Identification and characterization of two radical SAM enzymes involved in biosynthesis of the [Fe]-hydrogenase cofactor

Dissertation

zur Erlangung des Grades eines

Doktor der Naturwissenschaften

(Dr. rer. nat.)

des Fachbereichs Biologie der Philipps-Universität Marburg

Vorgelegt von

Francisco Javier Arriaza Gallardo

Aus Copiapó, Chile

Marburg/Lahn, 2022

Die vorliegende Dissertation wurde von Oktober 2018 bis August 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: ____25.01.2023_____

“Persevere then until thou shalt have made these things thy own, as the stomach which is strengthened makes all things its own, as the blazing fire makes flame and brightness out of everything that is thrown into it.”

- Marcus Aurelius

Table of Contents

1. Abstract

Most methanogenic archaea reduce CO_2 to methane using molecular hydrogen (H_2) as the electron donor. Specific metalloenzymes called hydrogenases catalyze oxidation of H_2 or hydride transfer from H_2 to provide electrons to the metabolism. Under nickel-limiting conditions, production of [Fe]-hydrogenase (Hmd) is up-regulated, and it functions as one of the major hydrogenases in hydrogenotrophic methanogens. Hmd catalyzes the reversible reduction of methenyl-tetrahydromethanopterin (methenyl-H_4MPT^+) to methylene-H_4MPT by transferring a hydride from H_2. This enzyme possesses a unique metallocofactor, the iron-guanylylpyridinol (FeGP) cofactor, in its active site. The FeGP cofactor contains an Fe-center, a pyridinol ring, and a GMP moiety. The iron center is coordinated with two CO ligands, one cysteine-sulfur, the nitrogen of the pyridinol ring, and an acyl-carbon of one of the substituents of the pyridinol ring. Biosynthesis of this cofactor requires the product of at least seven genes, which are named *hmd* co-occurring genes (*hcgA-G*). HcgC is a SAM-dependent methyltransferase that methylates the 3-position of 6-carboxymethyl-5-methyl-4-hydroxy-2-pyridinol (precursor **1**), to make precursor **2**. HcgB is a GTP-dependent enzyme that guanylylates precursor **2** to produce a guanylylpyridinol (GP or precursor **3**). HcgE is an ATP-dependent enzyme that adenylylates **3** and activates the carboxyl group. The adenylylated product is predicted to bind to HcgF forming a thioester bond with a cysteine residue. HcgD is a possible iron-trafficking protein. In this work, I elucidated the function of the remaining Hcg proteins (HcgA and HcgG). I expressed the *hcgA* gene in *Escherichia coli* and purified an active form of HcgA. By using a newly developed *in vitro* biosynthesis assay (Schaupp and Arriaza et al. 2022. Angew. Chem. Int. Ed. 61, e202200994), I determined HcgA as a radical SAM enzyme that produces precursor **1** from an unknown compound. Since HcgG tends to form inclusion bodies in the *Escherichia coli* cells, I expressed the *hcgG* gene homologously in *Methanococcus maripaludis*, and purified an active form of HcgG. *In vitro* biosynthesis experiments suggested that the HcgG is a radical SAM enzyme that catalyzes multiple catalytic reactions: (1) production of CO from an unknown cellular component, (2) biosynthesis of the CO ligands, (3) formation of the acyl-ligand, and (4) assembly of the FeGP cofactor including incorporation of Fe^{2+}. In addition, I investigated the reconstitution of the holoenzyme *in vitro* from the apoenzyme and the isolated FeGP cofactor. Based on the kinetic and structural analyses, I proposed a mechanism of the FeGP cofactor binding to the protein. Furthermore, I studied the function of the Hmd paralog (HmdII) by mutation analyses. These experiments suggested that HmdII is overproduced in the *M. maripaludis* strain (Δ*frh*) lacking F_{420}-reducing [NiFe]-hydrogenase activity and that HmdII negatively regulates the production of the FeGP cofactor.

2. Zusammenfassung

Die meisten methanogenen Archaeen reduzieren CO_2 mit Hilfe von Wasserstoff (H_2) zu Methan. Für diese Aufgabe verwenden sie spezielle Metalloenzyme, die Hydrogenasen. Die Expression der [Fe]-Hydrogenase (Hmd) wird dabei unter Nickel-limitierten Wachstumsbedingungen erhöht. Hmd katalysiert die Reduktion von Methenyl-H_4MPT$^+$ zu Methylene-H_4MPT mittels Transfer eines Hydrids von H_2. Im aktiven Zentrum des Enzyms befindet sich ein einzigartiger Kofaktor, der FeGP Kofaktor, der aus einem Eisenzentrum, einem Pyridinol und einer GMP-Gruppe aufgebaut ist. Das Eisen wird von zwei Kohlenmonoxidliganden, Schwefel eines Cysteines, dem Stickstoff und der Acylgruppe des Pyridinols koordiniert. Für die Biosynthese dieses Kofaktors sind mindestens sieben Gene erforderlich, die *hmd co-occurring genes* (*hcgA-G)* genannt werden. HcgC ist eine SAM-abhängige Methyltransferase, die die dritte Position des Vorläufermoleküls **1** (6-Carboxymethyl-4-hydroxyl-2-pyridinol) methyliert und das Vorläufermolekül **2** bildet. Dieser wird von HcgB guanylyliert und daraus entsteht Verbindung **3** oder auch genannt GP. Das ATP-abhängige Enzym HcgE aktiviert die Carbonsäuregruppe mittels Adenylylierung. Zusätzlich wurde eine anschließende Bildung eines Thioester mit einem Cystein von HcgF angenommen. Die Funktion von HcgD konnte noch nicht abschließend geklärt werden, aber die Kristallstruktur weist auf eine Funktion im Eisentransport hin. In dieser Arbeit konnte ich die Funktion der beiden anderen Proteine HcgA und HcgG aufklären. Ich habe das *hcgA* Gen in *Escherichia coli* produziert und das aktive Protein aufgereinigt. Mit Hilfe eines neu entwickelten *in vitro* Biosyntheseassay (Schaupp und Arriaza et al. 2022. Angew. Chem. Int. Ed. 61, e202200994) habe ich gezeigt, dass HcgA ein radikalbildendes SAM-abhängiges Enzym ist, welches das Vorläufermolekül **1** bildet. Da HcgG bei der Produktion in *E. coli* Einschlusskörper bildet habe ich das Protein in *Methanococcus maripaludis* produziert und konnte so aktives Protein aufreinigen. *In vitro* Biosyntheseexperimente zeigen, dass HcgG ebenfalls ein radikalbildendes SAM-abhängiges Enzym ist. Diese Experimente weisen darauf hin, dass HcgG erstens CO aus einer unbekannten, zellulären Substanz synthetisiert, zweitens während der Biosynthese die CO-Liganden aus gasförmigen CO formen kann, drittens die Acylgruppe bildet und viertens den FeGP Kofaktor assembliert inklusive dem Einbau des Fe^{2+}-Ions. Neben der Untersuchung der Biosynthese habe ich *in vitro* die Aktivierung des Holoenyzmes der [Fe]-Hydrogenase kinetisch untersucht und einen Mechanismus vorgeschlagen. Weiterhin habe ich die Funktion des Hmd paralogs (HmdII) mittels Mutationsanalyse untersucht und konnte zeigen, dass HmdII die [Fe]-Hydrogenaseaktivität negativ reguliert ohne die Produktion des Hmd-Proteins zu beeinflussen.

3. Publications

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

Arriaza-Gallardo, F. J.*, Schaupp, S.*, Zheng, Y.-C., Abdul-Halim, M. F., Pan, H.-j., Kahnt, J., Angelidou, G., Paczia, N., Hu, X. Costa, K., Shima, S. The function of two radical-SAM enzymes, HcgA and HcgG, in biosynthesis of the [Fe]-hydrogenase cofactor. Submitted.

Schaupp, S.*, **Arriaza-Gallardo, F. J.***, Pan, H., Kahnt, J., Angelidou, G., Paczia, N., Costa, K., Hu, X., Shima, S. (2022). In Vitro Biosynthesis of the [Fe]-Hydrogenase Cofactor Verifies the Proposed Biosynthetic Precursors. *Angew. Chem. Int. Ed.*, *61*, e202200994.

Huang, G., **Arriaza-Gallardo, F. J.**, Wagner, T., Shima, S. (2020). Crystal structures of [Fe]-hydrogenase from *Methanolacinia paynteri* suggest a path of the FeGP-cofactor incorporation process. *Inorganics*, *8*, 50.

*These authors contributed equally to the works.

4. Introduction

4.1. Methanogenesis

Methanogens are a group of microorganisms that produce methane as the catabolic product.[1] In many anaerobic habitats, such as ruminants, wetlands, paddy soils and landfills, methanogens produce in total approximately 1 billion tons (1 Giga ton, Gt) methane per year.[2] Methanogens produce methane from carbon dioxide (CO_2) and molecular hydrogen (H_2), methanol, methylthiols, methylamines and/or acetate.[2,3] Recently, methanogens that utilize methoxylated compounds were discovered.[4] Methoxylated compounds are found in large amounts in coal and lignin.[4] Most of the methane produced is consumed by anaerobic and aerobic methanotrophs, however, around 0.4 Gt are released to the atmosphere, where the majority is decomposed by photochemical reactions. Nevertheless, a huge amount of methane is accumulated in the atmosphere and contributes to the problem of global warming.[2,5] These methane-producing organisms belong to the domain of Archaea and they are classified into two groups: methanogens with and without cytochromes.[2] Most of the methanogens utilizing the aforementioned substrates, other than H_2 and CO_2, are found in the methanogens with cytochromes. Hydrogenotrophic methanogens, which belong to the methanogens without cytochromes, can grow on H_2 and CO_2, and many can also grow on formate. Some exceptions have been described, with one being *Methanosphaera stadtmanae*, which can only grow on methanol and H_2.[6]

Most methanogens can reduce CO_2 and fix the C_1 unit as a formyl group, and then metabolize it to methenyl, methylene and methyl groups. In these reduction steps, H_2 is used as a source of electrons.[2,3,7] The first reaction of this pathway is catalyzed by the formyl-MFR dehydrogenase, an iron-sulfur enzyme that contains tungsten (Fwd), or molybdenum (Fmd).[7,8] In the first active site of this enzyme complex, CO_2 is reduced with two high-energy electrons to produce formate, where the high-energy electrons are produced by a later step of methanogenesis. Formate produced by the first active site of Fmd/Fwd is transferred to the second active site of this enzyme to bind it to the C_1-carrier methanofuran (MFR) as a formyl group.[9] The formyl group of formyl-MFR is then transferred to the next C_1-carrier, tetrahydromethanopterin (H_4MPT) by a reaction catalyzed by the formylmethanofuran:H_4MPT formyltransferase (Ftr) to produce N^5-formyl-H_4MPT.[10] The next step is the formation of N^5,N^{10}-methenyl-H_4MPT^+ from formyl-H_4MPT by the N^5,N^{10}-methenyl-H_4MPT^+ cyclohydrolase (Mch).[7,11] N^5,N^{10}-methenyl-H_4MPT^+ is subsequently reduced to N^5,N^{10}-methylene-H_4MPT by two types of N^5,N^{10}-methylene-H_4MPT dehydrogenase:coenzyme F_{420}-dependent methylene-H_4MPT dehydrogenase (Mtd)[12] and, under nickel-limitation conditions, H_2-forming methylene-H_4MPT dehydrogenase (Hmd).[7,13–15] Hmd is also called [Fe]-hydrogenase. N^5,N^{10}-

methylene-H_4MPT is reduced to N^5-methyl-H_4MPT by the F_{420}-dependent N^5-N^{10}-methylene-H_4MPT reductase (Mer).[7,16] The coenzyme F_{420}, or simply F_{420}, acts as a hydride carrier. F_{420} is reduced to $F_{420}H_2$ using H_2 as the electron donor by the F_{420}-reducing hydrogenase (Frh), or when growing under formate, by the F_{420}-reducing formate dehydrogenase (Fdh).[7,17] Next, the sodium-translocating membrane-associated N^5-methyl-H_4MPT:coenzyme M methyltransferase (Mtr) catalyzes the methyl group transfer to coenzyme M (CoM-SH), forming methyl-S-CoM, which is coupled with Na^+ translocation.[2,7,18] The electrochemical gradient generated is used as the driving force for ATP synthesis.[2,19] Methane production is catalyzed by methyl-coenzyme M reductase (Mcr), which produces methane and a heterodisulfide of CoM-SH and coenzyme B (CoB-SH).[20] Finally, the heterodisulfide is reduced with H_2 to reproduce CoM-SH and CoB-SH, which is coupled to the reduction of ferredoxin via flavin based electron bifurcation (FBEB). These reactions are catalyzed by an enzyme complex of Hdr-associating hydrogenase (Mvh) and heterodisulfide reductase (Hdr) (Mvh-Hdr).[21–23] Costa *et al*. showed that Mvh-Hdr forms a supercomplex with Fwd (Mvh-Hdr-Fwd) in *Methanococcus maripaludis*,[24] where the authors also indicated the presence of an Fdh-Hdr-Fwd supercomplex complex. Recently, a Fdh-Hrd-Fmd complex was isolated from *Methanospirillum hungatei* and enzymologically and structurally characterized.[25] These results suggest physical and electronic connections between the first and the last step of methanogenesis.

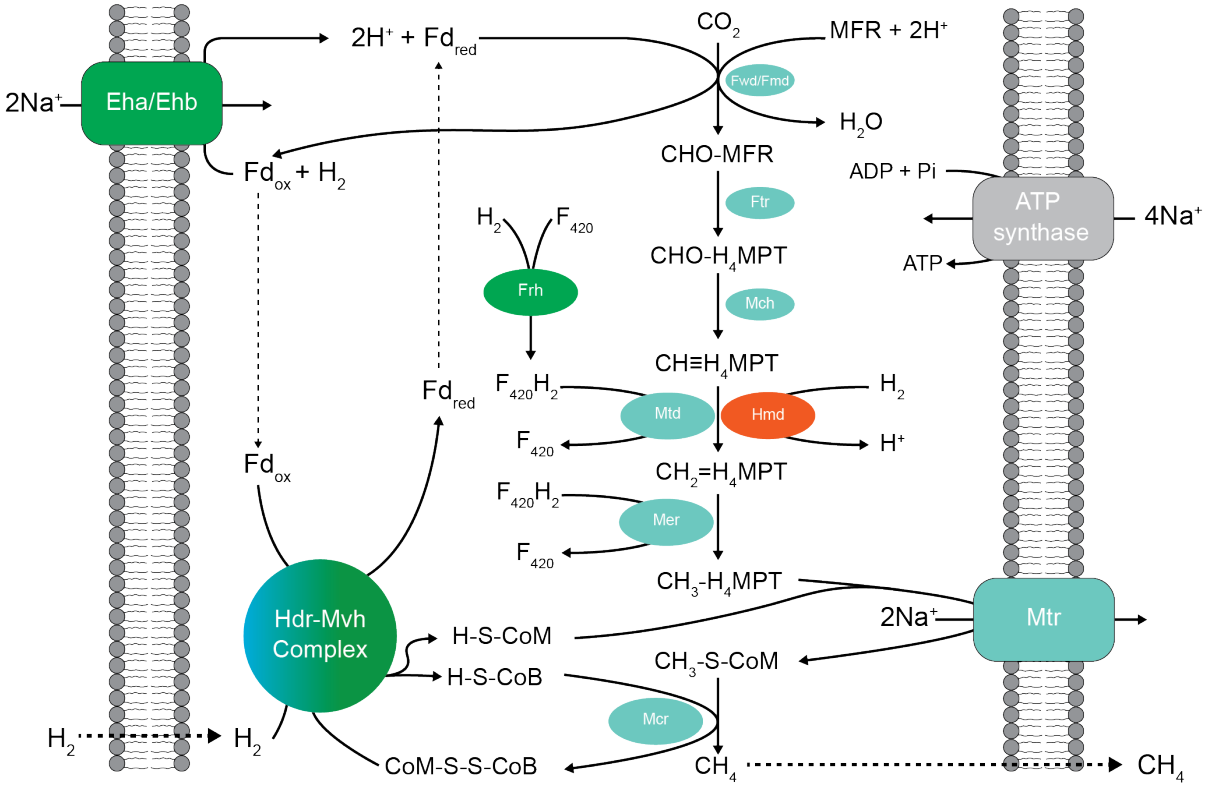


Figure 1. Hydrogenotrophic methanogenic pathway. Fwd: Tungsten-dependent formyl-MFR dehydrogenase; Fmd: molybdenum-dependent formyl-MFR dehydrogenase; Ftr: formyl-MFR:H_4MPT formyltransferase; Mch: methenyl-H_4MPT$^+$ cyclohydrolase; Frh: F_{420}-reducing [NiFe]-hydrogenase; Mtd: F_{420}-dependent methylene-H_4MPT dehydrogenase; Hmd: , H_2-forming methylene-H_4MPT dehydrogenase; Mer: F_{420}-dependent methylene-H_4MPT reductase; Mcr: methyl-coenzyme M reductase; Hdr-Mvh Complex: heterodisulfide reductase–[NiFe]-hydrogenase complex; Eha/Ehb: energy-converting membrane-associating [NiFe]-hydrogenases; Mtr: membrane-associating methyl-H_4MPT:coenzyme M methyltransferase; MFR: methanofuran; H_4MPT: tetrahydromethanopterin; CoB-SH: Coenzyme B; CoM-SH: Coenzyme M; CoM-S-S-CoB: heterodisulfide; Fd, ferredoxin; F_{420}: coenzyme F_{420}. [NiFe]-hydrogenases are highlighted in green. [Fe]-hydrogenase in orange.

4.2. Hydrogenases

Hydrogenases are a group of metalloenzymes that catalyze the reversible conversion of H_2 into electrons and protons, or the heterolytic cleavage of H_2 into a hydride and a proton.[3,7,26] Hydrogenases are classified into three phylogenetically distinct classes: [NiFe]-, [FeFe]-, and [Fe]-hydrogenases, characterized by a distinct metal center (Figure 2).[7,26,27]

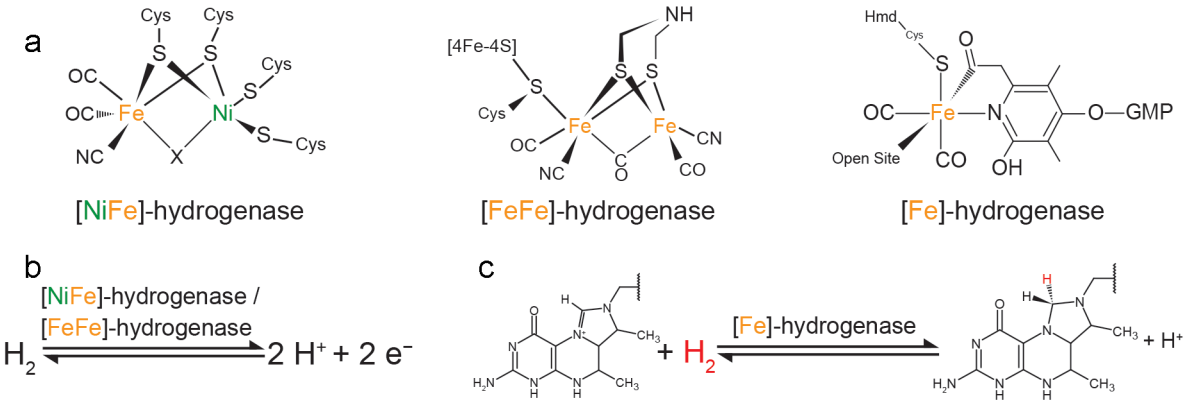


Figure 2. Structure of the hydrogenases' active sites and their respective reactions. (**a**) Metal centers of the different hydrogenases. (**b**) Reaction catalyzed by [NiFe]- and [FeFe]-hydrogenases. (**c**) Reversible heterolytic cleavage of H_2 into a hydride and a proton catalyzed by [Fe]-hydrogenase. The hydride transferred stereo-specifically to methylene-H_4MPT is shown in red.

4.2.1. [NiFe]-hydrogenases

As their name suggests, this type of hydrogenases possesses a binuclear metal center composed of Ni and Fe atoms.[26–28] There are four phylogenetically distinct groups: group 1 contains membrane-bound H_2-uptake hydrogenases; group 2 contains uptake- and sensory-hydrogenases; group 3 contains F_{420}-reducing, $NAD(P)^+$-reducing, heterodisulfide-associating and bidirectional $NAD(P)^+$-reducing hydrogenases; group 4 contains energy-converting hydrogenases. The first crystal structure of a [NiFe]-hydrogenase was reported by Volbeda *et al* in 1995.[29] Since then, extensive works of the structure of several [NiFe]-hydrogenases have been made, which is summarized in a review by Lubitz *et al.*[26]

The different [NiFe]-hydrogenases groups share a common folding architecture.[26] This basic structure of [NiFe]-hydrogenases consists of a large (~63 kDa) and a small (~29 kDa) subunit.[26,27] The small subunit contains three [Fe-S]-clusters that are responsible for the electron transfer from the active site to the physiological partner. These clusters are named proximal, medial, and distal depending on their distance from the active site. The large subunit is composed of two *αβ*-domains, a helical domain and two less-structured regions.[26,29] The [NiFe]-center is located in the large subunit and an additional metal that is assigned to Mg is located on the C-terminus.[30] The [NiFe]-center is the catalytic center of the enzyme. Four strictly conserved cysteine residues are coordinated to the Ni atom, two of them in a terminal fashion and two as bridging ligands to the Fe atom. The Fe ion is a low-spin Fe(II) in octahedral configuration with three ligands and two of the bridging cysteine thiolates. The ligands are two CN^- and one CO, and where first shown by Fourier-transform infrared spectroscopy (FTIR) studies.[31] A third bridging ligand is coordinated between the Ni and the Fe atoms, and its identity depends on the states of the active site (Figure 2a).[26]

4.2.2. [FeFe]-hydrogenases

This class of hydrogenases possesses two Fe atoms in its metal center. The first structure of a [FeFe]-hydrogenase was published in 1998, and revealed that the monomeric enzyme possesses four [4Fe-4S]-clusters and one [2Fe-2S]-cluster, in addition to the binuclear [FeFe]-center.[32] Many [FeFe]-hydrogenases are monomeric; however, some [FeFe]-hydrogenases possess additional subunits.[33] The domain that harbors the active site with the catalytic metal center is conserved among organisms and homologous enzymes.[33] This metal center is called H-cluster and is composed of a [4Fe-4S]-cluster that is coordinated by a single cysteine thiolate and bridged to a binuclear Fe center coordinated by CO, CN^- and a dithiomethylamine ligand (Figure 2a).[32,34] In addition to the H-cluster, the additional [4Fe-4S]-clusters form an electron transfer chain. There is an exception, [FeFe]-hydrogenase of *Chlamydomonas reinhardtii* does not contain any auxiliary [Fe-S]-clusters other than the H-cluster.[35,36]

4.2.3. [Fe]-hydrogenase (Hmd)

Both [NiFe]- and [FeFe]-hydrogenases catalyze the reversible cleavage of dihydrogen into protons and electrons (Figure 2b).[3,26,30] The third type of hydrogenase, [Fe]-hydrogenase, catalyzes the heterolytic cleavage of hydrogen into a proton and a hydride and subsequent transfer of the hydride to methenyl-H_4MPT^+ to form methylene-H_4MPT (Figure 2c).[13,15,37] Hmd is found in hydrogenotrophic methanogenic archaea without cytochromes.[3,7,15] This hydrogenase contains two moles Fe per mole homodimer but no [Fe-S]-clusters.[13] In the initial stage of the researches, Hmd was called "[Fe-S]-cluster free hydrogenase" and afterward the name was rephrased to "metal-free hydrogenase" based on the proposed catalytic mechanism without a redox-active metal center.[38,39] However, it was later found that Hmd contains the iron-guanylylpyridinol (FeGP) cofactor, as described below. From then on, Hmd is called [Fe]-hydrogenase.

Hmd production is upregulated under Ni-limiting conditions. In a methanogen, *Methanothermobacter marburgensis* growing under Ni-limitation conditions (< 200 nM Ni^+), the expression of Frh is down-regulated, and Mtd and Hmd are up-regulated.[15,40] Because Frh activity is decreased, the cells regenerate $F_{420}H_2$ from H_2 utilizing the coupled enzyme reaction of Hmd and Mtd.[15,41,42] Hmd catalyzes the H_2-dependent reduction of methenyl-H_4MPT^+ to methylene-H_4MPT, then Mtd catalyzes the dehydrogenation of methylene-H_4MPT to methenyl-H_4MPT^+ coupled to the reduction of F_{420}.[15,41,42] In the homodimeric Hmd, the central domain is composed of an intertwined helix domain made up by two C-terminal segments of the monomers, and two peripheral Rossmann fold-like domains at the N-terminus of each peptide (Figure 3a).[43,44] The expression of the *hmd* gene in *Escherichia coli* yielded inactive enzyme. This phenomenon was originally interpreted as an incorrect

folding during heterologous expression.[38] It was later determined that the heterologously produced enzyme can be activated by mixing it with the filtrate of denatured purified active Hmd from *M. marburgensis*, which suggested that the enzyme requires a cofactor for its activity and that the cofactor can be reversibly extracted from the holoenzyme.[45]

This cofactor is sensitive to UV-A/blue light and along with light inactivation and EDTA treatment, one mole Fe dissociates from one mole protein.[46] These findings suggested that Fe is crucial for the enzymatic activity. Mössbauer studies determined that the cofactor contains low spin Fe(II) or Fe(0),[47] and FTIR indicated the presence of two CO ligands at an angle of 90°.[48] X-ray absorption spectroscopy using mutated enzymes indicated that a cysteine is an Fe ligand of the cofactor.[49] The structure of the light-inactivated cofactor (GP) was determined by NMR spectroscopy and mass spectrometry as a guanylylpyridinol derivative (Figure 3b).[50] Together with the structural information obtained by the crystal structure analysis, the first model of the cofactor was proposed.[43] However, subsequent crystal structure analysis of a mutated enzyme lacking the ligand cysteine indicated that the cofactor contains an acyl-ligand, which is the 6-substituent of the pyridinol ring.[51] Due to the structure and composition of the major parts of this cofactor, this cofactor is called the iron-guanylyl-pyridinol (FeGP) cofactor. The structure of the final model of the FeGP cofactor is shown in Figure 3b. The FeGP cofactor is composed of an Fe complex fixed with pyridinol nitrogen and its 6-acylmethyl substituent. The pyridinol ring is highly substituted and connects a guanosine monophosphate moiety. In addition, the iron atom is coordinated with two CO ligands and one cysteine-thiolate described above. The Fe site has an open coordination site, which is proposed as the H_2 binding site.[52,53]

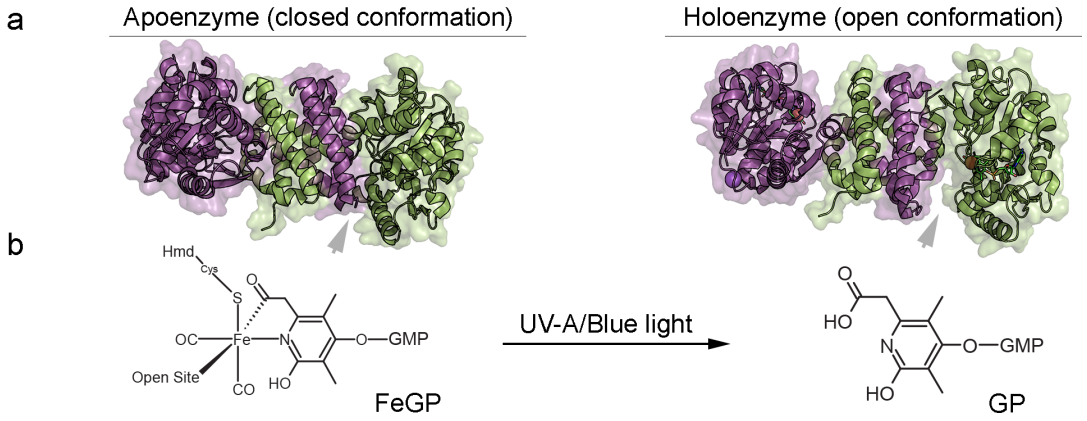


Figure 3. Structure of the [Fe]-hydrogenase Hmd from *Methanocaldococcus jannaschii*. (**a**) Structure of the apoenzyme (PDB: 2B0J) and the holoenzyme (PDB: 3F47) in the dimeric forms. The active site cleft formed by the binding of the FeGP cofactor is pointed with a grey arrow. Structures are in cartoon representation with transparent surface. (**b**) Structure of the FeGP cofactor and the light-inactivated product GP.

4.3. Maturation of [NiFe]- and [FeFe]-hydrogenases

4.3.1. [NiFe]-hydrogenase maturation

The requirement of at least six proteins is proposed for the maturation of this hydrogenase.[26,54] The biosynthesis sequence is composed of several stages: biosynthesis of the Fe center, incorporation of the Ni atom, and specific proteolysis at the C-terminus (Figure 4).[54] As described in Section 4.2, the Fe atom of this hydrogenase contains one CO and two CN^- ligands. Previous studies established that these ligands possess different metabolic origins. The CN^- ligand derives from carbamoyl phosphate[55–57] and is catalytically produced by HypF in complex with HypE.[54,57] The metabolic origin of the CO ligand is partially known. In the particular case of aerobic organisms, the CO ligand originates from the central C_1-pool, where HypX generates CO from the formyl group of N^{10}-formyl-tetrahydrofolate.[58] For anaerobic organisms without *hypX* in their genome, the origins of the CO ligand are still unknown. It has been proposed that HypC is involved in the formation of the CO ligand and interacts with HypD which assembles the final Fe center.[59,60] The Ni delivery is a process that is accomplished by at least three proteins, HypA, HypB and SlyD,[61–63] and occurs after the Fe-center is inserted in the hydrogenase.[54] After the [NiFe]-complex was formed in the hydrogenase, a specific protease cleaves a peptide in the C-terminal, inducing the correct folding, which renders the active [NiFe]-hydrogenase.[54]

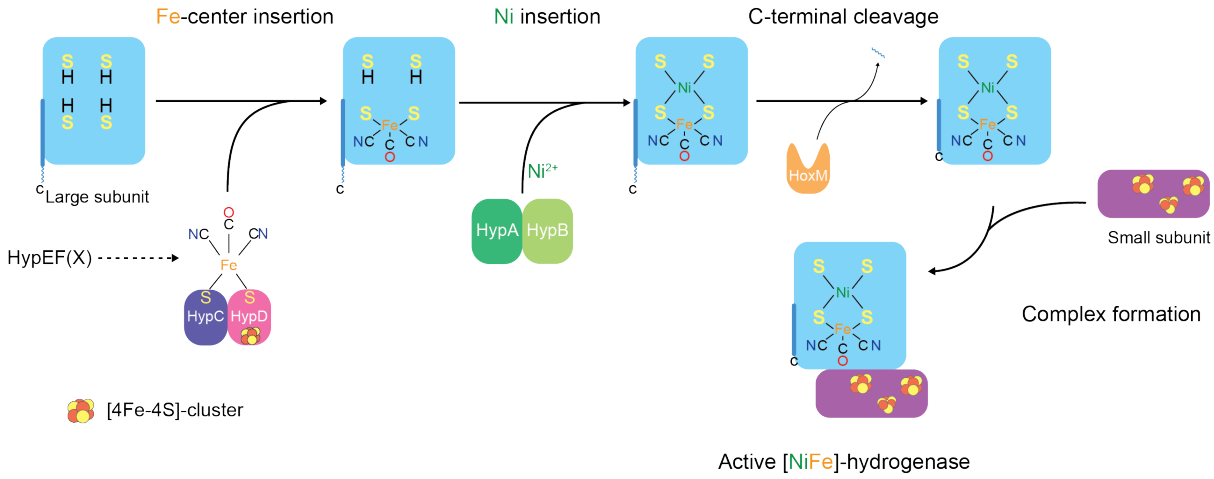


Figure 4. Maturation of the [NiFe]-hydrogenase. The maturation process occurs in several steps. The Fe-center is proposed to be formed in the HypCD complex by HypEF. In hydrogenases from aerobic organisms, CO is produced by HypX. After the Fe complex insertion, Ni is inserted in a GTP-dependent manner. The C-terminus is cleaved and only then the small subunit can bind and complex with the large subunit and form the active hydrogenase.

4.3.2. [FeFe]-hydrogenase maturation

This class of hydrogenase is expressed as an inactive apoenzyme with a [4Fe-4S]-cluster, to which the binuclear Fe site is inserted. As described before (Figure 2a), the H-cluster possesses a binuclear Fe-center with CO and CN^- ligands, in a similar fashion to the [NiFe]-hydrogenase, but the origins of these

ligands are different in these two types of hydrogenases. Another difference is the dithiomethylamine bridge only found in the H-cluster of the [FeFe]-hydrogenases, which connects the two Fe atoms. Maturation of [FeFe]-hydrogenase requires three specific proteins, HydG, HydE and HydF (Figure 5).[33,64] HydG is a radical *S*-adenosyl methionine (SAM) enzyme that forms the CO and CN^- ligands from tyrosine[65,66] and forms an [Fe(Cys)$(CO)_2$(CN)] synthon.[67] HydE is a radical SAM enzyme that was originally proposed to be involved in the formation of the dithiomethylamine bridge;[68] however, later studies have shown that HydE reduces the Fe(II) synthon from HydG to an adenosylated Fe(I) synthon that may serve as the precursor of the di-Fe center.[69] HydF is an GTPase that was proposed to be a scaffold protein that assembles the complete H-cluster and delivers it to the hydrogenase protein HydA.[64] Recently, the lipoyl-H-protein from the glycine cleavage system has been shown to be involved as a glycine donor which forms the bridging ligand in HydF.[70]

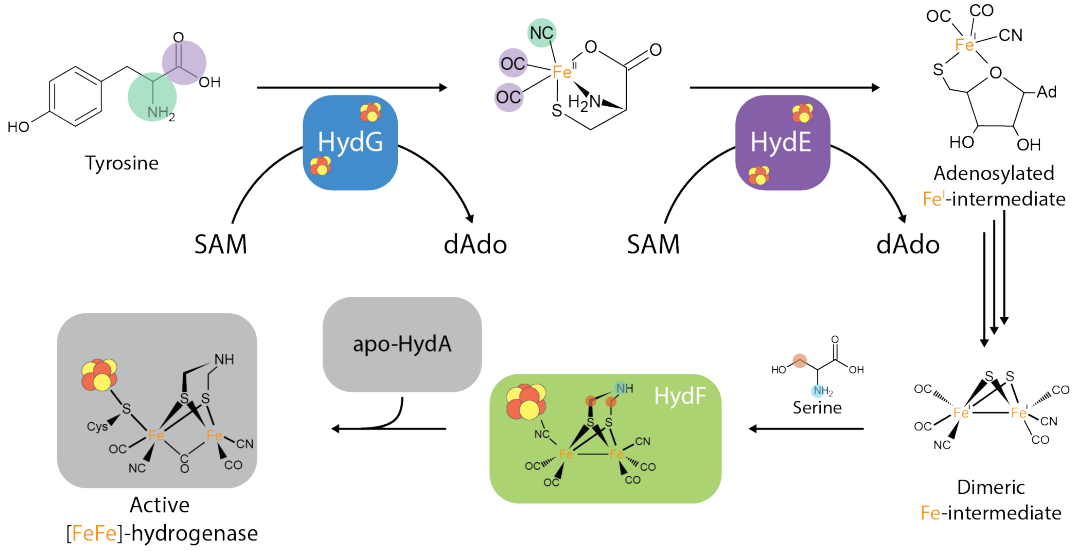


Figure 5. Maturation of the [FeFe]-hydrogenase. HydG produces CO and CN^- from tyrosine in a SAM-dependent radical reaction. Then HydE reduces the Fe-synthon produces to an adenosylated intermediate, which is then bound to HydF and the dithiomethylamine bridge is formed. This precursor is then transfer to the apoenzyme to produce the active [FeFe]-hydrogenase.

