Probing Iron-Sulfur Clusters Redox States by Incorporating an Unnatural Amino Acid



Jiaao Wei

Merton College

University of Oxford

Trinity Term 2021

A Thesis submitted to the Board of the Faculty of Physical Sciences, for the degree of

Doctor of Philosophy, University of Oxford

Abstract

Probing Iron-Sulfur Clusters Redox States by Incorporating an Unnatural Amino Acid

Jiaao Wei

Merton College, University of Oxford

Iron-sulfur clusters (FeS) are crucial redox cofactors in nature. They are involved in complex enzyme catalysis, acting as electron-relays. Techniques focusing on studying the electron loading on FeS-cluster relays during enzyme catalysis are rather scarce. Here, a novel approach is described to probe the redox state of FeS clusters by infrared (IR) spectroscopy, exploiting the versatility of this technique across a range of sample conditions and during turnover.

In this work, an unnatural amino acid *p*-cyano-L-phenylalanine (*p*CNF) was incorporated into the model protein spinach ferredoxin, which contains a [2Fe-2S] cluster. It is shown that *p*CNF can be used as infrared probe to prove the redox states of the [2Fe-2S] cluster. The integrity and correct functioning of the [2Fe-2S] cluster upon *p*CNF incorporation was confirmed by UV-Vis spectroscopy and electrochemistry. A small but reproducible shift in the streching frequency of the CN group of *p*CNF upon chemical or electrochemical reduction of the protein was observed by IR spectroscopy. Furthermore, detailed crystallographic characterisation was performed to compare the structures of the oxidised and reduced *p*CNF-containing ferredoxin, suggesting that the IR shift does not arise from a structural change but rather from the reduction of the [2Fe-2S] cluster.

Then this method was applied to more complex metalloenzymes containing FeS clusters. Experiments were then conducted on *E. coli* [NiFe] hydrogenase I to test whether an overexpression system could be suitable for extending the *p*CNF incorporation to Hyd-1. Although the overexpressed protein was found to behave normally comparing to the chromasomally encoded enzyme, it was decided not to continue with this enzyme due to difficulties in crystallising the overexpressed protein, unusual IR active site spectra features and limitation in yield.

Instead, the method was applied to [FeFe] hydrogenases from the sulfate reducing bacterium *Desulfovibrio desulfuricans* (*Dd*HydAB), to investigate how the FeS clusters interact with the diiron prosthetic group to allow efficient catalysis. *p*CNF was incorporated successfully close to the proximal cluster in *Dd*HydAB. From the preliminary IR results, there is a detectable difference in the band position of the cyano group in the F27CNF variant between the oxidised and reduced states on both apo- (lacking the binuclear site) and holo- (containing the full active site H-cluster) enzymes.

Overall, the research work in this thesis demonstrated a novel approach that applying an unnatural amino acid *p*CNF to metalloenzymes which contain FeS clusters to study the redox state of the FeS cluster by IR spectroscopy.

Acknowledgements

I would first like to express my sincere gratitude to Professor Kylie Vincent for giving me the opportunity to work in the lab on such an exciting project. You have been supportive and helpful all the time. Thanks for your invaluable advice, continuous support during the last four years.

Second, my thanks must go to Dr Miguel Ramirez. You guided and trained me on molecular biology research skills through the early days of my DPhil. Thanks for introducing me to the biology world, patiently teaching me every time and proofreading my thesis.

A huge thanks go to Dr Patricia Rodriguez. You not only provided me with endless guidance and advice on research, preparing *Dd*H, helping me to do FTIR and electrochemistry; but also being a friend outside the lab. You are a super woman and always work hard. Thanks so much for helping with my project and proofreading my thesis.

I am also really thankful to Dr Rhiannon Evans for all the help and advice, especially on protein film electrochemistry experiments, thesis proofreading. My special thanks go to Dr Stephen Carr, the crystallography expert in the Vincent group. Without you, my ferredoxin crystals cannot come into the world. Finally, I would like to thank all the group members past and present. I will never forget the time spending with all of you and cherish the memory working in the lab.

I would like to thank our collaborator Dr Amit Sachdeva from University of East Anglia, for kindly providing the plasmid used for incorporation of the unnatural amino acid. I would also like to acknowledge gratefully the financial support from China Scholarship Council.

I feel lucky to make a lot of friends during these 4 years PhD study, all of you made my spare time more colourful and meaningful at Oxford.

Additionally, the wholehearted, I am extreme grateful to my parents, grandparents, uncle and cousin for indispensable support and love. Mum and Dad, it was your unconditional love, trust that encouraged me to raise up and keep going. Mum, thanks for your endless support for whatever decisions I made and always being there for me, it is my constant sources of bravery and motivates me to keep forward. Finally, another piece of comfort and support comes from my girlfriend Yao Meng. Fate brought us together at Oxford, thanks for the company and encouragement when I was down. My life at Oxford cannot be the same without you. We have full of memories in Oxford UK, Europe, and China, I believe our story will continue forever.

Abbreviations

A	Amperes
Å	Angstrom (10^{-10} m)
BV	Benzyl viologen
CE	Counter electrode
CE Hyd-1	Chromosomally encoded *E.coli* NiFe Hydrogenase
CV	Cyclic voltammogram
°C	Degrees centigrade
Da	Dalton
*Dd*HydAB	[FeFe] hydrogenase from *Desulfovibrio desulfuricans*
E	Electrode potential
E°	Standard reduction potential
FdI	Spinach ferredoxin
IR	Infrared
I	Current
K_M	Michaelis constant; enzyme specific affinity constant
MB	Methylene blue
O.D.	Optical density
OE Hyd-1	Overexpressed *E.coli* NiFe Hydrogenase
*p*CNF	*p*-cyano-L-phenylalanine
PCR	Polymerase chain reaction
PFE	Protein film electrochemistry
PFIRE	Protein film infrared electrochemistry
SCE	Saturated calomel electrode
SEC	Spectroelectrochemistry
SHE	Standard hydrogen electrode
t	Time
T	Temperature
UV-vis	Ultraviolet-visible
V	Volts
WE	Working electrode
WT	Wild type

Contents

Chapter 1: Introduction 1

1.1 Iron-sulfur clusters 2

1.1.1 The role of iron-sulfur clusters 2

1.1.2 Common techniques to determine FeS cluster potential 3

1.1.3 Iron-Sulfur cluster proteins 4

1.1.4 Biosynthesis of FeS clusters 5

1.1.5 A model FeS protein: spinach ferredoxin 6

1.2 Learning from nature: Hydrogenases 7

1.2.1 Renewable energy technologies 7

1.2.2 Classification of hydrogenases 11

1.3 NiFe hydrogenase 12

1.3.1 The architecture of NiFe hydrogenase 13

1.3.2 The hydrogenase redox states and catalytic cycle 17

1.3.3 [NiFe] hydrogenase in *Escherichia coli* 21

1.4 FeFe hydrogenase 22

1.4.1 The structure of FeFe hydrogenase 23

1.4.2 FeFe hydrogenases catalytic cycle 27

1.5 Techniques for studying hydrogenases 28

1.5.1 Combining techniques to clarify enzyme mechanism 30

1.6 A novel IR vibrational reporter of protein environments 32

1.7 Aim and scope of this thesis 36

Chapter 2: Experimental Methods 39

2.1 Molecular biology methods 40

2.1.1 Strains, plasmids and oligonucleotide primers 40

2.1.2 Cell growth and culture 41

2.1.3 Methods for genetic engineering 42

2.2 Protein expression and purification 46

2.2.1 Bacteria growing and protein expression 47

2.2.2 Expression and purification of spinach ferredoxin 48

2.2.3 Expression and purification of [FeFe] hydrogenases from the sulfate reducing bacterium *Desulfovibrio desulfuricans* (*Dd*HydAB) 50

2.2.4 SDS-PAGE for protein separation 51

2.2.5 Western Blot 52

2.3 Solution assays for H_2 oxidation by hydrogenase 53

2.4 Spinach ferredoxin stability test 53

2.5 Electrochemistry 54

2.5.1 Three electrode system 54

2.5.2 Cyclic voltammetry and chronoamperometry 55

2.5.3 Electrochemical experiment set up for protein film electrochemistry 57

2.5.4 Electrolytic buffer solutions 58

2.6 Infrared spectroscopy 59

2.7 Spectroelectrochemical experiment set up 60

2.7.1 IR spectro-electrochemical cell 60

2.8 UV-Vis spectroscopy 61

2.9 Enzyme kinetics 61

2.9.1 Affinity constant 61

2.9.2 Henry's law and gas constants 64

2.9.3 Eyring equation and use 64

Chapter 3: Incorporating an Unnatural Amino Acid into the [2Fe-2S] Cluster Spinach Ferredoxin 65

3.1 Introduction and aim of the chapter 66

3.2 Spinach ferredoxin expression and characterisation 69

3.2.1 Spinach ferredoxin plasmid construction 69

3.2.2 Purification results 71

3.2.3 Basic and functional characterisation of the spinach ferredoxin 77

3.3 Test the tolerance of spinach ferredoxin to mutations close to the FeS cluster 80

3.4 Introduction of the amber codon in the targeted positions 85

3.5 FdI-CNF expression and characterisation 86

3.5.1 Expression and purification of FdI-CNF 88

3.5.2 Basic and functional characterisation of the unnatural ferredoxin I variants Y3CNF and Y37CNF 94

3.5.3 Infrared spectroscopy of the unnatural amino acid FdI variants Y3CNF and Y37CNF 96

3.6 Discussion and conclusion 97

3.6.1 Expression and purification of spinach ferredoxin I (FdI) 97

3.6.2 Characterisation of the overexpressed spinach ferredoxin I (FdI) 99

3.6.3 Selection of amino acids close to the FeS cluster for site-directed mutagenesis 99

3.6.4 Expression and purification of variants with *p*CNF 100

3.6.5 Characterisation of the FdI-CNF variants via electrochemistry and infrared spectroscopy 100

3.7 Conclusion 101

Chapter 4: Probing FeS Cluster Redox States by Incorporating an Unnatural Amino Acid 102

4.1 Introduction and aim of the chapter 103

4.2 Infrared spectroscopy on FdI-CNF variant 103

4.3 FdI crystallisation 108

4.4 Discussion and conclusion 113

4.4.1 Infrared spectroscopy on FdI-CNF variant 113

4.4.2 FdI crystallisation 114

4.4.3 Conclusion 115

Chapter 5: Characterisation of *E. coli* Ni-Fe Hydrogenase 1 Overexpressed in Native Host 117

5.1 Introduction and aim of the chapter 118

5.2 Aim of the chapter 123

5.3 Characterisation of overexpressed *E. coli* Hyd-1 124

5.3.1 Purification results 124

5.3.2 Solution assay 128

5.3.3 Michaelis constant and activation enthalpies 129

5.3.4 Electrocatalytic Window and E_{Switch} 135

5.3.5 Oxygen tolerance 138

5.4 Infrared spectroscopy of overexpressed Ni-Fe Hydrogenase 140

5.5 Site-directed mutagenesis of Hyd-1 ready for unnatural amino acid incorporation. 143

5.6 Discussion and conclusion 149

5.6.1 Purification results of overexpressed Hyd-1 149

5.6.2 Characterisation of overexpressed Hyd-1 150

5.6.3 Site-directed mutagenesis of Hyd-1 151

5.6.4 Conclusion 152

Chapter 6: Incorporating an Unnatural Amino Acid into [FeFe] Hydrogenase from *D. desulfuricans* 153

6.1 Introduction and aim of the chapter 154

6.1.1 [FeFe] hydrogenases 154

6.1.2 Biosynthesis of [FeFe] hydrogenase 155

6.1.3 Aim of this chapter 157

6.2 Overexpression of apo-*Dd*HydAB [FeFe] hydrogenase and in vitro (artificial) maturation 158

6.3 Incorporation of the unnatural amino acid *p*CNF around the proximal cluster in *Dd*HydAB: expression and purification 168

6.4 Characterisation of F27CNF *Dd*HydAB 170

6.5 *Dd*HydAB mutagenesis suitability tests around the proximal [4Fe-4S] cluster 177

6.6 Discussion and conclusion 182

6.6.1 Overexpression and characterisation of *Dd*HydAB 182

6.6.2 Expression and characterisation of F27CNF *Dd*HydAB 183

6.6.3 Testing the tolerance of A48 to mutagenesis close to the proximal FeS cluster 184

6.6.4 Conclusion 185

Chapter 7: Conclusion and Outlook 187

7.1 Conclusion 188

7.1.1 Incorporation of cyano-phenylalanine (*p*CNF) into the [2Fe-2S] cluster Spinach Ferredoxin I 188

7.1.2 Characterisation of *E.coli* [NiFe] hydrogenase 1 overexpressed in native host towards the incorporation of an unnatural amino acid *p*CNF 189

7.1.3 Incorporating an unnatural amino acid into [FeFe] hydrogenases from *D. desulfuricans* 190

7.2 Outlook 191

Bibliography 194

1

Chapter 1: Introduction

This project seeks to develop a method for measuring the redox state of FeS clusters in a large multi-redox centre enzyme, hydrogenase. A simple protein ferredoxin is first used as a model before moving on to hydrogenase.

1.1 Iron-sulfur clusters

Iron-Sulfur (FeS) clusters are evolutionarily ancient, ubiquitous and diverse prosthetic groups, which are involved in a variety of cellular processes that are central to life on earth, including respiration, photosynthesis and nitrogen fixation.[1] The structural similarities between biological and inorganic iron sulfide minerals, along with the prevalence of Fe and S in the earth's mineral record have driven numerous suggestions that these basic, abiotic precursor compounds may represent some of the earliest catalysts on the primordial earth, and have remained an essential component in living organisms.[2] Fe atoms participate in many other prosthetic groups and proteins as part of hemes, and mono- or di-iron non-heme centres, which is one of the reasons why Fe is an essential element for most organisms. Sulfur is also an essential element, required for the synthesis of amino acids cysteine and methionine.[3]

1.1.1 The role of iron-sulfur clusters

FeS clusters have critical roles in proteins from diverse organisms. They are functionally versatile and are used for a wide variety of essential electron transfer steps, as well as other biological processes including acid-base and radical catalysis and sensing. FeS clusters are ideal for electron transferring due to the redox potential of Fe, since the redox potential of the $Fe^{III/II}$ couple is compatible with many biologically processes. The potential is tuned by the cluster environment. The redox potential range of FeS cluster is from greater than 500 mV to less than –500 mV which is much larger than any other redox cofactor.[4] As an example of their utility in biological processes, relays of FeS clusters function as series of redox centres or electron reservoirs in the respiratory and photosynthetic electron transfer chain.[5] Electrons transfer can

occur at catalytically relevant rate with a distance of up to 14 Å between redox centre through the protein medium.[6]

1.1.2 Common techniques to determine FeS cluster potential

There are many techniques and methods to study and determine FeS cluster potential, electrochemistry and electron paramagnetic resonance (EPR) spectroscopy are two common techniques. Electrochemistry is widely used to study redox reactions. A single redox reaction can be isolated and analysed precisely using the electrochemical method, where oxidation occurs at an anode, and a complementary reduction occurs at a cathode. Cyclic voltammetry provides a method to control the redox chemistry through the applied potential and observe the FeS cluster transformations that are induced.[7] A pair of oxidation and reduction peaks act as electrochemical markers for identifying the status of FeS cluster within the protein. However, voltammorgrams cannot provide structure details, therefore, electron paramagnetic resonance (EPR) is always used together with electrochemistry to examine species generated in solution at corresponding potential.[8]

Electron paramagnetic resonance (EPR) spectroscopy is a spectroscopic technique to determine FeS cluster potential that detects species that have unpaired electrons.[9] The degree of interaction between the unpaired electrons and the perturbing nuclei is given by the variation of the spacing between the EPR spectral lines; therefore, a large amount of information can be obtained by analysis of these parameters in FeS cluster studies, such as the oxidation state and spin state of the metal. EPR spectroscopy has been instrumental in the early detection of the FeS centres.[10] However, there is a drawback of EPR spectroscopy that the experiments have to be performed under frozen conditions, it is difficult to determine the FeS cluster redox states and potentials under room temperature catalytic conditions. in some EPR spectroscopy has made it possible to obtain information on all the redox states and redox potentials of the FeS clusters, but the experimental conditions are on frozen samples, so not under active catalytic turnover.[11]

1.1.3 Iron-Sulfur cluster proteins

Iron sulfur clusters consist of iron atoms directly connected to the cysteinyl sulfurs of the associated protein, with additional inorganic sulfides. Fe and S can occur at different stoichiometries in proteins, which allows for a diverse array of FeS cluster structures that span a range of redox potentials. There are many types of FeS clusters in nature: the common cluster types are shown in Figure 1.1. The rhombic [2Fe-2S] and cubic [4Fe-4S] are the most common species in FeS proteins.[12] The [3Fe-4S] cluster can be considered as a cubic 4Fe cluster with one Fe missing. FeS proteins have a wide range of reduction potentials, thus FeS clusters can serve as excellent donors or acceptors of electrons in a large number of biological reactions.[12] The mid-range redox potential and unique features of these remarkably versatile cofactors allows the FeS proteins to transfer electrons, catalyse enzymatic reactions (as it displays strong Lewis acid properties) and function as regulatory proteins.[4,13]

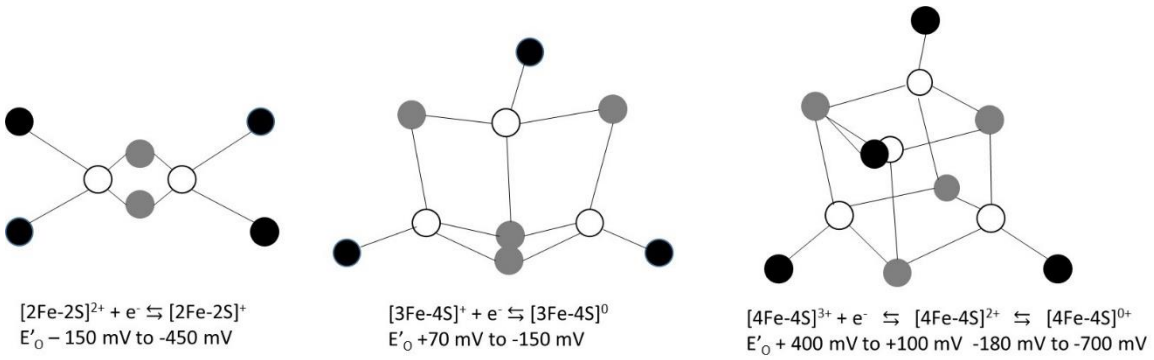


Figure 1.1 Models for structures of 2Fe, 3Fe, and 4Fe clusters. White = Fe, black = cysteine; grey = sulfide sulfur.[1]

The function of FeS clusters in electron transfer protein was first found in ferredoxins. These are involved in many fundamental processes of life such as photosynthesis, cell-signalling and respiration.[14] There are different types of FeS clusters in ferredoxins. These proteins are found in many organisms including archaea, bacteria and eukaryotes.[14] Spinach ferredoxin (FdI) is a simple protein containing a single [2Fe-2S] cluster, and it is the most abundant protein in spinach chloroplasts.[15]

Whilst ferredoxins are one of the simplest FeS proteins, there are many others in nature. Table 1.1 lists a few of the known FeS containing proteins that have diverse functions as in systems ranging from bacteria to mammals.

Table 1.1 Some examples of diverse iron-sulfur cluster proteins.

ISC protein	Function	Subcellular location	Ref
Respiratory complex I-III	Electron transport chain	Mitochondrial inner membrane	16
Mitochondrial aconitase	Citric acid cycle	Mitochondrial matrix	5
Ferrochelatase	Haem biosynthetiesis	Mitochondrial matrix	17
Ferredoxin-NADP reductase (FNR)	Regeneration of NADPH	Chloroplast stroma	18
[NiFe] Hydrogenase	H_2 metabolism	Bacteria, Archaea	19
Spinach ferredoxin	Electron carrier	Spinach chloroplast	20

1.1.4 Biosynthesis of FeS clusters

Iron and sulfur are toxic elements in the living cell but FeS clusters are indispensable for the cells, with roles in metabolic pathways, protein translation and DNA maintenance.[21] Therefore, it is important to understand the biosynthetic mechanism for FeS cluster formation in the cell. The majority of cellular FeS proteins are regulated by the iron-sulfur cluster assembly machinery (ISC system).[22] The isc homologue genes encode the fundamental FeS cluster biosynthesis pathways and have been identified in most of the sequenced genomes from bacteria and archaea to plants and humans.[21] Figure 1.2 shows the main components of the bacterial ISC machinery. IscU is a scaffold protein where the FeS clusters are assembled; where sulfur is provided by the cysteine sulfurase IscS and iron is offered by a still to be defined chaperone.[23] FdX is a [2Fe-2S]-cluster protein which can deliver electrons reduced by NADP reductase (Fenr). The assembled cluster is then transferred to late-acting components by chaperone system HscA and HscB with ATP (adenosine triphosphate) consumption. Grx4 acts as a relay to late-acting components, and

IscA has specialised functions in assembling the [4Fe-4S] clusters and inserting of [2Fe-2S] and [4Fe-4S] clusters into apo protein.[23,24]

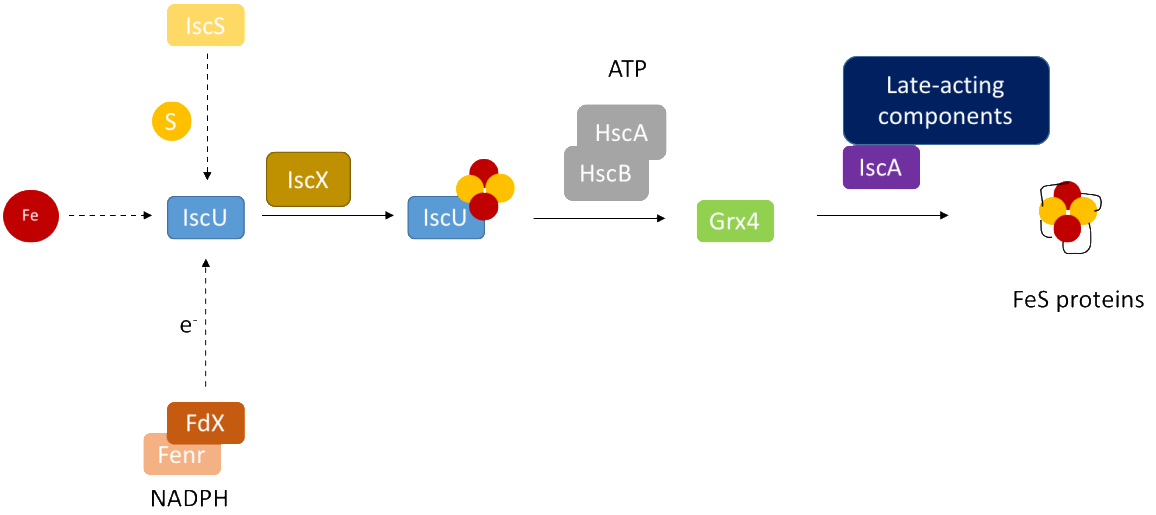


Figure 1.2 Main components of the bacterial ISC machineries. It contains proteins encoded by the *isc* operon (IscU, IscS, IscX, FdX, HscA, HscB and IscA).[23]

1.1.5 A model FeS protein: spinach ferredoxin

Ferredoxins are the first FeS proteins discovered in nature. The first ferredoxin was purified in 1962 from the anaerobic bacterium *Clostridium pasteurianum.* Their name arises from the fact that they were found to contain Fe and be redox active.[25] This large group of FeS proteins can be classified depending on the characteristics of their FeS clusters, as well as by sequence similarity, but they all share certain common properties: they are very small proteins, they function in electron transfer roles, and they are highly acidic proteins containing Fe and inorganic or acid-labile sulfurs. Ferredoxins can be found in nature as [2Fe−2S], [4Fe−4S], [3Fe−4S], [3Fe−4S][4Fe−4S], and [4Fe−4S][4Fe−4S] cluster containing proteins.[26] They are usually low potential proteins with an average redox potential of −400 mV (vs. SHE). This wide range of potentials enables ferredoxin to act as redox partners to a variety of molecules in many important biological reactions.[27]

Spinach ferredoxin (FdI) was first described in 1968 by Matsubara and Sasaki, and it was the first iron-sulfur protein with known function.[28] It is a small [2Fe-2S] protein of around 13 kDa, and it

is the most abundant isoform of photosynthetic-type ferredoxin present in the spinach chloroplast. The simple structure of the FdI makes it a commonly used model system to study FeS cluster proteins.[29] The rhombic [2Fe-2S] cluster in chloroplast ferredoxin (shown in Figure 1.3) is the electron carrier which transfers electrons received from PSI (photosystem I) to several enzymes. The most important electron-acceptors include the ferredoxin-$NADP^+$ reductase and ferredoxin-thioredoxin reductase.[30] The FeS cluster in the ferredoxin can accept or donate electrons, with the effect of a change in the oxidation state of the cluster.[20]

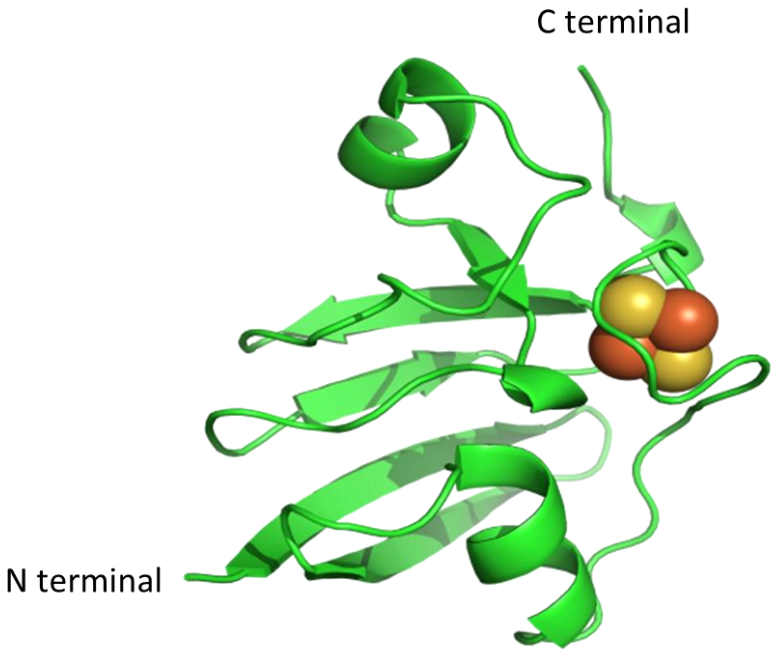


Figure 1.3 The structure of spinach ferredoxin E92K. The sphere in the middle represents the [2Fe-2S] cluster (PDB code: 1A70)

1.2 Learning from nature: Hydrogenases

FeS cluster-containing enzymes called hydrogenases are inspirational for teaching us about the sustainable use of H_2 as a fuel. They are effective biocatalysts for biological H_2 metabolism.

1.2.1 Renewable energy technologies

Nowadays, fossil fuels such as oil, coal and natural gas are the major energy sources used around the world, but they will not last forever, and the increasing use of these non-renewable

fuel sources also produces a large number of negative effects such as pollution, the greenhouse effect and the fast consumption of the finite fossil fuel reserves.[31] Therefore, it is important to find a clean energy alternative to meet the high energy demand. In principle, dihydrogen (H_2) could be an excellent solution for the energy crisis. Hydrogen is not only one of the vital clean energy carriers playing a significant role in the global energy transitions, but it is also important in industrial chemical processes such as hydrogenation of organic compounds.[32] In one application, crude oil is treated with H_2 to get rid of sulfur and nitrogen by means of hydrodesulfurisation and hydrodenitrogenation.[32] Hundreds of millions of tons of ammonia produced from H_2 and N_2 in the Haber-Bosch process can be used as fertiliser in the agriculture industry; catalytic hydrogenations are the most common chemical reactions in the world.[33]

However, due to its low density, there is only a small amount of H_2 that is easily available in the earth's atmosphere. Hydrogen on earth exists in molecular forms like organic compounds and water; hydrogen atoms form a strong bond with oxygen to form H_2O, a readily available resource.[34] The dihydrogen molecule contains a strong H-H bond with a bond dissociation enthalpy of 436 kJ/mol at 25 °C.[35] The reaction of H_2 with O_2 is able to release a large amount of energy (under exploding conditions, upon ignition).

$$2H_2 + O_2 \rightarrow 2H_2O\ (l) \qquad \Delta H^O = -286\ kJ\ mol^{-1}$$

For this reason, catalysts are needed to help break the H–H bond under milder conditions, by lowering the activation energy. In chemistry, the best catalysts for H_2 activation are precious metals, whereas in nature the catalysts for H_2 activation are enzymes containing Fe or Fe and Ni.

A significant disadvantage of H_2 is its low solubility in water, preventing its use as an energy source in aqueous solutions. Biology has circumvented this disadvantage by binding the hydrogen atom to an organic carrier. In living organisms, hydrogen is often stored and transferred as the hydride carrier NADH and NADPH (nicotinamide adenine dinucleotide and the phosphate derivative).[35]

In microorganisms, the oxidation of H_2 to generate protons and electrons is an important process.[36] Additionally, many prokaryotes are able to produce H_2 as a way of getting rid of excess reductant during fermentation. Other microorganisms can directly utilise H_2 as an energy source.[37] For example *E. coli* can use hydrogenases in both H_2 production and H_2 oxidation depending on its metabolic needs and environmental conditions.[38] Table 1.2 summarises the main metabolic activities involving H_2. Biological H_2 production is mainly due to fermentation and anaerobic CO oxidation, whilst the aerobic oxidation of H_2, methanogenesis and acetogenesis, are the most common processes to consume H_2 in microorganisms.[37]

Table 1.2 Summary of the main metabolic activities involving H_2 in prokaryotes.

Process	Chemical reaction	Class of enzyme involved
Fermentation	$CH_3CH_2CH_2COOH + 2H_2O \rightarrow 2CH_3COOH + 2H_2$	FeFe[39,40]
	$CH_3CH_2COOH + 2H_2O \rightarrow 2CH_3COOH+CO_2 + 3H_2$	
Anaerobic CO oxidation	$CO + H_2O \rightarrow CO_2 + H_2$	NiFe[41]
N_2 fixation	$N_2 + 8H^+ + 8\ e^- + 16MgATP \rightarrow 2NH_3 + H_2 +$ $16MgADP + 16P_i$	Mo nitrogenase[42]
Aerobic H_2 oxidation	$2H_2 + O_2 \rightarrow 2H_2O$	NiFe[36]
Methanogenesis	$4H_2 + CO_2 \rightarrow CH_4 + H_2O$	NiFe[43,44]
	$CH_3OH + H_2 \rightarrow CH_4 + H_2O$	
Acetogenesis	$2CO_2 + 4H_2 \rightarrow CH_3COOH + 2H_2O$	NiFe, FeFe[37]

Oxidation of H_2 by hydrogenases provides energy to the organism; the resulting electrons can be used to reduce O_2 or fumarate in the respiratory chain which is one of the most common routes of energy generation.[38] The low-potential electrons generated from hydrogenase-catalysed H_2 oxidation can also be used to fix inorganic carbon.[45] Additionally, certain organisms can use these electrons for the direct reduction of bio-cofactors such as NAD^+ or $NADP^+$ to NADH and NADPH (for example, by the soluble hydrogenase from *Ralstonia eutropha*).[46]

Since H_2 has been regarded as a clean energy carrier for the future, it is indispensable to develop new technologies for both production and oxidation of H_2. The main method for H_2 oxidation and electricity generation is via a platinum (Pt) electrode in a fuel cell.[47] However, due to the high cost of Pt, hydrogenases have attracted wide interest to gain an understanding of how the biological systems make these reactions happen efficiently in the cell with the abundant metals Fe and/or Ni.[46] Unfortunately, the stability of the hydrogenases is much lower than for Pt, and this is a significant barrier for their implementation outside cells. Nevertheless, understanding how hydrogenases work is very important for developing a new generation of efficient, low cost catalysts for H_2-based energy technologies.[35,48]

Additional applications of hydrogenases in H_2-dependent chemical reactions have been developed. For example, the group of my supervisor Professor Vincent recently developed a system that combines the selectivity of biocatalysis and atom efficiency of catalytic hydrogenation by using enzymes for H_2-driven NADH recycling, is shown in Figure 1.4.[49] This methodology, overcomes one of the most important challenges with biocatalysis: minimising waste production. In this HydRegen system, a hydrogenase and an NAD^+ reductase are co-immobilised on a carbon particle where electrons produced during H_2 oxidation are used for the enzymatic reduction of NAD^+. These enzyme-modified particles have been applied in multiple settings including batch and continuous processes. Additionally, these particles can be recycled multiple times, retaining high NADH production activity.[49]

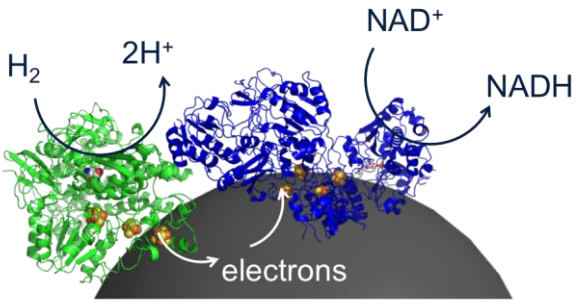


Figure 1.4 Scheme representation of H_2-driven cofactor generation system for NADH production. H_2 is oxidised by the hydrogenase (green, representative PDB code: 3USE) and the electrons transferred through the conductive particles for NAD^+ reduction by an NAD^+ reductase enzyme (blue, representative PDB code: 5XF9). Adapted with permission from Wiley-VCH publisher: *ChemCatChem*, vol 7, 3480–3487, copyright 2015.[49]

1.2.2 Classification of hydrogenases

Hydrogenases can be phylogenetically classified into [NiFe], [FeFe], and [Fe] hydrogenases based on the metal composition of the catalytic active site. The vast majority of the discovered hydrogenases belong to the [NiFe] and [FeFe] hydrogenase classes, with more than 100 types characterised genetically and/or biologically. Iron-only hydrogenases are only found in archaea, where their function is to transfer hydride (H^-) to a cofactor.[50,51] The distinct structures of the three different classes of proteins are the consequence of convergent evolution.[38,52]

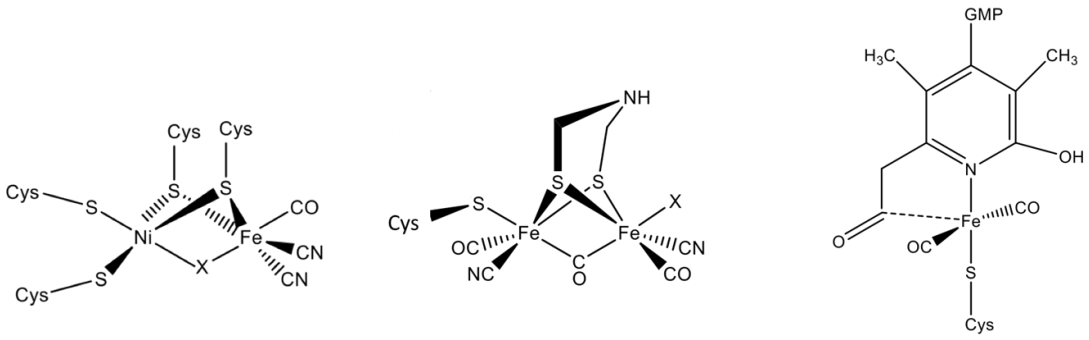


Figure 1.5 Chemical structures of active sites from the three different hydrogenases.[53]

Figure 1.5 shows the active sites of the three hydrogenase classes that each form a highly specialised complex centre in the protein active site. The active site provides space for the

catalytic reactions and is connected to the protein surface via channels to transfer substrates/products. It is notable that the active site of [NiFe] and [FeFe] hydrogenases share a commonality: they both contain CO and CN^- ligands linked to an iron (Fe) ion. These two inorganic ligands are able to stabilise the iron centre in a low oxidation state and low spin configuration, presumably due to their strong σ-donating and π-accepting character. The two metal ions in the active site, either NiFe or FeFe, are connected by sulfur atoms. The metals act as Lewis acids: the empty orbitals of the metal ions can accept electrons which makes the H_2 catalysis more effective.[54]

1.3 NiFe hydrogenase

[NiFe] hydrogenases can be classified into four groups according to their phylogeny. Group 1 contains membrane-bound H_2-uptake [NiFe] hydrogenases; group 2 encompasses cytosolic H_2-uptake [NiFe] hydrogenases; group 3 includes cytosolic bidirectional [NiFe] hydrogenases, and membrane-bound H_2-evolving [NiFe] hydrogenases belong to group 4.[45]

Although the [NiFe] class hydrogenases are sub-divided based on function, there is great interest in the differences between a subgroup in group 1 which are classified as O_2-tolerant hydrogenases, as distinct from their O_2-sensitive counterparts.[19] For example, there are three active [NiFe] hydrogenases (Hyd1-3) in *E. coli* K-12; Hyd-1 and Hyd-2 are both respiratory enzymes that oxidise H_2 with their active sites facing the periplasm. Interestingly, Hyd-1 is O_2-tolerant with very low activity towards proton reduction, and Hyd-2 is regarded as a 'standard' O_2-sensitive hydrogenase with reversible H_2 oxidation/proton reduction activity.[55] These two membrane-bound H_2-uptake [NiFe] hydrogenases from group 1 are subjects of some of the work in this thesis.

1.3.1 The architecture of NiFe hydrogenase

The 'core' modules of [NiFe] hydrogenase are composed of two subunits, a large subunit with a buried active site and a small subunit with an FeS cluster relay for electron transport from the protein surface to the catalytic active site.[19,55] Substrates and products are transported through specific routes to facilitate catalysis, as discussed below. Many of the [NiFe] hydrogenases are purified as heterodimers, i.e., just the small and large subunits.

Active site

[NiFe] hydrogenases have a highly optimised functional protein matrix with a deeply buried catalytic active site. Almost all the [NiFe] hydrogenases have similar active sites except for the subclass [NiFeSe].[19] The active site is located in the large subunit of [NiFe] hydrogenase and contains the Ni and Fe and four cysteine sulfurs linking the active site to the main protein structure (shown in Figure 1.5). Two of the cysteine residues connect to the Ni only in a terminal fashion, the other two bridge the Ni and Fe ions. The Fe is also coordinated by small inorganic ligands, CO and two CN^-, which are both strong-field diatomic ligands, first detected by FTIR spectroscopy and are thought to keep the Fe as low spin Fe(II).[56] An additional ligand (shown as ligand X in the [NiFe] active site in Figure 1.5) is located between the Ni and Fe, the identification of this ligand varies depending on the redox state of the enzyme, presence of substrate, or inhibitors.[57]

Iron-sulfur clusters

Since the catalytic active site of [NiFe] hydrogenase is deeply buried in the protein, the electrons used/released in catalysis and recovery of oxidised inactive states are dependent on the electron relay chain of FeS clusters. Standard [NiFe] hydrogenases (or so-called O_2 sensitive [NiFe]-hydrogenases) are inactivated by even small quantities of O_2, forming oxidised inactive states that can take minutes to hours to reductively activate. Interestingly, some membrane-

bound [NiFe] hydrogenases are O_2-tolerant, able to perform catalysis in the presence of ambient O_2.[58,59] This requires electron transfer back from the FeS clusters to the active site. Electrons must also flow to the active site during H_2 production and away from the active site during H_2 oxidation. The FeS cluster relays of [NiFe] hydrogenases generally contain FeS clusters located proximally ([4Fe-4S] or [4Fe-3S], see below), medially ([3Fe-4S]) and distally ([4Fe- 4S]) from the active site with a 12–14 Å distance between each redox site. It has been proposed that this minimal distance between the FeS clusters permits viable electron transfer without constraining the rate of catalysis of the enzyme.[60] A prominent difference between standard O_2-sensitive and O_2-tolerant hydrogenases comes from the cluster proximal to the active site, standard [NiFe] hydrogenases display a [4Fe-4S] cluster while O_2-tolerant possess a special [4Fe3S] cluster.[61,62] According to sequence analysis, modelling studies and crystal structures of O_2-tolerant hydrogenase, the proximal [4Fe3S] cluster is ligated by six cysteine residues, so two more cysteine residues are coordinated to the cluster than those of standard [NiFe] hydrogenases.[63,64]

Figure 1.6A shows the crystal structures of reduced cubane-shape form of the [4Fe-3S] cluster from *Hydrogenovibrio marinus* (Hm) MBH.[65] The formation of the special [4Fe-3S] cluster results from one of the thiolates of the cysteine residues (Cys25 in *Hm*) replacing the sulfide of the FeS cluster.[66] Figure 1.6B shows the asymmetric superoxidised state of [4Fe-3S] cluster, the deprotonated amide nitrogen of Cys26 being coordinated to the Fe results in a conformational change with breaking of the bond between Fe and the sulfide.[65]

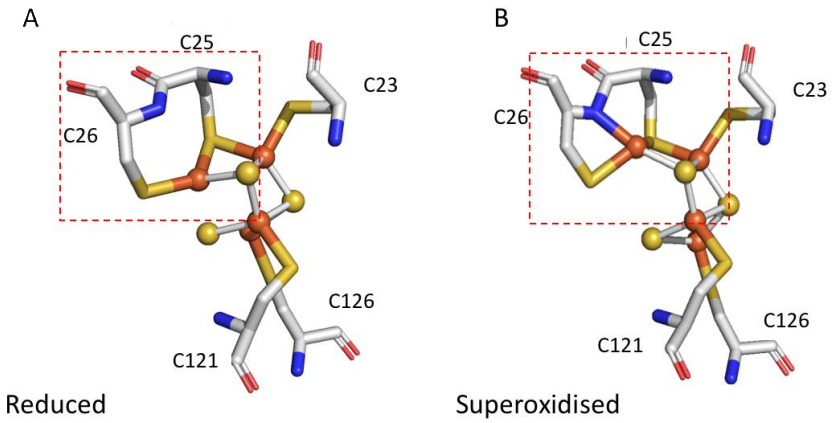


Figure 1.6 Crystallographic structure of the reduced [4Fe-3S]$^{3+}$ (A:PDB 3AYX) and superoxidised [4Fe3S]$^{5+}$ (B:PDB 3AYY) proximal cluster of the O_2-tolerant hydrogenase class, represented here by *Hydrogenovibrio marinus* (*Hm*) hydrogenase.

The unprecedented proximal [4Fe-3S] cluster is thought to be a key component in the O_2-tolerance of these hydrogenases. The special [4Fe-3S] cluster is able to access three redox states with two electrons rapidly transferring to the active site when O_2 is there so that O_2 can be fully reduced to a water molecule with an additional electron from the FeS cluster chain to give 4 electrons in total.[67]

With the help of electron mediators and chemical reductants and oxidants, the redox potential of the FeS clusters can be determined by using electrochemical and spectroscopic techniques.[68,69] Table 1.3 compares the redox potentials of FeS clusters from O_2-tolerant hydrogenases with the standard *D. vulgaris MF* hydrogenase.[70]

From Table 1.3, it is notable that the [4Fe-3S] proximal cluster redox transitions both occur at much higher potentials than the single redox transition of the [4Fe-4S] cubane of standard hydrogenases. The high potential of this cluster also has consequences for stabilization of states during catalysis. Since the cluster will be predominantly reduced at catalytic potentials and therefore less able to accept an electron during turnover.[10]

Table 1.3 Midpoint potentials of FeS clusters in [NiFe] Hydrogenases.

	Fe-S clusters reduction potential / V vs. SHE			
	$[4Fe\text{-}3S]^{5+/4+}$ proximal	$[4Fe\text{-}3S]^{4+/3+}$ proximal	Medial	Distal
E. coli Hyd-1[68,69]	230	30	130	176
A. aeolicus MBH[10]	232	98	–78	–65
Re-MBH[68]	160	–	25	–60

Water and proton channels

In several oxygen-tolerant [NiFe] hydrogenases, it has been shown that H_2 oxidation remains active under ambient O_2, and it is coupled to a continuous formation of water as a product of O_2 detoxification.[61,63,64,71,72] It is thought that the H_2O produced during catalysis cannot escape through the hydrophobic gas-diffusion channels.[73] The crystal structure of *Re* and *Hm* [NiFe] hydrogenases suggest that there is a hydrophilic water-filled cavity close to the [NiFe] active site which is able to transfer the water molecules to the protein surface.[63,65] Small residues replaced by bulky groups such as Val and Tyr in O_2-tolerant [NiFe] hydrogenases result in the formation of such cavities compared to the O_2-sensitive [NiFe] hydrogenases.[61]

In [NiFe] hydrogenases, the mechanisms of proton transfer have not been completely determined. There are several proposed routes based on experimental and computational studies. One of the hypotheses is that the first position to accept protons from H_2 splitting is the terminal cysteine which is coordinated to the Ni ion, and one of the routes for protons is that protons are transferred to the protein surface through the Glu residues located near the proximal FeS cluster.[74] Another hypothetical pathway involves the protons being transferred through the residue Arg located above the active site, and then reaching the protein surface by a His-rich motif.[75]

Gas channels

The enzyme structure is equipped with specific channels to allow diffusion of gaseous substrates including H_2 into the buried active site. However, other gaseous non-polar inhibitors of hydrogenases such as O_2 and CO can diffuse by means of these routes. For this reason, substrate transfer channel selectivity has been considered to be one of the origins of O_2-tolerance.[76] The analysis of the O_2-tolerant regulatory [NiFe] hydrogenase RH from *Ralstonia eutropha* H16 by site-directed mutagenesis indicates replacing bulky 'gating residues' (Isoleucine and Phenylalanine) for the small amino acid residues (Valine and Leucine) expands the putative gas channel, and the variants are sensitive to O_2, presumably due to the active site being exposed to oxygen. This observation illustrates constructing O_2-tolerant derivatives might be achieved by modification of the gas channel.[76,77] However, constricting gas channels of O_2 sensitive hydrogenases did not lead to a clear increase in O_2-tolerance.

1.3.2 The hydrogenase redox states and catalytic cycle

In the binuclear active site of [NiFe] hydrogenase, the Ni atom is coordinated to two terminal cysteines and another two bridging cysteines connect Ni to the Fe atom. Three native ligands, one CO and two CN^-, are further coordinated to the Fe atom.[76] There are two spectroscopically (EPR and FTIR) defined inactive states Ni-A and Ni-B which are also kinetically defined as the Unready and Ready state, respectively.[78] Ni-B and Ni-A are the most oxidised states, both containing a Ni (III) atom. The Ni-B state is formed both under aerobic and anaerobic conditions and it can be activated in a short time (usually within a minute), whereas the Ni-A state is probably only formed aerobically and requires a longer time for re-activation with H_2.[79]

Figure 1.7 shows the electron transfer model for O_2-tolerance of [NiFe] hydrogenases. Ni-B state (Ready) can be formed only if there are enough electrons immediately available when O_2 attacks, whilst Ni-A is formed if not enough electrons are provided and a very slow process for activation proceeds.[80]

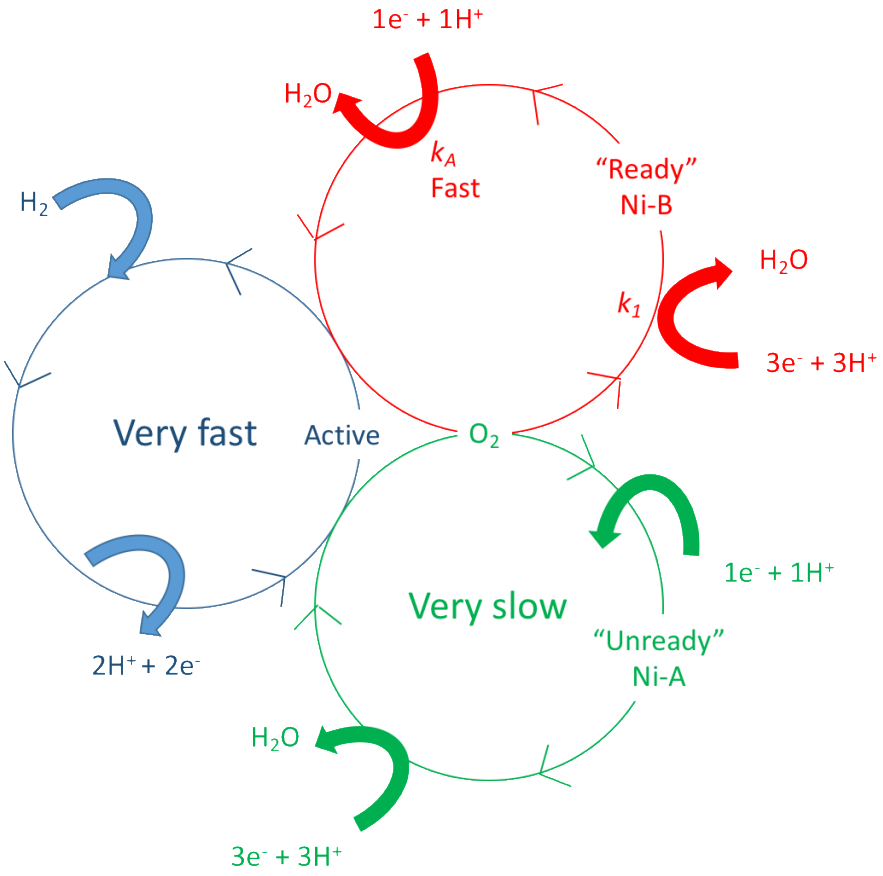


Figure 1.7 Electron transfer model from O_2 tolerance of [NiFe] hydrogenases. Blue indicates fast H_2 oxidation. Red indicates fast catalytic four-electron O_2 reduction producing Ni-B. Green indicates very slow cycle where O_2 is partly reduced, producing Ni-A.[80]

The crystal structure of the Ni-B state from [NiFe] hydrogenases has been determined by Higuchi *et al.,* to consist of a hydroxide ligand bridging the two metals.[81] However, one hypothesis for the Ni-A state was proposed by Volbeda *et al.,* who suggested it also comprises a hydroxide in the bridging position, although orientated differently to Ni-B. Additionally, one of the bridging cysteines is oxidised to sulfenic acid in Ni-A (shown in Figure 1.8).[82]



Figure 1.8 The structure detail of Ni-B and Ni-A active site states from *Desulfovibrio fructosovorans (Df)* [NiFe] hydrogenase.[82]

Figure 1.9 shows the catalytic cycle of [NiFe] hydrogenases. It starts at Ni_a-SI state, where the active site has Ni(II), and an unoccupied bridging coordination site, and this is an EPR-silent species.[66] The Ni_a-SI state is an important state for catalysis and is also the state to which H_2 is thought to bind and form a transient Michaelis complex.[83] When bound H_2 is cleaved during catalysis, a hydride bridge forms between the Ni and Fe giving a state known as Ni_a-R which is also EPR silent with Ni(II). Experimental evidence for these steps in the mechanism was obtained from the crystal structure of the [NiFe] hydrogenase from *Desulfovibrio vulgaris,* where the bridging hydride was first observed.[66] The proposed Ni_a-R state structure is consistent with the theoretical calculation study by Breglia *et al*.[84] Subsequently, the Ni_a-C state is formed by losing one electron from Ni_a-R, while retaining the bridging hydride ligand. Experimental evidence for this state has been obtained by EPR spectroscopy due to the S=1/2 spin state of the Ni(III) ion.[68] There are two pathways to recover Ni_a-SI from Ni_a-C. One route is through another proton coupled electron transfer step, and the other is through an intermediate called Ni_a-L. Ni_a-L is regarded as a tautomeric form of Ni_a-C with only migration of the hydride as a proton.[85] The Ni-L state is observed more readily in the O_2-tolerant hydrogenases, as the high potential FeS clusters (especially the proximal cluster) are likely to be predominantly in the reduced state under the low potential conditions required for Ni_a-C and Ni_a-L states, the retarded electron transfer to form Ni_a-SI results in the Ni_a-L state prevailing.[86] Another cycle of H_2 catalysis begins when Ni_a-SI is recovered.

Figure 1.9B shows one of the proposals for the H_2 activation step in [NiFe] hydrogenases. An arginine (Arg 509) is positioned close to the metal centre which acts as a general base. Hydrogen is polarized between the metal and the guanidine group of Arg 509, oxidising H_2 and generate the Ni-R state.[87]

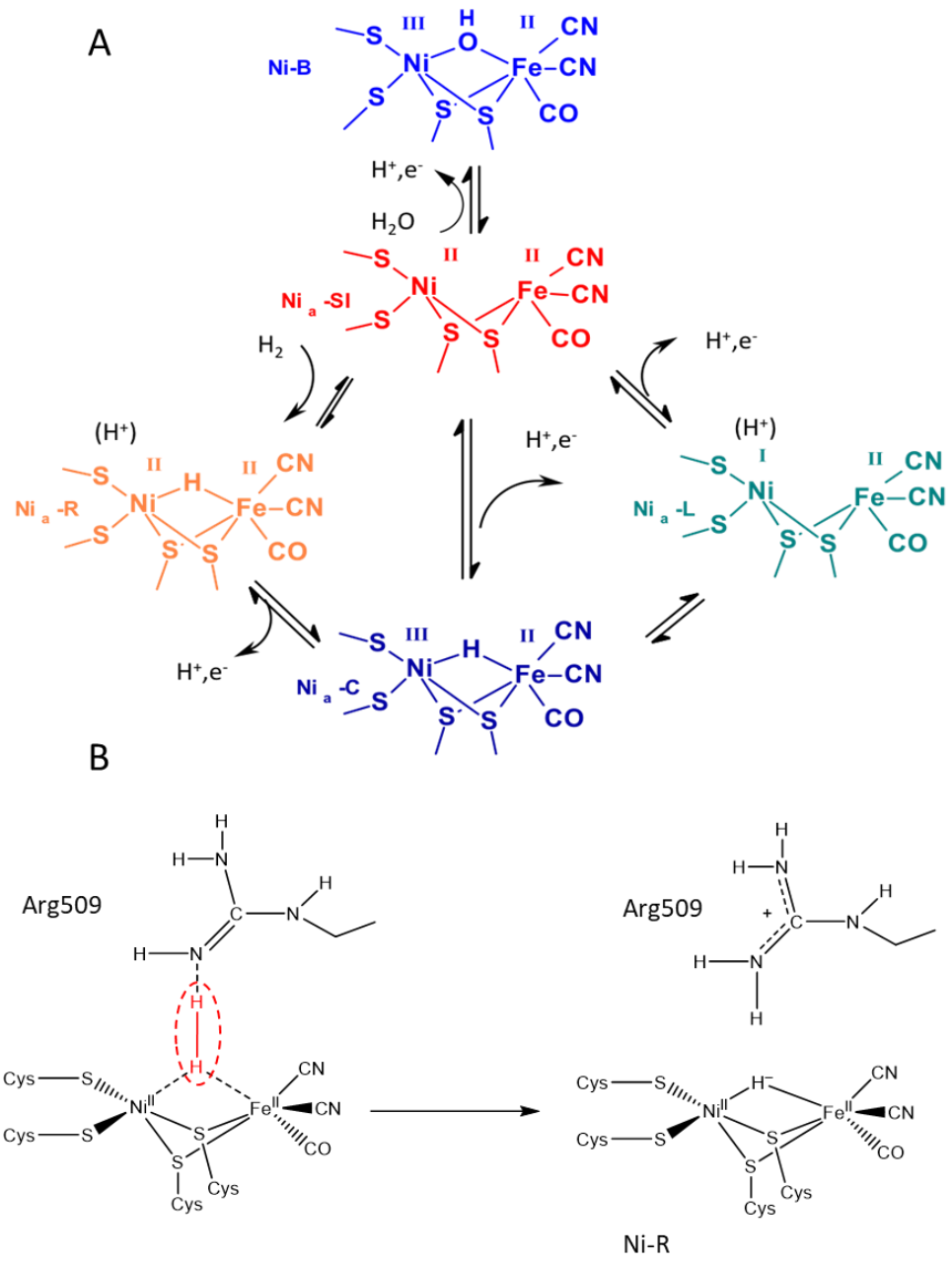


Figure 1.9 Panel A: the skeletal mechanism for H_2 oxidation and production at [NiFe] hydrogenase active sites; Panel B: Hydrogen activation at the catalytic centre of [NiFe] hydrogenases.[83,87]

1.3.3 [NiFe] hydrogenase in *Escherichia coli*

Escherichia coli (*E. coli*) is a facultative anaerobe, which can produce hydrogenases for both H_2 oxidation and H_2 evolution during metabolism.[87] There are three well characterised [NiFe] hydrogenases that have been isolated from *E coli*, named Hyd-1, Hyd-2 and Hyd-3.[88] The former two enzymes are periplasmic membrane-bound proteins, and can only be expressed under anaerobic conditions; while Hyd-3 possesses different metabolic and catalytic activities to both Hyd-1 and Hyd-2. Hyd-3 is an O_2-sensitive component of the membrane-bound cytoplasmic formate hydrogenase lyase complex in group 4a. This complex is active during anaerobic fermentation and mediates the reversible conversion of formate to CO_2, producing H_2 as a by-product. This conversion can balance the acidity of the cytoplasm, and produce dihydrogen.[89]

Hyd-1 is a membrane protein with a large subunit HyaA containing the Ni-Fe active site, a small subunit HyaB encompassing the FeS cluster relay, and an integral membrane *b*-type cytochrome (HyaC) shown in Figure 1.10.[66] Under slow growth conditions and fluctuating O_2 concentrations, H_2 oxidation in Hyd-1 is coupled to reduction of the quinone pool by means of HyaC, creating a proton gradient.[19,90] Based on this physiological feature, Hyd-1 shows almost no activity for H_2 production even at very low potentials except at low pH.[59,91]

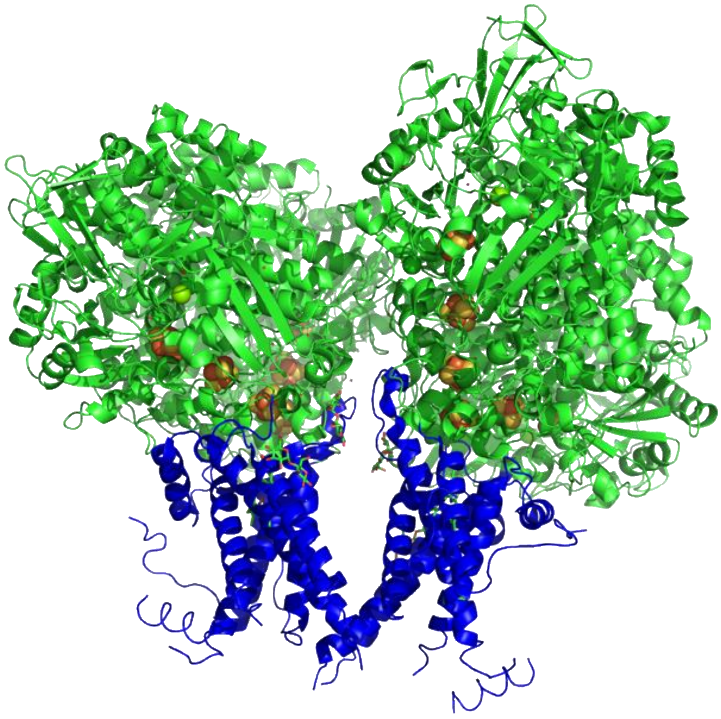


Figure 1.10 Structure of *E. coli* [NiFe] hydrogenase 1 (Hyd-1) in complex with cytochrome *b*. The green part is the Hyd-1 large and small subunits HyaA and HyaB, the blue helical protein is the cytochrome *b*. The FeS clusters and the NiFe active sites are shown in the middle represented by spheres (cartoon generated with PyMOL, PDB: 4GD3).

