ADVANCES IN INORGANIC CHEMISTRY, VOL. 38

EVOLUTION OF HYDROGENASE GENES

GERRIT VOORDOUW

Division of Biochemistry, Department of Biological Sciences, The University of Calgary, Calgary, Alberta, T2N 1N4 Canada

I. Introduction

Hydrogenases catalyze the reversible oxidation of hydrogen. The past 7 years have seen a surge in published nucleotide sequences of the structural genes encoding these enzymes (Table I) (*1–24*). As a result, translated primary structures from different families, the iron-only and nickel-containing hydrogenases (*25*), are now available: 16 published sequences for nickel-containing hydrogenases, both from eubacteria (*Desulfovibrio, Bradyrhizobium, Azotobacter, Rhodobacter, Rhodocyclus, Alcaligenes,* and *Rhizobium*) and archaebacteria (*Methanobacterium* and *Methanothermus*), and three different sequences for iron-only hydrogenases (all *Desulfovibrio vulgaris*). Comparison of these sequences allows conserved features to be identified, as will be discussed below. All hydrogenases studied with molecular biological techniques to date consist of at least two subunits. The structural genes encoding these subunits are organized in an operon with the gene order (5′ → 3′) as indicated in Table I. Nucleotide sequencing has indicated the presence of additional genes in the operon, e.g., hydrogenase 1 of *Escherichia coli* is encoded by an operon encoding six open reading frames (Table I). Two of these (*hyaA* and *hyaB*) encode the small and large subunits,

TABLE I

HYDROGENASES FOR WHICH THE PRIMARY STRUCTURE HAS BEEN DETERMINED BY NUCLEOTIDE SEQUENCING OF THEIR STRUCTURAL GENES

N[a]	Organism	Type	Gene name[b]	Ref.
1	*Desulfovibrio vulgaris* Hildenborough	[Fe]	*hydA* (46.0), *hydB* (9.6)	*1*
2	*Desulfovibrio vulgaris* Monticello	[Fe]	*hydA* (46.0), *hydB* (9.6)	*2*
3	*Desulfovibrio vulgaris* Hildenborough	[Fe]	*hydC* (65.8)	*3*[d]
4	*Desulfovibrio gigas*	[Ni–Fe]	*hynB* (28.4), *hynA* (61.3)	*4–6*
5	*Desulfovibrio vulgaris* Miyazaki F	[Ni–Fe]	*hynB* (28.8), *hynA* (62.5)	*7*
6	*Desulfovibrio fructosovorans*	[Ni–Fe]	*hynB* (28.4), *hynA* (61.6)	*8*
7	*Desulfovibrio baculatus*	[Ni–Fe–Se]	*hysB* (30.8), *hysA* (56.8)	*5, 9*
8	*Bradyrhizobium japonicum*	[Ni–Fe]	*orf1* (34.5), *orf2* (65.9), *orf3* (?[d])	*10*
9	*Rhizobium leguminosarum*	[Ni–Fe]	*hupS* (34.4), *hupL* (65.9)	*11, 12*
10	*Azotobacter chroococcum*	[Ni–Fe]	*hupS* (34.2), *hupL* (66.4), *orf3* (?)	*13*

