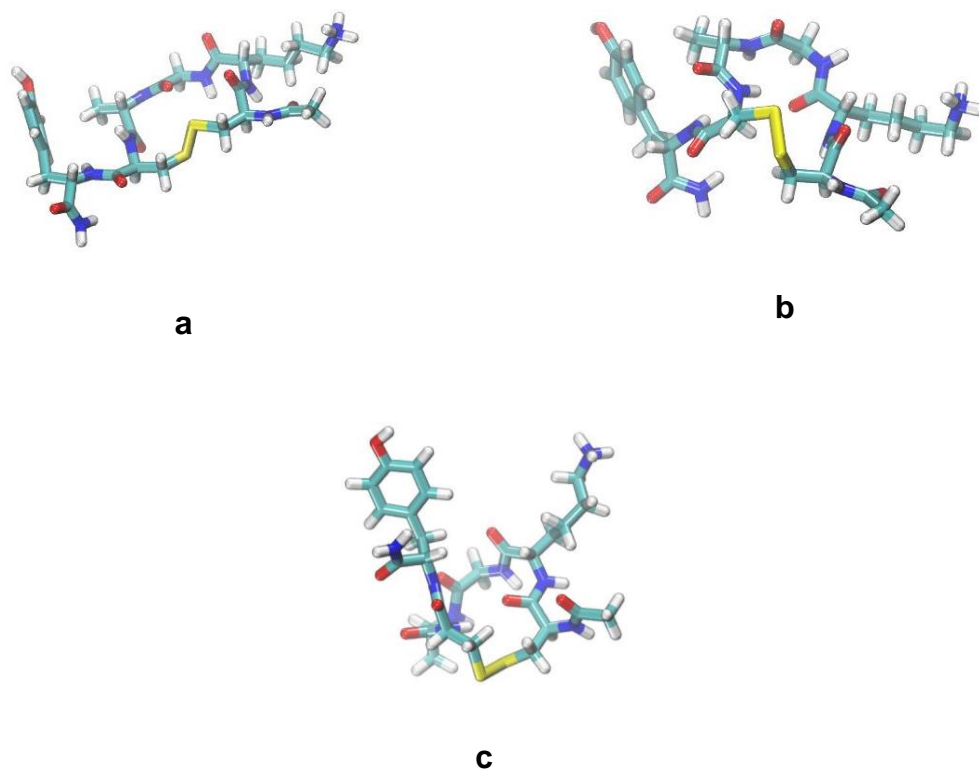
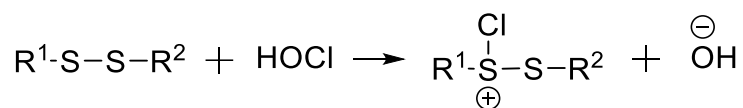


Figure 107 – Structures representing the cluster centres extracted from dominant clusters 3 (a), 11 (b), and 12 (c) for synthetic peptide CEGACY. The structure is visualised with VMD.

InfleCS clustering analysis provided the most stable configurations for each cyclised peptide, describing the atoms and bonds orientation in which the three peptides were most likely to exist.

### 6.3.1.2 Reaction Enthalpy Calculations

Reaction enthalpy ( $\Delta_r H^\ominus$ ) calculations were performed using quantum mechanics (QM). It is thought that oxidation of disulphide bonds begins with formation of a  $R^1S^+(Cl)SR^2$  intermediate (**Scheme 3**). Reaction between the disulphide bond and HOCl can be fast and stabilisation of the RS-Cl species can occur by a remote lone pair of electrons or neighbouring sulfur atoms.<sup>192</sup> The aim of these calculations was to understand if, and how, the three different amino acid residues played a role in the stabilisation of the intermediate. The reaction of interest is showed in Scheme 34.



Scheme 34 – Reaction investigated by  $\Delta_r H^\circ$  calculations.

To perform  $\Delta_r H^\circ$  calculations, four cluster centers were taken that best represented the InfeCS clustering plot of each peptide. An example is showed in Figure 108 for synthetic peptide CAGACY.

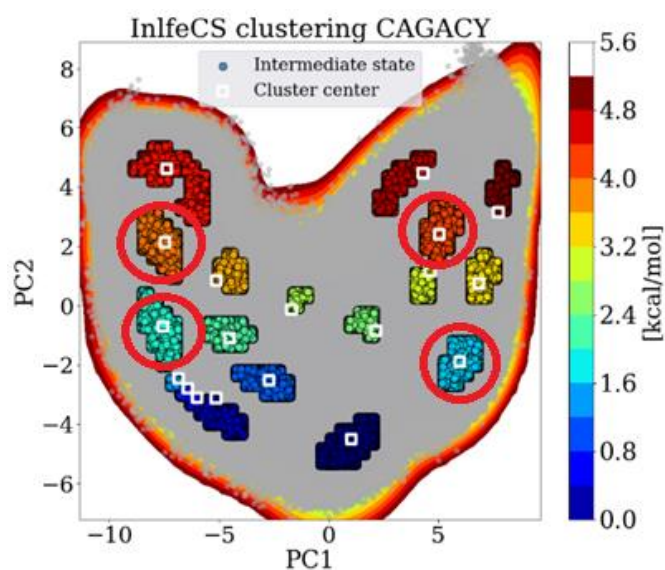


Figure 108 – Circled clusters from the InfeCS clustering were selected to sample the most diverse peptide conformations along both variables PC1 and PC2.

A chlorine atom was bound to the sulfur atoms involved in the disulphide bond in each of the four possible conformations, as showed in Figure 109.

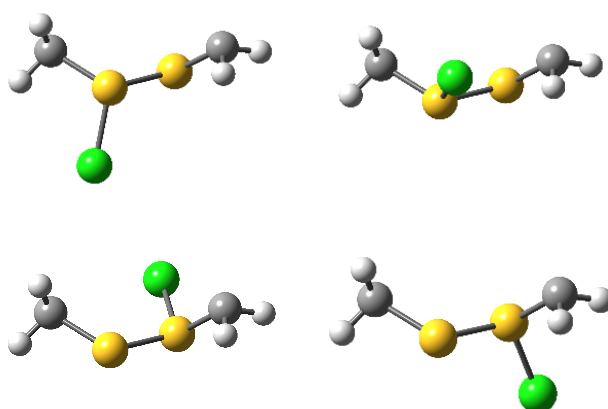


Figure 109 – Four different combinations for the S-Cl bond on the disulphide bridge. A total of 48 structures (16 for each peptide) were analysed. Green ball = Cl, yellow ball = S, grey ball = C, and white ball = H.

The energy enthalpy values for the S-Cl bond are reported in Table 5. These values were plotted in Figure 110 for an easier interpretation.

Table 5 –  $\Delta_r H^\circ_{298}$  (kJ/mol) values (highlighted numbers) given in kJ/mol for all the chlorinated structures of each peptide. The chosen notation was X-Y, where X is a letter, and Y is a number. The letter refers to the characteristic amino acid of each peptide (A = CAGACY, K = CKGACY, and E = CEGACY), whereas the number refers to the cluster centre. The double-digit numbers (1-1, 1-2, 2-1, and 2-2) refer to the four different combinations of S-Cl bond on the disulphide bridge (see Figure 110). All geometries and vibrational frequency calculations were obtained with the M05-2X/6-31G(d) method using implicit SMD water solvation model. Zero-point vibrational energies and enthalpic temperature corrections were obtained from scaled M05-2X/6-31G(d) vibrational frequencies using scaling factors from literature.<sup>8</sup> Single-point energies were obtained with M06-2X/6-311 + G(3df,2p) including SMD solvation and thermal corrections. Energies are given in kJ/mol, and they represent condensed-phase enthalpies at 298 K. All density functional QM calculations were carried out using Gaussian 16.<sup>193</sup>

<b>A-15</b>	1-1	<b>59.71</b>	<b>K-12</b>	1-1	<b>46.52</b>	<b>E-9</b>	1-1	<b>81.32</b>
	1-2	<b>67.95</b>		1-2	<b>65.84</b>		1-2	<b>75.49</b>
	2-1	<b>71.41</b>		2-1	<b>69.22</b>		2-1	<b>85.70</b>
	2-2	<b>84.24</b>		2-2	<b>54.54</b>		2-2	<b>75.81</b>
<b>A-8</b>	1-1	<b>44.54</b>	<b>K-3</b>	1-1	<b>43.52</b>	<b>E-4</b>	1-1	<b>92.04</b>
	1-2	<b>60.46</b>		1-2	<b>61.59</b>		1-2	<b>87.26</b>
	2-1	<b>74.22</b>		2-1	<b>43.09</b>		2-1	<b>80.43</b>
	2-2	<b>41.52</b>		2-2	<b>54.43</b>		2-2	<b>87.42</b>
<b>A-7</b>	1-1	<b>68.76</b>	<b>K-5</b>	1-1	<b>71.68</b>	<b>E-5</b>	1-1	<b>54.02</b>



	1-2	<b>75.63</b>		1-2	<b>56.91</b>		1-2	<b>52.45</b>
	2-1	<b>62.11</b>		2-1	<b>77.69</b>		2-1	<b>59.48</b>
	2-2	<b>88.64</b>		2-2	<b>50.76</b>		2-2	<b>48.07</b>
<b>A-16</b>	1-1	<b>74.07</b>	<b>K-11</b>	1-1	<b>61.83</b>	<b>E-13</b>	1-1	<b>75.26</b>
	1-2	<b>82.37</b>		1-2	<b>78.92</b>		1-2	<b>62.88</b>
	2-1	<b>58.52</b>		2-1	<b>65.38</b>		2-1	<b>72.92</b>
	2-2	<b>86.48</b>		2-2	<b>68.81</b>		2-2	<b>36.13</b>

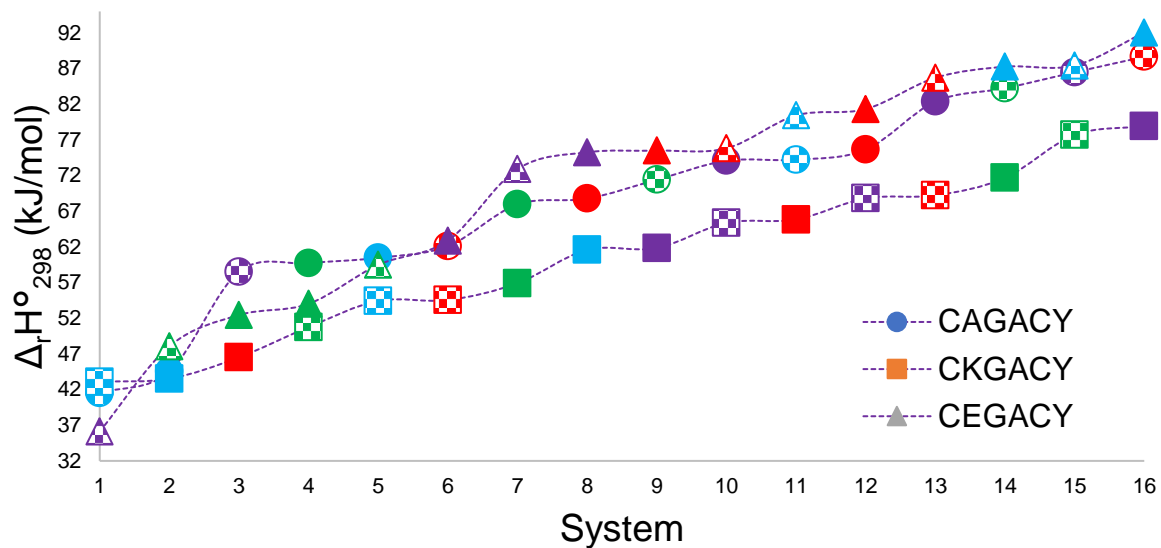


Figure 110 –  $\Delta_f H^\circ_{298}$  trend for chlorinated structures of CAGACY, CKGACY, and CEGACY peptides in water. Different colours represent the cluster from the PCA plot: red is top-left cluster, blue is bottom-left cluster, green is bottom-right cluster, and purple is top-right cluster. Different patterns represent the chlorine atom position on the disulphide bridge: solid fill is the sulfur atom closer to the C-terminus, and checkerboard pattern is the sulfur atom closer to the N-terminus.

For both cyclised peptides CAGACY and CKGACY the bottom-left cluster of the InfleCS clustering (green marker) seemed to be the most stable, with the first two structures being the lowest in reaction enthalpy. The rest of the structures did not seem to follow a specific trend. For CEGACY the cluster showing the lowest reaction enthalpy was the bottom-right one (green marker). This peptide also covered the widest range in reaction enthalpy, showing both the lowest and the highest overall value, respectively 36 and 92 kJ/mol.

### 6.3.1.3 Sulfur-Chlorine Bond Length

Once all structures were obtained for the three synthetic peptides, the length of the sulfur-chlorine bond was measured per each structure. The results are gathered in Table 6.

Table 6 – Bond (S-Cl) length (Å) values (highlighted numbers) given in Å for all the chlorinated structures of each peptide.

<b>A-15</b>	1-1	<b>2.07</b>	<b>K-12</b>	1-1	<b>2.18</b>	<b>E-9</b>	1-1	<b>2.03</b>
	1-2	<b>2.05</b>		1-2	<b>2.06</b>		1-2	<b>2.03</b>
	2-1	<b>2.17</b>		2-1	<b>2.02</b>		2-1	<b>2.03</b>
	2-2	<b>2.04</b>		2-2	<b>2.05</b>		2-2	<b>2.03</b>
<b>A-8</b>	1-1	<b>2.14</b>	<b>K-3</b>	1-1	<b>2.12</b>	<b>E-4</b>	1-1	<b>2.03</b>
	1-2	<b>2.06</b>		1-2	<b>2.05</b>		1-2	<b>2.03</b>
	2-1	<b>2.05</b>		2-1	<b>2.06</b>		2-1	<b>2.03</b>
	2-2	<b>2.07</b>		2-2	<b>2.04</b>		2-2	<b>2.03</b>

<b>A-7</b>	1-1	<b>2.06</b>	<b>K-5</b>	1-1	<b>2.03</b>	<b>E-5</b>	1-1	<b>2.07</b>
	1-2	<b>2.05</b>		1-2	<b>2.08</b>		1-2	<b>2.05</b>
	2-1	<b>2.05</b>		2-1	<b>2.04</b>		2-1	<b>2.05</b>
	2-2	<b>2.03</b>		2-2	<b>2.04</b>		2-2	<b>2.15</b>
<b>A-16</b>	1-1	<b>2.06</b>	<b>K-11</b>	1-1	<b>2.03</b>	<b>E-13</b>	1-1	<b>2.02</b>
	1-2	<b>2.03</b>		1-2	<b>2.04</b>		1-2	<b>2.05</b>
	2-1	<b>2.05</b>		2-1	<b>2.16</b>		2-1	<b>2.06</b>
	2-2	<b>2.03</b>		2-2	<b>2.04</b>		2-2	<b>2.40</b>

These values were also plotted in Figure 111 for an easier interpretation.

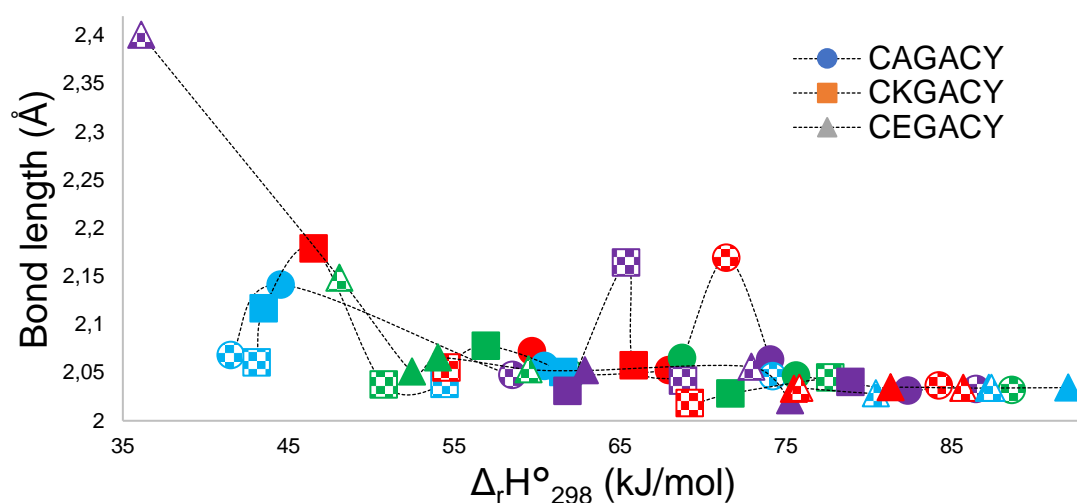


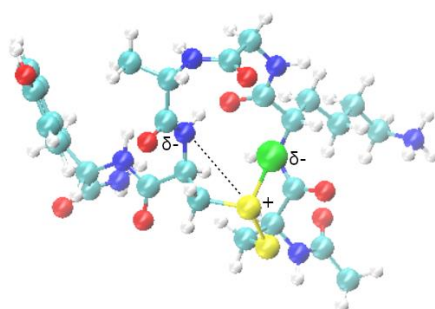
Figure 111 – S-Cl bond length (Å) vs  $\Delta_r H^\circ_{298}$  for chlorinated structures of CAGACY, CKGACY, and CEGACY peptides in water. Different colours represent the cluster from the PCA plot: red is top-left cluster, blue is bottom-left cluster, green is bottom-right cluster, and purple is top-right cluster. Different patterns represent the chlorine atom position on the disulphide bridge: solid fill is the sulfur atom closer to the C-terminus, and checkerboard pattern is the sulfur atom closer to the N-terminus.

The overall trend observed in Figure 111 indicated that the S-Cl bond distances negatively correlated with the calculated  $\Delta_rH^\circ$  values, namely the lowering of the reaction enthalpy corresponded to the increasing in the S-Cl bond distance. While most of the S-Cl distances were found to be roughly around 2.1 Å, the exception of 2.4 Å was observed in the case of one conformation in the CEGACY peptide. This outcome was probably due to the stabilising effect of the neighbouring oxygen and nitrogen atoms that are able to donate electrons to the empty orbital of the sulfur atom, which then results in a decrease of the reaction enthalpy. The formation of a bond between oxygen and sulfur would certainly benefit the reaction enthalpy of the first step, but might interfere with the overall reactivity of the intermediate in the following steps. In particular, this feature might be important when considering the next step of the mechanism, which is the hydrolysis of the intermediate (addition of a water molecule to the sulfur atom and cleavage of the S-Cl bond). Moreover, if the intermediate is too stabilised, it would be less likely to react during the following step. More computational data will be required for better understanding the overall forces in place.

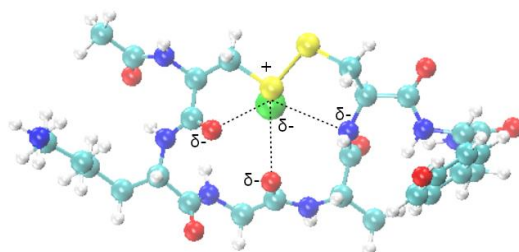
All of our calculated reaction enthalpies are relatively high (35-39 kJ/mol). The nature of the interactions stabilising the product is electrostatic and takes place between the nitrogen or the oxygen atoms from the peptide, which are partially negatively charged due to their electronegativity (respectively 3.04 and 3.44), and the positively charged sulfur atom involved in the disulphide bridge bonded to the chlorine atom (the S-Cl bond is polarised). The chlorine atom is involved in two different types of interaction. It is negatively charged due to its electronegativity (3.16), and thus interact with positively charged atoms. However, it also has an  $\sigma$ -hole, which involves a region of positive electrostatic potential on the extension of the covalent S-Cl bond, and thus can interact with negatively charged atoms.<sup>194</sup> Generally, the structures analysed showed the chlorine atom pointing away from the rest of the molecule, probably to avoid steric hindrance.

CKGACY showed overall lower  $\Delta_rH^\circ$  values with respect to the other two peptides. This could possibly be explained by the number of interactions generated (formation of stable macrocycles) between the negatively charged oxygen and nitrogen atoms surrounding the peptide that coordinate the positively charged sulfur atoms. As a result

of these interactions, the formation of a macrocycle was observed as the conformation with the lowest enthalpy (Figure 112 – bottom structure).



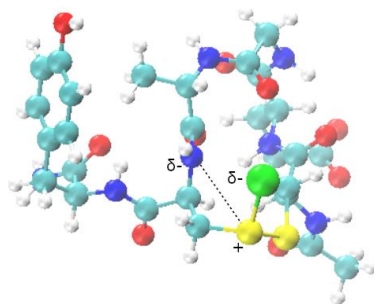
**78.92 kJ/mol**



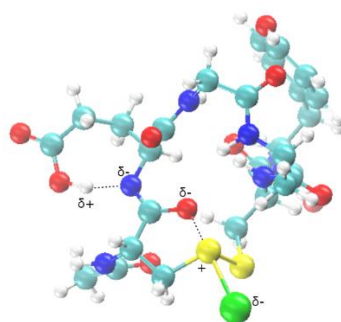
**43.09 kJ/mol**

*Figure 112 – CKGACY structures showing highest (top, K-11 1-2) and lowest (bottom, K-3 2-1) values of  $\Delta_r H^\ominus$ . Top structure: interaction with a nitrogen atom that leads to the formation of a four-membered heterocycle. Bottom structure: interaction with two oxygen atoms and one nitrogen atom that led to two macro cyclisation of respectively 10 and 7 atoms.*

The lowest  $\Delta_r H^\ominus$  value was 36.13 kJ/mol and belonged to CEGACY. The structure related to this value showed a strong interaction between the positively charged sulfur atom bonded to chlorine, and the negatively charged oxygen atom (Figure 113).



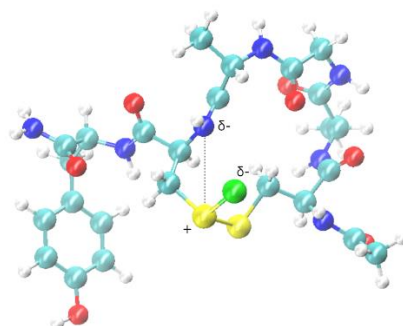
**92.04 kJ/mol**



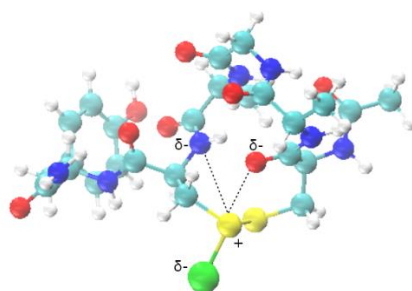
**36.13 kJ/mol**

Figure 113 – CEGACY structures showing highest (top, E-4 1-1) and lowest (bottom, E-13 2-2) values of  $\Delta_r H^\ominus$ . Top structure: interaction similar to that of CAGACY peptide with the lowest enthalpy. Bottom structure: strong interaction between the sulfur atom and an oxygen atom that leads to the formation of a five-membered heterocycle, as well as proton transfer from the peptide bond nitrogen to the oxygen of the carboxylic group on the Glu sidechain.

Despite the CKGACY peptide showed three interactions with the formation of two macrocycles and a rigid structure, its  $\Delta_r H^\ominus$  was higher (43.09 kJ/mol), meaning that the proximity of the oxygen atom to the sulfur atom of CEGACY induced a stronger stabilization of the peptide. The peptide CAGACY showed medium values of enthalpy for the S-Cl bond cleavage. The structure related to the highest and lowest enthalpies presented some degree of interaction between the positively charged sulfur and the surrounding oxygen and nitrogen atoms (Figure 114).



**88.64 kJ/mol**



**41.52 kJ/mol**

Figure 114 – CAGACY structures showing highest (top, A-7 2-2) and lowest (bottom, A-8 2-2) values of  $\Delta_r H^\circ$ . Top structure: only one interaction with a nitrogen atom that leads to the formation of a four-membered heterocycle. Bottom structure: interaction with the same nitrogen atom and the additional interaction with an oxygen atom that leads to the formation of a six-membered heterocycle.

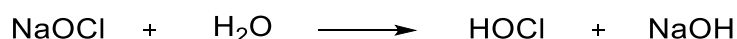
To conclude, CKGACY presented the lower average of reaction enthalpies, despite the lowest overall value was reported for synthetic peptide CEGACY. CAGACY showed the highest values of  $\Delta_r H^\circ$ , even though they did not differ too much from the other two peptides. These results suggest that either a negative (CEGACY) or a positive (CKGACY) charge neighbouring one of the Cys residues involved in the disulphide bond increases the overall stability of the intermediate.

## 6.3.2 Chemical Reduction of Disulphide Bonds

### 6.3.2.1 Optimisation of Reduction Conditions

Peptide CKGACY (**30b**) was selected for the initial optimisation of the disulphide bridge reduction. Four different concentrations of HOCl were tested in the presence of the cyclised peptide. This screening was aimed at understanding what concentration of oxidative agent was required to reduce a significant amount of disulphide bond to

be monitored by analytical HPLC. Too little HOCl could lead to a slow reaction or no reduction of the disulphide bond, whereas too much oxidative agent could result in a too quick cleavage of the S-S bridge to be observed via HPLC. No screening of solution buffer was performed as phosphate buffer 0.1 M at pH = 7.4 was found to be the best in previous literature.<sup>195</sup> NaOCl is stable in basic conditions, and it is commercially available at a pH of 10-11. Dissociation of NaOCl at pH ~ 7.0 leads to *in situ* formation of HOCl, as described in Equation 3.<sup>196</sup>



Equation 3 – Dissociation of NaOCl in water.

After addition of activated NaOCl solution, under the form of HOCl, to the peptide in phosphate buffer at pH = 7.4, the reaction mixture was left for 12 h and then analysed by HPLC. Four different concentrations of HOCl were tested (0.1 mM, 0.5 mM, 1 mM, and 5 mM) and the results are showed in Figure 115. The compound absorbing before 4.5 minutes corresponds to the cyclised peptide **30b**, while the peak at 5.0 minutes was recognised as a product of the disulphide bond cleavage.

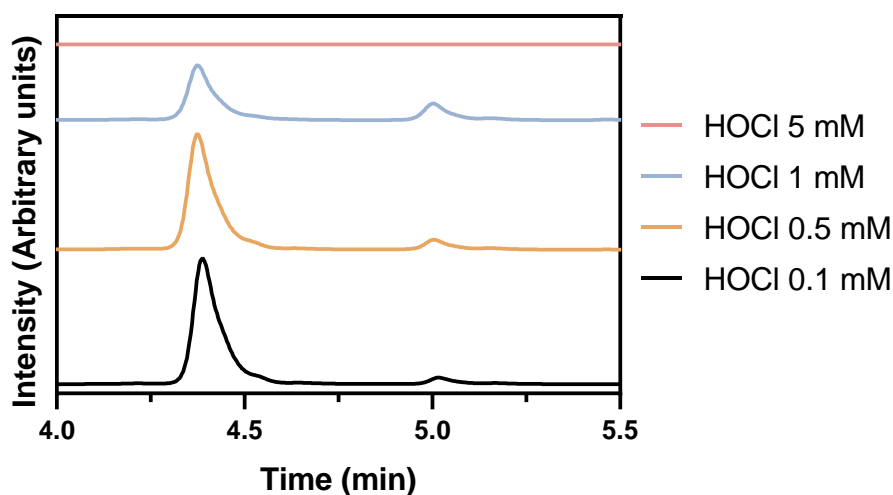


Figure 115 – Reaction optimisation for the reduction of the disulphide bond on Peptide CKGACY (**30b**) using HOCl.

The optimisation experiment run at concentration of 5 mM was repeated with different batches of both peptide and acid, but still showed no peaks in the analytical trace. This outcome was not expected as full hydrolysis of the peptide should only happen in presence of strong acids or bases and at high temperatures.<sup>197</sup> Despite the concentration of HOCl not being high enough to promote the full hydrolysis of the



peptide, even if the HOCl concentration was too high, this should have only cleaved the disulphide bond showing only one peak for the non-cyclised peptide. Even in the remote possibility where the acid was cleaving the amide bond resulting in a disruption of the peptide chain, Tyr is aromatic and should have been visible at 280 nm.

Concentrations of 0.1 and 0.5 mM were not strong enough to successfully cleave the disulphide bond and only a trace of the non-cyclised product was observed. HOCl concentration of 1 mM was selected as the best candidate as it showed a good cleavage profile for the disulphide bond. Most of the cyclised peptide was still present after 12 hours, but the cleavage profile was successful enough to proceed with a longer experiment and the characterisation of the disulphide bond cleavage products. Time point measurements were taken across 177 hours to explore the cleavage trend of cyclised peptide CKGACY, and the areas of both the cyclised and non-cyclised peaks in the HPLC traces were determined. The areas for the cleaved peptide were plotted against time to show the trend of the disulphide bond cleavage, as showed in Figure 116.

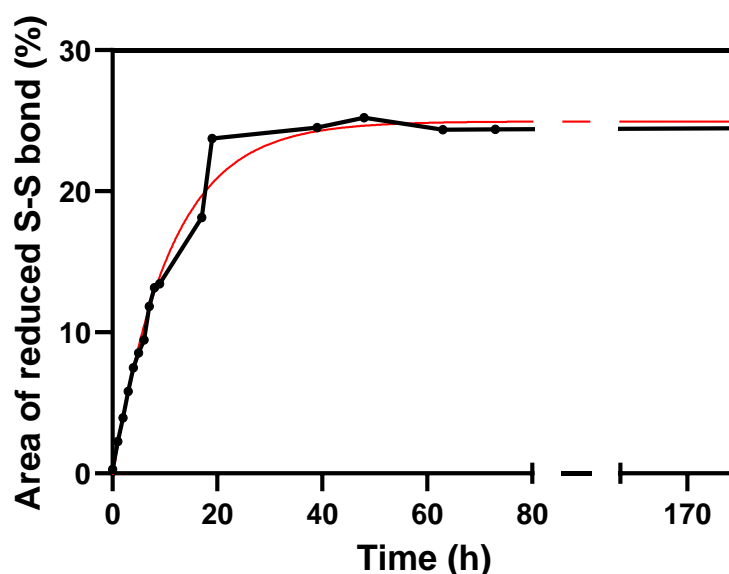


Figure 116 – Trend for the area of reduced disulphide vs. time for the reduction of disulphide bond on Peptide CKGACY (30b) using HOCl. Values from 80 h to 160 h were compressed for easier interpretation.

The data fitted with an exponential association trend (red curve). It was noticed that after 40 hours the oxidative agent was not able to further cleave the cyclised peptide and a plateau was reached in the relative area percentage of the oxidation product.

For this reason, it was concluded that the reaction monitoring of the disulphide bond cleavage on the three synthetic peptides would have been monitored over 40 hours with a scan every 60 minutes.

### 6.3.2.2 Monitoring of Disulphide Bridge Reduction in Synthetic Peptide Ac-CAGACY-NH<sub>2</sub>

Synthetic peptide CAGACY (**29b**) was the first to be analysed after treatment with the oxidative agent HOCl. The chromatograms for both the non-cyclised (**29a**) and cyclised (**29b**) peptides are showed in Figure 117, along with the respective structures. Despite a small difference in the elution times of both peptides, the cyclised peptide was eluted earlier, proving it to be the most polar species between the two.

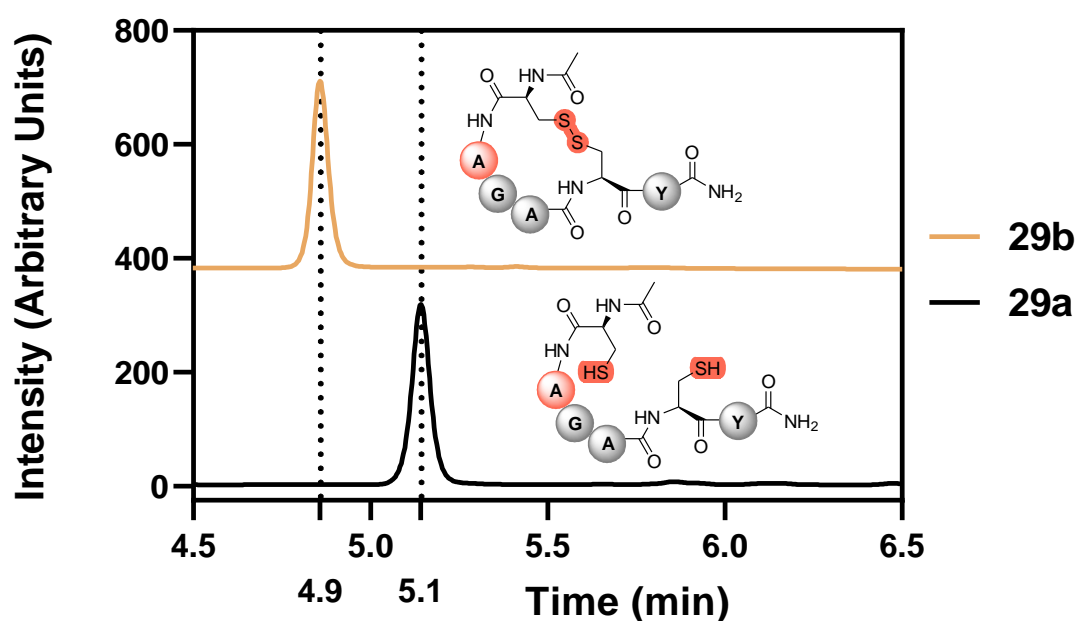


Figure 117 – Comparison of the UPLC traces for the reduction of the disulphide bond on peptide CAGACY. Reduced peptide **29a** eluted after 5.1 min, oxidised peptide **29b** eluted after 4.9 min. Analytical method 5-40% B over 5 min, 210 nm.

The reduction profile of the synthetic peptide CAGACY during the monitored period was characterised by a great number of peaks (Figure 118), implying high reactivity of the cyclised peptide towards the oxidative agent HOCl. However, likely due to the low concentration of the species present in solution, LC-MS was not successful in identifying all of the oxidation products, and only a few of them were investigated.

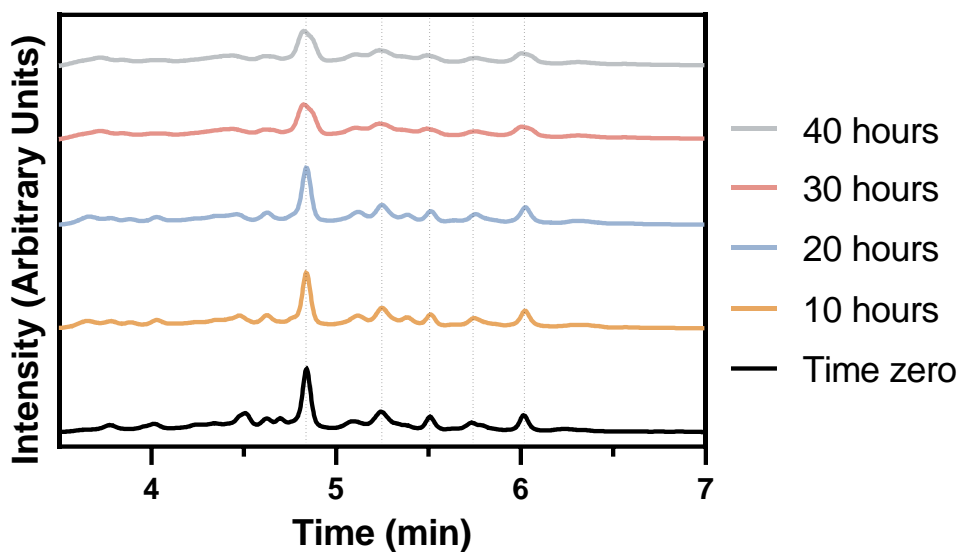


Figure 118 – UPLC traces showing the reaction profile for the reduction of the disulphide bond on peptide CAGACY (**29b**) from time zero (black trace) to 40 hours (light-grey trace).

Looking at the chromatograms showed in Figure 118, it is clear that the species eluting before 5 minutes was predominant throughout the monitored period. The other identified peaks had a much more symmetrical shape in the early stages of the monitored period, between time zero and 20 hours. The shape progressively flattened, and the peaks seemed to be splitting after 30 hours. The reason why this happened is not clear, as the mass spectrum of those peaks only showed one mass.

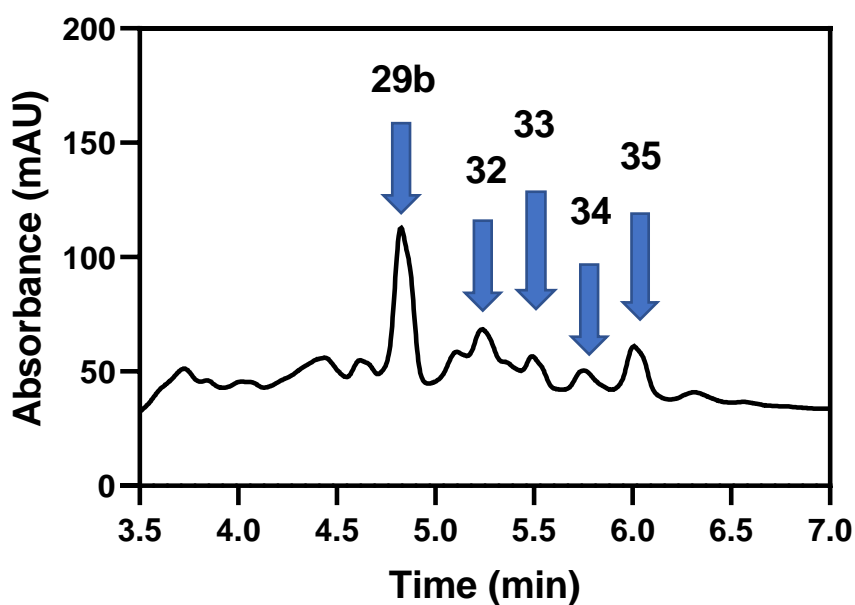


Figure 119 – UV-Vis trace for the reaction mixture after 40 hours, with the main products indicated by an arrow.

Figure 119 shows that many products were formed after exposure of the cyclised peptide to the oxidative environment formed in presence of HOCl. Despite the number of compounds noticed in the chromatogram, LC-MS was only able to detect a few of them. The five species characterised by mass spectrometry are highlighted with an arrow in Figure 119.

The first compound to be identified was unreacted starting material **29b**, the structure of which is showed in Figure 120. The mass spectrum obtained from LC-MS matched the peptide sequence.

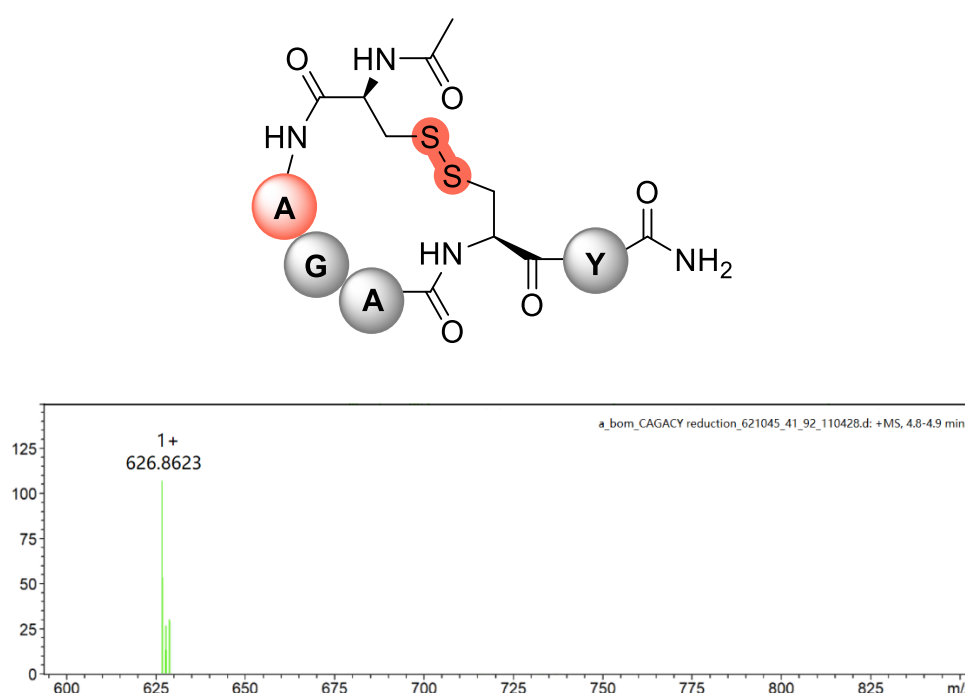


Figure 120 – Structure (top) and HRMS spectrum (bottom) for compound **29b**. Calculated mass: 625.20 [M+H]<sup>+</sup>.

The sulfinic acid **32** showed in Figure 121 was recognised by LC-MS as the second product to be eluted. The peak related to this product appeared more defined during the first 20 hours of monitoring, whereas its concentration decreased, and the peak flattened at the end of the monitored period. The oxidation state of the sulfur taking part in the sulfinic acid moiety is +2, whereas the sulfur atom involved in the disulphide bridge has an oxidation state of -1.

A similar trend was followed by the sulfonic acid (**33**), identified by LC-MS as the oxidation product eluting at approximately 5.5 minutes. The structure showed in Figure 122 matched the molecular weight observed in the respective chromatogram. While

the sulfur atom involved in the disulphide bond has an oxidation state of -1, the sulfur atom of the sulfonic acid has an oxidation state of +4.

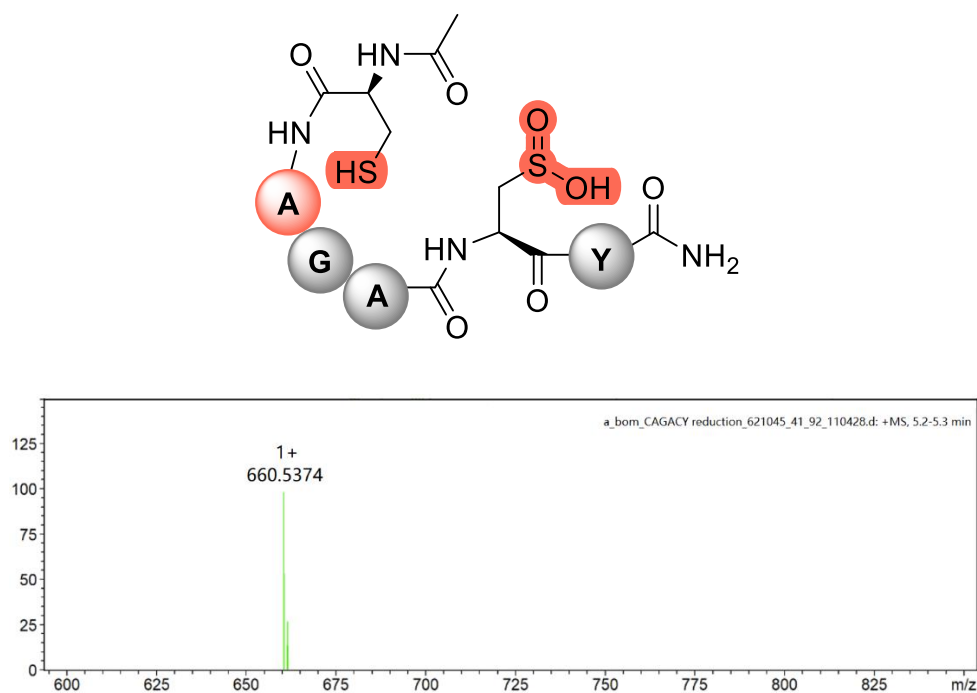


Figure 121 – Structure (top) and HRMS spectrum (bottom) for compound **32**. Calculated mass: 659.20 [M+H]<sup>+</sup>.

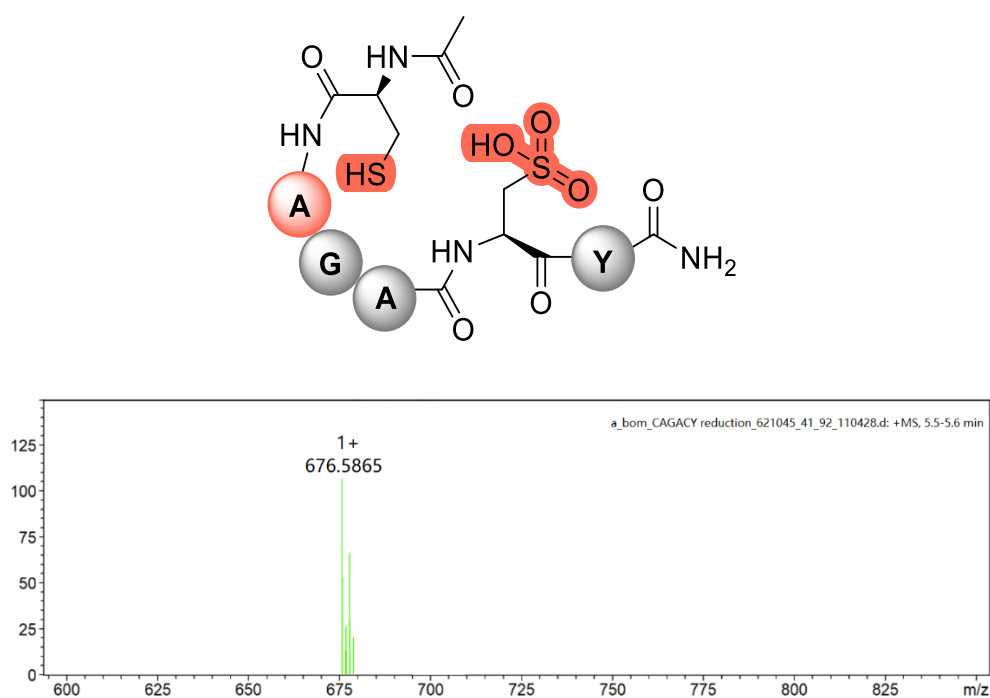


Figure 122 – Structure (top) and HRMS spectrum (bottom) for compound **33**. Calculated mass: 675.20 [M+H]<sup>+</sup>.

The last two compounds identified by LC-MS were oxidation products of the cyclised peptide in which the disulphide bridge remained intact. As previously described in Scheme 33, these species are considered to be intermediates between the disulphide bridge and the oxidation and consequent cleavage of the S-S bond. The first one to be eluted was the thiosulfinate (**34**) showed in Figure 123, the structure of which matched the molecular weight displayed in the respective mass spectrum. The sulfur atom involved in the thiosulfinate moiety has an oxidation state of +1.

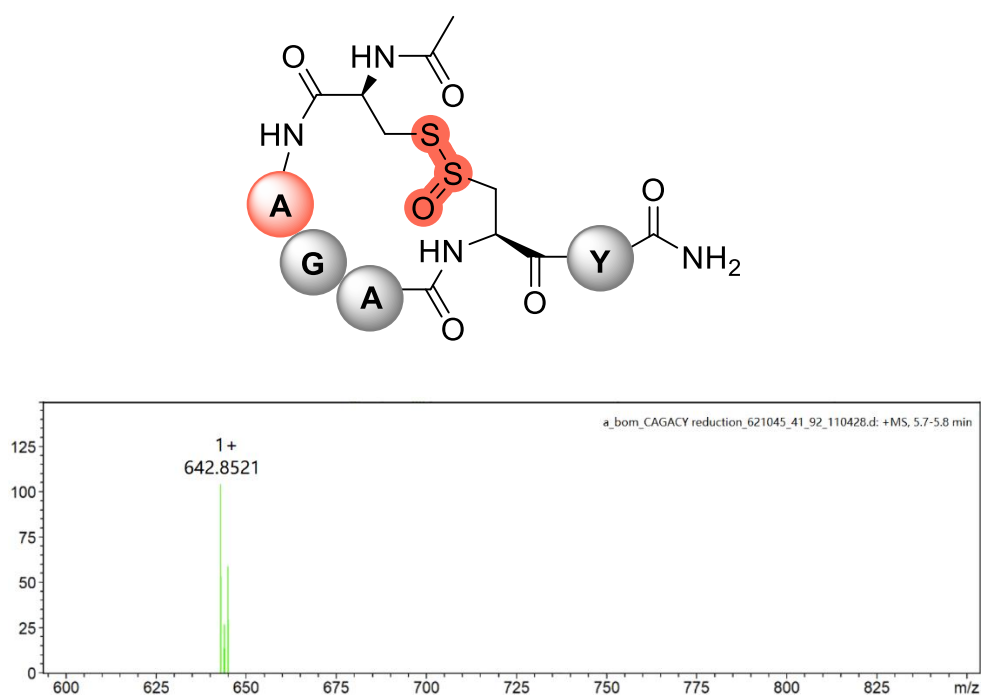


Figure 123 – Structure (top) and HRMS spectrum (bottom) for compound **34**. Calculated mass: 641.20 [M+H]<sup>+</sup>.

The last oxidation product to be characterised was the thiosulfonate **35** eluted at approximately 6 minutes. The mass detected by LC-MS matched the structure showed in Figure 124, and in this case the oxidation state of the sulfur atom is +3.

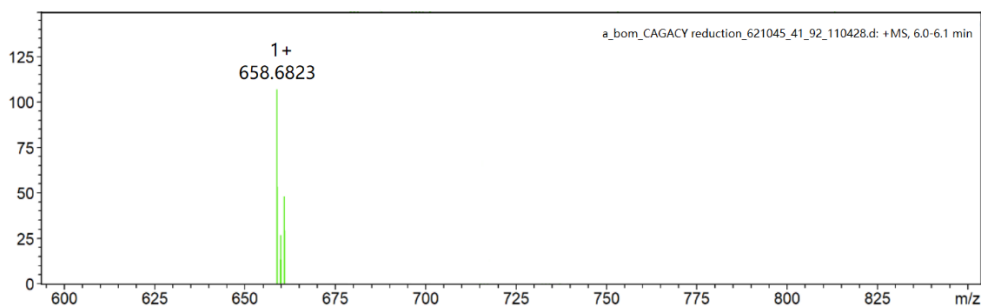
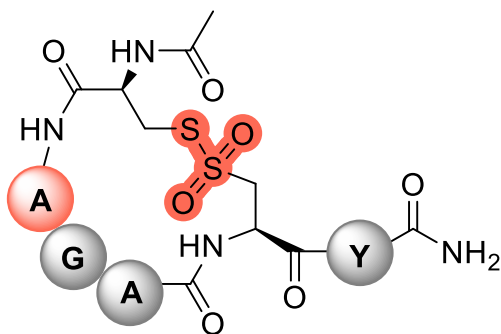


Figure 124 – Structure (top) and HRMS spectrum (bottom) for compound **35**. Calculated mass: 657.19 [M+H]<sup>+</sup>.

Figure 126 shows the trend of the relative area percentages for the characterised peaks over the 40 hours of monitoring. It is clear that all species presented did not deviate much from the original concentration observed at time zero.

