In Vitro Biosynthesis of the FeGP Cofactor of the [Fe]-Hydrogenase

Dissertation

zur Erlangung des Grades eines

Doktor der Naturwissenschaften

(Dr. rer.nat.)

des Fachbereichs Biologie der Philipps-Universität Marburg

Vorgelegt von

Sebastian Jakob Schaupp

Aus Pforzheim

Marburg an der Lahn, 2022

Die vorliegende Dissertation wurde von Oktober 2018 bis Juni 2022 am Max-Planck-Institut für terrestrische Mikrobiologie in der Abteilung „Microbial protein structure“ unter Leitung von Dr. Seigo Shima angefertigt.

Vom Fachbereich Biologie der Philipps-Universität Marburg (Hochschulkennziffer 1180) als Dissertation angenommen am _____

Erstgutachter(in): Dr. Seigo Shima

Zweitgutachter(in): Prof. Dr. Johann Heider

Tag der Disputation: _____

Table of content

Abbreviations

ADT	2-azapropane-1,3-dithiolate
AMP	Adenosine-mono-phosphate
AS	Ammonium sulfate
ATP	Adenosine-tri-phosphate
CN	Cyanide
CO	Carbon monoxide
CoA	Coenzyme A
CP	Carbamoyl phosphate
CTAB	Cetrimonium bromide
CV	Column volume
dAdo	5'-deoxyadenosine
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EPR	Electron-paramagnetic resonance
$F_{420}H_2$	Reduced F_{420}
FeGP	Ironguanylypyridinol
Frh	F_{420}-reducing [NiFe]-hydrogenase
GDP	Guanosine-di-phosphate
GMP	Guanosine-mono-phosphate
GTP	Guanosine-tri-phosphate
H protein	Lipoyl-H-Protein
H_2	Dihydrogen
H_4MPT	Tetrahydromethanopterin
IPTG	β-D-1-thiogalactopyranoside
IR	Infrared spectroscopic
jHcgF	HcgF from *Methanocaldococcus jannaschii*
jHmd	Apo Hmd from *M. jannaschii*
K_i	Inhibitor constant
M. jannaschii	*Methanocaldocuccus jannaschii*
M. marburgensis	*Methanothermobacter marburgensis*
M. maripaludis	*Methanococcus maripaludis*
MALDI-TOF-MS	Matrix assisted laser desorption ionization – time of flight mass spectrometry
Met	Methionine
mHcgE	HcgE from *Methanothermobacter marburgensis*
min	Minutes
MOPS	3-(N-morpholino)propanesulfonic acid
mpHcgB	HcgB from *Methanococcus maripaludis*
mpHcgC	*HcgC from Methanococcus maripaludis*
MS/MS	Tandem mass spectrometry
Mtd	Methylene-H_4MPT dehydrogenase
NMR	Nuclear magnetic resonance
rpm	Rotations per minute
rSAM	Radical SAM enzyme
RT	Room termperature

SAH	*S*-adenosyl homocysteine
SAM	*S*-adenosyl methionine
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
SHMT	Serine hydroxymethyltransferase
T protein	Aminomethyltransferase
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Triflouroacetic acid
TosMIC	Tosylmethylisocyanide
Tris	Tris(hydroxymethyl)aminomethane
XAD	380 mM Formiat/NaOH buffer pH 3
XAS	X-ray absorption spectroscopic

Abstract

Life requires many challenging chemical reactions, which are enabled by metallocofactors of enzymes. One such reaction is splitting of molecular hydrogen (H_2) by hydrogenases, which are classified based on their cofactors as [NiFe]-, [FeFe]- or [Fe]-hydrogenases. These enzymes are of special interest due to their potential application in solutions for future energy storage. The [Fe]-hydrogenase reversibly catalyzes the heterolytic cleavage of H_2 into a proton and a hydride. The latter is subsequent transferred to methenyl-tetrahydromethanopterin, which is involved in the hydrogenotrophic methanogenesis. The [Fe]-hydrogenase contains the iron-guanylylpyridinol (FeGP) cofactor as prosthetic group of this enzyme. This cofactor consists of a low-spin Fe(II), which is coordinated by two CO, one cysteine sulfur and one bidentate acyl-methylene-pyridinol ligands. In this work, the biosynthesis of the FeGP cofactor was investigated. The cofactor is presumably synthesized by the reactions of the HcgA-G proteins. In previous studies, the functions of HcgB–F have already been partially analyzed. To investigate the missing function of HcgA and HcgG, as well as to confirm the nature of the precursors of the FeGP cofactor, we developed a method of synthesizing the FeGP cofactor *in vitro* using a mixture of defined and undefined compounds in combination with proteins. The *in vitro* biosynthesis solution contains the [Fe]-hydrogenase apoenzyme, a possible precursor, e.g. 6-carboxyl-methylene-3,5-dimethyl-4-guanylyl-2-pyridinol (compound **3**), ATP/Mg^{2+}, *S*-adenosyl methionine, dithiothreitol and sodium dithionite, Hcg proteins, and cell extract from a *Methanococcus maripaludis Δhcg* mutant that lacks endogenous [Fe]-hydrogenase activity. Based on the results of the *in vitro* biosynthesis, we confirmed the structure of the pyridinol precursors, which were predicted based on investigations of the Hcg proteins. Importantly, we confirmed that a carboxyl group of these pyridinol precursors is converted into the acyl ligand of the FeGP cofactor. In addition, we found that the *in vitro* biosynthesis requires reducing equivalents, which can be generated from H_2 or formate. We then observed requirements of some small cellular components such as a possible CO precursor and an electron carrier for the biosynthesis reactions. We demonstrated that CO gas can also be incorporated into the CO ligands. Furthermore we confirmed that the reaction of HcgE generates a proposed adenylylated compound **3** in the presence of ATP, which was also necessary for the *in vitro* biosynthesis, and that the compound **3** is bound to HcgF**.** Moreover, we showed that HcgA catalyzes the biosynthesis of the initial pyridinol precursor, 6-carboxyl-methylene-4-hydroxyl-5-methyl-2-pyridinol. Further we demonstrated that HcgG catalyzes the biosynthesis of the FeGP cofactor from the guanylylpyridinol precursor **3** and the components from the cell extract of methanogens (see PhD thesis of F. Arriaza). In addition, in this thesis, I discuss the production of the FeGP cofactor from compound **3** in a fully defined protein mixture containing HcgE, HcgG and small components from the cell extract of *M. maripaludis* based on some preliminary results.

Zusammenfassung

Das Leben benötigt unterschiedlichste, komplexe, chemische Reaktionen, die zum Großteil nur durch das Zusammenspiel von Metallkofaktoren und Enzymen erreichbar sind. Ein Beispiel ist die Spaltung von Wasserstoff (H_2) durch Hydrogenasen, die nach der Metallzusammensetzung ihrer Kofaktoren als [NiFe]-, [FeFe]- oder [Fe]-Hydrogenasen klassifiziert werden. In meiner Doktorarbeit habe ich die Biosynthesynthese des Kofaktor der [Fe]-Hydrogenase untersucht, der auch Eisenguanylylpyridinol (FeGP) genannt wird. Er besteht aus einem Eisenatom, das von zwei Kohlenstoffmonoxid (CO) Liganden, dem Schwefel eines Cysteins und einem Acylmethylenepyridinol komplexiert wird. Es wird angenommen, dass die Proteine HcgA-G für die Biosynthese zuständig sind, wobei HcgB-F bereits teilweise strukturell und funktional untersucht wurden. Um diese Ergebnisse zu bestätigen, die Funktion von HcgA und HcgG aufzuklären und die Bildung des Acyl- und der CO-Liganden zu analysieren, habe ich einen *in vitro* Biosyntheseassay entwickelt. Diese Strategie wurde auch für die Untersuchung der Biosynthese der Metallkofaktoren der [FeFe]-Hydrogenase und der Nitrogenase verwendet. Der entwickelte Assay ermöglicht die Aktivierung der [Fe]-Hydrogenase Apoenzyms durch Zugabe eines Vorläufermoleküls 6-Carboxyl-methylene-3,5-dimethyl-4-guanylyl-2-pyridinol (**3**), chemischen Substraten und des Zellextrakts von Δ*hcg* Mutanten des Archaeons *Methanococcus maripaludis*. Dabei bestätigten die Ergebnisse aus den Experimenten mit der Δ*hcgB*Δ*hcgC* Mutante die Identität des Kofaktorvorläufermoleküls **3**. Insbesondere die Präsenz einer Carboxylgruppe konnte gezeigt werden, die in einem der folgenden Schritte zum Acylliganden des Kofaktors reduziert wird. Um mehr über die Biosynthesereaktionen herauszufinden habe ich mittels der *in vitro* Biosynthese und anderer biochemischer Methoden untersucht, welche Substrate für die Reaktionen notwendig sind. Hier habe ich gezeigt, dass die *in vitro* Biosynthese ATP abhängig ist. Außerdem benötigt der *in vitro* Biosyntheseassay Reduktionsequivalente, die aus Wasserstoff oder Formiat erzeugt wurden. Weiterhin wurde während der *in vitro* Biosynthese gasförmiges CO in die CO-Liganden des Kofaktors eingebaut. Durch die Verwendung von verschiedenen Mutanten für die Zellextrakterzeugung, habe ich gezeigt, dass die vorgeschlagene Reaktionsfolge von HcgB und HcgC korrekt ist, sowie das HcgF für die *in vitro* Biosynthese vermutlich nicht notwendig ist. Außerdem habe ich die Bildung eines vorgeschlagenen AMP-**3**-Zwischenproduktes in der HcgE-Reaktion durch eine gekoppelte, chemische Reduktion bestätigt und dass Molekül **3** an HcgF bindet. Die Funktion von HcgA, die Bildung des ersten Kofaktorvorläufermoleküls (**1**), und HcgG, die Umwandlung von Molekül **3** in den FeGP-Kofaktor, wurden außerdem aufgeklärt (Dissertation, F. Arriaza). Wir haben auch erste Erkenntnisse zur *in vitro* Biosynthese des FeGP-Kofaktors mit reinen Hcg-Proteinen und dem Filtrat des Zellextrakts von *M. maripaludis* gewonnen, wobei diese Experimente bisher nicht reproduziert werden konnten.

Parts of the studies presented in this thesis have been published or are in preparation for publication:

S. Schaupp,* F. J. Arriaza-Gallardo,* H.-j. Pan, J. Kahnt, G. Angelidou, N. Paczia, K. Costa, X. Hu, S. Shima, *In vitro* biosynthesis of the [Fe]-hydrogenase cofactor verifies the proposed biosynthetic precursors. *Angewandte Chemie International Edition* 2022, *61*, e202200994, DOI: 10.1002/anie.202200994.

F. J. Arriaza-Gallardo,* S. Schaupp* *et al.* , Functions of HcgA and HcgG in the [Fe]-hydrogenase cofactor biosynthesis. *To be submitted*.

***** These authors contributed equally.

1 Introduction

Metallocofactors

Enzymes are indispensable for life, since the enzymatic conversion of molecules is the basis of all metabolisms. The amino acid side chains and main chains of proteins can act as the catalytic site of a wide range of enzymes; however, some reactions are unachievable by protein based catalysis. To overcome this problem, nature evolved enzymes to use metals as prosthetic groups. Non-transition metals e.g. magnesium (Mg), calcium (Ca) and zinc (Zn), and also lanthanides can serve as the prosthetic group of enzymes.[1] Transition metals, such as iron (Fe), nickel (Ni), cobalt (Co), copper (Cu), manganese (Mn), molybdenum (Mo), and less frequently tungsten (W) and vanadium (V), are also used to enable difficult enzymatic reactions. The transition metals are mostly conjugated with non-protein organic and/or inorganic ligands to form a complex metallocofactor. The ligands tune the electronic properties of the transition metals in the metallocofactors. The amino acid side chains of the protein also serve as ligands of metallocofactors, which are responsible for the correct incorporation of the metallocofactor into the enzymes' active sites and modify the properties of the metal center.[2] Of the many functions catalyzed by enzymes containing metallocofactors, the ability to activate gasses, in particular H_2, is of great importance for microbial life. Hydrogenases are necessary for the utilization of H_2 as electron source and production of H_2 as electron sink.[3] The ability of these enzymes to reversibly split H_2 into a proton and a hydride relies on their metallocofactors. In my PhD project, I have studied the biosynthesis of the metallocofactor of the [Fe]-hydrogenase mainly using an *in vitro* biosynthesis method. In the following chapters, I outline structures and functions of hydrogenases and their unique metallocofactors, as well as the biosynthesis of those metallocofactors. In addition, I introduce nitrogenases with a focus on the machinery to biosynthesize their metallocofactors in order to analyze the methods used in these analyses and their use for the investigation of the FeGP cofactor biosynthesis.

Hydrogenases

Heterolytic splitting or formation of H_2 is the simplest chemical reaction (Equation 1) and is catalyzed by a group of enzymes, hydrogenases, which were first identified ninety years ago.[4] Hydrogenases are modular enzymes and serve in a range of different functions by forming complexes with other proteins. For example, by associating with electron carrier-binding proteins, hydrogenases can donate electrons from H_2 to several electron carriers (NAD^+, F_{420}, cytochromes, ferredoxins, etc.).[5] In complex with reductases, electrons from H_2 are provided for reduction reactions, including electron-bifurcating energy-coupling reactions.[6] Some hydrogenases form a complex with membrane proteins to catalyze energy conservation by mediating proton or sodium ion translocation. Hydrogenases are

also used to sense H_2 in the environment for regulatory functions.[7] In contrast to the versatile enzymatic functions of hydrogenases, the structures of the hydrogenase modules and the structure of the active site metallocofactors are common within each class of hydrogenase.[3]

$$H_2 \rightleftarrows 2H^+ + 2e^-$$
$$H_2 \rightleftarrows H^+ + H^- \quad (1)$$
$$H^- \rightleftarrows H^+ + 2e^- \quad (2)$$

Equation 1. Splitting of H_2 catalyzed by hydrogenases and thetwo half reactions. (1) Heterolytic cleavage into hydride (H^-) and proton (H^+). (2) Splitting of H^- into a H^+ and two electrons.

There are three phylogenetically different hydrogenases known today, which can be classified by the metal content of their active site clusters as [NiFe]-,[8] [FeFe]-[9] or [Fe]-hydrogenases.[10] In the metabolic pathways characterized so far, [NiFe]- and [Fe]-hydrogenases are mostly responsible for the uptake of H_2, whereas [FeFe]-hydrogenases are more often involved in H_2 release.[5] While some hydrogenases possess a catalytic bias, all hydrogenase catalyzed reactions are reversible.[5]

The iron-guanylylpyridinol (FeGP) cofactor of [Fe]-hydrogenase contains a single Fe(II),[11] which is ligated by two carbon monoxide (CO) ligands[12] and a bidentate pyridinol ligand, including an unusual iron-acyl ligand and one pyridinol nitrogen.[10] Additionally, a cysteine-sulfur ligand of the [Fe]-hydrogenase protein covalently anchors the cofactor to the protein (Figure 1a).[13] In the resting-state of the enzyme, a water molecule is bound to the iron center in *trans* to the acyl ligand. In the active form, the water ligand dissociates from the iron site.[10] A detailed description of this enzyme and the FeGP cofactor will be given below. In contrast to [Fe]-hydrogenases, the metallocofactor of the [FeFe]-hydrogenase (H-cluster) consists of a bimetallic cluster of two iron atoms connected via a cysteine-sulfur to a canonical [4Fe-4S]-cluster.[9] The two iron atoms are bridged by a unique 2-azapropane-1,3-dithiolate (ADT) ligand (Figure 1b).[14] The iron atoms of the bimetallic Fe center carry diatomic ligands, carbon monoxide (CO) and cyanide (CN) on each iron and an additional bridging CO ligand.[15] The active site cofactor of [NiFe]-hydrogenases does not contain any bridging organic ligands, instead the nickel and iron atom are bridged by two cysteine residues. The other two cysteine residues are terminally bound to the nickel atoms. In common with [FeFe]-hydrogenases the iron is coordinated by three diatomic ligands, two CN and one CO. Furthermore, an additional bridging ligand (X in Figure 1c) changes with oxidation state and is assumed to be the site of H_2 binding.[16]

The [Fe]-hydrogenase is unique among hydrogenases, since it catalyzes only the first part of the reaction that is catalyzed by [NiFe]- and [FeFe]-hydrogenases. The hydride obtained by heterolytic cleavage of H_2 is directly transferred to the hydride acceptor. The cofactor biosynthesis of these enzymes is described in the next chapter.

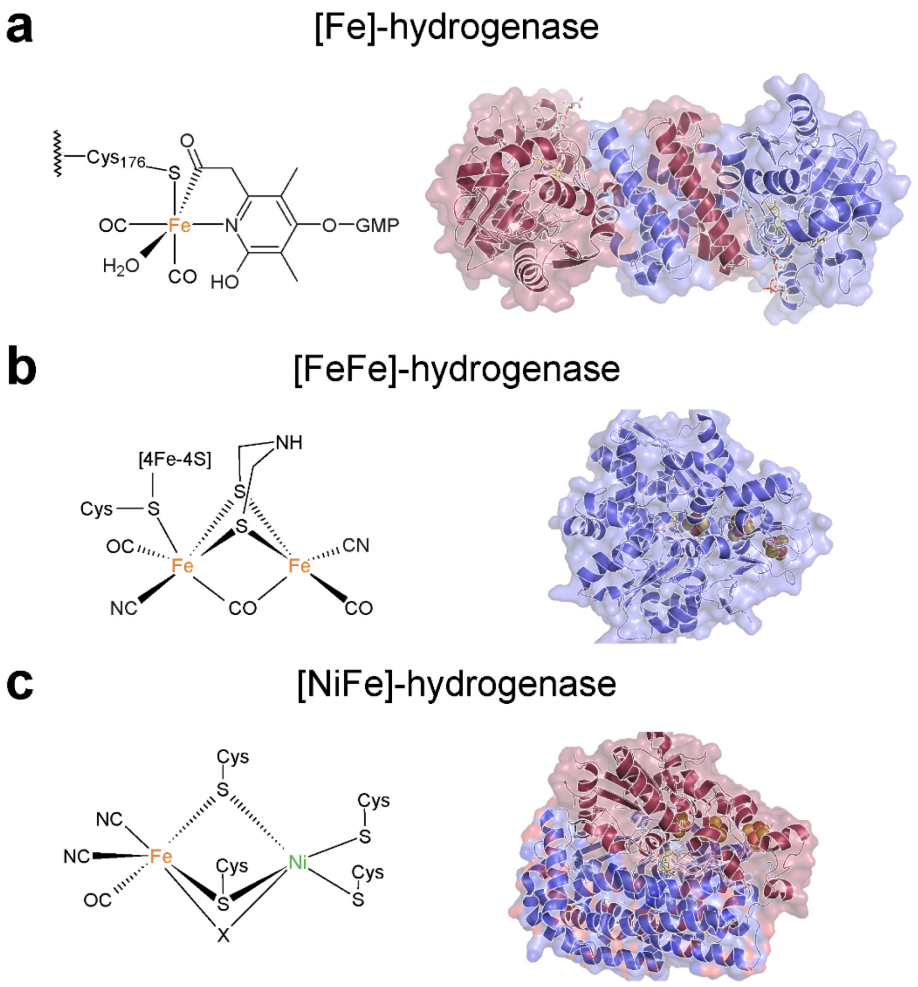

Figure 1. Metallocofactors (left column) of [Fe]- (a), [FeFe]- (b) and [NiFe]-hydrogenase (c) and their respective representative crystal structures (right column, PDBs: 6HAE, 1HFE and 1WUJ).

[Fe]-hydrogenase and the FeGP cofactor

Physiological context of the [Fe]-hydrogenase

Both [NiFe]- and [FeFe]-hydrogenases require a bimetallic metallocofactor to activate H_2.[8b, 9, 17] [Fe]-hydrogenase (or methylene-tetrahydromethanopterin dehydrogenase (Hmd)) in contrast carries a metallocofactor with only a single iron atom.[18] This enzyme is involved in the hydrogenotrophic methanogenic pathway,[19] where it catalyzes the activation of H_2 by splitting it into a proton and a hydride (Equation 1-1), which is transferred to methenyl-tetrahydromethanopterin (methenyl-H_4MPT$^+$) to form methylene-H_4MPT.[20] In contrast to other hydrogenases, the [Fe]-hydrogenase does not catalyzes the splitting of hydride into electrons and a proton (Equation 1-2). The production of the [Fe]-hydrogenase is not affected by the H_2 concentration in the culture. However, the expression of [Fe]-hydrogenase is up-regulated six-fold in a medium containing lower nickel concentration (0.2 μM Ni^{2+}) than the standard medium (5 μM Ni^{2+}).[21] In this condition, the F_{420}-dependent methylene-H_4MPT dehydrogenase (Mtd) is also up-regulated four-fold. Mtd catalyzes the same reaction as the [Fe]-hydrogenase, but uses the reduced form of F_{420} ($F_{420}H_2$) as electron donor instead of H_2. $F_{420}H_2$ is an electron donor and provides hydrides, for example for the two reduction reaction steps in the methanogenic reduction of CO_2 to methane.[22] In contrast to the [Fe]-hydrogenase and Mtd, under nickel-limiting conditions, the F_{420}-reducing [NiFe]-hydrogenase (Frh)[23] is down-regulated vastly (> 160 fold). Frh catalyzes the reduction of F_{420} with H_2 in the standard culture medium. The observed regulation of the [Fe]-hydrogenase, Mtd and Frh indicates that the function of Frh (reduction of F_{420}) is substituted by a coupled reaction of [Fe]-hydrogenase and Mtd under nickel-limiting conditions.[21b] Thus, the nickel concentration in the culture medium is crucial for the regulation of the F_{420}-reducing enzyme system and this finding suggests that the [Fe]-hydrogenase-Mtd system is the major source of reducing equivalents for the methanogenic pathway under lower nickel concentrations, which resemble the environmental nickel concentrations.[19]

Structure determination of the FeGP cofactor

When the [Fe]-hydrogenase was heterologously produced, the recombinant [Fe]-hydrogenase did not show any enzymatic activity.[24] This observation was the first hint that the [Fe]-hydrogenase requires a specific cofactor. The presence of a cofactor was substantiated by reconstitution of active holoenzyme by mixing heterologously produced apoenzyme and the extracted fraction from denatured [Fe]-hydrogenase holoenzyme from *Methanothermobacter marburgensis* (*M. marburgensis*) in the presence of 2-mercaptoethanol or dithiothreitol.[24] The reconstituted enzyme exhibited substantial enzymatic activity. However, in Buurman *et al.* [Ref 24], the enzymatic activity was not determined quantitatively and only trace amounts of iron and zinc were detected in the extracted cofactor

fractions. The amount of the intact cofactor in the fraction was too little to be unambiguously detected. This is the case because the [Fe]-hydrogenase cofactor is highly susceptible to UV-A and blue light.[18] The cofactor quickly loses the ability to reconstitute the active holoenzyme under room light. Later, from light-inactivated [Fe]-hydrogenase holoenzyme iron was extracted in the presence of EDTA, which is the first indication of the presence of a metallocofactor in the [Fe]-hydrogenase.[18] After finding the light sensitivity of the cofactor, the experiments of [Fe]-hydrogenase and the cofactor were performed anaerobically under red or yellow light in the presence of thiol reagents, where the reconstituted [Fe]-hydrogenase exhibited a specific activity comparable to that of the native enzyme.[13] The chemical analysis of the extracted cofactor indicated the presence of iron and phosphate in a 1:1 ratio.[18] Infrared spectroscopic (IR) analyses showed the presence of two CO ligands on the iron at an angle of 90°.[12] Mössbauer spectroscopy indicated that the electronic state of the iron is low-spin Fe(0) or Fe(II) and redox inactive.[11] X-ray absorption spectroscopic (XAS) analysis of the protein-bound and extracted cofactor showed the presence of sulfur and a nitrogen or oxygen ligand from the organic moiety of the cofactor.[13] Mutation analyses indicated that Cys176 (number in the enzyme from *Methanocaldococcus jannaschii* (*M. jannaschii*)) is an iron ligand.[13] The structure of the light inactivated cofactor was elucidated by using nuclear magnetic resonance (NMR) spectroscopy, mass spectrometry and other methods.[25] The light inactivated cofactor is 6-carboxyl-methylene-3,5-dimethyl-4-guanylyl-2-pyridone, which is in an equilibrium with its pyridinol form.[25] The whole structure of the intact cofactor was first proposed based on the crystal structure of the holoenzyme. Here the structure of the light inactivated cofactor, including the carboxyl group, was used to model the organic part.[26] In this paper, the cofactor was named FeGP cofactor. Unfortunately, the initial model of the FeGP cofactor contained a wrong interpretation due to the above mentioned use of the light-inactivated cofactor, which contained the 6-carboxyl-methylene group at the pyridinol ring, as the model for the organic moiety. This misinterpretation was realized during the crystal structure analysis of C176A mutant of the [Fe]-hydrogenase reconstituted with the FeGP cofactor.[27] The lack of the Cys176-S ligand resulted in the binding of a bulky, external dithiothreitol ligand at the iron site of the FeGP cofactor in the mutated holoenzyme. In this structure, there is not enough space for modeling of the 6-carboxyl group at the iron site. This discrepancy suggested that the carboxyl group does not exist in the intact FeGP cofactor. To explain this contradiction, a new model of the FeGP cofactor was constructed, in which the carboxyl group was substituted by an acyl group and the pyridinol ring was rotated by 180°.[27] The new model was proposed based on chemical precedence of conversion of acyl-ligands to carboxyl groups.[28] This new model is supported by the electron density obtained by X-ray crystallography and the XAS data.[27] Further evidence of the intact FeGP cofactor structure was obtained by high resolution mass spectrometric analysis. The exact mass and tandem mass spectrometry (MS/MS) data were in accordance with the structure of the acyl-ligand containing model.[29] In the same

paper, the presence of the infrared absorbance corresponding to the acyl ligand was reported. Additionally, the hydrolysis of the acyl ligand of the FeGP cofactor was confirmed by incorporation of [^{18}O]-H_2O into the 6-carboxyl-methylene group of the light-decomposed cofactor. The exact geometry of the FeGP cofactor was finally determined by atomic resolution crystal structure of [Fe]-hydrogenase from *Methanococcus aeolicus* reconstituted with the FeGP cofactor from [Fe]-hydrogenase from *M. marburgensis*.[10] The iron in this structure is also coordinated by a water molecule in the open conformation and the water ligand is absent in the closed conformation. The water binding site is thought to be the site of H_2 binding. The final structure of the FeGP cofactor is shown in Figure 1a, which contains a low-spin iron(II) ligated by two CO ligands, the nitrogen and acyl ligands from the pyridinol, and the cysteine-sulfur of the [Fe]-hydrogenase.

Properties of the FeGP cofactor

As described above, the FeGP cofactor can be extracted from denatured [Fe]-hydrogenase holoenzyme.[24] The isolated FeGP cofactor is stable in the presence of an external iron ligand (for example, 2-mercaptoethanol and acetate) under dark conditions with red/yellow light and can be separated from the denatured protein by ultrafiltration.[30] The extractability and stability of the FeGP cofactor is a distinct feature from the bimetallic cofactors of [NiFe]- and [FeFe]-hydrogenases. Because of these characteristics, the apoenzyme heterologously produced in *Escherichia coli* can be reconstituted by mixing it with isolated FeGP cofactor to form holo [Fe]-hydrogenase.[24] This is an important feature for the *in vitro* biosynthesis experiments described below.

Reconstitution processes of [Fe]-hydrogenase holoenzyme were analyzed using apoenzyme from *Methanolacinia paynteri* and FeGP cofactor isolated from [Fe]-hydrogenase of *M. marburgensis*.[31] The guanylyl moiety is probably involved in the initial incorporation of the cofactor into the protein.[31] Mutation analyses suggested that a lysine residue might have a guiding function for the pyridinol ring towards the exact position required for the covalent bond between the cysteine-sulfur and the iron of the FeGP cofactor.[31] The binding of the FeGP cofactor induces a change from closed to open conformation of the apoenzyme, which probably allows the substrate to bind to the active site.[31] [Fe]-hydrogenase holoenzyme can adopt an open and a closed conformation in the catalytic cycle.[10, 26, 32] The closed conformation of the holoenzyme is triggered by binding of the substrate, methenyl-H_4MPT^+. Activation of H_2 at the FeGP cofactor occurs in the active site in closed conformation, in which the active site cleft forms a hydrophobic environment and the water ligand at the iron site is removed.[10] H_2 is proposed to be supplied via a narrow channel and bound to the vacant iron coordination site after removal of the water ligand. H_2 is heterolytically cleaved at the iron site. The deprotonated 2-hydroxyl group of the pyridinol is proposed to act as the catalytic base.[10] In the closed conformation, the C14a of methenyl-H_4MPT^+ is in close

proximity to the proposed H_2-binding site of the FeGP cofactor, enabling hydride transfer.[10]

[Fe]-hydrogenase is inhibited by CN and CO by binding to the iron of the FeGP cofactor.[12] The other hydrogenases are also inhibited by CO, but the apparent K_i of the [Fe]-hydrogenase (20% (v/v) CO in the gas phase) is substantially higher than the K_i other hydrogenases.[18] Isocyanides have a chemical structure that is formally isoelectronic to CO and CN. Kinetic analyses indicated that isocyanides specifically inhibit the [Fe]-hydrogenase strongly with a K_i of 1−150 nM, although other hydrogenases are not inhibited by these compounds.[33] The inhibition kinetics data and the crystal structures of the [Fe]-hydrogenase inhibited by isocyanides suggested that the iron site *trans* to the acyl ligand is the hydrogen binding site.[34] In the enzyme-isocyanide complexes, the isocyano carbon binds to not only the iron but also to the deprotonated 2-hydroxyl group, which could be the reason for the strong binding affinity of isocyanides to the FeGP cofactor in the enzyme. The nature of this deprotonated 2-hydroxyl group may correlate with its predicted function as a catalytic base for H_2 activation.[10] The activation of H_2 by [Fe]-hydrogenase occurs only in the presence of substrate, as indicated by proton isotope (H/D) exchange and *para*-/*ortho*-H_2 exchange experiments.[35] The substrate dependence of the H_2 activation was explained to be due to the removal of the water ligand at the H_2 binding site upon binding of the substrate, methenyl-H_4MPT$^+$.[10] [Fe]-hydrogenases from the methanogens belonging to the *Methanothermobacter* genus are protected against oxygen- and light-induced inactivation by binding of an additional aspartate ligand to the predicted H_2-binding iron site of the FeGP cofactor in the hexameric state.[36] This finding emphasized the importance of the predicted H_2-binding site.

Biosynthesis of FeGP cofactor

Hcg proteins

Seven genes were commonly conserved in all microorganisms containing the gene of the [Fe]-hydrogenase structural gene (*hmd*).[19] These so-called ***h**md* **c**o-occurrin**g** genes (*hcgA*–*G*) are thought to be the minimal set of gene products needed for the biosynthesis of the FeGP cofactor. The *hcg* genes were first predicted by comparison of the gene cluster at the *hmd* gene in genomes of *Methanobacteriales* and *Methanopyrales* orders, and *Methanocorpusculum labreanum*, where *hcgA*–*G* and *hmd* genes are found in one region of the genome with several insertions of non-conserved genes.[19] Although in the methanogens belonging to *Methanococcales*, *hcgA* and *hcgG* genes co-localize with the *hmd* gene, the other *hcg* genes are conserved in other regions in the genome. Later, by *in silico* comparative genomic analysis, it was confirmed that all *hcg* genes coexist with the *hmd* gene in most methanogens containing the *hmd* gene.[37] Because in the case of the [NiFe]-hydrogenases gene cluster the bimetallic cluster biosynthesis genes are clustered with the structural gene of

[NiFe]-hydrogenases,[38] the *hcg* gene products could be the enzymes responsible for biosynthesis of the FeGP cofactor.[19] The deletion of the *hmd* gene in the *Methanococcus maripaludis ΔfrcΔfrh* strain leads to prolonged lag phase when grown with H_2 as sole energy source.[39] The same growth phenotype is observed in case of deletion of any of the *hcgA – G* genes, which suggested that HcgA–G are all required for the biosynthesis of the FeGP cofactor.[40] In addition, it was reported that deletion of the *hcgB* or *hcgC* gene resulted in loss of [Fe]-hydrogenase activity.[41]

The metabolic origin of the FeGP cofactor was investigated by *in vivo* isotope labelling.[42] The guanyl moiety is derived from the canonical GMP biosynthesis pathway. The labelling pattern of the pyridinol group is shown in Figure 2. The 3-methyl group originates from methionine, indicating involvement of a methyltransferase. In the pyridinol ring structure, two molecules of acetate are condensed to form the connected C5 and 5-methyl carbons, and C6 and 6-methylene carbons. The C2 of the FeGP cofactor also originates from C1 of acetate. Neither the CO nor the acyl ligand originated from acetate or pyruvate, instead they originate from CO or CO_2. This labeling pattern indicates that the 5-methyl carbon and 6-methylene carbon are originated from two acetate molecules; therefore, these two carbon atoms of the substituents are not synthesized by group transfer reactions and should exist in all pyridinol precursors. In contrast, the carboxyl group of the light-inactivated cofactor (or acyl ligand of the FeGP cofactor) was labeled by gaseous CO or CO_2 and not via any other metabolite tested. Those are critical points, while considering the structure of the pyridinol precursors.

