HEINRICH HEINE
UNIVERSITÄT
DÜSSELDORF

Genetic and Molecular Characterizations of Hydrogenases and Sulfate Metabolism Proteins of *Desulfovibrio vulgaris* Miyazaki F

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Amrit Pal Kaur
aus Amritsar (Indien)

Düsseldorf, January 2009

Zusammenfassung

Der hier untersuchte Organismus *Desulfovibrio vulgaris* Miyazaki F (DvMF) ist seit langem bekannt auf Grund detaillierter Untersuchungen an seiner [NiFe] Hydrogenase und an Sulfat-Stoffwechselproteinen sowie einiger anderer Metallo-Enzyme, die in Zusammenhang mit den Hydrogenasen stehen. Unabhängig von den relativ guten Ausbeuten, die man für diese Proteine erzielt, ist die funktionale und genetische Manipulation schwierig, da die Sequenzen der kodierenden Gene nicht bekannt waren sind. In dieser Arbeit wurden genetische und molekularbiologische Fortschritte erzielt und weitere Informationen über den Stoffwechsel der Hydrogenasen von DvMF erhalten. Das gesamte Genom von DvMF wurde in eine Cosmid-Bank kloniert. Die Verwendung von dioxygenin-markierten DNA-Sonden aus bekannten homologen Sequenzen von *D. vulgaris* Hildenborough (DvH) führte zur Identifizierung der gesuchten Gene. Sie kodieren zwei zusätzliche Hydrogenasen, eine aus zwei Untereinheiten bestehende [NiFeSe] Hydrogenase (*hysA* und *hysB*) und eine sechs Untereinheiten große Ech Hydrogenase (*echA* - *echF*). Zusätzlich wurden die Gensequenzen für fünf „Maturierungs"-Proteine aufgeklärt (*hypA* - *hypF*). Diese Methode wurde auch angewandt auf die Gene von Proteinen des Sulfatmetabolismus, 5'-Adenylyl Sulfatreduktase (*aprA* und *aprB*) und Desulfoviridin (*dsvA*, *dsvB*, *dsvD* und *DsvC*). Ein weiteres wichtiges Ergebnis ist der eindeutige Nachweis, dass ein [FeFe] Hydrogenase-kodierendes Gen in DvMF nicht vorhanden ist. Methodisch wurden sämtliche Gene sequenziert, indem mittels degenerierter "primer" aus identifizierten, konservierten Genbereichen heraus sequenziert wurde. Die Gene der [NiFe] - und der [NiFeSe] - Hydrogenasen wurden zusammen mit den Maturierungsgenen Epitop-getaggt. Die Verwendung von *E. coli* als Wirtszelle mit einem T7-Promotor Expressionsvektor ergab die Reinigung von Strep-getaggter [NiFe] Hydrogenase, die Integration von Selenocystein in die [NiFeSe] Hydrogenase und die gleichzeitige Expression von zwei-Gen operons (hypAB und hypDE). Hiermit wurden wichtige Fortschritte hin zu einer heterologen Expression von physiologisch aktiven Hydrogenasen erzielt. Zusätzlich zu den Untersuchungen an Hydrogenasen konnten in gemeinsamer Arbeit die Protein-Komplexe der 5'-Adenylyl-sulfatreduktase und von Desulfoviridin gereinigt und kristallisiert werden. Regulatorische Sequenzen der Gene wurden mit verschiedenen Programmen detektiert, ebenfalls wurden dreidimensionale Modelle für die Proteine generiert und eine phylogenetische Analyse durchgeführt.

Abstract

The studied organism *Desulfovibrio vulgaris* Miyazaki F (DvMF) has long been known from detailed investigations of its [NiFe] hydrogenase, of its proteins involved in sulfate metabolism proteins and of some other metallo-enzymes associated with hydrogenases. Although the yields of such proteins from DvMF are high, the functional and genetic manipulations have been a daunting task due to non-availability of the sequences of the coding genes. In this work, genetic and molecular biological advances have been made to add on information on hydrogenase metabolism of DvMF. The entire genome of DvMF was cloned into a cosmid genomic library. Probing by means of dioxygenin-labeled DNA probes derived from the known homologous sequences of *Desulfovibrio vulgaris* Hildenborough (DvH) led to the identification of several genes of focus. These genes included two additional hydrogenases, a two subunit [NiFeSe] hydrogenase (*hysA* and *hysB*) and a six subunit Ech hydrogenase (*echA*, *echB*, *echC*, *echD*, *echE* and *echF*), as well as five maturation genes *hypA*, *hypB*, *hypD*, *hypE* and *hypF*. This method was extended to sequence the genes coding for the sulfate metabolism proteins 5'-adenylyl sulfate reductase (*aprA* and *aprB*) and desulfoviridin (*dsvA*, *dsvB*, *dsvD* and *dsvC*). The presence of an [FeFe] hydrogenase could be ruled out. All of the identified genes were sequenced by designing degenerate primes from the conserved domains. The [NiFe]- and [NiFeSe]- hydrogenases along with maturation gene operons were epitope tagged by means of PCR and cloning. Using *E. coli* as a host along with a T7 promoter expression vector, the purification of strep-tagged [NiFe] hydrogenase, integration of selenocysteine into the [NiFeSe] hydrogenase and the simultaneous expression of two-gene operons (*hypAB* and *hypDE*) of *hyp* maturation genes happened to be significant observations for advancing in the direction of heterologous expression and modification of hydrogenases in physiological functional forms. In addition to the work aiming at hydrogenases, the protein complexes of 5'-adenylyl sulfate reductase and desulfoviridin were purified and crystallized in a collaborative effort. Using various software, the possible regulatory sequences for the sequenced genes, the three dimensional models for the translated proteins and the phylogenetic analysis based on amino acid sequence conservation were also successfully performed.

Curriculum Vitae

Personal Information

Name	Amrit Pal Kaur
Birth Date	9th Oct., 1978
Birth place	Amritsar, India

Education

1993	Basic Schooling
1995	12th standard (Physics, Chemistry, Biology)
1996-1999	Bachelors of Science (Industrial Microbiology) at G.N.D university, Amritsar, India.
1999-2001	Masters of Science (Biotechnology) at G.N.D university, Amritsar, India.
2000	Two months summer training research ‘’Molecular diversity analysis in *Capsicum annuum* using RAPD and AFLP profiling of morphological variants’’ carried under supervision of Dr. S. P. S. Khanuja at CIMAP, Lucknow, India.
2001	Master thesis ‘’Glutamate Excitotoxicity: Its correlation with neuronal plasticity in adult rats’’ carried under supervision of Prof. G Kaur, at G.N.D university, Amritsar, India.
08-12/2002	Research work ‘’Deletion of chromosome 3p14.2-p25 involving the VHL and FHIT genes in conventional renal cell carcinoma’’ carried under supervision of Prof. G Kovacs at R.K. Universität, Heidelberg, Germany.
2003-2004	Research work ‘Transcriptional fusions of regulatory sequences of lipase of *Pseudomonas aeruginosa* with GFP (green fluorescent protein) carried under supervision of Prof. K.E. Jäger, at IMET, FZ-Jülich, Germany.
07/2004-present	Working on doctoral thesis ‘‘Genetic and Molecular Characterizations of Hydorgenases and Sulfate metabolism proteins of *Desulfovibrio vulgaris* Miyazaki F’’ under supervision of Prof. W. Gärtner, at MPI-BAC, Mülheim a.d. Ruhr, Germany

Publications

Kaur G., Basu A., and Kaur A.P. (2003) Glutamate Excitotoxicity: Its Correlation with Neuronal Plasticity in Young Adult and Neonate Rats. *Journal of Neurological Sciences***, 20**(4:40)

Sukosad F., Kuroda N., Beothe T., Kaur A. P., and Kovacs G. (2003) Deletion of Chromosome 3p14.2-p25 involving the VHL and FHIT genes in Conventional renal carcinoma. *Cancer Res.* **63(2)**, 455-457.

In this work

Ogata, H., Goenka Agrawal, A., Kaur, A. P., Goddard, R., Gärtner, W., and Lubitz W. (2008) Purification, crystallization and preliminary X-ray analysis of adenylylsulfate reductase from *Desulfovibrio vulgaris* Miyazaki F. *Acta. Cryst.*, **64(11)**, 1010-1012.

Acknowledgement

I would like to extend my sincere gratitude to Prof. Dr. Wolfgang Gärtner for giving me the opportunity to work in his group and providing an excellent guidance for carrying this doctoral thesis research on hydrogenases. I am thankful for his insightful inputs and for getting a very liberal working environment where it was possible to learn new things and apply own ideas. Prof. Gärtner gave an indispensible help to edit and organize the scientific content of this thesis as well as improving the grammar. His help was again needed to translate the English text of the presented abstract to Deutsch language. I am equally thankful to Prof. Dr. Wolfgang Lubitz for arranging the basic and auxiliary funding of this project and for his constant interest in the development of the molecular biology of the hydrogenases at MPI-BAC, Mülheim an der Ruhr. I am also grateful to Prof. Dr. Karl-Erich Jäger, who has agreed to co-evaluate my thesis.

I am thankful to Dr. Hideaki Ogata and Dr. Aruna Goenka Agrawal for fruitful collaborations over purification and crystallization of APS reductase and desulfoviridin proteins. I am also thankful to Dr. Aruna Goenka Agrawal for providing a [NiFe] hydrogenase deletion mutant of *Desulfovibrio vulgaris* subsp Hildenborough.

I am grateful to Dr. Marcus Ludwig of (Department of Biology) Humboldt University, Berlin, for teaching me to perform hydrogenase activity assays. At the same time, I wish to thank Dr. Oliver Lenz of the same institute for sharing his experience with the molecular biology of *Ralstonia eutropha* H16. It has been very kind of Dr. Inês C. Pereira of ITQB, Portugal for providing the purified [NiFeSe] hydrogenase protein from *Desulfovibrio vulgaris* subsp Hildenborough.

My laboratory coworkers, Mrs. Shivani Sharda, Mr. Cao Zhen, Mr. Björn Zorn, Mrs. Rashmi Shah and Mr. Gopal Pathak do deserve special thanks for maintaining a cordial and work-friendly environment. Above all it is very appreciative of Mrs. Helena Steffen for maintaining the order and providing basic support for keeping things functional in the laboratory.

Last, but not the least, it would be incomplete without acknowledging the constructive support I received from my parents and extended family. I am indebted to my hard working parents for giving me a good education, which has been the actual foundation of writing this thesis and thank you note.

Abbreviations

aa	amino acid(s)
AP	alkaline phosphatase
ATP	adenosine-5'-triphosphate
BCIP	5-bromo-4-chloro-3-indolyl-phosphate
bp	base pairs
BSA	bovine serum albumin
BV	Benzyl viologen
CTP	cytosine-5’-triphosphate
dig	dioxygenin
DMSO	dimethyl sulfoxide
dNTP	deoxynucleoside-triphosphate mix
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
GTP	guanosine-5'-triphosphate
Hyp	hydrogenase pleiotropy
IPTG	isopropyl-b-D-thiogalactopyranoside
LB	Luria Bertani media
MCS	multiple cloning site
MOPS	3-(N-morpholino)propanesulfonic acid
MV	Methyl viologen
NBT	nitrobluetetrazolium
OD	optical density
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PB	postgate B medium
PC	postgate C medium
PCR	polymerase chain reaction
PE	postgate E medium
PMS	phenazine
PVDF	polyvinylidene difluoride
rpm	rotations per minute
RT	room temperature
SDS	sodium dodecyl sulphate
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
TTC	2, 3, 5-Triphenyl Tetrazolium Chloride
TTP	thymidine-5'-triphosphate
UTP	uridine-5’-triphosphate

INDEX

Chapter 1

Introduction

Microbes adopt vast metabolic strategies to sustain life even in the most extreme environments; for that, they explore almost every possible energy source and oxidant available. Anaerobic respiring prokaryotes are known to be capable of using carbon dioxide, sulfate, sulfur, nitrate and iron oxides as terminal acceptor for a respiratory electron transfer chain. Among these, the sulfate reducing bacteria (SRB) are widespread in the anaerobic zones of multiple terrestrial and aquatic environments and especially in habitats where sulfate is available in plenty.

The members of the *Desulfovibrio* genus belong to the δ-subdivision of the proteobacteria and carry out dissimilatory sulphate reduction in their natural habitats, a regular cause of foul smells. The biological sulfate reduction is an important part of the sulfur cycle as sulfate is a rather stable molecule and is less energetically favorable to be reduced as compared to nitrate or oxygen. A series of enzymes and electron transfer complexes present in the cytoplasm and periplasm of these organisms are involved in sulfate reduction and provide the link to the diverse hydrogen metabolisms of the cell.

Desulfovibrio are the best studied SRB and have considerable economic impact on the oil industry due to their involvement in bio-corrosion (Dinh et al, 2004, Hamilton, 1998, 2003), and souring of oil and gas deposits by the sulfides produced (Hamilton, 1998). The metabolism of these organisms is also being explored for their potential role in bioremediation of polluted aqueous environments (Lovley, 2003) as, e.g., in the decontamination of toxic metals, radionuclides (Wall and Krumholz 2006, Lloyd et al 2003), aromatic hydrocarbons and chlorinated compounds (Gibson and Harwood, 2002). Sulfur reducing organisms also have health implications since there is evidence suggesting they may have a role in inflammatory bowel diseases, due to the toxic effects of sulphide on colonic epithelial cells (Loubinoux et al, 2002, Pitcher and Cummings, 1996).

The most interesting features of the *Desulfovibrio* metabolism are their hydrogenase diversity and sulfate reduction capability. Historically, *Desulfovibrio* have remained a rich source of metalloproteins and have been intensely studied for the wide array of the metal-associated hydrogenases they possess. Reasonable advances have been made in understanding these proteins in some species, e.g., in *Desulfovibrio vulgaris* subsp Hildenborough (DvH) and *Desulfovibrio fructosovorans*, by means of genome sequencing or mutant construction. The study of these organisms is interesting, but also difficult due to obstacles associated with genetic manipulation of SRB. SRB are naturally resistant to many antimicrobials (Postgate, 1984) and only a small number of facile selectable markers can be used to select a genetically modified or transformed cell.

Another aspect which is motivating hydrogenase research is the proposed use of hydrogen as an environment-friendly fuel of the future, especially in wake of fossil fuel crisis and global warming. A bio-fuel cell containing a hydrogenase for the oxidation of hydrogen produced by photosynthetic enzymes on a cathode has been proposed (Vincent et al, 2005). However the oxygen sensitivity and the stability of the hydrogenase remains a challenge for contemporary biotechnologists and biophysicists. Therefore, a better understanding of these enzymes is needed to implement them in a more useful way.

Most of the organisms that are capable of metabolizing hydrogen usually possess more than one hydrogenase. The simultaneous presence of several hydrogenases, and their regulation and interplay with other components of hydrogen metabolism is highly complex and will be discussed in detail in the following chapter.

Our study group has been able to purify some of the interesting proteins from our study organism *Desulfovibrio vulgaris* subsp Miyazaki F (DvMF), such as the [NiFe] hydrogenase, APS (adenosine 5'-phosphosulfate) reductase and Desulfoviridin. The [NiFe] hydrogenase has been consistently characterized biochemically and spectroscopically (Lubitz et al, 2007). The other two protein complexes, APS reductase and the Desulfoviridin have initially been purified (by Dr. Arunka Goenka Agrawal) and crystallized (by Dr. Hideaki Ogata). Detailed characterization of the two sulfate metabolism-related proteins has been accomplished during this thesis. Also, the

three-dimensional structure for the APS reductase is resolved (Ogata et al, 2008). However, limited information about the genetics of these proteins and the expression and maturation of the hydrogenases for *Desulfovibrio vulgaris* subsp Miyazaki F (DvMF) was available at the beginning of this work. In this thesis, genetic, sequential and molecular characterizations of the two crystallized proteins involved in the sulfate metabolism have been performed. The existence and absence of additional hydrogenases has been detected and their sequence determined, as has been done with the proteins involved in the maturation of [NiFe] hydrogenase from *Desulfovibrio vulgaris* Miyazaki F (DvMF).

Chapter 2

Hydrogenases and Sulfate Reduction

2.1 Hydrogenases

Hydrogenases are oxidoreductases, which catalyze the reversible oxidation of hydrogen to protons and electrons. Purified [NiFe] hydrogenases are found to catalyse both H_2 evolution and -uptake depending on the oxidation state and presence of a suitable acceptor (A) or donor (D).

$$2H^+ + D_{red} \rightleftharpoons H_2 + D_{ox}$$
$$H_2 + A_{ox} \rightleftharpoons 2H^+ + A_{red}$$

Physiologically the hydrogenases combine the oxidation of a hydrogen molecule with the reduction of an electron acceptor coupled with the energy conservation, maintenance of intracellular pH and redox potential. Substrates for *Desulfovibrio* bacteria range from hydrogen or acetate to other small organic acids such as lactate, malate, pyruvate and ethanol etc. They are also capable of utilizing hydrogen as a sole energy source when given acetate and CO_2 as carbon sources are present.

2.2 Multiplicity of Hydrogenases and Hydrogen Cycling

In a cell, hydrogenases can also act as hydrogen sensors (e.g. *Ralstonia eutropha*), however, most hydrogenases function as enzymes and make use of hydrogen by coupling to other cellular reactions to complement the energy metabolism of the cell by oxidizing or producing hydrogen. Widely distributed in prokaryotes and most classes of Archaea and Bacteria, the hydrogenases are present in more than a single form and associated with different transistion metals. Three different phylogenetic classes of hydrogenases are discussed in section 2.3.

Initial studies carried out by Odom and Peck (1981), have highlighted the role of hydrogen in *Desulfovibrio* metabolism by proposing a hydrogen cycling hypothesis which suggests the cyclic transfer of hydrogen. For the transfer of hydrogen produced in the cytoplasm via the lactate

metabolism, it is transferred to the periplasm after oxidation by a periplasmic hydrogenase. The protons and electrons thus generated move through the membrane in a vectorial manner involving various redox partners to finally produce a proton gradient for ATP synthesis. Similarly, excessive protons are combined with electrons to produce hydrogen by a cytoplasm-located hydrogenase depending on the metabolic state of the cell. Thus, the regular cycling of hydrogen between the cytoplasm and periplasm plays an integral role in balancing the bioenergetics.

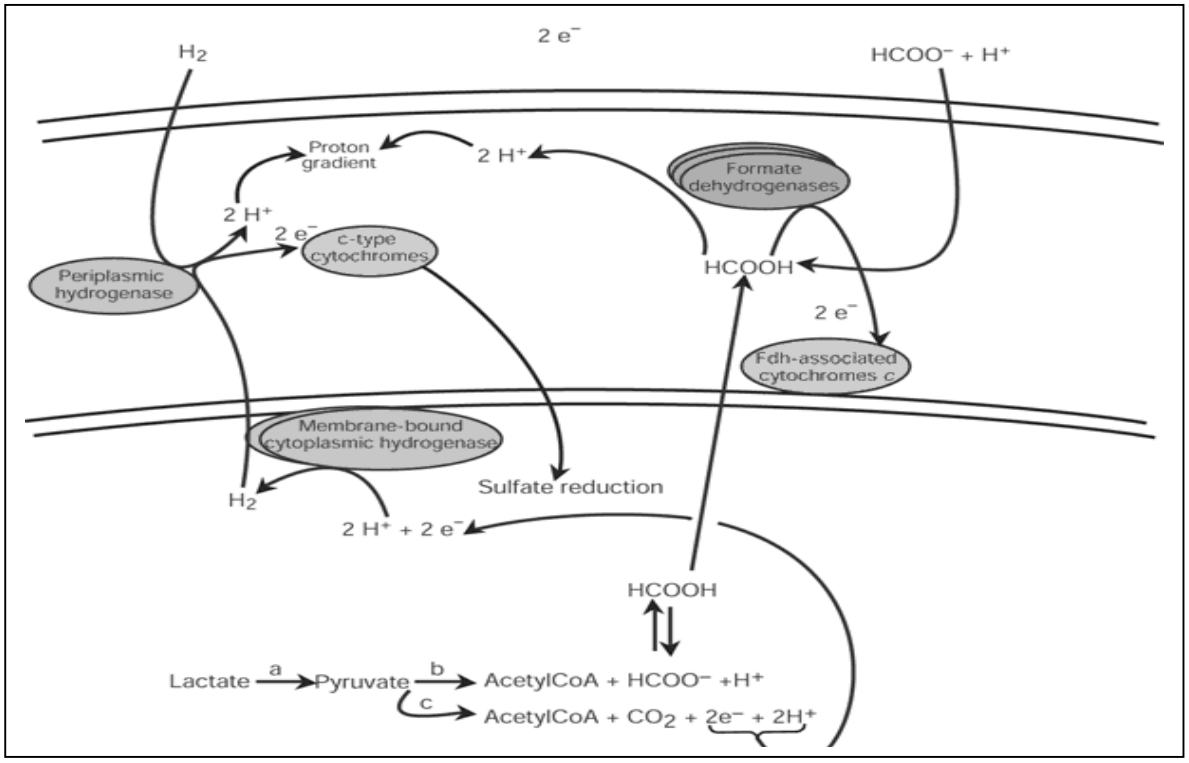


Fig 2.1 Diagrammatic representation of the proposed hydrogen and formate cycling in *D. vulgaris* Hildenborough (Heidelberg et al, 2004). Reducing equivalents ($2H^+ + 2e^-$) generated from lactate or pyruvate oxidation are the substrate for one of two cytoplasm oriented, membrane-bound hydrogenases. The hydrogen molecule can diffuse to the periplasm where one or more of periplasmic hydrogenases can oxidize the hydrogen to release electrons to the c-type cytochrome network. The electrons could then be channeled through the cytoplasmic membrane by one of several putative transmembrane protein channels. The protons generated may contribute to the proton gradient supporting transport processes and ATP synthesis.

In line with this hypothesis, the multiplicity of hydrogenases in the central metabolism of these bacteria should be corroborated by the presence of at least one cytoplasmic hydrogenase producing hydrogen while using energy from the carbon metabolism. However, with the

exception of *Desulfovibrio fructosovorans* (Casalot et al, 2002) no cytoplasm-located hydrogenase has been reported in this group of sulfate reducing bacteria. For the best studied species in this group, *Desulfovibrio vulgaris* Hildenborough (DvH), four of the six identified hydrogenases are reported to be involved in hydrogen oxidation and are periplasmic (Heidelberg et al, 2004), and the other two are predicted to be membrane associated. Amongst the periplasmic hydrogenases, there is a soluble [FeFe] hydrogenase (Hyd), two membrane-associated nickel-iron [NiFe] hydrogenase isozymes (Hyn1 and Hyn2), and one membrane-associated nickel-iron-selenium [NiFeSe] hydrogenase (Hys). Two more hydrogenases that have been described, Coo (CO induced) and Ech (energy converting hydrogenase), are considered to be membrane associated and cytoplasm oriented. These latter two hydrogenases also belong to the group of [NiFe] hydrogenases and are thought to be providing an alternate pathway to the required role of a cytoplasm-located hydrogenase as given in the hydrogen cycling hypothesis fig 2.1.

2.3 Classifications and Phylogenies of Hydrogenases

Three main approaches have been employed for the elemental studies of hydrogenases. At first, hydrogenases were biochemically purified and studied for their catalytic abilities. Secondly, the hydrogenases were approached by genetic tools that led to the discovery of new genes of hydrogenases as well as the accessory genetic machinery needed to bring them into a structurally and functionally active form. And finally, selected hydrogenases, which are easy to purify and crystallize, have been used as models to decipher the mode and optimization of catalysis.

Based on such structural studies and phylogenetic analysis of the protein sequences, three distinct phylogenetic classes of hydrogenases (Viganis and Billoud, 2007) have been assigned, those are [NiFe]-, [FeFe]- and Hmd (5,10-methenyltetrahydromethanopterin hydrogenase) or [Fe]- hydrogenases (Shima and Thauer, 2007). The [NiFe] hydrogenases are further distributed into uptake [NiFe] hydrogenases, cyanobacterial uptake [NiFe] hydrogenases and H_2 sensors, bidirectional heteromultimeric cytoplasmic [NiFe] hydrogenases and H_2-evolving, and energy-conserving, membrane associated hydrogenases (Viganis and Billoud, 2007). Some of these hydrogenases such as the [NiFe] species are widely distributed and, as discussed before, DvH and many of the other species of *Desulfovibrio* are known to possess [NiFe]- and [FeFe]-

hydrogenases. There is a large conservation in the protein sequences and also among the available crystal structures of the similar types of hydrogenases and their maturation proteins, and thus it is hard to determine the evolutionary history of these modular enzymes (Viganis and Billoud, 2007). The uptake [NiFe] and [FeFe] hydrogenases have probably separate origins and phylogeny, even in spite of similarities in the cellular locations and functions. However, the evloutionary relationship between uptake [NiFe]- and the selenocysteine containing variant [NiFeSe] hydrogenase is not that apparent.

2.4 Purification and Structures of the Hydrogenases

Three hydrogenases out of the many types of known hydrogenases from *Desulfovibrio* have been regularly purified and studied in various labs; those are the [FeFe]-, [NiFe]- and the [NiFeSe]-hydrogenases.

2.4.1 [NiFe] hydrogenase

The [NiFe] hydrogenase from DvMF is periplasmic and membrane attached. For purification purposes, a trypsin digestion (a trypsin site is natively present) step targeting the N-terminal of the large subunit of this hydrogenase can be employed in solublising this enzyme during purification (Yagi et al, 1976). This hydrogenase when purified under aerobic conditions is partially inactive and has to be activated by reducing with hydrogen or other chemicals. Crystal structures from the [NiFe] hydrogenase of DvMF (Higuchi et al, 1997) and *D. gigas* (Volbeda et al, 1995) hydrogenases are available, and comprehensive biochemical and spectroscopical studies are being pursued in recent times to resolve the dynamics of the catalytic steps and charge (protons and electrons) transfer from this molecule to its redox partner cytochrome.

Structurally, the [NiFe] hydrogenases are always heterodimeric proteins and there is a large degree of conservation in the structural domains of [NiFe] hydrogenase even among phylogenetically diverse organisms. The parallel crystal structure investigations from the study organism DvMF and from the closely related *Desulfovibrio gigas* (Dg) show the protein consisting of small (SSU) and large (LSU) subunits with a total metal content of 12 Fe and 1 Ni atom. The small subunit of a catalytically active hydrogenase is an approximately 28 kDa

polypeptide and contains three iron-sulphur clusters, two of [Fe_4S_4] type (proximal and distal) and one centered [Fe_3S_4] cluster; the large subunit contains a nickel-iron center (Volbeda et al, 1995).

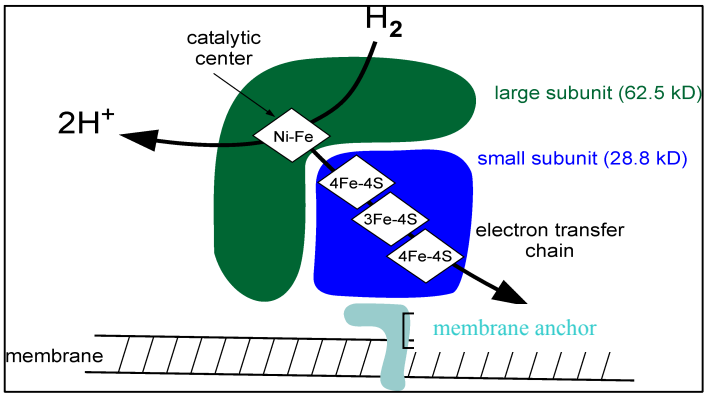


Fig 2.2 Electron transport chain from the [NiFe] active site to [FeS] clusters. Figure adapted from 'Hydrogen as a Fuel: Learning from nature', Robert Robson.

Similarly, the larger subunit of an active enzyme is an approximately 62.5 kDa polypeptide and contains the catalytically active site responsible for splitting di-hydrogen into two protons and two electrons. This active site has a bimetallic Ni-Fe arrangement; each of the two atoms is co-ordinated by two different cysteines. The Fe is also bound to two cyanide (CN) and one carbon monoxide (CO) ligands and the presence of some ligands bridging the two metals has been also proposed in various oxidized states (Higuchi, 2000, Volbeda et al, 2005). The nickel is paramagnetic in several functional states of the hydrogenases and gives different EPR (electron paramagnetic resonance) signals under different applied conditions. The first step of activation of the enzyme is the binding of H_2 to the metal centers, which polarizes the hydrogen bond and facilitates heterolytic cleavage.

$$H_2 \rightleftharpoons H^+ + H^-$$

This reaction can be used for studying the properties of Hydrogenase by deuterium and even tritium exchange reactions. Also, this hydrogenase is able to convert *para* to *ortho* hydrogen.

$$H_a\text{-}H_b\ (\uparrow\uparrow) + H^+_c \rightarrow H_a\text{-}H_c\ (\uparrow\downarrow) + H^+_b$$

Yet there has not been a precisely formulated and generally agreed catalytic mechanism for this enzyme. The flow of electrons and protons to the acceptors was inferred based on the arrangement of the iron sulfur clusters. The current approach in use is to delineate the exact sequence and sites of electron and proton transfer in the molecule by generating mutants and their subsequent activity studies.

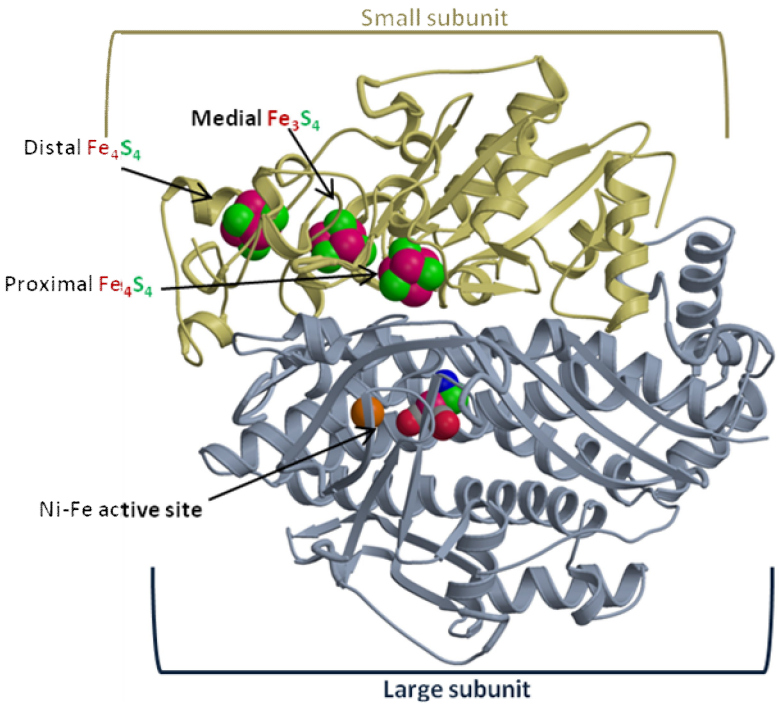


Fig 2.3 X-ray structure (resolution of 1.4 Å) of [NiFe] hydrogenase of *D. vulgaris* Miyazaki F in the reduced state [Here kindly provided by Dr. Hideaki Ogata]. Colour coding shows the two subunits, and the [NiFe] active site and the three [FeS] clusters are indicated.

Some mutants have been generated and characterized in *D. fructosovorans* (Df). With the help of site directed mutagenesis, proline-239 in the small subunit of the [NiFe] hydrogenase polypeptide was replaced by cysteine. This led to the conversion of the $[Fe_3S_4]$ cluster into an $[Fe_4S_4]$ cluster (Rousset et al, 1998). This cluster conversion resulted in a lowering of approximately 300 mV of the midpoint potential of the modified cluster (+65 mV to −250 mV), whereas no significant alteration of the spectroscopic and redox properties of the two native $[Fe_4S_4]$ clusters and the NiFe center occurred. The significant decrease of the midpoint potential of the intermediate Fe-S cluster had only a slight effect on the catalytic activity of the P238C

mutant as compared with the wild-type enzyme (Volbeda et al, 1995). It is possible that the spacing of the three clusters might be able to overcome the energetic barrier for electron transfer or the measured values might not be precise at microscopic level. Since the involvement of the median cluster is not the rate-limiting step for electron transfer, it was also speculated that there might be multiple pathways for electron transfer. It is also possible that the median cluster is only giving structural support to the channel governing the electron transfer.

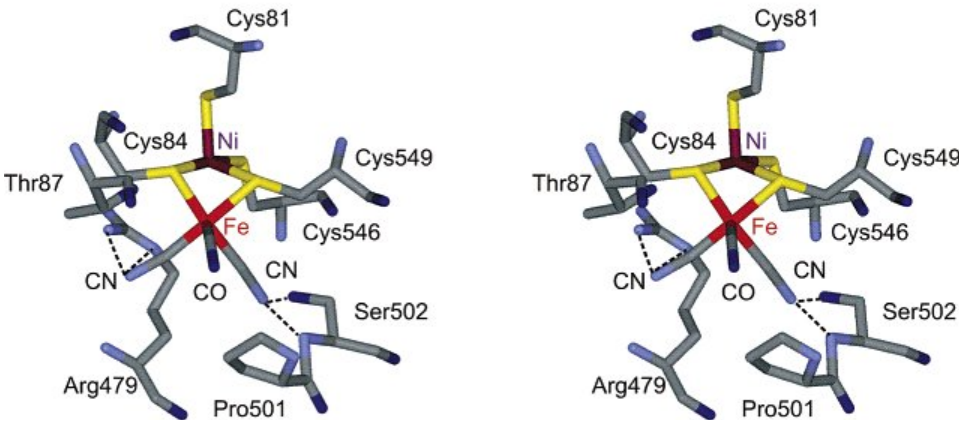


Fig 2.4 Stereoview of the active site and some of the surrounding amino acids in the reduced state of [NiFe] hydrogenase of DvMF (adapted from Fichtner et al, 2006). One CO and two CN^- ligands are bound to the iron atom. Arg479 and Ser502 of the large subunit form hydrogen bonds to one the two CN^- ligands, whereas the CO ligand is connected to Pro501 and Thr87 via van der Waals interactions. A third bridging ligand between Ni and Fe (oxygen- or sulfur-based) atom, which is present in the oxidized enzyme (Volbeda et al, 2005) is missing here.

The proton transfer to solvent is *via* a water molecule and some amino acids with acid-base properties such as glutamic acid, histidine and aspartic acid. Glu-25 of [NiFe] Hydrogenase of Df has been shown to play an important role in fast proton transfer from the active site to the solvent which is responsible for the high turnover rate of hydrogen oxidation (Dementin et al, 2004). Replacement of Glu-25 with glutamine does not affect the spectroscopic properties but cancels the catalytic activity except the para-H_2/ortho-H_2 conversion. Introducing a smaller residue, aspartic acid at Glu-25, does not abolish the catalytic activity completely but slows the transfer significantly.

2.4.2 Structure of [NiFeSe] hydrogenase

Besides the genes identified and sequenced for the [NiFeSe] hydrogenase in this work, the presence of this hydrogenase has been indicated in only three other organisms, DvH, *D. desulfuricans* and *D. baculatus*. A crystal structure from the [NiFeSe] hydrogenase of *D. baculatus* is available. Structurally this hydrogenase is quite similar to the NiFe hydrogenase of DvMF and *D. gigas*. The differences lie in the catalytic site of the large subunit and the medial [FeS] clusters it possesses in its small subunit.

In the [NiFeSe] hydrogenase of *D. baculatus*, a selenocysteine replaces one of the usually conserved cysteines co-ordinating the nickel atom in the active site of the [NiFe] hydrogenase of *D. gigas* and DvMF. Similarly, it possesses a [Fe_4S_4] medial cluster instead of a [Fe_3S_4] cluster found in the standard [NiFe] hydrogenases (Garcin et al, 1999).

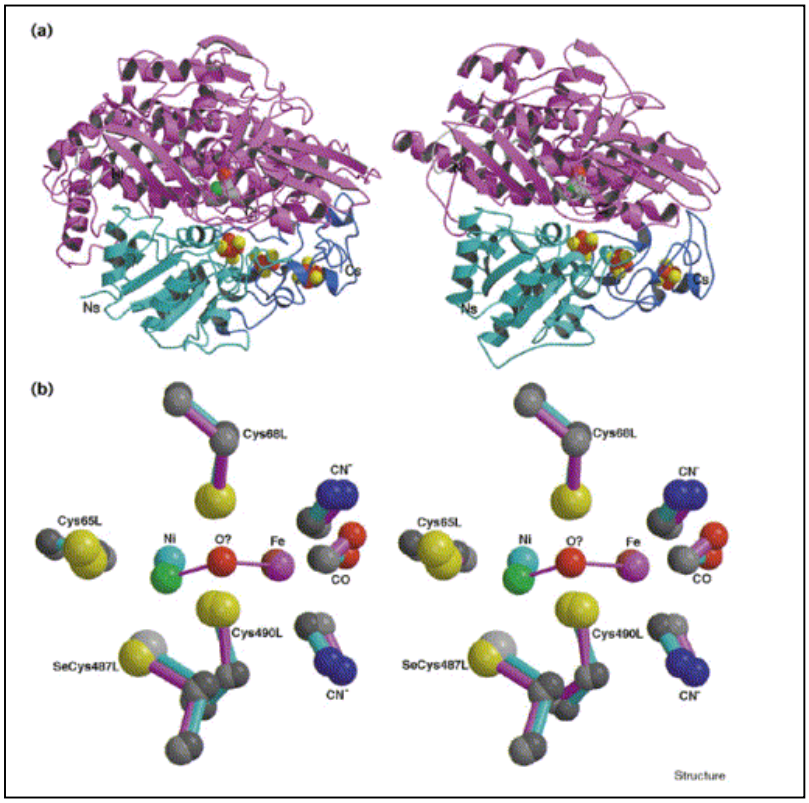


Fig 2.5 Structural comparison of the *D. baculatus* (PDB entery, 1CC1) and *D. gigas* hydrogenases (PDB, 1FRV). (a): 3D Structure. (b): A close-up stereoscopic view of the superposition of the nickel–iron active sites. (Garcin et al, 1999)

In their studies, Valente et al, 2005 have purified [NiFeSe] hydrogenase present in the DvH. The characterization revealed the presence of two isoforms of the detergent soluble [NiFeSe]

hydrogenase, one being cytoplasmic and the other membrane associated. The membrane associated species has a higher activity than the soluble one. It also shows a higher molecular weight (63 kDa for the large subunit and 35 kDa for the small) than the predicted molecular weight (54.4 kDa and 31.4 kDa respectively), owing to post-translational modifications. Other interesting features are the higher catalytic efficiency and lower oxygen sensitivity of [NiFeSe] hydrogenases as compared to [NiFe] hydrogenases. Spectroscopic studies and 3D modeling of [NiFeSe] hydrogenase of DvH have yielded a structural similarity to the [NiFeSe] hydrogenase of *D. baculatum* (Valente et al, 2005). In other studies involving the NiFeSe hydrogenase from DvH, it is predicted to be a lipoprotein (Valente et al, 2007), and selenium availability (Valente et al, 2006) may play significant role in its higher cellular presence and functions, thus dominating the charisma of [NiFe]-and [FeFe]- hydrogenases in the active metabolism of DvH.

2.5 Biosynthesis of [NiFe] Hydrogenases

After the translation of the two polypeptides of the [NiFe] hydrogenases, they undergo a variety of complex post-translational modifications to acquire the three [FeS] clusters, and the insertion of the active site metals Fe and Ni, as well as the two CN^- and one CO ligands of the Fe atom. The insertion of the Fe and Ni atoms and their ligands need the functions of at least five auxiliary proteins namely hydrogenase pleiotropy proteins or simply Hyp proteins (HypA, HypB, HypD, HypE and HypF) and also a C-terminal endopeptidase (targeting the large subunit), a chaperone and several proteins involved in the twin-arginine motif-dependent transport of the hydrogenase into the periplasm. These proteins have been well studied for the maturation of hydrogenase-3 of *E. coli* (see a detailed review, Böck et al, 2006). The Hyp proteins of *Desulfovibrio* are discussed in Chapter 5.

We will take a brief overview on the maturation of the large subunit by the Hyp proteins after its expression. HypF forms a complex with HypD through HypE, which in turn is in close association with HypC (the chaperone). HypF hydrolyzes carbamoyl phosphate and transfers a cyanide ligand to a cysteinyl residue present at the C-terminusof HypE via the formation of a thio-carboximide, while using energy from an ATP molecule hydrolysis (Paschos et al, 2002). HypE dehydrogenates the thio-carboximide bond to form a thiocyanate$^-$ possibly, through a carbamoyladenylate intermediate in an ATP-dependent manner (Blokesch et al, 2004).

The role of HypD in transfer of cyanide to the active site iron atom is speculative, as thiocyanate itself is known to be a good donor of iron (Reissmann et al, 2003). The origin of the CO ligand and the mechanism of its transfer are not yet clear (Forzi and Sawers, 2007). The nickel is transferred into the active site after the insertion of $Fe(CN)_2(CO)$ (Maier and Böck, 1996). HypA, which has zinc binding properties, forms a complex with HypB that is able to hydrolyze GTP for binding and delivering nickel into the active site (Gasper et al, 2006). A peptidyl-prolyl *cis/trans* isomerase SlyD was also shown to aid the kinetics of nickel insertion (Zhang et al, 2005). The nickel insertion is followed by the C-terminal cleavage of the large subunit by the endopeptidase HycI (Menon et al, 1991). This originally present C-terminal extension might act as an endo-chaperone, which maintains the active site in a particular conformation before nickel insertion, and after nickel insertion its removal seals off the active site (Magalon and Böck, 2006).

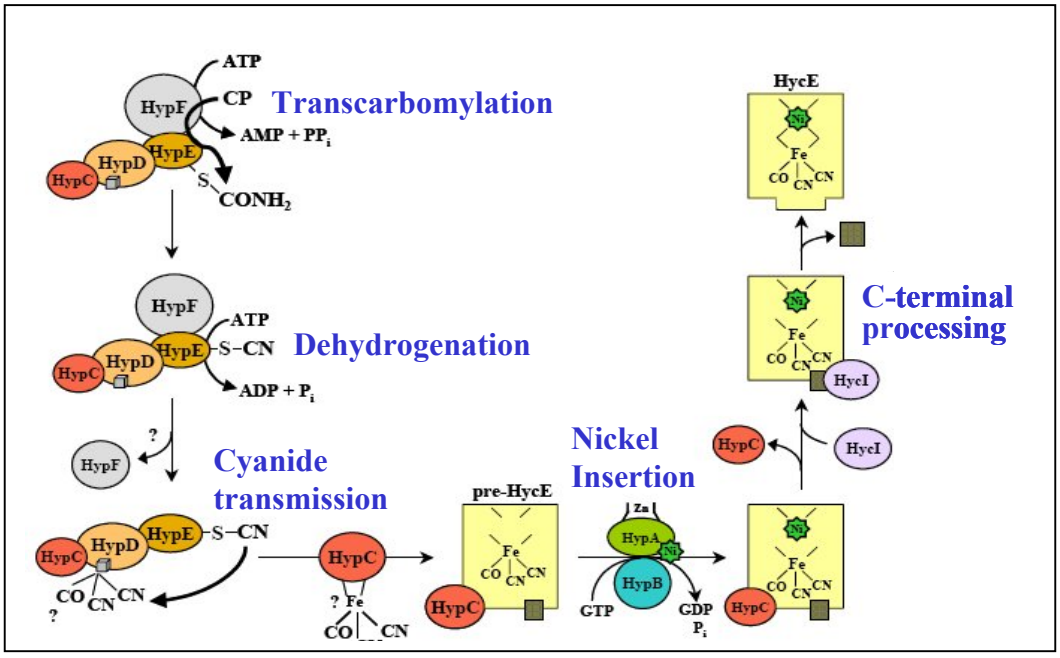


Fig. 2.6 The maturation pathway of the large subunit of hyd-3 from *E. coli*. Modified from the PhD thesis of Melanie Blokesch (2004).

In recent years most of these outlined proteins have been crystallized from various organisms, such as HypB from *Methanocaldococcus jannaschii* (Gasper et al, 2006) and HypC, HypD and HypE from *Thermococcus kodakaraensis* KOD1 (Watanabe et al, 2007). The details obtained by

the available crystal structures of maturation proteins and the structural comparison between these proteins and their functional homologs present in *Desulfovibrio* are discussed in chapter 5.

2.5.2 Membrane Targeting of Hydrogenases

The maturation of the hydrogenases takes place in the cytoplasm and then the two functionally active subunits are transported to the periplasm by a tat (twin-arginine) pathway. Such transport is determined by the presence of a conserved SRRxFLK motif (two arginines) in the N-terminus of the protein to be transported. A hydrophobic region generally follows this twin arginine motif, however, the N-terminal sequence present before this motif can be variable in number and composition. A sequence stretch S_{21}RRDFLK$_{27}$ is also present in the N-terminus of the small subunit of the [NiFe] hydrogenase of DvMF, and similar sequences have been found in this class of hydrogenases (Voordouw, 1992).

The twin-arginine signal peptide was originally found in periplasmic and membrane-bound hydrogenases (Voordouw, 1992, Wu and Mandrand, 1993). It was first reported that the twin-arginine signal peptide of the [NiFe] hydrogenase of *D. vulgaris* is capable of exporting ß-lactamase into the periplasm of *E. coli* (Nivière et al, 1992) and the twin-arginine motif was found to be essential for the translocation. The periplasmic transport is completed by cleavage of this N-terminus at a residue, situated post the hydrophobic region. As a result, the purified hydrogenase of DvMF is shorter by 50 amino acids at the N-terminus as compared to the length expressed from the genome, indicating a mechanistic tat-transport.

The tat-pathway transport of [NiFe] hydrogenase of *E. coli* involves multiple proteins including the proteins that form the pore through which the protein is transported and also some proteins that stabilize the folded hydrogenase (Berks et al, 2005). It is generally believed that a fine tuned interplay between the synthesis, folding, co-factor insertion and translocation ensures that only the catalytically active enzyme is transported to the periplasm (Voordouw, 2000).

A study encompassing the transport of the membrane-bound hydrogenase (MBH) of *Ralstonia eutropha* H16 has shown the elemental role of the tat protein motif in the interaction of the small subunit HoxK with two proteins HoxO and HoxQ, expressed downstream the structural genes

HoxK and HoxG of the MBH hydrogenase (Shubert et al, 2007). It is shown that these two proteins act as chaperones and protect the [FeS] clusters from oxidation inside the cell. However, equivalents of these proteins HoxO and HoxQ have not been found in the *Desulfovibrio* group of bacteria. The oxygen tolerance of *Ralstonia* may play a role in the absolute essentialness of these two proteins.

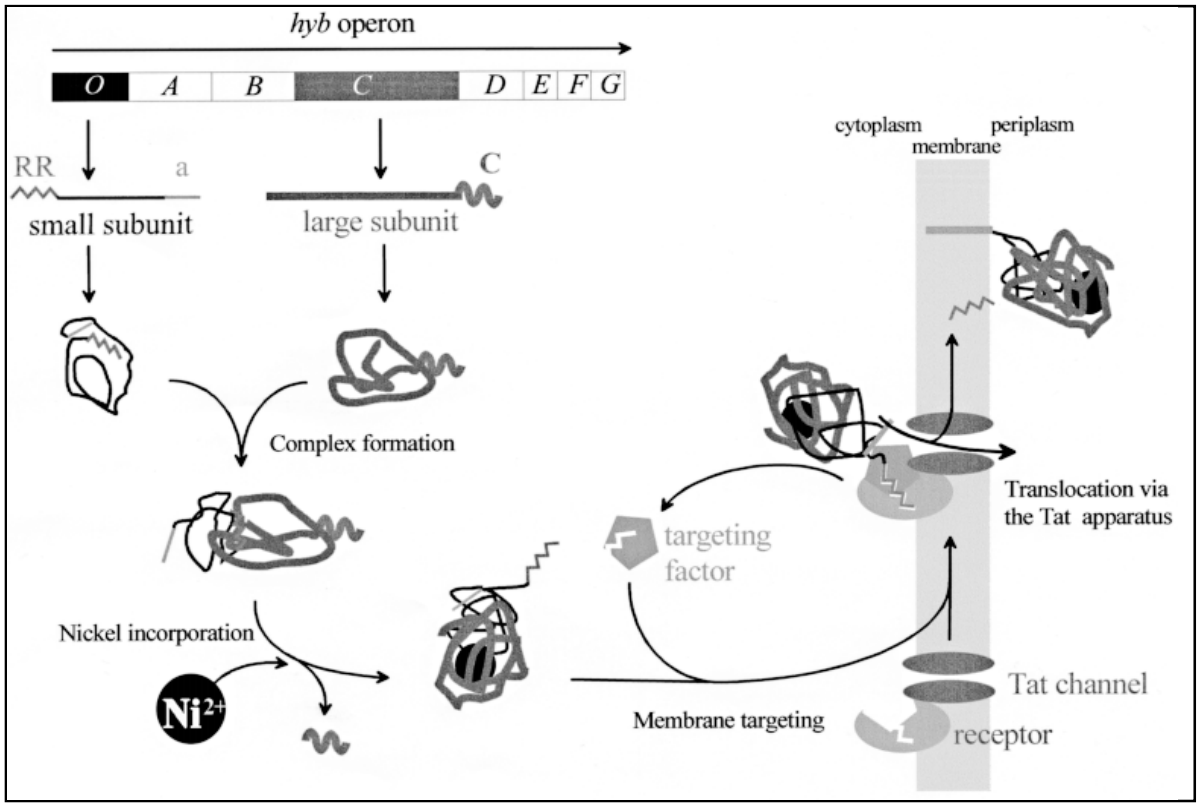


Fig 2.7 Mechanism of "Hitchhiker" co-translocation of the folded [NiFe] hydrogenase subunits in functionally active form across the cytoplasmic membrane. RR indicates the twin-arginine signal peptide, the C'- terminus anchor sequence of the small subunit of hydrogenase 2 is represented by *a* and the C-terminal extension sequence of the large subunit by C. Figure taken from Wu *et al*, 2000.

2.5.3 Requirement for a Heterologous Expression of [NiFe] Hydrogenase

As discussed above, the heterologous expression of [NiFe] hydrogenase requires more than the presence of structural proteins. Attempts to express hydrogenases by passing the structural genes from one organism to other are not generally successful (Goenka et al, 2006), or limited to a narrow range of closely related host cells. A successful expression of a functional hydrogenase of *D. fructosovorans* has been reported by introducing its structural genes into the closely related [NiFe] hydrogenase deletion mutant of *D. gigas* (Rousset et al, 1998).

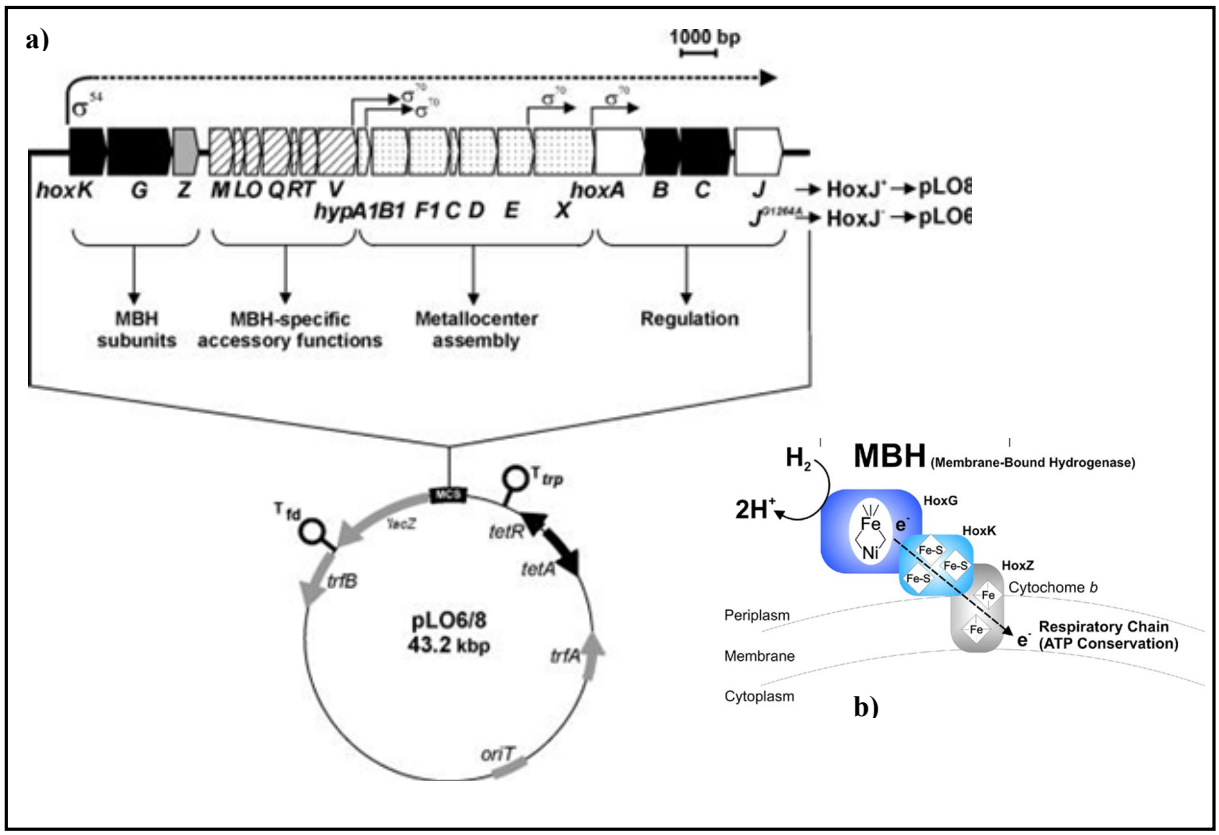


Fig 2.8 a) Map of broad-host-range plasmids pLO6 and pLO8 constucted by Lenz et al, 2005. The MBH operon of *R. eutropha* including the structural subunits and the accessory proteins needed for maturation are indicated T_{fd}, phage fd terminator; T_{trp}, *trpA* terminator; MCS, multiple cloning site. **b)** Cellular location of a functional MBH.

As in case of *Ralstonia eutropha* H16, the genetic information for the whole molecular-machinery needed for the expression of all of the three [NiFe] hydrogenase hydrogenases, i.e., the membrane bound (MBH), soluble (SH) and regulatory (RH) one, is present on one megaplasmid. Taking advantage of these tightly clustered genes, Lenz et al, 2005 have constructed broad-host-range recombinant plasmids, pLO6 and pLO8, which carried the entire membrane-bound hydrogenase (MBH) operon encompassing 21 genes. The pLO6 plasmid was carrying a functional hydrogen responsive operon, whereas pLO8 carried a nonfunctional hydrogen responsive operon due to a point mutation. The introduction of both of these plamids into a megaplasmid-free strain of *Ralstonia* i.e. *R. eutropha* HF210 restored the lithotrophical growth of the cell to normal level in hydrogen supplemented minimal media. Similarly, the

introduction of these constructs led to the detection of methylene blue reducing hydrogenase activity in the cell extracts of an otherwise hydrogenase free host *Pseudomonas stutzeri* (predominately respiratory bacteria).

2.6 Regulation of Hydrogenases

The main role of a hydrogenase is to maintain the bioenergetics of the cells by energy conservation

- By using hydrogen to generate reducing power (sulfate reduction, carbon fixing) and the production of electrons and protons for ATP synthesis.
- While growing in conditions with alternative source of reducing power (growth with lactate/ pyruvate/ acetate), converting this reducing power to hydrogen, thus producing hydrogen.

An investigation encompassing 24 species of *Desulfovibrio* led to invariable positive detection of the genes for [NiFe] hydrogenases, but the presence of other hydrogenases has not been found to be that concurrent (Voordouw et al, 1990). Expression studies with *D. vulgaris* and other microorganisms have recognized a number of regulatory factors triggered by the environment of the cell growth including the availability of the metal ions such as iron, nickel and selenium as well electron acceptors. Comparative growth studies of various knock-out mutants of different periplasmic hydrogenases of DvH have shown that the level of expression of a particular hydrogenase is regulated with regard to the availability and concentration of metal ions and hydrogen in different growth media/conditions (Pereira et al, 2007, Caffrey et al, 2007).

2.6.1 Regulatory Effect of Metals

Whereas the presence of nickel supports the expression of the periplasmic hydrogenases [NiFe] and [NiFeSe], the presence of selenium is involved in the negative regulation of the [NiFe] and [FeFe] hydrogenases in DvH (Valente et al, 2006). It has been shown that in the presences of 1 μM nickel and 1 μM selenium, [NiFeSe] is the dominantly expressed hydrogenase.

2.6.2 Electron Donor and Hydrogen Concentration

In case of DvH, [FeFe] hydrogenase is best expressed in the absence of selenium in a medium with pyruvate and sulfate as electron donor, and is independent of nickel presence. On the other hand, the [NiFe] hydrogenase has a foremost expression in presence of nickel in a hydrogen-sulfate medium (Valente et al, 2006). However, [NiFeSe] remains the dominant hydrogenase in presence of selenium, and is best expressed in a hydrogen-sulfate medium.

In similar studies, Caffrey et al, 2007 have shown that, when selenium and nickel are present in the growth medium, the [NiFeSe] hydrogenase has been found to be enhanced under conditions, where hydrogen is the only electron donor to the cell as compared to growth on lactate medium. This enhancement in the expression is slightly better in an atmosphere, which contains hydrogen present at lower concentration (5% vs 50%). The effect of hydrogen concentration is reverse in case of the [FeFe] hydrogenase, while the pattern of Ech hydrogenase expression is concurrent with the [NiFeSe] hydrogenase (Caffrey et al, 2007, Pereira et al, 2008).

2.6.3 Catalytic Activity

The level of activity varies strongly for different hydrogenases. The highest specific activity (up to 50,000 μM min^{-1} mg $protein^{-1}$) has been reported for the [FeFe] hydrogenase of DvH which also shows the lowest binding affinity (K_m= 100 μM) for hydrogen uptake, whereas the [NiFe] hydrogenase and the [NiFeSe] hydrogenases have somewhat lower specific activities (100 to 1000 μM min^{-1} mg $protein^{-1}$) and higher binding affinities (K_m =1 μM). Furthermore, the susceptibility of different hydrogenases to various inhibitors such as CO, NO and NO_2^- is different from each other. Here again the differential affinity and activity of various hydrogenases and their corresponding advantage in a stable or changing ecosystem cannot be overemphasized (Pereira et al, 2007; Caffrey et al, 2007).

Catalytic activities of Hases (in U/mg)							
DvH [NiFeSe] Hase		DvH [NiFe]1 Hase	DvH [FeFe] Hase	Dg [NiFe] Hase	Dmb [NiFese] Hase	Dmb NJ [NiFeSe] Hase	DvMF
H_2 production	6,908	174	4800	440	2,000	8,600	-
H_2 consumption	900	89	50,000	1500	-	-	-

Table 2.1 A collection of data of hydrogen production and consumption by various hydrogenases, the – sign means the activities that are not measured (Valente et al, 2005).

2.7 Sulfate Reduction in *Desulfovibrio vulgaris*

Sulfate reduction is one of the earliest types of energy metabolism used by ancestral organisms to sustain life. Despite extensive studies, the exact coordination of the respiratory sulfate reduction with energy conservation is yet to be elucidated. The *Desulfovibrio* species perform a dissmilatory sulfate reduction, which is understood as the reduction of sulfur from a higher oxidation state to lower. Dissimilatory sulfate reduction operates under strictly anaerobic conditions and represents an important element within the sulfur cycle on earth (Hansen, 1994, LeGall and Fauque, 1988).

Desulfovibrio can derive energy from the dissimilatory reduction of sulfate by using hydrogen or organic substrates as electron donors. The electron transport chain catalyzing this reaction involves periplasmic hydrogenases, several multi-heme cytochromes, such as cytochrome *c*3 and nonhaeme *c,* and other membrane-bound and cytoplasmic redox enzymes.

The pathway of this dissimilatory sulfate reduction at least involves four key enzymes: ATP sulfurylase, inorganic pyrophosphatase, adenylylsulfate reductase and sulfite reductase (Hansen, 1994). ATP sulfurylase utilizes two molecules of ATP to activate sulfate (-516 mV) to adenosine-5'-phosphosulfate (APS). This compound has a much lower negative redox potential (-60mV) and is better suited as a catabolic substrate for further reduction.

$$SO_4^{2-} + ATP \rightarrow APS^{2-} + PPi$$

$$PPi \rightarrow 2Pi$$

APS is the first electron acceptor of the respiratory chain and is reduced by APS reductase to sulfite by a two-electron reduction (Lampreia et al, 1990).

$$APS + 2e^- \rightarrow SO_3^{2-} + AMP$$

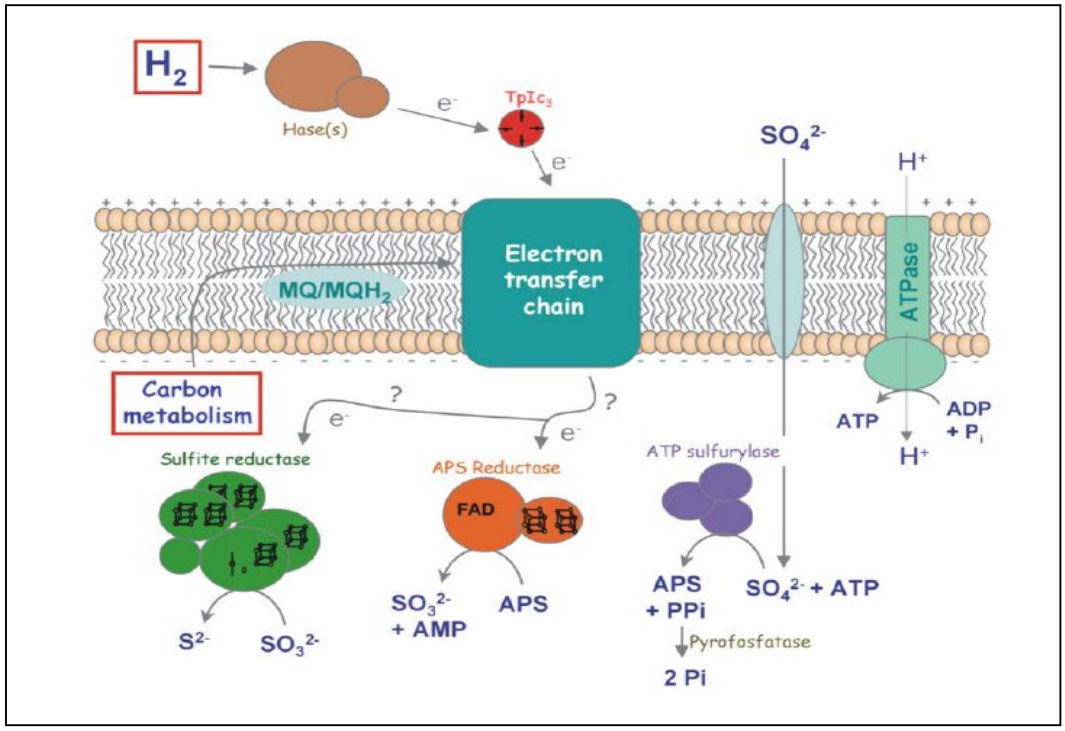


Fig 2.9 Schematic representation of the respiratory electron transfer chain in *Desulfovibrio*, with H_2 or organic compounds as energy source and sulfate as terminal electron acceptor. (Matias et al, 2002)

The dissimilatory sulfite reductase, also known as desulfoviridin or Dsr, further reduces sulfite to sulfide in a six-electron reduction reaction (Odom and Peck, 1984).

$$HSO_3^- + 6e^- + 6H^+ \rightarrow HS^- + 3H_2O$$

APS reductase and desulfoviridin are two iron sulfur protein complexes, which are present in the cytoplasm of the cell and accept the electrons from the electron transport chains located in the cell membrane.

2.7.1 APS Reductase

APS reductase (5'- adenylyl sulfate reductase) in general is located in the cytoplasm of several sulfate-reducing and sulfur-oxidizing organisms. It is an iron-sulfur cluster-associated flavoprotein, which catalyzes both the two-electron reduction of APS to sulfite and oxidation of sulfite and AMP to APS. The adenylylsulfate reductase from *Desulfovibrio vulgaris* Miyazaki F has been previously isolated and biochemically characterized (Yagi and Ogata, 1996). Catalytically, it mediates the transfer of two electrons from dihydroflavin coenzymes ($FADH_2$, $FMNH_2$ or dihydroflavin) to adenylyl sulfate (APS), and catalyzes flavin-mediated oxidation of ferrocytochrome c3 with APS. The APS reductase is not inhibited by ATP or GTP, but shows diminished acivity in the presence of AMP and sulfite. The low K_m for APS reduction (< 1μM) enables it to effectively reduce cytosolic concentrations of APS (usually in micromolar concentration). Furthermore, the inhibitory effect of sulfite enables DvMF to utilize an energetically favorable electron acceptor i.e. sulfite preferentially over APS which is produced from sulfate at the cost of ATP.