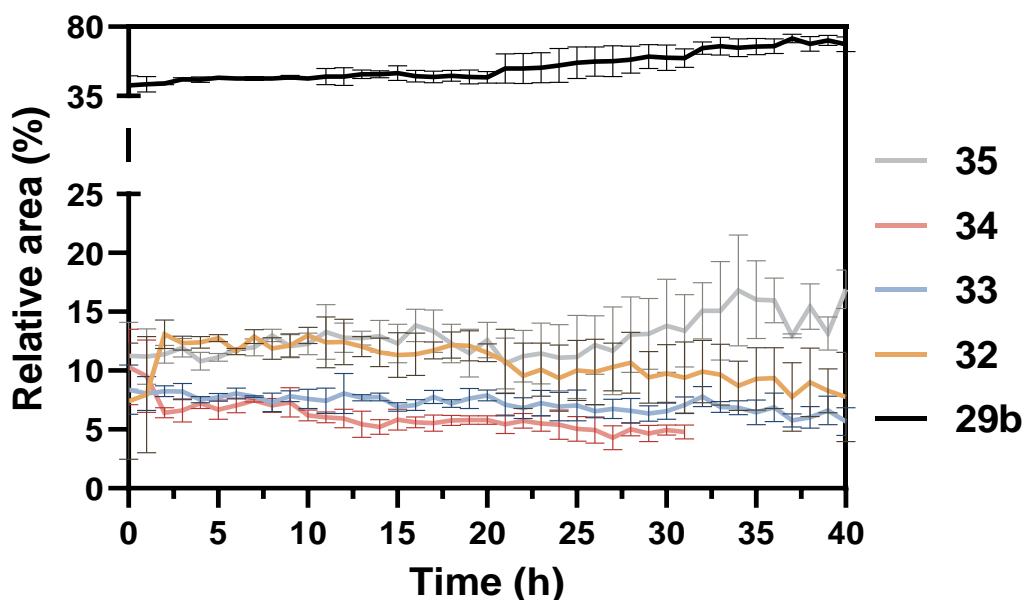


Figure 125 – Relative area (%) vs. time (h) for the five species characterised after 40 hours from exposure of the cyclised peptide CAGACY (**29b**) to HOCl.

The unreacted cyclised peptide (**29b**) was overall the most abundant species with a relative area that increased from 43% to 69%. The thiosulfinate (**34**) could not be distinguished from the baseline after approximately 30 hours, and its trend saw a decrease in the relative area from 10% to 5%. On the other hand, the concentration of the thiosulfonate (**35**) slightly increased from 11% to 17% over the monitored period. This outcome could be explained by the fact that the thiosulfonate is an oxidation product of the thiosulfinate and the concentration of the first is therefore expected to increase. Both oxidation products resulting from the cleavage of the disulphide bond showed a steady trend over the monitored period. The sulfinic acid (**32**), despite an irregular tendency, remained at a consistent 8% of relative area percentage. Sulfonic acid (**33**) experienced a slight decrease from an average of 8% to an average of 5%.

### *6.3.2.3 Monitoring of Disulphide Bridge Reduction in Synthetic Peptide Ac-CKGACY-NH<sub>2</sub>*

Synthetic peptide CKGACY (**30b**) was the second peptide to be monitored after addition of HOCl. Figure 127 shows the HPLC traces of the reduced (**30a**) and oxidised (**30b**) sequences and the corresponding structures. Elution times of the cyclised and non-cyclised peptides were similar, but the cyclised one was eluted slightly earlier than the non-cyclised one.

After addition of HOCl to the reaction mixture, the disulphide bond cleavage was monitored via HPLC over a period of 40 hours with an aliquot of solution being analysed every 60 minutes. Figure 128 shows five different time points taken at time zero and every 10 hours from the beginning of the reaction.



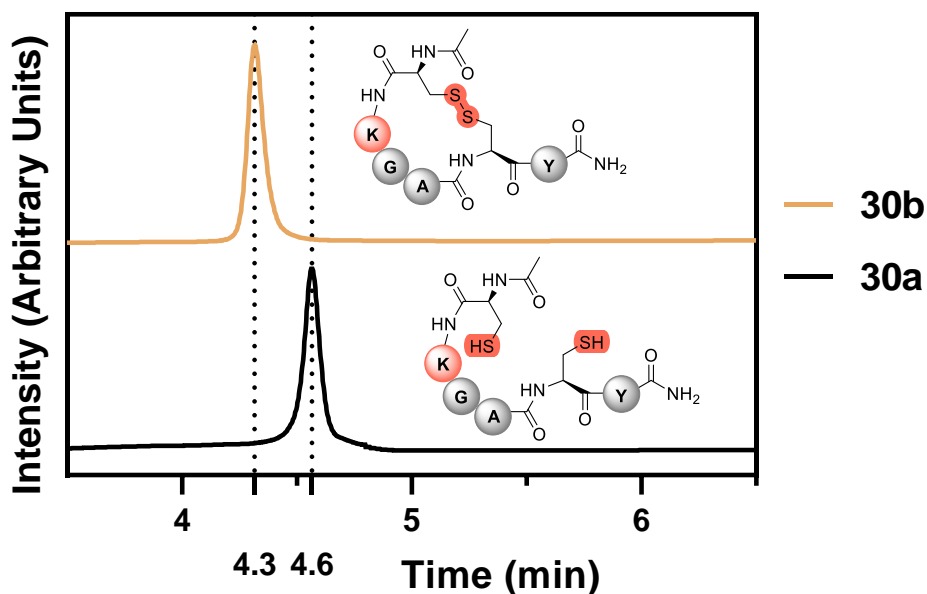


Figure 126 – Comparison of the UPLC traces for the reduction of the disulphide bond on peptide CKGACY. Reduced peptide **30a** eluted after 4.6 min, oxidised peptide **30b** eluted after 4.3 min. Analytical method 5-40% B over 5 min, 210 nm.

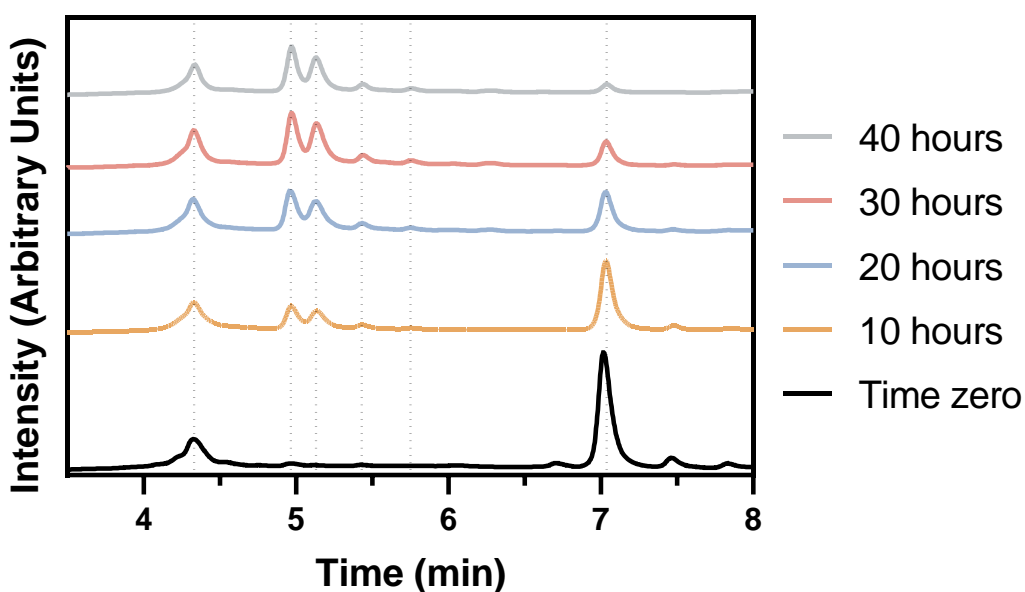
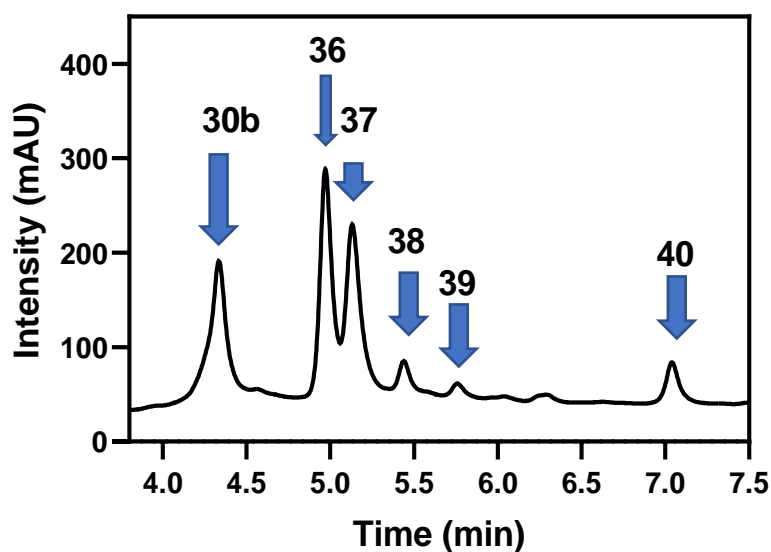


Figure 127 – UPLC traces showing the reaction profile for the reduction of the disulphide bond on peptide CKGACY (**30b**) from time zero (black trace) to 40 hours (light-grey trace).

As presented in Figure 127, after addition of HOCl, monitoring of the reduction process showed conversion of the oxidised peptide (eluting at 4.3 minutes) to a less polar intermediate that was eluted at 7 minutes of the analytical method. Over the analysed 40 hours, this intermediate was nearly completely consumed, and four different compounds appeared to be the stable products. These oxidation products of the disulphide bond were all eluted between minutes 5 and 6 of the analytical method and

showed decreased polarity with respect to both the cyclised and non-cyclised peptides. Other peaks were identified in the HPLC trace of the reduction monitoring after 40 hours; however, their concentration was too low to clearly distinguish them from the baseline and LC-MS was not successful in their characterisation. Therefore, efforts were focused on investigating the nature of the six peaks displayed in Figure 128.



*Figure 128 – UV-Vis trace for the reaction mixture after 40 hours, with the main products indicated by an arrow.*

The numbers in Figure 128 refer to the oxidation products identified after exposure of the synthetic peptide to the oxidative conditions exerted by HOCl. LC-MS was implemented to investigate the nature of these compounds. Due to incomplete conversion of the starting material to a single product, the presence of all intermediates at the end of the 40 hours led to clear characterisation.

The first compound to be eluted was identified with unreacted cyclic peptide (**30b**) as showed by the two matching chromatograms in Figure 129.

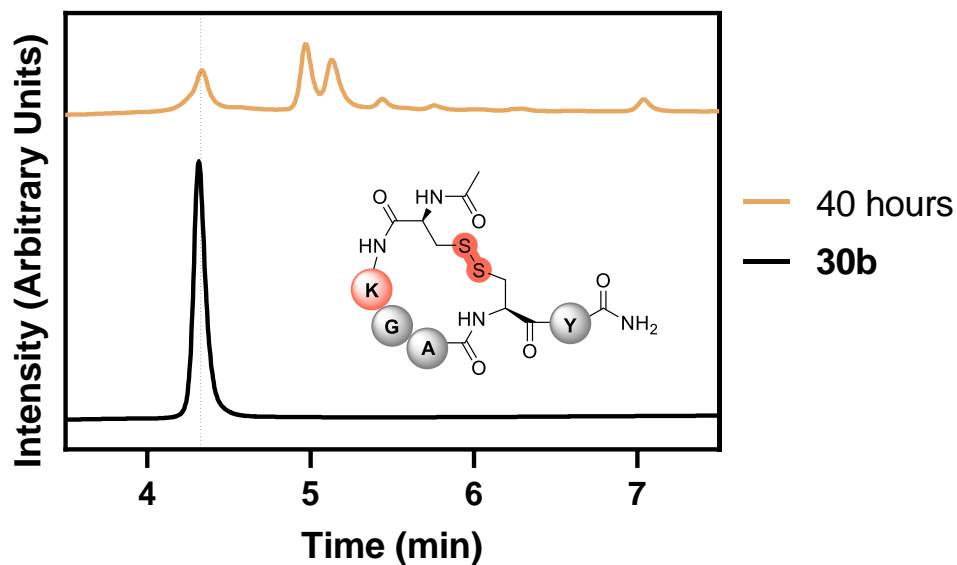


Figure 129 – Comparison of the UPLC traces for peptide **30b** (back trace) and the reaction mixture after 40 hours (gold trace).

The molecular weight was confirmed by the mass spectrum showed in Figure 130.

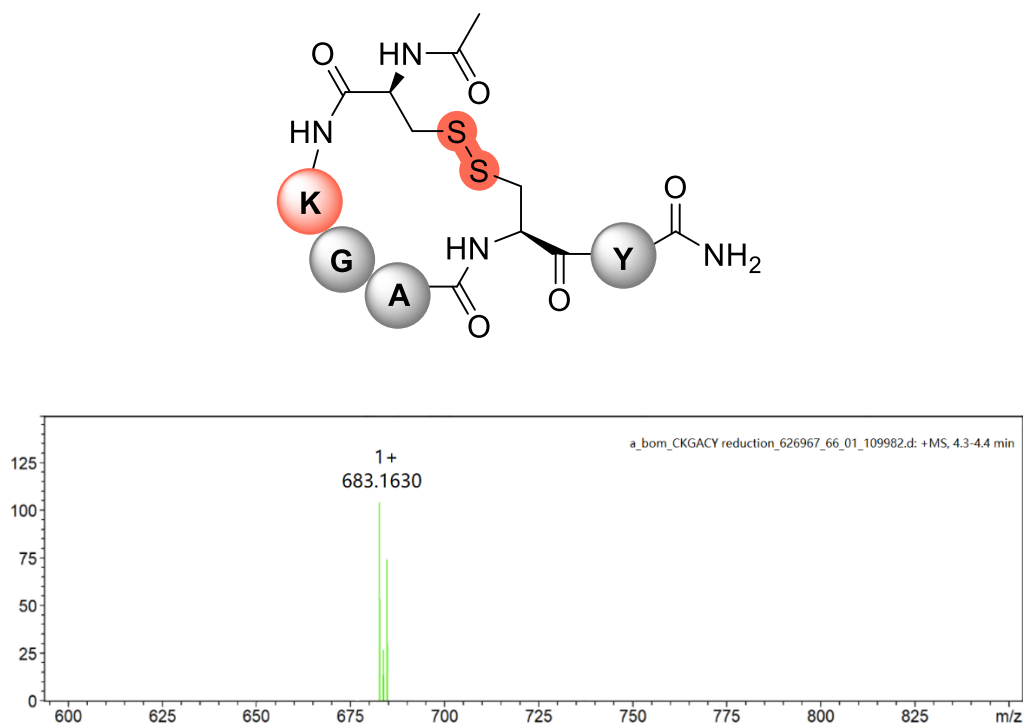


Figure 130 – Structure (top) and HRMS spectrum (bottom) for compound **30b**. Calculated mass: 682.26  $[M+H]^+$ .

Compound **40** was identified as the main component of the reaction mixture at time zero, but its concentration dropped over the monitored period, as showed by the decrease in its absorbance peak in Figure 127. This trend could imply that this species

is an intermediate of the reduction profile, but it is not stable and either progressively degrades into other products, or the presence of HOCl in solution further oxidises it. LC-MS helped in characterising this compound as the thiosulfinate **40** shown in Figure 131. The disulphide bond is retained in this compound, but one Cys residue is oxidised by HOCl. The oxidation state of sulfur goes from -1 when involved in the disulphide bond, to 0 when the disulphide-S-monoxide is formed. As previously described by Nybo *et al.*, the thiosulfinate is an intermediate of the oxidation process through which the disulphide bond undergoes before full cleavage of the S-S bridge.<sup>195</sup> This explains why its concentration experiences such a steep decline as this intermediate is further oxidised to yield other oxidation products.

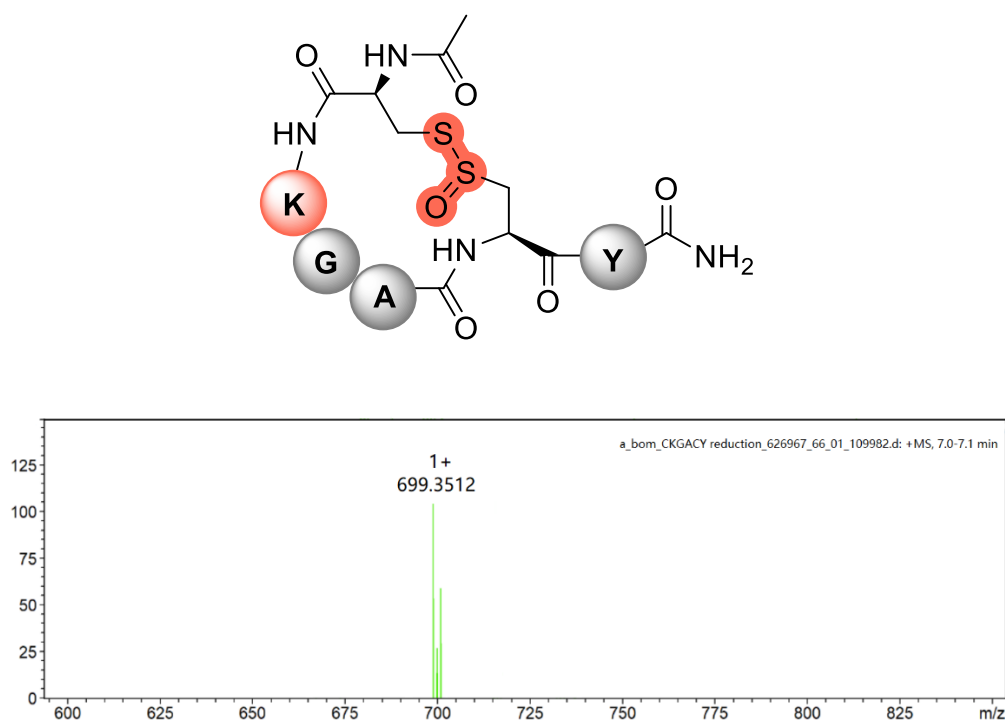


Figure 131 – Structure (top) and HRMS spectrum (bottom) for compound **40**. Calculated mass: 699.20 [M+H]<sup>+</sup>.

Further oxidation of the thiosulfinate led to formation of the thiosulfonate **36** shown in Figure 132, as proved by the mass recorded in the respective spectrum. The oxidation state of the sulfur atom involved in the thiosulfonate bond is +3, with respect to 0 when the sulfur atom was involved in the thiosulfinate, and -1 when it was involved in the disulphide bond.

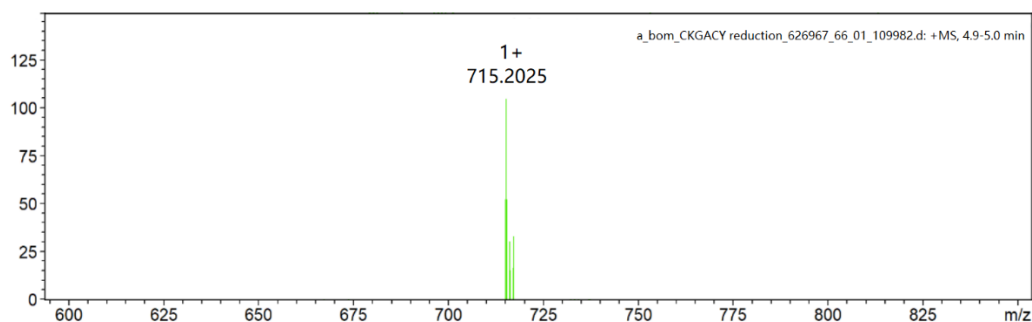
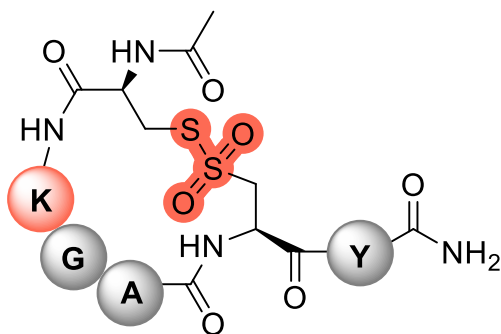


Figure 132 – Structure (top) and HRMS spectrum (bottom) for compound **36**. Calculated mass: 714.25 [M+H]<sup>+</sup>.

The other three peaks identified by LC-MS, were recognised as oxidation products obtained from the oxidation of one of the two Cys residues involved in the disulphide bridge after its cleavage. All these three products increased in concentration over the monitored time, meaning that the reaction equilibrium was shifted towards their formation as the reaction went to completion.

Compound **37** was isolated as the sulfonic acid showed in Figure 133, and its structure was confirmed by the respective mass spectrum.

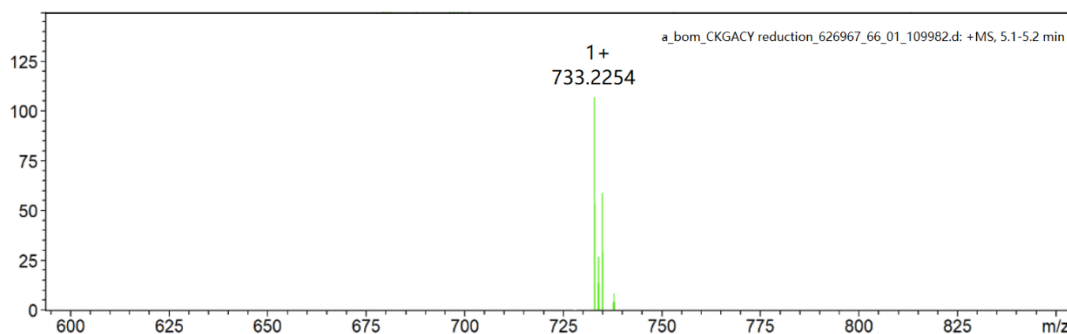
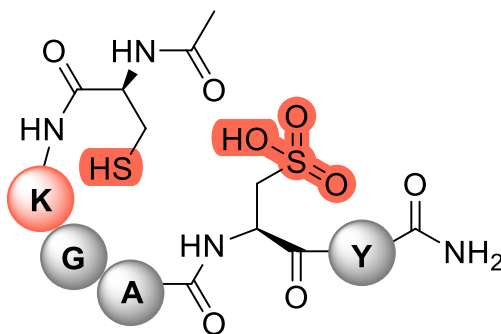


Figure 133 – Structure (top) and HRMS spectrum (bottom) for compound **37**. Calculated mass: 732.26 [M+H]<sup>+</sup>.

**37** was the oxidation product recorded in highest concentration at the end of the monitored period. This result means that the equilibrium of the disulphide bond cleavage and further oxidation of the participating Cys residues is driven towards full oxidation of the sulfur atom. The oxidation state goes indeed from -1 on the sulfur involved in the disulphide bond, to +4 on the sulfur taking part in the sulfonic acid moiety.

**38** was eluted at a similar time to the sulfonic acid, and it was identified as the sulfinic acid showed in Figure 134, as confirmed by the mass spectrum.

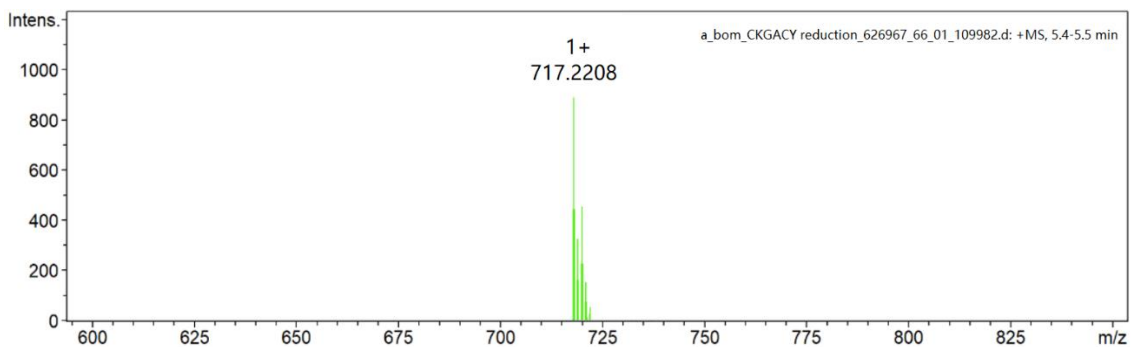
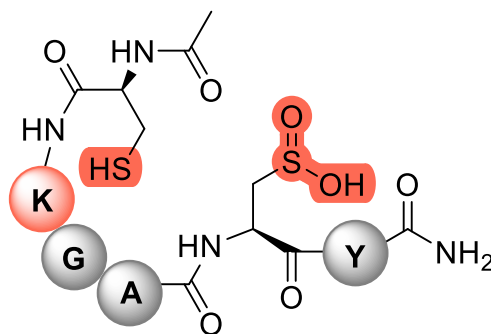


Figure 134 – Structure (top) and HRMS spectrum (bottom) for compound **38**. Calculated mass: 716.26 [M+H]<sup>+</sup>.

The oxidation state of sulfur in **38** goes from -1 when taking part in the disulphide bond, to +2 when the sulfenic acid is produced.

The final mass spectrum reported by LC-MS matched the structure of the sulfenic acid **39** showed in Figure 135. In this case, cleavage of the disulphide bridge led to oxidation of the sulfur atom from -1 to 0. Interestingly, this compound was not reported by Karimi *et al.*, where only the di-sulfenic acid was observed.<sup>190</sup>

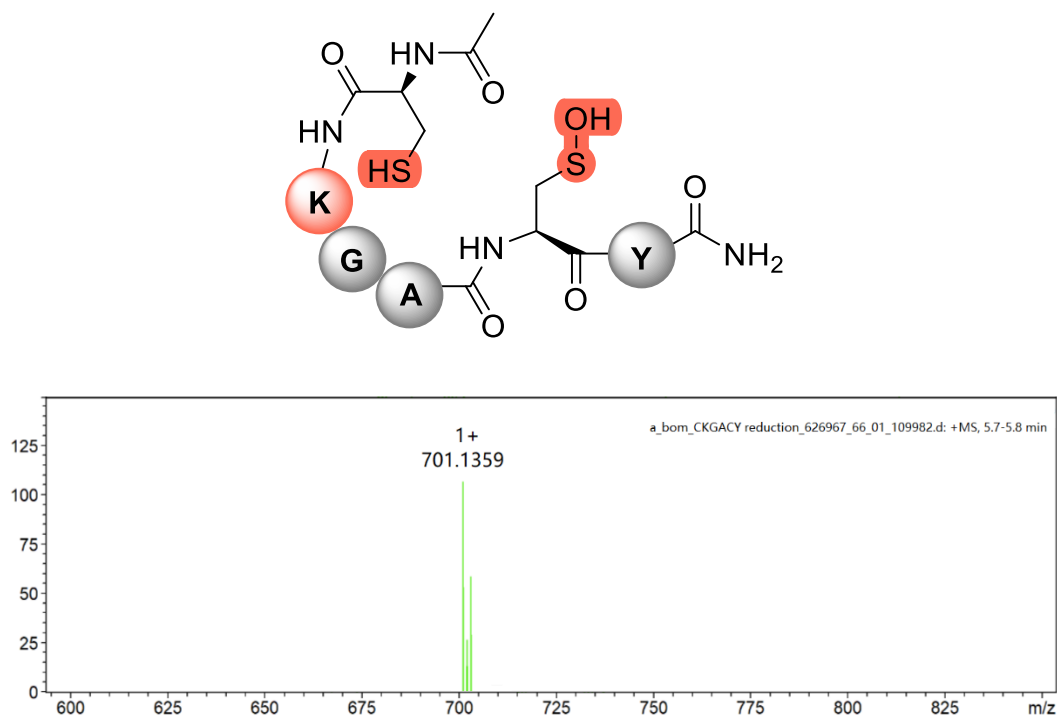


Figure 135 – Structure (top) and HRMS spectrum (bottom) for compound **39**. Calculated mass: 700.27 [M+H]<sup>+</sup>.

It is noteworthy to mention that compounds **36** and **37** were the most abundant oxidation products in the reaction mixture at the end of the monitored period and presented the two highest oxidation state on the sulfur atom, respectively +4 and +2. This outcome emphasises how disulphide bonds can be a critical tool to ensure effective quenching of reactive species that could hamper the correct functioning of living organisms. They provide an active nucleophile for all species containing an electrophilic oxygen that could potentially lead to oxidative damage if not blocked.

Figure 136 shows the relative area percentage trend for the five different main products identified at the end of the disulphide bond reduction monitoring.



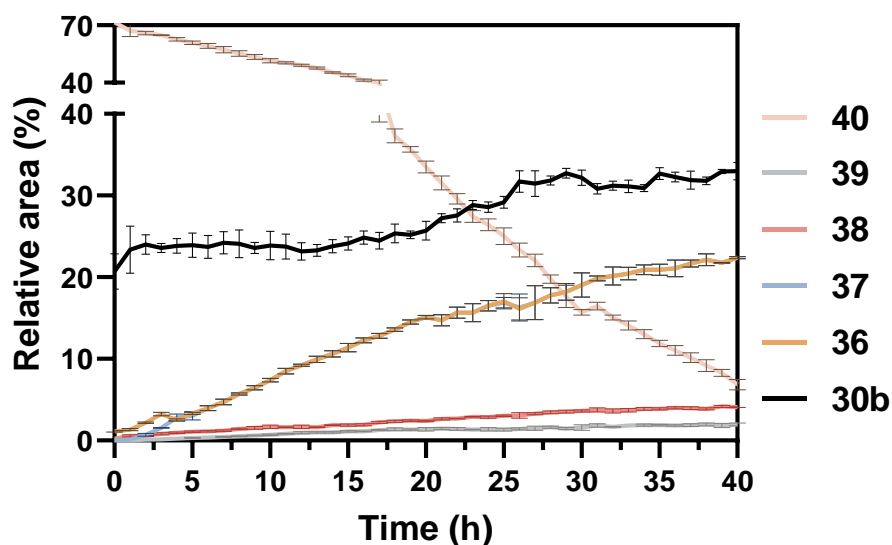


Figure 136 – Relative area (%) vs. time (h) for the five species characterised after 40 hours from exposure of the cyclised peptide **CKGACY (30b)** to HOCl.

As previously discussed, compound **36** was the most abundant at the beginning of the reaction monitoring, and its relative area progressively decreased from an average of 63% to an average of 7%. Compounds **32** and **33** followed a similar trend, being almost absent at the beginning of the reaction, but their relative areas gradually incremented to respectively 22% and 30% by the end of the monitored period, making them the second and third most abundant products in the reaction mixture. These two compounds also had the two highest oxidation states for the sulfur atom, respectively +3 for the thiosulfonate and +4 for the sulfonic acid. Therefore, it is plausible that the reaction equilibrium is shifted towards the most oxidised intermediates in the conditions we reported. However, due to the similar moiety of **32** and **36**, the thiosulfonate was expected to follow a similar trend to the thiosulfinate, and thus being in high concentration at the beginning of the reaction, before dropping in concentration to yield products where the disulphide bridge was cleaved. The sulfenic acid **34** was the second least abundant species throughout the course of the monitored period, and its final relative area was 4%. The products in lowest concentration were the sulfinic acid **35**, the relative area of which reached only 2% at the end of the monitored period. The cyclic peptide **30b** did not experience a noticeable fluctuation over the course of the reaction and the relative area overall increased by less than 10%. It is plausible that after the initial oxidation carried by HOCl to yield compound **36**, the cyclic peptide was not affected by the presence of the oxidative agent. The equilibrium for **30b** seemed to be slightly reversible, with the cyclic peptide being partially reformed, as

described by the moderate increase in relative area percentage that went from an average of 24% to an average of 33%.

#### 6.3.2.4 Monitoring of Disulphide Bridge Reduction in Synthetic Peptide Ac-CEGACY-NH<sub>2</sub>

The last experiment was performed using synthetic peptide CEGACY (**31b**) as the substrate for monitoring the oxidation activity of HOCl. An intramolecular disulphide bridge was installed as previously described for CAGACY (**29b**) and CKGACY (**30b**), and the reduction of the disulphide bond and further oxidation of the sulfhydryl moiety was investigated.

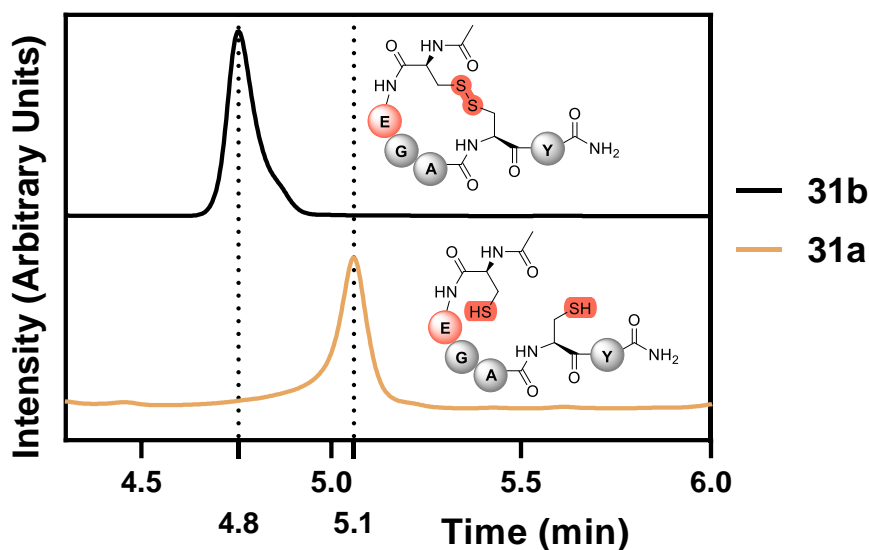


Figure 137 – Comparison of the UPLC traces for the reduction of the disulphide bond on peptide CEGACY (**31b**).

Figure 137 shows the chromatograms for both the reduced (**31a**) and oxidised (**31b**) peptides and their structures.

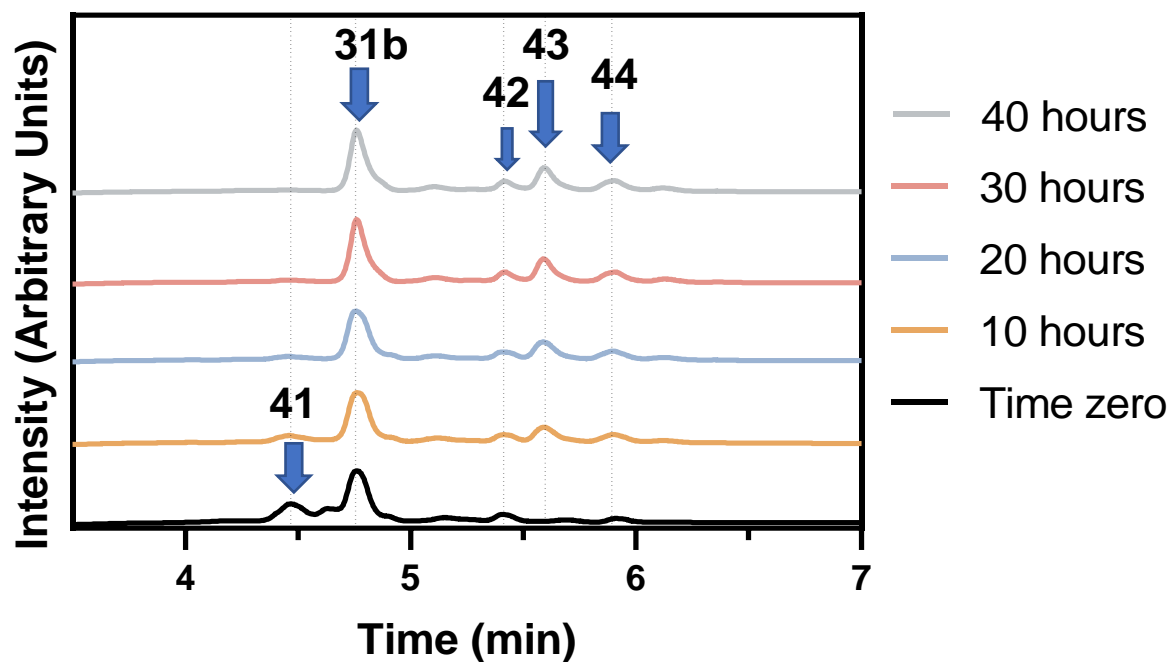


Figure 138 – UPLC traces showing the reaction profile for the reduction of the disulphide bond on peptide CEGACY (**31b**) from time zero (black trace) to 40 hours (light-grey trace).

The chromatogram at time zero showed in Figure 138 described a species eluting at approximately 4.5 minutes. However, the relative area of this compound progressively decreased over the course of the monitored period, and it was no longer present in the final sample after 40 hours. To characterise this product, the reaction was repeated and analysed immediately with LC-MS. The peak was identified with the thiosulfinate **41** showed in Figure 139. The structure of this intermediate was confirmed by the mass spectrum and the oxidation state of the sulfur atom increased from -1 in the disulphide bond to +1 in the thiosulfinate moiety.

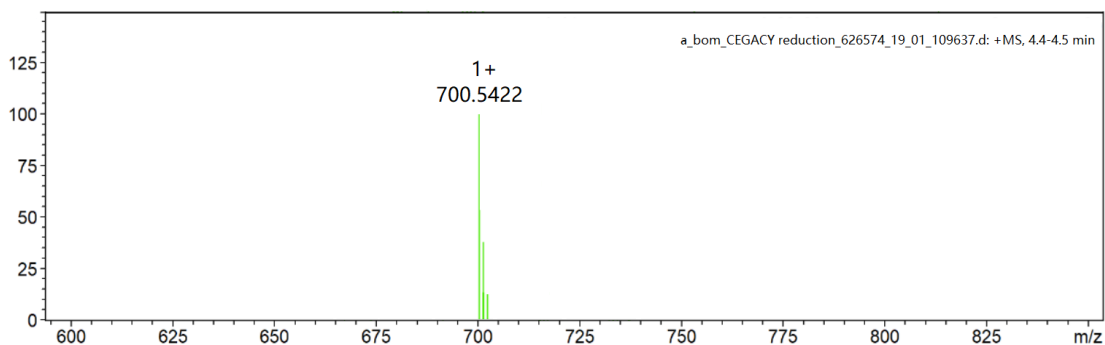
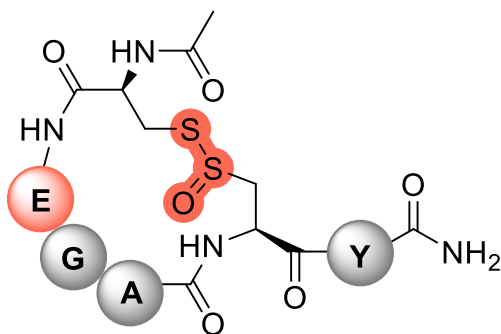


Figure 139 – Structure (top) and HRMS spectrum (bottom) for compound **41**. Calculated mass: 699.20 [M+H]<sup>+</sup>.

In contrast to the thiosulfinate **40** produced by synthetic peptide **CKGACY** (Figure 128), product **41** was eluted first in the analytical method and was therefore the most polar species. It must also be mentioned that compound **40** was the most abundant at the beginning of the monitored period, and despite its concentration progressively decreased, it was still observed after 40 hours. On the other side, compound **41** not only was second in concentration to the unreacted cyclised peptide **31b** but was also not stable enough to survive the 40 hours of monitoring. Because of the high similarity between the two oxidation products **40** and **41**, the difference that resulted in such an opposed outcome could only be addressed with the amino acid residue next to the Cys at the *N*-terminus of the synthetic peptide. It is plausible that Lys could interact with the thiosulfinate moiety to help and stabilise it, whereas the Glu residue cannot provide the same extent of interaction and stabilisation.

The second compound eluting in the chromatogram showed in Figure 138 was identified with unreacted cyclised peptide **31b**. The mass spectrum in Figure 140 confirmed the molecular weight of the oxidised peptide.

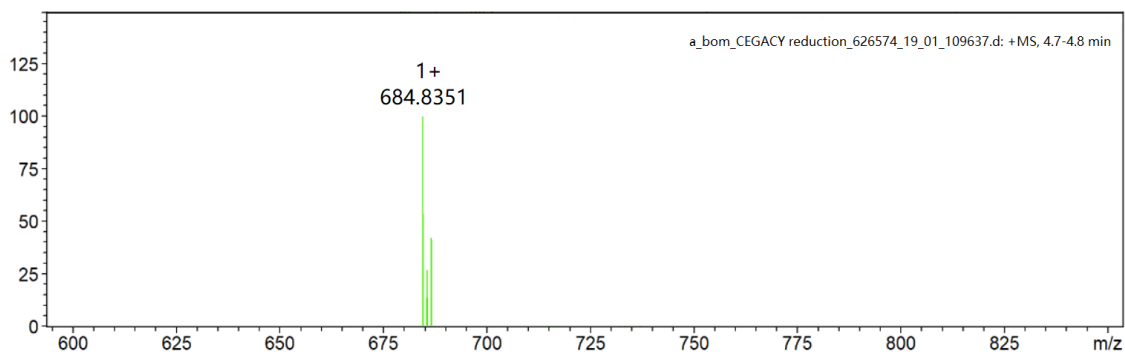
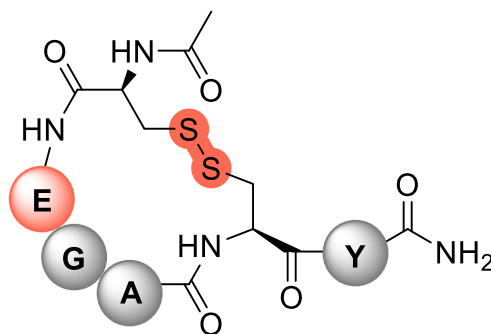


Figure 140 – Structure (top) and HRMS spectrum (bottom) for compound **31b**. Calculated mass: 683.20  $[M+H]^+$ .

The mass spectrum showed in Figure 141 matched the structure of the thiosulfonate **42**. The oxidation state of the sulfur atom involved in this intermediate is +3, with respect to an oxidation state of -1 when the sulfur atom takes part in the disulfide bridge.

The most abundant peak among the oxidised products showed in the chromatogram of Figure 138 was characterised as the sulfonic acid **43**, as proved by the mass spectrum in Figure 142. In this case the oxidation state of the sulfur atom increased to +4.

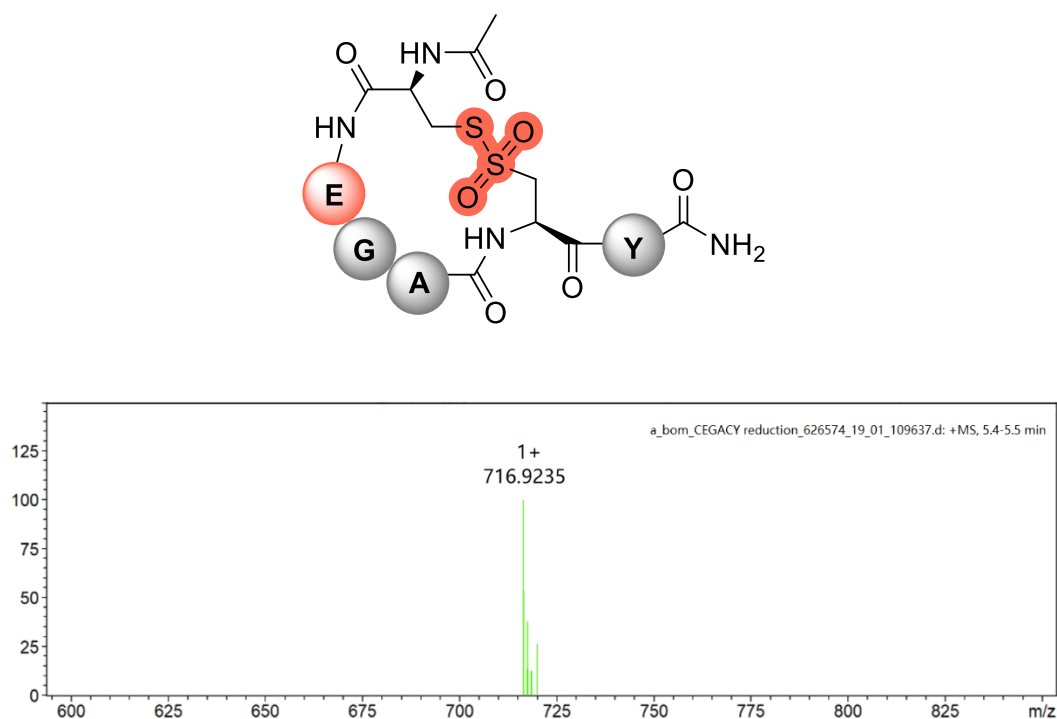


Figure 141 – Structure (top) and HRMS spectrum (bottom) for compound **42**. Calculated mass: 715.19 [M+H]<sup>+</sup>.

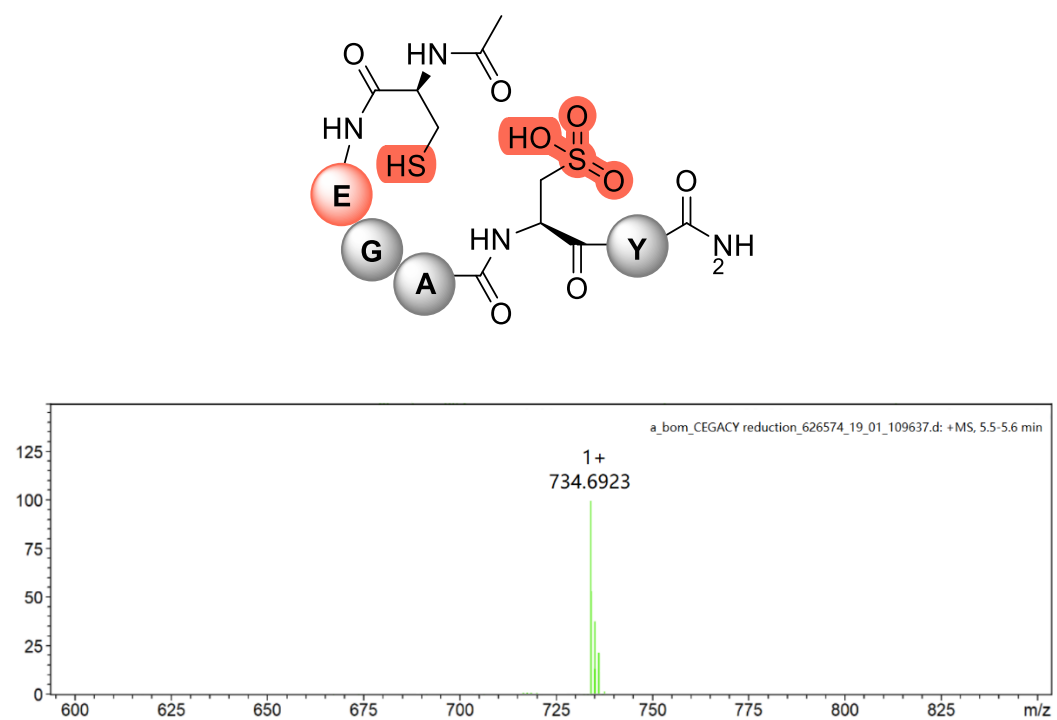


Figure 142 – Structure (top) and HRMS spectrum (bottom) for compound **43**. Calculated mass: 733.20 [M+H]<sup>+</sup>.

In agreement with the findings observed for synthetic peptide CKGACY (**30b**), the most stable oxidation product for synthetic peptide CEGACY (**31b**) was again the

sulfonic acid **43**. This outcome corroborated the hypothesis by which disulphide bridges represent an effective sink for electrophilic oxygen trapped in reactive species. Moreover, this species has one of the higher oxidation states for the sulfur atom (+3) and it is therefore plausible that the equilibrium of the reaction is shifted towards the most oxidised products.

The last compound to be eluted was the sulfinic acid **44**, and its structure was matched by the mass spectrum displayed in Figure 143. While the sulfur atom involved in the disulphide bond has an oxidation state of +1, the sulfur atom taking part in the sulfinic acid moiety shows an oxidation level of +2.

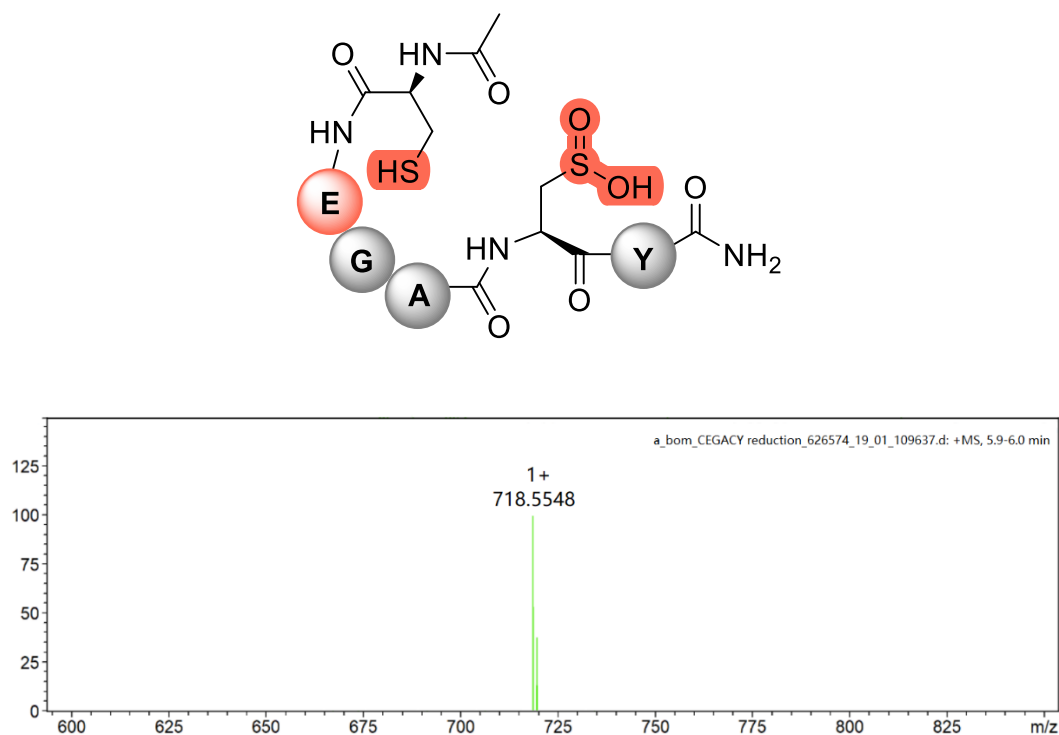


Figure 143 – Structure (top) and HRMS spectrum (bottom) for compound **44**. Calculated mass: 717.21 [M+H]<sup>+</sup>.

As displayed in Figure 144, the cyclised peptide **31b** did not experience a significant fluctuation in the relative area percentage and remained steady at approximately 55% throughout the monitored period.

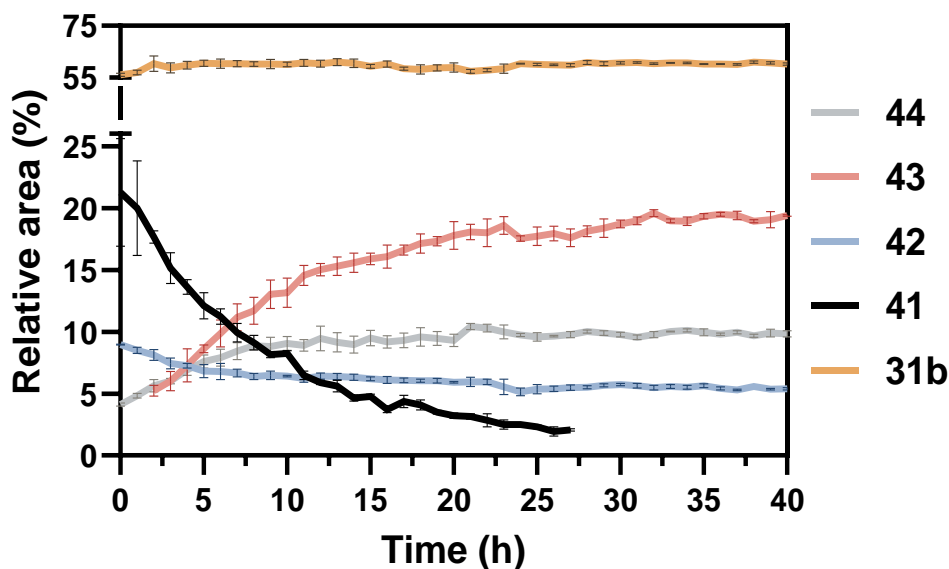


Figure 144 – Relative area (%) vs. time (h) for the five species characterised after 40 hours from exposure of the cyclised peptide CEGACY (**31b**) to HOCl.