11	*Azotobacter vinelandii*	[Ni–Fe]	*hoxK*, *hoxG*, *orf3* 34.2 66.8 27.7	*14*
12	*Rhodobacter capsulatus*	[Ni–Fe]	*hupS*, *hupL*, *orfX* 34.3 68.1 30.2	*15*
13	*Rhodocyclus gelatinosus*	[Ni–Fe]	*hupS*, *hupL* 34.6 68.5	*16, 17*
14	*Escherichia coli* (hydrogenase-1)	[Ni–Fe]	*hyaA*, *hyaB*, *hyaC*, *hyaD*, *hyaE*, *hyaF* 35.6 66.2 27.6 21.5 14.9 31.5	*18*
15	*Escherichia coli* (hydrogenase-3)	[Ni–Fe]	*orf1*, *orf2*, *orf3*, *orf4*, *orf5*, *orf6*, *orf7*, *orf8* 17.6 21.8 64.1 33.3 65.0 20.3 28.0 15.5	*19*
16	*Alcaligenes eutrophus*	[Ni–Fe]	*hoxF*, *hoxU*, *hoxY*, *hoxH* 66.8 26.0 22.9 54.7	*20*
17	*Methanobacterium thermoautotrophicum* ΔH	[Ni–Fe]	*mvhD*, *mvhG*, *mvhA*, *mvhB* 15.8 33.0 53.0 44.0	*21*
18	*Methanobacterium thermoautotrophicum* ΔH	[Ni–Fe]	*frhA*, *frhD*, *frhG*, *frhB* 44.7 17.6 25.7 30.7	*22*
19	*Methanothermus fervidus*	[Ni–Fe]	*mvhD*, *mvhG*, *mvhA*, *mvhB* ?[d] ?[d] 53.0 44.0	*23*
20	*Anabaena cylindrica*	?	*texP*[e] 41.1	*24*

[a] Number used to identify hydrogenases in Table III.

[b] Gene names as given in the referenced publications. The size (kDa) of the putative gene product is indicated below its name. Suggested structural genes are underlined. The gene order is that in the operon.

[c] It is not clear whether this gene product is expressed.

[d] Gene was not completely sequenced.

[e] Gene name given in this publication.

which constitute this hydrogenase as it is isolated from *E. coli* (*26*). The function of the products of the other four genes (*hyaC–hyaF*) has not yet been established. They could function in anchoring the two-subunit hydrogenase to the membrane and in conducting electrons to the membrane-bound electron transport chain (see below). Similarly, in the hydrogenase 3 operon of *E. coli, orf5* has been indicated as a structural gene because of its homology with other [NiFe] hydrogenase large-subunit genes (e.g., *hyaB*). Again, assignment of the roles of the other seven gene products is difficult in the absence of definitive biochemical data, although an electron transfer function is implicated for the products of *orf2* and *orf6*. Some or all can be expected to constitute the membrane-bound hydrogenase 3 complex (*27*). In view of these uncertainties, only *hyaA* and *hyaB*, and *orf2, orf5*, and *orf6*, have been indicated as possible structural genes for hydrogenase 1 and hydrogenase 3 in Table I.

The problem of gene nomenclature needs some consideration. The names for the hydrogenase genes of *Desulfovibrio* are as described elsewhere (*6, 28*). Those for hydrogenase genes of other organisms (Table I) are generally as in the references given, except where no names were provided by the authors. This creates the problem that multiple designations are used for genes encoding homologous proteins (e.g., the gene for the large subunit of [NiFe] hydrogenase has been named *hynA, orf2, hupL, hoxG, hyaB, hoxH, mvhA,* and *frhA*) and, more seriously, that the same gene name may refer to completely different proteins in different operons (e.g., *orf2* of *Bradyrhizobium japonicum* and of the hydrogenase 3 operon of *E. coli*). Introduction of a unifying gene nomenclature will not be attempted here (see Section V). The advantage of this diversity is that it is often not necessary to specify the organism (e.g., *hyaB* is an *E. coli* gene only). Ambiguous names will be fully specified.