1.4 FeFe hydrogenase

Another class of hydrogenases is the [FeFe] hydrogenases. These enzymes are found in different types of prokaryotes, such as *Clostridia* and sulfate-reducing bacteria, as well as some eukaryotes including *Cilates, Trichomonads* and some green algae and anaerobic fungi.[38,52] [FeFe] hydrogenases play a role in both the reductive and oxidative reactions of hydrogen. In green algae and *Clostridia*, these enzymes allow protons to be used as an electron sink during the oxidation of substrates. In sulfate reducing bacteria, for example *Desulfovibrio desulfuricans*, [FeFe] hydrogenases provide electrons for reduction reactions. These are the two main physiological functions of [FeFe] hydrogenase.[19]

1.4.1 The structure of FeFe hydrogenase

There are five subtypes of [FeFe] hydrogenases. The simplest one known is the M1-type, which is comprised of just the catalytic core complex (H-domain).[19] The H-cluster is buried inside the enzyme in the prokaryotic hydrogenases. While the second type (M2) requires the F-cluster; a protein moiety that contains additional FeS clusters (either [4Fe-4S] or [2Fe-2S]) connected with the H-cluster. The F-cluster serves as an electron transfer relay connecting the deeply buried H-cluster with the protein surface.[97] The third type (M3) of [FeFe] hydrogenases consist of two N-terminal cluster binding motifs. One motif contains a [4Fe-4S] cluster, while the other contains a [2Fe-2S] cluster.[92] In the M4 subtype, an additional C-terminal domain (NuoE) containing a single [2Fe-2S] cluster allows direct coupling of NADH/NAD^+-turnover with H_2/H^+ exchange. In the M5 subtype, an additional domain containing a [4Fe-4S] cluster, NuoF, is bound to NuoE.[93,94,95]

The M2-subtype *Desulfovibrio desulfuricans* [FeFe] hydrogenase (herein termed *Dd*HydAB) is one subject of study in this thesis. *Dd*HydAB contains two subunits: the small subunit is composed entirely of α-helices without any metal centres and is surrounded by the large subunit. The large subunit consists of the H-cluster and two additional [4Fe-4S] clusters (Figure 1.11).[92]

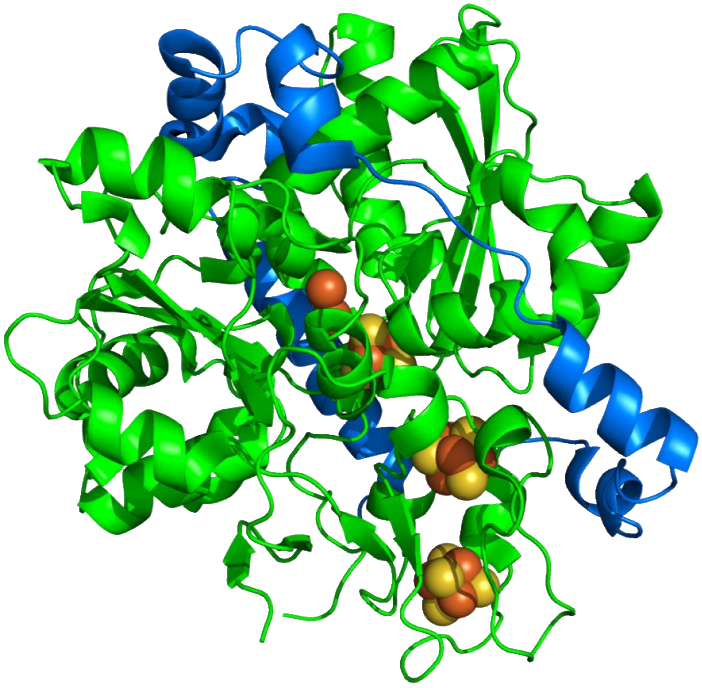


Figure 1.11 The overall structure of *D. desulfuricans* HydAB, the green cartoon shows the large subunit HydA and the blue cartoon shows the small subunit HydB. The FeS clusters and $[2Fe]_H$ are represented by spheres, with the yellow ones being the S atoms and the red ones being the Fe atoms. (PDB code: 1HFE).

Active site

All [FeFe] hydrogenases share the same common element: the H-domain which is about 40 kDa and harbours the catalytically active metal centre, H-cluster. The H-cluster comprises a unique $[2Fe]_H$ sub-cluster and a $[4Fe\text{-}4S]_H$ cluster.[95] In the $[2Fe]_H$ sub-cluster, the two iron ions are called proximal (Fe_p) and distal (Fe_d) according to the distance to the $[4Fe\text{-}4S]_H$ cluster, and they are bridged by a disulfide ligand with a N bridge-head (azadithiolate, $[2Fe]^{adt}$) and a CO ligand in some states.[93] Additionally, each iron is further coordinated by a CO and a CN^- terminal ligand (Figure 1.12).[93] The CO ligands are located in a hydrophobic pocket. One of the irons from the special [2Fe] sub-cluster Fe_p is connected to a canonical [4Fe-4S] cluster via a covalent sulfur bridge from a cysteine residue, which also connects the active site to the protein matrix.[92]

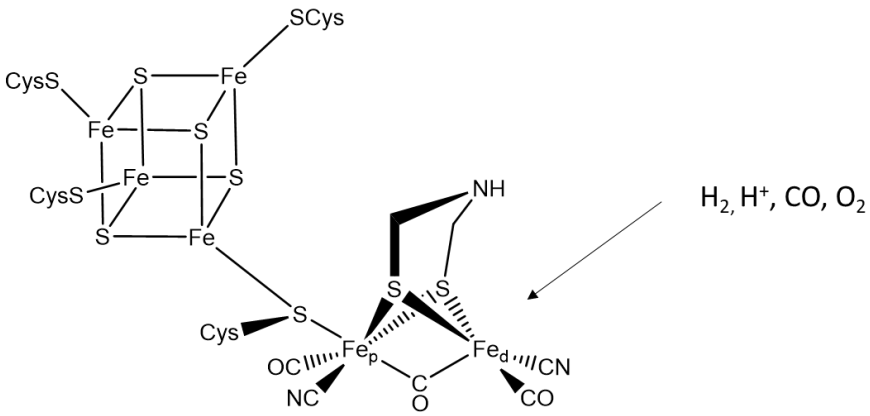


Figure 1.12 The structure of H-cluster in [FeFe] hydrogenases showing the four cysteines coordinating the [4Fe-4S] cluster. The arrow shows the open coordination site where inhibitors or substrate can bind.

Accessory FeS clusters as electron relay

All FeS clusters found in [FeFe] hydrogenases are situated within 11 Å of each other. This distance is suited for electron transfer, creating an electron transfer chain.[6] Dihydrogen can be cleaved in the active site and the resulting electrons are transferred through the FeS cluster relay to the protein surface, or the reverse.[96]

Figure 1.13 shows the schematic depiction of the arrangement of different FeS clusters in M1-M3 subtype [FeFe] hydrogenases. The smallest M1-subtype [FeFe] hydrogenases, e.g., *Cr*HydA1 (*Chlamydomonas reinhardtii*) found in a sub-class of green algae, contains only the H-domain.[97] For the M2-subtype [FeFe] hydrogenase such as *Desulfovibrio desulfuricans* (*Dd*HydAB), apart from the [4Fe-4S] cluster in the H-domain (Figure 1.12), there are two additional [4Fe-4S] clusters housed in the F-domain, termed the proximal cluster and distal cluster according their distance from the H-cluster.[98] M3-subtype [FeFe] hydrogenases such as *Clostridium pasteurianum* (*Cp*I) consist of two N-terminal cluster binding motifs: one contains a [4Fe-4S] cluster linked by three cysteines and one histidine residue while the other contains a plant-type ferredoxin [2Fe-2S] cluster.[94]

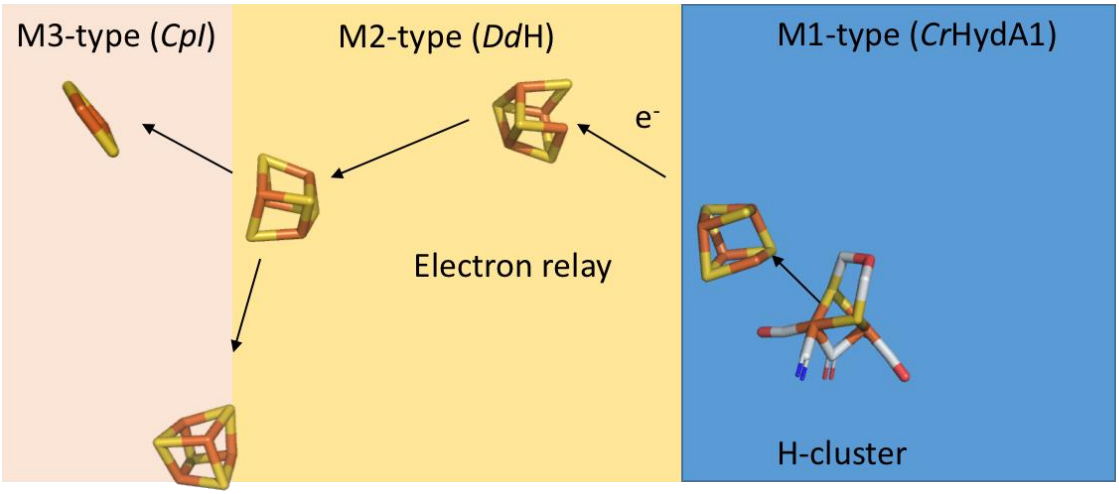


Figure 1.13 Schematic representation of *C. pasteurianum (Cp*I*)* [FeFe] hydrogenase showing the catalytic site with FeS cluster relay. The different colours represent the different sub-type of [FeFe] hydrogenase. PDB code: 3C8Y.

Water and proton channels

In [FeFe] hydrogenases, an efficient proton transfer pathway is required to transport protons between bulk water and the catalytic active site. Several putative proton transfer pathways have been investigated according to the structure of *Dd*HydAB and *Cp*I.[98] The results suggest that the protons are transferred via three conserved glutamates, one serine, a water molecule and a cysteine residue. This cysteine residue close to the nitrogen atom from the bridge-head of the H-cluster is the preliminary proton acceptor from the active site.[99] Duan *et al.* demonstrated that substituting any of these four conserved residues in the proton transfer pathway Cys299, Glu279, Ser319 and Glu282 (numbering corresponds to *Cp*I) caused a dramatic reduction in hydrogenase activity.[100]

Gas channels

Like the [NiFe] hydrogenases (see section 1.3.1 for more detail), [FeFe] hydrogenases have a buried active site; therefore, a gas channel is essential for the catalytic cycle.[92,101] The substrate H_2 or known inhibitors CO and O_2 access the catalytic centre of the protein through the

hydrophobic channel and then bind to the distal iron in the active site (shown in Figure 1.12).[102] The gas channel is shorter than that those found in the [NiFe] hydrogenases based on the crystal structure of [FeFe] hydrogenases, which could be one of the reasons for the higher turnover rate of the [FeFe] hydrogenases.[19]

1.4.2 FeFe hydrogenases catalytic cycle

The catalytic cycle of [FeFe] hydrogenases is still highly disputed between two proposed models.[103,104] Figure 1.14 shows one of the predominant models and summarises the key redox states in the cycle. The most well studied enzyme is *Cr*HydA1, and the main identified states come from IR spectroelectrochemical studies.[105] In the H-cluster, the redox-coupled and proton-coupled electron transfer (PCET) between the $[2Fe]_H$ and $[4Fe\text{-}4S]_H$ is essential for catalytic activity.[104] In the oxidised state H_{ox}, the binuclear site has a mixed valence redox state $Fe_p(II)Fe_d(I)$ with the oxidised form (2+) of the $[4Fe\text{-}4S]_H$ cluster. H_{red} can be formed by one electron reduction of the $[4Fe\text{-}4S]_H^{2+}$ cluster, where the binuclear active site remains in a mixed valence configuration. The H_{red} state undergoes protonation at the bridgehead N-atom of $[2Fe]^{adt}$, resulting in the formation of the $H_{red}H^+$ state which then undergoes electron transfer from the $[4Fe\text{-}4S]_H^+$ to the binuclear subsite ($Fe_p(II)Fe_d(I)$ to $Fe_p(I)Fe_d(I)$). Subsequent reduction of the $[4Fe\text{-}4S]_H^{2+}$ forms the so-called $H_{sred}H^+$ state.[106] The hydride-bound H_{hyd} state is believed to be formed from $H_{sred}H^+$ due to the transfer of the H^+ from the pendant amine base to Fe_d. The [4Fe-4S] cluster remains reduced whilst the [2Fe] sub-cluster gives two electrons to form a terminally bound hydride and ends up in a homovalent $Fe_p(II)Fe_d(II)$ state (denoted H_{hyd}). The final step of the cycle involves another proton-coupled electron transfer (PCET), where electron transfer from the [4Fe-4S] sub-cluster to the binuclear sub-cluster is coupled to protonation of the nitrogen bridge of $[2Fe]^{adt}$ followed by the proton and hydride combining to form H_2 and releasing from the active site, giving back the H_{ox} state for the next cycle to begin.[107]

Furthermore, the H_{ox} state can bind exogenous CO forming the inactive H_{ox}-CO state. Then, one electron reduction of the$[4Fe\text{-}4S]_H^{2+}$ cluster from the inactive H_{ox}-CO state yields the H_{red}-CO state.[108]

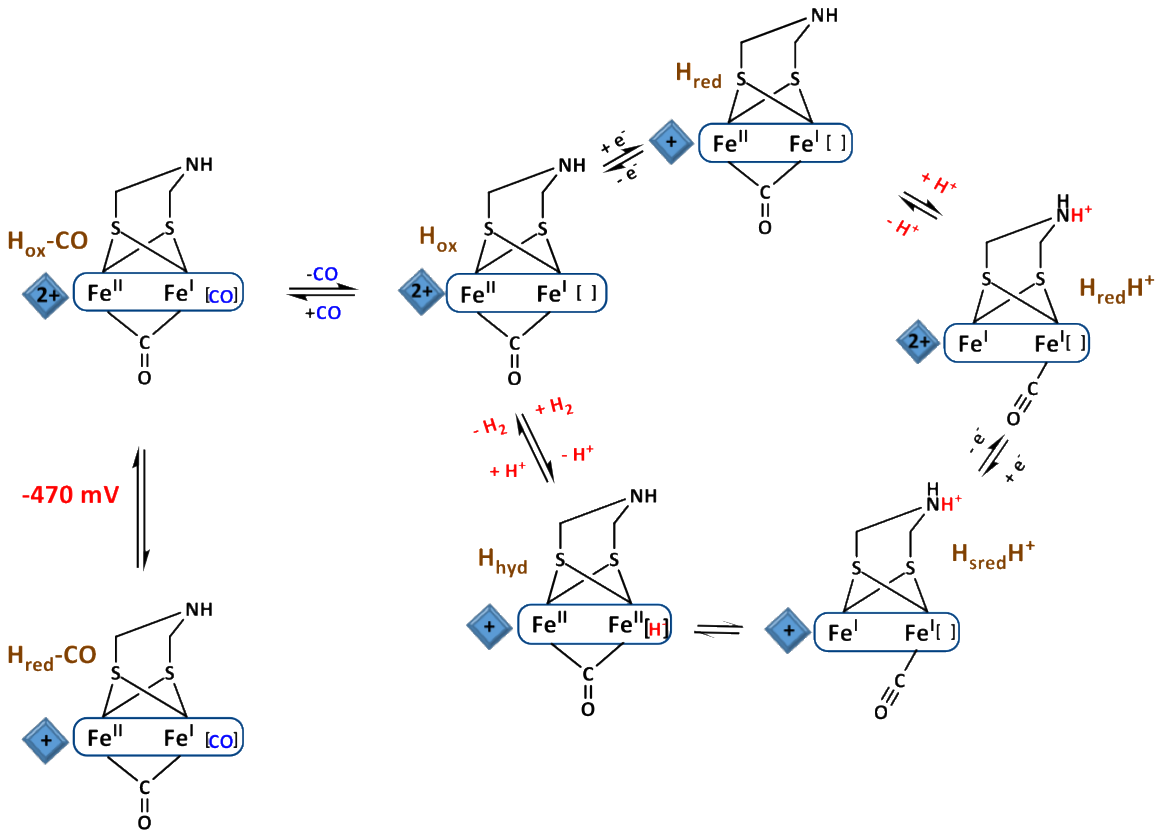


Figure 1.14 Catalytic cycle of [FeFe] hydrogenases showing the active states (H_{ox}, H_{red}, $H_{red}H^+$, $H_{sred}H^+$ and H_{hyd}), as well as the CO inhibited inactive states (H_{ox}-CO and H_{red}-CO). The $[4Fe\text{-}4S]_H$ cluster is represented as ◆, the additional CO and CN^- ligands are left off for simplicity.[19]

1.5 Techniques for studying hydrogenases

For the characterisation of catalytic steps and intermediate species in hydrogenase activation and catalysis, several biochemical and biophysical methods are used. All understanding including the structures and mechanisms of hydrogenases so far published in the literature have been revealed by various methods, as briefly discussed below. Within these methods, spectroscopy plays a significant role, in particularly for the study of electronic structure and dynamic aspects.

Infrared spectroscopy

Infrared spectroscopy, as applied to hydrogenases, is a biophysical technique based on the absorption of mid-infrared light in the range from 400 to 4000 cm^{-1}. The absorption of infrared radiation by proteins gives pivotal information on their structures and functions which is due to its sensitivity to the chemical composition and architecture of the molecules.[109] The most intense bands come from the overlap of all the peptide bond absorptions, for example amide I and amide II. Amide I (at about 1650 cm^{-1}) mainly arise from the amide C=O stretching vibration; amide II vibration absorbing near 1550 cm^{-1} mainly results from amide CN stretching and NH in-plane bend. These two vibrations are barely affected by the side chain.[110] Additionally, some metalloproteins contain prosthetic groups which have intrinsic ligands or bound substrates that enable intense infrared absorption at characteristic wavelengths.

For the study of hydrogenases, the active site contains built-in biologically unusual ligands CO and CN^- that have intense infrared absorption features. The ν_{CO} and ν_{CN} stretching bands are very sensitive to changes in the electron density in the surrounding environment, and are in the 2200–1800 cm^{-1} region.[111] Therefore, they are versatile IR probes for the detection of changes in coordination and electronic environment of the active site including protonation and deprotonation events.[66] Moreover, due to the technique being easily adapted for physiologically relevant conditions of IR measurements, i.e., at room temperature, neutral pH and ambient pressure environments, IR spectroscopy is regarded as an advantageous technique for the study of hydrogenases. This versatility in assaying conditions makes IR a highly useful technique for the study of hydrogenases, for which other techniques cannot be utilised. For example, the difficulties that exists in getting definable and homogeneous crystals for X-ray crystallography, or the fact that electron paramagnetic resonance (EPR) spectroscopy can only be conducted under frozen conditions for FeS proteins means that it is difficult to study transformations during catalytic turnover with these techniques.

1.5.1 Combining techniques to clarify enzyme mechanism

In order to gain a unifying understanding of the mechanisms of complex redox enzyme such as hydrogenases, kinetic information (e.g., turnover frequencies and inhibition rates) has to be considered and studied alongside structural and spectroscopic results. Often, it is difficult to reconcile information obtained by a variety of techniques on samples in diverse physical forms.

Measurements of the activity of an enzyme are often carried out using a biochemical assay in solution (such as H_2-oxidising activity of Hyd-1 with benzyl viologen) or with electrode-immobilised enzyme using PFE.[112,113] Additionally, spectroscopic techniques such as EPR are useful techniques to report and study the features of the metalloclusters in enzymes, but these require specific conditions (e.g. frozen samples). Likewise, the atomic-level structure is obtained from X-ray diffraction of the solid crystalline metalloprotein and is thus difficult to link to turnover conditions. Therefore, it is important to combine different techniques to clarify the understanding of the metalloprotein.[114]

Protein film IR electrochemistry (PFIRE)

The research group of Prof. Kylie Vincent (University of Oxford) has established a method to combine PFE at carbon-based electrodes with infrared spectroscopy: Protein Film IR Electrochemistry (PFIRE).[115] This technique uses a multiple-reflection attenuated total reflectance geometry to probe a redox protein immobilised on a high surface area carbon electrode for *in situ* study.[116] The enzyme absorbed carbon acts as a macroscopic working electrode, whose high electrode surface area is pivotal to provide enough protein in the IR sampling area for spectroscopic detection of the catalytic active site via the CO and CN^- vibrational bands. Incorporation of a three-electrode system allows precise potential control, while use of a liquid flow cell allows solution pH and solute concentrations to be controlled during the measurements. At the same time, electrochemistry can be used to control the redox states of the hydrogenases in the presence or absence of H_2 saturated buffer. Rapid flow

through the cell can be used to maintain fast catalysis by providing a supply of fresh substrate H_2.[115]

A limitation of PFE is that it is unable to provide information on non-catalytic processes in the redox enzyme unless individual electron transfers to the FeS clusters can be detected. PFIRE provides a significant mechanistic insight into hydrogenases: it can be used to identify and measure under steady state conditions and thus provide detail on redox states present during catalysis in the presence of H_2. Under different conditions, i.e. in the absence of H_2, it can be used to identify and measure information on the active site redox states and monitor their transitions.[114,116] The PFIRE technique has been used extensively with Hyd-1 as a case to demonstrate structure details of enzyme active site chemistry during both non-turnover and catalytic conditions.[117] PFIRE studies of Hyd-1 first revealed that the active site Ni_a-L state (Figure 1.9) can be generated reversibly in the dark at ambient temperature under both turnover and non-turnover conditions, and it is observed at all potentials during catalytic turnover.[116] Murphy *et al.* first demonstrated that the Ni_a-C/Ni_a-L interconversion in Hyd-1 under non-turnover conditions is a pH-dependent process in the dark via PFIRE and EPR.[85] The ability to reconcile the sustained presence of Ni_a-L with the reduced state of the high potential proximal cluster required EPR because PFIRE cannot tell the redox state of the FeS clusters.[118,120] However, due to the frozen environment of the EPR experiment, only non-turnover conditions were studied.

Single crystal micro-spectroscopy coupled to electrochemical control.

Infrared micro-spectroscopy permits spectra to be obtained from single protein crystals. These can provide higher signals in the IR because of the high effective protein concentration in the crystals compared to solution protein. Many previous structural studies have focused on inactive or inhibited states of redox enzymes.[119] The research group of Prof. Kylie Vincent has developed a custom-designed IR micro-spectroscopic electrochemical cell, enabling the generation of well-defined redox states in single crystals (more details of the cell design are given in Chapter 2).

The crystalline state can give insights into a variety of functions of redox enzymes. For Hyd-1, combining IR spectroscopy and electrochemical control allows probing chemical transformations of the [NiFe] active site within a single crystal. The electrochemical control over the system allows for fine potential control of redox states and allows equilibration of redox states of the active site under non-turnover conditions. More importantly, the peak positions and redox potentials of each redox state from the single crystal are consistent with those previously reported from samples studied in solution and PFIRE experiments.[109,114] Using this combination technique, fully reversible electrochemical reduction from the oxidised inactive Ni-B state to the fully reduced state Ni-R states in *E. coli* Hyd1 has been demonstrated.[109]

1.6 A novel IR vibrational reporter of protein environments

Infrared spectroscopy plays an indispensable role in the study of structure and dynamics in many biological systems. In proteins, the amide I mode which is between 1600–1700 cm^{-1} arises mainly from the C=O stretching vibration; however, it is difficult to provide site-specific information regarding the structural or environmental detail. To overcome this limitation, extensive efforts have been dedicated to the development and application of extrinsic site-specific infrared IR probes. In addition, the past few years have seen rapid development in the incorporation of unnatural amino acids into proteins via *in vitro* and *in vivo* methods. The genetic incorporation of unnatural amino acids into proteins has been proven a potent method for investigating the structure and function in biological systems. Unnatural amino acids expand the toolkit of mutagenesis, providing a novel approach to study the local protein structure and dynamics. The use of unnatural amino acids allows the incorporation of residues with unique side chains such as fluorescent groups, vibrational probes and NMR reporters which can be used in different fields.[120]

Unnatural amino acids which can be used as site-specific IR probes are listed in Table 1.4. Nitriles and azides are effective vibrational probes in part because of the position of the azide

asymmetric stretch vibration and the nitrile stretch vibration which occur in a relatively clear region of the infrared spectrum.[121]

Table 1.4 Overview of unnatural amino acids as site-specific IR probes

Name		Frequency range (cm^{-1})	Ref
Nitriles	*p*-cyano-phenylalanine	2220–2250	122
	Cyano-cysteine	2150–2180	123
	5-cyano-tryptophan	2210–2240	124
Azide	Azidohomoalanine	2100–2140	125
	p-azido-phenylalanine	2080–2160	126

The nitrile group C≡N has been utilised as a probe of ligand binding, local environment and electric field in proteins. The nitrile probe has received great attention experimentally and computationally, it can be used to characterise the microenvironments of small molecule ligands, free amino acids or peptides.[127,128] More importantly, the nitrile group stretching vibration is a sensitive IR reporter and enables the detection of subtle changes in local environment as well as alterations due to nitrogen interaction with metal ions.

Azides (R-N_3) are also actively used as a local environmental reporter. Compared to nitriles, the 2–4 fold higher extinction coefficient enables its application at lower protein concentrations and with a higher signal-to-noise ratio.[129] However, there are some drawbacks of the azido moiety as an IR reporter, e.g., azido probes usually have broad bands and are not as sensitive as nitrile to its environment. Additionally, because azides are thought to participate in Fermi resonance, the nitrile IR absorbance profile is generally less complicated to interpret.[129]

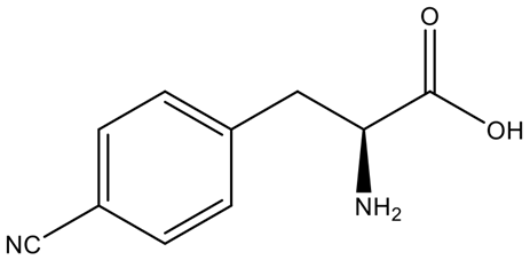


Figure 1.15 Structure of the unnatural amino acid *p*CNF.

Among three unnatural nitrilo amino acids in Table 1.4, the most widely used vibrational probe is *p*-cyano-phenylalanine (*p*CNF, (Figure 1.15) due to the higher expression yield and the *p*CNF having a strong absorption in a region of the spectrum (2200–2250 cm^{-1}) which is free of other vibrational absorptions such as those of amide groups in proteins and CO (carbon monoxide) and CN^- (cyanide) groups of the hydrogenase active site.[130] Schultz and co-workers have developed methods for the site-specific introduction and developed an infrared probe *p*CNF (Figure 1.15) into proteins and driven the expansion of the genetic code through amber codon suppression.[131] An unnautral amino acid can be effectively site-directly introduced at defined sites in proteins in bacteria by using an orthogonal nonsense suppressor tRNA/aminoacyl-tRNA synthetase (aaRS) pair derived from the *Methanococcus janannaschi (Mj)* tyrosyl tRNA/aminoacyl-tRNA synthetase (aaRS) pair was utilised to ensure that the unnatural amino acid is uniquely incorporated into a specific site.[132] This pair is able to recognise the amber nonsense codon UAG in the messenger RNA which was used to carry the unnatural amino acid *p*CNF into the ribosome in *E. coli*. As it is the rarely used stop codon, the presence of natural amber suppressors in some *E. coli* strains does not significantly affect bacterial growth rates.[133,134] Then the amino acid sequence is elongated to get the full-length protein. The process of protein synthesis with *p*CNF incorporated is shown in Figure 1.16. [127, 135, 136]

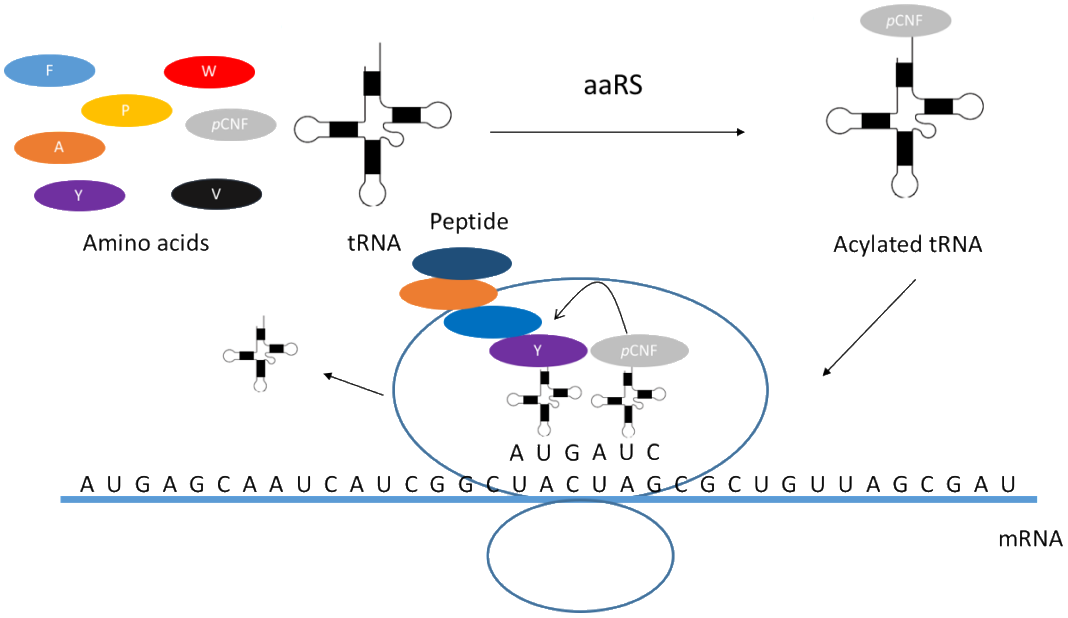


Figure 1.16 The process of protein synthesis in the ribosome with *p*CNF. Each amino acid is acylated to its cognate tRNA by a specific aminoacyl-tRNA synthesis (aaRS) and then transferred to the ribosome. According to the codon-anticodon interaction between mRNA and acylated tRNA, the tRNA bound *p*CNF is linked to the adjacent amino acid, extending the polypeptide and releasing the free tRNA.

Unnatural amino acids such as *p*CNF could be incorporated into proteins utilizing an orthogonal engineered aminoacyl-tRNA synthetase binding with the amber codon (TAG) in mRNA.[127,136]

Successful incorporation of *p*CNF has been demonstrated in several systems, such as myoglobin,[135] cytochrome *c*,[137] and super-folder green fluorescent protein (sfGFP)[122] and it has been shown that the $\nu(CN^-)$ IR absorptions are observable and sensitive to their protein environment. Only single positions were examined, and the function of the protein was not disturbed.

Völler *et al.* have probed the local electric field by incorporating *p*CNF into cytochrome *c* variants with the method developed by Schultz and co-workers. Their studies revealed that the $\nu(C\equiv N)$ stretch from *p*CNF was sensitive to the redox states of cytochrome *c*, with a 1.5 cm^{-1} shift between the oxidised and reduced state of the cytochrome *c* variants, thus indicating that the CN group is sensitive to changes in the electrostatics of cytochrome *c*.[127]

1.7 Aim and scope of this thesis

Iron-sulfur (FeS) clusters are essential cofactors in all types of life, functioning as redox centres in many significant biochemical reactions, including mitochondrial respiration. This thesis develops an approach to probe FeS cluster redox states and their potential-dependence by incorporating an unnatural amino acid close to the redox centre to allow the study of the cluster, indirectly, via IR spectroscopy. The ultimate goal is to apply this technique to complex redox enzymes such as hydrogenases. Hydrogenases are the most active molecular catalysts for H_2 activation and use earth-abundant metals. However, many open questions remain in hydrogenase chemistry and catalytic mechanism, demanding the development of novel techniques to resolve ambiguities. One of the important unknowns in hydrogenase catalysis is how the FeS clusters and the prosthetic group work together during catalytic turnover. IR is a useful tool for studying the active site of hydrogenases because of the active site CO and CN^- ligands (details are shown in sections 1.3 and 1.4). However, conventional methods to study FeS clusters like EPR and Mössbauer spectroscopy require frozen samples and therefore cannot be used for operando detection during the catalytic turnover. FeS cluster do not absorb in the mid-IR and give absorption bands towards the far-IR in a region which is complex to interpret. We therefore developed a novel approach to study FeS cluster by IR spectroscopy, to study how FeS cluster redox chemistry interplays with the active site chemistry in complex metalloenzymes like hydrogenases during catalysis via incorporation of the mid-IR absorbing amino acid *p*-cyano-L-phenylalanine (*p*CNF). Towards this end, several methods are employed herein.

In this thesis, several molecular biology techniques are implemented to develop protein over-production systems and mutagenesis of specific residues, achieving high yields of highly pure proteins and incorporating the IR-probe-containing unnatural amino acid into the protein. Physical chemistry methods such as protein film electrochemistry are used to characterise the native and variant proteins, and IR spectroscopy is used to study the system in detail to

determine whether the C≡N stretch of *p*CNF can be used to probe the oxidation state of FeS clusters.

In this thesis, a novel approach to probe the FeS clusters is developed and demonstrated, incorporating an unnatural amino acid containing IR reporter into an FeS protein, close to the FeS cluster in space. This method is applied to spinach ferredoxin which is a low molecular weight protein containing a single [2Fe-2S] cluster. The construction of an over-production system for spinach ferredoxin via an expression plasmid and using *E. coli* as the host cells enables improvement of the protein yield. Once an over-production method was established, incorporation of the *p*-cyano-phenylalanine into spinach ferredoxin by using molecular biology methods was possible (see Chapter 3 for more detail).

In Chapter 4, infrared spectroscopy was carried out on unnatural ferredoxin variant Y37CNF to determine if *p*CNF is able to respond to different redox states of the FeS cluster via IR spectroscopy. To study the successful incorporation of *p*CNF into spinach ferredoxin and confirm minimal structural disruption compared with wild type, particularly near the FeS cluster site, wild-type spinach ferredoxin and site-directed variants (incorporating both native and non-native amino acid resides) were crystallised. The crystal structure of wild type spinach ferredoxin was solved at high resolution, followed by that of variants Y3F, Y37F and a variant containing the unnatural *p*CNF residue. The crystallisation and refinement were carried out at the Research Complex at Harwell by Dr Stephen Carr (University of Oxford, Department of Chemistry).

To apply the same method of FeS redox probe into a complex redox enzyme, namely *E. coli* [NiFe] hydrogenase 1, it was first required that an over-production system for Hyd-1 was tested. Current methods to produce Hyd-1 rely on chromosomal expression under the control of native promoters. In the group of Professor Kylie Vincent, an over-expression of this enzyme was recently developed. This over-production method was adapted for the incorporation of unnatural amino acids. Chapter 5 discusses the characterisation of the over-expressed Hyd-1

using conventional solution assays, protein film electrochemistry and infrared spectroscopy. The over-expressed Hyd-1 was successfully produced with improved yield, and characterisation results illustrated that the function and properties of the overexpressed Hyd-1 are essentially the same as the natively expressed enzyme. Amino acid residues were then selected for incorporating the unnatural amino acid after testing tolerance to mutation of certain sites by site-directed mutagenesis. It was decided not to continue with the [NiFe] hydrogenase because expression yields are still relatively low and the over-expressed protein did not crystallise. [FeFe] hydrogenase was selected instead due to better expression.

In Chapter 6, *p*-cyano-L-phenylalanine (*p*CNF) was incorporated into the [FeFe] hydrogenase from the sulfate reducing bacterium *Desulfovibrio desulfuricans* (*Dd*HydAB). Since this enzyme contains two additional [4Fe-4S] clusters in addition to the active site H-cluster, residues around the proximal cluster were tested for mutagenesis suitability prior to *p*CNF incorporation. Subsequently, incorporation of the *p*CNF into *Dd*HydAB by using molecular biology tools (the same as in Chapter 3) was carried out. The catalytic active site and FeS cluster properties were studied using *p*CNF as a probe and are presented in Chapter 6.

Overall, this thesis establishes *p*CNF as a viable reporter for FeS cluster redox state in complex metalloenzymes.

2

Chapter 2: Experimental Methods

2.1 Molecular biology methods

2.1.1 Strains, plasmids and oligonucleotide primers.

Strains used in this work are listed in Table 2.1. Modifications to develop new expression strains were carried out by Jiaao Wei (JW), Dr Miguel Ramirez (MR), Dr Rhiannon Evans (RE) and Dr Patricia Rodriguez Macia (PM) as noted. Plasmid vectors used in the experiments are listed in Table 2.2.

Table 2.1 Strains used in this research and description

Strain	Description	Source
DH5α	High Efficiency Cloning strain	ThermoFisher Scientific, U.K.
BL21(DE3)	The strain contains the lambda DE3 prophage that carries the gene for T7 RNA polymerase under control of a lacUV5 promoter. Specifically constructed for high-level expression of recombinant proteins	ThermoFisher Scientific, U.K.
BL21(DE3)*ΔiscR*	BL21 cell with a mutation to knock out the Fe/S cluster synthesis regulation genes	From P.R. Jones and co-workers[138]

Table 2.2 Vector used in this research

Plasmid	Description	Source
pGEX-6P-1	Ampicillin, bacterial vector for expressing GST fusion proteins with a PreScission protease site.	Sigma Aldrich
pACYC-Duet-1	Chloramphenicol, pACYCDuet-1 is designed for the coexpression of two target genes. It carries the P15A replicon and the *lac*I gene and chloramphenicol resistance gene.	Merck Millipore

Oligonucleotide primers for amplification of DNA or mutagenesis in PCR were obtained from Invitrogen Life Sciences. On arrival the lyophilised oligonucleotides were dissolved in H_2O to a concentration of 100 μM and stored frozen at –20 °C for long term storage. Oligonucleotide primers used in polymerase chain reactions in this research are listed in Appendix A.

2.1.2 Cell growth and culture

Culture medium and antibiotics for cell growth

In this research, two main media compositions were used, Lysogeny Broth (LB) and 2xYT. The main medium used for cell growth is LB, a pre-made mixture purchased from Merck and prepared to a concentration of 25 g L^{-1} in distilled water, giving final concentrations of 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract and 10 g L^{-1} sodium chloride (NaCl). All LB media was autoclaved prior to use. To increase the expression level of the recombinant protein, LB media was replaced with the rich medium 2xYT (prepared in house), which contained additional carbon and nitrogen sources, specifically 16 g L^{-1} tryptone, 10 g $^{-1}$ yeast extract and 5 g L^{-1} NaCl. Solid media plates were made using LB medium, and agar was added to a concentration of 15 g L^{-1} prior to autoclaving. Where required, four different antibiotics were used: chloroamphenicol, ampicillin, kanamycin, and streptomycin. Chloroamphenicol (Duchefa Biochemie, > 98%) was used at a final concentration of 34 μg mL^{-1} from a stock solution of 34 mg mL^{-1} in 100 % ethanol (1000X). All other antibiotics were dissolved in water for a final concentration of 100 μg mL^{-1} for ampicillin, 50 μg mL^{-1} for kanamycin, and 75 μg mL^{-1} for streptomycin.

Glycerol Stocks

All strains produced for this thesis were preserved as glycerol stocks using the following protocol. Once a strain was produced (or obtained) and verified, a single colony was inoculated in 10 mL of LB medium supplemented with antibiotic when needed and grown at 37 °C with shaking. After overnight incubation (about 16 hours), the overnight culture was mixed with 50%

(*v*/*v*) glycerol (Thermo Fisher Scientific), which was prepared with distilled water and autoclaved before use. The mixture of cells and glycerol was flash-frozen in liquid N_2 and stored at –80 °C. All plasmids were also preserved in DH5α cells as glycerol stocks. A single transformant was picked into LB media supplemented with the appropriate antibiotic and grown at 37 °C with shaking overnight. The cells were preserved in glycerol as described above.

2.1.3 Methods for genetic engineering

Preparation of competent cells

For transformations using DNA from a plasmid(s) (Table 2.1), competent cells were prepared in-house. LB medium (10 mL) was inoculated with a single colony and grown overnight at 37 °C with shaking. This culture (50 μL) was used to inoculate fresh LB medium (50 mL) and grown to an optical density at 600 nm (OD600) of 0.4 to 0.6 in a sterile conical flask. This culture was incubated on ice for 30 minutes before harvesting the cells by centrifugation at 4 °C for 10 minutes at 3000 *xg*. The cell pellet was resuspended in buffer TFB I (Table 2.3) and incubated on ice for a minimum of 10 minutes. The cells were harvested using a further centrifugation step and resuspended in 1 mL buffer TFB I and then 0.1 mL DMSO was added slowly drop by drop. Finally, the cell suspension was incubated on ice for 10–15 minutes before flash freezing in liquid N_2 in 50 μL aliquots, which were stored at –80 °C.

Table 2.3 Buffer for preparation of competent cells. Buffer was prepared using distilled water and was filter-sterilised before use. Chemicals were purchased from Merck and Thermo Fisher Scientific.

Buffer Name	Composition
TFB I	10 mM PIPES (pH adjusted before adding other chemicals), pH 6.7; 250 mM KCl, 55 mM $MnCl_2$, 15 mM $CaCl_2$.

Plasmid construction

To generate the plasmids used throughout this thesis, the following procedure was carried out. The general process of plasmid construction is to digest DNA sequences with restriction enzymes and then ligate the DNA fragments. Two restriction sites should be found on both the vector and target protein genes. The reaction mixture recipe is shown in Table 2.4. Both the vector and target protein genes were digested by restriction enzymes respectively, followed by adding 1 μL of enzyme rSAP to the vector tube and incubating at 37 °C for 30 minutes. The final step was to inactivate the enzymes by incubating the tube at 65 °C for 5 minutes. Then an agarose gel electrophoresis was run (detail shown below) to separate the mixture of multiple DNA chains into bands of different size. The correct band based on the DNA size was cut out, and both the vector and insert DNA were purified with the gel extraction kit (detail shows below). Finally, a ligation reaction was set up as shown in Table 2.5 and incubated at 16 °C overnight (about 16 hours). The following day constructed plasmids are ready for transformation.

Table 2.4 Reaction mixture for plasmid digestion using restriction enzymes.

	Vector:	Insert DNA
DNA	3–5 μg:	1–2 μg
Cutsmart buffer (10X)	10 μL	5 μL
Restriction enzyme 1 (20 U μL^{-1})	1 μL	0.5 μL
Restriction enzyme 2 (20 U μL^{-1})	1 μL	0.5 μL
Ultrapure water	varied	varied
Total volume	100 μL	50 μL

Table 2.5 Reaction mixture for ligation.

	Ligation
Vector	50 ng
Insert	15 ng

10x T4 ligase buffer	1 μL
T4 Quick ligase	0.5 μL
Ultrapure water	–
Total volume	10 μL

Polymerase chain reaction for site-directed mutagenesis

The primers used in this research are described in appendix A. A set of complementary primers for each variant were used to create linear fragments from the plasmid. Each primer is around 40 base pairs long and the mutation is encoded in the middle of the primers. The reaction mixture and method are outlined in Table 2.6 and Table 2.7 respectively. Phusion High Fidelity DNA Polymerase (New England Biolabs, NEB) was used to minimise random mutations generated during PCR. DpnI (NEB) (1 μL) was then added to the PCR mixture to digest methylated template DNA and incubated at 37 °C for at least 2 hours and then heat inactivated by incubating at 80 °C for 20 minutes to prevent 'star activity' (non-specific cutting). When the GC% of a primer is over 50%, the reaction buffer was replaced by GC enhancer (NEB) to increase the primer binding affinity with the template DNA.

Table 2.6 Reaction mixture for a 50 μL PCR using Phusion polymerase. The reaction buffer was supplied with the Phusion DNA polymerase (NEB). dNTPs were also supplied by NEB.

Reaction Component	Volume / μL	Final Concentration
Reaction Buffer (5x)	10	1x
Deoxynucleotides (dNTPs) (10 mM each)	1	200 μM
Template DNA (10 ng μL^{-1})	5	1 ng μL^{-1}
Forward Primer (10 μM)	1.1	0.22 μM
Reverse Primer (10 μM)	1.1	0.22 μM

Phusion Polymerase (2 U μL^{-1})	0.5	0.02 (U μL^{-1})
Ultrapure water	29.8	–
Total Volume	50	

Table 2.7 Thermocycling conditions for site-directed mutagenesis with Phusion polymerase.

Step	**Temperature /** ℃	**Duration**
Initial Denaturation	98	10 seconds
Amplification:		20 cycles
Denaturation	98	30 seconds
Primer Annealing	60	20 seconds
Extension	72	30 seconds per kb
Final Extension	72	10 minutes
Cooling	4	Until further use

Transformation of plasmid DNA and PCR product

For a transformation procedure using either commercially purchased competent cells, or competent cells prepared locally, 1 μL of plasmid DNA or 5 μL ligation product was added to 50 μL of thawed competent cells. The cell-DNA mixture was incubated on ice for 10–20 minutes before being 'heat shocked' in a water bath (or heat block) at 42 °C for 45s and subsequently incubated on ice for 2 minutes. LB medium (1 mL) was added and the transformed cells allowed to recover at 37 °C with shaking. For transforming a plasmid, after 1 hour, 100 μL culture was spread onto an LB agar plate with antibiotics to select for transformants. For transforming a ligation reaction product: after 1 hour, the culture was centrifuged at 3000 *xg* for 5 minutes and most of the supernatant removed, leaving approximately 100 μL. The cell pellet was resuspended in the remaining LB medium spread onto an LB agar plate with antibiotics. The plate was grown at 37 °C for 16 hours and stored at 4 °C when not in use.

'Mini-prep' purification of plasmid DNA

Following transformation, a single colony was used to inoculate 10 mL of LB medium with antibiotics and grown for a minimum of 16 hours at 37 °C with shaking. The cells were harvested by centrifugation for 10 minutes at 10000 *xg*. Plasmid DNA was then extracted and purified using the Quiagen QIAquick Mini-prep kit and purified into 30 μL autoclaved ultrapure water.

Measuring DNA concentration

DNA samples were diluted 1:50 to a final volume of 100 μL, and absorbance was measured at 260 nm (A_{260}) using NanoVue (GE Healthcare) spectrophotometer. The concentration of DNA was calculated based on the assumption that a 50 μg mL^{-1} dsDNA solution has an A_{260} of 1.0.

Sanger sequencing of DNA

Plasmid products were checked for mutations using Sanger Sequencing provided by Source Bioscience. The concentration of purified DNA was determined using NanoVue (GE Healthcare) and was diluted using ultrapure water as required (100 ng μL^{-1}).

2.2 Protein expression and purification

For the majority of experiments in this thesis, proteins were over-expressed and purified before assays. Details on developing expression conditions and purification steps of each protein are discussed in the relevant chapters. In this section, the principles and theory of protein expression and purification are briefly introduced and described. Additionally, the tools used for detection, quantification and routine analysis of proteins are also described below.

2.2.1 Bacterial growth and protein production

In this research, there are several proteins expressed by using *E. coli* cells as a host. Bacterial growth is a proliferation of bacteria into two daughter cells, in a process called binary fission.[139] There are four phases during the bacterial growth curve shown in Figure 2.1, which are lag phase (A), log phase or exponential phase (B), stationary phase (C), and death phase (D). During the lag phase, bacteria adapt to the growth medium and conditions. Cell duplication is slow in this phase. Cells are ready to divide exponentially after being adapted in the lag phase, which occurs in the logarithmic growth phase (log phase). Cell concentration doubles in less than 30 minutes. Under typical growth conditions (LB medium, shaking at 37 °C) the optical density of the bacteria is about 0.4–0.8. At this point, cells are best suited for the over-expression of the proteins of interest.

In the stationary phase, the cell growth rate equals the death rate, and the cell density remains constant. Once some nutrients run out or the medium becomes toxic due to by-product accumulation, cells stop dividing and slowly undergo cell death. This process involves membrane rupture and a decrease in optical absorbance.

Therefore, during heterologous protein expression, it is crucial to monitor the optical density of the bacteria regularly to ensure that protein induction occurs during the log phase and before reaching the death phase.

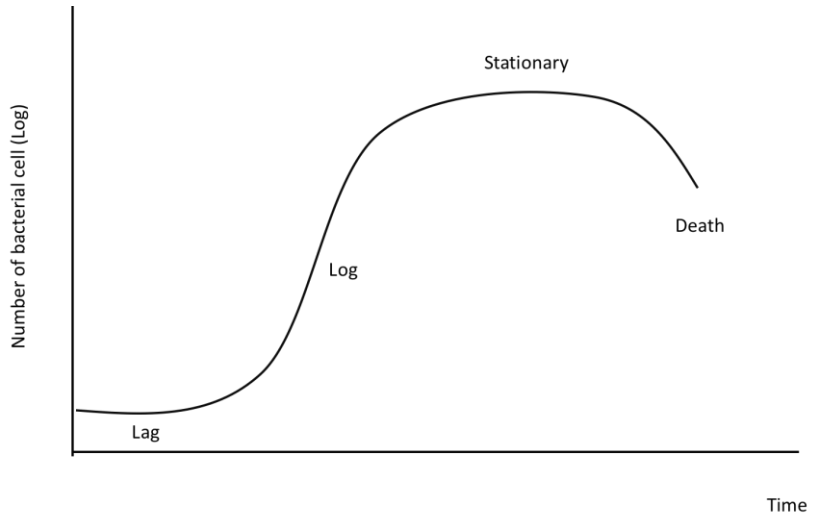


Figure 2.1 Graphic representation of typical bacterial growth curve in culture medium.[139]

2.2.2 Expression and purification of spinach ferredoxin

E. coli BL21(DE3) competent cells were transformed with the pGEX-FdI plasmid (see chapter 3 for more detail) in a Lysogeny Broth (LB) media containing 100 µg mL^{-1} ampicillin (the Ampicillin genes shown in the plasmid in Figure 3.2). LB is a nutritionally rich medium commonly used for growing bacteria. Afterwards, 1 µL of the *E. coli* stock culture bearing the pGEX-FdI plasmid was inoculated into a 15 mL Falcon tube containing 10 mL of sterile LB media as seed culture which was grown in the incubator at 220 rpm at 37 °C until the optical density (OD) of the culture reached 0.8. The OD is measured by following the absorbance at 600 nm wavelength. The OD measures the amount of light scattered by the bacterial cells within an optical component, thus giving a direct measurement of the concentration of the bacterial cells present in the culture, i.e., more scattered light indicates greater turbidity and thus higher cell density. A 600 nm wavelength (OD_{600}) is commonly used in spectrophotometry for this purpose. The seed culture was then diluted into a 2 L flask containing 1 L of new sterilised LB media containing ampicillin (100 µg mL^{-1}). The culture was incubated by shaking at the same speed and temperature as before until OD_{600} reached 0.6. Then, the culture was cooled in a water bath for 10 min before adding 100 µL of 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) (0.1 mM final concentration)

to induce protein production. The culture was left overnight at 16 °C and shaking, to allow for protein production.

Affinity chromatography

The cells were harvested by centrifugation at 8000 *xg* for 10 min. The pellet was resuspended in buffer W (100 mM Tris-HCl, 150 mM NaCl, pH 8.0) with DNase I and lysozyme. The cells were lysed by sonication (Fisherbrand™ Q500 Sonicator fitted with standard 0.5-inch probe; 2-second pulse, 5-second pause, total sonication of 5 minutes). The lysis mixture was centrifuged to separate out protein from the cell debris. Since the desired protein is soluble it will be in the cell soluble lysateThe supernatant was filtered through a 0.2 µm filter, to prevent any insoluble cell components from blocking the column, before being applied to a column for purification by GST affinity chromatography.

Removal of the GST tag

To prepare PreScission protease mixture, 40 µL of PreScission protease and 960 µL of PreScission cleavage buffer were mixed for each mL of Glutathione Sepharose 4B column volume. The tagged GST-FdI bound to the Glutathione Sepharose column was the washed by 10 column volumes of PreScission cleavage buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, pH 7.5). The PreScission protease mixture was loaded onto the column, and the column was sealed and incubated at 4 °C. Following incubation, the column was washed by PreScission cleavage buffer, using approximately 3 bed volumes, to elute the ferredoxin protein.

2.2.3 Expression and purification of [FeFe] hydrogenases from the sulfate reducing bacterium Desulfovibrio desulfuricans (*Dd*HydAB)

_Dd_HydAB Expression

A single colony of bacteria was grown in a preculture using LB-phosphate media (20 g L^{-1} Luria Bertani (LB), 5 g L^{-1} Na_2HPO_4, 10 g L^{-1} K_2HPO_4) and the antibiotics kanamycin (50 µg mL^{-1}) and chloramphenicol (34 µg mL^{-1}). The preculture was left at 37 °C overnight in the incubator with gentle shaking (200–300 rpm). The next day, 10 mL of the preculture were used to inoculate 1 L of LB-phosphate (20 g L^{-1} LB, 5 g L^{-1} Na_2HPO_4, 10 g L^{-1} K_2HPO_4) with kanamycin (50 µg mL^{-1}) and chloramphenicol (34 µg mL^{-1}) supplemented with 2 mM ammonium ferric citrate (the source of Fe), 0.5% w/v glucose and 2 mM cysteine (the source of sulphur). The cells were grown in the incubator at 37 °C and 200–300 rpm until OD_{600} reached about 0.6. Then the culture was deoxygenated by bubbling with argon for 1 h, followed by induction of protein expression with 1 mM IPTG and an addition of an additional 3 mM cysteine. The culture was left overnight on the bench for protein expression with stirring. The following day, the cells were harvested anaerobically by centrifugation (6,500 *xg*, 30 min, 4 °C).

_Dd_HydAB purification

The pellet was resuspended in buffer W (100 mM Tris-HCl 150 mM NaCl pH 8.0) containing deoxyribonuclease I (DNase I) to cleave bacterial DNA and lysozyme to break the bacterial cell wall. All this is again done under strict anaerobic conditions inside an anoxic glovebox. The cells were then lysed by sonication (Fisherbrand™ Q500 Sonicator fitted with standard 0.5-inch probe; 2-second pulse, 5-second pause, total sonication of 15 minutes) inside the glovebox and then an ultrafast centrifugation step is performed (45000 *xg* for 60 min) anaerobically to separate cell debris and supernatant. After that, the supernatant containing the apo-protein (containing the [4Fe–4S] clusters but lacking the [2Fe] subcluster) was loaded onto a strep-tactin column (iba-lifesciences) inside the glovebox followed by washing with buffer W to remove

impurities. The apo-*Dd*HydAB was eluted with the elution buffer containing desthiobiotin (100 mM Tris-HCl pH 8.0, 150mM NaCl, 2.5 mM desthiobiotin).

Artificial maturation

For the artificial maturation process, apo-DdHydAB was diluted to about 200 µM in buffer W in a 50 mL falcon tube and 1 mM $[2Fe]^{adt}$ was added (this corresponds to a 5-fold excess, 20 µL of a 50 mM stock solution prepared in Buffer W). The mixture was incubated at 35°C and 300 rpm shaking in the dark (covered with aluminium foil) for about 50 hours to allow all the apo-protein to be maturated with the $[2Fe]^{adt}$ cofactor.[140] Then, the maturated protein was loaded into a 10 mL desalting PD-10 column equilibrated with 25 mM Tris-HCl, 20 mM NaCl pH 8 buffer to remove the excess of unbound $[2Fe]^{adt}$.

2.2.4 SDS-PAGE for protein separation

A SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) is a discontinuous electrophoresis technique used to separate proteins by molecular mass. Sodium dodecyl sulfate and polyacrylamide gel can eliminate differences due to protein shape and charge. SDS acts as a surfactant, masking the intrinsic charge of proteins and conferring them very similar charge-to-mass ratios. The final separation of proteins is dependent almost completely on the differences in relative molecular mass of polypeptides.

10 µL sample of the protein was combined with 10 µL 2x loading buffer (detail shown in Table 2.13) and denatured at 95 °C for 3 minutes. 5 µL of the denatured mixture was loaded onto a pre-made SDS gel (Invitrogen) and electrophoresed for 45 minutes in 1x Glycine Buffer (Table 2.8 shows the ingredients of 5x SDS-PAGE Glycine Buffer) at 200 V; a pre-stained protein ladder (NEB) was included to allow determination of protein bands based on the molecular weight. The gel was stained using instant blue (Expedeon) for 15–30 minutes to reveal protein bands.

The protein concentration was determined by the Bradford assay using BSA as a standard.

Table 2.8 Buffers for SDS-PAGE. Buffer components supplied by Melford, Thermo Fisher Scientific.

Buffer Name	Composition
2x SDS-PAGE loading dye	4% SDS, 20% glycerol, 200 mM DTT, 0.01% bromo-phenol blue and 0.1 M Tris HCl, pH 6.8
5x SDS-PAGE running buffer (1L)	15 g Tris-HCl, 72 g Glycine, 5 g SDS

2.2.5 Western Blot

After completion of SDS-PAGE, the proteins were transferred to a nitrocellulose membrane (GE Healthcare Amersham™) using the Mini Trans-Blot Cell electrophoresis system (Bio-Rad) at 100 V for 90 minutes. The Trans-Blot Cell was filled with pre-chilled transfer buffer (25 mM Tris, 192 mM Glycine, 20% Methanol and 0.1% SDS) and left on ice during the electrophoretic transfer process. After the transfer step, nitrocellulose membranes were transferred to fresh blocking buffer (PBS containing 5% skimmed milk powder, 0.1% Triton X-100) to block for 1 hour at room temperature or overnight at 4 °C. After blocking, membranes were incubated with the primary antibody (HRP conjugated), diluted to an appropriate concentration in blocking buffer, in a 50 mL Falcon tube and incubated for 1 hour with rotation on a roller mixer. The membrane was subsequently washed 3 times, 10 minutes each time, with fresh PBST buffer (PBS containing 0.1% Triton X-100) to remove unbound antibodies.

The peroxidise signal was detected by a chemiluminescent detection system. 2 mL mixed ECL western blotting detection reagents (from GE Healthcare Amersham™, 1 mL of each detection regent in the kit) was added to one nitrocellulose membrane and left for 1 minute. After this ECL reagent was removed, membranes were fixed on a cassette by taped cling film and exposed on chemiluminescence films (GE Healthcare Amersham™). The film was then developed using an OPTIMAX® 2010 X-Ray Automatic Film Processor (PROTEC X-ray technology).

2.3 Solution assays for H_2 oxidation by hydrogenase

The H_2 activity of hydrogenases was measured using a conventional solution assay. All measurements were conducted at 25 °C in an anaerobic glove box (<3 ppm O_2, Belle Technologies), and enzymes were activated for 16–32 hours under a flow of H_2 prior to use. Methylene blue was made to a stock concentration of 50 μM in pH 6.0 potassium phosphate buffer in 50 mM K_2HPO_4 and KH_2PO_4 and diluted to a working concentration of 25 μM. The reduction of methylene blue (Aldrich) at 600 nm was monitored using an Ocean Optics S2000 fibre optic spectrometer controlled by OOIBase32 software (Ocean Optics). The working solution (5 – 10 mL) was continuously purged with H_2 in a sealed glass vial for a minimum of 10 minutes prior to the experiment. H_2-saturated methylene blue (1 mL) was then transferred to a sealed cuvette, and the change in absorbance monitored to ensure it was stable. At the same time, continuing to monitor the absorbance, 5 μL of the activated enzyme was injected into the sealed cuvette using a syringe. The solution was stirred continuously throughout to ensure a homogenous distribution of enzyme within the cuvette. The turnover frequencies were taken from the initial velocities, where the change in absorbance was linear, and the frequencies were expressed as μmol H_2 oxidised per minute per mg enzyme (μmol H_2 min^{-1} mg^{-1}) and apparent turnover rate *k* (s^{-1}). The extinction coefficient for methylene blue at 600 nm is 22400 cm^{-1} M^{-1}.

2.4 Spinach ferredoxin stability test

Protein stability was determined using nano-DSF (Differential Scanning Fluorimetry) using a Prometheus NT.48 (Nanotemper Technologies Ltd, UK). The protein samples to be analysed were drawn into 10 μL glass capillaries (0.5 mm internal diameter) by capillary action. Initial screening experiments being performed at sample concentrations of 0.1, 0.5 and 1.0 mg/ml to identify conditions that optimise the fluorescence signal from tryptophan residues in the sample while minimising the intensity of the incident laser. Once suitable conditions were identified samples were heated through a temperature gradient of 20–95 °C at a rate of

1 °C/minute. During heating the fluorescence signals at 330 nm and 350 nm were continuously measured throughout the melting process. The ratio of fluorescence at 350 nm and 330 nm (F350/330) was plotted against temperature to identify sample unfolding. All Fd1 samples were measured at a concentration of 1 mg/ml using a 50% laser intensity.