The thiosulfinate **41** was in moderate concentration at the beginning of the reaction monitoring with a relative area of 20%, but it dropped below a detectable level over 26 hours. Thiosulfonate **42** had a steady trend over the 40 hours of monitoring, with a final relative area of 5%. The two oxidation products formed after cleavage of the disulphide bond, respectively the sulfonic acid **43** and the sulfinic acid **44**, where the second and third most abundant species at the end of the monitored period. The average of the relative area percentage for **43** went from 0% to 19%, and that of **44** increased from 5% to 10%.

## 6.4 CONCLUSION AND FUTURE WORK

Molecular dynamics studies were employed to evaluate how the presence of different amino acid residues, and thus variations in local steric and electrostatic environment, affected the oxidation of an intra-molecular disulphide bond. The results were accompanied by experimental data, proving that the electronic environment that



surrounds the Cys residues involved in the disulphide bond plays a role in its stabilisation and cleavage in presence of the oxidative agent HOCl. In particular, the presence of a Lys or Glu residues next to one of the Cys involved in the disulphide bond decreased the enthalpy values for the respective intermediates. This means that those intermediates were overall more stable than those related to peptide CAGACY (red **29a** and ox **29b**).

The characterisation of the oxidation products formed after treatment of the cyclic peptides with HOCl clearly showed different trends depending on the nature of the amino acid residues proximal to the sulfur atoms involved in the disulphide bond. Cleavage of the disulphide bond in synthetic peptide CAGACY showed little conversion to the oxidation products. These results indicate that a neutral charge close to the disulphide bridge possibly hampers the activity of HOCl. The presence of a positive charge in the surroundings of the disulphide bond, as per synthetic peptide CKGACY (red **30a** and ox **30b**), led to complete conversion of the cyclic peptide to a thiosulfinate intermediate. The two most abundant species at the end of the monitored period, the thiosulfinate and the sulfonic acid, carried the highest oxidation states on the sulfur atom. This outcome demonstrates that oxidation of the disulphide bond is favoured in presence of a positive charge, and the equilibrium is shifted towards the highest oxidation state on the sulfur atoms. In case of synthetic peptide CEGACY (red **31a** and ox **31b**), a negative charge proximal to the disulphide bond seemed to have impeded the oxidation activity of HOCl. The thiosulfinate was again the oxidation intermediate formed at the early stages, but the cyclic peptide was overall the most abundant species at the end of the monitored period.

The results obtained experimentally were in partial agreement with the computational studies. The latter showed that presence of oxygen and nitrogen atoms on the side chain proximal to the sulfur atom involved in the disulphide bond, stabilises the S-S-Cl intermediate with electrostatic interactions. Despite synthetic peptide CKGACY showed a macro cyclisation that was favoured by the chlorine atom interacting with two oxygen atom and one nitrogen atom, synthetic peptide CEGACY recorded the lowest enthalpy due to formation of a five-membered ring that strongly stabilised the intermediate.

Overall, the results gathered in this chapter expanded our understanding regarding the reactivity of disulphide bonds in the presence of oxidative agents. Diverse reactivity was observed when amino acids bearing side chains with different charges were employed. The oxidation products characterised agreed with previously published literature.

## 7 THESIS CONCLUSIONS AND FUTURE WORK

In this thesis we have described some of the applications of Cys-containing peptides at the frontiers between chemistry and biology. Cys is an amino acid that, thanks to its unique reactivity, finds broad and diverse uses in biological systems. The aim of this thesis is to enlarge the number of these uses and understand how these can be implemented in everyday life to support development. Being Cys the main focus, we studied its applicability both in biological processes and in chemical modification reactions.

Chapter 2, 3, and 4 looked at the production and consequent application of a [4Fe-4S]-cluster maquette where the Cys residues of a synthetic peptide coordinate the iron ions of an inorganic cluster. This system was applied to a biological process where synchronized redox events involving the synthetic maquette, H<sub>2</sub>, and a hydrogenase led to the production of H<sup>+</sup>. These three chapters described the full process of a project; beginning by engineering a synthetic peptide that could successfully incorporate a [4Fe-4S]-cluster, moving through the analysis of the maquette to determine the redox potential, and finding a suitable application in a biological system. This case study demonstrates the application of a synthetic maquette in an electron transport process that could in future be further investigated and optimized to find an industrial application. Initial optimization of the method for the incorporation of a [4Fe-4S]-cluster in a synthetic peptide proved challenging, as described in Chapter 1. The method required various adjustments both in terms of peptide sequence and concentration of the different reagents to achieve satisfactory results. Despite successful synthesis of a [4Fe-4S]-cluster maquette being eventually obtained using a small library of peptides, the overall yields were modest as showed by the EPR data. Future efforts could be focused at optimising the reconstitution process to improve the yield of this methodology, perhaps exploring different buffer solutions, order of addition of the reagents, a different reducing agent or diverse peptide sequences. The experiments described in Chapter 3 looked at determining the redox potentials of the synthesized [4Fe-4S]-cluster maquettes. The measurements were obtained with success and used in Chapter 4 to accurately select a suitable electron transport pathway to catalyse. However, this experiment could be further explored in the future by analysing a sample reconstituted in presence of  $\beta$ ME as this would allow for a

higher concentration of [4Fe-4S]-cluster maquette. Chapter 4 described the use of a [4Fe-4S]-cluster maquette in combination with H<sub>2</sub> and a hydrogenase to generate H<sup>+</sup>. The results described in this chapter showcase the biotechnological application of a [4Fe-4S]-cluster maquette in the H<sub>2</sub> industry. Limitations of this methodology that could impact its scalability are the strict anaerobic conditions required to ensure the redox process is successful. Future works could be focused on determining whether the experiment is successful at higher concentration of O<sub>2</sub>, and how the redox reaction is affected at a bigger scale.

Chapter 5 looked at the use of Cys residues in peptides as handles to be modified and increase the chemical diversification for, as an example, drug discovery. In this chapter we proposed a novel photochemical modification of Cys residues that via desulfurisation of the thiol functional group yields a nitro-substituted Phe moiety. We envision that this methodology will find application in late-stage modification of Cys residues in both small peptides and proteins to quickly introduce functionalized hydrophobic groups. In the future, the scope of this methodology could be increased by extending the number of boronic acids used. The biologically relevant biotin-inspired boronic acid proved challenging to be synthesized and its preparation could not be achieved. Future efforts could be focused at exploring different synthetic route for its preparation, and the synthesis of other biologically relevant boronic acids to expand the scope of this methodology as a bioconjugation tool. Furthermore, more complex peptides and small proteins should be tested as a substrate for this synthetic method to prove its robustness in more elaborated systems.

In Chapter 6 we investigated the stability of intramolecular disulphide bonds in small peptides under exposure to oxidative conditions. Disulphide bonds are employed in cells as a defense against oxidative damage thanks to the wide oxidation state range of sulfur. In our project, various amino acids proximal to one of the Cys residues involved in the disulphide bond were installed to determine the effect of steric and electronic to the overall stability of the disulphide bridge. Molecular dynamics simulations were used to model the bond lengths and reaction enthalpies for the cleavage of the disulphide bond. Experimental data aimed at determining the main products of the disulphide bond cleavage. We hope that merging computational simulations and experimental data on a wider library of peptides containing

intramolecular disulphide bridges could produce a tool to aid synthesis of more stable peptides and proteins that can better respond towards oxidative damage. Further studies could be aimed at expanding the library of peptides analysed under oxidative conditions. For example, longer peptide sequences should be included to better mimic larger proteins. The intramolecular disulphide bridge could be installed between Cys residues positioned further away than three amino acids to investigate how the size of the cyclised peptide affects its stability under oxidative conditions. Computational studies could be further explored for instance by running molecular dynamics simulations of the peptide sequences in solution at physiological pH, to gather more information about the most stable conformation of the peptides.

Overall, we have positively impacted the common knowledge about Cys and Cys-containing peptides in both biology and chemistry. Crucial to this amino acid, we have described and further investigated biological processes and chemical modification that could have a meaningful repercussion on the scientific community and on our society.

## 8 EXPERIMENTAL

### 8.1 GENERAL METHODS

**Nuclear Magnetic Resonance (NMR) Spectroscopy:** Samples were analysed on a Bruker AVII 400 NMR system ( $^1\text{H}$ -NMR frequency 400MHz:  $^{13}\text{C}$ -NMR frequency 101MHz). Chemical shifts are given in parts per million (ppm) and coupling constant (J) are given in Hertz (Hz). All spectra are referenced to solvent residual signals:  $\text{CDCl}_3$  ( $\delta$  7.26 [ $^1\text{H}$ ]) ( $\delta$  77.16 [ $^{13}\text{C}$ ]),  $\text{DMSO-}d_6$  ( $\delta$  2.50 [ $^1\text{H}$ ]) ( $\delta$  39.52 [ $^{13}\text{C}$ ]),  $\text{CD}_3\text{OD}$  ( $\delta$  3.31 [ $^1\text{H}$ ]) ( $\delta$  49.00 [ $^{13}\text{C}$ ]).  $^1\text{H}$ -NMR data is reported as chemical shift ( $\delta$ ), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, dd = doublet of doublets, ddd = doublet of doublet of doublets, dt = doublet of triplets, or m = unassigned multiplet), relative integral, coupling constant, and assignment where possible. For  $^{19}\text{F}$  NMR an internal standard was used ( $\alpha,\alpha,\alpha$ -trifluorotoluene: -63.72 ppm).

**High-Resolution Mass Spectrometry (HRMS):** Samples were analysed on a Bruker MicrOTF Focus II by Electron Spray Ionisation (ESI) operating in positive or negative ionisation mode.

**Analytical High-Performance Liquid Chromatography (HPLC):** Samples were analysed on a Thermo Scientific Ultimate 3000 uHPLC system equipped with a photodiode array (eAPDA) detector ( $\lambda$  = 210, 254, 280, and 400 nm). Peptide samples were analysed using a Waters SunFire 5  $\mu\text{m}$ , 2.1 x 150 mm column (C-18) operating at a flow rate of 0.6 mL/min. The mobile phase was composed of 0.1% trifluoroacetic acid (TFA) in water (Solvent A) and 0.1% TFA in acetonitrile (Solvent B). The analysis of the chromatograms was conducted using Chromeleon 7 software.

**Preparative HPLC:** Samples were purified on a Waters 1525 binary pump HPLC equipped with a dual wavelength UV detector set to 210 nm and 280 nm. Purification was performed on a Waters SunFire 5  $\mu\text{m}$ , 19 x 150 mm preparative column (C-18) operating at a flowrate of 6 mL/min. The mobile phase was composed of 0.1% TFA in water (Solvent A) and 0.1% TFA in acetonitrile (Solvent B). The gradient used is specified for each compound.

**Semi-Preparative HPLC:** Samples were purified on a Waters 1525 binary pump HPLC equipped with a dual wavelength UV detector set to 210 nm and 280 nm.

Purification was performed on a Waters SunFire 5  $\mu\text{m}$ , 10 x 150 mm semi-preparative column (C-18) operating at a flowrate of 5 mL/min. The mobile phase was composed of 0.1% TFA in water (Solvent A) and 0.1% TFA in acetonitrile (Solvent B). The gradient used is specified for each compound.

**Electrochemical Measurements:** Experiments were performed under strictly anaerobic conditions at room temperature on a CH Model 600E Series Potentiostat/Galvanostat with a three-electrode cyclic voltammetry configuration. Saturated calomel electrode (SCE) was used as the reference electrode. A platinum wire was used as the counter electrode. Glassy carbon electrode was used as the working electrode. Prior to experiments, the working electrode was polished with alumina slurry starting with 1  $\mu\text{m}$ , followed by 0.3  $\mu\text{m}$ , and 0.05  $\mu\text{m}$  particles. 100 mM NaCl was added to all samples as a supporting electrolyte. CHI600E Electrochemical Analyzer was used for data acquisition and manipulation. Values were reported vs SCE. Redox experiments of the [4Fe-4S]-cluster maquette were evaluated by monitoring faradaic current for both oxidation and reduction processes over ten cycles.

**UV-Visible Spectroscopy:** Experiments were performed at room temperature on an Agilent Cary 5000 UV-Vis-NIR spectrophotometer, controlled by UVProbe software (Version 2.5). All measurements were performed in gas-tight quartz cuvettes with 0.5 cm optical path length. All blank and sample solutions were freshly prepared in anaerobic conditions and immediately analysed.

**Electron Paramagnetic Resonance (EPR) Spectroscopy:** All samples were analysed by Dr Muralidharan Shanmugam at the University of Manchester on a Bruker ELEXSYS-E580 X-band EPR spectrometer with the microwave power set to 30 dB (0.2 mW), the modulation amplitude set to 5 G, a time constant of 41 ms, a conversion time of 41 ms, a sweep time of 84 s, the receiver gain set to 60 dB and an average microwave frequency of 9.384 GHz. All annealing measurements were performed using a 1-propanol and liquid N<sub>2</sub> solvent mixture, and all samples were annealed for the specified times (see main text) at the temperature stated. All samples were measured as a frozen solution at 20 K, unless otherwise indicated. The analysis of the continuous wave EPR spectra was performed using EasySpin toolbox for the Matlab program package (Version 9.2.0.556344).

**Molecular Dynamics:** MobaXterm was used as Unix terminal for Windows to access the High-Performance Computing (HPC) provided with the Operating System (OS) Ubuntu 18.4. and cuda 10.1 (Nvidia Quadro K1200). The program used for MD was AMBER 16 (cuda). 3D visualisations of the MD were run using VMD 1.9.3 (cuda).

## Materials

Unless otherwise indicated, commercial materials were used as received. Amino acids, coupling reagents, and resins were obtained from CEM, GL Biochem, Fluorochem, or Novabiochem. Reagents that were not commercially available were synthesised as outlined for each specific compound.

## Solid Phase Peptide Synthesis (SPPS)

**Preloading Rink Amide Resin:** Rink amide resin was initially swollen in anhydrous  $\text{CH}_2\text{Cl}_2$  for 30 min and then washed with  $\text{CH}_2\text{Cl}_2$  (5 x 3 mL) and DMF (5 x 3 mL). This was followed by removal of the Fmoc group via treatment with 20% piperidine/DMF (3 mL, 2 x 5 min). The resin was washed with DMF (5 x 3 mL),  $\text{CH}_2\text{Cl}_2$  (5 x 3 mL) and DMF (5 x 3 mL). DIC (4 eq.) and Oxyma (4 eq.) were added to a solution of Fmoc-AA-OH (4 eq.) in DMF (final concentration 0.1 M of amino acid). After 5 min of pre-activation, the mixture was added to the resin and shaken at r.t. for 2 h. The resin was washed with DMF (5 x 3 mL),  $\text{CH}_2\text{Cl}_2$  (5 x 3 mL) and DMF (5 x 3 mL). The resin was consequently submitted to iterative peptide assembly (Fmoc-SPPS).

**Preloading 2-Chlorotrityl Chloride Resin:** 2-Chlorotrityl chloride (CTC) resin was initially swollen in anhydrous  $\text{CH}_2\text{Cl}_2$  for 30 min and then washed with  $\text{CH}_2\text{Cl}_2$  (2 x 3 mL). A solution of Fmoc-AA-OH (0.5 eq. relative to resin functionalisation) and *i*Pr<sub>2</sub>Net (2 eq. relative to resin functionalisation) in  $\text{CH}_2\text{Cl}_2$  (final concentration 0.1 M of amino acid) was added and the resin shaken at r.t. for 16 h. The resin was washed with DMF (5 x 3 mL) and  $\text{CH}_2\text{Cl}_2$  (5 x 3 mL). The resin was treated with a solution of  $\text{CH}_2\text{Cl}_2$ /MeOH/*i*Pr<sub>2</sub>Net (17:2:1 v/v/v, 3 mL) for 1 h and washed with DMF (5 x 3 mL),  $\text{CH}_2\text{Cl}_2$  (5 x 3 mL), and DMF (5 x 3 mL). The resin was consequently submitted to iterative peptide assembly (Fmoc-SPPS).

**Determination of Amino Acid Loading:** The resin was treated with 20% piperidine/DMF (3 mL, 2 x 3 min – the volume was accurately measured) and washed



with DMF (5 x 3 mL), CH<sub>2</sub>Cl<sub>2</sub> (5 x 3 mL) and DMF (5 x 3 mL). Both the 20% piperidine/DMF solutions were combined and collected; 20 µL were placed into a clean volumetric flask and diluted to 10 mL in 20% piperidine/DMF. The UV absorbance of the resulting fulvene-piperidine adduct ( $\epsilon = 7800 \text{ M}^{-1} \text{ cm}^{-1}$ ) was measured at  $\lambda = 301 \text{ nm}$  using a solution of 20% piperidine/DMF as the blank. The loading was determined using the following equation:

$$mol = \frac{A \times df \times V}{\epsilon}$$

Where A = absorbance, df = dilution factor, and V = volume of 20% piperidine/DMF.

### **Manual Iterative Peptide Assembly (Fmoc-SPPS)**

**General Amino Acid Coupling:** A solution of protected amino acid (4 eq.), DIC (4 eq.), and Oxyma (4 eq.) were dissolved in DMF (final concentration 0.1 M of amino acid). After 5 min of pre-activation, the mixture was added to the resin and shaken at r.t. for 1 h. The resin was washed with DMF (5 x 3 mL), CH<sub>2</sub>Cl<sub>2</sub> (5 x 3 mL) and DMF (5 x 3 mL).

**Capping:** Acetic anhydride/pyridine (1:9 v/v, 3 mL) was added to the resin and shaken for 3 min. The resin was washed with DMF (5 x 3 mL), CH<sub>2</sub>Cl<sub>2</sub> (5 x 3 mL) and DMF (5 x 3 mL).

**Fmoc Group Deprotection:** The resin was treated with 20% piperidine/DMF (3 mL, 2 x 3 min) and washed with DMF (5 x 3 mL), CH<sub>2</sub>Cl<sub>2</sub> (5 x 3 mL) and DMF (5 x 3 mL).

**Cleavage:** For peptide containing Cys: A mixture of TFA, triisopropylsilane (TIPS), thioanisole, and water (85:5:5:5 v/v/v/v) was added to the dried resin and shaken at r.t. for 3 h. The resin was washed with TFA (3 x 2 mL). For peptide containing Sec (U): a mixture of TFA, trimethylsilyl trifluoromethanesulfonate (TMSOTf), thioanisole, and *m*-cresol (66:18:11:5 v/v/v/v) was added to the dried resin and shaken at -4°C for 1 h. The resin was washed with TFA (3 x 2 mL). For all other peptides: A mixture of TFA, triisopropylsilane (TIPS), and water (90:5:5 v/v/v) was added to the dried resin and shaken at r.t. for 3 h. The resin was washed with TFA (3 x 2 mL).

**Work-Up:** The combined solutions were concentrated under a stream of N<sub>2</sub> to <5 mL. Diethyl ether (40 mL) was added to precipitate the peptide and the suspension was centrifuged at 4000 rpm for 10 min. The pellet was dissolved in a mixture of 0.1% TFA in water/acetonitrile, purified by preparative HPLC and analysed by HRMS.

### **Automated Peptide Assembly (Fmoc-SPPS)**

Automated Fmoc-SPPS was performed either on a Biotage Initiator<sup>+</sup> Alstra or on a CEM Liberty Blue 2.0 peptide synthesiser. General synthetic procedures were carried out in accordance with the specifications detailed by the manufacturers. Final peptide cleavage and work-up were performed as previously described for manual SPPS.

## **8.2 GENERAL PROTOCOLS**

General Protocols for Chapter Two

### **General Reconstitution Protocol for [4Fe-4S]-Cluster Maquette with $\beta$ ME in Anaerobic Conditions**

Methodologies for cluster assembly into the required apo-peptide were adapted from previous protocols.<sup>20,75</sup> All steps of the reconstitution procedure were performed, unless otherwise indicated, in strict anaerobic conditions in an MBraun UNIlab LMF Glovebox Workstation at O<sub>2</sub> concentrations  $\leq$  0.1 ppm. All solutions and buffers were degassed as described later in the text.

Peptide (1  $\mu$ mol) was equilibrated under gentle stirring with  $\beta$ ME (200  $\mu$ L, 10% v/v of final solution) and 50 mM HEPES 10 mM KCl buffer at pH = 8.0 (1560  $\mu$ L) at r. t. for 30 min. FeCl<sub>3</sub> (120  $\mu$ L of a 50 mM solution) was added in twenty increments (20 x 6  $\mu$ L) over a period of 30 min. Na<sub>2</sub>S (120  $\mu$ L of a 50 mM solution) was added in ten increments (10 x 12  $\mu$ L) over a period of 10 min. The reaction mixture was equilibrated under gentle stirring at r.t. for 1 hour, before being centrifuged (13 000 rpm for 5 min). The supernatant was discharged, and the solution was transferred into an Eppendorf tube.

After reconstitution, maquettes solutions were diluted with 50 mM HEPES 10 mM KCl buffer at pH = 8.0 prior to UV-Vis spectroscopy analysis to various concentrations, as detailed in the text.

### **General Reconstitution Protocol for [4Fe-4S]-Cluster Maquette with Dithiothreitol in Anaerobic Conditions**

Methodologies for cluster assembly into the required apo-peptide were adapted from previous protocols.<sup>20,75,92</sup> All steps of the reconstitution procedure were performed, unless otherwise indicated, in strict anaerobic conditions in an MBraun UNIlab LMF Glovebox Workstation at O<sub>2</sub> concentrations ≤ 0.1 ppm. All solutions and buffers were degassed as described later in the text.

Peptide (1 μmol) was equilibrated under gentle stirring in a solution of DTT (0.7 mg, 5 μmol, 5 eq. with respect to the peptide) in 50 mM HEPES 10 mM KCl buffer at pH = 8.0 buffer (1560 μL) at r. t. for 30 min. FeCl<sub>3</sub> (120 μL of a 50 mM solution) was added in twenty increments (20 x 6 μL) over a period of 30 min. Na<sub>2</sub>S (120 μL of a 50 mM solution) was added in ten increments (10 x 12 μL) over a period of 10 min. The reaction mixture was equilibrated under gentle stirring at r.t. for 1 hour, before being centrifuged (13 000 rpm for 5 min). The supernatant was discharged, and the solution was transferred into an Eppendorf tube.

After reconstitution, maquettes solutions were diluted with 50 mM HEPES 10 mM KCl buffer at pH = 8.0 prior to UV-Vis spectroscopy analysis to various concentrations, as detailed in the text.

### **General Reconstitution Protocol for [4Fe-4S]-Cluster Maquette without βME in Anaerobic Conditions**

Methodologies for cluster assembly into the required apo-peptide were adapted from previous protocols.<sup>20,75</sup> All steps of the reconstitution procedure were performed, unless otherwise indicated, in strict anaerobic conditions in an MBraun UNIlab LMF Glovebox Workstation at O<sub>2</sub> concentrations ≤ 0.1 ppm. All solutions and buffers were degassed as described later in the text.

Peptide (1  $\mu\text{mol}$ ) was equilibrated under gentle stirring in 50 mM HEPES 10 mM KCl buffer at pH = 8.0 buffer (1760  $\mu\text{L}$ ) at r.t. for 5 min.  $\text{FeCl}_3$  (120  $\mu\text{L}$  of a 50 mM solution) was added in twenty increments (20 x 6  $\mu\text{L}$ ) over a period of 30 min.  $\text{Na}_2\text{S}$  (120  $\mu\text{L}$  of a 50 mM solution) was added in ten increments (10 x 12  $\mu\text{L}$ ) over a period of 10 min. The reaction mixture was equilibrated under gentle stirring at r.t. for 1 hour, before being centrifuged (13 000 rpm for 5 min). The supernatant was discharged, and the solution was transferred into an Eppendorf tube.

After reconstitution, maquettes solutions were diluted with 50 mM HEPES 10 mM KCl buffer at pH = 8.0 prior to UV-Vis spectroscopy analysis to various concentrations, as detailed in the text.

### **General Protocol for Degassing of Buffer Solution – Freeze-Pump-Thaw Method**

The buffer used for the cluster-maquette reconstitution protocol was HEPES 50 mM KCl 10 mM at pH = 8.0. The buffer solution was prepared in aerobic conditions using Milli-Q water. Once prepared, the buffer had a pH  $\sim$  4.5, and was adjusted to the required value using a solution of NaOH 10 M in Milli-Q water. The buffer solution was transferred into a Schlenk round-bottom flask equipped with a glass stopcock and sealed with a glass lid. Vacuum grease was applied to both the stopcock and lid to ensure better sealing. The round-bottom flask was connected to the Schlenk line with the stopcock closed and immersed in liquid  $\text{N}_2$  to completely freeze the buffer solution (**Figure 148-A**). The stopcock was opened, and the frozen buffer solution was exposed to active vacuum for 5 minutes while the flask was kept in liquid  $\text{N}_2$  (**Figure 148-B**). After 5 minutes, the stopcock was closed, and the buffer solution was allowed to thaw (a warm-water bath was used to speed the process) (**Figure 148-C**). The presence of bubbles escaping from the solution while this is thawing indicates the presence of dissolved gases in solution (**Figure 148-D**). The freeze-pump-thaw procedure was repeated four times to ensure no gases were dissolved in the buffer solution. After the final repetition, the stopcock was closed, and the buffer solution was allowed to thaw. The flask was disconnected from the Schlenk line, and a plastic clip was used to hold the glass lid. Finally, the flask was transferred into the glovebox.



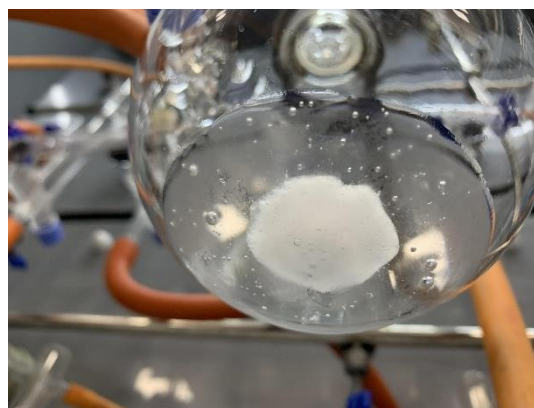
**A**



**B**



**C**



**D**

*Figure 145 – Freeze-Pump-Thaw method for degassing of buffer solutions.*

### **General Protocol for EPR Samples Preparation**

EPR samples were prepared, unless otherwise indicated, in strict anaerobic conditions in an MBraun UNIlab LMF Glovebox Workstation at  $O_2$  concentrations  $\leq 0.1$  ppm. All samples contained 10% glycerol as cryoprotectant.<sup>198</sup> Aliquots of freshly reconstituted maquettes (final concentration 250  $\mu\text{M}$  based on peptide) were reduced using a freshly prepared solution of DT (final concentration 500  $\mu\text{M}$ ). After 5 minutes of incubation with DT, samples were transferred into EPR tubes, capped with rubber septa, and flash frozen in liquid  $N_2$ . EPR samples were shipped to Manchester in a dry shipper and stored in liquid  $N_2$  or in a  $-86$  °C freezer prior to data collection.

## General Protocols for Chapter Three

### **General Protocol for Cyclic Voltammetry Cell Assembly**

[4Fe-4S]-cluster maquette samples were reconstituted in an MBraun UNIlab LMF Glovebox Workstation at O<sub>2</sub> concentrations  $\leq 0.1$  ppm as previously described. The samples were reconstituted to achieve a final volume of 4 mL and treated with NaCl to a final concentration of 100 mM. The cell was assembled as depicted in **Error! Reference source not found.** A 10 mL glass drum vial was charged with the [4Fe-4S]-cluster maquette solution supplemented with 100 mM NaCl, and the three electrodes were inserted in the vial. Particular care was paid to ensure the active components of the electrodes were fully submerged in the solution and not in contact with each other. Experiments were performed in the range between -1.2 and +0.6 V vs. SCE. More detailed parameters are expressed in the caption of each spectrum in the main text.

## General Protocols for Chapter Four

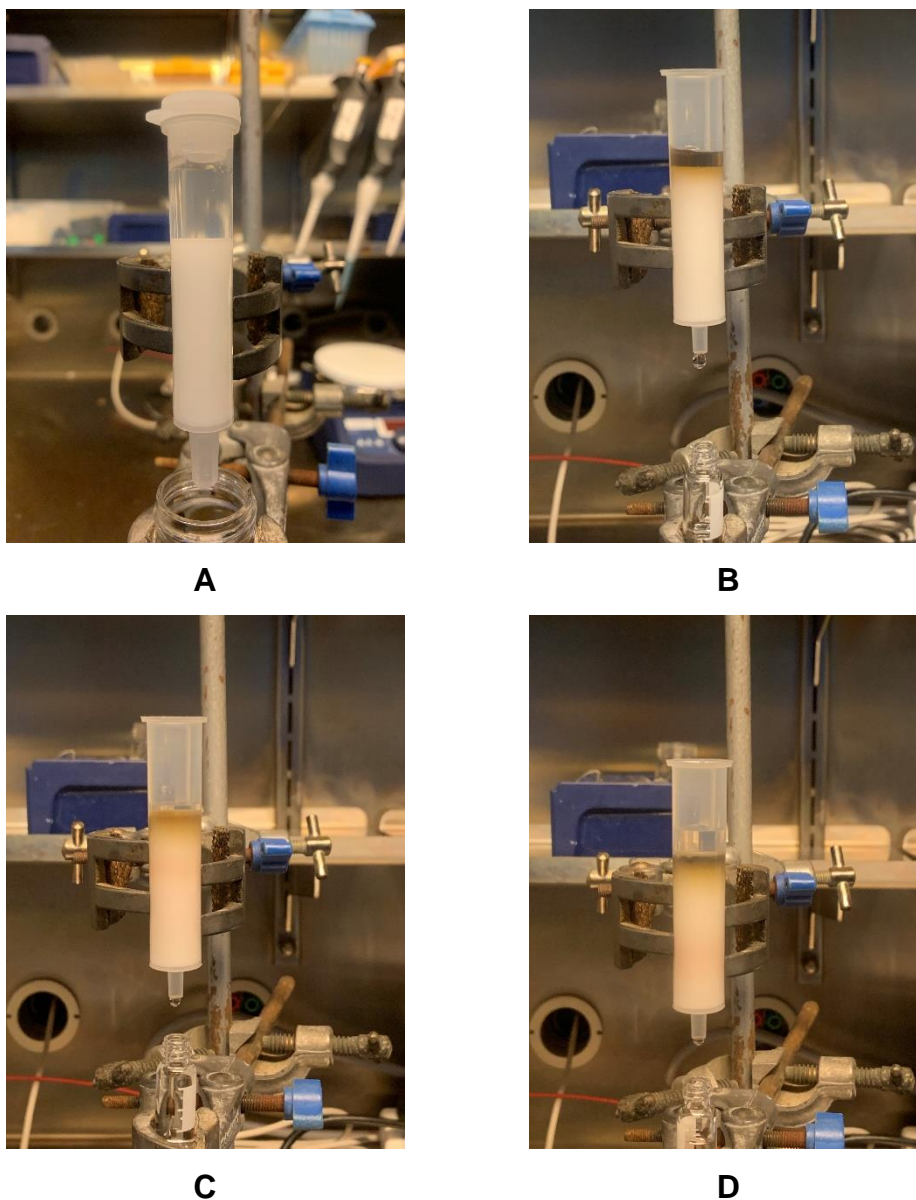
### **General Protocol for SET Process Catalysed by [4Fe-4S]-Cluster Maquette**

Samples were handled in a Whitley A85 TG Anaerobic Workstation and analysed on a Shimadzu UV-2600 Spectrophotometer. The hydrogenase used in these experiments was *Clostridium acetobutylicum* CaHydA. After reconstitution, [4Fe-4S]-cluster maquettes were diluted 4-folds (175  $\mu$ L of maquette in 525  $\mu$ L of buffer), transferred into a quartz cuvette and sealed with a rubber turnover stopper. Samples were purged with hydrogen for 5 minutes. Aliquots of CaHydA were prepared in an LC-MS vial sealed with a silicone screw cap. A gastight syringe (10  $\mu$ L) was used to transfer the stock solution of hydrogenase CaHydA to the cluster-maquette samples.

### **General Protocol for [4Fe-4S]-Cluster Maquette Purification by SEC Columns**

When indicated, reconstituted samples were purified using PD MidiTrap G-10 SEC columns in strict anaerobic conditions in an MBraun UNIlab LMF Glovebox Workstation at O<sub>2</sub> concentrations  $\leq 0.1$  ppm. The SEC column was equilibrated with degassed HEPES 50 mM, KCl 10 mM at pH = 8.0 for a total volume equivalent to 5

packed bed volumes (Figure 146, **A**). After equilibration, 1 mL of sample was loaded on top of the column (Figure 146, **B**) and allowed to completely enter the packed bed before addition of 0.7 mL of degassed buffer (Figure 146, **C**). When the sample completely entered the packed bed, additional buffer was added for elution (Figure 146, **D**). Fractions were collected in LC-MS vials or Eppendorf tubes.



*Figure 146 – SEC for [4Fe-4S]-cluster maquette purification. A) SEC column pre-loading of the sample. B) loading of the sample (brown solution) on top of the column. C) and D) progressive elution of the sample through the column by gravity.*

## General Protocols for Chapter Five

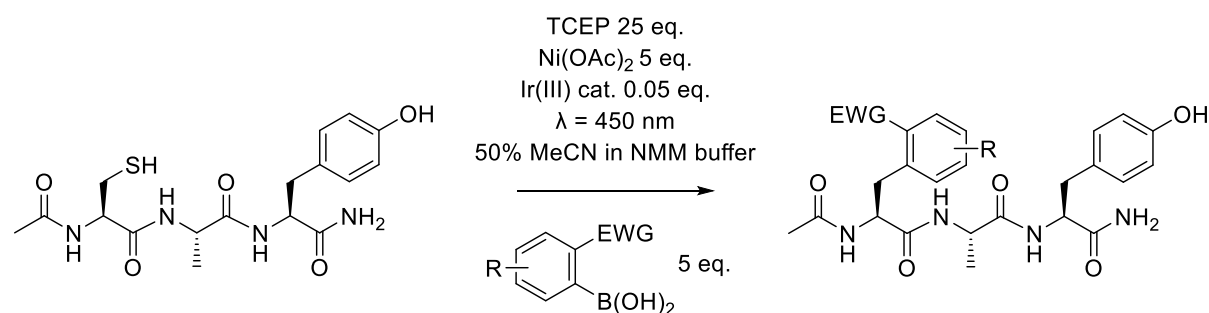
### Photochemical Reaction Setup

iNextStation blue LED strip lights wrapped inside a Pyrex dish were used as the light source (Figure 147). 8 mL screw cap vials were used as the reaction vessels. Reaction mixtures were prepared inside a standard fume hood, and transferred into the LED dish after the pH was adjusted.



Figure 147 – Photochemical reaction setup.

### Optimised General Protocol for Cysteine Desulfurisation and $\beta$ -Arylation



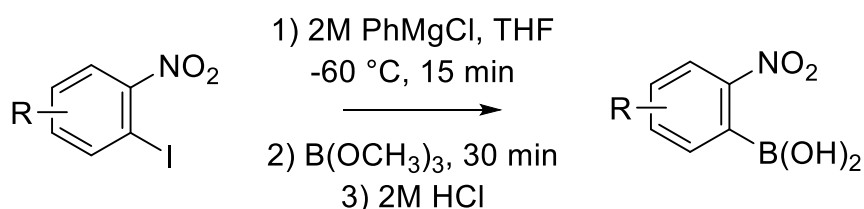
Scheme 35 – Schematic synthesis for the Cys desulfurisation and  $\beta$ -arylation reaction of small peptides.

To Ac-CAY-NH<sub>2</sub> (5 mg, 12.6  $\mu\text{mol}$ ) dissolved in MeCN : 0.01 M NMM in H<sub>2</sub>O (50 : 50, 2.52 mL), was added a solution of TCEP (25 eq., 0.5 M stock solution in H<sub>2</sub>O) and (Ir[dF(CF<sub>3</sub>)ppy]<sub>2</sub>(dtbpy))PF<sub>6</sub> (0.05 eq., 1 mM stock solution in MeCN). The reaction mixture was diluted to a final peptide concentration of 0.5 mM with MeCN : 0.01 M NMM in H<sub>2</sub>O (50 : 50, 1.13 mL) and the pH was adjusted to 7.8 – 8.2. Quickly, Ni(OAc)<sub>2</sub>



(5 eq., 0.5 M stock solution in H<sub>2</sub>O) followed by phenylboronic acid (5 eq.) were added. The pH of the reaction mixture was adjusted to 7.8 – 8.2 and the reaction vessel was placed into blue LEDs and left stirring at r.t. for 1 hour. After this time, the crude reaction mixture was filtrated and purified by preparative RP-HPLC

### General Protocol for the Synthesis of Nitro-Phenylboronic Acids

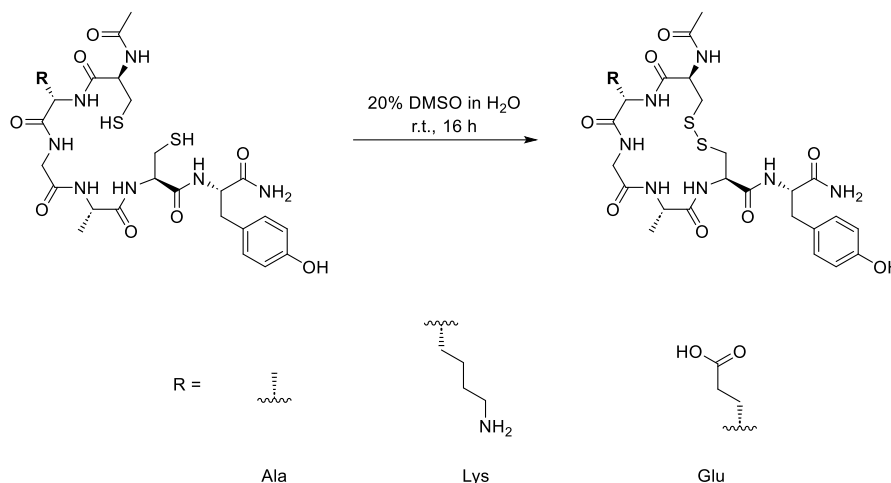


*Scheme 36 – Schematic synthesis of functionalised ortho-nitrophenyl boronic acids.*

Substituted nitrophenylboronic acids were synthesised following an adapted literature procedure.<sup>169</sup> A dry 25 mL two-neck round bottom flask equipped with a magnetic stirrer was charged with the corresponding aryl iodide (4 mmol). The round bottom flask was sealed with a rubber turnover stopper and connected to a Schlenk line. Dry THF (6 mL) was added, and the resulting mixture was cooled to -90 °C in an acetone/liquid N<sub>2</sub> bath. PhMgCl (2 M in THF, 2.2 mL, 4.4 mmol) was added dropwise and the reaction mixture was left stirring at -90 °C for 15 min. Trimethyl borate (4.8 mmol, 536 µL) was added dropwise and the reaction mixture was left stirring at -90 °C for 30 min. After 30 min the round bottom flask was removed from the acetone/liquid N<sub>2</sub> bath, and the reaction mixture quenched with 2 M aq. HCl (4 mL). The reaction was extracted with Et<sub>2</sub>O (3 x 20 mL), the organic layers were combined, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered, and concentrated under reduced pressure. The crude residue was recrystallised from ethyl acetate.

## General Protocols for Chapter Six

### General Protocol for the Inter-Molecular Formation of Disulphide Bond in Peptides



*Scheme 37 – Schematic synthesis for the intra-molecular disulphide bond formation on peptides **29a** (Ac-CAGACY-NH<sub>2</sub>), **30a** (Ac-CKGACY-NH<sub>2</sub>), and **31a** (Ac-CEGACY-NH<sub>2</sub>).*

Oxidised peptides (**29b**, **30b**, and **31b**) were synthesised following an adapted literature procedure.<sup>149</sup> After cleavage and work-up, the crude peptide (**29a**, **30a**, and **31a**) was dissolved in 20% DMSO/water (final concentration 1  $\mu$ M of peptide) and stirred at room temperature for 16 h. The progress of the reaction was monitored by analytical HPLC and HRMS. After completion of the reaction, the solution was directly subjected to purification by preparative HPLC. Cyclised peptides were purified using a linear gradient 5-40% Solvent B over 30 minutes.

### General Protocol for Disulphide Bond Oxidation by Hypochlorous Acid

To a 50  $\mu$ M solution of cyclised peptide (**29b**, **30b**, or **31b**) in phosphate buffer (phosphate buffer: 0.1 M, pH = 7.3) in a LC-MS vial, a solution of HOCl in phosphate buffer (0.1 M, pH = 7.3) was added (final concentration 50  $\mu$ M of peptide and 1 mM of HOCl). The mixture was quickly shaken, and the progress of the reaction was monitored by analytical HPLC using a linear gradient 5-40% Solvent B over 5 minutes. HOCl was generated *in situ* by diluting NaOCl in phosphate buffer (0.1 M, pH = 7.3).

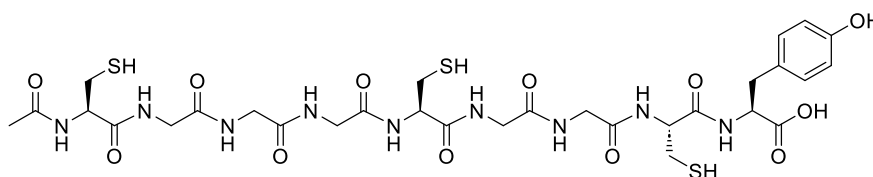
## General Protocol for Appendix 2

### **General Protocol for the Reduction from Sulfone to Sulfinic acid**

Synthetic peptides **46** and **47** after cleavage were lyophilised to a dry solid before being re-dissolved in dH<sub>2</sub>O : MeOH (1:1, 0.5 mL) and stirred at r.t.. A solution of NaBH<sub>4</sub> in dH<sub>2</sub>O (20 eq. with respect to peptide in max. 0.1 mL of solvent) were added and the mixture stirred at r.t. in an Eppendorf vial (2 mL) and monitored by HPLC until complete conversion was observed.

## 8.3 PEPTIDE SYNTHESIS

### Synthesis of Ac-CGGGCGGCGY-NH<sub>2</sub> (MQ1)



Peptide **MQ1** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol). The crude peptide was purified by preparative RP-HPLC (10-60% B over 30 min) and lyophilised to produce the desired peptide (**MQ1**, 175 mg, 0.114 mmol, 57% yield).

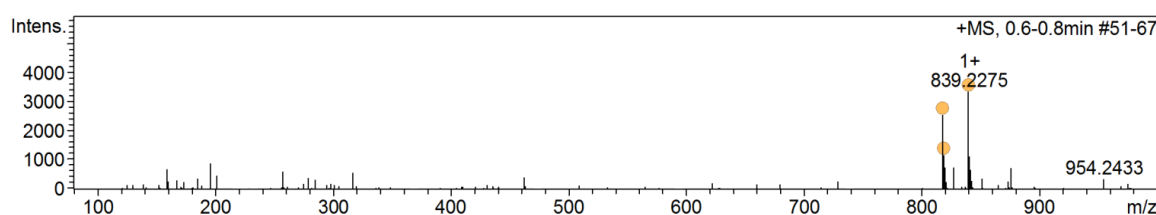
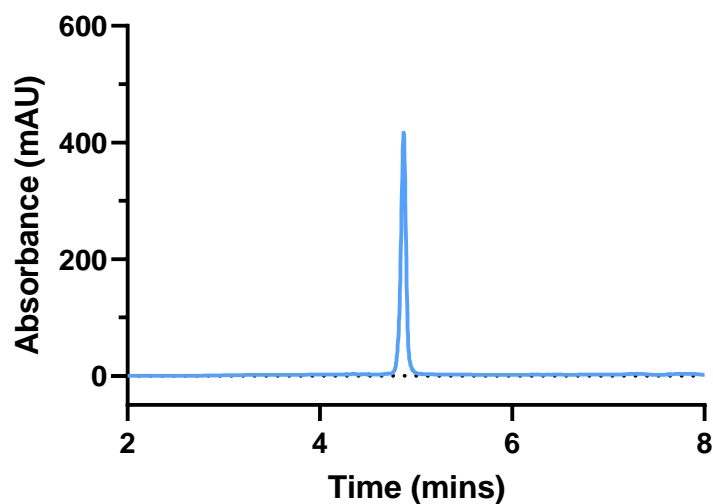
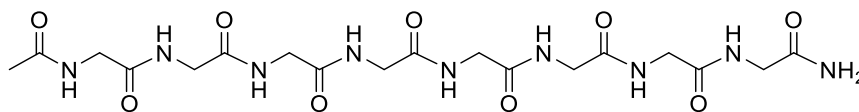


Figure 148 – Analytical HPLC trace and ESI MS of pure Ac-CGGGCGGCGY-NH<sub>2</sub> (**MQ1**). Analytical gradient 10-60% B over 30 min, 210 nm. Calculated Mass: 840.21 [M+Na]<sup>+</sup>, observed mass: 839.23 [M+Na]<sup>+</sup>.

## Synthesis of Ac-GGGGGGGGG-NH<sub>2</sub> (MQ2)



Peptide **MQ2** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol). The crude peptide was purified by preparative RP-HPLC (10-60% B over 30 min) and lyophilised to produce the desired peptide (**MQ2**, 49 mg, 0.086 mmol, 43% yield).

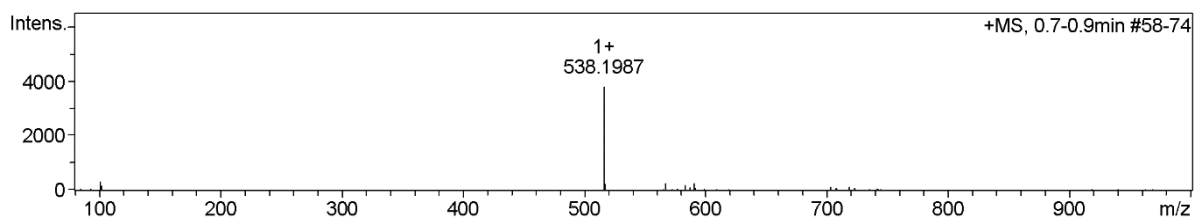
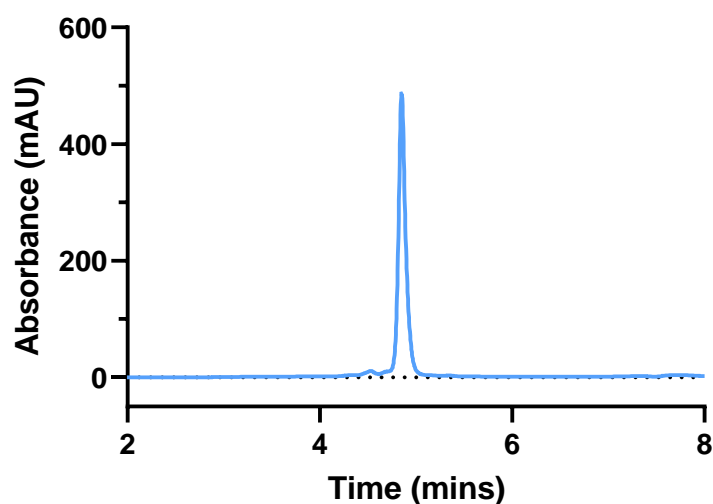
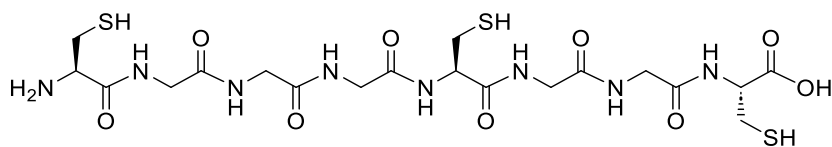


Figure 149 – Analytical HPLC trace and ESI MS of pure Ac-GGGGGGGGG-NH<sub>2</sub> (**MQ2**). Analytical gradient 5-80% B over 30 min, 210 nm. Calculated Mass: 538.20 [M+Na]<sup>+</sup>, observed mass: 538.20 [M+Na]<sup>+</sup>.

## Synthesis of H<sub>2</sub>N-CGGGCGGC-OH (MQ3)



Peptide **MQ3** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol). The crude peptide was purified by preparative RP-HPLC (10-60% B over 30 min) and lyophilised to produce the desired peptide (**MQ3**, 78 mg, 0.128 mmol, 64% yield).

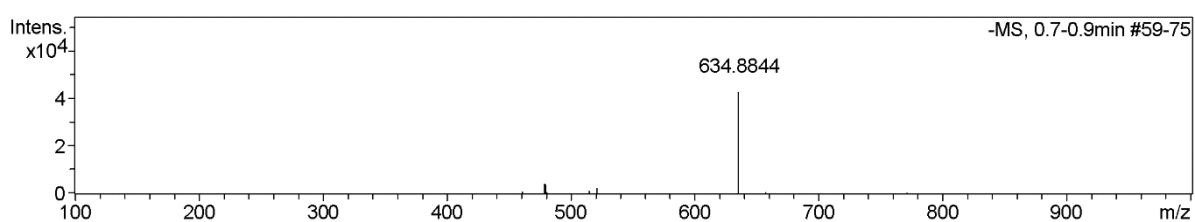
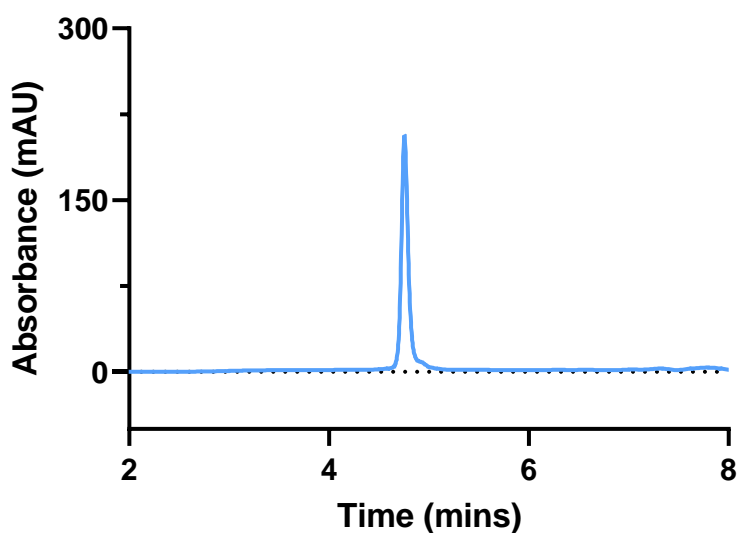
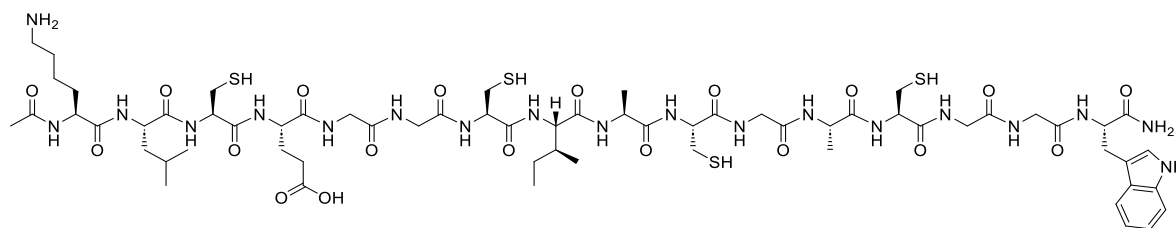


Figure 150 – Analytical HPLC trace and ESI MS of pure H<sub>2</sub>N-CGGGCGGC-OH (**MQ3**). Analytical gradient 10-80% B over 30 min, 210 nm. Calculated Mass: 635.14 [M+Na]<sup>+</sup>, observed mass: 634.88 [M+Na]<sup>+</sup>.

