NN08201, 1594

Molecular studies on iron-sulfur proteins in *Desulfovibrio*

Jack Stokkermans

CENTRALE LANDBOUWCATALOGUS
0000 0512 8885

Promotor : Dr C. Veeger
hoogleraar in de Biochemie

Co-promotor : Dr ir W. M. A. M. van Dongen
universitair docent

NN08201, 1579

Jack Stokkermans

Molecular studies on iron-sulfur proteins in *Desulfovibrio*

Proefschrift
ter verkrijging van de graad van doctor
in de landbouw- en milieuwetenschappen
op gezag van de rector magnificus, Dr H. C. van der Plas,
in het openbaar te verdedigen
op vrijdag 26 februari 1993
des namiddags te 13.30 uur in de
Aula van de Landbouwuniversiteit te Wageningen.

4095I

BIBLIOTHEEK
LANDBOUWUNIVERSITEIT
WAGENINGEN

Aan Jolande
Aan mijn moeder
ter nagedachtenis aan mijn vader

NN08201, 1594

Stellingen

1. De aanwezigheid van twee [6Fe-6S] clusters i.p.v. één in het prismaan eiwit van *Desulfovibrio desulfuricans* ATCC 27774, zoals wordt voorgesteld door Moura et al., wordt tegengesproken door de resultaten van Pierik et al. en hoofdstuk 5 en 6 van dit proefschrift.

   Moura, I., Tavares, P., Moura, J. J. G., Ravi, N., Huynh, B. -H., Liu, M. -Y. and LeGall, J. (1992) *J. Biol. Chem.* **267**, 4489-4496.
   Pierik, A. J., Hagen, W. R., Dunham, W. R. and Sands, R. H. (1992) *Eur. J. Biochem.* **206**, 705-719..
   Hoofdstuk 5 en 6 van dit Proefschrift.

2. In het onderzoek naar de rol van helpereiwitten in de biosynthese van metaal-cofaktor bevattende enzymen dient meer aandacht besteed te worden aan de karakterisering van het -vaak niet actieve- enzym dat akkumuleert in stammen waarin één of meerdere genen voor deze helpereiwitten gemuteerd zijn.

   Hoofdstuk 1 van dit proefschrift.

3. Het gebruik van "Miller units" in de β-galactosidase assays voor het vaststellen van promoter aktiviteit dient vermeden te worden.

   Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

4. Opheldering van het mechanisme van een aantal abiotisch reductiereacties (bv. nitraat tot ammonia; acetyleen tot ethaan en oxaloactaat tot fumaraat/succinaat), die, zoals recentelijk gevonden, onder gematigde kondities (druk, temperatuur en pH) kunnen verlopen in een milieu met FeS of FeS/H_2S, kan belangrijke inzichten opleveren voor de katalytische mechanismen in FeS-bevattende enzymen.

   Blöchl, E., Keller, M., Wächterhäuser, G. and Stetter, K. O. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 8117-8120.

5. Beweringen zoals o.a. in het artikel van Wynn et al. voorkomen, als zouden chaperones korrekte vouwing van ontvouwen of nieuw gesynthetiseerde eiwitten katalyseren, zijn onjuist.

   Wynn, R. M., Davie, J. R., Cox, R. P. and Chuang, D. T. (1992) *J. Biol. Chem.* **267**, 12400-12403.
   Georgopoulos, C. (1992) *Trends Biochem. Sci.* **17**, 295-299.

6. Het gebruik van antisense RNA technologie in de bestrijding van virale plantinfecties, zal, hoewel huidige resultaten bemoedigend zijn, meestal niet leiden tot een volledige resistentie van planten.

   Bejarano, E. R. and Lichtenstein, C. P. (1992) *Trends Biotechnol.* **10**, 383-388.

7. Gepopulariseerde nederlandse samenvattingen in proefschriften zijn zelden populair bij zowel wetenschappers als leken.

8. Het plan van de Vereniging van Samenwerkende Nederlandse Universiteiten (VSNU) en de Nederlandse Organisatie voor Wetenschappelijk onderzoek (NWO) om promovendi in de toekomst de status van student te geven, toont aan dat de beleidsmakers binnen beide organisaties in feite boekhouders zijn.

9. De koppeling van de financiering van universiteiten en HBO instellingen aan het aantal instromende en vooral aan het aantal doorstromende studenten, zal onherroepelijk een verlaging van de norm van deze onderwijsinstellingen tot gevolg hebben.

10. Er bestaan verrassend veel overeenkomsten tussen een promotor en een promoter.

Jack Stokkermans

Molecular studies on iron-sulfur proteins in *Desulfovibrio*

Wageningen, 26 februari 1993.

VOORWOORD

Een proefschrift is niet het werk van de promovendus alleen. Daarom wil ik allereerst een aantal mensen bedanken die op de een of andere manier een bijdrage aan de totstandkoming van dit proefschrift hebben geleverd.

Allereerst wil ik mijn promotor professor C. Veeger bedanken voor de supervisie van het onderzoek dat tot dit proefschrift heeft geleid. Daarnaast ben ik Walter van Dongen zeer erkentelijk voor zijn grote bijdrage aan dit proefschrift en ik wens hem veel sterkte in zijn eenzame strijd tegen het telkens weer oprukkende "radiogeweld" op het lab 7.

De samenwerking met de naaste collega's van lab 7, Willy van den Berg, Anita Kaan en Peter Fijneman heb ik niet alleen als zeer prettig ervaren, maar was onmisbaar voor het succesvol afronden van het experimentele gedeelte van dit proefschrift. Ook Pieter Houba, die in het kader van een afstudeervak een flinke bijdrage aan het prismaan eiwit onderzoek heeft geleverd, ben ik zeer erkentelijk.

Natuurlijk gaat veel dank uit naar de "zusterlab", lab 1 en 2. De hulp van Fred Hagen, Antonio Pierik, Marc Verhagen en Ronnie Wolbert, was onmisbaar voor het voltooien van hoofdstukken 4 t/m 6. Daarnaast ben ik ze dankbaar voor de vele "essays", uitgevoerd bij de koelkast van lab 1, die een welkome afsluiting waren van een zware wetenschappelijke dag.

Ook was de steun van de naaste collega's René Wientjens, Axel Berg, Martina Duyvis, Adrie Westphal en Jacques Benen, in goede, maar zeker ook in slechte tijden, enorm. Verder wil ik alle overige medewerkers van de vakgroep Biochemie bedanken voor de prima werksfeer en samenwerking.

Last but not least wil ik Jolande, mijn moeder en de rest van de familie, vrienden en kennissen bedanken die niet op een wetenschappelijke, doch ieder op hun eigen manier, hebben bijgedragen aan de totstandkoming van dit proefschrift.

The research described in this thesis was carried out at the Department of Biochemistry, Agricultural University, Wageningen, The Netherlands, under the direction of Prof. Dr C. Veeger.

The investigations were financially supported by the Netherlands Organisation for Scientific Research (N.W.O.) under the auspices of the Foundation for Chemical Research (S.O.N.).

CONTENTS

		Page
Abbreviations		2
Chapter 1.	General Introduction: Biosynthesis of redox-active metalloproteins	5
Chapter 2.	*Hyd* C, a gene from *Desulfovibrio vulgaris* (Hildenborough) encodes a polypeptide homologous with the periplasmic hydrogenase	41
Chapter 3.	Extensive sequence homology between domains of bacterial hydrogenases and the 75 kDa subunit of beef heart mitochondrial NADH:ubiquinone reductase: evidence for a new type of conserved element for electron transport.	51
Chapter 4.	The primary structure of a protein containing a putative [6Fe-6S] prismane cluster from *Desulfovibrio vulgaris* (Hildenborough).	67
Chapter 5.	The primary structure of a protein containing a putative [6Fe-6S] prismane cluster from *Desulfovibrio desulfuricans* (ATCC 27774)	85
Chapter 6.	Overproduction of prismane protein in *Desulfovibrio vulgaris* (Hildenborough): evidence for a second S=1/2-spin system in the one-electron reduced state.	95
Summary		113
Samenvatting		117
Curriculum vitae		125

Abbreviations

AMP, ADP and ATP	adenosine 5'-phosphate, 5'-diphosphate and 5'-triphosphate
bp	base pairs
Ci	curie
CoA	coenzyme A
Da	dalton
ddNTP	dideoxynucleoside triphosphate
DEAE	diethylaminoethyl
ΔG°	standard Gibbs free energy
dNTP	deoxynucleoside triphosphate
E_0'	redox potential measured at pH 7.0 under standard conditions
EDTA	ethylenediaminetetraacetate
E_m	midpoint redox potential
ENDOR	electron nuclear double resonance
EPR	electron paramagnetic resonance
EXAFS	extended X-ray adsorption fine structure
FAD	flavin-adenine dinucleotide
FMN	riboflavin 5'-phosphate
FPLC	fast protein liquid chromatography
GTP	guanosine 5'-triphosphate
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IEF	isoelectric focussing
IgG	immunoglobulin G
kb	kilobases
kDa	kilodalton
MCD	magnetic circular dichroism
M_r	relative molecular mass
N	an unspecified nucleoside
NAD, NAD^+ and NADH	nicotinamide-adenine dinucleotide and its oxidized and reduced forms
P_i	inorganic phosphate
SDS	sodium dodecyl sulfate
U	unit of activity: 1 μmole product per minute
V_m	maximal velocity, μmoles per mg enzyme per minute
Xaa	an unspecified amino acid

IPTG	isopropyl-β-D-thiogalactoside
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
TY	trypton yeast
PAGE	polyacrylamide gelelectrophoresis
Tris	tris (hydroxymethyl) aminomethane

CIP-Gegevens Koninklijke Bibliotheek Den Haag

ISBN 90-5485-092-2

Chapter 1

General Introduction

Biosynthesis of redox-active metalloproteins

I. REDOX-ACTIVE METALLOPROTEINS

Metalloproteins are very abundant in nature and fulfill a wide range of biological functions [1]. A large part of these metalloproteins acts as redox-active compounds participating in various oxidation reduction reactions. Some of them function as redox carrier proteins involved in electron transport (e.g. small blue copper proteins (also named cupredoxins) like azurin, amicyanin or plastocyanin; the heme-iron proteins such as cytochromes; the nonheme-iron proteins such as the ferredoxins). Other redox-active metalloproteins are involved in catalysis (e.g. succinate or xanthine dehydrogenase, superoxide dismutase, hydrogenases, nitrogenase). In these redox-active proteins the metal plays an essential role in the functioning of the protein. In redox-carrier proteins, electrons are transmitted by the metal cluster of the protein and in metalloproteins that act as enzymes, metals are part of the metal cluster or the metal cofactor that is essential for catalytic activity.

Until recently, very little knowledge existed on how these metals are incorporated into redox-carrier proteins or redox enzymes. Only in the past few years some insight has been obtained in the biosynthesis of these proteins, mainly by genetic studies: the construction of (mostly bacterial) mutants in which the synthesis of a specific redox-enzyme is impaired, followed by complementation of these mutants with a specific, cloned gene and analysis of the gene. From these genetic studies, the conclusion can be drawn, that many metal containing redox enzymes require a cluster comprising several genes for their biosynthesis. Such a gene cluster may contain:

1. Genes for the structural proteins of the enzyme;
2. Genes coding for proteins involved in regulation of the expression of the enzyme;
3. Genes encoding electron carrier proteins that are involved in electron transport to or from the enzyme;
4. Genes involved in the biosynthesis and incorporation of the metal (cluster) into the structural proteins.

Biosynthesis of these enzymes is often complex; evidence has been obtained that, at least in some cases, it requires more functions than are provided for by the genes in such a gene cluster. For instance chaperones or metal uptake- and processing systems with a more general function may be involved.

While genetic studies give more evidence about the proteins that are involved in biosynthesis of metal containing redox enzymes, in vitro reconstitution of a biosynthesic pathway with purified proteins has hardly been accomplished. Only a few examples have been described: incorporation of [2Fe-2S] or [4Fe-4S] clusters in ferredoxins [2,3] and repair of a defective [4Fe-4S] cluster in the Fe-protein of nitrogenase [4] with rhodanese

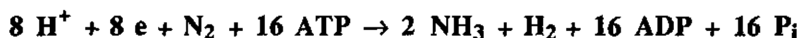
and the incorporation of preformed FeMo-cofactor in the MoFe-protein of nitrogenase with the Fe-protein [5,6].

In the following sections examples will be given of the biosynthesis of specific enzymes with redox-active metal(cluster)s. Lack of data on in vitro studies makes emphasis on the genetic aspects of biosynthesis of these proteins unavoidable.

II. NITROGENASES

Types of nitrogenase. Nitrogen fixation is the term for the enzymatic conversion of atmospheric N_2 to NH_3 that is performed by certain diazotrophic (N_2-consuming) bacteria (e. g. *Azotobacter*, *Rhizobium*, *Klebsiella* and *Rhodobacter*), which are the founding organisms of the food chain since nitrogen compounds are required for various metabolic processes. Nitrogen fixation is performed by a complex enzyme named nitrogenase. Nitrogenase consists of two interacting component proteins: the MoFe-protein (component 1) and the Fe-protein (component 2). In some nitrogenases the MoFe-cofactor of component 1 is replaced by a VFe-cofactor or a Fe-only cofactor. Hence the nitrogenases are classified as the Mo-dependent nitrogen fixing system (*nif*), the V-dependent nitrogen fixing system (*vnf*) and the alternative nitrogen fixing system (*anf*), respectively. The best studied enzymes are the nitrogenases in *Klebsiella pneumoniae*, *Azotobacter vinelandii* and *A. chroococcum* [7]. In *K. pneumoniae* only the Mo-dependent nitrogen fixing system, *nif*, is present, while the two *Azotobacter* species contain also the V-dependent, (*vnf*)-system and *A. vinelandii* contains also the third alternative (*anf*)-system. Later, the nitrogen fixing systems *vnf* and *anf* were also found in many other diazotrophic organisms [8].

Nitrogenase activity. The basic reaction that is catalyzed by the nitrogenases is:



$$8\ H^+ + 8\ e + N_2 + 16\ ATP \rightarrow 2\ NH_3 + H_2 + 16\ ADP + 16\ P_i$$

Component 1 contains the active site for N_2 reduction. Under optimal conditions, 8 electrons are required to reduce each N_2 molecule to NH_3 (Fig. 1, step 3). These electrons are provided by e.g. the oxidation of pyruvate (in *Klebsiella*) or, probably, the respiratory chain (in *Azotobacter*) and transferred by a low-potential flavodoxin [9] and component 2 to component 1 (Fig. 1, steps 1 and 2). As flavodoxin and component 2 catalyze one-electron transfer, eight consecutive one-electron transfer steps, each requiring 2 ATP molecules, are necessary to provide component 1 with the electrons for reduction of one N_2. A mechanism for the nitrogenase reaction has been proposed by Lowe and Thorneley [10,11] and the same mechanism has also been demonstrated for the V-dependent nitrogenase [12].

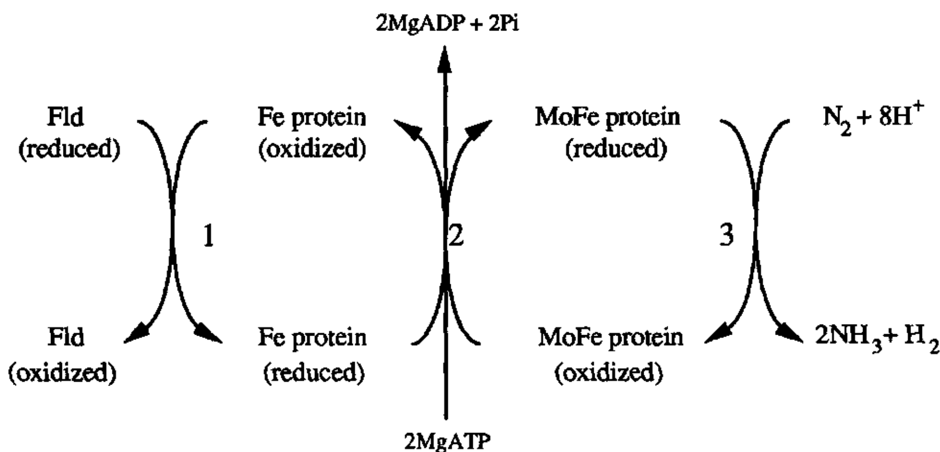


Fig. 1 Transfer of electrons through nitrogenase. Fld, Flavodoxin.

Composition. The oxygen-sensitive component 2 that is encoded by the *H*-gene (Fig. 4A), occurs as a γ_2 dimer with molecular mass of approximately 60 kDa and contains 4 Fe and 4 S^{2-} per dimer. These 4 irons are coordinated in a single [4Fe-4S] cluster that functions as a one-electron donor between the +1 and +2 oxidation levels. It was indicated from site-directed mutagenesis experiments [13] and later confirmed by the three dimensional structure of the Fe-protein [14], that the FeS cluster is ligated by two conserved Cys residues (Cys97 and Cys132) and that a single [4Fe-4S] cluster bridged the two subunits of the component 2 dimer. Cys97 and Cys132 mutants also lost component 1 activity, indicating that the Fe-protein may be involved in activation of component I [15]. In addition to this iron-sulfur ligation site, an ATP binding motif Gly-Xaa-Gly-Xaa_2-Gly-Lys-Ser near the NH_2-terminus is conserved in all component 2 sequences. The three dimensional structure of the Fe-protein (resolution 0.29 nm) has been solved by X-ray spectroscopy [16].

In the Mo nitrogenase, component 1 occurs as a tetrameric $\alpha_2\beta_2$ structure of about 220 kDa. The α and β subunits are encoded by the *nifD* and *-K* genes, respectively (Fig. 4A). The VFe-protein from *A. chroococcum* has an $\alpha_2\beta_2\delta_2$ subunit structure with the δ subunit being encoded by the *vnfG* gene (Fig. 4A) [17]. Component 1 contains 2 Mo, approximately 32 Fe and 32 S^{2-}-ions per tetramer. Spectroscopic studies together with cluster extrusion techniques, indicates that approximately half of the irons is present in the complex P-clusters (two or four clusters/tetramer), while the rest, together with the Mo-ions is present in the FeMo cofactor (two cofactors/tetramere) [17]. A model based on X-ray data at 0.29 nm resolution indicates that the P-clusters may consist of two closely linked, cubane-like [4Fe-4S] clusters that are bridged by the thiol groups of Cys88 of the α subunit and Cys95 of the β subunit (Fig. 2) [18]. Each thiol-group coordinates two irons, one from each cluster, resulting in an unusual coordination

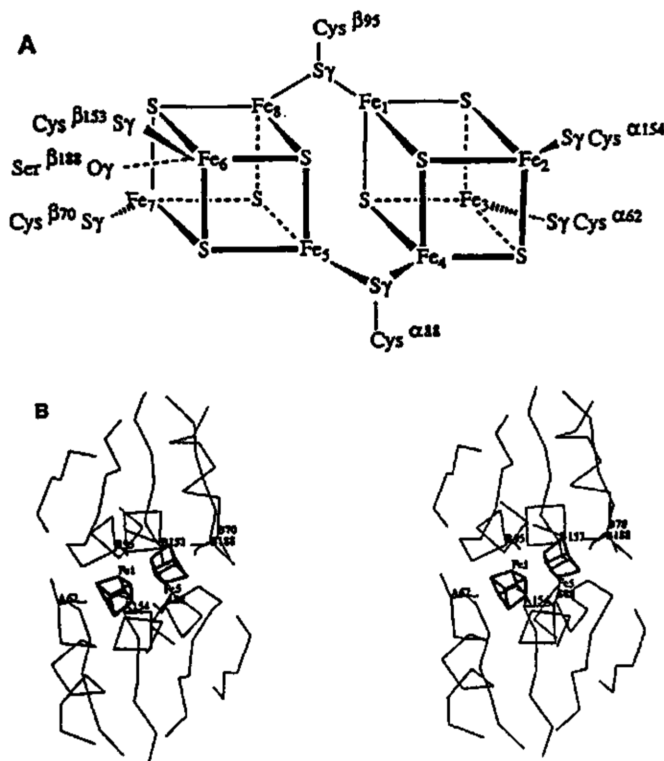


Fig. 2 (A) Schematic representation of the P-cluster model. (B) Stereoview of the P-cluster and surrounding protein model. The view is approximately along the direction of a twofold rotation axis approximately relating the α and β subunits [18].

number for the sulfur atom. EPR, MCD and Mössbauer studies all propose a high-spin Fe-S cluster with a S=7/2 spin system for the P-clusters [19].

The FeMo-cofactor has been reported to contain (R)-homocitrate as organic component [20]. Five possible model structures have been suggested based on EXAFS, ENDOR, EPR and Mössbauer data for the cofactor centers of Mo- and V-nitrogenases [7]. Each model structure has some attraction, but they all demonstrate convincingly that these data are inadequate for the prediction of the three dimensional structure. Also for this cofactor a model structure has been proposed recently based on X-ray data (Fig. 3) [18]. This model structure is different from those proposed by analysis of the spectroscopic data.

Nitrogenase gene clusters. With the development of molecular biology, the complexity of the nitrogen fixing system became apparent (Fig. 4A). The *nif* gene cluster in *K. pneumoniae* comprises, besides the three structural *nifHDK* genes, 17 additional genes. Two of them (*nifAL*) are involved in regulation, one (*nifF*) encodes a flavodoxin, another (*nifJ*) encodes a pyruvate-flavodoxin oxidoreductase, and the other genes are somehow involved in the biosynthesis of the enzyme, although this has not been

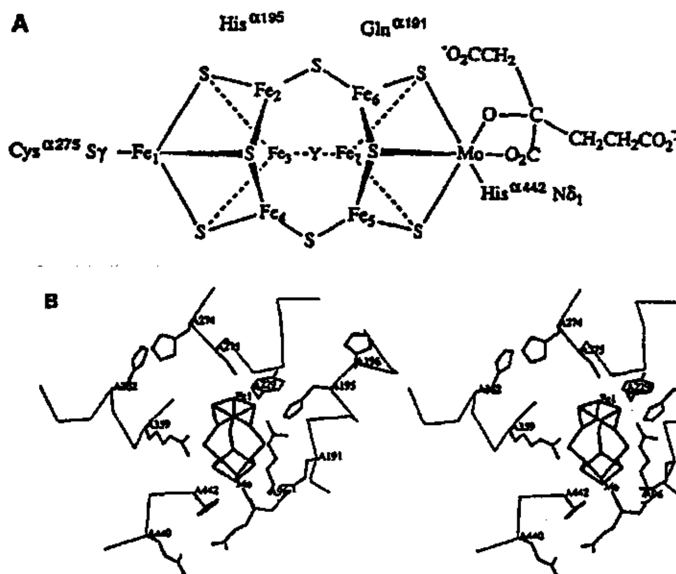


Fig. 3 (A) Schematic representation of the FeMo-cofactor model. Y represents the bridging ligand with relatively light electron density. (B) Stereoview of the FeMo-cofactor and surrounding protein environment [18].

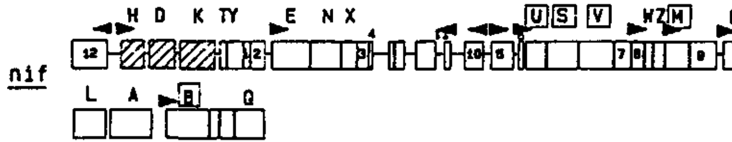
demonstrated for all of them. These genes are clustered on a 23 kb region of the chromosome [21]. The *nif* cluster in *Azotobacter* species contains besides 19 of the 20 *nif* genes of *Klebsiella*, additional open reading frames of unknown function interspersed between the *nif* genes. The two additional nitrogenase clusters, *vnf* and *anf*, that are both present in *A. vinelandii* contain besides the four structural *HDKG*-genes, only 3 and 1 additional genes, respectively. Deletion of the *nifUSVM* and *B*-genes in *A. vinelandii* results in a Nif^-, Vnf^- and Anf^- phenotype, indicating that these genes are involved in the biosynthesis of all three nitrogenases [22,23]. The *nifEN*-genes have homologues in the *vnf*-cluster, but not in the *anf*-cluster. The *vnfEN* genes are involved in biosynthesis of both V- and alternative nitrogenases [24].

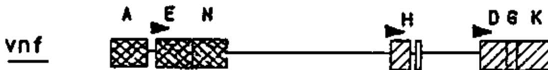
Regulation of *nif* gene expression. The expression of *nif* operons is regulated by the environment of the diazotroph. Nitrogenase production is repressed by NH_3 or O_2. Transcription of the *nif* operons starts from specific *nif* promoters that all have a common structure in *Klebsiella, Azotobacter* and *Rhizobium* (Fig. 4B) [26]. They all contain specific -12 and -24 sequences that are recognized by RNA polymerase containing an alternative sigma factor, σ^{54} [27]. In addition to the common σ^{54}-dependent promoter, *nif* gene expression is dependent on the sequence TGT-N_{10}-ACA, the upstream activator sequence (UAS), which is the site of interaction with the *nifA* gene

A

Klebsiella pneumoniae

Azotobacter vinelandii





0 10 20 27 kilobases

B

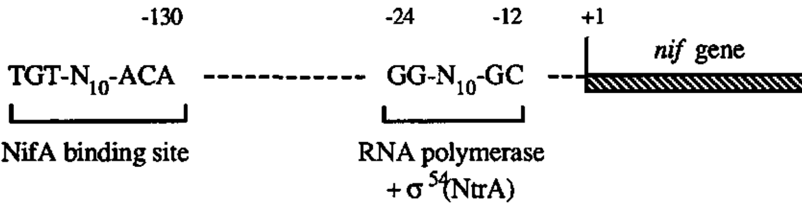


Fig. 4 (A) The organization of nitrogen fixation genes in *Klebsiella pneumoniae* and *Azotobacter vinelandii*. Arrowheads indicate the presence of σ^{54}(*rpo*N) dependent promoter sequences [25]. From ref. [7]. (B) Structure of NifA-dependent *nif* promoters

product. The *nifA* gene encodes a transcriptional activator for *nif* operons. The *nifA* gene is cotranscribed with a gene for another regulatory protein, *nifL*. In *Klebsiella*, NifL acts as a repressor for *nif*-gene transcription in the presence of O_2 and fixed nitrogen. Under these conditions, NifL interacts with NifA and prevents the activator function of NifA. The repressor function of NifL might be dependent on specific phosphorylation of the NifA protein or on the PII protein [27].

Transcription of the regulatory *nifLA* genes in *Klebsiella* is under control of a second two-component regulatory system: NtrB/NtrC. In the absence of NH_3, NtrC is phosphorylated by NtrB and acts as an activator of *nifLA* expression [27]. In *Azotobacter* and *Rhizobia* regulation of nif genes expression differs in some respects from *Klebsiella*. In *R. meliloti*, *nifA* expression is prevented by oxygen, while in *Bradyrhizobium japonicum* the NifA protein itself is inactivated by anaerobic growth [28,29].

The biosynthesis of MoFe-nitrogenase. The biosynthesis of MoFe-nitrogenase was studied by two methods. In *A. vinelandii* it was performed by studies on the physiological phenotypes of *A. vinelandii* strains that had deletions in specific genes from the *nif* cluster [22-24,30,31]. Biosynthesis of *K. pneumoniae* nitrogenase was studied by complementary expression of structural and additional *nif* genes in *E. coli* [32,33].

Biosynthesis of the Fe-protein. Both studies clearly demonstrated the involvement of *nifM* in the biosynthesis of component 2 (Fe-protein). In *E. coli*, co-expression of only *nifH* and *nifM* genes resulted in a fully active *K. pneumoniae* component 2, containing one [4Fe-4S] cluster that showed normal EPR spectra [32]. Expression of the *nifH* gene alone resulted in accumulation of inactive component 2 without Fe-S clusters [32].

Also in *A. vinelandii*, NifM is required for biosynthesis of the Fe-protein, as deletion of the *nifM* gene resulted in accumulation of inactive Fe-protein. However, also NifU and NifS seem to be involved in biosynthesis of component 2 in *A. vinelandii*: in *nifU* or *nifS* mutants, the specific activity of the Fe-protein was reduced to 5% of the wild-type activity and in a *nifUS* double mutant to 1%. NifU and NifS might be required for full activation of stability of component 2 in *Azotobacter*.

Biosynthesis of the MoFe-protein. The biosynthesis of the MoFe-protein requires besides the structural genes *nifDK*, at least the *nifH, nifE, nifN, nifU, nifS, nifV, nifW, nifZ, nifM, nifB*, and *nifQ* genes in both *A. vinelandii* and *K. pneumoniae* [23,30,31,33]. NifH, the Fe-protein, catalyzes the last step in biosynthesis of the MoFe-protein. *NifH* mutants of *A. vinelandii* accumulate FeMo-cofactor-less MoFe-protein that contains approximately 16 Fe ions [6,34]. Also an *E. coli* strain harboring the *K. pneumoniae nifDKTYVSWZMBQ* genes [5,33], accumulates FeMo-cofactor-less MoFe-protein. The FeMo-cofactor could be inserted into FeMo-cofactor-less MoFe-protein by incubation of the apoprotein with purified FeMo-cofactor in the presence of the Fe-protein, dithionite and ATP [5,33,34]. This resulted in fully active component 1. Even catalytically inactive Fe-protein was still effective in insertion of the FeMo-cofactor into the MoFe-apoprotein [33,35].

Deletion of the *nifEN* genes in the *K. pneumoniae* or *A. vinelandii nif* gene clusters also resulted in accumulation of inactive, FeMo-cofactor-less component 1 [33]. The

predicted polypeptides synthesized by the *nifE*- and *N* genes have sequence homology with the α and β subunits of the MoFe-protein [31] and it has been suggested that the NifEN proteins might provide a scaffold for synthesis of FeMo-cofactor [7,31]. The NifEN proteins have been purified as an $\alpha_2\beta_2$ tetramer [36].

Also deletion of the *nifV* gene resulted in production of FeMo-cofactor-less component 1 [31-33]. The Nif^- phenotype of *K. pneumoniae nifV* mutants can be cured by adding homocitrate to the culture medium [37]. This curing effect of homocitrate was not detected in *A. vinelandii nifV* mutants [30], but in vitro reconstitution experiments with ^3H-labeled homocitrate demonstrated that homocitrate or an organic component derived from it, was incorporated into the FeMo-cofactor of *A. vinelandii* component 1 [38]. Sequence comparison suggests that NifV is a homocitrate synthase [39]. NifV is required for all three nitrogenases in *A. vinelandii*, indicating that homocitrate is a common compound of their cofactors [22].

The involvement of NifB and Q in the biosynthesis of MoFe protein from *K. pneumoniae* was not analyzed by co-expression in *E. coli*, but by the study of *K. pneumoniae nifB* and *nifQ* mutants [5,40]. Characterization of extracts from *A. vinelandii nifB* mutants demonstrated the requirement for NifB in the biosynthesis of FeMo-cofactor, FeV-cofactor and the cofactor of the alternative nitrogenase and strongly suggested that the step catalyzed by the *nifB* gene product may be common in all three systems [23]. The FeMo-cofactor-less component 1 accumulated in *nifB* mutants had a different mobility in non-denaturating polyacrylamide gels than FeMo-cofactor-less component 1 from *nifH* or *nifE* mutants [41]. Extracts containing FeMo-cofactor-less component 1 from a *nifB* mutant could not be activated to wild-type activity with Fe-protein and purified FeMo-cofactor [41]. It has been suggested that the binding site for FeMo-cofactor in component 1 from *nifB* mutants is occupied by a precursor of the cofactor (e. g. MoS_4^{2-} [42]) that requires the NifB protein for further processing [41]. A similar suggestion has been made for the role of NifB in activation of *K. pneumoniae* component 1 [5].

Diazotrophic growth of *nifQ* mutants is very slow with the normal concentration of Na_2MoO_4 (1 μM) in the growth medium, but is increased to normal rates by elevation of the molybdenum concentration (100 μM for *K. pneumoniae* and 1 mM for *A. vinelandii*) [23,40]. It has been shown that *K. pneumoniae nifQ* mutants have the same affinity for molybdate than a wild-type strain [30]. Therefore, NifQ is probably not involved in Mo-uptake, but in accellerating a step in the processing of Mo.

Besides the genes absolutely required for biosynthesis of active component 1, others seem only to be necessary to obtain full activity of the MoFe-protein. The precise functioning of these genes is not yet clear. In *A. vinelandii*, deletions in *nifU,S,W,Z* or *M* genes resulted in 50-80% reduced component 1 activity [30]. This reduction was

higher (approximately 90%) in *nifUS* and *nifZM* double mutants. Since some of these genes are also involved in biosynthesis of the Fe-protein (*nifU, S, M*), which in its turn is required for the biosynthesis of component 1, it is not clear, whether mutations of *nifU, S* and *M* genes affect the biosynthesis of the MoFe-protein directly or indirectly.

Analysis of the function of above mentioned *K. pneumoniae nif* genes by expression in *E. coli* yielded slightly different results. First, it was shown that the *nifS* gene was absolutely required for biosynthesis of the apo MoFe-protein [33], contrary to what has been found in *A. vinelandii* [30]. Second, in addition to the *nifUSWZM* genes, also the *nifY* gene was required for expression of maximal component 1 activity in *Klebsiella* [33]; in *A. vinelandii nifY* mutants, both component 1 and 2 activities of the MoFe-nitrogenase were unaffected, but NifY was only required for full activity of the third, alternative nitrogenase [22]. Furthermore, it was shown by co-expression of different *K. pneumoniae nif* genes in *E. coli* that in the absence of the *nifM* gene, either *nifYU* genes or *nifWZ* genes are absolutely required for expression of active component 1 [33]. In *A. vinelandii nifM* mutants require only *nifZ* (or vice versa) [30].

Functions of the *nifT* and *nifX* genes in *K. pneumoniae* and *A. vinelandii* are unclear; deletion of these genes does not affect activities of components 1 and 2 [30,33].

In the biosynthesis of nitrogenases also other genes, encoding proteins with a more general function that are no part of the *nif* gene clusters, are involved. The sulfurtransferase rhodanese has been shown to reactivate component 2 that had lost part of the Fe-S cluster [4]. Other proteins might be involved in the assembly of the subunits. Expression of the *K. pneumoniae nif* gene cluster in an *E. coli* strain in which the genes for the chaperonin GroEL had been deleted, resulted in the production of only very low amounts of the nitrogenase subunits, with reduced stability [43]. Wild-type levels were restored, when this *E. coli* strain also received a plasmid with the cloned *groE* operon [43]. There are indications for transient binding of GroEL, not only with structural NifH and NifDK proteins, but also with the regulatory protein NifA.

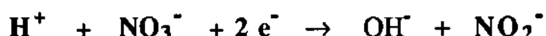
Another protein, with a mass of 20 kDa and apparently not encoded by the *nif* gene cluster, was found to be associated with the FeMo-cofactor-less component 1 of an *A. vinelandii nifB* mutant [36]. Its function has not been elucidated.

General systems involved in uptake and processing of metals are important for the biosynthesis of nitrogenases. In *K. pneumoniae*, a Mol^- mutant was identified as an equivalent to *chlD* mutants of *E. coli* [44]. The *chlD* gene encodes the Mo-uptake system in *E. coli* [44].

Contrary to the extended knowledge of the biosynthesis of the FeMo cofactor of nitrogenase component I, less is known about the biosynthesis of the P-clusters [7].

III. NITRATE REDUCTASE

Nitrate reductase activity. In anaerobic cultures of enterobacteria (e. g. *E. coli*), nitrate is the favored alternative electron acceptor in the absence of oxygen [44]. Two membrane-bound nitrate reductases (NRA and NRZ) exist in *Escherichia coli*. NRA catalyzes the reduction of nitrate to nitrite coupled to formate oxidation probably via quinones and *b*-type cytochromes [45,46]. The function of NRZ in the *E. coli* metabolism is not yet clear.



$$H^+ + NO_3^- + 2e^- \rightarrow OH^- + NO_2^-$$

Nitrate reductase composition. Both iso-enzymes, NRA and NRZ, are composed of three subunits that form an $\alpha\beta\gamma$ trimer. Both α subunits (α_a and α_z) have comparable molecular masses (M_r 150 kDa) and contain a Mo-cofactor (a molybdopterin-cofactor that is different from the FeMo-cofactor in nitrogenase) that forms the catalytic site for nitrate reduction [47]. The β subunits, βa and β_z (M_r 60 kDa), are electron-transfer units which are thought, on base of the four cysteine motifs in their amino acid sequence [48], to contain the one [3Fe-3S] cluster and three [4Fe-4S] clusters that have been proposed for the $\alpha\beta$ dimer according to observations made by EPR spectroscopy [49]. The two enzymes are anchored to the membrane by the membrane-bound *b*-type cytochrome subunits, γ_a and γ_z (M_r 25 kDa), which receive electrons from the quinone pool and transfer them to the FeS centers of the β subunit [50].

Regulation of nitrate reductase expression. The enzymes are encoded by different operons, *narA* and *narZ*, which are identically organized [51], but are regulated in opposite manners. NRA production is repressed under aerobic conditions and strongly induced by nitrate [52], while NRZ is anaerobically repressed and induced weakly by nitrate [53]. For induction of NRA synthesis under anaerobic conditions, three genes upstream of the *narA* operon are required, *narL*, *narK* and *narX* [54]. The *narK* gene enhances nitrate uptake and nitrite excretion in *E. coli* and its gene product is thought to be part of the transport system that mediates supply and drainage of substrate and product, respectively [55]. The *narL* gene product acts as a nitrate-responsive positive activator, that presumably binds to an upstream site in the control region of the *narA* operon [56,57]. The *narX* gene product was recently identified as a nitrate sensor [58], and it was proposed the *narL* and *narX* compose a two-component regulatory system not only to induce the *narA* operon expression, but also to repress that of enzymes involved in fermentative pathways (e.g. fumarate reductase) [59]. Another regulatory gene at a distinct locus, *narQ*, was recently isolated. From sequence homology and from

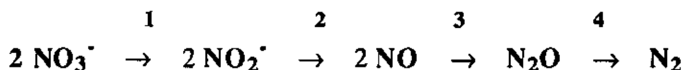
mutational analysis, it was concluded that *E. coli* possesses two distinct nitrate sensor-transmitter proteins, NarX and NarQ, that both are able to interact with NarL to regulate NRA expression [60].

Biosynthesis of nitrate reductase. In the biosynthesis of nitrate reductase, three genes of the *narA* and *narZ* gene clusters, *narG, H, I* and *narZ, Y, V*, encode the structural subunits $\alpha_a,\beta_a,\gamma_a$ and $\alpha_z,\beta_z,\gamma_z$ respectively [48,53]. It was demonstrated by the characterization of *narI* mutants that the γ subunit of the enzyme complex is not required for an active NRA enzyme. In these mutants an active form of the $\alpha\beta$ complex, which is water-soluble and resides in the cytoplasmic fraction, can be released after heat treatment [61]. In the two *nar* operons, *narGHJI* and *narZYWV*, the third genes (*narJ* and -*W*, respectively), encode a non-structural protein (M_r 26.5 kDa) that is required for the correct association of the α and β subunits, but is no part of the active complex [62,63]. Nitrate reductase isolated from *narJ* mutants exhibited low levels of activity and was less well protected against proteases and heat denaturation, but it still contained non-heme iron and the Mo-pterin cofactor. It was concluded that the protein was not involved in cluster biosynthesis or insertion but in the assembly of the enzyme complex, although the protein is not homologous with bacterial chaperonins [62].

The biosynthesis of nitrate reductase comprises also the uptake of molybdenum and the synthesis and insertion of the molybdopterin cofactor. These functions are fulfilled by genes from the *Chl* gene cluster. Genes from this cluster are involved Mo-cofactor biosynthesis of all molybdopterin containing enzymes (for example, formate dehydrogenase and dimethylsulfoxide-reductase). Mo-uptake is performed by the *ChlD* and *G* gene products [44], while synthesis of the Mo-pterin cofactor involves the gene products of *ChlA,E* and *B* [64,65]. There are no indications that an insertase is required for incorporation of the Mo-cofactor into the apoenzyme [66,67]: incubation of apoenzyme with Mo-cofactor results in spontaneous reconstitution of the active enzyme. However, it was indicated that free cofactor is extremely labile and is probably stabilized in vivo by association with specific proteins serving as carrier molecules [66,67].

IV. NITROUS OXIDE REDUCTASE

Nitrous oxide reductase activity. Nitrous oxide (N_2O) reductase from *Pseudomonas stutzeri* is a periplasmic enzyme that is part of a four step respiratory system that catalyzes the reduction of nitrate to molecular nitrogen:



$$2\,NO_3^- \xrightarrow{1} 2\,NO_2^- \xrightarrow{2} 2\,NO \xrightarrow{3} N_2O \xrightarrow{4} N_2$$

$$\text{(overall reaction: } 2\,NO_3^- + 10\,[H] + 2\,H^+ \rightarrow N_2 + 6\,H_2O\text{)}$$

The metalloenzymes that are part of this system are: nitrate reductase (*nar*, step 1), nitrite reductase or cytochrome cd_1(*nir*, step 2), nitric oxide reductase (*nor*, step 3) and nitrous oxide reductase (*nos*, step 4). This denitrifying system is induced under anaerobic conditions in the presence of nitrate or N_2O [68,69], although low levels of nitrous oxide reductase are produced constitutively [70].