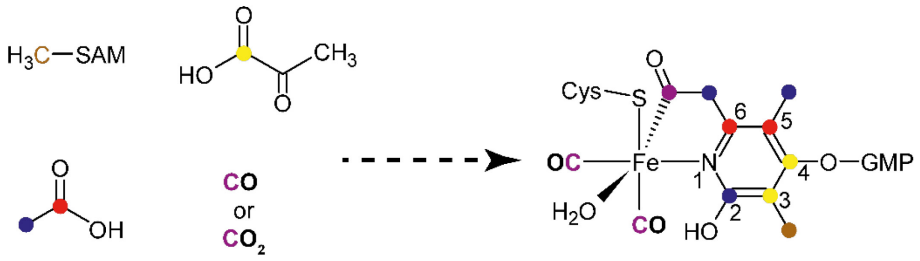

Figure 2. Labelling pattern of the carbon atoms of FeGP cofactor pyridinol and Fe center *in vivo*, adopted from [Ref 42]. The position of the ^{13}C label in the metabolites is marked by colors. The presence of a colored circle also marks a ^{13}C label.

Function of the Hcg proteins

The proposed reaction sequence of the FeGP cofactor biosynthesis based on previous studies is shown in Figure 3. 6-carboxyl-methylene-3,5-dimethyl-4-hydroxyl-2-pyridinol (**1**) is proposed as the initial pyridinol precursor. The C4 position is methylated to form compound **2**[41, 43] and then the 4-hydroxyl group is

guanylylated to form compound **3**[44]. The carboxyl group of compound **3** is activated by adenylylation and then conjugated with the sulfur of Cys9 of HcgF.[45] The thioester-bonded carbon is assumed to be the precursor of the acyl ligand.[45] HcgD is proposed as an iron-trafficking protein.[46] The Fe complex is completed by formation of the CO and acyl groups and transferred to [Fe]-hydrogenase apoenzyme to form the [Fe]-hydrogenase holoenzyme. The contributions of HcgB, HcgC, HcgE and HcgF in the reaction sequence are also shown in Figure 3. Thus, functions of these Hcg proteins and structures of all pyridinol precursors were proposed. However, the functions of HcgA and HcgG were not determined yet. In addition, as described below, the presence of the carboxyl group in the precursors **1**, **2** and **3** is not unambiguously verified by the previous studies (Figure 3).

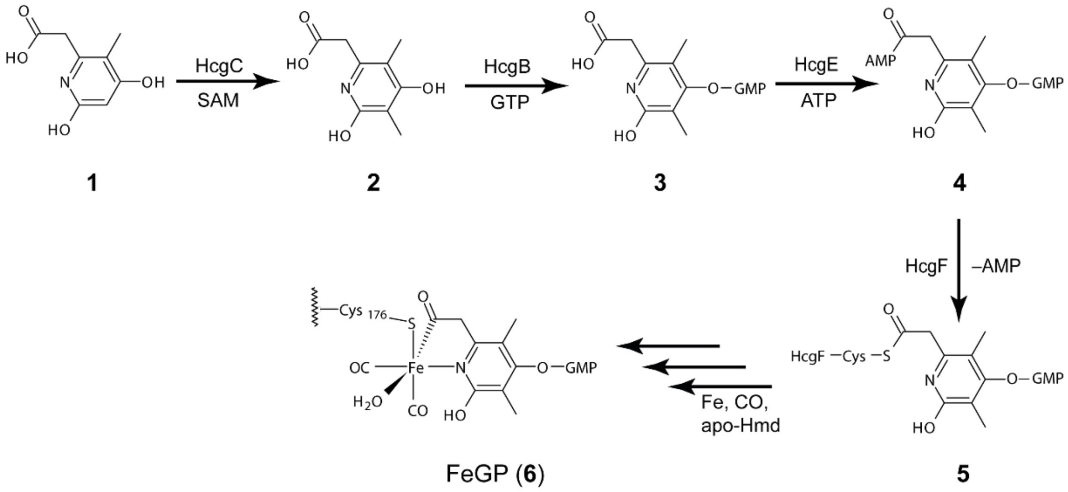

Figure 3. Biosynthesis of the FeGP cofactor as elucidated based on the prior structure-to-function studies.

The catalytic function of the Hcg enzymes and their substrates was studied using the structure-to-function strategy, in which the tertiary structure of Hcg proteins was compared with structures of enzymes with known functions. Based on the function of similar enzymes, the catalytic reaction and substrate structures were predicted. The hypothetical catalytic reactions were enzymatically tested using heterologously produced enzymes with commercially available substrate analogs and/or chemically synthesized substrates. The first result of the structure-to-function strategy of the Hcg proteins is the study of HcgB, by which the light inactivated cofactor, 6-carboxyl-methylene-3,5-dimethyl-4-guanylyl-2-pyridinol (**3**) is proposed as the physiological reaction product of the HcgB reaction.[44] This finding is the key point of elucidation of the structure of the biosynthetic precursors.

HcgB

The crystal structure of HcgB had been solved by a structural genomics project.[47] The tertiary structure showed similarities to nucleoside triphosphatases, which catalyze the cleavage of a pyrophosphate moiety from nucleoside triphosphates. Based on the structural similarity, it was assumed that HcgB catalyzes cleavage of GTP and conjugation of the guanylyl moiety to a pyridinol. Indeed, the chemical analysis of the HcgB reaction verified HcgB as a guanylyltransferase to produce a guanylylpyridinol. The HcgB catalyzed reaction is relatively promiscuous and can transfer a guanylyl group to a range of different 2,4-dihydroxyl-pyridines.[44] The structure of the physiological substrate of HcgB was predicted based on the structure of HcgB in complex with compound **3**, where this compound was prepared by light-inactivation of FeGP cofactor.[44] Furthermore, co-crystallization of HcgB with its substrate **2** indicated that even in the absence of the guanylyl moiety, **2** binds to the exact position that was observed in the case of co-crystallization with the product **3**.[37] Both structural analyses showed an interaction of Ser132 with the 6-carboxyl-methylene group of **2** and **3**, which suggested that this carboxyl group is present in the native pyridinol precursors of FeGP cofactor. However, HcgB could also be co-crystallized with 3,6-dimethyl-4-guanylyl-2-pyridinol, which was produced by using HcgB using 3,6-dimethyl-4-hydroxyl-2-pyridinol as the substrate and lacked the carboxyl group.[44] This result left the question of whether physiological pyridinol precursors have the carboxyl group open to further analysis.

HcgC

Based on the comparison of the crystal structure of HcgC with those of known enzymes, HcgC was hypothesized to be an *S*-adenosyl methionine (SAM) dependent methyl transferase. The structure of the possible substrate of HcgB (**2**) allowed the prediction of the substrate of HcgC as the 3-demethylated compound (**1**) that contains one 6-carboxyl-methylene group. Docking experiments suggested possible binding of compound **1** in the active site pocket of HcgC co-crystallized with *S*-adenosyl homocysteine (SAH), which is the demethylated product of SAM and the second product of HcgC. Biochemical characterizations using chemically synthesized compound **1** indicated that HcgC indeed catalyzes the methyl transfer reaction from SAM to **1**.[41] Based on the crystal structure of HcgC in complex with compound **1** and SAH, and mutational analyses, a catalytic mechanism involving keto-enol-tautomerism at the 2- and 4-hydroxyl groups was proposed.[43] Furthermore, an interaction of Thr176 (numbering based on *M. jannaschii* enzyme) with the 6-carboxyl-methylene group of compound **1** in co-crystal structure suggests that the carboxyl group is present in the pyridinol precursor and that **1** is the physiological precursor for biosynthesis of FeGP cofactor. The methyl group is thought to be transferred at the pyridinol level, as the binding site of HcgC is too narrow to accommodate the large GMP moiety[41] and HcgC is unable to catalyze the methylation of 6-carboxyl-methylene-5-methyl-

4-guanylyl-2-pyridinol, which is the 3-non-methylated version of **3**.[37] However, the HcgC reaction using a pyridinol precursor lacking the carboxyl group (3,5,6-trimethyl-4-hydroxyl-2-pyridinol) was not tested. Therefore, it cannot be excluded that the decarboxylated pyridinol compound can be the physiological substrate of the HcgC reaction.

HcgE

The primary structure of HcgE is similar to E1-like enzymes,[45] which catalyze the activation of the carboxyl terminus of ubiquitin or ubiquitin-like proteins by adenylylation with ATP.[48] Comparison of the crystal structures of the E1 enzyme in complex with ubiquitin-like protein and HcgE in complex with ATP suggested that HcgE catalyzes adenylylation of the possible precursor **3**. In the crystal structure of HcgE in complex with ATP and compound **3**, the carboxyl group of **3** locates near the alpha phosphate group of ATP in the active site. This finding suggested that the carboxyl group of the 6-carboxyl-methylene substituent of compound **3** is activated by adenylylation, which forms the activated intermediate **4**. The enzyme reaction was confirmed by mass spectrometry and the kinetic analyses of the production of pyrophosphate during the HcgE reaction.[45]

HcgF

HcgF shows structural similarity to nicotinamide mononucleotide deaminases,[49] which shows a propensity to bind compounds similar to **3**. Indeed a co-crystallization study showed compound **3** is bound to the protein and unexpectedly in only one subunit of the HcgF dimer in the crystal. The carboxyl group of compound **3** formed a thioester bond with the thiol of Cys9 (numbering based on the *M. jannaschii* HcgF). This finding indicated that the adenylylated product **4** reacts with HcgF to form the thioester bonded intermediate **5** although the thioester bond was formed even in absence of activation by HcgE in the crystal.[45] Activation steps catalyzed by HcgE and HcgF are similar to the two-step activation with adenylylation and subsequent transesterification reactions observed in ubiquitin and ubiquitin-like protein systems.[48, 50] The next step in this reaction sequence in the ubiquitin or ubiquitin-like system is usually the transfer of acyl to a strong nucleophile, e.g. hydride[51] or an amino group.[48] This may also be the case for the FeGP cofactor biosynthesis, but no study has been conducted yet to analyze this step, although the presence of an Fe(0)- or Fe(I)-carbonyl nucleophile to produce the acyl ligand was predicted.[45] These results support the presence of the carboxyl group in the pyridinol precursors, but they cannot finally prove its existence.

HcgD

HcgD belongs to the Nif3 protein family, which is wide spread in all domains of life, however their function is unknown currently.[52] The crystal structure of HcgD showed the presence of a di-nuclear iron center. One of the two iron atoms was labile to chelating agents, which could indicate the iron donating function in biosynthesis of FeGP cofactor. Change of UV/Vis spectra of the protein upon treatment with sodium dithionite indicated the presence of redox active iron.[46] The reduction increased the binding affinity of the iron to the protein, which was judged from the iron content after EDTA treatment. Another study indicated a possible DNA-binding activity of HcgD, but no target DNA sequence has been identified yet.[53] The function of HcgD might therefore be the transport of iron for the FeGP cofactor biosynthesis or regulatory function, but a definite proof has not been presented yet.

HcgA and HcgG

HcgA shows sequence similarity to BioB, a radical SAM (rSAM) enzyme;[19] however, the [4Fe-4S]-cluster binding motif of HcgA (CX_5CX_2C)[54] differs from canonical rSAM enzymes' binding motif (CX_3CX_2C or CX_2CX_4C).[55] The protein contains iron-sulfur clusters, but the electron-paramagnetic resonance (EPR) signal of this cluster did not unambiguously proof the structure of the iron-sulfur cluster. The EPR signal was strongly decreased by adding SAM. This finding together with the fact that HcgA can convert SAM to 5-deoxyadenosine (dAdo) in the presence of sulfur containing reducing equivalents indicates that HcgA is a rSAM enzyme.[54] HcgA shows sequence similarity to HydG, which is the protein responsible for the biosynthesis of the CO and CN ligands in [FeFe]-hydrogenase maturation and HydE, which is involved in formation of the bimetallic [FeFe]-hydrogenase metallocofactor.[56] HcgG is the most cryptic in the Hcg proteins and shows sequence similarity to fibrillarin-like proteins.[54]

As described in this chapter before, previous studies strongly suggested the most likely structure of biosynthetic pyridinol precursors and responsible Hcg enzymes. However, there is no direct evidence of incorporation of the precursors into the FeGP cofactor. In addition, the enzymes and their catalyzed reaction for biosynthesis of initial pyridinol precursor **1** and the final iron complex are unidentified. To answer these unresolved questions, we can learn from studies of biosynthesis of other metallocofactors, such as other hydrogenases and nitrogenases, since they also contain complex metallocofactors. In the next chapter, I will present the current knowledge of the biosynthesis of these cofactors, with a special focus on the methods and strategies used to obtain this knowledge, which might help us to investigate the FeGP cofactor biosynthesis.

Biosynthesis of other metallocofactors

The complex metallocofactors associated with hydrogenases and nitrogenases have beckoned the question how these molecules are assembled *in vivo* for a long time. These studies often combine protein biochemistry, structural biology and chemical analysis of cofactor intermediates. To further elucidate the FeGP cofactor biosynthesis, we can learn from studies on the biosynthesis of other metallocofactors from [NiFe]-, [FeFe]-hydrogenase and nitrogenases. In many cases, two major techniques have aided the understanding of biosynthesis pathways; identification of involved proteins by mutational analysis and development of an *in vitro* biosynthesis system to analyze intermediates and reactions catalyzed by each protein. During mutational analysis, strains are generated, which lack one or more potential biosynthesis genes. The absence of fully assembled metallocofactor in the mutated strains indicates an involvement of the gene in biosynthesis. In addition, detection of accumulated intermediates can give insights into the possible reaction catalyzed by the gene product. In a second step, the possible biosynthesis is reconstructed *in vitro* often by supplying a mutant cell extract with potential intermediates hypothesized to be produced by the deleted enzyme. In some cases, a fully defined *in vitro* biosynthesis is possible using heterologously produced biosynthesis enzymes and chemical substrates.

Biosynthesis of the cofactor of the [NiFe]-hydrogenase

The protein products of six genes, *hypA-F* and *slyD*, have been identified as being responsible for the formation of the [NiFe] cofactor.[5, 57] An overview of the suggested pathway is given in Figure 4. Briefly, the CN ligands are formed by HypE/F (**I**),[58] transferred to an iron on HypC/D (**III**)[59] and the CO is introduced (**II**).[60] The iron is transferred to hydrogenase apoenzyme (**IV**), the nickel is inserted by HypA/B and SlyD (**V**) [61] and finally maturation is completed by cleavage of the C-terminal extension by an endopeptidase.[62] The major common feature of [NiFe] compared to FeGP cofactor is the presence of diatomic ligands, which were also first identified by IR spectroscopy.[63] Insight into the origin of these ligands came first from the sequence of HypF, which showed a carbamoyl phosphate (CP) binding sequence and requirement for CP was proven by deletion of carbamoyl phosphate synthase, which abolished [NiFe]-cofactor biosynthesis.[64] HypF was predicted by sequence comparison to activate CP under ATP consumption, which suggests the production of CN and/or CO from CP.[65] The finding that HypF interacts with HypE opened the door to understand the role that CP plays in the biosynthesis of the [NiFe] cofactor.[66] The previously described activation of CP to AMP-C(O)NH_2 by HypF is followed up by transesterification to a cysteine residue in HypE. The precise residue was identified by MS/MS of digested HypE protein. The Cys-S-C(O)NH_2 group is then dehydrated by HypE under ATP consumption to form a Cys-S-CN group, from which the CN ligand can be transferred to the carrier.[58] Initially, it was believed that the CO ligand might also originate from CP. However, *in vivo* labelling experiments combined with IR spectroscopy showed

that incorporation of ^{13}C from acetate differs between CO and CN ligands[60]. Finally, using ^{13}C labelled CP produced *in vivo* from [^{13}C]-citrulline unambiguously demonstrated that CO and CN ligands have different metabolic origins.[67] The route of CO production in [NiFe]-hydrogenases from aerobic organisms has been elucidated, while it remains ambiguous in [NiFe]-hydrogenases from anaerobic organisms.[57a] In organisms growing under aerobic conditions an additional enzyme, HypX is present.[68] This enzyme shows sequence similarity to N^{10}-formyl-tetrahydrofolate- and coenzyme A (CoA)-dependent enzymes. The first compound was identified as origin of the CO ligands by *in vivo* labelling in combination with IR spectroscopy.[69] Indeed structural analysis of HypX showed that it contained CoA[70] and the CO production from N^{10}-formyl-tetrahydrofolate via formyl-CoA could be shown by measurement of CO release.[71] The investigation of the diatomic ligands of [NiFe]-hydrogenase cofactor helped to understand the biosynthesis of this metallocofactor drastically. Similarly, the investigation of these ligands also played a major aim in the investigation of the [FeFe]-hydrogenase cofactor biosynthesis described below.

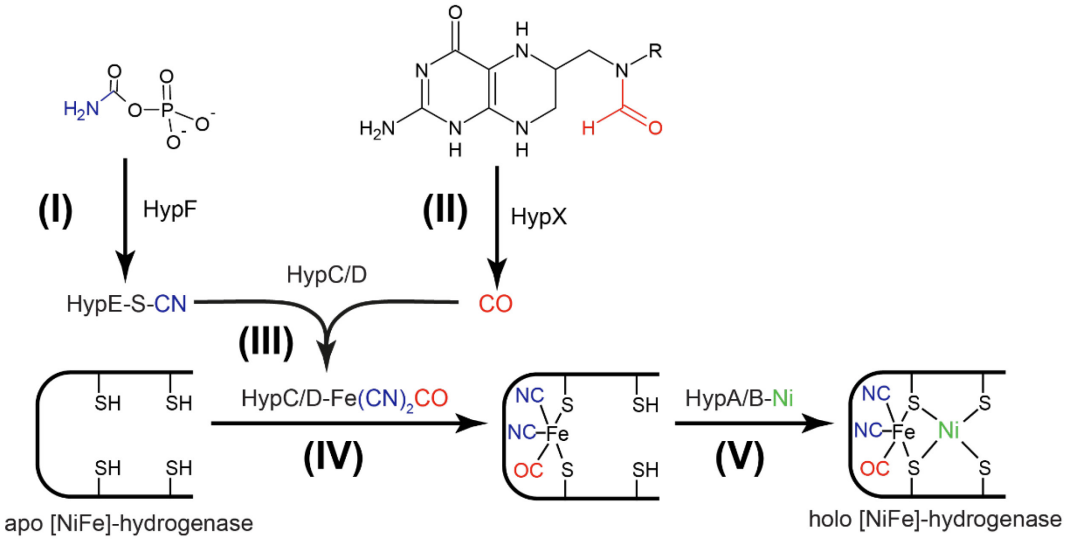

Figure 4. Schematic representation of the [NiFe]-hydrogenase cofactor biosynthesis in aerobic organisms. The final biosynthesis step, the endopeptidase catalyzed processing of the large subunit is not shown.

Biosynthesis of the H-cluster of [FeFe]-hydrogenases

[FeFe]-hydrogenase metallocofactor (H-cluster) contains two metal ions, an organic moiety and it is covalently linked to a conventional [4Fe-4S]-cluster. Initially, three biosynthesis enzymes were identified (HydE, HydF and HydG)[72] but in a recent report a fourth enzyme was indicated to play a critical role in biosynthesis of the H-cluster.[73] This biosynthesis starts with the rSAM enzyme HydG forming an Fe(II)[$(CO)_2$CN(cysteine)] synthon from tyrosine and cysteine (Figure 5 (**I**)).[74] Two of these complexes are fused by HydE[56b, 75] (Figure 5 (**II**)), ADT carbon and nitrogen are inserted by the H-protein[73] and finally the cofactor is inserted into hydrogenase apoenzyme (Figure 5 (**III**)). The function of HydF has not been fully understood to this moment, but it seems to be involved in these last steps.[56a] Two major method developments have defined the analysis of the biosynthesis of the [FeFe]-hydrogenase H-cluster, first an *in vitro* biosynthesis system for the H-cluster[76] and later the finding that chemically synthesized H-cluster analogs can spontaneously incorporated into hydrogenase apoenzyme.[15] The first indication of the function of HydG was given by sequence similarity to tyrosine lyases, like ThiH,[77] and the ability of HydG to split tyrosine into p-cresol.[78] Dehydroglycine, the second product of tyrosine lyases, was proposed to be split into CN and CO, and indeed CN and dehydroglycine production could be verified by biochemical experiments.[74e] The development of *in vitro* biosynthesis system employing heterologously produced HydG/E/F and *E. coli* cell extracts enabled testing this hypothesis in the context of biosynthesis of the H-cluster.[76b, 79] Tyrosine, but also cysteine, the function of which became apparent later, improved *in vitro* biosynthesis yields[76a] and using site specific labelled tyrosine combined with IR spectroscopy showed that the CO and CN ligands of the H-cluster originate from the carboxylic acid and amino group.[80] Utilizing isotopically labelled tyrosine in the *in vitro* biosynthesis in combination with spectroscopy was also used to analyze the mechanism of HydG, which was thought to result in the diatomic ligands bound to the [4Fe-4S]-cluster of HydG.[74b, 74c] The crystal structure of HydG showed that the ligands actually bind to a fifth iron.[74a] Subsequently cysteine,[81] whose involvement was already indicated by *in vitro* biosynthesis,[76a] was found to ligate this Fe(II)[$(CO)_2$CN(cysteine)] synthon. The involvement of this intermediate was strengthened by the replacement of HydG in the *in vitro* biosynthesis by a chemical analogue of this complex.[82]

The analysis of the second rSAM enzyme HydE followed a similar route, first *in vitro* biochemical experiments indicated that it can form carbon-sulfur bonds[83] and indeed the cysteine sulfur was shown to be the precursor of sulfur of the ADT bridge in a reaction catalyzed by HydE.[82] The enzyme was thought to generate bimetallic cluster by fusing two Fe(II)[$(CO)_2$CN(cysteine)] synthons provided by HydG[56b] and indeed HydE could be replaced by a chemically synthesized cluster in *in vitro* biosynthesis.[56a] In this final *in vitro* biosynthesis, only HydF was required, but still cell extract components of *E. coli* could not be omitted from the reaction. This led to further investigations of necessary compounds from the cell extracts. A

recent study showed the involvement of the glycine cleavage system in *in vitro* biosynthesis of the H-cluster, in which the aminomethyl group of this system was the precursor of nitrogen and carbon moieties of the ADT-bridge.[73] This last study showed for the first time a fully defined *in vitro* biosynthesis system for the H-cluster. Such an *in vitro* biosynthesis system had been described for FeMoCo of nitrogenases, which I will introduce below and has greatly aided the analysis of nitrogenase cofactor.

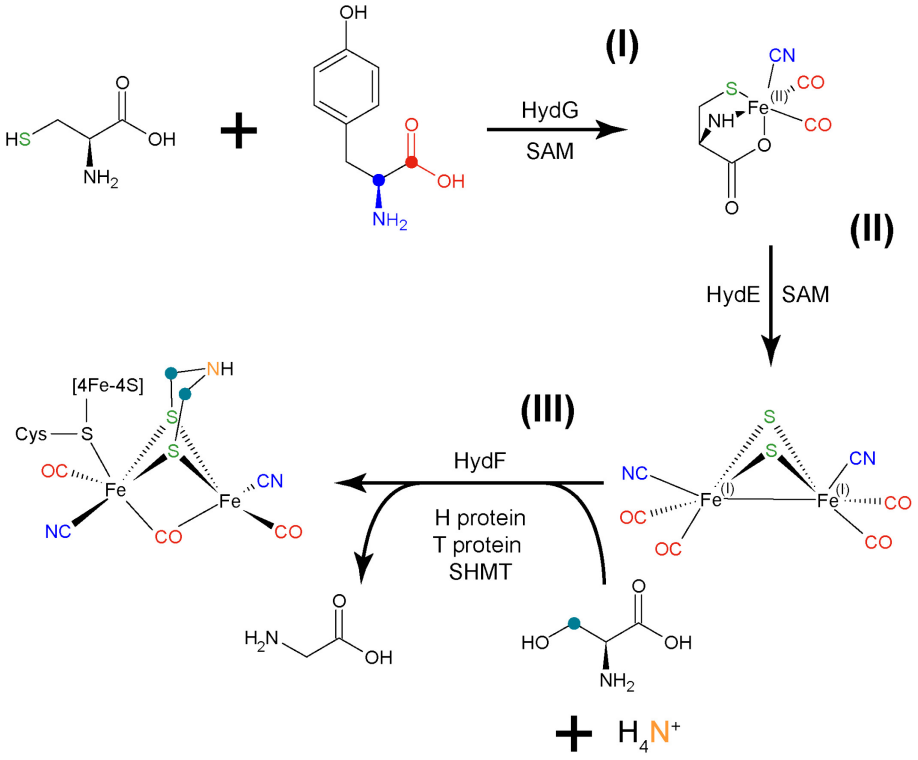

Figure 5. Biosynthesis of the H-cluster of [FeFe]-hydrogenase. The Lipoyl-H-Protein (H protein), aminomethyltransferase (T protein) and serine hydroxymethyltransferase (SHMT) are part of the glycine cleavage system.

Biosynthesis of FeMoCo of nitrogenase

FeMoCo of nitrogenases was first described as an acid-extractable compound from holo-nitrogenases, which activates nitrogenases apoenzyme from a *nifB* deletion strain,[84] similar to the finding of the FeGP cofactor.[24] The Fe-S-backbone[85] and central carbon[86] of this cluster is formed by NifB, transferred to NifEN, where Mo and homo-citrate are added.[87] The complete cofactor is then transferred to nitrogenase apoenzyme. *In vitro* biosynthesis systems of FeMoCo were developed almost 40 years ago by combining cell extracts of mutants, without nitrogenase activity, to recover nitrogenase activity.[88]

The development of a fully defined *in vitro* biosynthesis assay based on purified proteins and substrates enabled a wide range of studies.[89] Similar to

investigations of H-cluster biosynthesis, the proposed functions of biosynthetic enzymes could be tested by their replacement with their products in the *in vitro* biosynthesis assay.[90] One example is NifV enzyme, which provides the organic ligand, *R*-homocitrate. This compound was identified by using *in vitro* biosynthesis system to purify this compound from cell extract and it was then analyzed by NMR and mass spectrometry. The final proof was again based on *in vitro* biosynthesis, where adding chemically synthesized *R*-homocitrate could replace the need for cell extract containing this compound or NifV.[91] Similarly NifQ was shown to be the Mo donor, by replacing Mo in *in vitro* biosynthesis with heterologously purified NifQ.[92]

These investigations show that while the metallocofactors of [NiFe]-,[FeFe]-hydrogenase and nitrogenase have different structures the investigation of their biosynthesis pathways shows similar methodological trends. Based on these prior findings I will outline my aim for my study of the FeGP cofactor biosynthesis below.

Aim of this project

The investigation of structure and function of FeGP cofactor and its biosynthesis have advanced our understanding of this complex and fascinating molecule greatly in the past decade. Based on the studies, the function of more than half of biosynthetic proteins involved in the biosynthesis was elucidated and the structure of pyridinol precursors was proposed. These studies have mostly relied on the structure-to-function methodology employing detailed structural and biochemical studies of Hcg enzymes. However, to fully understand FeGP-cofactor biosynthesis, information about iron insertion, acyl- and CO ligand formation, as well as the synthesis of the initial pyridinol precursor **1** are lacking for instance. To address these questions, a complementary method to structure-to-function method is required. As the analysis of biosynthesis of nitrogenase and [FeFe]-hydrogenase metallocofactor was greatly aided by the development of an *in vitro* biosynthesis method, a similar technique might also help us to answer the questions mentioned above. In my PhD project, I developed an *in vitro* system for the study of biosynthesis of FeGP cofactor using a mixture of cell extracts of methanogenic archaea and precursor compounds. Using cell extract from mutated methanogens lacking *hcg* genes, I was able to detect the production of FeGP cofactor by mass spectrometry and by enzymatic activity of reconstituted [Fe]-hydrogenase in the *in vitro* system without interference by endogenous FeGP cofactor or [Fe]-hydrogenase. By using this method, I verified that the proposed pyridinol precursors really function for biosynthesis of the FeGP cofactor and that the carboxyl group of the pyridinols is the precursor of the acyl-ligand. In addition, it is also indicated that the biosynthesis utilizes CO gas and an unknown compound as the source of CO ligand and that the biosynthesis reactions require an unknown electron carrier in the cell extract. This *in vitro* biosynthesis method allowed us to verify the function of HcgA and HcgG.

2 Materials and Methods

2.1 Materials

2.1.1 Chemicals

All chemicals used in this work, if not noted otherwise, were obtained from Carl Roth GmbH + Co. KG (Karlsruhe) or Sigma-Aldrich Chemie GmbH (Darmstadt). The purity grade of the chemicals was stated “for analysis”, “for biochemistry” or “for synthesis”. Deionized water had a final conductivity of 18 mΩ.

2.1.2 Gases

Nitrogen (5.0) was procured from Air liquide Deutschland GmbH (Düsseldorf) or Westfalen AG (Münster). Carbon dioxide (4.5) 20 % / Hydrogen (5.0) 80 % gas mixture was obtained from Westfalen. Carbonmonoxide (4.7) was procured from Messer SE & Co. KGaA (Bad Soden am Taunus) and nitrogen 95 % / hydrogen 5 % mixture from Westfalen or Air liquide.

2.1.3 Stable isotopes

[^{18}O]-H_2O was obtained from Sigma-Aldrich Chemie GmbH with a purity of 97 % ^{18}O. [^{13}C]-CO was procured from Cambridge Isotope Laboratories Inc. (Andover, USA) with a purity of 99% ^{13}C and < 2 % ^{18}O. [^{15}N]-NH_4Cl was procured from Cambridge Isotope Laboratories Inc. at 99 % ^{15}N.

2.2 Anaerobic experiments

2.2.1 Preparation of anaerobic solutions

For the work with enzymes, compounds and organisms, which are affected by oxygen, the removal of oxygen from all used solutions is necessary. To eliminate the oxygen, the solutions were brought to boiling in an electric kettle, if no heat sensitive compounds were added. Then the solution was transferred to a glass bottle (Schott AG) and outgassed with nitrogen for at least five minutes. The solution was then closed air tight by a butyl rubber stopper, which was held in place by the glass bottle’s plastic cap. After cooling the bottles were transferred into the anaerobic chamber (see below) and stirred at 600 rotations per minute (rpm) to remove trace of oxygen still present for at least 12 h.

In case of heat sensitive solutions, water was prepared as described above and the substance dissolved in anaerobic water inside the anaerobic chamber.

2.2.2 Operation of the anaerobic chamber

The anaerobic chamber (Coy Laboratory Products Inc.) consists of a vinyl chamber, which is sealed air tight and connected to an air lock. Inside of the chamber the atmosphere is replaced by 95% N_2/ 5% H_2. Further a platinum catalyst converts trace oxygen with the hydrogen to form water, which is bound by silica gel. Both the silica gel and the catalyst need to be regenerated in regular intervals by heating at 150 °C and 180 °C for at least 3 h, respectively. Material transfer is conducted via the air lock, which is connected to a 95% N_2/ 5% H_2 gas line and a vacuum pump, by exchanging the atmosphere inside the airlock for at least three times. Inside the anaerobic chamber different machines, such as Aekta FPLC systems or centrifuges are used for protein purification and enzyme assay preparation.

2.3 Cloning

2.3.1 Construction of expression systems for Hcg proteins from various methanogens

The genes for Hcg proteins from the given methanogenic organisms were synthesized by Genscript Biotech Corp. Synthesized DNA fragments were cloned into the corresponding plasmids by given restriction enzymes (Table 1). The resulting plasmids were transformed into *Escherichia coli* BL21(DE3).

Table 1. Overview of expression systems for Hcg proteins.

Encoded Hcg proteins	Organism	Accession number	Plasmid	Restriction enzymes
HcgB	*Methanococcus maripaludis*	WP_011171441	pET 24b(+)	*Nde*I / *Xho*I
HcgB	*Methanocaldococcus jannaschii*	WP_064496510	pET 28b(+)	*Nco*I / *Xho*I
HcgC	*Methanococcus maripaludis*	WP_011171442	pET 24b(+)	*Nde*I / *Sal*I
HcgD	*Methanococcus maripaludis*	WP_181491542	pET 28b(+)	*Nde*I / *Xho*I
HcgE	*Methanothermobacter marburgensis*	WP_013296321	pET 24b(+)	*Nde*I / *Sal*I
HcgF	*Methanococcus maripaludis*	WP_181487750	pET 28b(+)	*Nde*I / *Hind*III
HcgF	*Methanocaldococcus jannaschii*	WP_010870765	pET 28b(+)	*Nde*I / *Hind*III
HcgF C9A	*Methanocaldococcus jannaschii*	Based on WP_010870765	pET 28b(+)	*Nde*I / *Hind*III
HcgF C119A	*Methanocaldococcus jannaschii*	Based on WP_010870765	pET 28b(+)	*Nde*I / *Hind*III

2.4 Heterologous production and purification of proteins

2.4.1 Heterologous production and purification of HcgB, HcgC and HcgD

The *E. coli* BL21(DE3) strain containing the HcgB gene from *Methanococcus maripaludis* was grown aerobically at 37 °C in 2 L LB medium with 50 µg/ml Kanamycin. The cell culture was agitated by stirring at 600 rpm). When the cell density reached an OD of 0.8 at 600 nm (OD_{600}), protein production was initiated by adding isopropyl β-D-1-thiogalactopyranoside (IPTG, final concentration of 1 mM) and the cultivation was continued at 37 °C for 3 – 4 h. The cells were harvested by centrifugation at 7,300 rpm with a Beckmann JLA 10.500 rotor at 4 °C for 20 min, washed once with phosphate buffered saline (PBS, 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 20 mM KH_2PO_4) and stored at −20 °C until usage.