## Synthesis of Ac-KLCEGGCIACGACGGW-NH<sub>2</sub> (MQ4)



Peptide **MQ4** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol). The crude peptide was purified by preparative RP-HPLC (10-60% B over 30 min) and lyophilised to produce the desired peptide (**MQ4**, 249 mg, 0.162 mmol, 81% yield).

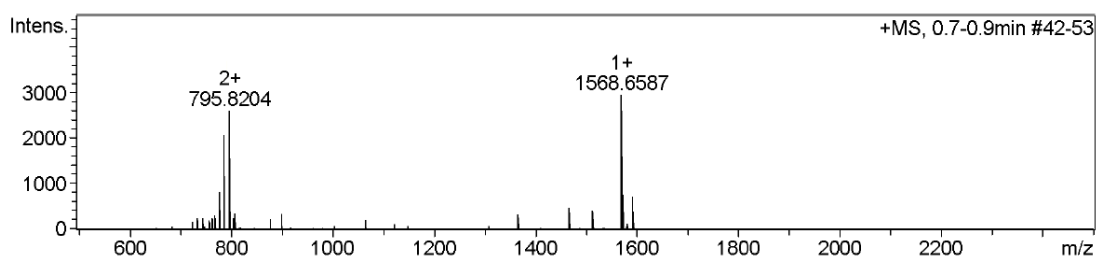
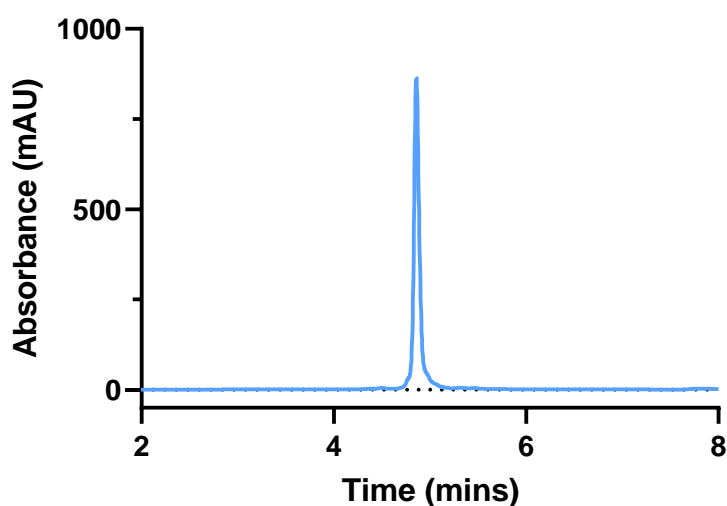
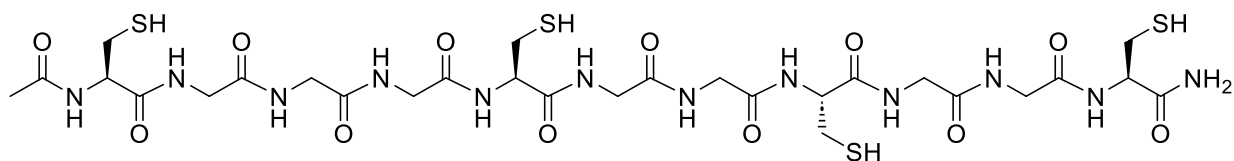


Figure 151 – Analytical HPLC trace and ESI MS of pure Ac-KLCEGGCIACGACGGW-NH<sub>2</sub> (**MQ4**). Analytical gradient 10-60% B over 5 min, 210 nm. Calculated Mass: 1568.62 [M+H]<sup>+</sup>, observed mass: 1568.66 [M+H]<sup>+</sup>.

## Synthesis of Ac-CGGGCGGCGGC-NH<sub>2</sub> (MQ5)



Peptide **MQ5** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol). The crude peptide was purified by preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired peptide (**MQ5**, 106 mg, 0.120 mmol, 61% yield).

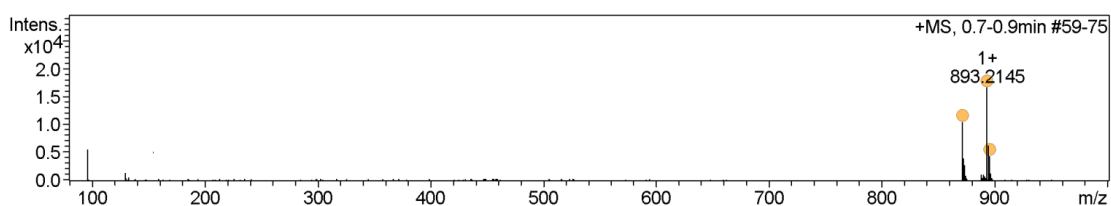
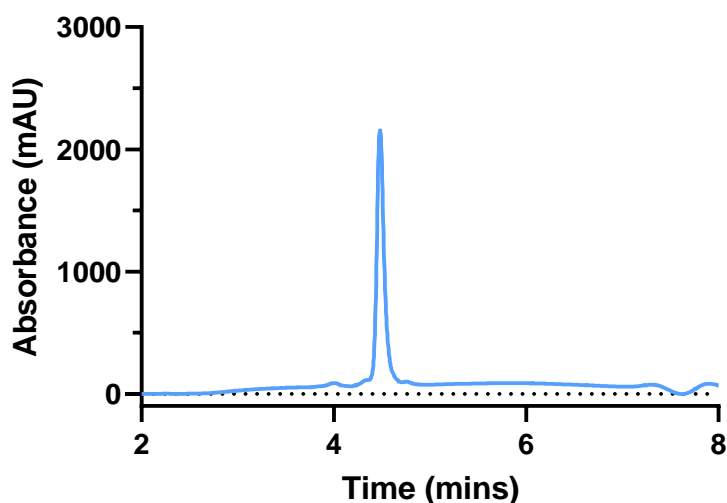
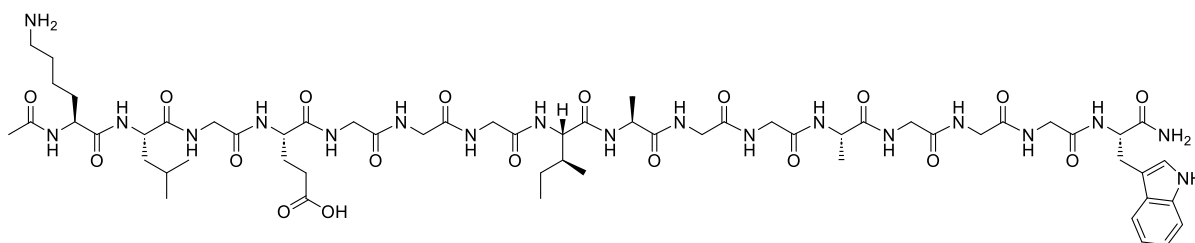


Figure 152 – Analytical HPLC trace and ESI MS of pure Ac-CGGGCGGCGGC-NH<sub>2</sub> (**MQ5**). Analytical gradient 2-40% B over 5 min, 210 nm. Calculated Mass: 894.20 [M+Na]<sup>+</sup>, observed mass: 893.21 [M+Na]<sup>+</sup>.



## Synthesis of Ac-KLGEGGGIAGGAGGGW-NH<sub>2</sub> (MQ6)



Peptide **MQ6** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol). The crude peptide was purified by preparative RP-HPLC (10-100% B over 30 min) and lyophilised to produce the desired peptide (**MQ6**, 150 mg, 0.110 mmol, 54% yield).

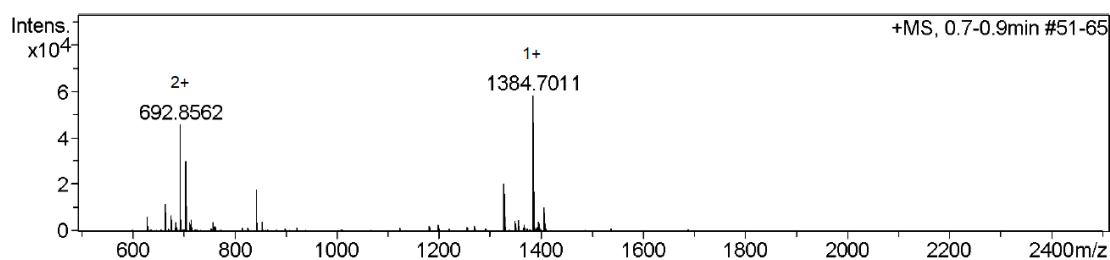
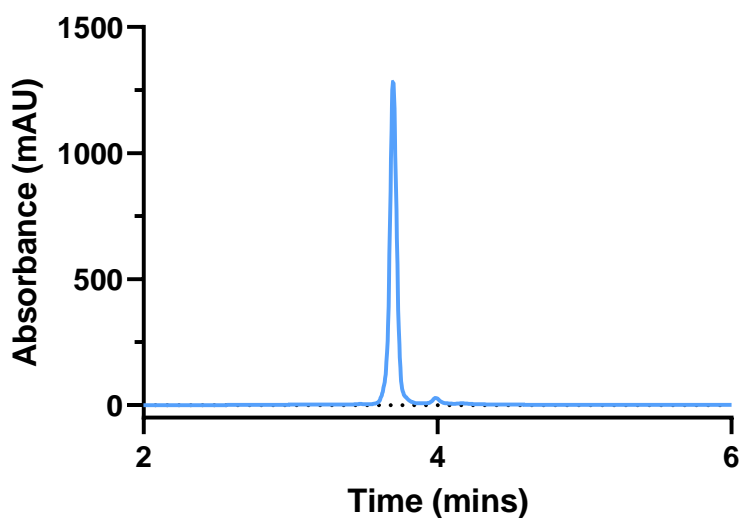
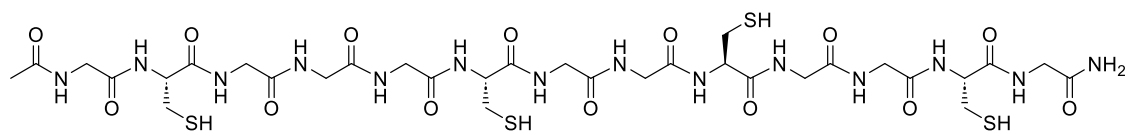


Figure 153 – Analytical HPLC trace and ESI MS of pure Ac-KLGEGGGIAGGAGGGW-NH<sub>2</sub> (**MQ6**). Analytical gradient 10-100% B over 5 min, 280 nm. Calculated Mass: 1384.67 [M+H]<sup>+</sup>, observed mass: 1384.70 [M+H]<sup>+</sup>.

## Synthesis of Ac-GCGGGCGGCGGCG-NH<sub>2</sub> (MQ7)



Peptide **MQ7** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol). The crude peptide was purified by preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired peptide (**MQ7**, 85 mg, 0.086 mmol, 43% yield).

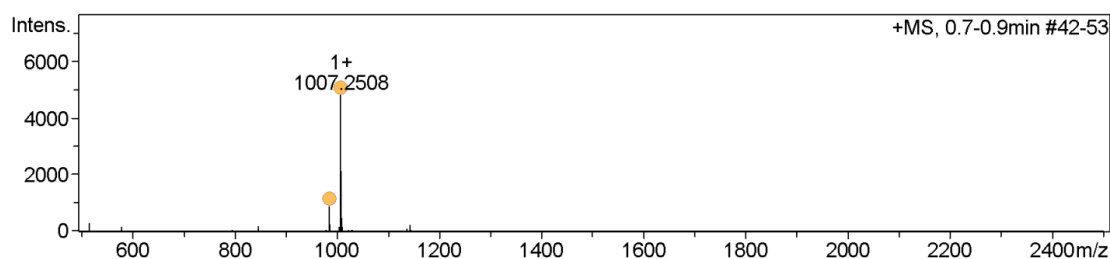
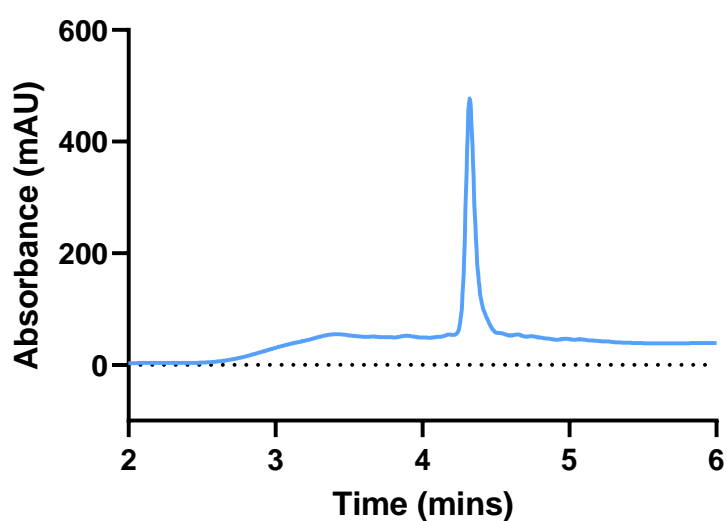
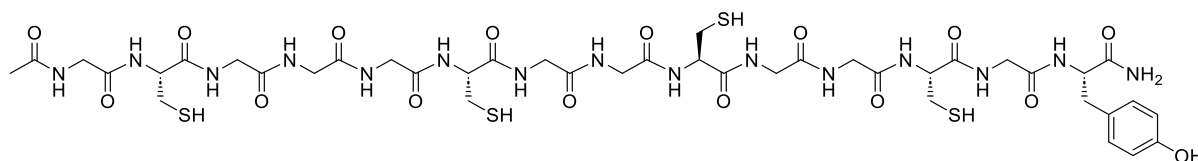
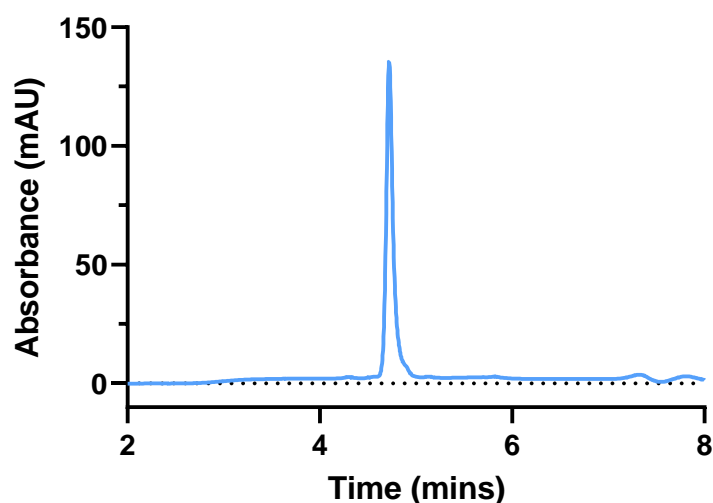


Figure 154 – Analytical HPLC trace and ESI MS of pure Ac-GCGGGCGGCGGCG-NH<sub>2</sub> (**MQ7**). Analytical gradient 2-40% B over 5 min, 210 nm. Calculated Mass: 1008.24 [M+Na]<sup>+</sup>, observed mass: 1007.25 [M+Na]<sup>+</sup>.

## Synthesis of Ac-GCGGGCGGCGGCGY-NH<sub>2</sub> (MQ8)



Peptide **MQ8** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol). The crude peptide was purified by preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired peptide (**MQ8**, 165 mg, 0.140 mmol, 72% yield).



+MS, 0.7-0.9min #41-52

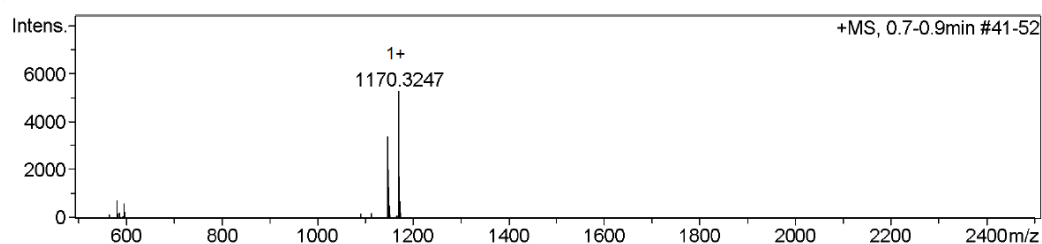
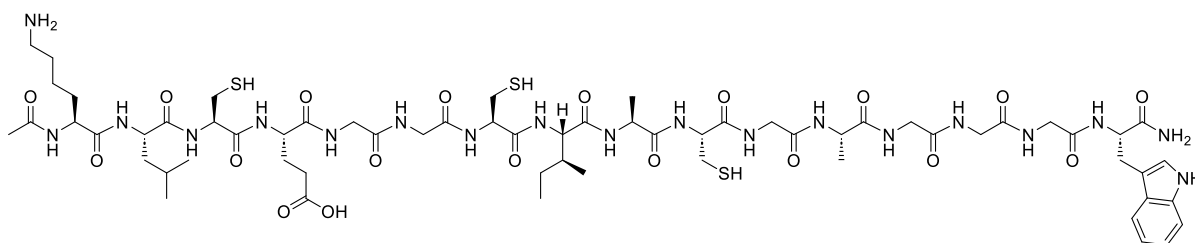


Figure 155 – Analytical HPLC trace and ESI MS of pure Ac-GCGGGCGGCGGCGY-NH<sub>2</sub> (**MQ8**). Analytical gradient 2-40% B over 5 min, 280 nm. Calculated Mass: 1171.30 [M+Na]<sup>+</sup>, observed mass: 1170.32 [M+Na]<sup>+</sup>.

## Synthesis of Ac-KLCEGGCIACGAGGGW-NH<sub>2</sub> (MQ9)



Peptide **MQ9** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol). The crude peptide was purified by preparative RP-HPLC (20-90% B over 30 h) and lyophilised to produce the desired peptide (**MQ9**, 158 mg, 1.04 mmol, 52% yield).

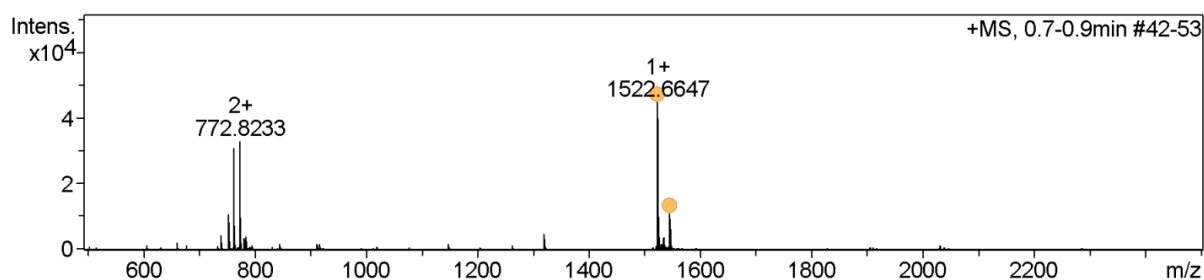
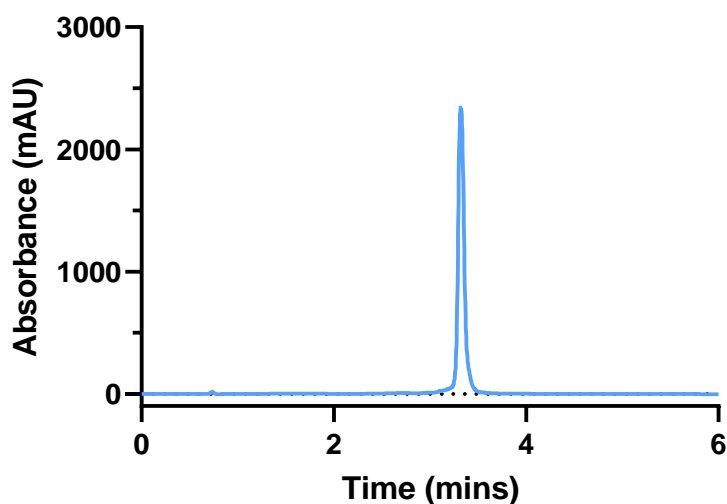
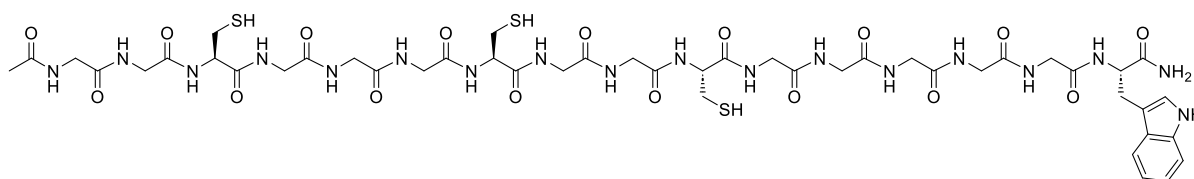
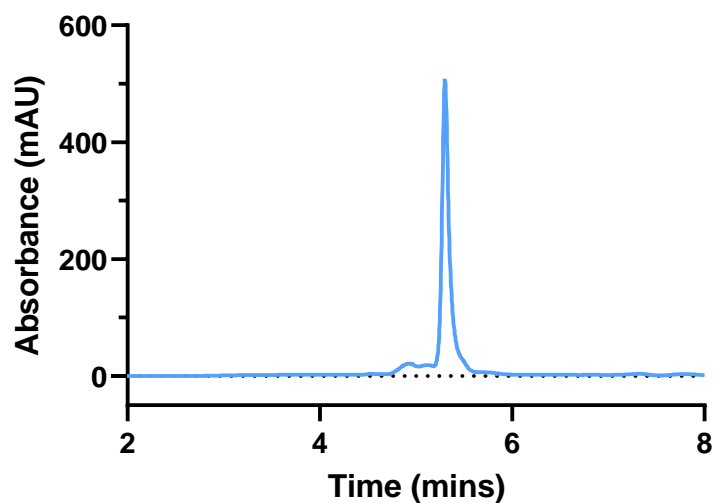


Figure 156 – Analytical HPLC trace and ESI MS of pure Ac- KLCEGGCIACGAGGGW-NH<sub>2</sub> (**MQ9**). Analytical gradient 20-90% B over 5 min, 280 nm. Calculated Mass: 1521.65 [M+H]<sup>+</sup>, observed mass: 1522.66 [M+H]<sup>+</sup>.

## Synthesis of Ac-GGCGGGCGGGGGW-NH<sub>2</sub> (MQ10)



Peptide **MQ10** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol). The crude peptide was purified by preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired peptide (**MQ10** 117 mg, 0.94 mmol, 47% yield).



+MS, 0.7-0.9min #41-52

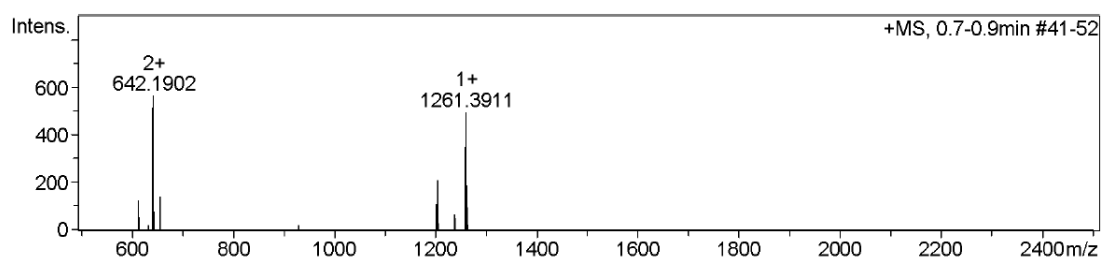
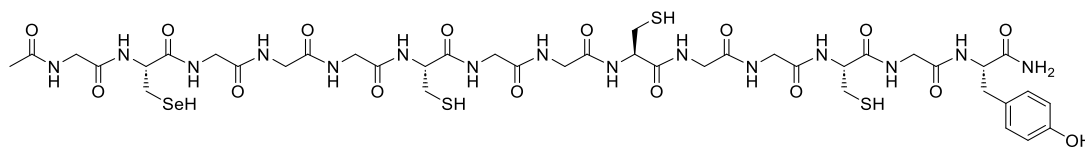


Figure 157 – Analytical HPLC trace and ESI MS of pure Ac-GGCGGGCGGGGGW-NH<sub>2</sub> (**MQ10**). Analytical gradient 2-40% B over 5 min, 280 nm. Calculated Mass: 1261.39 [M+Na]<sup>+</sup>, observed mass: 1261.39 [M+Na]<sup>+</sup>.

## Synthesis of Ac-GUGGGCGGCGGCGY-NH<sub>2</sub> (MQ11)



Peptide **MQ11** was synthesised using general automated synthesiser protocols with microwave assistance on rink amide resin (0.2 mmol) up to the 12<sup>th</sup> amino acid (Gly). The Sec building block (**45**) was synthesised following the procedure described earlier and manually coupled following general manual coupling protocol. The crude peptide was purified by semi-preparative RP-HPLC (2-95% B over 1 h) and lyophilised to produce the desired peptide (**MQ11** 15 mg, 0.013 mmol, 38% yield).

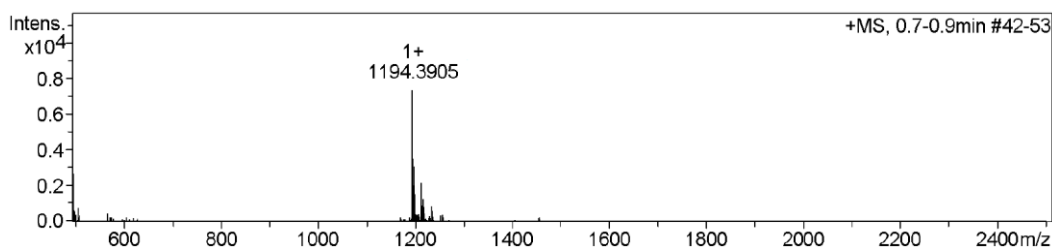
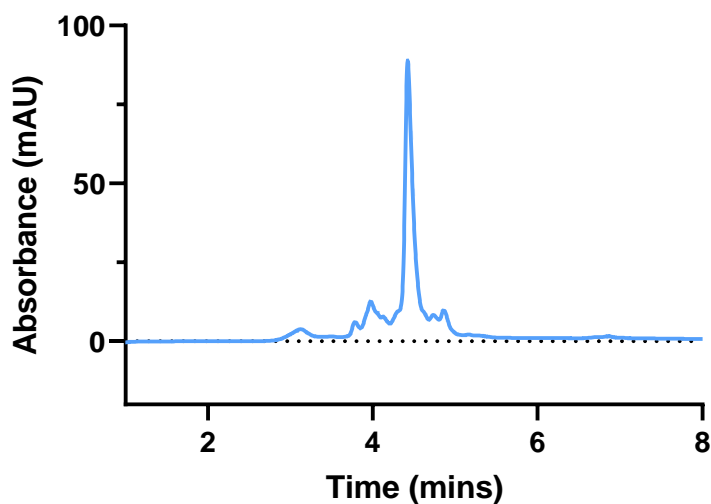
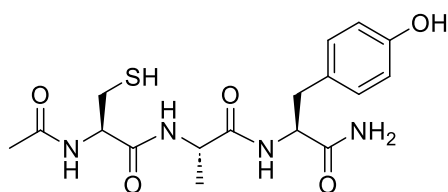


Figure 158 – Analytical HPLC trace and ESI MS of 90% pure Ac-GUGGGCGGCGGCGY-NH<sub>2</sub> (**MQ11**). Analytical gradient 2-40% B over 5 min, 280 nm. Calculated Mass: 1195.27 [M+H]<sup>+</sup>, observed mass: 1194.39 [M+H]<sup>+</sup>.

## Synthesis of Ac-CAY-NH<sub>2</sub> (1)



Peptide **1** was synthesised using general automated synthesiser protocols with microwave assistance. The crude peptide was purified by preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired peptide (**1** 60.6 mg, 0.15 mmol, 76% yield); <sup>1</sup>H NMR (400 MHz, DMSO) δ 9.15 (s, 1H, OH), 8.21 (d, J = 6.9 Hz, 1H, sec. amide), 8.11 (d, J = 7.8 Hz, 1H, sec. amide), 7.69 (d, J = 8.2 Hz, 1H, sec. amide), 7.26 (s, 1H, sec. amide), 7.06 (s, 1H, sec. amide), 6.98 (d, 2H, Ar-H), 6.63 (d, J = 8.2 Hz, 2H, Ar-H), 4.39 – 4.25 (m, 2H, CH<sub>2</sub>), 4.16 (p, J = 7.0 Hz, 1H, CH), 2.88 (dd, J = 13.9, 5.1 Hz, 1H, CH), 2.79 – 2.58 (m, 3H, CH<sub>3</sub>), 2.34 (t, J = 8.4 Hz, 1H, CH), 1.87 (s, 3H, CH<sub>3</sub>), 1.24 (s, 1H, SH), 1.15 (d, J = 7.1 Hz, 3H, CH<sub>3</sub>); <sup>13</sup>C NMR (101 MHz, DMSO) δ 172.78, 171.66, 169.92, 169.57, 155.77, 130.09, 127.73, 114.85, 55.02, 53.94, 48.78, 36.63, 26.14, 22.50, 17.75.

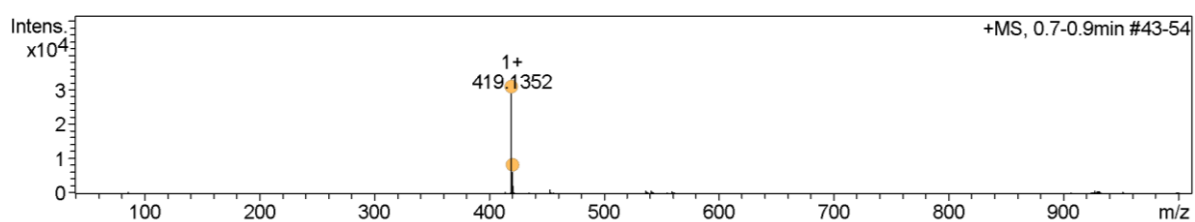
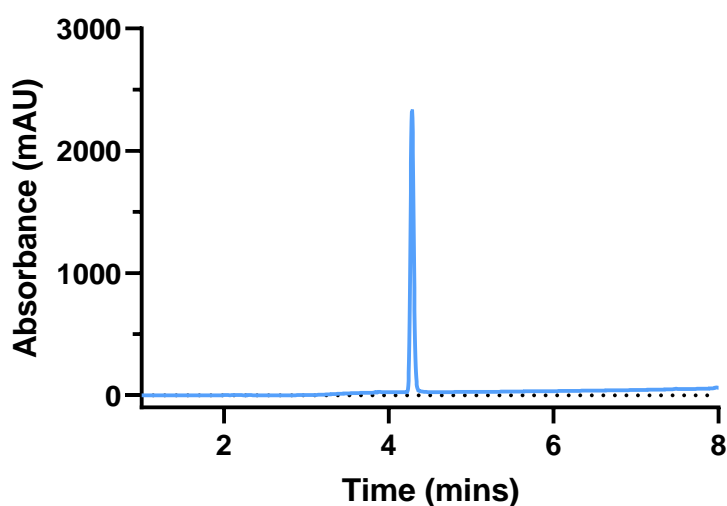
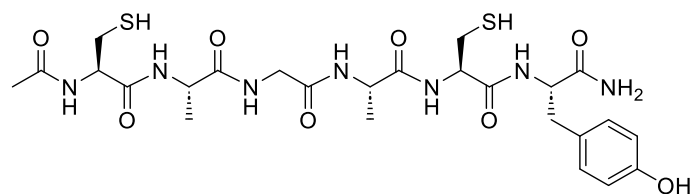


Figure 159 – Analytical HPLC trace and ESI MS of pure Ac-CAY-NH<sub>2</sub> (**1**). Analytical gradient 5-60% B over 5 min, 280 nm. Calculated Mass: 419.14 [M+Na]<sup>+</sup>, observed mass: 419.14 [M+Na]<sup>+</sup>.

## Synthesis of reduced Ac-CAGACY-NH<sub>2</sub> (**29a**)



Peptide **29a** was synthesised using general automated synthesiser protocols with microwave assistance. The crude peptide was purified by preparative RP-HPLC (5-40% B over 30 min) and lyophilised to produce the desired peptide (**29a** 103.1 mg, 0.16 mmol, 80% yield).

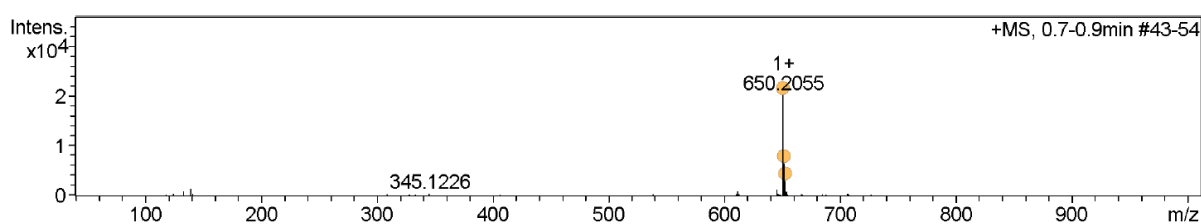
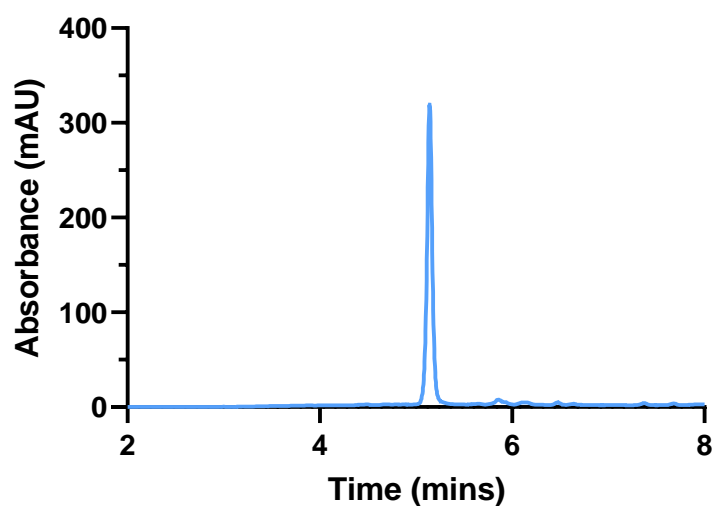
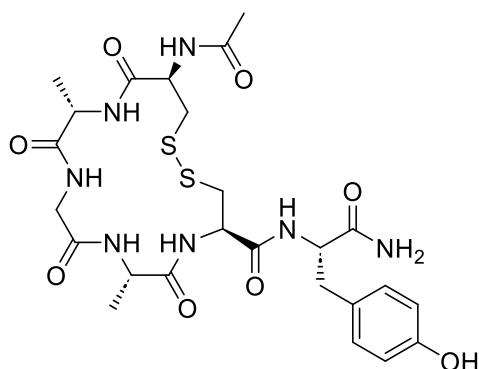


Figure 160 – Analytical HPLC trace and ESI MS of pure Ac-CAGACY-NH<sub>2</sub> (**29a**). Analytical gradient 5-40% B over 5 min, 210 nm. Calculated Mass: 651.19 [M+Na]<sup>+</sup>, observed mass: 650.21 [M+Na]<sup>+</sup>.



## Synthesis of oxidised Ac-CAGACY-NH<sub>2</sub> (**29b**)



Peptide **29b** was synthesised from **29a** following the protocol detailed in section *General Protocols for Chapter Six* (**29b** 75.2 mg, 0.12 mmol, 75% yield).

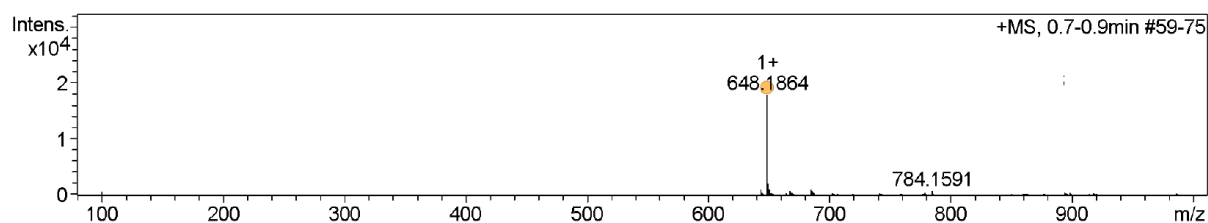
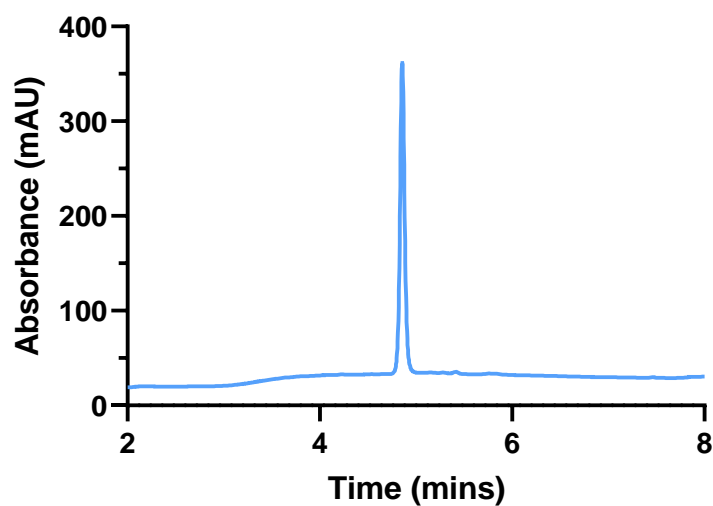
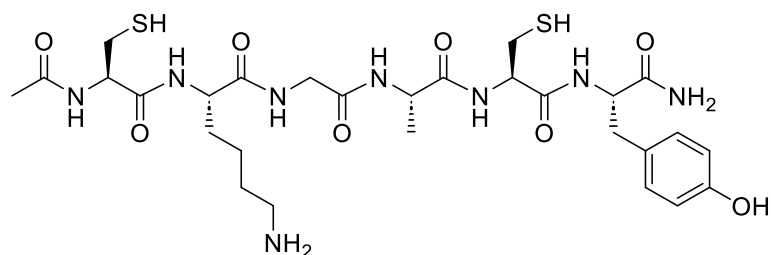


Figure 161 – Analytical HPLC trace and ESI MS of pure Ac-CAGACY-NH<sub>2</sub> (**29b**). Analytical gradient 5-40% B over 5 min, 210 nm. Calculated Mass: 649.17 [M+Na]<sup>+</sup>, observed mass: 648.19 [M+Na]<sup>+</sup>.

## Synthesis of reduced Ac-CKGACY-NH<sub>2</sub> (**30a**)



Peptide **30a** was synthesised using general automated synthesiser protocol with microwave assistance. The crude peptide was purified by preparative RP-HPLC (5-40% B over 30 min) and lyophilised to produce the desired peptide (**30a** 97.2 mg, 0.14 mmol, 89% yield).

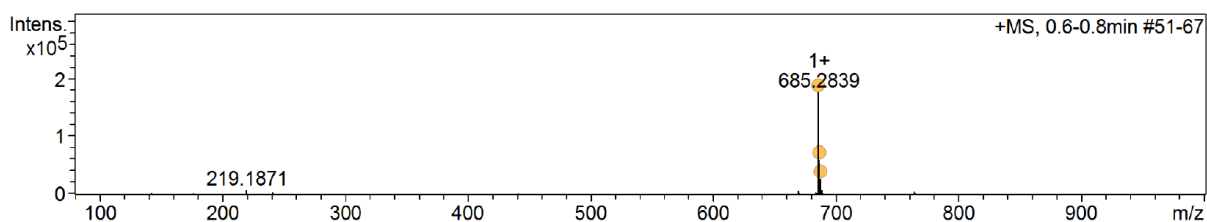
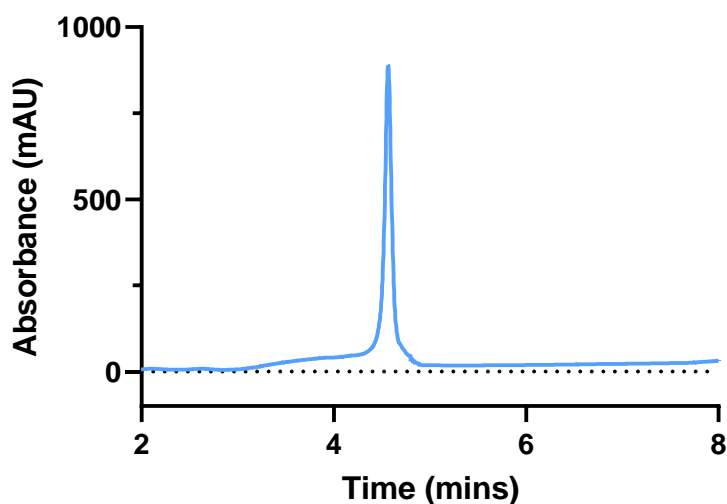
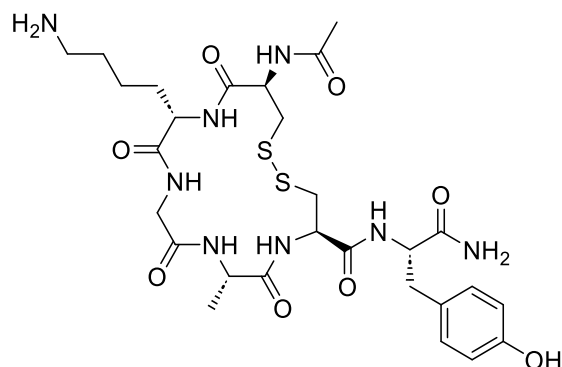


Figure 162 – Analytical HPLC trace and ESI MS of pure Ac-CKGACY-NH<sub>2</sub> (**30a**). Analytical gradient 5-40% B over 5 min, 210 nm. Calculated Mass: 685.26 [M+H]<sup>+</sup>, observed mass: 685.28 [M+H]<sup>+</sup>.

## Synthesis of oxidised Ac-CKGACY-NH<sub>2</sub> (**30b**)



Peptide **30b** was synthesised from **30a** following the protocol detailed in section *General Protocols for Chapter Six* (**30b** 71.6 mg, 0.10 mmol, 75% yield).

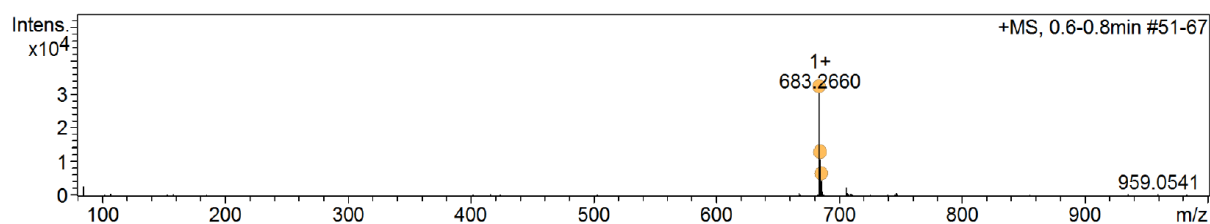
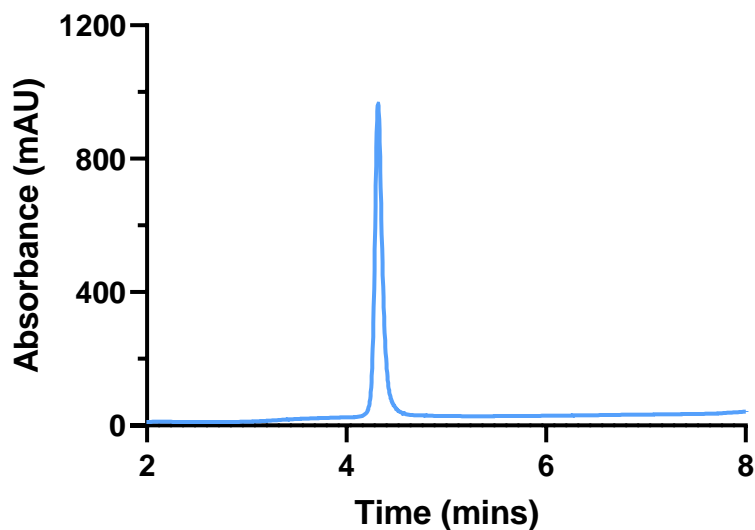
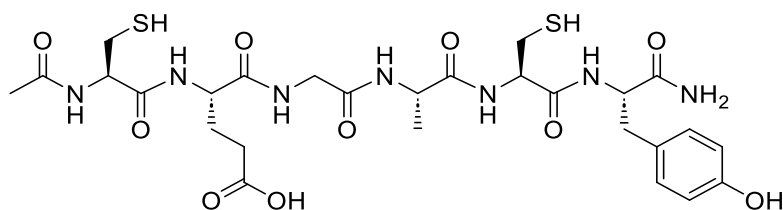


Figure 163 – Analytical HPLC trace and ESI MS of pure Ac-CKGACY-NH<sub>2</sub> (**30b**). Analytical gradient 5-40% B over 5 min, 210 nm. Calculated Mass: 683.24 [M+H]<sup>+</sup>, observed mass: 683.27 [M+H]<sup>+</sup>.

## Synthesis of reduced Ac-CEGACY-NH<sub>2</sub> (**31a**)



Peptide **31a** was synthesised using general automated synthesiser protocols with microwave assistance. The crude peptide was purified by preparative RP-HPLC (5-40% B over 30 min) and lyophilised to produce the desired peptide (**31a** 85.1 mg, 0.12 mmol, 62% yield).

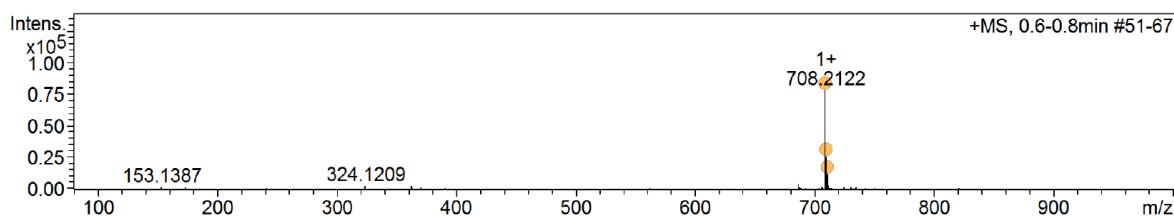
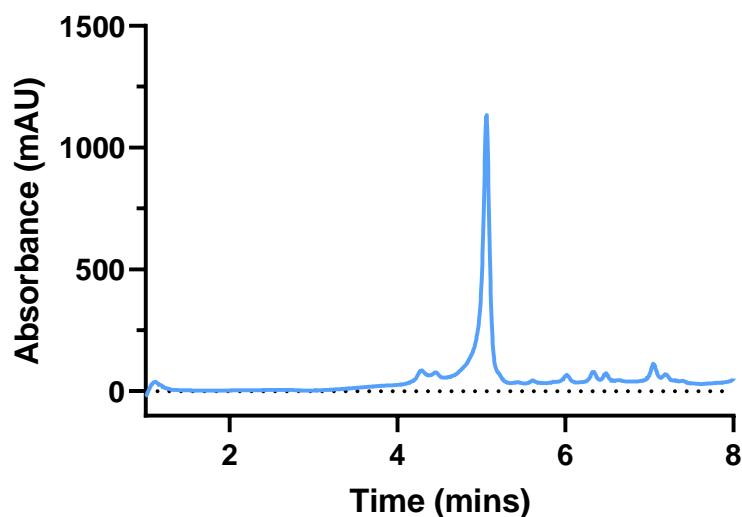
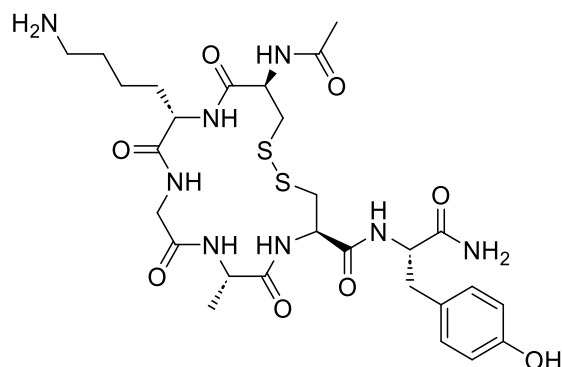


Figure 164 – Analytical HPLC trace and ESI MS of pure Ac-CEGACY-NH<sub>2</sub> (**31a**). Analytical gradient 5-40% B over 5 min, 210 nm. Calculated Mass: 709.19 [M+Na]<sup>+</sup>, observed mass: 708.21 [M+Na]<sup>+</sup>.

## Synthesis of oxidised Ac-CEGACY-NH<sub>2</sub> (**31b**)



Peptide **31b** was synthesised from **31a** following the protocol detailed in section *General Protocols for Chapter Six* (**31b** 77.3 mg, 0.11 mmol, 94% yield).

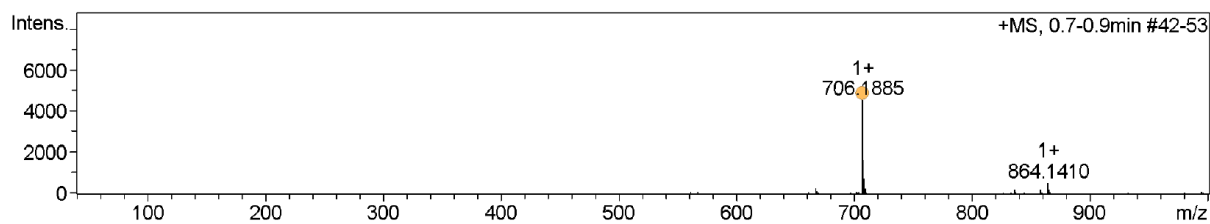
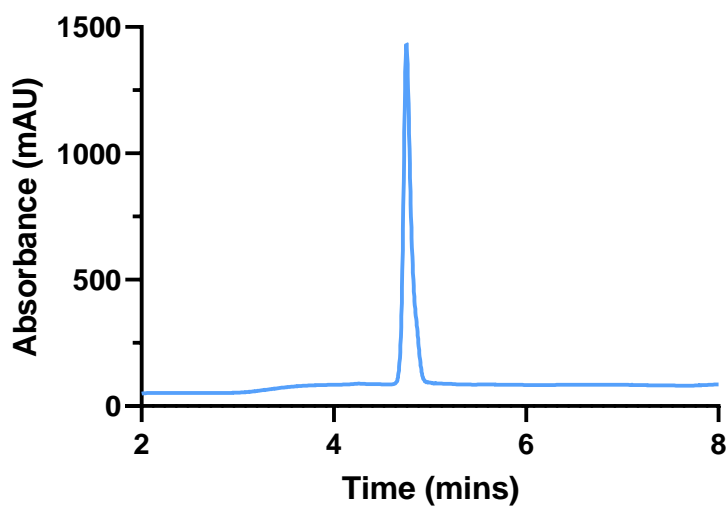
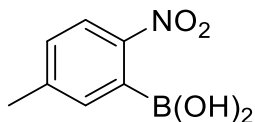


Figure 165 – Analytical HPLC trace and ESI MS of pure Ac-CEGACY-NH<sub>2</sub> (**31b**). Analytical gradient 5-40% B over 5 min, 210 nm. Calculated Mass: 706.23 [M+Na]<sup>+</sup>, observed mass: 706.19 [M+Na]<sup>+</sup>.

