



HAL
open science

Chemical maturation of hydrogenases : an insight into artificial and biohybrid systems

Giorgio Caserta

► **To cite this version:**

Giorgio Caserta. Chemical maturation of hydrogenases : an insight into artificial and biohybrid systems. Biochemistry [q-bio.BM]. Université Pierre et Marie Curie - Paris VI, 2016. English. NNT : 2016PA066701 . tel-01631276

HAL Id: tel-01631276

<https://theses.hal.science/tel-01631276>

Submitted on 9 Nov 2017

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

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

Université Pierre et Marie Curie

Ecole doctorale de Chimie Moléculaire de Paris-Centre

Laboratoire de Chimie des Processus Biologiques

CNRS UMR 8229 Collège de France, Paris

Chemical maturation of hydrogenases: an insight into artificial and biohybrid systems

Présentée par

Giorgio CASERTA

Pour l'obtention du grade de Docteur de l'Université Pierre et Marie Curie

Spécialité

Biochimie et Chimie Bio-inorganique

Dirigée par le Professeur Marc Fontecave

Présentée et soutenue publiquement le 7 Novembre 2016

Devant un jury composé de:

Dr. Sandrine OLLAGNIER DE CHOUDENS, Directeur de Recherche CNRS	Rapporteur
Pr. Bruno GUIGLIARELLI, Professeur à Aix-Marseille Université	Rapporteur
Pr. Solange LAVIELLE, Professeur à UPMC	Examinatrice
Dr. Jean-Maurice MALLET, Directeur de Recherche CNRS	Examineur
Dr. Mohamed ATTA, Directeur de Recherche CEA Grenoble	Examineur
Pr. Marc FONTECAVE, Professeur au Collège de France	Directeur de thèse



Université Pierre et Marie Curie

Ecole doctorale de Chimie Moléculaire de Paris-Centre

Laboratoire de Chimie des Processus Biologiques

CNRS UMR 8229 Collège de France, Paris

Chemical maturation of hydrogenases: an insight into artificial and biohybrid systems

Présentée par

Giorgio CASERTA

Pour l'obtention du grade de Docteur de l'Université Pierre et Marie Curie

Spécialité

Biochimie et Chimie Bio-inorganique

Dirigée par le Professeur Marc Fontecave

Présentée et soutenue publiquement le 7 Novembre 2016

Devant un jury composé de:

Dr. Sandrine OLLAGNIER DE CHOUDENS, Directeur de Recherche CNRS	Rapporteur
Pr. Bruno GUIGLIARELLI, Professeur à Aix-Marseille Université	Rapporteur
Pr. Solange LAVIELLE, Professeur à UPMC	Examinatrice
Dr. Jean-Maurice MALLET, Directeur de Recherche CNRS	Examineur
Dr. Mohamed ATTA, Directeur de Recherche CEA Grenoble	Examineur
Pr. Marc FONTECAVE, Professeur au Collège de France	Directeur de thèse



Learn from yesterday,

live for today,

hope for tomorrow.

The important thing is not to stop questioning.

(A. Einstein)

Summary

Chemical maturation of hydrogenases: an insight into artificial and biohybrid systems

There is a general agreement that the building of a sustainable H₂ economy relies on the availability of cheap, abundant and efficient catalysts. Nature has provided attractive solutions, hydrogenases. However, these enzymes are difficult to produce and so far only few HydAs have been completely characterized showing diversity despite the same active site. This core, H-cluster, is composed of a [4Fe-4S] cluster bound via a Cys to a diiron complex which has 3 CO, 2 CN and an azadithiolate ligands. Recently, it has been showed that hydrogenases can be easily produced through the insertion of a biomimetic [Fe₂(adt)(CO)₄(CN)₂]²⁻ complex inside the heterologously produced apo-enzyme, resulting in a full activation. Part of the PhD has been focused on the chemical maturation of new HydA from *Megasphaera elsdenii* and its truncated version, MeH-HydA, containing only the H-cluster. The assembly of all metal cofactors via the chemical reconstitution of the [Fe-S] clusters and the maturation through the [Fe₂(adt)(CO)₄(CN)₂]²⁻ complex has been carried out. Interestingly, HydF hybrids synthesized incorporating biomimetic [Fe₂(xdt)(CO)₄(CN)₂]²⁻ complexes onto the [4Fe-4S] cluster HydF protein, have a 6Fe core reminiscent of the H-cluster. HydFs from different organisms were purified and subsequently the [4Fe-4S] cluster has been reconstituted. For the first time, an X-ray structure of HydF with its [4Fe-4S] cluster has been obtained. The 6Fe cluster of HydF has been also prepared chemically with diiron complexes mimicking the active site of HydA. The metallo-cofactors have been spectroscopically characterized (EPR, FTIR, HYSCORE), hydrogenase activities evaluated.

Keywords: [FeFe]-hydrogenase; artificial hydrogenase; chemical maturation; hydrogen evolution; iron-sulfur cluster; protein crystallography.

Résumé

Maturation chimique des hydrogénases: une étude des systèmes artificiels et biohybrides

Le développement d'une économie basée sur l'hydrogène implique l'utilisation de catalyseurs efficaces et peu chers. Pour cela on peut s'inspirer de la nature qui a produit des métalloenzymes, les hydrogénases.

On considère que la maturation est réalisée en deux étapes. Dans un premier temps, les clusters $[4\text{Fe}-4\text{S}]$ sont assemblés par les systèmes ISC/SUF. Puis trois maturases, HydE/F/G réalisent la biosynthèse du 2Fe sous-cluster pour synthétiser le cluster-H.

Seules quelques hydrogénases à $[\text{FeFe}]$ ont été caractérisées montrant une grande diversité alors qu'elles possèdent le même centre catalytique, le cluster-H. Ainsi, une partie de cette thèse a porté sur l'utilisation de la «maturation chimique» pour activer de nouvelles enzymes apo-HydA. L'hydrogénase provenant de *Megasphaera elsdenii*, et sa version tronquée, MeH-HydA, contenant seulement le domaine du cluster-H ont été étudiées grâce à cette technique.

La stratégie mise en œuvre a été la reconstitution des cofacteurs métalliques par l'assemblage des clusters $[4\text{Fe}-4\text{S}]$ et la maturation des enzymes par le complexe biomimétique $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$.

Les hybrides HydF synthétisés en incubant la protéine contenant le cluster $[4\text{Fe}-4\text{S}]$ avec les complexes biomimétiques $[\text{Fe}_2(\text{xdt})(\text{CO})_4(\text{CN})_2]^{2-}$ possèdent un centre à 6 fers similaire au cluster-H. Cette grande similarité amène au dernier point traité dans cette thèse: la possible activité catalytique d'HydF en tant que «hydrogénase artificielle». Les hybrides de HydF ont été caractérisés et leur activité d'hydrogénase a été évaluée. De plus, une structure RX de la protéine contenant son cluster $[4\text{Fe}-4\text{S}]$ a été obtenue.

Mots-clés : $[\text{FeFe}]$ -hydrogénases; hydrogénase artificielle; maturation chimique; production d'hydrogène; cluster fer-soufre; cristallographie des protéines.

Table of contents

Abbreviations and Acronyms	7
Statement	11
Chapter I: Introduction	13
Section A	13
1.1 Metalloprotein: metal ions in living systems	13
1.2 Iron-sulfur cluster proteins	13
1.3 Biosynthesis of Fe–S clusters	16
1.4 Hydrogenase enzymes	16
Section B	19
1.5 The metal center of [FeFe]-hydrogenases	19
1.5.1 A protein machinery for [FeFe]-hydrogenase maturation	20
1.5.2 The Hyd proteins	22
1.5.3 The HydF protein	22
1.5.3.1 <i>HydF, a scaffold protein</i>	23
1.5.3.2 <i>HydF, an iron-sulfur protein</i>	25
1.5.3.3 <i>HydF, a nucleotide-binding protein</i>	26
1.5.3.4 <i>Structure of the HydF protein</i>	27
1.5.4 The HydE protein	29
1.5.4.1 <i>HydE, a radical-SAM enzyme</i>	29
1.5.4.2 <i>HydE: a second, non-essential, cluster</i>	30
1.5.4.3 <i>HydE partners</i>	31
1.5.4.4 <i>Structure of the HydE protein</i>	31
1.5.5 The HydG protein	33
1.5.5.1 <i>HydG catalyzes tyrosine conversion to CO and CN</i>	33
1.5.5.2 <i>HydG: a radical mechanism</i>	34
1.5.5.3 <i>HydG: an unprecedented [5Fe-4S] cluster</i>	35
1.5.5.4 <i>Structure of the HydG protein</i>	37
1.5.5.5 <i>HydG enzyme mechanism</i>	38
1.5.6 Mechanism of maturation of [FeFe]-hydrogenases	40
1.5.7 Redox states of the H-cluster of [FeFe]-hydrogenases	42
1.5.8 Maturase-free Chemical maturation: a unique technological tool	44

Section C	53
1.6 Artificial hydrogenases	53
1.6.1 Artificial hydrogenases based on synthetic diiron complexes	54
1.6.1.1 <i>Micelles</i>	56
1.6.1.2 <i>Dendrimers</i>	56
1.6.1.3 <i>Polymers</i>	57
1.6.1.4 <i>Oligo-/poly-saccharides</i>	58
1.6.1.5 <i>Metal-Organic-Frameworks</i>	59
1.6.2 Ni-based artificial hydrogenases	60
1.6.3 Artificial hydrogenases based on synthetic cobalt complexes	63
Section D	69
Aim of the project	69
Results and discussion	73
Chapter II: The [FeFe]-hydrogenase from <i>Megasphaera elsdenii</i>	73
2.1 Expression and aerobic purification of MeHydA	74
2.2 [Fe–S] cluster reconstitution of apo-MeHydA	75
2.3 Static light scattering of apo-MeHydA	77
2.4 EPR spectroscopic characterization of the iron-sulfur MeHydA	78
2.5 Chemical maturation	79
2.5.1 Incorporation of synthetic $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex	79
2.5.2 Incorporation of synthetic $[\text{Fe}_2(\text{pdt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex	80
2.6 FTIR and EPR spectroscopic characterization of holo-MeHydA	81
2.7 Conclusion and future perspectives: biotechnological application?	85
Chapter III: A truncated form of <i>M. elsdenii</i> [FeFe]-hydrogenase	91
3.1 Strategy 1: Cys to Ser mutations	93
3.1.1 Expression and purification of <i>CystoSerMeHydA</i> construct	94
3.2 Strategy 2: MeH-HydA	96
3.2.1 Cloning experiment	96
3.2.2 Expression and aerobic purification of apo-MeH-HydA	97
3.2.3 $[\text{4Fe-4S}]_{\text{H}}$ cluster reconstitution of MeH-HydA apoprotein	99
3.2.4 Spectroscopic characterization of FeS-MeH-HydA	101
3.2.5 Chemical maturation of FeS-MeH-HydA with $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ synthetic complex	103

3.2.5.1 <i>MeH-HydA: new maturation conditions led to higher maturation</i>	105
3.2.6 FTIR and EPR characterization of chemically matured MeH-HydA	107
3.2.7 Conclusion and future perspectives	110
Chapter IV: The [FeFe]-hydrogenase maturase HydF	115
4.1 HydF protein from <i>Thermotoga maritima</i>	116
4.1.1 Expression, purification and iron-sulfur reconstitution of HydF from <i>Thermotoga maritima</i>	117
4.1.2 Insertion of synthetic complexes 1 and 2 onto TmHydF: hydrogenase activity	118
4.1.3 X-ray structure of apo-TmHydF, the apoform of HydF from <i>Thermotoga maritima</i>	120
4.2 Preparation and characterization of iron-sulfur reconstituted HydF from <i>Thermosipho melanesiensis</i> and <i>Clostridium Thermocellum</i>	123
4.2.1 X-ray structure of TmeHydF: a [4Fe-4S] with an unexpected and exchangeable ligand	127
4.2.2 Preparation and characterization of 1 - and 2 -TmeHydF hybrids	136
4.2.3 Hydrogenase activity of 2 -TmeHydF hybrid protein	139
4.2.4 Site-directed mutagenesis of TmeHydF: E305C and E305H	140
4.3 Conclusion and future perspectives	144
General conclusions	149
Chapter V: Materials and Methods	153
5.1 Biologic material	153
5.1.1 Competent cells	153
5.1.2 Plasmids	154
5.1.3 Growth media	154
5.1.4 Molecular Biology	155
5.1.4.1 <i>Competent cells preparation</i>	155
5.1.4.2 <i>Transformation</i>	155
5.1.4.3 <i>Plasmid preparation</i>	156
5.2 Biochemical methods: Protein expression	157
5.2.1 TmHydF: HydF from <i>Thermotoga maritima</i>	157
5.2.2 HydF from <i>Thermosipho melanesiensis</i> (TmeHydF) and <i>Clostridium thermocellum</i> (CtHydF)	157

5.2.2.1	<i>E305C and E305H TmeHydF mutants</i>	158
5.2.3	MeHydA: the [FeFe]-hydrogenase from <i>Megasphaera elsdenii</i>	158
5.2.4	MeH-HydA: the [FeFe]-hydrogenase truncated form of <i>Megasphaera elsdenii</i>	159
5.2.5	CsdA: the Cysteine Desulfurase from <i>E. coli</i>	159
5.3	Proteins purification	160
5.3.1	TmHydF: HydF from <i>Thermotoga maritima</i>	160
5.3.2	HydF from <i>Thermosipho melanesiensis</i> (TmeHydF) and <i>Clostridium thermocellum</i> (CtHydF)	161
5.3.2.1	<i>E305C and E305H TmeHydF mutants</i>	163
5.3.3	CsdA from <i>E. coli</i> .	163
5.3.4	MeHydA: the [FeFe]-hydrogenase from <i>Megasphaera elsdenii</i>	164
5.3.5	MeH-HydA: the [FeFe]-hydrogenase truncated form of <i>Megasphaera elsdenii</i>	167
5.4	Biochemical methods: [Fe–S] cluster reconstitution of apo-proteins	167
5.4.1	Iron-sulfur reconstitution with ⁵⁷ Fe	168
5.4.2	[Fe–S] protein preparation for EPR and HYSCORE measurements	168
5.4.2.1	<i>TmeHydF, E305C and E305H mutants</i>	168
5.4.2.2	<i>MeHydA and MeH-HydA</i>	169
5.5	Chemical methods: Insertion of [Fe ₂ (adt/pdt)(CO) ₄ (CN) ₂] ²⁻ onto HydF and HydA	169
5.5.1	Synthesis of (Et ₄ N) ₂ [Fe ₂ (adt/pdt)(CO) ₄ (CN) ₂] complexes	169
5.5.2	Insertion of [Fe ₂ (adt/pdt)(CO) ₄ (CN) ₂] ²⁻ complex onto MeHydA	170
5.5.2.1	<i>H_{ox} state</i>	170
5.5.2.2	<i>H_{ox}-CO inhibited state</i>	170
5.5.2.3	<i>H_{red} state</i>	171
5.5.3	Insertion of [Fe ₂ (adt)(CO) ₄ (CN) ₂] ²⁻ complex onto MeH-HydA	171
5.5.3.1	<i>H_{ox}, H_{ox}-CO and H_{red} states</i>	171
5.5.4	Insertion of [Fe ₂ (adt/pdt)(CO) ₄ (CN) ₂] ²⁻ complex onto HydF proteins	171
5.6	Hydrogenase-like activity	172
5.6.1	Calibration curve for H ₂ quantitation	172
5.6.2	H ₂ detection via Gas Chromatography of MeHydA and MeH-HydA proteins	173
5.6.3	H ₂ detection via a miniaturized Clark-type hydrogen sensor	174

5.6.3.1	<i>Methyl viologen assay for TmHydF, TmeHydF and MeHydA proteins</i>	174
5.6.3.2	<i>Photocatalytic H₂ production assay driven by Ru(bpy)₃²⁺</i>	176
5.7	Determination of protein concentration	177
5.8	Determination of cofactor concentration	177
5.8.1	Iron quantitation: Fish method	177
5.8.2	Sulfur quantitation: Beinert method	178
5.9	Spectroscopic characterization	179
5.9.1	Uv-Visible spectroscopy	179
5.9.2	EPR spectroscopy	179
5.9.3	HYSCORE spectroscopy	183
5.9.4	FTIR spectroscopy	185
5.9.5	Mössbauer spectroscopy	186
5.9.6	X-ray Crystallography	188
5.9.6.1	<i>Refinement statistics</i>	193
	Acknowledgments	195
	Annex I	197

Abbreviations and Acronyms

[4Fe-4S]_{Aux}, auxiliary [4Fe-4S] cluster;
[4Fe-4S]_H, [4Fe-4S] cluster of the H-cluster;
[4Fe-4S]_{RS}, [4Fe-4S] binding Radical SAM;
1, [Fe₂(adt)(CO)₄(CN)₂]²⁻;
1-HydF, HydF loaded with **1**
2, [Fe₂(pdt)(CO)₄(CN)₂]²⁻;
2-HydF, HydF loaded with **2**
3, [Fe₂(odt)(CO)₄(CN)₂]²⁻;
4OB°, 4-oxidobenzyl radical;
Ado°, 5'-deoxyadenosyl radical;
Adt, azapropanedithiolate;
apo-MeHydA, MeHydA without FeS clusters;
APPA, 4-aminophenyl propionic acid;
Arg, arginine;
ATP, adenosine triphosphate;
Bpy, Bipyridine;
CaHydF, HydF from *Clostridium acetobutylicum*;
CoMyo, cobalt-protoporphyrin IX incorporated into myoglobin;
CpHydA, [FeFe]-hydrogenase from *Clostridium pasteurianum*;
CrHydA1, [FeFe]-hydrogenase from *Chlamydomonas reinhardtii*;
CW, Continuous Wave;
Dcdbt, terephthalate-2,3-dithiolate;
DdH, [FeFe]-hydrogenase from *Desulfovibrio desulfuricans*;
DFT, Density Functional Theory;
DHG, dehydroglycine;
DMSO, dimethylsulfoxide;
DTT, dithiothreitol;
DvH, *Desulfovibrio vulgaris* Hildenborough;
EDTA, Ethylenediaminetetraacetic acid;
ENDOR, Electron Nuclear Double resonance;
EPR, Electronic Paramagnetic Resonance;

FeS-MeHydA, MeHydA reconstituted with three [4Fe-4S] clusters;
FeS-MeHydA, MeH-HydA reconstituted with a [4Fe-4S] cluster;
FeS-TmeHydF, TmeHydF reconstituted with a [4Fe-4S] cluster;
FID, Free Induction Decay;
FMN, flavin mononucleotide;
FPLC, Fast protein liquid chromatography
FTIR, Fourier Transform Infrared;
GC, Gas chromatography;
Gly, glycine;
GTP, Guanosine-5'-triphosphate
Hb, Hemoglobin;
Hb-CO, CO-bound Hemoglobin;
His, Histidine;
Holo-MeHydA, FeS-MeHydA matured with $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$;
 H_{ox} , oxidized state of [FeFe]-hydrogenase;
 $\text{H}_{\text{ox}}\text{-CO}$, CO-inhibited oxidized state of [FeFe]-hydrogenase;
HPLC, High Pressure Liquid Chromatography;
 H_{red} , reduced state of [FeFe]-hydrogenase;
 H_{sred} , superreduced state of [FeFe]-hydrogenase;
HydF-2Fe(xdt), HydF hybrids with complexes harboring different dithiolate bridges;
HYSCORE, HYperfine Sublevel CORrElation;
ICP-AES, inductively coupled plasma-atomic emission spectrometry;
IPTG, Isopropyl β -D-1-thiogalactopyranoside;
LB, Lysogeny broth;
LMCT, Ligand to metal charge transfer;
MeHydA, [FeFe]-hydrogenase from *Megasphaera elsdnii*;
MOF, Metal-organic-Frameworks;
MV, methyl viologen;
MP11, microperoxidase-11;
NaDT, sodium dithionite;
NMR, Nuclear Magnetic Resonance;
Odt, oxopropanedithiolate;
PAA, Polyacrylic acid;
PCR, Polymerase Chain Reaction

Pdt, propanedithiolate;
PEG, Polyethylene glycol;
PSI, Photosystem I;
SAM, S-Adenosyl methionine;
SDS, Sodium dodecyl sulfate;
SDS-PAGE, polyacrylamide gel electrophoresis
SEC, size exclusion chromatography;
SHE, Standard hydrogen electrode;
SoHydG^{WT}, HydG from *Shewanella oneidensis*;
SwMb, Sperm-whale myoglobin;
TB, Terrific broth;
TEA, triethylamine;
TiHydG, HydG from *Thermoanaerobacter italicus*;
TmeHydF, HydF from *Thermosipho melanesiensis*;
TmHydE, HydE from *Thermotoga maritima*;
TmHydE^{ΔFeS}, HydE from *Thermotoga maritima* lacking the auxiliary [4Fe-4S] cluster;
TmHydF, HydF from *Thermotoga maritima*;
TnHydF, HydF from *Thermotoga neapolitana*;
TOF, turnover frequency;
TON, TurnOver Number;
Tris, tris(hydroxymethyl)-aminomethane;
x-HydF, HydF loaded with diiron complexes harboring different dithiolate bridges.
β-met, β-mercaptoethanol;

Statement

The work in this thesis is composed of five chapters. The introduction (Chapter I) has been divided in four distinct sections (**A**, **B**, **C** and **D**), focusing on the [FeFe]-hydrogenase maturation i) via the three maturases HydE, HydF and HydG and ii) via small synthetic complexes. The section **C**, instead, has been centred on artificial and biohybrid systems for H₂ evolution. In section **D** is presented the aim of the PhD project. This first chapter will help the reader in the comprehension of my PhD work.

The experimental results of this thesis are contained in chapters II, III and IV. Chapter V contains all the methodologies for protein purification, sample handling, enzyme activation and some theoretical principles about the techniques used in this study.

The investigations treated in this thesis were carried out jointly with other co-authors from different groups. Their contributions to this thesis include:

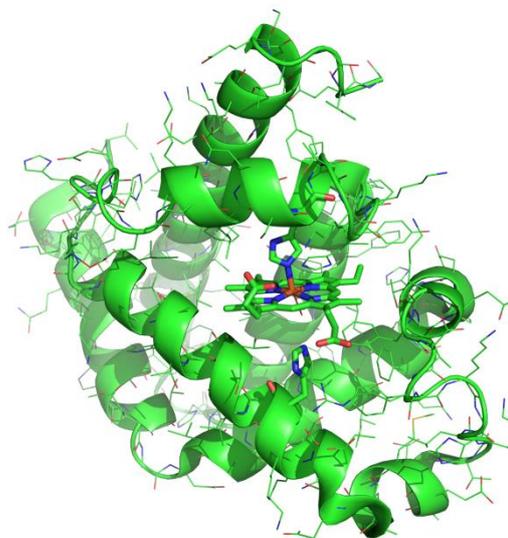
- i) Synthetic complexes preparation by Dr. V. Artero and Dr. S. Roy (CEA, Grenoble);
 - ii) Plasmids preparation of *Megasphaera elsdenii* HydA, truncated *Megasphaera elsdenii* HydA and *Thermotoga maritima* HydF by Dr. Mohamed Atta (CEA, Grenoble);
 - iii) EPR, HYSCORE, Mössbauer and FTIR measurements by Dr. A.V. Adamska, Dr. E. Reijerse and Prof. W. Lubitz (Max Plank Institute, Germany);
 - iv) Resolution of X-ray structures with the help of Dr. L. Pecqueur.
- All other work in this thesis is my own work.

Chapter I
Introduction

Section A

1.1 Metalloprotein: metal ions in living systems

Metalloproteins have fascinated scientists, particularly since the 1950s, when the first X-ray crystal structure of a protein, sperm whale myoglobin, was reported indicating the presence of an iron atom into the protein environment (Figure 1). Since then, this “term” was associated to a protein containing a metal cofactor. Several metals are involved in complex metabolic processes of living organisms with Fe, Cu, Ni, Mn, Co and Zn overall.¹ They have been selected over an evolutionary period of 2-3 billion years on the basis of their unique physicochemical properties and bioavailability.² Protein metal-binding sites are responsible for catalyzing important biological processes, such as photosynthesis, respiration, water oxidation, molecular oxygen reduction, nitrogen fixation and the diversity of function is made possible partly



because they can incorporate cofactors — such as small organic molecules, single metal atoms or clusters that contain metal and non-metal atoms — into their active sites. Although the reactions are based on metal centers, the protein matrix regulates reactivity and a huge effort has been applied to understanding the structure–function relationship of these systems.

These ions are usually coordinated by nitrogen, oxygen or sulfur atoms belonging to amino acids in the polypeptide chain and/or a macrocyclic ligand incorporated into the protein (porphyrin, corrin, chlorin, and so on). Notable examples are Fe, Mg, and Co ions, which are the component of heme, chlorophyll and cobalamins, respectively. Approximately 30-40% of the known protein structures contain metal ions, which are essential for their functions or structures.³ Among the most widespread and ubiquitous inorganic cofactors, the iron–sulfur clusters participate in electron transfer, catalysis and regulatory processes essential for physiological pathways in living cells.⁴

1.2 Iron-sulfur cluster proteins

Section A

[Fe–S] proteins are widely distributed in almost all organisms, leading scientists to theorize that iron–sulfur compounds had a significant role in the origin of life.⁵ They play different functions ranging from enzyme catalysis, protein structure, DNA repair, radical chemistry to biological electron transfer, based on the propensity of Fe to switch between +2 and +3 redox states. [Fe–S] centers can have redox potentials from -500 mV to +300 mV, thus serving as donors/acceptors in several biological reactions (photosystem I, respiration).⁶ In almost all [Fe–S] proteins, iron ions are in a tetrahedral geometry chelated by sulfide groups, two- or three-coordinated, and sidechains of peptide residues. The electron density is delocalized over both Fe and S atoms, making suitable these clusters in biological electron transport.

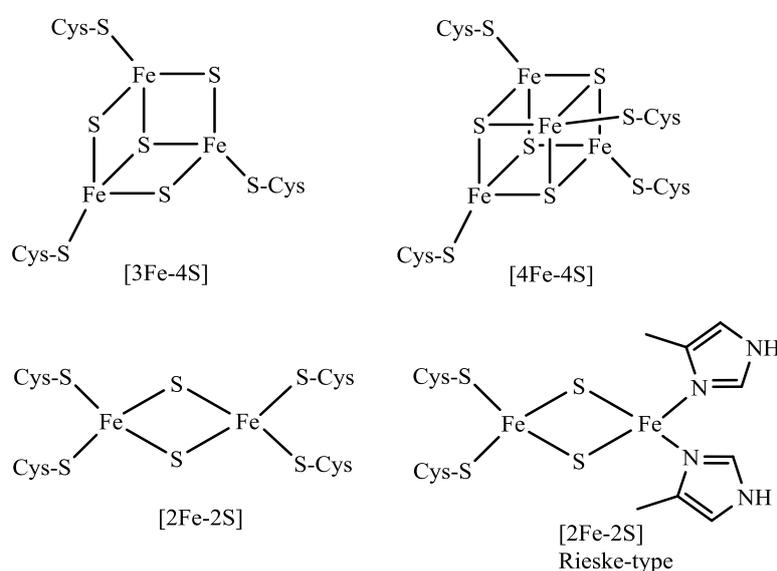


Figure 2: Structures of common [Fe–S] clusters found in nature.

The thiolate function of cysteine generally completes the coordination sphere of the iron ion. Histidine, serine, tyrosine, aspartate, exogenous molecule and/or backbone ligation at a unique Fe site are occasionally encountered in [Fe–S] cluster and generally have a specific role in modifying redox potential, electron transfer, substrate binding and catalysis.⁷ Intriguing is the case of the “Radical SAM” superfamily where the binding of S-adenosylmethionine (SAM) to a [4Fe–4S] cubane, via the amino and carboxylate groups of the methionine fragment, allows reductive cleavage and generation of the 5'-deoxyadenosyl radical.⁸ Different [Fe–S] centers have been reported, through intensive spectroscopic and structural characterizations, among them the [2Fe–2S], [3Fe–4S] and [4Fe–4S] core units (Figure 2). A description of common cluster type, geometry, ligands and cluster function is provided below:

a) [2Fe–2S]

The simplest iron-sulfur core is the [2Fe–2S] cluster, presenting two iron ions bridged by two sulfides and coordinated by four cysteines (Ferredoxins) or by two cysteines and two histidines (Rieske-type). The [2Fe–2S]²⁺ oxidized form has two Fe³⁺ ions, whereas the reduced form [2Fe–2S]⁺ contain a couple Fe³⁺/Fe²⁺. This cluster has access to a third redox state where both irons are Fe²⁺ with a net zero charge, [2Fe2S]⁰. This core is the most versatile and ubiquitous center involved mainly in electron transfer processes.

b) [3Fe–4S]

Iron-sulfur proteins can also contain [3Fe–4S] centers, which present one iron less than the usual [4Fe–4S] clusters. Three sulfur bridge two iron ions each, while the fourth S ion bridges three iron ions. This cluster has access to different oxidation states, varying from [3Fe–4S]⁺, where all the iron are Fe³⁺ to [3Fe–4S]²⁻, with all Fe²⁺. In some iron-sulfur proteins containing a [4Fe–4S] cubane, these clusters can be transformed reversibly into [3Fe–4S] systems by loss of one iron atom. The inactive form of aconitase has a [3Fe–4S] and is activated by reparation of the [4Fe–4S] cluster, adding Fe²⁺ and a reducing agent.⁹ Furthermore, some protein crystals containing a standard [4Fe–4S] cluster can be damaged by RX sources due to strong power amplitude, producing [3Fe–4S] forms of the original cluster.

c) [4Fe–4S]

This unit presents four iron ions and four sulfides placed at the vertices of a cubic structure. The typical ligands are cysteines of the polypeptide chain, although coordination of His, Tyr, Ser, Glu, Lys, Asp and some others are reported. In bacterial ferredoxins, the [4Fe–4S] cluster can adopt +2/+1 oxidation states with redox potentials ranging from -0.3 to -0.7 V. Some [4Fe–4S] proteins bind exogenous ligands (substrates) and function as enzymes (aconitase, radical SAM).⁷ HiPIP proteins have unique [4Fe–4S] clusters, shuttling between [4Fe–4S]²⁺ and [4Fe–4S]³⁺ oxidation states with a redox potential from 0.4 to 0.1 V.

d) Unusual geometries: insight into function

Section A

Particularly interesting are the [4Fe–3S] and [8Fe–7S] clusters, the latter called also P-cluster. The first has been found in [NiFe]-hydrogenases, while the second in nitrogenases.^{6,10} Nitrogenase plays a fundamental role in the nitrogen cycle of the earth, as it catalyzes the conversion of molecular nitrogen to ammonia whereas [NiFe]-hydrogenase catalyze the reversible reaction of proton reduction into hydrogen. Some members of [NiFe]-hydrogenases family are classified as O₂-tolerant enzymes, due to the presence of the [4Fe–3S] center. This asymmetric iron-sulfur cluster presents a unique six cysteine-coordination located close to the catalytic center and has three unusual accessible redox states: i) reduced [4Fe–3S]³⁺, ii) oxidized [4Fe–3S]⁴⁺ and iii) superoxidized [4Fe–3S]⁵⁺, which enable the consecutive transfer of two electrons within a physiological potential range (O₂ reduction to H₂O). The P-cluster [8Fe–7S] is one of the most complex biological metalcenters and takes form of two [4Fe–3S] cubes linked by a central sulfur atom. It is involved in the electron transfer to the FeMo proteins.

1.3 Biosynthesis of [Fe–S] clusters

The maturation of bacterial [Fe–S] proteins has been intensely studied, identifying three different systems for the biogenesis of bacterial iron-sulfur proteins: the NIF system, for specific maturation of nitrogenase, the ISC and SUF assembly systems, for the generation of housekeeping [Fe–S] proteins under normal and oxidative-stress conditions, respectively.⁶ The overall biogenesis process can be split into two main steps: the assembly of an [Fe–S] cluster on a scaffold protein, and the transfer of this core to target apoproteins and its subsequent assembly into the polypeptide chain. Each of these steps involves the participation of several proteins and cofactors, which execute specific biosynthetic partial reactions. All of them have in common a cysteine desulfurase, (NifS, IscS or SufS in bacteria) a pyridoxal 5'-phosphate (PLP)-dependent enzyme, which releases the sulfur required for [Fe–S] cluster formation from cysteine and thus produces alanine. A persulfide intermediate is formed on a conserved cysteine residue of these enzymes. This persulfide may then be transferred to conserved cysteine residues of helper proteins (SufE) or directly to scaffold proteins. Iron is delivered from scaffold/specific iron donors.

1.4 Hydrogenase enzymes

Chapter I

Among the iron-sulfur proteins, hydrogenases have a unique role since they are responsible for hydrogen metabolism in microorganisms. Recent remarkable review articles on this class of enzymes are available.^{11,12} These enzymes can catalyze both hydrogen uptake, furnishing the cell with low-potential electrons, and hydrogen evolution that re-oxidizes the pool of electron carriers participating in fermentation or photosynthesis, according to the following reaction:



There are two kinds of “true” hydrogen-metabolizing enzymes: the [NiFe]- and [FeFe]-hydrogenases which carry reaction 1 to completion. A third kind, the [Fe-S] cluster-free- or [Fe]-hydrogenase, is only found in methanogens where it transfers a hydride to methenyl-tetrahydromethanopterin according to the intermediate step in reaction 1. The latter uses a single Fe atom for catalysis, whereas the other two enzymes have catalytic binuclear metal centers in which either one nickel and one iron or two iron atoms are connected through two thiolate bridges. The bridging thiolates have different origins in the two enzymes: they are provided by two cysteine residues from the polypeptide chain in [NiFe]-hydrogenases, and by a small organic ligand, azapropanedithiolate (adt^{2-}), in [FeFe]-hydrogenases. X-ray crystallography and FTIR spectroscopy have shown that the three kinds of hydrogenase are characterized by active sites that have in common carbonyl ligands to the Fe ion, while only in the case of [NiFe]- and [FeFe]-hydrogenases, Fe atoms also bind CN^- ligands (Figure 3).

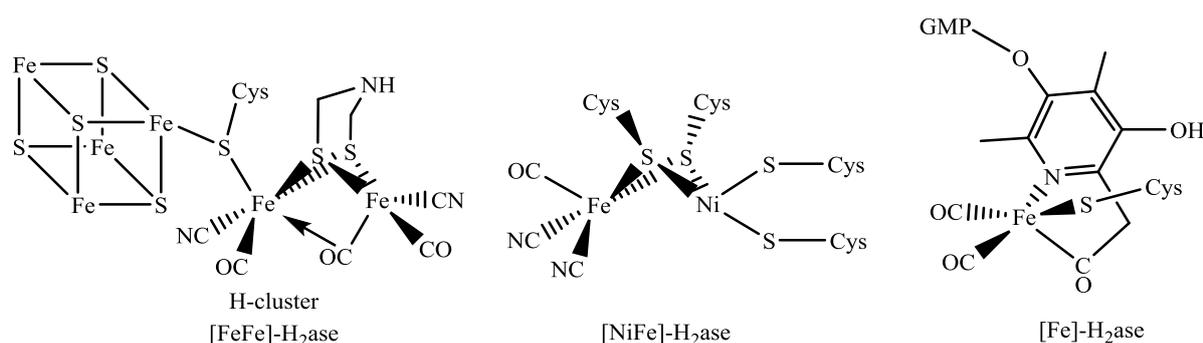


Figure 3: Schematic representation of the active sites of [NiFe]-, [FeFe]- and [Fe]-hydrogenase.

Section A

References:

- 1 W. Kaim, B. Schwederski and A. Klein, *Bioinorganic chemistry: inorganic elements in the chemistry of life: an introduction and guide*, Wiley, Chichester, West Sussex, United Kingdom, Second edition, 2013.
- 2 R. J. Williams, *Cell. Mol. Life Sci. CMLS*, 1997, **53**, 816–829.
- 3 K. J. Waldron and N. J. Robinson, *Nat. Rev. Microbiol.*, 2009, **7**, 25–35.
- 4 H. Beinert, *J. Biol. Inorg. Chem.*, 2000, **5**, 2–15.
- 5 M. J. Russell, R. M. Daniel, A. J. Hall and J. A. Sherringham, *J. Mol. Evol.*, 1994, **39**, 231–243.
- 6 D. C. Johnson, D. R. Dean, A. D. Smith and M. K. Johnson, *Annu. Rev. Biochem.*, 2005, **74**, 247–281.
- 7 M. Fontecave, *Nat. Chem. Biol.*, 2006, **2**, 171–174.
- 8 J. B. Broderick, B. R. Duffus, K. S. Duschene and E. M. Shepard, *Chem. Rev.*, 2014, **114**, 4229–4317.
- 9 M. C. Kennedy, M. H. Emptage, J. L. Dreyer and H. Beinert, *J. Biol. Chem.*, 1983, **258**, 11098–11105.
- 10 T. Goris, A. F. Wait, M. Saggi, J. Fritsch, N. Heidary, M. Stein, I. Zebger, F. Lenzian, F. A. Armstrong, B. Friedrich and O. Lenz, *Nat. Chem. Biol.*, 2011, **7**, 310–318.
- 11 W. Lubitz, H. Ogata, O. Rüdiger and E. Reijerse, *Chem. Rev.*, 2014, **114**, 4081–4148.
- 12 J. W. Peters, G. J. Schut, E. S. Boyd, D. W. Mulder, E. M. Shepard, J. B. Broderick, P. W. King and M. W. W. Adams, *Biochim. Biophys. Acta*, 2015, **1853**, 1350–1369.

Section B

1.5 The metal center of [FeFe]-hydrogenases

Rather detailed information regarding the active site of this class of enzymes is available from the X-ray crystallographic structures of the [FeFe]-hydrogenases from *Clostridium pasteurianum* and *Desulfovibrio desulfuricans* originally solved at 1.8 Å and 1.6 Å resolution, respectively.^{1,2} In all [FeFe]-hydrogenases, the dinuclear Fe center, from now named the [2Fe]-subcluster, shares a cysteine ligand with a classical [4Fe-4S] cluster, thus forming the buried so-called H-cluster, shown in Figure 4. This cysteine is the only protein ligand to the [2Fe]-subcluster. A combination of X-ray crystallographic and infrared spectroscopic investigations has established that each Fe atom has one CO and one CN⁻ ligand and connects to the other Fe through an additional bridging CO ligand.^{3,4} As indicated above, the coordination sphere is completed by a dithiolate bridging ligand, which has been unambiguously identified as azapropanedithiolate (adt²⁻) by spectroscopic and chemical characterization.⁵⁻⁸

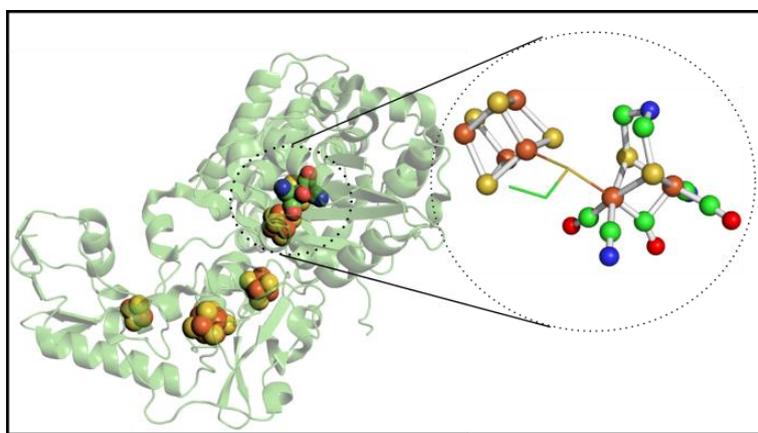


Figure 4: X-ray structure of CpHydA with its protein surrounding. PDB 4XDC. CpHydA has 1 [2Fe-2S], 2 [4Fe-4S] clusters and the H-cluster depicted in the circle. The picture is designed using the program PYMOL. Fe atoms are depicted in brown, S in yellow, C in green, N in blue and O in red.

In [FeFe]-hydrogenases from green algae including *Chlamydomonas reinhardtii*, *Scenedesmus obliquus* and *Chlorella fusca*, the protein only contains the catalytic H-cluster. These enzymes thus represent the simplest forms of [FeFe]-hydrogenases yet identified. So far no three-dimensional structures of fully matured members of that class have been obtained. The enzyme from *C. reinhardtii* has been structurally characterized however with only the [4Fe-4S] component of the H-cluster, named [4Fe-4S]_H.⁹ More generally, [FeFe]-

Section B

hydrogenases contain accessory [4Fe–4S] and [2Fe–2S] ferredoxin-like clusters that function as electron-transfer centers, electrically connecting the active site to the protein surface. At that surface, redox partners either provide electrons for hydrogen generation or capture them from H₂ oxidation for using them as a reducing power within the cell. For example, clostridial enzymes contain, in addition to the H-cluster, four accessory [Fe–S] centers. *D. desulfuricans* and *Megasphaera elsdenii* enzymes contain two ferredoxin-like [4Fe–4S] clusters but the former has two protein subunits while the latter has a single polypeptide chain chelating the H-cluster and the accessory clusters.^{1,3,4}

1.5.1 A protein machinery for [FeFe]-hydrogenase maturation

In order to be catalytically active after its synthesis, the [FeFe]-hydrogenase polypeptide encoded by the *hydA* gene, and thus named HydA in the following, has to incorporate the H-cluster and, when required, accessory [Fe–S] clusters. This post-translational process is extraordinarily complex as it involves a number of difficult reactions including: (i) the assembly of the accessory [Fe–S] clusters; (ii) the synthesis of CO, CN[−] and the adt^{2−} bridging ligand; (iii) the assembly of the [2Fe]-subcluster or a close precursor; (iv) its incorporation into the enzyme already containing the [4Fe–4S]_H component of the H-cluster. These reactions have to be tightly controlled because they produce the toxic CN[−] and CO molecules, as well as the hydrolytically sensitive adt^{2−} ligand. This suggests that within the cell the synthesis of these ligands, their binding to Fe, and the transfer of organometallic intermediates occur via concerted processes, probably within a multiprotein complex consisting of the maturases and the hydrogenase target.

It is well established, using the [FeFe]-hydrogenase from *C. reinhardtii*, that the [4Fe–4S]_H component of the H-cluster is assembled prior to [2Fe]-subcluster insertion into the hydrogenase.¹⁰ Already in 1987 the expression of the two structural genes of [FeFe]-hydrogenase from *Desulfovibrio vulgaris* Hildenborough (*DvH*) in *E. coli* was reported to yield an inactive enzyme, which contained about 12 Fe atoms/hydrogenase molecule.¹¹ At a time when the exact nature of the active metal center of [FeFe]-hydrogenases was unknown, only the two ferredoxin-like [4Fe–4S] clusters were supposed to have been assembled in the recombinant hydrogenase.

However, the reported number of Fe ions per protein molecule strongly suggests that the [4Fe–4S]_H cluster was also assembled in the *DvH* hydrogenase. Thus the currently accepted

model implies separate mechanisms, with a specific machinery being exclusively involved in the synthesis/assembly of the [2Fe]-subcluster and its transfer to an inactive [FeFe]-hydrogenase polypeptide containing all [Fe-S] clusters including the [4Fe-4S]_H cluster. The assembly of these [Fe-S] clusters is likely to depend on the ISC and SUF machineries which are essential for maturation and activation of [Fe-S] enzymes in all prokaryotes.¹²

The [FeFe]-hydrogenase maturation protein machinery, named HYD in the following, was initially discovered in the eukaryotic green algae *C. reinhardtii*, in 2004.¹³ The disruption of either the *hydEF* or *hydG* gene resulted in a mutant that proved to be unable to produce hydrogen, even though full-length hydrogenase accumulated. Furthermore, these genes and the two *C. reinhardtii* [FeFe]-hydrogenase genes, *hydA1* and *hydA2*, were shown to be co-regulated. Finally, heterologous expression of an active *C. reinhardtii* HydA1 protein, CrHydA1, in *E. coli* absolutely required the co-expression of both *hydEF* and *hydG* genes. From all these results it could be concluded that HydEF and HydG provide the minimal protein machinery involved in the synthesis and assembly of the [2Fe]-subcluster of CrHydA1 active site. Whether other proteins are required for an optimal maturation process seems to be excluded so far.

Analysis of the *hydEF* gene from *C. reinhardtii* demonstrated that it contains two domains, which are homologous to two distinct genes coding for the proteins HydE and HydF, which are found exclusively in prokaryotic organisms naturally expressing a [FeFe]-hydrogenase. Thus the genomic organization of the three *hyd* genes, *hydE*, *hydG*, *hydF*, involved in [FeFe]-hydrogenase maturation, varies from one organism to another with three possible patterns; (i) independent genes scattered in the genome (*C. acetobutylicum*, *C. perfringens*); (ii) fused genes (*C. reinhardtii*); (iii) genes organized in operons (*D. desulfuricans*, *B. thetaiotaomicron*).

Several reports of heterologous expression of active [FeFe]-hydrogenases have demonstrated that the HYD machinery from one organism can be successfully used for the maturation of a hydrogenase from another one. For example, expression of an active HydA1 enzyme from *C. reinhardtii* or *S. obliquus* has been shown to be possible using *C. acetobutylicum*, another [FeFe]-hydrogenase synthesizing organism, as the expression host.^{14,15}

Further evidence for the lack of selectivity of the HYD machinery came from the observation that co-expression of HydE, HydF and HydG from the bacterium *C. acetobutylicum* with various algal and bacterial [FeFe]-hydrogenases in *E. coli* resulted in purified enzymes with

Section B

specific activities that were not very different from those of their counterparts from native sources.¹⁶ Also, the bacterium *Shewanella oneidensis* proved to be an efficient system for expression and maturation of HydA1 from *C. reinhardtii*.¹⁷ As additional evidence of non-specificity within the maturases, incubation of *E. coli* extracts expressing inactive HydA from either *C. saccharobutylicum*, *C. reinhardtii* or *C. pasteurianum*, with *E. coli* extracts expressing the three maturation proteins from *C. acetobutylicum* was reported to result in the efficient activation of HydA.¹⁸

A very interesting *in vitro* HydA activation reaction was developed by Swartz and collaborators.^{19,20} First, a cell-free transcription/translation system allowed expression of functional algal and bacterial hydrogenases using cell extracts from *E. coli* expressing the HydG, HydE, HydF proteins from *S. oneidensis*.¹⁹ In a second system, translation and activation were separated into two distinct steps.²⁰ *C. reinhardtii* HydA1 was first heterologously produced in *E. coli*, an organism lacking the HYD machinery, and mixed with dialyzed cell extracts from *E. coli* expressing recombinant HydE, HydF, and HydG maturases from *Shewanella oneidensis* (the maturase extract). While CrHydA1 was partly activated, this system allowed identifying small molecules and salts having a stimulatory effect on the activity. Specifically, cysteine, tyrosine, iron, sulfide and S-adenosylmethionine (SAM) were shown to play an important role in the maturation process. While iron and sulfide are likely to participate to ferredoxin-like cluster assembly, the function of the other additives is now well understood and is discussed below.

1.5.2 The Hyd proteins

Analysis of amino acid sequences of the maturases immediately provided some keys regarding their function:¹³ both HydE and HydG display the CX₃CX₂C sequence characteristic of the so-called “Radical-SAM” superfamily^{21,22} whereas HydF contains all the signatures of a nucleotide-binding protein.¹⁶ These initial assignments were rapidly confirmed specifically through the structural, spectroscopic and functional characterization of the maturases from *Thermotoga maritima* heterologously produced in *E. coli*.^{23,24} Later, maturases from other organisms have also been used as models for investigating the properties of the three maturases and the mechanism of the [FeFe]-hydrogenase maturation.

1.5.3 The HydF protein

1.5.3.1 HydF, a scaffold protein

HydF from *T. maritima* was initially shown, by a variety of spectroscopic methods, to bind a unique [4Fe–4S] cluster, which is described further below.²³ Then, Peters and collaborators elegantly showed that HydF functions as a scaffold protein:²⁵ it provides the site for transiently assembling the [2Fe]-subcluster of HydA or a closely related precursor during a process which requires HydE and HydG action and subsequently transferring it to the [FeFe]-hydrogenase active site for maturation. Indeed, HydF from *C. acetobutylicum* isolated from an *E. coli* strain heterologously expressing the three maturation proteins from *C. acetobutylicum* is sufficient to confer hydrogenase activity *in vitro* to a purified inactive HydA protein heterologously expressed in the absence of HydE, HydG and HydF, thus lacking the [2Fe]-subcluster only, and in the following named apo-HydA.²⁶ This activation reaction is not affected by the addition of GTP (see below). In full agreement with a scaffold protein function, HydF from *C. acetobutylicum*, overexpressed either in *C. acetobutylicum*, a microorganism naturally expressing HydE and HydG, or in *E. coli* expressing the *C. acetobutylicum* maturases, was shown to contain indeed a [4Fe–4S] cluster (see below) but also a second, dinuclear, Fe center with CO and CN⁻ ligands, as demonstrated by FTIR spectroscopy.^{26,27} X-ray absorption spectroscopy at Fe K-edge confirmed this extra cluster to be a diiron complex.²⁸

These results indicate that when expressed in the presence of HydE and HydG, HydF is obtained in a form which contains all the requested chemical elements for activating HydA. This is in full agreement with the hypothesis that it functions as a scaffold protein upon which a [2Fe]-subcluster or a closely related precursor is synthesized.

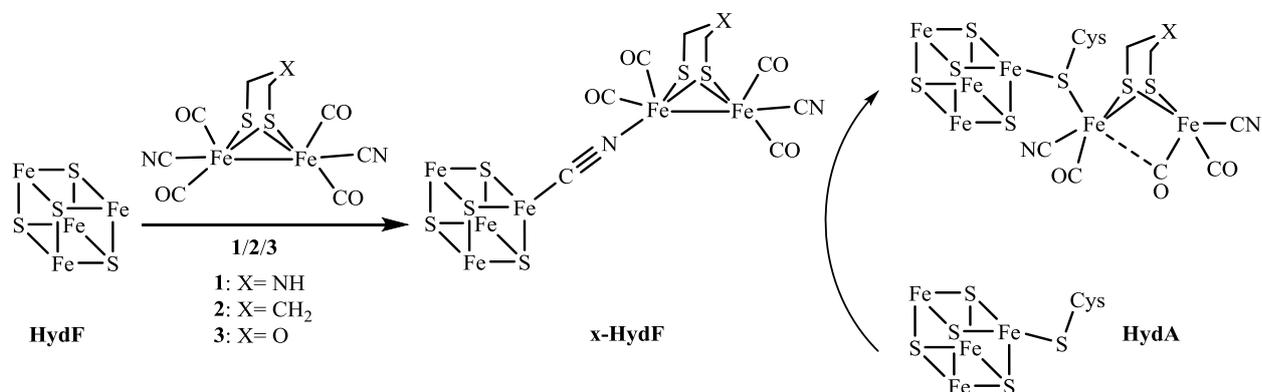


Figure 5: Chemical reconstitution of HydA [2Fe]-subcluster. HydF loaded with synthetic diiron complexes mimicking the [2Fe]-subcluster (complexes **1**, **2**, **3**) can transfer the diiron unit to HydA.

Section B

This hypothesis has been exquisitely confirmed through the collaboration of three different groups (M. Fontecave, W. Lubitz and T. Happe) where they showed that the three synthetic analogues (Figure 5) of the hydrogenase [2Fe]-subcluster (containing different bridging dithiolate ligands) **1**, $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$, **2**, $[\text{Fe}_2(\text{pdt})(\text{CO})_4(\text{CN})_2]^{2-}$ with pdt = propanedithiolate) and **3** $[\text{Fe}_2(\text{odt})(\text{CO})_4(\text{CN})_2]^{2-}$ with odt = oxopropanedithiolate) could be loaded onto *Thermotoga maritima* HydF generating a HydF-2Fe(xdt) hybrid species, named x-HydF (x=1-3).⁶ Elegant FTIR spectroscopy, pulsed EPR techniques and DFT calculations unambiguously demonstrated that the [4Fe-4S] cluster of HydF and the [2Fe]-subcluster analogues shared a cyanide ligand. The diiron component of the x-HydF hybrids can be efficiently transferred to apo-HydA, as shown by FTIR spectroscopy. The resulting HydA preparations proved fully active only when using the 1-HydF hybrid protein.^{6,7} More specifically, the FTIR spectra of x-TmHydF were similar to that of CaHydF purified from the native host (Figure 6). These data clearly: (i) confirm the presence of the adt^{2-} ligand in the active site of native HydA; (ii) support the hypothesis that HydF transiently binds a precursor of the active [2Fe]-subcluster of HydA; (iii) suggest that this precursor onto HydF is structurally close to complex **1**.²⁹ If this is correct, maturation of HydA by HydF thus implies some chemical transformations of complex **1**: (i) dissociation of one CO ligand; (ii) attachment of one iron (thus becoming the proximal Fe) to one cysteine ligand of the [4Fe-4S]_H cluster; (iii) conformational change allowing formation of a CO bridge and liberating a coordination site at the distal Fe.

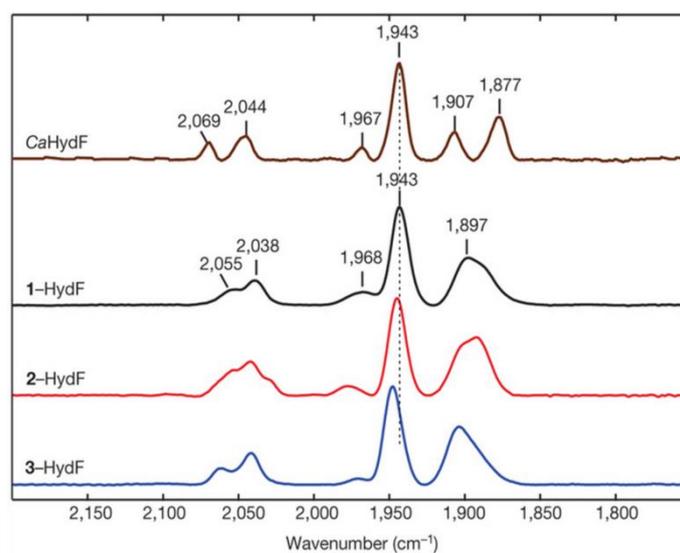


Figure 6: FTIR spectra of CaHydF, isolated from *C. acetobutylicum* host expressing all the clostridial maturases (HydEFG), and x-TmHydF (x=1-3) matured with synthetic diiron complexes shown in figure 5.

1.5.3.2 *HydF*, an iron-sulfur protein

Analysis of an alignment of HydF protein primary sequences from different [FeFe]-hydrogenase-containing microorganisms showed the complete conservation of three cysteine residues located in the C-terminal region of the protein, within a CxHx_nHCxxC motif which was proposed to chelate the reconstituted cluster. Cendron L. and coworkers reported the X-ray structure (paragraph 1.5.3.4) of HydF from *Thermotoga neapolitana* (TnHydF) but unfortunately in its apoform where the Cys residues were involved in intramolecular and intermolecular S-S bridges.³⁰ Those cysteines were found, by site-directed mutagenesis, to be essential for the assembly of the [4Fe-4S] center.³¹ Furthermore, mutation of the three cysteines into serine in HydF from *C. acetobutylicum* resulted in proteins inactive for *C. acetobutylicum* [FeFe]-hydrogenase maturation in an *E. coli* maturation system, demonstrating the cluster to be essential for HydF function.^{16,31} While the cysteines provide three ligands to the cluster, the ligand for the fourth Fe atom of the cluster has been a matter of controversy.

While it seemed to vary from one species to another, it has been well established by spectroscopic studies, in particular HYSCORE experiments, that the fourth coordination site of the cluster is readily exchangeable. Indeed, the cluster can bind exogenous imidazole,^{23,31} a histidine from the N-terminal His-tag of a tagged preparation of HydF,³² or a synthetic diiron complex mimicking the [2Fe]-subcluster of HydA, via a cyanide bridge.⁶ This has led, in the case of HydF from *T. neapolitana*, to the proposal that the exchangeable ligand is a solvent molecule (H₂O or OH⁻), coupled to other water molecules, as suggested by ESEEM and HYSCORE spectroscopy.^{33,34} HYSCORE spectroscopy and site-directed mutagenesis also allowed excluding unambiguously a histidine coordination from the CxHx_nHCxxC motif in HydF from *T. maritima* and *T. neapolitana*.^{23,31,32} In contrast, it seems that the HydF protein from *C. acetobutylicum* uses the C-terminal histidine H352 as a ligand of the cluster as shown by HYSCORE spectroscopy and site directed mutagenesis.³¹ On the other hand, the H352A mutant can still assemble a [4Fe-4S] cluster. All these observations suggest a weakly bound ligand, probably required for functional assembly of the transient [2Fe]-subcluster precursor in HydF. On the other hand, mutations of either histidine of the CxHx_nHCxxC motif in HydF from *C. acetobutylicum* resulted in severe impairment of HydA maturation suggesting that these histidines, which are structurally close to the cluster, play an important role in the chemical processes associated either in the assembly of the diiron complex in HydF or its

Section B

dissociation for transfer to apo-HydA, lacking the [2Fe]-subcluster.^{31,35} The impact of similar mutations in other HydFs on HydA activation has not been studied so that it is impossible to make a generalization of the role of these histidines so far.

In addition to the nature of the fourth ligand, the question of the number of clusters in HydF preparations produced in the absence of HydE and HydG remains a source of debate. *In vitro* reconstitution of clusters in apo-HydF, from *T. maritima*, through anaerobic addition of iron and sulfide, generates one [4Fe-4S] cluster exclusively.^{23,32} This is also the case with HydF from *T. neapolitana* isolated from *E. coli* lacking HydE and HydG.^{33,34} Finally it is also the case with HydF from *Shewanella oneidensis* heterologously expressed in a genetic background devoid of HydE and HydG.³⁶ In contrast and only in that case, HydF from *C. acetobutylicum* isolated from *E. coli* lacking the maturases was claimed to contain a second cluster, defined as a [2Fe-2S] cluster on the basis of its EPR spectroscopy characteristics.²⁶ However, these spectroscopic characterization studies have been carried out on protein preparations containing only 0.85 Fe per protein, only 14% of the theoretical number of 6 expected Fe atoms, thus reflecting a protein with highly degraded clusters. This casts serious doubts regarding the relevance of the observed signals.

Recently Broderick JB and coworkers reported on the existence of redox active [4Fe-4S] and [2Fe-2S] clusters bound to HydF from *C. acetobutylicum*, through UV-visible, circular dichroism and EPR spectroscopic analysis.³⁷ Here again, the investigated protein preparation contained from 2.24 ± 0.49 Fe/dimer up to 4.8 ± 0.8 Fe/dimer, far from the expected 12.0 Fe atoms/dimer for a dimeric HydF containing one [4Fe-4S] and one [2Fe-2S] cluster per monomer, indicating that the studied protein is too much under-metallated to provide relevant conclusions.

1.5.3.3 HydF, a nucleotide-binding protein

Primary sequence analysis of the N-terminal domain of HydF from *T. maritima* and all homologous proteins from various [FeFe]-hydrogenase-containing microorganisms revealed the presence of several conserved consensus sequences, similar to those involved in guanine nucleotide binding in Small-G proteins. The first motif is the (G/A)X₄GK(T/S) sequence responsible for the binding of α - and β -phosphate groups of the nucleotide (the P loop). Mutant HydF proteins from *C. acetobutylicum* in which either one of the glycines or the serine of the P-loop motif has been changed to alanine have been shown to be unable, in

combination with HydE and HydG, to activate HydA from the same organism.¹⁶ Other remarkable features present in HydF proteins are: (i) a conserved threonine residue that might correspond to the residue of the G2 loop involved in Mg^{2+} binding; (ii) the DX₂G sequence that represents the G3 loop involved in the interaction with the γ -phosphate and Mg^{2+} , and (iii) the G4 loop, (N/T)(K/Q)XD, which is supposed to interact with the nucleotide part of GTP. GTP binding to HydF was indeed demonstrated using fluorescence spectroscopy, with a K_d value of 3 μ M for the dissociation of the HydF protein-GTP complex, and furthermore HydF was shown to catalyze GTP hydrolysis to GDP in the presence of Mg^{2+} .²³ A GTPase activity has also been observed in the case of HydF from *Clostridium acetobutylicum*.^{26,27} This activity is affected by the nature of the monovalent cation in the reaction buffer with the highest activity obtained with K^+ and Rb^+ , suggesting that a binding pocket for K^+ exists near the active site. Furthermore, while it is not affected by the presence of HydA, it is stimulated by the presence of HydE and HydG. These results, together with the observed lack of effect of GTP on the activation of HydA by HydF, suggest that the GTPase activity of HydF plays no role in the transfer of the transient diiron complex from HydF to HydA but more likely contributes to assembling this complex on HydF, in concert with HydE and HydG activity. Finally, it was shown also that the GTPase activity was independent on the presence of the cluster.

It is interesting to note that HypB, a protein with GTPase activity and K_d values for GTP comparable to those displayed by the HydF protein, is involved in the maturation of [NiFe]-hydrogenase.^{38,39} It has been shown that HypB participates in the GTP-dependent insertion of the nickel atom into an hydrogenase form already containing the $Fe(CO)(CN)_2$ motif. It seems to be a common theme that nucleotidase activity is required in metal-containing active site biosynthesis.

In addition to the HydF and HypB proteins, one can mention the $CooC$ ^{38,40} and $UreG$ ⁴¹ proteins, involved in nickel insertion during maturation of carbon monoxide dehydrogenase and urease, respectively, and $NifH$, which functions in the ATP-dependent insertion of the FeMoco into apo-nitrogenase.⁴²

1.5.3.4 Structure of the HydF protein

The first crystal structure of TnHydF, from the hyperthermophilic organism *Thermotoga neapolitana*, however without its [4Fe-4S] cluster, refined at 3.0 Å resolution, has

Section B

been reported in 2011 by Cendron L. and coworkers.³⁰ The asymmetric unit contained a monomer, but the biological unit is a dimer, generated by a crystallographic 2-fold axis, or a tetramer, generated by the dimerization of dimers (Figure 7). A single monomer of TnHydF is composed of three domains. Residues 1-171 belong to the GTP binding domain (depicted in blue), residues 186-266 to the dimerization domain (depicted in yellow) and residues 267-398 to the iron-sulfur cluster binding domain (depicted in green). The GTP binding domain is linked to the dimerization domain via a stretch of some residues which likely provides flexibility. The three highly conserved cysteine residues (Cys-302, Cys-353, and Cys-356) that represent the [Fe-S] cluster-binding site, possibly along with a fourth residue, are spatially close together forming a superficial pocket.

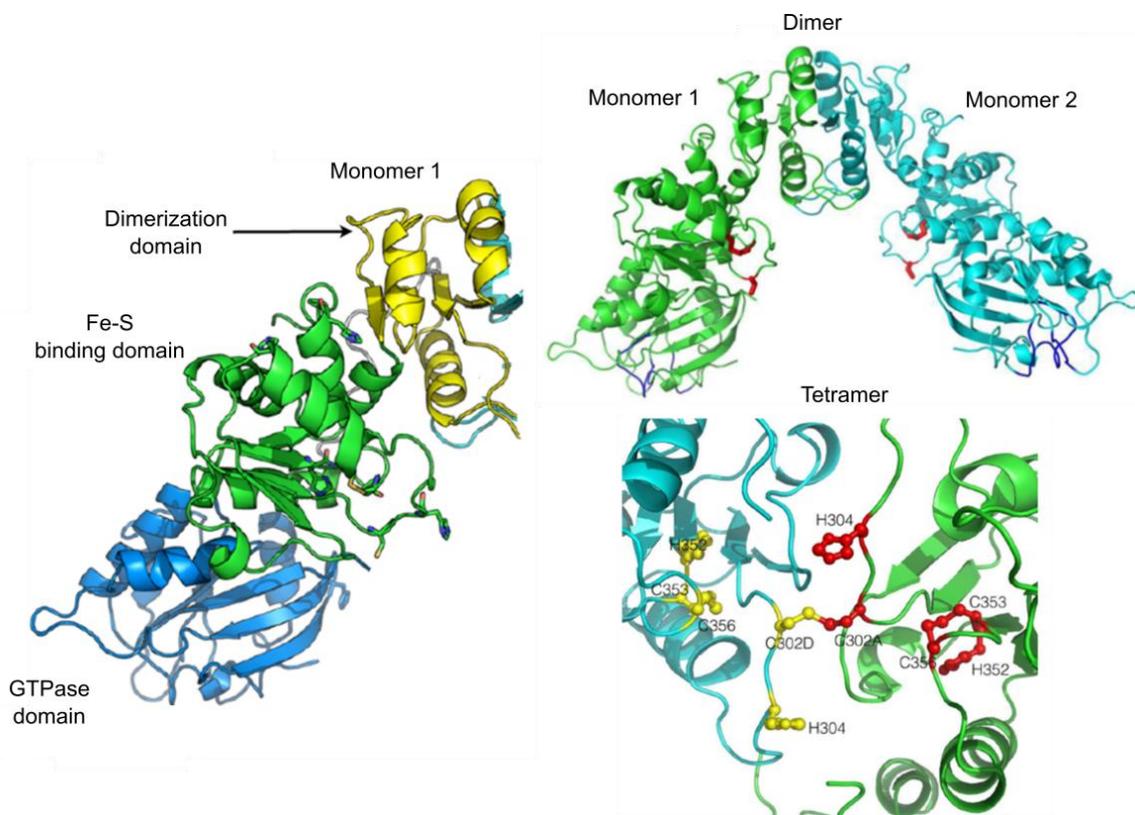


Figure 7: X-ray structure of apo-HydF from *Thermotoga neapolitana*. Representation of HydF monomer composed of three domains: dimerization domain, Fe-S binding domain and GTP binding domain. Two monomers interact producing a dimer, which is the biological unit, and two dimers can form a supramolecular tetrameric organization.

These residues of the CxHx_nHCxxC motif were involved in intramolecular and intermolecular disulfide bridges giving rise to a supramolecular tetrameric organization (dimer of dimers). Unfortunately, this structure could not provide any information on the iron-sulfur binding site.³⁰

1.5.4 The HydE protein

1.5.4.1 HydE, a Radical-SAM enzyme

Primary sequences of HydE proteins contain the CX₃CX₂C motif characteristic of “Radical-SAM” enzymes.²² The latter are involved in such diverse functions as enzyme activation, DNA repair and tRNA modification, cofactor (lipoic acid, molybdopterin) and vitamin (biotin, thiamin) biosynthesis, maturation of nitrogenase and they participate in many other metabolic and biosynthetic pathways.²² In all cases studied so far, “Radical-SAM” proteins use a [4Fe–4S]⁺ cluster to bind SAM and reduce it by a one-electron process which results in the homolytic cleavage of the S-C(deoxyadenosyl) bond. This cleavage generates the very reactive 5'-deoxyadenosyl catalytically competent radical intermediate, 5'-Ado[•]. The latter serves to activate the enzyme substrate (a small molecule or a glycine residue of a second protein) for subsequent radical transformations through the initial selective abstraction of a critical H atom.

