INVESTIGATIONS ON [FEFE]-HYDROGENASE

ACTIVE SITE BIOSYNTHESIS

by

Shawn Erin McGlynn

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Biochemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2010

APPROVAL

of a dissertation submitted by

Shawn Erin McGlynn

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency and is ready for submission to the Division of Graduate Education.

Prof. John W. Peters

Approved for the Department of Chemistry and Biochemistry

Prof. David J. Singel

Approved for the Division of Graduate Education

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Shawn Erin McGlynn

May 2010

DEDICATION

To my mother Debbie, who gave me wide eyes to explore the world.

ACKNOWLEDGEMENTS

In life there is so much given to us that it is difficult to accurately describe the depth and content of these gifts – nonetheless attribute these to specific persons. In light of this limitation, a number of people immediately come to mind whom which I am indebted. First, I thank John Peters my academic supervisor who has created an environment in his lab where the operation of science has the potential take the form of an intellectual playground. So much of teaching has to do with saying it's O.k. to ask, and the past years have been replete with questions. I am also grateful to Joan Broderick for guidance and support. I would like to thank Gordon Brittan Jr. for offering so much support and intellectual charity during the last years. Dr. Robin Gerlach has been instrumental in this work, and forms an inspiring basis from which to consider the nature of collaboration. I would also like to thank Pat Callis who's wide eyed and energetic conversations with me have been inspirational in the research that I have enjoyed so much. Finally, I thank Mike Russell at the Jet Propulsion Laboratory for intellectual conversations that are still continuing, growing, and like the chemical world around us – evolving.

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

LIST OF TABLES

Table Page

LIST OF FIGURES

LIST OF FIGURES CONTINUED

ABSTRACT

Hydrogenase enzymes, which catalyze the reversible oxidation of molecular hydrogen, occupy important roles as catalysts in microbial energy transfer and conservation. This seemingly simple reaction between protons and electrons necessitates the utilization of some of nature's most complicated organo-metallic cofactors. Remarkably, two evolutionarily independent types of enzymes capable of catalyzing this reaction exist – termed the [NiFe] and [FeFe]-hydrogenases. The biosynthesis of the cofactors harbored by these enzymes poses questions as to the assembly pathways involved in constructing hydrogen competent catalysts, and herein research as to the biosynthesis of the [FeFe]-hydrogenase active site is presented. Data as to the protein components involved in this process are presented which include the development of an *E.coli* based expression system for hydrogenase maturation protein factors, their isolation, and the first functional assignment of two of these proteins. The HydF protein is shown to be operative as an H-cluster intermediate bearing scaffold for [FeFe]-hydrogenase active site assembly, and the HydG protein is demonstrated to be responsible for the formation of cyanide and carbon monoxide from tyrosine. In addition, observations of a novel radical SAM enzyme is reported in conjunction with its putative involvement in the biosynthesis of the Hmd-hydrogenase found in methanogens. Together, these observations contribute to understanding biology's ability to construct complex organo-metallic cofactors, and lay a foundation for the consideration of the evolutionary events that led to the biological ability to assemble complex metallocofactors.

CHAPTER 1

INTRODUCTION

Iron-Sulfur Clusters and Biology

The existence of what is commonly referred to as ``life" has for its chemical basis a set(s) of interconnected reactions through which net energy dissipation may be coupled to the generation of local entropy minima[1]. Given this, a major requirement for the emergence, proliferation, and evolution of life is the presence of a repertoire of catalysts capable of allowing the occurrence of key chemical transitions in a time frame ``in sync" with other reactions of metabolic networks. In this context catalysts occupy a salient role, for in metabolism the role of the catalyst is that of the connector through which energy may be harnessed.

Protein bound iron-sulfur clusters, consisting of iron bridged by sulfide and ligated typically by cysteine thiols (Figure 1.1) fill a wide array of catalytic “niches” within the cell and are found in nearly all extant organisms and are structurally and chemically versatile biological co-factors[2,3]. In biology, these clusters are “wrapped” up by a polypeptide and the blending of the properties of the cluster together with those made possible by myriad amino acid sequences can result in clusters with varying reactivities and affinities. The properties of iron and sulfur together, namely that both can populate the d-electron orbitals, makes the iron-sulfur clusters proficient charge transfer agents that are able to cycle through different electronic states in reactions. Indeed, as discussed

---

This section contains text authored by McGlynn SE, Mulder DW, Shepard EM, Broderick JB, and Peters JW. *Dalton Trans.* 2009 Jun 14;(22):4274-85.

further below, a wide array of chemical functionality is made possible by the utilization of these clusters, and from the perspective of a network analysis approach of metabolism[4], iron-sulfur clusters are responsible for reactions that if not made possible by their presence would likely inhibit metabolism as a whole and therefore life. From their ubiquity in nature have come hypotheses that suggest these co-factors may have been among the first in nature to act in the chemistries of living systems[5]. Studies directed at the elucidation of the biological formation of these clusters may thus yield deep insight into the functional chemistry utilized by life, and the chemistry that provided for its emergence.

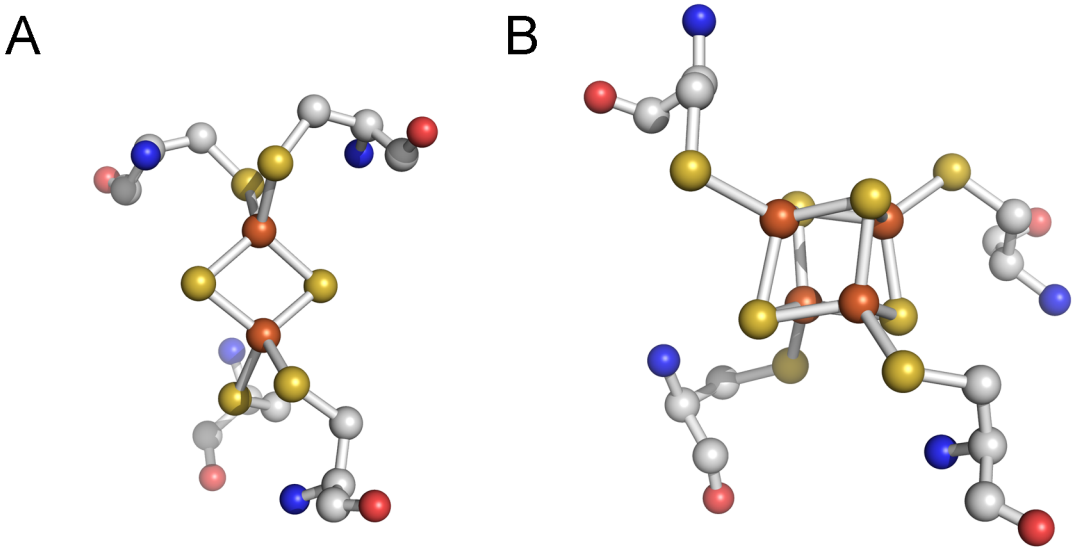

Figure 1-1 Structures of (A) the [2Fe-2S] from HydA (*C.p.*) ligated by cysteines 34, 46, 49, 62 and (B) the cysteine 147, 150, 153, 200 ligated [4Fe-4S] cluster (*C.p.*) from the 3C8Y structure.

Much progress has been made toward understanding both the physiological roles and the mechanisms of formation of these clusters and has indicated that the process of

metal cluster assembly and utilization in biology occurs is a highly coordinated and regulated process[6-11] . These studies have revealed that specialized enzymes take part in the biosynthesis of metal clusters in biology, with individual proteins mediating the metal transfer and delivery events that are required for cluster formation. Information resulting from these studies is pertinent to areas of research concerned with biological hydrogen production, the origins of life, and biological metal cluster assembly in general. Given the above considerations, progress in the field of the utilization and biogenesis of protein bound iron-sulfur clusters directly contributes to and furthers understanding of some of the foundational chemistry of living systems.

Iron-sulfur clusters in proteins have been observed to be operative in a number of roles including: structural, electron transfer reactions, enzyme catalysis, and regulation of gene expression[3,12,13]. These clusters occupy core positions in reactions of biological importance and allow organisms an expansive metabolic potential. The involvement of iron-sulfur clusters in a diverse array of redox reactions is made possible by the wide range of reduction potentials possible which range from -0.6 to 0.45 V depending on protein environment and coordination[14]. As such, iron-sulfur clusters are a centerpiece of biologically relevant reactions in that a great diversity of function and form is made possible by utilizing these co-factors. While iron-sulfur clusters of the now seemingly common, [2Fe2S], and [4Fe4S] type have been found to be capable of much diversity in reactivity and function, recent insights into the presence and operation of complex bridged metal cluster assemblies has furthered the understanding of how iron-sulfur clusters may be structurally modified in sometimes surprising ways that make accessible

an even broader array of chemical reactivities and functions. Among these modified cofactor containing proteins are nitrogenase, carbon monoxide dehydrogenase, acetyl coA synthase, and the two types of hydrogenases[15-18]. These proteins contain biologically unique co-factors, and are produced by the action of dedicated maturation protein systems that result in their assembly.

The Hydrogenase Enzymes

Hydrogenase activity, which can be defined as the reversible oxidation of molecular hydrogen, occupies an important place in the bio-energetics of microbial life across domains[19,20]. The enzymes that catalyze this reaction come in two forms with names corresponding to the metal content at their active sites. The [NiFe] and the [FeFe]-hydrogenases, while being distinct in overall fold and catalytic site composition are both able to catalyze reversible hydrogen activation. In a fascinating example of convergent evolution, these two protein classes are not evolutionary related yet are unified in the presence of CO and CN^- as ligands to iron and reactivity[21]. The [NiFe]-enzyme, which is found only in bacteria and archaea[1,2] has an active site cluster composed of a nickel ion bridged to an iron ion via two cysteine thiolates with the iron ligated by two cyanides and a carbon monoxide ligand[22]. In contrast, the [FeFe]-hydrogenase is found in bacteria and some lower eukarya and contains as an active site the "H-cluster" which consists of a [4Fe-4S] cubane linked via a bridging cysteine thiolate to a 2Fe-containing subcluster that is ligated by both carbon monoxide and cyanide ligands as well as an additional nonprotein dithiolate linkage[23] (see Figure 1.2).

These enzymes occupy the functional role within the cell of either comprising an avenue to regenerate reduced electron carriers via the reduction of protons or to couple hydrogen oxidation to energy yielding processes[19].

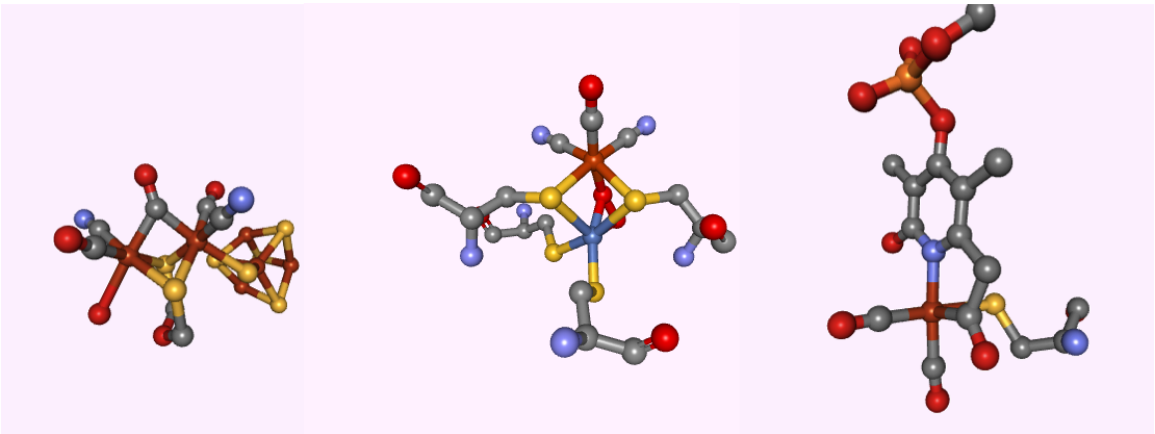

Figure 1-2 Active site structure of the hydrogenase enzymes. Left to right the [FeFe]-hydrogenase (PDB code 3C8Y), the [NiFe]-hydrogenase (PDB code 1YQ9)), and the Hmd hydrogenase (PDB code 3H65). Colors: red, oxygen; blue, nitrogen; grey, carbon; rust, iron; dark blue, nickel; yellow, sulfur; orange, phosphorous.

In addition to these two enzymes, a third enzyme which carries out a reaction with molecular hydrogen is often referred to as a hydrogenase. While not strictly catalyzing the reaction described above between protons and electrons to form molecular hydrogen, the enzyme does contain an active site reminiscent of the [NiFe] and [FeFe]-enzymes (see Figure 1.2), consisting of a single iron atom ligated by two CO molecules, a cysteine side chain, a guanylyl pyridinol cofactor (GP cofactor), and an unknown ligand suggested from crystallographic data to be an acyl group[24-30]. This hydrogenase reversibly splits molecular hydrogen to form a hydride that becomes bonded to the methenyl-tetrahydromethanopterin (H_4MPT^+) cofactor as a step in the reduction of CO_2

to methane in hydrogenotrophic methanogens and is thus termed the Hmd-hydrogenase[24] (see below for further detail).

The presence of these related cofactors in distinct lineages poses intriguing questions as to their evolution and biosynthesis. From an evolutionary perspective, the ability of microbes to utilize hydrogen as a mobile electron carrier enabled cells to form energetically coupled units in ecosystems, an example of which is seen in syntrophic communities where hydrogen is cycled between species so as to partially combine metabolic capabilities[31]. In addition, inter and intra-cellular hydrogen reduction and oxidation allows for energy conservation between cellular organelles and their membranes[19,20,32]. The wide utilization and relative simplicity of H_2 as an energy carrier has led to the suggestion that the evolution of hydrogenase activity may have been important in the origins and evolution of biological systems[33]. Indeed, on the early Earth, levels of hydrogen may have been 3-4 orders of magnitude higher than those that exist today[34], and local concentrations may have been much higher than that. These high concentrations of hydrogen may have thus been an attractive force of reducing power for emerging biology on the early earth and a parallel between hydrogen producing rock water interactions known as serpentinization[33,35]. Modern biological hydrogen production and utilization also suggests that hydrogen as a source of electron equivalents may have been important at a very early state in biological systems. That biology has continued to utilize hydrogen for energetic processes up until today is a testament to the versatility of the hydrogen redox couple and the ubiquity of protons in the aqueous environments in which life exists. The accomplishment of hydrogen activation through

the use of transition metal containing cofactor supports the notion that the electron transfer ability afforded by these metal centers may be requisite for biology. Aside from these early steps in the origin and evolution of life, hydrogen and the activity of hydrogenases may have also been important for the evolution of advanced multi-cellular life[36,37] through symbiosis.

Given that the hydrogenases comprise complex molecular structures, their formation within the cell may be expected to require the actions of devoted biosynthetic machinery. Moreover, the development of this machinery in evolutionary time could pose a bottleneck for the origination of the aforementioned activities of the hydrogenase enzymes. One avenue to gain insight into the evolution and utilization of the hydrogen redox couple in biology is in the research of the biosynthetic pathways by which the hydrogenase enzymes are made.

Biosynthesis of the Hydrogenase Enzymes

With the elucidation of the unique coordination environments of the active site structures of the hydrogenases, lines of research were begun aimed at gaining knowledge of the biosynthesis of these unique metal cofactors within the cell. Putative significant challenges in the assembly of these active sites include the trafficking of metal ions and the synthesis of the unique ligands. The fact that the three examples of hydrognenases are evolutionarily distinct is also coincident with the observation of independent evolutionary origins for of the biosynthetic proteins as discussed further below.

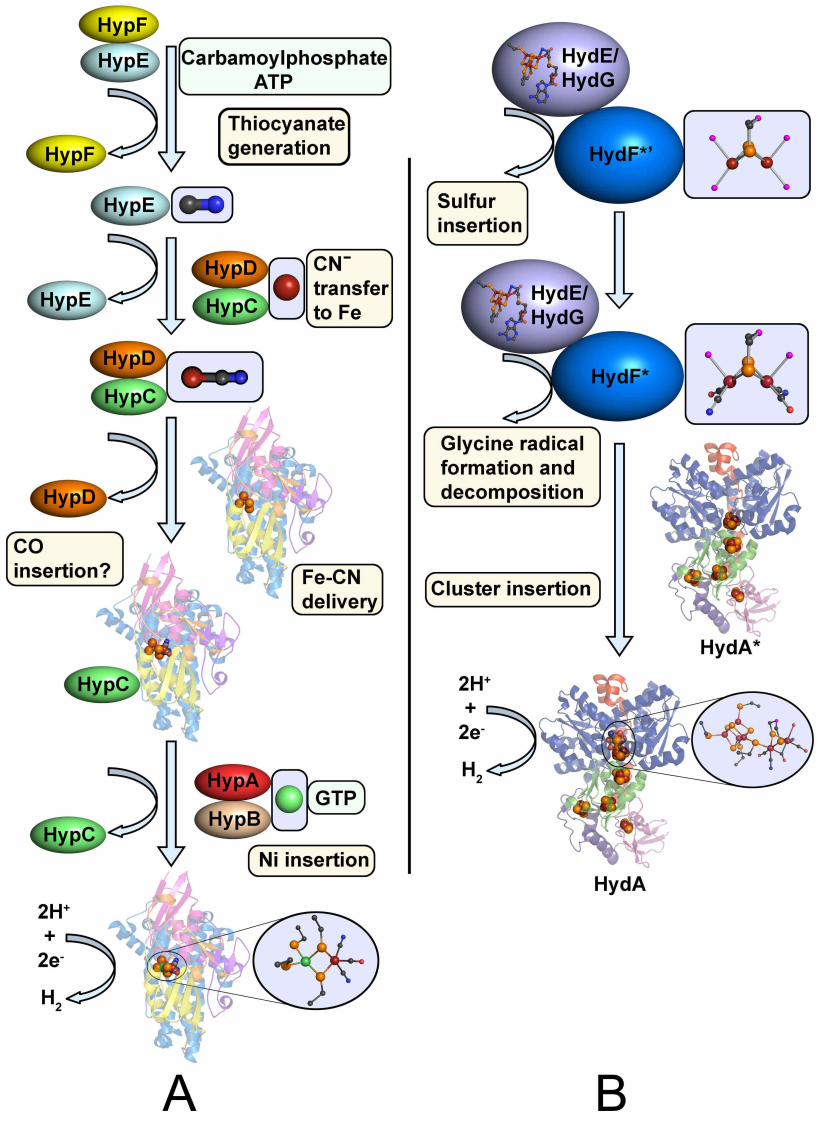

Figure 1-3 Overview of the metal cluster assembly processes for the (A) [NiFe]- and (B) [FeFe]-hydrogenases. The [NiFe]-hydrogenase maturation scheme is based on hydrogenase-3 in *E. coli* and the [FeFe]-hydrogenase maturation model is thought to be general to organisms containing this enzyme. See text for detail.

Progress in the area of the [NiFe]-hydrogenases has been made with the study of a number of organisms including *Ralstonia eutropha, Bradyrhizobium japonicum,*

Rhizobium leguminosarum, Azotobacter species, and *E. coli* [NiFe]-hydrogenase maturation proceeds by the action of at least six accessory proteins, with seven being required in cases in which proteolytic processing is required at the final step of maturation[38-40]. The overall process is summarized in Figure 1.3 A. The proposed roles of individual proteins begins with the generation of cyanide from carbamoylphosphate by the action of the two proteins HypF and HypE. In a two step process that requires two ATP hydrolyzing events, HypF transfers the carboxamide group of carbamoylphosphate to HypE as thiocarboxyamide where dehydration occurs to yield an unique enzyme bound thiocyanate, which can later be donated to iron[41-46]. Subsequent interaction of thiocyanate bound HypE with the HypD/HypC complex results in the donation of the CN^- ligand to an iron bound to HypD/HypC[43]. The CN^- ligated iron bound by the HypD/HypC complex is then transferred by HypC to the large subunit of the hydrogenase enzyme. Still in question is the metabolic source of the CO ligand and at which stage it is incorporated. Carbamoylphosphate has been ruled out as the source of CO by labeling studies[47,48] and the utilization of a pathway involving acetate has been suggested[45].

With CO and CN- coordinated Fe now present on the large subunit, Ni is inserted by the actions of the Ni binding GTPase HypB[49,50] and the Ni binding HypA[51]. The final step in [NiFe]-hydrogenase maturation involves proteolytic cleavage of a C-terminal extension, if present. Major challenges in the elucidation of the biosynthetic pathway for the [NiFe]-hydrogenases includes defining where the CO of the active site comes from, as well as understanding the regulation of nickel transport and delivery to the active site cluster. This line of research detailed above has revealed the activity of novel enzymes

being required in the formation of a unique and novel enzyme active site, and suggests the biologically unique hydrogenase enzyme active sites, contain within their biosynthetic pathways further novel and unprecedented chemistry.

In the case of the [FeFe]-hydrogenases, Posewitz and coworkers first identified the necessary gene products for [FeFe]-hydrogenase activity in the algae *Chlamydomonas reinhardtii (C.r.).* Through studies of deletion mutants, the gene products of *hydEF* and *hydG* were found to be required factors in the formation of active [FeFe]-hydrogenase[52,53]. In organisms other than *C.r.* the *hydEF* gene is found to exist as two separate genes *hydE* and *hydF*. These genes were observed to be conserved in all organisms containing an active hydrogenase and thus are thought to be sufficient factors for the maturation and assembly of the [FeFe]-hydrogenase H-cluster, outside of other widely distributed protein factors that are common to organisms. Sequence analysis of the three gene produces revealed that the genes belong to the radical SAM family of enzymes in the case of HydE and HydG, and that HydF is likely a GTPase capable of binding an iron-sulfur cluster. The notion that these three products could suffice for H-cluster assembly was bolstered by the successful expression of active [FeFe]-hydrogenases in *E.coli* – an organism devoid of an endogenous [FeFe]-hydrogenase - by the genetic complementation of the hydrogenase gene *hydA* with the aforementioned maturation factors[54]. In addition, this study indicated that the radical-S-adenosylmethionine (SAM) cluster binding domains of HydE and HydG, as well as the GTPase activity and iron-sulfur cluster binding ability of HydF were all requirements for H-cluster assembly.

The identification of these gene products as being involved in the process of H-cluster assembly and the demonstration that the gene products could successfully be expressed and manipulated in *E.coli* resulted in the ability to begin detailed biochemical investigation of these enzymes and the chemistry that they perform in the cell. In addition, knowledge as to the enzyme families to which the Hyd-maturation proteins belonged to allowed for the formation of hypotheses as to how chemical events - precedented by other enzymes – could be involved in the formation of the H-cluster.

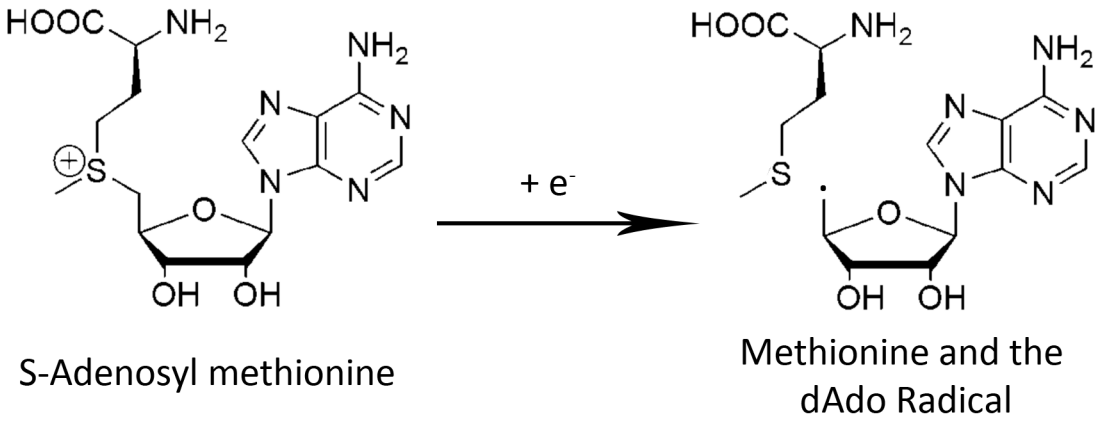

Figure 1-4 Reaction carried out by the radical SAM family of enzymes.

The first biochemical characterization of the enzymes came with work that corroborated inferences from sequence analysis, namely that the HydE and HydG proteins do behave as radical SAM enzymes insofar that they bind a 4Fe-4S cluster and cleave SAM into the reaction product 5'-deoxyadenosine (dAdo)[55] (see below and Figure 1.4). Shortly thereafter, the HydF protein was characterized and also found to behave as predicted from sequence analysis in that it bound a 4Fe-4S cluster and exhibited GTPase

activity[56]. Despite confirming membership in these enzyme classes, little could be said about the actual chemistry that the enzymes were catalyzing to form the H-cluster.

The radical SAM class of enzymes comprises a large family of iron-sulfur cluster containing enzymes capable of catalyzing a wide range of radical initiated reactions including carbon-carbon bond formation, isomerizations, and sulfur insertions[57-59] all being initiated by the formation of a radical carbon-sulfur bond via homolysis of the substrate SAM molecule (Figure 1.4). These enzymes utilize a mechanism by which the functionalization of very stable bonds is accomplished in biology, especially in anoxic microbes. Many of these reactions are challenging from a chemical standpoint. For instance, for the functionalization of C-H bonds, high bond dissociation energies and large pKa's make these moieties rather inert and the modification of these groups could be expected to be challenging[60,61]. The radical SAM enzymes exist as a biological response to this challenging chemistry, and by providing a unique environment which combines protein environment and transition metal based cluster reactivity, the enzymes are able to accomplish myriad types of reactions. The wide chemical repertoire of reactivities observed from this enzyme class suggests that they are extremely versatile and that electron transfer from a [FeS] cluster to form a carbon centered radical via homolysis of a sulfonium ion is a highly effective means from which to accomplish such a wide array of reactions.

The inclusion of the three maturation associated proteins into enzyme families gave rise to a hypothesis for the action of two radical SAM proteins and a Fe-S cluster binding GTPase as being involved in the formation of the H-cluster. Given the wide

diversity of reactivity from the radical SAM enzyme class, a number of chemistries may be envisaged as taking part in H-cluster biosynthesis. Based on pre-established reactivity of radical SAM enzymes, a hypothesis was proposed in which the activities of HydE and HydG, are proposed to act to generate the CO and CN^- ligands and the bridging dithiolate ligand[62].

In the hypothesis, the three maturation proteins are presumed to be responsible in total for the generation of the ligand set of the [FeFe]-subcluster[62]; the possibility that CO could be generated through the action of carbon monoxide dehydrogenase was excluded as [FeFe]-hydrogenases are found in organisms which do not contain this enzyme. The three gene products HydE, HydF, and HydG were thought to be directed solely at the synthesis of the 2Fe subunit of the [FeFe] active site, as the [4Fe-4S] moiety presumably would not require the action of a specific set of assembly proteins but instead be formed via the action of the endogenous host iron-sulfur cluster assembly machinery[7]. HydE or HydG was envisioned as acting to generate the bridging dithiolate linkage in a mechanism analogous to that of the radical SAM protein LipA, which catalyzes carbon–sulfur bond formation[63,64]; such C–S bond formation, occurring at a bridging sulfide of an iron–sulfur cluster, would be expected to shift the reactivity from the sulfur atoms to the irons of the cluster as this step would result in a greater electrophilicity of the iron atoms[65]. Following alkylation of the sulfides, a radical-initiated amino acid decomposition of glycine was proposed to occur by the action of HydE or HydG, resulting in the formation of CO and CN^-. The first step of this latter proposed reaction is analogous to the chemistry catalyzed by the radical SAM enzymes pyruvate formate-

lyase activating enzyme[66], lysine amino mutase[67], or ThiH involved in thiamine biosynthesis[68] which involves the generation of an amino acid radical intermediate. These ligand-forming events were hypothesized to result in assembly of the 2Fe subcluster of the H-cluster on one of the accessory proteins. The role of GTP binding and hydrolysis by HydF was postulated as being involved in cluster translocation from the HydF protein to the HydA protein. Observation that other metal cluster assembly systems make use of nucleotide hydrolysis suggests a similar reaction characteristic may be operative in a wider array of enzymatic systems (an overview of this process is depicted in Figure 1.3).