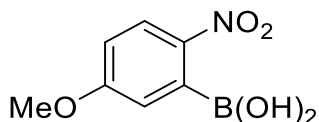
## 8.4 COMPOUNDS SYNTHESIS

### Synthesis of (5-methyl-2-nitrophenyl)boronic acid (10)



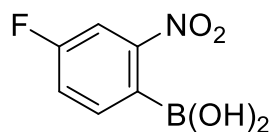
Synthesised using general procedure outlined in *General Protocol for the Synthesis of Ortho-Nitro-Phenylboronic Acids* in *General Protocols for Chapter Five* (**10** 580 mg, 3.21 mmol, 80% yield). White crystalline solid: mp 257-260 °C, <sup>1</sup>H NMR (400 MHz, DMSO) δ 8.04 (d, J = 8.4 Hz, 1H, Ar-H), 7.41 – 7.30 (m, 2H, Ar-H), 2.41 (s, 3H, CH<sub>3</sub>); <sup>13</sup>C NMR (101 MHz, DMSO) δ 148.24 (C), 145.37 (C), 134.53 (C), 133.09 (CH), 129.90 (CH), 123.14 (CH), 21.41 (CH<sub>3</sub>); HRMS calc. for C<sub>7</sub>H<sub>9</sub>BNO<sub>4</sub>: 180.95 [M-H]<sup>-</sup>, found: 180.97 [M-H]<sup>-</sup>.

### Synthesis of (5-methoxy-2-nitrophenyl)boronic acid (12)



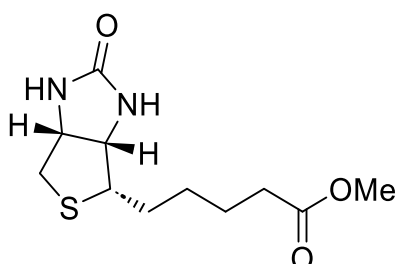
Synthesised using general procedure outlined in *General Protocol for the Synthesis of Ortho-Nitro-Phenylboronic Acids* in *General Protocols for Chapter Five* (**12** 678 mg, 3.44 mmol, 86% yield). White solid, mp 155-158 °C, <sup>1</sup>H NMR (400 MHz, DMSO) δ 7.41 (d, J = 9.2 Hz, 1H, Ar-H), 6.27 (dd, J = 9.2, 2.8 Hz, 1H, Ar-H), 6.17 (d, J = 2.8 Hz, 1H, Ar-H), 3.12 (s, 3H, CH<sub>3</sub>); <sup>13</sup>C NMR (101 MHz, MeOD) δ 164.86 (C), 143.27 (C), 132.66 (C), 125.22 (CH), 115.68 (CH), 114.32 (CH), 55.26 (CH<sub>3</sub>).

### Synthesis of (4-fluoro-2-nitrophenyl)boronic acid (**14**)



Synthesised using general procedure outlined in *General Protocol for the Synthesis of Ortho-Nitro-Phenylboronic Acids* in *General Protocols for Chapter Five General Protocols for Chapter Five* (**14** 675 mg, 3.65 mmol, 91% yield). White solid, mp 166-169 °C, <sup>1</sup>H NMR (400 MHz, DMSO) δ 7.20 (m, J = 9.0, 1.4 Hz, 1H, Ar-H), 6.77 (dd, J = 7.1, 1.5 Hz, 2H, Ar-H); <sup>13</sup>C NMR (101 MHz, DMSO) δ 162.76 (C), 160.30 (C), 151.10 (CH), 134.40 (C), 121.43 (CH), 110.30 (CH); <sup>19</sup>F NMR (376 MHz, DMSO) δ -111.81 (dt, J = 9.1, 7.0 Hz).

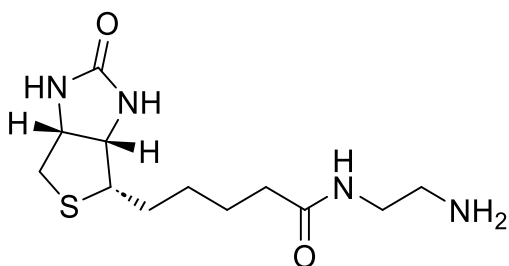
### Synthesis of methyl 5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanoate (**17**)



**17** was synthesised following an adapted literature procedure.<sup>172</sup> A solution of hydrogen chloride was firstly prepared by the dropwise addition of acetyl chloride (1.5 mL, 20.05 mmol) to ice cold MeOH (10 mL) over 10 min. The solution was stirred for 5 min and added to a suspension of d-biotin (**16**, 1.0 g, 4.1 mmol) in MeOH (10 mL). The r.m. was warmed to room temperature and stirred for 1 h (the solution changes from non-homogeneous to colourless) before being evaporated under reduced pressure. The crude was redissolved in MeOH : CH<sub>2</sub>Cl<sub>2</sub> (5 : 95, 50 mL) and washed with saturated aqueous NaHCO<sub>3</sub> (3 x 20 mL). The organic layer was isolated, dried over MgSO<sub>4</sub>, filtered, and evaporated under reduced pressure to yield the titled compound (0.7 g, 66 % yield) as a white solid: mp 153-157 °C; <sup>1</sup>H NMR (400 MHz, DMSO) δ 6.43 (s, 1H, NH), 6.35 (s, 1H, NH), 4.32 – 4.28 (m, 1H, CH), 4.13 (ddd, J =

7.7, 4.4, 1.9 Hz, 1H, CH), 3.58 (s, 3H, CH<sub>3</sub>), 3.10 (ddd, J = 8.6, 6.2, 4.4 Hz, 1H, CH), 2.82 (dd, J = 12.4, 5.1 Hz, 1H, CH), 2.57 (d, J = 12.4 Hz, 1H, CH), 2.30 (t, J = 7.5 Hz, 2H, CH<sub>2</sub>), 1.66 – 1.26 (m, 6H, CH<sub>3</sub>); <sup>13</sup>C NMR (101 MHz, DMSO) δ 173.31 (C), 162.70 (C), 61.02 (CH), 59.18 (CH), 55.33 (CH), 51.19 (CH<sub>3</sub>), 39.85 (CH<sub>2</sub>), 33.09 (CH<sub>2</sub>), 28.00 (CH<sub>2</sub>), 27.97 (CH<sub>2</sub>), 24.47 (CH<sub>2</sub>). These spectral properties agreed with previously reported literature.<sup>166</sup>

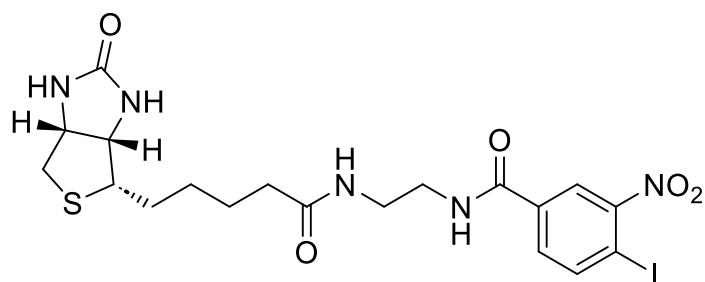
### Synthesis of N-(2-aminoethyl)-5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanamide (**18**)



To a solution of **17** (0.7 g, 2.7 mmol) in MeOH (24 mL) was added ethylenediamine (10 mL, 150 mmol) and the r.m. was stirred at 60 °C for 48 h before being evaporated under reduced pressure. The crude was triturated with excess ethyl acetate. Lyophilisation led to the titled compound (690 mg, 89% yield) as a white solid: <sup>1</sup>H NMR (400 MHz, DMSO) δ 7.75 (t, J = 5.6 Hz, 1H, sec. amide), 6.43 (s, 1H, NH), 6.36 (s, 1H, NH), 4.34 – 4.26 (m, 1H, CH), 4.16 – 4.08 (m, 1H, CH), 3.14 – 3.06 (m, 1H, CH), 3.02 (q, J = 6.2 Hz, 2H, CH<sub>2</sub>), 2.82 (dd, J = 12.4, 5.1 Hz, 1H, CH), 2.61 – 2.51 (m, 4H, CH and CH<sub>2</sub>), 2.06 (t, J = 7.4 Hz, 2H, NH<sub>2</sub>), 1.66 – 1.24 (m, 7H, CH<sub>2</sub> and CH); <sup>13</sup>C NMR (101 MHz, DMSO) δ 172.05 (C), 162.72 (C), 61.04 (CH), 59.19 (CH), 55.42 (CH), 45.99 (CH<sub>2</sub>), 42.15 (CH<sub>2</sub>), 41.37 (CH<sub>2</sub>), 35.21 (CH<sub>2</sub>), 28.22 (CH<sub>2</sub>), 28.03 (CH<sub>2</sub>), 25.29 (CH<sub>2</sub>); HRMS calc. for C<sub>12</sub>H<sub>22</sub>N<sub>4</sub>O<sub>2</sub>S: 286.15 [M+H]<sup>+</sup>, found: 287.15 [M+H]<sup>+</sup>.

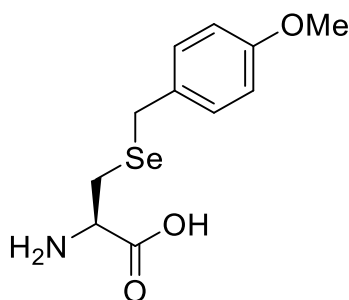


**Synthesis of 4-iodo-3-nitro-N-(2-(5-((3a*S*,4*S*,6a*R*)-2-oxohexahydro-1*H*-thieno[3,4-*d*]imidazol-4-yl)pentanamido)ethyl)benzamide (22)**



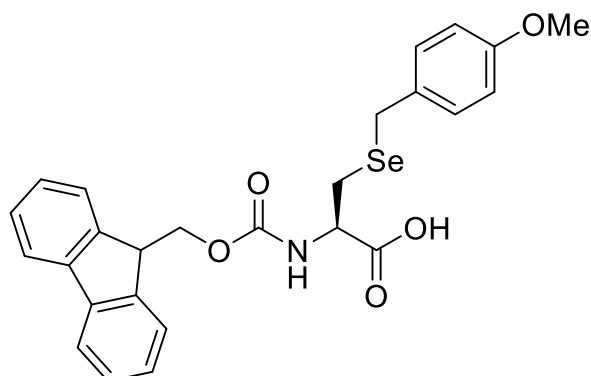
To a solution of 4-iodo-3-nitrobenzoic acid (0.2 g, 0.7 mmol) in DMF (2 mL) were added DIC (110  $\mu$ L, 0.7 mmol) and Oxyma (100 mg, 0.7 mmol). The resulting mixture was left stirring at r.t. for 5 min.. After this time, A solution of **18** (100 mg, 0.35 mmol) in DMF (2 mL) was added. The resulting mixture was left stirring at r.t. for 16 h. After this time, DMF was evaporated, and the crude orange oil was re-dissolved in ethyl acetate (50 mL). The organic layer was washed with H<sub>2</sub>O (5 x 25 mL) and brine (3 x 25 mL). The organic layer was separated, dried over MgSO<sub>4</sub>, and evaporated under reduced pressure to yield the titled compound (163 mg, 83%) as a white solid. <sup>1</sup>H NMR (400 MHz, DMSO)  $\delta$  8.93 (t, J = 5.4 Hz, 1H, sec. amide), 8.36 (t, J = 1.9 Hz, 1H, Ar-H), 8.23 (d, J = 8.2 Hz, 1H, Ar-H), 8.01 (t, J = 5.4 Hz, 1H, sec. amide), 7.90 – 7.82 (m, 1H, Ar-H), 6.39 (d, J = 20.0 Hz, 2H, NH), 4.33 – 4.25 (m, 1H, CH), 4.12 – 4.05 (m, 1H, CH), 3.33 – 3.28 (m, 2H, CH<sub>2</sub>), 3.22 (q, J = 6.2 Hz, 2H, CH<sub>2</sub>), 3.06 – 2.99 (m, 1H, CH), 2.79 (dd, J = 12.4, 5.1 Hz, 1H, CH), 2.57 (d, J = 12.4 Hz, 1H, CH), 2.06 (t, J = 7.4 Hz, 2H, CH<sub>2</sub>), 1.63 – 1.55 (m, 1H, CH), 1.54 – 1.46 (m, 2H, CH<sub>2</sub>), 1.46 – 1.37 (m, 1H, CH), 1.35 – 1.21 (m, 2H, CH<sub>2</sub>); <sup>13</sup>C NMR (101 MHz, DMSO)  $\delta$  172.36, 163.74, 162.70, 153.19, 141.44, 135.46, 131.88, 123.43, 92.14, 60.99, 59.19, 55.38, 39.84, 39.40, 37.96, 35.22, 28.14, 28.01, 25.21; HRMS calc. for C<sub>19</sub>H<sub>23</sub>IN<sub>5</sub>O<sub>5</sub>S: 560.05 [M-H]<sup>-</sup>, found: 560.05 [M-H]<sup>-</sup>.

## Synthesis of (R)-2-amino-3-((4-methoxybenzyl)selanyl)propanoic acid (48)



A 100 mL round-bottom flask was charged with selenocystine (1.8 g, 5.3 mmol) dissolved in 0.5 M NaOH (5 mL). NaBH<sub>4</sub> (1.7 g, 43 mmol) dissolved in water (10 mL) was added dropwise and the r.m. stirred in an ice-bath. After the vigorous reaction has subsided (the solution changes from yellow to colourless), the mixture was acidified to pH = 6.0 by the dropwise addition of glacial acetic acid. *p*-Methoxybenzyl chloride (1.44 mL, 10.6 mmol) was added dropwise, and the r.m. was stirred at room temperature for 1 h. After 1 h the r.m. was concentrated under reduced pressure and recrystallisation from hot water led to the title compound (2.7 g, 88% yield) as a white solid: <sup>1</sup>H NMR (400 MHz, DMSO) δ 8.67 (s, 2H, NH<sub>2</sub>), 7.30 – 7.21 (m, 2H, Ar-H), 6.87 – 6.77 (m, 2H, Ar-H), 4.23 – 4.14 (m, 1H, CH), 3.87 (s, 2H, CH<sub>2</sub>), 3.82 – 3.67 (m, 3H, CH<sub>2</sub>), 3.01 – 2.90 (m, 2H, CH<sub>2</sub>); <sup>13</sup>C NMR (101 MHz, DMSO) δ 178.38, 174.63, 169.67, 163.72, 159.42, 158.10, 155.88, 150.36, 145.93, 140.72, 130.83, 130.06, 129.31, 124.06, 117.43, 115.54, 113.86, 113.66, 102.91, 58.27, 55.07, 52.28, 49.17, 47.96, 40.15, 39.94, 39.73, 39.52, 39.31, 39.10, 38.89, 27.00, 22.32; HRMS calc. for C<sub>11</sub>H<sub>16</sub>NO<sub>3</sub>Se: 290.03 [M+H]<sup>+</sup>, found: 290.03 [M+H]<sup>+</sup>. These spectral properties agreed with previously reported literature.<sup>199,200</sup>

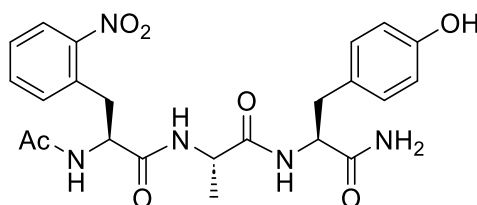
**Synthesis of (R)-2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-((4-methoxybenzyl)selenanyl)propanoic acid (49)**



**48** (2.2 g, 7.5 mmol) was dissolved in water (20 mL) to make a slurry and trimethylamine (1.1 mL, 7.5 mmol) was added. A solution of *N*-(9-fluorenylmethoxycarbonyloxy)succinimide (2.53 g, 7.5 mmol) in acetonitrile (20 mL) was added and a second equivalent of trimethylamine (1.1 mL, 7.5 mmol) was added. Acetonitrile was added to fully dissolve all solutes and the r.m. was stirred at room temperature for 1 h. After this time the r.m. was acidified to pH = 2 with 1M HCl (10 mL) and extracted with ethyl acetate (1 x 30 mL). The organic layer was washed with 1M HCl (3 x 20 mL). The combined aqueous layers were extracted with ethyl acetate (3 x 20 mL). The organic phases were combined, dried over MgSO<sub>4</sub>, filtered, and concentrated under reduced pressure. The crude residue was purified *via* flash chromatography (0-20% MeOH in CH<sub>2</sub>Cl<sub>2</sub>) to yield the titled compound (2.17 g, 57% yield) as a pale yellow solid: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.75 (dd, J = 7.6, 2.5 Hz, 2H, fluorene), 7.59 (q, J = 6.8 Hz, 2H, fluorene), 7.39 (td, J = 7.5, 3.4 Hz, 2H, fluorene), 7.35 – 7.26 (m, 2H, fluorene), 7.18 (d, J = 8.2 Hz, 2H, Ar-H), 6.80 (d, J = 8.0 Hz, 2H, Ar-H), 5.58 (d, J = 7.8 Hz, 1H, sec. amide), 4.66 (q, J = 6.0 Hz, 1H, fluorene), 4.41 (d, J = 7.1 Hz, 2H, CH<sub>2</sub>), 4.23 (t, J = 7.1 Hz, 1H, CH), 3.75 (s, 3H, CH<sub>3</sub>), 3.49 (s, 2H, CH<sub>2</sub>), 2.10 (s, 1H, CH), 2.00 (s, 1H, CH); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 172.26, 169.37, 158.73, 146.99, 143.74, 142.39 (d, J = 4.5 Hz), 141.82, 141.43, 140.84, 137.50, 130.53, 130.17, 127.91, 127.26, 125.23, 120.77, 120.15, 115.12, 114.21, 58.03, 55.39, 53.81, 50.95, 47.20, 27.79, 25.47; HRMS calc. for C<sub>26</sub>H<sub>25</sub>NO<sub>5</sub>SeNa: 534.08 [M+Na]<sup>+</sup>, found 534.08 [M+Na]<sup>+</sup>. These spectral properties agreed with previously reported literature.<sup>199,200</sup>

## 8.5 SYNTHESIS OF MODIFIED PEPTIDES

### Synthesis of (S)-2-acetamido-N-(((S)-1-(((S)-1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)-3-(2-nitrophenyl)propanamide (**2**)



Peptide **2** was synthesised using the general procedure outlined in Optimised *General Protocol for Cysteine Desulfurisation and  $\beta$ -Arylation* in *General Protocols for Chapter Five*. The crude r.m. was filtrated, purified by semi-preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired modified peptide (**2** 1.96 mg, 4.03  $\mu$ mol, 32% yield).

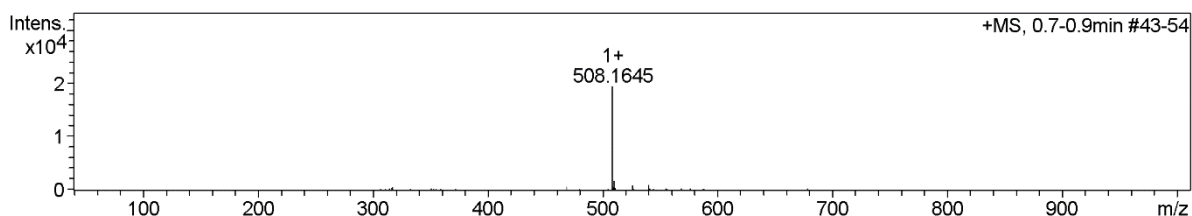
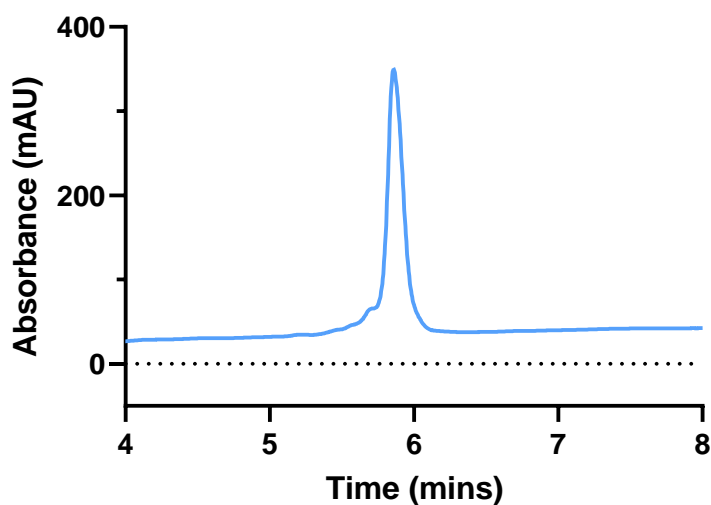
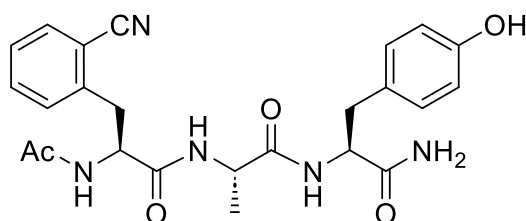


Figure 166 – Analytical HPLC trace and ESI MS of pure **2**. Analytical gradient 2-40% B over 5 min, 280 nm. Calculated Mass: 508.18 [M+Na]<sup>+</sup>, observed mass: 508.17 [M+Na]<sup>+</sup>.

**Synthesis of (S)-2-acetamido-N-(((S)-1-(((S)-1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)-3-(2-cyanophenyl)propenamamide (3)**



Peptide **3** was synthesised using the general procedure outlined in *Optimised General Protocol for Cysteine Desulfurisation and  $\beta$ -Arylation* in *General Protocols for Chapter Five*. The crude r.m. was filtrated, purified by semi-preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired modified peptide (**3** 1.47 mg, 3.15  $\mu$ mol, 25% yield).

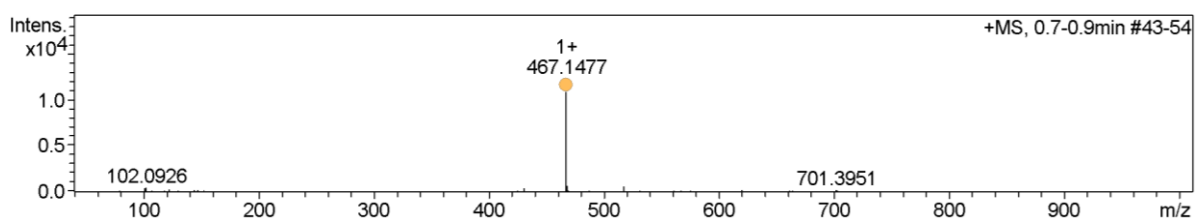
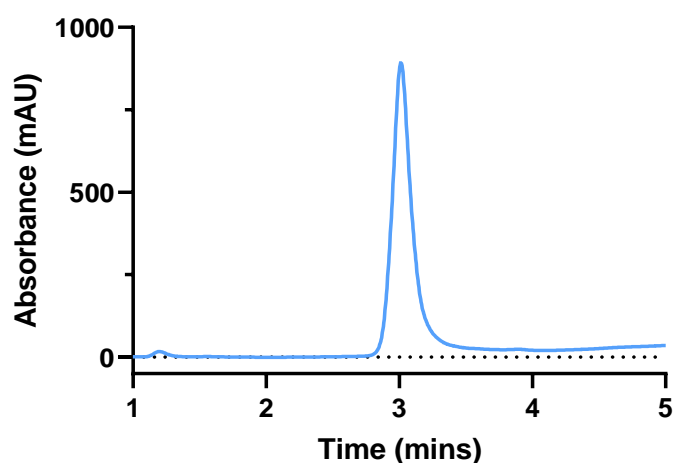
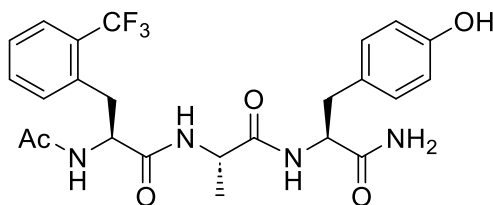


Figure 167 – Analytical HPLC trace and ESI MS of pure **3**. Analytical gradient 2-40% B over 5 min, 280 nm. Calculated Mass: 466.20 [M+H]<sup>+</sup>, observed mass: 467.15 [M+H]<sup>+</sup>.

**Synthesis of (S)-2-acetamido-N-((S)-1-(((S)-1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)-3-(2-(trifluoromethyl)phenyl)propenamide (4)**



Peptide **4** was synthesised using the general procedure outlined in *Optimised General Protocol for Cysteine Desulfurisation and  $\beta$ -Arylation in General Protocols for Chapter Five*. The crude r.m. was filtrated, purified by semi-preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired modified peptide (**4** 1.47 mg, 2.9  $\mu$ mol, 23% yield).

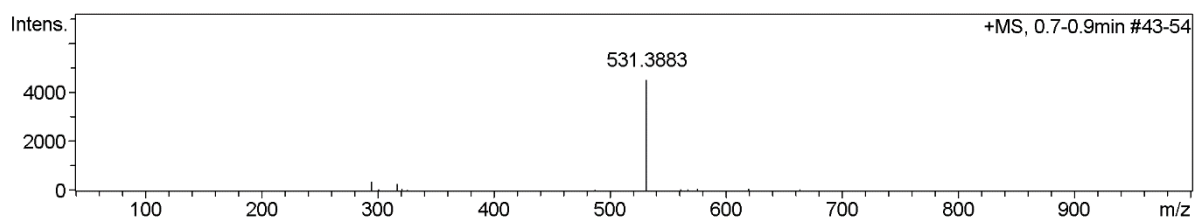
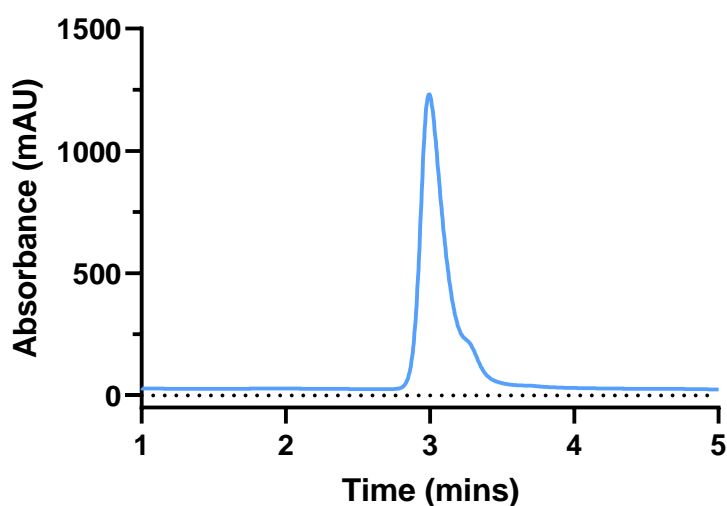
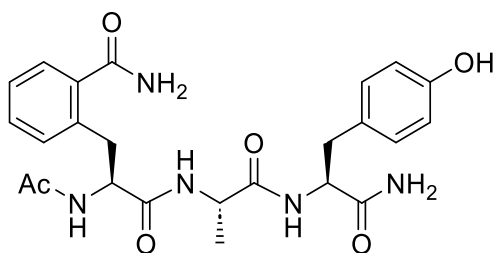


Figure 168 – Analytical HPLC trace and ESI MS of pure **4**. Analytical gradient 2-40% B over 5 min, 280 nm. Calculated Mass: 531.49 [M+Na]<sup>+</sup>, observed mass: 531.39 [M+Na]<sup>+</sup>.

**Synthesis of 2-((S)-2-acetamido-3-(((S)-1-(((S)-1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)amino)-3-oxopropyl)benzamide (5)**



Peptide **5** was synthesised using the general procedure outlined in *Optimised General Protocol for Cysteine Desulfurisation and  $\beta$ -Arylation* in *General Protocols for Chapter Five*. The crude r.m. was filtrated, purified by semi-preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired modified peptide (**5** 0.98 mg, 2.02  $\mu$ mol, 16% yield).

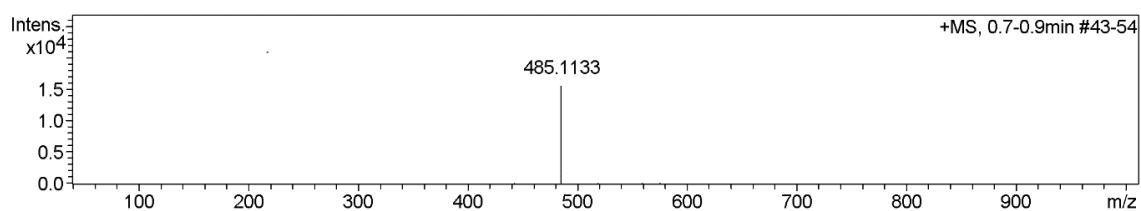
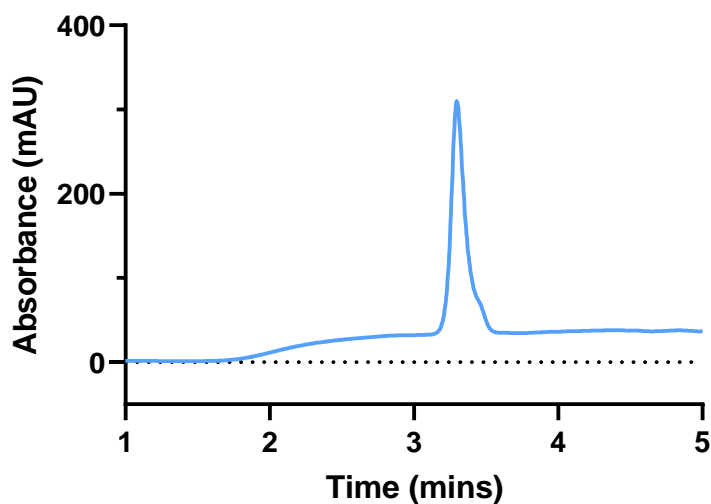
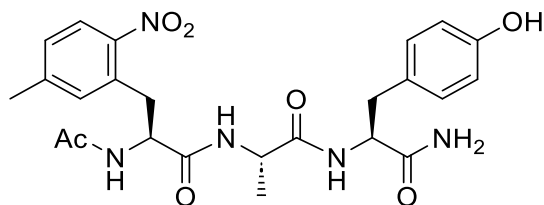


Figure 169 – Analytical HPLC trace and ESI MS of pure **5**. Analytical gradient 2-40% B over 5 min, 280 nm. Calculated Mass: 484.53 [M+H]<sup>+</sup>, observed mass: 485.11 [M+H]<sup>+</sup>.

**Synthesis of (S)-2-acetamido-N-((S)-1-(((S)-1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)-3-(5-methyl-2-nitrophenyl)propenamide (11)**



Peptide **11** was synthesised using the general procedure outlined in *Optimised General Protocol for Cysteine Desulfurisation and  $\beta$ -Arylation* in *General Protocols for Chapter Five*. The crude r.m. was filtrated, purified by semi-preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired modified peptide (**11** 1.57 mg, 3.15  $\mu$ mol, 25% yield).

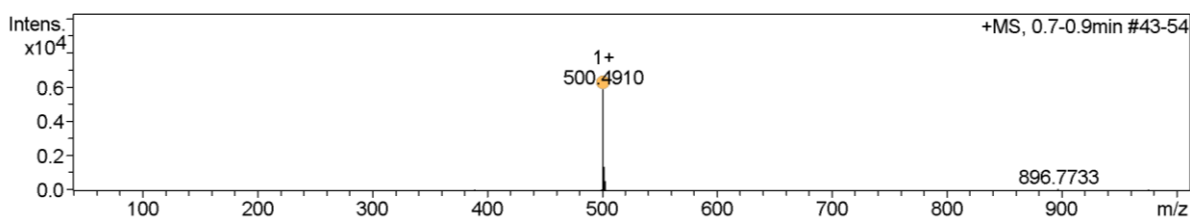
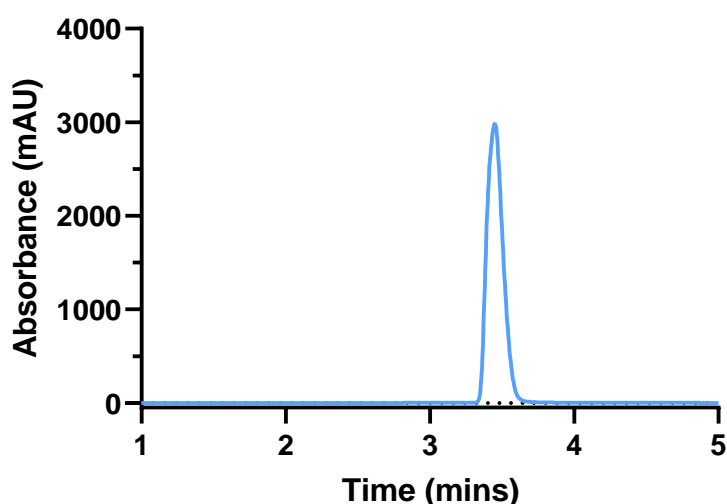
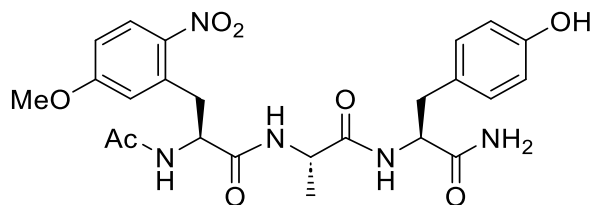


Figure 170 – Analytical HPLC trace and ESI MS of pure **11**. Analytical gradient 2-40% B over 5 min, 280 nm. Calculated Mass: 499.52  $[M+H]^+$ , observed mass: 500.49  $[M+H]^+$ .



**Synthesis of (S)-2-acetamido-N-((S)-1-(((S)-1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)-3-(5-methoxy-2-nitrophenyl)propenamide (13)**



Peptide **13** was synthesised using the general procedure outlined in *Optimised General Protocol for Cysteine Desulfurisation and  $\beta$ -Arylation* in *General Protocols for Chapter Five*. The crude r.m. was filtrated, purified by semi-preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired modified peptide (**13** 0.52 mg, 1.01  $\mu$ mol, 8% yield).

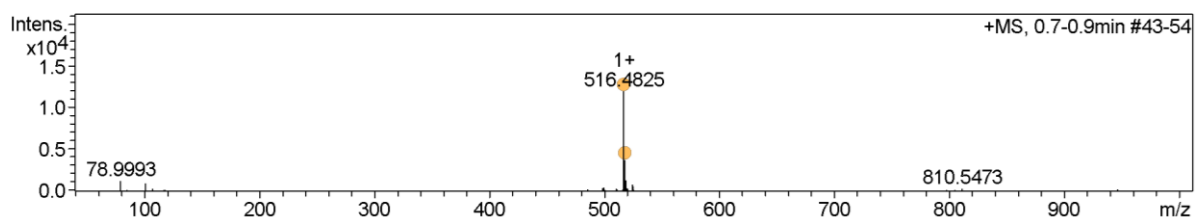
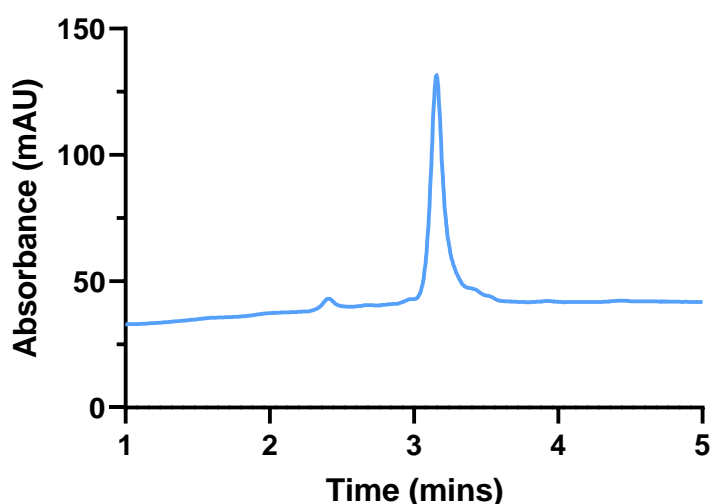
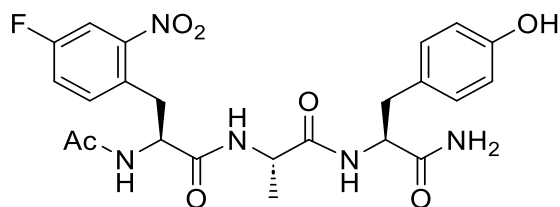


Figure 171 – Analytical HPLC trace and ESI MS of pure **13**. Analytical gradient 2-40% B over 5 min, 280 nm. Calculated Mass: 515.52 [M+H]<sup>+</sup>, observed mass: 516.48 [M+H]<sup>+</sup>.

**Synthesis of (S)-2-acetamido-N-((S)-1-(((S)-1-amino-3-(4-hydroxyphenyl)-1-oxopropan-2-yl)amino)-1-oxopropan-2-yl)-3-(4-fluoro-2-nitrophenyl)propenamide (**15**)**



Peptide **15** was synthesised using the general procedure outlined in *Optimised General Protocol for Cysteine Desulfurisation and  $\beta$ -Arylation* in *General Protocols for Chapter Five*. The crude r.m. was filtrated, purified by semi-preparative RP-HPLC (2-40% B over 30 min) and lyophilised to produce the desired modified peptide (**15** 2.09 mg, 4.16  $\mu$ mol, 33% yield).

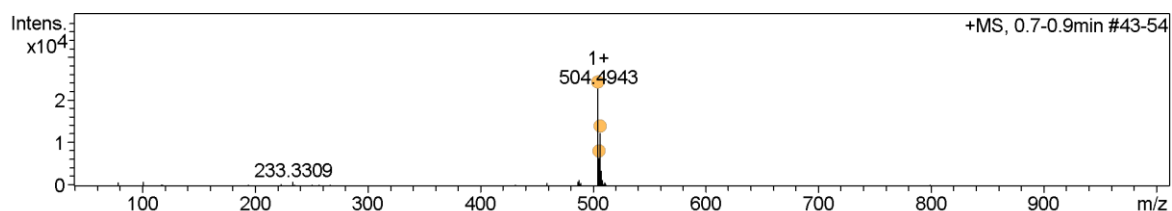
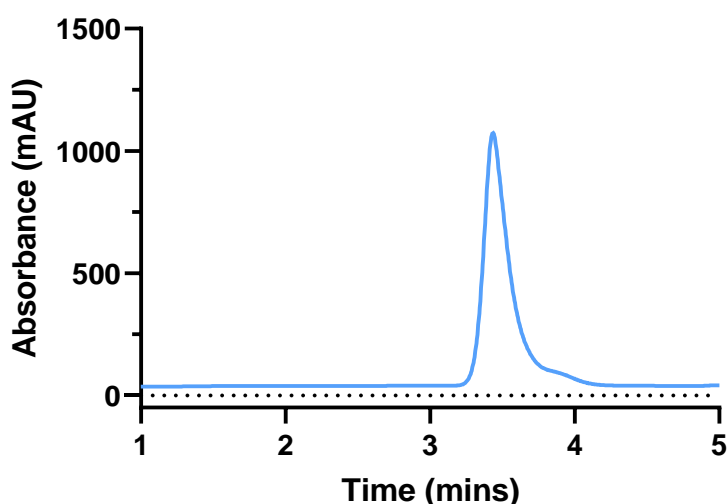


Figure 172 – Analytical HPLC trace and ESI MS of pure **15**. Analytical gradient 2-40% B over 5 min, 280 nm. Calculated Mass: 503.49 [M+H]<sup>+</sup>, observed mass: 504.49 [M+H]<sup>+</sup>.

## REFERENCES

- 1 P. Sánchez-Baracaldo and T. Cardona, *New Phytol.*, 2020, **225**, 1440–1446.
- 2 Y. Nicolet and J. C. Fontecilla-Camps, *Iron-Sulfur Clusters in Chemistry and Biology*, De Gruyter, Berlin, 2014.
- 3 H. A. H. Sands, *Biochem. Biophys. Res. Commun.*, 1960, **3**, 41–46.
- 4 J. Liu, S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, I. Petrik, A. Bhagi and Y. Lu, *Chem. Rev.*, 2014, **114**, 4366–4369.
- 5 C. Bonfio, L. Valer, S. Scintilla, S. Shah, D. J. Evans, L. Jin, J. W. Szostak, D. D. Sasselov, J. D. Sutherland and S. S. Mansy, *Nat. Chem.*, 2017, **9**, 1229–1234.
- 6 D. W. Bak and S. J. Elliott, *Curr. Opin. Chem. Biol.*, 2014, **19**, 50–58.
- 7 A. E. Boncella, E. T. Sabo, R. M. Santore, J. Carter, J. Whalen, J. D. Hudspeth and C. N. Morrison, *Coord. Chem. Rev.*, 2022, **453**, 1–71.
- 8 P. A. Frey, A. D. Hegeman and F. J. Ruzicka, *Curr. Opin. Chem. Biol.*, 2008, **1**, 63–88.
- 9 B. J. Landgraf, E. L. McCarthy and S. J. Booker, *Annu. Rev. Biochem.*, 2016, **85**, 485–514.
- 10 R. D. Finn, T. K. Attwood, P. C. Babbitt, A. Bateman, P. Bork, A. J. Bridge, H. Y. Chang, Z. Dosztanyi, S. El-Gebali, M. Fraser, J. Gough, D. Haft, G. L. Holliday, H. Huang, X. Huang, I. Letunic, R. Lopez, S. Lu, A. Marchler-Bauer, H. Mi, J. Mistry, D. A. Natale, M. Necci, G. Nuka, C. A. Orengo, Y. Park, S. Pesseat, D. Piovesan, S. C. Potter, N. D. Rawlings, N. Redaschi, L. Richardson, C. Rivoire, A. Sangrador-Vegas, C. Sigrist, I. Sillitoe, B. Smithers, S. Squizzato, G. Sutton, N. Thanki, P. D. Thomas, S. C. E. Tosatto, C. H. Wu, I. Xenarios, L. S. Yeh, S. Y. Young and A. L. Mitchell, *Nucleic Acids Res.*, 2017, **4**, 190–199.

- 11 H. Yang, E. C. McDaniel, S. Impano, A. S. Byer, R. J. Jodts, K. Yokoyama, W. E. Broderick, J. B. Broderick and B. M. Hoffman, *J. Am. Chem. Soc.*, 2019, **141**, 12139–12146.
- 12 J. B. Broderick, B. R. Duffus, K. S. Duschene and E. M. Shepard, *Chem. Rev.*, 2014, **114**, 4229–4317.
- 13 J. Knappe, F. A. Neugebauer, H. P. Blaschkowski and M. Ganzler, *Proc. Natl. Acad. Sci. U. S. A.*, 1984, **81**, 1332–1335.
- 14 J. Knappe, J. Schacht, W. Möckel, T. Höpner, H. Vetter and R. Edenharder, *Eur. J. Biochem.*, 1969, **11**, 316–327.
- 15 B. J. Levin, Y. Y. Huang, S. C. Peck, Y. Wei, A. Martínez-Del Campo, J. A. Marks, E. A. Franzosa, C. Huttenhower and E. P. Balskus, *Science (1979)*, 2017, **355**, 1–9.
- 16 J. B. Broderick, R. E. Duderstadt, D. C. Fernandez, K. Wojtuszewski, T. F. Henshaw and M. K. Johnson, *J. Am. Chem. Soc.*, 1997, **119**, 7396–7397.
- 17 J. Knappe and T. Schmitt, *Biochem. Biophys. Res. Commun.*, 1976, **71**, 1110–1117.
- 18 J. L. Vey and C. L. Drennan, *Chem. Rev.*, 2011, **111**, 2487–2506.
- 19 C. M. Jäger and A. K. Croft, *J. Biol. Inorg. Chem.*, 2018, **5**, 143–162.
- 20 A. Galambas, J. Miller, M. Jones, E. McDaniel, M. Lukes, H. Watts, V. Copié, J. B. Broderick, R. K. Szilagyí and E. M. Shepard, *J. Biol. Inorg. Chem.*, 2019, **24**, 793–807.
- 21 W. E. Broderick, B. M. Hoffman and J. B. Broderick, *Acc. Chem. Res.*, 2018, **51**, 2611–2619.
- 22 C. J. Walsby, D. Ortillo, J. Yang, M. R. Nnyepi, W. E. Broderick, B. M. Hoffman and J. B. Broderick, *Inorg. Chem.*, 2005, **44**, 727–741.
- 23 R. I. Sayler, T. A. Stich, S. Joshi, N. Cooper, J. T. Shaw, T. P. Begley, D. J. Tantillo and R. D. Britt, *ACS Cent. Sci.*, 2019, **5**, 1777–1785.

- 24 J. A. Kampmeier, *Biochemistry*, 2010, **49**, 10770–10772.
- 25 L. Flühe, T. A. Knappe, M. J. Gattner, A. Schäfer, O. Burghaus, U. Linne and M. A. Marahiel, *Nat. Chem. Biol.*, 2012, **8**, 350–357.
- 26 A. Caruso, L. B. Bushin, K. A. Clark, R. J. Martinie and M. R. Seyedsayamdost, *J. Am. Chem. Soc.*, 2019, **141**, 990–997.
- 27 W. M. Kincannon, N. A. Bruender and V. Bandarian, *Biochemistry*, 2018, **57**, 4816–4823.
- 28 G. A. Hudson, B. J. Burkhart, A. J. DiCaprio, C. J. Schwalen, B. Kille, T. V. Pogorelov and D. A. Mitchell, *J. Am. Chem. Soc.*, 2020, **141**, 8228–8238.
- 29 K. R. Schramma, L. B. Bushin and M. R. Seyedsayamdost, *Nat. Chem.*, 2015, **7**, 431–437.
- 30 I. Barr, J. A. Latham, A. T. Iavarone, T. Chantarojsiri, J. D. Hwang and J. P. Klinman, *J. Biol. Chem.*, 2016, **291**, 8877–8884.
- 31 K. R. Schramma, C. C. Forneris, A. Caruso and M. R. Seyedsayamdost, *Biochemistry*, 2018, **57**, 461–468.
- 32 L. B. Bushin, K. A. Clark, I. Pelczer and M. R. Seyedsayamdost, *J. Am. Chem. Soc.*, 2018, **140**, 17674–17684.
- 33 R. Ayikpoe, J. Salazar, B. Majestic and J. A. Latham, *Biochemistry*, 2018, **57**, 5379–5383.
- 34 L. Huo, S. Rachid, M. Stadler, S. C. Wenzel and R. Müller, *Chem. Biol.*, 2012, **19**, 1278–1287.
- 35 N. Mahanta, Z. Zhang, G. A. Hudson, W. A. Van Der Donk and D. A. Mitchell, *J. Am. Chem. Soc.*, 2017, **139**, 4310–4313.
- 36 M. F. Freeman, C. Gurgui, M. J. Helf, B. I. Morinaka, A. R. Uria, N. J. Oldham, H. G. Sahl, S. Matsunaga and J. Piel, *Science (1979)*, 2012, **338**, 387–390.

- 37 A. Benjdia, A. Guillot, P. Ruffié, J. Leprince and O. Berteau, *Nat. Chem.*, 2017, **9**, 698–707.
- 38 B. I. Morinaka, E. Lakis, M. Verest, M. J. Helf, T. Scalvenzi, A. L. Vagstad, J. Sims, S. Sunagawa, M. Gugger and J. Piel, *Science (1979)*, 2018, **359**, 779–782.
- 39 P. C. Nalam, N. N. Gosvami, M. A. Caporizzo, R. J. Composto, R. W. Carpick and A. Mechanics, *Science (1979)*, 2016, **11**, 8165–8178.
- 40 W. E. Broderick and J. B. Broderick, *J. Biol. Inorg. Chem.*, 2019, **24**, 769–776.
- 41 H. Matsubara and K. Saeki, *Arch. Biochem. Biophys.*, 1992, **474**, 283–291.
- 42 R. C. Valentine, *Bacteriol. Rev.*, 1964, **28**, 497–517.
- 43 W. Qi and J. A. Cowan, *Coord. Chem. Rev.*, 2011, **255**, 688–699.
- 44 J. C. Crack, J. Green, M. I. Hutchings, A. J. Thomson and N. E. Le Brun, *Antioxid. Redox Signaling*, 2012, **17**, 1215–1231.
- 45 J. C. Crack, J. Green, M. R. Cheesman, N. E. Le Brun and A. J. Thomson, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 2092–2097.
- 46 C. M. Jäger and A. K. Croft, *Biochemistry*, 2023, **62**, 241–252.
- 47 G. Xing and DeRose V. J., *Curr. Opin. Chem. Biol.*, 2001, **5**, 196–200.
- 48 F. Natri, M. Chino, O. Maglio, A. Bhagi-Damodaran, Y. Lu and A. Lombardi, *Chem. Soc. Rev.*, 2016, **45**, 5020–5054.
- 49 A. Nakamura and N. Ueyama, in *Metal Clusters in Proteins*, ed. Lawrence Que Jr, American Chemical Society, Washington, 19th edn., 1988, vol. 372, pp. 292–301.
- 50 M. Lotierzo, B. Tse, S. Bui, D. Florentin, F. Escalettes and A. Marquet, *Biochem. Soc. Trans.*, 2005, **33**, 820–823.
- 51 E. Choi-Rhee and J. E. Cronan, *Cell Chem. Biol.*, 2005, **12**, 461–468.

- 52 A. Caruso, R. J. Martinie, L. B. Bushin and M. R. Seyedsayamdost, *J. Am. Chem. Soc.*, 2019, **141**, 16610–16614.
- 53 P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K. D. Entian, M. A. Fischbach, J. S. Garavelli, U. Göransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Müller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. T. Reaney, S. Rebuffat, R. P. Ross, H. G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Süßmuth, J. R. Tagg, G. L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey and W. A. Van Der Donk, *Nat. Prod. Rep.*, 2014, **30**, 108–160.
- 54 K. Choi and A. D. Hamilton, *Coord. Chem. Rev.*, 2003, **240**, 101–110.
- 55 E. K. Sinner, M. S. Lichstrahl, R. Li, D. R. Marous and C. A. Townsend, *Chem. Comm.*, 2019, **55**, 14934–14937.
- 56 M. Kühner, P. Schweyen, M. Hoffmann, J. V. Ramos, E. J. Reijerse, W. Lubitz, M. Bröring and G. Layer, *Chem. Sci.*, 2016, **7**, 4633–4643.
- 57 A. D. Sheftel, C. Wilbrecht, O. Stehling, B. Niggemeyer, H. P. Elsässer, U. Mühlhoff and R. Lill, *Mol. Biol. Cell.*, 2012, **23**, 1157–1166.
- 58 W. H. Tong and T. A. Rouault, *Cell. Metab.*, 2006, **3**, 199–210.
- 59 D. Brancaccio, A. Gallo, M. Mikolajczyk, K. Zovo, P. Palumaa, E. Novellino, M. Piccioli, S. Ciofi-Baffoni and L. Banci, *J. Am. Chem. Soc.*, 2014, **136**, 16240–16250.
- 60 R. Hanscam, E. M. Shepard, J. B. Broderick, V. Copié and R. K. Szilagyi, *J. Comput. Chem.*, 2019, **40**, 515–526.