Structurally the APS reductase is a conserved protein. A crystal structure from the thermophilic aracheon *Archaeoglobus fulgidus* has been produced and a reaction mechanism was proposed (Fritz et al, 2002). The functionally active enzyme is a heterodimeric complex composed of an α-subunit with a molecular mass of 70-75 kDa and a β-subunit of 18-23 kDa. A comparative overview of the respective 3D structures of *Archaeoglobus fulgidus* and DvMF will be given in chapter 5.

2.7.2 Dissimilatory Sulfite Reductase

Sulfite reductases occur in diverse bacterial and archaeal lineages. Sulfite reductases are divided in two major classes depending on their functions, one is the assimilatory ones where these reductases help to assimilate the sulfur into the cellular components and second one are the dissmilatory reductases which allow the removal of the sulfur from the cell while combining with the cellular metabolism. Sulfite reductases are present in both sulfur oxidizing and reducing organisms. The assimilatory sulfite reductases are denoted as Asr or as aSiR and the dissimilatory reductase as Dsr or dSiR. Both of these types share a highly conserved domain 'C-

X_5-C-n-C-X_3-C' for binding sirohemes (Fig 2.10) and a ferredoxin, i.e., an iron sulfur cluster containing region that facilitates electron transfer to the substrate. The siroheme is the site of the catalytic reduction. Structurally, a siroheme is a reduced porphyrin of the isobacteriochlorin class. It forms covalent bonds with an iron sulfur cluster of the protein. The Dsr are multisubunit enzymes (167 to 225 kDa) that catalyze the six-electron reduction of sulfite to sulfide. They contain a siroheme and [Fe_4S_4] prosthetic centers and are classified according to their spectroscopic properties in four major groups (Lee and Peck, 1974) those are desulfoviridin, desulforubidin, *P*-582, and desulfofusicidin.

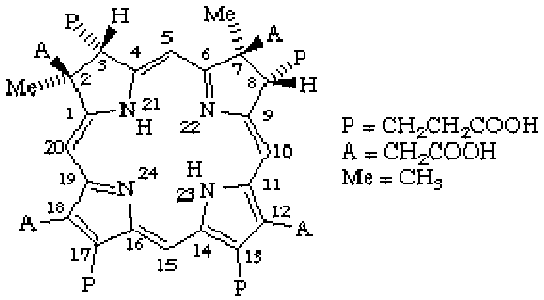


Fig 2.10 A heme-like prosthetic group found in a class of enzymes that catalyze the six-electron reduction of sulfite and nitrite to sulfide and ammonia, respectively.

An $\alpha_2\beta_2\gamma_2$ structure is proposed for the Dsr from *Desulfovibrio vulgaris* (Hildenborough), *D. vulgaris oxamicus* (Monticello), *D. gigas*, and *D. desulfuricans* ATCC 27774 (Pierik et al, 1974). The α and β subunits are coded by a conserved three gene operon *DsrABD* and the gene encoding the γ subunit is located separately in the genome. Recently a crystal structure of the DsrAB complexed with DsrC from DvH has been reported (Oliveira et al, 2008). The structure obtained has elucidated an $\alpha_2\beta_2\gamma_2$ assembly in which each monomer of DsrB and DsrA contains two [Fe_4S_4] clusters and one heme ring. The heme rings present in DsrB is a siroheme coupled with one of the [Fe_4S_4] clusters and acting as the catalytically active site. The heme ring present in DsrA is a de-metalled sirohydrochlorin associated with one of the [Fe_4S_4] clusters (Oliveira et al, 2008).

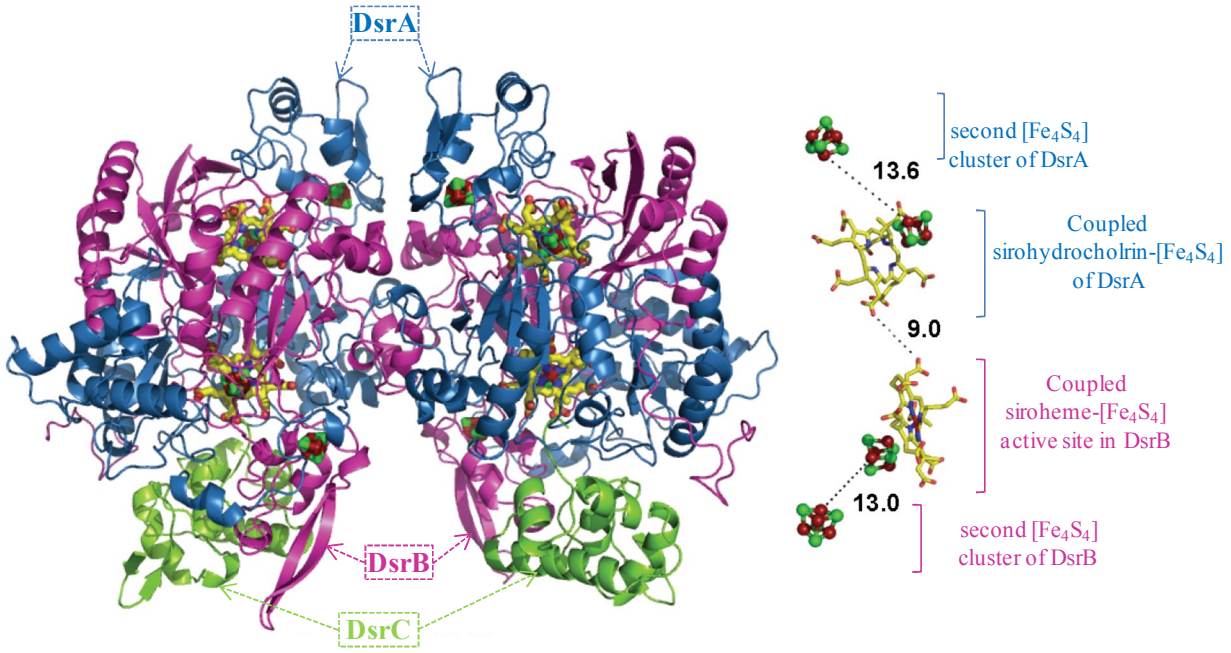


Fig 2.11 Schematic three dimensional model representation of the $\alpha_2\beta_2\gamma_2$ assembly of the DsrAB sulfitre reductase bound to DsrC, with the cofactors in ball and stock mode (figure here taken from Oliveira et al, 2008). DsrA is colored in blue, DsrB in magenta and DsrC in green. On the right side, the 3D distances between the cofactors of a single αβγ- unit are displayed.

A crystal structure from the desulfoviridin has been also reported from *Archaeoglobus fulgidus* (Schiffer et al, 2008), which shows that this protein is organized in a heterotetrameric $(\alpha\beta)_2$ complex composed of two catalytically independent αβ heterodimers. Since the desulfoviridin of DvMF is closely related to DvH and has been crystallized in a $\alpha_2\beta_2\gamma_2$ stoichiometery, it is likely that it is quite similar in its three dimensional structure as well.

2.8 Scope of this Thesis

In recent times, *Desulfovibrio* sps have become model organisms to study the expression and functional modulation of the hydrogenases. Their large number of hydrogenases and the availability of genomic and micro-array data have made these organisms an immense source of information. In times of shortage and high prices of fossil fuel, the quest of using a hydrogenase-based fuel cell in hydrogen fueled automotives and machineries has started a talented hunt for various hydrogenase candidates. Furthermore, the likelihood of using these organisms in bio-remediation efforts has raised their economic value. However, these organisms

are difficult to manipulate genetically because of their strict anaerobicity, uneven plating efficiencies, insufficiently characterized genetics, and inherently high antibiotic resistance.

On the other hand, *E. coli* and microaerobic organisms as *Ralstonia eutopha*, are easier to grow and manipulate and are genetically accessible. Thus, most of the maturation studies on hydrogenase have been carried in these two organisms. Nevertheless, efforts are continued to understand the genetics of *Desulfovibrio*, with the aim to manipulate it in a more useable manner.

Desulfovibrio vulgaris Miyazaki F is an organism, which has been extensively studied for the attempts to express large amounts of [NiFe] hydrogenase, for which detailed structural and spectroscopical data are available. Our effort is directed to proceed in the direction of understanding its genetics so that an adaptable hydrogenase expression system can be developed in the future for its heterologous expression. As to understand the stereochemistry and activity, expression and maturation network, this enzyme is planned to be expressed in large amounts in its native and mutant composition. However, its heterologous expression requires many details to be considered.

- The cellular expression and activity of this hydrogenase is oxygen sensitive.
- There is an array of maturation genes involved to bring the enzyme in a physiologically active state.
- The maturation genes are organism-specific and heterologous complementation has so far not been a convincing method of good expression.

Unlike *Ralstonia* the hydrogenase machinery is not clustered in any of the known genomes of *Desulfovibrio* species, but is scattered throughout the genome. No hydrogenase-related gene was detected in downstream sequencing of the *hyn* operon of DvMF (Goenka et al, 2005). Starting with this work, there was a need to search the equivalent genes needed for the proteins involved in the maturation of [NiFe] hydrogenase and also in the successful heterologous expression as in the case of *Ralstonia eutropha* H16 (Lenz et al, 2005).

The two structural genes for the [NiFe] hydrogenases of DvMF were first sequenced quite a while ago (Deckers et al, 1990), and were found to be quite similar to those from other sulfate reducing bacteria. In their work, restriction-digests of chromosomal DNA of DvMF were hybridized with [α-^{32}P] labeled probes derived from DvH. The positive signals were cloned and sequenced. Similar hybridization experiments for the iron hydrogenase pointed to an absence of an iron-only enzyme in DvMF (Voordouw et al, 1990). However, not only the complex maturation pathways, but also the fact that genomic sequences of any major sulfate reducing bacteria were not available, has strongly impaired the advancement of such studies. Even after crystallization and extensive spectroscopic studies, for a long time the molecular genetics of DvMF hydrogenase had stayed untouched, until two accessory maturation genes C- terminal peptidase, *hynC* and the chaperone *hynD*, present downstream directly following the structural genes, were further sequenced (Agrawal et al, 2006).

The comparisons of the available hydrogenase- and sulfate metabolism-related genes from several organisms have always revealed a high level of conservation among various *Desulfovibrio* sps. In 2004 the genome sequences from DvH and *Desulfovibrio desulfuricans* subsp G20 had been published. It provided us with a reliable framework to identify and then clone such genes out of the DvMF genome in absence of completed genomic data.

In this work, the genome of DvMF has been explored for specific genes by means of constructing a cosmid genomic library. This library has been screened for the presence of other hydrogenases genes as well as their possible *cis/trans* acting maturation genes. We are reporting and characterizing the sequences and expression of two hydrogenases for the first time, those are, a heterodimeric selenocysteine-containing [NiFeSe] hydrogenase denominated as HysAB and a six-subunit complex, Energy Converting Hydrogenase denominated as Ech. Also, the *hyp* maturation genes involved in the maturation of the standard [NiFe] hydrogenase (hydrogenase pleiotropy; *hypA*, *hypB*, *hypD*, *hypE* and *hypF*) have been sequenced and expressed in *E. coli* from their natural operons. A heterologous expression of the *hynABCD* operon in the [NiFe] hydrogenase deletion mutant (Goenka et al, 2005) has been also attempted. Lastly, the genomic characterization of the two sulfate reductase proteins, i.e., adenylyl sulfate reductase and

desulfovirdin has been performed together with their purification (partial collaboration with Dr. Aruna Goenka Agrawal) and crystallization (in collaboration with Dr. Hideaki Ogata).

Chapter 3

Materials and Methods

3.1 Bacterial strains, plasmids and vectors

Desulfovibrio vulgaris* subsp. *vulgaris

Miyazaki F	IAM 12604.
Hildenborough	NCIMB 8303; isolated from clay soil near Hildenborough, United Kingdom, source of the [NiFe] hydrogenase gene used here; Cm^S, Suc^R (Postgate, 1984).
Hyn 100	A [NiFe] hydrogenase gene deletion derivative of Hildenborough. The hydrogenase gene was interrupted by inserting a chloroamphenicol gene cassette from pUC19Cm; Km^R, Cm^R, Suc^R (Goenka et al, 2005).

***D. desulfuricans* subsp. *desulfuricans* ATCC 7757**

Escherichia coli

E. coli S17-1	*thi pro hsdR hsdM*$^+$ *recA* RP4-2 (Tc::Mu Km::Tn*7*) (Simon *et al*, 1983)
E. coli DH5α	*supE44 dlacU169* (*p80 lacZ*dM15*) hsdR17 recA1 endA1 gyrA96 thi-1 relA1* (Hanahan, 1985).
E. coli EPI100-$T1^R$	*F- mcrA D(mrr-hsdRMS-mcrBC) f80dlacZDM15 DlacX74 recA1endA1 araD139 D(ara, leu)7697 galU galK l- rpsL nupG*
E. coli XL1-Blue	*endA1 gyrA96(nal*^{R}*) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB*$^+$ *lacI*^{q} *Δ(lacZ)M15] hsdR17(*r_K^- m_K^+*)*
E .coli (BL21DE3RIL)	F^- *ompT hsdS* ($r_B^-m_B^-$) *dcm*$^+$ Tet^r *gal endA* Hte [*argU, ileY, leuW*, Cam^r].

3.2 Plasmids

pJRD215hynBACD	The *hynBACD* operon cloned into the SmaI cut pJRD215 (Davison et al, 1987) . Kanamycin and streptomycin resistance.
pUC19nptII	The SacII fragment of pBSL180 (Alexeyev and Shokolenko, 1995) containing the nptII (neomycin phosphotransferase) gene cloned into the SmaI cut pUC19. Ampicillin and kanamycin resistance.
pMOShynBACD	The HindIII flanked *hynBACD* operon cloned into the EcoRV site of pMOS vector
pMOShynBACDstrep	A strep tag followed by an XhoI site is inserted by PCR at the C-terminal of *hynB* of pMOShynBACD. Ampicillin resistance.
pMOShynBACDstrepNptII	The XbaI cut fragment of the puc19nptII is cloned into the XbaI-SpeI cut pMOShynBACDstrep vector. Ampicillin and kanamycin resistance.
pRSFhynBACDstrep	The reading frames of strep-tagged *hynBACD* operon cloned into the LIC (ligation independent cloning) compatible site of the T4 polymerase treated pRSF-2 Ek/LIC. *hynB* is also in frame with a N-terminal hexa-histidine tag and the *hynD* is in frame with a C-terminal S-tag. Kanamycin resistance.
pRSFhypAB	The reading frames of maturation genes *hypAB* operon cloned into the the LIC compatible site of the T4 polymerase treated pRSF-2 Ek/LIC. *hypA* is also in frame with an N-terminal hexa-histidine tag and the *hypB* is in frame with a C-terminal S-tag. Kanamycin resistance.
pRSFhypDE	The reading frames of maturation genes *hypDE* operon of DvMF cloned into the LIC compatible site of the T4 polymerase treated pRSF-2 Ek/LIC. *hypD* is in frame with a N-terminal hexa-histidine tag and the *hypE* is in frame with a C-terminal S-tag. Kanamycin resistance.

pRSFhysBA	The reading frames of the structural genes of [NiFeSe] hydrogenase *hysBA* operon of DvMF cloned into LIC compatible site of the T4 polymerase treated pRSF-2 Ek/LIC. *hysB* is in frame with a N-terminal hexahistidine tag and the *hysA* is in frame with a C-terminal S-tag. Kanamycin resistance.
pRSFhysA	The reading frame of the large subunit of the [NiFeSe] hydrogenase of DvMF cloned into LIC compatible site of the T4 polymerase treated pRSF-2 Ek/LIC. *hysA* is in frame with both of a N-terminal hexa-histidine tag and a C-terminal S-tag. Kanamycin resistance.
pRSFhysApreseleno	*hysA* reading frame before th selenocysteine coding TGC, was cloned into pRSF-2 Ek/LIC in frame with both a N-terminal hexa-histidine tag and a C-terminal S-tag.

3.3 Chemicals and Enzymes

3.3.1 Chemicals

All chemicals and reagents used throughout this work were of analytical grade or of maximal purity available and were obtained from Merck (Darmstadt), Sigma (Deisenhofen), Serva (Heidelberg), DIFCO (Detroit, USA), Gibco/BRL (Eggenstein), Biomol (Hamburg), Pharmacia (Freiburg), Roth (Karlsruhe, Germany), ICN Biomedicals (Aurora, USA) and BioRad (München). Chelex resin (used for metal exchange/removal from buffer and media for minimal medium preparation for labelling experiments) was from Biorad.

3.3.2 DNA modifying Enzymes

Restriction- and other DNA-modifying enzymes were from Fermentas (St. Leon-Rot), Stratagene (Heidelberg), GE Healthcare (München), Novagen from Merck Biosciences (Darmstadt) and NEB (New England Biolabs) (Schwalbach/Taunus). Oligonucleotides were custom-synthesized by Metabion (Martinsried).

3.3.3 Antibodies

Anti [NiFe] hydrogenase of DvMF	Polyclonal rabbit serum, custom synthesis by Eurogentec (Belgium).
Anti [NiFeSe] hydrogenase of DvH	Polyclonal rabbit serum, custom synthesis by Eurogentec (Belgium). Protein was donated by Dr. Inês Pereira (ITQB, Portugal)
Anti Ech hydrogenase of *M. barkeri*	Polyclonal rabbit serums specific to the was provided by Prof. R. Heddrich (Max Planck Institute for Terrestrial Microbiology, Marburg)
Anti- Histidine	Invitrogen
Anti- Strep tag	Novagen, Merck Biosciences
Anti- S-Tag	Novagen, Merck Biosciences
AP conjugated anti- DIG	Roche Applied Sciences
AP conjugated anti-mouse immunoglobulins	DAKO (Hamburg)
AP conjugated anti-rabbit immunoglobulins	DAKO (Hamburg)

3.3.4 Kits

Bradford assay kit	Biorad
Dneasy® Tissue kit	QIAGEN
GFX PCR DNA and Gel Band Purification Kit	GE Healthcare
Illustra PlasmidPrep Mini Spin	GE Healthcare
pMOS Blue Blunt Ended Cloning Kit	GE Healthcare
pRSF-2 Ek/LIC Cloning Kit	Novagen, Merck
pWEB-TNC™ Cosmid Cloning Kit	Epicentre, Biozym
QIAfilter Plasmid Maxi Kit	QIAGEN
QIAprep 96 Turbo Miniprep Kit	QIAGEN

3.4 Media and Solutions

The broths were prepared by mixing the contents to maximum homogeneity and sterilized by steam autoclaving in loose cap bottles or flasks at 121°C and 15 psi of steam pressure for 15 minutes. Millipore distilled threefold deionised water was used for preparing media and buffers. For the minimal medium preparation for *Desulfovibrio* cultures, chelex-cleaned-water was prepared by using a chelex resin column. Chelex resin binds metal ions selectively and thus cleans the solution of any metal contamination.

The autoclaved solid broths were poured in capped flat plastic plates (also called Petri dishes) to solidify and dry. The sterile growth media were allowed to cool to ambient temperatures before inoculating, streaking or plating. Filtration, if necessary, was done through 0.22 μm filter, to sterilise heat labile components.

3.4.1 *E. coli* Growth Medium

LB: 10 g/l NaCl, 5 g/l yeast extract 10 g/l tryptone

LB-agar: 10 g/l NaCl, 5 g/l yeast extract 10 g/l tryptone, 15 g/l agar

TB: 16 g/l Tryptone, 10 g/l yeast extract, 5 g/l NaCl and 100 ml of sterile 200 mM potassium phosphate buffer added after autoclaving the broth.

2xYT 16 g/l Tryptone, 10 g/l yeast extract, 5 g/l NaCl

SOC: 20 g/l Tryptone, 5 g/l yeast extract, 1% (v/v) 1 M NaCl, 0.25% (v/v) 1 M KCl, 1% (v/v) filter sterilized 2 M Mg^{2+} stock (1 M $MgCl_2$; 1 M $MgSO_4$), 1% (v/v) filter sterilized 2 M glucose stock. The filter-sterilized components of the SOC are added separately to the autoclaved broth prepared by mixing rest of the components.

3.4.2 *Desulfovibrio* Growth Medium

PB 0.5 g/l KH_2PO_4, 1 g/l NH_4Cl, 1 g/l $CaSO_4$, 2 g/l $MgSO_4.7\ H_2O$, 3.5 g/l sodium lactate (70%), 1 g/l yeast extract, 0.1 g/l ascorbic acid, 0.1 g/l thioglycolate, 0.5 g/l $FeSO_4.7\ H_2O$ (adjusted to pH 7.0-7.4)

Purpose: preparation of stock cultures.

PC 0.5 g/l KH_2PO_4, 1 g/l NH_4Cl, 4.5 g/l Na_2SO_4, 0.06 g/l $CaCl_2.6\ H_2O$, 0.06 g/l $MgSO_4.7\ H_2O$, 6 g/l sodium lactate (70%), 1 g/l yeast extract, 0.004 g/l $FeSO_4.7\ H_2O$, 0.3 g/l sodium citrate (adjusted to pH 7.5)

Purpose: routine DvMF culture growth broth.

PE 0.5 g/l KH_2PO_4, 1 g/l NH_4Cl, 1 g/l Na_2SO_4, 1 g/l $CaCl_2.6\ H_2O$, 2 g/l $MgCl_2.7\ H_2O$, 3.5 g/l sodium lactate (70%), 1 g/l yeast extract, 0.5 g/l $FeSO_4.7\ H_2O$, 15 g/l agar (adjusted to pH 7.6)

Purpose: routine culture growth plates.

PE/KNO_3 500 ml PE, 1.5 g KNO_3, 7.5 g agar

Purpose: plates for conjugation.

FM 4 g/l Na_2SO_4, 0.6 g/l $MgSO_4.7\ H_2O$, 0.4 g/l KH_2PO_4, 3 g/l Tryptone, 0.8 g/l yeast extract, 10 ml sodium lactate (50%), 0.01 g/l $FeSO_4.7\ H_2O$, 5 ml 10% antifoam to 8 L media

Purpose: Fermentation medium for protein extraction.

3.4.3 Antibiotic Stock Solutions

All water-soluble antibiotics were filter sterilized.

Ampicillin	25 mg/ml in H_2O
Chloramphenicol	20 mg/ml in ethanol
Kanamycin	10 mg/ml in H_2O
Streptomycin	10 mg/ml in H_2O

3.4.4 Southern Blot

Depurination solution	0.25 M HCl
Denaturation solution	0.5 M NaOH, 1.5 M NaCl
Neutralization solution	0.5 M Tris-HCl, 1.5 M NaCl ; pH 7.5
20xSSC	0.3 M sodium citrate, 3 M NaCl ; pH 7.0 (autoclaved)

Hybridization buffer	5x SSC, 0.1% N-lauroylsarcosine, 0.02% SDS, 1% blocking powder (Roche), 50% de-ionized formamide
Low stringency wash buffer	2x SSC, 0.1% SDS
High stringency wash buffer	0.5x SSC, 0.1% SDS
1x Maleic Acid buffer	100 mM maleic acid, 150 mM NaCl; pH 7.5

3.4.5 SDS-PAGE and Western Blot

2x Laemmli Buffer	100 mM Tris-HCl (pH 6.8), 200 mM DTT, 4 % (w/v) SDS, 0.2 % (w/v) Bromophenol Blue, 20 % (v/v) glycerol
Coomassie Stock	1 tablet of Phastgel blue R in 200 ml 60 % (v/v) methanol (gives a 0.2% Coomassie blue R solution)
Coomassie staining	1 volume Coomassie stock, 1 volume 20 % acetic acid
Transfer Buffer	25 mM Tris base, 150 mM glycine; pH 7.5, 20 % methanol
TBS	10 mM Tris-HCl, 150 mM NaCl; pH 7.5
TBST	20 mM Tris-HCl, 500 mM NaCl; pH 7.5, 0.05 % (v/v) Tween-20
Blocking Buffer	3 % (w/v) BSA in TBS buffer
NBT-BCIP	One Sigma-FAST BCIP-NBT tablet in 10ml H_2O or 200 µl of NBT-BCIP solution (Roche) in 10 ml detection buffer.
Detection buffer	100 mM Tris buffer, pH 9.5; supplemented with 50 mM $MgSO_4$ in case of protein detection

3.4.6 Native PAGE

TS buffer	150 mM NaCl, 10 mM Tris-HCl; pH 7.5, 10 % glycerol, 1% Tween 100
Phosphate buffer	50 mM Phosphate buffer (pH 7.0)
Triton detergent	2% Triton X-117
Loading buffer (5X)	60 mM Tris-HCl (pH 6.8), 25% glycerol, 0.1% Bromophenol blue

3.4.7 Hydrogenase activity assay

PMS	0.09 mM Phenazine methosulfate
NBT	0.06 mM Nitroblue Tetrazolium
BV	0.5 mM Benzyl viologen
MV	1 mM Methyl viologen
DTT	60 mM Sodium dithionite/ Sodium hydrosulfite
TTC	1 mM Triphenyl tetrazolium chloride

3.4.8 Electrochemical analysis buffers

For H_2 uptake	100 mM Tris-HCl (pH 8.5), 0.1 mM MV
For H_2 production	100 mM sodium acetate (pH 5.6), 0.1 mM MV

3.4.9 Other Solutions

P1	50 mM Tris-HCl pH 8.0, 10 mM EDTA, 100 µg/ml RnaseA
P2	0.2 M NaOH, 1%SDS
P3	60% (V/V) 5 M potassium acetate, 11.5% acetic acid
Phenol Chloroform	50 volumes of phenol, 50 volumes of chloroform
EP	0.5 mM HEPES (pH 7.6), 10 % glycerol
6x Loading Buffer	0.25 % (w/v) bromophenol blue, 0.25 % (w/v) xylene cyanol FF, 30% (v/v) glycerol
TE	10 mM Tris-HCl (pH 8.0), 1 mM EDTA
TAE	50 mM Tris acetate, 1 mM EDTA (pH 8.0)
Agarose Gel Staining	0.01% Ethidium bromide in 1x TAE

3.5 Cell Growth

3.5.1 *E. coli*

Depending on purpose, the bacteria were grown on suitable liquid broths or solid agar plates with appropriate antibiotic concentrations, wherever was applicable as follows: penicillin 100 µg/ml, chloramphenicol 33 µg/ml, kanamycin 50 µg/ml, streptomycin 50 µg/ml, tetracycline 20 µg/ml. The growth temperature was in general maintained at 37°C, and cultures growing in liquid medium were kept shaking at 180 rpm.

The temporary storage *E. coli* cells grown in colonies on solid agar plates were kept at -4°C up to 4 weeks. For long time storage, an overnight grown culture of *E. coli* was mixed with 1/15 parts of sterile glycerol or 1/7 parts of DMSO and stored at -80°C.

3.5.2 *Desulfovibrio* sps.

Fermentation The *Desulfovibrio* cells were at first grown by three steps, starting from the stock cultures, grown into 10 ml and 50 ml glass vials and then the cells were transferred into 500 ml screw-capped glass bottles; the centre of the lids was replaced by a septum. All inoculations were done in anaerobic glove boxes. For 10 L growth cultures (using a 10 L glass fermenter), three 500 ml cultures were grown for two days, i.e., the growth was extended almost to the end of the log phase. Such cultures were then used to inoculate a 10 Litres fermentor, containing 8 Litres of PC or FM growth medium. Under these conditions, the cells were grown for three days.

For temporary stocks, the cells were grown in PB medium in air-tight soft cork capped 5 ml glass vials. These cultures were stable up to 2 months at 4° C. For long time storage at -80°C, fully grown cells were centrifuged and re-suspended into sterile degassed glycerol (1/100 parts of original growth).

Anaerobic conditions were maintained by constantly bubbling nitrogen into the fermentor; 1 M H_2SO_4 was used to maintain the pH during growth at around 7.0-7.4. During the three days growth, 400 ml of sterilized sodium lactate (25% w/v in chelexed water) were added into

growing cultures on the second and the third day morning. The cells were harvested by centrifugation at 6000 rpm, 4°C, and 30 minutes. Usually, a wet cell-yield of 15-20 g per 10 L fermentor was obtained. Cells were kept frozen at –80°C, until lysed for protein purification. 60-80 g wet cell mass were used for one protein purification.

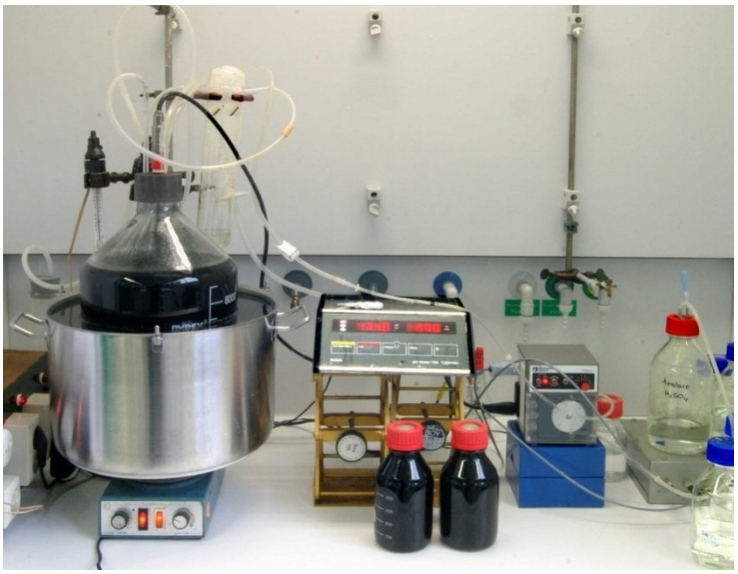


Fig 3.1 Home built 10-litre glass fermentor. The growing culture is kept anaerobic by bubbling nitrogen constantly and the temperature is maintained at 37°C using a water bath. 1M H_2SO_4 is fed via a manually controlled pump to maintain the pH in a range of 7.0 to 7.4.

Growth on plates The isolated colonies were grown on Postgate E medium and occasionally supplemented with potassium nitrate when mixed with *E. coli* under anaerobic conditions.

3.5.3. Conjugation

The *Desulfovibrio* sps cells were conjugated with *E. coli* S17-1 cells on PE plates supplied with nitrate under anaerobic conditions. Plasmid bearing RSF1010 replicon are capable of replication in both *E. coli* and *Desulfovibrio* sps. The plasmid pJRD215 possesses RSF1010 replicon along with two antibiotic resistance markers, kanamycin and streptomycin. To transfer the genes cloned in pJRD215 (chapter 5) from *E. coli* to *Desulfovibrio*, a filter-mating method was used for conjugation. Briefly, 1 ml of an O/N culture of *D. vulgaris* in medium C was mixed with 0.2 ml of an O/N culture of the *E. coli* donor cells in a 1.5 ml microfuge tube under anaerobic conditions (work performed inside the anaerobic hood). The cell mixture was spun down,

resuspended in ~20 µl medium C, spread onto a membrane filter (Millipore, 0.22 µm, Ø =25 mm) placed on a Medium E-nitrate mating plate which has been pre-dried to facilitate absorption of the liquid medium. The plates were incubated anaerobically at 37°C for 1 day. The mating filter was then transferred into a microfuge tube and the cells were resuspended in 1 ml of medium C. Aliquots of the cell suspension were plated onto medium E plates containing Km and Sm. The plates were incubated at 37°C for four to seven days to select for the Km^R Sm^R transconjugants of D. *vulgaris*. Colonies started appearing after four days, the clones were further analyzed by colony PCR and western blotting for transfer of genes and their expression.

3.6 Molecular Biological Techniques

3.6.1 Chromosomal DNA isolation

Small cell pellets of *Desulfovibrio sps* from, e.g., 2 ml overnight cultures were processed using QIAGEN Dneasy® Tissue kit according to the instructions given by the provider. Relatively larger cell pellets, e.g., up to 0.5 g were processed using QIAGEN Genomic-tip 500/G and QIAGEN genomic buffer set. In this case, the additionally needed solutions lysozyme (100 mg/ml) and Proteinase K (20 mg/ml) were prepared separately from the kit. Usually a yield of 2 µg/µl of genomic DNA was obtained, which was pure enough to be directly used for cosmid cloning.

3.6.2 Plasmid DNA isolation

For a small-scale analytical preparation, 2 ml of an overnight *E. coli* cell culture were centrifuged and resuspended in 100 µl buffer P1. The homogeneous cell suspension was lysed with 200 µl of buffer P2 for 1-4 minutes, followed by the addition of 150 µl of buffer P3. For a midi preparation of a high copy plasmid DNA, 200 ml of overnight LB grown culture were processed using 4 ml of each of buffer P1, P2 and P3 buffer in the same order. Occasionally, the volumes of these solutions were adjusted according to the amount of cells and scale of preparation of plasmid DNA. This preparation was centrifuged at 13000 rpm for 10 minutes. Sporadically, the quality of supernatant was improved by addition of one equal volume of a phenol: chloroform: iso-amyl alcohol mixture to the extraction for small scale preparations.

The DNA in the clear supernatant obtained after addition of buffer P3 was either centrifuged directly in for 30 minutes at 13000 rpm after addition of 0.7 volumes of isopropanol, or it was further purified by using the commercial silica columns for DNA binding (e.g. DNA binding columns at QIAGEN). Cosmids are relatively large plasmids and needed a concentration of 2 µg/µl for the purpose of sequencing. For isolation of such larger amounts of DNA, the cells were grown overnight in 2 liters of LB medium and the supernatant obtained after addition of P3 buffer was filtered using QIAGEN maxi cartridge and processed using QIAGEN-tip 2500. Adding 0.7 volumes of isopropanol into 30 ml glass vials, and centrifuging at 14,000 rpm in a JLA 16.225 rotor in Beckman coulter centrifuge for 1 hour precipitated the DNA in the elutions. The pellets obtained after isopropanol precipitations were washed with 70% ethanol and air-dried.

3.6.3 Agarose Gel electrophoresis

0.3-1.2 % agarose gels were prepared in 1x TAE buffer and run at 3-7 V/cm depending on the sizes of the DNA molecules to be separated. A 1 kb DNA ladder from New England Biolabs (NEB) and the lambda marker II from Hoffmann-LaRoche were used as standards to mark the size separation. Finished Gels were stained in 0.01% Ethidium Bromide solution in 1x TAE for 10 minutes and observed on a UV transilluminator. The pictures of the fluorescent DNA were taken using a Mitsubishi gel documentation machine.

3.6.4 Determination of DNA concentrations

The 1 kb ladder size marker from NEB has defined concentrations for each band of different size, allowing a DNA concentration to be estimated by comparison.

The most accurate method is the measurement of the absorbance at 260 nm, using for a pure DNA sample (A_{260}/A_{280} =1.8-2.0) a value of A_{260} =1 for 50 µg/ml.

Another method for evaluating the amount of DNA is based on the spotting of various dilutions of known concentration of a standard solution along with the unknown sample on an ethidium

bromide agar plate followed by incubation in the dark for one hour. When viewed in UV illumination, the respective intensities of bands can be used for concentration determination.

3.6.5 Polymerase chain reaction

3.6.5.1 Gene amplification from genomic DNA and Plasmid clones

Analytical PCR in volumes of 10-50 µl were performed using native Taq polymerase (Invitrogen), while longer and more precision PCR were tried using Pfx polymerase (Invitrogen), Pfu turbo (Stratagene), Hot taq (Metabion) and KOD Hifi (Novagen) in given buffers and other reagents according to the enzyme properties. The annealing temperatures were roughly calculated by the formula; T annealing = 2(A+T) + 4(G+C) - 6 and sometimes determined by parallel reactions using a temperature gradient in the PCR machine.

3.6.5.2 Generation of DIG probes

The probes for the detection of the maturation genes or hydrogenase DNA sequences in a complex mixture of restriction-digested chromosomal DNA or in multiple samples in the cosmid library were generated by incorporation of DIG-dUTP (Roche Diagnostics, Fig. 3.2) in PCR amplification using native Taq-polymerase from Invitrogen. The ratio of dNTP was kept at 2 mM dATP, 2 mM dGTP, 2 mM dCTP, 1.3 mM dTTP, 0.7 mM dUTP to ensure optimal labeling of the amplified product.

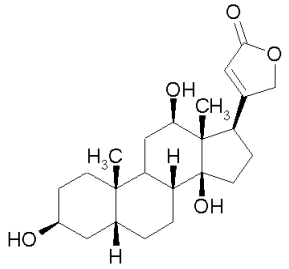


Fig 3.2 Chemical structure of Digoxigenin

3.6.5.3 Degenerate amplification of unknown sequences

Conserved stretches of 15- 30 nucleotides in DNA and of at least 5 amino acids in protein sequences were looked for in genes and proteins similar to the target DNA and its translated

protein. A primer pair of a forward and a reverse primer bearing optimal degeneracy and annealing were designed to these conserved residues. Parameters of the PCR amplification were optimized to get an amplification product in correct range that in the following was verified by sequencing.

3.6.6.3 Whole plasmid amplification for strep tag insertion

While needing to insert a large number of nucleotides coding for the strep tag, the pMOShynBACD was amplified by using back to back primers to the target site and one of which contained the strep tag coding nucleotides in frame to the C-terminus of the *hynA* (with appropriate overhangs). The PCR-amplified plasmid was self-ligated and transformed.

3.6.7 Restriction Digestion and end modification

The digestions of plasmids and PCR products were usually carried out at a concentration of 10-20 ng/µl with the recommended amount of the restriction enzyme for sufficient time, after which they were inactivated by raising the temperature to 65-75°C. Depending on the required applications, the cut plasmid or the amplified DNA products were dephosphorylated using Shrimp alkaline phosphatase, phosphorylated using polynucleotide kinase or were made blunt-ended using T7 polymerase. After the required modification, the DNA solution was desalted before ligation.

3.6.8 Desalting of DNA solutions

Desalting of DNA reaction mixtures for subsequent enzymatic reactions and removal of primers from PCR reactions was accomplished by the GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) using the manufacturer's protocol.

3.6.9 Ligation

A vector cut with different enzymes was ligated with a 2-3 fold molar excess of insert DNA, incubated in most cases at 22°C overnight with T4 ligase enzyme in the appropriate buffer.

3.6.10 Chemically Competent cells

Cells of *E. coli* XL, BL-21 and DH5alpha strains have been used. Chemically competent cells were prepared by $CaCl_2$-treatment of the cells. 1 ml of an overnight culture was used to inoculate 100 ml of LB broth. The cells were grown at 37°C till an A_{600} of 0.4-0.6 was reached. The cells were cooled on ice and centrifuged. Thereafter, 2-3 washes were given with 100 mM $CaCl_2$ by centrifuging at 4000 rpm and re-suspending. The cells were incubated on ice in the same 100 mM $CaCl_2$ solution for about 1 hour. After the incubation, the cells were centrifuged at 2000 rpm and finally re-suspended in 1/100 volume of 15% glycerol in 0.1 mM $CaCl_2$. The final volume was distributed into pre-cooled 1.5 ml tubes in 100 µl aliquots and immediately frozen at -80°C. All these steps were carried out on ice and using cold tips. Special care was taken for gentle re-suspension of cells, as the $CaCl_2$ treatment makes them very brittle.

3.6.11 Electro-competent cells

1 ml of an overnight culture was used to inoculate 1000 ml of LB broth. The cells were grown at 37°C till an A_{600} of 0.8-1.0 was reached. The cells were cooled on ice and centrifuged at 8000 rpm. At first, cells were washed with HEPES buffer and thereafter 3-4 washes were given with a sterile solution of 10% glycerol in distilled water by centrifuging at 5000 rpm and re-suspending. After a final wash, the cells were suspended in a 1/100 volume of the original growth volume of sterile 15% sterile glycerol solution.

3.6.12 Transformation

A 100 µl aliquot of chemically competent cells in the frozen vial was thawed on ice, and 10-100 ng of the DNA was added, followed by incubation on ice for at least 20 minutes. The vial was given a heat shock in a water bath maintained at a constant temperature of 42°C for 35-45 seconds and immediately transferred to ice. After waiting 2 minutes, 200 µl of SOC was added to this vial and the cells were grown at 37°C shaking at 180 rpm for at least 45 minutes before plating on LB-agar plates containing an appropriate antibiotic. For electro-transformation of the electro-competent cells, 1% or a lower volume of DNA in buffer of low ionic strength to the thawed cells was added and incubated on ice for at least 5 minutes before transferring into electroporation cuvettes. The Gene Pulser parameters for the transformation were set at 25 µF,

200 Ω and 2.50 kV when using 0.2 cm cuvettes (1.50 to 1.80 kV when using 0.1 cm cuvettes). Both of the pulse buttons were pressed together until it beeps. After the beep, the cuvette was removed and 1 ml of SOC broth was added immediately. The cells were plated after growing at 37°C in similar way as of the heat-shock transformed chemical competent cells.

3.6.13 Southern Hybridization

This technique is designed to detect a specific sequence of DNA in a complex mixture. In this study, it was used during the detection of homologous genes in restriction digests of chromosomal DNA as well as to screen for those genes in cosmid DNA genomic library.

3.6.13.1 Blotting

Chromosomal or plasmid DNA was restriction-digested with appropriate enzymes and separated by electrophoresis on 0.8% agarose gels. Gels were stained with ethidium bromide and observed briefly under the UV transilluminator. Later the gel was de-stained in distilled water for 1 hour.

The agarose gel was first de-purinated by shaking in 25 mM HCl for 10 minutes, usually indicated by a change of the color of xylene blue (loading dye) to yellow. This delicate gel was then washed with water (two to three times) before denaturating in 0.5 M NaOH/ 1.5 M NaCl for 20 minutes. In case 20x SSC was to be used as transfer buffer, the gel was neutralized with 1 M Tris/ 1.5 M NaCl (pH 8.0) after giving few washes with water.

The DNA from the treated gel was transferred to an equal-sized positively charged nylon membrane by capillary action. The pre-washed membrane (in transfer buffer 20X SSC or NaOH for 2 minutes) was sandwiched between the gel and few equal-sized Whatman paper filters, followed by a tall stack of ordinary paper towels. The gel was placed on top of a long Whatman paper filter dipped in transfer buffer. A heavy load was placed on the top to support the diffusion. This arrangement was left usually 6 hours to overnight.

3.6.13.2 DNA- DNA Hybridization and Detection

The blotted membrane was washed once in 2X SSC and UV irradiated for 3 minutes after drying. After fixation of DNA, the membrane was washed in pre-hybridization buffer for 2 hours at 42°C. The hybridization probe was prepared by boiling 10-20 µl of DIG labeled PCR product in 200 µl endonuclease-free water for 10 minutes. This probe was added to 10 ml of pre-hybridisation buffer and the incubation proceeded overnight.

Next day, the hybridized membrane was washed in 200 ml low stringency buffer for 5 minutes at room temperature. During this, the high stringency buffer was preheated to 60°C and two washes of 15 minutes each were performed at the same temperature.

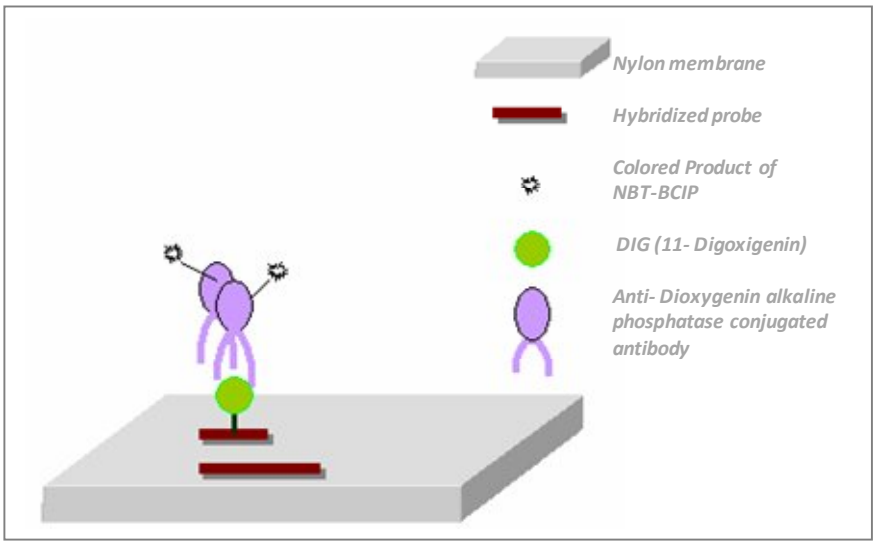


Fig 3.3 Schematic presentation of color development at the site of the homologous DNA-DNA hybridization by immune-detection of DIG bound to the UTP present in the probe DNA by an Anti-Digoxygenin alkaline phosphatase conjugated antibody.

The membrane was cooled and washed with 1x maleic acid buffer and blocked at room temperature for 30 minutes with 3% BSA in 1x maleic acid buffer. A 1/10000 dilution of DIG monoclonal alkaline phosphatase-conjugated antibody was added at room temperature and was incubated for 1 hour. After the final three washes with 1x maleic acid buffer supplemented with 0.001% Tween-20, the blot was color-developed in the dark using a NBT-BCIP solution for 1-5 hours.

3.7 Cosmid Library Generation

3.7.1 Cloning

The isolated genomic DNA of DvMF cells was repeatedly pipetted up and down in a 200 µl pipette tip or was expelled from a syringe through a small bore needle to generate a random distribution of DNA molecules. This sheering procedure will also generate a small portion of DNA which will be of the optimal size (30-45 kb) for ligation into the pWEB-TNC vector. The sheared DNA samples were then separated in a low-melting agarose gel running overnight at low voltages (30 volts). The gel area containg the fragments located in the desired range was cut and the DNA was purified from agarose by agarase digestion (given in the pWEB-TNC™ Cosmid Cloning Kit) and subsequently concentrated to yield a concentration of at least 1µg/µl.

2 µg/µl of this DNA was treated with a mixture of T4 polynucleotide kinase and T4 polymerase (PK reaction mix, GE Healthcare) in a 10 µl reaction. 1 µl of the blunt-cut and dephosphorylated pWEB-TNC vector (0.5 µg/µl) was added to end-treated chromosomal DNA molecules. The added enzymes were inactivated at 75°C for 20 minutes. 1 µl of T4 ligase (GE Healthcare) was added to this mixture and the solution was incubated overnight at 22°C. At the same day, 50 ml of LB broth supplemented with 10 mM $MgSO_4$ was inoculated with a single colony of EPI100-$T1^R$ cells and shaken overnight at 37°C.

Next day, 5 ml of the overnight culture were used to inoculate 50 ml of LB broth supplemented with 1 mM MgSO4 and 0.3% maltose, and shaken at 37°C until an A_{600} = 0.8-1.0 was reached. The cells were stored at 4°C until needed for transfection in next 24 hours. The ligation mixture was packaged into phage heads by using an *in vitro* packaging extract ('MaxPlax Packaging Extracts' pWEB-TNC™ Cosmid Cloning Kit). After thawing from -80° C, a 25 µl aliquot of this packaging extract was added and mixed well with the ligation reaction in a 1.5 ml plastic cap tube. This mixture was incubated at 30°C for 90 minutes, before adding another 25 µl of the packaging extract and extending the 30°C incubation for another 90 minutes.

On completion of the second incubation period, this mixture was diluted in 1 ml of phage dilution buffer (10 mM Tris-HCl; pH 8.3, 100 mM NaCl, 10 mM $MgCl_2$). After well-mixing, 25

µl of chloroform was added to it and the solution was vortexed for several seconds. The proteins and lipids present in the original phage will precipitate down at the bottom of the tube and the packaged cosmids will remain soluble.

The packed heads were added to cells prepared in LB, supplemented with 1 mM $MgSO_4$ and 0.3% maltose, in a 1:10 dilution and were incubated at 37°C for 20 minutes. The phage heads containing the cosmids will infect the *E. coli* cells and inject the cosmids inside the cells. The infected bacteria were plated on LB-ampicillin selection plates and incubated at 37°C overnight. Generally 25-40 colonies were obtained per 100 µl of plated cells.

3.7.2 Colony Picking and Cosmid DNA Isolation

Cosmid clones were randomly picked and grown in 2 ml deep-well plates available in 96-well graded format (Sigma) in 2xYT medium at 37°C and 180 rpm for 16-24 hours, in 3-6 replicas. Each cosmid clone was assigned a unique number and the replicate of each 96-well growth plate was frozen at -80°C for later use. Cells grown from various replicas were pooled to get sufficient amounts of cells for each cosmid and the DNA was isolated using the QIAprep 96 Turbo Miniprep Kit from QIAGEN. All the centrifugation steps were carried out using a rotor 1460 of the bench top centrifuge model Universal 320R (Hettich Zentrifugen, Tuttlingen). The purified cosmid DNA was stored at -4°C in the same pattern as of the original growth.

3.7.3 Dot blotting

For dot blot experiments, the cosmid DNA from eight clones present in one column of the 96 well plate was pooled in 100 µl tubes (usually 1-10 ng of each position) and was mixed with one volume 2x SSC. One volume of formamide was added to this mixture and the mixture was boiled for 10 minutes at 95°C for complete denaturation. 0.5-1 µl of these denatured DNA samples were spotted directly on a nylon membrane gridded with a lead pencil to give each sample a unique position. The nylon membrane was oven-dried at 50°C for 30 minutes and directly processed for hybridization.

3.8 Protein Chemistry Methods

3.8.1 Extraction of hydrogenase from DvMF cells

A frozen cell pellet was thawed and washed (1 g in 2 ml) with ice-cold 25 mM TrisCl, pH 7.0 in a JA 25.5 rotor (10,000 rpm, 30 minutes, 4°C). The washed cell pellet was again resuspended in the same buffer and 0.2 mg/ml DNase I was added. The cell membrane was lysed by sonication or using the French press. For sonification, 50 ml aliquots were sonicated; each aliquot was subjected to 3 rounds of 2.5 sec cycles, at 70% output, with 5-10 minutes of cooling in between. To collect the membrane fraction, broken cells were centrifuged at 50,000 rpm, 4°C, for 90 minutes. The pellet thus obtained was re-suspended in 25 mM TrisCl, pH 7.5 (1 g of wet pellet/4 ml buffer), and trypsin digested overnight at 4°C under argon atmosphere to release the hydrogenase from the membrane. A centrifugation (50,000 rpm at 4°C for 90 minutes) followed to separate the now soluble hydrogenase from the membrane fraction. The supernatant thus obtained contained the protein and was concentrated if required before loading on an ion exchange column for purification.

3.8.2 [NiFe] Hydrogenase purification

For protein purification, an Äctabasic 10/100 machine from Amersham Pharmacia Biotech (now General Electric) was used. Äcta^{TM}basic is an automated liquid chromatography system designed for method development and research applications. It consists of a compact separation unit and a personal computer running the UNICORN^{TM} control system version 3.0. DvMF hydrogenase was purified from the cell-free extract using a DEAE Toyopearl 650 S ion-exchange column (350 ml (XK 50/20 cm)), washed and run at a flow rate of 5ml/minute with buffer A (25 mM TrisCl, pH 7.4), containing 10 mM NaCl. A sample volume of 20-100 ml was injected and the adsorbed hydrogenase was eluted by a linear gradient of buffer B (1 M NaCl in H_2O) from 0-20% (depending on the column volume). The peaks obtained were assayed for the presence of hydrogenase by an activity assay and SDS-PAGE. The active fractions were pooled, and concentrated using AMICON filters (PM 30) to a final volume of 5-50 ml depending on the estimated protein concentration, and applied to a second column of Sephacryl S 200HR (480 ml (XK 26/100)). Buffer A was used to run the column with a flow rate of 2 ml/minute. The peaks thus obtained were subjected to another ion-exchange column DEAE Toyopearl 650 S (74 ml

(XK 2.6/14 cm)), run with buffer A with a flow rate of 1 ml/minute and eluted with a gradient of buffer B from 0-22.5% dependent on the total volume. This was followed by a Hi-Load column (318 ml (XK 26/60 cm)), run with 25 mM TrisCl (pH 7.4) with a flow rate of 1 ml/minute, to desalt the sample besides further purifying the protein.

3.8.3 Purification of APS reductase

The cytoplasmic fraction was collected and 0.25 mM of a serine protease inhibitor were added, Pefabloc SC (Biomol GmbH, Germany). The proteins were loaded on a DEAE-toyopearl 650S (Tosoh, Japan) anion exchange column (10 - 25 % linear gradient of 1 M NaCl). The eluted APS reductase fractions from the DEAE column were concentrated to 5 ml and were passed through a Sephacryl S200HR (GE Healthcare, Uppsala, Sweden) gel-filtration column. The eluted APS reductase fractions were again loaded without concentration on a DEAE-toyopearl 650S (Tosoh, Japan) anion exchange column (0 - 22.5 % linear gradient of 1 M NaCl as per the column volumes). As a final step in the purification, a gel-filtration column, Hiload 26/60 Superdex 200 prep grade (GE Healthcare, Uppsala, Sweden), was run with 25mM Tris-HCl buffer pH 7.4. After the purification, the samples were sealed in an air-tight tube and were stored in liquid nitrogen. The quality of the samples was estimated by SDS-PAGE.

3.8.4 Purification of Desulfoviridin

The protein was isolated from sulfate-reducing bacterum, *D. v.* Miyazaki F. The cell growth was carried out using the method described before (Yagi *et al*, 1968). Cells were lysed by sonication using a buffer containing 10 mM NaCl, 25 mM Tris-HCl pH 7.4 followed by the centrifugetion at 184000 g at 277 K for 90 m. All purification steps were performed at 277 K under aerobic condition. The cytoplasmic fraction was collected and 0.25 n*M* Pefabloc SC (Biomol GmbH, Germany) as serine protease inhibitor were added. The proteins were loaded on a DEAE-toyopearl 650S (Tosoh, Japan) anion exchange column (10 - 25 % linear gradient of 1 *M* NaCl). The eluted DsrABC fractions from DEAE column were concentrated to 5 ml and were passed through a Sephacryl S200HR (GE Healthcare, Uppsala, Sweden) gel-filtration column. The collected fractions from Sepahcryl column were brought on a Q-Sepharose (GE Healthcare, Uppsala, Sweden) strong anion exchange column (10 - 22.5 % linear gradient of 1 M NaCl). The

eluted DsrABC fractions without concentration were loaded on a Q-Sepharose column (10 - 22.5 % linear gradient of 1 M NaCl) again. As a final step in the purification, a gel-filtration column, Hiload 26/60 Superdex 200 prep grade (GE Healthcare, Uppsala, Sweden), in 25 m*M* Tris-HCl buffer pH 7.4 was used. After the purification, the samples were sealed in an anaerobic chamber and were stored in liquid nitrogen. The quality of the samples was estimated by SDS-PAGE.

3.8.5 N-terminal sequencing of the APS reductase and the Desulfoviridin

The purified proteins were separted by SDS-PAGE and stained by coomassie blue. The separated proteins bands were excised from gel and sent at university of Geißen for N-terminal sequencing. N-Terminal Sequencing uses a chemical process based on the technique developed by Pehr Edman in the 1950's where the N-terminal amino acid reacts with phenylisothiocyanate (PITC). The derivatizing process results in a phenylthiohydantoin (PTH) – amino acid and this amino acid is then sequentially removed while the rest of the peptide chain remains intact. Each derivatization process is a cycle and each cycle removes a new amino acid. The amino acids are sequentially analyzed to give the sequence of the protein or peptide**.**

3.8.6 Heterologous over-expression and purification of recombinant proteins in *E. coli*

The desired proteins are furnished with a hexa-histidine epitope, either at their C- or N- terminal. This histidine epitope has the property to complex nickel or cobalt and other bivalent cations. Agarose or similar matrix material which contains one of these two metals in immobilized form is used as resin, to affinity-purify the over-expressed recombinant protein from total cell lysates.

The desired open reading frame or frames were cloned in a T7 inducible promoter-containing vector pRSF-2 Ek/LIC, in frame with a hexameric his-tag-encoding sequence at the future N- or C-terminus of the expressed proteins. This vector is transformed into an *E. coli* strain that carries a T7 RNA polymerase gene, e.g., *E .coli* BL21DE3RIL. An overnight grown culture (1% by volume of the desired scale of growth) was used as inoculum in applicable antibiotic. The cells were grown at 37°C, shaking at 180 rpm until an A_{600} of 0.5-0.6 is achieved. At this phase, these cells were induced by 0.1-1 mM IPTG which activated the T7 promotor and enabled over-expression of the desired protein or proteins. Usually after the induction the growth temperature

is lowered from 37°C aiming at an optimal expression of the over-expressed protein in soluble form.

For collection, the cells were centrifuged at 8000 rpm for 20 minutes, usually after 4-12 hrs of growth after induction. Harvested cells were either frozen at -80°C for later use or were directly processed further. The cell pellet was dissolved in cell breaking buffer TS (10 mM Tris 150 mM NaCl, pH 7.5; 2-3 g of wet weight per ml of buffer) supplied with protease inhibitor. The cell suspension was sonicated in 2 rounds of 2.5 sec cycles, at 50% output. The sonication was evident by the frothing of the cell suspension. During sonication the cold temperature was maintained by keeping the cell suspension container immersed in ice.

Debris was removed from the sonicated cells by centrifugation at 5000 rpm for 15 minutes. The supernatant was ultra-centrifuged at 4.5x 10^4 rpm for 1 hour. In general, the over-expressed protein was present in the soluble fraction obtained after ultracentrifugation. This soluble part of the cell lysate was manually applied to a nickel- or cobalt-resin column, pre-washed with the TS buffer. The bound his-tagged recombinant protein was gradient-eluted by a stepwise increasing amount of imidazole (10 mM, 25 mM, 50 mM, 100 mM, and 150 mM) in the TS buffer. The various elutions were analysed by SDS-PAGE followed by Coomassie staining and in some cases by western blotting.

3.8.7 Determination of protein concentration

Protein concentrations of the protein were determined using the Bradford assay kit from Biorad. Protein dilutions ranging from 0-1.5 mg/ml, with assay buffers, were measured with a spectrophotometer at 700 nm. The spectrophotometer reading in the range of BSA standard was used to calculate the protein concentration.

3.8.8 Methyl Viologen activity assay

For hydrogenase activity assay, 50 µl of hydrogenase were added to 1 ml of a 25 mM TrisCl buffer (pH 7.4), containing 3 mM methyl viologen. This mixture was incubated at 37°C for 5-10

minutes with constant bubbling with H_2. A strong blue color indicates the reduction of methyl viologen and the presence of an active hydrogenase.

3.8.9 SDS-PAGE and coomassie staining

The Pharmacia phast gel system was used for the analytical protein preparations, with 4-20% gradient gels and alternatively, the New England Biolabs gel system with 10% and 12% homogeneous gels was used. To 25 µl of the cell pellet suspension, 25 µl of Laemmli buffer was added, then the mixture was boiled for 10 minutes, and centrifuged for 5 minutes. The supernatant was loaded and electrophoresis was run at 10 V/cm.

For Coomassie staining, the gel was fixed with two incubations of 5 minutes each in a solution of 40% ethanol and 10% acetic acid. It was then stained for 2-3 hours in Coomassie staining solution, followed by destaining in the same solution as for fixing to remove background staining. To preserve the gels, they were washed in 10% glycerol for 5 minutes and allowed to dry. Usually more clarity is obtained after incubating the gel in water after de-staining for 20 minutes.

3.8.10 Western Blotting

The protein samples were analyzed in a gel electrophoresis system from Invitrogen with 1x MOPS buffer using various concentrations of bis-Tris gels depending on the size of protein to be analyzed. The protein samples were boiled at 80°C with 3x SDS-PAGE buffer for 2 minutes. In cases where protein extracts from whole cells were to be directly loaded, a pellet size equivalent of 150 µl overnight-grown *E. coli* cells was boiled in 30 µl of 3x SDS-PAGE buffer for 15 minutes.

After finishing the electrophoresis, the separated proteins were blotted on a PVDF membrane using a constant current of 300 mV for 1 hour in the blotting chamber filled with transfer buffer. The membrane was placed toward the cathode and the gel toward the anode. Gel and membrane were sandwiched with Whatman paper (thickness of 3 mm) and were cushioned by blotting pads in the X-blot module of Invitrogen. The blotting apparatus was kept cool as much as possible using ice water mixture.

After finishing, the blotted membrane was gently removed from the gel. After giving a single wash with the transfer buffer, the membrane was blocked with 3% skim milk powder in TBS buffer. Three washes in TBST each for 10 minutes were given before the primary antibody in appropriate dilution (mixed in 3% BSA-TBS) was applied. Three more washes in TBST were given before adding the appropriate secondary antibody (prepared in 3% skim milk powder in TBS buffer). In some cases, when the anti-DIG monoclonal primary antibody was itself conjugated to alkaline phosphatase, the secondary antibody step was omitted. The color was developed after giving final three washes with TBST, using NBT-BCIP by alkaline phosphatase.

3.8.11 Native PAGE and *in-gel* Activity Assay

Desulfovibrio sp. cells grown in PC medium for 2 days were harvested and equal wet cell weights (1 g) were used for the experiment. Occasionally, the growth medium was supplied with either of $NiCl_2$ or Na_2Se or both at a final concentration of 1 μM. Cells were washed with 50 mM phosphate buffer (pH 7.0) and lysed in TS buffer using glass powder homogenisation in the IKA Ultra Turrax Tube Disperser from IKA® Werke GmbH & Co. KG, Staufen. Cell debris was separated from crude cell lysate by a slow spin at 4000 rpm for 15 minutes at 4°C. The supernatant was centrifuged at 13000 rpm for 1 hour at 4°C. This supernatant thus obtained contained both the soluble and the membrane fraction. Total protein content in each sample was quantified using the Bradford assay, and 200 μg of protein was loaded in one well.

Pre-prepared 3-8 % native gels were run at 3000 Volt per hour at 4°C, to separate the proteins on the gels. Gels were subsequently incubated in either 50 mM phosphate buffer (pH 7.0) or pH 5.5. The gels and solutions were H_2 saturated by evacuating and refilling with H_2 three times. After incubating at RT for 2 hours under H_2 bubbling, the gels were activity-stained either with PMS/NBT, BV/TTC or MV/TTC. For the PMS/NBT assay, 1 ml each of PMS and NBT stock solutions were mixed and injected in a rubber-tight bottle, containing the gel in 200 ml phosphate buffer. The hydrogenase reacts with H_2, the electrons are transferred to PMS, and the reduced Phenazine Methosulfate ($PMSH_2$) reacts with NBT to give a blue colour. The bottles were incubated at 37°C till the blue colour developed. In case of *Desulfovibrio sp.* it took ~ 10-20 minutes. To stop the reaction, gels were washed in excess water.

For the MV/TTC or BV/TTC assay, 0.5 ml of a H_2-saturated BV or MV solution was added to 1 ml of sodium dithionite (60 mM). 200 µl of this solution were injected into a rubber-tight bottle containing the gel in 200 ml phosphate buffer. Almost immediately, the blue colour development could be seen. To fix the bands, 100 µl of TTC was added to the buffer upon which the colour changed into pink. After a uniform colour change, washing the gels in excess water stopped the reaction.

3.8.12 MALDI-TOF MS molecular weight analysis (matrix assisted laser desorption ionisation-time of flight mass spectrometry)

MALDI-TOF MS is used to identify the subunit composition and the molecular mass of a purified protein. MALDI-TOF MS involves embedding the analyte in a solid matrix, which absorbs the energy generated by a LASER beam. This energy absorption, which can be up to the order of 10^6 watts/cm², leads to intense heating and generation of a plume of ejected material that rapidly expands and undergoes cooling. Since this process takes place within the high vacuum of the mass spectrometer, the protein is gradually stripped from matrix and solvent molecules and can then be separated according to its apparent mass to charge ratio.

For the experiments conducted, a Voyager-DE PRO Workstation and a Voyager-DE™ PRO Biospectrometry Workstation from Applied Biosystems (USA) were used. The matrices used were α-cyano-4-hydroxycinnamic acid (CHCA)(M1), 2,5-dihydroxybenzoic acid (2,5-DHB)(M2), Sinapinic acid and a mixture of 2,5- dihydroxybenzoic acid and 5-methoxysalicyclic acid (DHBs).

3.9 Software

3.9.1 Primer Design

The primers for PCR reactions were designed by manual gene alignments and frequent use of vector NTI program as well as Fast PCR program. The primers for site directed mutagenesis were designed using the primeX program available at Stratagene website.

3.9.2 Gene Sequence Analysis

The sequenced genes were analyzed to find promoters by BPROM program at www.softberry.com and BDGP neural network program at the website of Berkeley Drosophila Genome Project i.e. www.fruitfly.org.