Nitrous oxide reductase composition. The nitrous oxide reductase, which was shown to be a homodimeric multi copper enzyme, contains a Cu chromophore with Cu(I)-Cu(II) interaction that is thought to be similar to that of cytochrome-c oxidase, as evidenced by its primary structure and by several spectroscopical studies (e.g. EPR, X-ray and MCD) [71-73].

Biosynthesis of nitrous oxide reductase. In *Pseudomonas stutzeri* a DNA fragment of 30 kb was found that harbors the *nir, nor* and *nos* gene clusters (Fig. 5) [74]. The nitrous oxide reductase (*nos*) is encoded by a 8 kb gene cluster, which besides the structural gene (*nosZ*) contains a presumed regulatory gene (*nosR*) and functions that are involved in the biosynthesis of the Cu chromophore (*nosDFY*, the chromophore region) [75]. Mutagenesis of each of these genes resulted in a catalytically inactive enzyme with a low metal content. Expression of the structural gene *nosZ* alone in *E. coli* resulted in a Cu-less apoenzyme [72]; after addition of Cu^{2+} ions to the apoenzyme, some ions were incorporated, but enzyme activity was not restored and EPR spectra of the reconstituted apoprotein were different from those of the active enzyme [76]. These observations favor the idea of a catalyzed insertion of Cu into the protein and the involvement of *nosDFY* therein [75]. The three genes have been sequenced and their predicted products have been analyzed [75]. NosD (M_r 48.2 kDa; mature protein 45.4 kDa) carries an export signal and is thought to be a periplasmic component. The *nosF* encoded protein belongs to the superfamily of ATP/GTP-binding proteins involved in prokaryotic and eukaryotic import-export processes of small molecules (e. g. molybdate, sulfate, iron, arabinose) [77]. The NosF protein is a hydrophilic component with no predicted transmembrane helices and since it lacks a predicted export signal, it is presumed to be cytoplasmic [75]. The NosY protein is extremely hydrophobic and its hydrophobicity profile is very similar to that of the membrane-bound components of bacterial transport systems [78].

In Fig. 6 a model for the location of the components of the N_2O reductase system is presented. The chromophore genes *nosDFY* resemble those of bacterial transport operons for small molecules (e. g. maltose-, arabinose-, iron- or inorganic-phosphate-uptake), which usually are composed of a periplasmic binding protein, two intrinsic membrane proteins and a nucleotide binding protein [79]. An additional gene, *nosA*, has been identified, which is strongly repressed by Cu-ions in the culture medium. Mutagenesis of

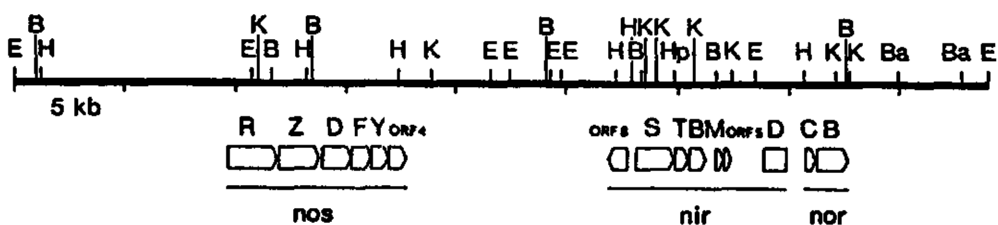


Fig. 5 Physical map of the gene clusters encoding nitrite reductase (nir), nitricoxide reductase (nor) and nitrous oxide reductase (nos) in *P. stutzeri* [74].

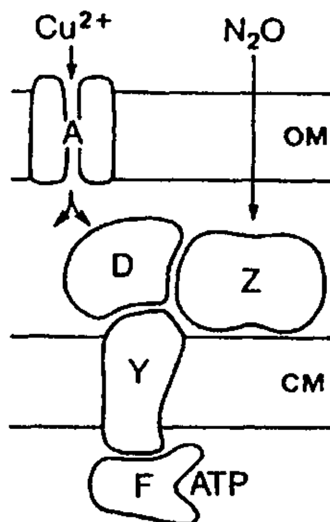
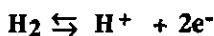


Fig. 6 Model for the localization and interaction of the Cu-processing proteins encoded by the chromophore region (NosD, F and Y), the outer-membrane protein NosA, N_2O reductase (NosZ) and the cytoplasmic membrane in *P. stutzeri* [75].

the *nosA* gene results in the production of apo-N_2O reductase. NosA has been found in the outer membrane and acts as a channel-forming and Cu-binding protein; the function of NosA might be Cu-uptake and/or processing [80,81].

V. HYDROGENASES

Hydrogenase activity. Important metabolic conversions in different organisms are the oxidation of molecular hydrogen / the reduction of protons to molecular hydrogen, which are catalyzed by hydrogenases [82].



Hydrogenases have been divided in uptake hydrogenases that catalyze in vivo the oxidation of molecular hydrogen, and hydrogenases that catalyze H_2 production.

The physiological functions of hydrogenases can be threefold [83]:

1. The oxidation of H_2 is coupled to the reduction of several electron acceptors (e.g. O_2, NO_3^-, SO_4^{2-}, CO_2, NAD^+).
2. The production of H_2 is part of fermentative processes and functions in the removal of excess electrons (for example, coupled to lactate or pyruvate fermentation).
3. The oxidation of H_2 generates a proton gradient across the cytoplasmic membrane, which is used for ATP-production.

Hydrogenases are divided into two major groups on base of the composition of their metal centers : the Fe-hydrogenases that contain only iron-sulfur clusters (discussed later) [82] and the NiFe-hydrogenases that contain besides non-heme iron a redox-active Ni in their catalytic centers [84]. Some members of the latter group contain also selenium in the form of selenocysteine and have been classified as NiFeSe-hydrogenases [85].

Distribution of the NiFe-hydrogenases is, contrary to the Fe-hydrogenases, not restricted to the strictly anaerobic bacteria, but they are also found in facultative anaerobes and aerobes. Till now, approximately 15 Ni-containing hydrogenases have been sequenced [83]. The structures of the operons encoding the NiFe-hydrogenases from a variety of bacteria have been elucidated and insight in the biosynthesis is emerging rapidly, although much is still unclear . The genetically best studied examples of NiFe-hydrogenases are those present in *E. coli* and *Alcaligenes eutrophus* and hence the discussion here will be restricted to the NiFe-hydrogenases present in these two organisms.

V.I.I. NIFE-HYDROGENASES IN *E. COLI*

NiFe-hydrogenase activities in *E. coli*. *E. coli* possesses at least three NiFe-hydrogenases, hydrogenases-1, -2 and -3 [86-89]. Hydrogenases-1 and -2 are membrane-bound uptake hydrogenases. Hydrogenase-2 catalyzes H_2 oxidation coupled to the anaerobic reduction of fumarate; the physiological role of hydrogenase-1 is not clear [87]. The physiological function of hydrogenase-3 is that of a H_2-evolving hydrogenase part of the formate hydrogenlyase (FHL) complex that catalyzes formate oxidation coupled to proton reduction [90].

NiFe-hydrogenase composition. Biochemically, hydrogenase-3 is the least characterized isoenzyme because it has been proved difficult to purify due to its

instability. Therefore, its subunit structure and its metal content remain unknown. On the other hand, both hydrogenases-1 and -2 have been purified and extensively characterized at the biochemical level [87,88]. These two isoenzymes are both membrane-bound and consist of two subunits composing a $\alpha_2\beta_2$ heterodimeric complex. Both large α subunits have a molecular mass of 60 kDa, while the two small β subunits have molecular masses of 32 kDa and 35 kDa for hydrogenase-1 and -2, respectively [87,88]. Hydrogenase-1 contains approximately 11 non-heme iron atoms and 0.64 atoms of nickel per mol of enzyme [87]. Spectroscopic analysis of the *E. coli* hydrogenases has been hampered by limited amounts of the hydrogenases that are produced by the *E. coli* strains. However, it has been shown that hydrogenases from different species (*Desulfovibrio, Azotobacter*) are quite homologous with the *E. coli* enzymes, and especially the *Desulfovibrio* hydrogenases have been studied extensively with spectroscopic techniques (reviewed in [83]).

Biosynthesis of NiFe-hydrogenases in *E. coli*. The clusters coding for hydrogenase-1 and -2 (the *hya* and *hyb* cluster, respectively) contain both genes for the two subunits and four and five additional genes, respectively (Fig. 7). The large subunits encoded by *hyaB* and *hybC*, are not only highly conserved (approximately 70% conserved amino acids), but they are also highly homologous with the large subunits of NiFe-hydrogenases from *A. chroococcum* [91], *A. vinelandii* [92], *Br. japonicum* [93], *Rhodobacter capsulatus* [94] and *Rhodocyclus gelatinosus* [95]. The small subunit of hydrogenase-1, encoded by *hyaA*, has some homology with other small subunits (*Azotobacter, Rhizobia* and *Rhodobacter*) [83]. The small subunit of hydrogenase-2 is probably encoded by the *hybA* gene [88,96]; this subunit has not much homology with small subunits of other hydrogenases. The predicted small subunits, encoded by *hyaA*, and *hybA*, contain short signal peptides for export, which is consistent with the membrane localization of both hydrogenases [88,97].

It was suggested that the *hyc* cluster, encoding hydrogenase-3, contains only one structural gene, *hycE*, based on its homology with the genes for the large subunits of hydrogenase-1 and -2. The other genes in the *hyc* operon do not exhibit homology with genes found in the *hya* and *hyb* operons [83]. It is not clear, whether hydrogenase-3 has a small subunit.

Some hydrogenases, for example hydrogenase-1 of *E. coli*, can be isolated in monomeric form, comprising only the large subunit. Although it is not clear, if these monomeric forms have any physiological relevance, they are catalytically active and have been found to contain the nickel [83]. This indicates that the nickel binding site is located in the large subunits. EXAFS studies with *Desulfovibrio* hydrogenase reveal that the nickel is penta- or hexa coordinate and is ligated to sulfur and nitrogen or oxygen ligands (reviewed in [83]). In the primary structures of the large subunits, 2 motifs were

Hydrogenase 1

hya A hya B hya C hya D hya E hya F

Hydrogenase 2

hyb A hyb B hyb C hyb D hyb E hyb F hyb G

Hyp operon

hyp E hyp D hyp C hyp B hyp A

Hydrogenase 3

hyc A hyc B hyc C hyc D hyc E hyc F hyc G hyc H

Small subunit sequence Large subunit sequence Ferredoxin-like sequence

Fig. 7 Outline structures for operons of *E. coli* nickel-containing hydrogenases and for the *hyp* operon, involved in activation of all three hydrogenases.

identified that occur in all 15 hydrogenases: R-Xaa-C-Xaa_2-C at the N-terminus and D-P-C-Xaa_2-C at the C-terminus [83,107]. Mutagenesis of residues in both motifs in the large subunit of hydrogenase-1 from *E. coli* confirms that these sites are essential for catalytic activity of hydrogenase-1 [83,98].

Studies concerning the biosynthesis of the three NiFe-hydrogenases has started only recently. Although additional genes present in each operon have been identified, studies to establish their functioning are still in progress [83]. Additional genes that are present in the *hya* operon are conserved in NiFe-hydrogenase operons from other bacteria [97].

In the hydrogenase-1 encoding *hya* operon, *hyaC* codes for a hydrofobic protein with four putative membrane-spanning regions. It has been suggested that HyaC is a membrane-bound protein that is involved in membrane anchoring of the hydrogenase [97]. Genes homologous with *hyaC* have been found in *Rh. capsulatus* [99], *Wolinella succinogenes* [100], *A. vinelandii* [92]. *R. leguminosarum* [101] and *A. eutrophus* [102]. Deletions in *hyaC* do not affect the benzyl-viologen linked hydrogenase activity and both subunits of the hydrogenase are still membrane-bound, but hydrogenase-1 in these mutants shows multiple forms with different mobilities in non-denaturing gels [103]. These results suggest that HyaC increases the stability of the complex [103]. More insight in what may be the function of HyaC was derived from studies of the

homologous counterparts of HyaC in *Rh. capsulatus* and *W. succinogenes*, HupM and HydC [99,100]. Deletion of the *hupM* gene in *Rh. capsulatus* did not affect the in vitro viologen-linked hydrogenase activity, but in vivo electrons released by H_2 oxidation were not transferred to the respiratory chain [99]. In *W. succinogenes* , the homologous counterpart of HyaC, HydC, was identified as a cytochrome-*b* involved in electron transfer between hydrogenase and menaquinone and forms the third subunit of the enzyme complex that was co-purified with the large and small subunits after Triton X-100 extraction [100].

Deletion of other additional genes of the *hya* operon revealed that the gene products of *hyaD* and *hyaE* are involved in the processing of the nascent hydrogenase-1 structural subunits and that *hyaF* enhances nickel incorporation into the nascent enzyme [103]. These mutational studies clearly demonstrated that wildtype levels of hydrogenase were only reached when all six genes of the *hya* operon were present, indicating that all genes are functional for the synthesis of fully active hydrogenase-1, but only *hyaA* (encoding the large subunit) and *hyaE* are absolutely required for hydrogenase that has in vitro viologen-linked activity [103].

Additional proteins encoded by the *hyb* operon, *hybB* and *hybD-hybG* have no homology with proteins encoded in the *hya* operon with one exception [96]. The protein encoded by *hybB*, is similar to the *hyaC* gene product: both are very hydrofobic and might contain several membrane-spanning helices [96]. Mutational analysis of the *hyb* operon has not yet been performed.

Like the additonal genes of the *hyb* operon, the seven additional genes of the *hyc* operon have no homology with those present in the *hya* operon [83]. However, two genes, *hycB* and *hycF*, code for proteins rich in cysteinyl residues in C-Xaa_2-C-Xaa_2-C-Xaa_3-C motifs [92], which indicates the presence of [4Fe-4S] clusters of the ferredoxin type [104]. Therefore, they were proposed to serve as electron transfer proteins. The gene products of *hycC* and *hycD* are homologous with subunits (chain 4 and 1, respectively) of eukaryotic mitochondrial NADH-ubiquinone oxidoreductases, while the product of *hycG* is somehow similar to the protein G from the chloroplast redox chain of *Zea mais* [106].

Recently, in frame deletions were introduced into each of the eight genes of the *hyc* operon of *E. coli* , [105]. Since hydrogenase-3 is part of the formate hydrogen lyase (FHL_H) system, both the H_2-dependent benzyl viologen reduction (a measure for total hydrogenase activity) and the formate-dependent benzyl viologen reduction (which is a measure of formate dehydrogenase (FDH_H) activity) were analyzed in cell-free extracts of the mutants. These mutants were grown under fermentative conditions which favour the expression of genes of the FHL_H, while hydrogenases-1 and -2 are only produced in minor amounts. From these experiments four observations were made:

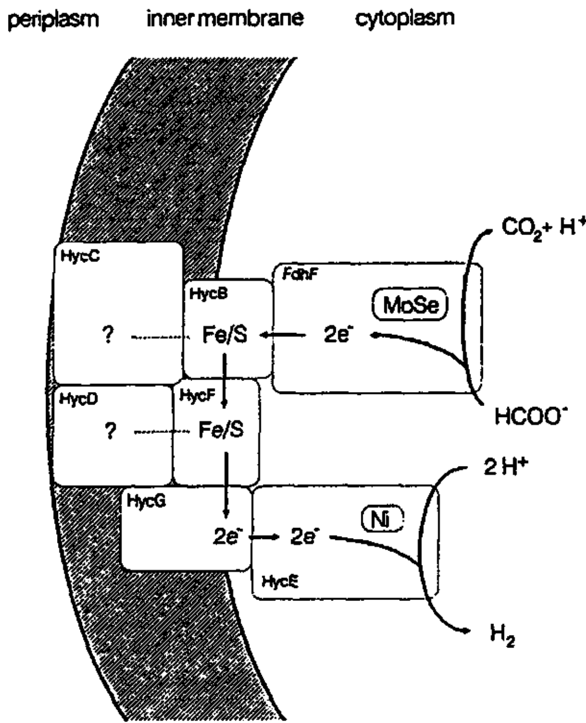


Fig. 8 Schematic model depicting the arrangement of the formate hydrogenlyase structural components and the possible path of electron flow from formate to H^+ [105]. FdhF is formate dehydrogenase that is encoded by a different operon.

1. all genes, except *hycA*, were required for total hydrogenase activity;
2. all genes, except *hycA*, *hycC* and *hycD*, were required for FDH_H activity;
3. non of the deletions in the *hyc* operon affected hydrogenase-1 or -2 activity;
4. deletion of *hycA* results in increased total hydrogenase and FDH_H activity

The fourth observation suggests a regulatory role for the *hycA* gene. This suggestion was substantiated by the observation that overproduction of HycA nearly abolished both activities. Immunoblotting demonstrated that the structural proteins reponsible for both activities, HycE (hydrogenase) and FdhF (FDH_H),were strongly reduced in the *hycA* overproducing strain [105].

The effect of *hyc* deletions on FDH_H activity is indirect. Therefore, it was proposed that the reduced FDH_H activity may be the result of a lack of attachment of formate dehydrogenase to the membrane, which decreases the efficiency of electron flow to the artifical acceptor, benzyl viologen. The attachment could be accomplished by the gene

products of *hycB*, *hycC*, *hycD*, *hycF* and *hycG*, since they were found to be membrane-bound proteins [105]. Based on the knowledge gathered about the hydrogenase-3 and the FHL complex, a model is postulated as a working hypothesis (Fig. 8) [105]. In this model HycB,C,D,F and G function as carrier proteins attached to the membrane that transport electrons produced in formate oxidation to the H_2-evolving hydrogenase-3. HycH is not present in this model. It was suggested from the different mobility on SDS gels of hydrogenase-3 from *hycH* mutants that HycH is involved in the processing of hydrogenase-3 [105].

Besides the three operons that contain the genes for hydrogenases-1, -2, and -3 (*hya*, *hyb* and *hyc*, respectively) a fourth operon was identified that is necessary for the activity of all three hydrogenases and contains five genes (Fig. 7) [83,97,106,107] . Mutations in this operon affect the activity of all three *E. coli* hydrogenases in a pleiotropic manner and hence the operon was designated *hyp* [108-114]. Characterization of the *hyp* operon was performed by sequencing and complementation analysis of *hyp* deletion mutants [107]. From these experiments it was indicated that *hypA* is a regulatory gene, that *hypB* has a role in nickel sequestering or cofactor synthesis and that *hypCD* are involved in the post-translational modification of hydrogenase-1 and -2.

V.I.II. NIFE HYDROGENASES IN *A. EUTROPHUS*

NiFe-hydrogenase acivities in *A. eutrophus*. *Alcaligenes eutrophus* strain H16, a gram-negative CO_2 assimilating, lithoautotrophic aerobic bacterium, contains two NiFe-hydrogenases [115]. Contrary to the *E. coli* hydrogenases, the two hydrogenases from *A. eutrophus* are easily isolated in large amounts, allowing biochemical and spectroscopic studies of the wild-type enzymes and of enzymes produced in strains that have deletions in one or more of the additional genes involved in biosynthesis of these hydrogenases [116-118].

Composition of the NiFe-hydrogenases in *A. eutrophus*. The first enzyme was characterized as a cytoplasmic NAD-reducing NiFe-hydrogenase that contains FMN and is composed of four subunits, which form a heterotetrameric $\alpha\beta\gamma\delta$ complex. The NAD-reducing hydrogenase catalyzes reduction of NAD^+ by H_2 [117]. The complex dissociates into 2 dimers , the $\beta\delta$ or 'hydrogenase' dimer and the $\alpha\gamma$ or 'diaphorase' dimer, as a response to low nickel concentrations, which leads to the suggestion that one of the two nickel atoms present in this enzyme may play a role in the maintenance of its tetrameric structure [116]. The hydrogenase complex contains 2 Ni atoms/mol, 1 FMN/mol and 12 iron plus sulfide/mol that are organized in [4Fe-4S], [3Fe-4S] and [2Fe-2S] non-heme iron clusters [117]. The 'hydrogenase' dimer, composed of a large β (56 kDa) and small δ (26 kDa) subunit that are homologous with subunits in other NiFe-

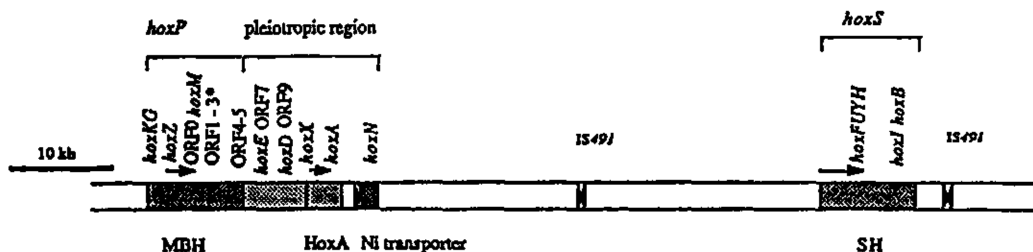


Fig. 9 Structure of the hox gene clusters on megaplasmid pHG1 of *A. eutrophus* H16 [102].

hydrogenases, contains, besides the catalytically active nickel, the non-heme irons and it catalyses hydrogenase activity with artifical electron acceptors. The 'diaphorase' dimer is composed of a large α (63 kDa) and a small γ (30 kDa) subunit. These subunits are homologous with the 51 and 24 kDa subunits, respectively, of mitochondrial NADH-ubiquinone reductase. The 'diaphorase' dimer contains FMN, one [2Fe-2S] and two [4Fe-4S] clusters [119].

The second NiFe-hydrogenase is a heterodimeric membrane-bound uptake hydrogenase that is coupled to the electron transport chain via an unknown acceptor molecule, generating ATP [118]. The enzyme is composed of a small (31 kDa) and a large (65 kDa) subunit that are highly (more than 80%) homologous with the corresponding polypeptides from *B. japonicum* [93] and *R. capsulatus* [94].

Biosynthesis of the NiFe-hydrogenases in *A. eutrophus*. The genes involved in the biosynthesis of both hydrogen oxidizing enzymes from *A. eutrophus* H16, are located on a 100 kilobase DNA region on the 450 kilobase conjugative plasmid pGH1 (Fig. 9) [120]. The *hoxP* locus contains genes for the membrane-bound hydrogenase and the *hoxS* locus for the soluble NAD-reducing hydrogenase. Besides this, the *hoxC* or the pleiotropic region has been identified, containing genes required for synthesis of both hydrogenases.

The *hoxS* cluster contains four structural genes, *hoxF,U, Y,* and *H*, that code for the α, γ, δ and β subunits, respectively, of the NAD-reducing hydrogenase [120]. In addition to this, the *hoxS* locus contains two distal localized genes, *hoxJ* and *hoxB*. Mutations in these two genes abolish NAD-linked hydrogenase activity, but their function remains unclear [102].

The *hoxP* locus contains besides the structural genes, *hoxK* and *G*, that code for the small and large subunits of the membrane-bound hydrogenase, respectively, at least four

additional open reading frames, *hoxZ, hoxM*, *orf2* and *orf3*, which are homologous with the helper genes that are present in the *hya* operon encoding hydrogenase-1 in *E. coli* [102].

The 'pleiotropic' or *hoxC* region contains genes that are required for the biosynthesis of both hydrogenases: *hoxA, hoxE, hoxD*, *orf7, orf9*, *hoxX* and *hoxN* [102]. Expression of both hydrogenases is strictly dependent on the *hoxA* gene, which responds to the physiological signals energy limitation [121] and temperature [122]. Sequence analysis further confirmed that HoxA is a member of the family of regulator proteins that respond to external stimuli [102,123]. It was further suggested that *hoxX* might be involved in the signal respons of *hoxA*, although this has not been confirmed by experimental evidence [102]. Genes *hoxE, hoxD, orf7* and *orf9* display a striking homology with the genes of the pleiotropic *E. coli hyp* operon [107]. Mutants carrying deleted *hoxED* genes produce inactive hydrogenases [124]. The *hoxD* mutants exhibit a decreased nickel accumulation; in these mutants no Ni was incorporated in both hydrogenases [125]. Mutation of *hoxN* results in a reduced activity of all Ni containing enzymes (e. g. hydrogenases and urease), and can be complemented by high levels of nickel in the culture medium. In these mutants, nickel uptake is mediated by the non-specific high-capacity magnesium transport system [126]. Sequence analysis of the *hoxN* gene revealed that it encodes an extremely hydrofobic protein. It was shown that HoxN is an intrinsic cytoplasmic-membrane protein that functions as a high-affinity nickel transporter [127,128].

V.II. FE-HYDROGENASES

Fe-hydrogenase activities. Contrary to the widespread distribution and the extensive knowledge of the molecular biology of the NiFe-hydrogenases, the Fe-hydrogenases are so far only detected in some strictly anaerobic bacteria. Because of this, only little is known about their molecular biology [82]. Five Fe-hydrogenases have been characterized so far: from *Megasphaera elsdenii* [129-132], *Desulfovibrio vulgaris* (Hildenborough) [133-137], *Desulfovibrio desulfuricans* (ATCC 7757) [137a,137b] and two from *Clostridium pasteurianum* [138-143].

Composition of the Fe-hydrogenases. The characteristics of the five hydrogenases are summarized in Table 1. All contain two or more ferredoxin-like [4Fe-4S] clusters (F-clusters) that are proposed to act in electron transfer. It has been proposed, first for the *D. vulgaris* enzyme [136], later also for the clostridial enzymes [82] that the active site might contain a [6Fe-6S] cluster (H-cluster).

The *D. vulgaris* and the *D. desulfuricans* Fe-hydrogenases are both heterodimeric enzymes that are located in the periplasm [136,137a,137b,144,145]. The large or α

Table 1 Properties of iron-hydrogenases [62,137b]

	Organism				
	Clostridium pasteurianum		*Megasphaera elsdenii*	*Desulfovibrio vulgaris* (Hildenb.)	*Desulfovibrio desulfuricans* (ATCC 7757)
	I	II			
Location:	cytoplasmic	cytoplasmic	cytoplasmic	periplasmic	periplasmic
Purification conditions:	anaerobic	anaerobic	anaerobic	aerobic	aerobic
O_2 sensitivity as isolated:	extremely	extremely	extremely	insensitive	insensitive
Molecular weight:	61944[a]	54738[a]	58000[b]	45820+10000[c]	42500+11000[b]
V_m, H_2 evolution[d]:	5500	10	7000	10400	8200
V_m, H_2 oxidation[e]:	24000	34000	9000	50000	62000
Fe/mole	20.1±0.7	13.8±0.4	15.6±2.7	9-15	14.0
S^{2-}/mole	17.8±1.2	11.4±0.2	15.5±2.4	≈13	14.4

[a] Based on amino acid composition.

[b] Based on sedimentation equilibrium analysis.

[c] Based on amino acid (gene) sequence.

[d] Expressed as µmol of H_2 evolved/min per mg using dithionite-reduced methyl viologen as electron donor

[e] Expressed as µmol of H_2 evolved/min per mg using methylene blue or benzyl viologen as electron donor

subunit (46 kDa) of the *D. vulgaris* enzyme is proposed to contain the three iron sulfur clusters; the two F-clusters are ligated to cysteines that occur in a motif also found in 2[4Fe-4S] ferredoxins and there is evidence that also the H-cluster is ligated to residues in the large subunit [145,146]. The small or β subunit is translated as a 13 kDa precursor and contains a presequence that is cleaved upon export across the cytoplasmic membrane [147]. The α subunit lacks a leader sequence for export and therefore it was suggested that the β subunit is involved in the export of both hydrogenase subunits across the membrane [148].

Biosynthesis of the Fe-hydrogenases. Although *D. vulgaris* Fe-hydrogenase was the first hydrogenase that was cloned and sequenced, further study of its biosynthesis has been hampered by the poor knowledge of the molecular biology of anaerobes.

Indications that additional genes may be required for assembly, export and metal incorporation were derived from studies in which the genes encoding the subunits of the Fe-hydrogenase were expressed in *E. coli*. The recombinant hydrogenase purified from these *E. coli* cells was inactive (< 0.1 U/mg protein) [149,150]. EPR studies of the purified recombinant enzyme indicated that the two ferredoxin-like F-clusters were present in the protein, but the third H-cluster was not assembled at all [149]. Moreover, subcellular localisation studies showed that in spite of the production of the small β subunit, which harbors the signal peptide for export, the major part (about 80%) of the overproduced hydrogenase subunits resides in the membrane fraction and only a minor part (about 20%) is transported to the periplasm and assembled into an αβ dimer in *E. coli* [148,150].

Overproduction of the subunits of the Fe-hydrogenase in its natural host, *D. vulgaris*, did not result in a proportional increase in hydrogenase activity in these cells, and resulted in accumulation of the precursor of the small subunit, and incomplete export of the subunits [151]. A similar observation was made when hydrogenase-1 from *E. coli* was overexpressed in *E. coli* cells. These results indicate that export, assembly and metal incorporation of the Fe-hydrogenase might require additional factors that are not provided by *E. coli* and occur in limiting amounts in *D. vulgaris*.

VI. CONCLUSIONS

Only recently it is realized, that biosynthesis of metal proteins requires expression of a set of genes, not only genes for the structural subunits of the proteins, but also genes encoding helper proteins involved in regulation, metal uptake, its incorporation and processing of the enzymes. Until now, evidence for such systems has been obtained primarily by genetic methods: construction and analysis of mutants impaired in

biosynthesis. Only a few reports exist about the in vitro reconstitution of part of a biosynthetic pathway with purified proteins (e. g. incorporation of FeMo cofactor in nitrogenase component 1 by component 2)

Genetic analysis of the biosynthetic pathways of several enzymes presented above, is still in a preliminary state. However, some conclusions can be drawn from the above mentioned experiments with the different systems:

1. Biosynthesis of homologous enzymes occurs with a similar set of helper genes. For example, the nitrogenase gene clusters of *Klebsiella* and *Azotobacter* encode similar proteins, as do hydrogenase gene clusters in *Alcaligenes* and *E. coli* (*hoxP* and *hya*, *hoxC* and *hyp*).

2. Until now, no similarity has been found between genes involved in biosynthesis of different proteins (e.g. between genes involved in nitrogenase, hydrogenase and N_2O-reductase biosynthesis). However, it is to early to state that similar principles underlying biosynthesis of different metalloenzymes do not occur.

3. Incorporation of a metal-cofactor in an enzyme is in some cases possible without an 'insertase' (e.g. Mo-cofactor in nitrate reductase, but also [2Fe-2S] and [4Fe-4S] clusters in ferredoxins, that can be incorporated both chemically and with rhodanese), but there is strong evidence that insertases are required for e.g. N_2O-reductase, hydrogenase and nitrogenase.

4. There is clearly lack of information on the effect of deletion of a gene for a helper protein on the molecular properties of the (apo-)enzyme. This requires purification of the (inactive) enzyme. A difficulty in this respect is that deletion of helper proteins often results in (inactive) enzyme with strongly declined stability, difficult to purify.

5. Construction of an in vitro system for reconstitution of the biosynthesis pathway of redox active metal proteins, starting from the apoproteins, will be a formidable task, especially when membrane transport or membrane-localized components are involved in the activation (hydrogenase, N_2O-reductase). It should be realized that such a system might also require proteins with a more general function (e.g. chaperonins).

VII. OUTLINE OF THIS THESIS

The very first hydrogenase genes isolated were those for the periplasmic Fe-hydrogenase of *D. vulgaris* [144,145], which is one of the two most active hydrogenases known till now [82,137b]. Very soon after isolation of the genes for the subunits of this enzyme, it became apparent that expression of these genes in *E. coli* under a variety of conditions (aerobe/anaerobe, with excess iron in the culture medium, et cetera) did not result in the production of active enzyme. Only a very small amount of these subunits were assembled into an αβ dimer. This dimer contained the two ferredoxin-like [4Fe-4S] clusters involved in electron transport, but not the active-site FeS cluster [149]. Moreover, the majority of the subunits produced was not exported to the periplasm [148]. These observations resulted in the hypothesis that helper genes are required encoding proteins for export of the subunits across the cytoplasmic membrane and insertion of the active-site metal cluster.

In order to verify this hypothesis, and for a first analysis of the factors that could be involved in biosynthesis of active Fe-hydrogenase, a genetic approach was chosen: the construction of *D. vulgaris* mutants (by transposon insertion or marker exchange) with impaired hydrogenase biosynthesis, followed by analysis of the genes involved. As genes serving a single pathway are often clustered in the genome, this project was started by isolation of large DNA fragments surrounding the genes for the hydrogenase subunits. Surprisingly, on one such a fragment, a gene was found encoding a polypeptide with homology in primary structure to the α and β subunits of the hydrogenase (**chapter 2**), and it was speculated that this gene encodes either an alternative Fe-only hydrogenase or a helper protein, involved in processing of the hydrogenase. Subsequently, it was shown that this gene also contained homology with the subunits of NADH-ubiquinone reductase, with subunits of a NiFe-hydrogenase and with another Fe-hydrogenase. The implications of this are discussed in **chapter 3**.

In the mean time, experiments were started for establishing a system for genetic manipulation of *Desulfovibrio* (including marker exchange and/or transposon mutagenesis), which was not available at that moment for bacteria belonging to the delta subdivision of purple bacteria. Although this project resulted in development of a system for introduction of plasmids into *Desulfovibrio* [151], which has been applied subsequently for the overproduction of the [6Fe-6S] prismane protein (**chapter 6**), it has not yet been possible to develop this system further for directed mutagenesis of genes in the *Desulfovibrio* genome by marker exchange. One of the main reasons seems to be the high endogenous resistance of *Desulfovibrio* against many of the common antibiotics, which are required for the selection of mutants. The numerous laborous experiments on this subject are not described in this thesis.

Lack of a system for transposon mutagenesis or marker exchange has prevented further analysis of genes required for biosynthesis of the Fe-hydrogenase in *D. vulgaris*. Subsequently, similar methods have been applied for the analysis of the biosynthetic pathway of NiFe-hydrogenases in easier-to-manipulate bacteria than *Desulfovibrio*. This indeed indicated that at least for the NiFe-hydrogenases several proteins, encoded by clusters of genes, were required for biosynthesis of active enzyme. Results obtained on the investigation of a biosynthetic pathway of the NiFe-hydrogenases are discussed earlier in this Introduction.

Indications have been obtained that the Fe-hydrogenase may contain a [6Fe-6S] active-site cluster [82,136,152]. Such clusters had not been described previously in enzymes. Clearer indications for the existence of such novel Fe-S clusters in proteins were obtained after the isolation and spectroscopic characterization of the 'prismane protein' from *D. vulgaris* [153-155]. The low iron content (approximately six irons/mol protein, i. e. only sufficient for one [6Fe-6S] cluster) and the stability under aerobic conditions made this protein a model protein for studying the redox properties of these complex Fe-S clusters [155] and prompted to try to crystallize the protein in order to obtain information about structure and ligandation of the cluster. Therefore, the primary structure of the protein is required. The isolation of the gene for the prismane protein and the determination of the amino acid sequence from the nucleotide sequence are described in **chapter 4**. In the mean time, Moura et al. [160] had isolated a similar protein from another *Desulfovibrio* strain, *D. desulfuricans* ATCC 27774. Although the EPR spectra of the *D. desulfuricans* protein in the one- and three-electron-reduced states were quite similar to that of the *D. vulgaris* protein, there was a difference in some details of the spectra. Moreover, interpretation of EPR and Mössbauer spectra led Moura et al. [160], to assume two [6Fe-6S] clusters in their protein, while strong evidence for only one such a cluster existed for the *D. vulgaris* prismane protein (Pierik et al. [159]). To solve this controversy, it would be very helpful to know, wether both proteins really were homologous or not. Therefore, the gene encoding the *D. desulfuricans* protein was isolated and its primary structure was shown to be very similar to that of the *D. vulgaris* prismane protein (**chapter 5**).

The prismane protein is produced only in low amounts in *D. vulgaris*. In order to facilitate purification of large amounts of protein for spectroscopy and crystallization, overproduction of the protein in *D. vulgaris* was attempted by application of the previously described cloning procedure for *D. vulgaris* (**chapter 6**). Experiments with the overproduced *D. vulgaris* prismane protein gave some insight in the nature of the differences that exist in the EPR spectra of *D. vulgaris* and *D. desulfuricans* proteins (**chapter 6**).

VIII. REFERENCES

1 Harrison, P. M. (ed.) (1985) *Metalloproteins* part 1, Verlag Chemie, Weinheim.

2 Bonomi, F., Pagani, S. and Kurtz, D. M., Jr (1985) *Eur. J. Biochem.* **148**, 67-73.

3 Pagani, S., Bonomi, F. and Cerletti, P. (1984) *Eur. J. Biochem.* **142**, 361-366.

4 Pagani, S., Eldridge, M. and Eady, R. R. (1987) *Biochem. J.* **244**, 485-488.

5 Filler, W. A., Kemp, R. M., NG, J. C., Hawkes, T. R., Dixon, R. A. and Smith, B. E. (1986) *Eur. J. Biochem.* **160**, 371-377.

6 Paustian, T. D., Shah, V. K. and Roberts, G. P. (1990) *Biochemistry* **29**, 3515-3522.

7 Smith, B. E. and Eady, R. R. (1992) *Eur. J. Biochem.* **205**, 1-15.

8 Robson, R. L. (1991) in *Biology and biochemistry of nitrogen fixation* (Dilworth, M. J. and Glenn, A. R., eds) pp. 142-161, Elsevier, Amsterdam, New York.

9 Thorneley, R. N. F. and Deistung, J. (1988) *Biochem. J.* **253**, 587-595.

10 Thorneley, R. N. F. and Lowe, D. J.. (1985) in *Molybdenum Enzymes,* Ch. 5, ed. Spiro, T. G. (Wiley), pp. 221-284.

11 Hageman, R. V. and Burris, R. H. (1979) *J. Biol. Chem.* **254**, 11189-11192.

12 Bergstrom, J., Eady, R. R. and Thorneley, R. N. F. (1988) *Biochem. J.* **251**, 165-169.

13 Howard, J. B., Davis, R., Moldenhauer, B., Cash, V. L. and Dean, D. (1989) *J. Biol. Chem.* **264**, 11270-11274.

14 Georgiadis, M. M., Chakrabarti, P. and Rees, D. C. (1990) in *Nitrogen fixation: achievements and objectives* (Gresshof, P. M., Roth, E. L., Stacey, G. and Newton, W. E., eds) pp. 111-116, Chapman and Hall, New York, London.

15 Smith, B. E. (1990) in *Nitrogen fixation: achievements and objectives* (Gresshof, P. M., Roth, E. L., Stacey, G. and Newton, W. E., eds) pp. 3-13, Chapman and Hall, New York, London.

16 Georgiadis, M. M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J. J. and Rees, D. C. (1992) *Science* **257**, 1653-1659.

17 Eady, R. R. (1990) in *Vanadium in biological systems* (Chasteen, N. E., ed.) pp 99-127, Kluwer Academic Publishers, Dordrecht.

18 Kim, J. and Rees, D. C. (1992) *Science* **257**, 1677-1682.

19 Hagen, W. R., Wassink, H., Eady, R. R., Smith, B. E. and Haaker, H. (1987) *Eur. J. Biochem.* **169**, 457-465.

20 Hoover, T. R., Imperial, J., Ludden, P. W. and Shah, V. K. (1988) *Biofactors* **1**, 199-206.

21 Arnold, W., Rump, A., Klipp, W., Priefer, U. B. and Pühler, A. (1988) *J. Mol. Biol.* **203**, 715-738.

22 Kennedy, C. and Dean, D. (1992) *Mol. Gen. Genet.* **231**, 494-498.

23 Joerger, R. D. and Bishop, P. E. (1988) *J. Bacteriol.* **170**, 1475-1487.

24 Bishop, P. E., MacDougal, S. I., Wolfinger, E. D. and Shermer, C. L. (1990) in *Nitrogen fixation: achievements and objectives* (Gresshof, P. M., Roth, E. L., Stacey, G. and Newton, W. E., eds) pp. 789-795, Chapman and Hall, New York, London.

25 Bishop, P. E. and Joerger, R. D. (1990) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **41**,109-125.

26 Kennedy, C. (1989) in *Genetics of bacterial diversity* (Hopwood, D. A. and Chater, K. F., eds.) pp. 107-127, Academic press, London.

27 Merrick, M. J. (1988) in *The nitrogen and sulphur cycles* (Cole, J. A. and Ferguson, S., eds.) pp. 331-361, Cambridge University Press, Cambridge.

28 Ditta, G., Virts, E., Palomares, A. and Choong-Hyun, K. (1987) *J. Bacteriol.* **169**, 3217-3223.

29 Morett, E., Fischer, H. M. and Hennecke, H. (1991) *J. Bacteriol.* **173**, 3478-3487.

30 Jacobson, M. R., Cash, V. L., Weiss, M. C., Laird, N. F., Newton, W. E. and Dean, D. R. (1989) *Mol. Gen. Genet.* **219**, 49-57.