4.4. [Fe]-hydrogenase maturation and FeGP cofactor biosynthesis

Biosynthesis of the FeGP cofactor of [Fe]-hydrogenase is a complex pathway that involves at least seven different genes.[3,71,72] These genes, named *hmd co-occurring* genes (*hcgA-G*), were initially identified by comparative genomic analysis of the genome of methanogens that possess the *hmd* gene.[73] Deletion of the *hmd* gene in a strain *Methanococcus maripaludis* without the *frcA* and *fruA* genes, which is referred to as the *Δfrh* strain in my thesis, leads to a phenotype with an increased lag phase in the growth under H_2/CO_2.[72] In the same studies using the Δ*frh* mutant, strains with deletion of each of the *hcg* genes were made, and it was observed that the *hcg* mutants possess the same phenotype under H_2/CO_2, suggesting that the *hcg* genes could be related to the Hmd activity in the cell.[72]

Stable isotope labelling studies were done to elucidate the biosynthetic origins of the FeGP cofactor. *Methanothermobacter marburgensis or Methanobrevibacter smithii* were grown with ^{13}C- and ^{2}H-labelled precursors (acetate, pyruvate, methionine, CO_2, CO etc.). The CO ligands and the acyl-ligand were determined to originate from either $^{13}CO_2$ or ^{13}CO, while the methyl group in the 3-position of the pyridinol ring is derived from the methyl group of methionine via *S*-adenosyl methionine (SAM).[71] The pyridinol ring itself exhibits a complex labelling pattern with labels originating from three C-1 of acetate, two C-2 of acetate, and two C-1 of pyruvate (Figure 6).

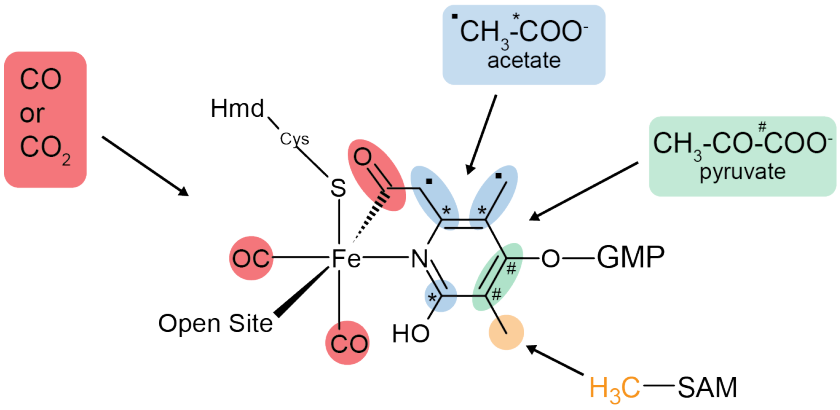


Figure 6. Stable isotopic labelling of the FeGP cofactor.[71]

Some of the Hcg proteins have been characterized structurally and biochemically. This has led to a proposal of the biosynthesis pathway.[73–78] The biosynthesis pathway (Figure 7) starts from 6-carboxymethyl-5-methyl-4-hydroxy-2-pyridinol (Precursor **1**). HcgC is SAM-dependent methyltransferase that transfers a methyl group of SAM to the C-3 of **1** to form precursor **2**.[74,78] HcgB is a guanylyltransferase that hydrolyses guanosine triphosphate (GTP) and transfers the guanosine monophosphate (GMP) moiety to precursor **2** to form precursor **3.**[76] The precursor **3** is called guanylylpyridinol (GP) and it is the light-decomposition product of the FeGP cofactor as described above.[50,79] The carboxymethyl group of **3** is activated by adenosine triphosphate (ATP)-dependent adenylylation catalyzed by HcgE**.**[77] The AMP-activated precursor is transferred to HcgF, where GP is bound to the Cys9 thiolate by thioester bonding (precursor **4**).[77] The steps after this process, namely the formation of the iron center, including the formation of CO ligands and the acyl-ligand, and the subsequent transfer to the apoenzyme are still unclear. HcgD has been proposed to be an iron trafficking protein or an iron chaperone because it harbors a binuclear iron center, in which one of the iron is strongly bound and the other can be removed by chelating agents.[75] An alternative hypothesis presented by another research group is that this protein could be a regulator;[80] however, this hypothesis is unlikely because the DNA binding properties presented are unspecific.

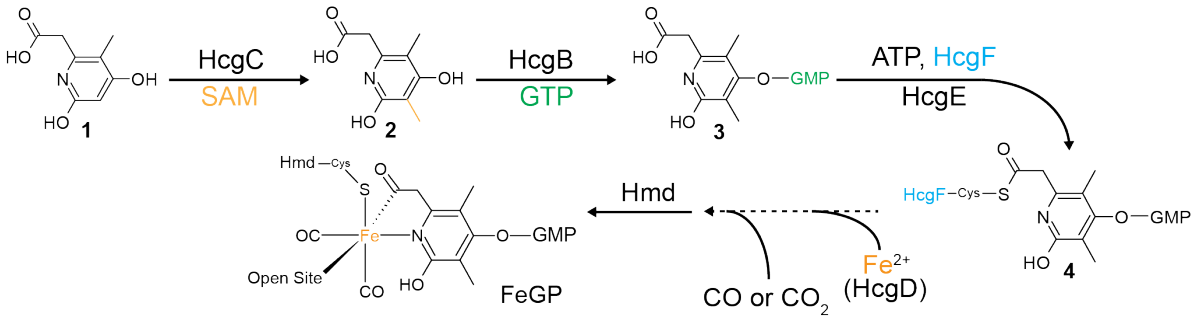


Figure 7. Proposed biosynthesis pathway of the FeGP cofactor before this work.

The functions of HcgA and HcgG are still unclear. HcgA from *M. maripaludis* has been found to be a possible radical SAM enzyme.[81] The primary structure of HcgA bears homology to HydE and HydG, both of which are [FeFe]-hydrogenase maturation proteins.[68] HydE and HydG are responsible for the formation of the binuclear Fe-center of the [FeFe]-hydrogenase[67,68,82] as described in the section before (See section 4.3). From the similarity to these enzymes, HcgA was proposed to be involved in the formation of the CO ligands of the FeGP cofactor.[73] The crystal structure of HcgA was not experimentally studied; nonetheless, the homology with HydE and HydG presents an opportunity to obtain a homology model (Figure 8b). HcgA does not contain the classical [Fe-S]-cluster binding motif CX_3CX_2C; instead, it has a unique CX_5CX_2C motif, suggesting that this protein is a new member of the radical SAM family.[81,83] HcgG is annotated as a fibrillarin-like protein that could bind SAM;[3] however, the function of HcgG is obscure and all attempts to heterologously produce this enzyme resulted in insoluble protein.[73] The final step of the biosynthesis is the incorporation of the FeGP cofactor into the Hmd apoenzyme. The Hmd apoenzyme can be reconstituted *in vitro* by incubation with the extracted and purified cofactor;[84,85] however, the requirement of a scaffold protein *in vivo* cannot be discarded.

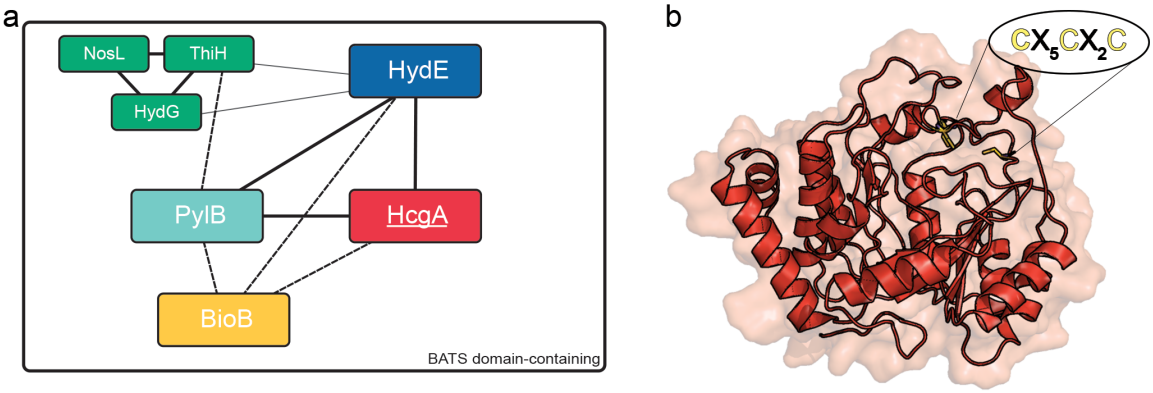


Figure 8. Homology network and structure of HcgA. (**a**) Similarity network of the subfamily of biotin and thiamin synthesis-associated (BATS) domain-containing radical SAM enzymes. Lines represent the homology level: thick black lines represent close homology, dashed thick lines represent medium homology and thin lines represent more distant homology. (**b**) Homology model of HcgA from *Methanocaldococcus jannaschii* constructed by the Robetta web server.[86] Cysteines in the predicted [Fe-S]-cluster binding motif are shown in stick models.

4.5. The role of radical SAM enzymes in the maturation of hydrogenases

As described in the previous sections, the biosynthesis of the [Fe]-hydrogenase requires one or more radical SAM enzymes. Radical SAM enzymes are a diverse family of enzymes that catalyze a wide range of reactions.[83,87,88] Although this class of enzymes catalyzes diverse reactions, they all share a common mechanism involving the production of radicals from SAM. In these enzyme reactions, SAM is reductively cleaved to produce a 5'-deoxyadenosyl radical (dAdo·) that abstracts a hydrogen from the substrate, forming 5'-deoxyadenosine (dAdo)[83,88,89] and the product (Figure 9). Although some members of this family show only limited sequence identity, they are structurally similar. All radical SAM enzymes characterized so far possess an incomplete or full triose-phosphate isomerase (TIM) barrel architecture, where three cysteines coordinate a [4Fe-4S]-cluster with a characteristic motif CX_3CX_2C,[83,88] although some members harbor a different binding motif. The conversion of SAM to dAdo· occurs even in the absence of substrate; however, it is greatly enhanced by incubation with the substrate.[68,81,83,88] This behavior is useful in the determination of the substrates where they cannot be determined by whole enzyme reactions. Radical SAM enzymes are involved in the biosynthesis of several cofactors (Figure 10).[69,88,90–92] In the specific case of hydrogenases' maturation, HydG and HydE are involved in the formation of the H-cluster in [FeFe]-hydrogenase.[66,69,93]

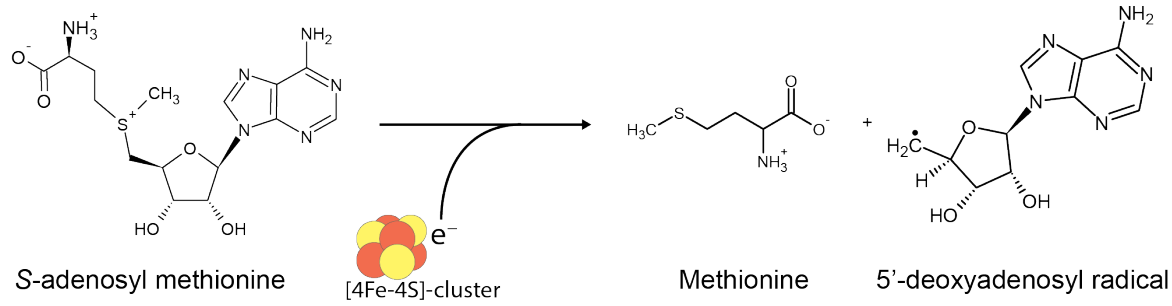


Figure 9. Production of the 5'-deoxyadenosyl radical from SAM requiring one electron.

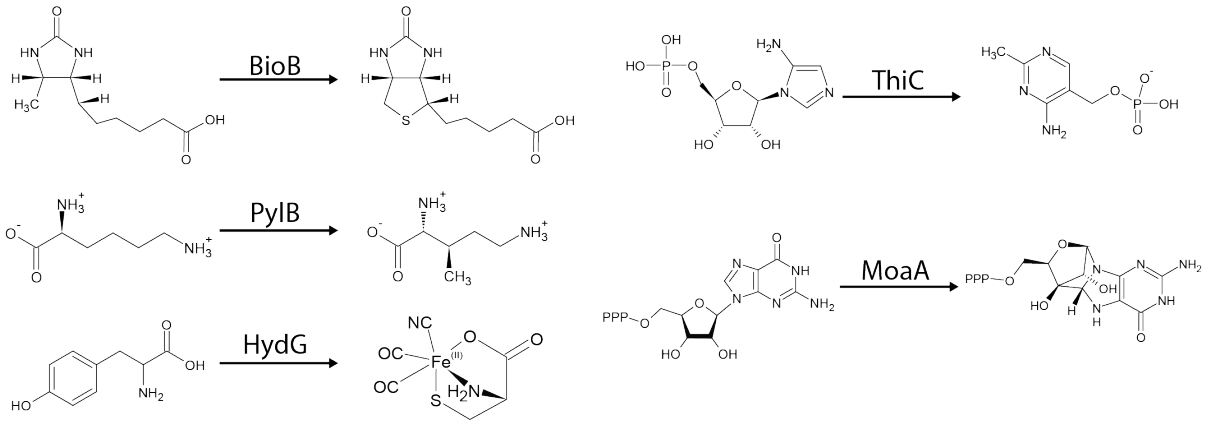


Figure 10. Selected examples of radical SAM reactions involved in cofactor biosynthesis. BioB (biotin synthase) catalyzes the production of biotin. ThiC (HMP-P synthase) catalyzes the formation of an intermediate in the biosynthesis of thiamine. PylB (methylornithine synthase) catalyzes the isomerization of lysine to methylornithine. MoaA (a molybdenum cofactor biosynthesis protein) catalyzes the formation of an intermediate from GTP that becomes the pyranopterin moiety. HydG catalyzes the formation of the CO and CN^- ligands and the Fe-center precursor in [FeFe]-hydrogenase maturation.

4.5.1. HydG

This radical SAM enzyme is responsible for the formation of the CO and CN^- ligands in the H-cluster of [FeFe]-hydrogenase.[66] HydG belongs to the biotin and thiamin synthesis-associated (BATS) domain-containing subgroup of radical SAM enzymes,[94] and possesses a classical radical SAM [4Fe-4S]-cluster and an auxiliary [4Fe-4S]-cluster.[66,93,95] Observation of *in vitro* maturation in the cell extract of *E. coli* determined that incubation with tyrosine and cysteine improved the [FeFe]-hydrogenase activity by enhancing the maturation.[96] It was later determined that HydG produces the CO and CN^- ligands from tyrosine in a SAM-dependent fashion.[66,97] Spectroscopic studies on the catalytic mechanisms of HydG determined that the enzyme cleaves tyrosine to produce CO and CN^-, which are then bound to an Fe atom in the second [4Fe-4S]-cluster to produce a $[Fe(Cys)(CO)_2(CN)]$ synthon.[67] After the crystal structure was obtained, it was determined that the synthon is formed in an additional Fe atom (dangler Fe) coordinated to the second [4Fe-4S]-cluster and an external cysteine ligand.[65]

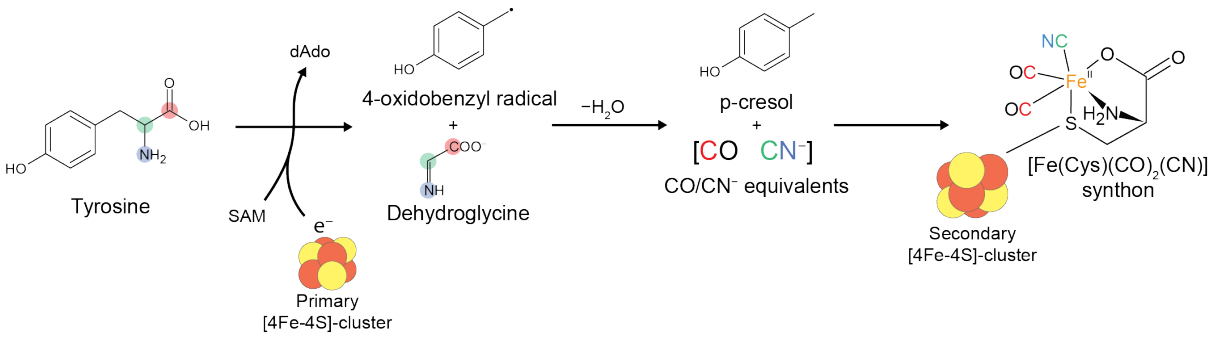


Figure 11. Proposed formation of the $[Fe(Cys)(CO)_2(CN)]$ synthon by HydG. The reaction starts by reductive cleavage of SAM and production of dehydroglycine from tyrosine. This dehydroglycine is then transformed into the CO and CN^- equivalents and subsequently bound to a dangler Fe coordinated to the secondary [4Fe-4S]-cluster via an external cysteine.

4.5.2. HydE

The second radical SAM enzyme involved in the maturation of the [FeFe]-hydrogenase is HydE, which also belongs to the BATS domain-containing subgroup of radical SAM enzymes.[94] In some organisms, the protein is expressed as a fusion product with HydF.[93,98] The enzyme possesses the canonical [4Fe-4S]-cluster binding motif CX_3CX_2C, and spectroscopic analysis suggested that this enzyme may possess an additional [Fe-S]-cluster,[93,99] which was later confirmed by determination of the crystal structure.[99,100] It was first proposed that this enzyme was responsible for the formation of the dithiomethylamine bridging ligand of the H-cluster,[68] however later studies spectroscopically detected that the enzyme forms an adenosylated Fe(I) intermediate that serves as a precursor for the binuclear Fe-center of the H-cluster (Figure 12).[69] More recently, it was determined that the dithiomethylamine bridging ligand is formed in HydF using aminomethyl-lipoyl-H-protein as the substrate.[70]

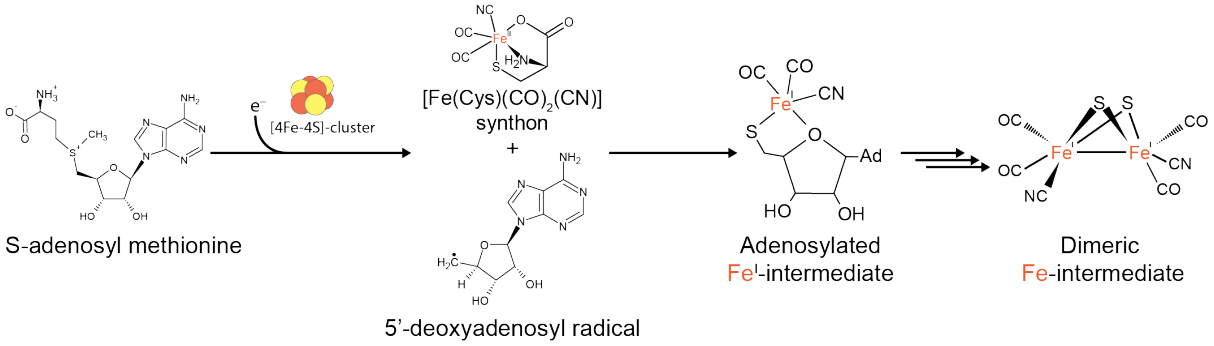


Figure 12. Overview of the proposed reaction pathway involving HydE in the formation of the H-cluster. First, 5'-deoxyadenosine radical (dAdo·) is formed from SAM, which then reacts with the $[Fe(Cys)(CO)_2(CN)]$ precursor from HydG. This forms an adenosylated intermediate, which then dimerizes into the proposed dimeric Fe-intermediate.

4.6. *In vitro* biosynthesis of the FeGP cofactor

Studies of the biosynthesis of the FeGP cofactor were initiated from a "structure to function" approach, where the structure of the HcgB, HcgC, HcgD, HcgE, and HcgF proteins were determined, and from the structure, the reactions were inferred.[73–77] The estimated enzymatic activities of the Hcg proteins other than HcgD were demonstrated by enzymological and/or structural methods. Although this approach was effective in determining the function of these enzymes, it cannot be excluded that these enzymes catalyze similar reactions using slightly different substrates. The *in vitro* maturation of [FeFe]-hydrogenases have been used to elucidate the function of the maturation enzymes,[101] where mostly cell extract is used in combination with heterologously produced maturation proteins to drive the *in vitro* biosynthesis of the cofactors and incorporation into the hydrogenase. However, before our study, no *in vitro* biosynthesis of the FeGP cofactor of [Fe]-hydrogenase was available.[102,103]

At first, we developed an *in vitro* biosynthesis assay for the FeGP cofactor using the cell extract of the Δ*hcgB*Δ*hcgC* strain of *Methanococcus maripaludis*.[102,103] This *in vitro* system contained [Fe]-hydrogenase apoenzyme, possible precursors, ATP/Mg^{2+}, SAM, dithiothreitol (DTT) and sodium dithionite, and the Δ*hcgB*Δ*hcgC* strain cell extract lacking endogenous [Fe]-hydrogenase activity. This system was used to confirm the proposed precursors **1**, **2**, and **3**, and elucidated additional requirements for the biosynthesis of the FeGP cofactor, namely CO-precursor (as CO gas or a CO precursor-compound) and reducing equivalents, which can be generated from H_2, formate and/or CO.[102,103]

The development of this *in vitro* biosynthesis assay opened the door to the study of the components of the biosynthesis of the FeGP cofactor such as reducing equivalents, CO precursor and other unknown components, and also it allowed the study of the function of unknown proteins (i.e., HcgA and HcgG).

4.7. Aim of this work

Several steps of the biosynthesis of the FeGP cofactor of the [Fe]-hydrogenase are still unknown: the formation of the first pyridinol precursor, the formation of the CO and acyl ligands, and the assembly and transfer of the final cofactor to the apoenzyme. The uncharacterized enzymes, HcgA and HcgG, have been historically difficult to work with on the identification of the functions. However, the involvement of two possible radical SAM enzymes, HcgA and HcgG, in the FeGP cofactor is reasonably predicted. Genomic analysis of methanogens showed that only the HcgA-G proteins are involved in the biosynthesis; therefore, we hypothesized that either HcgA or HcgG must be responsible for the production of the first pyridinol precursor **1**. Initially, due to the homology of HcgA to HydE and HydG, we hypothesized HcgA could produce the CO-ligands and make the Fe-center, and that HcgG is responsible for the formation of **1**.[80]

During my PhD project, I was able to heterologously and homologously produce and purify HcgA and HcgG, respectively. The amount of the purified enzymes was sufficient for their biochemical and functional characterization. Using the *in vitro* biosynthesis assay, I determined that HcgA is involved in a step before the formation of precursor **1**, and further enzyme assays using MS analysis of the reaction with protein-free filtrate of the cell extract revealed that the enzyme catalyzes the production of **1** from an unknown substrate. *In vitro* biosynthesis assays revealed that HcgG plays a role after the production of precursor **3**. Further studies showed that HcgG contributes to the formation of the CO-ligands from CO in the gas phase or from an unknown small-molecule precursor. I also discussed the possible CO precursors based on the CO-forming experiments using purified HcgG. In addition, I studied the crystal structures of Hmd from *Methanolacinia paynteri* and kinetic analysis of

incorporation of the FeGP cofactor into the protein. In combination with mathematical models and the use of a mutated enzyme, the data revealed a role for a conserved lysine residue in the binding path of the FeGP cofactor to the apoenzyme. Furthermore, I studied the role of an Hmd paralog (HmdII). Mutation studies and proteome analysis suggested that HmdII could be involved in the regulation of the Hmd activity. In conclusion, this work expanded our knowledge of the biosynthesis of the FeGP cofactor, incorporation of the FeGP cofactor and regulation of Hmd production.

5. Results and Discussion

5.1. Expression, purification, and characterization of HcgA

5.1.1. Heterologous production of HcgA

HcgA has been previously expressed and purified by our group and other researchers;[80,81] however, the function of the enzyme is unknown. A His-tagged version of HcgA from *Methanocaldococcus jannaschii* was produced and purified as previously described,[80] but this preparation rapidly aggregated and precipitated after purification. To solve the solubility problem, the previously described Tris/HCl pH 8.0[80] buffer was exchanged to sodium phosphate buffer pH 7.4 for the Ni-affinity purification. After this, the protein was desalted in MOPS/NaOH pH 7.4 containing 500 mM NaCl and 10% glycerol according to the system previously published for HcgA from *M. maripaludis*.[81] This slightly improved the stability during the purification; however, overnight storage or freeze-thawing still led to aggregation. Assuming the protein from *M. jannaschii* was unstable under these conditions, I changed the source organism to the mesophilic archaeon *M. maripaludis*. A His-tagged version of HcgA from *M. maripaludis* was produced in *E. coli* containing the helper plasmid pRKISC for the overexpression of [Fe-S] cluster-containing proteins.[104] HcgA from *M. maripaludis* was more stable in MOPS/NaOH pH 7.4 containing 500 mM NaCl and 10% glycerol under the purification, in which there was lower aggregation and precipitate formation under regular storage conditions. However, repeated freeze-thaw cycles increased the aggregation. To avoid this, the protein was aliquoted, quickly frozen in liquid N_2, and stored at −20 °C. The thawed sample was used only one time. The initial iron content was determined to be 2.6 ± 0.9 moles Fe/moles HcgA, although a previous paper indicated HcgA contains one [4Fe-4S]-cluster per protein.[81] Substoichiometric iron content is often observed in heterologously expressed and purified [Fe-S] cluster-containing proteins.[81,83,105] UV-visible spectroscopy showed a peak at approximately 410 nm, which corresponds to [Fe-S]-clusters as observed before[81] and it is typical for radical SAM enzymes (Figure 13b). Incubation with 2 mM sodium dithionite is expected to reduce the [Fe-S]-clusters, reducing the absorbance of this peak; however in this case, there was no significant difference when dithionite was added. This suggests an incomplete assembly of the [Fe-S]-clusters. It is known that [Fe-S]-clusters can be reconstituted by incubating proteins in DTT, sodium sulfide and ferrous iron.[81,106,107] By this method, I reconstituted the [Fe-S]-cluster of HcgA. The reconstitution experiments were performed with six equivalents of Fe^{2+} and S^{2-} to the purified HcgA protein, which resulted in an increased Fe content. However, the [Fe-S]-cluster was not reducible with dithionite (Figure 13b).

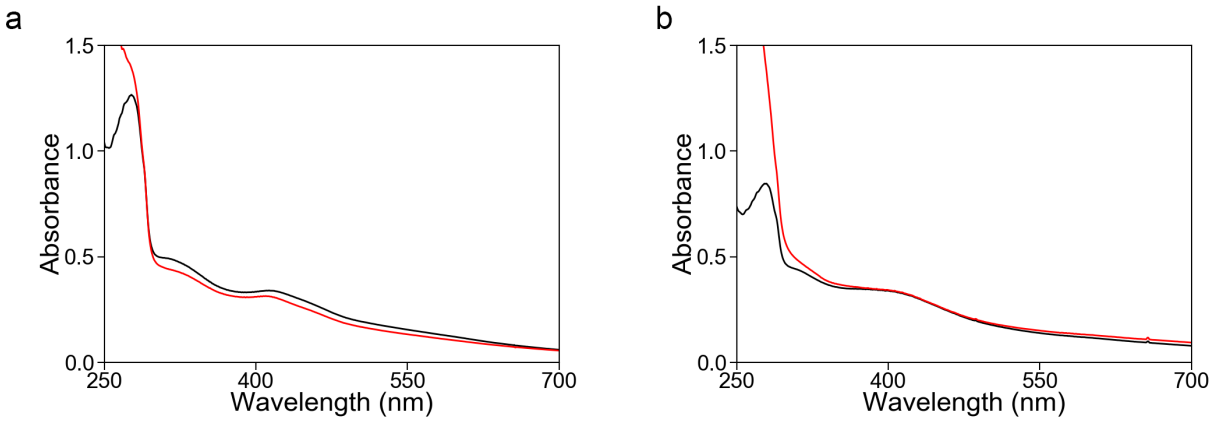


Figure 13. UV-visible spectra of HcgA. (**a**) Spectra of as-isolated HcgA (3mg/ml). (**b**) Spectra of [Fe-S] cluster-reconstituted HcgA (2 mg/ml). Samples were measured before (black line) or after (red line) incubation with 2 mM sodium dithionite. The measurement was performed 3-mm light path quartz cuvette.

5.1.2. Production of HcgA with a reducible [Fe-S]-cluster

In one purification, expression of the gene of HcgA was started at OD_{600} = 1 instead of the usual 0.6 to 0.8 and induction continued at 20 °C overnight. This preparation had an Fe content of 4.0 ± 0.38 moles Fe/moles HcgA. Additional [Fe-S]-cluster reconstitution of this protein led to changes in the Fe content of the protein, as shown in Table 1. Increasing added Fe^{2+} and S^{2-} equivalents led to a higher final Fe content; however, the excess amount of Fe could be explained by unspecific loading of Fe to the protein, which is supported by the similar UV-visible spectra of these preparations (Figure 14a). More importantly, these [Fe-S]-clusters were reducible with 2 mM sodium dithionite (Figure 14b). As this protocol reproducibly produced protein with Fe content of approximately four and a reducible [4Fe-4S]-cluster, I used this protocol for the HcgA preparation. For the experiments described in this thesis, I routinely purified HcgA by Ni-affinity chromatography and subsequent gel permeation chromatography. For the protocol, see section 8.1.

Table 1. HcgA reconstitution of [Fe-S]-clusters with different equivalents of Fe^{2+} and S^{2-}.

	Equivalents of Fe^{2+} and S^{2-} [a]				
	0[b]	5	10	20	As isolated
moles Fe/moles HcgA	2.59 ± 0.04[c]	3.40 ± 0.35	5.74 ± 1.29	6.84 ± 0.26	4.00 ± 0.38

[a] The concentration of Fe^{2+} and S^{2-} was adjusted to make 0, 5, 10, 20 molar equivalents to HcgA.
[b] The sample was reconstituted in the buffer without Fe^{2+} and S^{2-}.
[c] Standard deviation of two measurements.

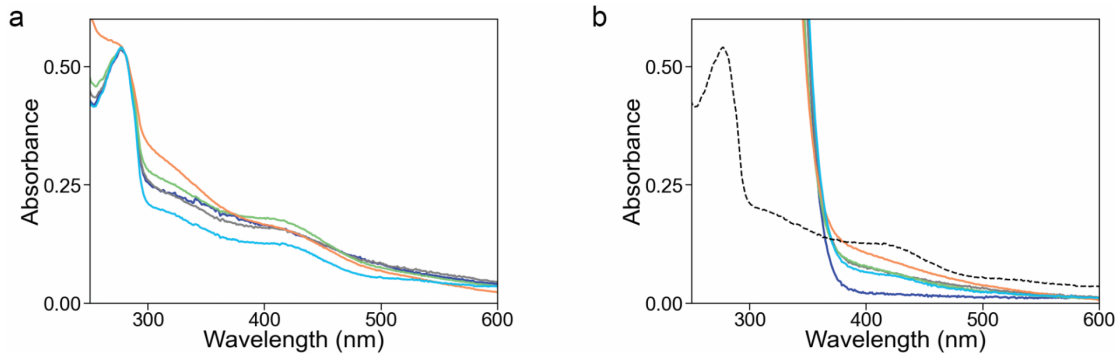


Figure 14. UV-visible spectra of HcgA obtained by the improved method (see Table 6): reconstituted with 0 (blue), 5 (gray), 10 (green) and 20 (orange) equivalents of Fe^{2+} and S^{2-}, and as isolated (cyan). The spectrum of the samples were normalized to the spectrum of as isolated sample by protein concentration. Spectra of the HcgA samples before (**a**) and after incubation with 2 mM sodium dithionite (**b**). In panel **b**, the UV-visible spectrum of the HcgA sample as isolated (before sodium dithionite reduction) (dashed black line) is shown for a comparison. The measurement was performed in a 3-mm light path quartz cuvette under N_2 atmosphere.

The HcgA protein eluted as a single peak from the HisTrap column with a relatively high degree of purity (Figure 15a). To remove small amounts of the contaminating proteins, HcgA was further purified by gel permeation chromatography. HcgA eluted at 70 mL in the Sephacryl S-300 HR column, which corresponds to an apparent molecular mass of 66 kDa that is possibly a dimer (monomeric mass 38 kDa) (Figure 15b). Notably, a previous publication indicated that HcgA from *M. maripaludis* is monomer but in the report, the methods and data were not described.[81] The [Fe-S]-cluster was stable and reducible after gel permeation chromatography (Figure 16).

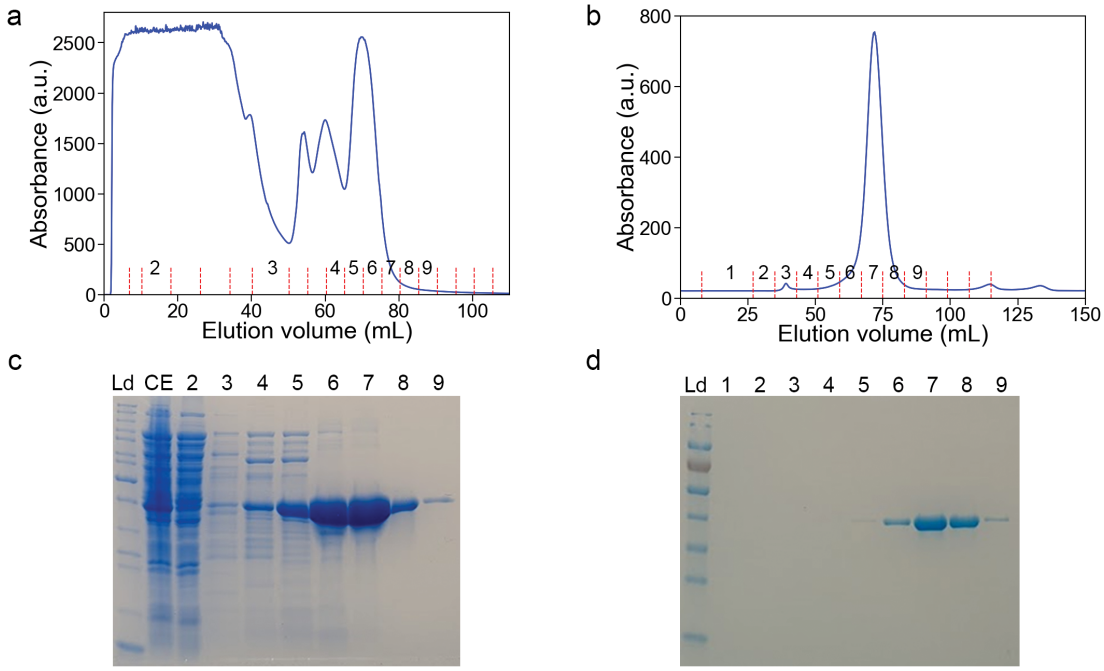


Figure 15. Purification of HcgA with a full [Fe-S]-cluster. (**a**) Chromatogram of the purification by Ni-affinity chromatography. (**b**) Chromatogram of the gel permeation chromatography. (**c**) SDS-PAGE of the fractions from Ni-affinity chromatography. (**d**) SDS-PAGE of the fractions from the gel permeation chromatography. Numbers correspond to the fractions analyzed by the corresponding SDS-PAGE in panels **c** and **d**. Ld = molecular size ladder. CE = cell extract. The measurement was performed in a 3-mm light path quartz cuvette.

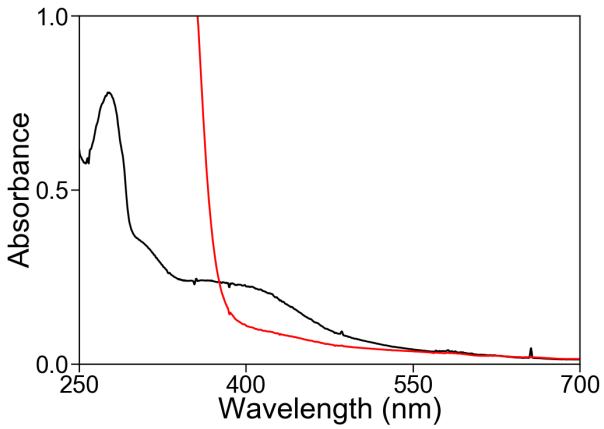


Figure 16. UV-visible spectra of purified HcgA after gel pemeation chromatography. As isolated protein is shown in black line. The protein concentration was 1.5 mg/ml. The peak at 410 nm is reduced after the addition of 5 mM sodium dithionite (red line). The measurement was performed in a 3-mm light path quartz cuvette under N_2 atmosphere.