2.5 Electrochemistry

2.5.1 Three electrode system

There are three electrodes in a simple electrochemical experiment, a working electrode (WE) where the redox processes of interest happen, a counter electrode (CE) where the opposite half-cell reaction occurs and a reference electrode (RE). The potential of the WE is monitored relative to that of RE. RE should be an ideal, non-polarisable electrode and has a fixed, known potential, and only a small current passes through the reference electrode. If significant currents are passed through the reference electrode circuit, the reference electrode undergoes polarisation (change the concentration of redox species), making the potential inaccurate.

In the three-electrode arrangement, a voltmeter is added with the reference in parallel with the working electrode, as shown in Figure 2.2. Thus, the current is passed mainly between the working electrode and counter electrode and measured by the ammeter.

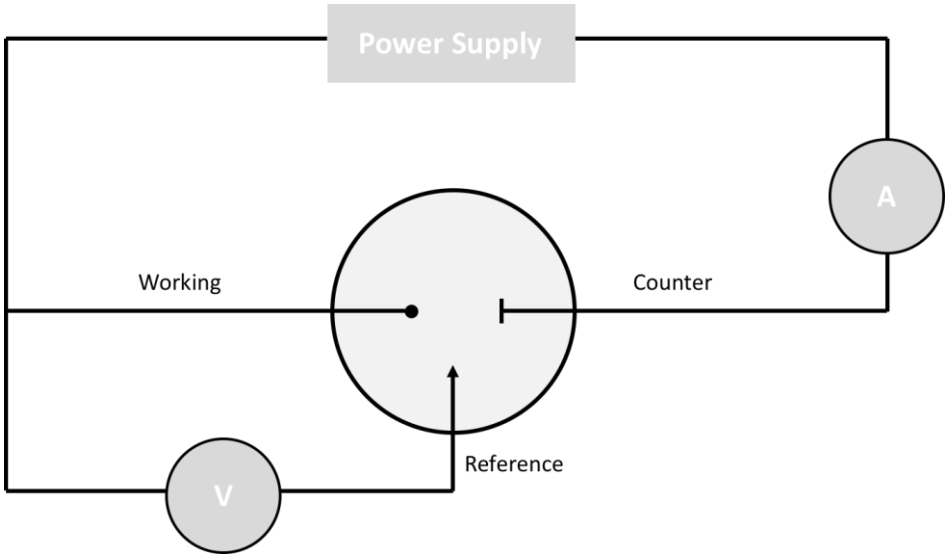
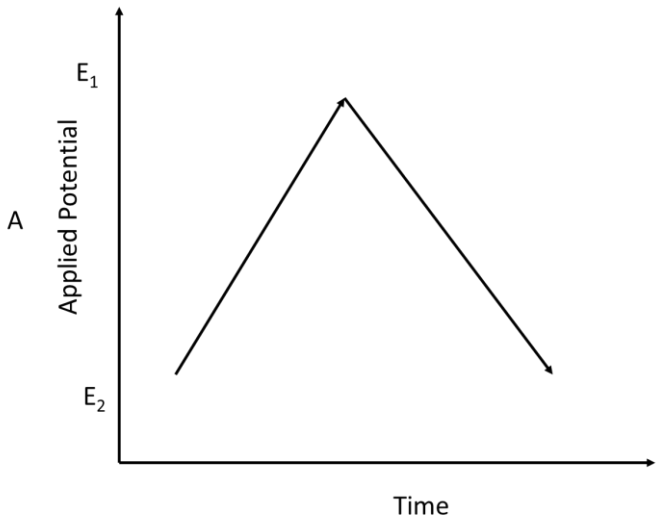


Figure 2.2 A schematic circuit for a three-electrode cell showing the working electrode (WE), counter electrode (CE), reference electrode (RE), voltmeter (V) and ammeter (A).

2.5.2 Cyclic voltammetry and chronoamperometry

Cyclic Voltammetry

In cyclic voltammetry (CV), an applied potential to the working electrode is swept between two potentials, E_1 and E_2 as shown in Figure 2.3A. The scan direction of CV is reversed again to complete a cyclic. The range of the potential E_1 and E_2 should cover the reduction potential of the studied species.



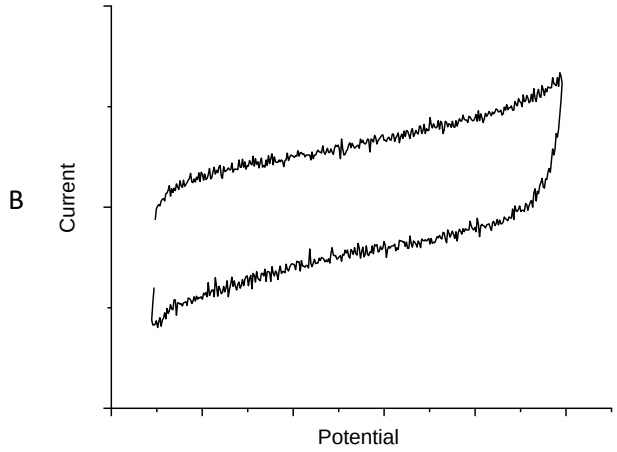


Figure 2.3 (A) Potential applied as functions of time during cyclic voltammetry. (B) Current response at a pyrolytic graphite edge electrode as a result of the variation in applied potential.

A CV experiment is shown in Figure 2.3B performed at a pyrolytic graphite edge without redox species at the electrode. The ions are rearranged within the double layer of the electrode when the working electrode's applied potential is changed, which results in the variation of the current. The parallelogram-shaped current response results from the non-Faradaic phenomenon of charging and discharging the double layer, known as the 'discharge capacitance, 'C_d' mentioned in section 2.5.3. The discharge capacitance is described by the equation:

$$C_d = \frac{i_d}{v} \tag{1}$$

The value of C_d is dependent on the scan rate, v. The higher the scan rate, the higher the discharge current, i_d contribution to the cyclic voltammogram. Before carrying out the experiment, the pyrolytic graphite working electrode was polished until the discharge capacitive current was low (for example less than 0.1 μA). Due to the high capacitive current in the first scan, this scan was always discarded to allow equilibration of the electrode and solution.

Chronoamperometry

Cyclic voltammetry may not be appropriate when a particular state is targeted, or a particular state needs to be avoided. In these cases, chronoamperometry is used instead. In the chronoamperometry experiments, the electrode is held at a series of specific potentials and the catalytic current is continuously recorded during a given period. The current of each step is described by the Cottrell equation:

$$i = \frac{nFAc_j^0\sqrt{D_j}}{\sqrt{\pi t}} \quad (2)$$

Where A is the area of electrode surface (cm^2), c_j^0 is an initial concentration of reducible analyte j (mol/cm^3), D_j is the diffusion coefficient for species j (cm^2/s).

2.5.3 Electrochemical experiment set up for protein film electrochemistry

All the protein film electrochemistry (PFE) experiments were performed in an anaerobic glovebox (Belle Technology) under an N_2 environment (<3 ppm O_2). The electrochemical set-up for protein film electrochemistry is shown in Figure 2.4. An autolab potentiostat (PGSTAT128N) controlled by Nova software (Metrohm Autolab) was used for electrochemical measurements. A pyrolytic graphite edge (PGE) rotating disk electrode (RDE) was used as the working electrode and rotated at a constant speed (ω = 1000–3500 rpm) by an electrode rotator (Metrohm Autolab). Before carrying out the experiment, the PGE working electrode was polished with sandpaper and then 1 µL or 2 µL of the enzyme was added to the electrode and allowed to sit for 1 or 2 minutes. The enzyme is adsorbed on the PGE electrode by non-covalent interactions and allows for electron transfer directly from the electrode. The counter electrode was a platinum wire (Surepure Chemetals, 99.95%, 0.4 mm diameter) that took the current from the

working electrode to complete the circuit. The three-electrode system was completed with a saturated calomel electrode (SCE) as the reference electrode.

All the protein film electrochemistry experiments on hydrogenases were carried out in a mixed buffer system (Table 2.9) prepared using ultrapure water (Millipore, >18.2 MΩ cm^{-1}).[141] Precise gas mixtures (BOC gases) were produced using mass flow controllers (Sierra Instruments).

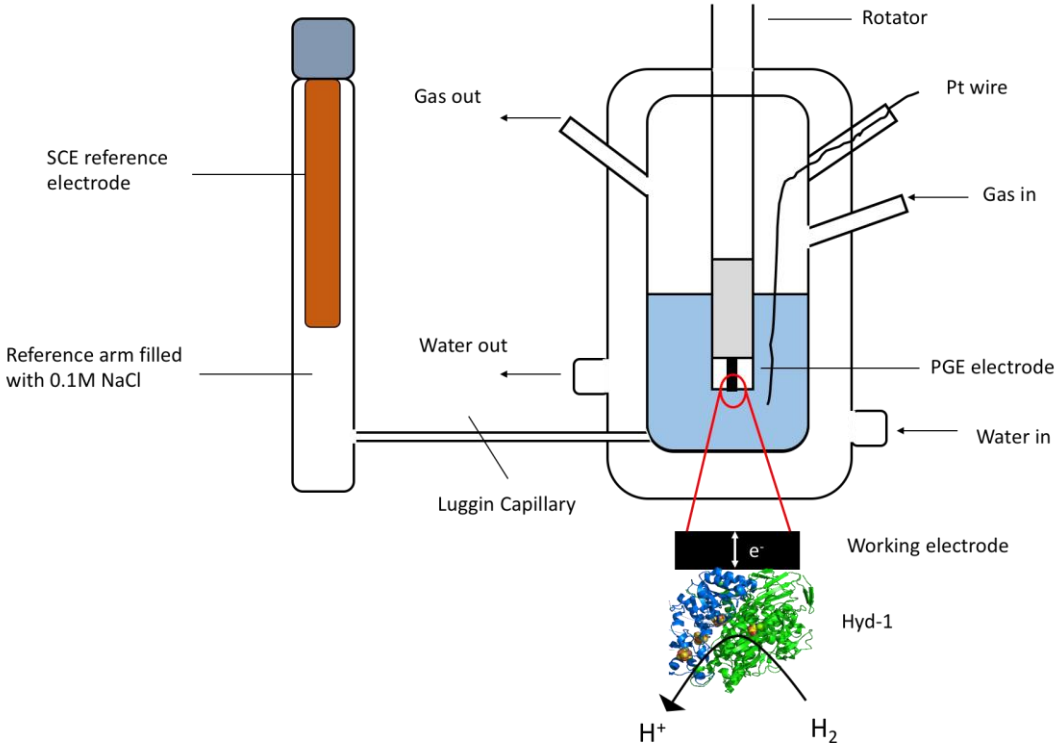


Figure 2.4 Electrochemical set-up for protein film electrochemistry. Three-electrode set up is completed by a PGE working electrode, a platinum counter electrode and the SCE reference electrode. The reference electrode is separated from the main cell by a reference arm filled with 0.1 M NaCl. The glass cell provides a sealed, anaerobic atmosphere. A water jacket is used to control the temperature.

2.5.4 Electrolytic buffer solutions

Table 2.9 Mixed buffer system for electrochemical experiments

Component	Concentration/mM
Sodium acetate	15
2-(Cyclohexylamino)ethanesulfonic acid	15

(CHES)	
N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) (HEPES)	15
2-(*N*-morpholino)ethanesulfonic acid (MES)	15
Sodium chloride	100
N-[Tris(hydroxymethyl)methyl]-3-aminopropanesulfonic acid (TAPS)	15

2.6 Infrared spectroscopy

_p_-Cyano-L-phenylalanine

The unnatural amino acid *p*-cyano-L-phenylalanine (*p*CNF) is used in this research and was incorporated into proteins close to an FeS cluster. The structure of *p*CNF is shown in Figure 2.5. The cyano group is on the phenyl group of side chain. The C≡N stretching vibration (ν_{CN}) is a useful IR probe since this substituent is small and has a strong absorption in a spectral region (~2200 cm^{-1}) free from other vibrational absorptions typically presented in proteins. The CN probes have received significant attention being studied both experimentally and computationally. *p*CNF exhibit a sharp ν_{CN} at 2200–2000 cm^{-1}.[142] and its frequency is sensitive to subtle changes in its local environment and local electric field.[143,144] The CN group is among the first side chain-based extrinsic IR probes used as a probe of ligand binding, local environment and electric field in different proteins to study biological processes.[124,125]

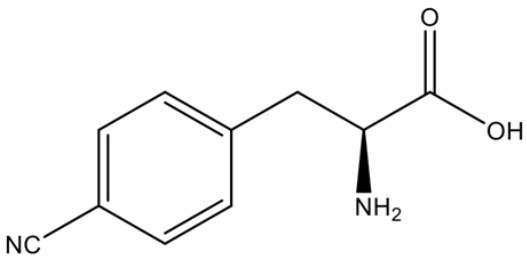


Figure 2.5 The structure of _p_-cyano-L-phenylalanine

2.7 Spectroelectrochemical experiment set up

2.7.1 IR spectroelectrochemical cell

The IR spectroelectrochemical (SEC) cell used in this research was designed in house and made it possible to study small quantities of protein solution in reflection mode. The cell houses a 3-electrode system: a 1 mm diameter polished glassy carbon WE surrounded by a glassy carbon counter electrode ring (CE), and a dismountable Ag/AgCl reference electrode (RE). Figure 2.6 shows a schematic representation of the SEC cell.

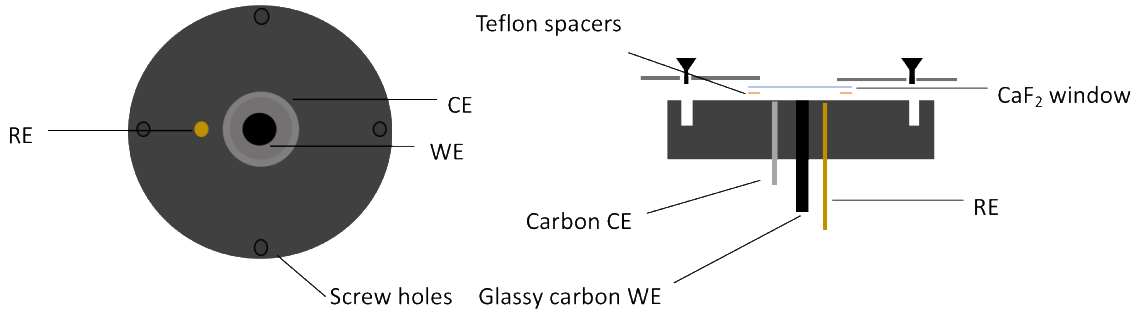


Figure 2.6 Schematic representation of the small volume IR micro-spectroelectrochemical cell. Showing the arrangement of working, counter and reference electrode (WE, CE, RE).

Electrochemical control of protein solution can be achieved in the SEC cell. Theoretically, the rate-limiting step of electron transfer is the diffusion of the electroactive species to and from electrodes, especially for large proteins. Depending on the spacer thickness separating the electrode from the window, the electrochemical response in this cell can be diffusion controlled, or a thin layer-type response where there is minimal exchange with the bulk solution. Having a small diameter working electrode helps to give a diffusion-controlled response. For SEC measurements, a small pathlength (spacer thickness) is important to minimise the strong background absorption from water.

2.8 UV-Vis spectroscopy

For proteins, the conjugated double bonds of the aromatic side chains of residues such as tryptophan, tyrosine, and phenylalanine give rise to a band at 280 nm in the UV-Vis spectrum. The features characteristic of FeS cluster arise from charge transfer from ligand to metal (Fe^{3+}); this process is called ligand to metal charge transfer (LMCT), which are observed. The LMCT absorption process in the FeS cluster happens between 300 and 600 nm. Different types of FeS clusters give variable peak shapes and positions. $[4Fe\text{-}4S]^{2+}$ clusters have the main absorption band around 410 nm with ε_{410} of 9–16 $mM^{-1}cm^{-1}$, while $[2Fe\text{-}2S]^{2+}$ clusters displays multiple bands at 320, 410, 450–460 nm and a small shoulder at 550 nm with ε_{452} of 7–10 $mM^{-1}\,cm^{-1}$.[145] The absorption ratio (A_{410}/A_{280}) can be used to evaluate the extent of FeS cluster incorporation.

2.9 Enzyme kinetics

Enzymes play an important role in chemical catalysis, and they are regarded as 'green chemistry catalysts' as the biological environments are at neutral pH, ambient temperature and aqueous phase. A catalytic reaction takes place at the active site of the enzyme. The compound bound with the active site is named as 'substrate', and a simple enzyme reaction can be written as:

$$E + S \rightleftharpoons ES \rightleftharpoons E + P$$

where E is the enzyme, S is the substrate, ES is the enzyme-substrate complex, and P is the product.

2.9.1 Affinity constant

The initial rate V_0 is a significant parameter in a catalytic reaction, and is proportional to the substrate concentration [S] over a range of substrate concentration.[146]

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

The Michaelis-Menten equation can be written as:

$$V_0 = \frac{V_{max}[S]}{K_M + [S]} \quad (3)$$

The Michaelis-Menten equation shows the relationship between the initial velocity, maximum velocity, and the initial substrate concentration. V_{max} represents the maximum rate achieved by the system at saturating substrate concentration for a given enzyme concentration. The value of the Michaelis constant, K_M, is numerically equal to the substrate concentration at which the reaction rate is half of V_{max}. K_M is an inverse measure of the substrate's affinity for the enzyme, as the smaller K_M indicates higher affinity.[147] K_M can be determined as the concentration of the substrate [S] at the point of ½ V_{max} according to the equation and the plot in Figure 2.7. However, the K_M is dependent on both the identity of enzyme and substrates. At low substrate concentration, K_M >> [S], the Michaelis-Menten equation can be rewritten as:

$$V_0 = \frac{V_{max}}{K_M}[S] \quad (4)$$

When K_M << [S], the Michaelis-Menten equation can be rewritten as:

$$V_0 = V_{max} \quad (5)$$

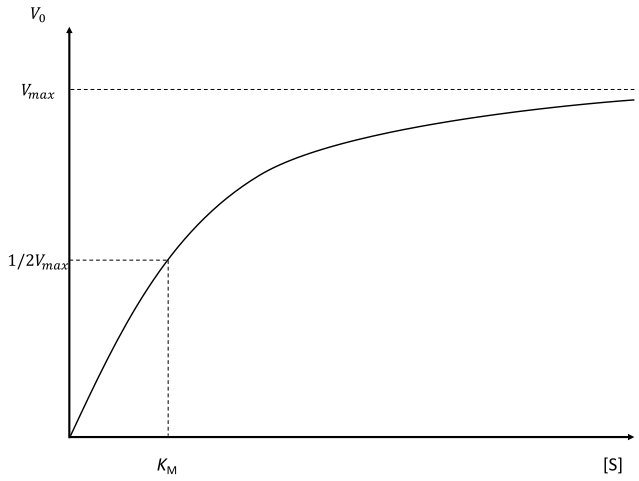


Figure 2.7 shows the Michaelis-Menten curve indicating the relationship between the concentration of substrate, velocity and Michaelis-Menten constant.

The Hanes-Woolf analysis

The Hanes-Woolf representation was used to determine K_M here in this study in order to ensure consistency with previous studies, as this has been used for Hyd-1 and variants in the past. A Hanes-Woolf plot uses a rearrangement of the Michaelis-Menten equation:

$$\frac{[S]}{v} = \frac{[S]}{V_{max}} + \frac{K_M}{V_{max}} \quad (6)$$

Using the current i under different $[H_2]$ as the velocity v, this becomes:

$$\frac{[S]}{i} = \frac{[S]}{i_{max}} + \frac{K_M}{i_{max}} \quad (7)$$

Thus, plotting $[H_2]/_i$ versus $[H_2]$ gives a slope of $1/_{V_{max}}$ and an intercept of $K_M/_{V_{max}}$. For Hyd-1 and variants discussed in this thesis, this was performed at several potentials. A Hanes-Woolf analysis requires the data to be linear and, in some cases, particularly when the K_M was found to be high, the data at low $[H_2]$ was omitted.

2.9.2 Henry's law and gas constants

Henry's law states that at a constant temperature, the amount of a given gas dissolved in a given type and volume of liquid is directly proportional to the partial pressure of that gas in equilibrium with that liquid.[148] Therefore, at constant temperature:

$$p = K_H \cdot c \tag{8}$$

where p is the partial pressure of the solute in the atmosphere above the solution, c is the concentration of the solute, and K_H is the Henry's constant, which depends on the solute, the solvent and the temperature. The Henry's constant K_H of H_2 is 1282.05 (L·atm)·mol^{-1}, the Henry's constant K_H of O_2 is 769.23 (L·atm)·mol^{-1}.

2.9.3 Eyring equation and use

The Eyring equation is derived from transition state theory which describes reaction kinetics:

$$\ln\frac{k}{T} = \ln\frac{k_B}{h} - \frac{\Delta H^{\ddagger}}{RT} + \frac{\Delta S^{\ddagger}}{R} \tag{9}$$

This form can be used to determine the activation enthalpy $\Delta H^{\ddagger}$ and entropy $\Delta S^{\ddagger}$ by using the current i in place of the rate constant k and plotting $\ln i/T$ versus $1/T$. The gradient gives $\Delta H^{\ddagger}/R$ and the intercept is $\Delta S^{\ddagger}/R + \ln k_B/h$.

3

Chapter 3: Incorporating an Unnatural Amino Acid into the [2Fe-2S] Protein Spinach Ferredoxin

3.1 Introduction and aim of the chapter

FeS clusters are important cofactors in nature as they have roles in a variety of fundamental cellular processes such as respiration, amino-acid and cofactor metabolism, DNA metabolism and regulation of environmental responses. In the respiratory chain complex I, FeS clusters act as redox carriers, and in the chloroplast, they transfer electron from photosynthesis.[4] Lots of research has been done on FeS clusters for quite a long time. Generally, the techniques used to study FeS clusters are UV-visible spectroscopy, electrochemistry, electron paramagnetic resonance and X-ray absorption spectroscopy among others.[1]

Infrared spectroscopy is a potent method to study the redox state of the active site of hydrogenases, as there are two different inorganic ligands CN^- and CO which are unusual cofactors in biology, and their vibrational absorption bands appear in a part of the IR spectrum that does not overlap with other protein bands. More importantly, the frequencies of these bands are very sensitive to the oxidation state or electron density of the metal to which they are coordinated, so that we can follow the oxidation state of the active site. However, it is not easy to directly detect the oxidation state of FeS clusters by IR spectroscopy. This chapter will explore a new method to study the oxidation state of the FeS clusters in metalloenzymes. It consists of incorporating an artificial amino acid containing an IR probe (RCN) close enough to the FeS cluster that it can sense its oxidation state, but at a distance which does not alter chemistry of the active site or the FeS cluster or disrupt the biology of the enzyme (meaning its expression and protein folding). This approach should make it possible to observe by IR spectroscopy how an FeS clusters chain works to facilitate fast catalysis at the active site in enzymes such as hydrogenases.

FeS clusters are essential cofactors in all forms of life from respiration to protein assembly.[2,4] The ability to delocalise electron density over both Fe and S atoms can make the iron-sulfur

clusters ideal groups in mediating fast biological electron transport, and these cofactors are the major components in the electron transport pathways in the photosynthetic and respiration chains.[22]

Initially identified in mammalian succinate dehydrogenase, FeS clusters were subsequently found in a variety of proteins such as hydrogenases and nitrogenases, which have different and complex functions. In this study, we focus on the FeS proteins involved in electron transfer and redox catalysis. Spinach ferredoxin has been selected as a simple electron transfer model protein to incorporate the unnatural amino acid containing an IR probe group. The long-term goal of the project is to understand how the catalytic active site works together with the FeS clusters during the catalysis in the redox enzymes such as hydrogenase by using IR spectroscopy. Ferredoxins were the first FeS proteins discovered in nature. The first ferredoxin was purified in 1962 from the anaerobic bacterium *Clostridium pasteurianum*.[25] Their name arises from the fact that they were found to contain Fe and be redox active. This large group of FeS proteins can be classified depending on the nature of their FeS clusters as well as by sequence similarity but they all share certain common properties: they are very small proteins with a low mass, they all function as electron transfer centres in processes and they are highly acidic proteins containing Fe and inorganic or acid-labile sulfurs. Ferredoxins can be found in nature as low-potential [2Fe−2S], [4Fe−4S], [3Fe−4S], [3Fe−4S][4Fe−4S], and [4Fe−4S][4Fe−4S] containing proteins. They are usually low potentials proteins with an average redox potential of −400 mV, vs. Standard Hydrogen Electrode (SHE) and extending over a range of 800 mV depending on the protein structure, cluster types, ligands to the Fe, hydrogen bonding network and water accessibility of the cluster. This wide range of potentials enable ferreodoxin to act as redox partners to a variety of molecules in many important biological reactions.[26] Usually, the FeS cluster is located in a hydrophobic pocket and it is the one responsible for the distinctive dark-brown/red colour that the ferredoxins exhibit.[149] Upon reduction, all ferredoxins experience a partial decrease in

absorbance in the UV-Visible region but the pattern of absorption bands depends on the method and extent of reduction.[27]

Spinach ferredoxin (FdI) was the first FeS protein discovered with known function.[28] It is a small [2Fe-2S] protein with a molecular weight of around 13 kDa and it is the most abundant form of the photosynthetic-type ferredoxin present in the spinach chloroplast.[150] The rhombic [2Fe-2S] cluster in chloroplast ferredoxin (shown in Figure 3.1) is the electron carrier which transfers electrons received from photosystem I (PSI) to several enzymes such as ferredoxin-NAD^+ reductase.[20] By changing the oxidation state of the FeS cluster, the the ferredoxin can accept or discharge electrons.

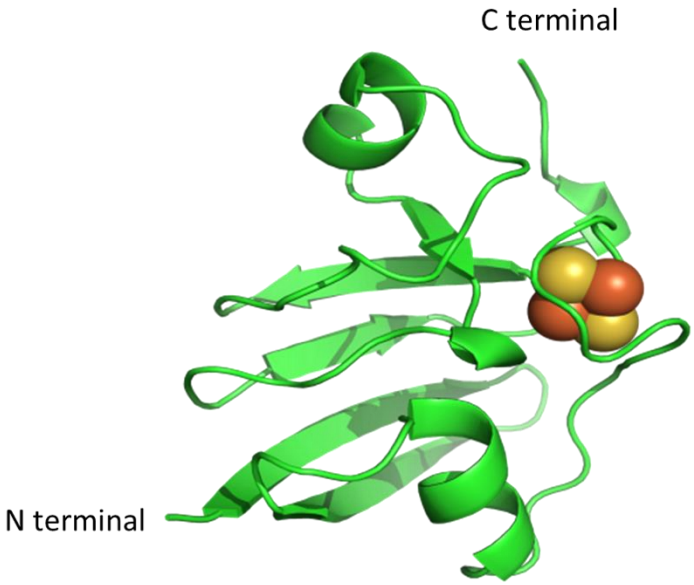


Figure 3.1. **Structure of the spinach ferredoxin I. The [2Fe-2S] cluster is shown in the middle represented by spheres, with the yellow ones being the S atoms and the brown ones the Fe atoms (cartoon representation generated with PyMOL PDB: 1A70).**

Protein dynamics are often significant for biological function. They are not only directly concerned with the transition between states, but also involve the nature and equilibrium populations of the states themselves. Recent efforts have focused on using sensitive IR probes to characterize local protein dynamics.[136] The unnatural amino acid *p*-cyano-L-phenylalanine

(*p*CNF) is one of the most commonly used IR probes (see chapter 1 for more detail). By following the C≡N stretching vibration (ν_{CN}) of the nitrile group the changes in the local protein environment can be indirectly observed, as ν_{CN} is extremely sensitive to changes in the electron density, oxidation state and coordination environment of metal centres in proteins.[151] The CN probes have received great attention being studied both experimentally and computationally to a great extent.[122,142]

In this chapter, spinach ferredoxin will be employed as a simple model to study the incorporation of the the unnatural amino acid *p*CNF and to test whether it can be an IR probe to detect the redox state of the [2Fe-2S] cluster.

3.2 Spinach ferredoxin expression and characterisation

3.2.1 Spinach ferredoxin plasmid construction

The spinach ferredoxin gene was purchased from Fisher, codon-optimised for *E. coli*. A XhoI restriction site was added to the 3'-terminus of the gene by the polymerase chain reaction (PCR) (details are described in section 2.1.3). The amplified DNA sequence was verified by DNA sequencing.

The first attempt on over-expression of spinach ferredoxin (FdI) followed the procedures reported by Piubelli and co-workers. The FdI gene was inserted into the pET-11d vector (named pETFdI plasmid) by restriction digest and ligation. *E. coli* BL21(DE3) was transformed with the pETFdI plasmid and grown on LB plates containing ampicillin. 2×YT media was used for liquid cultures which were incubated at 37 °C until OD_{600} reached 3.5. FdI gene expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation after 2 further hours of growth in the presence of inducer.[20] Due to the highly hydrophilic and acidic character of FdI, the main purification step was ammonium sulfate fractionation coupled to adsorption on Sepharose 6B. However, no FdI was expressed by this protocol; the other

drawback of this old protocol was the lack of an affinity tag for protein purification. Therefore, developing an over-expression and purification system of FdI was the first target in this project.

The pGEX 6P-1 vector was chosen for several reasons, one of the main reasons was that this plasmid consisted of the GST tag gene fused to the gene of the target gene to be expressed. The GST tag not only enabled increased expression levels of the target recombinant protein, but it can be used as an affinity tag for purification. The map of the glutathione S-transferase (GST) vector can be seen in Figure 3.2, showing the reading frames and main features. GST follows the *tac* promoter and the *lac* operator in the map. It folds rapidly into a stable and highly soluble protein upon translation. Inclusion of the GST tag often promotes higher expression and increased solubility of recombinant proteins compared with expression without the tag. In addition, GST-tagged fusion proteins can be purified or detected based on the ability of GST to bind its substrate, glutathione (GSH). Therefore, the spinach ferredoxin DNA fragments together with the pGEX vector containing the ampicillin resistant gene, were digested by BamHI and XhoI endonucleases. Both products were separated by agarose gel electrophoresis and isolated by gel extraction kits. Finally, the DNA segment and vector were ligated by T4 DNA ligase, and the product was verified by both electrophoresis and DNA sequencing. This new plasmid, denoted as pGEX-FdI, was transformed into *E. coli* DH5α/Top 10 competent cells for DNA amplification and storage, and the BL21 (DE3) strain for protein production.

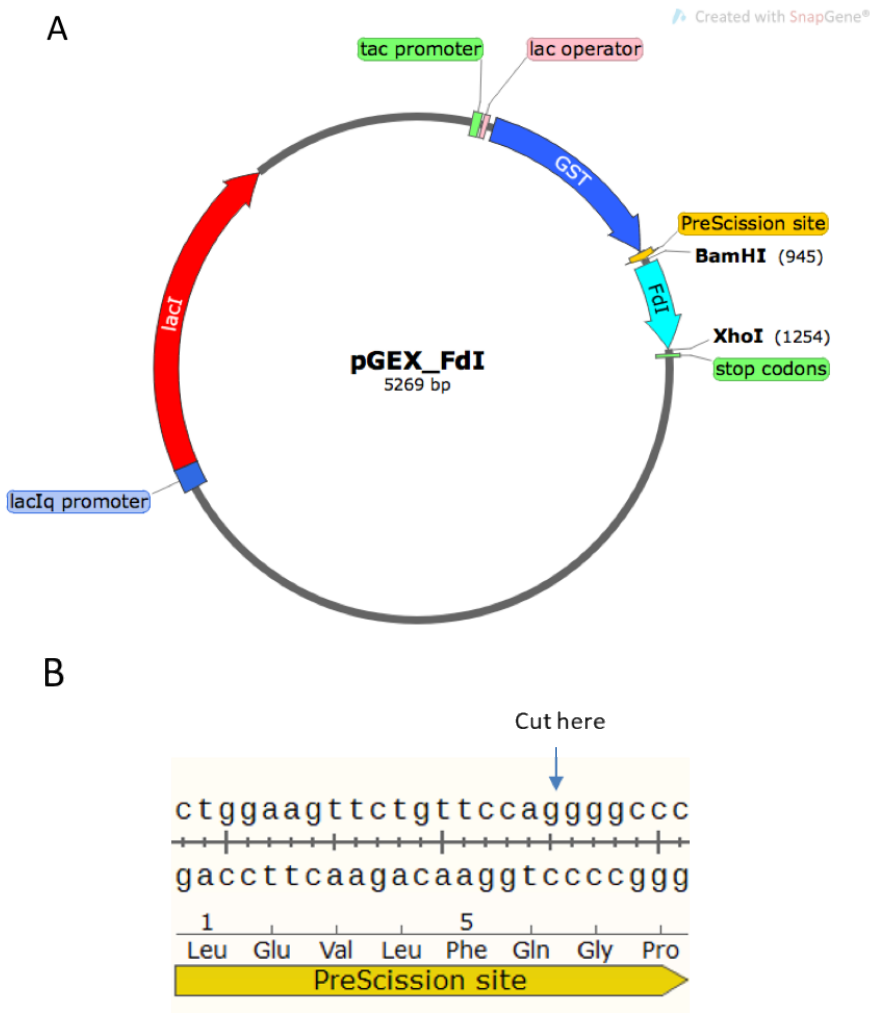


Figure 3.2 Panel A: Graphical map of the plasmid pGEX-FdI. The turquoise blue arrow shows the inserted spinach ferredoxin gene, which is under the control of the *tac* promoter. Map generated by SnapGene. Panel B: The PreScission protease site enables cleavage between Gln and Gly in the linker.

3.2.2 Purification results

Affinity chromatography:

The AKTA protein purification system FPLC was used to purify the ferredoxin (see section 2.2.2). The lysate was loaded onto a pre-equilibrated GST-Trap column with Glutathione Sepharose 4B (GE Healthcare) and washed with Buffer W to remove non-bound cell components.

Removal of the GST tag

The next step was to remove the GST tag linked with the ferredoxin in the GST-Trap Column rather than eluting the GST-FdI. The molecular weight of GST-linked FdI is about 39.6 kDa, while the free FdI is about 12 kDa. In Figure 3.2A, the graphic structure of plasmid pGEX-FdI shows that there is a Protease recognition sequence Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro between the GST tag sequence and the targeted ferredoxin. PreScission protease (GE Healthcare) cleaves between Gln and Gly residues in the sequence at low temperature (shown in Figure 3.2B).

The GST-tagged purification process is shown in Figure 3.3A. Each fragment was analysed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) The picture of the SDS-PAGE is shown in Figure 3.3B.

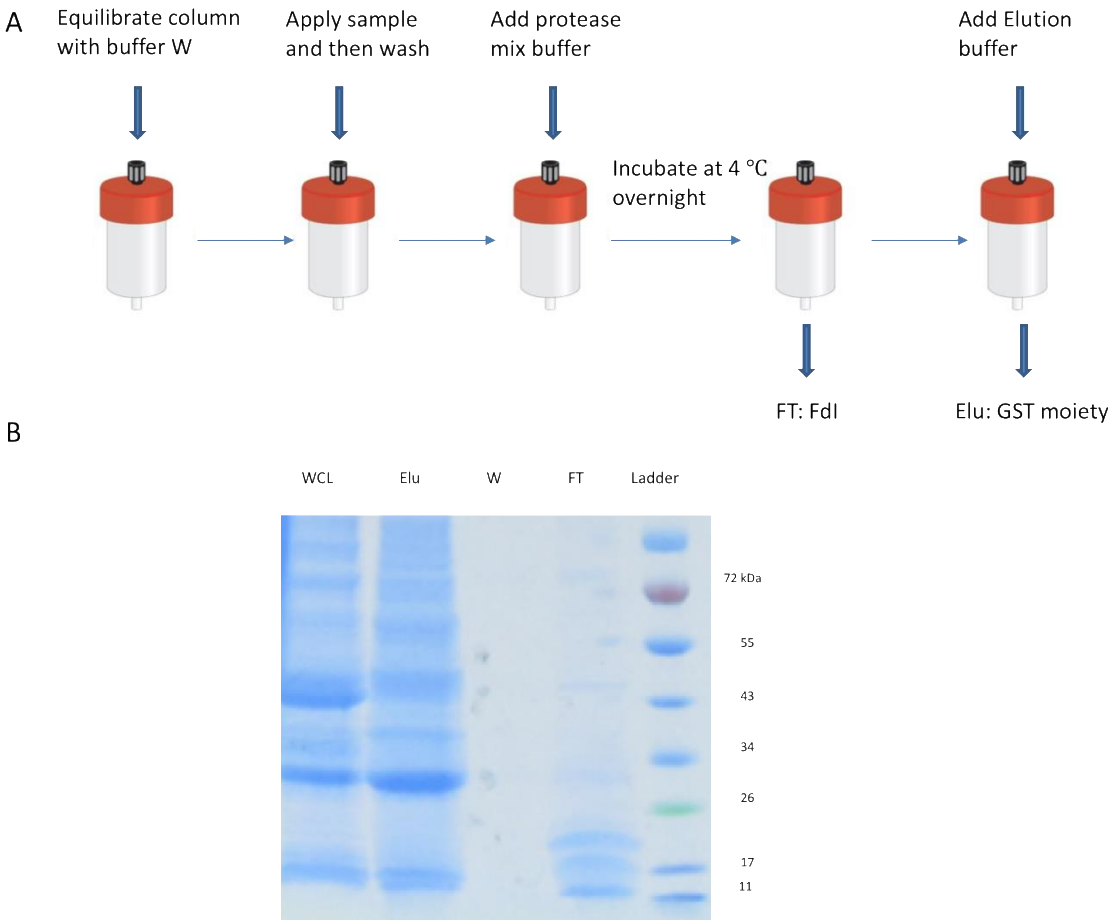


Figure 3.3. Panel A: The process of protein purification of GST-FdI using a GST-Trap column. Panel B: SDS-PAGE of the recombinant FdI after purification. The ladder is the marker, for which molecular weights of the bands are indicated on the right. WCL: whole cell lysate, Elu: the collected eluent after introducing elution buffer (containing free GST moiety), W: eluent from washing of bound-protein with buffer, FT: flow through from the GST-Trap column when the protease mix buffer was applied to the column (containing FdI).

If the protein was pure, it was expected that there will be only one band in the eluent (Elu), at the molecular weight of the protein of interest. The size of GST tagged ferredoxin is about 39.6 kDa. A band around this molecular weight can be found in the whole cell lysate (WCL). During purification, PreScission protease was added to cleave the GST tag from FdI, and the free GST-tag (molecular weight of around 26 kDa) was observed in the flow through (FT). However, the

'Elu' sample showed more than one band between the ladder bands 11 kDa and 26 kDa; the ferredoxin is the band at about 12 kDa. Therefore, further purification was required.

Ion-exchange chromatography

Anion exchange chromatography was carried out to further purify the ferredoxin. As its isoelectric point (pI) is about 3.6, the surface of FdI is predominantly negatively charged at neutral pH. It binds to an anion exchange resin and can be, but it can be eluted with high concentration salt. Then, the protein eluted with a salt gradient, by linearly increasing the concentration of NaCl.

Figure 3.4A shows the results after the second purification step of the recombinant FdI by anion exchange chromatography. From the elution traces, it is clear that at low concentration of salt, some impurities were eluted at low NaCl concentration whereas FdI was eluted at about 0.55 M NaCl. As shown by the SDS-PAGE analysis in Figure 3.4, an apparent single band can be observed in the P5 lanes, meaning that the impurities have been mostly separated from the FdI, yielding a much purer protein.

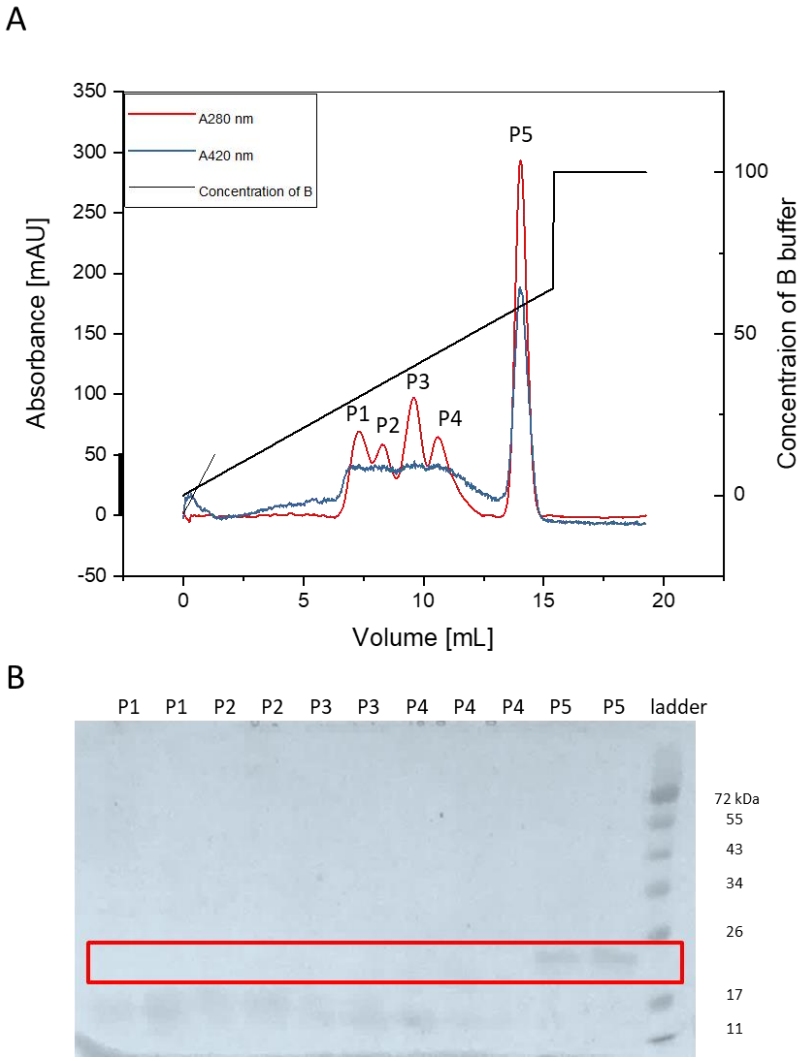


Figure 3.4 Further purification of FdI via anion exchange chromatography. Panel A: The anion exchange eluting traces. The absorbance (A_{280} and A_{420}) is shown over the elution volume. Conditions: 0.5 mL/min flow rate, Q Sepharose column. Buffer A: Tris-HCl pH 7.4 with no salt, Buffer B: Tris-HCl pH 7.4, 1M NaCl. Panel B: SDS-PAGE of the Ferredoxin I after anion exchange chromatography purification. Peak numbers from SDS-PAGE lanes correspond to the Peak numbers from the AKTA traces in Panel A.

Size exclusion chromatography

FdI was further purified by gel filtration chromatography which separates proteins by their molecular weight. It yielded a highly pure preparation of FdI, suitable for most of the subsequent structural and functional studies. All buffers were degassed. A Superdex 200 Increase 10/300 GL column (GE Healthcare) was utilised and the concentrated protein (the final volume had to be less than 1 mL) was injected onto the equilibrated column and washed with

Buffer W. The fractions were collected according to the peaks shown in the UV-Vis traces at 280 nm and FeS cluster absorbance at 420 nm.

Figure 3.5 shows the result of the last step of purification after the size exclusion chromatography. In the absorption trace (A) a large and well-defined peak at 16 mL and a much smaller peak at about 21 mL can be observed in the absorbance at 280 nm (A_{280}). However, the absorbance arising from the iron-sulphur cluster is at 420 nm (A_{420}), and it was only observed in the large peak at 16 mL. The purity of the recombinant FdI was also checked by running the proteins on a 12% SDS-PAGE (Figure 3.5B). An apparent single band at about 22 kDa in lanes 1, 2 and 3 indicated a higher size than expected (12.77 kDa). However, it has been previously observed that in SDS-PAGE, the ferredoxin gives a diffuse band with a higher apparent molecular mass (M_r) than the real value.[20]

The traces from AKTA of the 3-step purified over-expressed FdI are identical to the data from literature on natively expressed FdI from spinach.[20] An A_{420}/A_{280} ratio of 0.50 can be calculated, in close agreement with the reported value of 0.48. This confirmed the high purity and the 1:1 FeS cluster to protein stoichiometry of the purified protein.[152]

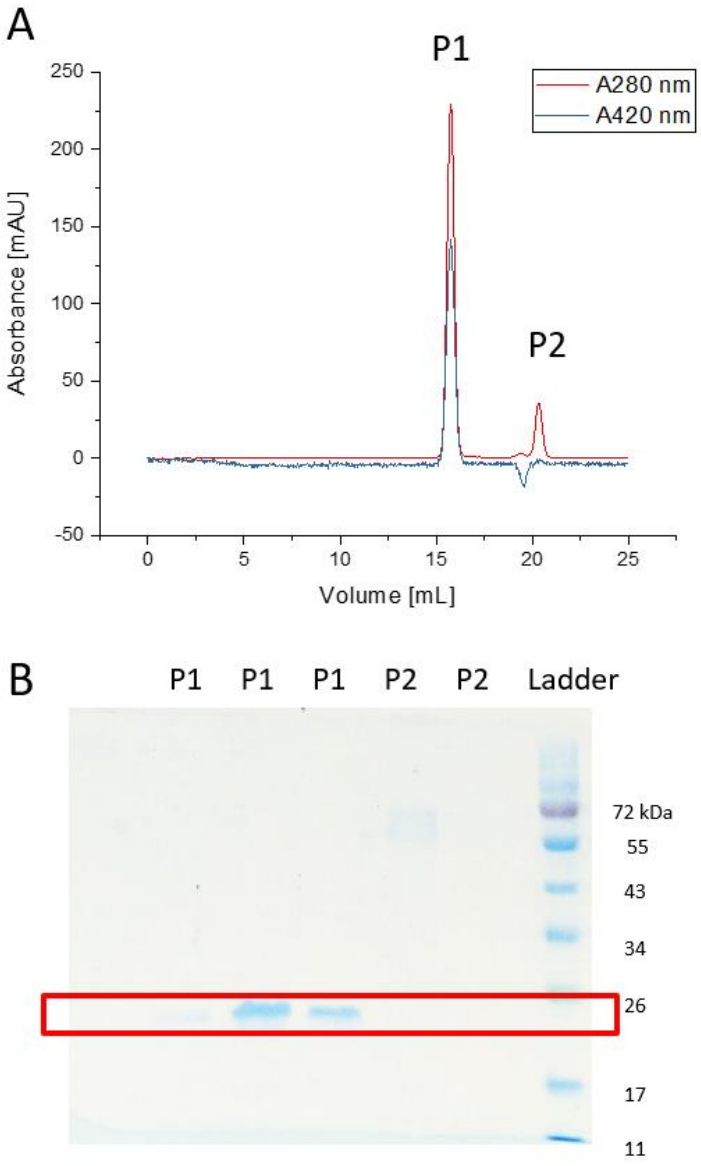


Figure 3.5 The elution trace (A) and SDS-PAGE (B) after size exclusion chromatography. The absorbances at 420 and 280 nm are shown over the elution volume with a flow rate of 0.3 mL/min. P1 lanes from SDS-PAGE are from the peak 1 (P1) while P2 lanes corresponded to the second peak (P2).

3.2.3 Basic and functional characterisation of the spinach ferredoxin

The nature of the [2Fe-2S] cluster was investigated by UV-Visible spectroscopy. According to Armengaud *et al.*, FdI exhibits two peaks at around 420 nm and 460 nm and a shoulder at around 320 nm, all arising from the oxidised $[2Fe\text{-}2S]^{2+}$ cluster.[153] Figure 3.6A shows the UV-Vis spectrum of ferredoxin linked to the glutathione-S-transferase (GST) tag after purification by

affinity chromatography. The broad characteristic absorption peaks corresponding to the oxidised [2Fe-2S] cluster, at around 320 nm, 420 nm and 460 nm, can be clearly seen. The characteristic absorption peaks can be observed only in the oxidised sample as it arises from a ligand-to-metal (Cys-S/Fe^{3+}) charge-transfer. To confirm that the absorption peaks correspond to the $[2Fe\text{-}2S]^{2+}$ cluster, the $[2Fe\text{-}2S]^{2+}$ cluster was reduced by adding a stoichiometric amount of sodium dithionite which resulted in a complete loss of the UV-Vis absorption peaks. Figure 3.6B shows the untagged spinach ferredoxin purified further by anion exchange and size exclusion chromatography. In the oxidised spectrum the characteristic peaks belonging to the $[2Fe\text{-}2S]^{2+}$ cluster are clearly detectable and are consistent with peaks of oxidised GST-tagged FdI and the natively expressed FdI, as reported in the literature.[150] After the $[2Fe\text{-}2S]^{2+}$ cluster was reduced with sodium dithionite, these bands again vanished.

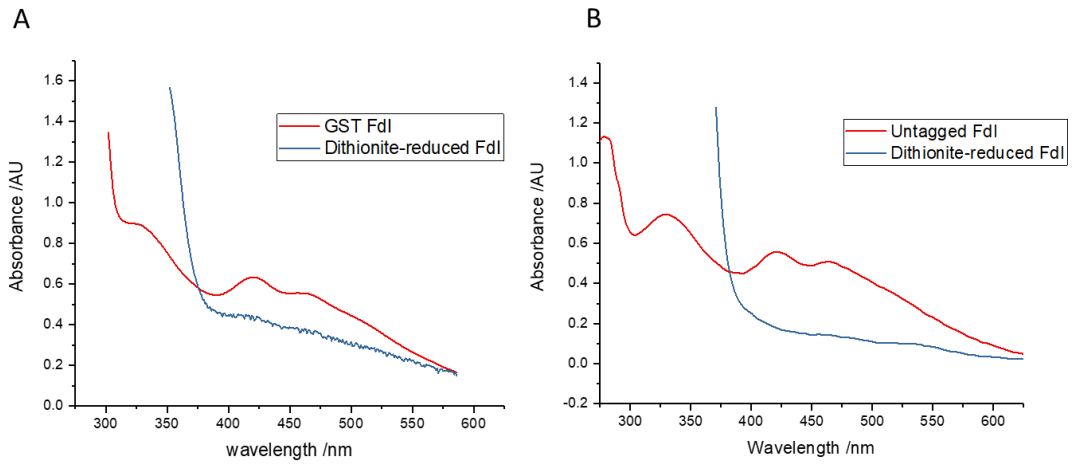


Figure 3.6 UV-Visible (UV-Vis) absorption spectra of the oxidised and reduced spinach ferredoxin I (FdI) recombinantly expressed and purified by GST-Affinity chromatography; A: FdI with GST tag; B: FdI without GST tag. The red traces correspond to the oxidised ferredoxin in 100 mM Tris-HCl pH 8 buffer; the blue traces represent the dithionite-reduced protein also in 100 mM Tris-HCl pH 8 buffer.

Next, the redox activity of the over-expressed spinach ferredoxin was tested by cyclic voltammetry. In cyclic voltammetry, the applied potential is swept linearly versus time in a cyclic fashion while recording the generated current at the working electrode. In a cyclic

voltammogram, the current is plotted vs. the applied potential. The redox potential of spinach ferredoxin reported by Verhagen *et al.* was –0.40 V *vs*. the standard hydrogen electrode (SHE).[154]

In this experiment, Ag/AgCl (3 M KCl) was utilised as a reference electrode. Its potential was calibrated and found to be –0.20 V *vs.* SHE at 25 °C. FdI proteins were measured in 150 mM Tris pH 8 buffer containing 100 mM NaCl. Glassy carbon was chosen as the working electrode. Well-defined oxidation and reduction peaks from spinach ferredoxin were observed (Figure 3.7). The peak-to-peak separation of about 100 mV indicated a fairly reversible process.[155] A cathodic peak assigned to the $[2Fe\text{-}2S]^{2+}/[2Fe\text{-}2S]^{+}$ reduction is observed at –0.45 V (*vs.* SHE) and an anodic peak at around –0.35 V is observed on the reverse scan. The redox potential of the [2Fe-2S] cluster is estimated from the average of the anodic and cathodic peak potentials, and it is found to be –0.405 V (*vs.* SHE).

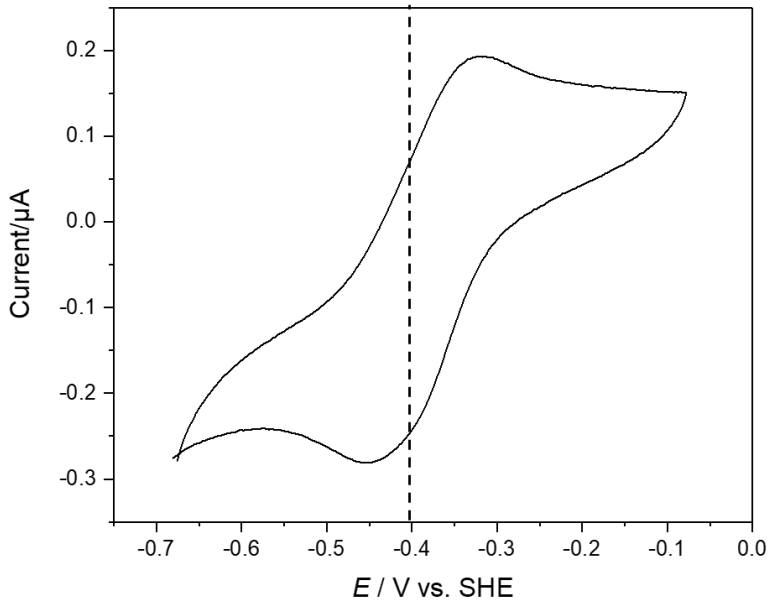


Figure 3.7 Cyclic voltammograms of over-expressed FdI measured in solution with a glassy carbon working electrode in 100 mM Tris-HCl pH 8, 150 mM NaCl at 25 ^{O}C.

In conclusion, the pGEX vector with GST fusion system allowed over-expression of spinach ferredoxin I in *E. coli*. The protein was successfully purified and characterised through UV-Vis

spectroscopy and electrochemistry. The results were in good agreement with data previously reported in the literature.[156] Because the excised FdI and the GST-linked FdI showed the same spectroscopic and electrochemical properties, to reduce the loss of the protein during the multiple purification steps, the GST-linked form of the protein was used in subsequent studies.

Next, in order to select feasible amino acid residues for incorporation of the *p*CNF, site-directed mutations and characterisation will be done.

3.3 Test the tolerance of spinach ferredoxin to mutations close to the FeS cluster.

Having improved the expression level of spinach ferredoxin so that enough pure protein can be produced, some suitable amino acid positions were chosen to incorporate the unnatural amino acid. As the structures of phenylalanine (Phe, F) and tyrosine (Tyr, Y) are similar to *p*CNF (Figure 3.8), Phe and Tyr residues were targeted to avoid disruption of the protein structure and the cluster environment. Residues were chosen at different distances from the FeS cluster. First, mutation to Phe and Tyr was performed at each chosen site to see if the FeS cluster can still be incorporated and if the protein can be properly folded.

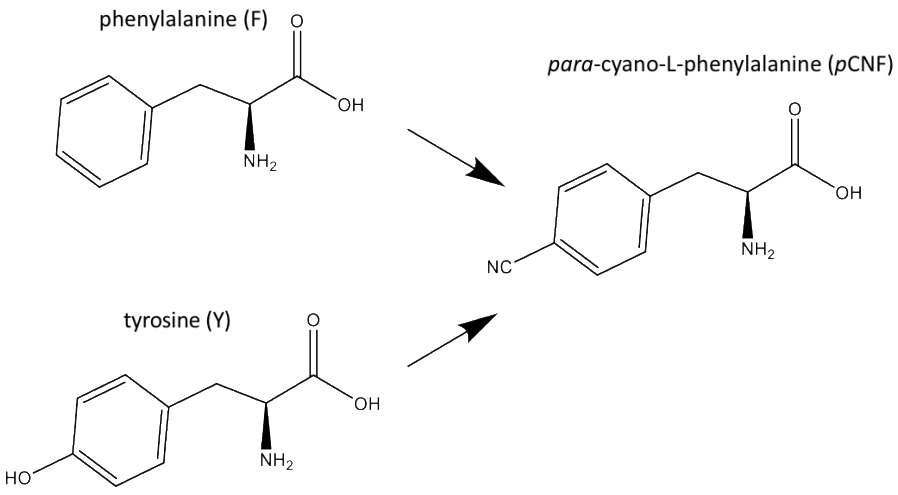


Figure 3.8 Comparison of the structure of the unnatural amino acid *p*-cyano-L-phenylalanine (*p*CNF) with the amino acids phenylalanine (F) and tyrosine (Y).

Four residues were chosen as potential candidates for the incorporation of *p*CNF: Y23, Y37, F63 and Y80 (see Figure 3.9). The targeted sites are not on the protein surface and less than 15 Å from the FeS cluster. The most promising candidates, Y37 and F63, are 7.4 Å and 6.7 Å respectively, from the [2Fe-2S] cluster. Y3, which is far away from the cluster, was selected as a control.

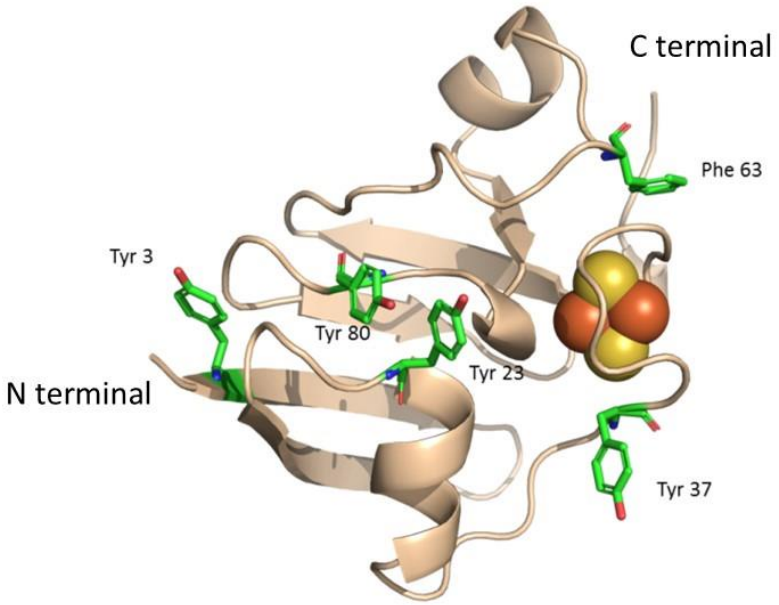


Figure 3.9 Residues chosen for mutagenesis shown in green, the [2Fe-2S] cluster is shown as spheres and the protein structure is shown as a cartoon in the background, the targeted amino acids are shown in stick form. (Figure prepared using the PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. The PDB code is 1A70).[157]

As mentioned above, firstly, variants were produced by site-directed mutagenesis with conservative mutations to natural amino acids. Thus, phenylalanine (F) was mutated to tyrosine (Y), tryptophan (W) or alanine (A), whereas tyrosine (Y) was mutated to phenylalanine, tryptophan and alanine. The nomenclature of the variant is based on the abbreviation and sequence number of the amino acid, for example, phenylalanine 63 (F63) mutated to Tyrosine (Y) is named F63Y.

Site directed mutagenesis was carried out via the PCR-based QuickChange method and each variant generated was confirmed by DNA sequencing. When the sequencing results matched the required plasmid sequence, the mutated plasmids were then transformed into BL21 (DE3) *E. coli* cells. Each variant was over-expressed following the protocol developed for wild type spinach ferredoxin I explained in the previous section, and as stated, only GST affinity chromatography was carried out to purify the protein prior to characterisation of the [2Fe-2S] cluster within the protein via UV-Vis spectroscopy and electrochemistry. The results for each variant are shown in Table 3.1.