31 Brigle, K. E., Weiss, M. C., Newton, W. E. and Dean, D. R. (1987) *J. Bacteriol.* **169**, 1547-1553.

32 Howard, K. S., McLean, P. A., Hansen, F. B., Lemley, P. V., Koblan, K. S., Orme-Johnson, W. H. (1986) *J. Biol. Chem.* **261**, 772-778.

33 Harris, G. S., White, T. C., Flory, J. E. and Orme-Johnson, W. H. (1990) *J. Biol. Chem.* **265**, 15909-15919.

34 Robinson, A. C., Chun, T. W., Li, J. and Burgess, B. K. (1989) *J. Biol. Chem.* **264**, 10088-10095.

35 Robinson, A. C., Dean, D. R. and Burgess, B. K. (1987) *J. Biol. Chem.* **262**, 14327-14332.

36 Paustian, T. D., Shah, V. K. and Roberts, G. P. (1990) *Biochemistry* **29**, 3515-3522.

37 Hoover, T. R., Imperial, J., Ludden, P. W. and Shah, V. K. (1988) *J. Bacteriol.* **170**, 1978-1979.

38 Shah, V. K., Hoover, T. R., Imperial, J., Paustian, T. D., Roberts, G. P. and Ludden, P. W. (1988) in *Nitrogen fixation: hunderd years after* (Bothe, H., de Bruijn, F. J. and Newton, W. E., eds) pp. 115-120, Gustav Fischer, Stuttgart, New York.

39 Dean, D. R. and Jacobson, M. R. (1992) in *Biological nitrogen fixation* (Stacey, G., Evans, B. J. and Burris, R. H., eds) pp. 763-834, Chapman and Hall, New York.

40 Imperial, J., Ugalde, R. A., Shah, V. K. and Brill, W. J. (1984) *J. Bacteriol.* **158**,187-194.

41 Tal, S., Chun, T. W., Gavini, N. and Burgess, B. K. (1991) *J. Biol. Chem.* **266**, 10654-10657.

42 Shah, V. K., Ugalde, R. A., Imperial, J. and Brill, W. J. (1985) *J. Biol. Chem.* **260**, 3891-3894.

43 Govezensky, D., Greener, T., Segal, G. and Zamir, A. (1991) *J. Bacteriol.* **173**, 6339-6346.

44 Stewart, V. (1988) *Microbiol. Rev.* **52**, 190-232.

45 Iobbi, C., Santini, C. -L., Bonnefoy, V. and Giordano, G. (1987) *Eur. J. Biochem.* **168**, 451-479.

46 Wallace, B. J. and Young, I. J. (1977) *Biochim. Biophys. Acta* **461**, 84-100.

47 Chaudhry, G. R. and Mac Gregor, C. H. (1983) *Biol. Chem.* **258**, 5819-5827.

48 Blasco, F., Iobbi, C., Giordano, G., Chippaux, M. and Bonnefoy, V (1989) *Mol. Gen. Genet.* **218**, 249-256.

49 Johnson, M. K., Bennett, D. E., Morningstar, J. E., Adams, M. W. W. and Mortenson, L. E. (1985) *J. Biol. Chem.* **260**, 5456-5463.

50 Chaudry, G. R. and Mac Gregor, C. H. (1983) *J. Biol. Chem.* **258**, 5819-5827.

51 Blasco, F., Iobbi, C., Ratouchniak, J., Bonnefoy, V. and Chippaux, M. (1990) *Mol. Gen. Genet.* **222**, 104-111.

52 Showe, M. K. and DeMoss, J. A. (1968) *J. Bacteriol.* **95**, 1305-1313.

53 Iobbi-Nivol, C., Santini, C., Blasco, F. and Giordano, G. (1990) *Eur. J. Biochem.* **188**, 679-687.

54 Stewart, V. and Berg, B. (1988) *J. Bacteriol.* **170**, 4437-4444.

55 DeMoss, J. A. and Hsu, P. -Y. (1991) *J. Bacteriol.* **173**, 3303-3310.

56 Li, S and DeMoss, J. A. (1987) *J. Bacteriol.* **169**, 4614-4620.

57 Li, S., Rabi, T. and DeMoss, J. A. (1985) *J. Bacteriol.* **164**, 25-32.

58 Collins, L. A., Egan, S. M. and Stewart, V. (1992) *J. Bacteriol.* 174, 3667-3675.

59 Stewart, V., Parales, J. and Merkel, S. M. (1989) *J. Bacteriol.* **171**, 2229-2234.

60 Chiang, R. C., Cavicchioli and Gunsalus, R. P. (1992) *Mol. Microbiol.* **6**, 1913-1923.

61 MacGregor, C. H. (1976) *J. Bacteriol.* **126**, 122-131.

62 Dubourdieu, M. and DeMoss, J. A. (1992) *J. Bacteriol.* **174**, 867-872.

63 Blasco, F., Pommier, J., Augier, V., Chippaux, M. and Giordano, G. (1992) *Mol. Microbiol.* **6**, 221-230.

64 Johnson, M. K., Rajagopalan, K. V. (1987) *J. Bacteriol.* **169**, 117-125.

65 Johnson, J. L., Indermaur, L. W. and Rajagopalan, K. V. (1991) *J. Biol. Chem.* **266**, 12140-12145.

66 Johnson J. L., Jones, H. P. and Rajagopalan, K. V. (1977) *J. Biol. Chem.* **252**, 4994-5003.

67 Rajagopalan, K. V. and Johnson, J. L. (1992) *J. Biol. Chem.* **267**, 10199-10202.

68 Körner, H. and Zumft, W. G. (1989) *Appl. Environ. Microbiol.* **55**, 1670-1676.

69 Zumft, W. G. and Kroneck, P. M. H. (1990) *FEMS Symp.* **56**, 37-55.

70 Körner, H. and Zumft, W. G. (1989) *Appl. Environ. Microbiol.* **55**, 1670-1676.

71 Kroneck, P. M. H., Antholine, W. A., Riester, J. and Zumft, W. G. (1988) *FEBS Lett.* **242**, 70-74.

72 Viebrock, A. and Zumft, W. G. (1988) *J. Bacteriol.* **170**, 4658-4668.

73 Scott, R. A., Zumft, W. G., Coyle, C. L. and Dooley, D. M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4082-4086.

74 Braun, C. and Zumft, W. G. (1992) *J. Bacteriol.* **174**, 2394-2397.

75 Zumft, W. G., Viebrock-Sambale, A. and Braun, C. (1990) *Eur. J. Biochem.* **192**, 591-599.

76 Riester, J., Zumft, W. G. and Kroneck, P. M. H. (1989) *Eur. J. Biochem.* **178**, 751-762.

77 Higgins, C. F., Callagher, M. P., Mimmack, M. L. and Pearce, S. R. (1988) *BioEssays* **8**, 111-116.

78 Hiles, I. D., Gallagher, M. P., Jamieson, D. J. and Higgins, C. F. (1987) *J. Mol. Biol.* **195**, 125-142.

79 Ames, G. F. -L. (1986) *Annu. Rev. Biochem.* **55**, 397-425.

81 Mokhele, K., Tang, Y. J., Clark, M. A. and Ingraham, J. L. (1987) *J. Bacteriol.* **169**, 5721-5726.

81 Lee, H. S., Hancock, R. E. W. and Ingraham, J. L. (1989) *J. Bacteriol.* **171**, 2096-2100.

82 Adams, M. W. W. (1990) *Biochim. Biophys. Acta* **1020**, 115-145.

83 Przybyla, A. E., Robbins, J., Menon, N. and Peck, H. D. Jr. (1992) *FEMS Microbiol. Rev.* **88**, 109-136.

84 LeGall, J., Ljungdahl, P. O., Moura, I., Peck, H. D. Jr., Xavier, A. V., Moura, J. J. G., Teixeira, M., Huynh, B. -H. and DerVartanian, D. V. (1982) *Biochem. Biophys. Res. Commun.* **106**, 610-616.

85 Fauque, G., Peck, H. D. Jr., Moura, J. J. G., Huynh, B. -H., Berlier, Y., DerVartanian, D. V., Teixeira, M., Przybyla, A. E., Lespinat, P. A., Moura, I. and LeGall, J. (1988) *FEMS Microbiol. Rev.* **54**, 299-344.

86 Sawers, R. G., Ballantine, S. P. and Boxer, D. H. (1985) *J. Bacteriol.* **164**, 1324-1331.

87 Sawers, R. G. and Boxer, D. H. (1986) *Eur. J. Biochem.* **156**, 256-275.

88 Ballantine, S. P. and Boxer, D. H. (1986) *Eur. J. Biochem.* **156**, 277-284.

89 Ballantine, S. P. and Boxer, D. H. (1985) J. Bacteriol. 163, 454-459.

90 Peck, H. D. and Gest, H. (1957) *J. Bacteriol.* **73**, 706-721.

91 Ford, C. M., Garg, N., Garg, R. P., Tibelius, K. H., Yates, M. G., Arp, D. J. and Seefeldt, L. C. (1990) *Mol. Microbiol.* **4**, 999-1009.

92 Menon, A. L., Stults, L. W., Robson, R. L. and Mortenson, L. E. (1990) *Gene* **96**, 67-74.

93 Sayavedra-Soto, L. A., Powell, G. K., Evans, H. J. and Morris, R. O. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8395-8399.

94 Leclerc, M., Colbeau, A., Cauvin, B. and Vignais, P. M. (1988) *Mol. Gen. Genet.* **214**, 97-107 [erratum (1989) **215**, 368].

95 Uffen, R. L., Colbeau, A., Richaud, P. and Vignais, P. M. (1990) *Mol. Gen. Genet.* **221**, 49-58.

96 Przybyla, A. E., Menon, N. K., Robbins, J., DerVartanian, L. and Peck, H. D. Jr. (1991) in *Abstracts of the 3rd International Conference on Molecular Biology of Hydrogenases* pp. 18-22, Troia, Portugal.

97 Menon, N. K., Robbins, J., Peck, H. D. Jr., Chatelus, C. Y., Choi, E. -S. and Przybyla, A. E. (1990) *J. Bacteriol.* **172**, 1969-1977.

98 Peck, H. D. Jr., Robbins, J., Menon, N. K. and Przybyla, A. E. (1991) in *Abstracts of the 3rd International Conference on Molecular Biology of Hydrogenases* pp. 213, Troia, Portugal.

99 Cauvin, B., Colbeau, A. and Vignais, P. M. (1991) *Mol. Microbiol.* **5**, 2519-2527.

100 Dross, F., Geisler, V., Lenger, R., Theis, F., Krafft, T., Fahrenholz, F., Kojro, E., Duchene, A., Triper, D., Juvenal, K. and Kröger, A. (1992) *Eur. J. Biochem.* **206**, 93-102.

101 Hidalgo, E., Palacios, J. M., Murillo, J. and Ruiz-Argüeso, T. (1992) *J. Bacteriol.* **174**, 4130-4139.

102 Friedrich, B. (1991) in *Abstracts of the 3rd International Conference on Molecular Biology of Hydrogenases* pp. 113-116, Troia, Portugal.

103 Menon, N. K., Robbins, J., Wendt, J. C., Shanmugam, K. T. and Przybyla, A. E. (1991) *J. Bacteriol.* **173**, 4851-4861.

104 Bruschi, M. and Guerlesquin, F. (1988) *FEMS Microbiol. Rev.* **54**, 155-176.

105 Sauter, M., Böhm, R. and Böck, A. (1992) *Mol. Microbiol.* **6**, 1523-1532.

106 Böhm, R., Sauter, M. and Böck, A. (1990) *Mol. Microbiol.* **4**, 231-243.

107 Lutz, S., Jacobi, A., Schlensog, V., Böhm, R., Sawers, G. and Böck, A. (1991) *Mol. Microbiol.* **5**, 123-135.

108 Lee, J. H., Patel, P., Sankar, P. and Shanmugam, K. T. (1985) *J. Bacteriol.* **162**, 344-352.

109 Sankar, P. and Shanmugam, K. T. (1985) *J. Bacteriol.* **162**, 353-360.

110 Waugh, R. and Boxer, D. H. (1986) *Biochimie* **68**, 157-166.

111 Wu, L. F. and Mandrand-Berthelot, M. -A. (1986) *Biochimie* **68**, 167-179.

112 Sankar, P. and Shanmugam, K. T. (1988) *J. Bacteriol.* **170**, 5446-5451.

113 Sankar, P. and Shanmugam, K. T. (1988) *J. Bacteriol.* **170**, 5433-5439.

114 Schlensog, V., Birkmann, A. and Böck, A. (1989) *Arch. Microbiol.* **152**, 83-89.

115 Friedrich, C. G., Schneider, K. and Friedrich, B. (1982) *J. Bacteriol.* **152**, 42-48.

116 Schneider, K. and Schlegel, H. G. (1976) *Biochim. Biophys. Acta.* **748**, 353-361.

117 Schneider, K., Cammack, R., Schlegel, H. G. and Hall, D. O. (1979) *Biochim. Biophys. Acta* **578**, 445-461.

118 Schink, B. and Schlegel, H. G. (1979) *Biochim. Biophys. Acta* **567**, 315-324.

119 Pilkington, S. J., Skehel, J. M., Gennis, R. B. and Walker, J. E. (1991) *Biochemistry* **30**, 2166-2175.

120 Friedrich, B. (1990) *FEMS Microbiol. Rev.* **87**, 425-430.

121 Friedrich, C. G. (1982) *J. Bacteriol.* **149**, 1144-1149.

122 Friedrich, B. and Friedrich, C. G. (1983) *J. Bacteriol.* **153**, 176-181.

123 Schwartz, E., Tran-Betcke, A., Eberz, G., Gewinner, P. and Friedrich, B. (1991) in *Abstracts of the 3rd International Conference on Molecular Biology of Hydrogenases* pp. 130-132, Troia, Portugal.

124 Eberz, G. and Friedrich, B. (1992) *J. Bacteriol.* **173**, 1845-1854.

125 Lorenz, B., Dernedde, J. and Friedrich, B.in *Abstracts of the 3rd International Conference on Molecular Biology of Hydrogenases* pp. 121-122, Troia, Portugal.

126 Eberz, G., Eitinger, T. and Friedrich, B. (1989) *J. Bacteriol.* **171**, 1340-1345.

127 Eitinger, T. and Friedrich, B. (1991) *J. Biol. Chem.* **266**, 3222-3227.

128 Eitinger, T., Wolfram, L., Eberz, G. and Friedrich, B. (1991) in *Abstracts of the 3rd International Conference on Molecular Biology of Hydrogenases* pp. 165-168, Troia, Portugal.

129 Grande, H. J., Dunham, W. R., Averill, B. A., van Dijk, C. and Sands, R. H. (1983) *Eur. J. Biochem.* **136**, 201-207.

130 Van Dijk, C., Mayhew, S. G., Grande, H. J. and Veeger, C. (1979) *Eur. J. Biochem.* **102**, 317-330.

131 Van Dijk, C., Grande, H. J., Mayhew, S. G. and Veeger, C. (1980) *Eur. J. Biochem.* **107**, 251-261.

132 Fillipiak, M., Hagen, W. R. and Veeger, C. (1989) *Eur. J. Biochem.* **185**, 547-553.

133 Huynh, B. -H., Czechowski, M. H., Krüger, H. J., DerVartanian, D. V., Peck, H. D. Jr. and LeGall, J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3728-3732.

134 Grande, H. J., van Berkel-Arts, A., Bregh, J., van Dijk, K. and Veeger, C. (1983) *Eur. J. Biochem.* **131**, 81-88.

135 Van der Westen, H. M., Mayhew, S. G. and Veeger, C. (1978) *FEBS Lett.* **86**, 122-126.

136 Hagen, W. R., van Berkel-Arts, A., Krüse-Wolters, K. M., Voordouw, G. and Veeger, C. (1986) *FEBS Lett.* **203**, 59-63.

137 Patil, D. S., Moura, J. J. G., He, S. H., Teixeira, M., Pickril, B. C., DerVartanian, D. V., Peck, H. D. Jr., LeGall, J. and Huynh, B. -H. (1988) *J. Biol. Chem.* **263**, 18732-18738.

137a Glick, B. R., Martin, W. G. and Martin, S. M. (1980) *J. Can Microbiol.* **26**, 1214-1223.

137b Hatchikian, E. C., Forget, N., Fernandez, V. M., Williams, R. and Cammack, R. (1992) *Eur. J. Biochem.* **209**, 357-365.

138 Chen, J. -S. and Mortenson, L. E. (1974) *Biochim. Biophys. Acta.* **371**, 283-298.

139 Chen, J. -S. and Blanchard, D. K. (1978) *Biochem. Biophys. Res. Commun.* **84**, 1144-1150.

140 Adams, M. W. W. and Mortenson, L. E. (1984) *J. Biol. Chem.* **259**, 7045-7055.

141 Chen, J. -S., Mortenson, L. E. and Palmer, G. (1976) in *Iron and Copper Proteins* (Yasunobu, K. T., Mower, H. F. and Hayaishi, O., eds), pp. 68-83, Plenum, New York.

142 Adams, M. W. W., Ecleston, E. and Howard, J. B. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4932-4936.

143 Adams, M. W. W. and Mortenson, L. E. (1984) *Biochim. Biophys. Acta* **755**, 51-61.

144 Voordouw, G., Walker, J. E. and Brenner, S. (1985) *Eur. J. Biochem.* **148**, 509-514.

145 Voordouw, G. and Brenner, S. (1985) *Eur. J. Biochem.* **148**, 515-520.

146 Meyer, J. and Gagnon, J. (1991) *Biochemistry* **30**, 9697-9704.

147 Pickrill, B. C., Czechowski, M. H., Przybyla, A. E., Peck, H. D. and LeGall, J. (1986) *J. Bacteriol.* **167**, 722-725.

148 Van Dongen, W., Hagen, W. R., van den Berg, W. and Veeger, C. (1988) *FEMS Microbiol. Lett.* **50**, 5-9.

149 Voordouw, G., Hagen, W. R., Krüse-Wolters, K. M., van Berkel-Arts, A. and Veeger, C. (1987) *Eur. J. Biochem.* **162**, 31-36.

150 Pierik, A. J., Hagen, W. R., Redeker, J. S., Wolbert, R. B. G., Boersma, M., Verhagen, M. F. J. M., Grande, H. J., Veeger, C., Mutsaers, P. H. A., Sands, R. H. and Dunham, W. R. (1992) *Eur. J. Biochem.* **209**, 63-72.

151 Van den Berg, W. A. M., Stokkermans, J. P. W. G. and van Dongen, W. M. A. M. (1989) *J. Biotechnol.* **12**, 173-184.

152 Hagen, W. R., Pierik, A. J., Wolbert, R. B. G., Verhagen, M. F. J. M. and Veeger, C. (1991) in *Abstracts of the 3rd International Conference on Molecular Biology of Hydrogenases* pp. 72-75, Troia, Portugal.

153 Hagen, W. R., Pierik, A. J. and Veeger, C. (1989) *J. Chem. Soc. Faraday Trans. I* 85, 4083-4090.

154 Pierik, A. J., Wolbert, R. B. G., Mutsaers, P. H. A., Hagen, W. R. and Veeger, C. (1992) *Eur. J. Biochem.* **206**, 679-704.

155 Pierik, A. J., Hagen, W. R., Dunham, W. R. and Sands, R. H. (1992) *Eur. J. Biochem.* **206**, 705-719.

156 Moura, I., Tavares, P., Moura, J. J. G., Ravi, N., Huynh, B. -H., Liu, M. -Y. and LeGall, J. (1992) *J. Biol. Chem.* **267**, 4489-4496.

Chapter 2

HydC, a gene from *Desulfovibrio vulgaris* (Hildenborough) encodes a polypeptide homologous with the periplasmic hydrogenase

Jack Stokkermans, Walter van Dongen, Anita Kaan, Willy van den Berg and Cees Veeger

Published in: *FEMS Microbiology Letters* (1989) **58**, 217-222.

Note: The nucleotide sequence reported in this chapter has been submitted to EMBL and is available under accesion number X57838.

I. Summary.

Downstream of the genes for the structural α and β subunits of the periplasmic *Desulfovibrio vulgaris* (Hildenborough) hydrogenase a DNA fragment was detected with sequence homology to these genes. This fragment was cloned in *Escherichia coli* and the nucleotide sequence was determined. A gene was detected on the fragment with coding capacity for a 65.8 kDa polypeptide, HydC. The central part of HydC has an remarkable high degree of homology with the α subunit and the C-terminal part has similarity with the β subunit. These results strongly suggest that the three genes for HydC and the α and β subunits derive from one common ancestor gene. We succeeded in the identification of the translational product of this gene in *E. coli*, but were unable to determine the function of HydC after expression in *E. coli*.

II. Introduction.

The periplasmic hydrogenase of the anaerobic sulphate-reducer *D. vulgaris* (Hildenborough) catalyzes bidirectionally the oxidation and production of molecular hydrogen [1]. The enzym consists of two subunits, a 46 kDa α subunit and a 10 kDa β subunit [2], and contains three iron-sulphur clusters. Two of these are contained in a ferredoxin-like domain [3]; the third cluster probably forms the catalytic domain.

Genes for both subunits have been cloned in *E. coli* [2] and the amino acid sequences of both proteins were derived from the nucleotide sequences of the genes [3]. Although the subunits were expressed in *E.coli* from the cloned genes, no enzymatic activity was detected with the recombinant hydrogenase [4]; EPR studies showed, that the FeS cluster in the catalytic domain was not formed in *E.coli* [4].

Here we report on the detection and cloning of another *D. vulgaris* gene, that is situated in the immediate vicinity of the genes for the α and β subunits and encodes a protein with a striking similarity in amino acid sequence to both the α and β subunits.

III. Materials and Methods.

Bacterial strains and plasmids. *E.coli* TG2 [5] was used as a host for recombinant plasmids and for propagation of recombinant DNA in M13-derived vectors. Minicells were isolated from *E.coli* DS410 [6]. Recombinant plasmids were introduced in a hydrognase-negative mutant of *E.coli* HB101 [7] for measuring plasmid-encoded hydrogenase activity (this mutant was a gift of Dr. K. Stoker, Free University,

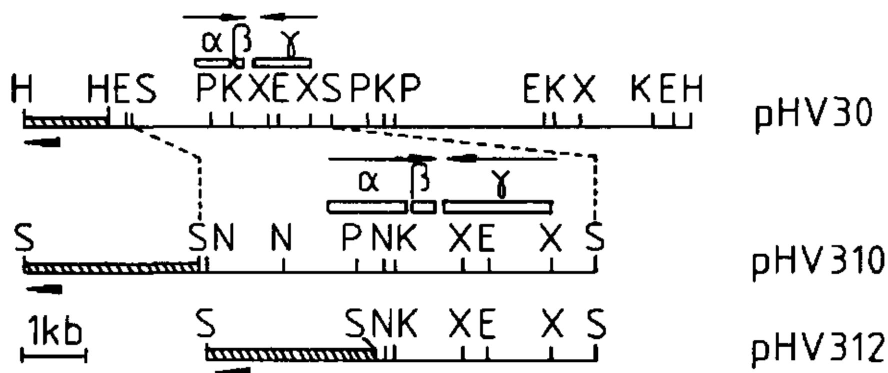


FIGURE 1: Recombinant plasmids used in this study. The positions of the *hyd* A, -B and -C genes, encoding the α and β subunits and *hyd* γ, respectively, are indicated above the recombinant plasmids; arrows show the direction of transcription. Vector (hatched) for all plasmids is pUC9. Arrowheads indicate the position and direction of the *lac*Z promoter on the vector. Plasmids pHV311 and -313 have the same inserts as pHV310 and -312, respectively, but with the vector and *lac*Z promoter in opposite orientation. pHV310 is drawn on a twofold reduced scale compared to the other plasmids. Restriction sites shown are those for *Eco*RI (E), *Hind*III (H), *Kpn*I (K), *Nru*I (N), *Pst*I (P), *Sal*I (S) and *Sma*I (X).

Amsterdam). Vectors used for cloning were pUC9 [8] and M13mp9, -18 and -19 [9]. All bacteria were grown in TYmedium [10] at 37°C.

Construction of recombinant plasmids. Recombinant plasmids are shown in Fig. 1. Plasmid pHV150 [4] with genes for the α and β subunits of *D. vulgaris* (Hildenborough) hydrogenase was used as a hybridization probe. All other procedures involved in cloning were as described [11].

Nucleotide sequence analysis. Nucleotide sequences were determined using the dideoxy chain termination method [12]; in some experiments 7'-deaza-dGTP (Boehringer) was used to prevent compression. All sequences were read from both strands. Sequences were analyzed with computer programs developed by Staden [13,14].

Minicells. Minicells of *E.coli* DS410 containing recombinant plasmids were isolated and labelled with a mixture of tritiated L-amino acids (New England Nuclear, 50 μCi per incubation) as described previously [15]. Lysates of labelled minicells were run on a 12.5 % polyacrylamide gel containing 0.1% SDS [16] and labelled proteins were visualized on Kodak XAR film after fluorography with 2,5-diphenyloxazide for ten days at -80°C.

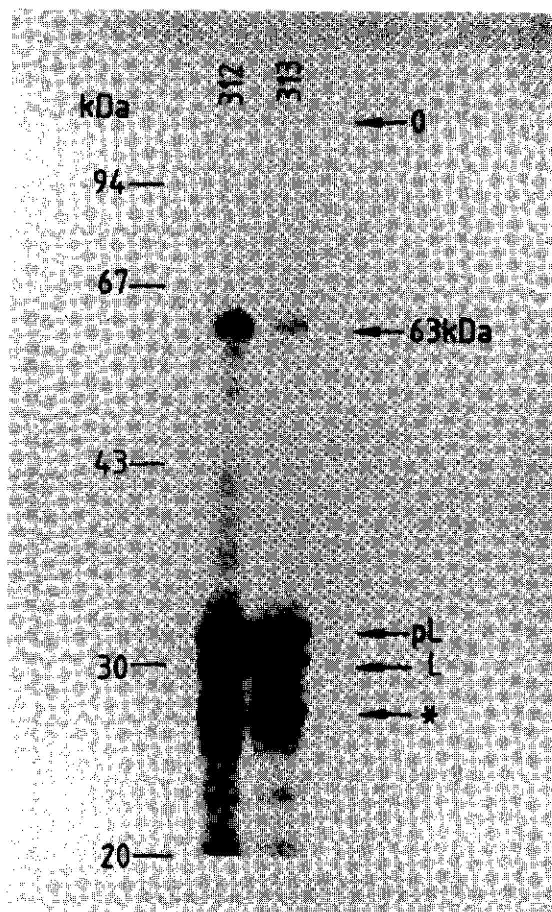


FIGURE 2: Fluorograph of a SDS-containing 12.5% polyacrylamide gel with tritiated proteins synthesized in E.coli minicells. The minicells contained plasmids pHV312 and -313, as indicated at the top of the lanes. The positions of the insert-encoded 63-kDa polypeptide and the vector encoded β-lactamase (L) and its precursor (pL), are indicated on the right (* *denotes an unidentified vector-encoded protein*); marker proteins are indicated on the left.

IV. Results and Discussion

In our search for genes encoding helper proteins which might be involved in the activation of the periplasmic *D. vulgaris* hydrogenase, we have cloned large DNA fragments of the *D.vulgaris* genome surrounding the genes for the structural α and β subunits. One such a recombinant plasmid, pHV30 (Fig. 1), contains in addition to the genes for the structural α and β subunits, another region that hybridizes with these genes

on a 18.5 kb insert. This region is located just downstream of the gene for the β subunit.

This region was subcloned on a 3.3 kb *NruI/SalI* fragment. Plasmids with this 3.3 kb insert in both orientations towards the vector-encoded *lacZ* promoter (pHV312 and -313) were introduced in a minicell-producing *E.coli* strain. Both minicells with pHV312 and with pHV313 produce a protein with an apparent molecular mass of approximately 63 kDa under direction of the cloned insert (Fig. 2). The level of expression is highly dependent on the orientation of the insert towards the *lacZ* promoter: it is very low in clones with pHV313 (not well reproducible on photograph), but much higher in clones with pHV312.

The nucleotide sequence of the DNA in this fragment was determined. It was found to contain an open reading frame of 2060 bp with a gene with coding capacity for a 65.8 kDa protein (Fig. 3) that has a remarkable homology with the hydrogenase α and β subunits (Fig. 4); no other reading frames of similar size were detected. The putative ATG-startcodon of this gene is preceded by a sequence (GGAGA) with homology to the consensus sequence for ribosome binding [17]. A weak promoter should be present on the cloned fragment, as this gene is expressed in both orientations towards the vector-encoded *lacZ* promoter (Fig. 2). Nevertheless, no homology with *E. coli* consensus promoter sequences could be detected. Transcription of the gene might terminate just downstream of the coding region at the same palindromic sequence that has been proposed as a terminator for the convergently transcribed genes for the α and β subunits [3].

The protein encoded by this gene, *hydC*, has a remarkable homology with both the α and β subunits of the hydrogenase. As shown in Fig. 4, 52% of the α subunit sequence between Ala^{107} and Pro^{389} is conserved in *hydC* (Ala^{235}-Pro^{518}). This region of the α subunit probably contains the residues, presumed to be cysteines, that are involved in coordination of the catalytic FeS cluster; 5 out of 10 cysteines in this region of the α subunit are conserved in *hydC*. Little sequence conservation exists between the N-terminal parts of the α subunits and HydC, except for two short stretches that contain the sequence Cys-Xaa_2-Cys-Xaa_2-Cys-Xaa_3-Cys. These sequences are highly homologous to the sequences known to be involved in the coordination of two [4Fe-4S] clusters in bacterial ferrodoxins. Indeed, two ferrodoxin-like FeS clusters, that might be involved in electron transport, have been demonstrated in the α subunit [18] and it is tempting to speculate about their presence in *hydC*.

The C-terminal part of HydC is homologous with the β subunit: 38% of the β subunit sequence between Asp^{54} and Val^{99} is conserved in HydC (Asp^{539}-Val^{584}). The β subunit is synthesized as a precursor protein with a 34-amino acid leader sequence for

```
                                                            M  N  A  F  I  N  G  K  E  V  R  C
 CATCCCGCAGTGCGACGACTGACGGGCCGACGAAGCGGAGAGCACATGAACGCGTTCATCAATGGCAAGGAAGTCCGGTG
         420       430       440       450       460       470       480       490

  E  P  G  R  T  I  L  E  A  A  R  E  N  G  H  F  I  P  T  L  C  E  L  A  D  I  G
TGAACCGGGCAGGACGATACTTGAGGCCGCACGCGAGAACGGGCACTTCATTCCCACGTTGTGCGAACTTGCCGACATCGG
        500       510       520       530       540       550       560       570

  H  A  P  G  T  C  R  V  C  L  V  E  I  W  R  D  K  E  A  G  P  Q  I  V  T  S  C
TCATGCACCCGGGACGTGCCGGGTCTGTCTGGTCGAGATATGGCGTGACAAGGAGGCCGGGCCGCAGATTGTCACCTCCTG
       580       590       600       610       620       630       640       650

  T  T  P  V  E  E  G  M  R  I  F  T  R  T  P  E  V  R  R  M  Q  R  L  Q  V  E  L
TACGACCCCCGTCGAGGAGGGAATGCGCATCTTCACGCGTACCCCTGAAGTACGCAGGATGCAGCGGCTACAGGTCGAACT
      660       670       680       690       700       710       720       730

  L  L  A  D  H  D  H  D  C  A  A  C  A  R  H  G  D  C  E  L  Q  D  V  A  Q  F  V
GCTGCTGGCCGACCATGACCATGACTGCGCAGCCTGCGCCCGTCATGGAGACTGCGAGTTGCAGGATGTGGCACAATTCGT
     740       750       760       770       780       790       800       810

  G  L  T  G  T  R  H  H  F  P  D  Y  A  R  S  R  T  R  D  V  S  S  P  S  V  V  R
GGGTCTTACCGGTACGCGTCACCATTTTCCGGACTATGCCCGCAGCCGCACCCGTGATGTCTCTTCGCCGTCCGTCGTGCG
    820       830       840       850       860       870       880       890

  D  M  G  K  C  I  R  C  L  R  C  V  A  V  C  R  N  V  Q  G  V  D  A  L  V  V  T
CGACATGGGCAAGTGCATCAGGTGCCTGCGCTGTGTCGCCGTGTGCCGCAACGTACAGGGCGTCGATGCCCTCGTGGTGAC
   900       910       920       930       940       950       960       970

  G  N  G  I  G  T  E  I  G  L  R  H  N  R  S  Q  S  A  S  D  C  V  G  C  G  Q  C
GGGAAACGGCATCGGCACCGAAATCGGGCTGCGGCACAATCGTAGCCAGAGTGCGTCCGACTGTGTGGGCTGTGGCCAGTG
  980       990      1000      1010      1020      1030      1040      1050

  T  L  V  C  P  V  G  A  L  A  G  R  D  D  V  E  R  V  I  D  Y  L  Y  D  P  E  I
CACATTGGTCTGCCCTGTGGGGGCATTGGCTGGACGGGACGACGTGGAGCGTGTCATCGACTATCTCTACGACCCCGAAAT
1060      1070      1080      1090      1100      1110      1120      1130      1

  V  T  V  F  Q  F  A  P  A  V  R  V  G  L  G  E  E  F  G  L  P  P  G  S  S  V  E
CGTCACCGTGTTCCAGTTCGCCCCGGCGGTGCGGGTGGGCCTCGGTGAGGAGTTCGGGCTGCCTCCCGGTTCAAGCGTGGA
140      1150      1160      1170      1180      1190      1200      1210      122

  G  Q  V  P  T  A  L  R  L  L  G  A  D  V  V  L  D  T  N  F  A  A  D  L  V  I  M
AGGGCAGGTGCCCACGGCCTTGCGCCTTCTCGGGGCAGACGTGGTACTCGATACCAACTTCGCAGCCGACCTCGTCATCAT
0      1230      1240      1250      1260      1270      1280      1290      1300

  E  E  G  T  E  L  L  Q  R  L  R  G  G  A  K  L  P  L  F  T  S  C  C  P  G  W  V
GGAGGAGGGCACCGAACTCCTGCAACGTCTTCGGGGCGGGGCGAAGCTGCCGCTCTTCACCTCCTGCTGCCCCGGCTGGGT
       1310      1320      1330      1340      1350      1360      1370      1380

  N  F  A  E  K  H  L  P  D  I  L  P  H  V  S  T  T  R  S  P  Q  Q  C  L  G  A  L
GAATTTCGCCGAGAAGCACCTCCCCGACATCCTGCCGCATGTCTCGACCACACGCTCGCCTCAGCAGTGCCTTGGCGCATT
      1390      1400      1410      1420      1430      1440      1450      1460

  A  K  T  Y  L  A  R  T  M  N  V  A  P  E  R  M  R  V  V  S  L  M  P  C  T  A  K
GGCCAAGACCTATCTTGCGCGCACCATGAACGTCGCACCGGAGAGGATGCGCGTCGTATCGTTGATGCCCTGCACGGCGAA
     1470      1480      1490      1500      1510      1520      1530      1540

  K  E  E  A  A  R  P  E  F  R  R  D  G  V  R  D  V  D  A  V  L  T  T  R  E  F  A
GAAGGAAGAGGCCGCACGGCCCGAATTCAGGCGCGACGGTGTCCGGGATGTGGACGCAGTGCTCACCACGCGTGAGTTCGC
    1550      1560      1570      1580      1590      1600      1610      1620

  R  L  L  R  R  E  G  I  D  L  A  G  L  E  P  S  P  C  D  D  P  L  M  G  R  A  T
CCGTCTTCTCCGGCGTGAGGGCATAGACCTCGCCGGACTCGAACCCTCGCCCTGCGACGACCCCCTGATGGGGCGGGCAAC
   1630      1640      1650      1660      1670      1680      1690      1700
```

```
  G  A  A  V  I  F  G  T  T  G  G  V  M  E  A  A  L  R  T  V  Y  H  V  L  N  G  K
CGGGGCGGCTGTCATCTTCGGTACGACAGGCGGGGTAATGGAGGCGGCACTGCGTACGGTCTACCATGTGCTGAACGGCAA
 1710      1720      1730      1740      1750      1760      1770      1780

  E  L  A  P  V  E  L  H  A  L  R  G  Y  E  N  V  R  E  A  V  V  P  L  G  E  G  N
GGAACTCGCCCCAGTAGAACTGCATGCCCTGCGCGGATACGAGAACGTGCGTGAGGCTGTCGTCCCGCTTGGTGAGGGTAA
1790      1800      1810      1820      1830      1840      1850      1860      1

  G  S  V  K  V  A  V  V  H  G  L  K  A  A  R  Q  M  V  E  A  V  L  A  G  K  A  D
CGGTTCCGTGAAGGTCGCCGTGGTGCATGGGCTCAAGGCTGCCCGGCAGATGGTCGAGGCGGTTCTTGCAGGGAAGGCCGA
870       1880      1890      1900      1910      1920      1930      1940      19

  H  V  F  V  E  V  M  A  C  P  G  G  C  M  D  G  G  G  Q  P  R  S  K  R  A  Y  N
CCATGTGTTCGTGGAGGTCATGGCATGCCCGGGTGGATGCATGGACGGAGGCGGTCAGCCGAGGTCGAAGCGCGCCTACAA
50        1960      1970      1980      1990      2000      2010      2020      203

  P  N  A  Q  A  R  R  A  A  L  F  S  L  D  A  E  N  A  L  R  Q  S  H  N  N  P  L
CCCCAACGCGCAGGCGCGACGTGCCGCCCTTTTCTCGCTCGATGCGGAAAACGCACTGCGGCAGTCGCACAACAATCCGCT
0         2040      2050      2060      2070      2080      2090      2100      2110

  I  G  K  V  Y  E  S  F  L  G  E  P  C  S  N  L  S  H  R  L  L  H  T  R  Y  G  D
CATCGGCAAGGTCTACGAATCATTCCTTGGCGAGCCCTGTTCGAATTTGTCTCACCGTCTGCTGCACACCCGGTATGGCGA
      2120      2130      2140      2150      2160      2170      2180      2190

  R  K  S  E  V  A  Y  T  M  R  D  I  W  H  E  M  T  L  G  R  R  V  R  G  D  S  D
CCGCAAGAGCGAAGTCGCCTACACCATGCGCGACATCTGGCATGAGATGACCCTTGGCAGGCGGGTACGGGGCGACTCTGA
     2200      2210      2220      2230      2240      2250      2260      2270

  *                                 ------- ---> <---   -------
TTGATCGTGGACAGGTGCTGACGTTTTCCGCCCCCCTGCGCTGCGCGTCGGGGGGCTTTTCTGTCGCCGCGCCACAAG
      2280      2290      2300      2310      2320      2330      2340      2350
```

FIGURE 3: Nucleotide sequence of the *hydC* gene and the derived amino acid sequence of its product, HydC. Numbering is from the rightmost *Sal*I site in pHV312. A putative ribosome binding site around position 450 is underlined; arrows denote a region of dyad symmetry that might function as a terminator for transcription.

export [19,20]. When the sequences of the α and β subunits are aligned alongside the HydC sequence to obtain maximal homology (Fig. 4), it is clear that the homology between the α and HydC sequences terminates almost exactly at the same position (corresponding to Pro^{518} in HydC), where in the aligned sequence of the β subunit the cleavage site for leader peptidase can be found. This observation suggests, that the three genes for HydC and for the α and β subunits derive from one common ancestor gene, that has been duplicated; one of the copies remained intact and evolved in the present *hydC* gene, while the other was split and evolved into the present *hydA* and *hydB* genes, the latter of these being provided with leader peptide encoding sequences.