5.2. Modeling of HcgG

I constructed several models of the tertiary structure of HcgG using different web servers: Robetta,[86] RaptorX,[108] Swiss-MODEL,[109,110], I-TASSER[111,112], and Phyre2.[113] These methods heavily rely on sequence similarity to proteins with known structures. The modeled structures are variable, and no common fold was observed as shown in Table 2. All the programs produced incomplete models with a variable sequence coverage, except for the I-TASSER program (Figure 17a), which covered the whole protein. I used this model for the similarity search against known PDB files using the MADOKA web server[114] to find related proteins and infer a possible function. AP2, a clathrin adaptor protein (Table 2) showed the highest scoring hits. Clathrin proteins are involved in the formation of vesicles,[115] but bear no similarity to other proteins with known enzymological activity. An important note is that the Robetta and RaptorX indicated that small N-terminal regions of HcgG have low similarity to HydE and HydG, both of which are involved in the maturation of the [FeFe]-hydrogenase. This prompted me to do a deeper analysis of the sequence of HcgG.

Table 2. Comparison of the I-TASSER models against the PDB database by MADOKA web server.

PDB ID	TM-Score	Name
4UQI	0.835165	AP2 Clathrin Adaptor
2JKT	0.826870	AP2 Clathrin Adaptor Core
1W63	0.819696	AP1 Clathrin Adaptor Core
2XA7	0.778390	AP2 Clathrin Adaptor
2JKR	0.747327	AP2 Clathrin Adaptor
4NEE	0.738766	AP-2 alpha/simga2 complex bound to HIV-1 Nef

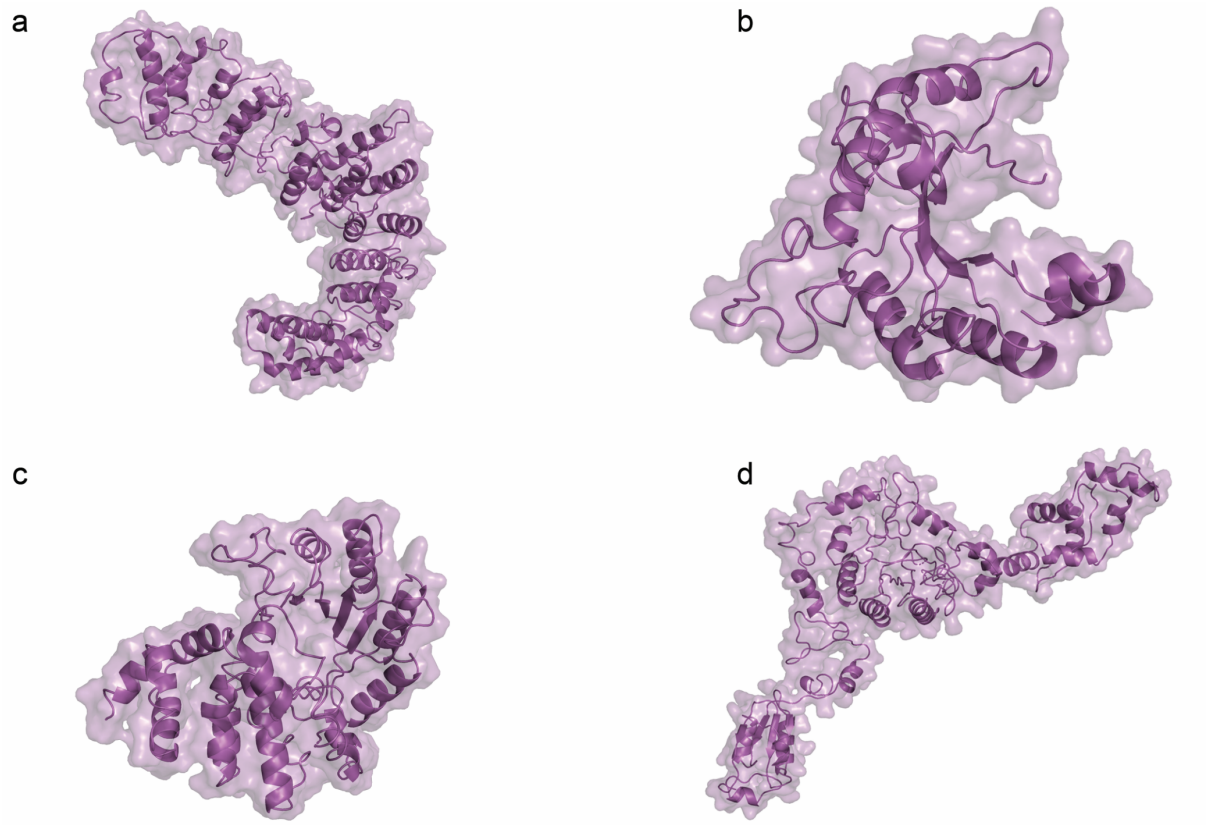


Figure 17. Models of HcgG. The models were created by the programs I-TASSER (**a**), Robetta (**b**), RaptorX (**c**), and Phyre2 (**d**).

The sequence of HcgG from several methanogens were analyzed by multiple sequence alignment with T-COFFEE,[116] which indicated two highly conserved regions connected with less conserved peptides (Figure 18). These two regions of HcgG are designated as N-terminal and C-terminal regions. Modelling of the N- and C-terminal regions was performed separately with the Robetta program to predict the structures. The MADOKA web server showed structural similarities of the N-terminal region to radical-SAM enzymes, HydG and NosL (Table 3). The similarity of the N-terminal to HydG hints that HcgG may be a radical SAM enzyme and could be involved in a similar process in the biosynthesis of the FeGP cofactor. The C-terminal region of HcgG showed similarity to DNA-binding regulatory proteins, which suggests that the C-terminal region of HcgG forms a nucleotide-binding domain (Table 3).

Table 3. Structure comparison of the I-TASSER models for the N-terminal and C-terminal against the PDB database by MADOKA web server**.**

PDB ID	TM-Score	Name
N-terminal HcgG model		
4WCX	0.745648	HydG[a]
4RTB	0.730309	HydG
4MWA	0.711512	GCPE
5C6M	0.709985	Deoxyribose-phosphate aldolase
4R34	0.704463	Tryptophan lyase NosL[a]
1SPQ	0.704111	Triosephosphate isomerase
C-terminal HcgG model		
4NIC	0.695828	RstA BeF3-activated N-terminal receiver domain[b]
3QZC	0.685407	Periplasmic stress response protein CpxP
3ILM	0.68425	Alr3790
3W9S	0.682344	N-terminal domain of Response Regulator PmrA[b]

[a] Radical-SAM enzymes.

[b] Nucleotide-binding protein

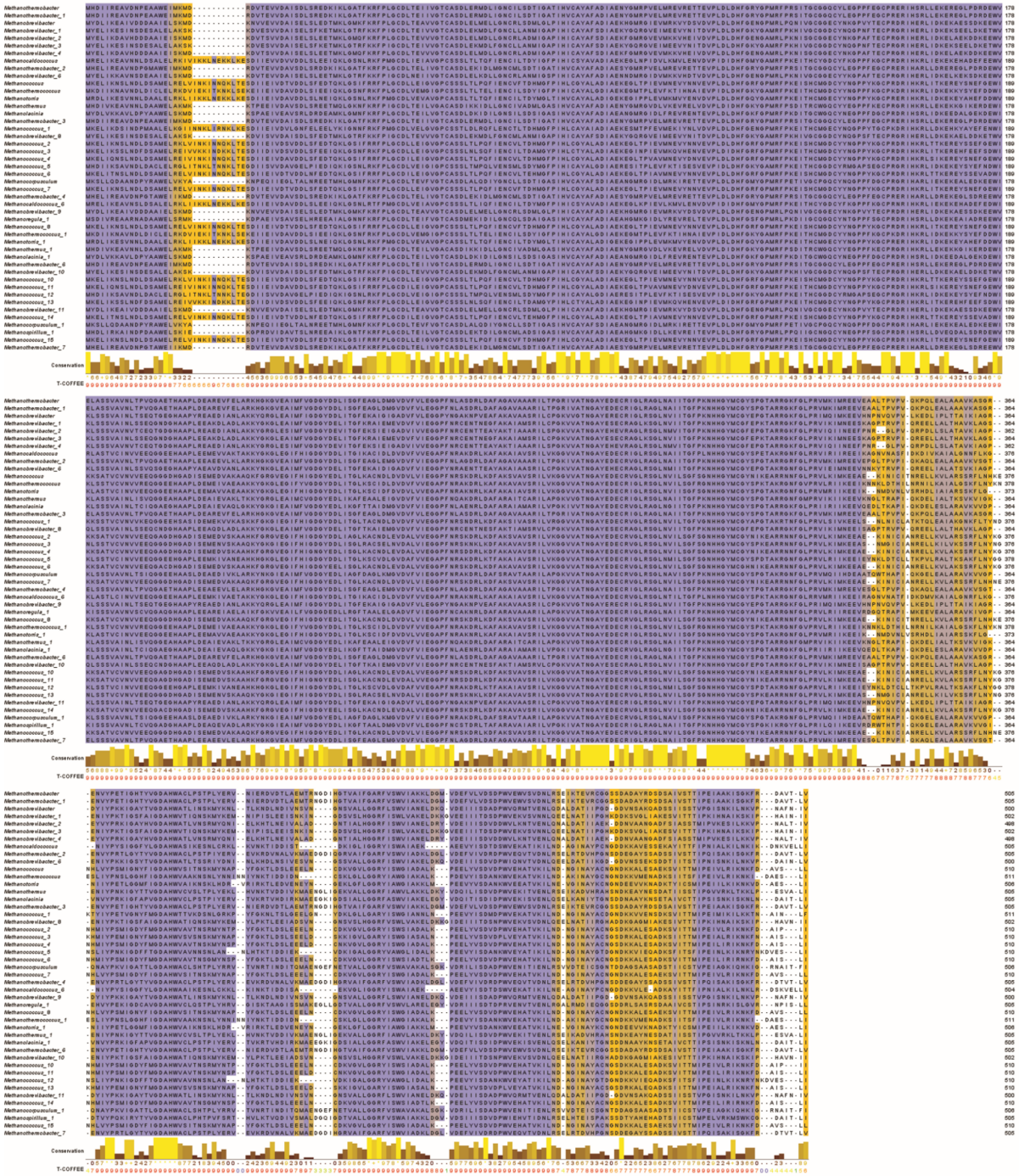


Figure 18. Multiple alignment of HcgG protein sequences. Color code by M-COFFE scores: residues in highly conserved regions are highlighted in purple (Score of 9) and residues colored in yellow are less conserved (score 7 or lower).

During the time of my PhD project, advances in the prediction of protein structures by neural networks, such as AlphaFold[117] and RoseTTAFold,[118] have dramatically improved the accuracy and quality of the predicted protein models; in addition, these modeling programs do not require templates for the prediction. The structure model of HcgG from *Methanocaldococcus jannaschii* has already been analyzed and is available in the AlphaFold database.[119] As I predicted, the model shows a two-domain architecture, where the N-terminal domain presents a TIM-barrel fold, with a loop containing three cysteine residues. The C-terminal domain has a Rossmann fold-like structure (Figure 19a). Using this model, both domains were compared against the PDB using the Dali web server.[120] The N-terminal domain showed similarities to HydG (Figure 19c and d), HydE, NosL and other radical SAM enzymes. Comparison of the C-terminal domain with the PDB database showed similarities with several nucleotide-binding proteins, including the Hmd paralog HmdII from *M. jannaschii* (Figure 19 b). HmdII has been shown to bind the FeGP cofactor.[121]

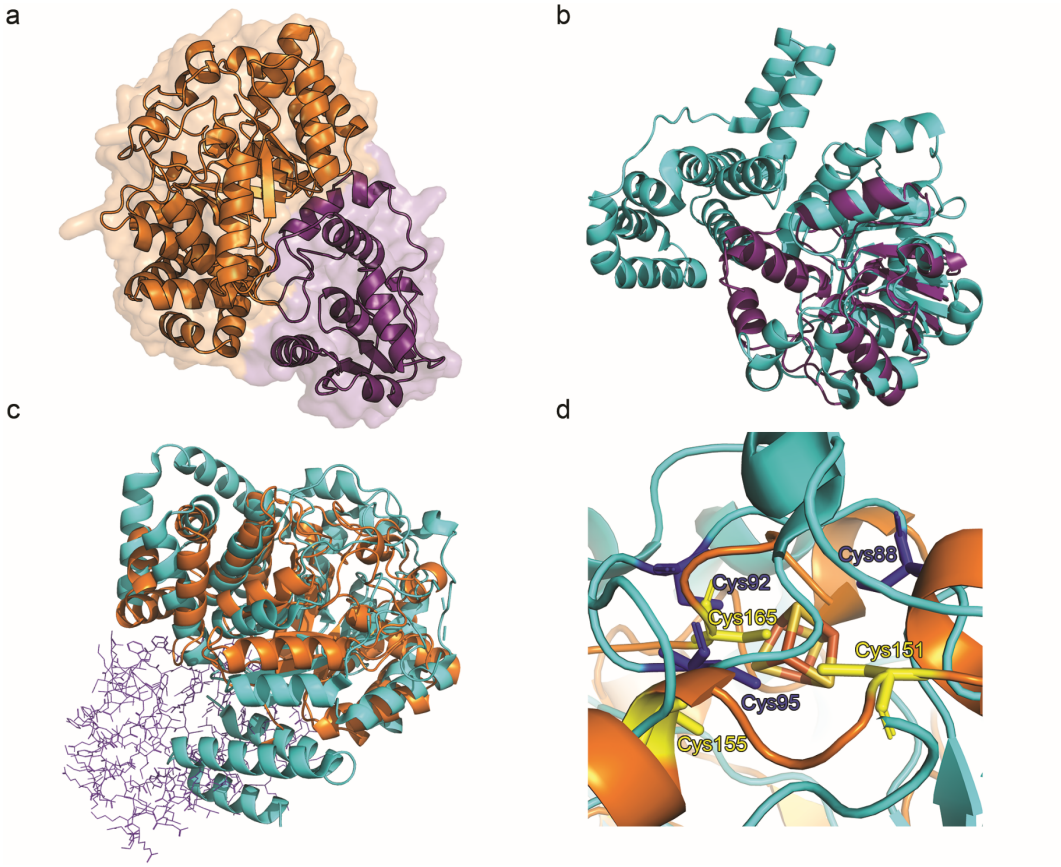


Figure 19. Structural prediction of HcgG by AlphaFold. (**a**) The whole structure of the AlphaFold model of HcgG from *Methanocaldococcus jannaschii*, the N-terminus (orange) and C-terminus (purple) are shown in cartoon model with transparent surface. (**b**) Structural alignment of the C-terminal domain of the HcgG model (purple) and the crystal structure of HmdII from *M. jannaschii* (cyan) (PDB ID: 6hux) with RMSD = 2.99 Å. (**c**) Alignment of the N-terminal AlphaFold model of HcgG from *M. jannaschii* (orange) to the crystal structure of HydG from *Carboxydothermus hydrogenoformans* (PDB ID: 4RTB) (cyan). The C-terminus of the HcgG model is depicted as purple lines. (**d**) The zoom-up view of the [4Fe-4S]-cluster binding region of HydG (cyan) with the three binding cysteine residues in blue, which is aligned with the AlphaFold model of HcgG (orange) with the conserved cysteine residues at the possible [4Fe-4S]-binding site in yellow.

5.3. Expression, purification, and characterization of HcgG

5.3.1. Optimization of the culture conditions

As stated before, HcgG has not been able to be produced in a soluble form in *E. coli*;[80] therefore, an alternative production method of HcgG was needed. In collaboration with Dr. Kyle Costa (University of Minnesota), a *M. maripaludis* strain was prepared to produce a His-tagged version of HcgG (See Materials and Methods for the construction of the strain). Initially, cells were grown in 5-liter bottles containing formate medium, as described previously for obtaining the *hcg* mutants' cell extract.[102] Cells grew until OD = 0.6−0.8. The protein was purified by Ni-affinity chromatography and eluted with an imidazole gradient. The protein eluted in two different peaks from the column (Figure 20a) and co-eluted with several smaller molecular mass bands as seen in the SDS-PAGE (Figure 20b). The total yield was less than 1 mg of protein per 5 grams of cells. Approximately 0.2% yield from the total protein was estimated.

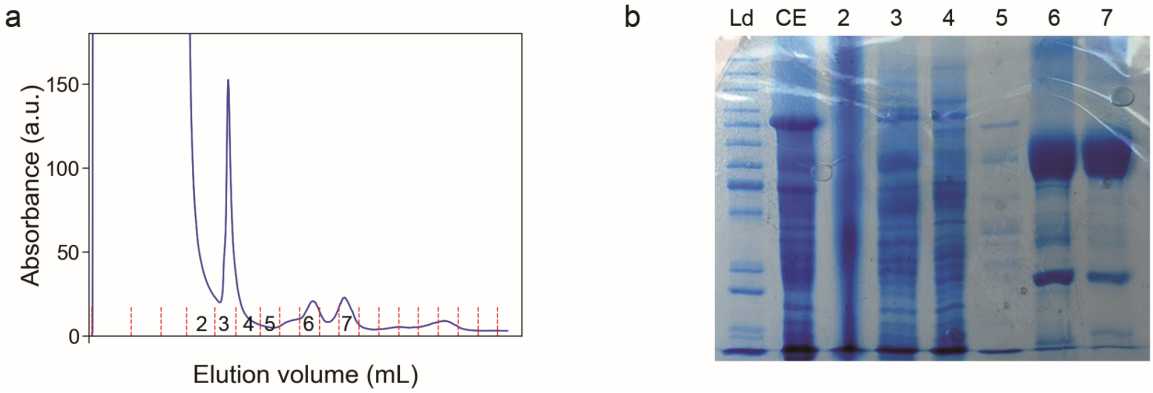


Figure 20. Purification of HcgG by Ni-affinity chromatography. (**a**) Chromatogram of the Ni-affinity purification. (**b**) SDS-PAGE of the fractions of the Ni-affinity chromatography. The lanes of the molecular size marker (Ld) and the cell extract (CE) were indicated. The fraction numbers of the Ni-affinity chromatography are indicated.

The His_6-tagged HcgG was constructed under the *M. voltae* histone promoter. Previous studies showed expressions up to 1% of the total protein using this promoter.[122] However, the yield of HcgG obtained in the initial culture condition was 0.2%, which was much lower than this; therefore, it is possible that unknown factors limit the expression of HcgG. The amount of the protein obtained with these conditions was a limiting factor to get pure protein in adequate quantity for the biochemical characterization. In a personal communication, Dr. Kyle Costa suggested that 100 mM Tris buffer in the media could chelate metals required for the overexpression of HcgG.[123–125] In addition, as a past report has shown that sulfur of the [Fe-S]-clusters are derived from S^{2-} in the media and not from cysteine, the sulfide concentration should be kept at a relatively high concentration.[126] Incomplete incorporation of [Fe-S]-cluster may cause unfolding and aggregation of the protein, which can result in a decrease of the protein yield. Considering its toxicity, it is difficult to increase the Na_2S concentration in the batch medium for the formate-dependent methanogenesis. Therefore, I changed

the cultivation method to cultivation in a 10-liter fermenter using a continuous flow of $H_2/CO_2/H_2S$ (80%/20%/0.1%) gas, which could keep a reasonable sulfide concentration without altering the pH. In my experiments, cells under formate grew until OD_{660} = 0.6−0.8, while the cells growing under $H_2/CO_2/H_2S$ grew to OD_{660} = 2−2.5 (Figure 21). This cultivation method also solved the problem of the requirement for the Tris buffer because the culture with H_2/CO_2 does not need the extra buffer.

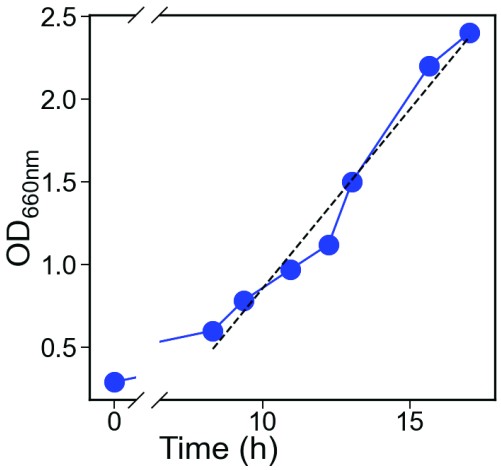


Figure 21. *M. maripaludis* growth under $H_2/CO_2/H_2S$. *M. maripaludis* containing a plasmid for the expression of a His-tagged version of HcgG was grown in a 10-liter fermenter with continuous $H_2/CO_2/H_2S$ (80%/20%/0.1%) gas supply with 500 rpm stirring.

5.3.2. Production and purification of an active form of HcgG

Cells were harvested by centrifugation, suspended in buffer, and frozen at −20 °C. Due to the halophilic nature of the marine microorganism *M. maripaludis*, freezing and thawing of the cells in a low salt buffer disrupted most of the cells. To ensure the complete disruption after freeze and thaw, cells were subjected anaerobically to French press. The membrane fraction, unbroken cells and cell debris were removed by ultracentrifugation, and the extract was loaded onto a HisTrap column and eluted by a linear imidazole gradient (Figure 22a). The fractions from the Ni-affinity purification were concentrated to ~2 mL and loaded onto a Sephacryl S-300 HR gel permeation chromatography column equilibrated with 50 mM MOPS buffer pH 7.4 with 500 mM NaCl and 20% glycerol and eluted in the same buffer. The final yield after gel permeation chromatography was between 7 and 10 mg of total protein from 14 grams of cells (from a 20-liter culture). This amount is approximately 0.5 to 0.7% of the estimated proteins from the harvested cells. Considering the loss of proteins during purification, the content of HcgG is estimated to be ~1% of the cellular protein, which is a reasonable value using this expression system.[122] In the gel permeation chromatography, the protein eluted at 66.5 mL, which corresponds to an apparent molecular mass of 99 kDa, suggesting a dimer structure (calculated monomeric molecular mass = 57 kDa) (Figure 22b and d).

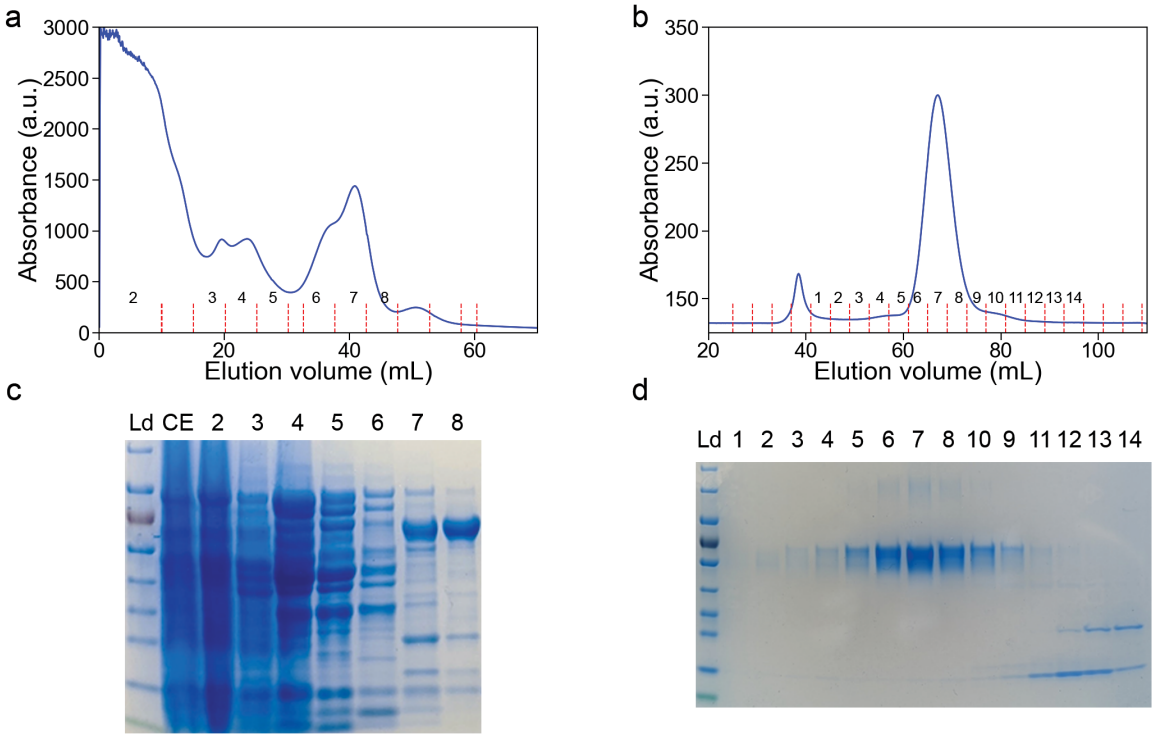


Figure 22. Purification of HcgG from *M. maripaludis* cells grown under $H_2/CO_2/H_2S$. (**a**) Chromatogram of the Ni-affinity chromatography. (**b**) Chromatogram of the gel permeation chromatography. (**c**) SDS-PAGE of the Ni-affinity chromatography. (**d**) SDS-PAGE of the gel permeation chromatography. The lanes of the molecular size marker (Ld) and the cell extract (CE) were indicated. The fraction numbers of the Ni-affinity chromatography are indicated.

The purified protein was dark green to brown in color and presented several features in the UV-visible spectra: a shoulder at 300 nm, a peak at 410 nm and another peak at 600 nm (Figure 23a). The peak at 410 nm suggested the presence of an [Fe-S]-cluster. Incubation with sodium dithionite reduced the absorbance, which indicated the presence of a redox active chromophore (Figure 23b).

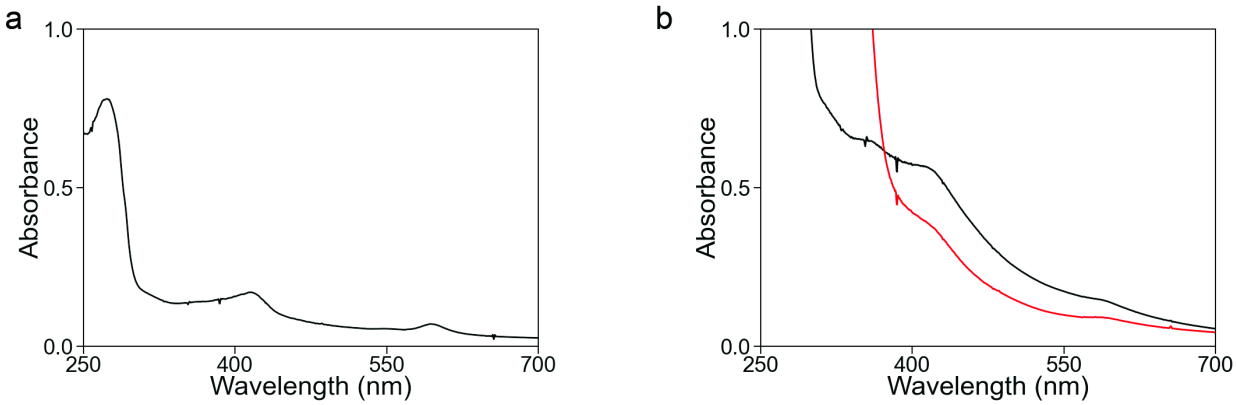


Figure 23. UV-visible spectra of HcgG. (**a**) UV-visible spectrum of 1 mg/ml HcgG as isolated. (**b**) UV-visible spectra of HcgG (2.5 mg/ml) without (black) and with (red) 5 mM sodium dithionite. The absorbance spectra were recorded using 3 mm light pass quartz cuvette under N_2 atmosphere.

5.4. Effects of the deletion of the *hcg* genes in *M. maripaludis*

5.4.1. Optimization of the buffer system in the culture media

To obtain the cell extract from *M. maripaludis* for *in vitro* biosynthesis, cells containing high concentration of the Hcg proteins are required. It can be assumed that the cells exhibiting high Hmd activity could have high concentrations of the Hcg proteins. Production of Hmd in *M. marburgensis* is upregulated when cells are grown under Ni-limitation conditions.[42] *M. maripaludis* shows a similar Ni-dependent behavior.[127] In addition, the *hmd* gene in *M. maripaludis* is regulated by H_2 supply, in which rapid cell growth under H_2-non-limiting conditions was found to increase Hmd production compared to H_2-limiting conditions.[128] Based on these previous findings, Ni-limiting and H_2-non-limiting conditions are the best for protein production in the cells. *M. maripaludis* grows under H_2/CO_2 gas mixture as the energy and carbon source. However, as the Δ*hcg* mutants of *M. maripaludis* are also mutated in the F_{420}-reducing hydrogenase (Frh) genes (see Materials and methods Table 9 for the full genetic background of the strains), the growth of these strains on H_2/CO_2 is impaired; the mutants have a very long lag phase.[72] The other possible substrate of this methanogen is formate. Lie, *et al* showed that the *M. maripaludis* cells grow rapidly in the formate media.[72] We speculated that *M. maripaludis* cells grown rapidly with sodium formate would have high Hmd activity and, accordingly, high levels of Hcg protein. However, one disadvantage of the formate culture is that the pH of the media increases as the cells grow, impairing growth.[41,129] To solve this problem, I added 100 mM Tris buffer to stabilize the pH with less effects on the cell growth (Figure 24).

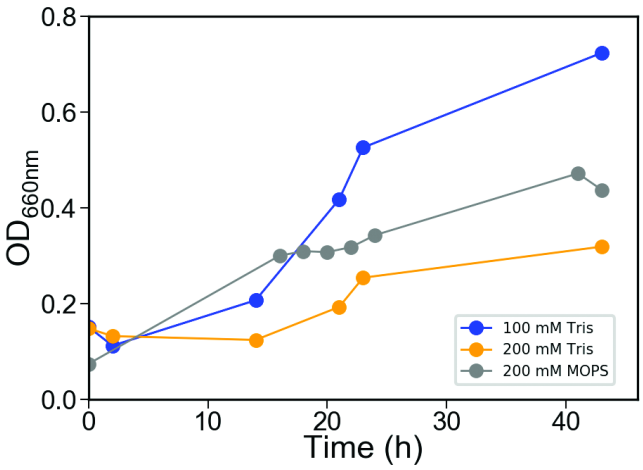


Figure 24. The growth of the *M. maripaludis* S2 strain using formate as the growth substrate in the medium amended with different buffer compounds. Cells were grown in 1-liter bottles containing 500 mL of media at 37 °C. pH was kept at approximately 7.5 by neutralization with formic acid. The standard McCas medium for *M. maripaludis* cultivation contains 200 mM MOPS.[130]

5.4.2. Hmd activity in the Δ*hcg* mutants

The Hmd activity obtained from the Ni-non-limiting formate culture of the wild type S2 strain was ~2.5 U/mg that was higher than the activity observed in the same strain cultivated under the nickel-limiting H_2/CO_2 culture (~1.5 U/mg).[127] I compared the Hmd activity of the wild-type and the mutants strains (Figure 25). Unexpectedly, the Δ*frh* strain, which is the basal strain for all Δ*hcg* and Δ*hmd* strains, showed no Hmd activity. All Δ*hcg* strains lost the Hmd activity except for the Δ*hcgD* strain, which showed 10% activity. The loss of Hmd activity might result from the deletion of the *hmd* or *hcg* genes but it can also be a basic property of the Δ*frh* strain. In the latter case, the loss of Hmd activity in the Δ*frh*, Δ*hcg*, and Δ*hmd* strains contradicts the previous growth phenotype experiments using these mutants, which show substantial differences in the growth properties of the basal Δ*frh* and other strains.[72] This unexpected feature of the Δ*frh* strain is further discussed in Section 5.10.

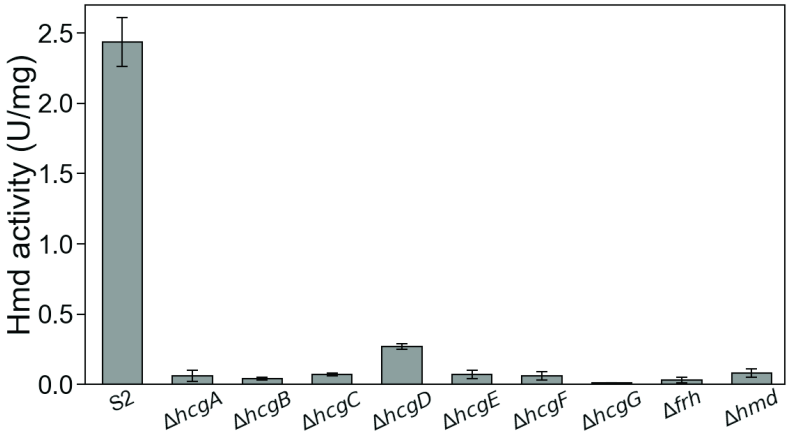


Figure 25. Hmd activity of the wild type and mutants grown in formate medium. Bars represent the standard deviation of two independent measurements. The genotypic features of the strains are described on the abscissa. The genotypes are listed in Table 9.

5.4.3. Proteomic analysis of the Δ*hcg* mutants

To check if the Δ*hcg* mutations caused unexpected changes in the genome while constructing the mutants, I analyzed the size of the encoding regions of the *hcg* genes by PCR. The PCR fragments showed the expected size, which indicated that there are no additional insertions and/or deletions (data not shown). The deletion of certain genes can influence the expression of other genes and/or complete pathways. To study if this is true for the *hcg* genes, we analyzed the total proteome of the strains. Table 4 shows the peptide spectrum matches of the expected proteins in the mutants. HcgF was not detected in any of the samples analyzed; however, because of the similar lag phase to the Δ*hmd* strain,[72] I assume the enzyme must be produced.

Table 4. Peptide Spectrum Matches (PSMs) of the Hmd-related proteins in the proteome of the *M. maripaludis* mutants. The proteins over- or under- expressed compared to the wild type S2 are underlined.

	Strains										
Protein	*S2*	*Δfrh*	*ΔhmdII*	*Δhmd*	*ΔhcgA*	*ΔhcgB*	*ΔhcgC*	*ΔhcgD*	*ΔhcgE*	*ΔhcgF*	*ΔhcgG*
FrhA	72	0	0	0	0	0	0	0	0	0	0
Hmd	161	22	78	0	82	35	35	81	34	30	29
HmdII	31	230	0	24	19	8	8	20	9	7	7
HcgA	5	15	5	7	0	2	2	6	4	2	2
HcgB	5	7	3	4	4	0	0	4	2	3	3
HcgC	23	34	11	13	10	0	0	10	8	7	5
HcgD	5	14	4	4	5	4	3	0	4	3	3
HcgE	3	9	3	3	3	2	2	4	0	2	2
HcgG	13	36	6	10	9	3	4	6	4	3	0

5.4.4. Effect of Δ*hcgB* deletion in the phenotype of the Hcg proteins

In the proteome analysis, Δ*hcgB* and Δ*hcgC* mutants share a common phenotype, where HcgB and HcgC proteins are absent in both mutants. In the genome of most methanogens, both genes are next to each other, and even make a fusion protein in a group of bacteria.[131] The ribosome binding site (5'-GGTG-3')[132] of the *hcgC* gene of *M. maripaludis* locates within the 3' region of the *hcgB* gene. Therefore the deletion of the *hcgB* gene results in the deletion of the ribosome binding site of the *hcgC* gene. This explains the absence of HcgC in the proteome of the Δ*hcgB* mutant (Figure 26). However, in the case of the Δ*hcgC* mutant, no changes happen upstream of the *hcgB* gene. It might be possible that the deletion of the *hcgC* gene has effects on the upstream *hcgB* gene, which inhibits the production of HcgB or destabilize HcgB. Because of the shared phenotype, we further used the Δ*hcgB* mutant with the name of Δ*hcgB*Δ*hcgC* strain for *in vitro* biosynthesis experiments as described below.

Another interesting feature detected in the proteome analysis is that a homolog of Hmd (HmdII), is strongly up-regulated in the Δ*frh* strain. It might be related to the observation of reduction of the Hmd activity in the cell extract of the *M. maripaludis* basal strain (Mm1280) with Δ*frh* genotype (Figure 25). A correlation between the HmdII and Frh encoding genes is discussed in section 5.10.