The cells were resuspended in 50 mM potassium phosphate buffer pH 7 containing 500 mM KCl and 20 mM imidazole and cell lysis was conducted by sonication using a SONOPULS GM200 (Bandelin) with a KS 76 probe at 50% cycle and 160 W for 1 min for a total of ten times on ice, with 1 min pauses in between. Cell debris and particulate proteins were removed by ultracentrifugation using a Sorvall T647.5 rotor at 30,000 rpm at 4 °C for 30 min. The supernatant was filtered by a 0.45 µm filter and loaded onto a HisTrap column (GE Healthcare) in 50 mM potassium phosphate buffer pH 7 containing 500 mM KCl and 20 mM imidazole. The column was washed with 10 column volumes (CV) of the same buffer and the protein was eluted in a gradient of 20 to 500 mM imidazole in 50 mM potassium phosphate buffer pH 7 containing 500 mM KCl over 8 CV. The protein fractions were collected and desalted by a HiPrep 26/10 desalting column in 50 mM potassium phosphate pH 7 containing 300 mM KCl. The protein was finally concentrated and frozen in liquid N_2 and stored at −75 °C. HcgC from *M. maripaludis* was prepared and stored analogously.

HcgD from *M. maripaludis* was produced in an *E. coli* BL21 (DE3) strain harboring the plasmid containing the HcgD gene (Table 1). The cells were grown at 37 °C in 2 L LB medium with 50 µg/ml Kanamycin until an OD_{600} of 1.0 is reached as described above. The protein production was initiated by adding to a final concentration of 1 mM and the cells were grown at 37 °C for an additional 3 – 4 h. After harvesting by centrifugation (7,300 rpm in a Beckmann JLA 10.500 rotor at 4 °C for 20 min) the cells were washed once by PBS and stored at −20 °C. Protein purification was conducted as described above for the purification of HcgB.

2.4.2 Heterologous production and anaerobic purification of HcgF wild type and its C119A mutated proteins from *Methanocaldococcus jannaschii*

The strain of *E. coli* BL21(DE3) containing the respective plasmid of HcgF from *Methanocaldococcus jannaschii* was cultivated in 2 L LB medium supplemented with 50 µg/ml kanamycin with stirring at 600 rpm until an OD_{600} of 1.0 is reached

and the protein production was initiated by adding 1 mM IPTG final concentration. The cells were cultivated at 37 °C for 3 – 4h and harvested by centrifugation (7,300 rpm with Beckmann JLA 10.500 rotor at 4 °C for 20 min). The cells were washed once with PBS and stored frozen at –20 °C until use. To avoid oxidation of the cysteine proposed to be involved in FeGP cofactor biosynthesis[45] the following steps were conducted using anaerobic buffers in the anaerobic chamber (95% N_2/ 5% H_2).

The cells were resuspended in 50 mM Tris/HCl pH 7 containing 500 mM KCl, 20 mM imidazole and freshly added 1 mM Tris(2-carboxylethyl)phosphine (TCEP). Cell lysis was conducted by sonication on ice using a KS76 probe (SONOPULS GM200) at 50% for 1 min with 1 min pause, which was repeated ten times. Cell debris and membrane fragments were removed by ultracentrifugation using a Sorvall T647.5 rotor at 30,000 rpm at 4 °C for 30 min. The supernatant was applied to a 5 ml HisTrap column (GE healthcare) in the buffer as described above at a flow of 2 ml/min. The column was washed by 10 CV of the same buffer at 5 ml/min and the protein was then eluted using 20 to 500 mM imidazole gradient in 50 mM Tris/HCl pH 7 containing 500 mM KCl and 1 mM TCEP in 10 CV at 5 ml/min. The protein profiles of the eluted fractions were evaluated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and the fractions containing the corresponding HcgF protein were pooled and desalted using a HiPrep 26/10 desalting column with 50 mM Tris/HCl pH 7 containing 300 mM KCl and 1 mM TCEP. The desalted solution was concentrated, frozen in liquid N_2 and stored at –75 °C.

The *hcgE* gene from *M. marburgensis* on a pET28b(+) plasmid was expressed in an *E. coli* BL21 (DE3) strain. The cells were grown at 37 °C in 2 L LB media and 50 µg/ml kanamycin with stirring at 600 rpm. When an OD_{600} of 0.6 is reached, 1 mM (final concentration) IPTG was added. After further incubation at 37 °C for 4 – 6 h, the cells were harvested by centrifugation (7,300 rpm with Beckmann JLA 10.500 rotor at 4 °C for 20 min).

The following steps for protein preparation were conducted in an anaerobic chamber using anaerobic buffers. The cells were resuspended in 50 mM potassium phosphate pH 7 containing 500 mM KCl, 20 mM Imidazole and 1 mM TCEP. Cell lysis was conducted by sonication with a KS76 tip at 50% cycle and 160 W for 1 min with 1 min of pause. This step was repeated for a total number of ten times on ice. After ultracentrifugation (Sorvall T647.5 rotor at 30,000 rpm at 4 °C for 30 min) the supernatant was applied to a 5 ml HisTrap column (GE Healthcare) equilibrated with the same buffer. The column was washed with 5 CV of the loading buffer and the protein eluted with a gradient of 20 to 500 mM imidazole in 50 mM potassium phosphate pH 7 containing 500 mM KCl and 1 mM TCEP over 20 CV. The protein containing fractions were analyzed by SDS-PAGE and the fractions containing HcgE were pooled. Afterwards the pooled fractions were desalted with 50 mM potassium phosphate buffer at pH 7 containing 500 mM

KCl and 1 mM TCEP with HiPrep 26/10 desalting columns. The desalted solutions were concentrated with ultrafiltration (30 kDa cut-off). The concentrated samples were frozen in liquid N_2 and stored at −75 °C.

2.4.3 Construction of expression system for [Fe]-hydrogenase apoenzymes

The gene of the [Fe]-hydrogenase apoenzyme from *M. jannaschii* was amplified by PCR from extracted genomic DNA of *M. jannaschii* and cloned into a pET28b plasmid.[24] The *hmd* gene from *M. maripaludis* (WP_011170071.1 or Genbank accession number NC_005791.1) was synthesized and cloned into a pET28b(+) vector using *Nco*I and *BamH*I by Genscript. The construct contained an N-terminal strep tag and was transformed into *E. coli* BL21(DE3).

2.4.4 Heterologous production and purification of [Fe]-hydrogenase apoenzyme from *M. jannaschii*

The *E. coli* BL21(DE3) cells harboring the described plasmid were grown in 2 L tryptone-phosphate (TP) medium with 50 µg/ml Kanamycin at 37 °C with agitation of 600 rpm.[93] When OD_{600} of 0.6 is reached, protein production was induced by addition of 1 mM (final concentration) IPTG and further grown at 37 °C for 3 h. The cells were harvested by centrifugation using a Beckmann JLA 10.500 rotor at 7,300 rpm for 20 min at 4 °C and stored at −20 °C. The cells were resuspended with 50 mM 3-(N-morpholino)propanesulfonic acid (MOPS)/KOH pH 7, containing 2 mM dithiothreitol (DTT) and lysed twice by sonication with a KS76 probe at 50 % cycle and 160 W for 5 min with 5 min pause on ice. Membrane fragments and cell debris were removed by ultracentrifugation using a Sorvall T647.5 rotor at 30,000 rpm at 4 °C for 30 min. The samples were incubated at 70 °C for 15 min and the precipitated proteins were removed by centrifugation (JA 25-50 rotor at 13,000 rpm at 4 °C for 20 min). Ammonium sulfate was added to a final concentration of 2 M and removal of precipitate by centrifugation was repeated. The supernatant was loaded onto a self-packed Phenyl Sepharose column at 1 ml/min in 50 mM MOPS/KOH pH 7 containing 2 M ammonium sulfate, 2 mM dithiothreitol (DTT). The column was washed with 10 CV of the same buffer and the protein was eluted with a gradient of 2 to 0 M ammonium sulfate in 50 mM MOPS pH 7 containing 2 mM DTT over 10 CV. The protein containing fractions were identified by SDS-PAGE and fractions with [Fe]-hydrogenase were desalted in 50 mM MOPS pH 7 by ultrafiltration or a HiPrep 26/10 desalting column. The protein fractions were pooled and finally concentrated and stored at -75 °C after freezing in liquid N_2.

2.4.5 Purification of Strep-tagged [Fe]-hydrogenase apoenzyme from *M. maripaludis*

The *E. coli* BL21(DE3) strain transformed with the expression vector containing the Strep-tagged *hmd* gene from *M. maripaludis* was grown in 2 L TP medium containing 50 µg/ml kanamycin with agitation at 600 rpm. When OD_{600} of 1.0 was

reached, the gene expression was induced by addition of 1 mM (final concentration) IPTG and cell growth was continued at 37 °C for 4 – 6 h. The cells were collected by centrifugation with a Beckmann JLA 10.500 rotor at 7,300 rpm and 4 °C for 20 min and stored at −20 °C until further use.

The cells were resuspended in 150 mM Tris/HCl pH 8 containing 100 mM NaCl, 2 mM DTT and 5 % (v/v) glycerol and lysed by sonication. The sonication was conducted on ice with a KS76 tip and a SONOPULS GM200 at a 50 % cycle with 160 W for ten times 1 min with 1 min pauses. After ultracentrifugation (Sorvall T647.5 rotor at 30,000 rpm at 4 °C for 30 min). The supernatant was applied to two connected 5 ml Streptrap columns (GE Healthcare) equilibrated with 150 mM Tris/HCl pH 8 containing 100 mM NaCl, 2 mM DTT and 5 % (v/v) glycerol at 1 ml/min. After washing the column with 10 CV of the same buffer, the protein was eluted by 5 CV of the same buffer containing 2.5 mM desthiobiotin. The buffer was then exchanged to 150 mM Tris/HCl pH 8 containing 100 mM NaCl, 2 mM DTT and 5 % (v/v) glycerol. The desalted solution was concentrated with ultrafiltration (30 kDa cut-off) and frozen in liquid N_2 and stored at −75 °C.

2.5 Cultivation of methanogens

2.5.1 Cultivation of *Methanococcus maripaludis*

Methanococcus maripaludis S2 ΔuptΔfrcΔfruΔhcgB (Mm1334) was obtained from Kyle Costa (University of Minnesota).[40] The strain lacks the ability to produce HcgC due to the deletion of the ribosome binding site by deletion of the overlapping hcgB gene. Therefore, this strain is referred as *ΔhcgBΔhcgC* strain in the following text. The *M. maripaludis ΔhcgBΔhcgC* strain was cultivated using a modified Balch medium (Table 2) at 37 °C with formate as the substrate and 80% N_2/ 20% CO_2 atmosphere in 0.5 or 5 l scale.[94] The cells were grown until OD_{600} of 0.6 – 0.7 was reached and pH was adjusted to 7 by adding formic acid at least every 8 h as well as directly before harvesting. The cells were harvested with a continuous flow centrifuge (Heraeus 3049 continuous flow rotor at 15,000 rpm at 4 °C) and the cell pellet was resuspended in cell medium and centrifuged a second time with Beckmann JLA 10.500 rotor at 7,300 rpm and 4 °C.

To store the intact cells the cell pellet was dissolved in 1 g cell wet weight per 2 ml of medium and centrifuged in these aliquots in a Megafuge 16 (Thermo Scientific) at 4,000 rpm at 8 °C for 20 min. The supernatant was discarded and the cells were stored at −20 °C. For the cell lysis 2 ml of 50 mM Tris/HCl pH 7.5 containing 5 mM $MgCl_2$ and 2.5 U/ml DNase I (Invitrogen) were added per 1 g of cell wet weight to the frozen cells, where the marine methanogen cells were disrupted by osmotic shock.

For storage of lysed cells they were resuspended in 50 mM Tris/HCl pH 7.5 containing 5 mM $MgCl_2$ and 2.5 U/ml DNase I (Invitrogen) after the centrifugation

in the Beckman rotor mentioned above in a ratio of 1 g cell wet weight to 2 ml buffer. The lysed cells were then frozen in liquid N_2 and stored at −20 °C.

Table 2. Components of a modified Balch medium for *M. maripaludis* culture.[94]

Compound	Concentration
Amino acids[1]	0.15 g/l of each amino acid
Cysteine	0.5 g/l
$NaHCO_3$	60 mM
Sodium acetate	10 mM
Sodium formate	200 mM
NaCl	180 mM
$FeSO_4$	30.0 µM
K_2HPO_4	0.8 mM
KCl	4.5 mM
$MgCl_2$	13.5 mM
$MgSO_4$	14.0 mM
$CaCl_2$	1.0 mM
NH_4Cl	9.5 mM
Tris/HCl pH 7	100 mM
$NiCl_2$	1 µM
Resazurin	20 mM
Na_2S	1 mM
Trace mineral solution[2]	0.1 %
Vitamin solution[3]	1.0 %

[1]Nineteen proteinogenic amino acids excluding cysteine.
[2]See Table 3.
[3]See Table 4.

Table 3. Components of trace mineral solution for cultivation of *M. maripaludis.*

Components	Concentration (mg/l)
$Na_3Citrate*2H_2O$	21
$MnSo_4*2H_2O$	5
$CoSO_4$	1
$ZnSO_4.5H_2O$	1
Na_2SeO_3	2
$V(III)Cl_3$	0.1
$Na_2WO_4*2H_2O$	0.033

Table 4. Components of vitamin solution for cultivation of *M. maripaludis.*

Component	Concentration (mg/l)
Biotin	2
Folic acid	2
Pyridoxine-HCl	10
Thiamine-HCl	5
Riboflavin	5
Nicotinic acid	5
D-Ca-pantothenate	5
Vitamin B12	0.1
p-Aminobenzoic acid	5
Lipoic acid	5

2.5.2 Cultivation of *Methanothermobacter marburgensis*

M. marburgensis was cultivated in a 10 l fermenter with a minimal medium containing 50 mM KH_2PO_4, 40 mM NH_4Cl, 24 mM Na_2CO_3, 0.5 mM nitrilotiracetic acid, 0.2 mM $MgCl_2$, 1 mM $CoCl_2$, 1 mM Na_2MoO_4, 50 mM $FeCl_2$ and 20 mM resazurin. Ni^{2+}-sufficient conditions additionally contained 5 µM $NiCl_2$. The media was prepared aerobically and reduced by bubbling with 80% H_2/ 20% CO_2/ 0.1% H_2S at 65 °C with a gas flow rate of 1000 ml/min and an agitation speed of 1000 rpm.[30] Once the medium was reduced and the temperature reached 65 °C, 150 ml stored culture (see below) was inoculated with a 50 ml plastic syringe. The cells grew for around 20 hours until an OD_{600} of 6 – 7 was reached and then the culture was cooled down to ca. 4 °C. The cells were harvested by continuous flow centrifugation using Heraeus 3049 continuous flow rotor at 15,000 rpm at 4 °C under N_2 atmosphere. After harvesting, the gas tight rotor was transferred to the anaerobic chamber. The turbid supernatant (OD_{600} = ca. 10) in the rotor was transferred to a glass bottle and stored at 8 °C under 80% H_2/ 20% CO_2. The cell suspension was active for at least three months and was also used as an inoculum for later cultures. The cell pellets were transferred into glass bottles and stored under N_2 at −75 °C.

2.6 Native purification of [Fe]-hydrogenase from the nickel-limiting cells of *M. marburgensis*

All steps are performed in the anaerobic chamber using anaerobic buffers under yellow light. *M. marburgensis* cells from a nickel-limiting culture were resuspended in 2 ml 50 mM potassium phosphate buffer pH 7 per 1 g of cells. The cells were lysed by sonication using a SONOPULS GM200 with a KE76 tip in ice water at 50% cycle and 160 W for 8 min. After a pause of 7 min this step was repeated for a total number of six times. After ultracentrifugation using a Sorvall T647.5 rotor at 30,000 rpm at 4 °C for 30 min, ammonium sulfate powder was added to the

supernatant to 60 % of ammonium sulfate saturation, while the solution was agitated by stirring on ice. After incubation on ice without agitation for 20 min, the aggregated protein was removed by centrifugation at 20,000 rpm in a Sorvall T647.5 rotor at 4 °C for 30 min. Ammonium sulfate powder was added to the supernatant to 90 % of ammonium sulfate saturation with agitation by stirring and then incubated without agitation for 20 min. After the centrifugation step under the same condition for 60% of ammonium sulfate saturation precipitation, the protein pellet was dissolved in 50 mM MOPS pH 7. This solution was then dialyzed against 50 mM sodium citrate buffer pH 5 with a molecular weight cut off of 30 kDa at 8 °C overnight. The dialyzed protein solution was centrifuged at 20,000 rpm in a Sorvall T647.5 rotor at 4 °C for 30 min. The supernatant was applied to a Source 30Q column equilibrated with 50 mM sodium citrate buffer pH 5 and washed with 250 ml of the equilibrating buffer containing 200 mM NaCl. Proteins were eluted in a gradient from 200 to 500 mM NaCl in 50 mM sodium citrate buffer pH 5 over 500 ml. The protein fractions containing [Fe]-hydrogenase were neutralized with 1 M MOPS pH 7 and NaOH solution, concentrated, frozen in liquid N_2 and stored at −75 °C.

2.7 Extraction of the FeGP cofactor from [Fe]-hydrogenase

2.7.1 MeOH/2-ME extraction

For the extraction of the FeGP cofactor from [Fe]-hydrogenase the enzyme was used at a maximal concentration of 4 mg/ml. To the protein 60% MeOH, 1 mM 2-mercaptoethanol and 1% (v/v) NH_3 (final concentrations) were added. After incubation at 40 °C for 15 min denatured protein was removed by ultrafiltration (10 kDa cut-off). The sample was concentrated by evaporation at 4 °C and the dried material was dissolved in 10 mM $(NH_4)_2(CO_3)$ pH 9 containing 1 mM 2-mercapothanol.

2.7.2 Acetic acid extraction

The [Fe]-hydrogenase at a final concentration of less than 1 mg/ml was diluted with 1:1 with acetic acid. The precipitate was removed by centrifugation at 14.000 rpm for 10 min. Then the residual denatured protein was removed by ultrafiltration (10 kDa). After concentration by evaporation at 4 °C the sample was dissolved in water or 50 mM sodium acetate pH 4.6.

2.8 Single enzyme assays

2.8.1 [Fe]-hydrogenase activity assay

The [Fe]-hydrogenase activity was photometrically determined by measuring the conversion of methylene-tetrahydromethanopterin (methylene-H_4MPT) to methenyl-H_4MPT^+. 20 µM methylene-H_4MPT was added to 0.7 ml of 120 mM

potassium phosphate buffer at pH 6 containing 1 mM EDTA in a 1 ml cuvette (1 cm light pass length) under N_2 atmosphere. The enzyme reaction was started by adding 10 – 20 µl enzyme solution to the assay solution. The rate of increased absorbance at 336 nm was recorded caused by forming of methenyl-H_4MPT$^+$ (ε_{336} nm = 21.6 $mM^{-1} \cdot cm^{-1}$). One U was defined as the conversion of 1 µmol of methylene-H_4MPT to methenyl-H_4MPT$^+$ per min. The protein concentration was determined by the Bradford assay with bovine serum albumin as a standard using the assay kit (Bio-Rad Protein Assay (#5000006) from Bio-Rad). The activities are given as U/mg protein.

2.8.2 HcgB activity assay

To test the activity of heterologously produced HcgB, the conversion of **2** to **3** was quantified by HPLC based on UV-detection of standards both compounds. The compounds' structures are shown in Figure 3. The standard assay reaction solution contained 100 µM **2**, 5 mM GTP/$MgCl_2$ and 1 µM HcgB from *M. maripaludis* (mpHcgB) in 50 mM MOPS/KOH pH 7 in a volume of 0.1 ml. The reaction was conducted at 40 °C for 1 h and was stopped by addition of 40 µM acetic acid (final concentration) and incubation on ice for 20 min. The denatured protein was removed by centrifugation (14,500 rpm in an Eppendorf MiniSpin plus) and filtration (0.2 µm pore size). The reaction was analyzed on an HPLC equipped with a Synergi 4m Polar RP 80A column (250 mm × 4.6 mm, Phenomenex) with acidified water (HCl, pH 3)/methanol gradient (Table 1) detected via UV absorbance.

Table 5. HPLC gradient used for the purification of **3** and quantification of **1** and **2**.

Time (min)	MeOH moiety / % (v/v)
0	0
9	0
36	98
41	98
46	0
61	0

2.8.3 Competition assay of the reaction catalyzed by HcgB

The activity of HcgB with two similar substrates, 6-carboxyl-methenyl-5-methyl-2-pyridinol (**1**) and **2**, was measured in a mixture of 100 µM **1**, 100 µM **2**, 5 mM GTP, 5 mM $MgCl_2$, and 1 µM mpHcgB in 50 mM MOPS/KOH pH 7 at 40 °C for 1 h. The sample was prepared and analyzed by HPLC as described for the enzyme assay of HcgB (see chapter 3.8.2).

2.8.4 HcgC reaction

The conversion of pyridinol compound **1** to its 3-methylated product **2** was measured by HPLC to detect the activity of heterologously produced HcgC. The standard assay reaction contained 100 µM **1**, 1 mM *S*-adenosyl methionine (SAM) and 1 µM HcgC from *M. maripaludis* (mpHcgC) in 50 mM MOPS/KOH pH 7 in a volume of 0.1 ml. The reaction mixture was incubated at 40 °C for 1 h. The following steps for the HPLC analysis were conducted in the same manner as described for the HcgB reaction in chapter 3.8.2.

2.8.5 Combined HcgB and HcgC reactions

For monitoring combined activity of HcgB and HcgC on the pyridinol compounds, a mixture containing 100 µM pyridinol compound, 5 mM GTP, 5 mM $MgCl_2$, 1 mM SAM, 1 µM mpHcgC and 1 µM mpHcgB in 50 mM MOPS/KOH pH 7 in a volume of 0.2 ml was used. The assay was carried out analogously to the HcgB assay (chapter 3.8.2). Possible products were analyzed by matrix assisted laser desorption ionization − time of flight mass spectrometry (MALDI-TOF-MS) using α cyano-4-hydroxycinnamic acid as matrix.

2.8.6 Analysis of the binding of compound 3 to HcgF by ultrafiltration

The interaction of compound **3** with HcgF was verified via ultrafiltration using reaction solutions containing 150 µM jHcgF form *M. jannaschii*, 75 or 150 µM **3**, 10 µM mHcgE from *M. marburgensis*, 10 mM ATP, 5 mM $MgCl_2$ in 25 mM 2-(*N*-morpholino) ethanesulfonic acid (MES)/KOH pH 7, 25 mM potassium phosphate buffer pH 7, 0.15 M KCl and 1 mM DTT.

After incubation at 65 °C for 15 min, the reaction solution was filtrated with 10 kDa cut-off (Millipore Amicon-Ultra 0.5 10 kDa). The concentration of **3** in the filtrate and concentrate was determined by measuring the UV-Vis absorbance at 300 nm (ε_{300} = 9 cm^{-1} mM^{-1}).[25]

2.8.7 Analysis of the HcgE catalyzed reaction

The activation of HcgE by adenylation and the proposed subsequent transfer to HcgF to form a thioester were analyzed by chemical reduction of the intermediates. For these analyses, jHcgF and mHcgE were purified under anoxic conditions and the reactions were conducted in an anaerobic chamber. The reaction mixture contained 0−150 µM HcgF, 0−150 µM mHcgE, 0−100 µM compound **3** and 0.1−10 mM ATP/$MgCl_2$ were incubated at 65 °C for up to 60 min in a total volume of 20 µl. After the enzyme reaction, 0.6 µl of 1 M $NaBH_4$ in 1 M NaOH was added (28 mM final concentrations) to reductively cleave the potential thioester bond between compound **3** and HcgF. Following, the same volume of 1 M HCl was added to neutralize the solution and activate the reduction agent $NaBH_3$. After incubation at 18 °C for at least 30 min, the reaction tubes were removed from the anaerobic

chamber. The samples were acidified by adding 500 µl 0.1% triflouroacetic acid (TFA) and 500 µl 5% formic acid and applied to a C_{18} reverse-phase cartridge. The samples were eluted from the cartridge with 0.35 ml 50 % acetonitrile in 0.1% TFA and then concentrated and analyzed by MALDI-TOF-MS using α cyano-4-hydroxycinnamic acid as matrix. For quantification of the reaction product, the peak heights of compound **3** (HGP) and its alcohol derivative (HAD) were used. The reaction yield was calculated by Equation 2 assuming equal ionization in the MALDI-MS.

$$reaction\ yield = \frac{HAD}{(HGP + HAD) * initial\ concentration\ of\ \mathbf{3}}$$

Equation 2. Amount of reduced compound **3** (reaction yield) produced by the reduction with $NaBH_4$. HAD is the height of the alcohol peak and HGP the height of the compound **3** peak in the MALDI-MS spectra.

2.9 Purification of coenzymes from methanogens

2.9.1 Purification of tetrahydromethanopterin from *M. marburgensis*

Tetrahydromethanopterin (H_4MPT) was purified from 100 g of *M. marburgensis* grown under nickel-sufficient conditions. All following steps were conducted in anaerobic chamber and light protected conditions. The cells were resuspended in 200 ml 50 mM MOPS/KOH pH 6.8 and heated to 60 °C in a water bath. 92.5 ml of preheated 5% cetrimonium bromide (CTAB) solution, 25% (m/m) of the cell dry weight as volume, was added and the solution was incubated at 60 °C for 6 min. After cooling for 15 min in ice water, the pH was adjusted to pH 3 by adding formic acid and the solution was centrifuged with a T647.5 rotor at 20,000 rpm at 4 °C for 30 min. The supernatant was loaded onto a Serdolit PAD II resin with a volume of 50 ml in 380 mM Formate / NaOH buffer pH 3 (XAD buffer) with 4 ml/min flow rate. The column was washed with 6 CV XAD buffer and the coenzyme was eluted with 10 CV 15 % MeOH in XAD buffer à 100 ml fractions. The fractions containing H_4MPT were pooled and dried down by evaporation at room temperature. The coenzymes were dissolved in XAD buffer and were adjusted to pH 3 with formic acid. Further, the solution was loaded onto a column with Serdolit PAD I resin with a volume of 50 ml in 0.1 % formic acid in water and washed with 6 CV 0.1 % formic acid before H_4MPT was eluted with 2 CV 30% MeOH in 0.1 % formic acid. H_4MPT containing fractions were pooled and dried down by evaporation at room temperature. The final H_4MPT solution was dissolved in water at a concentration of 2.6 mM and stored at −20 °C in aliquots of 250 µl.

2.9.2 Purification of guanylylpyridinol compound 3 from *M. marburgensis*

M. marburgensis cells obtained under nickel-limiting conditions were suspended in 50 mM potassium phosphate buffer pH 7. The cell extract was prepared by sonication using SONOPULS GM200 with KE76 tip in ice water with a 50% cycle

and 160 W for six times for 8 min with 7 min pauses. To obtain a partially purified [Fe]-hydrogenase, ammonium sulfate fractionating was conducted as described in chapter 3.6 for purifying of [Fe]-hydrogenase from *M. marburgensis*. The protein fraction with 90% ammonium sulfate precipitation was dialyzed against water using a dialysis tube with a molecular weight cut off of 30 kDa at 8 °C over night. From this solution the FeGP cofactor was extracted by adding 60% MeOH, 1% NH_3 and 1 mM 2-mercaptoethanol (final concentrations) with a final protein concentration of 4 mg/ml. After incubating at 40 °C for 30 min, the solution was filtered using a membrane with a molecular weight cut off of 10 kDa at 8 °C under air and room light to remove denatured proteins. Simultaneously, the FeGP cofactor was light inactivated. The filtrate was evaporated to dryness at room temperature and dissolved in 0.01% NH_3. This solution was applied to three connected Q-trap HP columns (GE healthcare, total 15 ml column volume) equilibrated with 0.01% NH_3. After washing with 5 CV of 0.01% NH_3, the compounds were eluted by a gradient of 0 to 1 M NaCl in 1% NH_3 over 20 CV. Elution of compound **3** was detected at ca. 0.35 M NaCl via absorbance at 280 nm, which was confirmed by HPLC with Synergi 4m Polar RP 80A column (250 mm × 4.6 mm, Phenomenex) using a acidified water (HCl, pH3)/methanol gradient. Compound **3** in the fraction was further purified by HPLC under the same conditions used for the analysis of compounds **1** and **2** in chapter 3.8.2. The purified compound was dried down by evaporation, dissolved in water and stored at −20 °C.

2.9.3 Preparation of $^{18}O_1$ labelled compound 3

The FeGP cofactor was extracted from 15 mg of *M. marburgensis* [Fe]-hydrogenase at a final concentration of 2 mg/ml by adding 60 % MeOH, 1 % NH_3 and 1 mM 2-mercaptoethanol (all final concentrations). The solution was incubated at 40 °C for 15 min and denatured protein was removed by ultra-filtration with a 10 kDa cut-off membrane. The cofactor was dried by evaporation at room temperature and dissolved in [^{18}O]-H_2O containing 1 mM 2-Mercaptoethanol. The cofactor was then inactivated by cold light (3000k 4D with Schott KL 2500 LCD) over night. The final preparation of compound **3** was analyzed by HPLC and mass spectrometry as described in chapter 3.12.2.

2.10 Enzymatic synthesis of pyridinol compounds

2.10.1 Synthesis and purification of 3,5,6-trimethyl-4-guanylyl-2-pyridinol (3')

Compound **3'** was prepared by enzymatic conversion of 3,5,6-trimethyl-2-pyridinol (**2'**) by HcgB from *M. maripaludis* with GTP. A mixture of 100 µM **2'**, 5 mM GTP, 5 mM $MgCl_2$ and 10 µM HcgB was incubated in 50 mM MPOS/KOH pH 7 in a volume of 3 ml at 40 °C for 48 h. The protein was precipitated by adding 40 µM acetic acid (final concentration) and removed by centrifugation at 14,500 rpm in an Eppendorf MiniSpin plus centrifuge at room temperature for 20 min followed by

filtration with 0.22 µM pore size. The sample was dried by evaporation at 4 °C and **3'** was separated from **2'** and other contaminants by HPLC with a Synergi 4m Polar RP 80A column (250 mm × 4.6 mm, Phenomenex) employing an acidified water (HCl, pH3)/methanol gradient. The fraction containing **3'** was collected, dried by evaporation and dissolved in water.

2.11 *In vitro* biosynthesis reactions

2.11.1 *In vitro* biosynthesis using *Methanothermobacter marburgensis* cell extract

The *M. marburgensis* cell extract was prepared as described in chapter 3.6. In the anaerobic chamber 1 mM DTT and 1 mM $Fe(SO_4)_2(NH_4)_2$ were added to the cell extract in a 1.5 ml plastic reaction tube. After 5 min, 2 mM sodium dithionite, 60 µM (2.3 mg/ml) [Fe]-hydrogenase from *M. jannaschii*, 20 mM ATP/$MgCl_2$ and 30 µM **3** were added. The solution was transferred to a glass vial sealed with rubber stopper containing 50% H_2/ 50% CO atmosphere and 0.5 mM Na_2S was added. The final total cell extract protein concentration in the assay was 16 mg/ml. The solution was incubated at room temperature and a 10 µL aliquot was withdrawn and used for the [Fe]-hydrogenase activity assay after the indicated time.

2.11.2 *In vitro* biosynthesis of FeGP cofactor using cell extract from *M. maripaludis* mutants

The cell extract was thawed or the frozen cells were lysed as described in chapter 3.5.1. Then cell debris and membrane segments were removed by ultracentrifugation with a Sorvall TFT-80.4 rotor at 35,500 rpm at 4 °C for 1 h. The *in vitro* biosynthesis assay was prepared in the anaerobic chamber with a master mix containing Fe^{2+}, sodium dithionite, DTT, ATP, $MgCl_2$ and SAM added to 200 µl of cell extract and supplemented with 10 µM apo [Fe]-hydrogenase from *M. jannaschii* (jHmd) and 10 µM of the appropriate precursor compound. The final concentration of all compounds and the standard conditions are summarized in Table 6. ATP and SAM were added to the master mix directly before usage to avoid spontaneous decomposition during storage. The *in vitro* assay reaction mixture was then transferred into 6 ml amber serum bottles with the appropriate gas phase indicated in the results section and incubated at room temperature or 40 °C for various durations.

For the filtration experiments, the cell extract was anaerobically filtrated by a 3- or 10-kDa centrifugal filter (Millipore Amincon Ultra-0.5) at 8 °C until around 10% of the initial volume. The first filtrate was stored separately on ice. The concentrate was then diluted to the initial volume using 50 mM Tris/HCl pH 7.5 containing 5 mM $MgCl_2$ and the concentration/dilution cycle was repeated for three times. Finally, the concentrate was diluted to the initial volume by either the first filtrate or by

50 mM Tris pH 7.5 containing 5 mM $MgCl_2$ and used for the *in vitro* biosynthesis assay.

In the case of *in vitro* biosynthesis employing heterologously purified Hcg proteins, the above mentioned cell extract was replaced with the filtrate, obtained by ultrafiltration with a 3 kDa cut-off. The heterologously purified enzymes were added between the master mix and [Fe]-hydrogenase apoenzyme.