- 61 R. Malkin and J. Rabinowitz, *Biochem. Biophys. Res. Commun.*, 1966, **23**, 822–827.
- 62 Y. Dong, S. Zhang and L. Zhao, *Chin. J. Chem.*, 2022, **40**, 1478–1491.
- 63 D. B. Weiler, M. C. Brück, I. Kothe, E. Bill, R. Lill and U. Mühlhoff, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 20555–20565.
- 64 L. Que, J. R. Anglin, A. Bobrik, A. Davison and R. H. Holm, *J. Am. Chem. Soc.*, 1974, **96**, 6042–6048.
- 65 A. Tomita, H. Hirai, S. Makishima, A. Tomita, H. Hirai and S. Makishima, *Inorg. Nucl. Chem. Lett.*, 1967, **4**, 715–718.
- 66 B. R. Gibney, S. E. Mulholland, F. Rabanal and P. L. Dutton, *Biochemistry*, 1996, **93**, 15041–15046.
- 67 A. Hoppe, M. E. Pandelia, W. Gärtner and W. Lubitz, *Biochim. Biophys. Acta Bioenerg.*, 2011, **1807**, 1414–1422.
- 68 S. E. Mulholland, B. R. Gibney, F. Rabanal and P. Leslie Dutton, *J. Am. Chem. Soc.*, 1998, **120**, 10296–10302.
- 69 S. E. Mulholland, B. R. Gibney, F. Rabanal and P. L. Dutton, *Biochemistry*, 1999, **38**, 10442–10448.
- 70 M. R. Ghadiri, J. R. Granja and L. K. Buehler, *Nature*, 1994, **369**, 301–304.
- 71 A. K. Brisdon, in *Inorganic Spectroscopic Methods*, ed. Oxford University Press, Oxford University Press, Oxford, 1st edn., 1998, pp. 70–73.
- 72 C. J. Adams, A. Baber, S. Boonyuen, N. G. Connelly, B. E. Diosdado, A. Kantacha, A. G. Orpen and E. Patrón, *Dalton Trans.*, 2009, **44**, 9746–9758.
- 73 J. LeGall, M. Y. Liu, C. M. Gomes, V. Braga, I. Pacheco, M. Regalla, A. V. Xavier and M. Teixeira, *FEBS Lett.*, 1998, **429**, 295–298.
- 74 S.-A. Freibert, B. D. Weiler, E. Bill, A. J. Pierik, U. Mühlhoff and R. Lill, *Methods Enzymol.*, 2018, **599**, 197–226.



- 75 B. Guigliarelli and P. Bertrand, *Adv. Inorg. Chem.*, 1999, **47**, 421–497.
- 76 I. R. Vassiliev, M. L. Antonkine and J. H. Golbeck, *Biochim. Biophys. Acta Bioenerg.*, 2001, **1507**, 139–160.
- 77 A. Winkler, S. J. Papadakis, De Poortere E P and M. Shayegan, in *Advances in Solid State Physics*, ed. B. Kramer, Springer, Heidelberg, 1st edn., 2001, vol. 41, pp. 211–223.
- 78 W. R. Hagen, *Biomolecular EPR Spectroscopy*, CRC Press, Boca Raton, 1st edn., 2008.
- 79 H. Edelhofer, *Biochemistry*, 1967, **6**, 1948–1954.
- 80 S. Scheinok, P. Leveque, P. Sonveaux, B. Driesschaert and B. Gallez, *Free Radic. Res.*, 2018, **52**, 1182–1196.
- 81 S. Ciurli and F. Musiani, *Photosynth. Res.*, 2005, **85**, 115–131.
- 82 S. G. Mayhew, *Eur. J. Biochem.*, 1978, **85**, 535–547.
- 83 A. Edgar and Y. Haider, *J. Phys. C: Solid State Phys.*, 1978, **11**, 2595–2605.
- 84 Z. Klencsár and Z. Köntös, *J. Phys. Chem. A*, 2018, **122**, 3190–3203.
- 85 O. Tiedt, J. Fuchs, W. Eisenreich and M. Boll, *J. Biol. Chem.*, 2018, **293**, 10264–10274.
- 86 C. Verduyn, R. Van Kleef, J. Frank, H. Schreuder, J. P. Van Dijken and W. Alexander Scheffers, *Biochem. J.*, 1985, **226**, 669–677.
- 87 D. Bermejo-Velasco, A. Azémar, O. P. Oommen, J. Hilborn and O. P. Varghese, *Biomacromolecules*, 2019, **20**, 1412–1420.
- 88 M. Shanmugam, M. Quareshy, A. D. Cameron, T. D. H. Bugg and Y. Chen, *Angew. Chem., Int. Ed.*, 2021, **60**, 4529–4534.
- 89 Y. Orii, T. Miki and K. Kakinuma, *Photochem. Photobiol.*, 1995, **61**, 261–268.

- 90 J. R. Rumble, *CRC Handbook of Chemistry and Physics*, CRC Press, Boca Raton, 102nd edn., 2017.
- 91 A. G. Cațianis, B. Virgolici, B. C. Dogaru, H. Virgolici and M. Mohora, *Acta Medica Transilvanica*, 2020, **25**, 56–59.
- 92 J. Meyer, J.-M. Moulis and J. Gaillard, *Adv. Inorg. Chem.*, 1992, **38**, 73–115.
- 93 T. D. Weatherill, T. B. Rauchfuss and R. A. Scott, *Inorg. Chem.*, 1986, **25**, 1466–1472.
- 94 T. M. Buscagan, J. T. Kaiser and D. C. Rees, *Elife*, 2022, **11**, 1–14.
- 95 C. P. Owens and F. A. Tezcan, in *Methods in Enzymology*, Academic Press Inc., 2018, vol. 599, pp. 355–386.
- 96 W. W. Cleland, *Biochemistry*, 1964, **3**, 480–482.
- 97 E. G. Goh, X. Xu and P. G. McCormick, *Scr. Mater.*, 2014, **78–79**, 49–52.
- 98 E. J. F. Dickinson, J. G. Limon-Petersen, N. V Rees and R. G. Compton, *J. Phys. Chem. C*, 2009, **113**, 11157–11171.
- 99 M. Yonemura, T. Sekine and H. Ueda, *J. Phys. Chem.*, 1986, **90**, 3003–3005.
- 100 N. Elgrishi, K. J. Rountree, B. D. McCarthy, E. S. Rountree, T. T. Eisenhart and J. L. Dempsey, *J. Chem. Educ.*, 2018, **95**, 197–206.
- 101 P. Spitzer, S. Wunderli, K. Maksymiuk, A. Michalska, A. Kisiel and Z. Galus, in *Handbook of Reference Electrodes*, eds. G. Inzelt, A. Lewenstam and F. Scholz, Springer, Heidelberg, 1st edn., 2013, vol. 1, pp. 77–143.
- 102 D. Polcari, P. Dauphin-Ducharme and J. Mauzeroll, *Chem. Rev.*, 2016, **116**, 13234–13278.
- 103 D. Wu, L. Wu, Q. Hu, Q. Mahmood and A. Khan, *J. Chem. Soc. Pak.*, 2011, **33**, 65–70.
- 104 P. T. Kissinger and W. R. Heineman, *J. Chem. Educ.*, 1983, **60**, 702–706.

- 105 F. Marken, A. Neudeck and A. M. Bond, in *Electroanalytical Methods*, ed. F. Scholz, Springer, Heidelberg, 1st edn., 2005, vol. 1, pp. 51–97.
- 106 A. Roy, D. J. Sommer, R. A. Schmitz, C. L. Brown, D. Gust, A. Astashkin and G. Ghirlanda, *J. Am. Chem. Soc.*, 2014, **136**, 17343–17349.
- 107 K. Wijeratne, PhD Thesis, Linköping University, 2018.
- 108 J. T. Atkinson, I. Campbell, G. N. Bennett and J. J. Silberg, *Biochemistry*, 2016, **55**, 7047–7064.
- 109 M. L. Kennedy and B. R. Gibney, *J. Am. Chem. Soc.*, 2002, **124**, 6826–6827.
- 110 F. Li, P. B. Lutz, Y. Pepelyayeva, E. S. J. Arnér, C. A. Bayse and S. Rozovsky, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 6976–6981.
- 111 B. J. Byun and Y. K. Kang, *Biopolymers*, 2011, **95**, 345–353.
- 112 C. Jacob, G. I. Giles, N. M. Giles and H. Sies, *Angew. Chem., Int. Ed.*, 2003, **42**, 4742–4758.
- 113 E. S. J. Arnér, *Exp. Cell. Res.*, 2010, **316**, 1296–1303.
- 114 J. T. Atkinson, I. Campbell, G. N. Bennett and J. J. Silberg, *Biochemistry*, 2016, **55**, 7047–7064.
- 115 P. M. Vignais, B. Billoud and J. Meyer, *FEMS Microbiol. Rev.*, 2001, **25**, 455–501.
- 116 P. W. King, M. C. Posewitz, M. L. Ghirardi and M. Seibert, *J. Bacteriol.*, 2006, **188**, 2163–2172.
- 117 R. Roskoski, in *Reference Module in Biomedical Sciences*, Elsevier, 2015.
- 118 M. Demuez, L. Cournac, O. Guerrini, P. Soucaille and L. Girbal, *FEMS Microbiol. Lett.*, 2007, **275**, 113–121.
- 119 O. Guerrini, B. Burlat, C. Léger, B. Guigliarelli, P. Soucaille and L. Girbal, *Curr. Microbiol.*, 2008, **56**, 261–267.

- 120 M. McFarland, in *Biosolids Engineering*, McGraw-Hill, New York, 1st edn., 2001, vol. 1.
- 121 Mach H., Middaugh C. R. and Lewis R. V., *Anal Biochem.*, 1992, **200**, 74–80.
- 122 J. M. Antosiewicz and D. Shugar, *Biophys. Rev.*, 2016, **8**, 163–177.
- 123 L. P. Dutton, *Methods Enzymol.*, 1978, **54**, 411–435.
- 124 N. Stephanopoulos and M. B. Francis, *Nat. Chem. Biol.*, 2011, **7**, 876–884.
- 125 L. Stryer, J. Berg, J. Tymoczko and G. Gatto, *Biochemistry*, W. H. Freeman, New York, 9th edn., 2019.
- 126 T. A. King, J. Mandrup Kandemir, S. J. Walsh and D. R. Spring, *Chem. Soc. Rev.*, 2021, **50**, 39–57.
- 127 J. Kalia and R. T. Raines, *Curr. Org. Chem.*, 2010, **14**, 138–147.
- 128 S. B. Gunnoo and A. Madder, *ChemBioChem.*, 2016, **17**, 529–553.
- 129 J. M. Chalker, G. J. L. Bernardes, Y. A. Lin and B. G. Davis, *Chem. Asian J.*, 2009, **4**, 630–640.
- 130 M. D. Simon, F. Chu, L. R. Racki, C. C. De La Cruz, A. L. Burlingame, B. Panning, G. J. Narlikar and K. M. Shokat, *Cell*, 2007, **128**, 1003–1012.
- 131 Greg T. Hermanson, *Bioconjugate Techniques*, Elsevier Science, Amsterdam, 3rd edn., 2013.
- 132 M. E. B. Smith, F. F. Schumacher, C. P. Ryan, L. M. Tedaldi, D. Papaioannou, G. Waksman, S. Caddick and J. R. Baker, *J. Am. Chem. Soc.*, 2010, **132**, 1960–1965.
- 133 A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y. S. Lin and B. L. Pentelute, *J. Am. Chem. Soc.*, 2013, **135**, 5946–5949.

- 134 M. Kasper, M. Glanz, A. Stengl, M. Penkert, S. Klenk, T. Sauer, D. Schumacher, J. Helma, E. Krause, M. C. Cardoso, H. Leonhardt and C. P. R. Hackenberger, *Angew. Chem.*, 2019, **131**, 11751–11756.
- 135 A. Abbas, B. Xing and T. P. Loh, *Angew. Chem., Int. Ed.*, 2014, **53**, 7491–7494.
- 136 M. J. Matos, C. D. Navo, T. Hakala, X. Ferhati, A. Guerreiro, D. Hartmann, B. Bernardim, K. L. Saar, I. Compañón, F. Corzana, T. P. J. Knowles, G. Jiménez-Osés and G. J. L. Bernardes, *Angew. Chem.*, 2019, **131**, 6712–6716.
- 137 B. Bernardim, P. M. S. D. Cal, M. J. Matos, B. L. Oliveira, N. Martínez-Saéz, I. S. Albuquerque, E. Perkins, F. Corzana, A. C. B. Burtoloso, G. Jiménez-Osés and G. J. L. Bernardes, *Nat. Commun.*, 2016, **7**, 1–9.
- 138 C. E. Hoyle and C. N. Bowman, *Angew. Chem., Int. Ed.*, 2010, **49**, 1540–1573.
- 139 C. Zhang, E. V. Vinogradova, A. M. Spokoyny, S. L. Buchwald and B. L. Pentelute, *Angew. Chem.*, 2019, **131**, 4860–4892.
- 140 E. V. Vinogradova, C. Zhang, A. M. Spokoyny, B. L. Pentelute and S. L. Buchwald, *Nature*, 2015, **526**, 687–691.
- 141 S. Schreiber, *Therap. Adv. Gastroenterol.*, 2011, **4**, 375–389.
- 142 M. R. Lewis and J. E. Shively, *Bioconjug. Chem.*, 1998, **9**, 72–86.
- 143 J. Q. Liu, A. Shatskiy, B. S. Matsuura and M. D. Kärkäs, *Synthesis (Stuttg)*, 2019, **51**, 2759–2791.
- 144 C. Bottecchia and T. Noël, *Chem. - Eur. J.*, 2019, **25**, 26–42.
- 145 T. P. Yoon, M. A. Ischay and J. Du, *Nat. Chem.*, 2010, **2**, 527–532.
- 146 Y. Su, K. P. L. Kuijpers, N. König, M. Shang, V. Hessel and T. Noël, *Chem. - Eur. J.*, 2016, **22**, 12295–12300.
- 147 Merck Photocatalysis Chart, <https://macmillan.princeton.edu/wp-content/uploads/Merck-Photocatalysis-Chart.pdf>, (accessed 19 August 2023).

- 148 L. Chen, I. Annis and G. Barany, *Curr. Protoc. Protein Sci.*, 2001, **23**, 1–19.
- 149 J. P. Tam, C. R. Wu, W. Liu and J. W. Zhang, *J. Am. Chem. Soc.*, 1991, **113**, 6657–6662.
- 150 E. Calce, R. M. Vitale, A. Scaloni, P. Amodeo and S. De Luca, *Amino Acids*, 2015, **47**, 1507–1515.
- 151 C. Bottecchia, N. Erdmann, P. M. A. Tijssen, L. G. Milroy, L. Brunsveld, V. Hessel and T. Noël, *ChemSusChem*, 2016, **9**, 1781–1785.
- 152 A. Talla, B. Driessen, N. J. W. Straathof, L. G. Milroy, L. Brunsveld, V. Hessel and T. Noël, *Adv. Synth. Catal.*, 2015, **357**, 2180–2186.
- 153 E. L. Tyson, Z. L. Niemeyer and T. P. Yoon, *J. Org. Chem.*, 2014, **79**, 1427–1436.
- 154 E. L. Tyson, M. S. Ament and T. P. Yoon, *J. Org. Chem.*, 2013, **78**, 2046–2050.
- 155 A. Meister and M. E. Anderson, *Annu. Rev. Biochem.*, 1983, **52**, 711–760.
- 156 B. A. Vara, X. Li, S. Berritt, C. R. Walters, E. J. Petersson and G. A. Molander, *Chem. Sci.*, 2018, **9**, 336–344.
- 157 GSK Solvent Selection Guide, <https://www.rsc.org/suppdata/gc/c0/c0gc00918k/c0gc00918k.pdf>, (accessed 19 August 2023).
- 158 C. Bottecchia, M. Rubens, S. B. Gunnoo, V. Hessel, A. Madder and T. Noël, *Angew. Chem., Int. Ed.*, 2017, **56**, 12702–12707.
- 159 H. A. Beard, J. R. Hauser, M. Walko, R. M. George, A. J. Wilson and R. S. Bon, *Commun. Chem.*, 2019, **2**, 1–9.
- 160 S. Sato and H. Nakamura, *Angew. Chem.*, 2013, **125**, 8843–8846.
- 161 L. Z. Yan and P. E. Dawson, *J. Am. Chem. Soc.*, 2001, **123**, 526–533.
- 162 X. F. Gao, J. J. Du, Z. Liu and J. Guo, *Org. Lett.*, 2016, **18**, 1166–1169.

- 163 Q. Wan and S. J. Danishefsky, *Angew. Chem., Int. Ed.*, 2007, **46**, 9248–9252.
- 164 R. C. Griffiths, F. R. Smith, D. Li, J. Wyatt, D. M. Rogers, J. E. Long, L. M. L. Cusin, P. J. Tighe, R. Layfield, J. D. Hirst, M. M. Müller and N. J. Mitchell, *Chem. - Eur. J.*, 2023, **29**, 1–8.
- 165 R. C. Griffiths, F. R. Smith, J. E. Long, D. Scott, H. E. L. Williams, N. J. Oldham, R. Layfield and N. J. Mitchell, *Angew. Chem., Int. Ed.*, 2022, **61**, 1–6.
- 166 R. C. Griffiths, F. R. Smith, J. E. Long, H. E. L. Williams, R. Layfield and N. J. Mitchell, *Angew. Chem., Int. Ed.*, 2020, **59**, 23659–23667.
- 167 K. Hanaya, J. Ohata, M. K. Miller, A. E. Mangubat-Medina, M. J. Swierczynski, D. C. Yang, R. M. Rosenthal, B. V. Popp and Z. T. Ball, *Chem. Comm.*, 2019, **55**, 2841–2844.
- 168 M. S. Oderinde, M. Frenette, D. W. Robbins, B. Aquila and J. W. Johannes, *J. Am. Chem. Soc.*, 2016, **138**, 1760–1763.
- 169 S. E. Collibee and J. Yu, *Tetrahedron Lett.*, 2005, **46**, 4453–4455.
- 170 J. H. T. Luong and S. K. Vashist, *ACS Omega*, 2020, **5**, 10–18.
- 171 D. B. McConnell, *J. Med. Chem.*, 2021, **64**, 16319–16327.
- 172 M. B. Sims, J. J. Lessard, L. Bai and B. S. Sumerlin, *Macromolecules*, 2018, **51**, 6380–6386.
- 173 J. B. Diccianni and T. Diao, *Trends Chem.*, 2019, **1**, 830–844.
- 174 J. Wang, T. Qin, T. G. Chen, L. Wimmer, J. T. Edwards, J. Cornella, B. Vokits, S. A. Shaw and P. S. Baran, *Angew. Chem., Int. Ed.*, 2016, **55**, 9676–9679.
- 175 D. Andreu, F. Albericio, N. A. Solé, M. C. Munson, M. Ferrer and G. Barany, *Peptide Synthesis Protocols*, Humana Press, Totowa, 1st edn., 1994, vol. 35.
- 176 M. Karimi, M. T. Ignasiak, B. Chan, A. K. Croft, L. Radom, C. H. Schiesser, D. I. Pattison and M. J. Davies, *Sci. Rep.*, 2016, **6**, 1–12.

- 177 Entropy, <https://www.britannica.com/science/entropy-physics>., (accessed 5 June 2022).
- 178 M. J. Feige and L. M. Hendershot, *Curr. Opin. Cell Biol.*, 2011, **23**, 167–175.
- 179 M. A. Clark, J. Choi and M. Douglas, *Biology*, OpenStax, Houston, 2nd edn., 2018, vol. 1.
- 180 H. Sies and D. P. Jones, *Nat. Rev. Mol. Cell. Biol.*, 2020, **21**, 363–383.
- 181 B. M. Babior, *Am. J. Med.*, 2000, **109**, 33–44.
- 182 W. K. Alderton, C. E. Cooper and R. G. Knowles, *Biochem. J.*, 2001, **357**, 593–615.
- 183 M. Brunori, A. Giuffrè, P. Sarti, G. Stubauer and M. T. Wilson, *Cell. Mol. Life Sci.*, 1999, **56**, 549–557.
- 184 R. Radi, G. Peluffo, M. N. Alvarez, M. Naviliat and A. Cayota, *Free Radic. Biol. Med.*, 2001, **30**, 463–488.
- 185 S. J. Klebanoff, A. J. Kettle, H. Rosen, C. C. Winterbourn and W. M. Nauseef, *J. Leukoc. Biol.*, 2013, **93**, 185–198.
- 186 W. J. Wedemeyer, E. Welker, M. Narayan and H. A. Scheraga, *Biochemistry*, 2000, **39**, 4207–4216.
- 187 C. Wiedemann, A. Kumar, A. Lang and O. Ohlenschläger, *Front. Chem.*, 2020, **8**, 1–8.
- 188 T. J. Bechtel and E. Weerapana, *Proteomics*, 2017, **17**, 1–21.
- 189 T. N. Vinther, M. Norrman, U. Ribel, K. Huus, M. Schlein, D. B. Steensgaard, T. A. ° Pedersen, I. Pettersson, S. Ludvigsen, T. Kjeldsen, K. J. Jensen and F. Ek Hub Alek, *Protein Sci.*, 2012, **22**, 296–305.
- 190 M. Karimi, B. Crosset, S. J. Cordwell, D. I. Pattison and M. J. Davies, *Free Radic. Biol. Med.*, 2020, **154**, 62–74.

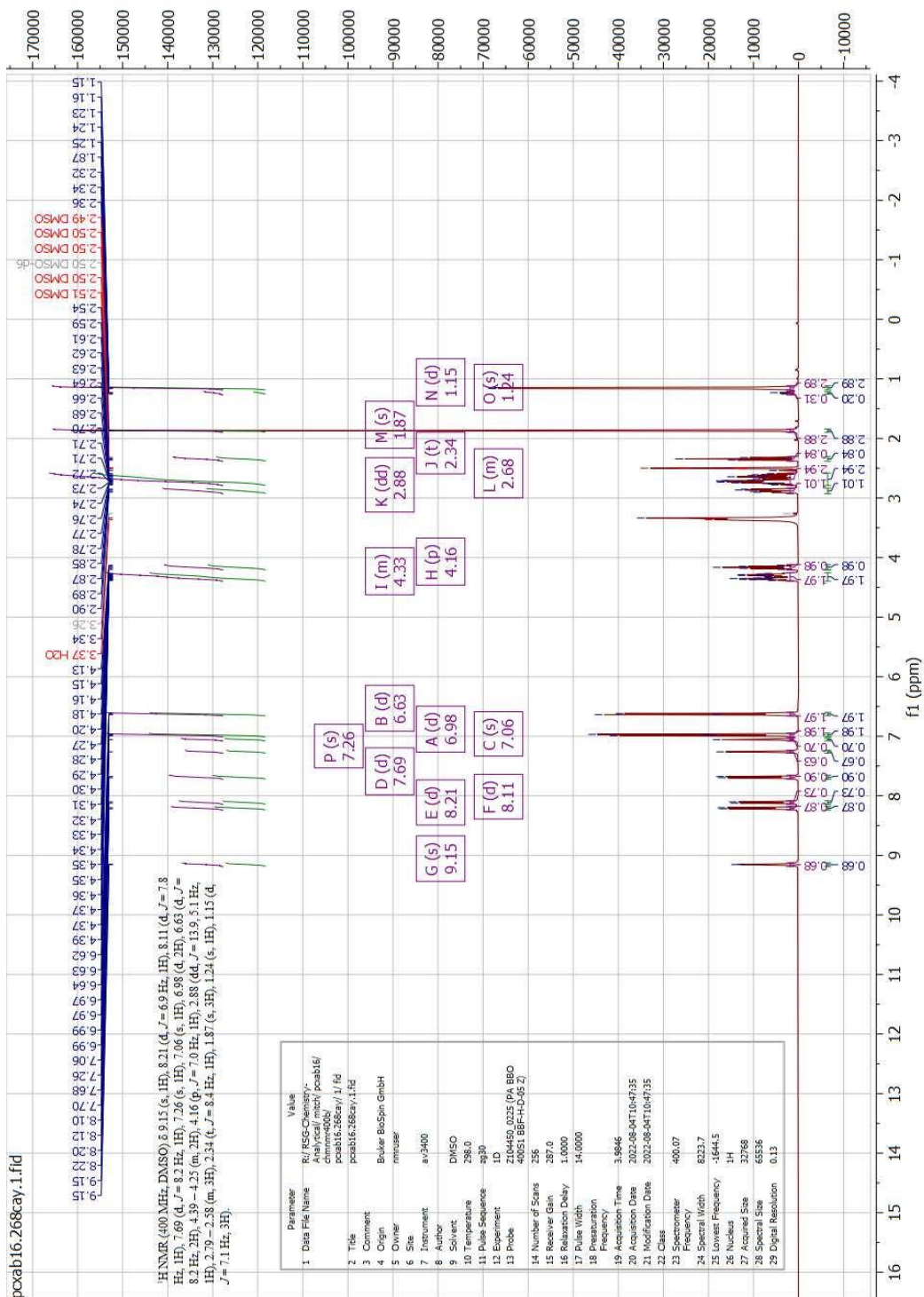


- 191 I. T. Jolliffe and J. Cadima, *Philos. Trans. R. Soc., A*, 2016, **374**, 1–16.
- 192 C. L. Hawkins, *Essays Biochem.*, 2021, **64**, 75–86.
- 193 J. Tirado-Rives and W. L. Jorgensen, *J. Phys. Chem. A*, 2019, **123**, 5713–5717.
- 194 T. Clark, M. Hennemann, J. S. Murray and P. Politzer, *J. Mol. Model.*, 2007, **13**, 291–296.
- 195 T. Nybo, S. Dieterich, L. F. Gamon, C. Y. Chuang, A. Hammer, G. Hoefler, E. Malle, A. Rogowska-Wrzesinska and M. J. Davies, *Redox Biol.*, 2019, **20**, 496–513.
- 196 W. Stephanie Da, C. Nizer, V. Inkovskiy and J. Overhage, *Microorganisms*, 2020, **8**, 1–27.
- 197 A. Radzicka and R. Wolfenden, *J. Am. Chem. Soc.*, 1996, **118**, 6105–6109.
- 198 I. Serra, I. García Rubio and S. Van Doorslaer, *Appl. Magn. Reson.*, 2021, **53**, 1105–1119.
- 199 T. Koide, H. Itoh, A. Otaka, H. Yasui, M. Kuroda, N. Esaki, N. Soda and N. Fujii, *Chem. Pharm. Bull.*, 1993, **41**, 502–506.
- 200 R. J. Hondal and R. T. Raines, *Methods Enzymol.*, 2002, **347**, 70–83.
- 201 M. J. Holdsworth and D. J. Gibbs, *Curr. Biol.*, 2020, **30**, R362–R369.
- 202 S. Vilà, C. Camõ, E. Figueras, E. Badosa, E. Montesinos, M. Planas and L. Feliu, *European J. Org. Chem.*, 2014, **2014**, 4785–4794.
- 203 G. Tong, J. M. Lawlor, G. W. Tregear and J. Haralambidis, *J. Org. Chem.*, 1993, **58**, 2223–2231.
- 204 N. Corpuz and J. P. Schwans, *Bioorg. Med. Chem. Lett.*, 2017, **27**, 2410–2414.
- 205 W. M. Hussein, T. Y. Liu, P. Maruthayanar, S. Mukaida, P. M. Moyle, J. W. Wells, I. Toth and M. Skwarczynski, *Chem. Sci.*, 2016, **7**, 2308–2321.

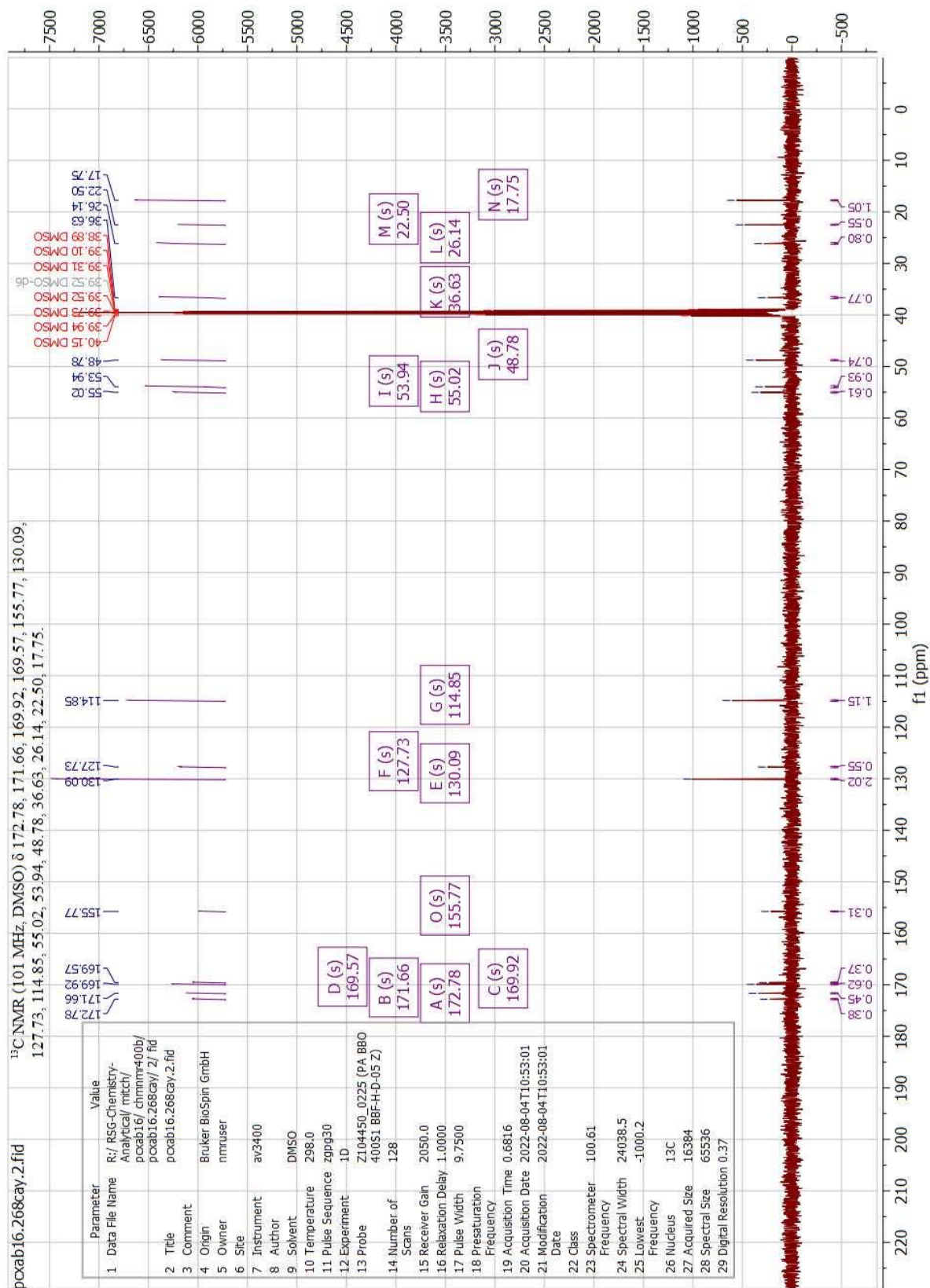
# APPENDIX 1

## Ac-CAY-NH<sub>2</sub> (1)

### <sup>1</sup>H-NMR

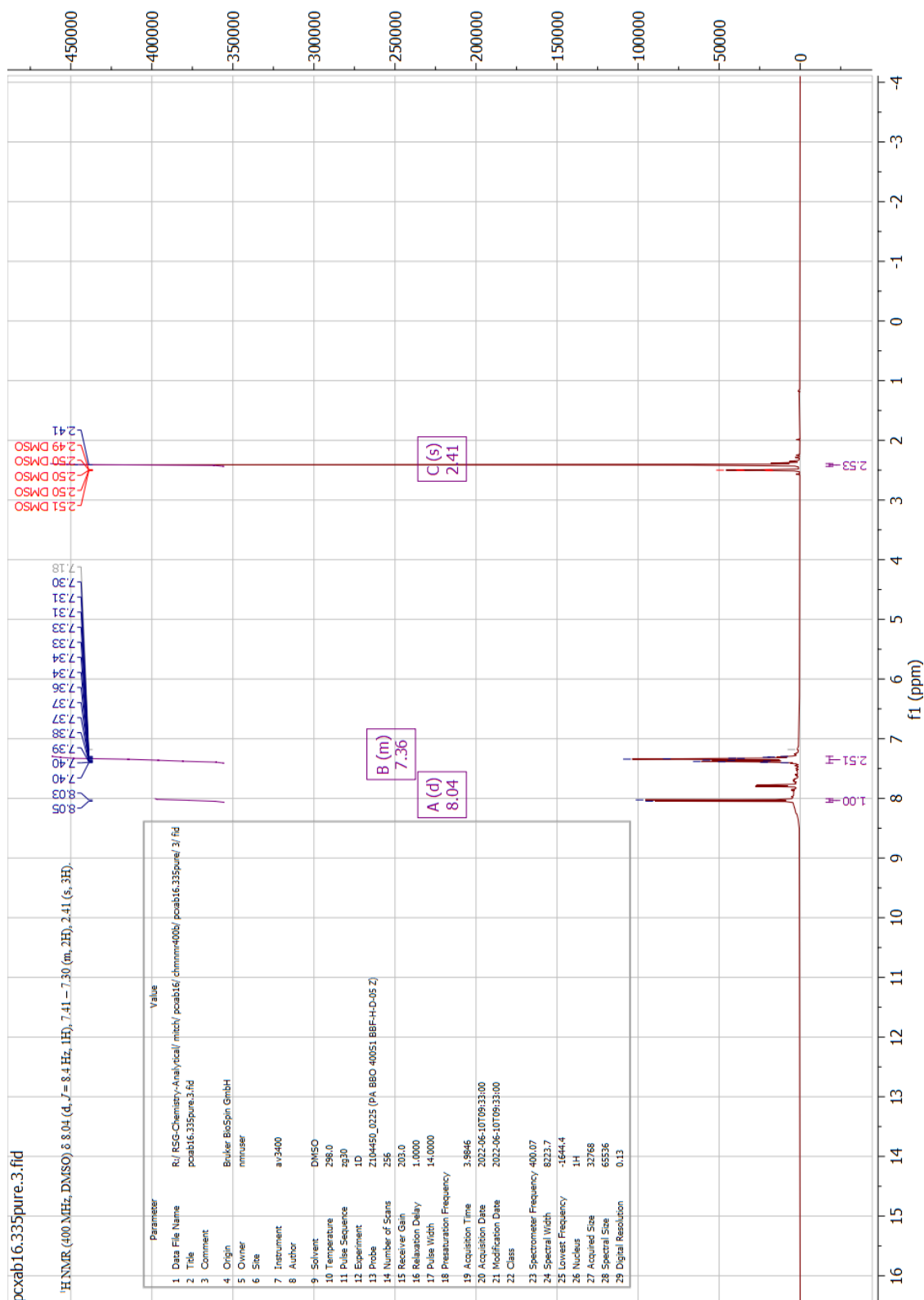


<sup>13</sup>C-NMR

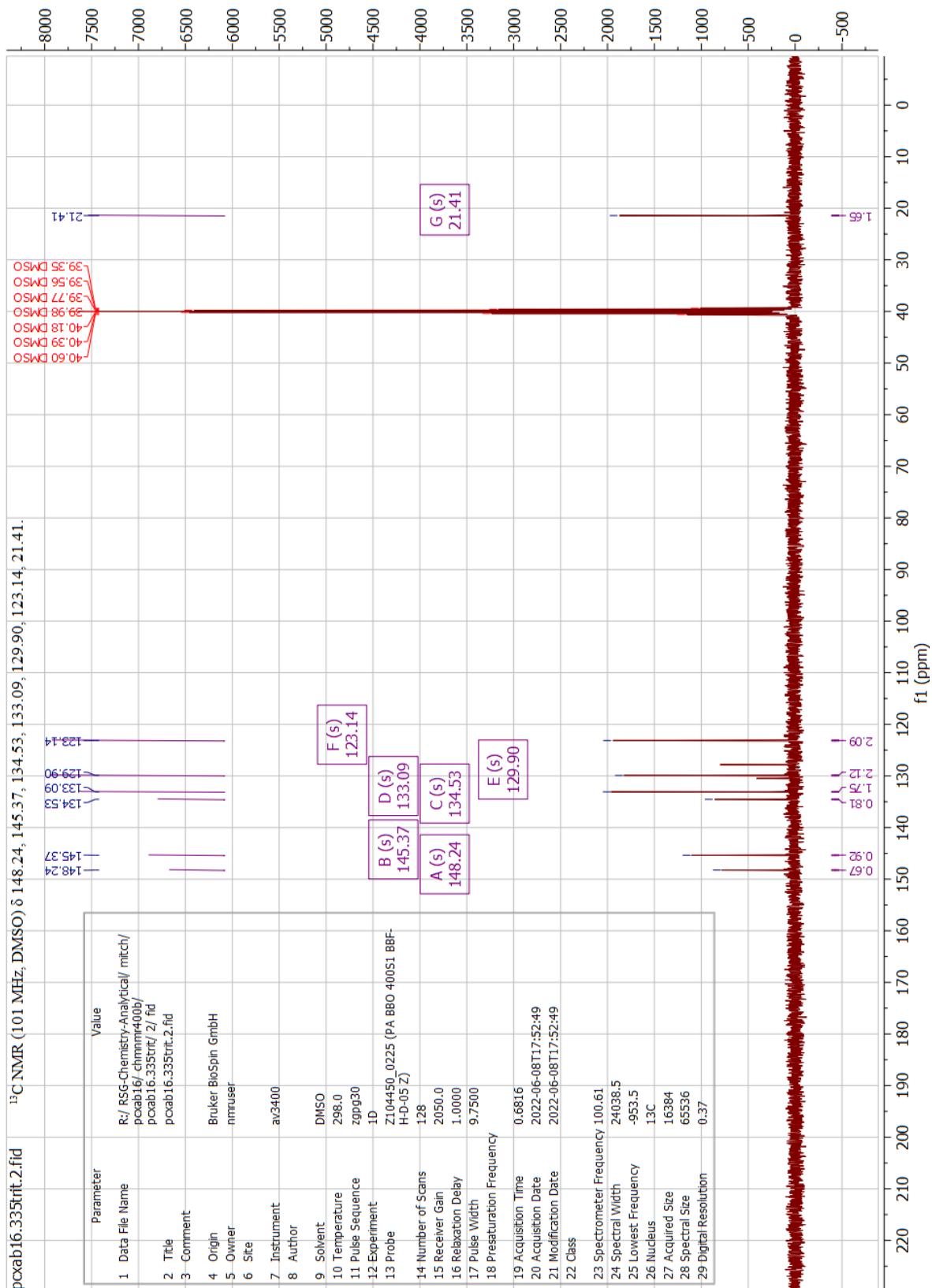


# (5-methyl-2-nitrophenyl)boronic acid (10)

## <sup>1</sup>H-NMR

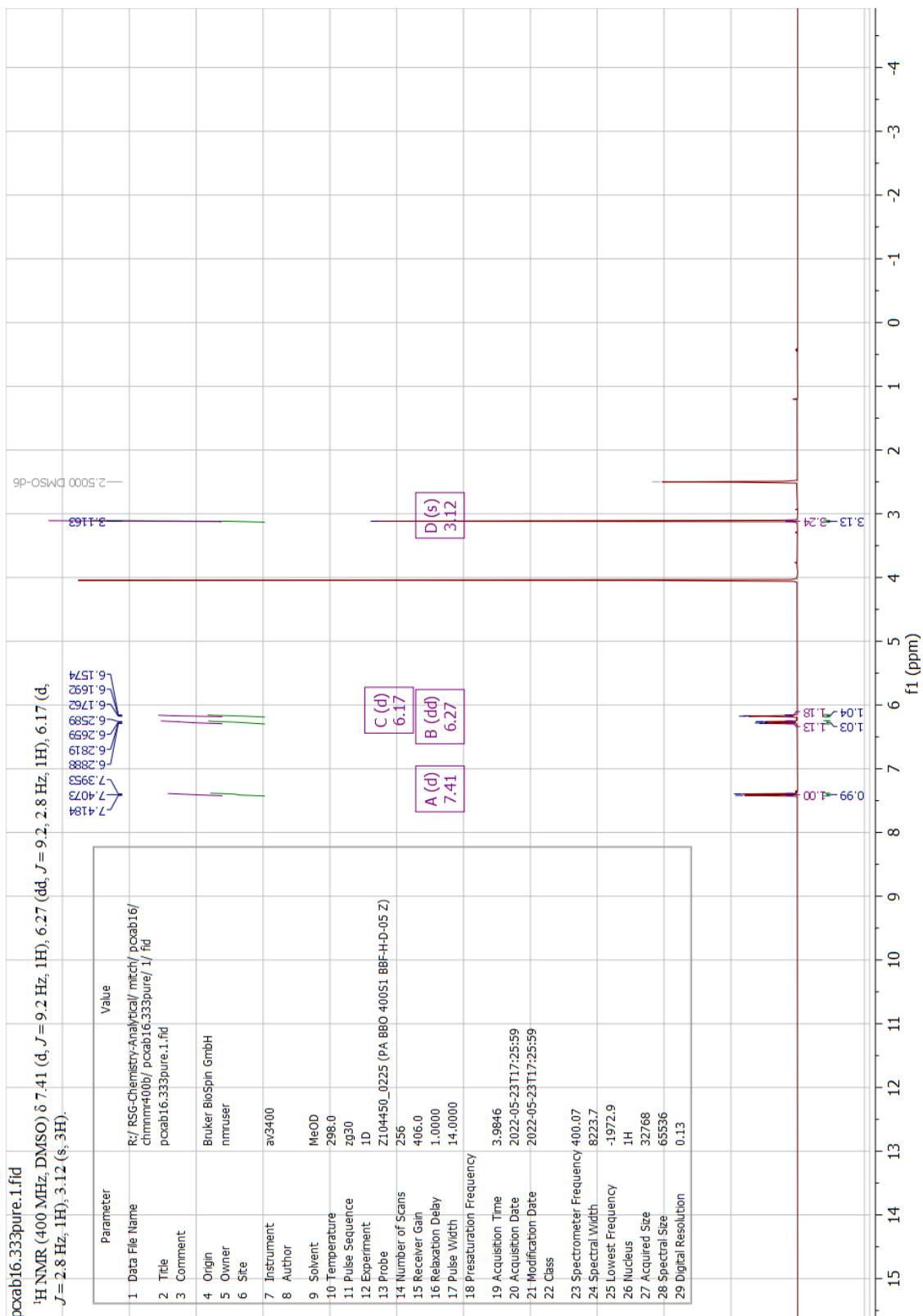


<sup>13</sup>C-NMR

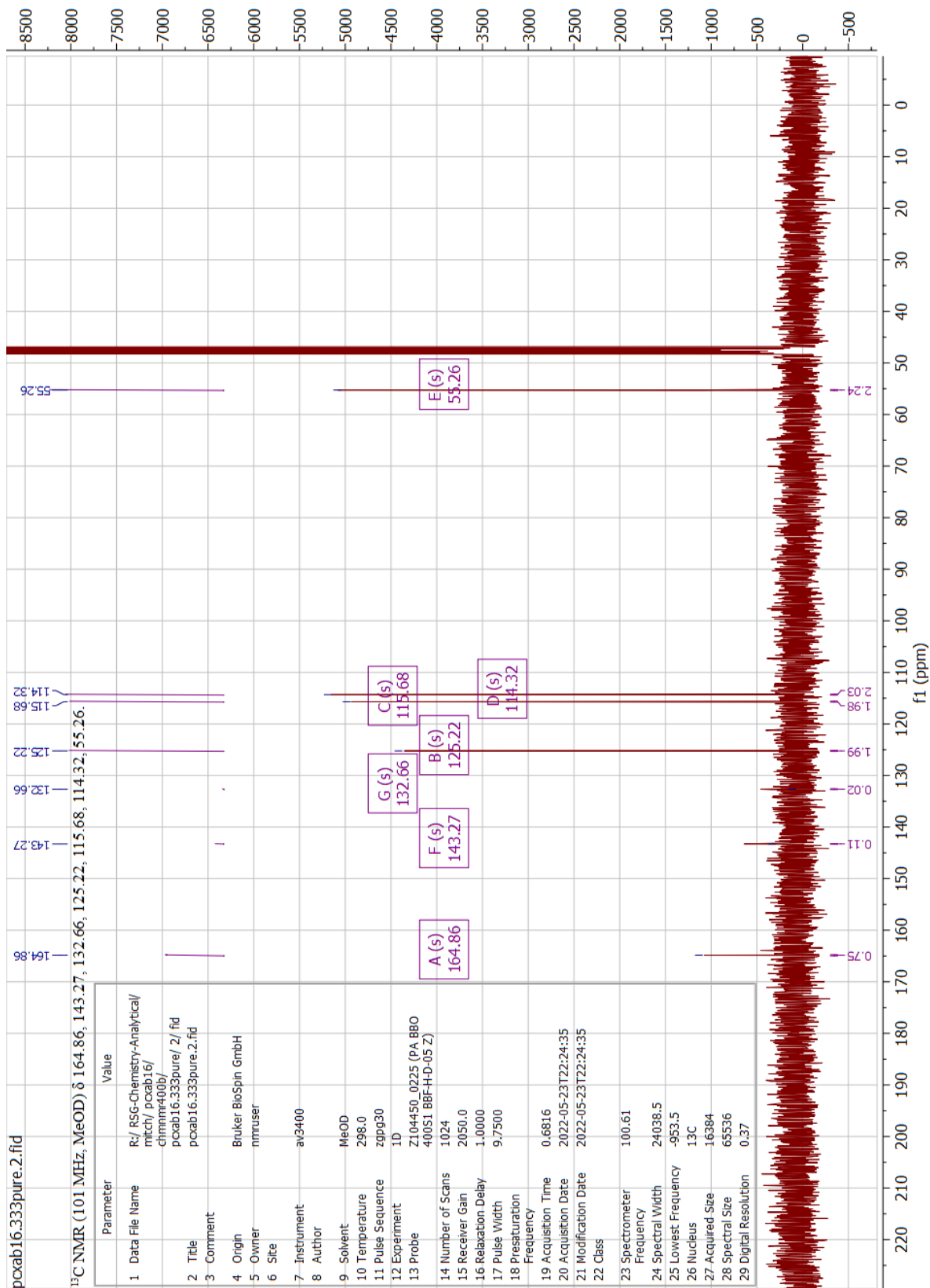


# (5-methoxy-2-nitrophenyl)boronic acid (12)

## <sup>1</sup>H-NMR

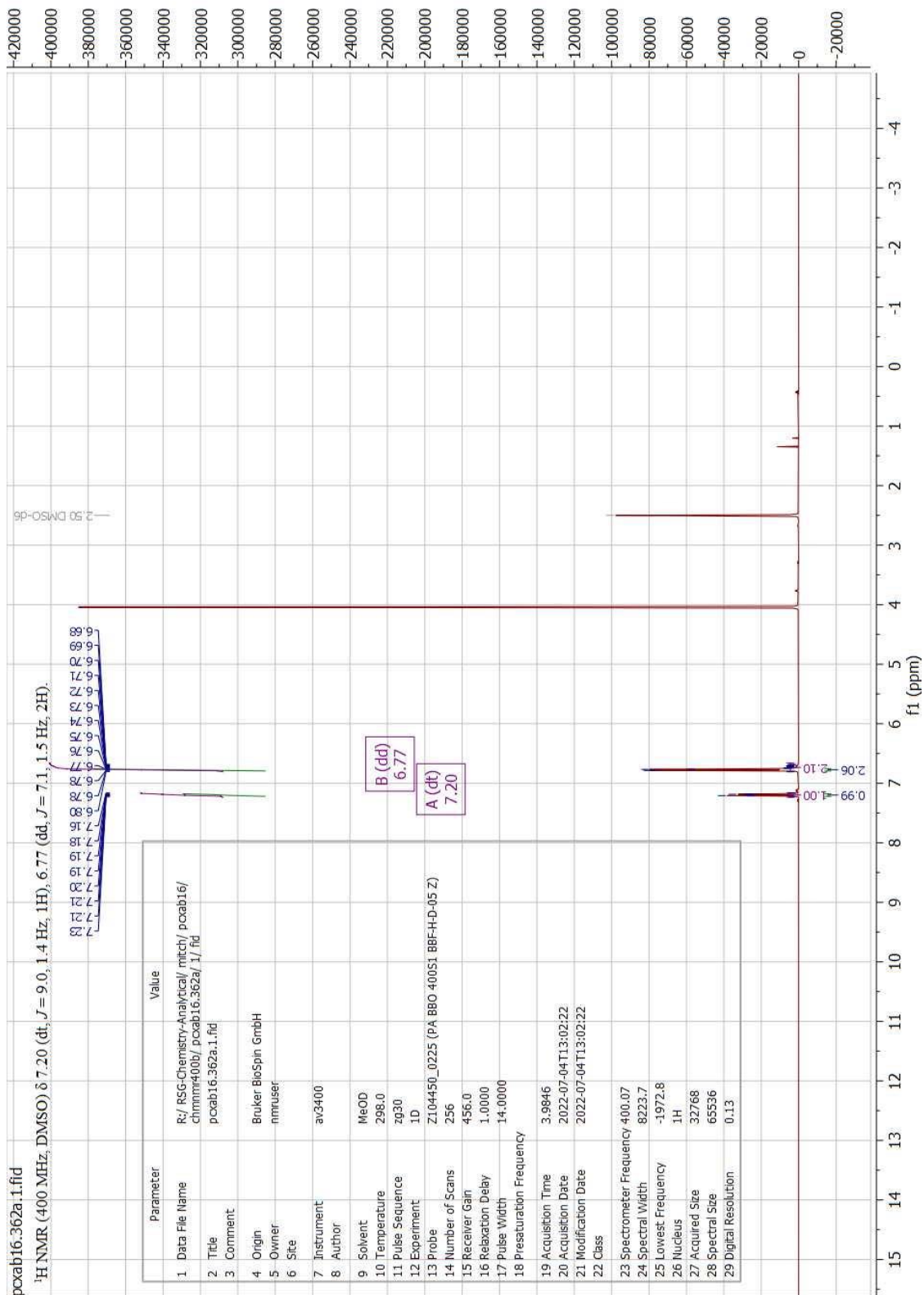


# <sup>13</sup>C-NMR



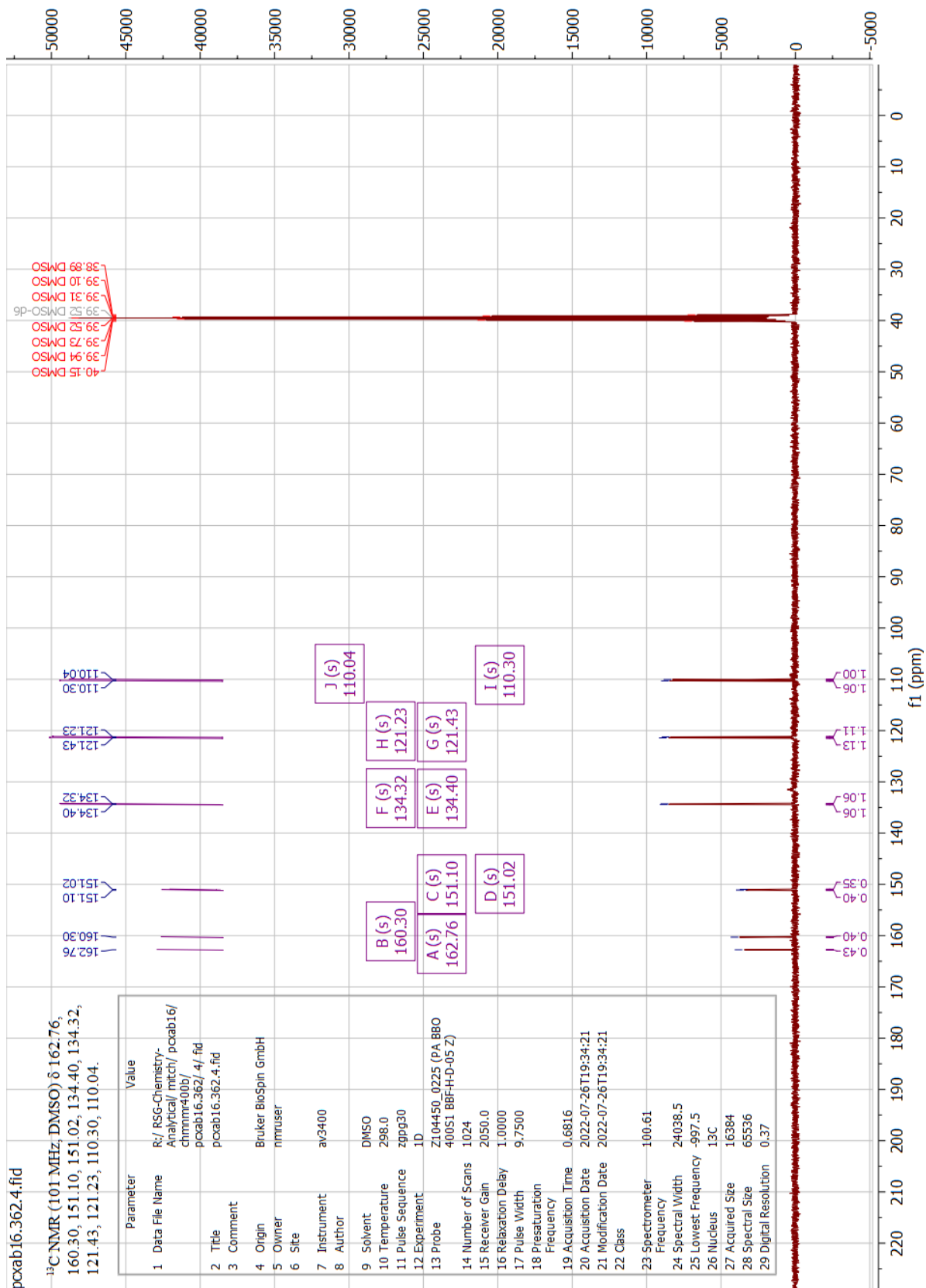
# (4-fluoro-2-nitrophenyl)boronic acid (14)