3.9.3 Protein Sequence Analysis

The protein sequences of the sequenced genes in this work were compared to known sequences by using NCBI BLAST. The level of similarity of the alignment with various sequences was used to calculate the tree and its bootstrap values in newick (an alogrithm form Arthur Cayley, 1857, now standard for representing trees in computer-readable form that makes use of the correspondence between trees and nested parentheses) format by using PhyML method (phylogenies by maximum likelihood). FigTree software (Author, Andrew Rambaut, 2007) was used to draw the tree.

The PSIPRED secondary structure prediction method (Jones DT, 1999) was used to predict the secondary structure of hydrogenase maturation protein HypA, HypB, HypD and HypE of DvMF. The three-dimensional structure modelling of the maturation proteins HypB, HypD, HypE, HynD and [NiFeSe] hydrogenase was carried out by using comparative modelling program at the ExPasy protein structure repository site http://swissmodel.expasy.org/. The figures of three dimensional models for printing were generated using PyMOL software from www.pymol.org.

3.10 Miscellaneous

Sequencing services at the ADAC of MPI-Köln and also from the Sequiserve, Vaterstetten, were used to sequence the DNA.

Chapter 4

Sequencing of the DvMF Hydrogenases and Sulfate Metabolism Genes

The genomic sequence of *Desulfovibrio vulgaris* Hildenborough (DvH) has revealed genes coding for six hydrogenases (Heidelberg et al, 2004), four of them are periplasmic and other two are membrane-associated, cytoplasmic-oriented hydrogenases (see section 2.3 for details). Among the periplasmic hydrogenases, the hydrogenase *Hyn*1 is present in all of the known *Desulfovibrio* species, and also [FeFe]- and [NiFeSe]- hydrogenases have been reported in most of the subsps of *Desulfovibrio desulfuricans*. The Ech [energy converting hydrogenase] hydrogenase has been reported in DvH and *D. gigas*. But, the hydrogenase isozyme Hyn2 and the Coo (CO induced) hydrogenase has been reported only in DvH. The Hyn1 hydrogenase and the known [NiFe] hydrogenase of DvMF are homologous and present in similar operon structures (Goenka et al, 2006). Both of these operons contain two structural genes coding for the small and the large subunit followed by two maturation genes, the endopeptidase (HynC) and the chaperone (HynD) coding genes (Goenka et al, 2006).

Similarly, additional maturation genes present in the DvH and later in the *Desulfovibrio desulfuricans* subsp G20 (Dd) genomes were deduced based on sequence homologies of these genes to the gene sequences of the maturation genes of *E. coli* (Table 4.1). Compared to multiple maturation genes for multiple hydrogenases, there are only *hyp* (hydrogenase pleiotropy) genes present in the DvH genome. These maturation genes of DvH, namely *hypA*, *hypB*, *hypD*, *hypE* and *hypF* are present in three different operons as shown in figure 4.9. Compared to that, for DvMF only one operon coding for [NiFe] hydrogenase and two *cis* coded proteins were known (Chapter 2). Given a high similarity between the protein sequences of [NiFe] hydrogenase of DvMF with DvH (89% in the large and 85% in the small subunit) and Dd (*Desulfovibrio desulfuricans*) G20 (79% in the large and 80% in the small subunit), we assumed that such a similarity would also be found in the distribution pattern and the numbers of genes present in DvMF for the maturation of the [NiFe] hydrogenase.

Species / Function	*E. coli*	*Desulfovibrio*				
		Dg	*Df*	DvMF	DvH	Dd G20
	hyd 2	*hynABC*	*hynABCD*	*hynBACD*	*hynABCD*	*hynABCD*
Small Subunit (SSU)	*hybO*	*hynB*	*hynB*	*hynB*	*hynB*	*hynB*
Large Subunit (LSU)	*hybC*	*hynA*	*hynA*	*hynA*	*hynA*	*hynA*
C-terminal peptidase	*hybD*	*hynC*	*hynC*	*hynC*	*hynC*	*hynC*
Ni incorporation/ maturation	*hybF*			*hypA*	*hypA*	*hypA*
Ni insertion	*hypB*			*hypB*	*hypB*	*hypB*
Chaperone/maturation	*hybG*	*hynD*		*hynD*	*hynD*	*hynD*
Ni incorporation/ maturation	*hypD*			*hypD*	*hypD*	*hypD*
Purine derivative binding	*hypE*			*hypE*	*hypE*	*hypE*

Table 4.1 Functional homologues of *E. coli* hydrogenase 2 in *Desulfovibrio* sps. Dg (*Desulfovibrio gigas*), Df (*Desulfovibrio fructosovorans*), DvMF, DvH and Dd G20. Complete genetic information is available for *D. vulgaris* Hildenborough and *D. desulfuricans* G20. Genes highlighted in blue are reported in this work for DvMF.

Given the possibility of existing more than one hydrogenase in most of the known *Desulfovibrio* species and to facilitate the search and corroboration of the *hyp* maturation genes in the DvMF genome, we created and screened a cosmid genomic library. The positively detected clones were sequenced for the desired genes by degenerate amplification. This search was aiming at the hydrogenase genes and was further extended for the *hyp* genes. Also, this library was probed to detect gene operons coding for some of the sulfate metabolism proteins, as Adenylyl sulfate reductase and desulfoviridin (Chapter 2, section 2.6).

4.1 Genomic Libraries

A genomic library is a collection of clones, large enough in number to preferentially contain every single gene present in a particular organism. Over the years many vector and phage systems have been developed for this purpose and have led to successful genome sequencing of many organisms employing shot gun, lambda, cosmid, fosmid, BAC (Bacterial Artificial Chromosome) and YAC (yeast Artificial Chromosome) libraries. One of the major frontiers of microbial genome sequencing, the Joint Genome Institute of Department of Energy, USA, employs strategic generation of three types of libraries of three distinct insert sizes, in order to give a complete coverage of the whole genome. The randomly sheared chromosomal DNA is isolated from analytical agarose gels for smaller (3 kb), mid-size (8 kb) and large size (40 kb). After end repairing, the DNA is cloned into suitable vectors to a shot gun library of insert size (3 kb), a middle sized insert library (8 kb) and a cosmid library (40 kb).

Clones of a genomic library are usually designed to include as much genomic DNA as possible in order to minimize the number of clones required to be isolated. A major target of the genomic library is to generate enough clones for a practically complete (99%) coverage of any and every gene. It has to be kept in mind that the larger the insert size, the smaller is the number of clones one needs to screen. However, further sequence identification of genes has to cope with inserts of large size.

Among the current vectors used to clone eukaryotic genomes, the BAC and YAC vectors are linear vectors, which possess the bacterial and yeast origins of replication. Both of these vectors can harbor very large inserts (See table 4.2 for comparison). YAC have advantage over BAC as it can be used to express proteins, which can be functional in eukaryotic cells only, however BAC has a higher stability. The bacterial genomes are more frequently cloned into plasmids, Lambda phage and cosmids. Plasmids are used for creating a shotgun library with smallest insert size, normally anything between 100 bases to 3-4 kb. There are plasmids such as pUC19, which can be stable at a very high copy number. However a large number of colonies is needed to give a full coverage to genome. On the other hand, the Lambda phages are linear vectors, which can be used for cloning relatively larger inserts of size between 22-28 kb. However, these vectors

tend to be unstable and have a lower copy number than plasmids. P1 phage is another *E. coli* virus that is capable of packaging even larger inserts of DNA and is more stable.

Approximate maximum length of DNA that can be cloned into these vectors.		**Characteristics**
Vector type	**Cloned DNA (kb)**	
Plasmid	0.1-10	Very stable, used for shot gun cloning; requires a large number of clones.
lambda phage	25	Stable, large number of clones needed for bacterial genomes
Cosmid	45	Stable, moderate number of clones for bacterial genomes, large amounts of DNA can be purified.
P1 phage	100	Stable, often used to clone eukaryotic genomes.
BAC (Bacterial Artificial Chromosome)	300	Stable, can be used to genetically manipulate mammalian genomes.
YAC (yeast Artificial Chromosome)	1000	Variable stability, difficult to isolate in intact form, clone- mating possible.

Table 4.2 Comparison of the characteristics of various types of genomic libraries in terms of size, stability and their advantages and disadvantages.

Cosmid vectors are hybrids of plasmid (a small ~5 kb plasmid containing the plasmid origin of replication, an antibiotic resistance gene such as amp^r and a suitable restriction site) and the bacteriophage lambda DNA (COS sequence). The contribution of the lambda phage is limited to the certain specific DNA sequences, namely COS sites. The COS DNA sequences are approximately 200 base pairs long and responsible for packaging of genomic DNA of lambda (48502 base pairs) into phage heads. The DNA packing mechanism of the phage is not particular about the DNA sequence present together with COS site; only the size of the insert is important (comparable to its genome size, allowing an insert size between 37 - 53 kb). The total size of the recombinant DNA after cloning the 'genomic DNA insert' in to the cosmid vector should be comparable to:

Size of insert DNA = Size of insert that can be packed – Size of the cosmid vector

The recombinant cosmid DNA is packed into phage heads using in vitro cell extracts prepared from phage-infected *E. coli*. Such a packaging extract, which contains empty phage heads and additional needed proteins, selectively packs the COS sequence-containing DNA molecules lying in the packagable range. Since the recombinant DNA does not encode any lambda proteins, the cosmids do not form any viral particles (or plaques) but rather form large circular plasmids and the colonies that arise can be selected on antibiotic plates, like other plasmid DNA transformants.

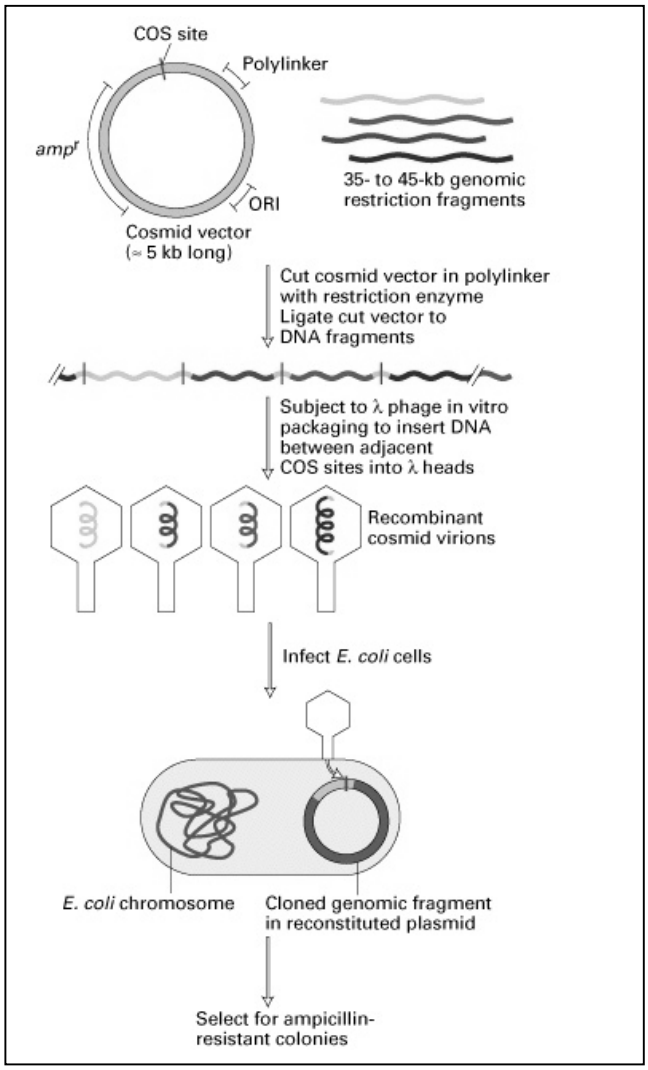


Fig 4.1 Schematic presentation of the mechanism of construction of a cosmid genomic library. Figure taken from Lodish et al, Molecular Cell Biology, 1999, 4th edition,

Thus such cosmid clones have the advantages of both a plasmid (high copy number, antibiotic resistance, non-lytic, non lysogenic function and multiple cloning sites) and a bacteriophage

(stably accommodating large inserts, can be packaged into phage heads and transfects at low concentration).

Preparation of a representative un-biased cosmid library begins with generating the chromosomal DNA in correct clonable range (35- 45 for approximately 5 kb cosmid vector). It is achieved either by partially digesting the genomic DNA with a frequent cutter or via randomly shearing by some mechanical force. The cosmid vectors are cleaved at a compatible cloning site with a restriction enzyme. In case of using mechanical sheared chromosomal DNA, the chromosomal DNA is treated with a combination of DNA-modifiying enzymes such as a polymerase and exo-nucleases, which fill-in or removes the overhangs, and thus blunt and phophorylate the DNA ends. The treated chromosomal DNA ends are ligated in a similarly blunted but dephosphorylated cosmid vector. This ligation reaction is subjected to packaging *in vitro* into the phage extracts and is then used to transfect healthy phage resistant *E. coli* cells

4.2 Cosmid Cloning of DvMF

4.2.1 Preparation of Insert

Different strategies have been employed to generate inserts in desirable range. Ample amounts of genomic DNA were isolated from DvMF using Qiagen midi-preparation kit. During the DNA preparation step where the isopropanol is added, the relatively intact DNA precipitates instantaneously in form of fibers, which were spooled immediately. The rest of the DNA was precipitated by centrifugation. Different strategies were employed to get the maximum number of fragments in clonable range (39-53 kb). The high molecular weight DNA was subjected to partial digestion with Sau3a. In a second attempt, the high molecular weight DNA was randomly sheared by passing through a thin bored needle. The cut or sheared DNA fragments were then separated in a low melting agarose gel running overnight at low voltages (30 volts).

Easiest of all the approaches used, the isopropanol-precipitated DNA prepared after centrifugation (Fig 4.2) was directly used for cloning and usually gave the highest number of clones (Chapter 3).

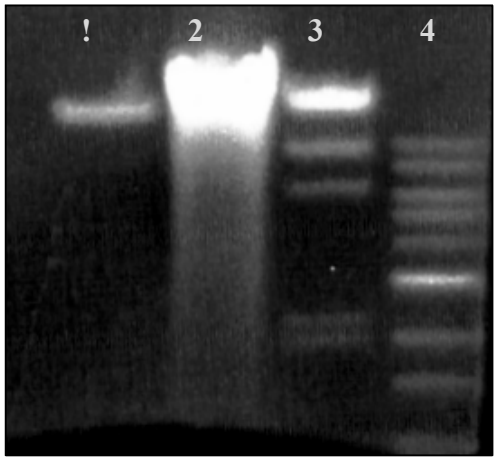


Fig 4.2 Agarose gel electrophoresis of genomic DNA of DvMF, isolated by using the QIAGEN Genomic-tip 500/G, on 0.8% agarose. Lane 1: 100 ng of T7 phage genome marker (40 kb); Lane 2 : 2µl of genomic DNA (~ 5 µg); Lane 3: 5 µl of λ marker Roche [23 kb, 9.4 kb, 6.5 kb, 2.3 kb and 0.5 kb (0.5 kb not visible here)] and lane 4 : NEB 1 kb marker

4.2.2 The cosmid vector pWEB-TNC™

For our purpose we used a cosmid cloning vector pWEB-TNC™ (Fig. 4.3, Epicentre Biotechnologies, Biozym). This is a 5812 base pairs plasmid with a colE1 origin of replication from *E. coli* and possesses ampicillin and chloramphenicol resistance markers. It has been linearized at the SmaI position and dephosphorylated. This cloning site is double flanked by BamHI and EcoRI as well as by a rare cutter NotI, which can be useful in subsequent manipulation.

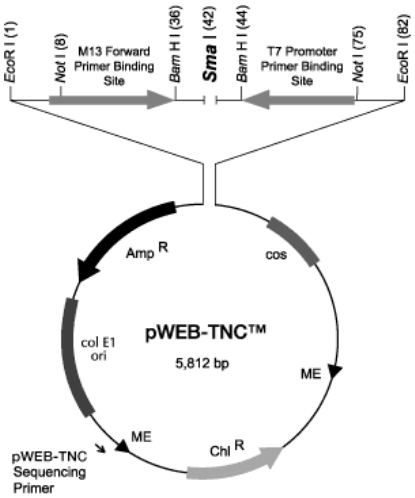
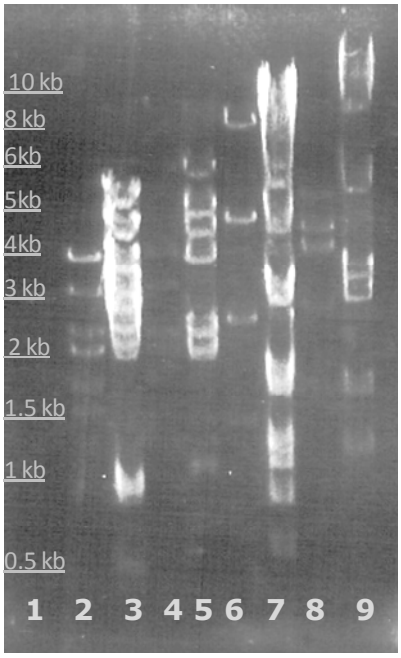


Fig 4.3 Map & cloning region of the pWEB-TNC™ Vector. The SmaI cloning site is flanked by pairs of BamHI, EcoRI, and NotI sites to aid in the excision and mapping of insert DNA.

An amount of 500 ng of this vector was used to clone the isopropanol-precipitated genome DNA and transfect maltose- and $MgSO_4$-treated *E. coli* cells (Chapter 3). After infection, the cells were plated to pick the transfected colonies. Some randomly picked cosmid clones were grown

to isolate DNA. The isolated cosmid DNA were restriction-digested to determine the heterogeneity of the cosmid libray.



1. NEB 1 Kb marker
2. Cosmid clone 1 EcoRI digested
3. Cosmid clone 2 EcoRI digested
4. Cosmid clone 3 EcoRI digested
5. Cosmid clone 4 EcoRI digested
6. Cosmid clone 1 PstI digested
7. Cosmid clone 2 PstI digested
8. Cosmid clone 3 PstI digested
9. Cosmid clone 4 PstI digested

Fig 4.4 Restriction digestion pattern of four different cosmid clones with EcoRI and PstI, resolved on a 1% agarose gel

4.3 Screening and Sequencing

The number of cosmid clones in a genomic library, which can reasonably ensure that any given DNA sequence is contained in the library can be determined by using the following formula (Wahl et al, 1987).

$$N = \ln (1-P) / \ln (1-f)$$

Where P is the probability of cloning a particular gene (expressed as a fraction between 0 and 1); f is the proportion of the genome contained in a single clone i.e. the size of insert divided by the total size (expected or known) and N is the required number of cosmid clones. For example, to

calculate the number of clones required ensuring a 99% probability of a given DNA sequence of *E. coli*, the formula yields:

$$N = \ln (1-0.99) / \ln (1- [4 \times 10^4 \text{ bases} / 4.7 \times 10^6 \text{ bases}]) = -4.61 / -0.01 = 461 \text{ clones}$$

The genomic length of DvMF (3.58 mega bp) is expected to be smaller than that of *E. coli* (4.7 mega bp), indicating that the cosmid DNA preparations from less than 461 clones and subsequent screening could be a reasonable approach for searching homologous genes.

4.3.1 Cosmid DNA Isolation and Dot Blot Hybridization

From the generated and plated genome library, almost 1.5 times of needed, a total of 768 cosmid clones were randomly picked and grown in three to six replicates into 2 ml multi-well deep plates in 96 well formats. The DNA isolated in 96 well format purification kit and purified cosmid DNA were also stored and labeled in 96 well formats. The cosmid DNA clones were dot blotted on positively charged nylon membrane to be probed by various DIG-labeled probes by usual southern hybridization (See Chapter 3 for details).

4.3.2 Sequencing of Positive Clones

A cosmid clone contains an insert of approximately 40 kb of genetic information and the desired gene may be located anywhere inside. Though it is possible to start sequencing from the primers designed to the vector, yet it was preferred to stay focused on the intended gene. Therefore, a strategy to perform the degenerate amplification for straight amplification and sequencing of the required genes was employed.

4.3.3 Degenerate Amplification

As discussed before, a high degree of conservation has been found in the known amino acid sequences of the hydrogenases and related sequences in the *Desulfovibrio sps*, both at the intra-generic and inter-species level. When these sequences are aligned, short stretches of amino acids showing almost perfect conservation can be readily detected. These conserved stretches of amino acids are often integral to a functional or structural domain of an enzyme or a structural protein. In this study, whenever available, the protein and gene sequences from the organisms closely

related to DvMF, such as DvH, *Dd* G20, *Dv* subsp *vulgaris* DP4, *Desulfovibrio gigas* (Dg) and *Desulfomicrobium baculatus* (Db) have been aligned to find such domains. The conserved domains were back-translated into the nucleotide sequences and guess primers of optimal degeneracy were designed out of these sequences by incorporating degenerate bases at the place of variable nucleotides.

Standardized key symbols for degenerate nucleotides are: R = A+G; M = A+C; W = A+T; K = G+T; S = G+C; Y = C+T; H = A+T+C; B = G+T+C; D = G+A+T; N = A+C+G+T; V = G+A+C

Having been designed, the degenerate primer pairs were used to amplify DNA both, directly from genomic DNA and from identified clones. In case of positive identification, the clones were sequenced from the desired gene sequence by primer walking till the end of the required open reading frames.

4.4 Sequencing of Hydrogenases

4.4.1 [NiFeSe] hydrogenase

The [NiFeSe] hydrogenase has two structural genes and has been sequenced from very few organisms, e.g., DvH, *Dv* subsp *vulgaris DP4*, DdG20 and *D. baculatum*.

4.4.1.1 [NiFeSe] Hydrogenase detection

Presence of a gene coding for a [NiFeSe] hydrogenase in the DvMF genomic library was confirmed by dot blot hybridization using a DIG-dUTP labeled probe. This probe was amplified from the homologue of the large subunit of similar hydrogenase in the genome of DvH, employing a forward primer NiFeSeFr 5'-caccgcctccgtcctcgcactcgacg-3'and a reverse primer NiFeSeRV 5'-gtgcttcttgccgctgttgtcgagcg -3'.

4.4.1.2 Degenerate amplification of *hysA*

A comparison with the nucleotide sequence of DvH, *Desulfovibrio vulgaris* DP4 and *Desulfovibrio baculatus* was performed to design optimal primers with minimal degeneracy to

the regions of maximal conservation. The genomic DNA of DvMF was used as template for PCR amplification with a forward primer [NiFeSe]fr230 5'-tcbcagatcgtacagcgyatctgcggcgt-3' (which codes for the domain SQIVQRICGV) and a reverse primer [NiFeSe]Rv900 5'-tcggggagccttgacgaagctg-3' (antisense reverse sequence codes for SFVKAPR). The PCR reaction gave a strong amplification of an approximately 650 base-pairs product with a significant translational homology to the large subunit of [NiFeSe] hydrogenase of DvH. The same primer pair was then used to cross-check the cosmid clone and carrying out the desired sequencing in both the directions by primer walking.

```
                          5'-tcbcagatcgtacagcgyatctgcggcgt-3'[NiFeSe]fr230
       G  R  D  P  R  D  A  S  Q  I  V  Q  R  I  C  G  V  C  P  T
1992234 gggcgcgacccgcgcgacgcttcgcagatcgtacagcgcatctgcggcgtgtgccccact DvH
2182357 gggcgcgatccgcgcgatgcgtctcagattgtacagcgtatctgcggcgtatgccccact DD G20
1195               agggattcttcccagatcgtacagcgtatctgcggcgtgtgccccacg D. bacu
                    R  D  S  S  Q  I  V  Q  R  I  C  G  V  C  P  T


                       NiFeSeRv900  3'-gtcgaagcagttccgaggggct-5'
G K/Q  T  I  P  A  P  D  K  A  G  A  Y  S  F  V  K  A  P  R
ggcaagaccattcctgcccctgacaaggccggtgcgtacagcttcgtcaaggctccccgc 1993013 DvH
ggccagaccatacctgaccccgaaaagcccggcgcatacagcttcgtcaaggctccccgc 2183136 DDG20
                        aaacccggtgcgtacagcttcgtcaaggctccccgt 2177   D. bacu
                         K  P  G  A  Y  S  F  V  K  A  P  R
```

Fig 4.5 Sequence alignment of the [NiFeSe] large subunit genes from DvH (gi|46451220), DD G20 (gi|78217452) and *D. baculatus* (gi|145101) coding for regions maximally conserved at DNA and protein level for designing primers to amplify similar sequences in the genomic DNA and cosmid library of DvMF. The numbers given at the beginning and end of the sequences are the automated numbers given in the NCBI database.

4.4.1.3 Genomic proximity of [NiFe] hydrogenase and [NiFeSe] hydrogenase

The downstream sequencing of the positively identified cosmid yielded another open reading frame (Fig 4.6 and Fig 4.22) at a distance of 409 bases from the 3' end of hysB, encoding for an approximately 14 kDa polypeptide homologous to hupG (putative hydrogenase maturation protein) of DvH and the peptidase M2 of *D. desulfuricans*. This protein is annotated as hydrogenase maturation gene in the DvH genome. Further downstream sequencing of the M2-encoding gene present in cosmid N2 led into the beginning of the transcription unit of the [NiFe] hydrogenase that has formerly been characterized in detail (Goenka et al, 2006). Similar arrangement is present in the DvH and Dd G20 genome also.

4.4.1.4 The [NiFeSe] transcription unit

The operon containing the [NiFeSe] hydrogenase has been completely sequenced from a single cosmid clone. It contains two open reading frames corresponding to the two structural genes, namely the small subunit *hysB* (954 bases) and the large subunit (1527 bases) *hysA* (submitted to NCBI; accession number EU327629). The coding region of the large subunit contains a selenocysteine coding TGA codon.

A comprehensive analysis of the genome fragment carrying both the [NiFeSe] and [NiFe] hydrogenases with BPROM and by BDGP neural network was carried as shown in fig 4.6. It shows all the possible promoter regions, putative ribosome binding sites, the start and stop of seven coding regions namely *hysB*, *hysA*, *hupG*, *hynB*, *hynA*, *hynC* and *hynD*.

BPROM indicated three strong sigma 70 promoters deriving the expression of these two hydrogenases and the two accessory genes present downstream of the [NiFe] hydrogenase. BPROM showed that individual promoter sequences are present upstream of each of the small subunit of both hydrogenase, s1 for the [NiFeSe] and s2 for the [NiFe] hydrogenase. Another sigma 70 promoter is present upstream of *hynC*. The -35 and -10 regions of these sigma 70 promoters have been given in bold.

The BDGP neural network indicated four possible promoter sites present upstream of these two hydrogenase. Sequence B1 is present 35 bp upstream of the start codon of *hysB*. Two overlapping sequences B2 and B3 (also overlep with sequence s2) along with one more sequence B4 are present in front of *hynB* coding region of [NiFe] hydrogenases. Any strong promoter in between the coding region of the structural genes of these two hydrogenases was not found.

Putative ribosome binding sites including the conserved motif GGAG, which is also present in the structural genes of the [NiFe] hydrogenases of most *Desulfovibrio* sps (Goenka et al, 2006), **aaggagg** and **aaggaggaaa** could be identified 5 bp and 3 bp up-stream of the starting codons of the small subunit and the large subunit of [NiFeSe] as well. A phopholipase coding gene is present at the end of the *hynD* coding area.

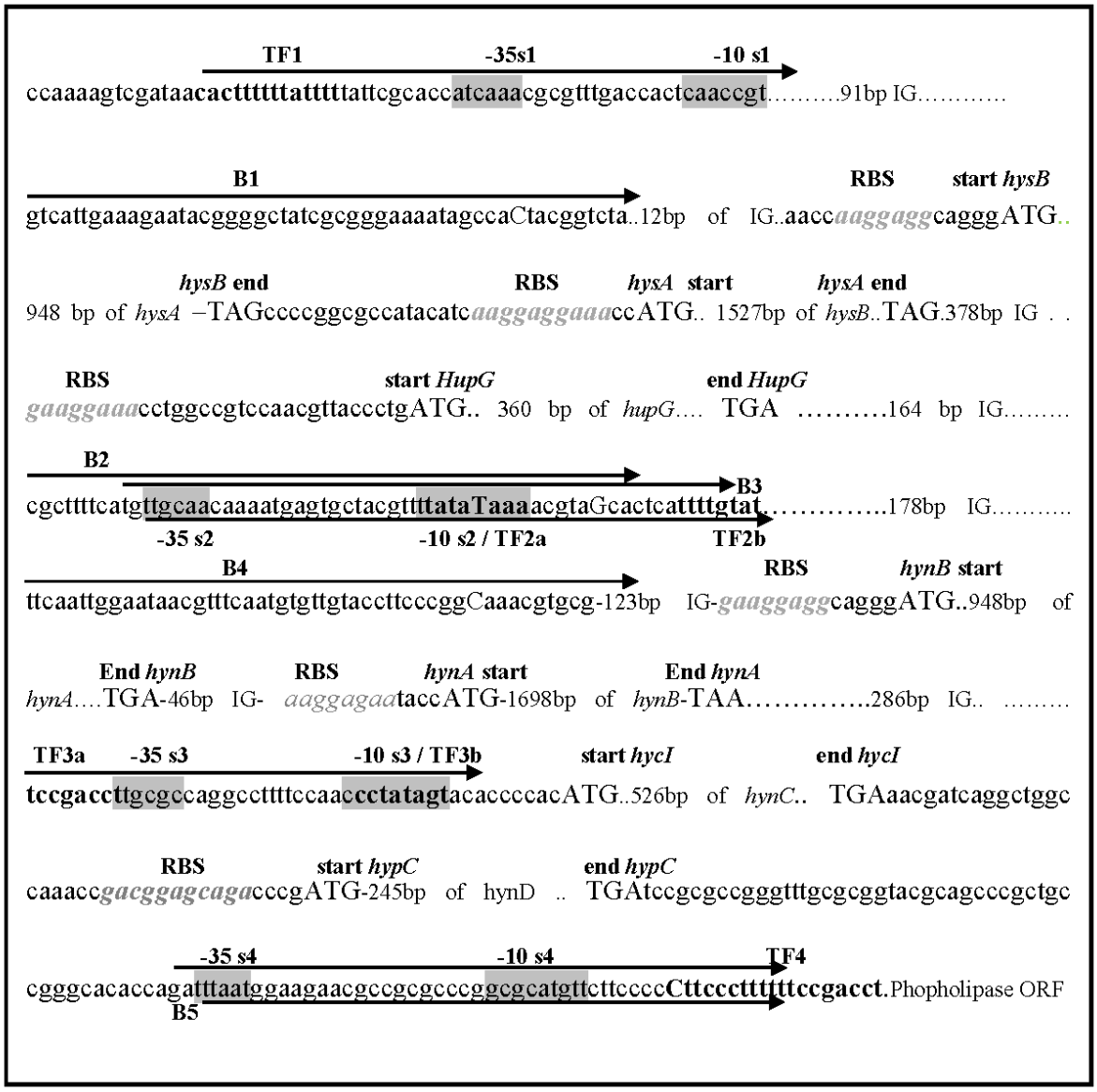


Fig 4.6: Proposed regulatory structure of the [NiFeSe] and [NiFe] hydrogenases encoding genes *hysB*, *hysA*, *hupG*, *hynB*, *hynA, hycI*, and *hypC* of DvMF. Three σ70 promoters were identified by BPROM, s1, s2, and s3 (marked by arrows).The -35 and -10 regions are underlaid in grey. The respective transcription factor binding sites in bold text are, TF1 (carP: CACTTTTT, argR: TTTTTAT, argR2: TTTTTATT, ihf: TTTTATTT, argR2: TTTATTTT, argR2: TTTTTATT, carP: CACTTTTT), TF2a (lexA: TATATAAA, lexA: ATATAAAA), TF2b (fis: CTCATTTT, rpoD17: ATTTTGTA, rpoD17: TTTTGTAT) TF3a (araC: TCCGACCT) and TF3b (rpoD17: CCTATAGT). Another sigma-70 promoter derives the gene expression of a phospholipase present downstream of *hynD* i.e. s4 with transcription binding sites in sequence TF4 (rpoH3: CTTCCCTT, rpoD17: TCCCTTTT, araC: TCCGACCT). The promoter sequences found by BDGP program, B1, B2, B3, B4 and B5 are also indicated by arrows, the captilized letter under these arrows indicates the possible transcription start site. The start and stop codons of the ORFs are given in capitals and the potential RBS (ribosome binding sites) are indicated in the grey italics.

The protein. The small subunit of [NiFeSe] hydrogenase of DvMF translates into a 34 kDa (317 amino acids) polypeptide that includes an N-terminal tat signal sequence (Wu et al, 2002), indicative of the cross-membrane transport of the mature protein into the periplasm. The large subunit hysA (56.3 kDa, 509 amino acids) includes a selenocysteine following proline486. Comparing the sequences of the [NiFeSe]- and the [NiFe]-hydrogenase, both from DvMF, one finds 38% identity and 55% similarity between the primary structures of small subunit and 37% identity and 52% similarity between the large subunit of [NiFeSe] and [NiFe] hydrogenase of DvMF, respectively.

4.4.2 Ech hydrogenase

Purification studies from Archaea have revealed an Ech hydrogenase as a multi-subunit, membrane-bound enzyme complex consisting of at least four hydrophilic and two integral membrane proteins (Meuer et al, 1999). In *Desulfovibrio*, the Ech hydrogenase has been earlier sequenced from DvH and *D. gigas* (Rodrigues et al, 2003).

4.4.2.1 Ech hydrogenase Detection

The presence of sequences of Ech hydrogenase in the DvMF genomic library was probed by a similar method as described above, using a probe amplified from the *EchE* gene present in the DvH genome. The primers employed were (forward primer EcH-Fr) 5‘-ctgcaccgtatccacagc cacctg-3’and a reverse primer Ech-Rv 5‘-tcgaagtcgagttcgtcgaaggccgcgt-3’.

4.4.2.2 Degenerate amplification of *EchE*

As described above for the [NiFeSe]-hydrogenase, the Ech hydrogenase sequences of DvH and *Desulfovibrio gigas* were compared to design primers with a balance between the annealing and degeneracy. A primer pair Ech1076rv 5‘- ckgtgcagcagctgatgcasgggtcgatggac-3’ and Ech319fr 5‘- tggctcggyctcttygccgacgscttcgg-3’ was used to amplify the genomic DNA and the positive clone (identified by dot blot hybridization), which gave a correct PCR product of 600 bp homologous to the corresponding *echE* reading frame of the *Ech* operon of DvH.

4.4.2.3 The *echABCDEF* Transcription Unit

Primer walking led to the gene sequencing of the cosmid clone and revealed an arrangement for the six open reading frames *echA*, *echB*, *echC*, *echD*, *echE* and *echF* in a single operon (SubmittedEU796885) comprising the multi-subunit Ech hydrogenase complex, in a similar manner to that present in the *D. gigas* and DvH genomes (Rodrigues et al, 2003; Heidelberg et al, 2004).

The nucleotide sequence of the complete operon was analyzed by the web-based BDGP neural network program and BPROM as before. Figure 4.7 shows the all the possible promoter regions, putative ribosome binding sites, the start and stop of six coding regions namely *echA*, *echB*, *echC*, *echD*, *echE* and *echF*. BPROM indicated seven sigma 70 promoters present in the sequenced operon of *echABCDEF*. Among these promoter sequences s1 is present...bp upstream and s2 is present 551 bp downstream of the start codon of *echA*. Another promoter internal prmoter s3 is present 441bp downstream to s2. The promoter s4 sequence includes the start codon of *echB* and the s5 is present 6 bp downstream of the start codon of *echC*. Transcription factor binding sequence of promoter s6 is present 12 bp upstream of the start codon of *echE*. One internal promoter is present 260 bp upstream of the stop codon of *EchE*.

The BDGP neural network indicated three possible promoter sequences P1, P2 and P3 present in the *echABCDEF* operon. Sequences P1 and P2 are present just upstream of the *echA* while P3 sequence is present in an overlapping with promoter s7.

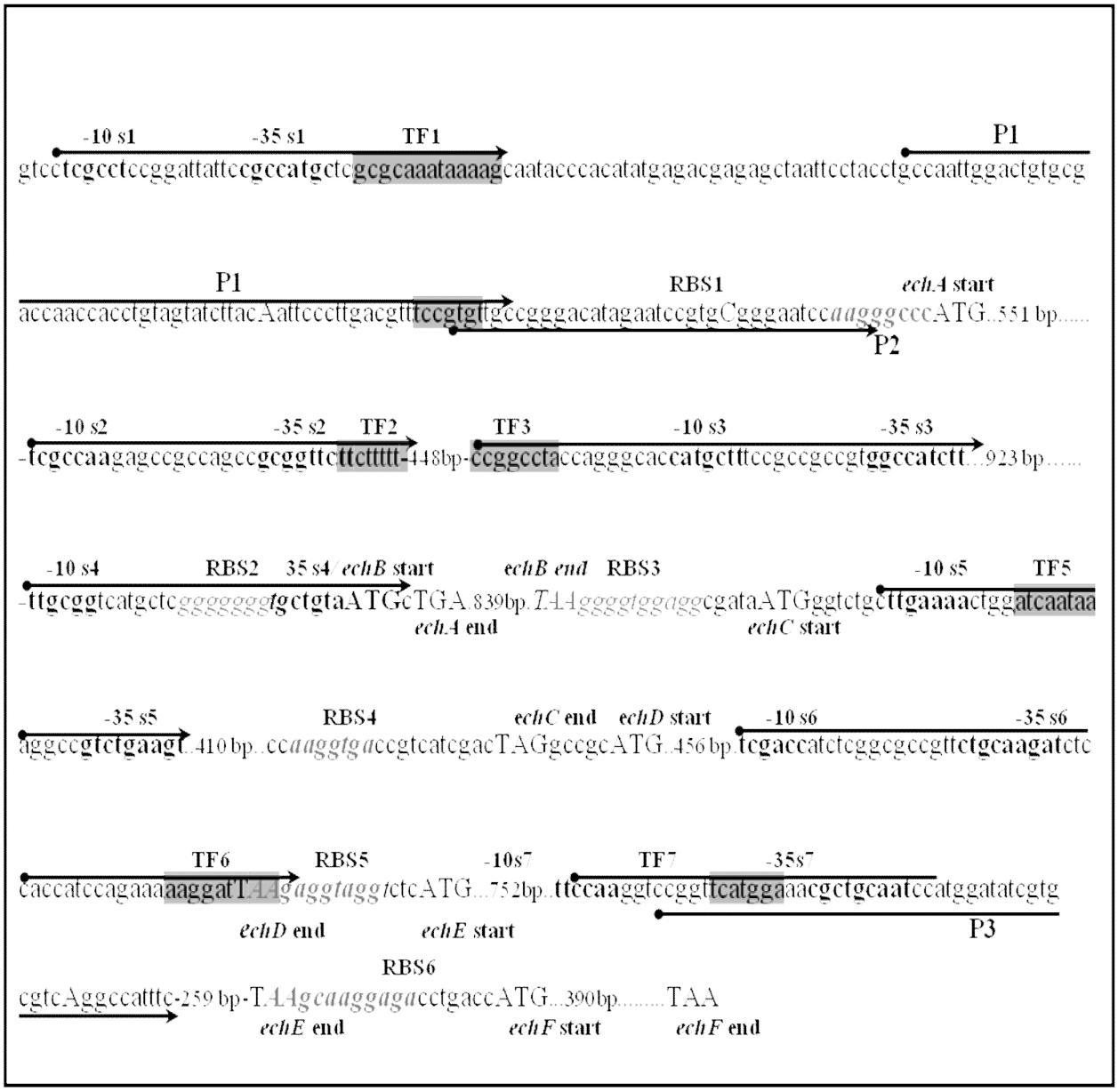


Fig 4.7 Proposed regulatory structure of the Ech hydrogenase containing genes *echA*, *EchB*, *EchC*, *EchD* amd *echE*. Seven sigma-70 promoters were identified by BPROM, s1, s2, s3, s4, s5, s6 and s7, which are marked by arrows. The -35 and -10 regions are underlaid in grey. The respective transcription factor binding sites in bold text are, TF1 (argR2: GCAAATAA, ihf: AAATAAAA, phoB: AATAAAAG), TF2 (ompR: TTCTTTTT), TF3 (ihf: CCGGCCTA), TF4 (tyrR: TGTAATTT), TF5 (fnr: ATCAATAA), TF6 (rpoD15: AGGATTAA) and TF7 (deoR: CATGGA). The promoter sequences found by BDGP program P1, P2 and P3 are also indicated by arrows, the captilized letter under these arrows indicate the possible transcription start site. The start and stop codons of the ORFs are indicated in capitals and the possible RBS (ribosome binding sites are indicated in the grey italics.

The polypeptides. A comparative analysis of the translated polypeptide Ech hydrogenase subunits of the DvMF has been made with the subunits of DvH and *D. gigas* (table 4.3)

EchA is a 68 kDa polypeptide (648 amino acids), containing at least 13 hydrophobic trans-membrane segments as revealed by the hydropathy index (Kyte and Doolittle, 1982).

EchB of DvMF is a 31 kDa polypeptide (282 amino acids) that also shows several hydrophobic trans-membrane domains as demonstrated by the hydropathy index.

EchC is a 22 kDa polypeptide (207 amino acids) showing a conserved domain structure potentially harboring iron sulfur clusters. It exhibits an approximately 34 % similarity with the small subunit of [NiFe] hydrogenase and a 45% similarity with the small subunit of [NiFe] hydrogenase. However, no twin arginine motif has been detected in the N-terminus of any of the known sequences, so presumably this subunit of the Ech hydrogenase may be located in the cytoplasm (Rodrigues et al, 2003).

EchD is a 15.2 kDa hydrophilic polypeptide (135 amino acids) with slightly hydrophobic N-terminal and C-terminal domains.

EchE is a 40.13 kDa hydrophilic **(**358 amino acids) showing approximately 22% and 32% identity to the large subunits of the [NiFe] hydrogenase and of [NiFeSe] hydrogenase, respectively. These similarities are largely located in the N-terminal region of these proteins.

EchF is a 14.39 kDa (129 amino acids) hydrophilic $[Fe_4S_4]$-type ferredoxin protein.

Most probably due to the unstable nature of this enzyme complex in these bacteria (that has been reported also for similar proteins from other organisms), this hydrogenase could not be purified so far and only preliminary functional information has been reported (Rodrigues et al, 2003). Sequence homologies have shown the possible functional conservation in the *Desulfovibrio* group of bacteria including DvMF. In this work, we find a higher similarity between the catalytic

subunits of the Ech hydrogenase and the [NiFeSe] hydrogenase than with the [NiFe] hydrogenase.

DvMFGene Similarities	**Organism**	**Characterized function on sequence homologies**	**Identities**	
EchA (68kD)	*DvH* *Dg* *M. barkeri* *T tenecogenesis*	Transmembrane protein, Anchoring of the complex to membrane and proton transfer.	74 51 44 46	86 66 62 63
EchB (31kD)	*DvH* *Dg* *M. barkeri* *T tenecogenesis*	Transmembrane protein, Anchoring of the complex to membrane and proton transfer.	86 62 45 49	92 82 66 74
EchC (21kD)	*DvH* *Dg* *M. barkeri* *T tenecogenesis*	Fe-S cluster protein similar to the small subunit of [NiFe] hydrogenase	94 68 65 68	98 74 88 86
EchD (16kD)	*DvH* *Dg* *M. barkeri* *T tenecogenesis*	Hydrophilic protein Function unknown	75 54 39 41	88 67 60 60
EchE (40kD)	*DvH* *Dg* *M. barkeri* *T tenecogenesis*	Fe-S cluster protein similar to the large subunit of [NiFe] hydrogenase; hydrogen splitting domain of complex	84 75 54 64	92 86 74 80
EchF (14kD)	*DvH* *Dg* *M. barkeri* *T tenecogenesis*	Hydrophilic [Fe_4S_4] type ferredoxin protein, Electron transfer to and from cytoplasm	81 54 41 46	88 69 56 60

Table 4.3 Functions assigned to the various subunits of the Ech hydrogenase subunits of DvMF, based on the structural homologies to various organisms, percentage similarties given at DNA and protein level

4.4.3 Search for [FeFe] and Coo hydrogenases

No positive signal was obtained from the cosmid genomic library, when it was tried to screen the presence of [FeFe]- and Coo hydrogenase, by similar methods as above. The presence of the Coo hydrogenase was not clear from the faint signals obtained from the genomic restriction digestion hybridization profile (data not shown). However, the presence of a [FeFe] hydrogenase could clearly ruled out.

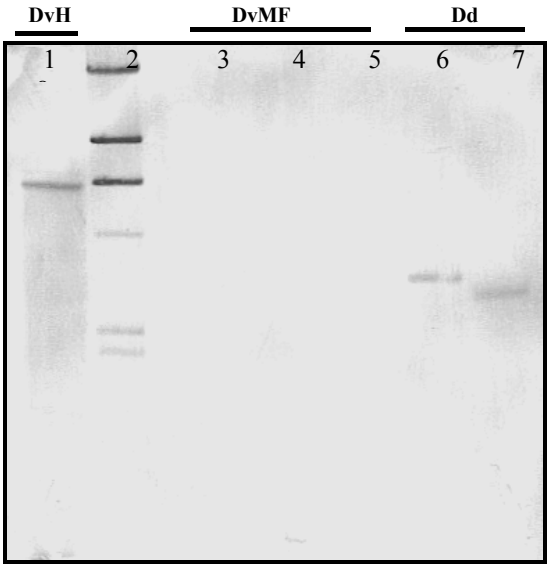


Fig 4.8 Detection of the large subunit of the [FeFe] hydrogenase in DvH, DvMF and *D. desulfuricans* (ATCC 7757) as shown. The order of restriction enzymes used is lane 1: HincII (DvH), HincII, lane 2: λ-DIG marker II, lane 3: HIncII (DvMF), lane 4: NcoI (DvMF). lane 5: EcoRI (DvMF), lane 6 : BamH1 (DvMF), lane 7: HincII (Dd), lane 8: NcoI (Dd). The southern blot was probed against the DIG-labeled probe of 500 bp derived from PCR amplification from *hydA* region of DvH. whereas positive signals were

4.5 Sequencing of Maturation Genes

E. coli, a gamma proteobacterium, possesses three types of nickel iron hydrogenases. These three proteins, hydrogenase 1 (*hyd1*), hydrogenase 2 (*hyd2*) and hydrogenase3 (*hyd3*) are organized in three different operons. Each of the three hydrogenases has some distinct maturation machinery, which is encoded by genes present within the structural gene operons. Yet they all share activities of some other maturation genes (*hyp* genes), present in the two other operons upto a variable extent as per the variable growth environments (Hube et al, 2006).

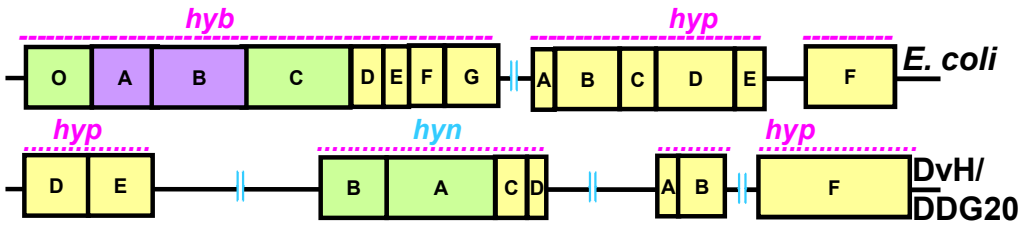


Fig 4.9 A comparison of the clustering of the various hydrogenase related genes in *E. coli* and DvH (green : structural genes; yellow : maturation. Some cytochrome genes are also clustered together in *E. coli* shown in purple.

4.5.1 Detection of the hyp genes in the Chromosomal Digest of DvMF

As the *hyp* maturation genes of DvH and *D. desulfuricans* do not occur in a single operon arrangement, but in three distinct operons (Fig 4.9). To find out the presence and pattern of occurrence, the genomic DNA of the DvMF was restriction-digested with various endonucleases (PvuII, SacI, PstI, SphI, EcoRI and BamHI) and probed by southern blot by using DIG labeled hyp gene probes (derived from one of the genes present in the three known operons, those are *hypB*, *hypE* and *hypF*) amplified from the 'DvH genome' (table 4.4).

Primer name	**Sequence of the primer**	**Base pairs amplified**
hypBfr	5'- acctcggtgcttgaacgcactctcac -3'	~ 240 bp of *hypB*
hypBrv	5'- acggggccaacccctcgccgct -3	
hypEfr	5'-ttcatcgtgacgcgtgataccaaggtcgt-3'	~ 500 bp of hypE
hypErv	5'-cagcggagcagttcgtgcaggtc-3'	
hypFfr	5'- aagggactcgggggcttccacctcctcgcgtcc-3'	~ 500 bp of hypF
hypFrv	5'- accaccgaatcgtccgtgcggatgagaatgtc-3'	

Table 4.4 Primer details used for amplifying the DIG labeled probe from DvH genome for the detection of homologues in the genomic digests of DvMF.

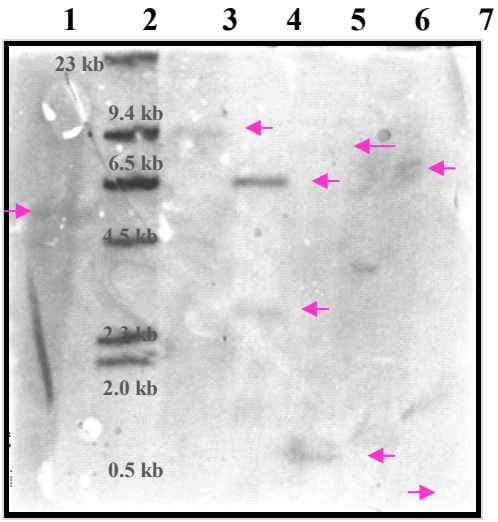


Fig 4.10 Detection of *hypB* sequences in the genomic digests of DvMF as shown.in the order of loading the restriction-digested DNA is lane1: PvuII, lane 2: ; lane 3: PstI; lane 4: SphI; lane 5; EcoRI and lane 6: BamHI and lane 7: probe itself. The southern blot of this gel was probed against the DIG-labeled probe of 252 nucleotide base pairs derived from PCR amplification

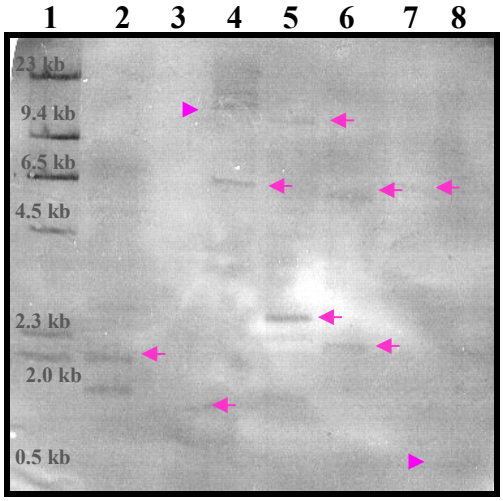


Fig 4.11 Detection of *hypE* in the genomic digests of DvMF. The order of loading restriction- digested DNA are lane 2: PvuII, lane 3: SacI, lane 4: PstI, lane 5: SphI, lane 6: EcoRI, lane 7: BamHI and lane 8: probe. The southern blot of this gel was probed against the DIG labeled probe of 500 nucleotide base pairs derived from PCR amplification from *hypE* region of DvH.

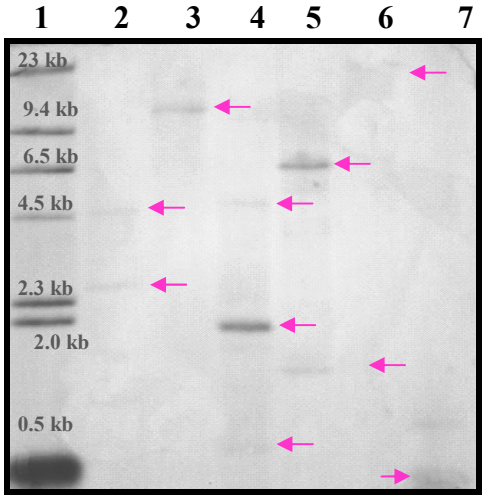


Fig 4.12 Detection of *hypF* in the genomic digests of DvMF. The order of loading genomic DNA restriction digestions were lane 1: λ DIG marker II, lane 2: PvuII, lane 3: PstI, lane 4: SphI, lane5: EcoRI, lane 6: BamHI and lane 7: probe. The southern blot was probed against the DIG labeled probe 500 nucleotide base pairs derived from PCR

A close comparison of the location of the signals obtained in the region of high molecular weight range indicated that these genes i.e. *hypB*, *hypE* and *hypF* are located somewhat far from each other in genome. For instance signals obtained from PstI digested genomic DNA, the signal for *hypB* lies in a region parallel to 9.5 kb (Fig 4.10, lane 4) and those for *hypE* are at 6.5 kb (Fig 4.11, lane 4) and for *hypF* are over 10 kb (Fig 4.12, lane 3) respectively. It means at the closest these three genes are present in a 30 kb genome area, so it is unlikely that these genes are present in one genome.

4.5.2 Sequencing of *hypAB* Operon

4.5.2.1 Degenerate Amplification of *hypB*

A BLAST of the HypB polypeptide sequence of DvH at NCBI database revealed a high degree of conservation among at least two domains of the protein at the intra- and inter-generic level, which have been recently also shown as elemental (chapter 5) for the function of HypB protein (Gasper et al, 2006). The conserved SSPG(A/S)GKT domain (Fig 4.13) was used to design a forward primer hypB-CDF '5- cagttcgcccggckcgggcaagacc -3' and the conserved ENVGNLVCPV domain (Fig 4.13) was used to design a reverse primer hypB-CDR '5-cgggcagaccaggttgcccacgttctc-3'. The 241 base pair PCR amplification product thus obtained gave an 86% translational homology with *hypB* of DvH.

The same set of primers was used to confirm the presence of the *hypB* gene in the positively detected cosmid clones. Extended sequencing of one of the cosmid clones employing various internal primers designed to the DNA sequences generated by the degenerate amplification led to the complete sequencing of two open reading frames of this operon of DvMF that are *hypA* and *hypB* (Submitted EU127916).

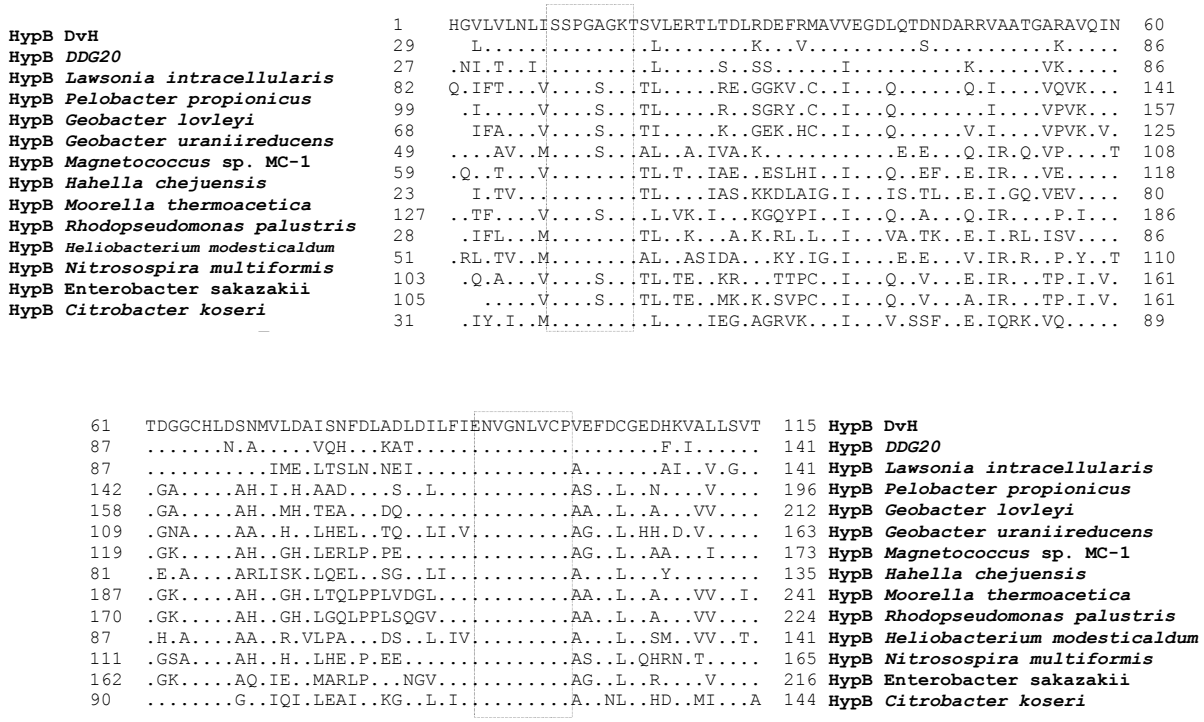


Fig 4.13: Sequence alignment of the conserved domains SSPG(A/S)GKT and ENVGNLVCPV of the HypB in various bacteria. Only the variations are shown and the identities are indicated by dots.

4.5.2.2 The *hypAB* transcription unit

The *hypAB* operon contains two open reading frames *hypA* and *hypB* genes. The *hypA* codes for a 12.8 kDa protein (117 residues) and *hypB* codes for a 24.4 kDa protein (218 residues). The start codon GTG of hypA was designated based on the sequence homology it shares with HypA of DvH and *D. desulfuricans* G20. An analysis by various search programs did not lead to the detection of any promoter sequence present upstream of the *hypA* coding region in the obtained sequence. Also, no typical purine rich ribosome-binding site could be detected. The *hypB* start codon followed just after a 10 bases separation from the stop codon of *hypA*, which is possibly containing the ribosome-binding site.

4.5.3 Sequencing of *hypDE* operon

4.5.3.1 Degenerate Amplification of *hypE*

The polypeptide sequence of HypE of DvH was aligned with BLAST to find conserved domains at the intra- and inter-generic level. Unlike HypB, the HypD and HypE polypeptide sequences did not show a high level of inter-generic conservation, so the search was limited to closely related species such as DvH, *D. vulgaris* subsp *vulgaris* DP4 and *D. desulfuricans* G20. When limited to only these few bacteria, the search for conservation was extended to nucleotide sequences as well.

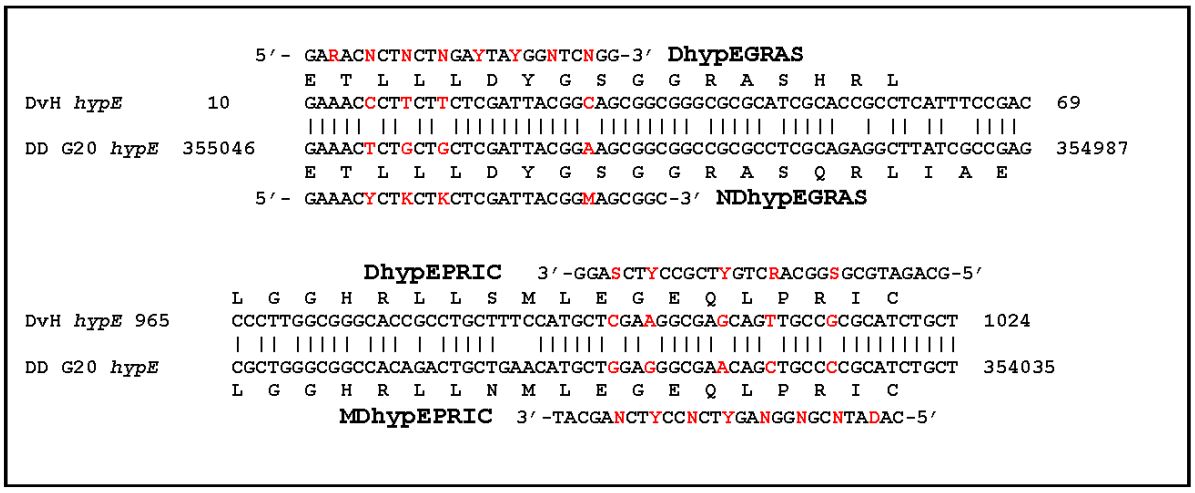


Fig 4.14 Sequence alignments of the *hypE* genes from DvH (gi|46451220) and DD G20 (gi|78355410) of regions maximally conserved at DNA and protein level. Primers derived therefrom are also shown.

The conserved domain ETLLLDYGSG was used to design two forward primers of different degree of degeneration, DhypEGRAS '5-garacnctnctngaytayggntcngg-3' and NDhypEGRAS 5'-gaaacyctkctkctcgattacggmagcggc-3', and in a similar way the conserved MLEGEQLPRIC domain was used to design two reverse primers DhypEPRIC- '5-ggasctyccgctygtcracggsgcgtagacg -3' and MDhypEPRIC 5'- tacganctyccnctyganggngcntadac-3'. Correct annealing of at least one forward and one reverse primer should yield a PCR product of 1 kb. Carrying out the PCR reaction using the genomic DNA as well as the positively detetcted clone, a combination of primers DhypPRIC and NDhypEGRAS gave the highest amount of DNA in correct range from the cosmid clone (Fig 4.15). This reaction was repeated and the

DNA fragment of expected size was gel-purified for sequencing. The sequence obtained gave a 75% translational homology with *hypE* of DvH. The positively detected cosmid clone was confirmed by a similar PCR reaction. Extended sequencing of this cosmid clone in both directions gave the complete ORF sequences of *hypE* and *hypD* with *hypD* being located upstream of *hypE* (Submitted EU127915).

1. 1 Kb marker NEB
2. MDhypEPRIC + NDhypEGRAS
3. MDhypEPRIC+ DhypEGRAS
4. DhypEPRIC + DhypEGRAS
5. DhypEPRIC+ NDhypEGRAS
6. MDhypEPRIC+ NDhypEGRAS
7. MDhypEPRIC+ DhypEGRAS
8. DhypEPRIC + DhypEGRAS
9. DhypEPRIC+NDhypEGRAS

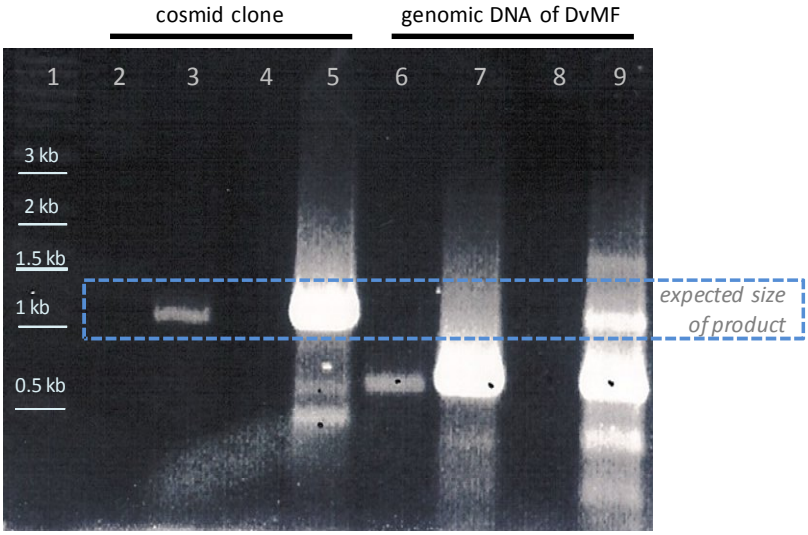


Fig 4.15 PCR amplification of the *hypE* gene from the positively identified cosmid (lane 2, 3, 4 and 5) and from DvMF genomic DNA (lane 6, 7, 8 and 9).

4.5.3.2 The *hypDE* transcription unit

The *hypDE* operon contains two genes *hypD* and *hypE*, being separated by only 3 bases. HypD is a 39.6 kDa protein consisting of 365 amino acids and HypE is a 36.1 kDa protein consisting of 341 amino acids. The start codon GTG of *hypD* was assigned based on the homology with sequences present in the DvH and DD G20. As seen for the region of *hypA* and *hypB* (see 4.5.2.2), also for this operon no strong promoter could be detected, but there are two DnaA type transcription factor-binding consensus sequences present, the first, ttttccacg, present 423 bases upstream and a second motif, ttctccacg, is present 9 bases downstream of the GTG start codon. The possible ribosome binding sites are, one, aggaaga, being present 5 bases upstream of GTG of

hypD and another, gtggagg, present 9 bases upstream the start codon atg of *hypE* and 2 bases before the stop codon of *hypD*.

Fig 4.16: Schematic presentation of proposed partial regulatory sequences present in the *hypDE* operon of DvMF.