II. Hydrogenase Operons

A. Iron-Only Hydrogenases

Periplasmic [Fe] hydrogenases have so far only been isolated from sulfate-reducing bacteria of the genus *Desulfovibrio*. The *hydA,B* operon of *D. vulgaris* Hildenborough encodes a simple two-subunit hydrogenase (Ref. *1*; reviewed in Refs. *6, 29–32*). Its large α subunit (M_r = 46K, encoded by the *hydA* gene) contains 18 cysteine residues, whereas the mature small β subunit (M_r = 9.6K) lacks cysteine. The small

subunit is synthesized in precursor form (M_r = 13.5K for unprocessed pro-β). The three iron–sulfur clusters known to be present in this hydrogenase (*33*) must thus all coordinate to large-subunit cysteine residues. The NH_2-terminal amino acid sequence of α (residues 1–105) is homologous with 8Fe–8S ferredoxins: two groups of 4 cysteine residues are present in the characteristic pattern C-X-X-C-X-X-C-X-X-X-C-P (Table II, elements 1 and 2), which have been shown to bind two 4Fe–4S clusters in the ferredoxin of *Peptococcus aerogenes* (*34*). The COOH-terminal part of the α sequence (residues 106–420) contains the remaining 10 cysteine residues. Comparison of the amino acid sequences of the subunits of periplasmic [Fe] hydrogenase from *D. vulgaris* and *D. vulgaris* ssp. *oxamicus* Monticello (*1, 2*) with that of the putative 65.8-kDa *hydC* gene product (*3*) has indicated that 5 of these 10 cysteine residues, located in elements 3, 4, and 7 (Table II), are conserved and these could therefore serve as ligands to the third active site cluster, which is thought to have an unusual composition (6Fe–6S; Ref. *33*). This conclusion cannot be based on comparison of subunit sequences for the two [Fe] hydrogenases alone, since these are too homologous (~75% identity) to allow firm conclusions, with respect to critically conserved residues, to be drawn. The amino acid sequence of the putative *hydC* gene product is only 20–40% identical. Some caution should be exercised in making comparisons with the *hydC* gene product sequence: this gene has so far only been found in *D. vulgaris* Hildenborough and two closely related strains (*28*). The *hydC* gene, which can be regarded as an in-frame fusion of the *hydA* and *hydB* genes, is present

TABLE II

SOME SEQUENCE ELEMENTS IN THE α AND β SUBUNITS OF PERIPLASMIC HYDROGENASE FROM *D. vulgaris* HILDENBOROUGH THAT HAVE BEEN CONSERVED IN THE *hydC* GENE PRODUCT[a]

Element	Subunit	Sequence[b]	Possible function
1	α	34-KCIgCdtCsqyCptaaif	Coordination 4Fe–4S cluster (F cluster)
2	α	66-aCinCGQClthCPenaiy	Coordination 4Fe–4S cluster (F cluster)
3	α	172-LPqFTSCCPGW	Coordination active site H cluster
4	α	229-VSiMPCiAKKyE	Coordination active site H cluster
5	α	291-GgAtIFGvTGGVMEAALR	?
6	α	346-VKVAVVHGaK	?
7	α	372-FiEyMACPGGCvcGGGQP	Coordination active site H cluster
8	β	80-PlghkSHdLLHThwfDkskgV	?

[a] Conserved residues are indicated by uppercase letters.

[b] The number indicates the position of first residue in the amino acid sequence.

immediately downstream from the *hydA,B* operon. Expression of *hydC* in *D. vulgaris* Hildenborough has yet to be demonstrated and it has been suggested that this gene has been deleted from other *Desulfovibrio* strains (*2*). The *hydC* gene is transcribed convergently with respect to the direction of transcription of the *hydA,B* genes: both genes may share the same transcription terminator (*2, 3*). Recent data by van den Berg *et al.* (*35*) indicated a size of 1.9 kilobases (kb) for the *hydA,B* mRNA transcript in *D. vulgaris* Hildenborough. This size is sufficient to accommodate the *hydA* (1.27 kb) and *hydB* (0.37 kb) genes and confirms that the suggested terminator functions *in vivo*.