The very first studies of HydE from *Thermotoga maritima* established that the protein bound two distinct [Fe–S] clusters.^{24,43,44} Spectroscopic, crystallographic and site-directed mutagenesis characterization clearly demonstrated the presence of a [4Fe–4S] cluster chelated by the conserved CX₃CX₂C motif. In its fully metallated purified form, HydE is able to catalyze the *in vitro* reductive cleavage of S-adenosylmethionine (SAM) into methionine and 5'-deoxyadenosine (AdoH), as expected for a “Radical-SAM” enzyme. Furthermore, mutation of each of the three characteristic cysteines to serine in *C. acetobutylicum* HydE showed that these residues are essential for *in vivo* maturation of *C. acetobutylicum* [FeFe]-hydrogenase, when using an *E. coli* heterologous expression system.¹⁶

It is thus likely that also in the case of HydE, an Ado[•] radical intermediate is formed and reacts with a low-molecular weight substrate to generate a substrate radical which subsequently undergoes a selective transformation. The scientific community favors the hypothesis that this substrate is the precursor of the azapropanedithiolate ligand of the [2Fe]-subcluster, but this substrate has not been identified yet. Its identification is currently one of the most pressing issues.

Recently, it was shown that a variety of low-molecular weight thiols, including cysteine, increased the rate of SAM cleavage catalyzed by CaHydE (from *C. acetobutylicum*).⁴⁵ These thiols promoted incorporation of deuterium from D₂O into deoxyadenosine, proving direct

Section B

abstraction of the H atom of the thiol by the radical Ado^o. These data were used to speculate that the substrate of HydE is a cysteine-like thiol which would get converted, via a radical mechanism, into thioformaldehyde. Two molecules of thioformaldehyde would then condense together with ammonia for generating azapropanedithiolate. The authors pointed the genetic association of *aspA* with *hydAEFG* genes as supporting their hypothesis. Indeed AspA is the aspartate ammonialyase which converts aspartic acid into ammonia and fumarate. However, no direct evidence of participation of this enzyme in H-cluster biosynthesis has ever been obtained so far.

Identification of HydE substrate could be theoretically obtained via labeling experiments. With the aim of establishing a specific spectroscopic probe for the adt^{2-} ligand of the H-cluster in HydA, CrHydA1 was activated with three isotopically labeled variants of the synthetic $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex, containing either C^{15}N^- ligands, $^{15}\text{N}\text{-adt}^{2-}$ or both ligands.

Using HYSCORE spectroscopy, it was found that $^{15}\text{N}\text{-adt}^{2-}$ signals were well-defined, easy to detect and clearly distinguishable from those of C^{15}N^- ligands.⁸ This represents a unique probe for monitoring the presence of ^{15}N -labeled adt^{2-} within the active site of HydA and provide a strategy to screen ^{15}N -labeled potential precursors of adt^{2-} using for example the *in vitro* cell-free [FeFe]-hydrogenase maturation methodology developed by Swartz and coworkers.^{19,20}

1.5.4.2 HydE: a second, non-essential, cluster

The nature and function of the second cluster in HydE are less well established. First, in the case of HydE from *T. maritima*, a CysX₇CysX₂Cys motif in the C-terminal region of the protein was shown to chelate a variety of cluster forms, with 2 Fe, 3 Fe or 4 Fe atoms, by high resolution X-ray crystallography and site-directed mutagenesis.^{16,24,43,44} Intriguingly, this motif is present in several, but not all, HydE orthologs, and there is a second class of HydE proteins, which lack the CysX₇CysX₂Cys motif. Mutation of any of the cysteines of this motif into alanine in *C. acetobutylicum* HydE, belonging to the first class, did not prevent it from activating the [FeFe]-hydrogenase from the same organism, heterologously expressed in *E. coli*. In addition, in similar experiments, replacement of *C. acetobutylicum* HydE by *Bacteroides thetaiotaomicron* HydE, which belongs to the subset lacking the cysteine ligands of the second cluster, did not affect hydrogenase maturation.⁴³ All these data clearly show that the cysteine-containing sequence that coordinates the second cluster, found in HydE from *T.*

maritima and some of its orthologs, is not essential for maturation, raising the question of the function of the second cluster, if any.

1.5.4.3 *HydE* partners

HydE was shown to have a privileged physical connection with HydF while there was no evidence for an interaction between HydE and HydG.⁴⁶ HydE stimulates the rate of GTP hydrolysis in HydF²⁵ and the rate of dissociation of HydE from HydF was increased in the presence of GTP. A direct and strong interaction between the two proteins was also established by surface plasmon resonance experiments.⁴⁶ Finally in some organisms HydE and HydF are fused in a single protein encoded by a single gene.¹³

1.5.4.4 *Structure of the HydE protein*

The 1.35 Å resolution crystal structure of HydE from *T. maritima*, depicted in Figure 8, came out in 2008 from a collaboration between M. Fontecave group and J. Fontecilla-Camps one.⁴³ Later on, as mentioned above, J. Fontecilla-Camps and coworkers reported three-dimensional structures of TmHydE revealing the second cluster in a variety of forms, with different nuclearities.⁴⁴ As the other “Radical SAM” proteins, HydE has the common structural core built around a $(\beta\alpha)_6$ three quarter TIM barrel.²² As was expected from previous amino acid sequence alignments, the HydE structure is very similar to that of BioB⁴⁷ with its C-terminal extension defining a complete $(\beta\alpha)_8$ TIM-barrel.⁴³ HydE can thus be classified within the “complete barrel Radical SAM” protein sub-class. This, in turn, confirms that its substrate is a small molecule. Indeed, many of the conserved amino acids in HydE are found on the β -strands that, in BioB and other members of this family, define the substrate-binding cavity inside the barrel, below the [4Fe–4S]-SAM complex site. Accordingly, mutations of these residues around the inferred substrate-binding site produced HydE mutants with little or no maturation activity assayed via *in vivo* tests.⁴³ Furthermore, HydE was shown to have a high affinity for methionine and 5'-deoxyadenosine.⁴⁸

In the HydE crystal structure, three linearly disposed sites bind halide ions as shown by the anomalous difference peaks observed when displacing putative Cl^- by Br^- in soaking experiments.⁴³ Two of these sites are found in the putative substrate-binding region mentioned

Section B

above whereas the third one is located at the bottom of the internal cavity, relatively far from SAM.

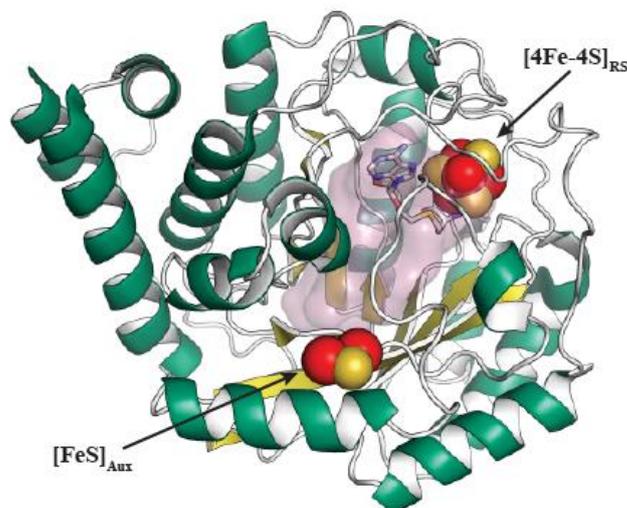


Figure 8: Structure of TmHydE. The Fe-S clusters are shown as spheres with sulfur colored in yellow and iron in red, SAM is displayed as sticks. HydE contains one [4Fe-4S] cluster associated with Radical-SAM activity ([4Fe-4S]_{RS}). A second cluster is shown ([Fe-S]_{Aux}) but it is of variable constitution and can be removed without affecting the enzyme activity. The substrate binding site cavity is shown as transparent purple surface.

Additional crystal soaking experiments showed that SCN^- binds with very high affinity to this latter site, displacing the putative Cl^- ion of the original structure. Interestingly, mutations around the SCN^- binding site also resulted in a significant drop in maturase activity.⁴³ This array of binding sites all along the barrel might suggest that the substrate reacts at the top of the barrel near the SAM binding site and the product travels down to the bottom of the barrel for transfer to another protein, HydF for example. The *in silico* screening for docking of about 60,000 small molecules indicated that the two anion-binding sites found in the putative substrate-binding region of HydE could also bind two carboxylate moieties.⁴³ For example, Nicolet Y. and collaborators reported a crystal structure of TmHydE ^{ΔFeS} (in which the cysteines of the CX₇CX₂C motif have been mutated into serines and the second cluster is absent) containing a 2-methyl-1,3-thiazolidine-2,4-dicarboxylic acid sitting in the substrate binding site.⁴⁹ The two carboxylate moieties are located in the previously identified anion-binding sites. This molecule is not the physiological substrate of HydE and indeed was unable to enhance [FeFe]-hydrogenase activation.

In conclusion, even though there is too little information regarding the chemistry of HydE so far, the current scenario is that this maturase is responsible for the synthesis of the azapropionedithiolate ligand of the [2Fe]-subcluster of H-cluster. Due to the affinity of HydE

for HydF it is tempting to suggest that the product of the reaction is directly transferred to HydF for assembling the transient diiron complex, precursor of the [2Fe]-subcluster of the H-cluster.

There are still a number of issues to solve before we understand how the full process occurs, for example regarding the substrate of HydE, the function of the non-essential accessory cluster and the structure of the HydE-HydF complex.

1.5.5 The HydG protein

1.5.5.1 HydG catalyzes tyrosine conversion to CO and CN

In contrast to HydE, detailed knowledge of HydG structure and function has resulted from exquisite biochemical, spectroscopic and structural studies in the last few years. It has been very early established that HydG proteins are “Radical-SAM” enzymes which contain two [4Fe–4S] clusters, both absolutely required for activity.²⁴ The first one, [4Fe–4S]_{RS}, is chelated by the characteristic “Radical-SAM” CysX₃CysX₂Cys motif while the second, [4Fe–4S]_{AUX}, uses the cysteines from another conserved C-terminal CysX₂CysX₂₂Cys motif. The six cysteines were shown by site-directed mutagenesis to be essential for activity.¹⁶ Thus both clusters have a free coordination site, the first one for binding SAM and the second for binding a key mononuclear Fe complex, as discussed below.

An amino acid sequence comparison of HydG with other members of the “Radical-SAM” protein family indicated that tyrosine lyases (ThiH) were among the most closely related to HydG with 27% identity. 35% of the identities between HydG and ThiH correspond to strictly conserved residues in all the known amino acid sequences of both families. ThiH catalyzes the conversion of tyrosine into para-cresol and dehydroglycine (DHG), a precursor in thiazole biosynthesis.⁵⁰ This genomic analysis and the remarkable homology between ThiH and HydG led to discover that tyrosine is the substrate of HydG.⁵¹ This result extends the sequence homology of HydG and ThiH to a remarkable functional similarity, which provided a solid basis for subsequent mechanistic investigations.

P. Roach and coworkers then showed that CO and CN[−] ligands are produced from SAM-dependent radical cleavage and decomposition of tyrosine.^{52,53} Cyanide was detected and quantified after derivatization with a fluorescent probe then analyzed by HPLC and LCMS.

Section B

Experiments carried out with labeled tyrosine confirmed that it is indeed the source of cyanide.⁵³ CO production from tyrosine was detected by using deoxyhemoglobin as a reporter and monitored by visible spectroscopy since carboxyhemoglobin (HbCO) spectrum displays characteristic absorption bands.⁵² Hb¹³C formation in assays using labeled ¹³C-tyrosine was confirmed by FTIR spectroscopy.⁵² Finally, the cell-free hydrogenase maturation system developed by Swartz J. and coworkers, using a mixture of HydEFG maturases and inactive CpHydA, heterologously expressed in *E. coli*, in the presence of labeled tyrosine, allowed confirming, by FTIR spectroscopic analysis, that CO and CN⁻ ligands of the H-cluster were labeled and thus derived from the tyrosine. ¹³CO and C¹⁵N⁻ were formed when tyrosine was selectively labeled at the carboxylic and the amino function, respectively.⁵⁴

1.5.5.2 HydG: a radical mechanism

A first mechanism involving a radical intermediate and several iron species derived from tyrosine has been proposed in 2013 by Britt D. and coworkers.⁵⁵ A combination of EPR, hyperfine sublevel correlation (HYSCORE) and electron nuclear double resonance (ENDOR) spectroscopies allowed identifying key reaction intermediates using HydG from *Shewanella oneidensis* (SoHydG^{WT}) and ²H-, ¹³C- and ¹⁵N-labeled tyrosine isotopologs. Consistent with HydG being a Radical-SAM enzyme, a mechanism was proposed in which Ado^o radical, derived from SAM cleavage at the N-terminal [4Fe-4S]_{RS} cluster, abstracts a H atom at the tyrosine substrate. The resulting tyrosyl radical then enjoys heterolytic cleavage at the C α -C β bond generating a transient 4-oxidobenzyl (4OB^o) radical as well as a dehydroglycine bound to the [4Fe-4S]_{AUX} cluster. Reduction and protonation of 4OB^o generates p-cresol, a final product of the reaction, while dehydroglycine is activated further for conversion into CO and CN⁻ which end up bound to the [4Fe-4S]_{AUX} cluster.⁵⁵

In a second report, D. Britt and coworkers using time-resolved FTIR and ENDOR spectroscopy provided evidence for the formation of two distinct Fe(CO)_x(CN) species: complex A, with the Fe(CO)(CN) stoichiometry, forms first followed by complex B, the final product, with the Fe(CO)₂(CN) stoichiometry.⁵⁶ The presence of a second CO ligand in complex B thus implies the requirement for HydG to fragment two tyrosines and the release of one CN⁻ ion. In addition, exquisite ⁵⁷Fe-labeling experiments clearly established HydG as the source of Fe for the [2Fe]-subcluster of HydA.⁵⁶ This thus excludes HydF as being the source of Fe, in the form of a [2Fe-2S] cluster as proposed by some reports.^{26,37} From these

studies, it was thus established that HydG serves to convert tyrosine into an organometallic $\text{Fe}(\text{CO})_2(\text{CN})$ synthon that is ultimately transferred to the hydrogenase to form the [2Fe]-subcluster, presumably via HydF. At that time the $\text{Fe}(\text{CO})_2(\text{CN})$ unit was proposed to be part of the cubane $[\text{4Fe-4S}]_{\text{AUX}}$ cluster, with the CO and CN^- ligands bound to one corner iron of the cluster, the three other iron atoms being coordinated by three cysteine residues. This configuration implied mobilization of this $\text{Fe}(\text{CO})_2(\text{CN})$ site for transfer to HydF and then HydA and thus partial degradation of the cluster into a [3Fe-4S] species during the process. However, recent structural and spectroscopic studies now exclude such a model.

1.5.5.3 HydG: an unprecedented [5Fe-4S] cluster

As a matter of fact, the three-dimensional structure of HydG (described in paragraph 1.5.5.4) showed a structurally unprecedented $[\text{5Fe-5S}]_{\text{AUX}}$ cluster, at the C-terminus, consisting of the complete $[\text{4Fe-4S}]_{\text{AUX}}$ cubane bridged to a mononuclear Fe(II) center, so-called “dangling Fe”, via a sulfide bridge (Figure 9c).⁵⁷ The fifth Fe is further coordinated by a histidine residue, water molecules and a nonproteinaceous amino acid which could not be identified due to structural disorder (Figure 9c). Support for such a cluster came from detailed EPR spectroscopic characterization of HydG. In particular a $S = 5/2$ species was detected and assigned to the $[\text{5Fe-5S}]_{\text{AUX}}$ cluster. Accordingly, the $S = 5/2$ signal is absent from EPR spectra of HydG mutants that cannot bind the C-terminal cluster, as a consequence of mutation of Cys residues that coordinate this cluster into serine.⁵⁵ Finally, Britt D. and coworkers further characterized the dangling Fe by showing that it contains L-Cysteine as a ligand rather than sulfide. It is thus proposed to be attached to the $[\text{4Fe-4S}]_{\text{AUX}}$ cluster indeed via a sulfur bridge but provided by this cysteine ligand and not by a sulfide.^{58,59}

Section B

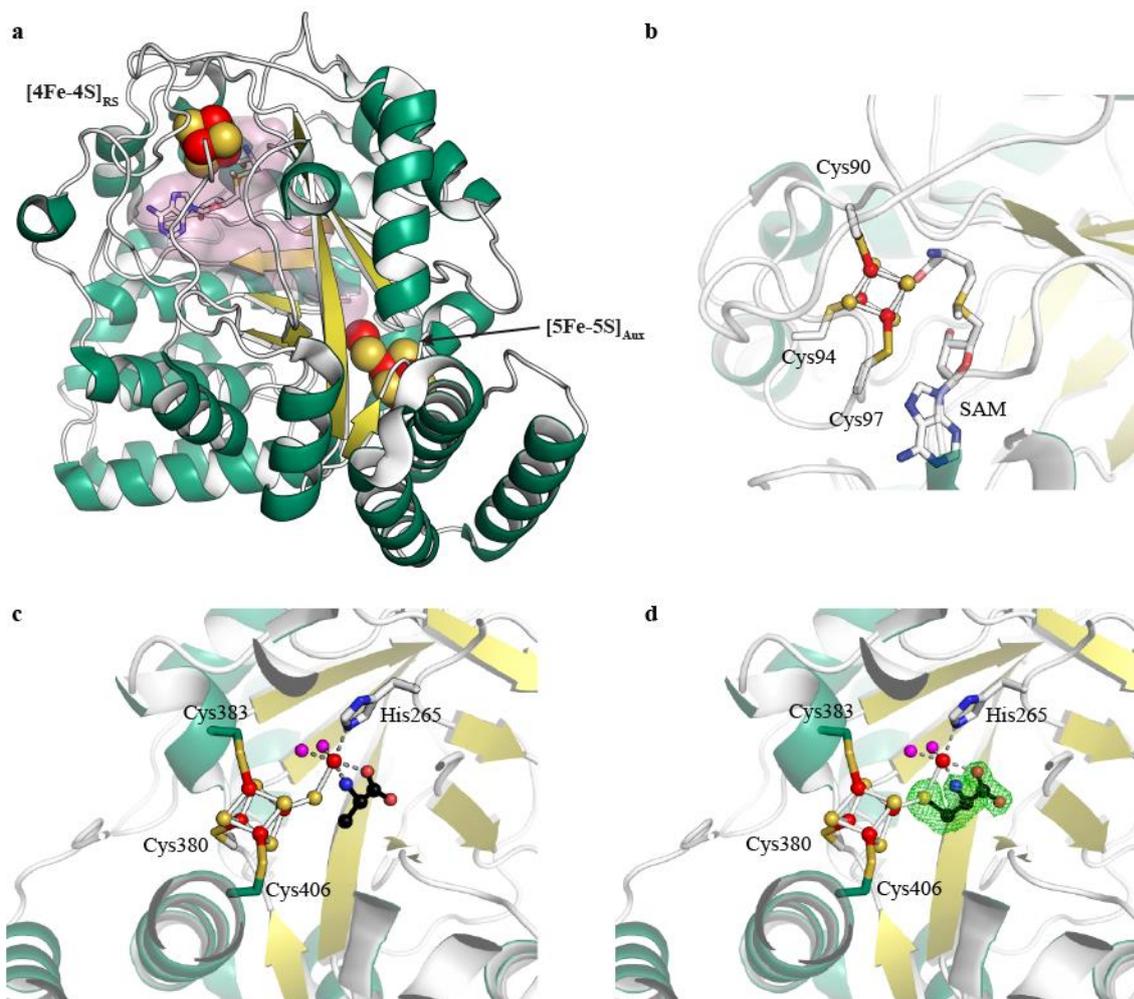


Figure 9: Structure of TiHydG. (a) Overall structure of TiHydG (pdb 4wcx). The two iron-sulphur clusters are depicted as spheres with sulphur and iron colored in yellow and red respectively. SAM is represented as sticks. The substrate binding site cavity linking the two clusters is shown as transparent purple surface. (b) [4Fe-4S] cluster associated with Radical-SAM where tyrosine is converted to CO and CN⁻ (c) Auxiliary [Fe-S] in its [5Fe-5S] form. The fifth iron (in red) is coordinated by His265, two water molecules (magenta), a sulfide and an unidentified amino acid (in sticks with carbons and bonds in black, nitrogen and oxygen in blue and light red respectively). (d) Reinterpretation of the auxiliary [Fe-S] cluster. The unidentified amino acid and sulfide in TiHydG were replaced by a cysteine and the model was refined. After removal of the cysteine and an additional cycle of refinement a (F₀-F_c) omit map was calculated and is shown at the 3σ level (green). The cysteine fits nicely in the map, which is consistent with the results published by Suess et al.⁵⁹

In this configuration, the cluster is rather a [5Fe-4S]_{AUX} cluster, likely with a [(κ³-Cys)Fe(H₂O)₂] dangling unit. It is noteworthy that the structure reported by P. Roach and coworkers could be reinterpreted by filling the electron density around the dangling Fe with homocysteine⁶⁰ or L-cysteine (Figure 9d) rather than with sulfide + alanine. EPR spectroscopic analysis showed that Fe binding is reversible. Treatment with the metal chelating agent EDTA converts the S = 5/2 [5Fe-4S]_{AUX} cluster to a S=1/2 [4Fe-4S]_{AUX}

cluster, in which the cysteine stays as a ligand. Accordingly, addition of Fe^{2+} , only, regenerates the $S = 5/2$ EPR signal.⁵⁸ Such a dangling Fe is well organized to receive two CO and one CN^- ligands to generate a labile $[(\kappa^3\text{-Cys})\text{Fe}(\text{CO})_2(\text{CN})]^-$ unit absolutely required for HydA maturation. Accordingly, formation of complex B likely containing a $[5\text{Fe-4S}]_{\text{AUX}}$ cluster with a $[(\kappa^3\text{-Cys})\text{Fe}(\text{CO})_2(\text{CN})]^-$ unit was observed by stopped-flow IR spectroscopy, indeed only when enzyme preparations contained a dangling Fe bound to cysteine.⁵⁹ Using labeled tyrosine and HYSORE spectroscopy, a $[4\text{Fe-4S}]_{\text{AUX-CN}}$ derivative was also observed, following formation of complex A and complex B, during catalysis, presumably via liberation of $[(\kappa^3\text{-Cys})\text{Fe}(\text{CO})_2(\text{CN})]^-$.⁵⁸ It has been proposed that the released CN^- ion from one of the two tyrosines converted by HydG might serve to displace $[(\kappa^3\text{-Cys})\text{Fe}(\text{CO})_2(\text{CN})]^-$ from the $[5\text{Fe-4S}]_{\text{AUX}}$ cluster. Complex B could not be observed when D-cysteine, L-homocysteine, L-serine and alanine plus sulfide were used in place of L-cysteine.

It is important, finally, to note that a HydG mutant lacking the C-terminal cluster produces only cyanide and not CO.^{59,60} Formate instead is produced, showing that the cluster is absolutely essential for controlled concomitant production of cyanide and carbon monoxide. More specifically, a HydG mutant, in which binding of the dangling Fe is prevented by mutation of the histidine ligand, does not produce any CO (but formate), supporting the idea that this Fe site is critical for CO synthesis.⁶⁰

1.5.5.4 Structure of the HydG protein

A X-ray crystallographic structure at 1.59 Å resolution of HydG from *Thermoanaerobacter italicus* (TiHydG) has been recently solved by Roach PL and coworkers (Figure 9a).⁵⁷ With two HydG monomers per asymmetric unit, it consists of a complete $(\beta\alpha)_8$ TIM barrel fold with an N- and C-terminal helical extensions. A comparison with structures of other Radical-SAM enzymes, HydE, the biotin synthase BioB, the tryptophan lyase NosL and the methylornithine synthase, showed that TiHydG has an extra C-terminal extension of 80 amino acids chelating the auxiliary cluster.⁵⁷ The distance between the $[4\text{Fe-4S}]_{\text{RS}}$ cluster and the auxiliary cluster is about 24 Å. In the first monomer, the $[4\text{Fe-4S}]_{\text{RS}}$ cluster is chelated by the $\text{CysX}_3\text{CysX}_2\text{Cys}$ motif and a methionine via its α -amino and α -carboxy groups, probably derived from SAM. In the second monomer, SAM is bound via α -amino and α -carboxy groups, with the S atom located 3.4 Å away from the proximal iron (Figure 9b).

Section B

Regarding the auxiliary cluster, placed at the end of the TIM barrel fold in the C-terminal helical domain, one of the 2 monomers showed the presence of a [5Fe–5S] cluster. The fifth iron had occupancy of 0.73 and adopted an approximately octahedral geometry with a histidine ligand (His 265) coordinating the iron via the N ϵ of the imidazole ring. The coordination sphere of the iron is completed by two H₂O molecules and an undefined nonproteinaceous amino acid bound in a bidentate manner via the α -amino and α -carboxy functions, modelled as an alanine in the structure (Figure 9c).⁵⁷ In the second monomer, the dangling Fe is absent and the density is assigned to a [4Fe–5S] cluster, in which a S atom is bound to one corner Fe of the cubane.

While the tyrosine substrate was not present in the structure it is proposed to lie close to the [4Fe–4S]_{RS} cluster that would allow H atom abstraction. Superimposition of HydG crystal structure with that of NosL allowed modelling the tyrosine in the same pose of tryptophan in NosL, in a cavity between the two clusters.^{57,61} With this assumption, in the putative binding site the α -amino function of tyrosine is oriented towards the [4Fe–4S]_{RS} cluster and the phenol group pointing towards the dangling Fe.⁵⁷

1.5.5.5 HydG enzyme mechanism

Taken together these results allow proposing a detailed catalytic mechanism for CO and CN⁻ production by HydG. The mechanism in Figure 10 combines consistent proposals from different studies.^{57,59,60}

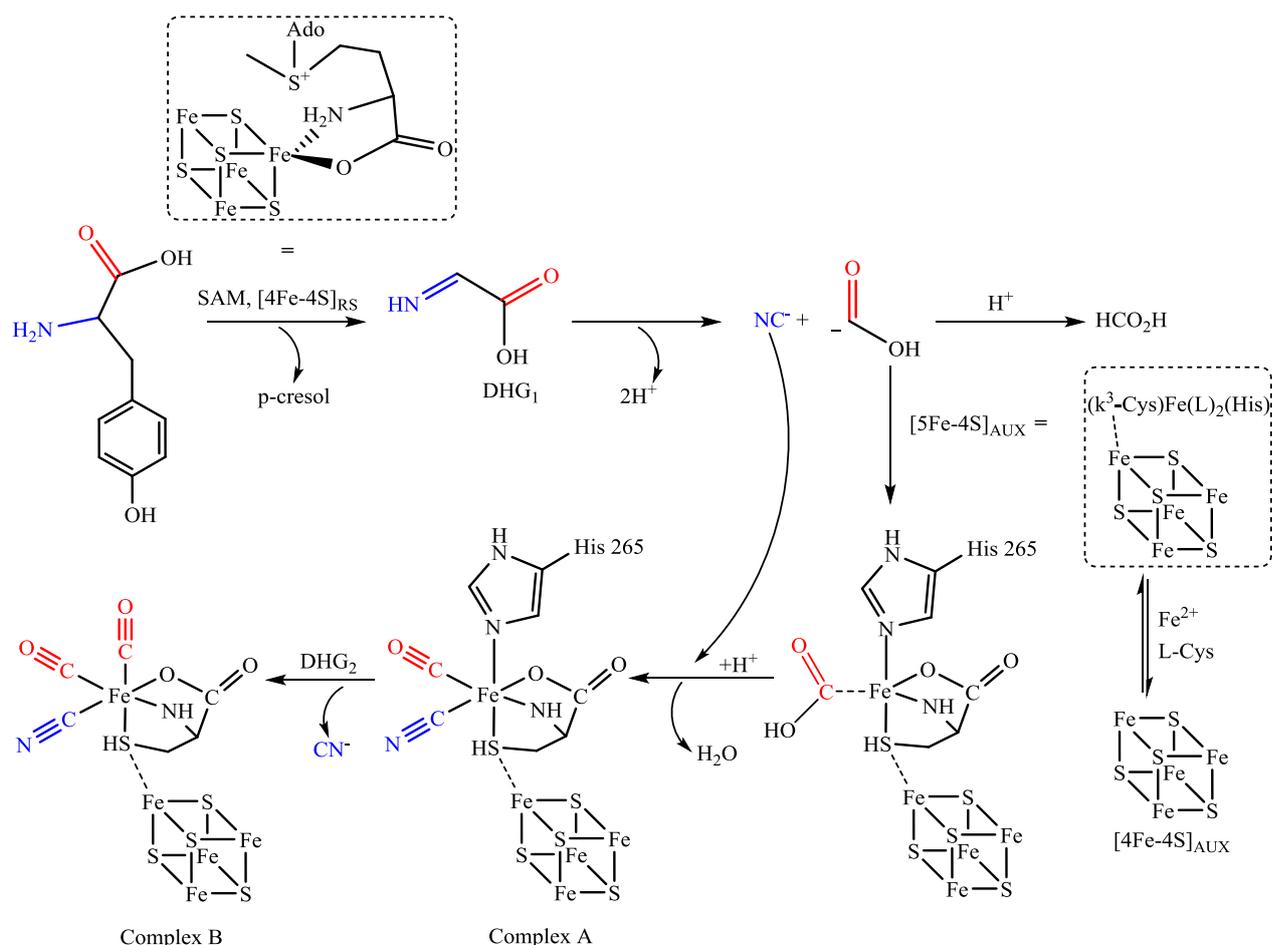


Figure 10: Proposed mechanism of HydG. Tyrosine is fragmented into p-cresol and dehydroglycine (DHG). DHG is then converted into CO (in red) and CN⁻ (in blue) which are coordinated by the “dangling Fe” of the [5Fe-4S]_{AUX} cluster. Two molecules of DHG (DHG₁ and DHG₂) are needed to provide the final product (complex B) with a Fe(CO)₂(CN) unit.

First, tyrosine binds close to the [4Fe-4S]_{RS} cluster. Radical H atom abstraction fragments it into the observed EPR active hydroxybenzyl radical, which ends up as p-cresol, and dehydroglycine (DHG). DHG migrates within the cavity from the [4Fe-4S]_{RS} cluster site to the auxiliary cluster where it undergoes an elimination reaction that simultaneously generates CN⁻ and the ⁻CO₂H intermediate. The latter binds to the dangling Fe, attached to the [4Fe-4S]_{AUX} cluster via its cysteine ligand, which catalyzes its conversion to CO and a water/OH⁻ molecule in a manner reminiscent of CO synthesis by anaerobic carbon monoxide dehydrogenase.⁶² Then, the CN⁻ ion displaces the water/OH⁻ ligand, generating complex A intermediate containing a Fe(CO)CN unit.⁵⁶ In the absence of the dangling Fe, the ⁻CO₂H intermediate is simply protonated to give formate. A second reaction cycle generates complex B, containing a [Fe(CO)₂(CN)] unit, the final product of the reaction. This unit is the key

Section B

component provided by HydG to HydF to assemble the diiron precursor of the [2Fe]-subcluster of HydA. The mechanism by which it is liberated from HydG is still unknown but might involve the second CN^- ion, used to displace the $[\text{Fe}(\text{CO})_2(\text{CN})]^-$ unit, for example in the form of a $[(\kappa^3\text{-Cys})\text{Fe}(\text{CO})_2(\text{CN})]^-$ complex, in agreement with the observation of a $[4\text{Fe}-4\text{S}]_{\text{AUX-CN}}$ cluster.

1.5.6 Mechanism of maturation of [FeFe]-hydrogenases

[FeFe]-hydrogenases, HydA, contain a unique active site, the H-cluster, as well as accessory [Fe-S] clusters. In cells they are first produced as a metal-free polypeptide chain. Maturation occurs in two steps. First the accessory clusters and the $[4\text{Fe}-4\text{S}]_{\text{H}}$ cluster are assembled thanks to the action of the ISC and SUF machineries specific for the biosynthesis of [Fe-S] clusters. Introduction of the [2Fe]-subcluster of the H-cluster occurs in a second step.

Three specific maturases, i.e. HydE, HydF and HydG, are responsible for this second maturation step. Due to the large amount of information available, described in this chapter, it is now possible to draw a general mechanism for this complex process (Figure 11). But it is unfortunately still incompletely defined. In particular, not only the substrate of HydE, the precursor of the adt^{2-} ligand, is unknown, as discussed above, but also the Hyd maturases have been mainly studied isolated. However, since small fragile organic and organometallic species are transferred from one protein into another, it is very likely that these transfers occur within, probably transient, complexes, which have, for the essential, not been investigated. This is probably the main direction to explore in further studies.

HydG is essential for assembling a first synthon, the $[\text{Fe}(\text{CO})_2(\text{CN})]^-$ unit, in which the CO and CN^- ligands derive from SAM-dependent radical decomposition of tyrosine. This process occurs within the coordination sphere of a dedicated [Fe-S] cluster, the $[4\text{Fe}-4\text{S}]_{\text{AUX}}$ cluster.

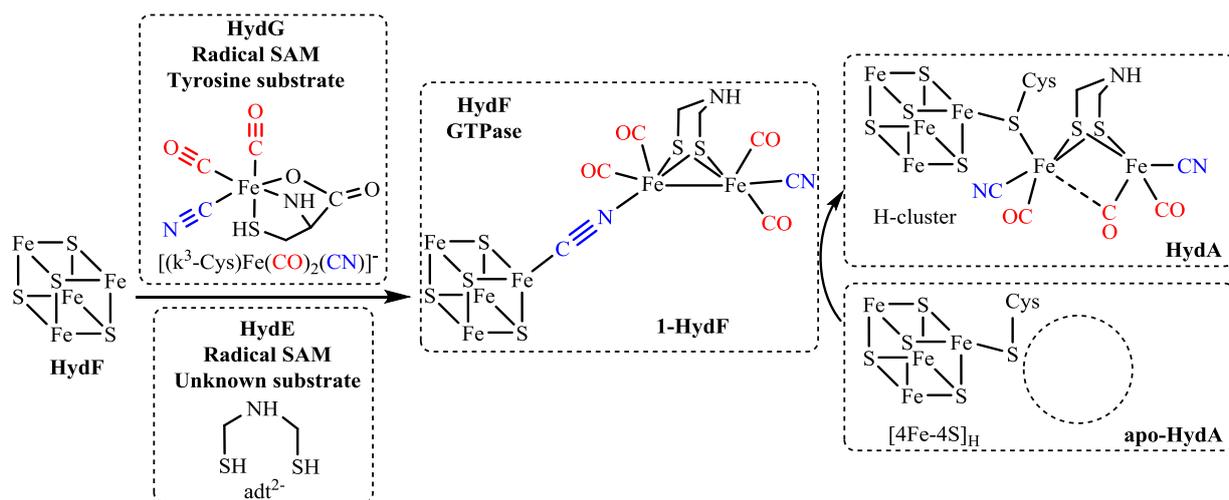


Figure 11: [FeFe]-hydrogenase (HydA) maturation process: assembly of the [2Fe]-subcluster. HydG and HydE produce the $[(\kappa^3\text{-Cys})\text{Fe}(\text{CO})_2(\text{CN})]^-$ and the adt^{2-} synthons, respectively. HydF captures these synthons, assembles a diiron precursor (presumably close to **1**) and transfers it to HydA where it converts into the catalytically active [2Fe]-subcluster.

Since HydF contains a [4Fe-4S] cluster which has also the ability to bind an organometallic species, such as the $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex,⁶ we propose that the $[\text{Fe}(\text{CO})_2(\text{CN})]^-$ unit from HydG is transferred to HydF, moving from the $[\text{4Fe-4S}]_{\text{AUX}}$ of HydG to the [4Fe-4S] cluster of HydF. Since HydF assembles a diiron complex, it is likely that the process described above occurs twice to generate a transient $[\text{Fe}_2(\text{CO})_4(\text{CN})_2]$ complex onto HydF. Then HydE catalyzes the formation of the last adt^{2-} component and transfers it to HydF to assemble the final $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex, in agreement with the observation of transient HydE-HydF complexes as well with their association within a single gene in some organisms.^{13,46} As discussed above, we indeed believe that the final product in HydF is indeed really close to the synthetic biomimetic complex **1** (Figure 11) since the **1**-HydF complex activates HydA in the absence of HydE and HydG. This is one possible scenario but it is still unknown at which step, early or late, adt^{2-} is attached to the Fe complex. For example, it has been shown that under certain conditions HydA could be matured in the absence of HydF.³⁶

This might indicate that HydE has also the ability to transfer adt^{2-} to Fe intermediate species bound to HydG, generating a $[\text{Fe}(\text{adt})(\text{CO})_2(\text{CN})]$ unit, subsequently transferred to HydF.

Finally, the assembled diiron unit in HydF is delivered to apo-HydA (lacking the [2Fe]-subcluster), leading to a fully matured and active enzyme. This also likely occurs within a

Section B

transient HydF-HydA complex which allows channelling the diiron complex from the [4Fe–4S] cluster of HydF to the [4Fe–4S]_H cluster of HydA. During that reaction: (i) a CO ligand is displaced allowing a cysteine ligand of the [4Fe–4S]_H cluster to bind to one Fe atom of the diiron complex; (ii) a rotation occurs leading to the formation of a CO bridge and liberating an open coordination site at the distal Fe for catalysis.

A striking and unprecedented feature of this machinery is the critical importance of [4Fe–4S] clusters for assembling and anchoring organometallic species, probably as a mechanism for protecting them from hydrolysis but also for allowing displacement from one cluster to another (Figure 12).

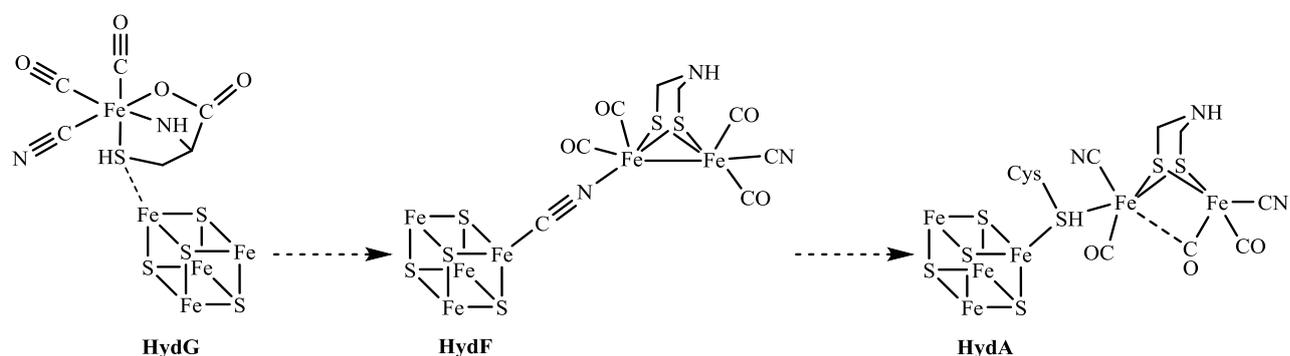


Figure 12: A common theme in HydA maturation: a [4Fe–4S] cluster controls organometallic species in HydG, HydF and HydA.

1.5.7 Redox states of the H-cluster of [FeFe]-hydrogenases

As described above, the H-cluster of [FeFe]-hydrogenases is composed of a [4Fe–4S] cluster, linked to a [2Fe]-subcluster through a cysteine bridge. Each Fe atom of the binuclear part has one CO and one CN[−] ligand and connects to the other Fe through an additional bridging CO ligand and an azadithiolate ligand. The two irons are called *proximal* (Fe_p) and *distal* (Fe_d) respectively.³

Both the cubane and the binuclear part of the H-cluster are redox active and during catalysis they can change their oxidation states. In particular the cubane has access to the redox states [4Fe–4S]⁺² and **[4Fe–4S]⁺** while the [2Fe]-subcluster can have different configurations, Fe_p(II)Fe_d(II), **Fe_p(I)Fe_d(II)** and Fe_p(I)Fe_d(I). Some of them are paramagnetic (shown in bold) and observable by EPR techniques.⁶³

Up to now only two [FeFe]-hydrogenases, HydA from *Desulfovibrio desulfuricans* (DdH) and from *C. reinhardtii* (CrHydA1), have been thoroughly studied for their redox

behavior and a combination of FTIR, EPR and Mössbauer techniques allowed the isolation and identification of various states of the H-cluster, which differ from each other substantially.

CO and CN ligands give rise to FTIR stretches in a zone free of other protein bands and their position is sensitive to the charge density on the binuclear core and can be used to observe the various H-cluster states.⁶⁴ In particular, the reduced states present lower frequencies of stretches due to the higher component of back-bonding. Moreover, the paramagnetic states of the H-cluster produce very characteristic EPR spectra facilitating their identification.

Not all the redox states are observable in all species. In the particular case of the aerobically purified DdH HydA, the H-cluster is found in the inactive air-stable H_{ox}^{air} state, where the cubane is in the $[4Fe-4S]^{2+}$ state and the binuclear in the Fe(II)Fe(II) state, thus EPR silent.

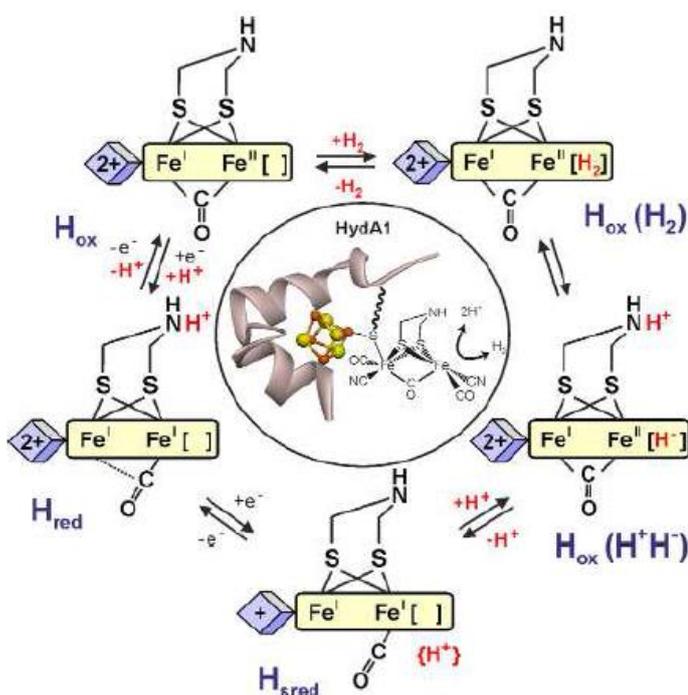
Mild reduction conditions lead to an intermediate redox state called H_{trans} , where the electron is stored on the cubane giving rise to the $[4Fe-4S]^+-Fe(II)Fe(II)$, which shows the typical spectrum of reduced $[4Fe-4S]^+$ cluster.⁶⁵ This redox state is isoelectronic with the H_{ox} state, $[4Fe-4S]^{2+}-Fe(I)Fe(II)$, which has peculiar FTIR and EPR signatures. The FTIR spectra of H_{ox} of various [FeFe]-hydrogenases present a characteristic signal for the bridging CO around 1800 cm^{-1} .³ The diiron subcluster, Fe(I)Fe(II), is paramagnetic ($S=1/2$) with similar g values ($g= [2.10, 2.04, 2.00]$) for all [FeFe]-hydrogenases while the $[4Fe-4S]_H$ is oxidized ($2+$) and thus EPR silent.

Upon inhibition by external CO gas, a new state is formed with the same electronic structure of H_{ox} , where a CO molecule binds the free coordination site on the 2Fe-subcluster. This H_{ox} -CO inhibited state gives rise to significant changes in the FTIR and EPR spectra. The EPR ($S=1/2$) signal is axial and is often observed in samples exposed to light and/or oxygen, where the CO ligands dissociated from the destroyed H-cluster are captured by active sites still intact (cannibalization).⁶⁶ While the H_{ox}^{air} and H_{trans} are not observed in any other [FeFe]-hydrogenase, the H_{ox} state and its CO inhibited version H_{ox} -CO are found in all [FeFe]-hydrogenases.³

Upon further reduction of the H_{ox} state, the H_{red} state is formed. Here, the cubane is in the $[4Fe-4S]^{2+}$ state and the binuclear in the Fe(I)Fe(I) and thus both are EPR silent. The FTIR stretches shift to lower wavenumbers confirming the reduction of Fe ions in the [2Fe]-subcluster.

Section B

In bacterial hydrogenases, containing accessory iron-sulfur clusters, the H_{red} state has EPR signals originated from reduced F-clusters. Recently, it has been proved that for CrHydA1 the H_{red} state can be additionally reduced to H_{sred} state (superreduced) giving rise to $[4Fe-4S]^+-Fe(I)Fe(I)$ configuration. This redox state is EPR active since the additional electron is stored on the cubane, producing a typical reduced $[4Fe-4S]^+$ cluster. The FTIR spectrum of CrHydA1- H_{sred} is similar to H_{red} of bacterial hydrogenases where the bridging CO becomes terminal.⁶³ This H_{sred} state has not been observed in hydrogenases with multiple iron-sulfur clusters and only in DdH HydA, a further reduction of H_{red} to H_{sred} is incomplete and irreversible while for CrHydA1 the reduction is complete and reversible. Thus, it has been proposed that H_{sred} state occurs in all hydrogenases but cannot be stabilized due to the electron redistribution between F and H-cluster, thermodynamically favored.



Both the active H_{ox} and H_{red} , observed in all hydrogenases, have been proposed as components of the catalytic cycle. The recent study on CrHydA1 led to the speculation that H_{sred} state takes part in the mechanism since it was shown to be a reversible and complete transition.⁶³ However it was observed only for the CrHydA1, thus it needs to be further elucidated.

1.5.8 Maturase-free Chemical maturation: a unique technological tool

As discussed above, HydF complexed with the synthetic $[Fe_2(adt)(CO)_4(CN)_2]^{2-}$ compound can fully activate apo-HydA without the need for HydE and HydG. Even simpler, T. Happe, W. Lubitz and M. Fontecave group found that *in vitro* treatment of apo-HydA, only lacking the $[2Fe]$ -subcluster, directly with the $[Fe_2(adt)(CO)_4(CN)_2]^{2-}$ complex (1), in the

Chapter I

absence of HydF, also results in full maturation and activation (Figure 14).⁷ The chemically matured HydA is indistinguishable from the same hydrogenase purified from its native host, as shown by FTIR and EPR spectroscopy as well by X-ray crystallography.^{7,67} This simple methodology opens a wide variety of potential applications.

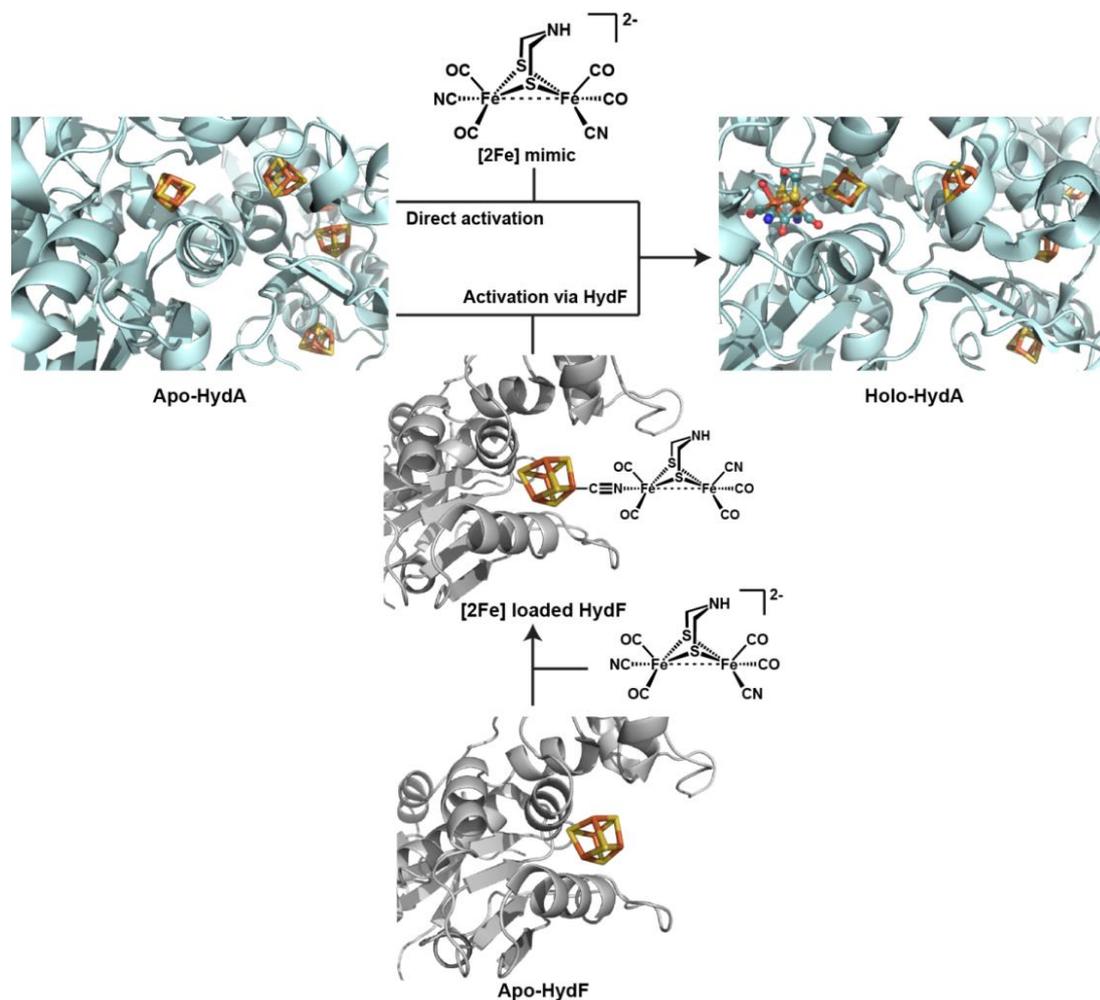


Figure 14: Activation of apoHydA by direct incorporation of [2Fe] model complex (top route) or via HydF maturase loaded with the [2Fe] synthetic mimic (bottom route). [4Fe–4S] clusters are shown as cubanes. Color code: iron (orange), sulfur (yellow), carbon (pale cyan), oxygen (red), and nitrogen (blue) (PDB code 3C8Y for holoHydA with modification at the bridgehead atom from oxygen to nitrogen).

First, it may greatly simplify the exploration of the biodiversity of hydrogenases and allow the investigation of a much larger ensemble of purified enzymes. Rather than using difficult anaerobic protocols to prepare low amounts of pure O₂-sensitive hydrogenases either from native hosts, often uneasy microorganisms to cultivate, or from microorganisms engineered for expression of the maturase complex system, it now becomes possible to express any *hydA* gene in *Escherichia coli*, lacking the HYD machinery, purify the inactive

Section B

recombinant protein anaerobically, and finally activate it in one step for assembly of the [2Fe]-subcluster. It has been shown to be efficient for the production of large amounts of fully active CrHydA1 and CpHydA.

Second, this methodology has allowed specific labelling of the [2Fe]-subcluster of HydA, which was otherwise unreachable. This has been used for incorporating $C^{15}N^-$, $^{13}CN^-$, ^{13}CO and N^{15} -labeled adt^{2-} ligands as well as ^{57}Fe atoms, thus allowing a detailed characterization of the electronic properties of this unique active site and the identification of specific spectroscopic probes for further characterization of the maturation process (for example the identification of the substrate of HydE discussed above).^{6,8,67,68}

Third, it allows the introduction, in the active site of HydA, of a variety of synthetic diiron organometallic complexes. This is a unique tool for understanding the function of each of the biological ligands. It also offers the opportunity to create new hydrogenases and to develop artificial hydrogenases.²⁹ For example, it has been possible to introduce a modified [2Fe]-subcluster in which the N atom of the adt^{2-} ligand has been changed to C and O. This nicely confirms the requirement for a nitrogen atom within the dithiolate bridge for activity, and also shows that the redox properties of the H-cluster and its reactivity with regard to CO can be tuned via synthetic modifications of this ligand.^{6,69,70} For example HydA with a pdt^{2-} ligand can access only two redox states (and not three with an adt^{2-} one) that can be characterized by FTIR spectroscopy. Furthermore, such a modified H-cluster is more resistant at high redox potential and in the presence of CO, raising opportunities regarding the design of new biomimetic catalysts. Other modifications have been introduced thanks to the chemical maturation methodology.⁶⁹ This includes further modifications of the dithiolate (thiopropanedithiolate, N-methylazapropanedithiolate, dimethyl-azapropanedithiolate) as well as variations in the CO/CN ratio within the coordination sphere of the synthetic complex. However, in all cases, the resulting HydA variants showed much less hydrogenase activity than the enzyme reconstituted with the $[Fe_2(ad)(CO)_4(CN)_2]^{2-}$ complex **1**, thus revealing a limited coordination flexibility of the [2Fe]-subcluster.⁶⁹ Chemical assembly of [Fe-S] clusters also allowed the incorporation of Se atoms within the $[4Fe-4Se]_H$ cluster, with very little effect on the activity.⁷¹

To become a significant component of the new “hydrogen economy” [FeFe]-hydrogenase enzymes need to be further investigated. In particular, their production at large scale and their O_2 -tolerance are issues to be addressed. A partial response resides in a still better understanding of the fascinating and highly complex process of [FeFe]-hydrogenase maturation, which has nevertheless greatly increased during the very last years. Indeed,

optimization of expression systems will depend on the engineering of the associated HYD maturation protein machinery. On the other hand, the discovery of the chemical maturation technology is a major breakthrough and might contribute to solve some of these issues, via simplification of hydrogenases expression/maturation methods, modification of the chemical structure of H-cluster of HydA and the development of artificial hydrogenases. The chemical maturation process might even have more impact if the discovery by Berggren and coworkers that apo-HydA can be matured *in cellulo*, within *E. coli* cells treated by the $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex **1** is confirmed (G. Berggren, personal communication), opening a new illustration of the potential of synthetic biology.

References:

- 1 J. W. Peters, W. N. Lanzilotta, B. J. Lemon and L. C. Seefeldt, *Science*, 1998, **282**, 1853–1858.
- 2 Y. Nicolet, C. Piras, P. Legrand, C. E. Hatchikian and J. C. Fontecilla-Camps, *Structure*, 1999, **7**, 13–23.
- 3 W. Lubitz, H. Ogata, O. Rüdiger and E. Reijerse, *Chem. Rev.*, 2014, **114**, 4081–4148.
- 4 J. W. Peters, G. J. Schut, E. S. Boyd, D. W. Mulder, E. M. Shepard, J. B. Broderick, P. W.
- 5 A. Silakov, B. Wenk, E. Reijerse and W. Lubitz, *Phys. Chem. Chem. Phys.*, 2009, **11**, 6592–6599.
- 6 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.
- 7 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.
- 8 A. Adamska-Venkatesh, S. Roy, J. F. Siebel, T. R. Simmons, M. Fontecave, V. Artero, E. Reijerse and W. Lubitz, *J. Am. Chem. Soc.*, 2015, **137**, 12744–12747.
- 9 D. W. Mulder, E. S. Boyd, R. Sarma, R. K. Lange, J. A. Endrizzi, J. B. Broderick and J. W. Peters, *Nature*, 2010, **465**, 248–251.
- 10 D. W. Mulder, D. O. Ortillo, D. J. Gardenghi, A. V. Naumov, S. S. Ruebush, R. K. Szilagyi, B. Huynh, J. B. Broderick and J. W. Peters, *Biochemistry*, 2009, **48**, 6240–6248.
- 11 G. Voordouw, W. R. Hagen, K. M. Krüse-Wolters, A. van Berkel-Arts and C. Veeger, *Eur. J. Biochem. FEBS*, 1987, **162**, 31–36.

Section B

- 12 B. Roche, L. Aussel, B. Ezraty, P. Mandin, B. Py and F. Barras, *Biochim. Biophys. Acta Bioenerg.*, 2013, **1827**, 923–937.
- 13 M. C. Posewitz, P. W. King, S. L. Smolinski, L. Zhang, M. Seibert and M. L. Ghirardi, *J. Biol. Chem.*, 2004, **279**, 25711–25720.
- 14 G. Vonabendroth, S. Stripp, A. Silakov, C. Croux, P. Soucaille, L. Girbal and T. Happe, *Int. J. Hydrog. Energy*, 2008, **33**, 6076–6081.
- 15 L. Girbal, G. von Abendroth, M. Winkler, P. M. C. Benton, I. Meynial-Salles, C. Croux, J. W. Peters, T. Happe and P. Soucaille, *Appl. Environ. Microbiol.*, 2005, **71**, 2777–2781.
- 16 P. W. King, M. C. Posewitz, M. L. Ghirardi and M. Seibert, *J. Bacteriol.*, 2006, **188**, 2163–2172.
- 17 K. Sybirna, T. Antoine, P. Lindberg, V. Fourmond, M. Rousset, V. Méjean and H. Bottin, *BMC Biotechnol.*, 2008, **8**, 73.
- 18 S. E. McGlynn, S. S. Ruebush, A. Naumov, L. E. Nagy, A. Dubini, P. W. King, J. B. Broderick, M. C. Posewitz and J. W. Peters, *J. Biol. Inorg. Chem.*, 2007, **12**, 443–447.
- 19 M. E. Boyer, J. A. Stapleton, J. M. Kuchenreuther, C.-W. Wang and J. R. Swartz, *Biotechnol. Bioeng.*, 2008, **99**, 59–67.
- 20 J. M. Kuchenreuther, J. A. Stapleton and J. R. Swartz, *PloS One*, 2009, **4**, e7565.
- 21 H. J. Sofia, G. Chen, B. G. Hetzler, J. F. Reyes-Spindola and N. E. Miller, *Nucleic Acids Res.*, 2001, **29**, 1097–1106.
- 22 J. B. Broderick, B. R. Duffus, K. S. Duschene and E. M. Shepard, *Chem. Rev.*, 2014, **114**, 4229–4317.
- 23 X. Brazzolotto, J. K. Rubach, J. Gaillard, S. Gambarelli, M. Atta and M. Fontecave, *J. Biol. Chem.*, 2006, **281**, 769–774.
- 24 J. K. Rubach, X. Brazzolotto, J. Gaillard and M. Fontecave, *FEBS Lett.*, 2005, **579**, 5055–5060.
- 25 S. E. McGlynn, E. M. Shepard, M. A. Winslow, A. V. Naumov, K. S. Duschene, M. C. Posewitz, W. E. Broderick, J. B. Broderick and J. W. Peters, *FEBS Lett.*, 2008, **582**, 2183–2187.
- 26 E. M. Shepard, S. E. McGlynn, A. L. Bueling, C. S. Grady-Smith, S. J. George, M. A. Winslow, S. P. Cramer, J. W. Peters and J. B. Broderick, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 10448–10453.
- 27 I. Czech, A. Silakov, W. Lubitz and T. Happe, *FEBS Lett.*, 2010, **584**, 638–642.
- 28 I. Czech, S. Stripp, O. Sanganas, N. Leidel, T. Happe and M. Haumann, *FEBS Lett.*, 2011, **585**, 225–230.

Chapter I

- 29 V. Artero, G. Berggren, M. Atta, G. Caserta, S. Roy, L. Pecqueur and M. Fontecave, *Acc. Chem. Res.*, 2015, **48**, 2380–2387.
- 30 L. Cendron, P. Berto, S. D'Adamo, F. Vallese, C. Govoni, M. C. Posewitz, G. M. Giacometti, P. Costantini and G. Zanotti, *J. Biol. Chem.*, 2011, **286**, 43944–43950.
- 31 P. Berto, M. Di Valentin, L. Cendron, F. Vallese, M. Albertini, E. Salvadori, G. M. Giacometti, D. Carbonera and P. Costantini, *Biochim. Biophys. Acta*, 2012, **1817**, 2149–2157.
- 32 G. Berggren, R. Garcia-Serres, X. Brazzolotto, M. Clemancey, S. Gambarelli, M. Atta, J.-M. Latour, H. L. Hernández, S. Subramanian, M. K. Johnson and M. Fontecave, *J. Biol. Inorg. Chem.*, 2014, **19**, 75–84.
- 33 M. Albertini, L. Galazzo, L. Maso, F. Vallese, P. Berto, E. De Rosa, M. Di Valentin, P. Costantini and D. Carbonera, *Top. Catal.*, 2015, **58**, 708–718.
- 34 M. Albertini, P. Berto, F. Vallese, M. Di Valentin, P. Costantini and D. Carbonera, *J. Phys. Chem. B*, 2015, **119**, 13680–13689.
- 35 N. Joshi, E. M. Shepard, A. S. Byer, K. D. Swanson, J. B. Broderick and J. W. Peters, *FEBS Lett.*, 2012, **586**, 3939–3943.
- 36 J. M. Kuchenreuther, R. D. Britt and J. R. Swartz, *PloS One*, 2012, **7**, e45850.
- 37 E. M. Shepard, A. S. Byer, J. N. Betz, J. W. Peters and J. B. Broderick, *Biochemistry*, 2016, **55**, 3514–3527.
- 38 M. J. Lacasse and D. B. Zamble, *Biochemistry*, 2016, **55**, 1689–1701.
- 39 M. Blokesch, A. Paschos, E. Theodoratou, A. Bauer, M. Hube, S. Huth and A. Böck, *Biochem. Soc. Trans.*, 2002, **30**, 674–680.
- 40 W. B. Jeon, J. Cheng and P. W. Ludden, *J. Biol. Chem.*, 2001, **276**, 38602–38609.
- 41 M. B. Moncrief and R. P. Hausinger, *J. Bacteriol.*, 1997, **179**, 4081–4086.
- 42 R. M. Allen, M. J. Homer, R. Chatterjee, P. W. Ludden, G. P. Roberts and V. K. Shah, *J. Biol. Chem.*, 1993, **268**, 23670–23674.
- 43 Y. Nicolet, J. K. Rubach, M. C. Posewitz, P. Amara, C. Mathevon, M. Atta, M. Fontecave and J. C. Fontecilla-Camps, *J. Biol. Chem.*, 2008, **283**, 18861–18872.
- 44 Y. Nicolet, R. Rohac, L. Martin and J. C. Fontecilla-Camps, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 7188–7192.
- 45 J. N. Betz, N. W. Boswell, C. J. Fugate, G. L. Holliday, E. Akiva, A. G. Scott, P. C. Babbitt, J. W. Peters, E. M. Shepard and J. B. Broderick, *Biochemistry*, 2015, **54**, 1807–1818.
- 46 F. Vallese, P. Berto, M. Ruzzene, L. Cendron, S. Sarno, E. De Rosa, G. M. Giacometti and P. Costantini, *J. Biol. Chem.*, 2012, **287**, 36544–36555.

Section B

- 47 F. Berkovitch, Y. Nicolet, J. T. Wan, J. T. Jarrett and C. L. Drennan, *Science*, 2004, **303**, 76–79.
- 48 Y. Nicolet, P. Amara, J.-M. Mouesca and J. C. Fontecilla-Camps, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 14867–14871.
- 49 R. Rohac, P. Amara, A. Benjdia, L. Martin, P. Ruffié, A. Favier, O. Berteau, J.-M. Mouesca, J. C. Fontecilla-Camps and Y. Nicolet, *Nat. Chem.*, 2016, **8**, 491–500.
- 50 M. Kriek, F. Martins, M. R. Challand, A. Croft and P. L. Roach, *Angew. Chem. Int. Ed.*, 2007, **46**, 9223–9226.
- 51 E. Pilet, Y. Nicolet, C. Mathevon, T. Douki, J. C. Fontecilla-Camps and M. Fontecave, *FEBS Lett.*, 2009, **583**, 506–511.
- 52 E. M. Shepard, B. R. Duffus, S. J. George, S. E. McGlynn, M. R. Challand, K. D. Swanson, P. L. Roach, S. P. Cramer, J. W. Peters and J. B. Broderick, *J. Am. Chem. Soc.*, 2010, **132**, 9247–9249.
- 53 R. C. Driesener, M. R. Challand, S. E. McGlynn, E. M. Shepard, E. S. Boyd, J. B. Broderick, J. W. Peters and P. L. Roach, *Angew. Chem. Int. Ed.*, 2010, **49**, 1687–1690.
- 54 J. M. Kuchenreuther, S. J. George, C. S. Grady-Smith, S. P. Cramer and J. R. Swartz, *PLoS One*, 2011, **6**, e20346.
- 55 J. M. Kuchenreuther, W. K. Myers, T. A. Stich, S. J. George, Y. Nejatyjahromy, J. R. Swartz and R. D. Britt, *Science*, 2013, **342**, 472–475.
- 56 J. M. Kuchenreuther, W. K. Myers, D. L. M. Suess, T. A. Stich, V. Pelmeshnikov, S. A. Shiigi, S. P. Cramer, J. R. Swartz, R. D. Britt and S. J. George, *Science*, 2014, **343**, 424–427.
- 57 P. Dinis, D. L. M. Suess, S. J. Fox, J. E. Harmer, R. C. Driesener, L. De La Paz, J. R. Swartz, J. W. Essex, R. D. Britt and P. L. Roach, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 1362–1367.
- 58 D. L. M. Suess, I. Bürstel, L. De La Paz, J. M. Kuchenreuther, C. C. Pham, S. P. Cramer, J. R. Swartz and R. D. Britt, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 11455–11460.
- 59 D. L. M. Suess, C. C. Pham, I. Bürstel, J. R. Swartz, S. P. Cramer and R. D. Britt, *J. Am. Chem. Soc.*, 2016, **138**, 1146–1149.
- 60 A. Pagnier, L. Martin, L. Zeppieri, Y. Nicolet and J. C. Fontecilla-Camps, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 104–109.
- 61 Y. Nicolet, L. Zeppieri, P. Amara and J. C. Fontecilla-Camps, *Angew. Chem. Int. Ed.*, 2014, **53**, 11840–11844.
- 62 P. Amara, J.-M. Mouesca, A. Volbeda and J. C. Fontecilla-Camps, *Inorg. Chem.*, 2011, **50**, 1868–1878.