4.5.4 Sequencing of *hypF*

4.5.4.1 Degenerate amplification of *hypF*

The *hypF* gene of DvH and other known *Desulfovibrio* genomes is present as a single unit in the chromosome. Of all the *hyp* maturation genes identified in hydrogenase maturation, this is least conserved at the DNA level. A comparison among the gene sequences of *hypF* of various *Desulfovibrio* species shows somewhat higher variability than found in other maturation genes. Again the nucleotide and protein sequences of DvH and *D. desulfuricans* G20 were compared to detect a sequence conserved both at nucleotide and protein level. Two 5' terminal primers hypF200FR 5-tcrtctcsggacaggtacagggwgtgsgctt-3' and hypF590Fr 5'-acccctttyaccaactgyaccaactgcg-3' were designed to the peptide domains V(S/I)GQVQGVG and PFTNCTNCG respectively (Fig 4.17a). Similarly, 3' terminal primers hypF1300rv 5'-cgctgtaagagtaggcgtgsctgctragcca-3' and hypF2300rv 5'-ccacccygaaa ggccsccgcactacgtctt-3' were designed for the peptide domains RDILIRTDSV (Fig 4.17b). After extensive optimization of the PCR conditions, the primer combination hypF590Fr and hypF2300rv gave a combination of more than one faint DNA products, one of which was positively confirmed by sequencing.

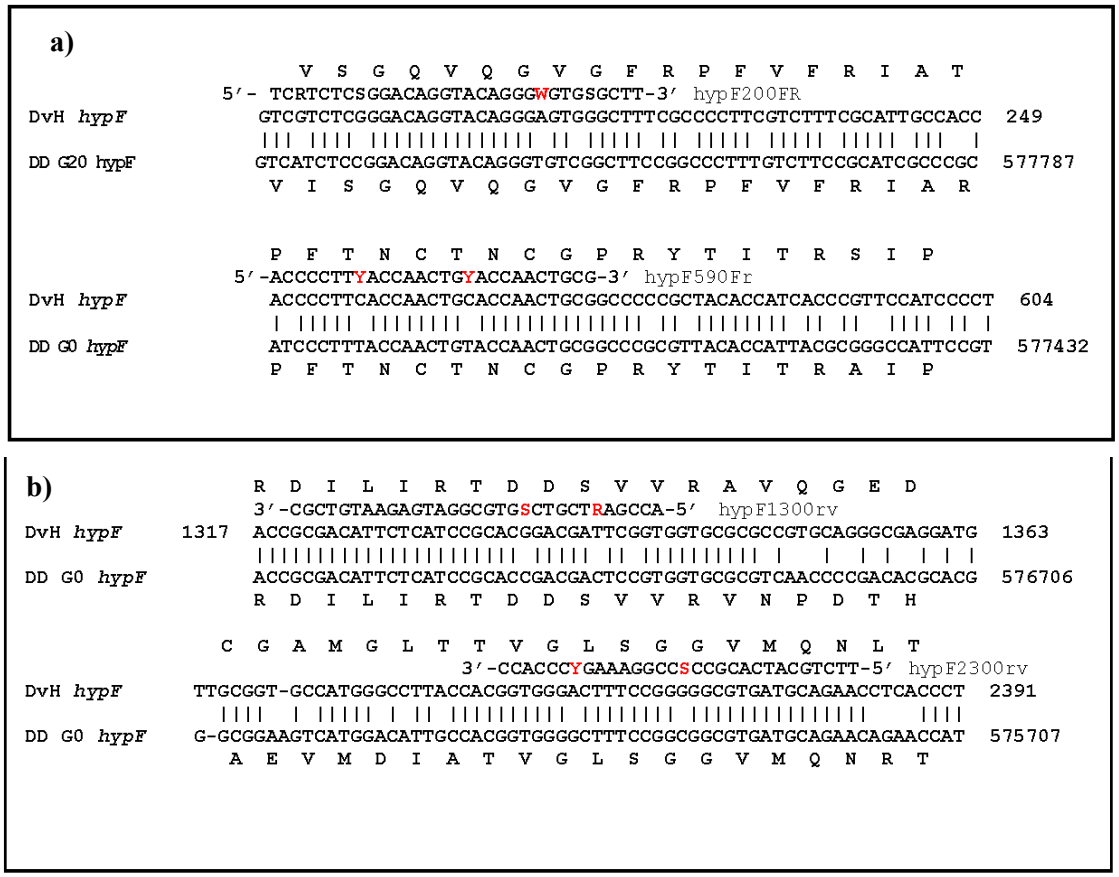


Fig 4.17 a) Sequence alignments of the *hypF* subunit genes from DvH (gi|46581470) and DD G20 (gi|78355602) coding for regions maximally conserved at DNA and protein level for designing forward primers. **b)** Sequence alignments of the *hypF* subunit genes from DvH (gi|46581470) and DD G20 (gi|78355602) coding for regions maximally conserved at DNA and protein level for designing reverse primers.

4.5.4.2 The ORF of *hypF* of DvMF

The ORF of *hypF* of DvMF codes for an 853 amino acids polypeptide (Submitted EU327693) with a theoretically calculated molecular weight 91.4 kDa protein. It shows 61% identity with the corresponding protein from DvH and 62% identity with that from DD G20. The start codon CTG of *hypF* has been assigned based on the homologies between *hypF* of the sequenced genes of DvMF and DvH and DD G20. Interestingly, the analysis of the nucleotide sequence by the BPROM software showed a sigma 70 promoter which includes this start codon.

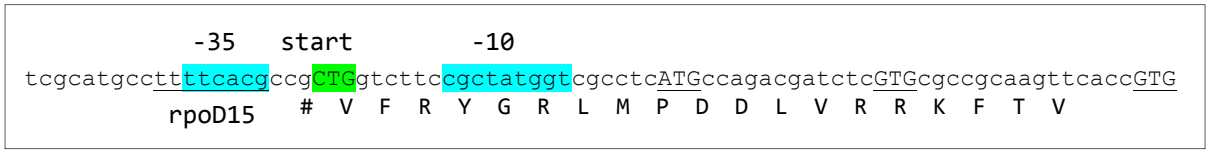


Fig 4.18 Schematic representation of the regulatory sequences present above the *hypF* transcription unit. The area underlaid in blue is the promoter area and the underlined sequence is the transcription factor rpoD15 binding site. The possible start codons are capitalized and the given start codon CTG is highlighted in green.

4.6 The "Green" and the "Brown" proteins

4.6.1 Identification of Adenylylsulfate (APS) reductase and the Dissimilatory sulfite reductase (DSR)

During routine preparations of the [NiFe] hydrogenases, two distinct proteins have been purified from the soluble extracts of the DvMF cell lysates, one with a green and the other with brown color. These proteins were preliminary identified by using MALDI and N-terminal sequencing (see details in chapter 5).

The "Green" protein was found to be consisting of three subunits by SDS-PAGE, having molecular masses of 48, 42, and 11 kDa, also confirmed by MALDI-TOF mass spectrometry. Two subunits were further analyzed and identified by advanced MALDI, the 48 kDa band was assigned as the alpha-subunit of the dissimilatory sulfite reductase of *D. vulgaris* and the 42 kDa band as the beta-subunit of this protein. N-terminal sequencing of the green protein 11 kDa band gave the sequence AEVTYKGKSF and allowed assignment as the gamma subunit. These proteins are also called desulfoviridins.

The "Brown" protein was found to be consisting of two subunits by SDS-PAGE, having molecular masses of approximately 74 kDa and 20 kDa. These subunits were identified with MS to be the alpha and beta subunits of the APS reductase (Chapter 2). The N-terminal sequencing of the 74 kDa subunit gave the sequence PMIPVKEQPK that is similar to the highly conserved N-terminus of the APS reductase of *Desulfovibrio vulgaris* except of the N-terminal methionine,

which might be missing in the processed form. The sequencing of the 20 kDa subunit of the brown protein yielded the sequence PTYVDPSKCDGC. This sequence is highly similar to the N-terminal of the adenylylsulfate reductase beta subunit of *D. desulfuricans.*

4.6.2.1 Detection of APS reductase from the cosmid library

A forward primer ADSRfr 5'-cggttctgaaatcatgcccaccgagcct-3' and a reverse primer ADSrv 5'-tcccagcagcggag cagttcgtgcaggtc-3' were used to amplify a DIG-dUTP labeled PCR probe from the APS reductase alpha gene region of DvH genome (gi|46579260). This PCR probe was used to screen and yielded more than one clearly positive signal among the isolated cosmid DNA from the genomic library of DvMF by dot blot southern assay.

4.6.2.2 Degenerate amplification and complete sequencing of the APS reductase

Two forward primers were designed complimentary to the alpha subunit of APS reductase, one with maximum degeneracy to the sequenced N-terminus that is DbrownFR 5'-cgatgccnatgathccngtnaargarcc-3' (corresponding to the identified sequence MPMIPVKE) and another one was a non-degenerate primer NDBrownFR 5'- atgccgatgattcccgttaagg-3' (related to identified peptide MPMIPVK) assuming that the homologous DNA sequences present in the DvH genome will be perfectly conserved in DvMF as well. Similarly, one non-degenerate reverse primer was designed complimentary to a conserved domain near the C-terminal of the alpha subunit, NDBrownRv 5'-cttcttgaagatcttggtctcgccagtgg-3' (antisense reverse frame TGETKIFKK). The two forward primers were combined to pair with the reverse primer in two separate reactions and successfully amplified a PCR product of size approximately 1.7 kb, which upon DNA sequencing led to positive identification.

The same primer combinations were used to confirm the presence of the APS reductase operon in the cosmid clones. Extended sequencing of one of the cosmid clones by means of various primers designed to the sequenced DNA (obtained by the degenerate amplification) led to the complete sequencing of two open reading frames of the APS reductase subunits of DvMF (Submitted EU 127913).

Fig 4.19 The APS reductase operon of DvMF contains two open reading frames. Arrows indicate the approximate location of the degenrate primers used to amplify the coding DNA sequences.

4.6.2.3 The APS reductase transcription unit

The two gene operon (AprAB) of APS contains the gene *AprA* which codes for the large or alpha subunit polypeptide of 665 amino acids (theoretical molecular mass 75.7 kDa) and the *AprB* gene coding for a smaller or the beta subunit polypeptide of 168 amino acids (molecular mass 18.6 kDa). BPROM was used to detect promoters present upstream of the coding region of AprAB. Besides some potential promoter sequences within the gene, two sigma70 promoter were detected present upstream. The first promoter, S1, was approximately 0.5 kb upstream of the start codon of *aprA* with a crp type TF (transcription factor) binding site, and a second promoter S2 included the start codon of the *aprA* (see Fig 4.20).

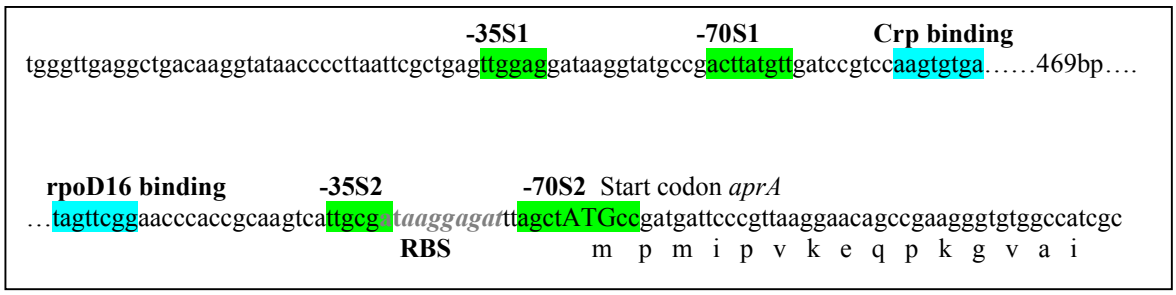


Fig 4.20 Schematic presentation of the possible regulatory structures present upstream of the start of the *AprAB* operon (see text for details).

4.6.3.1 Detection of DSR genes in the cosmid library

A primer pair consisting of a forward primer DsvFR 5'-ggaggtaaggcaatggcgaaacggagg -3' and a reverse primer DsvR 5'-actt gccgaggcacgattcaggggtacg -3' was used to amplify a DIG-dUTP

labeled PCR probe from the *dsvA* gene of the *dsvABD* operon from DvH genome and to screen the cosmid library.

4.6.3.2 Degenerate amplification and complete sequencing of the DSV operon

Two forward primers was designed complimentary to the N-terminus of the alpha subunit DSV reductase of DvH, one with maximum degeneracy to the obtained protein sequence that is DGreenFR 5'-tatgacncaytggaarcayggnggnathgt-3' (MTHWKHGGI) and one other non-degenerate primer NDGreenFF 5'-ggacccactggaagcacggcggcatcg-3' (THWKHGGI) considering that the homologous DNA sequences from the DvH will be perfectly conserved in the DvMF as well. Similarly, a set of two reverse primers was designed complimentary to the sequenced N-terminal part of the gamma subunit, a non-degenerate one, NDGreenRV 5'-gctcttgcccttgtagtgacttcag -3' (antisense reverse frame to EVTYKGKS) and a degenerate one, DGreenRV 5'-tccyttrtang tnacyttcngccat-3' (antisense reverse frame to WxKVTYKG). Each of the forward and the reverse primer pairs was used to amplify a PCR product of approximately 2 kb, which upon DNA sequencing led to the positive identification of *dsvA* and *dsvB* genes of DvMF.

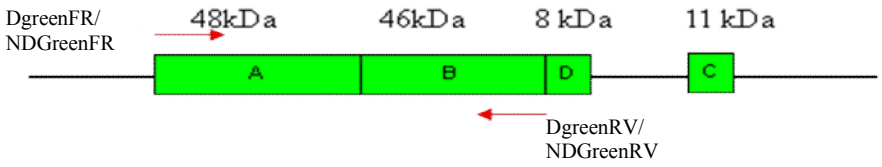


Fig 4.21 The desulfoviridin protein of DvMF contains at least four open reading frames arranged in two transcription units. Arrows indicate the approximate location of the degenrate primers used to amplify the coding DNA sequences.

The same primer combinations were used to confirm the presence of the *DsvA* gene in the positively detected cosmid clones. Later, the complete operon of *DsvABD* reductase was elucidated (submitted EU 127914) by means of primer walking in both directions to the known DNA sequences (obtained by degenerate amplification).

4.6.3.3 The DSV reductase transcription unit

The mature desulfoviridin protein consists of at least three subunits, dsvA, dsvB and dsvC. These three subunits are coded from two different transcription units; one three gene operon (*dsvABD*) contains the α-subunit *dsvA* coding for a polypeptide of 437 amino acids (49.5 kDa) and the β-subunit *dsvB* coding for a polypeptide of 381 amino acids (42.7 kDa). Neural Network Promoter Prediction (Reese 1980) identified two possible promoter regions, one being present as aagctgttgtcattccccgcatggtgggctacaagcaaggAagtgtcgcc (where the capitalized letter stands for the possible transcription start), 390 bases upstream the start codon, and the second one, tccttttaggaatggcgttactttgtccaaaaaatcacgaGttttccggg, is 122 bases upstream of the start codon of the *dsvA*. The *dsvC* subunit was partially sequenced by degenerate amplification, but the full gene sequence could not be detected from the cosmid genomic library.

4.7 Discussion

The used strategy to clone the genomic DNA in randomly selected large inserts (40 kb) into a cosmid library. The subsequent screening for some selected genes including diverse new hydrogenases and *hyp* maturation genes, by means of southern blot screening based on the existing high homologies among various *Desulfovibrio* genomes was successful. We sequence-identified two new hydrogenases, the four genes coding for the mature desulfoviridin protein present in two loci, a two gene APS reductase operon and five *hyp* maturation genes present in three operons. The data produced from the sequencing of the desulfoviridin and APS reductase helped to analyse the crystal structure of these proteins (Ogata et al, 2008). The DNA sequences for the [FeFe] hydrogenase were not found in this cosmid library and its presence in the genome also was clearly ruled out by genomic hybridization detection.

Among the hydrogenases, the two structural genes coding for the [NiFeSe] hydrogenase, *hysA* and *hysB* of DvMF, are located just 1 kb upstream of those of the [NiFe] hydrogenase (Fig 4.22). This arrangement is almost identical to those found in the genomes of DvH and Dd G20. It is proposed that the expressions of [NiFeSe] and [NiFe] hydrogenase are probably co-regulated by a common mechanism (Pereira et al, 2006). The obtained DNA sequences of the [NiFeSe] hydrogenase as well the adjacent [NiFe] hydrogenase, together with the connecting region are

analyzed to present a single regulon structure possessing multiple promoters (Fig 4.6). As introduced before, the [NiFeSe] hydorgenase is dominant in various growth conditions in DvH, the further expression and characterization of [NiFeSe] hydrogenase have been carried in Chapter 6.

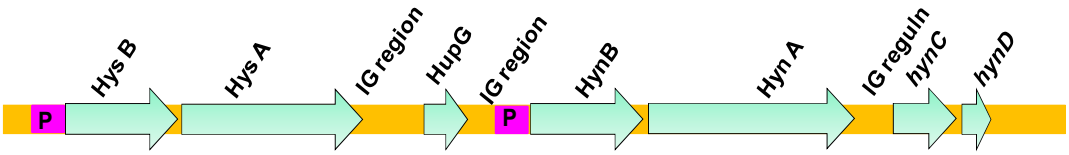


Fig 4.22 Schematic presentation of the genes coding for [NiFe] and [NiFeSe] hydrogenases in the DvMF genome as found on the cosmid clone. The size of arrow is approximately proportional to the gene size and P indicates the promoter present before the coding sequences.

Further, an operon of a multisubunit, membrane-associated Ech hydrogenase comprising of six genes (*EchA, EchB, EchC, EchD, EchE* and *EchF*) has been sequenced and characterized for its regulatory structure. Among the three known hydrogenases, the Ech hydrogenase has the largest number of genes present in one operon. Again, the Ech hydrogenase contains upstream and internal promoters in multiple numbers (Fig 4.7). RT-PCR detection for the transcription of this hydrogenase from total RNA of *D. gigas* could not verify the presence of a transcript that codes for all the genes together at once (Rodrigues et al, 2003). Functions of various subunits are listed (table 4.4), based on the structural homologies to some known proteins: EchE and EchC being the catalytic units, similar to the large and smaller subunit of a heterodimeric hydrogenase. The different level of cellular presence of the various subunits under given physiological conditions has been also confirmed by a microarray study (Caffrey et al, 2007, Pereira et al, 2008), indicating a relatively higher up-regulation of EchE over the other subunits.

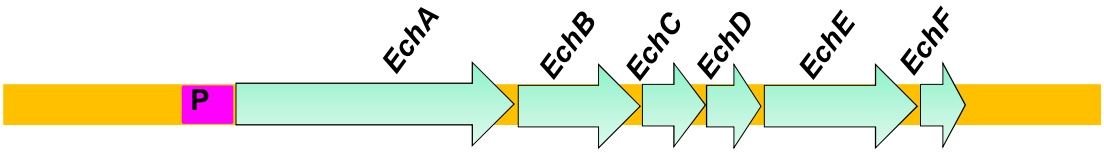


Fig. 4.23 Schematic presentation of the sequenced subunits of *echABCDEF* operon and their arrangement in the DvMF genome. The EchC subunit from DvMF shows approximately 34% similarity with the small subunit of its [NiFe] hydrogenase and 45% similarity with the small subunit of the [NiFeSe] hydrogenase, but lacks a tat signal sequence. Similarly, EchE shows approximately 22% and 32% identity to the large subunits of the [NiFe] hydrogenase and of the [NiFeSe] hydrogenase.

Growth studies of DvH in recent times employed microarrays, to see the levels of expression in order to give functions to various hydrogenases. Data from that approach added information to an advanced hydrogen cycling hypothesis (Caffrey et al, 2007, Pereira et al, 2008). A growth study of DvH with either hydrogen or a lactate-pyruvate combination (Pereira et al, 2008) has led us to assume that the main role of the periplasmic hydrogenases is to fuel the metabolism of *Desulfovibrio* in a growth condition deficient of other electron-donating organic compounds. Further, even among the periplasmic hydrogenase themselves the level of expression of each hydrogenase is differently affected by presence of some metal ions and partial pressure of hydrogen in the growth environment (chapter 2).

On the other hand, the cytoplasmic oriented hydrogenases are upregulated under conditions when the concentration of electron-donating organic compounds such as lactate or pyruvate is sufficiently high (Pereira et al, 2008); under such situation these hydrogenases are mainly involved in producing hydrogen out of that energy.

In conclusion, we have confirmed based on the sequencing work that DvMF contains a [NiFe]-, one [NiFeSe], one Ech hydrogenase and possibly also one Coo hydrogenase. In other words, it contains several periplasmic and at least one cytoplasm oriented hydrogenases. Even in the absence of an [FeFe] hydrogenase, these facts fit very much with former data and with the new hydrogen cycling principle of *Desulfovibrio* bioenergetics.

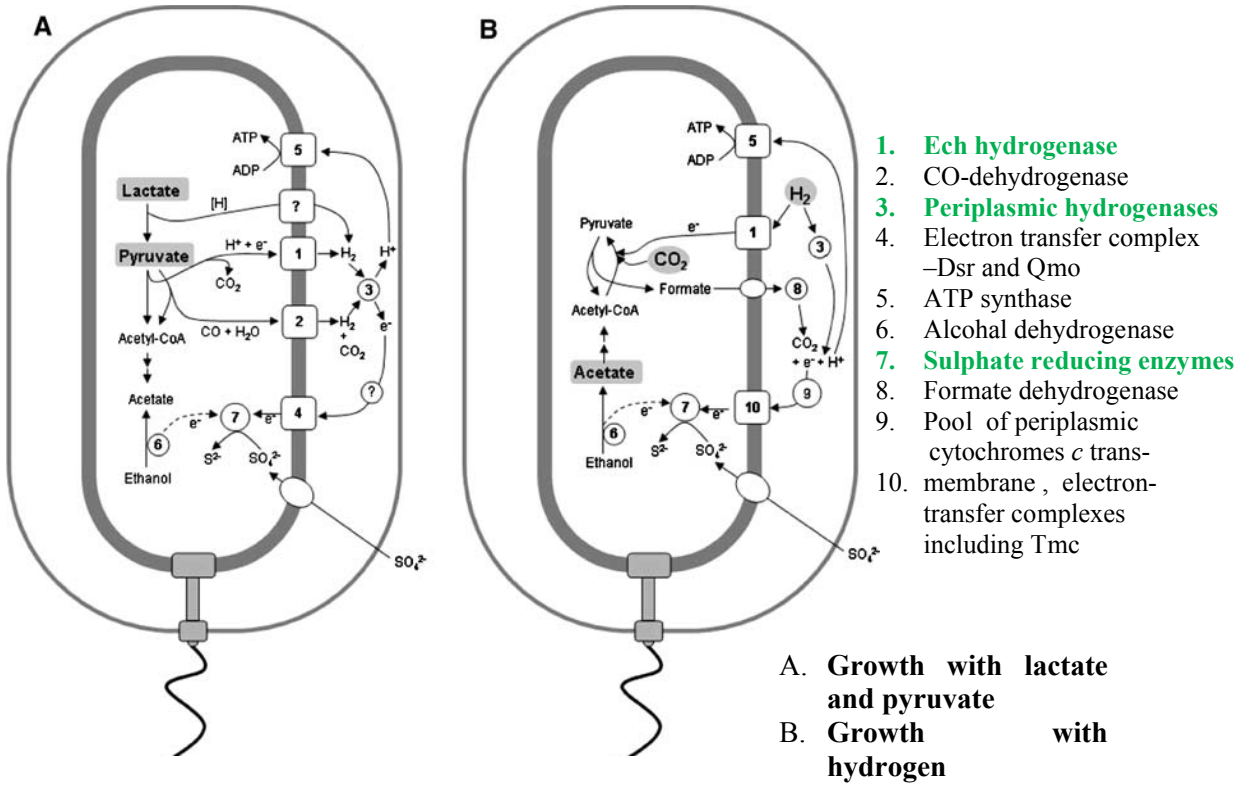


Fig 4.24 Schematic presentation of a DvH cell, indicating the role of various hydrogenases when grown with both lactate and pyruvate or with hydrogen only (Figure taken from Pereira et al, 2008). The captions in green font indicate to the similar gene sequences for DvMF, found only in this work.

Not much has been done on elucidating the pathways of the maturation of the [NiFe] hydrogenase in *Desulfovibrio*, believing that the presence of *hyp* gene sequences is an indication enough for the existence of a maturation cycle similar to *E. coli.* However, despite high conservation, the maturation cycle of each organism is likely to have some distinct aspects. This might be demonstrated by the results from attempts to heterologously express the four gene hyn operon of *D. frutosovorans* in the [NiFe] deletion mutant of *D. gigas*, which resulted in only one sixth of the original hydrogenase activity (Rousset et al, 1998).

It is noteworthy that the hydrogenase expression and purification is much easier in *Desulfovibrio*, which has otherwise difficult genetics (Bender et al, 2006). Moving on to find information on the maturation of the [NiFe] hydrogenase, five hydrogen pleiotropy genes, also abbreviated as *hyp* genes (*hypA, hypB, hypD, hypE* and *hypF*), have been sequenced in this work, present in three

different operons. The pattern of the operon structure of these genes in DvMF is similar to the other known *hyp* genes of *Desulfovibrio* sps. With the availability of these gene sequences, efforts can be made in direction of a heterologous expression of the [NiFe] hydrogenase.

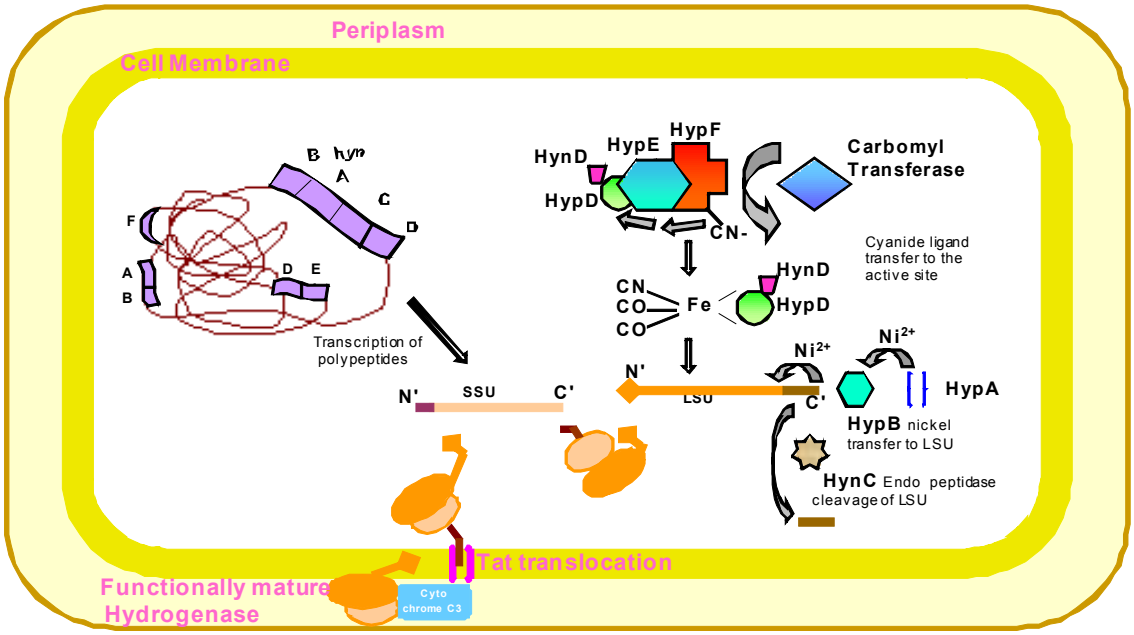


Fig 4.25 A proposed model of expression and maturation of [NiFe] hydrogenase and the Hyp protein in the DvMF cell, parallel to that for the hydrogenase 2 of *E. coli*.

Chapter 5

Molecular characterization of Hydrogenases and Sulfate Reducing Proteins of DvMF

As introduced in chapter 2, the catalytic units of the [NiFe] hydrogenases belonging to *Desulfovibrio* are usually heterodimers, consisting of small (SSU) and large (LSU) subunits. The highly conserved polypeptide sequences of the small subunits contain three iron-sulfur clusters while the large subunits contain a nickel-iron centre. A general tat-transport pathway along with the Hyp proteins (HypA, HypB, HypD, HypE, HypF) as well as the endopeptidase HynC (also named as HycI) and the chaperone HynD (also named as HypC) work together to bring [NiFe] hydrogenases into a functional state. For the [NiFeSe] hydrogenases, one of the Ni-bound cysteine residues is replaced by selenocysteine. Among the three main classes of hydrogenases, [NiFe], [FeFe] and [Fe] hydrogenases, the [NiFeSe] hydrogenases are placed into the [NiFe] family, based on their sequence similarity.

The few crystal structures of [NiFe] hydrogenases we have by now derive all from DvMF and other *Desulfovibrio* sps (table 5.1). Most of them exhibit similar basic structures.

Class	**Organism**	**PDB entry**	**Reference**
[NiFe]	*Desulfovibrio vulgaris* Miyazaki F	1H2R	Higuchi et al, 1997
[NiFe]	*Desulfovibrio gigas*	1FRV	Volbeda et al, 1995
[NiFe]	*Desulfovibrio fructovorans*	1YQW	Volbeda et al, 2005
[NiFe]	*Desulfovibrio desulfuricans* ATCC27774	1E3D	Matias et al, 2001
[NiFeSe]	*Desulfomicrobium baculatus*	1CC1	Garcin et al, 1999

Table 5.1 Available crystal structures of [NiFe] hydrogenases and their PDB codes.

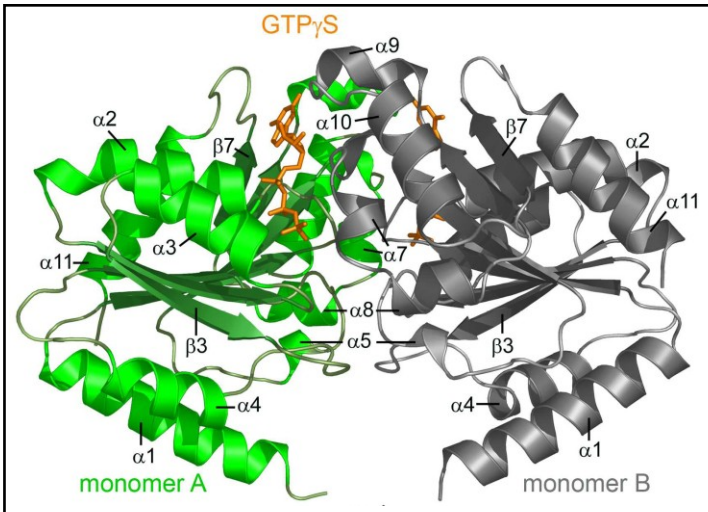
Considering the hydrogenase maturation proteins, a similar situation is seen. A score of these proteins from *E. coli* has been well characterized in terms of genetics and their biochemical role

in the maturation (Chapter 2 for introduction). Also, at least one functional homolog of such proteins from various microbes has been crystallized to resolve their three-dimensional architectures, which have furthered our understanding of the maturation process at the molecular level. A list of available structures of these maturation proteins has been given in table 5.2. As these maturation proteins are highly conserved in a broad range of microbes, the crystallized structure resolutions obtained for HypD, HypE and HypC of *Thermococcus kodakaraensis* KOD1 have been helpful to propose a common reaction mechanism (Fig 5.2).

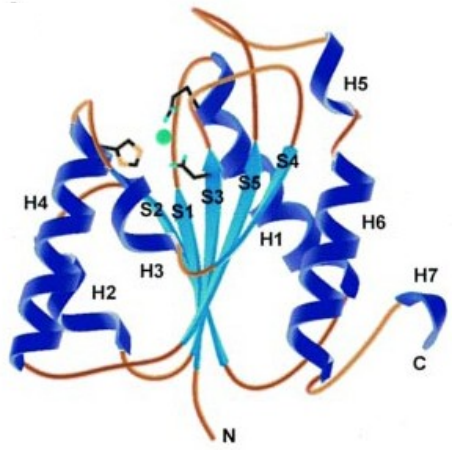
The overall structure of HypC of *T. kodakaraensis* consists of a β-barrel (oligonucleotide/oligosaccharide binding (OB)-fold) domain and a C-terminal α helix (Fig 5.1E, Watanabe et al, 2007a). The N-terminal residues assume an extended conformation to allow the essential cysteine residue (Cys2) to make contact with the solvent. The structures of HypE of *T. kodakaraensis* and also that from DvH (Shmoura et al, 2007) consist of two α/β domains (domains A and B) and a C-terminal tail (Fig. 5.1C, Watanabe et al, 2007b). These structures are similar to that of PurM proteins, whose members are characterized by a common motif for ATP hydrolysis. HypE of *T. kodakaraensis* was obtained in form of a dimer (Watanabe et al, 2007a), which showed four conserved motifs (Table 5.2), assembled around the active site. This arrangement has led to the suggestion that the HypE dimer is a functional unit for ATP-dependent dehydration activity (Watanabe et al, 2007b). The crystal structures prepared by co-crystallization with ATP have shown that the C-terminal tail of HypE exists in an ATP-dependent dynamic equilibrium between two conformations termed outward and inward. In the ATP-unbound state, the C-terminal tail of HypE takes the outward conformation, which is appropriate for receiving carboxamide from a carbamoylphosphate by HypF or for transferring thiocyanate to the HypCD complex. Binding of ATP to HypE will shift the equilibrium toward the formation of the inward conformation, in which HypE dehydrates the S-carboxamide moiety to yield thiocyanate (Watanabe et al, 2007a).

Protein	Function	Organism	**PDB entry and Reference**	**Main Characteristics**	**Conserved Structural domains**
HybD	Endopeptidase Functional homolog of HycI	*E. coli*	**1CFZ** (Fritsche et al, 1999)	Show nickel specific cleavage Not inhibited by EDTA or EGFA	36VxxxDGG42 41GGTA44
HypC	Chaperone Interacts with HypD to enable cynaide tranfer to the active site	*Thermococcus kodakaraensis* KOD1	**2Z1C** (Watanabe et al, 2007b)	Hydrophobic cleft at N-terminus with an acidic molecular surface at the opposite side.	C2 H45
HypB	Nickel transfer to active site	*Methanoclado-* *--coccus jannaschii*	**2BEK** (Gasper et al, 2006)	SIMBI class of protein GTP binding and hydrolyzing Nickel Binding, may contain a histidine rich N-terminus	ENVGNLVCP GSGKTLLIEKL
HypD	Forms a complex with HypE and HypC for cyanide transfer to active site.	*Thermococcus kodakaraensis* KOD1	**2Z1D** (Watanabe et al, 2007b)	FeS clusters Four disulfide bridges	148GFETT152 GPGCPVC CGXH PXHVS
HypE	Forms a complex with HypF, HypD and HypC for cyanide transfer to active site	*Thermococcus kodakaraensis* KOD1 *Desulfovibrio vulgaris* subsp Hildenborough *E. coli*	**2Z1F** (Watanabe et al, 2007b) **2Z1T** (Shomura et al, 2007) **2I6R** (Rangarajan et al, 2008)	purM class of protein, ATP hydrolyzing	68FPGG(D/N)IGxLA(V/I)xGxXNDLAxxGAxP92 130(T/A)GDTKV135 220(K/R)D(P/A)TR(A/G)G226 270ANEG(K/R)274 PR(V/I)C

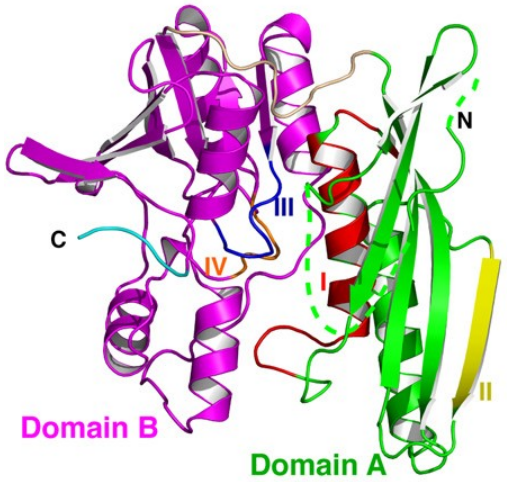
Table 5.2 A list of crystal structures of proteins involved in the maturation of NiFe hydrogenase. Cellular functions and structural features along with conserved domains are also briefly mentioned.



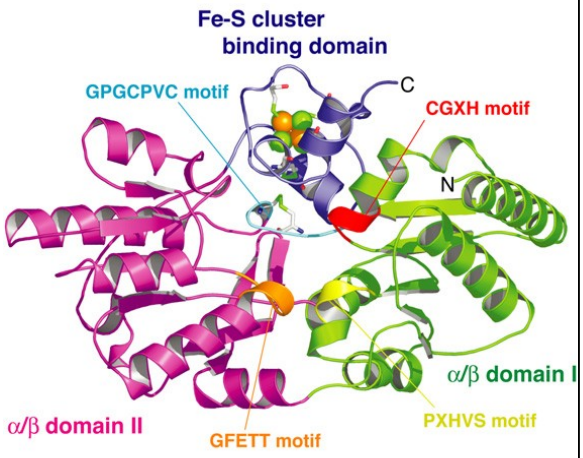
A) HypB of *M. janaschii*



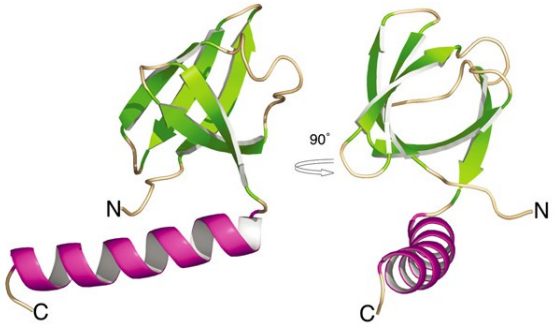
B) HybD of *E. coli*



C) HypE of *T. kodakaraensis*



D) HypD of *T. kodakaraensis*



E) HypC of *T. kodakaraensis*

Fig 5.1 Representative three-dimensional structures of Hyp proteins crystallized from various organisms: **A)** HypB of *M. janaschii* **B)** HypD of *E. coli* **C)** HypE of *T. kodakaraensis* **D)** HypD of *T. kodakaraensis* **E)** the two conformations of HypC of *T. kodakaraensis*: The important residues are marked and references given in table 5.2. The functions are discussed in following sections.

The resolved crystal structure of HypD of *T. kodakaraensis* consists of three major domains: an α/β domain I, an α/β domain II, and an Fe-S cluster binding domain (Fig. 5.1D, Watanabe et al, 2007b). The Fe-S cluster-binding domain carries a [Fe_4S_4] cluster. These three domains form a cleft at the center of the molecule. The residues participating in and around the central cleft are very much conserved. In particular, four conserved motifs are assembled at the bottom of the cleft (Fig. 5.1D, Watanabe et al, 2007b) and are likely to be forming the catalytically active part of the HypD. The disulfide bond SS1 (Cys66-Cys69) lies close to the SS2 (Cys325-Cys354) which in turn is located very close to the [Fe_4S_4] cluster (Fig. 5.1D, Watanabe et al, 2007b), and the sulfur atom of Cys325 in the reduced form makes a close contact with the iron atom. The [Fe_4S_4] cluster environment of HypD as a whole was found to be quite similar to that of ferredoxin:thioredoxin reductase (FTR), and led to the suggestion that a redox cascade is present between the [Fe_4S_4] cluster and two disulfide bonds (Watanabe et al., 2007a and Watanabe et al, 2007b).

The available structures of HypC, HypD and HypE have provided deep insights into the cyanation reaction. In particular, the assembly of conserved residues of HypD around the center cleft suggested that the coordination of the cyanide ligand to the Fe atom takes place at the conserved motifs (Fig. 5.1D and Fig. 5.2). The conserved motifs present in the regions around the central cleft of HypD might be the providing the interface to other Hyp proteins. Fig 5.2 shows a mechanism proposed by Watanabe et al, 2007b, which outlines the stepwise-reaction taking place at the active site of these Hyp proteins.

Further, the conserved motifs of HypC and HypE allow the essential cysteine residues of both proteins to interact with the active site of HypD. The uncovered FTR-like redox cascade implies that the cyanation reaction is catalyzed by a unique thiol redox signaling in the HypCDE complex.

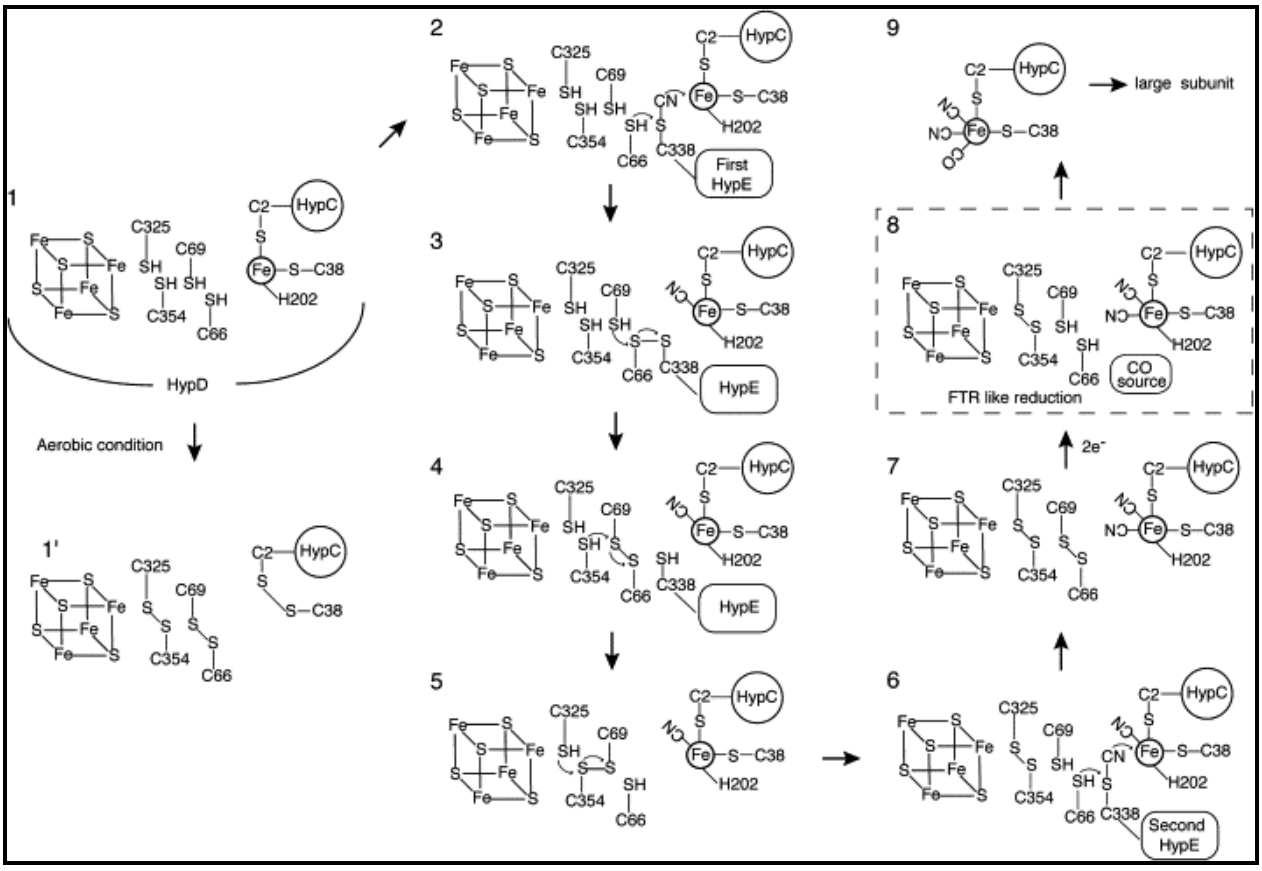


Fig 5.2 A schematic presentation of the mechanism of cyanide transfer at the molecular level as proposed by Watanabe et al, 2007. **(1)** Cys-38 and His-202 residues of HypD trap an iron atom together with the conserved N-terminal cysteine residue of HypC. **(2)** The C-terminal cysteine of HypE is charged with a thiocyanate and interacts with the Cys-66 of HypD in the HypD-HypC complex. **(3)** The cyanide group is transferred to the iron atom. **(4 and 5)** The C-terminal cysteine of HypE is reduced and released. **(6 and 7)** The disulfide bond present in the conserved CPCV motif of HypD is reduced by disulfide bond SS2, and SS1 participates in the next reaction. **(8 and 9)** The second cyanide is transferred in the same manner as the first CN ligand. According to this proposed mechanism, HypD can mediate the CN-transfer reaction twice without the $[Fe_4S_4]$ cluster. In the final steps a CO ligand is transferred to the iron atom from an unknown donor and possibly through the reduction of SS1 by the $[Fe_4S_4]$ cluster in an FTR (ferredoxin:thioredoxin reductase) -like manner.

As in cases of other *Desulfovibrio* sps, the known genes for the maturation for the [NiFe] hydrogenase are present at four different locations in the genome. An attempt to produce the

[NiFe] hydrogenase heterologously should at the minimum need the expression of all these genes together. Alternatively, a very closely related host could be employed to express the structural genes and the corresponding maturation could be complemented via the host machinery. Otherwise all the maturation genes can be co-expressed together with the structural genes. However, such attempts have to ensure that all of these genes are expressed optimally.

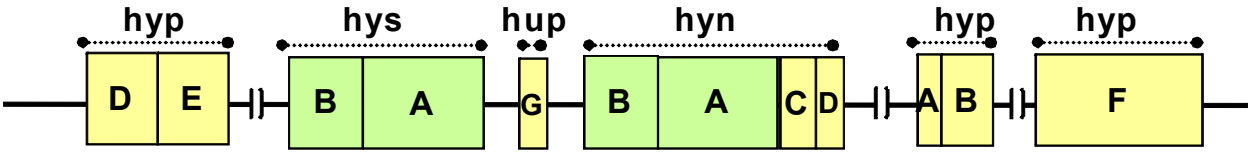


Fig 5.3 Schematic presentation of the location of the *hyn, hys* and *hyp* genes of DvMF. The yellow colored boxes show the maturation genes of DvMF, which have been found to be arranged on four different operons in the genome. The interrupts indicate that these genes are well separated from each in the genome as in case of DvH.

In this work, the maturation proteins of DvMF, those are HynD, HypA, HypB, HypD, and HypF are characterized by their sequence and their secondary structure prediction, and they are compared to their crystallized homologous proteins. Further attempts are reported on expressing the [NiFe] hydrogenase in a heterologous fashion. Firstly, the *hyn1*- [NiFe] deletion mutant of DvH (Goenka et al, 2005) has been tried for such an expression by providing the four gene *hynBACD* operon. Later a strep tag is inserted at the 3' end of the *hynB* gene and also the other two gene operons *hypDE* and *hypAB* have been epitope tagged for expression and purification in *E. coli* cells.

5.1.1 Hydrogenases of DvMF

In this work, genes coding for three hydrogenases are reported in *Desulfovibrio vulgaris* Miyazaki F. Two of them are periplasmic, one [NiFe] - and a [NiFeSe] hydrogenase (Chapter 4). The third one is a membrane associated, hexameric energy converting [Ech] hydrogenase complex (Chapter 4). The structural characteristics of the [NiFe] hydrogenase of DvMF have been discussed in chapter 2, and the structures of [NiFeSe] hydrogenase sequenced in this work

have been characterized in this work by sequential homology (Fig 5.5 and Fig 5.6) and 3D modeling (Fig 5.4).

The amino acid sequences of the selenium-containing hydrogenases yield high homologies with the sequence of the [NiFe] hydrogenases, especially among the cysteines involved in formation of iron sulfur clusters and the residues surrounding them. A three-dimensional model has been created for the [NiFeSe] hydrogenase of DvMF by using the PyMOL program and energy minimization (Fig 5.4). The comparisons between the structures of [NiFe] and [NiFeSe] hydrogenases of DvMF showed an overall similar architecture except for some additional external loops in the small subunit of the [NiFeSe] hydrogenase.

Except of DvMF, [NiFeSe] hydrogenase gene sequences are available from just three organisms so far: *Desulfovibrio vulgaris* subsp Hildenborough (DvH), *Desulfovibrio desulfuricans* subsp G20 (Dd), and *Desulfomicrobium baculatus* (Db). Converted in the corresponding polypeptides of the two subunits of the [NiFeSe] hydrogenases (Hys) are compared with the [NiFe] hydrogenases (Hyn) sequences of DvH and DvMF. A simple sequence alignment of the small subunit with the clustalW program showed a perfect conservation among the cysteines involved in the formation of proximal and distal [Fe_4S_4] clusters among the Hyn and Hys hydrogenases (Fig 5.5a and Fig 5.5b). The number and positioning of cysteines involved in the formation of medial clusters of Hys hydrogenases were similar among themselves, but they are different from those of Hyn. This was as expected an observation as the presence of a [Fe_4S_4] medial cluster in Hys instead [Fe_3S_4] medial cluster of Hyn has been confirmed by structural and EPR data of Hys hydrogenases of DvH (Valente et al, 2005) and Db (Garcin et al, 1999).

However, interestingly the positioning of one of the conserved cysteine residue (Cys-289) of the [NiFeSe] hydrogenase of DvMF which is expected to participate in the formation of the medial cluster is slightly different from the other HysA sequences (Fig 5.5a). In all the HysA sequences except DvMF the first two cysteine participating in the medial clusters are separated by 11 residues, e.g., DvH (Cys-281 and Cys-293), DdG20 (Cys-268 and Cys-280) and Db (Cys-278 and Cys-290) as well as in the Hyn hydrogenases of DvMF (Cys-281 and Cys-293) and DvH (Cys-281 and Cys-293). The corresponding two cysteines present in HysA of DvMF are

separated by 13 residues (Cys-279 and Cys-293), whereas a serine (Ser-281) residue replaces the conserved location of a cysteine (Fig 5.5a). The possible participation of Cys-279 of HysA of DvMF could not be obvious, and moreover, serine has been also known to form iron sulfur clusters, but the nature of these iron sulfur clusters thus formed and the impact on the redox potential needs to be determined.

Further, a similar sequence alignment of the large subunit showed high homologies between Hys and Hyn hydrogenases and an almost perfect conservation among the Hys (Fig 5.6A). The selenocysteine residues present at the C-terminal ends of the Hys hydrogenases are known as the functional replacements of one of the cysteine residue participating in the catalytic active site of the large subunits of [NiFe] hydrogenases (Fig 5.6B).

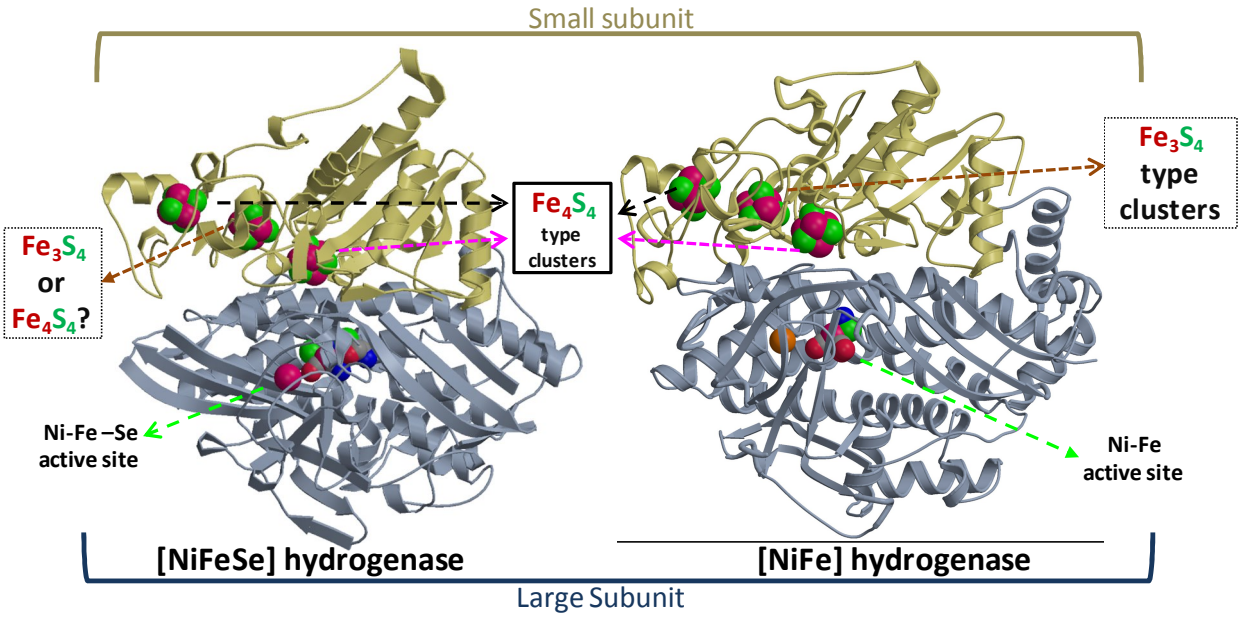


Fig 5.4: A comparison of the three-dimensional structures of the [NiFeSe]- and [NiFe] hydrogenases of DvMF (1HR). The two hydrogenases show an overall high similarity in the subunit architecture. Both hydrogenases have two [Fe_4S_4] clusters (distal and proximal). The [NiFe] hydrogenase possesses a [Fe_3S_4] medial cluster and the nature of the bonding present in the medial cluster in the FeS cluster present in the NiFeSe] hydrogenase is not clear yet.

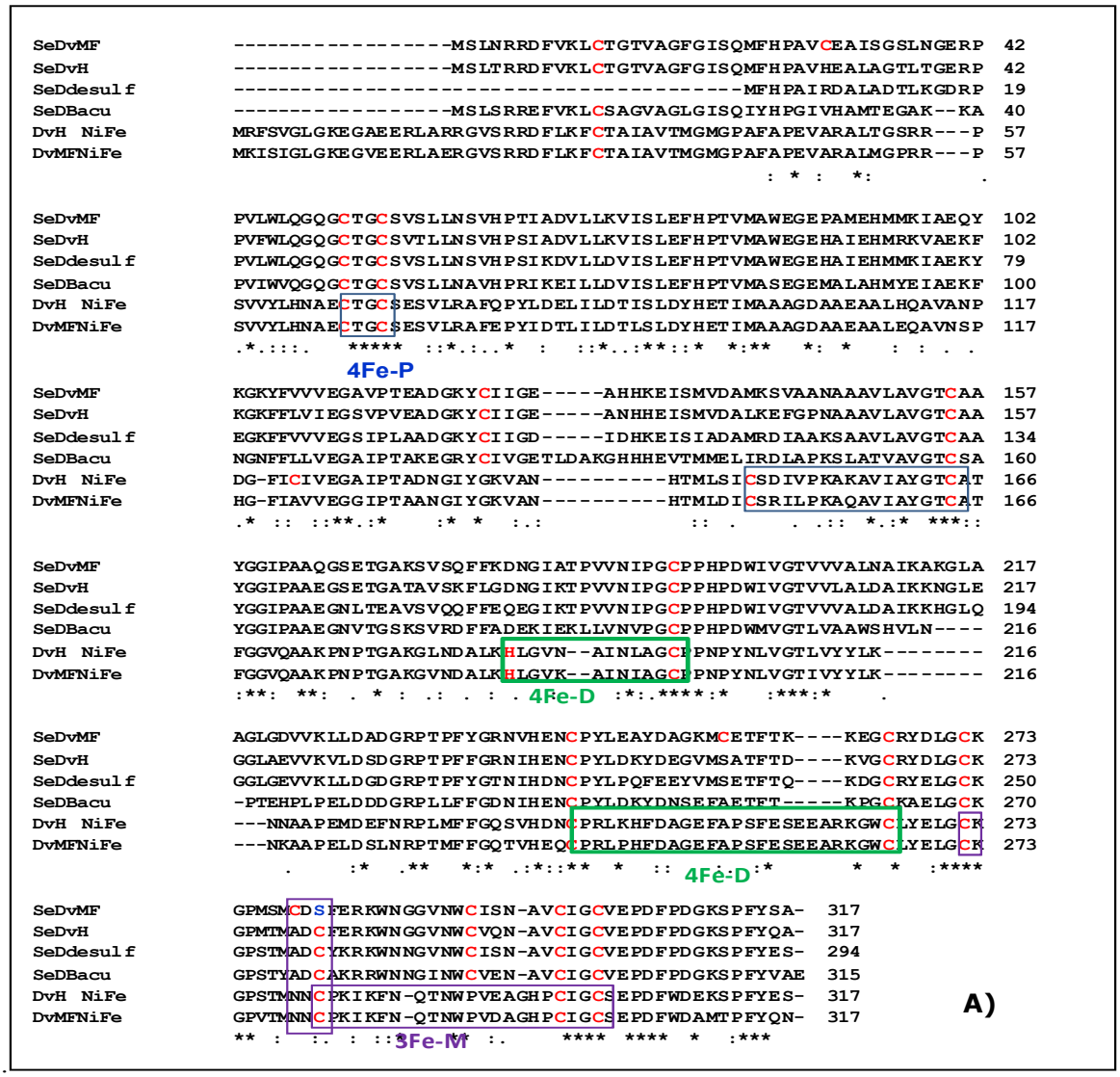
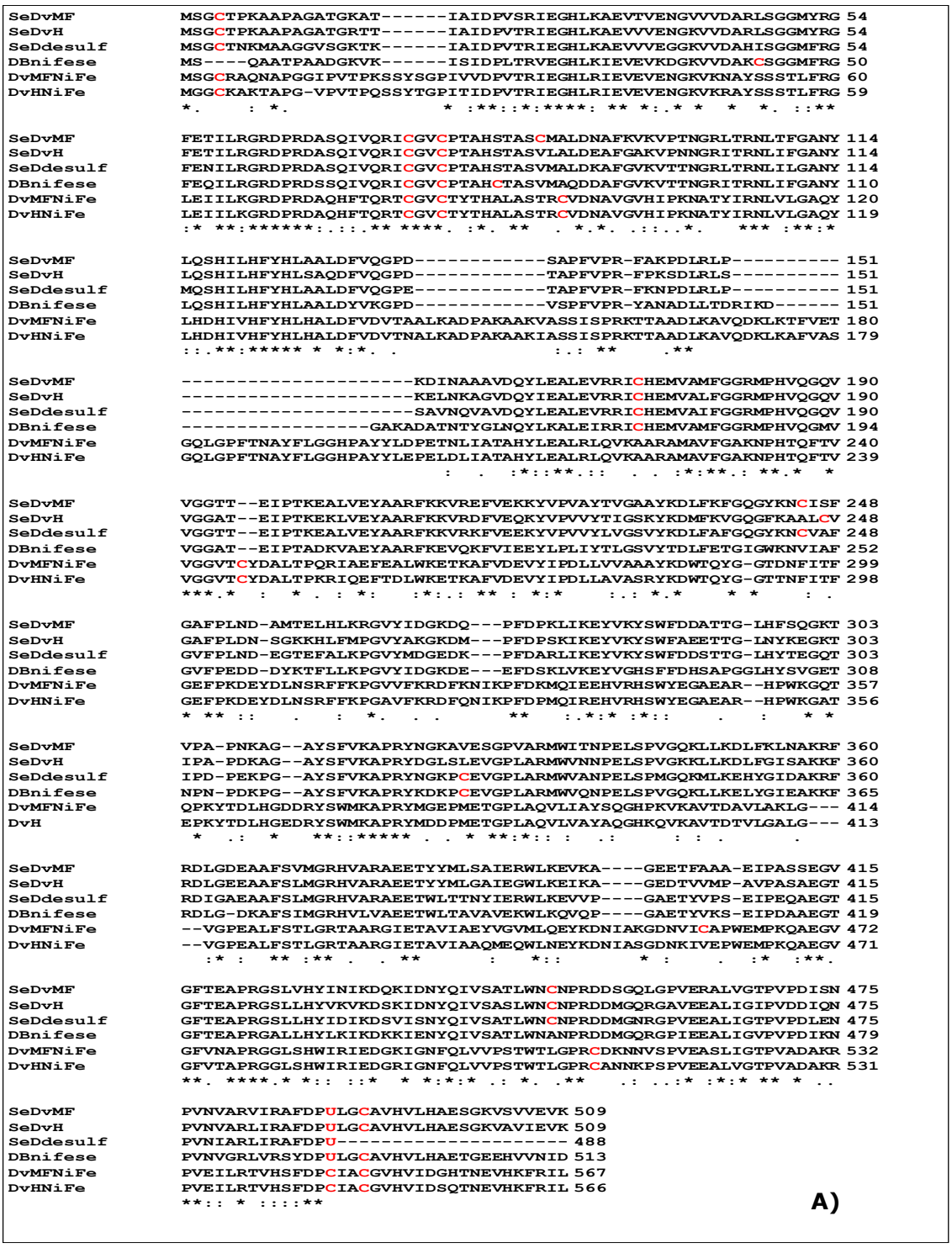
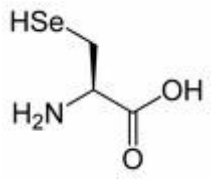
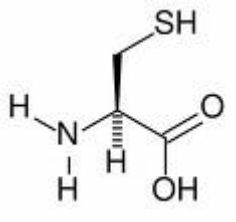


Fig 5.5 A) Polypeptide sequence alignment of the small subunits of the known [NiFeSe] hydrogenases with some [NiFe] hydrogenases. Selenium containing hydrogenases sequences are abbreviated as: SeDvMF– DvMF; SeDvH- DvH; SeDdesulfcn- *Desulfovibrio desulfuricans*; SeDBacu- *Desulfomicrobium baculatus*; NiFe DvHnife- from DvH; DvMFNiFe- from DvMF **B)** Schematic presentation of the ligand structure of $[Fe_4S_4]$- and $[Fe_3S_4]$-type of clusters and the usual participating amino acids in the FeS cluster containing proteins.