The amino acid sequences of the periplasmic [Fe] hydrogenase subunits should ideally be compared with those of the cytoplasmic [Fe] hydrogenases from gram-positive bacteria, such as *Clostridium* and *Megasphaera* (*36, 37*). *Clostridium pasteurianum* has two hydrogenases (I and II). Hydrogenase II is thought to contain three iron–sulfur clusters, two 4Fe–4S ferredoxin-like (F) clusters, and one hydrogenase (H) cluster, thought to contain six Fe. This composition is similar to that of periplasmic [Fe] hydrogenase from *Desulfovibrio*. Hydrogenase I is thought to contain one H-cluster and four F-clusters (*36*) and could thus contain an extra 8Fe–8S ferredoxin domain. Both of these hydrogenases consist of a single polypeptide, which must combine the essential features of both the α and β subunit sequences, e.g., as in the putative *hydC* gene product.

The function of the β subunit in periplasmic [Fe] hydrogenase from *Desulfovibrio* is twofold: (1) it functions in the export of this hydrogenase to the periplasm via its complex 34-amino acid signal sequence (see Section III), and (2) despite the fact that it does not participate in cluster coordination via cysteine residues, it must contribute some essential catalytic function, possibly via element 8 (Table II), which contains several conserved His residues. These could participate as proton conductors to or from the hydrogenase active site during hydrogen production or consumption. Such assignments are, of course, entirely speculative in the absence of an [Fe] hydrogenase structure or biochemical data obtained for mutants, generated by oligonucleotide-directed mutagenesis.

The structure of periplasmic [Fe] hydrogenase from *Desulfovibrio* can thus be interpreted in terms of a simple model. The NH_2-terminal domain functions as the electron conductor via its two F-clusters, whereas the larger COOH-terminal domain contains the H-cluster. The three clusters are spatially arranged to allow rapid flow of electrons from the H-cluster to the F-clusters during hydrogen consumption and vice versa during hydrogen production. The physiological redox partner

of [Fe] hydrogenase in *Desulfovibrio* is thought to be cytochrome c_3, which is exclusively present in the periplasm (*38*) and has four *c*-type hemes in a polypeptide chain of 13 kDa. The structure of this cytochrome is known (*39, 40*). Expression of the *hydA,B* genes in a heterologous system has not yet been successful. Expression in *E. coli* gives rise to an $\alpha\beta$ dimer that does contain part of the F-clusters but lacks the H-cluster (*41*). However, even overexpression in the homologous host *D. vulgaris,* made possible by recently developed bacterial conjugation technology (*42–44*), is difficult: increasing the gene dosage by presenting the *hydA,B* genes on a multiple-copy broad-host-range vector in addition to the chromosomal copy of these genes leads to increased synthesis of α and β subunits but hardly to an increase in enzyme activity (*35, 42*). The conclusions drawn from these experiments are that insertion of the H-cluster and possibly the F-clusters is mediated by the activity of specific protein factors that become limiting when overexpression in *D. vulgaris* is attempted and that are absent from a heterologous host such as *E. coli.* Combined with the need to export periplasmic [Fe] hydrogenase via a specific mechanism (Section III), these features provide an unusually complex assembly mechanism. Structure–function studies by oligonucleotide-directed mutagenesis may not be easily achieved in this system.

B. Nickel-Containing Hydrogenases

The nickel-containing [Ni–Fe] and [Ni–Fe–Se] hydrogenases contain 1 mol of Ni per mole of enzyme and have been characterized by gene cloning and sequencing from a wide variety of sources (Table I). Sequence elements with conserved features are compared in Tables III and IV for 16 different nickel-containing hydrogenases. The hydrogenases in groups I–III are two-subunit enzymes, consisting of a small β (28–35 kDa) and a large α (56–68 kDa) subunit. Those in groups I and II are encoded by simple, two-gene-containing (bicistronic) operons, whereas those in group III are encoded by more complex, polycistronic operons. The enzymes in groups IV and V have a more complex, or unknown, subunit composition: homology with the small subunit sequence of groups I–III is fading in these hydrogenases. It is therefore appropriate to discuss the hydrogenases in groups I–V first of all with reference to the sequences of the "large subunits," since this polypeptide has been found in every nickel-containing hydrogenase sequenced to date. From the discussion below it will become clear that the "large subunit" binds the active site nickel, whereas the "small subunit" has an electron transfer function. Since the designation "small" and "large"