Chapter I

- 63 A. Adamska, A. Silakov, C. Lambertz, O. Ruediger, T. Happe, E. Reijerse and W. Lubitz, *Angew. Chem.-Int. Ed.*, 2012, **51**, 11458–11462.
- 64 Y. Nicolet, A. L. de Lacey, X. Vernède, V. M. Fernandez, E. C. Hatchikian and J. C. Fontecilla-Camps, *J. Am. Chem. Soc.*, 2001, **123**, 1596–1601.
- 65 S. P. J. Albracht, W. Roseboom and E. C. Hatchikian, *J. Biol. Inorg. Chem.*, 2006, **11**, 88–101.
- 66 W. Roseboom, A. L. De Lacey, V. M. Fernandez, E. C. Hatchikian and S. P. J. Albracht, *J. Biol. Inorg. Chem.*, 2006, **11**, 102–118.67 J. Esselborn, N. Muraki, K. Klein, V. Engelbrecht, N. Metzler-Nolte, U.-P. Apfel, E. Hofmann, G. Kurisu and T. Happe, *Chem Sci*, 2016, **7**, 959–968.
- 68 R. Gilbert-Wilson, J. F. Siebel, A. Adamska-Venkatesh, C. C. Pham, E. Reijerse, H. Wang, S. P. Cramer, W. Lubitz and T. B. Rauchfuss, *J. Am. Chem. Soc.*, 2015, **137**, 8998–9005.
- 69 J. F. Siebel, A. Adamska-Venkatesh, K. Weber, S. Rumpel, E. Reijerse and W. Lubitz, *Biochemistry*, 2015, **54**, 1474–1483.
- 70 A. Adamska-Venkatesh, D. Krawietz, J. Siebel, K. Weber, T. Happe, E. Reijerse and W. Lubitz, *J. Am. Chem. Soc.*, 2014, **136**, 11339–11346.
- 71 J. Noth, J. Esselborn, J. Güldenhaupt, A. Brünje, A. Sawyer, U.-P. Apfel, K. Gerwert, E. Hofmann, M. Winkler and T. Happe, *Angew. Chem. Int. Ed.*, 2016, **55**, 8396–8400.

Section B

Section C

This section of chapter I has been centered on biohybrid and supramolecular systems for catalytic hydrogen production and uptake. All the reported examples are part of a recently published review article where we described, first, systems based on amino acids/peptides/proteins but then also expand to non-biological polymers, which have been used as scaffolds for the assembly of artificial enzymes. All the reported systems displayed hydrogenase-like catalytic activities.

1.6 Artificial hydrogenases

As soon as structural and functional details of the active site of hydrogenases were known, notably through X-ray crystallography, production of biomimetic and bioinspired catalysts became a major field of research in bioinorganic chemistry.¹ This has led to hundreds of easily accessible synthetic diiron and Ni-based complexes mimicking the organometallic active sites of the two major classes of hydrogenases, the [FeFe]- and [NiFe]-hydrogenases, respectively (Figure 15).

Even though no natural hydrogenase containing cobalt in its active site is known, a number of cobalt complexes have also been described as hydrogenase mimics and should be considered as well.² Unfortunately and in spite of remarkable progress, this field has not generated the perfect catalyst yet. As a matter of fact these complexes are in general less active (larger overpotentials and/or lower catalytic rates) than the corresponding enzymes, generally insoluble in aqueous solutions and often unstable during catalysis.³

This has recently led to the emergence of projects aiming at developing artificial hydrogenases. The concept of artificial enzymes finds its origin in the pioneering work of G. Whitesides⁴ and has already been applied in many different contexts.⁵ In the specific case of hydrogenases, it implies that a well-chosen, easily accessible host protein or polypeptide can be combined to a synthetic complex with proton reduction/H₂ oxidation catalytic activity, in order to assemble a cheap, stable, water-soluble and active hybrid system. It is indeed expected that the protein will provide an environment of functional groups (acids, bases, redox centers...) around the synthetic complex facilitating proton and electron transfers. One advantage of such a system resides in the possibility to rationally tune and thus optimize the catalytic performances through manipulation of both the synthetic component, by synthetic

Section C

variations of the ligand, and the interactions with the protein cavity, in which the latter is embedded, by modifying amino acid side chains of the protein through site-directed mutagenesis. There is hope that, through this approach, a new, easily affordable metalloenzyme could be generated with excellent hydrogenase activity and stability, thus with potential for application in technological devices. It is important to note that whereas the choice of the catalytic synthetic component is trivial (since a number of such catalytic, bioinspired or not, complexes are available), there is no clear rationale for guiding the choice of the protein, besides the fact that the protein should contain a cavity or chelating lateral chains preorganized to selectively bind the catalyst.

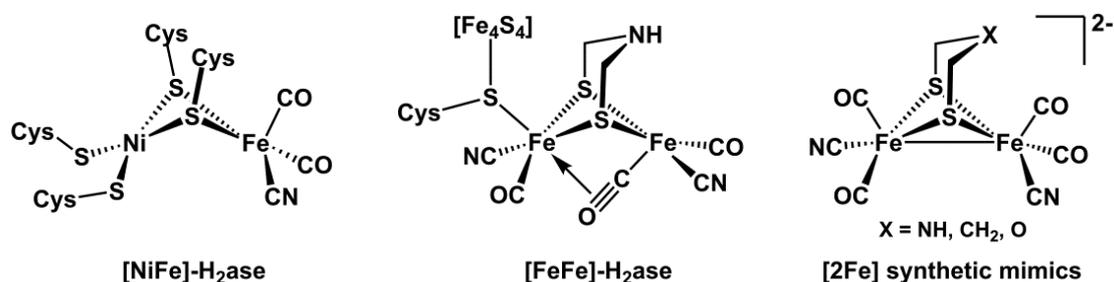


Figure 15: Schematic representation of the active sites of [NiFe]-hydrogenase, [FeFe]-hydrogenase (H-cluster), and [2Fe] synthetic mimics.

1.6.1 Artificial hydrogenases based on synthetic diiron complexes

Synthetic diiron complexes mimicking the active site of [FeFe]-hydrogenases have been incorporated into other peptides and proteins. It should be noted that, in most cases, these studies were based on diiron complexes lacking both cyanide ligands and a pendant basic site in the close vicinity of the iron atoms, either as an amine bridgehead group of the dithiolate ligand or in any other position. The presence of such proton relays is crucial for the high catalytic rate observed for the [FeFe]-hydrogenases and its absence in peptide-diiron complexes is expected to result in substantially lower catalytic activity. For example Ghirlanda and collaborators synthesized an artificial amino acid containing a 1,3-propanedithiol side chain (dithiol amino acid, Dt) and incorporated it into an alanine-rich 19-mer peptide, known to generate a helical structure.⁶ The dithiolate group served to anchor a {Fe₂CO₆} unit into the helical peptide (Figure 16a).

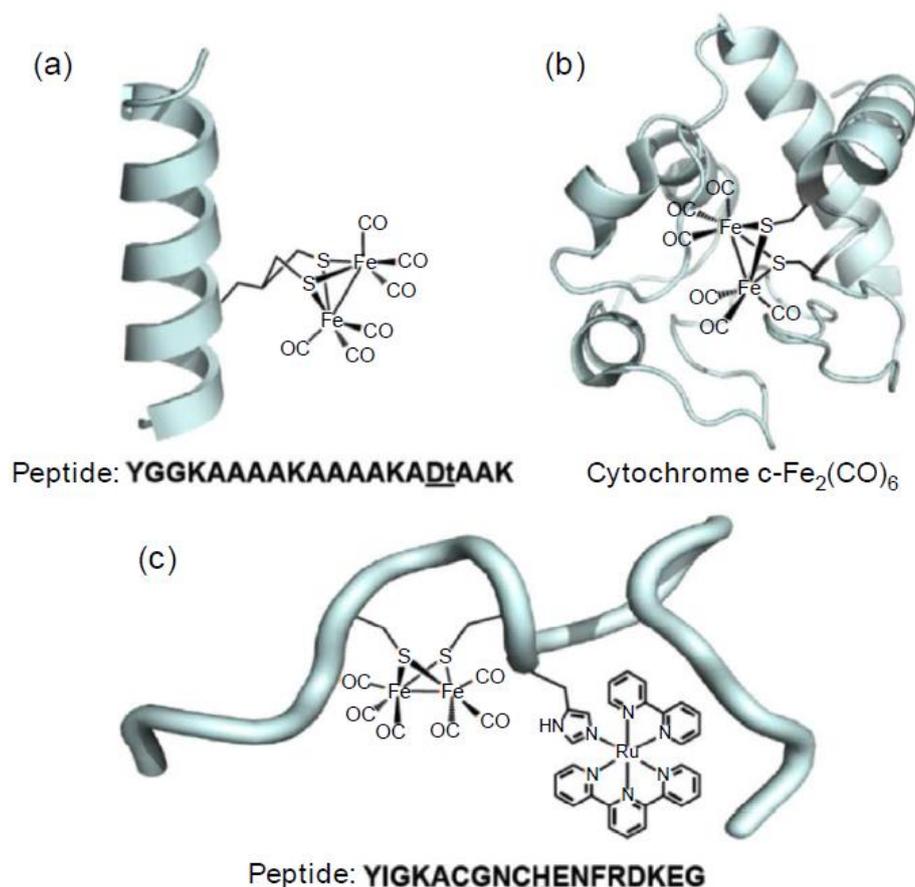


Figure 16: Model representations of [2Fe] biohybrids where the diironhexacarbonyl complex is attached to (a) a synthetic helical peptide, (b) apo-cytochrome c, and (c) an octapeptide derived from cytochrome c.

The resulting hybrid system features a pH-dependent irreversible reduction process at -1.1 V vs. SHE. It was shown to catalyze light-driven H₂ evolution at pH 4.5 in aqueous buffer and in the presence of [Ru(bpy)₃]²⁺ as the photosensitizer and ascorbate as the sacrificial electron donor, resulting in a total TON of 84 (with regard to the catalyst) in 2.3 hours. Similarly, Hayashi and coworkers demonstrated that hybrid systems could be prepared simply by treating apo-cytochrome c (Cyt c) with [Fe₂(CO)₉].⁷ Mass spectrometry and FTIR spectroscopy indicated a $\{(\mu\text{-S}_{\text{cys}})_2\text{Fe}_2(\text{CO})_6\}$ structure, thus showing that the diiron unit binds to two adjacent cysteine residues of cytochrome c (Figure 16b). This biohybrid complex solution displayed catalytic activity for hydrogen production from water (82 turnovers in two hours at pH 4.7) when irradiated in the presence of [Ru(bpy)₃]²⁺ and ascorbate. The activity is dependent on the pH of the solution: only 5.6 turnovers are reported when the pH value is raised to 7.4. Much lower activity (TON = 10 at pH 4.7) was also obtained when cytochrome c was replaced by a cytc-derived heptapeptide fragment containing the Cys-X-X-Cys sequence, illustrating the importance of a complete protein environment. In another study, an

octadecapeptide fragment derived from cytochrome c556 and containing a Cys-X-X-Cys-His sequence was used to attach both the diiron carbonyl cluster to cysteines and a Ru-based photosensitizer, $[\text{Ru}(\text{tpy})(\text{bpy})(\text{His})]^{2+}$, to histidine (Figure **16c**).⁸ In fact, the direct coordination of the Ru complex next to the diiron complex was found to be crucial for photocatalysis. Visible light irradiation of the $[\text{2Fe}][\text{Ru}]$ -peptide at pH 8.5 in the presence of ascorbate resulted in catalytic H_2 evolution with a total TON of 9 in two hours, at which point H_2 production stopped. When the reaction was performed under similar conditions using the same peptide containing only the diiron complex and free $[\text{Ru}(\text{tpy})(\text{bpy})(\text{Imidazole})]^{2+}$ as the photosensitizer, no H_2 production was observed.

Diiron mimics have also been frequently evaluated in non-protein confined environments. Some of the best characterized systems are briefly described below.

1.6.1.1 Micelles

Sodium dodecyl sulfate (SDS) micelles, which are single lipid layer systems featuring a negatively charged surface and a hydrophobic interior, have been found useful for solubilizing and activating such catalysts.⁹⁻¹² For example $[(\mu\text{-bdt})\text{Fe}_2(\text{CO})_6]$ ($\text{bdt}^{2-} = 1,2$ -benzenedithiolate), a slow but robust catalyst for H_2 evolution in organic solvents was shown to become water soluble and to display higher activity in the presence of SDS. Electrocatalytic H_2 evolution was observed at -0.7 V vs. SHE (500 mV overpotential) in a pH 3.3 aqueous micellar solution and controlled potential electrolysis at -0.66 V vs. SHE at pH 3, led to 52 turnovers during the first hour with no noticeable degradation of the catalyst.¹¹ In a follow-up study, micellar solutions of $[(\mu\text{-bdt})\text{Fe}_2(\text{CO})_6]$ were evaluated for H_2 evolution in fully aqueous solution under photocatalytic conditions with Eosin Y and rose bengal as photosensitizers and triethylamine (TEA) as the sacrificial donor, with significant stability enhancement due to the inclusion into SDS micelles.¹³

1.6.1.2 Dendrimers

A $[(\mu\text{-SR})_2\text{Fe}_2(\text{CO})_6]$ core has also been anchored within the hydrophobic pocket of a dendrimer by Yang and Li and co-workers.¹⁴ This was achieved by covalent attachment of the sulfide ligands of the organometallic core to a Fréchet-type dendritic moiety, consisting of aryl-benzyl ethers, to yield Hy-G4 (Figure **17**).

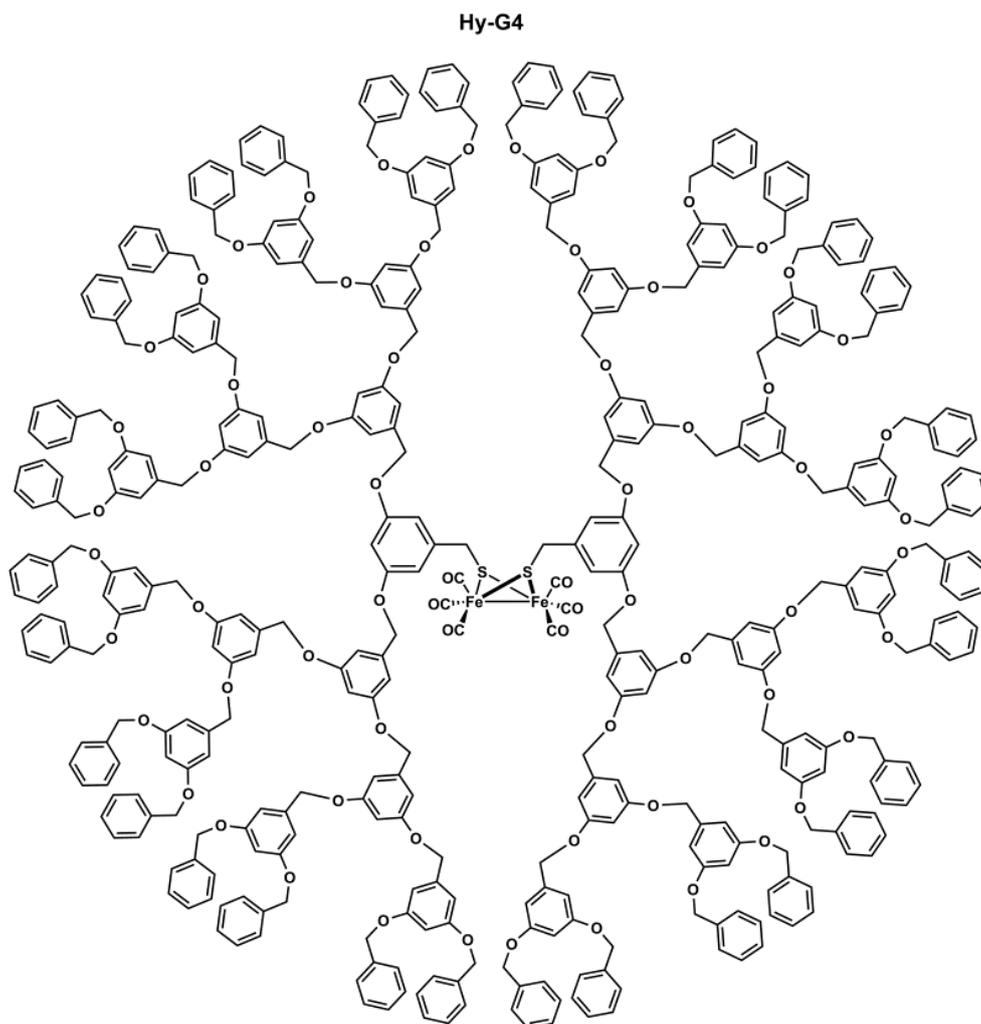


Figure 17: Structure of the dendrimeric [2Fe]mimic Hy-G4 featuring a $[(m-SR)_2Fe_2(CO)_6]$ cluster.

Photocatalytic hydrogen production was performed in an acetone/water mixture (9:1) in the presence of TEA as sacrificial electron donor and $[Ir(ppy)_2(bpy)]^+$ as the photosensitizer. The TOF and total TON values of the system increased with the size of the dendritic structure. For the largest structure, the system had an impressive TON of 22200, with an initial TOF of 2 s^{-1} during the first hour of the reaction.

1.6.1.3 Polymers

Polyacrylic acid (PAA) is a hydrophobic polymer, in which the carboxyl groups not only enhance water solubility but also provide modification sites for functionalization. Using an aromatic linker containing an amino group at one end for an amide linkage and an isocyanide group at the other end as a ligand to iron, Wu and collaborators were able to attach a diiron mimic to PAA generating a catalyst with exceptional performances.¹⁵ The successful

Section C

incorporation of the complex was verified using NMR, FTIR, UV/Vis and inductively coupled plasma-atomic emission spectrometry (ICP-AES). As expected, the PAA framework improved the water solubility of the complex and allowed the photocatalytic studies to be conducted in aqueous solution. In the presence of 0.1 M ascorbic acid (at an initial pH of 4.0) and CdSe quantum dots employed as photosensitizers, the functionalized polymer was found to produce H₂ under irradiation at 430 nm (TON = 27,000 after 8 h, with an initial TOF value of 3.6 s⁻¹). It should be noted that the observed efficiency is likely attributable not only to improved catalytic efficiency of the diiron unit, but also to some extent to the interaction between the PAA and the CdSe photosensitizer, which prevents aggregation of the nanoparticles and ensures close contact between the sensitizer and the catalyst.

1.6.1.4 Oligo-/poly-saccharides

Dinuclear iron complexes in which the bridgehead nitrogen atom has been functionalized to incorporate an aryl sulfonate group are prone to inclusion within β -cyclodextrins.^{16,17}

Cyclodextrins are cyclic oligosaccharides, consisting of 6-8 glucose units (denoted α -, β - and γ -cyclodextrins respectively), generating a hydrophobic cavity with a hydrophilic rim. The adducts formed between β -cyclodextrins and various [FeFe]-hydrogenase mimics were characterized by different methods including X-ray crystallography, making this class of guest-host complexes one of the few systems discussed herein for which crystal structures have been reported.¹⁶ The effect of cyclodextrins was clearly shown for example in a photocatalytic system using Eosin Y as a photosensitizer and TEA as an electron donor at pH 10 in the case of $[(\mu-(\text{SCH}_2)_2\text{NC}_6\text{H}_4\text{SO}_3)\text{Fe}_2(\text{CO})_6]^-$: addition of a 10-fold excess of γ -cyclodextrin led to a 9-fold increase of the TON value (from 8 to 75 mol H₂·mol catalyst⁻¹).¹⁸ More recently, a polysaccharide, chitosan, was shown to greatly stimulate the activity of diiron mimics, such as $[\text{Fe}_2(\text{CO})_6(\mu-(\text{SCH}_2)_2\text{NCH}_2\text{C}_6\text{H}_5)]$.¹⁹ Chitosan contains a significant content of primary amines and hydroxyl groups. When the amines are protonated by acids, chitosan behaves as a polycation. In this form it binds not only the complex but also the negatively charged 3-mercaptopropionic acid-capped CdTe quantum dots used as a photosensitizer. The system displays H₂ production, using ascorbic acid as a sacrificial electron donor, in H₂O/methanol (3:1, v/v) pH 4.5 with a catalytic activity 4000-fold higher

than that of the same system without chitosan (TON and initial TOF values of 50000 and 1.4 s⁻¹ respectively) under visible light irradiation.

1.6.1.5 Metal-Organic-Frameworks

Metal-organic-Frameworks (MOFs) have emerged as an interesting class of microporous crystalline materials with potential applications in catalysis. Unlike other porous materials, the organic ligand allows for functionalization of internal channels and cavities and for incorporation of homogeneous catalysts. Very recently this strategy has been used to incorporate dinuclear iron complexes within MOFs and thus to generate hydrogen evolution heterogeneous catalysts, which have the advantage of being easily separated, reusable and more stable. For example $[\text{Fe}_2(\text{dcdbt})(\text{CO})_6]^{2-}$ (dcdbt= terephthalate-2,3-dithiolate) has been incorporated into a highly robust Zr(IV)-based MOF. The functionalized MOF proved to be more active than the molecular precursor in photochemical H₂ production in water at pH 5, in the presence of a Ru photosensitizer and ascorbate. However the system produces only a few turnovers and stops after about one hour.²⁰

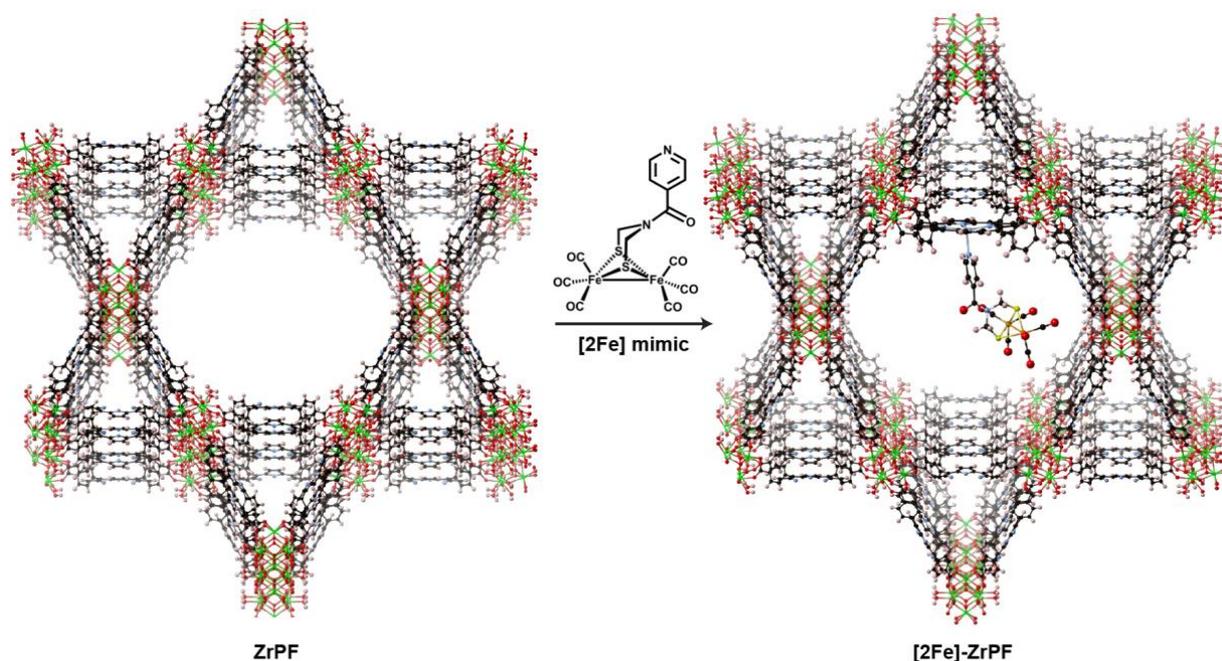


Figure 18: Model structures of zirconium-porphyrin MOF before (ZrPF) and after ([2Fe]-ZrPF) incorporation of [2Fe] mimic.

In a different configuration a diiron complex was coordinatively attached to the Zn-porphyrin linker, used as an integrated photosensitizing unit, of a Zr(IV)/Zn-porphyrin MOF

Section C

(Figure 18). Under visible light irradiation in acetate buffer and in the presence of ascorbate the system produces hydrogen during about 2 hours.²¹ Within the MOF the diiron catalyst proved much more active and stable than alone in solution.

1.6.2 Ni-based artificial hydrogenases

So far very few active synthetic dinuclear Ni-Fe mimics have been reported.^{1,22} On the other hand, mononuclear $[\text{Ni}(\text{P}^{\text{R}}_2\text{N}^{\text{R}'})_2]^{2+}$ complexes bearing diphosphine $\text{P}^{\text{R}}_2\text{N}^{\text{R}'}$ ligands, developed by DL. DuBois, are among the most efficient bioinspired catalysts for hydrogen production and furthermore form the only series of complexes capable of catalytic H_2 oxidation.²³ Some of these catalysts are shown in Figure 19.

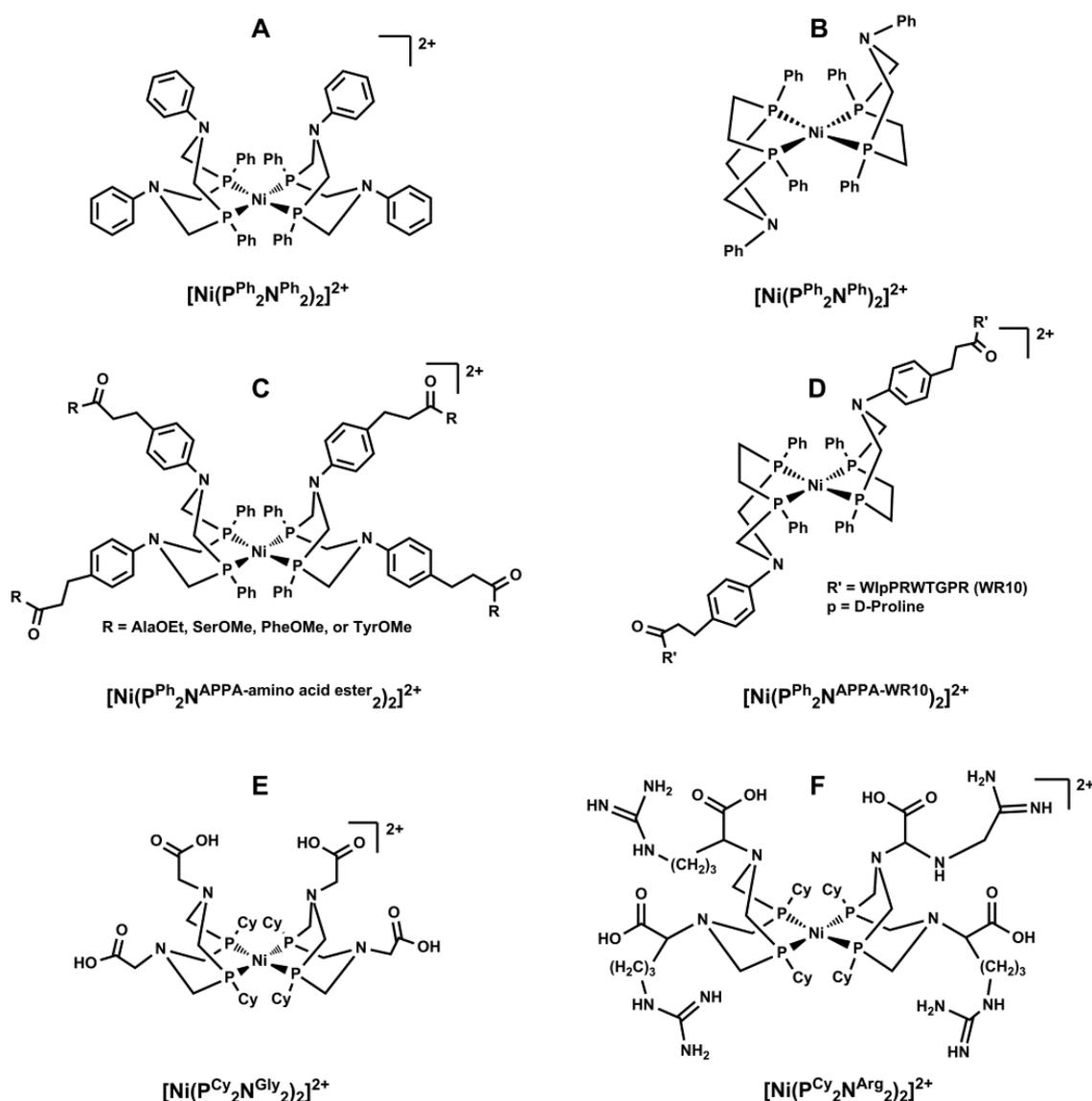


Figure 19: Schematic representation of selected nickel complexes bearing $\text{P}^{\text{R}}_2\text{N}^{\text{R}'}$ ($n = 1$ or 2) ligands.

The pendant amine groups incorporated into the diphosphine ligand of the Ni ion, inspired by the N atom of the bridging dithiolate ligand at the active site of [FeFe]-hydrogenases, act as essential proton relays allowing high catalytic rates for H^+/H_2 interconversion at moderate overpotentials. This is probably one of the most brilliant illustrations of the importance of the presence of bioinspired functional groups in the second coordination sphere of the metal center in mimics of hydrogenases. These complexes have recently been exploited by W. Shaw and collaborators as platforms to study more systematically the effect of enzyme-like functionalities, such as amino acids and amino-ester derivatives as well as peptides, in the outer coordination sphere on the catalytic rates.²⁴

Starting from the $[\text{Ni}(\text{P}^{\text{Ph}}_2\text{N}^{\text{Ph}}_2)_2]^{2+}$ complex, the N-phenyl has been replaced by 4-aminophenyl propionic acid (APPA) and the terminal acid was exploited for the introduction of diverse amino acids and their esters, thus allowing a structure-function relationships study (Figure **19C**).^{25–28} In some cases, complexes with acidic or basic side chains ($[\text{Ni}(\text{P}^{\text{Ph}}_2\text{N}^{\text{APPA-Lys}}_2)_2]^{2+}$, 3350 s^{-1}) displayed up to five times increase in the rate of electrocatalytic H_2 evolution in water/acetonitrile mixtures compared to the unmodified parent complex ($[\text{Ni}(\text{P}^{\text{Ph}}_2\text{N}^{\text{Ph}}_2)_2]^{2+}$, 720 s^{-1}).^{26,29} Similarly, terminal propionic acid functions were used to introduce a highly structured, sterically constrained, β -hairpin 10-mer peptide (denoted as WR10 in Figure **19D**) in the structurally related $[\text{Ni}(\text{P}^{\text{Ph}}_2\text{N}^{\text{Ar}})_2]^{2+}$ complex (the diphosphine ligands used here contain only one amine group). Whereas the peptide system was shown to operate at twice the rate (100000 s^{-1}) of the parent peptide-free complex in a 90/10 mixture of CH_3CN and water, it proved too flexible to allow consistent studies of the correlation between the activity and the nature of the protein-based second coordination sphere.²⁸ The above mentioned derivatives, with phenyl-phosphine ligands, are only active for H_2 evolution in organic solvents with overpotential requirements ranging from 300 mV to 540 mV.^{27,28}

In this series of complex however, modifications in the first and second coordination spheres allow to tune the driving force for H_2 addition to the Ni complex.³⁰ As an example, introduction of cyclohexyl-phosphine ligands together with a second set of proton relays in the outer coordination sphere of the Ni complexes has been shown to generate catalysts with the ability to function for both H_2 evolution and H_2 oxidation.^{31,32} This has been first achieved by incorporating the amine group of glycine directly into the $\text{P}^{\text{Cy}}_2\text{N}^{\text{R}}_2$ ligand, thus generating a fully water soluble complex, active for hydrogen oxidation from pH 0.1 to 9, with a maximum rate of 33 s^{-1} and minimum overpotential requirement of 150 mV at pH 0.7–1.0 (Figure **19E**).³¹ Such low pH conditions correspond to those at work in proton-membrane fuel cells. It is proposed that the combination of a carboxylate of glycine and the pendant amine groups

Section C

provides a two-relay proton channel optimizing proton transfers coupled to electron transfers between the Ni center and the surface of the complex, as in the respective enzymes.

The same strategy has been used to generate complexes in which glycine was replaced by arginine, as guanidinium groups might play the role of a third proton relay (Figure 19F).³² The arginine analogue $[\text{Ni}(\text{P}^{\text{Cy}}_2\text{N}^{\text{Arg}}_2)_2]^{8+}$ can operate as a hydrogen oxidation catalyst over a wider pH range, from pH 0-14, with higher turnover frequencies (210 s^{-1} at 1 atm H_2 with an overpotential of 180 mV). It is so far the most active catalyst for hydrogen oxidation in this family of biomimics. It is however unlikely that the rate enhancement is exclusively due to the extra proton relay. Indeed intramolecular interactions between guanidinium groups (Arg-Arg pairing) operating at very low pH might also contribute to more subtle structural and electrostatic effects, favoring H_2 addition and heterolytic cleavage of the H-H bond.

Remarkably, at elevated temperatures (348 K), both glycine and arginine analogues have recently been demonstrated to perform as fully reversible H_2 production/oxidation catalysts in acidic aqueous solutions (pH = 0-6) with catalytic rates close to those of hydrogenase enzymes.³³ The fact that the arginine derivative, $[\text{Ni}(\text{P}^{\text{Cy}}_2\text{N}^{\text{Arg}}_2)_2]^{8+}$, displays higher catalytic rates in both directions (300 s^{-1} H_2 production and 20 s^{-1} H_2 oxidation; pH = 1, 348 K, 1 atm 25% H_2/Ar) compared to the glycine analogue, $[\text{Ni}(\text{P}^{\text{Cy}}_2\text{N}^{\text{Gly}}_2)_2]^{8+}$ (35 s^{-1} H_2 production and 7 s^{-1} H_2 oxidation), and an arginine-methyl ester analogue, $[\text{Ni}(\text{P}^{\text{Cy}}_2\text{N}^{\text{Arg-OMe}}_2)_2]^{8+}$ (3 s^{-1} H_2 production and 4 s^{-1} H_2 oxidation), suggests that both guanidinium and carboxyl groups play critical roles in the catalytic cycle by facilitating fast and reversible H_2 binding and release. Nevertheless this work provides a nice illustration of the positive impact of specific functional groups (here guanidiniums) in the second coordination sphere of the active metal ion on hydrogenase activity.

There is one example in which this class of complex has been incorporated into a protein scaffold and shown to display excellent activity. Because the objective was to develop a photocatalytic H_2 production system, using Photosystem I as the photosensitizing component, LM. Utschig, D. Tiede and collaborators chose apoflavodoxin as the host protein. This flavin mononucleotide (FMN) protein has indeed been shown to interact specifically with PSI and shuttle photogenerated reducing equivalents from PSI to several metabolic pathways.³⁴ While incompletely characterized, the hybrid system seems to use the FMN binding site for binding the $[\text{Ni}(\text{P}^{\text{Ph}}_2\text{N}^{\text{Ph}}_2)_2]^{2+}$ complex, which is furthermore partly exposed to the solvent. Assayed in association with PSI, an initial H_2 production rate of 1.25 mol H_2 /mol PSI/s (and a total turnover of 2825 mol H_2 /mol PSI in 4h) was obtained, using ascorbate as the reducing agent and cytochrome c6 as a redox mediator upon illumination with visible light

in aqueous solution. The PSI-flavodoxin $[\text{Ni}(\text{P}_2^{\text{Ph}}\text{N}_2^{\text{Ph}})_2]$ hybrid is the most active photocatalytic system reported so far using this class of Ni catalysts. However this system shows instability since it is functional only for 4 hours, for reasons that have not been clearly identified. In the same paper it is shown that the $[\text{Ni}(\text{P}^{\text{Ph}}_2\text{N}^{\text{Ph}}_2)_2]^{2+}$ complex can bind directly to PSI, in the absence of the flavodoxin scaffold. However, the system is less well defined, less active and less stable.

1.6.3 Artificial hydrogenases based on synthetic cobalt complexes

A great variety of cobalt complexes have been reported as promising electrocatalysts for hydrogen production.² Only few of them were also studied in the context of biohybrid systems.

Substituting Co for Fe in microperoxidase-11 (MP11), a water-soluble heme-undecapeptide (sequence: VQKCAQCHTVE) derived from horse cytochrome c, K. Bren and collaborators prepared a very stable biocatalyst for proton reduction in aqueous pH 7.0 solution (Figure 20A).³⁵

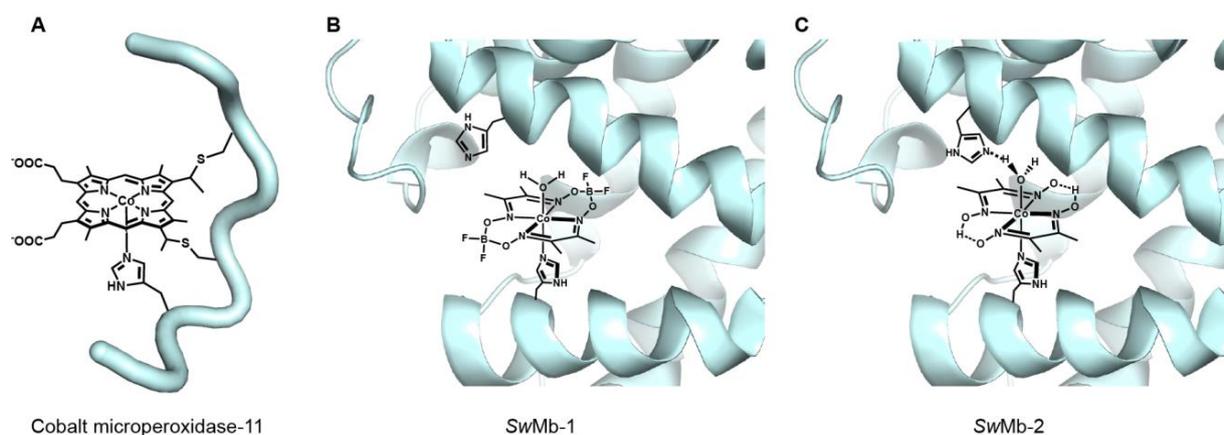


Figure 20: Model representation of cobalt biohybrids featuring (A) a cobalt-heme type complex attached to an undecapeptide and (B,C) cobaloxime type complexes (1: $[\text{Co}(\text{dmgBF}_2)_2(\text{H}_2\text{O})_2]$, and 2: $[\text{Co}(\text{dmgH})_2(\text{H}_2\text{O})_2]$) anchored inside the heme binding cavity of SwMb.

Hydrogen is evolved with nearly quantitative Faradaic efficiency and a turnover frequency of $\sim 6.7 \text{ s}^{-1}$ at 850 mV overpotential. With a total turnover number of 2.5×10^4 , this system takes benefit of the high stability of porphyrin ligands although the catalytic activity of the system clearly shows some decrease in activity after 15 min. Recently Ghirlanda and co-workers have demonstrated that cobalt-protoporphyrin IX incorporated into myoglobin

Section C

(CoMyo) can also efficiently catalyse electrochemical and photochemical proton reduction at neutral pH.³⁶ The activity of CoMyo is pH dependent with maximal photoinduced hydrogen production (518 turnovers in twelve hours) observed at pH 7, when irradiated in the presence of [Ru(bpy)₃]³⁺ and ascorbate. Relative to Co-MP11, in which the Co-porphyrin unit is completely exposed to the solvent, the protein environment of myoglobin offers a greater shielding of the Co-porphyrin cofactor resulting in an improved stability of the metalloprotein. Moreover, mutation of one or more histidine residue(s) in close proximity of the active site of CoMyo to alanine (H64A, H97A, and H64/97A) caused significant variation in the catalytic activity of the CoMyo mutants, demonstrating the importance of outer coordination sphere interactions in tuning the catalytic property of biohybrids. However, the low energy efficiency (overpotential above 800 mV) of these cobalt-porphyrin systems (CoMP11 and CoMyo) stands as a barrier against future technological developments. Indeed, cobalt porphyrins and phthalocyanine complexes have been extensively studied as catalysts for H₂ evolution^{37,38} in the past and attempts to significantly lower their overpotential requirement through remote structural modifications of the ligand have systemically failed, as also observed for other Co tetra-imine catalysts.^{2,39}

By contrast, cobaloximes and cobalt diimine-dioxime complexes are known to operate at relatively mild overpotentials and have been exploited in electrochemical and photochemical systems.^{2,3} As they display significant structural similarities with metalloporphyrins, it is not surprising that they have been combined with hemoproteins in the context of artificial hydrogenases. *Sperm-whale* myoglobin (SwMb) is a host protein of choice since it is known to readily bind Schiff-base complexes of chromium(III), manganese(III), iron(III) and copper(II), through axial coordination to a histidine residue, mostly, His93, the natural ligand of hemin in holo-SwMb.⁴⁰ Furthermore cobaloximes possess square-planar geometry similar to Schiff-base complexes as well as good affinity in their axial position for N-heterocyclic bases. It has been recently reported that cobaloxime complexes, such as {Co(dmgH)₂} or {Co(dmgBF₂)₂} (dmgH₂=dimethylglyoxime), can bind to the heme pocket of apoSwMb through coordination to a single histidine giving a mononuclear Co protein (Figure 20B and 20C).⁴¹ QC/MM docking calculations and EXAFS measurements establish that the cobalt ion is hexacoordinated with a water molecule as the sixth ligand. The hydrophobic environment provided by the heme binding site is reflected in the g-values extracted from EPR spectra of the biohybrids and explains the 100 mV negative shift observed for the redox potential of the cobaloxime Co(II)/Co(I) couple upon incorporation into the heme pocket of the protein. Using assays based on either thermal proton reduction by

Eu(II) complexes or light-driven proton photoreduction in the presence of a photosensitizer, catalytic H₂ evolution activity was observed in aqueous neutral solutions.

Furthermore low overpotential requirement (~200 mV) was established. However the stability of these cobaloxime-based artificial hydrogenases is limited (up to 5 turnovers). This might be improved in the future through active site engineering or through enhanced electron communication with the electron donor. Probably, the low stability of the cobaloxime-myoglobin biohybrids observed upon turnover conditions originates from hydrogenation of the glyoximato ligand by the catalytic Co(III)-H hydride intermediate. Faster reductive conversion of the latter to the H₂-evolving Co(II)-H intermediate might favor H₂ production over ligand reduction.⁴² Such a mechanism might be at work in the Photosystem I /cobaloxime assemblies reported by Tiede, Utschig and coworkers.⁴³ These self-assembled systems, even if they are still poorly characterized, produce H₂ upon illumination with visible light in the presence of ascorbate and cytochrome c6.⁴³ Samples containing 2-4 Co atoms per PSI display impressive activities (TOF = 170 mol H₂/mol PSI/min during the first 10 min; TON= 5200 after 1.5 h).

In conclusion, despite the efforts the field of artificial hydrogenases is still in its infancy.

This is not surprising considering the complexity of the problem, as the reactions at work imply electron transfers, proton transfers and gas transport which need to be highly coordinated.

However, the possibility of converting an inefficient catalyst into a highly active artificial hydrogenase by embedding it into a protein environment has been clearly demonstrated. In the absence of a clear rationale for the choice of the host macromolecular (proteic or non-proteic) system so far, several exciting biohybrid or supramolecular systems, namely those based on biomimetic diiron complexes, nickel diphosphines and cobaloxime derivatives, described here, constitute powerful platforms to progress towards efficient artificial enzymes, with high catalytic rates at low overpotentials. By well-designed modifications of the second and further coordination spheres, through site-directed mutagenesis of the host protein or synthetic modification of peptides and non-biological polymers, not only this goal may be achieved but also a wealth of basic knowledge, at a molecular level, will be provided regarding how a metal catalyst can be affected by the electrostatic, hydrophobic, and steric properties of the scaffold.

References:

Section C

- 1 T. R. Simmons, G. Berggren, M. Bacchi, M. Fontecave and V. Artero, *Coord. Chem. Rev.*, 2014, **270**, 127–150.
- 2 V. Artero, M. Chavarot-Kerlidou and M. Fontecave, *Angew. Chem.-Int. Ed.*, 2011, **50**, 7238–7266.
- 3 V. Artero and J.-M. Saveant, *Energy Environ. Sci.*, 2014, **7**, 3808–3814.
- 4 M. E. Wilson and G. M. Whitesides, *J. Am. Chem. Soc.*, 1978, **100**, 306–307.
- 5 M. Duerrenberger and T. R. Ward, *Curr. Opin. Chem. Biol.*, 2014, **19**, 99–106.
- 6 A. Roy, C. Madden and G. Ghirlanda, *Chem. Commun.*, 2012, **48**, 9816–9818.
- 7 Y. Sano, A. Onoda and T. Hayashi, *Chem. Commun.*, 2011, **47**, 8229–8231.
- 8 Y. Sano, A. Onoda and T. Hayashi, *J. Inorg. Biochem.*, 2012, **108**, 159–162.
- 9 H.-Y. Wang, W.-G. Wang, G. Si, F. Wang, C.-H. Tung and L.-Z. Wu, *Langmuir ACS J. Surf. Colloids*, 2010, **26**, 9766–9771.
- 10 F. Quentel, G. Passard and F. Gloaguen, *Chem.-Eur. J.*, 2012, **18**, 13473–13479.
- 11 F. Quentel, G. Passard and F. Gloaguen, *Energy Environ. Sci.*, 2012, **5**, 7757–7761.
- 12 F. Wang, W.-G. Wang, H.-Y. Wang, G. Si, C.-H. Tung and L.-Z. Wu, *Acs Catal.*, 2012, **2**, 407–416.
- 13 C. Orain, F. Quentel and F. Gloaguen, *Chemsuschem*, 2014, **7**, 638–643.
- 14 T. Yu, Y. Zeng, J. Chen, Y.-Y. Li, G. Yang and Y. Li, *Angew. Chem.-Int. Ed.*, 2013, **52**, 5631–5635.
- 15 F. Wang, W.-J. Liang, J.-X. Jian, C.-B. Li, B. Chen, C.-H. Tung and L.-Z. Wu, *Angew. Chem.-Int. Ed.*, 2013, **52**, 8134–8138.
- 16 M. L. Singleton, J. H. Reibenspies and M. Y. Darensbourg, *J. Am. Chem. Soc.*, 2010, **132**, 8870–8871.
- 17 M. L. Singleton, D. J. Crouthers, R. P. Duttweiler, J. H. Reibenspies and M. Y. Darensbourg, *Inorg. Chem.*, 2011, **50**, 5015–5026.
- 18 X. Li, M. Wang, D. Zheng, K. Han, J. Dong and L. Sun, *Energy Environ. Sci.*, 2012, **5**, 8220–8224.
- 19 J.-X. Jian, Q. Liu, Z.-J. Li, F. Wang, X.-B. Li, C.-B. Li, B. Liu, Q.-Y. Meng, B. Chen, K. Feng, C.-H. Tung and L.-Z. Wu, *Nat. Commun.*, 2013, **4**, 2695.
- 20 S. Pullen, H. Fei, A. Orthaber, S. M. Cohen and S. Ott, *J. Am. Chem. Soc.*, 2013, **135**, 16997–17003.
- 21 K. Sasan, Q. Lin, C. Mao and P. Feng, *Chem. Commun.*, 2014, **50**, 10390–10393.
- 22 T. R. Simmons and V. Artero, *Angew. Chem.-Int. Ed.*, 2013, **52**, 6143–6145.

Chapter I

- 23 D. L. DuBois, *Inorg. Chem.*, 2014, **53**, 3935–3960.
- 24 B. Ginovska-Pangovska, A. Dutta, M. L. Reback, J. C. Linehan and W. J. Shaw, *Acc. Chem. Res.*, 2014, **47**, 2621–2630.
- 25 A. Jain, S. Lense, J. C. Linehan, S. Raugei, H. Cho, D. L. DuBois and W. J. Shaw, *Inorg. Chem.*, 2011, **50**, 4073–4085.
- 26 A. Jain, M. L. Reback, M. L. Lindstrom, C. E. Thogerson, M. L. Helm, A. M. Appel and W. J. Shaw, *Inorg. Chem.*, 2012, **51**, 6592–6602.
- 27 M. L. Reback, B. Ginovska-Pangovska, M.-H. Ho, A. Jain, T. C. Squier, S. Raugei, J. A. S. Roberts and W. J. Shaw, *Chem.-Eur. J.*, 2013, **19**, 1928–1941.
- 28 M. L. Reback, G. W. Buchko, B. L. Kier, B. Ginovska-Pangovska, Y. Xiong, S. Lense, J. Hou, J. A. S. Roberts, C. M. Sorensen, S. Raugei, T. C. Squier and W. J. Shaw, *Chem.-Eur. J.*, 2014, **20**, 1510–1514.
- 29 U. J. Kilgore, J. A. S. Roberts, D. H. Pool, A. M. Appel, M. P. Stewart, M. R. DuBois, W. G. Dougherty, W. S. Kassel, R. M. Bullock and D. L. DuBois, *J. Am. Chem. Soc.*, 2011, **133**, 5861–5872.
- 30 W. J. Shaw, M. L. Helm and D. L. DuBois, *Biochim. Biophys. Acta-Bioenerg.*, 2013, **1827**, 1123–1139.
- 31 A. Dutta, S. Lense, J. Hou, M. H. Engelhard, J. A. S. Roberts and W. J. Shaw, *J. Am. Chem. Soc.*, 2013, **135**, 18490–18496.
- 32 A. Dutta, J. A. S. Roberts and W. J. Shaw, *Angew. Chem.-Int. Ed.*, 2014, **53**, 6487–6491.
- 33 A. Dutta, D. L. DuBois, J. A. S. Roberts and W. J. Shaw, *Proc. Natl. Acad. Sci.*, 2014, **111**, 16286–16291.
- 34 S. C. Silver, J. Niklas, P. Du, O. G. Poluektov, D. M. Tiede and L. M. Utschig, *J. Am. Chem. Soc.*, 2013, **135**, 13246–13249.
- 35 J. G. Kleingardner, B. Kandemir and K. L. Bren, *J. Am. Chem. Soc.*, 2014, **136**, 4–7.
- 36 D. J. Sommer, M. D. Vaughn and G. Ghirlanda, *Chem. Commun.*, 2014, **50**, 15852–15855.
- 37 R. Kellett and T. Spiro, *Inorg. Chem.*, 1985, **24**, 2378–2382.
- 38 R. Kellett and T. Spiro, *Inorg. Chem.*, 1985, **24**, 2373–2377.
- 39 V. Artero and M. Fontecave, *Coord. Chem. Rev.*, 2005, **249**, 1518–1535.
- 40 F. Yu, V. M. Cangelosi, M. L. Zastrow, M. Tegoni, J. S. Plegaria, A. G. Tebo, C. S. Mocny, L. Ruckthong, H. Qayyum and V. L. Pecoraro, *Chem. Rev.*, 2014, **114**, 3495–3578.
- 41 M. Bacchi, G. Berggren, J. Niklas, E. Veinberg, M. W. Mara, M. L. Shelby, O. G. Poluektov, L. X. Chen, D. M. Tiede, C. Cavazza, M. J. Field, M. Fontecave and V. Artero, *Inorg. Chem.*, 2014, **53**, 8071–8082.

Section C

42 E. S. Andreiadis, P.-A. Jacques, P. D. Tran, A. Leyris, M. Chavarot-Kerlidou, B. Jusselme, M. Matheron, J. Pécaut, S. Palacin, M. Fontecave and V. Artero, *Nat. Chem.*, 2013, **5**, 48–53.

43 L. M. Utschig, S. C. Silver, K. L. Mulfort and D. M. Tiede, *J. Am. Chem. Soc.*, 2011, **133**, 16334–16337.

Section D

Aim of the project

The field of hydrogenases and artificial systems is very active and fascinating. It has provided new tools, inorganic complexes and biohybrids for hydrogen evolution and uptake. However these systems are still quite inefficient and so far platinum is still the best catalyst for H₂ production. More work is needed on hydrogenases in order to optimize potential artificial hydrogenases. Indeed, still a number of questions regarding hydrogenases remain to be addressed such as the diversity of redox states observed in HydA enzymes, the function of accessory F clusters, the molecular basis of oxygen resistance/sensitivity, the biological mechanism of active site assembly (maturation), etc... Furthermore, more X-ray structures from new hydrogenases are needed in order to identify key structural features that should be implemented into an artificial hydrogenase.

Here we explore: (i) a new [FeFe]-hydrogenase, from *Megasphaera elsdenii*; (ii) a fully chemical methodology for maturing/activating such enzymes; (iii) the protein HydF, as an artificial enzyme, because of its potential to bind a 6Fe cluster

The points addressed in this thesis are the following:

First, only few [FeFe]-hydrogenases have been completely characterized showing that, despite the same catalytic core (H-cluster), the activities, the number of accessory clusters and redox behaviour differ from each other substantially. There is thus a need to get easier access to more members of that class of enzymes and the tool developed by Berggren G. and coworkers (*i.e* chemical activation of [FeFe]-hydrogenase by treating the apoprotein with the synthetic complex **1**, [Fe₂(adt)(CO)₄(CN)₂]²⁻) could allow progressing in this direction.

Thus, part of my PhD work has been focused on the use of the “chemical maturation” tool to activate new apo-HydA enzymes and to study their spectroscopic and catalytic properties. The enzyme that has been studied is the HydA from *Megasphaera elsdenii*, partially characterized in the past. Chapter II deals with *Megasphaera elsdenii* HydA enzyme. This protein harbours two [4Fe–4S] F centers and the catalytic H-cluster. The implemented strategy was the heterologous protein overexpression in *E. coli*, aerobic purification of the enzyme (purified without any cluster, since all the metallo-cofactors are destroyed in the

Section D

presence of oxygen) and subsequently the reconstitution of F- and H-clusters via the insertion of iron, sulfide and complex **1**.

The enzyme has been spectroscopically characterized, using EPR and FTIR, and the catalytic properties have been studied.

Second, the simplest form of [FeFe]-hydrogenases is found in algae, which contains only the H-cluster and no F-clusters.

One of these proteins, HydA from *Chlamydomonas reinhardtii*, is the most studied enzyme and has been recently exploited for biochemical and spectroscopic characterization. Unfortunately there is no X-ray structure available of an active enzyme belonging to this class. This protein lacks the additional [Fe–S] clusters observed in most bacterial [FeFe]-hydrogenases, that complicate the direct examination of the catalytic core.

Thus, in order to study directly the H-cluster in bacterial [FeFe]-hydrogenases without the influence of the F-clusters, a new construct of *Megasphaera elsdenii* HydA has been realized, where the domain harbouring the accessory clusters has been removed.

Chapter III presents the results obtained with the truncated *Megasphaera* enzyme. Similarly to *Megasphaera* HydA full length protein, aerobic purification followed by the anaerobic reconstitution of the H-cluster ([4Fe–4S]_H and [2Fe]-subcluster) has been carried out. Finally, the spectroscopic properties of the H-cluster have been studied by multiple techniques (FTIR, EPR and Mössbauer spectroscopy) and the catalytic properties have been evaluated.

Third, the interesting work of G. Berggren and coworkers showed that it is possible to prepare HydF hybrids having a 6Fe metallo-cofactor relevant to the catalytically competent H-cluster. Such a capacity of HydF to bind a di-iron organometallic complex is obviously related to its biological role as scaffold protein during maturation of hydrogenases. A [4Fe–4S] cluster binds a diiron complex via a cyanide bridge (HydF), generating a 6Fe cluster with a structure displaying remarkable analogies with that of the H cluster, while obviously not similar. It was thus tempting to evaluate the catalytic properties of such a hybrid metalloprotein, get access to structural information on this system not available yet, thus opening a new direction towards an original artificial hydrogenase, unexplored before.

HydF proteins from different organisms were purified aerobically and subsequently with similar protocols used for HydA enzymes, the iron-sulfur cluster has been reconstituted. For the first time a three-dimensional structure of HydF with its [4Fe–4S] cluster, has been obtained. The 6 Fe cluster of HydF has been also prepared chemically with di-iron complexes

Chapter I

mimicking the active site of HydA. The metallo-cofactors have been spectroscopically characterized (EPR, FTIR, HYSCORE) and hydrogenase activities evaluated.

These results are described in chapter IV.

Section D

Chapter II

Results and discussion

The [FeFe]-hydrogenase
from *Megasphaera elsdenii*

Chapter II: The [FeFe]-hydrogenase from *Megasphaera elsdenii*

As described in chapter I, [FeFe]-hydrogenases are unique biocatalysts for the interconversion between protons and dihydrogen.¹ These metalloenzymes display extremely large catalytic reaction rates at very low overpotential and they have been well studied from structural and mechanistic perspectives.²⁻⁴ Furthermore they are considered as potential alternative catalysts to noble metals for the development of novel bio-fuel cells and bio-electrolyzers as well as bio-photoelectrochemical cells.⁵⁻⁸ Studies on [FeFe]-hydrogenases are unfortunately limited by the fact that these enzymes are difficult to prepare. First, they have to be purified and manipulated only under strict anaerobic conditions. Second, their maturation, the process by which the [Fe-S] clusters and the [2Fe]-subcluster are synthesized and assembled, depends on complex and specific protein machineries (chapter I, section **B**). In particular, the system involved in the biosynthesis of the [2Fe]-subcluster is still incompletely characterized.⁹ Recently, it has been shown that recombinant [FeFe]-hydrogenases can be prepared anaerobically in an inactive form containing only the [Fe-S] clusters and subsequently fully matured by reaction with the synthetic $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex, called **1** in this thesis.^{10,11} The resulting active site has been shown to display EPR and FTIR characteristics identical to those of the naturally matured enzyme, as illustrated in the case of HydA1 from *Chlamydomonas reinhardtii* (CrHydA1). In case of HydA from *Clostridium pasteurianum* (CpHydA) the definitive confirmation that the chemically synthesized active site was identical to the biosynthesized one came from the observation of a remarkable identity of the two three-dimensional structures.¹²

This synthetic maturation methodology has already led to a number of very interesting applications which would have been otherwise unreachable, such as specific labeling of the active site or preparation of artificial hydrogenases.^{4,13-17} In this thesis, we go a step further towards the facile preparation of active hydrogenases in large amounts from standard *E. coli* strains lacking the HydEFG machinery. It is sufficient to express the corresponding gene in *E. coli*, purify the apoenzyme aerobically and activate it chemically, simply via, first, incorporation of the [Fe-S] clusters by treatment with iron and sulfide and, second, incorporation of the [2Fe]-subunit by reaction with the complex **1**, anaerobically (Figure **21**). This methodology is illustrated here with the preparation of HydA from the fermentative anaerobic rumen bacterium *Megasphaera elsdenii*. This enzyme, named MeHydA, has been previously isolated from a *M. elsdenii* strain but only partially characterized in the 1980's,

The [FeFe]-hydrogenase from *Megasphaera elsdenii*

when the structure of these enzymes was not yet known.¹⁸⁻²¹ Furthermore, we characterized completely MeHydA using EPR and FTIR spectroscopy, isolating H_{ox} , H_{ox} -CO and H_{red} redox states and compared it with the most studied enzymes.²²⁻²⁴ This part of the project has been done in collaboration with the group of Prof. W. Lubitz at the Max Plank Institute in Mülheim (Germany) where I spent a couple of weeks studying MeHydA.

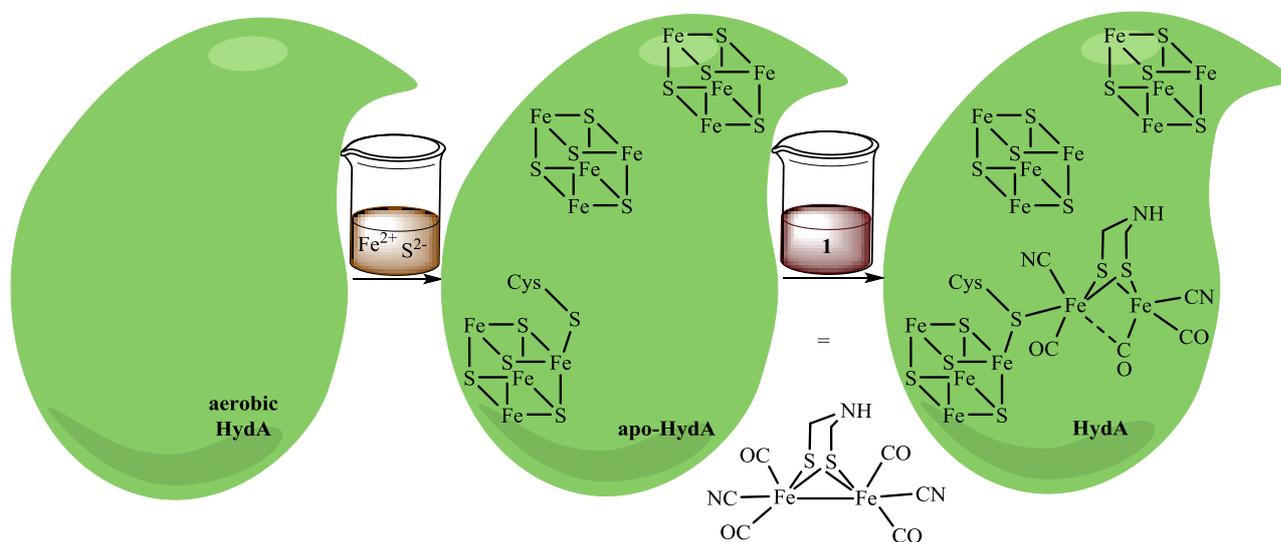


Figure 21: Chemical maturation of F- and H-clusters of apo-MeHydA

2.1 Expression and aerobic purification of MeHydA

Expression of MeHydA in *E. coli* resulted in a large production of soluble recombinant protein. As the protein is His-tagged, it could be purified to homogeneity with only two chromatographic steps (Figure 22, chapter V for detailed purification protocols). After a His tag affinity column and a gel filtration step using a Superdex S200 chromatographic column the protein was purified to homogeneity and was finally concentrated to 5-7 mg/ml. Overall 10-15 mg of MeHydA could be obtained from 1 L of culture.

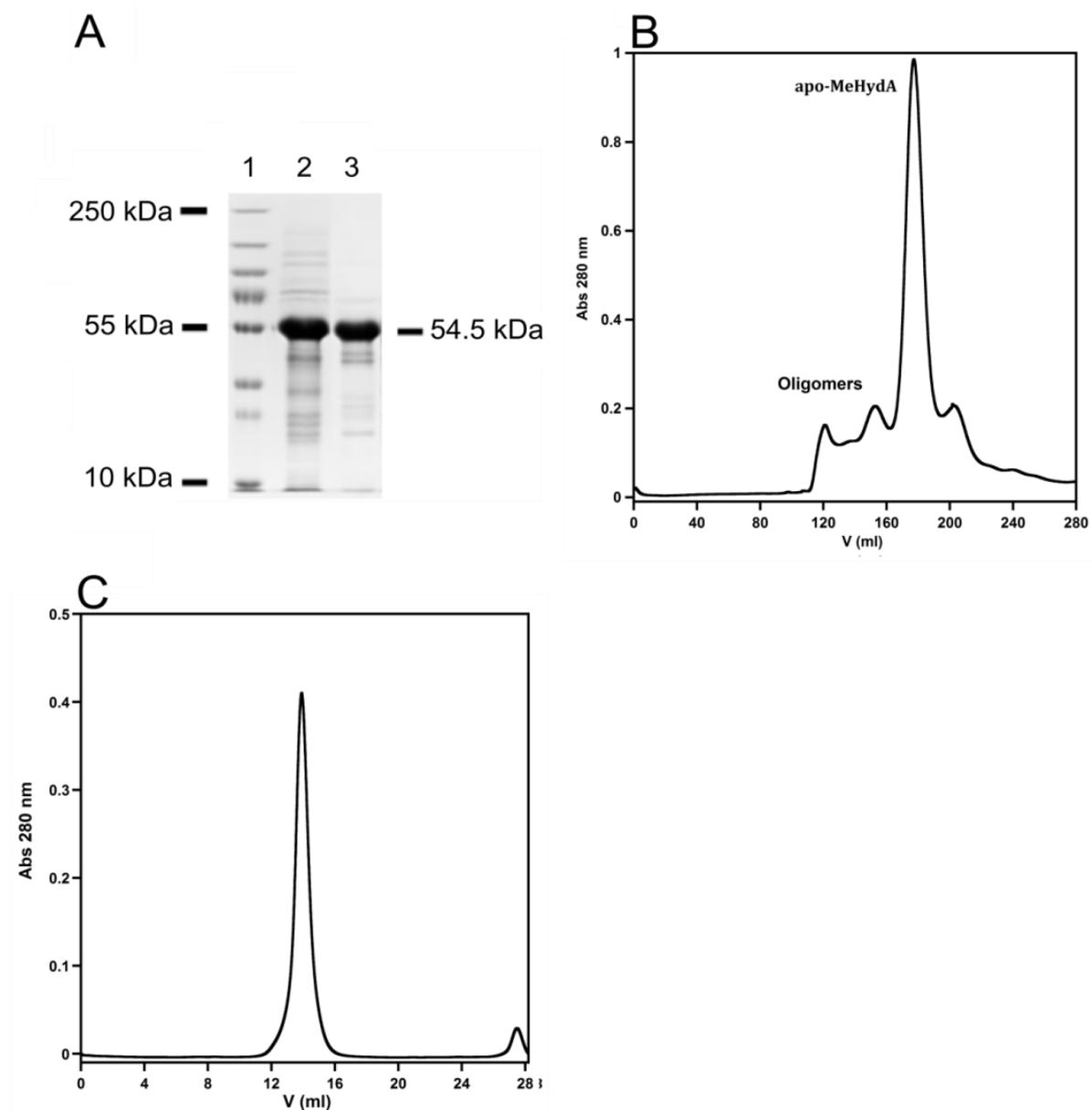


Figure 22: Aerobic purification of the hydrogenase from *Megasphaera elsdenii*: A) SDS-PAGE of the protein eluted with an imidazole gradient using a His trap column (lane 2). Protein obtained after the size exclusion chromatography step (lane 3). MW Protein Standard (lane 1). B) Size exclusion chromatography elution profile of apo-MeHydA obtained after the His trap purification step. C) Size exclusion chromatography elution profile of apo-MeHydA used for chemical reconstitution of the [4Fe-4S] clusters. The elution volume of the apo-MeHydA is 13.7 mL.

All the purification procedures were carried out aerobically. β -mercaptoethanol and dithiothreitol (DTT) were needed in the buffers in order to limit the formation of oligomers, certainly via disulfide bridge reduction. While a very small amount of iron was detected in this preparation the protein was mainly in the apo form and was named apo-MeHydA.

2.2 [Fe-S] cluster reconstitution of apo-MeHydA

The [FeFe]-hydrogenase from *Megasphaera elsdenii*

To incorporate the iron-sulfur clusters in apo-MeHydA, a standard protocol was used, consisting of the anaerobic incubation of the protein with an excess of iron ammonium sulfate and cysteine in the presence of DTT, as reducing agent. The reaction was initiated by addition of catalytic amounts of CsdA, a cysteine desulfurase, and monitored by light absorption spectroscopy since [Fe-S] clusters display a characteristic charge-transfer absorption band at 410 nm (Figure 23 and 26A). Since apo-MeHydA should contain three [4Fe-4S] clusters, according to previous studies¹⁸⁻²¹, several reconstitution experiments were carried out varying the excess of the reactants and monitoring the amount of iron and sulfur incorporated inside the protein.

Finally, the protein proved to be fully reconstituted using a 17-18 molar excess of Fe and L-cys compared to the apoprotein concentration.