SeDvMF MSGCTPKAAPAGATGKAT------IAIDPVSRIEGHLKAEVTVENGVVVDARLSGGMYRG 54
SeDvH MSGCTPKAAPAGATGRTT------IAIDPVTRIEGHLKAEVVVENGKVVDARLSGGMYRG 54
SeDdesulf MSGCTNKMAAGGVSGKTK------IAIDPVTRIEGHLKAEVVVEGGKVVDAHISGGMFRG 54
DBnifese MS----QAATPAADGKVK------ISIDPLTRVEGHLKIEVEVKDGKVVDAKCSGGMFRG 50
DvMFNiFe MSGCRAQNAPGGIPVTPKSSYSGPIVVDPVTRIEGHLRIEVEVENGKVKNAYSSSTLFRG 60
DvHNiFe MGGCKAKTAPG-VPVTPQSSYTGPITIDPVTRIEGHLRIEVEVENGKVKRAYSSSTLFRG 59
FETILRGRDPRDASQIVQRICGVCPTAHSTASCMALDNAFKVKVPTNGRLTRNLTFGANY 114
SeDvH FETILRGRDPRDASQIVQRICGVCPTAHSTASVLALDEAFGAKVPNNGRITRNLIFGANY 114
SeDdesulf FENILRGRDPRDASQIVQRICGVCPTAHSTASVMALDKAFGVKVTTNGRLTRNLILGANY 114
DBnifese FEQILRGRDPRDSSQIVQRICGVCPTAHCTASVMAQDDAFGVKVTTNGRITRNLIFGANY 110
DvMFNiFe LEIILKGRDPRDAQHFTQRTCGVCTYTHALASTRCVDNAVGVHIPKNATYIRNLVLGAQY 120
DvHNiFe LEIILKGRDPRDAQHFTQRTCGVCTYTHALASTRCVDNAVGVHIPKNATYIRNLVLGAQY 119
LQSHILHFYHLAALDFVQGPD----------SAPFVPR-FAKPDLRLP---------- 151
SeDvH LQSHILHFYHLSAQDFVQGPD----------TAPFVPR-FPKSDLRLS---------- 151
SeDdesulf MQSHILHFYHLAALDFVQGPE----------TAPFVPR-FKNPDLRLP---------- 151
DBnifese LQSHILHFYHLAALDYVKGPD----------VSPFVPR-YANADLLTDRIKD------ 151
DvMFNiFe LHDHIVHFYHLHALDFVDVTAALKADPAKAAKVASSISPRKTTAADLKAVQDKLKTFVET 180
DvHNiFe LHDHIVHFYHLHALDFVDVTNALKADPAKAAKIASSISPRKTTAADLKAVQDKLKAFVAS 179
----------KDINAAAVDQYLEALEVRRICHEMVAMFGGRMPHVQGQV 190
SeDvH ----------KELNKAGVDQYIEALEVRRICHEMVALFGGRMPHVQGQV 190
SeDdesulf ----------SAVNQVAVDQYLEALEVRRICHEMVAIFGGRMPHVQGQV 190
DBnifese ----------GAKADATNTYGLNQYLKALEIRRICHEMVAMFGGRMPHVQGMV 194
DvMFNiFe GQLGPFTNAYFLGGHPAYYLDPETNLIATAHYLEALRLQVKAARAMAVFGAKNPHTQFTV 240
DvHNiFe GQLGPFTNAYFLGGHPAYYLEPELDLIATAHYLEALRLQVKAARAMAVFGAKNPHTQFTV 239
VGGTT--EIPTKEALVEYAARFKKVREFVEKKYVPVAYTVGAAYKDLFKFGQGYKNCISF 248
SeDvH VGGAT--EIPTKEKLVEYAARFKKVRDFVEQKYVPVVYTIGSKYKDMFKVGQGFKAALCV 248
SeDdesulf VGGTT--EIPTKEALVEYAARFKKVRKFVEEKYVPVVYLVGSVYKDLFAFGQGYKNCVAF 248
DBnifese VGGAT--EIPTADKVAEYAARFKEVQKFVIEEYLPLIYTLGSVYTDLFETGIGWKNVIAF 252
DvMFNiFe VGGVTCYDALTPQRIAEFEALWKETKAFVDEVYIPDLLVVAAAYKDWTQYG-GTDNFITF 299
DvHNiFe VGGVTCYDALTPKRIQEFTDLWKETKAFVDEVYIPDLLAVASRYKDWTQYG-GTTNFITF 298
GAFPLND-AMTELHLKRGVYIDGKDQ---PFDPKLIKEYVKYSWFDDATTG-LHFSQGKT 303
SeDvH GAFPLDN-SGKKHLFMPGVYAKGKDM---PFDPSKIKEYVKYSWFAEETTG-LNYKEGKT 303
SeDdesulf GVFPLND-EGTEFALKPGVYMDGEDK---PFDARLIKEYVKYSWFDDSTTG-LHYTEGQT 303
DBnifese GVFPEDD-DYKTFLLKPGVYIDGKDE---EFDSKLVKEYVGHSFFDHSAPGGLHYSVGET 308
DvMFNiFe GEFPKDEYDLNSRFFKPGVVFKRDFKNIKPFDKMQIEEHVRHSWYEGAEAR--HPWKGQT 357
DvHNiFe GEFPKDEYDLNSRFFKPGAVFKRDFQNIKPFDPMQIREHVRHSWYEGAEAR--HPWKGAT 356
VPA-PNKAG--AYSFVKAPRYNGKAVESGPVARMWITNPELSPVGQKLLKDLFKLNAKRF 360
SeDvH IPA-PDKAG--AYSFVKAPRYDGLSLEVGPLARMWVNNPELSPVGKKLLKDLFGISAKKF 360
SeDdesulf IPD-PEKPG--AYSFVKAPRYNGKPCEVGPLARMWVANPELSPMGQKMLKEHYGIDAKRF 360
DBnifese NPN-PDKPG--AYSFVKAPRYKDKPCEVGPLARMWVQNPELSPVGQKLLKELYGIEAKKF 365
DvMFNiFe QPKYTDLHGDDRYSWMKAPRYMGEPMETGPLAQVLIAYSQGHPKVKAVTDAVLAKLG--- 414
DvH EPKYTDLHGEDRYSWMKAPRYMDDPMETGPLAQVLVAYAQGHKQVKAVTDTVLGALG--- 413
RDLGDEAAFSVMGRHVARAEETYYMLSAIERWLKEVKA----GEETFAAA-EIPASSEGV 415
SeDvH RDLGEEAAFSLMGRHVARAEETYYMLGAIEGWLKEIKA----GEDTVVMP-AVPASAEGT 415
SeDdesulf RDIGAEAAFSLMGRHVARAEETWLTTNYIERWLKEVVP----GAETYVPS-EIPEQAEGT 415
DBnifese RDLG-DKAFSIMGRHVLVAEETWLTAVAVEKWLKQVQP----GAETYVKS-EIPDAAEGT 419
DvMFNiFe --VGPEALFSTLGRTAARGIETAVIAEYVGVMLQEYKDNIAKGDNVICAPWEMPKQAEGV 472
DvHNiFe --VGPEALFSTLGRTAARGIETAVIAAQMEQWLNEYKDNIASGDNKIVEPWEMPKQAEGV 471
GFTEAPRGSLVHYINIKDQKIDNYQIVSATLWNCNPRDDSGQLGPVERALVGTPVPDISN 475
SeDvH GFTEAPRGSLLHYVKVKDSKIDNYQIVSASLWNCNPRDDMGQRGAVEEALIGIPVDDIQN 475
SeDdesulf GFTEAPRGSLLHYIDIKDSVISNYQIVSATLWNCNPRDDMGNRGPVEEALIGTPVPDLEN 475
DBnifese GFTEAPRGALLHYLKIKDKKIENYQIVSATLWNANPRDDMGQRGPIEEALIGVPVPDIKN 479
DvMFNiFe GFVNAPRGGLSHWIRIEDGKIGNFQLVVPSTWTLGPRCDKNNVSPVEASLIGTPVADAKR 532
DvHNiFe GFVTAPRGGLSHWIRIEDGRIGNFQLVVPSTWTLGPRCANNKPSPVEEALVGTPVADAKR 531
PVNVARVIRAFDPULGCAVHVLHAESGKVSVVEVK 509
SeDvH PVNVARLIRAFDPULGCAVHVLHAESGKVAVIEVK 509
SeDdesulf PVNIARLIRAFDPU---------- 488
DBnifese PVNVGRLVRSYDPULGCAVHVLHAETGEEHVVNID 513
DvMFNiFe PVEILRTVHSFDPCIACGVHVIDGHTNEVHKFRIL 567
DvHNiFe PVEILRTVHSFDPCIACGVHVIDSQTNEVHKFRIL 566
A)

B)





Selenocysteine **Cysteine**

Fig 5.6 A) Sequence alignment of the large subunit of known [NiFeSe] hydrogenases with some of [NiFe] hydrogenases. Selenium containing hydrogenases sequences are abbreviated as: SeDvMF- DvMF; SeDvH- DvH; SeDdesulfcn- *Desulfovibrio desulfuricans*; SeDBacu- *Desulfomicrobium baculatus*; Nickel Iron hydrogenase sequences: DvHnife- DvH; DvMFNiFe- DvMF **B)** Chemical structures of selenocysteine and cysteine molecules.

5.1.2 Expression detection of [NiFe] and [NiFeSe] hydrogenase

We have been regularly using a rabbit serum polyclonal antibody (anti-Hyn) to detect the two subunits of [NiFe] hydrogenase. This antibody is able to detect [NiFe] hydrogenases of other *Desulfovibrio* sps such as DvH, *D. gigas* and also *D. desulfuricans*. Another antibody anti-Hys has been prepared to the purified [NiFeSe] hydrogenase of DvH (the enzyme being a generous gift by Dr. Ines A.C. Pereira of ITQB, Portugal) in a similar way. Western blot detection of the [NiFeSe] hydrogenase of DvMF and DvH total cell lysates showed a relatively poor detection of the two subunits of the [NiFeSe] hydrogenase in these organisms (Fig 5.7). This might be due to the small amount and relatively low purity of the protein used for antibody preparation.

A characterization study of the [NiFeSe] hydrogenase of DvH has previously shown that the two subunits present in the membrane fraction usually show a higher molecular weight than predicted. The detected masses are 63 kDa for the LSU and 35 kDa for the SSU, while the predicted masses are only 54.4 kDa and 31.4 kDa for the LSU and SSU, respectively (Valente et al, 2005). The predicted molecular masses for the two unprocessed subunits for DvMF are 34 kDa for small subunit and 56 kDa for the large subunit (Chapter 3). The western blot of total cell lysates with anti-Hys antibody did detect a protein band in the total cell lysate of DvMF cells, which is parallel to the preferentially detected small subunit of DvH. These antibody-reactive protein bands were also located at higher molecular mass positions (approx 39 kDa) than the predicted mass for both DvH and DvMF.

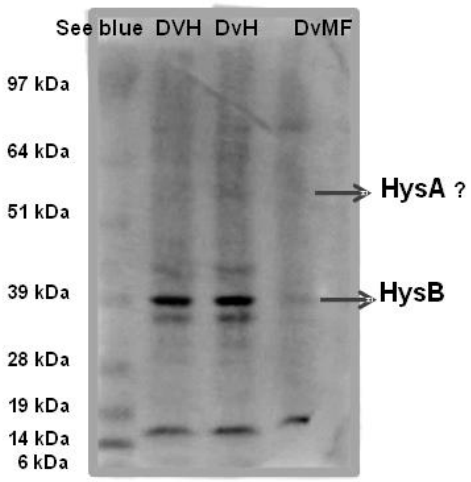


Fig 5.7 Rabbit serum was raised against purified [NiFeSe] hydrogenase from DvH (gift from Ines A.C. Pereira; ITQB, Portugal) and used to detect the expression in DvMF. A preliminary examination showed a lower level of protein in case of DvMF as compared to DvH.

This increase in the molecular mass of processed [NiFeSe] hydrogenase has been considered to be part of its strong interaction and/or covalent association with the lipids, and in fact this protein has been considered as a lipoprotein by some groups (Valente et al, 2005). Biosynthesis of the bacterial lipoproteins is carried out via a highly conserved and distinct pathway in prokaryotes. Conserved signal peptide sequences direct the export of the prolipoprotein, which are processed by the enzyme prolipoprotein diacylglycerol transferase (Lgt). The Lgt uses phospholipid substrates and catalyses the addition of a diacylglycerol unit onto the thiol of an elemental, conserved cysteine which is located within the 'lipobox' motif at the cleavage region of the prolipoprotein signal peptide. After the lipid-modification, a lipoprotein-specific signal peptidase SpaseII, cleaves the signal peptide from the lipoprotein at the conserved cysteine present at the C-terminal end of the lipobox, which later forms the N-terminus of the mature lipoprotein (Hayashi and Wu, 1990).

Some well documented lipobox sequence motifs at the ExPASy-PROSITE, characteristic for lapidating, are described as {DERK}(x6)-[LIVMFWSTAG}(x2)-[LIVMFYSTAGCQ]-[AGS]-C. The permitted lipobox amino acids preceding the invariant cysteine at positions -1 to -4 are indicated by the lack of charged residues (no D, E, R, K) within the immediate hydrophobic region following. A study by Valente et al (2007) has suggested [NiFeSe] hydrogenase as a lipoprotein as evident by globomycin induced transcription inhibition. Also there is a partial lipobox of four conserved hydrophobic residues MSGC in the N-terminus of the large subunit,

but still it is more likely to be transported via tat pathway (Fig 5.9), as it lacks any typical N-terminal signal sequence for secretory pathways.

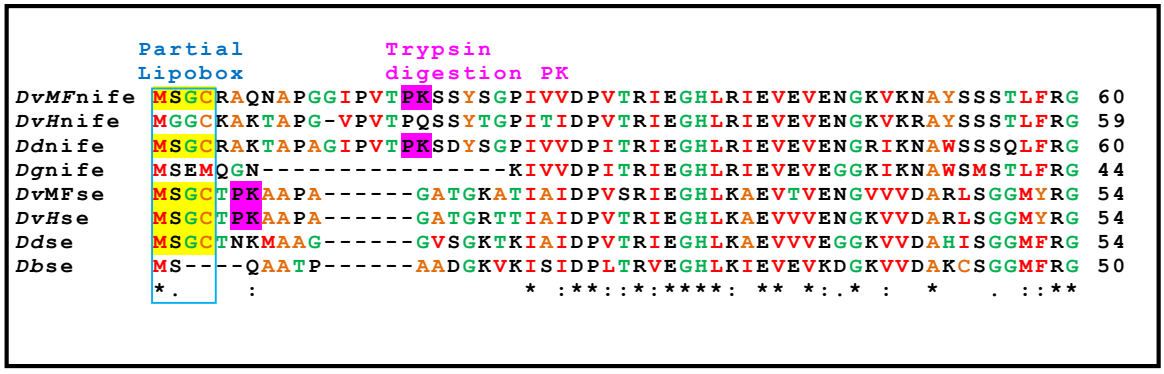


Fig 5.8 Schematic comparison of the N-terminal sequences of the large subunits of the periplasmic-hydrogenases from some *Desulfovibrio* species. These ends may possess a consensus sequence, MSGC (underlaid in yellow and squared in blue), which can be considered as a partial lipo-box. The N-terminals for most of these sequences are rich in hydrophobic residues. The highly hydrophobic residues are marked in red (Phe, Ile, Trp, Leu, Val, Met), lesser in orange (Tyr, Cys, Ala) and the neutral ones (Thr, His, Gly, Ser, Gln) in green; the hydrophobic residues may play an important role in the anchoring of the hydrogenase to the plasma membrane. The trypsin digestion positions (PK) are highlighted in pink, which in the case of the [NiFe] hydrogenase of DvMF is used for isolation of the protein. Abbreviations used: *DvMFnife*, [NiFe] Hase of DvMF; *DvHnife*,[NiFe] Hase of DvH; *Ddnife*,[NiFe] Hase of *D. desulfuricans* G20, Dgnife,[NiFe] Hase of *D. gigas*; *DvMFse*,[NiFeSe] Hase of DvMF; *DvHse*,[NiFeSe] Hase of DvH; *Ddse*, [NiFeSe] Hase *of D. desulfuricans* G20, *Dbse*, [NiFe] Hase of *D. baculatus*.

The tat and the four residue-conserved lipobox proposed for DvH [NiFeSe] hydrogenase is also conserved in the DvMF [NiFe] and [NiFeSe] hydrogenase. Yet a hydrophilic residue (arginine-5) is present following cysteine (Cys-4) in case of [NiFe] hydrogenase (Fig 5.8), which further lowers the chances to be recognized by a lipid peptidase. We have also compared the tat pathway sequences present in the N-terminal part of the small subunit with other conserved sequences and also with the N-terminus of the small subunits of the [NiFeSe] hydrogenase of DvMF (Fig 5.9). Although the RRDFLK sequence present in the N-terminal of the [NiFe] hydrogenase is exactly conserved in the [NiFeSe] hydrogenase, there are only four residues present before the twin arginines of [NiFeSe] hydrogenase, which makes its N-terminal extension 19 residues shorter than that of the [NiFe] hydrogenase.

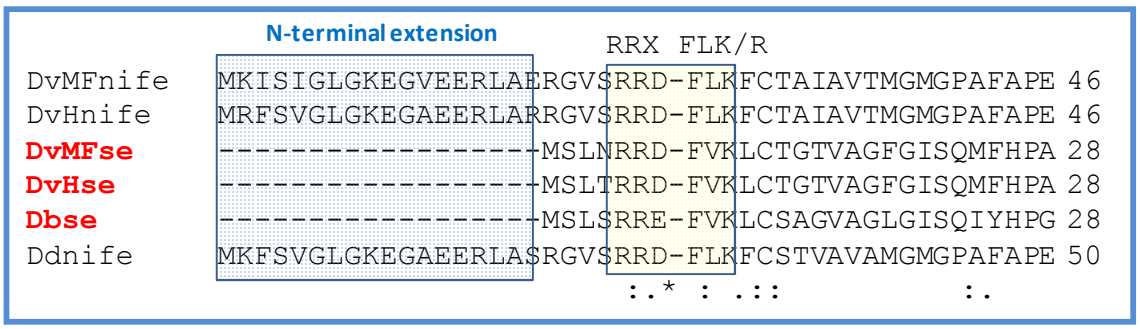


Fig 5.9 The N-terminal parts of the small subunit of the [NiFeSe] hydrogenases of various Desulfovibrio species are compared for the conservation of the twin arginine motif RRxFLK as well as the N-terminal extension present before this conserved motif. Shown in order are DvMFnife [NiFe] Hase of DvMF; *DvHnife*, [NiFe] Hase of DvH; *Ddnife*, [NiFe] Hase of *D. desulfuricans* G20; *DvMFse*, [NiFeSe] Hase of DvMF; *DvHse*, [NiFeSe] Hase of DvH; *Ddse*, [NiFeSe] Hase *of D. desulfuricans* G20; *Dbse*, [NiFe] Hase of *D. baculatus*.

5.1.3 The regulatory effect of selenium on periplasmic hydrogenases

It was shown for DvH that the [NiFeSe] hydrogenase becomes the dominant hydrogenase when the growth medium is simultaneous supplied with 1 µM each of selenium and nickel. Though to a lesser extent, the expression studies on the transcription level have shown a similar up-regulation of [NiFeSe] hydrogenase, when hydrogen at a concentration lower than 5% is the only available electron donor, even without additional supply of nickel and selenium (Caffrey 2007, Pereira et al, 2007, Pereira et al, 2008). Microarray analyses have shown a complex network of gene regulation responsible for the up-regulation of various hydrogenases with different growth media and hydrogen concentrations (Pereira et al, 2008). Whereas the genes-encoding for the [NiFeSe] and [NiFe] Hases have been found to be adjacently placed in DvH, *D. desulfuricans*, a possible co-regulation of the genes is holding more ground with detection of similar genetic arrangements of these two hydrogenases in DvMF. In fact, we have found that the role of cellular availability of selenium and its regulatory effect is also true for DvMF.

The effect of the availability of nickel and selenium on the cellular expression of the periplasmic hydrogenases of DvMF level was compared in parallel to the periplasmic hydrogenases of *D. desulfuricans,* subsp. *desulfuricans* ATCC 7757, in this work. The cells were grown in Postgate C media prepared from chelexed water thus devoid of any salt overload. The growth media were

occasionally supplemented with either $NiCl_2$ (1 µM) or Na_2Se (1 µM) or both. Conspicuously, the availability of these elements caused a significant variation in the expression of [NiFe] hydrogenase: when nickel was supplied, the level of [NiFe] hydrogenase increases in both DvMF (Fig. 5.10B, lane 3) and *D. desulfuricans* (Fig. 5.10B, lane 7). However, the additional supplementation with selenium even at the level of the nickel availability decreases the expression of the [NiFe] hydrogenases to a level even lower than the basal expression level that was detectable without the additional supplementation with nickel (Fig 5.10B, lanes 4 and 8). The expression level is similarly lowered, when only selenium is supplied. The effect of supplied metals was almost similar in both DvMF (Fig. 5.10B, lane 4) and *D. desulfuricans* (Fig. 5.10B, lane 9).

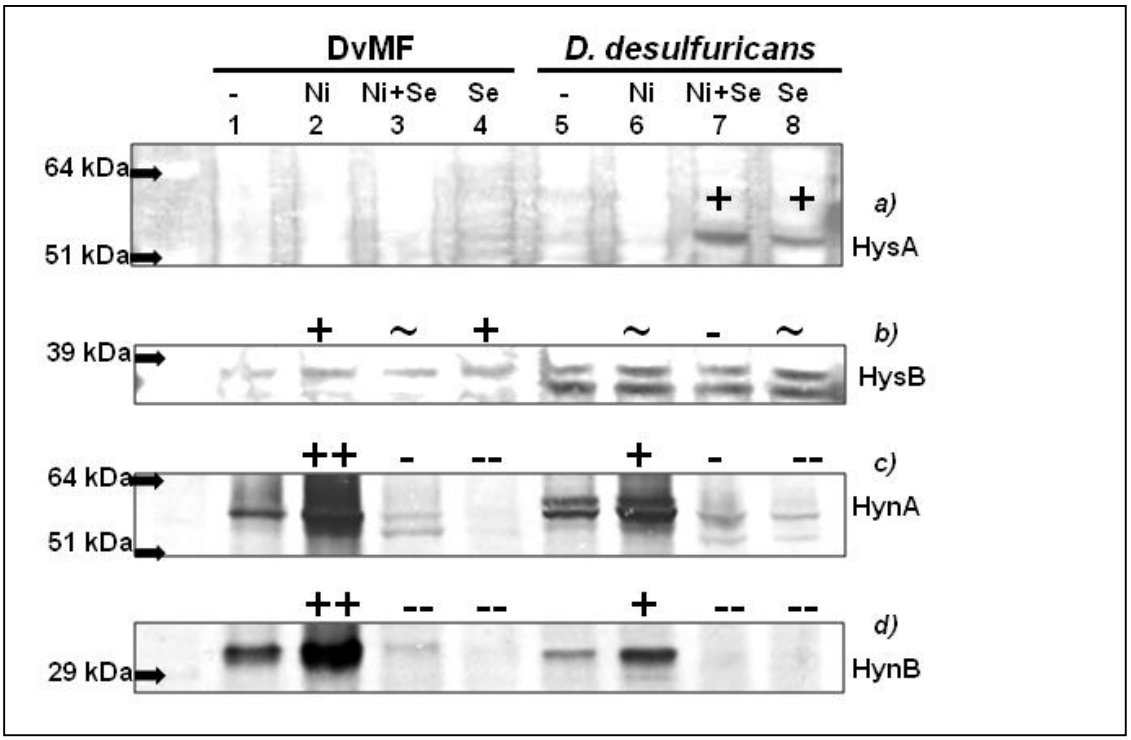


Fig 5.10 Western blot detections of the total cell lysates of the DvMF (lanes 1-4) and *D. desulfuricans* (lanes 5-8), grown on Postgate C medium supplemented with 1 µM $NiCl_2$ (lanes 2, 3, 4 and 7) and 1 µM Na_2Se (lanes 3, 4, 7 and 8). (a & b) A small to moderate increase in the expression of the two subunits of [NiFeSe] hydrogenase, with nickel and selenium supplementation is seen for both DvMF and *D. desulfuricans*. (c & d) There was a clear up-regulation in the expression of the [NiFe] hydrogenase when supplemented with nickel in both the organisms. On the other hand, there is a clear down-regulation when supplied with 1 µM selenium both with and without the presence of nickel.

The same protein samples were tested against the anti-Hys antibody in a parallel western blot. It did not indicate any up-regulation of the small subunit of the [NiFeSe] hydrogenase in DvMF, when either of nickel and selenium or both was supplied as compared to the basal level detected (Fig 5.10A, lanes 2-5). In case of the expression detection in *D. desulfuricans*, there was increased expression in the presence of selenium, irrespective of the presence or absence of nickel (Fig 5.10A lanes 6-9). It is noteworthy that the western blot detected only the small subunit of the [NiFeSe] hydrogenase of DvMF in all the conditions tested; the large subunit of this hydrogenase from *D. desulfuricans* was also detectable when selenium was supplied. In any of these experiments, any strong changes in the band intensities were not that conspicuous as was in case of the detection with the anti-Hyn antibody. However, the poor sensitivity of the anti-Hys antibody was consistent.

5.1.4 Native-PAGE/Hydrogenase activity assay

In order to establish the existence and the regulation of the two periplasmic hydrogenases (by selenium) independent of antibody sensitivity, the crude cell lysates of DvMF were assessed by MV-TTC assay after native gel electrophoresis. The DvMF cells were grown in lactate supplement minimal Postgate C medium prepared with chelexed water. One growth culture was supplement with 1 μM $NiCl_2$ and a parallel second growth culture was supplemented with 1 μM $NiCl_2$ and 1 μM Na_2Se.

The cell lysates were prepared with glass bead homogenization in a buffer supplied with 1% triton-X100, and the total protein extracts were separated by electrophoresis in native form in a 3-8% gradient tris-acetate gel (details given in chapter 3). The detergent was included to ensure that both the membrane and soluble fractions are solubilized and assayed together in this experiment. The catalytically active hydrogenase shows up due to the reduction of methyl viologen yielding blue colored bands, and the reduced methyl viologen was oxidized by TTC (triphenyl tetrazolium chloride) to yield pink bands (Fig 5.11).

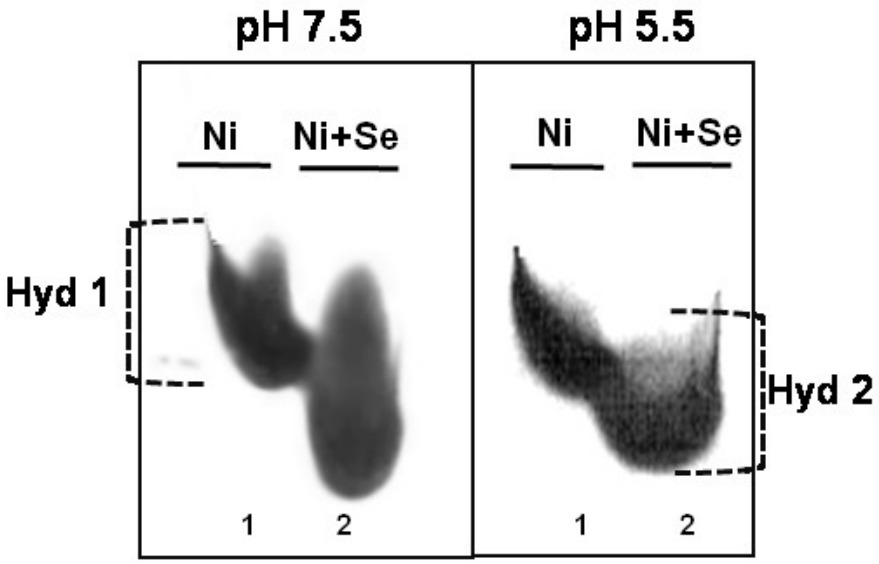


Fig 5.11 Native gel hydorgenase BV/TTC assay of DvMF cell lysate at pH 7.5 and pH 5.5. Crude cell lysates of DvMF cells were prepared by homogenizing the cells grown on a minimal medium postgate C. The samples loaded in lane 1 were from cells which were supplied with 1 µM $NiCl_2$, and the samples loaded in lane 2 were supplied with both 1 µM $NiCl_2$ and 1 µM Na_2Se. The activity assays were performed at pH 7.5 and pH 5.5, and always under an atmosphere of hydrogen. The hydrogenase activity was tested by supplying 1 µM Methylviologen in 50 mM Phosphate buffer and the reaction was stopped adding 1 µM TTC. The hydrogen-active band (indicated as hydrogenase 1) was down-regulated in presence of selenium while hydrogenase 2 was detected only in presence of selenium.

In case of the protein extracts prepared from cells grown only with nickel supplement, only one activity band appears (marked as hydrogenase 1). The intensity of the hydrogenase 1 activity band was reduced significantly in protein extracts prepared from cells grown both with nickel and selenium supplements. More importantly, however, a second activity band appears in this case (marked as hydrogenase 2). These data excellently corroborate with the results obtained from the western blotting (Fig 5.10). Taking into account the clear proof for the absence of a [FeFe] hydrogenase-coding gene in the DvMF genome or in the cosmid library, it is concluded that hydrogenase 1 and hydrogenase 2 are the [NiFe] and [NiFeSe] hydrogenases of DvMF (Fig 5.11), respectively.

5.2. Heterologous expression of [NiFe] hydrogenase in *Desulfovibrio*

An attempt to generate a [NiFe] hydrogenase deletion mutant of DvMF has not been successful previously; however such a deletion mutant of the [NiFe] hydrogenase, hyn1 for DvH is available (Goenka et al, 2005). Since these two organisms are closely related, the maturation machinery of DvH should be able to complement the structural gene products of the [NiFe] hydrogenase of DvMF. The four-gene operon *hynABCD* was cloned in the HindIII restriction site of a RSF1010 replicon-based vector pJRD215 to produce the recombinant plasmid pJRD215hynABCD (Fig 5.12). The pJRD215 plasmid harbors a kanamycin and a streptomycin resistance and is one of the few vectors available, which are capable of replicating in various *Desulfovibrio* sps.

The recombinant vector pJRD215hynABCD was conjugated with the deletion mutant via *E. coli* S17-1 cells (Simon et al, 1991). Though the *Desulfovibrio* sps are known to have high antibiotic resistance, it was assumed that the conjugated deletion mutant might have some selective advantage while growing on high concentrations of kanamycin and streptomycin. After conjugation, the cells were plated on PE media containing 500 µg/µl kanamycin and 500 µg/µl streptomycin. The ex-conjugate colonies started appearing on PE plates after three days of incubation at 37°C. The colonies were selected and grown in PC medium. Attempts to test them by PCR and western blotting for the presence of the plasmid and its expression did not yield any positive results. It appears that these cells are naturally resistant to high antibiotic concentration. Even as later the concentrations of the two antibiotics were increased up-to 600 µg/µl, it did not allow selection of cells harboring the plasmid.

As the introduction of the kanamycin and streptomycin resistance carrying plasmid did not give any selection advantage to the recipient DvH deletion mutant, the use of an alternate host was anticipated.

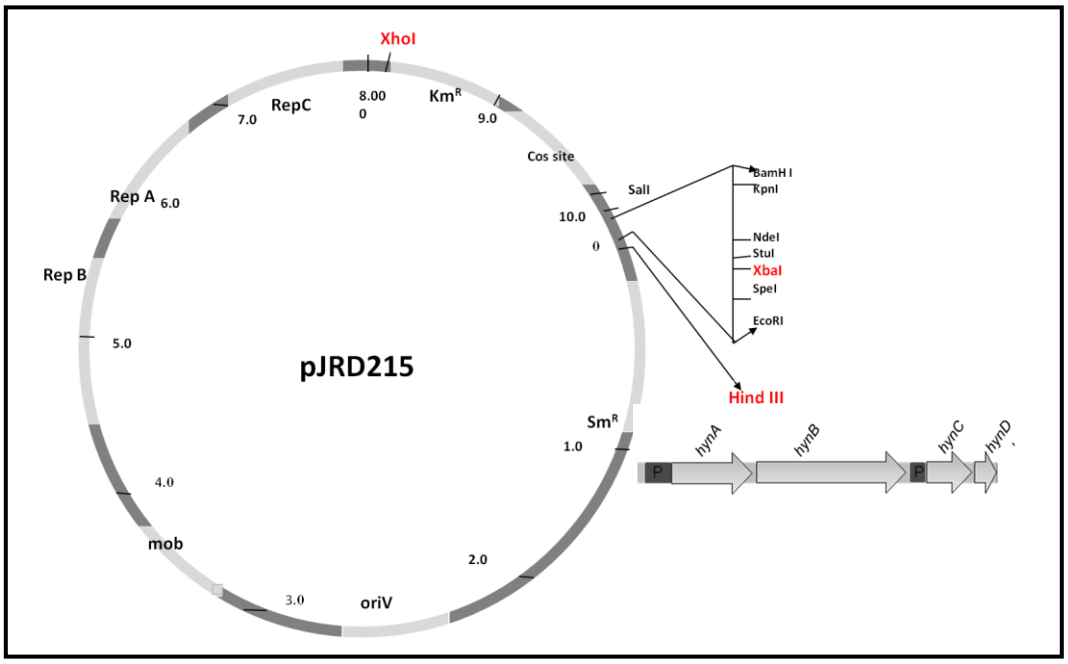


Fig 5.12 Schematic presentation of the plasmid pJRD215 and the four-gene operon *hynABCD* cloned into the HindIII site to generate pJRD215hynABCD.

To overcome this, an expression system needs to be developed which can lead to the expression in a host that is easy to handle and to manipulate such as *E. coli*. First of all, the proteins coded by the Hyp maturation genes sequenced in this work (Chapter 4) were analyzed and compared sequentially with their better characterized homologs. Secondly, the naturally occurring operons in the DvMF genome were expressed separately in *E. coli* by employing a strong 'T7- derived promoter' carrying plasmid.

5.3 Maturation Proteins of DvMF

5.3.1 The nickel-transferring proteins HypA and HypB

5.3.1.1 HypA of DvMF

The HypA protein of *E. coli* is involved in the maturation of hydrogenase 3 while its homolog HybF protein is involved in the maturation of hydrogenase 2 and hydrogenase 1 (Hube et al, 2002). The [NiFe] hydrogenase of DvMF is sequentially more related to hydrogenase 1 and 2 of *E. coli* rather than to the hydrogenase 3. The HybF polypeptide of *E. coli* also shares significant

similarities with HypA (48% homologies and 32% identities) of DvMF. Both HypA and HybF of *E. coli* are cysteine-rich proteins which show some nickel- or zinc binding properties (Atanassova and Zamble, 2005). The HypA of DvMF has conserved most of the cysteines (Fig 5.13).

```
HypA DvMF      MHEMSVASSVISIVQEELAKHSATRLLLVRVRFGALANLVPEAMEFAFEAMTIGTDFEGA 60
               MHE+S+   S + I+Q +   +H    R+   V +   GAL+ +    A+ F+FE +   GT   +G
HybF E. coli.  MHELSLCQSAVEIIQRQAEQHDVKRVTAVWLEIGALSCVEESAVRFSFEIVCHGTVAQGC 60

HypA DvMF      KLELAMAPVKLQCGGCGKEFEPQGKELLFAPCPAC-GEEAGHHVLAGRELYVEHIEAE  117
                  +    P +   C   C +   E     +    A CP C GE      V   G   L V+ IE E
HybF E. coli.  DLHIVYKPAQAWCWDCSQVVEIHQHD---AQCPICHGERL--RVDTGDSLIVKSIEVE  113
```

Fig 5.13 Primary sequence alignment of the HypA of DvMF and the HybF protein of *E. coli* show the conserved cysteines and histidines.

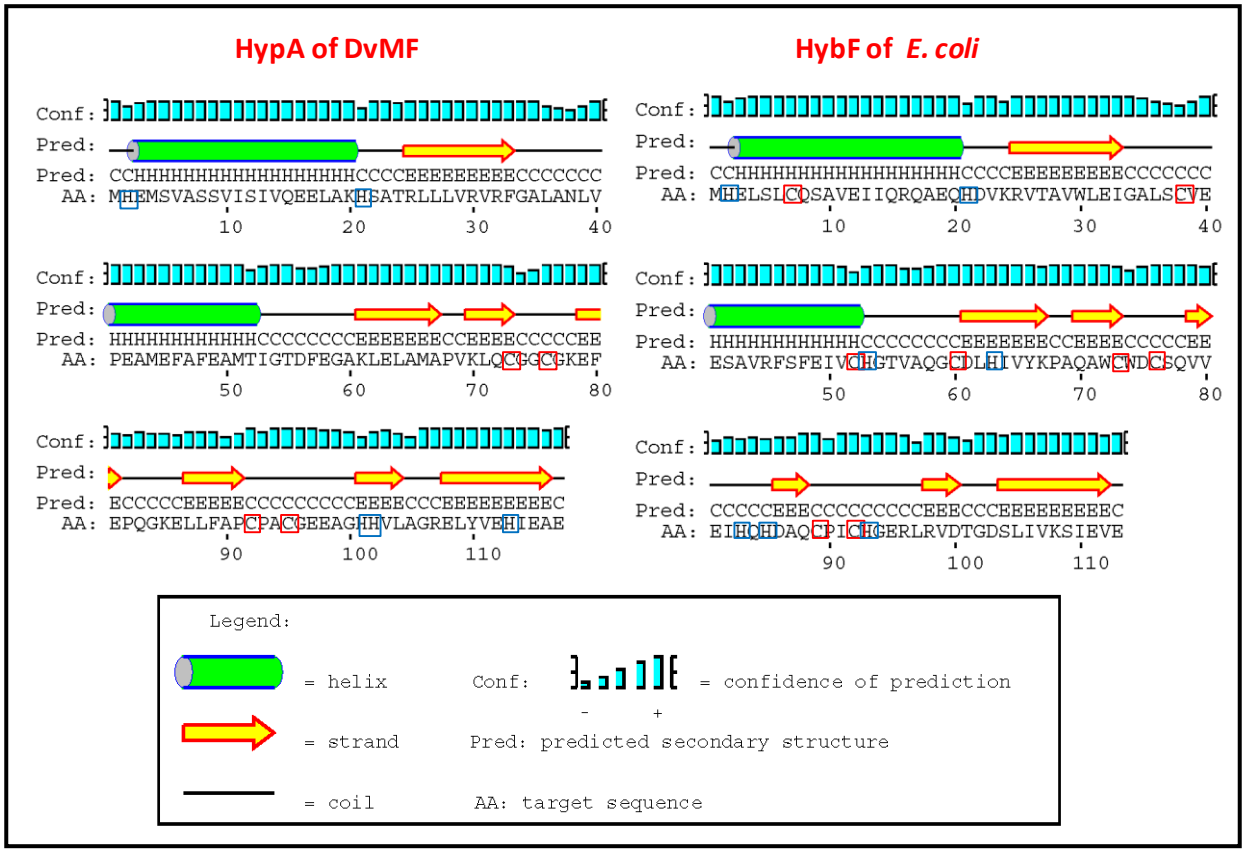


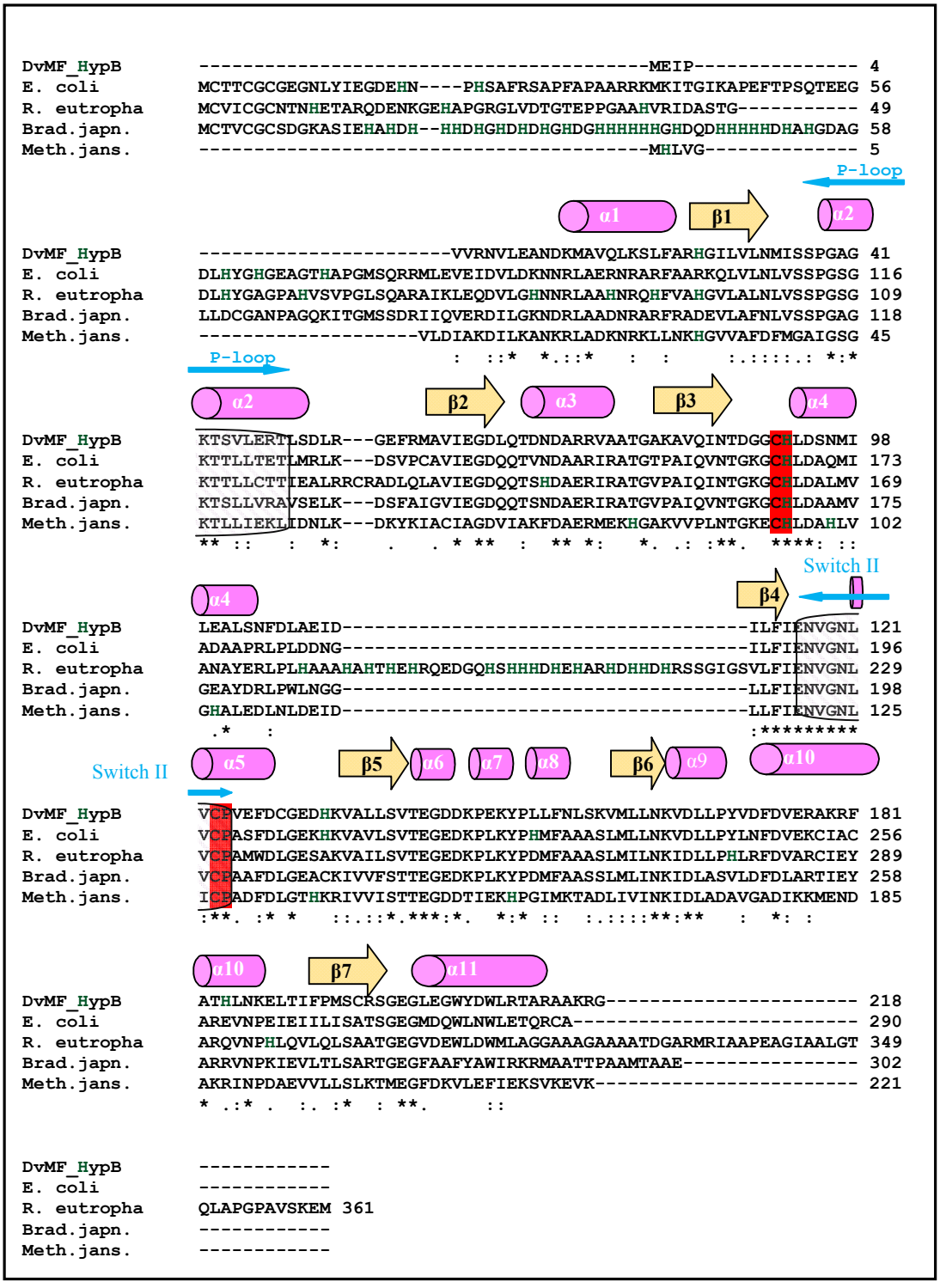
Fig 5.14 Secondary structure predictions for HypA of DvMF and HybF of *E. coli* using PsiPred. The cysteines are marked in red boxes and histidine in blue. The result showed a similar arrangement of α-helical and β-strand domains as found for hybF from *E. coli*.

5.3.1.2 HypB of DvMF

HypB purification and sequential characterization studies from various organisms have revealed that this protein consists of two main domains, one is a metal binding and the other is a GTPase- or G-domain. Further, based on sequence comparison and the metal binding properties, the HypB belonging to prokaryotic organisms is divided into three different subgroups (Gasper et al, 2006). First group, found, e.g. in *Bradyrhizobium japonicum*, possesses an N-terminal poly-histidine stretch in addition to the G-domain. The second group, which also includes the HypB from *E. coli*, has an extended N-terminus, but does not contain any poly-histidine stretches. The third group is mainly composed of HypB proteins from thermophilic organisms e.g. *Methanocaldococcus jannaschii*; here this protein lacks the N-terminal part and only a short stretch of approximately twenty amino acids is present before the G-domain.

Based on the N-terminus, the HypB from DvMF falls in the category of *B. japonicum*, however, it does not possess any poly-histidine stretches. A ClustalW sequence alignment of HypB polypeptide sequences of DvMF, *E. coli*, *Ralstonia eutropha*, *Bradyrhizobium japonicum* and *Methanocladococcus janaschii* was performed (Fig 5.15). This alignment was also compared with the secondary structure of HypB of *M. janaschii* (Gasper et al, 2006). The overall primary sequence alignment revealed that barring the length of the N-terminal extension, HypB of DvMF possesses highest homology with the HypB of *E. coli*.

Overall, there are high homologies in the conserved structural and functional domains in all of the compared HypB sequences including that of DvMF. The high homology between these proteins allowed to prepare a three-dimensional model for the HypB of DvMF (Fig 5.16). These proteins share a topology to the SIMBI (Signal Recognition Particle, MinD and BioD) class of GTP-binding proteins, which as a feature undergo nucleotide dependent dimerization. The nucleotides (one per molecule) are present at the interface of such a dimer. The functionally important motifs indicated in the 3D structure of *M. janaschii* have been well conserved in HypB of DvMF, which included metal binding sites $_{91}$CH$_{92}$ and $_{123}$CP$_{124}$, the characterized (Gasper et al, 2006) switch II relay motif ENVGNLVCP as well as the loop forming GAGKTSVLERT motifs.




```
DvMF_HypB       ----------------------------------------MEIP--------------- 4
E. coli         MCTTCGCGEGNLYIEGDEHN----PHSAFRSAPFAPAARRKMKITGIKAPEFTPSQTEEG 56
R. eutropha     MCVICGCNTNHETARQDENKGEHAPGRGLVDTGTEPPGAAHVRIDASTG----------- 49
Brad.japn.      MCTVCGCSDGKASIEHAHDH--HHDHGHDHDHGHDGHHHHHHGHDQDHHHHHDHAHGDAG 58
Meth.jans.      ----------------------------------------MHLVG-------------- 5

DvMF_HypB       ----------------------VVRNVLEANDKMAVQLKSLFARHGILVLNMISSPGAG 41
E. coli         DLHYGHGEAGTHAPGMSQRRMLEVEIDVLDKNNRLAERNRARFAARKQLVLNLVSSPGSG 116
R. eutropha     DLHYGAGPAHVSVPGLSQARAIKLEQDVLGHNNRLAAHNRQHFVAHGVLALNLVSSPGSG 109
Brad.japn.      LLDCGANPAGQKITGMSSDRIIQVERDILGKNDRLAADNRARFRADEVLAFNLVSSPGAG 118
Meth.jans.      --------------------VLDIAKDILKANKRLADKNRKLLNKHGVVAFDFMGAIGSG 45
                                     :  ::*  *.::*   :  :     :.::::.: *:*

DvMF_HypB       KTSVLERTLSDLR---GEFRMAVIEGDLQTDNDARRVAATGAKAVQINTDGGCHLDSNMI 98
E. coli         KTTLLTETLMRLK---DSVPCAVIEGDQQTVNDAARIRATGTPAIQVNTGKGCHLDAQMI 173
R. eutropha     KTTLLCTTIEALRRCRADLQLAVIEGDQQTSHDAERIRATGVPAIQINTGKGCHLDALMV 169
Brad.japn.      KTSLLVRAVSELK---DSFAIGVIEGDQQTSNDAERIRATGVPAIQVNTGKGCHLDAAMV 175
Meth.jans.      KTLLIEKLIDNLK---DKYKIACIAGDVIAKFDAERMEKHGAKVVPLNTGKECHLDAHLV 102
                ** ::   :  *:    .   . * **  :  ** *:   *. .: :**.  ****: ::

DvMF_HypB       LEALSNFDLAEID-------------------------------------ILFIENVGNL 121
E. coli         ADAAPRLPLDDNG-------------------------------------ILFIENVGNL 196
R. eutropha     ANAYERLPLHAAAHAHTHEHRQEDGQHSHHHDHEHARHDHHDHRSSGIGSVLFIENVGNL 229
Brad.japn.      GEAYDRLPWLNGG-------------------------------------LLFIENVGNL 198
Meth.jans.      GHALEDLNLDEID-------------------------------------LLFIENVGNL 125
                 .*   :                                            :*********

DvMF_HypB       VCPVEFDCGEDHKVALLSVTEGDDKPEKYPLLFNLSKVMLLNKVDLLPYVDFDVERAKRF 181
E. coli         VCPASFDLGEKHKVAVLSVTEGEDKPLKYPHMFAAASLMLLNKVDLLPYLNFDVEKCIAC 256
R. eutropha     VCPAMWDLGESAKVAILSVTEGEDKPLKYPDMFAAASLMILNKIDLLPHLRFDVARCIEY 289
Brad.japn.      VCPAAFDLGEACKIVVFSTTEGEDKPLKYPDMFAASSLMLINKIDLASVLDFDLARTIEY 258
Meth.jans.      ICPADFDLGTHKRIVVISTTEGDDTIEKHPGIMKTADLIVINKIDLADAVGADIKKMEND 185
                :**. :* *   ::.::*.***:*.  *:* ::  :.::::**:**   :  *: :

DvMF_HypB       ATHLNKELTIFPMSCRSGEGLEGWYDWLRTARAAKRG---------------------- 218
E. coli         AREVNPEIEIILISATSGEGMDQWLNWLETQRCA-------------------------- 290
R. eutropha     ARQVNPHLQVLQLSAATGEGVDEWLDWMLAGGAAAGAAAATDGARMRIAAPEAGIAALGT 349
Brad.japn.      ARRVNPKIEVLTLSARTGEGFAAFYAWIRKRMAATTPAAMTAAE---------------- 302
Meth.jans.      AKRINPDAEVVLLSLKTMEGFDKVLEFIEKSVKEVK----------------------- 221
                * .:* .  :. :*  : **.     ::

DvMF_HypB       ------------
E. coli         ------------
R. eutropha     QLAPGPAVSKEM 361
Brad.japn.      ------------
Meth.jans.      ------------
```

Fig 5.15 The multiple sequence alignments of HypB from DvMF, *E. coli*, *Ralstonia eutropha*, *Bradyrhizobium japonicum* and *Methanocladococcus janaschii*, positioned parallel to the secondary structures domains as found in the crystal structure of *M. janaschii*. The two hatched boxes show the conserved domains GAGKTSVLERT and ENVGNLVCP, the red shaded residues show the metal binding sites.

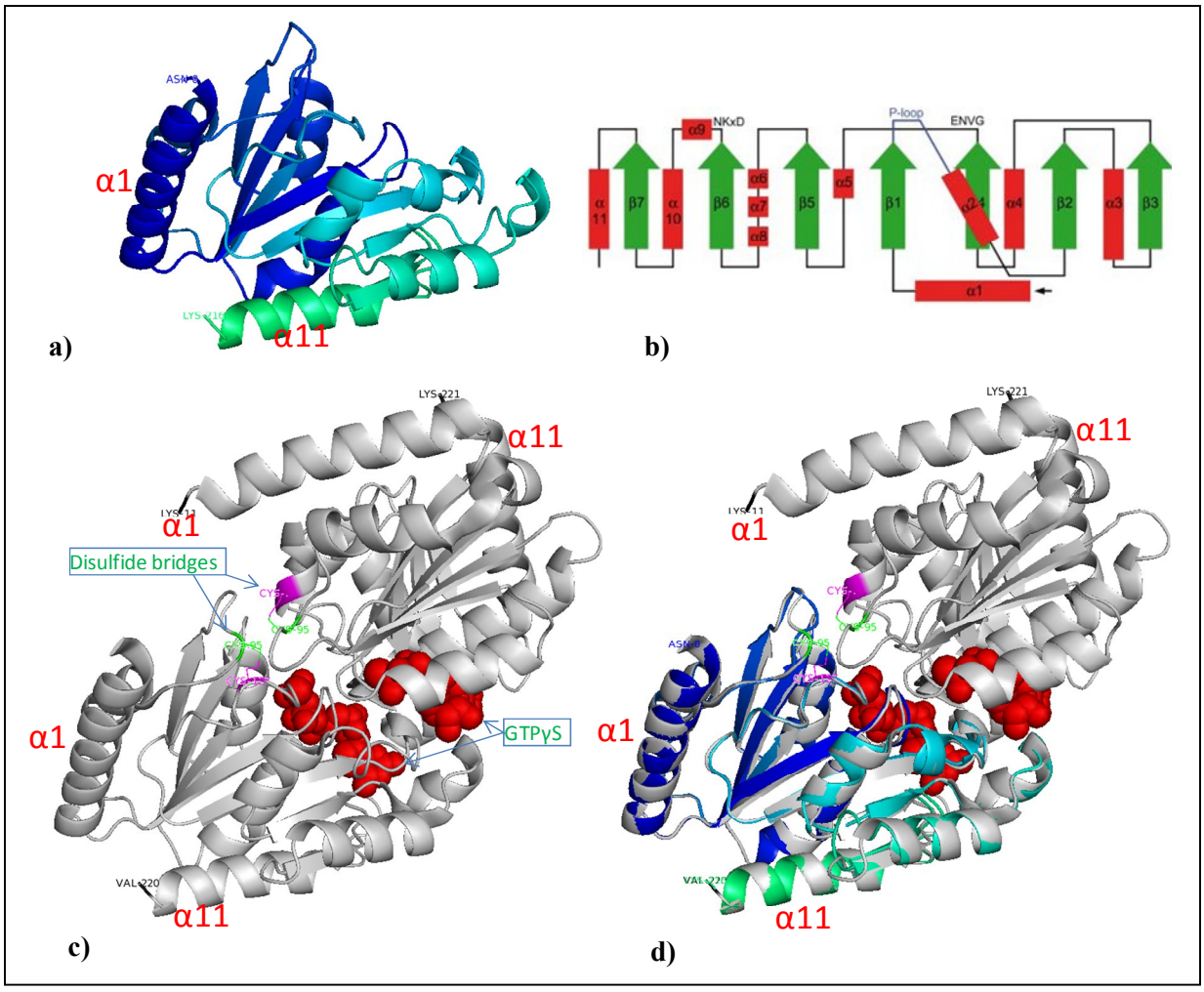


Fig 5.16 (a) The predicted three dimensional structure of HypB from DvMF by automatic homology modelling using HypB of *M. janaschii*, **(b)** topology diagram for *M. janaschii* **(c)** the dimeric three dimensional structure of the HypB of *M. janaschii* also showing the disulfide bonds and the cofactor GTP **(d)** the overlap of HypB of DvMF and *M. janaschii* shows an almost perfect fitting.

5.3.2 Ligand transfer maturation proteins

5.3.2.1 HypD of DvMF

Following, a secondary structure prediction of the HypD polypeptide sequence of DvMF in comparison to that of HypD from *Thermococcus kodakaraensis* was performed by means of a one vs one secondary structure alignment using the SSEA server. Fig 5.17 shows this comparison along with positions of the α-helices and β- sheets. Also a three dimensional model has been generated by homology modeling (Fig 5.18)

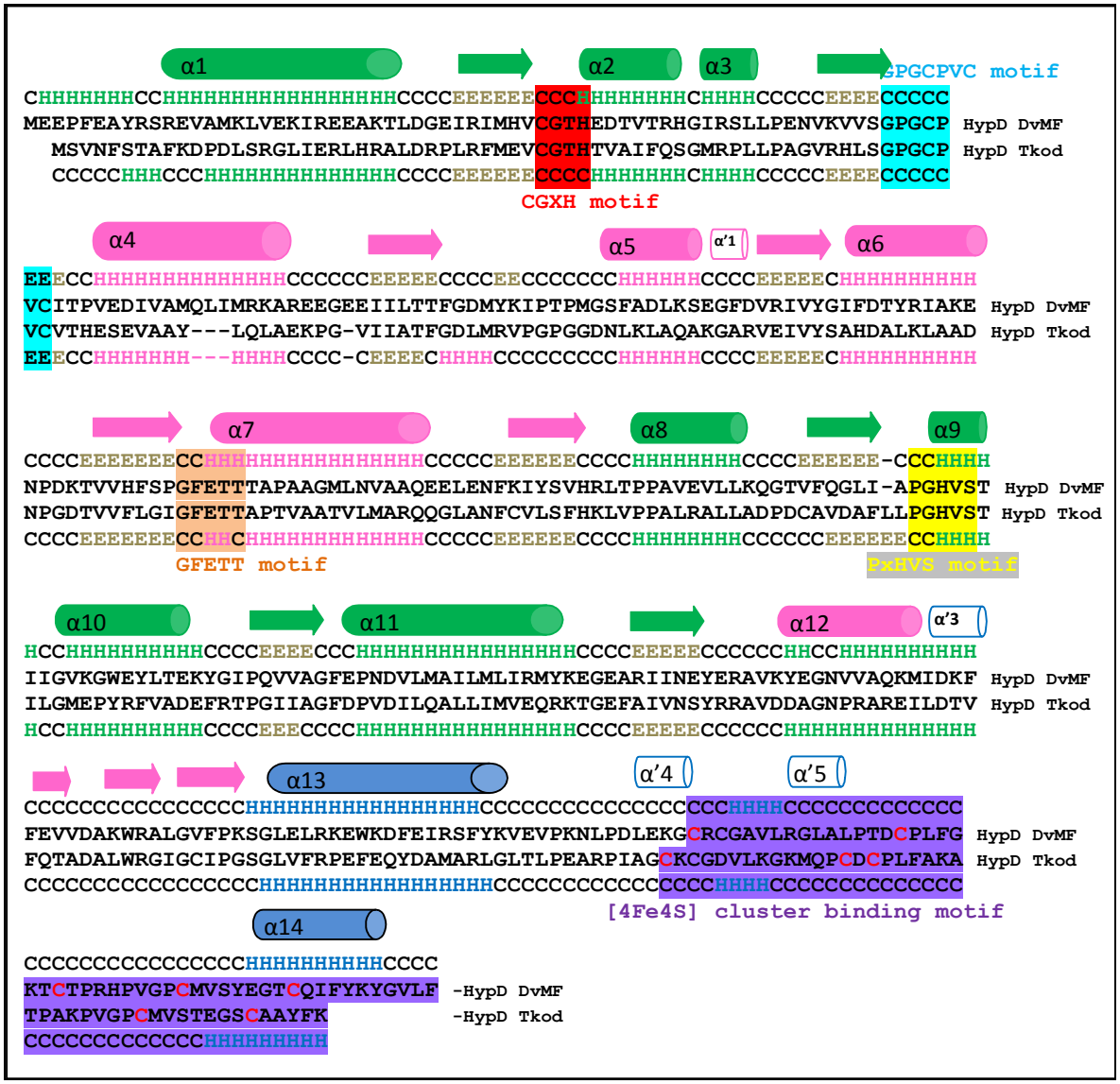


Fig 5.17 Secondary structure comparison of HypD of DvMF and HypD of *Thermococcus kodakaraensis*.

The HypD from both these organisms show high homologies at primary, secondary and tertiary level. The structurally and functionally important motifs CGTH (complex formation with HypC by oxidation of a cysteine), GPGCPVC (metal binding CxxC domain, characteristic of thiol redox proteins), GFETT (stabilization of the central cleft), and PGHVS (stabilization of the central cleft) are conserved. The Fe-S cluster forming C'- terminal sequences ($CX_{14}CX_6CX_{16}C$) are also conserved.

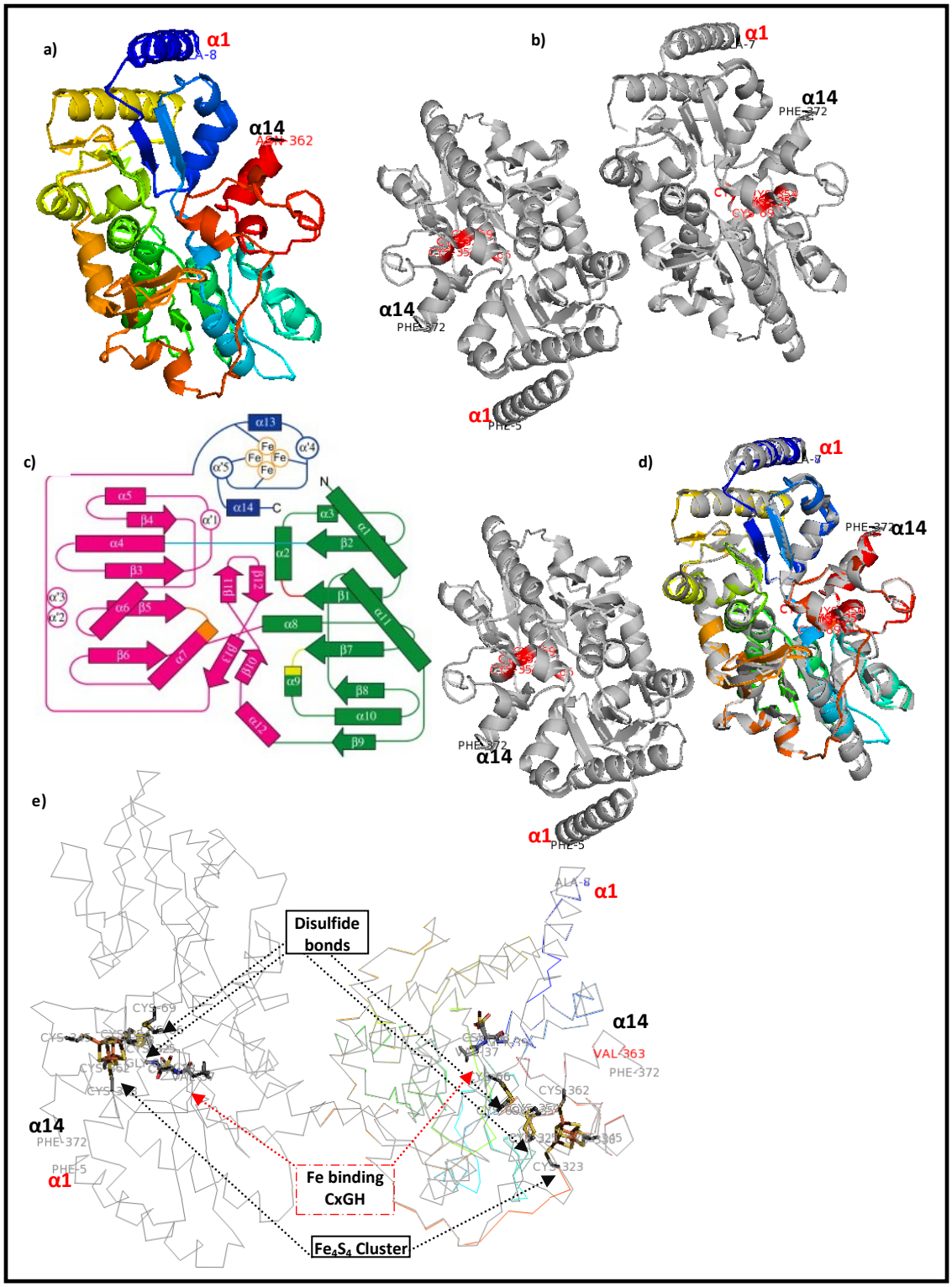


Fig 5.18 a) The three dimensional structure of HypD of DvMF as predicted by a automatic homology modelling using **b)** the dimeric three dimensional structure of the HypD of *T. kodakaraensis* **(c)** the topology diagram of HypD of *T. kodakaraensis* **(d)** HypD of *T. kodakaraensis* (grey) and the overlapped monomer (colored) of HypD of DvMF **(e)** another view showing the overlapping of two proteins to illustrate the ligands (catalytically important disulfide bonds, Fe binding motif and the [Fe_4S_4] cluster).

5.3.2.2 HypE of DvMF

The HypE maturation proteins from various bacteria have been assigned to the aminoimidazole ribonucleotide synthetase (PurM) family on the basis of the amino acid sequence analysis (Blokesch et al, 2004). The members of this family are in general involved in an ATP-dependent conversion of acyl compounds into various products. The three dimensional structures of HypE from *T. kodakaraensis* (Watanabe et al, 2007b), *E. coli* (Rangarajan et al, 2008) and also the sulfate-reducing bacterium DvH (Shmoura et al, 2007) are available (Table 5.2). In our work, the amino acids sequences of HypE of DvMF, DvH and *T. kodakaraensis* are compared, together with an approximate locations of the α helices and ß domains as found in the tertiary structure of DvH (Fig 5.19).

All the three crystal structures of HypE that are available are in general similar to each other. The overall structure is similar to other PurM proteins and the FGAM synthase domain of PurL. The four conserved motifs $_{68}$FPGG(D/N)IGxLA(V/I)xGxXNDLAxxGAxP$_{92}$ (motif I of *T. kodakaraensis*), $_{130}$(T/A)GDTKV$_{135}$ (motif II of *T. kodakaraensis*), $_{220}$(K/R)D(P/A)TR (A/G)G$_{226}$ (motif III of *T. kodakaraensis*) and $_{270}$ANEG(K/R)$_{274}$ (motif IV of *T. kodakaraensis*) form the central cleft and especially the conserved sequence Dx$_4$GAxP is a common signature of ATP binding domains (Anand et al, 1999). Also the C-terminal tail having the conserved motif PR(V/I)C has been found to change its conformation dependent on the presence of ATP (Watanabe et al., 2007b and Shmoura et al, 2007).

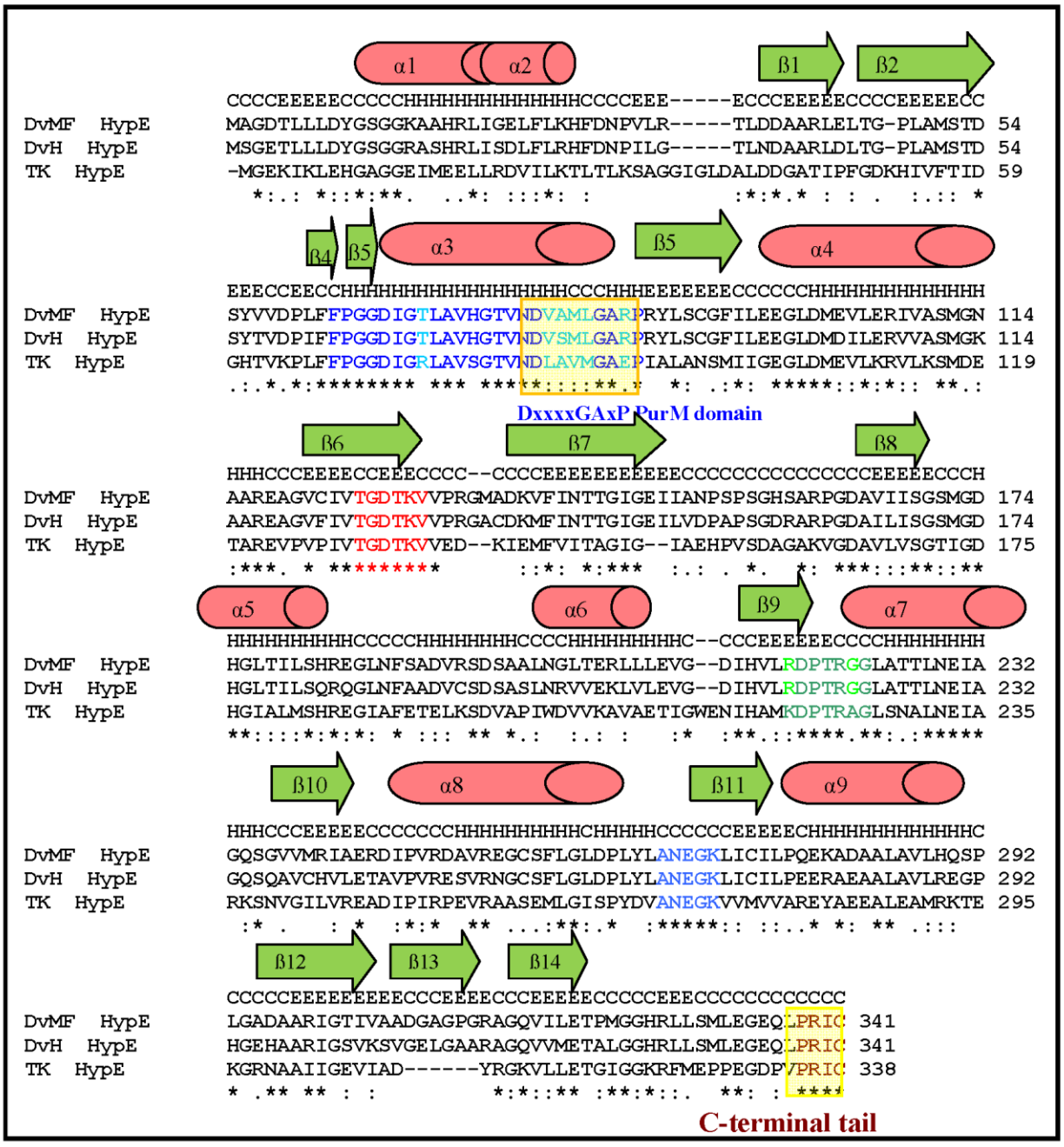


Fig 5.19 Primary and secondary structure comparison of HypE of DvMF, DvH and *Thermococcus kodakaraensis*.