Table 6. Components of the *in vitro* biosynthesis assay

Compound	concentration
ATP	5 mM
GTP*	5 mM
SAM	2 mM
$Fe(NH_4)_2(SO_4)_2$	1 mM
$MgCl_2$	5 mM or 10 mM*
Dithiothreitol	1 mM
Sodium dithionite	2 mM
Precursor compound: **1**, **2** , **2'**, **3** or **3'**	10 µM
[Fe]-hydrogenase apoenzyme	10 µM
mpHcgB and/or mpHcgC*	1 µM
jHcgF, mpHcgD, mHcgE#	10 µM

* When compounds **1** or **2** were used as precursors.

When heterologously purified proteins and filtrate was used instead of cell extract.

2.12 Mass spectrometric analyses

2.12.1 Preparation of the FeGP cofactor from in vitro biosynthesis solution for mass spectrometric analysis

To extract the FeGP cofactor from the *in vitro* biosynthesis assay solution, the Strep-tagged apoenzyme from *M. maripaludis* was used in a larger *in vitro* biosynthesis assay in a volume of 5 ml instead of the apoenzyme from *M. jannaschii*. The assay solution was stored at 4 °C in N_2 atmosphere after completion of the assay overnight. To adjust the pH the solution was diluted to 50 % of the initial volume in 150 mM Tris/HCl pH 8 containing 100 mM NaCl. Removing excess biotin from the media was conducted by adding 1 U of Avidin per ml of *in vitro* solution. The holo [Fe]-hydrogenase was purified by Strep trap affinity chromatography using two connected 5 ml StrepTrap column (GE healthcare) in 150 mM Tris/HCl containing 100 mM NaCl and 5% glycerol at 1 ml/min. The column was washed with 5 CV loading buffer and the protein was

eluted with 150 mM Tris/HCl containing 100 mM NaCl and 2.5 mM desthiobiotin. The FeGP cofactor was extracted from the [Fe]-hydrogenase by the MeOH/2-ME method (see chapter 3.7) and analyzed by high resolution mass spectrometry described in following chapters.

2.12.2 Electron spray ionization mass spectrometric analysis of the FeGP cofactor

The following electrospray ionization mass spectrometric (ESI-MS) analysis was conducted in the Metabolomics core facility of the Max-Planck-Institute of terrestrial microbiology, Marburg. The subsequent isotopologue analysis was conducted by myself. The description of the mass spectrometric analysis was first published in Schaupp *et al.,* 2022. [95]

Mass spectrometric determination of the FeGP cofactor was performed using a HRES-LC-MS. The chromatographic separation was performed on a Thermo Scientific Vanquish HPLC System using a polymer based ZICpHilic (Sequant, 150 × 2.1 mm, 5 µm, Merck) equipped with a 20 × 2.1 mm guard column of similar specificity at a constant eluent flow rate of 0.25 ml/min and a column temperature of 40 °C with eluent A being 10 mM ammoniumhydroxyde in water adjusted to a pH of 9.8 and eluent B being acetonitrile (Honeywell) The injection volume was 2 µl. The elution profile consisted of the following steps and linear gradients: 0 – 3 min constant at 95 % B; 3–10 min from 95 to 20 % B; 10 – 20 min constant at 20 % B; 20 – 20.1 min from 20 to 95 % B; 20.1 – 30 min constant at 95 % B. Ionisation was performed using a high temperature electro spray ion source at a static spray voltage of 3300 V, Sheath gas at 35 (Arb), Auxilary Gas at 7 (Arb), and Ion transfer tube and Vaporizer at 300 and 275 °C. Full Scan measurements were conducted applying an orbitrap mass resolution of 240 000 without using quadrupole isolation in a mass range of 100−642. Data was saved in full profile mode. Targeted fragmentations measurements were performed at similar chromatography and ionisation settings, but using a quadrupole isolation of the target ion in a window of 0.4 *m/z*. Collision induced dissociation was performed in the ion routing multipole with a relative collision energy of 5 %. Fragments were detected using the orbitrap at a predefined mass resolution of 60,000 in the range between 100 and 640. For the analysis of the amount of incorporated ^{13}C in the [^{13}C]-CO labelling experiments the peak area of the extracted ion chromatogram of the calculated ion mass ± 5 ppm was integrated for each isotopologue and subsequently the natural isotope abundance of ^{13}C was subtracted using the IsoCor software. [96]

2.12.3 Proteome analysis

The proteome analysis and the MALDI-TOF-MS were conducted in the proteomics core facility of the Max-Planck-Institute of terrestrial microbiology, Marburg. The

description of the mass spectrometry methods proteome analysis was first published in Schaupp *et al.,* 2022. [95]

Cell pellets were lysed and reduced by tris(2-carboxylethyl)phosphine (TCEP) in the presence of deoxycholate (DOC) at 90 °C for 10 min. After that it was incubated at 25 °C for 30 min in ammonium bicarbonate pH 8.2 and iodoacetic acid (IAA) and then digested overnight at 30 °C with trypsin, MS approved (Serva). Before LC-MS analysis, samples were desalted using C18 microspin columns (Nest Group) according to the manufacturer's instructions. Dried and resuspended peptides were then analyzed using liquid-chromatography-mass spectrometry carried out on an Orbitrap Exploris 480 instrument connected to an Ultimate 3000 RSLC nano and a nanospray ion source (all Thermo Scientific). Peptide separation was performed on a reverse phase HPLC column (75 µm x 40 cm) packed in-house with C18 resin (2.4 µm; Dr. Maisch) with a 135 min gradient (formic acid / acetonitrile). MS data were searched against an in-house *Methanococcus maripaludis* S2 protein database using SEQUEST embedded into Proteome Discoverer 1.4 software (Thermo Scientific).

2.12.4 Matrix assisted laser desorption - time of flight mass spectrometry

The samples were acidified by adding 500 µl 0.1% triflouroacetic acid (TFA) and 500 µl 5% formic acid and applied to a C18 reverse-phase cartridge. The samples were eluted with 0.1% TFA and then concentrated and analyzed via MALDI-TOF-MS using α-cyano-4-hydroxycinnamic acid as matrix employing a 4800 Proteomics Analyzer (Applied Biosystems/MDS Sciex).

3 Results and Discussion

3.1 *In vitro* biosynthesis of the FeGP cofactor

3.1.1 *In vitro* biosynthesis assays based on *M. marburgensis*

Initially, two *in vitro* biosynthesis assays were tested, which are based on the cell extract of *M. marburgensis*. The cells are known to produce high amounts of [Fe]-hydrogenase under Ni-limiting conditions[21] and therefore have a large capacity for the biosynthesis of the FeGP cofactor. The cells show a [Fe]-hydrogenase activity of 20 U per mg of cellular protein and grew with a doubling time of ~10 h[21b, 30], which indicates that the biosynthesis machinery of the cells synthesizes the FeGP cofactor at least 1 U per mg of cellular protein per hour during growth. The first experiments were based on previous work in our laboratory, in which the FeGP cofactor was partially inactivated (50% decrease of [Fe]-hydrogenase activity) by illumination with strong white light. By incubation under 50 % H_2 / 50 % CO or 100 % CO, the [Fe]-hydrogenase activity increased in some experiments (Figure 6a). However, the reproducibility of this method was very low and finally this biosynthesis activity could not be reliably reproduced. This might be due to three major problems: the variability of co-substrate concentrations in the cell extract, different concentration of the possible precursors, and finally the high initial background activity of 10 U/mg, since the expected biosynthesis capacity (see above) is close to the error margin of this method (5 to 10 % of the activity). Especially the last finding made it difficult to precisely assess the [Fe]-hydrogenase activity produced during the assay. The variability in co-substrates and precursors could be addressed by purifying the precursors and adding a mixture of possible small components based on previous studies on the FeGP cofactor biosynthesis and other hydrogenase *in vitro* biosynthesis assays (Figure 6b).[45, 76a] But also in this case, the high background activity made it impossible to detect the effect of the addition of the precursor. Based on this experience, I realized that for a successful *in vitro* biosynthesis assay, a system with no or very little background of [Fe]-hydrogenase activity was needed.

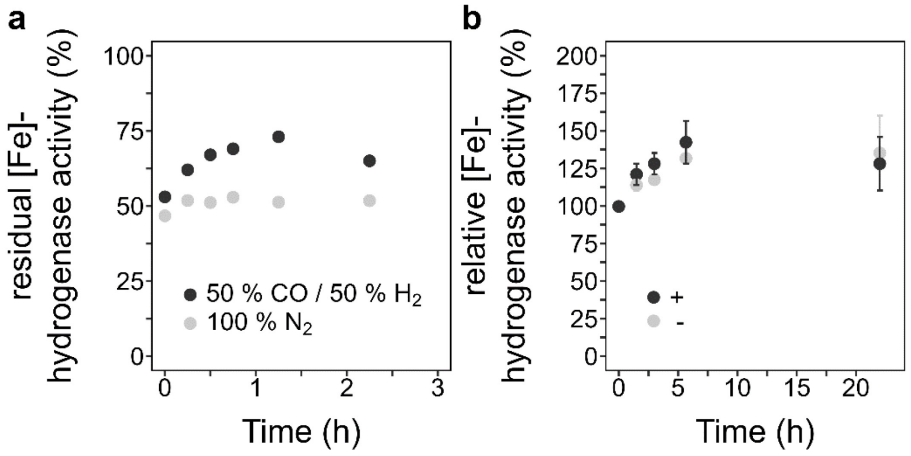

Figure 6. Biosynthesis assays of the FeGP cofactor based on cell extract from *M. marburgensis*. (a) Change of the [Fe]-hydrogenase activity in partially light inactivated cell extract of *M. marburgensis* incubated under 100 % N_2 (grey) or 50 % CO / 50 % H_2 (black) atmosphere. In this assay no external precursor was added. The residual activity refers to the activity of the *M. marburgensis* cell extract [Fe]-hydrogenase activity before the light inactivation, see material & methods for details on the light inactivation procedure. (b) Time-dependent change of the [Fe]-hydrogenase activity in cell extract of *M. marburgensis* with ATP, SAM, Fe^{2+}, DTT and DTH added. (+) compound **3** was added to the assay as a precursor (black), (−) no precursor was added (grey). The relative [Fe]-hydrogenase activity refers to the *M. marburgensis* cell extract [Fe]-hydrogenase activity before the start of the assay.

3.1.2 *In* vitro biosynthesis using the *M. maripaludis* cell extract

To overcome the problems described above, I chose the *M. maripaludis Δhcg* mutants since they were proposed as [Fe]-hydrogenase inactive and therefore even small amounts of biosynthesis activity can be easily detected by measuring the [Fe]-hydrogenase activity generated. The group of John A. Leigh (Washington University) has previously shown that deletion of each *hcgA–G* gene in *M. maripaludis* caused an extended lag-time of growth under H_2, which is similarly observed in the *Δhmd* strain.[40] The growth properties were restored by complementation of the respective gene, which indicated that expression of the *hcg* genes is not disturbed by the deletion of other genes. In this case, addition of the missing enzymes or precursors to the cell extract of the mutants should complement the FeGP cofactor biosynthesis activity to the cell extract. I chose the Δ*hcgB* strain for the initial *in vitro* biosynthesis experiments. In this work, this strain is referred to as the Δ*hcgB*Δ*hcgC* strain because the *hcgB* gene was deleted together with the ribosome binding site of *hcgC*, which led to a mutant that produces neither HcgB nor HcgC (Table S 1). The *hcg* mutants of *M. maripaludis* were obtained from Kyle Costa (University of Minnesota). The Δ*hcgB*Δ*hcgC* strain did not show [Fe]-hydrogenase activity. In the *in vitro* biosynthesis using the Δ*hcgB*Δ*hcgC* strain, the biosynthesis activity should be restored by addition of HcgB and HcgC, or addition of the proposed guanylylpyridinol compound **3**, based on the proposed model of the FeGP cofactor biosynthesis shown in Figure 3. The proposed physiological precursors **1** and **2** were chemically synthesized by the group of Xile Hu (École polytechnique fédérale de Lausanne).

3.1.3 *In vitro* biosynthesis from 3 with the *M. maripaludis* ∆*hcgB*∆*hcgC* strain

I tested the production of holo [Fe]-hydrogenase in a mixture of *M. maripaludis ΔhcgBΔhcgC* cell extract, the possible precursor **3** with the reaction components including ATP, $MgCl_2$, SAM, DTT, sodium dithionite and the [Fe]-hydrogenase apoenzyme (for more information, see Materials & Methods). Under an atmosphere of 100 % N_2, even after extended incubation (> 1 h) at room temperature, no [Fe]-hydrogenase activity was detected; however, when I performed the assay under an H_2 atmosphere or with formate in the solution, some [Fe]-hydrogenase activity was observed. This activity increased four-fold when 50% CO / 50% H_2 was used instead (Figure 7). The necessity of H_2 and/or formate indicated the need for reducing equivalents, since H_2 and formate are used by hydrogenases and formate dehydrogenases as electron donors, respectively. CO could also supply reducing equivalents via carbon monoxide dehydrogenases or it might be a precursor of the CO ligands as indicated by M. Schick *et al.*[42]

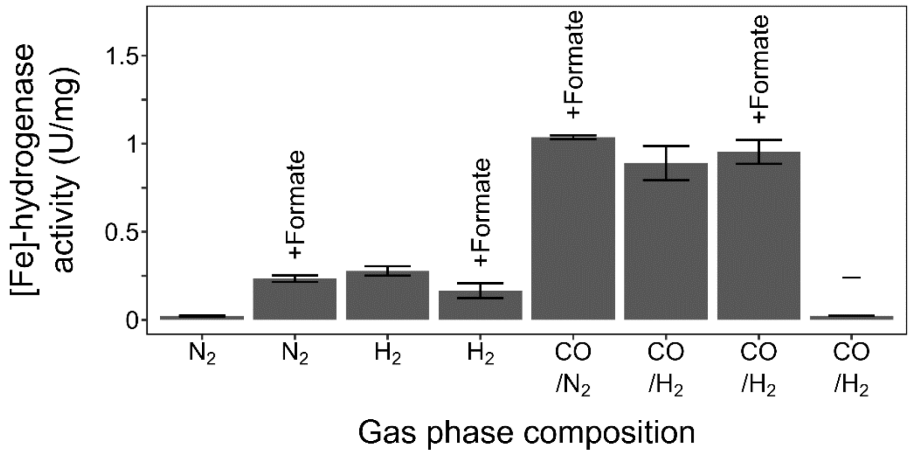

Figure 7. Effect of the gas phase and addition of formate in the assay solution on the *in vitro* biosynthesis activity. The composition of the gas phase is given on the abscissa. Addition of formate is shown above each bar. (–) indicates the absence of **3** in the assay.

The 50 % H_2 / 50 % CO gas phase condition is used as the standard in the *in vitro* biosynthesis assay in the following text. In the absence of guanylylpyridinol precursor **3** in the standard assay condition, the activity was not detected. To confirm that the observed activity is indeed [Fe]-hydrogenase activity, I determined the [Fe]-hydrogenase activity of the *in vitro* biosynthesis solution by measuring the UV-Vis spectrum change caused by the enzyme reaction (Figure 8a). The spectrum change was consistent with the conversion of methylene-H_4MPT to methenyl-H_4MPT$^+$, which is observed in the standard assay in the presence of [Fe]-hydrogenase purified from *M. marburgensis* (Figure 8a). The difference of the spectrum in Figure 8a in the region of 250 – 300 nm after addition of the *in vitro*

biosynthesis solution can be attributed to the absorbance of the proteins in the cell extract (Figure 8a, green spectrum). This change in the spectrum was time dependent (Figure 8b, c) and the speed was dependent on the amount of *in vitro* biosynthesis solution added to the reaction. [Fe]-hydrogenase is specifically inhibited by tosylmethylisocyanide (TosMIC) [33] and the same is true for the [Fe]-hydrogenase activity produced by the *in vitro* biosynthesis assay (Figure 8d). This indicates that the active [Fe]-hydrogenase holoenzyme was produced by activation of apo [Fe]-hydrogenase with the assay by the FeGP cofactor synthesized in the assay. The [Fe]-hydrogenase activity in the *in vitro* biosynthesis solution increased during incubation at room temperature (Figure 8e). The [Fe]-hydrogenase activity formed in the assay is called *in vitro* biosynthesis activity in the following text. The FeGP cofactor can be extracted by thiols, such as 2-mercaptoethanol, and acids, such as acetic acid, from the holo-enzyme upon denaturation of the protein with methanol, urea, guanidinium chloride or by reduction of the pH by addition of acid.[24, 29] This extracted cofactor can then be used to reconstitute the [Fe]-hydrogenase holoenzyme. This effect was observed with the cofactor from the *in vitro* biosynthesis, which could be extracted by 2-mercaptoethanol/MeOH, acetic acid, H_2SO_4 or formic acid and subsequently used for reconstitution of apoenzyme (For more information, see Materials and Methods) (Figure 8f). These experiments showed for the first time that the FeGP cofactor was produced *in vitro* using the cell extract of *M. maripaludis ΔhcgBΔhcgC* and that the reaction is dependent on the presence of **3**.

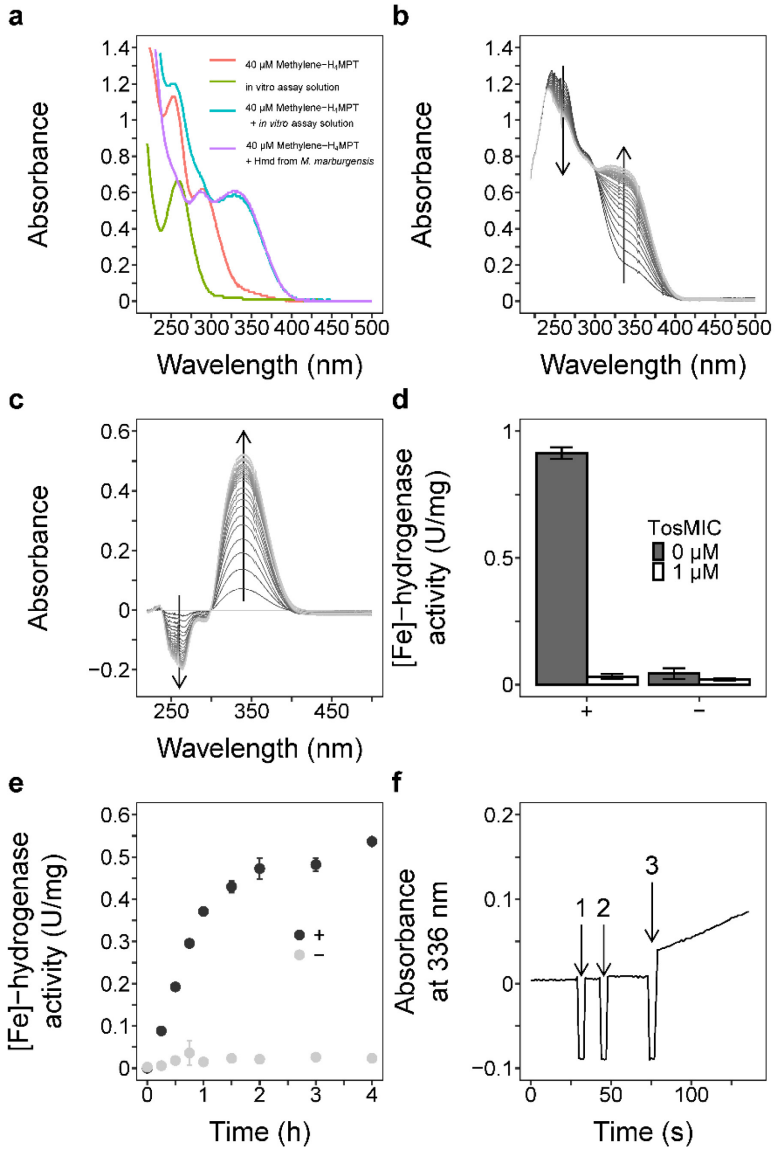

Figure 8. Conversion of the UV-Vis spectrum upon dehydrogenation of methylene-H_4MPT catalyzed by the [Fe]-hydrogenase produced in the *in vitro* biosynthesis assay. (a) UV-Vis spectra of methylene-H_4MPT (red), after the reaction with [Fe]-hydrogenase formed in the *in vitro* biosynthesis assay (cyan), after the reaction with [Fe]-hydrogenase purified from *M. marburgensis* as a positive control (violet), and the *in vitro* biosynthesis solution without methylene-H_4MPT (green). The peak observed at 336 nm indicates the presence of methenyl-H_4MPT formed in the solutions. (b) Time-resolved (10s) UV-Vis spectrum changes by the reaction of [Fe]-hydrogenase produced in *the in vitro* biosynthesis assay. (c) Differential spectra (calculated by subtraction of the initial spectrum) of the [Fe]-hydrogenase activity detected in (b). Arrows indicate the direction of the change. The reaction was started by addition of 25 µl 50-fold diluted *in vitro* biosynthesis solution to a 1 cm light pass cuvette with 40 µM methylene-H_4MPT. (d) [Fe]-hydrogenase activity of the *in vitro* biosynthesis solution of the positive control with the precursor **3** (+) and without addition of **3** (–). The activity in the presence of 0 µM and 1 µM tosylmethylisocyanide (TosMIC) is shown. (e) Change of the [Fe]-hydrogenase activity in the *in vitro* biosynthesis solution during incubation at room temperature in the presence (+) and absence (–) of **3**. (f) Reconstitution of [Fe]-hydrogenase in the enzyme assay cuvette with the FeGP cofactor extracted by acetic acid from the *in vitro* biosynthesis solution. Arrows indicate the addition of the corresponding substance (1: 10 µg [Fe]-hydrogenase, 2: FeGP cofactor, 3: 40 µM methylene-H_4MPT) to the assay solution.

3.1.4 Analysis of the FeGP cofactor from the *in vitro* biosynthesis by mass spectrometry

The structural characterization of the FeGP cofactor has so far exclusively relied on FeGP cofactor prepared from *M. marburgensis* and the structure of FeGP cofactor from *M. maripaludis* has not been reported. Above described enzymological assay of FeGP cofactor, which was produced in *in vitro* biosynthesis assay using cell extract from the *M. maripaludis* strain, does not provide evidence of the structure of FeGP cofactor. To obtain structural information of FeGP cofactor produced in the assay, high resolution mass spectrometric analyses were performed. For this purpose, [Fe]-hydrogenase apoenzyme from *M. maripaludis* with N-terminal Strep-tag was used (Figure S 1). FeGP cofactor was extracted from the purified Strep-tag [Fe]-hydrogenase holoenzyme by either acetic acid or MeOH/2-ME method.

The cofactor was analyzed by N. Paczia (MPI Marburg, Metabolomics und Kleinmolekül-Massenspektrometrie), as described in Material & Methods. FeGP cofactor was detected as $[M\text{-}2H]^-$ ($C_{21}H_{20}FeN_6O_{12}P$) with a monoisotopic mass of 635.0207 *m/z* (4.3 ppm deviation) in negative mode. (Figure 9a and b). The natural isotope pattern of measured cofactor contains the characteristic ion species containing the ^{54}Fe isotope rather than ^{56}Fe (5.8% natural abundance)[97] proving the existence of a single iron ion. Furthermore, detected fragment ions at MS/MS level match already reported fragments of FeGP cofactor of *M. marburgensis* verifying its identity (Figure 9c).[29] Apart from not fragmented precursor, two fragments are detected at 607 and 551 m/z corresponding to the loss of one and three CO units, respectively. The later fragment shows the highest abundance. Since they have the same exact mass, the MS/MS data does not allow the distinction between CO and acyl groups regarding the formation of the ion species at 607 m/z. The existence of three CO equivalent units (i.e. two CO and acyl ligands) in FeGP cofactor was already shown before.[29] To exclude the possibility that this mass originates from another molecule in the *in vitro* biosynthesis assay, the same purification was performed for an *in vitro* biosynthesis assay solution lacking precursor **3**. The purified Strep-tag [Fe]-hydrogenase did not show any [Fe]-hydrogenase activity and no FeGP cofactor could be detected by ESI-MS (Figure 9d, red). These findings show that FeGP cofactor is produced in the *in vitro* biosynthesis only in presence of precursor **3** and that the structure of biosynthesized cofactor is identical to the native FeGP cofactor from *M. marburgensis*. These results show that the [Fe]-hydrogenase activity described above arises from FeGP cofactor produced by *in vitro* biosynthesis. Therefore, this method can be used to test different precursors, as well as enzymes in the *in vitro* biosynthesis reaction. Furthermore, this shows that FeGP cofactor of *M. marburgensis* and *M. maripaludis* likely have the same structure.

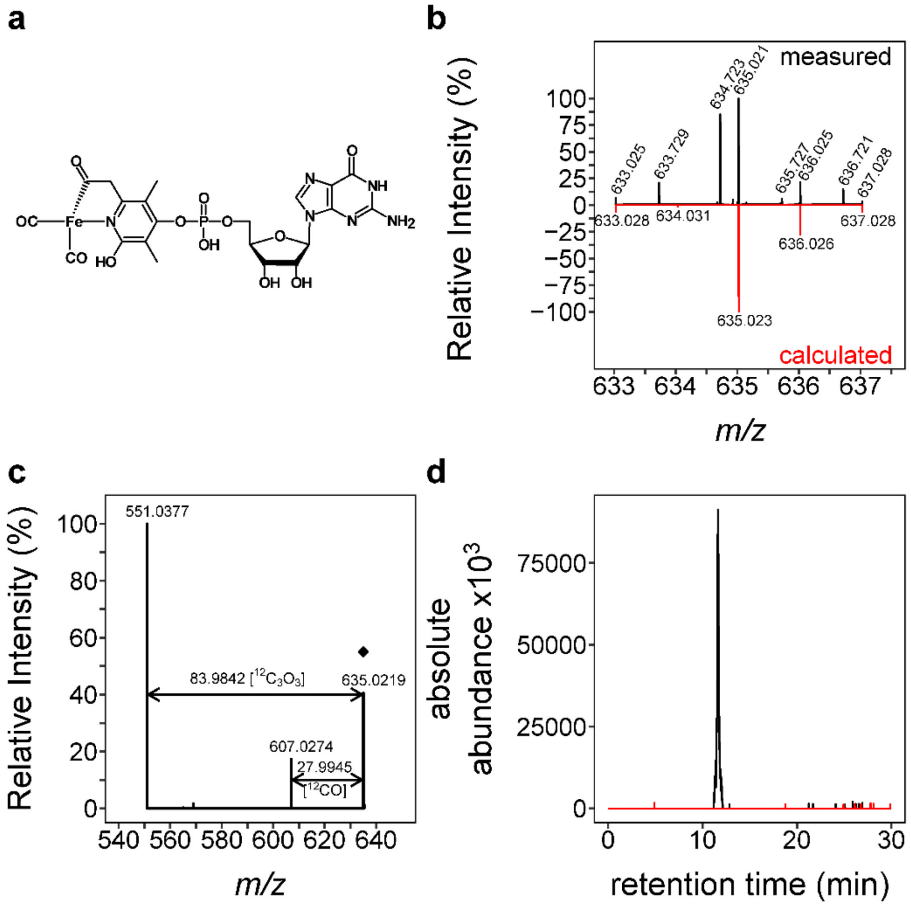

Figure 9. Mass spectrometric analysis of the FeGP cofactor purified from [Fe]-hydrogenase isolated from *in vitro* biosynthesis. (a) Predicted structure of FeGP cofactor before ionization based on previously published data of cofactor from *M. marburgensis*. The ion resulting from deprotonation is detected by mass spectrometry as a negative species $[M\text{-}2H]^-$. (b) Calculated mass spectrum (red) of the FeGP ion based on the predicted structure and the measured spectrum (black) of FeGP cofactor extracted from *in vitro* biosynthesis. The deviation between calculated and measured mass is 4.3 ppm. (c) Tandem mass spectrometric analysis of 635 *m/z* precursor ion (M_0) of FeGP cofactor. A diamond indicates the precursor ion and fragment peaks (551 and 607 *m/z*) from the precursor ion are shown. (d) Extracted ion chromatogram of the cofactor (635.0200 - 635.0264 *m/z*) extracted from [Fe]-hydrogenase purified from *in vitro* biosynthesis containing compound **3** (black) or without precursor (red).

3.1.5 Origin of the CO and acyl-ligands in the *in vitro* biosynthesis

One question that arose in the initial stage of *in vitro* biosynthesis investigation was the role of CO in the gas phase of the assay. Addition of CO into the gas phase (50 % H_2 / 50 % CO) increased [Fe]-hydrogenase activity four-fold compared to the activity with 100 % H_2 in the assay (Figure 7). The *in vitro* biosynthesis activity increased with partial pressure of CO in the gas phase and CO partial pressure was approximately 0.1 bar for the half-maximum activity (Figure 10a). An obvious interpretation of this result is that the CO gas molecule is bound to an Fe ion in the biosynthesis of the CO ligands of FeGP cofactor. However, since CO is a toxic gas

and inhibits some metalloenzymes in methanogenic archaea, the presence of high concentrations of CO gas in the cell, as used in the *in vitro* biosynthesis, is unlikely.

To answer the question about the function of CO gas in *in vitro* biosynthesis assay, I conducted the standard *in vitro* biosynthesis assay in presence of [^{13}C]-CO. For the analysis of this data, it was necessary to distinguish between the fraction of isotopic labelling resulting from natural isotopic abundance and from introduced labelled molecules. For this purpose, raw data was corrected by IsoCor, a software to calculate and remove the contribution of natural isotopic abundance to the isotopologue distribution.[96] Mass spectrometric analysis detected several ^{13}C isotopologues of FeGP cofactor, when FeGP cofactor was synthesized under 50 % [^{13}C]-CO / 50 % H_2 by *in vitro* biosynthesis. The IsoCor calculation indicates that 18 % of FeGP carried two ^{13}C and 31 % carried one ^{13}C, while 51 % was not labelled by ^{13}C (Figure 10b). ^{13}C from [^{13}C]-CO was predicted to be incorporated into CO and/or acyl ligands based on previous studies[42] and this hypothesis is discussed in-depth below and in chapter 4.2. While the finding that ^{13}C from [^{13}C]-CO was incorporated into FeGP cofactor indicates that gaseous CO can be the precursor of CO and/or acyl ligands, the majority (67 %, assuming all ^{13}C was incorporated into the CO ligands and not the acyl ligand) of the CO ligands originated from another CO source. A possible explanation of this finding is the presence of a molecule in the cell extract, which can be converted into the CO ligands enzymatically. The presence of such a physiological precursor was predicted since the presence of CO-donating compounds was described for [FeFe]- and [NiFe]-hydrogenase maturation.[70-71, 74d, 98]

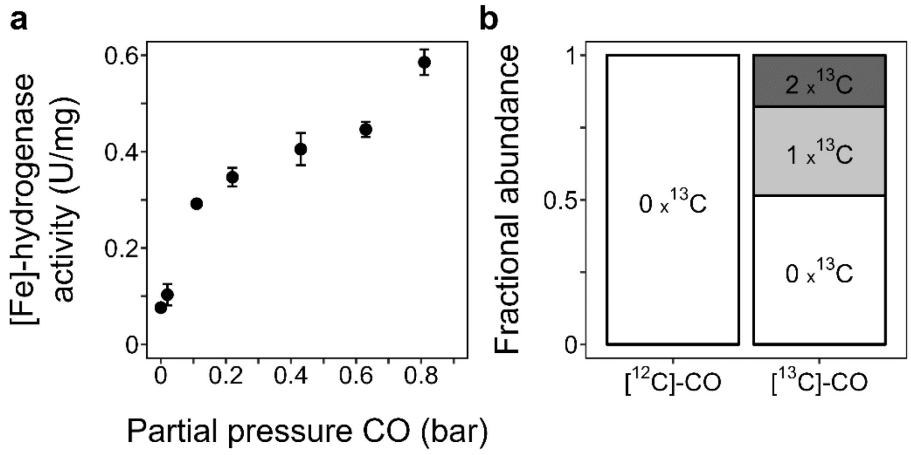

Figure 10. Effect of CO in the gas phase on the *in vitro* biosynthesis assay. (a) Effect of CO partial pressure on [Fe]-hydrogenase activity produced in *in vitro* biosynthesis (total pressure in the assay was 1 bar). (b) Incorporation of [^{13}C]-CO from the gas phase into FeGP cofactor as indicated by ^{13}C enrichment in the molecule. The fractional abundance corresponds to experimentally observed isotope distribution, which was corrected for natural isotope abundance contribution using IsoCor.[96]

MS/MS experiments were performed to determine whether [^{13}C]-CO is incorporated into CO and/or acyl ligands of FeGP cofactor, since the analysis ofcomplete FeGP cofactor could not answer this question. In contrast to the

analysis of incorporation of [^{13}C]-CO into FeGP cofactor, IsoCor cannot correct the isotopologue distribution of the fragmentation ions for the natural abundance of ^{13}C. This correction is not possible since signal intensity is not high enough due to low abundance of the precursor ions of carrying one or two ^{13}C atoms. Therefore, a more complex analysis is needed. The different isotopologues containing different numbers of ^{13}C were denoted in this text as $^{13}C_0$, $^{13}C_1$ or $^{13}C_2$, which contain zero, one or two ^{13}C carbon atoms. In this case, non-carbon atoms in the isolated molecules are all the most abundant isotope (^{14}N/^{16}O/^{56}Fe/^{31}P). In these experiments, an isotopologue of FeGP cofactor, with a given number of ^{13}C atoms, is isolated at MS level and then fragmented by MS/MS. The fragmentation pattern of FeGP cofactor shows two main fragmentation ions, in which CO units (28 *m/z*) are split off. These CO units are proposed to correspond to CO ligands and/or acyl ligands, which cannot be distinguished based on their *m/z* since they have the same exact mass. The first fragment ion lacks one CO unit (FeGP–$[CO]_1$, 607 *m/z*) and the second ion lacks three CO units (FeGP–$[CO]_3$, 551 *m/z*).[29] To analyze if the CO from the [^{13}C]-CO is incorporated into the CO and/or acyl ligands I analyzed the isotope distribution of FeGP–$[CO]_3$ fragment. Here I compared fragments from FeGP cofactor produced in presence of [^{12}C]-CO or of [^{13}C]-CO. The MS/MS spectra of $^{13}C_0$ from FeGP cofactor obtained by *in vitro* biosynthesis in presence of [^{12}C]-CO and [^{13}C]-CO are identical because both do not contain any ^{13}C atoms (Figure 11a, b). FeGP–$[CO]_3$ species in these spectra show a single ion at 551 *m/z* indicating no ^{13}C is present (Figure 11a, b). When $^{13}C_1$ of the FeGP cofactor prepared in presence of [^{12}C]-CO is analyzed, the FeGP–$[CO]_3$ species shows a signal at 552 *m/z* (Figure 11c). This signal originates from molecules, in which this ^{13}C atom is not present in either acyl or CO ligands (Figure 11c). This carbon atom originates from the natural abundance of ^{13}C (1.1%).[97] Theoretically, one would expect two peaks at 551 *m/z* and 552 *m/z* with a signal relation of 15% to 85% to account for the presence of ^{13}C from natural abundance. However, signal intensity is too low to detect the 551 *m/z* signal, since the initial abundance of $^{13}C_1$ from [^{12}C]-CO FeGP is too low (Figure 11c). A similar result is obtained for the FeGP–$[CO]_3$ species from $^{13}C_2$ of the FeGP cofactor obtained under [^{12}C]-CO atmosphere, which leads to a major signal at 553 *m/z* with a theoretical distribution of 1.4 %:15 %:83 % for 551 *m/z* to 552 *m/z* to 553 *m/z* signal intensities (Figure 11e). Again, the lower signals at 552 *m/z* and 551 *m/z* cannot be detected due to low initial intensity of the $^{13}C_2$ ion (Figure 11e). In case of FeGP cofactor synthesized by *in vitro* biosynthesis in presence of [^{13}C]-CO, $^{13}C_1$ shows a major signal at 551 *m/z* and a lower signal at 552 *m/z* (Figure 11d), where presence of 551 *m/z* signal indicates that the ^{13}C originating from [^{13}C]-CO is lost upon fragmentation. Therefore, this ^{13}C from [^{13}C]-CO is incorporated into the CO and/or acyl ligands (Figure 11d). The low signal at 552 *m/z* originates from natural abundance of ^{13}C, similar to the effect observed for FeGP cofactor synthesized in presence of [^{12}C]-CO (Figure 11c). The same analysis can be done for $^{13}C_2$, which detected signals at 551 *m/z*, 552 *m/z* and 553 *m/z* for FeGP–$[CO]_3$ species (Figure

11f) indicating that two [^{13}C]-COs are incorporated into CO and/or acyl ligand of FeGP cofactor during *in vitro* biosynthesis.