Table 3.1 Summary of the over-expression, purification and characterisation of the different variants generated. 'WT-like' indicates that the variant was indistinguishable from the 'native-sequence FdI' wild type (WT) protein in terms of its UV-visible spectroscopic and electrochemical features.

Residues	Variants	Results (Determined by UV-Vis and electrochemistry)
Tyr, Y3	Y3F	WT like
	Y3W	WT like
	Y3A	WT like
Tyr, Y23	Y23F	Incorrect cluster
	Y23W	Incorrect cluster
	Y23A	Incorrect cluster
Tyr, Y37	Y37F	WT like
	Y37W	WT like
	Y37A	WT like
Phe, F63	F63Y	WT like
	F63W	WT like
	F63A	WT like
Tyr, Y80	Y80F	No Expression
	Y80W	No Expression
	Y80A	No Expression

According to the results shown in Table 3.1, Y37, F63 and Y3 were suitable sites to introduce unnatural amino acid mutations. The clear characteristic peaks in the UV-visible spectra show that the [2Fe-2S] cluster has been correctly formed. Figure 3.10A shows the UV-Vis spectra of one variant as an example (Y37F). For this variant, the characteristic peaks of the $[2Fe\text{-}2S]^{2+}$ cluster are observed at about 320 nm, 420 nm and 460 nm respectively (red trace). The fully reduced (by sodium dithionite) spectrum (blue trace) shows how the peaks are lost upon reduction, consistent with the UV-vis spectrum of wild-type spinach ferredoxin. Figure 3.10B shows the cyclic voltammogram of Y37F. The midpoint redox potential estimated from the

voltammogram as the average of the anodic and cathodic peak potentials was–0.405 V (*vs.* SHE). However, some variants showed abnormal behaviour and their UV-Vis spectra were different from that of the wild type. For example, for Y23F shown in Figure 3.11, there was only a shoulder peak at about 410 nm, which might indicate that the protein had not folded properly.

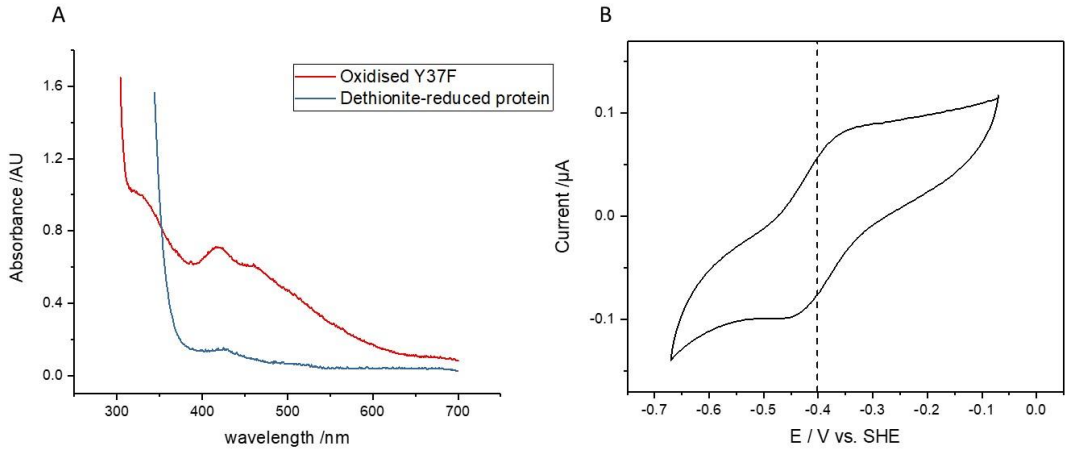


Figure 3.10 Panel A. UV-visible spectra of the variant Y37F. The red trace represents the oxidised form in 100 mM Tris-HCl pH 8 buffer; the blue trace represents the dithionite-reduced ferredoxin. Panel B. Cyclic voltammogram of the Y37F spinach ferredoxin variant measured in solution with a glassy carbon working electrode in 100 mM Tris-HCl 150 mM NaCl buffer pH 8, 25 ^{O}C.

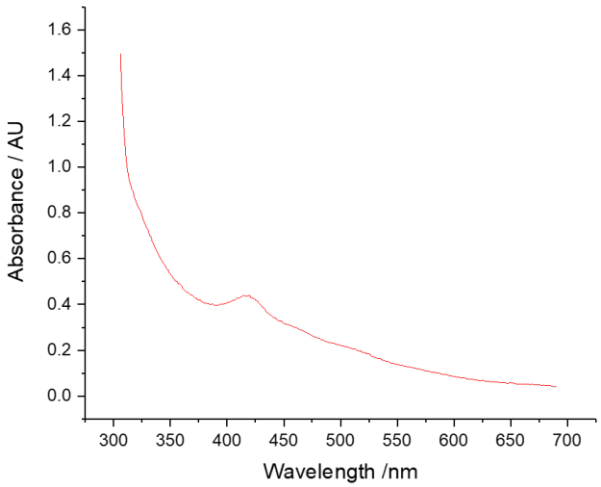


Figure 3.11 UV-Visible spectrum of the oxidised Y23F spinach ferredoxin variant in 100 mM Tris-HCl pH 8 buffer.

3.4 Introduction of the amber codon in the targeted positions.

The codons of each site were individually replaced by site-directed mutagenesis with the amber stop codon (TAG), which would be used for incorporating the unnatural amino acid in the next step. The mutations were also carried out via quick change PCR the plasmid of each variant was sent for sequencing. Before incorporating the unnatural amino acid *p*CNF, the Tyr-3, Tyr-37 and Phe-63 amber codon variants (Y3amber, Y37amber and F63amber) were expressed without adding the unnatural amino acid as a negative control to demonstrate that the FeS cluster cannot be formed. These amber codon variants should stop expression at the amber codon and, therefore, end up truncated. In the UV-Vis spectra shown in Figure 3.12, the [2Fe–2S] cluster peaks are absent for the Y3amber and Y37amber variants, consistent with the translation in the ribosome finishing at amber codon 3 and 37 in the ferredoxin genes respectively. Thus, the truncated protein cannot form the [2Fe-2S] cluster. For the F63amber variant, there is a small peak at about 420 nm. This might be because three of the four cysteines for cluster binding have been translated and hence, there may be partial binding of Fe atoms. However, this can still be utilised as 'non-correctly assembled cluster' negative control, to compare it with the protein containing the correctly assembled cluster.

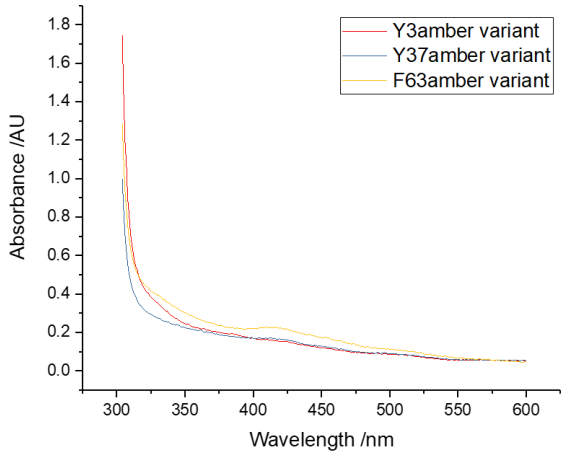


Figure 3.12 UV-Visible absorption spectra of the Y3amber, Y37amber and F63amber variants purified by GST-tag affinity chromatography. All variants were purified in 100 mM Tris-HCl pH 8 buffer. The blue trace, Y37amber, was purified in 100mM Tris-HCl pH 8 buffer. The yellow trace, F63amber, was purified in 100mM Tris-HCl pH 8 buffer.

3.5 FdI-CNF expression and characterisation

The ability to genetically incorporate an unnatural amino acid into a protein in living cells provides an important tool to probe the structure and function of the protein. The unnatural amino acid of interest is selectively incorporated into the protein in response to a reassigned nonsense codon by an aminoacyl-tRNA synthetase (aaRS)/tRNA pair that is orthogonal to the endogenous translational machinery of the host.[158] For expression of the spinach ferredoxin with *p*CNF incorporated (FdI-CNF), the corresponding *E. coli* strain was transformed with the arabinose-inducible pULTRA vector containing the genes for the unnatural aaRS/tRNA pair shown in Figure 3.13B.

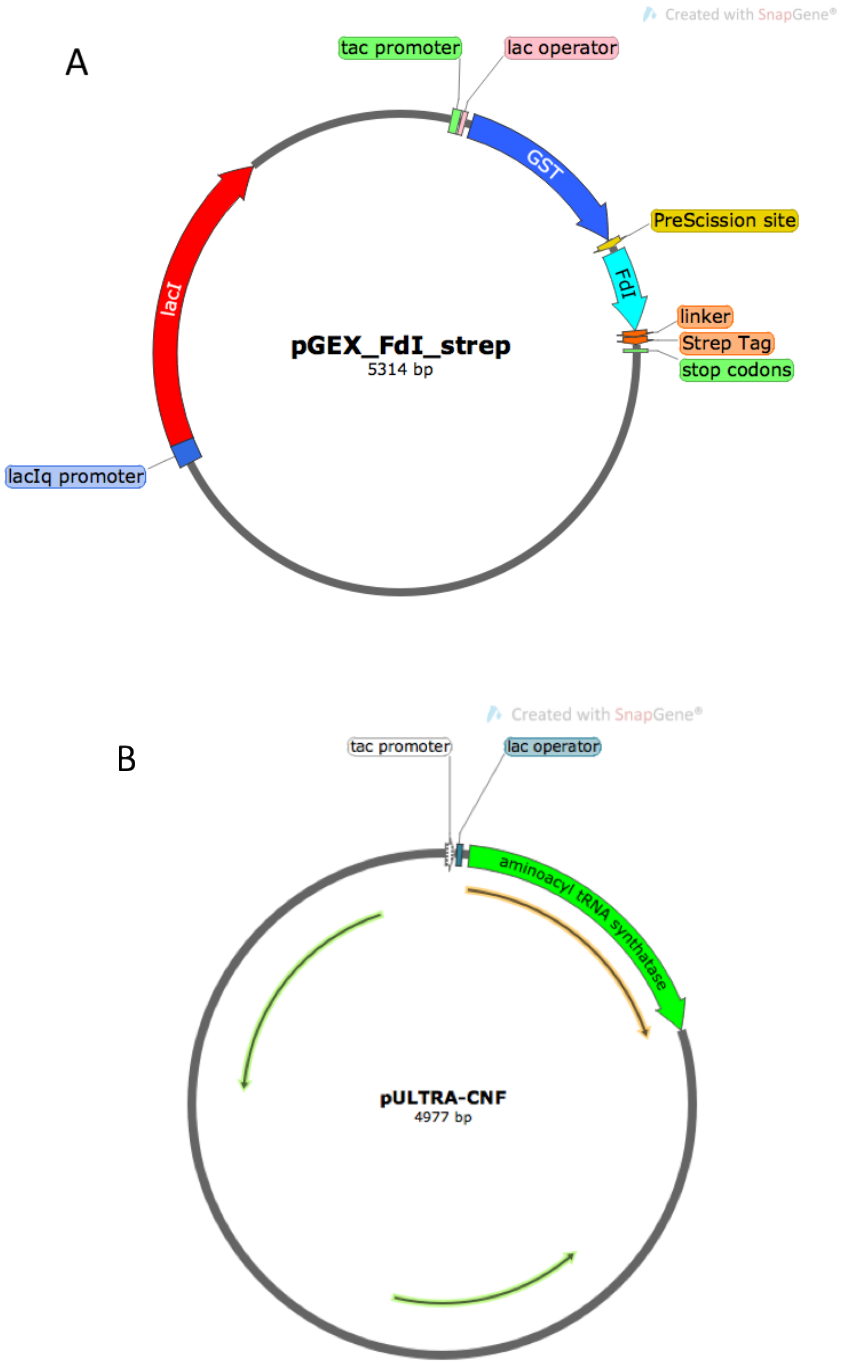


Figure 3.13 Panel A. Graphical structure of the plasmid pGEX_FdI_Strep. The turquoise arrow shows the inserted spinach ferredoxin gene, which is under the tac promoter control. Panel B. The graphical structure of the pULTRA-CNF plasmid, the green arrow shows the aminoacyl tRNA synthetase gene. (bp: base pairs.)

3.5.1 Expression and purification of FdI-CNF

The FdI-CNF over-expression with the pULTRA-CNF plasmid in *E. coli* showed much lower growth rate and final cell density. Obvious causes of such a reduced cell growth are the selection pressure due to the introduction of two antibiotics (ampicillin and spectinomycin). In previous studies, to overcome the resulting low yield of protein per unit culture volume, a second plasmid was used to over-express the simple target protein. However, in our case, the presence of the second promoter in the pULTRA-CNF vector is expected to mitigate this complication. Furthermore, LB media was replaced with the richer 2x YT media to help cell growth and gene expression. After several attempts varying the induction time and the concentration of the *p*CNF, an optimal protocol was found, which yielded sufficient pure protein to perform electrochemical and IR studies.

Figure 3.13A shows the new expression plasmid pGEX_FdI_Strep. A strep tag was added at the C-terminus by a polymerase chain reaction (PCR). During the expression of ferredoxin, both the truncated and full-length proteins are expressed, with the truncated protein in the majority, but it is difficult to separate them with GST-affinity chromatography, as both of them have the GST-tag at the N-terminal. In contrast, adding a strep tag at the C-terminal of the protein enables the use of Strep-Tactin affinity chromatography to purify the target protein since only the full-length spinach ferredoxin with *p*CNF incorporated contains the strep tag which can bind to the column. Therefore, it is easy to separate the truncated protein from the full-length spinach ferredoxin using this method. Before inoculation, the new plasmid was verified by sequencing. Both plasmids, the pGEX_FdI_Strep with amber codon (TAG) and the pULTRA-CNF, were transformed into the same BL21 (DE3) strain. The selected positions to mutate to the amber stop codon were Y3, Y37 and F63. The over-expression procedure is a modification of that of the wild type spinach ferredoxin I described in section 2.2.2. The *E. coli* strain BL21(DE3) competent cells bearing the pGEX_FdI_Strep plasmid with amber codon and pULTRA-CNF plasmids were

cultured in LB media containing 100 μg mL^{-1} ampicillin and 75 μg mL^{-1} spectinomycin. Subsequently, 1 μL of the stock culture was inoculated into a 15 mL sterilised Falcon-tube containing 10 mL of LB media and grown as the seed culture. Then, the seed culture was placed into the shaker and left overnight at 220 rpm and 37 °C. The next morning, the seed culture was diluted into a 2 L flask containing 1 L fresh 2x YT media with 100 μg mL^{-1} ampicillin and 75 μg mL^{-1} spectinomycin. The culture was incubated at the same speed and temperature as previously, until the OD_{600} reached 0.6. The culture was then cooled down in a water bath for 10 min, and *p*-cyano-L-phenylalanine (*p*CNF) dissolved in 0.5 M NaOH was added to the main culture to yield a final concentration of 8 mM. The same volume of 0.5 M HCl was also added to neutralise the pH of the culture, followed by addition of 0.1 mM final concentration of IPTG to induce protein production. The culture was left overnight at 16 °C and at 200 rpm to produce protein. After 20 hours of protein expression, the cells were harvested by centrifugation at 8000 *xg* for 10 minutes.

Due to the high cost of the *p*CNF, together with the low yield of the FdI-CNF protein, a strategy to recycle the media was developed. Thus, the media was tested to be utilised for another batch of FdI-CNF expression. The procedure is a modification of the standard ferredoxin preparation. The same amounts of tryptone, yeast and NaCl were added and dissolved into the recycled media, as done when preparing the fresh media. Then the recycled 2x YT media containing the fresh added tryptone, yeast and NaCl, was sterilised by filtering it through a 0.2 μm filter into a sterilised flask. Then the cell growth was started by placing the culture in the incubator at 37 °C and 220 rpm. Due to the slow growth of the bacteria in the recycled media, it usually takes 7 – 8 hours until the OD reaches 0.6. Then, as done before, the culture was cooled down in a water bath and the protein production was induced by adding IPTG to a final concentration of 0.1 mM. Finally, the culture was left in the incubator at 16 °C and 200 rpm for 20 – 24 hours.

It was found that the media cannot be utilised a third time as the bacteria cannot grow anymore shown in Table 3.2. It is assumed that this is due to accumulation of a large amount of toxic waste produced by *E. coli* cells during the cell growth and expression.

Table 3.2 Optical density (OD) of *E. coli* cells grown in the 2xYT media.

Times of media reusable	OD after 4 hours
1st time expression	0.6 – 0.8
2nd time expression	0.2 – 0.25
3rd time expression	< 0.1

To purify the protein, the culture was first centrifuged to harvest the cells. Then, the cell pellet was re-suspended in buffer W (100 mM Tris with 150 mM NaCl, pH 8.0) with deoxyribonuclease I (DNase I) to digest bacterial DNA and lysozyme to break the bacterial cell wall. Afterwards, the cells were lysed by sonication, followed by centrifugation again to separate the cell debris from the supernatant containing the protein. The soluble lysate was then filtered using a 0.2 μm filter to prevent the presence of any insoluble cell components that can block the column. Instead of using the GST-Trap column to purify the ferredoxin, strep-tactin affinity chromatography was carried out so that only the full-length protein with a strep tag at the C-terminus can bind to the column. In this way, the truncated proteins and other impurities can be separated from the targeted proteins. The supernatant was loaded onto a column (which was previously equilibrated by buffer W), washed with buffer W and then eluted with the elution buffer containing desthiobiotin (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 2.5 mM desthiobiotin). The fractions were analysed by SDS-PAGE to see if the correct protein had been eluted and to check its purity.

Figure 3.14A and Figure 3.14C show fractions from the purification of FdI expressed in fresh media and recycled media for Y37CNF and Y3CNF respectively. From the whole cell lysate and

soluble lanes, the majority of the expressed protein is truncated and the level of expression for the full-length ferredoxin (43.2 kDa) is very low. However, after loading the supernatant on to the Strep-tactin column (IBA Lifesciences), most of the full-length protein should bind to the column and can, therefore, be separated from the truncated proteins and other impurities. In the elution lane, it is clearly seen that there are two intense bands at about 43 kDa and 27 kDa, which are consistent with the ferredoxin with strep-tag at the C-terminus (43.2 kDa) and the truncated protein (26.5 kDa) respectively.

Figure 3.14B and Figure 3.14D show the purification results of Y37CNF and Y3CNF from the recycled media. There is not a big difference between the protein expressed in the fresh media and recycled media. After Strep-tactin affinity chromatography purification, the full-length ferredoxin with the incorporated *p*CNF could be separated from the truncated proteins and other impurities as well. Two intense bands are also visible in the elution lane. However, In Figure 3.14B, there is a middle band between the full-length and truncated protein which does not appear in Figure 3.14A, the characterisation results between the proteins from the fresh and recycle media illustrated that this band did not influence the property of the FeS cluster and function of the FdI. Further information can be obtained from the Western blot. The strep tag antibody was used to specifically bind with the full-length protein with the strep tag. If the full-length protein was expressed successfully, there would be a single band, and nothing would be shown in the Western Blot if there are only truncated proteins. Figure 3.15 shows the result of the Western blot of the Y37CNF variant. There is a single band at about 43 kDa that corresponds to the full-length protein.

Another selected residue, Phe63, did not give any full-length protein under the same expression and purification conditions. One reason might be that there is no space at position 63 in the FdI structure for the nitrile group in the *p*CNF which prevents the protein from folding properly, as the *p*CNF is bulkier than phenylalanine.

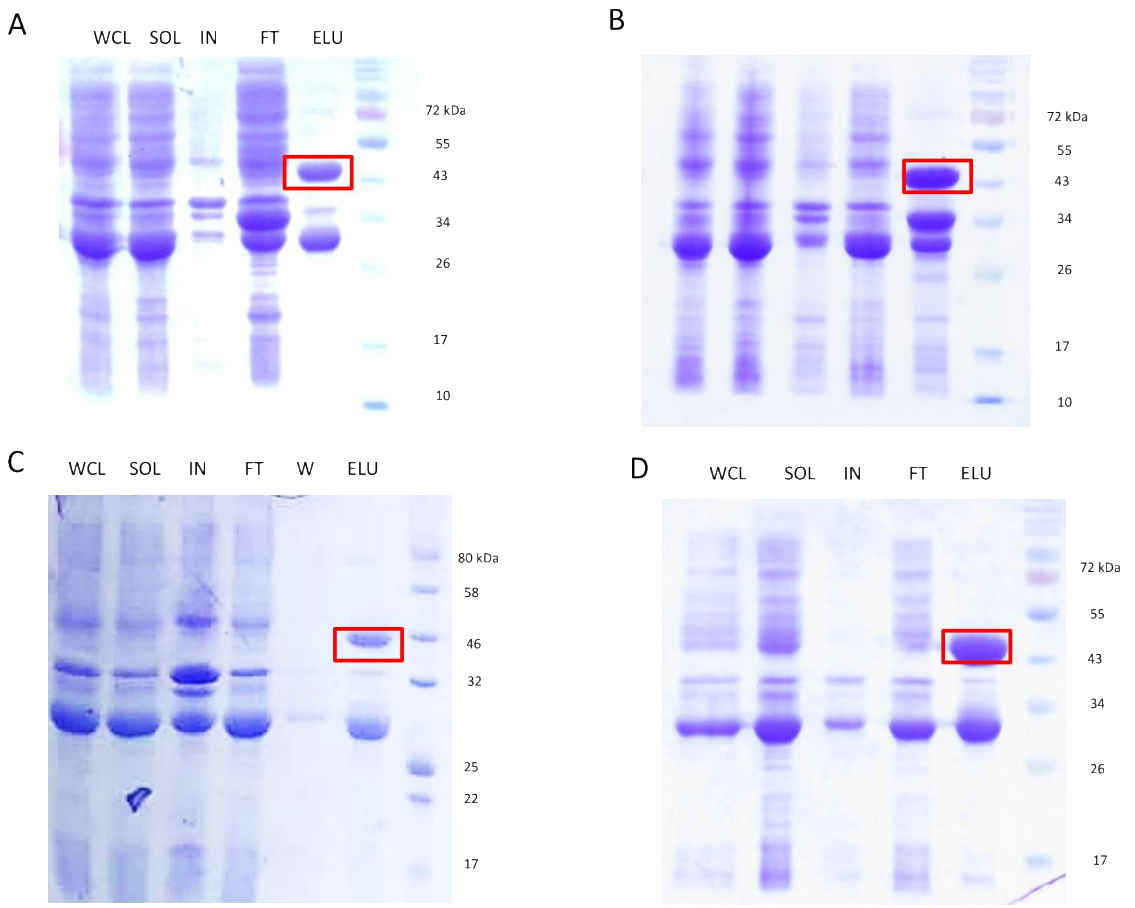


Figure 3.14 SDS-PAGE analysis. A: Y37CNF FdI expression in fresh media and purification. B. Y37CNF FdI variant expression in recycled media and purification. C. Y3CNF FdI expression in fresh media and purification. D. Y3CNF FdI variant expression in recycled media and purification. WCL: the whole cell lysate of *E. coli* cells overexpressing spinach ferredoxin. SOL: the supernatant. IN: insoluble fraction. FT: the supernatant flowing through the Strep-Tactin column. ELU: the collected fraction washed with desthiobiotin buffer.

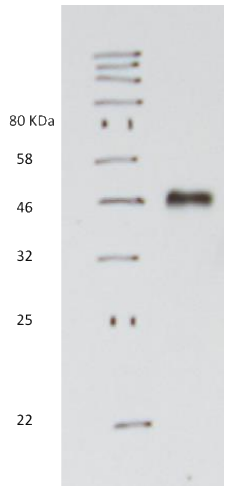


Figure 3.15 The Western blot of the GST-Linked Y37CNF FdI.

Table 3.3 shows the yield of Y3CNF and Y37CNF purified only by Strep-Tactin affinity chromatography, determined by absorption at 420 nm (A_{420}) using the molar extinction coefficient of the [2Fe-2S] cluster mentioned in section 3.2. From Table 3.3, It is clear to see that the protein yield from recycled media is higher than that from fresh media, and combined with the observation of the lower OD shown in Table 3.2, it suggests that slow growth of bacteria may help the protein expression.

Table 3.3 Yield of the Y3CNF and Y37CNF variants.

	Yield (from 1 L media)	**% Compared to native-sequence (wild type) GST-Tagged FdI** (from 1 L media)
Y3CNF (from fresh media)	1.51 mg	20.6%
Y3CNF (from recycled media)	2.52 mg	34.5%
Total	4.03 mg	
Y37CNF (from fresh media)	2.06 mg	28.2%
Y37CNF (from recycled media)	2.20 mg	30.2%
Total	4.26 mg	

3.5.2 Basic and functional characterisation of the unnatural ferredoxin I variants Y3CNF and Y37CNF

After successfully expressing and purifying the two variants with the unnatural amino acid *p*CNF incorporated (Y3CNF and Y37CNF), the next step was to characterise them. It is important to confirm that the *p*CNF unnatural amino acid can be incorporated without significantly affecting the [2Fe-2S] cluster structure or the overall fold of the protein. Thus, the presence of the [2Fe-2S] cluster was detected, and its behaviour was examined by UV-Vis spectroscopy and electrochemistry. Electrochemistry was also used to calculate the redox potential of the cluster and compare it to that of the 'native-sequence' (wild type) FdI.

The absorption spectrum of the two variants with the unnatural amino acid are shown in Figure 3.16. The characteristic peaks at around 320 nm, 420 nm and 460 nm for the oxidised protein, can be clearly seen in the "as isolated" spectra (red line) from Figure 3.16. The [2Fe-2S] cluster could be reduced by addition of sodium dithionite, causing all the peaks in the UV-Vis spectra to vanish. This is consistent with the 'native-sequence' (wild type) FdI behaviour. This indicates that the incorporation of the *p*CNF into the FdI did not affect the overall structure of the [2Fe-2S] cluster.

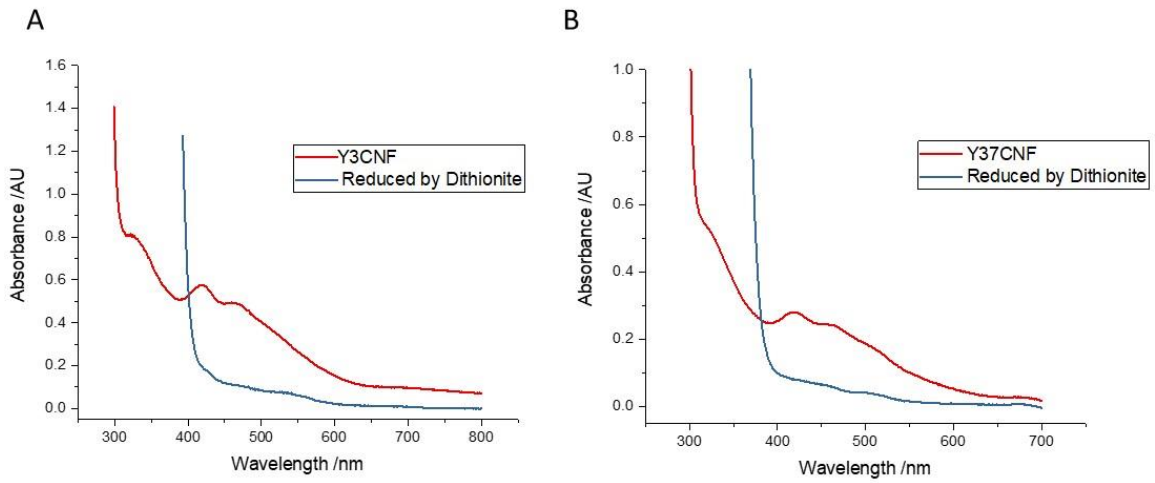


Figure 3.16 UV-visible absorption spectra of A: Y3CNF spinach Ferredoxin I purified by Strep-tactin affinity chromatography. B. Y37CNF spinach Ferredoxin I purified by Strep-tactin affinity chromatography. The red traces correspond to the oxidised ferredoxin in 100 mM Tris-HCl pH buffer; the blue traces represent the dithionite-reduced protein in 100 mM Tris-HCl pH 8 buffer.

Figure 3.17 shows the cyclic voltammogram of Y3CNF and Y37CNF compared with the 'native-sequence' (wild type) spinach ferredoxin. For each, a reversible wave is observed with a potential of –0.405 V. Therefore, the redox potential of these two variants is the same as in wild type FdI, which indicates that the *p*CNF incorporation into the protein did not disturb the redox properties of the cluster.

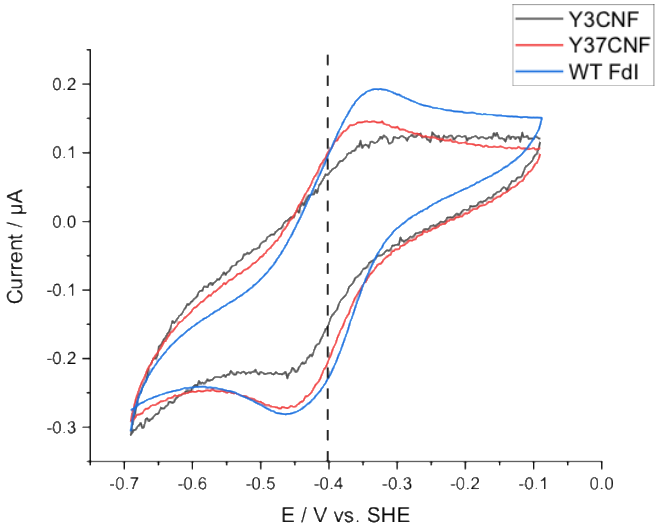


Figure 3.17 Cyclic voltammograms of the 'native-sequence' wild type (WT) spinach Ferredoxin (blue line), Y3CNF variant (black line) and Y37CNF variant (red line) measured in solution with a glassy carbon working electrode in 100 mM Tris-HCl 150 mM NaCl buffer pH 8, 25 ^{O}C.

3.5.3 Infrared spectroscopy of the unnatural amino acid FdI variants Y3CNF and Y37CNF

The FdI variants with the *p*CNF incorporated were expressed and purified successfully. Once the properties of the [2Fe-2S] cluster were characterised by UV-Vis spectroscopy and electrochemistry, infrared (IR) spectroscopy was carried out to determine whether *p*CNF could be used as a probe of the [2Fe-2S] cluster oxidation state.

The C≡N stretching vibration (ν_{CN}) has a strong absorption in a region of the spectrum (about 2200 cm^{-1}) free from other vibrational absorptions typically present in proteins.

A negative control experiment with the native Y37F variant was first carried out in a transmission IR cell and then compared with Y37CNF. Figure 3.18A shows raw IR spectra of the the Y37F and Y37CNF FdI variants. 2nd derivative IR spectra (shown Figure 3.18B) clearly indicate a difference between Y37F and Y37CNF FdI. Y37F has no nitrile group and therefore no CN stretching vibration in its IR spectrum (red trace). However, when the *p*CNF is incorporated in the FdI Y37CNF variant, the signal of the nitrile group at 2237.0 cm^{-1} can be clearly seen (blue

trace), and the band appears in the normal region for *p*CNF stretching (between 2229 cm^{-1} and 2236 cm^{-1}).[159] This also demonstrates that the unnatural amino acid (*p*CNF) has been successfully incorporated into the FdI.

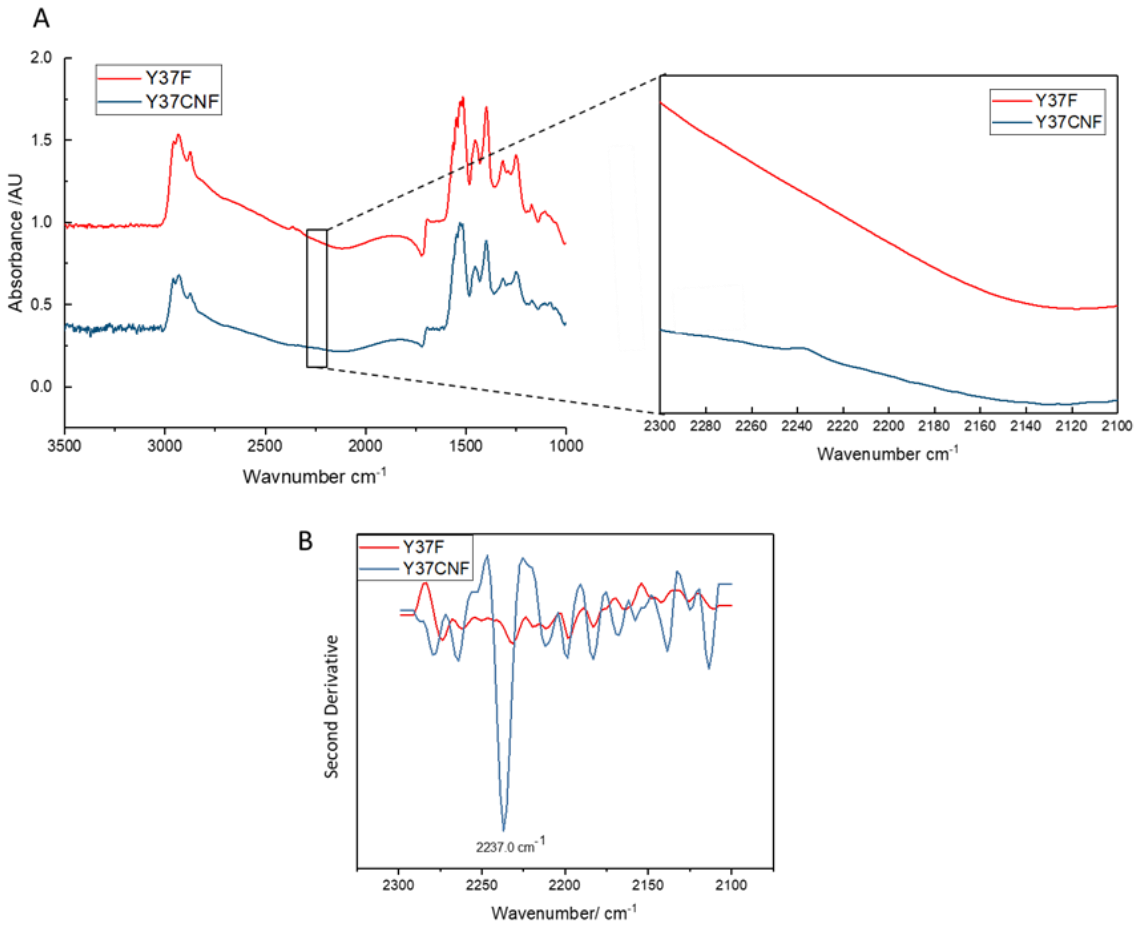


Figure 3.18 Panel A: Raw data of transmission-mode infrared spectra of Y37F and Y37CNF FdI, zoomed in between 2100-2300 cm^{-1}. Panel B: The transmission-mode infrared spectra of the Y37F variant and Y37CNF variant displayed in the second derivative. Conditions: 100 mM Tris-HCl 150 mM NaCl buffer pH 8, 25 ^{O}C, 4 cm^{-1} resolution. Protein concentration: 1.5 mM.

3.6 Discussion and conclusion

3.6.1 Expression and purification of spinach ferredoxin I (FdI)

The first goal to achieve in this project was the overproduction of FdI in *E. coli*, as much more protein would be needed for incorporating the unnatural amino acid (*p*CNF). However, the

spinach ferredoxin protein was challenging to over-express in *E. coli.* After several attempts with different vector systems, the pGEX system was the one which yielded the best results. Because GST folds rapidly into a stable and highly soluble protein upon translation, inclusion of the GST tag often promotes greater expression and solubility of recombinant proteins than expression without the tag.[160] Therefore, the GST genes in the front of the FdI genes were able to improve the expression level of FdI. Another advantage of the pGEX system is that the GST tag could be used to purify the protein by affinity chromatography. Thus, affinity chromatography was the first purification step followed by anion exchange chromatography. As the isoelectric point of the FdI (3.6) is much lower than for *E. coli* proteins (average is 6.0), anion exchange chromatography is an effective purification method for the FdI since it binds tightly to the positively charged resin. The protein can then be eluted by increasing the concentration of salt (NaCl). By increasing the NaCl concentration, the electrostatic attraction between the negatively charged protein and the positively charged resin are disrupted so that the protein is finally eluted. The FdI was eluted at about 0.55 mM NaCl in Tris buffer pH 7.4. After this step, most of the impurities had been removed. Further purification was achieved by size exclusion chromatography. Every purification step was monitored by SDS-PAGE.

At this stage, a reproducible protocol to over-express and purify spinach ferredoxin I was established. This same protocol was then used for expression and purification of different variants generated by side-directed mutagenesis. However, the FdI band shown in the SDS-PAGE shows higher molecular weight than expected. This may be explained by the fact that the content of acidic amino acids is higher than that of basic amino acids and thus the acidic character of FdI prevents proper binding of SDS which makes the FdI not diffuse appropriately in the SDS-PAGE.[161]

The concentration of recombinant FdI could be determined by UV-Vis spectroscopy by the absorption at 420 nm (A_{420}) with a molar extinction coefficient of 9400 $cm^{-1}M^{-1}$ for the [2Fe-2S]

cluster.[162] The final yield of the over-expressed spinach FdI generated by the GST fusion system is about 7.3 mg protein for 1 L of culture.

3.6.2 Characterisation of the overexpressed spinach ferredoxin I (FdI)

After developing a protocol for over-expression of FdI, UV-Vis spectroscopy and electrochemistry were carried out to characterise the purified protein. The results are shown in section 3.2.2. From the UV-Vis spectra, the characteristic peak of $[2Fe\text{-}2S]^{2+}$ could be detected, in agreement with previously reports in the literature.[28] Then, the protein could be successfully reduced by addition of sodium dithionite. This showed that the [2Fe-2S] cluster in the ferredoxin functions properly. The comparison between the pure recombinant FdI and the GST linked FdI elucidated that the GST tag did not affect the properties and function of the FdI. Afterwards, electrochemistry was also carried out with the GST linked FdI to further test the redox behaviour of the [2Fe-2S] cluster in the Ferredoxin. The redox behaviour of the over-expressed GST-FdI in the electrochemistry looked completely normal, and the midpoint redox potential extracted from the cyclic voltammetry (–0.405 V, vs. SHE) is in good agreement with what has been published previously for the native FdI.[163]

As the GST tag did not affect the function and properties of the FdI, it was decided to only purify the different variants generated afterwards by GST affinity chromatography to further perform spectroscopic and electrochemical investigations. From the UV-Vis spectra of both GST-tagged and untagged FdI, we can conclude that there is no difference on the [2Fe-2S] cluster behaviour. In this way, protein loss during further subsequent purification steps was avoided.

3.6.3 Selection of amino acids close to the FeS cluster for site-directed mutagenesis

Six residues close to the [2Fe-2S] cluster were selected to perform site-directed mutagenesis. First, the selected amino acid was mutated to an alternative similar amino acid to test the

tolerance of each site toward mutagenesis and to see if the protein could fold properly. The results are shown in the Table 3.1 section 3.3.1. They indicate that some residues had a big impact on the properties of the cluster and more importantly on the protein itself as some of the variants could not be expressed as the holoprotein. Finally, three residues (Tyr 3, Tyr 37 and Phe 63) were selected for mutagenesis as they showed the same characteristics as the wild type FdI by UV-Vis spectroscopy and electrochemistry.

3.6.4 Expression and purification of variants with *p*CNF

After the three residues were selected for mutagenesis, the next step was to incorporate the unnatural amino acid *p*CNF into the FdI (FdI-CNF) at these positions. The plasmid pULTRA-CNF was transformed into the *E. coli* cells in response to a reassigned nonsense or frameshift codon by an aminoacyl-tRNA synthetase (aaRS)/tRNA pair that is orthogonal to the endogenous translational machinery of the host. For easy purification of the artificial amino acid FdI (FdI-CNF), a strep tag was added at the C-terminus of the FdI gene, which can be separated from the truncated protein by Strep-tactin affinity chromatography. The over-expression protocols for the variants were slightly modified from that for the native protein. After testing several conditions, a successful protocol to overexpress the FdI-CNF variants was developed. The media with *p*CNF can be used twice for over-expression FdI-CNF variants. The yield of the FdI-*p*CNF variants was much lower than for the wild type FdI. However, the Y37CNF and Y3CNF variants were successfully expressed and purified, and they were used for the IR spectroscopy experiments.

3.6.5 Characterisation of the FdI-CNF variants via electrochemistry and infrared spectroscopy

Before carrying out infrared spectroscopy on the FdI-CNF variants, electrochemistry was performed to test if the artificial amino acid incorporated in the protein affects the redox properties of the [2Fe-2S] cluster. The results shown in Figure 3.17 of section 3.4.2 indicate that the redox potential of both variants, Y37CNF and Y3CNF, was –0.405 V (*vs.* SHE), identical to the

midpoint redox potential found for the WT FdI. Thus, the incorporation of the *p*CNF did not cause any functional perturbation in the cluster and the whole protein.

Transmission-mode IR experiments were done on both Y37CNF and Y37F. The native Y37F variant has no nitrile group and therefore no C≡N stretching vibration can be seen in its IR spectrum while there is a clear peak at 2237 cm^{-1} in the IR spectrum on the FdI Y37CNF variant. Hence, *p*CNF was successfully incorporated into FdI and the nitrile group of *p*CNF in the protein had a relatively strong IR signal.

3.7 Conclusion

In this chapter, a novel method to study the long-range effect and the redox states of FeS cluster in a protein is established, by incorporating an unnatural amino acid. The amino acid *p*-cyano-L-phenylalanine (*p*CNF) was used in this research, since the nitrile group in this compound is a sensitive reporter in infrared spectroscopy. A simple FeS protein spinach ferredoxin (FdI) was selected as a model system. The overproduced system and purification protocol of FdI were established. This same protocol was used for the expression and purification of different variants generated by site-directed mutagenesis. *p*CNF could be successfully incorporated into the protein, giving two variants Y3CNF and Y37CNF. Y37 is close to the FeS cluster (7.4 Å) while Y3 is far away (22.6 Å) and hence can act as a negative control. The preliminary IR results illustrated that *p*CNF in FdI can show an IR signal. Further studies on these variants are described in Chapter 4.

In addition, various nitrile-containing unnatural amino acids (as the CNP) are also commercially available.

4

Chapter 4: Probing FeS Cluster Redox States by Incorporating an Unnatural Amino Acid

4.1 Introduction and aim of the chapter

Incorporation of unnatural amino acids into a protein provides a novel approach to study local protein structure and dynamics. There are a variety of unnatural amino acids with functionalised side chains which can act as spectroscopic reporters, NMR probes and fluorescent groups, and these have all been incorporated into protein systems successfully.[132,164] The field of protein dynamics is not only concerned with the transitions between states, but also the nature and equilibrium populations of the states themselves and recent efforts have focused on using sensitive IR-absorbing probes to characterise local protein dynamics.[136] The two commonly used unnatural amino acids which contain vibrational reporters to study local protein structure and dynamics are *p*-cyano-L-phenylalanine (*p*CNF) and *p*-azido-L-phenylalanine (*p*N_3).[128,135,165] Incorporation of *p*CNF allows the stretching frequency of the nitrile group to be used as a probe of the local protein environment. The structure of *p*CNF is quite similar to tyrosine and phenylalanine, therefore, tyrosine and phenylalanine are ideal target sites to incorporate the unnatural amino acid *p*CNF.[136] By following the CN stretching vibration the changes in the local protein environment can be indirectly observed, as ν_{CN} is extremely sensitive to changes in the electron density or oxidation state in proteins.[151,166,167]

From the preceding chapter, a system for over-producing spinach ferredoxin was developed in *E. coli*, and two unnatural amino acid variants Y3CNF and Y37CNF, were expressed successfully and characterised by UV-Vis spectroscopy and electrochemistry. In this chapter, Infrared spectroscopy was carried out on Y37CNF FdI to probe whether the CN absorption band position responds to the different oxidation states of the FeS cluster.

4.2 Infrared spectroscopy on FdI-CNF variant

Transmission mode IR spectroscopy was performed on oxidised and reduced samples of Y37CNF, to see if the vibrational frequency of the CN substituent shifts upon reduction of the [2Fe-2S]

cluster, to assess whether CNF can be used as an IR probe to determine the oxidation state of a cluster.

UV-visible spectroscopy was used to independently verify the redox state of the cluster. The protein used for UV-vis spectroscopy was 100 times diluted from the concentrated protein used for IR studies (about 2 mM). Figure 4.1A and B show the UV-Vis spectra and the base-line subtracted IR spectra respectively of the oxidised Y37CNF ferredoxin variant. The characteristic peaks of the oxidised [2Fe-2S]$^{2+}$ cluster can be seen in the UV-Vis spectrum, thus confirming that the protein is in its oxidised state. In the transmission IR spectrum (Figure 4.1B), a sharp peak at 2237 cm^{-1} was observed, consistent with the CN stretch of the *p*CNF observed previously for the as-isolated protein. The protein was chemically reduced by addition of 10–20 mM sodium dithionite. Excess reducing agent was removed by buffer exchange with Buffer W (100 mM Tris-HCl, 150 mM NaCl pH8), and the reduced protein was concentrated to 2 mM. To confirm that it was fully reduced, a UV-Vis spectrum was also measured on the reduced form (Figure 4.1C). The disappearance of the three peaks at around 320 nm, 420 nm and 460 nm confirmed that the protein had been reduced. Subsequently, the reduced protein was loaded into the transmission IR cell and an IR spectrum was recorded (Figure 4.1D). Figure 4.1B and Figure 4.1D show that there must be a different absorption coefficient for ν_{CN} of oxidised and reduced states of FdI since the IR signals from the same concentration of oxidised and reduced FdI are of quite different intensities. There is a small but observable shift in the peak position between the oxidised and reduced Y37CNF samples (2237.0 cm^{-1} in the oxidised form *vs.* 2235.1 cm^{-1} in the reduced form). Thus, the C≡N stretching frequency differs by 1.9 cm^{-1} between the oxidised and reduced forms (Figure 4.1B and D), suggesting that ν_{CN} is sensitive to a change in the local electrostatic environment of the [2Fe-2S] cluster in these two redox states. Similar shifts (1.5 cm^{-1}) were previously observed in cytochrome *c* variants carrying the unnatural amino acid *p*CNF when the haem centre was oxidised and reduced, and in studies with myoglobin variants

incorporating *p*CNF.[131] The raw data of oxidised and reduced Y37CNF whole spectra are shown in Figure 4.1E, with zoomed in spectra between 2210 – 2270 cm^{-1}.

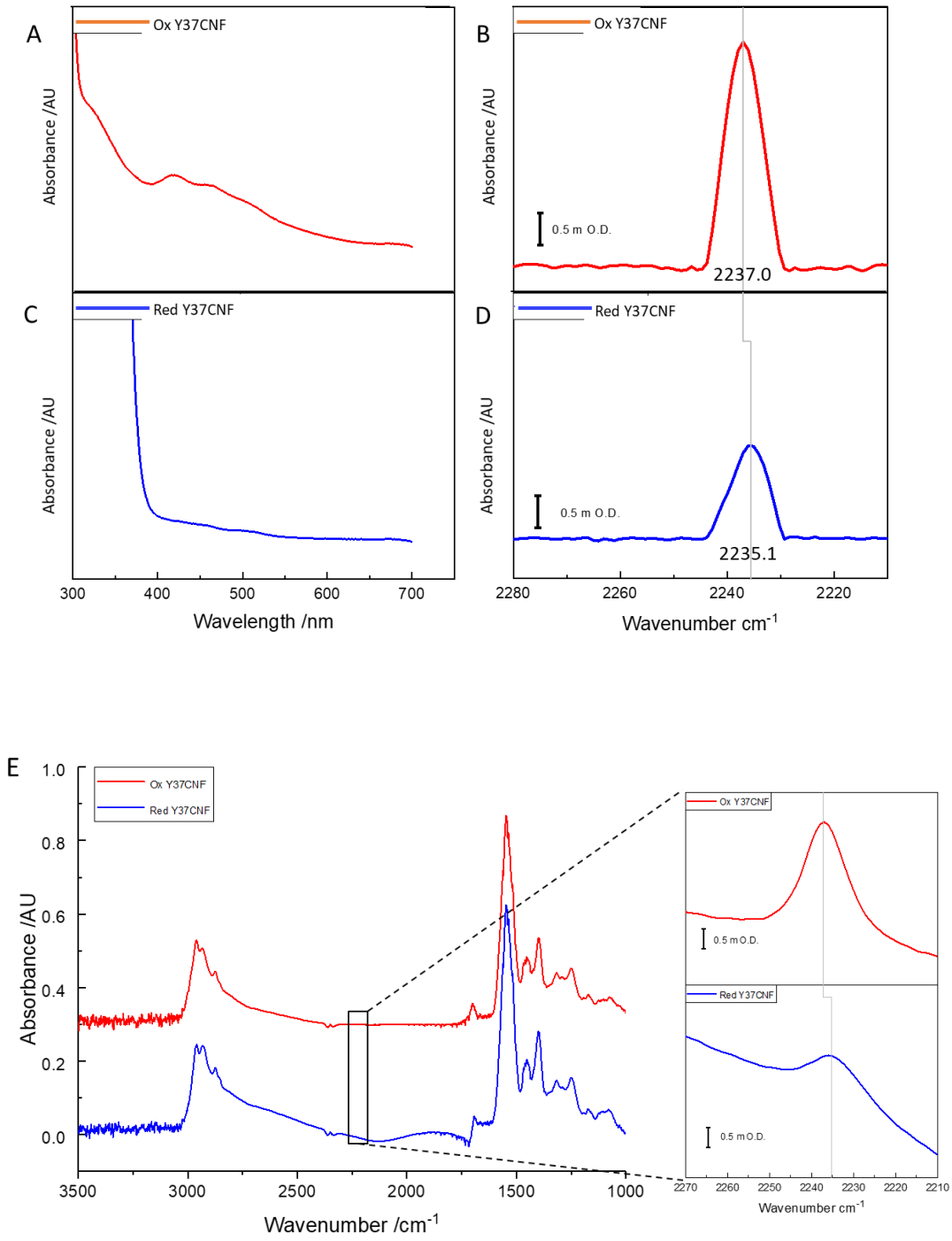


Figure 4.1 Panel A: UV-Visible spectrum of the oxidised Y37CNF; Panel B: Base-line subtracted IR spectrum of the oxidised Y37CNF ferredoxin showing in the nitrile stretching region. Panel C: UV-Visible spectrum of the reduced Y37CNF; Panel D: Baseline subtracted IR spectrum of the reduced Y37CNF

ferredoxin showing in the nitrile stretching region. Panel E: Raw data of transmission-mode infrared spectra of both oxidised and reduced Y37CNF FdI, zoomed in between 2210-2270 cm^{-1}. Conditions: 100 mM Tris-HCl 150 mM NaCl buffer pH 8, 25 °C. 2 cm^{-1} resolution, 1024 scans.

Reflection mode Infrared microspectroscopy was also carried out on the Y37CNF FdI variant. The cell was designed by the Vincent group (see section 2.7.1 for more detail). It is a spectroelectrochemical cell consisting of a small polished glassy carbon electrode surrounded by a carbon counter electrode and a nearby hole to connect a small Ag/AgCl commercial reference electrode. The detail and 3-dimensional structure of this cell has been described in section 2.7.1. The advantage of using this cell is that electrochemical control can be used to oxidise or reduce the sample at the same time that IR is being measured, using a small volume (about 10 μL) of protein. According to the redox potential estimated from the cyclic voltammetry (–405 mV *vs.* SHE), applying a potential of –650 mV (*vs.* SHE) should reduce the protein and a potential of –5 mV (*vs.* SHE) should oxidise the protein.

First, a negative control experiment with just the amino acid *p*CNF in solution was performed to ensure that the ν_{CN} from the free *p*CNF amino acid in solution does not shift with the different applied potentials. Figure 4.2 shows the IR spectra recorded at two different potentials (–650 mV and –5 mV) for the free *p*CNF in Tris pH 8 buffer. The signal of the cyano probe appears at 2237.0 cm^{-1} and it does not change with the different applied potentials. Thus, the C≡N stretching vibration ν_{CN} of the free cyano-phenylalanine is not influenced by the applied potential.

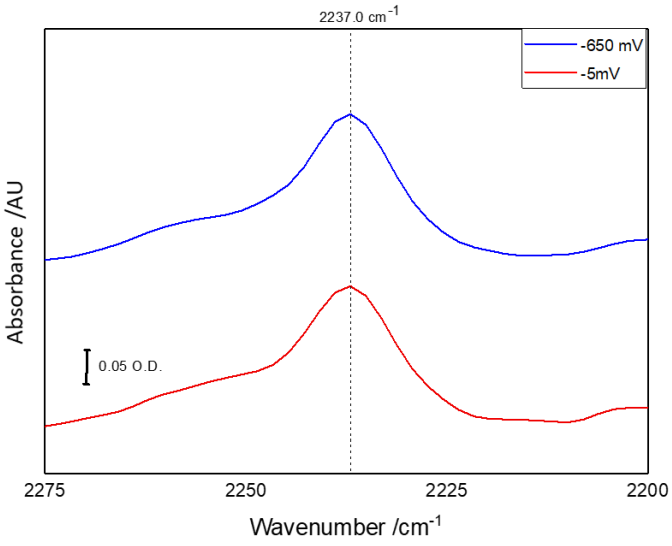


Figure 4.2 Base-line subtracted IR spectra of 100 mM *p*CNF in 100 mM Tris-HCl 150 mM NaCl buffer pH 8 measured in the reflection mode at 25 °C. 2 cm^{-1} resolution, 1024 scans.

It was found to be important to be careful when concentrating the FdI since too high a concentration results in a gel-like texture and complete loss of the cluster redox signals. Thus, the Y37CNF variant was concentrated to about 2 mM. The two potentials, –650 mV and –5 mV were applied to reduce and oxidise the Y37CNF ferredoxin variant, respectively. The potential steps were repeated twice to make sure that the cluster was not disturbed by the applied potential and that the oxidation/reduction process was reversible. Figure 4.3 shows the second derivative of the IR spectra of the Y37CNF variant. A weak but clearly distinguishable peak in the C≡N stretching region can be seen. Although the peaks are not very intense, the second derivatives allow for a reliable wavenumber determination. The second derivative spectrum revealed a 2237.0 cm^{-1} peak position for the oxidised form and 2235.1 cm^{-1} for the reduced form. Thus, 1.9 cm^{-1} shift towards lower energy is observed in the reduced form. This agrees with the previous results obtained in the IR transmission cell after chemical reduction (Figure 4.1). The wavenumber differences between the two redox states of the cluster indicate that the local electric field differences from the redox change extended towards the environment of the

cyano group of Y37CNF. Both frequencies (for oxidised and reduced forms) fall in the range of a non-hydrogen bonded nitrile functional group.[167]

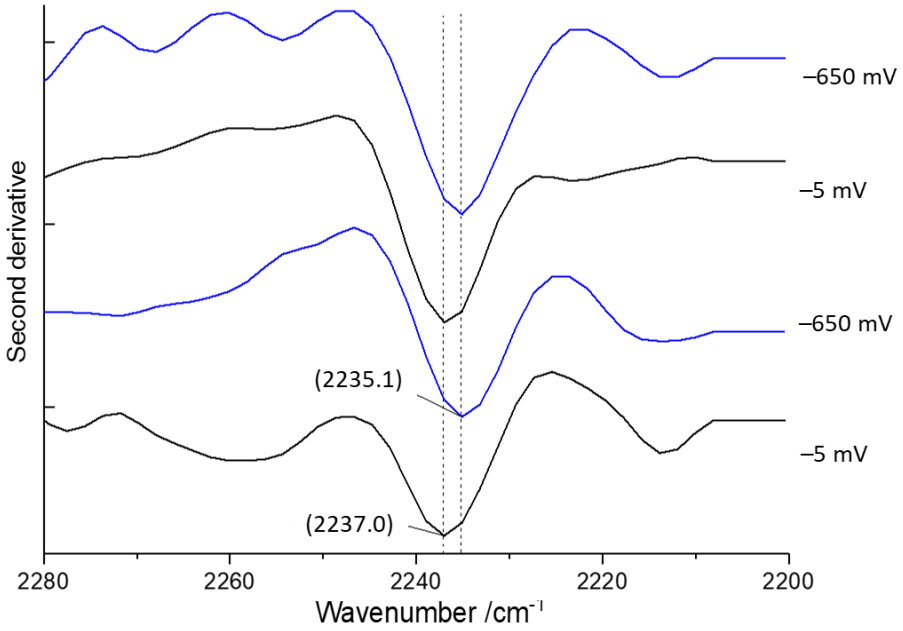


Figure 4.3 IR spectra of Y37CNF measured in the reflection SEC cell displayed in the second derivative mode. Conditions: 100 mM Tris-HCl 150 mM NaCl buffer pH 8, 25 °C. 2 cm^{-1} resolution, 1024 scans. Protein concentration: 2 mM.

The other variant, Y3CNF was intended as the negative control for this experiment as it is far away (22.6 Å) from the FeS cluster. However, insufficient protein was isolated for full IR spectroscopic characterisation of this variant so far.

4.3 FdI crystallisation

Next, crystallisation was carried out on wild type spinach ferredoxin and variants to confirm (i) there are no major structural changes between oxidised and reduced FdI, so that it can be sure the shift in ν_{CN} is not from a structural change; (ii) the mutations at position 3 and 37 do not alter the structure, and also to measure the distance between *p*CNF and the FeS cluster. Methods and crystal structure parameters are shown in Appendix B. Crystallisation of FdI and variants, X-ray

diffraction and solution of protein structures was carried out by Dr. Stephen Carr (University of Oxford).

Given the high resolution of the X-ray diffraction data, the resulting electron density maps were of exceptional quality and clearly displayed each of the introduced mutations (Figure 4.4). Superposition of the structures in Figure 4.5 shows that changing the identity of the amino acid at either position 3 or 37 has no effect on the structure of the protein either at a global or local level, each Y/F/CNF side chain adopted the same conformation. Root-mean-square deviation (RMSD) is used for measuring of the average distance between the atoms of the superimposed protein, and the summary of RMSD between wild type FdI and variants are shown in Table 4.1.

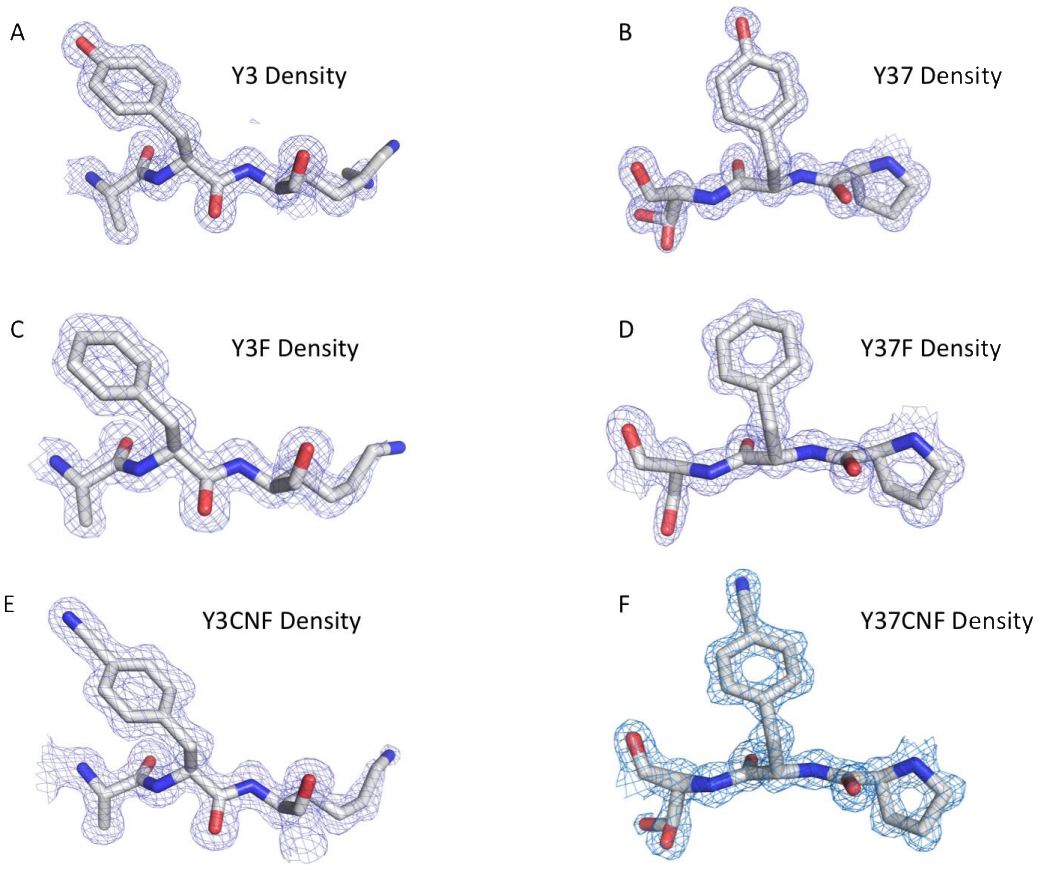


Figure 4.4 Electron density for specific residues in wild type FdI and variants. A: Y3 of wild type FdI; B: Y37 of wild type FdI; C: F of Y3F variant; D: F of Y37F variant; E: *p*CNF of Y3CNF variant; F: *p*CNF of Y37CNF variant. Crystallisation of FdI and variants, data collection, and structural solution and refinement were carried out by Dr Stephen Carr.