TABLE III

CONSERVED SEQUENCE ELEMENTS IN THE NICKEL-BINDING α SUBUNIT OF Ni-CONTAINING HYDROGENASES

Group	N^a	Gene	Element 1	Element 2	Element 3	Element 4[b]
I	7	*hysA*	44-FRGFEQILRGRDPRDSSQ-IVQRICGVCPTAHCN	110-YLQSHILHFYHLAALDYVK	417-GTGFTEAPRGALLHYL	491-RLVRSYDP*LGCAVHVLHAE
II	4	*hynA*	42-FRGLEMILKGRDPRDAQH-FTQRACGVCTYVHAL	80-YMHDHLVHFYHLHALDWVN	454-GVGEVNAPRGMLSHWI	522-RTVHSYDPCIACGVHVIDPE
	5	*hynA*	58-FRGLEIILKGRDPRDAQH-FTQRTCGVCTYTHAL	120-YLHDHIVHFYHLHALDFVD	469-GVGFVNAPRGGLSHWI	538-RTVHSFDPCIACGVHVIDGH
	6	*hynA*	49-FRGLEIILKGRDPRDAQH-FTQRACGVCTYVHAL	111-YLHDHLVHFYHLHALDWVD	468-GVGLADAPRGSLSHWI	535-RTVHAFDPCIACGVHVIEPE
III	8	*orf2*	52-WRGIEVILKNRDPRDAWA-FTERICGVCTGTHAL	114-QVHDHVVHFYHLHALDWVD	500-GVGFTEAPRGALAHWI	567-RTIHSFDPCLACSTHVMSPD
	9	*hupL*	52-WRGIEVILKNRDPRDAWA-FTERICGVCTGTHAL	114-QVHDHVVHFYHLHALDWVD	500-GVGFTEAPRGALAHWI	567-RTIHSFDPCLACSTHVMSPD
	10	*hupL*	51-WRGLEVILKGRDPRDAWA-FVERICGVCTGTtrw[c]	113-QVqDHIVpFYHLlrLDWVN	505-GVGINEAPRGRSAHWI	572-RTLHSFDPCLACSTHVMSPD
	11	*hoxG*	51-WRGLEVILKGRDPRDAWA-FVERICGVCTGTHAL	113-QVHDHIVHFYHLHALDWVN	508-GVGINEAPRGALGHWI	573-RTLHSFDPCLACSTHVMSPD
	12	*hupL*	51-WRGLEVILKGRDPRDAWA-FTERICGVCTGTHAL	113-QIHDHIVHFYHLHALDWVN	500-GVGMTEAPRGALGHWV	567-RTLHSFDPCLACSTHVMSAE
	13	*hupL*	51-WRGLEVILKGRDPRDAWA-FVERICGVCTGCHAL	113-QVHDHAVHFYHLHALDWVD	521-GVGTVAAPRGMLGHWI	588-RTLHSFDPCLACSTHVMSED
	14	*hyaB*	53-FRGLEIILQGRDPRDARA-FVERICGVCTGVHAL	115-WCHDHLVHFYQLAGMDWID	501-GVGFTEAPRGALGHWA	568-RTLHSFDPCLACSTHVLGDD
IV	16	*hoxH*	38-FRGFEKFVQGHPFWEAPM-FLQRICGICFVSHHL	106-MLQSHTTAYFYLIVPEMLF	382-GVGVVEAPRGTLLHHY	449-VGIRAYDPCLSCATHALGQM
	17	*mvhA*	38-FRGFEKFLQGRPIEEAPR-IVPRICGICDVQHHL	103-YMHSHGLHFYFLAAPDFIA	368-GVGIVEAPRGTLTHHY	434-MVIRAYDPCLSCATHTIDSQ
	18	*frhA*	40-VRGLEKIVTGKAPETAPV-IVQRICGVCPIPHTL	101-HVNSHAIH-HFLIAPDFVP	300-GVGAIEGPRGLDVHMA	372-HVIRAYDPCLSCATHVMVVD
	19	*mvhA*	38-FRGFEKFLQGRRIEEAPR-IVPRICGICGVQHHL	103-YVHSHGLHFYFLAAPDFIG	367-GVGIVEAPRGTLIHHY	434-MVIRAYDPCLSCATHTVDGK
V	15	*orf5*	216-HRGMEKLAETMRGYNEVTFLSDRVCGICGFAHST	279-RLHSHLLNLGLACHFTGFD	469-ALGFAEAPRGDDIHWS	522-LIIGSLDPCYSCTDRMTVVD
	CONSENSUS 1		RG E R CG C H	H	G PRG H	DPC C