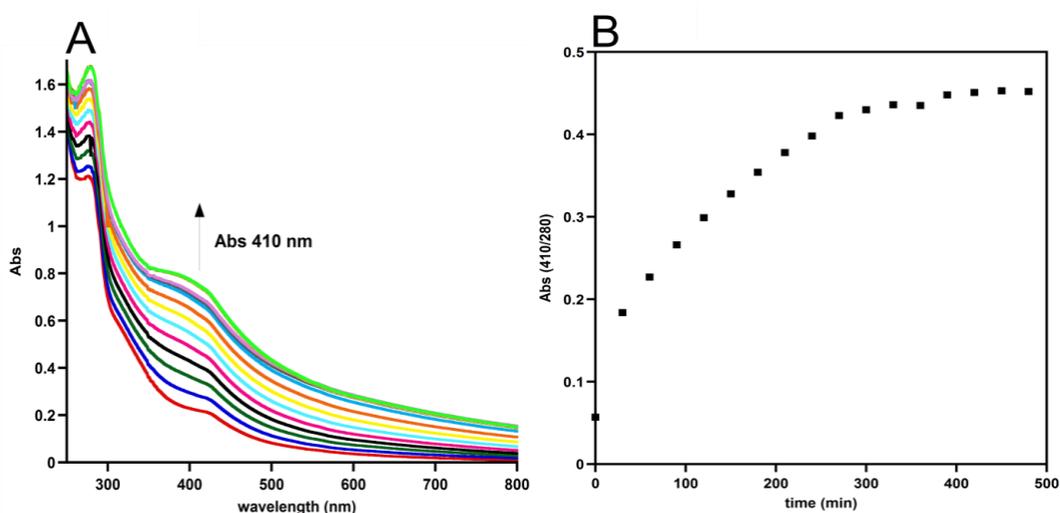


Figure 23: In vitro reconstitution of the [Fe-S] clusters in MeHydA. A) UV-visible absorption spectra of 100 μ M MeHydA in 50 mM Tris pH 8, 300 mM NaCl, 10% v/v glycerol, 5 mM DTT using a 1 mm pathlength cuvette recorded every 30 minutes. B) Plot of $A_{410/280}$ absorption ratio as a function of time. The [4Fe-4S] cluster reconstitution is completed after 300 minutes.

Visible absorption spectra were recorded on a Cary 100 spectrophotometer (Agilent) connected to the cell holder located in a glovebox with optical fibers. A huge band in the 400-410 nm region appeared due to the *in vitro* reconstitution of the three [4Fe-4S] clusters. Plotting the absorption ratio $A_{400/280}$ as a function of the time, it was possible to monitor the kinetic of the reconstitution process (Figure 23B). Gel filtration showed the formation of large oligomers (Figure 24A), which were enzymatically inactive. Only the last protein fraction was collected and further characterized. The elution volume of apo-MeHydA (13.7 mL) and reconstituted MeHydA (named FeS-MeHydA, 14.7 mL) were different and led us to

hypothesize a dimer/monomer transition upon reconstitution of the [4Fe–4S] clusters (Figure 24B).

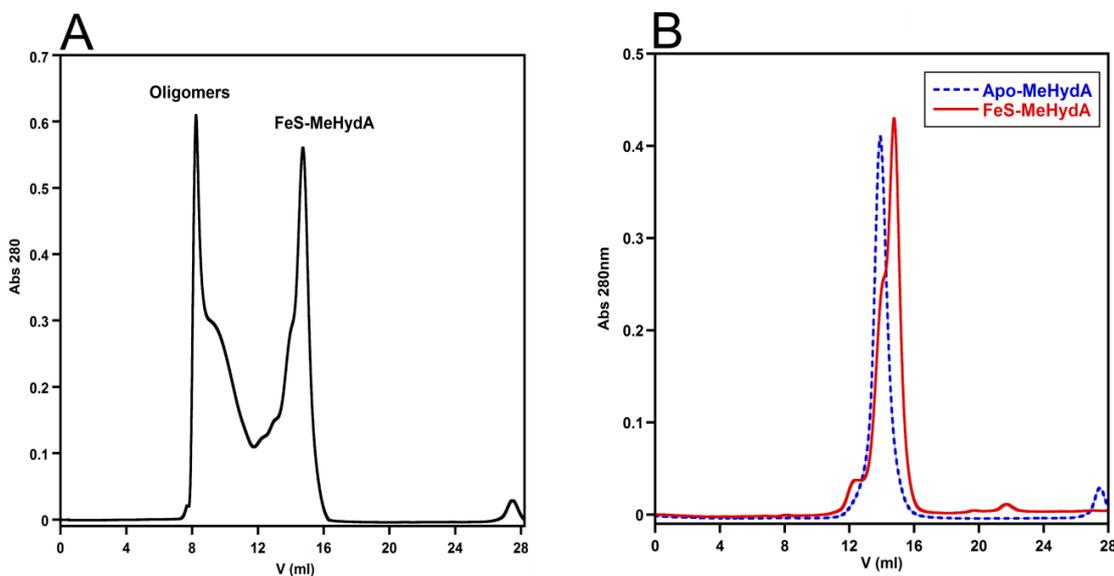


Figure 24: Purification of reconstituted FeS-MeHydA. **A)** SEC elution profile on a S200 10/300 GL of MeHydA after overnight [4Fe–4S] clusters reconstitution. The peaks at low elution volumes (8-10 mL) correspond to oligomers. **B)** S200 elution profile of apo-MeHydA (dashed blue) and FeS-MeHydA (red).

2.3 Static light scattering of apo-MeHydA

To verify if the [Fe–S] cluster reconstitution was coupled to a dimer/monomer transition (Figure 25C), we used static light scattering in batch mode to assess the molecular weight of apo-MeHydA using the Zimm equation and the Debye plot (Figure 25D). This experiment showed that apo-MeHydA had an apparent molecular weight of 62.7 ± 0.8 kDa excluding it being a dimer (MW_{th} : 54.5 kDa). As a consequence, after [Fe–S] cluster reconstitution of MeHydA, the elution volume increased due to a change of the protein conformation, likely becoming more compact.

Static light scattering measurements were performed in batch mode in a quartz cuvette on a Zetasizer ZSP (Malvern) using toluene as a standard. The measurements were done on apo-MeHydA freshly purified. A dn/dc of 0.187 at 633 nm, calculated from the protein sequence using the program SEDFIT was used.²⁵ The molecular weight was extracted from the intercept of the linear fit with the Zimm equation for an isotropic scatterer, $Kc/R_{\theta} = 1/M + 2A_2c$, of the Debye Plot: where K is an optical constant, c is the concentration of the sample, R_{θ} is the Rayleigh ratio, M is the molecular weight and A_2 the second virial coefficient.

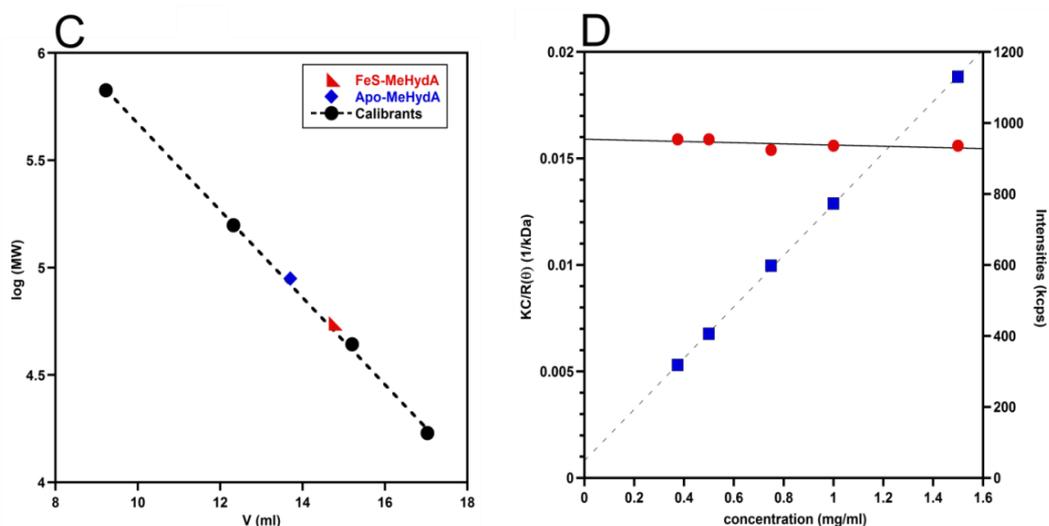


Figure 25: Apo-MeHydA behaviour. **C)** Superdex S200 10/300 GL calibration curve using Thyroglobulin (670 kDa, 9.22 mL), γ -globulin (158 kDa, 12.32 mL), Ovalbumin (44 kDa, 15.2 mL) and Myoglobin (17 kDa, 17.03 mL). The molecular weights derived for apo-MeHydA and FeS-MeHydA are 90 kDa and 54.5 kDa respectively, suggesting apo-MeHydA might be a dimer. **D)** Debye plot of apo-MeHydA at different concentrations in 50 mM Tris pH 8, 300 mM NaCl (red filled circles). The intensity scattered for each concentration of apo-MeHydA (blue squares) is shown as a control for sample quality. The estimated molecular weight extracted from the linear fit with the Zimm equation of the Debye plot is 62.7 ± 0.8 kDa.

2.4 EPR spectroscopic characterization of the iron-sulfur MeHydA

The quantitation of Fe (10.5 ± 0.2 per protein) and S (11.0 ± 0.2 per protein) as well as the UV-Visible spectrum (Figure 26A) of the protein were in good agreement with the expected presence of three [4Fe-4S] clusters within a single polypeptide chain. To our knowledge it is the first time that this procedure proves effective for full reconstitution of three [4Fe-4S] clusters within a single polypeptide chain. The type of clusters present in the protein was further confirmed by Q-band EPR spectroscopy (Figure 26B). The reconstituted oxidized protein was EPR silent. However, upon anaerobic reduction using sodium dithionite, it became EPR active with overlapping signals characteristic of reduced $S=1/2$ [4Fe-4S]⁺ clusters (Figure 26B). The g values obtained from a simulation (2.055, 1.930, 1.885) are typical for [4Fe-4S]⁺ clusters. The sharp signal at $g \approx 2.0$ marked with an asterisk was attributed to a radical originating from the aerobic purification of apo-MeHydA in the presence of DTT.²⁶

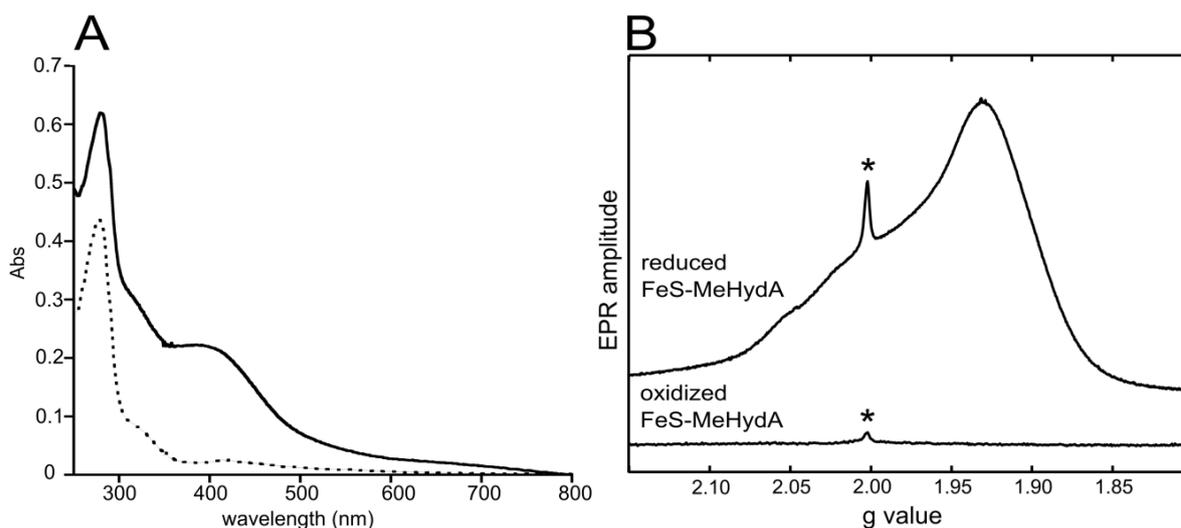


Figure 26: Spectroscopic characterization of reconstituted FeS-MeHydA. A) UV-visible spectra of 68 μM apo-MeHydA (dashed line) and FeS-MeHydA (solid line) in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% v/v glycerol, 5 mM DTT. B) Q-band FID-detected EPR spectra (10K) of 170 μM FeS-MeHydA, before (20K) and after (10K) reduction with sodium dithionite, in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% v/v glycerol, 5 mM DTT.

These data show that the reconstitution process leads exclusively to $S=0$ $[4\text{Fe}-4\text{S}]^{2+}$ oxidized clusters and exclude the presence of $[3\text{Fe}-4\text{S}]^+$ and $[2\text{Fe}-2\text{S}]^{2+}$ clusters. FeS-MeHydA is an interesting novel form of *M. elsdenii* HydA as it may allow studying the electron-transfer chain specifically. MeHydA indeed contains, in addition to the H-cluster, two additional $[4\text{Fe}-4\text{S}]$ clusters bound to the same polypeptide chain.

2.5 Chemical maturation

2.5.1 Incorporation of synthetic $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex

Applying the methodology of the chemical maturation¹⁰, FeS-MeHydA was treated with an excess of the synthetically prepared diiron complex **1** for one hour anaerobically and then purified by desalting (NAP 10) in order to remove the excess of the chemical. Evidence for protein maturation, thus named holo-MeHydA, was obtained by enzymatic assay. Using a standard assay based on the reduction of protons by dithionite/methyl viologen and quantification of dihydrogen by gas chromatography, we showed that holo-MeHydA was highly active. A specific hydrogenase activity of $600 \pm 50 \mu\text{mol H}_2 \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$, that corresponds to a TOF of 550 turnovers per second (MW 54.5 kDa), was obtained for holo-

The [FeFe]-hydrogenase from *Megasphaera elsdenii*

MeHydA (Figure 27). In the presence of an excess of the mimic, the measured hydrogenase activity was lower by ~33% ($400 \pm 50 \mu\text{mol H}_2 \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$), possibly because of release of CO, an enzyme inhibitor, from the complex **1** in excess.

An activity of $400 \mu\text{mol H}_2 \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$, for native MeHydA has been reported in literature, proving the efficiency of the chemical maturation.¹⁸

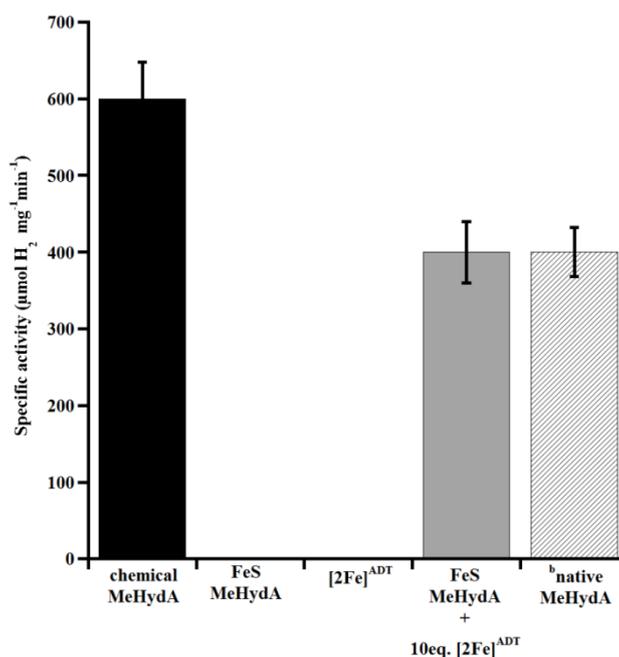


Figure 27: Specific hydrogenase activity. H_2 production was measured in the presence of methyl viologen (10 mM) and sodium dithionite (100 mM) in 100 mM potassium phosphate pH 6.8. Chemical-MeHydA, FeS-MeHydA, complex **1**, FeS-MeHydA + 10 eq. complex **1**, native MeHydA. The data shown are mean values \pm s.d. Complex **1** is indicated as $[\text{2Fe}]^{\text{ADT}}$.

2.5.2 Incorporation of synthetic $[\text{Fe}_2(\text{pdt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex

Applying the same methodology, MeHydA has been matured with $[\text{Fe}_2(\text{pdt})(\text{CO})_4(\text{CN})_2]^{2-}$ synthetic complex, called **2**, where the nitrogen atom of the adt^{2-} bridge has been replaced by a carbon atom. The new hybrid system formally represents an artificial hydrogenase. Evidence for maturation of the protein, named **2**-MeHydA, was obtained by enzymatic assay. The H_2 evolution quantified by GC was sensibly lower (close to the detection limit) compared to the maturation with synthetic complex **1** (which recreates the H-cluster). Thus, an electrochemical H_2 microsensor (Unisense H_2 microsensor) was used, registering a current that is function of the H_2 produced in solution.

A diluted solution of **2**-MeHydA (0.2 nmol) was transferred into a small chamber (1.15 mL) containing methyl viologen, as sacrificial electron donor, in 100 mM potassium phosphate buffer pH 6.8. The H₂ evolution started adding a reducing agent, sodium dithionite. The activities of **2** and FeS-MeHydA were also evaluated and compared to the **2**-MeHydA and holo-MeHydA (matured with the complex **1**).

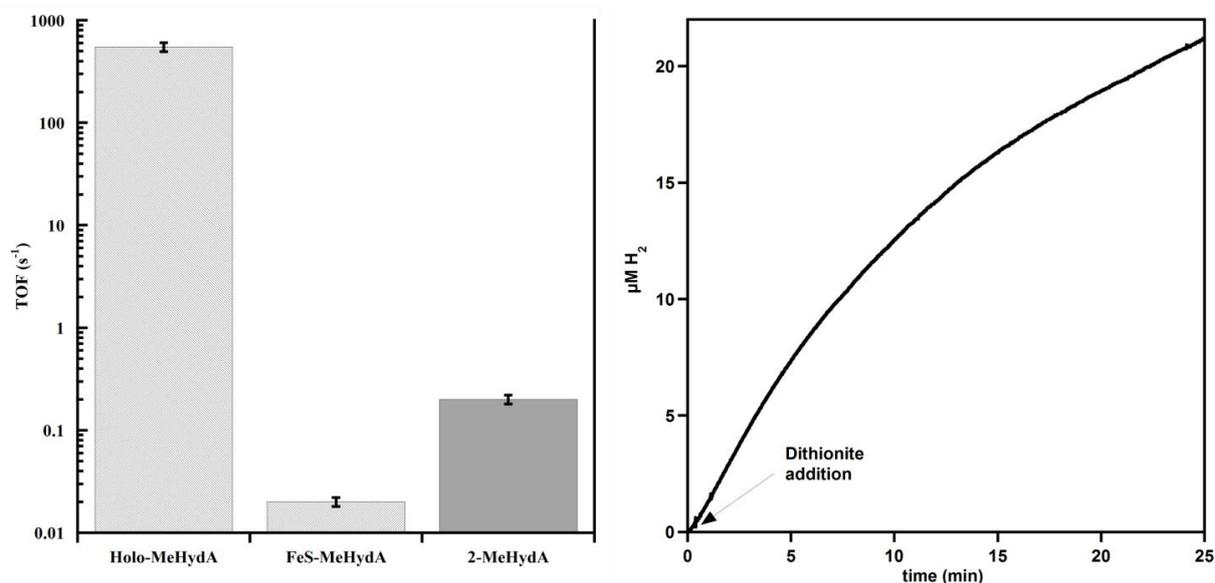


Figure 28: Left: TOF (s⁻¹) of FeS-MeHydA, **2**-MeHydA and **1**-MeHydA (holo-MeHydA) chemically matured enzymes. The y axis is logarithmic. Right: H₂ production (μM) of **2**-MeHydA after the addition of sodium dithionite (Initial TOF ≈ 10-12 min⁻¹). The data shown are mean values ± s.d.

As shown in Figure **28**, **2**-MeHydA produces H₂ with a TOF of 10-12 min⁻¹, much lower compared to the 550 s⁻¹ of holo-MeHydA. The lower H₂ production of **2**-MeHydA could be explained by the absence of the nitrogen atom in the bridgehead position. The N-bridgehead atom is not only fundamental as Brønsted base but, its flexibility is essential for the function; indeed, in a report of Lubitz and coworkers, a simple methylation of N-bridgehead atom decreased of 100-fold the hydrogen evolution.¹⁷

The FeS-MeHydA, containing only the three [4Fe-4S] clusters, showed some activity with a TOF of ≈ 1-1.5 min⁻¹, much lower compared to the matured protein with complexes **1** and **2**.

2.6 FTIR and EPR spectroscopic characterization of holo-MeHydA

In order to confirm that the H-cluster has been assembled correctly but also to provide a full characterization of active *M. elsdenii* HydA, we used Fourier Transform Infrared (FTIR) spectroscopy for characterization (Figure **29A**, Table **1**). FTIR is appropriate for identifying

The [FeFe]-hydrogenase from *Megasphaera elsdenii*

the CO (1800-2020 cm^{-1}) and CN (2040-2100 cm^{-1}) vibrations associated with the CO and CN^- ligands present in the [2Fe]-subcluster.²⁷ These signals can be clearly distinguished from those of complex **1** in solution which are much broader (Figure 29A, top).¹¹ The narrow CO and CN FTIR bands of the H-cluster allow differentiating between the various states usually coexisting in hydrogenase preparations.

The paramagnetic states of the H-cluster (as well as the $[\text{4Fe-4S}]^+$ clusters in the unmaturation enzyme) were also studied using EPR.

In the as-isolated sample (after chemical maturation), a mixture of mainly H_{ox} and $\text{H}_{\text{ox-CO}}$ states, all expectedly containing a CO bridge revealed by the 1803/1804 cm^{-1} feature, is observed.

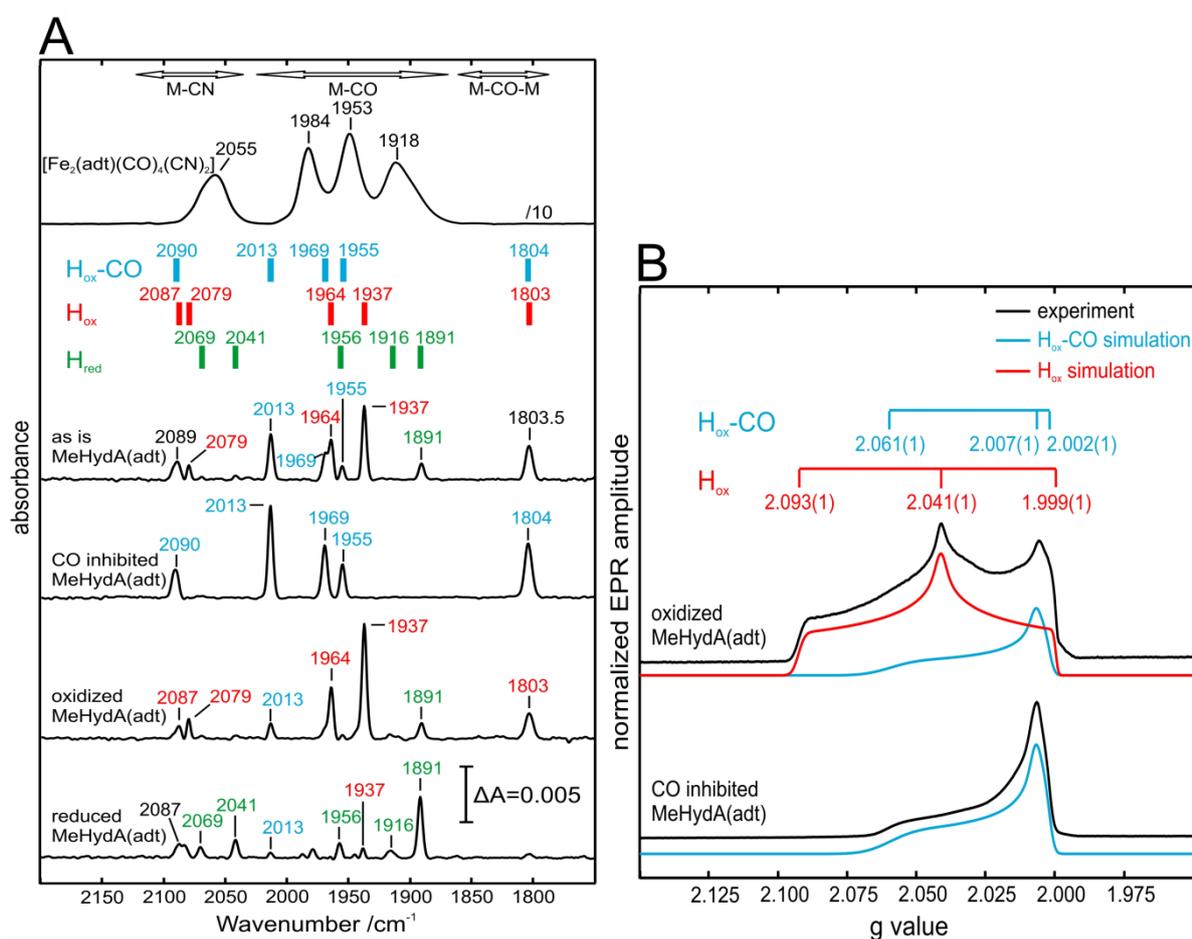


Figure 29: EPR and FTIR spectra of chemically matured MeHydA. **A)** FTIR spectra of $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$, holo-MeHydA as isolated, H_{ox} , $\text{H}_{\text{ox-CO}}$ and H_{red} states in 100 mM potassium phosphate pH 6.8. **B)** Q-band FID-detected EPR-spectra ($T = 10\text{K}$) of 0.5 mM holo-MeHydA in the $\text{H}_{\text{ox-CO}}$ and H_{ox} states in 100 mM potassium phosphate pH 6.8 and their simulations. Information about FTIR and EPR signal positions for each redox state is presented at the top of the panels (**A**, **B**) and in Table 1 and Table 2, respectively.

Under synthetic maturation, the supernumerary CO is released from the complex **1** ($[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$) resulting in a substantial amount of the $\text{H}_{\text{ox}}\text{-CO}$ state. This signal is also often present in $[\text{FeFe}]$ -hydrogenase preparations due to the so-called ‘‘cannibalization process’’ in which the CO ligands, from light- or oxygen-damaged H-clusters, are released and captured by H-clusters that are still intact.^{14,22,28} Additional CO is ‘‘leaking’’ from the excess binuclear complex present in the maturation buffer further increasing the $\text{H}_{\text{ox}}\text{-CO}$ population (Figure **29A**).

To obtain a pure CO inhibited state, the sample was flushed for 20 minutes with CO gas. The FTIR spectrum is shown in Figure **29A** and is in good agreement with previously obtained data from different $[\text{FeFe}]$ -hydrogenases (Table **1**). The $\text{H}_{\text{ox}}\text{-CO}$ state can be converted to the H_{ox} state by flushing the sample with argon gas for 1 hour. The FTIR spectrum presented in Figure **29A** again shows the typical signals of the H_{ox} state (Table **1**) with a small contamination of $\text{H}_{\text{ox}}\text{-CO}$ and H_{red} .

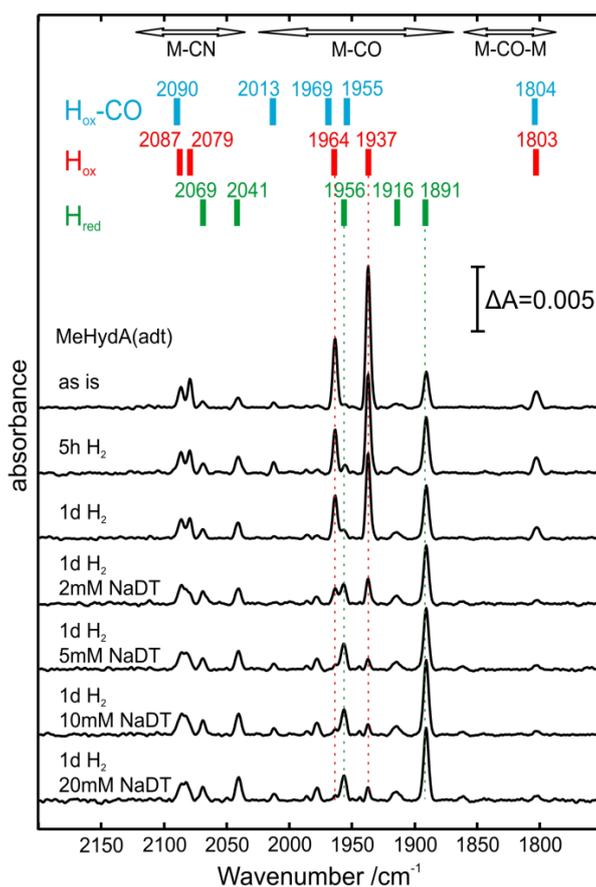


Figure 30: Progressive reduction of matured MeHydA(adt) observed by FTIR. 0.5 mM holo-MeHydA in 50 mM Tris-HCl pH 8.0, 300 mM NaCl were flushed with H₂ for 24 h, followed by adding different amounts of sodium dithionite to obtain the H_{red} state.

The [FeFe]-hydrogenase from *Megasphaera elsdenii*

Finally, upon reduction, H_{ox} is converted into the H_{red} state. It turns out that MeHydA cannot be fully reduced by H₂ flushing, thus additional sodium dithionite treatment (5-10 mM) was needed (Figure 30). The FTIR signature of MeHydA-H_{red} and DdH-H_{red} are very similar but different from CrHydA1-H_{red} which still contains a bridging CO ligand. Instead the H_{red} state in bacterial [FeFe] hydrogenases has an FTIR spectrum resembling that of CrHydA1-H_{sred} (Table 1). It should be noted that before the present work, an FTIR spectrum of HydA isolated from *M. elsdenii* was reported.²⁹ However, while some (not all) CO vibrational features characteristic of H_{ox} and H_{red} were present, no CN vibrations could be observed. This raises uncertainties regarding the relevance of these data.

Figure 29B shows the EPR spectra of the oxidized forms described above. The EPR spectrum of the pure H_{ox}-CO state is characterized by a nearly axial signal originating from the low spin mixed-valence configuration of the 2Fe-subcluster Fe(I)Fe(II) while the [4Fe-4S] cluster is in the oxidized 2+ state.

Table 1. Collection of FTIR vibrations associated with the CO and CN ligands present in the [2Fe] subcluster for different HydAs.

Organism	Species	Fe-CN	Fe-CO	Fe-CO-Fe	Ref
<i>M. elsdenii</i>	H _{ox}	2087, 2079	1964, 1937	1803	This work
<i>D. desulfuricans</i>	H _{ox}	2093, 2079	1965, 1940	1802	Ref 22
<i>C. pasteurianum</i>	H _{ox}	2086, 2072	1971, 1948	1802	Ref 30
<i>C. reinhardtii</i>	H _{ox}	2088, 2072	1964, 1940	1800	Ref 24
<i>C. acetobutylicum</i>	H _{ox}	2082, 2070	1969, 1946	1801	Ref 31
<i>M. elsdenii</i>	H _{ox} -CO	2090	2013, 1969, 1955	1804	This work
<i>D. desulfuricans</i>	H _{ox} -CO	2096, 2088	2016, 1971, 1963	1810	Ref 22
<i>C. pasteurianum</i>	H _{ox} -CO	2095, 2077	2017, 1974, 1971	1810	Ref 30
<i>C. reinhardtii</i>	H _{ox} -CO	2092, 2084	2013, 1970, 1964	1810	Ref 24
<i>C. acetobutylicum</i>	H _{ox} -CO	2090, 2075	2015, 1973, 1967	1806	Ref 31
<i>M. elsdenii</i>	H _{red}	2069, 2041	1956, 1916, 1891		This work
<i>D. desulfuricans</i>	H _{red}	2079, 2041	1965, 1916, 1894		Ref 22
<i>C. acetobutylicum</i>	H _{red}	2053, 2040	1899		Ref 31
<i>C. reinhardtii</i>	H _{red}	2083, 2070	1935, 1891	1793	Ref 24
<i>C. reinhardtii</i>	H _{sred}	2070, 2026	1954, 1919, 1882		Ref 23,24

Although the redox states of H_{ox} and H_{ox}-CO are the same, the EPR signal originating from the 2Fe-subcluster in H_{ox} is rhombic. Only the rhombic signal has been reported before in the literature.³² Finally, the H_{red} state is EPR silent in agreement with the [2Fe]-subcluster being in the S=0 Fe(I)Fe(I) state.

The different states of the H-cluster H_{ox} -CO, H_{ox} , and H_{red} in MeHydA have thus been prepared following procedures similar to those reported for other bacterial [FeFe]-hydrogenases.^{22,33} The spectral features of the H-cluster in these three states can be compared with those found for HydA from the organisms *D. desulfuricans*, *C. pasteurianum*, and *C. reinhardtii* (Table 1 and Table 2). It turns out that the FTIR bands and EPR g-values of H_{ox} and H_{ox} -CO are very similar in all species. Only the g_1 -value of H_{ox} -CO shows some variability where *M. elsdenii* is closest to *D. desulfuricans*. Also the FTIR signature of the EPR silent H_{red} state of MeHydA is very similar to that of the other bacterial [FeFe]-hydrogenases. While for *C. reinhardtii* (lacking the extra F clusters) the (semi) bridging CO ligand is still present, it is a terminal CO in hydrogenases from *M. elsdenii* and the other bacterial species. On the other hand, the doubly reduced state H_{sred} observed in algae (CrHydA1) is not stable and undetectable in bacterial [FeFe]-hydrogenases.

Table 2. Collection of EPR g values of the H cluster for HydA in different species.

Organism	Species	g1	g2	g3	Ref
<i>M. elsdenii</i>	H_{ox}	2.093	2.041	1.999	This work
<i>D. desulfuricans</i>	H_{ox}	2.100	2.040	1.998	Ref 28
<i>C. pasteurianum</i>	H_{ox}	2.097	2.039	1.999	Ref 34
<i>C. reinhardtii</i>	H_{ox}	2.104	2.042	1.998	Ref 10
<i>C. acetobutylicum</i>	H_{ox}	2.089	2.036	1.995	Ref 31
<i>M. elsdenii</i>	H_{ox} -CO	2.061	2.007	2.002	This work
<i>D. desulfuricans</i>	H_{ox} -CO	2.065	2.007	2.001	Ref 28
<i>C. pasteurianum</i>	H_{ox} -CO	2.072	2.006	2.006	Ref 34
<i>C. reinhardtii</i>	H_{ox} -CO	2.054	2.009	2.009	Ref 10
<i>C. acetobutylicum</i>	H_{ox} -CO	2.075	2.009	2.009	Ref 31
<i>C. reinhardtii</i>	H_{sred}	2.073	1.935	1.880	Ref 10
<i>D. desulfuricans</i>	H_{trans}	2.060	1.960	1.890	Ref 28

2.7 Conclusion and future perspectives: biotechnological application?

In this thesis has been developed an efficient method for producing [FeFe]-hydrogenases by heterologous expression in *E. coli* of the structural gene, aerobic purification in two steps, i) reconstitution of [Fe-S] clusters and subsequent ii) chemical maturation with a synthetic organometallic complex (**1**).³⁵

Straightforward high-yield aerobic purification afforded the apoenzyme without any metallic cofactors. Subsequent reconstitution of the (oxygen sensitive) [4Fe-4S] clusters is carried out chemically from iron salt and sulfide under anaerobic conditions. We showed in the case of MeHydA that this is efficient even for the case of a protein containing three [4Fe-

The [FeFe]-hydrogenase from *Megasphaera elsdenii*

4S] clusters. Treatment with the $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex (**1**) completes the activation process, as it generates a fully assembled catalytically competent H-cluster – including the [2Fe]-subcluster.

In the case of MeHydA this procedure is highly efficient since we obtained for this enzyme a high specific activity ($600 \mu\text{mol H}_2.\text{mg}^{-1}.\text{min}^{-1}$). FTIR and EPR spectroscopic characterization confirmed that the redox states of the MeHydA enzyme are very similar to the bacterial [FeFe]-hydrogenases from *D. desulfuricans* and *C. pasteurianum*. With this proof of concept we hope that this methodology can be applied to other hydrogenases from other living organisms for large scale production of an active enzyme. It is one of the many possible applications of the discovery that simple chemical reactants, iron and sulfide as well as $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$, can fully replace the complex metallofactor biosynthetic machineries within the same hydrogenase, the [Fe–S] cluster assembly machinery and the HydEFG machinery, respectively.

Moreover, this methodology enables the easy insertion of different diiron catalysts inside the hydrogenase scaffold, including those with multiple metal cofactors. For example, MeHydA has been matured with another biomimetic complex, $[\text{Fe}_2(\text{pdt})(\text{CO})_4(\text{CN})_2]^{2-}$, where the bridgehead atom of the dithiolate bridge is a carbon atom. The artificial enzyme, **2**-MeHydA, proved active towards protons reduction with a TOF of *c.a* $10\text{-}12 \text{ min}^{-1}$. In another report of Lubitz and coworkers showed similar results for CrHydA1 matured with the same synthetic complex.¹⁴

Interestingly, **2**-CrHydA1 can access only two redox states (and not three as the protein matured with the complex containing the biologically relevant adt^{2-} ligand), is more stable to higher potentials and is not inhibited by carbon monoxide. A similar behavior for **2**-MeHydA can be envisaged and further experiments can be hypothesized such as O_2 and CO inhibition of holo-MeHydA and **2**-MeHydA. The O_2 irreversible inactivation of these enzymes remains the first drawback for [FeFe]-hydrogenases application in biotechnological devices. Indeed, they need to be manipulated under strict anaerobic conditions due to H-cluster degradation.

Future studies on holo-MeHydA will be directed towards the study of the oxygen sensitivity of this enzyme, since only few proteins have been studied. An electrochemical study is fundamental to investigate and compare the properties of this enzyme with other hydrogenases, since in a recent study of Léger C. and coworkers high difference in terms of O_2 resistance is observed between the HydA from *Clostridium acetobutylicum* (CaHydA), containing additional F clusters, and CrHydA1 containing just the $[\text{4Fe-4S}]_{\text{H}}$.³⁶ CaHydA is

more stable to short exposure of O₂ compared to CrHydA1 and was not excluded the participation of the F clusters in the mechanism of protection against oxygen. Since MeHydA has a different number of accessory F clusters compared to CaHydA, a direct comparison between the two enzymes is needed.

In conclusion, it has been shown how synthetic chemistry and biochemistry can work together for the facile preparation of active [FeFe]-hydrogenases. The powerful tool of the “chemical maturation” has gained consensus in the scientific community and has been proved to be one of the first examples where a chemical manipulation of a protein leads to an *in vitro* enzyme as efficient as the *in vivo* one.

References:

- 1 W. Lubitz, H. Ogata, O. Rüdiger and E. Reijerse, *Chem. Rev.*, 2014, **114**, 4081–4148.
- 2 T. Lautier, P. Ezanno, C. Baffert, V. Fourmond, L. Cournac, J. C. Fontecilla-Camps, P. Soucaille, P. Bertrand, I. Meynial-Salles and C. Léger, *Faraday Discuss*, 2011, **148**, 385–407.
- 3 D. W. Mulder, E. M. Shepard, J. E. Meuser, N. Joshi, P. W. King, M. C. Posewitz, J. B. Broderick and J. W. Peters, *Structure*, 2011, **19**, 1038–1052.
- 4 V. Artero, G. Berggren, M. Atta, G. Caserta, S. Roy, L. Pecqueur and M. Fontecave, *Acc. Chem. Res.*, 2015, **48**, 2380–2387.
- 5 T. W. Woolerton, S. Sheard, Y. S. Chaudhary and F. A. Armstrong, *Energy Environ. Sci.*, 2012, **5**, 7470–7490.
- 6 J. A. Cracknell, K. A. Vincent and F. A. Armstrong, *Chem. Rev.*, 2008, **108**, 2439–2461.
- 7 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.
- 8 G. Caserta, S. Roy, M. Atta, V. Artero and M. Fontecave, *Curr. Opin. Chem. Biol.*, 2015, **25**, 36–47.
- 9 E. M. Shepard, F. Mus, J. N. Betz, A. S. Byer, B. R. Duffus, J. W. Peters and J. B. Broderick, *Biochemistry*, 2014, **53**, 4090–4104.
- 10 J. Esselborn, C. Lambertz, A. Adamska-Venkatesh, T. Simmons, G. Berggren, J. Nothl, J. Siebel, A. Hemschemeier, V. Artero, E. Reijerse, M. Fontecave, W. Lubitz and T. Happe, *Nat. Chem. Biol.*, 2013, **9**, 607–609.
- 11 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.

The [FeFe]-hydrogenase from *Megasphaera elsdenii*

- 12 J. Esselborn, N. Muraki, K. Klein, V. Engelbrecht, N. Metzler-Nolte, U.-P. Apfel, E. Hofmann, G. Kurisu and T. Happe, *Chem Sci*, 2016, **7**, 959–968.
- 13 A. Adamska-Venkatesh, T. R. Simmons, J. F. Siebel, V. Artero, M. Fontecave, E. Reijerse and W. Lubitz, *Phys. Chem. Chem. Phys.*, 2015, **17**, 5421–5430.
- 14 A. Adamska-Venkatesh, D. Krawietz, J. Siebel, K. Weber, T. Happe, E. Reijerse and W. Lubitz, *J. Am. Chem. Soc.*, 2014, **136**, 11339–11346.
- 15 A. Adamska-Venkatesh, S. Roy, J. F. Siebel, T. R. Simmons, M. Fontecave, V. Artero, E. Reijerse and W. Lubitz, *J. Am. Chem. Soc.*, 2015, **137**, 12744–12747.
- 16 R. Gilbert-Wilson, J. F. Siebel, A. Adamska-Venkatesh, C. C. Pham, E. Reijerse, H. Wang, S. P. Cramer, W. Lubitz and T. B. Rauchfuss, *J. Am. Chem. Soc.*, 2015, **137**, 8998–9005.
- 17 J. F. Siebel, A. Adamska-Venkatesh, K. Weber, S. Rumpel, E. Reijerse and W. Lubitz, *Biochemistry*, 2015, **54**, 1474–1483.
- 18 M. Filipiak, W. R. Hagen and C. Veeger, *Eur. J. Biochem. FEBS*, 1989, **185**, 547–553.
- 19 C. Van Dijk, S. G. Mayhew, H. J. Grande and C. Veeger, *Eur. J. Biochem. FEBS*, 1979, **102**, 317–330.
- 20 C. van Dijk, H. J. Grande, S. G. Mayhew and C. Veeger, *Eur. J. Biochem. FEBS*, 1980, **107**, 251–261.
- 21 C. DIJK and C. VEEGER, *Eur. J. Biochem.*, 1981, **114**, 209–219.
- 22 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.
- 23 A. Adamska, A. Silakov, C. Lambertz, O. Ruediger, T. Happe, E. Reijerse and W. Lubitz, *Angew. Chem.-Int. Ed.*, 2012, **51**, 11458–11462.
- 24 A. Silakov, C. Kamp, E. Reijerse, T. Happe and W. Lubitz, *Biochemistry*, 2009, **48**, 7780–7786.
- 25 H. Zhao, P. H. Brown and P. Schuck, *Biophys. J.*, 2011, **100**, 2309–2317.
- 26 L. E. Netto and E. R. Stadtman, *Arch. Biochem. Biophys.*, 1996, **333**, 233–242.
- 27 Y. Nicolet, A. L. de Lacey, X. Vernède, V. M. Fernandez, E. C. Hatchikian and J. C. Fontecilla-Camps, *J. Am. Chem. Soc.*, 2001, **123**, 1596–1601.
- 28 S. P. J. Albracht, W. Roseboom and E. C. Hatchikian, *J. Biol. Inorg. Chem.*, 2006, **11**, 88–101.
- 29 T. M. van der Spek, A. F. Arendsen, R. P. Happe, S. Yun, K. A. Bagley, D. J. Stufkens, W. R. Hagen and S. P. Albracht, *Eur. J. Biochem. FEBS*, 1996, **237**, 629–634.

- 30 Z. Chen, B. J. Lemon, S. Huang, D. J. Swartz, J. W. Peters and K. A. Bagley, *Biochemistry*, 2002, **41**, 2036–2043.
- 31 S. Morra, S. Maurelli, M. Chiesa, D. W. Mulder, M. W. Ratzloff, E. Giamello, P. W. King, G. Gilardi and F. Valetti, *Biochim. Biophys. Acta BBA - Bioenerg.*, 2016, **1857**, 98–106.
- 32 P. J. van Dam, E. J. Reijerse and W. R. Hagen, *Eur. J. Biochem. FEBS*, 1997, **248**, 355–361.
- 33 D. W. Mulder, M. W. Ratzloff, E. M. Shepard, A. S. Byer, S. M. Noone, J. W. Peters, J. B. Broderick and P. W. King, *J. Am. Chem. Soc.*, 2013, **135**, 6921–6929.
- 34 B. Bennett, B. J. Lemon and J. W. Peters, *Biochemistry*, 2000, **39**, 7455–7460.
- 35 G. Caserta, A. Adamska-Venkatesh, L. Pecqueur, M. Atta, V. Artero, R. Souvik, E. Reijerse, W. Lubitz and M. Fontecave, *Biochim. Biophys. Acta BBA - Bioenerg.*, 2016.
- 36 C. Orain, L. Saujet, C. Gauquelin, P. Soucaille, I. Meynial-Salles, C. Baffert, V. Fourmond, H. Bottin and C. Léger, *J. Am. Chem. Soc.*, 2015, **137**, 12580–12587.

Chapter III

Results and discussion

A truncated form of
M. elsdenii
[FeFe]-hydrogenase

Chapter III: A truncated form of *M. elsdenii* [FeFe]-hydrogenase

The development of industrially relevant energy conversion with (photo)electrochemical systems, such as electrolyzers, photoelectrochemical cells and fuel cells, will depend on the identification and optimization of cheap, stable and efficient catalysts for the redox reactions at the electrodes.^{1,2} Among them, the [FeFe]-hydrogenases are extensively studied (as for potential bioelectrodes) as they display remarkable performances in particular with respect to turnover frequencies (TOFs), largely exceeding those displayed by synthetic catalysts.^{3,4} The simplest [FeFe]-hydrogenases are observed in algae, including *Chlamydomonas reinhardtii* (CrHydA1), *Chlorella fusca*, *Scenedesmus obliquus* etc... which express enzymes consisting of only the H-cluster without F-cluster domains.^{5,6}

One of these proteins, CrHydA1, has been exploited more recently for biochemical and spectroscopic characterization because it lacks the additional iron-sulfur clusters observed in most bacterial [FeFe]-hydrogenases, like the hydrogenases from *Megasphaera elsdenii* and *Clostridium pasteurinum*, which can make hard the direct examination of the catalytic core.^{7,8} These extra clusters are fundamental for electron shuttling between the catalytic core and the protein surface and in the previous chapter we showed that they were also important for the correct protein folding of *M. elsdenii* HydA protein. CrHydA1 has been widely characterized with different techniques, X-ray crystallography included. Indeed, an X-ray structure of the inactive apo-form of CrHydA1 presenting only the [4Fe–4S]_H cluster has been solved. The study of this protein provided lots of information regarding the mechanism of H-cluster assembly, the active site destruction and poisoning in the presence of some inhibitors like oxygen, carbon monoxide and formaldehyde.^{9–11} Indeed, in 2010 Mulder and co-workers reported the three dimensional structure of CrHydA1 protein heterologously expressed in *E. coli*., showing that the H-cluster assembly proceeds in a stepwise manner. The first step is the assembly of the [4Fe–4S] cluster, by ISC and SUF machineries, and subsequently the [2Fe]-subcluster thanks to the cooperation of HydE, HydF and HydG proteins.^{10,11}

CrHydA1, due to its simplicity, has been deeply studied presenting unusual redox properties. Indeed, A. Adamska et al. reported that for this enzyme it was possible to isolate the so-called *superreduced* state (H_{sred}), in which both the [4Fe–4S]_H and the [2Fe]-subcluster of the H-cluster are in the reduced configuration.¹² A combination of FTIR, EPR and Spectroelectrochemistry characterized this redox intermediate which is assumed to take part to the catalytic mechanism as described in chapter I (Paragraph 1.5.7). This allowed

A truncated form of *M. elsdenii* [FeFe]-hydrogenase

postulating that for bacterial HydA containing additional iron-sulfur centers, the H_{red} state is not stabilized due to the electron redistribution towards F clusters which is thermodynamically favoured.

In the previous chapter has been developed an efficient method for producing the [FeFe]-hydrogenase from *Megasphaera elsdenii* by first, the heterologous expression in *E. coli* of the structural gene, followed by the reconstitution of [Fe-S] clusters and subsequent chemical maturation with a synthetic organometallic complex (**1**), [Fe₂(adt)(CO)₄(CN)₂]²⁻. UV-Vis, Fe/S content, EPR, FTIR and hydrogenase activity clearly proved the efficient assembly of all metal cofactors onto the protein scaffold, which was converted in a fully active enzyme.⁸

On the other hand, [FeFe]-hydrogenases are fraught with a number of drawbacks which still need to be addressed before they can be implemented into electrodes. One is their limited tolerance to molecular oxygen and CrHydA1, in particular, is one of the most unstable.¹³⁻¹⁵ Another drawback is their size which limits enzyme density at the surface of the electrodes. For example, [NiFe]-hydrogenases contain two subunits, with a total mass of almost 100 kDa, while [FeFe]-hydrogenases (HydA) contain a conserved ~ 350 residue H-cluster domain that harbors a [4Fe-4S] cluster bridged to a catalytic [2Fe]-subcluster via a cysteine thiol and several terminal domains that bind ferredoxin-type iron-sulfur clusters for electron transfer.³

Inspired by the simplicity of CrHydA1, we studied a truncated form of HydA from *Megasphaera elsdenii*, lacking the ferredoxin domains indicated as F clusters, as an example to illustrate the concept that it is possible to engineer hydrogenases to generate smaller and still active enzyme versions. The characterization of this new “*chlamydomonas-like*” hydrogenase is here presented. This new construct could provide information about the mechanism of HydAs from all different organisms, and could help to understand why CrHydA1, harbouring only the H-cluster, is so unstable in the presence of oxygen while *Clostridium acetobutylicum* and *Clostridium pasteurianum* HydAs, with multiple [Fe-S] cores, showed a higher resistance.¹⁶

The aim of the project was to remove the two iron-sulfur F clusters obtaining a form of *Megasphaera elsdenii* holding only the H-cluster; thus two strategies were ideated.

First, it has been shown that some site-directed mutations of key residues can allow a selective [4Fe-4S] assembly, isolating a single iron-sulfur core. For example in previous studies, Fontecilla J. and coworkers were able to purify and crystallize many different forms of HydE maturase containing either two [4Fe-4S] clusters, one [4Fe-4S] and a [2Fe-2S] core

or a single [4Fe–4S] cluster.^{17,18} A similar approach for the HydG maturase enzyme allowed the comprehension of the role of both [4Fe–4S] clusters in the protein scaffold. Indeed, two constructs harboring either the [4Fe–4S] Radical Sam or the [4Fe–4S] auxiliary cluster were realized showing different spectroscopic properties for the clusters.¹⁹

These studies, clearly, provided evidence that is possible to assemble selectively a [4Fe–4S] cluster without repercussions on the protein folding. The hydrogenase from *Megasphaera elsdenii*, called in the last chapter MeHydA, possesses two CxxCxxC motifs both harboring an F cluster.^{8,20} The first strategy was to mutate 2 Cys residues of each binding region to prevent the [4Fe–4S] cluster assembly. This new construct should have the same size (≈ 55.4 kDa) of the wild type protein but with a single [4Fe–4S] center, that of the H-cluster.

Second, a different strategy could be the deletion of the entire domain holding the two F extra clusters. Most [FeFe]-hydrogenases consist of a single polypeptide chain where two domains can be distinguished. The C-terminal domain, called “*Hdomain*” harbors the H-cluster while the accessory F-domains hold a set of Ferredoxin-type [4Fe–4S] clusters (in CpHydA structure there is also a [2Fe–2S] cluster at the N-terminal region).⁷ These additional [Fe–S] centers form an electric wire, shuttling electrons from the protein surface to the active site. This part of the project was conducted in Grenoble under the supervision of Dr. M. Atta (CEA researcher, LCBM). I spent 4 months in Grenoble where I have been trained in biochemistry, protein chemistry and iron-sulfur cluster reconstitution assays.

3.1 Strategy 1: Cys to Ser mutations

Aligning MeHydA with *Clostridium pasteurianum* HydA, I identified the Cys residues of both CxxCxxC motifs, harboring the two F clusters. A new construct with the **Cys31-34-62-65Ser** punctual modifications, called *CystoSerMeHydA*, was ordered to Genscript and cloned into the vector pet22b (Figure **31**). We replaced the cysteine residues with the isosteric serines (-SH replaced by -OH) and to facilitate the purification procedure, a hexahistidine tag has been fused at the N-terminus.

A truncated form of *M. elsdenii* [FeFe]-hydrogenase

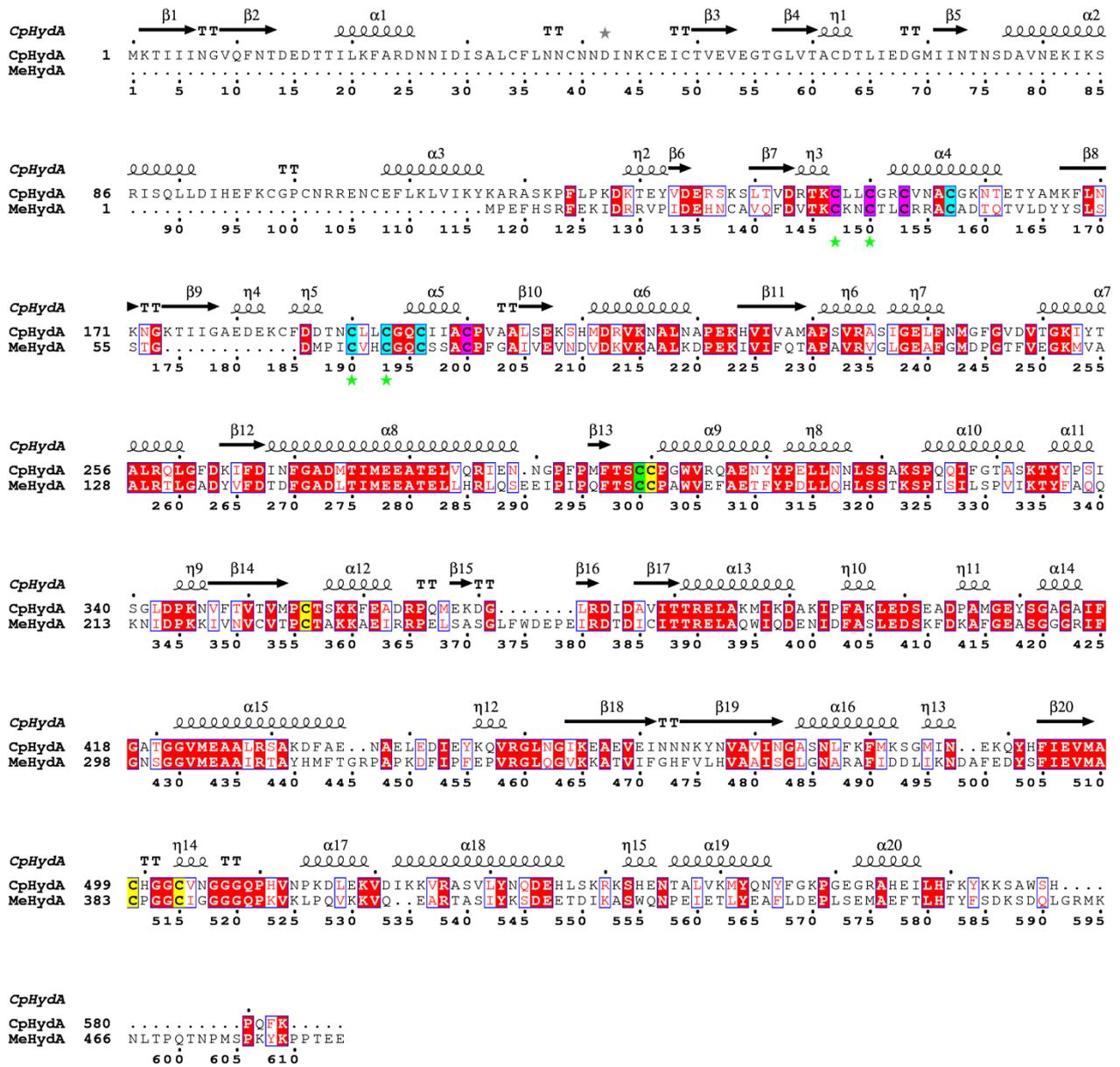


Figure 31: Sequence alignment of the [FeFe]-hydrogenase from *Clostridium pasteurianum* with the [FeFe]-HydA from *Megasphaera elsdenii*. Conserved secondary structures are calculated using the available X-ray structure of CpHydA (pdb 3c8y). The cysteine residues harboring the [4Fe-4S]_H cluster are highlighted in yellow, the conserved cys residues which bind the F1 and F2 [4Fe-4S] clusters are highlighted in pink and light blue, conserved residues between the two hydrogenases are highlighted in red and residues with similar features are outlined in blue boxes. The cys in green is involved in H⁺ transfer in CpHydA and is conserved in MeHydA. With stars are indicated the four Cys residues of two CxxCxxC motifs mutated to serines with the aim of recreating a *Chlamydomonas*-like HydA enzyme. The figure was generated with ESPript3.²¹

3.1.1 Expression and purification of *CystoSerMeHydA* construct

BL21DE3 competent cells transformed with the *pet22b-CystoSerMeHydA* plasmid were grown in similar conditions of MeHydA wild type protein. Unfortunately, the protein was not anymore soluble and a huge overexpression band of the desired protein was found into inclusion bodies (Figure 32). Several essays changing the growth medium, temperature, IPTG concentration did not lead to a soluble fraction of the target protein.

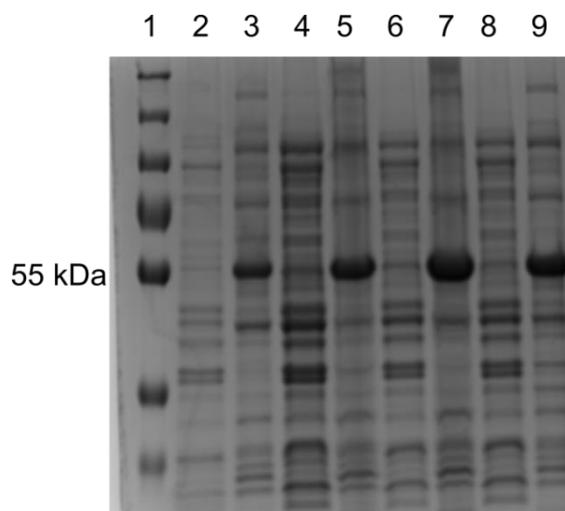


Figure 32: SDS page of *CystoSerMeHydA* protein. Lanes 2, 4 and 6 represent the soluble fraction after 2, 3 and 4 h expression in LB using BL21DE3 competent cells at 37°C after induction with 0.5 mM IPTG. Lane 8 represents the soluble fraction after 3h expression using BL21DE3 competent cells at 20°C. Lanes 3, 5, 7 and 9 are the not soluble fractions. Experiments carried out with lower IPTG concentrations and different media led to similar results.

Additional denaturation/renaturation process with an 8 M urea solution and sequent dialysis to renature the protein, led to a small soluble fraction since almost 95% of the protein precipitated during dialysis. This small amount of apo-protein was used for reconstitution and maturation assays in order to determine if the protein was able to catalyze protons reduction into hydrogen. A standard reconstitution protocol with iron, L-Cys and DTT was used to assemble the unique [4Fe-4S] cluster that should be present, but the protein was not able to bind either the [4Fe-4S]_H or [2Fe]-subcluster, normally matured via chemical insertion of the synthetic biomimetic complex **1**, [Fe₂(adt)(CO)₄(CN)₂]²⁻. Thus, the Cys to Ser mutations highly affected the protein folding and the new construct of MeHydA was not able to bind any cofactor.

The extra [4Fe-4S] clusters seemed to have not only a redox role, shuttling electrons from the surface to the catalytic H cluster, but they resulted fundamental in the folding of the N-terminal domain of the protein. A proof of the structural role of these [4Fe-4S] cluster has been indirectly provided in the last chapter with MeHydA protein. The gel filtration analysis

A truncated form of *M. elsdenii* [FeFe]-hydrogenase

of the apo and iron-sulfur reconstituted MeHydA proteins showed a high difference in terms of retention volumes, initially hypothesized as a dimer to monomer transition upon [Fe-S] cluster insertion into the protein scaffold. However, Static light scattering experiments (chapter II, paragraph 2.3) confirmed the nature of the monomeric nature of both forms of MeHydA leading to the conclusion that the three dimensional conformations of the apo- and reconstituted MeHydA were different. It is not the first time that a metal cofactor/prosthetic group can affect the protein conformation, inducing a conformational change once incorporated into the proteins scaffold.^{22,23}

3.2 Strategy 2: MeH-HydA

Failing to obtain a soluble and homogeneous form of the *H-domain* with the first strategy, the second idea pointed to the deletion of the domain holding the F clusters. In Figure 33 is shown an alignment of the sequences of HydA from *M. elsdenii* and HydA from *Clostridium pasteurianum*, which display a high degree of identity. Both contain two ferredoxin-like clusters, within CxxCxxC motifs, in the N-terminal domain and the latter has been well-characterized by X-ray crystallography, allowing a precise identification of potential sites for truncation. Based on this information we chose to start the new protein, named MeH-HydA, at residue 80 of MeHydA between the β -strand 10 and α -helix 6 of CpHydA (corresponding to the residue 208, Figure 33). The basic idea was to preserve the most important secondary structure elements which characterize normally the catalytic *H-domain* of [FeFe]-hydrogenases. This new protein, fused with a hexahistidine tag at the N-terminus, was called MeH-HydA. This new “*Chlamydomonas-like*” construct should have a molecular weight of 45.8 kDa, representing the smallest hydrogenase enzyme ever characterized and could provide key structural elements to build an artificial hydrogenase.

3.2.1 Cloning experiment

The plasmid pT7-7-MeH-HydA has been provided by Dr. Mohamed Atta. The PCR fragment was purified with the High Pure PCR kit (Roche Applied Science), double-digested with NdeI and HindIII (Fermentas), and gel-purified before direct cloning into the pT7-7 expression vector, leading to the pT7-7-MeH-HydA plasmid (Amp resistance, 6His tagged).

DH5 α cells were transformed with the plasmid and one clone containing an insert with the expected size was selected for sequencing.

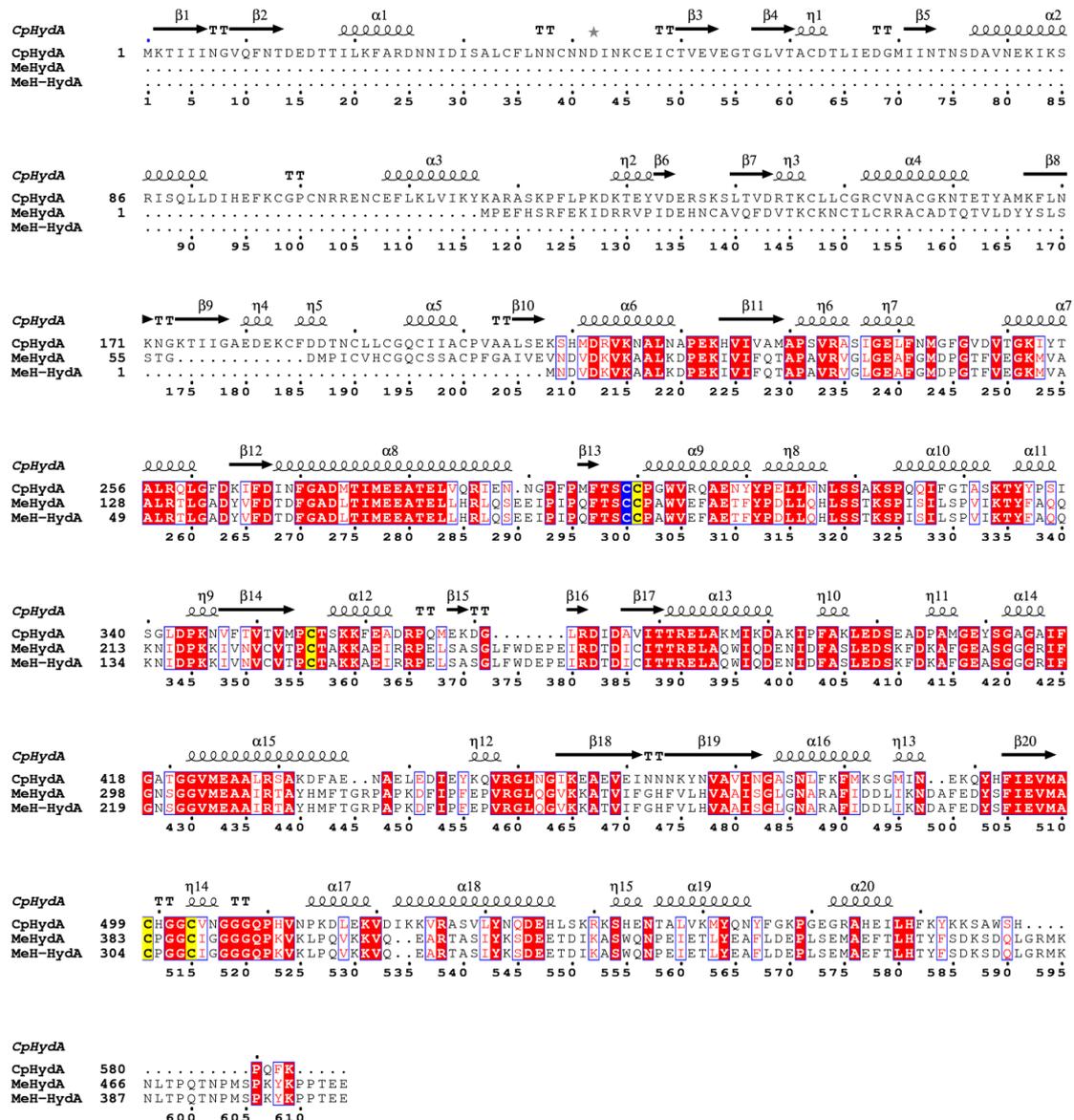


Figure 33: Sequence alignment of CpHydA with MeHydA and truncated MeH-HydA protein. Conserved secondary structures are calculated using the available X-ray structure of CpHydA (pdb 3c8y). The cysteine residues harboring the [4Fe-4S]_H cluster are highlighted in yellow, conserved residues are highlighted in red and residues with similar features are outlined in blue boxes. The cys in blue is involved in H⁺ transfer in CpHydA and is conserved in MeHydA. The figure was generated with ESPrnt3.²¹

3.2.2 Expression and aerobic purification of apo-MeH-HydA

This strategy revealed to be successful. Indeed, expression of MeH-HydA in *E. coli* resulted in a large production of soluble recombinant protein. TunerDE3 competent cells were found to be a perfect host for the expression of a soluble form of MeH-HydA and IPTG

A truncated form of *M. elsdenii* [FeFe]-hydrogenase

concentration, temperature and different grow medium essays provided an optimized protocol for high yield protein production. Details are reported in chapter V (paragraph 5.2.4). MeH-HydA could be purified to homogeneity with only two chromatographic steps (Figure **34**); after a His tag affinity column (Hi Trap column, Figure **34a**) and a size exclusion chromatography through a Superdex S200 16/600 (Figure **34b**), the protein was obtained in a pure and homogeneous form. An SDS page allowed the identification of the fractions containing the MeH-HydA protein (MW \approx 45.8 kDa), which were concentrated to 5-10 mg/ml and stored at -80 °C. Overall 15-20 mg could be obtained from 1 L of culture.