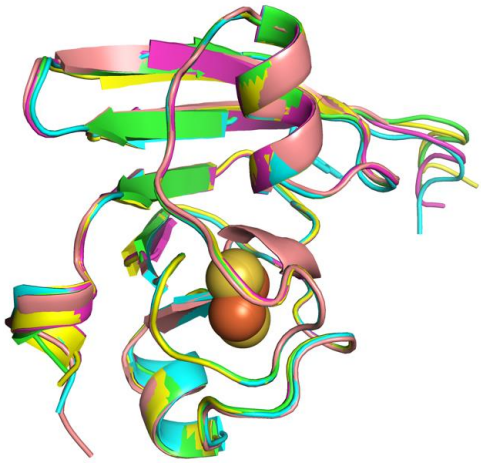


Figure 4.5 Superposition of wild type (green), Y3F variant (cyan), Y37F variant (yellow), Y3CNF variant (magenta), Y37CNF variant (orange). The structure work was credited to Dr Stephen Carr.

Table 4.1 Superposition differences between wild type FdI and variants.

Proteins	Root-mean-square deviation (R.M.S.D.)
WT vs Y3F	0.31 Å
WT vs Y3CNF	0.35 Å
WT vs Y37F	0.31 Å
WT vs Y37CNF	0.60 Å

Melting temperature is a method to test protein stability. A solution of protein is heated until the point of its complete denaturation and melting temperature is defined as the temperature at which the native and denatured states are evenly distributed at equilibrium.[168] Comparing the melting temperature between wild type and variants makes it possible to investigate the effect of specific residues on protein stability. The melting temperature data of wild type spinach

ferredoxin and variants shown in Table 4.2 indicated that the solution stability of variants and wild type FdI are also very similar.

Table 4.2 Melting temperature of wild type FdI and variants.

Proteins	Melting Temperature (duplicate measurements measured in Celsius degrees)
WT	44.0, 44.1
Y3F	42.1, 42.5
Y37F	42.2, 42.3
Y3CNF	43.2, 43.3
Y3CNF	44.8, 44.4

The root-mean-square deviation (RMSD) is 0.6 Å between the Y37CNF and wild type FdI. The difference is due to changes in conformation of a linker (bottom left) and loop near the N-terminus (bottom right) shown in Figure 4.6. The small structural change in Y37CNF is due to H-bonding to a neighbouring protein molecule (Y37 is at the edge of a crystal contact) and Y37 forms a H-bond to E71 in the other molecule in wild type FdI. The CN group in Y37CNF cannot do this; instead it interacts with Q68, resulting in positional shift, while the position of the main chain is not affected. The FeS cluster environment of Y37CNF variant is similar to that of the WT FdI, the similarity in the cluster environment is consistent with the native-like UV-Vis spectra and redox potential of the Y37CNF.

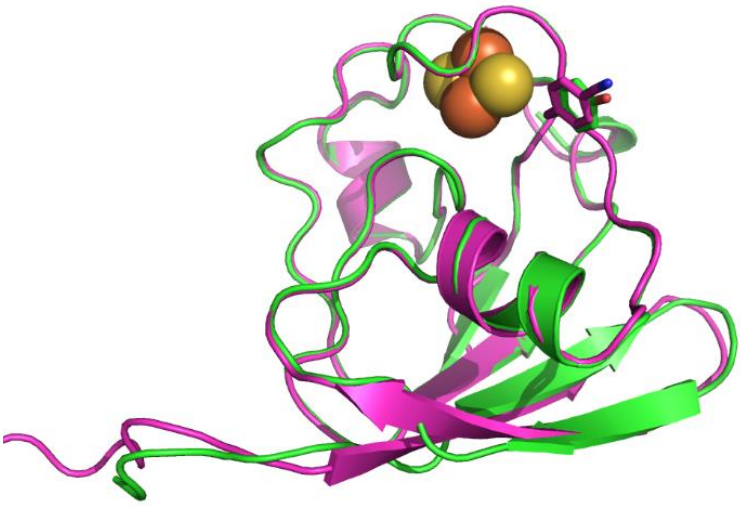


Figure 4.6 Superposition of wild type (green) and Y37CNF variant (magenta). The structure work was credited to Dr Stephen Carr.

Furthermore, to confirm that the shift in ν_{CN} for the oxidised and reduced variant does not arise from a structural change accompanying oxidation and reduction, structures were performed on both oxidised and reduced WT FdI, while insufficient Y37CNF was available to obtain a reduced structure. The crystal structures are essentially identical with small differences in mobile side-chains at the protein surface (shown in Figure 4.7) which indicates that changing the oxidation state of the FeS cluster has no effect on the position of the side chains relative to the cluster. The RMSD between Cα positions between the oxidised and reduced FdI is 0.06 Å.

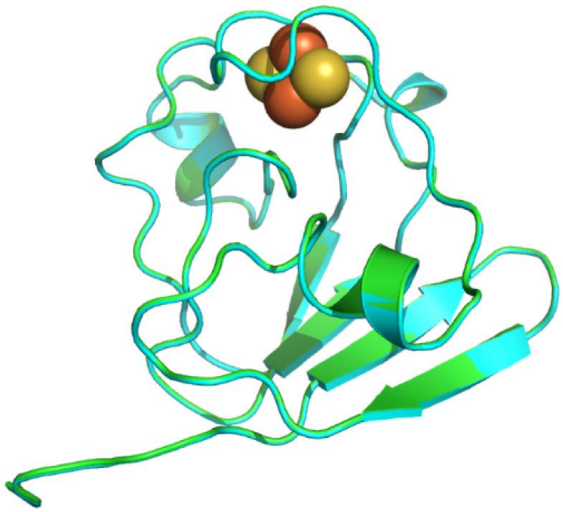


Figure 4.7 Superposition of oxidised (green) and reduced (cyan) wild type FdI. The structure work was credited to Dr Stephen Carr.

4.4 Discussion and conclusion

4.4.1 Infrared spectroscopy on FdI-CNF variant

A clear signal at 2237.0 cm^{-1} is seen for the as-isolated form of Y37CNF (oxidised) in IR experiments. This signal is in the nitrile stretching region thus belonging to the CN group of the unnatural amino acid cyano-phenylalanine within the protein (confirmed by UV-Vis spectroscopy). After reducing the protein, the peak was red-shifted to 2235.1 cm^{-1}. Therefore, there is a distinguishable difference in the vibrational frequency of the oxidised and reduced forms of Y37CNF ferredoxin, which can be associated with the oxidation state of the [2Fe-2S] cluster, suggesting long range sensitivity of the CN mode to cluster redox state.

The shift was confirmed separately by IR spectroelectrochemistry (SEC). The Y37CNF variant was measured in a reflection mode SEC cell. As the nitrile signal was not as intense as in the transmission mode IR, only second derivative spectra are displayed. In the second derivative spectra the 1.9 cm^{-1} peak shift between the oxidised and reduced forms of Y37CNF can be distinguished. This shift is associated with the change in the electric field strength along the

nitrile bond (2.5×10^6 V m^{-1}) upon $[2Fe\text{-}2S]^{2+}$ cluster reduction.[127] Moreover, ν_{CN} of the oxidised protein is indistinguishable from the free amino acid in solution, illustrating that the protein does not perturb the CN bond in the oxidised state; this is also consistent with the observation that incorporation of *p*CNF does not perturb the redox state of the FeS cluster.

More importantly, the identical shift in ν_{CN} observed for chemical and electrochemical reduction shows that this shift is not due to a change in solution potential or an electric field effect from the electrode but must be associated with the cluster redox state.

4.4.2 FdI crystallisation

Additional residues at the N-terminus, that remain after cleavage of the GST-tag and some of which are visible in the $2F_o\text{-}F_c$ electron density maps, are the most likely cause of the different unit cell parameters between the native protein and previously reported structure (PDB code: 1A70). Changing the identity of amino acids on the surface of the protein can alter both the charge distribution and hydrophobicity of the protein surface resulting in different crystal packing and hence the observed changes in the unit cell and symmetry.

Positions 3 and 37 in FdI are tolerant of mutations, with both sites able to accommodate F and *p*CNF at this position with negligible effect on the structure of the protein or the loading of co-factor suggesting FdI makes an ideal system to test the feasibility of monitoring transitions between redox states of FeS clusters by the addition of extrinsic IR probes. The FeS cluster environment of these variants is similar to the WT FdI, which is consistent with the native-like UV-Vis spectra and redox potential.

Ultra-high resolution structures of the High-Potential Iron-sulfur Protein (HiPIP) show Fe-S bond lengths of the the [4Fe-4S] cluster change only slightly upon oxidation/reduction (0.005 – 0.01 Å).[169] This suggests small changes likely occur in the [2Fe-2S] cluster of FdI, but are not

observed in the structures presented since the changes are less than the precision of the atomic coordinates (0.02 Å).

4.4.3 Conclusion

In Chapter 3 and 4, a novel method to study the redox states of FeS clusters in ferredoxins is developed by incorporating an unnatural amino acid *p*-cyano-L-phenylalanine (*p*CNF) into a model protein. The nitrile group in this compound is a sensitive reporter in infrared spectroscopy. A simple FeS protein spinach ferredoxin (FdI) was selected as a model system to test. An overproduction system and purification protocol for FdI were established. This same protocol was also used for expression and purification of different variants generated by site-directed mutagenesis. *p*CNF then could be successfully incorporated into the protein and gave two variants Y3CNF and Y37CNF. Whereas Y37 is close to the FeS cluster (7.4 Å), Y3 is far away (22.6 Å). The Y3CNF protein was intended as a negative control variant but insufficient protein was obtained for IR spectroscopic study. Based on the two different modes of infrared spectroscopy (reflection mode and transmission mode), it is clear to detect the difference in the spectra of nitrile group in Y37CNF between the oxidised and reduced states and the consistent results from two different modes of IR spectroscopy indicates that the shift of ν_{CN} is associated from different redox states rather than a variation in electrode field effect from electrode. Finally, according to the crystal structures and melting temperatures of wild type and variants, it shows that the FeS cluster environment of variants are similar to the wild type, which is consistent with the results from UV-vis spectra and electrochemistry. As such, we conclude that with the model protein spinach ferredoxin, it could be verified that the unnatural amino acid *p*CNF can be used as an IR probe to detect the redox states of the FeS cluster by means of infrared spectroscopy. Next, this novel approach will be extended to the hydrogenases to establish a method which could be used to study the redox state of the electron transfer clusters

and how the prosthetic group and FeS clusters work together during catalysis via infrared spectroscopy.

5

Chapter 5: Characterisation of *E. coli* Ni-Fe Hydrogenase 1 Overexpressed in Native Host

5.1 Introduction and aim of the chapter

Hydrogenases are highly active enzymes that catalyse the oxidation of H_2, or reduction of H^+ to produce H_2 at high turnover frequency and low overpotential. Hydrogenases function under ambient conditions yet work with high thermodynamic efficiency comparable to precious metals such as Pt,[47] despite containing only first-row, earth abundant transition metals such as Ni and Fe. The *E. coli* Ni-Fe 'hydrogenase 1' (Hyd-1) is of great interest as it displays O_2-tolerance, defined as being capable of continuous H_2 oxidation during prolonged O_2 exposure.[170] The physiological role of Hyd-1 is likely to involve operation at the anaerobic to aerobic switch when *E. coli* is transitioning from an anaerobic to aerobic environment.[59] The structure of Hyd-1 and the mechanism of how it works have been discussed in Chapter 1.

[NiFe] hydrogenase consists of a complex array of functional moieties including a protein matrix with a deeply buried catalytic active site in a hydrophobic pocket. The active site is located in the large subunit of [NiFe] hydrogenase and includes Ni and Fe atoms, linked to the protein backbone by four cysteine sulfurs (shown in Figure 5.1). Two of the cysteines connect to the Ni only in a terminal fashion, the other two bridge the bimetallic centre Ni and Fe ions. The low spin Fe(II) with low spin state is coordinated by two small inorganic ligands as well, CO and CN^-, which are both strong-field diatomic ligands and were first detected in hydrogenases by FTIR spectroscopy.[56]

Standard [NiFe] hydrogenases (or so-called O_2 sensitive [NiFe]-hydrogenases) are inactivated under an aerobic environment, while they can be reactivated after gaseous O_2 is replaced with H_2, or by reducing the inactivated active site via protein film electrochemistry.[107] Interestingly, certain [NiFe] hydrogenases are O_2-tolerant, capable of maintaining catalytic activity in the presence of ambient O_2.[58,59]

Since the catalytic site of [NiFe] hydrogenase is deeply buried within the protein structure, the electrons used during catalysis are dependent on an electron relay chain of FeS clusters. The

iron-sulphur cluster relay contains the proximal [4Fe-3S], medial [3Fe-4S] and distal [4Fe-4S] clusters with 12–14 Å spacing between each other.[65] All three are contained within the small subunit HyaA. A prominent difference between O_2-sensitive and O_2-tolerant hydrogenases comes from the cluster proximal to the active site: standard [NiFe] hydrogenases display a [4Fe-4S] cluster while O_2-tolerant hydrogenases possess a special [4Fe-3S] cluster.[171] According to sequence analysis, modelling studies and crystal structures, the proximal [4Fe-3S] cluster is bound by six cysteine residues, where two of the cysteine residues are coordinated to the cluster in addition to the conserved cysteines of other [NiFe] hydrogenases.

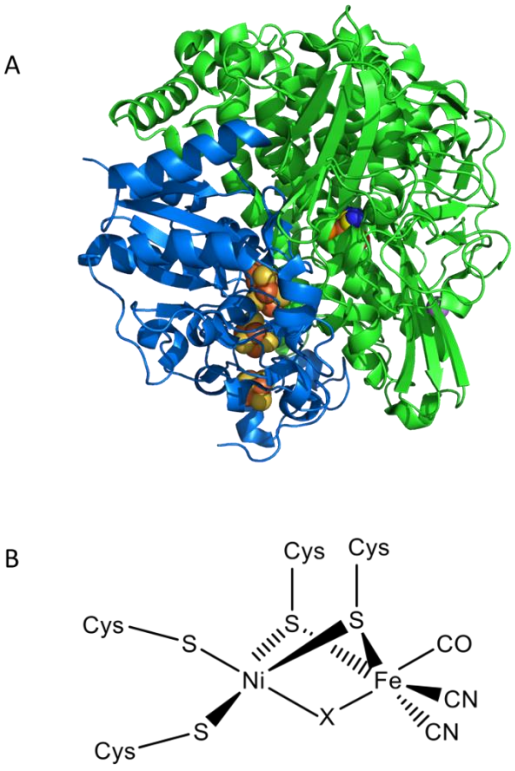


Figure 5.1 Panel A: Structure of *E. coli* [NiFe] hydrogenase 1 (Hyd-1). The large subunit HyaB is in green and the small subunit HyaA is in blue. The FeS clusters and the NiFe active sites are shown as spheres (cartoon representation generated with PyMOL; PDB code: 3USE). Panel B: The representation of the catalytic active site of Hyd-1.

Within the *E. coli* genome, Hyd-1 is encoded by a six-gene operon *hyaABCDEF* which was first reported in 1990 and indicated to encode a NiFe hydrogenase small subunit (HyaA), a NiFe hydrogenase large subunit (HyaB), an integral *b*-type membrane cytochrome (HyaC), and a

system-specific protease (HyaD).[172] Transcription of the operon is controlled by anaerobiosis.[173] According to sequence analysis of the *hya*ABCDEF operon, the core NiFe hydrogenase component would be a heterodimeric complex consisting of small subunit HyaA and large subunit HyaB anchored to the bacterial membrane, as shown following purification of Hyd-1 with detergent to solubilise the bacterial membrane.[172] The large subunit binds to the small subunit which in turn anchors the catalytic core (large + small) to the membrane via the small trans-membrane helix subunit. The 6x His tag (used for purification, mentioned below) is at the C terminus of the trans-membrane helix. Several groups attempted to over-express Hyd-1. kim *et al.* tried to over-express Hyd-1 with genes for these core enzyme subunits (*hya*A and *hya*B) of the complete *hya* operon in *E. coli*. However, the majority of enzyme was produced as inclusion bodies detected in the insoluble cell debris fraction.[174]

Therefore, apart from the six-gene operon *hyaABCDEF* required to express the Hyd-1 complex, additional genes located at the *hypABCDE* operon are also essential for the biosynthesis of the NiFe cofactor. The operon is called '*hyp-*', where 'hy' represents 'hydrogenase' and 'p' represents 'pleiotropic' which means hydrogenase pleiotropic operon. The nonproteinaceous ligand CN^- for the Ni-Fe(CO)$(CN^-)_2$ cofactor is biosynthesised from a carbamoyl phosphate molecule.[175] The initial stage of the carbamoyl phosphate modification involves an important protein HypF.[176] The HypF protein interacts with HypE: there is a special cysteine residue at the C-terminus of HypE and due to the close interaction between HypE and HypF, the product of the HypF reaction, carbamoyl adenylate, is able to react with the HypE special cysteine to form an S-carbamoyl modified HypE. The HypE protein has carbamoyl dehydratase activity and next uses ATP hydrolysis to modify the bound carbamoyl moiety to isothiocyanate.[177] The isothiocyanate-loaded HypE is able to donate CN^- to Fe(II) in the biosynthesis of the NiFe cofactor. The second nonproteinaceous ligand in the active site, CO, is first loaded to the Fe, and it has been suggested the CO may come from CO_2 that can bind to an Fe(II) ion held in place by the HypCD

complex.[177] The HypCD complex holds the Fe-CO moiety and the HypE-bound cyanate is deposited onto the Fe-CO repeatedly until the $Fe(CO)(CN^-)_2$ species is formed.[177]

The significant difference between O_2-tolerant hydrogenases and standard hydrogenases is the unusual [4Fe-3S] cluster that is ligated by six cysteines rather than the four cysteines which ligate the standard in [4Fe-4S] proximal cluster in the O_2-sensitive hydrogenases. As shown in Figure 5.2, the unique proximal cluster is able to undergo two consecutive one-electron transfers to/from the active site at similar potentials, and this is thought to be a significant factor in detoxification of O_2.[66] The redox variation at the Cys_6[4Fe-3S] cluster results from the structural rearrangements which involve a ligand exchange at one of the Fe atoms shown in Figure 5.2.

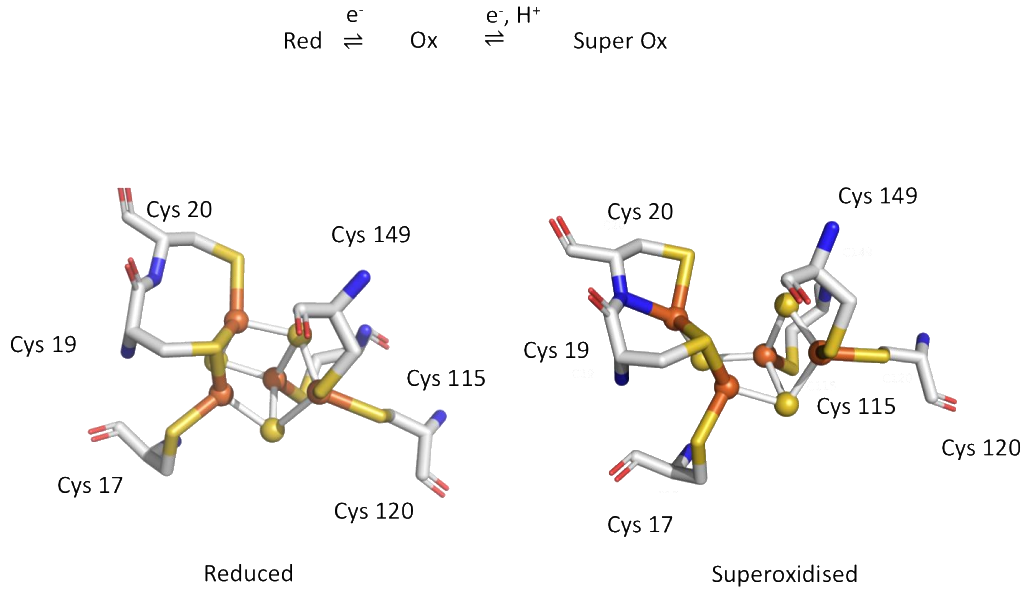


Figure 5.2 Structure of the Cys_6[4Fe-3S] centre of membrane-bound hydrogenase (MBH)-like enzymes in the reduced (A:PDB 3RGW) and superoxidised (B:PDB 4TTT) in O_2-tolerant hydrogenases from *Ralstonia eutropha* MBH *(R. eutropha* MBH*)*.

Goris and co-workers have demonstrated the role of additional Cys ligands in the proximal cluster by mutating two additional Cys residues in *R. eutropha* MBH (Cys 19 and Cys 120) to Gly residues. Strikingly, the *R. eutropha* MBH became an O_2-sensitive enzyme.[63] Electron

paramagnetic resonance (EPR) spectroscopy on this variant showed a significant change in the electronic properties of the proximal FeS cluster centre.[178] Further studies have been carried out using protein film electrochemistry and EPR spectroscopy, which have revealed additional details on the role of the two additional Cys residues. These studies showed that Cys120 plays a less important role in O_2 tolerance than Cys19 which comes from the result that the C120G Hyd-1 mutant (Cys-120 mutated to Gly) is unstable and not able to sustain H_2-dependent growth under an O_2 environment, whilst mutant C19G is still able to mediate growth at low O_2 conditions.[68]

Further studies indicated that the bridging Cys19 is important for destabilising the Fe4-sulfide bond, therefore, during the two one-electron transfers in the proximal [4Fe-3S] cluster, the bond between Fe4 and the sulfide can be broken and a new link between Fe4 and the peptide amino-N ligand of Cys20 is established.[72] Similar conclusions were drawn from Mössbauer spectroscopy on O_2-tolerant Hyd-1 from the hyperthermophile *Aquifex aeolicus*, and X-ray absorption spectroscopic analysis of oxidised *R. eutropha* MBH.[10,179]

Site-Directed Mutagenesis

Site-directed mutagenesis (SDM) is widely used to probe the role of specific amino acids in proteins and a potent method to study the mechanism of [NiFe] hydrogenases. Previous SDM studies has been applied to address the following problems in hydrogenase catalysis: the distinct features of O_2-tolerant hydrogenases which enable them to continue catalysis under aerobic conditions; the mechanism of proton and electron transfer to the active site; and the role of some amino acid residues close to the active site. Therefore, site-directed mutations have been focused on three regions in the [NiFe] hydrogenases: gas channels, FeS cluster relays, and the NiFe active site. For example, Evans and co-workers illustrated that a highly conserved glutamate (residue E28 in *E. coli* Hyd-1 numbering) which is located within 4 Å of the NiFe active site, is significant for proton transfer during catalysis. Substituting glutamate (E) with glutamine

(Q) gives the E28Q variant which has greatly suppressed proton-transfer ability. The results were corroborated by protein film electrochemistry, protein film infrared electrochemistry and X-ray crystallography.[180] By investigating the variant E28Q, it has been suggested that the glutamate 28 acts as the essential "gating" residue for proton transfer during the catalysis as the active site redox chemistry or structure do not change at all. Combining IR spectroscopic studies with protein film electrochemistry on E28Q variant Hyd-1 has shown that E28 is not involved in the initial activation/deprotonation of H_2 but facilitates H^+ exit later in the catalytic cycle to regenerate the initial oxidised active state, assumed to be Ni_a-SI. These results have crucial implications for the control of long-range proton-coupled electron transfer in hydrogenases.[114]

Up to now, very few studies focused on mutations close to the FeS cluster relay to investigate electron transfer through the FeS clusters of the small subunit of [NiFe] hydrogenases. Rousset and co-workers investigated the function of the high potential medial [3Fe-4S] cluster in the small subunit of Class I [NiFe]-hydrogenases, the substitution of proline 238 with cysteine in *D. fructosovorans* [NiFe] hydrogenase led to a [4Fe-4S] cluster being bound in this position. Proline 238 aligns in several sequence alignments with a coordinating cysteine residue which binds the [4Fe-4S] cluster, therefore, this residue is assumed to occupy the position of a potential ligand to the empty fourth Fe site of the [3Fe-4S] cluster.[181] The crystal structure obtained during these experiments indicated that the P238C variant contained three [4Fe-4S] clusters and the midpoint potential of the medial cluster was –0.25 V, whereas the midpoint potential of wild type is +0.065 V higher.[182]

5.2 Aim of the chapter

Even though the biosynthetic route for chromosomally encoded Hyd-1 (CE Hyd-1) has been determined, a big challenge when working with this enzyme is the low yield following native expression, with yields of around 0.05 mg of enzyme per litre of culture.[59] To facilitate

biotechnological applications as well as to further engineer this sophisticated enzyme, it is highly desirable to explore over-expression methods of Hyd-1. Over-expression methods would enable the use of this O_2-tolerant hydrogenase as a catalyst in industry, for example: the 'HydRegen' program (see section 1.2.1 for more detail). As the genes to express Hyd-1 are already known and the over-expression system has been established by Dr Miguel Ramirez (University of Oxford),[91] the aim of this chapter's work was to characterise the recombinant Hyd-1 via conventional solution activity assays, protein film electrochemistry and infrared spectroscopy to analyse its behaviour alongside the native chromosomally encoded Hyd-1 to assess the integrity of the system.

Site-directed mutagenesis is a crucial tool to study the catalytic mechanism of [NiFe] hydrogenases and has been used to identify specific function/roles for certain amino acid residues. Based on the overexpression system of Hyd-1 described in Chapter 5, Some novel site-directed variants of Hyd-1 are designed, selecting amino acid residues close to the FeS clusters as preparation for incorporating an unnatural amino acid.

5.3 Characterisation of over-expressed *E. coli* Hyd-1

5.3.1 Purification results

To increase production yields of the *E. coli* hydrogenase-1 (Hyd-1), an over-expression system was developed using two plasmids containing the essential genes from the *hyaABCDEF* and *hypABCDEF* operons. After Hyd-1 over-expression in anaerobic bacterial growth, the protein was isolated following published protocols.[49]

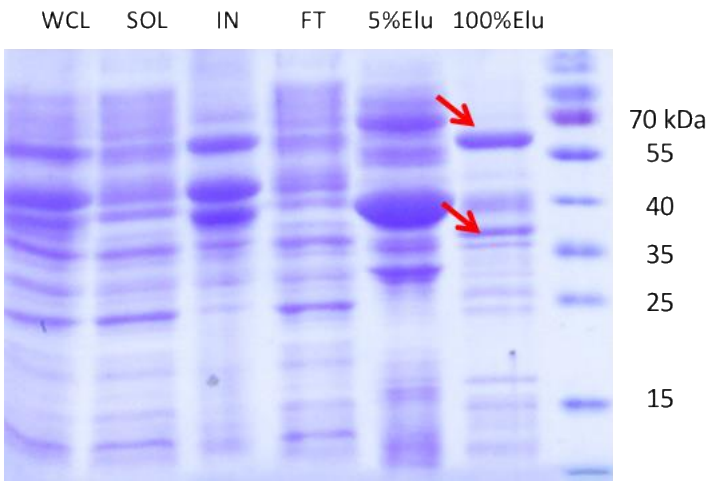


Figure 5.3 SDS-PAGE analysis of Ni^{2+}-affinity chromatography of over-expressed Hyd-1. WCL: whole cell lysate, SOL: soluble lysate. IN: Insoluble. FT: the flowing through from the His-Trap column. W: washing buffer through the column. 5%Elu: the collected elution at 5% of Buffer B (75 mM imidazole). 100%Elu: the collected elution at 100% of Buffer B (750mM of imidazole). The two arrows shown in the '100%Elu' illustrated the two subunits of Hyd-1, the large subunit HyaB (approx. 66 kDa) and small subunit HyaA (approx. 36 kDa).

After over-expression, the Hyd-1 was purified by Ni^{2+}-affinity chromatography. The majority of non-specifically bound protein was eluted at a low concentration (75 mM) of imidazole ("5%Elu" lane in Figure 5.3). Increasing the concentration of imidazole to 750 mM resulted in a small eluting peak corresponding to the specifically-bound proteins, as judged by SDS-PAGE (Figure 5.3). In the "100%Elu" lane, the two main bands are around 65 kDa and 37 kDa, corresponding to the large subunit HyaB (66.5 kDa) and small subunit HyaA (36.5 kDa), indicating that Hyd-1 was over-expressed successfully and isolated and purified to a reasonable degree by Ni-affinity chromatography.

Although Hyd-1 was expressed successfully, there were still many impurities (additional bands in the "100%Elu" lane); therefore, an additional purification step of size-exclusion chromatography was carried out. The fractions containing the target protein were pooled and concentrated to 1 mL volume using an Ultra Centrifugal Filter Unit (50 kDa cut-off molecular weight, Amicon™). A Superdex 200 Increase 10/300 GL column (GE Healthcare) was used with degassed solutions. The column was first equilibrated with buffer of composition 50 mM Tris, 350 mM NaCl, 1 mM

DTT, at pH 7.2. Concentrated protein following Ni-affinity purification as outlined above, was injected onto the equilibrated column using the AKTA protein purification system. The column which was continuously washed with the same buffer and fractions were collected according to the peaks shown in the trace at 280 nm and 420 nm, for probing protein content via tryptophan and tyrosine, and Fe-S clusters and b-type haems respectively.[183]

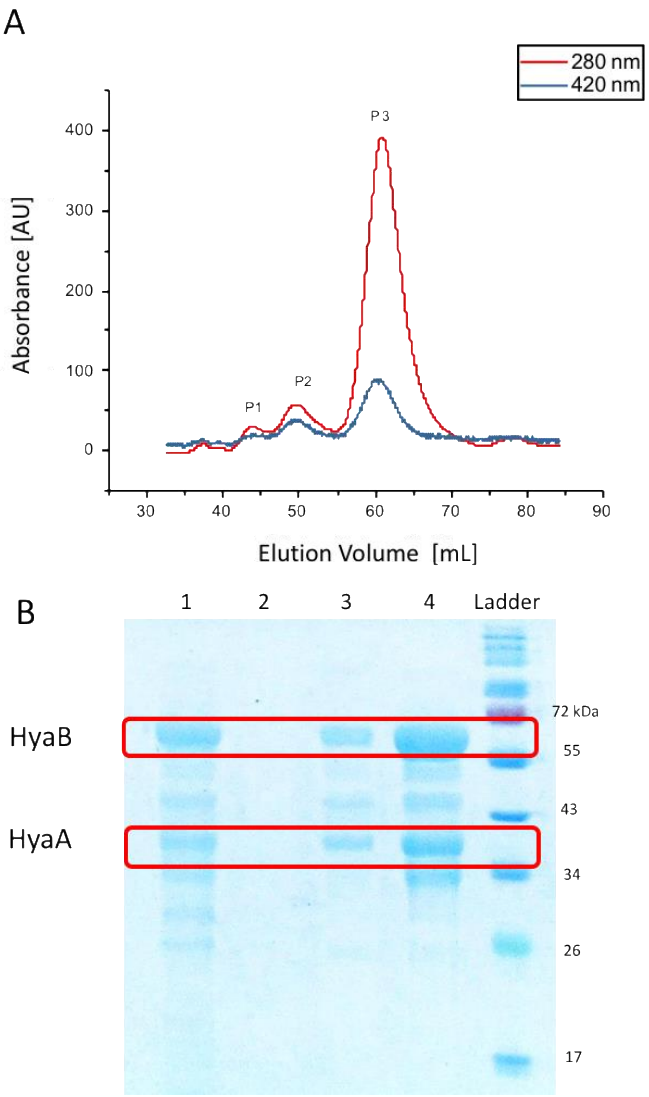


Figure 5.4 Panel A: The AKTA elution traces from the HiLoad Superdex 200 column. Panel B: The SDS-PAGE after size exclusion chromatography on overexpressed NiFe Hydrogenase 1. Lane 1: Hyd-1 purified by Ni^{2+} column. Lane 2: peak 1 (P1) from AKTA elution trace of 280 nm absorbance. Lane 3: peak 2 (P2) from AKTA eluting trace of 280 nm absorbance. Lane 4: peak 3 (P3) from AKTA elution trace of 280 nm absorbance. The two boxes illustrated the two subunits of Hyd-1, big subunit HyaB (approx. 66 kDa) and small subunit HyaA (approx. 36 kDa).

From the eluting trace from the HiLoad Superdex 200 column in Figure 5.4 it is clear that the protein was separated into three fractions as judged by the absorbance at 280 nm (red line), only two of which had a corresponding absorbance at 420 nm (blue line) which indicates the presence of oxidised Fe-S clusters. More information and details can be found in the SDS-PAGE gel in Figure 5.5B: in lane 3 and lane 4 which are associated with the peak 2 and peak 3 in the eluting trace respectively, the large subunit HyaB and small subunit HyaA can be seen at approximately 66 kDa and 36 kDa respectively. However, there are some other bands in the peak 2 and peak 3 sample and it is hard to predict what these impurities are at this stage, as these could also be product of protein degradation or electroporation artifacts. Nevertheless, the level of purity achieved after size exclusion was satisfactory for subsequent characterisation steps.

In the next step, both peak 2 Hyd-1 (called P2) and peak 3 Hyd-1 (called P3) were characterised and compared with the chromosomally encoded (CE) Hyd-1 (see later, Section 5.3). (All the data of CE Hyd-1 adapted from Brooke *et al*, Biochemistry, 56, 2017, 132-142.)[184]

The concentration of overexpressed Hyd-1 was measured via the Bradford Assay. As shown in Table 5.1, the yield of over-expressed Hyd-1 is about 10 times higher than chromosomally encoded Hyd-1.

Table 5.1 The yield comparison between chromosomally encoded (CE) and overexpressed (OE) Hyd-1.

	Yield (mg/L culture)
Chromosomally encoded Hyd-1 fused to a 6x histidine tag (CE Hyd-1)	0.05 mg/L
Overexpressed Hyd-1 fused to a 6x histidine tag (OE Hyd-1)	0.62 mg/L

Following the successful overexpression of Hyd-1, Traditional solution-based activity assays, protein film electrochemistry and infrared spectroscopic methods were carried out to establish whether the OE Hyd-1 behaves the same as the CE Hyd-1 and to compare the kinetic, thermodynamic and structural properties of the enzymes.

5.3.2 Solution assay

Simple activity test with benzyl viologen

As an initial crude assay to test for H_2 oxidation activity of Hyd-1 benzyl viologen was used as the terminal electron acceptor and provides a colorimetric response detectable by eye. Oxidised benzyl viologen is colourless while reduced benzyl viologen is purple. The purified Hyd-1 oxidises dihydrogen to protons and electrons, transferring the electrons through its Fe-S cluster chain to reduce benzyl viologen in solution.

To quickly determine whether the purified overexpressed Hyd-1 was capable of catalysing the oxidation of H_2, a benzyl viologen assay was carried out as follows: benzyl viologen was added to 1 mL of Tris buffer with final concentration 5 mM (50 mM Tris-HCl, pH7.5) and then sparged with H_2 for 30 min to ensure the solution was fully saturated with H_2. Hyd-1 was added to the assay solution and compared with a negative control, the assay solution without Hyd-1. The two solutions were continuously bubbled with H_2. The mixture containing Hyd-1 turned purple, while the control reaction remained colourless. This indicated that the overexpressed Hyd-1 can catalyse the oxidation of H_2 and reduction of benzyl viologen.

Quantitative assay with methylene blue

On confirmation that the Hyd-1 from the overexpression system was active for H_2 oxidation, the steady state rate of H_2 oxidation for overexpressed Hyd-1 at pH 6 and 25 °C was measured by conventional solution assay in an H_2-saturated solution with methylene blue as an electron acceptor. The extinction coefficient for methylene blue is 22400 cm^{-1} M^{-1}. For this experiment,

enzyme concentration was varied between 0.1 – 5 mg mL^{-1}. The simple colorimetric test outlined above, used benzyl viologen as the soluble electron acceptor. For quantitative analysis of Hyd-1 activity, methylene blue (–0.45 V *vs*. SHE) is preferred. This compound has a more negative reduction potential than that of the benzyl viologen (–0.38 V *vs*. SHE).[185,113] Initial velocity measurements were utilised to estimate both the turnover frequency k_{cat} (s^{-1}) and specific activity in µmol H_2 oxidised $min^{-1}mg^{-1}$ enzyme (see section 2.3 for more detail). Table 5.2 shows the kinetic data for both chromosomally encoded and overexpressed Hyd-1 (both P2 and P3, as shown on the AKTA trace and SDS-PAGE gel in Figure 5.4). The rate of overexpressed Hyd-1 elution peak 2 (k_{cat} = 126.5 s^{-1}) and peak 3 (k_{cat} = 240.4 s^{-1}) are within experimental error of the rate of chromosomally encoded (CE) Hyd-1.

Table 5.2 Kinetic data for chromosomally encoded (CE) and overexpressed (OE) Hyd-1. The turnover frequency is given as k_{cat} (s^{-1}) and specific activity (µmol H_2 oxidised min^{-1} mg^{-1} enzyme). The data represent averages from at least three repeats. The CE Hyd-1 data were produced by Emily Brooke and Dr Rhiannon Evans (University of Oxford).[184]

	CE WT Hyd-1	OE WT Hyd-1 P2	OE WT Hyd-1 P3
solution assay [µmol of H_2 min^{-1} (mg of enzyme)$^{-1}$]	153 ± 18	79 ± 6	143 ± 9
solution assay turnover frequency [(s^{-1})]	257 ± 30	126 ± 12	240 ± 14

5.3.3 Michaelis constant and activation enthalpies.

Michaelis constant

Michaelis-Menten kinetics is one of the best known models of enzyme kinetics; it describes the rate of enzyme reactions by relating reaction rate (rate of formation of product [P] to the concentration of a substrate, S. The Michaelis constant K_M represents the concentration of substrate at half the maximum rate is observed of binding, unbinding and k_{cat}. Protein film electrochemistry (PFE) (see section 2.5 for more detail of electrochemistry and PFE), cyclic

voltammograms (CVs) and CV were used to determine the Michaelis constant for H_2 $K_M^{H_2}$. The current is equivalent to the H_2 oxidation activity in this experiment which is an average of forward and back scans with the scanned range between –0.65 V and +0.24 V (*vs*. SHE) at a scan rate of 5 mV s^{-1}. The H_2 gas concentration was varied in the range of 3–100% while Ar was used as a carrier gas, as shown in Figure 5.5A. Precise gas mixtures were created using mass flow controllers (Sierra Instruments). The working electrode was rotated at a speed of 1000 rpm to ensure mass transport of substrate and product to and from the enzyme film was not rate limiting. The experiments were conducted at pH 6 and 30 °C in mixed buffer (the components of mixed buffer were shown in section 2.5.8, Table 2.14). Prior to recording each CV, equilibration at each H_2 concentration was monitored using chronoamperometry at 0 V, a potential where H_2 oxidation current is visible at all concentrations of H_2 and there is no negligible anaerobic inactivation of Hyd-1.[184] The concentration of H_2 was returned to 100% H_2 periodically to enable monitoring of the general deterioration of the enzyme film (film loss) during the experiments. Since current is proportional to activity of the enzyme film, a Hanes-Woolf analysis of P2 and P3 (Figure 5.4) was then utilised to determine K_M values shown in Figure 5.5B and C. (see section 2.9.2 for more detail).

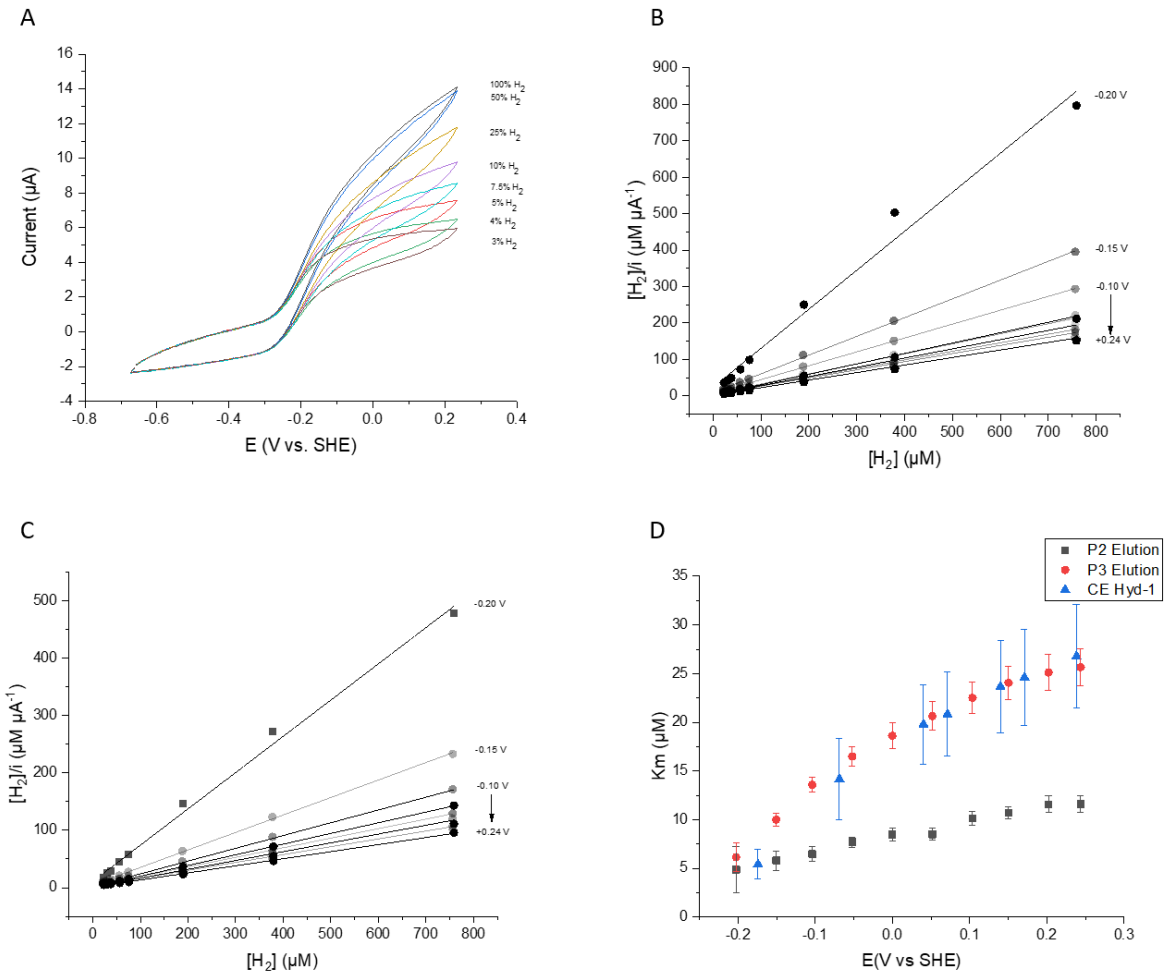


Figure 5.5 Panel A: One example of cyclic voltammograms (P3 OE Hyd-1) in different H_2/Ar mixes. These were used to generate the Hanes-Woolf plot. Panel B and C: The Hanes-Woolf analysis for the P2 and P3 OE Hyd-1 respectively, the data were used in a H_2/i vs. [H_2] plot at different potentials. A linear regression was used to determine V_{max} and K_M. Panel D: Potential dependence of $K_M^{H_2}$ for both overexpressed (OE) Hyd-1 P2 and P3 elution comparing with chromosomally encoded (CE) Hyd-1. Cyclic voltammograms (CVs) were measured at a scan rate of 5 mV s^{-1} in the potential range -0.65 V to + 0.24 V with varying H_2 from 3-100% Ar as a carrier gas. To measure the film loss during the experiments, the headspace atmosphere was returned to 100% H_2 periodically. $K_M^{H_2}$ was determined by using a Hanes-Woolf analysis. The condition to measure $K_M^{H_2}$ was 30 °C, mixed buffer pH 6, total H_2 gas flow = 1000 scc min^{-1} (standard cubic centimetres per minute), ω = 4000 rpm. Error is the standard error of the mean calculated from three separate experiments. The CE Hyd-1 data is from Brooke *et al.*[184]

It is clear to see that in Figure 5.5D the minor increase with potential of P2 elution Hyd-1 while there is a significant increase with potential in P3 elution, with 5-fold from –0.2 V to +0.25 V, but the trend of both P2 and P3 elution are very similar. $K_M^{H_2}$ is potential dependent which can be

explained by the equation $K_M^{H_2} = \frac{(k_{cat}+k_{off})}{k_{on}}$, where k_{on} and k_{off} refer to H_2 binding and dissociation, the value of k_{cat} increases with potential, as it is the net turnover frequency based on all elementary steps, which includes interfacial electron transfer, therefore, $K_M^{H_2}$ should increase with potential as well.[113] The $K_M^{H_2}$ of P2 and P3 at 0 V are 7.1 µM and 18.6 µM respectively. While the $K_M^{H_2}$ of chromosomally encoded (CE) Hyd-1 at 0 V is 19.8 µM which is very close to that of the P3 elution.[184] It indicates that both P2 and P3 are in the same order of magnitude with that of the CE Hyd-1 and therefore not significantly different from each other.

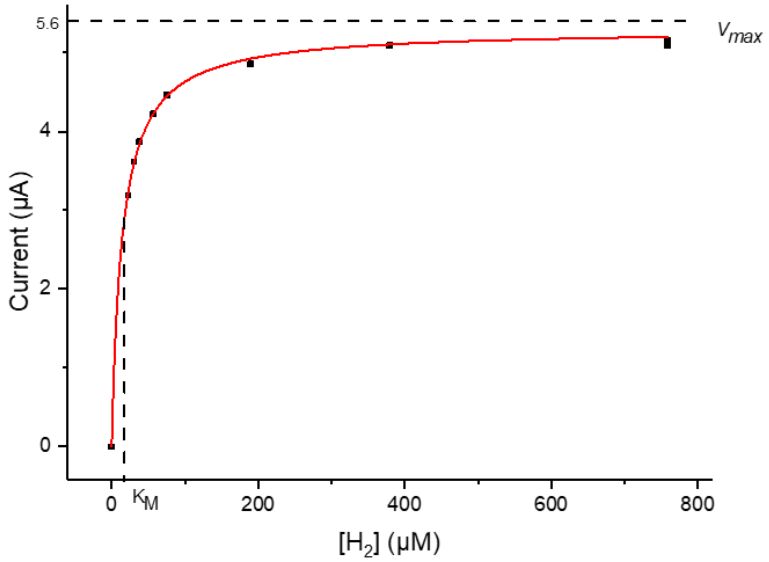


Figure 5.6 The Michaelis-Menten plot for the P3 OE Hyd-1. The data were used in a current vs. [H_2] plot at –0.15 V potential.

The plot of the catalytic current against the substrate concentration [H_2] at –0.15 V potential for P3-OE Hyd-1 is shown in Figure 5.6. The data were fitted to the Michaelis-Menten equation in the form of a rectangular hyperbola by OriginPro. $K_M^{H_2}$ was determined to be 15.5 µM (R^2 = 0.9967).

Activation enthalpies

The enthalpy of activation, $\Delta H^{\ddagger}$, for H_2 oxidation was determined for peaks 2 and 3 of over-expressed Hyd-1 (P2 and P3) using the temperature-dependence of the H_2 oxidation current during CV scans. The CV scans were measured between –0.65 V and +0.24 V at a scan rate of 0.005 V s^{-1}. The temperature was altered between 1 and 37 °C, and accurately measured by an internal thermometer placed inside the working compartment of the electrochemical cell (see section 2.5.7 for more detail). Equilibration of the buffer upon a change in temperature was monitored by recording CVs at a scan rate 0.03 V s^{-1} until the current of each CV overlapped. Stable CVs at 30 °C were recorded periodically to measure film loss during the experiment and allowed for film loss correction. The working electrode was rotated at 3000 rpm to ensure mass transport of substrate and product was not rate limiting. The Eyring equation is derived from transition state theory that describes reaction kinetics, it can be used to quantify the temperature dependence of the activity of overexpressed Hyd-1 in terms of enthalpy $\Delta H^{\ddagger}$. Using the data from cyclic voltammograms (one example for P3 is shown in Figure 5.7A), an Eyring plot of the ln(i/T) against 1/T (Figure 5.7B and C) allowed the determination of $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$.

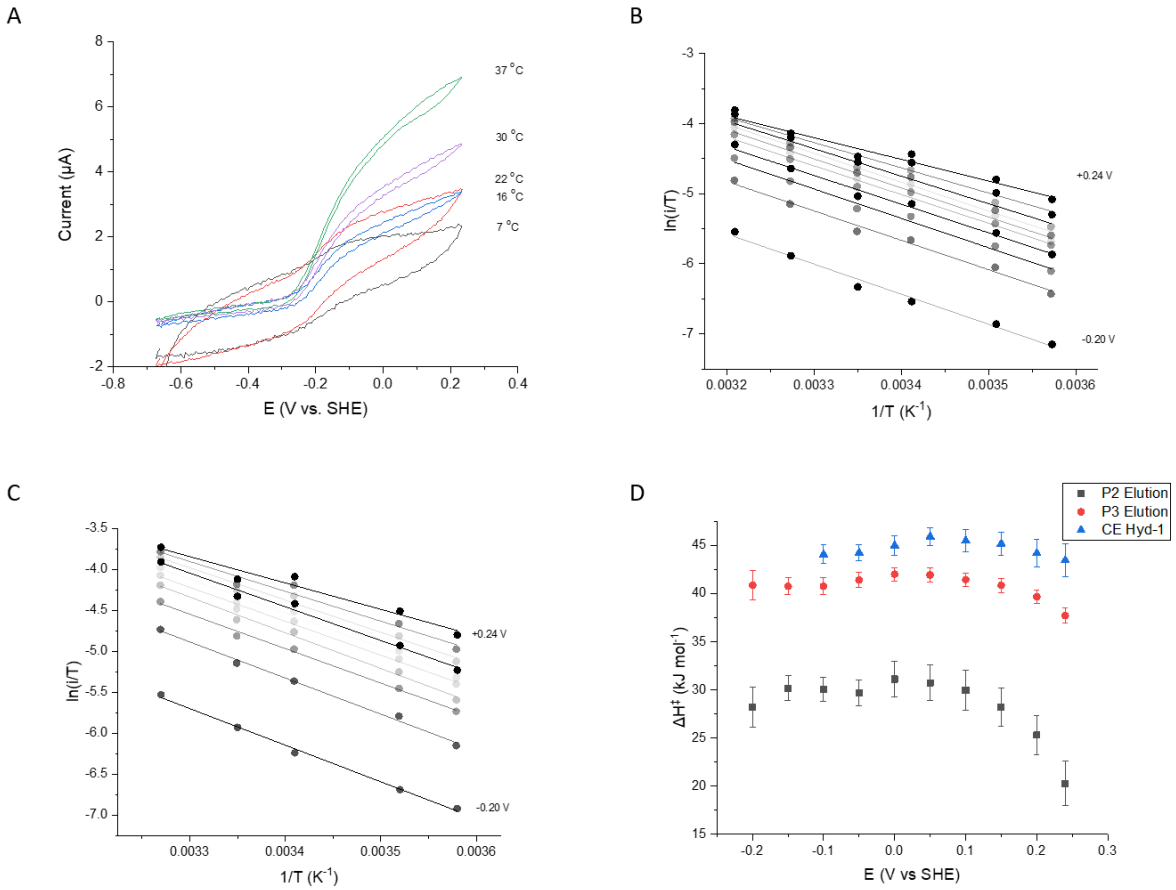


Figure 5.7 Panel A: One example of cyclic voltammograms (P3 OE Hyd-1) at different temperatures. Panel B and C: The Eyring plot for the P2 and P3 respectively. The data were used to produce a plot of $\ln {}^{i}/_{T}$ vs. $1/T$ at different potentials. A linear regression was used to determine $\Delta H^{\ddagger}$ and ${}^{\Delta S^{\ddagger}}/_{R} + \ln({}^{k_B}/_{h})$. Panel D: Potential dependence of $\Delta H^{\ddagger}$ for both overexpressed Hyd-1 P2 and P3 elution. Cyclic voltammograms in the potential range –0.65 V to +0.24 V were measured at a scan rate of 5 mV s^{-1}, whilst varying the temperature between 7°C and 37°C. Equilibration at each temperature was monitored by recording CVs at 30°C periodically so film loss could be measured. $\Delta H^{\ddagger}$ values were calculated by Eyring analysis. Other conditions: mixed buffer pH 6, total gas flow = 1000 scc min^{-1}, ω = 1000 rpm. Error is the standard error of the mean calculated from three separate experiments in which the potential steps were repetitions in the same experiment. (The CE Hyd-1 data is adapted from Brooke *et al,* 2017)[184]

Figure 5.7D shows the potential dependence of the activation enthalpy $\Delta H^{\ddagger}$ for over-expressed Hyd-1 P2 and P3 elutions and the chromosomally encoded Hyd-1. The activation enthalpy of P3-OE Hyd-1 appears largely potential independent but with a slight decrease at higher potential. The decrease at higher potential was more marked for the P2 elution. The activation enthalpy of

P2 and P3 Hyd-1 is 31.1 kJ mol^{-1} and 42.1 kJ mol^{-1} respectively; the value for P3 was more similar to that of chromosomally encoded Hyd-1 (44.7 kJ mol^{-1})[184] at 0 V.

According to the Michaelis constant, activation enthalpies and steady-state solution assay rates of both P2 and P3 (Figure 5.4) over-expressed Hyd-1 elutions, the properties of both P2 and P3 elution are consistent with chromosomally encoded Hyd-1. However, as the yield of P3 elution is much higher than P2 elution, further characterisation and experiments only focused on the P3 elution.

5.3.4 Electrocatalytic Window and E_{Switch}

At high potential, [NiFe]-hydrogenases are anaerobically inactivated as the Ni-B state is formed (see section 1.3.2 for more detail).[115] Ni-B (the 'ready' state) is the most oxidised form of the [NiFe] metal centre and able to be activated via one-electron reduction with proton transfer. The formation of Ni-B under anaerobic conditions is slow and the rate is potential independent because the rate determining step is a chemical step (ligand binding of OH_2) whereas electrochemical oxidation of Ni^{II} to Ni^{III} is fast. The reverse is true for re-activation of Ni-B, where first Ni^{III} is reduced to Ni^{II} (rate determining step) and then OH^- is lost in a fast chemical step. Thus, at a given potential both the formation of Ni-B and reactivation of Ni-B are competing processes and will be happening at the same time, and so it is the relative difference in the rate of formation of Ni-B and the rate of reactivation of Ni-B which determines the extent of the overall process.

Hyd-1 is not able to oxidise H_2 in the Ni-B state and the H_2 oxidation current shows a significant fall at high potential as Ni-B forms. Therefore, a 'window' for H_2 oxidation is observed which is defined by the potential window between the onset potential for H_2 oxidation, and the 'E_{Switch}' potential. E_{Switch} has been defined as the potential at which the rate of change of current is greatest during reductive activation of inactivated states, found as the inflection point of the first derivative of the reductive scan.[35,113] E_{Switch} is sometimes used as a proxy for the ease of

activation of the Ni-B state, but it is not an equilibrium potential reflecting activation of Ni-B to Ni_a-SI since the reduced enzyme cycles through multiple states rapidly. The empirical value of E_{Switch} is dependent on the scan rate of the CV used in its determination.[186] This is due to the fact that Ni-B must first be formed in order for its re-activation to be reflected in E_{Switch}. The slower the CV scan rate the greater the proportion of Ni-B formed on the scan to high potential. In determining E_{Switch} therefore, it is important to ensure that most the enzyme is in the Ni-B state first, before scanning to lower potentials slowly to monitor the reductive re-activation process. It is also important to specify the scan rate at which E_{Switch} is determined.

To determine a potential of E_{Switch}, the enzyme films were first completely inactivated by applying a potential of +0.39 V in pH 9 and 100% Ar environment for 10,000 seconds to inactivate the enzyme. Hyd-1 is inactivated to a greater extent at low H_2 and high pH. The cell solution and headspace were then returned to pH 6 and 100% H_2 respectively, all the time maintaining electrochemical connection (see section 2.5.8 for more detail). Then the potential was scanned at a very slow rate of 0.1 mV s^{-1} from +0.39 V to –0.6 V.