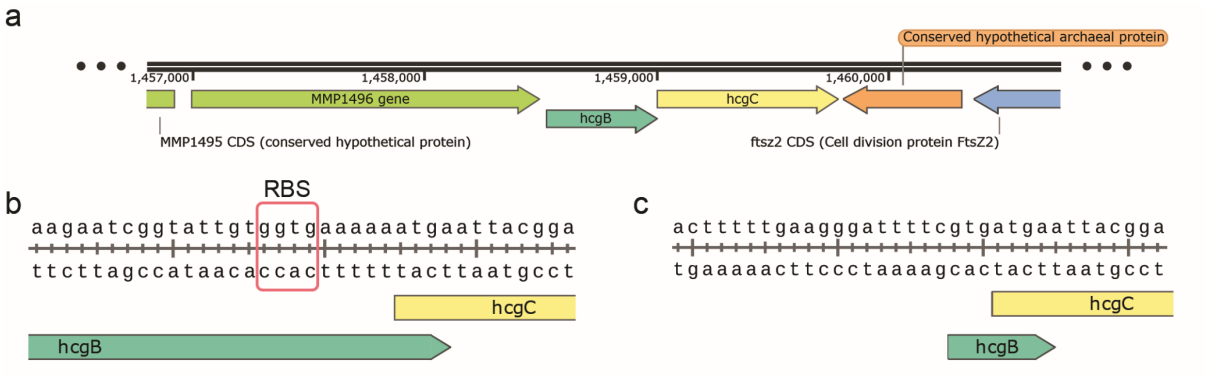


Figure 26. Genome of the wild type and Δ*hcgB* mutant strains. (**a**) Genetic context of the *hcgB* and *hcgC* genes. (**b**) Sequence of the interface region of the *hcgB* and *hcgC* genes, where the ribosome binding site (RBS) (in red open box) locates within the 3'-region of the *hcgB* gene. (**c**) The sequence of the Δ*hcgB* mutant, in which the RBS of the *hcgC* gene was deleted.

5.5. *In vitro* biosynthesis of the FeGP cofactor using the *M. maripaludis* mutants

Figure 27 shows the *in vitro* biosynthesis activity using the cell extract from the *Δhcg* mutants. The *in vitro* biosynthesis activity of the cell extract from the *ΔhcgBΔhcgC* strains was complemented with addition of precursor **3**, as shown in a previous report.[102] These results indicated that HcgB and HcgC catalyze reactions located before formation of precursor **3**, in agreement with the characterized functions of HcgB and HcgC (see Figure 7). In contrast, precursor **3** did not complement *in vitro* biosynthesis using the cell extract from the Δ*hcgE*, Δ*hcgF* and Δ*hcgG* strains. The *in vitro* biosynthesis of Δ*hcgE* and Δ*hcgG* strains were complemented by addition of purified HcgE and HcgG, respectively (the experiment of *in vitro* complementation with HcgG is described in detail in section 5.8.1), which indicates that the inability of *in vitro* biosynthesis using the Δ*hcgE* and Δ*hcgG* strains was caused by lack of the respective Hcg enzymes. In addition, these results indicate that HcgE- and HcgG-catalyzed biosynthesis reactions occur after formation of **3**, which agrees with the predicted reaction scheme of the biosynthesis. On the other hand, *in vitro* biosynthesis using the Δ*hcgF* strain was not complemented by addition of HcgF from *M. maripaludis*; therefore, we need more careful examinations including the activity of the purified HcgF fraction to draw a conclusion. In the case of the experiments using the Δ*hcgD* strain, the small endogenous Hmd activity (Figure 25) could not be strengthened by addition of **3**. HcgD is predicted to be an Fe-trafficking protein, which should function using **3** or its derivative. Therefore, the complementation experiment with the Δ*hcgD* strain is in accordance to the predicted function of HcgD.

The *in vitro* biosynthesis assay using the cell extract from Δ*hcgA* and Δ*hcgG* strains provided information on these two enzymes, whose function has never been indicated by any methods. The *in vitro* biosynthesis assay using the Δ*hcgA* strain was complemented by precursor **3**, which indicated

that HcgA catalyzes a reaction before the formation of **3**. Because biosynthesis of **3** from **1** via **2** is confirmed by the previous enzyme assays,[102] the catalytic function of HcgA is the biosynthesis of **1** or a reaction before the formation of **1**. The *in vitro* biosynthesis using the Δ*hcgG* strain was not complemented by the addition **3.** This result suggests that the catalytic function of HcgG occurs after the formation of **3**. If the predicted functions of HcgE and HcgF are correct, HcgG might catalyze the last reactions for completion of the full FeGP cofactor using the activated thioester-bonded GP on HcgF. The characterization of HcgA and HcgG is described in the following chapters.

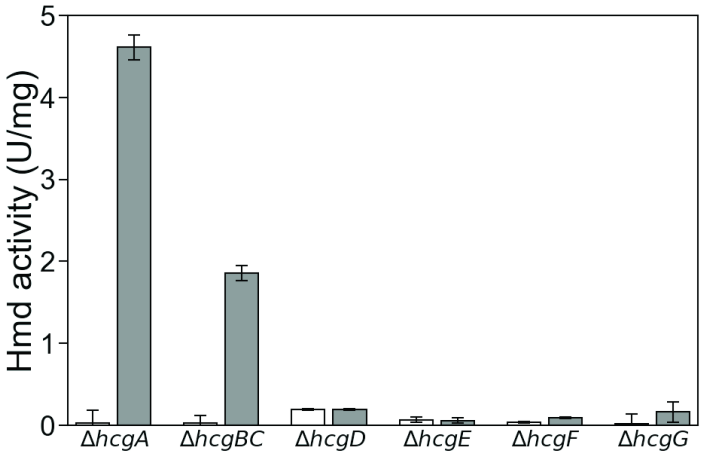


Figure 27. *In vitro* biosynthesis of the FeGP cofactor using the cell extract from the Δ*hcg* mutants. The Hmd activity in the *in vitro* biosynthesis assay was determined. The activity was complemented by addition of precursor **3** only in the test with the Δ*hcgA* and Δ*hcgB*Δ*hcgC* mutants. Left bar (white) shows the *in vitro* biosynthesis activity without addition of any precursors. Right bar (gray) shows the *in vitro* biosynthesis activity with precursor **3**.

5.6. Role of HcgA in the biosynthesis of the FeGP cofactor

5.6.1. 5'-deoxyadenosine-forming reaction of HcgA to screen the substrate

Once the heterologously produced HcgA was purified in a redox-active form, I tried to find its substrate by detecting 5'-deoxyadenosine (dAdo) from SAM. This is based on that radical SAM enzymes share a common half-reaction, the formation of a 5'-deoxyadenosine radical and subsequent conversion to dAdo. The possible substrates can be screened by detection of increased dAdo formation from SAM.[68] In the case of HcgA, no clear reaction was observed (Figure 28). In the meantime, in the process of *in vitro* biosynthesis experiments, I found that HcgA is a catalyst for the formation of precursor **1** from an unknown compound, and suggested several candidates as the substrate by mass spectrometry (See section 5.7 and 6).

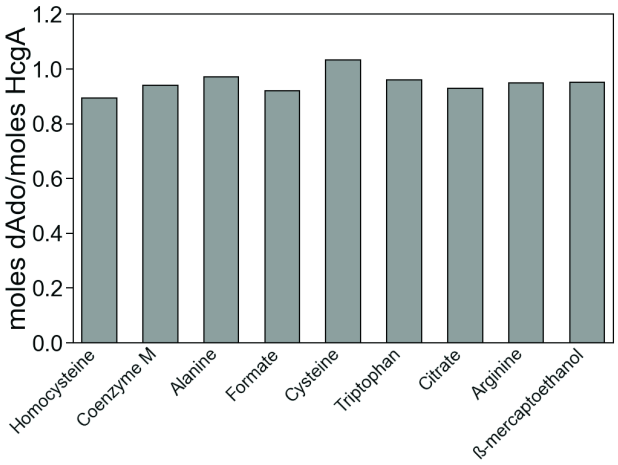


Figure 28. HPLC determination of 5'-deoxyadenosine (dAdo) production. HcgA reaction was measured by dAdo production from SAM.

5.6.2. Characterization of the HcgA-catalyzed reaction using *in vitro* biosynthesis

In *in vitro* biosynthesis of the FeGP cofactor using the cell extract from the Δ*hcgA* mutant, active Hmd was obtained from precursor **3** (Figure 29). Optimization of the conditions of *in vitro* biosynthesis improved the Hmd activity in the assay, which led to a more sensitive assay.[102] To identify the reaction catalyzed by HcgA in the biosynthesis pathway, I tested the chemically synthesized precursors **1** and **2**. Hmd activity was obtained after performing *in vitro* biosynthesis using the Δ*hcgA* cell extract from these precursors. In the case of *in vitro* biosynthesis from **1** and **2**, I added GTP as the substrate of the HcgB reaction into the *in vitro* biosynthesis in addition to the standard assay solution. This finding further confirmed that HcgA plays a role in a step before precursor **1** (Figure 30a).

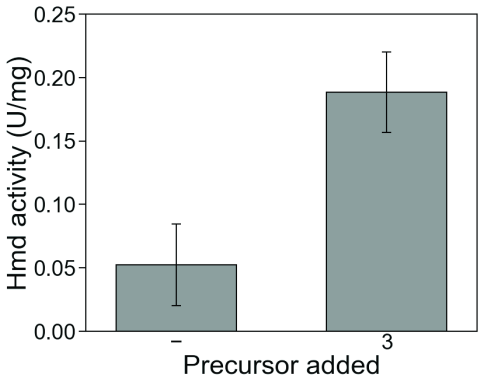


Figure 29. Production of Hmd activity in the *in vitro* biosynthesis assay using the cell extract of the *M. maripaludis* Δ*hcgA* strain without (−) and with precursor **3** (3). These experiments were performed under the conditions before optimization; therefore, the activity was lower than the values shown in Figure 30.

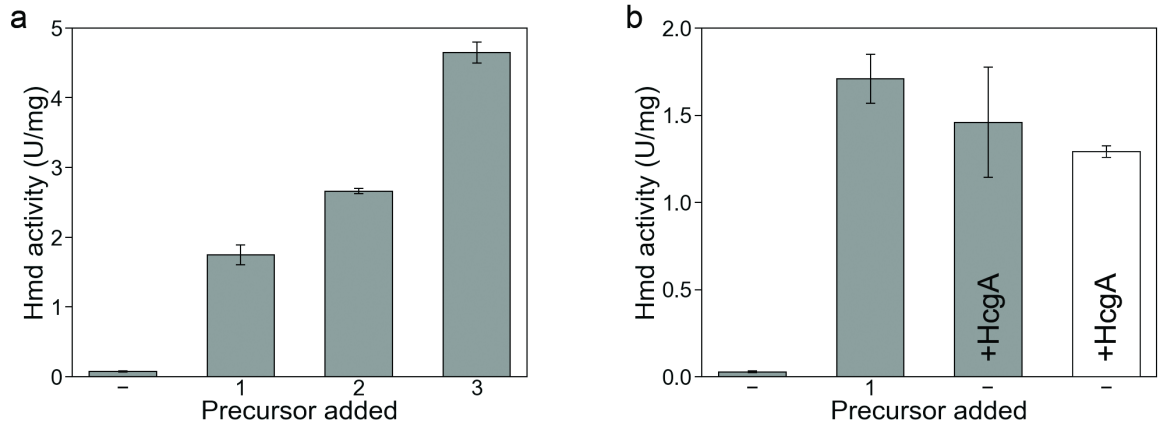


Figure 30. *In vitro* biosynthesis using cell extract from the *M. maripaludis* Δ*hcgA* strain. (**a**) Hmd activity obtained in *in vitro* biosynthesis from precursors **1**, **2** and **3**. The precursor added are described on abscissa. A negative control did not contain any precursors (-). (**b**) Complementation of *in vitro* biosynthesis. The *in vitro* biosynthesis was complemented by addition of precursor **1** (1) or no precursor (-) and 20 µM purified HcgA (indicated in the bars). The *in vitro* biosynthesis in the presence of HcgA was not changed by removal of SAM from the *in vitro* biosynthesis assay (open bar).

I could also indicate the enzymatic activity of the purified HcgA by complementation of *in vitro* biosynthesis with purified HcgA. The *in vitro* biosynthesis was performed in the absence of any precursor and in the presence of 20 µM HcgA in the assay with the cell extract from the Δ*hcgA* strain. A similar activity was obtained both from the assay complemented with either precursor **1** or the purified HcgA. This result indicated that the heterologously produced and purified HcgA is active and that the substrates of the HcgA reaction for the formation of **1** are present and accumulated in the cell extract in enough quantities to produce approximately 2 µM of FeGP cofactor (Figure 30b). Additionally, we tested whether removing SAM from the assay could impact the activity; however, removal of exogenous SAM had only a minor impact on the amount of activity produced. A possible explanation is that the cell extract already has enough SAM to produce precursor **1** or possesses an active methionine adenosyltransferase that could reproduce SAM from methionine and ATP.

5.6.3. The HcgA catalyzed reaction using the protein-free extract

Considering the comparative genomic results, HcgA-G proteins are the only enzymes responsible for biosynthesis of the FeGP cofactor, this restriction suggests that HcgA should catalyze the formation of precursor **1**. To test this hypothesis, a protein-free extract of the Δ*hcgA* strain was prepared by filtrating the cell extract through a 3 kDa membrane. The protein-free extract was then incubated with HcgA, SAM and sodium dithionite as previously described for other radical SAM enzymes.[66,87,93]

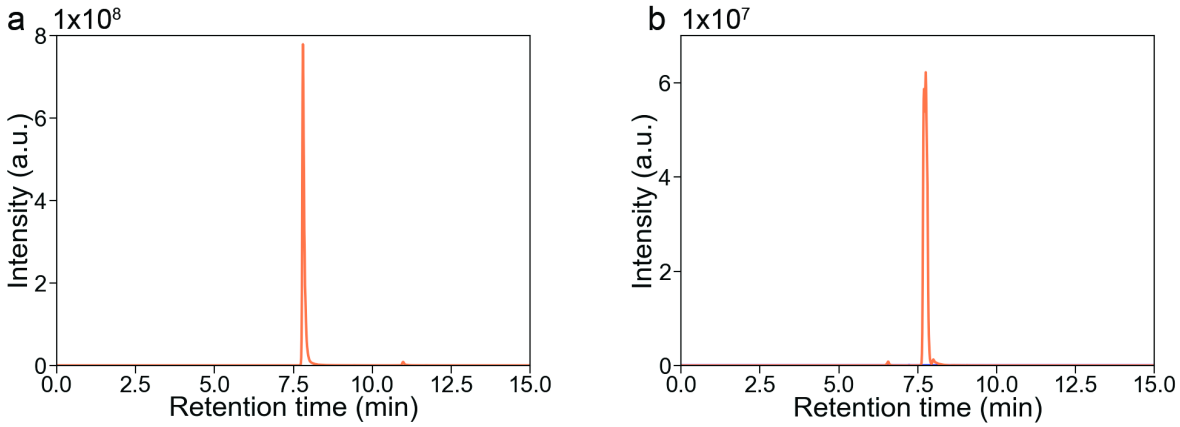


Figure 31. HPLC-MS of HcgA reaction product. (**a**) Chromatogram of the 184.06 *m/z* chemically synthesized compound **1**. (**b**) HcgA reaction product in *in vitro* biosynthesis using the protein-free extract from the Δ*hcgA* strain. The 184.06 *m/z* product was not observed in the reaction solution without HcgA.

HcgA was found to produce a 184.0605 *m/z* compound with the same retention time as the authentic standard precursor **1** (Figure 31). This indicated that the observed compound is precursor **1** and HcgA is responsible for the formation of precursor **1** from compound(s) in the protein-free extract. Additionally, it shows that no other protein is needed to form this compound from the accumulated unknown compound(s) in the cell extract.

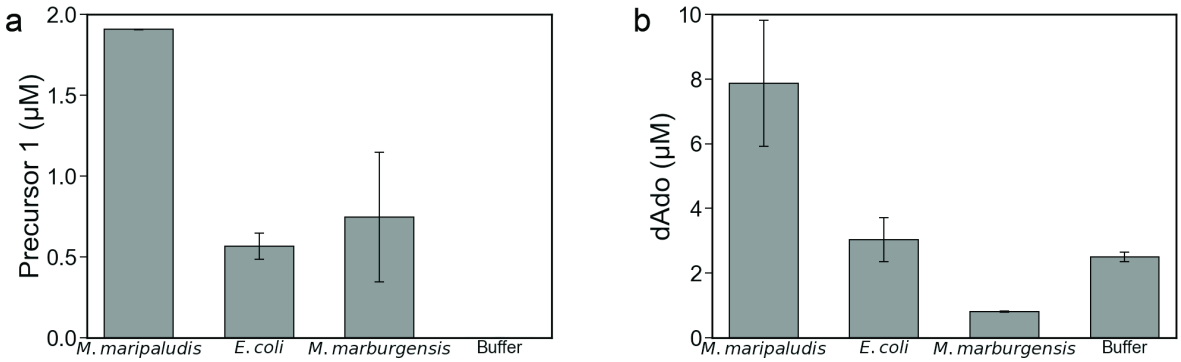


Figure 32. Determination of the reaction products catalyzed by HcgA. The reaction was performed in the presence of the protein-free extract from the *M. maripaludis* Δ*hcgA* strain (*M. maripaludis*), *E. coli* BL21 (*E. coli*), *M. marburgensis* (*M. marburgensis*) or only the reaction buffer solution (Buffer). Production of precursor **1** (**a**) and 5'-deoxyadenosine (**b**) were measured by HPLC-MS. Error bars correspond to the standard deviations of three independent measurements.

The formation of precursor **1** was observed in an HcgA-dependent manner from the protein-free extracts obtained from *E. coli* and *M. marburgensis* in addition to the *M. maripaludis* Δ*hcgA* strain (Figure 32). It is likely that the substrate or substrates are common molecules shared between bacteria and archaea. The effect of the addition of the protein-free extract on the production of dAdo was also tested (Figure 33). The results indicated that the addition of the protein-free extract increased the amount of dAdo formed in the assay and that the substrate(s) in the protein-free extract is converted using the radical SAM reaction of HcgA. To further confirm the production of dAdo along with

precursor **1**, a time-dependent study of the HcgA reaction was performed. Both compounds increased in a time-dependent fashion (Figure 34) although not stoichiometric manner.

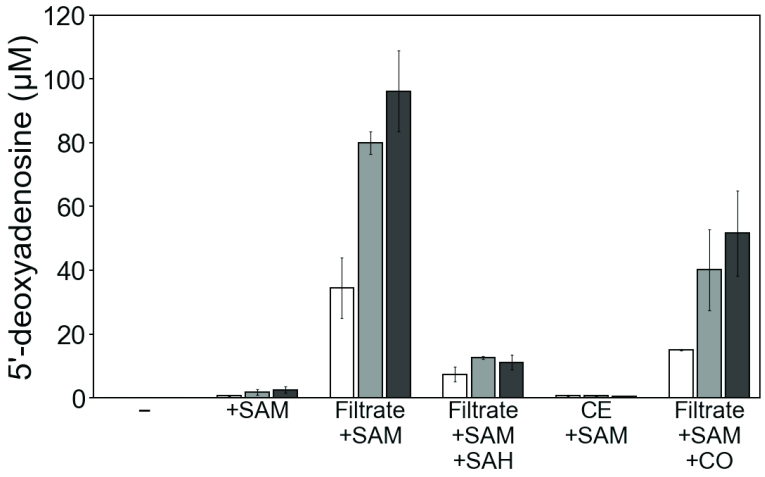


Figure 33. HPLC-MS quantification of 5'-deoxyadenosine (dAdo) after incubation with HcgA. Incubation was performed for 0 h (open bar), 1 h (light gray bar) and 6 h (black bar) in the presence of SAM, the cell extract (CE) of the Δ*hcgA* strain, its 3-kDa filtrate (filtrate) and/or SAH. The experiments were performed under 95% N_2 /5% H_2 atmosphere except for one condition with the filtrate and SAM under 47.5% N_2/2.5% H_2/50% CO atmosphere. Error bars correspond to the standard deviations of three independent measurements.

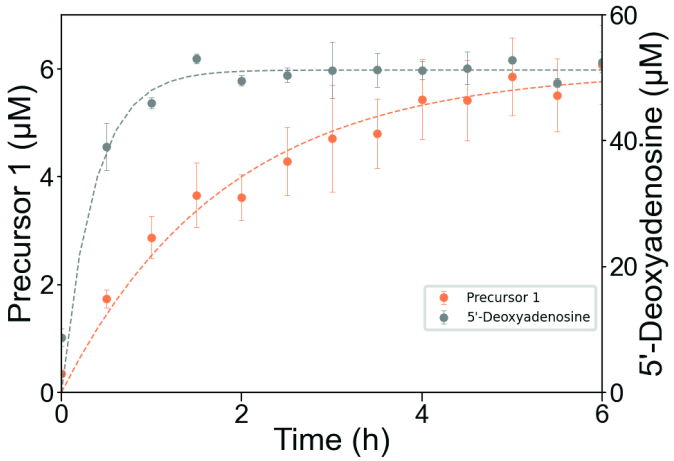


Figure 34. Time-dependent formation of the HcgA reaction products, precursor **1** and 5'-deoxyadenosine (dAdo). Dashed lines represent the best fit to an exponential function. Bars represent the standard deviation of three separate measurements.

To confirm the HcgA reaction product, MS/MS analysis was performed, which indicated the same signals as those observed in the MS/MS analysis of chemically synthesized authentic precursor **1** (Figure 35).[74,78] Furthermore, I checked the conversion of the HcgA reaction product by using HcgC. HcgC is the enzyme responsible for the formation of precursor **2** from precursor **1**. HcgA produced a product of 184.06 *m/z* corresponding to precursor **1**. The subsequent reaction with HcgC converted the HcgA reaction product to a compound with a 193.07 *m/z*, which corresponds to the signal of precursor **2** (Figure 36). This experiment enzymologically supported the nature of the HcgA product as precursor **1**.

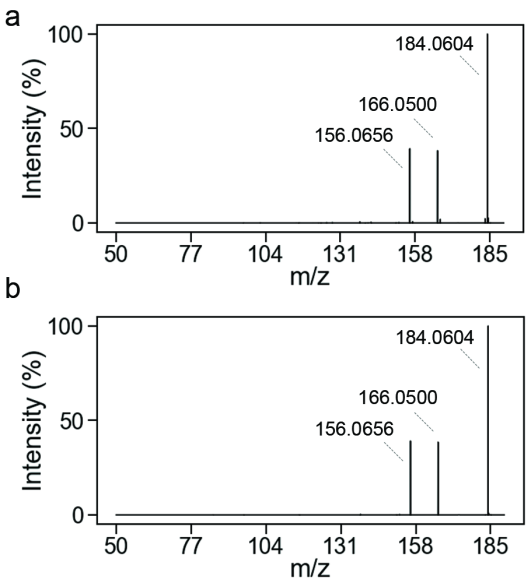


Figure 35. ESI-MS/MS analysis of the compound with 184.0604 *m/z* produced by the reaction with HcgA (**a**) and the chemically synthesized precursor **1** (**b**).

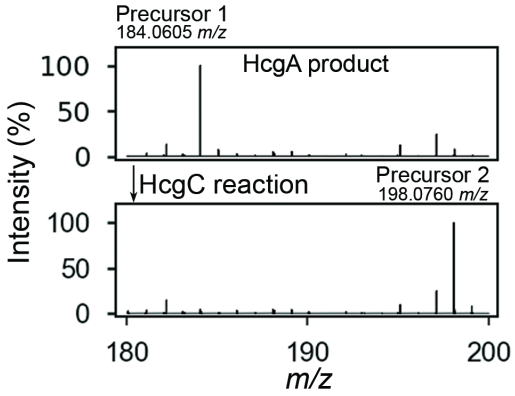


Figure 36. Confirmation of the nature of the HcgA reaction product by the HcgC reaction. HcgA reaction was performed overnight in the absence or presence of HcgC. The reaction with only HcgA produced a 184.06 *m/z* product that corresponds to precursor **1**. The addition of HcgC lead to a formation of a peak of 193.07 *m/z* that corresponds to precursor **2**.

A characteristic feature of the radical SAM superfamily is that the enzymes are inhibited by products and product analogues, such as dAdo, SAH, methionine. To determine the inhibitory effects, the activity of HcgA that produces precursor **1** was measured in the presence of the potential inhibitors (Figure 37). The removal of SAM did not abolish the activity probably because of the presence of SAM in the protein-free extract. Further addition of SAH to the sample without SAM revealed a significant inhibition by SAH. None of the other products inhibited the reaction at concentrations up to 1 mM. It is known that some radical SAM enzymes are resistant to some potential inhibitors.[89,133] These results show for the first time that HcgA is responsible for the production of precursor **1** even though CO production was considered previously from the similarity to HydG.[3,81,83]

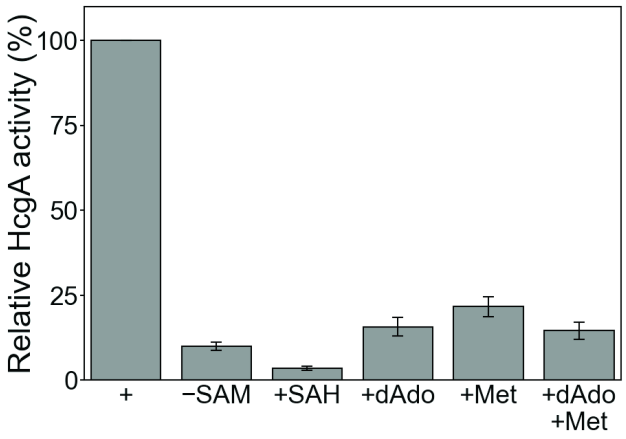


Figure 37. Inhibition of the HcgA reaction. The HcgA activity that produces precursor **1** was measured. The HcgA activity without externally added SAM (−SAM) was compared with the positive control, full reaction (+). On the assay without SAM, the effect of the 1 mM potential inhibitors: *S*-adenosyl homocysteine (+SAH), 5'-deoxyadenosine (+dAdo), methionine (+Met) and both methionine and 5'-deoxyadenosine (+dAdo +Met) were tested. Error bars represent the standard deviation of three MS measurements.

In a previous study, methanogenic archaea were cultivated in media containing ^{13}C- or ^{2}H-labelled substrates, and it was found that five carbons of the pyridinol ring are derived from acetate, two from pyruvate, and one from CO_2.[71] Based on the labelling pattern, the authors proposed that the pyridinol ring is formed by condensation of a ß-alanine or aspartate with a five-carbon sugar-like compound. To test this hypothesis, I performed the HcgA reaction with a mixture of twenty essential [$^{15}N_1$]-labelled amino acids, or [$^{13}C_3$,$^{15}N_1$]-β-alanine. Because these compounds could be an indirect precursor of the HcgA reaction substrate, the cell-free extract was first incubated with the labelled compounds for 2 hours and then the HcgA assay mixture was prepared. In addition, considering the potential requirement of the second substrate from the cell extract, we also tested the HcgA reaction with the protein-free extract with the labeled compounds. MS analysis of the precursor **1** measured after the reaction showed no increase in the mass outside the natural isotopic occurrence (Figure 38 and Figure 39).

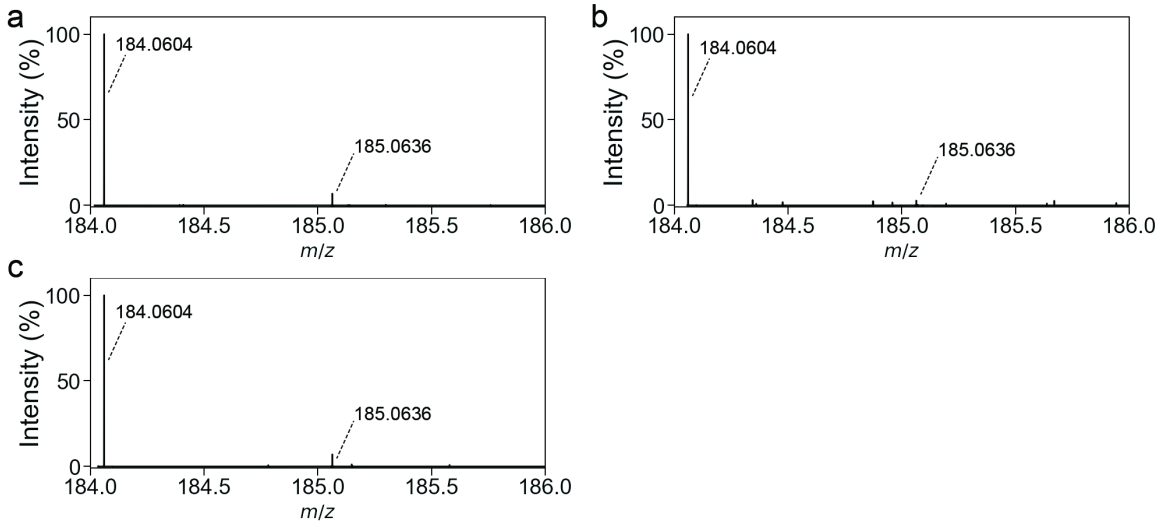


Figure 38. Precursor **1** produced by the HcgA reaction with ^{15}N-labelled amino acids mixture. MS spectra at the retention time of precursor **1** on HPLC incubated without labelled compounds (**a**), with [$^{15}N_1$]-labelled amino acids and the cell extract filtrate (**b**), or with [$^{15}N_1$]-labelled amino acids after pre-incubation in the cell extract (**c**). The 184.060 *m/z* signal corresponds to precursor **1**. The 185.063 *m/z* and 185.057 *m/z* signals correspond to [$^{13}C_1$]-precursor **1**, and [$^{15}N_1$]-labelled precursor **1**, respectively.

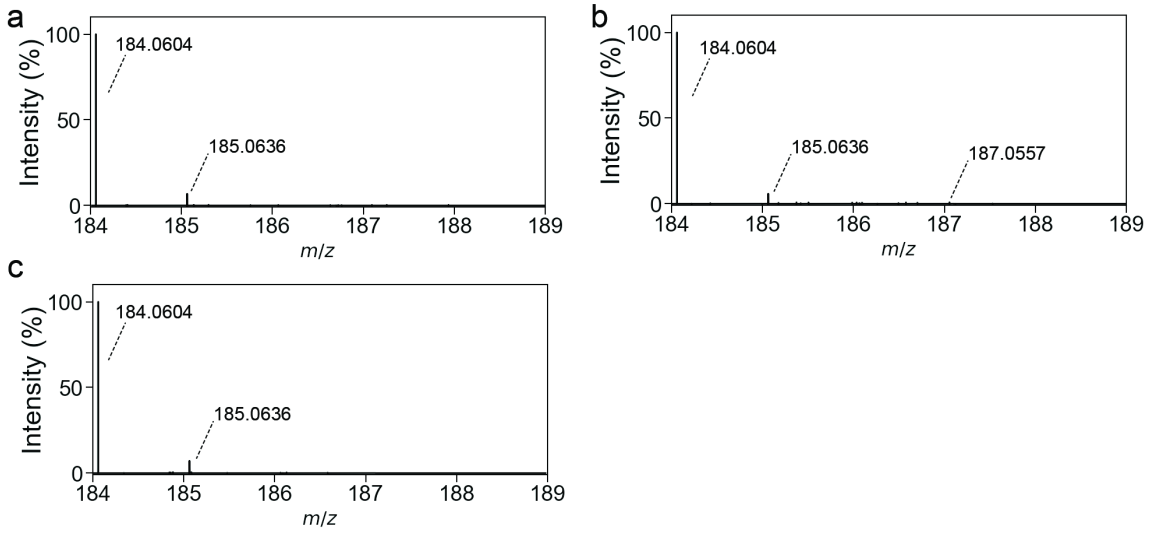


Figure 39. Precursor **1** produced by the HcgA reaction with ^{13}C,^{15}N-labelled ß-alanine. No different masses compared to the non-labelled reactions were found. A full incorporation of labelled [$^{13}C_3$,$^{15}N_1$]-ß-alanine should yield a 188.068 *m/z*. The 185.063 *m/z* corresponds to the natural abundance of [$^{13}C_1$]-Precursor **1**. MS spectra at the retention time of precursor **1** on HPLC incubated without labelled compounds (**a**). The HcgA reaction product with [$^{13}C_3$,$^{15}N_1$]-ß-alanine and the cell extract filtrate (**b**), and that with [$^{13}C_3$,$^{15}N_1$]-ß-alanine pre-incubated with the cell extract (**c**) were tested. None of the conditions increased the isotopic distribution compared to the non-labelled condition (**a**).

5.7. Search of the possible substrate(s) of the HcgA reaction

5.7.1. Chromatographic fractionation

To screen the possible substrate compounds of HcgA, I fractionated the protein-free extract. The protein-free extract was adjusted to the physiological pH of 7.4 and loaded onto an anion exchange HiTrap Q HP column or to a cation exchange SP XL column and eluted with a linear gradient of NaCl. By using cation- or anion-exchange chromatography, several compounds from the protein-free extract were fractionated and the HcgA reaction producing precursor **1** was analyzed by mass spectrometry. From the analysis of the anion and cation exchange chromatography, the substrate(s) were not found in the eluted fractions, instead they were found mostly in the flow through.

5.7.2. Untargeted MS analysis

Because the HcgA assay shows a time-dependent increase in precursor **1**, all detectable compounds in the protein-free extract were analyzed by untargeted MS analysis in a time-dependent manner to determine the correlation of the decrease in compounds to the increase in **1**. Because the behavior of the compounds in the cell extract was variable, I analyzed the trends of these compounds by fitting a linear regression and comparing the slopes (Figure 40). Although a linear regression does not fit the obtained data perfectly, it is a reasonable and fast approximation to compare all metabolites detected by MS.

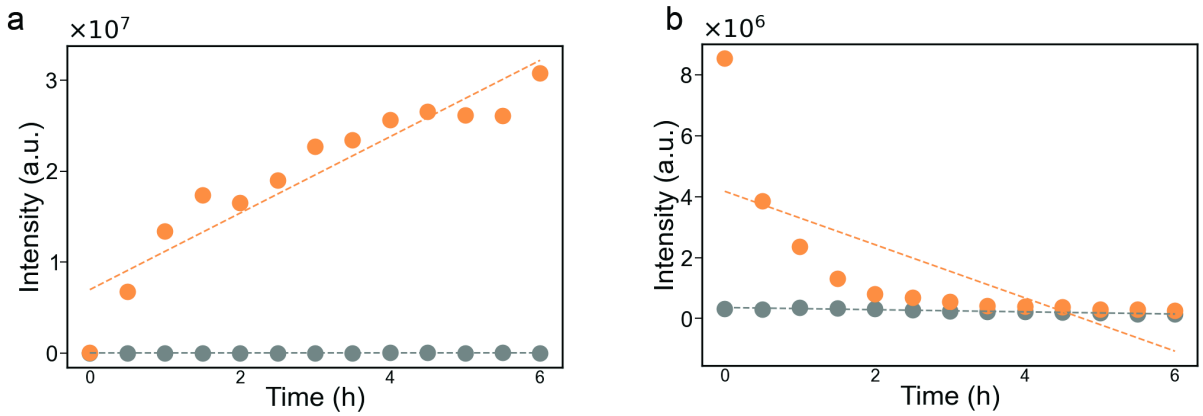


Figure 40. Example of an increased (**a**) and a decreased (**b**) compound in HcgA reaction. The dashed lines represent the calculated linear fit used for the comparisons.

The protein-free extracts from *M. maripaludis* and *E. coli* were reacted with 20 µM HcgA for 6 hours and samples were taken each 30 minutes. After quenching and extraction with 80% methanol, samples were analyzed by HPLC-MS. A reverse-phase chromatography was performed and analyzed in tandem with an Orbitrap in negative mode. Over 1100 compounds were identified in an *m/z* window of 50 to 500. The intensities over time of the compounds were adjusted to a linear regression as stated. All compounds with a positive slope (threshold $m_{reaction}/m_{control} > 20$) are shown in Figure 41 and

additional information such as retention time, calculated mass, mass error can be found in section 8.3. All compounds that decrease threshold $m_{reaction}/m_{control} < -10$) are shown in Figure 42 and the detailed information in section 8.4. Because the substrate(s) must be a common molecule between *M. maripaludis* and *E. coli*, I used the *E. coli* metabolome database (https://ecmdb.ca/[134]) to identify the compounds by comparing the calculated masses of the detected compounds (with an error of 10 ppm) (Table 5 for increasing compounds and Table 6 for decreasing compounds). dAdo production was observed by this method. This approach is systematic and comprehensive but no compound was directly identified as a substrate. However, some compounds detected deserve special attention, which are pyridoxal and pyridoxine. These compounds could be either an *in situ* fragmentation of pyridoxal-5'-phosphate, which is a cofactor of other radical SAM enzymes.[135] The total number of increasing and decreasing compounds is over 100. As the identification is limited by the databases (ChemSpider and ECMDB), several calculated masses could not be identified. If it is possible to decrease the amounts of the candidate compounds by the other methods, this could be a powerful method to identify a substrate.