The function of *hydC* is still obscure. We investigated, after expression of the polypeptide in *E.coli*, if HydC might be an alternative hydrogenase, and, by co-expression of HydC and the α and β subunits, if it might be a helper protein, involved in the activation of the periplasmic hydrogenase, for example by insertion of the FeS cluster in the catalytic centre. Helper proteins with a high degree of homology to the structural

```
C:  MNAFINGKEV RCEPGRTILE AARENGHFIP TLCELADIGH APGTCRVCLV EIWRDKEAGP
                                                         50

A:                                                              MSR

C:  QIVTSCTTPV EEGNRIFTRT PEVRRMQRLQ VELLLADHDH DCAACARHGD CELQDVAQFV
                                             100

                                                   50
A:  TVNERIEYEK HTPDPKADPD KLHFVQIDEA KCIGCDTCSQ YCPTAAIFGE NGE-------

C:  GLTGTRHHFP DYARSRTRDV SSPSVVRDMG KCIRCLRCVA VCRNVQGVDA LVVTGNGIGT
                                 150

                                                        100
A:  -----PHSIP HIEACINCGQ CLTHCPENAI YEAQSWVPEV EKKLKDGKVK CIAMPAPAVR

C:  EIGLRHNRSQ SASDCVGCGQ CTLVCPVGAL AG-RDDVERV IDYLYDPEIV TVFQFAPAVR
                  200

                                          150
A:  YALGDAPGMP VGSVTTGKML AALQKLGFAH CWDTEFTADV TIWEEGSEFV ERLTKKSDMP

C:  VGLGEEPGLP PGSSVEGQVP TALRLLGADV VLDTNFAADL VIMEEGTELL QRLRG-GA-K
          250

                            200
A:  LPQFTSCCPG WQKYAETYYP ELLPHFSTCK SPIGMNGALA KTYGAERMKY DPKQVYTVSI

C:  LPLFTSCCPG WVNFAEKHLP DILPHVSTTR SPQQCLGALA KTYLARTMNV APERMRVVSL
    300                                                   350

                250
A:  MPCIAKKYEG LRPELKSSGM RDIDATLTTR ELAYMIKKAG IDFAKL-PDG KRDSLMGEST

C:  MPCTAKKEEA ARPEFRRDGV RDVDAVLTTR EFARLLRREG IDLAGLEPSP CDDPLNGRAT
                                                400

          300
A:  GGATIFGVTG GVMEAALRFA YEAVTGKKPD SWDFKAVRGL DGIKEATVNV G--GTDVKVA

C:  GAAVIFGTTG GVMEAALRTV YHVLNGKELA PVELHALRGY ENVREAVVPL GEGNGSVKVA
                                     450

   350                                                400
A:  VVHGAKRFKQ VCDDVKAGKS PYHFIEYMAC PGGCVCGGGQ PVMPGVLEAM DRTTTRLYAG

C:  VVHGLKAARQ MVEAVLAGKA DHVFVEVMAC PGGCHDGGGQ PRSKRAYNPN AQARRAALFS
    480

B:          MQ IASITRRGFL KVACVTTGAA LIGIRMTGKA VAAVKQIKDY MLDRINGVVG
                                                              50

A:  LKKRLAMASA NKA*
                           500
C:  LDAENALRQS HNNPLICKVY ESFLGEPCSN LSHRLLHTRY GDRKSEVAYT MRDIWHEMTL

B:  ADAKFPVRAS QDNTQVKALY KSYLEKPLGH KSHDLLHTHW PDKSKGVKEL TTAGKLPNPR
                                                   100

    600
C:  GRRVRGDSD*

B:  ASEFEGPYPYE*
```

FIGURE 4: Alignment of the amino acid sequences of HydC and the α and β hydrogenase structural subunits according to homology (lines marked C, A and B, respectively). Each sequence is numbered from its N-terminus. Asterisks indicate identical residues. Cysteines are indicated by arrowheads and ferredoxin-like FeS coordination centres are boxed. The arrow indicates the cleavage site for leader peptidase in the β subunit.

proteins have been found, for example, also in nitrogenase: *nif*E and *nif*N gene products

insert the MoFe cofactor into the structural MoFe proteins and are highly homologous with these proteins (*nif*D and *nif*K gene products, respectively) [21]. No hydrogenase activity could be detected, however, in cell extracts of *E.coli* HB101 clones that produce either HydC or HydC together with the α and β subunits. Yet we do not exclude the possibility that in *Desulfovibrio hydC* has one of both functions, hydrogenase or helper protein, but like nitrogenase, requires additional gene products for activation that are not provided for by the *E.coli* background in which the above mentioned experiments were performed.

V. Acknowledgements

This investigation was supported by the Netherlands Organization for Chemical Research (S.O.N.) with financial aid from the Netherlands Organization for Scientific Research (N.W.O.). The authors are indebted to Mr. J. Benen for advice with nucleotide sequence determination.

VI. References

1 Van der Westen, H.M., Mayhew, S.G. and Veeger, C. (1978) *FEBS Lett.* **86**, 122-126.

2 Voordouw, G., Walker, J.E. and Brenner, S. (1985) *Eur. J. Biochem.* **148**, 509-514.

3 Voordouw, G. and Brenner, S. (1985) *Eur. J. Biochem.* **148**,514-520.

4 Voordouw, G., Hagen, W.R., Krüse-Wolters, K.M., Van Berkel-Arts, A. and Veeger, C. (1987) *Eur. J. Biochem.* **162**, 31-36.

5 Gibson, T.J. (1984) PhD Thesis, Univ. of Cambridge.

6 Dougan, G. and Sherratt, D. (1977) *Mol. Gen. Genet.* **151**, 151-160.

7 Boyer, H.W. and Rouland-Dussoix, D. (1969) *J. Mol. Biol.* **41**, 459-472.

8 Vieira, J. and Messing, J. (1982) *Gene* **19**, 259-268.

9 Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* **33**, 103-119.

10 Miller, J.H. (1977) Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

11 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

12 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

13 Staden, R. (1982) *Nucl. Acids Res.* **10**, 4731-4751.

14 Staden, R. (1984) *Nucl. Acids Res.* **12**, 499-503.

15 Mooi, F.R., Harms, N., Bakker, D. and De Graaf, F.K. (1981) *Infect. Immun.* **32**, 1155-1163.

16 Laemmli, U.K. (1970) *Nature (Lond.)* **227**, 680-685.

17 Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer,B.S. and Stormo, G. (1981) *Ann. Rev. Microbiol.* **35**, 118-135.

18 Hagen, W.R., Van Berkel-Arts, A., Krüse-Wolters, K.M., Voordouw, G. and Veeger, C. (1986) *FEBS Lett.* **203**, 59-63.

19 Prickrill, B.C., Czechowski, M.H., Przybyla, A.E., Peck, H.D. and LeGall, J. (1986) *J. Bacteriol.* **167**, 722-725.

20 Van Dongen, W., Hagen, W.R., Van den Berg, W. and Veeger, C. (1988) *FEMS Microbiol. Lett.* **50**, 5-9.

Chapter 3

Extensive sequence homology between domains of bacterial hydrogenases and the 75 kDa subunit of beef heart mitochondrial NADH:ubiquinone reductase: evidence for a new type of conserved element for electron transport.

Jack P.W.G. Stokkermans, Anita Kaan, Walter M.A.M. van Dongen and Cees Veeger

Note: The nucleotide sequence reported in this chapter has been submitted to EMBL and is available under accesion number X 57838.

Abbreviations: Nur: 75 kDa subunit of beef heart mitochondrial NADH dehydrogenase; HoxF, HoxU: α and γ subunits of the NAD-reducing hydrogenase from *Alcaligenes eutrophus* H16; hydA, hydB: α and β subunits of Fe-hydrogenase from *Desulfovibrio vulgaris* (Hildenborough); HydC: HydC polypeptide from *D. vulgaris*. Genes for these polypeptides are in italics.

I. SUMMARY

The N-terminus of the HydC polypeptide of *Desulfovibrio vulgaris* (Hildenborough) has been mapped. Similarity in primary structure between the central and C-terminal parts of HydC with the α and β subunits of the Fe-hydrogenase was noticed already previously. Here we demonstrate, that a domain containing more than 200 aminoacids at the N-terminus of this polypeptide shows extensive homology with one of the subunits of a different hydrogenase, the γ subunit of the NAD^+-reducing hydrogenase from *Alcaligenes eutrophus* H16 and, surprisingly, also with the N-terminal sequence of the 75 kDa subunit of beef heart mitochondrial NADH:ubiquinone reductase (NADH dehydrogenase). Among the conserved residues are 11 cysteines; four of these are contained in 2 Cys-Xaa_2-Cys motifs and another four in a Cys-Xaa_2-Cys-Xaa_2-Cys-Xaa_n-Cys motif as found in coordination sites for ferredoxin-like [4Fe-4S] cubanes. This homology suggests similarity in structure (e.g. as coordination sites of metal clusters) and function of these domains. The function of HydC is not yet known. In the NAD^+-reducing hydrogenase and in NADH-dehydrogenase, the conserved domains are probably involved in electron transport between subunits of these multimeric enzymes -containing FMN as a prosthetic group to catalyze the oxidation/reduction of NAD(H))- and the different subunits that have metal clusters as catalytic centers -to catalyze the oxidation of H_2 and the reduction of ubiquinone, respectively-. The structural homology between parts of these enzymes suggests, that the relatively simple NAD-reducing hydrogenase might be useful as a model to elucidate part of the catalytic mechanism and electron transport pathway in the extremely complex NADH-dehydrogenase.

II. INTRODUCTION

In some redoxenzymes, the reduction of NAD^+ on the oxidation of NADH is coupled, via FMN, to the oxidation/reduction of metal clusters elsewhere in the enzyme. Here we present evidence, that at least some of these enzymes contain a highly conserved domain, that might be involved in electron transfer from and to FMN. This conserved domain is found in the 75 kDa subunit of beef heart mitochondrial $NADH^+$ dehydrogenase and a NAD-dependent hydrogenase from *A. eutrophus* H16. It is also found in the HydC polypeptide of *D. vulgaris*, for which until now no function has been determined.

NADH dehydrogenase (complex I) is the most complex of the components of the mitochondrial respiratory chain. It catalyzes the transfer of electrons from NADH to ubiquinone, which is coupled to proton translocation. Although subject of intensive research,

nowledge of structure and catalytic mechanism of this enzyme is rather limited because of ts extreme complexity (26 -30 subunits). The 75 kDa subunit is the largest subunit of the omplex. Four iron ions and four acid-labile sulfide ions have been detected in the isolated ubunit (Ragan and Hatefi, 1986), that are either arranged into two [2Fe-2S] clusters Ragan, 1987) or into one [4Fe-4S] cluster (Runswick et al., 1989). The aminoacid equence of this subunit has been recently derived from the nucleotide sequence of the cDNA Runswick et al., 1989).

The NAD^+-reducing hydrogenase from *Alcaligenes eutrophus* H16 consists of 4 lifferent subunits ($\alpha_1\beta_1\gamma_1\delta_1$)and catalyzes the transfer of protons and electrons from H_2 to NAD^+ (Schneider and Schlegel, 1976). The subunits are arranged in two dimers, kept ogether by nickel. The hydrogenase dimer consists of the β and δ subunits and catalyzes the xidation of H_2 in the presence artificial electron acceptors. This dimer contains nickel and FeS clusters. The diaphorase dimer, formed by the α and γ subunits, catalyzes the reduction of NAD^+ and contains one [2Fe-2S] and two [4Fe-4S] clusters and FMN (Schneider et al., 1979, 1984). The amino acid sequence of all four subunits has recently been derived from the nucleotide sequences of the cloned genes (Tran-Betcke et al., 1990).

The function of the HydC polypeptide of the anaerobe sulfate reducer *Desulfovibrio vulgaris* (Hildenborough) is not yet known. The gene for this protein is located directly upstream of the *hydA* and *hydB* genes encoding the subunits of the heterodimeric Fe-hydrogenase that catalyzes H_2 production linked with lactate oxidation. The primary structure of the C-terminal two-thirds of the HydC polypeptide is highly homologous with that of the Fe-hydrogenase (Stokkermans et al., 1989, see also Fig. 3) and can be considered as being derived from a fusion between the α subunit and the mature part (i.e. without leader sequence for export) of the β subunit of this hydrogenase.

Here we demonstrate the high protien sequence homology of the three proteins.

II. MATERIALS AND METHODS.

Antiserum against the HydC polypeptide. A 0.4 kb *Eco*RI/*Sma*I fragment of pHV312, containing the sequence encoding Phe^{372}-Pro^{508} of HydC (Stokkermans et al., 1989) was ligated in frame with the truncated *lac*Z gene in vector pEX2 (Stanley and Luzio, 1984). The resulting plasmid was transformed in *E. coli* JM101 harbouring plasmid pCI857 (Remaut et al., 1983) with a gene for the temparature-sensitive p_L-repressor. Induction of expression of the cro-LacZ-HydC fusion protein by a temperature shock and isolation of inclusion bodies containing this fusion protein was as described (Stanley and Luzio, 1984). After a final purification step by preparative SDS-polyacrylamide gel electrophoresis, the

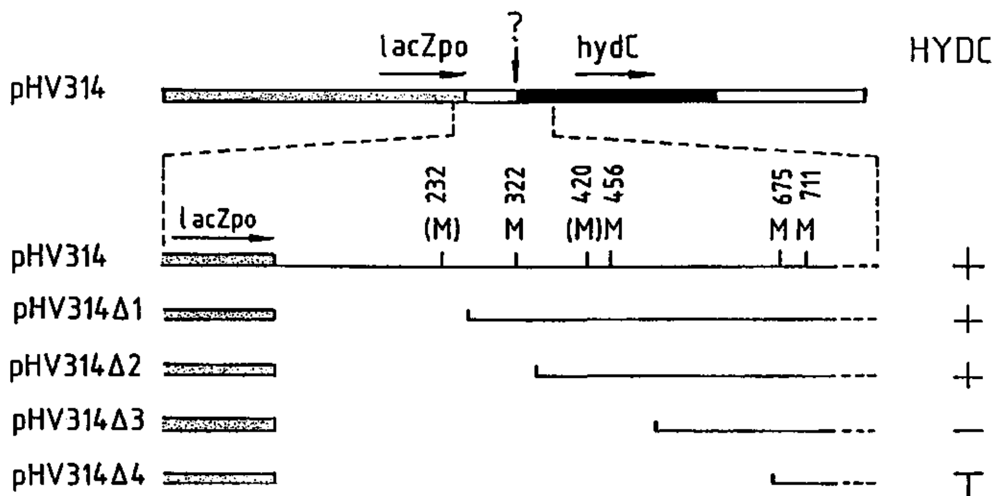
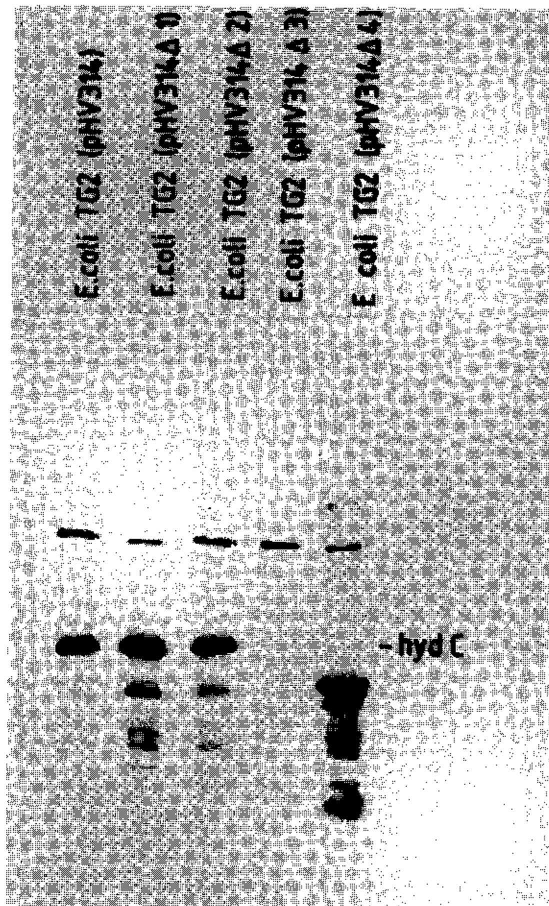


Figure 1 A : Mapping of the N-terminus of HydC. Left: Map of plasmid pHV314 with the vector pUC9 (stippled) and the approximate position of the *hydC* gene (black). Arrows show the direction of transcription of the *hydC* gene and the orientation of the *lac*Z-promoter. M and (M) represent positions of ATG- and GTG-codons, respectively, that might serve as startcodons, with the distance from the end of the insert nearest to *lac*Zpo (in basepairs). pHV314Δ1-4 represent plasmids with deletions in the upstream region made with exonuclease III.

fusion protein was used to elicit antibodies in rabbits. Antibodies to the *cro-lacZ* domains of the fusion protein were removed from the antiserum by absorption with a concentrated extract of *E. coli* JM101 (pEX2). Antibodies against this part of HydC did not react with the homologous part of HydA.

Mapping of the N-terminus of the HydC polypeptide. A 3.3 kb *Sal*I/*Nru*I fragment of pHV312 (Stokkermans et al., 1989), containing the *hydC* gene together with up- and downstream DNA sequences, was inserted downstream of the *lac*Z-promoter of vector pUC9 (Vieira and Messing, 1982), resulting in the recombinant plasmid pHV314 (Fig. 1). This plasmid was digested with *Pst*I and *Sal*I, which both cut the recombinant plasmid once in the polylinker of the vector, between the *lac*Z-promoter and the N-terminus of the *hydC* gene (the *Sal*I-site proximal to *hydC*). Unidirectional deletions were made in the linearized plasmid by digestion with exonuclease III; at several time points during the exonuclease digestion, samples were removed and single-strand protruding ends were removed with nuclease S_1 (Henikoff, 1987). The plasmids were recircularized and transformed to *E. coli* TG2 (Gibson, 1984). The extent of the deletions was determined with nucleotide sequencing. Production of the intact 63 kDa HydC polypeptide in several clones with



igure 1 B : SDS-polyacrylamide gel with polypeptides produced by *E. coli* clones with he respective plasmids, detected with antiserum against HydC. (+): expression of the ntire 63 kDa HydC polypeptide; (-): no expression of this polypeptide; (T): expression f a truncated HydC polypeptide.

lifferent deleted plasmids was checked by SDS-polyacrylamide gel electrophoresis of lysates of the respective clones and immunodetection of the 63 kDa protein.

V. RESULTS AND DISCUSSION.

Mapping of the N-terminus of HydC. The gene for the HydC polypeptide of *D. vulgaris* (Hildenborough) has previously been identified by its high degree of homology with he genes for the α and β subunits of the Fe-hydrogenase (Stokkermans et al., 1989). The evel of expression of this protein in *D. vulgaris* is very low. The HydC polypeptide is

expressed from the cloned gene in *E. coli*, but accumulates in an insoluble form in this host and is not readily purified. Expression of the *hydC* gene in *E. coli* minicells resulted in production of a polypeptide with an apparent molecular mass of 63 kDa. Based on the size of this polypeptide, we tentatively assigned the ATG-codon at position 456 in the cloned fragment as the translational start site defining the N-terminus of *hydC*. However, several other ATG and GTG codons, some of these preceded by putative ribosome binding sites, are present in this part of the sequence (Fig. 1). In order to map the N-terminus of HydC more accurately, the *hydC* gene was cloned upstream of the *lac*Z promoter of vector pUC9 and a series of deletion mutants was constructed (Fig. 1, left panel). Deletions upstream of the startcodon will result in expression of the entire 63 kDa HydC protein; deletions extending beyond the startcodon will either result in the production of a truncated HydC polypeptide (if the codons for the N-terminal aminoacids for β-galactosidase encoded by the vector are in frame with the codons for the remaining part of HydC) or in complete loss of HydC synthesis (in case of a frameshift). It appears from Fig. 1 (right) that deletions extending as far as the ATG codon around position 322 of the insert (pHV314Δ1 and -2) still result in production of the entire 63 kDa polypeptide. On the other hand, deletions extending beyond the ATG-codon around position 456 result either in the production of truncated polypeptides (pHV314Δ4) or in complete loss of protein production (pHV314Δ3). This leaves the GTG-codon around position 420 or the ATG-codon around position 456 as putative translational start sites. The presence of a sequence (GGAGA) with homology to the consensus sequence for ribosome binding indicates, that the ATG-codon at position 456 represents the translational start site, as has been predicted previously (Stokkermans et al.,1989).

Homology between HydC and other polypeptides. The central part of the HydC polypeptide was previously shown to be highly homologous with the α subunit of this hydrogenase, 52% of the residues being conserved; in the C-terminal part of HydC, 38% of the residues occurring in the mature part of the β subunit of the Fe-hydrogenase are conserved (Stokkermans et al., 1989). A third conserved domain was detected in HydC by screening the PIR-database using the FASTA program (Pearson and Lipman, 1988). The N-terminal sequence of HydC is highly homologous with the γ subunit of the NAD^+-reducing hydrogenase of *A. eutrophus* H16 and, surprisingly, also with the N-terminal sequence of the 75 kDa subunit of beef heart mitochondrial NADH dehydrogenase (excluding the 23 aminoacid leader sequence) (Fig. 2). The homology with the NADH dehydrogenase subunit (Nur) extends over 214 amino acids, of which 72 residues (34%) are identical in both sequences and another 27 residues (13%) have homologous counterparts. The entire γ subunit of the *A. eutrophus* hydrogenase (HoxU) has homology with the N-terminal 245

```
                                      1                                            50
NUR.BOS:  23 aa leader...   TATAASNLIEVFFDGQSVMVEPGTTVLQACEKVGMQIPRFCYHERLSVA-GNCRMCLVEI--EK--APKVVAACAMPVN

HYDC.DVU:                       MNAFINGKEVRCEPGRTILEAARENGHFIPTLCELADIGHAPGTCRVCLVEIWRDKEAGPQIVTSCTTPVE

HOXU.ALC:                      MSIQITIDGKTLTTEEGRTLVDVAAENGVYIPTLCYLKDKPCL-GTCRVCSV-----KVNG-NVAAACTVRVS


                                                 100                                         150
NUR.BOS:  KGWNILTNSEKTKKAREGVMEFLLANHPLDCPICDQGGECDLQDQSMMFGSDRSR--FLEGKRAVEDKNIGGPLVKTIMTR CIQCTRCIRFA SE

HYDC.DVU: EGMRIFTRTPEVRRMQRLQVELLLADHDHDCAACARHGDCELQDVAQFVGLTGTRHHFPDYARS-RTRDVSSPSVVRDMGK CIRCLRCVAVC RN

HOXU.ALC: KGLNVEVNDPELVDMRKALVEFLFAEGNHDCPSCEKSGRCQLQAVGYEVDMNVSR--FP-YRFPVRVVDHASEKIWLERDR CIFCQRCVEFI RD


                                           200
NUR.BOS:  IA-GVDDLGTTGRGNDMQVGTYIEK-MFMSE L---GNIIDIC PVGALTSK...493 aa

HYDC.DVU: VQ-GVDALVVTGNGIGTEIGLRHNRSQSASD CVGCGQCTLVC PVGALAGRDDVERVIDYLYDPEIVTVFQFAPAVRVGLGEEF...360 aa

HOXU.ALC: KASGRKIFSISHRGPESRIEIDAELANANP- PEQVKEAVAIC PVGTILEKR-------VGYDDPIGRRKYEIQSVRARALEGEDK*
```

Figure 2: Alignment of the N-terminal amino acid sequence of HydC from *D. vulgaris* (Hildenborough) (HYDC.DVU) with the N-terminal sequence of the 75 kDa subunit of beef heart mitochondrial NADH-Ubiquinone oxidoreductase (NUR.BOS) and the g subunit of the (NAD+)-reducing hydrogenase from *A. eutrophus* H16 (HOXU.ALC). The arrow indicates the splicing site for leader peptidase in NUR.BOS. (|): identical residues; (.): homologous residues: ILV; ST; SA; FWY; RKH; DE; NQ. Numbers refer to the HydC sequence.

amino acids of HydC, with 75 (32%) identical residues and 32 (13%) conservative replacements. Also the, not earlier recognized, homology between HoxU and Nur is high: 61 identical residues (29%) are found in the aligned sequences in Fig. 2; 40 residues are identical in all three sequences. The high cysteine content in the homologous sequences is striking: 12 cysteines are found in Nur and HoxU and 16 in HydC. Eleven of these cysteines are conserved in all three polypeptides. Fig. 3 gives a schematic representation of the homology between HydA, HydB, HydC, HoxU and Nur.

Especially interesting in Fig. 3 is the segment containing residues 120 - 220 of HydC. This segment contains two clusters of cysteine residues (Cys-Xaa_2-Cys-Xaa_2-Cys-Xaa_3-Cys). Similar cysteine motifs occur twice in [8Fe-8S]ferredoxins with two magnetically interacting [4Fe-4S] cubanes. It is known from the three-dimensional structure of these ferredoxins, that the first three cysteines from the first cysteine motif and the fourth cysteine from the second motif coordinate together one [4Fe-4S] cluster, the second cluster being coordinated by the remaining cysteines (Bruschi and Guerlesquin, 1988). Two similar motifs of clustered cysteines are apparent in the N-terminal region of the α subunit of the Fe-hydrogenase (Voordouw and Brenner, 1985). Evidence for the presence of two magnetically interacting [4Fe-4S] cubanes in the Fe-hydrogenase has been obtained by EPR spectroscopy (Voordouw et al., 1987); these [4Fe-4S] clusters are involved in electron transport to the catalytic site of the enzyme. Except for these cysteine motifs, the region between residues 120 and 220 of HydC has no sequence homology with the N-terminal region of HydA (Stokkermans et al., 1989).

It is clear from Figs. 2 and 3, that only one of these [4Fe-4S] coordination centers, comprising the first three cysteines from the first motif and the fourth from the second motif, is conserved in the Nur and HoxU sequences. This difference suggests, that the conserved element is involved in one-electron transfer in Nur and HoxU -the redox states of a [4Fe-4S] cluster shuttles between 2^+, oxidized, and 1^+, reduced,- and in two-electron transfer in HydC. Although the second [4Fe-4S] coordination center is apparently absent in Nur and HoxU, the HydC-Nur and HydC- HoxU homologies are higher in this region than the HydC-HydA homology.

The diaphorase dimer of the NAD-reducing hydrogenase of *A. eutrophus* and its highly homologous counterpart in *Nocardia opaca* 1b contain one [2Fe-2S] and two [4Fe-4S] clusters (Schneider et al., 1979, 1984). A coordination center for a second [4Fe-4S] cluster in this dimer can be assigned to a cysteine motif in the C-terminal region of the α subunit encoded by *hox*F (Tran-Betcke et al., 1990). It cannot be excluded, that some of remaining

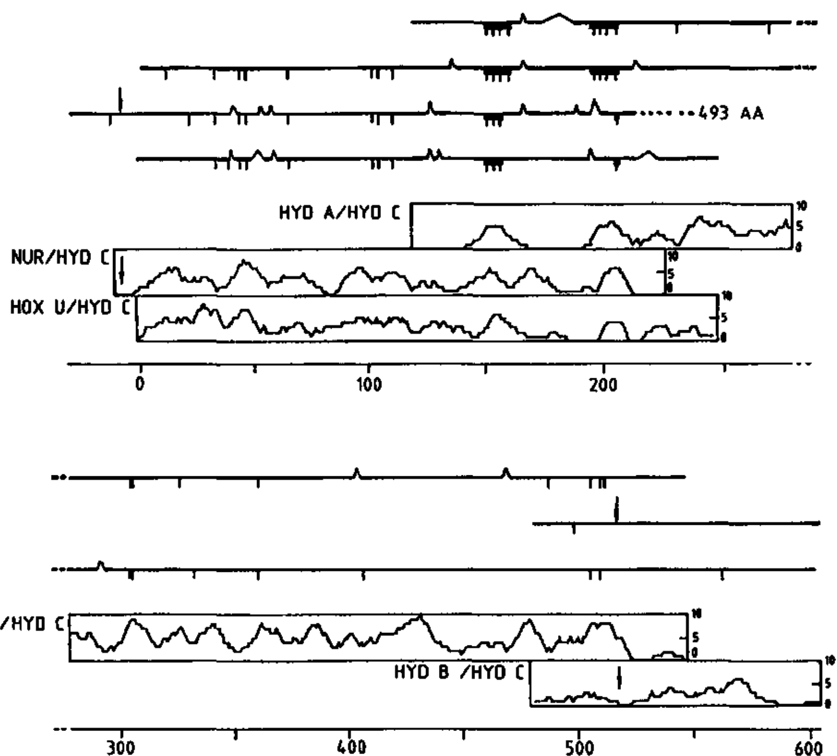


FIGURE 3:.Schematic representation of the homology between HydC and the α and β subunits of the Fe-hydrogenase from *D. vulgaris* (HYDA and HYDB, respectively), the γ subunit of the NAD+-reducing hydrogenase of *A. eutrophus* H16 (HOXU) and the N-terminal sequence of the 75 kDa subunit of beef heart mitochondrial NADH-dehydrogenase (NUR). Upper lines show the alignment of the sequences; gaps to be made for optimal alignment are shown as interruptions in the lines. Downward pointing arrows indicate splicing sites for leader peptidases in NUR and HYDB. Downward ticks represent positions of cysteines; black boxes represent ferredoxin-like cysteine motifs. Homology between HydC and the other sequences is shown in the boxes below the aligned sequences. Shown is the amount of identical residues in a ten-aminoacid-long window that is shifted along the aligned sequences. Numbers refer to the HydC sequence.

seven conserved cysteines in HoxU and Nur are involved in coordination of a [2Fe-2S] cluster. Four of these remaining cysteines occur in two Cys-Xaa_2-Cys motifs.

The Fe-hydrogenase contains a third, novel type of FeS cluster [6Fe-6S] in its active site (Hagen et al., 1986, Adams, 1990). Although the high degree of homology between the C-terminal part of HydC and the subunits of the Fe-hydrogenase might be suggestive for the presence of a similar cluster in HydC, failure to isolate HydC has prevented till now

verification of this hypothesis. It should be noted, that there is no homology between the C-terminal part of HydC or the subunits of the Fe-hydrogenase on the one hand and the subunits of the hydrogenase dimer of the NAD^+-reducing hydrogenase from *A. eutrophus* on the other hand. Contrary to the *D. vulgaris* Fe-hydrogenase, the *Alcaligenes* dimer belongs to the class of hydrogenases that have, besides of iron-sulfur clusters, also nickel in the active site (Schneider et al., 1979, 1984).

Functional relevance of the conserved domains. The sequence homology between HydC, HoxU and Nur suggests similarity in structure and function of the conserved domains. Several lines of evidence indicate an even more extended structural relationship between the diaphorase dimer of the NAD^+-dependent hydrogenase and the site of NADH oxydation of NADH dehydrogenase.

The diaphorase dimer of the hydrogenase catalyzes the reduction of NAD^+ with electrons derived from the oxydation of H2 in the hydrogenase subunit. The diaphorase dimer contains FMN, a [2Fe-2S] and two [4Fe-4S] cubanes as well as a NAD^+ binding-site (Schneider et al., 1984).

In NADH dehydrogenase, the site of NADH oxydation has been assigned to a subfraction of the enzyme (the FP fragment), that contains 3 subunits, of 51, 24 and 10 kDa (Ragan and Hatefi, 1986). The 75 kDa subunit is not isolated in the FP fragment, but is closely associated with it, as shown by chemical cross-linking (Cleeter et al., 1985). Foto-affinity labelling with NAD^+-analogs labelled both the 75 kDa and the 51 kDa subunits (Chen and Guillory, 1981, 1984), indicating that the NADH binding-site might be a hydrophobic cleft formed between both subunits. The 51 kDa subunit has been assigned as the site for FMN reduction (Chen and Guillory, 1981). Besides the 4 Fe ions and the 4 acid-labile sulfide ions in the 75 kDa subunit, another 6 Fe and 6 S ions have been detected in the FP fragment (Ragan and Hatefi, 1986, Ragan, 1987). The temperature dependence of the EPR spectra of the reduced, paramagnetic metal clusters in the FP fragment indicate, that they probably occur as a [2Fe-2S] cluster (which copurifies with the 24 kDa subunit) and a [4Fe-4S] cluster (associated with the 51 kDa subunit) (Ragan and Hatefi, 1986). Of the five Fe-S clusters that become paramagnetic upon reduction NADH dehydrogenase (Kowal et al., 1986), two, a [2Fe-2S] and a [4Fe-4S] cluster are found in the FMN containing fragment and a second [4Fe-4S] cluster might be present at a close distance in the 75 kDa subunit. The [2Fe-2S] cluster in the FP fragment is the first cluster to be reduced after addition of NADH to the complex (Bakker and Albracht, 1986, Krishnamoorthy and Hinkle, 1988). Further evidence for structural relationship between the diaphorase dimer and the site of NADH-oxidation in eukaryotic NADH dehydrogenase is given by the conservation of the primary structure of the N-terminal 200 amino acids of the α subunit of the diaphorase dimer

ran-Betcke et al., 1990) in the 24 kDa subunits of the FP fragments from several ıkaryotic sources (human, bovine, rat) (Pilkington and Walker, 1989, Nishikimi et al., 988, Von Bahr-Lindstrom et al., 1983, Chomyn and Tsai Lai, 1989). These observations ıggest a structural similarity between the diaphorase subunit of the NAD^+-reducing ydrogenase and part of NADH-dehydrogenase, comprising the N-terminal part of the 75 Da subunit, the 24 kDa subunit and presumably the 51 kDa subunit (although no sequence ata are available for this subunit). The conserved domain in Nur and HoxU might serve a milar function: the transfer of electrons between FMN and metal-containing reaction centers sewhere in the enzyme, i.e. linking NAD^+ reduction/NADH oxidation to duction/oxidation of metal clusters.

It should be noted, that besides these homologies also differences exist. In the NAD^+-ependent hydrogenase, the oxidation of H_2 (E_0' = -420 mV) is linked with the reduction of AD^+ (E_0' = -320 mV). NADH-dehydrogenase, on the other hand, catalyzes the oxidation f NADH and reduces ubiquinone (E_0' = -40 mV). In other words, under standard onditions, the electron flow in NADH dehydrogenase is reversed compared to that in the AD^+-dependent hydrogenase. In agreement with this, midpoint potentials of the iron-sulfur lusters in NADH dehydrogenase are roughly in between those of the NAD/NADH and /QH_2 couples (E_m -30 - -318 mV, Krishnamoorthy and Hinkle, 1988;E_m -20 - -245 mV, agan, 1987). Midpoint potentials of the Fe-S clusters in NAD^+-dependent hydrogenase are ower (E_m -325 - -445 mV, Schneider et al., 1979), in between those of the H_2/H^+ and AD/NADH couples. Although the subunit structure and composition around the site of AD-reduction/NADH-oxidation probably have considerable homology in both enzymes, e differences in the midpoint potentials of the Fe-S clusters in both enzymes are indicative r differences in the protein structure around the clusters.

Although the HydC polypeptide has not yet been purified, the presence of domains omologous with the subunits of the Fe-hydrogenase on the one side and a domain with milarity to polypeptides occuring in NAD^+-reducing/NADH oxidizing enzymes at the other de, suggests that this polypeptide might be a subunit of an enzyme involved in the oduction of H_2 from NADH or in the oxidation of H_2 by NAD^+. Indirect evidence for the ccurence of a H_2 producing NADH dehydrogenase in *Desulfovibrio* spp. exists (Kremer, 989). However, only a few catabolic pathways are known in *Desulfovibrio* that result in the oduction of NADH. No NADH production occurs in the most common catabolic pathway, e oxidation of lactate or pyruvate to acetate: neither lactate dehydrogenases, nor pyruvate hydrogenase are NAD^+-dependent enzymes in *Desulfovibrio* (Ogata et al., 1981, Stams d Hansen, 1982, Ogata and Yagi, 1986). It has been shown that protons and electrons rived from at least one of the two oxidation steps in the conversion of lactate to acetate are

directed to H_2 production in a pathway that involves the Fe-hydrogenase (Van den Berg, W.A.M, Van Dongen, W.M.A.M. and Veeger, C., 1991), and possibly cytochromes as electron donors for the hydrogenase. The only pathway known in *D. vulgaris* (Hildenborough) to result in NADH production is the conversion of glycerol to acetate: glyceraldehyde-3-phosphate dehydrogenase is an NAD^+-dependent enzyme in *D. vulgaris* (Kremer and Hansen, 1987). However, further investigation of the putative role of HydC in the glycerol metabolism of *D. vulgaris* is prevented by the very slow growth rates (generation time 1 - 2 weeks) and extremely low growth yields of *D. vulgaris* in glycerol medium.

V. ADDED NOTE

Recently, the primary structure of Fe-hydrogenase I from *Clostridium pasteurianum* has been published (Meyer and Gagnon, 1991). The high similarity in sequence, size and cysteine topology between *C. pasteurianum* hydrogenase I and the product encoded by *hydC* suggests that HydC is an alternative Fe-hydrogenase in *D. vulgaris*.

VI. REFERENCES.

ADAMS, M.W.W.: The structure and mechanism of iron-hydrogenases. *Biochim. Biophys. Acta* **1020** (1990) 115-145.

BAKKER, P.T.A. AND ALBRACHT, S.P.J.: Evidence for two independent pathways of electron transfer in mitochondrial NADH:Q oxidoreductase. *Biochim. Biophys. Acta* **850** (1986) 413-422.

BRUSCHI, M. AND GUERLESQUIN, F.: Structure, function and evolution of bacterial ferredoxins. *FEMS Microbiol. Rev.* **54** (1988) 155-176.

CHEN, S. AND GUILLORY, R.J.: Studies on the interaction of arylazido-β-alanyl NAD^+ with the mitochondrial NADH-dehydrogenase. *J. Biol. Chem.* **256** (1981) 8318-8332.

CHEN, S. AND GUILLORY, R.J.: Identification of the NADH-NAD^+ transhydrogenase peptide of the mitochondrial NADH-CoQ reductase (Complex I). *J. Biol. Chem.* **259** (1984) 5124-5131.

CHOMYN, A. AND TSAI LAI, S.S.-A.: cDNA of the 24 kDa subunit of the bovine respiratory chain NADH dehydrogenase: high sequence conservation in mammals and tissue-specific and growth-dependent expression. *Curr. Genet.* **16** (1989) 117-125.

CLEETER, M.W.J., BANISTER, S.H. AND RAGAN, C.I.: Chemical cross-linking of mitochondrial NADH dehydrogenase from bovine heart. *Biochem. J.* **227** (1985) 467-474.

GIBSON, T.J.: Studies on the Epstein-Bar virus genome Academic Thesis (1984), University of Cambridge, UK.

HAGEN, W.R., VAN BERKEL-ARTS, A., KRÜSE-WOLTERS, K.M., VOORDOUW, G. AND VEEGER, C.: The iron-sulfur composition of the active site of hydrogenase from *Desulfovibrio vulgaris* (Hildenborough) deduced from its subunit structure and total iron-sulfur content. *FEBS Letters* **203** (1986) 59-63.

HENIKOFF, S.: Unidirectional digestion with exonuclease III in DNA sequence analysis. *Meth. Enzymol.* **155** (1987) 156-165.

KREMER, D.R.: The role of NADH and NADPH in dissimilatory processes in *Desulfovibrio* strains (1989). Academic Thesis, Rijksuniversiteit Groningen, NL, p. 92.

KREMER, D.R. AND HANSEN, T.A.: Glycerol and dihydroxyacetone dissimilation in *Desulfovibrio* strains. *Arch. Microbiol.* **147** (1987) 249-256.

KOWAL. A.T., MORNINGSTAR, J.E., JOHNSON, M.K., RAMSAY, R.R. AND SINGER, T.P.: Spectroscopic characterization of the number and type of iron-sulfur clusters in NADH:ubiquinone oxidoreductase. *J. Biol. Chem.* **261** (1986) 9239 - 9245.

KRISHNAMOORTHY, G. AND HINKLE, P.C.: Studies on the electro transfer pathway, topography of iron-sulfur centers, and site of coupling in NADH-Q oxidoreductase. *J. Biol. Chem.* **263** (1988) 17566-17575.

MEYER, J. AND GAGNON, J.: Primary structure of hydrogenase I from Clostridium pasteurianum. *Biochemistry* **30** (1991) 9697-9704.

NISHIKIMI, M., HOSOKAWA, Y., TODA, H., SUZUKI, H. AND OZAWA,T.: The amino acid sequence of the 24 kDa subunit, an iron-sulfur protein, of rat liver mitochondrial NADH dehydrogenase deduced from cDNA sequence. *Biochim. Biophys. Res. Comm.* **157**(1988) 914-920.

OGATA, M., ARIHARA, K AND YAGI, T.: D-lactate dehydrogenase of *Desulfovibrio vulgaris. J. Biochem.* **89** (1981) 1423-1431.

OGATA, M. AND YAGI, T.: Pyruvate dehydrogenase and the path of lactate degradation in *Desulfovibrio vulgaris* Miyazaki F. *J. Biochem.* **100** (1986) 311-318.

PEARSON, W.R. AND LIPMAN, D.J.: Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85** (1988)2444-2448.

PILKINGTON, S.J. AND WALKER, J.E.: Mitochondrial NADH-ubiquinone reductase: complementary DNA sequences of import precursors of the bovine and human 24 kDa subunits. *Biochemistry* **28** (1989) 3257-3264.

RAGAN, C.I.: Structure of NADH-Ubiquinone reductase (Complex I). *Curr. Top. Bioenerg* **15** (1987) 1-36.

RAGAN, C.I. AND HATEFI, Y.: Isolation of the iron-sulfur containing polypeptides of NADH:oxidoreductase ubiquinone.*Meth. Enzymol.* **126** (1986) 360-369.

REMAUT, E., TSAO, H. AND FIERS, W.: Improved plasmid vectors with a thermoinducible expression and temperature-regulated runaway replication. *Gene* **22** (1983) 103-113.

RUNSWICK, M.J., GENNIS, R.B., FEARNLEY, I.M. AND WALKER, J.E.: Mitochondrial NADH-Ubiquinone reductase: complementary DNA sequence of the import precursor of the bovine 75 kDa subunit. *Biochemistry* **28** (1989) 9432-9439.

SCHNEIDER, K., CAMMACK, R., SCHLEGEL, H.G. AND HALL, D.O.: The iron-sulphur centres of soluble hydrogenase from *Alcaligenes eutrophus*. *Biochim. Biophys. Acta* **578** (1979) 445-461.