All the purification steps were carried out aerobically and dithiothreitol (DTT) was needed during the gel filtration step in order to limit the formation of protein oligomers, certainly produced by S-S bridges. A calibration with protein standards of defined molecular weight conducted in anaerobic conditions, inside a glove box, suggested that the protein is monomeric and the SEC profile provided a clear evidence of homogenous sample (Figure **34c** and **34d**).

The purified protein contained some negligible Fe (0.2 ± 0.1 Fe per polypeptide chain), thus was mainly in the apo-form and was called apo-MeH-HydA. A UV-Visible spectrum has been registered and proved the small absorption around 400-410 nm, normally assigned to the presence of [Fe-S] clusters (Figure **35a**).

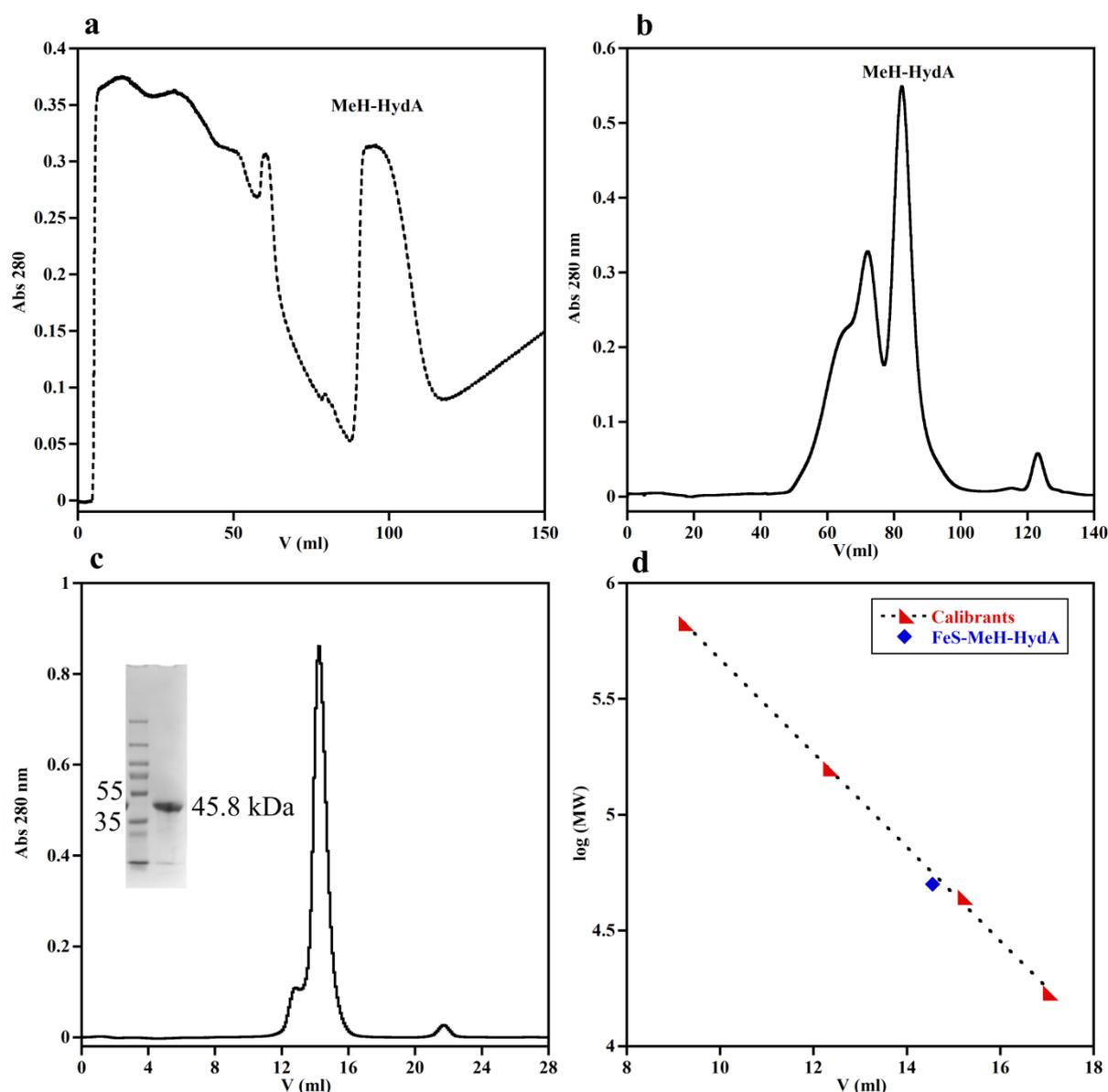


Figure 34: Apo-MeH-HydA was purified to homogeneity in 2 chromatographic steps. **a)** Ni-NTA chromatographic step. **b)** Size exclusion chromatography of apo-MeH-HydA on a Superdex S200 16/600 in 50 mM Tris-HCl buffer pH 8.0, 200 mM NaCl, 5 mM DTT. **c)** Anaerobic size exclusion chromatography elution profile of apo-MeH-HydA on an analytical Superdex S200 10/300 in 50 mM Tris-HCl pH 8.0 containing 300 mM NaCl, 10% glycerol and 5 mM DTT. The elution volume of the apo-MeH-HydA is 14.55 mL. **d)** Superdex S200 10/300 GL calibration curve using Thyroglobulin (670 kDa, 9.22 mL), γ -globulin (158 kDa, 12.32 mL), Ovalbumin (44 kDa, 15.2 mL) and Myoglobin (17 kDa, 17.03 mL). The molecular weight derived for apo-MeH-HydA suggests that the protein is monomeric.

3.2.3 $[4\text{Fe}-4\text{S}]_{\text{H}}$ cluster reconstitution of MeH-HydA apoprotein

As described for MeHydA, the incorporation of the $[4\text{Fe}-4\text{S}]$ cluster into MeH-HydA apoprotein has been carried out using a standard protocol, consisting of the anaerobic incubation

A truncated form of *M. elsdenii* [FeFe]-hydrogenase

of the protein with an excess of iron ammonium sulfate (Mohr salt) and L-cysteine in the presence of a reducing agent, the dithiothreitol (DTT). The reaction was initiated by the addition of a catalytic amount of CsdA, a cysteine desulfurase, and monitored by UV-Visible spectroscopy (Figure 35b).²⁴ Unlike MeHydA wild type, MeH-HydA should hold a single [4Fe-4S] cluster, that of the H-cluster, thus the molar excess of Fe and L-cys was decreased to 5-5.5 (see chapter V, paragraph 5.4).

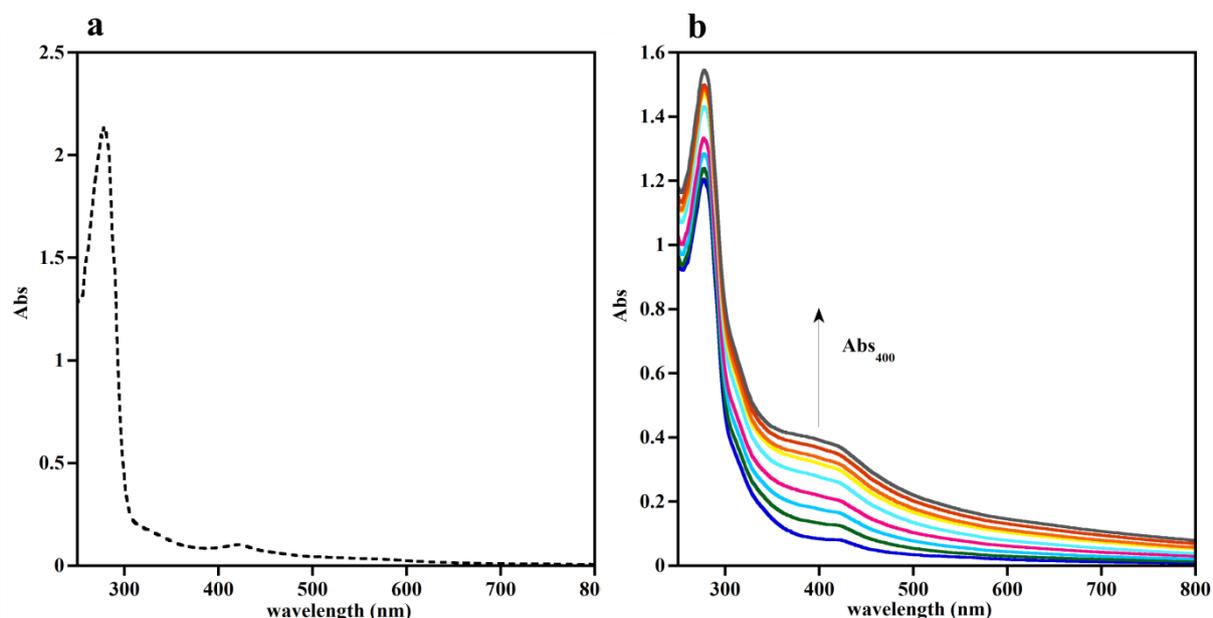


Figure 35: *In vitro* [Fe-S] reconstitution of apo-MeH-HydA protein. **a)** UV-visible spectrum of 25 μM apo-MeH-HydA in 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM DTT in a 1 cm pathlength cuvette. **b)** UV-visible absorption spectra of 100 μM MeH-HydA in 50 mM Tris-HCl pH 8, 200 mM NaCl 5 mM DTT using a 1 mm pathlength cuvette recorded every 20 minutes.

Size exclusion chromatography showed the formation of some oligomers, which could be removed due to the different elution volumes (Figure 36, left). Only the protein fraction corresponding to the brownish monomer, confirmed by a calibration curve, was collected, concentrated and further characterized. The deletion of the N-terminal domain decreased the amount of oligomeric fraction produced during the reconstitution process compared to the full length MeHydA protein (Figure 23). Moreover, the elution volumes of apo and iron-sulfur reconstituted form (called FeS-MeH-HydA) of MeH-HydA were exactly the same (14.55 ml), leading to the hypothesis that the different retention volumes (Figure 36, right) of apo-MeHydA and FeS-MeHydA observed in the last chapter were due to the presence of the unfolded N-terminal domain.

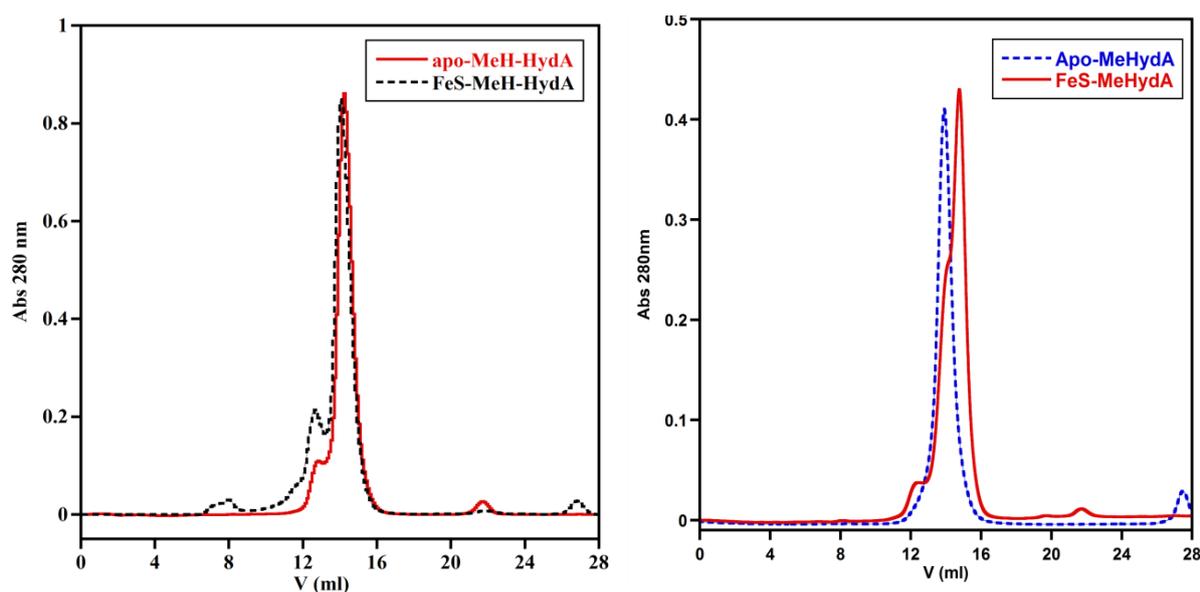


Figure 36: S200 elution profiles of apo-MeH-HydA (red, left) and FeS-MeH-HydA (dashed black, left) and apo-MeHydA (dashed blue, right) and FeS-MeHydA (red, right) in the same buffer condition, 50 mM Tris-HCl buffer pH 8.0 containing 300 mM NaCl, 10% v/v glycerol and 5 mM DTT.

3.2.4 Spectroscopic characterization of FeS-MeH-HydA

A single [4Fe-4S] cluster, that of the H-cluster, could be nicely chemically reconstituted (Figure 37, left) as shown by UV-Visible spectroscopy. The UV-Vis spectrum in the 400-410 region shows a new absorption band originated by a LMCT transition (Ligand to Metal Charge Transfer) with an absorption ratio $A_{400}/A_{280} = 0.24$. The iron and sulfide quantitation (Fe: 3.7 ± 0.2 ; S: 3.6 ± 0.2) pointed to the assembly of a single [4Fe-4S] cluster into the protein. FeS-MeH-HydA was further characterized using Q-band EPR and Mössbauer spectroscopies, in collaboration with the group of Prof. W. Lubitz where I spent a couple of weeks at the Max Plank Institute in Mülheim (Germany). FeS-MeH-HydA, as isolated, is EPR silent consistent with a $S=0$ [4Fe-4S]²⁺ state, and thus excluding other kinds of clusters, while upon anaerobic reduction with dithionite it could be converted into an EPR active state. The rhombic signal ($g = 2.043, 1.924, 1.892$) is characteristic of a $S=1/2$ [4Fe-4S]⁺ cluster (Figure 37, right). Moreover, at a g value of 2.0 a radical was detected. This radical was attributed to the aerobic purification of this protein conducted in the presence of DTT. As already observed in *Megasphaera elsdenii* full length protein, DTT in the presence of O₂ and Fe traces can produce radical species.

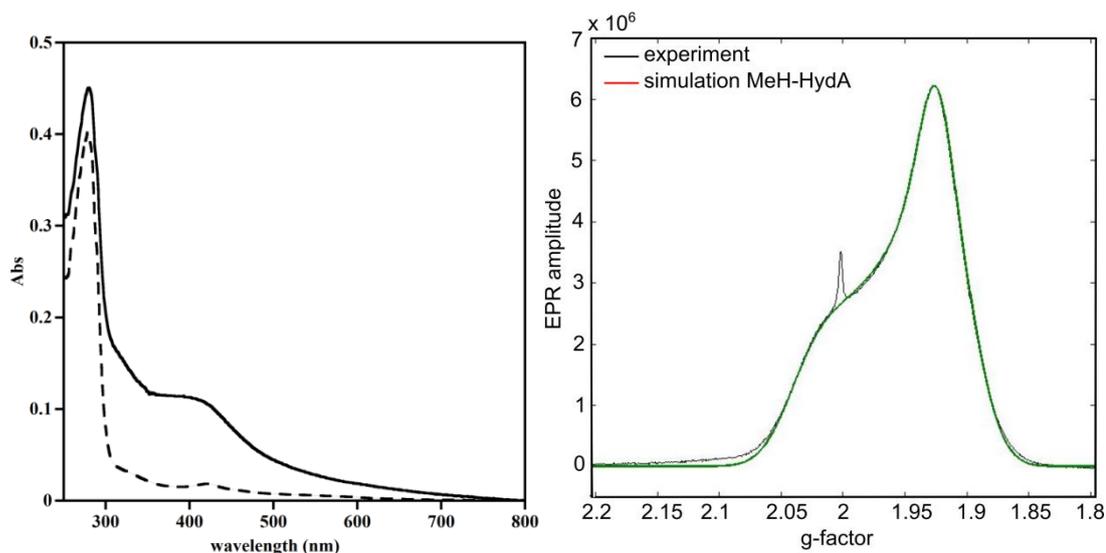


Figure 37: Spectroscopic characterization of reconstituted FeS-MeH-HydA. UV-visible spectra (left) of 80 μM apo-MeH-HydA (dashed line) and FeS-MeH-HydA (solid line) in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% v/v glycerol, 5 mM DTT using a 1 mm pathlength cuvette. The UV-Visible spectrum of apo-MeH-HydA has been registered in 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM DTT. Q-band FID-detected EPR spectrum (10K) of 270 μM FeS-MeH-HydA after reduction with 10 mM sodium dithionite (right).

In addition, the protein was characterized by Mössbauer spectroscopy after reconstitution of the cluster with ^{57}Fe . The $^{57}\text{Fe}(\text{SO}_4)$ replaced Mohr salt and the reconstitution was executed in the same conditions. $^{57}\text{Fe}(\text{SO}_4)$ has been provided by Dr. E. Mulliez from the *Laboratoire de Chimie et Biologie des Métaux* in Grenoble as 0.62 M solution in sulfuric acid and it has been stored in glove box. Before starting the reconstitution, the ^{57}Fe solution was diluted to 0.023 M in 1.0 M Tris-HCl buffer pH 8.0 to basify the solution. The reconstitution and purification of MeH-HydA protein was then carried out as described in paragraph 3.2.3. The Mössbauer spectrum is shown in Figure 38. It can be simulated with a single pure $[\text{4Fe-4S}]^{2+}$ cluster, having two valence-delocalized $[\text{Fe}_2\text{S}_2]^+$ pairs of the same intensity. The chemical shift, δ , and the quadrupole splitting, ΔQ , for both components are reported in panel **b**.

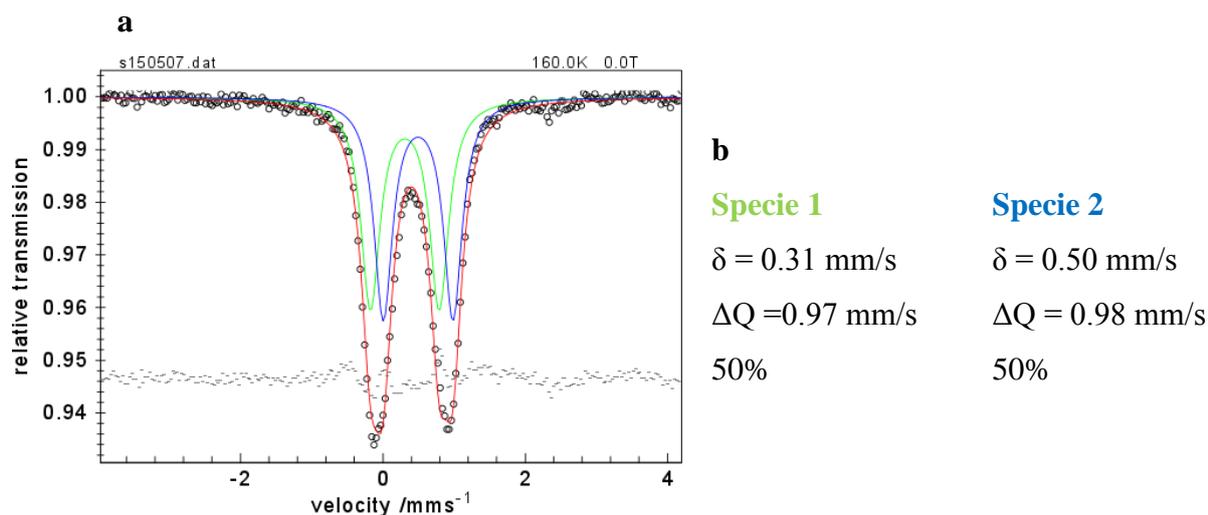


Figure 38: Mössbauer spectrum of $^{57}\text{FeS-MeH-HydA}$ protein. The experimental spectrum (black trace) was recorded at 160 K. In green and in blue is shown the simulation with two valence-delocalized $[\text{Fe}_2\text{S}_2]^+$ pairs. The protein concentration is 0.6 mM. The isomer shift and quadrupole splitting parameters are listed in panel b.

The combined results of Fe and S quantitation, UV-Visible, EPR and Mössbauer spectroscopies allow concluding that the new construct MeH-HydA can hold a single $[\text{4Fe-4S}]$ cluster, which can be isolated in the +2/+1 redox states. These data clearly establish that, while lacking the whole N-terminal domain, MeH-HydA can properly assemble a redox-active $[\text{4Fe-4S}]_{\text{H}}$ cluster within the H-domain. In the following paragraphs, we show that the cluster is also functional as it can anchor the catalytically active $[\text{2Fe}]$ -subcluster.

3.2.5 Chemical maturation of FeS-MeH-HydA with $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ synthetic complex

Applying the same methodology which led to a successful maturation of MeHydA full length in the last chapter, FeS-MeH-HydA was treated with an excess of the synthetically prepared diiron complex **1** for one hour anaerobically and then purified by desalting (NAP 10) in order to remove the excess of the chemical.^{25,26} The reaction mixture was then used for H_2 production assays to verify the effective insertion of the chemical into the apo-enzyme. H_2 production was determined using the standard methyl viologen/sodium dithionite assay in 100 mM potassium phosphate pH 6.8 via GC chromatographic system. The amount of hydrogen produced by MeH-HydA was quantified using a calibration curve. A specific hydrogenase activity of $5.5 \pm 1.0 \mu\text{mol H}_2 \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$, that corresponds to a TOF of 5.0 turnovers per

A truncated form of *M. elsdenii* [FeFe]-hydrogenase

second (MW 45.8 kDa), was obtained for holo-MeH-HydA (Table 3). The activity was 120-fold lower compared to holo-MeHydA.

Why truncated MeH-HydA is not as active as the full length protein? The lower activity could be due to either a low incorporation of the synthetic complex or to the absence of the F clusters, which could have compromised the electron shuttling between surface and H-cluster.

Fe and S content measured after the chemical maturation (Fe: 3.7 ± 0.3 , S: 3.8 ± 0.2) proved that very little complex was incorporated inside the protein. Indeed, no FTIR bands could be detected analysing 1.0 mM protein solution in the zone of CO ($1800\text{-}2020\text{ cm}^{-1}$) and CN ($2040\text{-}2100\text{ cm}^{-1}$) vibrations, where normally different states of the H-cluster are detected in hydrogenase preparations. Thus, the incorporation of the biomimetic complex was too small to be detected by FTIR spectroscopy.

Table 3: Hydrogenase activity assays.

Experimental conditions	Specific activity ($\mu\text{mol H}_2 \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$)
MeH-HydA + 10 molar excess of complex 1 , 2 h at 21°C pH 6.8	5.5 ± 1.0
MeH-HydA + 10 molar excess of complex 1 , 16-24h at 21°C pH 6.8	3.5 ± 0.5
MeH-HydA + 10 molar excess of complex 1 , 2h at 37°C pH 6.8	2.5 ± 0.4
MeH-HydA + 10 molar excess of complex 1 , 16-24h at 37°C pH 6.8	1.5 ± 0.6
MeH-HydA + 10 molar excess of 1 -HydF, 2h at 21 °C pH 6.8	4.3 ± 0.5

In bold the best condition. The data shown are mean values \pm s.d. In bold the best activity assay.

Different parameters were varied during the maturation process: temperature, incubation time and excess of the synthetic complex $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ but anyone of these parameters could enhance the hydrogen evolution activity. Indeed the increase of temperature (from 21 to 37°C), the incubation time (from 2 to 16-24 h) and the excess of the synthetic complex was greatly detrimental to the reaction. HydF protein was also used as it acts as the *in vivo* scaffold and carrier of a precursor of the [2Fe]-subcluster. An HydF hybrid (described in the next chapter) synthesized through the incorporation of the synthetic complex, **1**, onto the HydF from *Thermosipho melanesiensis* (**1**-TmeHydF) was incubated with the immature MeH-HydA without any effect on the hydrogen evolution (Table 3, Table 4).²⁶

3.2.5.1 MeH-HydA: new maturation conditions led to higher maturation

Any of the parameters varied in the last paragraph allowed a higher incorporation of the synthetic complex into MeH-HydA. Why the biomimetic complex is not entering onto the protein scaffold?

EPR and Mössbauer characterizations of FeS-MeH-HydA showed that the [4Fe-4S] cluster was correctly assembled and proved to be redox active since it was reduced in the presence of sodium dithionite. Using the software *PHYRE2 Protein fold recognition server*, a model of the wild type MeHydA protein was realized and was used to calculate the charge distribution at pH 7.0, as the chemical maturation is performed in neutral pH conditions (Figure 39). Interestingly, the protein model resulted to be highly negative charged and with *ExPasy – ProtParam* tool software a theoretical pI of 5.3 was estimated according to the protein sequence. The N-terminal domain of MeHydA, according to the model, seems to be positively charged on the surface (blue zone in Figure 39) and has been deleted in the new construct MeH-HydA, since it harbored the extra F [4Fe-4S] clusters. Thus, the surface charge distribution of MeHydA protein helped to rationalize the results of maturation of the truncated protein, where almost no hydrogenase activity was found by GC methods, and FTIR and Fe/S content were not able to evidence any incorporation of $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex.

The following modifications provided a much higher incorporation of the synthetic complex.

First, replacing the reconstitution buffer (50 mM Tris-HCl buffer pH 8.0, 200 mM NaCl, 5 mM DTT) with 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10% v/v glycerol, 5 mM DTT which represent the reconstitution conditions of the full length MeHydA and executing the chemical insertion of the synthetic complex $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ in the usual conditions (100 mM Potassium phosphate buffer pH 6.8) led to an enhancement of the activity to $25 \pm 2 \mu\text{mol H}_2 \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$. However, this activity was not enough for FTIR characterization since we did not observe any CO and CN stretch from this protein preparation, thus indicating still low amount of complex inside the protein scaffold.

A truncated form of *M. elsdenii* [FeFe]-hydrogenase

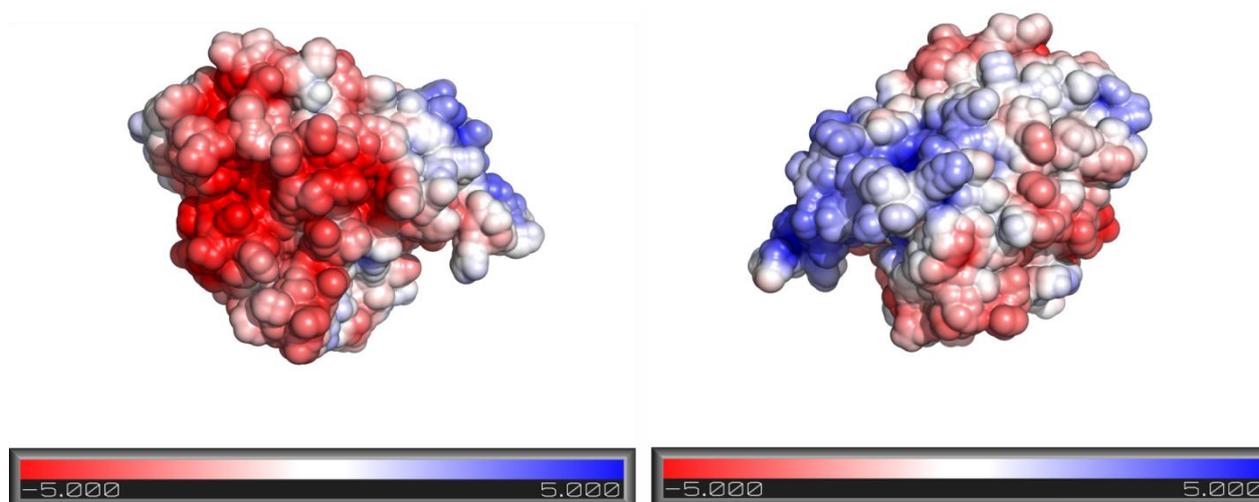


Figure 39: Charge distribution profile of MeHydA at pH 7.0 using a model produced by Phyre software using as template the X-ray structure of CpHydA (pdb 3c8y).

Second, since the N-terminal domain (positively charged) has been removed, the theoretical pI of the protein moved from 5.3 to 4.9 (as calculated using *ExPasy – ProtParam* software). The chemical maturation is normally executed at pH of 6.8, where the protein should be highly negatively charged. Moreover, the biomimetic complex has a negative charge (2-) and coulombic repulsions between the protein and the diiron biomimetic complex could prevent the enzyme maturation.

Thus, we decided to vary the pH in order to study the charges effect on the maturation conditions while the pH of the hydrogen evolution assays was kept constant (pH 6.8) to direct compare all the activities. A study of maturation conditions from pH 8 to 5 allowed finding new condition with increased complex incorporation. The protein solutions matured in different pH were assayed for hydrogen evolution in 100 mM potassium phosphate pH 6.8, which identified the best experimental conditions. As reported in Table 4, at pH 8.0 the insertion of $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ was lower compared to the conditions reported earlier with a specific activity of $2.2 \pm 0.3 \mu\text{mol H}_2.\text{mg}^{-1}.\text{min}^{-1}$. Moving the pH to 6.8 led to the activity of $23 \pm 3.5 \mu\text{mol H}_2.\text{mg}^{-1}.\text{min}^{-1}$, indicating a better incorporation at slightly more acidic conditions. Finally, at pH 6.0 a huge hydrogen evolution was observed with a specific activity of $125 \pm 15 \mu\text{mol H}_2.\text{mg}^{-1}.\text{min}^{-1}$, only 4-fold lower compared to the full length MeHydA protein. Longer incubation times or higher temperature did not help and larger excesses of the synthetic complex were greatly detrimental to the reaction (Table 4). The specific activity was slightly improved to $135 \pm 10 \mu\text{mol H}_2.\text{mg}^{-1}.\text{min}^{-1}$ when the excess of the chemical was removed, proving the inhibition effect of CO molecules released by the excess

of free complex **1** in solution. Finally, iron quantitation proved the H-cluster assembly (Fe: 4.6 ± 0.2), with almost 50-70% of [2Fe]-subcluster incorporation.

Further increase of the pH to 5.5 and 5 values led to the precipitation of MeH-HydA protein.

Table 4: Specific hydrogenase activity of MeH-HydA in different maturation conditions.

pH of maturation	pH activity	T (°C)	Specific activity ($\mu\text{mol H}_2 \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$)	Incubation time (h)
6.8	6.8	21	23 ± 3.5	2
6.8*	6.8	21	$20 \pm 2.0^*$	2
8	6.8	21	2.2 ± 0.3	16-24
6.8	6.8	21	17.5 ± 2.5	16-24
6.8	6.8	37	3.2 ± 0.8	16-24
6.8	6.8	37	2.5 ± 0.5	2
6	6.8	21	125 ± 15	2
6**	6.8	21	80 ± 10	2
6***	6.8	21	135 ± 10	2
5.5****	/	/	/	/
5****	/	/	/	/

The data shown are mean values \pm s.d. * The maturation has been carried out with 10 molar excess of the hybrid **1**-HydF in the same buffer conditions. ** The excess of $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ was increased to 100 folds the protein concentration. In all the other tests the excess of the complex is 10-fold the protein concentration. *** In bold the best condition where the excess of the chemical has been removed. **** Protein precipitated.

3.2.6 FTIR and EPR characterization of chemically matured MeH-HydA

In order to confirm that the H-cluster has been correctly assembled, but also to provide a full characterization of this “new” hydrogenase, we used Fourier Transform Infrared (FTIR) spectroscopy for characterization.

FTIR is appropriate for identifying the CO ($1800\text{-}2020 \text{ cm}^{-1}$) and CN ($2040\text{-}2100 \text{ cm}^{-1}$) vibrations associated with the CO and CN⁻ ligands present in the [2Fe]-subcluster. These signals can be clearly distinguished from those of free complex **1** in solution which are much broader and appear in a zone free of other protein signals.²⁶ The narrow CO and CN FTIR

A truncated form of *M. elsdenii* [FeFe]-hydrogenase

stretches of the H-cluster, in addition, allow differentiating between the various states usually coexisting in hydrogenase preparations.^{3,26}

The new maturation conditions, described in the last paragraph, were used to prepare active MeH-HydA for spectroscopic studies. Once activated to better compare this new hydrogenase to the wild type MeHydA, the protein solution was changed to 100 mM potassium phosphate pH 6.8.

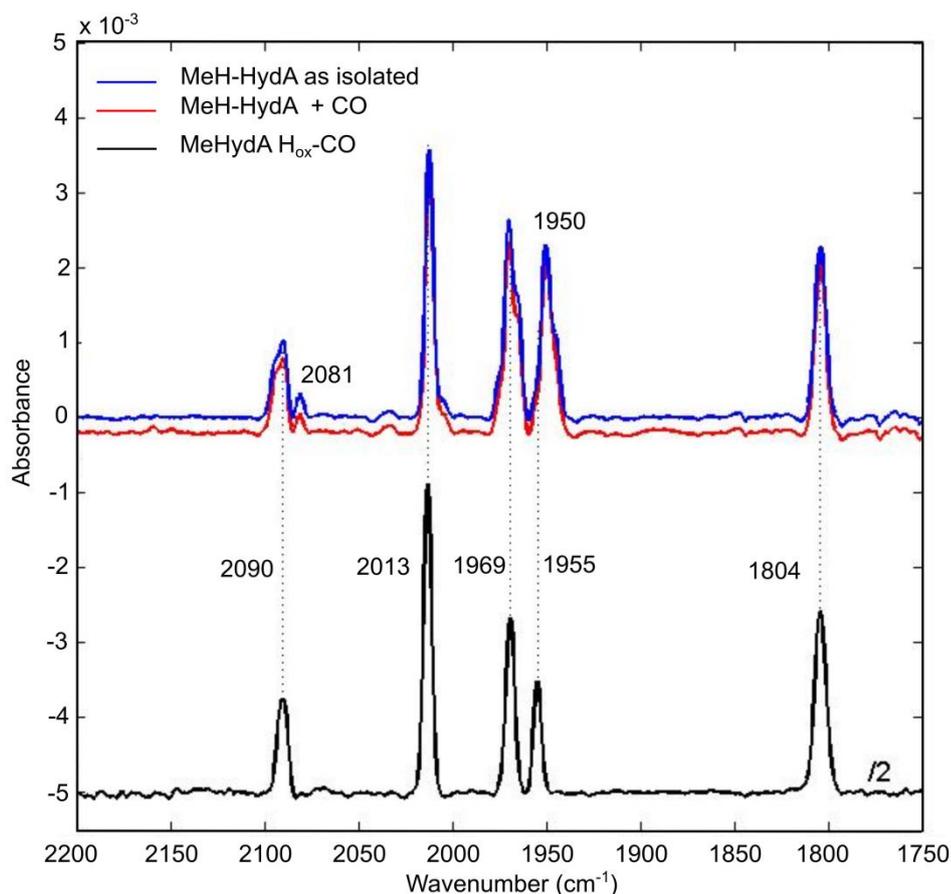


Figure 40: FTIR spectra of chemically matured MeH-HydA. FTIR spectra of 1.0 mM as isolated MeH-HydA and its H_{ox}-CO inhibited state in 100 mM potassium phosphate pH 6.8 (top). H_{ox}-CO inhibited state of chemically matured MeHydA (bottom) in the same buffer conditions.

The FTIR spectrum of as isolated sample essentially showed a H_{ox}-CO inhibited state (Figure 40). Under synthetic maturation, the supernumerary CO is released from the complex **1** ([Fe₂(adt)(CO)₄(CN)₂]²⁻) resulting in a substantial amount of the H_{ox}-CO state. This signal is also often present in [FeFe]-hydrogenase preparations due to the so-called “cannibalization process” in which the CO ligands, from light- or oxygen-damaged H-clusters, are released and captured by still intact H-clusters.²⁷⁻²⁹ Thus, additional CO is “leaking” from the excess

binuclear complex present in the maturation buffer further increasing the H_{ox} -CO population. In addition, in the new maturation conditions (100 mM potassium phosphate pH 6.0) the synthetic complex proved to be quite unstable and decomposed in some hours in water, pH 6.0, conditions.

The sample was flushed for 20 minutes with CO gas and the obtained spectrum was exactly the same of the as isolated protein sample. A comparison of this H_{ox} -CO inhibited state with the equivalent obtained for the full length MeHydA protein shows similar features although some stretches were slightly broader (Figure 40, 1950 cm^{-1} and 1969 cm^{-1}). Probably the stretches at 2081 cm^{-1} and that broader around 1950 cm^{-1} are coming from the H_{ox} redox state contribution.

These are only preliminary results and further FTIR experiments on other redox states are still in progress. However, it clearly confirms that the H-cluster is correctly assembled although it seems that the latter is more sensitive to CO binding than the corresponding full length MeHydA⁸

MeH-HydA protein was also analyzed by Q-band EPR spectroscopy which can easily distinguish between various oxidations states coexisting in hydrogenase preparations. Figure 41 shows the EPR spectrum of the as-isolated MeH-HydA which could be simulated using the H_{ox} and H_{ox} -CO spectra (Figure 41 left). This showed that MeH-HydA was mainly in the H_{ox} -CO inhibited state. A nearly isotropic signal ($g = [2.02, 2.009, 2.007]$) is originated from the low spin mixed-valence configuration of the [2Fe]-subcluster Fe(I)Fe(II) while the [4Fe-4S] cluster is in the oxidized 2+ state (EPR silent). The spectrum showed also the presence of a small amount of H_{ox} redox state (rhombic signal with $g = [2.090, 2.033, 1.996]$). The simulated spectrum using these two spectra perfectly fit the experimental spectrum of the as isolated sample.

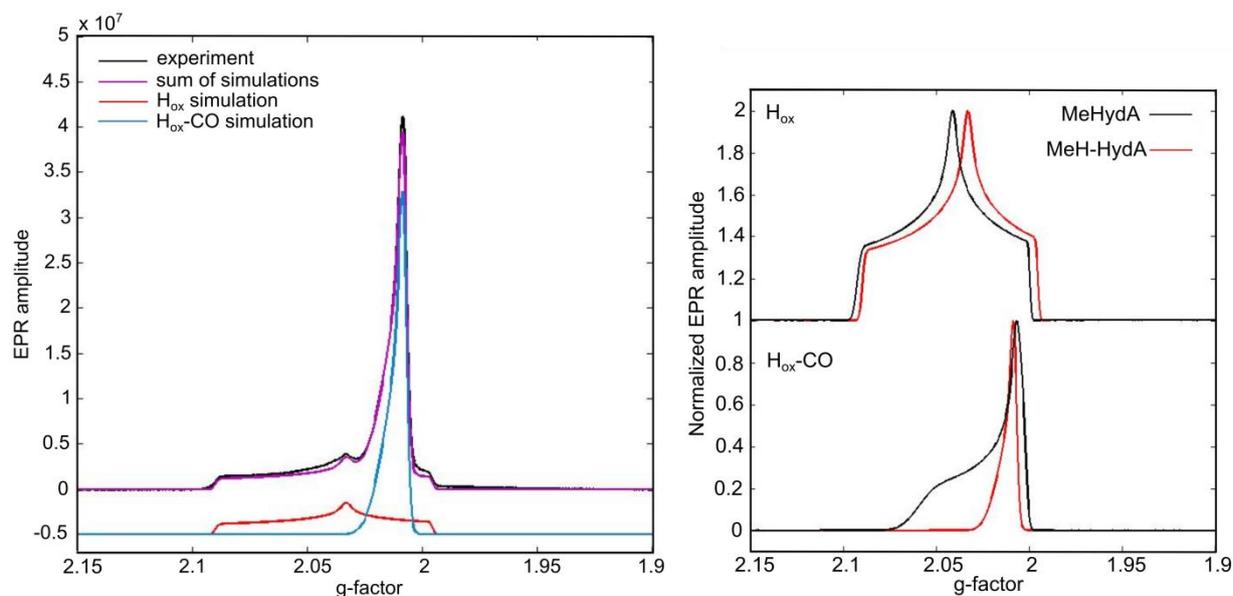


Figure 41: Q-band FID-detected EPR spectra ($T=10\text{K}$) of chemically matured MeH-HydA. EPR spectrum of 1.0 mM as isolated protein sample (left), EPR spectra of H_{ox} , and $\text{H}_{\text{ox}}\text{-CO}$ redox states (right) and comparison with EPR spectra of the same redox states in MeHydA. Protein samples are prepared in 100 mM potassium phosphate pH 6.8.

Intriguingly, the $\text{H}_{\text{ox}}\text{-CO}$ state of MeH-HydA has a nearly isotropic shape (Figure 41) whereas the corresponding state in the full length protein presents a nearly axial signal as in hydrogenases from other organisms.³ It is possible that the removal of the N-terminal domain in the truncated MeH-HydA protein changed the protein folding so that the cavity of the H-cluster resulted to be slightly affected. This could have produced a change of the symmetry of the [2Fe]-subcluster spin system. Also the H_{ox} state changed slightly compared to wild type MeHydA with a small shift of the g-factor; however, it retained the same rhombic shape.⁸

3.2.7 Conclusion and future perspectives

A very simple method for producing a “new” *Chlamydomonas*-like hydrogenase has been developed by heterologous expression in *E. coli* of the truncated MeHydA gene, MeH-HydA, aerobic purification in two steps and chemical reconstitution of the H-cluster with Fe, sulfur and the synthetic organometallic complex **1**. The successful strategy resulted into the deletion of the N-terminal domain; indeed, point mutations of the cysteines of the CxxCxxC motifs holding the two F clusters resulted in the accumulation of the protein inside inclusion bodies. Trials of denaturation and renaturation did not produce a soluble protein.

The iron-sulfur reconstituted protein, FeS-MeH-HydA, has been characterized by different spectroscopic techniques (Uv-Vis, EPR and Mössbauer) proving the effective insertion of a single [4Fe-4S] cluster, which was able to shift between +2/+1 redox states in the presence of a reducing agent.

Initial FTIR, Fe/S quantitation and H₂ evolution assays showed that only a small portion of the H-cluster of MeH-HydA could be assembled in the standard maturation conditions.

Structural model studies, using the known three-dimensional structure of CpHydA, led to the hypothesis that in the normal pH 6.8 conditions coulombic repulsions between the negatively charged biomimetic complex and the protein scaffold, negatively charged at pH 6.8, could prevent the enzyme maturation.

Finally, a shift of the pH to 6.0 increased the incorporation of the synthetic biomimetic complex as shown by Fe/S quantitation, FTIR and overall activity assays. Huge H₂ production was observed in the new maturation conditions with a specific activity of $135 \pm 10 \mu\text{mol H}_2 \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$. This high activity is only 4-fold smaller than MeHydA full length.

The preliminary spectroscopic characterization (EPR and FTIR) showed that the as isolated sample is mainly in the H_{ox}-CO inhibited configuration. This is quite unusual since all other hydrogenases matured chemically showed the presence of different oxidation states. Further studies are still in progress and could answer if this new hydrogenase protein can access to the H_{red} and H_{sred} redox states. The observation of the latter state could shed light on the mechanism of catalysis of hydrogenases containing multiple [Fe-S] clusters. Indeed, CrHydA1 which holds only the H-cluster, has access to the H_{sred} state which, instead, has never been observed in bacterial hydrogenases. This point led to the speculation that the H_{sred} state is not visible due to the electron redistribution between the H-cluster and the extra F clusters resulting in an overlapping between the H_{red} and H_{sred} states.^{3,12}

Moreover, MeH-HydA represents one of the first examples which illustrate the concept that it is possible to engineer hydrogenases. Indeed, hydrogenases despite the high efficiency, using simple iron and/or nickel, are fraught with a number of drawbacks which need to be addressed before these enzymes can be implemented into electrodes. One is their size which limits enzyme density at the surface of the electrodes. This novel hydrogenase clearly shows that is possible to modify hydrogenases to generate smaller but still active proteins.

The second drawback is oxygen sensitivity since these enzymes need to be manipulated under strictly anaerobic conditions.

A truncated form of *M. elsdenii* [FeFe]-hydrogenase

Future studies on MeH-HydA will be directed towards the analysis of the oxygen sensitivity of this enzyme, since only few proteins have been studied showing a high diversity. An electrochemical study is fundamental to investigate and compare the properties of this enzyme with the parent full length protein, MeHydA.

In a recent study of Léger C. and coworkers high difference in terms of O₂ resistance is observed between the HydA from *Clostridium acetobutylicum* (CaHydA), containing additional F clusters, and CrHydA1 containing just the [4Fe–4S]_H. CaHydA is more stable to short exposure of O₂ compared to CrHydA1 suggesting that the presence of a chain of auxiliary [Fe–S] clusters may help the deliver electrons to the active site used for O₂ reduction to H₂O.^{16,30} For that reason, a direct comparison of the oxygen resistance of MeHydA and its truncated, MeH-HydA, is absolutely required. Indeed, it might finally clarify the importance of the [Fe–S] cluster network.

References:

- 1 T. W. Woolerton, S. Sheard, Y. S. Chaudhary and F. A. Armstrong, *Energy Environ. Sci.*, 2012, **5**, 7470–7490.
- 2 J. A. Cracknell, K. A. Vincent and F. A. Armstrong, *Chem. Rev.*, 2008, **108**, 2439–2461.
- 3 W. Lubitz, H. Ogata, O. Rüdiger and E. Reijerse, *Chem. Rev.*, 2014, **114**, 4081–4148.
- 4 J. W. Peters, G. J. Schut, E. S. Boyd, D. W. Mulder, E. M. Shepard, J. B. Broderick, P. W. King and M. W. W. Adams, *Biochim. Biophys. Acta*, 2015, **1853**, 1350–1369.
- 5 C. Kamp, A. Silakov, M. Winkler, E. J. Reijerse, W. Lubitz and T. Happe, *Biochim. Biophys. Acta*, 2008, **1777**, 410–416.
- 6 S. T. Stripp and T. Happe, *Dalton Trans. Camb. Engl.* 2003, 2009, 9960–9969.
- 7 J. W. Peters, W. N. Lanzilotta, B. J. Lemon and L. C. Seefeldt, *Science*, 1998, **282**, 1853–1858.
- 8 G. Caserta, A. Adamska-Venkatesh, L. Pecqueur, M. Atta, V. Artero, S. Roy, E. Reijerse, W. Lubitz and M. Fontecave, *Biochim. Biophys. Acta*, 2016, **1857**, 1734–1740.
- 9 A. Silakov, C. Kamp, E. Reijerse, T. Happe and W. Lubitz, *Biochemistry*, 2009, **48**, 7780–7786.
- 10 D. W. Mulder, E. S. Boyd, R. Sarma, R. K. Lange, J. A. Endrizzi, J. B. Broderick and J. W. Peters, *Nature*, 2010, **465**, 248–251.
- 11 D. W. Mulder, D. O. Ortillo, D. J. Gardenghi, A. V. Naumov, S. S. Ruebush, R. K. Szilagyi, B. Huynh, J. B. Broderick and J. W. Peters, *Biochemistry*, 2009, **48**, 6240–6248.

- 12 A. Adamska, A. Silakov, C. Lambertz, O. Ruediger, T. Happe, E. Reijerse and W. Lubitz, *Angew. Chem.-Int. Ed.*, 2012, **51**, 11458–11462.
- 13 K. D. Swanson, M. W. Ratzloff, D. W. Mulder, J. H. Artz, S. Ghose, A. Hoffman, S. White, O. A. Zadornyy, J. B. Broderick, B. Bothner, P. W. King and J. W. Peters, *J. Am. Chem. Soc.*, 2015, **137**, 1809–1816.
- 14 D. W. Wakerley and E. Reisner, *Energy Env. Sci.*, 2015, **8**, 2283–2295.
- 15 M. L. Ghirardi, *Photosynth. Res.*, 2015, **125**, 383–393.
- 16 C. Orain, L. Saujet, C. Gauquelin, P. Soucaille, I. Meynial-Salles, C. Baffert, V. Fourmond, H. Bottin and C. Léger, *J. Am. Chem. Soc.*, 2015, **137**, 12580–12587.
- 17 Y. Nicolet, R. Rohac, L. Martin and J. C. Fontecilla-Camps, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 7188–7192.
- 18 R. Rohac, P. Amara, A. Benjdia, L. Martin, P. Ruffié, A. Favier, O. Berteau, J.-M. Mouesca, J. C. Fontecilla-Camps and Y. Nicolet, *Nat. Chem.*, 2016, **8**, 491–500.
- 19 J. M. Kuchenreuther, W. K. Myers, T. A. Stich, S. J. George, Y. Nejatyjahromy, J. R. Swartz and R. D. Britt, *Science*, 2013, **342**, 472–475.
- 20 M. Atta and J. Meyer, *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.*, 2000, **1476**, 368–371.
- 21 X. Robert and P. Gouet, *Nucleic Acids Res.*, 2014, **42**, W320-324.
- 22 N. A. Bushmarina, *Protein Sci.*, 2006, **15**, 659–671.
- 23 B. Gubernator, J. Króliczewski, T. Kallas and A. Szczepaniak, *Biochim. Biophys. Acta*, 2006, **1764**, 735–742.
- 24 L. Loiseau, S. Ollagnier-de Choudens, D. Lascoux, E. Forest, M. Fontecave and F. Barras, *J. Biol. Chem.*, 2005, **280**, 26760–26769.
- 25 J. Esselborn, C. Lambertz, A. Adamska-Venkatesh, T. Simmons', G. Berggren, J. Nothl, J. Siebel, A. Hemschemeier, V. Artero, E. Reijerse, M. Fontecave, W. Lubitz and T. Happe, *Nat. Chem. Biol.*, 2013, **9**, 607–609.
- 26 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.
- 27 A. Adamska-Venkatesh, D. Krawietz, J. Siebel, K. Weber, T. Happe, E. Reijerse and W. Lubitz, *J. Am. Chem. Soc.*, 2014, **136**, 11339–11346.
- 28 S. P. J. Albracht, W. Roseboom and E. C. Hatchikian, *J. Biol. Inorg. Chem. JBIC Publ. Soc. Biol. Inorg. Chem.*, 2006, **11**, 88–101.

A truncated form of *M. elsdenii* [FeFe]-hydrogenase

- 29 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.
- 30 A. Kubas, C. Orain, D. De Sancho, L. Saujet, M. Sensi, C. Gauquelin, I. Meynial-Salles, P. Soucaille, H. Bottin, C. Baffert, V. Fourmond, R. B. Best, J. Blumberger and C. Léger, *Nat. Chem.*, 2016.

Chapter IV

Results and discussion

The [FeFe]-hydrogenase maturase HydF

Chapter IV: The [FeFe]-hydrogenase maturase HydF

In this chapter are presented the results obtained working on different HydF proteins. Several questions are still missing about this maturase of the [FeFe]-hydrogenase and overall structural information of the metal binding site.

As introduced in chapter I, extensive work from various laboratories has established that at least three maturases are required for [FeFe]-hydrogenase maturation.¹ HydG, a Radical-SAM enzyme, synthesizes CO and CN⁻ from tyrosine and assembles a Fe(CO)₂(CN) unit which ends up into HydA.²⁻⁴ HydE, also a Radical-SAM enzyme, has been well characterized by X-ray crystallography^{5,6} and is assumed to be responsible for the synthesis of adt²⁻, whose precursor is still unknown.^{7,8} Finally, HydF, a [4Fe-4S] protein with GTPase activity,⁹ combines the components provided by HydG and HydE and assembles a diiron unit, which is subsequently transferred to the hydrogenase for activation.¹⁰ Thus HydF functions as a scaffold/carrier protein. Accordingly, HydF from *C. acetobutylicum* isolated from an *E. coli* strain heterologously expressing the three maturation proteins from *C. acetobutylicum* is sufficient to confer hydrogenase activity to a purified inactive HydA protein, named apoHydA, heterologously expressed in the absence of HydE, HydG and HydF, thus lacking the [2Fe]-subcluster.¹¹ These results indicate that when expressed in the presence of HydE and HydG, HydF is obtained in a form which contains all the requested chemical elements for activating HydA.

The exact nature of the key transient diiron unit in HydF is unknown but obviously is likely to resemble the final [2Fe]-subcluster in HydA. Recently, Berggren G. and coworkers showed that HydF, from *Thermotoga maritima*, was able not only to bind the synthetic complex [Fe₂(adt)(CN)₂(CO)₄]²⁻ (**1**) mimicking the [2Fe]-subcluster (Figure **42b**), generating the so-called “**1**-HydF hybrid” (Figure **42c**), but also to transfer it to apo-HydA and fully activate it.^{12,13} A structure of HydA after chemical maturation unambiguously showed that a correct [2Fe]-subcluster has been assembled.¹⁴ Thus even though **1** is not identical to the [2Fe]-subcluster, it can undergo specific modifications during the transfer and it is tempting to suggest that **1** is identical to the natural transient diiron unit in HydF. However, this hypothesis needs to be further experimentally validated and has been addressed in this thesis.

The [FeFe]-hydrogenase maturase HydF

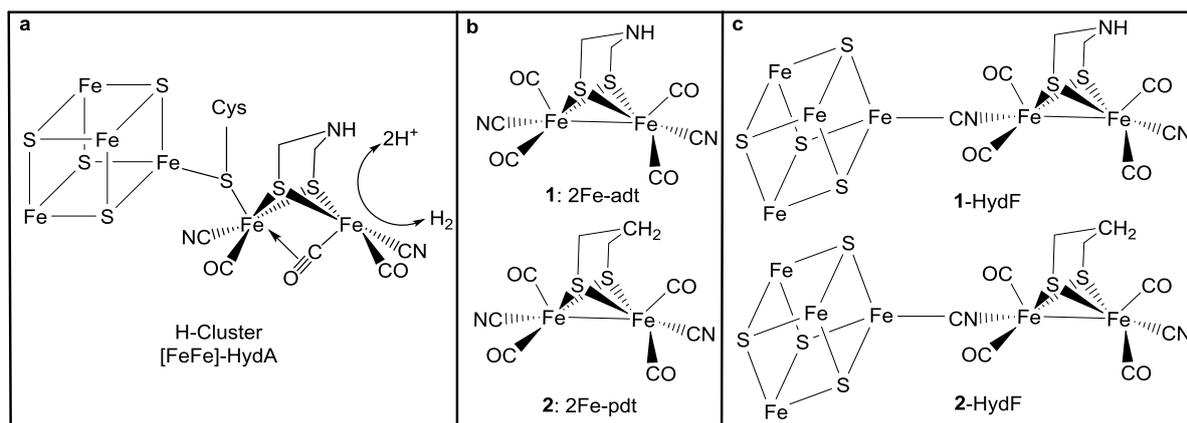


Figure 42: Structures of the H-cluster, chemical analogs of the 2Fe-subcluster and HydF hybrids described in this study. (a) Structure of the H-cluster of [FeFe]-hydrogenases; (b) Structure of complex **1** = [Fe₂(adt)(CO)₄(CN)₂]²⁻ and complex **2** = [Fe₂(pdt)(CO)₄(CN)₂]²⁻; (c) Structure of the 6Fe-units in **1**-HydF and **2**-HydF hybrids.

While HydE⁵ and HydG³ have been structurally characterized with their metallic cofactors, this structural information is still missing for HydF. Thus the [4Fe–4S] cluster coordination in HydF is unknown and, as a consequence, has been a matter of intense speculation. For the same reason, the mechanism driving assembly of the intermediate diiron unit and its transfer to HydA is also unknown. The only three-dimensional structure available is that of apo-HydF, lacking the [4Fe–4S] cluster, from *Thermotoga neapolitana*, however in an extensively oxidized state since the putative cysteine ligands are involved in inter- and intra-molecular disulfide bridges (Figure 7).⁹

In this chapter, I describe the results obtained with HydF proteins from the thermophilic organisms *Thermotoga maritima*, *Thermosipho melanesiensis* and *Clostridium thermocellum*, the latter two studied for the first time. In particular, we succeeded to obtain the crystal structure of the iron-sulfur form of HydF from *Thermosipho melanensiensis* organism (TmeHydF), showing unusual coordination to the [4Fe–4S] cluster. This original structure is compared to those of apo-HydF from *T. neapolitana* (Tn) and from *T. maritima* (Tm), the latter also solved in this work and reported in the chapter. Furthermore, in collaboration with the group of Professor W. Lubitz, we characterized TmeHydF protein using EPR, HYSCORE and FTIR spectroscopies. These results, described in the following, provide new insights into the HydA maturation process and for the first time the conversion of HydF into an artificial hydrogenase.

4.1 HydF protein from *Thermotoga maritima*

HydF from *Thermotoga maritima* is one of the best characterized HydF protein and was shown to bind $[\text{Fe}_2(\text{xdt})(\text{CN})_2(\text{CO})_4]^{2-}$ synthetic complexes. Since the purification protocol was well-established and the protein could be obtained with high yield in a homogenous form, TmHydF was chosen for crystallization trials and as scaffold for anchoring diiron catalysts evaluating their protons reduction activity. The results are described in the following.

4.1.1 Expression, purification and iron-sulfur reconstitution of HydF from *Thermotoga maritima*

A plasmid of HydF from *Thermotoga maritima* was available in the laboratory (provided by Dr. M. Atta) and the protein (TmHydF) was overexpressed and purified as previously described,¹⁵ replacing the HiPrep 26/10 desalting column (GE Healthcare) with a Superdex S200 26/600 equilibrated with 25 mM Tris-HCl pH 8.0, 200 mM NaCl. The last chromatographic step allowed the separation of the dimer from a small tetrameric fraction, for subsequent concentration up to 5-8 mg/ml and storage at -80 °C until use. Details of protein purification are described in chapter V (paragraph 5.3.1).

All the purification procedures were carried out aerobically. While a very small amount of iron was detected in this preparation, the protein was mainly in the apo form and was named apo-TmHydF. To incorporate the iron-sulfur cluster in apo-HydF protein, a standard protocol was used, consisting of the anaerobic incubation of the protein with an excess of iron ammonium sulfate and cysteine in the presence of DTT, as reducing agent. The reaction was initiated by addition of catalytic amounts of CsdA, a cysteine desulfurase. Visible absorption spectrum was recorded on a Cary 100 spectrophotometer (Agilent) connected to the cell holder located in a glovebox with optical fibers. A new band in the 400-410 nm region appeared after the *in vitro* reconstitution of the [4Fe-4S] cluster. After a gel filtration purification step, only the protein fractions corresponding to the dimer were collected and further characterized (Figure 43).

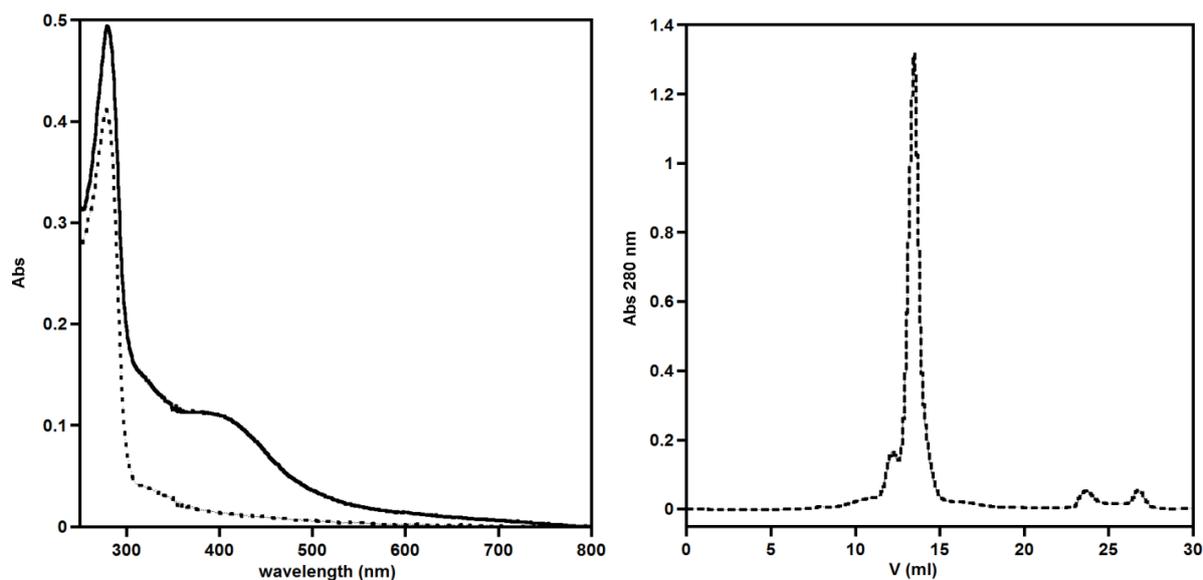


Figure 43: UV-Visible absorption spectra of 100 μM apo (dashed line) and FeS-TmHydF (black line) in 25 mM Tris-HCl Buffer pH 8.0, 200 mM NaCl, 5mM DTT (left) and SEC of apo-TmHydF on Superdex S200 10/300 in 25 mM Tris-HCl Buffer pH 8.0, 200 mM NaCl, 5mM DTT (right). TmHydF is mainly dimeric with a small tetrameric contamination, which was removed by gel filtration.

The iron-sulfur reconstituted protein (FeS-TmHydF) was previously shown to possess a [4Fe-4S] by multiple spectroscopic analyses (UV-Vis, EPR and Mössbauer)^{15,16} and the iron and sulfide quantitation, Fe (3.9 ± 0.4 per monomer) and S (3.6 ± 0.2 per monomer), were consistent with a single [4Fe-4S] cluster.

4.1.2 Insertion of synthetic complexes **1** and **2** onto TmHydF: hydrogenase activity

As already shown by Berggren G. and coworkers, iron-sulfur reconstituted TmHydF was reacted with an excess of the synthetic complexes **1** and **2**, generating HydF hybrids (Figure 42) and assayed for H₂ evolution. The protein was incubated 1h with an excess of the two chemicals and the excess of the compounds was removed via a desalting column. Iron and sulfur quantitation proved the incorporation of the two synthetic complexes inside the TmHydF protein scaffold (Table 5). Interestingly, **2**-TmHydF displayed a small but significant hydrogenase activity, while the **1**- analogue proved not active although it contained the biologically relevant adt²⁻ ligand.

Table 5 : Iron content of FeS-TmHydF before and after maturation with synthetic complexes **1** and **2**.

Specie	Fe content	S content
FeS-TmHydF	3.9 ± 0.4	3.6 ± 0.2
1 -TmHydF	5.0 ± 0.2	4.9 ± 0.2
2 -TmHydF	5.3 ± 0.3	5.1 ± 0.3

The data shown are mean values on three different protein reconstitutions and maturations.

This was shown in a standard chemical assay using sodium dithionite and methyl viologen as source of electrons (Figure **44A**), as well as in a photochemical assay, using a metal-organic photosensitizer ($[\text{Ru}(\text{bpy})_3]^{2+}$) and a sacrificial electron donor (sodium ascorbate) under visible light irradiation (Figure **44B**, $\lambda > 400$ nm). For that purpose we used a miniaturized Clark-type hydrogen microsensor, allowing direct detection of hydrogen produced in solution and convenient for measuring small activities, undetectable with common gas chromatographic methods. **2**-TmHydF allowed H_2 production with an initial TOF (turnover frequency) of $0.5\text{-}1 \text{ min}^{-1}$, about 30 times larger than compound **2** free in solution.¹³ These results, here presented, showed that **2**-TmHydF can produce some hydrogen, although this protein has been design by nature to carrier and transfer a diiron complex to apo-HydA protein.

Hydrogen evolution started adding a sodium dithionite solution (reducing agent) to a solution containing the hybrid protein and methyl viologen (electron mediator) then leveled off after few minutes, while the system was still active after 5 min reaction. The initial slope was quite high (TOF 0.5 min^{-1}) and some minutes later it started to decrease reaching a slope similar to free complex **2** in solution. This could mean that: (i) the complex is dissociated from the protein scaffold, reflecting the HydF role, or (ii) the complex is destroyed by the strong reducing conditions. I support the first hypothesis since the reached slope during H_2 evolution is similar to the one of free complex **2** in solution.

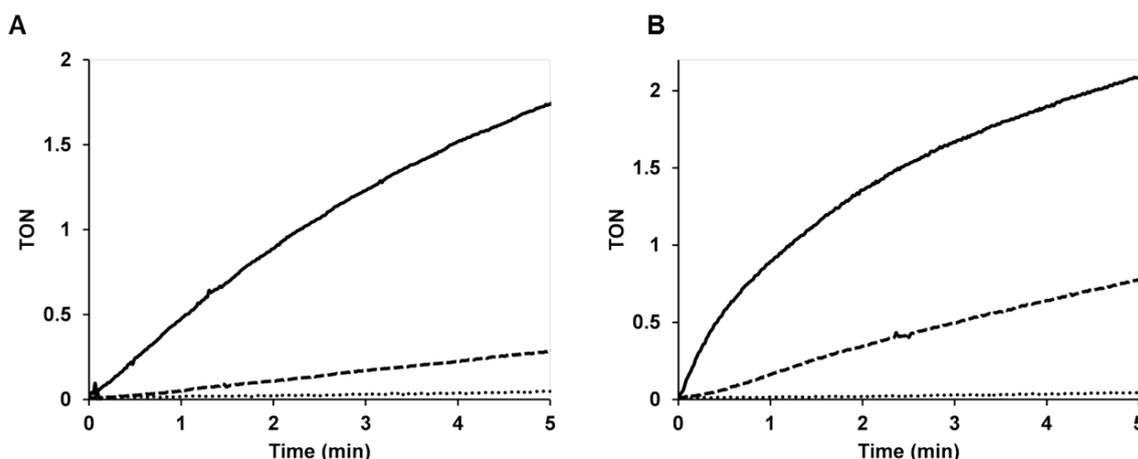


Figure 44: Hydrogen evolution catalyzed by the 2-TmHydF hybrid. **(A)** Chemical assay using sodium dithionite/methyl viologen as the reducing agent. **(B)** Photochemical assay using $[\text{Ru}(\text{bpy})_3]^{2+}$ as photosensitizer and sodium ascorbate as sacrificial electron donor. Experimental conditions are described in materials and methods. 2-HydF hybrid (full), free complex 2 (dashed), and apo-HydF (dotted). TON is defined as $\text{mol}(\text{H}_2) \cdot \text{mol}(\text{cat})^{-1}$.

Thus, the complex seems to be dissociated after few minutes of activity. This observation leads to the conclusion that a structural characterization of the active site is absolutely required in order to i) understand how the metallo-cofactors are anchored inside the protein scaffold and ii) how stabilize the biomimetic complex for a better hydrogenase activity.