The electrocatalytic window for H_2 oxidation by overexpressed Hyd-1 P3 elution is shown in Figure 5.8A which shows the chronoamperometry to detect the whole process of inactivation, buffer exchange and equilibration. The electrocatalytic current at pH 9 in the Ar environment decreases over time as the H_2 is flushed from the system and the concentration of H_2 dips below K_M for the enzyme. After 10,000 seconds inactivation, the buffer was exchanged to pH 6 and the H_2 atmosphere returned, a brief period was given to allow temperature and H_2 equilibration of the exchanged buffer (approx. 600 seconds) during which time the electrocatalytic current increased which shows the reactivation begins, even at +0.39 V. The slow CV scan was begun and is shown in Figure 5.8B. E_{Switch} is achieved from the potential at which the minimum in the first derivative of the voltammogram is detected, as shown in Figure 5.8C. The E_{Switch} obtained in the overexpressed Hyd-1 was +0.23 V, very similar to the E_{Switch} of the CE Hyd-1 of +0.24 V.[184]

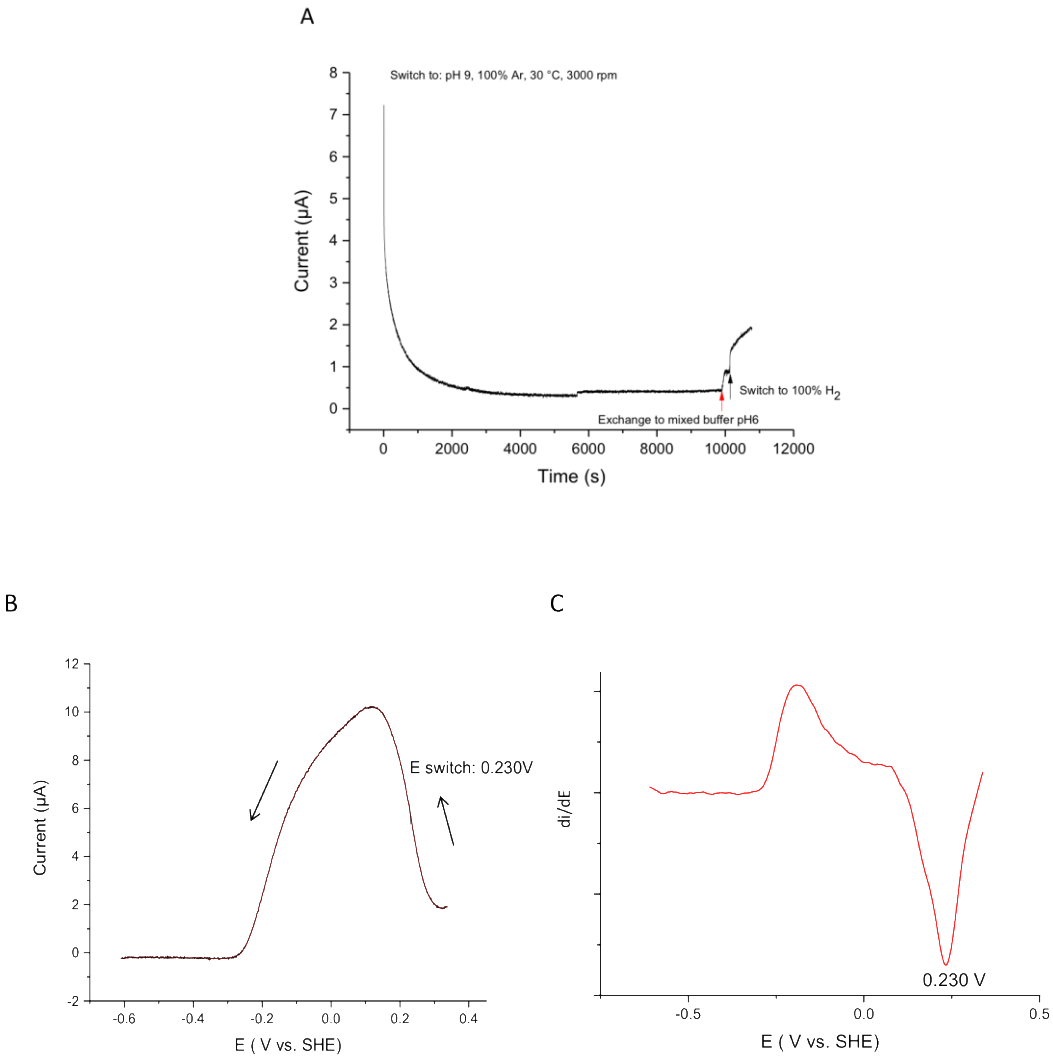


Figure 5.8 Determination of the ease of reactivation of anaerobically-generated inactive states in Hyd-1. Panel A: The chronoamperometry trace recorded during the process of inactivation of the enzyme. Before the start of this trace, the enzyme film was first activated at low potential, pH 6, 30 °C, 100% H_2 to activate as-isolated states (not shown). The potential was then poised at +0.39 V at the start of the trace shown, and the gas switched to 100% Ar and the buffer exchanged to pH 9. The H_2 oxidation current begins to fall as H_2 is flushed from the system and the enzyme begins to inactivate. This process was continued for 10,000 seconds, after which time the buffer was exchanged to the mixed buffer pH 6 with H_2 to reactivate (red arrow) and H_2 was re-introduced (black arrow). Panel B shows the reductive reactivation electrocatalytic profile for H_2 oxidation by overexpressed Hyd-1 following the inactivation procedure described in Panel A. The potential was then scanned from +0.39 V to -0.60 V at a scan rate of 0.1 mV s^{-1}. Panel C is the first derivative of the current scan, on which E_{Switch} is the minimum.

5.3.5 Oxygen tolerance

There is a great interest in Hyd-1 due to its ability to oxidise H_2 in the presence of O_2. These O_2-tolerant NiFe hydrogenases have an unusual [4Fe-3S] cluster near the active site as part of a redox chain that mediates intramolecular electron transfer. The general consensus is that this cluster is responsible for resistance to O_2.[187] It is therefore essential to test if the overexpressed Hyd-1 is still capable of continuous H_2 oxidation in an O_2 environment. A film was prepared of the overexpressed Hyd-1 P3 (Figure 5.4) and was fully activated. The potential was scanned from –0.40 V to +0.24 V and back, at a scan rate of 0.5 mV s^{-1} (Figure 5.9). The first CV was scanned to record a control 'before' scan prior to introduction of O_2. On the following scan, an aliquot of O_2-saturated buffer was injected at +0.03 V, a potential well below E_{Switch} and where anaerobic inactivation is therefore minimal. This potential also avoids O_2 reduction at the electrode.[188] The O_2-saturated buffer injection results in an initial O_2 concentration in the electrochemical cell of 154 µM (see section 2.9.1 for more detail about Henry's law). The constant H_2 flow into the cell headspace, rapid rotation and slow scan rate ensured all O_2 is flushed from the system by the time the reverse scan begins.

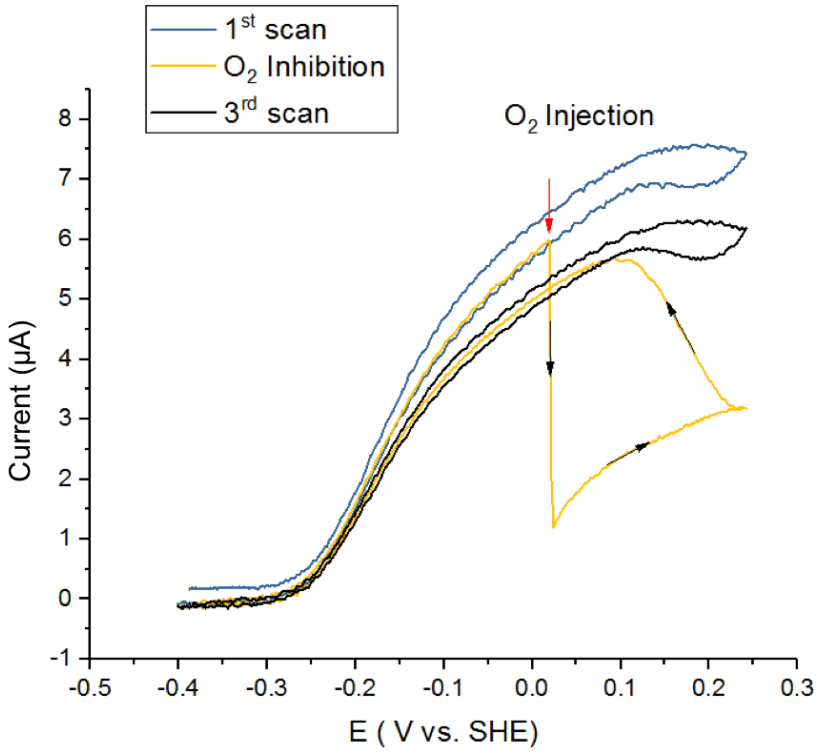


Figure 5.9 Cyclic Voltammograms (CV) to test the O_2 tolerance of overexpressed Hyd-1. Response of overexpressed Hyd-1 to a transient burst of O_2 during H_2 oxidation. The blue curve is the 1st scan before O_2-saturated buffer was injected; the yellow curve is the response of the Hyd-1 when O_2 is introduced at +0.03 V (red arrow), black arrows indicate the direction of the current response immediately following O_2 exposure; the black curve is the 3rd scan following O_2 exposure. All experiments were performed as follows: CVs were scanned from –0.40 V to + 0.24 V at a scan rate of 0.5 mV s^{-1}; at the specified potential, +0.03 V, an aliquot of O_2-saturated buffer was injected so that the initial concentration of O_2 in the electrochemical cell was 154 µM. Other conditions: mixed buffer pH 6, 100% H_2, total gas flow = 1000 scc min^{-1}, ω = 3000 rpm.

Figure 5.9 shows the results of a test for tolerance to transient exposure to O_2 for the overexpressed Hyd-1. After injecting the O_2-saturated buffer at +0.03 V, the current drops immediately but there is still some current remaining (*ca*. 75%). Furthermore, after the initial drop in current, activity begins to recover before the scan direction is reversed, at which point there is approximately 75% of the current compared to that of the first scan at the same potential (+0.24 V). The current recovers when the O_2 is gradually replaced by the flow of H_2 and enzyme reactivation. In the third scan, the normal H_2 oxidation current is restored (not withstanding film loss) which illustrates that the attack of O_2 did not disturb the Hyd-1 on the

timescale of the exposure and the enzyme is still able to catalyse H_2 oxidation and it behaves the same as chromosomally encoded Hyd-1.

5.4 Infrared spectroscopy of overexpressed Ni-Fe Hydrogenase.

After characterising overexpressed Hyd-1 via conventional steady-state solution activity assays and protein film electrochemistry, Infrared spectroscopy was carried out to reveal the potential-dependent structural (ligand) details of the enzyme active site. Infrared spectroscopy is ideal for probing hydrogenase active site details due to the presence of the strong IR absorbers, CO and CN^-, in the active site (Figure 5.1B).[83] The intrinsic CO and CN^- ligands give rise to vibrational bands in the mid-IR, the frequency of which is highly sensitive to changes in electron density at the active site.[189]

The transmission mode FTIR was carried out on both as-isolated and H_2-reduced Hyd-1. The overexpressed Hyd-1 was concentrated to 1 mM and buffer exchanged into Tris-HCl, pH 8. To reduce Hyd-1, the enzyme was subjected to 2 bar H_2 gas for 4 days 25 °C in a pressure vessel (Tinyclave steel, Buchi AG). Aliquots (2 μL) of as-isolated and H_2-reduced enzyme were individually loaded into the transmission cell and the Infrared spectra recorded. All the manipulations and infrared spectral collection were carried out in an anaerobic glovebox (Glove Box Technology Ltd) with O_2 levels <2 ppm.

Based on the structure of the catalytic active site (Figure 5.1B), there are two CN^- and one CO ligands to the Fe. There are the same corresponding number of peaks in the ν_{CN} and ν_{CO} region of the spectra for each NiFe active site state present. Sometimes more than one state can have the same/overlapping stretch in the ν_{CN} region so unequivocally assigning the ν_{CN} stretches to an individual state is difficult, hence we tend to focus on the ν_{CO} region. Figure 5.10 shows the transmission mode FTIR spectra of both as-isolated and H_2-reduced overexpressed Hyd-1 zoomed in between 2150-1850 cm^{-1}.

The as-isolated overexpressed Hyd-1 shows a large peak at 1943 cm^{-1}, which is associated with the inactive Ni-B (Ready) or Ni-A (Unready) state (see section 1.3.2 for more detail); however, Ni-B and Ni-A state cannot be distinguished via IR spectroscopy; EPR spectroscopy has been used to test the difference between Ni-A and Ni-B in chromosomally encoded Hyd-1 and confirmed that Hyd-1 forms Ni-B in the as-isolated form.[59] There is another small peak at 1923 cm^{-1} in the as-isolated state, which does not seem to disappear with reduction, and may be from inactivated protein. In the reduced state of overexpressed Hyd-1 shown in Figure 5.10, there are three new peaks at 1950 cm^{-1}, 1914 cm^{-1} and 1877 cm^{-1} which correspond to the Ni-C, Ni-R_{III} and Ni-L_{II} respectively. Comparing the spectra of the reduced and as-isolated enzymes, there is an increase in the relative proportions of the Ni-R_{II} and Ni-R_{III} states at 1922 cm^{-1} and 1914 cm^{-1} respectively, while the Ni-B state is no longer present. The catalytic cycle and the structure of each state in the active site are shown in Figure 1.9.[109]

The experiments above demonstrate that the active site of overexpressed Hyd-1 functions as expected with regards to the infrared-detected states formed and in the as-isolated and H_2-reduced forms of the enzyme with the exception of the 1923 cm^{-1} species.[19] However, it was not possible to detect overexpressed Hyd-1 when adsorbed on a carbon particle electrode for protein film infrared electrochemistry (see section 1.5.2 for more detail).

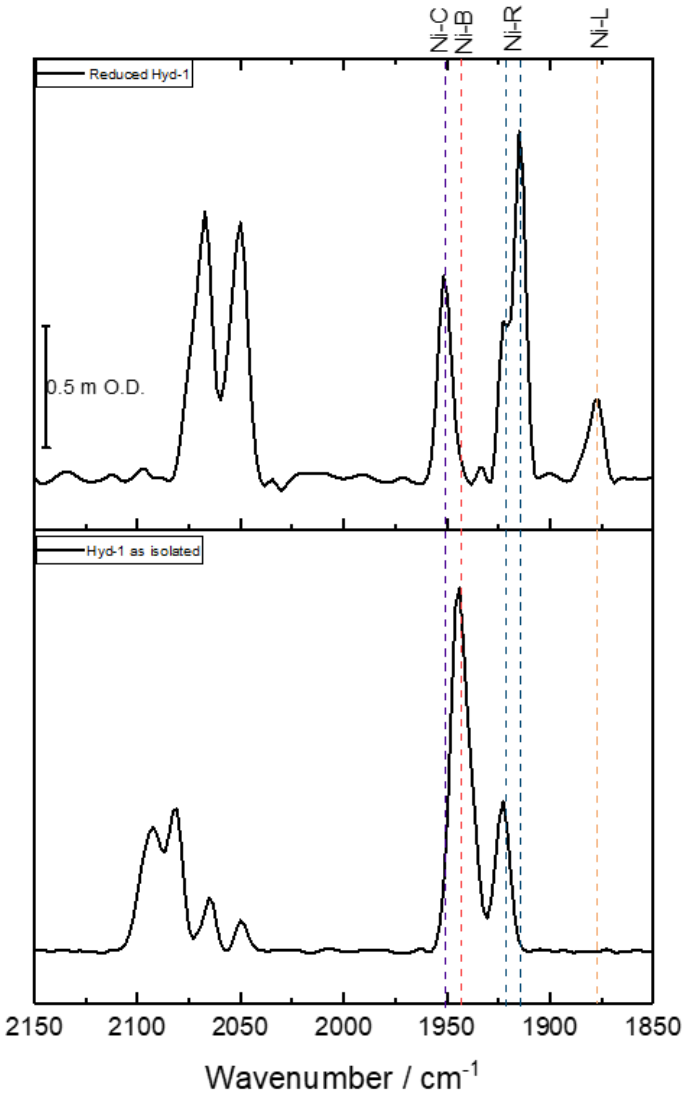


Figure 5.10 Base-line corrected transmission mode infrared spectra of overexpressed as-isolated and H_2-reduced Hyd-1 in the ν_{CO} and ν_{CN} region. Dashed lines show the wavenumber positions of the intrinsic bands of the Ni-B (red), Ni-R (blue), Ni-L (orange) and Ni-C (purple) states of overexpressed Hyd-1.[116]

The spectra resulting from overexpressed Hyd-1 were then compared with the published spectra of chromosomally encoded Hyd-1 (peak assignment for ν_{CO} shown in Table 5.3).[83] The overexpressed Hyd-1 shows the same IR behaviour as chromosomally encoded Hyd-1.

Table 5.3 Infrared peak positions (in cm^{-1}) of the intrinsic active site ν_{CO} band for chromosomally encoded Hyd-1 in different redox or coordination states.[83]

	Ni-A	Ni-B	Ni_a-SI	Ni-C	Ni-R_{II}	Ni-R_{III}	Ni-L_I	Ni-L_{II}	Ni-L_{III}
E. coli Hyd-1	-	1943	1927	1949	1922	1914	1898	1877	1867

5.5 Site-directed mutagenesis of Hyd-1 ready for unnatural amino acid incorporation.

The next step was to select suitable amino acid residues close to the FeS clusters for incorporation of the *p*CNF IR probe. To prevent mutagenesis from having unintended effects on the structural/catalytic functions of the FeS clusters, key cysteine residues for coordination of the clusters or residues within 5 Å were excluded. Figure 5.11 shows five amino acid residues as potential mutation site candidates. These residues were chosen with a distance less than 10 Å to either the [3Fe-4S] proximal cluster or the [4Fe-4S] medial cluster, but all of them are far away from the NiFe active site (more than 15 Å). The selected amino acid residues were first mutated to phenylalanine (Phe, F) to test the tolerance of the enzyme to residues with aromatic side chains at these positions.

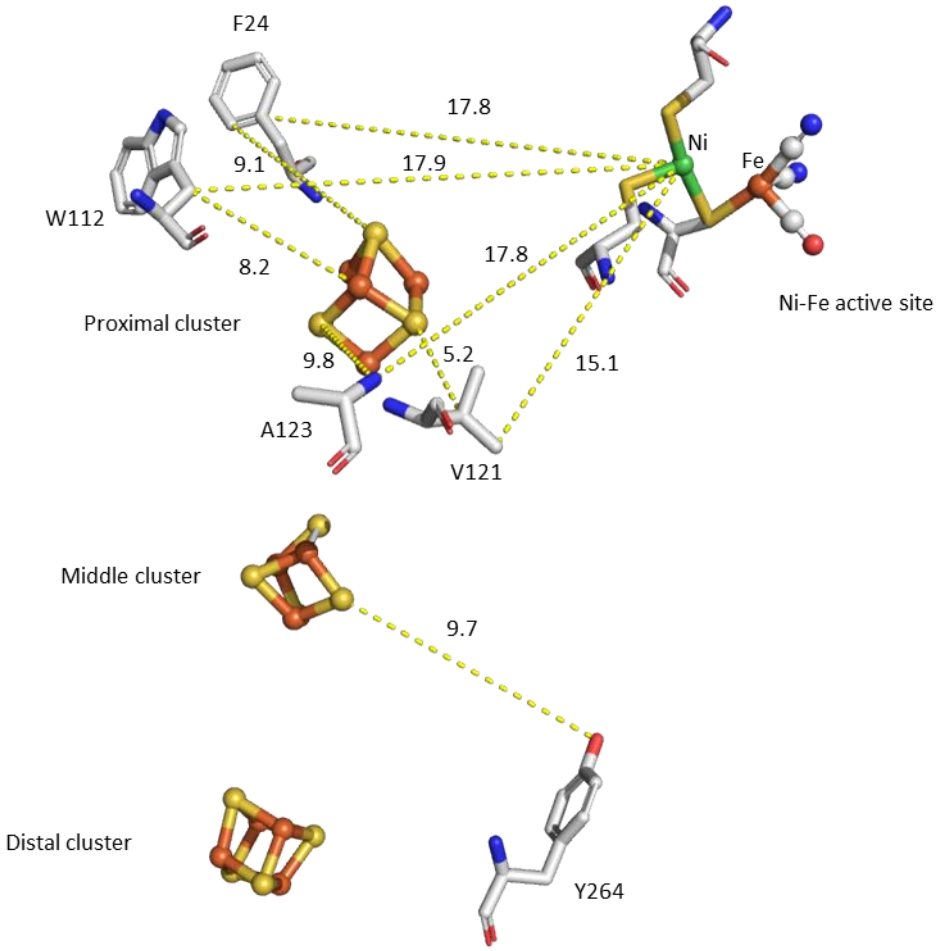


Figure 5.11 Amino acid residues (shown in grey sticks) around either proximal cluster or middle cluster (shown in yellow and orange sticks) of *E. coli* Hyd-1. Unit of distance: Å (Figure prepared using PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC, PDB code is 3USE)

Before carrying out experiments to test the tolerance of each site to Phe incorporation, hypothetical structures of the variants were generated using PyMOL. Figure 5.12 shows the modelled structure data of each variant. Among the five residues, it seems Phe-24 is the most promising position to mutate, as phenylalanine to tyrosine does not cause too much variation in the structure and it has predicted 81.9% chance of succeeding. The other three variants W112F, V121F and A123F display a strong clash to the neighbouring residues, especially V121 and A123, as shown in Figure 5.12 Panel B,C,D respectively. This may be due to the side chain volumes between valine, alanine and phenylalanine and differences in hydrophobicity and aromaticity. However, regarding the structure data of these three residues, as they are all buried in the

complex enzyme and not too close to the FeS cluster (between 5-10 Å), the possibilities of protein folding are 51%, 61% and 68% for these three variants, respectively.

As Y264 is on the surface of the protein, even though the structure of tyrosine is similar to phenylalanine, mutating tyrosine to phenylalanine may cause the protein to fold incorrectly, Figure 5.12 shows there are some clashes to the neighbouring residues in the Y264F variant and this mutation is predicted to have only 11.2% chance of success. All five sites were then tested experimentally by carrying out site-directed mutagenesis. Each variant was over-expressed and isolation attempted following the protocol of wild type *E. coli* Hyd-1.

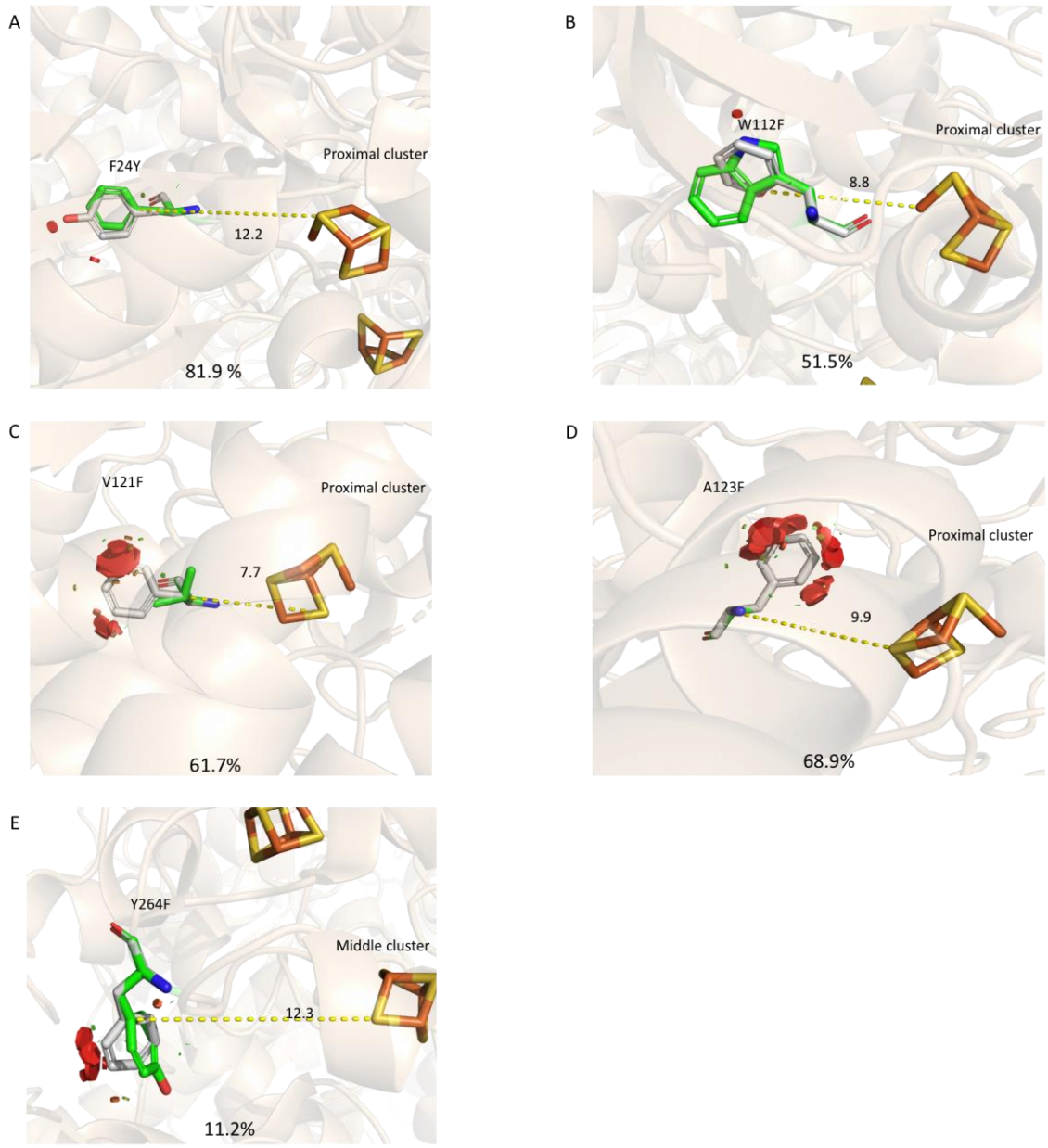


Figure 5.12 The structure data of each variant; the original amino acid residues shown in green, and in silico mutated amino acid shown in grey; FeS clusters are shown in yellow and orange sticks and the red dots represent the clashes between selected residue and neighbours (Figure prepared using the PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC, The PDB code is 3USE).

Subsequently, expression and purification were carried out to test the tolerance of the sites. As the integrity of overexpressed Hyd-1 has been verified, the Hyd-1 variants can be constructed through plasmids rather than the complex protocol for mutating the *chromosomal copies* of the Hyd-1 genes.[46]

Site directed mutagenesis was carried out via quick change PCR (mechanisms and procedures are discussed in section 2.1.3) and then each variant generated was confirmed by Sanger DNA sequencing (section 2.1.3). Each variant was overexpressed following the protocol of wild type *E. coli* Hyd-1 and the resulting proteins were isolated via Ni^{2+}-affinity chromatography. Only those variants that successfully folded could be isolated.

Table 5.4 summarises the results of overexpression, purification, and characterisation of each variant. The variants F24Y, V121F and A123F can be successfully expressed. A quick colorimetric assay with benzyl viologen was carried out to determine if these variants maintained their catalytic properties (oxidation of H_2). The colorimetric assays indicated that these three variants can catalyse the oxidation of H_2 and reduce benzyl viologen. For the other two variants W112F and Y264F, the benzyl viologen colour did not change in the activity assay test, which is consistent with the hypothesis from structure data shown by PyMOL in Figure 5.12.

Table 5.4 Summary of the overexpression, purification and characterisation of the different *E. coli* Hyd-1 variants generated.

Variants	Purification Results	Benzyl viologen colourimetric assay
F24Y	Protein obtained	Purple
W112F	No protein	Colourless
V121F	Protein obtained	Purple
A123F	Protein obtained	Purple
Y264F	No protein	Colourless

To further characterise the activity of *E. coli* Hyd-1 variants, protein film electrochemistry was carried out in each enzyme variant. Panel A and Panel C in Figure 5.13 show the cyclic voltammetry of F24Y and A123F which are similar to that of the wild type *E. coli* Hyd-1.[59,190] When the potential reached –0.3 V, catalytic current was observed which indicated that F24Y

and A123F variants had high H_2 oxidation activity. However, the variant V121F (Panel B) has much lower activity compared with the wild type, F24Y and A123F *E. coli* Hyd-1. The preliminary CV indicated that valine mutation to phenylalanine decreased the catalytic activity. However, further experiments need to be done to support this. Therefore, F24 and A123 residues were selected for incorporating *p*CNF.

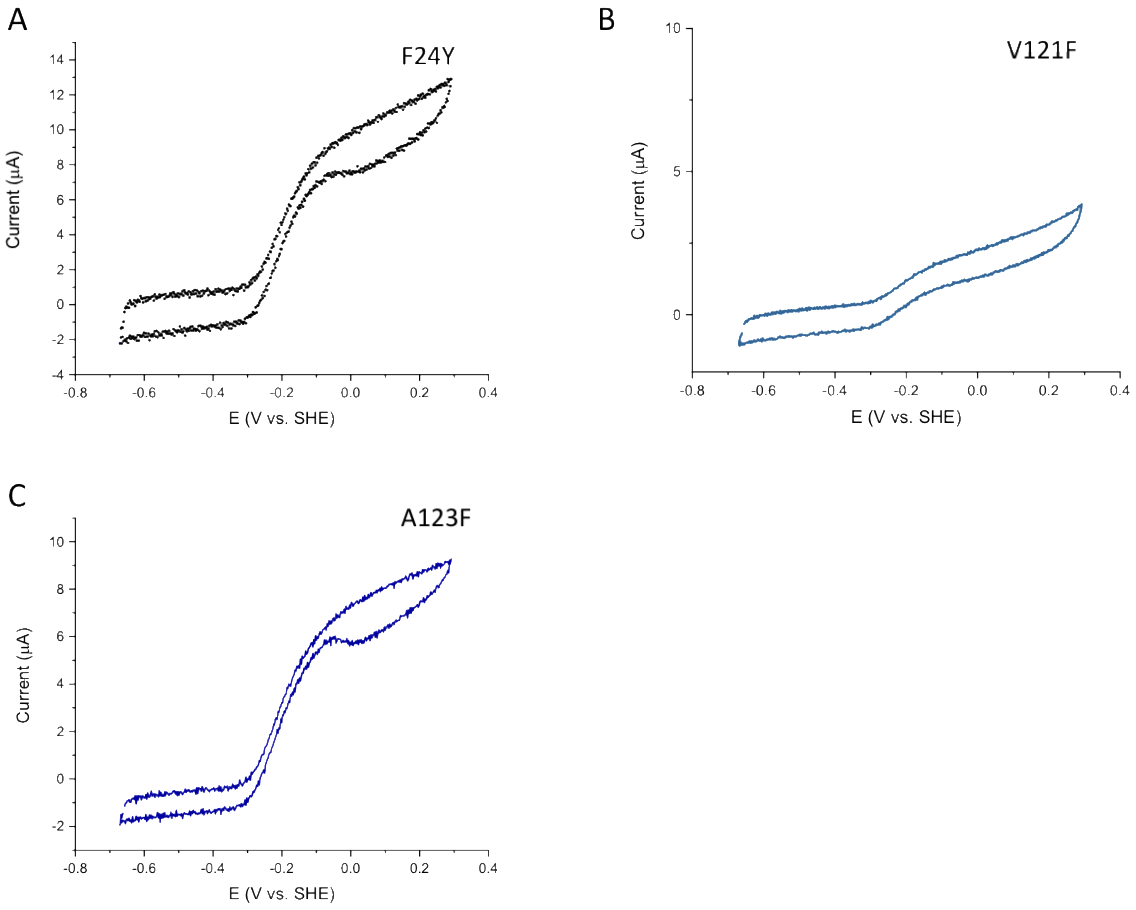


Figure 5.13 Cyclic voltammograms of protein film electrochemistry A: F24Y Hyd-1; B: V121F Hyd-1; C: A123F Hyd-1. Working electrode: pyrolytic graphite edge (PGE), rotating rate: 1000 rpm, 30°C, mixed buffer pH 6, 100% H_2.

5.6 Discussion and conclusion

5.6.1 Purification results of overexpressed Hyd-1

The aim of this chapter's work was to establish whether an overexpression system for Hyd-1 developed in the Vincent group could be adapted for expression of Hyd-1 incorporating unnatural amino acid. This is because the yield of Hyd-1 produced natively (chromosomally encoded Hyd-1) is very low, and large amounts of proteins are often required for some physical chemistry techniques, especially for infrared spectroscopy experiments.[59] On the other hand, the yield is quite low for incorporation of an unnatural amino acid, it is therefore essential to have high expression level system of Hyd-1 for incorporation unnatural amino acid. A plasmid-based overexpression system has been established in the Vincent group and the overexpressed Hyd-1 was successfully overexpressed with a highly increased yield compared to chromosomally encoded Hyd-1 (Table 5.1).

Purification steps previously published for chromosomally encoded Hyd-1, i.e., Ni^{2+}-affinity chromatography and size exclusion chromatography, were carried out to purify the overexpressed Hyd-1.[49] These processes were able to remove most of the impurities following Ni^{2+} affinity chromatography. Following size exclusion chromatography, two peaks that displayed Fe-S cluster absorbance at 420 nm in the UV-Vis spectrum (Figure 5.4A) were separated. Both peaks (P2 and P3) were shown to have the two subunits of the catalytic core of Hyd-1, with the large subunit (HyaB) and small subunit (HyaA) by SDS-PAGE (Figure 5.4B), suggesting that both two elution peaks contain Hyd-1 in different quaternary conformations. These conformations could arise from those having the cytochrome subunit (HyaC) still bound to the other subunits (P2), and those without the cytochrome (P3). To test this, further purification steps including separation via a hydroxyapatite column may be needed, but these are not required for the experiments described herein. Additional techniques such as analytical ultracentrifugation, light-scattering, or protein crystallisation are required to support this hypothesis.[176]

5.6.2 Characterisation of overexpressed Hyd-1

The biophysical characterisation of the overexpressed Hyd-1 was performed using conventional steady-state solution activity assays, protein film electrochemistry and IR spectroscopic techniques. The data from steady-state rate determination, Michaelis constant for H_2, and activation enthalpies indicates that within the likely errors, these samples have similar behaviour to the chromosomally encoded Hyd-1. Since P3 represented more material, it was used for spectroscopic experiments.

E_{Switch} is achieved from the potential at which the minimum in the first derivative of the voltammogram is detected. The E_{Switch} obtained in the overexpressed Hyd-1 was +0.23 V which is a little bit lower than the chromosomally encoded Hyd-1 with E_{Switch} about +0.242 V. The slightly difference of E_{Switch} is more likely due to the experiments error rather than the more reductive power needed on overexpressed Hyd-1 to activate Ni-B. Therefore, it indicates that the E_{Switch} of overexpressed Hyd-1 is consistent with the natively chromosomally encoded Hyd-1.

From the O_2 tolerance test of overexpressed Hyd-1, it behaves normally as there is an immediate drop in H_2 oxidation current following O_2 saturated buffer injection and the enzyme begins to reactivate before gas exchange is completed. This suggests that overexpressed Hyd-1 is O_2 tolerant, and Ni-B is the majority inactive state in the presence of O_2 rather than Ni-A. However, this is not the ultimate test of O_2 tolerance, further experiments such as testing the ability to continue to do H_2 oxidation in the presence of O_2 by performing chronoamperometry can be done.

Development of an overexpression system provides a big advantage for performing infrared spectroscopy, because the technique requires much higher protein concentrations than solution-based assays of activity or protein film electrochemistry.

Using transmission mode IR spectroscopy, the difference between the as-isolated and H_2-reduced states of overexpressed Hyd-1 is evident: there is a significant increase in the Ni-R, Ni-C and Ni-L states in the reduced form of Hyd-1, as expected, and the wavenumber positions for these states match those of natively chromosomally encoded Hyd-1.[116]

5.6.3 Site-directed mutagenesis of Hyd-1

Introduction of IR-probe unnatural amino acid could potentially be used to study the redox state of the FeS clusters during catalytic turnover. For example, this could confirm the role of the proximal cluster in O_2-tolerance Hyd-1. The aim of this work is to select several amino acid residues close to the FeS cluster relay in Hyd-1. Subsequently, mutants were created at these positions and the proteins were expressed to test the effect on enzyme structure and catalytic activity.

Five amino acids were chosen with a distance between 5–10 Å to either proximal or middle cluster in Hyd-1. The expression and purification result of each variant was consistent with the prediction model generated using structural data in PyMOL. Three variants did not fold correctly, as the mutation caused in a strong clash with neighbours. The other three variants F24Y, V121F and A123F were expressed and folded correctly. Additionally, protein film electrochemistry was used to test the catalytic properties of each variant. F24Y and A123F retained catalytic behaviour similar to the wild type enzyme, indicating that these two residues could be suitable for incorporating the unnatural amino acid *p*CNF.

The V121F variant had lower activity in the protein film electrochemistry. Based on its location, the mutation may disturb the FeS cluster relay for electron transfer. Additional experiments are needed to test this hypothesis, which are beyond the scope of this thesis.

5.6.4 Conclusion

The experiments performed in this chapter show that Hyd-1 can be overexpressed successfully. Two different Hyd-1 elution peaks were obtained and separated during size exclusion chromatography. These peaks show slightly different characteristics in the turnover frequency, Michaelis constant and activation enthalpies compared with natively produced Hyd-1 (chromosomally encoded Hyd-1). P3 was taken forward because it was the larger peak.

Two amino acid residues F24 and A123 were selected for incorporating the unnatural amino acid *p*CNF after testing tolerance by site-directed mutagenesis. Further experiments on incorporation of *p*CNF was put on hold because the overexpressed protein seemed to have some unusual IR active site spectra features in the protein film infrared electrochemistry and did not crystallise in preliminary trials. Also, the moderate level of overexpression, about 10-fold of native expression, would make this enzyme difficult to obtain in the *p*CNF variants in sufficient quantity for spectroscopic study.

On the other hand, as the overexpression system of [FeFe] hydrogenases from *Desulfovibrio desulfuricans* has been established, the apo-protein can be produced in *E. coli* with high yield and purity and maturated with the synthetic [2Fe] cofactor mimic ($[Fe_2(adt)(CO)_4(CN)_2]^{2-}$) to generate fully active holo-enzyme. It is interesting to see if *p*CNF can be used as an IR-probe to detect different oxidisation states on both apo- and holo-enzyme. Thus, the next chapter explores incorporation of the unnatural amino acid into [FeFe] hydrogenases.

6

Chapter 6: Incorporating an Unnatural Amino Acid into [FeFe] Hydrogenase from *D. desulfuricans*

6.1 Introduction and aim of the chapter

6.1.1 [FeFe] hydrogenases

[FeFe] hydrogenases catalyse proton reduction and molecular hydrogen (H_2) oxidation with very high turnover rates (catalytic turnover up to 10^4 μmol / min / mg) at low over-potentials (i.e. with essentially no energy waste) but they are extremely sensitive toward molecular O_2, getting inactivated even by traces of O_2.[191] The most well studied [FeFe] hydrogenases are: the [FeFe] hydrogenases from the green alga *Chlamydomonas reinhardtii* (*Cr*HydA1) representing the M1-type (see section 1.4 for more detail), the [FeFe] hydrogenases from the sulfate reducing bacterium *Desulfovibrio desulfuricans* (*Dd*HydAB) representative of the M2-type class, and M3-type [FeFe] hydrogenases from the bacterium *Clostridium pasteurianum* (*Cp*I).[19] The enzyme worked on in this project is *Dd*HydAB which is one of the most active bidirectional heterodimeric [FeFe] hydrogenases.[192]

*Dd*HydAB is a typical M2 type [FeFe] hydrogenase, consisting of two subunits: a large subunit HydA with the catalytic active site (44 kDa) and a small subunit (10 kDa) HydB which ´wraps´ the big subunit and which is essential for stability (Figure 1.12). In the *D. desulfuricans* bacterium, *Dd*H, [FeFe] hydrogenases can be found in the periplasm.[98]

In the HydA large subunit we can find the active site H-cluster as well as the cohort of FeS clusters, forming the electron relay (Figure 6.1). The H-cluster contains a [2Fe] subcluster ($[2Fe]_H$) and a [4Fe-4S] cluster ($[4Fe\text{-}4S]_H$) covalently linked through a cysteine-thiolate.[19] The $[2Fe]_H$ subcluster has two iron ions bridged by the sulfides of an aza-propane-1,3-di-thiolate (ADT) ligand, a bridging carbon monoxide ligand and two terminally bound cyanide and carbon monoxide. Cyanide is a strong σ donor but weak π acceptor, while carbon monoxide is a weak σ donor but strong π acceptor. The Fe centre closer to the $[4Fe\text{-}4S]_H$ subcluster is coordinatively-saturated and it is called proximal Fe (Fe_p). The other Fe is further away from the $[4Fe\text{-}4S]_H$ subcluster and therefore called distal Fe (Fe_d); it is unsaturated and therefore possessing an

open coordination site where the substrate (molecular hydrogen, protons) and inhibitors (such as CO and O_2) can bind.[102] *Dd*HydAB has two additional [4Fe-4S] clusters (named proximal and distal F-cluster according to their distance to the H-cluster) separated from each other by about 9 Å, whereas the distance between the proximal [4Fe-4S] cluster and the H-cluster is about 8 Å (Figure 6.1).[6,104]

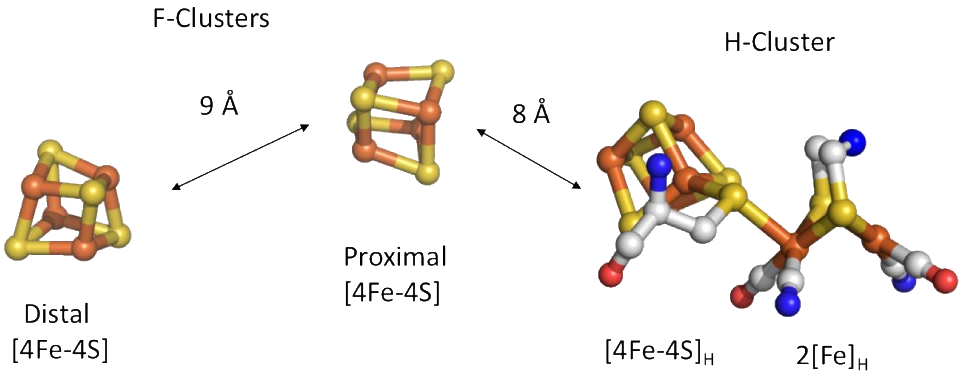


Figure 6.1 Arrangement of cofactors in *Dd*HydAB showning the F-clusters and the H-cluster. Yellow spheres represent sulfur atoms, brown spheres represent iron atoms, grey spheres represent carbon atoms, red spheres represent oxygen atoms, blue spheres represent nitrogen atoms. (PDB code: 1HFE.)

6.1.2 Biosynthesis of [FeFe] hydrogenase

During the expression of [FeFe] hydrogenases, the [4Fe-4S] clusters are built by the iron-sulfur cluster (ISC) assembly machinery of the cells (see section 1.1.3 for more detail). Figure 6.2 shows the scheme of in *vivo* [FeFe] hydrogenase H-cluster assembly. The synthesis and insertion of $[2Fe]_H$ sub-cluster into the apo-enzyme (the enzyme lacking the $[2Fe]_H$ sub-cluster) requires three essential FeS protein maturases: HydE, HydF and HydG.[193] HydE and HydG are both members of the radical S-adenosyl methionine (SAM) family. HydE is used for the synthesis of the bridging dithiolate (ADT), while HydG is responsible for producing the CO and CN^- ligands of the $[2Fe]_H$ sub-cluster using tyrosine as a substrate.[194] HydF acts as a scaffold protein for the $[2Fe]_H$ sub-cluster formation, binding and transferring the assembled cluster to the apo HydA enzyme to complete the assembly process of the active H-cluster.[193]

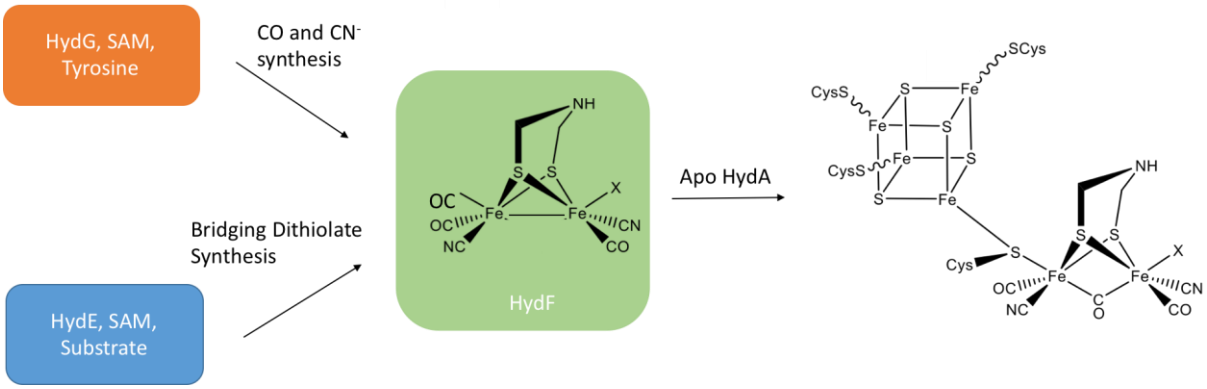


Figure 6.2 Maturation scheme and sequential model for the roles of the HydG, HydE and HydF maturases in the biosynthesis of [FeFe] hydrogenase H-cluster, SAM: radical S-adenosyl methionine.[194]

The low protein yield and low purity samples from expression in the native organism were the main challenges when working with [FeFe] hydrogenases. In recent years, a novel approach was developed to improve the expression level (Figure 6.3) and purity of the sample. The inactive 'apo-hydrogenase' (lacking the binuclear site and only containing the FeS cluster/clusters) is overexpressed in an alternative host system, *E. coli*.[195] However, as *E. coli* cells do not natively express [FeFe] hydrogenases, they lack the three specific maturases HydE, HydF and HydG (mentioned above) to synthesise the $[2Fe]_H$ sub-cluster. However, it has been shown that the inactive 'apo-hydrogenase' can be maturated artificially *in vitro* by adding a chemically synthesised $[2Fe]_H$ sub-cluster mimic (i.e., $[2Fe]^{adt}$ which stands for $[Fe_2(adt)(CO)_4(CN)_2]^{2-}$) with or without the scaffold HydF protein, generating a fully active holo-hydrogenase which is essentially identical to the native enzyme. Figure 6.3 shows the artificial maturation procedure in the absence of HydF.[195,196] To date, several [FeFe] hydrogenases have been successfully generated using this artificial maturation method including *Cp*I, *Cr*HydA1 and *Dd*HydAB among others.[140,197] This approach provides a strategy for the simple production of recombinant [FeFe] hydrogenases simplifying the purification step as well as the generation of variants by site-directed mutagenesis.

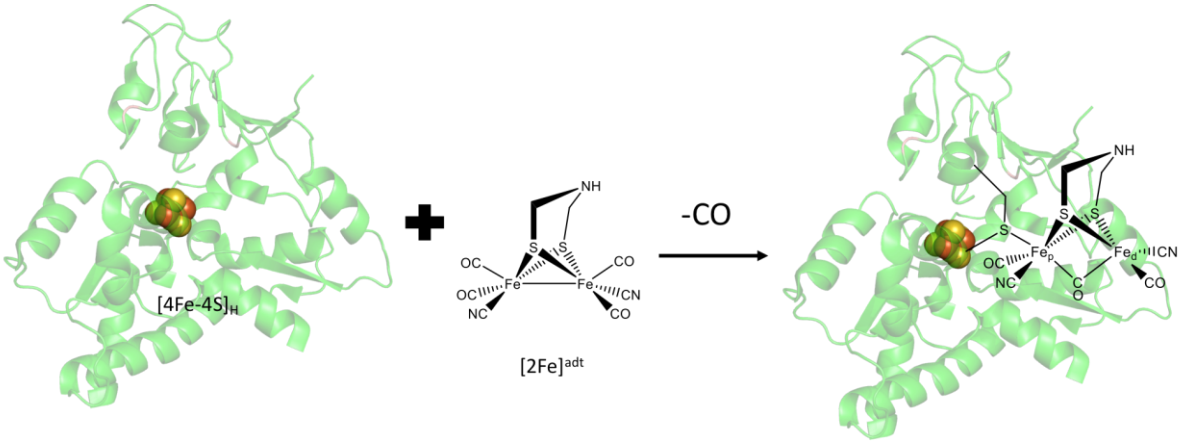


Figure 6.3 Scheme of the artificial maturation without assistance of the HydF maturase. The apo-protein shown in this scheme belongs to *Cr*HydA1 that only contains one [4Fe-4S] cluster (PDB code: 1HFE.)

6.1.3 Aim of this chapter

Following the preceding research on incorporating an unnatural amino acid *p*-cyano-L-phenylalanine (*p*CNF) close to the FeS cluster into the spinach ferredoxin to demonstrate that the *p*CNF can act as an IR reporter of FeS cluster redox state. In this chapter, *p*CNF will be incorporated into the [FeFe] hydrogenase from the sulfate reducing bacterium *Desulfovibrio desulfuricans* (*Dd*HydAB). This enzyme contains two additional [4Fe-4S] clusters in addition to the active site H-cluster. It was decided to incorporate the *p*CNF close to the proximal cluster in *Dd*HydAB, as the CO and CN^- ligands in the binuclear site $[2Fe]_H$ already serve as an IR probe to sense the redox state of $[4Fe\text{-}4S]_H$, thus it makes sense to target the proximal [4Fe-4S] cluster to be able to investigate how this cluster interplays with the catalytic active site using infrared spectroscopy. First, a recombinant system for *Dd*HydAB in *E. coli* will be established. Prior to *p*CNF incorporation, residues around the proximal cluster will be tested for tolerance to mutagenesis to find the most suitable position to introduce the non-natural amino acid *p*CNF.

6.2 Overexpression of apo-*Dd*HydAB [FeFe] hydrogenase and *in vitro* (artificial) maturation

The first step was to establish a heterologous overexpression system for *Dd*HydAB [FeFe] hydrogenase in the laboratory. Based on the method reported by Birrell *et al.* (Isr. J. Chem., 2016), using a codon optimised *Dd*HydAB plasmid and *E. coli* BL21 (DE3)*ΔiscR* strain. As explained above, as *E. coli* lacks the maturase system for synthesising and inserting the [2Fe] subcluster, so overexpression of *Dd*HydAB in *E. coli* generates the apo-hydrogenase (lacking $[2Fe]_H$ but containing the F-clusters) which artificially maturated with the synthetic active site mimic $[2Fe]^{adt}$ ($[Fe_2(adt)(CO)_4(CN)_2]^{2-}$).[140] First the plasmid pACYCDuet-1 HydAB for apo-*Dd*HydAB shown in Figure 6.4 was transformed into the *E. coli* BL21(DE3)*ΔiscR* cells; both the plasmid and strain were provided by Dr. J. Birrell and Prof. S. DeBeer (MPI-CEC, Germany). The plasmid had codon optimised DNA sequences encoding the HydA and HydB proteins lacking the N-terminal (HydB) and C-terminal (HydA) periplasmic targeting sequences and containing a C-terminal Strep tag on HydA. BL21(DE3)*ΔiscR* cells have a mutation in which the gene encoding the repressor IscR is deleted, leading to improved synthesis of FeS clusters and therefore better suited to generate proteins with many FeS clusters such as *Dd*HydAB.[138]

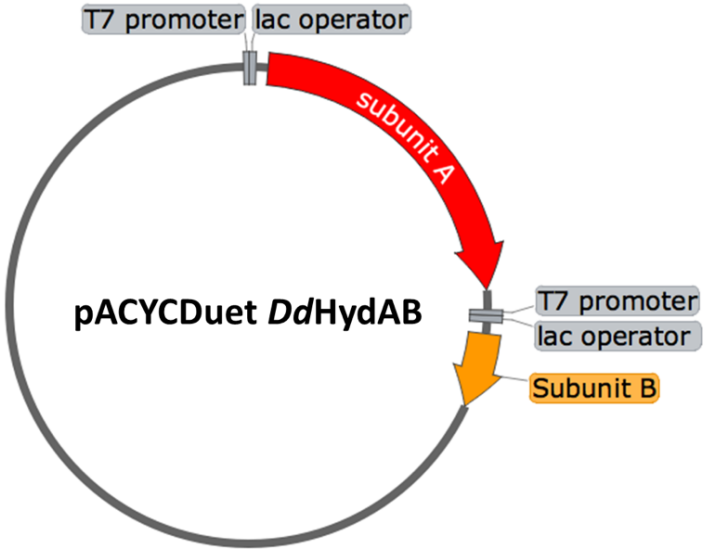


Figure 6.4 Graphic structure of the plasmid pACYCDuet-1 *Dd*HydAB. The red arrow shows the large subunit HydA gene, the orange arrow shows the small subunit HydB genes. Both are under the control of the T7 promoter. Map generated by SnapGene.

The procedures of *Dd*HydAB expression and purification were described in section 2.2.3. Fractions were analysed by SDS-PAGE as shown in Figure 6.5. The supernatant lane ('SUP') showed the presence of the two subunits (HydA at 44 kDa and HydB at 11 kDa). The eluted protein ('ELU' lane) was very pure.

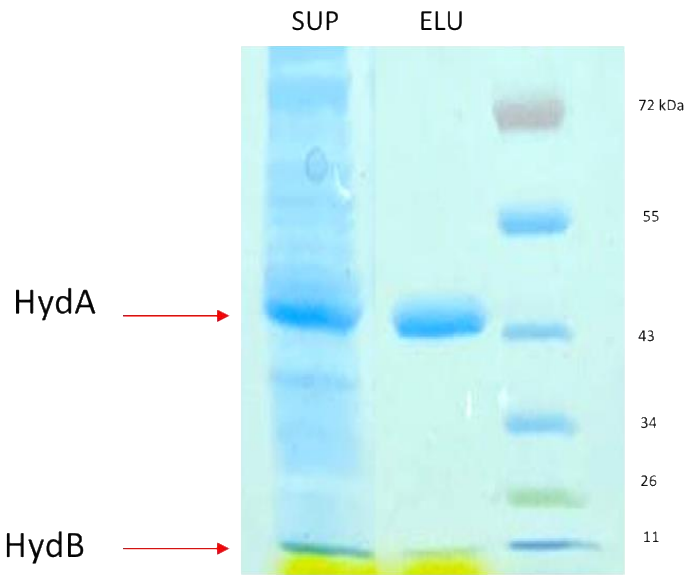


Figure 6.5 The SDS-PAGE of the apo-*Dd*HydAB [FeFe] hydrogenase after strep-tactin affinity chromatography. Red arrows indicate the desired bands for the protein at 44 kDa (large subunit HydA) and 12 kDa (small subunit HydB). SUP: the supernatant. ELU: the collected fraction after eluting with desthiobiotin.

To confirm the integrity of the FeS clusters in the apo-protein, UV-Visible spectroscopy was performed (Figure 6.6). The broad band at around 400 nm corresponds to the oxidised [4Fe-4S] cluster (arising from the metal-ligand charge transfer). The concentration of recombinant apo-*Dd*HydAB could be determined by UV-Vis spectroscopy using the Beer-Lambert law from the absorption at 400 nm (A_{400}), knowing that the molar extinction coefficient at this wavelength is 52500 cm^{-1} M^{-1} for the [4Fe-4S] cluster.[162] With that, the final yield of wild-type (WT) apo-*Dd*HydAB was found to be about 17 mg per litre of culture.

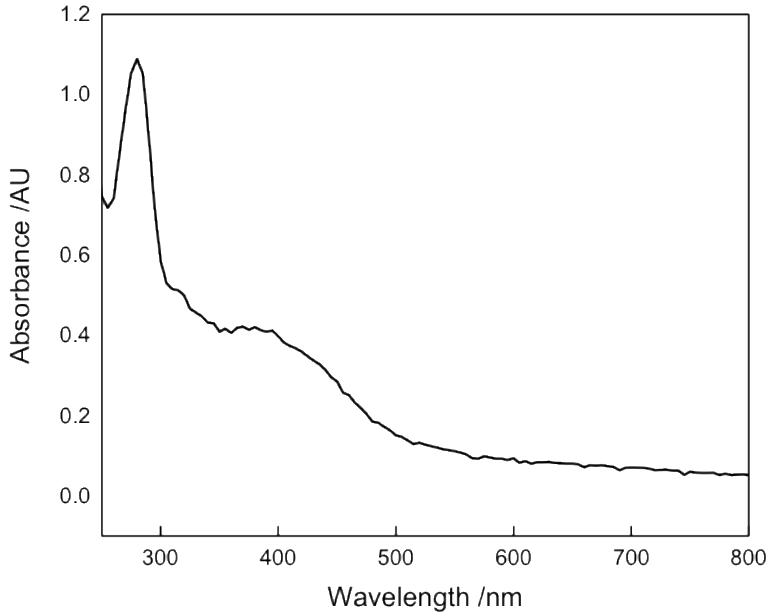


Figure 6.6 The UV-Visible (UV-Vis) absorption spectrum of purified apo-*Dd*HydAB. This absorbance corresponds to a 6 µM apo-*Dd*HydAB in 100 mM Tris-HCl, 150 mM NaCl, pH 8 measured aerobically in a Cary 60 spectrophotometer using a 1 cm pathlength UV-Vis cuvette.

Next, the apo-*Dd*HydAB was artificially maturated with the chemically synthesised $[Fe_2(adt)(CO)_4(CN)_2]^{2-}$ active site mimic under strict anaerobic conditions, to generate the holo-enzyme. The artificial maturation procedure was based on that reported by Birrell *et al*.[196] *Dd*H maturation is much slower than all the other [FeFe] Hydrogenases generated by this method and it is yet not understood why.[196] Thus, to compensate for disintegration of the $[2Fe]^{adt}$ over a long maturation time we have to utilise higher excess of $[2Fe]^{adt}$ (5-fold excess instead of 2-fold for others).[196] The procedure was carried out in a gloveBox (O_2 <2 ppm) under dim light because the holo-*Dd*HydAB is light sensitive (see chapter 2 for more detail).[198] The scheme of how the artificial maturation is believed to occur is shown above in Figure 6.3.

Figure 6.7 shows infrared spectra measured on the holo-*Dd*HydAB after being artificially maturated for 50 hours and then eluted through a desalting column (as explained above). Table 6.1 shows the summary of the FTIR bands with their assignment to previously identified states of

*Dd*HydAB. In Figure 6.7A, six bands can be observed (see also Table 6.1). The bands at 1963 cm^{-1} and 1940 cm^{-1} are typical CO stretching modes of the most oxidised state H_{ox} (the structure of H_{ox} is shown in Figure 6.7C). The 2016 and 1812 cm^{-1} bands have been assigned to the H_{ox}-CO state (the structure of H_{ox}-CO is shown in Figure 6.7C). The band at 2016 cm^{-1} is associated with the extrinsic CO ligand while the one at 1812 cm^{-1} corresponds to the bridging CO.[199] It can be concluded that after 50 hours of maturation, the predominant states in holo-*Dd*HydAB is the H_{ox}-CO state, which is due to a terminal CO on Fe_p rotating to a bridging position, followed by a CO released from the Fe_d during the artificial maturation process. This enzyme has been shown to have very high affinity for CO, thus the H-cluster binds the released CO molecules very tightly inhibiting the enzyme and forming the well-known H_{ox}-CO state.[200] In order to release the CO, the enzyme needs to be left to shake under a N_2 atmosphere for around 48 hours.[201] After incubating the sample over the weekend (around 48–50 h) inside the glovebox (N_2 atmosphere) at 15 °C (to prevent protein degradation) in the dark and with gentle shaking (300 rpm), another IR spectrum was measured (shown in Figure 6.7B). After CO release, the most prominent band is at 1940 cm^{-1} which corresponds to the H_{ox} state. The 1802 cm^{-1} band corresponds to the bridging CO of the H_{ox} state. Small peaks at 2017 and 1972 cm^{-1} can still be seen, which correspond to the bands of the two vibrationally coupled terminally-bound CO ligands of each Fe ion in the H_{ox}-CO state. In any case, this indicates that most of the H_{ox}-CO has disappeared after the CO-release treatment and the sample is now mainly in the H_{ox} state, in agreement with what has been observed in previous studies.[111,200,201]

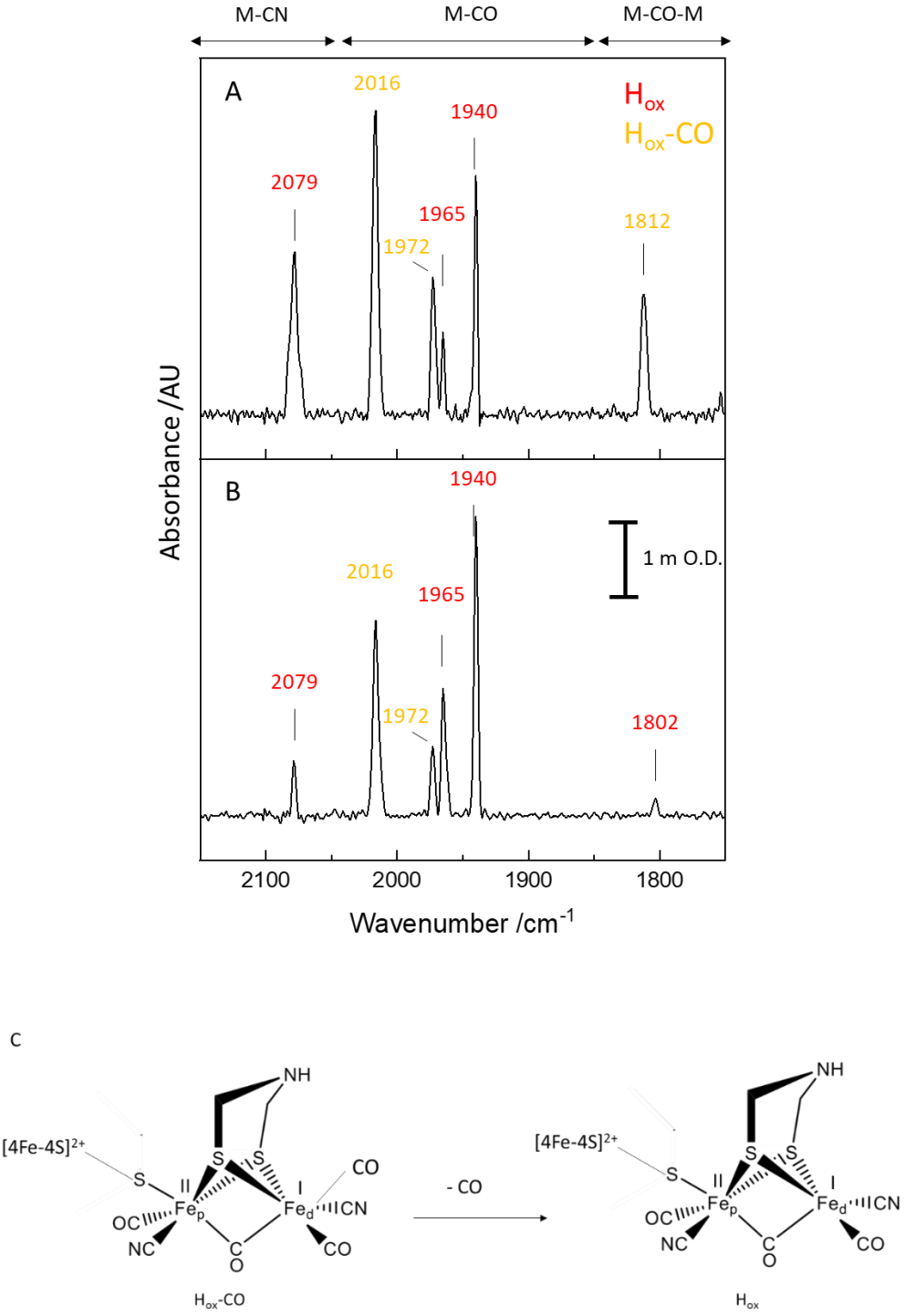


Figure 6.7 Panel A: FTIR spectrum of holo-*Dd*HydAB in 25 mM Tris-HCl, 20 mM NaCl pH 8 directly after being run through the desalting column. Panel B: FTIR spectrum of holo-*Dd*HydAB in 25 mM Tris-HCl, 20 mM NaCl pH 8 after 50 hours of CO release inside the glovebox (N_2 atmosphere) at 15 °C, shaking at 300 rpm and in the dark. The protein concentration is 2 mM. Peak values corresponding to the H_{ox}, H_{ox}-CO states are coloured in red and yellow respectively. 2 cm^{-1} resolution, 1024 scans. Panel C: The structure of active site stated in H_{ox} and H_{ox}-CO states.