These speculations motivate the notion the three Hyd proteins themselves could be responsible in total for the synthesis of the unique features of the [FeFe]-hydrogenase enzymes and that outside of the core iron-sulfur cluster assembly pathways employed by organisms, these three enzymes would then occupy the roles of generating the diatomic as well as thiolate ligands of the H-cluster. It is of interest to note that in the case of the [NiFe]-hydrogenase biosynthetic pathway, no radical-SAM chemistry has been observed to be required. In addition CN^- coordination to iron is accomplished utilizing carbomyl phosphate as a precursor in a multistep process where oxygen is eliminated with a phosphorylation event[41]. Still unknown in this process is the formation of the CO ligand, which could still prove to be derived from a radical intermediate.

Finally, in the case of the biosynthesis of the Hmd-hydrogenase active site there has been no identification or suggestion as to what enzymatic process and proteins may be involved in the generation of the active site of that enzyme. In this case as well as the aforementioned [FeFe]-hydrogenases the ability of radical-SAM enzymes to accomplish

a diverse set of chemical functionalizations would make them possible candidates in this process. It is of interest to note the diversity of metabolic processes that employ the operation of hydrogenases in microbial energetics, and in the case of the Hmd-hydrogenase this notion is furthered in the observation of it being operative in carbon dioxide reduction to methane in the hydrogenotrophic methanogenisis pathway discussed below.

The origins and operation of the Hmd-hydrogenases themselves exist as an evolutionary and metabolic puzzle in that the reaction they accomplish can occur by two independent enzyme catalyzed reactions. In the reduction of CO_2 via hydride transfer to methane in the hydrogenotrophic methanogens, carbon is bonded to an organic cofactor (H_4MPT) that serves as a scaffold upon which the CO_2 is sequentially reduced by a

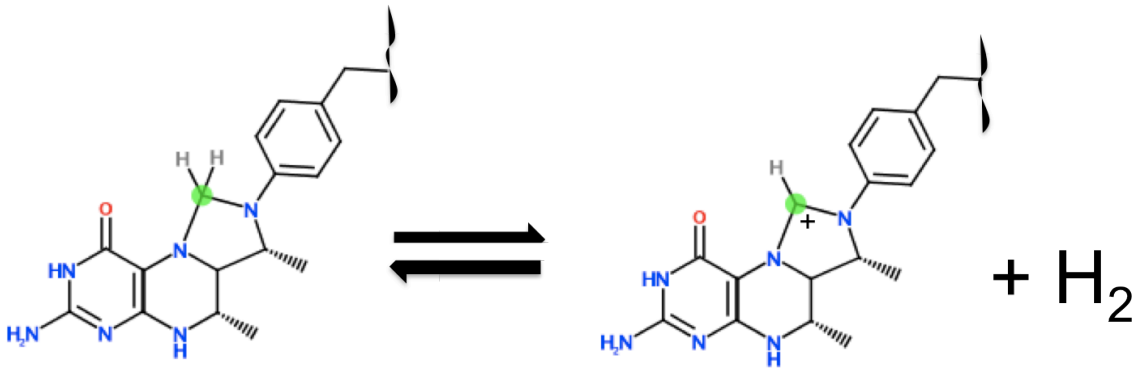

Figure 1-5 Reaction carried out by the Hmd-hydrogenase. A hydride derived from H_2 is reversibly transferred to the methylenetetrahydromethanopterin. The green circled carbon indicates the carbon position that is derived from CO_2 in steps prior to the activity of the Hmd-hydrogenase. The full molecule is not drawn, and where it is abbreviated is shown by the presence of a wavy line (top right).

complex process involving multiple enzymes including a number of [NiFe]-hydrogenases[69]. At the second stage of this reduction, either the Hmd-hydrogenase or a F_{420}-reducing [NiFe]-hydrogenase is utilized to accomplish reduction of the carbon atom. This reaction is depicted in Figure 1.5, where the CO_2 derived carbon is highlighted with a green circle. Thus the Hmd-hydrogenase takes part in a fundamentally different reaction involving the activation of hydrogen than in the case of the [NiFe]- and [FeFe]-hydrogenases as it functions as a catalyst for hydride transfer to a carbon center.

Proteomic studies where experiments were conducted under nutrient limitation indicated that the Hmd-hydrogenase is active when H_2 in cultures is not limiting, whereas the F_{420}-reducing [NiFe]-hydrogenase dominates at limiting H_2 supporting the hypothesis that the two have varying affinities for the reaction[70]. Together with previous observations where the Hmd-hydrogenase has been observed to be active only in the case of nickel limitation[71], utilization of the Hmd-hydrogenase appears to be a means by which the organisms that employ it can reduce overall nickel requirements but still accomplish the hydride transfers required for the overall cellular energetics of methanogenesis.

The uniqueness of the chemical activity of the Hmd-hydrogenase of hydride reduction of a carbon atom, together with the evolutionary uniqueness of the three classes of hydrogenase enzymes themselves suggests that that the mechanisms of generating the Fe-CO ligands may also be unique amongst these divergent proteins. It is therefore difficult to hypothesize possible routes to the Hmd-hydrogenase Fe-CO active site, especially in the absence of known protein assembly factors that could form the basis for

hypothesis generation. Aside of the generation of Fe-CO in the active site of the enzyme, the biosynthesis of the Hmd-hydrogenase requires the formation of a modified pyridine ring of which there is little precedence for. Thus the biosynthesis of this unique bi-dentate ligand of the Fe atom of the active site itself holds questions as to the origin and evolution of Hmd-hydrogenases in hydrogenotrophic methanogens. Furthermore, knowledge gained of the processes utilized in generating this unique coordination sphere will yield insight into the requirements and determinants of methanogenic energetics, as well as provide a chemical context to relate to the other classes of hydrogenases.

Research Directions

This work carries on from a body of knowledge gained on the utilization of metals in biology – specifically transition metal based catalysts, and exists as an endeavor to begin elucidation of the biosynthesis of the [FeFe]- and Hmd-hydrogenases. In the case of the former, work begins with seminal work from Posewitz and co-workers[52] discussed above which resulted in the identification of gene products required for the maturation of the [FeFe]-hydrogenases. In addition to this work, that of King et al.[54] provided a major advance in the context of studies undertaken herein in that it laid the groundwork for the heterologous expression of the [FeFe]-hydrogenases and the maturation proteins. These two studies serve as a sort of starting point from which research detailed here began. Indeed, the identification of specific proteins in biological processes often exists as a starting point from which to base detailed, mechanistic studies where the precise mode of molecular action can be delineated. This research proceeds then by the application of

previously gained knowledge to the generation of novel hypotheses about systems which can be addressed experimentally. In the case of the [FeFe]-hydrogenases, questions that arose from the aforementioned works include what are the individual proteins doing and how do they function together to produce the enzyme active site. By isolating the proteins individually with these questions and their associated hypotheses in mind, progress (detailed in the subsequent chapters) has been made towards answering these questions.

In the case of the Hmd-hydrogenase, very little is known as to the mechanistic details by which the unique co-factor harbored by the enzyme is synthesized. Herein, a combination of biochemical and bioinformatic approaches are utilized in an effort to begin to understand which proteins in the cell may take part in this process. These studies provide evidence as to which proteins may be involved in the assembly process of the Hmd-hydrogenases and what type of chemistry is involved, resulting in an incremental but perhaps significant advance in the understanding of this enzyme and the factors that control its formation.

Of course with knowledge often comes more questions and indeed the process of gaining awareness is in part about being able to ask questions of increasing levels of detail than previously would have been allowed in the absence of gained knowledge. In some senses, the job of the scientist is and has been construed as being the job of detective, where more and more evidence for or against a particular case is gained by observation, and the scientist forms expectations of experimental outcomes accordingly[72].

Below, some questions on the biosynthesis of the [FeFe]- and Hmd-hydrogenases are addressed, and in turn many raised.

CHAPTER 2

IN VITRO ACTIVATION OF [FEFE]-HYDROGNENASES

Chapter Introduction

The assignment of specific proteins to the process of [FeFe]-hydrogenase biosynthesis, together with the demonstrated ability to express these heterologously in an *E.coli* based system gave rise to the unprecedented ability to begin probing the pathway of H-cluster biosynthesis systematically. In light of other metal cluster assembly pathways, one possibility for H-cluster assembly would involve the use of a specific protein scaffold upon which the H-cluster or an intermediate of this would be assembled. Iron-sulfur cluster assembly involves the operation of scaffolds upon which clusters are assembled[6-9], and in the case of complex metal cluster assembly, the nitrogenase system exists as an example of a biosynthetic route to a complex metal cluster which occurs via the action of specific proteins. In the case of nitrogenase, the catalysts of enzymes, carrier proteins, and scaffolds have been observed[73,74], the operation of which gives rise to the notion that the process of nitrogenase cofactor assembly takes place in a coordinated and specific manner. Drawing inspiration from these observations, it is logical to hypothesize that analogous functionalities may be present in the case of [FeFe]-hydrogenase biosynthesis and in particular, that the Hyd maturation proteins themselves may occupy the function of enzymes and scaffold/carrier proteins. If a scaffold protein is involved in H-cluster assembly, it may be the case that it is isolatable as a stable

---

This section contains text authored by McGlynn SE, Ruebush SS, Naumov A,Nagy LE, Dubini A, King PW, Broderick JB, Posewitz MC, and Peters JW. *J Biol Inorg Chem.* 2007 May;12(4):443-7.

intermediate. As a beginning step, splitting the process into two (or more) steps from which to probe intermediates and processes would give biochemical control and fidelity to studies aimed at isolating intermediates in the reaction pathway and elucidating the actions of proteins leading to these intermediates.

Recent studies demonstrate that the *Clostridium acetobutylicum* maturation proteins (HydE, HydF, and HydG), when expressed in concert with a hydrogenase structural protein (HydA), are sufficient to support the expression of active hydrogenases from a variety of sources[54,75,76]. Previously work by King et al. found that the coexpression of *Clostridium saccharobutylicum* HydA with the *C. acetobutylicum* maturation proteins provided an effective reporter of hydrogenase activity in *E.coli*. These constructs were therefore used to initiate studies examining the mechanism of hydrogenase maturation.

Materials and Methods

The proteins HydA, HydE, HydF, and HydG were expressed either singly or in concert in *E.* coli BL21(DE3) under inducible control of the T7 polymerase using a variety of combinations of constructs from the Novagen Duet™ plasmid suite. Cultures were grown in supplemented minimal media[77] with appropriate antibiotics to OD_{600} of 0.5 and protein expression was then induced by the addition of 1 mM IPTG. Cultures were then shaken for 2 hours at 37^{o}C and following this initial induction period, cultures were incubated at 4^{o}C overnight under a N_2 sparge. Western blot analysis and Coomassie stain were used to visualize the expressed proteins after SDS-PAGE. Cells were

harvested anaerobically, transferred to a Coy™ anaerobic chamber, and suspended in anaerobic buffer A for washing (50 mM Tris 150 mM NaCl with 1 mM dithiothreotol (DTT)). Anaerobic cell suspensions were transferred to sealed centrifuge tubes and sedimented at 10,000x*g*. Following centrifugation the cell pellet was resuspended in 5 volumes of anaerobic buffer B (50 mM HEPES pH 7.5, 200 mM NaCl, 5% (w/v) glycerol, 1% (w/v) Triton X-100, 10 mM $MgCl_2$, 1mM DTT). DTT was used as a reductant instead of sodium dithionite as it was observed that sodium dithionite concentrations greater than 0.05 mM are inhibitory toward *in vitro* activation. Cells were lysed using a pressure cell bomb by exposing cells to ~100 bar pressure for one hour prior to controlled pressure release. The cell lysate was then clarified by centrifugation at 30,000*xg* for 40 min at 4°C and decanted into sealed vials in a Coy™ anaerobic chamber. Protein concentrations were determined by using the BCA™ Protein Assay kit.

Hydrogenase activity assays using cell-free extracts were performed by monitoring H_2 production in reaction mixtures containing hydrogenase, methyl viologen, and dithionite[78]. Dithionite, added following *in vitro* combination serves as a reductant of the newly activated hydrogenase. Assay mixtures to monitor the *in vitro* activation of $HydA_{\Delta}^{EFG}$ were prepared in a Coy™ anaerobic chamber by mixing cell free extracts of $HydA_{\Delta}^{EFG}$ (5µL) and HydEFG (5 – 1000 µL) in 3 ml sealed Wheaton™ vials. A solution of sodium dithionite (final concentration 20mM) in 50 mM Tris pH 7.5, 150 mM NaCl in a total volume of 2 ml was added to the extracts. Methyl viologen (10 mM final concentration) was added to initiate the reaction and samples of headspace gas were

analyzed using a Shimadzu GC-8A gas chromatograph with a Supelco 80/100 Porapak N column (6 ft. x 1/8th in.).

Results

In vitro Activation of [FeFe]-Hydrogenases

To test whether the Hyd assembly proteins themselves could confer or produce an H-cluster intermediate without the presence of the structural protein HydA, two cell lines were created one which expressed only HydA, and a second which expressed HydE, HydF, and HydG[79]. Upon combination of the soluble fraction of lysates from these cell lines, it was found that the inactive HydA gene product expressed in a background devoid of the H-cluster maturation gene products can be activated *in vitro* in the presence of cell extracts containing the maturation proteins HydE, HydF, and HydG heterologously

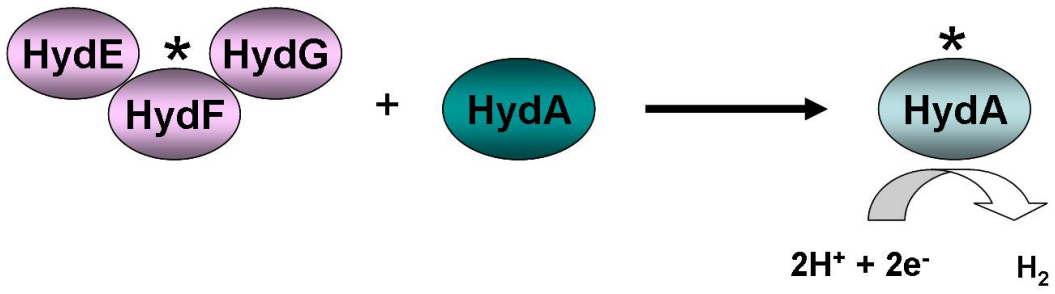

Figure 2-1 Cartoon scheme for the activation of [FeFe]-hydrogenases (HydA) by the maturation proteins HydE, HydF, and HydG.

expressed in concert (HydEFG). The features of *in vitro* activation support a model for hydrogenase active site maturation in which a preformed cluster fragment synthesized by the activity of the hydrogenase accessory proteins is readily transferred to an inactive, formative HydA gene product completing the H-cluster active site (Figure 2.1).

Hydrogenase activity was observed after extracts of heterologously expressed $HydA^{\Delta EFG}$ (6.3mg protein/ml) were combined with extracts in which the hydrogenase maturation proteins HydEFG (*C.a.*) (11mg protein/ml) were expressed in concert (Table 2.1). Hydrogenase activity was measured using a methyl viologen and dithionite assay with H_2 evolution measured by gas chromatography[78]. Control experiments consisting of cell extracts containing $HydA^{\Delta EFG}$ alone, and HydEFG alone, or accessory proteins

Table 2.1: Observed FeFe-hydrogenase activation (±) after combining indicated combinations of coexpressed hydrogenase accessory proteins HydE (*Clostridium acetobutylicum*), HydF (*C. acetobutylicum*), and HydG (C. *acetobutylicum*) with hydrogenase structural protein HydA heterologously expressed singly in *Escherichia coli* ($HydA^{\Delta EFG}$) (*C. saccharobutylicum*)

Extract combination	Activity observed (±)
$HydA^{\Delta EFG}$ + HydEFG	+
$HydA^{\Delta EFG}$ + HydEF + HydG	–
$HydA^{\Delta EFG}$ + HydEG + HydF	–
$HydA^{\Delta EFG}$ + HydFG + HydE	–
$HydA^{\Delta EFG}$ + HydE + HydF + HydG	–

HydEFG hydrogenase maturation proteins HydE, HydF, and HydG coexpressed heterologously in *E. coli*

expressed individually resulted in no detectable H_2 evolution. Hydrogenase activity was also not observed when the three accessory proteins were expressed separately or in varying combinations in independent cell lines and combined *in vitro* with $HydA^{\Delta EFG}$.

These results indicate that the accessory protein components must be expressed in concert and that they act together to affect the formation of active hydrogenase in the *in vitro* activation assay (Table 1). *In vitro* activation of HydA$^{\Delta EFG}$ appeared to be complete within five minutes and proceeded without supplying additional reactants. The addition of SAM (2mM), GTP, and ATP (1 mM) to reaction mixtures did not stimulate activation.

Preliminary Fractionation of the [FeFe]-Hydrogenase Activating Component

The possibility that expression of HydE, HydF, and HydG results in the formation of a small molecule precursor that is spontaneously inserted into the hydrogenase active

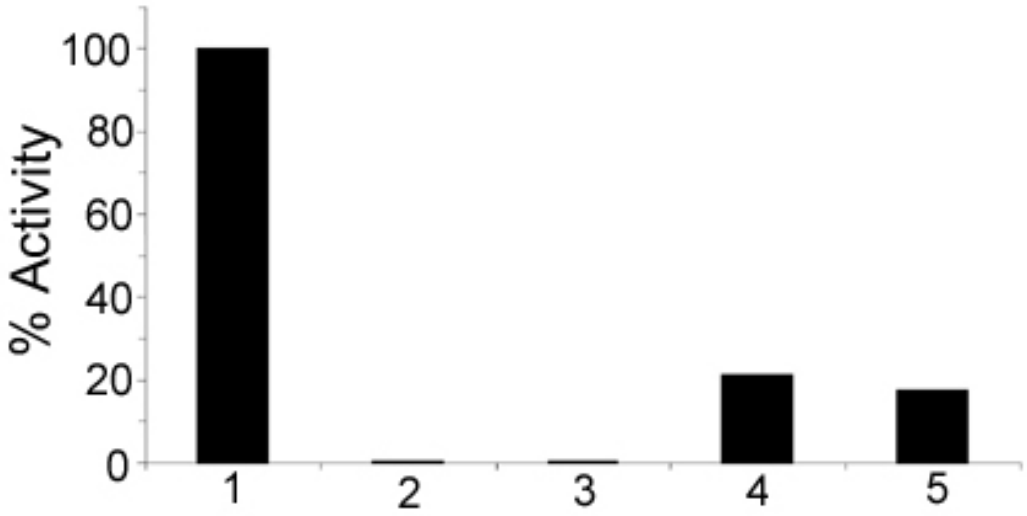

Figure 2-2 Hydrogen evolution by HydA$^{\Delta EFG}$ (*C. saccharobutylicum*) activated by extracts of hydrogenase maturation proteins HydE, HydF, and HydG coexpressed heterologously in *E. coli* (HydEFG) (*C. acetobutylicum*). Bars from left to right: 1 HydA (C. *saccharobutylicum*) plus HydEFG (C. *acetobutylicum*), 2 HydA (*C. saccharobutylicum*) plus HydEFG (*C.* acetobutylicum) heat- treated, 3 HydA (*C. saccharobutylicum*) plus re-addition of column flow through small molecules, 4 HydA (*C. saccharobutylicum*) plus HydEFG (*C. acetobutylicum*) washed, 5 HydA (*C. saccharobutylicum*) plus HydEFG (*C. acetobutylicum*) washed plus re-addition of column flow through. The absence of either HydA$^{\Delta EFG}$ or HydEFG resulted in no detectable hydrogen evolution. Percentages are based on a maximal activation curve.

site without the involvement of accessory proteins was also explored. To analyze this possibility, extracts of *E. coli* in which HydEFG were expressed in concert were subsequently heat treated at 95° C for ten minutes to denature proteins. These extracts were not competent to activate HydA as analyzed by hydrogenase activity assays. In addition, the HydEFG extracts were filtered using a 10kDa MW spin column and the filtrate was also unable to activate HydA$^{\Delta EFG}$ (Figure 2.2). Moreover, when the small molecule fraction from HydEFG cell-free extracts was added to a reaction mixture containing both HydA$^{\Delta EFG}$ and HydEFG extracts, there was no observable enhancement of activity, indicating that components of the extracts less than 10kDa alone could not affect activation. In contrast, the protein fraction that was retained upon filtration conferred activity even after multiple washes (Figure 2.2). These experiments indicate that the cluster precursor is a protein-associated component. This is highly suggestive that one or more of the accessory proteins themselves serve as the scaffold for the assembly of the cluster precursor. The observed decrease in activity under these experimental conditions is likely a result of the labile nature and instability of the precursor itself or the assembly scaffold, which is presumed to be one of the accessory proteins or perhaps a complex of maturation proteins.

Characteristics of [FeFe]-Hydrogenase Activation

Probing the maximal activity of cell free HydA$^{\Delta EFG}$ (*C.s.*) activated extracts was accomplished by titration with increasing amounts of HydEFG (*C.a.*) cell-free extract, with full activation of the hydrogenase observed at high HydEFG/HydA$^{\Delta EFG}$ ratios

(~80:1 total protein ratio) (Figure 2. 3). The relationship between the ratio observed for activation *in vitro* and that required for full activation *in vivo* is unknown, however, the

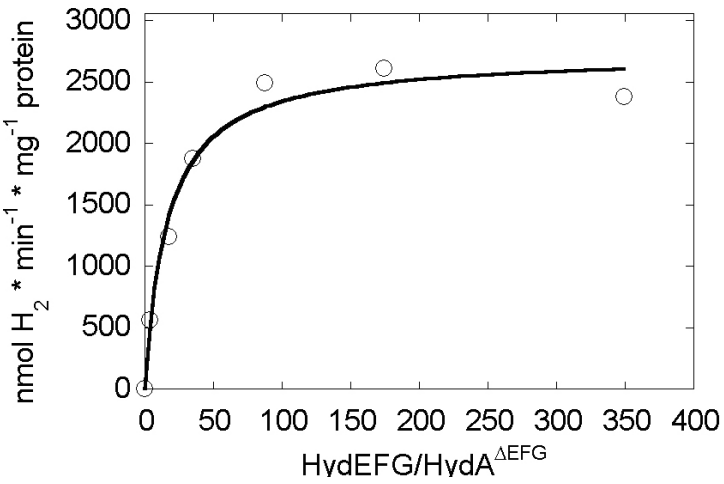

Figure 2-3. Rate of hydrogen evolution per minute per milligram of protein in HydA$^{\Delta EFG}$ (*C. saccharobutylicum*) extract titrated with HydEFG (*C. acetobutylicum*) extract. Protein amounts and ratios represent total protein content. Maximal activation is observed at HydEFG-to-HydA protein ratios of approximately 80:1.

high titer volumes required for full activation of the singly expressed HydA$^{\Delta EFG}$ may be a result of differential rates of protein expression observed between single and multiple expression schemes. As may be expected with expression systems where either single or multiple co-expressions are made, the single expression is often observed in the higher quantity. This phenomenon is operative in this activation scheme, where high ratios of the co-expression of HydEFG extract are required to affect full activation of HydA$^{\Delta EFG}$. Unknown at present is the percent of HydA$^{\Delta EFG}$ capable of being activated and the percent amount of HydEFG capable of effecting activation in our preparations; titer ratios observed are a complex function of the amount of protein

expressed and the protein available for activation. Maximal activation was observed to occur rapidly and incubation of the activation mixture for time periods of up to 2.5 hours (Figure 2.4A) resulted in no significant increase in the rate of hydrogenase activity over what is initially observed.

It was previously shown that active HydA can be obtained when co-expressed in a background of HydE, HydF, and HydG[52,53]. In these studies, it was observed that

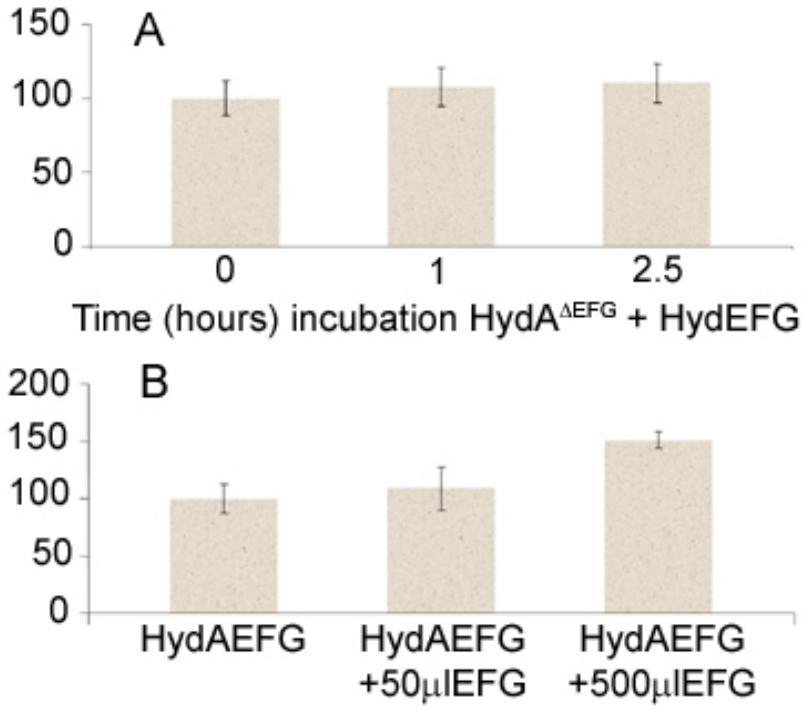

Figure 2-4 Hydrogenase activity (percentage activity) after cell extracts containing HydA$^{\Delta EFG}$ were combined with extracts containing HydEFG in ratios corresponding to partial activation (see Fig. 2b) and incubated for the time indicated. B. Rates of H_2 evolution (percentage activity) observed when extracts of cells expressing HydEFG (*C. acetobutylicum*) were added to extracts of cells expressing the full complement of hydrogenase proteins in concert: HydA (*C. saccharobutylicum*), HydE (*C. acetobutylicum*), HydF (*C. acetobutylicum*), HydG (C. *acetobutylicum*). Hydrogenase activity was increased by approximately 50% with the addition of accessory protein containing extracts. HydEFG (*C. acetobutylicum*) extracts alone in both cases exhibited no detectable hydrogenase activity.

specific activities of the resulting hydrogenases were lower than anticipated when the appropriate comparisons were possible. Here, it was observed that hydrogenase activities of *E. coli* crude extracts containing *C. saccharobutylicum* HydA expressed in a background of HydEFG (*C.a.*) can be increased by up to 50% by the addition of crude extracts containing HydEFG (Figure 2.4B). The observation that the HydEFG extracts support the supplemental activation of heterologous co-expressions of HydA (*C.s.*) together with HydEFG (*C.a.*) suggest that the inability to achieve the specific activities observed for hydrogenase enzymes expressed in their surrogate host is a result of incomplete activation of heterologously expressed HydA *in vivo*.

In the above, HydA (*C.s.*) was activated utilizing extracts of HydEFG (*C.a.*). To determine the flexibility of the system in activating proteins from different species, activation experiments were performed with different HydA's as well as different HydF's. In the case of alternate HydA proteins, the proteins from *Chlamydomonas reinhardtii* and *Clostridium pasteurianum* were found to be suitable for activation with the *C.a*. accessory proteins. In the case of HydF, the *Thermotoga Maritima* (*T.m*) accessory proteins were not able to activate HydA (*C.p.*). This was overcome however by co-expression of the *T.m* HydF protein with HydE and HydG from *C.a.,* which resulted in the ability to accomplish HydA(*C.p.*) activation (Figure2.5). Importantly, this activation was observed only in the case of expressions carried out in the BL21DE3 cell line, and not in the BL21DE3-RIL cell line in which higher expression levels of the *T.m.* proteins are achieved.