4.1.3 X-ray structure of apo-TmHydF, the apof orm of HydF from *Thermotoga maritima*

The first crystallization trials were conducted on the iron-sulfur reconstituted form of TmHydF. FeS-TmHydF was concentrated to 17-18 mg/ml in 25 mM Tris-HCl Buffer pH 8.0, containing 200 mM NaCl and 5mM DTT and used for crystallization trials automated with an Oryx Nano crystallization robot (Douglas Instruments) installed in a temperature and hygrometry regulated anaerobic chamber. Hits were obtained in the PGA screen (Molecular Dimensions). While the protein prepared for crystallization contained a single [4Fe-4S] cluster, the latter was lost during crystallization as indicated by the formation of colorless crystals (Figure 45). Final crystallization conditions were 0.3 M sodium malonate dibasic monohydrate, 0.1 M sodium acetate pH 5.0, 8 % w/v γ -Poly glutamic acid, which favored

iron mobilization. Apo-TmHydF crystals were cryoprotected by soaking with a reservoir solution implemented with 30% glycerol.

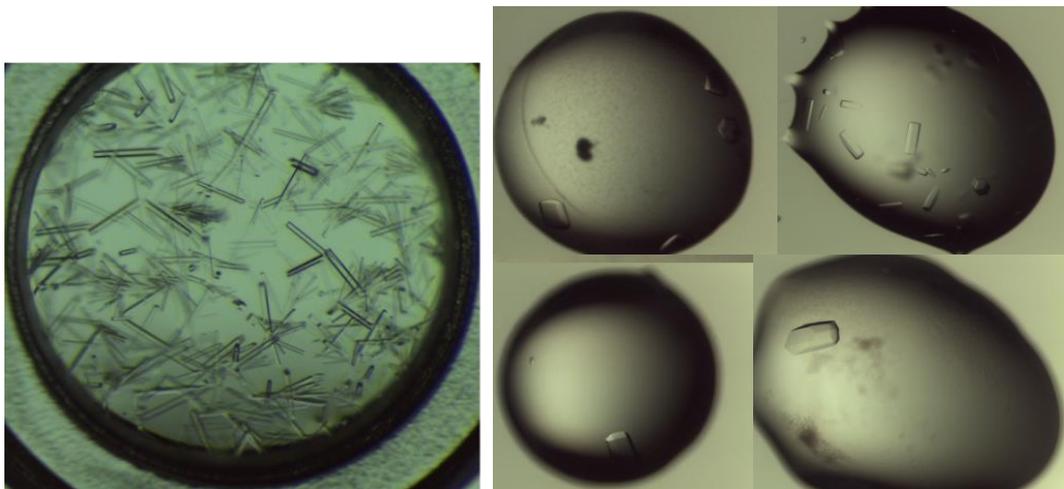


Figure 45: Pictures of apo-TmHydF crystals. The crystals on the left were reproduced manually (hanging drop) while the crystals on the right are produced by the Oryx Nano crystallization robot (sitting drop).

The crystal structure of apo-TmHydF (Figure 46) was determined at 3.0 Å resolution in the crystalline space group I4 by Dr L. Pecqueur. The structure was solved by molecular replacement using the structure of apo-HydF from *Thermotoga neapolitana* as search template.

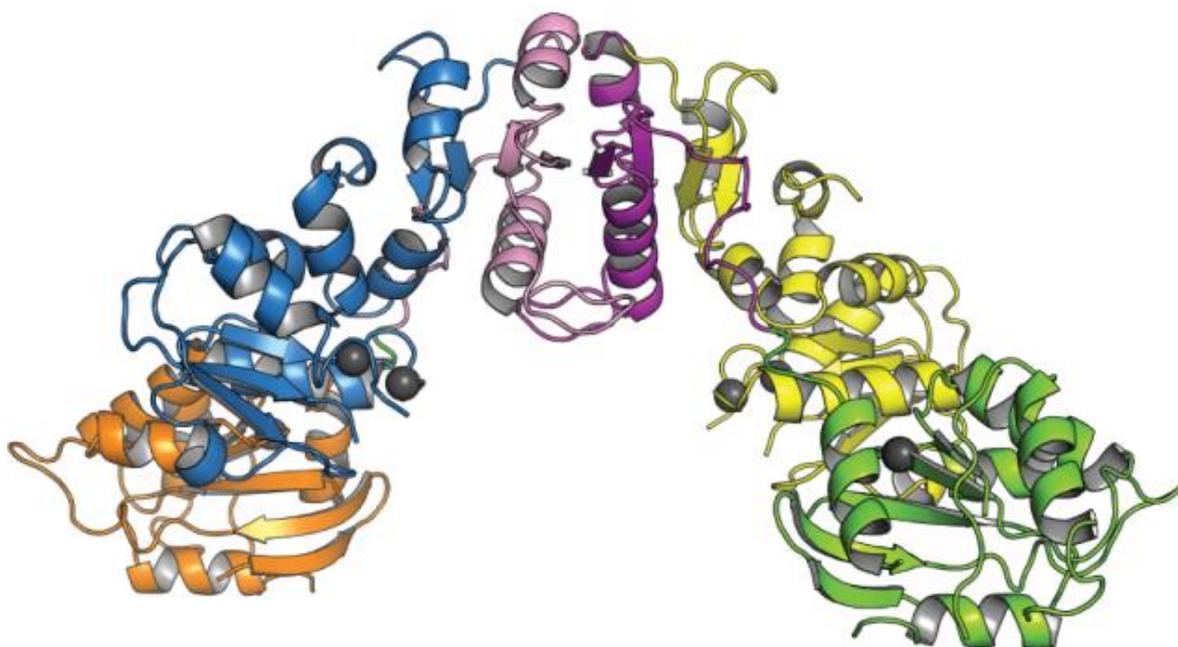


Figure 46: RX structure of apo-TmHydF. The dimer is represented with the GTPase domains of each chain colored in green and orange, the dimerization domains in magenta and light pink and the cluster binding domain in yellow and blue.

The [FeFe]-hydrogenase maturase HydF

As expected from the very high sequence identity (97%), apo-TmHydF structure is very close to the published structure of apo-TnHydF⁹ with an overall r.m.s.d. of 1.26 Å (359 C α) which decreases to 0.94 Å (216 C α) when superimposing the dimerization and [4Fe–4S] binding domains only, thus excluding the GTP-binding domain. The asymmetric unit contains one molecule and the crystal packing consists of dimers (Figure 47a). A disulfide bond between C353 and C356 was assumed due to the low resolution of the data. C302 was not observed. This is in contrast with apo-TnHydF for which a supramolecular assembly consisting of a dimer of dimers formed via a disulfide bond between C302 of different subunits was observed in the crystal, resulting in a different crystal packing (Figure 47c).

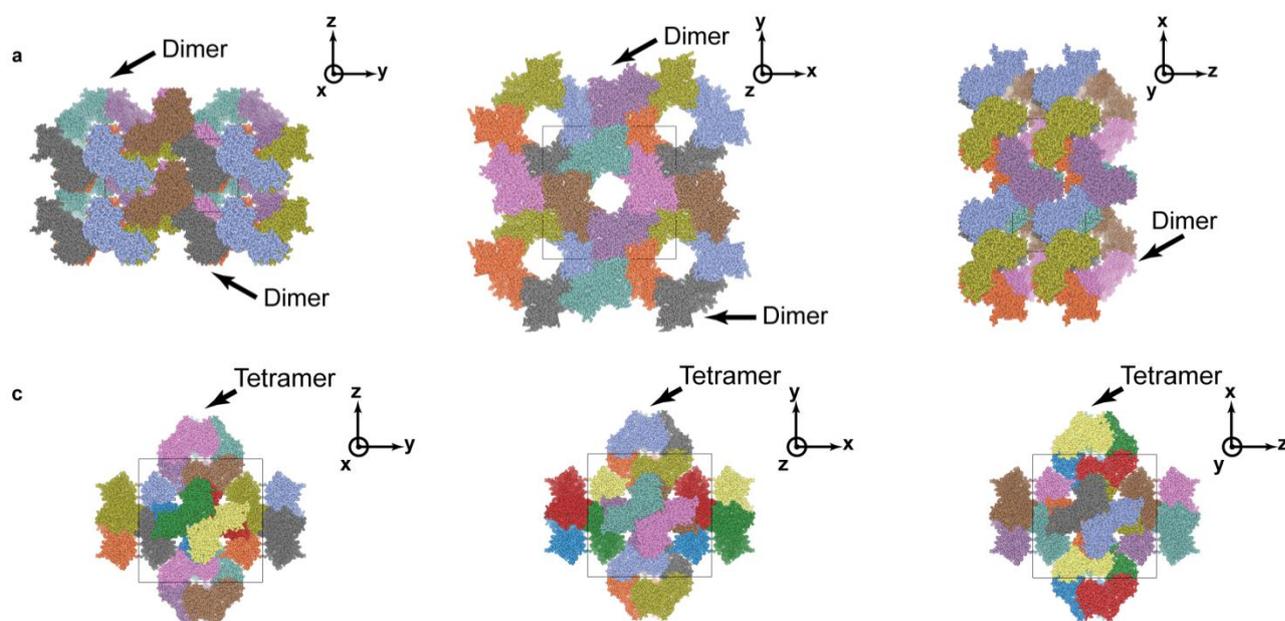


Figure 47: Representation of the crystal packing of apo-TmHydF (a) and apo-TnHydF (c). Each view corresponds to a face of the unit cell. **a)** Crystal packing of apo-TmHydF. There is only one molecule in the asymmetric unit. Subunits in violet/cyan, pink/brown, orange/blue and grey/green represent one dimer respectively. **c)** Crystal packing of apo-TnHydF. One tetramer in each orientation is indicated.

During the purification of apo-TmHydF, a mixture of dimer and tetramer could be observed by size exclusion chromatography with a tetramer/dimer ratio of \approx 1:10. However, when the tetramer was removed, a stable dimer was obtained when kept under anaerobiosis or in the presence of a reducing agent, even after concentration. As a consequence, the tetrameric form observed for apo-TnHydF might be a purification and crystallization artifact. In apo-TmHydF structure no electron density could be observed for residues 302-307 indicating that this region is likely disordered in the crystal while the corresponding residues (303-310) in

The [FeFe]-hydrogenase maturase HydF

Commercially supplied codon optimized synthetic plasmids of *Thermosipho melanesiensis* HydF (pet22b-TmeHydF, ampicillin-resistant) and *Clostridium thermocellum* HydF (pet22b-CtHydF, ampicillin-resistant) were obtained from Genscript.

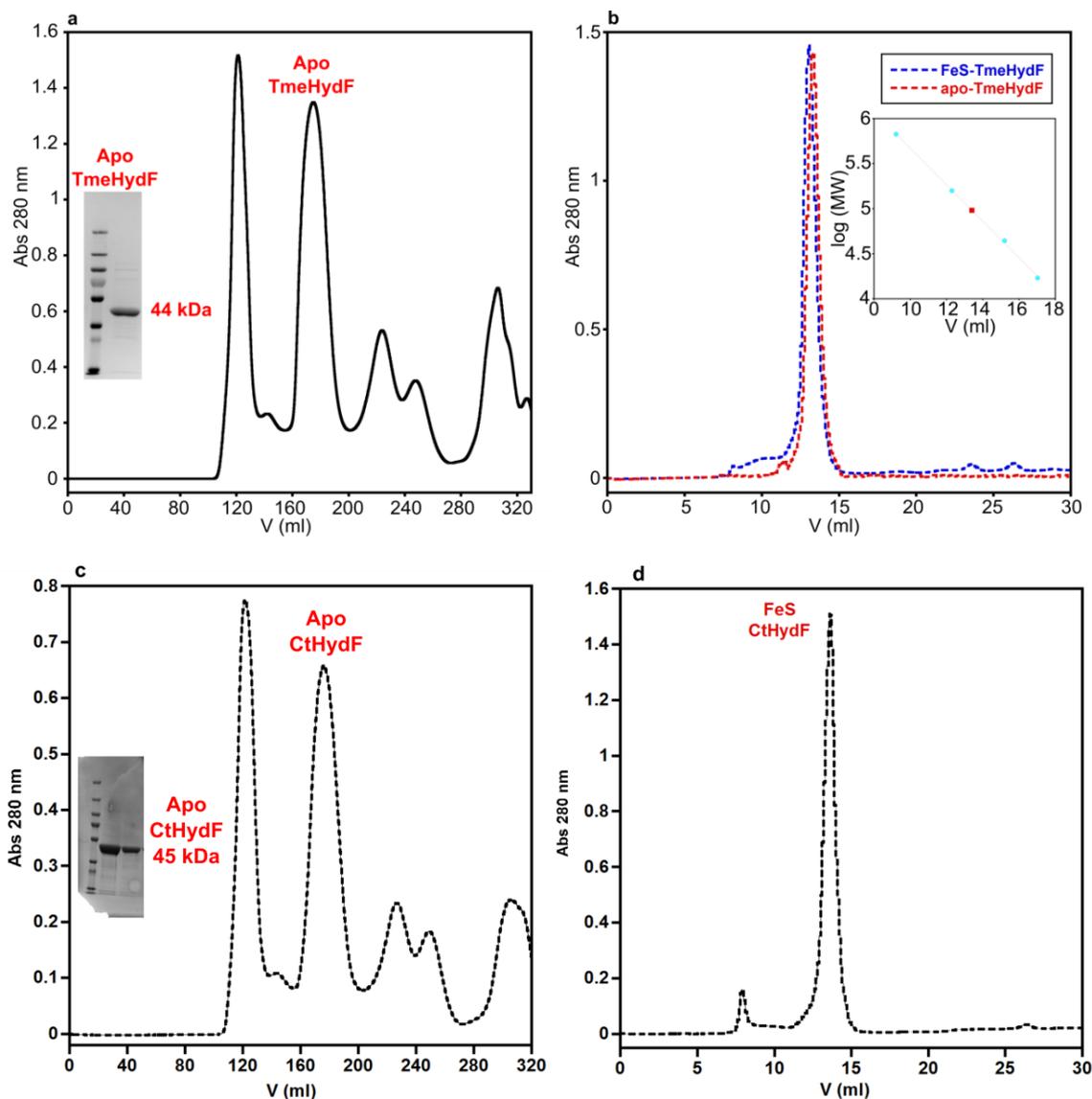


Figure 49: a,c) SEC elution profiles of apo-TmeHydF and apo-CtHydF on Superdex S200 26/600 in 25 mM Tris-HCl Buffer pH 8.0, 200 mM NaCl and SDS electrophoresis gels of purified proteins. b) Superimposition of anaerobic SEC profiles of apo- and FeS-TmeHydF on a Superdex S 200 10/300 and calibration curve using Thyroglobulin (670 kDa, 9.22 mL), γ -globulin (158 kDa, 12.32 mL), Ovalbumin (44 kDa, 15.2 mL) and Myoglobin (17 kDa, 17.03 mL). The molecular weight derived for apo-TmeHydF is 88.6 kDa, confirming apo-TmeHydF is a dimer. d) Anaerobic SEC profiles of FeS-CtHydF in the same conditions as c. Only for TmeHydF (b) protein is shown the superimposition of apo- and iron-sulfur forms since for CtHydF was found exactly the same behavior.

Recombinant expression of TmeHydF and CtHydF in *E. coli* resulted in a large production of soluble proteins. As both proteins originate from thermophilic organisms, they

could be purified to homogeneity via three aerobic purification steps (Figure 49): (i) a 50 °C heating step for TmeHydF and 60 °C for CtHydF protein; (ii) precipitation with 30% w/w of ammonium sulfate; (iii) gel filtration chromatographic column. Overall, 15-20 mg of pure proteins, lacking the cluster (< 0.1 Fe per protein), could be obtained from 1 L of culture and were called apo-TmeHydF and apo-CtHydF. Detailed protocols of proteins purification are provided in materials and methods (Chapter V, paragraph 5.3.2).

Once purified the apo proteins, the next step was the reconstitution of the iron-sulfur cluster. Anaerobic incubation of the proteins with an excess of iron ammonium sulfate and L-cysteine in the presence of dithiothreitol (DTT) and catalytic amounts of CsdA, a cysteine desulfurase, led to the incorporation of one [4Fe-4S] cluster per polypeptide chain.

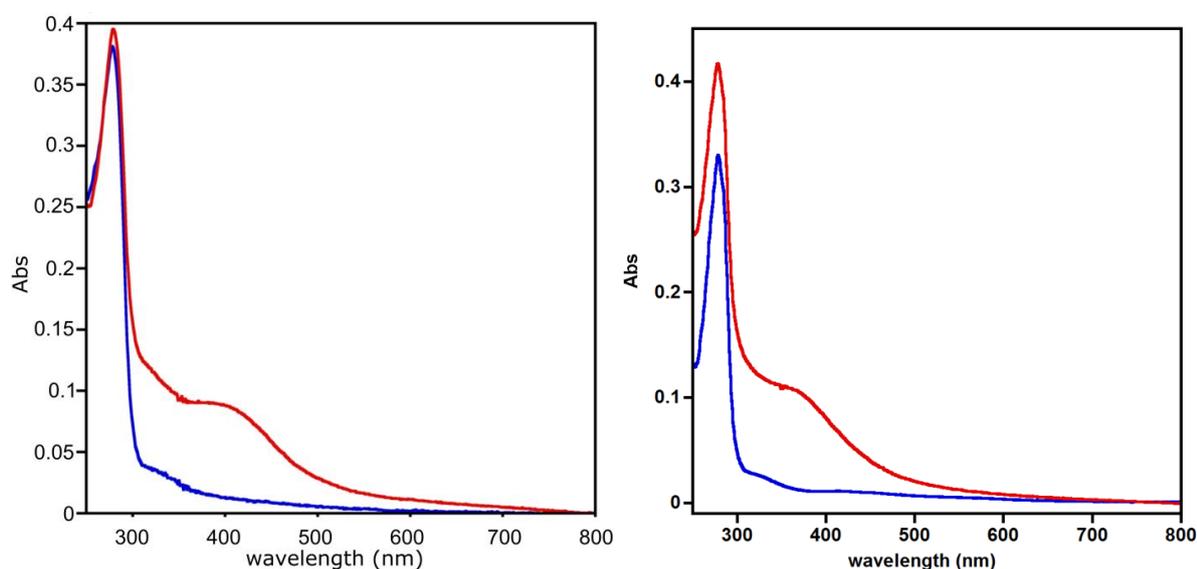


Figure 50: UV-Visible absorption spectra of 100 μ M apo (blue line, left) and FeS-TmeHydF (red line, left), apo (blue line, right) and FeS-CtHydF (red line, right). The spectra of apo forms was recorded in 25 mM Hepes buffer pH 7.5, 200 mM NaCl while the [4Fe-4S] species contained also 5 mM DTT.

This was monitored by UV-visible spectroscopy, focusing on the [4Fe-4S] cluster characteristic absorption band centered at 405-410 nm (Figure 50), and iron (3.5 ± 0.1 per monomer) and sulfide (3.7 ± 0.2 per monomer) quantitation after a gel filtration step. The reconstituted proteins, FeS-CtHydF and FeS-TmeHydF, behaved as a dimer in solution and the incorporation of the cluster did not alter the overall shape of the protein as confirmed by identical elution volumes of apo and iron-sulfur reconstituted forms (Figure 49b). Fresh preparations of both proteins containing the [4Fe-4S] cluster were used for crystallization trials automated with an Oryx Nano crystallization robot (Douglas Instruments) installed in a temperature and hygrometry regulated anaerobic chamber ($O_2 < 1.5$ ppm). Hits were obtained

The [FeFe]-hydrogenase maturase HydF

only for FeS-TmeHydF (Figure 51), thus we pursued the study of this new HydF protein combining the crystallographic approach to the spectroscopic study of the protein.

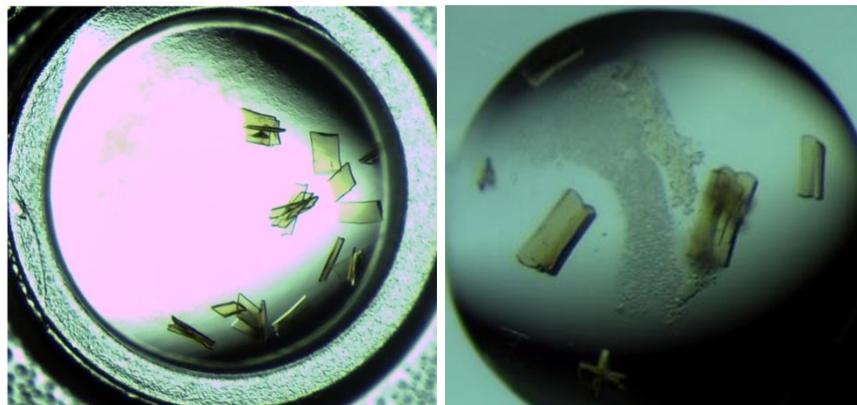


Figure 51: Pictures of FeS-TmeHydF crystals. The crystals on the left were reproduced manually (hanging drop) while the crystals on the right are produced by the Oryx Nano crystallization robot (sitting drop).

FeS-TmeHydF was characterized using EPR, HYSCORE and FTIR spectroscopies, in collaboration with the group of Prof. W. Lubitz where I spent a couple of weeks at the Max Plank Institute in Mülheim (Germany). In particular, EPR is the best spectroscopic technique that allows confirming the presence of a [4Fe-4S] cluster; indeed, it can easily distinguish between different iron-sulfur centers and redox states. FeS-TmeHydF was analyzed through X-band EPR spectroscopy (Figure 52) and these experiments were realized with the help of Dr. Edward Reijerse and Dr Agnieszka Adamska-Venkatesh.

FeS-TmeHydF proved to be EPR silent in the as-purified form, in agreement with a $S=0$ [4Fe-4S]²⁺ state, while it became EPR active after reduction with sodium dithionite exhibiting signals with $g= (2.046, 1.90, 1.86)$ characteristic for a reduced $S=1/2$ [4Fe-4S]⁺ cluster (Figure 52). These data thus exclude the presence of other kinds of [Fe-S] clusters as instead raised recently by Broderick JB and coworkers.^{17,18}

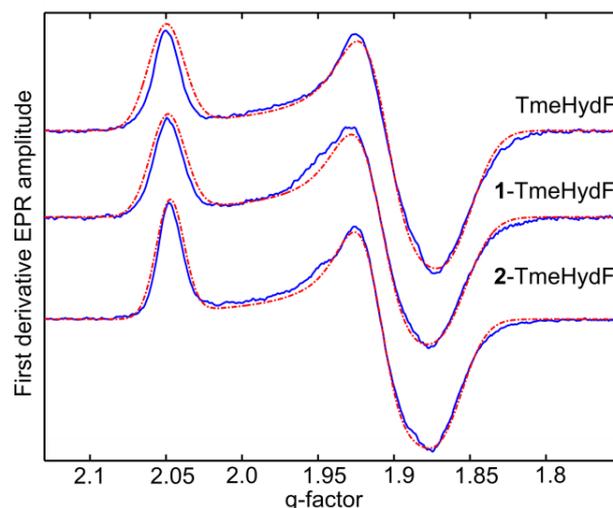


Figure 52: X-band (9.5 GHz) pseudo-modulated 2-pulse echo detected EPR spectra of reduced FeS-TmeHydF and hybrid proteins 1-TmeHydF and 2-TmeHydF at 10K. Experimental conditions: $T = 10\text{K}$, $\tau = 200\text{ns}$, $\pi/2$ pulse = 16ns. The spectral simulations (dotted red lines) were performed using the parameters listed in Table 7. A description of samples preparation has been provided in materials and methods.

In their studies, they reported on the existence of redox active [4Fe–4S] and [2Fe–2S] clusters bound to HydF from *C. acetobutylicum* although the investigated protein preparation contained from 2.24 ± 0.49 Fe/dimer up to 4.8 ± 0.8 Fe/dimer. The amount of iron is far from the expected 12.0 Fe atoms for a dimeric HydF (containing one [4Fe–4S] and one [2Fe–2S] cluster per monomer). Thus, the studied protein was too much under-metallated to provide a clear evidence of this second [2Fe–2S] cluster.¹⁷

Our study, on the other hand, confirms what was already reported for other HydF proteins, *i.e* the existence of a single [4Fe–4S] cluster when either i) HydF is expressed and anaerobically purified in a background devoid of HydG and HydE or ii) it is expressed, aerobically purified and its [4Fe–4S] cluster reconstituted chemically.

4.2.1 X-ray structure of TmeHydF: a [4Fe–4S] with an unexpected and exchangeable ligand

As shown in the last paragraph, only FeS-TmeHydF crystallized during screening conditions. Indeed, FeS-TmeHydF concentrated to 13 mg/ml ($\approx 290 \mu\text{M}$) in Hepes 25 mM pH 7.5, DTT 5 mM, NaCl 200 mM crystallized in 3 weeks using the screening kit Peg Rx1 (Hampton research) and conditions were then manually optimized (Figure 51). Final

The [FeFe]-hydrogenase maturase HydF

crystallization conditions were Hepes 25 mM pH 7.5, PEG 8K 1.5%. Crystals were briefly soaked in a cryoprotectant solution consisting of mother liquor components supplemented with 30% v/v glycerol before flash freezing in liquid propane in anaerobic conditions. The flash freezing anaerobic condition consisted in the preparation of a small metallic cylindrical chamber (built to harbor a loop containing a single crystal) filled with liquid propane, which is frozen in liquid nitrogen. The next step was the transfer of this device inside the glove box where in two minutes solid propane liquefies allowing the flash freezing of protein crystals. Once transferred, the crystals are stored in liquid nitrogen and analyzed on the Synchrotron lines (Soleil synchrotron in Saint-Aubin, PX1 and PX2 lines).

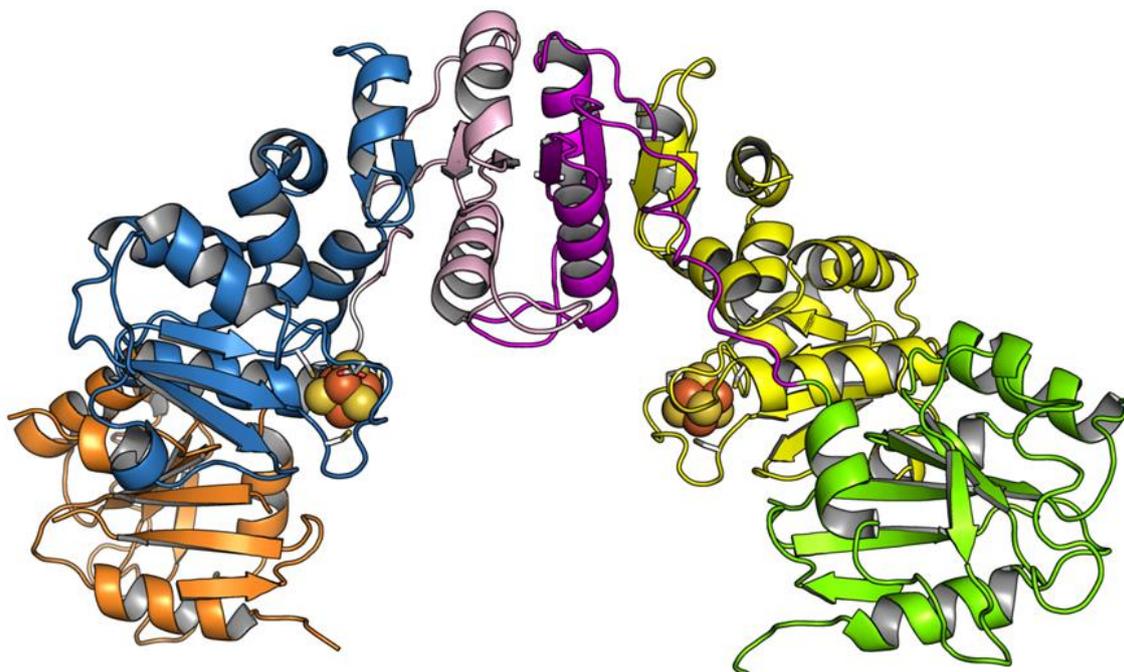


Figure 53: X-ray structure of iron-sulfur reconstituted HydF from *Thermosiphon melanesiensis*. Cartoon representation of FeS-TmeHydF. The dimer is represented with the GTP domains of each chain colored in green and orange, the dimerization domains in magenta and light pink and the cluster binding domain in yellow and blue. The [4Fe-4S] clusters are represented as spheres with the protein ligands represented as sticks.

FeS-TmeHydF crystallized in the monoclinic space group $P2_1$ and about 200 crystals were screened in order to solve the structure. The structure of FeS-TmeHydF was refined to 2.8 Å resolution (Figure 53). The asymmetric unit contains four monomers and the crystal is formed by the packing of dimers as apo-TmHydF. As previously described for the structure of *T. neopolitana* apo-HydF, the monomer is composed of three domains. Residues 7-166 belong to the GTP-binding domain, residues 172-262 to the dimerization domain and residues 263-395 to the cluster-binding domain. The GTP-binding domain is linked to the dimerization

domain via a stretch of about 5 residues which likely provides flexibility. The folding of GTPase domain, like in apo-TnHydF and apo-TmHydF, is composed of five parallel and one anti-parallel β -strands with five α -helices delineating this sheet (Figure 54, magenta). The B factors of this domain are higher than for the other two domains. An explanation might reside in the low number of crystal contacts originating from the GTP-binding domain.

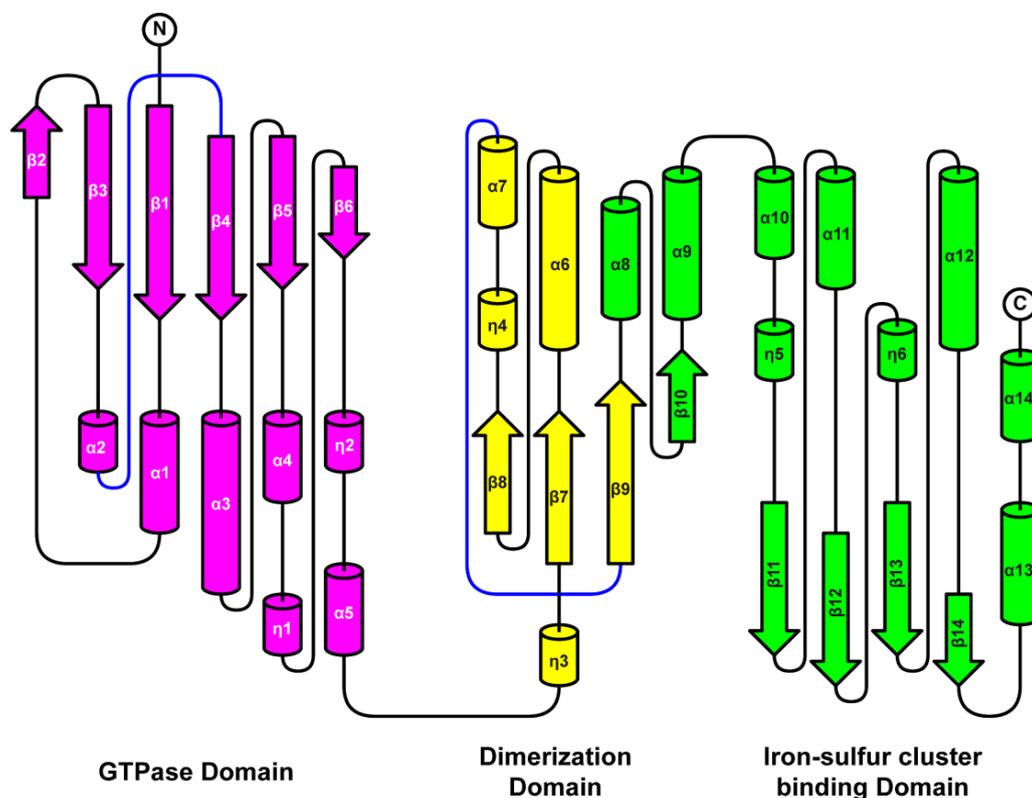


Figure 54: Topology diagram of FeS-TmeHydF. Structural elements of the GTP binding, dimerization and iron-sulfur cluster binding domains are colored in magenta/yellow and green respectively. Figure was generated with TOPDRAW.¹⁹

The dimerization domain is composed of three α -helices and four-stranded parallel β -sheet. $\beta 22$ forms an antiparallel β -sheet with its counterpart in the second subunit. Additional contacts are made between $\alpha 16$ with the loop connecting $\beta 15$ to $\alpha 16$ of the other subunit. The cluster-binding domain is composed of a sheet of 4 parallel β -strands and 5 α -helices. The presence of the [4Fe-4S] cluster allowed the stabilization of the loop (residues 299-306), which provided coordination to the cluster. In the structure of apo-TmHydF these residues are not observed (Figure 46) while in apo-TnHydF it is stabilized by inter-dimer contacts in the tetramer (Figure 7).⁹

EPR spectroscopy clearly showed that the protein contained a single [4Fe-4S], and its location in the protein structure was confirmed by anomalous signals at the iron edge within

The [FeFe]-hydrogenase maturase HydF

data set. Single-wavelength anomalous diffraction (SAD) is a very effective method that relies entirely on the measurement of the anomalous diffraction produced by one or more scattering atoms in the crystal. Data sets were recorded from the same crystal at the wavelength around the X-ray absorption edge of Iron (7.15 keV). This procedure allowed confirming the position of the [4Fe–4S] cluster inside the protein structure.

Thus, the protein contains only one [4Fe–4S] cluster and the three-dimensional structure does not show any obvious site for adjacent binding of a second, [2Fe–2S] cluster, as proposed by JB. Broderick and collaborators.^{17,18} In particular no additional, even tiny, anomalous signal for iron could be observed using a dataset collected at the iron edge. Observation of spectroscopic signatures for a [2Fe–2S] cluster has been reported only in the case of *C. acetobutylicum*. However these data were obtained with incompletely cluster-loaded and certainly partially degraded forms of HydF, since in one case the studied protein contained 0.85-0.98 Fe per protein while in a more recent case it contained 2.24 Fe per dimer, thus 1.12 Fe per protein.^{17,18}

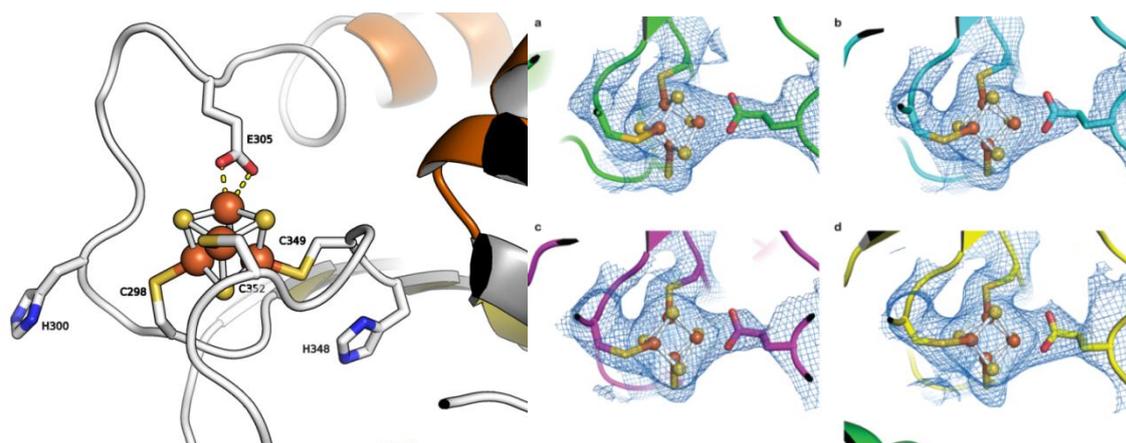


Figure 55: Representation of the cluster binding site (left), in which 3 conserved cysteines and a glutamate chelate the [4Fe–4S] cluster. The two conserved Histidine residues of the CxH_xnHCxxC motif are also shown. Iron-sulfur cluster binding sites in with the 2mF₀-DF_{calc} electron density map contoured at the 1σ level for the chain **a**, **b**, **c**, **d** respectively (right).

The [4Fe–4S] cluster of FeS-TmeHydF is coordinated by three cysteines (C298, C349, C352) (Figure 55), which are strictly conserved among the members of the HydF family. An alignment conducted on 227 bacterial HydFs showed the complete conservation of three cysteine residues located in the C-terminal region of the protein, within a CxH_xnHCxxC motif, which are confirmed now to chelate the unique [4Fe–4S] cluster.

The ligand for the fourth Fe atom of the cluster in HydF proteins has been a matter of controversy. Some studies reported the possibility of a histidine 4th coordination

(C_xH_{x_n}HC_{xx}C motif) to the iron-sulfur site. This is the case of *Clostridium acetobutylicum* HydF (CaHydF) where HYSCORE spectroscopy showed that His352 was a ligand of the [4Fe–4S] cluster.²⁰ A His to Ala mutation removed this coordination, however, the H352A mutant could still assemble a [4Fe–4S] cluster. All these observations suggest a weakly bound ligand, probably required for functional assembly of the transient [2Fe]-subcluster precursor in HydF. Indeed, it has been shown by spectroscopic studies that the fourth coordination site of the cluster is readily exchangeable. The cluster can bind exogenous imidazole,^{16,20} as in the case of HydFs from *Thermotoga maritima* and *Thermotoga neapolitana*, a histidine from the N-terminal His-tag of a tagged preparation of HydF,¹⁶ or a synthetic diiron complex via a cyanide bridge both cases shown for TmHydF protein.¹² This has led, in the case of HydF from *T. neapolitana*, to the proposal that the exchangeable ligand is a solvent molecule (H₂O or OH⁻), coupled to other water molecules, as suggested by ESEEM and HYSCORE spectroscopy.^{21,22} HYSCORE spectroscopy and site-directed mutagenesis also allowed excluding unambiguously a histidine coordination from the C_xH_{x_n}HC_{xx}C motif in HydF from *T. maritima* and *T. neapolitana*, as instead observed in CaHydF.^{16,22}

Finally the X-ray structure of FeS-TmeHydF shows that the only iron not coordinated by a cysteine is coordinated by the carboxylate group of a glutamate, E305 (Figure 55). However, whether glutamate coordination is mono or bidentate cannot be inferred from the structure due to the resolution of the data.

This position is not strictly conserved when analyzing the alignment of 227 bacterial HydFs but is an acidic residue (Asp or Glu) in 98% of the cases as deduced from our alignment (Figure 56b).

Carboxylate residues (aspartate or glutamate) are rarely used as ligands of iron-sulfur clusters. An aspartate is found in the [4Fe–4S] cluster of IscA, a cluster carrier protein involved in iron-sulfur cluster assembly.²⁴ Interestingly, in IscA the cluster is only transiently attached to the protein. An aspartate is also a ligand of the [4Fe–4S] cluster in ferredoxin from *Pyrococcus furiosus* as well as in the light-independent protochlorophyllide reductase.^{25,26} To our knowledge, the only example of a [4Fe–4S] protein with three Cys and a Glu ligand is the IspG protein from *Aquifex aeolicus* involved in the biosynthesis of the universal isoprenoid precursors (isopentenylpyrophosphate and dimethylallylpyrophosphate).²⁷ In IspG the glutamate protects the fourth iron in the absence of the substrate and is exchanged when the enzyme is involved in catalysis. A similar function for the glutamate of TmeHydF could be envisaged and the experiments described in the following pages go towards this direction.

The [FeFe]-hydrogenase maturase HydF

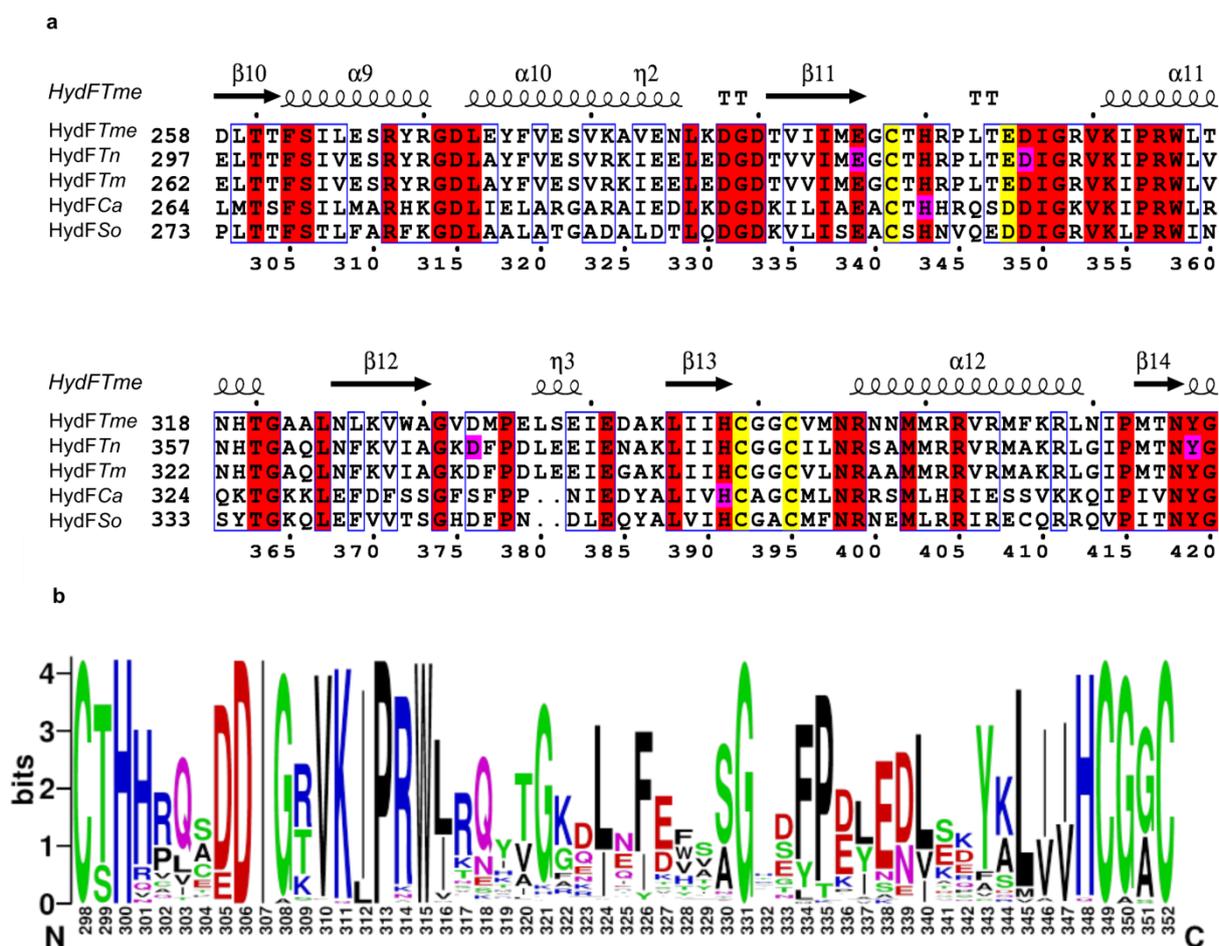


Figure 56: Alignment of the cluster binding region of the five HydFs studied in the literature with ESPrpt3⁴ (a) and logo generated from the alignment of 227 bacterial HydFs (b).

a) The residues chelating the [4Fe–4S] cluster are highlighted in yellow, conserved residues are highlighted in red, residues with similar features are outlined in blue boxes and residues highlighted in pink for *Clostridium acetobutylicum* and *Thermotoga neapolitana* HydF are those postulated as ligands of the [4Fe–4S] cluster and mutated in alanine with no effect on *in vivo* [4Fe4S] cluster assembly.²⁰ Mutations of either conserved histidine (pink) in HydF from *C. acetobutylicum* resulted in severe impairment of HydA maturation suggesting that these histidines, which are structurally close to the cluster (Figure 53), play an important role in the chemical processes associated either in the assembly of the diiron precursor in HydF or its dissociation for transfer to HydA.²⁰

b) Numbering on the x axis corresponds to the residue number in the sequence of TmeHydF. Figure was generated with WebLogo.²³ Residues of TmeHydF involved in the coordination of the FeS cluster are C298 / E305 / C349 and C352.

It is intriguing that the three best characterized HydFs would have three different cluster coordinations. However, the OH₂/OH⁻ ligand in *T. neapolitana* HydF requests, as discussed in the introduction, further confirmation.²¹ Our crystal structure furthermore shows that His348 in TmeHydF (equivalent to His352 in *C. acetobutylicum* HydF) is close to the cluster but not appropriately oriented and too much buried in the interior of the protein to

interact with the fourth Fe atom. Considering that in sequence alignment of 227 bacterial members of the HydF family, the equivalent position of E305, while not strictly conserved, is in $\approx 98\%$ of the sequences an acidic residue, it is likely that HydF cluster coordination is in most cases made of 3 cysteines and one glutamate or aspartate, as in TmeHydF.

Furthermore the structure shows that the cluster is exposed to the solvent. These features are in line with the presumed function of HydF cluster, which has to transiently assemble a diiron complex from precursors provided by HydG and HydE, thus requesting a flexible and labile coordination site, and then transfer it to HydA.

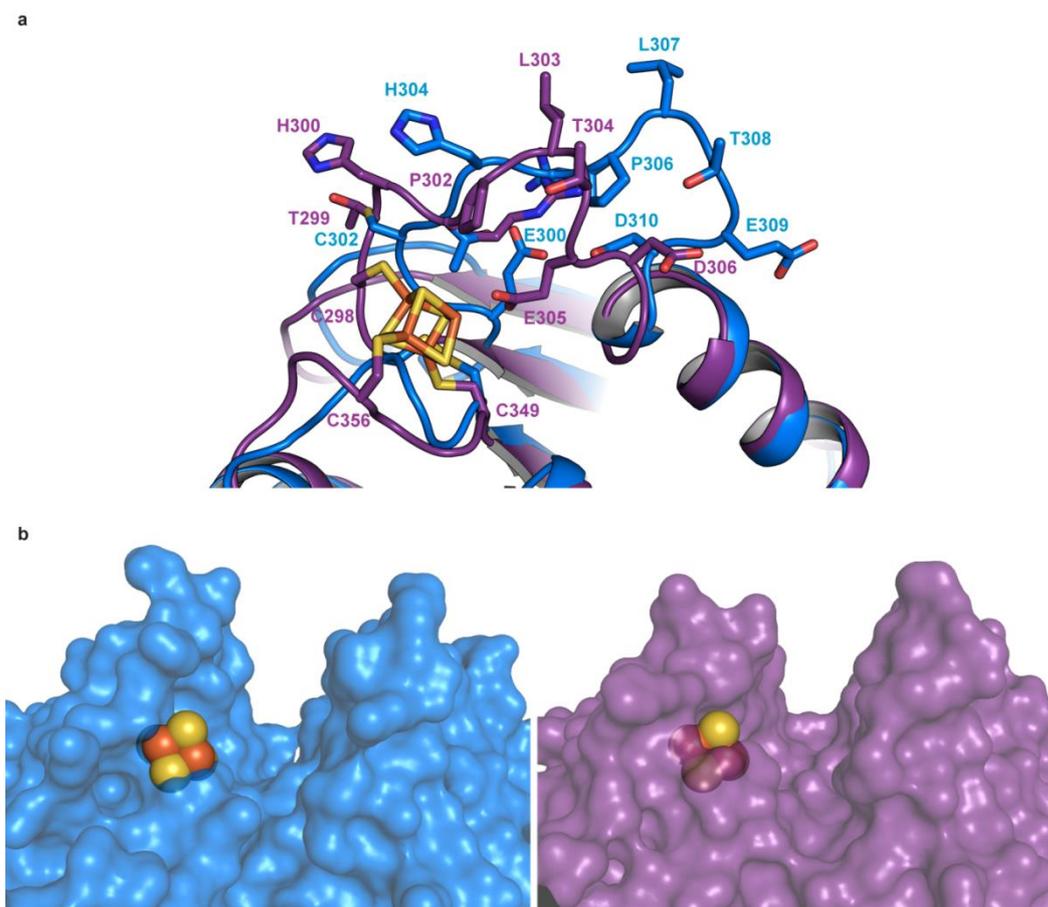


Figure 57: Potential conformational changes around the iron-sulfur cluster binding site. **a)** Comparison of the conformation of the loop around the $[4\text{Fe}-4\text{S}]$ cluster in apo-TnHydF (blue) and FeS-TmeHydF (purple). Only the residues of the loop are displayed as sticks. **b)** Surface representation around the $[4\text{Fe}-4\text{S}]$ cluster using the same color scheme as in **a)**. In apo-TnHydF (left) the iron sulfur cluster has been added via pasting the $[4\text{Fe}-4\text{S}]$ cluster of FeS-TmeHydF and is displayed only as a guide to the eye. The conformation of the loop connecting Xn and Yn observed in apo-TnHydF changes the shape of the cavity near the $[4\text{Fe}-4\text{S}]$ cluster binding region. This loop was not observed in apo-TnHydF which is indicative of conformational heterogeneity in this region.

The X-ray structure shows also that the $[4\text{Fe}-4\text{S}]$ cluster is adjacent to a cavity which can accommodate the diiron complex (Figure 57). However, with the glutamate E305

The [FeFe]-hydrogenase maturase HydF

coordinated to the cluster and because of the protection provided by the loop (residues 299-306) containing E305, obviously the diiron complex cannot easily get access to the cluster. It is thus likely that transient binding of the diiron complex requires a significant conformational change of that loop which not only de-coordinates the glutamate residue, but also increases the access to the cluster.

The structure of apo-TnHydF provides a view of such a possible open conformation. Indeed, the structural superimposition of the loop from FeS-TmeHydF and apo-TnHydF shows that in the second structure the glutamate (E309) is turned inside out. The surface representations show that if the loop would adopt the conformation as in apo-TnHydF (Figure 57, left) the cluster would be much more accessible than in FeS-TmeHydF (Figure 57, right).

In this context it is significant that the region of the protein harboring the iron-sulfur cluster is characterized by the presence of an area of positively charged residues (Figure 58). This might serve as a specific binding site for protein partners (HydA, HydG, HydE) and interestingly, it also fits with the need for the protein to bind a negatively charged diiron complex, as shown for TmHydF, chemically matured with biomimetic complexes, and CaHydF, purified in a background of HydE and HydG naturally synthesizing the binuclear compound.

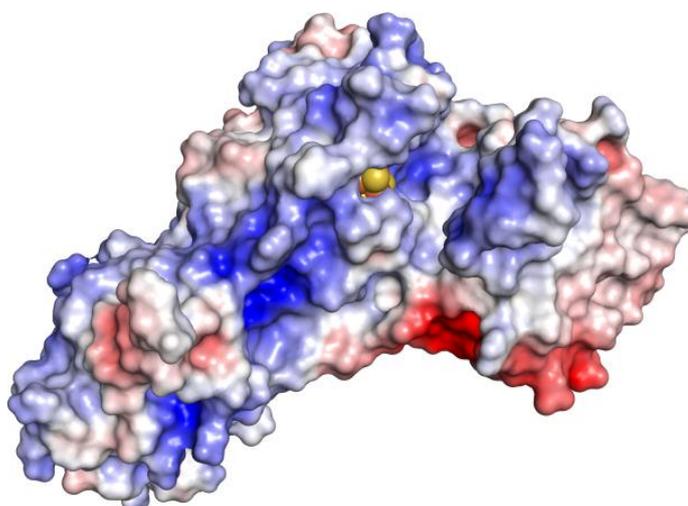


Figure 58: Electrostatic potential surface calculated with APBS²⁸ at ± 5 kT/e for one monomer of FeS-TmeHydF. The FeS cluster is shown as spheres (Fe: orange, S: yellow). Negative and positive potentials are displayed in red and blue respectively.

The presence of a glutamate is unexpected and intriguing. To confirm the exchangeability of this coordination site, we checked whether exogenous ligands could bind to the cluster, in spite of the presence of the glutamate ligand. In previous studies, it has been

shown that imidazole can coordinate the iron-sulfur cluster of HydF proteins from *Thermotoga maritima* and *Thermotoga neapolitana*. Indeed, upon incubation of proteins with an imidazole solution, a new coordination to the iron-sulfur cluster is observed by HYSCORE spectroscopy. HYSCORE, a two-dimensional pulsed EPR technique, allows detection of hyperfine coupling of nuclei with a low gyromagnetic moment. In the case of a strong coupling between an $S = 1/2$ cluster ($[4\text{Fe}-4\text{S}]^+$) with an $I = 1$ nuclear spin nucleus such as ^{14}N , it is possible to observe a characteristic pattern in the $(+, -)$ quadrant with so-called double quantum-double quantum correlation features. Such features have been observed in only a few cases, such as proteins in which histidine is ligated to a $[2\text{Fe}-2\text{S}]$ cluster,²⁹ some Radical SAM proteins^{30,31} and in the specific case of HydF proteins (TmHydF and TnHydF) treated with imidazole.¹⁶

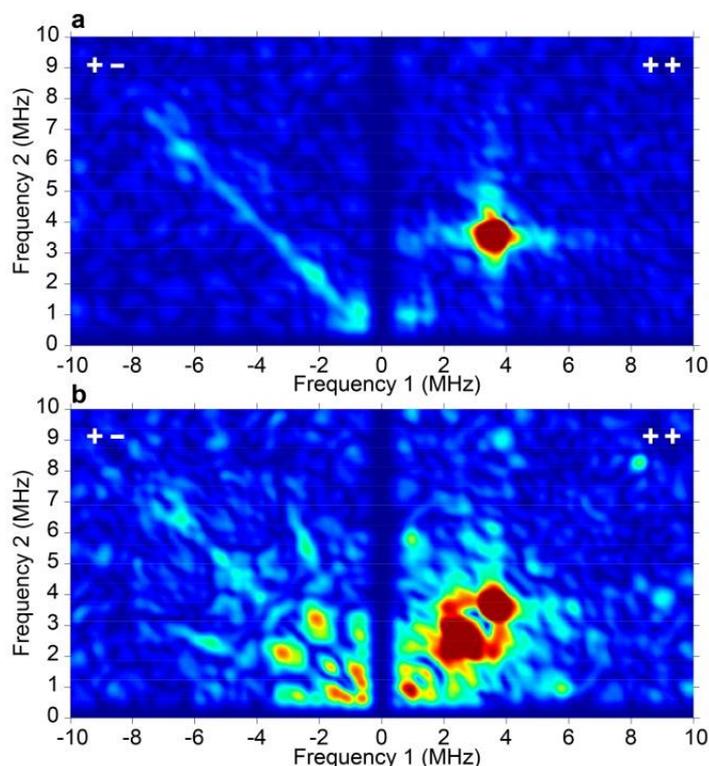


Figure 59: X-band (9.4 GHz) HYSCORE spectra of (a) reduced TmeHydF and (b) TmeHydF + imidazole. Experimental conditions: 10K, $\tau=136$ ns, magnetic field was set to the maximum echo amplitude at 352.5 mT (near g_2). See paragraph 5.9.3 for details.

The HYSCORE spectrum of the reduced TmeHydF presented no cross-coupling peaks except those corresponding to distant ^{13}C atoms present in natural abundance and ^1H atoms (not shown), thus excluding a nitrogen atom in the coordination sphere of the cluster. Indeed, the protein structure showed the glutamate E305 as 4th ligand of the $[4\text{Fe}-4\text{S}]$. On the other hand, HYSCORE spectrum of reduced TmeHydF treated with imidazole was drastically different.

The [FeFe]-hydrogenase maturase HydF

The latter contained a pattern of peaks in the (+, -) quadrant characteristic of a coupled nitrogen atom (Figure 59). Double quantum features occur in the (+, -) quadrant at frequency (5.5, 2.4) MHz indicative of a ^{14}N hyperfine interaction of 3.1 MHz in combination with a quadrupole parameter K between 1.35 and 1.55 MHz. As these features are absent in the HYSCORE spectrum of FeS-TmeHydF in the absence of imidazole, this shows that imidazole can displace the labile glutamate and bind to the cluster.

4.2.2 Preparation and characterization of 1- and 2-TmeHydF hybrids

The combined structural and spectroscopic study of FeS-TmeHydF proved the existence of a single [4Fe-4S] cluster inside the protein scaffold with an unusual and exchangeable coordination site. The next step towards the realization of an “artificial hydrogenase”, as in the case of TmHydF (paragraph 4.1.2), is the binding of $[\text{Fe}_2(\text{xdt})(\text{CO})_4(\text{CN})_2]^{2-}$ biomimetic complexes which should displace the glutamate ligation, as shown for the imidazole ligation.

As previously shown for HydF from *Thermotoga maritima*,¹² FeS-TmeHydF was reacted with an excess of the synthetically prepared complex **1** or **2**, for one hour anaerobically, and then purified by desalting for removing the excess of chemicals.

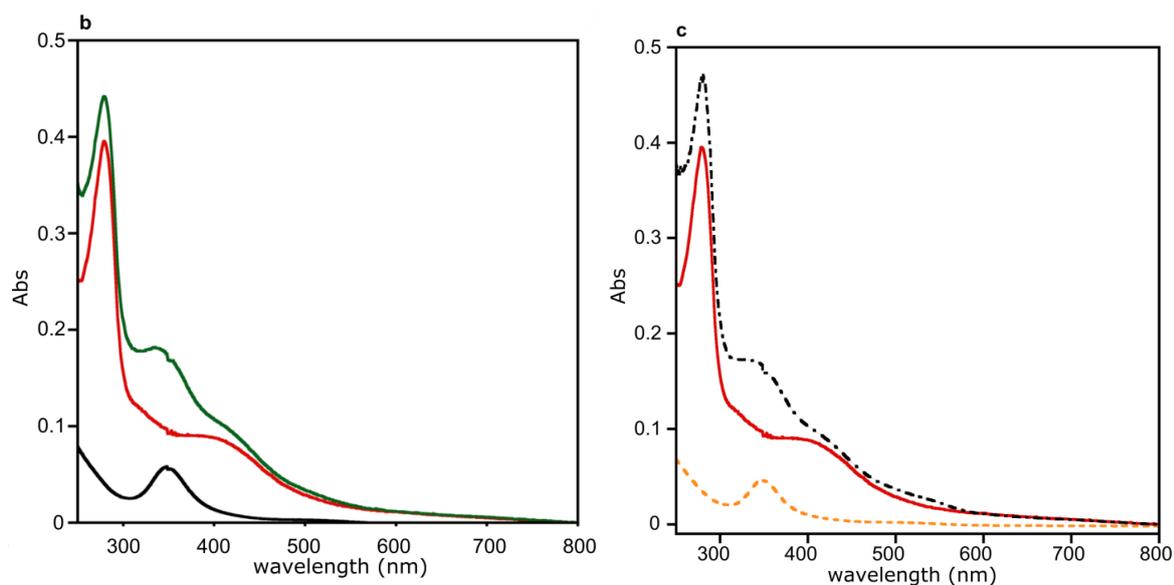


Figure 60: Superimposition of UV-Visible absorption spectra of 100 μM 2-FeS-TmeHydF, FeS-TmeHydF and **2** (b), and 100 μM 1-FeS-TmeHydF, FeS-TmeHydF and **1** (c). All the spectra were recorded in 25 mM Hepes

buffer pH 7.5, 200 mM NaCl while the FeS-TmeHydF sample contained also 5 mM DTT. Color code: FeS-TmeHydF (red), 2-FeS-TmeHydF (green), **2** (black), 1-FeS-TmeHydF (black dashed), **1** (dashed orange).

Evidence for incorporation of the complexes into the protein was first obtained by Fe quantitation, which showed around 60-80 % incorporation of both synthetic complexes. Indeed, 2-TmeHydF contained 4.9 ± 0.1 Fe while 1-TmeHydF has 4.6 ± 0.1 Fe per monomer, as compared to 3.5 ± 0.1 Fe per monomer in the case of FeS-TmeHydF. Furthermore, the UV-visible spectra of the hybrid proteins displayed a new absorption band at 350 nm characteristic of the protein-bound diiron complex (Figure 60). This new absorption band is originated by a LMCT (Ligand to Metal Charge Transfer) transition due to the electron transfer from CO and CN ligands to the iron of the binuclear cluster.

In order to confirm the specific binding of both synthetic complexes inside the protein scaffold, notably via the [4Fe-4S], EPR and FTIR spectroscopy were applied. The EPR spectrum of the reduced 1-TmeHydF and 2-TmeHydF showed a very small upshift in the g_x and g_y values indicative of a possible interaction with the added diiron complex (Figure 52). The final confirmation was given by Fourier transform infrared spectra of both HydF hybrids. CO and CN ligands give rise to FTIR stretching bands in a zone free of other protein bands. CN-stretching bands between 2,000 and 2,100 cm^{-1} and CO-stretching bands in the 1,800–2,000 cm^{-1} range resulted to be much narrower than those observed for the free complexes **1** and **2** in solution,¹² and indicated protein-bound complexes (Figure 61a).

FTIR spectra are reminiscent of those reported previously for the HydF hybrid from *T. maritima*, prepared similarly.¹² Reduction using sodium dithionite caused only slight shifts of the four terminal CO stretches indicating a Fe(I)Fe(I) valence state for the diiron complexes in the protein, as well as in the free compounds. Thus, the small shift and the better resolution of the FTIR bands of HydF hybrids in the presence of sodium dithionite has to be attributed to the reduction of the [4Fe-4S] cluster. The reduction increased slightly the wavenumbers of FTIR stretches proving, as expected, the close vicinity of the diiron biomimetic complex to the iron-sulfur cluster.

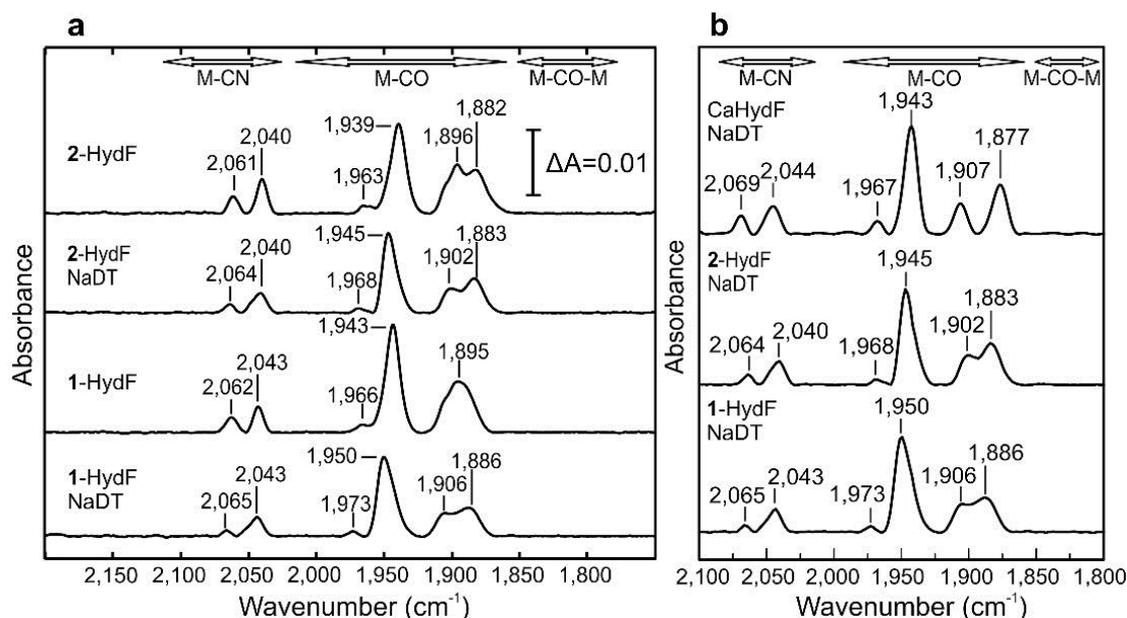


Figure 61: a) FTIR absorption spectra of x-TmeHydF (x=1 or x=2) produced by chemical maturation with and without sodium dithionite (NaDT) incubation. b) Superimposition of FTIR spectra of HydF from *C. acetobutylicum*¹¹ produced from the native host expressing all the clostridial maturases (HydEFG) and x-TmeHydF. In all spectra the FTIR bands in the range 1800-2000 cm⁻¹ are assigned to CO stretching bands whereas the bands above 2000 cm⁻¹ belong to CN stretching bands.

As extra proof for the correct insertion of the synthetic complex within the protein, 1-TmeHydF was shown to function efficiently as a maturase of FeS-HydA. After incubation with FeS-HydA from *Megasphaera elsdenii* (FeS-MeHydA), the enzyme was assayed for hydrogenase activity. Using a standard dithionite/methyl viologen assay³² and quantifying the amount of produced H₂ by gas chromatography, we showed that HydA was fully matured as it displayed a high specific hydrogenase activity comparable to that obtained after maturation of FeS-MeHydA directly with the synthetic complex **1** ($600 \pm 50 \mu\text{mol H}_2 \cdot \text{mg}^{-1} \cdot \text{min}^{-1}$).³³ Thus, it indicates correct transfer of the synthetic biomimetic complex from HydF scaffold to HydA protein, which proved completely matured.

Finally, it is important to note how strikingly similar the FTIR spectra of TmeHydF hybrids (Figure 61b) are to that of HydF from *C. acetobutylicum* purified from the native host, homologously expressing the clostridial maturation enzymes.¹¹ In addition, both proteins are able to fully activate apo-HydA. These observations strongly support the conclusion that complex **1** is the diiron unit naturally assembled in HydF (in agreement with HydG providing two [Fe(CN)(CO)₂] units) and the true precursor of the [2Fe]-subcluster in HydA, as shown in figure 11 (Chapter I, paragraph 1.5.6).

Obviously this requires final confirmation notably via structural characterization of naturally and chemically produced active HydFs. Soaking experiments of synthetic complexes inside FeS-TmeHydF crystals led protein crystals melting. Again, it points towards the change of the coordination of the [4Fe-4S] cluster where the glutamate is replaced by the synthetic diiron complex.

Crystallization trials of purified HydF hybrids have not produced hits yet.

4.2.3 Hydrogenase activity of **2**-TmeHydF hybrid protein

In the last paragraph, I proved the correct insertion of synthetic complexes **1** and **2** onto the protein scaffold HydF from the new organism *Thermosipho melanesiensis*. As for HydF from *Thermotoga maritima*, TmeHydF hybrids were assayed for hydrogenase activity. **1**-TmeHydF, with the biologically relevant adt^{2-} ligand, proved not active as the same hybrid from *Thermotoga maritima*. On the contrary, **2**-TmeHydF, matured with the $[\text{Fe}_2(\text{pdt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex, showed hydrogen evolution under reducing conditions. The activities of HydF hybrids required a miniaturized Clark-type hydrogen microsensor (Figure **62**), allowing direct detection of hydrogen produced in solution.

2-TmeHydF was shown to produce H_2 with an initial TOF (turnover frequency) of 0.3 min^{-1} , about 30 times larger than compound **2** free in solution. Such an activity is comparable to the same hybrid protein synthesized from TmHydF, reported in above (Paragraph 4.1.2).¹³ The reaction started immediately after the addition of dithionite (Figure **62b**) then levelled off after a few minutes, while the system was still active after 5 min reaction time. For comparison FeS-MeHydA (the iron-sulfur form of *Megasphaera elsdenii* HydA containing only the three [4Fe-4S] clusters and no [2Fe]-subcluster) synthetically matured with **2** (Chapter II, paragraph 2.5.2) displayed a TOF of $\approx 10 \text{ min}^{-1}$, only 20-30 times larger than the TOFs obtained for the **2**-HydF hybrids (Figure **62a**).