Table 6.1 Summary of the FTIR bands for *Dd*HydAB [FeFe]-hydrogenases. The two iron ions are called proximal (Fe_p) and distal (Fe_d) according to the distance to the $[4Fe\text{-}4S]_H$ cluster (Reference: J.A. Birrell, *Isr. J. Chem.*, 2016, 56, 852–863).[140]

Source	Recombinant in *E. coli*					
FTIR	Fe_d-CN	Fe_p-CN	Fe_p-CO	Fe_d-CO	Fe_p-CO	μ-CO
H_{ox} /cm^{-1}	2093	2079	–	1965	1940	1802
H_{ox}-CO/cm^{-1}	2096	2089	2016	1972	1963	1812
H_{red}/cm^{-1}	2079	2041	1964	1916	1894	–
H_{sred}/cm^{-1}	–	–	1955	1932	1883	–

Next, the catalytic activity of the artificially maturated wild type (WT) *Dd*HydAB for H_2 oxidation was assessed by means of solution activity assays. This is done by saturating the buffer with H_2 in the presence of benzyl viologen as electron acceptor. Briefly, the reduction of 1 mM of benzyl viologen by 0.1 nM enzyme is followed in the UV-Vis (more detail can be seen in section 2.8).[196] The absorbance of benzyl viologen was followed at 600 nm (ε_{600} = 7.0 mM^{-1} cm^{-1}) for up to 2 minutes and initial rates were calculated by measuring the gradient over the first 10–20 s. The activity of WT *Dd*HydAB for H_2 oxidation was 51,000 $s^{-1} \pm$ 9500 s^{-1} (47,000 U/mg $\pm$ 8500 U/mg) which is consistent with the values published in the literature for the artificially maturated WT *Dd*HydAB (63,000 $s^{-1} \pm$ 9000 s^{-1}).[140] The activities of the different enzymes are compared in Figure 6.8.[202,203]

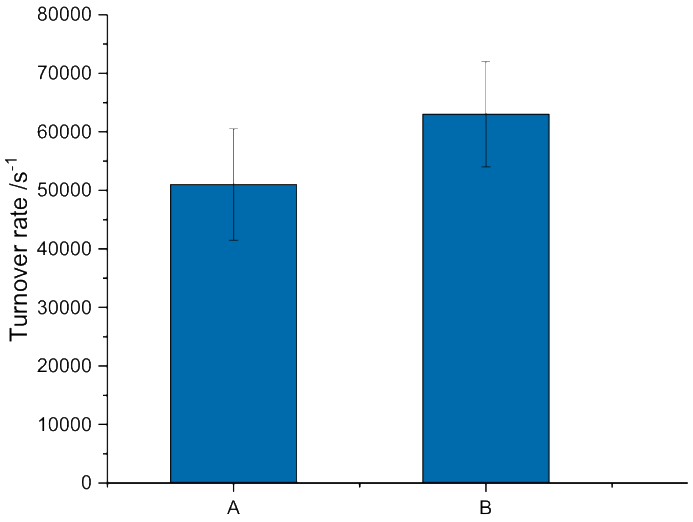


Figure 6.8 Activity assay of H_2 oxidisation measured following the reduction of 1 mM benzyl viologen by *Dd*HydAB. A: Artificially maturated *Dd*HydAB in this research; B: Artificially maturated *Dd*HydAB in other work.[140] 6 repetitions on each protein and the thin vertical line corresponds to the calculated error, the error bar represents ± 1 standard deviation.

Subsequently, to investigate the catalytic behaviour of the enzyme for both H_2 oxidation and production, protein film electrochemistry was performed on the artificially maturated *Dd*HydAB at pH 6 and 25°C. The cyclic voltammogram (CV) resembles previous published reports on the native and artificially maturated *Dd*HydAB enzyme.[107] The CV can be separated into three regions: H_2 production, H_2 oxidation, and inactivation at high potential (HPI) are indicated by blue, orange and green boxes, respectively, Figure 6.9. H_2 production is at potentials more negative than the thermodynamic potential (negative currents), the H_2 oxidation at potentials is more positive than the thermodynamic potential (positive currents). The negative current is associated with H^+ reduction to produce H_2, as the enzyme takes electrons from the electrode to reduce protons to H_2. The current crosses zero at about –0.37 V vs. SHE, which is consistent with the thermodynamic potential for the $2H^+/H_2$ couple (which is –0.354 V as calculated from the Nernst Equation for pH 6 and 1 bar H_2). When the potential is more positive than the thermodynamic potential of $2H^+/H_2$ at pH 6, the enzyme is oxidising H_2 (positive current). At very

oxidative potentials ($E_{applied}$ > 0 V vs. SHE), the catalytic H_2 oxidation current drops due to an inactivation of the enzymes, a process which is so far not well understood. The current can be recovered by reducing the enzyme on the reverse scan. The potential at which the current recovers is called the switch potential (E_{switch}) which is associated with anaerobic interconversion between inactive and active states and is defined as the steepest point in the reductive activation direction in the cyclic voltammetry.[204] The switch potential (E_{switch}) of *Dd*HydAB is about 0.03 V at pH 6, 25°C (scan rate 0.02 V/s).

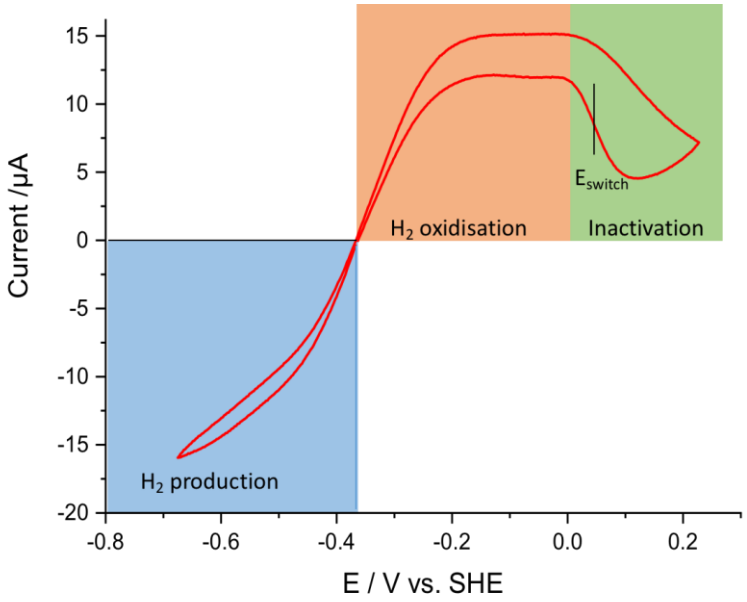


Figure 6.9 Protein film electrochemistry of holo-*Dd*HydAB adsorbed on a pyrolytic graphite electrode. Cyclic voltammetry was performed under 1 atm H_2, at 25°C in a 10 mM MES/HEPES, 100 mM NaCl, pH 6 buffer at 20 mV s^{-1}, and with electrode rotation at 2000 rpm. The regions of the voltammogram corresponding to H_2 production, H_2 oxidation, and high-potential inactivation (HPI) are indicated in blue, orange and green boxes, respectively.

It can be concluded that a heterologous overexpression system for wild type *Dd*HydAB has been reproduced. *Dd*HydAB can be expressed and purified successfully in high yields and purity. Following the same strategy as for spinach ferredoxin (see section 3.3 for more detail), the next step is to identify amino acid residues close to the proximal [4Fe-4S] will be selected for potential incorporation of the unnatural amino acid *p*CNF. The targeted residues should be at an

appropriate distance (about 5-10 Å) to the proximal [4Fe-4S] cluster. If *p*CNF is incorporated very far away from the FeS cluster, then the ν_{CN} stretch in the IR will not be able to respond to the change in electron density in the cluster; while if *p*CNF is too close to the cluster (less than 5 Å), it may disturb the function of the cluster.

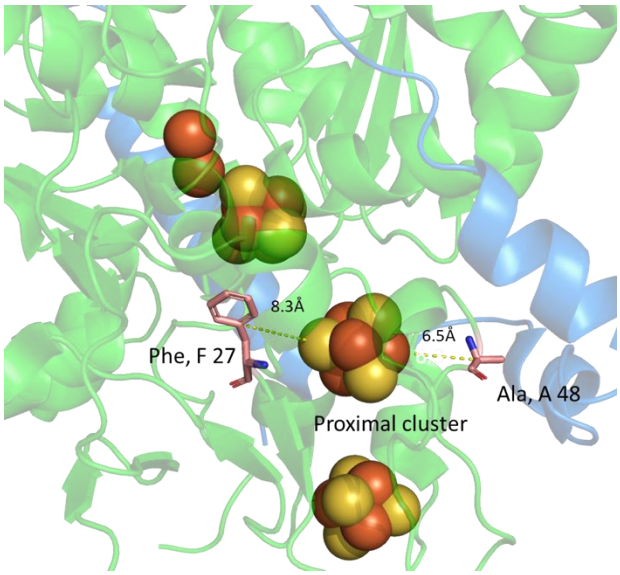


Figure 6.10 The overall structure of *D. desulfuricans* HydAB zoomed in around the proximal cluster. The green cartoon represents the large subunit HydA and the blue cartoon represents the small subunit HydB. [4Fe-4S] clusters and the $[2Fe]_H$ subcluster are represented by spheres, with the yellow ones being the S atoms and the red ones being the Fe atoms. Phenylalanine 27 and Alanine 48 are shown in sticks (PDB code: 1HFE).

Figure 6.10 shows the two targeted amino acids close to the proximal cluster, phenylalanine-27 (Phe, F27) and alanine (Ala, A48) which are at a distance of 8.3 Å and 6.5 Å from the proximal cluster respectively. For the F27 residue, as the structure of the phenylalanine is very similar to that of *p*CNF (**Error! Reference source not found.**A), the incorporation of *p*CNF will be attempted directly without trying any native amino acid mutation test beforehand (**Error! Reference source not found.**A). On the other hand, as the structure of alanine differs quite a lot from that of *p*CNF, the alanine (A) site was fist tested for native mutagenesis suitability by mutating it first to tyrosine (Y) and phenylalanine (F), as tyrosine and phenylalanine are the most similar residues to *p*CNF. The first point is to confirm that mutating residue 48 does not disrupt

the formation of the [4Fe-4S] cluster or the correct folding of the protein, as it involves a significant structural change in this residue. If the native mutagenesis proves successful, then this residue will be mutated to the unnatural amino acid *p*CNF.

6.3 Incorporation of the unnatural amino acid *p*CNF around the proximal cluster in *Dd*HydAB: expression and purification

In this section it will be described how *p*CNF was incorporate in apo-*Dd*HydAB in the targeted position. As mentioned above, for the position F27, mentioned in section 6.2 as the structure of phenylalanine is very similar to that of *p*CNF, it was incorporated directly without previous native amino acid mutation tests.

For the F27CNF, the codons of F27 (TTT) were replaced with the amber stop codon (TAG) by site-directed mutagenesis, which was then used to incorporate the unnatural amino acid *p*CNF as explained in Chapter 3. The side-directed mutagenesis was carried out via quick change PCR and then the plasmid of each variant was sent for sequencing. When the sequencing results matched the required plasmid sequence the pACYCDuet HydAB plasmid (with the amber codon (TAG) at the targeted positions), then the plasmid was transformed into the BL21(DE3)*ΔiscR* cells together with the second plasmid (pULTRA-CNF), which contain the genes for the unnatural aaRS/tRNA pair.

The over-expression procedures for *Dd*HydAB-CNF mutants are slightly different from that described in section 6.2. Pre-culture of the bacteria was left to grow overnight at 37 °C at 300 rpm. Then 10 mL of the bacteria preculture were added to 1 L of LB-phosphate broth media (20 g L^{-1} Luria Bertani (LB), 5 g L^{-1} Na_2HPO_4, 10 g L^{-1} K_2HPO_4) containing three antibiotics: 50 µg mL^{-1} kanamycin, 34 µg mL^{-1} chloramphenicol, 75 µg mL^{-1} spectinomycin. The medium was supplemented with 2 mM ammonium ferric citrate, 0.5% w/v glucose and 2 mM cysteine. When the OD_{600} reached about 0.6, the culture was cooled down in a water-bath for 10 min. Then, *p*-

cyano-L-phenylalanine (*p*CNF) was dissolved in 0.5 M NaOH and added to the main culture to yield a final concentration of 3 mM. Afterwards, the culture was bubbled with argon for 1 hour to make it anaerobic, followed by induction of protein expression by adding 1 mM IPTG and an additional 3 mM cysteine. Finally, the culture was placed in the incubator overnight at 16 °C with gentle shaking. The following morning, the cells were then harvested by centrifugation (6500 *xg*, 30 min, 4°C) under strictly anaerobic conditions, ready for protein isolation and purification. Strep-tactin affinity chromatography was carried out to purify the protein again, and protein purification steps follow the protocol for *Dd*HydAB explained in the previous section.

Figure 6.11 shows the SDS-PAGE result of F27CNF, and a comparison between adding *p*CNF into the media during expression. Lane 1 is the purified elution which did not have additional *p*CNF during the protein expression (as a negative control experiment); no full-length protein had bound and eluted from the strep-tactin column, as expected because only the truncated protein which does not have the strep tag at C-terminal end is expressed. On the other hand, lane 2 shows the result of purification from cells with 3 mM final concentration of *p*CNF added into the culture during protein expression. There are only two bands in the elution lane which correspond in molecular weight to the large subunit HydA (44 kDa) and small subunit HydB (12 kDa) respectively.

The UV-vis spectrum of F27CNF apo-*Dd*HydAB is identical to wild type apo-*Dd*HydAB (Figure 6.6) which shows FeS clusters seemed correctly assembled. The concentration of F27CNF apo-*Dd*HydAB can be determined by UV-Vis spectroscopy by the absorption at 400 nm (A_{400}) with a molar extinction coefficient of 52500 cm^{-1} M^{-1} for the [4Fe-4S] cluster.[162] The final yield for the F27CNF variant is about 3.5 mg per litre of culture, which is about 30% of the wild type yield.

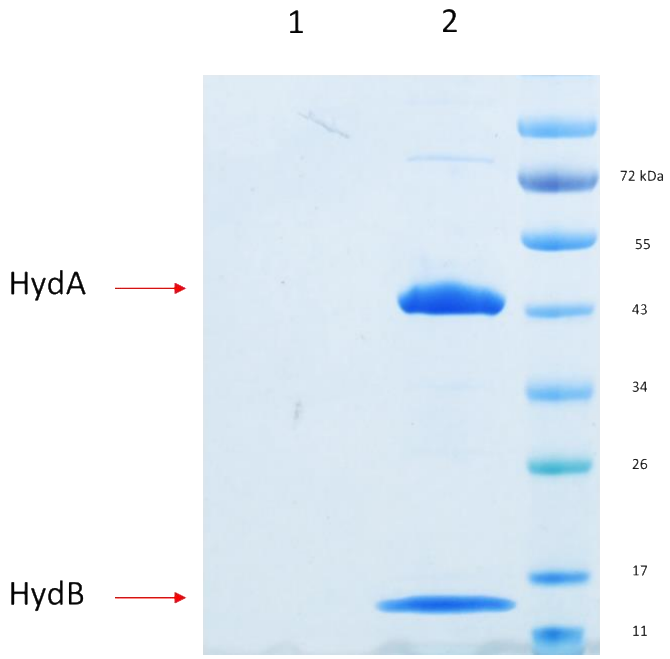


Figure 6.11 The SDS-PAGE analysis of apo-*Dd*HydAB [FeFe] hydrogenase after the strep-tactin affinity chromatography. Red arrows indicate the desired bands for the protein at 44 kDa (large subunit HydA) and 12 kDa (small subunit HydB). Lane 1: Elution lane of purified protein with no *p*CNF added into the main culture during protein expression (negative control). Lane B: Purified F27CNF with 3 mM final concentration of *p*CNF in the culture.

6.4 Characterisation of F27CNF *Dd*HydAB

Before maturating the apo-F27CNF *Dd*HydAB with the active site cofactor (ADT), transmission-mode IR spectroscopy was performed on oxidised and reduced apo-samples to see whether the the CN stretch in *p*CNF responded to the different oxidation states of proximal FeS cluster in the apo-protein without the catalytic active site. Around 5 µL of about 2 mM apo-F27CNF *Dd*HydAB were loaded into the transmission-mode IR cell. Figure 6.12A shows the raw data of transmission-mode infrared spectra of both oxidised and reduced F27CNF apo-*Dd*HydAB and zoomed in between 2220–2250 cm^{-1}. Baseline subtracted IR spectra and peak fitting IR spectra are shown in Figure 6.12B and C. In the oxidised sample shown in Figure 6.12B, a peak at 2233.9 cm^{-1} could be observed in the right region for the aromatic nitrile substituent. The UV-Vis spectrum of the apo-*Dd*HydAB suggests that all the FeS clusters in the 'as isolated protein' are oxidised. Then, the protein was chemically reduced by addition of 10 mM Eu(II)-DTPA which is a

very powerful reducing agent ($E_{1/2}$= –1.09 V *vs*. SHE).[205] The Eu(II)-DTPA reduced IR spectrum is shown in Figure 6.12C. There is a small (about 1.1 cm^{-1}) but observable shift in the peak position between the oxidised and reduced F27CNF samples (2233.9 cm^{-1} in the oxidised form *vs.* 2232.8 cm^{-1} in the reduced form, the wavenumbers of both oxidised and reduced forms are from the fitted peaks). This seems to indicate that the ν_{CN} in *p*CNF responds to the oxidation state of the proximal cluster.

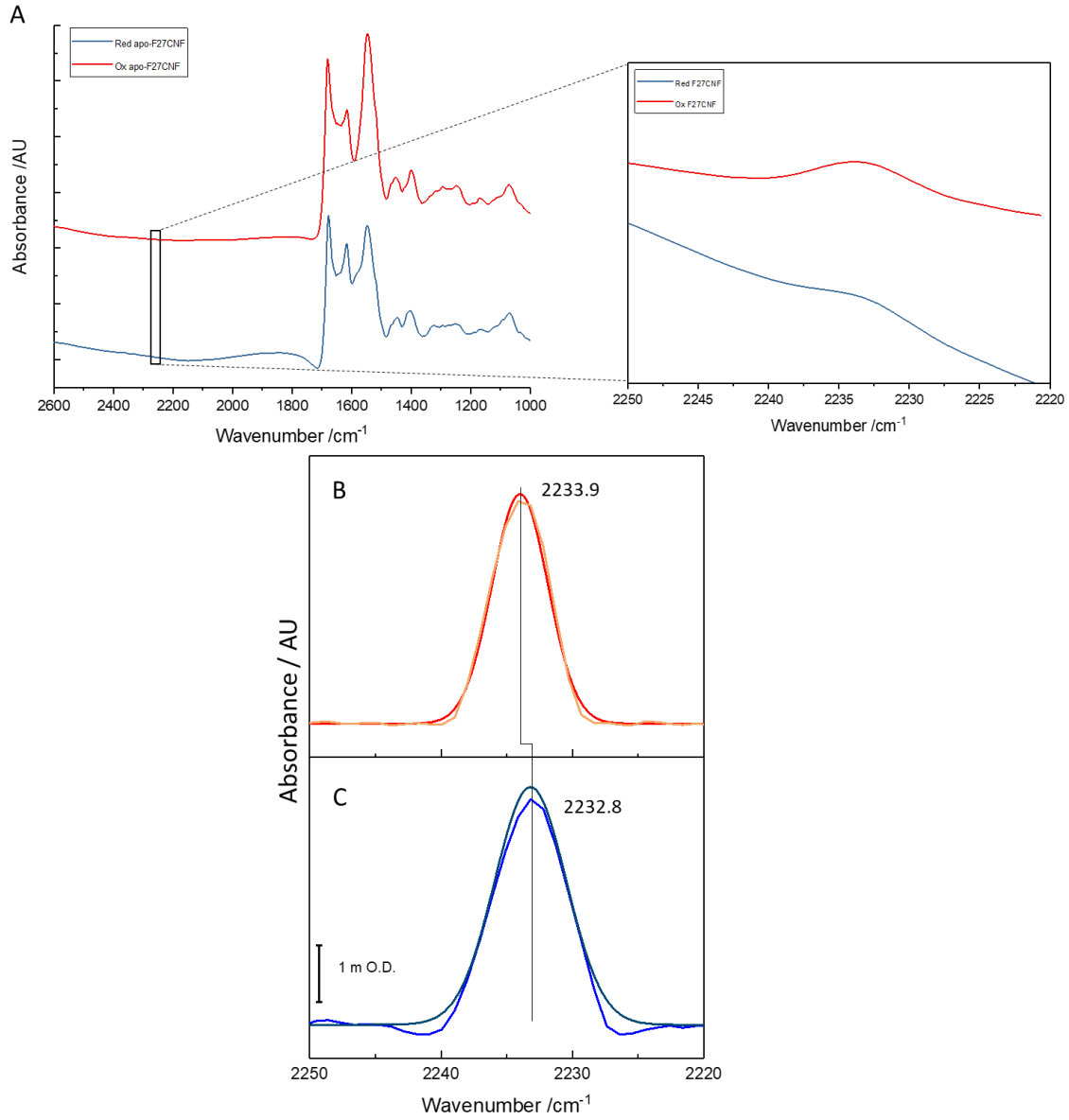


Figure 6.12 Panel A: Raw data of transmission-mode infrared spectra of both oxidised and reduced F27CNF apo-*Dd*HydAB, zoomed in between 2220-2250 cm^{-1}. Panel B: IR spectra of oxidised F27CNF apo-*Dd*HydAB; baseline subtracted IR spectrum is shown in red and the fitted peak with Gaussian functions is shown in orange colour. Panel C: IR spectra of reduced F27CNF apo-*Dd*HydAB; base-line subtracted IR spectrum is shown in light blue colour and the fitted peak with Gaussian functions is shown in dark blue colour. Close-up view of the 2220 to 2250 cm^{-1} region, highlighting the cyanide group in *p*CNF peak shifts. Protein concentration: 2.0 mM. Conditions: 100 mM Tris-HCl 150 mM NaCl buffer pH 8, 25 °C, 2 cm^{-1} resolution, 1024 scans.

Figure 6.13 shows the activity of the artificially maturated F27CNF *Dd*HydAB measured by the H_2-dependent reduction of 1 mM benzyl viologen by about 0.1 nM enzyme (following the same procedure as for wild type *Dd*HydAB, more detail in section 2.4).[196] The activity of F27CNF *Dd*HydAB for H_2 oxidation was 12,000 $s^{-1} \pm$ 3000 s^{-1} (11,000 $\pm$ 2750 U/mg), which is about a quarter of the H_2 oxidation activity of wild type. This could indicate that the *p*CNF might block the $[2Fe]^{adt}$ sub-cluster insertion channel and only a fraction of the protein is maturated.

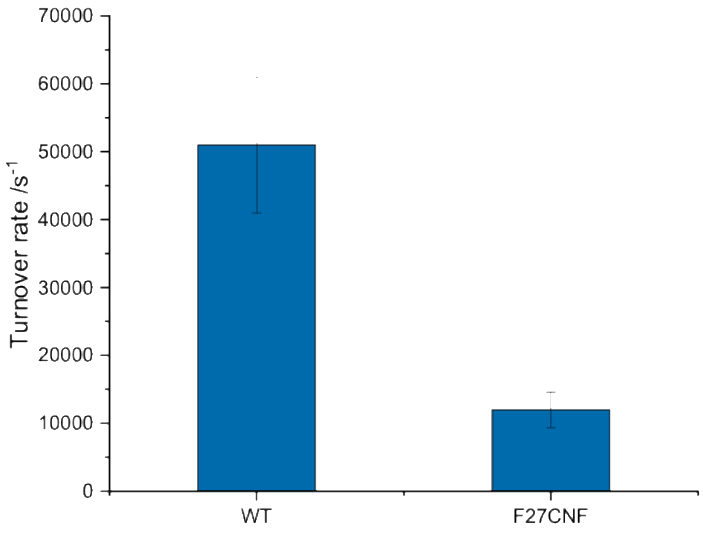


Figure 6.13 Comparison of the H_2 oxidation activity of wild type *Dd*HydAB and F27CNF variant. 6 repetitions on each protein and the thin vertical line corresponds to the calculated error, the error bar represents $\pm$ 1 standard deviation.

To investigate the catalytic behaviour of the F27CNF *Dd*HydAB for both H_2 oxidation and production, protein film electrochemistry was carried out as shown in Figure 6.14, by adsorbing the protein onto a pyrolytic graphite electrode. The catalytic behaviour shown in the CV is very similar to the wild type *Dd*HydAB with a switch potential E_{switch} of about 30 mV at 25 °C.[107,140] It can also be separated into three regions: H_2 production (blue box), H_2 oxidation (orange box) and the high-potential inactivation region (green box).

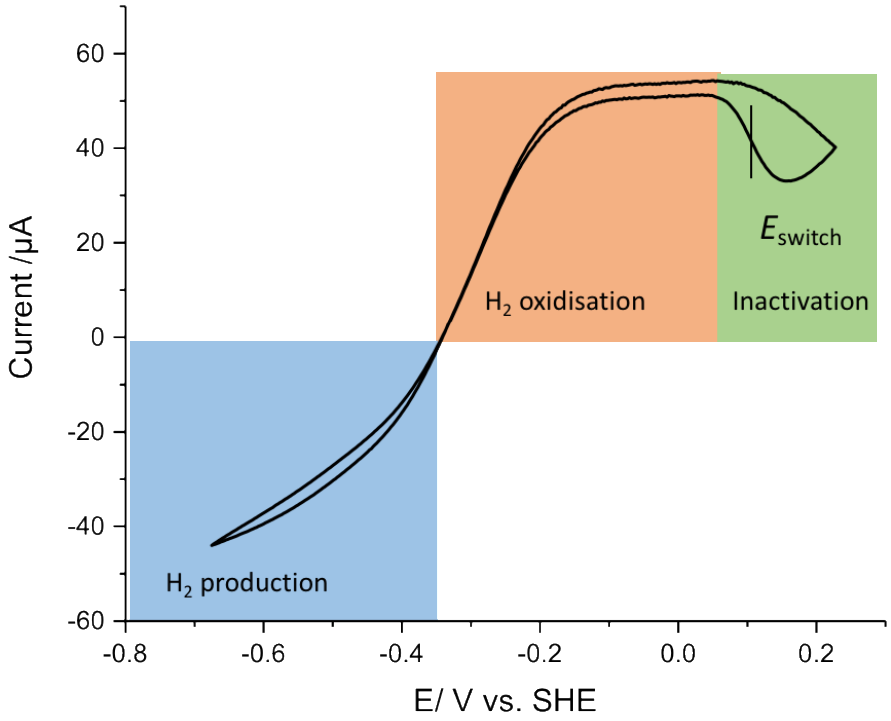


Figure 6.14 Protein film electrochemistry of holo-F27CNF *Dd*HydAB adsorbed to a pyrolytic graphite electrode. Cyclic voltammetry was performed under 1 atm H_2, at 25 °C in a 10 mM MES/HEPES, 100 mM NaCl, pH 6 buffer at 20 mV s^{-1}, with electrode rotation at 2000 rpm. The regions of the voltammogram corresponding to H_2 production, H_2 oxidation, and inactivation at high potential are indicated by blue, orange and green boxes, respectively.

Subsequently, to see whether the vibrational mode of the CN in *p*CNF is able to respond to the different oxidation states of the proximal FeS cluster in the holo-protein, FTIR spectra of holo-F27CNF *Dd*HydAB under oxidising and reducing conditions are shown in Figure 6.15A and Figure 6.15B. In the FTIR spectrum of the oxidised protein, the most prominent band is 1940 cm^{-1} which corresponds to the H_{ox} state (the most oxidised state). The small peaks at 1812 cm^{-1} and 2017 cm^{-1} correspond to the H_{ox}-CO state. In addition to the bands of the CO and CN^- of the active site, the band at higher energy (2232 cm^{-1}) corresponding to the CN in *p*CNF.

Next, the holo-F27CNF *Dd*HydAB was reduced by addition of Eu-DTPA(II) (10 mM). In Figure 6.15B, the most intense peak in the FTIR spectrum at 1894 cm^{-1} is from the H_{red} state shown in

Figure 6.16. However, some low intensity peaks (1940 cm^{-1} and 2093 cm^{-1}) belonging to the H_{ox} state are also observed, which can be explained by the fact that the active holo-enzymes can reduce protons to H_2 leading to re-oxidation to the H_{ox} state.[200] The cyano band of *p*CNF in the holo-enzyme is not very intense but second derivatives allow for wavenumber determination (shown in Figure 6.15C). From Figure 6.15D, the intensity of the cyano band from *p*CNF is much weaker than the bands from the active site ligands. Figure 6.15C shows the comparison of the FTIR spectra of F27CNF in oxidised and reduced states displayed in the second derivative mode and zoomed in the nitrile stretching region. There is a small but clear shift between the main peaks in the oxidised H_{ox} state (ν_{CN} = 2233.3 cm^{-1}) and reduced H_{red} (ν_{CN} = 2232.8 cm^{-1}) state ($\Delta\nu_{CN}$ = 0.5 cm^{-1}), which can be associated with the change in the oxidisation state of the proximal FeS cluster. This result is also consistent with the apo-F27CNF *Dd*HydAB oxidised/reduced experiments shown in Figure 6.12, which suggests that ν_{CN} of *p*CNF directly responds to a change in the local electrostatic environment of the proximal FeS cluster. However, the shift in holo-F27CNF ($\Delta\nu_{CN}$ = 0.5 cm^{-1}) is smaller than in the apo-F27CNF ($\Delta\nu_{CN}$ = 1.1 cm^{-1}), which is probably because the reduced holo-enzyme is active in proton reduction, and some of the enzyme is re-oxidised. In this scenario two ν_{CN} bands of different intensities and separated by about 1 cm^{-1} should be observed. One possibility is that the apparent shift of 0.5 cm^{-1} is due to the overlap of the cyano bands of oxidised and reduced forms of the holo-enzyme.

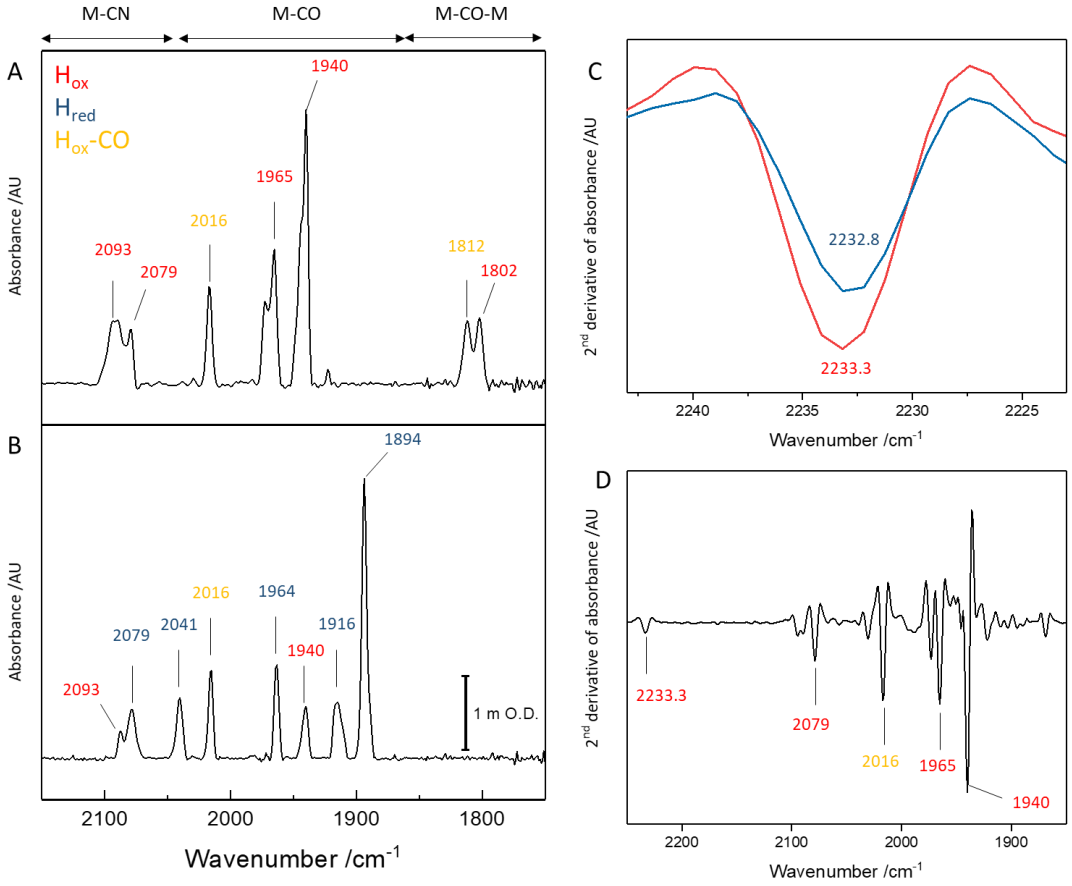


Figure 6.15 Panel A: FTIR spectrum of oxidised holo-F27CNF *Dd*HydAB. Panel B: FTIR spectrum of reduced holo-F27CNF *Dd*HydAB. Bands corresponding to the H_{ox}, H_{red}, and H_{ox}-CO states are coloured in red, blue and yellow respectively. Panel C: FTIR spectra of F27CNF zoomed in the nitrile stretching region displayed in the second derivative mode, red spectrum is from oxidised holo-F27CNF and blue spectrum is from the reduced holo-F27CNF. Panel D: FTIR spectra of oxidised holo-F27CNF between 1850 cm^{-1} to 2250 cm^{-1} displayed in the second derivative mode. Conditions: 100 mM Tris-HCl 150 mM NaCl buffer pH 8, 25°C. Protein concentration: 2 mM. 2 cm^{-1} resolution, 1024 scans.

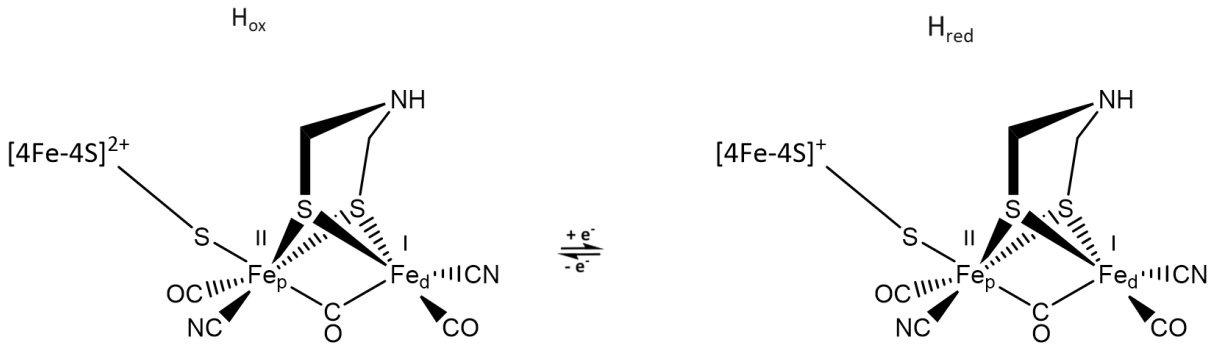


Figure 6.16 The structure of active site in the H_{ox} and H_{red} states of FeFe hydrogenases.

6.5 *Dd*HydAB mutagenesis suitability tests around the proximal [4Fe-4S] cluster.

As previously mentioned that the structure of alanine is quite different to *p*CNF, therefore, site directed mutagenesis on the residue alanine 48 (6.5 Å away from the proximal cluster) was carried via quick change PCR (see section 2.1.3 for more detail). Each generated variant (A48F and A48Y) was sent for sequencing. The nomenclature of the variant is based on the abbreviation and sequence number of the targeted amino acid, e.g., alanine 48 (A48) mutated to phenylalanine is called A48F while alanine 48 (A48) mutated to tyrosine (Y) is named A48Y. When the sequencing results matched the correct plasmid sequence, then the mutated plasmids were transformed into BL21(DE3)*ΔiscR* cells to start with the overexpression procedure.

Following the protocol established for wild type *Dd*HydAB. Strep-tactin affinity chromatography was used to purify the protein prior to artificial maturation of each variant inside the glovebox using 5-fold excess of $[2Fe]^{adt}$. The variants were left 48–50 hours for CO release followed by detailed characterisation via infrared spectroscopy, activity solution assay and protein film electrochemistry (PFE). SDS-PAGE analysis (Figure 6.17) showed that both the A48F and A48Y variants were expressed successfully and the isolated proteins were very pure, with only two bands corresponding to the large HydA subunit at 44 kDa and the small HydB subunit at 12 kDa. UV-vis spectra of both apo-A48F and apo-A48Y (not shown) were identical to apo-WT (Figure 6.6), indicating that the Fe-S clusters correctly assembled and inserted in both these variants.

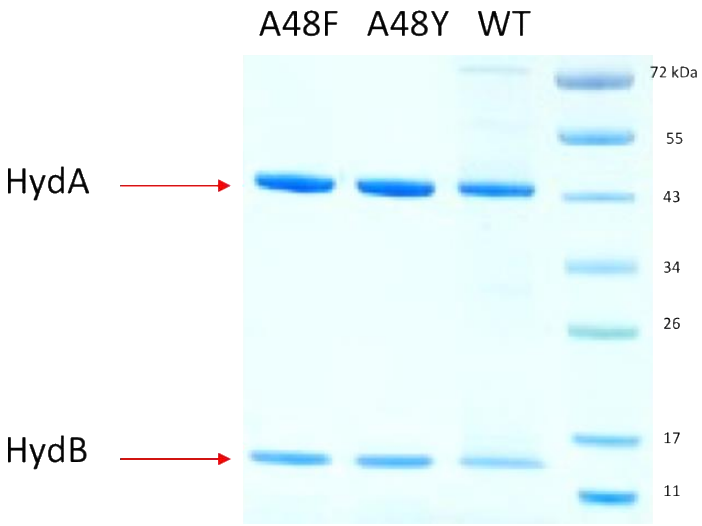


Figure 6.17 The SDS-PAGE analysis of two native variants amino acid A48F, A48Y and wild type apo-*Dd*HydAB [FeFe] hydrogenase after the strep-tactin affinity chromatography. Red arrows indicate the desired bands for the protein at 44 kDa (large subunit HydA) and 12 kDa (small subunit HydB).

Artificial maturation was carried out for these two native variants A48F and A48Y, by following the procedures mentioned in section 6.2. Figure 6.18 shows the FTIR spectra of the artificially maturated A48Y and A48F holo-*Dd*HydAB under oxidising conditions after CO-release (no H_2 or reductant). The spectra are essentially identical to the wild type *Dd*HydAB (Figure 6.7B), showing mainly the H_{ox} state (most oxidised states) with a small proportion of the H_{ox}-CO state (see Table 6.1 for *Dd*HydAB FTIR bands). This seems to indicate that the mutations A48F and A48Y do not disturb the active site or the incorporation of the $[2Fe]^{adt}$ sub-cluster.

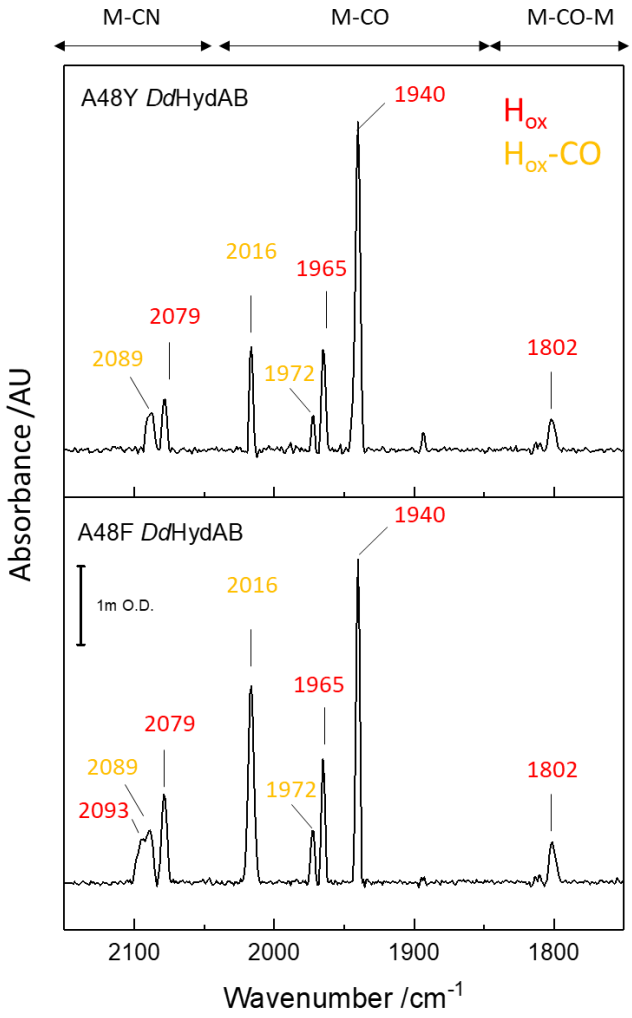


Figure 6.18 FTIR spectrum of A: 'as prepared' holo-A48Y *Dd*HydAB in 25 mM Tris-HCl, 20 mM NaCl pH 8. B: 'as prepared' holo-A48Y *Dd*HydAB in 25 mM Tris-HCl, 20 mM NaCl pH 8. Protein concentration: 2 mM. Peak values corresponding to the H_{ox}, H_{ox}-CO states are coloured in red and yellow respectively. 2 cm^{-1} resolution, 1024 scans.

Next, the activity of the artificially maturated A48Y and A48F *Dd*HydAB mutants was measured by solution activity assays. The catalytic H_2 oxidation activity was assessed by following by UV-Vis the reduction of 1 mM benzyl viologen by about 0.1 nM of enzyme in a H_2 saturated buffer (same procedure as for wild type *Dd*HydAB) as shown in Figure 6.19.[196] The activity of A48Y and A48F *Dd*HydAB in H_2 oxidation was 39,000 s^{-1} ± 8000 s^{-1} (35700 ± 7000 U/mg) and 32,000 ± 10,000 s^{-1} (29,300 ± 3500 U/mg) respectively, compared with 51,000 s^{-1} ± 9500 s^{-1} (47,000 U/mg ± 8500 U/mg) for wild type *Dd*HydAB. The H_2 oxidation activity of the variants is slightly lower

compared with the wild type enzyme. These suggests that there could have been some perturbation to the catalytic properties, perhaps a slight modification to the redox potential of the proximal cluster.

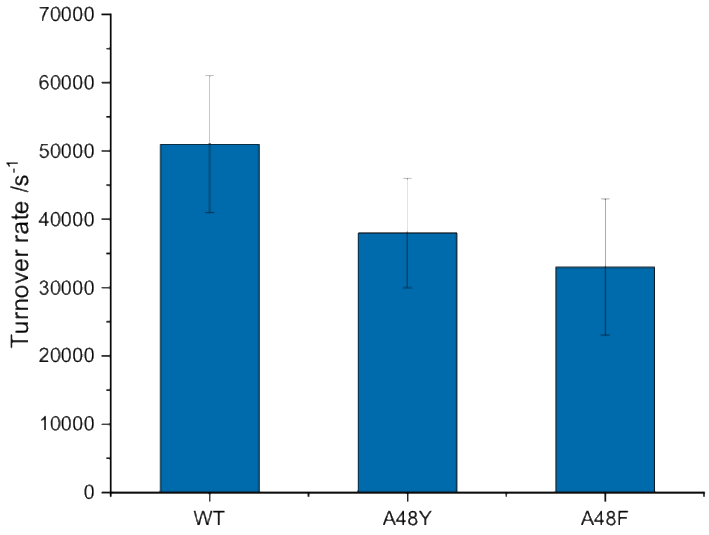


Figure 6.19 Comparison of the H_2 oxidisation activity of wild type *Dd*HydAB and the variants A48Y and A48F variants. 6 repetitions on each protein and the thin vertical line corresponds to the calculated error, the error bar represents ± 1 standard deviation.

To further investigate the catalytic behaviour of the variants and whether the mutations have affected the catalytic bias, protein film electrochemistry was carried out on A48Y and A48F *Dd*HydAB. As it can be seen in Figure 6.20, the cyclic voltammograms (CVs) of A48Y (A) and A48F (B) show that these variants are active for both H_2 oxidation and proton reduction.[107,140] The main difference regarding the catalytic behaviour of the variants compared with wild type seems to be that the variants suffer less high-potential inactivation. At this stage, it is difficult to find an explanation for this observation, and further experiments are needed to dig into this finding.

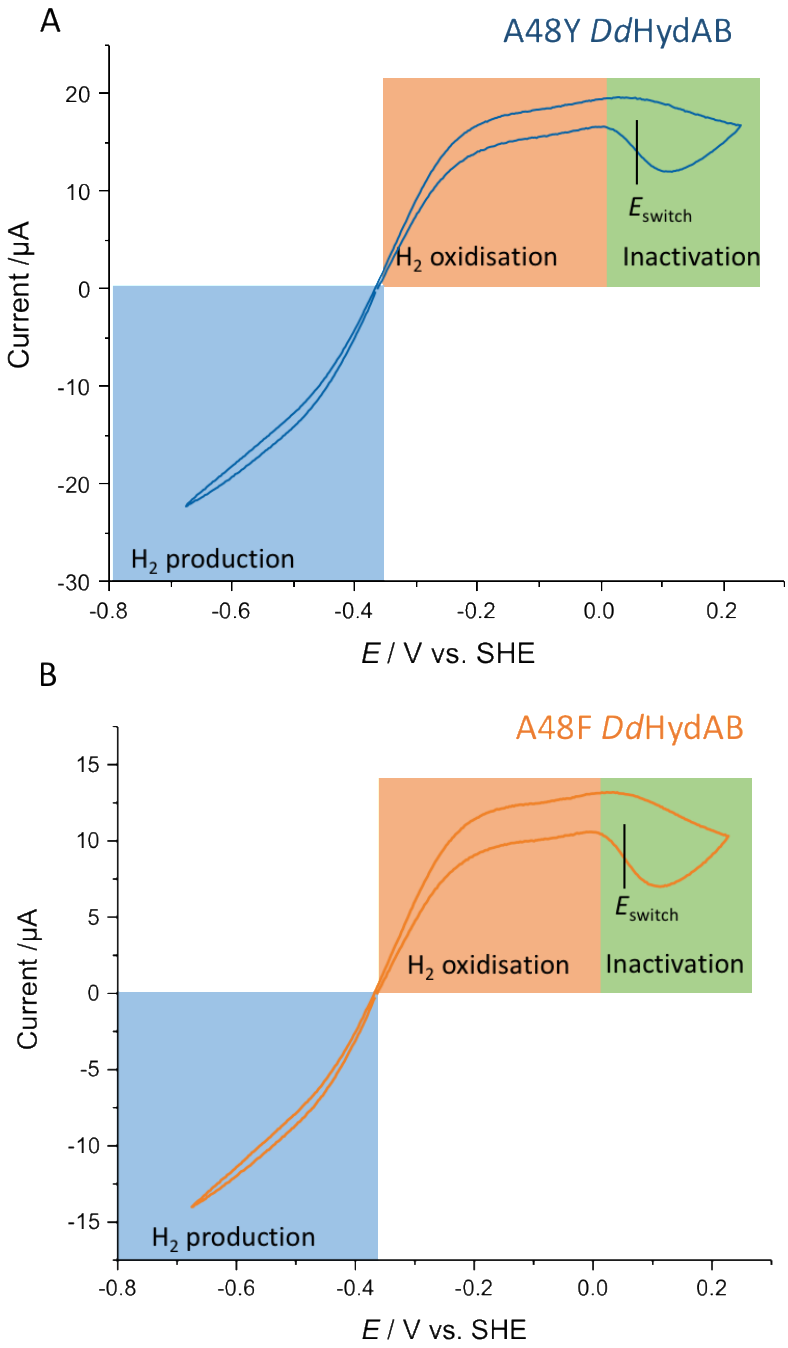


Figure 6.20 A: Protein film electrochemistry of A48Y holo-*Dd*HydAB adsorbed onto a pyrolytic graphite electrode. B: Protein film electrochemistry of A48F holo-*Dd*HydAB adsorbed onto a pyrolytic graphite electrode. Cyclic voltammetry was performed under 1 atm H_2, at 25°C in a 10 mM MES/HEPES, 100 mM NaCl, pH 6 buffer at 20 mV s^{-1}, and with electrode rotation at 2000 rpm. The regions of the voltammogram corresponding to H_2 production, H_2 oxidation, and high-potential inactivation (HPI) are indicated in blue, orange and green boxes, respectively.

It can be concluded that the mutations A48F and A48Y were successful. The two variants are nicely over-expressed in *E. coli* and the protein folded correctly which indicates that there is enough space around the residue alanine 48 to accommodate a bulkier amino acid. All this seems to indicate that this position may be suitable for the incorporation of the unnatural amino acid *p*CNF. However, because of time limitation, it was not possible to perform further mutagenesis to introduce *p*CNF have not been done in this research. In the conclusion section, it will be discussed how the research can be taken further towards to *p*CNF incorporation.

6.6 Discussion and conclusion

6.6.1 Overexpression and characterisation of *Dd*HydAB

The primary aim of this chapter was to establish an heterologous over-expression system in *E. coli* for the [FeFe] hydrogenase from *Desulfovibrio desulfuricans* (*Dd*HydAB) (based on the procedure reported by Birrell *et al*. 2016), which produces high yield of very pure protein, and to test the feasibility of mutagenesis for *p*CNF incorporation.[140] The heterodimeric wild type apo-hydrogenase, which is an inactive protein containing only the Fe-S clusters, was first produced in *E. coli* BL21 (DE3)Δ*iscR* cells and purified via strep-tactin affinity chromatography, achieving a yield of 17 mg per litre of culture of very pure protein. The apo-*Dd*HydAB was then artificially maturated by incubation with the synthetic cluster mimic $[Fe_2(adt)(CO)_4(CN)_2]^{2-}$ under strict anaerobic conditions and in the dark.

Characterisation via protein film electrochemistry and FTIR spectroscopy showed that the spectroscopic and electrochemical features of artificially maturated holo-*Dd*HydAB are essentially the same as the enzyme isolated from the native organism, as previously reported.[140] The turnover frequency for H_2 oxidation (51,000 s^{-1}) was in line with previous reports. This is an important step for further studies on the interesting *Dd*HydAB system.

After *Dd*HydAB was over-expressed and purified, two amino acid residues (phenylalanine 27 and alanine 48) were proposed as possible candidates to incorporate the unnatural amino acid *p*CNF close to the proximal FeS cluster.

6.6.2 Expression and characterisation of F27CNF *Dd*HydAB

In place of F27 (8.3 Å to the proximal FeS cluster), *p*CNF was incorporated directly, as the structure of phenylalanine is very similar to *p*CNF. The plasmid pULTRA-CNF was co-transformed into the *E. coli.* BL21 (DE3)Δ*iscR* cells in response to a reassigned amber codon by an aminoacyl-tRNA synthetase (aaRS)/tRNA pair that is orthogonal to the endogenous translational machinery of the host. The heterologous over-expression procedure in *E. coli* for the unnatural amino acid variant was slightly modified by adding 2 mM *p*CNF to the main culture before bubbling argon to make it anaerobic. In the spinach ferredoxin system, it was found that higher concentration of *p*CNF was able to improve the yield while in the *Dd*HydAB enzyme, higher concentrations of *p*CNF could not be dissolved in the cell culture and this resulted in the protein not expressing at all. This is because the pH decreases during *Dd*HydAB enzyme expression and *p*CNF can only be dissolved in a basic environment. After several attempts, it was found that 2 mM *p*CNF added to the main culture prior to induction of protein expression yielded the best result. The same purification procedure (strep-tactin affinity chromatography) as wild type was employed for the F27CNF *Dd*HydAB variants. The yield of the unnatural amino acid variant F27CNF was lower than the wild type (30% compared with WT).

Before artificially maturating the apo-protein with the $[2Fe]^{adt}$ sub-cluster, FTIR spectroscopy was carried out on apo-F27CNF. A clear band at 2233.9 cm^{-1} was observed for the as-isolated apo-F27CNF (oxidised). This band is different to the ν_{CN} peak (2237.0 cm^{-1}) in oxidised/as-isolated spinach ferredoxin but it is still in the nitrile stretching region. This suggests a different electrostatic environment around the residue in the hydrogenase. After reducing the protein, the peak red-shifted to 2232.8 cm^{-1}, thereby giving a distinguishable difference in the vibrational

frequency of the oxidised and reduced forms of apo-F27CNF *Dd*HydAB, which can be tentatively associated with the oxidation state of the proximal [4Fe-4S] cluster, as it was also seen in the spinach ferredoxin case, but the magnitude of the shift on reduction in apo-F27CNF ($\Delta\nu_{CN}$ = 1.1 cm^{-1}) is smaller than in the spinach ferredoxin ($\Delta\nu_{CN}$ = 1.9 cm^{-1}).

Apo-F27CNF was then maturated with $[2Fe]^{adt}$. Solution activity assays for H_2 oxidation showed that turnover frequency of holo-F27CNF *Dd*HydAB for H_2 oxidation was about one third of the wild type enzyme. On the other hand, protein film electrochemistry (PFE) showed F27CNF *Dd*HydAB to be highly active. This may indicate that *p*CNF might affect the $[2Fe]^{adt}$ insertion channel so that only a fraction of the protein is maturated, as the activity assay is very sensitive to the concentration of active holo-enzyme, while PFE only displays the activity of active enzyme and does not really depend on the concentration of active holo-enzyme. However, due to lack of time, more maturation experiments to investigate cofactor insertion and more repetitions of activity assays were not possible. Thus, it cannot be concluded with certainty whether this variant is completely maturated or the low activity in the solution assays is due to other factors. Further experiments are required to understand the reasons.

FTIR spectroscopy on holo-F27CNF showed that the IR bands from the active site (in both oxidised and reduced states) are essentially identical to those of the WT, which illustrates that incorporating *p*CNF did not disturb the catalytic active site. Focusing on the nitrile stretching region in Figure 6.15C, there was a small shift between the oxidised and reduced forms, which seems to indicate that *p*CNF is acting as an IR probe responding to the oxidation state of the FeS cluster relay.

6.6.3 Testing the tolerance of A48 to mutagenesis close to the proximal FeS cluster

The alanine 48 residue, which is about 6.5 Å to the proximal FeS cluster and far away from the other two clusters (more than 20 Å) was one of the selected residues. As the structure of alanine

is very different to *p*CNF, native amino acid site-directed mutagenesis was firstly carried out to test the tolerance of this site to mutagenesis, especially involving bulky aromatic amino acid residues. Two native amino acid variants, A48F and A48Y, were successfully generated and showed very similar characteristics with wild type *Dd*HydAB in FTIR spectroscopy, protein film electrochemistry and studies, although activity was slightly lower. This seems to indicate that there is enough space around the alanine 48 for incorporating a bulky amino acid such as *p*CNF and therefore this residue could be a potential candidate for the incorporation of the unnatural *p*CNF amino acid.

However, in this thesis project, due to the COVID-19 pandemic, there was no time for trying to incorporate *p*CNF into position 48, but it has been demonstrated that it is a suitable residue for mutating to a bulky amino acid like *p*CNF.

6.6.4 Conclusion

In this chapter, a novel method to study the redox states of FeS clusters in complex enzyme is developed by incorporating an unnatural amino acid *p*CNF into the [FeFe] hydrogenase (*Dd*HydAB) from *Desulfovibrio desulfuricans*. An over-expression system and purification protocol for *Dd*HydAB were established. This same protocol was also used for expression and purification of different variants generated by site-directed mutagenesis. *p*CNF was then successfully incorporated into the protein and gave the variant F27CNF. From the preliminary IR results, there is a clear detectable difference in the band position of the cyano group in F27CNF between the oxidised and reduced states on both apo- and holo-enzyme. The smaller shift in the holo-enzyme is likely to be due to incomplete reduction when the enzyme is catalytically active and can reduce protons which would re-oxidise centres in the protein. The result indicates that *p*CNF could be used as an IR probe to detect the redox states of FeS clusters in hydrogenases. More experiments such as protein film IR electrochemistry can be carried out to further study

the mechanism on how the FeS cluster work together with the active site during catalytic turnover.

7

Chapter 7: Conclusion and Outlook

7.1 Conclusion

Iron-sulfur (FeS) clusters are essential cofactors in all types of life with the function as redox centers in many significant reactions in life like mitochondrial respiration. This project has consisted of developing an approach to probe FeS cluster redox states by incorporating an unnatural amino acid which allows the study of iron-sulfur clusters via mid-IR spectroscopy. The ultimate goal is to study how FeS clusters interplay with the active site during catalysis by complex metalloenzymes like hydrogenases.

7.1.1 Incorporation of cyano-phenylalanine (*p*CNF) into the [2Fe-2S] cluster Spinach Ferredoxin I

Chapter 3 focused on selecting a FeS containing model protein to test this approach. Spinach ferredoxin I (FdI) is the most abundant form of photosynthetic-type ferredoxin present in spinach chloroplasts, and contains only one [2Fe-2S] cluster.[20] it was selected as a model system to investigate whether it is possible to incorporate the non-natural amino acid *p*CNF and use the cyano group as an IR probe to detect the redox states of the FeS cluster.