Table 5. Increased compounds identified in the *E. coli* metabolome database (ECMDB).[134] Index represents the position of the compounds in the full list (See Supplementary information)

Name	Molecular weight (Da)	Retention time (min)	Index
Adenine	135.0543	3.14	20
5'-Deoxyadenosine	251.10118	3.14	30
Precursor 1	183.05284	4.61	92
4-Hydroxyphenylacetaldehyde	136.05116	4.1	241
p-Aminobenzoic acid	137.04746	4.62	549
Pyridoxine	169.07362	4.61	557
Succinic acid	118.02761	3.15	583
Pyridoxine	169.0736	3.11	616
L-Homocysteine	135.0354	3.14	937

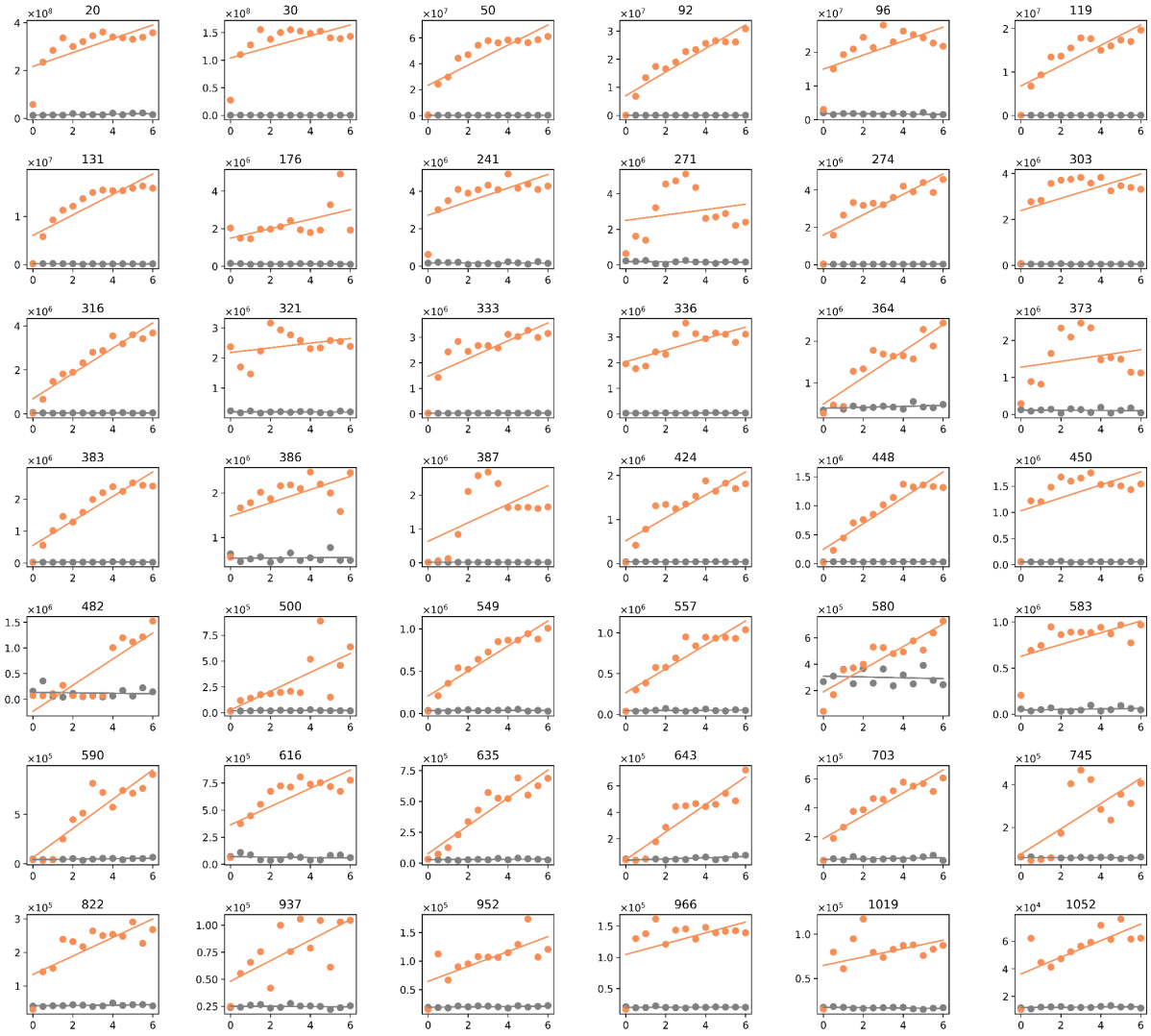


Figure 41. Trend of the MS intensity of selected compounds that increase over time in the HcgA reaction. The HcgA reaction with the protein-free extract from *M. maripaludis* Δ*hcgA* strain was tested. The abscissa shows time in hours, and the ordinate MS Intensity. Each subplot title shows the index number from the complete list (See Supplementary information). Compounds are plotted as the average of triplicate measurements, with the non-enzymatic control in gray and the HcgA reaction in orange. Lines represent the linear regressions used for the comparison, threshold $m_{reaction}/m_{control} > 20$.

Table 6. Decreasing compounds identified in the *E. coli* metabolome database (ECMDB).[134] Index represents the position of the compounds in the full list (See Supplementary information).

Name	Molecular weight (Da)	Retention time (min)	Index
3,4-Dihydroxy-L-phenylalanine	197.06823	5.28	65
Triethanolamine	149.10493	1.50	70
Pipecolic acid	129.07877	1.75	103
Iminobutyrate	129.07877	1.75	103
3-Acetamidobutanal	129.07877	1.75	103
Glycerol	92.04727	1.59	190
Pyridoxine	169.07364	5.01	281
p-Aminobenzoic acid	137.04749	5.01	289
Pyridoxine	169.07358	5.28	414
p-Aminobenzoic acid	137.04748	5.28	574
1,6-Anhydro-N-acetylmuramate	275.0999	4.93	985
Phe-Glu	294.12124	5.01	25
Pro-Ser	202.09514	1.80	396

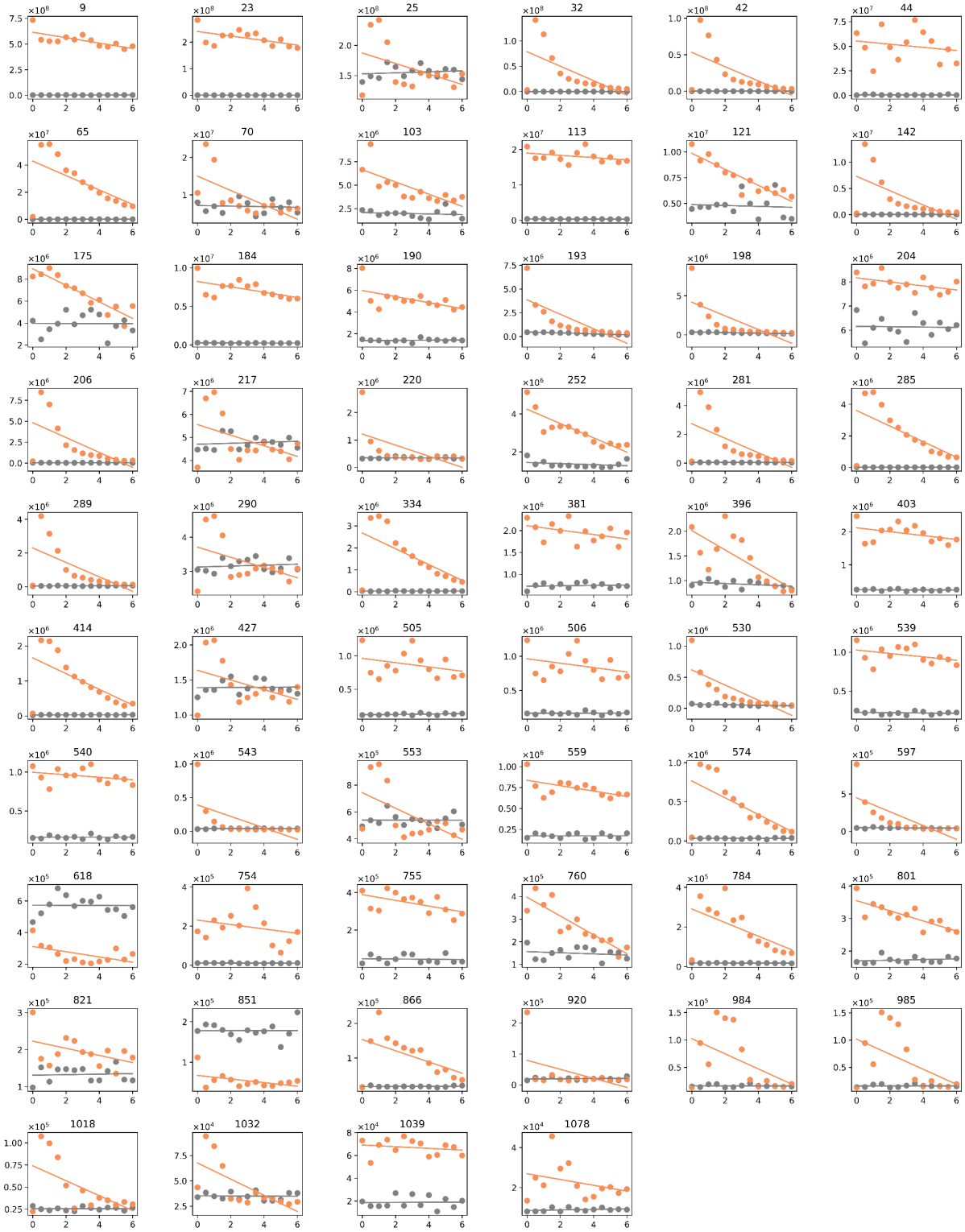


Figure 42. Trend of the MS intensity of selected compounds that decrease over time in the HcgA reaction. The HcgA reaction with the protein-free extract from *M. maripaludis* Δ*hcgA* strain was tested. The abscissa shows time in hours, and the ordinate MS Intensity. Each subplot title shows the index number from the complete list (See Supplementary information). Compounds are plotted as the average of triplicate measurements, with the non-enzymatic control in gray and the HcgA reaction in orange. Lines represent the linear regressions used for the comparison, threshold $m_{reaction}/m_{control} < -10$.

5.8. Role of HcgG in the biosynthesis of the FeGP cofactor

5.8.1. *In vitro* biosynthesis using purified HcgG

All Hcg proteins except HcgG have been biochemically and/or structurally characterized. As described above, in this PhD project, I identified HcgA as a radical SAM enzyme responsible for the production of precursor **1**. The organic backbone of the FeGP cofactor is biosynthesized by the reaction of the previously described enzymes; however, the key steps in its biosynthesis have not been clarified, which are the formation of the CO-ligands, the formation of the Fe-center, the reduction of the carboxylic acid to an acyl-ligand, the assembly of the full cofactor, and the delivery to the [Fe]-hydrogenase apoenzyme. In the section 5.5, I described the results of *in vitro* biosynthesis using the *M. maripaludis* Δ*hcgG* strain with precursor **3** and speculated that the HcgG functions for completion of the full FeGP cofactor using the activated thioester-bonded GP on HcgF. The updated proposed biosynthetic pathway is shown in Figure 43.

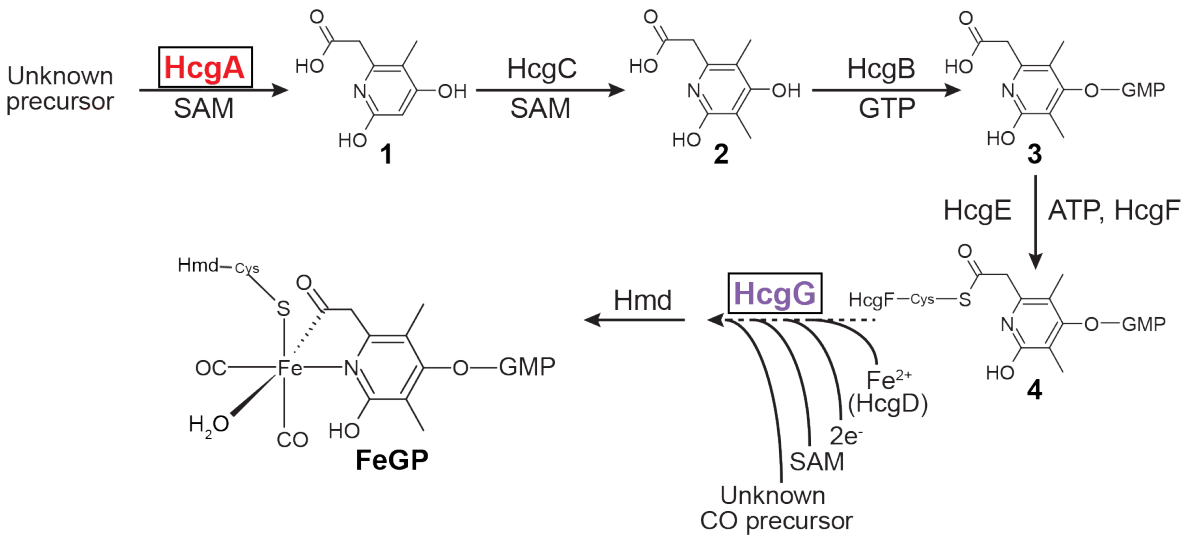


Figure 43. Proposed pathway for the biosynthesis of the FeGP cofactor including new information in this work. The enzyme reactions discussed in this work are highlighted: HcgA (red) produced the precursor **1** and HcgG (purple) is possibly involved in the last step.

Because HcgG appears to be involved in the last step of the biosynthesis sequence, combining the cell extract of strains with each gene deleted could complement each other, and produce active FeGP cofactor. Combining the cell extract from the two strains *ΔhcgA* and Δ*hcgG* did not produce Hmd activity by themselves; however, adding the incubation mixture (Fe^{2+}, Mg^{2+}, ATP, GTP, SAM, DTT, sodium dithionite, under H_2/CO atmosphere) produced Hmd activity without any added precursor (Figure 44). This indicated that the *ΔhcgA* and Δ*hcgG* cell extracts complement each other and provide the required enzymes to synthesize the FeGP cofactor, though they need additional small compounds included in the standard assay solution as co-substrates.

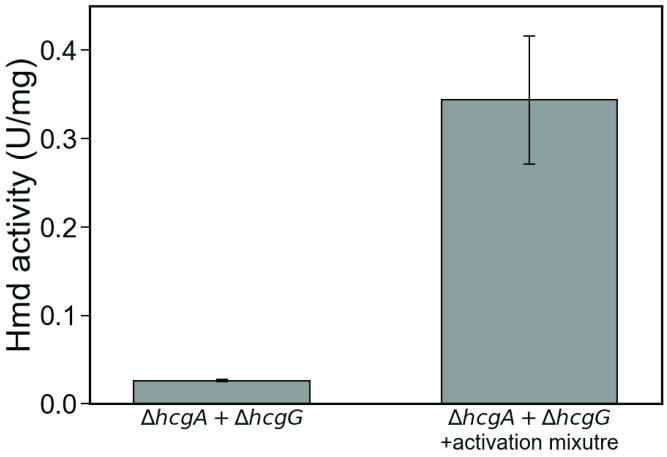


Figure 44. Biosynthesis of the FeGP cofactor in *in vitro* biosynthesis using the *M. maripaludis* Δ*hcgG* and Δ*hcgA* strains. *In vitro* biosynthesis performed using the mixed cell extracts from the Δ*hcgA* and Δ*hcgG* strains with the *in vitro* biosynthesis standard mixture composed of Fe^{2+}, Mg^{2+}, ATP, GTP, SAM, DTT and sodium dithionite under 50% H_2/ 50% CO (Δ*hcgG* + Δ*hcgG* + activation mixture) and without reaction mixture under N_2 (Δ*hcgA* + Δ*hcgG*).

To characterize the function of HcgG, an active form of HcgG was needed. Dr. Kyle Costa (University of Minnesota) generated a *M. maripaludis* strain that is manipulated for the production of a 6xHis-tagged version of HcgG. Initially, cells were grown in the formate medium as described previously for obtaining the *hcg* mutants' cell extract.[102] The protein was loaded onto a Ni-affinity chromatography column and eluted with a linear imidazole gradient. The protein with the size of 55 kDa eluted in two different peaks from the Ni-affinity column as shown before (Figure 20a) and the proteins in the two peaks were collected separately. SDS-PAGE indicated the presence of several smaller molecular weight bands (Figure 20b). Although the yield was low, the amount was enough to perform initial *in vitro* biosynthesis assays. The *in vitro* biosynthesis experiments indicated that the purified HcgG from the two fractions complemented the *in vitro* biosynthesis of the FeGP cofactor with the Δ*hcgG* strain cell extract from precursor **3** (Figure 45). The fractions did not show any Hmd activity, which indicated that Hmd is not coeluted with HcgG, although I speculated that HcgG might have a binding affinity to Hmd from the predicted function of HcgG.

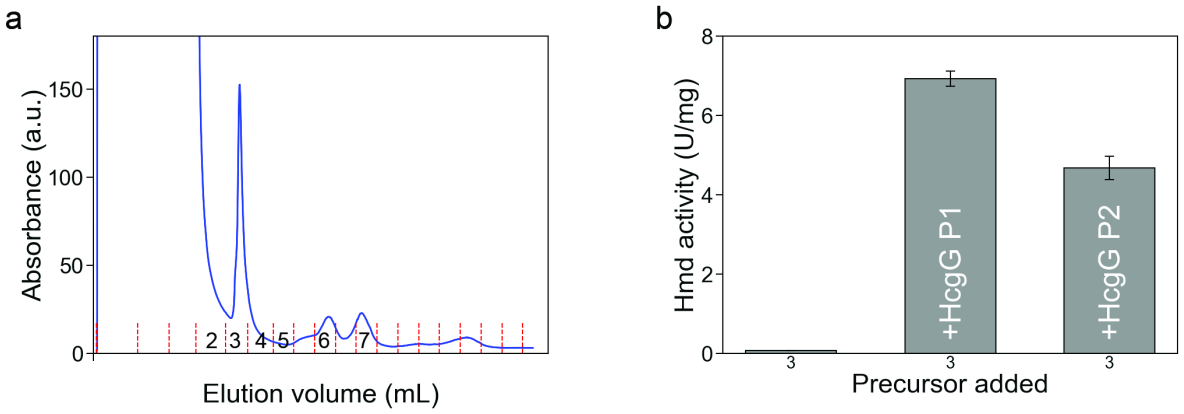


Figure 45. Complementation of the *in vitro* biosynthesis assay by addition of the purified HcgG. *In vitro* biosynthesis using the *M. maripaludis* Δ*hcgG* strain cell extract from precursor **3** under the standard condition with 50% H_2/50% CO was complemented with the protein in each fraction from the Ni-affinity chromatography. The fraction 6 and 7 in the chromatography in panel **a** correspond to HcgG P1 and HcgG P2 added in *in vitro* biosynthesis in panel **b**, respectively.

To improve the yield of HcgG from the *M. maripaludis* cells, I changed the cultivation method to fermentation with $H_2/CO_2/H_2S$ continuous flow as described in section 5.3, which led to higher yields of HcgG, where the protein eluted as a single peak in Ni-affinity chromatography. To remove the contaminated proteins in the Ni-affinity chromatography fraction, the fraction was further purified using gel permeation chromatography (Figure 22), which slightly reduced the *in vitro* biosynthesis activity (Figure 46). The reduction of activity could be attributed to a slight inactivation of HcgG due to the overnight storage time at room temperature after the affinity chromatography. However, I thought that removal of the contaminated protein could also reduce the HcgG activity; therefore, I performed a proteome analysis of the contaminated proteins. The data indicated that the contaminated proteins are IMP dehydrogenase, [NiFe]-hydrogenase maturase HypB, and the 30S ribosomal protein 53ae. It is unlikely that these proteins influence the biosynthetic activity of HcgG.

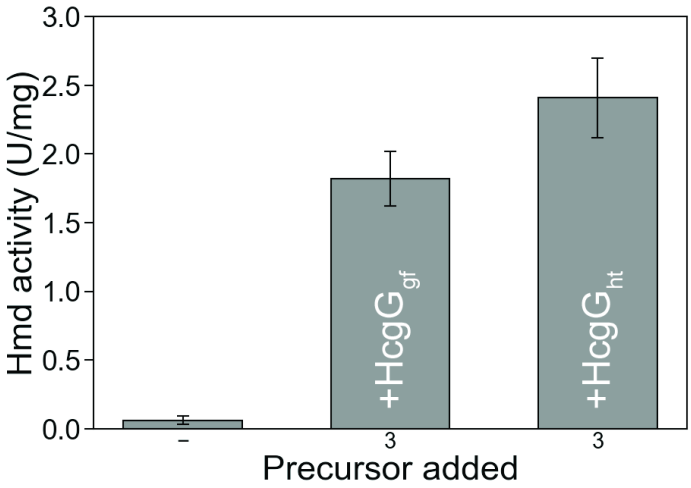


Figure 46. *In vitro* biosynthesis using cell extract from the *M. maripaludis* ΔHcgG strain with HcgG purified by Ni-affinity column ($HcgG_{ht}$) and further by gel permeation column ($HcgG_{gf}$) with (3) and without (–) precursor **3**.

5.8.2. Variable activities in *in vitro* biosynthesis using different batches of the cell extract

The *in vitro* biosynthesis activity of the FeGP cofactor was highly reproducible when the same batch of the cell extract from *M. maripaludis* was used. In contrast, in the assays using different batches of the cell extract, the deviation of the *in vitro* biosynthesis activity was large. Some of the assays produced up to 22 U/mg cell extract protein in the assay, which is equivalent to 210 U/mL assay solution (Figure 47c and d). In this *in vitro* biosynthesis activity assays, the specific activity of Hmd in the assay is ~500 U/mg, which is fully active. As this assay contains 10 μM apoenzyme, I can estimate that all the precursor **3** added (10 μM) in the assay was converted to the FeGP cofactor. In contrast, I sometimes observed a very low activity, for example, 3 U/mg (58 U/mL) that corresponds to a 30% conversion of precursor **3** to the FeGP cofactor (Figure 47a). Therefore, a series of *in vitro* biosynthesis experiments were routinely performed using the same batch of the cell extract. Otherwise, the values cannot be compared. The high variability of the assay could be attributed to the quality of the cell extract rather than the concentration of the extract because the final activity is normalized by the protein concentration in each assay. The condition of the cells prior to harvesting may alter the content of unknown precursors and electron carriers in the cell extract. Dilution of the cell extract dramatically decreased the *in vitro* biosynthesis activity in a non-linear fashion;[103] therefore, I could not dilute the cell extract to unify the activity. This dilution effect might be correlated to the variability of the activity of the cell extract.

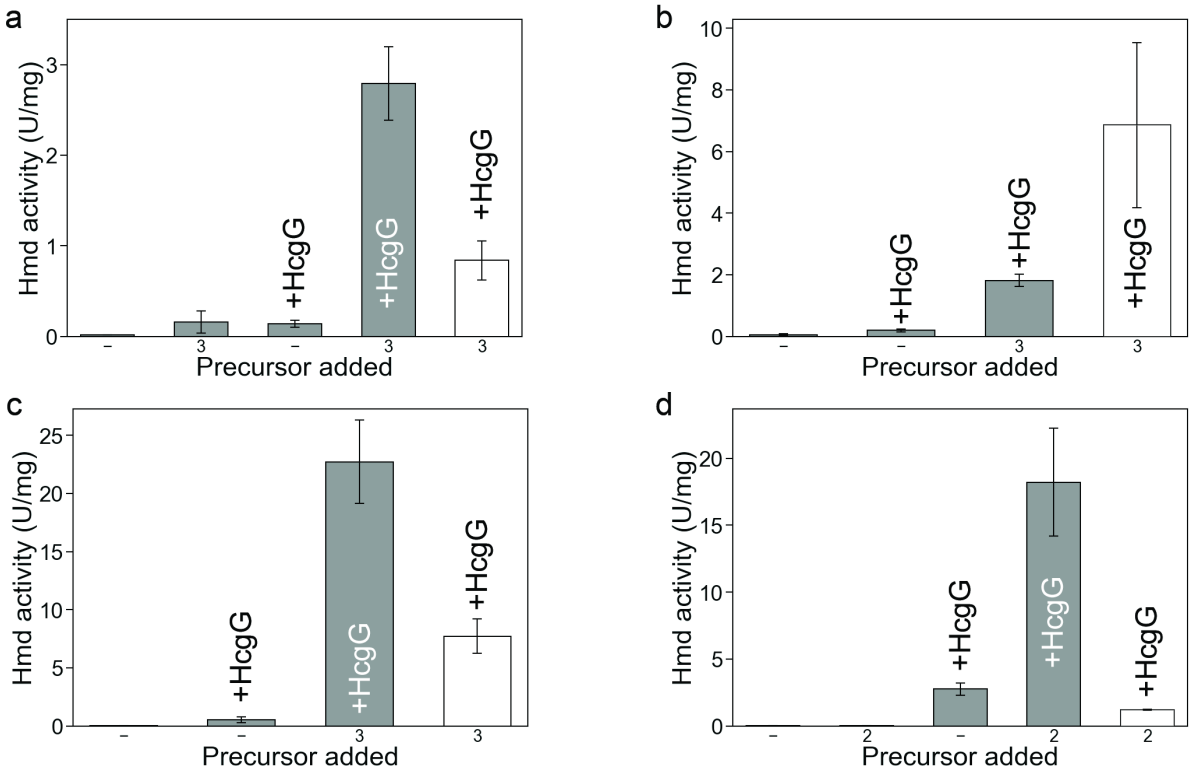


Figure 47. Large variation of the *in vitro* biosynthesis activity using different batches of the cell extract. *In vitro* biosynthesis under the standard assay condition using the cell extract from *M. maripaludis* Δ*hcgG* strain and precursor **3** (**a**, **b**, and **c**) or **2** (**d**) was complemented with purified HcgG. As a negative control, assay without precursor was also tested (–). The samples with HcgG are indicated by a label at the bar (+HcgG). The samples without HcgG are not indicated by a label. Four *in vitro* biosynthesis assays using different cell extracts are shown in each panel. The error bars represent the standard deviation of three separate measurements.

I tested whether HcgG catalyzes the production of dAdo because HcgG is proposed as a radical SAM enzyme and the AlphaFold model shows structural similarity to the radical SAM enzyme HydG (see section 5.2). However, no dAdo was detected in the *in vitro* biosynthesis assay parallel to the increase of the activity of the [Fe]-hydrogenase (Figure 48). This might be due to the fact that the assay was performed with the cell extract, where dAdo might be degraded by dAdo-degrading enzymes present in the cell extract.[136] To test this possibility, a SAM-cleavage assay was performed without the cell extract (Figure 49). dAdo was produced by HcgG from SAM in buffer and was increased by addition of the protein-free extract. The addition of SAH to the assay inhibited the production of dAdo. These experiments indicated that HcgG catalyzes a radical SAM reaction. Unexpectedly, the addition of CO to the gas phase inhibited the cleavage of SAM (Figure 49). CO inhibition is not seen in other radical SAM reactions. The CO inhibition property might be related to the possible CO-forming activity of HcgG described in the next section.

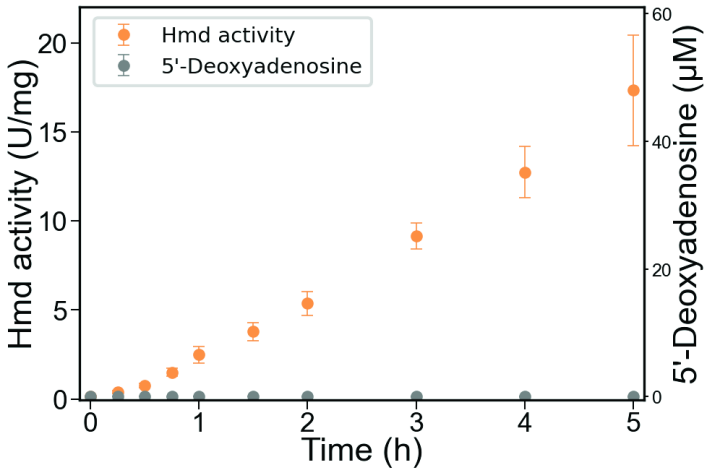


Figure 48. Time-dependent *in vitro* biosynthesis of the FeGP cofactor. Biosynthesis was performed from precursor **2** with all small components in the standard assay solution and GTP. The Hmd activity equivalent to 16 µM of the FeGP cofactor was produced. No dAdo was detected parallel to the increase of the Hmd activity.

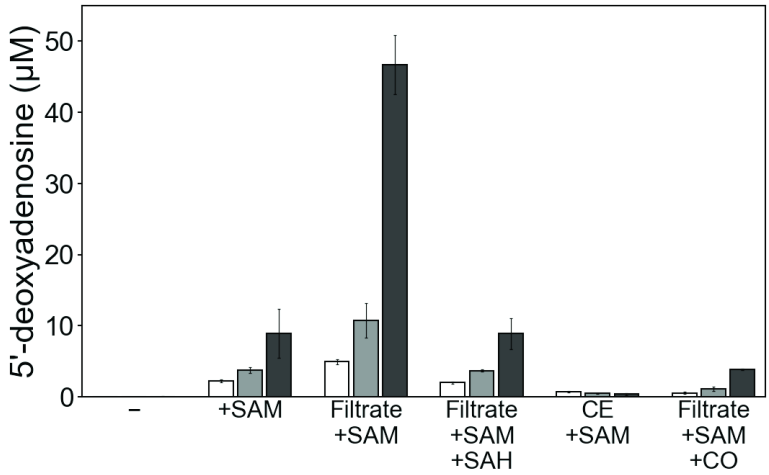


Figure 49. Production of 5'-deoxyadenosine (dAdo) by HcgG in different reaction conditions under N_2/H_2 (95%/5%), with an exception under N_2/H_2/CO (45.5%/2.5%/50%) gas phase as indicated (+CO). Bars represent the incubation time at 0 hours (open bar), at 1 hour (grey) and at 6 hours (dark grey). No dAdo was detected without SAM (–), while in the presence of only SAM, a slight production of dAdo was observed (+SAM). The dAdo-forming activity was stimulated by the addition of protein-free extract (3 kDa cut-off filtrate of the cell extract) (+Filtrate +SAM). Under the CO gas phase the reaction was inhibited (Filtrate +SAM +CO). In the presence of the cell extract and SAM did not show dAdo formation. Error bars indicate the standard deviation of three separate measurements.

5.8.3. Production of CO by HcgG

To study whether HcgG can produce CO, the protein-free extract of the Δ*hcgG* strain was incubated with HcgG in the presence of sodium dithionite, SAM, and myoglobin in 50 mM Tris/HCl pH 7.4. Binding of CO to heme in myoglobin was detected by a shift of the UV-visible spectrum[137] (Figure 50a). The same shift of the UV-visible spectrum of myoglobin was observed in the previous experiments of detection of the CO production in the HydG reaction assay.[138] This result indicated that CO is produced in this assay with HcgG and the CO produced was bound to myoglobin. A time-dependent change in the spectra of myoglobin was determined by incubation with HcgG (Figure 50b). The amount

of CO produced by the HcgG reaction was calculated using the extinction coefficient of carboxymyoglobin (ε_{423} = 183 mM^{-1} cm^{-1}).[137] Incubation of myoglobin with another radical SAM enzyme (HcgA) resulted in no change in the spectrum of myoglobin, which indicated that this CO-producing activity is specific for HcgG. The CO-producing activity was dependent on SAM added to the assay. However, the activity was also observed in the buffer solution with SAM in the absence of the protein-free extract, which suggested that Tris could be a substrate of the HcgG reaction. Titration of the CO-producing activity of HcgG with increasing concentrations of Tris indicated that the CO-producing activity increased by addition of Tris (Figure 51b).

The production of CO by HcgG and the possible structural similarity with HydG suggest that HcgG might use an amino acid as the substrate because HydG catalyzes the CO production from tyrosine. As it is known that tyrosine is not the substrate of the CO ligand from the previous isotope-labeling work, I tested the myoglobin assay in the presence of casamino acids, which is a mixture of amino acids, in the Tris buffer (Figure 51a). Addition of casamino acids in this buffer stimulated the CO-producing activity, which supports involvement of an amino acid as the substrate. Tryptophan could be exempted from the candidates of the amino acids because tryptophan is destroyed in the production process of casamino acids.

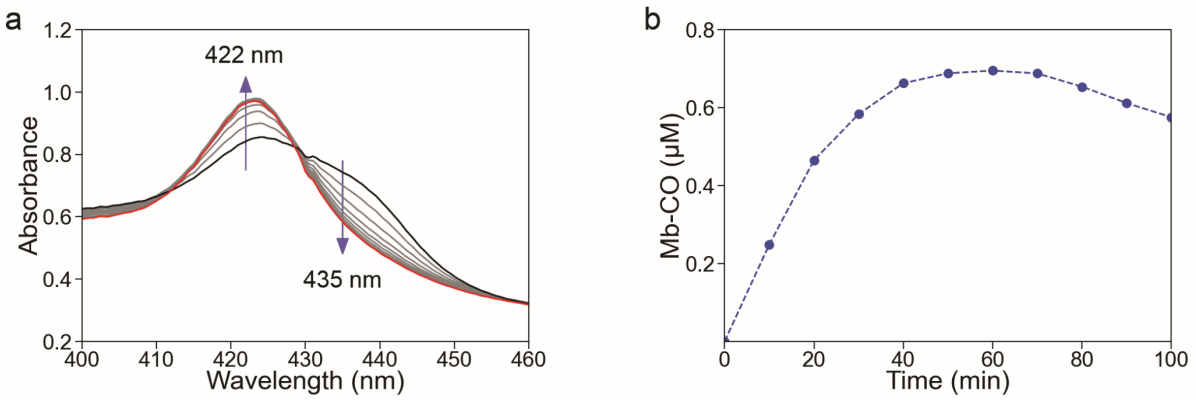


Figure 50. Change of the UV-visible spectrum of myoglobin in the presence of SAM, the *M. maripaludis* Δ*hcgG* strain protein-free extract and HcgG in 50 mM Tris/HCl pH 7.4 under N_2/H_2 (95%/5%) gas atmosphere. (**a**) The change of the UV-Vis spectrum during the reaction for 120 min was recoded at 10 min intervals. The spectrum was recorded using 3-mm light-path quartz cuvette. (b) The absorbance at 422 nm was plotted time-dependent manner.

It is difficult to imagine Tris as the substrate of HcgG from its chemical structure because Tris contains no carbonyl group. To date, there are no reports describing that Tris is used by a radical SAM enzyme as the substrate although it is known that Tris binds to the [Fe-S]-cluster of a radical SAM methylthiotransferase (RimO).[139] To draw a conclusion, further experiments are required.

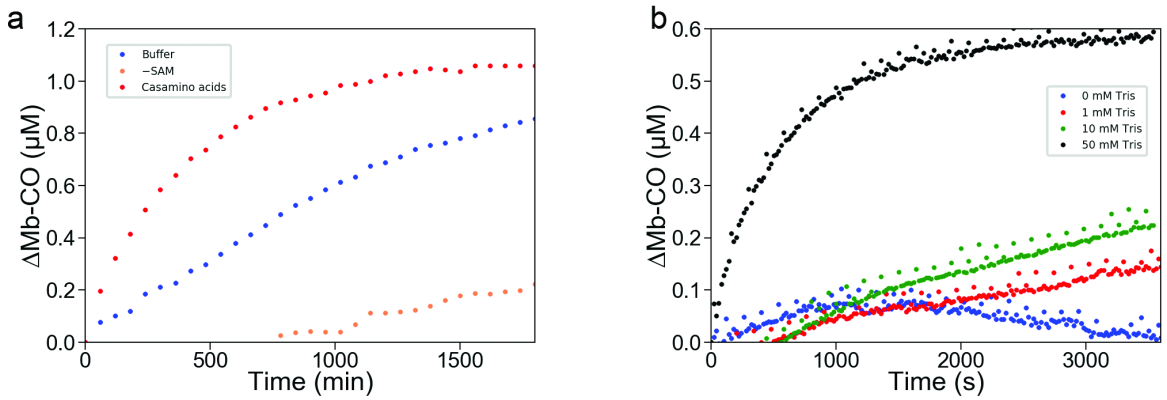


Figure 51. CO production activity of by HcgG detected by the myoglobin assay. The CO-production was detected the change of the spectrum of myoglobin at ΔA_{423}. (**a**) The HcgG assay in the presence (Buffer) and absence of SAM (–SAM), and in the presence of 2.5 mM SAM and 2 mg/ml casamino acid in the 50 mM Tris/HCl pH 7.4. (**b**) The HcgG assay in 50 mM MOPS buffer pH 7.4 with different concentration of Tris.