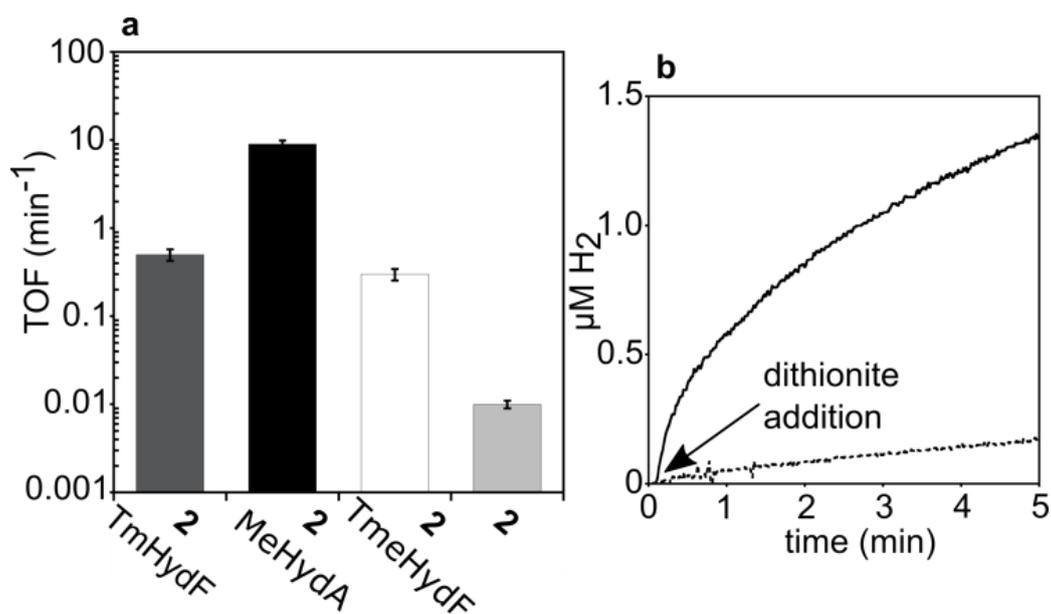


Figure 62: Hydrogen evolution activity of 2-HydF/A hybrids. **a)** TOF (min^{-1}) of 2-TmHydF(*), 2-TmeHydF, 2-MeHydA and 2. The y axis is shown as a logarithmic scale. **b)** H_2 evolution as a function of time using 2-TmeHydF (solid line) and 2 (dashed line). The data shown in a are mean values \pm s.d. based on three different experiments. *2-TmHydF activity is shown in paragraph 4.1.2.¹³

As for TmHydF hybrids (paragraph 4.1.2), again the complex seems to dissociate after few minutes of reaction. I support the hypothesis that the complex dissociates from the protein scaffold since the slope, initially higher, diminishes reaching values similar to the ones of free complex 2 in solution. This observation, again, points to the necessity of a structural characterization of the HydF hybrid in order to find the protein structural elements responsible for the binding and dissociation of the biomimetic complex.

Finally, although these activities are very low as compared to those of fully active HydA (maturated with synthetic complex 1 or purified in the presence of HYD protein machinery), they provide a unique opportunity to progress towards active artificial hydrogenases, with unnatural active sites however nicely mimicking the natural one.

4.2.4 Site-directed mutagenesis of TmeHydF: E305C and E305H

TmeHydF produces some hydrogen when artificially maturated with the organometallic complex 2. However, this complex dissociates after few minutes, proving that HydF has been built to carrier a diiron complex to the hydrogenase apoprotein. In the absence of a structure of the HydF hybrid, we analysed the crystal structure of iron-sulfur form, designing some mutants with respect to the fourth iron coordination site. This part of the

project has been done together with a master student, Cecilia Papini, which spent six months in the laboratory.

Two mutants have been prepared in pure form: E305H and E305C, in which the glutamate has been changed into a histidine and a cysteine, respectively.

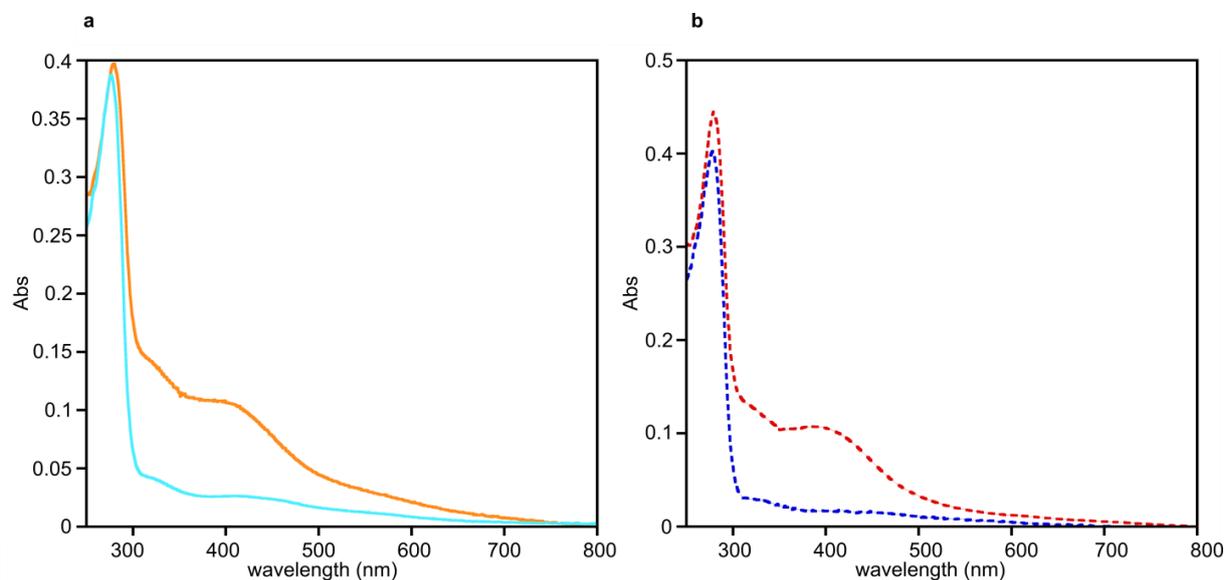


Figure 63: UV-Visible absorption spectra of 100 μ M apo- and FeS-TmeHydF-E305C mutant (a), 100 μ M apo- and FeS-TmeHydF-E305H mutant (b). The spectra of apoproteins were done in 25 mM Hepes buffer pH 7.5, 200 mM NaCl while the FeS-TmeHydF-E305C/H contained also 5 mM DTT. Color code: apo-TmeHydF-E305C (light blue), FeS-TmeHydF-E305C (orange), apoTmeHydF-E305H (dashed blue), FeS-TmeHydF-E305H (dashed red).

The mutation E305C was designed with the particular aim to hypothetically recreate the H-cluster of HydA. In fact the idea was that cysteine might bind the [4Fe-4S] and the diiron complex, acting as a bridge as it is observed in the HydA H-cluster. This way the insertion of the [2Fe]-subcluster could potentially result in an increased hydrogenase activity. The choice of histidine is due to the fact that this latter is a very common ligand for the Fe of the iron-sulfur cluster. Furthermore it has the advantage of being detected in spectroscopic techniques like HYSCORE spectroscopy, so it would have been easy to detect the binding between the histidine's imidazole and the iron of the cluster. A better comprehension of this binding site is absolutely required for the identification of the key structural elements responsible for the binding of the diiron unit and its release during catalytic conditions.

The [FeFe]-hydrogenase maturase HydF

Table 6: Iron content of FeS-TmeHydF and mutants before and after maturation with 2.

Species	Fe
FeS-TmeHydF-E305C	$3,4 \pm 0.3$
FeS-TmeHydF-E305H	$3,6 \pm 0.2$
FeS-TmeHydF	$3,5 \pm 0.1$
2-TmeHydF-E305C	$3,9 \pm 0.2$
2-TmeHydF-E305H	$4,1 \pm 0.3$
2-TmeHydF	$4,9 \pm 0.2$

The data shown are mean values on three different proteins reconstitution and maturation.

After the aerobic purification of both mutants, the iron-sulfur clusters were chemically reconstituted with protocols adopted for the wild type protein (Paragraph 5.4 for detailed protocols). Correct incorporation of the cluster is confirmed by UV-Visible spectroscopy (Figure 63) and iron quantitation (Table 6).

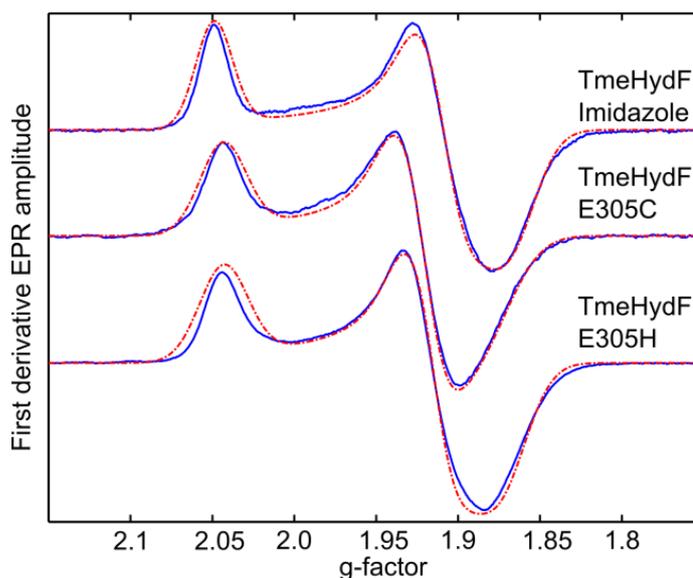


Figure 64: X-band (9.4 GHz) pseudo modulated 2-pulse echo detected EPR spectra of TmeHydF variants E305C, E305H as well as WT + imidazole. Experimental conditions: 10K, $\tau=200\text{ns}$, $\pi/2$ pulse =16ns. The spectral simulations (dotted red lines) were performed using the parameters listed in Table 7.

Here again, X-band EPR spectroscopy was applied to confirm the correct assembly of the [4Fe-4S] cluster inside the protein scaffolds. The EPR spectra evidenced that both mutants are able to assemble the cluster, although the 4th coordination site was mutated. In the presence of sodium dithionite a canonical spectrum of reduced [4Fe-4S]⁺ cluster was observed with small shifts in the g parameters (Figure 64, Table 7). It might reflect small

changes in the symmetry of the iron-sulfur cluster due to the replacement of the glutamate ligand (E305).

Table 7: G-values (g_1 , g_2 , g_3) and Gstrain parameters (gs_1 , gs_2 , gs_3) used to simulate the EPR spectra in Figure 52 and Figure 64.

Species	g_1	g_2	g_3	gs_1	gs_2	gs_3
TmeHydF	2.0455	1.8994	1.8641	0.020	0.020	0.047
1-TmeHydF	2.0484	1.9084	1.8671	0.0192	0.0306	0.0385
2-TmeHydF	2.0468	1.9084	1.8685	0.0208	0.0309	0.0369
TmeHydF	2.0485	1.9085	1.8681	0.0184	0.0269	0.0351
+ Imidazole						
TmeHydF	2.0430	1.9197	1.8851	0.0274	0.0266	0.0455
E305C						
TmeHydF	2.0430	1.9153	1.8738	0.0292	0.0263	0.0367
E305H						

Form the EPR spectra was not possible to state that the cysteine and histidine are ligands of the [4Fe-4S] cluster, thus we used HYSCORE spectroscopy which is a powerful technique for the detection of strong coupling between an $S=1/2$ cluster ([4Fe-4S]⁺) with an $I=1$ nuclear spin nucleus such as ¹⁴N of histidine. This His ligation should in fact produce a characteristic pattern in the (+, -) quadrant with so-called double quantum-double quantum correlation features, indicating a nitrogen interaction with the reduced iron-sulfur cluster, as already observed for wild type HydF in the presence of imidazole.¹⁶ The HYSCORE spectrum of the E305H mutant (Figure 65) only showed a weak diagonal feature around 3.5 MHz possibly indicating weakly interacting “distant” ¹⁴N nuclei. Therefore, in this mutant the introduced histidine does not seem to directly bind to the cluster.

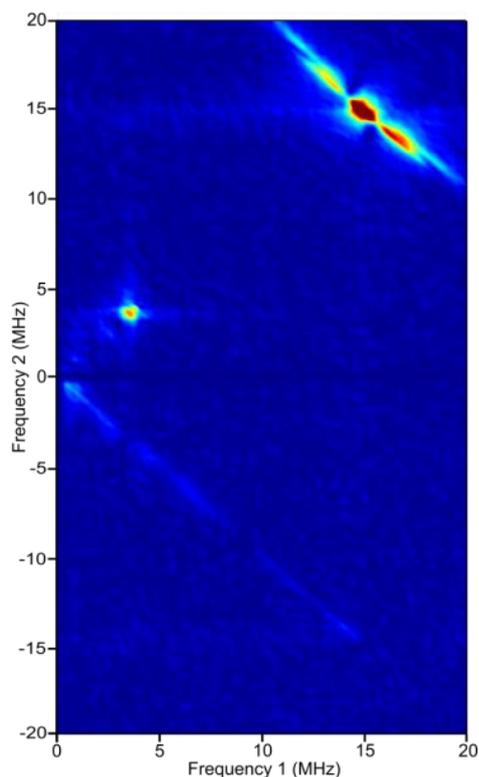


Figure 65: X-band (9.4 GHz) HYSCORE spectrum of E305H TmeHydF mutant. Experimental conditions: $T = 10\text{K}$, $\tau = 136\text{ ns}$, magnetic field was set to the maximum echo amplitude at 350 mT (near g_2).

Regarding the cysteine mutation, we have no direct evidence of its coordination to the iron-sulfur cluster, although the incorporation of synthetic complexes in the mutant showed interesting features. Reacting both mutants with the synthetic complexes, **1** and **2**, we could not find conditions to attach the complexes to the protein, as shown by Fe quantitation (Table **6**). While it is not clear whether the introduced cysteine is a ligand of the cluster or just adjacent (as in the case of the histidine in the E305H mutant), a cysteine at this position is much less appropriate than glutamate for the cluster to bind the complex.

The same experiment carried out with the E305H mutant led to similar observations (Table **6**). The percentage of diiron incorporation is quite low, confirming how critical this position is for assembling a 6Fe transient center. Both mutants proved to incorporate a single [4Fe–4S] cluster but could not bind the binuclear part. Thus, the flexibility of the glutamate E305 is fundamental to allow either binding to the iron-sulfur cluster or displacement for access of the binuclear complex to the cluster.

4.3 Conclusion and future perspectives

In conclusion, this first structural characterization of the TmeHydF maturase enzyme has provided unexpected observations that give new insights into its function and in the mechanism of hydrogenase maturation. The combination of X-ray structural results together with the FTIR, EPR and HYSCORE characterizations allowed establishing that the ligands of the [4Fe–4S] cluster in TmeHydF are provided by three cysteines and a glutamate residue (E305). The cluster is solvent-exposed, with a single loop (E299-D306) providing some protection from the solvent. This significant accessibility fits nicely with the function of HydF: the cluster presumably receives two Fe(CO)₂(CN) moieties from HydG, adt²⁻ from HydE and assembles the precursor of the [2Fe]-subcluster.

Furthermore, exchangeability of E305 has been well established since treatment of the FeS-TmeHydF protein with imidazole results in its binding to the cluster as unambiguously demonstrated by HYSCORE spectroscopy. As a further illustration of the accessibility of the cluster, as previously shown with FeS-TmHydF, we characterized the binding of complexes **1** and **2** into the iron-sulfur form of TmeHydF. FTIR clearly showed the presence of CO and CN narrow stretches assigned to cluster-bound biomimetic complexes. Moreover, the FTIR features were really reminiscent of that of HydF from *Clostridium acetobutylicum*, purified from native host leading us to conclude that the diiron complex synthesized by the HYD protein machinery is complex **1**.

E305 to histidine or cysteine mutation greatly impairs formation of the 6Fe-unit further supporting a critical role of the labile glutamate residue for controlling the access of the mimics and their binding to the cluster. Further characterizations should be done to get better insight into the active site environment of these mutants.

This thesis work also opens new avenues for addressing unsolved issues, such as the structure of protein complexes with the other maturases, HydE and HydG, the function of the GTP binding domain and the mechanism of the assembly in HydF of the key diiron precursor of the active site of hydrogenases. Finally it provides a unique basis towards the development of artificial hydrogenases based on HydF.

The efficacy of **2**-TmeHydF and **2**-TmHydF, as artificial hydrogenases, is quite low, but might be improved through a series of manipulations. One of the main limits of HydF as an artificial hydrogenase, as discussed above, is its labile and transient binding of the binuclear complex. One of the possible approaches could be the site-directed mutagenesis of the amino acids surrounding the iron-sulfur cluster looking for a stabilization of the diiron complex onto the [4Fe–4S]. Many mutations could be designed on the E305 residue or on other amino acids to get further insights on the binding site environment. The main focus

The [FeFe]-hydrogenase maturase HydF

remains the crystallization of the HydF hybrid protein and its conversion into an active enzyme. Of particular interest will be the glutamate-to-alanine mutation which should create an open coordination to the iron-sulfur cluster allowing the binding of synthetic complexes. Indeed, the alanine sidechain cannot bind the iron-sulfur cluster which could be free to interact with synthetic complexes as the wild type HydF protein. Experiments along this direction are in progress.

Moreover, previous studies reported that mutations of either conserved histidine of the CxH_{x_n}HC_{xx}C motif (Figure 56) in HydF from *C. acetobutylicum* resulted in severe impairment of HydA maturation suggesting that these histidines, which are structurally close to the cluster (Figure 53), play an important role in the chemical processes associated either in the assembly of the diiron precursor in HydF or its dissociation for transfer to HydA. The impact of similar mutations in other HydFs on HydA activation has not been studied so that it is impossible to make a generalization of the role of these histidines. They might be involved in H-bonding interactions with the binuclear part (notably with CO and CN acceptors) and this leads to a second possible idea, the mutation of these His with stronger ligands (Cys) which could better stabilize the binuclear compound. Obviously, we cannot mutate randomly all these residues and the X-ray structure of a HydF hybrid is fundamental for further studies.

In conclusion, the basic idea of converting an inactive enzyme into an artificial hydrogenase is still at its infancy and a lot of work directions can be found. For example other water soluble complexes can be thought for the incorporation inside protein scaffolds: Fe/Co-porphyrin, cobaloxime, mononuclear Fe complexes or other non-noble metals complexes can be envisaged. Experiments towards this direction are fascinating and might lead to interesting protein hybrids with unexpected activities.

References:

- 1 W. Lubitz, H. Ogata, O. Rüdiger and E. Reijerse, *Chem. Rev.*, 2014, **114**, 4081–4148.
- 2 D. L. M. Suess, I. Bürstel, L. De La Paz, J. M. Kuchenreuther, C. C. Pham, S. P. Cramer, J. R. Swartz and R. D. Britt, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 11455–11460.
- 3 P. Dinis, D. L. M. Suess, S. J. Fox, J. E. Harmer, R. C. Driesener, L. De La Paz, J. R. Swartz, J. W. Essex, R. D. Britt and P. L. Roach, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 1362–1367.

- 4 J. M. Kuchenreuther, W. K. Myers, D. L. M. Suess, T. A. Stich, V. Pelmeshnikov, S. A. Shiigi, S. P. Cramer, J. R. Swartz, R. D. Britt and S. J. George, *Science*, 2014, **343**, 424–427.
- 5 Y. Nicolet, J. K. Rubach, M. C. Posewitz, P. Amara, C. Mathevon, M. Atta, M. Fontecave and J. C. Fontecilla-Camps, *J. Biol. Chem.*, 2008, **283**, 18861–18872.
- 6 Y. Nicolet, R. Rohac, L. Martin and J. C. Fontecilla-Camps, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 7188–7192.
- 7 J. N. Betz, N. W. Boswell, C. J. Fugate, G. L. Holliday, E. Akiva, A. G. Scott, P. C. Babbitt, J. W. Peters, E. M. Shepard and J. B. Broderick, *Biochemistry*, 2015, 150205170738009.
- 8 R. Rohac, P. Amara, A. Benjdia, L. Martin, P. Ruffié, A. Favier, O. Berteau, J.-M. Mouesca, J. C. Fontecilla-Camps and Y. Nicolet, *Nat. Chem.*, 2016, **8**, 491–500.
- 9 L. Cendron, P. Berto, S. D’Adamo, F. Vallese, C. Govoni, M. C. Posewitz, G. M. Giacometti, P. Costantini and G. Zanotti, *J. Biol. Chem.*, 2011, **286**, 43944–43950.
- 10 E. M. Shepard, F. Mus, J. N. Betz, A. S. Byer, B. R. Duffus, J. W. Peters and J. B. Broderick, *Biochemistry*, 2014, **53**, 4090–4104.
- 11 I. Czech, A. Silakov, W. Lubitz and T. Happe, *FEBS Lett.*, 2010, **584**, 638–642.
- 12 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.
- 13 V. Artero, G. Berggren, M. Atta, G. Caserta, S. Roy, L. Pecqueur and M. Fontecave, *Acc. Chem. Res.*, 2015, **48**, 2380–2387.
- 14 J. Esselborn, N. Muraki, K. Klein, V. Engelbrecht, N. Metzler-Nolte, U.-P. Apfel, E. Hofmann, G. Kurisu and T. Happe, *Chem Sci*, 2016, **7**, 959–968.
- 15 X. Brazzolotto, J. K. Rubach, J. Gaillard, S. Gambarelli, M. Atta and M. Fontecave, *J. Biol. Chem.*, 2006, **281**, 769–774.
- 16 G. Berggren, R. Garcia-Serres, X. Brazzolotto, M. Clemancey, S. Gambarelli, M. Atta, J.-M. Latour, H. Hernández, S. Subramanian, M. Johnson and M. Fontecave, *J. Biol. Inorg. Chem.*, 2014, **19**, 75–84.
- 17 E. M. Shepard, A. S. Byer, J. N. Betz, J. W. Peters and J. B. Broderick, *Biochemistry*, 2016.
- 18 E. M. Shepard, S. E. McGlynn, A. L. Bueling, C. S. Grady-Smith, S. J. George, M. A. Winslow, S. P. Cramer, J. W. Peters and J. B. Broderick, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 10448–10453.

The [FeFe]-hydrogenase maturase HydF

- 19 C. S. Bond, *Bioinforma.*, 2003, **19**, 311–312.
- 20 P. Berto, M. Di Valentin, L. Cendron, F. Vallese, M. Albertini, E. Salvadori, G. M. Giacometti, D. Carbonera and P. Costantini, *Biochim. Biophys. Acta BBA - Bioenerg.*, 2012, **1817**, 2149–2157.
- 21 M. Albertini, P. Berto, F. Vallese, M. Di Valentin, P. Costantini and D. Carbonera, *J. Phys. Chem. B*, 2015, **119**, 13680–13689.
- 22 M. Albertini, L. Galazzo, L. Maso, F. Vallese, P. Berto, E. De Rosa, M. Di Valentin, P. Costantini and D. Carbonera, *Top. Catal.*, 2015, **58**, 708–718.
- 23 G. E. Crooks, G. Hon, J.-M. Chandonia and S. E. Brenner, *Genome Res.*, 2004, **14**, 1188–1190.
- 24 H. Jiang, X. Zhang, C. Ai, Y. Liu, J. Liu, G. Qiu and J. Zeng, *J. Microbiol. Biotechnol.*, 2008, **18**, 1070–1075.
- 25 L. Calzolari, C. M. Gorst, Z. H. Zhao, Q. Teng, M. W. Adams and G. N. La Mar, *Biochemistry*, 1995, **34**, 11373–11384.
- 26 N. Muraki, J. Nomata, K. Ebata, T. Mizoguchi, T. Shiba, H. Tamiaki, G. Kurisu and Y. Fujita, *Nature*, 2010, **465**, 110–114.
- 27 M. Lee, T. Gräwert, F. Quitterer, F. Rohdich, J. Eppinger, W. Eisenreich, A. Bacher and M. Groll, *J. Mol. Biol.*, 2010, **404**, 600–610.
- 28 N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 10037–10041.
- 29 R. J. Gurbiel, P. E. Doan, G. T. Gassner, T. J. Macke, D. A. Case, T. Ohnishi, J. A. Fee, D. P. Ballou and B. M. Hoffman, *Biochemistry (Mosc.)*, 1996, **35**, 7834–7845.
- 30 W. Wang, J. Li, K. Wang, T. I. Smirnova and E. Oldfield, *J. Am. Chem. Soc.*, 2011, **133**, 6525–6528.
- 31 P. Perche-Letuvé, V. Kathirvelu, G. Berggren, M. Clemancey, J.-M. Latour, V. Maurel, T. Douki, J. Armengaud, E. Mulliez, M. Fontecave, R. Garcia-Serres, S. Gambarelli and M. Atta, *J. Biol. Chem.*, 2012, **287**, 41174–41185.
- 32 J. Esselborn, C. Lambertz, A. Adamska-Venkatesh, T. Simmons', G. Berggren, J. Nothl, J. Siebel, A. Hemschemeier, V. Artero, E. Reijerse, M. Fontecave, W. Lubitz and T. Happe, *Nat. Chem. Biol.*, 2013, **9**, 607–609.
- 33 G. Caserta, A. Adamska-Venkatesh, L. Pecqueur, M. Atta, V. Artero, R. Souvik, E. Reijerse, W. Lubitz and M. Fontecave, *Biochim. Biophys. Acta BBA - Bioenerg.*, 2016.

General conclusions

There is a general agreement that the building of a sustainable hydrogen economy, based on the combined development of photoelectrolyzers (for energy storage through water splitting into H₂ and O₂) and of fuel cells (for converting H₂ and O₂ back into electricity) in part relies on the availability of cheap and efficient catalysts. The discovery of the hydrogenases and subsequent understanding of their structure and chemistry have been major breakthroughs towards novel catalytic strategies aimed at replacing platinum currently used in the technological devices mentioned above. However, only few [FeFe]-hydrogenases have been structurally and catalytically characterized showing high variability despite the same catalytic core.

Recently, it has been shown that recombinant [FeFe]-hydrogenases can be prepared anaerobically in an inactive form containing only the [Fe–S] clusters and subsequently fully matured by reaction with the synthetic [Fe₂(adt)(CO)₄(CN)₂]²⁻ biomimetic complex (**1**).

In this thesis we went a step further towards the facile preparation of active hydrogenases. Indeed it is sufficient i) to express heterologously the structural gene in *E. coli*, ii) purify the protein aerobically, iii) reconstitute the [Fe–S] clusters and subsequent iv) mature chemically the enzyme with a synthetic organometallic complex (**1**). Thus, it becomes possible to assemble all the metal cofactors using simple chemicals like iron and sulfide salts and a synthetic complex (**1**), showing that this procedure is efficient even for the case of a protein containing three [4Fe–4S] clusters. EPR and FTIR characterizations proved that the chemically matured HydA from *Megasphaera elsdenii* is virtually indistinguishable from enzymes purified from the native hosts. H_{ox}, H_{red} and H_{ox}-CO states were indeed isolated and characterized for the first time for this protein. Moreover, MeHydA, as other hydrogenases, displays remarkable performances in particular with respect to turnover frequencies (TOFs).

Despite the incredible catalytic performances, these enzymes are fraught with a number of drawbacks which still need to be addressed before they can be implemented into biotechnological devices. One is their limited tolerance to molecular oxygen and CrHydA1 (from *Chlamydomonas reinhardtii*), in particular, is one of the most unstable. Another drawback is their size which limits enzyme density at the surface of the electrodes; thus smaller and robust enzymes are absolutely required.

The simplest and most studied hydrogenase is CrHydA1, which has been exploited more recently for biochemical and spectroscopic characterization because it lacks the additional

General conclusions

iron-sulfur clusters observed in most bacterial [FeFe]-hydrogenases, which makes hard the characterization of the catalytic core. Inspired by CrHydA1, we have realized a truncated HydA from *M. elsdenii* lacking the ferredoxin domains, as an example to illustrate the concept that it is possible to engineer hydrogenases to generate smaller and still active enzyme versions.

The methodology implemented for MeHydA full length protein has been also applied for this new construct, called MeH-HydA. The protein was aerobically purified with high yields and the unique [4Fe-4S] was correctly assembled as shown by the EPR and Mössbauer characterization. Initial attempts of maturation with the $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex showed little complex incorporation which gave rise to small activity. Further studies allowed finding new experimental conditions, which increased the complex incorporation up to 50-70%. Once activated, the protein showed high activity towards protons reduction only 4-times lower compared to the wild type protein despite the deletion of the N-terminal domain harboring two F [4Fe-4S] clusters.

The preliminary FTIR and EPR characterization shows that the H-cluster of as isolated MeH-HydA is mainly in the $\text{H}_{\text{ox}}\text{-CO}$ configuration. This signal is often present in [FeFe]-hydrogenase preparations due to the so-called “cannibalization process” in which the CO from light- or oxygen-damaged H-clusters are released and captured by still intact H-clusters; however the $\text{H}_{\text{ox}}\text{-CO}$ populations resulted to be unusually higher than other hydrogenases prepared using the chemical maturation tool.

Other experiments are still in progress to elucidate all the other accessible oxidation states of this new “chlamydomonas-like” hydrogenase.

Finally, the chemical insertion of biomimetic complexes onto HydF protein scaffold described by Berggren and coworkers showed that HydF hybrids have a 6Fe metallo-cofactor reminiscent of the catalytically competent H-cluster. This close similarity between the H-cluster and the metal cofactor of α -HydF hybrids led to the idea that HydF could serve as good model system for the realization of artificial hydrogenases. New HydF proteins have been efficiently produced and their [4Fe-4S] cluster assembled correctly via the methodology described above. Moreover, we succeeded to crystallize the first HydF protein with its [4Fe-4S] cluster (HydF from *Thermosipho melanesiensis*) showing an unusual and exchangeable coordination to the [4Fe-4S] cluster. The combination of X-ray structural results together with the spectroscopic characterization (EPR, HYSORE) allowed establishing that the ligands of the [4Fe-4S] cluster in TmeHydF are provided by three cysteines and a glutamate residue (E305). The cluster is solvent-exposed, with a single loop providing some protection

General conclusions

from the solvent. This significant accessibility fits nicely with the function of HydF: the cluster presumably receives two $\text{Fe}(\text{CO})_2(\text{CN})$ moieties from HydG, adt^{2-} from HydE and assembles the precursor of the $[\text{2Fe}]$ -subcluster.

Furthermore, exchangeability of E305 has been well established since treatment of the FeS-TmeHydF protein with imidazole results in its binding to the cluster as unambiguously demonstrated by HYSORE spectroscopy. As a further illustration of the accessibility of the cluster we characterized the binding of complexes **1** and **2** into TmeHydF. FTIR clearly showed the presence of CO and CN narrow stretches assigned to cluster-bound biomimetic complexes. Moreover, the FTIR features were really reminiscent of that of HydF from *Clostridium acetobutylicum*, purified from native host leading us to conclude that the diiron complex synthesized by the HYD protein machinery is complex **1**.

HydF hybrids were tested for the first time for proton reduction assays with a miniaturized Clark-type microsensor proving that HydF chemically matured with complex **2**, $[\text{Fe}_2(\text{pdt})(\text{CO})_4(\text{CN})_2]^{2-}$, is able to catalyze hydrogen evolution with an initial turnover rate of 0.5 min^{-1} about 30 times larger than the compound free in solution. The reaction started immediately after the addition of dithionite then levelled off after a few minutes, while the system was still active after 5 min reaction time. However, the efficacy of **2**-HydFs is quite low and one of the main limits of HydF as an artificial hydrogenase is its labile and transient binding of the binuclear complex. A possible solution could be the site-directed mutagenesis of the amino acids surrounding the iron-sulfur cluster looking for a stabilization of the diiron complex onto the $[\text{4Fe-4S}]$. This is now possible with the HydF structure reported in this manuscript even though an X-ray structure of the HydF hybrid, with its 6Fe cluster, is absolutely required for the success of the project.

In conclusion, the basic idea of converting an inactive enzyme into an artificial hydrogenase is still at its infancy and a lot of work directions can be found. As said above, not only HydF can be modified by rational mutagenesis but also other synthetic complexes might be incorporated onto HydF scaffold. Experiments towards this direction are fascinating and might lead to interesting protein hybrids with unexpected activities.

This three year work has been focused on the better comprehension of the hydrogenase field. Different projects have been carried out with aim of elucidating the complex world of hydrogenases. Still a lot of work remains to be done and we are still far for a biotechnological and industrial application of these enzymes for the H_2 evolution and/or uptake, but I hope that this PhD thesis might serve as an interesting starting point for other scientists who want to join/pursue their study on these incredible enzymes.

General conclusions

Chapter V
Materials and Methods

Chapter V: Materials and Methods

5.1 Biological material

5.1.1 Competent cells

- *E. coli* DH5 α

These competent cells have been used to prepare plasmid stock solutions.

- *E. coli* BL21(DE3)

Strains used for the expression of TmeHydF and CtHydF proteins. The *LacI* repressor inhibits the expression of genes coding for proteins involved in the metabolism of lactose. These genes are repressed when lactose is not available to the cell. When lactose is present allolactose binds to the *lac* repressor that is unable to bind to its operator and this effect induces expression of the metabolic genes. *In vitro*, Isopropyl β -D-1 thiogalactopyranoside (IPTG) is a commonly used allolactose mimic which can be used to induce transcription of target genes.

- *E. coli* BL21 Codonplus® (DE3)-RIL

E. coli strains (chloramphenicol resistance) used for the expression of TmHydF protein.

These strains contain extra copies of the *argU*, *ileY*, and *leuW* tRNA genes. These genes encode tRNAs that recognize the arginine codons AGA and AGG, the isoleucine codon AUA, and the leucine codon CUA, respectively. The CodonPlus-RIL strains have available the tRNAs that most frequently restrict translation of heterologous proteins from organisms that have AT-rich genomes.

- *E. coli* Tuner(DE3)

Materials and Methods

Strains used for the expression of MeH-HydA. *E. coli* strains with a *lacZY* deletion, which enable adjustable levels of protein expression throughout all cells in a culture. The *lac* permease (*lacY*) mutation allows uniform entry of IPTG into all cells in the population, which produces a concentration-dependent, homogeneous level of induction. By adjusting the concentration of IPTG, expression can be regulated from very low levels up to the robust, fully induced levels. Lower-level expression may enhance the solubility and activity of difficult target proteins.

- *E. coli* Tuner(DE3)pLysS

Strains (chloramphenicol resistance) used for the expression of MeHydA. These strains have the same characteristics of Tuner(DE3). The extra pLysS strain expresses T7 lysozyme, which further suppresses basal expression of T7 RNA polymerase prior to induction, thus stabilizing pET recombinants encoding target proteins that affect cell growth and viability.

5.1.2 Plasmids

Table 8: Plasmids used in this thesis.

pT7-7-TmHydF	HydF from <i>Thermotoga maritima</i>
pT7-7-MeH-HydA	<i>HDomain</i> of HydA from <i>Megasphaera elsdenii</i> (N _{ter} 6His tag)
pT7-7-MeHydA	HydA from <i>Megasphaera elsdenii</i> (N _{ter} 6His tag)
pet22b-CystoSerMeHydA	HydA from <i>Megasphaera elsdenii</i> with serines mutations (N _{ter} 6His tag)
pet22b-TmeHydF	HydF from <i>Thermosipho melanesiensis</i>
pet22b-CtHydF	HydF from <i>Clostridium thermocellum</i>
pet22b(+)-EcCsdA	CsdA from <i>E. coli</i> (6His tag)
pet22b-TmeHydF(E305H)*	E305H mutant HydF from <i>Thermosipho melanesiensis</i>
pet22b-TmeHydF(E305C)*	E305C mutant HydF from <i>Thermosipho melanesiensis</i>

* Site directed mutagenesis and purifications of TmeHydF mutants was conducted in collaboration with Cecilia Papini during her master internship at Collège de France.

5.1.3 Growth media

LB (*Lysogeny broth*) and TB (*Terrific broth medium*) media were used for protein expression. Commercially available powders were dissolved in water and sterilized at 120°C using an autoclave.

5.1.4 Molecular Biology

5.1.4.1 Competent cells preparation

Competent cells are ready to use bacterial cells that possess more easily altered cell walls by which foreign DNA can be passed through easily. Most types of cells cannot take up DNA efficiently unless they have been exposed to special chemical or electrical treatments to make them competent. The standard method for making the bacteria permeable to DNA involves treatment with calcium ions. Brief exposure of cells to an electric field also allows the bacteria to take up DNA and this process is called as electroporation. The CaCl₂ treatment has been applied. All *E. coli* strains were grown overnight at 37 °C in LB medium supplemented with the correct antibiotic. The overnight culture was used to inoculate 10 mL of fresh LB growth medium and incubated at 37 °C (200 rpm) until the optical density at 600 nm reached 0.4. The solution was cool down with an ice bath for 10 minutes and then cells were harvested by centrifugation (10 minutes at 5000 rpm, 4°C). The paste cells was re-suspended in 5 mL of 0.1 M calcium chloride solution (sterile) and incubated in an ice bath. After 30 minutes, cells were harvested by centrifugation (10 minutes, 5000 rpm, 4°C) and re-suspended in 0.7 mL of 0.1 M calcium chloride solution (sterile) supplemented with 10% glycerol. Aliquots of 100/200 µl were prepared and frozen in liquid nitrogen (storage at -80 °C).

5.1.4.2 Transformation

One aliquot of *E. coli* competent cells is reacted with 1-3 µl of related plasmid (50 µg). After incubation in a water/ice bath for 30 minutes, the solution was heated at 42 °C for 2 minutes. 0.9 mL of fresh LB growth medium is added to the cell solution, incubating at 37 °C for 1h (400-500 rpm). The solution was then spread on a Petri dish containing LB-Agar growth medium supplemented with the correct antibiotic solution and incubated at 37 °C

overnight. The day after the Petri dish, containing several colonies, is stored at 4°C or directly used for expression tests.

5.1.4.3 Plasmid preparation

Plasmids stock solutions have been prepared using an *E. coli* DH5 α strain transformed with plasmids listed in Table 8. After the transformation procedure (Paragraph 5.1.4.2) one colony has been used to inoculate 10/50 mL of LB growth medium and incubated at 37 °C overnight (200 rpm). A kit for plasmid DNA purification (Quick-Start Protocol) provided by QUIAGEN has been used.

The protocol consists in the following steps:

- 1- Pellet 1-5 mL bacterial overnight culture by centrifugation at >8000 rpm (6800 xg) for 3 minutes at room temperature (15-25°C).
- 2- Resuspend pelleted bacterial cells in 250 μ l Buffer P1 and transfer to a microcentrifuge tube.
- 3- Add 250 μ l Buffer P2 and mix thoroughly by inverting the tube 4-6 times until the solution becomes clear. Do not allow the lysis reaction to proceed for more than 5 min. if using LyseBlue reagent, the solution will turn blue.
- 4- Add 350 μ l Buffer N3 and mix immediately and thoroughly by inverting the tube 4-6 times. If using the LyseBlue reagent, the solution will turn colorless.
- 5- Centrifuge for 10 min at 13.000 rpm (17900 x g) in a table-top microcentrifuge.
- 6- Apply 800 μ l supernatant from step 5 to the QIAprep 2.0 spin column by pipetting. Centrifuge for 30-60 s and discard the flow-through.
- 7- Wash the QIAprep 2.0 spin column by adding 0.5 ml Buffer PB. Centrifuge for 30-60 s and discard the flow-through.
- 8- Wash the QIAprep 2.0 spin column by adding 0.75 ml Buffer PE. Centrifuge 30-60 s and discard the flow-through. Transfer the QIAprep 2.0 spin column to the collection tube.
- 9- Centrifuge for 1 min to remove residual wash buffer.
- 10- Place the QIAprep 2.0 column in a clean 1.5 ml microcentrifuge tube. To elute DNA, add 50 μ l Buffer EB or water to the center of the QIAprep 2.0 column, let stand for 1 min, and centrifuge for 1 min.

The plasmid stock solution was analysed by UV-Vis absorption spectroscopy (Figure 66), monitoring the absorbance at 260 nm (Thermo Nanodrop), determining the DNA concentration.

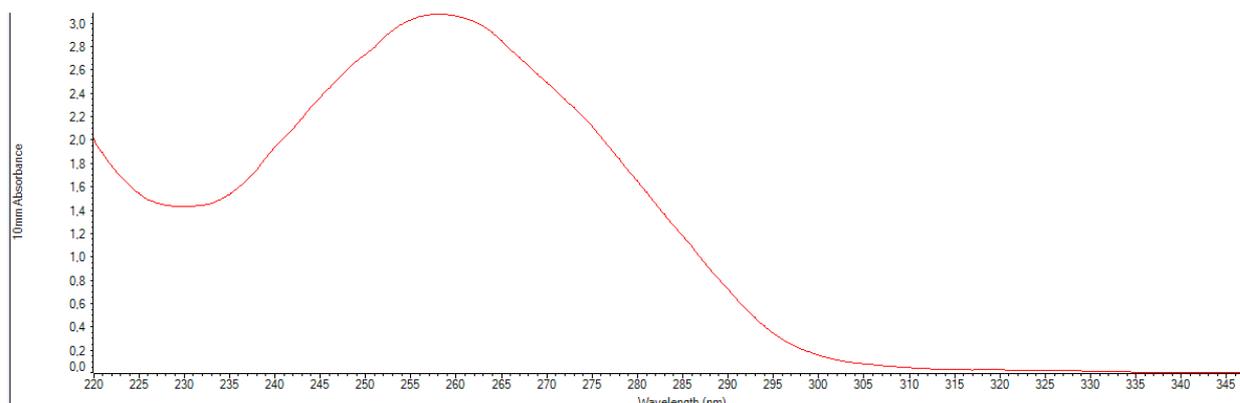


Figure 66: UV-Visible spectrum of MeHydA DNA after purification.

In order to assure that the purification led to the correct plasmids without any mutation, small aliquots have been used for DNA sequencing tests (Eurofins MWG Operon Company).

5.2 Biochemical methods: Protein expression

5.2.1 TmHydF: HydF from *Thermotoga maritima*

As *T. maritima* genes use numerous rare codons that code for arginine, isoleucine, or leucine residues, *E. coli* BL21-CodonPlus(DE3)-RIL cells (Stratagene) were used for the expression of the target protein. Cells transformed with the pT7-7-TmHydF plasmid were grown in LB medium supplemented with ampicillin and chloramphenicol at 37 °C, until the optical density at 600 nm reached 0.5. Protein expression was induced by addition of isopropyl β -D-thiogalactopyranoside to a final concentration of 0.5 mM and cells were grown for additional 5 h at 37°C. Cells were pelleted and stored at -80°C until use.^{1,2}

5.2.2 HydF from *Thermosipho melanesiensis* (TmeHydF) and *Clostridium thermocellum* (CtHydF)

Materials and Methods

Commercially supplied codon optimized synthetic plasmids of *Thermosipho melanesiensis* and *Clostridium thermocellum* HydF (pet22b-TmeHydF and pet22b-CtHydF, Ampicillin resistance) were ordered from Genscript.

Chemically competent *E. coli* BL21(DE3) cells were transformed with pet22b-TmeHydF and pet22b-CtHydF and used for proteins expression. 100mL (LB medium) overnight cultures were used to inoculate a larger scale cell culture (6L) at 37 °C until the optical density at 600 nm reached 0.5. Protein expression was induced by the addition of isopropyl β -D-thiogalactopyranoside to a final concentration of 0.5 mM. Cells were grown for an additional 4 h at 37 °C and were harvested by centrifugation and stored at -80°C until use.

5.2.2.1 E305C and E305H TmeHydF mutants

Site directed mutagenesis was performed using the QuickChange® strategy. Mutant sequences were checked by DNA sequencing. In Table 9 are reported the respective primers used for the generation of TmeHydF mutants. Expression and purification was carried out as described for wild type TmeHydF.

Table 9: Oligonucleotides used for the generation of E305C and E305H mutants.

Primer Name	Sequence
TmeHydF E305H forward	5' CGCATCGTCCGCTGACCCACGATATTGGTCGTG 3'
TmeHydF E305H reverse	5' CACGACCAATATCGTGGGTCAGCGGACGATGCG 3'
TmeHydF E305C forward	5' CGCATCGTCCGCTGACCTGTGATATTGGTCGTGTTAAAATTCCG 3'
TmeHydF E305C reverse	5' CGGAATTTTAAACACGACCAATATCACAGGTCAGCGGACGATGCG 3'

The 4 primers were purchased from Eurogentec.

5.2.3 MeHydA: the [FeFe]-hydrogenase from *Megasphaera elsdenii*

TunerDE3pLysS cells transformed with the pT7-7-MeHydA plasmid were grown in Terrific Broth medium supplemented with ampicillin and chloramphenicol at 37 °C. 100 mL

overnight cultures were used to inoculate a larger scale cell culture (6L) at 37 °C until the optical density at 600 nm reached 0.5. Protein expression was induced by the addition of isopropyl β -D-thiogalactopyranoside to a final concentration of 0.5 mM. Cells were grown for an additional 5 h with decreasing the temperature to 20 °C to avoid the formation of inclusion bodies. Cells were harvested by centrifugation and stored at -80°C until use.³

A second protocol with enhanced protein yield has been developed. TunerDE3pLysS cells transformed with the pT7-7-MeHydA plasmid were grown in LB medium supplemented with 5% glucose and 30 mM potassium phosphate buffer pH 7.6, ampicillin and chloramphenicol at 37 °C, until the optical density at 600 nm reached 0.5. Cells were grown overnight decreasing the temperature to 18°C to avoid formation of inclusion bodies.

5.2.4 MeH-HydA: the [FeFe]-hydrogenase truncated form of *Megasphaera elsdenii*

TunerDE3 cells transformed with the pT7-7-MeH-HydA plasmid were grown in Terrific Broth medium supplemented with ampicillin and chloramphenicol at 37 °C. 100mL overnight cultures were used to inoculate a larger scale cell culture (6L) at 37 °C until the optical density at 600 nm reached 0.5. Protein expression was induced by the addition of isopropyl β -D-thiogalactopyranoside to a final concentration of 0.5 mM. Cells were grown for an additional 5-6 h with decreasing the temperature to 20 °C to avoid the formation of inclusion bodies. Cells were harvested by centrifugation and stored at -80°C until use.

5.2.5 CsdA: the Cysteine Desulfurase from *E. coli*

Chemically competent *E. coli* BL21(DE3) cells were transformed with pet22b(+)-CsdA, containing an hexahistidine tag (Ampicillin resistance), and used for protein expression. 100 mL (LB medium) overnight cultures were used to inoculate a larger scale cell culture (4L) at 37 °C until the optical density at 600 nm reached 0.5. Protein expression was induced by the addition of isopropyl β -D-thiogalactopyranoside to a final concentration of 0.5 mM. Cells were grown for an additional 3-4h at 37 °C and were harvested by centrifugation and stored at -80°C until used.⁴

5.3 Proteins purification

Depending on the protein target, different purification strategies were chosen. HydF proteins belong to thermophilic organisms. HydAs and CsdA proteins have a 6His tag to facilitate protein purification using Ni-NTA columns.

5.3.1 TmHydF: HydF from *Thermotoga maritima*

As already reported in literature,^{1,2} cells were re-suspended and lysed by discontinuous sonication in a water/ice bath in the presence of 0.5 % Triton X solution (detergent). Cellular extracts were centrifuged 1h at 193.000 g leading to a soluble fraction of HydF protein. The resulting supernatant was heated at 75 °C for 12 min and denatured *E. coli* proteins were removed by centrifugation at 7.000g for 10 min at 4 °C. Ammonium sulfate was slowly added to the supernatant at 4 °C to a final concentration of 65 % (w/v). The cloudy solution was stirred for 60 min, centrifuged at 7.000g for 10 min at 4°C and the supernatant was discarded. The resulting off-white pellet was dissolved in a 50 mM Tris–HCl buffer pH 8, 150 mM NaCl, 1 M ammonium sulfate buffer (buffer A) under stirring. The first chromatographic step was performed on a butyl Sepharose 4 FF column (GE Healthcare). After the column had been washed with buffer A, removing all nucleic acids and remaining *E. coli* contaminants, TmHydF protein was eluted with a gradient from 100 % buffer A to 100 % buffer B 50 mM Tris–HCl, pH 8, 150 mM NaCl (Figure 67). Eluted fractions were analyzed by SDS-PAGE and the remaining ammonium sulfate was removed using a desalting column equilibrated with Tris–HCl 50 mM, 150 mM NaCl buffer.

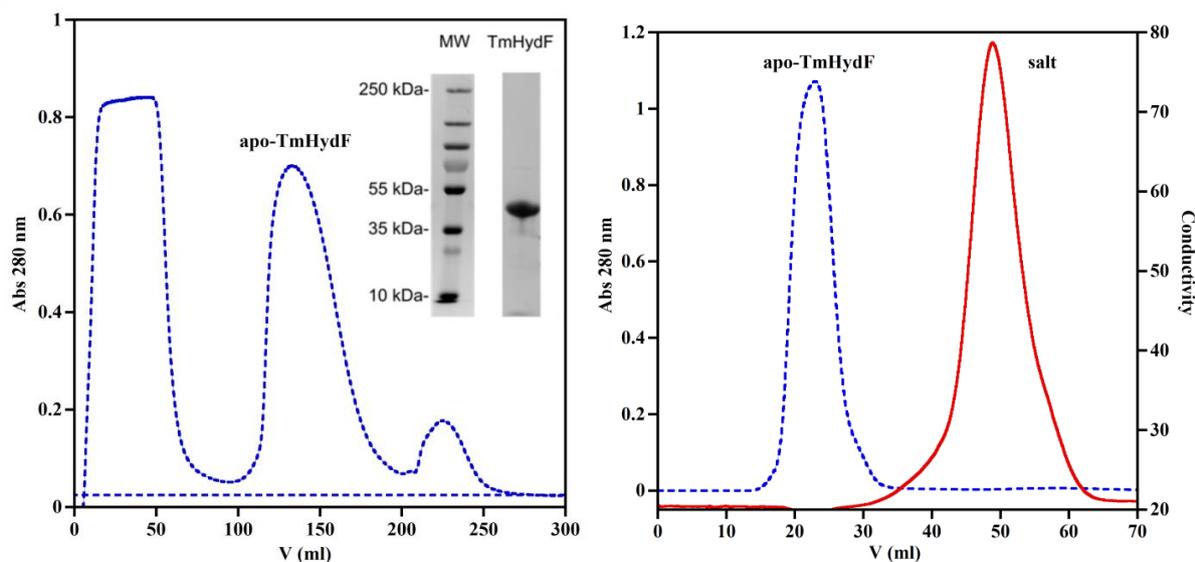


Figure 67: Butyl Sepharose chromatographic step (left) of apo-TmHydF, followed by desalting purification step (right) to remove the excess of ammonium sulfate. SDS page in denaturing conditions shows a single band corresponding to the HydF monomer of 45 kDa.

Eluted fractions containing the brownish (due to the presence of substoichiometric [Fe-S] clusters) TmHydF protein were pooled and concentrated using Amicon 30-kDa centrifugal filters up to 5-10 mg/ml. The apo-proteins were finally aliquoted, flash-frozen, and stored at -80°C . HydF protein prepared with this protocol was mainly dimeric containing a small amount of tetrameric fraction. Thus, the desalting step was replaced by size exclusion chromatographic step on a Superdex S200 26/600 in 25 mM Tris-HCl, pH 8, 200 mM NaCl which allowed removing the small tetrameric fraction. The eluted dimer was concentrated to 5-10 mg/ml, flash-frozen in liquid nitrogen and stored at -80°C until used. The oligomeric state of the protein was determined in anaerobic conditions by analytical gel filtration on a Superdex S 200 10/300 column equilibrated with 25 mM Tris-HCl buffer pH 8.0, 200 mM NaCl, 5 mM dithiothreitol (DTT).

5.3.2 HydF from *Thermosipho melanesiensis* (TmeHydF) and *Clostridium thermocellum* (CtHydF)

Cells were re-suspended in Tris-HCl Buffer 25 mM pH 8.0 containing 200 mM NaCl and discontinuously sonicated for 10 min in a water/ice bath. Cellular extracts were centrifuged 75 minutes at 193.000 g leading to a soluble fraction of both HydF proteins.

Materials and Methods

As both proteins belong to two thermophilic organisms, the first step was to study the temperature resistance in order to establish the optimal purification procedure. Heating the supernatant in the range 45-70 °C for 15 minutes allowed finding the best temperatures for TmeHydF (50°C) and CtHydF proteins (60°C). The heating step denatured and precipitated many *E. coli* proteins, which were removed by centrifugation (10 minutes, 7.000g at 4°C). Some essays of protein precipitation with ammonium sulfate in the range 20-70 % w/v allowed finding the correct salt concentration responsible for the precipitation of target proteins (Figure 68). An SDS page of protein pellets showed that at 30% w/v of ammonium sulfate TmeHydF is the only protein prone to precipitate.

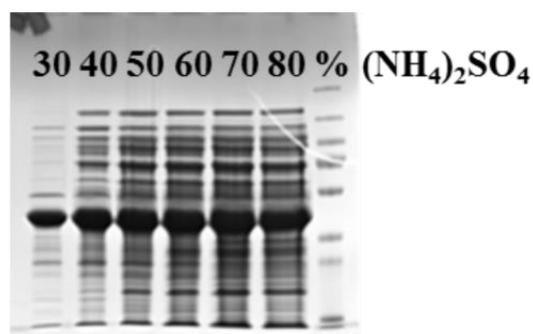


Figure 68: SDS page of TmeHydF protein pellets produced by different ammonium sulfate concentrations.

The brownish protein solution was filtered and the precipitated incubating the solution with 174 g (NH₄)₂SO₄ per L of protein solution (30 % w/v) for 60 minutes at 4°C.

The solution was centrifuged (10 minutes, 7000 rpm at 4°C) and the protein pellet was re-suspended in Tris-HCl buffer 25 mM pH 8.0, containing 200 mM NaCl. The last and only chromatographic step was performed on a Superdex S200 26-600 equilibrated in 25 mM Tris-HCl pH 8.0, 200 mM NaCl. SDS page allowed the identification of the apo-TmeHydF fractions from all the other contaminants (nucleic acids, remaining *E. coli* proteins and (NH₄)₂SO₄ coming from the last precipitation step, Figure 49a) and a pure and dimeric protein was collected, concentrated to 5-10 mg/ml and stored at -80°C until used.

CtHydF was purified in a similar manner; after a heating step at 60°C for 15 minutes to remove the *E. coli* proteins, the target protein was precipitated with ammonium sulfate (35%) and purified on a Superdex S 200 26/600 equilibrated with 25 mM Tris-HCl buffer pH 8.0 containing 200 mM NaCl. The dimeric apo-protein was concentrated to 5-10 mg/ml, frozen in liquid nitrogen and stored at -80°C.

As for TmHydF protein, the oligomeric state of HydFs from both organisms was determined in anaerobic conditions by analytical gel filtration on a Superdex S 200 10/300 column equilibrated with 25 mM Hepes buffer pH 7.5, 200 mM NaCl, 5 mM dithiothreitol (DTT).

5.3.2.1 E305C and E305H TmeHydF mutants

E305C and E305H were purified using the same strategy for wild type TmeHydF. The purity was assessed by SDS page (Figure 69) and the oligomeric state by size exclusion chromatography. As shown by a calibration on a Superdex S 200 10/300 in 25 mM Hepes pH 7.5, 200 mM NaCl and 5 mM DTT both purified mutants behaved as dimeric species with a retention volume of 13.2 mL (Figure 69). The molecular weight derived for both apo-TmeHydF mutants is 88.6 kDa.

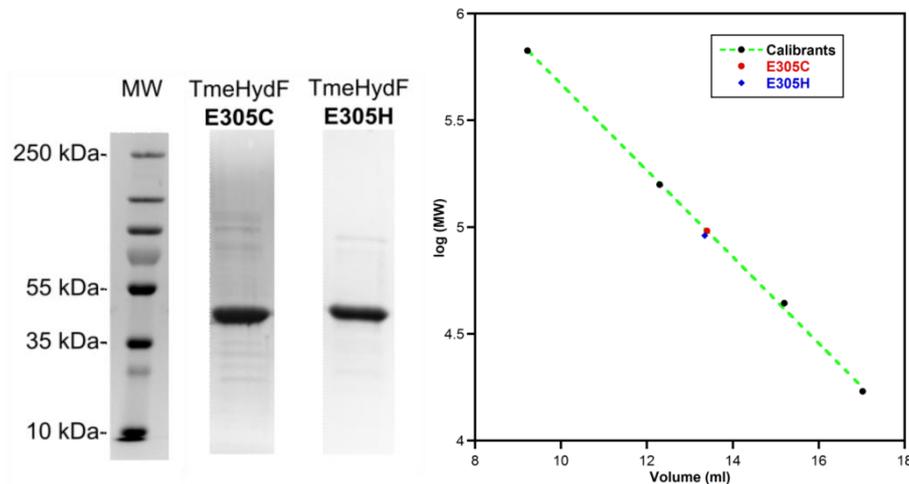


Figure 69: SDS page (left) of E305C/H TmeHydF mutants showing high protein purity (>99 %) and calibration curve using Thyroglobulin (670 kDa, 9.22 mL), γ -globulin (158 kDa, 12.32 mL), Ovalbumin (44 kDa, 15.2 mL) and Myoglobin (17 kDa, 17.03 mL). The molecular weight derived for TmeHydF mutants is 88.6 kDa, confirming the dimeric conformation for both proteins.

5.3.3 CsdA from *E. coli*.

A protocol for the purification of this protein was already reported.⁴ Cells were re-suspended in 50 mM Tris-HCl buffer pH 8.0 supplemented with 150 mM NaCl (buffer A) and were disrupted by discontinuous sonication in a water/ice bath. After ultracentrifugation at 196.000 g for 1h at 4 °C, soluble protein was loaded onto a nickel Hi-trap column.

Materials and Methods

The column was subsequently washed with buffer A supplemented with 10 mM imidazole to remove nonspecifically bound *E. coli* proteins. The desired hexahistidine tagged protein was eluted from the Ni-NTA column with a linear imidazole gradient from 10 to 500 mM. Eluted fractions were analyzed by SDS-PAGE (Figure 70) and imidazole was removed by a desalting column equilibrated with 50 mM Tris-HCl buffer pH 8.0, 150 mM NaCl. An SDS page confirmed the purity of the target protein (> 99 %).

A I B I MW Fractions collected

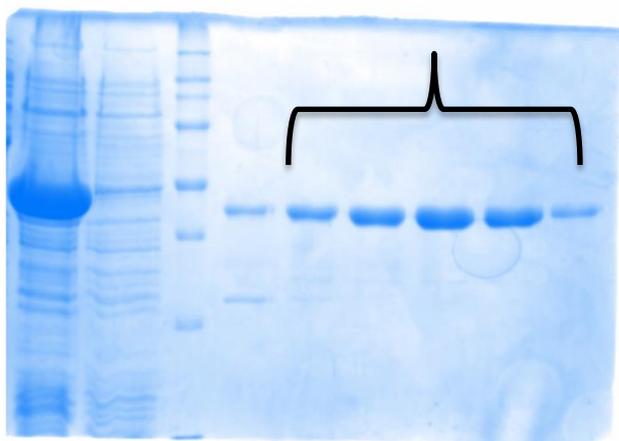


Figure 70: SDS page of CsdA protein after purification showing high purity (>99 %). In the first and second lines are reported the protein expression before and after the addition of IPTG.

Yellow protein fractions (due to the PLP cofactor) were concentrated to 1.2 mM, aliquoted and stored at -80 °C.

5.3.4 MeHydA: the [FeFe]-hydrogenase from *Megasphaera elsdenii*

The [FeFe]-hydrogenase from *Megasphaera elsdenii* possess a high number of Cys residues due to the presence of three [4Fe-4S] clusters each of them coordinated by four cysteines. Thus to avoid S-S bridges during the aerobic purification, reducing agents were added to the purification buffers. Indeed, some preliminary essays conducted in the absence of DTT and β -mercaptoethanol showed the presence of large oligomers and only a small (5-10%) monomeric fraction (Figure 71) once the protein was analyzed by size exclusion chromatography on a Superdex S200 10/300 in 50 mM Tris-HCl pH 8.0, 300 mM NaCl.

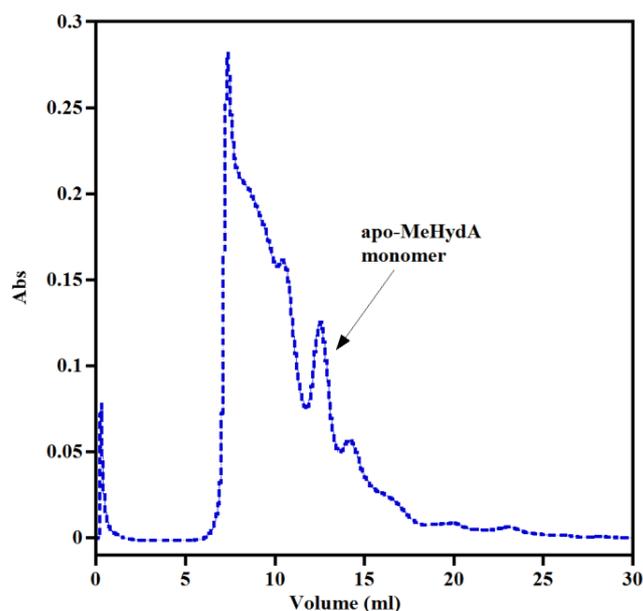


Figure 71: SEC profile of MeHydA on a Superdex S 200 10/300 in 50 mM Tris–HCl pH 8.0, 300 mM NaCl after a Ni-NTA column without reducing agents.

The optimized procedure for MeHydA purification leading to highest yield and most homogenous samples is the following. Cells were re-suspended in Tris–HCl buffer 50 mM pH 8.0 containing 5 % v/v glycerol, 1 % v/v Triton, 1 mM dithiothreitol (DTT), 5mM β -mercaptoethanol (β -met) and discontinuously sonicated for 10-12 min in a water/ice bath. Cellular extracts were centrifuged 1h at 193.000 g leading to a soluble fraction of MeHydA protein.

The first chromatographic step was performed on a His Trap column (GE-healthcare, resistant to reducing agents) equilibrated with 50 mM Tris–HCl pH 8.0, 300 mM NaCl, 10 % v/v glycerol, 5 mM β -met, 1 mM DTT (buffer A). After loading, the column was extensively washed ten column volumes with 50 mM Tris–HCl pH 8.0, 300 mM NaCl, 10 % v/v glycerol, 5 mM β -met, 1 mM DTT, 10 mM imidazole (to remove nonspecifically bound *E. coli* proteins) then eluted with a linear gradient of buffer A/buffer A supplemented with 500 mM imidazole (Figure 72).

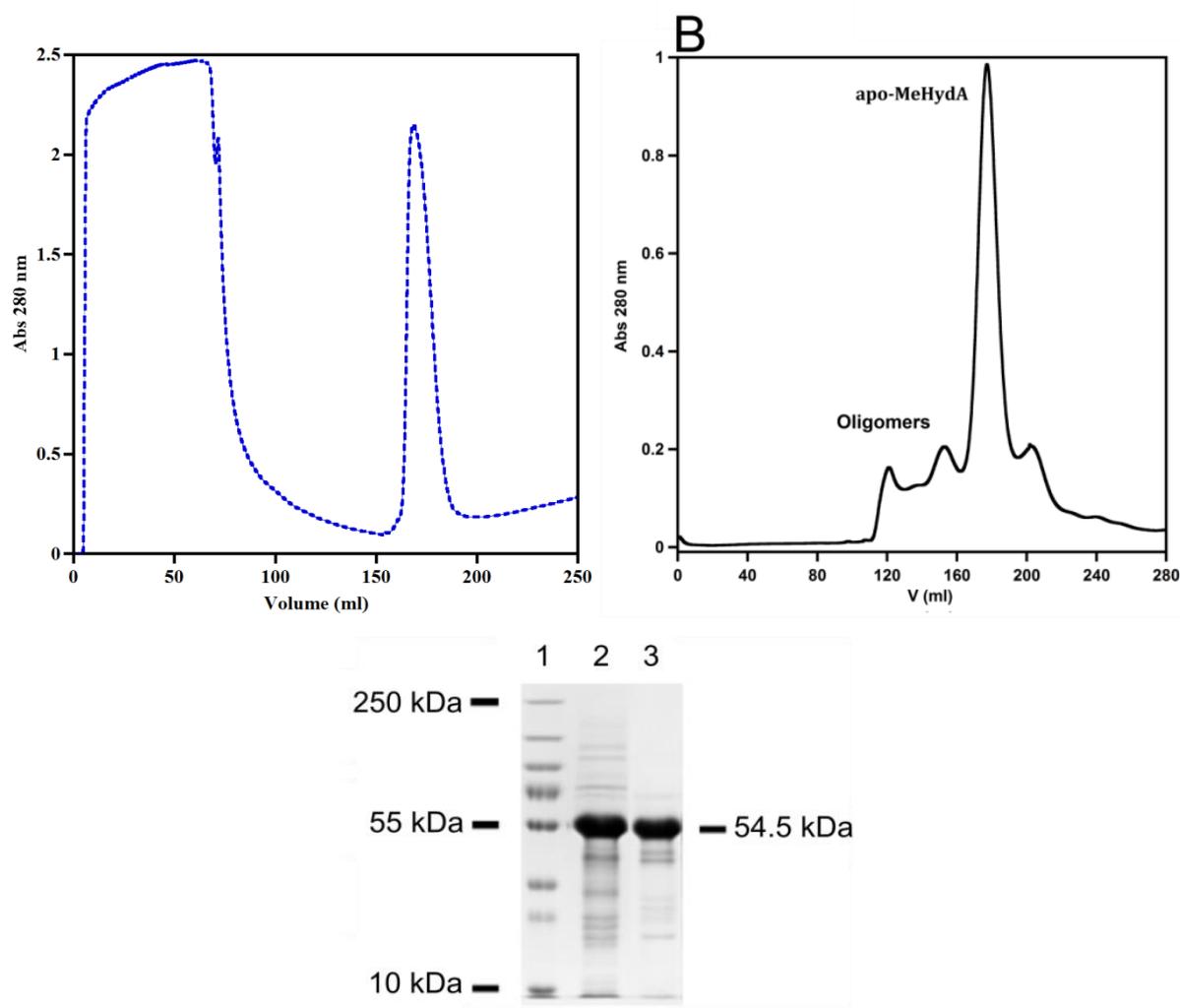


Figure 72: His Trap chromatographic profile of MeHydA (top left), SEC profile of MeHydA (top right) on a Superdex S200 26/600 in 50 mM Tris–HCl pH 8.0, 300 mM NaCl, 10 % v/v glycerol, 5 mM DTT and SDS page of MeHydA after Ni-NTA (2) and SEC (3) purifications.