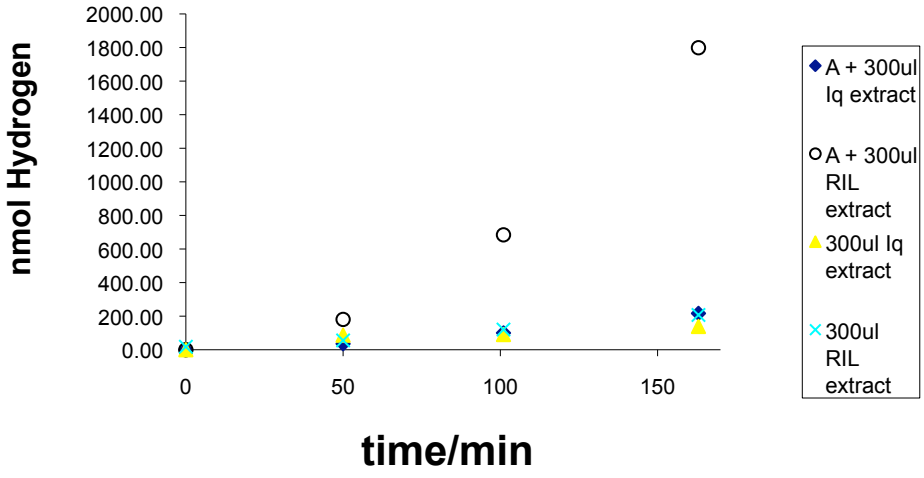

Figure 2-5 Activation in cellular extracts of HydA(*C.p.*) by HydE(*C.a.*), HydF(*T.m.*), and HydG (*C.a.*). In the legend, "A" indicates the HydA containing extract, and Iq denotes the BL21DE3 cell line, whereas RIL refers to the RIL cell line in which rare *E.coli* codons are supplemented (Novagen).

The observation that HydA$^{\Delta EFG}$ activation occurs by the addition of HydEFG as a reagent, and in the absence of small molecules is suggestive that the accessory proteins themselves are likely to harbor an H-cluster precursor that upon interaction with HydA produces active hydrogenase. These results in sum are consistent with a working model in which the biosynthesis of an H-cluster precursor occurs on one or a combination of the accessory proteins, with the accessory proteins themselves, and not the hydrogenase structural gene (HydA) serving as a physical scaffold for cluster assembly. This proposed mechanism, in which a cluster fragment is transferred to activate a formative structural gene product of a complex metalloenzyme, is analogous to the activation of nitrogenase in which the FeMo-cofactor is formed on a scaffold and in a final step transferred to an FeMo-cofactor deficient nitrogenase for full activation[73,80].

Summary and Perspectives

Although no direct experimental evidence exists as to the composition of the H-cluster precursor that is transferred upon activation, it is likely to be an intact version of the 2Fe subcluster. In this context, it is reasonable to suggest that the [4Fe-4S] subcluster of the H-cluster is formed on the HydA$^{\Delta EFG}$ without the requirement of the accessory proteins since *E. coli* possesses the host machinery for the synthesis and insertion of these clusters[41]. Preliminary data supporting the presence of a [4Fe-4S] cluster on HydA$^{\Delta EFG}$ prior to activation is provided by preliminary examination of purified *C. reinhardtii* HydA$^{\Delta EFG}$ heterologously expressed in *E. coli.* In contrast to the [FeFe]-hydrogenases from the *Clostridia*, the *C. reinhardtii* [FeFe]- hydrogenase does not contain accessory [FeS] clusters. All the Fe atoms associated with the enzyme are associated with the 6Fe H-cluster. Although the purified protein is somewhat unstable, preliminary characterization of the metal content indicated that the protein contains an average of 3.2 ± 0.13 Fe atoms per protein monomer with an EPR signal consistent with a reduced [4Fe-4S]$^{+1}$ cluster in the presence of dithionite. These results have been corroborated by detailed spectroscopic studies, which demonstrate that the state of HydA prior to its activation is that with a [4Fe-4S] cluster, and that this cluster is required for activation[81].

An alternative mechanism whereby a nearly complete H-cluster present on HydA is "decorated" with the addition of ligands by the maturation proteins is unlikely. The 2Fe subcluster is coordinated to the rest of the H-cluster (and thus to HydA) via a single bridging cysteine thiolate to the proximal Fe atom, whereas the distal Fe atom is linked

only through the bridging dithiolate and CO ligands. Thus the 2Fe subcluster has only a single protein-based ligand, and as such is unlikely to be bound to HydA in the absence of the cluster ligands (the CO, CN, and dithiolate moiety). It is therefore difficult to conceive how precursors of the 2Fe center would be tethered to the structural enzyme during maturation.

Data on the in vitro activation of [FeFe]-hydrogenases demonstrates that inactive HydA expressed in a background devoid of HydE, HydF, and HydG ($HydA^{\Delta EFG}$) can be activated in vitro by adding extracts of *E. coli* containing the three maturation proteins HydE, HydF, and HydG expressed in concert[79]. The observed in vitro activation of $HydA^{\Delta EFG}$ was not restricted to the *C. saccharobutylicum* enzyme as activation was also observed by combining HydEFG extracts with cell-free extracts containing HydA from *Chlamydomonas reinhardtii* and *Clostridium pasteurianum*, consistent with previous reports[76].

In the context of an overall mechanism of H-cluster biosynthesis and hydrogenase activation, the results presented herein provide a significant advancement in our understanding of [FeFe]- hydrogenase maturation. These results have shown that hydrogenase activation requires the presence of all three accessory proteins expressed in concert, that it is a non-catalytic process in the current measurable time frame, and that under current conditions, activation is not increased by the presence of small molecules from the cell extract. The non-catalytic nature of HydA activation described herein is perhaps surprising, but is consistent with observations on some other radical-SAM proteins (e.g. biotin synthase), where observed turnover *in vitro* has been limited[82].

These experiments also make clear that aside from these maturation proteins, *E.coli* contains all of the requisite factors required for the maturation of [FeFe]-hydrogenases. This means that the heterologously expressed enzymes are likely acted upon by the endogenous iron-sulfur cluster assembly system, and importantly, that the substrates required for the enzymatic processes carried out by the assembly proteins are present within *E.coli.* Since *E.coli.* lacks a [FeFe]-hydrogenase and the associated maturation proteins, this observation could be interpreted as indicating that the substrates required for [FeFe]-hydrogenase maturation are present in a wide array of organisms.

The lack of enhanced activation in the presence of small molecules such as GTP and SAM, as well as the observation of rapid activation of HydA by HydEFG, suggests that the radical SAM chemistry associated with cluster assembly has already occurred during co-expression of HydEFG, such that the cluster precursor is made and requires only transfer to HydA for activation. This result would be consistent with the operation of one or more of the three of the maturation proteins acting as scaffolds or carriers upon which a cluster intermediate is assembled and then transferred to the target protein. Thus these studies beg further investigation of the nature of the activation and what proteins specifically are responsible for this event.

From these experiments it becomes possible to further dissect and probe the nature of the proteins involved in [FeFe]-hydrogenase biosynthesis. It is of interest to note that activation was not achieved in the cases of expressing the proteins singly and then combining cellular extracts, which hints that some metabolic activity not achieved in vitro is required for activation. The creation of the robust *E.coli* based *in vitro*

experimental platform presented here paves the way for investigations of specific functions of the individual accessory proteins and the fascinating catalytic reactions presumably involved in H-cluster assembly. The system and approaches described herein will be highly valuable in defining a biochemical pathway for H-cluster biosynthesis, allowing for the investigation of defined reaction mixtures where the mode of action of the accessory proteins may be studied systemically.

CHAPTER 3

FUNCTIONAL ASSIGNMENT OF A SCAFFOLD PROTEIN IN H-CLUSTER BIOSYNTHESIS.

Chapter Introduction

Following the development of an in vitro activation framework for [FeFe]-hydrogenase activation, the identification of the activating component present in the cell extract was pursued. Given that the three maturation proteins HydE, HydF, and HydG were shown to be sufficient for the activation of [FeFe]-hydrogenase, one hypothesis for H-cluster biosynthesis would involve one of these proteins acting as a delivery or scaffold protein from which an H-cluster intermediate could be delivered to HydA. Purification of the three individual maturation proteins from strains in which all three Hyd proteins were expressed in concert revealed that HydF from this background ($HydF^{EG}$) mediates [FeFe]-hydrogenase activation, and HydF expressed in the absence of HydE and HydG ($HydF^{\Delta EG}$) is not able to effect activation. The activation of HydA expressed in a genetic background devoid of HydE, HydF, and HydG ($HydA^{\Delta EFG}$) by $HydF^{EG}$ was observed to occur in a manner consistent with the previously described *in vitro* activation of $HydA^{\Delta EFG}$, but does not require the presence of HydE and HydG during the activation process[83].

This section contains text authored by McGlynn SE, Shepard EM, Winslow MA, Naumov AV, Duschene KS, Posewitz MC, Broderick WE, Broderick JB, and Peters JW. *FEBS Lett.* 2008 Jun 25;582(15):2183-7.

An Expression System for the Purification of Hyd Accessory Proteins

Results from the in vitro activation scheme detailed above revealed that the activation component was most likely a protein. Given that the protein difference between normal *E.coli* and that whose extract is able to accomplish activation of HydA, it is likely that one (or more) of the accessory proteins themselves is participating in the

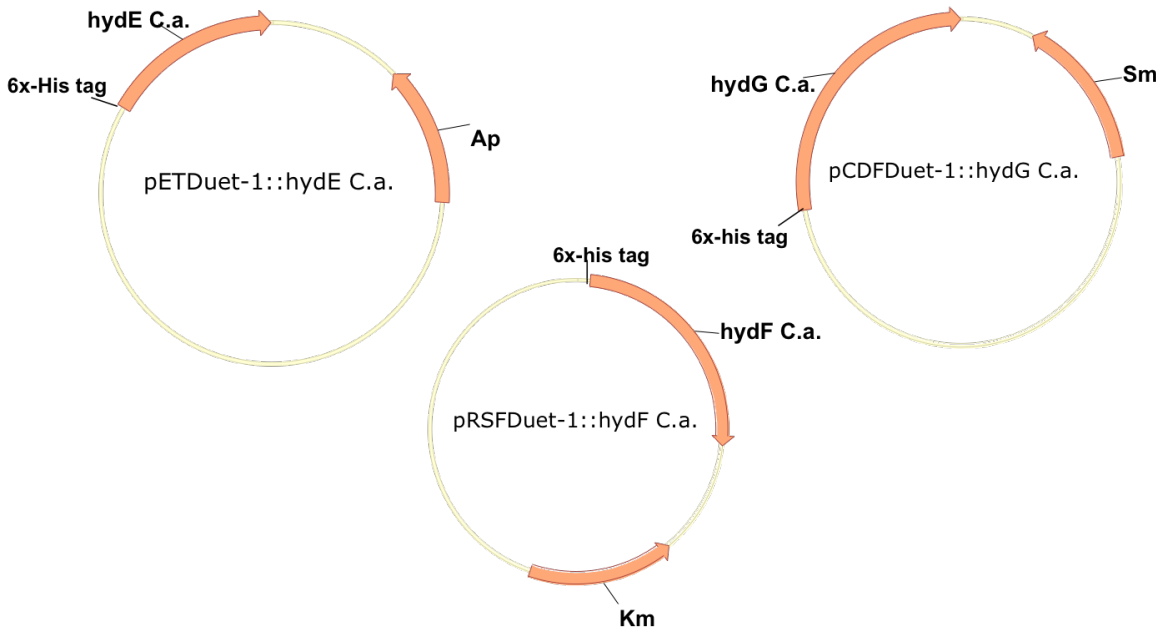

Figure 3-1 Plasmids created for the expression of Hyd proteins in *E.coli.*

activation event. To investigate these possibilities, expression vectors suitable for the isolation of the maturation proteins were created with the presence of 6x-histidine tags to allow facile purification (Figure 3.1). Using the Duet™ series of vectors, constructs were made to allow the simultaneous expression of all three of the maturation proteins where

one, two, or all three of them contained the affinity tag (Figure 3.2). With these constructs it was possible to purify the proteins from various genetic backgrounds.

Materials and Methods

The *hyd* genes were PCR amplified by the Expanded Hi-Fidelity Plus PCR system (Roche) from purified *C. acetobutylicum* or *T. maritima* chromosomal DNA utilizing gene specific primers designed to contain suitable restriction sites allowing for the presence or absence of a 6x-histidine affinity tag (His_6) and for cloning into the multiple cloning sites of Novagen Duet™ vector. Constructs were confirmed by DNA sequence analysis.

BL21 (DE3) (utilized for *C. acetobutylicum* and *C. pasteurianum* gene containing constructs) or BL21 Codon Plus (DE3) RIL (utilized for expression of *T. maritima*-derived genes) (Stratagene) cell lines were transformed with the appropriate plasmid constructs to create individual expression strains in which *hyd* gene products could be expressed singly or in tandem with or without 6x-histidine tags for affinity purification of a single protein. Transformation was followed by selection with appropriate antibiotics. Growth of the cell lines and protein expression was performed as in the case of the HydE, HydF, and HydG genes.

The genes for HydE (GenBank accession no. CAC1631) and HydG (GenBank accession no. CAC1356) were engineered to contain BamHI and SalI, while the gene for HydF (GenBank accession no. CAC1651) was engineered to contain BamHI and HindIII sites. These restriction sites allow for the introduction of a N-terminal 6x-histidine tag on each of the proteins for purification when cloned into the MCSI on the Novagen Duet™

series of vectors. The genes for HydE, HydF, and HydG were cloned into the pETDuet, pRSFDuet, and pCDFDuet vectors, respectively. The following primers were utilized:

HydE-F:5'- GGATCCGATGGATAATATAATAAAGTTAA,

HydE-R:5'- GTCGACTTAACCAATAGATTCTTTGTAGCT,

HydF-F:5'-GGATCCGATGAATGAACTTAACTCAACAC,

HydF-R:5'- AAGCTTTTAGTTCCTACTCGATTGATTAAATATTCTATCA,

HydG-F:5'-GGATCCGATGTATAATGTTAAATCTAAAG,

HydG-R:5'- CAGCTGTTAGAATCTAAAATCTCTTTGTCC.

The PCR products were gel purified and the correct size fragments isolated and ligated into the pGEM vector. Plasmid DNA from these colonies was isolated, digested with the appropriate restriction enzymes, and this fragment cloned into an expression vector (as indicated above), with sequencing information confirming the presence of the correct inserts.

Constructs encoding Hyd maturation proteins from *C. acetobutylicum* and *T. maritima* were transformed into *E. coli*- BL21 (DE3) or BL21 Codon Plus (DE3) RIL (Stratagene) cells, respectively, for protein expression. Single colonies obtained from these transformations were grown overnight in LB media prior to inoculation of larger cultures. Two liter flasks containing 1 L of low salt (5 g/L) LB broth containing 5 g/L glucose and 50 mM potassium phosphate buffer pH 7.5 were inoculated with 5 ml/L inoculums and grown at 37 ˚C and 250 rpm shaking. These cultures were allowed to grow to an OD_{600}= 0.5 before being induced by the addition of IPTG to a final concentration of 1 mM. At the point of induction, 0.075 g/L ferrous ammonium sulfate

(FAS) was also added. The cultures were grown an additional 2 hours, either at room temperature or 37˚C at which time an additional aliquot of 0.075g/L FAS was added. The cultures were then transferred to a 10˚C cold box and purged with N_2 overnight. Cells were harvested by centrifugation and the resulting cell pellets were stored at – 80˚C until further use.

All procedures were carried out under anaerobic conditions in a Coy chamber or in sealed bottles. Cell pellets were thawed and resuspended in a lysis buffer containing 20 mM sodium phosphate or 10 mM Hepes, pH 7.4, 0.5 M NaCl, 5% glycerol, 20 mM imidazole, 20 mM $MgCl_2$, 1 mM PMSF, 1% TritonX-100, 0.07 mg DNAse and RNAse per gram cell, ~ 0.6 mg lysozyme per gram cell. The lysis mixture was stirred for one hour, after which time the lysate was centrifuged in gas tight bottles at 38,000 x g for 30 minutes. The resulting extracts were loaded onto a 5 mL HisTrap^{TM} HP Ni^{+2}-affinity column (GE Healthcare) or a gravity flow column loaded with HisTrap^{TM} HP Ni^{+2}-affinity resin. In the case of utilizing the pre-packed column, the resin was pre-equilibrated with 20 mM sodium phosphate buffer, pH 7.4, 0.5 M NaCl, 5% glycerol, 10 mM imidazole (buffer A). The column was subsequently washed with 15 column volumes of buffer A. Protein elution was accomplished by increasing the imidazole concentration in a stepwise manner from 20% to 50% to 100% buffer B (20 mM sodium phosphate buffer, pH 7.4, 0.5 M NaCl, 5% glycerol, 500 mM imidazole), with each stage washing for 5 column volumes. Fractions of interest were pooled, dialyzed into 50 mM HEPES, pH 7.4, 0.5 M NaCl, and 5% glycerol, and concentrated using an Amicon concentrator fitted with a YM-10 membrane. For the gravity flow columns, a buffer with

composition 20 mM Hepes, pH 7.4, 0.5 M NaCl, 20 mM imidazole (buffer C) was utilized for column equilibration and washing. Following column washing with 15- 20 column volumes buffer C, protein elution was accomplished by increasing the imidazole concentration in a stepwise manner from 20% to 50% to 100% buffer D (20 mM Hepes buffer, pH 7.4, 0.5 M NaCl, 5% glycerol, 500 mM imidazole). Protein was flash frozen in liquid N_2 and stored at – 80 ˚C or in liquid N_2 until further use. Protein concentrations were determined by the Bradford method using bovine serum albumin as the standard[84]. Iron content was evaluated spectrophotometrically by the method of Fish[85]. For UV and visible absorption experiments, samples were transferred to an anaerobic cuvette within an MBraun glove box. Room temperature UV/vis absorption data were acquired using a Cary 6000i UV/vis/near-IR spectrophotometer (Varian). Reduced samples were prepared by the addition of dithionite (2 mM final concentration). UV/vis spectra were collected at a data interval of 0.5 nm and a scan rate of 60 nm/min.

Purification of HydF from cellular lysates over SP sepharose was accomplished using a column equilibrated with 50mM hepes 50mM NaCl 1mM DT 5% glycerol pH=7.4 buffer. 30ml lysate was loaded at 1.8ml/min, then washing with buffer until 280 decreased at which time a linear gradient was initiated to a 1000mM NaCl buffer over 90 min (other components the same).

Determination of the ability of a given Hyd maturation protein to activate $HydA^{\Delta EFG}$ was tested by combining purified inactive $HydA^{\Delta EFG}$ with the respective purified Hyd protein in addition to the assay reagents methyl viologen and dithionite. In this scheme, dithionite serves as an electron donor with methyl viologen acting as an

electron conduit to HydA. Assay mixtures were prepared in a Coy anaerobic chamber at 25˚C by mixing $HydA^{\Delta EFG}$ (4.4 µg) with an excess of purified maturation protein (typically ~350 µg) in 3 ml sealable glass vials. Each 2 ml reaction was performed in 50 mM Hepes, pH 7.4, 500 mM NaCl in the presence of 20 mM sodium dithionite. The reactions were sealed, removed from the anaerobic chamber, and degassed to remove residual hydrogen from the headspace. Hydrogenase assays were performed at 25˚C and initiated by the addition of oxidized methyl viologen (10mM final concentration) and the production of H_2 was monitored using gas chromatography as described previously [8]. Experiments with the *T. maritima*-derived *hydF* gene product were performed as described above, with 0.5 ml cellular extract added to 4.4 µg $HydA^{\Delta EFG}$.

Experiments directed at determining the maximal activation of $HydA^{\Delta EFG}$ achievable and the effects of increasing the amount of $HydF^{EG}$ were performed by adding increasing amounts of $HydF^{EG}$ (from 0.07-35 nmol) to a constant amount of $HydA^{\Delta EFG}$ (0.07 nmol). Assays were performed as described above following a 20 minute incubation period and raw data were fit using linear regression to obtain hydrogen evolution rates. The quantity of $HydA^{\Delta EFG}$ activated was estimated by comparison to the activity of native [FeFe] hydrogenase purified from *Clostridium pasteurianum* (CpI).

Although it was previously shown that the *in vitro* activation of hydrogenases from a variety of sources is possible[75], CpI was chosen as a reporter of hydrogenase activity in these studies as the purified wild type enzyme is readily available and allows for reliable activity comparison. Heterologously isolated HydA co-expressed with the requisite accessory proteins was not utilized for these studies because previous work showed that

preparations are comprised of enzyme in which maturation is not fully complete[54] and therefore does not provide an accurate benchmark for full hydrogenase activity. Under the conditions described above the activity of CpI was determined to be 304 μmol H_2*min^{-1}*mg^{-1} which was defined as 100% activity. In order to determine the proportion of $HydF^{EG}$ competent to activate $HydA^{\Delta EFG}$, a constant amount of $HydF^{EG}$ (0.01 nmol) was titrated with increasing amounts of $HydA^{\Delta EFG}$ (0.07, 0.14, 0.35, 0.69, 1.38 and 3.46 nmol). Assays were prepared in triplicate as described above, and the observed rates of H_2 evolution were converted to an activated fraction of $HydA^{\Delta EFG}$ present in assays.

FTIR samples were prepared under anaerobic conditions in a Coy anaerobic chamber containing ~ 3% H_2 atmosphere. FTIR spectra were measured using a Bruker IFS/66s FTIR spectrometer interfaced to a home- built stopped-flow drive system with the sample cuvette and drive system maintained inside an anaerobic chamber (O2 < 1.1 ppm) as described elsewhere[86]. The IR cuvette was thermostated at 25°C. For these measurements, one drive syringe contained the protein[87] sample in 50 mM HEPES pH 7.4 buffer containing 500 mM NaCl and 500 mM imidazole while the other syringe was loaded with buffer only. Spectra were measured at 4 cm-1 resolution. The IR cuvette path length was calibrated at 47.6 μm. The spectrum required an arbitrary background correction to make it flat.

Results

Purification of the HydA Activating Component HydF

The 6x-histidine tagged proteins eluted from the Ni^{+2} affinity column between 100 mM and 250 mM imidazole. In cases when the three Hyd accessory proteins were coexpressed, co-purification of non-His-tagged Hyd proteins was observed at low to intermediate imidazole concentrations (60-100 mM). For example, HydF co-eluted with HydE-His_6 and HydG-His_6, and HydE co-eluted with HydF-His_6. At higher imidazole concentrations, however, pure 6x-histidine tagged protein was obtained.

With these constructs it was possible to investigate which protein or combination of proteins was directly responsible for the activation of HydA. Isolation of the proteins and

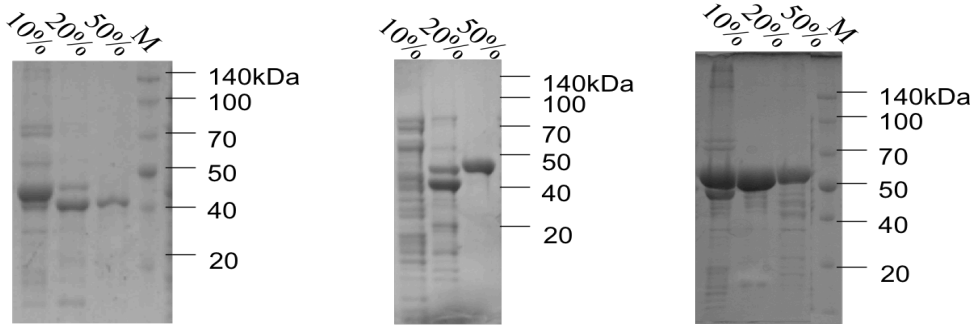

Figure 3-2 SDS-PAGE gels showing purification of 6x-histidine tagged FeFe-hydrogenase maturation proteins from heterologous expressions in E.coli. Gels left to right HydE, HydF, and HydG. Percentages above the lanes indicate the percentage of a 500mM imidizole containing buffer. "M" denotes markers, with masses in kDa shown on the right.

thus allowed for the isolation of the activating factor. Incubation of each of the Hyd maturation proteins (HydE, HydF, or HydG), purified from cellular extracts containing all three of these Hyd proteins, with inactive purified $HydA^{\Delta EFG}$ revealed that neither

$HydE^{FG}$ nor $HydG^{EF}$ alone has the ability to activate $HydA^{\Delta EFG}$. Purified $HydF^{EG}$, however, was competent to activate $HydA^{\Delta EFG}$. The ability to activate $HydA^{\Delta EFG}$ increased upon $HydF^{EG}$ purification, pointing to $HydF^{EG}$ as the sole species responsible for activation (Figure 3.3A). In contrast, $HydF^{\Delta EG}$ was not competent to activate

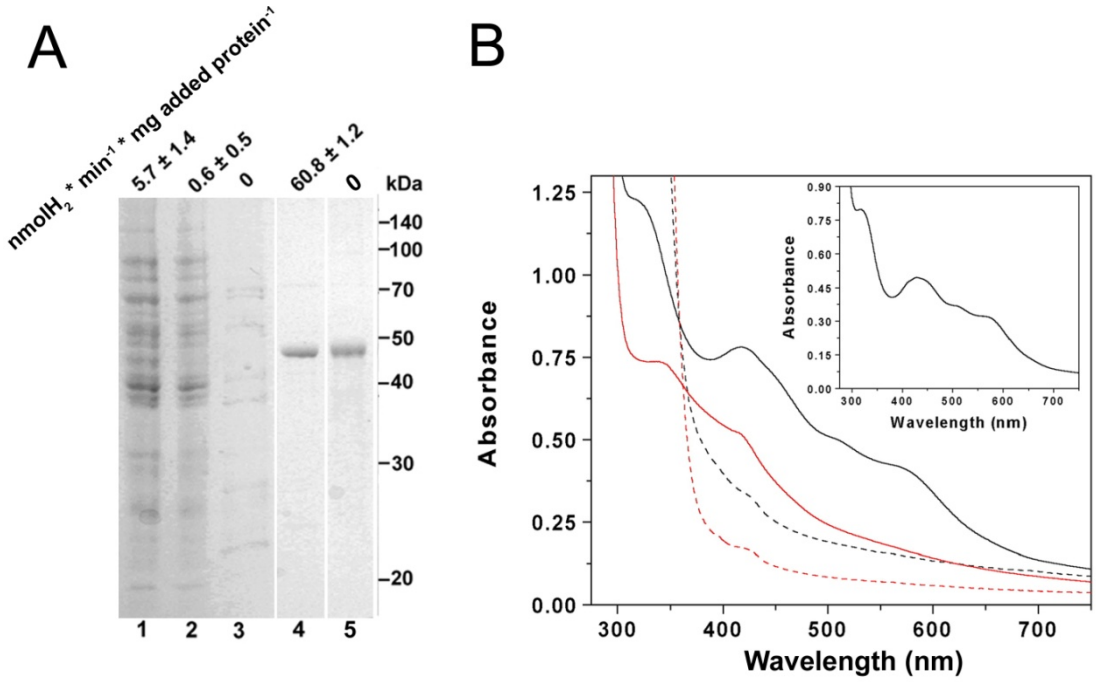

Figure 3-3 (A) SDS/PAGE analysis (10% gel) showing IMAC purification of HydF co-expressed with HydE and HydG (lanes 1–4) and purified HydF expressed in the absence of HydE and HydG (lane 5). Lane contents left to right: 1 – clarified cell lysate, 2 – column flow through, 3 – 60 mM imidazole elution, 4 – 250 mM imidazole elution. Lane 5 – purified HydF expressed in the absence of HydE and HydG. Activation of $HydA^{\Delta EFG}$was achieved by combining aliquots obtained during purification corresponding to the respective lanes with 4.4 µg $HydA^{\Delta EFG}$. Observed hydrogen evolution rates normalized to the amount of added protein are indicated above each lane. (B) UV–visible spectra of HydF expressed either singly (black trace, 175 µM) or in the presence of HydE and HydG (red trace, 170 µM). Dotted lines represent spectra obtained upon the addition of sodium dithionite. Inset: difference spectrum obtained by subtraction of iron-content-normalized spectra of the as-isolated proteins ($HydF^{\Delta EG}$, $HydF^{EG}$) highlighting the differences between the two isolated forms of HydF.