[a] See Table I.

[b] The asterisk in the *hysA* sequence indicates selenocysteine.

[c] Sequences that could be in error are indicated in lowercase letters.

TABLE IV

CONSERVED CYSTEINE-CONTAINING SEQUENCE ELEMENTS IN THE ELECTRON-TRANSFERRING β SUBUNIT OF THE Ni-CONTAINING HYDROGENASES

Group	N[a]	Gene	Element 5[b]	Element 6	Element 7	Element 8	Element 9	Element 10
I	7	*hysB*	48-GCTGCSVSLLNAVHP	155-VGTCSAYGGIPAA	192-NVPGCPPHP	240-HENCPYLDKY	262-GCKAEL--GCKGPSTYADCA	294-AVCIGCVEPDFPDGKSPFY
II	4	*hynB*	66-ECTGCSESLLRTVDP	159-IGTCATYGGVQAA	194-NIAGCPPNP	235-HDNCPRLKHF	262-YCLYEL--GCKGPDTYNNCP	294-HPCIACSEPNFWDLYSPFY
	5	*hynB*	66-ECTGCSESVLRAFEP	161-YGTCATFGGVQAA	196-NIAGCPPNP	238-HEQCPRLPHF	265-WCLYEL--GCKGPVTMNNCP	297-HPCIGCSEPDFWDAMTPFY
	6	*hynB*	66-ECTGCTEAAIRTIKP	157-cirh1phGGVQkA	192-NIPGCPPNP	233-HDNCPRLPHF	260-FCLYEL--GCKGPVTYNNCP	292-HPCLGCSEPDFWDTMTPFY
III	8	*orf1*	62-ECTCCS--FIRSAHP	158-WGACASWGCVQAA	191-KVPGCPPIA	233-HDKCYRRPHF	260-YCLYKM--GCKGPTTYNACS	293-HGCIGCSEDGFWD-KGSFY
	9	*hupS*	61-ECTCCSESFIRSAHP	157-WGACASWGCVQAA	190-KVPGCPPIA	232-HDKCYRRPHF	259-YCLYKM--GCKGPTTYNACS	292-HGCIGCSEDGFWD-NGSFY
	10	*hupS*	52-ECTCCSESFIRSGDP	144-WGSCASWGCVQAA	176-KVPGCPPIA	218-HDKsYRRPHF	245-YCLYKV--GCKGPTSYNACS	278-HGCIGCSEDGFWD-KGSFY
	11	*hoxK*	61-ECTCCSESFIRSAHP	157-WGSCASWGCVQAA	190-KVPGCPPIA	232-HDKCYRRPHF	259-YCLYKV--GCKGPTSYNACS	292-HGCIGCSEDGFWD-KGSFY
	12	*hupS*	62-ECTCCSESFIRSAHP	158-WGACASYGCVQAA	191-KVPGCPPIA	233-HDKCYRRPHF	260-YCLYKM--GCKGPTTYNACS	293 HGCIGCSEDGFWD-QGSFY
	13	*hupS*	58-ECTCCSESFIRSAHP	153-WGSCASWGCVQAA	186-KVPGCPPIA	228-HDKCYRRPHF	255-FCLYKV--GCKGPTTYNACS	288-HGarr-SEDGFWD-KGSFY
	14	*hyaA*	61-ECTCCTESFIRSAHP	157-WGTCASWGCVQAA	191-KVPGCPPIP	232-HDKCYRRAHF	259-YCLYKM--GCKGPTTYNACS	292-HGCLGCAENGFWD-RGSFY
IV	17	*mvhG*	14-GCSGCHLSIADFHGK	86-YGTCAVYGGIPGL	149-EVPGCPPRS	181-CEVCPREKPP	184-LCLIPQGLICMGPATVSICG	238-IPCRGCYGPTARVEDQGAK
	19	*mvhG*	n.d[c]	n.d	n.d	181-?EECEREKPP	184-LCLIAQGLVCMGPATTSICG	238-IPCQGCYGPTKAVEDQGAK
		CONSENSUS 2	C C	G C G	GCPP	C	C C GP C	C C
	16	*hoxY*	39-GCWGCTLSFLDMDER	109-VGACAVWGGVPAM	172-FIPGCPPDG	n.p[d]	n.p	n.p
		CONSENSUS 3	C C	G C G	GCPP			
	18	*frhG*	(2 Cys)	90-FGSCAQTGCFTRY	128-AIPGCPPSP	(8 Cys in 8Fe-8S ferredoxin sequence)		
		CONSENSUS 4		G C G	GCPP			
V	15	*orf2*	(16 Cys in 2 8Fe-8S ferredoxin motifs)					
	15	*orf6*	(8 Cys in 8Fe-8S ferredoxin motif; 6 additional Cys)					