SCHNEIDER, K., CAMMACK, R. AND SCHLEGEL, H.G.: Content and localization of FMN, Fe-S clusters and nickel in the NAD-linked hydrogenase of *Nocardia opaca* 1b. *Eur. J. Biochem.* **142** (1984) 75-84.

SCHNEIDER, K. AND SCHLEGEL, H.G.: Purification and properties of soluble hydrogenase from *Alcaligenes eutrophus* H16.*Biochim. Biophys. Acta* **452** (1976) 66-80.

STAMS, A.J.M. AND HANSEN, T.A.: Oxygen-labile L(+) lactate dehydrogenase activity in *Desulfovibrio desulfuricans*. *FEMS Microbiol. Lett.* **13** (1982) 389-394.

STANLEY, K.K. AND LUZIO, J.P.: Construction of a new family of high efficiency bacterial expression vectors: identification of cDNA clones coding for human liver proteins. *EMBO J.* **3** (1984) 1429-1434.

STOKKERMANS, J., VAN DONGEN, W., KAAN, A., VAN DEN BERG, W. AND VEEGER, C.: *HydC*, a gene from *Desulfovibrio vulgaris* (Hildenborough) encodes a polypeptide homologous with the periplasmic hydrogenase. *FEMS Microbiol. Letters* **58** (1989) 217-222.

TRAN-BETCKE, A., WARNECKE, U., BÎCKER, C., ZABOROSCH, C. AND FRIEDRICH, B.: Cloning and nucleotide sequences of the genes for the subunits of NAD-reducing hydrogenase of *Alcaligenes eutrophus* H16. *J. Bacteriol.* **172** (1990) 2920-2929.

VAN DEN BERG, W.A.M., VAN DONGEN W.M.A.M. AND VEEGER, C.: Reduction of the amount of periplasmic hydrogenase in *Desulfovibrio vulgaris* (Hildenborough) with antisense RNA: direct evidence for an important role of this hydrogenase in lactate metabolism. *J. Bacteriol.* **173** (1991) 3688-3694.

VIEIRA, J. AND MESSING, J.: The pUC plasmids, a M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19** (1982) 259-268.

VON BAHR-LINDSTROM, H., GALANTE, Y.M., PERSSON, M. AND JORNVALL, H.: *Eur. J. Biochem.* **134** (1983) 145-150.

VOORDOUW, G. AND BRENNER, S.: Nucleotide sequence of the gene encoding the hydrogenase from *Desulfovibrio vulgaris* (Hildenborough). *Eur. J. Biochem.* **148** (1985) 515-520.

VOORDOUW, G., HAGEN, W.R., KRÜSE-WOLTERS, K.M., VAN BERKEL-ARTS, A. AND VEEGER, C.: Purification and characterization of *Desulfovibrio vulgaris* (Hildenborough) hydrogenase expressed in *Eschrichia coli*. *Eur. J. Biochem.* **162** (1987) 31-36.

Chapter 4

The primary structure of a protein containing a putative [6Fe-6S] prismane cluster from *Desulfovibrio vulgaris* (Hildenborough).

Jack P. W. G. Stokkermans, Antonio J. Pierik, Ronnie B. G. Wolbert, Wilfred R. Hagen, Walter M. A. M. van Dongen and Cees Veeger.

Published in: *European Journal of Biochemistry* (1992) **208** , 435-442.

Note: The nucleotide sequence reported in this chapter has been submitted to EMBL and is available under accesion number Z11707.

I. Summary

The gene encoding a protein containing a putative [6Fe-6S] prismane cluster, has been cloned from *Desulfovibrio vulgaris* (Hildenborough) and sequenced. The gene encodes a polypeptide composed of 553 amino acids (M_r = 60161 Da). The DNA-derived amino acid sequence was partly confirmed by N-terminal sequencing of the purified protein and of fragments of the protein, generated by CNBr cleavage. Furthermore, the C-terminal sequence was verified by digestion with carboxypeptidase A and -B. The polypeptide contains nine Cys residues. Four of these residues are gathered in a Cys-Xaa_2-Cys-Xaa_7-Cys-Xaa_5-Cys motif that is located towards the N-terminus of the protein. No relevant sequence similarity was found with other proteins, including those with high-spin Fe-S clusters (nitrogenase, hydrogenase), with one significant exception: the stretch containing the first four Cys residues spans two submotifs, Cys-Xaa_2-Cys and Lys-Gly-Xaa-Cys-Gly, separated by 11 residues, that are also present in high-spin Fe-S cluster containing CO dehydrogenase. Western-blot analysis demonstrates cross-reactivity of antibodies raised against the purified protein both in *Desulfovibrio* strains and other sulfate-reducing bacteria. Hybridization of the cloned gene with genomic DNA of several other *Desulfovibrio* species indicates that homologous sequences are generally present in the genus *Desulfovibrio*.

II. Introduction

The prismane protein is a protein containing a putative [6Fe-6S] prismane cluster from the anaerobic sulfate reducer *Desulfovibrio vulgaris* (Hildenborough). The EPR spectrum of the dithionite-reduced protein is similar to that of inorganic model compounds containing the [6Fe-6S] prismane core [1,2]. This spectrum lacks signals that are characteristic of classical [2Fe-2S], [3Fe-4S] and [4Fe-4S] clusters. The prismane protein is also the first example of a biological system in which S=9/2 paramagnetism has been discovered [3,4]. It was shown that the Fe-S cluster in the prismane protein may occur in four different redox states: the three-electron reduced state $[6Fe\text{-}6S]^{3+}$ (S=1/2), $[6Fe\text{-}6S]^{4+}$ (S= even), $[6Fe\text{-}6S]^{5+}$ (S=1/2 and S=9/2), and the fully oxidized $[6Fe\text{-}6S]^{6+}$ (S=0) state that shows no EPR spectrum [3-5]. The aerobically isolated prismane protein occurs predominantly in the one electron reduced $[6Fe\text{-}6S]^{5+}$ state [5]. In this intermediate redox state the cluster exists in two magnetic forms: approximately 10% is low spin (S=1/2), and the major part occurs in the unusual high-spin state (S=9/2). Studies of the protein in this state by EPR and Mössbauer spectroscopies further suggested that two irons, at opposite ends of the prismane cluster, might be ligated by nitrogen [5]. A

protein similar to the prismane protein has recently been isolated from *Desulfovibrio desulfuricans* (ATCC 27774) [6,7]. This protein showed a $S=9/2$ EPR signal in the one-electron reduced state similar to that reported by Pierik et al. [3-5], however, it was proposed that the *D. desulfuricans* protein may contain two [6Fe-6S] clusters; one with $S=0$, the other with $S=9/2$ [6,7].

The biochemical characterization of the protein (purification, metal analysis, cellular localization, amino acid composition and N-terminal amino acid sequence) was further described by Pierik et al. [8]. The protein was described as a cytoplasmic, monomeric protein with a molecular mass of 52 kDa, as determined by sedimentation equilibrium centrifugation. The iron and acid-labile sulfur content were both estimated at approximately 6 atoms/polypeptide.

In the past few years, similar very high spin EPR signals were found in a number of enzymes: dissimilatory sulfite reductase from *D. vulgaris* (Hildenborough) containing 20-24 Fe plus two sirohemes [6], carbon monoxide dehydrogenase from *Methanothrix soehngenii* containing 18-30 Fe and two Ni [7,8] and the P-cluster of the nitrogenase MoFe-protein from *Azotobacter vinelandii* containing 24-28 Fe plus 2 Mo [9,10]. In these enzymes, the high-spin clusters are thought to be essential for the enzyme activity. Unfortunately, these proteins all contain several iron-sulfur clusters (and/or heme iron), which makes analysis of their EPR and Mössbauer spectra complicated. The prismane protein of *D. vulgaris* on the other hand, may serve as a relatively simple model protein to investigate this type of complex cluster, because it contains only six Fe arranged into one [6Fe-6S] cluster [5,8].

The structure of these so called superclusters in proteins will ultimately be given by X-ray crystallography. This requires knowledge of the primary structure of the protein containing the supercluster. In this paper we investigate the primary structure of the prismane protein of *D. vulgaris* (Hildenborough).

III. Materials and methods

Materials. Restriction endonucleases, DNA polymerase I, large fragment of DNA polymerase I (Klenow fragment), *Taq* DNA polymerase, T4 DNA ligase, S1 nuclease, RNA markers (0.24 - 9.5 kb) and exonuclease III were purchased form Bethesda Research Laboratories. Calf intestinal phosphatase, dNTP, ddNTP and protein A were obtained from Boehringer. [α-^{32}P]dATP (3000 Ci/mmol) and Na^{125}I was purchased from Amersham. Iodo-beads iodination reagent was from Pierce. Universal M13 sequencing primer was ordered from New England Biolabs. Bovine serum albumin, carboxypeptidases A and B (diisopropylfluorophosphate treated) and iodoacetamide were

from Sigma. 5,5'-Dithio-bis(2-nitrobenzoic acid) was ordered from Merck. Poly(vinylidene difluoride) and nitrocellulose membranes were purchased from Millipore. Goat anti-(rabbit IgG) serum conjugated to alkaline phosphatase was from Promega Biotec and goat anti-(mouse IgG) serum conjugated to alkaline phosphatase was from Bio-Rad. All chemicals used were of analytical grade.

Bacterial strains and plasmids. *Escherichia coli* strain TG2, (*recA*$^-$, *Δ(lac-pro), thi, supE,* [Res$^-$ Mod$^-$ (k)], F' (*traD36* , *proA*$^+$*B*$^+$, *lacI*^{q} ZΔM15) was used as a host for cloning [14]. Vectors used for cloning were pUC9 [15] and the phages M13mp18 and M13mp19 [16]. *D. vulgaris* (Hildenborough) (NCIMB 8303), *D. vulgaris* (Monticello) (NCIMB 9442), *D. gigas* (NCIMB 9332), *D.desulfuricans* strains ATCC 27774, G200 (isolated by J. D. Wall) and Norway 4 (NCIMB 8310) were grown under conditions as described [17]. *Desulfovibrio baarsii* (DSM 2075) [18], *Desulfosarcina variabilis* (DSM 2060) and *Desulfobacterium autotrophicum* (DSM 3382) were grown under standard conditions [18, 19].

Molecular biology techniques. For the construction of a library containing chromosomal DNA fragments of *D. vulgaris* (Hildenborough), a large amount of chromosomal DNA was isolated as described by Meade et al. [20]. This genomic DNA was partially digested with restriction enzyme *Sau*3AI and fragments were separated in a 0.7% agarose gel. Fragments of 5 - 10 kb were isolated by electroelution and ligated into *Bam*HI cleaved cloning vector pUC9. The recombinant plasmids were introduced into *E. coli* TG2 by electroporation in a Bio-Rad Gene pulser system, as reported in [21]. Transformed cells were spread onto agar plates containing 100 μg/ml ampicillin, 40 μg/ml 5-bromo-4-chloro-3-indolyl-b-D-galactoside and 20 μg/ml isopropyl-thio-β–D-galactoside. Replicas of these plates were made on nitrocellulose membranes and were used for immunodetection with antibodies against the prismane protein and with ^{125}I-labeled protein A, as described by Voordouw et al. [22]. An extract of *E. coli* TG2 containing (pUC9) was used to remove aspecifically reacting antibodies.^{125}I-labelled protein A was prepared according to Markwell [23]. The four different polyclonal antisera (two from rabbit and two from mouse) used in this screening were the same as those described previously [8].

Further standard DNA operations were performed as described in [24]. DNA sequence reactions, using DNA fragments cloned in M13mp18 and M13mp19 as a template, were performed according to the dideoxy-chain-termination method of Sanger [25]. Either large fragment of DNA polymerase I (Klenow fragment) or *Taq* DNA polymerase [26] was used as sequencing enzyme. Sequences were analyzed with computer programs developed by Staden [27, 28].

Isolation of *D. vulgaris* (Hildenborough) RNA, electrophoresis of the RNA in denaturating formaldehyde gels and Northern (RNA) blotting on Immobilon-N membrane (Millipore) were performed as described [17]. Genomic DNA of several *Desulfovibrio* strains was isolated on a small scale according to Meade et al. [20], digested with either *Eco*RI or *Pst*I and blotted on Immobilon-N membrane (Millipore) [24]. Hybridization with nick-translated insert DNA of plasmid pJSP9 was performed under standard conditions [29].

Protein purification, modification and analysis. The prismane protein of *D. vulgaris* (Hildenborough) was purified as described by Pierik et al. [8]. Protein concentrations were determined with the microbiuret method [30] after trichloroacetic acid/deoxycholate precipitation [31].

For determination of the the amino acid composition of the prismane protein, approximately 10 nmol of the purified protein was desalted and denatured on a Bio-Gel-P column (1 cm x 12 cm) equilibrated with 5% (by vol.) formic acid. The sample was lyophilized and hydrolyzed in 6M HCl for 24 h (110 °C). Amino acid analysis was carried out on a Pharmacia LKB Alpha Plus system. Values for threonine and serine were upgraded by 7.5% to correct for degradation during hydrolysis. The cysteine and tryptophan residues could not be detected by this method and were determined separately. The cysteine content was determined with 5,5'-dithio-bis(2-nitrobenzoic acid) according to the procedure of Ellman [32], employing the modifications of Habeeb [33], after removal of acid-labile sulfur of the Fe-S cluster on a Bio-Gel-P column equilibrated with 5% (by vol.) aqueous formic acid. Tryptophan was determined fluorimetrically according to the method of Pajot [34].

Fragments for N-terminal amino acid determination were generated by CNBr digestion [35]. Prior to CNBr treatment, the cysteine residues of the protein were alkylated with iodoacetamide as described in [35]. After lyophilization, the peptide mixture was separated in a 0.1% SDS, 20% polyacrylamide gel [36] and transferred to a poly(vinylidene difluoride) membrane according to Towbin et al. [37]. The fragments were excised and N-terminal amino acid sequencing of these fragments was carried out with a gas-phase sequencer (Dr Amons, Sylvius laboratory, University Leiden, The Netherlands and Dr. Bär, Department of Biochemistry, School of Medicine, University of Leipzig, Germany).

For C-terminal amino acid sequencing approximately 10 nmol purified protein was denaturated on a Bio-Gel-P column equilibrated with 5% (by vol.) formic acid as described above. The protein was lyophilized and dissolved in 0.4 ml 0.2 M N-ethylmorpholine acetate and 0.5% SDS, pH 8.5 [38]. A mixture of carboxypeptidases A and B was added, to give a molar ratio of enzyme/substrate of 1:100 for both peptidases.

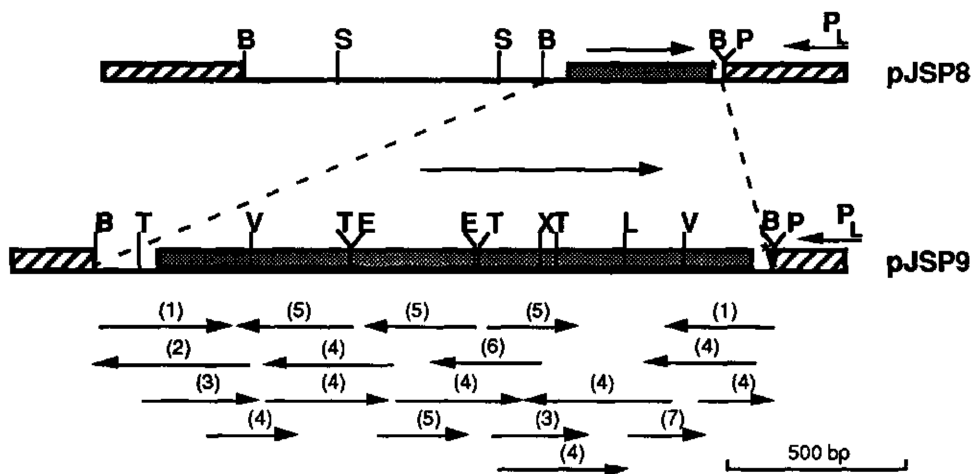


FIGURE 1 : Physical map of the plasmids containing the gene for the prismane protein. The gene for the prismane protein (shadowed) and the pUC9 vector (hatched) are shown as well as, the direction of transcription of the cloned gene and the *lacZ* promoter (P_L) of the vector (arrows) and the following restriction sites : (B) *Bam*HI, (S) *Sal*I, (P) *Pst*I, (T) *Taq*I, (E) *Eco*RI, (V) *Eco*RV, (L) *Ecl*XI and (X) *Sau*3AI. The gel readings and directions are presented below by arrows. Fragments were obtained by (1) *Bam*HI-*Pst*I, (2) *Bam*HI-*Eco*RV, (3) *Taq*I, (4) exonuclease III/S_1 nuclease [64], (5) *Eco*RI, (6) *Sau*3AI/*Eco*RV and (7) *Ecl*XI/*Bam*HI digests. *Bam*HI sites B* are lost after ligation to compatible ends.

The solution was incubated at 37°C, and aliquots containing 2.5 nmol protein were removed after 0, 0.5, 1 and 16 h. The samples were deproteinized by precipitation with 3% sulphosalicylic acid [39] followed by centrifugation. Amino acid analysis of the supernatant fluid was carried out on a Biotronic LC 6000E analyzer.

Cell lysates of several bacterial strains were electrophoresed in a 0.1% SDS 15% polyacrylamide gel [36]. Proteins were transferred onto a nitrocellulose membrane as described [37]. The membranes were incubated with 1000-fold diluted rabbit or mouseantisera against purified prismane protein. Immunodetection of bound antibodies was with goat anti-(rabbit IgG) serum or anti-(mouse IgG) serum conjugated to alkaline phosphatase.

IV. Results and Discussion

Cloning of the gene encoding the prismane protein. Approximately 5000 *E. coli* clones, containing a library of 5 - 10 kb fragments of *D. vulgaris* (Hildenborough) DNA were screened with four different polyclonal antisera for expression of the prismane protein. Two clones appeared positive and they contained the same recombinant plasmid,

pJSP8, with a 5.5 kb insert (Fig. 1). The gene encoding the prismane protein was located on a 1.9 kb *Bam*HI-*Pst*I fragment, subcloned in pJSP9 for expression studies and nucleotide sequencing.

Sequence of the prismane protein. Both strands of the insert of pJSP9 were sequenced (Fig. 2). A large open reading frame was found between the ATG codon around position 176 and a TAG codon around position 1835, which might encode a protein of 553 amino acids (60161 Da). Translation of the codons between nucleotides 176 and 217 (Fig. 2) resulted in an amino acid sequence that corresponds to the N-terminal sequence derived by direct peptide sequencing [8]. Two residues, identified only with some ambiguity by peptide sequencing (residues 3 and 6), appeared to be cysteines. This agreement confirmed that the cloned DNA fragment indeed encodes the prismane protein.

However, in some respects the DNA-derived polypeptide sequence, shown in Fig. 2 is not completely consistent with earlier observations. The amino acid composition as derived from the nucleotide sequence deviates from that determined earlier by analysis of the purified prismane protein, e. g. Cys, Pro, Ser and Met content [8]. Furthermore, the molecular mass of the protein as calculated from the nucleotide sequence (60.1 kDa) is considerably higher than that determined previously for the purified protein (52±0.9 kDa) [8]. It should be noted, however, that the latter figure was calculated from sedimentation equilibrium experiments. Gel filtration of the iodoacetylated, denaturated protein and electrophoresis in SDS-polyacrylamide gels resulted in estimations of the molecular mass that were more in agreement with that determined from the nucleotide sequence (62±6 kDa and 59.2±1.8 kDa, respectively).

Finally, the G/C content of the 1.9 kb *Bam*HI/*Pst*I fragment is very low (G+C=54.7%) compared to the average G/C content of *D. vulgaris* (Hildenborough) DNA (G+C=65%) [40]. This is reflected in a different codon usage of the gene compared to the codon usage of all other *D. vulgaris* genes sequenced till now (8 genes; 2536 codons) [41-48]. Table 1 shows that many codons with A or T in the wobble position are used relatively much more often in the gene encoding the prismane protein (e. g. AAA,TTT, AAT and TTA, etc.).

As these inconsistencies with earlier observations [8] may be indicative of mistakes in the nucleotide and the derived amino acid sequences, a series of control experiments was performed on the purified prismane protein. First, the amino acid composition of the purified protein was redetermined with two different purified protein preperations. This indicated, that numbers of the above mentioned amino acids were indeed overestimated or underestimated in the first determination. The average values of the three determinations

```
                                                   σ54 dependent promoter
                                                     -24         -12
GTGCTGTGTAAGGGGTTGCCCACAGACGACACGTACGGTGCCCTGGATCGTCTGGCGGATCGATTGCCGCCAAGCACAAGGAAAGCGGCT
        10        20        30        40        50        60        70        80        90

                                                -35                 -10     rbs        M F
TTCAGAGGTAGTTTTTTTCTAGTATCGATTTGAGGTGCGTGCATTGTCACGCGATCGGCAACAGTAATTATAAAGGAGGTCACTATGTT
       100       110       120       130       140       150       160       170       180

 C  F  Q  C  Q  E  T  A  K  N  T  G  C  T  V  K  G  M  C  G  K  P  E  E  T  A  N  L  Q  D
TTGCTTCCAGTGTCAGGAAACAGCGAAGAATACGGGATGCACGGTCAAAGGTATGTGCGGCAAGCCGGAAGAAACCGCCAATTTGCAGGA
       190       200       210       220       230       240       250       260       270

 L  L  I  F  V  L  R  G  I  A  I  Y  G  E  K  L  K  E  L  G  Q  P  D  R  S  N  D  D  F  V
TCTGCTCATTTTTGTGCTGCGTGGCATCGCCATTTATGGCGAAAAATTAAAGGAGTTGGGACAGCCGGACCGCTCTAATGACGACTTCGT
       280       290       300       310       320       330       340       350       360

 L  Q  G  L  F  A  T  I  T  N  A  N  W  D  D  A  R  F  E  A  M  I  S  E  G  L  A  R  R  D
GCTCCAGGGATTATTTGCGACCATCACCAATGCTAACTGGGATGATGCCCGATTTGAGGCTATGATATCGGAAGGCTTGGCTCGGCGTGA
       370       380       390       400       410       420       430       440       450

 K  L  R  N  A  F  L  A  V  Y  K  A  K  N  G  K  D  F  S  E  P  L  P  E  A  A  T  W  T  G
CAAGTTGAGAAATGCTTTTCTGGCTGTTTACAAGGCAAAAAATGGCAAAGATTTCAGCGAGCCGTTGCCCGAAGCCGCGACTTGGACTGG
       460       470       480       490       500       510       520       530       540

 D  S  T  A  F  A  E  K  A  K  S  V  G  I  L  A  T  E  N  E  D  V  R  S  L  R  E  L  L  I
GGACTCTACCGCTTTTGCTGAAAAGGCCAAGAGCGTAGGCATTCTGGCAACTGAGAACGAAGATGTGCGTTCTTTGCGCGAACTGCTCAT
       550       560       570       580       590       600       610       620       630

 I  G  L  K  G  V  A  A  Y  A  E  H  A  A  V  L  G  F  R  K  T  E  I  D  E  F  M  L  E  A
AATCGGTCTGAAAGGTGTCGCGGCCTATGCCGAGCACGCTGCTGTACTGGGCTTTCGCAAGACCGAAATCGACGAATTCATGCTTGAGGC
       640       650       660       670       680       690       700       710       720

 L  A  S  T  T  K  D  L  S  V  D  E  M  V  A  L  V  M  K  G  R  R  M  A  V  T  T  M  A  L
CCTGGCTTCAACGACCAAGGATTTGTCCGTGGATGAAATGGTCGCGCTGGTCATGAAGGGCCGGCGCATGGCCGTAACCACCATGGCCCT
       730       740       750       760       770       780       790       800       810

 L  D  E  A  N  T  T  T  Y  G  N  P  E  I  T  Q  V  N  I  G  V  G  K  N  P  G  I  L  I  S
GCTGGACGAGGCCAACACCACTACCTACGGCAACCCGGAAATCACCCAGGTCAACATCGGCGTGGGCAAGAACCCCGGCATCCTCATCAG
       820       830       840       850       860       870       880       890       900

 G  H  D  L  K  D  M  A  E  L  L  K  Q  T  E  G  T  G  V  D  V  Y  T  H  G  E  M  L  P  A
CGGCCATGACCTGAAGGATATGGCCGAGTTGCTGAAGCAGACCGAGGGCACAGGCGTGGACGTTTACACCCACGGCGAAATGCTTCCCGC
       910       920       930       940       950       960       970       980       990

 N  Y  Y  P  A  F  K  K  Y  P  H  F  V  G  N  Y  G  G  S  W  W  Q  Q  N  P  E  F  E  S  F
CAACTATTATCCGGCGTTCAAGAAGTACCCGCATTTCGTGGGGAACTACGGCGGTTCCTGGTGGCAGCAGAATCCCGAATTCGAGTCCTT
      1000      1010      1020      1030      1040      1050      1060      1070      1080

 N  G  P  I  L  L  T  T  N  C  L  V  P  L  K  K  E  N  T  Y  L  D  R  L  Y  T  T  G  V  V
CAACGGCCCCATCCTGCTGACCACAAACTGTCTGGTGCCGCTCAAAAAAGAGAACACCTATCTGGACCGTCTCTACACCACTGGCGTGGT
      1090      1100      1110      1120      1130      1140      1150      1160      1170

 G  Y  E  G  A  K  H  I  A  D  R  P  A  G  G  A  K  D  F  S  A  L  I  A  Q  A  K  K  C  P
TGGGTATGAAGGGGCCAAGCACATCGCTGACCGGCCTGCGGGAGGAGCCAAGGATTTCTCGGCGCTGATCGCACAGGCCAAGAAATGTCC
      1180      1190      1200      1210      1220      1230      1240      1250      1260

 P  P  V  E  I  E  T  G  S  I  V  G  G  F  A  H  H  Q  V  L  A  L  A  D  K  V  V  E  A  V
TCCGCCCGTTGAGATCGAGACTGGTAGCATCGTCGGCGGTTTCGCCCATCATCAGGTACTAGCCTTGGCCGACAAGGTGGTTGAGGCTGT
      1270      1280      1290      1300      1310      1320      1330      1340      1350

 K  S  G  A  I  K  R  F  V  V  M  A  G  C  D  G  R  Q  K  S  R  S  Y  Y  T  E  V  A  E  N
CAAGTCCGGCGCGATCAAGCGCTTTGTGGTCATGGCCGGGTGTGACGGTCGGCAGAAGTCTCGTTCCTATTACACTGAAGTGGCCGAAAA
      1360      1370      1380      1390      1400      1410      1420      1430      1440

 L  P  K  D  T  V  I  L  T  A  G  C  A  K  Y  R  Y  N  K  L  N  L  G  D  I  G  G  I  P  R
TTTGCCCAAGGACACCGTGATCTTGACGGCCGGTTGCGCCAAGTACCGCTATAACAAACTGAATCTCGGTGACATCGGTGGCATTCCACG
      1450      1460      1470      1480      1490      1500      1510      1520      1530

 V  L  D  A  G  Q  C  N  D  S  Y  S  L  A  V  I  A  L  K  L  K  E  V  F  G  L  D  D  I  N
TGTGCTGGACGCAGGGCAGTGCAACGATTCCTATTCCCTGGCCGTCATTGCTCTGAAACTCAAAGAAGTCTTCGGCCTGGATGACATCAA
      1540      1550      1560      1570      1580      1590      1600      1610      1620

 D  L  P  V  S  Y  D  I  A  W  Y  E  Q  K  A  V  A  V  L  L  A  L  L  F  L  G  V  K  G  I
CGATCTGCCCGTATCTTACGATATCGCCTGGTACGAGCAGAAGGCCGTGGCCGTGCTTCTGGCCCTCCTATTCTTAGGTGTGAAGGGAAT
      1630      1640      1650      1660      1670      1680      1690      1700      1710

 R  L  G  P  T  L  P  A  F  L  S  P  N  V  A  K  V  L  V  E  N  F  N  I  K  P  I  G  T  V
CCGCCTCGGACCCACGCTTCCGGCATTCCTGTCCCCGAACGTGGCCAAGGTACTGGTTGAAAATTTCAACATCAAACCAATCGGAACAGT
      1720      1730      1740      1750      1760      1770      1780      1790      1800

 Q  D  D  I  A  A  M  M  A  G  K  *       ρ-independent terminator
GCAGGATGACATCGCGGCCATGATGGCGGGTAAATAGATGAAGGGGGGCTTCGGCTCCCCTTTTTTCATCAATAGCTGATC
      1810      1820      1830      1840      1850      1860      1870      1880
```

FIGURE 2 : Nucleotide sequence of the 1883 bp *Bam* HI-*Pst* I fragment of plasmid pJSP9. The sequence lacks the initial *Bam*HI site. Putative promoter motifs are underlined in the upstream sequence. A possible ribosome binding site, just ahead of the translation-start site (position 176), is overlined. The amino acid sequence is written above the coding region. Amino acids that have been identified by peptide sequencing of CNBr fragments are overlined (including Met). In the downstream sequence a putative ρ-independent terminator is underlined.

aa	codon	I	II	aa	codon	I	II	aa	codon	I	II	aa	codon	I	II
F	TTT	36	5	S	TCT	25	6	Y	TAT	47	21	C	TGT	44	19
F	TTC	64	95	S	TCC	40	34	Y	TAC	53	79	C	TGC	56	81
L	TTA	5	0	S	TCA	5	4	*	TAA	0	0	*	TGA	0	13
L	TTG	19	8	S	TCG	10	28	*	TAG	100	87	W	TGG	100	100
L	CTT	7	22	P	CCT	9	7	H	CAT	57	25	R	CGT	32	23
L	CTC	18	38	P	CCC	39	62	H	CAC	43	75	R	CGC	37	42
L	CTA	4	1	P	CCA	9	2	Q	CAA	0	4	R	CGA	5	3
L	CTG	47	31	P	CCG	43	29	Q	CAG	100	96	R	CGG	21	21
I	ATT	17	7	T	ACT	21	2	N	AAT	38	14	S	AGT	0	3
I	ATC	76	82	T	ACC	52	67	N	AAC	62	86	S	AGC	20	25
I	ATA	7	11	T	ACA	12	5	K	AAA	33	7	R	AGA	5	1
M	ATG	100	100	T	ACG	15	26	K	AAG	67	93	R	AGG	0	10
V	GTT	15	5	A	GCT	22	7	D	GAT	44	21	G	GGT	25	24
V	GTC	24	45	A	GCC	30	62	D	GAC	56	79	G	GGC	47	55
V	GTA	15	11	A	GCA	8	13	E	GAA	57	56	G	GGA	16	7
V	GTG	46	39	A	GCG	19	18	E	GAG	43	44	G	GGG	12	14

Table 1 : Codon usage of the prismane protein gene compared to other *Desulfovibrio vulgaris* (Hildenborough) genes. The normalized codon usage in the gene of the prismane protein [I] (553 codons), is compared with that of the eight genes of *D. vulgaris* (Hildenborough) [II] (2536 codons) sequenced so far [41-48]. For the amino acids (aa) the single-letter code is used. *, amino acid not identified.

(including that in [8]) agree with the composition as derived from the nucleotide sequence (Table 2).

N-terminal peptide sequencing experiments were performed with CNBr cleaved fragments of the purified protein. The determined sequences (Fig. 2) all confirm the protein sequence derived from nucleic acid data. Finally, the C-terminal end of the purified protein was identified after hydrolysis by a mixture of carboxypeptidases A and B. A stoichiometric quantity of lysine and substoichiometric quantities of glycine and alanine were released during this experiment. This proves that the C-terminus is a lysine residue. The C-terminal lysine is probably preceded by a residue that is released very slowly by carboxypeptidase A or B digestion, e. g. glycine [38]. This confirms the C-terminal amino acid sequence Ala-Gly-Lys. Thus, the control experiments all confirm the amino acid sequence derived from the nucleotide sequence as presented here.

Transcription signals. The gene is preceded by a motif homologous with the consensus sequence of σ^{54} dependent promoters (Fig. 2) [49]. Transcription might terminate around position 1865 (Fig. 2), at a putative ρ-independent terminator ($\Delta G°$= -46.9 kJ/mol) [50]. This would result in a 1.8-kb mRNA. Hybridization of size-fractionated *D. vulgaris* RNA with the insert of pJSP9 showed an intensely labeled mRNA species of approximately 1.8 kb, but also a faint band with lower mobility (Fig.

Amino acid	From amino acid analysis[a]	From nucleotide Sequence
D+N	55.0±0.8	58
Q+E	46.3±2.6	50
S	21.7±3.3	20
T	32.0±0.8	33
P	30.0±5.0	23
H	8.3±1.3	7
K	39.7±1.7	40
R	20.0±3.3	19
G	52.7±0.5	49
A	62.3±2.5	59
V	41.0±0.8	41
I	26.0±2.2	29
L	58.7±4.0	57
Y	17.3±1.7	19
F	21.7±1.3	22
W	5[b]	5
M	11.0±3.6	13
C	10[c]	9

Table 2 : Amino acid composition of the prismane protein of *D. vulgaris* (Hildenborough)

[a]values represent average (± standard deviation) of determinations of 3 purified protein preps including that in [8]; values have been calculated for 553 residues (M_r 60161 Da)

[b]From fluorimetric analysis

[c]From 5,5'-dithio-bis(-2-nitrobenzoate) determination

3). This indicates that the majority of the mRNA is monocistronic, but some may be transcribed from a larger transcription unit.

Sequence comparisons with other proteins. Cysteines are commonly involved in coordination of Fe-S clusters. As there is evidence that two of the six irons in the prismane protein have nitrogen atoms as ligands [5], some four irons may be ligated to cysteines.

Nine cysteines are found in the primary structure of the prismane protein (shown in bold face in Fig. 2). Four of these are clustered in the N-terminal 21 amino acids, in a Cys-Xaa_2-Cys-Xaa_7-Cys-Xaa_5-Cys motif (Fig. 2). The other five residues are scattered through the C-terminal part of the primary structure.

Cysteines involved in ligation of the Fe-S clusters often occur in conserved motifs in the amino acid sequence, for example the Cys-Xaa_2-Cys-Xaa_2-Cys(-Xaa_3-Cys) motifs in ferredoxins with one (two) [4Fe-4S] cluster(s) [51,52] or the Cys-Xaa_2-Cys-Xaa_n-Cys-

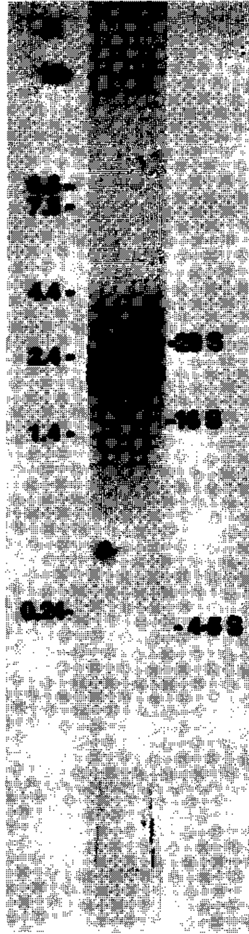


FIGURE 3 : Hybridization of size fractionated *D. vulgaris* RNA with the insert of pJSP9. Approximately 30 μg total RNA isolated from *D. vulgaris* (Hildenborough) was fractionated in a 1% agarose gel and hybridized with nick-translated insert DNA of plasmid pJSP9. RNA size markers and rRNA sizes are indicated on the left and the right, respectively.

Xaa_2-Cys motifs, which ligate a single iron in rubredoxins and probably in rubrerythrin [48,53]. In order to detect proteins with sequences homologous with that of the prismane protein, with special attention for conservation of the aforementioned Cys motif, the EMBL, GenBank and Swissprot data bases were screened with the FASTA program of Lipman and Pearson [54]. No protein sequences were found that were considered to have significant homology, except for the Cys-Xaa_2-Cys region. Also a second screening of the database with only the first 35 amino acid residues, containing the unusual Cys motif, did not reveal sequences with a more extensive similarity.

```
Prismane protein    (D.vu)     1    MFCFQCQETAKN-TGCTV---KGMCG
                                      |  |        | |     | | | |
CO dehydrogenase    (M.so)    65    DTCTLCTYGPCDLTGNK----KGACG
                                      |  |  | | |        : |  | |
CO dehydrogenase    (C.th)    61    ICCRFCMAGPCRIKATDGPGSRGICG

Consensus                           CXXC<--X11-14--> KGXCG
```

FIGURE 4 : Similarity between the prismane protein from *D. vulgaris (D. vu)* and CO dehydrogenases from *M. soehngenii (M. so)* and *C. thermoaceticum (C. th.)*. Residues identical in the aligned sequences are represented by (|); homologous residues K and R are represented by (:). Numbers refer to the corresponding positions in the proteins.

Mössbauer spectra of the as isolated and dithionite reduced prismane protein samples indicate that two of the six irons show increased quadrapole splitting and isomer shift upon reduction, and might have a more polar group as a ligand, e. g. a nitrogen atom [5]. These nitrogen atoms might be from histidines. Histidines function as coordination groups for heme iron and have also been indicated as iron ligands in iron-sulfur proteins [51] (see also the recent discussion of the evidence for His as a ligand of the FeMo cofactor in nitrogenase in [55]).

A more detailed sequence comparison, using Staden's DIAGON-program [56], was made with proteins that are known to contain high-spin Fe-S clusters: the MoFe-protein from *A. vinelandii* nitrogenase [12,57] and CO dehydrogenase from *M. soehngenii* [11,58]. Also included was the Fe-hydrogenase from *D. vulgaris* [46], for which a catalytic [6Fe-6S] cluster has been postulated [59]. No significant similarity was found with nitrogenase or hydrogenase . However, the N-terminal Cys motif spans two submotifs, comprising three Cys residues (Cys-Xaa_2-Cys and Lys-Gly-Xaa-Cys-Gly, Fig. 4) that were also found in the sequence of the α subunit of CO dehydrogenase (these motifs were not detected by screening the data bases with the FASTA program). In both proteins, these submotifs are separated by 11 amino acids containing the fourth Cys. The position of this fourth Cys is not conserved. These submotifs were also found in a slightly modified form in the only other CO dehydrogenase for which the primary structure is known, that of *Clostridium thermoaceticum* [60]. Arginine replaces lysine and the spacing between the Cys-Xaa_2-Cys and the Arg-Gly-Xaa-Cys-Gly sub-motifs is slightly larger (14 residues) in the *C. thermoaceticum* enzyme. Remarkably, this is one of the few conserved regions in the two CO dehydrogenases, which otherwise have a quite different primary structure [58, 60].

No clear indications exist yet about the geometry of the clusters, the ligation of the clusters and the protein structure around the clusters in the different proteins with

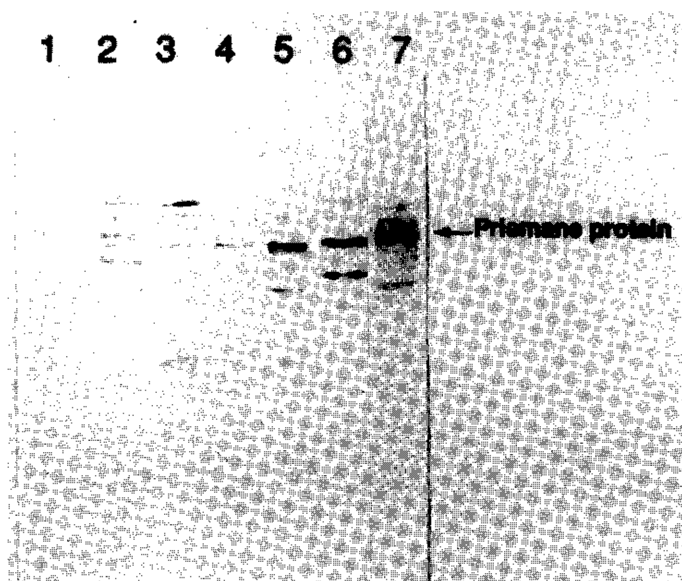


FIGURE 5 : Immunoblot screening for a prismane protein in sulfate reducing bacteria. A Western blot was immunostained after incubation with 1000-fold diluted rabbit antiserum raised against the prismane protein from *D. vulgaris* (Hildenborough). Approximately 15 μg cellular protein of *D. autotrophicum* (lane 1), *D. variabilis* (lane 2), *D. baarsii* (lane 3), *D. gigas* (lane 4), *D. desulfuricans* ATCC 27774 (lane 5), *D. desulfuricans* G200 (lane 6) and *D. vulgaris* (Hildenborough) (lane 7) were loaded on the gel.

superspin clusters. Based on the available data (amino acid sequence and EPR spectra of the superspin clusters), it is tempting to speculate that there is some similarity in the clusters and the cluster ligation of CO dehydrogenase and the prismane protein. For both proteins, a S=9/2 spin system has been proposed, based on the effective g values of the EPR signals of the high-spin clusters, but the redox behaviour of the clusters is quite different [11,4].