$HydA^{\Delta EFG}$, demonstrating an essential role for HydE and HydG in forming the activation-competent form of HydF. These experiments thus reveal a direct functional interaction between the HydA and HydF proteins, and lay a foundation for future experiments on the form of HydF from alternate genetic backgrounds including $HydF^{E}$ and $HydF^{G}$ and the clusters that may be harbored by these forms.

These results, which demonstrate the ability of purified $HydF^{EG}$ to activate $HydA^{\Delta EFG}$ in the absence of any other proteins or small molecules, support the hypothesis that HydF serves as a scaffold for assembly of a cluster that is subsequently transferred to $HydA^{\Delta EFG}$, thus converting it to an active holoenzyme. The observation that HydE and HydG must be coexpressed with HydF, but are not required in the $HydA^{\Delta EFG}$ activation assays, supports the hypothesis that although HydE and HydG serve to assemble a cluster precursor on HydF, they are not required for the subsequent transfer of the precursor to HydA. Furthermore, although HydF purified in these studies exhibits GTPase activity (Eric Shepard unpublished data), the presence of GTP was observed not to affect the amount of activated $HydA^{\Delta EFG}$ produced by $HydF^{EG}$, arguing against a direct role for GTP hydrolysis in cluster transfer to $HydA^{\Delta EFG}$.

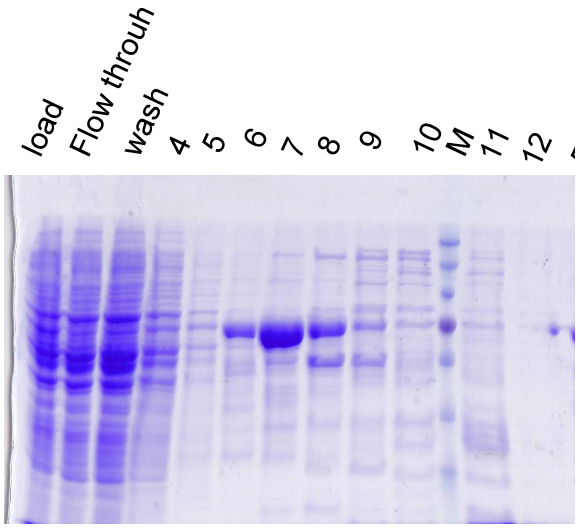

Figure 3-4 Purification of HydF by cation exchange on SP-Sepharose resin running NaCl gradient from 50-1000mM. HydF fractions of purity were observed to elute between 20-35% of the 1000mM NaCl containing buffer. Numbers above gel correspond to fraction collected.

As an alternative to the histidine tag approach for the purification of HydF, a purification scheme was developed in which the protein could be purified by via cationic exchange on SP-sepharose. HydF(*C.a.*) is predicted by sequence analysis to have a pKa of ~ 8.0 and thus exists as a potential target for facile ion exchange purification. From cell lysates in which the three maturation proteins HydE, HydF, and HydG were co-expressed, the HydF protein was observed to elute as a relatively pure band as judged by SDS-PAGE (Figure 3.4), and these fractions were observed to exhibit comparable activation ability as the histidine tagged variant, demonstrating ion exchange as a viable means to achieve purification of the HydF scaffold.

The involvement of HydF as a specific scaffold protein in H-cluster biosynthesis is analogous to other complex cluster assembly pathways, in particular that of

nitrogenase, where the FeMo cofactor is synthesized on a NifEN scaffold protein prior to being transferred to apo-nitrogenase to accomplish activation[73]. In addition, the Isc and Suf systems for iron-sulfur cluster assembly utilize specific scaffold proteins for the delivery of metal clusters to target apo-proteins[80,88,89]. This observation adds to what may be a common theme of complex metal cluster assembly occurring on a surrogate host until transfer occurs to form a holo-enzyme.

Initial Characterization of HydF

Purified HydF contains iron, however the stoichiometry depends on whether it is expressed in the absence (0.9 Fe/protein ± 0.2) or presence (1.6 Fe/protein ± 0.1) of a HydE/HydG background. UV-vis spectroscopic analysis of the as-isolated form of $HydF^{\Delta EG}$ shows a protein-centered band at 280 nm with additional features occurring at ~320, 420, 510 and 575 nm (Figure 3.3B, solid black line). UV-vis analysis of $HydF^{EG}$ shows a protein-centered band at 280 nm with shoulders occurring at ~335 and 415 nm (Figure 3.3B, solid red line). The spectroscopic features present in both $HydF^{\Delta EG}$ and $HydF^{EG}$, which decrease in intensity upon addition of dithionite (Figure 3.3B, dashed lines), are attributed to ligand to metal charge transfer bands characteristic of iron-sulfur clusters. This observed ability of HydF to bind iron-sulfur clusters is consistent with a previous report of the initial characterization of HydF from *T. maritima*[56]. The difference spectrum ($HydF^{\Delta EG}$- $HydF^{EG}$,Figure 3.3B, inset) highlights the spectroscopic differences between the purified HydF proteins from the two different backgrounds. Notably, the difference spectrum exhibits features characteristic of [2Fe-2S] clusters, such as those

previously characterized in BioB[82]. These data indicate clear differences in cluster composition between $HydF^{\Delta EG}$ and $HydF^{EG}$, thus suggesting that HydE and HydG affect the cluster composition of HydF. Further spectroscopic studies are underway to elucidate the differences between $HydF^{\Delta EG}$ and $HydF^{EG}$.

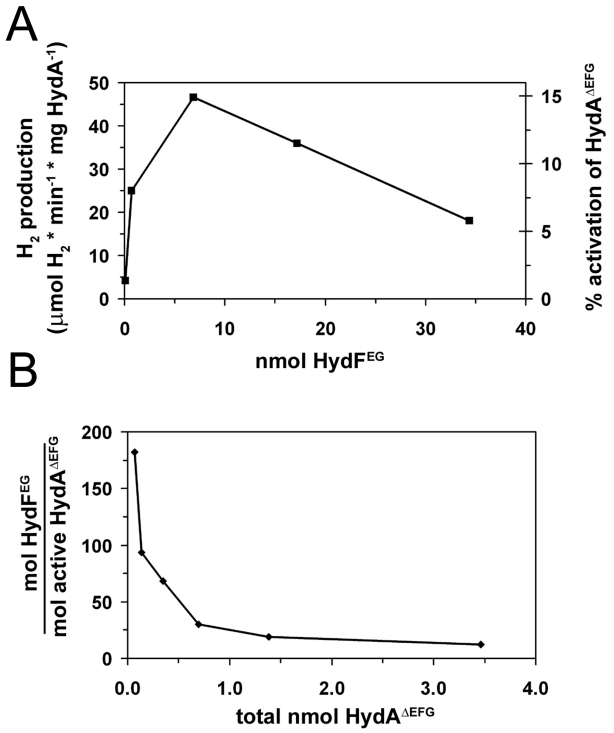

Figure 3-5 Rates of H_2 evolution and percent of $HydA^{\Delta EFG}$ activation as compared to the wild type enzyme activity are plotted as a function of the amount of $HydF^{EG}$ present in an assay mixture (0.07 to 35 nmol) at constant $HydA^{\Delta EFG}$(0.07 nmol). (B) Effect of increasing the amount of $HydA^{\Delta EFG}$ present in hydrogen production assays while holding the amount of $HydF^{EG}$ constant. Purified $HydF^{EG}$ (0.01 nmol) was titrated with increasing amounts of $HydA^{\Delta EFG}$ (0.07 to 3.46 nmol). The rates of hydrogen evolution were converted to a molar quantity of activated $HydA^{\Delta EFG}$ present in the reaction mixtures by comparing to the wild type isolated enzyme. The ratio of $HydF^{EG}$ to activated $HydA^{\Delta EFG}$ was plotted as a function of total $HydA^{\Delta EFG}$ present.

Maximal activation of $HydA^{\Delta EFG}$ by $HydF^{EG}$ was investigated by performing assays in which a constant amount of $HydA^{\Delta EFG}$ was titrated with increasing amounts of

$HydF^{EG}$ (Figure 3.5A). Based on the known activity of 304 μmol H_2*min^{-1}*mg^{-1} for native [FeFe]-hydrogenase purified from *Clostridium paterianum*, these experiments indicate that approximately 15% of the heterologously expressed $HydA^{\Delta EFG}$ is capable of being activated in this process. The inability to accomplish full activation of $HydA^{\Delta EFG}$ may result from metal content heterogeneity and/or improper folding in purified $HydA^{\Delta EFG}$, either of which might result from heterologous expression in the absence of the full complement of maturation proteins. In addition, a decreased percentage of $HydA^{\Delta EFG}$ activation is observed at higher $HydF^{EG}$ to $HydA^{\Delta EFG}$ ratios which is similar to effects observed for other systems in which scaffold proteins operate including the Isc and Nif systems[90,91].

Activation Stoichiometry of HydA Activation by HydF

To determine the fraction of $HydF^{EG}$ in preparations capable of activating $HydA^{\Delta EFG}$, a constant amount of $HydF^{EG}$ was titrated with increasing amounts of $HydA^{\Delta EFG}$, and the resulting amount of activated $HydA^{\Delta EFG}$ was determined. In this analysis, the lowest observed ratio of $HydF^{EG}$ to activated $HydA^{\Delta EFG}$ provides an estimate of the fraction of $HydF^{EG}$ molecules that are capable of activating $HydA^{\Delta EFG}$. As can be seen in Figure 3.5B, the ratio of $HydF^{EG}$ to activated $HydA^{\Delta EFG}$ decreases as the total $HydA^{\Delta EFG}$ present in the assay increases; the ratio approaches a limit of ~10 $HydF^{EG}$ per activated $HydA^{\Delta EFG}$ (Figure 3.5B). The presence of excess $HydA^{\Delta EFG}$ presumably favors the transfer of the activating component from $HydF^{EG}$ to $HydA^{\Delta EFG}$, and prevents the decrease in maturation observed at high ratios of $HydF^{EG}$ to $HydA^{\Delta EFG}$

as seen in Figure 3.5A. These results reveal that ~10% of the $HydF^{EG}$ protein present in these preparations is competent to activate $HydA^{\Delta EFG}$, assuming a 1:1 activation stoichiometry. The observation that only ~10% of $HydF^{EG}$ is competent for activation may suggest the existence of multiple cluster intermediates at differing stages of synthesis on $HydF^{EG}$. This would be consistent with the role of HydF as a scaffold for the entire sequential assembly process. Unfortunately, given the partial occupancy of the activating component and the presumed heterogeneity of the system it is not possible to assign spectroscopic signatures of the activating component that would allow parallel characterization and quantification. Efforts are underway to increase the occupancy of the activating component bound to HydF to pave the way for detailed analysis.

One possible mode in which the cluster occupancy of HydF might be increased in

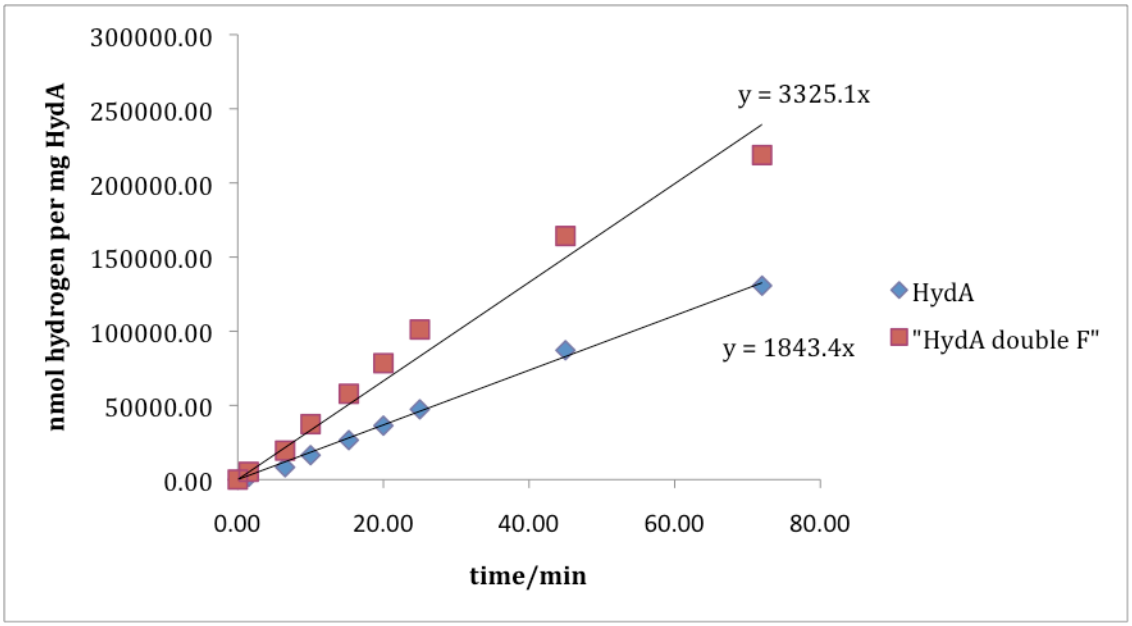

Figure 3-6 Hydrogenase evolution from HydA purified from cell lines expressing HydA, HydE, HydF and HydG. The *hydF* gene was present on two individual plasmids in the case of "HydA double F" (red squares). Equations obtained from linear regression are displayed above and below the traces for HydA purified from double and single *hydF* gene copy cell lines.

heterologous expressions is in the modulation of gene dosage in the cell, which could be expected to change the protein level accordingly. The observation of HydF behaving as a scaffold suggests that higher occupancy may be afforded by increasing the cellular levels of HydE and HydG – the proteins that presumably act on HydF to load it. In addition, HydF as a scaffold for H-cluster assembly itself may be a bottleneck for HydA maturation in heterologous expression systems and varying the intracellular level of HydF may have the potential to affect levels of active HydA in the cell. In an effort to investigate the latter possibility, a BL21DE3 *E.coli* expression cell line was created which contained plasmids such that two copies of the *hydF* gene were present. This was accomplished utilizing plasmids obtained from Paul King at NREL, in conjunction with the pRSF::*hydF* plasmid created in this work.

As seen in Figure 3.6, expression of the HydA protein with HydE and HydG being present singly on respective plasmids, and two gene copies of HydF present on

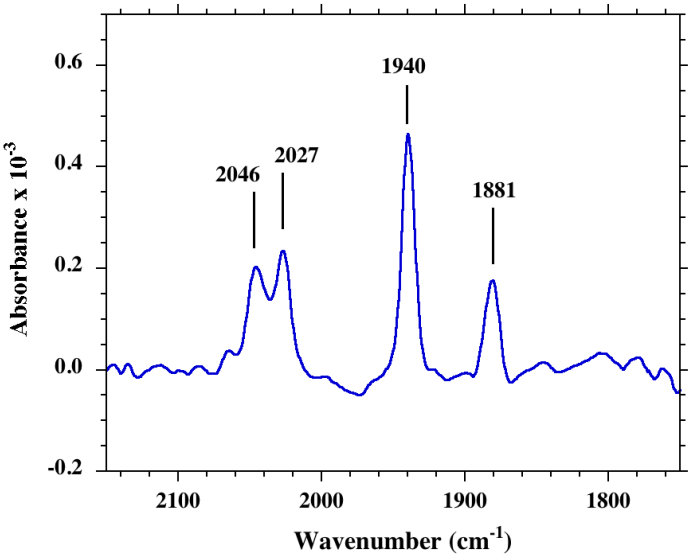

Figure 3-7 FTIR spectrum of a $HydF^{EG}$ sample (36 mg/mL). The major observed bands are indicated. By comparison to the H-cluster FTIR, bands at 1940 and 1881 cm^{-1} can be attributed stretching vibrations of coordinated carbonyl groups while the bands at 2046 and 2027 cm^{-1} are most likely vibrations of bound cyanide.

respective plasmids, yields purified HydA with an approximate 2 fold increase in the hydrogenase activity compared with an expression in which only one gene copy of each maturation protein is present. This demonstrates that the protein level of HydF in the cell is a key factor in the loading level of HydA, and that HydF – as could be expected from its role as a scaffolding protein for H-cluster assembly – is a bottleneck in the assembly process. Further, this observation suggests that further increases in HydA activity may be feasible from other perturbations which effect protein expression levels.

FTIR Analysis of the HydF Scaffold

The observation that the HydF protein is able to confer hydrogenase activity to HydA implies that the protein may be involved in the delivery of a subcluster which contains iron ligated by CO and CN^-. To test this possibility, the $HydF^{EG}$ protein was purified, concentrated, and subjected to analysis by FTIR, which is capable of detecting Fe-CO and Fe-CN^- stretching frequencies. Upon analysis, features consistent with this type of iron ligation are observed, thus indicating that the scaffold does indeed harbor the CO/CN^- coordinated iron which is found in the active site of the holo-hydrogenase (Figure 3.6). In comparison with previous studies on the holo-hydrogenase protein[86,92] the vibrational bands compare well with two distinct CO stretches and two distinct CN^- stretches. However, the bridging CO stretch, which is observed at lower wave numbers than the two aforementioned features, is not present. This may be due either to its absolute absence, or that the feature is low intensity and the sample in the study is not sufficiently concentrated to observe this.

Summary and Conclusions

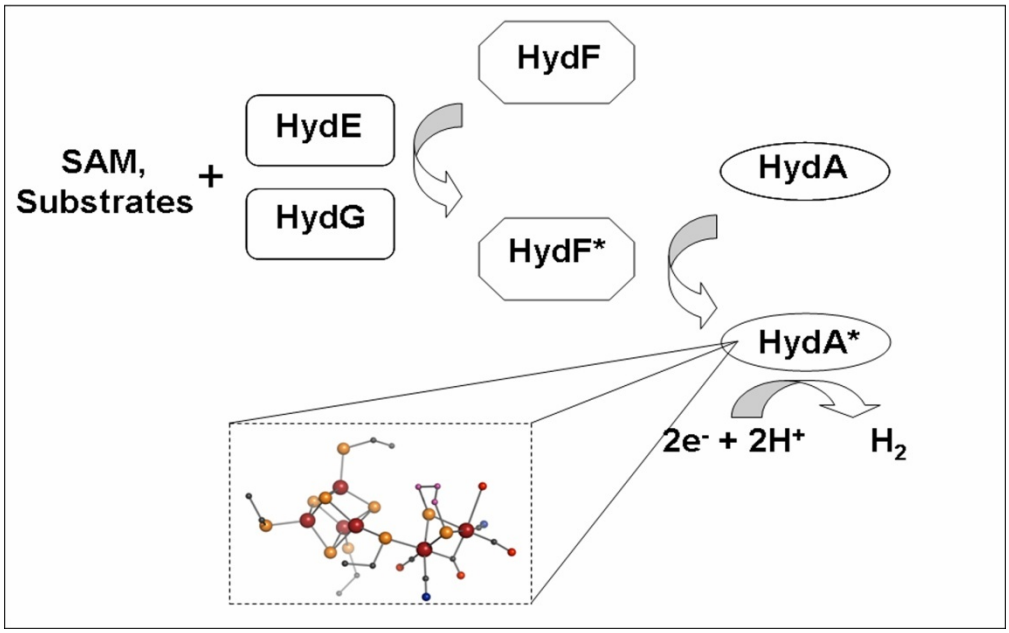

Figure 3-8 Cartoon diagram of FeFe-hydrogenase active site cluster proceeding by the actions of the radical-SAM proteins HydE and HydG on the HydF scaffold followed by delivery of an H-cluster intermediate to HydA.

These results provide the first functional analysis of [FeFe]- hydrogenase maturation in a system using purified proteins, and identify HydF as acting in the terminal step of this process. Given the difference in iron content and spectroscopic features between $HydF^{\Delta EG}$ and $HydF^{EG}$, the ability of purified $HydF^{EG}$ alone to activate $HydA^{\Delta EFG}$, and the apparent functional similarity between HydF and other characterized scaffolds for iron-sulfur cluster assembly, it is concluded that HydF operates as a scaffold or carrier protein in H-cluster assembly. A proposed model for [FeFe]- hydrogenase maturation involves HydE and HydG acting on HydF to assemble a cluster precursor, which is subsequently transferred to HydA to effect activation (Figure 3.8). This working model provides a strong foundation from which to base further experiments

directed at elucidating the contributions of HydE, HydG, and other endogenous proteins on the formation of the remarkable H-cluster active site.

Sill not addressed with the above characterization of HydF is the chemistry associated with the nucleotide hydrolyzing ability of the protein, as well as the specific actions of the radical SAM proteins HydE and HydG on HydF. Presumably, HydA activation competent HydF exists with the 2Fe subcluster of the H-cluster and contains the diatomic CO and CN^- ligands, as well as the dithiolate linker. Future studies on the HydE and HydG protein will thus likely result in the observation of these enzymes to catalyze chemistry consistent with the deposition of these ligands on HydF. The HydF protein itself may well harbor chemical ability in the synthesis and trafficking of these ligands as well, and future studies aimed at elucidating the specific role of GTPase activity and the FeS cluster harbored by the protein will further aid in addressing these roles.

Further studies will also be important in defining the nature of the cluster harbored by HydF prior to activation of the HydA protein. Open questions include oxidation state, the role of this in transfer to the HydA protein, and what the state of the cluster is prior to modification. Presumably the HydF protein in its loaded form contains the entire unique ligand set (dithiolate linker and CO/CN^- molecules), and the deposition of these could be expected to occur in a sequential process. The previously enumerated hypothesis for H-cluster assembly suggests that the first step to cluster modification occurs with the ligation of the dithiolate linker[79]. The results and techniques presented here provide a framework for addressing this and other questions in that isolatable

intermediates of HydF are obtainable, for example HydF that has been co-expressed in the presence of HydG and not HydE. These intermediates will be valuable means by which to further dissect and partition the assembly of the H-cluster active site, and thus with this work a precedence exists for biochemical experimentation that addresses questions of greater and greater detail.

CHAPTER 4

THE ROLE OF HYDG IN THE MATURATION OF THE [FEFE]-HYDROGENASES

Chapter Introduction

Previous work[93] indicated that the HydE and HydG proteins were closely related evolutionarily to other radical SAM proteins LAM, BioB, and ThiH. Indeed, HydE and HydG are commonly annotated as BioB and ThiH respectively. Biochemical characterization of these Hyd accessory proteins showed that they both are capable of cleaving SAM to dAdo[55,56]. Aside from this little has been reported as to the possible biochemical mechanisms by which these proteins may contribute to assembly of the H-cluster. The close similarity of HydG to ThiH suggests that HydG may be involved in cleavage of tyrosine to the reaction products *p*-cresol and dehydroglycine[68], and indeed this has been shown to be the case insofar as the production of *p*-cresol was observed, and a possible mechanism by which these products further react to form the H-cluster dithiolate linker has been proposed stemming from the presumed formation of dehydroglycine by the enzyme[94]. Continuing from this, this chapter discusses the relationships between the HydG and ThiH proteins and shows evidence that the reaction products from cleavage of tyrosine by HydG include cyanide and carbon monoxide, thus indicating that the enzyme contributes these to the active site H-cluster from the same reaction intermediate.

This section contains text authored by Driesener RC, Challand MR, McGlynn SE, Shepard EM, Boyd ES, Broderick JB, Peters JW, Roach PL. Angew Chem Int Ed Engl. 2010 Feb 22;49(9):1687-90 and Shepard EM, Duffus BR, McGlynn SE, Challand SE, Swanson K, Roach PL, Peters JW, Broderick JB. Submitted Feb 2010

Materials and Methods

Protein expression and purification was performed as for Hyd accessory proteins described in previous chapters, and alternatively with triton-X being omitted from lysis solutions and cells were grown in 2YT media. Lysis in these cases was achieved by sonication (3 x 10 min, 1 s bursts, 20W) and the lysate cleared by centrifugation (SLA1500, 13000 rpm, 40 min, 4 °C). The protein was reconstituted as described above for the HmdB protein. Activity assays were prepared by adding the substrates AdoMet and tyrosine (both to final concentrations of 1 mM) to chemically reconstituted HydG (60 – 200 μM). The solution was equilibrated at 37 C and the HydG catalyzed reaction initiated by addition of sodium dithionite (1 mM). Reactions were stopped by precipitating the protein by the addition of acid (20% perchloric acid (15 μL)) followed by centrifugation (10 min, Eppendorf 5415D microcentrifuge, maximum speed). The solution was neutralized by immediate addition of 1 M phosphate buffer pH 7.5 (20 ml) and 1 M NaOH (30 ml) before removal of the precipitated protein by centrifugation (10 min, Eppendorf 5415D microcentrifuge, maximum speed). To establish a cyanide calibration plot, potassium cyanide standards at a range of concentrations (25 - 400 mM) in 100 mM phosphate pH 7.5 were prepared and subjected to the same conditions as HydG assays, with detection performed as described previously[95]. Particular care was taken in preparing these cyanide standards, so that they lacked tyrosine but contained HydG and dithionite. They were prepared in triplicate, incubated for reaction times of 3 min, 12.5 min and 45 min and subjected to the same precipitation and neutralization conditions as assay samples. A calibration curve for cyanide concentration was

constructed from average values obtained from these three sets of cyanide standards. Detection and quantification of dAdo, *p*-cresol, and glyoxylate were performed as described previoulsy[68].

Assays for HydG-catalyzed CO formation were carried out in an anaerobic chamber (Mbraun, O2 < 1 ppm) and in sealed 1.4 ml anaerobic cuvettes (Spectrocell Inc). Reconstituted samples of HydG were assayed for CO production using deoxyhemoglobin as a reporter molecule. The protein was reconstituted as described for the HmdB protein above. Assays contained HydG (10 - 89 μM) in buffer (50 mM HEPES, pH 7.4, 0.5 M KCl, 5 % glycerol) supplemented with 1 – 4.8 mM dithionite. Deoxyhemoglobin was added to a final concentration of less than 10 μM. UV-visible spectra were recorded before and after addition of deoxyhemoglobin; given the low extinction coefficient of reduced HydG in the visible region, this procedure allowed for confirmation of the presence of deoxyhemoglobin, which shows characteristic λmax values at 430 nm (Soret band) and 555 nm (visible band). Tyrosine was then added, and the assay was initiated by addition of AdoMet. Assay volume was maximized to 1330 μl to limit the amount of available headspace in the 1.4 mL cuvettes in an attempt to minimize loss of CO from the aqueous phase. Spectral changes were then monitored as a function of time on a Cary 6000i UV/vis/near-IR spectrophotometer (Varian) at a data interval of 1 nm. CO formation assays were carried out in 50 mM HEPES pH 7.4, 0.5 M KCl, 5% glycerol. Raw data were corrected for the contribution of reduced HydG to the visible absorption. The ΔA419 and Δε419 were used to calculate the concentration of Hb-CO at each time point in order to determine the rate of CO formation. Data was fit to a biphasic

exponential, and k_{cat} values were calculated as previously described[96]. Control experiments lacking AdoMet were carried out using 36 μM HydG at 7.57 ± 0.26 Fe/protein in 3.22 mM DT and 10.5 μM deoxyhemoglobin all in 50 mM HEPES pH 7.4, 5% glycerol buffer. Tyrosine (1 mM) was added and spectra were monitored for 2 hrs at 30 °C. No spectral changes were observed. At the 2 hr mark, the sample was transferred back to the MBraun box and 1.3 mM SAM was added. Within 15 minutes spectral changes associated with the Soret band consistent with Hb-CO formation were observed.