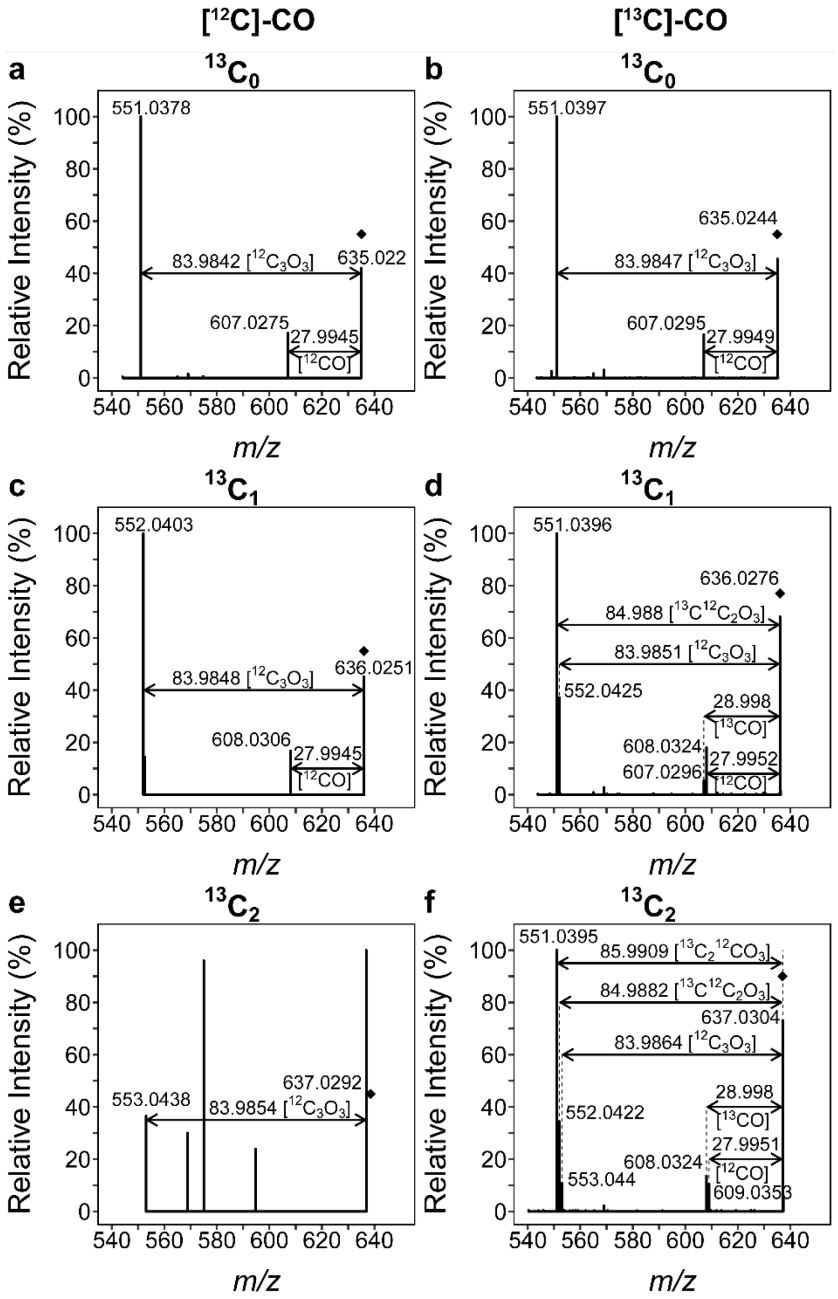

Figure 11. MS/MS analysis of [^{13}C]-CO labelled FeGP cofactor. MS/MS spectra of $^{13}C_0$ (a and b), $^{13}C_1$ (c and d) and $^{13}C_2$ (e and f) isotopologues of [^{12}C]-CO labeled FeGP cofactor (left panels) in panels a, b and c, respectively. The same analysis of FeGP cofactor produced in presence of [^{13}C]-CO (right panels). The fragments leading to 607 *m/z* and 551 *m/z* ions are shown with arrows and mass difference and corresponding molecule is given on arrows.

While the analysis of FeGP–$[CO]_3$ (551 *m/z*) shows CO from the gas phase is incorporated during *in vitro* biosynthesis into acyl and/or CO ligands, it cannot distinguish between incorporation into the CO ligands or into the acyl ligand as discussed before. To address this issue, I prepared FeGP cofactor by *in vitro* biosynthesis under [^{13}C]-CO and then decomposed the cofactor by light-exposure. The light decomposition transforms the acyl ligand into carboxyl group of compound **3** and the CO ligands dissociate.[29] If resulting compound **3** still contains ^{13}C, it indicates that the label was originally incorporated into the acyl ligand since this carboxyl group is formed by hydrolysis of the acyl ligand of FeGP cofactor upon light inactivation.[29] Analysis of light-decomposed product **3** by ESI-MS showed labelling with ^{13}C in 10 % of resulting **3** (Figure 12a, corrected for natural abundance by IsoCor). MS/MS was performed to confirm this ^{13}C atom is present in the carboxyl group. The loss of carboxyl group as CO_2 was observed previously in MS/MS of compound **3**, leading to a signal at 497 *m/z*.[29] The fragmentation of $^{13}C_1$ of compound **3** originating from FeGP cofactor produced in presence of [^{12}C]-CO (Figure 12b, red) shows only a single peak still containing a ^{13}C atom (498 *m/z*), which originates from the natural abundance of ^{13}C. In contrast, the same analysis of $^{13}C_1$ from [^{13}C]-CO showed two fragments (497 *m/z* and 498 *m/z*), indicating the presence of^{13}C in the carboxyl group of compound **3**. This finding indicates that a part of ^{13}C label from ^{13}CO can end up in the acyl ligand, which is in accordance with previous *in vivo* labelling experiments.[42]

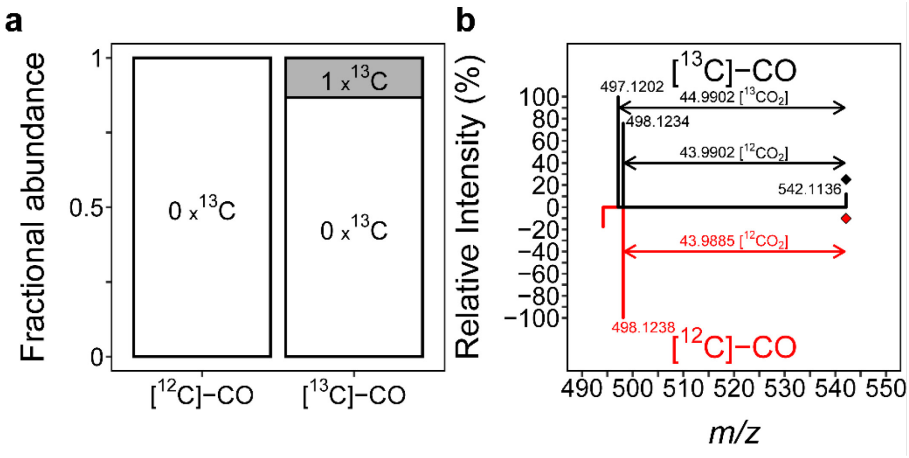

Figure 12. ^{13}C incorporation into compound **3** prepared by light-decomposition of FeGP cofactor. (a) Fractional abundance of ^{13}C isotopologues of **3** prepared from FeGP cofactor obtained from *in vitro* biosynthesis under [^{12}C]-CO or [^{13}C]-CO. The fractional abundance corresponds to experimentally observed isotope distribution corrected for the natural isotope abundance contribution using IsoCor.[96] (b) MS/MS fragmentation of compound **3**, which dissociated the carboxyl group of **3** prepared from the [^{13}C]-CO labeled FeGP cofactor as [^{12}C]-CO_2 and [^{13}C]-CO_2 molecules (black). The same experiment with **3** from non-labeled FeGP cofactor ([^{12}C]-CO) indicated dissociation of only [^{12}C]-CO_2 (red).

However, at this stage, a conclusion about the origin of the acyl ligand cannot be made. Either [^{13}C]-CO was incorporated into the acyl ligand, as the result of the biosynthesis of the FeGP cofactor or during the light inactivation, which produced **3** from the FeGP cofactor. If it was incorporated during the *in vitro* biosynthesis assay three ^{13}C should be incorporated (two CO and one acyl ligand) in one molecule. However I only detect an incorporation of two ^{13}C (see Figure 10b). In chapter 4.2, I performed further experiments to address this issue.

3.2 Incorporation of isotopically labelled 3 into the FeGP cofactor during the *in vitro* biosynthesis

By using the *in vitro* biosynthesis method, I investigated the biosynthetic precursor **3** of the FeGP cofactor. As shown in chapter 4.1, the addition of **3** was necessary to biosynthesize the [Fe]-hydrogenase holoenzyme in the assay using the cell extract from the *M. maripaludis* Δ*hcgB*Δ*hcgC* strain. This suggested that **3** is the real physiological precursor and the FeGP cofactor is biosynthesized from externally added **3**. This hypothesis is supported by the finding that the final [Fe]-hydrogenase activity produced by the *in vitro* biosynthesis assay is dependent on the concentration of **3** added (Figure 13a). To test incorporation of **3** into the FeGP cofactor, I prepared ^{15}N-labeled compound **3** ([$^{15}N_6$]-**3**), which contains six ^{15}N, and used it as the substrate of the *in vitro* biosynthesis. MS analysis of the FeGP cofactor produced in this assay carried six ^{15}N and the isotopic labelling pattern was identical to that of the precursor [$^{15}N_6$]-**3** (Figure 13b), which indicated that external pyridinol precursor **3** was indeed incorporated into the FeGP cofactor by the *in vitro* biosynthesis. Similarly, if the *ΔhcgBΔhcgC M. maripaludis* strain was grown in the presence of [^{15}N]-NH_4Cl and non-labelled **3** was added, the produced FeGP cofactor did not show any incorporation of ^{15}N, which again showed that the **3** used in the *in vitro* biosynthesis originates from the added pyridinol precursors (Figure 13c). The incorporation of [^{15}N]-NH_4Cl into the metabolites of this strain was shown for guanosine monophosphate as a positive control (Figure 13d).

This experiment shows that **3** is incorporated into the FeGP cofactor, but it cannot answer the question raised in Chapter 4.1.5, why the carboxyl group of **3** prepared from the ^{13}C-labeled FeGP cofactor contains a [^{13}C]-CO-derived ^{13}C. Two possible origins of the acyl ligand of the FeGP cofactor can be postulated: the carboxyl group of **3** is converted to the acyl ligand of the FeGP cofactor during the biosynthesis reaction or the acyl ligand can be formed from external CO, similar to the CO ligands. To answer this question, one of the two oxygen atoms of the carboxyl group of **3** was labelled with ^{18}O by hydrolysis of the acyl ligand during light decomposition in [^{18}O]-H_2O (see material & methods). If ^{18}O at the carboxyl group is incorporated into the acyl group, 50% of the acyl group will be labelled with the oxygen atom from the carboxyl group of ^{18}O-labeled **3**. This is because one of the two oxygen atoms of the carboxyl group should be lost during reduction of the carboxyl group to the acyl group and this process should be stochastic, as

the two oxygen atoms are chemically identical. If the carboxyl group is replaced by external CO during the *in vitro* biosynthesis, the ^{18}O label should be fully lost in the *in vitro* biosynthesis reaction. The ESI-MS analysis of the FeGP cofactor produced from 60 % $^{18}O_1$-labeled **3** showed about 30 % ^{18}O labelling, which fits well with the predicted behavior if the acyl ligand is produced by conversion of the carboxyl group (Figure 14a).

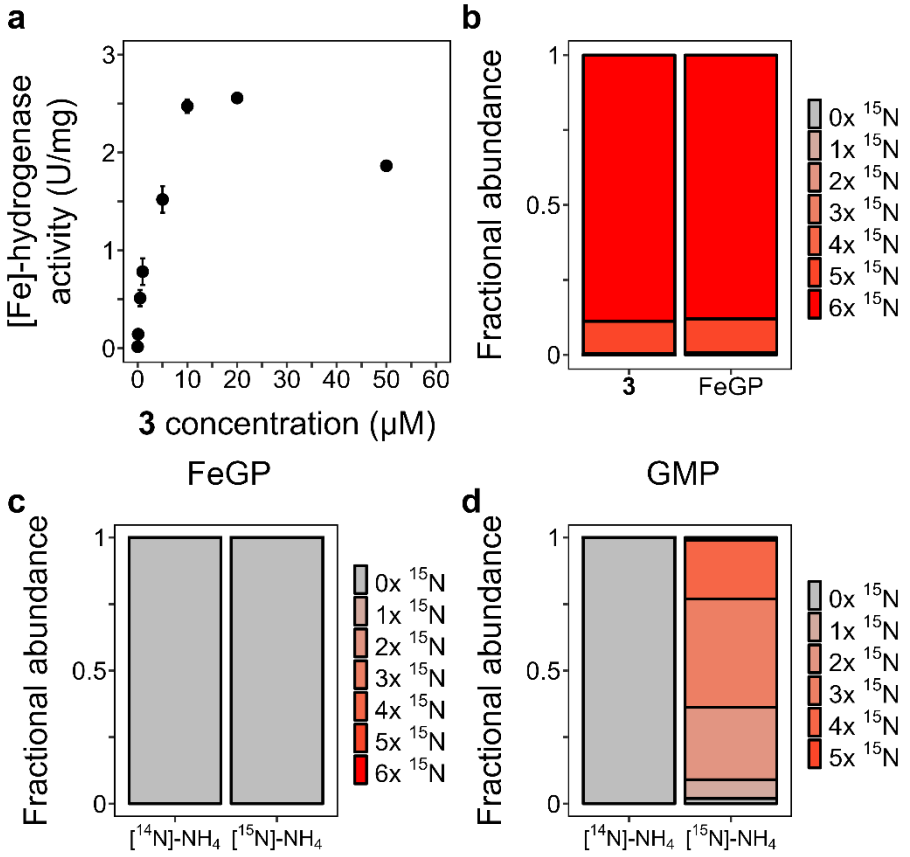

Figure 13. Incorporation of **3** into the FeGP cofactor during the *in vitro* biosynthesis. (a) [Fe]-hydrogenase activity dependence on the concentration of **3** in *in vitro* biosynthesis assy. (b) Incorporation of ^{15}N-labelled **3** into the FeGP cofactor by *in vitro* biosynthesis. Isotopologue fraction of the ^{15}N label in **3** from [Fe]-hydrogenase purified from *M. marburgensis* grown with [^{15}N]-NH_4Cl and the FeGP cofactor prepared by *in vitro* biosynthesis using the ^{15}N-labelled **3** as a precursor are shown. Approximately 90% of the molecules of both compounds were fully occupied with ^{15}N of the six nitrogen atoms. The regions of 0×, 1×, 2×, 3× and 4× ^{15}N species are too small to be seen in the bar graphs. (c) Incorporation of components from ^{15}N-labelled *M. maripaludis ΔhcgBΔhcgC* cell extract into the FeGP cofactor during the *in vitro* biosynthesis. (d) Incorporation of ^{15}N into GMP during the cultivation of *M. maripaludis ΔhcgCΔhcgB* shows the labelling of nitrogen containing compounds in this cell extract.

Another interesting question is the character of the first CO unit which is dissociated during the fragmentation of the FeGP cofactor during the MS/MS experiments, leading to the FeGP–$[CO]_1$ species (607 *m/z*). Based on the results

for the [^{13}C]-CO labelled FeGP (see Chapter 4.1.5), it was not possible to determine if this CO unit is the acyl or one of the CO ligands. The following analysis will follow the same procedure as the description of MS/MS experiments of the [^{13}C]-CO labelled FeGP in chapter 4.1.5. When the $^{18}O_1$ isotopologue of the $^{18}O_1$-FeGP was fragmented, the loss of one [^{12}C]-CO unit rather than the [^{18}O]-CO unit was observed, which indicates the acyl ligand is still part of the FeGP-$[CO]_1$ (609 *m/z*) from the $^{18}O_1$ isotopologue (Figure 14c). This observation indicates that the FeGP–$[CO]_1$ fragment is formed by dissociation of one of the two CO ligands and that the acyl ligand and the CO ligands are not scrambled during the *in vitro* biosynthesis reaction. This leaves the question of the origin of the ^{13}C label found in the **3** prepared by the light inactivation of [^{13}C]-CO labelled FeGP (Figure 12a). To test if scrambling can happen during the light induced inactivation of the FeGP cofactor, which was shown for other CO containing metalloenzymes, we determined the isotopologues of the **3** produced from this ^{18}O-labelled FeGP cofactor. Indeed, the isotopologue $^{18}O_1$ is decreased to about one-third of the FeGP cofactor after light inactivation, which indicated that the CO and acyl ligands are scrambled upon light decomposition. The liberation of ^{18}O from the carboxyl group can also occur during incubation in aqueous solutions at acidic pH. However these exchanges are slow and at neutral pH the ^{16}O/^{18}O exchange is < 5% during a time period of three days at room temperature for amino acids, which should not impact the measurements shown here.[99]

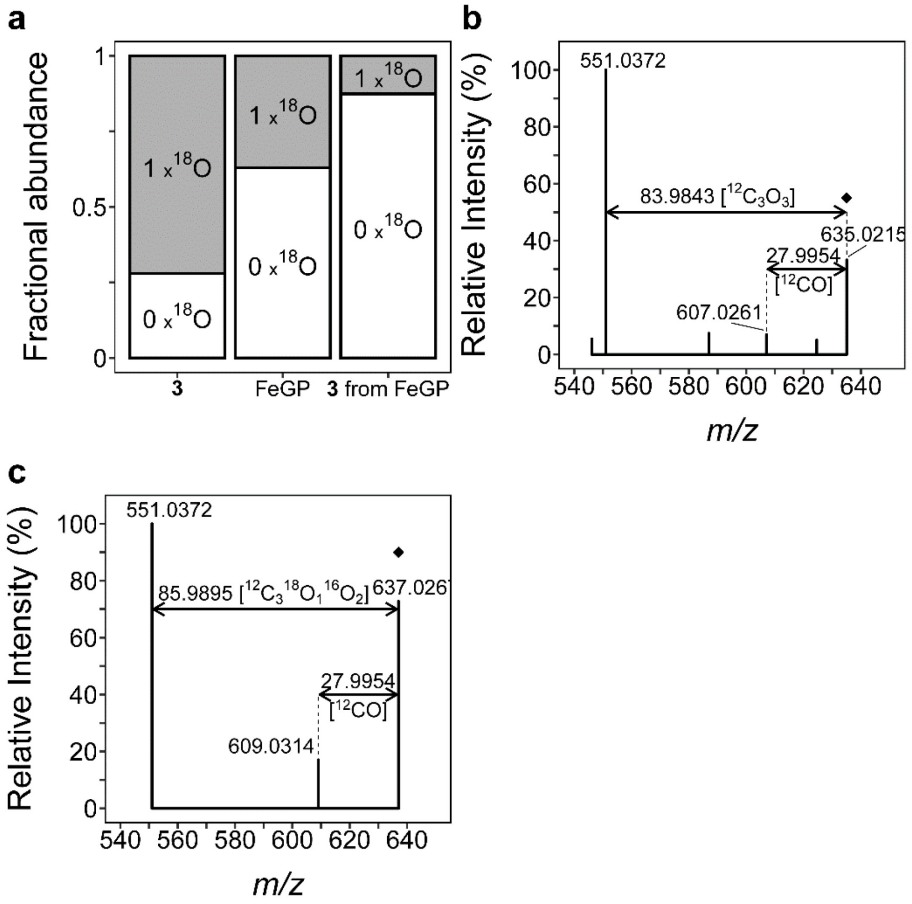

Figure 14. Incorporation of ^{18}O from [$^{18}O_1$]-carboxyl-**3** into the FeGP cofactor by in vitro biosynthesis. (a) Fractional abundance of the $^{18}O_0$ and $^{18}O_1$ isotopologues in **3** produced by light inactivation of the FeGP cofactor in [^{18}O]-H_2O, the FeGP cofactor synthesized by *in vitro* biosynthesis from [$^{18}O_1$]-carboxyl-**3** (FeGP) and the light decomposed product **3** from the [$^{18}O_1$]-FeGP cofactor. (b-c) MS/MS of the $^{18}O_0$ and $^{18}O_1$ isotopologues of the [$^{18}O_1$]-FeGP cofactor.

3.3 Investigation of cell extract components

One of the major components of the *in vitro* biosynthesis assay is the cell extract of the *ΔhcgBΔhcgC* strain and the other *M. maripaludis* mutants. Since the composition of the cellular materials of the cell extract could be changed dependent on the growth phase, the cell lysis method and other variables, their effect on the *in vitro* biosynthesis assay were tested. In this chapter I describe these analyses, which were required to improve the *in vitro* biosynthesis assay and what we can learn about the uncharacterized reactions needed to produce the FeGP cofactor.

3.3.1 pH controlled cultivation data

Initially, the *M. maripaludis* cell were grown with formate under 20% CO_2 / 80% N_2 in 500 ml bottle batch cultures. During growth on formate (Figure 15a), pH of the culture medium increased due to consumption of formate by the cells and growth

is inhibited at pH >7.5 – 8. Therefore, formic acid was added every hour to neutralize the pH. If the cells were harvested at an alkaline pH, the *in vitro* biosynthesis activity was greatly reduced. Therefore, the culture was neutralized before harvesting. Furthermore, the *in vitro* activity of the cell extract was dependent on the OD_{660} of the cell extract. The highest *in vitro* biosynthesis activity was detected at an OD_{660} of 0.6, which corresponds to the mid-exponential phase of the cells (Figure 15b).

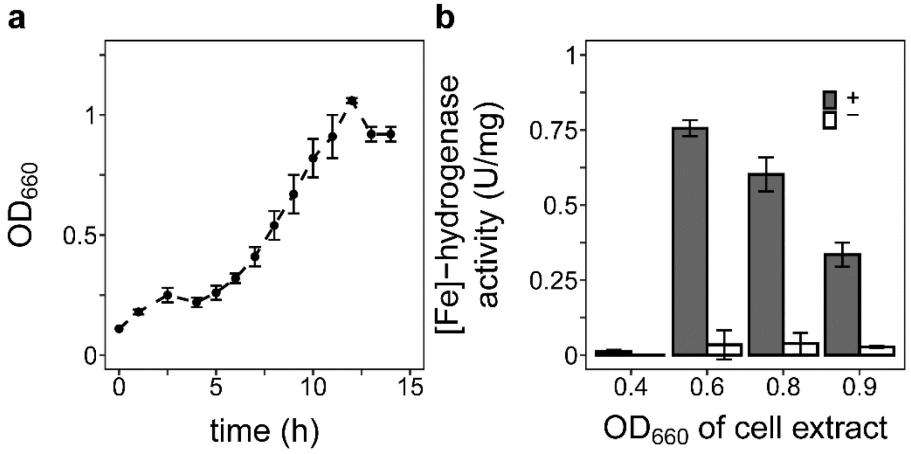

Figure 15. Growth of the Δ*hcgB*Δ*hcgC* strain of *M. maripaludis*. (a) Growth of the methanogen measured by OD_{660}. The culture was neutralized after each hour by addition of formic acid. (b) [Fe]-hydrogenase activity produced by *in vitro* biosynthesis using the cell extract with different growth phases. Aliquots of the culture after 9, 10, 12 and 13 hours were harvested, where the OD_{660} were 0.4, 0.6, 0.8 and 0.9, respectively. The OD_{660} values are shown on the abscissa. The *in vitro* biosynthesis in the presence (+) or absence (−) of pyridinol precursor **3** was performed.

To overcome the limitation of low cell mass in the batch bottle culture, I used a 5 l batch culture (see material & methods). In this case, the adjustment of pH could not be performed during the growth of the culture, because it is cultivated overnight and the medium pH during growth did not affect the *in vitro* biosynthesis activity. Before harvesting, the pH was adjusted to 7.0 by addition of formic acid, since harvesting the cells at high pH decreased the *in vitro* biosynthesis activity. The culture was cooled down in ice water before harvesting. The final OD_{660} in the 5 l culture was 0.6 − 0.8. This cultivation method was used to obtain the cells for the *in vitro* biosynthesis assay in this work.

3.3.2 Stability of the cell extract components of *M. maripaludis* for the *in vitro* biosynthesis

The cell extract provides the biosynthetic enzymes and possible non-protein components, which are needed for the *in vitro* biosynthesis. To obtain reproducible *in vitro* biosynthesis results and gain insights into the characteristic of these components, we analyzed the stability of these materials. Because many of the possible precursors of the FeGP cofactor biosynthesis could be unstable, the effect

of incubation at room temperature was tested. When the cell extract from the *M. maripaludis ΔhcgBΔhcgC* strain was incubated as it was under 95 % N_2 / 5 % H_2 at room temperature for 1 h and then used for *in vitro* biosynthesis, the *in vitro* biosynthesis activity was reduced to 70% (Figure 16a, black). When the standard *in vitro* biosynthesis assay solution was added to the cell extract and then incubated under 100% N_2 the *in vitro* biosynthesis activity was reduced to 20 % (Figure 16a, grey). The assay was started by transferring the solution to a 50 % CO / 50 % H_2 gas phase. The reduction of the *in vitro* biosynthesis activity indicates that some of the cell extract components were inactivated by incubation at room temperature, but the complete *in vitro* biosynthesis mixture with the cell extract is inactivated faster. To fully understand this effect, we need to also incubate the *in vitro* biosynthesis mixture alone and then start the reaction by addition of cell extract in future experiments. One of the explanations for this is the ATPase activity of the cell extract, which will deplete ATP in the assay solution. ATP is an essential component in the *in vitro* biosynthesis as described in chapter 4.5.1.

One of the major unsolved questions in the FeGP cofactor biosynthesis is the origin of the CO ligands and the CO forming reaction. One possibility is that the cell extract might contain an organometallic iron carbonyl compound as the substrate for the CO ligands. To test the presence of such a compound, the susceptibility of the cell extract to strong cold white light (3000K) was investigated since some organometallic iron carbonyl compounds can be decarbonylated by visible light, although decomposition under UV light is more common.[100] For the FeGP cofactor, it was found that it readily decomposes under UV and visible blue light,[18] therefore a possible precursor might also be sensitive to the same treatment. As shown in Figure 16b, exposure to cold visible light did not affect the *in vitro* biosynthesis activity, which suggests that light-sensitive compounds in the cell extract might not be involved in *in vitro* biosynthesis of the FeGP cofactor.

Another question is the electron source for the FeGP cofactor biosynthesis. If Fe^{2+} is used for *in vitro* biosynthesis, two-electrons are required for a reduction of the carboxyl group for the acyl ligand formation. For the reduction reaction, involvement of reducing substrate in the conversion of **3** to the FeGP cofactor was predicted.[45] Molecular oxygen (O_2) or reactive oxygen species formed from O_2 might oxidize the reducing substrates and deplete them. As shown in Figure 16b, exposure of the cell extract to air before addition of the *in vitro* biosynthesis solution resulted in approximately 50 % reduction of the activity, which indicates some of the components can be oxidized by O_2 or reactive oxygen species. This inactivation might also be due to the oxygen sensitive proteins, for example, possible radical SAM enzymes like HcgA and HcgG. However, it is difficult to draw a conclusion, because incubation for 1 h under anaerobic condition at room temperature also led to a 30% reduction in produced [Fe]-hydrogenase activity. In conclusion the investigation of the stability of the cell extract components showed

some oxygen, but no light sensitivity as well as a negative effect of prolonged incubation of the reaction mixture at room temperature. To better understand the components, which play a crucial role in the *in vitro* biosynthesis a more focused approach is needed. Therefore, I analyzed the effect of the precursor molecules (chapter 4.4) and substrates (chapter 4.5), ATP and SAM, added to the *in vitro* biosynthesis next.

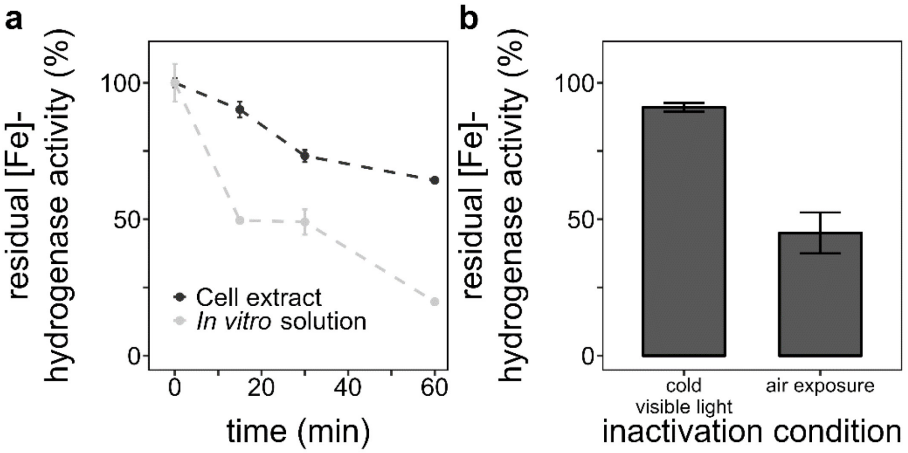

Figure 16. Stability of the cell extract from the *M. maripaludis* Δ*hcgB*Δ*hcgC* strain for *in vitro* biosynthesis. (a) The cell extract was incubated for the indicated time under N_2 and then mixed with *in vitro* biosynthesis assay solution for the *in vitro* biosynthesis assay (black). The cell extract containing the *in vitro* biosynthesis assay solution was incubated under N_2 and then the *in vitro* biosynthesis was started by exchanging the gas phase to 50% H_2 / 50% CO (grey). The residual [Fe]-hydrogenase activity is calculated by comparison to an *in vitro* biosynthesis experiment without incubation of the cell extract or the *in vitro* solution. (b) Effect of oxygen and light. The cell extract was incubated under air and dark condition or under 95% N_2 / 5 % H_2 under white light for 1 h and the *in vitro* biosynthesis activity was tested. These experiments were performed on ice for the light inactivation and at 8°C for the air exposure.