## <sup>1</sup>H-NMR

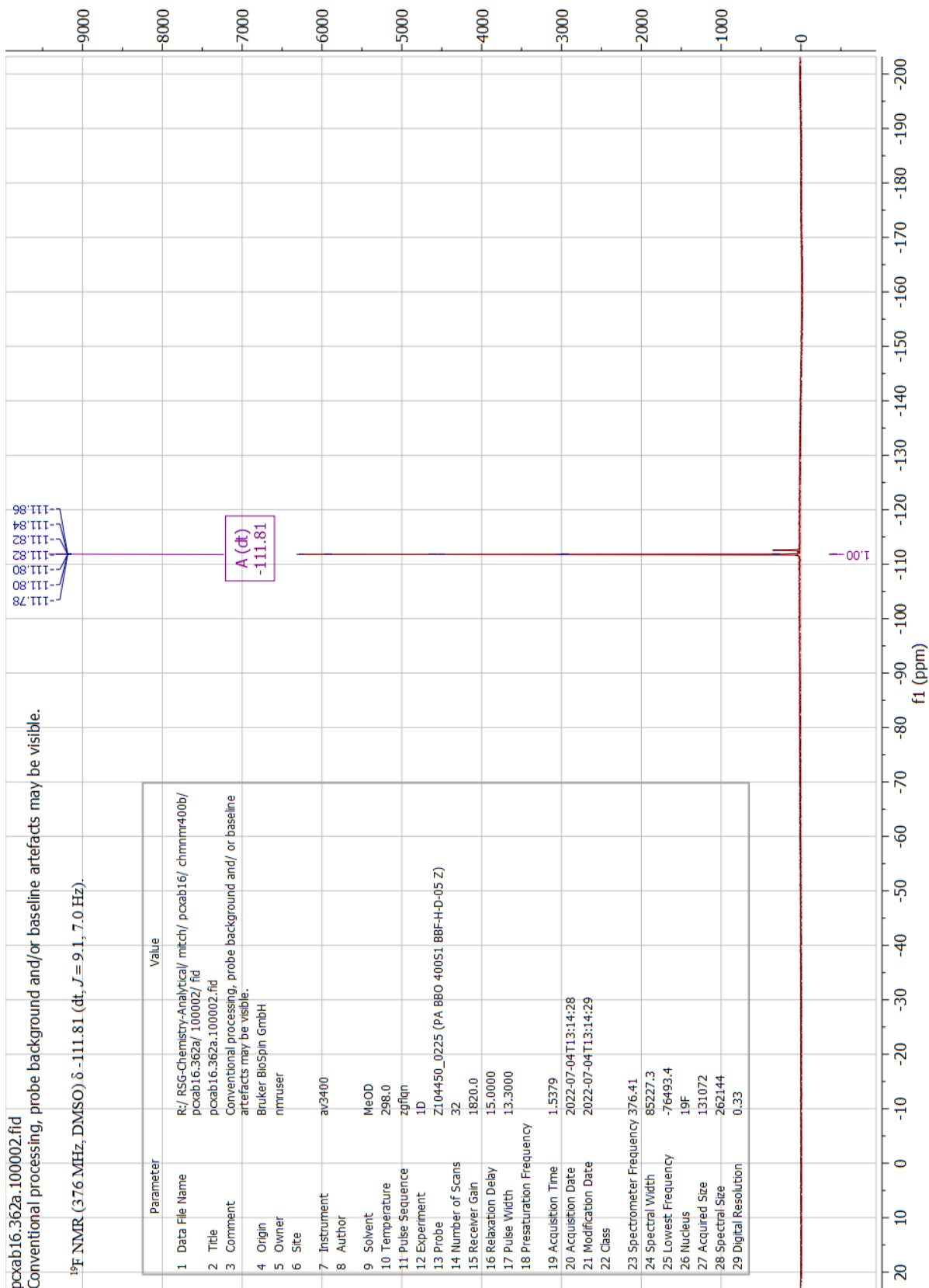




<sup>13</sup>C-NMR

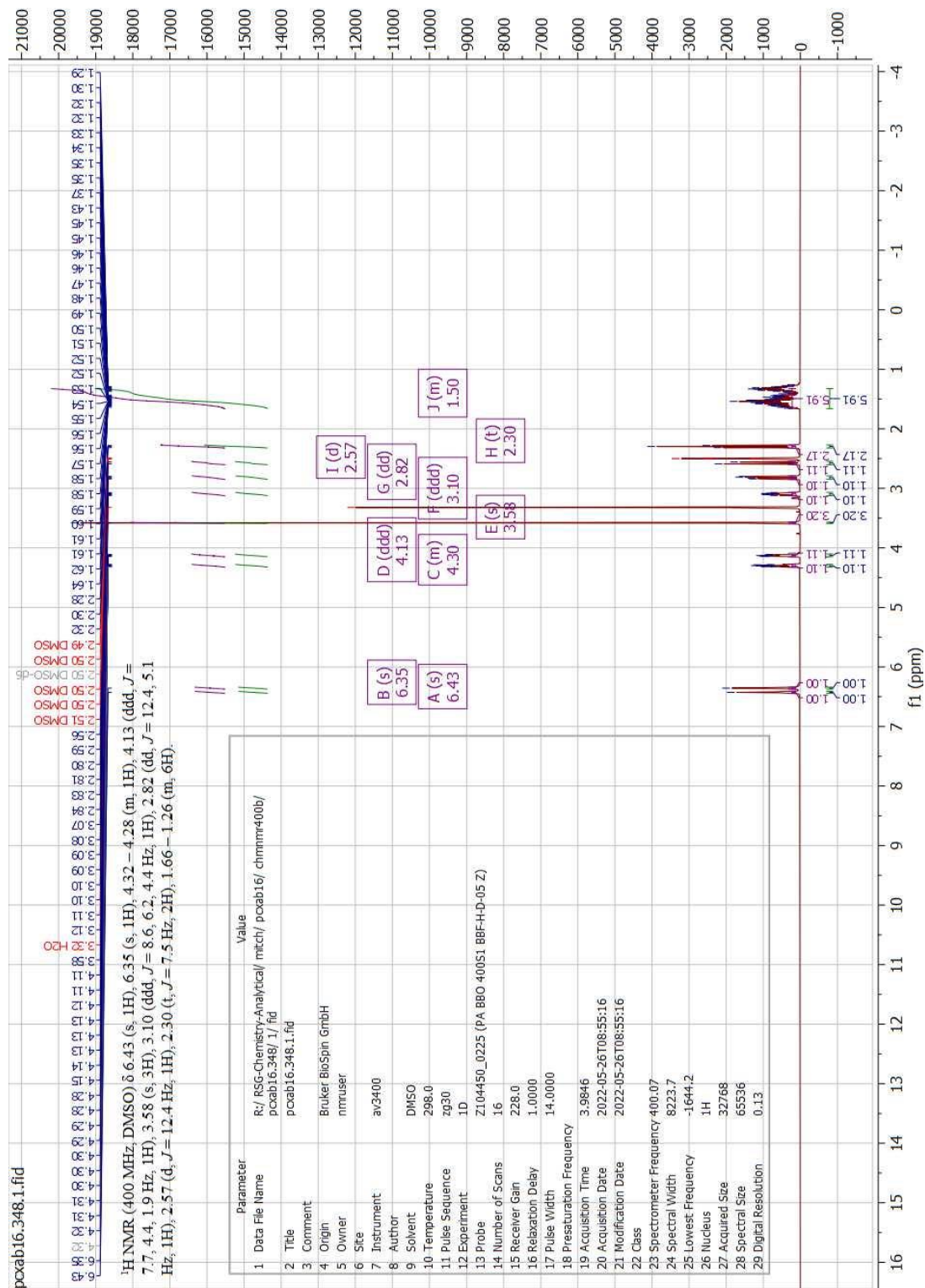


**<sup>19</sup>F-NMR**

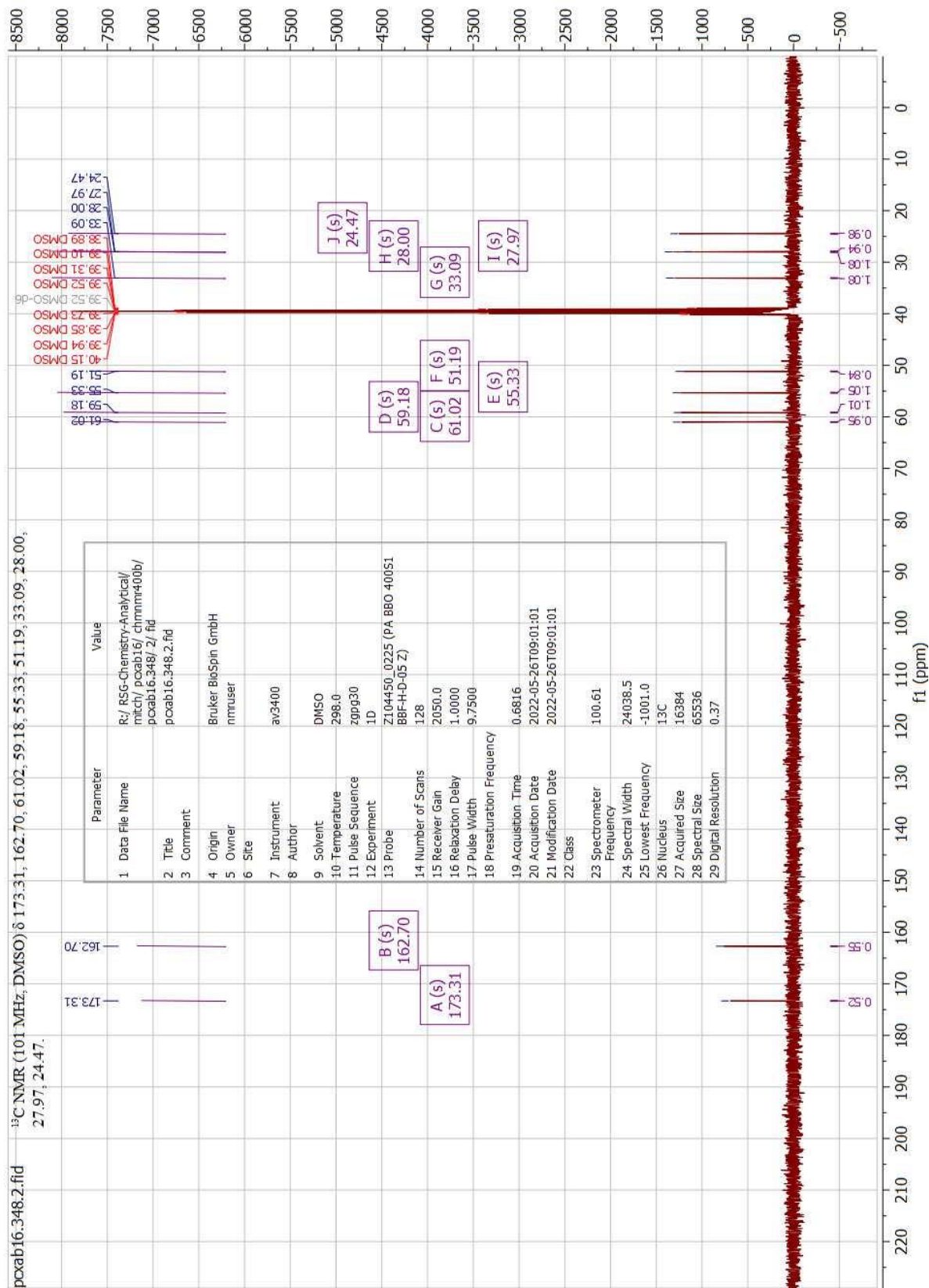


# Synthesis of methyl 5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoate (17)

## <sup>1</sup>H-NMR

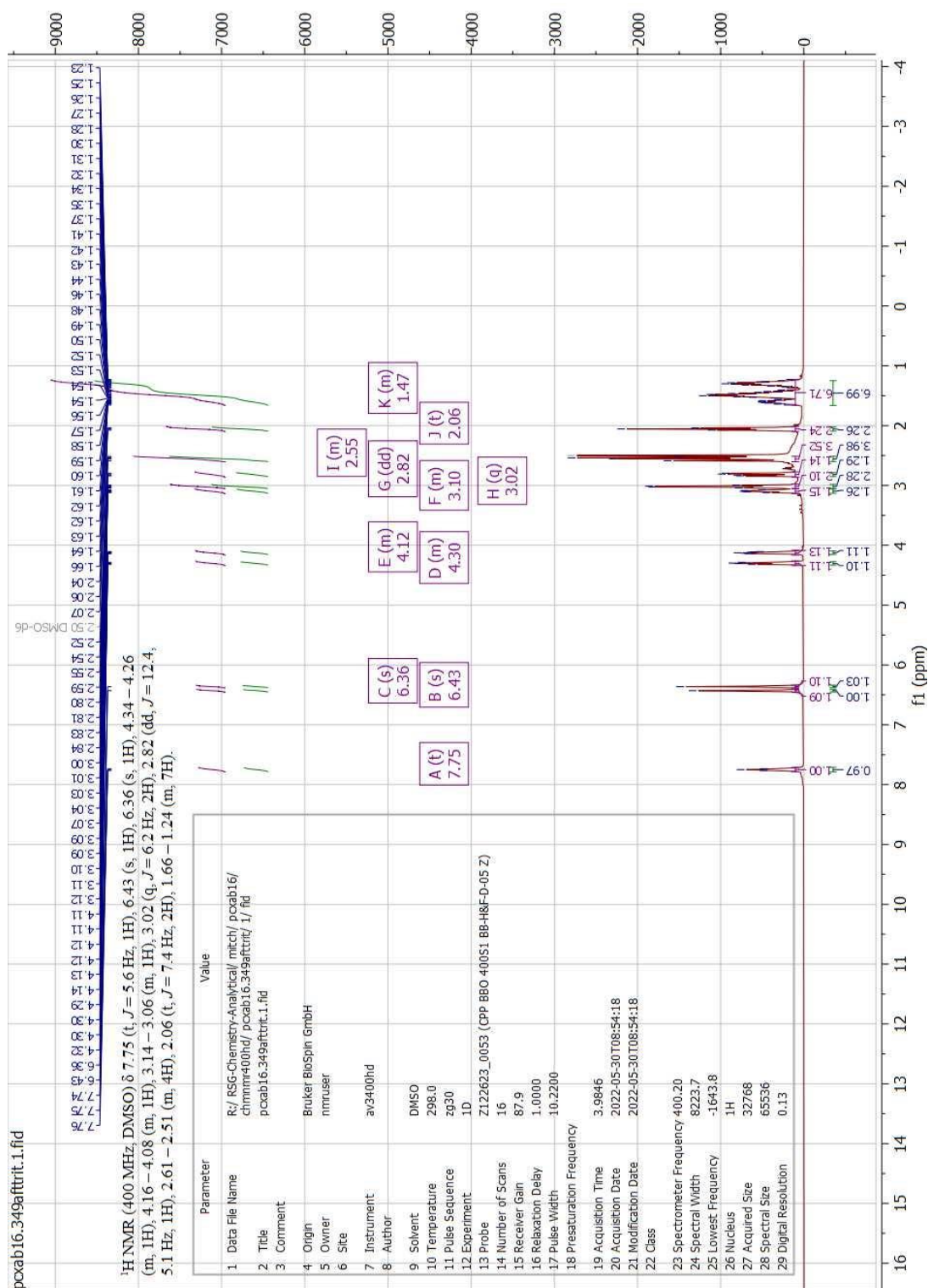


<sup>13</sup>C-NMR

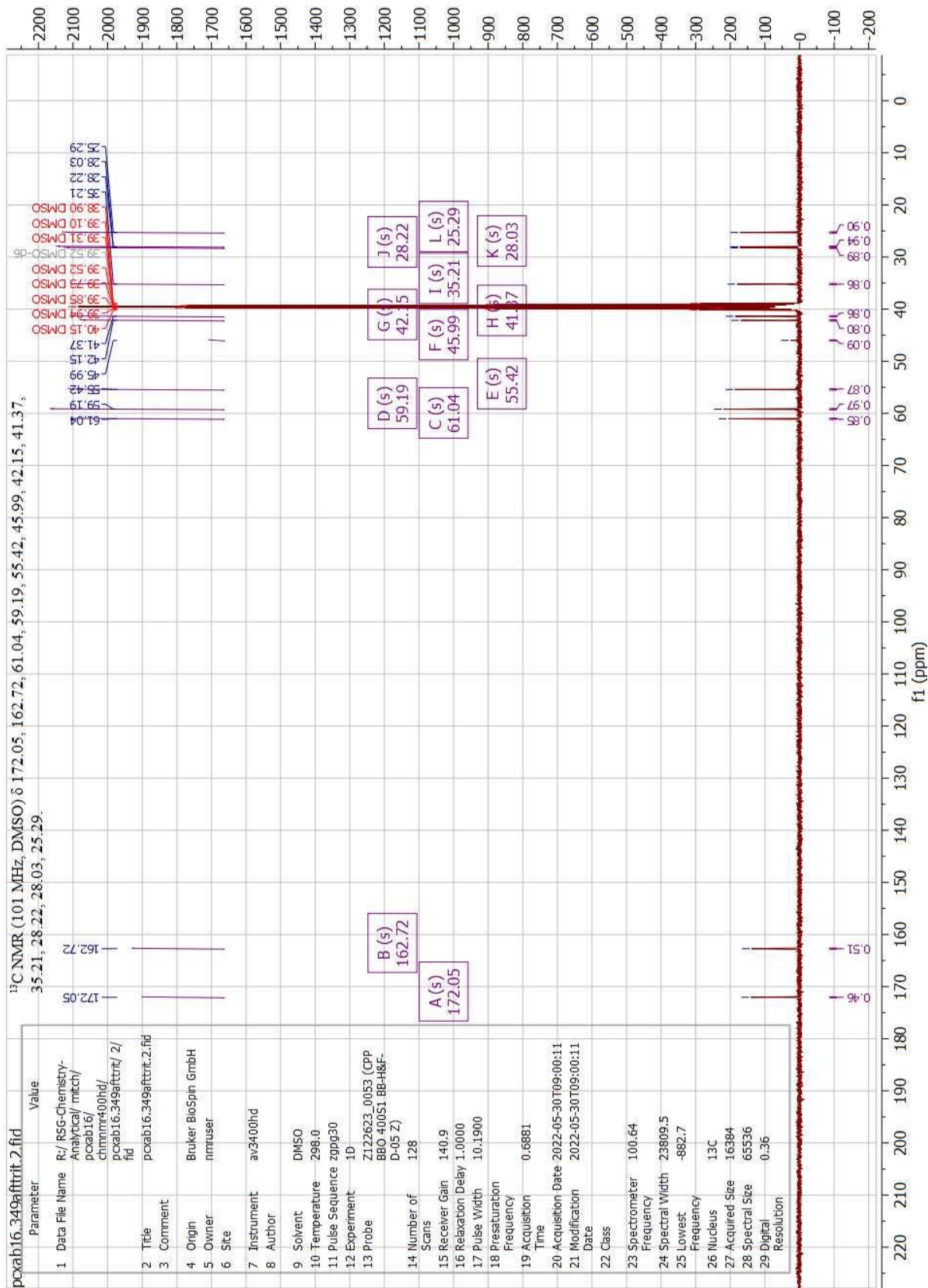


# Synthesis of N-(2-aminoethyl)-5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (18)

## <sup>1</sup>H-NMR

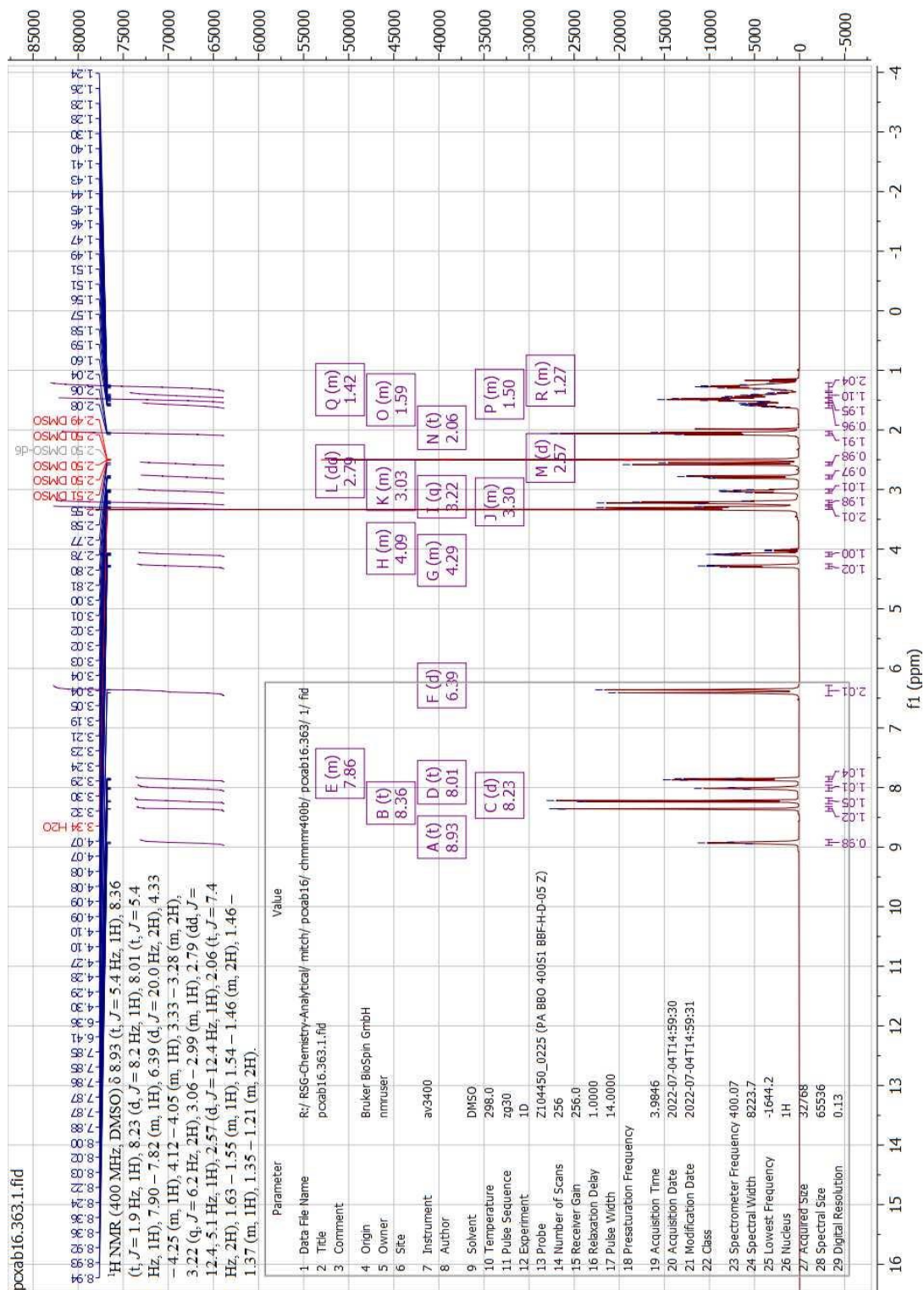


**<sup>13</sup>C-NMR**

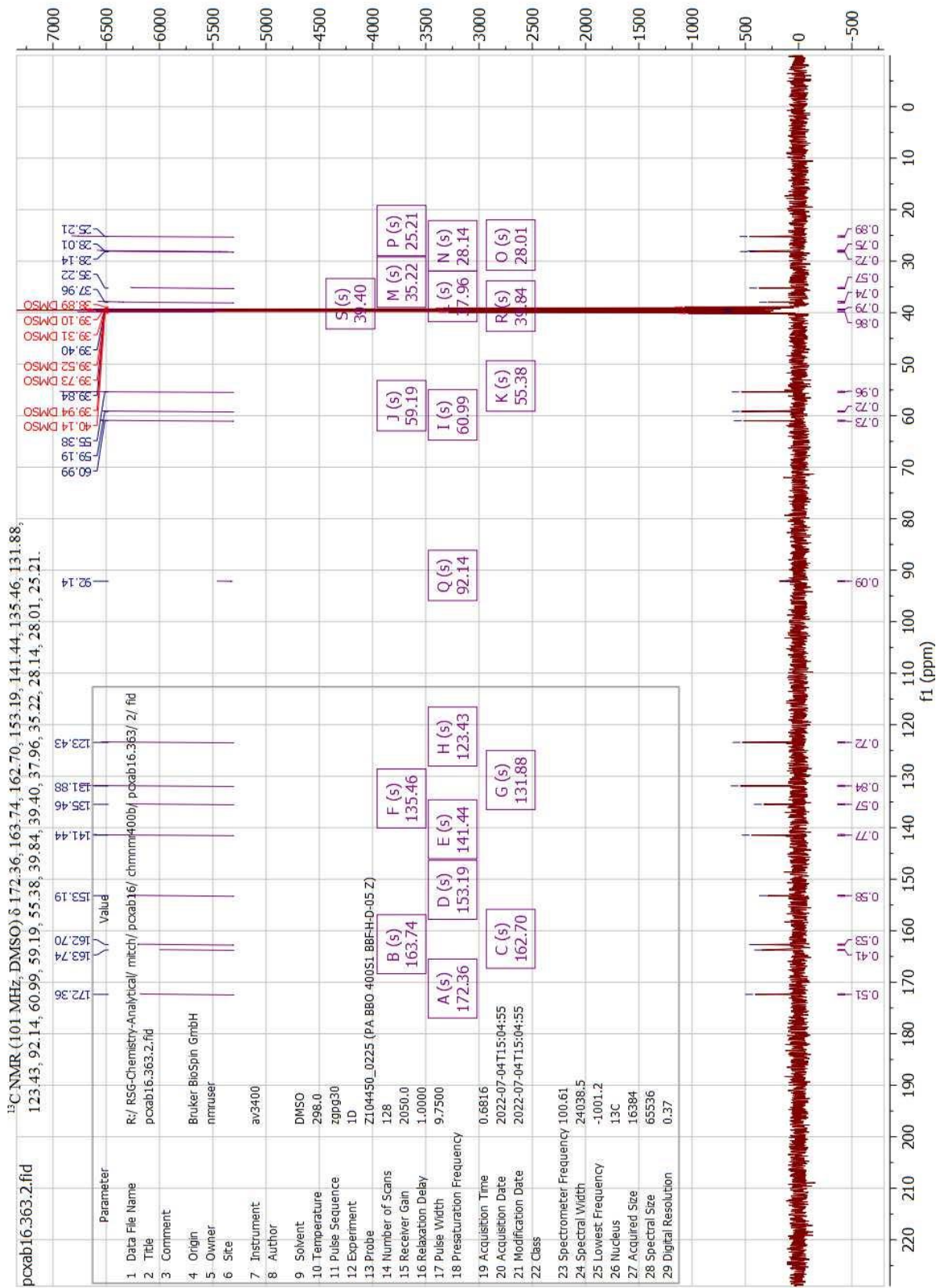


# Synthesis of 4-iodo-3-nitro-N-(2-(5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamido)ethyl)benzamide (22)

## <sup>1</sup>H-NMR



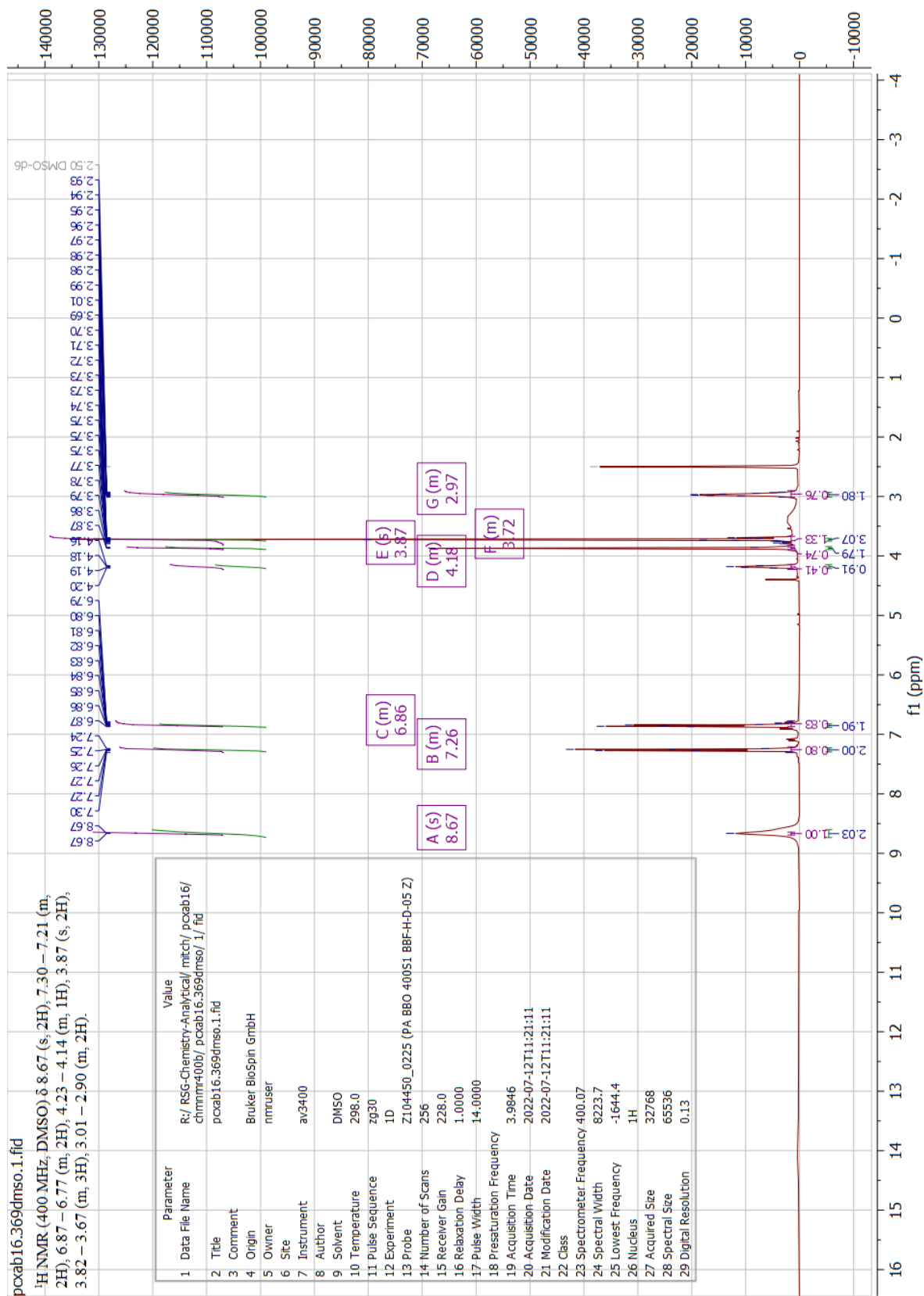
**<sup>13</sup>C-NMR**



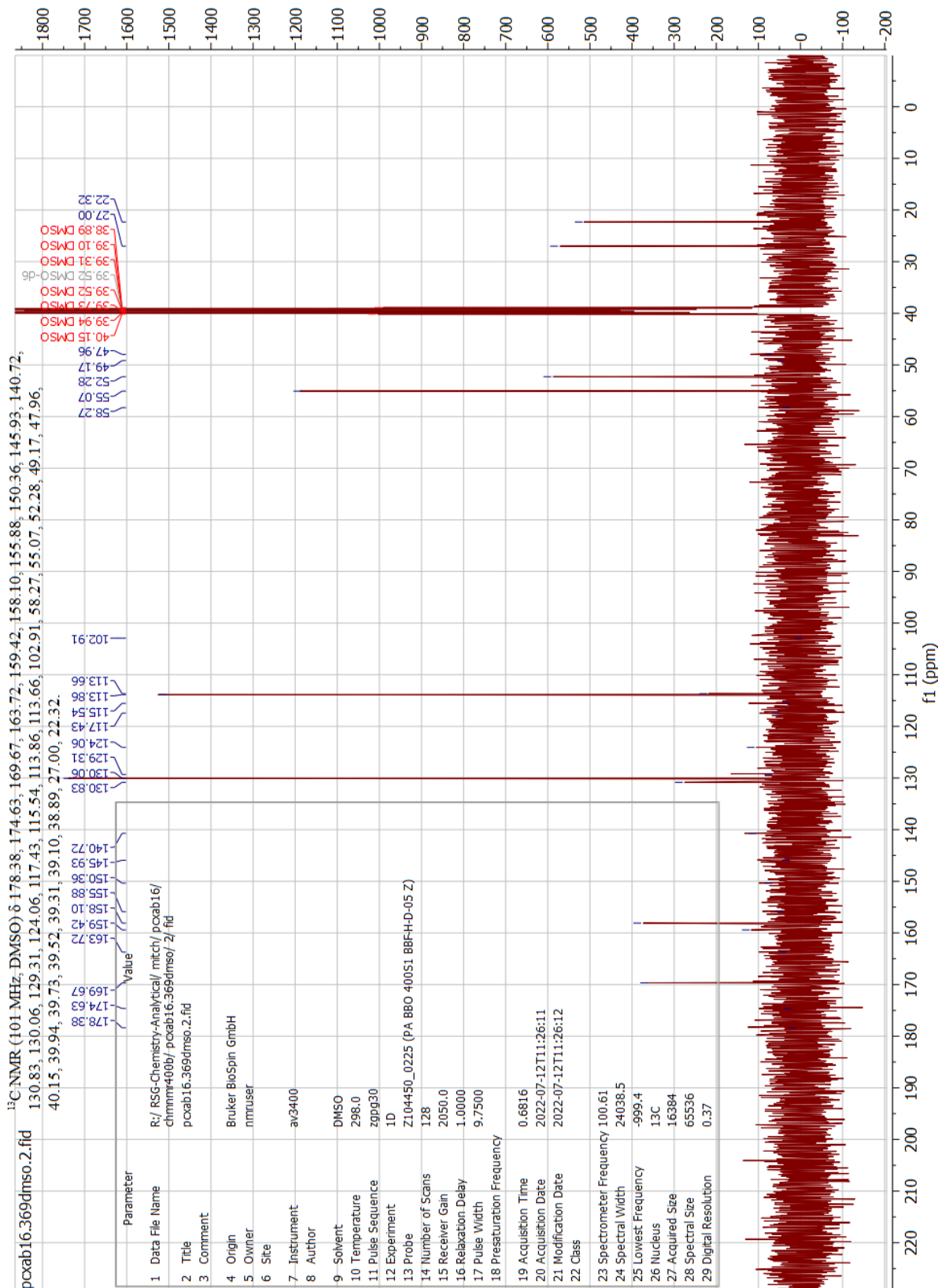


# (R)-2-amino-3-((4-methoxybenzyl)selenyl)propanoic acid (48)

## <sup>1</sup>H-NMR

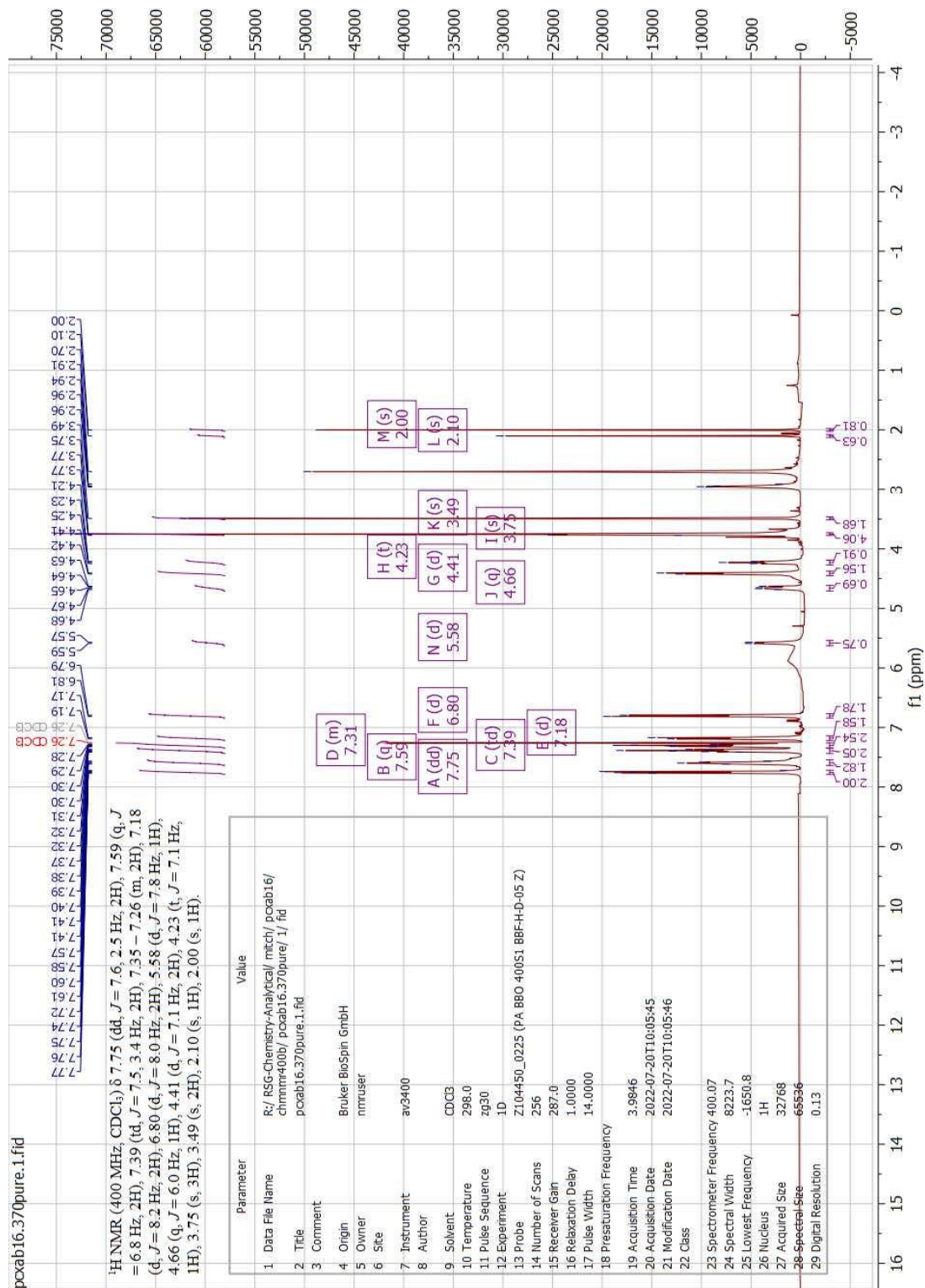


<sup>13</sup>C-NMR

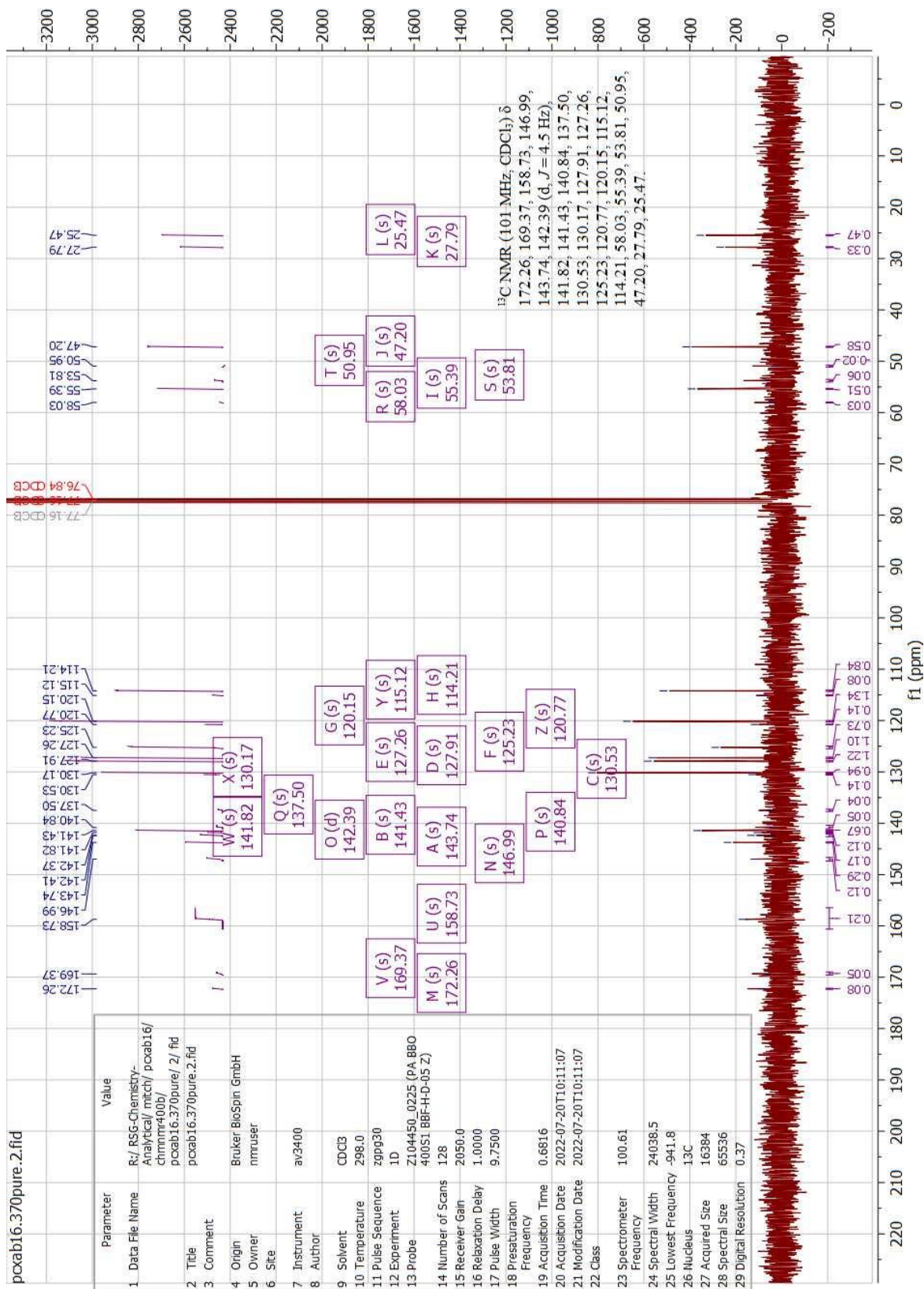


# Synthesis of (R)-2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-((4-methoxybenzyl)selenyl)propanoic acid (49)

## <sup>1</sup>H-NMR



<sup>13</sup>C-NMR



## APPENDIX 2

This appendix describes efforts focused on the synthesis of three peptides for a collaboration with Prof. Michael Holdsworth and Prof. Neil Oldham from the University of Nottingham. The focus of this research sought to better understand the enzymatic activity that governs oxygen-sensing in plants. It was hypothesized that specific enzymes in plants can add oxygen to the sulfinic acid moiety at the *N*-terminus Cys residue of certain proteins to trigger its arginylation and subsequent proteasomal degradation.<sup>201</sup> This process is thought to play a key role in the way plants regulate the concentration of oxygen and therefore respond in high or low levels of oxygen in the organism. To explore this oxygen sensing pathway, the synthesised peptides will be exposed to a plant extract containing the enzymes that add oxygen to the sulfinic acid moiety at the *N*-terminus of the synthetic sequence.

The first peptide to be synthesised was H<sub>2</sub>N-CGGAIISDYAPLVTK(Biotin)-OH (**45**, Figure 173). The synthesis for building block **51** used to afford **45** is described later in the text of APPENDIX 2. This sequence was required as a control, to investigate whether the enzymes in the plant extract would add oxygen on a Cys residue, even if the latter would not be in the oxidised form of a sulfinic acid.

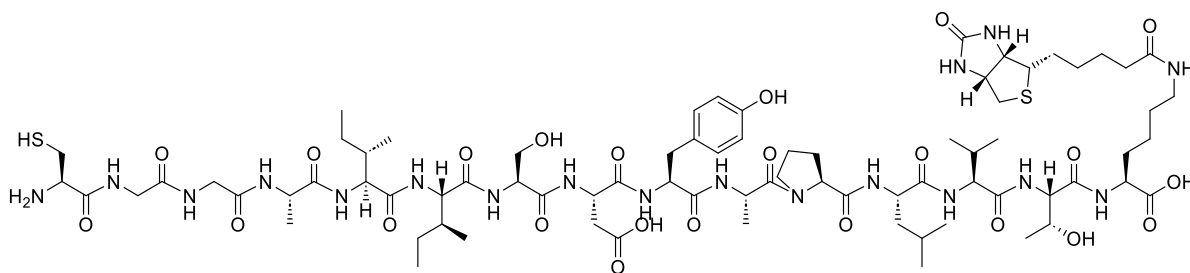


Figure 173 – Peptide sequence of **45**.

The second peptide to be synthesised was H<sub>2</sub>N-C(SO<sub>2</sub>)GGAIISDYAPLVTK(Biotin)-OH (**46**, Figure 174). This peptide sequence is the representation of the protein substrate on which the enzymes presented in the plant extract are meant to add oxygen to trigger the arginylation of the oxidised *N*-terminus.

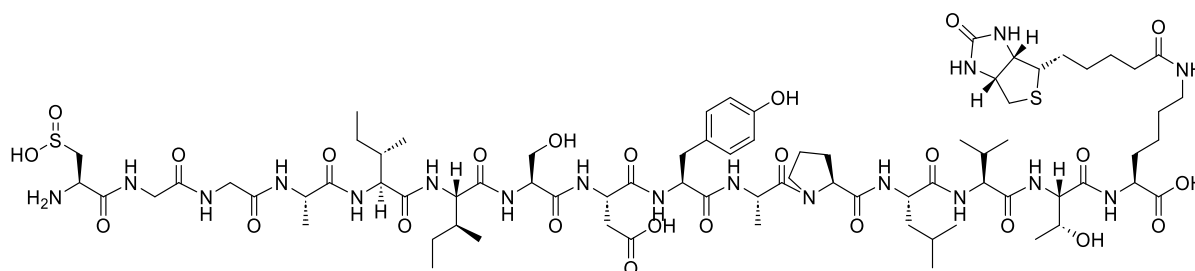


Figure 174 – Peptide sequence of **46**.

Peptide sequence **46** presents a sulfonic acid moiety on the *N*-terminus. The synthesis of building blocks **51** and **54** used for the synthesis of modified peptide **46** is described later in the text of APPENDIX 2. After cleavage of the peptide from the resin (see 8.1 GENERAL METHODS for more details), the sulfone moiety was reduced to the sulfonic acid by treatment with sodium borohydride (see 8.2 GENERAL PROTOCOLS for more details).

Finally, peptide  $\text{H}_2\text{N-RC}(\text{SO}_2)\text{GGAIISDYAPLVTK}(\text{Biotin})\text{-OH}$  (**47**, Figure 175) was synthesised as a further control. In this case, the control experiment would be to investigate whether the plant extract adds oxygen to the peptide even if the Arg residue is already present at the *N*-terminus.