For comparison of HydG to ThiH, sequences were obtained from NCBI BLAST using the *Clostridium kluyveri DSM 555* HydG (YP_001397177.1) protein as a query. ThiH sequences were obtained using the *E.coli* ThiH protein as a query in an analogous search (NP_418417.1). Sequences were aligned using the CLUSTALX (2.0) program.

Results

HydG is Closely Related Protein to the ThiH Protein

Biochemical investigation of the ThiH and HydG proteins has revealed that the two proteins share in common the activity of coupling the generation of the dAdo radical to the alpha-beta bond cleavage of tyrosine, and that the two enzymes share in common the products *p*-cresol and dehydroglycine, which is detected via quantification of the hydrolysis product glyoxylate[68,97]. Inspection of ThiH and HydG sequences reveals that the proteins share in common a number of conserved residues that may be responsible for binding of SAM and the tyrosine substrate (Figure 5.1A). At the sequence level, the major difference between the two proteins appears to be at the C-terminus, where HydG

has an approximately 80 amino acids in length compared to ThiH, and the presence of a conserved CxxC motif near this terminus (where "x" denotes any amino acid, Figure 5.1B). This difference in sequence similarity suggests that the two proteins may be involved in a similar chemical process, and specifically motivates the hypothesis that HydG generates the same reaction intermediates as ThiH. The C-terminal extension of residues on HydG compared to ThiH could be envisaged as being responsible for further reactivity related to H-cluster biosynthesis. Previous mutational studies[54] indicated that a point mutation in either of the two cysteines in the C-terminal CxxC motif resulted in the

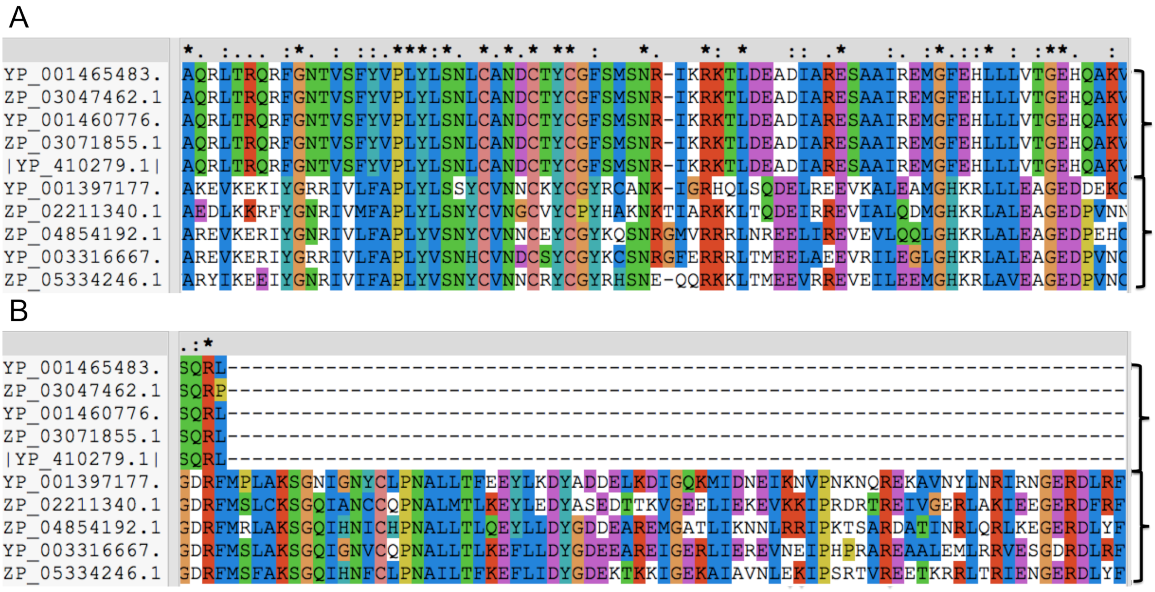

Figure 4-1. Sequence alignments of (A) the 4Fe-4S cluster binding domains of ThiH and HydG and (B) the C-terminal domains of ThiH and HydG. Residues with an asterisk above the column indicate conservation. ThiH sequences are indicated by the top bracket, and HydG sequences are indicated by the bottom bracket (both panes). The Sequences were aligned with ClustalX using sequences the top sequences obtained from a BLAST search using the *E.coli* ThiH protein (YP_001465483.1) and the *Thermotoga Maritima* HydG protein (YP_001397177.1). Accesssion numbers for the sequences are listed on the left.

formation of non-active HydA in heterologous expression studies where the three maturation proteins were co-expressed with HydA, indicating that within this domain chemistry that is critical for H-cluster assembly is taking place. In this case, it could be that the mutated HydG is incapable of forming H-cluster related products, and is operating effectively as ThiH in the cell to produce dehydroglycine and *p*-cresol.

Given the close grouping of the HydG and ThiH proteins upon phylogenetic analysis (Figure 4.2C), possible routes to their differentiation could be envisaged including gene duplication followed by divergence, gene fusion in a species (to add the

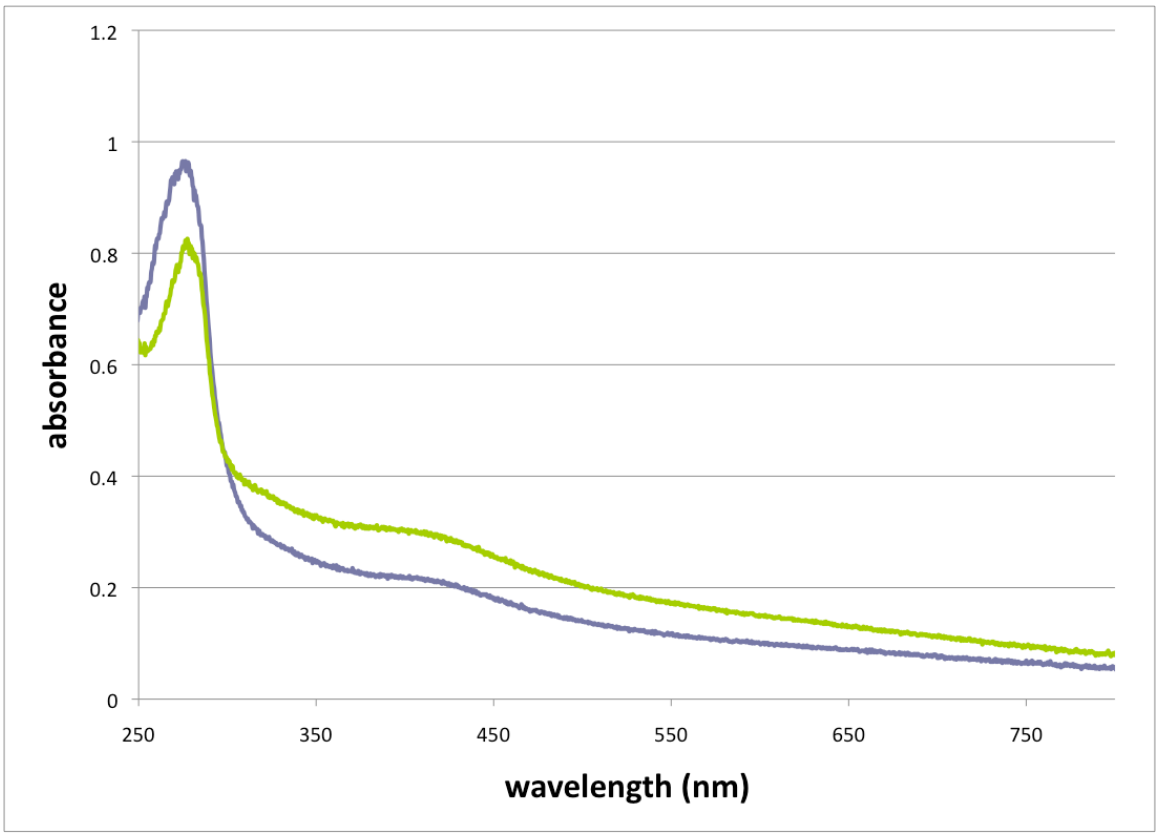

Figure 4-2. UV/visible spectrum of the HydG (blue trace) and HydG^{C386S} mutant (green trace) proteins. Protein concentrations were measured by the Bradford method to be 105µM and 104µM for the two respectively.

C-terminal domain of HydG to ThiH), or possibly gene fission to produce ThiH from HydG. In an evolutionary context, it would be likely that the HydG protein came after ThiH, as the ThiH protein is required in anaerobic organisms to form thiamine, an essential cofactor for many reactions[98,99].

One possible biochemical function of the C-terminal cysteines unique to HydG is that they take part in the binding of a second Fe-S cluster – which is supported by a report that the protein could bind ~6Fe/protein upon reconstitution[97]. The activity of this cluster could be involved with either the degradation of dehydroglycine into CO and CN^-, or alternatively may serve as a chemical platform upon which these ligands are deposited, which would indicate that the HydG protein is operative as a scaffold. Of course, a combination of these two roles could also be possible. In terms of the mechanism of the reaction catalyzed by ThiH and HydG, mutational studies that have been carried out on the ThiH protein may provide insight as to the residues involved in the formation of *p*-cresol and dehydroglycine by the two[100].

To begin probing the activity of this CxxC motif, the gene from *C.a.* in which the first of these cysteines (C386) was mutated to a serine was cloned into the pCDFDuet vector and expressed in *E.coli*, with purification carried out as in the cases of the other Hyd maturation proteins. Despite lacking this residue – which could be involved in binding an Fe-S cluster – upon reconstitution with iron and sulfide, the protein was observed to bind 9.1 ± 0.1 Fe/protein (in comparison to 8.2 ± 0.1 for the native protein in similar conditions). In addition, UV/Vis spectroscopy showed very similar spectra (Figure 5.2). These results indicate that either this cysteine does not take part in the

binding of an Fe-S cluster, or alternatively that a cluster is still able to be formed, albeit with a different or even no ligand coordinating one of the iron atoms. In addition, it should be noted that the chemical reconstitution employed to generate this form of the HydG protein could feasibly result in the binding of a cluster that would not exist under physiological conditions. It is of interest that these spectra, which show that the protein concentration of the mutant protein is less than that of the non-mutant (as judging by A_{280}) shows a larger absorbance peak at the FeS cluster corresponding wavelength at 420nm. The reason for this observation is currently unknown, but given the similarity in iron number for the two proteins, this difference could be attributed to a difference in cluster coordination environment between the two. Presumably, given the results of King *et.al* [54] this mutation alters some mode of the biochemistry undertaken by the enzyme, although the current data is insufficient yield insight as to what this constitutes. Future studies investigating the ability of the mutant protein to form the reaction products will be valuable in discerning the functionality of this residue. Specifically, experiments probing the reaction stoichiometry between the reaction products dAdo, *p*-cresol, and dehyroglycine would be informative as to the mechanistic consequences of the C-386 mutation of the HydG protein.

Detection of Cyanide Production in HydG Assays

The formation of *p*-cresol by HydG was previously reported by Pilet *et al.* [94] and may serve as indication that the HydG enzyme also produces dehydroglycine, which could be further reacted to carbon monoxide and cyanide in a second enzymatic step. In an effort to explore this possibility and define the reaction products of the HydG enzyme

and its role in H-cluster assembly, assays of HydG were prepared containing the enzyme, the substrates tyrosine and SAM, and reductant in the form of dithionite.

These assays revealed that the enzyme catalyzed the cleavage of tyrosine into *p*-cresol and dehydroglycine as detected from the hydrolysis product glyoxylate. In addition, upon derivatization of assay mixtures for detection of any possible cyanide in these mixtures, cyanide was detected in near equivalent amounts to *p*-cresol (Figure 5.3). This result indicates that the enzyme cleaves tyrosine to form the reaction products *p*-cresol and dehydroglycine, and that the dehydroglycine is likely reacted further to yield cyanide. In experiments where tyrosine was replaced by a ^{13}C and ^{15}N labeled variant, a

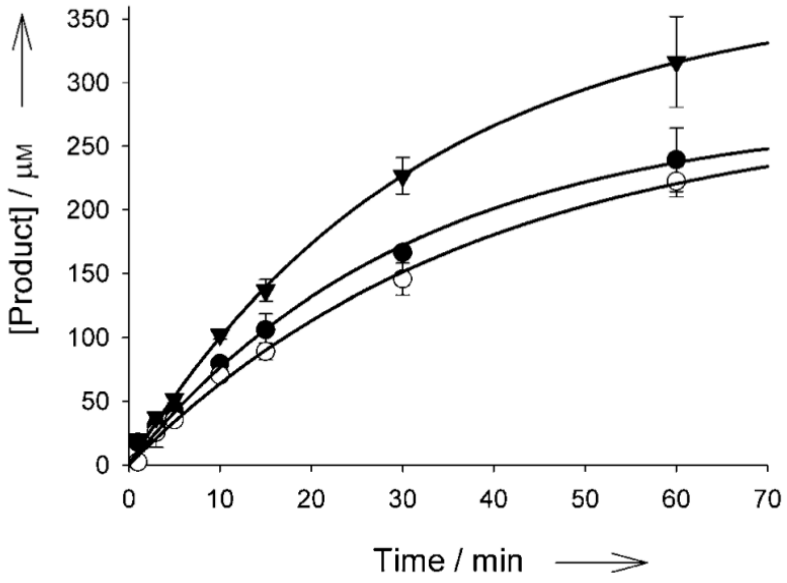

Figure 4-3. Time course of DOA (▼), p-cresol (●), and cyanide (○) formation from a HydG assay containing HydG (62 mm), tyrosine (1 mm), AdoMet (1 mm), and sodium dithionite (1 mm).

mass shift of 2 amu was observed of the derivatized cyanide reaction product, demonstrating that tyrosine is the species which serves as the precursor to cyanide. Importantly, these assays show that the molar ratio of *p*-cresol and cyanide are nearly identical, and track closely to the observed formation of dAdo.

In these assays, dehydroglycine (as the hydrolysis product glyoxylate) was detected at substoichiometric amounts compared to *p*-cresol. This, compared with the observation that in the case of the related ThiH protein equivalent amounts of these products are observed motivates the hypothesis that dehydroglycine serves as the intermediate in the formation of cyanide from tyrosine (Figure 5.4). Also notable in these experiments is the production of greater amounts of dAdo in assays than of both *p*-cresol and CN^-, which could be interpreted to mean that there is an amount of uncoupled turnover occurring in the enzyme where the formation of the dAdo radical does not result in the bond cleavage event that is responsible for generating *p*-cresol and cyanide. A similar type of activity has been observed in the case of the ThiH enzyme, and thus these results are further

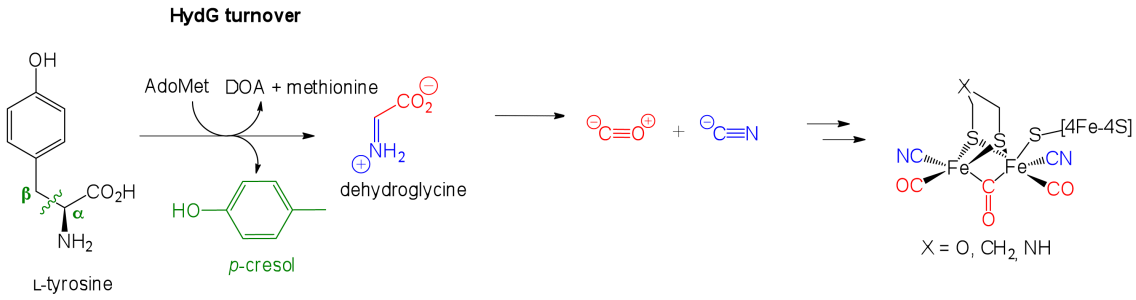

Figure 4-4. Schematic representation of the reaction carried out by HydG, showing the production of CN^- and CO from dehydroglycine and incorporation of these into the H-cluster active site.

indication that the two enzymes are operating in similar ways.

Detection of Carbon Monoxide Production in HydG Assays

The reaction scheme shown in Figure 5.4 suggests that with the generation of CN^- by the HydG enzyme, CO could also be produced concomitantly. The generation of both of the diatomic ligands by the enzyme would be sufficient for accounting for these on the

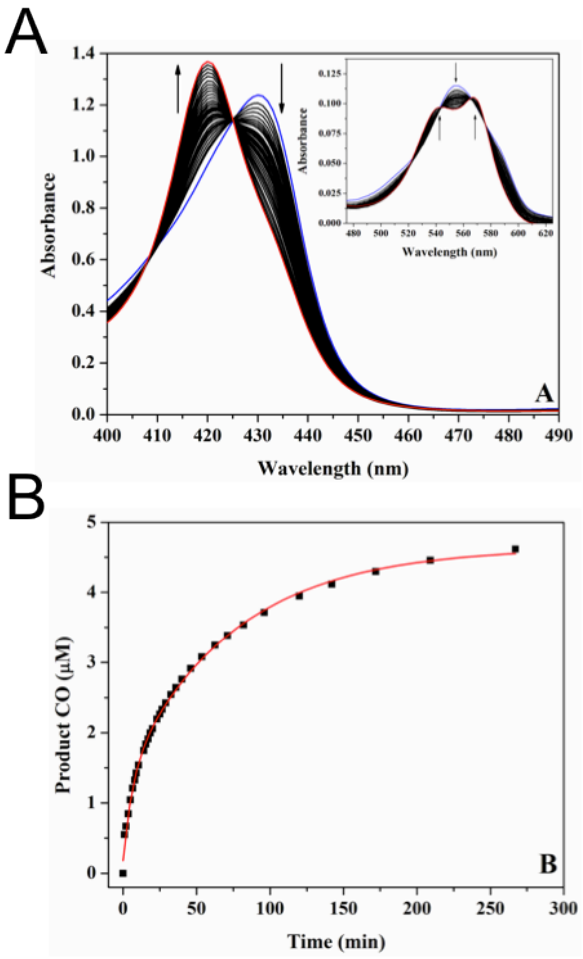

Figure 4-5. HydG-dependent CO formation. A. Time-dependent conversion of deoxy-Hb to Hb-CO as monitored by changes in the Soret band and the visible band (inset). B. CO formation as a function of time, as calculated using the data in panel A and as described in the materials and methods. Reaction contained 35 μM HydG, 4 mM dithionite, 1 mM Tyr, 1.1 mM SAM, 9.2 μM Hb.

active site H-cluster, and to investigate this possibility, assays in which HydG was combined with tyrosine, SAM, and dithionite as a reductant were performed in the presence of hemoglobin, which would exhibit a spectroscopic shift in the Soret band upon binding of any CO produced by the HydG enzyme[101].

The results of these assays reveal rapid formation of Hb-CO, dependent on the presence of HydG, tyrosine, and AdoMet (Figure 5.5). The formation of Hb-CO is evidenced by the shift in the Soret band (430 nm to 419 nm) and by the shift and splitting of the 555 nm visible band, which together give rise to a visible spectrum characteristic of Hb-CO. The data was fit to a biphasic exponential, with a burst phase corresponding to a k_{cat}= 3.5 (± 0.5) x 10^{-4} s^{-1} and a slow phase with k_{cat} = 3.0 (±0.2) x 10^{-5} s^{-1}. The fast rate is consistent with the rates observed above for above cyanide formation[97] and the rate of Hb-CO formation is dependent on the concentration of HydG present, as well as its iron content.

Discussion

The reaction by which HydG makes usage of the dehydroglycine to form both diatomic ligands of the H-cluster and poses questions as to the ways these ligands may be transferred to an iron atom which eventually becomes incorporated into the active site. Indeed, the apparent coordination of a second [4Fe-4S] cluster on the HydG protein makes possible a number of scenarios. This cluster itself could form the basis for CO and CN^- coordination, which would result in the HydG protein acting as a scaffold.

Alternatively, this second cluster may be involved in an electron transfer that may be involved in the formation of the diatomic ligands from CO and CN^-.

A major implication of the observation of HydG acting to produce CO and CN^- ligands is that HydE may be involved in the formation of the H-cluster dithiolate linkage. Sequence relationship of the HydE protein to the LipA and BioB proteins suggest an activity related to the sulfur insertion activities of these proteins, an event that in the case of H-cluster biosynthesis may involve not a sacrificial use of an FeS cluster to add sulfur to a carbon atom, but rather the alkylation of a FeS cluster to produce the thiolate as in Figure 1.3B.

The functional assignment of the HydG protein in H-cluster biosynthesis makes possible new lines of research that will further understanding of the process of cluster assembly. One outstanding question is whether the HydG protein itself serves as a scaffold for the CO or CN^- ligands. One possible route to testing this would be to conduct assays as described above, then separate the protein from the small molecules of the assay mixture by gel filtration, and analyze the protein to see if CN^- and CO could be detected. Another possibility along these lines would be to incubate the HydG protein with the HydF protein in known CN^- and CO producing conditions and after this, separate the two proteins and investigate whether HydF contained the diatomic ligands bound to Fe. In these experiments, the utilization of FTIR in conjunction with ^{13}C and ^{15}N tyrosine labeling experiments would be helpful in full assignment of the source and transfer of these ligands to the HydF protein.

The presumable production of one CO and one CN^- from a single molecule of tyrosine poses a question as to how the H-cluster active site, which contains three CO molecules and two CN^- molecules acquires the difference in the number of CO and CN^- ligands. Future work where the stoichiometry of CO and CN^- production from these enzymes will be of value in this vein, as it may be that the enzyme does not produce these two molecules in an equal amount, which could explain how the H-cluster stoichiometry occurs. Experiments with the $HydG^{C386S}$ protein will also be of value towards this goal, as it may be that this ligand disrupts the formation of either or both of the diatomic species produced by the enzyme.

From a microbiological standpoint, a possible route to investigate the function of the $HydG^{C386S}$ mutant protein is in conducting rescue experiments using a *E.coli* strain in which ThiH has been inactivated by mutation to investigate if the mutant HydG protein would be capable of replacing the activity of the ThiH protein. Along these lines, performing a research project such as that conducted by Downs and co-workers[100] where mutational analysis of the ThiH protein in *Salmonella enterica* in a ThiH *E.coli* strain would be valuable for discerning which residues of the protein are critical for function and which ones in particular are unique outside of those presumably shared with the ThiH protein to facilitate SAM and tyrosine binding.

CHAPTER 5

IDENTIFICATION AND CHARACTERIZATION OF A NOVEL MEMBER OF THE RADICAL SAM ENZYME SUPERFAMILY AND IMPLICATIONS FOR THE BIOSYNTHESIS OF THE HMD-HYDROGENASE ACTIVE SITE COFACTOR

Chapter Introduction

The HmdA enzyme is found in hydrogenotrophic methanogens where it functions in the reversible reduction of methenyl-tetrahydromethanopterin (H_4MPT^+) to methylene-H_4MPT and H^+, an intermediary step in CO_2 reduction to methane using H_2 as electron donor[24] (Figure 1.5). Biochemical characterization of the enzyme has revealed the presence of a unique active site metal cofactor which consists of an Fe ion ligated by two CO molecules, a cysteine side chain, a guanylyl pyridinol cofactor (GP-cofactor), and an unknown ligand suggested from crystallographic data to be an acyl group[26] (Figure. 1.1). The active site features and activity of the HmdA protein, namely the CO coordinated Fe and the ability to evolve H_2, bear similarity to the [NiFe]- and [FeFe]-hydrogenases. While HmdA does not partake in the reversible reaction between protons and electrons to form molecular hydrogen, as is the case with the [NiFe]- and [FeFe]-hydrogenases, the Hmd protein is often referred to as a hydrogenase.

Phylogenetic analysis of the three classes of hydrogenase suggest that they have independent evolutionary origins (see introduction). [FeFe]-hydrogenases are present only in anaerobic bacteria and a few anaerobic eukaryotes, but have yet to be identified in Archaea. In contrast, [NiFe]-hydrogenases are present in a diversity of bacteria and

---

This section contains text authored by McGlynn SE, Boyd ES, Shepard EM, Lange RK, Gerlach R, Broderick JB, and Peters JW. J. Bacteriology 2010 Jan;192(2):595-598.

archaea, but have yet to be identified in eukarya. Finally, the HmdA class of hydrogenase is found only in hydrogenotrophic methanogens and thus is most restricted in its distribution amongst species. That these divergent enzymes are unified by similar iron coordination environments suggests that the unique active sites exist as examples of convergent evolution, and points to an evolutionary pressure to develop the capability to generate proficient catalysts for hydrogen activation. Given the evolutionary un-relatedness of the different hydrogenases, it is likely that all three of these contain active sites whose biosynthesis is not related. Indeed, this is the case for the [NiFe]- and [FeFe]-hydrogenases, and may also be true in the case of the Hmd-hydrogenases. Alternatively, the Hmd-hydrogenase biosynthetic apparatus may exist as having recruited functionality from either the [FeFe]- or [NiFe]-hydrogenase maturation systems.

The genes involved in active site cluster biosynthesis in [FeFe]- and [NiFe]-hydrogenases are often co-localized with structural genes on the chromosome (see introduction). Thus, as a first step towards identifying proteins that may be involved in active site cluster biosynthesis in HmdA hydrogenase, all available genome sequences were screened for the presence of HmdA and the genes flanking *hmdA* were noted[102]. Independently, available genome sequences were screened for the presence of proteins homologous to those thought to be involved in the CO diatomic ligand formation in [FeFe]-hydrogenase, namely the radical SAM proteins HydE and HydG. These two independent approaches led to the identification of a conserved protein-encoding gene that was co-localized with *hmdA*. This gene product, referred to herein as HmdB, exhibits sequence homology with members of the radical SAM superfamily of enzymes,

but differs in that it contains a unique CX_5CX_2C motif that deviates from the CX_3CX_2C, and CX_4CX_2C[103,104] motifs characteristic of all other known members of the superfamily[93]. A second open reading frame which is related to a eukaryotic methytransferase, designated here as HmdC, was also co-localized with *hmdA*. To confirm that the identified radical SAM homolog exhibited properties associated with the enzyme family, HmdB was cloned, expressed, purified and characterized enzymatically and spectroscopically. The biochemical characterization of HmdB and its genetic association with a methyltransferase allows for the discussion of a plausible role for radical SAM chemistry and methyltransferase activity in the assembly of the Hmd active site cofactor.

Materials and Methods

Homologs of the radical SAM enzymes HydE and HydG were identified in the genomes of methanogenic archaea using the Department of Energy-Integrated Microbial Genomes (DOE-IMG) BLASTp server using sequences from *Chlamydomonas reinhardtii* as queries. The genetic context of putative radical SAM sequences was examined using the gene neighborhood viewer on the DOE-IMG server. Independent of the above analyses, regions flanking *hmdA* in the chromosomes of methanogenic archaea were examined for the presence of conserved genes. Both approaches independently identified a putative radical SAM protein annotated as 'BioB', herein referred to as HmdB, in close proximity (0-2 ORFs) to *hmdA*. In addition, a second protein-encoding gene annotated as a 'eukaryotic fibrillarin', referred to herein as *hmdC*, was found in

close proximity to *hmdA* and *hmdB*. HmdABC homologs were compiled and were individually aligned with ClustalX (version 2.0.8)[105] employing the Gonnet protein weight matrix and default alignment parameters. The best-fit protein evolutionary models for each of the three proteins was determined with ProtTest[106], with the WAG+I+G model selected for HmdBC sequences and the CpREV+G+F selected for HmdA sequences. The phylogenetic position of HmdA, B, and C was evaluated using PhyML (version 3.0)[107]with default settings and either the WAG (HmdBC) or the CpREV (HmdA) evolutionary model with a proportion of variable sites (HmdBC) or fixed amino acid frequencies (HmdA) and gamma-distributed rate variation across sites (HmdABC). The outgroup used for reconstructing the phylogeny of HmdB was a putative biotin synthase from *Desulfitobacterium hafniense* Y51 (YP_520064) whereas the outgroup used for reconstructing the phylogeny of HmdA was the uncharacterized HmdA2 protein from *Methanococcus maripaludis* S2 (CAF31272). The outgroup for reconstructing the phylogeny of HmdC was the ribosomal protein L11 methyltransferase from Finegoldia magna ATCC 29328. A consensus phylogram for each protein was projected from 100 trees using FigTree (version 1.2.2).