3.4 Pyridinol precursors and the specificity of HcgB and HcgC

3.4.1 Substrate promiscuity of HcgB and HcgC

The origin and the structure of the precursors of the FeGP cofactor, as well as the enzymes responsible for biosynthesis of the pyridinol precursors have previously been proposed based on *in vivo* isotopic labeling, structure-to-function and enzymological studies (Figure 18).[29, 41, 43-45] The proposed physiological substrates of HcgB and HcgC are supported by structural and biochemical data.[37, 41, 43-44] The presence of the 6-carboxyl-methylene group of **3** was indicated by the incorporation of the $^{18}O_1$-**3** during *in vitro* biosynthesis (see chapter 4.2).In order to analyze the reactivity of HcgB and HcgC to artificial pyridinol compounds, I used commercially available and chemically synthesized 2-pyridinol compounds. The products synthesized by the HcgB or HcgC reactions were analyzed by MALDI-

TOF-MS (Figure 17). It is to note, that the compounds without a 5-methyl and 6-methyl or carboxyl-methylene substitution (**4**, **5**, **6** and **7**) are highly unlikely to be the natural pyridinol precursors. This is due to the fact that the carbons at these positions originate from two molecules of acetate (see Figure 2). Compound **8** is a pyridinol compound with a 6-amino group and is not a candidate for the physiological precursor. While these compounds cannot be considered as pyridinol precursors, they can enable us to learn more about the promiscuity of the HcgB and HcgC reaction.

At first, conversion of compound **1** to **2** by HcgC, and **2** to **3** by HcgB were confirmed by this method. It was found that HcgB catalyzes guanylylation of **1**, which is proposed as the substrate of HcgC. HcgB catalyzed guanylylation of pyridinols with different 5- and 6-moieties (**2'**, **4** and **8**, Figure 17 and Figure S 2) and produced the guanylylated products **3'**, **11** and **19**. Although guanylylation of **5** by HcgB was reported,[44] this reaction was not detected by this method. The reaction of HcgB with the 6-amino-2-pyridinol (**8**) was not detected using the enzyme from *M. maripaludis*, but it was detected by using HcgB from *M. jannaschii* (Figure S 3). For the 3, 5-non-methylated compound **6**, trace amounts of the guanylylated products (**16**) were detected. I also tested the HcgB reaction in the presence of HcgC to check the reactivity to **12** and **15**; however, the predicted products (**14** and **17**) were not detected. Of particular interest is the reactivity of 3, 5, 6-trimethyl-4-hydroxyl-2-pyridinol (**2'**) with HcgB. The structure of **2'** is identical to the proposed physiological precursor **2** except for the missing carboxyl group at the 6-substitutent. Previous *in vivo* labelling experiments of the FeGP cofactor showed that the acyl ligand, which is proposed to be formed from the carboxyl group,[45] originates from both CO and CO_2 (Figure 2), which suggested that the acyl ligands can originate from CO rather than the carboxyl group of the pyridinol precursors.[42] This *in vivo* finding is in contrast with the finding that the carboxyl group of **3** is the precursor of the acyl ligand in the *in vitro* biosynthesis (chapter 4.2). Therefore it is of interest to check if 6-methyl pyridinol **2'** can also be incorporated into the FeGP cofactor in the *in vitro* biosynthesis.

The reactivity of HcgC was also analyzed in this experiment. This methyltransferase catalyzed the methylation of the pyridinol compounds (**1**, **5**, and **6**), which do not carry the 3-methyl group (Figure 17 and Figure S 4). MALDI-TOF-MS indicated that only one position was methylated by the HcgC reaction (Figure S 5). The reaction with **6** was very slow as indicated by the low intensity of the MALDI-TOF-MS signals (Figure S 5). These data indicated that HcgC has also promiscuous properties on the substrate utilization. Since these compounds have three possible methylation sites at 3-, 5- and 6-positions, it is interesting to investigate which of these can be methylated by HcgC although based on the proposed catalytic mechanism of HcgC, these pyridinol compounds would be methylated at the 3 position.[43]

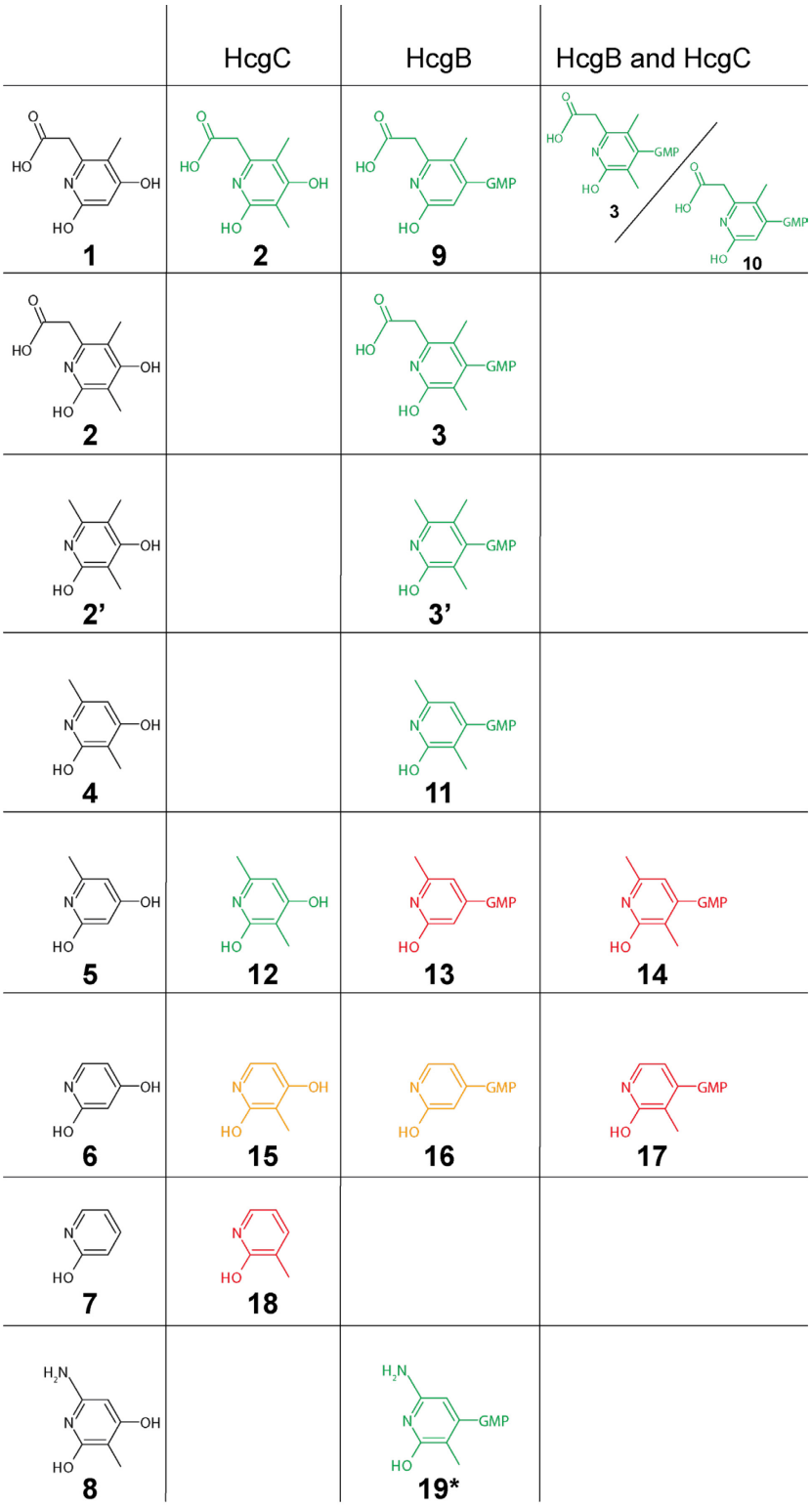

Figure 17. Reactivity of HcgB and/or HcgC to different pyridinol compounds, as detected by MALDI-TOF-MS. (Green) Detected with considerable intensity. (Orange) Detected with very low signal intensity and signal-to-noise ratio. (Red) Not detected by this method. The mass spectra for the reaction of 1, 5 and 6 with HcgB and HcgC simultaneously are shown in Figure S 6. *Detected only when HcgB from *M. jannaschii* was used.

3.4.2 Preference of the pyridinol substrates of promiscuous HcgB

HcgC and HcgB catalyze two consecutive reactions to produce guanylylpyridinol **3** in the biosynthetic sequence of the FeGP cofactor (Figure 18). As described above, HcgB reacts with various 4-hydroxy-2-pyridinols.[44] This finding raised the question whether HcgB catalyzes guanylylation of **1** in addition to the reaction with **2** in the methanogen cells. The guanylylation product **9** could interfere with biosynthesis of the FeGP cofactor. To answer this question, an HcgB assay using equimolar amounts of compounds **1** and **2** was conducted (Figure 19). HPLC analysis indicated that compound **2** was almost completely consumed within 5 min. On the other hand, only ~30% of **1** was converted to **9** after 105 min. This result clearly indicated that HcgB reacts with **2** much faster than **1**. The different reaction rates could explain why only the FeGP cofactor with the 3-methyl moiety is produced under the physiological condition. It is important to note that a bacterium, *Desulfurobacterium thermolithotrophicus*, which biosynthesizes the FeGP cofactor, contains a fused gene of *hcgB* and *hcgC*,[101] which suggests that the fusion protein complex might channel the HcgC reaction product to the active site of HcgB to avoid the unfavored reaction forming **9**. HcgB and HcgC are encoded in distinct genes in the genome of methanogens and previous blue-native PAGE analysis showed the absence of a heterodimer of HcgB and HcgC;[37] however, it cannot be excluded that HcgB and HcgC might produce a transient complex of HcgB and HcgC in the cell, in which **2** produced in HcgC is channeled to HcgB.

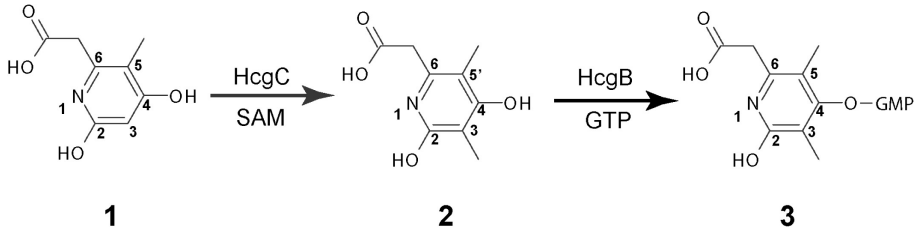

Figure 18. Proposed reaction sequence of HcgB and HcgC with pyridinol precursors as substrates and products.

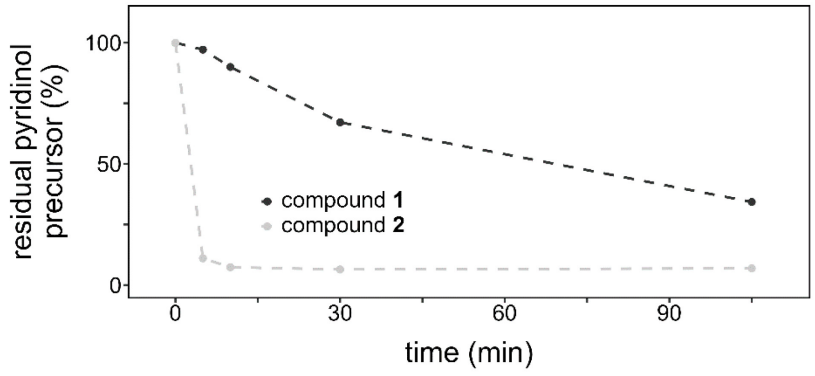

Figure 19. Competition assay for 3-non-methylated pyridinol (**1**) and 3-methylated pyridinol (**2**) in the HcgB reaction. The consumption of **1** and **2** were detected by HPLC**.**

3.4.3 *In vitro* biosynthesis from various pyridinol derivatives

With the *in vitro* biosynthesis assay, I was able to analyze the reaction of HcgB and HcgC also in the context of the full biosynthesis. The use of the *ΔhcgBΔhcgC M. maripaludis* strain, which does not produce HcgB or HcgC (Table S 1) allowed me to test the conversion of **1** and **2** to the FeGP cofactor by the reactions catalyzed by heterologously produced HcgB and HcgC. When **2** was used, the *in vitro* biosynthesis activity required the presence of HcgB and its co-substrate GTP. When **1** was used, the activity was detected only in the presence of HcgB and HcgC with GTP in the standard assay solution (Figure 20a). The co-substrate of HcgC is SAM, which was added to the standard assay solution. These results support the proposed reaction sequence, in which **1** was converted to **3** via **2** by the HcgC and HcgB reactions, and finally converted to the FeGP cofactor (Figure 18).

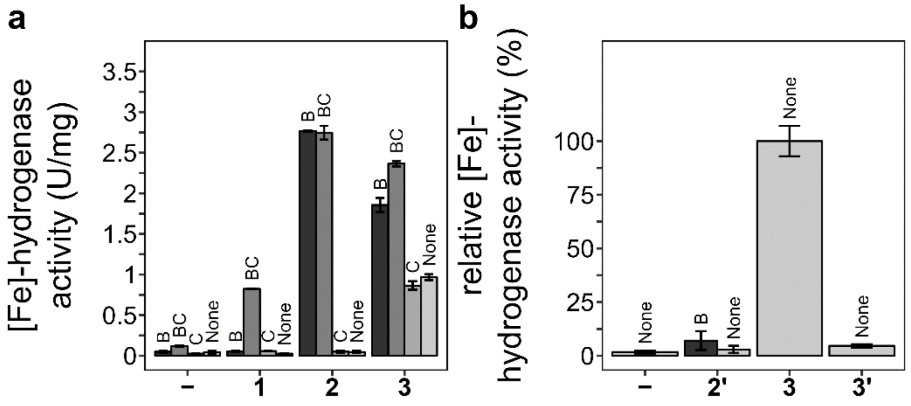

Figure 20. *In vitro* biosynthesis of the FeGP cofactor. (a) Activity obtained from the precursors by complementation with HcgB (B) and/or HcgC (C). (b) *In vitro* biosynthesis from **2'**, **3'** or **3**. The pyridinol precursors used in the assay are shown on the abscissa. The [Fe]-hydrogenase activity was normalized to the activity obtained from the reaction using **3** as a precursor. Compounds **1**, **2**, **2'**, **3**, **3'** or no precursor (−) were added to the assay (shown in abscissa). The Hcg enzymes added are indicated above the bar. (None) No Hcg enzymes added.

To test the importance of the carboxyl group at the 6-carboxyl-methylene substituent of the proposed pyridinol precursor **3**, I performed an *in vitro* biosynthesis experiment using derivatives of **2** and **3**, which contained a 6-methyl group rather than carboxyl-methylene group (**2'** and **3'**, Figure 17). *In vitro* biosynthesis with compound **2'** did not produce the [Fe]-hydrogenase activity in the standard *in vitro* biosynthesis assay containing additionally HcgB and GTP (Figure 20b), which suggests that **2'** is not functional in biosynthesis of the FeGP cofactor. However, this result can also be attributed to the slow reaction speed of HcgB in converting **2'** (5 pmol/min at 40°C, 5 mM GTP) compared to **2** (≥30 pmol/min at 40°C, 5 mM GTP) in the assay condition. The reaction speed of HcgB

with **2** is likely much higher than 30 pmol/min, because the reaction was already finished after the first measurement and Figure 19 shows HcgB can convert the same amount of **2** to **3** within 10 min with 10 mM GTP. To circumvent the possibility that conversion of **2'** by HcgB might be too slow to detect the *in vitro* biosynthesis activity, **3'** was synthesized by the reaction with HcgB for extended time (Figure S 7) and the produced **3'** was purified and used for the *in vitro* biosynthesis assay. The *in vitro* biosynthesis with **3'** again showed no activity, which indicates that the carboxyl group is required for the biosynthesis of the FeGP cofactor (Figure 20b).

These experiments confirmed the structure of the pyridinol precursors in the FeGP cofactor biosynthesis, which were proposed based on the results from the investigation of HcgB and HcgC.[37, 41, 43-44] Since HcgB preferably converts **2** instead of **1** and HcgC cannot convert **9** into **3** based on structural restrains [41] the order of the biosynthetic reactions shown in Figure 18 could be confirmed. Furthermore, the presence of the carboxyl group of the 6-substituent was necessary for the *in vitro* biosynthesis indicating the carboxyl group is the precursor of the acyl ligand, which was also indicated in this study by the mass spectrometric analysis (chapter 4.2). For this conversion, the action of HcgE and HcgF are thought to be crucial. In the next chapter, I therefore investigated the function of these two enzymes using the biochemical and *in vitro* biosynthesis assays.

3.5 Role of HcgE and HcgF in biosynthesis of the FeGP cofactor

3.5.1 Activation of the carboxyl group of the guanylylpyridinol precursor 3 by HcgE

The first step of the conversion of the carboxyl-group of **3** is thought to be catalyzed by HcgE. It has been previously proposed that HcgE adenylylates the 6-carboxyl-methenylene group of **3** and then the carboxyl group is bound covalently to HcgF by a thioester bond to the Cys9 residue (amino acid numbering is based on the *M. jannaschii* enzyme).[45] This activation of the carboxyl group enables the reduction from a carboxylic acid to an aldehyde that is the same oxidation state with the acyl ligand in the FeGP cofactor. Similar reductions of activated carboxyl groups have already been described in the catalytic cycle of carboxylic acid reductase.[51]

To test if the thioester adduct of **3** on HcgF (**5**) can be reduced, I conducted a reduction assay using $NaBH_4$ to reduce **5** (Figure 21a). After incubation in the presence of HcgE, HcgF and ATP, the reaction product was reduced with $NaBH_4$. I predicted formation of an 6-acetaldehyde compound and/or 6-hydroxyl-ethyl compound (**3^{OH}**). MALDI-TOF-MS indicated the formation of only a compound with 529 *m/z*, which corresponds to **3^{OH}** (Figure 21b). Addition of HcgF decreased the yield of **3^{OH}**, indicating it is not required for the activation of **3** and subsequent

chemical reduction by $NaBH_4$. (Please refer to the results indicated in Figure 22a and the associated text in 4.5.2.) The reduction reaction was dependent on time (Figure 21c) as well as the concentration of ATP and HcgE (Figure 21d, e), which indicates that the not-activated carboxyl group of **3** cannot be reduced by $NaBH_4$. In general, $NaBH_4$ cannot reduce carboxylic acids alone, but can reduce activated carboxylic acids.[102] Due to the large excess of $NaBH_4$ used and the higher reactivity of the aldehyde lead to the reduced product as only an alcohol.[102] In contrast, the enzymatic reduction by reductases will form primarily the aldehyde and not the alcohol.[51] These results indicate that the AMP-**3** adduct formed by the HcgE reaction is reduced by $NaBH_4$. The involvement of ATP in the biosynthesis of the FeGP cofactor could also be confirmed in the *in vitro* biosynthesis assay. This reaction is dependent on the addition of ATP (Figure 21f), possibly due to the HcgE reaction.

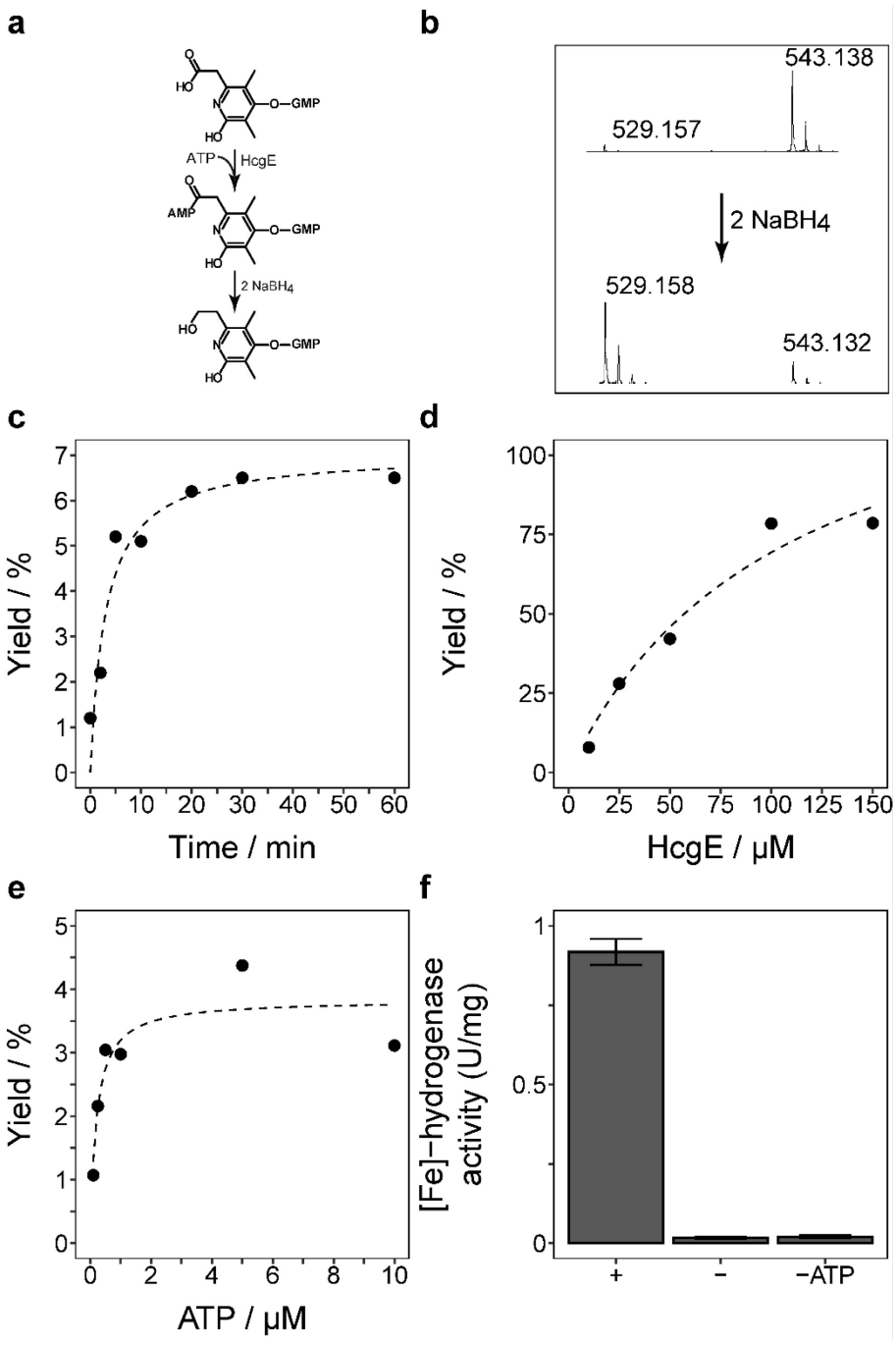

Figure 21. Analysis of the activated products of **3** formed by HcgE. (a) Proposed reactions of HcgE,[45] and the $NaBH_4$ reduction product from the adduct formed by the enzyme reaction. (b) MALDI-TOF-MS of 6-hydroxyl-ethylated **3** ($\mathbf{3^{OH}}$) produced by reduction with $NaBH_4$ after an assay containing **3**, HcgE, HcgF and ATP. (c) Kinetics of $\mathbf{3^{OH}}$ formation during incubation of the assay mixtures at room temperature. HcgE concentration used in this assay was 100 µM. Yield of $\mathbf{3^{OH}}$ from **3** was presented as the conversion rate (%) determined from the relative intensities of the mass peaks at 529 and 543 *m/z* (see Materials & Methods for details). Effect of HcgE (d) and ATP (e) on the production of $\mathbf{3^{OH}}$ in the assay. Incubation time of the reaction was 30 min. (f) The *in vitro* biosynthesis assay using the Δ*hcgB*Δ*hcgC* strain with **3** as the precursor in the presence (+) and absence of ATP (−ATP). (−) A negative control with ATP in the absence of **3**.

3.5.2 HcgF is not required for *in vitro* biosynthesis of the FeGP cofactor

It was thought that the formation of a thioester with HcgF to form **5** is needed to stabilize the activated carboxylic acid in **3** prior to reduction to the acyl moiety. As shown in Figure 22a, HcgF is not required for the reduction by $NaBH_4$ and omission of HcgF can increase the amount of produced $\mathbf{3^{OH}}$ by reduction with $NaBH_4$. Next it was investigated if HcgF can bind **3** in activated (with addition of ATP and HcgE) or non-activated conditions, since it was reported that HcgF can covalently bind **3** without activation by HcgE.[45] For this analysis, an ultrafiltration method was used, where **3** is retained in the concentrate if it binds to the protein or migrates into the filtrate in case no binding occurs. As shown in Figure 22b, **3** is found in the concentrate, independent on the addition of HcgE and ATP. The enrichment of **3** in the concentrate, which was dependent on addition of HcgF (Figure 22d), indicated that it is retained by HcgF in this assay. This result indicates that HcgF binds **3** independent on the activation by HcgE. This is in agreement to the crystal structure of HcgF, in which **3** was bound to HcgF without activation by HcgE. Although this filtration experiment does not show whether **3** is covalently bound to Cys9, a notable information was obtained by UV-Vis spectra of the concentrate containing the complex of HcgF and **3**, in which the absorbance peak of **3** was shifted from 300 nm to 310 nm. This shift was also observed in the absence of ATP (Figure 22c). This result could be attributed to the formation of the thioester adduct with HcgF independent on the presence of ATP. It is important to note that for the crystal structure analysis and the filtration assay, HcgF was purified under aerobic conditions.[45] However, it cannot be excluded that the behavior of Cys9 can be different under anoxic/aerobic conditions because under aerobic conditions, Cys9 can be oxidized as indicated by mass spectrometric analysis and no oxidation adducts to Cys9 were detected in anoxically purified HcgF (Figure S 8). However, only two peptide single matches were detected from the anoxically purified HcgF, which makes conclusive analysis difficult.

The above finding prompted me to assess the function of HcgF in the biosynthesis of the FeGP cofactor *in vivo* and *in vitro*. To check the *in vivo* biosynthesis, the *M. maripaludis* Δ*hcgF* strain was used. The Δ*hcgF* strain has previously been shown to have similar growth properties to the Δ*hmd* strain under H_2 / CO_2 atmosphere.[40] According to this finding, this strain did not show any [Fe]-hydrogenase activity, which indicated that HcgF is essential in the biosynthesis of the FeGP cofactor *in vivo*. Unexpectedly, even in the absence of HcgF, some [Fe]-hydrogenase activity was detected by *in vitro* biosynthesis (Figure 22d) and addition of **3** to the assay increased the *in vitro* biosynthesis activity. Addition of HcgF only slightly increased the [Fe]-hydrogenase activity in the assay. The C9A mutation of HcgF provided almost no effect on the [Fe]-hydrogenase activity and a similar effect was observed, when the C119A mutated enzyme was used (Figure 22d). The Cys119 in contrast to the Cys9 is not thought to be involved in the binding of the FeGP cofactor to HcgF.[45] These results indicated that HcgF is not required for the *in vitro* biosynthesis under the tested

conditions. Further evidence for this finding and a discussion of possible reasons is given in chapter 4.6.2.

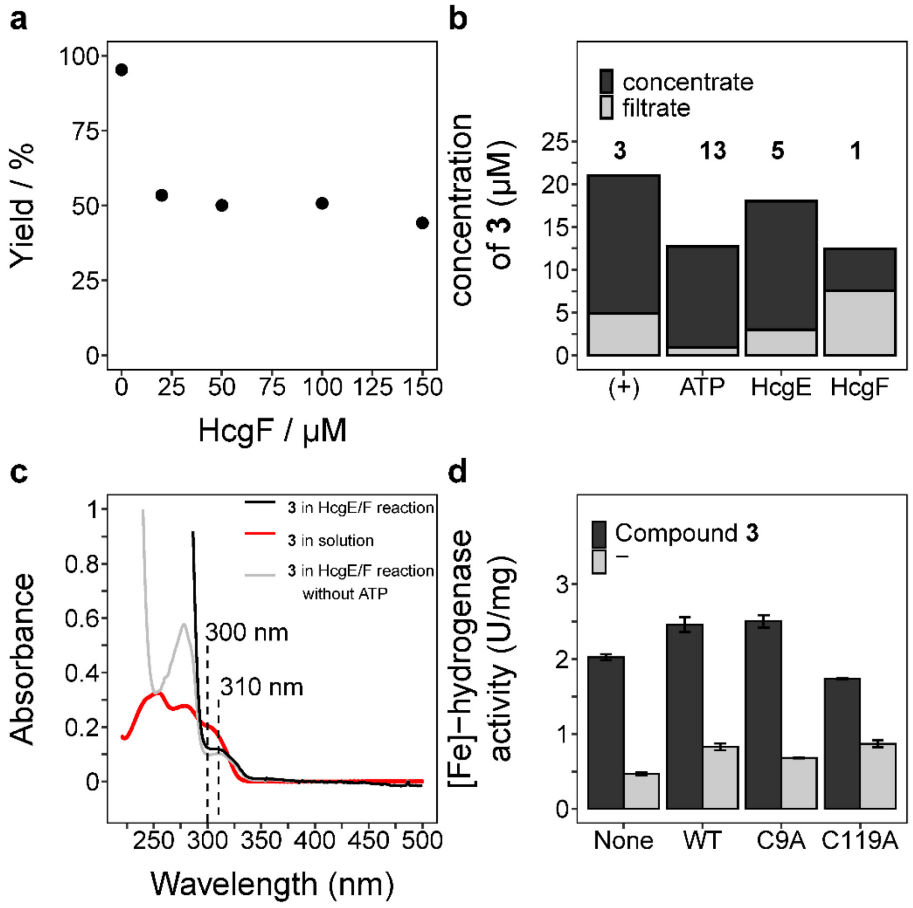

Figure 22. Function of the HcgF protein in the biosynthesis of the FeGP cofactor. (a) Effect of HcgF on the production of the activated adduct **3** formed by HcgE and/or HcgF in the presence of ATP. The adduct formation was evaluated by formation of 6-hydroxyl-ethylated **3** (**3**^{OH}) by reduction with $NaBH_4$. Yield of **3**^{OH} from **3** was presented as the conversion rate (%) determined by MALDI-TOF-MS. (b) Ultrafiltration experiment to determine the binding of **3** to HcgF. The concentration of compound **3** in the filtrate (grey) and the concentrate (black) is shown. The enrichment of **3** in the concentrate compared to the filtrate is shown above each bar. The compound missing from the full assay (HcgE, HcgF, ATP and **3**) is shown in the abscissa. (+) no compound was omitted. (c) UV/Vis spectrum of **3** in the concentrate of the HcgE/F ultrafiltration assay (black and red) compared to the spectrum in solution (grey), adopted from.[25] (d) *In vitro* biosynthesis in the Δ*hcgF* strain with (black) or without **3** (grey). The addition of HcgF and its C9A and C119A variants is indicated on the abscissa.

3.5.3 External *S*-adenosyl methionine is not required for the *in vitro* biosynthesis of the FeGP cofactor

Apart from the utilization in the HcgE reaction, ATP could also be used to regenerate other co-substrates, which are required for the *in vitro* biosynthesis. One such compound is SAM, which is known to be involved in biosynthesis of the [FeFe]-hydrogenase cofactor, especially in the synthesis of the CO and CN ligands and preparation of the iron center.[98b] In the case of the FeGP cofactor, SAM is the methyl-donor substrate of HcgC and might be a co-substrate of HcgA and HcgG (see below). To assess if SAM is necessary for the conversion of **3** to the FeGP cofactor, we omitted SAM and also added inhibitors of SAM dependent enzymes in the absence of SAM. Neither condition showed a reduction in the activity of *in vitro* biosynthesis. This might indicate that no SAM dependent enzyme is required; however, in the *in vitro* biosynthesis assay an excess of ATP is present, which can be used by the cell extract protein to regenerate SAM.[103] The cell extract might not only contain the necessary Hcg proteins, but also other small compounds and maybe even enzymes, which are required for the *in vitro* biosynthesis. In the next chapter I therefore analyze the need for the protein and small compounds for the *in vitro* biosynthesis as well as the minimum proteins required to catalyze the conversion from **3** to the FeGP cofactor.

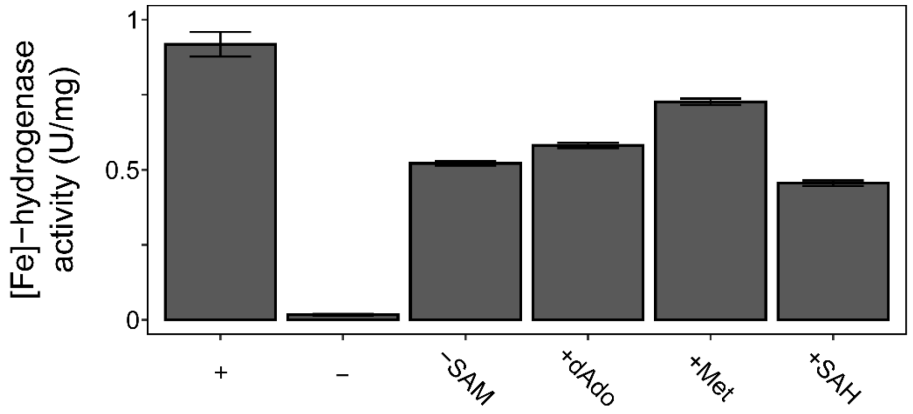

Figure 23. Effect of SAM and inhibitors for SAM dependent enzymes on *in vitro* biosynthesis. The *in vitro* biosynthesis assay was performed using the *ΔhcgBΔhcgC* strain with **3** as the pyridinol precursor. Control experiments in the presence (+) and absence (−) of **3**, and in the absence of SAM (−SAM). Inhibitors for SAM dependent enzymes: 5'-deoxyadenosine (dAdo), methionine (Met), and *S*-adenosyl homocysteine (SAH) were tested in the absence of SAM.