5.8.4. Metabolomic analysis of the substrate of HcgG.

In a similar fashion to the method for HcgA, the compounds were separated by reverse phase HPLC and analyzed by Orbitrap in negative mode with an *m/z* window from 50 to 500. Because of the behavior of the compounds in the cell extract, the threshold for the comparison of the regressions was decreased to $m_{reaction}/m_{control} > 10$ for the increasing compounds and $m_{reaction}/m_{control} < -2$ for the decreasing compounds. Because of the low threshold, many outlier compounds were selected and were manually removed. Similar to the case of HcgA, no clear compound could be identified as the substrate of HcgG, however certain compounds are of note. The detection of *p*-cresol (index 397) in the decreasing compounds is interesting, because it is the by-product of the reaction of HydG when producing the CO and CN^- ligands.[140] In addition, a slight decrease over time of Tris (index 35) was detected, which could relate to the previous CO-releasing experiment; however, the correlation is not good and it may due to deviations in the quantification of the buffer. The full lists of compounds that increase is shown in Table 7 and Figure 52; however, it should be noted that the annotations are done automatically by predicting the estimated formula and comparing it to the ChemSpider database;[141] therefore, many detected compounds have no identified name. The list of compounds that decrease over time is shown in Table 8 and Figure 53 . To simplify the reaction mixture, it is necessary to develop a new assay without the cytosolic proteins. For this aim, we need to develop an *in vitro* biosynthesis assay containing only the protein-free extract, Hcg proteins, apo-Hmd and the standard reaction solution. Sebastian Schaupp, has reported success with such systems, but unfortunately the reproducibility is poor.[103] In the future, it will be necessary to study conditions to improve the reproducibility of *in vitro* biosynthesis systems that do not use cell extracts.[103]

Table 7. Compounds that increased over time in *in vitro* biosynthesis using the *M. maripaludis* Δ*hcgG* strain with HcgG. Index represents the position of the compounds in the full list (See Supplementary information).

Name	Predicted formula	Molecular weight (Da)	RT (min)	Index
	$C_7H_{11}N_3O_3P_2$	247.02771	1.598	105
	$C_{15}H_{33}N_3O_5P_2$	397.18816	1.62	217
Guanine	$C_5H_5N_5O$	151.04907	4.939	261
Timonacic	$C_4H_7NO_2S$	133.01953	2.239	354
Miglitol	$C_8H_{17}NO_5$	207.1115	2.399	429
	$C_8H_6N_4O_3$	206.04307	1.422	483
	-	179.13972	1.489	579
Hydroxyhexanoycarnitine	$C_{13}H_{25}NO_5$	275.17291	4.921	583
	$C_4H_7N_7$	153.0764	1.372	590
	$C_{11}H_{15}N_5O_4S$	313.08406	4.938	639
	$C_5H_{15}N_4O_7P$	274.06752	4.211	729
Isatin	$C_8H_5NO_2$	147.03168	6.466	932
	$C_3H_{11}ClN_3OP$	171.03289	1.383	963
	$C_5H_{12}N_8O_6$	280.08784	6.304	989
	$C_3H_3N_2O_6P$	193.97278	7.53	1180
	-	206.01152	4.147	1188
	$C_7H_3ClN_6$	206.01114	8.586	1189
	$C_7H_3ClN_6$	206.0114	4.631	1191

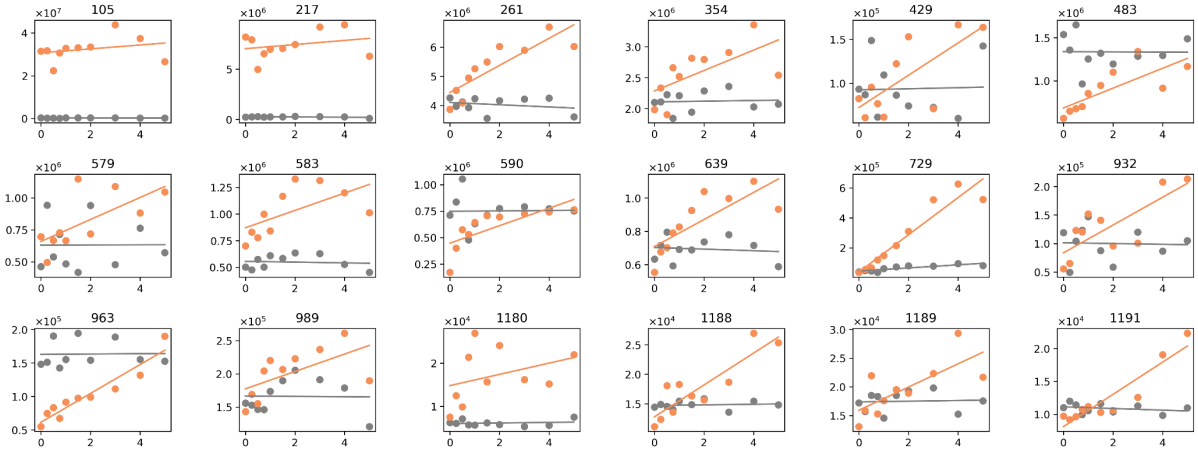


Figure 52. Trend of the MS intensity of selected compounds that increased over time in *in vitro* biosynthesis using the *M. maripaludis* Δ*hcgG* strain with HcgG. The index in the title represents the position of the compounds in the full list.

Table 8. Compounds that decreased over time in *in vitro* biosynthesis using the *M. maripaludis* Δ*hcgG* strain with HcgG. Index represents the position of the compounds in the full list (See Supplementary information).

Name	Predicted formula	Molecular weight (Da)	RT (min)	Index
	$C_7H_{15}NO_4S$	209.0719	1.624	16
	$C_5H_9N_7O_2S$	231.0538	1.601	18
Tris	$C_4H_{11}NO_3$	121.0738	1.318	35
	$C_6H_{12}N_5OPS$	233.0496	1.6	191
	C_7H_6	90.04678	6.189	243
Sulfurolpropionate	$C_9H_{13}NO_2S$	199.0663	6.189	288
	$C_{12}H_{20}N_6O_4S$	344.1264	1.683	330
DH3325000	$C_7H_6O_2S$	154.0085	6.185	389
	-	152.012	6.186	394
p-Cresol	C_7H_8O	108.0573	6.188	397
	$C_{11}H_{22}O_4$	218.1513	6.74	557
	$C_9H_{16}N_6O_2$	240.1332	6.743	558
Phenylethylalcohol	$C_8H_{10}O$	122.0729	6.188	560
	-	107.0494	6.188	564
	$C_7H_{15}NO_5S$	225.0667	1.716	596
9-Hydroxy-10-undecenoicacid	$C_{11}H_{20}O_3$	200.1408	6.743	665
Hexyl2-furoate	$C_{11}H_{16}O_3$	196.1095	5.809	715
	$C_{19}H_{41}NO_2$	315.3132	6.598	722

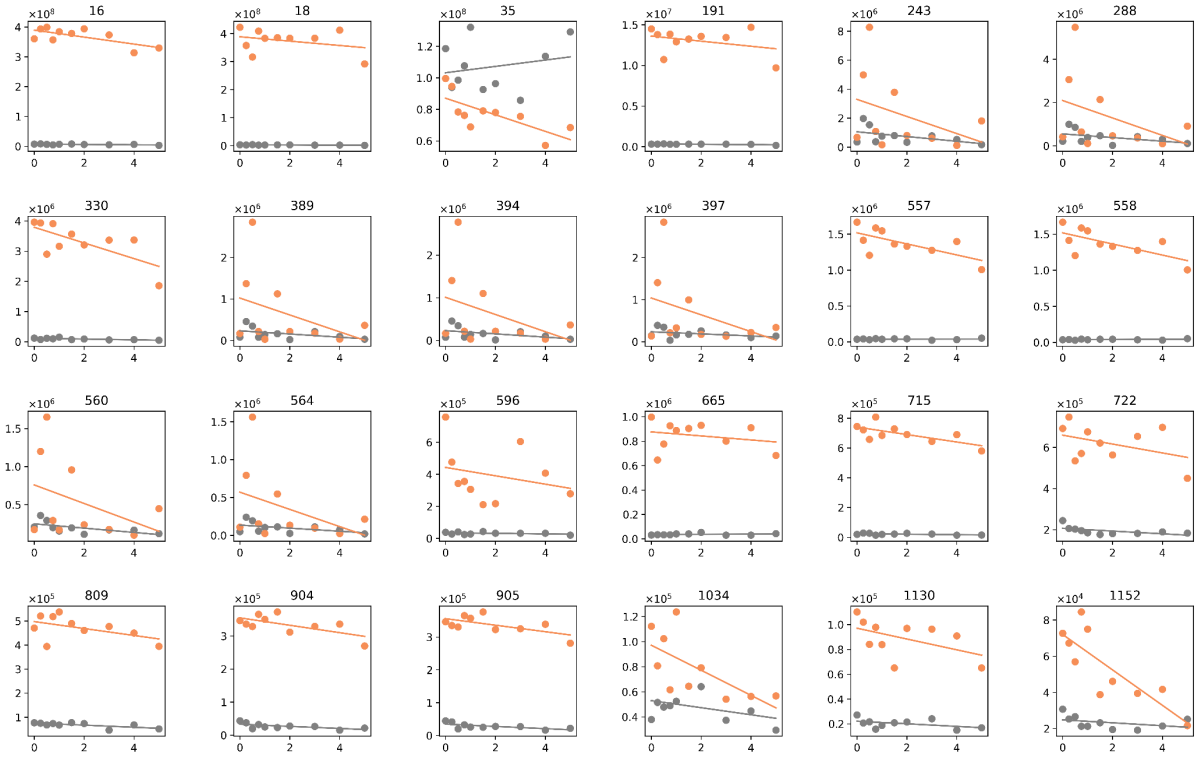


Figure 53. Trend of the MS intensity of selected compounds decreased over time in *in vitro* biosynthesis using the *M. maripaludis* Δ*hcgG* strain with HcgG. The index in the title represents the position of the compounds in the full list.

5.9. The binding trajectory of the FeGP cofactor to the [Fe]-hydrogenase

The Hmd apoenzymes can be heterologously produced in *E. coli*; however, the heterologously-produced enzymes are inactive due to a lack of the FeGP cofactor.[84] The FeGP cofactor can be extracted from the purified native enzyme from methanogens (i.e., *M. marburgensis*) and used to reconstitute the holoenzyme by mixing it with the apoenzyme without a chaperone system.[84] Crystal structures of the apoenzyme from *M. jannaschii* showed that the enzyme without cofactor is folded in a closed conformation;[44] however, upon incubation with the isolated FeGP cofactor, this cofactor binds to the active site cleft and produces an active holoenzyme in an open conformation (Figure 3).[45] Crystal structures of the holoenzyme have always been observed in an open conformation, where the FeGP cofactor is exposed to the solvent.[43] Upon binding of the substrate methenyl-H_4MPT^+, the enzyme undergoes a conformational change, where the enzyme changes to a closed conformation. In the closed conformation, the water molecule bound to the open coordination site of the FeGP cofactor is expelled, which activates the Fe active site. The remaining empty coordination site is proposed to be the H_2 binding site.[53]

Several mimic compounds have been synthesized by creating new scaffolds for the Fe-center and/or changing the active metal to other metals, for example, Mn.[142–147] Some of these mimic compounds have been shown to be incorporated into the Hmd apoenzyme. The semisynthetic [Fe]- and [Mn]-hydrogenases show enzyme activity, albeit very low.[145] In those studies, the addition of GMP to the reconstitution process enhanced the activity of the reconstituted protein.[143,145] The GMP moiety does not play an active role in the catalysis;[53] however, the simultaneous binding with the mimic compounds improved the specific activity of the Hmd-mimic complex.[145]

We obtained the crystal structure of Hmd from *Methanolacinia paynteri* after reconstitution with the FeGP cofactor, where we observed an asymmetric binding of the cofactor to the apoenzyme dimer. One monomer was obtained in the closed conformation without the FeGP cofactor, while the second monomer was observed in the open conformation with the FeGP cofactor bound to the protein (Figure 54). This structure indicated that the open/close conformational change in the [Fe]-hydrogenase occurs in each monomer independently.

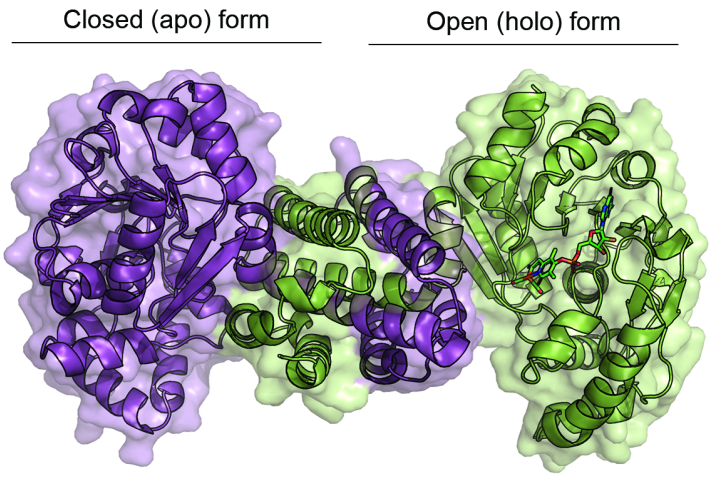


Figure 54. Asymmetric crystal structure of [Fe]-hydrogenase of *M. paynteri* homodimer. One monomer in apo-form is shown in purple and another in holo-form with the FeGP cofactor in green (PDB: 6YKA). The FeGP cofactor is shown by a stick model.

The N-terminal domains of the apo- and holo-forms of the asymmetric homodimer are very similar to the corresponding apo- and holo-forms of the [Fe]-hydrogenase from *M. jannaschii* (Figure 55a); however, in the apo-form, there is a difference in the loop involved in the binding of the FeGP cofactor (Figure 55b). A conserved lysine (Lys150) in the loop points to the active site in the structure from *M. paynteri*, while it moves outside the active site in the structure of Hmd from *M. jannaschii*. This lysine is strictly conserved in all Hmd genes except in genomes from *Methanobrevibacter* species, where a glutamate is found in its place.

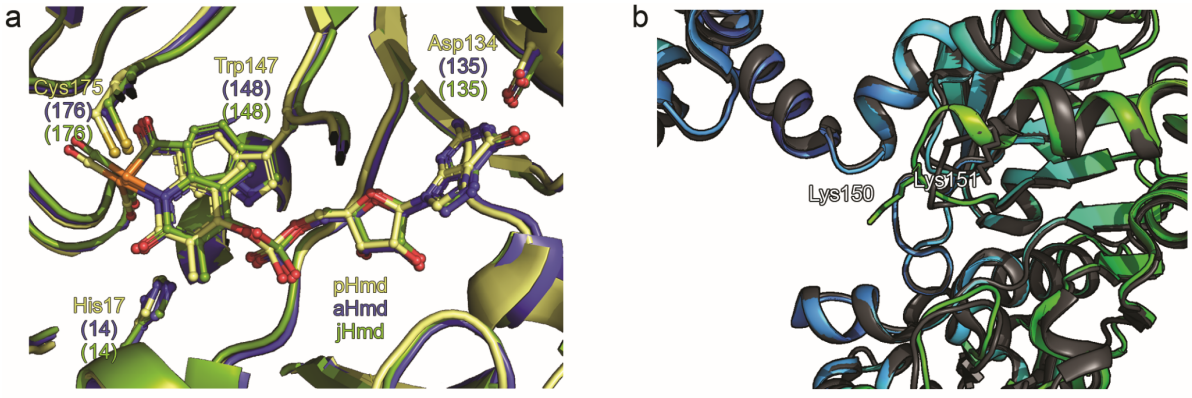


Figure 55. Structures of the FeGP cofactor binding site in Hmd. (**a**) The FeGP cofactor binding sites of pHmd (PDB: 6YKA) was compared with Hmd from *Methanococcus aeolicus* (aHmd) (PDB: 6HAC) and jHmd (PDB: 3F47). The protein parts are shown in cartoon model. The side chain of residues corresponding to the FeGP cofactor binding and the FeGP cofactor are shown in ball and stick model. (**b**) Comparison of the loop structure involved in the FeGP cofactor coordination in jHmd and pHmd. Apo-jHmd was colored in black. Apo-pHmd part of the pHmd asymmetric homodimer was indicated by colors (dark blue with low B-factor to green with higher B-factor).

An additional crystalline form of pHmd was found when the protein was co-crystallized with a mimic compound of the FeGP cofactor and GMP. No mimic compound was observed in the crystal structure, but GMP was bound in the same position as the GMP moiety of the FeGP cofactor in the holo-form.

In this GMP binding form of pHmd, Lys150 was observed in the same conformation as in the apo-form. If this Lys150 conformation is conserved in the holo-form, the lysine clashes with the phenyl part of methenyl-H_4MPT^+ in the catalytic reaction. In the ternary complex of Hmd from *M. aeolicus*, Lys150 interacts with the substrate via a water network, which indicates that this conformational change should affect the substrate-binding properties (Figure 56).

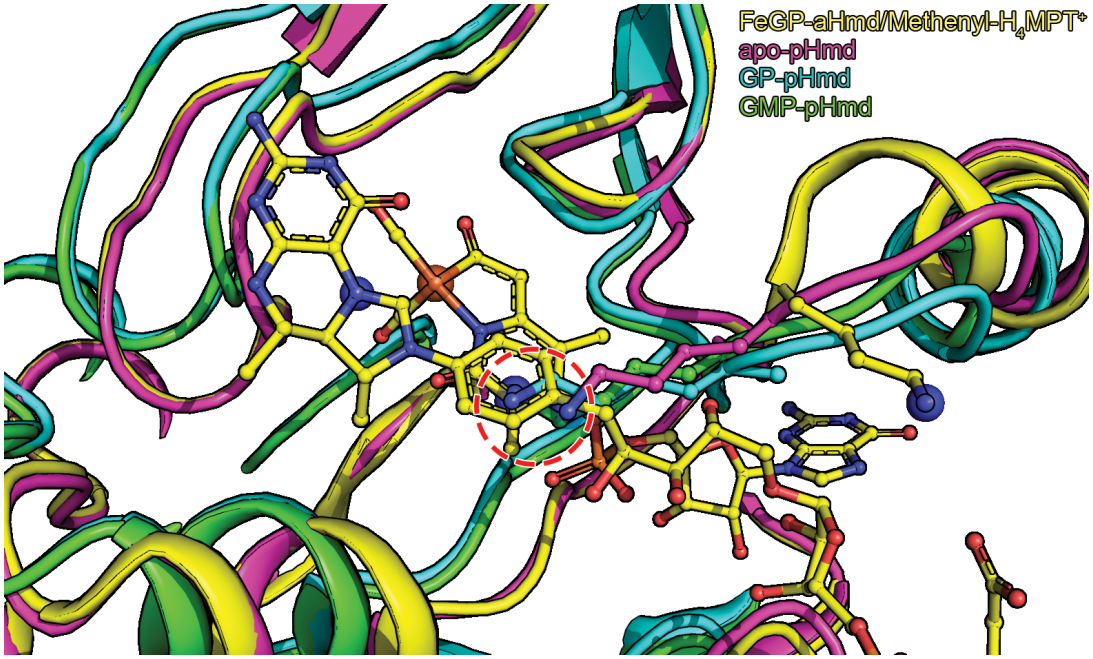
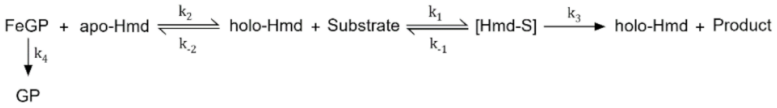


Figure 56. Superposition of the active sites of the apo- and holo-forms of Hmd. Lys150 is shown in ball and stick model. The clash point between Lys150 side chain and the phenyl ring part of methenyl-H_4MPT^+ is highlighted by a red dash open circle. The holoenzyme of Hmd from the *Methanococcus aeolicus* (aHmd) holo-form in complex with methenyl-H_4MPT^+ (yellow, FeGP-aHmd/methenyl-H_4MPT^+), the apo-form of the asymmetric homodimer of pHmd (purple, apo-pHmd), the GMP-bound form of pHmd (green, GMP-pHmd) and the GP-bound form of pHmd (cyan, GP-pHmd) are shown.

The observed structures suggested that Lys150 has a role in the binding of the FeGP cofactor. To test this hypothesis, a Lys150Ala mutant of the pHmd enzyme was produced and the kinetic parameters were measured. At pH 6 in the oxidation reaction of methylene-H_4MPT, the Lys150Ala variant exhibited much higher V_{max} and K_m values (640 U/mg and 160 µM) than the wild-type (66 U/mg and 6 µM). In the reduction reaction of methenyl-H_4MPT^+ at pH 7.5, the wild-type exhibited a V_{max} of 1300 U/mg and a K_m of 62 µM, while the mutant had a V_{max} of 820 U/mg and a K_m of 110 µM. The increase of the K_m values in the Lys150Ala mutant is consistent with the observation of the contact of the Lys150 residue with the substrate (Figure 56). Since this lysine is involved in a conformational change upon the binding of the FeGP cofactor, mathematical models of the reactions catalyzed by Hmd were constructed to further study the binding dynamics. As the enzyme is only active with the FeGP cofactor, we can measure the apparent activity with an excess of substrate and a limiting amount of the FeGP cofactor. The reaction was simplified and the binding of H_2 was not considered in the model as it is in excess. The full set of ordinary differential equations is shown in Figure 57.

Enzymatic Reaction:



Full set of ODEs:

$$\frac{d[FeGP]}{dt} = -k_2 \cdot [FeGP](t) \cdot [apo\text{-}Hmd](t) + k_{-2} \cdot [holo\text{-}Hmd](t) - k_4 \cdot [FeGP]$$

$$\frac{d[apo\text{-}Hmd]}{dt} = -k_2 \cdot [FeGP](t) \cdot [apo\text{-}Hmd](t) + k_{-2} \cdot [holo\text{-}Hmd](t)$$

$$\frac{d[holo\text{-}Hmd]}{dt} = k_2 \cdot [FeGP](t) \cdot [apo\text{-}Hmd](t) - k_2 \cdot [holo\text{-}Hmd](t) - k_1 \cdot [S](t) \cdot [holo\text{-}Hmd](t) + k_{-1} \cdot [Hmd\text{-}S]$$

$$\frac{d[Substrate]}{dt} = -k_1 \cdot [S](t) \cdot [holo\text{-}Hmd](t) + k_{-1} \cdot [Hmd\text{-}S]$$

$$\frac{d[Hmd\text{-}S]}{dt} = k_1 \cdot [S](t) \cdot [holo\text{-}Hmd](t) - k_{-1} \cdot [Hmd\text{-}S] - k_3 \cdot [Hmd\text{-}S](t)$$

$$\frac{d[Product]}{dt} = k_3 \cdot [Hmd\text{-}S]$$

Figure 57. Enzymatic reaction and full set of ordinary differential equations (ODE) derived from law of mass action. To calculate the apparent V_{max} and apparent K_m, we fixed the concentration of H_2 and H^+. Therefore, the rate of the changes of H_2 and H^+ is zero and the model was approximated to the Michaelis-Menten (single substrate/product) equation. Modified from Huang, *et al.*[85]

Numerical integration of these equations allows us to predict the enzymatic reaction, as seen in Figure 58. Besides the changes in the kinetic parameters, we can use the model to predict the effect of the Lys150Ala mutation in the binding of the FeGP cofactor to Hmd. At pH 7.5, the calculated binding constant is 10-fold lower for the mutant (0.01 $\mu M^{-1}s^{-1}$ vs 0.002 $\mu M^{-1}s^{-1}$), while there was no change at pH 6 (0.10 $\mu M^{-1}s^{-1}$ vs 0.13 $\mu M^{-1}s^{-1}$). These results suggest that the binding of the FeGP cofactor is a pH dependent process and that the Lys150 residue plays a key role in the conformational change that happens at physiological conditions at pH 7.5 (Figure 59). The use of mathematical tools provides an effective way to predict and test different hypothesis that can be kinetically measured and can complement other techniques like isothermal titration calorimetry, crystallography, and even other simulation techniques such as quantum mechanics/molecular mechanics (QM/MM).

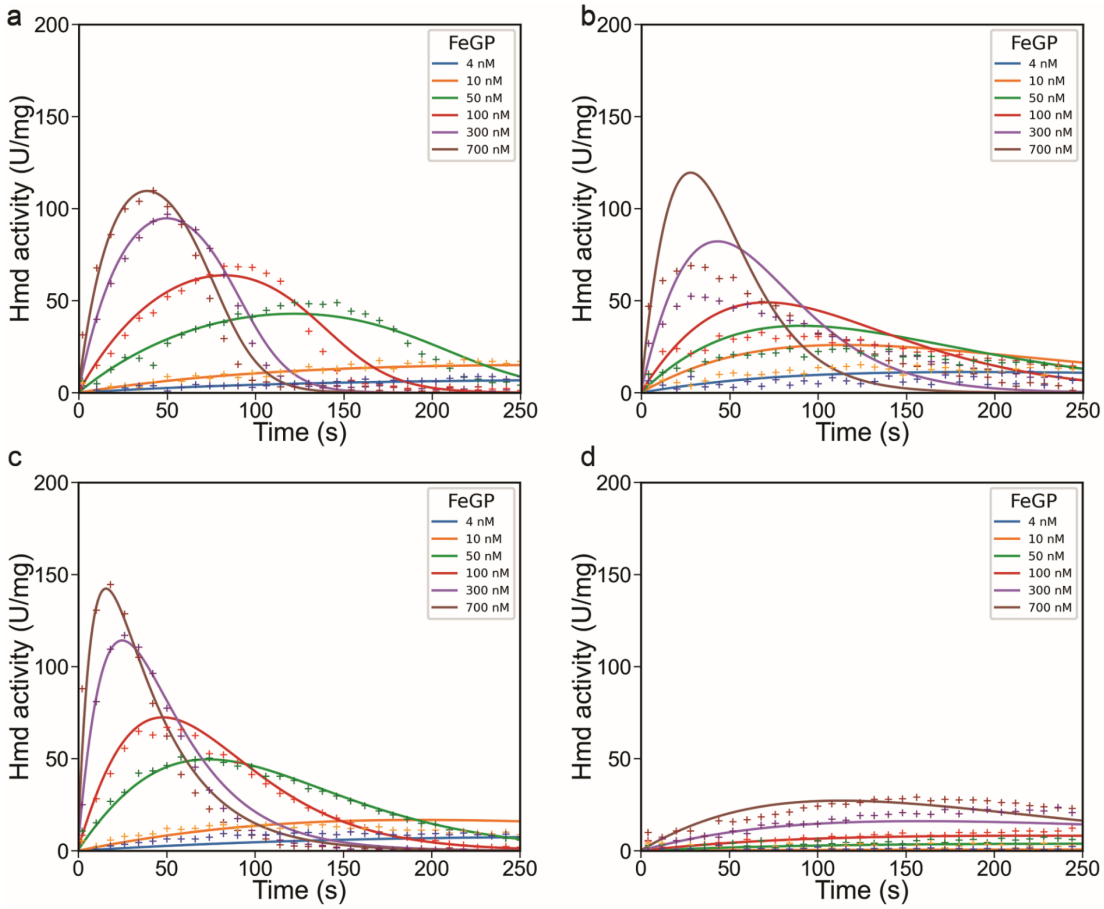


Figure 58. Simulation of the change of the activity in the reconstitution assay of pHmd. Experimental data is shown in (+) and simulated data in solid lines. Because the occupancy of the FeGP cofactor in the protein was variable under the standard conditions. The V_{max} value was adjusted by the ratio between the specific activity of the reconstituted enzyme at 18 µM substrate concentrations and that of the maximum activity obtained in the reconstitution kinetic assays. Reduction of methenyl-H_4MPT^+ at pH 7.5 from the wild type (**a**) and Lys150Ala mutant (**b**) is shown in the upper panels. Oxidation of methylene-H_4MPT at pH 6 from the wild type (**c**) and Lys150Ala mutant (**d**) is shown in the lower panels. Modified from Huang, *et al.*[85]

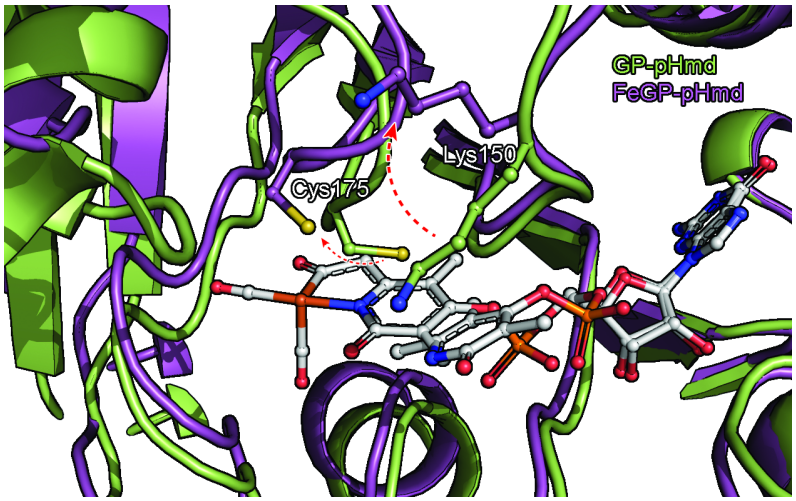


Figure 59. The binding trajectory of the FeGP cofactor. First the GP part of the cofactor binds to the protein (GP-pHmd, green), which triggers a conformational change and the Lys150 residue moves out of the binding pocket, allowing the Fe-center to bind and align with the Cys175. The proposed movements of the amino acid residues are shown with red arrows.

5.10. A possible role of HmdII in the regulation of the [Fe]-hydrogenase activity

The disappearance of the Hmd activity of the *M. maripaludis* Δ*frh* strain prompted me to study the inactivation mechanism. Proteome analysis indicated that the amount of Hmd and Hcg proteins in the cell extract did not change by this mutation (See Table 4). The presence of the Hcg proteins suggested that the production of the FeGP cofactor in the Δ*frh* strain is not damaged. Indeed, I could extract the FeGP cofactor from the cell extract of the Δ*frh* strain and the amount of the FeGP cofactor detected was sufficient to support the Hmd activity in the wild type strain (Figure 60). These observations indicated that the Δ*frh* strain is able to produce the Hmd and Hcg proteins; therefore, the Hcg proteins produce the FeGP cofactor. In this case, a possibility of reduction of the Hmd activity is that Hmd cannot bind the FeGP cofactor to make the active Hmd holoenzyme. In the proteome analysis, I could see that an Hmd paralog (HmdII) is overproduced in the Δ*frh* strain. HmdII has a binding affinity to the FeGP cofactor. However, the HmdII bound with the FeGP cofactor exhibits only very low enzymatic activity.[131] Because the estimated amount of HmdII in the cell extract of the Δ*frh* strain is ten-fold higher than the amount of Hmd, HmdII could bind most of the FeGP cofactor and therefore diminish the Hmd activity in the cell.

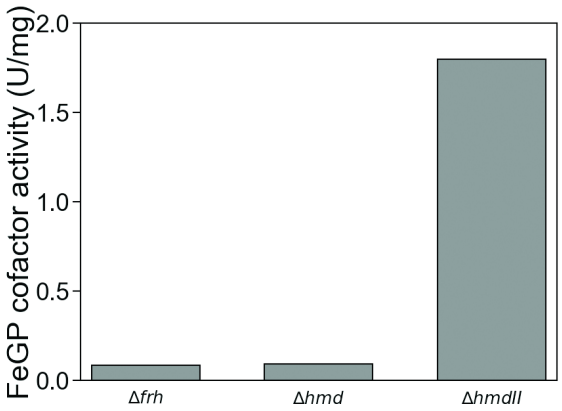


Figure 60. FeGP cofactor extracted from the *M. maripaludis* mutants.

To further test the physiological function of HmdII, I tested the effect of the deletion of the *hmdII* gene from the Δ*frh* strain. This mutation resulted in a five-fold increase of the Hmd activity in the Δ*hmdII* strain over the wild type S2 strain (Figure 61). Unexpectedly, deletion of the *hmdII* gene increased the amount of the FeGP-cofactor that is extractable from the cell extract. This finding suggests that HmdII regulates the Hmd activity by tuning the biosynthesis of the FeGP cofactor. However, the proteome analysis of the Δ*hmdII* strain indicated no change in the Hmd and Hcg proteins, which suggests that HmdII might regulate the other factors involved in biosynthesis of the FeGP cofactor, for example, regulation of the activity of Hcg proteins.

We do not know the reason why HmdII is over-produced in the Δ*frh* strain and whether this phenomenon is physiologically relevant; however, it is important to note that the *hcg* deletion mutants were only obtained using the Δ*frh* strain as the parent strain (J. A. Leigh personal communication). The deletion of the *hcg* genes causes dramatic changes in the metabolism by elimination of the Hmd activity in the cell. In contrast, as the Δ*frh* strain does not have Hmd activity, further deletion of the *hcg* genes does not give any physiological impact on the cells. For adaptation to physiological stresses, cells require time for tuning up the metabolism, which might cause a delay in growth and it might be detrimental for colony formation in mutant screenings.[72] Thus, the behavior of HmdII in the Δ*frh* strain could be the reason that the *hmd* and *hcg* mutations were only possible on the Δ*frh* strain.

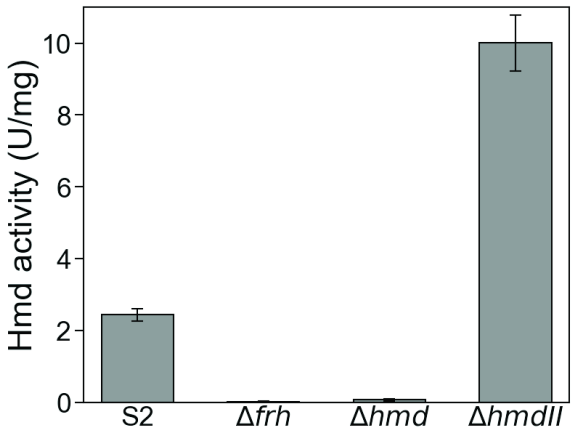


Figure 61. Hmd activity in the cell extract of *M. maripaludis* mutant strains. S2 corresponds to the wild type. All mutants were created on the Mm901 strain in which the *upt* gene was deleted. Δ*frh* corresponds to a Δ*frcAfruA* mutant. Both Δ*hmd* and Δ*hmdII* mutations are created in the Δ*frh* background.

6. Conclusions and Outlook

In this study, I was able to express and purify the biosynthetic proteins HcgA and HcgG. By biochemical characterization, I determined that HcgA produces the pyridinol precursor **1** from an unknown substrate present in the cell extracts of *M. maripaludis*, *M. marburgensis* and *E. coli*. Although the substrates of HcgA cannot be inferred from the primary structure, untargeted metabolomic studies of the reaction shows several compounds that increase or decrease over time in correlation with the production of precursor **1** (Figure 41 and Figure 42). Because of the similarity of some of these compounds to precursor **1**, they could be candidates for the HcgA substrate (Figure 62). Radical SAM enzymes have been shown to catalyze difficult and versatile reactions, such as rearrangements, cyclization of organic compounds, insertion of sulfur atoms and other reactions. The detected compounds could be the precursor in the formation of precursor **1**; however, there are substantial differences in the chemical structures between these compounds and precursor **1**.