Genes and organisms utilized in phylogenetic analysis of HmdA enzymes and homologs are listed below, in addition to accession numbers of HmdB and HmdC gene accession numbers.

Table 5.1. Genes used in this study.

		Hmd Gene Accession Numbers				
Numeral[a]	Taxon	HmdAI	HmdAII	HmdAIII	HmdB	HmdC
1	*Methanobacterium thermoautotrophicum* ΔH	AAB85631	AAB85010, AAB85987		AAB85632	AAB85626
2	*Methanococcus voltae* A3	ZP_02193338			ZP_02193350	ZP_02193361
3	*Methanocaldococcus jannaschii* DSM 2661	NP_247770	NP_247700, NP_248340		CAF29682	CAF29681
4	*Methanococcus maripaludis strain S2*	CAF29683			NP_247771	NP_247774
5	*Methanocorpusculum labreanum* Z	YP_001031172			YP_001031166	YP_001031173
6	*Methanococcus maripaludis* C5	YP_001098062	YP_001098201		YP_001098063	YP_001098064
7	*Methanobrevibacter smithii* ATCC 35061	YP_001273145			YP_001273146	YP_001273144
8	*Methanococcus vannielii* SB	YP_001323644	YP_001323533		YP_001323643	YP_001323642
9	*Methanococcus aeolicus* Nankai-3	YP_001325215	YP_001324889		YP_001325216	YP_001325218
10	*Methanococcus maripaludis* C7	YP_001330343	YP_001330209		YP_001330342	YP_001330341
11	*Methanococcus maripaludis* C6	YP_001548874	YP_001549002		YP_001548875	YP_001548876
12	*Methanopyrus kandleri* AV19	AAM01230		AAM02335, AAM02578	AAM01233	AAM01234
NA	*Methanospirillum M.hungatei* JF-1					YP_501722

Representative protein sequences from several members of the radical SAM superfamily with known biochemical functions (HydE, HydG, PflA, BchE, BchQ, BchR, NirJ, LAM, NifB, MoaA, LipA, ThiH, BioB) were compiled using query sequences outlined in Sofia *et al*.,[93] with the DOE-IMG BLASTp server. Individual sets of compiled sequences were aligned as described above and subjected to phylogenetic reconstruction using PhyML. Five to ten sequences representing the primary lineages for each locus were identified and compiled. Compiled radical SAM sequences were aligned as described above. Due to the modular nature of the various members of the radical SAM superfamily of proteins, it was often necessary to trim portions of sequence in order

to achieve a good alignment. Thus, the final alignment contains few gapped characters. The final alignment can be attained through correspondence with the authors. The best-fit protein evolutionary model was determined with ProtTest. The phylogenetic position of HmdB in relation to representative members of the Radical SAM superfamily of proteins was evaluated using a Bayesian approach to likelihood (MrBayes version 3.1.2)[108,109] with the default settings and the WAG model with gamma-distributed rate variation across sites (I+G in evolutionary model). A total of $2e^7$ generations were calculated for the data set and a phylogram with mean branch lengths was obtained with a parameter burn-in of $1.25e^7$ (standard deviation of split trees was <0.05). Phylograms were drawn from 750 trees sampled at stationary with FigTree.

To investigate whether HmdB was capable of catalyzing chemistry characteristic of the radical SAM superfamily, the gene was cloned and the protein heterologously expressed. Two hundred microliters of a log-phase culture of *Methanococcus maripaludis* str. S2 was pelleted by centrifugation (10,000x*g*, 10 min., 4ºC) and was re-suspended in ~ 10 µl of media. One microliter of cell-suspension was added to 22.5 µL water containing 1.0 µM of each forward and reverse primer. Primers used in PCR were HmdF (5'-AAAAAAAAAGAATTCGATATTTTCAAAAATAAGAGGCAATT TC -3') and

HmdR (5'-AAAAAAAAAAAGCTTTTAATTTAGT TTGGGGTTAATCCAAGTTC -3').

Primers were designed to include an N-terminal 6x-histidine tag, an EcoRI restriction site, and a C-terminal HindIII site to facilitate cloning into MCSI of the Novagen Duet™

vector series. PCR was run with an initial heat denaturation at 94°C (4 min) followed by 30 cycles of PCR consisting of denaturation at 94°C (1 min), annealing at 58°C (1 min), and elongation at 72°C (2min) with a final extension step of 5 min at 72°C. The *hmdB* PCR product was purified using the Qiaquick PCR kit (Qiagen), digested with EcoRI and HindIII, and ligated into a previously digested Duet vector (Novagen). The integrity of the construct was confirmed by DNA sequencing. Expression, cell lysis, and purification of the protein from *E.coli* was performed as previously described above for the cases of the Hyd maturation proteins using a Ni^{2+} gravity flow column (GE Healthcare). The protein was further purified to remove imidazole by gel filtration using a PD-10 desalting column (GE Healthcare) with the exchange buffer consisting of 50 mM Hepes and 500 mM NaCl in 5% glycerol (pH 7.4). The purified protein was concentrated with a Microcon B-15 concentrator (Millipore). Protein concentration was determined by the method of Bradford.

Iron concentrations were determined by inductively coupled plasma mass spectrometry (ICP/MS) using an Agilent 7500ce ICP-MS using He as a collision gas to reduce interferences of argon oxide ions. Iron standards were obtained from Agilent Technologies (Environmental Calibration Standard 5183-4688). All samples and standards were injected in 5% HNO_3 (Trace Metal Grade, Fisher). Aliquots of protein were mixed with nitric acid (5%) to achieve a final volume of 500 μL. Samples were heated and vortexed periodically for ~ 40 minutes, then centrifuged at 16,000xg for 30 minutes. Four hundred μL of supernatant was collected and nitric acid (5%) was added

to achieve a final volume of 5.0 mL. Samples were subsequently analyzed by inductively coupled plasma-mass spectrometry (ICP/MS).

Room temperature UV/visible absorption data were acquired for HmdB under anaerobic conditions using a Cary 6000i UV/vis/ near-IR spectrophotometer (Varian). Reduced HmdB was prepared by the addition of dithionite (2 mM final concentration) to 95 μM protein. UV/visible spectra were collected at a data interval of 0.5 nm and a scan rate of 60 nm/min.

Low temperature EPR spectra were collected using a Bruker ESP300E X-band spectrometer equipped with a liquid helium cryostat and temperature controller (Oxford Instruments). The protein concentration was 395 μM. EPR parameters were: sample temperature, 12 K; microwave frequency, 9.24 GHz; modulation amplitude, 10 G; time constant, 0.25 seconds; microwave power, 2 mW. Reduced samples were prepared by the addition of dithionite to a final concentration of 1 mM from a 0.1 M solution in an anaerobic chamber. For experiments where SAM was added to EPR samples, SAM was added just after dithionite from a 0.1 M solution. Protein concentrations between samples were kept constant by adding buffer in cases where either SAM or dithionite were not added. Samples were frozen in liquid nitrogen within minutes after addition of dithionite/SAM.

Enzymatic assays were performed in sealed and anaerobic vials containing purified HmdB protein (20 μM), SAM (2 mM), DT (1 mM), in 50 mM Hepes and 500 mM NaCl in 5% glycerol (pH 7.4) in a total reaction volume of 1.0 mL. At time points of 10, 20, 40, 60, and 90 minutes, triplicate aliquots (150 μL) were collected using a

syringe and aliquots were immediately boiled for 2 minutes to inhibit further enzymatic activity. Boiled solutions were passed through 10,000 kDa molecular weight cut off filters (Millipore) and the flow through was frozen at -20°C until utilized in further analysis.

Frozen aliquots were thawed and analyzed via liquid chromatography (HPLC) using an Agilent 1100 series HPLC equipped with a C-18 Spherisorb column (160x2.5mm). The reaction mixtures were injected (10 μl) onto the column and were eluted using 0.05% trifluoroacetic acid (TFA) in water at a flow rate of 1 mL per minute for a total duration of 8 min, followed by a 35 minute linear gradient to 95% acetonitrile with 0.05% TFA while maintaining a flow rate of 1.0 mL per minute. The elution times of SAM and 5'dAdo were determined by comparison to authentic standards. Finally, mass spectrometry was used to confirm the identity of the components in each elution profile of interest. The concentration of dAdo formed as a function of time was quantified by comparison to a standard curve relating integrated peak area and 5'dAdo concentration.

Results

Distribution of Hmd in Genome Sequences

A total of 24 *hmdA* genes were identified in the 70 archaeal and 1790 bacterial genomes available to date (July 2009). The 24 sequences were identified in the genomes of twelve Euryarchaea, specifically in the orders *Methanopyrales*, *Methanobacteriales*, and *Methanococcales*. Homologs of HmdA were not identified (BLASTp cutoff of e^{-5})

in any of the available bacterial genomes, nor were they identified in the genomes of archaea including the more recently evolved acetoclastic methanogens that includes the orders *Methanomicrobiales* and *Methanosarcinales*. Multiple copies of *hmdA* exhibiting significant sequence homology were identified in nine of the twelve Euryarchaeal genomes containing *hmdA,* suggesting that they derive from gene duplication. The multiple copies of *hmdA* provide a root for each other in an unrooted phylogram since these protein-encoding genes resulted from gene duplication[110]. The 24 HmdA proteins resolved into three distinct lineages, herein referred to HmdAI, HmdAII, and HmdAIII. HmdAI sequences from *Methanopyrus kandleri* str. AV19 branch early with respect to

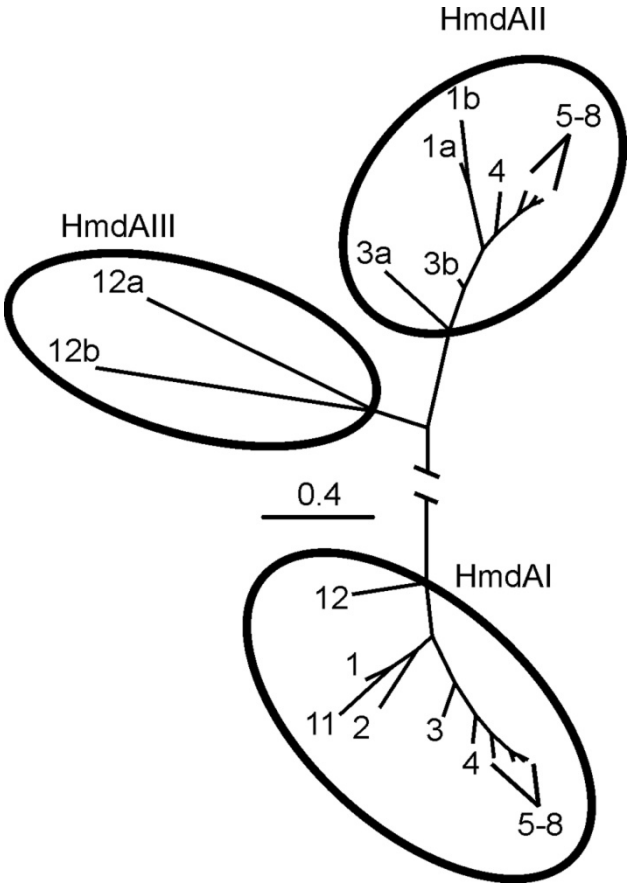

Figure 5-1 Phylogenetic relationship of identified Hmd-hydrogenase paralogs. Organisms corresponding to numbers can be found in Table 4.1.

other HmdAI sequences. Taxa with *hmdAI* contained a copy of *hmdAII* in 9 out of 12 taxa, with two copies of *hmdAII* in the genomes of *Methanobacterium thermoautotrophicum* str. Delta H and *Methanocaldococcus jannaschii* str. DSM 2661. HmdAII sequences from *M. jannaschii* diverged early with respect to HmdAII proteins from the other nine taxa present in this lineage. The HmdAIII lineage, which contains two constituent and divergent sequences from the genome *Methanopyrus kandleri* str. AV19, branches distinctly between the HmdAI and HmdAII lineages. Taken together, these observations suggest that HmdAI emerged in a methanogen that was most closely related to *M. kandleri*.

Hmd Genetic Context

The regions of the chromosome flanking HmdAI-III were examined for the presence of conserved ORFs that may provide insight into the biosynthesis of the Hmd cofactor. A consistent pattern of gene conservation was not identified in regions flanking HmdIII-encoding genes. Similarly, patterns of gene conservation flanking the HmdAII-encoding gene were generally not observed in the genomes of earlier branching lineages (*M. jannaschii*, *M. thermoautotrophicum, M. aeolicus*). However, two conserved genes flanking the more recently evolved HmdAII sequences derived from *Methanococcus vannielli* and *Methanococcus maripaludis* str. C5/C6/C7/S2 were identified. These genes were annotated as a 'type 12 methyltransferase' and a short 'GTP-binding protein'.

In the case of *hmdAI,* two conserved genes were observed to co-localized with the *hmdAI* gene in hydrogenotrophic methanogens. These genes were annotated as the radical SAM protein 'BioB' and the methylase protein 'eukaryotic fibrillarin', referred to

here for simplicity as HmdB and HmdC, respectively. *hmdB* was proximal to *hmdAI* in 11 of the 12 genomes and was located 5 open reading frames (ORFs) upstream in the twelfth genome (Fig 2A). *hmdC* was generally found in close association with *hmdAI* and its context in relation to *hmdAI* in the chromosome generally followed the phylogenetic position of the HmdAI protein (Figure 4.2A). For example, *hmdC* is located proximal to *hmdAB* in the genomes of the more recently evolved *Methanococcales* with the exception of *Methanococcus aeolicus* str. Nankai-3 where HmdC was separated from *hmdAB* by a single ORF. In the remaining genomes with HmdAI sequences that diverged early, *hmdC* was separated from *hmdAB* by no more than 4 ORFs. Both *hmdB* and *hmdC* were absent in all archaeal and bacterial genomes that did not contain an HmdA homolog, the exception being the genome of *Methanospirillum hungatei* str. JF-1 which contains an HmdC homolog but not HmdA or HmdB. This apparent directional evolution of *hmdABC* toward synteny with increasing evolutionary time suggests that these genes participate in a similar biological function.

Hmd Evolutionary History

To further investigate the likelihood that HmdA, B, and C proteins function in related biological processes, an evolutionary history of HmdAI, HmdB, and HmdC was constructed. HmdAI proteins formed three distinct lineages, with HmdAI from *M. kandleri* forming a well-supported early-diverging lineage, suggesting these proteins to have evolved early with respect to other HmdAI proteins, including those from the *Methanobacteriales*/*Methanomicrobiales* and *Methanococcales*. The two remaining lineages are comprised of proteins from *Methanobacteriales*/*Methanomicrobiales* and

Methanococcales. The HmdA topology is nearly mirrored in the HmdB phylogeny, with the primary exception being the sequence from *M. kandleri* which forms a basal branch in the *Methanobacteriales*/*Methanomicrobiales* lineage that is sister to HmdB from the *Methanococcales* (Figure 4.2B). Thus, the topology of the HmdC phylogram is similar to that observed in the HmdB phylogram. Similar tree topology for each of the individual

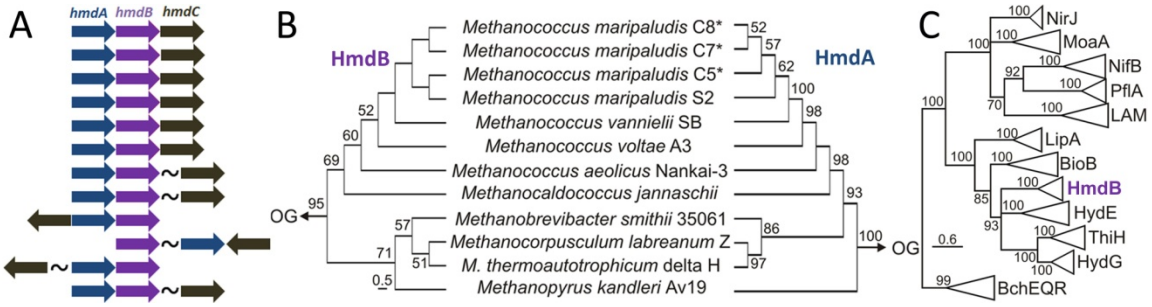

Figure 5-2 (A) Gene context of hmdA, hmdB, and hmdC. hmdA, blue; hmdB, purple; hmdC, black. (B) Chronogram of HmdB rooted with a homolog identified in the genome of *Desulfitobacterium hafniense Y51* and HmdAI rooted with HmdAII from *Methanococcus maripaludis S2*. The scale bar represents 0.5 substitutions per site. (C) Phylogram of representative members of the radical SAM superfamily of enzymes. The depth of the collapsed clades is proportional to the diversity of the individual enzyme class. The scale bar represents 0.6 substitutions per site.

HmdABC phylograms suggests that the protein-encoding genes have co-evolved.

The evolutionary relationship of HmdB to other members of the radical SAM superfamily was determined by reconstructing the evolutionary history of several members of the superfamily, including all closely related paralogs and those families that have defined biochemical function. Three distinct stem lineages were observed in the radical SAM topology, with one internally branching lineage comprised of proteins involved in the biosynthesis of bacteriochlorophyll in anoxygenic phototrophs[111]

(BchEQR) (Figure 4.2C). Interestingly, proteins that are thought to be involved in sulfur insertion (LipA, BioB) and the generation of unique diatomic ligands (HydE, HydG)[53,112,63] formed a distinct and well-supported stem lineage. HmdB along with HydE and HydG formed a distinct and well-supported sublineage within this stem lineage, suggesting that these proteins may perform similar reactions. Importantly, the Bayesian phylogram generated in the present study is generally consistent with the hierarchical clustering tree generated previously by Sofia et al., 2001[93], with the exception of NifB proteins which branch early in the previously generated cluster tree.

Identification of a Unique CX_5CX_2C Motif

HmdB sequences lack the $CX_{3-4}CX_2C$ motif that defines the radical SAM superfamily of enzymes. Rather, sequence alignment of HmdB reveals a high degree of amino acid conservation in the vicinity of a novel CX_5CX_2C motif (Figure 4.3). This motif is reminiscent of the conserved CX_3CX_2C motif characteristic of the radical SAM superfamily of enzymes. This particular motif is restricted to hydrogenotrophic

Figure 5-3 Sequence alignment of HmdB proteins showing the conserved presence of a unique $C_{XXXXX}C_{XX}C$ motif.

methanogens and a number of uncharacterized proteins identified in the genomes of *Desulfitobacterium hafniense* Y51 (YP_520064), *Hyphomicrobium denitrificans* ATCC 51888, *Syntrophomonas wolfei subsp. wolfei* str. Goettingen, *Alkaliphilus oremlandii* OhILAs, *Treponema denticola* ATCC 35405, and *Gemmatimonas aurantiaca T-27*. Phylogenetic analysis of these proteins reveals that two major clades are formed by the methanogens, one of which comprises the sequences analyzed in the present study, and another of currently unknown biochemical and genomic attribute (Figure 4.4). Branching early in this representation are sequences from *Desulfitobacterium hafniense*, *Syntrophomonas wolfei*, and *Hyphomicrobium denitrificans*, members of the *Clostrdia* and *alpha-proteobacteria* respectively. Interestingly, the *Alkaliphilus oremlandii* OhILAs, *Treponema denticola* and *Gemmatimonas aurantiaca T-27* sequences appear to bridge the two methanogenic clades (Figure 4.4), perhaps indicating that a duplication of one of these genes resulted in the two clades found in the methanogens. These sequences, along with the other methanogenic sequences may exist as potential targets for future biochemical and genomic investigations of the distribution and functionality of the CX_5CX_2C motif.

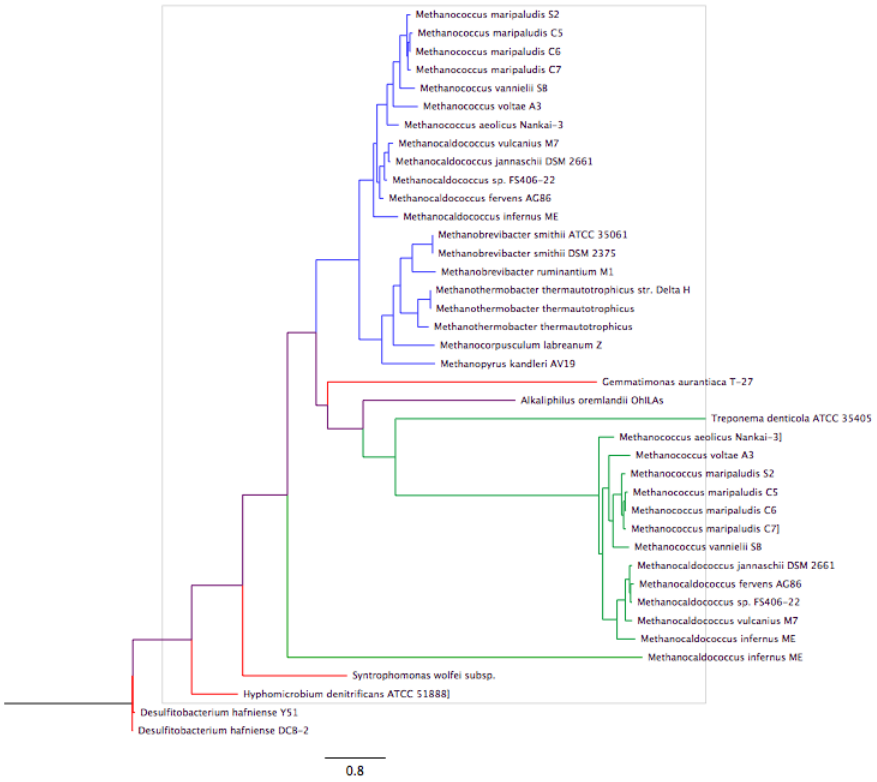

Figure 5-4 Unrooted phylogram of CX_5CX_2C motif containing sequences obtained through PHI-blast of the motif with the *M.m.* (NP_987246.1) sequence as a query. In blue, the sequences appearing in close proximity to the Hmd hydrogenase on genomes. In green, a related family also found in methanogens. In red, bacterial radical-SAM sequences with the CX_5CX_2C motif. The tree was constructed with the PhyML webserver using the maximum likelihood method and LG substitution matrix and SH-like branch support.

The presence of these putative radical SAM proteins with their unique CX_5CX_2C motifs poses intriguing questions as to their biochemical attributes and functions within these organisms. The phylogenetic clustering of HmdB sequences with the radical SAM proteins HydE and HydG, which are known to be involved in [FeFe]-hydrogenase maturation, suggests that the proteins may be involved in catalyzing similar chemistry.

Purification of the HmdB Enzyme

To investigate the chemical propensity of HmdB to conduct chemistry that would warrant its inclusion in the radical SAM superfamily, the protein was cloned, expressed heterologously, and purified. When the lysate of *E.coli* expressing HmdB was passed

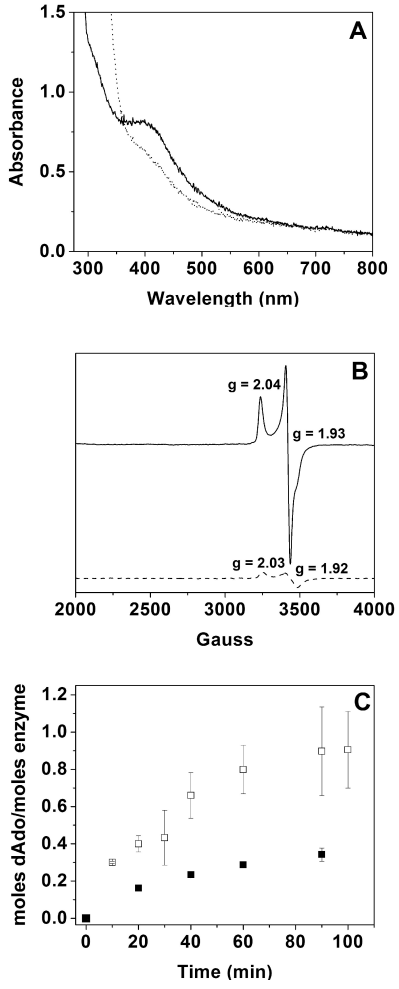

Figure 5-5 (A) UV-visible spectra of HmdB obtained in the presence (dotted) or absence (solid) of dithionite; the protein concentration was 95 μM. (B) EPR spectra of dithionite-reduced HmdB in the presence (dashed) or absence (solid) of SAM; the protein concentration was 395 μM. (C) Production of dAdo by HmdB (20 μM) in the presence of 2 mM SAM and 1 mM DT. Time points correspond to assays performed in triplicate which contained (open squares) or lacked (closed squares) 5 mM DTT.

through a Ni^{2+} column, the column turned dark brown, an observation which persisted after copious washing with increasing amounts of imidazole-containing buffer up to 50 mM. At 50 mM imidazole, the brown protein eluted to purity. SDS-PAGE indicated that the eluted protein was homogeneous. During gel filtration chromatography, HmdB migrated in a manner consistent with it being a monomer.

HmdB binds an Iron-Sulfur Cluster

Using ICP-MS, the as-isolated protein was observed to bind 2.2 ± 0.3 iron atoms. UV-visible spectroscopy revealed a peak centered at ~410 nm, a feature indicative of the presence of an iron-sulfur cluster (Figure 4.5A). Upon addition of dithionite, the peak at ~410 nm exhibited a marked decrease in intensity, presumably due to reduction of the cluster from a $[4Fe\text{-}4S]^{2+}$ to a $[4Fe\text{-}4S]^{1+}$ state. EPR of the as-isolated protein revealed a small isotropic signal characteristic of $[3Fe\text{-}4S]^{1+}$ clusters, although spin integration versus a copper standard indicated this signal was attributed to less than 1% of bound iron (data not shown). Upon addition of dithionite, the isotropic signal disappeared and a strong axial signal corresponding to a S=1/2 $[4Fe\text{-}4S]^{1+}$ cluster appeared which accounted for approximately 31% of bound iron as calculated from spin integration (Figure 4.5B solid line). Power and temperature dependence indicated the presence of a mixture of [4Fe-4S] and [2Fe-2S] clusters bound by the protein (Figure 4.6A). To investigate possible interaction of the cluster with SAM, SAM (8 mM final concentration) was added to the reduced protein immediately prior to freezing it for EPR. The resulting EPR spectrum revealed a markedly decreased signal intensity

(approximately 5% of total iron), as well as slightly altered g values (Figure 4.5B dashed line). The decrease in signal, which is similar to that observed previously when the radical SAM enzyme spore photoproduct lyase was incubated in the presence of radical SAM[113,114], could be the result of reductive cleavage of SAM (which would be accompanied by cluster oxidation); alternatively, a change in spin state of the $[4Fe\text{-}4S]^{1+}$ cluster could also give rise to a decrease in the S=1/2 EPR signal. In addition to the decrease in signal intensity, the change in g-values observed with the addition of SAM may be indicative of interaction with the cluster. From these studies, it is clear that HmdB binds an iron-sulfur cluster(s) and the cluster(s) likely interacts with radical SAM.