Since the HypE of DvMF resembles most closely that of DvH (78% identities and 89% similarities), a three dimensional structure model for HypE was generated by homology modeling using the tertiary structure from DvH. The three dimensional structure of DvH has been determined in an apo and an ATP-bound form (Shomura et al, 2007). The three dimensional overlap of the HypE protein structures of DvH and DvMF shows clearly the

conservation of the ATP binding sites and the proximity to the C'-terminal residue (Fig 5.20) which accepts the carbamoyl group of carbamoyl-AMP from HypF. Subsequently, HypE catalyzes the dehydration of the carboxamide in an ATP-dependent manner.

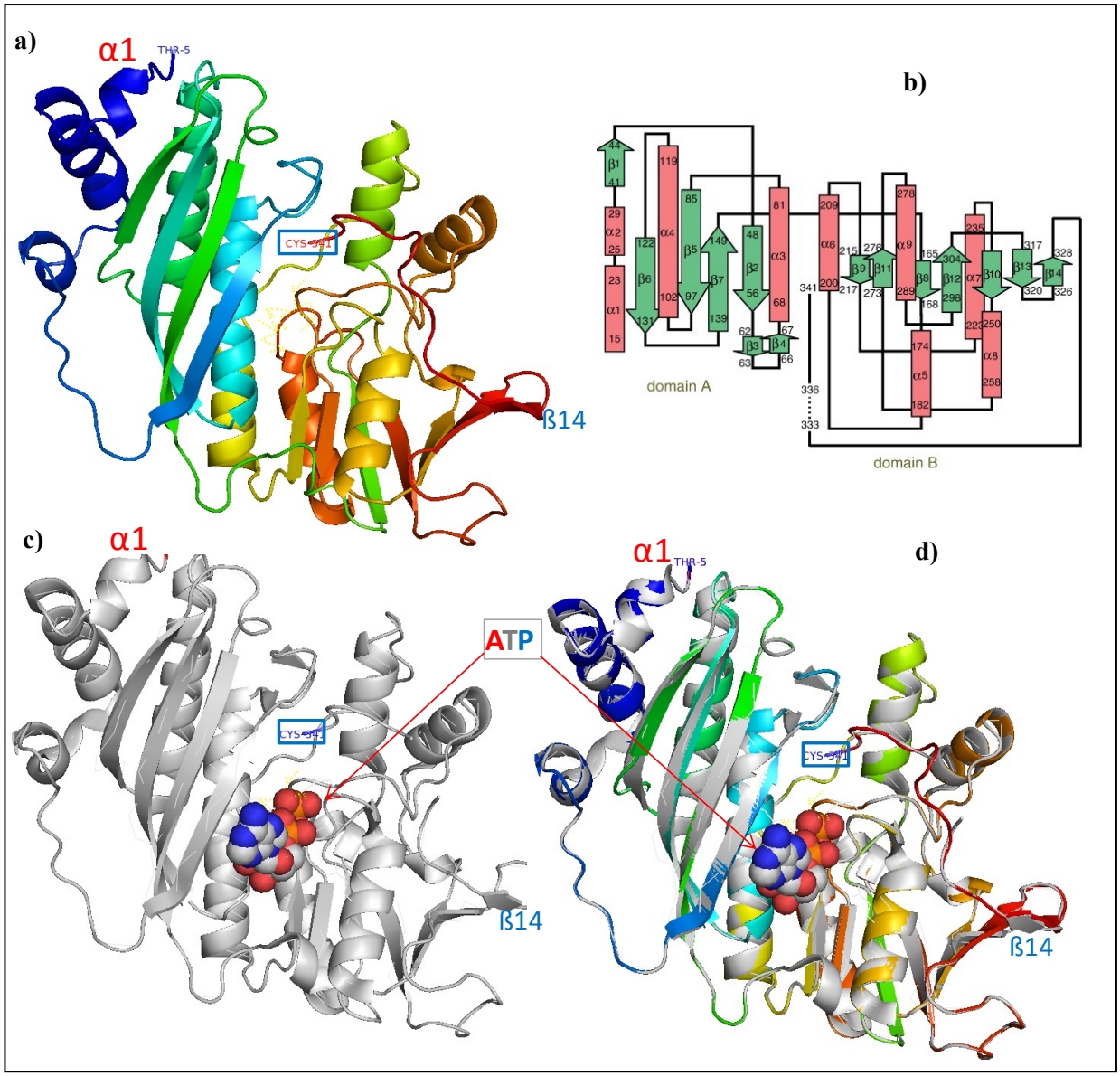


Fig 5.20 (a) The three dimensional structure of HypE of DvMF as predicted by automatic homology modelling using HypE of DvH, **(b)** topology diagram, and **(c)** the three dimensional structure of HypE from DvH in the ATP-bound form. The overlap **(d)** of HypE of DvMF and DvH shows a perfect fitting; also the positioning of the terminal cysteine in regard to the ATP binding domain is identical.

5.3.2.3 HynD of DvMF

HynD which is mostly known as HypC in other organisms, is a small 8 kDa chaperone protein that forms a complex with HypD during the maturation of the [NiFe] hydrogenase. The crystallographic asymmetric unit of *T. kodakaraensis* contains three HypC molecules. The α-helix present at the C- terminal and the residues present at the N-terminus show flexibility and can adopt different conformations that gives the molecule a hydrophobic or acidic molecular surface (Watanabe et al, 2007). It was proposed that the residues Cys-2 and His-45 may play a greater role in the conformational changes and the interactions of HypC with HypD. The homologous model for HynD of DvMF (Fig 5.21) shows a perfect conservation of the overall structure and the catalytically important residues (Cys-2 and His-49).

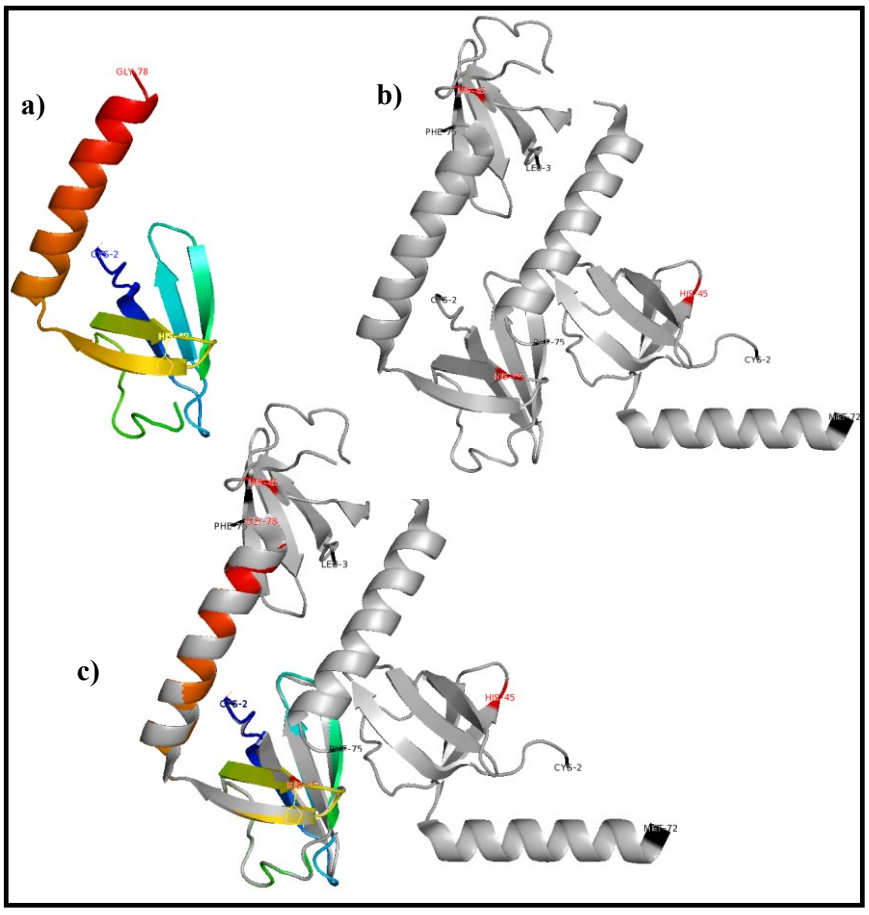


Fig 5.21 (a) The three dimensional structure of HynD from DvMF as predicted by automatic homology modeling using HypC of *T. kodakaraensis*, **(b)** and the three dimensional trimeric structure of the HypC of *T. kodakaraensis* **(c)** the overlapping of HynD of DvMF and HypC of *T. kodakaraensis* shows ideal fitting.

5.3.2.4 HypF of DvMF

HypF is involved in the synthesis of the CN ligands of [NiFe] hydrogenase. The HypF protein accepts carbamoyl phosphate (CP) as a substrate and catalyzes both a CP phosphatase reaction in the absence of any other substrate and a CP-dependent hydrolysis of ATP into AMP and inorganic pyrophosphate (PPi). Because the latter reaction is reversible, as shown by a CP-dependent pyrophosphate-ATP exchange, it was postulated that an adenylated CP derivative is an intermediate in the reaction and that the CP phosphatase activity may reflect a side reaction followed in the absence of ATP (Paschos at el, 2005). The 82.1 kDa HypF protein from *E. coli* shows significant similarity in the N-terminus to the acylphosphatases of eukaryotes, followed by two classical identical zinc finger motifs, and the C- terminal part shows significant similarity to the O-carbamoytransferases (Paschos et al, 2005).

A comparison of the primary sequences of HypF proteins from DvMF, DvH, DDG20, *Ralstonia eutropha* and *E. coli* showed a perfect conservation of the acylphosphate phosphatase ($IXGXV_{17}QGV_{20}XFR$ of *E. coli*), O-carbamoyltransferase ($VXH_{475}H_{476}XAH_{479}$ of *E. coli*) and two zinc finger domains ($C_{109}X_2C_{112}X_{18}CX_2C$ and $C_{159}X_2C_{162}X_{18}CX_2C$ of *E. coli*) (Fig 5. 22). Whereas the conservation of the functional domain of HypF even across phylogenetically widely spaced organisms is remarkable, the HypF of DvMF has been found to possess at least two long stretches (G_{214}- T_{243} and G_{336}-L_{267}) of amino acids which are not present in DvH and neither in DDG20.

5.3.2.5 HynC of DvMF

HynC of DvMF has been compared with the three dimensional structure of HybF of *E. coli* (Goenka et al, 2006) and has been found to conserve most of the features needed for the endopeptidase activity (Fritsche et al, 1999), followed by nickel insertion by HypA and HypB.

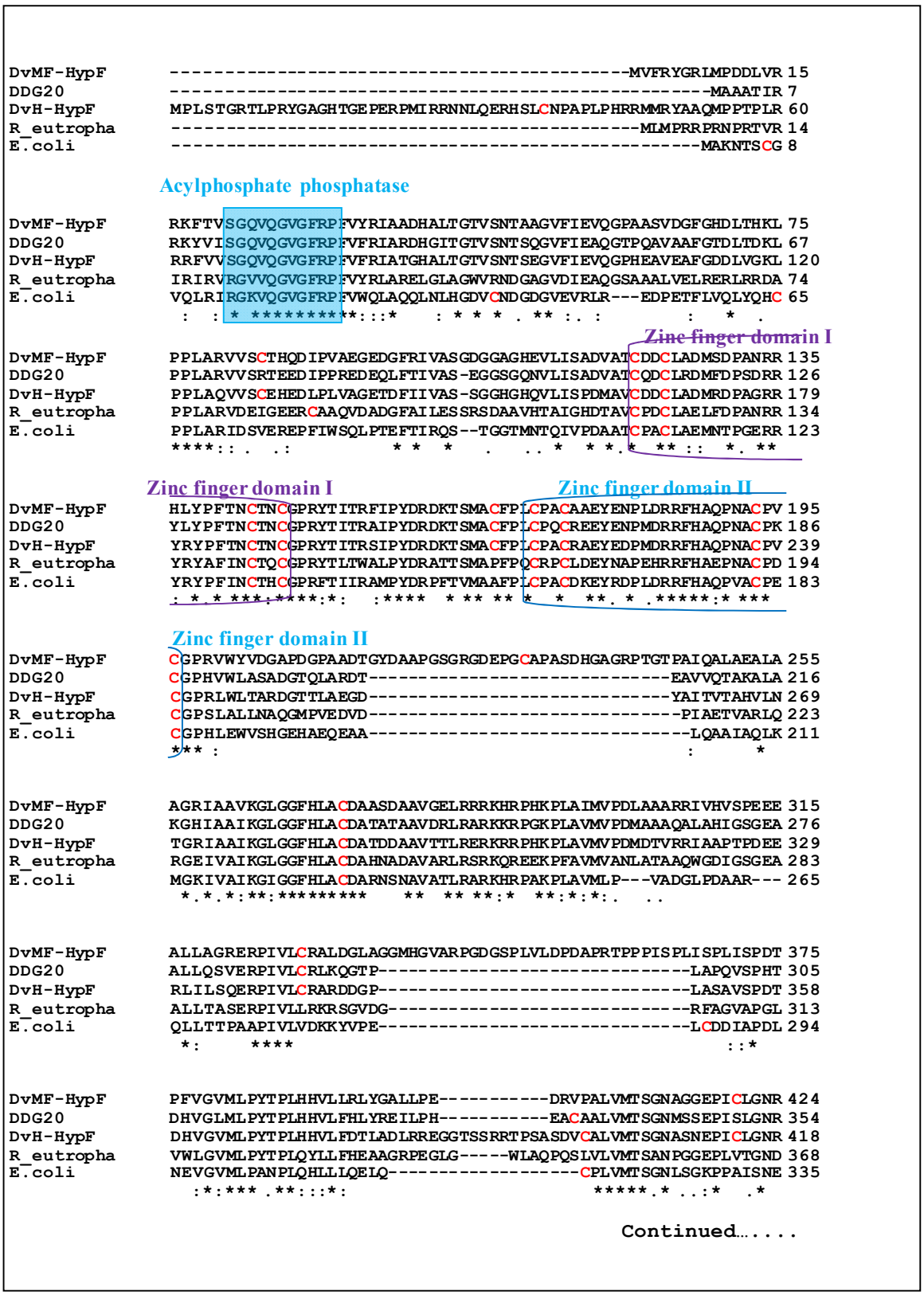

Acylphosphate phosphatase
Zinc finger domain I
Zinc finger domain I
Zinc finger domain II
Zinc finger domain II
Continued……

Fig 5.22 Schematic representation of the primary sequence conservation of HypF of DvMF, DDG20, DvH, *Ralstonia eutropha* and *E. coli*, indicating the position of the acylphosphatase, the two zinc fingers and the carbamoyltransferase motifs.

5.4.2 Expression of hynA-strepBCD operon in *E. coli*

The strep-tagged *hynB-strepACD* operon excluding the promoter region was amplified by PCR using the primer pair hynBpRSF2B 5'-gacgacgacaagatgaaaatctcgatcggtctcggcaa-3' and hynDpRSF2B 5'-gaggagaagccccgtatgaacccggcgggctggcctcc-3' and cloned into the vector pRSF-2 Ek/LIC to give a construct pRSF2hynB-strepBCD (Km^R). pRSF-2 Ek/LIC is a T7 promoter derived expression vector from Novagen and is constructed in a way that the first ORF is in frame with an N-terminal histidine tag and the last ORF can be put in a frame with a C- terminal S-tag. When induced, a functional construct should express a double tagged HynB (N- terminal histidine and C-terminal strep), HynA, HynC and a C- terminal, S-tag harboring HynD (Fig 5.24).

Fig 5.24 Schematic presentation showing the coding arrangement of the pRSF2hynABCD operon and the location of the epitope-coding sequences.

BL-21 E. *coli* cells, transformed with pRSF2hynB-strepACD, were grown at 37°C and constant shaking at 180 rpm in order to express these proteins. The cells were grown to an O.D. of 0.6 before inducing by 0.7 mM IPTG. After induction, the cells continued to grow overnight at room temperature. A 50 µl aliquot of these cells was centrifuged and boiled in 50 µl of SDS-PAGE buffer. The resultant samples were tested by western blotting using four antibodies.

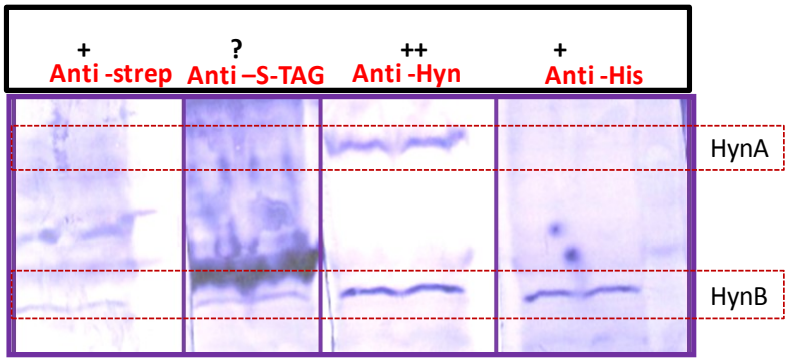
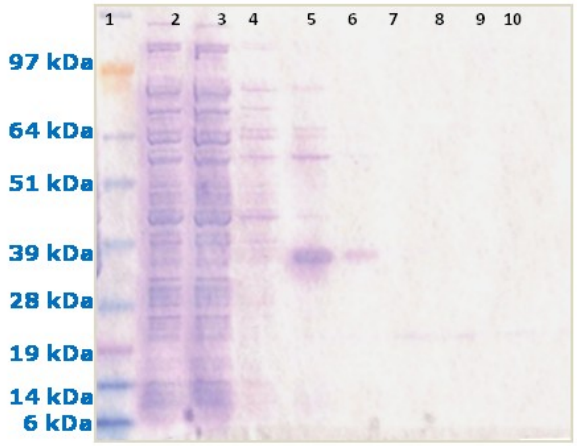


Fig 5.25 Western blot detections of hydrogenase expression from the *hynBACD* operon cloned in the pRSF2hynB-strepACD vector using anti-strep, anti-S-Tag, anti-Hyn and anti-His antibodies. Two red boxes mark the locations of the small subunit HynA and the large subunit of HynB.

The anti-Strep (to detect HynB), anti-His (to detect HynB) and anti-Hyn (to detect HynB and HynA) antibodies showed the expression of the double tagged small subunit, while the presence of the large subunit was also confirmed by the anti-Hyn antibody. The S-tag antibody could not identify any specific bands at the expected 11 kDa location. However, it reacted with the small subunit along with some other nonspecific identification. The sequence analysis has shown the presence of the promoter region before the *hynCD* genes of the *hynBACD* operon, and in the construct pRSF2hynBACD a strong T7 promoter might have given additional support. Here we think that the non-detection of HynD might be due to its small size or due to its capability to form strong complexes with other proteins.

The final volumes of prepared cells were harvested by centrifugation and sonicated in 10 mM Tris, 150 mM NaCl buffer. The cell lysate was centrifuged at 13000 rpm for 1 hour to separate the soluble fraction from the cell debris and other insoluble fractions. The soluble fraction was purified by a strep-tag purification sepharose column (IBA-technologies) following the manufacturer's instructions. The various fractions were analyzed by gel electrophoresis by coomassie staining and also by western blotting using anti-Hyn antibody (Fig 5.27).



1. See plus Marker
2. Cell lysate soluble fraction
3. Flow through
4. Column wash
5. Elution 1
6. Elution 2
7. Elution 3
8. Elution 4
9. Elution 5
10. Elution 6

Fig 5.26 Protein gel electrophoresis (SDS-PAGE) of protein extract prepared from T7-induced *hynA-strepBCD* operon over-expressed from the pRSF2hynA-strepBCD operon and its purification through the strep tag purification column.

The strep-tag purification profile showed the presence of two major protein bands (Fig 5.26, lane 5 and Fig 5.27 lanes 4-9), one corresponding to the molecular weight of the unprocessed small

subunit of the DvMF [NiFe] hydrogenase (Fig 5.27, enclosed in the blue box) and a second protein band corresponding to the large subunit (Fig 5.27, enclosed in the yellow box).

1. Total cell lysate/cell boiling
2. See plus Marker
3. DvMF total cell lysate/cell boiling
4. Cell lysate soluble fraction
5. Flow through
6. Column wash
7. Column wash
8. Elution 1
9. Elution 2

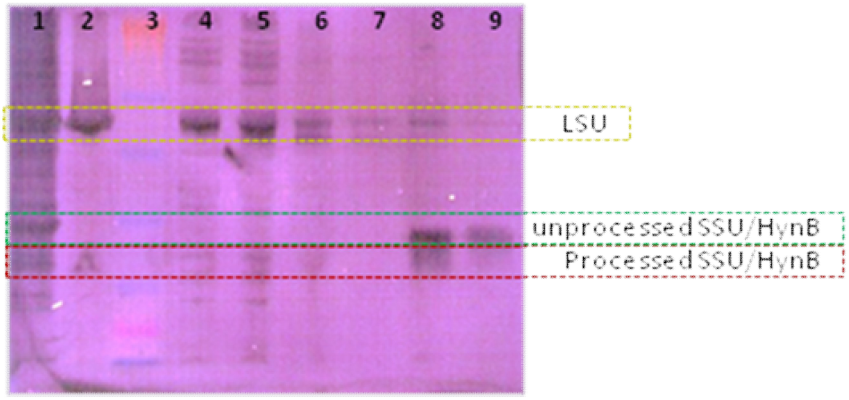


Fig 5.27 Western blot detection of [NiFe] hydrogenase expression and purification by anti-Hyn antibody. The protein extract was prepared from T7-induced *hynA-strepBCD* operon and purified using the strep tag purification column.

Given that the strep tag was present only in the small subunit, the large subunit was co-purified along with the small subunit. Similarly, the western blotting did show the presence of two subunits in the final purification products. It is noteworthy that in the unpurified total cell lysates, there was a protein band reacting with the anti-Hyn antibody, being present at about the same location of a processed small subunit of NiFe hydrogenase of DvMF. Also, the soluble fraction and washes did not show the presence of the small subunit as the major over-expressed protein (see purification gel, Fig 5.26, lane 2, 3 and 4; western blot Fig 5.27 lane 4, 5, 6 and 7), and it became visible only in elutions (Fig 5.26, lane 5 and 6; Fig 5.27, lane 8 and 9). Moreover, the intensity of the large subunit, which is second in position to the T7 promoter, showed a much higher intensity in the total soluble fraction and in the washes as compared to that in the final elution. Putting this information in one perspective, the proportions of the unprocessed small subunit was highest in the final elution while the amounts for the large subunit and the apparently 'processed small subunit' decreased as the purification proceeds.

A plausible explanation for this could be that a part of the expressed small subunit polypeptide was transported to the periplasm via the *E. coli* tat transport pathway machinery and the rest of the unprocessed small subunit remained in the cytoplasm. The expressed small subunit and the

large subunit polypeptides did show some interaction, which was of weaker nature and was lost during the washing steps. Nevertheless, a certain amount of the large subunit polypeptide was co-purified with the small subunit.

5.4.3 Cloning and expression of two-gene Hyp and Hys operons

The operons *hypAB*, *hypDE* and *hysBA* were amplified by appropriate primers for cloning into the pRSF2 Ek/LIC vector as described before (see table 5.3 for list of primers). The final constructs pRSF2hypAB, pRSF2hypDE and pRSF2hysBA were transformed separately into BL-21 cells. Two more constructs, pRSF2-hysA and pRSF2-hysA-preseleno were prepared; both of these carried the large subunits of the [NiFeSe] hydrogenase, i.e. HysA in frame with an N-terminal-His- and a C-terminal S-tag of pRSF-2 ek/LIC. The construct pRSF2-hysA carried the entire HysB coding gene sequence including the selenocysteine, and the second construct pRSF2-hysA-preseleno carried the HysA coding gene sequence before the selenocysteine.

Name of the construct, operon and the primers used	**Primer sequences 5'-3'**
pRSF2-hypAB : [*operon hypAB*] **Forward-**RSFhyPA-Fr **Reverse-** hypB-STAG	**5'-GACGACGACAAGATGCATGAAATGTCCGTCGCATC-3'** **5'-GAGGAGAAGCCCGGTAACCCGCGCTTGGCGGCCCGCGCGGTGC-3'**
pRSF2-hypDE: [*operon hypDE*] **Forward-** RSFhypD-Fr **Reverse-** hynE-STAG	**5'-GACGACGACAAGATGGTGAGCGTGAATTTCTCCACGGCCTT-3'** **5'-GAGGAGAAGCCCGGTATGAACCCGGCGGGCTGGCCTT-3'**
pRSF2-hysAB : [*operon hysAB*] **Forward** RSFhysB-Fr **Reverse-** hysA-STAG	**5'-GACGACGACAAGATGAGTCTCAACAGGCGTGATTTCG-3'** **5'-GAGGAGAAGCCCGGTCTTCACTTCGACGACGGAGA-3'**
pRSF2-hysB: [*hysB*] **Forward-** RSFhysA-Fr **Reverse** hysA-STAG	**5'-GACGACGACAGAATGTCCGGCTGTACACCCAAGGC-3'** **5'-GAGGAGAAGCCCGGTCTTCACTTCGACGACGGAGA-3'**
pRSF2-hysA-preseleno : [*hysB*] **Forward-** RSFhySA-Fr **Reverse-** hysASTAG-preCYS	**5'-GACGACGACAGAATGTCCGGCTGTACACCCAAGGC-3'** **5'-GAGAGAAAGCCCGGTTTCGGGTCGGGGTCGAAGGCGCGGATC-3'**

Table 5.3 Nucleotide sequences of the primers used to amplify operons *hypAB*, *hypDE*, *hysAB* and the large subunit of [NiFeSe] hydrogenase of DvMF (*hysB*).

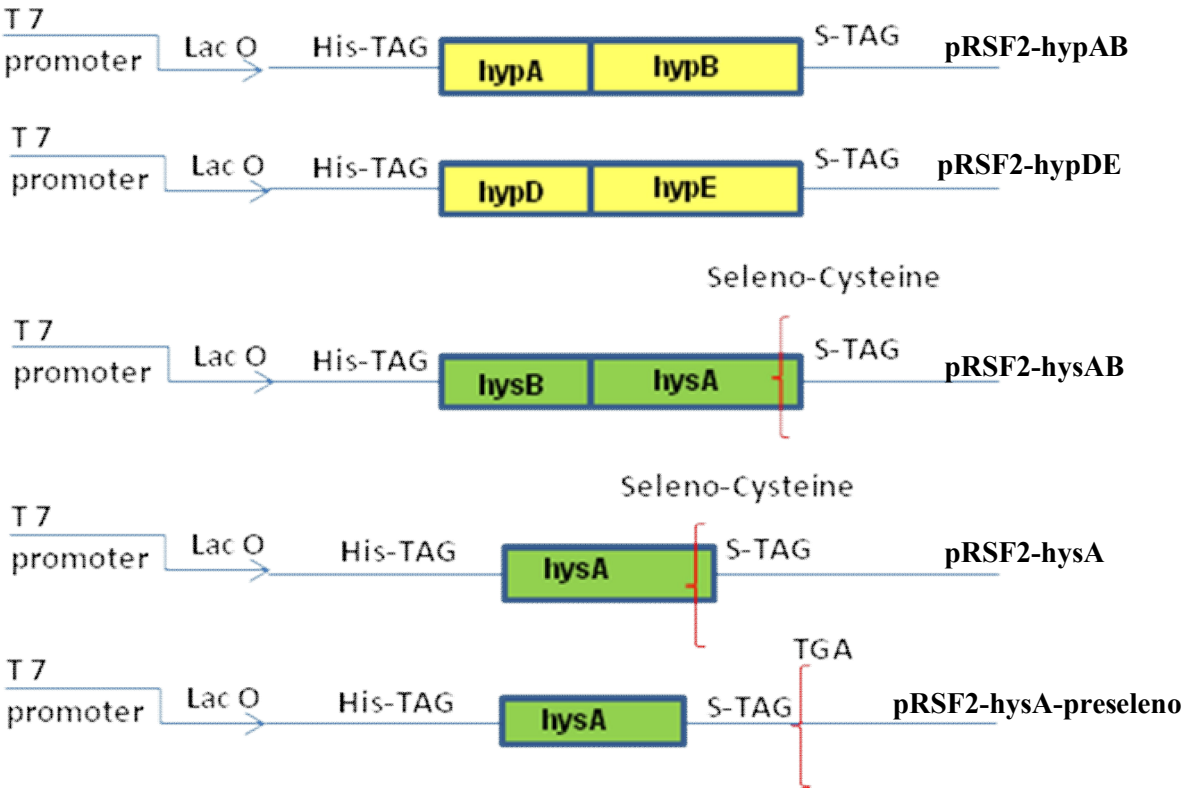


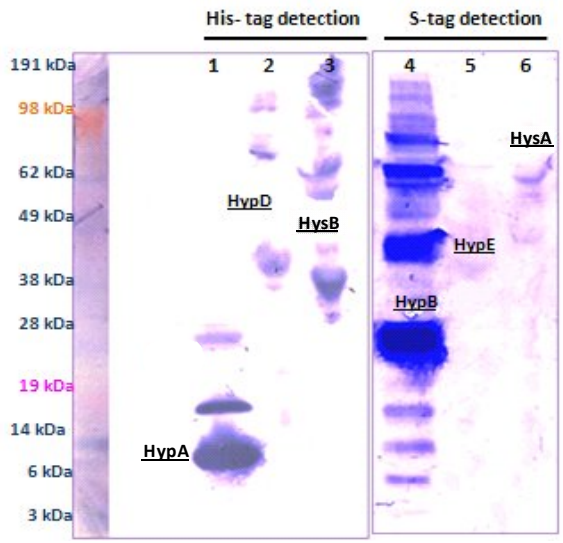
Fig 5.28 Construction of the plasmids carrying the maturation gene operons *hypAB* (pRSF2-hypAB), *hypDE* (pRSF2-hypDE), and the [NiFeSe] hydrogenase structural gene operon *hysAB* (pRSF2-hysAB). The large subunit of the [NiFeSe] hydrogenase *hysB* (pRSF2-hysB) contains the complete coding region and the pRSF2-hysB-preseleno construct contains the coding region before selenocysteine). The red bracket represents the selenocysteine codon.

Western blotting detection

The *E. coli* cells harboring these recombinant expression vectors were grown at 37°C and shaking at 180 rpm until an A_{600} = 0.6, when they were induced by 0.6 mM IPTG and left to grow overnight at room temperature. Next day, 50 µl aliquots of each of the growth cultures were tested by western blotting by anti-His and anti-S-tag antibodies as before. The anti-His antibody detected the expression of HypA, HypD and HysB at the expected molecular weight position (Fig 5.29). The S-tag antibody detected HypB very well, however the HypE and HysA showed poor reactivity (Fig 5.29). Interestingly, the HysA showed immunodetection by the S-tag antibody only in cases when its seleno-cysteine codon was read as selenocysteine codon (in *E. coli*) and not as termination codon.

Plasmid construct	Expected proteins	Expected molecular weight	Detection by westren blotting
pRSF2-hypAB	His-HypA HypB-S-tag	1.6 kDa + 12.8 kDa 24.4 kDa + 3.1 kDa	+ +
pRSF2-hypDE	His-HypD HypE-S-tag	1.6 kDa + 39.6 kDa 36.1 kDa + 3.1 kDa	+ +
pRSF2-hysBA	His-HysB HysA-S-tag	1.6 kDa + 34 kDa 56.3 kDa + 3.1 kDa	+ +
pRSF2-hysA	His- HysA-S-tag	1.6 kDa + 56.3 kDa + 3.1 kDa	+
pRSF-hysA-preseleno	His- HysA-S-tag	1.6 kDa + 56.3 kDa + 3.1 kDa	+

Table 5.4 A tabular presentation of the expected molecular weights of the expressed proteins from the constructs prepared by cloning the operons *hypAB* (pRSF2-hypAB), *hypDE* (pRSF2-hypDE), *hysBA* (pRSF2-hysBA) and the gene hysA (pRSF2-hysAand pRSF2-hysA-preseleno) in pRSF-2 Ek/LIC vector. The molecular weights are calculated by translating the proteins, considering the histidine tag contributes 1.6 kDa, while the S-tag should contribute 3.1 kDa. The + (plus) sign indicates positive detection of various proteins by western blotting in our work.



See Blue Marker

Anti-His detection

1) pRSF2hypAB; induced
2) pRSF2hypDE; induced
3) pRSF2hysBA; induced

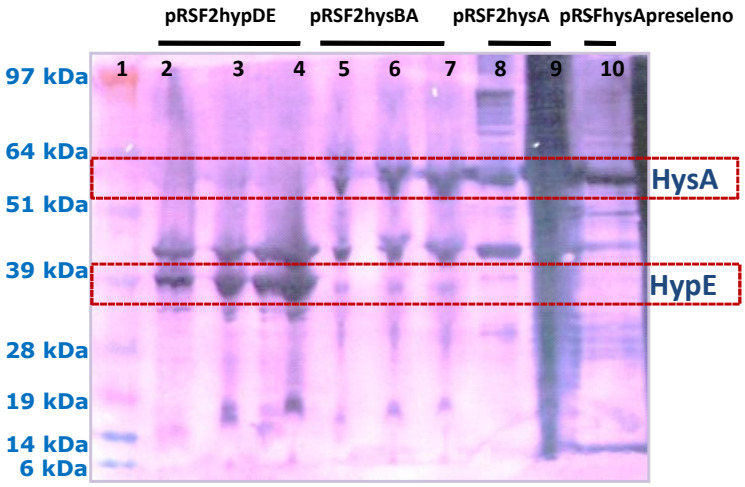
Anti-S-tag detection

4) pRSF2hypAB; induced
5) pRSF2hypDE; induced
6) pRSF2hysBA; induced

Fig 5.29 Western blotting detections of the HypA, HypD and HysB proteins by an anti-His antibody and detection of HypB, HypE and HysA using an anti-S-tag antibody.

In order to improve the quality of the detections with pRSF-hypDE and pRSF-hysAB, the IPTG induction steps were slightly modified and the expression of HysA in the two additional constructs pRSF2-hysA and pRSF-hysA-preseleno was also checked. All of these constructs were again transformed in freshly prepared BL-21 cells and grown similarly as before, however, this time they were induced at A_{600} = 0.8 and by 1.2 mM IPTG. After overnight growth at room temperature, a small amount of each, prepared by boiling in SDS-PAGE buffer, was analyzed by western blotting using an anti S-TAG antibody (Fig 5.29).

The anti-S-tag antibody detected a protein band slightly above 39 kD band in each sample. Barring this unspecific detection, the HypE was well detected in the pRSF2hynDE cell lysates both in induced and un-induced samples. Also the HysA was detected in all the three constructs, pRSF-hysBA, pRSF-hysA and pRSFhysA-preseleno. Though, a careful comparison of the western blot profile showed a small difference between the locations of these bands. This could be very well attributed to the presence of the His-tag in pRSF-hysA and pRSF-hysA-preseleno and also to the absence of the post seleno-cysteine polypeptide sequence in pRSF-hysA-preseleno.



1. See Blue Marker
2. pRSF2hypDE uninduced
3. pRSF2hypDE induced
4. pRSF2hypDE induced
5. pRSF2hysBA uninduced
6. pRSF2hysBA induced
7. pRSF2hysBA induced
8. pRSF2hysA induced- partially purified via his tag
9. pRSF2hysA induced
10. pRSF2hysA- preseleno- induced partially purified via his tag

Fig 5.30 Western blot using an anti-S-tag antibody detects HypE from constructs pRSF2hypDE (lane 2, 3 and 4) and HysA from pRSF2hysBA (lane 5, 6 and 7), pRSF2hysA (lane 8 and 9) and pRSF2hysA-preseleno (lane 10).

As it appears from these experiments, the ribosome binding sites of the transcripts originating from DvMF gene sequences have been recognized in *E. coli*. Thus, the two- or four- gene operons that exist in the known DvMF hydrogenase genetic machinery are expressed as multiple proteins simultaneously from a single promoter in *E. coli*. Accordingly, an attempt to clone and co-express these genes needs not to consider each gene individually. A second and very interesting aspect is the integration of the selenocysteine in the large subunit HysA of [NiFeSe] hydrogenase polypeptide in the *E. coli* cells.

5.4.4 Co-cloning of strep tagged *hynBACD* operon with npt2 gene

An antibiotic marker npt2 (neomycin phosphor transferase) has been recently reported to enable the selection of transformed DvH using a kanamycin analogoue G418 (Caffrey et al, 2007). For applying this information while manipulating DvMF, the npt2 coding gene sequences have been co-cloned in to the strep tagged strep tagged *hynBACD* operon present in pMOShynBACDstrep construct. The npt2 gene was isolated from pBSL180 (Fig 9.1 in appendix) by digesting with doubly present *SacI* site and cloned into SmaI site of pUC19 to generate pUC19nptII. It was further sub-cloned into *xbaI-speI* digested pMOShynBACDstrep construct after cutting with doubly present *xbaI*. The resultant pMOShynBACDstrepNptII was both ampicillin and kanamycin resistance. Experiments are on way to transfer this combination to a pRSF1010 based replicon to be tried in DvMF.

5.5 Purification and Structural Characterisation of Sulfate Metabolism Proteins from *Desulfovibrio vulgaris* Miyazaki F

During the routine preparations of [NiFe] hydrogenase, two soluble proteins which were present in abundance were purified further. They were named Green and Brown proteins respectively based on the visual colors (work by Dr. Aruna Goenka Agrawal). Afterwards these proteins were identified by N-terminal sequencing and high energy MALDI and the data obtained was used to sequence the corresponding genes in DvMF genome (Chapter 4)

5.5.1 Purification and Characterization of Brown protein

The brown protein was purified from the soluble fraction of the DvMF cell lysate using gel exclusion and anion exchange columns (Chapter 3). The purified protein was analysed on SDS gels and showed two subunits (Fig 5.31), which by MALDI analysis gave two distinct peaks one at 74 kDa and other at 20 kDa (Fig 5.32).

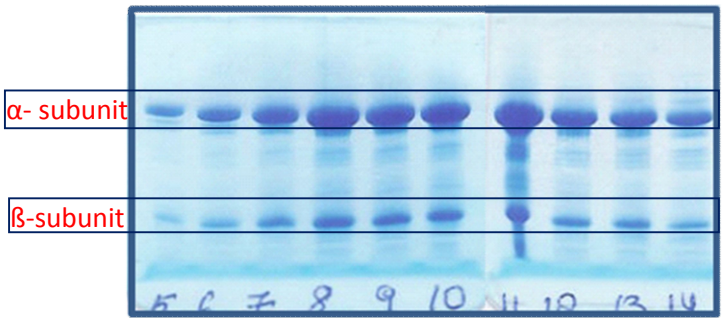


Fig 5.31 SDS-PAGE resolution of various fractions of purified Brown protein showed two subunits which were identified as α- and ß- subunits of 5'-Adenine phospho reductase.

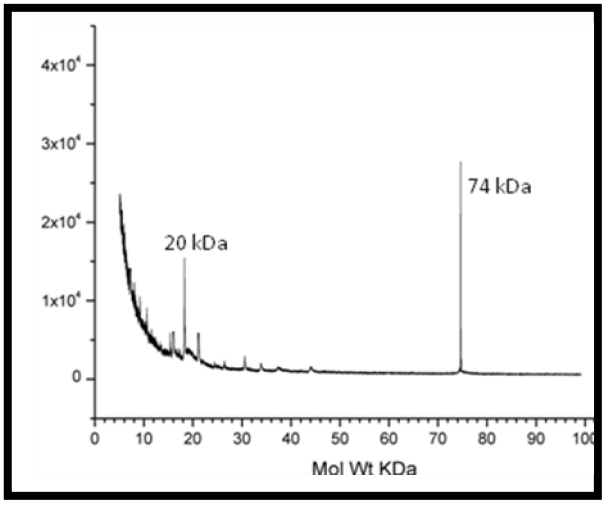


Fig 5.32 MALDI-TOF of purified Brown protein showed two distinct peaks at 74 kDa and 20 kDa corresponding to the α- and ß-subunits of 5'-adenine phospho reductase.

The purified protein subunits were gel-isolated and sent for N-terminal sequencing (U. Giessen). They were also analysed for their peptide profile through high energy MALDI at a complete subunit (Fig 5.33). The N-terminal sequence of the 74 kDa subunit led to the identification of MIPVKEQPK sequence (identical for the α- subunit of APS reductase of DvH) and of the 20 kDa residues PTYVDPSKCDGC (identical for the ß-subunit of APS reductase of DvH and DDG20). Further, the MALDI gave 27% sequence coverage of the 74 kDa subunit and data fitted the α-subunit for a possible APS reductase sequence (Fig 5.34). Using the N-terminal

sequences and other informations, the corresponding two genes present in the DvMF genome were sequenced utilizing the cosmid library (Chapter 4).

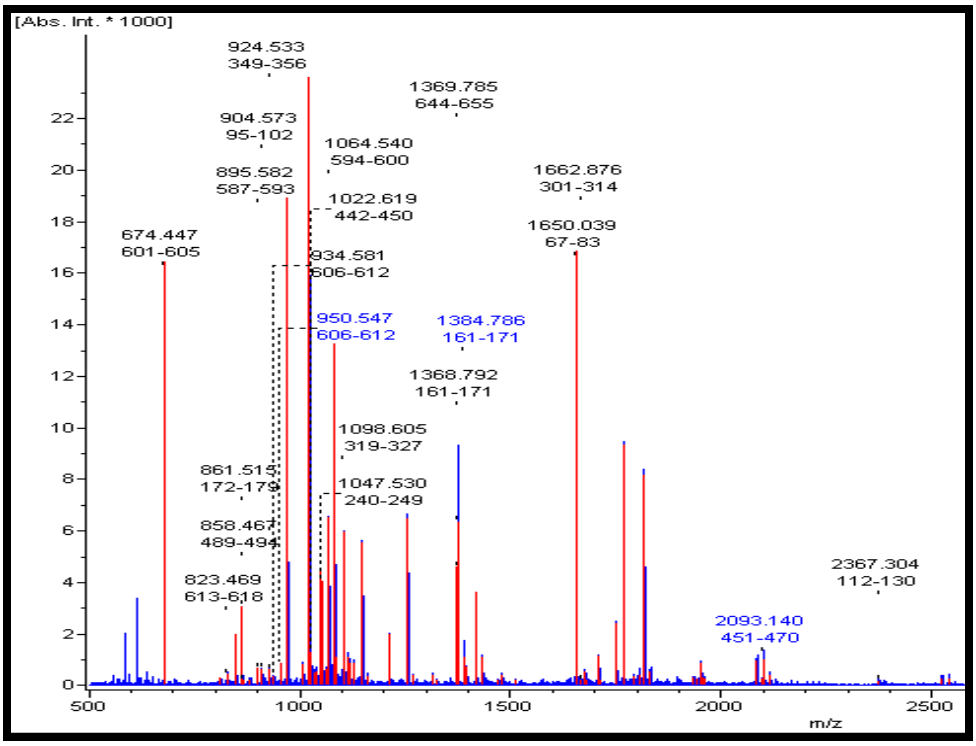


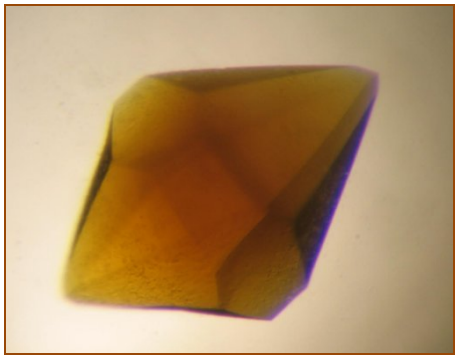
Fig 5.33 Peptide profile of the 74 kDa subunit of APS reductase generated by high energy fragmentation of the polypeptide by using MALDI-MS, the molecular mass were calculated by MALDI and the residue numbers were given by protein aligment with available sequences

```
        10         20         30         40         50         60         70         80         90        100
MPMIPVKEQP KGVAIAEPTV KEHDVDLLIV GGGMGACGTA FEAVRWADKY APELKILLID KASLERSGAV AQGLSAINTY LGKNDADDYV RMVRTDLMGL
       110        120        130        140        150        160        170        180        190        200
VREDLIFDLG RHVDDSVHLF EEWGLPCWIK DEHGHNLDGA QAKAAGKSLR NGDDPVRSGR WQIMINGESY KCIVAEAAKN ALGEARIMER IFIVKLLLDA
       210        220        230        240        250        260        270        280        290        300
NTPNRVAGAV GFNLRANEVH IFRSNAMLVA CGGAVNVYKP RSTGEGMGRA WYPVWNAGST YTMCAQVGAE MTMMENRFVP ARFKDGYGPV GAWFLLFKAK
       310        320        330        340        350        360        370        380        390        400
ATNYKGEDYC ATNRAMLKPY EDRGYAKGHV IPTCLRNHMM LREMREGRGP IYMDTKTALQ STFANMTPEQ QKHLESEAWE DFLDMCVGQA NLWASMNIQP
       410        420        430        440        450        460        470        480        490        500
EERGSEIMPT EPYLLGSHSG CCGIWVSGPD EKWVPEDYKV RASNGKIYNR MTTVEGLWTC ADGVGASGHK FSSGSHAEGR ICGKQMVRWC LDHKDYKPAI
       510        520        530        540        550        560        570        580        590        600
KESADELVKL IYRPYYNYME GKAASTDPVV NPSYITPKNF MMRLVKCTDE YGGGVGTYYT TSAAALDTGF SLLGMLEEDS LKLAARDLHE LLRCWENYHR
       610        620        630        640        650        660        670
LWTVRLHMQH IRFREESRYP GFYYRADFMG LDDSKWKCFV NSKYDPATGE TKIFKKAYYQ IIPE
```

Fig 5.34 Fitting the data generated by peptide profiling of the 75 kDa subunit of APS reductase by MALDI-MS into the available sequence database of NCBI by means of BLAST program.

5.5.2 Crystallization of 5'-adenine phospho reductase of DvMF

After optimization, crystals suitable for diffraction experiments (Ogata et al, 2008) were obtained using the hanging-drop vapour-diffusion method with the following conditions: 0.1 M HEPES pH 7.5, 0.1 M NaCl, 2.0 M ammonium sulfate and 15% (v/v) glycerol. The crystal dimensions were typically 0.5 x 0.5 x 0.5 mm (Fig 5.35).



X-ray source	BL14.1 at BESSYII
Wavelength (Å)	0.91841
Space group	$P3_1$
Unit-cell parameters (Å)	125.93, 125.93, 164.24
Resolution (Å)	51.78-1.70 (1.81-1.70)
No. of observed reflections	1499388
No. of unique reflections	317895
R_{merge}	0.100 (0.402)
Completeness (%)	99.8 (98.9)
$\langle I/\sigma(I)\rangle$	4.6 (1.8)
V_M	4.02

Fig 5.35 Crystal of APS reductase from DvMF and the initial X-ray diffraction data (Ogata et al, 2008). Values in parentheses are for the highest resolution shell (1.81-1.70 Å). $R_{merge} = \Sigma |I_i - \langle I \rangle| / \Sigma I$, where I_i is the intensity of the ith observation and $\langle I \rangle$ is the mean intensity of the reflections.

5.5.3 Structure determination

The crystal structure of APS reductase from DvMF was solved by the molecular-replacement method using the program CNS (Brünger et al., 1998). As coordinates for the search model, the data for APS reductase from *A. fulgidus* (PDB code 1jnr) were used, which has 48% (α-subunit) and 65% (β-subunit) amino acid sequence identity to APS reductase from *D. vulgaris* Miyazaki F. After the calculation of the electron-density map using the molecular replacement solution, two heterodimers in the asymmetric unit that form an $\alpha_2\beta_2$ heterotetramer were observed. The electron-density map shows one FAD in the α-subunit and two $[Fe_4S_4]$ clusters in the β-subunit.

Since the APS reductase is highly conserved across a wide range of bacteria, the mode of catalytic activity is likely to be preserved. Nevertheless, work is being done to propose an exact reaction mechanism for the APS reductase of DvMF.

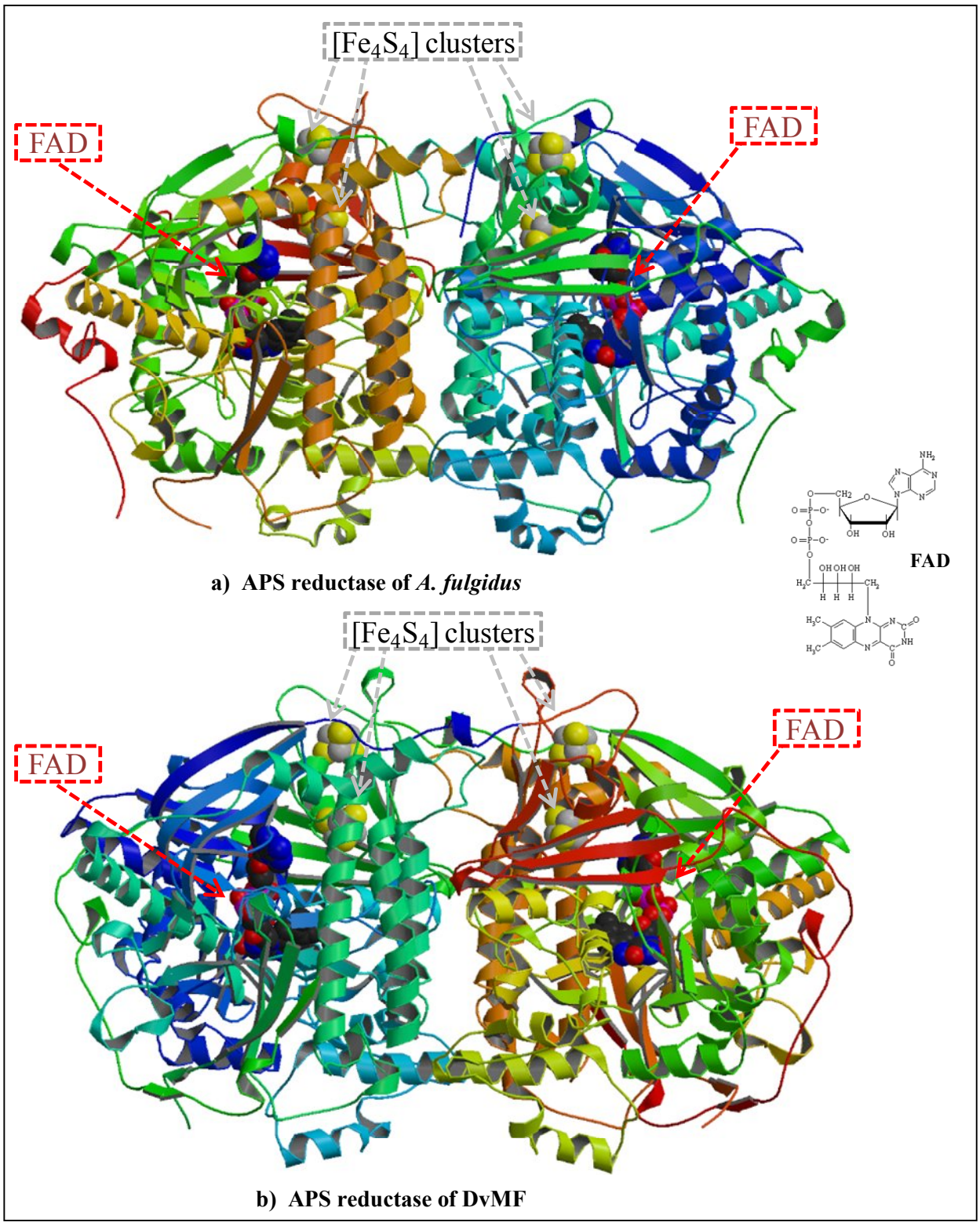


Fig 5.36 Schematic representation of the three dimensional models of (a) $\alpha_2\beta_2$ heterotetramer of APS reductase of *A. fulgidus* and (b) $\alpha_2\beta_2$ heterotetramer of APS reductase of DvMF (Ogata et al, unpublished). Also shown the chemical structure of Flavin adenine dinucleotide (FAD).

5.5.4 Purification and Characterization of Green protein

The Green protein was also purified from the soluble fraction of the DvMF cell lysate using gel exclusion and anion exchange columns (Chapter 3). The SDS-PAGE gel analysis of the finally purified fraction showed three subunits (Fig 5.37) which on MALDI analysis gave three corresponding peaks at 48 kDa, 42 kDa and at 11 kDa (Fig 5.38).

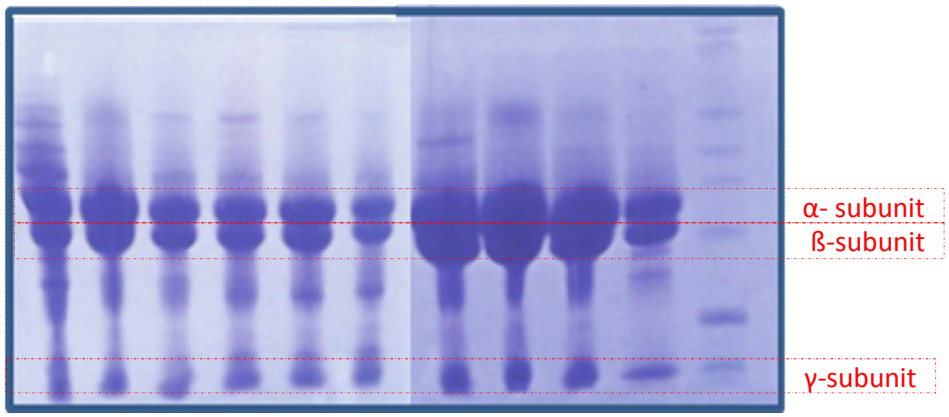


Fig 5.37 SDS-PAGE resolution of various fractions of purified Brown protein showed three major subunits which were later identified as α- and ß- and γ- subunits of Sulfite reductase or Desulfoviridin of DvMF.

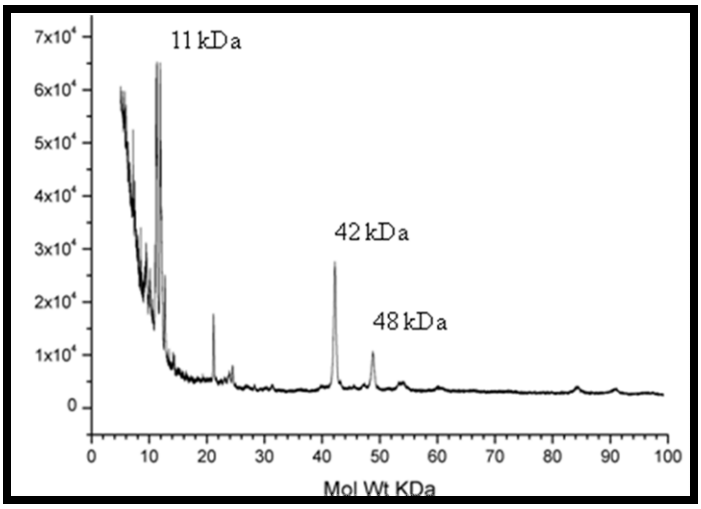


Fig 5.38 MALDI-TOF of the purified Green protein showed three peaks corresponding to the protein bands visible with SDS-PAGE; those were at 48 kDa, 42 kDa and at 11 kDa.

As for the Brown protein, the purified protein subunits of the Green protein were also gel-isolated and sent for N-terminal sequencing (University of Giessen). Also the peptide profile by MALDI-MS were noted and compared with available sequences. The N-terminal sequencing of the 11 kDa subunit was successful and gave the sequence AEVTYKGKSF, which is identical to the γ-subunits of the desulfoviridin proteins of DvH and DDG20. Further, the MALDI-MS

peptide profiling also gave some levels of sequence coverage which was compared with accessible sequences by NCBI-BLAST, 51% for the 48 kDa (Fig 5.39 and Fig 5.40), 56% for the 42 kDa (Fig 5.41 and Fig 5.42) subunit and 24% for the 11 kDa subunit (Fig 5.43 abd Fig 5.44). The deduced sequences indicated in synergy that the Green protein is the Desulfoviridin protein of DvMF, which should consist of at least these three subunits. Using the information thus generated, the corresponding three genes coding for the α-, β-, and γ- subunits, being present in two operons in the DvMF genome, were sequenced by means of cosmid library (Chapter 4).

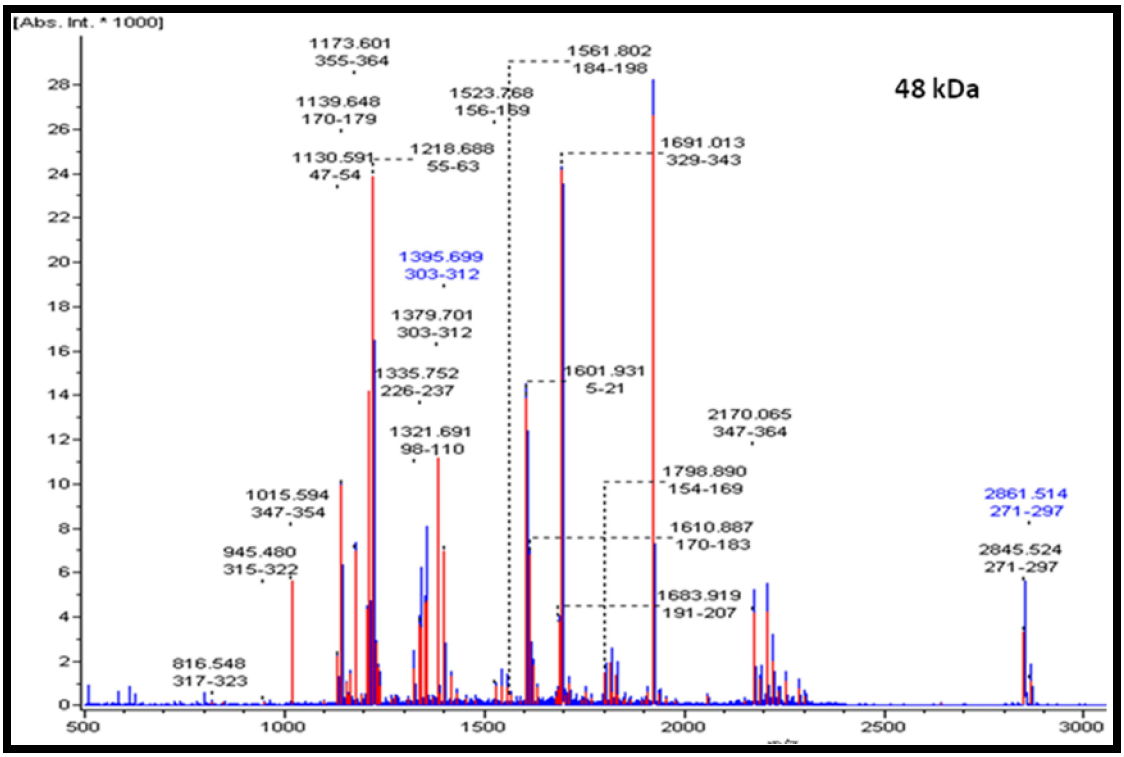


Fig 5.39 Peptide profile of the 48 kDa subunit of Desulfoviridin generated by high energy fragmentation of the polypeptide by using MALDI-MS, the molecular mass were calculated by MALDI and the residue numbers were given by protein aligment with available sequences.

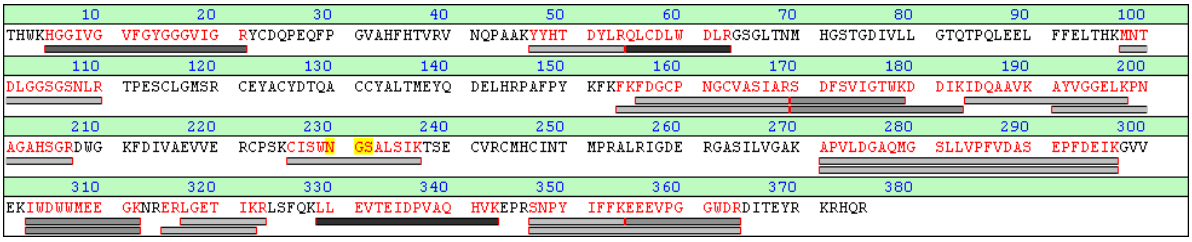


Fig 5.40 Fitting the data generated by peptide profiling of the 48 kDa α-subunit of Desulfoviridin by MALDI-MS into the available sequence database of NCBI by means of BLAST program.

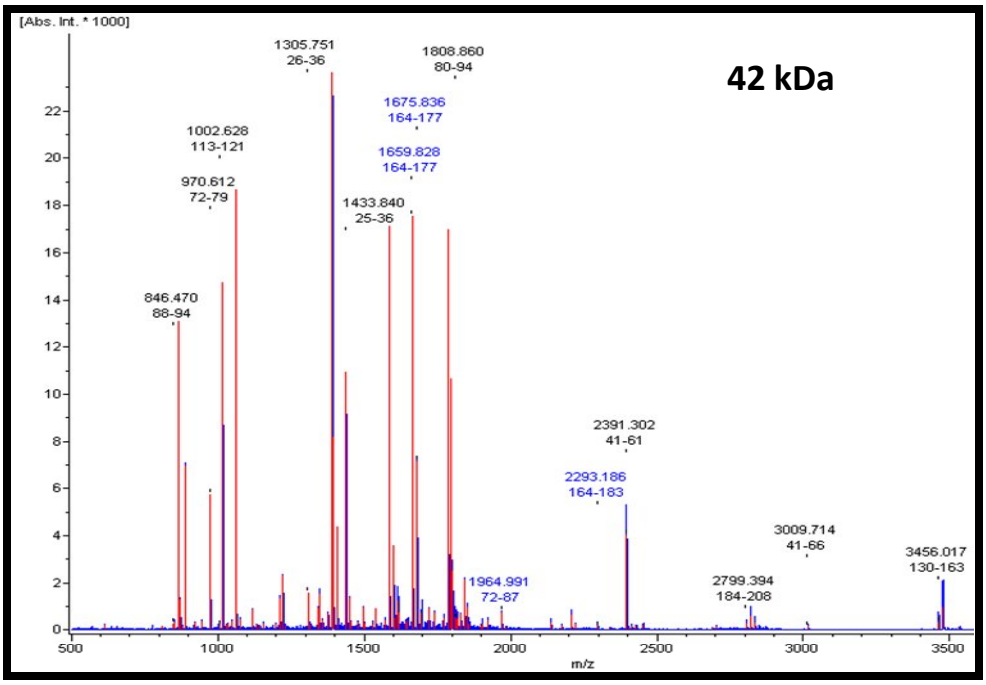


Fig 5.41 Peptide profile of the 42 kDa ß-subunit of Desulfoviridin generated by high energy fragmentation of the polypeptide by using MALDI-MS, the molecular mass were calculated by MALDI and the residue numbers were given by protein aligment with available sequences

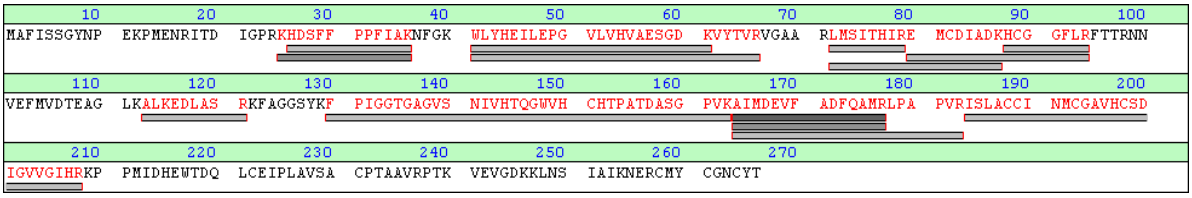


Fig 5.42 Fitting the data generated by peptide profiling of the 42 kDa ß-subunit of Desulfoviridin by MALDI-MS into the available sequence database of NCBI by means of BLAST program.