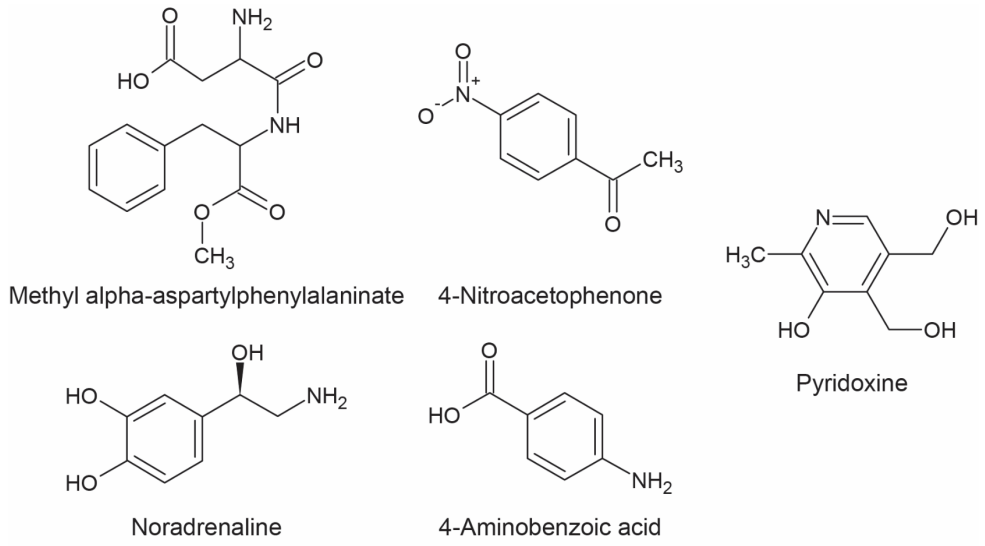


Figure 62. Selected compounds found in the MS analysis of the HcgA reaction. Note that this compounds all present an aromatic ring, with nitrogen and carboxylic acid or similar elements.

To further determine the substrate of HcgA, fractionation of the cell extract should be performed using several approaches, like as organic extraction of components and other chromatographic methods such as gel permeation column. Structural studies of HcgA have been so far unsuccessful as no crystals have been obtained. Although the enzyme can now be routinely produced with the proper Fe content per protein using this new protocol, the purified HcgA protein could behave better for crystallization due to proper incorporation of [4Fe-4S]-clusters. Different reconstitution methods to improve the ability of crystal formation could be tested. Co-crystallization of the protein with either substrates or products is another technique that could improve the likelihood of obtaining good crystals. The formation of compounds like pyridinol **1** (2-pyridones) is an important biotechnological

target because some of the derivatives are bioactive compounds.[148–150] However, all the reactions for 2-pyridone derivatives so far do not include a radical SAM enzyme in the respective pathways. The determination of the reaction mechanisms of HcgA would expand the toolbox of the radical SAM superfamily and increase the understanding of these remarkably interesting enzymes.

In this work, I successfully expressed and produced HcgG in a soluble form. The enzyme was homologously produced in *M. maripaludis* in enough quantities for biochemical characterization. *In vitro* complementation analysis indicated that HcgG catalyzes biosynthetic reactions after the biosynthesis of precursor **3**, which probably include formation of the CO and acyl ligands, Fe insertion and completion of the FeGP cofactor. By using purified HcgG, I determined that HcgG is a radical SAM enzyme that is inhibited by CO, and that this enzyme catalyzes the formation of CO from Tris and/or a casamino acid, suggesting that an amino acid could be the physiological CO donor. As the next step, I will identify the CO donor substrate. The study of the catalytic mechanism will be initiated by analyzing the product formation with stable isotope labelled compounds. I hope that we will be able to clarify the mechanism of this new radical SAM enzyme reaction using spectroscopic and structural analyses in future.

The untargeted metabolomic approach in this work showed no candidate compounds. Because of this, other strategies must be considered, especially further developing a defined *in vitro* assay for the biosynthesis of the FeGP cofactor that does not require the cell extract. In the latest results for [FeFe]-hydrogenase maturation, the reaction and precursor of the dithiomethylamine bridge of the H-cluster was identified using a defined lysate-free assay.[70] *In vitro* biosynthesis using pure proteins will improve the effectiveness of an untargeted metabolomic assay and open the door for further studies in a similar fashion to those proposed for HcgA.

This study showed that HcgG is a radical SAM enzyme, whose primary structure shows no similarity to any other proteins described; however, the structure prediction by AlphaFold shows a TIM-barrel structure with a Rossmann fold domain. This needs to be verified by structural studies, such as crystallography and/or cryo-electron microscopy. The structural characterization will expand the knowledge about the radical SAM superfamily.

In addition to the work in HcgA and HcgG, this project dealt with the roles of HmdII and the final step of the maturation of Hmd, the incorporation of the FeGP cofactor into the apoenzyme. HmdII appears to be a regulator of the Hmd activity in the cell, which could further expand the knowledge of the regulation of biosynthesis of metallocofactors and metalloenzymes. The binding trajectory of the FeGP cofactor to the apoenzyme was studied by structural and modelling studies. This led to propose a

trajectory where Hmd must undergo several local conformational changes. A specific residue, Lys150 in Hmd was determined to play an important role in the binding of the FeGP cofactor under physiological pH. Further studies of the residues that interact with the substrate or the FeGP cofactor need to be done as this knowledge could improve the activity of mimic compounds like Mn-derived cofactors,[143] which show very little activity probably due to unspecific binding of the cofactor and their misalignment to the substrate.[143]

Biochemical characterization of hydrogenases and other metalloenzymes including the biosynthesis machinery can expand our knowledge of the complex chemistry in microorganisms' metabolisms and the mechanisms of the enzymes conserving energy using e.g., CO_2 and H_2.

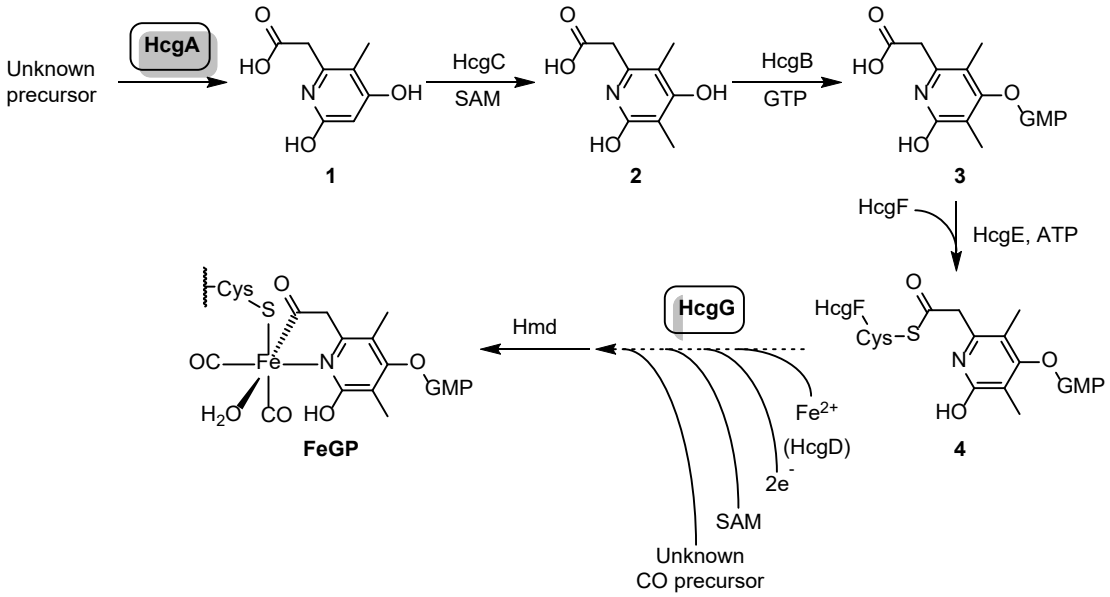


Figure 63. Updated biosynthesis pathway of the FeGP cofactor.

7. Materials and methods

7.1. Anaerobic experiments

7.1.1. Anaerobic solutions

Solutions were boiled under a constant stream of 100% N_2 for 15 minutes, transferred into a glass bottle and sealed with a butyl rubber stopper. After the solutions cooled to room temperature, they were transferred to an anaerobic chamber (Coy Laboratories, Grass Lake, MI, USA) containing N_2/H_2 (95%/5%) with a platinum catalyst to remove traces of oxygen. The solutions were stirred overnight with a magnetic stirrer at 500 rpm to remove traces of oxygen. For oxygen-sensitive compounds, the solids were first weighted, transferred into the anaerobic chambers and dissolved with oxygen-free water prepared as described above.

7.2. Expression and production of proteins

7.2.1. Cultivation of *M. maripaludis*

The strains of *M. maripaludis* shown in Table 9 were grown in a modified formate media in an atmosphere of N_2/CO_2 (80%/20%) with 100 mM Tris as buffer shown in Table 10.[129] Cultivation was performed in 5 L or 500 mL scale until an optical density (OD) at 660 nm = 0.6−0.8 as described previously.[102] The actively growing cells in the exponential growth phase were anaerobically harvested by a continuous-flow centrifuge (Heraeus 3049) with a continuous flow rotor at 15,000 rpm at 4°C until all the media flowed into the rotor, resuspended in the medium and sedimented by centrifugation with an Avanti JXN-26 centrifuge using a Beckman JLA 10.500 rotor at 7,300 rpm at 4°C for 30 min. The use of the culture medium for resuspension aimed to avoid lysis of the cells of a marine methanogen *M. maripaludis* in the low salt concentration buffer solutions. The cell pellets were finally anaerobically resuspended in a low salt concentration lysis buffer: 50 mM Tris/HCl pH 7.5 containing 5 mM $MgCl_2$ and 2.5 U/mL DNase I to a final concentration of 0.5 g cells/mL buffer. One mL aliquots were frozen in liquid N_2 and stored until use at −20°C. The frozen samples were anaerobically thawed on ice, which disrupted most of the cells. Unbroken cells and membrane particles were removed by ultracentrifugation with a Sorvall WX+ Ultracentrifuge using a TFT-80.4 rotor at 37,000 rpm and 4°C for 0.5 h. This supernatant is designated as the cell extract and used for *in vitro* biosynthesis assay.

Table 9. Strains of *M. maripaludis* used in this work.

Genotype	Strain name	Reference
Wild type *M. maripaludis*	S2	[151,152]
S2 Δ*upt*	Mm901	[24]
Mm901 Δ*frcAfruA*[a]	Mm1280	[153]
Δ*frcAfruA ΔhmdII*	Mm1281	[153]
Δ*frcAfruA ΔhcgA*	Mm1328	[72]
Δ*frcAfruA ΔhcgB*	Mm1329	[72]
Δ*frcAfruA ΔhcgC*	Mm1330	[72]
Δ*frcAfruA ΔhcgD*	Mm1331	[72]
Δ*frcAfruA ΔhcgE*	Mm1332	[72]
Δ*frcAfruA ΔhcgF*	Mm1333	[72]
Δ*frcAfruA ΔhcgG*	Mm1334	[72]
Δ*frcAfruA Δhmd*	Mm1335	[72]

[a]The deletions of the *hcg* genes were made in the strain Mm1280, in which both copies of the A subunit of the F_{420}-reducing hydrogenase isoenzymes genes, *fruA* and *frcA*, were deleted.[72,153] The Δ*fruA*Δ*frcA* double mutation is referred to as Δ*frh* in this study. The *hcg* gene mutations in the wild type strain are unstable by unknown reasons (J. A. Leigh, personal communication).

Table 10. Modified media for *M. maripaludis.*

Component	Amount per liter
General Salts	
KCl	0.34 g
$MgCl_2 \cdot 6H_2O$	2.75 g
$MgSO_4 \cdot 7H_2O$	3.45 g
$CaCl_2 \cdot 2H_2O$	0.14 g
NH_4Cl	0.50 g
Solids	
$NaHCO_3$	5.0 g
NaCl	10.5 g (<u>22 g</u>)
Na-acetate·$3H_2O$	1.4 g
Na-formate	13.6 g (<u>0 g</u>)
Liquids	
Buffer (2 M Tris pH 7.0)	50 mL (<u>0 mL</u>)
K_2HPO_4 solution (14 g/L)	10 mL
$FeSO_4 \cdot 7H_2O$ (1.9 g/L in 10 mM HCl)	5 mL
Trace mineral solution (Table 11)	1 mL
Nickel chloride (1 mM) Solution	1 mL
Resazurin 0.2%	0.5 mL
Amino acid solutions (Table 12)	15 mL each (<u>0 mL</u>)
Add before boiling with N_2	
Cysteine	0.50 g
Casamino acids	0 g (<u>2 g</u>)
Add before inoculation	
Vitamin solution (Table 13)	10 mL
Sodium sulfide (1M)	1 mL

For H_2/CO_2 growth use <u>underlined</u> values instead

Table 11. Trace mineral solution for *M. maripaludis*.

Component	Concentration (mg/L)
Na_3-citrate·$2H_2O$ [a]	21
$MnSO_4·2H_2O$	5
$CoSO_4$ [b]	1
$ZnSO_4·5H_2O$	1
$CuSO_4·5H_2O$	0.1
$AlK(SO_4)_2$	0.1
H_3BO_3	0.1
$Na_2MoO_4·2H_2O$	1
Na_2SeO_3	2
$V(III)Cl_3$	0.1
$Na_2WO_4·2H_2O$	0.033

[a] Adjust pH to 6.5 immediately afterwards

[b] Can be replaced by $CoCl_2·6H_2O$

Table 12. Amino acids solutions for *M. maripaludis.*

Component	Concentration (mg/L)
Amino acid Solution #1 (pH 7)	
Asparagine monohydrate	12.5
Aspartic Acid	10
Glutamine	10
Glutamic Acid	10
Methionine	10
Amino acid Solution #2 (pH 7)	
Isoleucine	10
Leucine	10
Alanine	10
Glycine	10
Valine	10
Amino acid Solution #3 (pH 7)	
Lysine monohydrochloride	12.5
Arginine	10
Serine	10
Threonine	10
Amino acid Solution #4 (pH 11)	
Tyrosine	25
Proline	10
Phenylalanine	10
Tryptophan	10
Histidine	10

Table 13. Vitamin solution for *M. maripaludis*.

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

7.2.2. Cultivation of *M. marburgensis*

M. marburgensis was cultivated in 2- or 10-liters glass fermenters with continuous gas flow (1.5 L/min) of $H_2/CO_2/H_2S$ (80%/20%/0.1%) at 65 °C using the previously described media[84], which contains 2.12 g/L NH_4Cl; 6.8 g/L KH_2PO_4; 2.54 g/L Na_2CO_3; 0.3 mL resazurin (0.2%); and 1 mL/L trace element solution, which contains 90 g/L Nitrilotriacetic acid; 40 g/L $MgCl{\cdot}6H_2O$; 10 g/L $FeCl_2{\cdot}6H_2O$; 0.2 g/L $CoCl_2{\cdot}6H_2O$; 1.2 g/L $NiCl_2{\cdot}6H_2O$; 0.2 g/L $NaMoO_2{\cdot}2H_2O$. For Ni-limitation conditions, $NiCl_2$ was omitted from the media. Cells were grown to an OD_{600} between 6 and 7.5. The fermenters were then cooled with ice water and the cells harvested by continuous-flow centrifugation using a Heraeus 3049 continuous flow rotor at 15,000 rpm at 4 °C under N_2 atmosphere. The rotor was transferred to the anaerobic tent and cells were stored in glass bottles sealed with butyl rubber stoppers. The usual yield was approximately 100g per 10-liter fermentation. Cells were stored at -75° C under 100% N_2.

7.3. Protein production and purifications

7.3.1. Heterologous production and purification of HcgA from *M. jannaschii*

Escherichia coli C41(DE3) cells were transformed with a plasmid for overexpression of [Fe-S]-cluster proteins (pRKISC),[104] and then further transformed with the expression plasmid pET24b(+) containing *M. jannaschii hcgA* gene with a N-terminal His_6-tag, whose codon usage was optimized for expression in *E. coli*. The recombinant *E. coli* cells were grown at 37 °C in lysogeny broth (LB) supplemented with 50 µg/mL kanamycin and 10 µg/mL tetracycline until OD_{600} = 0.6−0.8 and the gene expression was induced with 1 mM isopropyl ß-D-thiogalatopyranoside (IPTG). After the incubation for 2 hours, the cells were harvested by centrifugation using Avanti JXN-26 centrifuge with JLA-10.500 rotor at 7,300 rpm for 30 min at 4 °C and stored at –20 °C. HcgA was purified as described before:[80] frozen cell pellet was resuspended in 20 mM tris(hydroxymethyl)aminomethane (Tris)/HCl pH 8.0 containing 0.5 M NaCl and 20 mM imidazole (buffer A) and disrupted on ice by sonication for 10 min (1 min sonication, 1 min pause, five times) using a SONOPULS GM200 (Bandelin) with KE76 tip with 50% amplitude. The

supernatant was collected by centrifugation in a Sorvall WX+ Ultracentrifuge with a T-647.5 rotor at 30,000 rpm for 30 min at 4 °C and loaded onto a Ni^{2+}-charged HiTrap chelating column (Cytiva, Freiburg im Breisgau, Germany) equilibrated with buffer A. The column was washed with buffer A and eluted with a linear gradient (5 column volumes) of imidazole from 20 to 500 mM in 20 mM Tris/HCl buffer (pH 8) containing 0.5 M NaCl. The protein fractions were concentrated with a centrifugal filter unit (Merck Millipore, Darmstadt, Germany) and loaded into a HiPrep 26/10 desalting column (GE Healthcare, Uppsala, Sweden) equilibrated with 20 mM Tris/HCl buffer pH 8.0. The purity of the final sample was evaluated by 8−16% SDS-PAGE and the protein concentration was measured by the Bradford Method using bovine serum albumin as a standard.

7.3.2. Heterologous production and purification of HcgA from *M. maripaludis* in *E. coli*

E. coli C41(DE3) cells were transformed with a plasmid (pRKISC) for overexpression of [Fe-S]-cluster proteins,[104] and then further transformed with the expression plasmid pET28b(+) harboring the synthesized *M. maripaludis hcgA* gene with a N-terminal His_6-tag (GenScript), whose codon usage was optimized for expression in *E. coli* and inserted between the XhoI and NdeI restriction sites. The transformant was grown at 37 °C in Terrific Broth (1.2% Tryptone, 2.4% yeast extract, 0.5% glycerol and 89 mM potassium phosphate) supplemented with 50 µg/mL kanamycin and 10 µg/mL tetracycline. At OD_{600} = 1.0, the expression of the *hcgA* gene was induced with 1 mM IPTG and supplemented with 0.12 g/L cysteine dihydrochloride, 0.1 g/L iron (II) sulfate, 0.1 g/L iron citrate and 0.1 g/L ferric ammonium citrate (final concentrations) and cultivated for 18 hours at 20 °C. After the incubation, the cells were harvested by centrifugation using Avanti JXN-26 centrifuge with JLA-10.500 rotor at 7,300 rpm for 30 min at 4 °C and stored at –20 °C. All purification steps were performed anaerobically in the anaerobic chamber. The frozen cells containing HcgA were resuspended in 20 mM sodium phosphate buffer pH 7.4 containing 0.5 M sodium chloride, 5% glycerol and 20 mM imidazole (buffer A) and disrupted on ice by sonication for 10 min (1 min sonication, 1 min pause, five times) using a SONOPULS GM200 (Bandelin) with KE76 tip with 50% amplitude. The supernatant was collected by centrifugation in a Sorvall WX Ultra centrifuge with a T-647.5 rotor at 30,000 rpm for 30 min at 4 °C and loaded onto a Ni^{2+}-charged HiTrap chelating column (Cytiva, Freiburg im Breisgau, Germany) equilibrated with buffer A. The column was washed with buffer A and eluted with a linear gradient (10 column volumes) of imidazole from 20 to 500 mM in 20 mM sodium phosphate buffer pH 7.4 containing 0.5 M NaCl and 5% glycerol (w/v). The protein fractions were concentrated with a centrifugal filter unit (Merck Millipore, Darmstadt, Germany) and further purified by gel permeation chromatography using a HiPrep Sephacryl S-300 HR (GE Healthcare, Uppsala, Sweden) equilibrated with 50 mM MOPS/NaOH buffer pH 7.4 containing 0.5 M NaCl and 20% glycerol (w/v). The fractions

were pooled and concentrated using a 30 kDa centrifugal filter unit. The protein was frozen in liquid N_2 and stored at −20 °C.

HcgD, HcgE and HcgF from M. maripaludis were heterologously overexpressed as a His_6-tagged enzyme and purified with the methods described previously.[75,77]

7.3.3. Construction of a *M. maripaludis* strain for the expression of His_6-tagged HcgG

M. maripaludis hcgG gene was amplified with the primer pairs 0125_Fg (CCATCACATCGAAGGTCGTGGGCCCATGAAAGAACTCATAAAAAATTCATTAAATG) and 0125_Rg (TTTTATGACCTACAGATCTCCTAGGTTAAAGTAATGATACGGCATC), and cloned into pLW40neo digested with ApaI and AvrII as previously described.[154] The resulting plasmid pWL40neo*HishcgG* was transformed into *M. maripaludis Δupt* (Mm901) by the polyethylene glycol (PEG) method as previously described with PEG 8000 (Millipore Sigma).[155]

7.3.4. Homologous production and purification of HcgG from *M. maripaludis*

Cells harboring the pWL40neoHishcgG plasmid were grown in 10-litres glass fermenters with 9-litres McCas medium (Table 10)[24] under constant gas flow of $H_2/CO_2/H_2S$ (80%/20%/0.1%) at 1.5 L/min. Cells were grown to an OD_{600} between 2 and 3. The fermenters were then cooled with ice water and the cells were harvested anaerobically by a continuous-flow centrifuge equipped with a Heraeus 3049 continuous flow rotor at 15,000 rpm at 4 °C under N_2 atmosphere. The rotor was then transferred to an anaerobic chamber and the cells were resuspended in the residual medium in the rotor, then the cells were further centrifuged using an Avanti JXN-26 centrifuge with JLA-10.500 rotor at 15,000 rpm for 30 min at 4 °C. The cells were weighted and resuspended in 2-fold volume per weight of lysis buffer (50 mM Tris pH 7.4, 5 mM MgCl2, 0.5−2.5 U/mL DNAse I), aliquoted, flash frozen in liquid N_2 and stored anaerobically at −75 °C. All purification steps were done anaerobically. The cell resuspension was thawed and disrupted using a Thermo IEC FRENCH® Press with a 40K cell at 20,000 psi for five cycles. Clear cell-free extracts were obtained by centrifugation in a Sorvall WX Ultra centrifuge with a T-647.5 rotor at 30,000 rpm for 30 min at 4 °C and subsequent filtration by 0.45 µm filters. The resulting supernatant was loaded onto two 5-mL Ni^{2+}-charged HiTrap chelating column (Cytiva, Freiburg im Breisgau, Germany) connected to make a total 10 ml volume and equilibrated with 20 mM sodium phosphate buffer pH 7.4 containing 0.5 M sodium chloride, 5% glycerol (w/v) and 20 mM imidazole (buffer A). The column was washed with buffer A and eluted with a linear gradient of imidazole from 20 to 500 mM in 20 mM sodium phosphate buffer pH 7.4 containing 0.5 M NaCl and 5% glycerol (w/v). The protein fractions were concentrated with a centrifugal filter unit (Merck Millipore, Darmstadt, Germany) and loaded onto a HiPrep Sephacryl S-300 HR (GE Healthcare, Uppsala, Sweden) equilibrated with 50 mM MOPS/NaOH buffer pH 7.4 containing 0.5 M NaCl, 20%

glycerol (w/v), and 2 mM dithiothreitol. After 8−16% SDS-PAGE, the fractions containing HcgG were pooled and concentrated using a 30 kDa centrifugal filter unit. The sample was frozen in liquid N2 and stored at −75 °C.

7.3.5. Production and purification of the apo-Hmd from *M. jannaschii* in *E. coli*

The apoenzyme of Hmd from *M. jannaschii* was heterologously produced in *E. coli* BL21(DE3) as described previously.[53,84,85] The *E. coli* cells were grown in tryptone-phosphate (TP) media supplemented with 50 µg/mL kanamycin at 37 °C with agitation until OD_{600} = 0.6 and protein production was induced with 1 mM of IPTG for 2 hours. Cells were harvested by centrifugation with an Avanti JXN-26 centrifuge 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 in a 50 mM MOPS/KOH buffer pH 7 containing 1 mM DTT and disrupted on ice by sonication for 10 min using SONOPULS GM200 (Bandelin) with KE76 tip with 50 cycles. The cell debris and the unbroken cells were removed by centrifugation using an Avanti JXN-26 centrifuge with a Beckman JA-25.50 rotor at 7,300 rpm for 30 min at 4 °C. Ammonium sulfate (2 M final concentration) was slowly added to the supernatant with constant stirring. Precipitates were removed by centrifugation using an Avanti JXN-26 centrifuge with a Beckman JA-25.50 rotor at 7,300 rpm for 30 min at 4 °C. The supernatant was loaded on a Phenyl Sepharose High Performance column (75 mL, GE Healthcare Life Sciences, Solingen, Germany) and eluted with a linear gradient of ammonium sulfate from 2 M to 0 M in 50 mM MOPS/KOH buffer pH 7.0 containing 1 mM DTT. Fractions containing the apoenzyme were collected and concentrated by using Amicon Ultra-4 centrifugation filters (30-kDa cut-off). To further purify the apoenzyme, the concentrated apoenzyme sample was loaded to a HiPrep 16/60 Sephacryl S-200 HR gel filtration column (120 mL, GE Healthcare Life Sciences) using 25 mM Tris/HCl buffer pH 7.5 containing 150 mM NaCl, 5% glycerol and 2 mM DTT. To increase the purity of protein, the gel filtration repeated two times using the same conditions.[53] Finally, the purified apoenzyme was concentrated to 50–100 mg/mL and stored at −75 °C until further use.

7.3.6. Purification of native Hmd from *M. marburgensis*

All procedures were done under strict anaerobic conditions inside an anaerobic tent under yellow light or inside amber glass vials sealed with rubber stoppers. Approximately 100 g *M. marburgensis* cells were disrupted on ice by sonication using SONOPLUS GM200 (Bandelin) with VS-70-T tip using 90% amplitude and 50 cycles for 1.5 h (8 minutes of sonication with 7 minutes pause, repeat 6 times). The samples were centrifuged using a Sorvall WX Ultra centrifuge with a T-647.5 rotor at 30,000 rpm for 30 min at 4 °C to separate the soluble proteins from the membrane and cell debris. Ammonium sulfate was added to the supernatant up to 60% saturation and then gently agitated for 10 min on ice. After

incubation without agitation for 20 minutes, the sample was centrifuged using a Sorvall WX+ Ultracentrifuge with a T-647.5 rotor at 20,000 rpm for 30 min at 4 °C. The supernatant was collected and ammonium sulfate was added to 90% saturation under constant stirring for 10 min and on ice, and then incubated without agitation for 20 minutes. The sample was centrifuged using a Sorvall WX+ Ultracentrifuge with a T-647.5 rotor at 20,000 rpm for 30 min at 4 °C and the supernatant was discarded. The precipitate was then dissolved in 50 mM MOPS/KOH pH 7.0. This solution was dialyzed against 50 mM sodium citrate buffer pH 5 in a dialysis membrane with molecular size 30 kDa cut off at 8 °C overnight. After dialysis, the solution was centrifuged in a Sorvall WX+ Ultracentrifuge with a T-647.5 rotor at 20,000 rpm for 30 minutes at 4 °C. 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 equilibration buffer supplemented with 200 mM NaCl. Proteins were eluted with a linear gradient from 200 to 500 mM NaCl over 500 mL. The protein fractions containing Hmd were neutralized with 1 M MOPS pH 7 and NaOH solution. The fractions were then pooled, concentrated, frozen in liquid N_2 and stored at −75 °C.

7.4. Extraction of the FeGP cofactor from native Hmd

7.4.1. Extraction with methanol, ammonia and 2-mercaptoethanol

To extract the FeGP cofactor from the natively purified [Fe]-hydrogenase enzyme or cell free extract from *M. marburgensis*, 99.8% methanol, 100 mM 2-mercaptoethanol and 32% ammonia solutions were added to the Hmd solution to make the final concentration of 4 mg/ml (or lower) of protein(s), 60% methanol, 1 mM 2-mercaptoethanol and 1 % v/v NH_3. The sample was incubated at 40 °C for 15 minutes and denatured protein was removed by ultrafiltration (10 kDa cut-off) under anaerobic conditions. The sample was evaporated by vacuum at 4 °C, the dried material was resuspended in 10 mM ammonium carbonate pH 9.0 containing 1 mM 2-mercaptoethanol. The cofactor solution was stored in amber bottles, frozen in liquid N_2 and stored in Dry Shippers (CX-100).

7.4.2. Acetic acid extraction

The native purified Hmd solution or the cell free extract from *M. marburgensis* was diluted to 1 mg/mL or less and an equal volume of acetic acid was mixed in. The precipitate was removed by centrifugation at 14,000 rpm for 10 minutes. The remaining precipitates were removed by filtration with a 10 kDa cut-off filter. The filtrate was dried by evaporation at 4 °C and resuspended in 50 mM sodium acetate pH 4.6. The solution was aliquoted in amber bottles and kept under liquid N_2 in Dry Shippers (CX-100).

7.4.3. Extraction and purification of the FeGP cofactor and GP from *M. marburgensis*

For GP extraction, semi-purified Hmd was used (after dialysis step, see above). The FeGP cofactor was extracted first with the methanol/ammonia/mercaptoethanol method. However, for the preparation of GP, the 10 kDa filtration was performed aerobically under room light to decompose the FeGP cofactor. The filtrate was evaporated until dry at room temperature and resuspended in 0.01% NH_3. This solution was applied to 5-mL Q-trap HP columns (GE healthcare), in which three 5-mL columns were connected to make total a 15-mL column volume. These connected columns were equilibrated with 0.01% NH_3. After washing with 5 column volumes (CV) of 0.01% NH_3, the compounds were eluted with a 20 CV gradient from 0 to 1 M NaCl in 1% NH_3. Elution of **3** was confirmed by HPLC using Synergi 4m Polar RP 80A column (250 mm D 4.6 mm, Phenomenex, Aschaffenburg) using a acidified water (pH 3 by HCl)/methanol gradient. For further purification, the same HPLC procedure was repeated. The purified compound was dried by evaporation at 4 °C and then dissolved in water. The final purity was evaluated by HPLC and UV-visible spectrophotometry, and stored at −20 °C.

7.5. Enzyme assays

7.5.1. Hmd activity assay

H_4MPT and methenyl-H_4MPT$^+$ were purified as previously described.[156] Methylene-H_4MPT was synthesized from H_4MPT by reaction with formaldehyde.[84] As a routine Hmd activity assay, formation of methenyl-H_4MPT$^+$ from methylene-H_4MPT was measured as an increase of absorbance at 336 nm. The reaction was performed in 120 mM potassium phosphate buffer pH 6.0 containing 1 mM EDTA in a 1-mL quartz cuvette containing 700 µL of assay mixture under a 100% N_2 atmosphere at 40 °C. Measurement of the reverse reaction: reduction of methenyl-H_4MPT$^+$ to methylene-H_4MPT was performed in 120 mM potassium phosphate pH 7.5 containing 1 mM EDTA under a 100% H_2 atmosphere at 40 °C. The concentration of the substrate was 20 µM and the reaction was started by adding 10 µL of the enzyme solutions, which are the *in vitro* biosynthesis products or the purified enzyme. The activity was calculated by using the extinction coefficient of methenyl-H_4MPT$^+$ (ε_{336} = 21.6 $mM^{-1}cm^{-1}$).[157,158] One unit of the enzyme activity is defined as the amount of the enzyme producing or consuming 1 µmol of methenyl-H_4MPT$^+$ per one minute in the assays.

7.5.2. Determination of the FeGP cofactor by measuring reconstituted Hmd

To determine the amount of the FeGP cofactor in the samples, the Hmd holoenzyme was reconstituted in the activity assay cuvette prior to the addition of the reaction substrates. In 700 µL assay buffer in a quartz cuvette, 10 µL of 1 mg/mL of apoenzyme was added and then the FeGP cofactor sample was added. In the presence of excess amount of the apoenzyme, the FeGP cofactor is almost immediately captured by the apoenzyme to reconstitute the holoenzyme. The reaction was

started by addition of 20 µM of the substrates and the change of absorbance at 336 nm was assayed as described above. The sample containing the FeGP cofactor was diluted with 10 mM ammonium carbonate buffer pH 9 supplemented with 1 mM 2-mercaptoethanol.

For the kinetic study of the reconstitution of the FeGP cofactor (Section 5.9), the order of addition of the component was altered: 20 µM of substrate was added first, then apoenzyme was added and finally the reaction was started by adding the FeGP cofactor. The specific activity increased in a time dependent manner. I calculated the binding constants from the kinetics of the specific activity in this assay.

7.5.3. Production of 6-carboxymethyl-4-hydroxy-2-pyridinol (Precursor **1**) by HcgA

To prepare the protein-free extract, proteins in the cell free extract were removed by filtrating the cell free extract with a 3 kDa cut-off filter. The protein free extract was incubated under strict anaerobic conditions with 5 mM sodium dithionite, 2.5 mM *S*-adenosyl methionine and 20 µM HcgA at 37 °C for 1 hour or the time period described in the figures. After incubation, 50 µL of the sample were quenched with 80% MeOH (final concentration) and incubated at 40 °C for 15 minutes. The sample was filtered with a 0.22 µm filter to remove precipitated proteins and then dried using a Heraeus Centrivac aerobically. After the sample was dried, it was resuspended in 50 µL of distilled water. The solutions were stored at −20 °C until analysis.

7.6. *In vitro* biosynthesis of the FeGP cofactor

The *in vitro* biosynthesis of the FeGP cofactor was performed as previously described.[102,103] In short, 200 µL cell extract of *M. maripaludis* was incubated with 1 mM $Fe(SO_4)_2(NH_4)_2$, 1 mM DTT, 2 mM sodium dithionite, 5 mM $MgCl_2$, 2 mM SAM, 5 mM ATP, 10 µM precursors and 10 µM [Fe]-hydrogenase apoenzyme (final concentrations). The solution was transferred to a vial sealed a rubber stopper and containing 50% H_2/50% CO or otherwise described atmosphere. The reaction mixtures were incubated at 40 °C for 1 hour or at 20 °C for 6 hours, and the Hmd activity was determined as previously described (See section 7.5.1).

7.7. Mass spectrometric methods

7.7.1. Proteome analysis

Cell pellets were lysed and reduced by 5 mM tris(2-carboxyethyl)phosphine (TCEP) in the presence of 2% deoxycholate (DOC) at 90 °C for 10 minutes. After that, this solution was incubated at 25 °C for 30 minutes in 100 mM ammonium bicarbonate pH 8.2 and 10 mM iodoacetamide (IAA) and then digested overnight at 30 °C with trypsin (MS approved) from Serva. Before LC-MS analysis, samples were

desalted using C18 microspin columns (Nest Group) according to the manufacturer's instructions. Dried and reconstituted peptides were then analyzed using liquid-chromatography-mass spectrometry carried out on a Orbitrap Exploris 480 instrument connected to an Ultimate 3000 RSLC nano and a nanospray ion source (Thermo Scientific). Peptide separation was performed on a reverse phase HPLC column (75 µm x 42 cm) packed with C18 resin (2.4 µm; Dr. Maisch) with a 90-minute 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).

7.7.2. Determination of metabolites and cofactors

Determination of the FeGP cofactor biosynthesis precursors were performed using two different types of HRES-LC-MS. The chromatographic separation was performed using a Kinetex EVO C18 column (150 × 1.7 mm, 3 µm particle size, 100 Å pore size, Phenomenex) connected to a guard column of similar specificity (20 × 2.1 mm, 5 µm particle size, Phenomoenex) a constant flow rate of 0.2 mL/min with mobile phase A being 0.1% Formic acid in water and phase B being 0.1% formic acid methanol (Honeywell, Morristown, New Jersey, USA) at 40 °C. The injection volume was 5 µL. The mobile phase profile consisted of the following steps and linear gradients: 0 – 0.5 min constant at 5% B; 0.5 – 4 min from 5 to 90% B; 4 – 5 min constant at 90% B; 5 – 5.1 min from 90 to 5% B; 5.1 – 10 min constant at 5% B.

For the data obtained in QTOF, an Agilent 6550 ion funnel Q-ToF mass spectrometer was used in positive full scan mode (50 – 1200 m/z [Da/eV]) applying electrospray ionization, and the following settings: ESI spray voltage 1000 V, nozzle voltage 200 V, sheath gas 150° C at 8 L/min, nebulizer pressure 50 psi and drying gas 200 °C at 11 L/min.

For all other determinations, a Thermo Scientific I-DX Orbitrap mass spectrometer was used. Ionization was performed using a high temperature electro spray ion source at a static spray voltage of 3300 V, Sheath gas at 50 (Arb), Auxiliary Gas at 25 (Arb), and Ion transfer tube and Vaporizer at 325 and 300 °C.