Studies where iron-sulfide reconstitution of the HmdB protein was performed by incubation with 25 equivalents dithiothreitol (DTT), 6.5 equivalents $FeCl_3$, and 6.5 equivalents Na_2S showed that the protein was capable of binding 3.91± 0.4 Fe

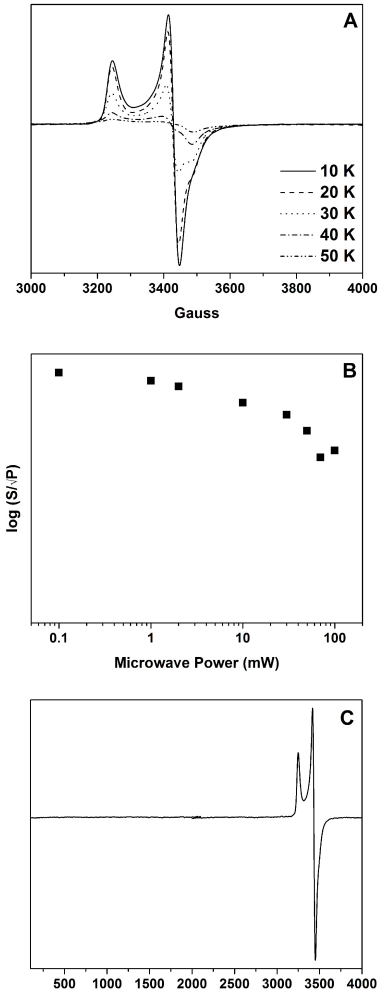

Figure 5-6 A) EPR temperature dependence of the $[4Fe4S]^+$ signal of the reduced HmdB protein. B) Power dependence of the $[4Fe4S]^+$ signal of the reduced HmdB protein. EPR measurements were performed with the same sample as used in the generation of the reduced EPR spectrum presented in Figure 4.5.

atoms/protein, and EPR/UV-vis spectroscopic properties were similar (but with greater intensity). Importantly, these observations suggests that the protein is capable of binding a single 4Fe-4S cluster. Future studies in which the conserved CX_5CX_2C cysteine residues in the protein are mutated would definitively assign these as cluster binding ligands.

HmdB Cleaves A-adenosylmethionine to 5'deoxyadenosine

HmdB was incubated in the presence of SAM under reducing conditions and the production of the reaction product dAdo was monitored as a function of time. In the presence of HmdB, the concentration of dAdo increased over time. dAdo was not observed in incubations performed in the absence of HmdB, demonstrating that dAdo production due to cleavage of SAM was catalyzed by HmdB, presumably through the formation of a 5'-dAdo radical (Figure 4.5 C). An approximate two-fold increase in reaction rate was observed in reactions containing DTT, a result coincident with the radical SAM enzymes HydE and HydG, where an increase in reaction rate has been attributed to DTT acting as a radical acceptor, shifting the reaction equilibrium[55]. Under these conditions, maximal dAdo production of ~1 mole product per mole HmdB enzyme was observed.

Summary and Discussion

Metallocluster assembly involves a number of proteins which are often co-localized with the structural gene on the chromosome (see introduction and above). The present study identified two protein-encoding genes that are co-localized on the

chromosome with the Hmd structural gene in all organisms that harbor Hmd. One protein, referred to here as HmdB, exhibited significant sequence identity to members of the radical SAM superfamily of enzymes. This protein exhibits a novel CX_5CX_2C motif rather than the characteristic $CX_{3-4}CX_2C$ motif characteristic of known members of the radical SAM superfamily. Nevertheless HmdB was shown to bind a [4Fe-4S] cluster and catalyze the reductive cleavage of SAM to dAdo, biochemical features that are consistent with the phylogenetic placement of this enzyme within the radical SAM superfamily. The second enzyme, referred to here as HmdC, exhibits sequence homology to the fibrillarin class of methyltransferase proteins present in the genomes of several eukaryotes. Collectively, these results suggest a role for HmdB and HmdC in the maturation of HmdA.

The presence of two distinct sequences containing the CX_5CX_2C motif in methanogens with sequence homology to the radical SAM protein family suggests that a gene duplication may have occurred in the methanogens to produce these two varieties. It is interesting that related sequences are also found in a small number of bacteria, and that in some of these branch between the two methanogenic clades in phylograms (Figure 4.4) These observations may mean that the sequence first appeared either in one of these bacterial taxa, or alternatively in *Methanocaldococcus infernus ME*, (which appears to be the earliest methanogen harboring these sequence) and was subsequently distributed amongst the hydrogenotrophic methanogens and bacteria. Further phylogenetic probing of this unique family of sequences would be valuable in addressing these types of questions. Within this line of inquiry, the question of why the vast majority of radical

SAM proteins utilize the CxxxCxxC FeS cluster binding motif and only a small number utilize that observed to be operative with HmdB arises (where "x" denotes any amino acid). What the function (if any) of these extra residues spaced between two of the cluster binding cysteine residues could help to address the reasons behind the broad adoption of the CxxxCxxC motif as opposed to the latter.

While the exact role of HmdB and HmdC proteins is unknown, it is tempting to speculate on the function of these proteins based on the chemistry that they are likely to catalyze as predicted by phylogenetic affiliations to known protein families. The activity of specialized protein machinery could be expected to be involved in the synthesis of the iron-carbonyl linkage in the Hmd cofactor, as well as in the synthesis of the GP-cofactor, which exists as a modified nucleoside. The fibrillarin homolog HmdC belongs to a class of enzymes involved in RNA maturation, including enzymes which catalyze nucleoside modification including the methylation of target nucleotides, using SAM as a substrate[115,116]. In addition to methylation activity using SAM as a substrate, this class of enzymes is thought to be involved in methoxycarboxylation of the tricyclic nucleoside wybutosine[117,118]. Given the structural similarity of the GP cofactor and nucleosides, it is possible that HmdC is involved in the methylation of the GP cofactor on the pyridinol ring; however, the demonstration of this activity is beyond the scope of the current study.

The biosynthesis of many active site metal clusters rely on radical SAM chemistry, including NifB in nitrogenase[74] and HydE/HydG in [FeFe]-hydrogenases[53]. The radical SAM superfamily is comprised of enzymes capable of catalyzing a myriad of reactions, all of which begin with the controlled formation of the 5'deoxyadenosyl radical

(see introduction). Some examples of this reactivity include the ability to catalyze carbon-nitrogen bond rearrangement as seen in LAM[119] and sulfur insertion as in the case of LipA[64] and BioB[112]. Of particular interest is the observation that members of the radical SAM superfamily that catalyze similar chemistry often cluster together phylogenetically (Fig. 4.2). HmdB forms a well-supported cluster with both HydE and HydG, two proteins which have been implicated in the formation and addition of CO and CN^- to the active site cluster of the [FeFe]-hydrogenase. The close phylogenetic relationship of HmdB to HydE and HydG and the fact that the active sites of both Hmd and the [FeFe]-hydrogenase contain iron-carbonyl linkages suggests that HmdB may be involved at this level of cofactor biosynthesis.

The production of dAdo in the absence of any putative substrate supports the hypothesis that the enzyme utilizes SAM as a substrate *in vivo*. Accumulation of only one equivalent of 5'dAdo indicated that while capable of cleaving SAM, HmdB does not behave catalytically in the assay conditions employed herein. The observation of non-catalytic dAdo production associated with the enzyme may be an indication that the current assays preclude the re-reduction of the cluster following one SAM cleavage event, and/or that the assays lack the physiological substrate which would provide the hydrogen to be abstracted by the generated dAdo radical. This would be consistent with other known radical SAM family members where decreased or no cleavage of SAM has been observed with the omission of substrate[113,120,121].

Recent computational analysis of the HmdA enzymes and paralogs by Goldman *et al*., 2009[122] confirms the phylogenetic relationships of HmdA homologs presented

herein. Further, the results of Goldman *et al*[122]. indicate that while the homologs differ in amino acid substitutions, the functional ability to bind the Hmd cofactor among the homologs is likely maintained, leading the authors to speculate that the homologs may function in the cell as either reservoirs for the HmdA cofactor or as scaffolds upon which assembly may occur. Such hypotheses are coincident with cofactor biosynthesis in other metallo-proteins including nitrogenase where the scaffold NifEN, which is paralogous to the structural protein NifDK, supports cofactor assembly[123] and where the NafY protein operates as an intercellular reservoir of the Mo-Fe cofactor[73].

The presence of a number of homologs of HmdA in organisms is a shared feature between Hmd and [FeFe]-hydrogenases; many organisms containing the [FeFe]-hydrogenase contain multiple copies of the structural gene *hydA*, but only single copies of the maturation proteins *hydE*, *hydF*, and *hydG*[32]. The multiple copies of *hydA* are often scattered throughout the genome, whereas *hydE, hydF, and hydG* are often co-localized, often in association with a *hydA* paralog (Boyd E.S., unpublished data). Thus, the single occurrence of both *hmdB* and *hmdC* in the genomes of hydrogenotrophic methanogens and the presence of multiple copies of *hmdA* may suggest an analogous role for *hmdA* and *hydA* homologs as cofactor reservoirs or as scaffolds where cofactors are built.

The observation of members of the radical SAM family involved in both Hmd and [FeFe]-hydrogenase biosynthesis but absence in the case of the [NiFe]-hydrogenases poses interesting questions as to the evolution and ecological distribution of the hydrogenases and their associated maturation proteins. The phylogenetic distribution of Hmd and [FeFe]-hydrogenase are similar in that they are only found in anaerobic bacteria

and in the case of [FeFe]-hydrogenase, several lower order eukarya such as the algae *C. reinhardtti*[32]. In contrast, [NiFe]-hydrogenases are common among both anaerobes and aerobes[19]. The radical SAM chemistry associated with the biosynthesis of the Hmd and [FeFe]-hydrogenase active site metalloclusters is oxygen sensitive[124] and thus it may be expected that the presence or absence of oxygen may influence the recruitment of particular enzymes that are best suited for given conditions. The partitioning of functional enzymes due to the presence of oxygen is visible in ribonucleotide synthesis[125] where ribonucleotide reductases (RNR) in anaerobes (class III RNR) utilize oxygen-sensitive radical SAM chemistry whereas RNR in aerobes (class I RNR) utilize SAM-independent and oxygen tolerant chemistry. In addition to RNR, previous studies have shown that bacteriochlorophyll biosynthesis in anaerobes utilize SAM methylating enzymes[111], whereas oxygenic phototrophs catalyze this same chemistry using non-oxygen sensitive enzymes that do not utilize radical SAM chemistry. Since oxygenic phototrophs are clearly derived from anoxygenic phototrophs[126], this observation may indicate that the ancestral bacteriochlorophyll methylation step was catalyzed by a radical SAM enzyme, only to be replaced by non-oxygen sensitive enzymatic catalysts with the ensuing rise of oxygen. Similar arguments have been made with the utilization of FeS cluster containing proteins in glycolytic pathways where differences in enzymes used in the Entner-Doudoroff pathways of Archaebacteria and Eubacteria differ in the utilzation of iron containing enzymes led to the suggestion that primitive organisms may have relied more on oxygen sensitive co-factors whereas later organism adopted a greater utilization of oxygen tolerant enzymes[127]. Similarly, the observation that the maturation

of [NiFe]-hydrogenase active sites in aerobes and anearobes alike do not utilize radical SAM chemistry may suggest that these enzymes emerged later than Hmd and/or [FeFe]-hydrogenase. Alternatively, these observations may suggest that the Hmd and [FeFe]-hydrogenase and maturation enzymes evolved in anoxic environments whereas the [NiFe]-hydrogenase and associated maturation enzymes evolved in a more oxidizing environment. Taken with these types of arguments, it must be noted that virtually all organisms utilize oxygen sensitive co-factors to some degree, thus the presence or absence of particular cofactors in organisms from different niches (oxygenic/anoxygenic) may not correspond in a definitive way to the evolutionary path in which the enzymes themselves followed.

The combination of bioinformatics and biochemical approaches presented here represents a means by which information as to protein function, distribution, and evolution may be gained. Here, this combinatorial approach led to the discovery of two genes located proximal to the *hmdA* structural gene that may be implicated in Hmd cofactor biosynthesis. The phylogenetic distribution, phylogenetic affiliation, and evolutionary history of these gene products suggest a role in the maturation of the HmdA cofactor. Heterologous expression and biochemical characterization of HmdB confirms bioinformatic predictions that this enzyme belongs to the radical SAM superfamily of enzymes. In the context of the structure of the HmdA active site cofactor, these observations collectively point to a role for HmdB and HmdC in the maturation of the HmdA cofactor. Further spectroscopic and biochemical characterization of HmdA cofactor synthesized in a background lacking HmdB and/or HmdC, in addition to further

characterization of the HmdB protein *in vitro* will continue to provide insight into the specific role of these enzymes in the maturation of HmdA. In addition, the implementation of a high throughput assay by which the HmdB enzyme substrate specificity could be identified would be a useful avenue by which to identify function of the enzyme via product analysis.

CONCLUDING REMARKS

The utilization of transition metals in biology, and the complex structures that these are made into via dedicated biosynthetic pathways represents a fascinating example of the chemical proficiency and reaction control afforded by evolving living systems. The hydrogenase enzymes, as a subset of metal cluster bearing enzymes, form a protein family from which to base inquiry on fundamental processes in biology. Due to the widespread microbial utilization of hydrogen in microbial energetics, and the necessity of utilizing protein bound transition metal cofactors to accomplish hydrogen activation, observation of the different facets of the hydrogenase enzymes yields insight into a wide range of issues and areas including and spanning the fields of bio-inorganic chemistry, molecular evolution, and origins of life research. Research into these enzymes thus has wide import.

The results presented herein, which represent the first functional assignments of two of the [FeFe]-hydrogenase active site cluster assembly proteins (HydF and HydG) marks a significant advance in the understanding of the assembly of the active site cofactor, and expand notions of the type of chemistry may be carried out by [Fe-S] cluster containing enzymes. This seminal work, in which the partitioning of biochemical functionality is accomplished provides motivation for the establishment of a number of lines of continuing research and allows for the formation of testable hypotheses which will probe the mechanistic, evolutionary, and ecological attributes of the [FeFe]-hydrogenases and their associated assembly proteins.

Starting with the development of an in vitro assay by which the HydA protein can be activated, the system as detailed in Chapter 2 provided the first notion that the three maturation proteins themselves were responsible for the assembly of a intermediate cofactor outside of the presence of the HydA protein and led to speculation that a specific protein scaffold may exist that is a surrogate host of an H-cluster precursor prior to activation. These results created the opportunity for the pursuit of two distinct but complementary lines of research from which to probe H-cluster assembly. These include i) investigation of the activating factor created by the HydE, HydF, and HydG proteins and in what way this is formed and ii) investigation of the state of the HydA protein before and after activation by the maturation proteins insofar that prior to interaction with the maturation proteins, the HydA protein exists as a stable, activatable and characterizable entity.

The characteristics of the in vitro [FeFe]-hydrogenase activation system developed herein holds promise as a platform from which further inquiry can probe the ability to accomplish H-cluster assembly in vitro from purified components. Recently, a variant of this system by which [FeFe]-hydrogenases can be activated in vitro has been utilized to demonstrate that both tyrosine and cysteine stimulate activation in analogous mixtures[128]. Together with results that show that CN^- and CO are derived from tyrosine by the action of HydG[97], this result provides motivation for the consideration of cysteine as a possible source of the dithiolate ligand of the H-cluster. Thus the synthesis of the H-cluster in vitro from purified starting components may be observed in the not to distant future as has been observed in the case of nitrogenase Mo-Fe cofactor biosynthesis[90].

Investigation of the mode of HydA activation by the maturation proteins resulted in the identification of the HydF protein as acting as a scaffold and demonstrates that the protein and the biochemistry associated with it are at the nexus of the assembly process where the protein serves to mediate the radical SAM chemistry of HydE and HydG to that of HydA to effect activation and full formation of the H-cluster. The assignment of the HydF protein as bearing CO and CN^- coordinated iron atoms represents a novel instance of this ligand combination observed in biology outside of the [NiFe]- and [FeFe]-hydrogenases themselves and with the precise coordination of this cluster currently unknown, further spectroscopic and crystallographic characterization represent enticing research prospects.

The activation characteristics of the HydF protein reported here indicate that HydF as isolated from heterologous expressions does not exist in a fully loaded form, and motivates the hypothesis that the non HydA activating HydF isoforms contain cluster intermediates resulting from interaction of the endogenous [Fe-S] cluster assembly proteins of *E.coli*, as well as the HydE and HydG proteins. This observation provides an avenue for future research aimed at isolating and characterizing these cluster intermediates, and data obtained on these will illuminate the step wise nature of cluster assembly on the HydF protein. Further characterization of the HydF protein thus provides basis for dissecting the order and timing of ligand deposition on the nascent H-cluster. The genetic and expression system described in Chapter 3 will be valuable for these studies and by being able to express and purify the HydF protein from different genetic backgrounds, different forms of the protein will be obtainable that will further

define the assembly process. For example, HydF may be purified with methods developed here from expressions where HydE but not HydG is present, thus presenting the opportunity to investigate whether HydE affects the cluster bound by HydF in the absence of HydG. These different forms of HydF will likely harbor unique intermediates, and their elucidation will further expand notions of the diversity and manner by which metal clusters are assembled in biology and how the assembly process proceeds in the case of the [FeFe]-hydrogenases.

The formation of cyanide and carbon monoxide by the HydG protein starting from tyrosine exists as an unprecedented biochemical transformation and is unique in mechanism and starting material from the case of cyanide formation in [NiFe]-hydrogenase biosynthesis[41]. It is of interest to note that cyanide is produced in organisms outside of the biosynthesis of the hydrogenases by some *Pseudomonas* species, and that in these the substrate for the reaction is glycine, which through non-radical chemistry (dehydration) is converted to dehydroglycine, which then is further reacted to form cyanide and carbon dioxide[129,130]. Given this it may be that these two enzymes share common steps following the formation of dehydroglycine that lead to cyanide production.

The involvement of radical SAM proteins in the [FeFe]-hydrogenase active site assembly pathway, and as discussed in Chapter 4 likely the Hmd hydrogenase pathway, points to the chemical versatility of this family of enzymes in accomplishing a diverse set of functionalities. This flexibility is exemplified by the case of HydG and the demonstration that it forms CN^- and CO from tryrosine in a multistep process that involves the formation of dehydroglycine followed by the degradation of this to CN^- and

CO. The HydG like enzyme ThiH catalyzes the first part of this reaction to generate dehydroglycine which is incorporated into thiazole in thiamine biosynthesis and HydG differs from ThiH in that it carries this first reaction product further by an unknown mechanism to the diatomic ligands that are found in the H-cluster active site[97,68]. Continued studies of these enzymes will further the understanding of alpha-beta bond cleavage of amino acids, and the observation that this often occurs by the utilization of amino acids with aromatic side chains[97].

Of outstanding function and mechanism are the HydE and HmdB proteins, which may act to form the non-protein dithiolate harbored by the [FeFe]-hydrogenases or part of the Hmd hydrogenase active site respectively (see Figure 1.1). With the observation that the HydG protein acts in multiple steps to result in the formation of the diatomic CO and CN^- ligands, the activity of HydE is implied to be the synthesis of the dithiolate linker. Obfuscating this question is that the precise atomic composition of the H-cluster bridging dithiolate is unknown; thus with neither the products nor the reactants of the reaction known, a reasonable hypothesis for possible substrates which the HydE enzyme may utilize is difficult to form. This problem is analogous to that arising from the identification of the HmdB radical SAM enzyme (Chapter 4). The HydE and HmdB enzymes are closely related in sequence but differ in a number of significant aspects, most notable is that the HmdB protein harbors a unique CxxxxxCxxC motif. This deviation from the CxxxCxxC "consensus" sequence for the radical SAM proteins suggests that a number of motifs may be capable of performing radical SAM associated chemistry so long as the peptide is capable of binding a site differentiated (3Cys-[4Fe-

4S]) cluster and coordinating SAM in a way amenable to formation of the dAdo radical intermediate. These proteins thus are of interest for future studies which may better define the requirements for homolysis of a sulfonium ion via electron transfer from a metal sulfur cluster.

Taken together and in conjunction with results which have indicated the presence of a pre-formed [4Fe-4S] cluster in the HydA protein for activation[81], results presented here are consistent with the hypothesis that the three maturation proteins HydE, HydF, and HydG are the specific protein requirement to achieve synthesis of the H-cluster, and that organisms which contain the four proteins together could be expected to be capable of assembling and utilizing the [FeFe]-hydrogenase as part of a metabolic circuit. Thus, outside of the iron-sulfur cluster assembly machinery ubiquitous in living systems[7], these protein factors appear to exist as unique, evolved, and specific entities which are capable of performing the chemical transformations necessary for H-cluster biosynthesis.

The observation of three separate modes by which Nature assembles and utilizes Fe-CO containing complexes suggests that these confer a high degree of positive evolutionary selection on organisms harboring them. The evolution of these entities poses questions as to the adaptation of different groups of organisms to environmental conditions and metabolic pressure. The apparent uniqueness in mechanism through which ligand arrangement is derived in the three cases supports the hypothesis that given a chemical potential from which organisms can derive energy, it will eventually be utilized and in cases where this requires a complex series of events (as in the hydrogenases), multiple systems can evolve from different genetic backgrounds and

starting points. In this light, the hydrogenase enzymes and the assembly processes associated with them are compatible with Gould and Vrba's notion[131] of evolution operating as a continual process of exaptation and reinforces the idea that evolution proceeds in part by the modification and adaptation of pre-existing functionality into new roles. This situation of microbes deriving alternative modes to accomplish similar functions is in contrast to what is largely observed in animals, where it has been posited that the emergence of new traits suited for a particular niche occurs as a very rare event and sometimes even singly[132].

One possible avenue of further study is in the evolution of the [FeFe]-hydrogenase cluster assembly system and how the three requisite proteins were recruited into the assembly process. That the HydE and HydG proteins are members of the radical-SAM family of enzymes suggests that they may have been recruited from pre-existing functionality; the nearest relationship of the two being BioB and ThiH respectively[93] (Figure 4.2C). In contrast, sequence analysis of the HydF protein reveals that the protein is comprised of two distinct domains and exists with a unique two domain sequence. Proteins belonging to this group of domain families contain a GTPase domain, in addition to a C-terminal domain. In the case of HydF, this C-terminal region contains conserved cysteine residues that may be involved in the coordination of the Fe-S cluster harbored by the protein ($CxHx_{46–53}CxxC$, where "x" denotes any amino acid)[54]. This arrangement of domains appears to be unique to HydF, and hints that the protein itself may be the result of a gene fusion between a preexisting GTPase protein and a Fe-S cluster binding protein. In addition, co-occurrence of this unique domain architecture to HydF proteins that are

present in organisms with [FeFe]-Hydrogenases may be interpreted as a strong evolutionary pressure for the presence of HydF proteins to be operative in conjunction with HydA.

Given that the HydE and HydG, proteins are closely related to other proteins that have presumably been present in genomes for a long period of time (possibly arising from BioB, and ThiH respectively) [19,32,93], these genomic observations suggest that the origination of the HydF protein may have been a determinant in the formation of [FeFe] – hydrogenases in microbial communities. This argument that the radical SAM contribution to the synthesis of the H-cluster is perhaps the oldest of the system is bolstered by the fact that both thiamine and biotin are required for all life[133,134], and that the synthesis of these requires the existence of the BioB and ThiH proteins respectively, whereas hydrogenase activity is not absolutely required for life. Thus the closely related BioB, ThiH, HydE and HydG proteins may share a common ancestor of an ancient protein family, but in contrast the HydF sequence appears to be unique to this protein and may suggest that the evolution of the HydF protein was a evolutionary bottleneck in the formation of the H-cluster in microbes. This notion is corroborated by biochemical evidence that the HydF protein serves as a molecular scaffold to activate the HydA protein, mediating the radical SAM chemistry of HydE and HydG to HydA and begs the questions of where (in what species) the event that lead to the formation of the HydF protein may have occurred.

A BLAST search with either the HydF protein sequence or the GTPase domain of HydF reveals that the domain is most closely related to the Era/EngA family of proteins.

Of still unknown biochemical function, this class of GTPases are ubiquitous in bacteria, but not found in Eukaryota or Archaea implying a bacterial origin for the domain[135], an observation that may be of importance considering the restricted distribution of [FeFe]-hydrogenases to bacteria and some lower eukarya[32] . The C-terminal domain of the protein shares little in common with other known sequences. However, the C-terminal CxxC region does harbor some similarity to the heterodisufide reducatase class of proteins found in archaea which bind a 4Fe-4S cluster and are involved in the reduction of CoM-S-S-CoB in methanogenic archea[136]. Future studies of the evolution of these two domains can be expected to further the understandings of the evolution of the H-cluster biosynthetic apparatus, and in addition shed light on the phylogenetic distribution of metal cluster assembly associated NTP-ases in biology.

Still not addressed with the HydF protein is the role of its observed GTPase activity. Many metal-cluster assembly pathways often employ nucleotide triphosphate hydrolysis in their assembly pathways. Indeed, the proteins UreG, NifH, HypB, CooC, Ind1, HscA/B, and Cdf1-Nbp35 all have requisite NTP-ase activity associated with their operations of nickel insertion[137,138], FeMo-co biosynthesis[80], and irons-sulfur cluster biogenesis[7,139] respectively. The elucidation of whether this activity serves a common function between the different cluster assembly pathways or is unique to individual systems will be a defining set of observations for the field, and future work as to the phylogenetic distribution of these NTP-ases in biology will perhaps aid in this effort.

From an applied science perspective, the possibility of utilizing [FeFe]-hydrogenases found in microbial metabolic circuits for anthropic energy production could

be furthered from the biochemical studies presented herein. Due in part to the well studied nature of *E.coli*, studies directed at elevating hydrogen yields from this organism are underway in a number of labs, where through the creation of artificial metabolic pathways in which hydrogenases are incorporated increased hydrogen production is observed in bacterial expression systems[140-143]. These studies are operative at the metabolic stage of pyruvate oxidation and aim to couple this process to hydrogen production by introducing [FeFe]-hydrogenases for hydrogen production, artificial electron transfer proteins such as ferredoxin to shuttle electrons to the hydrogenase, and by deleting genes whose products siphon electronic equivalents from electron transfer to a hydrogenase protein. By better understanding the means and determinants by which hydrogenases are formed within the cell, these studies could perhaps be furthered taking into account biochemical advances presented herein such that increased hydrogen yields would be possible. One possible avenue for this would be in modulating the expression level of the Hyd accessory proteins and as presented in Chapter 3, gene dosage of the HydF scaffold within *E.coli* can vary the level of hydrogenase activity by as much as 2 fold and thus could potentially result in a more active (hydrogen producing) recombinant *E.coli* strain if utilized in conjunction with the metabolic engineering studies referenced above.

The continued pursuit of research directed at the many facets of the hydrogenase enzymes will better the collective scientific community's understandings of biology as a whole and the underlying features of it insofar as transition metals appear to be required the operation of biology. The synthesis of the H-cluster in vitro from purified starting

components may be observed in the not to distant future as has been observed in the case of nitrogenase Mo-Fe cofactor biosynthesis[90], and the realization of this will be an achievement of pinnacle importance in the understanding of H-cluster biosynthesis. In ways, the focused study of this class of enzymes can be projected such that inquiry into multiple disciplines may be gained from the observation of this single system and this research may thus form a footing from which to base inquiry as to the mechanistic aspects of more global processes. These include the mechanisms by which biology utilizes and modifies metals to accommodate metabolic diversity, the evolution of this ability, and as briefly mentioned above, possible routes in which this understanding may be applied to issues facing the human population such as technological energy transduction. The dual relationship of the hydrogenase enzymes to the early evolution of biological systems and to future anthropogenic energetic considerations is striking, and indicates that this research contributes to an overall research paradigm directed at asking the perennial questions at the molecular level; “What are we, where have we come from, and where are we going?”

REFERENCES CITED

1. Wicken JS. A Thermodynamic Theory of Evolution. *J. Theo. Biol.* 1980;87(1):9-23.

2. Beinert H, Kiley PJ. Fe-S proteins in sensing and regulatory functions. *Curr. Opin. Chem. Biol.* 1999;3(2):152-157.

3. Beinert H, Holm RH, Munck E. Iron-sulfur clusters: nature's modular, multipurpose structures. *Science*. 1997;277(5326):653-659.