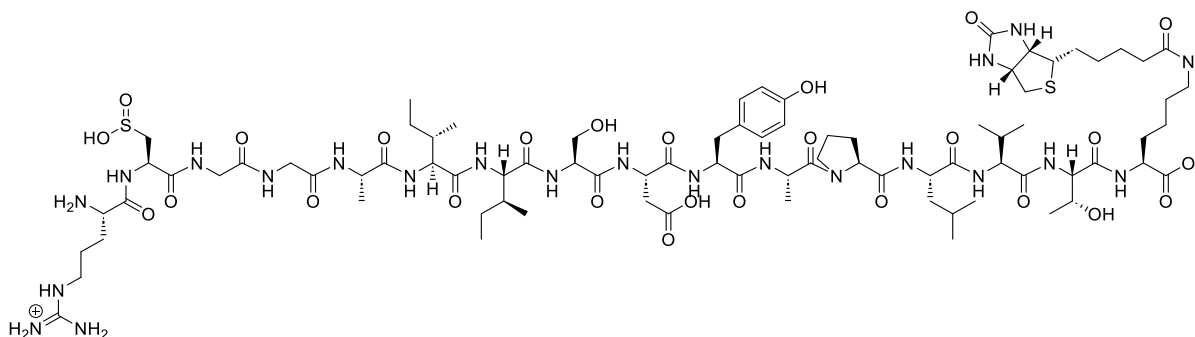


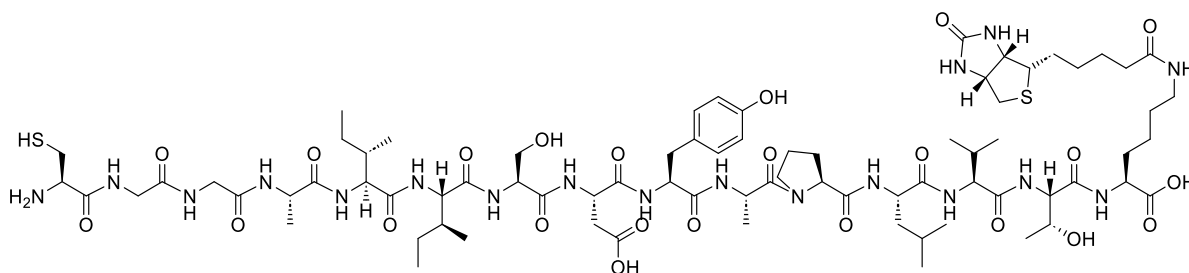
Figure 175 – Peptide sequence of **47**.

Again, the synthesis for building blocks **51** and **54** used for achieving modified peptide **47** is described in 8.4 COMPOUNDS SYNTHESIS. All three peptides presented a biotinylated Lys residue at the *C*-terminus so that the sequence could be bound to streptavidin magnetic beads. Streptavidin binds to biotin with high affinity and this will allow the peptide to be easily isolated from the plant extract for characterisation.

At the time of writing, the pull-down assay to purify the peptides from the plant extract is on-going. Once this stage is optimised, MS techniques will be implemented to fully

characterise the peptides and to understand if the expected arinylation or other modification were triggered by the enzymatic extract.

## Synthesis of H<sub>2</sub>N-CGGAIISDYAPLVTK(Biotin)-OH (**45**)



Peptide **45** was synthesised by manually loading the first amino acid (**51**) following general manual coupling protocol on CTC resin (0.2 mmol). The rest of the peptide sequence was synthesised using general automated synthesiser protocols with microwave assistance. The crude peptide was purified by semi-preparative RP-HPLC (5-60% B over 30 min) and lyophilised to produce the desired peptide (**45** 196 mg, 0.113 mmol, 57% yield).

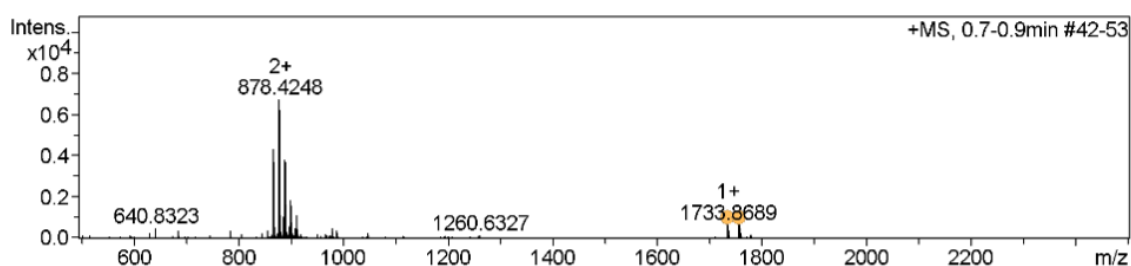
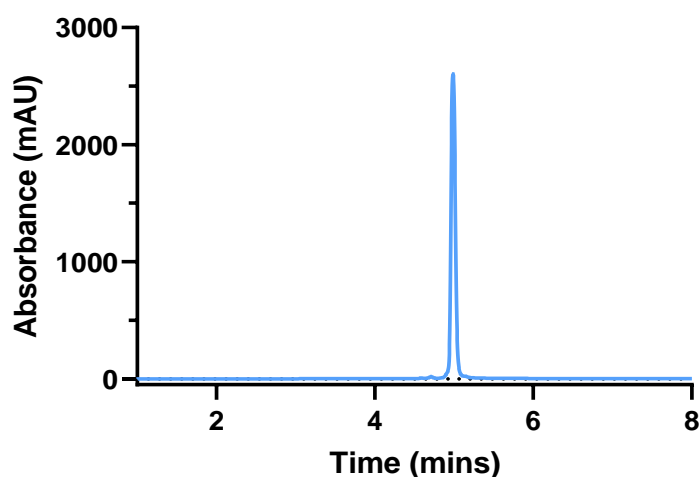
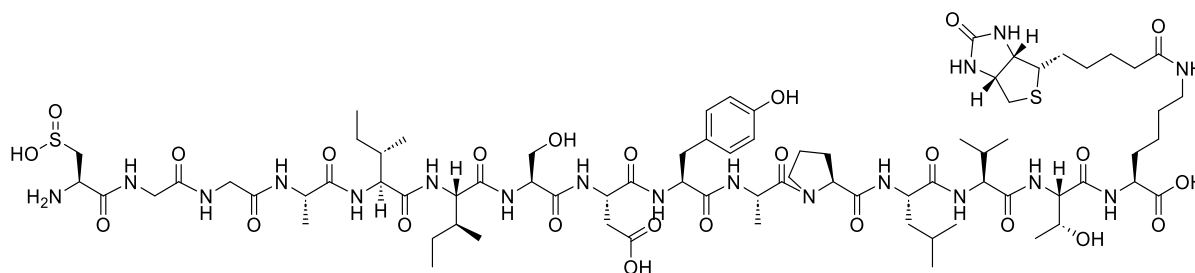


Figure 176 – Analytical HPLC trace and ESI MS of pure (**45**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass: 1732.85 [M+H]<sup>+</sup>, observed mass: 1733.87 [M+H]<sup>+</sup>.



## Synthesis of H<sub>2</sub>N-C(SO<sub>2</sub>)GGAIISDYAPLVTK(Biotin)-OH (**46**)



Peptide **46** was synthesised by manually loading the first (Lys-biotin, **51**) and final (Cys-SO<sub>2</sub>, **54**) amino acids following general manual coupling protocol on CTC resin (10 μmol). The rest of the peptide sequence was synthesised using general automated synthesiser protocols with microwave assistance. After cleavage, the crude peptide was treated with NaBH<sub>4</sub> (2 eq.) in EtOH (final concentration of peptide = 1mM) for 3 h to achieve the benzothiazole deprotection and reduction to the sulfinic acid. The crude peptide was purified by semi-preparative RP-HPLC (5-60% B over 30 min) and lyophilised to produce the desired peptide (**46** 2.2 mg, 1.25 μmol, 12% yield).

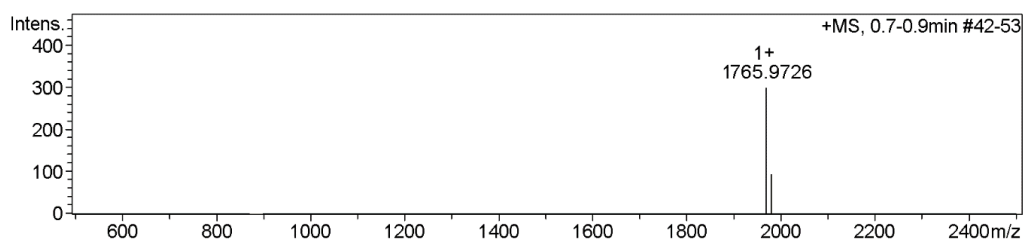
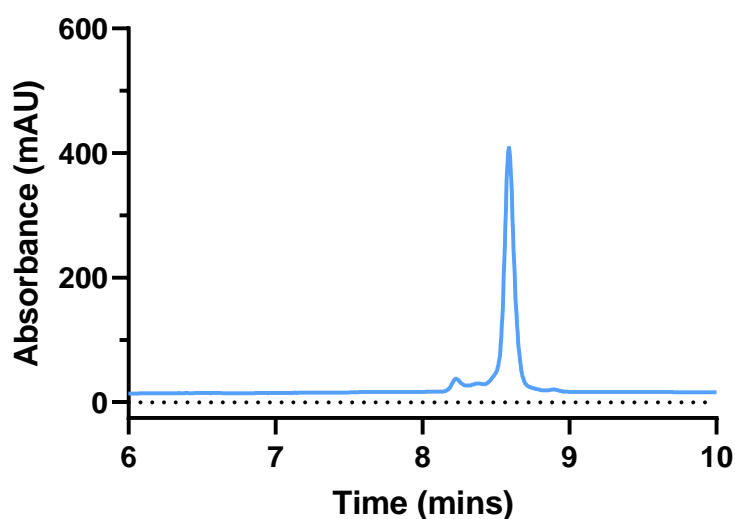
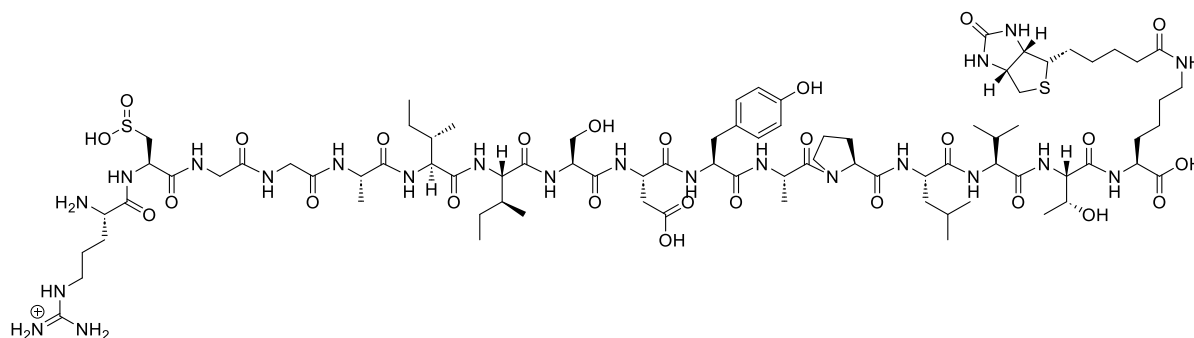


Figure 177 – Analytical HPLC trace and ESI MS of pure (**46**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass: 1764.84 [M+H]<sup>+</sup>, observed mass: 1765.97 [M+H]<sup>+</sup>.

## Synthesis of H<sub>2</sub>N-RC(SO<sub>2</sub>)GGAIIISDYAPLVTK(Biotin)-OH (**47**)



Peptide **47** was synthesised by manually loading the first (Lys-biotin, **51**), and final two (Cys-SO<sub>2</sub> **54**, and Arg) amino acids following general manual coupling protocols on CTC resin (10  $\mu$ mol). The rest of the peptide sequence was synthesised using general automated synthesiser protocol with microwave assistance. After cleavage, the crude peptide was treated with NaBH<sub>4</sub> (2 eq.) in EtOH (final concentration of peptide = 1mM) for 3 h to achieve the benzothiazole deprotection and reduction to the sulfinic acid. The crude peptide was purified by semi-preparative RP-HPLC (5-60% B over 30 min) and lyophilised to produce the desired peptide (**47** 1.6 mg, 0.83  $\mu$ mol, 8% yield).

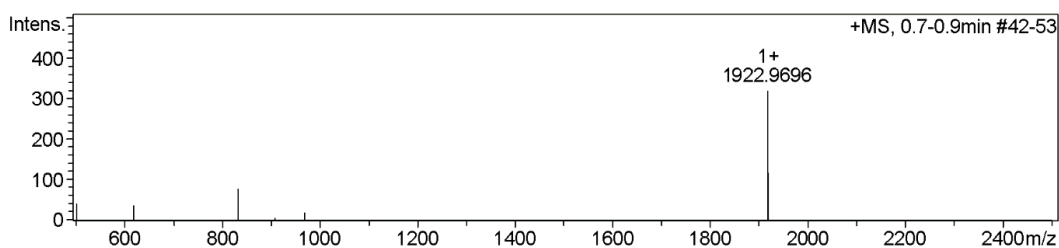
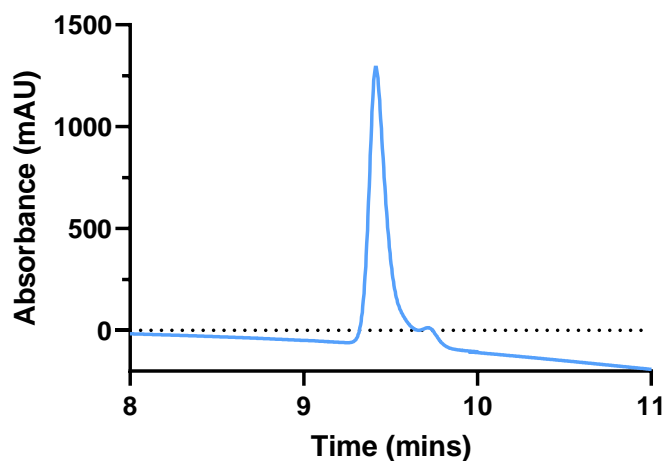
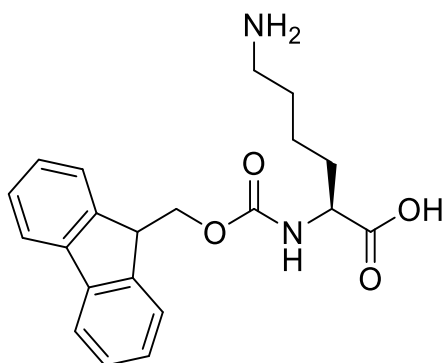


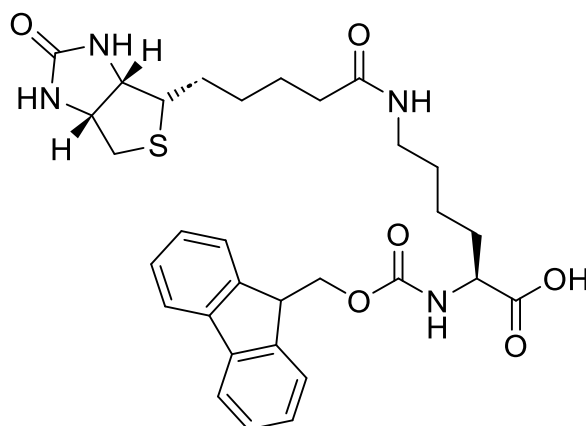
Figure 178 – Analytical HPLC trace and ESI MS of pure (**47**). Analytical gradient 5-60% B over 10 min, 280 nm. Calculated Mass: 1921.95 [M+H]<sup>+</sup>, observed mass: 1922.97 [M+H]<sup>+</sup>.

## Synthesis of (((9H-fluoren-9-yl)methoxy)carbonyl)-L-lysine (**50**)



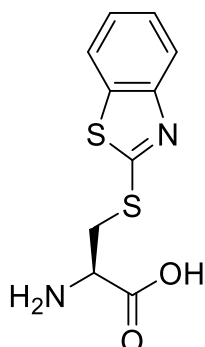
**50** was synthesised following an adapted literature procedure.<sup>202</sup> Fmoc-Lys(Boc)-OH (1 g, 2.1 mmol) was dissolved in a solution of TFA:H<sub>2</sub>O:TIPS (90:5:5 v/v/v, 5 mL) and stirred at r.t. for 1 h. Completion of the reaction was followed by HRMS. After 1 h, the solvent was evaporated under reduced pressure, water was added and washed with Et<sub>2</sub>O. Lyophilisation led to the titled compound (0.71 g, 92% yield) as a white powder that was utilised without further purification: <sup>1</sup>H NMR (400 MHz, DMSO) δ 7.90 (d, J = 7.5 Hz, 2H, fluorene), 7.78 (s, 1H, carboxylic acid), 7.73 (dd, J = 7.5, 3.6 Hz, 2H, fluorene), 7.63 (d, J = 8.0 Hz, 1H, sec. amide), 7.42 (td, J = 7.5, 1.1 Hz, 2H, fluorene), 7.33 (td, J = 7.4, 1.2 Hz, 2H, fluorene), 4.35 – 4.27 (m, 2H, CH<sub>2</sub>), 4.23 (q, J = 6.9, 6.2 Hz, 1H, CH), 3.93 (ddd, J = 9.7, 8.0, 4.6 Hz, 1H, fluorene), 2.78 (tt, J = 7.9, 3.6 Hz, 2H, CH<sub>2</sub>), 1.77 – 1.48 (m, 4H, CH<sub>2</sub>), 1.38 (td, J = 11.9, 11.1, 6.0 Hz, 2H, amine); <sup>13</sup>C NMR (101 MHz, DMSO) δ 173.87, 173.87, 156.22, 143.86, 143.81, 140.75, 129.00, 127.67, 127.09, 127.09, 125.28, 125.28, 120.15, 120.15, 65.59, 53.62, 47.55, 30.19, 26.53, 22.58, 16.16; HRMS calc. for C<sub>21</sub>H<sub>25</sub>N<sub>2</sub>O<sub>4</sub>: 369.18 [M+H]<sup>+</sup>, found: 369.18 [M+H]<sup>+</sup>.

**Synthesis of N2-(((9H-fluoren-9-yl)methoxy)carbonyl)-N6-(5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanoyl)-L-lysine (51)**



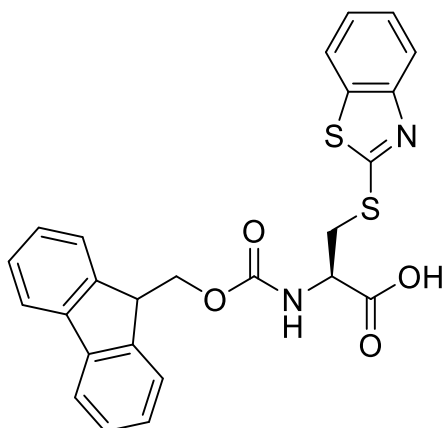
**51** was synthesised following an adapted literature procedure.<sup>203</sup> To biotin (594 mg, 2.4 mmol) in DMF (5 mL) was added HATU (1.08 g, 2.9 mmol) and *i*Pr<sub>2</sub>Net (0.85 mL, 4.9 mmol). After 5 min of pre-activation, the mixture was added to a solution of **50** (300 mg, 0.8 mmol) in DMF (3 mL), and the r.m. was left stirring at room temperature for 2 h. After this time, cold aq. HCl (pH = 2.0, 500 mL) was added, the precipitate was filtered and washed with cold aq. HCl (pH = 2.0, 3 x 200 mL), and H<sub>2</sub>O (3 x 200 mL). Lyophilisation led to the titled compound (438 mg, 93% yield) as a white fluffy solid that was utilised without further purification: <sup>1</sup>H NMR (400 MHz, DMSO) δ 7.89 (d, J = 7.5 Hz, 2H, fluorene), 7.76 (d, J = 5.7 Hz, 1H, sec. amide), 7.73 (d, J = 7.8 Hz, 2H, fluorene), 7.61 (d, J = 8.0 Hz, 1H, sec. amide), 7.42 (t, J = 7.4 Hz, 2H, fluorene), 7.36 – 7.30 (m, 2H, fluorene), 6.42 (s, 1H, CH), 4.30 – 4.20 (m, 4H, fluorene, CH<sub>2</sub>, CH), 4.13 – 4.10 (m, 1H, CH), 3.91 (ddd, J = 9.6, 8.0, 4.7 Hz, 1H, CH), 3.10 – 3.06 (m, 1H, CH), 3.04 – 2.98 (m, 2H, CH<sub>2</sub>), 2.84 – 2.77 (m, 1H, CH), 2.57 (dd, J = 12.4, 4.0 Hz, 1H, CH), 2.20 (t, J = 7.4 Hz, 1H, CH), 2.04 (t, J = 7.4 Hz, 2H, CH<sub>2</sub>), 1.60 (dq, J = 9.6, 4.7 Hz, 2H, CH<sub>2</sub>), 1.53 – 1.44 (m, 4H, CH<sub>2</sub>), 1.40 – 1.28 (m, 6H, CH<sub>2</sub>); <sup>13</sup>C NMR (101 MHz, DMSO) δ 174.01, 171.86, 162.74, 156.18, 151.19, 143.86, 143.81, 140.72, 128.92, 127.66, 127.09, 125.31, 120.77, 120.13, 65.61, 61.06, 59.21, 55.43, 53.78, 46.68, 38.19, 35.24, 33.50, 30.44, 28.76, 28.24, 28.13, 28.05, 25.33, 24.55, 23.12; HRMS calc. for C<sub>31</sub>H<sub>37</sub>N<sub>4</sub>O<sub>6</sub>S: found: 593.24 [M-H]<sup>-</sup>, found: 593.24 [M-H]<sup>-</sup>.

## Synthesis of S-(benzo[d]thiazol-2-yl)-L-cysteine (52)



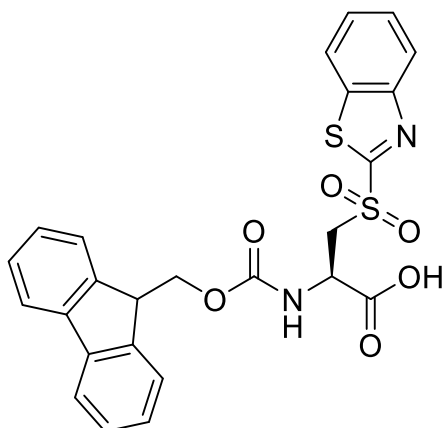
NaOH (1.96 g, 49.8 mmol) was dissolved in MeOH (60 mL). L-cysteine (2 g, 16.6 mmol) was added, and the r.m. was stirred under argon at r.t. for 10 min. 2-Chlorobenzothiozole (2.56 mL, 16.6 mmol) was added and the mixture was heated to reflux under argon for 10 min. The r.m. was allowed to cool to r.t. and stirred under argon for 16 h. After this time, water (100 mL) was added, and methanol was evaporated under reduced pressure. NaOH was added to completely dissolve the yellow-white solid. The r.m. became yellow and was washed with Et<sub>2</sub>O (2 x 20 mL). The aqueous layer was acidified to pH = 6.0 by the drop wise addition of glacial acetic acid. A white precipitate was formed and washed with ice-cold 100 mM ammonium acetate (1 L, pH = 6.7). Lyophilisation led to the titled compound (3.65 g, 86% yield) as a white solid that was used without further purification: <sup>1</sup>H NMR (400 MHz, DMSO) δ 7.89 (d, J = 8.0 Hz, 1H, CH), 7.84 (d, J = 8.1 Hz, 1H, CH), 7.42 (t, J = 7.7 Hz, 1H, CH), 7.31 (t, J = 7.6 Hz, 1H, CH), 4.05 – 3.96 (m, 1H, CH), 3.85 (dd, J = 8.1, 4.1 Hz, 1H, CH), 3.59 (dd, J = 14.0, 8.2 Hz, 1H, CH); <sup>13</sup>C NMR (101 MHz, DMSO) δ 169.99 (C), 167.07 (C), 153.34 (C), 135.68 (C), 126.97 (CH), 125.27 (CH), 122.19 (CH), 121.91 (CH), 54.34 (CH), 34.72 (CH<sub>2</sub>). HRMS calc. for C<sub>10</sub>H<sub>11</sub>N<sub>2</sub>O<sub>2</sub>S<sub>2</sub>: 255.03 [M+H]<sup>+</sup>, found: 255.03 [M+H]<sup>+</sup>. These spectral properties agreed with previously reported literature.<sup>204</sup>

## Synthesis of N-(((9H-fluoren-9-yl)methoxy)carbonyl)-S-(benzo[d]thiazol-2-yl)-L-cysteine (**53**)



**53** was synthesised following an adapted literature procedure.<sup>205</sup> A solution of N-(9-fluorenylmethoxycarbonyloxy)succinimide (2 g, 5.9 mmol) in acetonitrile (20 mL) was added to a solution of **52** (1.5 g, 5.9 mmol) in 9% Na<sub>2</sub>CO<sub>3</sub> (20 mL). The r.m. was stirred at room temperature for 2 h. Completion of the reaction was followed by HPLC. After 2 h, water (100 mL) was added, and the solution was acidified to pH = 2.0 with 12M HCl and extracted with ethyl acetate (3 x 100 mL). The combined organic layers were washed with water (2 x 50 mL) and brine (2 x 50 mL), dried over MgSO<sub>4</sub>, and evaporated under reduced pressure. The crude residue was purified *via* flash chromatography (0-10% MeOH in CH<sub>2</sub>Cl<sub>2</sub>) to give product as a white solid (1.93 g, 69% yield): <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.84 – 7.77 (m, 1H, Ar-H), 7.72 – 7.64 (m, 3H, Ar-H), 7.44 (dd, J = 11.9, 7.5 Hz, 2H, Ar-H), 7.36 (ddd, J = 8.3, 7.3, 1.3 Hz, 1H, Ar-H), 7.33 – 7.25 (m, 3H, Ar-H), 7.12 (td, J = 7.6, 3.6 Hz, 2H, Ar-H), 4.68 (dd, J = 7.5, 3.9 Hz, 1H, NH), 4.33 – 4.22 (m, 2H, CH<sub>2</sub>), 4.07 (t, J = 7.2 Hz, 1H, Fmoc), 3.88 (dd, J = 14.4, 3.9 Hz, 1H, CH), 3.74 (dd, J = 14.3, 7.4 Hz, 1H, CH); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ 177.81, 171.83, 167.27, 165.79, 164.58, 158.05, 157.45, 156.49, 152.35, 143.85, 143.82, 141.32, 135.42, 127.75, 127.11, 126.57, 125.21, 124.94, 121.52, 121.26, 120.00, 67.26, 54.87, 47.16, 35.30; HRMS calc. for C<sub>25</sub>H<sub>21</sub>N<sub>2</sub>O<sub>4</sub>S<sub>2</sub>: 477.09 [M+H]<sup>+</sup>, found: 477.09 [M+H]<sup>+</sup>. These spectral properties agreed with previously reported literature.

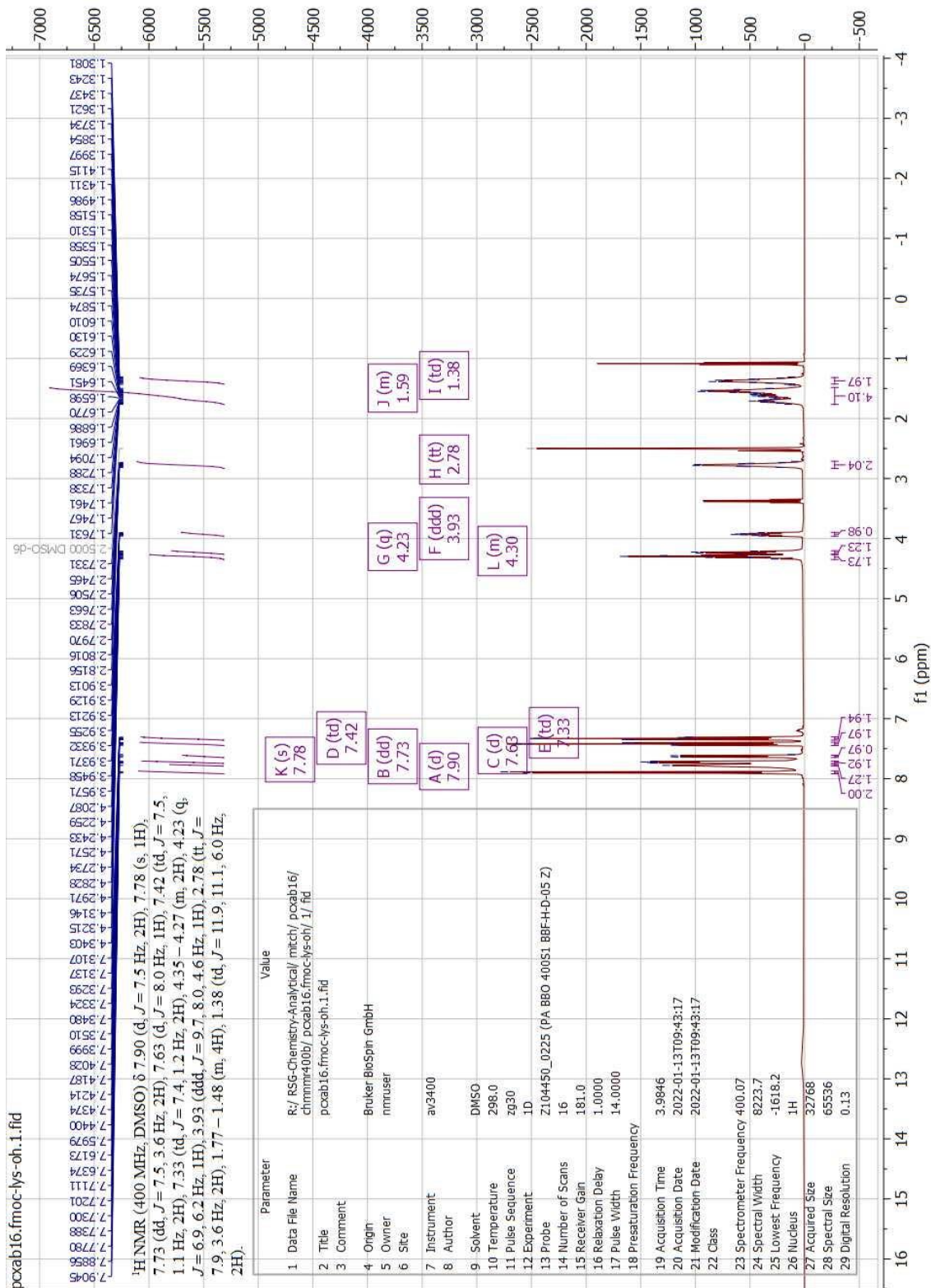
**Synthesis of (((9H-fluoren-9-yl)methoxy)carbonyl)(benzo[d]thiazol-2-ylsulfonyl)-D-alanine (54)**



To a solution of **53** (300 mg, 0.63 mmol) in dioxane (19 mL) mCPBA (543 mg, 3.14 mmol) was added in portions over 30 min. The r.m. was stirred at room temperature for 6 h. Completion of the reaction was followed by HPLC. After 6 h the r.m. was concentrated under reduced pressure and purified *via* flash chromatography (0-10% MeOH in CH<sub>2</sub>Cl<sub>2</sub>) to give product as a white solid (196 mg, 58% yield): <sup>1</sup>H NMR (400 MHz, DMSO) δ 13.34 (s, 1H, carboxylic acid), 8.24 (dd, J = 8.2, 4.1 Hz, 2H, Ar-H), 7.93 (d, J = 8.4 Hz, 1H, NH), 7.87 (d, J = 7.6 Hz, 2H, Ar-H), 7.65 (t, J = 7.6 Hz, 1H, Ar-H), 7.62 – 7.52 (m, 3H, Ar-H), 7.41 (t, J = 7.5 Hz, 2H, Ar-H), 7.34 – 7.24 (m, 2H, Ar-H), 4.59 (td, J = 8.0, 5.2 Hz, 1H, CH), 4.22 – 4.15 (m, 2H, CH<sub>2</sub>), 4.03 (dt, J = 11.0, 5.6 Hz, 1H, CH), 3.93 (q, J = 5.7, 4.1 Hz, 2H, CH<sub>2</sub>); <sup>13</sup>C NMR (101 MHz, DMSO) δ 170.42, 166.17, 155.38, 152.15, 143.61, 143.49, 140.64, 136.43, 128.05, 127.85, 127.66, 127.09, 127.06, 125.23, 125.19, 124.86, 123.39, 121.70, 120.70 – 120.64 (m), 120.08, 65.79, 54.80, 48.99, 46.30, 42.90; HRMS calc. for C<sub>25</sub>H<sub>20</sub>N<sub>2</sub>O<sub>6</sub>S<sub>2</sub>Na: 531.07 [M+Na]<sup>+</sup>, found: 531.06 [M+Na]<sup>+</sup>. These spectral properties agreed with previously reported literature.<sup>205</sup>

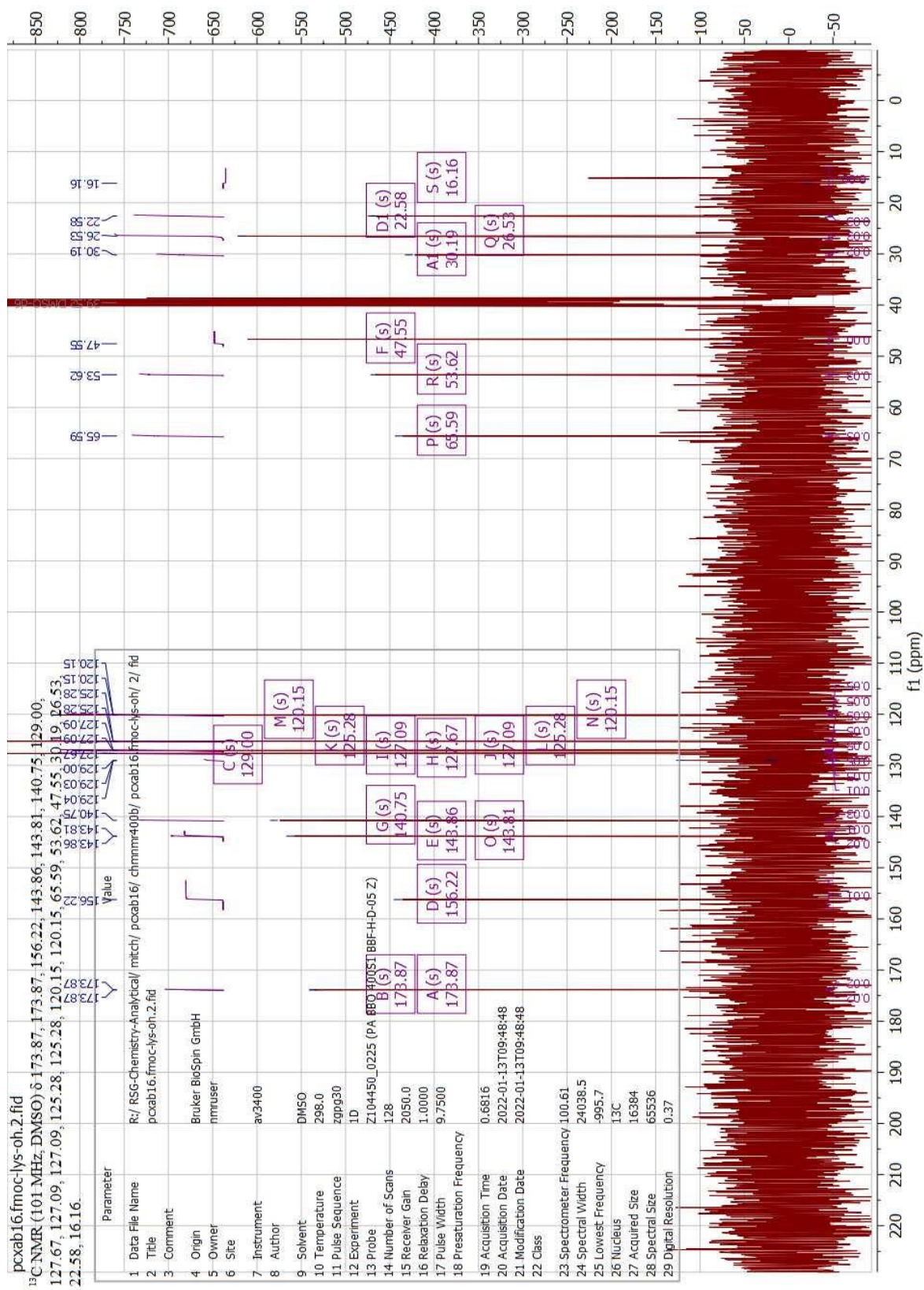
(((9H-fluoren-9-yl)methoxy)carbonyl)-L-lysine (50)

<sup>1</sup>H-NMR



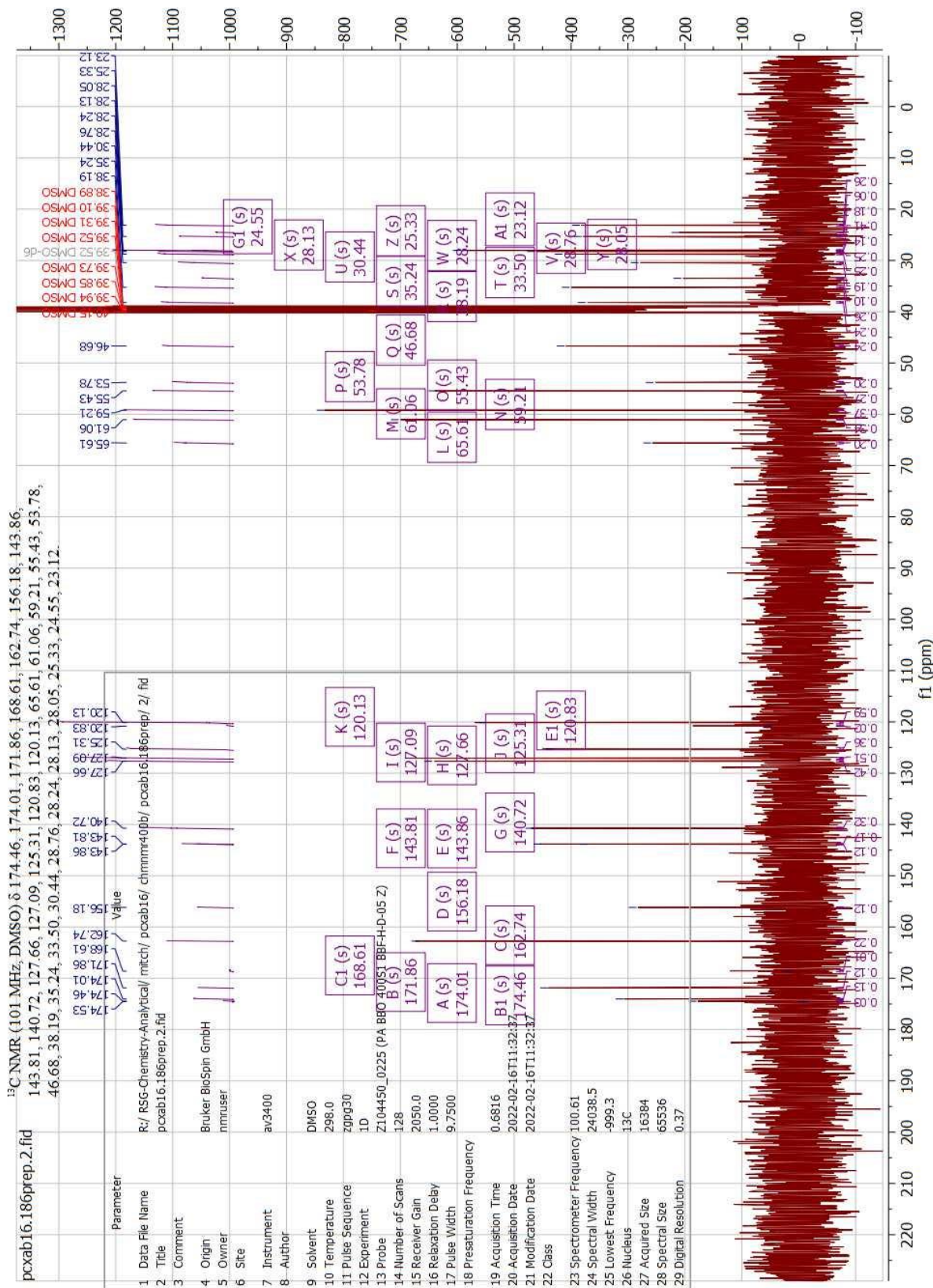


**<sup>13</sup>C-NMR**



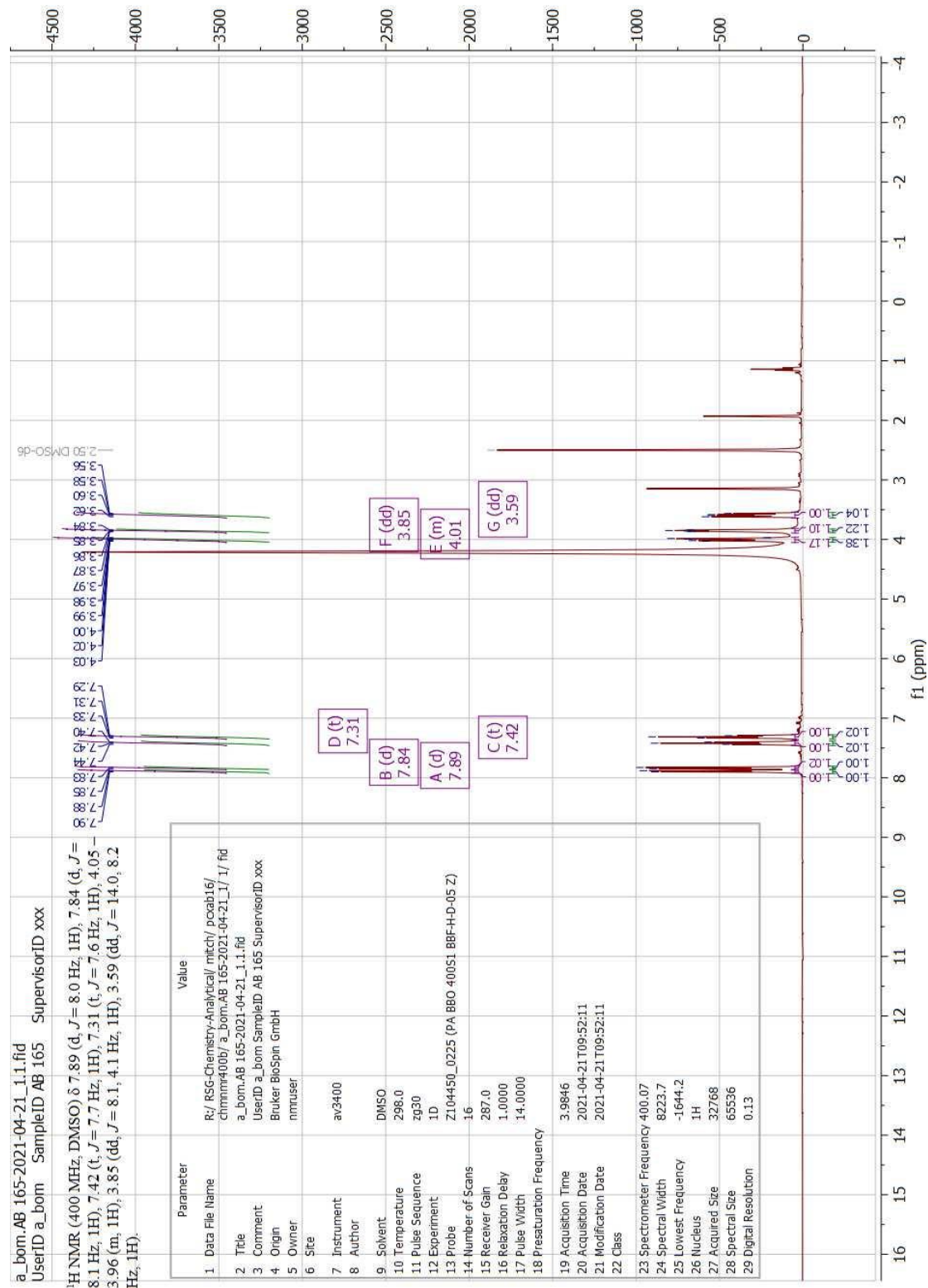


<sup>13</sup>C-NMR



# S-(benzo[d]thiazol-2-yl)-L-cysteine (52)

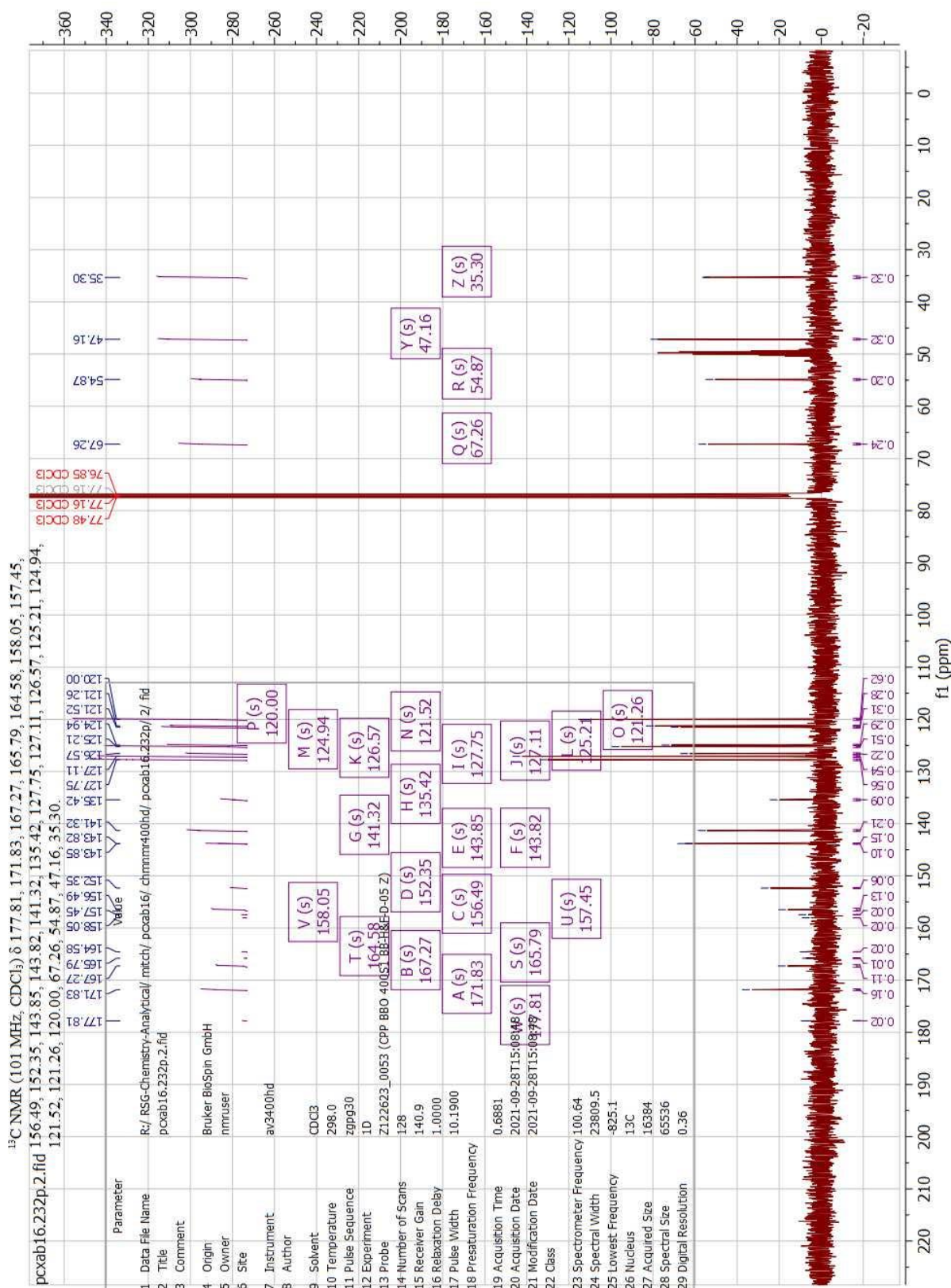
## <sup>1</sup>H-NMR





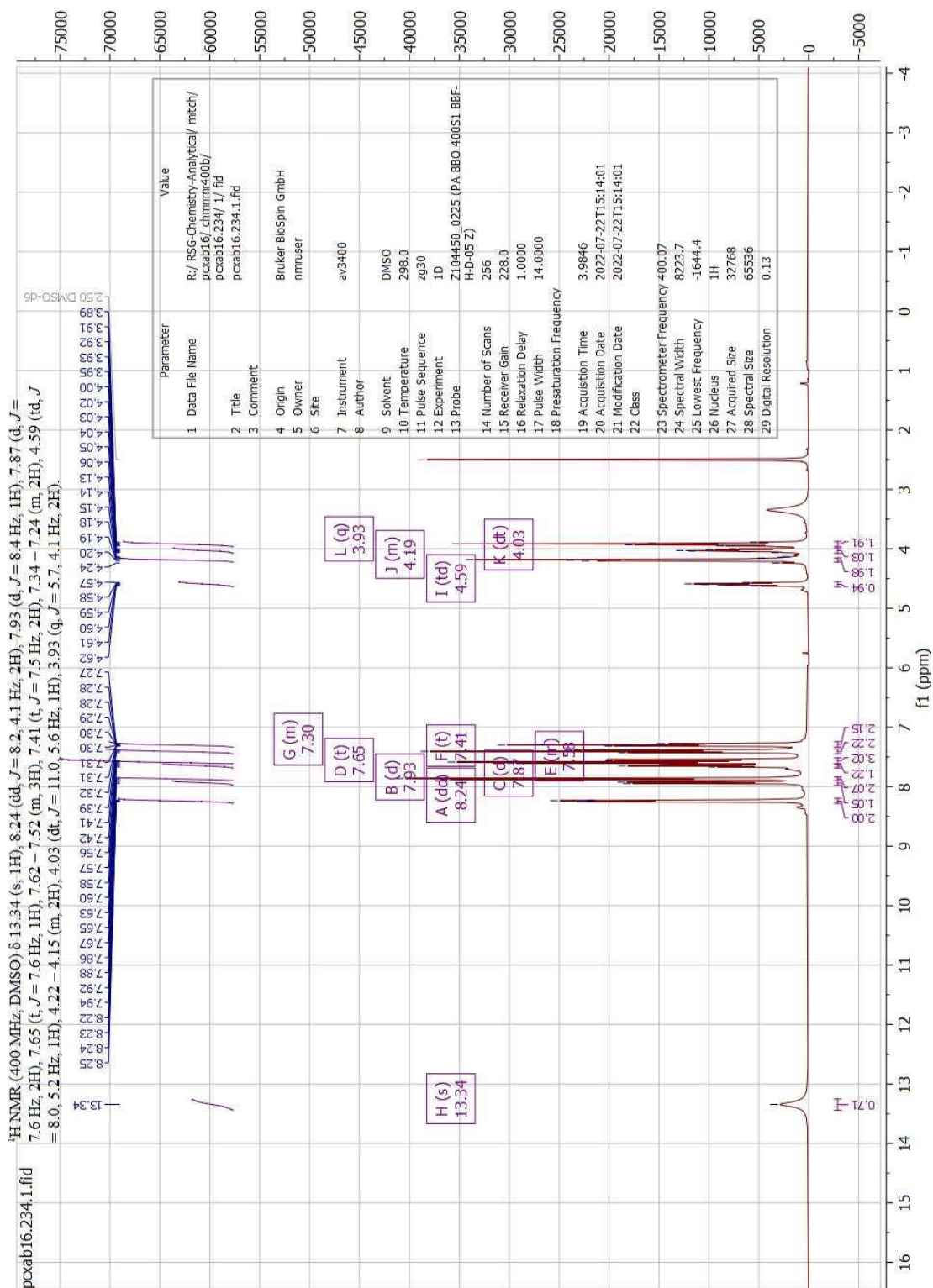


<sup>13</sup>C-NMR



**(((9H-fluoren-9-yl)methoxy)carbonyl)(benzo[d]thiazol-2-ylsulfonyl)-D-alanine**  
**(54)**

**<sup>1</sup>H-NMR**





<sup>13</sup>C-NMR