Data dependent MS/MS measurements were conducted applying an Orbitrap mass resolution of 120000 using quadrupole isolation in a mass range of 100 – 600 and combining it with a high energy collision dissociation (HCD). HCD was performed on the five most abundant ions per scan with a relative collision energy of 15%. Fragments were detected using the Orbitrap mass analyzer at a predefined mass resolution of 30 000. Dynamic exclusion with and exclusion duration of 2.5 seconds after 1 scan with a mass tolerance of 10 ppm was used to increase coverage.

7.7.3. Determination of the FeGP cofactor

In vitro biosynthesis of the FeGP cofactor biosynthesis was performed as previously described.[102] Afterwards, the solution was diluted to <1 mg/mL and diluted with 1 volume of acetic acid. The precipitate was removed by centrifugation at 14,000 rpm for 10 minutes in a 1.5 mL Eppendorf tube. The remaining precipitates were removed by filtration with a 10 kDa cut-off filter. The filtrate was dried by evaporation at 4 °C and resuspended distilled water and analyzed via ESI-MS.

The chromatographic separation was performed on an Thermo Scientific Vanquish HPLC System using a polymer based ZICpHilic (Sequant, 150 × 2.1mm, 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 ammonium hydroxide 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 – 3min 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. Ionization was performed using a high temperature electro spray ion source at a static spray voltage of 3300 V (-), Sheath gas at 35 (Arb), Auxiliary 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 630−645. The data was saved in full Profile mode. Targeted fragmentations measurements were performed at similar chromatography and ionization 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.

7.8. Computational methods

7.8.1. Generation of protein models

Protein models were generated using the default settings with the web servers from Robetta (https://robetta.bakerlab.org/),[86,118] I-TASSER (https://zhanggroup.org/I-TASSER/),[111] RaptorX (http://raptorx.uchicago.edu/),[108] and SWISS-MODEL (https://swissmodel.expasy.org/).[110]

Protein models were visualized with PyMOL 2.4.1 software (Schrödinger, LLC).

7.8.2. Simulation of enzyme kinetics data

Michaelis-Menten kinetic parameters were obtained by simulation of the substrate consumption and product formation by numerical integration of the equations derived from mass action kinetics. The reconstitution rate with the FeGP cofactor and simulated enzymatic activities were calculated by numerical integration using the equations shown in Figure 57. All simulations were coded in Python 3.7 using Spyder 4.1 development environment and the following libraries: SciPy,[159] NumPy,[160] Matplotlib[161] and pandas.[162]

8. Supplementary information

8.1. Protocol for production and purification of an active form of HcgA

Buffer 3A (1 liters) pH 7.4

- **20 mM $Na(PO)_4$**
  - 15.48 mL 1M Na_2HPO_4
  - 4.52 mL 1M NaH_2PO_4
- **0.5 M NaCl**
  - 29.22 g NaCl
- **20 mM Imidazole**
  - 1.36 g Imidazole
- **5% Glycerol**
  - 50 g Glycerol

Buffer 3B (1 liter) pH 7.4

- **20 mM $Na(PO)_4$**
  - 15.48 mL 1M Na_2HPO_4
  - 4.52 mL 1M NaH_2PO_4
- **0.5 M NaCl**
  - 29.22 g NaCl
- **500 mM Imidazole**
  - 34.04 g Imidazole
- **5% Glycerol**
  - 50 g Glycerol

Buffer 3Cv2 (1 liter) pH 7.4

- **50 mM MOPS**
  - 10.46 g MOPS
- **0.5 M NaCl**
  - 29.22 g NaCl
- **20% Glycerol**
  - 200 g Glycerol
- **2 mM DTT**
  - 300 mg

- Grow *E. coli* in 200 mL LB with Kanamycin and Tetracycline (2-fold higher concentration) as pre inoculum.
- Prepare TB media as described, add Tetracycline and Kanamycin (1-fold)
- Inoculate 2 liters of media with 200 mL of the preculture
- Grow until OD = 0.8−1 at 700 rpm 37°C
- Add as powder:
  0.12 g/l cysteine
  0.1 g/l iron ammonium citrate
  0.1 g/l iron citrate
  0.1 g/l iron sulfate
- Grow until OD = 1; induce with 1 mM IPTG
- Reduce speed to 200 RPM
- Stir at 20°C overnight
- Cool down cells and harvest
- Resuspend cells in Buffer 3A (5 to 10-fold) and break by sonication
- Spin down cell debris (100,000 x g).
- Equilibrate HisTrap column with Buffer 3A.
- Load cell extract into column
- Wash with Buffer 3A
- Elute with linear gradient with increasing Buffer 3B concentration.
- Analyze fractions via SDS-PAGE.
- Collect fractions and concentrate.
- Desalt with Desalting Column equilibrated with Buffer 3Cv2.
- If needed, size exclusion with Buffer 3Cv2

8.2. Protocol for production and purification of an active form of HcgG

Buffer 3A (1 liters) pH 7.4

- **20 mM $Na(PO)_4$**
  - 15.48 mL 1M Na_2HPO_4
  - 4.52 mL 1M NaH_2PO_4
- **0.5 M NaCl**
  - 29.22 g NaCl
- **20 mM Imidazole**
  - 1.36 g Imidazole
- **5% Glycerol**
  - 50 g Glycerol

Buffer 3B (1 liter) pH 7.4

- **20 mM $Na(PO)_4$**
  - 15.48 mL 1M Na_2HPO_4
  - 4.52 mL 1M NaH_2PO_4
- **0.5 M NaCl**
  - 29.22 g NaCl
- **500 mM Imidazole**
  - 34.04 g Imidazole
- **5% Glycerol**
  - 50 g Glycerol

Buffer 3Cv2 (1 liter) pH 7.4

- **50 mM MOPS**
  - 10.46 g MOPS
- **0.5 M NaCl**
  - 29.22 g NaCl
- **20% Glycerol**
  - 200 g Glycerol
- **2 mM DTT**
  - 300 mg

- Grow *M. maripaludis* in 10-liter fermenters are described in Materials and Methods (See 7.3.2)
- Harvest cells anaerobically and resuspend in 50 mM Tris buffer pH 7.4, 5 mM $MgCl_2$, 1−2 U/mL DNAse I.
- Break cells by anaerobic French press.
- Spin down cell debris (100,000 x g).
- Equilibrate HisTrap column with Buffer 3A.
- Load cell extract into column.
- Wash with Buffer 3A.
- Elute with linear gradient with increasing Buffer 3B concentration.
- Analyze fractions via SDS-PAGE.
- Collect fractions and concentrate.
- Perform size exclusion chromatography with Buffer 3Cv2.

8.3. Full list of compounds that increase over time in the HcgA reaction

Index	Name	Calculated MW	Mass error	Retention time
20	Adenine	135.0543	-1.48	3.143
30	2'-Deoxyadenosine	251.1012	-2.64	3.145
50		182.0687	-2.43	3.108
92	Precursor 1	183.0528	-1.71	4.61
96		283.0735	-1.43	4.219
119	AM9627000	165.0422	-2.19	3.107
131	2,3-dihydro-3-hydroxyanthranilic acid	155.0578	-2.77	3.106
176		209.0719	-1.27	3.331
241		136.0512		3.131
271	4-Aminophenol	109.0526	-1.95	4.105
274		329.122	4.77	4.249
303		208.0481	-1.47	4.8
316	2,3-dihydro-3-hydroxyanthranilic acid	155.0578	-2.72	4.611
321		344.1264	0.52	1.74
333		208.0481	-1.67	4.24
336		344.1262	-1.32	1.654
364		265.0983		2.036
373	2,3-dihydro-3-hydroxyanthranilic acid	155.0578	-2.84	4.109
383	AM9627000	165.0423	-1.55	4.611
386		278.1474	1.62	1.525
387		240.0199	-2.25	5.199
424		268.1055	1.64	3.421
448		286.1161	-1.21	3.423
450		309.0959	-0.48	4.798
482	Triethylamine	101.1202	-2.29	1.665
500		257.0896	-1.49	4.57
549	4-Aminobenzoic acid	137.0475	-1.62	4.612
557	noradrenaline	169.0736	-1.59	4.614
580		297.1066	-0.8	2.55
583		118.0276	-2.83	3.115
590		283.1086	-2.51	1.649
616	noradrenaline	169.0736	-1.71	3.105
635		265.098	3.71	4.157
643		247.0874	-1.87	1.651
703	gaboxadol	140.0583	-1.72	3.5
745		228.0562	-1.31	4.762
822		208.0482	3.73	2.706
937	DL-Homocysteine	135.0354	-0.02	3.138
952		208.0481	2.71	5.638
966		208.0479	1.85	6.13
1019		135.073		3.153
1052		208.0481	2.38	5.853

8.4. Full list of compounds that decrease over time in the HcgA reaction

Index	Name	Calculated MW	Mass error	Retention time
9		209.0719	2.16	1.583
23		231.0539	1.71	1.569
25	Methyl alpha-aspartylphenylalaninate	294.1212	0.26	5.007
32		208.0481	2.13	5.006
42	Dinoseb	240.0743	3.44	5.007
44		209.0719	-0.76	1.75
65	dihydroxyphenylalanine	197.0682	3.24	5.279
70	Triethanolamine	149.1049	4.41	1.504
103	Vigabatrin	129.0788	0.15	1.753
113		247.0278	-1.13	1.571
121		171.0869	-0.75	1.48
142	AM9627000	165.0424	-1.78	5.006
175		133.9858	2.65	5.457
184		233.0495	4.95	1.565
190	glycerin	92.04727	0.21	1.585
193		163.9991	-0.74	1.237
198		145.9884		1.239
204	Acetylcadaverine	144.126	-1.31	1.626
206	2,3-dihydro-3-hydroxyanthranilic acid	155.058	0.12	5.007
217		277.0947	-3.83	5.007
220		216.0971	-1.12	1.486
252	Thiirane	60.00326	-0.82	1.761
281	noradrenaline	169.0736	-1.65	5.007
285	AM9627000	165.0424	-1.2	5.279
289	4-Aminobenzoic acid	137.0475	-1.61	5.007
290	Benzoyldehydro-2,3-dihydroxy-benzone	231.0892	-1.85	5.007
334	2,3-dihydro-3-hydroxyanthranilic acid	155.0579	-1.09	5.279
381		201.0142	-3.03	1.484
396	L-coprine	202.0951	-2.03	1.8
403		253.0355	2.74	1.536
414	noradrenaline	169.0736	1.57	5.28
427		185.0837	3.04	5.007
505		418.1443	-2.05	1.582
506		440.126	-1.3	1.575
530		136.0042	-0.91	1.238
539		348.1331	-1.13	1.571
540	dazopride	326.151	-1.35	1.577
543		224.0023	-1.68	1.238
553	Ethanal tetramer	176.1045	-1.16	5.004
559		215.9151	-1.51	1.464
574	4-Aminobenzoic acid	137.0475	-1.34	5.28
597		117.9936	-1.54	1.241
618		199.9653	3.82	1.246
754	tolonidine	209.0716	-2.28	4.037
755		196.1095	-2.91	5.803
760		103.9753	-1.35	5.461
784		208.0481	-2.02	4.51
801		240.1332	-1.87	6.737
821		205.8864	-1.45	1.474

851		207.9729	-0.51	3.714
866		362.0603	-1.76	5.375
920		355.2718	-1.82	5.922
984		277.0978	3.75	4.931
985		275.0999	3.16	4.929
1018		209.0686	-1.26	5.578
1032	4-Mercapto-5-methyl-3(2H)-thiophenone	145.9857	3.75	5.042
1039	tolonidine	209.0717	-1.45	6.531
1078		261.103	-1.07	6.397

9. References

[1] D. R. Boone, W. B. Whitman, P. Rouvière, in *Methanogenesis Ecol. Physiol. Biochem. Genet.* (Ed.: J.G. Ferry), Springer US, Boston, MA, **1993**, pp. 35–80.

[2] R. K. Thauer, A.-K. Kaster, H. Seedorf, W. Buckel, R. Hedderich, *Nat. Rev. Microbiol.* **2008**, *6*, 579–591.

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

[4] D. Mayumi, H. Mochimaru, H. Tamaki, K. Yamamoto, H. Yoshioka, Y. Suzuki, Y. Kamagata, S. Sakata, *Science* **2016**, *354*, 222–225.

[5] S. H. Zinder, in *Methanogenesis Ecol. Physiol. Biochem. Genet.* (Ed.: J.G. Ferry), Springer US, Boston, MA, **1993**, pp. 128–206.

[6] W. F. Fricke, H. Seedorf, A. Henne, M. Krüer, H. Liesegang, R. Hedderich, G. Gottschalk, R. K. Thauer, *J. Bacteriol.* **2006**, *188*, 642–658.

[7] S. Shima, G. Huang, T. Wagner, U. Ermler, *Annu. Rev. Microbiol.* **2020**, *74*, 713–733.

[8] J. A. Vorholt, R. K. Thauer, *Eur. J. Biochem.* **1997**, *248*, 919–924.

[9] T. Wagner, U. Ermler, S. Shima, *Science* **2016**, *354*, 114–117.

[10] A. A. DiMarco, K. A. Sment, J. Konisky, R. S. Wolfe, *J. Biol. Chem.* **1990**, *265*, 472–476.

[11] A. A. DiMarco, M. I. Donnelly, R. S. Wolfe, *J. Bacteriol.* **1986**, *168*, 1372–1377.

[12] B. W. te Brömmelstroet, C. M. H. Hensgens, J. T. Keltjens, C. van der Drift, G. D. Vogels, *Biochim. Biophys. Acta BBA - Gen. Subj.* **1991**, *1073*, 77–84.

[13] C. Zirngibl, R. Hedderich, R. k. Thauer, *FEBS Lett.* **1990**, *261*, 112–116.

[14] C. Zirngibl, W. Van Dongen, B. Schwörer, R. Von Bünau, M. Richter, A. Klein, R. K. Thauer, *Eur. J. Biochem.* **1992**, *208*, 511–520.

[15] G. Huang, T. Wagner, U. Ermler, S. Shima, *Nat. Rev. Chem.* **2020**, *4*, 213–221.

[16] K. Ma, R. K. Thauer, *Eur. J. Biochem.* **1990**, *191*, 187–193.

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

[18] G. Gottschalk, R. K. Thauer, *Biochim. Biophys. Acta BBA - Bioenerg.* **2001**, *1505*, 28–36.

[19] K. Schlegel, V. Müller, *Biochem. Soc. Trans.* **2013**, *41*, 421–426.

[20] R. K. Thauer, *Biochemistry* **2019**, *58*, 5198–5220.

[21] W. Buckel, R. K. Thauer, *Biochim. Biophys. Acta BBA - Bioenerg.* **2013**, *1827*, 94–113.

[22] W. Buckel, R. K. Thauer, *Chem. Rev.* **2018**, *118*, 3862–3886.

[23] T. Wagner, J. Koch, U. Ermler, S. Shima, *Science* **2017**, *357*, 699–703.

[24] K. C. Costa, P. M. Wong, T. Wang, T. J. Lie, J. A. Dodsworth, I. Swanson, J. A. Burn, M. Hackett, J. A. Leigh, *Proc. Natl. Acad. Sci.* **2010**, *107*, 11050–11055.

[25] T. Watanabe, O. Pfeil-Gardiner, J. Kahnt, J. Koch, S. Shima, B. J. Murphy, *Science* **2021**, *373*, 1151–1156.

[26] W. Lubitz, H. Ogata, O. Rüdiger, E. Reijerse, *Chem. Rev.* **2014**, *114*, 4081–4148.

[27] P. M. Vignais, B. Billoud, *Chem. Rev.* **2007**, *107*, 4206–4272.

[28] P. M. Vignais, B. Billoud, J. Meyer, *FEMS Microbiol. Rev.* **2001**, *25*, 455–501.

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

[30] J. C. Fontecilla-Camps, A. Volbeda, C. Cavazza, Y. Nicolet, *Chem. Rev.* **2007**, *107*, 4273–4303.

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

[32] J. W. Peters, W. N. Lanzilotta, B. J. Lemon, L. C. Seefeldt, *Science* **1998**.

[33] D. W. Mulder, E. M. Shepard, J. E. Meuser, N. Joshi, P. W. King, M. C. Posewitz, J. B. Broderick, J. W. Peters, *Structure* **2011**, *19*, 1038–1052.

[34] 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.

[35] T. Happe, J. D. Naber, *Eur. J. Biochem.* **1993**, *214*, 475–481.
[36] A. Silakov, C. Kamp, E. Reijerse, T. Happe, W. Lubitz, *Biochemistry* **2009**, *48*, 7780–7786.
[37] S. Shima, R. K. Thauer, *Chem. Rec.* **2007**, *7*, 37–46.
[38] R. K. Thauer, A. R. Klein, G. C. Hartmann, *Chem. Rev.* **1996**, *96*, 3031–3042.
[39] A. Berkessel, R. K. Thauer, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2247–2250.
[40] C. Afting, E. Kremmer, C. Brucker, A. Hochheimer, R. K. Thauer, *Arch. Microbiol.* **2000**, *174*, 225–232.
[41] E. L. Hendrickson, J. A. Leigh, *J. Bacteriol.* **2008**, *190*, 4818–4821.
[42] C. Afting, A. Hochheimer, R. K. Thauer, *Arch. Microbiol.* **1998**, *169*, 206–210.
[43] 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.
[44] O. Pilak, B. Mamat, S. Vogt, C. H. Hagemeier, R. K. Thauer, S. Shima, C. Vonrhein, E. Warkentin, U. Ermler, *J. Mol. Biol.* **2006**, *358*, 798–809.
[45] G. Buurman, S. Shima, R. K. Thauer, *FEBS Lett.* **2000**, *485*, 200–204.
[46] E. J. Lyon, S. Shima, G. Buurman, S. Chowdhuri, A. Batschauer, K. Steinbach, R. K. Thauer, *Eur. J. Biochem.* **2004**, *271*, 195–204.
[47] S. Shima, E. J. Lyon, R. K. Thauer, B. Mienert, E. Bill, *J. Am. Chem. Soc.* **2005**, *127*, 10430–10435.
[48] E. J. Lyon, S. Shima, R. Boecher, R. K. Thauer, F. W. Grevels, E. Bill, W. Roseboom, S. P. J. Albracht, *J. Am. Chem. Soc.* **2004**, *126*, 14239–14248.
[49] M. Korbas, S. Vogt, W. Meyer-Klaucke, E. Bill, E. J. Lyon, R. K. Thauer, S. Shima, *J. Biol. Chem.* **2006**, *281*, 30804–30813.
[50] 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.* **2004**, *43*, 2547–2551.
[51] 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.
[52] S. Shima, U. Ermler, *Eur. J. Inorg. Chem.* **2011**, 963–972.
[53] G. Huang, T. Wagner, M. D. Wodrich, K. Ataka, E. Bill, U. Ermler, X. Hu, S. Shima, *Nat. Catal.* **2019**, *2*, 537–543.
[54] M. J. Lacasse, D. B. Zamble, *Biochemistry* **2016**, *55*, 1689–1701.
[55] L. Forzi, P. Hellwig, R. K. Thauer, R. G. Sawers, *FEBS Lett.* **2007**, *581*, 3317–3321.
[56] O. Lenz, I. Zebger, J. Hamann, P. Hildebrandt, B. Friedrich, *FEBS Lett.* **2007**, *581*, 3322–3326.
[57] S. Reissmann, E. Hochleitner, H. Wang, A. Paschos, F. Lottspeich, R. S. Glass, A. Böck, *Science* **2003**, *299*, 1067–1070.
[58] I. Bürstel, E. Siebert, S. Frielingsdorf, I. Zebger, B. Friedrich, O. Lenz, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 14722–14726.
[59] B. Soboh, S. T. Stripp, C. Bielak, U. Lindenstrauß, M. Braussemann, M. Javaid, M. Hallensleben, C. Granich, M. Herzberg, J. Heberle, R. G. Sawers, *FEBS Lett.* **2013**, *587*, 2512–2516.
[60] B. Soboh, S. T. Stripp, E. Muhr, C. Granich, M. Braussemann, M. Herzberg, J. Heberle, R. Gary Sawers, *FEBS Lett.* **2012**, *586*, 3882–3887.
[61] C. Pinske, F. Sargent, R. G. Sawers, *Metallomics* **2015**, *7*, 683–690.
[62] M. Khorasani-Motlagh, M. Noroozifar, K. Kerman, D. B. Zamble, *BioMetals* **2019**, *32*, 521–532.
[63] M. Senger, S. T. Stripp, B. Soboh, *J. Biol. Chem.* **2017**, *292*, 11670–11681.
[64] M. Bortolus, P. Costantini, D. Doni, D. Carbonera, *Int. J. Mol. Sci.* **2018**, *19*, DOI 10.3390/ijms19103118.
[65] A. Pagnier, L. Martin, L. Zeppieri, Y. Nicolet, J. C. Fontecilla-Camps, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 104–109.
[66] 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.
[67] J. M. Kuchenreuther, W. K. Myers, D. L. M. 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.

[68] J. N. Betz, N. W. Boswell, C. J. Fugate, G. L. Holliday, E. Akiva, A. G. Scott, P. C. Babbitt, J. W. Peters, E. M. Shepard, J. B. Broderick, *Biochemistry* **2015**, *54*, 1807–1818.
[69] L. Tao, S. A. Pattenaude, S. Joshi, T. P. Begley, T. B. Rauchfuss, R. D. Britt, *J. Am. Chem. Soc.* **2020**, *142*, 10841–10848.
[70] 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.* **2022**, *n/a*, e202203413.
[71] M. Schick, X. Xie, K. Ataka, J. Kahnt, U. Linne, S. Shima, *J. Am. Chem. Soc.* **2012**, *134*, 3271–3280.
[72] T. J. Lie, K. C. Costa, D. Pak, V. Sakesan, J. A. Leigh, *FEMS Microbiol. Lett.* **2013**, *343*, 156–160.
[73] 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.
[74] L. Bai, T. Wagner, T. Xu, X. Hu, U. Ermler, S. Shima, *Angew. Chem. Int. Ed.* **2017**, *56*, 10806–10809.
[75] T. Fujishiro, U. Ermler, S. Shima, *FEBS Lett.* **2014**, *588*, 2789–2793.
[76] T. Fujishiro, H. Tamura, M. Schick, J. Kahnt, X. Xie, U. Ermler, S. Shima, *Angew. Chem. Int. Ed.* **2013**, *52*, 12555–12558.
[77] T. Fujishiro, J. Kahnt, U. Ermler, S. Shima, *Nat. Commun.* **2015**, *6*, 6895.
[78] T. Fujishiro, L. Bai, T. Xu, X. Xie, M. Schick, J. Kahnt, M. Rother, X. Hu, U. Ermler, S. Shima, *Angew. Chem. Int. Ed.* **2016**, *55*, 9648–9651.
[79] S. Shima, M. Schick, J. Kahnt, K. Ataka, K. Steinbach, U. Linne, *Dalton Trans.* **2011**, *41*, 767–771.
[80] L. Bai, Charakterisierung Der an Der Biosynthese Des Cofaktors Der [Fe] -Hydrogenase Hmd Beteiligten Hcg-Proteine, PhD Thesis, Philipps-Universität Marburg, **2017**.
[81] 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.
[82] D. L. M. Suess, J. M. Kuchenreuther, L. De La Paz, J. R. Swartz, R. D. Britt, *Inorg. Chem.* **2016**, *55*, 478–487.
[83] J. B. Broderick, B. R. Duffus, K. S. Duschene, E. M. Shepard, *Chem. Rev.* **2014**, *114*, 4229–4317.
[84] S. Shima, M. Schick, H. Tamura, in *Methods Enzymol.* (Eds.: A.C. Rosenzweig, S.W. Ragsdale), Academic Press, **2011**, pp. 119–137.
[85] G. Huang, F. J. Arriaza-Gallardo, T. Wagner, S. Shima, *Inorganics* **2020**, *8*, 50.
[86] Y. Song, F. DiMaio, R. Y.-R. Wang, D. Kim, C. Miles, T. Brunette, J. Thompson, D. Baker, *Struct. Lond. Engl. 1993* **2013**, *21*, 1735–1742.
[87] S. J. Booker, T. L. Grove, *F1000 Biol. Rep.* **2010**, *2*, 52.
[88] Y. Nicolet, *Nat. Catal.* **2020**, *3*, 337–350.
[89] M. J. Hiscox, R. C. Driesener, P. L. Roach, *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2012**, *1824*, 1165–1177.
[90] B. M. Hover, A. Loksztejn, A. A. Ribeiro, K. Yokoyama, *J. Am. Chem. Soc.* **2013**, *135*, 7019–7032.
[91] J. A. Wiig, Y. Hu, C. C. Lee, M. W. Ribbe, *Science* **2012**, *337*, 1672–1675.
[92] K. Yokoyama, E. A. Lilla, *Nat. Prod. Rep.* **2018**, *35*, 660–694.
[93] J. K. Rubach, X. Brazzolotto, J. Gaillard, M. Fontecave, *FEBS Lett.* **2005**, *579*, 5055–5060.
[94] E. Akiva, S. Brown, D. E. Almonacid, A. E. Barber 2nd, A. F. Custer, M. A. Hicks, C. C. Huang, F. Lauck, S. T. Mashiyama, E. C. Meng, D. Mischel, J. H. Morris, S. Ojha, A. M. Schnoes, D. Stryke, J. M. Yunes, T. E. Ferrin, G. L. Holliday, P. C. Babbitt, *Nucleic Acids Res.* **2014**, *42*, D521–D530.
[95] P. Dinis, D. L. M. Suess, S. J. Fox, J. E. Harmer, R. C. Driesener, L. De La Paz, J. R. Swartz, J. W. Essex, R. D. Britt, P. L. Roach, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 1362–1367.
[96] J. M. Kuchenreuther, J. A. Stapleton, J. R. Swartz, *PLoS ONE* **2009**, *4*, DOI 10.1371/journal.pone.0007565.
[97] 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.* **2010**, *49*, 1687–1690.
[98] M. C. Posewitz, P. W. King, S. L. Smolinski, L. Zhang, M. Seibert, M. L. Ghirardi, *J. Biol. Chem.* **2004**, *279*, 25711–25720.

[99] Y. Nicolet, P. Amara, J.-M. Mouesca, J. C. Fontecilla-Camps, *Proc. Natl. Acad. Sci.* **2009**, *106*, 14867–14871.
[100] Y. Nicolet, J. K. Rubach, M. C. Posewitz, P. Amara, C. Mathevon, M. Atta, M. Fontecave, J. C. Fontecilla-Camps, *J. Biol. Chem.* **2008**, *283*, 18861–18872.
[101] R. D. Britt, G. Rao, L. Tao, *Chem. Sci.* **2020**, *11*, 10313–10323.
[102] S. Schaupp, F. J. Arriaza-Gallardo, H. Pan, J. Kahnt, G. Angelidou, N. Paczia, K. Costa, X. Hu, S. Shima, *Angew. Chem. Int. Ed.* **2022**, *61*, e202200994.
[103] S. Schaupp, In Vitro Biosynthesis of the FeGP Cofactor of the [Fe]-Hydrogenase, PhD Thesis, Philipps-Universität Marburg, **2022**.
[104] M. Nakamura, K. Saeki, Y. Takahashi, *J. Biochem. (Tokyo)* **1999**, *126*, 10–18.
[105] 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.
[106] N. B. Ugulava, B. R. Gibney, J. T. Jarrett, *Biochemistry* **2000**, *39*, 5206–5214.
[107] E. L. McCarthy, S. J. Booker, in *Methods Enzymol.*, Elsevier, **2018**, pp. 217–239.
[108] M. Källberg, H. Wang, S. Wang, J. Peng, Z. Wang, H. Lu, J. Xu, *Nat. Protoc.* **2012**, *7*, 1511–1522.
[109] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, *Nucleic Acids Res.* **2018**, *46*, W296–W303.
[110] N. Guex, M. C. Peitsch, T. Schwede, *ELECTROPHORESIS* **2009**, *30*, S162–S173.
[111] A. Roy, A. Kucukural, Y. Zhang, *Nat. Protoc.* **2010**, *5*, 725–738.
[112] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, *Nat. Methods* **2015**, *12*, 7–8.
[113] L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, *Nat. Protoc.* **2015**, *10*, 845–858.
[114] L. Deng, G. Zhong, C. Liu, J. Luo, H. Liu, *BMC Bioinformatics* **2019**, *20*, 662.
[115] R. M. Twyman, in *Encycl. Neurosci.* (Ed.: L.R. Squire), Academic Press, Oxford, **2009**, pp. 1013–1017.
[116] C. Notredame, D. G. Higgins, J. Heringa, *J. Mol. Biol.* **2000**, *302*, 205–217.
[117] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583–589.
[118] M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch, R. D. Schaeffer, C. Millán, H. Park, C. Adams, C. R. Glassman, A. DeGiovanni, J. H. Pereira, A. V. Rodrigues, A. A. van Dijk, A. C. Ebrecht, D. J. Opperman, T. Sagmeister, C. Buhlheller, T. Pavkov-Keller, M. K. Rathinaswamy, U. Dalwadi, C. K. Yip, J. E. Burke, K. C. Garcia, N. V. Grishin, P. D. Adams, R. J. Read, D. Baker, *Science* **2021**, *373*, 871–876.
[119] M. Varadi, S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Žídek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi, M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis, S. Velankar, *Nucleic Acids Res.* **2021**, *50*, D439–D444.
[120] L. Holm, *Nucleic Acids Res.* **2022**, *50*, W210–W215.
[121] T. Fujishiro, K. Ataka, U. Ermler, S. Shima, *FEBS J.* **2015**, *282*, 3412–3423.
[122] W. L. Gardner, W. B. Whitman, *Genetics* **1999**, *152*, 1439–1447.
[123] R. Nakon, *Anal. Biochem.* **1979**, *95*, 527–532.
[124] E. A. Lance, C. W. Rhodes, R. Nakon, *Anal. Biochem.* **1983**, *133*, 492–501.
[125] J. M. Pope, P. R. Stevens, M. T. Angotti, R. Nakon, *Anal. Biochem.* **1980**, *103*, 214–221.
[126] Y. Liu, M. Sieprawska-Lupa, W. B. Whitman, R. H. White, *J. Biol. Chem.* **2010**, *285*, 31923–31929.
[127] A.-K. Kaster, Hochregulation Der [Fe]-Hydrogenase-Synthese in Methanococcus Maripaludis Und Methanocaldococcus Jannschii Bei Wachstum Unter Nickel-Mangelbedingungen, Diplomarbeit, **2007**.

[128] E. L. Hendrickson, A. K. Haydock, B. C. Moore, W. B. Whitman, J. A. Leigh, *Proc. Natl. Acad. Sci.* **2007**, *104*, 8930–8934.
[129] F. Long, L. Wang, B. Lupa, W. B. Whitman, *Archaea Vanc. BC* **2017**, *2017*, 7046026.
[130] K. C. Costa, S. H. Yoon, M. Pan, J. A. Burn, N. S. Baliga, J. A. Leigh, *J. Bacteriol.* **2013**, *195*, 1456–1462.
[131] T. Watanabe, T. Wagner, G. Huang, J. Kahnt, K. Ataka, U. Ermler, S. Shima, *Angew. Chem. Int. Ed.* **2019**, *58*, 3506–3510.
[132] E. Torarinsson, H.-P. Klenk, R. A. Garrett, *Environ. Microbiol.* **2005**, *7*, 47–54.
[133] M. R. Challand, T. Ziegert, P. Douglas, R. J. Wood, M. Kriek, N. M. Shaw, P. L. Roach, *FEBS Lett.* **2009**, *583*, 1358–1362.
[134] A. C. Guo, T. Jewison, M. Wilson, Y. Liu, C. Knox, Y. Djoumbou, P. Lo, R. Mandal, R. Krishnamurthy, D. S. Wishart, *Nucleic Acids Res.* **2013**, *41*, D625–D630.
[135] P. A. Frey, *Chem. Rec.* **2001**, *1*, 277–289.
[136] J. Rapp, K. Forchhammer, *Microb. Physiol.* **2021**, *31*, 248–259.
[137] E. R. Henry, J. H. Sommer, J. Hofrichter, W. A. Eaton, M. Gellert, *J. Mol. Biol.* **1983**, *166*, 443–451.
[138] A. J. Atkin, J. M. Lynam, B. E. Moulton, P. Sawle, R. Motterlini, N. M. Boyle, M. T. Pryce, I. J. S. Fairlamb, *Dalton Trans.* **2011**, *40*, 5755.
[139] T. Molle, M. Clémancey, J.-M. Latour, V. Kathirvelu, G. Sicoli, F. Forouhar, E. Mulliez, S. Gambarelli, M. Atta, *JBIC J. Biol. Inorg. Chem.* **2016**, *21*, 549–557.
[140] 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.
[141] H. E. Pence, A. Williams, *J. Chem. Educ.* **2010**, *87*, 1123–1124.
[142] S. A. Kerns, A.-C. Magtaan, P. R. Vong, M. J. Rose, *Angew. Chem. Int. Ed.* **2018**, *57*, 2855–2858.
[143] H.-J. Pan, G. Huang, M. D. Wodrich, F. F. Tirani, K. Ataka, S. Shima, X. Hu, *Nat. Chem.* **2019**, *11*, 669–675.
[144] T. Xu, C.-J. M. Yin, M. D. Wodrich, S. Mazza, K. M. Schultz, R. Scopelliti, X. Hu, *J. Am. Chem. Soc.* **2016**, *138*, 3270–3273.
[145] S. Shima, D. Chen, T. Xu, M. D. Wodrich, T. Fujishiro, K. M. Schultz, J. Kahnt, K. Ataka, X. Hu, *Nat. Chem.* **2015**, *7*, 995–1002.
[146] B. Hu, D. Chen, X. Hu, *Chem. – Eur. J.* **2014**, *20*, 1677–1682.
[147] L.-C. Song, L. Zhu, B.-B. Liu, *Organometallics* **2019**, *38*, 4071–4075.
[148] G. Darland, B. Arison, L. Kaplan, *J. Ind. Microbiol.* **1991**, *8*, 265–271.
[149] J. M. Dickinson, J. R. Hanson, P. B. Hitchcock, N. Claydon, *J. Chem. Soc. Perkin 1* **1989**, 1885–1887.
[150] T. Weber, K. J. Laiple, E. K. Pross, A. Textor, S. Grond, K. Welzel, S. Pelzer, A. Vente, W. Wohlleben, *Chem. Biol.* **2008**, *15*, 175–188.
[151] T. J. Lie, J. A. Leigh, *Appl. Environ. Microbiol.* **2007**, *73*, 6595–6600.
[152] W. B. Whitman, J. Shieh, S. Sohn, D. S. Caras, U. Premachandran, *Syst. Appl. Microbiol.* **1986**, *7*, 235–240.
[153] T. J. Lie, K. C. Costa, B. Lupa, S. Korpole, W. B. Whitman, J. A. Leigh, *Proc. Natl. Acad. Sci.* **2012**, *109*, 15473–15478.
[154] D. R. Fonseca, M. F. A. Halim, M. P. Holten, K. C. Costa, *J. Bacteriol.* **2020**, *202*, DOI 10.1128/JB.00355-20.
[155] D. L. Tumbula, R. A. Makula, W. B. Whitman, *FEMS Microbiol. Lett.* **1994**, *121*, 309–314.
[156] S. Shima, R. K. Thauer, in *Methods Enzymol.*, Academic Press, **2001**, pp. 317–353.
[157] S. Shima, M. Schick, H. Tamura, *Preparation of [Fe]-Hydrogenase from Methanogenic Archaea*, Elsevier Inc., **2011**.
[158] A. A. DiMarco, T. A. Bobik, R. S. Wolfe, *Annu. Rev. Biochem.* **1990**, *59*, 355–394.
[159] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N.

Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, *Nat. Methods* **2020**, *17*, 261–272.

[160] S. van der Walt, S. C. Colbert, G. Varoquaux, *Comput. Sci. Eng.* **2011**, *13*, 22–30.

[161] J. D. Hunter, *Comput. Sci. Eng.* **2007**, *9*, 90–95.

[162] W. McKinney, *Proc. 9th Python Sci. Conf.* **2010**, 56–61.

10. Acknowledgments

11. *Curriculum vitae*

12. Erklärung

Ich versichere, dass ich meine Dissertation mit dem Titel „Identification and characterization of two radical SAM enzymes involved in biosynthesis of the [Fe]-hydrogenase cofactor“ 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 26.09.2022

Francisco J. Arriaza G.