4. Taffs R, Aston JE, Brileya K, et al. In silico approaches to study mass and energy flows in microbial consortia: a syntrophic case study. *BMC Systems Biology*. 2009;3(1):114.

5. Wachtershauser G. On the chemistry and evolution of the pioneer organism. *Chemistry & Biodiversity*. 2007;4(4):584-602.

6. Ayala-Castro C, Saini A, Outten FW. Fe-S cluster assembly pathways in bacteria. *Microbiol. Mol. Biol. Rev*. 2008;72(1):110-125.

7. Lill R. Function and biogenesis of iron-sulphur proteins. *Nature*. 2009;460(7257):831-838.

8. Lill R, Kispal G. Maturation of cellular Fe-S proteins: an essential function of mitochondria. *Trends Biochem. Sci*. 2000;25(8):352-356.

9. Lill R, Mühlenhoff U. Iron-sulfur protein biogenesis in eukaryotes: components and mechanisms. *Annu. Rev. Cell Dev. Biol*. 2006;22:457-486.

10. Lill R, Mühlenhoff U. Maturation of iron-sulfur proteins in eukaryotes: mechanisms, connected processes, and diseases. *Annu. Rev. Biochem*. 2008;77:669-700.

11. Xu XM, Møller SG. Iron-sulfur cluster biogenesis systems and their crosstalk. *Chembiochem*. 2008;9(15):2355-2362.

12. Fontecave M. Iron-sulfur clusters: ever-expanding roles. *Nat Chem Biol*. 2006;2(4):171-174.

13. Outten FW. Iron-sulfur clusters as oxygen-responsive molecular switches. *Nat Chem Biol*. 2007;3(4):206-207.

14. Capozzi F, Ciurli S, Luchinat C. Coordination sphere versus protein environment as determinants of electronic and functional properties of iron-sulfur proteins. *Metal Sites in Proteins and Models Redox Centres*. 1998;(90):127–160.

15. Drennan CL, Peters JW. Surprising cofactors in metalloenzymes. *Curr. Opin. Struct.*

Biol. 2003;13(2):220-226.

16. Ragsdale SW. Life with carbon monoxide. *Crit. Rev. Biochem. Mol. Biol*. 2004;39(3):165-195.

17. Peters JW, Szilagyi RK. Exploring new frontiers of nitrogenase structure and mechanism. *Curr Opin Chem Biol*. 2006;10(2):101-108.

18. Fontecilla-Camps JC, Volbeda A, Cavazza C, Nicolet Y. Structure/function relationships of [NiFe]- and [FeFe]-hydrogenases. *Chem. Rev*. 2007;107(10):4273-4303.

19. Vignais PM, Billoud B, Meyer J. Classification and phylogeny of hydrogenases. *FEMS Microbiol. Rev*. 2001;25(4):455-501.

20. Vignais PM, Billoud B. Occurrence, classification, and biological function of hydrogenases: an overview. *Chem. Rev*. 2007;107(10):4206-4272.

21. McGlynn SE, Mulder DW, Shepard EM, Broderick JB, Peters JW. Hydrogenase cluster biosynthesis: organometallic chemistry nature's way. *Dalton Transactions*. 2009;(22):4274-4285.

22. Volbeda A, Charon MH, Piras C, et al. Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. *Nature*. 1995;373(6515):580-587.

23. Peters JW, Lanzilotta WN, Lemon BJ, Seefeldt LC. X-ray crystal structure of the Fe-only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 angstrom resolution. *Science*. 1998;282(5395):1853-1858.

24. Shima S, Thauer RK. A third type of hydrogenase catalyzing H2 activation. *Chem Rec*. 2007;7(1):37-46.

25. Hiromoto T, Warkentin E, Moll J, Ermler U, Shima S. The crystal structure of an [Fe]-hydrogenase-substrate complex reveals the framework for H2 activation. *Angew. Chem. Int. Ed. Engl*. 2009;48(35):6457-6460.

26. Hiromoto T, Ataka K, Pilak O, et al. The crystal structure of C176A mutated [Fe]-hydrogenase suggests an acyl-iron ligation in the active site iron complex. *FEBS Lett*. 2009;583(3):585-590.

27. Lyon EJ, Shima S, Boecher R, et al. Carbon monoxide as an intrinsic ligand to iron in the active site of the iron-sulfur-cluster-free hydrogenase H2-forming methylenetetrahydromethanopterin dehydrogenase as revealed by infrared spectroscopy. *J. Am. Chem. Soc*. 2004;126(43):14239-14248.

28. Pilak O, Mamat B, Vogt S, et al. The crystal structure of the apoenzyme of the iron-sulphur cluster-free hydrogenase. *J. Mol. Biol*. 2006;358(3):798-809.

29. Shima S, Lyon EJ, Sordel-Klippert M, et al. The cofactor of the iron-sulfur cluster free hydrogenase hmd: structure of the light-inactivation product. *Angew. Chem. Int. Ed. Engl*. 2004;43(19):2547-2551.

30. Shima S, Pilak O, Vogt S, et al. The crystal structure of [Fe]-hydrogenase reveals the geometry of the active site. *Science*. 2008;321(5888):572-575.

31. Stams AJ, Plugge CM. Electron transfer in syntrophic communities of anaerobic bacteria and archaea. *Nature Reviews Microbiology*. 2009;7(8):568–577.

32. Meyer J. [FeFe] hydrogenases and their evolution: a genomic perspective. *Cell. Mol. Life Sci*. 2007;64(9):1063-1084.

33. Russell MJ. The alkaline solution to the emergence of life: Energy, entropy and early evolution. *Acta Biotheoretica*. 2007;55(2):133-179.

34. Tian F, Toon OB, Pavlov AA, De Sterck H. A hydrogen-rich early Earth atmosphere. *Science*. 2005;308(5724):1014-1017.

35. Hoehler TM. Biogeochemistry of dihydrogen (H2). *Met Ions Biol Syst*. 2005;43:9-48.

36. López-Garćia P, Moreira D. Metabolic symbiosis at the origin of eukaryotes. *Trends Biochem. Sci*. 1999;24(3):88-93.

37. Martin W, Müller M. The hydrogen hypothesis for the first eukaryote. *Nature*. 1998;392(6671):37-41.

38. Böck A, King PW, Blokesch M, Posewitz MC. Maturation of hydrogenases. *Adv. Microb. Physiol*. 2006;51:1-71.

39. Posewitz MC, Mulder DW, Peters JW. New Frontiers in Hydrogenase Structure and Biosynthesis. *Current Chemical Biology*. 2008;2:178-199.

40. Leach MR, Zamble DB. Metallocenter assembly of the hydrogenase enzymes. *Curr Opin Chem Biol*. 2007;11(2):159-165.

41. Reissmann S, Hochleitner E, Wang H, et al. Taming of a poison: biosynthesis of the NiFe-hydrogenase cyanide ligands. *Science*. 2003;299(5609):1067-1070.

42. Blokesch M, Paschos A, Bauer A, et al. Analysis of the transcarbamoylation-dehydration reaction catalyzed by the hydrogenase maturation proteins HypF and HypE.

Eur. J. Biochem. 2004;271(16):3428-3436.

43. Blokesch M, Böck A. Maturation of [NiFe]-hydrogenases in *Escherichia coli*: the HypC cycle. *J. Mol. Biol*. 2002;324(2):287-296.

44. Blokesch M, Albracht SPJ, Matzanke BF, et al. The complex between hydrogenase-maturation proteins HypC and HypD is an intermediate in the supply of cyanide to the active site iron of [NiFe]-hydrogenases. *J. Mol. Biol*. 2004;344(1):155-167.

45. Roseboom W, Blokesch M, Böck A, Albracht SPJ. The biosynthetic routes for carbon monoxide and cyanide in the Ni-Fe active site of hydrogenases are different. *FEBS Lett*. 2005;579(2):469-472.

46. Paschos A, Glass RS, Böck A. Carbamoylphosphate requirement for synthesis of the active center of [NiFe]-hydrogenases. *FEBS Lett*. 2001;488(1-2):9-12.

47. Forzi L, Hellwig P, Thauer RK, Sawers RG. The CO and CN(-) ligands to the active site Fe in [NiFe]-hydrogenase of Escherichia coli have different metabolic origins. *FEBS Lett*. 2007;581(17):3317-3321.

48. Lenz O, Zebger I, Hamann J, Hildebrandt P, Friedrich B. Carbamoylphosphate serves as the source of CN(-), but not of the intrinsic CO in the active site of the regulatory [NiFe]-hydrogenase from *Ralstonia eutropha*. *FEBS Lett*. 2007;581(17):3322-3326.

49. Olson JW, Maier RJ. Dual roles of *Bradyrhizobium japonicum* nickel in protein in nickel storage and GTP-dependent Ni mobilization. *J. Bacteriol*. 2000;182(6):1702-1705.

50. Waugh R, Boxer DH. Pleiotropic hydrogenase mutants of *Escherichia coli K12*: growth in the presence of nickel can restore hydrogenase activity. *Biochimie*. 1986;68(1):157-166.

51. Böhm R, Sauter M, Böck A. Nucleotide sequence and expression of an operon in *Escherichia coli* coding for formate hydrogenlyase components. *Mol. Microbiol*. 1990;4(2):231-243.

52. Posewitz MC, King PW, Smolinski SL, et al. Discovery of two novel radical S-adenosylmethionine proteins required for the assembly of an active [Fe] hydrogenase. *J. Biol. Chem*. 2004;279(24):25711-25720.

53. Posewitz MC, King PW, Smolinski SL, et al. Identification of genes required for hydrogenase activity in *Chlamydomonas reinhardtii*. *Biochem. Soc. Trans*. 2005;33(Pt 1):102-104.

54. King PW, Posewitz MC, Ghirardi ML, Seibert M. Functional studies of [FeFe]

hydrogenase maturation in an *Escherichia coli* biosynthetic system. *J. Bacteriol*. 2006;188(6):2163-2172.

55. Rubach JK, Brazzolotto X, Gaillard J, Fontecave M. Biochemical characterization of the HydE and HydG iron-only hydrogenase maturation enzymes from *Thermatoga maritima*. *FEBS Lett*. 2005;579(22):5055-5060.

56. Brazzolotto X, Rubach JK, Gaillard J, et al. The [Fe-Fe]-hydrogenase maturation protein HydF from *Thermotoga maritima* is a GTPase with an iron-sulfur cluster. *J. Biol. Chem*. 2006;281(2):769-774.

57. Frey PA, Hegeman AD, Ruzicka FJ. The Radical SAM Superfamily. *Crit. Rev. Biochem. Mol. Biol*. 2008;43(1):63-88.

58. Wang SC, Frey PA. S-adenosylmethionine as an oxidant: the radical SAM superfamily. *Trends Biochem. Sci*. 2007;32(3):101-110.

59. Duschene KS, Veneziano SE, Silver SC, Broderick JB. Control of radical chemistry in the AdoMet radical enzymes. *Curr Opin Chem Biol*. 2009;13(1):74-83.

60. Frey PA. Importance of organic radicals in enzymic cleavage of unactivated carbon-hydrogen bonds. *Chemical Reviews*. 1990;90(7):1343-1357.

61. Booker SJ. Anaerobic functionalization of unactivated C-H bonds. *Curr Opin Chem Biol*. 2009;13(1):58-73.

62. Peters JW, Szilagyi RK, Naumov A, Douglas T. A radical solution for the biosynthesis of the H-cluster of hydrogenase. *FEBS Lett*. 2006;580(2):363-367.

63. Miller JR, Busby RW, Jordan SW, et al. Escherichia coli LipA is a lipoyl synthase: in vitro biosynthesis of lipoylated pyruvate dehydrogenase complex from octanoyl-acyl carrier protein. *Biochemistry*. 2000;39(49):15166-15178.

64. Reed KE, Cronan JE. Lipoic acid metabolism in Escherichia coli: sequencing and functional characterization of the lipA and lipB genes. *J. Bacteriol*. 1993;175(5):1325-1336.

65. Schwab DE, Tard C, Brecht E, et al. On the electronic structure of the hydrogenase H-cluster. *Chem. Commun. (Camb.)*. 2006;(35):3696-3698.

66. Wagner AF, Frey M, Neugebauer FA, Schäfer W, Knappe J. The free radical in pyruvate formate-lyase is located on glycine-734. *Proc. Natl. Acad. Sci. U.S.A*. 1992;89(3):996-1000.

67. Moss M, Frey PA. The role of S-adenosylmethionine in the lysine 2,3-aminomutase reaction. *J. Biol. Chem*. 1987;262(31):14859-14862.

68. Kriek M, Martins F, Challand MR, Croft A, Roach PL. Thiamine biosynthesis in Escherichia coli: identification of the intermediate and by-product derived from tyrosine. *Angew. Chem. Int. Ed. Engl*. 2007;46(48):9223-9226.

69. Thauer RK, Kaster A, Goenrich M, et al. Hydrogenases from Methanogenic Archaea, Nickel, a Novel Cofactor, and H(2) Storage. *Annu Rev Biochem*. 2010.

70. Xia Q, Wang T, Hendrickson EL, et al. Quantitative proteomics of nutrient limitation in the hydrogenotrophic methanogen Methanococcus maripaludis. *BMC Microbiol*. 2009;9:149.

71. Afting, Hochheimer, Thauer. Function of H2-forming methylenetetrahydromethanopterin dehydrogenase from *methanobacterium thermoautotrophicum* in coenzyme F420 reduction with H2. *Arch. Microbiol*. 1998;169(3):206-210.

72. Sintonen M. Reasoning to Hypotheses: Where Do Questions Come? *Foundations of Science*. 2004;9(3):249-266.

73. Rubio LM, Ludden PW. Maturation of nitrogenase: a biochemical puzzle. *J. Bacteriol*. 2005;187(2):405-414.

74. Rubio LM, Ludden PW. Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Annu. Rev. Microbiol*. 2008;62:93-111.

75. Girbal L, von Abendroth G, Winkler M, et al. Homologous and heterologous overexpression in *Clostridium acetobutylicum* and characterization of purified clostridial and algal Fe-only hydrogenases with high specific activities. *Appl. Environ. Microbiol*. 2005;71(5):2777-2781.

76. Nagy LE, Meuser JE, Plummer S, et al. Application of gene-shuffling for the rapid generation of novel [FeFe]-hydrogenase libraries. *Biotechnol. Lett*. 2007;29(3):421-430.

77. Krebs C, Henshaw TF, Cheek J, Huynh BH, Broderick JB. Conversion of 3Fe-4S to 4Fe-4S Clusters in Native Pyruvate Formate-Lyase Activating Enzyme: Mossbauer Characterization and Implications for Mechanism. *J. Am. Chem. Soc*. 2000;122(50):12497-12506.

78. Peck HD, Gest H. A new procedure for assay of bacterial hydrogenases. *J. Bacteriol*. 1956;71(1):70-80.

79. McGlynn SE, Ruebush SS, Naumov A, et al. In vitro activation of [FeFe] hydrogenase: new insights into hydrogenase maturation. *J. Biol. Inorg. Chem*. 2007;12(4):443-447.

80. Dos Santos PC, Dean DR, Hu Y, Ribbe MW. Formation and insertion of the nitrogenase iron-molybdenum cofactor. *Chem. Rev*. 2004;104(2):1159-1173.

81. Mulder DW, Ortillo DO, Gardenghi DJ, et al. Activation of HydA(DeltaEFG) requires a preformed [4Fe-4S] cluster. *Biochemistry*. 2009;48(26):6240-6248.

82. Ugulava NB, Gibney BR, Jarrett JT. Biotin synthase contains two distinct iron-sulfur cluster binding sites: chemical and spectroelectrochemical analysis of iron-sulfur cluster interconversions. *Biochemistry*. 2001;40(28):8343-8351.

83. McGlynn SE, Shepard EM, Winslow MA, et al. HydF as a scaffold protein in [FeFe] hydrogenase H-cluster biosynthesis. *FEBS Lett*. 2008;582(15):2183-2187.

84. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem*. 1976;72:248-254.

85. Fish WW. Rapid colorimetric micromethod for the quantitation of complexed iron in biological samples. *Meth. Enzymol*. 1988;158:357-364.

86. Chen Z, Lemon BJ, Huang S, et al. Infrared studies of the CO-inhibited form of the Fe-only hydrogenase from *Clostridium pasteurianum* I: examination of its light sensitivity at cryogenic temperatures. *Biochemistry*. 2002;41(6):2036-2043.

87. Thorneley R, George S. *Prokaryotic nitrogen fixation: a model system for the analysis of a biological process*. Ed. Triplett EW, Horizon Scientific Press, Wymondham, UK; 2000.

88. Frazzon J, Dean DR. Formation of iron-sulfur clusters in bacteria: an emerging field in bioinorganic chemistry. *Curr Opin Chem Biol*. 2003;7(2):166-173.

89. Johnson DC, Dean DR, Smith AD, Johnson MK. Structure, function, and formation of biological iron-sulfur clusters. *Annu. Rev. Biochem*. 2005;74:247-281.

90. Curatti L, Hernandez JA, Igarashi RY, et al. In vitro synthesis of the iron-molybdenum cofactor of nitrogenase from iron, sulfur, molybdenum, and homocitrate using purified proteins. *Proc. Natl. Acad. Sci. U.S.A*. 2007;104(45):17626-17631.

91. Bonomi F, Iametti S, Ta D, Vickery LE. Multiple turnover transfer of [2Fe2S] clusters by the iron-sulfur cluster assembly scaffold proteins IscU and IscA. *J. Biol.*

Chem. 2005;280(33):29513-29518.

92. Pierik AJ, Hulstein M, Hagen WR, Albracht SP. A low-spin iron with CN and CO as intrinsic ligands forms the core of the active site in [Fe]-hydrogenases. *Eur. J. Biochem*. 1998;258(2):572-578.

93. Sofia HJ, Chen G, Hetzler BG, Reyes-Spindola JF, Miller NE. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res*. 2001;29(5):1097-1106.

94. Pilet E, Nicolet Y, Mathevon C, et al. The role of the maturase HydG in [FeFe]-hydrogenase active site synthesis and assembly. *FEBS Letters*. 2009;583(3):506-511.

95. Tracqui A.[1], Raul J.S.[1], Gé, et al. Determination of Blood Cyanide by HPLC-MS. *Journal of Analytical Toxicology*. 2002;26:144-148.

96. Challand MR, Martins FT, Roach PL. Catalytic activity of the anaerobic tyrosine lyase required for thiamine biosynthesis in Escherichia coli. *J. Biol. Chem*. 2010;285(8):5240-5248.

97. Driesener RC, Challand MR, McGlynn SE, et al. [FeFe]-hydrogenase cyanide ligands derived from S-adenosylmethionine-dependent cleavage of tyrosine. *Angew. Chem. Int. Ed. Engl*. 2010;49(9):1687-1690.

98. Jurgenson CT, Begley TP, Ealick SE. The structural and biochemical foundations of thiamin biosynthesis. *Annu. Rev. Biochem*. 2009;78:569-603.

99. Bettendorff L, Wins P. Thiamin diphosphate in biological chemistry: new aspects of thiamin metabolism, especially triphosphate derivatives acting other than as cofactors. *FEBS J*. 2009;276(11):2917-2925.

100. Martinez-Gomez NC, Robers M, Downs DM. Mutational analysis of ThiH, a member of the radical S-adenosylmethionine (AdoMet) protein superfamily. *J. Biol. Chem*. 2004;279(39):40505-40510.

101. Antonini E, Brunori M. *Hemoglobin and Myoglobin in their Reactions with Ligands*. North-Holland Pub. Co.; 1971.

102. McGlynn SE, Boyd ES, Shepard EM, et al. Identification and characterization of a novel member of the radical AdoMet enzyme superfamily and implications for the biosynthesis of the Hmd hydrogenase active site cofactor. *J. Bacteriol*. 2010;192(2):595-598.

103. Chatterjee A, Li Y, Zhang Y, et al. Reconstitution of ThiC in thiamine pyrimidine biosynthesis expands the radical SAM superfamily. *Nat. Chem. Biol*. 2008;4(12):758-765.

104. Martinez-Gomez NC, Downs DM. ThiC is an [Fe-S] cluster protein that requires AdoMet to generate the 4-amino-5-hydroxymethyl-2-methylpyrimidine moiety in thiamin synthesis. *Biochemistry*. 2008;47(35):9054-9056.

105. Larkin MA, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23(21):2947-2948.

106. Cuff JA, Birney E, Clamp ME, Barton GJ. ProtEST: protein multiple sequence alignments from expressed sequence tags. *Bioinformatics*. 2000;16(2):111-116.

107. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol*. 2003;52(5):696-704.

108. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*. 2003;19(12):1572-1574.

109. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*. 2001;17(8):754-755.

110. Baldauf SL, Palmer JD, Doolittle WF. The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny. *Proc. Natl. Acad. Sci. U.S.A*. 1996;93(15):7749-7754.

111. Chew AGM, Bryant DA. Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity. *Annu. Rev. Microbiol*. 2007;61:113-129.

112. Duin EC, Lafferty ME, Crouse BR, et al. [2Fe-2S] to [4Fe-4S] cluster conversion in *Escherichia coli* biotin synthase. *Biochemistry*. 1997;36(39):11811-11820.

113. Buis JM, Cheek J, Kalliri E, Broderick JB. Characterization of an active spore photoproduct lyase, a DNA repair enzyme in the radical S-adenosylmethionine superfamily. *J. Biol. Chem*. 2006;281(36):25994-26003.

114. Rebeil R, Nicholson WL. The subunit structure and catalytic mechanism of the *Bacillus subtilis* DNA repair enzyme spore photoproduct lyase. *Proc. Natl. Acad. Sci. U.S.A*. 2001;98(16):9038-9043.

115. Chow CS, Lamichhane TN, Mahto SK. Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications. *ACS Chem. Biol*. 2007;2(9):610-619.

116. Kozbial PZ, Mushegian AR. Natural history of S-adenosylmethionine-binding proteins. *BMC Struct. Biol*. 2005;5:19.

117. Noma A, Kirino Y, Ikeuchi Y, Suzuki T. Biosynthesis of wybutosine, a hyper-modified nucleoside in eukaryotic phenylalanine tRNA. *EMBO J*. 2006;25(10):2142-2154.

118. Suzuki Y, Noma A, Suzuki T, Ishitani R, Nureki O. Structural basis of tRNA modification with CO2 fixation and methylation by wybutosine synthesizing enzyme TYW4. *Nucleic Acids Res*. 2009;37(9):2910-2925.

119. Ruzicka FJ, Lieder KW, Frey PA. Lysine 2,3-aminomutase from Clostridium subterminale SB4: mass spectral characterization of cyanogen bromide-treated peptides and cloning, sequencing, and expression of the gene kamA in *Escherichia coli*. *J. Bacteriol*. 2000;182(2):469-476.

120. Frey M, Rothe M, Wagner AF, Knappe J. Adenosylmethionine-dependent synthesis of the glycyl radical in pyruvate formate-lyase by abstraction of the glycine C-2 pro-S hydrogen atom. Studies of [2H]glycine-substituted enzyme and peptides homologous to the glycine 734 site. *J. Biol. Chem*. 1994;269(17):12432-12437.

121. Mulliez E, Padovani D, Atta M, Alcouffe C, Fontecave M. Activation of class III ribonucleotide reductase by flavodoxin: a protein radical-driven electron transfer to the iron-sulfur center. *Biochemistry*. 2001;40(12):3730-3736.

122. Goldman AD, Leigh JA, Samudrala R. Comprehensive computational analysis of Hmd enzymes and paralogs in methanogenic Archaea. *BMC Evol. Biol*. 2009;9:199.

123. Soboh B, Igarashi RY, Hernandez JA, Rubio LM. Purification of a NifEN protein complex that contains bound molybdenum and a FeMo-Co precursor from an *Azotobacter vinelandii* Delta nifHDK strain. *J. Biol. Chem*. 2006;281(48):36701-36709.

124. Imlay JA. Iron-sulphur clusters and the problem with oxygen. *Mol. Microbiol*. 2006;59(4):1073-1082.

125. Poole AM, Logan DT, Sjöberg B. The evolution of the ribonucleotide reductases: much ado about oxygen. *J. Mol. Evol*. 2002;55(2):180-196.

126. Xiong J, Fischer WM, Inoue K, Nakahara M, Bauer CE. Molecular evidence for the early evolution of photosynthesis. *Science*. 2000;289(5485):1724-1730.

127. Daniel RM, Danson MJ. Did primitive microorganisms use nonhem iron proteins in place of NAD/P? *Journal of Molecular Evolution*. 1995;40(6):559-563.

128. Kuchenreuther JM, Stapleton JA, Swartz JR. Tyrosine, cysteine, and S-adenosyl methionine stimulate in vitro [FeFe] hydrogenase activation. *PLoS ONE*. 2009;4(10):e7565.

129. Laville J, Blumer C, Von Schroetter C, et al. Characterization of the hcnABC Gene Cluster Encoding Hydrogen Cyanide Synthase and Anaerobic Regulation by ANR in the Strictly Aerobic Biocontrol Agent *Pseudomonas fluorescens* CHA0. *J. Bacteriol.* 1998;180(12):3187-3196.

130. Blumer C, Haas D. Mechanism, regulation, and ecological role of bacterial cyanide biosynthesis. *Arch. Microbiol*. 2000;173(3):170-177.

131. Gould SJ, Vrba ES. Exaptation-A Missing Term in the Science of Form. *Paleobiology*. 1982;8(1):4-15.

132. Diamond J. *The third chimpanzee*. HarperPerennial; 1993.

133. Chapman-Smith A, Cronan JE. Molecular biology of biotin attachment to proteins. *J. Nutr*. 1999;129:477S-484S.

134. Makarchikov AF, Lakaye B, Gulyai IE, et al. Thiamine triphosphate and thiamine triphosphatase activities: from bacteria to mammals. *Cell. Mol. Life Sci*. 2003;60(7):1477-1488.

135. Leipe DD, Wolf YI, Koonin EV, Aravind L. Classification and evolution of P-loop GTPases and related ATPases. *J. Mol. Biol*. 2002;317(1):41-72.

136. Hamann N, Mander GJ, Shokes JE, et al. A cysteine-rich CCG domain contains a novel [4Fe-4S] cluster binding motif as deduced from studies with subunit B of heterodisulfide reductase from Methanothermobacter marburgensis. *Biochemistry*. 2007;46(44):12875-12885.

137. Soriano A, Hausinger RP. GTP-dependent activation of urease apoprotein in complex with the UreD, UreF, and UreG accessory proteins. *Proc. Natl. Acad. Sci. U.S.A*. 1999;96(20):11140-11144.

138. Zambelli B, Stola M, Musiani F, et al. UreG, a chaperone in the urease assembly process, is an intrinsically unstructured GTPase that specifically binds Zn2+. *J. Biol. Chem*. 2005;280(6):4684-4695.

139. Larry E. Vickery, Jill R. Cupp-Vickery. Molecular Chaperones HscA/Ssq1 and HscB/Jac1 and Their Roles in Iron-Sulfur Protein Maturation. 2007;42(2): 95-111.

140. Maeda T, Sanchez-Torres V, Wood TK. Enhanced hydrogen production from

glucose by metabolically engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol*. 2007;77(4):879-890.

141. Akhtar MK, Jones PR. Deletion of iscR stimulates recombinant clostridial Fe-Fe hydrogenase activity and H2-accumulation in *Escherichia coli BL21(DE3)*. *Appl. Microbiol. Biotechnol*. 2008;78(5):853-862.

142. Akhtar MK, Jones PR. Construction of a synthetic YdbK-dependent pyruvate:H2 pathway in *Escherichia coli BL21(DE3)*. *Metab. Eng*. 2009;11(3):139-147.

143. Agapakis CM, Ducat DC, Boyle PM, et al. Insulation of a synthetic hydrogen metabolism circuit in bacteria. *J Biol Eng*. 2010;4:3.