On the other hand, the high-spin clusters of nitrogenase and hydrogenase might be less corresponding. EPR spectroscopic data on the putative [6Fe-6S] cluster of hydrogenase do not resemble that from the prismane protein [61]. Also the P-clusters of nitrogenase have quite different EPR- and redox properties and probably a different iron content (eight instead of six irons/cluster), compared to the cluster of the prismane protein [5, 12]. The release of more primary structures and spectroscopic and X-ray crystallographic data of supercluster containing proteins is necessary for the elucidation of a possible common geometrical structure of these clusters.

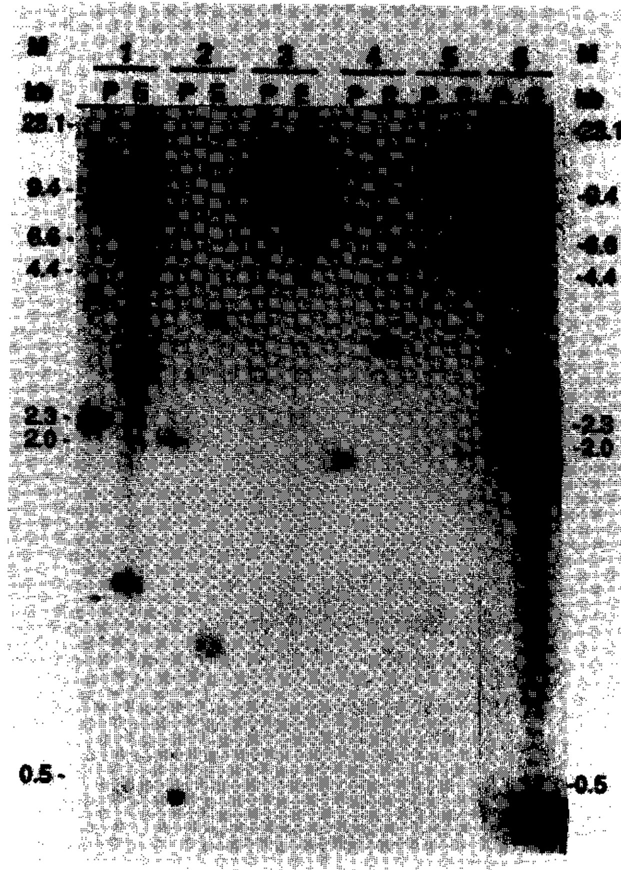


FIGURE 6 : Hybridization of digested DNA from several *Desulfovibrio* strains with the insert of pJSP9. 2.5 μg chromosomal DNA was digested with restriction endonuclease *Eco*RI (E) or *Pst*I (P), electrophoresed and blotted onto an Immobilon-N [poly(vinylidene difluoride)] membrane. This membrane was hybridized with nick-translated insert DNA of plasmid pJSP9. DNA is from *D. desulfuricans* Norway 4 (lane 1), *D. desulfuricans* G200 (lane 2), *D. desulfuricans* ATCC 27774 (lane 3), *D. gigas* (lane 4), *D. vulgaris* (Monticello) (lane 5) and *D. vulgaris* (Hildenborough) (lane 6). Sizes of marker fragments are shown. In order to visualize weakly labeled fragments, the blot exposed for an extended period of time; therefore, some lanes are overexposed.

Occurrence of the prismane protein and its structural gene in other sulfate reducers. Proteins with mobility similar to that of the *D. vulgaris* prismane protein in SDS/PAGE gel and cross-reacting with antibodies against this protein were detected in extracts of a variety of sulfate reducers (for classification, see [62]), not only in those belonging to the genus *Desulfovibrio* [8], but also in *D. autotrophicum*, *D.variabilis* and *D. baarsii* (three sulfate-reducing species that are able to degradate acetate by the acetyl-CoA route) [63] (Fig. 5). This observation suggests, that the prismane protein is common in sulfate reducers.

The occurrence of a gene for the prismane protein in other *Desulfovibrio* species can also be inferred from the hybridization of specific fragments of digested DNA from these species with the insert of pJSP9 (Fig. 6). The intensities of the hybridizing signals indicate, that the genes from the *D. desulfuricans strains* ATCC 27774 and Norway 4 have the highest similarity to the *D. vulgaris* (Hildenborough) gene. We are presently attempting the isolation and sequencing of genes from other sulfate-reducing species, in order to compare the primary structures of the proteins and to identify conserved regions that may have functional relevance (e. g. for coordination of the Fe-S cluster). Extensive simlarity (especially with the N-terminal Cys-motif) is found with the corresponding protein from *D. desulfuricans* ATCC 27774 (unpublished results).

V. Acknowledgements

We like to thank Mr. P. Houba for technical assistance, S. Arendsen and Dr. A. Stams (Department of Microbiology, Argicultural University, Wageningen, The Netherlands) for providing cultures of *D. baarsii*, *D. variabilis* and *D. autotrophicum* and Mr. L. Wennekes (Department of Genetics, Agricultural University, Wageningen, The Netherlands) for help with electroporation. We are indebted to Dr. W. de Jong (Department of Biochemistry, University of Nijmegen, The Netherlands) and Mr. W. Roelofsen (Department of Microbiology, Agricultural University, Wageningen, The Netherlands) for performing the amino acid analyses. This investigation was supported by the Netherlands Foundation for Chemical Research (SON), with financial aid from the Netherlands Organization for Scientific Research (NWO).

VI. References

1 Hagen, W. R., Pierik, A. J., and Veeger, C. (1989) *J. Chem. Soc. Faraday Trans. I* **85**, 4083-4090.

2 Kanatzidis, M. G., Hagen, W. R., Dunham, W.R., Lester, R. K. and Coucouvanis, D. (1985) *J. Am. Chem. Soc.* **107**, 953-961.

3 Pierik, A. J., Hagen, W. R. and Veeger, C. (1990) *EBEC Reports* **6**, 16.

4 Hagen, W. R., Pierik, A. J. and Veeger, C. (1990) *Ital. Biochem. Soc. Trans.* **1**, 85.

5 Pierik, A. J., Hagen, W. R., Dunham, W. R. and Sands, R. H. (1992) *Eur. J. Biochem.* **206**, 705-719.

6 Ravi, N., Moura, I., Tavares, P., LeGall, J., Huynh, B. -H. and Moura, J. J. G. (1991) *J. Inorg. Biochem.* **43**, 252.

7 Moura, I., Tavares, P., Moura, J. J. G., Ravi, N., Huynh, B. -H., Liu, M. -Y. and LeGall, J. (1992) *J. Biol. Chem.* **267**, 4489-4496.

8. Pierik, A. J., Wolbert R. B. G., Mutsaers P. H. A., Hagen, W. R. and Veeger, C., (1992) *Eur. J. Biochem.* **206**, 697-704.

9 Pierik, A. J. and Hagen, W. R.(1991) *Eur. J. Biochem.* **195**, 505-515.

10 Jetten, M. S. M., Hagen, W. R., Pierik, A. J., Stams, A. J. M. and Zehnder, A. J. B. (1991) *Eur. J. Biochem.* **195**, 385-395.

11 Jetten, M. S. M., Pierik, A. J. and Hagen, W. R. (1991) *Eur. J. Biochem.* **202**, 1291-1297.

12 Hagen, W. R., Wassink, H, Eady, R. R., Smith, B. E. and Haaker, H. (1987) *Eur. J. Biochem.* **169**, 457-465.

13 Hagen, W. R., Pierik, A. J., Wolbert, R. B. G., Wassink, H., Haaker, H., Veeger, C., Jetten, M. S. M., Stams, A. J. M. and Zehnder, A. J. B. (1991) *J. Inorg. Chem.* **43**, 237.

14 Gibson, T. J. (1984) Thesis, University of Cambridge.

15 Vieira, J.and Messing, J. (1982) *Gene (Amst.)***19**, 259-268.

16 Yanish-Perron, C., Vieira, J. and Messing, J. (1985) *Gene (Amst.)* **33**, 103-119.

17 Van den Berg, W. A. M., van Dongen, W. M. A. M. and Veeger, C.(1991) *J. Bacteriol.* **173**, 3688-3694.

18 Widdel, F. (1980) Thesis, University of Göttingen.

19 Byrsch, K., Schneider, C., Fuchs, G. and Widdel, F. (1987) *Arch. Microbiol.* **148**, 264-274.

20 Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E. and Ausubel, F. M. (1982) *J. Bacteriol.* **149**, 114-122.

21 Böttger, E. C.(1988), *Biotechniques* **6**, 878.

22 Voordouw, G., Walker, J. E. and Brenner, S. (1985) *Eur. J. Biochem.* **148**, 509-514.

23 Markwell, M. A. K. (1982) *Anal. Biochem.* **125**, 427-432.

24 Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1987) *Current Protocols in molecular biology*, John Wiley and Sons, New York.

25 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA.* **74**, 5463-5467.

26 D'Alessio, J. M., Kullman, J., Salyards, A. J. and Campbell, J. H. (1989) *Focus* **11:2**, 30-31.

27 Staden, R. (1982) *Nucleic Acids Res.* **10**, 4731-4751.

28 Staden, R. (1984) *Nucleic Acids Res.* **12**, 499-503.

29 Maniatis, T. , Fritsch, E. F. and Sambrook, J. (1982) *Molecular cloning: a laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

30 Goa, J. (1953) *Scand. J. Clin. Lab. Invest.* **5**,218-222.

31 Bensadoun, A. and Weinstein, D. (1976) *Anal. Biochem.* **70**, 241-250.

32 Ellman, G. L. (1959) *Arch. Biochem. Biophys.* **82**,70-77.

33 Habeeb, A. F. S. A. (1972) *Methods Enzymol.* **25**,457-464.

34 Pajot, P. (1976) *Eur. J. Biochem.* **63**, 263-269.

35 Matsudaira, P. T. (1989)*A practical guide to protein and peptide purification for microsequencing* , Academic Press Inc., London

36 Laemmli, U. K. (1970) *Nature* **227**, 680-685.

37 Towbin, H., Staehelin, T. and Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA.* **76**, 4350-4354.

38 Ambler, R. P. (1972) *Methods Enzymol.* **25**, 143-154.

39 Mondino, A., Bongiovanni, G., Fumero, S. and Rossi, L. (1972) *J. Chromatogr.* **74**, 255-263.

40 Postgate, J. R. (1984) in *The sulphate-reducing bacteria*, **2nd ed.**, p 11, Cambridge University Press, Cambridge.

41 Voordouw, G. and Brenner, S. (1985) *Eur. J. Biochem.* **148**, 515-520.

42 Krey, G. D., Vanin, E. F. and Swenson, R. P. (1988) *J. Biol. Chem.* **263**, 15436-15443.

43 Brumlik, M. J. and Voordouw, G. (1989) *J. Bacteriol.* **171**, 4996-5004.

44 Voordouw, G. (1988) *Gene (Amst.)* **69**, 75-83.

45 Curley, G. P. and Voordouw, G.(1988) *FEMS Microbiol. Lett.* **49**, 295-299.

46 Voordouw, G. and Brenner, S. (1986) *Eur. J. Biochem.* **159**, 347-351.

47 Stokkermans, J., van Dongen, W., Kaan, A., van den Berg, W. and Veeger, C. (1989) *FEMS Microbiol. Lett.* **58**, 217-222.

48 Prickril, B. C., Kurtz, D. M., LeGall, J. and Voordouw, G. (1991) *Biochemistry* **30**, 11118-11123.

49 Kustu, S., Santero, E., Keener, J., Popham, D. and Weiss, D. (1989) *Microbiol. Rev.* **53**, 367-376.

50 Tinoco, I., Borer, P. N.,Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., Gralla, J. (1973) *Nature New Biol.* **246**, 40-41.

51 Beinert, H. (1990) *FASEB J.* **4**, 2433-2491.

52 Bruschi, M. and Guerlesquin, F. (1988) *FEMS Microbiol. Rev.* **54**, 155-176.

53 Meyer, J., Gagnon, J., Sieker, L. C., van Dorsselaer, A. and Moulis, J. M. (1990) *Biochem. J.* **271**, 839-841.

54 Pearson, W. R. and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444-2448.

55 Smith, B. E., Eady, R. R. (1992) *Eur. J. Biochem.* **205**, 1-15.

56 Staden, R. (1982) *Nucleic Acids Res.* **10**, 2951.

57 Brigle, K. E., Newton, W. E. and Dean, D. R. (1985) *Gene (Amst.)* **37**, 37-44.

58 Eggen, R. I. L., Geerling, A. C. M., Jetten, M. S. M. and de Vos, W. M. (1991) *J. Biol. Chem.* **266**, 6883-6887.

59 Hagen, W. R., van Berkel-Arts, A., Krüse-Wolters, K. M., Voordouw, G. and Veeger, C. (1986) *FEBS Lett.* **203**, 59-63.

60 Morton, T. A., Runquist, J. A., Ragsdale, S. W., Shanmugasundaram, T., Wood, H. G. and Ljungsdahl, L. G. (1991) *J. Biol. Chem.* **266**, 23824-23828.

61 Hagen, W. R., Van Berkel-Arts, A., Krüse-Wolters, K. M., Dunham, W. R. and Veeger, C. (1986) *FEBS Lett.* **201**, 158-162.

62 Devereux, R., Delaney, M., Widdel, F. and Stahl, D. A. (1989) *J. Bacteriol.* **171**, 6689-6695.

63 Spormann, A. M. and Thauer, R. K. (1988) *Arch. Microbiol.* **150**, 374-380.

64 Henikoff, S. (1987) *Methods Enzymol.* **155**, 156-165.

Chapter 5

The primary structure of a protein containing a putative [6Fe-6S] prismane cluster from *Desulfovibrio desulfuricans* (ATCC 27774)

Jack P. W. G. Stokkermans, Willy A. M. van den Berg, Walter M. A. M. van Dongen and Cees Veeger.

Published in: *Biochimica Biophysica Acta* (1992) **1132**, 83-87.

Note: The nucleotide sequence reported in this paper has been submitted to EMBL and is available under accession number Z11975

I. Summary

The gene encoding a protein containing a novel iron sulfur cluster ([6Fe-6S]) has been cloned from *Desulfovibrio desulfuricans* ATCC 27774 and sequenced. An open reading frame was found encoding a 545 amino acid protein (M_r 58496). The amino acid sequence is highly homologous with that of the corresponding protein from D. vulgaris (Hildenborough) and contains a Cys-motif that may be involved in coordination of the Fe-S cluster.

II. Introduction

In the past five years Fe-S clusters in proteins have been described that exhibit unusual EPR spectra resulting from high-spin systems ($S \geq 7/2$) [1-14]. In these proteins, which function mostly as redox enzymes, these so called "superclusters" are thought to be essential for enzymatic activity [1-6]. Recently, a protein was isolated from *D. vulgaris* that contains only six irons, probably coordinated in a [6Fe-6S] cluster [7-12]. This protein, named "prismane protein", for which no physiological function has been found yet, is considered as a model protein for the study of superclusters. It can occur in four different redox states. In the "as isolated" protein, the supercluster is in the one-electron-reduced $[6Fe\text{-}6S]^{5+}$ state and exists in two magnetic forms: 10% is low spin (S=1/2) and approximately 90% is in an unusual high-spin state (S=9/2). In the full-reduced $[6Fe\text{-}6S]^{3+}$ state, the cluster exhibits a S=1/2 EPR spectrum that is similar to that of a [6Fe-6S] prismane core model compound [15]. The two electron-reduced $[6Fe\text{-}6S]^{4+}$ and the fully oxidized states are EPR-silent and have S=0 or even spin systems. From Mössbauer experiments it was indicated that two irons, at opposite ends of the prismane-cluster may be ligated by nitrogen [10]. The protein was further characterized as a monomeric, cytoplasmic protein with a molecular mass of 60161 Da [11,12]. The amino acid sequence has been determined from the nucleotide sequence of the gene [12].

A similar protein has recently been isolated from *D. desulfuricans* (ATTC 27774) by Moura et al. [13,14] This protein also showed a S=9/2 EPR signal in the one electron reduced state. However, these authors claimed that the *D. desulfuricans* protein contains two [6Fe-6S] clusters; one with a S=0, the other with S=9/2 spin system in the "as isolated" state [14].

```
         10        20        30        40        50        60        70        80        90
CATGCACAGATCATACGCAACATGCTTTCGGTATTCGCCCGCAAGGCTTTTTTCTGACGCGCCGCGTATCACTGCTTTCATCGGGCAG
CTGCGTCGCAAGATCGCCTCTTTACTGCTTGAGCACCGCAACCCTGACGGCTCGCTGAATCTGCTCATGAACCGTGAACAGATGGCCGA
TTTCTGGGCGTCACGCGGCCCTCGCTTTCACGCGAGCTGGCAGCCATGCGTGATGAAGGCCTGCTGCACATCAACGGAAAAAACATCTC
CTGCCTCATCCCGACCGCCTGAAAGAATTTTACGATTATTCTGTTCCGTAACAATTGTTACGGACGTCCACACTCCCCCTTGGTATACA
GAATCATCAAAAGCAAGGAGGGAACCATGCAACGCCTGCTCAAGAATCTGGATTACGCCACCGCGCTGCCCCTGGCCGCGCAGGTGGAT
GCCAGCCGGGCCAGGTAGCCAGCAAGACGCTTGTTCAGAATGATGCCGTGGGTATAACCCTTTTCGCCTTTGACAAGGGCGAAAAAATC
GCGCCCACACCTCCACGGGCGACGCCTTTGTGCTGGCCCTTGAAGGACAGGGGCAAGTGACCATCAACGGACAGACCAGCCCCCTCAAG
CCGGAGAGTCCATCATCATGCCCGCAGGGCAGCCCCATTCCGTCAGCGCTGTGGAGCGCTTCAAGATGCTGCTGGTTGTCATTTTTCCG
CGGAGTAACGCCAGCCCGCCCACCGGCGGCACAAAAGCGGCCCGGACAGTTTTCATACCGTAAACATTTTTTCATCCCCGATGCCCGCC
        730       740       750       760       770       780       790       800       810
                                                            M  S  N  A  M  F  C  Y  Q  C  Q
GCTGCGGGTTTCACCCAAGCCTCCGCCGCAACGGAGCCATATAAGGACAAGACACCATGAGTAATGCCATGTTCTGCTACCAGTGCCAG
        820       830       840       850       860       870       880       890       900
E  T  V  G  N  K  G  C  T  Q  V  G  V  C  G  K  K  P  E  T  A  A  L  Q  D  A  L  I  Y  V
AAACCGTGGGTAACAAAGGCTGCACCCAGGTAGGCGTGTGCGGCAAAAAGCCTGAAACAGCCGCCCTTCAGGACGCGCTGATCTATGTG
        910       920       930       940       950       960       970       980       990
T  K  G  L  G  Q  I  A  T  R  L  R  A  E  G  K  A  V  D  H  R  I  D  R  L  V  T  G  N  L
CCAAGGGCCTCGGCCAGATCGCCACGCGCCTGCGCGCCGAAGGCAAGGCCGTTGACCACAGGATAGACCGCCTGGTTACCGGCAACCTG
       1000      1010      1020      1030      1040      1050      1060      1070      1080
F  A  T  I  T  N  A  N  F  D  D  D  I  L  A  E  R  V  R  M  T  C  A  A  K  K  E  L  A  A
TTGCCACCATCACCAATGCCAACTTTGACGACGACATCCTTGCCGAGCGTGTGCGCATGACCTGTGCCGCCAAAAAGGAACTGGCCGCG
       1090      1100      1110      1120      1130      1140      1150      1160      1170
S  L  T  D  K  S  G  L  S  D  A  A  L  W  E  A  S  E  K  S  A  M  L  A  K  A  G  T  V  G
CCCTTACCGACAAGAGCGGCCTCAGCGATGCAGCCTTGTGGGAAGCATCCGAAAAGTCCGCCATGCTGGCCAAGGCCGGAACCGTAGGC
       1180      1190      1200      1210      1220      1230      1240      1250      1260
V  M  A  T  T  D  D  D  V  R  S  L  R  W  L  I  T  F  G  L  K  G  M  A  A  Y  A  K  H  A
GTTATGGCCACCACCGATGATGATGTGCGCTCCCTGCGCTGGCTCATCACCTTTGGGCTCAAGGGCATGGCGGCCTACGCCAAACATGCG
       1270      1280      1290      1300      1310      1320      1330      1340      1350
D  V  L  G  K  H  E  N  S  L  D  A  F  M  Q  E  A  L  A  K  T  L  D  D  S  L  S  V  A  D
GATGTGCTTGGCAAGCATGAAAACAGCCTTGACGCCTTCATGCAGGAAGCCCTTGCCAAAACCCTGGATGACAGCCTGAGCGTGGCCGAC
       1360      1370      1380      1390      1400      1410      1420      1430      1440
L  V  A  L  T  L  E  T  G  K  F  G  V  S  A  M  A  L  L  D  A  A  N  T  G  T  Y  G  H  P
CTTGTGGCCCTGACCCTTGAAACGGGCAAGTTCGGCGTATCGGCCATGGCCCTGCTGGATGCTGCCAATACCGGTACCTACGGCCACCCA
       1450      1460      1470      1480      1490      1500      1510      1520      1530
E  I  T  K  V  N  I  G  V  G  S  N  P  G  I  L  I  S  G  H  D  L  R  D  L  E  M  L  L  K
GAAATTACCAAGGTCAACATCGGCGTGGGCAGCAATCCCGGCATCCTCATTTCCGGGCATGACCTGCGCGACCTTGAAATGCTGCTCAAG
       1540      1550      1560      1570      1580      1590      1600      1610      1620
Q  T  E  G  T  G  V  D  V  Y  T  H  S  E  M  L  P  A  H  Y  Y  P  A  F  K  K  Y  A  H  F
CAGACCGAAGGCACAGGCGTTGACGTGTACACCCACTCTGAAATGCTGCCCGCCCATTACTACCCTGCCTTCAAGAAGTACGCGCACTTC
       1630      1640      1650      1660      1670      1680      1690      1700      1710
K  G  N  Y  G  N  A  W  W  K  Q  K  E  E  F  E  S  F  N  G  P  V  L  L  T  T  N  C  L  V
AAGGGCAACTACGGCAATGCATGGTGGAAACAGAAAGAAGAATTTGAAAGCTTTAACGGCCCCGTGCTGCTGACCACCAACTGCCTTGTG
       1720      1730      1740      1750      1760      1770      1780      1790      1800
P  P  K  D  S  Y  K  D  R  V  Y  T  T  G  I  V  G  F  T  G  C  K  H  I  P  G  E  I  G  E
CCGCCCAAGGACAGCTACAAGGACCGCGTGTACACCACCGGCATCGTGGGTTTTACGGGCTGCAAGCATATCCCCGGTGAAATCGGCGAA
       1810      1820      1830      1840      1850      1860      1870      1880      1890
H  K  D  F  S  A  I  I  A  H  A  K  T  C  P  A  P  T  E  I  E  S  G  E  I  I  G  G  F  A
CACAAGGACTTCAGCGCCATCATCGCCCATGCCAAGACCTGTCCCGCGCCTACGGAAATCGAATCCGGCGAAATCATCGGCGGCTTCGCG
       1900      1910      1920      1930      1940      1950      1960      1970      1980
H  N  Q  V  L  A  L  A  D  K  V  I  D  A  V  K  S  G  A  I  K  K  F  V  V  M  A  G  C  D
CACAATCAGGTACTGGCCCTGGCCGACAAGGTGATTGACGCGGTCAAATCCGGCGCCATCAAAAAGTTCGTGGTCATGGCCGGCTGCGAC
       1990      2000      2010      2020      2030      2040      2050      2060      2070
G  R  A  K  S  R  S  Y  Y  T  D  F  A  E  G  L  P  K  D  T  V  I  L  T  A  G  C  A  K  Y
GGCCGCGCCAAGTCCCGCAGCTACTACACCGACTTTGCCGAAGGCCTGCCCAAAGACACGGTCATCCTTACCGCCGGTTGCGCCAAATAT
       2080      2090      2100      2110      2120      2130      2140      2150      2160
R  Y  N  K  L  N  L  G  D  I  G  G  I  P  R  V  L  D  A  G  Q  C  N  D  S  Y  S  L  A  V
CGCTACAACAAGCTCAACCTGGGTGACATCGGCGGCATCCCGCGCGTACTGGACGCCGGGCAGTGCAACGACTCCTACTCCCTGGCCGTC
       2170      2180      2190      2200      2210      2220      2230      2240      2250
I  A  L  K  L  K  E  V  F  G  L  E  D  V  N  D  L  P  I  V  Y  N  I  A  W  Y  E  Q  K  A
ATCGCCCTCAAGCTCAAGGAAGTATTCGGCCTCGAGGACGTCAACGACCTGCCCATCGTCTACAATATCGCCTGGTACGAGCAGAAGGCC
       2260      2270      2280      2290      2300      2310      2320      2330      2340
V  I  V  L  L  A  L  L  S  L  G  V  K  N  I  H  L  G  P  T  L  P  A  F  L  S  P  N  V  A
GTTATCGTGCTGCTGGCCCTGCTGAGCCTCGGCGTGAAGAATATCCACCTCGGACCGACGCTGCCCGCCTTCCTTTCGCCCAACGTGGCC
```

```
     2350      2360      2370      2380      2390      2400      2410      2420      2
 K  V  L  V  E  Q  F  N  I  G  G  I  T  S  P  Q  D  D  L  K  A  F  F  G  *
AAGGTGCTGGTGGAACAGTTCAACATCGGCGGCATCACCAGTCCGCAGGACGACCTCAAGGCGTTCTTCGGCTAACCCTTCCCACAA
     2440      2450      2460      2470      2480      2490      2500      2510      2
ACCCAAACAAGCAGAAGGCGCGCACCGGAGTGCGCGCCTTTTATTTCTTTGGTCCAGATTAACGCTGGAATGTGAAGTATTTCAACG
TCATTCTGCCCAAACTTGCAATATCCGGCAGAATCCACGCCGTATTGCGGCGAGCTGCACTCTCGTGCAGCGTTAGCCCATGTAAAC
TTCATAAGGTAAACAGCGCTACCGCAGTTCGCCTTGCAAATAACCGACGATTACGCCCAGACAAGGCACACCTGCGCCAATCCGCCG
GGCCTTGCCCTCAAGCCCGTGCGCGCCCTTCAGGCATAGCATGC
     2800      2810      2820      2830
```

FIGURE 1 : Nucleotide sequence of the 2834 bp *Sph*I fragment of plasmid pWBP81 and derived amino acid sequence for the prismane protein from D. desulfuricans (ATCC 27774). A possible ribosome binding site, just ahead of the translational start (position 868) is underlined.

III. Results and Discussion

Here we describe the cloning and sequencing of the gene coding for the *D. desulfuricans* protein. Genomic DNA from *D. desulfuricans* ATCC 27774 was cleaved with several restriction endonucleases and hybridized with a 1.9 kb *Bam*HI/*Pst*I fragment containing the gene for the prismane protein from *D. vulgaris*. [12]. The *D. desulfuricans* gene was located on a 2.8 kb *Sph*I fragment , that was cloned in plasmid pUC9 and denoted by pWBP81. Fragments for sequencing were generated by either exonuclease III and S_1 nuclease [15] or restriction endonuclease cleavage. These fragments were subcloned into M13 vectors and sequenced by the dideoxy method [16] using *Taq* DNA polymerase (Gibco-BRL) [17]. Both strands were sequenced entirely. Sequences were analyzed with computer programs developed by Staden [18,19].

In Fig. 1., the sequence of the 2.8 kb *Sph*I fragment is shown. It contained an open reading frame between the ATG-codon around position 868 and a TAA-codon around position 2503, encoding a protein of 545 amino acids (M_r 58496 Da). This large open reading frame is preceded by a putative ribosome binding site (AGGA) around position 855. The molecular mass of the derived polypeptide, its amino acid composition and N-terminal amino acid sequence are in agreement with data published for the purified protein [14] (Fig. 2). From this its can be concluded that the gene encodes the protein from *D. desulfuricans* (ATCC 27774) described by Moura et al. [14].

A comparison of the amino acid sequences of the *D. vulgaris* and *D. desulfuricans* proteins shows a high degree of conservation, except for the region between 40-140 (Fig. 3). In the two proteins, 66% of the residues are identical. This proves that the proteins are homologous. The homology is highest in the C-terminal two thirds (residues 140-540) and in the first 45 residues.

2A

Amino acid	From Amino acid analysis[a]	From Nucleotide Sequence
D+N	56	57
Q+E	47	40
S	28	26
T	31	31
P	18	18
H	11	13
K	38-39	41
R	14	14
G	54	50
A	60	61
V	38	39
I	26	30
L	55	56
Y	15	17
F	16	20
W	n.d.[b]	5
M	13	11
C	11-12	11

2B

```
                 1             5                  10
Protein   (Met/Ser)-Asn-Ala-Met-Phe-Ala-Tyr-Gln-(   )-Gln
                 |   |   |   |   |       |   |       |
DNA       Met-Ser -Asn-Ala-Met-Phe-Cys-Tyr-Gln-Cys-Gln

              11              15                  20
Protein       Glu-Thr-Val-Gly-Asn-Lys-Gly-(   )-Thr-Gln
               |   |   |   |   |   |   |       |   |
DNA           Glu-Thr-Val-Gly-Asn-Lys-Gly-Cys-Thr-Gln

              21              25                  30
Protein       Val-Gly-Val-Ala-Gly-Lys-Lys-Pro-Glu-Thr
               |   |   |       |   |   |   |   |   |
DNA           Val-Gly-Val-Cys-Gly-Lys-Lys-Pro-Glu-Thr

           31                 35
Protein       Ala-(   )-Leu-Gln-(Asp/Lys)-Ala
               |         |   |      |      |
DNA           Ala-Ala-Leu-Gln-     Asp    -Ala
```

FIGURE 2 : Comparison of the amino acid composition (2A) and N-terminus (2B) of the native protein of *D. desulfuricans* ATTC 27774 and the DNA-sequence-derived polypeptide. ([a])Data from the native protein were taken from Moura et al. [14]; ([b]) not determined.

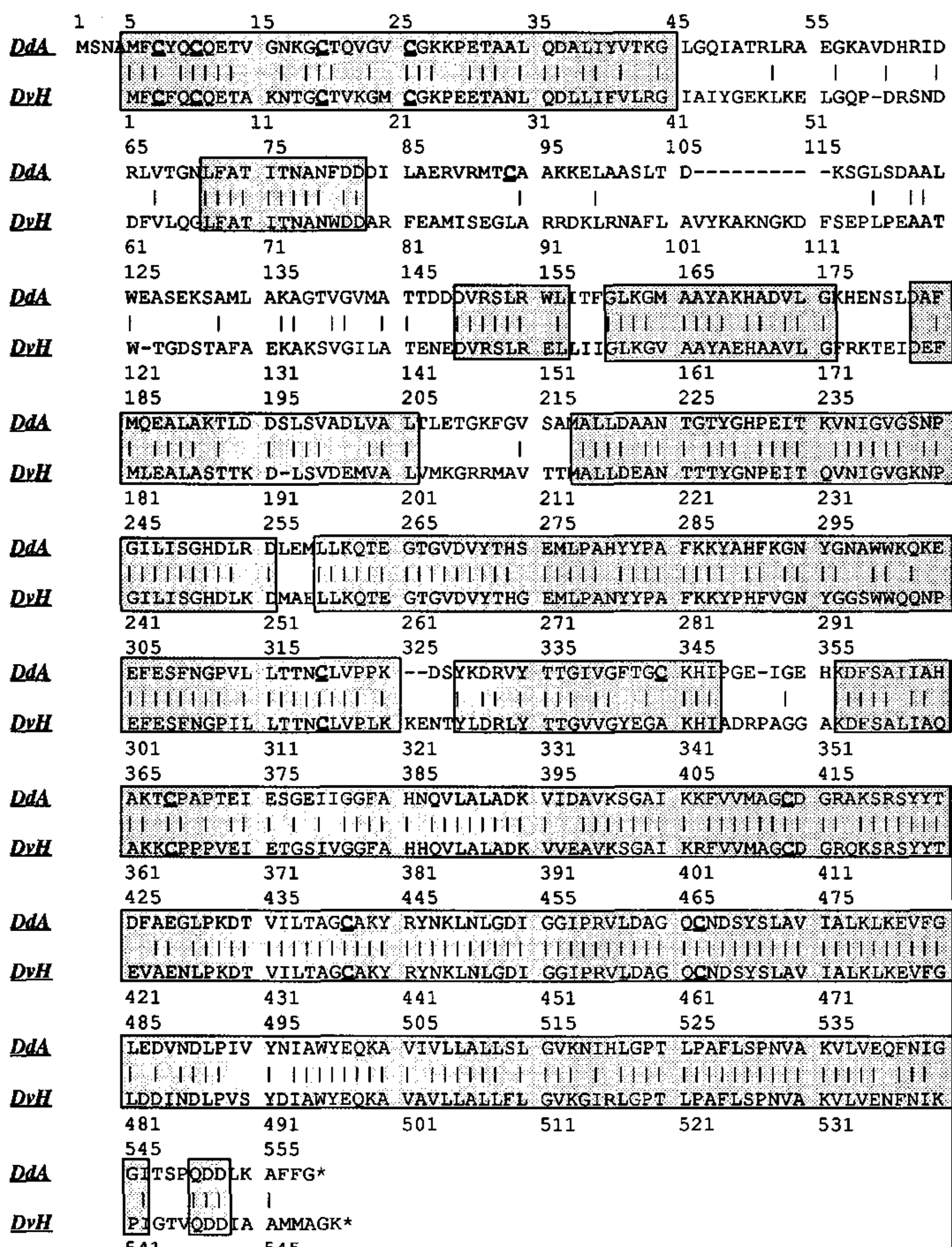


FIGURE 3 : Comparison of the prismane protein from *D. desulfuricans* (ATCC 27774) and *D. vulgaris* (Hildenborough). Alignment of the *D. desulfuricans* (ATCC 27774) (denoted by <u>*DdA*</u>) and the *D. vulgaris* (Hildenborough) (denoted by <u>*DvH*</u>) protein sequences. Identical residues are presented by (|). Gaps are presented by (-). Cysteines are marked bold and underlined. Homologous segments are presented by grey boxes.

```
Prismane pr. (D. de)    5    MFCYQCQETVGN-KGCTQ---VGVCG
                             ||| |||||  |  |||     | ||
Prismane pr. (D.vu)     1    MFCFQCQETAKN-TGCTV---KGMCG
                               |  |       ||      || ||
CO-dehydr.   (M.so)    65    DTCTLCTYGPCDLTGNK----KGACG
                               |  |  |||          :| ||
CO-dehydr.   (C.th)    61    ICCRFCMAGPCRIKATDGPGSRGICG

Consensus                      CXXC<-- X12-14 -->KGXCG
```

FIGURE 4: Homology between the prismane proteins from *D. vulgaris* (Hildenborough) and *D. desulfuricans* (ATCC 27774) and CO dehydrogenases from *M. soehngenii* and *C. thermoaceticum*. Residues identical in the aligned sequences are represented by (|). Numbers refer to the corresponding positions in the proteins.

Cysteines and histidines may be involved in the coordination of the Fe-S cluster [10,14]. The nine cysteines in the *D. vulgaris* protein are all conserved in the *D. desulfuricans* protein; this latter protein contains two additional cysteines (residues 92 and 342). The conserved N-terminal domain of both proteins contains a Cys-motif (Cys-Xaa_2-Cys-Xaa_{12}-Gly-Xaa-Cys-Gly) that is also found in the α subunit of CO dehydrogenase from *Methanothrix soehngenii* [20] and occurs in a slightly modified form in the α subunit of CO dehydrogenase from *Clostridium thermoaceticum* (Cys-Xaa_2-Cys-Xaa_{14}-Gly-Xaa-Cys-Gly) [21] (Fig. 4). The sequence (Cys-Xaa_2-Cys) is also present in [4Fe-4S] cluster binding metalloproteins, but is in these proteins part of a (Cys-Xaa_2-Cys-Xaa_2-Cys(-Xaa_3-Cys)) motif. Remarkably, also for the CO dehydrogenase from *M. soehngenii* a supercluster with a S=9/2 spin system was proposed, based on EPR signals with effective g-values of 14.5 and 5.5 [5,6]. It is tempting to speculate that the Cys-motif is involved in ligation of the prismane-like Fe-S cluster. Involvement of histidines in the coordination of this cluster was suggested by Mössbauer spectra of the prismane proteins, which indicated that two of the six irons are ligated to a nitrogen atom [10,14]. Six out of 13 histidines from the *D. desulfuricans* protein are conserved in the *D. vulgaris* protein (Fig. 3), but we have no indications which of these histidines might be involved in iron coordination.

Combination of data from EPR spectroscopy and chemical determinations of iron and acid-labile sulfur contents of the *D. vulgaris* protein led Pierik et al. [10,11] to conclude that this protein contains one Fe-S cluster (probably [6Fe-6S]) that may occur in four different redox states. Moura et al. [14] also measured 6.5±0.9 iron atoms in the *D. desulfuricans* protein by chemical analysis. However, in order to explain their Mössbauer data (the occurrence of a diamagnetic and a paramagnetic system in the "as isolated"

protein), these authors assumed that the iron content of the protein was underestimated twofold, and they proposed two [6Fe-6S] clusters per protein molecule. Based on the high homology in the primary structures of both proteins as presented here, it is unlikely that they contain a different number of Fe-S clusters. Quantitative analysis of EPR spectra measured at different redox potentials by Pierik et al [10] showed that the data can be fully explained with only one complex Fe-S cluster per protein molecule. Therefore, we propose that also the *D. desulfuricans* protein contains only one Fe-S cluster.

III. References

1 Hagen, W. R., Van Berkel-Arts, A., Krüse-Wolters, K. M., Voordouw, G. and Veeger, C. (1986) *FEBS Lett.* **203**, 59-63.

2 Pierik, A. J. and Hagen, W. R.(1991) *Eur. J. Biochem.* **195**, 505-515.

3 Hagen, W. R., Wassink, H, Eady, R. R., Smith, B. E. and Haaker, H. (1987) *Eur. J. Biochem.* **169**, 457-465.

4 Hagen, W. R., Pierik, A. J., Wolbert, R. B. G., Wassink, H., Haaker, H., Veeger, C., Jetten, M. S. M., Stams, A. J. M. and Zehnder, A. J. B. (1991) *J. Inorg. Chem.* **43**, 237.

5 Jetten, M. S. M., Hagen, W. R., Pierik, A. J., Stams, A. J. M. and Zehnder, A. J. B (1991) *Eur. J. Biochem.* **195**, 385-395.

6 Jetten, M. S. M., Pierik, A. J. and Hagen, W. R. (1991) *Eur. J. Biochem.* **202**, 1291-1297.

7 Hagen, W. R., Pierik, A. J., and Veeger, C. (1989) *J. Chem. Soc. Faraday Trans. I* **85**, 4083-4090.

8 Pierik, A. J., Hagen, W. R. and Veeger, C. (1990) *EBEC Reports* **6**, 16.

9 Hagen, W. R., Pierik, A. J. and Veeger, C. (1990) *Ital. Biochem. Soc. Trans.* **1**, 85.

10 Pierik, A. J., Hagen, W. R., Dunham, W. R. and Sands, R. H. (1992) *Eur. J. Biochem.*, **206**, 705-719.

11 Pierik, A. J., Wolbert R. B. G., Mutsaers P. H. A., Hagen, W. R. and Veeger, C., (1992) *Eur. J. Biochem.* **206**, 697-704.

12 Stokkermans, J. P. W. G., Pierik, A. J., Wolbert, R. B. G., Hagen, W. R., van Dongen, W. M. A. M. and Veeger, C. (1992) *Eur. J. Biochem* **208**, 435-442.

13 Ravi, N., Moura, I., Tavares, P., LeGall, J., Huynh, B. -H. and Moura, J. J. G. (1991) *J. Inorg. Biochem.* **43**, 252.

14 Moura, I., Tavares, P., Moura, J. J. G., Ravi, N., Huynh, B. -H., Liu, M. -Y. and LeGall, J. (1992) *J. Biol. Chem.* **267**, 4489-4496.

15 Henikoff, S. (1987) *Meth. Enzymol.* **155**, 156-165.

16 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467.

17 D'Alessio, J. M., Kullman, J., Salyards, A. J. and Campbell, J. H. *Focus* **11:2**, 30-31.

18 Staden, R. (1982) *Nucl. Acids Res.* **10**, 4731-4751.

19 Staden, R. (1984) *Nucl. Acids Res.* **12**, 499-503.

20 Eggen, R. I. L., Geerling, A. C. M., Jetten, M. S. M. and de Vos, W. M. (1991) *J. Biol. Chem.* **266**, 6883-6887.

21 Morton, T. A., Runquist, J. A., Ragsdale, S. W., Shanmugasundaram, T., Wood, H. G. and Ljungsdahl, L. G. (1991) *J. Biol. Chem.* **266**, 23824-23828.

Chapter 6

Overproduction of prismane protein in *Desulfovibrio vulgaris* (Hildenborough): evidence for a second $S=1/2$-spin system in the one-electron reduced state.