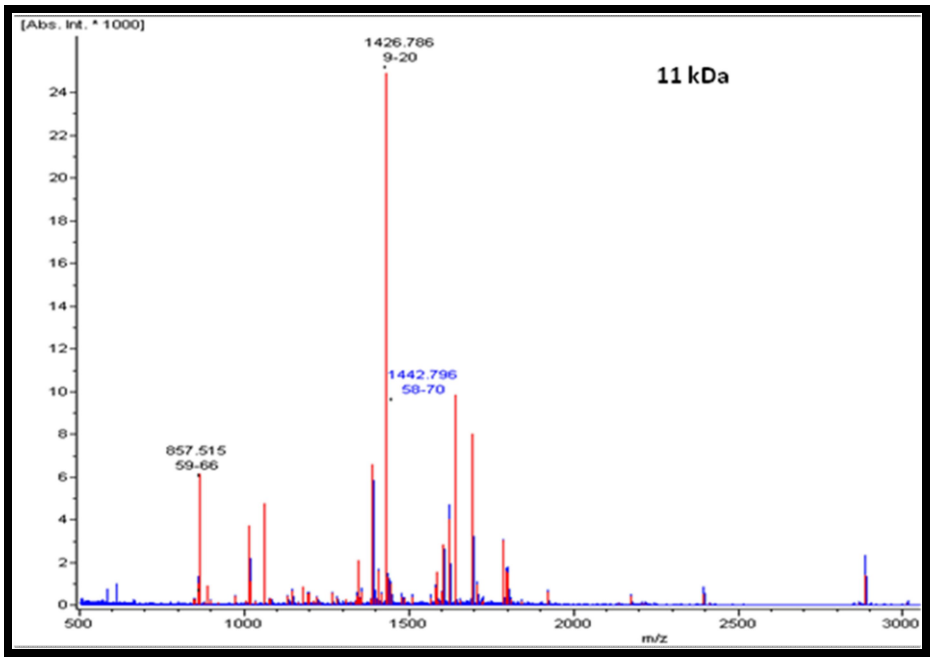


Fig 5.43 Peptide profile of the 11 kDa γ-subunit of Desulfoviridin generated by high energy fragmentation of the polypeptide by using MALDI-MS, the molecular mass were calculated by MALDI and the residue numbers were given by protein aligment with available sequences

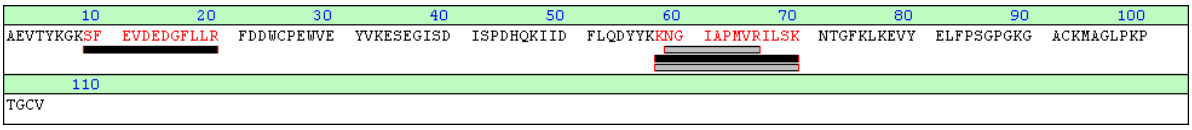


Fig 5.44 Fitting the data generated by peptide profiling of 11 kDa, γ-subunit of Desulfoviridin by MALDI-MS into the available sequence database of NCBI by means of BLAST program.

5.5.5 Crystallization of sulfite reductase of DvMF

After optimization, crystals suitable for diffraction experiments (Ogata et al, 2008) were obtained using the hanging-drop vapour-diffusion method with the following conditions: 20% PEG1000, 100 mM Imidazole pH 8.0, 200 mM Calcium Acetate. The crystal dimensions were of size: 0.2 x 0.1 x 0.1 mm (Fig 5.45). The structure determination using these crystals is underway.

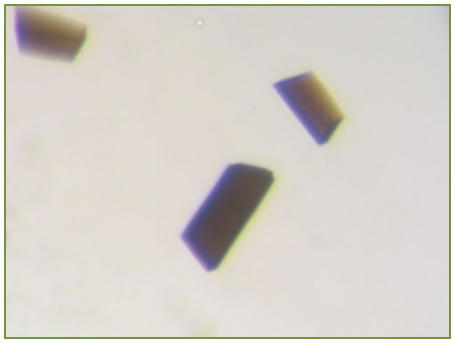


Fig 5.45 Crystal of Dsr complex from DvMF: Size: 0.2 x 0.1 x 0.1 mm (Ogata et al, unpublished)

5.6 Discussion

The template-based three dimensional structure models for the maturation proteins HypD, HypE, HypB and HynD as well as for the seleno-cysteine containing hydrogenase [NiFeSe] were presented. Also the two periplasmic hydrogenases of DvMF were characterized biochemically. The maturation proteins HypA and HypF were analysed by primary sequence conservation. The characterization of the maturation Hyp proteins of DvMF has shown that these proteins are well conserved in this organism and exhibit all the possible features so far identified in the [NiFeSe] hydrogenase maturation. On the other hand, there is a possiblity that the [NiFeSe] hydrogenase may exhibit some difference from other [NiFeSe] hydrogenases of the *Desulfovibrio* sps in the number or type of iron and sulfur atoms in its medial cluster. Further charcterization at the sequential level has shown that the [NiFeSe] hydrogenase retains features that may indicate an involvement of the lipoprotein synthesis pathway in its final maturation and transport.

The cellular expression of the two hydrogenases of DvMF was detected in presence of excess of nickel and selenium (this work) and was found to be in functional agreement with the observation of up-regulation of both hydrogenases in the presence of nickel, and the down-regulation of the expression of [NiFe] hydrogenase of DvH (Valente et al, 2006) in the presence of selenium. This observation was also extended to *D. desulfuricans* (this study). The expression of the [NiFeSe] hydrogenase was enhanced only slightly when selenium was supplied in presence of nickel, however, this investigation was compromised by the relatively low sensitivity

of the antibody. The effect of repression looked starker in case of DvMF as compared to *D. desulfuricans* (this study) and as reported for DvH (Valente et al, 2006). Further the presence and up-regulation of the [NiFeSe] hydrogenase in the DvMF cell extract, prepared from the cells grown in excess of selenium and nickel, was established with Native-PAGE BV/TTC assays. The role of selenium in the regulation of the non-selenium enzyme has been thought to be a result of a secondary regulation that means not the free selenium but a selenium associated cellular component is responsible for the down or up-regulationas of nonselenium and selenium containing enzymes respectively (Rother et al, 2003). In a study on the archaeon *M. maripaludis,* a nonsense mutation in the *selB* reverses the suppression of non-selenium enzymes (Rother et al, 2003) and usually the level of *selB* is directly associated with the availability of free selenium. Thus selenium is likely to be indirectly involved in the regulation of [NiFe] hydrogenase by a secondary mechanism involved in the synthesis and integration of selenocysteine into the selenium containing polypeptides.

Advancing the efforts in the direction of heterologous expression, the expression of the hydrogenases and the maturation gene operons in *E. coli* has been investigated. The two two-gene maturation gene operons *hypAB* and *hypDE* were cloned in a T7 promoter expression vectors to generate constructs pRSF2hypAB and pRSF2hypDE, respectively. The western blot analysis of total protein extracts showed that the conjunction of the two genes present in a single operon is functional in *E. coli*, when each one of these two operons is individually placed in behind of a single T7 promoter. The N-terminal of the small subunit of the [NiFe] hydrogenase undergoes a tat-pathway dependent cleavage, while HynC cleaves the C-terminal of the large subnit after the insertion of nickel. Purification studies and sequential characterizations have shown that the N-terminus of the large subunit is closely associated with the plasma membrane (Yagi et al, 1976, Higuchi et al., 1997). Thus, probably the C-terminal end of the small subunit is the only available free end that can be epitope-tagged and used for purification. In this work, the four-gene operon of [NiFeSe] hydrogenase was marked by tagging a strep-tag in the C-terminal region of the small subunit and was later cloned and expressed from a T7-promoter.

The [NiFe] hydrogenase, expressed under the control of the T7-promoter of pRSF2hynA-strepBCD in *E. coli*, also carried an N-terminal histidine tag in the HynB protein and a C-

terminal tag in the HynD protein. The total protein extract prepared from these *E. coli* cells was probed by western blotting using different antibodies. The results showed an unambiguous expression of the two subunits of hydrogenase and gave also hints to a transfer of some proportion of the small subunit to the periplasm. Purification of the soluble fraction of the total protein extract showed a competent refinement of the hydrogenase small subunit using the strep-tag and a noticeable co-purification of the two subunits.

The two subunits of the [NiFeSe] hydrogenase were also expressed in *E. coli* together in the T7 derived plasmid pRSF2hysBA, and later the large subunit alone was expressed in the constructs pRSFhysA and pRSFhysApreseleno. Agreeing with the construction, the recombinant plasmid pRSF2hysBA (Fig 5.28) led to the expression of a histidine tag in the N-terminal of the HysB (Fig 5.29) as well as also showed a presence of S-tag in the C-terminal of HysA (Fig 5.29 and Fig 5.30). Similarly, the expression of a C-terminal S-tag was observed in the western blot detection of cell lysates of pRSF-hysA (Fig 5.30) and pRSF-hysA-preseleno (Fig 5.30) as well. Such an expression of the S-tag in the large subunit of the [NiFeSe] hydrogenase, when expressed from pRSF2-hysBA and pRSF-hysA indicated an integration of selenocysteine corresponding to the TGC codon present in the C-terminal of the *hysA* gene instead of termination of the translation in *E. coli*.

The integration of selenocysteine in selenium-containing proteins is signaled by the presence of special secondary structures in the mRNA known as SECIS (Berry et al, 1993). In *E. coli*, a special tRNA (refer to fig 9.1 in appendix) encoded by *SelC* is responsible for the integration of selenocysteine. The *selC* gene product is charged first by a seryl-tRNA synthetase to yield seryl-tRNA. The serine residue in seryl-tRNA is converted to selenocysteine by the *SelA* gene product, a selenocysteine synthase. This reaction needs selenium-donating species, monoseleniumphosphate, being a product of a cellular acting seleniumphosphate synthase, a SelD gene product.

These four gene products responsible for the selenocysteine integration, SelA, SelB, *SelC* and seryl-tRNA synthetase are very specific in action and their compatibility with coded RNA secondary structures is very species-specific. An attempt to express the selenocysteine-

containing large subunit of the hydrogenase of *Desulfomicrobium baculatus* when fused with lacZ, has failed in *E. coli* (Tormay and Böck, 1997), attributed to the SelB specificity. Apart from the specificity of the SelB-mRNA interaction, a structural compatibility of the quaternary complex with the ribosome is required. The large subunit of the [NiFeSe] hydrogenase of DvMF has a very similar structure with that of *D. baculatus*; however, we have clearly shown the successful incorporation of selenocysteine by the detection of the S-tag by western blotting. Since this antibody is raised against the peptide following the selenocysteine, it indicates that the selenocysteine codon is not read as termination signal. Nonetheless these studies are in preliminary stages and the proteins needs to be purified to confirm any non-specific detection by antibody combined with some spectroscopic measurements. It is also possible to check the presence of the S-tag by an agarose RNAase assay (Novagen).

For the sulfate metabolism-related proteins of DvMF, the three dimensional structure of the APS reductase is determined at 1.7 Å resolution. This structure is the first one in the class of sulfate reducing bacteria and work is continued to determine and propose a reaction mechanism. Further work is being carried out to determine the structure and as well as to improve the quality and resolutions for the crystals obtained for the desulfoviridin complex of DvMF.

Chapter 6

Phylogenetic Analysis of DvMF Hydrogenases as Representatives of *Desulfovibrio* sps.

The biological consumption or evolution of hydrogen for the sake of producing cellular energy has been possibly invented at the earliest form of life. The origin of the proton-dependent chemiosmotic mechanism for ATP synthesis may also be seen in conjunction with the formation of proton gradients created by hydrogenases on either side of the cytoplasmic membrane.

Even today, diverse micro-organisms have retained the capability to use hydrogen to fuel their metabolism. Where animal systems are limited to use oxygen as oxidizing agent, bacterial systems have adapted to exploit alternative oxidizing components. H_2 metabolism has been detected and investigated in a variety of prokaryotes belonging to wide physiological ranges. This includes either of the aerobes and anaerobes, autotrophs and heterotrophs, prokaryotic and eukaryotic photosynthetic organisms, "knallgas" bacteria, methanogens, sulphate reducers, N_2-fixers, fermentative organisms, hyperthermophiles, parasitic protozoans, and anaerobic fungi. Also characterized for hydrogen metabolism are the Eubacteria, Cyanobacteria and the Proteobacteria (purple bacteria), in which all five subdivisions (α, β, δ, ε, & γ) are among the best studied ones. Then there are those microbes which are less well studied and are spread in widely divergent classes. These groups include the Spirochaetes, the Cytophaga/Flexibacter/Bacteroides group, Fibrobacters, and both high (*Frankia* sps) and low G+C content (saccharolytic *Clostridia* sps) Gram positive bacteria, the green non-sulfur bacteria (GNS) and the planctomyces (Robert Robsen, *Hydrogen as Fuel*).

6.1 Phylogenetic Studies of Hydrogenases

Hydrogenases are considered to be modular enzymes which have retained hydrogen metabolizing capability even after frequent gene exchange and reshuffling during the course of evolution. Detailed evolution studies encompassing hydrogenase gene content, co-occurrence

and transmission have revealed that most of the hydrogenase genes have evolved by vertical transmission, though some horizontal gene transfers have been proposed from Archaea to bacteria, or bacteria to bacteria and from bacteria to anaerobic protists (Vignais and Billoud, 2007).

Using the simplest approach, an insight into the evolution of H_2 metabolism can be gained by the comparison of different hydrogenases and their accessory proteins at the primary sequence level. The choice of proteins to be subjected to a comparison is influenced by the number of homologous and analogous sequences available, their occurrence and the presumption of a functional analogy such as the potential similarity between the various hydrogenase- and maturation-encoding genes. The relationships among these proteins are studied here via their phylogenetic relation. A phylogenetic tree is created by computer-aided alignments of the amino acid sequences (or also of the nucleotide sequences of the encoding genes) and the calculation of the degree of similarity between the sequences. This resulting similarity matrix is used to create a schematic tree showing the relatedness in a diagrammatic way. Either complete protein sequences could be used, or some degree of editing of the sequences is performed to maximise the alignment. Alternatively, only those segments of molecules which have high sequence identity, i.e., showing few or no gaps can be compared to represent important conserved domains within a protein.

The availability of a larger number of truly homologous sequences and the comparison of such tree with a life tree of 16S/18S ribosomal RNA adds to the significance of such study. One of the two basic types of trees is a rooted tree that computes or assumes a common ancestor. However, the rooting of trees can be erroneous when the sequences are quite divergent, and the functionality is not analogous or is unknown. This applies in most cases when considering both the hydrogenase structural and accessory gene products. In such cases, the construction of an un-rooted type of trees does not have to expain for many of such assumptions. Figure 6.1 is showing a rooted phylogenetic relationship between the hydrogenase-possessing species belonging to the various classes of bacteria and arachea based on their 16s ribosomal RNA sequences.

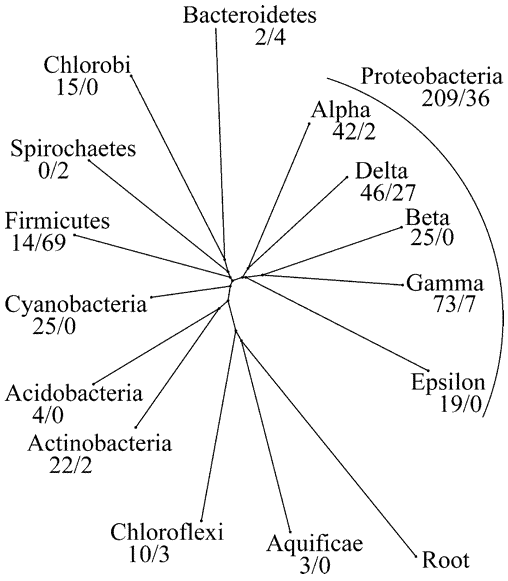


Fig 6.1 Phylogenetic tree of bacteria derived from the 16s rRNA sequences using data from Archaea and Eukaryota for the root. Numbers at the ends of the branches represent the number of hydrogenase genes known in species of that group. The figure on the left of the slash represents the number of [NiFe]-hydrogenases and the one on the right the number of [FeFe]-hydrogenases. (Adapted from Vignais and Billoud, 2007)

6.2 Hydrogenases of DvMF

Many present day species are known to contain more than one type of [NiFe] hydrogenases and some also contain [FeFe] along with that. Whereas the [NiFe] hydrogenase presence is widespread, the [FeFe] hydrogenases are present mainly in Gram-positive bacteria and in species belonging to the γ- and δ- divisions of proteobacteria.

Whole genome sequencing of some δ-proteobacteria (*Desulfovibrio vulgaris* subsp Hildenborough, *Desulfovibrio desulfuricans* G20, *Geobacter uraniumreducens* RF4, *Geobacter sulfurreducens*) have revealed a high degree of diversity in occurrence and co-occurrence of the various types of hydrogenases and also the cellular locations. Screening a cosmid genomic library has led to the identification of three additional hydrogenases other than the standard [NiFe] hydrogenases in the DvMF genome (Chapter 4). Two of them are now completely sequenced and are included in these phylogenetic comparisons.

6.2.1 Phylogenetic tree of [NiFe] and [NiFeSe] hydrogenases of DvMF

A protein sequence blast of the large subunit of [NiFeSe] hydrogenase of DvMF at the NCBI database pulls out sequences from [NiFe] and [NiFeSe] hydrogenases of some other bacteria belonging to proteobacterial subdivisions as well as some firmicutes. A phylogenetic tree (Fig 6.2) based on the top hit sequences from the representative organisms belonging to these different divisions, shows the relationship among various classes and subgroups of bacteria. The *Desulfovibrio* group of bacteria are amongst the only few known genus of the Desulfovibrionales order to possess a [NiFeSe] hydrogenase.

In the phylogram (Fig 6.2), constructed using the translated DNA sequences from the large subunit of the [NiFeSe] hydrogenase, i.e. HysA, all of the few [NiFeSe] hydrogenases sequences available till date are clustered on one node which is well separated from the node containing [NiFe] hydrogenase sequences present in the same organisms. The degree of relation among similar types of hydrogenases, i.e. among the [NiFeSe] sequences of organisms belonging to the order Desulfovibrionales and then among the non-selenium standard [NiFe] hydrogenase of the same organisms is almost identical.

The [NiFe] hydrogenase sequences of the organisms belonging to δ-proteobacteria other than the *Desulfovibrio* sps are present on two major branches. *Delta-proteo MLMS-1* and *Desulfotalea psychrophila* form one collective out-group and the other node contains mainly *Geobacter* sps. The *Geobacter* species hold a relatively closer propinquity with the α- and γ- subdivisions of proteobacteria as well as with the Green Non-Sulfur bacteria as compared to the *Desulfovibrio* sps. Assessing in an order of relatedness, the order of phylogenetic proximity of *Geobacter* sps is; α-proteobacteria > GNS (Green Non Sulfur) bacteria> γ –proteobacteria.

Fig 6.2 Phylogenetic tree for HysA. The alignment with various sequences was carried out by using NCBI blast, and the PhyML method (phylogenies by maximum likelihood) was used to calculate the tree. FigTree software (Author, Andrew Rambaut, 2007) was used to draw the tree, and the root lengths have been automatically calculated. The abbreviations used are; *nife* for [NiFe] hydrogenases, *nifese* for [NiFeSe] hydrogenase, DvH, *Desulfovibrio vulgaris* subsp Hildenborough, DvMF for *Desulfovibrio vulgaris* subsp Miyazaki F, DDG20 for *Desulfovibrio desulfuricans* G20

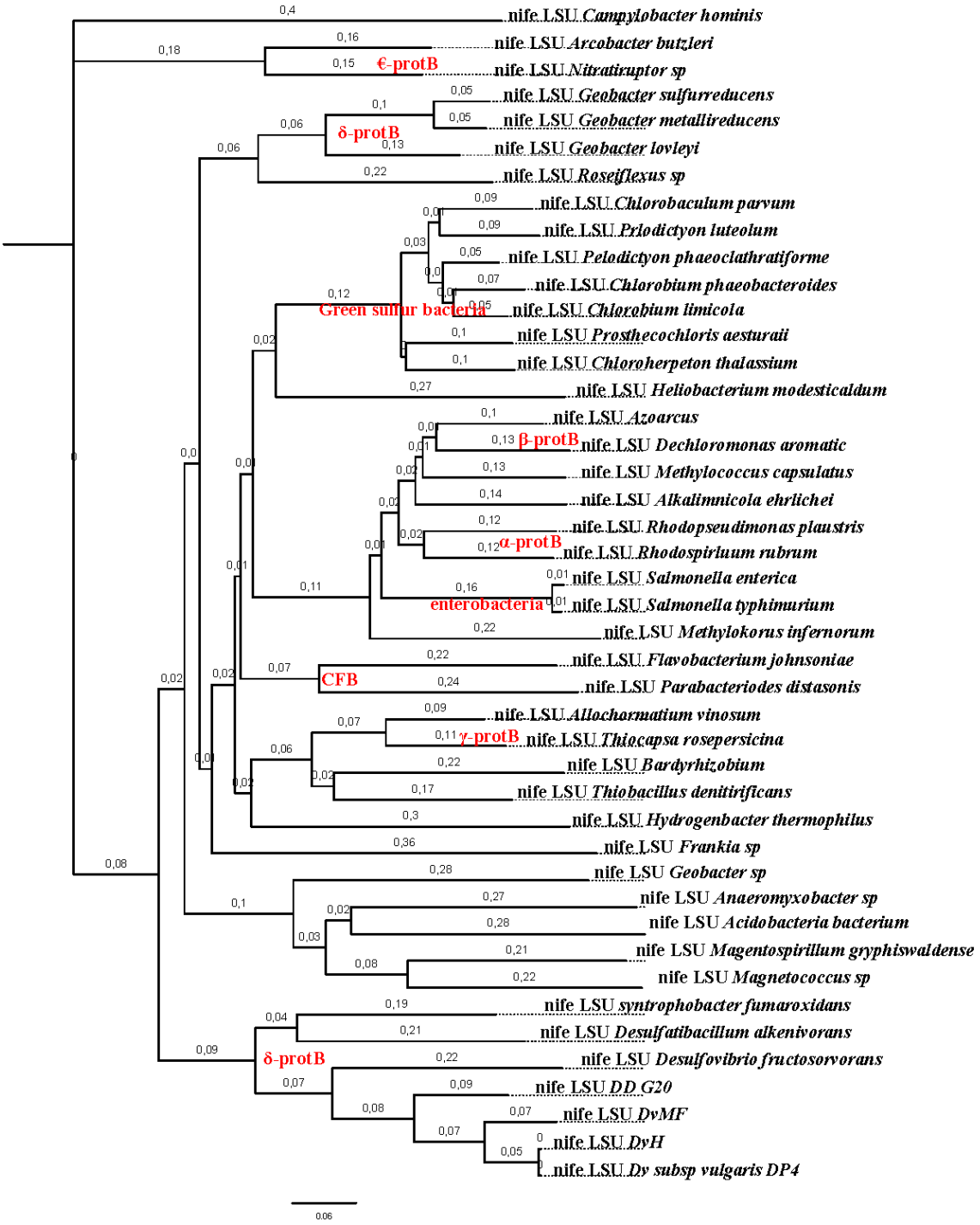


Fig 6.3 Phylogenetic tree for HynA. The alignment with various sequences was carried by using NCBI blast, and to draw the tree (similarly as in Fig 6.2). The abbreviations used are; *nife* for [NiFe] hydrogenases, *nifese* for [NiFeSe] hydrogenase, DvH, *Desulfovibrio vulgaris* subsp Hildenborough, DvMF for *Desulfovibrio vulgaris* subsp Miyazaki F, DDG20 for *Desulfovibrio desulfuricans* G20

To give an even closer look at the phylogenetic relationship among the [NiFe] hydrogenases, i.e. HynA, one more phylogenetic tree is constructed based on the polypeptide sequences of the large subunit of [NiFe] hydrogenase of DvMF (Fig 6.3). The usual aerobic or microaerobic organisms belonging to the α-proteobacteria subdivision form a separate node connected with the node containing the β-proteobacteria.

Giving an overall look to this tree, the α-proteobacteria are related to the other groups in an order of proximity, β-proteobacteria> enetrobacteriaece (γ-proteobacteria) > green sulphur bacteria > CFB (Cytophaga-Flexibacteria-Bacteroides) group of bacteria> photosynthetic γ-proteobacteria. Again the *Geobacter* group of δ-proteobacteria is closely placed to the nodes containing other proteobacterial subdivisions rather than *Desulfovibrio* group of δ- proteobacteria.

6.2.2 Phylogenetic tree of Energy-Conserving Hydrogenases (Ech)

Ech hydrogenases are H_2-evolving, energy conserving, and membrane associated multimeric enzymes that are placed among the group 4 of hydrogenases (six or more subunits). This class of hydrogenases is known to reduce protons coming from the anaerobic oxidation of low potential organic compounds such as acetate or formate to H_2.

The majority of hydrogenases assigned to group 4 have been found in Archaea including *Methanosarcina barkeri*, *Methanothermobacter marburgensis* and *Pyrococcus furious*. The Ech hydrogenase operon found in *Desulfovibrio* sps and *Thermoanaerobacter tengcongensis* are thought to be acquired by the horizontal gene transfer from an archaebacterium belonging to *Methanosarcina* clade (Calteau et al, 2005).

The Ech hydrogenase operon of DvMF consists of six subunits *EchABCDEF*. The structural characteristics have been discussed in the chapter 4. Phylogenetic trees are created using the polypeptide sequences of EchA (larger of the two membrane anchoring subunits; Fig 6.4), EchE (larger of the two subunits involved in the hydrogen splitting; Fig 6.5) and EchF (Ferredoxin like protein; Fig 6.6).

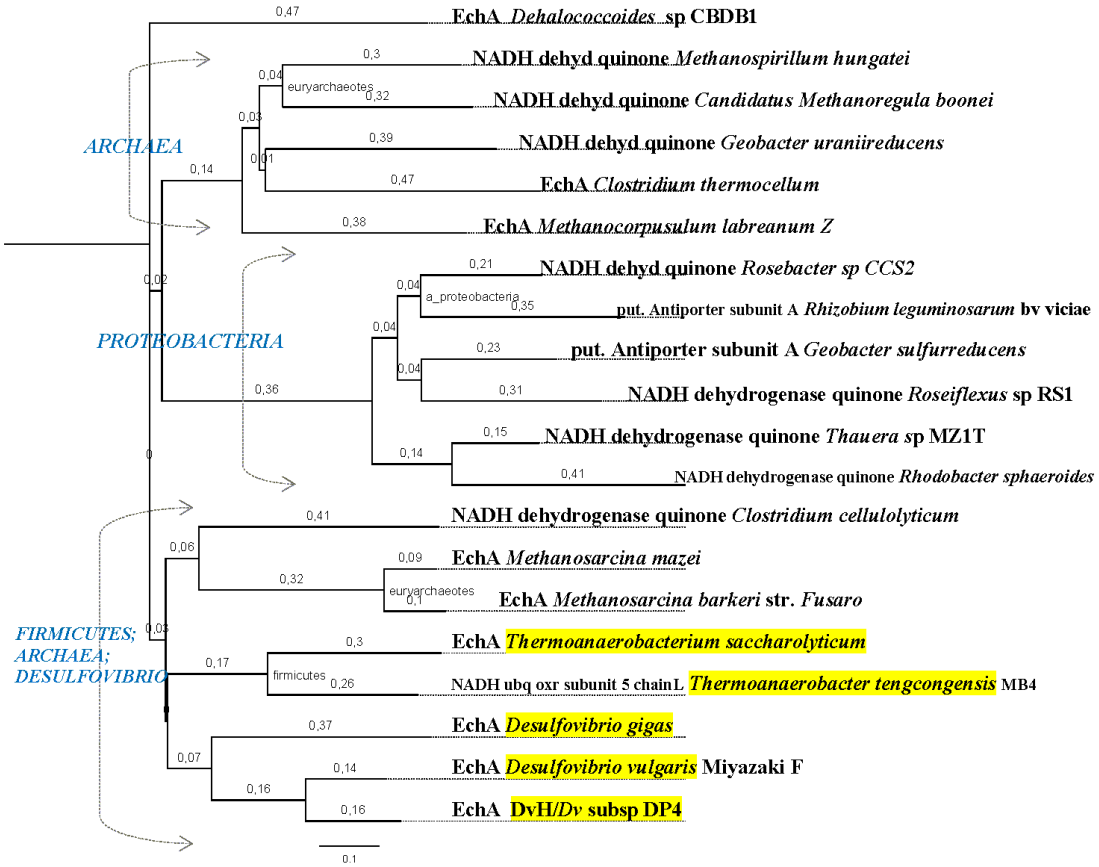


Fig 6.4 Phylogenetic tree for EchA. The alignment with various sequences was carried by using NCBI blast, and used to draw the tree as before. The abbreviations used are; oxr: oxidoreductase, Ubq: Ubiquinone, DvH, *Desulfovibrio vulgaris* subsp Hildenborough, DvMF for *Desulfovibrio vulgaris* subsp Miyazaki F. The sequences closely related to DvMF are highlighted in yellow.

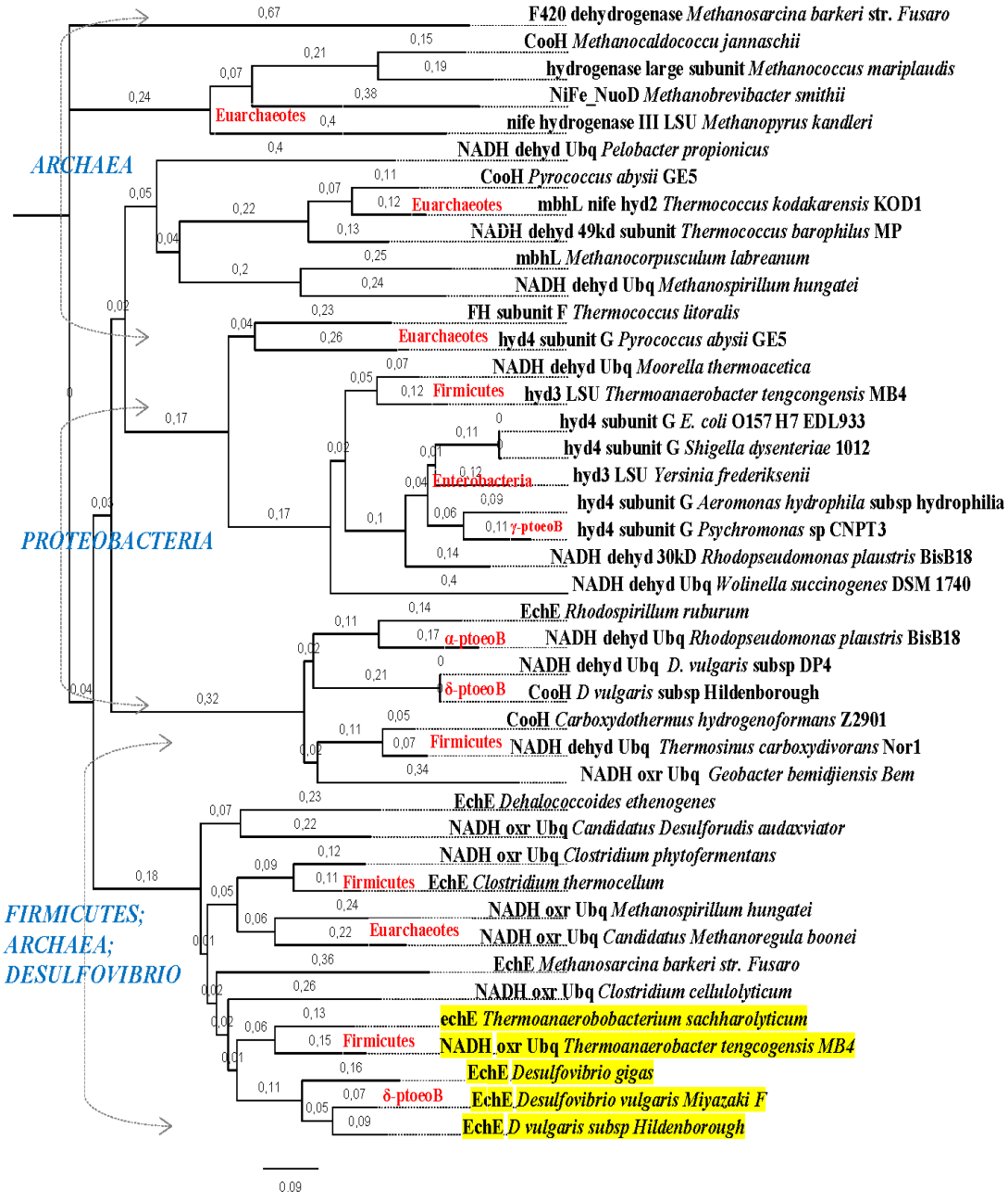


Fig 6.5 Phylogenetic tree for EchE. The alignment with various sequences was carried by using NCBI blast, and was used to draw the tree as before. The abbreviations used are; oxr: oxidoreductase, Ubq: Ubiquinone, DvH, *Desulfovibrio vulgaris* subsp Hildenborough, DvMF for *Desulfovibrio vulgaris* subsp Miyazaki F. The sequences closely related to DvMF are highlighted in yellow.

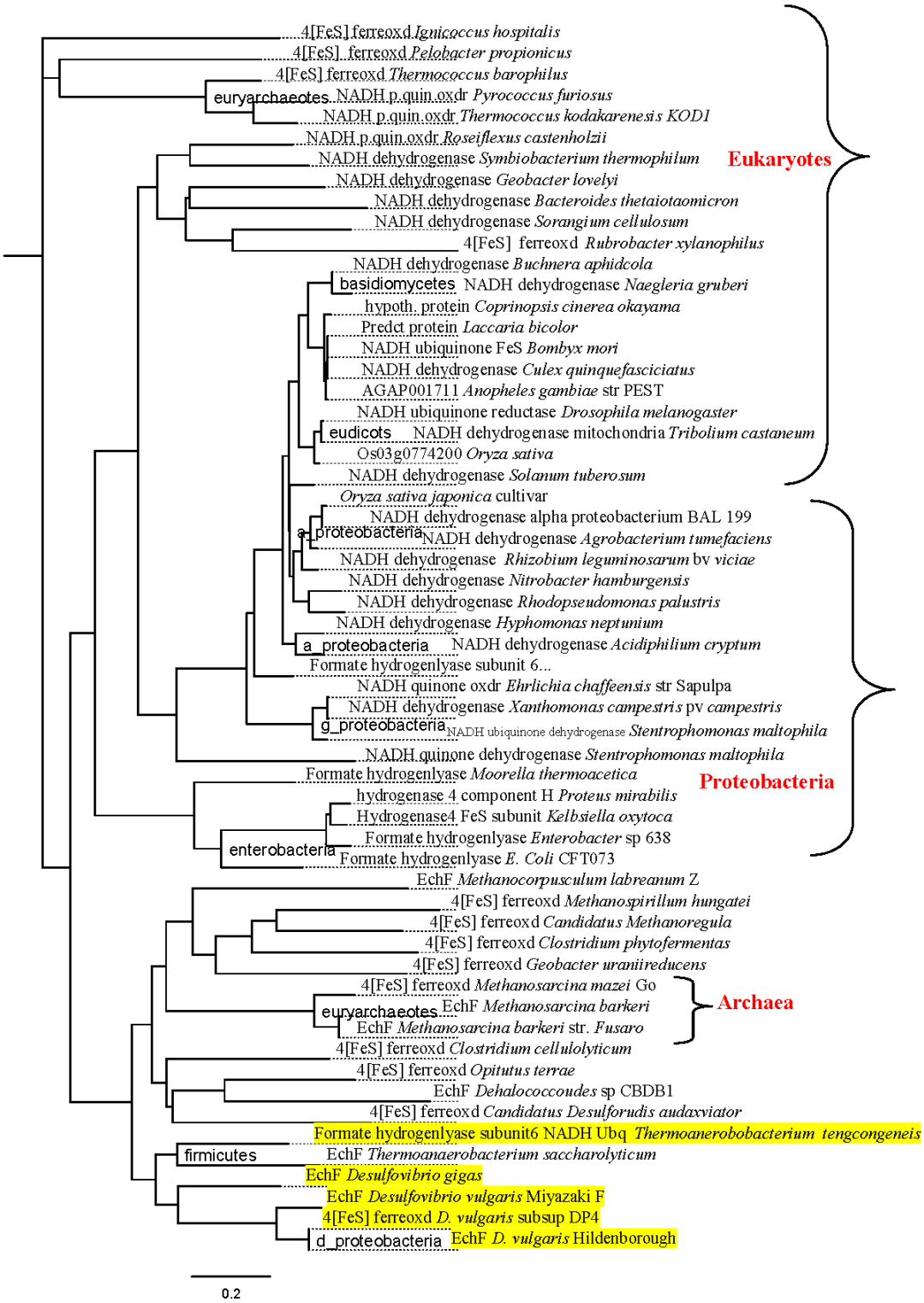


Fig 6.6 Phylogenetic tree for EchF. The alignment with various sequences was carried out by using NCBI blast, and was used to draw the tree as before. The abbreviations used are; oxr: oxidoreductase, Ubq: Ubiquinone, DvH, *Desulfovibrio vulgaris* subsp Hildenborough, DvMF for *Desulfovibrio vulgaris* susbsp Miyazaki F. The sequences closely related to DvMF are highlighted in yellow.

6.3 Phylogeny of Hydrogenase Pleiotropy Proteins (hyp)

Recent advances in the genome sequencing of various groups of microbes have provided new information to the comprehensive studies of the distribution and inter-relation among various groups of bacteria. The study of hydrogenase genetics and phylogeny now includes many proteins, in addition to regulatory and structural proteins, which are required for the functioning of the various hydrogenases. These accessory proteins are required for the uptake and insertion of Ni and the synthesis and insertion of metal clusters and their ligands into the nascent enzymes. In bacteria like *Ralstonia eutropha*, the whole hydrogenase machinery consisting of multiple polycistronic operons is present on a single megaplasmid, whereas in some other bacteria such as *E. coli* the structural and maturation genes are present at the most in two separate polycistronic operons. On the other hand, in case of δ-Proteobacteria, the hydrogenase machinery is present in several polycistronic operons scattered throughout the genome (Chapter 4).

Five Hyp proteins (HypA, HypB, HypC, HypD and HypE) are found in the DvH and other sequenced *Desulfovibrio* genomes. These proteins are thought to be involved in hydrogenase maturation in a similar way to the homologous proteins studied in *E. coli* (Chapter 4). The presence of *hyp* genes in a genome has been invariably associated with the presence of [NiFe] and [NiFeSe] hydrogenases. The evolutionary relationships among the *hyp* genes have been found to be more or less consistent with that of the hydrogenases.

6.3.1 Phylogeny of HypA and HypB

HypA and HypB are expressed from one operon and involved in the nickel insertion into the active sites of the [NiFe] and [NiFeSe] hydrogenases. Two separate phylograms are constructed using an alignment of the HypA and HypB protein sequences deduced by the translation of the sequenced hypAB operon of DvMF. The evolutionary relationships shown by these two phylograms (Fig 6.7 and Fig 6.8) portray a similar degree of relatedness among various groups and classes of bacteria.

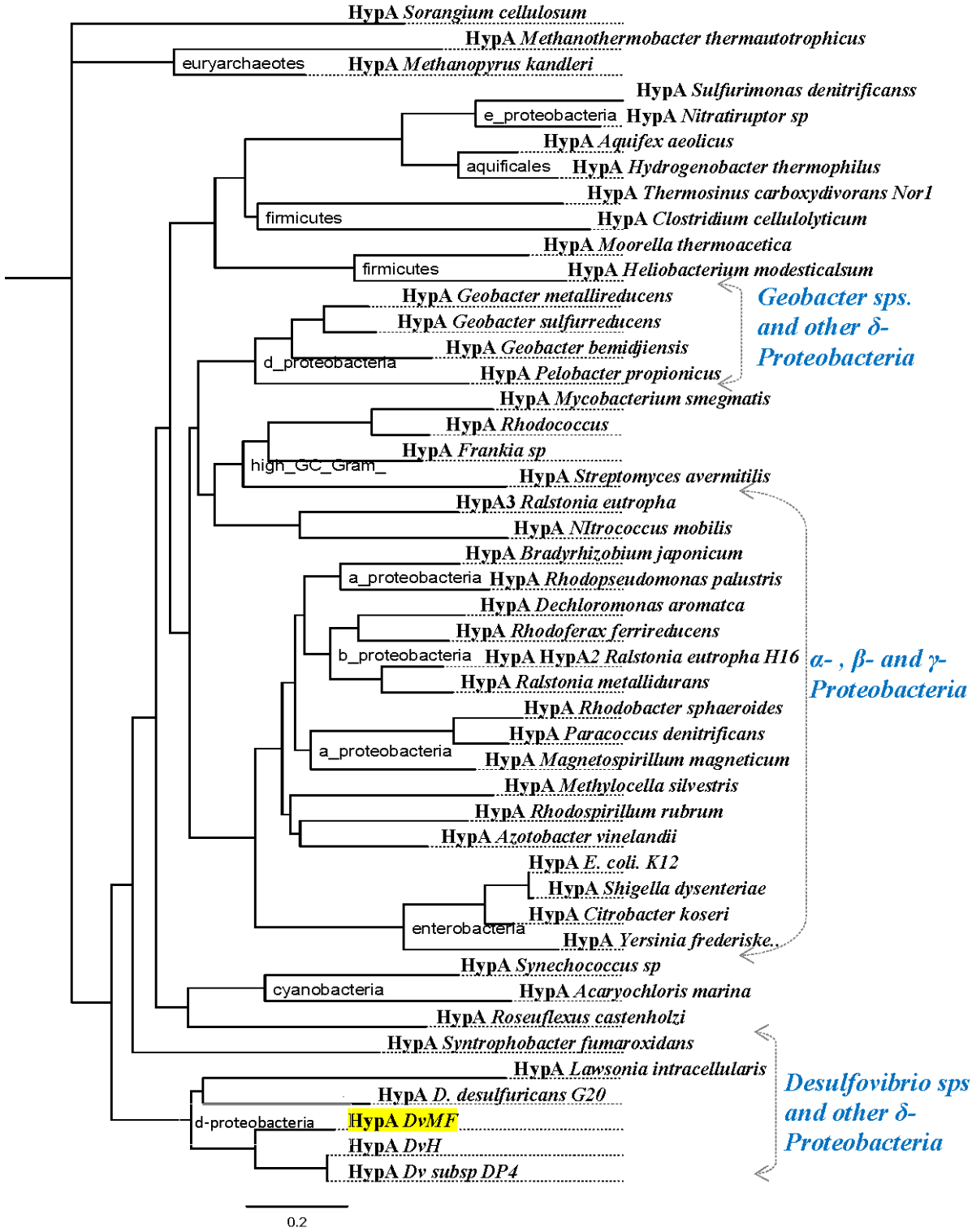


Fig 6.7 Phylogenetic tree for HypA. The alignment with various sequences was carried by using NCBI blast, and was used to draw the tree as before. The abbreviations used are; DvH for *Desulfovibrio vulgaris* subsp Hildenborough, DvMF for *Desulfovibrio vulgaris* subsp Miyazaki F.

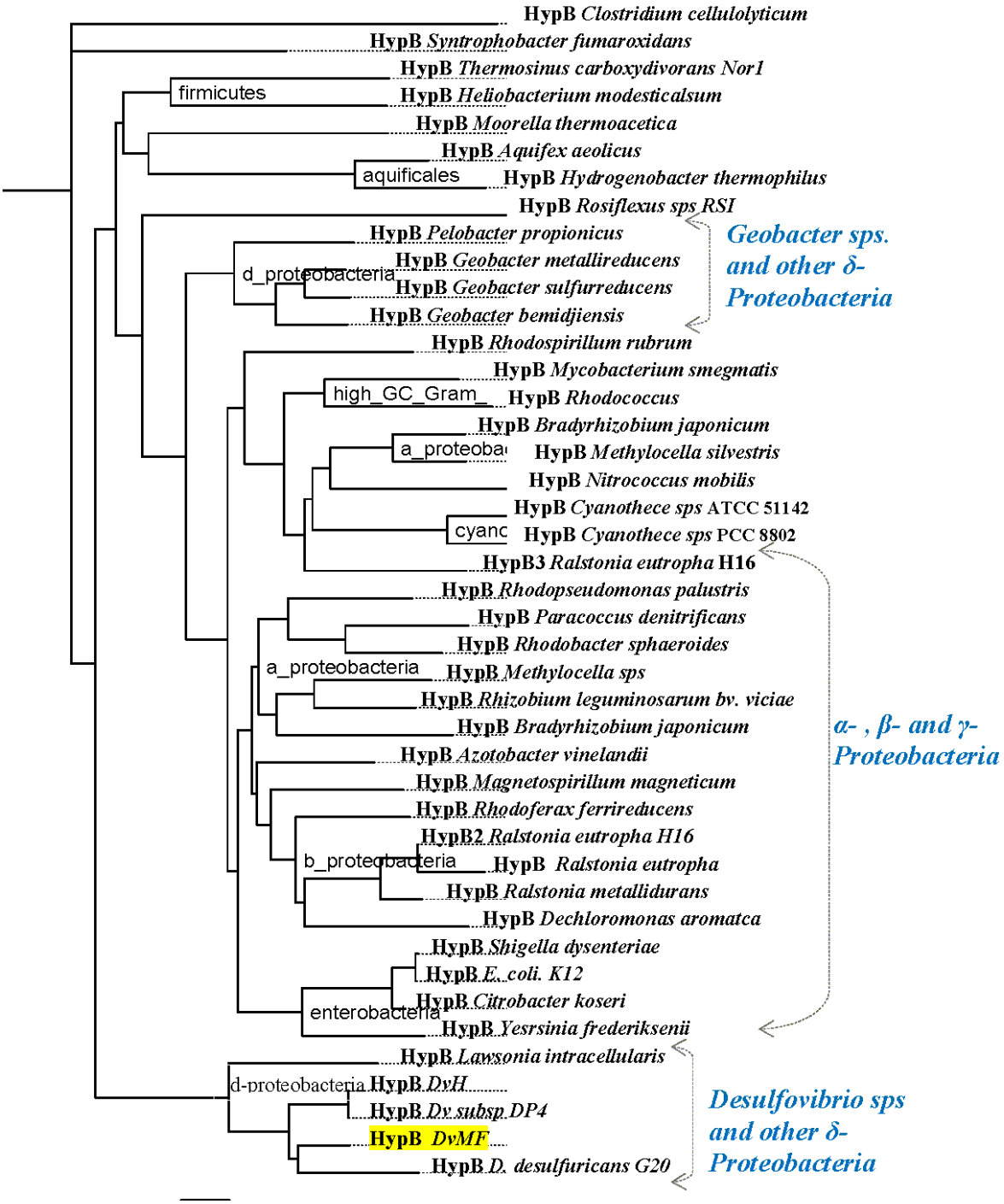


Fig 6.8 Phylogenetic tree for HypB. The alignment with various sequences was carried by using NCBI blast, and was used to draw the tree as before. The abbreviations used are; DvH for *Desulfovibrio vulgaris* subsp Hildenborough, DvMF for *Desulfovibrio vulgaris* subsp Miyazaki F.

Amongst the δ-proteobacteria, the *Desulfovibrio* sps together with *Lawsonia intracellularis* form an out-group located on a separate branch from *Geobacter* sps. The nodes containing other subdivisions of proteobacteria are more closely related to *Geobacter* sps than *Desulfovibrio* sps. Further, the α- and β- proteobacteria are more closely related to each other than to γ-proteobacteria.

6.3.2 Phylogeny of HypD and HypE

HypD and HypE are two proteins transcribed from the same operon with overlapping reading frames. Hydrogenase maturation studies (Böck et al, 2006) and crystal structures of these proteins (Watanabe et al, 2007) have shown that HypD together with HynC is responsible for iron insertion into the catalytic centre of [NiFe] hydrogenase, and HypE cooperates with HypF for the cyanation of the iron centre (Watanabe et al, 2007, Böck et al, 2006)

The phylograms for HypD and HypE are constructed using the polypeptide sequences deduced from the translation of sequenced *hypDE* operons. The species belonging to the δ-proteobacteria are distributed in several nodes on the phylogram for hypD. The *Desulfovibrio* sps are again placed on one node, in close proximity to *Desulfatibacillum alkenivorans* and *Desulfotalea psychrophila*. The other subdivisions of proteobacteria are again placed nearer to the *Geobacter* sps. *Syntrophobacter fumaroxidans* and *Syntrophus acidotrophicus* form another out-grouped node of similar order as that of *Geobacter* sps.

The phylogram for hypE shows an almost similar distribution of different species belonging to proteobacteria. However, the node containing the *Desulfovibrio* sps rather than the *Geobacter* sps are closely placed to other proteobacteria. The *Desulfovibrio* sps along with *Lawsonia intracellularis* are present on one node which is present at the similar level of the singularly placed *Syntrophobacter fumarioxidans* and also the node branching into closely related species belonging to α- and β- proteobacteria.

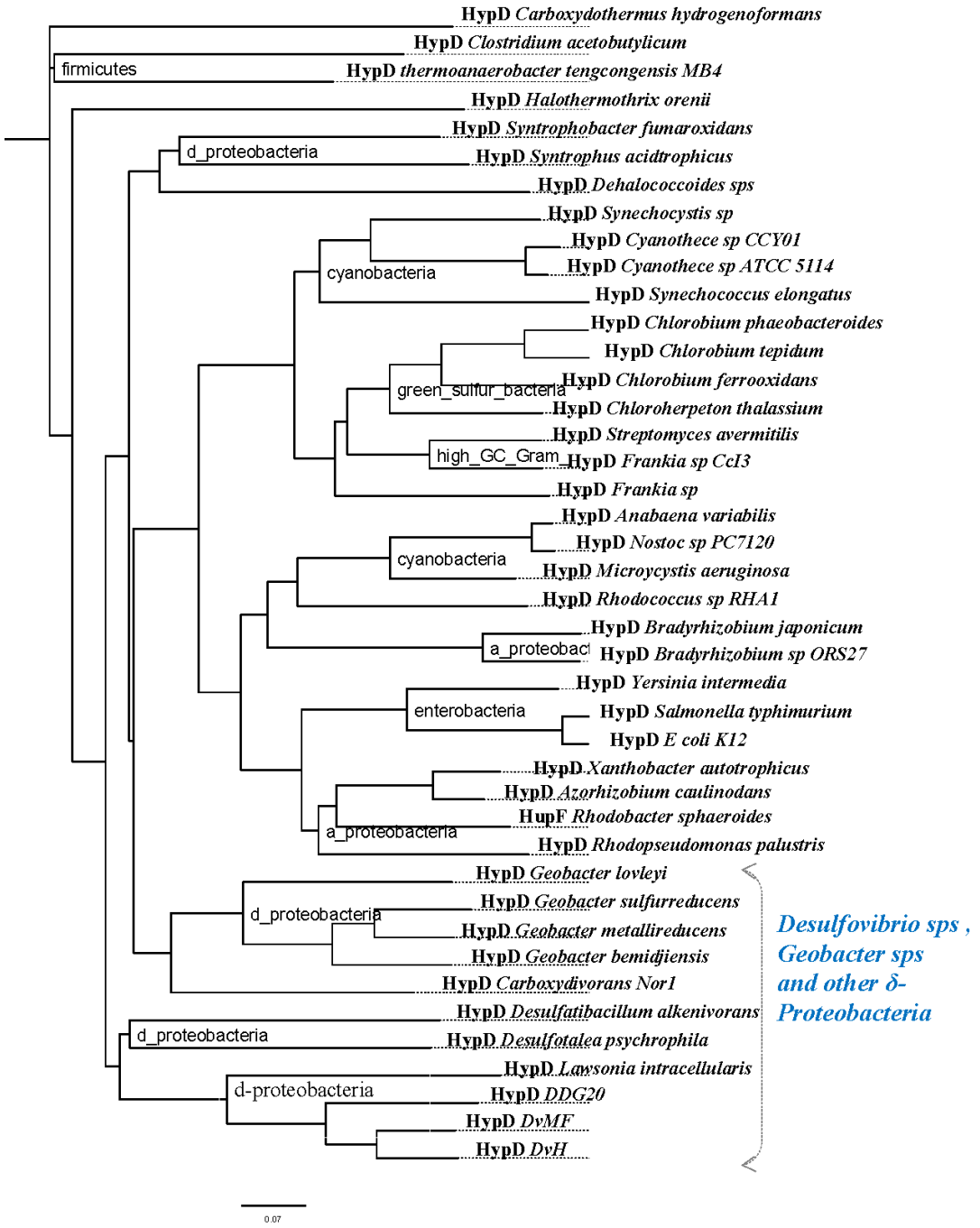


Fig 6.9 Phylogenetic tree for HypD. The alignment with various sequences was carried out by using NCBI blast, and was used to draw the tree as before. The abbreviations used are; DvH for *Desulfovibrio vulgaris* subsp Hildenborough, DvMF for *Desulfovibrio vulgaris* subsp Miyazaki F, DDG20 for *Desulfovibrio desulfuricans* G20

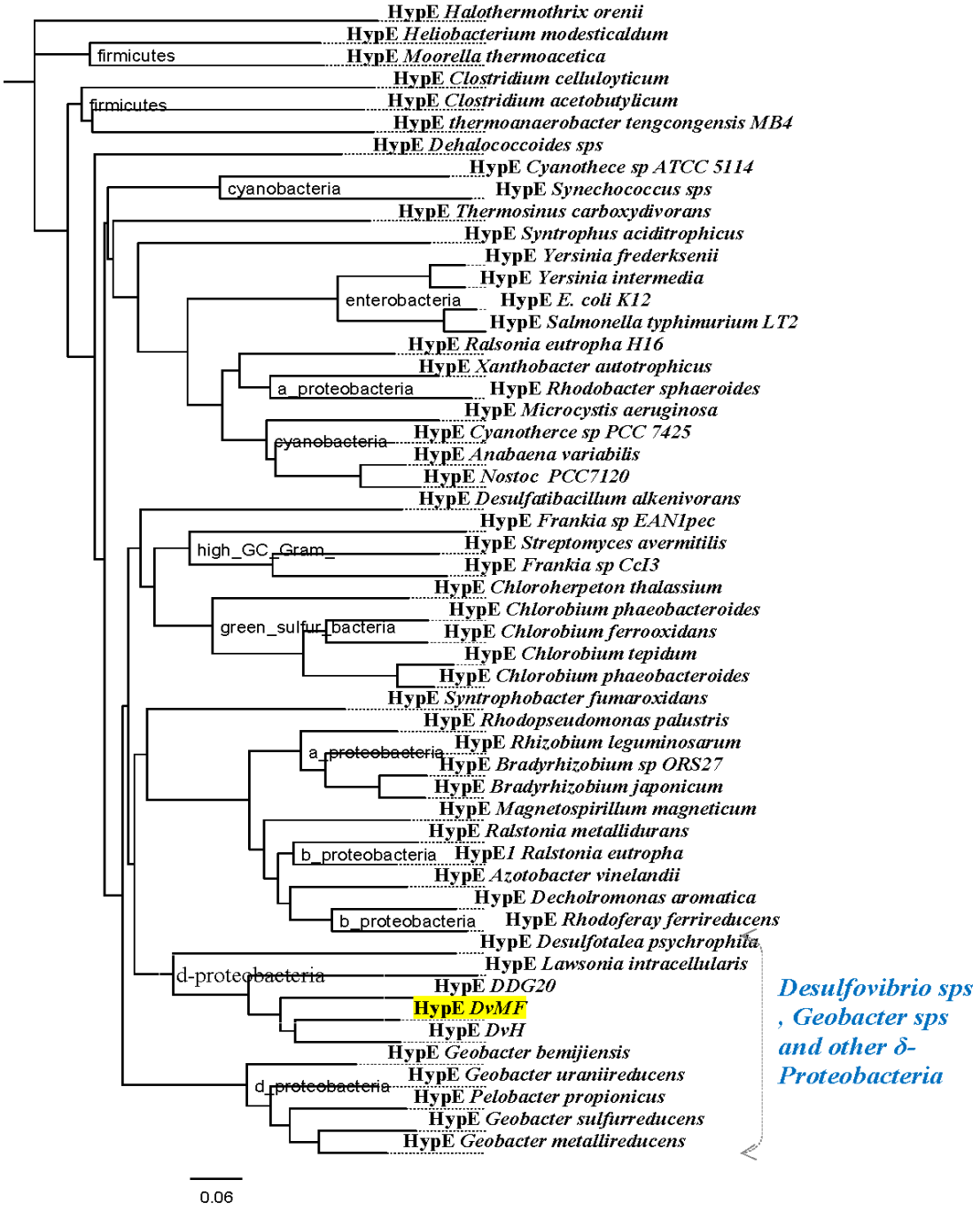


Fig 6.10 Phylogenetic tree for HypE. The alignment with various sequences was carried out by using NCBI blast, and was used to draw the tree as before. The abbreviations used are; DvH for *Desulfovibrio vulgaris* subsp Hildenborough, DvMF for *Desulfovibrio vulgaris* subsp Miyazaki F, DDG20 for *Desulfovibrio desulfuricans* G20

6.3.3 Phylogeny of HypF

The HypF protein is associated with the synthesis of the cyanide ligands of the catalytic site iron atom of the [NiFe] hydrogenase. The capacity to synthesize the cyanide ligands is widely distributed among various organisms, ranging from bacteria to plants (Knowles, 1976).

A phylogram is constructed based on the HypF protein sequence (Fig 6.11) deduced from the translation of the sequenced hypF gene of DvMF and its homologues extracted from NCBI database by BLAST. The distribution of various classes and subgroups of bacteria is very similar to that of the [NiFe] hydrogenase. The species belonging to the δ-proteobacteria subdivision are clustered upon three separate nodes of different levels. The *Desulfovibrio* sps are grouped on one node, located well separated from the *Geobacter* sps. Again the *Desulfatibacillum alkenivorans* and *Desulfotalea psychrophila* like in case of the phylogram of HypD form an outgroup, but are placed closer to the *Geobacter* sps this time. The branches containing α- and β-proteobacteria are closer to the *Desulfovibrio* rather than the nodes containing the other two δ-proteobacteria clusters.

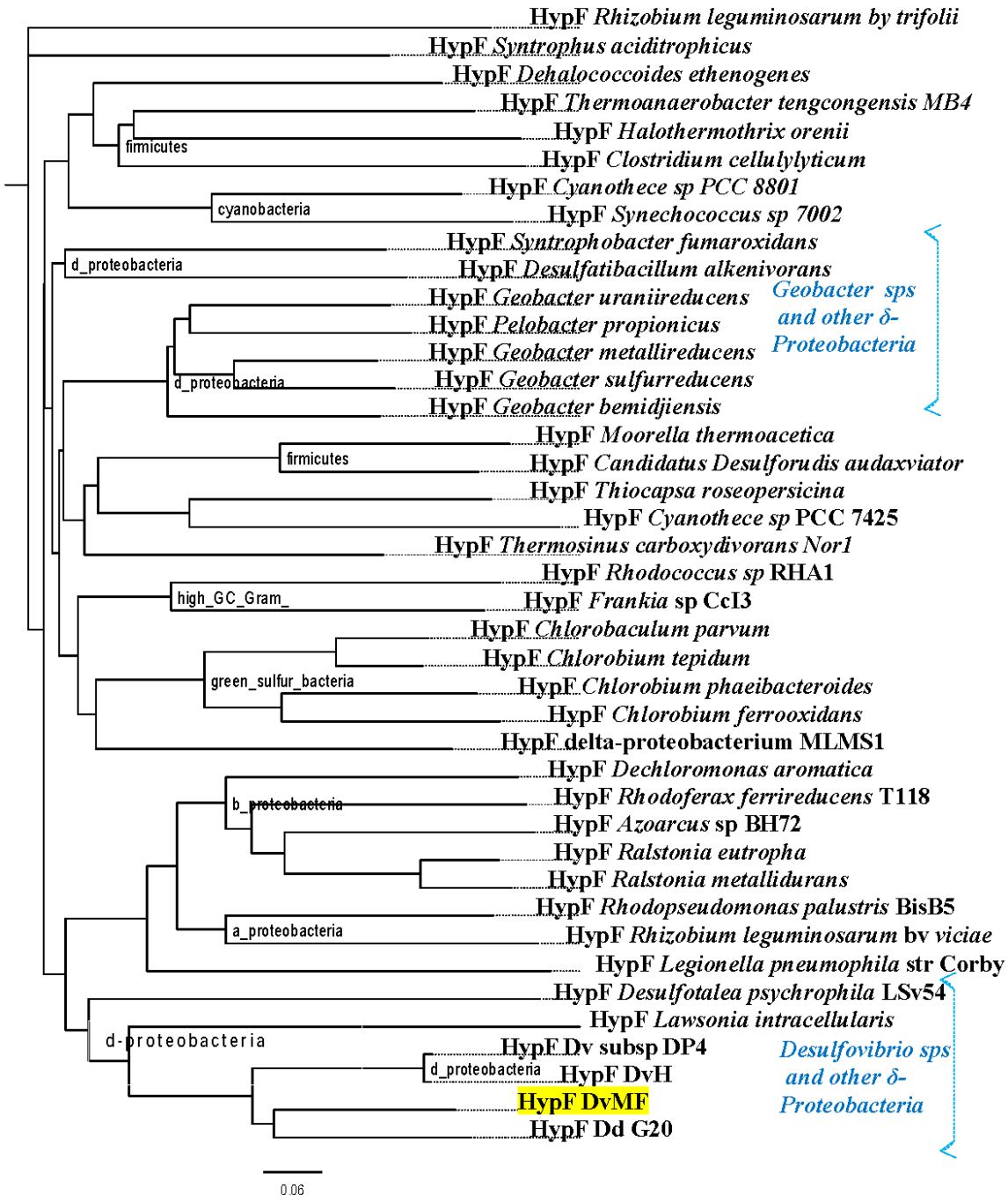


Fig 6.11 Phylogenetic tree for HypF. The alignment with various sequences was carried out by using NCBI blast, and was used to draw the tree as before. The abbreviations used are; DvH for *Desulfovibrio vulgaris* subsp Hildenborough, DvMF for *Desulfovibrio vulgaris* subsp Miyazaki F, DDG20 for *Desulfovibrio desulfuricans* G20

6.4 Co-evolution of Hydrogenases and Sulfate Metabolism

The oxidation and reduction of sulphur-containing compounds plays a vital role in the natural sulfur cycle. The utilization of such compounds for energy-yielding processes seems to have developed during a very early phase of prokaryotic evolution (Skyring and Donnelly et al, 1982). Discovery of a bi-functional sulfhydrogenase found in *Pyrococcus furiosus* (a combination of subunits for hydrogenase and sulfite reductase) which can dispose an excess of reductants either by the reduction of protons to H_2 or S_0 to H_2S (Ma et al, 1993, Pedroni et al, 1995) has led to the speculations that coupling of H_2 and sulphur metabolisms was also of fundamental importance in the origin of life.

6.4.1 The Dissimilatory Sulfate Metabolism Genes

Further aspects studied in this work are the phylogeny of the enzymes involved in sulfate metabolism, based on proteins of that function being found and identified in DvMF (Chapter 3). Where utilization of sulfur compounds as electron donor or acceptor takes place, the sulfate reducing bacteria (SRB) are broadly classified into four major groups: Gram negative mesophilic SRB (delta-proteobacteria), Gram positive pore forming SRB (*Desulfatomaculum*, *Bacillus* and *Clostridium* sps), thermophilic bacterial SRB (*Thermodesulfobacterium commune* and *Thermodesulfovibrio yellowstonii*) and thermophilic aracheal SRB (*Archaeoglobus* sps., *Halobacterium halobium*, *Thermoplasma acidophilum*) (Castro et al, 2000). The phylogenetic analysis of the enzymes participating in the dissimilatory sulfur reduction and oxidation of reduced sulfur compounds is capable of providing insights into the evolution of such metabolism. Two of such enzymes are studied in this work, the 5'-Adenylylsulfate (APS) reductase (AprAB) and the Dissimilatory sulfite reductase (DSR complex), also named Desulfoviridin (DsrABCD) of DvMF.

6.4.2 Phylogeny of 5'-Adenylylsulfate (APS) Reductase (AprAB)

The purified APS reductase from DvMF (Ogata et al, 2008) and of most of the other organisms studied till date is a conserved heterodimeric flavoprotein. In our phylogenetic analysis the polypeptide sequences of AprA of DvMF have been used to BLAST at the NCBI database to pull out similar sequences from the representatives of four main groups possessing sulfate

metabolism, as there are Green Sulfur bacteria, hyperthermophilic eubacteria, proteobacteria and Archaea. All these have been brought together on one phylogram (Fig 6.12). The branches containing the sulfate reducing δ-proteobacteria taxons are distributed on two separate, but interconnected nodes, where one node comprises solely the *Desulfovibrio* sps. The node connecting the δ-proteobacteria to other subdivisions of proteobacteria branches into three sub-nodes, one β- and two clusters of γ-proteobacteria. The proteobacteria class as an entire group is closely placed to the nodes containing three clusters comprising of euryarchaeotes, hyperthermophilc bacteria and firmicutes. The organisms belonging to the green sulfur bacteria and the hyperthermophilic bacteria group form two distinct nodes which separate much before the division between the Eubacteria and Archaea. The Aps reductase of DvMF is most closely related to *D. termitidis* (97% identities) followed by DvH (92% identities) and *D. vulgaris* subsps DP4 (92% identities).