3.6 Hcg proteins and cellular compounds required for the in *vitro* biosynthesis

3.6.1 The *M. maripaludis* cell extract contains small components required for the *in vitro* biosynthesis

The cell extract of *M. maripaludis* contains small molecules, which might be required for the *in vitro* biosynthesis of the FeGP cofactor. To test this hypothesis, the cell extract from the Δ*hcgB*Δ*hcgC* strain was separated by ultrafiltration into a filtrate, which contains components with a molecular weight of <3 kDa, and a concentrate, which contains molecules with a molecular weight >3 kDa, mostly proteins. Neither the concentrate, which was diluted with buffer to the initial volume, nor the filtrate catalyzed the *in vitro* biosynthesis reaction. However, when the two solutions were mixed again, the *in vitro* biosynthesis activity was restored (Figure 24a). The concentrate contained the Hcg proteins, which are known to be required for the reaction. The requirement of the filtrate indicated that some filtratable components are needed for *in vitro* biosynthesis. This finding is in accordance with the finding that the CO ligands originate from a component in the cell extract (see Chapter 4.1.5 and Figure 10b). Furthermore, the reduction of the carboxyl group to the acyl ligand likely requires a two electron reduction step, which might also involve a small molecule (<3 kDa) to provide electrons. The need for a reduction is also indicated by the fact that H_2 or formate are needed in the absence of CO for the *in vitro* biosynthesis (see chapter 4.1.2 and Figure 7). The activation of the carboxyl group was already pointed out in the experiments on the activation by HcgE in chapter 4.5.1.

Another question is why the *in vitro* biosynthesis reaction cannot completely convert the pyridinol precursors supplied in the assay solution. The yield of the FeGP cofactor is lower than 10% of the pyridinol precursors added. One possible explanation is that another compound, possibly the above mentioned small components, is limiting under these conditions. To test this hypothesis, the *in vitro* biosynthesis assay solution after completion of the assay was filtered to obtain the filtrate, in which the essential compounds could be depleted by the first round of *in vitro* biosynthesis. When this filtrate was supplemented with fresh concentrate, the *in vitro* biosynthesis activity was comparable to the initial activity (Figure 24b, reused F+C). This indicates that the proteins in the concentrate were inactivated during the first reaction but the filtrated small compounds did not lose the activity. The instability of the protein fraction might be the major reason for the incomplete conversion of **3** to the FeGP cofactor.

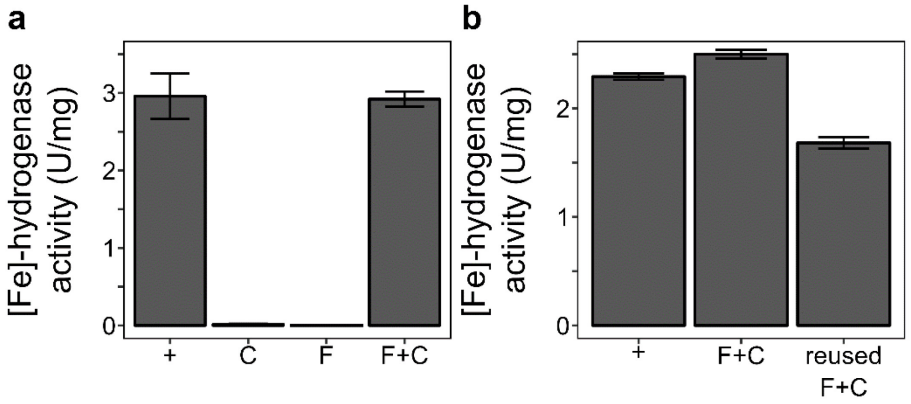

Figure 24. Analysis of the cell extract compounds required for *in vitro* biosynthesis of the FeGP cofactor. (a) Fractionation of the essential components in the cell extract of the *M. maripaludis* Δ*hcgB*Δ*hcgC* strain by 3 kDa ultrafiltration. *In vitro* biosynthesis using pyridinol precursor **3** with the whole cell extract (+), with the concentrate (C), the filtrate (F) and the mixture of the filtrate and concentrate (F+C) are shown. (b) Effect of reused filtrate. The *in vitro* biosynthesis assay with the whole cell extract (+), with filtrate and concentrate (F+C), and with the filtrate from the first assay and fresh concentrate (reused F+C).

3.6.2 HcgE and HcgG are the only Hcg proteins required for *in vitro* biosynthesis from 3

Comparative genomics analysis of methanogens revealed that only the Hcg genes and their respective protein products were found to be responsible for the biosynthesis of the FeGP cofactor. To test if only the Hcg proteins are required as proteins for the *in vitro* biosynthesis, we developed an updated version of the *in vitro* assay, in which the cellular protein was replaced with heterologously produced Hcg proteins. This work was conducted in collaboration with F. Arriaza, who homologously overproduced HcgG in an *M. maripaludis* strain, which contains an expression vector with a His_6 *hcgG* fusion gene. The strain was constructed and provided by Kyle Costa (University of Minnesota). We tested the conversion of **3** to the FeGP cofactor by using heterologously produced HcgD, HcgE, HcgF and HcgG in the presence of the filtrate prepared from the cell extract from *M. maripaludis* Δ*hcgA*. The Δ*hcgA* strain cell free extract and filtrate was chosen since HcgA is responsible for the synthesis of **1** (F. Arriaza, PhD thesis) and therefore no pyridinol precursors should be present in the filtrate of this strain. The need for the filtrate of *M. maripaludis* cell extract was predicted based on experiments in chapter 4.6.1, which indicated components from the filtrate are necessary for the biosynthesis of the FeGP cofactor. Furthermore, the same mixture of possible co-substrates (ATP, DTT, dithionite, Fe^{2+} and SAM), which was used for the *in vitro* biosynthesis based on the cell extract of *M. maripaludis* Δ*hcgB*Δ*hcgC* was added. In addition, the cell extract was reduced by incubation with 100% H_2 for 1 h at 40°C before the filtration to provide reducing equivalents to the reaction because my preliminary experiments indicated that the pre-reduction with H_2 improved the *in vitro* biosynthesis activity.

Interestingly, addition of CO to the *in vitro* assay reduced the *in vitro* biosynthesis activity substantially, which indicates that at least one reaction to transform **3** into the FeGP cofactor is sensitive to CO (Figure 25a). Furthermore, the addition of filtrate and HcgG was required, indicating both are important for the conversion of **3** to the FeGP cofactor. The stimulatory effect of the pre-treatment with H_2 of the cell extract might be caused by the reducing equivalents that can be generated by hydrogenase activity in the cell extract under these conditions. This finding indicates that another enzyme is required for providing the reducing equivalents to the system, apart from the Hcg proteins (Figure 25b).

To identify the minimum set of Hcg proteins needed, we tested the omission of each Hcg protein. The lack of HcgF did not affect the *in vitro* biosynthesis and omission of HcgD decreased the yield by 50%, but was not necessary for this reaction (Figure 25c). Thus, only HcgG and HcgE are required for the *in vitro* biosynthesis reaction from **3** using the filtrate of the reduced cell extract (Figure 25c). In the case of HcgF this finding is in accordance with the *in vitro* biosynthesis assay based on Δ*hcgF* strain (see chapter 4.5.2). The mutant was unable to produce FeGP cofactor *in vivo*, however, in the *in vitro* biosynthesis assay the omission of HcgF did not stop the production of [Fe]-hydrogenase activity (Figure 22). This finding might indicate that HcgF is only needed *in vivo*. HcgF might be unnecessary to stabilize the activated AMP-**3** (**4**) *in vitro* because the ATP concentration in the *in vitro* assay is 40-fold higher compared to the *in vivo* concentration of other methanogens.[104] Since the thioester intermediate and the mixed anhydride of the AMP-acyl are thermodynamically similar, the following reduction could be promiscuous towards **4** or HcgF-S-**3** (**5**).[105] Another possibility is that **4** can spontaneously react with thiol compounds present in the assay, such as DTT, which is known for other Acyl-AMP compounds, e.g. AMP-biotin, which was activated by biotin ligase.[106] This thioester might then be able to be used for the following reactions by HcgG.

These initial results show that the *in vitro* biosynthesis can be further updated to gain more information on the biosynthetic reactions from **3** to FeGP. However, reproducibility of these interesting results in this section was not high, probably because the *in vitro* biosynthesis activity was strongly dependent on the quality of the HcgG protein preparation. In the next chapter I will discuss possible further experiments and also summarize the here presented results and our conclusion regarding the biosynthesis of this unique and exciting cofactor.

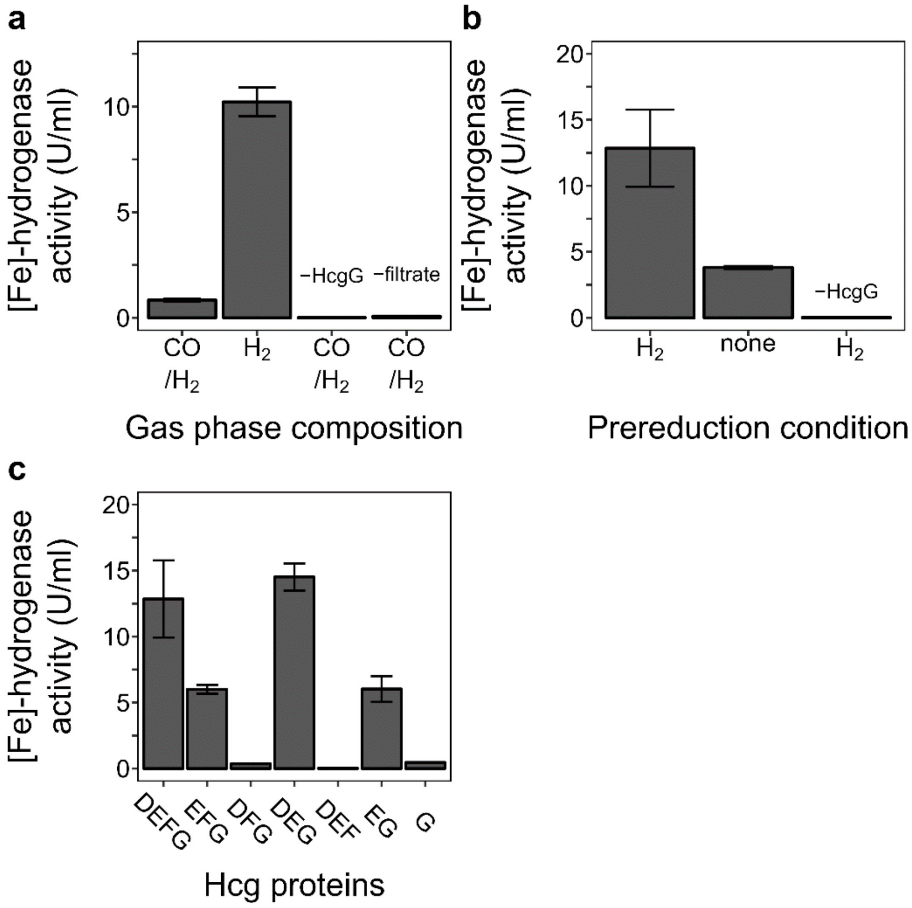

Figure 25. *In vitro* biosynthesis of the FeGP cofactor using heterologously produced HcgDEFG proteins. (a) *In vitro* biosynthesis under 50% H_2 / 50% CO and 100% H_2, and negative controls without HcgG or the filtrate under 50% H_2 / 50% CO. (b) The effect of pre-reducing the cell extract by H_2 prior to preparation of the filtrate. (−) Negative control without HcgG. (c) Requirement of the Hcg proteins in *in vitro* biosynthesis from pyridinol precursor **3**. Full *in vitro* biosynthesis was performed with the reduced filtrate and HcgDEFG. To test the function of Hcg proteins in *in vitro* biosynthesis Hcg proteins were omitted from the assay solution. The Hcg proteins added to the solution are shown in the abscissa. The *in vitro* biosynthesis reaction was performed under 100 % H_2.

4 Conclusion and outlook

In this study I developed an *in vitro* biosynthesis assay for the FeGP cofactor of the [Fe]-hydrogenase. This assay system can start from chemically synthesized precursors, which I used to show that compounds **1**, **2**, and **3** are precursors of the FeGP cofactor. The conversion of these precursors to the FeGP cofactor was verified by incorporation of isotopically labelled precursors into the cofactor during the *in vitro* biosynthesis. I showed that heterologously produced and purified HcgB and HcgC protein can complement deletion of their genes in *M. maripaludis ΔhcgBΔhcgC* strain that was used to prepare the cell extract for the *in vitro* biosynthesis. This finding indicates that the assay can also assess the activity of heterologously/homologously produced and purified Hcg proteins. Analyzing the function of HcgB more in-depth, I could also show that compound **1** is not converted in the time frame of our *in vitro* experiment to an active derivate of the FeGP cofactor in the absence of HcgC. This is further supported by our finding that HcgB preferably catalyzes the guanylylation of compound **2** instead of compound **1**, which was proposed based on previous structure-to-function analysis of HcgC.[37, 41, 43] Furthermore, we showed that the carboxyl group of these precursors is reduced to the acyl ligand of the FeGP cofactor. In summary, the *in vitro* assay can be used to investigate three types of questions regarding the biosynthesis of the FeGP cofactor: 1) to test the incorporation of the possible precursors including isotopically labelled precursor moieties; 2) to predict the reaction sequence of the biosynthesis; and 3) to verify the activity of heterologously produced Hcg proteins in the biosynthesis of the FeGP cofactor.

We were able to expand the *in vitro* biosynthesis assay, using the cell extract of the *ΔhcgBΔhcgC* mutant, to other mutant strains and to no include cell extract proteins. For example the *ΔhcgA* mutant showed that HcgA is involved in the first step of the biosynthesis, the formation of precursor **1**. Furthermore the analysis we showed that HcgG is involved in the insertion of the iron center into the framework of the guanylylpyridinol **3** (See PhD Thesis of F. Arriaza).[107] I replaced the protein components of the cell extract with purified Hcg proteins and conducted the *in vitro* biosynthesis assay without cell extract proteins. Although the results are not stably reproducible, the initial experiments suggest that only HcgE and HcgG are necessary for the *in vitro* biosynthesis from compound **3**. Addition of HcgD improves the FeGP cofactor yield, while HcgF has no effect on this assay's results. The important role of HcgE was further confirmed by the fact that the *in vitro* biosynthesis is highly dependent on ATP when using all cell extract components. The *in vitro* reaction also depends on the presence of small molecules from cell extract, which may involve a CO ligands precursor and an electron carrier, the latter of which is used for the reduction of the carboxyl group to the acyl ligand.

This results in an updated model for the *in vitro* biosynthesis, as shown in Figure 26. To prove this model and answer the remaining questions a range of different

experiments are necessary. Especially the functional characterization of HcgA, HcgD, HcgF and HcgG is of substantial importance which have not been fully understood yet. Hence, the investigation of HcgA and HcgG has a high priority since they catalyze the crucial first and last steps of the biosynthesis of the FeGP cofactor.

In the case of HcgA, the substrate identity, as well as the structure and mechanism of the protein, need to be elucidated. The substrate was found to be present in the cell extract of bacteria and archaea, indicating it is a common metabolic compound (PhD Thesis, F. Arriaza). Fractionation of cell extract combined with mass spectrometry may reduce the list of possible compounds, which can be realistically evaluated based on the HcgA catalyzed reaction. The structure and mechanism can subsequently be addressed by biochemical and crystallographic work. This investigation may reveal a new radical mechanism to synthesize highly modified pyridine compounds, which is a mayor improvement for biotechnological synthesis of complex chemicals. In contrast to HcgA, the function of HcgG is not very well understood and the elucidation of its substrates is important for understanding the final steps of the FeGP cofactor biosynthesis. HcgG is thought to be involved in three major reactions; the synthesis of CO from an organic precursor, the binding of the CO ligands to the iron atom and the reduction of the carboxyl to the acyl ligand. To investigate the function of this highly complex enzyme, a range of different studies are suggested. At least two substrates apart from compound **3** and SAM are required in addition to HcgG, the electron carrier and the precursor of the CO ligands. Since cell extract fractionation is experimentally challenging if two substrates are unknown, identification of the electron carrier for HcgG is required prior to investigating the CO precursor. Using purified proteins in the absence of other cell extract proteins, the *in vitro* biosynthesis assay can identify the electron carrier. In these conditions the reduction of the cell extract's small molecules by H_2 is necessary and adding the reduced electron carrier will remove this prerequisite. Therefore this assay can be used to identify the electron carrier of this reaction. Subsequently, the CO precursor can be identified analogously to the substrate of HcgA. Instead of product detection by mass spectrometry, the *in vitro* biosynthesis can be used as a readout. Independent of the substrate, the structural analysis of HcgG is of great interest, since it lacks homologues. Together with spectroscopic analysis of the likely Fe-S cluster of HcgG and IR analysis of possible CO containing intermediates the structural analysis will play a major role in determining the mechanism of this unique enzyme.

The proteins HcgD and HcgF have already been structurally characterized and possible functions have been proposed.[45-46] The new results from the *in vitro* biosynthesis enable a more in-depth analysis of the function of HcgD and HcgF, similar to the analysis of HcgB and HcgC in this work. In the case of HcgD, the transfer of iron from this enzyme to the FeGP cofactor can be investigated by loading this enzyme with isotopically labelled iron. The transfer of this iron to the FeGP cofactor might then be tracked by mass spectrometric analysis after the *in*

vitro biosynthesis in the presence of ^{57}Fe-loaded HcgD. Hence the function of a Nif3-like protein would be conclusively described across all domains of life for the first time. Similarly, to HcgD the function of HcgF was partially uncovered in a previous study. HcgF was proposed to be a chaperon, binding the activated compound **3** after the HcgE reaction. To investigate this function, it will be important to test the formation of HcgF-Cys9-S-**3** using Cys9 mutants. Furthermore, the function of HcgF-Cys9-S-**3** and AMP-**3** conjugates as a substrate of the *in vitro* biosynthesis in the absence of HcgE. Since thioester bonds are proposed to be highly labile the isolation of this protein conjugate might be experimentally challenging. A possible alternative is the chemical synthesis of an activated ester version of **3**, which is more stable under the experimental conditions.

The development of the here presented *in vitro* biosynthesis enables many exciting new studies to improve our understanding of the FeGP cofactor biosynthesis. Similar to *in vitro* biosynthesis methods developed for other metallocofactors, this method can help the future analysis of this unique and fascinating molecule.

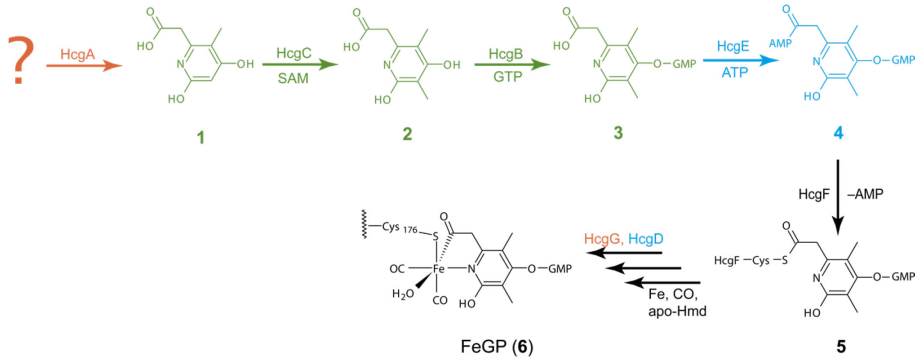

Figure 26. Updated scheme of the FeGP biosynthesis. In green reaction and intermediates are shown, which were proposed before and were confirmed by the *in vitro* biosynthesis assay developed in this work. In orange enzymes are displayed, which function was elucidated by F. Arriaza based on the *in vitro* biosynthesis assay system developed here. The function of HcgD and HcgE (cyan) are strengthened by the here presented work but not fully understood yet.

5 Supplementary figures

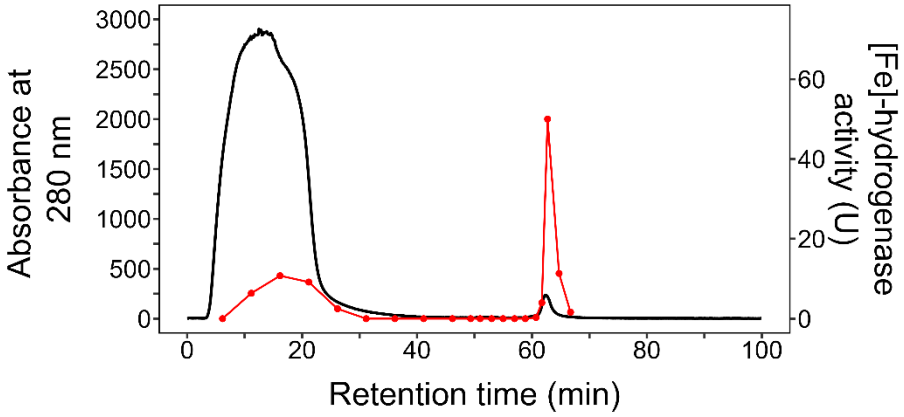

Figure S 1. Purification of holo [Fe]-hydrogenase from *M. maripaludis* from the in vitro biosynthesis solution by strep tag affinity chromatography.

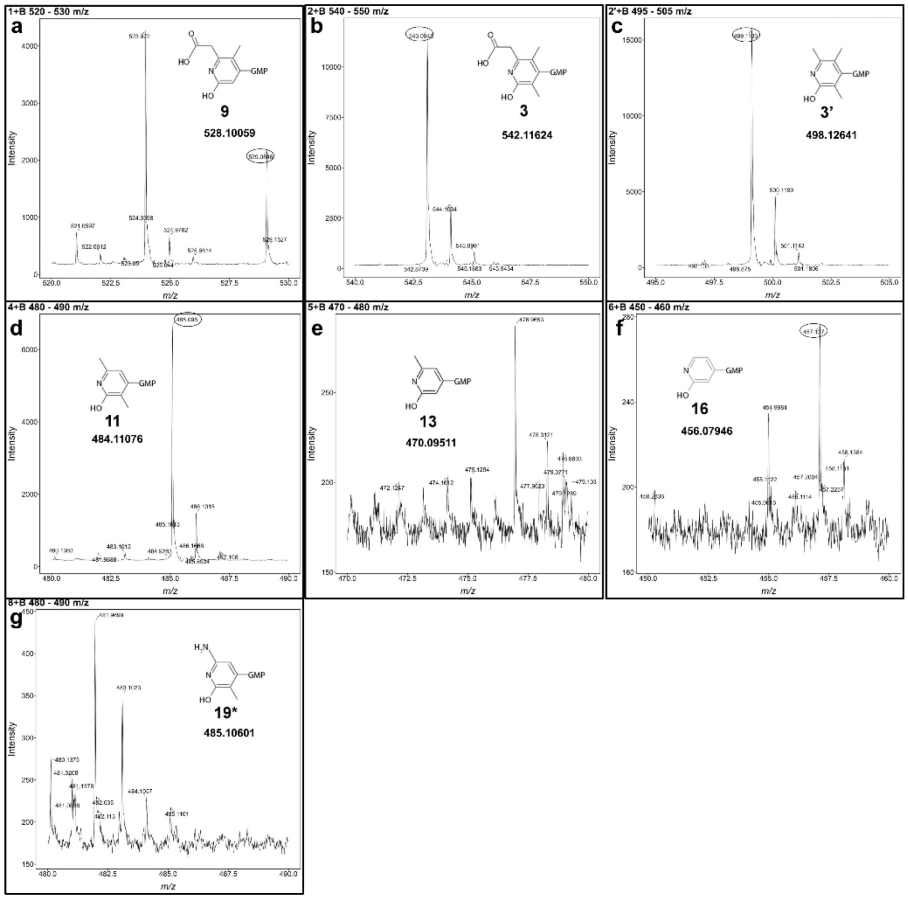

Figure S 2. Reaction of pyridinol compounds **1**, **2**, **2'**, **4**, **5**, **6**, **7** and **8** with HcgB. Masses which correspond to the shown structural formula marked by a black circle. The molecular mass of the compound is given below the structure.

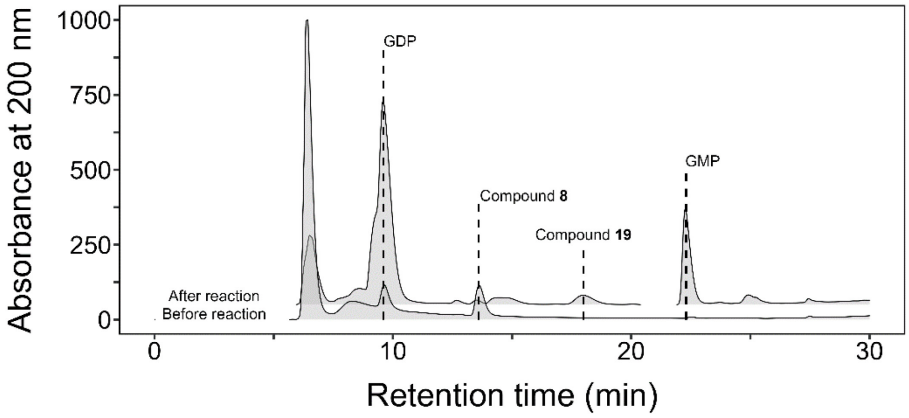

Figure S 3. Conversion of **8** to **19** as catalyzed by HcgB from *M. jannaschii*. The compounds of the main peaks in the HPLC are indicated by dashed lines and the compound is described above. Guanosine-di-phosphate (GDP) and Guanosine-mono-phosphate (GMP) are side products from the decomposition of GTP.

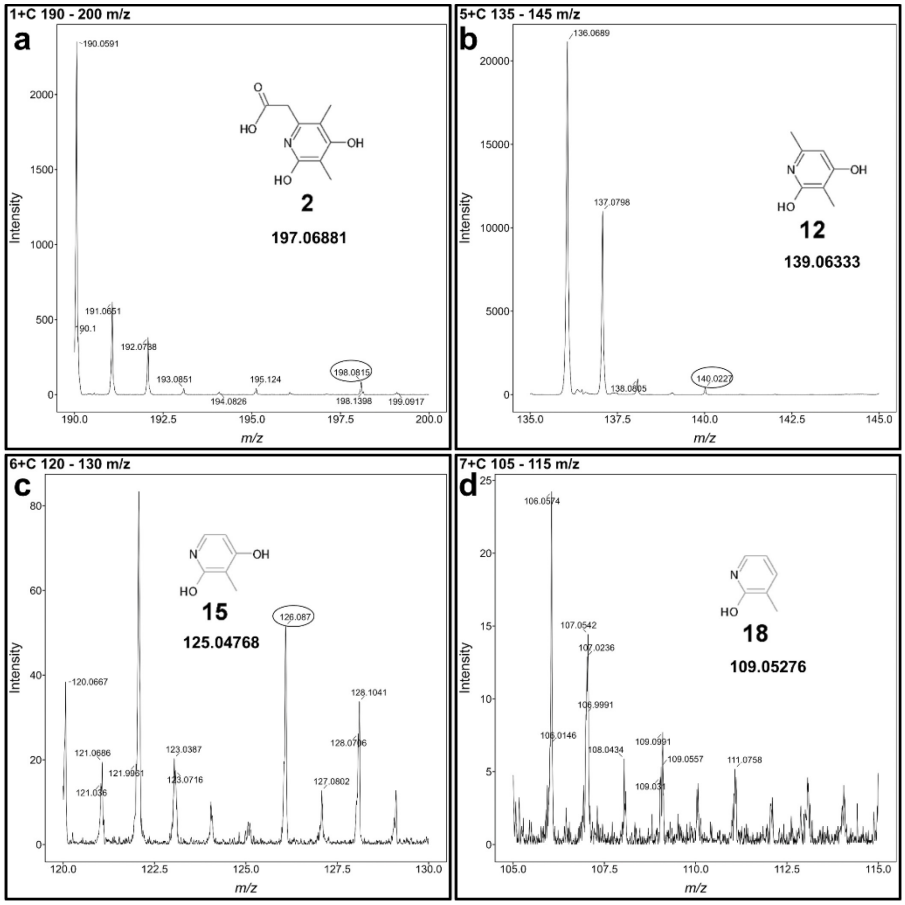

Figure S 4. Reaction of pyridinol compounds **1**, **5**, **6** and **7** with HcgC. Masses which correspond to the shown structural formula marked by a black circle. The molecular mass of the compound is given below the structure.

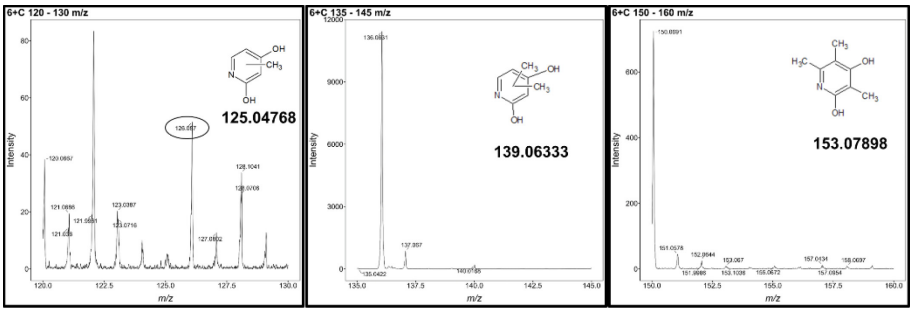

Figure S 5. Reaction of pyridinol compounds **6** with HcgC leading to possible multiple methylations. Masses which correspond to the shown structural formula marked by a black circle. The molecular mass of the compound is given below the structure.

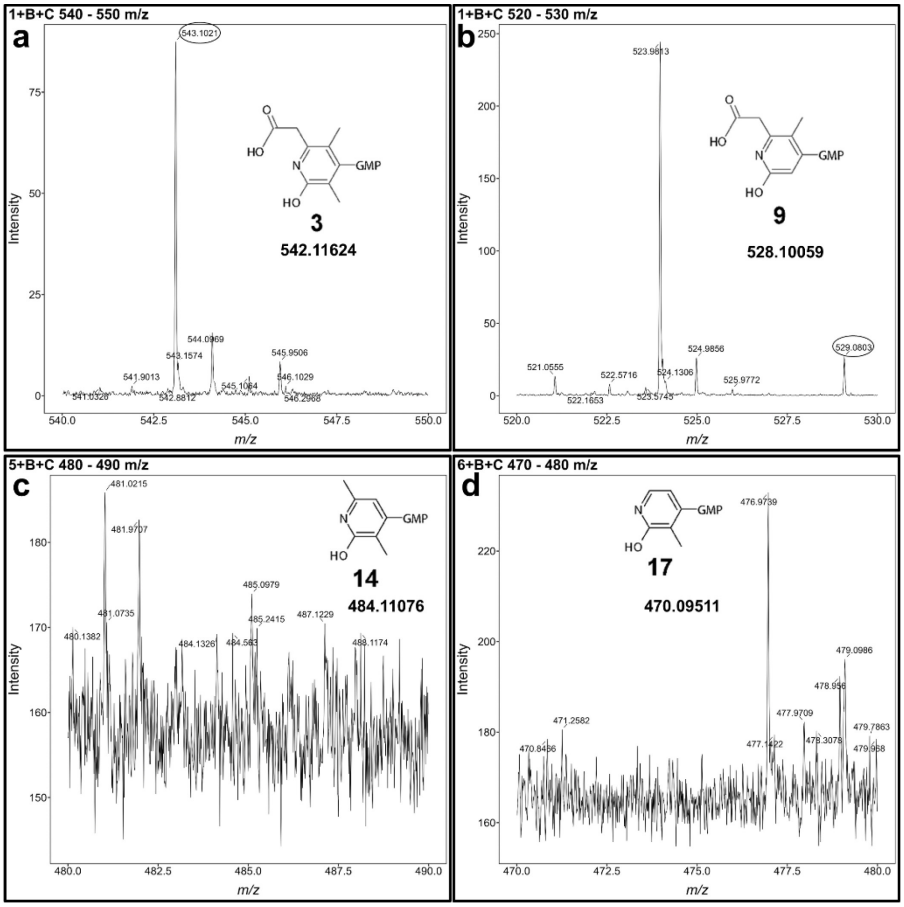

Figure S 6. Reaction of pyridinol compounds **1**, **5**, and **6** with HcgB and HcgC. Masses which correspond to the shown structural formula marked by a black circle. The molecular mass of the compound is given below the structure.