Jack P. W. G. Stokkermans, Pieter H. J. Houba, Antonio J. Pierik, Wilfred R. Hagen, Walter M. A. M. van Dongen and Cees Veeger

Published in: *European Journal of Biochemistry* (1992) **210**, 983-988.

I. Summary

The gene encoding the prismane protein from *Desulfovibrio vulgaris* (Hildenborough) was inserted into broad-host-range vector pSUP104. The recombinant plasmid, pJSP104, was transferred to *D. vulgaris* by conjugal plasmid transfer. In the transconjugant *D. vulgaris* cells the prismane protein was 25-fold overproduced. The overproduced prismane protein was characterized by molecular mass, isoelectric point, iron content and spectroscopical properties. Both the iron content and the ultraviolet/visible spectrum are identical to the wildtype protein indicating that iron incorporation in the overproduced protein is complete. EPR spectra of the dithionite-reduced form of the overproduced protein indicated that the Fe-S cluster might occur in a similar structure as found in inorganic model compounds containing a [6Fe-6S] prismane core. The as-isolated overproduced protein showed the presence of a second S=1/2 spin system that was also detected in the corresponding prismane protein from *D. desulfuricans* (ATCC 27774), but not in the protein from wild-type *D. vulgaris*. This additional signal was irreversibly transformed to the "wild-type" high and low spin systems upon two reduction/re-oxidation cycles. It is shown that the EPR-spectroscopy of the overproduced prismane protein is very similar to that of the *D. desulfuricans* enzyme, and, with exception to the second (S=1/2) spin system, to that of the prismane protein from wild-type *D. vulgaris*. Contrary to what has been claimed for the *D. desulfuricans* protein, it is shown here that all data can be fully explained assuming a single [6Fe-6S] cluster, that might be titrated into four different redox states and occurs in up to three different spin systems in the one-electron reduced state.

II. Introduction

The prismane protein from *D. vulgaris* (Hildenborough) was the first example of a protein in which a Fe-S cluster with S=9/2 paramagnetism was discovered [1-3]. Since then, high-spin ("superspin") Fe-S clusters with spin systems S>5/2 were described in a number of enzymes: dissimilatory sulfite reductase from *D. vulgaris* (Hildenborough) [4], CO dehydrogenase from *Methanothrix soehngenii* [5,6] and the nitrogenase MoFe-protein (the P-clusters in this protein) from *Azotobacter vinelandii* [7,8]. However, these latter enzymes all contain large numbers of iron atoms (20-30), arranged in multiple, different clusters. This makes a study of the redox properties and the spectroscopic analysis of the high-spin clusters in these proteins rather complicated. The prismane protein, on the other hand, was shown to contain only six irons and six acid-labile sulfide ions [9].Therefore, this protein was used as a relatively simple model protein for the study of these high-spin

Fe-S clusters, although its function is unknown yet [10]. EPR and Mössbauer spectroscopy of the prismane protein indicate, that the iron and sulfide ions are arranged in a [6Fe-6S] cluster, that might be titrated into four different redox states: the EPR-silent, fully oxidized $[6Fe\text{-}6S]^{6+}$ or 6+ state, the paramagnetic $[6Fe\text{-}6S]^{5+}$ or 5+ state, in which the enzyme is found after aerobic isolation (as-isolated state), the integer-spin $[6Fe\text{-}6S]^{4+}$ or 4+ state and the paramagnetic fully reduced $[6Fe\text{-}6S]^{3+}$ or 3+ state, with midpoint potentials for the transitions of +285, -5 and -165 mV, respectively [10]. In the 5+ state, the cluster occurs in two magnetic forms: approximately 90% of the clusters occurs in a high-spin S=9/2 system and 10% in a S=1/2 spin system. In the 3+ state, the cluster occurs in a S=1/2 spin system; the EPR spectrum of the protein in the 3+ state is reminiscent of inorganic model compounds containing the [6Fe-6S] prismane core [11]. Hence the name "prismane protein" was given to this protein.

Recently, a similar protein was isolated from another *Desulfovibrio* strain, *D. desulfuricans* (ATCC 27774) [12,13]. The *D. vulgaris* and *D. desulfuricans* proteins have very similar EPR spectra in the fully reduced state and for the S=9/2 system in the 5+ state. The S=1/2 EPR of the *D. vulgaris* protein in the 5+ state was characterized by unusual low g-values (1.971, 1.951 and 1.898) [1,10] and the signals broadened above 20 K. Besides a S=1/2 system with signals at similar g-values, the *D. desulfuricans* protein in the 5+ state contains a second S=1/2 spin system, with g-values rather typical for iron-sulfur S=1/2 systems (2.02, 1.98, 1.95). Broadening of this latter spin system was only detected at much higher temperatures [13]. It was shown, that primary structures of the *D. desulfuricans* and the *D. vulgaris* proteins, as derived from the nucleotide sequence of the genes, are homologous, with 66% identical residues [14,15].

Until now, progress in the spectroscopic analysis and crystallization of the prismane protein has been hampered by the limited availability of the protein: only approximately 5 mg of pure protein are obtained from a 240 liter culture of *D. vulgaris* (approximately 200 grams of cells) [9].In recent years several genes that have been isolated from anaerobic sulfate reducing *Desulfovibrio* strains have been cloned in *Escherichia coli* [16-23]. However, in this host some of the proteins that are encoded by these genes, are not expressed in an active form [16,17]. In one such a protein, the Fe-hydrogenase of *D. vulgaris* (Hildenborough), the third, putative [6Fe-6S] cluster, thought to be essential for the enzymatic activity, is not incorporated in the produced protein [16,23].

Here, we describe a system for 25-fold overproduction of the prismane protein in *D. vulgaris*. Paramagnetic properties of the overproduced protein are shown to be different from those of the protein isolated from wildtype *D. vulgaris* (i.e. the S=1/2 system of the 5+ state). A simple reduction/re-oxidation step suffices to convert the overproduced protein into the wildtype form.

FIGURE 1 : Physical map of the recombinant plasmid pJSP104. Presented are the gene encoding the prismane protein from *D. vulgaris* (Hildenborough) (shadowed) and its direction of transcription (arrow), the flanking region of broad-host-range vector pSUP104 (hatched) and the direction of both the *tet* promoter (P_{tet}) from pSUP104 and the putative promoter of the gene for the prismane protein (P_p)

III. Materials and methods

Materials. Restriction endonucleases and T4 DNA ligase were purchased from Bethesda Research Laboratories. DEAE-Sephacel and Superdex G-75 were from Pharmacia, Biogel HTP from BioRad and ferene from Aldrich. Other chemicals were obtained from Merck. All chemicals used were of analytical grade.

Bacterial strains and growth conditions. *Escherichia coli* strain TG2, (*recA$^-$*, *Δ(lac-pro), thi, supE,* [Res$^-$ Mod$^-$ (k)], F' (*traD36 , proA$^+$B$^+$, lacI^q* ZΔM15) was used as a host for the construction of recombinant plasmids [24]. *E. coli* strain S17-1, (*thi, pro, hsdR$^-$, hsdM$^+$, recA,* RP4-2(Tc::Mu, Km::Tn7) [25] was used for conjugal transfer of the plasmids. This strain contains, integrated in its genome, the trans-acting mobilization functions that are required for conjugal plasmid transfer [26]. *E. coli* strains were grown in TY medium at 37°C [27]. *Desulfovibrio vulgaris* (Hildenborough) (NCIMB 8303) was cultured anaerobically as described by van den Berg et al. [28].

Molecular biology techniques. All standard DNA operations were performed as described [27]. Plasmid pJSP104 (Fig. 1) was constructed by ligating the insert of plasmid pJSP9 containing the gene encoding the prismane protein [14] into the IncQ vector pSUP104 [26]. Therefore, plasmid pJSP9 was digested with *Bam*HI and *Sal*I and the 1.9 kb insert was isolated after electrophoresis in a 1% (by vol.) agarose gel [27]. Vector pSUP104 was digested with the same restriction enzymes and the 1.9 kb fragment was ligated into pSUP104. The resulting recombinant plasmid, pJSP104, was introduced into *E. coli* S17-1 [25]. *E. coli* S17-1 (pJSP104) was mated with *D. vulgaris* (Hildenborough) as described previously [29]. The plasmid-containing strain, *D. vulgaris* (pJSP104) was further grown in media containing 10 µg/ml chloramphenicol [28].

Protein purification and analysis. The prismane protein of *D. vulgaris* (pJSP104) was purified in a four step FPLC purification method involving an ion-exchange step on DEAE-Sephacel, gel-filtration on Superdex G-75 , hydroxyapatite chromatography on Biogel HTP and ion-exchange on MonoQ according to Pierik et al.

[9]. Protein concentrations were determined with the microBiuret method [30] after trichloroacetic acid/deoxycholate precipitation [31]. The purity of the isolated protein was checked by ultraviolet/visible spectroscopy (A_{400}/A_{280} ratio), electrophoresis in 0.1% SDS, 15% polyacrylamide gels [32] and flat bed isoelectric focussing on a LKB Ultrophor electrophoresis unit at 4°C using the protein markers described in [33]. The iron content of the purified protein was determined with ferene as described previously [9].

Spectroscopy. Ultraviolet/visible spectra of protein samples were recorded with an Aminco DW-2000 spectrophotometer interfaced with an IBM computer. Single wavelength absorbances for the determination of A_{400}/A_{280} ratios of purified protein samples and iron-content determinations were made with a Zeiss M4QIII spectrophotometer with a PI-2 logarithmic converter.

Normal-mode X-band EPR spectroscopy on a Bruker 200 D EPR spectrometer with peripheral instrumentation and data acquisition/quantification, were performed as described before [10]. Redox titrations were performed anaerobically at 25°C. The as-isolated protein in 25 mM Hepes, pH 7.5 was incubated with the same mixture of mediators as described previously [10] and freshly prepared sodium dithionite and $K_3Fe(CN)_6$ (2-200 mM stock solutions) in 0.5 M anaerobic Hepes, pH 7.5 were added as reductant and oxidant, respectively.

The reduction/re-oxidation step on the as-isolated protein in 25 mM Hepes, pH 7.5 was also performed anaerobically at 25°C by adding freshly prepared sodium dithionite (to 1 mM), after 20 minutes followed by addition of N-methyl phenazonium methosulfate (to 10 mM; Sigma) in 0.5 M anaerobic Hepes, pH 7.5. Afterwards, the protein sample was desalted aerobically on a Biogel P-6DG column and concentrated on a Centricon microconcentrator (Amicon).

IV. Results and Discussion

Overproduction of prismane protein in *D. vulgaris* (pJSP104) and purification of the overproduced prismane protein. In an initial attempt to obtain production of large amounts of the prismane protein, the gene encoding this protein was cloned in *E. coli* downstream of the *lac*Z promoter of vector pUC9 [34]. Induction of the *lac*Z promoter with isopropyl-thio-β-D-galactoside indeed resulted in the production of high amounts of the protein; however the protein accumulated in *E. coli* in insoluble form (inclusion bodies) without the iron-sulfur cluster incorporated as was demonstrated by EPR spectroscopy (data not shown).

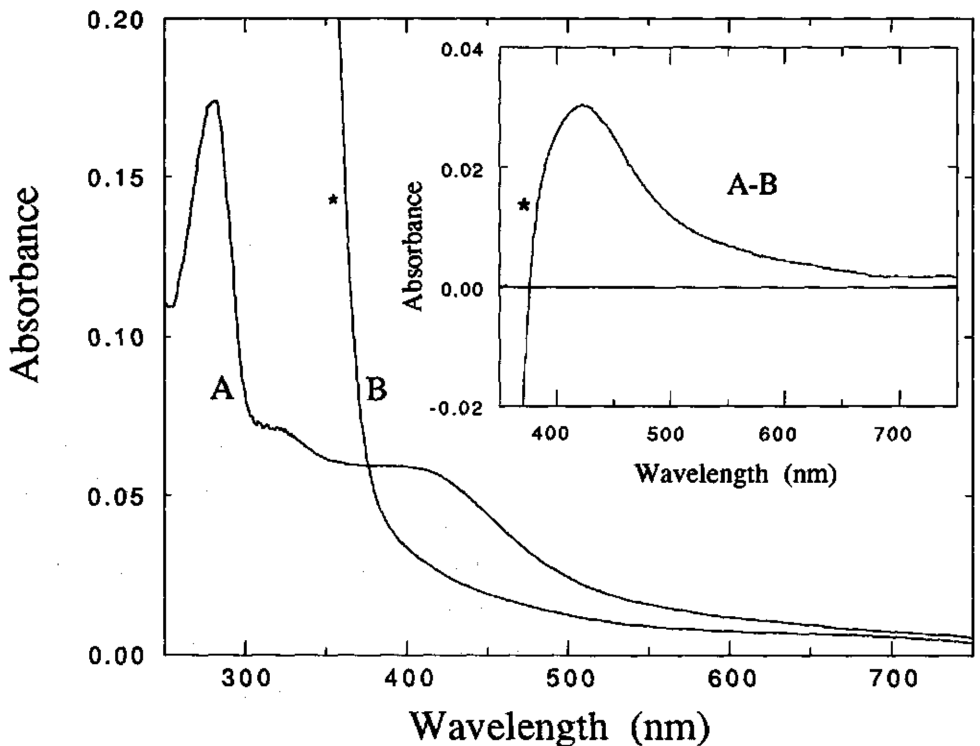


FIGURE 2 : Ultraviolet/visible absorption spectra of the overproduced prismane protein. Shown are (A) the as-isolated protein (0.11 mg/ml, 50 mM Hepes, pH 7.5), (B) the dithionite reduced protein and isolated minus dithionite-reduced (inset, A-B). Asteriks denote spectral contributions from excess dithionite and its decomposition products.

Therefore, overproduction in *D. vulgaris* was attempted. The gene for the prismane protein was introduced in the broad-host-range vector pSUP104, downstream of the *tet* promoter, that has been shown to promote transcription in *D. vulgaris* [28]. The gene for the prismane protein in the resulting recombinant plasmid pJSP104 (Fig. 1) is preceded by its own promoter and the tet promoter in tandem. pJSP104 was introduced into *D. vulgaris* by conjugal transfer. Western blotting of cell extracts of wild-type *D. vulgaris* and *D. vulgaris* (pJSP104) indicated, that introduction of the recombinant plasmid resulted in considerable overproduction of the prismane protein in *D. vulgaris* (pJSP104); the overproduced protein accumulates in a fully soluble form (not shown). The prismane protein produced in *D. vulgaris* (pJSP104) was isolated from 20-liter cultures of this strain and characterized by a number of methods, with special attention to the incorporation of the full complement of iron ions. (In next sections, the prismane protein

isolated from wild-type *D. vulgaris* and from *D. vulgaris* (pJSP104) are referred to as the "wild-type" and the "overproduced" prismane protein, respectively). Two isolations of the protein from independent cultures of *D. vulgaris* (pJSP104) yielded approximately 60 mg pure prismane protein per 100 g cells. This yield is approximately 25-fold higher than that obtained after isolation of the protein from wild-type *D. vulgaris*. Assuming 50% recovery for the purification [9], the prismane protein comprises approximately 2.5% of the soluble protein in *D. vulgaris* (pJSP104). The mobility of the overproduced protein in SDS-polyacrylamide gels and in iso-electric focussing gels (pI= 4.9±0.1) was indistinguishable from that of the wild-type protein.

Iron content of the prismane protein. Fig. 2 shows the ultraviolet/visible absorption spectra of as isolated and dithionite-reduced overproduced prismane protein. The general appearance of both spectra is very similar to that of the wild-type protein [9], with one peak at 280 nm and two shoulders at 320 and 400 nm for the as-isolated protein. The A_{400}/A_{280} ratio of the as-isolated overproduced protein (0.324±0.019; n=2) is comparable to that published for the wild-type *D. vulgaris* protein (0.307±0.031; n=11) [9], suggesting that the overproduced protein might contain the full complement of iron and sulfide ions.

This was reinforced by direct determination of the iron content of the overproduced prismane protein with ferene. A value of 7.2±0.8 (n=2) mol iron/mol protein was determined. Virtually all of this iron is specifically bound in the protein: the amount of adventitious iron was estimated from the intensity of the g=4.3 signal in the EPR spectrum of the as-isolated protein to be only a minor impurity (less than 0.01 spin per molecule).

The value of 7.2±0.8 mol iron/mol protein is higher than that published previously for the wild-type protein (6.3±0.4; n=9) [9]. Calculation of the former value is based on a molecular mass of the protein of 60.1 kDa, as determined from the nucleotide sequence of the gene [14], while the latter value was based on a molecular mass of 52 kDa, as estimated from sedimentation equilibrium experiments. Recalculation of the previously published iron content of the wild-type protein with the molecular mass of the protein as determined by the amino acid sequence, results in an iron content of the wild-type protein of 7.3±0.5 mol iron/mol protein. Consequently, the iron content of overproduced and wild-type protein is similar (7.2 and 7.3 mol/mol protein, respectively). Assuming an inaccuracy of at most 20% in the determination of the iron content, the amount of Fe-ions per molecule protein is 6, 7 or 8. These figures suggest that the prismane protein may contain, besides the six irons in the [6Fe-6S] prismane cluster, one or possibly two additional irons. EPR data show no evidence for redox-active FeS_4 centers (as found in rubredoxins [35], desulforedoxin [36], rubrerythrin [37]) or for [2Fe-2S] clusters.

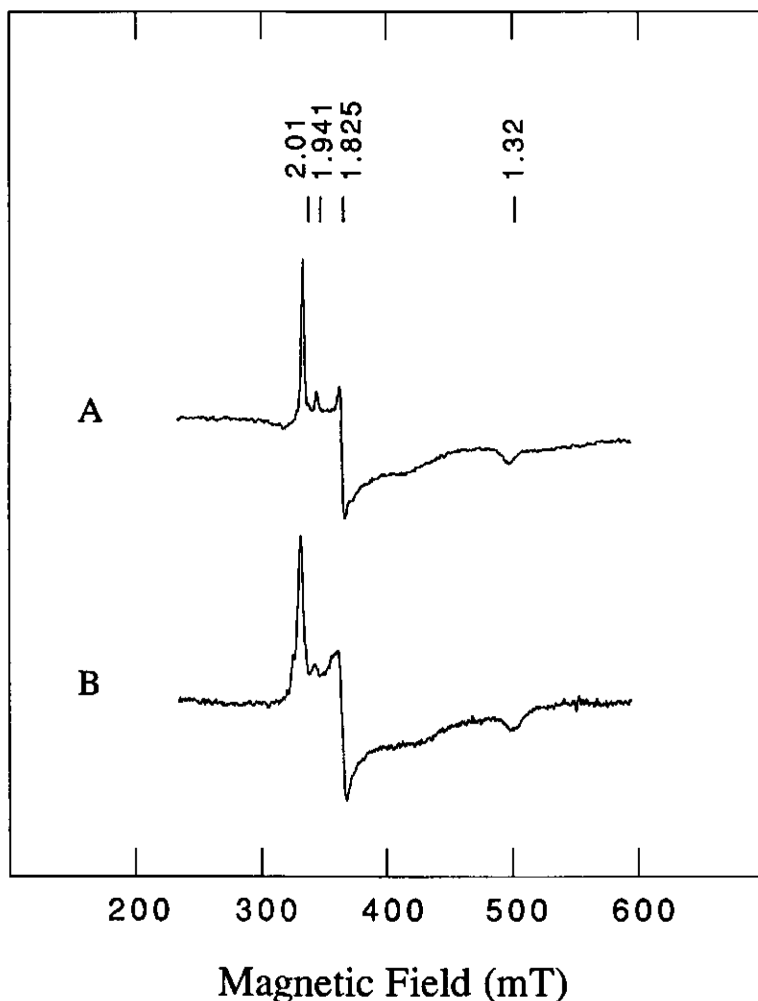


FIGURE 3 : EPR spectra of the dithionite-reduced wild-type (A) and overproduced (B) prismane protein from *D. vulgaris* (3+ state). Experimental conditions : Microwave power, 200 mW; microwave frequency, 9.30 GHz; modulation frequency, 100 kHz; modulation amplitude, 1.6 mT; temperature 22 K; protein concentration 532 μM (trace A) and 106 μM (trace B); buffer 50 mM Hepes, pH 7.5. The receiver gain in trace B was 5-fold higher.

Additional iron(s) might be ligated to the prismane core as proposed for the two Mo atoms in a capped prismane model compound $[Fe_6S_6(Cl)_6\{Mo(CO)_3\}_2]^{3-}$ [38]. It was proposed that $Mo(CO)_3$-moiety did not change the magnetic ground state and the EPR spectra of a typical S=1/2 signal of the prismane core at 10 K. On the other hand, the existence of

additional irons, ligated to the prismane core, was not confirmed by Mössbauer spectra of the wild-type protein [10].

EPR studies of the overproduced prismane protein. Fig. 3 shows EPR spectra of the dithionite-reduced prismane protein from wild-type *D. vulgaris* (trace A) and from *D. vulgaris* (pJSP104) (trace B). In the high-field region both proteins show a rhombic S=1/2 signal with g-values at 2.010, 1.941, 1.825 and 1.32, which have been extensively documented for the wild-type protein [10]. Both spectra exhibit the same "prismane-core specific" pattern, which suggests that the iron-sulfur cluster of the overproduced prismane protein, in the dithionite-reduced state (with the iron-sulfur cluster in the fully reduced $[6Fe\text{-}6S]^{3+}$ state), has the same geometrical conformation as proposed for both the inorganic model core compound [11] and the wild-type protein [1].

Also the S=9/2 spectra of the wild-type and overproduced prismane proteins in the 5+ state are similar (Fig. 4), although the resonances at g-values 8.1 and 6.6 are more intense in the overproduced protein. Variability in the intensities of these signals has also been detected in different purifications of the wild-type protein and has been attributed to two forms of the S=9/2 system with different rhombicities (E/D=0.061 for resonances at g~15.3 and 5.2-5.7 and E/D=0.108 for the g=8.1 and 6.6 lines) [10].

Contrary to the S=1/2 spectra of the fully reduced proteins and the S=9/2 spectra of the proteins in the 5+ state, the S=1/2 spectra of the 5+ state are different for the wild-type and the overproduced as-isolated protein (Fig. 5, traces A and B). Besides the resonance at g=1.898 belonging to the wild-type-like S=1/2 species with g =1.971, 1.951 and 1.898 $(S=1/2)_A$, an overlapping additional axial S=1/2 signal $(S=1/2)_B$ is seen in the overproduced protein, with g-values at 2.007 and 1.975. The temperature dependent relaxation behaviour of these two S=1/2 species is different: $(S=1/2)_A$ broadens significantly at temperatures above 20 K and is not visible above 50 K [10], while set $(S=1/2)_B$ can be observed up to at least 112 K (Fig. 6, trace A). At 112 K the resolution of the $(S=1/2)_B$ g=2.007 and 1.975 features is lost due to broadening.

Contrary to the wild-type prismane protein of *D. vulgaris*, where only the $(S=1/2)_A$ system has been detected in 11 different isolates [10], the corresponding protein from *D. desulfuricans* (ATCC 27774) has also two sets of (S=1/2) spin systems in the 5+ state, with similar g-values and temperature-dependent relaxation behaviour as found for the overproduced prismane protein from *D. vulgaris* [13].

In order to determine the redox behaviour of the $(S=1/2)_B$ system, a redox titration of the overproduced prismane protein was attempted. Therefore, a cocktail of redox mediators [10] was added to the protein. We noticed, however, that the signals belonging to the $(S=1/2)_B$ spin system disappeared immediately after addition of the cocktail, and a spectrum identical to that of the wild-type protein remained. Also after anaerobic reduction

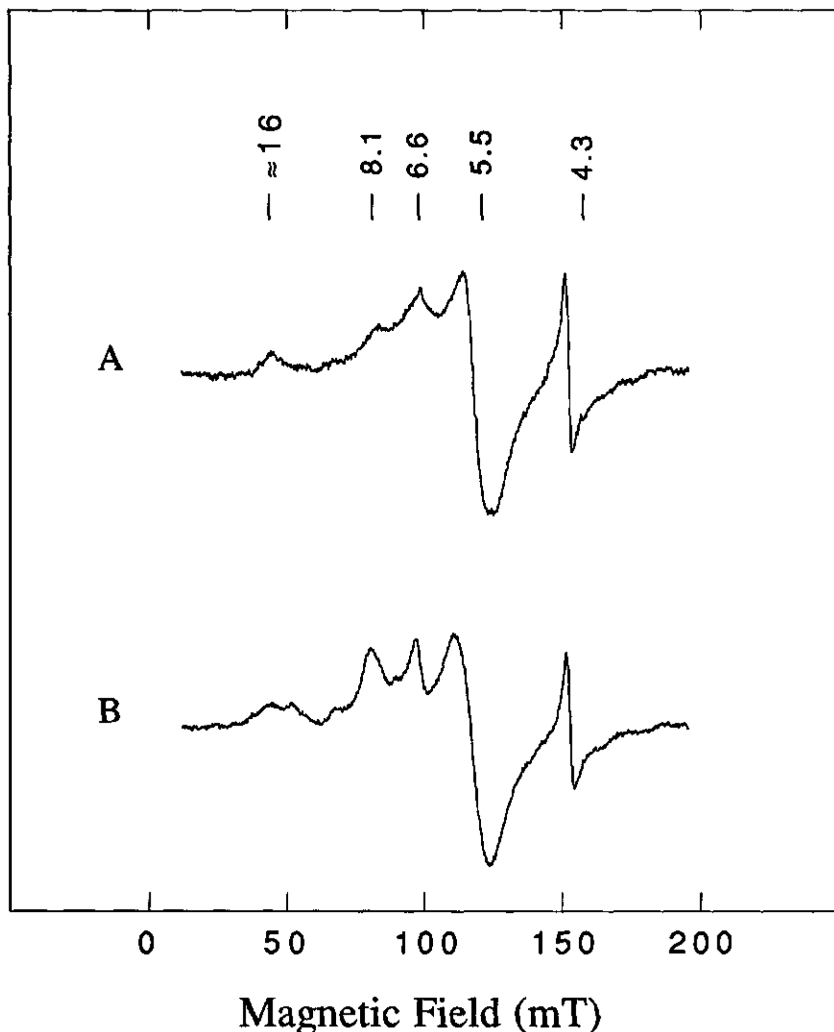


FIGURE 4 : Low-field S=9/2 EPR spectra of the as-isolated wild-type (A) and overproduced (B) prismane protein (5+ state). The spectra were recorded at 18 K and the protein concentrations were 61 μM and 225 μM for the wild-type and the overproduced prismane protein, respectively. All other conditions were as in Fig. 3. The signal at g=4.3 represents adventitious ferric ions.

of the overproduced protein with dithionite, followed by anaerobic re-oxidation to the as-isolated (5+) redox state with N-methyl phenazonium methosulfate (E_m= +80mV), the amplitude of the $(S=1/2)_B$ signal diminished: two reduction/re-oxidation cycles were required to obtain complete conversion to the wild-type spectrum (Fig. 5, traces C and D). Spin intensity of the $(S=1/2)_B$ system is probably redistributed, after reduction/re-

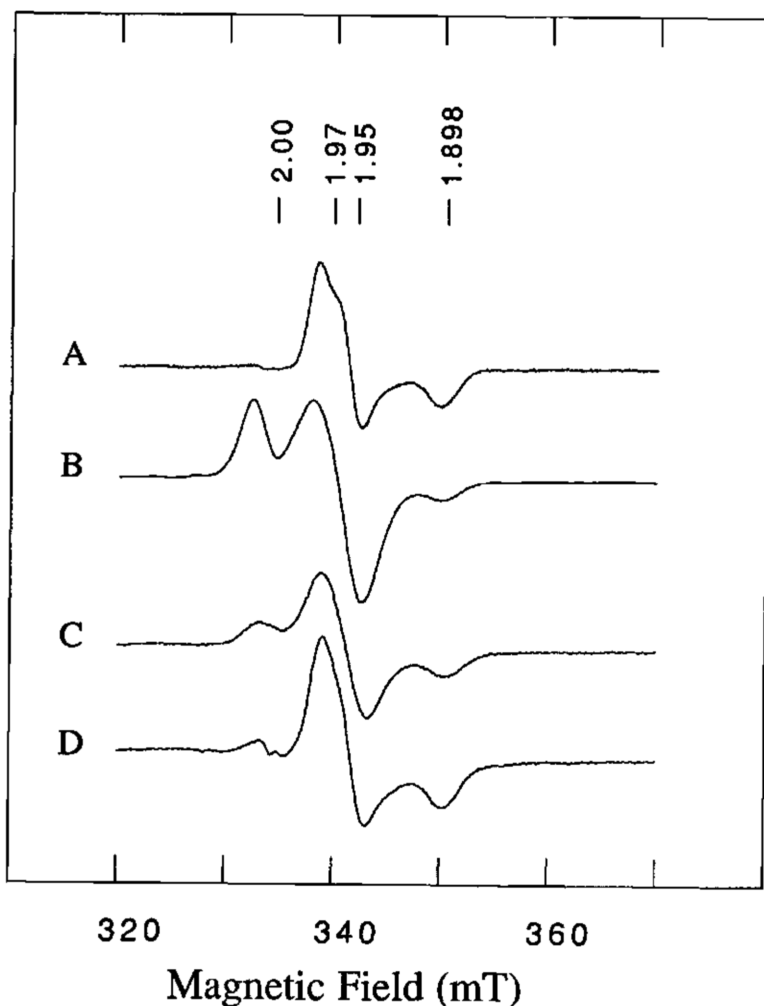


FIGURE 5 : High-field S=1/2 EPR spectra at 12K of one-electron-reduced prismane protein (5+ state). For all spectra, spectrometer settings were the same as in Fig. 3, except for the temperature (12 K) and microwave power (20 mW). Presented are spectra of (A) the as-isolated wild-type prismane protein (protein concentration 532 μM; receiver gain x1), (B) the as-isolated overproduced prismane protein (protein concentration 110 μM; receiver gain x2), (C) the overproduced prismane protein after one reduction/re-oxidation cycle (protein concentration 20 μM; receiver gain x20), (D) the overproduced prismane protein after two reduction/re-oxidation cycles (protein concentration 23 μM; receiver gain x16).

oxidation, over the S=9/2 and the $(S=1/2)_A$ systems, although it is difficult to determine the exact amount of spins/mole protein before and after the reduction/re-oxidation step, because of the inaccuracy to quantify the spins in the S=9/2 system [10]. We estimate that

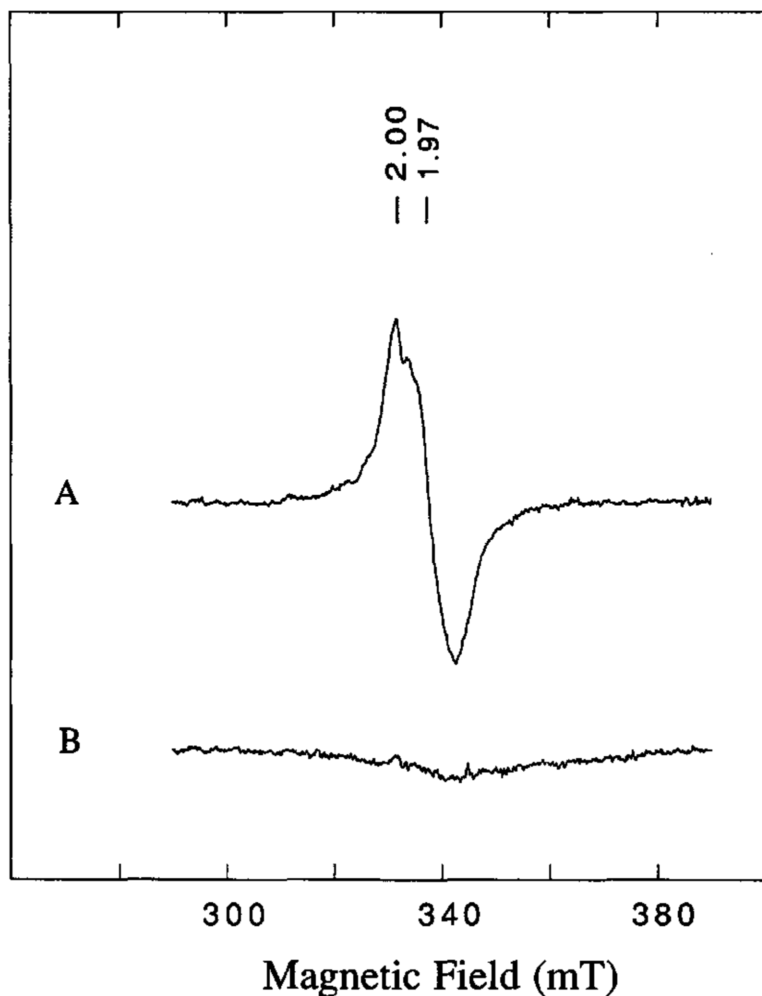


FIGURE 6 : S=1/2 EPR spectra of the overproduced prismane protein in the one-electron-reduced (5+) state at 112 K. The overproduced prismane protein was recorded at 112 K in the as-isolated state (A) and after two reduction/re-oxidation cycles (B). Samples were identical to those in Fig. 5, B and D, respectively. All conditions, except for the temperature, were the same as in Fig. 3.

the as-isolated overproduced protein in the 5+ state has 0.37 spin/mole protein in the S=1/2 systems. Quantitation of the spectrum in Fig. 6 A gives 0.23 spin/mole for the $(S=1/2)_B$ signal; therefore, 0.14 spin/mole should be present in the $(S=1/2)_A$ system. The

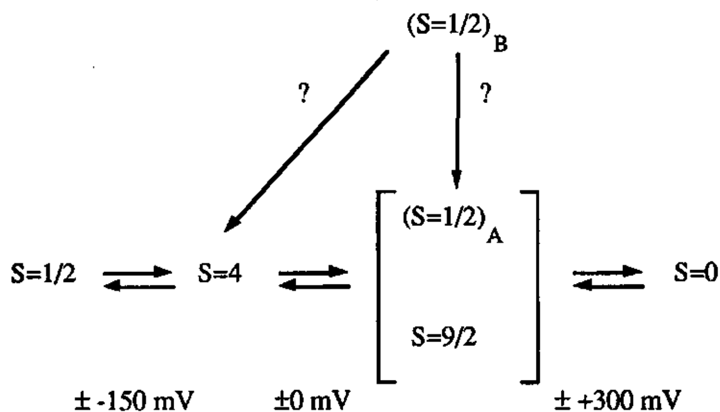


FIGURE 7 : Redox-transitions and accompanying transitions between spin systems in the *D. vulgaris* prismane protein. The $(S=1/2)_B$ system is only present in the overproduced enzyme.

S=9/2 system has 0.6-0.9 spin/mole. After reduction/re-oxidation, spins in the $(S=1/2)_B$ system disappeared, 0.22 spins/mole are found in the $(S=1/2)_A$ system, and 0.6-0.9 in the S=9/2 system. (For wild-type protein in the 5+ state, 0.1 spin/mole in the $(S=1/2)_A$ and 0.9 spin/mol in the S=9/2 system were measured [10]). There is no evidence, that part of the Fe-S clusters is destroyed after reduction/re-oxidation: ultraviolet/visible spectra and A_{400}/A_{280} ratios are identical before and after the reduction/re-oxidation steps. Fig. 7 presents a scheme for the observed transitions between spin systems in the overproducedprotein.

The behaviour of the $(S=1/2)_B$ signal upon reduction/re-oxidation is strongly reminiscent to that of transition metals in hydrogenases (Fe-S clusters in Fe-hydrogenase; Ni in NiFe-hydrogenase). Anaerobic reduction/re-oxidation of both Fe- and Ni-hydrogenases results in the disappearance of EPR signals that are associated with a "resting" or "unready" form of the enzyme and the appearance of new signals associated with a "ready" form of the hydrogenase [39-43]. Unfortunately, as we have not yet found an enzyme activity for the prismane protein, we are not able to conclude, whether the disappearance of the $(S=1/2)_B$ signal upon reduction/re-oxidation is associated with transition of the protein from a resting to an active form. In NiFe-hydrogenases, the transition of the "unready" into the "ready" form and the accompanying spectral changes can be rapidly reversed by addition of O_2 to the "ready" enzyme. This is not the case for the prismane protein: exposure of the reduced/re-oxidized prismane protein to oxygen does not result in reappearance of the $(S=1/2)_B$ signal.

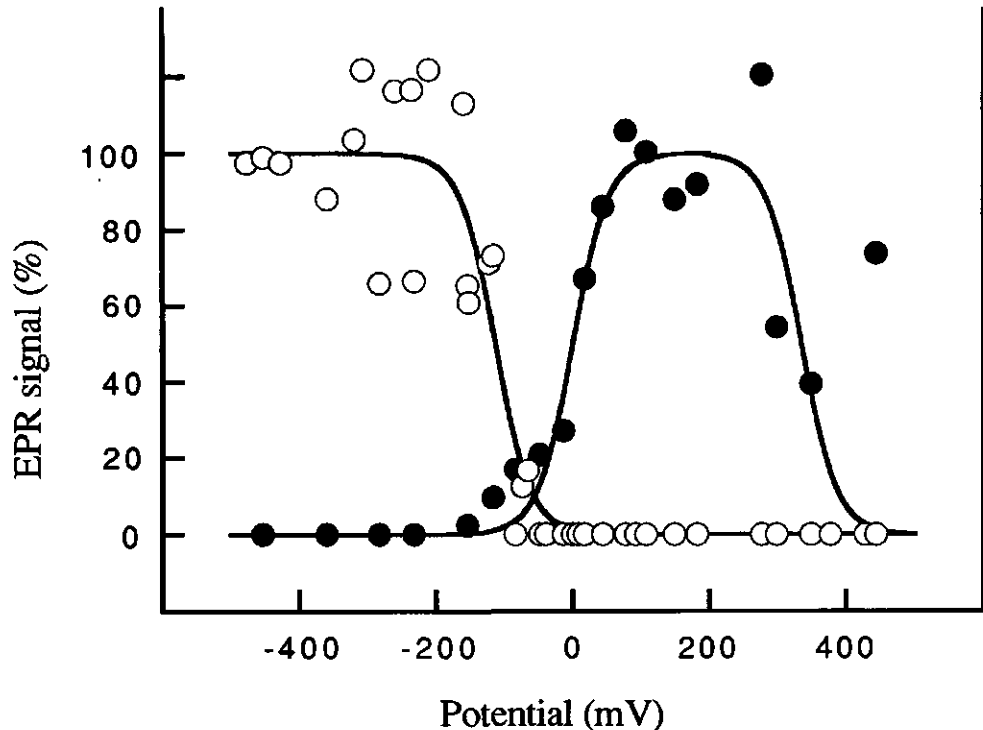


FIGURE 8 : Oxido-reductive titration at pH 7.5 of S=1/2 components in overproduced D. vulgaris prismane protein with dithionite and ferricyanide in the presence of mediators. Open circles represent the prismane-fingerprint signal (the 3+ state) monitored at g= 1.82; black circles represent the signal with all g< 2 (5+ state) monitored at g= 1.951, $(S=1/2)_A$. Titrations were performed in two directions starting from the as-isolated protein, which corresponds approximately to the top of the bell-shaped curve, and starting from the fully pre-oxidized state.

A redox titration of the overproduced prismane protein is shown in Fig. 8. The four redox states described for the wild-type [10] are also found in the overproduced protein. The titrations are fully reversible. Midpoint potentials for the transitions were calculated after fitting the points in Fig. 8 with Nernst equations with n=1. The observed midpoint potentials for the three transitions (-113, 0 and +335 mV) were slightly different than obtained for the wild-type protein (-165, +5 and 285 mV, respectively). These differences are not considered to be significant.

Concluding remarks. The results indicate that the unique Fe-S centre of the prismane protein is still properly assembled even at a 25-fold overproduction of the polypeptide. Physical properties, iron content and ultraviolet/visible spectrum are identical for overproduced and wild-type protein. Although the overproduced protein has a second (S=1/2) spin system in the 5+ state that is not found in the wild-type protein, reduction/re-

oxidation of the overproduced protein converts it into a form, that is spectroscopically indistinguishable from wild-type protein.

The original data published on wild-type *D. vulgaris* and *D. desulfuricans* prismane protein showed differences in the EPR spectroscopy (i.e. two instead of one (S=1/2) spin system for the as-isolated protein for the *D. desulfuricans* protein) [12,13]. Moreover, the interpretation of EPR and Mössbauer data was different: Pierik et al. [10] explained their data assuming only one [6Fe-6S] cluster (which approximately fits the chemically determined Fe and S-content of the protein), while Moura et al. [13] concluded, that the *D. desulfuricans* protein had to contain two [6Fe-6S] clusters (assuming they had underestimated the iron and sulfide content twofold). We previously showed the primary structures of the *D. vulgaris* and *D. desulfuricans* prismane proteins to be homologous [14,15], which renders it unlikely that the two proteins contain a different number of clusters. Our present work supports the view of Pierik et al. [10], that there is only one [6Fe-6S] cluster that occurs in the as-isolated state in two magnetic forms: 60- 90% as a S=9/2 and the remainder as a S=1/2 spin system.