The overexpression system and purification protocol of FdI have been established first. This same protocol was also used for expression and purification of different variants generated by site-directed mutagenesis. *p*CNF is selectively incorporated into the protein in response to a reassigned nonsense codon (amber stop codon) by an aminoacyl-tRNA synthetase (aaRS)/tRNA pair that is orthogonal to the endogenous translational machinery of the host.[131,135] *p*CNF then could be successfully incorporated into the protein at Y37 site (7.4 Å from the FeS cluster) and gave the Y37CNF variant. The integrity and correct functioning of the [2Fe-2S] cluster upon *p*CNF incorporation was confirmed by UV-Vis spectroscopy and electrochemistry. From the preliminary IR result in this chapter, it was illustrated that *p*CNF in FdI can show an IR signal which looks promising for viability of the method.

In chapter 4, based on the two different modes of infrared spectroscopy (reflection-mode and transmission mode), it is clear to detect the difference in the IR band position of the nitrile group in Y37CNF between the oxidised and reduced states. The consistent results from two different modes of IR spectroscopy, and from chemical vs electrochemical reduction, provides support for assigning the shift in ν_{CN} to the effect of cluster redox state.

Additionally, detailed crystallographic characterisation was performed to solve and compare the structures of the oxidised and reduced *p*CNF-containing ferredoxin. According to the crystal structures of wild type and variants, it shows that the FeS cluster environment of variants are similar to the wild type which is consistent with the result from UV-vis spectra and electrochemistry. As such, we conclude that with the model protein spinach ferredoxin, it could be verified that the IR shift does not arise from a structural change but rather from the reduction of the [2Fe-2S] cluster and unnatural amino acid *p*CNF can be used as an IR probe to detect the redox states of the FeS cluster by means of Infrared spectroscopy.

7.1.2 Characterisation of *E.coli* [NiFe] hydrogenase 1 overexpressed in native host towards the incorporation of an unnatural amino acid *p*CNF

Since the concept has been proven with the FdI, experiments were planned to test whether this approach could be applied to complex metalloenzymes. One target is *E. coli* [NiFe] hydrogenase 1 (Hyd-1), as the genes to express Hyd-1 are already known and the overexpression system has just been established by Dr Miguel Ramirez (University of Oxford).[91] essential In chapter 5, characterisation experiments were carried out to detect if the overexpressed Hyd-1 behaves the same as the natively chromosomally encoded Hyd-1. Protein film electrochemistry, traditional solution-based activity assays and infrared spectroscopic methods were used to analyse the kinetic and thermodynamic properties of the enzymes. Protein film electrochemistry also demonstrated that the O_2 attack had a similar effect on the overexpressed Hyd-1 as on the natively expressed enzyme and that, the enzyme is still able to oxidise H_2 in the presence of O_2.

Then, site-directed mutagenesis was performed to select amino acid residues for incorporating an unnatural amino acid *p*CNF. Five amino acids were chosen with a distance between 5-10 Å to either the proximal or middle cluster in *E. coli* Hyd-1. Three variants did not fold correctly, as the mutation appeared to cause in strong clash with neighbouring residues. The other three variants F24Y, V121F and A123F expressed and folded correctly. Additionally, protein film electrochemistry was used to test the catalytic properties of each variant. F24Y and A123F retained catalytic behaviour similar to the wild type, indicating that these two residues could be suitable for incorporating *p*CNF.

Further experiments on incorporation of *p*CNF were put on hold because the overexpressed protein seemed to have some unusual IR active site spectra features in the protein film infrared electrochemistry and did not crystallise in preliminary trials. Also, the moderate level of overexpression, about 10 X native expression, would make this enzyme difficult to obtain in the *p*CNF variants in sufficient quantity for spectroscopic study.

7.1.3 Incorporating an unnatural amino acid into [FeFe] hydrogenases from *D. desulfuricans*

In chapter 6, *p*CNF was incorporated into the [FeFe] hydrogenase from the sulfate reducing bacterium *Desulfovibrio desulfuricans* (*Dd*HydAB). This enzyme contains two additional [4Fe-4S] clusters in addition to the active site H-cluster. An overexpression system and purification protocol for *Dd*HydAB were followed, according to previous published work and characterisation via protein film electrochemistry and FTIR spectroscopy showed that the spectroscopic and electrochemical features of holo-*Dd*HydAB are essentially the same as the enzyme isolated from the native organism and the enzyme artificially maturated as previously reported.[140]

Then *p*CNF was incorporated into a selected position, replacing phenylalanine (F) 27 close to the proximal 4Fe-4S cluster, to give the variant F27CNF *Dd*HydAB. Before artificially maturating the apo-protein (inactive enzyme without catalytic site) with $[2Fe]^{adt}$ cluster mimic, FTIR

spectroscopy was carried out on apo-F27CNF. It can be seen that there is a small but observable shift in the peak position between the oxidised and reduced F27CNF samples (2233.9 cm^{-1} in the oxidised form vs. 2232.8 cm^{-1} in the reduced form). Thus, the C≡N stretching frequency differs by about 1.1 cm^{-1} in the oxidised and reduced forms. After that, apo-F27CNF was maturated with $[2Fe]^{adt}$ and gave the holo-F27CNF. The lower turnover frequency of holo-F27CNF *Dd*HydAB for H_2 oxidisation compared to the wild type enzyme indicates that the *p*CNF might block the way of the ADT insertion channel so that only a portion of enzyme maturated.

FTIR spectroscopy on holo-F27CNF shows the IR bands from the active site are identical to the wild type, which demonstrated that *p*CNF incorporation did not alter the electronic characteristics of the catalytic active site. Focusing on the cyano stretching region, the shift was still clear between the oxidised and reduced form, about 0.5 cm^{-1}, which can be associated with a partial change in the oxidation state of the proximal [4Fe-4S] cluster. The smaller shift in the holo-F27CNF is probably due to part of active enzymes being re-oxidised by proton reduction. The preliminary result indicates that *p*CNF could be used as an IR probe to detect the redox states of FeS clusters in *Dd*HydAB.

Overall, the work presented in this thesis demonstrated a novel approach for applying an unnatural amino acid *p*CNF to metalloenzymes which contain FeS clusters to study the redox state of the FeS cluster by IR spectroscopy. This approach has been proven with a model FeS protein spinach ferredoxin and then extended to the [FeFe] hydrogenase from *Desulfovibrio desulfuricans* (*Dd*HydAB), thus exploiting the versatility of this method across a range of proteins.

7.2 Outlook

To continue the research done in this thesis, detailed studies on the unnatural variants of the [FeFe] hydrogenase *Dd*HydAB will contribute new understanding of the FeS cluster relay. By

using the home-built reflection IR-SEC cell, apo- (lacking the binuclear site) and holo- (containing the full active site H-cluster) could be studied under electrochemical control, allowing the sample to be oxidised and reduced using electrochemistry while simultaneously measuring IR *in situ*. In this way, electrochemical reduction/oxidation of the unnatural apo- and holo- variants can be directly compared with chemical oxidation/reduction done in this thesis, to see if they yield the same results. Afterwards, protein film IR electrochemistry (PFIRE) can be applied on holo-F27CNF *Dd*HydAB to investigate how the proximal [4Fe-4S] clusters work together with the catalytic active site during catalytic turnover.

More unnatural amino acid variants should be studied such as A48CNF, as the residue alanine 48 (about 6.5 Å to the proximal FeS cluster) has been tested by site-directed mutagenesis, and it has been demonstrated that it is a suitable residue for mutating to a bulky amino acid like *p*CNF. Therefore, *p*CNF can be incorporated into this residue as well, to investigate how the distance between *p*CNF and FeS cluster influence the IR response.

Furthermore, the unnatural amino acid approach can be extended to the overexpressed *E. coli* Hyd-1 system studied in this thesis. As the special [3Fe-4S] proximal cluster in Hyd-1 has been suggested to be crucial for the O_2-tolerance of this enzyme, it is thought to give 2 electrons to the active site in a narrow potential window, thus supplying quickly enough the electrons required to reduce O_2 to water.[58] Therefore, it would be specially interesting to incorporate *p*CNF into Hyd-1 close to the proximal FeS cluster and investigate how this special [3Fe-4S] proximal cluster works together with the active site during catalysis by using protein film IR electrochemistry (PFIRE). Longer term goal would be to create variants with unnatural amino acids near each of the FeS clusters in turn, so that the electron-loading of each cluster during catalysis or O_2 response could be studied.

Approaches are emerging for incorporation of more than one unnatural amino acid into a single protein, so it may even be possible to use other IR-active unnatural amino acids such as azido-phenylalanine in tandem with *p*CNF to probe two clusters in hydrogenase simultaneously.

Bibliography

1 T. A. Rouault and W. H. Tong, *Nat. Rev. Mol. Cell Biol.*, 2005, **6**, 345–351.

2 E. M. Shepard and J. B. Broderick, in *Comprehensive Natural Products II: Chemistry and Biology*, Elsevier Ltd, 2010, vol. 8, pp. 625–661.

3 J. Przybyla-Toscano, M. Roland, F. Gaymard, J. Couturier and N. Rouhier, *J. Biol. Inorg. Chem.*, 2018, **23**, 545–566.

4 M. Fontecave, *Nat. Chem. Biol.*, 2006, **2**, 171–174.

5 B. E. Schultz and S. I. Chan, *Annu. Rev. Biophys. Biomol. Struct.*, 2001, **30**, 23–65.

6 C. C. Page, C. C. Moser, X. Chen and P. L. Dutton, *Nature*, 1999, **402**, 47–52.

7 J. L. Breton, J. L. C. Duff, J. N. Butt, F. A. Armstrong, S. J. George, Y. Pétillot, E. Forest, G. Schäfer and A. J. Thomson, *Eur. J. Biochem.*, 1995, **233**, 937–946.

8 F. A. Armstrong, H. A. Heering and J. Hirst, *Chem. Soc. Rev.*, 1997, **26**, 169–179.

9 W. R. Hagen, *J. Biol. Inorg. Chem.*, 2018, **23**, 623–634.

10 W. Pandelia, Maria Nitschke, P. Infossi, E. Giudici-Orticoni and W. Lubitz, *Proc. Natl. Acad. Sci.*, 2011, **108**, 6097–6102.

11 M. M. Roessler, M. S. King, A. J. Robinson, F. A. Armstrong, J. Harmer and J. Hirst, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 1930–1935.

12 I. M. Ibrahim, H. Wu, R. Ezhov, G. E. Kayanja, S. D. Zakharov, Y. Du, W. A. Tao, Y. Pushkar, W. A. Cramer and S. Puthiyaveetil, *Commun. Biol.*, 2020, **3**, 1–11.

13 D. C. Johnson, D. R. Dean, A. D. Smith and M. K. Johnson, *Annu. Rev. Biochem.*, 2005, 74, 247–281.

14 F. B. Silvia Pagani, Giuseppe Vecchio, *Biochem. Biophys. Acta*, 1986, **870**, 538–544.

15 H. Huang, L. Hu, W. Yu, H. Li, F. Tao, H. Xie and S. Wang, *Protein Expr. Purif.*, 2016, **121**, 1–8.

16 R. Z. Zhao, S. Jiang, L. Zhang and Z. Bin Yu, *Int. J. Mol. Med.*, 2019, **44**, 3–15.

17 G. C. Ferreira, R. Franco, S. G. Lloyd, I. Moura, J. J. G. Moura and B. H. Huynh, *J. Bioenerg. Biomembr.*, 1995, **27**, 221–229.

18 E. Balconi, A. Pennati, D. Crobu, V. Pandini, R. Cerutti, G. Zanetti and A. Aliverti, *FEBS J.*, 2009, **276**, 3825–3836.

19 W. Lubitz, H. Ogata, O. Ru and E. Reijerse, *Chem. Rev.*, 2014, **114**, 4081–4148.

20 L. Piubelli, A. Aliverti, F. Bellintani and G. Zanetti, *Protein Expr. Purif.*, 1995, **6**, 298–304.

21 S. S. Mansy and J. A. Cowan, *Acc. Chem. Res.*, 2004, **37**, 719–725.

22 B. Blanc, C. Gerez and S. Ollagnier De Choudens, *Biochim. Biophys. Acta*, 2015, **1853**, 1436–1447.

23 B. Srour, S. Gervason, B. Monfort and B. Dautreaux, *Inorganics*, 2020, **8**, 1–38.

24 C. J. Schwartz, J. L. Giel, T. Patschkowski, C. Luther, F. J. Ruzicka, H. Beinert and P. J. Kiley, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 14895–14900.

25 J. C. L Mortenson, R Valentine, *Biochem. Biophys. Res. Commun.*, 1962, **7**, 448–452.

26 J. Liu, S. Chakraborty, P. Hosseinzadeh, Y. Yu, S. Tian, I. Petrik, A. Bhagi and Y. Lu, *Chem. Rev.*, 2014, **114**, 4366–4369.

27 G. N. Hutber, A. J. Smith and L. J. Rogers, *FEMS Microbiol. Lett.*, 1978, **4**, 11–14.

28 H. Matsubara, R. M. Sasaki and R. K. Chain, *J. Biol. Chem.*, 1968, **243**, 1725–1731.

29 D. J. Crispin, G. Street and J. E. Varey, *Food Chem.*, 2001, **72**, 355–362.

30 M. Droux, M. Jacquot, P. Gadal, J. Huet, N. Crawford and B. Yee, *Arch. Biochem. Biophys.*, 1987, **252**, 426–439.

31 T. R. Cook, D. K. Dogutan, S. Y. Reece, Y. Surendranath, T. S. Teets and D. G. Nocera, *Chem. Rev.*, 2010, **110**, 6474–6502.

32 B. Widera, *Therm. Sci. Eng. Prog.*, 2020, **16**, 100460.

33 S. Millet, P.; Grigoriev, in *Water Electrolysis Technologies*, Eds. Elsevier: Amsterdam, 2013.

34 G. A. Olah, A. Goeppert and G. K. S. Prakash, in *Beyond Oil and Gas: The Methanol Economy: Second Edition*, Wiley-VCH, 2009, pp. 1–334.

35 K. A. Vincent, A. Parkin and F. A. Armstrong, *Chem.Rev.*, 2007, **107**, 4366–4413.

36 H. Nishihara, Y. Miyashita, K. Aoyama, T. Kodama, Y. Igarashi and Y. Takamura, *Biochem. Biophys. Res. Commun.*, 1997, **232**, 766–770.

37 E. Schwartz, J. Fritsch and B. Friedrich, in *The Prokaryotes: Prokaryotic Physiology and Biochemistry*, Springer-Verlag Berlin Heidelberg, 2013, pp. 119–199.

38 P. M. Vignais and B. Billoud, *Chem.Rev.*, 2007, **107**, 4206–4272.

39 M. Adams, E. Eccleston and J. B. Howard, *Proc. Natl. Acad. Sci. U. S. A.*, 1989, **86**, 4932–4936.

40 W. Adams, S. Lal Menon and J M Verhagen, *J. Bacteriol.*, 2001, **183**, 716–724.

41 J. D. Fox, R. L. Kerby, G. P. Roberts and P. W. Ludden, *J. Bacteriol.*, 1996, **178**, 1515–1524.

42 B. M. Hoffman, D. Lukoyanov, D. R. Dean and L. C. Seefeldt, *Acc. Chem. Res.*, 2012, **46**,

587–595.

43 U. Deppenmeier, *Biochem. Biophys. Res. Commun.*, 2004, **36**, 55–64.

44 R. Hedderich, *Biochem. Biophys. Res. Commun.*, 2004, **36**, 65–75.

45 C. Greening, A. Biswas, C. R. Carere, C. J. Jackson, M. C. Taylor, M. B. Stott, G. M. Cook and S. E. Morales, *ISME J.*, 2015, **10**, 761–777.

46 T. Burgdorf, O. Lenz, T. Buhrke, E. Van Der Linden, A. K. Jones, S. P. J. Albracht and B. Friedrich, *J Mol Microbiol Biotechnol*, 2005, **10**, 181–196.

47 A. K. Jones, E. Sillery, S. P. J. Albracht and F. A. Armstrong, *Chem. Commun*, 2002, **56**, 866–867.

48 F. A. Armstrong, N. A. Belsey, J. A. Cracknell, G. Goldet, A. Parkin, E. Reisner, K. A. Vincent and A. F. Wait, *chem. Soc. Rev.*, 2009, **38**, 36–51.

49 H. A. Reeve, L. Lauterbach, O. Lenz and K. A. Vincent, *ChemCatChem*, 2015, **7**, 3480–3487.

50 R. K. Thauer, A.-K. Kaster, M. Goenrich, M. Schick, T. Hiromoto and S. Shima, *Annu. Rev. Biophys. Biomol. Struct.*, 2010, **79**, 507–536.

51 M. Korbas, S. Vogt, W. Meyer-Klaucke, E. Bill, E. J. Lyon, R. K. Thauer and S. Shima, *J. Biol. Chem.*, 2006, **281**, 30804–30813.

52 P. M. Vignais, B. Billoud and J. Meyer, *FEMS Microbiogy Rev.*, 2001, **25**, 455–501.

53 L. Casalot and M. Rousset, *TRENDS Microbiol.*, 2001, **9**, 228–237.

54 A. Furstner, *ACS Cent. Sci.*, 2016, **2**, 778–789.

55 A. Volbeda, M. H. Charon, C. Piras, E. C. Hatchikian, M. Frey and J. C. Fontecilla-Camps, *Nature*, 1995, **373**, 580–587.

56 R. P. Happe, W. Roseboom, A. J. Plerlk, S. P. J. Albracht and K. A. Bagley, *Nature*, 1997, **385**, 126.

57 K. A. Vincent, A. Parkin and F. A. Armstrong, *Chem.Rev.*, 2007, **107**, 4366–4413.

58 M. Ludwig, J. A. Cracknell, K. A. Vincent, F. A. Armstrong and O. Lenz, *J. Biol. Chem.*, 2009, **284**, 465–477.

59 M. J. Lukey, A. Parkin, M. M. Roessler, B. J. Murphy, J. Harmer, T. Palmer, F. Sargent and F. A. Armstrong, *J. Biol. Chem.*, 2010, **285**, 3928–3938.

60 A. Contributions, *Nat Chem Biol*, 2013, **9**, 607–609.

61 J. Fritsch, P. Scheerer, S. Frielingsdorf, S. Kroschinsky, B. Friedrich, O. Lenz and C. M. T. Spahn, *Nature*, 2011, **479**, 249–253.

62 S. Frielingsdorf, J. Fritsch, A. Schmidt, M. Hammer, Y. Rippers, F. Lendzian, I. Zebger, C. Teutloff, M. Kaupp, R. Bittl, P. Hildebrandt, B. Friedrich, O. Lenz and P. Scheerer, *Nat. Chem. Biol.*, 2014, **10**, 378–385.

63 T. Goris, A. F. Wait, M. Saggu, J. Fritsch, N. Heidary, M. Stein, I. Zebger, F. Lendzian, F. A. Armstrong, B. Friedrich and O. Lenz, *Nat. Chem. Biol.*, 2011, **7**, 310–318.

64 O. Lenz, M. Ludwig, T. Schubert, I. Burstel, S. Ganskow, T. Goris, A. Schwarze and B. Friedrich, *ChemPhysChem*, 2010, **11**, 1107–1119.

65 Y. Shomura, H. Yoon and Y. Higuchi, *Nature*, 2011, **479**, 253–257.

66 H. Ogata, W. Lubitz and Y. Higuchi, *J. Biochem*, 2016, **160**, 251–258.

67 P. Wulff, C. C. Day, F. Sargent and F. A. Armstrong, *Proc. Natl. Acad. Sci.*, 2014, **111**, 6606–6611.

68 M. M. Roessler, R. M. Evans, R. A. Davies, J. Harmer and F. A. Armstrong, *J. Am. Chem.*

Soc., 2012, **134**, 15581–15594.

69 H. Adamson, M. Robinson, J. J. Wright, L. A. Flanagan, J. Walton, D. Elton, D. J. Gavaghan, A. M. Bond, M. M. Roessler and A. Parkin, *J. Am. Chem. Soc*, 2017, **139**, 10677–10686.

70 M. Asso, B. Guigliarelli, T. Yagi and P. Bertrand, *Biochim. Biophys. Acta (BBA)/Protein Struct. Mol.*, 1992, **1122**, 50–56.

71 J. A. Cracknell, A. F. Wait, O. Lenz, B. Rbel Friedrich and F. A. Armstrong, *Proc. Natl. Acad. Sci.*, 2009, **106**, 20681–20686.

72 A. Volbeda, P. Amara, C. Darnault, J.-M. Mouesca, A. Parkin, M. M. Roessler, F. A. Armstrong and J. C. Fontecilla-Camps, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 5305–5310.

73 M. Yael, V. Anne, V. Xavier and H. Claude, *Nat. strucutral Biol.*, 1997, **4**, 523–526.

74 I. Fdez. Galvan, A. Volbeda, J. C. Fontecilla-Camps and M. J. Field, *Proteins Struct. Funct. Bioinforma.*, 2008, **73**, 195–203.

75 E. Szori-Doroghazi, G. Maroti, M. Szori, A. Nyilasi, G. Rakhely and K. L. Kovacs, *PLoS One*, 2012, **7**, 34666.

76 O. Duche, S. Elsen, L. Cournac and A. Colbeau, *FEBS J.*, 2005, **272**, 3899–3908.

77 T. Buhrke, O. Lenz, N. Krauss and B. Friedrich, *J. Biol. Chem.*, 2005, **280**, 23791–23796.

78 A. Volbeda, A. E. Lydie, M. Ae, C. C. Michaei, M. Ae, B. W. Faber, A. E. Winfried, R. Simon, P. J. Albracht, A. E. Elsa, G. Ae and J. C. Rousset, *J. Biol. Inorg. Chem.*, 2005, **10**, 239–249.

79 S. Kurkin, S. J. George, R. N. F. Thorneley and S. P. J. Albracht, *Biochemistry*, 2004, **43**, 6820–6831.

80 F. A. Armstrong, R. M. Evans, S. V Hexter, B. J. Murphy, M. M. Roessler and P. Wulff, *Acc.*

Chem. Res., 2016, **49**, 884–892.

81 H. Ogata, S. Hirota, A. Nakahara, H. Komori, N. Shibata, T. Kato, K. Kano and Y. Higuchi, *Structure*, 2005, **13**, 1635–1642.

82 A. Volbeda, L. Martin, E. Barbier, O. Gutierrez-Sanz, A. L. De Lacey, P. Liebgott, S. Dementin, M. Rousset and J. Fontecilla-Camps, *J. Biol. Inorg. Chem.*, 2015, **20**, 11–22.

83 P. A. Ash, R. Hidalgo and K. A. Vincent, *ACS Catal.*, 2017, **7**, 2471–2485.

84 R. Breglia, C. Greco, P. Fantucci, L. De Gioia and M. Bruschi, *Phys. Chem. Chem. Phys*, 2018, **20**, 1693.

85 B. J. Murphy, R. Hidalgo, M. M. Roessler, R. M. Evans, P. A. Ash, W. K. Myers, K. A. Vincent and F. A. Armstrong, *J. Am. Chem. Soc.*, 2015, **137**, 8484–8489.

86 M. Brecht, M. van Gastel, T. Buhrke, B. rbel Friedrich and W. Lubitz, *J. Am. Chem. Soc.*, 2003, **125**, 13075–13083.

87 A. Dubini, R. L. Pye, R. L. Jack, T. Palmer and F. Sargent, *Int. J. Hydrogen Energy*, 2002, **27**, 1413–1420.

88 G. Sawers, *Antonie Van Leeuwenhoek*, 1994, **66**, 57.

89 J. S. Mcdowall, B. J. Murphy, M. Haumann, T. Palmer, F. A. Armstrong and F. Sargent, *Proc. Natl. Acad. Sci.*, 2014, **111**, E3948–E3956.

90 M. J. Lukey, A. Parkin, M. M. Roessler, B. J. Murphy, J. Harmer, T. Palmer, F. Sargent and F. A. Armstrong, *J. Biol. Chem.*, 2010, **285**, 3928–3938.

91 F. Sargent, *Adv. Microb. Physiol.*, 2016, **68**, 433–507.

92 Y. Nicolet, C. Piras, P. Legrand, C. E. Hatchikian and J. C. Fontecilla-Camps, *Structure*, 1999, **7**, 13–23.

93 D. W. Mulder, E. S. Boyd, R. Sarma, R. K. Lange, J. A. Endrizzi, J. B. Broderick and J. W. Peters, *Nature*, 2010, **465**, 248–252.

94 M. Winkler, J. Esselborn and T. Happe, *Biochim. Biophys. Acta - Bioenerg.*, 2013, **1827**, 974–985.

95 A. S. Pandey, T. V Harris, L. J. Giles, J. W. Peters and R. K. Szilagyi, *J. Am. Chem. Soc.*, 2008, **130**, 4533–4540.

96 M. Mccullagh and G. A. Voth, *J. Phys. Chem. B*, 2013, **117**, 4062–4071.

97 D. W. Mulder, M. W. Ratzloff, M. Bruschi, C. Greco, E. Koonce, J. W. Peters and P. W. King, *J. Am. Chem. Soc*, 2014, **136**, 15394–15402.

98 G. Hong, A. J. Cornish, E. L. Hegg and R. Pachter, *Biochim. Biophys. Acta - Bioenerg.*, 2011, **1807**, 510–517.

99 Y. Mu, H. Liang, J. Hu, L. Jiang and L. Wan, *J.Phys.Chem.B*, 2005, **109**, 22212–22216.

100 A. J. Cornish, K. Gartner, H. Yang, J. W. Peters and E. L. Hegg, *J. Biol. Chem.*, 2011, **286**, 38341–38347.

101 J. Cohen, K. Kim, P. King, M. Seibert and K. Schulten, *Structure*, 2005, **13**, 1321–1329.

102 H. Gongyi and R. Pachter, *ACS Chem. Biol.*, 2012, **7**, 1268–1275.

103 A. Adamska-Venkatesh, D. Krawietz, J. Siebel, K. Weber, T. Happe, E. Reijerse and W. Lubitz, *J. Am. Chem. Soc.* , 2014, **136**, 11339–11346.

104 P. Rodriguez-Macia, N. Breuer, S. DeBeer and J. A. Birrell, *ACS Catal.*, 2021, **15**, 13084–13095.

105 E. J. Reijerse, C. C. Pham, V. Pelmenschikov, R. Gilbert-Wilson, A. Adamska-Venkatesh, J. F. Siebel, L. B. Gee, Y. Yoda, K. Tamasaku, W. Lubitz, T. B. Rauchfuss and S. P. Cramer, *J.*

Am. Chem. Soc, 2017, **139**, 4306–4309.

106 C. Sommer, A. Adamska-Venkatesh, K. Pawlak, J. A. Birrell, O. Ru, E. J. Reijerse and W. Lubitz, *J. Am. Chem. Soc*, 2017, **139**, 1440–1443.

107 A. Adamska, A. Silakov, C. Lambertz, O. Rudiger, T. Happe, E. Reijerse and W. Lubitz, *Angew. Chemie - Int. Ed.*, 2012, **51**, 11458–11462.

108 P. Rodriguez-Macia, E. Reijerse, Macia, W. Lubitz, J. A. Birrell and O. Ru, *J. Phys. Chem. Lett*, 2017, **8**, 49.

109 P. A. Ash, S. B. Carr, bc H. A Reeve, A. Skorupskaite, J. S. Rowbotham, R. Shutt, M. D. Frogley, R. M. Evans, G. Cinque, F. A. Armstrong and K. A. Vincent, *Chem. Commun*, 2017, **53**, 5858.

110 A. Barth, *Biochim. Biophys. Acta - Bioenerg.*, 2007, **1767**, 1073–1101.

111 Y. Nicolet, A. L. De Lacey, X. Verne, V. M. Fernandez, E. C. Hatchikian and J. C. Fontecilla-Camps, *J. Am. Chem. Soc.*, 2001, **123**, 1596–1601.

112 D. J. Arp and R. H. Burris, *Biochemistry*, 1981, **20**, 2234–2240.

113 F. A. Armstrong, R. M. Evans, S. V Hexter, B. J. Murphy, M. M. Roessler and P. Wulff, *Acc. Chem. Res.*, 2016, **49**, 884–892.

114 P. A. Ash, S. E. T. Kendall-Price and K. A. Vincent, *Acc. Chem. Res*, 2019, **52**, 3120–3131.

115 P. A. Ash, R. Hidalgo and K. A. Vincent, *J. Vis. Exp*, 2017, **130**, 55858.

116 R. Hidalgo, P. A. Ash, A. J. Healy and K. A. Vincent, *Angew. Chemie Int. Ed.*, 2015, **54**, 7110–7113.

117 A. L. De Lacey, V. M. Fernandez, M. Rousset and R. Cammack, *Chem. Rev.*, 2007, **107**, 4304–4330.

118 S. H. Bell, D. P. E. Dickson, R. Rieder, R. Cammack, D. S. Patil, D. O. Hall and K. K. Rao, *Eur. J. Biochem.*, 1984, **145**, 645–651.

119 S. E. J. Bowman, J. Bridwell-Rabb and C. L. Drennan, *Acc. Chem. Res.*, 2016, **49**, 695–702.

120 N. Maurici, N. Savidge, B. U. Lee, S. H. Brewer and C. M. Phillips-Piro, *Acta Crystallogr. Sect. D Struct. Biol.*, 2018, **74**, 650–655.

121 X. S. Gai, B. A. Coutifaris, S. H. Brewer and E. E. Fenlon, *Phys. Chem. Chem. Phys.*, 2011, **13**, 5926.

122 C. G. Bazewicz, J. S. Lipkin, E. E. Smith, M. T. Liskov and S. H. Brewer, *J. Phys. Chem. B*, 2012, **116**, 10824–10831.

123 A. T. Fafarman, L. J. Webb, J. I. Chuang and S. G. Boxer, *J Am Chem Soc*, 2006, **128**, 13356–13357.

124 J. Ma, I. M. Pazos, W. Zhang, R. M. Culik and F. Gai, *Annu. Rev. Phys. Chem*, 2015, **66**, 357–377.

125 S. Yoshikawa, D. H. Keeffe and W. S. Caughey, *J. Biol. Chem.*, 1985, **260**, 3518–3528.

126 M. C. Thielges, J. Y. Axup, D. Wong, H. S. Lee, J. K. Chung, P. G. Schultz and M. D. Fayer, *J Phys Chem B*, 2011, **115**, 11294–11304.

127 J. Voller, H. Biava, B. Koksch, P. Hildebrandt and N. Budisa, *ChemBioChem*, 2015, **16**, 742–745.

128 A. B. Dippel, G. M. Olenginski, N. Maurici, M. T. Liskov, S. H. Brewer and C. M. Phillips-Piro, *Res. Pap. Acta Cryst*, 2016, **72**, 121–130.

129 A. Creon, I. Josts, S. Niebling, N. Huse and H. Tidow, *Struct. Dyn.*, 2018, **5**, 064701.

130 C. Kearney, L. T. Olenginski, T. D. Hirn, G. D. Fowler, D. Tariq, S. H. Brewer and C. M.

Phillips-Piro, *RSC Adv.*, 2018, **8**, 13503–13512.

131 T. S. Young, I. Ahmad, J. A. Yin and P. G. Schultz, *J. Mol. Biol.*, 2010, **395**, 361–374.

132 J. Xie and P. G. Schultz, *Nat. Rev. Mol. Cell Biol.*, 2006, 7, 775–782.

133 S. Benzer and S. Champe, *J. Am. Chem. Soc*, 1959, **88**, 562.

134 E. Freese, J. Champe, S. Benzer ; Garen and S. Benzer, *Proc. Natl. Acad. Sci.*, 1962, **48**, 1149.

135 K. C. Schultz, L. Supekova, Y. Ryu, J. Xie, R. Perera and P. G. Schultz, *J. Am. Chem. Soc.*, 2006, **128**, 13984–13985.

136 J. Zimmermann, M. C. Thielges, Y. J. Seo, P. E. Dawson and F. E. Romesberg, *Angew. Chemie - Int. Ed.*, 2011, **50**, 8333–8337.

137 J. Zimmermann, M. C. Thielges, Y. J. Seo, P. E. Dawson and F. E. Romesberg, *Angew. Chemie - Int. Ed.*, 2011, **50**, 8333–8337.

138 M. K. Akhtar and P. R. Jones, *Appl. Microbiol. Biotechnol.*, 2008, **78**, 853–862.

139 L. Wang, D. Fan, W. Chen and E. M. Terentjev, *Sci. Rep.*, 2015, **5**, 1–11.

140 J. A. Birrell, K. Wrede, K. Pawlak, P. Rodriguez-Macia, O. Rudiger, E. J. Reijerse and W. Lubitz, *Isr. J. Chem.*, 2016, **56**, 852–863.

141 A. K. Jones, S. E. Lamle, H. R. Pershad, K. A. Vincent, S. P. J. Albracht and F. A. Armstrong, *J. Am. Chem. Soc.*, 2003, **125**, 8505–8514.

142 K. Nakamoto, in *Handbook of Vibrational Spectroscopy*, ed. P. R. Griffiths, John Wiley & Sons, Ltd, Chichester, UK, 2006.

143 J. R. Reimers and L. E. Hall, *J. Am. Chem. Soc.*, 1999, **121**, 3730–3744.

144 R. A. Nyquist, *Appl. Spectrosc.*, 1990, **44**, 1405–1407.

145 S. Gervason, D. Larkem, A. Ben Mansour, T. Botzanowski, C. S. Muller and B. D'Autreaux, *Nat. Commun.*, 2019, **10**, 11470–11479.

146 W. W. Chen, M. Niepel and P. K. Sorger, *Genes Dev.*, 2010, 24, 1861–1875.

147 H. Lineweaver and Burk, *J. Am. Chem. Soc.*, 1934, **56**, 658–666.

148 D. E. Lide, *CRC Handbook of Chemistry and Physics* , CRC Press, 1997.

149 Y. Jouanneau, C. Meyer, J. Gaillard and P. M. Vignais, *Biochem. Biophys. Res. Commun.*, 1990, **171**, 273–279.

150 J. Armengaud, C. Meyer and Y. Jouanneau, *Biochem. J.*, 1994, **300**, 413–8.

151 C. G. Bazewicz, J. S. Lipkin, K. A. Lozinak, M. D. Watson and S. H. Brewer, *Tetrahedron Lett.*, 2011, **52**, 6865–6868.

152 U. Tagawa and D. 1 Arnon, *Biochim. Biophys. Acta*, 1968, **153**, 602–613.

153 Y. Lu, F. Qiao, Y. Li, X. H. Sang, C. R. Li, J. D. Jiang, X. Y. Yang and X. F. You, *Appl. Microbiol. Biotechnol.*, 2017, **101**, 7201–7212.

154 M. F. J. M. Verhagen, T. A. Link and W. R. Hagen, *FEBS Lett.*, 1995, **361**, 75–78.

155 N. N. Elgrishi, K. J. Rountree, B. D. Mccarthy, E. S. Rountree, T. T. Eisenhart and J. L. Dempsey, *J. Chem. Educ.*, 2017, **95**, 197–206.

156 C. D. Crawley and F. M. Hawkridge, *Biochem. Biophys. Res. Commun.*, 1981, **99**, 516–522.

157 C. Binda, A. Coda, A. Aliverti, G. Zanetti and A. Mattevi, *Int. Union Crystallogr. Acta Crystallogr. Sect. D Acta Cryst*, 1998, **54**, 1353–1358.

158 A. Chatterjee, S. B. Sun, J. L. Furman, H. Xiao and P. G. Schultz, *Biochemistry*, 2013, **52**,

1828–1837.

159 H. Biava, T. Schreiber, S. Katz, S. Vo, M. Stolarski, C. Schulz, N. Michael, N. Budisa, J. Kozuch, T. Utesch and P. Hildebrandt, *J. Phys. Chem. B*, 2018, **122**, 8330–8342.

160 J. V. Frangioni and B. G. Neel, *Anal. Biochem.*, 1993, **210**, 179–187.

161 J. G. Huismanf, A. F. M. Moormanx and F. N. Verkley, *Biochem. Biophys. Res. Commun.*, 1978, **82**, 1121–1131.

162 J. Rawlings, O. Siiman and H. B. Gray, *PNAS*, 2019, **71**, 125–127.

163 P. Bianco and J. Haladjian, *Electrochim. Acta*, 1997, **42**, 587–594.

164 M. M. Waegele, R. M. Culik and F. Gai, *J. Phys. Chem. Lett*, 2011, **2**, 2598–2609.

165 S. Ye, T. Huber, R. Vogel and T. Sakmar, *Nat Chem Biol*, 2009, **5**, 397–399.

166 C. M. Pradier, M. Salmain and S. Boujday, in *Biointerface Characterization by Advanced IR Spectroscopy*, Elsevier, 2011, pp. 167–216.

167 B. A. Lindquist, K. E. Furse and S. A. Corcelli, *Phys. Chem. Chem. Phys.*, 2009, **11**, 8119.

168 E. A. Franzosa, K. J. Lynagh and Y. Xia, *Proc. 4th Int. Beilstein Symp. Exp. Stand. Cond. Enzym. Charact.*, 2010, 99–106.

169 H. Ohno, K. Takeda, S. Niwa, T. Tsujinaka, Y. Hanazono, Y. Hirano and K. Miki, *PLoS One*, 2017, **12**, 1–13.

170 R. M. Evans, A. Parkin, M. M. Roessler, B. J. Murphy, H. Adamson, M. J. Lukey, F. Sargent, A. Volbeda, J. C. Fontecilla-Camps and F. A. Armstrong, *J. Am. Chem. Soc.*, 2013, **135**, 2694–2707.

171 C. Hunte, V. Zickermann and U. Brandt, *Science (80-.).*, 2010, **329**, 448–451.

172 H. M. Deckers, F. R. Wilson and G. Voordouw, *J. Gen. Microbiol.*, **136**, 2021–2028.

173 L. Brondsted, *J. Bacteriol.*, 1994, **176**, 5423–5428.

174 J. Y. Kim, B. Hoon Jo and H. Joon Cha, *Microb. Cell Fact.*

175 S. Reissmann, E. Hochleitner, H. Wang, A. Paschos, F. Lottspeich, R. S. Glass and A. Bock, *Science (80-.).*, 2003, **299**, 1067–1070.

176 F. Sargent, *Adv. Microb. Physiol.*, 2016, **68**, 433–507.

177 S. T. Stripp, U. Lindenstrauss, R. G. Sawers and B. Soboh, *PLoS One*, 2015, **10**, 1–12.

178 J. Fritsch, O. Lenz and B. Friedrich, *Nat. Rev. Microbiol.*, 2013, **11**, 106–114.

179 J. Fritsch, O. Sanganas, E. Siebert, I. Zebger, M. Stein, M. Ludwig, A. L. De Lacey, H. Dau, A. Friedrich, O. Lenz and M. Haumann, *Biochemistry*, 2011, **50**, 5858–5869.

180 R. M. Evans, P. A. Ash, S. E. Beaton, E. J. Brooke, K. A. Vincent, S. B. Carr and F. A. Armstrong, *J. Am. Chem. Soc.*, 2018, **140**, 10208–10220.

181 M. Rousset, Y. Montet, B. Guigliarelli, N. Forget, M. Asso, P. Bertrand, J. C. Fontecilla-Camps and E. Claude Hatchikian, *Proc. Natl. Acad. Sci.*, 1998, **95**, 11625–11630.

182 R. Bingemann and A. Klein, *Eur. J. Biochem.*, 2000, **267**, 6612–6618.

183 U. Breuer, *Eng. Life Sci.*, 2006, **6**, 603–603.

184 E. J. Brooke, R. M. Evans, S. T. A. Islam, G. M. Roberts, S. A. M. Wehlin, S. B. Carr, S. E. V Phillips and F. A. Armstrong, *Biochemistry*, 2017, **56**, 132–142.

185 P. Wardman, *Free Radic. Res.*, 1991, **14**, 57–67.

186 V. Fourmond, P. Infossi, M.-T. se Giudici-Orticoni, P. Bertrand and C. Le ger, *J. Am. Chem. Soc.*, 2010, **132**, 4848–4857.

187 M. Del Barrio, C. Guendon, A. Kpebe, C. Baffert, V. Fourmond, M. Brugna and C. Leger, *ACS Catal.*, 2019, **9**, 4084–4088.

188 A. Parkin, L. Bowman, M. M. Roessler, R. A. Davies, T. Palmer, F. A. Armstrong and F. Sargent, *FEBS Lett.*, 2012, **586**, 536–544.

189 M. M. Roessler, R. M. Evans, R. A. Davies, J. Harmer and F. A. Armstrong, *J. Am. Chem. Soc.*, 2012, **134**, 15581–15594.

190 C. Leger, A. K. Jones, S. P. J. Albracht and F. A. Armstrong, *J Phys Chem B*, 2002, **106**, 13058–13063.

191 P. Rodriguez, M. Macia, K. Pawlak, O. Ru, E. J. Reijerse, W. Lubitz and J. A. Birrell, *J. Am. Chem. Soc*, 2017, **139**, 15122–15134.

192 J. C. Fontecilla-Camps, A. Volbeda, C. Cavazza and Y. Nicolet, *chem. Soc. Rev.*, 2007, **107**, 4273–4303.

193 R. Britt, G. Rao and L. Tao, *Chem. Sci.*, 2020, **11**, 10313–10323.

194 E. M. Shepard, F. Mus, J. N. Betz, A. S. Byer, B. R. Duffus, J. W. Peters and J. B. Broderick, *Biochemistry*, 2014, **53**, 36.

195 G. Berggren, A. Adamska, C. Lambertz, T. R. Simmons, J. Esselborn, M. Atta, S. Gambarelli, J. M. Mouesca, E. Reijerse, W. Lubitz, T. Happe, V. Artero and M. Fontecave, *Nature*, 2013, **499**, 66–69.

196 J. Esselborn, C. Lambertz, A. Adamska-Venkatesh, T. Simmons, G. Berggren, J. Noth, J. Siebel, A. Hemschemeier, V. Artero, E. Reijerse, M. Fontecave, W. Lubitz and T. Happe, *Nat. Chem. Biol.*, 2013, **9**, 607–609.

197 C. F. Megarity, J. Esselborn, S. V Hexter, F. Wittkamp, T. Happe and F. A. Armstrong, *J.*

Am. Chem. Soc, 2016, **138**, 2021.

198 P. Rodriguez-Macia, J. A. Birrell, W. Lubitz and O. Rudiger, *Chempluschem*, 2017, **82**, 540–545.

199 A. J. Pierik, M. Hulstein, W. R. Hagen and S. P. J. Albracht, *Eur. J. Biochem.*, 1998, **258**, 572–578.

200 W. Roseboom, A. L. De Lacey, V. M. Fernandez, E. C. Hatchikian and S. P. J. Albracht, *J. Biol. Inorg. Chem.*, 2006, **11**, 102–118.

201 A. L. De Lacey, C. Stadler, C. Cavazza, E. Claude Hatchikian and V. M. Fernandez, *J. Am. Chem. Soc*, 2000, **122**, 11232–11233.

202 H. J. Grande, A. Berkel-Arts, J. Bregh, K. Dijk and C. Veeger, *Eur. J. Biochem.*, 1983, **131**, 81–88.

203 H. M. van der Westen, S. G. Mayhew and C. Veeger, *FEBS Lett.*, 1978, **86**, 122–126.

204 K. A. Vincent, A. Parkin, O. Lenz, J. C. Albracht, R. Cammack, B. Friedrich and F. A. Armstrong, *J. Am. Chem. Soc.*, 2005, **127**, 18179–18189.

205 N. Mccoy and P. Delahay, *J. Am. Chem. Soc*, 1936, **88**, 1577.

206 G. Winter, M. Waterman, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter and G. Evans, *Acta Crystallogr. Sect. D Struct. Biol.*, 2018, **74**, 85–97.

207 J. Beilsten-Edmands, G. Winter, R. Gildea, J. Parkhurst, D. Waterman and G. Evans, *Acta Crystallogr. Sect. D Struct. Biol.*, 2020, **76**, 385–399.

208 C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack and G. Bricogne, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2011, **67**, 293–302.

209 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J.

Read, *J. Appl. Crystallogr.*, 2007, **40**, 658–674.

210 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 2010, **66**, 486–501.

211 G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 1997, **53**, 240–255.

212 D. Liebschner, P. V. Afonine, M. L. Baker, C. J. Bunkoczi and P. D. Adams, *Acta Crystallogr. Sect. D Struct. Biol.*, 2019, **75**, 861–877.

213 C. J. Williams, V. B. Headd, S. Jain, S. M. Lewis, W. B. Arendall, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson and D. C. Richardson, *Protein Sci.*, 2018, **27**, 293–315.

214 C. Binda, A. Coda, A. Aliverti, G. Zanetti and A. Mattevi, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 1998, **54**, 1353–1358.

Appendix A

Spinach ferredoxin

Name	Sequence (5’ to 3’)	Description
EcoRI_FdI_F	cacacagaattcattaaagaggagaaattaactatgg	Amplification of the FdI gene to insert in pGEX vector digested with EcoRI at 5' end
XhoI_FdI_R	cggccgctcgagttatgcggtcagttcctcttct	Reverse amplification of the FdI gene to insert in pGEX vector digested with XhoI at 3' end
FdI_Y3F_F	cccggaattcgcagcaTTCaaagttaccctggtta	Site-directed mutagenesis for FdI at tyrosine 3 to phenylalanine forward primer
FdI_Y3F_R	taaccagggtaactttGAAtgctgcgaattccggg	Site-directed mutagenesis for FdI at tyrosine 3 to phenylalanine reverse primer
FdI_Y3W_F	cccggaattcgcagcaTGGaaagttaccctggtta	Site-directed mutagenesis for FdI at tyrosine 3 to tryptophan forward primer
FdI_Y3W_R	taaccagggtaactttCCAtgctgcgaattccggg	Site-directed mutagenesis for FdI at tyrosine 3 to tryptophan reverse primer
FdI_Y3A_F	cccggaattcgcagcaGCAaaagttaccctggtta	Site-directed mutagenesis for FdI at tyrosine 3 to alanine forward primer
FdI_Y3A_R	taaccagggtaactttTGCtgctgcgaattccggg	Site-directed mutagenesis for FdI at tyrosine 3 to alanine reverse primer
FdI_Y23F_F	tgtccggatgatgtgTTCattctggatgcagcaga	Site-directed mutagenesis for FdI at tyrosine 23 to phenylalanine forward primer
FdI_Y23F_R	tctgctgcatccagaatGAAcacatcatccggaca	Site-directed mutagenesis for FdI at tyrosine 23 to phenylalanine reverse primer
FdI_Y23W_F	tgtccggatgatgtgTGGattctggatgcagcaga	Site-directed mutagenesis for FdI at tyrosine 23 to tryptophan forward primer
FdI_Y23W_R	tctgctgcatccagaatCCAcacatcatccggaca	Site-directed mutagenesis for FdI at tyrosine 23 to tryptophan reverse primer
FdI_Y23A_F	tgtccggatgatgtgGCGattctggatgcagcaga	Site-directed mutagenesis for FdI at tyrosine 23 to alanine forward primer
FdI_Y23A_R	tctgctgcatccagaatCGCcacatcatccggaca	Site-directed mutagenesis for FdI at tyrosine 23 to alanine reverse primer
FdI_Y37F_F	ggtattgatctgccgTTTagctgtcgtgcaggt	Site-directed mutagenesis for FdI at tyrosine 37 to phenylalanine forward primer
FdI_Y37F_R	acctgcacgacagctAAAcggcagatcaatacc	Site-directed mutagenesis for FdI at tyrosine 37

		to phenylalanine reverse primer
FdI_Y37W_F	ggtattgatctgccgTGGagctgtcgtgcaggt	Site-directed mutagenesis for FdI at tyrosine 37 to tryptophan forward primer
FdI_Y37W_R	acctgcacgacagctCCAcggcagatcaatacc	Site-directed mutagenesis for FdI at tyrosine 37 to tryptophan reverse primer
FdI_Y37A_F	ggtattgatctgccgGCGagctgtcgtgcaggt	Site-directed mutagenesis for FdI at tyrosine 37 to alanine forward primer
FdI_Y37A_R	acctgcacgacagctCGCcggcagatcaatacc	Site-directed mutagenesis for FdI at tyrosine 37 to alanine reverse primer
FdI_Y37amber_F	ggtattgatctgccgTAGagctgtcgtgcaggt	Site-directed mutagenesis for FdI at tyrosine 37 to amber stop codon forward primer
FdI_Y37amber_R	acctgcacgacagctCTAcggcagatcaatacc	Site-directed mutagenesis for FdI at tyrosine 37 to amber stop codon reverse primer
FdI_F63Y_F	caggatgatcagagcTATctggatgatgaccag	Site-directed mutagenesis for FdI at phenylalanine 63 to tyrosine forward primer
FdI_F63Y_R	ctggtcatcatccagATAgctctgatcatcctg	Site-directed mutagenesis for FdI at phenylalanine 63 to tyrosine reverse primer
FdI_F63W_F	caggatgatcagagcTGGctggatgatgaccag	Site-directed mutagenesis for FdI at phenylalanine 63 to tryptophan forward primer
FdI_F63W_R	ctggtcatcatccagCCAgctctgatcatcctg	Site-directed mutagenesis for FdI at phenylalanine 63 to tryptophan reverse primer
FdI_F63A_F	caggatgatcagagcGCGctggatgatgaccag	Site-directed mutagenesis for FdI at phenylalanine 63 to alanine forward primer
FdI_F63A_R	ctggtcatcatccagCGCgctctgatcatcctg	Site-directed mutagenesis for FdI at phenylalanine 63 to alanine reverse primer
FdI_F63amber_F	caggatgatcagagcTAGctggatgatgaccag	Site-directed mutagenesis for FdI at phenylalanine 63 to amber stop codon forward primer
FdI_F63amber_R	ctggtcatcatccagCTAgctctgatcatcctg	Site-directed mutagenesis for FdI at phenylalanine 63 to amber stop codon reverse primer
FdI_Y80F_F	tctgacctgtgcagccTTCccggttagtgatgtta	Site-directed mutagenesis for FdI at tyrosine 80 to phenylalanine forward primer
FdI_Y80F_R	taacatcactaaccggGAAggctgcacaggtcaga	Site-directed mutagenesis for FdI at tyrosine 80 to phenylalanine reverse primer
FdI_Y80W_F	tctgacctgtgcagccTGGccggttagtgatgtta	Site-directed mutagenesis for FdI at tyrosine 80 to tryptophan forward primer

FdI_Y80W_R	taacatcactaaccggCCAggctgcacaggtcaga	Site-directed mutagenesis for FdI at tyrosine 80 to tryptophan reverse primer
Y80A_F	tctgacctgtgcagccGCGccggttagtgatgtta	Site-directed mutagenesis for FdI at tyrosine 80 to alanine forward primer
Y80A_R	taacatcactaaccggCGCggctgcacaggtcaga	Site-directed mutagenesis for FdI at tyrosine 37 to alanine reverse primer

***E.coli* Hyd-1**

Name	Sequence (5’ to 3’)	Description
Hyd-1 F24Y_F	tgctgtaccgaatctTACatccgctccgctcac	Site-directed mutagenesis for Hyd-1 at phenylalanine 24 to tyrosine forward primer
Hyd-1 F24Y_R	gtgagcggagcggatGTAagattcggtacagca	Site-directed mutagenesis for Hyd-1 at phenylalanine 24 to tyrosine reverse primer
Hyd-1 W112F_F	agcgcgattatcgccTTCggaacctgcgcgtcc	Site-directed mutagenesis for Hyd-1 at tryptophan 112 to phenylalanine forward primer
Hyd-1 W112F_R	ggacgcgcaggttccGAAggcgataatcgcgct	Site-directed mutagenesis for Hyd-1 at tryptophan 112 to phenylalanine reverse primer
Hyd-1 V121F_F	gcgtcctggggctgcTTTcaggccgcgcgaccc	Site-directed mutagenesis for Hyd-1 at valine 121 to phenylalanine forward primer
Hyd-1 V121F_R	gggtcgcgcggcctgAAAgcagccccaggacgc	Site-directed mutagenesis for Hyd-1 at valine 121 to phenylalanine reverse primer
Hyd-1 A123F_F	tggggctgcgtgcagTTCgcgcgacccaatccg	Site-directed mutagenesis for Hyd-1 at alanine 123 to phenylalanine forward primer
Hyd-1 A123F_R	cggattgggtcgcgcGAActgcacgcagcccca	Site-directed mutagenesis for Hyd-1 at alanine 123 to phenylalanine reverse primer
Hyd-1 Y264F_F	gatcgcggttcgttcTTCagccgcgtggtcgat	Site-directed mutagenesis for Hyd-1 at tyrosine 264 to phenylalanine forward primer
Hyd-1 Y264F_R	atcgaccacgcggctGAAgaacgaaccgcgatc	Site-directed mutagenesis for Hyd-1 at tyrosine 264 to phenylalanine reverse primer

***Dd*HydAB:**

Name	Sequence (5’ to 3’)	Description
*Dd*H F27amber_F	gatccggacaaactgcacTAGgtgcagattgatgaagcaa	Site-directed mutagenesis for *Dd*H at phenylalanine 27 to amber stop codon forward primer
*Dd*H F27amber_R	ttgcttcatcaatctgcacCTAgtgcagtttgtccggatc	Site-directed mutagenesis for *Dd*H at phenylalanine 27 to amber stop codon reverse primer
*Dd*H A48F_F	agtcaatactgcccgacgTTTgccattttcggtgaaatgg	Site-directed mutagenesis for *Dd*H at alanine 48 to phenylalanine forward primer
*Dd*H A48F_R	ccatttcaccgaaaatggcAAAcgtcgggcagtattgact	Site-directed mutagenesis for *Dd*H at alanine 48 to phenylalanine reverse primer
*Dd*H A48Y_F	agtcaatactgcccgacgTACgccattttcggtgaaatgg	Site-directed mutagenesis for *Dd*H at alanine 48 to tyrosine forward primer
*Dd*H A48Y_R	ccatttcaccgaaaatggcGTAcgtcgggcagtattgact	Site-directed mutagenesis for *Dd*H at alanine 48 to tyrosine reverse primer

* Caption letters shown in the sequence are the mutated base pairs in the site-directed mutagenesis.

Appendix B: Spinach Ferredoxin Crystallisation

Purified protein was concentrated to 35-40 mg mL^{-1} using an Amicon centrifugal filtration unit (Merck Millipore, Germany) with a molecular weight cutoff of 10 kDa. Crystallisation, using the sitting drop vapour diffusion technique, was performed by mixing 100 nL of protein solution with an equal volume of crystallisation solution (100 mM sodium, potassium phosphate pH 7-8, 2.6-3.6 M ammonium sulphate) using a mosquito crystallisation robot (SPT Labtech, UK) followed by incubation at 277 K. Crystals were transferred to cryoprotecting solution (100 mM sodium, potassium phosphate pH 7.5, 3.6 M ammonium sulphate, 5% v/v glycerol) for about 15 seconds then flash cooled by plunging into liquid nitrogen. Crystals of reduced ferredoxin were produced by transferring them to an anaerobic chamber (N_2 atmosphere, O_2 <0.3 ppm) followed by soaking in O_2-free cryoprotectant also containing 10 mM dithionite for several minutes prior to flash cooling.

X-ray diffraction

X-ray diffraction data were collected at beamlines I03 and I24 (Diamond Light Source, UK) with details summarised in Table and reduced using either DIALS or autoPROC.[206,207,208] Phase estimates were generated by Molecular Replacement using Phaser[209] and native spinach ferredoxin-I (pdb accession code 1A70) as a search model followed by iterative rounds of rebuilding in Coot and refinement with either Refmac or Phenix.[210,211,212] The diffraction data for Y37F FdI displays a significant level of pseudo-merohedral twinning, which was accounted for during structure refinement. All resultant models were validated using the Molprobity server.[213]

Table 1: X-ray data collection parameters

Spinach Ferredoxin variants	Native	Y37F	Y37CNF	Y3F	Y3CNF
Beamline	I03	I03	I03	I03	I24

Wavelength	0.77 Å	0.l62 Å	0.l62 Å	0.62 Å	0.97 Å
Detector	Eiger2 XE 16M	Eiger2 XE 16M	Eiger2 XE 16M	Eiger2 XE 16M	Pilatus3 6M
Resolution	0.92 Å	1.10 Å	1.20 Å	1.10 Å	1.17 Å
Data reduction software	Dials	Autoproc	Autoproc	Dials	Autoproc
Refinement software	Phenix_refine	Phenix_refine	Phenix_refine	Refmac5	Refmac5

Crystal forming

Crystals appeared overnight with some variants displaying differing morphology: wild type, Y3F and Y37CNF produced large rhombohedral crystals, whereas variant Y3CNF, containing the non-natural amino acid *p*CNF, produced clusters of thin rods. Variant Y37F produced a mixture of small rhombohedra and clusters of fine needles, with the former used for x-ray analysis. Although the crystallisation conditions were similar to those previously published the crystal symmetry and unit cell parameters varied significantly between each variant, however, crystal size and morphology had little impact on the diffraction quality of the crystals (Table).[214]

Table 2: X-ray data collection and refinement statistics for spinach ferredoxin variants

Spinach Ferredoxin variant	Native	Y37F	Y37CNF	Y3F	Y3CNF
Space group	$P2_12_12_1$	$P2_1$	$P2_12_12_1$	$P4_32_12$	$P6_1$
Unit cell dimensions (Å)	a=30.30, b=51.41, c=61.33	a=59.23, b=61.80, c=59.30, β=119.1	a=28.89, b=46.72, c=63.33	a=62.23, b=62.23, c=85.53	a=61.47, b=61.47, c=60
Resolution (Å)	39.4-0.92 (0.94-0.92)	51.8-1.10 (1.13-1.10)	37.63-1.20 (1.22-1.20)	39.13-1.10 (1.12-1.10)	53.24-1.17 (1.19-1.17)
Total reflections	835,290 (33,211)	1,081,068 (51,269)	533,096 (15,902)	1,776,223 (77,187)	413,624 (14,278)
Unique reflections	67,214 (3,289)	156,460 (7,783)	26,730 (1,168)	68,768 (3,330)	43,870 (2,193)
Completeness (%)	100 (100)	100 (99.9)	96.7 (87.7)	100 (100)	100 (100)
Multiplicity	12.4 (10.1)	6.9 (6.6)	20.7 (13.6)	25.8 (23.2)	9.4 (6.5)
$\langle I/\sigma \rangle$	18.2 (0.8)	11.4 (1.0)	18.6 (2.6)	20.2 (1.0)	5.9 (0.9)

R_{merge}	0.049 (1.19)	0.077 (2.0)	0.089 (0.998)	0.071 (1.85)	0.22 (2.23)
R_{pim}	0.014 (0.37)	0.03 (0.78)	0.026 (0.38)	0.02 (0.56)	0.078 (0.94)
$CC_{1/2}$	1.0 (0.8)	0.999 (0.3)	0.999 (0.825)	1.0 (0.596)	0.994 (0.40)
Refinement					
R_{work}/R_{free} (%)	12.6/13.6	13.9/15.9	15.3/18.2	13.5/15.2	16.3/19.5
No. of atoms	940	3386	890	943	867
Macromolecule	793	3111	820	793	775
Solvent	147	275	70	150	92
Average B-factors					
Macromolecule	16.9	14.7	13.3	18.4	13.2
Solvent	29.9	21.2	17.9	29.0	22.5
RMSD bond lengths (Å)	0.013	0.018	0.019	0.016	0.017
RMSD angles (°)	1.45	2.3	2.1	1.6	2.1
Ramachandran plot favoured / outliers	99/0	97.7/0	99/0	99/0	98.9/0
Clashscore (percentile)	1.95 (95^{th})	1.67 (97^{th})	1.82 (98^{th})	3.91 (84^{th})	1.34 (98^{th})
Molprobity score (percentile)	0.96 (98^{th})	1.01 (98^{th})	1.32 (89^{th})	1.21 (92^{nd})	0.92 (99^{th})