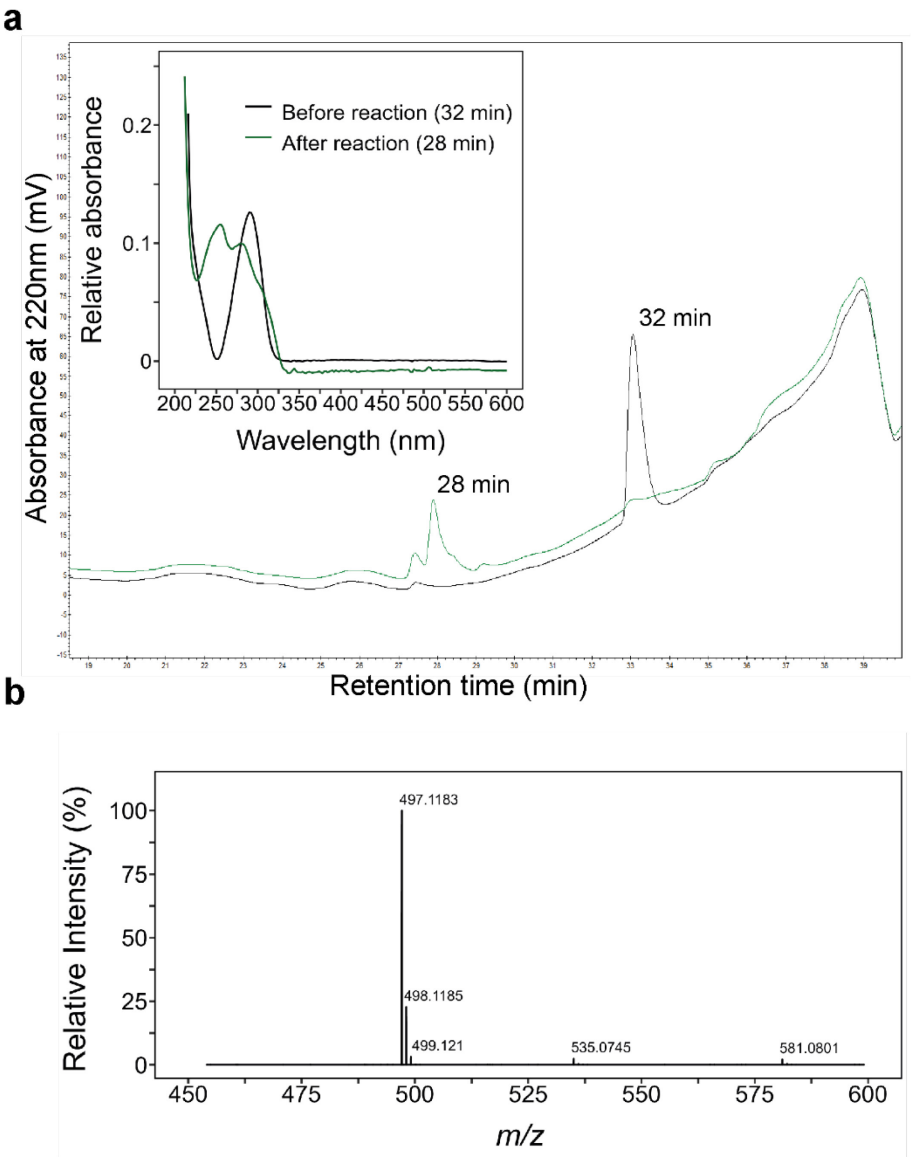

Figure S 7. Synthesis and purification of **3'**. (a) HPLC chromatogram of **2'** before (black) and after (green) reaction with HcgB. The conversion into **3'** is indicated by the shift in the retention time and the changed UV/Vis spectrum (inset). (b) Confirmation of the mass of **3'** ESI-MS (calculated mass for the $[M-H]^-$ = 497.119134 *m/z*, measured 497.1183 *m/z*, 1.678 ppm error).

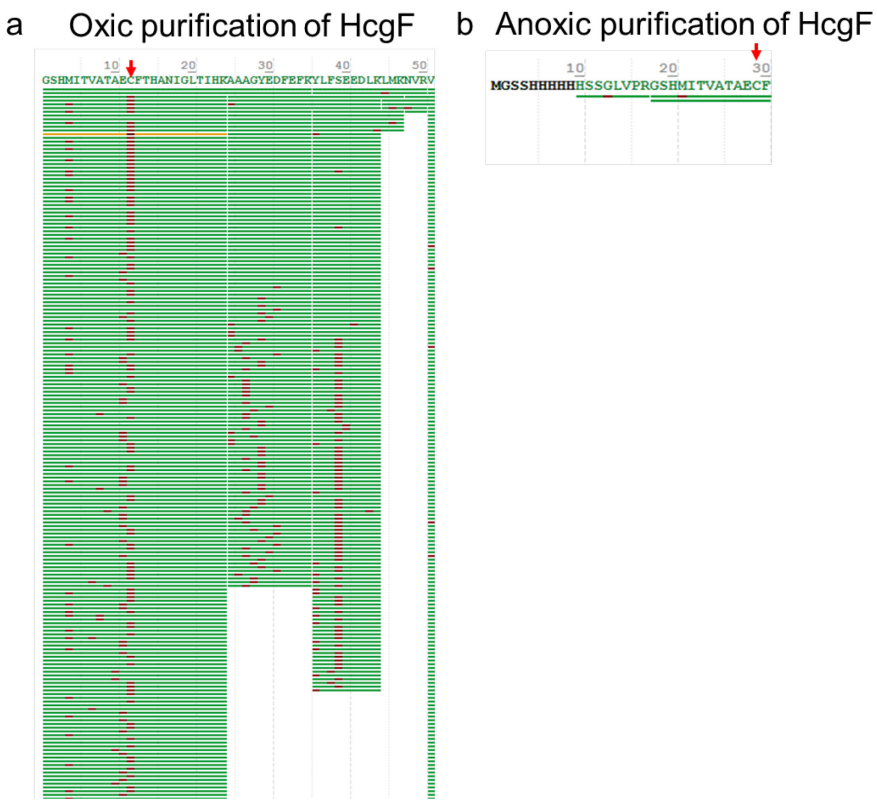

Figure S 8. Analysis of modifcations found on Cys9 of HcgF from M. jannaschii purified under oxic (a) or anoxic (b) conditions. The site of Cys9 is marked with a red arrow. Each green bar represents a individual peptide identified and red marks show modifications at the given aminoacid positions. The peptides and modifications were identified by Byonic by J. Kahnt.

Table S 1. Detection of Hcg proteins in the cell extract of Δ*hcgB*Δ*hcgC M. maripaludis* mutant by proteomics.

Hcg protein (Accession number)	Coverage	Number of peptides (unique peptides)
HcgA (WP_011170070.1)	91.79%	41 (10)
HcgB (WP_011171441.1)	-	-
HcgC (WP_011171442.1)	-	-
HcgD (WP_011169997.1)	86.99%	25 (13)
HcgE (WP_013999640.1)	81.86%	27 (6)
HcgF (WP_011170879.1)	-*	-*
HcgG (WP_011170069.1)	78.04%	48 (18)
Hmd (WP_011170071.1)	84.18%	43 (28)

* HcgF was not detected by method. As described in Schaupp *et al,* 2020 [95] PCR experiments confirmed the presence of HcgF in the Δ*hcgB*Δ*hcgC* and previous experiments showed that deletion of *hcgF* leads to the same growth phenotype as observed in the Δ*hmd* mutant.[40]

6 References

[1] E. Skovran, C. Raghuraman, N. C. Martinez-Gomez, *Curr. Issues Mol. Biol.* **2019**, *33*, 101-116.

[2] a) M. Stein, W. Lubitz, *Phys Chem Chem Phys* **2001**, *3*, 5115-5120; b) D. W. Bak, S. J. Elliott, *Curr. Opin. Chem. Biol.* **2014**, *19*, 50-58.

[3] C. Greening, A. Biswas, C. R. Carere, C. J. Jackson, M. C. Taylor, M. B. Stott, G. M. Cook, S. E. Morales, *ISME J.* **2016**, *10*, 761-777.

[4] a) D. E. Green, L. H. Stickland, *Biochem. J.* **1934**, *28*, 898-900; b) M. Stephenson, L. H. Stickland, *Biochem. J.* **1931**, *25*, 205-214.

[5] W. Lubitz, H. Ogata, O. Rudiger, E. Reijerse, *Chem. Rev.* **2014**, *114*, 4081-4148.

[6] a) A. K. Kaster, J. Moll, K. Parey, R. K. Thauer, *Proc. Natl. Acad. Sci. U S A* **2011**, *108*, 2981-2986; b) G. J. Schut, M. W. Adams, *J. Bacteriol.* **2009**, *191*, 4451-4457.

[7] a) N. Chongdar, J. A. Birrell, K. Pawlak, C. Sommer, E. J. Reijerse, O. Rudiger, W. Lubitz, H. Ogata, *J. Am. Chem. Soc.* **2018**, *140*, 1057-1068; b) A. J. Shaw, D. A. Hogsett, L. R. Lynd, *J. Bacteriol.* **2009**, *191*, 6457-6464; c) O. Lenz, B. Friedrich, *Proc. Natl. Acad. Sci. U S A* **1998**, *95*, 12474-12479.

[8] a) Y. Higuchi, T. Yagi, N. Yasuoka, *Structure* **1997**, *5*, 1671-1680; b) A. Volbeda, M. H. Charon, C. Piras, E. C. Hatchikian, M. Frey, J. C. Fontecilla-Camps, *Nature* **1995**, *373*, 580-587.

[9] J. W. Peters, W. N. Lanzilotta, B. J. Lemon, L. C. Seefeldt, *Science* **1998**, *282*, 1853-1858.

[10] G. F. Huang, T. Wagner, M. D. Wodrich, K. Ataka, E. Bill, U. Ermler, X. L. Hu, S. Shima, *Nat Catal* **2019**, *2*, 537-543.

[11] S. Shima, E. J. Lyon, R. K. Thauer, B. Mienert, E. Bill, *J. Am. Chem. Soc.* **2005**, *127*, 10430-10435.

[12] E. J. Lyon, S. Shima, R. Boecher, R. K. Thauer, F. W. Grevels, E. Bill, W. Roseboom, S. P. Albracht, *J. Am. Chem. Soc.* **2004**, *126*, 14239-14248.

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

[14] a) J. Esselborn, C. Lambertz, A. Adamska-Venkates, T. Simmons, G. Berggren, J. Noth, J. Siebel, A. Hemschemeier, V. Artero, E. Reijerse, M. Fontecave, W. Lubitz, T. Happe, *Nat. Chem. Biol.* **2013**, *9*, 607-609; b) 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, M. Fontecave, *Nature* **2013**, *499*, 66-69.

[15] a) W. Lubitz, E. Reijerse, M. van Gastel, *Chem. Rev.* **2007**, *107*, 4331-4365; b) J. C. Fontecilla-Camps, A. Volbeda, C. Cavazza, Y. Nicolet, *Chem. Rev.* **2007**, *107*, 4273-4303.

[16] H. Ogata, K. Nishikawa, W. Lubitz, *Nature* **2015**, *520*, 571-574.

[17] a) D. Sippel, O. Einsle, *Nat. Chem. Biol.* **2017**, *13*, 956-960; b) J. Kim, D. C. Rees, *Nature* **1992**, *360*, 553-560; c) J. R. Chisnell, R. Premakumar, P. E. Bishop, *J. Bacteriol.* **1988**, *170*, 27-33.

[18] E. J. Lyon, S. Shima, G. Buurman, S. Chowdhuri, A. Batschauer, K. Steinbach, R. K. Thauer, *Eur. J. Biochem.* **2004**, *271*, 195-204.

[19] R. K. Thauer, A. K. Kaster, M. Goenrich, M. Schick, T. Hiromoto, S. Shima, *Annu. Rev. Biochem.* **2010**, *79*, 507-536.

[20] C. Zirngibl, W. Van Dongen, B. Schworer, R. Von Bunau, M. Richter, A. Klein, R. K. Thauer, *Eur. J. Biochem.* **1992**, *208*, 511-520.

[21] a) C. Afting, E. Kremmer, C. Brucker, A. Hochheimer, R. K. Thauer, *Arch Microbiol* **2000**, *174*, 225-232; b) C. Afting, A. Hochheimer, R. K. Thauer, *Arch Microbiol* **1998**, *169*, 206-210.

[22] L. D. Eirich, G. D. Vogels, R. S. Wolfe, *Biochemistry* **1978**, *17*, 4583-4593.

[23] J. A. Fox, D. J. Livingston, W. H. Orme-Johnson, C. T. Walsh, *Biochemistry* **1987**, *26*, 4219-4227.

[24] G. Buurman, S. Shima, R. K. Thauer, *FEBS Lett.* **2000**, *485*, 200-204.

[25] S. Shima, E. J. Lyon, M. Sordel-Klippert, M. Kauss, J. Kahnt, R. K. Thauer, K. Steinbach, X. Xie, L. Verdier, C. Griesinger, *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 2547-2551.
[26] S. Shima, O. Pilak, S. Vogt, M. Schick, M. S. Stagni, W. Meyer-Klaucke, E. Warkentin, R. K. Thauer, U. Ermler, *Science* **2008**, *321*, 572-575.
[27] T. Hiromoto, K. Ataka, O. Pilak, S. Vogt, M. S. Stagni, W. Meyer-Klaucke, E. Warkentin, R. K. Thauer, S. Shima, U. Ermler, *FEBS Lett.* **2009**, *583*, 585-590.
[28] S. Nakanishi, T. Yamamoto, N. Furukawa, Y. Otsuji, *Synthesis* **1994**, *6*, 609-612.
[29] S. Shima, M. Schick, J. Kahnt, K. Ataka, K. Steinbach, U. Linne, *Dalton Trans.* **2012**, *41*, 767-771.
[30] S. Shima, M. Schick, H. Tamura, *Methods Enzymol.* **2011**, *494*, 119-137.
[31] G. F. Huang, F. J. Arriaza-Gallardo, T. Wagner, S. Shima, *Inorganics* **2020**, *8*, 50.
[32] S. Shima, S. Vogt, A. Gobels, E. Bill, *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 9917-9921.
[33] S. Shima, K. Ataka, *FEBS Lett.* **2011**, *585*, 353-356.
[34] H. Tamura, M. Salomone-Stagni, T. Fujishiro, E. Warkentin, W. Meyer-Klaucke, U. Ermler, S. Shima, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 9656-9659.
[35] a) S. Vogt, E. J. Lyon, S. Shima, R. K. Thauer, *J. Biol. Inorg. Chem.* **2008**, *13*, 97-106; b) G. C. Hartmann, E. Santamaria, V. M. Fernández, R. K. Thauer, *J. Biol. Inorg. Chem.* **1996**, *1*, 446-450.
[36] T. Wagner, G. Huang, U. Ermler, S. Shima, *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 15056-15059.
[37] L. Bai, T. Fujishiro, G. Huang, J. Koch, A. Takabayashi, M. Yokono, A. Tanaka, T. Xu, X. Hu, U. Ermler, S. Shima, *Faraday Discuss.* **2017**, *198*, 37-58.
[38] A. Jacobi, R. Rossmann, A. Bock, *Arch. Microbiol.* **1992**, *158*, 444-451.
[39] T. J. Lie, K. C. Costa, B. Lupa, S. Korpole, W. B. Whitman, J. A. Leigh, *Proc. Natl. Acad. Sci. U S A* **2012**, *109*, 15473-15478.
[40] T. J. Lie, K. C. Costa, D. Pak, V. Sakesan, J. A. Leigh, *FEMS Microbiol. Lett.* **2013**, *343*, 156-160.
[41] T. Fujishiro, L. Bai, T. Xu, X. Xie, M. Schick, J. Kahnt, M. Rother, X. Hu, U. Ermler, S. Shima, *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 9648-9651.
[42] M. Schick, X. Xie, K. Ataka, J. Kahnt, U. Linne, S. Shima, *J. Am. Chem. Soc.* **2012**, *134*, 3271-3280.
[43] L. Bai, T. Wagner, T. Xu, X. Hu, U. Ermler, S. Shima, *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 10806-10809.
[44] T. Fujishiro, H. Tamura, M. Schick, J. Kahnt, X. Xie, U. Ermler, S. Shima, *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 12555-12558.
[45] T. Fujishiro, J. Kahnt, U. Ermler, S. Shima, *Nat. Commun.* **2015**, *6*, 6895.
[46] T. Fujishiro, U. Ermler, S. Shima, *FEBS Lett.* **2014**, *588*, 2789-2793.
[47] K. Y. Hwang, J. H. Chung, S. H. Kim, Y. S. Han, Y. Cho, *Nat. Struct. Biol.* **1999**, *6*, 691-696.
[48] B. A. Schulman, J. W. Harper, *Nat. Rev. Mol. Cell. Biol.* **2009**, *10*, 319-331.
[49] L. Galeazzi, P. Bocci, A. Amici, L. Brunetti, S. Ruggieri, M. Romine, S. Reed, A. L. Osterman, D. A. Rodionov, L. Sorci, N. Raffaelli, *J. Biol. Chem.* **2011**, *286*, 40365-40375.
[50] P. Venkitasubramanian, L. Daniels, J. P. Rosazza, *J. Biol. Chem.* **2007**, *282*, 478-485.
[51] P. Venkitasubramanian, L. Daniels, J. P. N. Rosazza, *Journal of Biological Chemistry* 2007, *282*, 478-485.
[52] C. J. Reed, G. Hutinet, V. de Crecy-Lagard, *Biomolecules* **2021**, *11*, 1282.
[53] S. C. Chen, C. H. Huang, C. S. Yang, S. M. Kuan, C. T. Lin, S. H. Chou, Y. Chen, *Biomed. Res. Int.* **2014**, *2014*, 171263.
[54] S. E. McGlynn, E. S. Boyd, E. M. Shepard, R. K. Lange, R. Gerlach, J. B. Broderick, J. W. Peters, *J. Bacteriol.* **2010**, *192*, 595-598.
[55] a) P. A. Frey, A. D. Hegeman, F. J. Ruzicka, *Crit. Rev. Biochem. Mol. Biol.* **2008**, *43*, 63-88; b) N. C. Martinez-Gomez, D. M. Downs, *Biochemistry* **2008**, *47*, 9054-

9056; c) A. Chatterjee, Y. Li, Y. Zhang, T. L. Grove, M. Lee, C. Krebs, S. J. Booker, T. P. Begley, S. E. Ealick, *Nat. Chem. Biol.* **2008**, *4*, 758-765.

[56] a) Y. Zhang, L. Tao, T. J. Woods, R. D. Britt, T. B. Rauchfuss, *J. Am. Chem. Soc.* **2022**, *144*, 1534-1538; b) L. Tao, S. A. Pattenaude, S. Joshi, T. P. Begley, T. B. Rauchfuss, R. D. Britt, *J. Am. Chem. Soc.* **2020**, *142*, 10841-10848.

[57] a) R. D. Britt, G. Rao, L. Tao, *Nat. Rev. Chem.* **2020**, *4*, 542-549; b) J. W. Peters, G. J. Schut, E. S. Boyd, D. W. Mulder, E. M. Shepard, J. B. Broderick, P. W. King, M. W. Adams, *Biochim. Biophys. Acta* **2015**, *6*, 1350-1369.

[58] S. Reissmann, E. Hochleitner, H. Wang, A. Paschos, F. Lottspeich, R. S. Glass, A. Bock, *Science* **2003**, *299*, 1067-1070.

[59] a) M. Blokesch, S. P. Albracht, B. F. Matzanke, N. M. Drapal, A. Jacobi, A. Bock, *J. Mol. Biol.* **2004**, *344*, 155-167; b) S. Watanabe, R. Matsumi, H. Atomi, T. Imanaka, K. Miki, *Structure* **2012**, *20*, 2124-2137.

[60] W. Roseboom, M. Blokesch, A. Bock, S. P. Albracht, *FEBS Lett.* **2005**, *579*, 469-472.

[61] a) A. M. Sydor, J. Liu, D. B. Zamble, *J. Bacteriol.* **2011**, *193*, 1359-1368; b) M. Denic, E. Turlin, V. Michel, F. Fischer, M. Khorasani-Motlagh, D. Zamble, D. Vinella, H. de Reuse, *PLoS Pathog.* **2021**, *17*, e1009193; c) J. W. Zhang, G. Butland, J. F. Greenblatt, A. Emili, D. B. Zamble, *J. Biol. Chem.* **2005**, *280*, 4360-4366; d) D. Sasaki, S. Watanabe, R. Matsumi, T. Shoji, A. Yasukochi, K. Tagashira, W. Fukuda, T. Kanai, H. Atomi, T. Imanaka, K. Miki, *J. Mol. Biol.* **2013**, *425*, 1627-1640; e) M. J. Lacasse, C. D. Douglas, D. B. Zamble, *Biochemistry* **2016**, *55*, 6821-6831; f) S. Watanabe, T. Kawashima, Y. Nishitani, T. Kanai, T. Wada, K. Inaba, H. Atomi, T. Imanaka, K. Miki, *Proc. Natl. Acad. Sci. U S A* **2015**, *112*, 7701-7706; g) T. Maier, F. Lottspeich, A. Bock, *Eur. J. Biochem.* **1995**, *230*, 133-138.

[62] E. Theodoratou, R. Huber, A. Bock, *Biochem. Soc. Trans.* **2005**, *33*, 108-111.

[63] R. P. Happe, W. Roseboom, A. J. Pierik, S. P. Albracht, K. A. Bagley, *Nature* **1997**, *385*, 126.

[64] A. Paschos, R. S. Glass, A. Bock, *FEBS Lett.* **2001**, *488*, 9-12.

[65] A. Paschos, A. Bauer, A. Zimmermann, E. Zehelein, A. Bock, *J. Biol. Chem.* **2002**, *277*, 49945-49951.

[66] J. C. Rain, L. Selig, H. De Reuse, V. Battaglia, C. Reverdy, S. Simon, G. Lenzen, F. Petel, J. Wojcik, V. Schachter, Y. Chemama, A. Labigne, P. Legrain, *Nature* **2001**, *409*, 211-215.

[67] O. Lenz, I. Zebger, J. Hamann, P. Hildebrandt, B. Friedrich, *FEBS Lett.* **2007**, *581*, 3322-3326.

[68] L. Rey, D. Fernandez, B. Brito, Y. Hernando, J. M. Palacios, J. Imperial, T. Ruiz-Argueso, *Mol. Gen. Genet.* **1996**, *252*, 237-248.

[69] I. Burstel, E. Siebert, S. Frielingsdorf, I. Zebger, B. Friedrich, O. Lenz, *Proc. Natl. Acad. Sci. U S A* **2016**, *113*, 14722-14726.

[70] N. Muraki, K. Ishii, S. Uchiyama, S. G. Itoh, H. Okumura, S. Aono, *Commun. Biol.* **2019**, *2*, 385.

[71] A. C. Schulz, S. Frielingsdorf, P. Pommerening, L. Lauterbach, G. Bistoni, F. Neese, M. Oestreich, O. Lenz, *J. Am. Chem. Soc.* **2020**, *142*, 1457-1464.

[72] P. W. King, M. C. Posewitz, M. L. Ghirardi, M. Seibert, *J. Bacteriol.* **2006**, *188*, 2163-2172.

[73] A. Pagnier, B. Balci, E. M. Shepard, H. Yang, D. M. Warui, S. Impano, S. J. Booker, B. M. Hoffman, W. E. Broderick, J. B. Broderick, *Angew. Chem. Int. Ed. Engl.* **2022**, *61*, e202203413.

[74] a) P. Dinis, D. L. Suess, S. J. Fox, J. E. Harmer, R. C. Driesener, L. De La Paz, J. R. Swartz, J. W. Essex, R. D. Britt, P. L. Roach, *Proc. Natl. Acad. Sci. U S A* **2015**, *112*, 1362-1367; b) J. M. Kuchenreuther, W. K. Myers, D. L. Suess, T. A. Stich, V. Pelmenschikov, S. A. Shiigi, S. P. Cramer, J. R. Swartz, R. D. Britt, S. J. George, *Science* **2014**, *343*, 424-427; c) J. M. Kuchenreuther, W. K. Myers, T. A. Stich, S. J. George, Y. Nejatyjahromy, J. R. Swartz, R. D. Britt, *Science* **2013**, *342*, 472-475;

d) 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, J. B. Broderick, *J. Am. Chem. Soc.* **2010**, *132*, 9247-9249; e) R. C. Driesener, M. R. Challand, S. E. McGlynn, E. M. Shepard, E. S. Boyd, J. B. Broderick, J. W. Peters, P. L. Roach, *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 1687-1690.

[75] R. Rohac, L. Martin, L. Liu, D. Basu, L. Tao, R. D. Britt, T. B. Rauchfuss, Y. Nicolet, *J. Am. Chem. Soc.* **2021**, *143*, 8499-8508.

[76] a) J. M. Kuchenreuther, J. A. Stapleton, J. R. Swartz, *PloS one* **2009**, *4*, e7565; b) M. E. Boyer, J. A. Stapleton, J. M. Kuchenreuther, C. W. Wang, J. R. Swartz, *Biotechnol. Bioeng.* **2008**, *99*, 59-67.

[77] M. Kriek, F. Martins, M. R. Challand, A. Croft, P. L. Roach, *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 9223-9226.

[78] E. Pilet, Y. Nicolet, C. Mathevon, T. Douki, J. C. Fontecilla-Camps, M. Fontecave, *FEBS Lett.* **2009**, *583*, 506-511.

[79] S. E. McGlynn, S. S. Ruebush, A. Naumov, L. E. Nagy, A. Dubini, P. W. King, J. B. Broderick, M. C. Posewitz, J. W. Peters, *J. Biol. Inorg. Chem.* **2007**, *12*, 443-447.

[80] J. M. Kuchenreuther, S. J. George, C. S. Grady-Smith, S. P. Cramer, J. R. Swartz, *PloS one* **2011**, *6*, e20346.

[81] D. L. Suess, I. Burstel, L. De La Paz, J. M. Kuchenreuther, C. C. Pham, S. P. Cramer, J. R. Swartz, R. D. Britt, *Proc. Natl. Acad. Sci. U S A* **2015**, *112*, 11455-11460.

[82] G. Rao, S. A. Pattenaude, K. Alwan, N. J. Blackburn, R. D. Britt, T. B. Rauchfuss, *Proc. Natl. Acad. Sci. U S A* **2019**, *116*, 20850-20855.

[83] R. Rohac, P. Amara, A. Benjdia, L. Martin, P. Ruffie, A. Favier, O. Berteau, J. M. Mouesca, J. C. Fontecilla-Camps, Y. Nicolet, *Nat. Chem.* **2016**, *8*, 491-500.

[84] a) H. H. Nagatani, V. K. Shah, W. J. Brill, *J. Bacteriol.* **1974**, *120*, 697-701; b) T. MacNeil, D. MacNeil, G. P. Roberts, M. A. Supiano, W. J. Brill, *J. Bacteriol.* **1978**, *136*, 253-266.

[85] a) J. Wilcoxen, S. Arragain, A. A. Scandurra, E. Jimenez-Vicente, C. Echavarri-Erasun, S. Pollmann, R. D. Britt, L. M. Rubio, *J. Am. Chem. Soc.* **2016**, *138*, 7468-7471; b) J. A. Wiig, Y. Hu, M. W. Ribbe, *Proc. Natl. Acad. Sci. U S A* **2011**, *108*, 8623-8627; c) A. W. Fay, M. A. Blank, C. C. Lee, Y. Hu, K. O. Hodgson, B. Hedman, M. W. Ribbe, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 7787-7790; d) K. Tanifuji, C. C. Lee, N. S. Sickerman, K. Tatsumi, Y. Ohki, Y. Hu, M. W. Ribbe, *Nat. Chem.* **2018**, *10*, 568-572.

[86] J. A. Wiig, Y. Hu, C. Chung Lee, M. W. Ribbe, *Science* **2012**, *337*, 1672-1675.

[87] J. T. Kaiser, Y. Hu, J. A. Wiig, D. C. Rees, M. W. Ribbe, *Science* **2011**, *331*, 91-94.

[88] V. K. Shah, J. Imperial, R. A. Ugalde, P. W. Ludden, W. J. Brill, *Proc. Natl. Acad. Sci. U S A* **1986**, *83*, 1636-1640.

[89] L. Curatti, J. A. Hernandez, R. Y. Igarashi, B. Soboh, D. Zhao, L. M. Rubio, *Proc. Natl. Acad. Sci. U S A* **2007**, *104*, 17626-17631.

[90] S. Buren, E. Jimenez-Vicente, C. Echavarri-Erasun, L. M. Rubio, *Chem. Rev.* **2020**, *120*, 4921-4968.

[91] T. R. Hoover, A. D. Robertson, R. L. Cerny, R. N. Hayes, J. Imperial, V. K. Shah, P. W. Ludden, *Nature* **1987**, *329*, 855-857.

[92] J. A. Hernandez, L. Curatti, C. P. Aznar, Z. Perova, R. D. Britt, L. M. Rubio, *Proc. Natl. Acad. Sci. U S A* **2008**, *105*, 11679-11684.

[93] J. T. Moore, A. Uppal, F. Maley, G. F. Maley, *Protein Expr. Purif.* **1993**, *4*, 160-163.

[94] a) W. E. Balch, G. E. Fox, L. J. Magrum, C. R. Woese, R. S. Wolfe, *Microbiol. Rev.* **1979**, *43*, 260-296; b) F. Long, L. Wang, B. Lupa, W. B. Whitman, *Archaea* **2017**, *2017*, 7046026.

[95] S. Schaupp, F. J. Arriaza-Gallardo, H. J. Pan, J. Kahnt, G. Angelidou, N. Paczia, K. Costa, X. Hu, S. Shima, *Angew. Chem. Int. Ed. Engl.* **2022**, *61*, e202200994.

[96] P. Millard, B. Delépine, M. Guionnet, M. Heuillet, F. Bellvert, F. Létisse, *Bioinformatics* **2019**, *35*, 4484-4487.
[97] J. Meija, T. B. Coplen, M. Berglund, W. A. Brand, P. De Bievre, M. Groning, N. E. Holden, J. Irrgeher, R. D. Loss, T. Walczyk, T. Prohaska, *Pure Appl Chem* **2016**, *88*, 293-306.
[98] a) L. Forzi, P. Hellwig, R. K. Thauer, R. G. Sawers, *FEBS Lett.* **2007**, *581*, 3317-3321; b) R. D. Britt, G. Rao, L. Tao, *Chem. Sci.* **2020**, *11*, 10313-10323.
[99] R. C. Murphy, K. L. Clay, *Biomed. Mass Spectrom.* **1979**, *6*, 309-314.
[100] M. A. Wright, J. A. Wright, *Dalton Trans.* **2016**, *45*, 6801-6811.
[101] T. Watanabe, T. Wagner, G. Huang, J. Kahnt, K. Ataka, U. Ermler, S. Shima, *Angew. Chem. Int. Ed. Engl.* **2019**, *58*, 3506-3510.
[102] J. W. Simek, T. Tuck, K. C. Bush, *J Chem Educ* **1997**, *74*, 107-108.
[103] K. A. Cornell, M. K. Riscoe, *Biochim. Biophys. Acta* **1998**, *1396*, 8-14.
[104] a) M. Bott, R. K. Thauer, *Eur. J. Biochem.* **1987**, *168*, 407-412; b) D. O. Mountfort, E. Morschel, D. B. Beimborn, P. Schonheit, *J. Bacteriol.* **1986**, *168*, 892-900.
[105] R. Brückner, in *Reaktionsmechanismen*, 3 ed., Springer Spektrum Berlin, Heidelberg, **2004**, p. 270.
[106] D. I. Kim, K. J. Roux, *Trends Cell. Biol.* **2016**, *26*, 804-817.
[107] F. J. Arriaza-Gallardo, S. Schaupp, M. F. A. Halim, H. Pan, G. Angelidou, N. Pazcia, X. Hu, K. Costa, S. Shima, *in prepration* **2022**.

Acknowledgements

The work presented in this would not have been possible without the support from many people in my scientific and private life. I would like to take the time here to thank all of these amazing persons for their support during these last years and before.

First I would express my deepest appreciation to Seigo Shima for guiding me during my PhD project and showing me the fascinating world of the [Fe]-hydrogenase and science in general. Our many interesting and productive discussions furthered my scientific understanding and expanded my knowledge greatly.

I'm extremely grateful for Prof. Heider's support as my second reviewer and for the interesting discussions in my thesis advisory committee.

For their participation in my thesis committee I would like to extend my sincere thanks to Prof. Essen und Prof. Randau.

I would like to express my deepest gratitude for the support of Prof. Thauer in my thesis advisory committee as well as for the interesting input and fruitful consultations concerning my project.

For her help in measuring and analyzing the MS data I would like to extend my sincere thanks to Nicole Paczia and Georgia Angelidou. Without the chemically synthesized pyridinols this work would not have been possible, for this I would like to thank Hui-jie Pan and Xile Hu from the EPFL. Similarly, I would like to thank Kyle Costa for providing the *M. maripaludis* mutants. I am also grateful for the support of Jörg Kahnt concerning the MALDI-MS and proteomics analysis.

Many thanks also go to all lab members of the AG Shima, present and past, with whom I had the pleasure to work with and without whom the work would have been a lot less fun. I would like to thank Jürgen, Gangfeng and Tomohiro for training me how to work with anaerobic enzymes and all the associated techniques, which were required for this project. In particular my thanks go to Francisco and Manuel, who share this exciting journey with me and without whom the last years would have been a lot less fun. Thank you for the support, aid and many caffeine fueled brainstorming sessions. I would also like to thank the AG Brune for their great neighborhood and cooperation in keeping both labs running efficiently, as well as the shared coffee breaks to unwind and relax our minds.

Special thanks goes to Jose, without her support and her keeping my back free this would have not been possible. Finally, I would like to thank my parents who set me on my path to become who I am now and supported me all the way.

Curriculum vitae

This page contains personal information and is not part of the online publication.

Erklärung

Ich versichere, dass ich meine Dissertation mit dem Titel „*In vitro* biosynthesis of the FeGP cofactor of the [Fe]-hydrogenase“ selbstständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe.

Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 24.06.2022

Sebastian Jakob Schaupp