Now, we have also detected a second S=1/2 spin system in the overproduced *D. vulgaris* prismane protein similar to that found in the *D. desulfuricans* protein. The data in this paper indicate, that this second S=1/2 system represents a third form of the same [6Fe-6S] cluster.

V. Acknowledgements

We like to thank Mr. R.B.G. Wolbert and Mr. M. F. J. M. Verhagen for technical assistance and advice. This investigation was supported by the Netherlands Foundation for Chemical Research (SON), with financial aid from the Netherlands Organization for Scientific Research (NWO).

VI. References

1 Hagen, W. R., Pierik, A. J., and Veeger, C. (1989) *J. Chem. Soc. Faraday Trans. I* **85**, 4083-4090.

2 Pierik, A. J., Hagen, W. R. and Veeger, C. (1990) *European BioEnergetic Conference Reports* **6**, 16.

3 Hagen, W. R., Pierik, A. J. and Veeger, C. (1990) *Ital. Biochem. Soc. Trans.* **1**, 85.

4 Pierik, A. J. and Hagen, W. R.(1991) *Eur. J. Biochem.* **195**, 505-516.

5 Jetten, M. S. M., Hagen, W. R., Pierik, A. J., Stams, A. J. M. and Zehnder, A. J. B. (1991) *Eur. J. Biochem.* **195**, 385-391.

6 Jetten, M. S. M., Pierik, A. J. and Hagen, W. R. (1991) *Eur. J. Biochem.* **202**, 1291-1297.

7 Hagen, W. R., Wassink, H., Eady, R. R., Smith, B. E. and Haaker, H. (1987) *Eur. J. Biochem.* **169**, 457-465.

8 Hagen, W. R., Pierik, A. J., Wolbert, R. B. G., Wassink, H., Haaker, H., Veeger, C., Jetten, M. S. M., Stams, A. J. M. and Zehnder, A. J. B. (1991) *J. Inorg. Biochem.* **43**, 237.

9 Pierik, A. J., Wolbert, R. B. G., Mutsaers, P. H. A., Hagen, W. R. and Veeger, C., (1992) *Eur. J. Biochem.* **206**, 697-704.

10 Pierik, A. J., Hagen, W. R., Dunham, W. R. and Sands, R. H. (1992) *Eur. J. Biochem.* **206**, 705-719..

11 Kanatzidis, M. G., Hagen, W. R., Dunham, W.R., Lester, R. K. and Coucouvanis, D. (1985) *J. Am. Chem. Soc.* **107**, 953-961.

12 Ravi, N., Moura, I., Tavares, P., LeGall, J., Huynh, B. -H. and Moura, J. J. G. (1991) *J. Inorg. Biochem.* **43**, 252.

13 Moura, I., Tavares, P., Moura, J. J. G., Ravi, N., Huynh, B. -H., Liu, M. -Y. and LeGall, J. (1992) *J. Biol. Chem.* **267**, 4489-4496.

14 Stokkermans, J. P. W. G., Pierik, A. J., Wolbert, R. B. G., Hagen, W. R., van Dongen, W. M. A. M. and Veeger, C. (1992) *Eur. J. Biochem.* **208**, 435-442.

15 Stokkermans, J. P. W. G., van den Berg, W. A. M., van Dongen, W. M. A. M. and Veeger, C. (1992) *Biochim. Biophys. Acta* **1132**, 83- 87.

16 Voordouw, G., Hagen, W. R., Krüse-Wolters, K. M., van Berkel-Arts, A. and Veeger, C. (1987) *Eur. J. Biochem.* **162**, 31-36.

17 Menon, N. K., Peck, H. D., LeGall, J. and Przybyla, A. E. (1987) *J. Bacteriol.* **169**, 5401-5407.

18 Li, C., Peck, H.D., LeGall, J. and Przybyla, A. E. (1987) *DNA* **6**, 539-551.

19 Brumlik, M. J. and Voordouw, G. (1989) *J. Bacteriol.* **171**, 4996-5004.

20 Voordouw, G. (1988) *Gene* **69**, 75-83.

21 Curley, G. P. and Voordouw, G.(1988) *FEMS Microbiol. Lett.* **49**, 295-299.

22 Stokkermans, J., van Dongen, W., Kaan, A., van den Berg, W. and Veeger, C. (1989) *FEMS Microbiol. Lett.* **58**, 217-222.

23 Van Dongen, W., Hagen, W., van den Berg, W. and Veeger, C. (1988) *FEMS Microbiol. Lett.* **50**, 5-9.

24 Gibson, T. J. (1984) Thesis, University of Cambridge.

25 Simon, R., Priefer, U. and Pühler, A. (1983) *Biotechnology* **1**, 784-791.

26 Priefer, U., Simon, R. and Pühler, A. (1984) in *IIId Eur. Congress on Biotechnology*, vol III, Eur. Fed. of Biotechnology, Verlag Chemie, Weinheim, pp. 207-212.

27 Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1987) *Current Protocols in Molecular Biology*, John Wiley and Sons, New York.

28 Van den Berg, W. A. M., van Dongen, W. M. A. M. and Veeger, C. (1991) *J. Bacteriol.* **173**, 3688-3694.

29 Van den Berg, W. A. M., Stokkermans, J. P. W. G. and van Dongen, W. M. A. M. (1989) *J. Biotechnol.* **12**, 173-184.

30 Goa, J. (1953) *Scand. J. Clin. Lab. Invest.* **5**, 218-222.

31 Bensadoun, A. and Weinstein, D. (1976) *Anal. Biochem.* **70**, 241-250.

32 Laemmli, U. K. (1970) *Nature* **227**, 680-685.

33 Wolbert, R. B. G., Hilhorst, R., Voskuilen, G., Nachtegaal, H., Dekker, M., van't Riet, K. and Bijsterbosch, B. H. (1989) *Eur. J. Biochem.* **184**, 627-633.

34 Vieira, J. and Messing, J. (1982) *Gene* **19**, 259-268.

35 Frey, M., Sieker, L. C., Payan, F., Haser, R., Bruschi, M., Pepe, G. and LeGall, J. (1987) *J. Mol. Biol.* **197**, 525-541.

36 Moura, I., Huynh, B. H., Hausinger, R. P., LeGall, J., Xavier, A. V. and Münck, E. (1980) *J. Biol. Chem.* **255**, 2493-2498.

37 LeGall, J., Pickril, B. C., Moura, I., Xavier, A. V., Moura, J. J. G. and Huynh, B. H. (1988) *Biochemistry* **27**, 1636-1642.

38 Coucouvanis, D., Kanatzidis, M. G. (1985) *J. Am. Chem. Soc.* **107**, 5005-5006.

39 Van Dijk, C., van Berkel-Arts, A. and Veeger, C. (1983) *FEBS Lett.* **156**, 340-344.

40 Hagen, W. R., van Berkel-Arts, A., Krüse-Wolters, K. M., Dunham, W. R. and Veeger, C. (1986) *FEBS Lett.* **201**, 158-162.

41 Pierik, A. J., Hagen, W. R., Redeker, J. S., Wolbert, R. B. G., Boersma, M., Verhagen, M. F. J. M., Grande, H. J., Veeger, C., Mutsaers, P. H. A., Sands, R. H. and Dunham, W. R. (1992) *Eur. J. Biochem.* **209**, 63-72.

42 Fernandez, V. M., Hatchikian, E. C., Patil, D. S. and Cammack, R. (1986) *Biochim. Biophys. Acta* **883**, 145-154.

43 Coremans, J. M. C. C., van der Zwaan, J. W. and Albracht, S. P. J. (1989) *Biochim. Biophys. Acta* **997**, 256-267.

Summary

Desulfovibrio vulgaris **(Hildenborough).** The organism described in this thesis, is an anaerobic gram-negative sulfate reducing bacterium (SRB). Its natural environments are the anaerobic sediments in lower levels of lakes and pools. This habitat is rich in sulfate that is used as terminal electron-acceptor by the organism and by performing this, *D. vulgaris* contributes to the important sulfur-cycle in nature. *D. vulgaris* can both utilize lactate (by anaerobic oxidation) and molecular hydrogen as energy source. The oxidation of lactate to acetate and CO_2 occurs in the cytoplasm or at the cytoplasmic membrane and results in the production of ATP and the release of protons and electrons. When *D. vulgaris* uses molecular hydrogen as substrate, the oxidation of the hydrogen occurs at the periplasmic side of the inner membrane. This creates a proton motive force that drives ATP synthesis.

Periplasmic Fe-hydrogenase. Molecular hydrogen has been shown to play an important role in both energy-evolving systems described above. So far, three hydrogenases, catalyzing the reversible H_2 oxidation and reduction of protons have been identified in *D. vulgaris* (Hildenborough). The precise physiological function of each of these hydrogenases remains unclear. Two of these enzymes are localized in the cytoplasmic membrane and contain nickel in addition to iron-sulfur clusters as cofactor. The third enzyme contains only iron-sulfur clusters as cofactors and resides in the periplasmic space of the bacterium. This enzyme exhibits one of the highest catalytic activities ever described for hydrogenases. It occurs as a heterodimer that is composed of a large α subunit (46 kDa) and a small β subunit (10 kDa). Only the small subunit is translated as a precursor (13 kDa) with a cleavable signal sequence for export that is probably involved in the export of both hydrogenase subunits across the cytoplasmic membrane. The catalytically active enzyme contains three iron-sulfur clusters as cofactors. Two of them are typical ferredoxin-like [4Fe-4S] clusters (F-clusters) involved in electron transport. The third (H-cluster) cluster contains six iron and sulfide ions coordinated in an unknown structure and is part of the catalytic center of the enzyme.

The gene encoding the Fe-hydrogenase was the first hydrogenase gene that was isolated and expressed in *E. coli*. From these expression studies it became apparent that only very small amounts of α and β subunits were assembled into an $\alpha\beta$ dimer and transported across the membrane. Also the iron-sulfur cluster incorporation was incomplete in the recombinant enzyme. The enzyme contained sub-stoichiometric amounts of F-clusters, while the H-cluster was not incorporated at all. These results indicated that the assembly and export of hydrogenase generating a catalytically active enzyme, are not spontaneously occurring processes, but involve specific helper components, as has been shown for other enzymes with redox-active metal clusters (reviewed in Chapter 1).

Studies regarding the biosynthesis of Fe-hydrogenase: the *hydC* gene. As genes serving a single pathway are often clustered in the genome, the identification of genes encoding these additional activating components, was started by the isolation of large DNA fragments surrounding the structural hydrogenase genes. Surprisingly, one of the large isolated DNA fragments contained a gene, *hydC* (**Chapter 2**) with homology in primary structure to the α and β subunits of the Fe-hydrogenase.

HydC has a high degree of similarity with both the α subunit of the Fe-hydrogenase (in its central part) and with the β subunit, minus the leader peptide (in its C-terminal part). Analogous to the FeMo co-factor insertion in nitrogenase component 1 which involves genes (*nifEN*) with high similarities to the structural subunits, it was speculated that the *hydC* gene might code for a helper protein that is involved in the processing of the hydrogenase. The primary structure of *hydC* contains a N-terminus with no homology with one of the hydrogenase subunits. Subsequently, it was found that this N-terminal segment has homology with mitochondrial NADH-ubiquinone reductase, with subunits of a NAD^+-reducing NiFe-hydrogenase from *Alcaligenes eutrophus* and with the Fe-hydrogenase I from *Clostridium pasteurianum* (**Chapter 3**). On the basis of what is known about iron-sulfur cluster contents of these three enzymes and the conservation of cysteine motifs in these proteins, it was suggested that these motifs coordinate [2Fe-2S] clusters. Unfortunately, the HydC protein could not be purified from *D. vulgaris*, because no growth conditions were found resulting in a sufficient production of HydC protein. This hampered a further biochemical and spectroscopical characterization of the protein. On the other hand, the high degree of homology with the *C. pasteurianum* Fe-hydrogenase, strongly suggests that HydC is a second alternative Fe-hydrogenase and not a helper protein involved in the processing of Fe-hydrogenase.

Numerous attempts have been made to exchange the genes for the subunits of the Fe-hydrogenase and the *hydC* gene in the *D. vulgaris* genome with inactivated, interrupted copies of the genes. This type of marker exchange experiments would also be very useful for the identification of genes involved in biosynthesis of hydrogenase. One of the requirements for marker exchange is a system for the introduction of plasmids into *Desulfovibrio*. Such a plasmid transfer system has been developed, but subsequent experiments to apply it for marker exchange have been unsuccessful.

The prismane protein. The inability to design a system for marker-exchange mutagenesis in *Desulfovibrio* blocked further study of the biosynthesis of the Fe-hydrogenase. Therefore, investigations on another protein from *D. vulgaris*, the prismane protein, were started that are described in the second part of this thesis. As mentioned earlier, some indications were obtained that the H-cluster of Fe-hydrogenase is a [6Fe-6S] cluster.

Stronger indications for the existence of such "supercluster" were obtained by Hagen and Pierik in our department for another iron-sulfur containing protein from *D. vulgaris*, the prismane protein. They isolated a protein containing six irons and sulfide ions coordinated in only one [6Fe-6S] cluster. The putative [6Fe-6S] prismane cluster occurs in four different redox-states: the three-electron reduced state $[6Fe\text{-}6S]^{3+}$ (S=1/2), $[6Fe\text{-}6S]^{4+}$ (S = even), $[6Fe\text{-}6S]^{5+}$ (S = 1/2 and S = 9/2) and the fully oxidized $[6Fe\text{-}6S]^{6+}$ (S=0) that shows no EPR spectrum. **Chapter 4** and **5** describe the isolation of the genes for the prismane proteins from *D. vulgaris* (Hildenborough) and *D. desulfuricans* (ATCC 27774) and the determination of the amino acid sequence. Both proteins are highly conserved (66% identical residues), except for a 100 residues segment (residue 50-150). Besides this, both proteins contain typical cysteine motifs at the N-terminus. These motifs have also been found in the sequence of the α subunit of CO dehydrogenase from *Methanothrix soehngenii* and, in a slightly modified form, in that of CO dehydrogenase from *Clostridium thermoaceticum*. Also for the CO dehydrogenase from *M. soehngenii* a supercluster has been proposed. Therefore, it is tempting to speculate about the involvement of this motif in the ligation of the [6Fe-6S] prismane cluster.

Prismane protein is produced only in small amounts in *D. vulgaris*. Since large amounts of purified prismane protein are required for X-ray crystallography and Mössbauer studies, efforts were made for overproduction of the protein (**Chapter 6**). In a first attempt, the protein was overproduced in *E. coli*. In this host, a high production of prismane protein was obtained, but no iron-sulfur cluster was incorporated into the protein. The overproduced protein occurred as large insoluble protein-complexes. A second attempt for the overproduction of prismane protein was performed in *D. vulgaris* by using the aforementioned cloning system. A 25-fold overproduction of prismane protein was obtained by the introduction of extra copies of the gene encoding the prismane protein on a stable plasmid. Biochemical and spectroscopic properties of the protein overproduced in *D. vulgaris* were shown to be identical to wild-type prismane with one exception: in the as-isolated, one-electron-reduced state the protein shows EPR signals belonging to a second S=1/2 spin system that was not observed in the wild-type protein. These additional signals were also described for the wild-type prismane protein purified from *D. desulfuricans* by Moura and co-workers in Portugal. EPR signals belonging to this second S=1/2 spin system disappear upon reduction/re-oxidation of the overproduced prismane protein, indicating that this spin system represents a different magnetic form of the [6Fe-6S] cluster. There are no indications for a second cluster as proposed by Moura et al. Determination of the three dimensional structure by X-ray crystallography and further Mössbauer spectroscopy of the overproduced prismane protein are subject for further study in our department and will ultimately lead to

insight into the structure of this novel iron-sulfur cofactor.

Samenvatting

Redoxreacties. Redoxreacties spelen een essentiële rol in de natuur. Zij spelen een rol zowel in spontaan verlopende processen (zoals bijvoorbeeld het roesten (=oxyderen) van niet-edelmetalen aan de lucht), als in de stofwisseling van alle in de natuur voorkomende organismen (van relatief eenvoudige eencelligen, zoals bijvoorbeeld bacteriën, tot complex meercelligen, zoals bijvoorbeeld zoogdieren). Redoxreacties kunnen worden beschouwd als twee deelreacties die ook wel worden aangeduid met de naam redox-halfreacties. Hierbij is sprake van elektronenoverdracht van de gereduceerde vorm (=reductor) van de elektron-donerende halfreactie (of koppel) naar de geoxydeerde vorm (=oxydator) van de elektron-accepterende halfreactie. Doordat een electron van een verbinding met een relatief hoge energie inhoud, de reductor, naar een verbinding met een relatief lage energie, de oxydator, gaat, komt bij redoxreacties energie vrij (zoals bijvoorbeeld bij het ontladen van een accu in de vorm van electriciteit).

Halfreactie 1:	**reductor 1 $\leftrightarrows$ oxydator 1 + electron**
Halfreactie 2:	**oxydator 2 + electron $\leftrightarrows$ reductor 2**
Redoxreactie:	**reductor 1+ oxydator 2 $\leftrightarrows$ oxydator 1 + reductor 2**

Sommige stoffen, zoals bijvoorbeeld zuurstof in waterig milieu, nemen heel gemakkelijk elektronen op. Hierbij wordt de oxydator in het geval van zuurstof onder andere omgezet in het zuurstof anion (O^{2-}), dat zeer instabiel en daardoor zeer reactief is en onmiddellijk met aanwezige protonen H_2O vormt. In dit geval is sprake van een redox-koppel (H_2O/O_2) met een hoge redoxpotentiaal onder standaardcondities (E_0 = +800 mV bij pH 7). Andere stoffen, zoals bijvoorbeeld waterstof (H_2), fungeren daarentegen gemakkelijker als reductor en staan daarbij elektronen af (waarbij in het geval van waterstof, H^+ ionen (protonen) ontstaan). Hierdoor heeft een H_2/H^+ redox-koppel een lage redoxpotentiaal onder standaardcondities (E_0 = -400mV bij pH 7). Naast deze redoxkoppels komen in de natuur nog veel meer voorbeelden voor van stoffen die in één of meer stappen, één of meerdere elektronen opnemen of afstaan. Zo kan ijzer, Fe, twee elektronen afstaan en vóórkomen als Fe^{2+}, bijvoorbeeld als zwart gekleurd FeS dat aanwezig is in poelslib. Daarnaast kan Fe^{2+}, met name onder invloed van zuurstof, overgaan in Fe^{3+} wat voorkomt in roest. Een ander voorbeeld is zwavel: zwavel kan voorkomen als S^{6+} in sulfaten (bv. in kunstmeststoffen) en kan na opname van 2 elektronen overgaan in S^{4+} (sulfiet dat door slagers of wijnboeren als anti-oxydant en desinfectans wordt gebruikt); S^{4+} kan 4 elektronen opnemen, waarbij het

element zwavel (S^0) ontstaat, dat nogmaals 2 elektronen kan opnemen waarbij S^{2-} ontstaat dat in de vorm van het giftige en stinkende H_2S gas voorkomt in gassen die ontwijken uit moerassen en rottende eieren.

De hoeveelheid energie die vrijkomt bij verschillende redoxreacties kan sterk variëren en is gecorreleerd aan het verschil in redoxpotentiaal van beide redoxkoppels. Zo is de hoeveelheid energie die vrijkomt bij het oxyderen van metalen aan de lucht zeer gering, terwijl de hoeveelheid energie die vrijkomt bij ontlading van een accu zo hoog is, dat hij kan worden gebruikt voor het starten van een automotor. De energie die vrijkomt bij redoxreacties die optreden in cellen van levende organismen is onmisbaar voor het voortbestaan van deze organismen. Hierbij wordt de energie die vrijkomt niet uitsluitend omgezet in warmte, maar een groot gedeelte wordt gebruikt voor het aanmaken van allerlei energie bevattende verbindingen (o.a. nucleïnezuren, eiwitten en lipiden) die essentieel zijn voor het voortbestaan van de cel of het organisme. Zo maken deze verbindingen o.a. voortplanting, groei en ontwikkeling van organismen mogelijk.

Redoxenzymen. Redoxreacties verlopen normaal zeer langzaam onder de in de cel aanwezige condities. In de cel worden katalysatoren (=enzymen) geproduceerd die een reactie aanzienlijk kunnen versnellen (tot één miljoen maal sneller). Deze enzymen versnellen de redoxreacties dusdanig dat de vrijkomende energie voldoet aan de vraag van het organisme (indien voldoende uitgangsstoffen aanwezig zijn). De vrijkomende energie wordt door redoxenzymen omgezet in hoog-energetische verbindingen. Enzymen bestaan uit één of meerdere eiwitketens (=subeenheden) die opgebouwd zijn uit 20 verschillende aminozuren. De volgorde van deze aminozuren wordt bepaald door het erfelijk materiaal van de cel, het DNA. Hierin komen vier verschillende componenten (=nucleotiden) voor, waarvan de volgorde vertaald wordt in een eiwitketen. Elke kombinatie van drie nucleotiden vormt hierbij de kode voor een bepaald aminozuur dat vervolgens aan de eiwitketen wordt toegevoegd. Hierdoor ontstaat een eiwit met een unieke keten van aminozuren.

Na de synthese vouwt het eiwit zich in de juiste drie-dimensionale structuur en worden additionele componenten (ook wel "coenzymen" of "cofactoren" genaamd) geïnserteerd. Redoxenzymen hebben altijd de beschikking over cofactoren die redox-aktief zijn. Deze redox-actieve groepen kunnen organische verbindingen zijn (zoals bijvoorbeeld vitamine B_2 dat o.a. is opgebouwd uit koolstof, waterstof en zuurstof atomen), maar daarnaast komen ook anorganische verbindingen (bijvoorbeeld ijzer, nikkel of koper bevattende cofactoren) als redox-actieve groepen voor. In ijzer-zwavel bevattende redox-enzymen komen cofactoren voor die opgebouwd zijn uit ijzer en sulfide ionen die gerangschikt kunnen zijn in een kenmerkende struktuur (een Fe-S cluster). Deze struktuur kan worden vastgesteld met behulp van röntgen diffractie analyse van eiwitkristallen. Daarnaast zijn er spectroscopische

technieken zoals EPR (electron paramagnetische resonantie) en Mössbauer, die een "fingerprint" van een dergelijke cluster in de vorm van een specifiek spectrum kunnen geven. Zo zijn de vier ijzer en zwavel atomen van een [4Fe-4S] cluster gerangschikt in een kubus-achtige structuur. Dit type komt o.a. voor in bacteriële ferredoxines. De ijzer en zwavel atomen die voorkomen in de [6Fe-6S] cluster van het eiwit uit de laatste drie hoofdstukken van dit proefschrift, worden verondersteld in een prismaan structuur voor te komen zoals is afgebeeld op de kaft van dit proefschrift.

IJzer-zwavel clusters zijn vaak gebonden aan cysteine residuen in de eiwit keten. Cysteine is een aminozuur dat beschikt over een sulfide groep, die gebonden kan worden aan de ijzeratomen uit de cluster. Deze cysteines komen vaak in een speciale volgorde voor in de eiwitketen, waaruit men de mogelijke bindingsplaats voor een ijzer-zwavel cluster kan voorspellen. Daarnaast worden ook steeds meer ijzer-zwavel clusters in eiwitten beschreven die door andere aminozuren (bijvoorbeeld histidine) gebonden worden.

Desulfovibrio vulgaris **(Hildenborough).** Het organisme, dat gebruikt werd voor het in dit proefschrift beschreven onderzoek, is een sulfaatreducerende, anaëroob levende bacterie d.w.z. het organisme kan alleen groeien in afwezigheid van zuurstof. Als substraat kan door het organisme lactaat of waterstof gebruikt worden. Electronen die vrijkomen bij de oxydatie van deze substraten, worden niet gebruikt voor de reductie van zuurstof, zoals bij hogere organismen en aëroob levende bacteriën het geval is, maar voor de reductie van sulfaat. *D. vulgaris* leeft normaal in onderlagen van meren en poelen; een zuurstofloze omgeving die rijk is aan sulfaat. De bacterie levert zo haar bijdrage aan de sulfaatreductie die een onderdeel vormt van de zwavel-cyclus. De zwavel-cyclus is één van de belangrijke cycli in de natuur en, naar algemeen wordt aangenomen, zelfs de voornaamste redox-cyclus voordat zuurstof aanwezig was in de atmosfeer.

Fe-hydrogenase van *D. vulgaris* (Hildenborough). Tot de groep van ijzerbevattende redoxenzymen behoren onder andere hydrogenases. Hydrogenases zijn enzymen die de produktie (en/of oxydatie) van waterstof katalyseren. Waterstof is een belangrijke sleutelcomponent, die van groot belang is voor de energiehuishouding van anaërobe sulfaatreducerende bacteriën. Het Fe-hydrogenase uit *D. vulgaris* (Hildenborough), dat het uitgangspunt vormde voor onderzoek beschreven in dit proefschrift, heeft één van de hoogste activiteiten in waterstof productie van alle hydrogenases die bekend zijn. Gezien de grote belangstelling die bestaat voor waterstof als schone energiebron, wordt veel onderzoek aan hydrogenases gedaan. Met name de rol van metalen in de katalyse van de waterstof productie wordt onderzocht (behalve het ijzer-hydrogenase zijn er ook hydrogenases met nikkel in het actieve centrum).

Uit voorafgaand onderzoek dat is uitgevoerd op de vakgroep Biochemie van de Landbouwuniversiteit te Wageningen, is komen vast te staan dat Fe-hydrogenase van *D. vulgaris* drie ijzer-zwavel clusters als cofactoren bezit. Twee van deze clusters (F-clusters) bevatten vier ijzer en vier sulfide ionen, die geordend zijn in een al eerder genoemde kubusachtige structuur. Deze twee F-clusters worden verondersteld betrokken te zijn bij transport van elektronen van of naar het katalytische centrum van het enzym waar de oxydatie van waterstof of de reductie van protonen plaatsvindt. De derde cluster (H-cluster), die wordt geacht deel uit te maken van dit katalytische centrum bestaat uit ongeveer 6 ijzer en 6 sulfide ionen die mogelijk geordend zijn in een [6Fe-6S] cluster met een tot nu toe nog onbekende struktuur. Daarnaast zijn aanwijzingen verkregen dat de assemblage van het Fe-hydrogenase, dat o.a. de inbouw van deze derde H-cluster omvat, niet "spontaan" verloopt, maar dat daarvoor weer andere helper-eiwitten of -enzymen nodig zijn (**Hoofdstuk 1**).

Biosynthese studies. In het eerste deel van het onderzoek is geprobeerd te achterhalen, hoe de assemblage van het Fe-hydrogenase verloopt. Dit kan op verschillende manieren onderzocht worden. Allereerst kan dit gebeuren op een manier die samengevat kan worden onder de term "in vitro reconstitutie". Hierbij worden alle mogelijke componenten die betrokken kunnen zijn bij de assemblage van het enzym, uit de cel gezuiverd en vervolgens wordt bepaald hoe het niet-actieve enzym geactiveerd kan worden in een reageerbuis onder toevoeging van de gezuiverde componenten. In de praktijk blijkt dit niet gemakkelijk te zijn wanneer véél componenten betrokken zijn bij de activering. Een andere manier van aanpak, die veel geschikter leek, is het opsporen langs genetische weg van genen die coderen voor de helper-eiwitten die betrokken zijn bij de assemblage van Fe-hydrogenase. Deze methode gaat uit van een groepering (=clustering) van genen die koderen voor helpereiwitten in de assemblage, rondom de genen die coderen voor het hydrogenase enzym zelf (de zogenaamde structurele genen), zoals dat ook voor andere bacteriële enzymen is aangetoond. Hierbij worden grote DNA fragmenten, die zowel de genetische informatie voor het enzym als die voor helper-eiwitten kunnen bevatten, geïsoleerd. Vervolgens worden veranderingen (mutaties) in deze DNA fragmenten aangebracht. Daarna worden deze fragmenten ingebracht in een gastheer die zelf geen actief Fe-hydrogenase enzym produceert. Wanneer een helpergen op zo'n fragment gemuteerd is, resulteert dit in de productie van inactief enzym. Bestudering van het inactieve enzym van deze bacteriën geeft dan informatie omtrent de assemblagestap die verstoord is.

Meestal wordt in de genetische aanpak gekozen voor *Escherichia coli* als gastheer, omdat dit de bacterie is waarvan genetisch het meeste bekend is. Tot nu toe is echter nooit expressie van actief *D. vulgaris* hydrogenase in *E. coli* verkregen (ook niet onder anaërobe condities). Mede daarom is gekozen voor de natuurlijke gastheer, *Desulfovibrio vulgaris*

(Hildenborough) zelf. Het is dan echter wel noodzakelijk de Fe-hydrogenase genen van deze gastheer uit te schakelen. Het opzetten van een systeem voor deze genetische manipulatie voor *D. vulgaris*, is slechts ten dele gelukt. Het is mogelijk gebleken om stukjes circulair DNA (=plasmiden), die in staat zijn zichzelf te vermenigvuldigen in de gastheer, met daarin geïnserteerd stukken DNA die de structurele genen voor het Fe-hydrogenase bevatten, in te brengen in *D. vulgaris*. Vervolgens is aangetoond dat dit resulteerde in een overproductie van Fe-hydrogenase in deze bacteriën (gepubliceerd in van den Berg, W.A.M., Stokkermans, J.P.W.G. and van Dongen (1989) *J. Biotechnol.* **12**, 173-184). Getracht werd dit systeem daarna ook te gebruiken voor het aanbrengen van mutaties in het bacteriële genoom. Hierbij werd een geïsoleerd *D. vulgaris* gen gemuteerd door insertie van het chlooramphenicol gen (=marker), dat de gastheer resistent maakt tegen dit antibioticum. Vervolgens werd dit defecte gen op de reeds eerder beschreven methode ingebracht in *D. vulgaris*. Door middel van uitwisseling ("marker-exchange") van het defecte gen en het intacte gen, dat gelegen is op het genoom van de gastheer, zouden gerichte mutaties aangebracht kunnen worden. Deze bewerkelijke, en gecompliceerde experimenten, die tot nu toe geen resultaat hebben opgeleverd, namen een groot gedeelte van het onderzoeksproject in beslag en zijn niet opgenomen in dit proefschrift.

Het *hydC* gen. Daarnaast zijn grote stukken DNA van *D. vulgaris* gekloneerd, die behalve de genen die coderen voor de grote component (α subeenheid) en de kleine component (β subeenheid) die samen het hydrogenase vormen, ook grote stukken uit de omgeving bevatten. Op een van deze stukken werd een gen gevonden (*hydC*) dat kodeert voor een eiwit met aminozuur volgorde die een hoge mate van overeenkomst vertoont met de α en β subeenheden van het Fe-hydrogenase (**Hoofdstuk 2**). Deze overeenkomstige gebieden (domeinen) die in de α en β subeenheden afzonderlijk voorkomen, komen in het HydC eiwit in één eiwitketen voor. Bovendien bevat deze HydC eiwitketen aan het begin (N-terminus) nog een extra domein dat niet aanwezig is in beide Fe-hydrogenase subeenheden. Later bleek dat de aminozuur volgorde van dit domein zeer sterke gelijkenis vertoont met eiwit componenten of gedeelten daarvan in andere enzymen: het NAD^+ reducerende NiFe-hydrogenase uit *Alcaligenes eutrophus*, NADH-dehydrogenase uit runderhart en het Fe-hydrogenase-1 uit *Clostridium pasteurianum* (**Hoofdstuk 3**). De strukturen van de ijzer-zwavel clusters, die voorkomen als cofactoren in deze drie enzymen, zijn gekarakteriseerd. Zowel de ketens van deze drie eiwitten als die van het HydC eiwit, bevatten specifieke aminozuur volgordes (motieven). In deze motieven zijn de cysteines dusdanig gepositioneerd dat hieruit kan worden voorspeld dat dit mogelijke bindingsplaatsen zijn voor [2Fe-2S] clusters. Helaas kon het eiwit HydC niet geïsoleerd worden uit *D. vulgaris*, omdat geen condities gevonden konden worden, waaronder dit eiwit voldoende geproduceerd wordt. Dit

maakte een verdere karakterisering van het eiwit zelf onmogelijk. De grote mate van overeenkomst in aminozuur volgorde van HydC met die van hydrogenase-1 van *Clostridium pasteurianum* doet vermoeden dat HydC waarschijnlijk niet fungeert als helper-eiwit betrokken bij de biosynthese van Fe-hydrogenase, maar dat het mogelijk een alternatief tweede Fe-hydrogenase is.

Het prismaan eiwit. Doordat de eerder beschreven "marker-exchange" experimenten zonder resultaat bleven, was het onmogelijk mutaties aan te brengen in het DNA van *D. vulgaris*. Hierdoor bleek de opheldering van de biosynthese van Fe-hydrogenase langs genetische weg binnen de termijn van het onderzoeksproject onmogelijk. In het laatste deel van het onderzoek werd overgeschakeld naar een ander onderzoeksproject dat raakvlakken heeft met het hydrogenase onderzoek. Zoals eerder vermeld, waren er aanwijzingen voor de aanwezigheid van een [6Fe-6S] cluster (de H-cluster) in het actieve centrum van het Fe-hydrogenase. Dergelijke metaalclusters waren, tot dan toe, slechts bekend uit de anorganische chemie, maar zij waren nog nooit aangetoond als cofactor in metaal-eiwitten. Een onomstotelijk bewijs dat het hydrogenase een dergelijk [6Fe-6S] cluster heeft, is echter nog niet geleverd.

Veel sterkere aanwijzingen voor de aanwezigheid van een dergelijk "supercluster" werden door Hagen en Pierik van onze vakgroep wel verkregen voor een ander ijzer-zwavel bevattend eiwit van *D. vulgaris*. Zij isoleerden een eiwit dat in totaal 6 ijzer-ionen en 6 sulfide-ionen bevat die geordend waren een [6Fe-6S] cluster. Doordat de aanwezige ijzer in twee toestanden, Fe^{2+} en Fe^{3+}, kan voorkomen en sulfide lading -2 heeft, zijn in totaal zeven redox toestanden van het cluster, variërend van 6+ tot 0, mogelijk. Door het eiwit in verschillende redox-toestanden te brengen konden vier van deze toestanden , variërend van de meest gereduceerde 3+ toestand tot de meest geoxydeerde toestand, 6+, aangetoond worden. Elk van deze redox-toestanden kan door het spectrum dat met de al eerder genoemde spectroscopische EPR techniek wordt opgenomen, te vergelijken met die van een gesynthetiseerde ijzer-zwavel model-verbinding, gekarakteriseerd worden. Doordat het EPR spectrum van het eiwit met de cluster in de meest gereduceerde 3+ toestand, overeenkomt met het spectrum van een zogenaamde [6Fe-6S] prismaan model cluster, werd dit eiwit "prismaan eiwit" genoemd door Hagen en Pierik. Karakteristiek voor EPR-spectra zijn de plaatsen waar pieken (resonanties) voorkomen waaruit de zogenaamde *g*-waarden berekend kunnen worden. Sommige van de *g*-waarden zijn voor het eiwit met de cluster in de 5+ toestand, veel hoger dan tot dan toe waargenomen was voor eiwitten met "klassieke" [4Fe-4S], [3Fe-4S] of [2Fe-2S] clusters. Voor dit fenomeen is door Hagen en Pierik, de naam "superspin" bedacht. Superspins zijn later ook gevonden in andere ijzer-zwavel cluster bevattende eiwitten gevonden: CO dehydrogenase uit *Methanotrix soehngenii* en sulfietreductase uit *D. vulgaris.*

Voor het ophelderen van bindingsplaatsen die betrokken zijn bij de binding van deze superclusters was het noodzakelijk dat het gen dat codeert voor het prismaan eiwit geïsoleerd werd en dat hieruit de aminozuur volgorde bepaald werd (**Hoofdstuk 4**). De domeinen in een eiwitketen, die belangrijk zijn voor het functioneren van een eiwit zijn vaak in hoge mate overeenkomstig met die van corresponderend (=homoloog) eiwit uit een nauw verwant organisme. Het vergelijken van twee aminozuur volgordes (=primaire strukturen) van twee homologe eiwitten, geeft daarom een indicatie omtrent belangrijke gebieden, zoals bijvoorbeeld cofactor bindingsplaatsen, in een eiwit. In de aminozuur volgorde van het *D. vulgaris* prismaan eiwit bleek een motief voor te komen dat ook gevonden werd in de aminozuur volgordes van CO dehydrogenases uit *M. soehngenii* en *Clostridium thermoaceticum*. Het CO dehydrogenase van *M. soehngenii* bevat ook een ijzer-zwavel cluster met een superspin en daarom zouden de aminozuren (en met name de cysteines hierin) in dit motief mogelijk betrokken kunnen zijn bij binding van de prismaan supercluster. De primaire struktuur van een homoloog prismaan eiwit uit *D. desulfuricans* ATTC 27774, is beschreven in **Hoofdstuk 5**. Op een stukje van ±100 aminozuren in het begin van de keten (van residu 50 tot 150) na, blijken de primaire strukturen van beide eiwitten in grote mate overeen te komen. Ook het cysteine motief dat mogelijk betrokken is bij clusterbinding komt in dit eiwit voor.

Het prismaan eiwit wordt in *D. vulgaris* slechts in geringe mate geproduceerd. Aangezien voor de bestudering van een eiwit (o.a. de opheldering van de drie-dimensionale structuur van het eiwit), grote hoeveelheden nodig zijn, is getracht het eiwit tot overproductie te brengen. In eerste instantie werd dit geprobeerd door het gen dat codeert voor het prismaan eiwit, op een plasmied in *E. coli* in te brengen. Dit resulteerde in een overproductie van prismaan eiwit. In deze gastheer werd echter geen ijzer-zwavel cluster ingebouwd in het eiwit, dat in de vorm van grote eiwit-aggregaten voorkwam in de cel. Dit deed ons besluiten om over te stappen op overproductie van het prismaan eiwit in *D. vulgaris*. Hiertoe werden extra kopieën van het prismaan eiwit coderende gen, met behulp van het reeds eerder besproken kloneringssysteem in *D. vulgaris* ingebracht (**Hoofdstuk 6**). In *D. vulgaris* werd een 25-voudige overproductie van prismaan eiwit gerealiseerd. Het geïsoleerde eiwit bleek volledig ingebouwde ijzer-zwavel clusters te bevatten en was identiek aan het normaal geproduceerde prismaan eiwit in *D. vulgaris* op één uitzondering na. In één van de redox-toestanden van het eiwit, werd een extra signaal gevonden dat niet aanwezig was in het normaal geïsoleerde *D. vulgaris* eiwit, maar dat wel voorkwam in het homologe *D. desulfuricans* eiwit dat was beschreven door de groep van Moura in Portugal. Naar aanleiding van de Mössbauer spectra van het *D. desulfuricans* eiwit, concludeerde deze groep dat in prismaan eiwit twee [6Fe-6S] clusters aanwezig waren, hoewel dit niet strookte met het

door hun gevonden aantal van 6 ijzer atomen per eiwit molecule (dat overeenkomt met het *D. vulgaris* eiwit). Voorts was voor het *D. vulgaris* eiwit slechts één cluster per eiwit molecule voorgesteld door Hagen en Pierik naar aanleiding van de interpretatie van hun spectroscopische data en biochemische karakterisering van het eiwit. Experimenten met het overgeproduceerde *D. vulgaris* eiwit tonen duidelijk aan dat het extra signaal niet afkomstig is van een tweede ijzer-zwavel cluster, maar dat dit veroorzaakt wordt door één en hetzelfde cluster waarvan bij isolatie een gedeelte in een iets andere vorm voorkomt, die verdwijnt wanneer het eiwit gereduceerd en vervolgens weer geoxydeerd wordt. De opheldering van de drie-dimensionale struktuur met behulp van röntgen diffractie van gekristalliseerd prismaan eiwit en Mössbauer studies van, onder andere, overgeproduceerd *D. vulgaris* prismaan eiwit, waaraan binnen onze vakgroep wordt gewerkt, zal het uiteindelijke inzicht in de struktuur van deze nieuwe ijzer-zwavel cofactor geven.

Curriculum vitae

Jacobus Petrus Wilhelmus Gerardus Stokkermans werd geboren op 23 december 1960 te Tilburg. In 1979 behaalde hij het diploma Atheneum B aan het Dr Möllercollege te Waalwijk. In 1980 begon hij met de studie biologie (studie richting B4 met scheikunde als bijvak) aan de Katholieke Universiteit te Nijmegen. Het kandidaatsexamen (B4) werd in 1983 behaald. Het doctoraalexamen werd in 1987 afgelegd en omvatte het hoofdvak Biochemie (Prof. Dr H. Bloemendal) en de bijvakken Microbiologie (Prof. Dr G. D. Vogels) en Moleculaire Biologie (Prof. Dr J. G. G. Schoenmakers).

Van 1 augustus 1987 tot 1 februari 1992 was hij als promovendus verbonden aan de afdeling Biochemie van de Landbouwuniversiteit te Wageningen. Het in deze periode uitgevoerde onderzoek heeft geleid tot dit proefschrift.