Elution fractions containing the target MeHydA were pooled and a pure and homogeneous protein was obtained after a gel filtration step (Figure 72) with a Superdex S200 26/600 equilibrated in 50 mM Tris–HCl pH 8.0, 300 mM NaCl, 10 % v/v glycerol, 5 mM DTT.

SDS page proved the purity of MeHydA protein samples. The oligomeric state of the brownish protein was determined in anaerobic conditions via analytical gel filtration using a Superdex S200 10/300 column equilibrated in 50 mM Tris–HCl buffer pH 8.0, 300 mM NaCl, 10% v/v glycerol, 5 mM DTT (see chapter II, paragraph 2.3 for details about characterization of apo-MeHydA protein by SLS and SEC analysis).

5.3.5 MeH-HydA: the [FeFe]-hydrogenase truncated form of *Megasphaera elsdenii*

Cells were re-suspended in Tris-HCl buffer 50 mM pH 8.0, containing 200 mM NaCl and discontinuously sonicated for 10-12 min in a water/ice bath. Cellular extracts were centrifuged 75 min at 193.000 g leading to a soluble fraction of MeH-HydA.

The first chromatographic step was performed on a Hi Trap column (GE-healthcare) equilibrated with 50 mM Tris-HCl pH 8.0, 200 mM NaCl (buffer A). After loading, the column was extensively washed with 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 10 mM imidazole and then eluted with a linear gradient of buffer A/buffer A supplemented with 500 mM imidazole. Elution fractions containing apo-MeH-HydA were pooled and a pure and homogeneous protein was obtained after a gel filtration step with a Superdex S200 16-600 equilibrated in 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM DTT (Chapter III, paragraph 3.2.2 Figure 34c). The apo-protein has been concentrated to 5-10 mg/ml, frozen in liquid nitrogen and stored at -80 °C until used.

The oligomeric state of the protein was determined in anaerobic conditions via analytical gel filtration using a Superdex S200 10/300 column equilibrated in 50 mM Tris-HCl buffer pH 8.0, 300 mM NaCl, 10% v/v glycerol, 5 mM DTT.

5.4 Biochemical methods: [Fe-S] cluster reconstitution of apo-proteins

All the proteins studied in this PhD contain [Fe-S] clusters, which are damaged and destroyed due to the aerobic purifications. Thus once purified, they contain some negligible iron.

The [Fe-S] cluster(s) reconstitution was conducted in anaerobic condition inside a glove box ($O_2 < 0.5$ ppm) under nitrogen atmosphere. The apo-proteins, normally 50-100 μ M, have been incubated with 10 mM DTT for 15 min at 20°C, followed by an excess of ferrous ammonium sulfate $[(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O]$ and L-cysteine. The excess of the two reactants depended on the number of [Fe-S] clusters that had to be reconstituted.

MeHydA has three [4Fe-4S] cores and a 17-18 molar excess was found to be essential to reconstitute completely all the three clusters. All the other apo-proteins studied are

characterized by a single [4Fe–4S] cubane and a 5-5.5 molar excess of Fe and L-cys was sufficient for the assembly of the metal cofactor.

The reaction is normally initiated by the addition of a catalytic amount of the *E. coli* cysteine desulfurase CsdA (1–2 % molar equivalent) and the protein solution is monitored by recording UV–Vis absorption spectra every 20-30 min on a Cary 100 spectrophotometer (Agilent) connected to the cell holder located in a glovebox with optical fibers. All the reactions were normally conducted overnight. The reconstituted proteins were then centrifuged 30 min at 12000 rpm and purified on Superdex S200 10/300 GL equilibrated with the reconstitution buffer (supplemented with 5 mM DTT). The pure and homogenous FeS-proteins were concentrated with Amicon Ultra 30-kDa centrifugal filters and stored in liquid nitrogen or at -80 °C in sealed vials until used.

5.4.1 Iron-sulfur reconstitution with ^{57}Fe

Mössbauer spectroscopy requires samples selectively labelled with ^{57}Fe . This sample is exposed to a beam of gamma radiation and a detector measures the intensity of the beam transmitted through the sample. The atoms in the source emitting the gamma rays must be of the same isotope as the atoms in the sample absorbing them. For that reason the reconstitution of the [Fe–S] centers was conducted with a ^{57}Fe source.

The $^{57}\text{Fe}(\text{SO}_4)$ replaced Mohr salt and the reconstitution was executed in the same conditions. $^{57}\text{Fe}(\text{SO}_4)$ has been provided by *Laboratoire de Chimie et Biologie des Métaux* in Grenoble as 0.62 M solution in sulfuric acid and it has been stored in glove box. Before starting the reconstitution the ^{57}Fe solution was diluted to 0.023 M in 1.0 M Tris–HCl buffer pH 8.0 to basify the solution. The reconstituted protein was then purified with normal procedures described above.

The protein sample was then concentrated to 0.6-0.7 mM, placed in a Mössbauer cylindrical chamber and frozen inside the glove box with liquid iso-pentane/liquid nitrogen.

5.4.2 [Fe–S] protein preparation for EPR and HYSCORE measurements

5.4.2.1 *TmeHydF*, *E305C* and *E305H* mutants

In order to confirm the presence of the [4Fe-4S] cluster onto TmeHydF, 100 μ l of 0.3-0.4 mM of freshly reconstituted FeS-TmeHydF were incubated for 30 minutes with a 10 mM sodium dithionite solution prepared in 25 mM Hepes pH 7.5, 200 mM NaCl. As a blank, 100 μ l of the same protein concentration without sodium dithionite were used to measure the as isolated TmeHydF sample. The same protocol was used for E305C/H mutants. The proteins were then placed in EPR tubes and flash frozen in liquid nitrogen.

The preparation of the FeS-TmeHydF for HYSCORE analysis was slightly different. 100 μ l of 1.0 mM FeS-TmeHydF were incubated with 25-fold excess of imidazole solution for 30 minutes and then treated with a 15 mM sodium dithionite solution in 25 mM Hepes buffer pH 7.5, 200 mM NaCl. To prove the absence of the “imidazole-like” binding to the [Fe-S] cluster of the wild type protein, 100 μ l of 1.0 mM FeS-TmeHydF was incubated with 15 mM sodium dithionite. The same procedure was used to prepare samples of E305C/H mutants.

5.4.2.2 *MeHydA and MeH-HydA*

In order to prove the correct assembly of metal cofactors inside FeS-MeHydA and FeS-MeH-HydA proteins Q-band EPR spectroscopy was used.

100 μ l of 0.2-0.3 mM freshly reconstituted samples were incubated for 30 minutes with a 10 mM sodium dithionite solution and then flash frozen in liquid nitrogen. As isolated iron-sulfur proteins (without sodium dithionite) were also analyzed (blank) to prove correct assembly of the [4Fe-4S] clusters.

5.5 Chemical methods: Insertion of $[\text{Fe}_2(\text{adt/pdt})(\text{CO})_4(\text{CN})_2]^{2-}$ onto HydF and HydA

5.5.1 Synthesis of $(\text{Et}_4\text{N})_2[\text{Fe}_2(\text{adt/pdt})(\text{CO})_4(\text{CN})_2]$ complexes

This project has been carried out in collaboration with the Solhycat research group (Dr. V. Artero) that provided the synthetic complexes to maturate HydA and HydF proteins. The synthetic complexes, $(\text{Et}_4\text{N})_2[\text{Fe}_2(\text{adt/pdt})(\text{CO})_4(\text{CN})_2]$, were prepared as described in literature, dissolved in DMSO (80 mM stock solution) and stored at -80 °C in sealed vials

until used. Before incubation with protein samples they have been diluted in the protein buffer to prevent denaturation of samples due to the high DMSO concentration.

5.5.2 Insertion of $[\text{Fe}_2(\text{adt/pdt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex onto MeHydA

In a standard experiment, the protein solution was extensively washed to remove any DTT traces, to avoid DTT- dependent decomposition of the organometallic chemicals, and buffer exchanged to 100 mM potassium phosphate pH 6.8.

50-100 μM FeS-MeHydA was incubated for 1h with 10 molar equivalents of $[\text{Fe}_2(\text{adt/pdt})(\text{CO})_4(\text{CN})_2]^{2-}$ biomimetic complex (called **1**=adt and **2**=pdt in this thesis work) in 100 mM potassium phosphate pH 6.8 at room temperature (19-21 $^\circ\text{C}$).

The excess of the biomimetic complexes, **1** or **2**, was removed by a desalting column (NAP-10, GE healthcare) and the protein was concentrated to 0.5-1 mM and stored in liquid N_2 or directly used for spectroscopic and activity measurements.

Once the protein has been matured, different redox states were isolated and analyzed through EPR and FTIR spectroscopies.

5.5.2.1 H_{ox} state

50 μl solution of as isolated Holo-MeHydA (0.5-1 mM) in 100 mM potassium phosphate pH 6.8 has been flushed 1-2 h with argon gas in order to oxidize the $[\text{2Fe}]$ -subcluster of the H-cluster. 10 μl were used for FTIR characterization. 40-45 μl were needed for the Q-band EPR measurement.

5.5.2.2 $H_{\text{ox}}\text{-CO}$ inhibited state

50 μl solution of as isolated Holo-MeHydA (0.5-1 mM) in 100 mM potassium phosphate pH 6.8 has been flushed 20 minutes with carbon monoxide to obtain a pure $H_{\text{ox}}\text{-CO}$ inhibited state of the H-cluster. 10 μl were used for FTIR characterization. 40-45 μl were needed for the Q-band EPR measurement.

5.5.2.3 H_{red} state

50 μ l solution of as isolated Holo-MeHydA (0.5-1 mM) in 100 mM potassium phosphate pH 6.8 was flushed with H_2 for 24 h and subsequently treated with different concentration of sodium dithionite. Indeed, MeHydA could not be reduced completely with H_2 and a sodium dithionite treatment was required. 10 μ l were needed for each FTIR spectrum.

5.5.3 Insertion of $[Fe_2(adt)(CO)_4(CN)_2]^{2-}$ complex onto MeH-HydA

Initial attempts with standard chemical maturation conditions (100 mM potassium phosphate pH 6.8) as described in paragraph 5.5.2 led to small incorporation of the synthetic biomimetic complex. The optimized conditions are described in the following.

FeS-MeH-HydA reconstituted in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 % v/v glycerol, 5 mM DTT was intensively washed with 100 mM potassium phosphate pH 6.0. 50-100 μ M protein solution was incubated for 2h with 10 molar equivalents of $[Fe_2(adt)(CO)_4(CN)_2]^{2-}$ complex. The excess of the chemical was removed by a desalting column (NAP 10) and the protein was concentrated using Amicon 30-kDa centrifugal filters up to 1.0 mM. For spectroscopic measurements the protein buffer has been exchanged to 100 mM potassium phosphate pH 6.8 to direct compare MeHydA and MeH-HydA proteins.

5.5.3.1 H_{ox} , H_{ox} -CO and H_{red} states

For the preparation of different redox states, similar procedures, illustrated in paragraphs 5.5.2.1, 5.5.2.2 and 5.5.2.3 were used.

5.5.4 Insertion of $[Fe_2(adt/pdt)(CO)_4(CN)_2]^{2-}$ complex onto HydF proteins

FeS-HydF proteins from *Thermotoga maritima* and *Thermosipho melanesiensis* were washed with 25 mM Hepes Buffer pH 7.5, 200 mM NaCl to remove DTT traces from the

protein solution. In a standard experiment 50-100 μM protein solutions of freshly reconstituted HydFs were incubated for 1h with 10 molar equivalents of $[\text{Fe}_2(\text{adt/pdt})(\text{CO})_4(\text{CN})_2]^{2-}$ complex in 25 mM HEPES buffer pH 7.5 containing 200 mM NaCl.

The excess of the chemicals, $[\text{Fe}_2(\text{adt/pdt})(\text{CO})_4(\text{CN})_2]^{2-}$, was removed by a desalting column (NAP-10, GE healthcare), then the proteins were concentrated to 0.5-1 mM and stored in sealed vials or directly used for spectroscopic characterizations, hydrogenase maturation and activity essays.

5.6 Hydrogenase-like activity

5.6.1 Calibration curve for H_2 quantitation

Using a Gas chromatographic system, a calibration curve with pure hydrogen has been realized. Different concentrations of pure H_2 (Linde) or a mixture of H_2/N_2 have been injected on the GC. The column temperature has been setup at 50°C with a flowrate of nitrogen gas carrier of 5 ml/min. Plotting the area of known hydrogen concentration as a function of the amount of hydrogen injected, a linear fit allowed the determination of the slope ($5.8091 \cdot 10^{11}$, blue trace in Figure 73) that was used to quantify the unknown amount of hydrogen produced by HydA enzymes. The calibration has been repeated every 6 months with a daily check performed using a known hydrogen concentration.

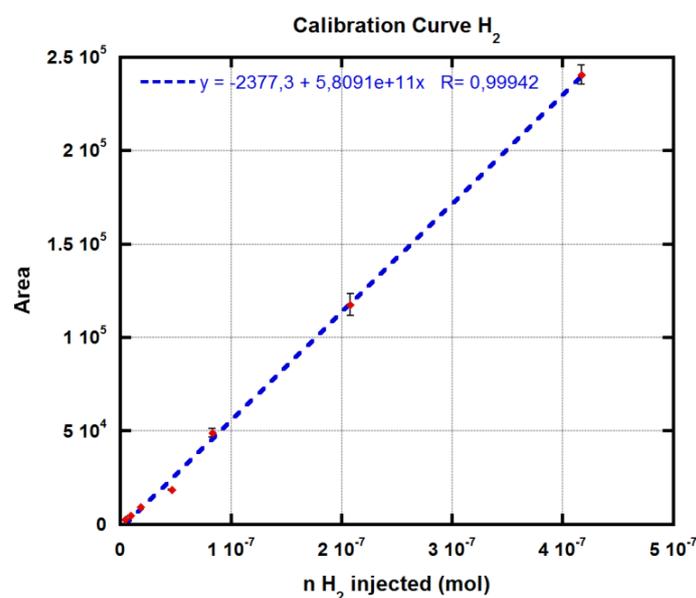


Figure 73: Calibration curve on a Shimadzu GC-2014 system for H_2 quantification

5.6.2 H₂ detection via Gas Chromatography of MeHydA and MeH-HydA proteins

H₂ production has been determined using a well-known procedure consisting in using methyl viologen (MV) as electron mediator and sodium dithionite (NaDT) as reducing agent.⁵ This methodology is the most commonly used method for the evaluation of hydrogenase activities, reporting the specific activity as $\mu\text{mol H}_2$ per minute per mg of enzyme.

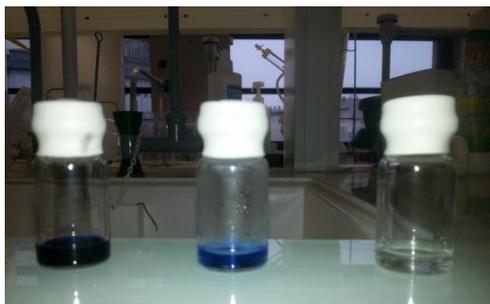


Figure 74: Sealed vials containing hydrogenase protein solution in the presence of 100 mM potassium phosphate pH 6.8, 100 mM sodium dithionite and 10 mM methyl viologen (MV). The violet color is due to MV which is reduced to the violet $\text{MV}^{\bullet+}$ in the presence of sodium dithionite. Over time sodium dithionite, which provides electrons, is consumed and the solution becomes colorless.

Briefly, holo-MeHydA (5–10 μL) corresponding to 20 pmol of active hydrogenase was added to a total amount of 1.11 mL of 100 mM potassium phosphate pH 6.8, 100 mM sodium dithionite, and 10 mM methyl viologen in a 10 mL vial sealed (Figure 74) under anaerobic conditions (M Braun glove box with $\text{O}_2 < 0.5$ ppm). MeH-HydA was assayed in the same conditions: 3–5 μl of protein (stock solution c.a 30 μM) corresponding to 0.2 nmol were added to a total amount of 1.11 mL of 100 mM potassium phosphate pH 6.8, 100 mM sodium dithionite, and 10 mM methyl viologen in a 10 mL vial. Gas chromatograms (Figure 75) were recorded on a GC System (Shimadzu GC-2014 with a thermal conductivity detector and a Quadrex column) and the amount of H₂ was quantified using the calibration curve reported in the last paragraph. $[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ and FeS-MeHydA and FeS-MeH-HydA were assayed as controls and did not show any detectable activity. All activities in this study were measured at least 3 times on 3 different protein preparations.

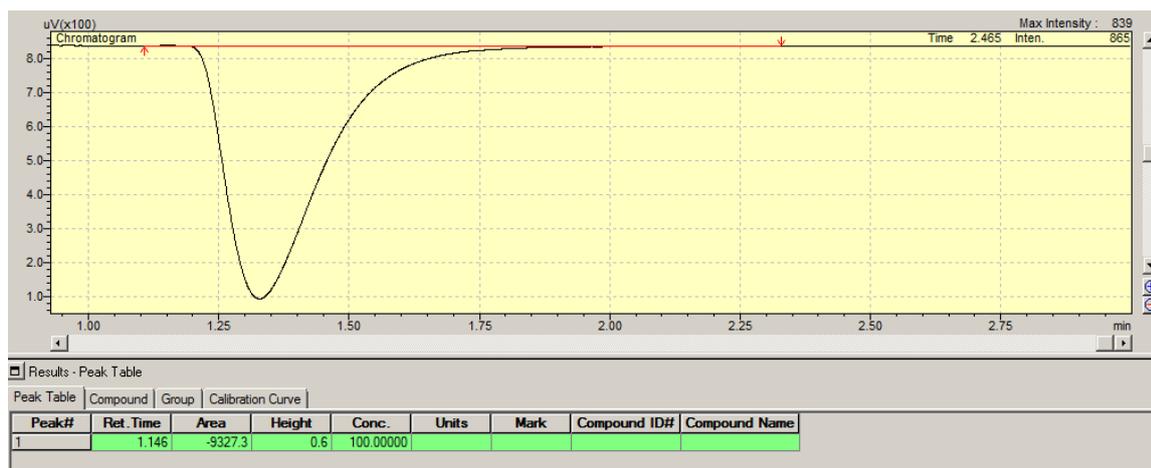


Figure 75: Gas chromatogram of hydrogen produced by MeHydA ($t_R=1.33$ min) using a nitrogen flowrate of 5 ml/min and a Temperature of 50 °C.

5.6.3 H₂ detection via a miniaturized Clark-type hydrogen sensor

Since the laboratory recently bought a miniaturized electrochemical H₂ microsensor, it was used to quantify low amount of hydrogen, normally undetectable with common GC systems. This device is a miniaturized Clark-type hydrogen sensor with an internal reference electrode and a sensing anode. The sensor must be connected to a high-sensitivity picoammeter where the anode is polarized against the internal reference (+1000 mV). Driven by the external partial pressure, hydrogen from the environment will pass through the sensor tip membrane and will be oxidized at the platinum anode surface. The picoammeter converts the resulting oxidation current to a signal.

This sensor resulted to be useful in detecting low amount of H₂, using less sample compared to the GC system, allowing direct detection of the hydrogen produced in solution by the catalyst.

The solubility of H₂ strongly depends on temperature and salinity and a calibration with pure hydrogen was done in the same experimental conditions of catalysis. The hydrogen concentration is reported in μmol/L.

5.6.3.1 Methyl viologen assay for TmHydF, TmeHydF and MeHydA proteins

Chapter V

As shown by spectroscopic analysis and Fe/S quantitation, TmHydF and TmeHydF proteins were able to bind a [4Fe-4S] cluster and diiron biomimetic complexes, called **1** and **2** ($[\text{Fe}_2(\text{adt})(\text{CO})_4(\text{CN})_2]^{2-}$ **1**=adt; **2**=pdt), generating HydF hybrids (**1**-HydF and **2**-HydF). Both hybrids were used for studies of hydrogen evolution. Interestingly, **1**-HydF from *Theromotoga maritima* and *Thermosipho melanesiensis* did not show any detectable hydrogen, although both hybrids had the biologically relevant azadithiolate bridge. **2**-HydF showed some hydrogen production in the following conditions.⁷

The activity was determined by transferring 4.6 nmol of **2**-HydF to a 1.15 mL reaction mixture containing 4 mM methyl viologen solution, as electron mediator, in 100 mM potassium phosphate buffer pH 6.0 (Figure 76). The H₂ evolution started adding 15 μl of a 400 mM sodium dithionite solution (5.2 mM). The activities of free **2** complex and apo-HydF were also evaluated transferring 4.6 nmol of each to the same reaction mixture as described above.

MeHydA matured with the synthetic complex **2** showed a higher activity; thus, the optimized non saturating conditions are the following. 70-80 μl of 2-3 μM **2**-MeHydA (corresponding to 0.173 μM) were mixed with 4 mM of methyl viologen in 100 mM potassium phosphate pH 6.8. The reaction started with the addition of 15 μl of a 400 mM sodium dithionite solution (5.2 mM).

2-MeHydA produces hydrogen with an initial TOF of $\approx 10 \text{ min}^{-1}$ comparable to previous results carried on **2**-HydA from *Chlamydomonas reinhardtii*.⁶

FeS-MeHydA was also evaluated in a similar manner transferring the same number of mol of protein into the vessel, in the presence of 4 mM of methyl viologen and 5.2 mM sodium dithionite in 100 mM potassium phosphate pH 6.8.



Figure 76: Clark-type H₂ microsensor immersed in a solution containing methyl viologen, sodium dithionite and 2-MeHydA. The blue/violet color is due to reduced methyl viologen.

5.6.3.2 Photocatalytic H₂ production assay driven by $Ru(bpy)_3^{2+}$

A light irradiation provided by a 300 W Xenon lamp (Oriel, ozone free) was used for these chemical essays. The lamp was operating at 280 W coupled with a water-filled Spectra Physics 6123NS liquid filter and a Spectra Physics 59472 UV cut-off filter ($\lambda > 400$ nm).

The performance of matured 2-TmHydF for H₂ evolution was evaluated transferring 4.6 nmol of HydF hybrid to a 1.15 mL reaction mixture containing 140 μ M $Ru(bpy)_3^{2+}$ and 100 mM ascorbate, as sacrificial electron donor, in 100 mM potassium phosphate buffer pH 6.0.⁷ The H₂ evolution started through irradiation by the xenon light source (Figure 77). The activities of free complex 2 and apo-HydF were also evaluated using the same protocol described for the matured protein.

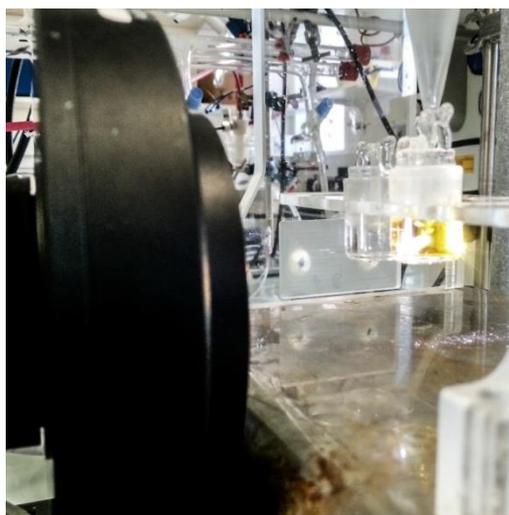


Figure 77: Clark-type H₂ microsensors immersed in a solution containing Ru(bpy)₃²⁺, sodium ascorbate as sacrificial electron donor and 2-TmHydF. The yellow color is due to ruthenium compound.

5.7 Determination of protein concentration

The Bradford protein assay is one of the most common methods used to determine the total protein concentration of a sample. The method is based on the proportional binding of the dye *Coomassie* to proteins. Within the linear range of the assay, the more protein present, the more *Coomassie* binds. The assay is colorimetric, measuring the absorption of the complex (Protein-*Coomassie*) at 595 nm. The protein concentration of an unknown sample is determined by comparison to that of a series of protein standards known to reproducibly exhibit a linear absorbance profile in this assay. Although different protein standards can be used, I have chosen the most widely used protein - Bovine Serum Albumin (BSA).

The concentration determined via Bradford is compared to that determined using the theoretical molar extinction coefficient at 280 nm ($\epsilon_{280 \text{ nm}}$). ExPASy *Bioinformatics Resource Portal* (*ProtParam* tool) was used to calculate the theoretical $\epsilon_{280 \text{ nm}}$, giving as input the protein sequence.

5.8 Determination of cofactor concentration

5.8.1 Iron quantitation: Fish method

Materials and Methods

The iron amount of iron-sulfur and chemically matured proteins was determined using the Fish method.⁸ A commercial standard solution of Fe (1000 mg/ml \pm 4 mg/ml) was diluted 36 times in H₂O in order to obtain a 500 μ M iron solution. This stock solution was used to prepare a calibration curve in the range 0-10 nmol of iron. In a similar way, proteins were assayed in the range 0.25-2 nmol depending on the expected amount of [Fe-S] clusters. For each sample, the Fe content was determined 3 times for 2 different protein concentrations.

All the samples (65 μ l) were incubated with 45 μ l of 1.0 M perchloric acid for 45 minutes in order to denature protein samples. After centrifugation (10000 rpm for 10 minutes), 90 μ l of the solution were mixed with 72 μ l of Bathophenanthrolinedisulfonic acid disodium salt (stock solution 1.7 mg/ml), 36 μ l of sodium ascorbate solution (stock solution 38 mg/ml) and 27 μ l of ammonium acetate solution (saturated solution diluted 1/3). After 30 minutes incubation (room temperature) the solution was analyzed by UV-Vis, measuring the absorption difference $Abs_{(535-680nm)}$. The iron content was determined using a calibration curve.

5.8.2 Sulfur quantitation: Beinert method

The sulfur content was determined using the Beinert method.⁹ A crystal of sodium sulfide (Na₂S*9H₂O) is dissolved in 0.03M sodium hydroxide to have a 20 mM stock solution. A 0.2 mM solution in 0.03M NaOH is prepared from the last solution and is used to prepare standards in the range 0-20 nmol of sulfur. Proteins were assayed in the range 0.5-4 nmol depending on the expected amount of [Fe-S] clusters, diluting the protein solution to 100 μ l in H₂O.

For each sample, the S content was determined 3 times for 2 different protein concentrations. The 100 μ l samples are mixed with 300 μ l of 1% w/v Zn(OAc)₂*2H₂O solution and 15 μ l of 3.0 M NaOH. This step allowed the immediate protein denaturation and precipitation of zinc hydroxide. Samples are mixed and incubated for 15 minutes at room temperature. A simultaneous addition of DMDP (75 μ l of 1mg/ml in HCl 5M) and FeCl₃ (16 μ l, 23 mM) solutions allowed the production of Methylene blue. Samples were incubated for 3h at room temperature, followed by centrifugation at 10000 rpm for 5 minutes. The sulfur content was determined by Uv-Vis measuring the absorption difference, $Abs_{(670-850 nm)}$, that corresponds to the methylene blue using a calibration curve.

5.9 Spectroscopic characterization

This paragraph contains some basic principles of the spectroscopic techniques used to characterize the [Fe-S] cluster proteins studied in this PhD.

5.9.1 UV-Visible spectroscopy

UV-Visible spectroscopy was used to measure protein and cofactor concentrations as described in paragraphs 5.7 and 5.8. Moreover, it was useful for monitoring the [Fe-S] cluster(s) assembly inside apo-protein scaffolds. Once reconstituted, FeS-proteins have a typical absorption band around the 400-410 nm. This LMCT transition is due to sulfur to iron charge transfer.

The absorption ratio, $Abs_{400/280}$, is also used to estimate the [Fe-S] cluster incorporation. HydF and MeH-HydA, with one [4Fe-4S] core, have typical absorption ratios varying from 0.21 to 0.24. While MeHydA with three [4Fe-4S] cluster has a $Abs_{400/280}$ of 0.36.

The laboratory is equipped with a Cary 100 spectrophotometer (Agilent) connected to the cell holder located in a glove box with optical fibers.

5.9.2 EPR spectroscopy

Electronic Paramagnetic Resonance spectroscopy can be applied to any paramagnetic specie. This includes a wide range of samples (organic and inorganic compounds), including complex biomolecules. This technique is used to identify molecules within a sample and characterize the environment of the unpaired electron.

The EPR spectroscopy studies the interaction of unpaired electrons in the sample with a magnetic field; this effect is called the Zeeman Effect. The magnetic field B_0 produces two energy levels for the magnetic moment of the electron ($\vec{\mu}$). The unpaired electron will have a state of lowest energy in which the moment of the electron ($\vec{\mu}$) is aligned with the magnetic field and a level of highest energy where $\vec{\mu}$ is aligned against the magnetic field. The energy of each orientation is the product of μ and B_0 and because the electron is a spin $1/2$ particle, the energy of each orientation (Figure 78) is given by the following equation, $\vec{\mu} = m_s g_e \beta$. Where m_s is $\pm 1/2$ and depends on the orientation of the electron moment, β is the Bohr magneton, g_e the g-factor of the electron (≈ 2.00). Therefore, the energies for an electron are given by:

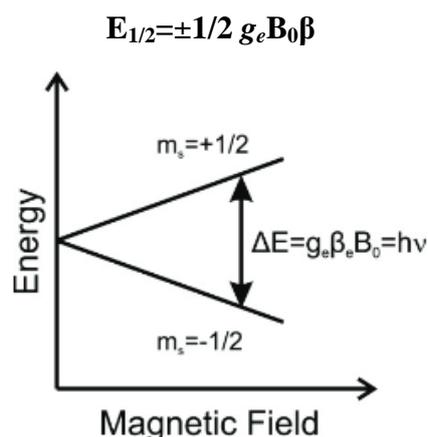


Figure 78: Splitting of electron spin states

When the electron is inside a molecule, there is also the spin-orbit coupling to consider that causes shifts in an electron's atomic energy levels due to electromagnetic interaction between the electron's spin and the magnetic field generated by the electron's orbit around the nucleus. The total angular momentum is the contribution of the intrinsic spin angular momentum (\vec{S}) and the orbital angular momentum (\vec{L}).

Thus, the equation becomes $\vec{\mu} \propto g_e \vec{S} + \vec{L}$

It's possible to consider that the spin-orbit contribution is proportional to \vec{S} and summarize with the following equation $\vec{\mu} \propto g \vec{S}$, where g is the combination of the g_e parameter with the spin-orbit coupling contribution.

During an EPR experiment, a sample with unpaired electrons is placed in a magnetic field that can split the energy levels of the ground state by a quantity given by the following equation:

$$\Delta E = h\nu = g B_0 \beta,$$

Knowing the energetic difference through irradiation of the sample with microwaves with a set frequency and sweeping the magnetic field, the strength of the magnetic field and the Bohr magneton (constant) is possible to calculate the g factor:

$$g = h\nu / \beta B_0$$

The g factor can be considered as a fingerprint of the molecule, containing the chemical information on the nature of the bond between the electron and the molecule.

Chapter V

The intensity of the EPR signal is closely linked to the temperature of the sample and the power of microwaves entering the microwave cavity. These two parameters must be considered and optimized together to maximize signal intensity. The temperature dependence of the EPR signals follows the *Curie's law* (valid for all paramagnetic samples) according to which, the magnitude of the signal tends to be inversely proportional to the temperature ($1/T$, expressed in Kelvin). EPR signal intensity also tends to be proportional to the microwave power, $\sqrt{P_{mw}}$. A compromise between the T and the P_{mw} has to be found.

Molecules with fully filled electronic shells are EPR silent; thus EPR spectroscopy is particularly suitable for the study of proteins prosthetic groups in solution that have a non-null electronic spin. It resulted particularly useful for metalloprotein studies, nevertheless not all metallic ions and redox states are suitable for study by EPR. In the particular case of iron, manganese and cobalt containing systems, EPR spectroscopy is particularly useful due to the fact that these metals have redox states with unpaired electrons. Among iron containing proteins, iron-sulfur clusters represent some of the most efficient and abundant electron-transfer groups and historically they were first detected using EPR spectroscopy. They have a wide range of functions, including enzyme catalysis, regulation, structural, scaffold systems but the most frequent activity is in biological electron transfer.

The most commonly occurring clusters, described in chapter I (section A) are the [2Fe–2S], [3Fe–4S] and [4Fe–4S] types. In general each cluster is able to transfer one electron changing its redox state, so can be EPR active ($S = 1/2$) in at least one state, either oxidized or reduced.

Their EPR spectra are usually observed at cryogenic temperatures ($T < 100$ K) and typically have a rhombic or near-axial lineshape. Cryogenic temperatures are reached using a cryostat; liquid and gaseous helium will flow through the system, reaching low temperature up to the 4.2 K limit.

EPR can provide information about the structure and biological function of the FeS centers, including:

- the type of [Fe–S] cluster ([2Fe–2S], [3Fe–4S], [4Fe–4S], etc.);
- the electronic structures thanks to observation of hyperfine interactions with stable isotopes nuclei such as ^{57}Fe and ^{33}S ;
- the redox states of the clusters;
- the orientation of the clusters (spin–spin interactions which can be observed as increased spin relaxation rates, and/or splitting of EPR signals).

Materials and Methods

The [4Fe–4S] clusters of most proteins have a ground state $S = 0$ in the oxidized ($[4\text{Fe–4S}]^{2+}$) state, and are reduced by one electron to a $[4\text{Fe–4S}]^+$ state with $S = 1/2$. The less common *high-potential-iron-sulfur proteins* (HiPIPs) contain a similar $[4\text{Fe–4S}]^{2+}$ cluster, but they can be oxidized to a $[4\text{Fe–4S}]^{3+}$ state ($S = 1/2$).

In Table 10 are summarized typical [Fe–S] clusters and their oxidation states studied by EPR technique and in black the types of clusters that we analyzed in this PhD work.

Table 10: Typical [Fe–S] clusters and their redox states. In bold are shown the clusters studied in this PhD work.

Type of cluster	EPR silent	EPR active
[2Fe–2S]	+2 (oxidized)	+1 (reduced)
[2Fe–2S] _{RIESKE type}	+2 (oxidized)	+1 (reduced)
[3Fe–4S]	0 (reduced)	+1 (oxidized)
[4Fe–4S]	+2 (oxidized)	+1 (reduced)
[4Fe–4S] _{HiPIP}	+2 (reduced)	+3 (oxidized)

Pulsed EPR was used to study the [4Fe–4S] clusters of MeHydA, MeH-HydA, TmeHydF and TmeHydFE305C/H mutants.

In a Cw-EPR experiment the sample is continuously illuminated with microwave radiation at a fixed frequency and the strength of the applied magnetic field is swept, observing changes in microwave absorption. Cw-EPR is the most basic experiment performed.

In pulsed experiments microwaves are applied to the sample as a series of nanoseconds-long pulses. The magnetic field strength is kept constant, and the sample is pulsed with microwave energy. The parameters of the pulse pattern can be changed and utilized in such a way that the hidden features of a transition observed at specific field strength in a Cw-EPR spectrum can be observed.

While Cw-EPR shows the transition between one spin state to another. Pulsed EPR spectra have the capability of showing the transitions that occur within a spin state, which are caused by the interaction of the unpaired electron with the surrounding nuclei. In short, pulse EPR allows the observation of transitions within the transitions seen in Cw spectra.

This part of the project was done in collaboration with the group of Professor W. Lubitz at the Max Planck Institute in Mülheim (Germany). They have a unique expertise in the characterization of FeS proteins with particular focus to [FeFe]-hydrogenases. Samples for EPR analysis were prepared inside a M Braun glove box ($[\text{O}_2] < 0.5$ ppm) using freshly

reconstituted/matured samples in the range of 0.2-1.0 mM. They are placed in EPR tubes and flash frozen in liquid nitrogen or frozen inside the glove box using liquid iso-pentane cooled with liquid nitrogen.

Q-band EPR spectra were recorded using free induction decay (FID) detected EPR with a microwave pulse length of 1 μ s. All pulse experiments were performed on a Bruker ELEXSYS E580 Q-band spectrometer with a SuperQ-FT microwave bridge and home built resonator. Cryogenic temperatures (10-20 K) were obtained by an Oxford CF935 flow cryostat.

Field swept X-band EPR studies were performed on a Bruker Elexsys II E580 FT-EPR spectrometer equipped with an ASE 1kW TWT amplifier. Spectra were recorded in the pulsed mode using 2-pulse echo detection with $\pi/2$ pulse =16ns and τ =200ns. After a pseudomodulation transformation, the spectra obtained in this way are comparable to those using Cw-EPR. The 30 μ l samples were contained in 2.8 mm outer diameter tubes accommodated in a 3 mm Bruker (MS3) split-ring resonator.

All the experimental measurements and simulations have been done by Dr. Agnieszka Adamska-Venkatesh and Dr. E. Reijerse.

5.9.3 HYSCORE spectroscopy

In addition to cysteine residues as the most favored ligand and providing sulfur coordination, oxygen-based (aspartate, tyrosinate, glutamate) and nitrogen-based (histidine, arginine) residues have also been observed as ligands to the clusters.

Spectroscopic analysis of reduced paramagnetic forms of the clusters, in particular electron paramagnetic resonance (EPR), electron-nuclear double resonance (ENDOR), and Hyperfine Sublevel Correlation spectroscopy (HYSCORE), has been successfully used to identify the ligands to the iron atoms, thus providing important insights into the electronic and geometric properties of the clusters as well as into the mechanisms of a number of reactions catalyzed by those clusters.

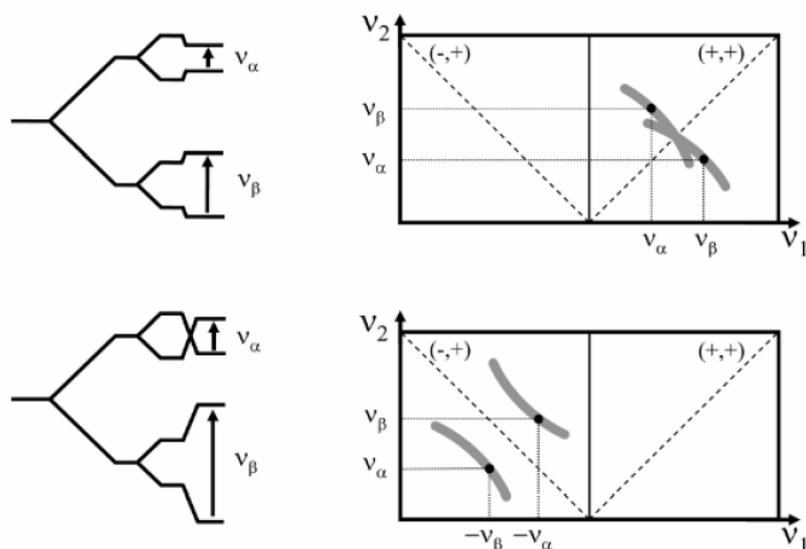


Figure 79: $I = 1/2$, $S = 1/2$ system. Upper panel represents the weak coupling case with its energy level diagram. Lower panel indicates the strong coupling case and its energy level diagram

HYSCORE spectroscopy, Hyperfine Sublevel Correlation, is a two-dimensional pulsed EPR technique that allows the detection of hyperfine coupling (between the electron spin and the nuclear spin) of nuclei with a low gyromagnetic moment. It allows the study of the hyperfine coupling between a $[4\text{Fe}4\text{S}]^+$ cluster with species having nuclear spin $\neq 0$ (^1H , ^2H , ^{13}C , ^{14}N , ^{15}N , ^{33}S). This technique resulted to be extremely useful in the study of the exchangeability of the 4th ligand of the $[4\text{Fe}4\text{S}]$ cluster into HydF proteins.

One of the main advantages of this spectroscopic technique is the ability to sort three types of nuclei: the strongly, weakly coupled and the “distant” nuclei characterized by very low hyperfine constants. In the latter case the corresponding peaks lie on the diagonal of the $(+, +)$ quadrant, whereas the strongly and weakly coupled nuclei appear off the diagonal in the $(-, +)$ and $(+, +)$ quadrants respectively (Figure 79). Moreover, with a strongly coupled $I=1$ nucleus such as ^{14}N , it is possible to observe a characteristic pattern in the $(-, +)$ quadrant with so-called double quanta-double quanta (dqdq) correlation features.

Hyperfine sublevel correlation spectroscopy (HYSCORE) experiments were performed to extract the ^{14}N hyperfine interactions using the standard HYSCORE pulse sequence: $[\pi/2]-t-[\pi/2]-t_1-[\pi]-t_2-[\pi/2]-t-(\text{ESE})$. The length of the microwave $[\pi/2]$ pulse was set to 16ns. The delay between the first two pulses (τ) was adjusted to 136-140 ns in order to suppress the amplitude of the ^1H correlation ridges. The starting t_1 and t_2 delays in all measurements were 100 ns. To cancel the effect of unwanted echoes, a four step phase cycling of the microwave pulses was used.

5.9.4 FTIR spectroscopy

Infrared (IR) spectroscopy is a very common tool used to identify and describe chemical compounds. It takes advantage of the fact that the radiation in the IR range has a frequency close to the frequency of molecular vibrations which are characteristic for the structure of the molecule.

An infrared spectrum represents a fingerprint of a sample with absorption peaks which correspond to the frequencies of vibrations between the bonds of the atoms making up the material. Because each different material is a unique combination of atoms, no two compounds produce the exact same infrared spectrum. If the frequency of the absorbed radiation matches the transition energy of the molecule, the radiation is adsorbed and oscillating electric dipole is created. A permanent dipole is not necessary, as the rule requires only a change in dipole moment ($\Delta\mu \neq 0$). The principle of the Fourier Transform Infrared experimental techniques is based on the Michelson interferometer. Fourier Transform Infrared (FT-IR) spectrometry was developed in order to overcome the IR limitations. The main difficulty was the slow scanning process. A method for measuring all of the infrared frequencies simultaneously, rather than individually, was found.

A solution was developed which employed a very simple optical device called an interferometer. The interferometer produces a unique type of signal which has all of the infrared frequencies “encoded” into it. The signal can be measured very quickly, usually on the order of one second or so. Thus, the time element per sample is reduced to a matter of a few seconds rather than several minutes. Most interferometers use a beamsplitter which takes the incoming infrared beam and divides it into two optical beams. One beam reflects off a flat mirror which is fixed in place. The other beam reflects off a flat mirror which is on a mechanism which allows this mirror to move a very short distance away from the beamsplitter. The two beams reflect off of their respective mirrors and are recombined when they meet back at the beamsplitter. Because the path that one beam travels is a fixed length and the other is constantly changing as its mirror moves, the signal which exits the interferometer is the result of these two beams “interfering” with each other. The detector records the signal for each mirror position (as a function of retardation) and a digital plot of light intensity versus the mirror position called *interferogram* is produced.

In some metalloproteins, like [FeFe]-hydrogenases, the active sites present particular ligands (CO and CN groups) that are easily detectable in the infrared. In such case, the FTIR

can be considered as a unique tool for the characterization of the structure-function relationship of these enzymes. Moreover these enzymes are characterized by different redox states (H_{ox} , H_{red} , H_{sred}) and a combination of FTIR, EPR, and spectroelectrochemical studies provide a lot of information about their mechanism.

In order to study the active sites of HydA and HydF proteins, we used *Fourier Transformed Infrared* (FTIR) spectroscopy for characterization. Indeed, FTIR is appropriate for identifying the CO (1800-2020 cm^{-1}) and CN (~ 2090 cm^{-1}) vibrations associated with the CO and CN⁻ ligands present in the 2Fe-subcluster or related mimics. These signals can be clearly distinguished from those of the $[Fe_2(adtpdt)(CO)_4(CN)_2]^{2-}$ complex in solution which are much broader. The narrow CO and CN FTIR bands of the H-cluster allow differentiating between the different redox states usually coexisting in HydA preparations. Moreover, the comparison of the FTIR bands of the HydF hybrids with the native cofactor could allow the identification of the native precursor, dealt in this PhD work.

5.9.5 Mössbauer spectroscopy

Mössbauer spectroscopy is widely used to study the nuclear structure with the absorption and emission of γ -rays. A sample is exposed to a beam of γ radiation, and a detector measures the intensity of the beam transmitted through the sample. The atoms in the source emitting the gamma rays must be of the same isotope as the atoms in the sample to analyze.

^{57}Fe has is the perfect isotope for Mössbauer spectroscopy and the γ radiation is produced by decaying ^{57}Co nuclei. The technique uses a combination of the *Mössbauer effect* and *Doppler shifts* to probe the hyperfine transitions between the excited and ground states of a nucleus. If the emitting and absorbing nuclei were in identical chemical environments, the nuclear transition energies would be exactly equal and resonant absorption would be observed. This would produce a single absorption line.

As the environment of the nuclei will certainly be different to our source, the hyperfine interactions between the nucleus and its environment will change the energy of the nuclear transition. To detect this, we need to change the energy of our probing gamma-rays. During Mössbauer absorption spectroscopy, the source is accelerated through a range of velocities to produce a Doppler effect, changing the energy of the gamma ray slightly. For convenience the energy scale of a Mössbauer spectrum is thus reported in terms of the velocity. At velocities

corresponding to the resonant energy levels of the sample, a fraction of the gamma rays are absorbed, resulting in a drop in the measured intensity and a corresponding peak in the spectrum.

The number, positions, and intensities of the peaks provide information about the chemical environment of the absorbing nuclei and can be used to characterize the sample. Typically, there are three types of nuclear interactions that are observed: isomer shift, quadrupole splitting and hyperfine splitting. **Isomer shift** (δ) is a relative measure describing a shift in the resonance energy of a nucleus due to the transition of electrons within its s orbital. The whole spectrum is shifted in either a positive or negative direction depending upon the s electron charge density. The isomer shift is useful for determining oxidation state, valency states, ligand bonding states etc. For example Fe^{3+} ions have lower isomer shifts than Fe^{2+} ions because s electron density at the nucleus of ferric ions is greater due to a weaker screening effect by d electrons.

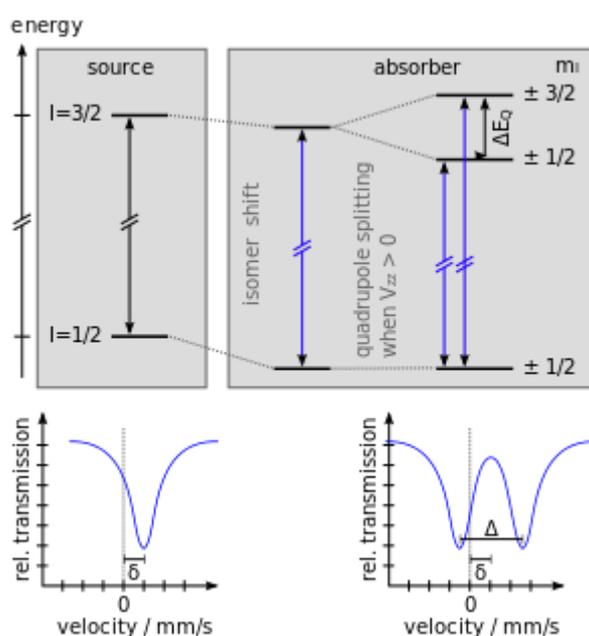


Figure 80: Chemical shift and quadrupole splitting of the nuclear energy levels and corresponding Mössbauer spectra

The **quadrupole splitting** is produced by the interaction between the nuclear energy levels and surrounding electric field gradient. It is observed in a non-spherical charge distribution in the presence of an asymmetrical electric field (like an asymmetric electronic charge distribution or ligand arrangement). The effect is a split of the nuclear energy levels and in the case of an isotope with a $I=3/2$ excited state, such as ^{57}Fe , the $3/2$ to $1/2$ transition is split into two substates $m_I=\pm 1/2$ and $m_I=\pm 3/2$. These appear as two specific peaks in a

spectrum, sometimes referred to as a 'doublet' (Figure 80). The Quadrupole splitting can be used for determining oxidation state, spin state, site symmetry and the organization of ligands.

The third effect, **magnetic splitting**, is a result of the interaction between the nucleus and any surrounding magnetic field. A nucleus with spin, I , splits into $2I + 1$ sub-energy levels in the presence of magnetic field. The three Mössbauer parameters: isomer shift, quadrupole splitting, and hyperfine splitting can often be used to identify a particular compound by comparison to spectra for standards.

In the absence of magnetic splitting, two equal-intensity lines centered on a velocity called "isomer shift", noted δ , and separated by a gap called "quadrupole splitting", noted ΔEQ are found. These two parameters are characteristic of the oxidation state of the iron and of its immediate environment.

This technique has been used to characterize the $[4Fe-4S]^{2+}$ cluster of MeH-HydA. Biological samples usually contain several types of iron species, each of which has a specific Mössbauer signature. The Mössbauer parameters of $[Fe-S]$ clusters are clearly distinct from those of other biological iron centers, like hemes and non-heme centers different ligands.

Iron in the most common $[Fe-S]$ clusters has the formal oxidation state iron(II) or iron(III) with quasi-tetrahedral coordination environment. These centers bear a charge and their Mossbauer spectra depend both on the type of cluster and on its net charge. Their quadrupolar parameters (δ and ΔEQ) depend mostly on the oxidation state of each iron site: Fe^{2+} , Fe^{3+} , or delocalized $Fe^{+2.5}$.

The $[4Fe-4S]^{2+}$ are mixed-valent in all oxidation states and show valence delocalization, predominantly occurring in pairs of formal $Fe(II)-Fe(III)$ ions, producing equal charge densities at the four iron nuclei.

This kind of clusters, as described early, can have different redox states: $[4Fe-4S]^{0/+1/+2/+3}$ along which the $+2/+1$ are the most common. Clusters in the $+2$ state consist of two delocalized ($2Fe^{+2.5}$) pairs, antiferromagnetically coupled into an $S=0$ ground state. The resulting spectrum is the convolution of four quadrupole doublets with $\delta \approx 0.44$ mm/s and $\Delta EQ \approx 1$ mm/s that appear as a single, sometimes asymmetric doublet.

MeH-HydA sample, after ^{57}FeS reconstitution as described in paragraph 5.4, was concentrated to 0.67 mM (0.4 ml) and flash frozen in the glove box using liquid iso-pentane cooled with liquid nitrogen.

5.9.6 X-ray Crystallography

X-ray crystallography is the most powerful spectroscopy to study and identify the molecular structure of a crystal, in which the crystalline atoms cause a beam of incident X-rays to diffract into many specific directions. By analyzing the angles and intensities of the beams, it's possible to find a 3D picture of the electron density within the crystal. The electron density map can provide the positions of the atoms, their chemical bonds and supramolecular organization.

This high resolution microscopy enables the visualization of the protein structure at the atomic level, giving a direct view of the protein function. Moreover, it enables to study how proteins interact with other molecules (ligands, substrates or other biologic molecules), how they undergo conformational changes, and how they perform catalysis in the case of enzymes. The first step is to obtain a pure sample of the target protein that is reached, as described early, after different steps: isolation of the gene, cloning into a high expression system, expression and purification to get a pure and homogenous protein sample. Dynamic light scattering can provide useful information about the homogeneity and monodispersity of the protein sample. It could also help in the choice of the correct protein buffer.

The concept is to prepare a solution in which the protein is soluble, homogenous and changing some conditions (protein or precipitant concentration, pH etc.) the protein slowly becomes insoluble and comes out of solution as crystals rather than as a precipitate. The protein solution has to reach a point of supersaturation, a metastable state which depends on the solubility of the protein. At this point, if you are lucky the nucleation is driven by this supersaturation and moves towards an equilibrium between crystals, protein solution and other solutes which can change the protein solubility.

A lot of parameters have to be adjusted (Figure 81) to find conditions for crystals growth and among them: protein concentration, pH, precipitant concentration, ions or effectors and temperature.

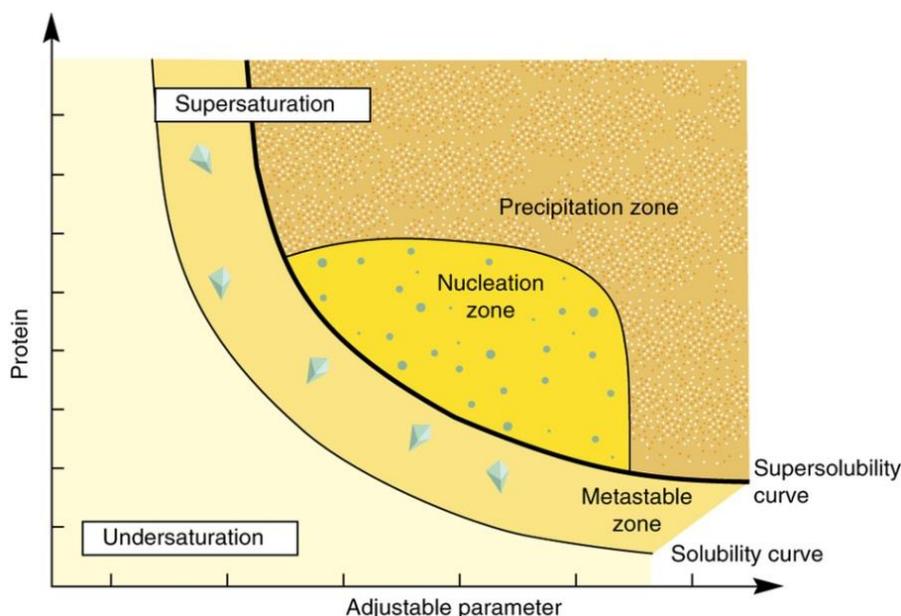


Figure 81: Schematic illustration of a protein crystallization phase diagram.

High throughput methods have been developed to facilitate and speed up the large number of experiments required to explore the various conditions that are necessary for successful crystal growth. There are numerous commercial kits available for order which apply preassembled “ingredients” in systems guaranteed to produce successful crystallization. Liquid-handling robots can be used to set up and automate large number of crystallization experiments simultaneously.

In the case of oxygen-sensitive proteins, the crystallization requires a glove box and all of the solutions have to be oxygen-free.

The “Laboratoire de Chimie de Processus Biologiques” in Collège de France where I worked during the PhD, has recently bought and equipped a glove box for protein crystallization trials with the intent to crystallize protein sensitive to oxygen (Figure 82). The glove box is equipped with an Oryx Nano crystallization robot (Douglas Instruments) installed in a temperature and hygrometry regulated anaerobic chamber (Figure 82 left). This part of the project has been done under the supervision of Dr. Ludovic Pecqueur, which is the responsible of the protein crystallographic part of the laboratory.



Figure 82: Pictures of glove boxes for protein crystallography. The chamber on the left side contains the Oryx Nano crystallization robot.

There are different crystallization methods and in our case we focused on vapor diffusion sitting drop and hanging drop.

A protein solution containing a precipitating agent at a concentration below the concentration needed for precipitation is equilibrated by vapor diffusion with a large volume of more concentrated precipitant solution, called reservoir.

Overtime water evaporates through vapor diffusion and the salt concentration slowly increases until it reaches equilibrium with the reservoir. At one point, the protein will have reached its saturation point and, hopefully, crystallize. If the precipitant concentration is too low, saturation may never be reached, while at high concentration, saturation may occur so quickly that protein precipitates appear. The optimal concentration is somewhere between these two opposite extremes. All this process is protein dependent; crystals can appear on a scale time ranging from few days to weeks and months.

The proteins studied in this PhD are oxygen sensitive and all crystallization trials have been conducted in strict anaerobic conditions. All the proteins have been freshly reconstituted and purified in anaerobic conditions on a Superdex S 200 10/300. Once purified, they have been transferred, via sealed vials, in a second glove box for crystallographic studies.

96-wells plates are used for crystallization trials and filled with reservoir solutions commercially available. Around 500-600 conditions are tested for each protein concentration. Crystallization trials are automated with an Oryx Nano crystallization robot (Douglas Instruments) installed in a temperature and hygrometry regulated anaerobic chamber. Protein drops of 0.2 μl are mixed with equal volumes of reservoir solution and are covered by a tape to allow reaching equilibrium between the liquid and gas phase. Once hits are obtained they are reproduced manually with the hanging drop method. Once you succeed to obtain and reproduce crystals, it's time for the next step.

Materials and Methods

The crystals have to be fished and mounted in a loop of the correct size and flash-cooled in liquid nitrogen. In the glove box the procedure is slightly different due to the impossibility to enter liquid nitrogen inside the anaerobic chamber. A home-made cylinder for crystals flash-cooling containing liquid propane (solidification temperature, $-187\text{ }^{\circ}\text{C}$) is frozen in liquid nitrogen and entered inside the glove box (Figure 83).



Figure 83: Loop (left) and device for crystal freezing inside the glove box (right). The cylindrical chamber is filled with liquid propane, frozen in liquid nitrogen and entered in the anaerobic chamber.

The liquid propane liquefied in a couple of minutes and the crystals are flash-frozen in liquid propane. Once back outside the glove box, the crystal is transferred in liquid nitrogen and stored until synchrotron experiments.

The crystals were analyzed at *SOLEIL* synchrotron facility in GIF-Sur-Yvette.

The crystals are attached to a goniometer head at the synchrotron line, which enables the sample to be accurately positioned in the X-ray beam. A cold nitrogen gas stream keeps the crystal at 100 K throughout the experiment. Focused X-rays emerge from a collimator and strike the crystal to produce a diffraction pattern that is recorded on the X-ray detector. Due to the high intensity of synchrotron radiation, the crystals are more prone to radiation damage, even at cryogenic temperatures. The perfect crystal should have sharp spots and no evidence of salt or ice which gave very strong spots or rings. It should be an ordered array of unit cells that continuously repeat by translation in all the dimensions. The unit cell is the smallest repeated element that generates the crystal, and is defined by three distances (a, b, c) and three angles (α, β, γ).

The data analysis required the help of Dr. L. Pecqueur that solved the RX structures of two HydF proteins (apo and reconstituted with the 4Fe4S cluster). The structures were solved by molecular replacement.

In some cases it can be useful to test the anomalous scattering of some atoms (close to their X-ray absorption edges) to gain phase information. Single-wavelength anomalous

diffraction (SAD) is a very effective method that relies entirely on the measurement of the anomalous diffraction produced by one or more scattering atoms in the crystal. Data sets were recorded from the same crystal at the wavelength around the X-ray absorption edge of Iron (7.15 keV). This procedure allowed confirming the position of the [4Fe–4S] cluster inside the protein structure.

5.9.6.1 Refinement statistics

The refinement statistics of the HydF structures solved in this thesis work have been summarized in Table 11.

Table 11: Data collection and refinement statistics of apo-TmHydF and FeS-TmeHydF proteins.

Dataset	Apo-TmHydF	FeS-TmeHydF
PDB code	5LAD	5KH0
Data collection		
Cell parameters (Å)	126.14 126.14 68.84 90.0 90.0 90.0	72.01 173.48 89.32 90.0 110.0 90.0
Space group	I4	P2 ₁
Wavelength (Å)	0.97934	0.9786
Resolution (Å)	44.60-3.00 (3.10-3.00)	47.62-2.80 (2.90-2.80)
R_{meas} (%)	5.1 (156.7)	6.7 (80.6)
I/σ	16.42 (0.99)	11.42 (1.57)
CC_{1/2}	0.999 (0.524)	0.998 (0.549)
Completeness (%)	0.974 (0.91)	0.95 (0.77)
Multiplicity	4.5 (4.4)	2.9 (2.8)
Wilson B (Å²)	125.23	92.27
Refinement Statistics		
R_{cryst}/R_{free}	0.244 / 0.280	0.234 / 0.262
CC_{work} / CC_{free}	0.93 / 0.90	0.92 / 0.89
No. of reflections	10940	48183
Work/Test set	9633 / 1050	46229 / 1948
Rms deviation from ideal geometry		
Bonds	0.011	0.09
Angles	1.19	1.14
No. of refined atoms		
Protein	2775	11339
Iron atoms	-	4
Ramachandran plot		
Preferred / allowed / outliers	92 / 8 / 0	94 / 5 / 1
B-factor values (Å²)		
Protein (overall)	115.15	86.47
GTPase domain	131	107.5
Dimerisation domain	100	67.5
Cluster binding	107	76.9
domain		
Fe₄S₄ cluster	-	66.15

References:

- 1 X. Brazzolotto, J. K. Rubach, J. Gaillard, S. Gambarelli, M. Atta and M. Fontecave, *J. Biol. Chem.*, 2006, **281**, 769–774.
- 2 G. Berggren, R. Garcia-Serres, X. Brazzolotto, M. Clemancey, S. Gambarelli, M. Atta, J.-M. Latour, H. Hernández, S. Subramanian, M. Johnson and M. Fontecave, *JBIC J. Biol. Inorg. Chem.*, 2014, **19**, 75–84.
- 3 G. Caserta, A. Adamska-Venkatesh, L. Pecqueur, M. Atta, V. Artero, S. Roy, E. Reijerse, W. Lubitz and M. Fontecave, *Biochim. Biophys. Acta*, 2016, **1857**, 1734–1740.
- 4 L. Loiseau, S. Ollagnier-de Choudens, D. Lascoux, E. Forest, M. Fontecave and F. Barras, *J. Biol. Chem.*, 2005, **280**, 26760–26769.
- 5 J. Esselborn, C. Lambertz, A. Adamska-Venkatesh, T. Simmons', G. Berggren, J. Nothl, J. Siebel, A. Hemschemeier, V. Artero, E. Reijerse, M. Fontecave, W. Lubitz and T. Happe, *Nat. Chem. Biol.*, 2013, **9**, 607–609.
- 6 V. Artero, G. Berggren, M. Atta, G. Caserta, S. Roy, L. Pecqueur and M. Fontecave, *Acc. Chem. Res.*, 2015, **48**, 2380–2387.
- 7 J. F. Siebel, A. Adamska-Venkatesh, K. Weber, S. Rumpel, E. Reijerse and W. Lubitz, *Biochemistry*, 2015, **54**, 1474–1483.
- 8 W. W. Fish, *Methods Enzymol.*, 1988, **158**, 357–364.
- 9 H. Beinert, *Anal. Biochem.*, 1983, **131**, 373–378.

Acknowledgments

Acknowledgments

The work presented in this thesis could not come without support from the following people to whom I would like to express my gratitude.

First of all, I would like to thank Professor Marc Fontecave for giving me the possibility to work in the *Laboratoire de Chimie de Processus Biologiques* at Collège de France. His interest in science, supervision and the scientific support were fundamental for all the work in this thesis. I still remember the first day in the new lab at Collège de France, Marc said: “*We are lucky to work in this institution in the center of Paris but we have a great responsibility since French people pay taxes for that and the money must not be wasted*”. I’m also grateful to Dr. M. Atta (Grenoble, CEA), who introduced me in the wonderful world of iron-sulfur proteins, and Dr. E. Mulliez (Grenoble, CEA) for helping and fruitful discussions.

I thank all the co-investigators which contributed to this work and in particular: Professor W. Lubitz, Dr. E. Reijerse and Dr. A. Adamska Venkatesh (Mülheim, MPI) who performed EPR, FTIR HYSCORE and Mössbauer experiments.

Dr. V. Artero and Dr. S. Roy (Grenoble, CEA), who provided the synthetic compounds for maturation experiments.

I want to thank also Dr. L. Pecqueur for helping in the crystallographic part and for the time (nights) spent together at the synchrotron.

I have also to thank the whole CPB lab who created a nice atmosphere for working. In particular I thank former members who helped me for my integration (Matt, Noemie, Alex, JP, Lucie and Mimi) in Paris and in Collège de France. I express my gratitude also to Charles, Cameron, Pierre, Ornella for helpful discussion and nice coffee breaks. I want to thank also my co-user of glove box, Simon, for all the discussions on reconstitution, iron-sulfur, EPR etc...I express my gratitude also to Maria for her friendship and the very good *paella*.

I thank also Xia, Vincent, Tanja, Greg, Victor, Safwan and Huan who still think that I’m a biologist. I’m also very grateful to Dr. P. Simon who was there for any technical problem and to Jonny for not working (not true).

I thank the chemistry coloc (Thibault, Gauthier and Quentin) for nice time spent together and the help and corrections of my french. Overall, a personal thank to Thibault who contributed to the French summary of this thesis and Gauthier, who corrected Thibault’s french.

Acknowledgments

I thank also my **first** student, Cecilia Papini, who resisted working under my conditions. I wish the best for her. It has been an honor to be part of you Master defense. Moreover, you Napolitan is improving...keep going.

I express gratitude to all the people I met during these three years in Grenoble, Marseille and Paris (Mahmoud, Alberto, Matteo, Carlotta). Thanks for all the help and the nice time spent together.

I'm grateful also to my friends in Italy and abroad, chemists and not, who always supported and encouraged me. I do not have to write their name, they know that I love them.

The biggest thank goes to my family, my mom and my brother, who always are with me. I'm really proud to have them.

After the whole cycle, the final and last "merci" is for me.

Paris 07/11/2016

Giorgio Caserta

Annex I

Publications on the scope of the thesis

- The [FeFe]-hydrogenase maturation protein HydF: Structural and Functional Characterization.
Caserta G., Pecqueur L., Adamska-V. A., Papini C., Roy S., Artero V., Atta M., Reijerse E., Lubitz W., Fontecave M. **Under revision** (*Nat. Chem. Bio.*)

- Metallocluster assembly: Maturation of [FeFe]-hydrogenases.
Caserta G., Pecqueur L., Papini C. and Fontecave M. *Encyclopedia of Inorganic and Bioinorganic Chemistry*. **Manuscript accepted.**

- Chemical assembly of multiple metal cofactors: The heterologously expressed multidomain [FeFe]-hydrogenase from *Megasphaera elsdenii*.
Caserta G., Adamska-Venkatesh A., Pecqueur L., Atta M, Artero V, Roy S, Reijerse E, Lubitz W, Fontecave M. *Biochim Biophys Acta- Bioenergetics*. 2016 Jul 13;1857(11):1734-1740. **DOI:** 10.1016/j.bbabo.2016.07.002.

- From enzyme maturation to synthetic chemistry: the case of hydrogenases.
Artero V, Berggren G, Atta M, **Caserta G**, Roy S, Pecqueur L, Fontecave M. *Acc Chem Res*. 2015 Aug 18;48(8):2380-7. **DOI:** 10.1021/acs.accounts.5b00157. Epub 2015 Jul 13.

- Artificial hydrogenases: biohybrid and supramolecular systems for catalytic hydrogen production or uptake.
Caserta G, Roy S, Atta M, Artero V, Fontecave M. *Curr Opin Chem Biol*. 2015 Apr;25:36-47. **DOI:** 10.1016/j.cbpa.2014.12.018. Epub 2014 Dec 29.

Other publications

- An artificial heme-enzyme with enhanced catalytic activity: evolution, functional screening and structural characterization.

Vitale R, Lista L, Cerrone C, **Caserta G**, Chino M, Maglio O, Nastri F, Pavone V, Lombardi A. *Org Biomol Chem.* 2015 May 7;13(17):4859-68. **DOI:** 10.1039/c5ob00257e.

- Reactivity of the excited states of the H-cluster of FeFe hydrogenases.

Sensi M., Baffert C., Greco C., **Caserta G.** Gauquelin C., Saujet L., Fontecave M., Roy S., Artero V., Soucaille P., Meynial-Salles I., Bottin, H., De Gioia L., Fourmond V., Léger, C., Bertini. **DOI:** 10.1021/jacs.6b06603