6.4.3 Phylogeny of Desulfoviridin (DSR) Proteins

The purified Desulfoviridin protein from *Desulfovibrio vulgaris* Miyazaki F consists of three polypeptides DsrA, DsrB and DsrC, whereas the presence of DsrD cannot be confirmed (Ogata et al, unpublished). Genetically, the α- (DsrA), β- (DsrB) and δ- (DsrD) subunits are present in consecutive positions possibly transcribed by a single operon, and the gamma (DsrC) subunit is present elsewhere in the genome. Yet, the purification studies of Dsr proteins from various *Desulfovibrio* sps have indicated a higher structural association of alpha and beta to the gamma subunit than to the delta subunit (Pierik et al, 1992, Laue et al, 2001).

In our study we have carried out the phylogenetic analysis comparing the tree topologies of the phylograms of DsrB (Fig 6.13), DsrC (Fig 6.14) and DsrD (Fig 6.15). DsrA and DsrB are two paralogous proteins and DsrB is likely to be the catalytically active subunit of the functional complex of Desulfovirdin (Oliveira et al, 2008). DsrD has been sequenced mainly from δ-proteobacteria and does not generate a detailed picture of the inter-relationship among various groups of bacteria. On the other hand, DsrC and its homologue proteins have been sequenced from several bacterial and Archaeal species.

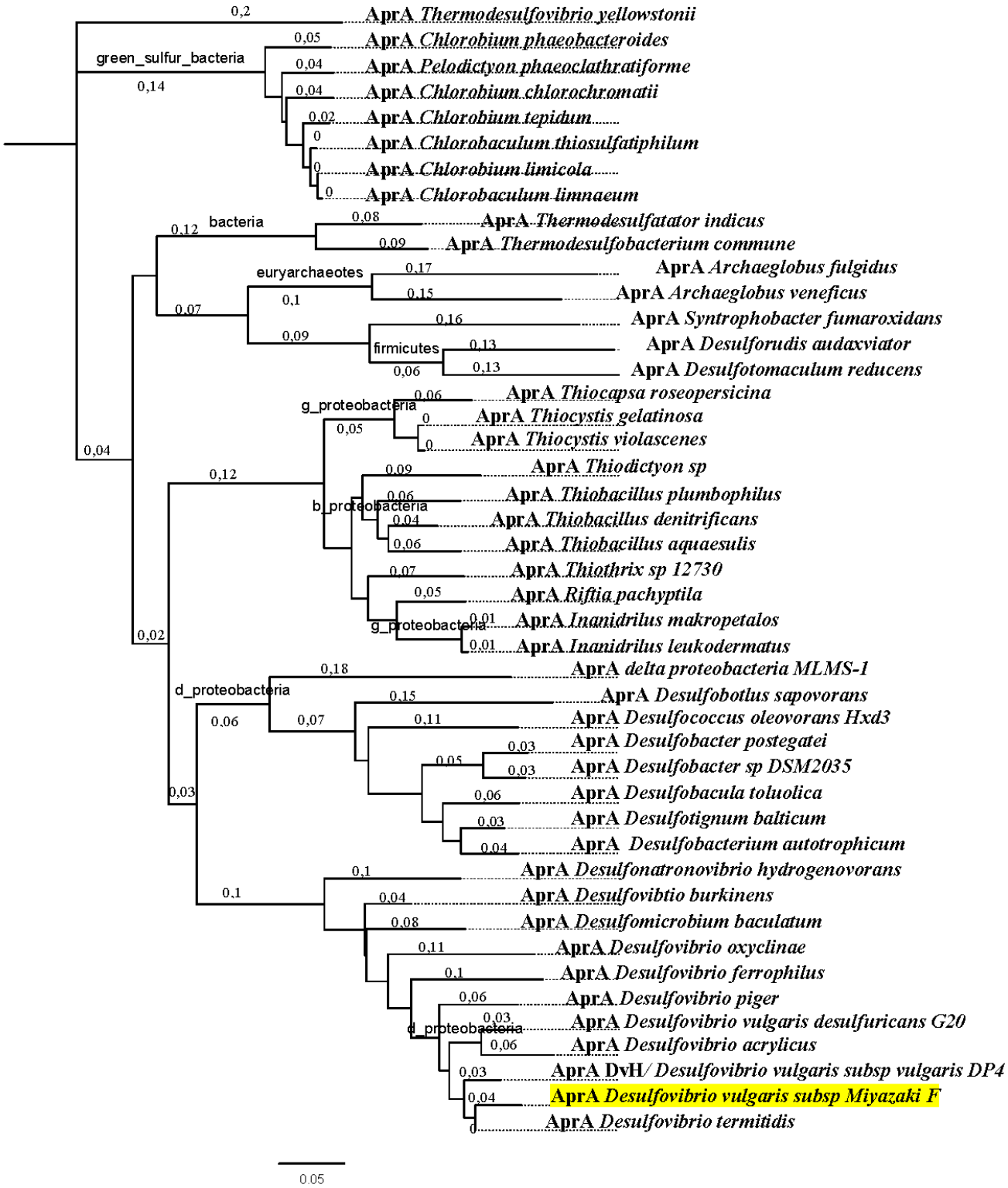


Fig 6.12 Phylogenetic tree for α-subunit, AprA of Adenylyl sulfate reductase. The alignment with various sequences was carried out by using NCBI blast, and was used to draw the tree as before.

The DsrC and its homologues present in the γ-proteobacteria are closely related to the corresponding proteins from Green Sulfur bacteria, and those of δ-proteobacteria are closer to the Firmicutes. The Archaeal nodes (Crenarchaeotes and Euryarchaeotes) segregate well before leading into the node containing other bacterial divisions and subgroups.

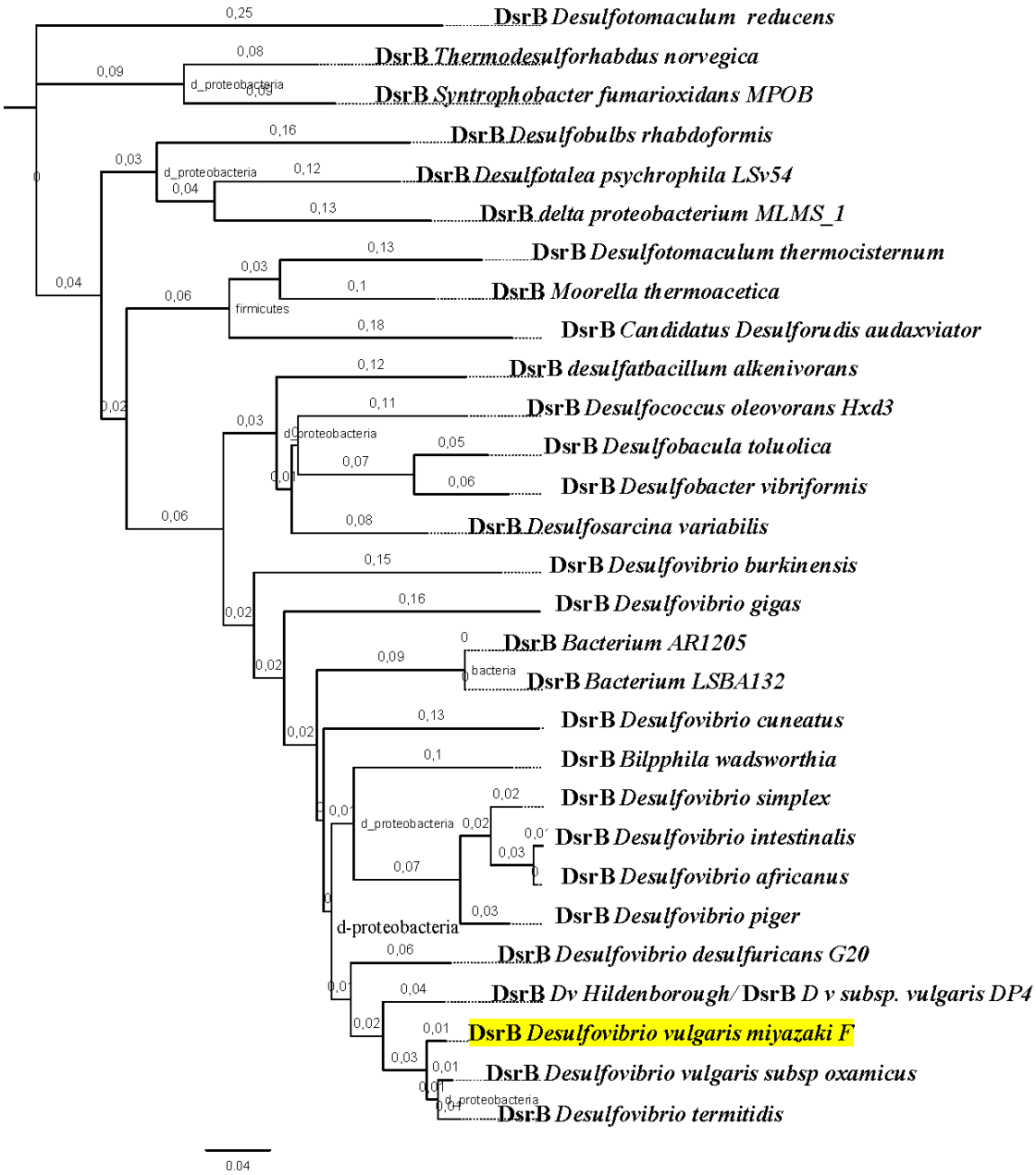


Fig 6.13 Phylogenetic tree for β-subunit, DsrB of desulfoviridin. The alignment with various sequences was carried out by using NCBI blast, and was used to draw the tree as before.

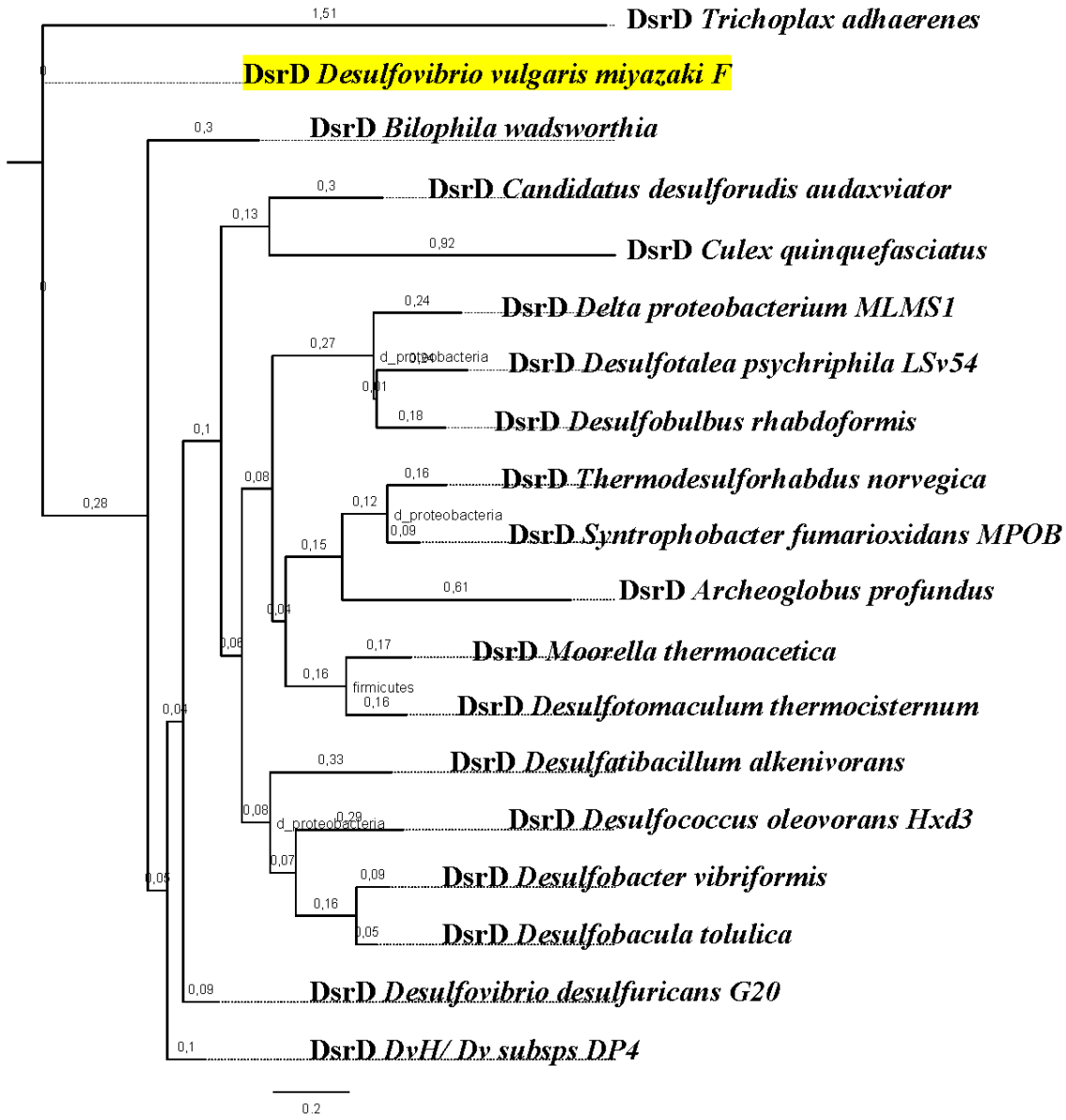


Fig 6.14 Phylogenetic tree for δ-subunit, DsrD of desulfoviridin. The alignment with various sequences was carried out by using NCBI blast, and was used to draw the tree as before. The abbreviations used are; DvH for *Desulfovibrio vulgaris* subsp Hildenborough and Dv for *Desulfovibrio vulgaris*

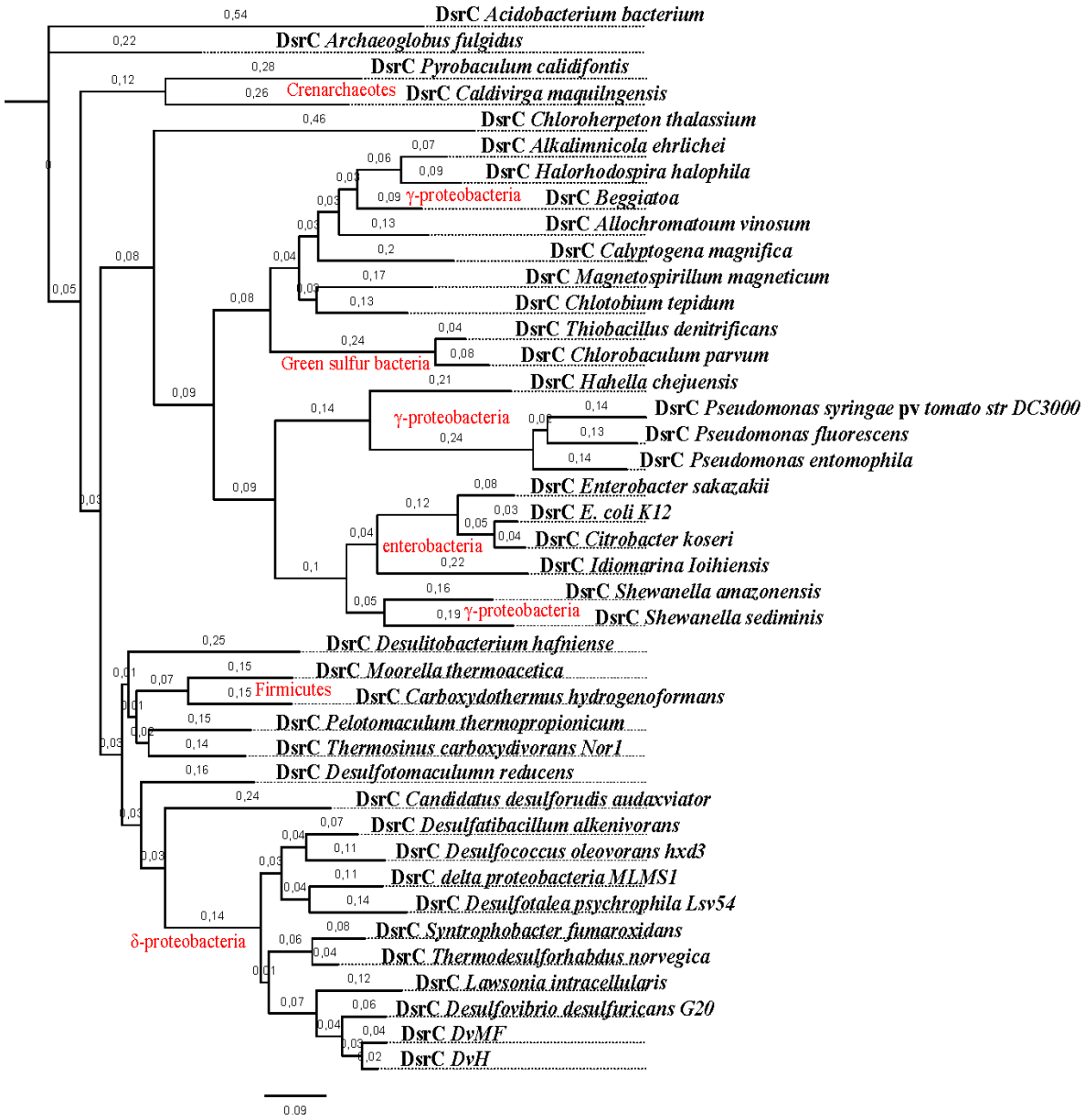


Fig 6.15 Phylogenetic tree for γ-subunit, DsrC of desulfoviridin. The alignment with various sequences was carried out by using NCBI blast, and was used to draw the tree as before. The abbreviations used are; DvH for *Desulfovibrio vulgaris* subsp Hildenborough and DvMF for *Desulfovibrio vulgaris* subsp Miyazaki F.

6.5 Discussion

All the known hydrogenases have been assigned to three distinct phylogenetic classes, the [NiFe]-hydrogenases, the [FeFe]-hydrogenases and the iron-sulfur free hydrogenase i.e. [Fe]-hydrogenases, based on the sequence alignments (Vignais, 2001). The [NiFeSe] hydrogenases are classified here also into the [NiFe]-group. Most of the hydrogenases have been found and characterized from the Archaeal and Bacterial domains and few from the Eukarya (Vignais, 2001, Vignais and Billoud, 2007). In addition to the hydrogenase genes, the additional genetic machinery involved in biosynthesis, maturation and processing of these complex molecules has also been characterized biochemically and phylogenetically for many species (Casalot, 2001, Böck et al, 2006). Microbial genome sequential data have further primed the understanding of the hydrogenase distribution and evolution.

The origin of hydrogenases may date back to the evolution of membrane complexes and compartmentalization of the cellular organelles, for associating proton translocation with the generation and conservation of energy (Vignais, 2008). Similarities between the sequences of various types of hydrogenases and the energy converting NADH-ubiquinone oxidoreductases of mitochondria and bacteria have been figured out in many studies (Friedrich and Scheide, 2000, Albracht and Hedderich, 2000). A further role for hydrogenases has been proposed for the origin of eukaryotic cells in two new hypotheses, which are the hydrogen hypothesis (Martin and Müller, 1998) and the syntrophic hypothesis (Moreira and López-García, 1998). The hydrogen hypothesis proposes an endo-symbiotic relationship of a heterotrophic alpha-proteobacterium that produces CO_2 and H_2 as waste products into an autotrophic Archaebacterium, leading to the division into a prokaryotic and eukaryotic cell with loss of anaerobic environment. The syntrophic hypothesis is based on a similar idea involving the metabolic participation between a delta proteobacterium and an archaebacterium showing syntrophy, i.e., one organism survives exclusively on the byproducts of the other rather than encountering an endosymbiotic relationship in which one organism is taken inside of the other and both are contributing to the survival.

In addition to the standard [NiFe] hydrogenase of DvMF, a [NiFeSe] hydrogenase and a multimeric six subunit containing Ech hydrogenase that have been found in this work (Chapter 3) are analyzed in terms of their phylogenetic proximity to other domains and divisions of bacteria. Though much less evenly distributed, the [NiFeSe] hydrogenase belongs to the same phylogenetic class as the standard [NiFe] hydrogenases. Interestingly, the [NiFeSe] hydrogenase of DvMF is closely related to the [NiFeSe] hydrogenases of other *Desulfovibrio* sps and to the [NiFe] hydrogenases of some other bacteria rather than to the [NiFe] hydrogenase present in the same cell. The overall topologies of the phylogenetic trees of these two hydrogenase forms show great similarity to the topology of the 16S rRNA tree of evolution. Though, the separation of the *Desulfovibrio* sps from other proteobacterial subdivisions, and also within the δ-proteobacteria indicates the horizontal transfer of the hydrogenase machinery into the ancestors of these species much after the genesis of these subdivisions. In agreement with this argument, the phylogenetic tree topologies of the Hyp maturation proteins of the [NiFe] hydrogenase are in agreement with such a proposal.

Another aspect is the co-occurrence of the [NiFeSe] hydrogenase genes together with the [NiFe] hydrogenase. This proximity has been detected in all the *Desulfovibrio* species in which the genes for the [NiFeSe] have been sequenced (Heidelberg et al, 2004, and this work). A primary structure alignment between the [NiFe] and the [NiFeSe] hydrogenases has been carried out (Chapter 5) and it depicts high similarities both between the small (38% identities and 55% similarities) and large subunits (37% identities and 52% similarities) of these two hydrogenases present in same cell. Furthermore, a significant conservation among the functional and structural domains has been also detected (Chapter 2 & 5). Considering this high preservation of the structure and function of these proteins and additionally the proximity of the coding genes for these hydrogenases in the genome, one may suggest that this is the result of a gene duplication. As the existence of the [NiFeSe] hydrogenase genes has been detected mainly in few of the known *Desulfovibrio* species, and in even lesser number of organisms also belonging to other Desulfovibrionales order, this gene duplication could be a much more recent event. However, the lack of a sufficient amount of sequences of the [NiFeSe] hydrogenases prohibits a clear association with any lateral gene transfer to or from, before or after the speciation in the *Desulfovibrio* group.

The Ech hydrogenase has been initially isolated from some of the Archaeal species, and later similar genes have been identified in several *Desulfovibrio* species. These hydrogenases are also placed in the phylogenetic class including the standard [NiFe] hydrogenase. The sequence and the phylogeny of the Ech hydrogenase of DvMF and those identified in other classes show that this multi-subunit complex is closely related to the multi subunit NADH ubiquinone oxidoreductase of *T. tengcongenesis* (Calteau et al, 2005). The high similarity between these two proteins from distant genomic clades is explained by one or more horizontal gene transfer processes in which such genes from archaea had been transferred to the ancestor of this bacterial lineage. Following, these genes are transferred from this ancestor to the ancestor of other bacteria.

The catalytic subunits of the Ech hydrogenase share slightly more sequence similarity with the [NiFeSe] hydrogenase rather with the [NiFe] hydrogenase. Again, in the phylograms of the catalytic EchE subunit of DvMF, this protein along with the EchE of other *Desulfovibrio* sps is closely grouped with the EchE subunit or the NADH ubiquinone oxidoreductase of some of the bacteria belonging to the Firmicutes division and the Euryarchaeotes sub-domain. The phylogram of the EchB and EchF show congruence with this observation.

Dissimilatory sulfate reduction is found in only three primary bacterial lineages and is restricted to a single archaeal genus, *Archaeoglobus*. The phylograms based on the subunit polypeptide sequences of the APS reductase and the desulfoviridin functional proteins do show topologies similar to each other. Furthermore, in these phylograms the topologies of relatedness among the *Desulfovibrio* and other δ-proteobacteria remain parallel to that derived by the 16s rRNA, but the topologies of relatedness to other domains are different. The APS reductase in terms of its characteristic sequence, structure and other physicochemical properties has remained highly conserved in various organisms even when they are placed across widely different classes.

A phylogenetic relationship, derived by Hipp et al, 1997, between the polypeptide sequences of the Aps and Dsr of *Chromatium vinosum*, *A. fulgidus* and *D. vulgaris*, has concluded that these genes have not developed convergently but are true homologues. The tree topologies shown by the comparison between the Aps reductase and Dsr proteins were the same, but were not

conformant with the topology of the 16S-rRNA. These topologies were better explained by considering these genes as paralogous instead of orthologous, which means that these genes were present before the domain separation and later have followed a similar pattern of evolution. By this way, the homologies between these proteins are a result of functional conservation and not of a lateral gene transfer among these prokaryotes.

In another study presented by Wagner et al, 1998, a single pair of degenerate primers complimentary to some highly conserved regions has been used, and has led to amplification of a 1.9 kb DNA region encoding most of the *dsrA* and *dsrB* genes from the genomic DNA of a range of bacteria. The entire DNA sequences obtained thereby exhibited high similarities with each other (49 to 89% identity), and the inferred evolutionary relationships between them were nearly identical to those inferred on the basis of 16S rRNA. This finding led to the conclusion that the high similarity of bacterial and archaeal DSRs is due to their common origin from a conserved DSR. This ancestral DSR should already been present before the division between the domains *Bacteria*, *Archaea*, and *Eukarya* or might be laterally transferred between *Bacteria* and *Archaea* soon after the domain divergence. Thus the ancestors of *Bacteria* and *Archaea* already possessed a key enzyme of sulfate and sulfite respiration and this supports the early origin of sulfate respiration and the physiological role of the DSR remained constant over time.

Chapter 7

Synopsis and Conclusions

Increasing numbers of hydrogenase and related gene sequences are being reported with the advancements of genomic sequencing. At the same time the diversities and significances of these metallo-enzymes, which catalyze the reversible oxidation of hydrogen, are examined by means of biochemical, growth and micro-array investigations. The study of these proteins is also important due to the potential application of hydrogenases for biohydrogen production and Biomimetic Chemistry. This thesis describes the diversities and characterizations of hydrogenases and their related maturation pathway proteins as well as some of the sulfate metabolism proteins and their corresponding genes belonging to *Desulfovibrio vulgari*s Miyazaki F (DvMF).

A general introduction to the DvMF, which belongs to the sulfate reducing δ-proteobacteria and the importance of studying this group of bacteria, has been given in **chapter 1**. In **chapter 2**, the hydrogenases are introduced in a greater detail regarding their structures, cellular functions, biosynthesis, in addition to the role of the associated Hyp maturation proteins. Furthermore, some examples for heterologous expression of [NiFe] hydrogenases, as in case of *Desulfovibrio gigas* and *Ralstonia eutropha*, were discussed briefly. The basic structures and functions of two sulfate metabolism enzymes, which were purified and crystallized in collaborative efforts in this work, 5'-adenylyl sulfate reductase and desulfoviridin, were also discussed. **Chapter 3** outlined the molecular, biochemical and computational methods used for the investigation. The major results are summarized in the following.

The [NiFe] hydrogenase was the only known hydrogenase from DvMF before this investigation was started. A cosmid genomic library was generated to probe for the presence of additional hydrogenases and related genes by means of southern hybridization **(Chapter 4)**. After identification, the desired genes were sequenced employing a degenerate PCR strategy based on conserved domain homologies between the gene and protein sequences from closely related organisms. A total of seven stretches of genomic DNA of DvMF, cloned in distinct cosmid

clones were sequenced. Two of these fragments corresponded to two operons, ***hysBA*** and ***echABCDEF*** coding for two additional hydrogenases, a [NiFeSe] hydrogenase and a six-subunit Ech hydrogenase, respectively. The *hysBA* operon was present just upstream of the [NiFe] coding hydrogenase operon *HynBACD*, separated by a single maturation gene *hupG*. The five Hyp maturation proteins, HypA, HypB, HypD, HypE and HypF were found to be coded from three different operons those are ***hypAB***, ***hypDE*** and ***hypF***. The two subunits of the flavin associated protein 5'-adenylyl sulfate reductase (AprAB) are coded from a single operon i.e. ***aprAB***, while the desulfiviridin (DsrABC) is coded from two separate locations in the genome. The α- (DsrA coded by *dsvA*) and β- (DsrB coded by *dsvB*) subunits-coding gene sequences of desulfoviridin were present together with the δ-subunit (DsrD coded by *dsvD*) coding sequences in one operon i.e. ***dsrABD.*** The γ-subunit (DsrC coded by *dsrC*) was present elsewhere denoted as ***dsrC*** gene and was partly sequenced from the genome by degenerate PCR. Trying simultaneously on similar lines, the presence of an [FeFe] hydrogenase was ruled out.

The sequenced gene products were characterized structurally and biochemically in **chapter 5**. The [NiFeSe] hydrogenase and the [NiFe] hydrogenase of DvMF have significant similarities with each other in their three dimensional architecture. Both of these hydrogenases are hetero-dimeric and possess three co-linear [FeS] clusters in the small subunit, whereas the large subunit holds the catalytically active bimetallic Ni-Fe center. The [NiFe] hydrogenase has two peripheral [Fe_4S_4] clusters and one [Fe_3S_4] medial cluster. The [NiFeSe] hydrogenase also has two peripheral [Fe_4S_4] clusters, but the nature of the medial cluster could not be ascertained by sequence homologies due to the displacement of the predictable cysteine (Cys-279) to two positions earlier than the usually conserved place. However a serine (Ser- 281) is present at the usual location and may be participating in the iron-sulfur cluster.

Further, the expression of [NiFeSe] hydrogenase was tested by means of western blotting and native-PAGE hydrogenase activity assays, also the co-regulation of the [NiFe] hydrogenase and the [NiFeSe] enzyme by the availability of metals like nickel and selenium was demonstrated. Where, the 1 μM nickel supplement supported both types of enzymes, presence of 1 μM selenium down-regulated the [NiFe] hydrogenase significantly. The structural genes for the [NiFeSe] hydrogenase were expressed in *E. coli* and the possibility of integration of a

selenocysteine into the transcript of the large subunit (HysA) was indicated by means of detecting an epitope (S-tag) present behind the selenocysteine.

The thee-dimensional structures of the Hyp maturation proteins were modelled based on the templates of the crystallized functional homologous proteins and the conservation was affirmatively compared. The two maturation two-gene operons *hypAB* and *hypDE* were separately expressed in *E. coli* by means of inducible T7-promoter carrying constructs. Simultaneous expression of *hypA* and *hypB* from a single promoter in pRSF2-hypAB as well as of *hypD* and *hypE* from another promoter in pRSF2-hypDE were detected. This observation holds importance in terms of minimizing the logistics needed for the co-expression of the maturation and structural genes of hydrogenase in *E. coli*.

Taking the advantage of elucidated maturation pathways in *E. coli* as well as the genetic advancements made by us, the efforts to advance the heterologous expression of [NiFe] hydrogenase were made by tagging the small subunit *hynB* by a strep-coding sequence in its C-terminus by means of PCR, while keeping the operon *HynBACD* intact. This strep-tagged operon was expressed by means of a single T7-promoter in *E. coli*. Though not functionally active, the structural proteins were successfully co-purified using this construct, thus opening the way for future purification when the corresponding maturation machinery would be optimized.

The 5'-adenylyl sulfate reductase and desulfoviridin proteins of DvMF were purified and crystallized in collaboration with Dr. Aruna Goenka Agrawal and Dr. Hideaki Ogata. These proteins were identified by means of high-energy MALDI-TOF MS de-fragmentation and N-terminal sequencing. This data was previously used to sequence these genes. A resolution of 1.7 Å for the $\alpha_2\beta_2$ structure of 5'-adenylyl sulfate reductase has been achieved and work is ongoing to resolve a $\alpha_2\beta_2\gamma_2$ crystal structure of desulfoviridin.

Lastly, the phylogenies of the three hydrogenases of DvMF and the Hyp maturation proteins along with 5'-adenylyl sulfate reductase and desulfoviridin of DvMF were studied by amino acids similarities to their homologs (**Chapter 6**). The [NiFe] and [NiFeSe] hydrogenases of DvMF and other *Desulfovibrio* form two separate outgroups, and the [NiFeSe] hydrogenases

appeared closely related to [NiFe] hydrogenases of non-*Desulfovibrio* groups of bacteria. Despite high homologies, a genesis of less abundant [NiFeSe] hydrogenases from the rather broadly distributed [NiFe] hydrogenase, as a result of an event of gene duplication could not be confirmed owing to lack of [NiFeSe] sequences in sufficient numbers. Nevertheless, the overall topologies of the phylograms derived for the [NiFe] hydrogenase and all of the Hyp maturation proteins of DvMF conformed to a similar relationship between the 16s rRNA sequences between various groups. Moreover, the horizontal transfer of the hydrogenase machinery was evident due to segregated conservation in between the δ-proteobacteria. The phylograms of various subunits of Ech hydrogenase (EchA, EchE and EchF) showed close relation to the multi subunit NADH ubiquinone oxidoreductase of *T. tengcongenesis*, as has been described for other *Desulfovibrio* sps as well.

The topologies of phylograms for 5'-adenylyl sulfate reductase and desulfoviridin of DvMF were quite similar to each other, but different from the 16s rRNA tree of evolution, a phenomenon which is in literature attributed to their paralogous relationship.

Conclusions

Additional hydrogenases other than the well characterized [NiFe] hydrogenase are present in DvMF. Their detection has affirmed the hydrogen cycling principle of *Desulfovibrio* metabolism. This information can form a basis to generate various hydrogenase deletion mutants and compare growth kinetics by means of molecular and biochemical characterizations. As the role of selenium in the regulation of hydrogenase expression has been figured out, additional factors can be identified to control the hydrogen metabolism of the cell in a desirable manner.

The [NiFeSe] hydrogenase of DvMF has been reported for the first time, a viable protocol for the purification of this hydrogenase needs to be procured for subsequent structural, biochemical and spectroscopic characterizations. This may be also helpful to ascertain the types of bonding present in the iron sulfur clusters of the small subunit. The integration of the selenocysteine, when expressed in *E. coli*, is another interesting topic to study.

The [NiFe] hydrogenase of DvMF, when expressed from a plasmid can be now isolated from non-native hosts by means of a strep-tag, in a highly pure form. This could be very helpful while employing some of the closely related heterologous host strains and even when expressed homologously, especially for carrying mutagenesis studies. The experiments to try new antibiotic markers in *Desulfovibrio* sps are on the way and also the logistics for co-expressing the structural together with the maturation genes sequenced for DvMF in this work, using *E. coli* as host appear to be feasible. As the [NiFe] - and [NiFeSe] hydrogenases are quite comparable structurally and biochemically, similar trials can be extended to the [NiFeSe] hydrogenase of DvMF that are epitope tagging and co-expression with the maturation machinery.

Chapter 8

Bibliography

Agrawal A. G., Voordouw G., and Gärtner W. (2006) Sequential and structural analysis of [NiFe]-hydrogenase-maturation proteins from *Desulfovibrio vulgaris Miyazaki F*. *Antonie Van Leeuwenhoek*, **90(3)**, 281-290.

Albracht S. P., and Hedderich R. (2000) Learning from hydrogenases: location of a proton pump and of a second FMN in bovine NADH-ubiquinone oxidoreductase (Complex I). *FEBS Lett*., **485(1)**, 1-6.

Albracht, S. P. (1994) Nickel hydrogenases: in search of the active site. *Biochim. Biophys. Acta.*, **1188**, 167-204.

Alexeyev M. F., and Shokolenko I. N. (1995) Mini-Tn10 transposon derivatives for insertion mutagenesis and gene delivery into the chromosome of gram-negative bacteria. *Gene*, **160(1),** 59-62.

Atanassova A., and Zamble D. B. (2005) *Escherichia coli* HypA is a zinc metalloprotein with a weak affinity for nickel. *J. Bacteriol.*, **187(14)**, 4689-4697.

Bender, K. S., Yen, H. C., and Wall. J. D. (2006) Analysing the metabolic capabilities of *Desulfovibrio* species through genetic manipulation. *Biotechnol. Genet. Eng. Rev.,* **23**, 157–174.

Blokesch M., Paschos A., Theodoratou E., Bauer A., Hube M., Huth S., and Böck A. (2002) Metal insertion into NiFe-hydrogenases. *Biochem Soc Trans.*, **30(4)**, 674-680.

Blokesch, M. (2004) [NiFe]-Hydrogenasen von *Escherichia coli*: Funktionen der am Metalleinbau beteiligten Proteine. PhD thesis, Ludwig-Maximilians-Universität, München.

Blokesch, M., Albracht, S. P., Matzanke, B. F., Drapal, N. M., Jacobi, A., and Böck, A. (2004) 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.*, **344(1)**, 155-167.

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

Blokesch, M., Magalon, A., and Böck, A. (2001) Interplay between the specific chaperone-like proteins HybG and HypC in maturation of hydrogenases 1, 2, and 3 from *Escherichia coli*. *J. Bacteriol.*, **183(9)**, 2817-2822.

Blokesch, M., Paschos, A., Bauer, A., Reissmann, S., Drapal, N., and Böck, A. (2004) Analysis of the transcarbamoylation-dehydration reaction catalyzed by the hydrogenase maturation proteins HypF and HypE. *Eur. J. Biochem.*, **271**, 3428-3436.

Böck A., King P. W., Blokesch M., and Posewitz, M. C. (2006) Maturation of hydrogenases. *Adv. Microb. Physiol.*, **51**, 1-71.

Böck, A., Forchhammer, K., Heider, J., Leinfelder, W., Sawers, G., Veprek B., and Zinoni, F. (1991) Selenocysteine: the 21st amino acid. *Mol. Microbiol.*, **5**, 515-520.

Caffrey S. M., Park H. S., Voordouw J. K., He Z., Zhou J., and Voordouw, G. (2007) Function of periplasmic hydrogenases in the sulfate reducing bacterium *Desulfovibrio vulgaris* Hildenborough. *J. Bacteriol.*, **189**, 6159–6167.

Calteau A., Gouy M., and Perrière, G. (2005) Horizontal transfer of two operons coding for hydrogenases between bacteria and archaea. *J. Mol. Evol.*, **60(5)**, 557-565.

Cammack, R., Frey, M., and Robson R. (2001) Hydrogen as a Fuel: Learning from nature. Taylor and Francis Press, London and New York.

Casalot L., De Luca G., Dermoun Z., Rousset M., and de Philip, P. (2002) Evidence for a fourth hydrogenase in *Desulfovibrio fructosovorans*. *J. Bacteriol.*, **184(3)**, 853-856.

Casalot, L., and Rousset, M. (2001) Maturation of the [NiFe] hydrogenases. *Trends Microbiol.*, **9**, 228-237.

Deckers, H. M., Wilson, F. R., and Voordouw, G. (1990) Cloning and sequencing of a [NiFe] hydrogenase operon from *Desulfovibrio vulgaris* Miyazaki F. *J. Gen. Microbiol.* **136(10)**, 2021-2028.

Dementin S., Burlat B., De Lacey A. L., Pardo A, Adryanczyk-Perrier G., Guigliarelli B., Fernandez V. M., and Rousset M. (2004) A glutamate is the essential proton transfer gate during the catalytic cycle of the [NiFe] hydrogenase. *J. Biol. Chem.*, **279(11)**, 10508-10513.

Dinh H. T., Kuever J., Mussmann M., Hassel A. W., Stratmann M., and Widdel F. (2004) Iron corrosion by novel anaerobic microorganisms. *Nature*, **427**, 829–832.

Drapal, N., and Böck, A. (1998) Interaction of the hydrogenase accessory protein HypC with HycE, the large subunit of *Escherichia coli* hydrogenase 3 during enzyme maturation. *Biochemistry*, **37(9)**, 2941-2948.

Fichtner C., Laurich C., Bothe E., Lubitz W. (2006) Spectroelectrochemical characterization of the [NiFe] hydrogenase of *Desulfovibrio vulgaris* Miyazaki F. *Biochemistry*, **45(32)**, 9706-9716.

Fontana, P., Bindewald, E., Toppo, S., Velasco, R., Valle, G., and Tosatto, S. C. E. (2005) The SSEA server for protein secondary structure alignment. *Bioinformatics*, **21**, 393-395.

Friedrich, T., and Scheide, D. (2000) The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Lett.*, **479(1-2)**, 1-5.

Fritsche, E., Paschos, A., Beisel, H. G., Bock, A., and Huber, R. (1999) Crystal structure of the hydrogenase maturating endopeptidase HYBD from *Escherichia coli*. *J. Mol. Biol.*, **288(5)**, 989-998.

Fritz G., Roth A., Schiffer A., Büchert T., Bourenkov G., Bartunik H. D., Huber H., Stetter K. O., Kroneck P. M., and Ermler, U. (2002) Structure of adenylylsulfate reductase from the hyperthermophilic *Archaeoglobus fulgidus* at 1.6 A° resolution. *Proc. Natl. Acad. Sci.*, **99(4)**, 1836-1841.

Garcin, E., Vernede, X., Hatchikian, E. C., Volbeda, A., Frey, M., and Fontecilla-Camps, J. C. (1999) The crystal structure of a reduced [NiFeSe] hydrogenase provides an image of the activated catalytic center. *Structure*, **7**, 557-566.

Gibson, J. S., and Harwood, C. (2002) Metabolic diversity in aromatic compound utilization by anaerobic microbes. *Annu Rev Microbiol.* **56,** 345-369.

Goenka, A., Voordouw, J. K., Lubitz, W., Gärtner, W., and Voordouw, G. (2005) Construction of a [NiFe]-hydrogenase deletion mutant of *Desulfovibrio vulgaris* Hildenborough. *Biochem. Soc. Trans.*, **33**, 59–60.

Guex, N., and Peitsch, M. C. (1997) SWISS-Model and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*, **18**, 2714-2723.

Hamilton, W. A. (1998) Bioenergetics of sulphate-reducing bacteria in relation to their environmental impact. *Biodegradation*, **9**, 201–212.

Hamilton, W. A. (2003) Microbially influenced corrosion as a model system for the study of metal microbe interactions: a unifying electron transfer hypothesis. *Biofouling*, **19(1)**, 65-76.

Hanahan, D. (1985) Techniques for transformation of *E. coli*, p.109-136. In DNA cloning vol. I. (ed. D.M. Glover), IRL Press, Washington, D.C.

Hansen, T. A. (1994) Metabolism of sulfate-reducing prokaryotes. *Antonie Van Leeuwenhoek.* **66(1-3)**, 165-185.

Hayashi S. and Wu H. C. (1990) Lipoproteins in bacteria. *J. Bioenerg. Biomembr.*, **22**, 451-471.

Hedderich, R. and Forzi, L. (2005) Energy converting [NiFe] hydrogenases: more than just H_2 activation. *J. Mol. Microbiol. Biotechnol.*, **10**, 92–104.

Heidelberg J. F., Seshadri R., Haveman S. A., Hemme C. L., Paulsen I. T., Kolonay J. F., Eisen J. A., Ward N., Methe B., Brinkac L. M., Daugherty S. C., Deboy R. T., Dodson R. J., Durkin A. S., Madupu R., Nelson W. C., Sullivan S. A., Fouts D., Haft D. H., Selengut J., Peterson J. D., Davidsen T. M., Zafar N., Zhou L. W., Radune D., Dimitrov G., Hance M., Tran K., Khouri H., Gill J., Utterback T. R., Feldblyum T. V., Wall J. D., Voordouw G., and Fraser, C. M. (2004) The genome sequence of the anaerobic, sulfate-reducing bacterium *Desulfovibrio vulgaris* Hildenborough. *Nat. Biotechnol*, **22**, 554–559.

Higuchi, Y., Ogata, H., Miki, K., Yasuoka, N., and Yagi, T. (1999) Removal of the bridging ligand atom at the Ni-Fe active site of [NiFe] hydrogenase upon reduction with H_2, as revealed by X-ray structure analysis at 1.4° A resolution. *Structure*, **7(5)**, 549-556.

Higuchi, Y., Yagi, T., and Yasuoka, N. (1997) Unusual ligand structure in Ni-Fe active center and an additional Mg site in hydrogenase revealed by high resolution X-ray structure analysis. *Structure*, **5**, 1671-1680.

Hipp, W. M., Pott, A. S., Thum-Schmitz, N., Faath, I., Dahl, C., and Trüper, H. G. (1997) Towards the phylogeny of APS reductases and sirohaem sulfite reductases in sulfate-reducing and sulfur-oxidizing prokaryotes. *Microbiology*, **143(9)**, 2891-2902.

Hube, M., Blokesch, M., and Böck, A. (2002) Network of hydrogenase maturation in *Escherichia coli*: role of accessory proteins HypA and HybF. *J. Bacteriol.* **184**, 3879-3885.

Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.*, **292**, 195-202.

Kunkel J., Vorholt J., Thauer R., and Hedderich, R. (1998) An *Escherichia coli* hydrogenase-3-type hydrogenase in methanogenic archaea. *Eur. J. Biochem*. **252**, 467–476.

Lampreia, J., Moura, I., Teixeira, M., Peck, H. D., Jr, Legall, J., Huynh, B. H., and Moura, J. J. (1990) The active centers of adenylylsulfate reductase from *Desulfovibrio gigas*. Characterization and spectroscopic studies. *Eur. J. Biochem*., **188(3)**, 653-664.

Laue, H., Friedrich M., Ruff J., and Cook, A. M. (2001) Dissimilatory sulfite reductase (desulfoviridin) of the taurine-degrading, non-sulfate-reducing bacterium *Bilophila wadsworthia* RZATAU contains a fused DsrB-DsrD subunit. *J. Bacteriol.* **183(5)**, 1727-1733.

LeGall, J., and Fauque, G. (1988) Dissimilatory reduction of sulfur compounds. In: Zehnder A.J.B. (ed) Biology of anaerobic microorganisms, chapter 11. John Wiley and Sons, New York London.

Lenz, O., Gleiche, A., Strack, A., and Friedrich, B. (2005) Requirements for heterologous production of a complex metalloenzyme: the membrane-bound [NiFe] hydrogenase. *J Bacteriol.* **187(18)**, 6590-6595.

Lloyd, J. R. (2003) Microbial reduction of metals and radionuclides. *FEMS Microbiol. Rev.* **27**, 411–425

Loubinoux, J., Bisson-Boutelliez, C., Miller, N., and Le Faou, A. E. (2002) Isolation of the provisionally named *Desulfovibrio fairfieldensis* from human periodontal pockets. *Oral Microbiol. Immunol.*, **17(5)**, 321-323.

Lovley, D. R. (2003) Cleaning up with genomics: applying molecular biology to bioremediation. *Nat. Rev. Microbiol.*, **1(1)**, 35-44.

Lubitz W., Reijerse E., van Gastel M. (2007) [NiFe] and [FeFe] hydrogenases studied by advanced magnetic resonance techniques. *Chem. Rev.*, **107(10)**, 4331-65.

Ma, K., Schicho, R. N., Kelly, R. M., and Adams, M. W. W. (1993) Hydrogenase of the hyperthermophile *Pyrococcus furiosus* is an elemental sulfur reductase or sulfhydrogenase: Evidence for a sulfur-reducing hydrogenase ancestor. *Proc. Natl. Acad. Sci.***, 90**, 5341-5344.

Maier, T., and Böck, A. (1996) Generation of active [NiFe] hydrogenase *in vitro* from a nickel-free precursor form. *Biochemistry*, **35(31)**, 10089-10093.

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

Matias, P. M., Coelho, A. V., Valente, F. M., Plácido, D., LeGall, J., Xavier, A. V., Pereira, I. A., and Carrondo, M. A. (2002) Sulfate respiration in *Desulfovibrio vulgaris* Hildenborough. Structure of the 16-heme cytochrome c HmcA at 2.5-Å resolution and a view of its role in transmembrane electron transfer. *J. Biol. Chem.*, **277(49)**, 47907-47916.

Matias, P. M., Soares, C. M., Saraiva, L. M., Coelho, R., Morais J., Le Gall, J., and Carrondo, M. A. (2001) [NiFe] hydrogenase from *Desulfovibrio desulfuricans* ATCC 27774 gene sequencing, three-dimensional structure determination and refinement at 1.8 Å

and modelling studies of its interaction with the tetrahaem cytochrome c3. *J Biol Inorg Chem.*, **6(1)**, 63-81.

McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics*, **16**, 404-405.

Meuer, J., Bartoschek, J., Koch, A., Kunkel, R., and Hedderich, R. (1999) Purification and catalytic properties of Ech hydrogenase from *Methanosarcina barkeri*. *Eur. J. Biochem.* **265**, 325–335.

Meuer, J., Kuettner, H. C., Zhang, J. K., Hedderich, R., and Metcalf, W. W. (2002) Genetic analysis of the archaeon *Methanosarcina barkeri* Fusaro reveals a central role for Ech hydrogenase and ferredoxin in methanogenesis and carbon fixation. *Proc. Natl. Acad. Sci.*, **99**, 5632–5637.

Moreira, D., and Lopez-Garcia, P. (1998) Symbiosis between methanogenic archaea and delta-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. *J. Mol. Evol.* **47(5)**, 517-530.

Odom, J. M., and Peck, H. D. (1981) Hydrogen cycling as a general mechanism for energy coupling in the sulphate-reducing bacterium *Desulfovibrio* sp. *FEMS Microbiol. Lett*., **12**, 47–50.

Odom, J. M., and Peck, J. H. D. (1984) Hydrogenase, electron-transfer proteins, and energy coupling in the sulfate-reducing bacteria *Desulfovibrio*. *Annu. Rev. Microbiol.* **38**, 551-592.

Odom, J. M., and Peck, J. H. D. (1981) Hydrogen cycling as a general mechanism for energy coupling in the sulfate-reducing bacteria, *Desulfovibrio* sp. *FEMS Microbiology Letters*, **12**, 47-50.

Ogata, H., Goenka Agrawal, A., Kaur, A. P., Goddard, R., Gärtner, W., and Lubitz, W. (2008) Purification, crystallization and preliminary X-ray analysis of adenylylsulfate reductase from *Desulfovibrio vulgaris* Miyazaki F. *Acta. Cryst.*, **64(11)**, 1010-1012.

Oliveira, T. F., Vonrhein, C., Matias, P. M., Venceslau, S. S., Pereira, I. A., and Archer, M. (2008) Purification, crystallization and preliminary crystallographic analysis of a dissimilatory DsrAB sulfite reductase in complex with DsrC. *J. Struct. Biol.*, **164(2),** 236-239.

Oliveira, T. F., Vonrhein, C., Matias, P. M., Venceslau, S. S., Pereira, I. A., and Archer, M. (2008) The Crystal Structure of *Desulfovibrio vulgaris* Dissimilatory Sulfite Reductase Bound to DsrC Provides Novel Insights into the Mechanism of Sulfate Respiration. *J. Biol. Chem.* **283(49)**, 34141-34149.

Olson, J. W., Fu, C., and Maier, R. J. (1997) The HypB protein from *Bradyrhizobium japonicum* can store nickel and is required for the nickel-dependent transcriptional regulation of hydrogenase. *Mol. Microbiol.* **24**, 119-128.

Pedroni, P., Della Volpe, A., Galli, G., Mura, G. M., Pratesi, C., and Grandi G. (1995) Characterization of the locus encoding the [Ni-Fe] sulfhydrogenase from the archaeon *Pyrococcus furiosus*: evidence for a relationship to bacterial sulfite reductases. *Microbiology* **141**, 449-458.

Pereira, I. A. C., Haveman, S. A., and Voordouw, G. (2007) Biochemical, genetic and genomic characterization of anaerobic electron transport pathways in sulphate-reducing delta-proteobacteria. In: Bolton LL, Hamilton WA (eds) Sulphate-reducing bacteria: environmental and engineered systems. Cambridge University Press, Cambridge, UK.

Pierik, A. J., Wolbert, R. B., Mutsaers, P. H., Hagen, W. R., and Veeger, C. (1992) Purification and biochemical characterization of a putative [6Fe-6S] prismane-cluster-

containing protein from *Desulfovibrio vulgaris* (Hildenborough). *Eur. J. Biochem.* **206(3)**, 697-704.

Pitcher, M. C. and Cummings, J. H. (1996) Hydrogen sulphide: a bacterial toxin in ulcerative colitis? *Gut*. **39(1)**, 1-4.

Postgate, J.R. (1984) The sulphate-reducing bacteria. Cambridge University Press, Cambridge, United Kingdom.

Rangarajan, E. S., Asinas, A., Proteau, A., Munger, C., Baardsnes, J., Iannuzzi, P., Matte, A., and Cygler, M. (2008) Structure of [NiFe] hydrogenase maturation protein HypE from *Escherichia coli* and its interaction with HypF. *J. Bacteriol*., **190(4)**, 1447-1458.

Reissmann, S., Hochleitner, E., Wang, H., Paschos, A., Lottspeich, F., Glass, R. S., and Böck, A. (2003) Taming of a poison: biosynthesis of the NiFe-hydrogenase cyanide ligands. *Science*, **299(5609)**, 1067-1070.

Rodrigue, A., Chanal, A., Beck, K., Muller, M., and Wu, L. F. (1999) Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. *J. Biol. Chem.*, **274**, 13223-13228.

Rodrigues, R., Valente, F. M. A., Pereira, I. A. C., Oliveira, S., and Rodrigues-Pousada, C. (2003) A novel membrane-bound Ech [NiFe] hydrogenase in *Desulfovibrio gigas*. *Biochem Biophys. Res. Commun.*., **306**, 366–375.

Rother, M., Mathes, I., Lottspeich, F., and Böck, A. (2003) Inactivation of the *selB* gene in *Methanococcus maripaludis*: effect on synthesis of selenoproteins and their sulfur-containing homologs. *J. Bacteriol.*, **185(1),** 107-114.

Rousset, M., Magro, V., Forget, N., Guigliarelli, B., Bélaich, J. P., and Hatchikian, C. (1998) Heterologous expression of *Desulfovibrio gigas* [NiFe] hydrogenase in *Desulfovibrio fructosovorans*. *J. Bacteriol.*, **180**, 4982-4986.

Rousset, M., Montet, Y., Guigliarelli, B., Forget, N., Asso, M., Bertrand, P., Fontecilla-Camps, J. C. and Hatchikian, E. C. (1998) [3Fe-4S] to [4Fe-4S] cluster conversion in *Desulfovibrio fructosovorans* [NiFe] hydrogenase by site-directed mutagenesis. *Proc. Natl. Acad. Sci. USA*, **95**, 11625-11630

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: a laboratory manual (2nd ed.), Cold Spring Harbour Laboratory Press, New York.

Shima, S. and Thauer, R. K. (2007) A third type of hydrogenase catalyzing H_2 activation. *Chem. Rec.*, **7(1)**, 37-46.

Shomura, Y., Komori, H., Miyabe, N., Tomiyama, M., Shibata, N., and Higuchi, Y. (2007) Crystal structures of hydrogenase maturation protein HypE in the Apo and ATP-bound forms. *J. Mol. Biol.*, **372(4)**, 1045-1054.

Simon, R., Priefer, U., and Pühler, A. (1983) A broad-host-range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram negative bacteria. *Bio Technol*, **1**, 784-791.

Skyring, G. W., and Donnelly, T. H. (1982) Precambrian sulfur isotopes and a possible role for sulfite in the evolution of biological sulfate reduction. *Precambrian Res*, **17**, 41–61

Tormay, P., and Böck, A. (1997) Barriers to heterologous expression of a selenoprotein gene in bacteria. *J. Bacteriol.*, **179(3)**, 576-582.

Valente, F. A. A., Almeida C. C., Pacheco I., Carita J., Saraiva L. M. and Pereira I. A. C. (2006) Selenium is involved in regulation of periplasmic hydrogenase gene expression in

Desulfovibrio vulgaris Hildenborough. *J. Bacteriol.*, **188**, 3228–3235.

Valente, F. M., Oliveira, A. S., Gnadt, N., Pacheco, I., Coelho, A. V., Xavier, A. V., Teixeira, M., Soares, C. M., and Pereira, I. A. (2005) Hydrogenases in *Desulfovibrio vulgaris* Hildenborough: structural and physiologic characterisation of the membrane-bound [NiFeSe] hydrogenase. *J. Biol. Inorg. Chem.*, **10**, 667–682.

Vignais, P. M. (2008) Hydrogenases and H($^{+}$)-reduction in primary energy conservation. *Results Probl Cell Differ.*, **45**, 223-252.

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

Vignais, P. M., Billoud B., and Meyer, J. (2001) Classification and phylogeny of hydrogenases. *FEMS Microbiol. Rev.*, **25**, 455-501.

Vincent, K. A., Cracknell, J. A., Parkin, A., and Armstrong, F. A. (2005) Hydrogen cycling by enzymes: electrocatalysis and implications for future energy technology. *Dalton. Trans.*, **21**, 3397-403.

Volbeda, A., Charon, M. H., Piras, C., Hatchikian, E. C., Frey, M., and Fontecilla-Camps, J. C. (1995) Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. *Nature,* **373**, 580-587.

Volbeda, A., Martin, L., Cavazza, C., Matho, M., Faber, B. W., Roseboom, W., Albracht, S. P., Garcin, E., Rousset, M., and Fontecilla-Camps, J. C. (2005) Structural differences between the ready and unready oxidized states of [NiFe] hydrogenases. *J Biol Inorg Chem.*, **10(3)**, 239-249.

Voordouw, G., Niviere, V., Ferris, F. G., Fedorak, P. M. and Westlake, D. W. (1990) Distribution of Hydrogenase Genes in *Desulfovibrio* spp. and their Use in Identification of Species from the Oil Field Environment. *Appl Environ Microbiol.*, **56(12)**, 3748-3754.

Wagner, M., Roger, A. J., Flax, J. L., Brusseau, G. A. and Stahl, D. A. (1998) Phylogeny of dissimilatory sulfite reductases supports an early origin of sulfate respiration. *J. Bacteriol.* **180(11),** 2975-2982.

Wall, J. D., and Krumholz, L. R. (2006) Uranium Reduction. *Ann Rev Microbiol* **60**, 149–166.

Watanabe, S., Matsumi, R., Arai, T., Atomi, H., Imanaka, T., and Miki, K. (2007) Crystal structures of [NiFe] hydrogenase maturation proteins HypC, HypD, and HypE: insights into cyanation reaction by thiol redox signaling. *Mol. Cell.*, **27(1)**, 29-40.

Wu, L. F., and Mandrand, M. A. (1993) Microbial hydrogenases: Primary structure, classification, signatures and phylogeny. *FEMS Microbiol. Rev.*, **104**, 243-269.

Wu, L. F., Chanal, A., and Rodrigue, A. (2000) Membrane targeting and translocation of bacterial hydrogenases. *Arch. Microbiol.*, **173**, 319-324.

Yagi, T., and Ogata, M. (1996) Catalytic properties of adenylylsulfate reductase from *Desulfovibrio vulgaris* Miyazaki F. *Biochimie*. **78(10)**, 838-846.

Zhang, J. W., Butland, G., Greenblatt, J. F., Emili, A., and Zamble, D. B. (2005) A role for SlyD in the *Escherichia coli* hydrogenase biosynthetic pathway. *J Biol Chem.*, **280(6)**, 4360-4366.

Chapter 9

Appendix

Sequences submitted

Proteins	**Accession number**	**Genes**	**Protein ID**
APS reductase	EU127913.1 GI:160863271	*aprA* *gi:160863272*	ABX51937.1
		aprB *gi:160863272*	ABX51938.1
DSV reductase	EU127914	*dsvA*	ABX51939
		dsvB	ABX51940
		dsvC	ABX51941
Ech hydrogenase	EU796885.1 GI:193479932	*echA* *gi:193479933*	ACF17986.1
		echB *gi:193479934"*	ACF17987.1
		echC *gi:193479935*	ACF17988.1
		echD *gi:193479936*	ACF17989.1
		echE *gi:193479937*	ACF17990.1
		echF *gi:193479938*	ACF17991.1
[NiFeSe] hydrogenase	EU327692 GI:165940632	*hysA* *gi:165940633*	ABY75251.1
		hysB *gi:165940633*	ABY75252.1

Proteins	Accession number	Genes	Protein ID
HypAB	EU127916 GI:160863281	*hypA* *gi:160863282*	ABX51944.1
		hypB *gi:163676556*	ABY40453.1
HypD and HypE	EU127915 GI:160863278	*hypD* *gi:160863279*	ABX51942.1
		hypE *gi:160863280*	ABX51943.1
HypF	EU327693 GI:165940635	*hypF* *gi:165940636*	ABY75253.1

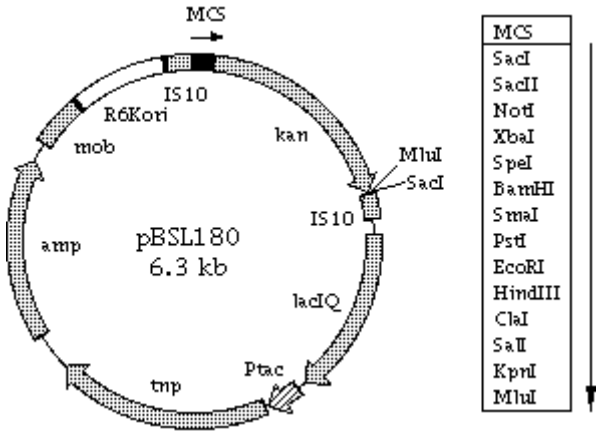


Fig 9.1 Map of pBSL180, showing the multiple cloning site and *npt2* gene (Alexeyev and Shokolenko, 1995).

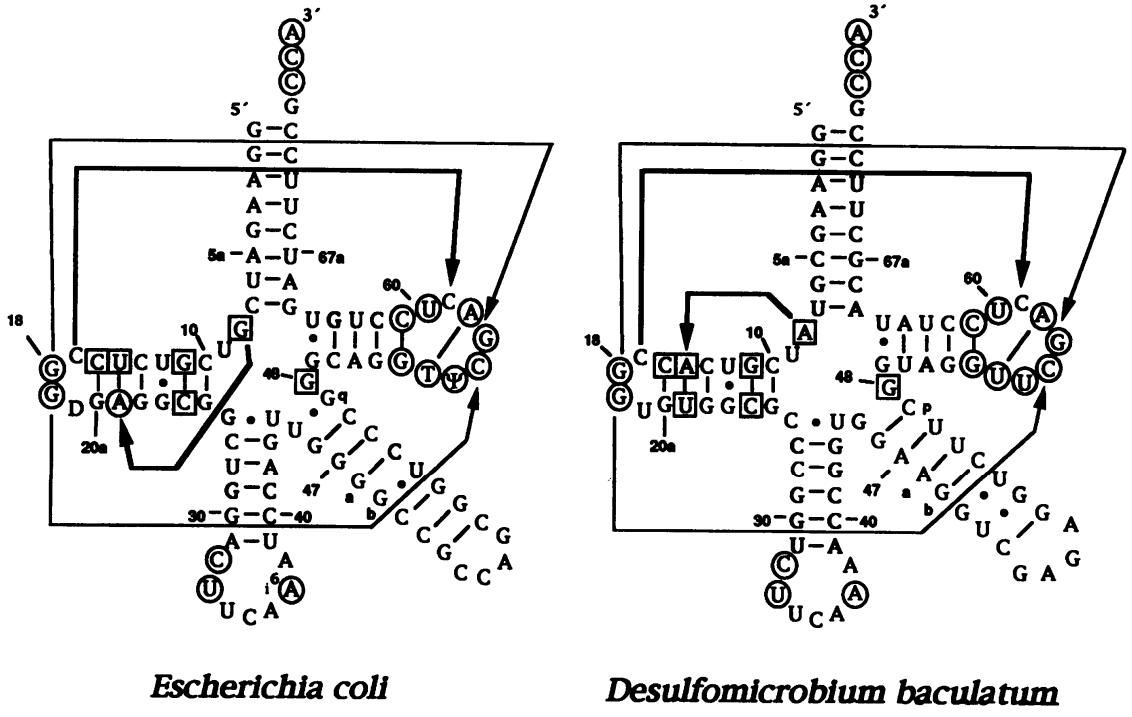


Fig 9.2 Cloverleaf structure of $tRNA^{Sec}$ species from prokaryotic sources identified thus far. The proposed tRNASec sequences from *E. coli* and *D. baculatum*. Tertiary interactions specific for $tRNA^{Sec}$ are indicated by bold arrows; interactions shared by all canonical tRNAs are represented by thin arrows. Nucleotides conserved in all tRNAs are circled; deviations from the canonical tRNA sequence are boxed. Figure adapted from Tormay et al, 1994.

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Unterschrift Düsseldorf, den ...../...../20...

Aus dem Max-Planck Institut für
Bioanorganische Chemie in Mülheim an der Ruhr

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Wolfgang Gärtner
Koreferent: Prof. Dr. Karl-Erich Jäger

Tag der mündlichen Prüfung:

(23-01- 2009)