[a] See Table I.
[b] Numbering continued from Table III.
[c] n.d., Not determined.
[d] n.p., Not present.

becomes meaningless in the multisubunit hydrogenases of groups IV and V, these polypeptides will be referred to as electron-transferring and nickel-binding (or simply β and α) subunits, respectively.

1. Enzyme Groups and Nickel-Binding Subunit Sequences

As is clear from Table III, the sequences for the α subunits form five groups, with a high degree of sequence homology within each group. Group I comprises only one sequence, that of the [Ni–Fe–Se] hydrogenase of *Desulfovibrio baculatus* (*5, 9*). It is distinct from the three sequences in group II, which represent the [Ni–Fe] hydrogenases of *Desulfovibrio gigas, D. vulgaris* Miyazaki F, and *Desulfovibrio fructosovorans* (*4, 5, 7, 8*). The sequences for both the α and β hydrogenase subunits in this group are highly homologous and have been found to share 65–70% overall sequence identity. This is reflected in the nearly identical sequences for elements 1–4 (Table III).

The genes for the [Ni–Fe] hydrogenase of *D. gigas* (*4, 5*) and the [Ni–Fe–Se] hydrogenase of *D. baculatus* (*5, 9*) were the first to be cloned and sequenced for the class of nickel-containing hydrogenases. The sequence of the [Ni–Fe–Se] enzyme from *D. baculatus* shares only 30–40% overall sequence identity with that of the group II [Ni–Fe] hydrogenases. The 3′ end of *hysA* (Table III, element 4) was found to have an unusual codon, TGA, which normally signals translation termination, but has been shown to encode selenocysteine in formate dehydrogenase from *E. coli* (*45*) and glutathione peroxidase from mouse cells (*46*). The homologous codon in the group II *hynA* sequences is TGC, which codes for cysteine. Spectroscopic studies have indicated coordination of the selenocysteine residue to the active site nickel (*47, 48*). The sequence comparison of [Ni–Fe] and [Ni–Fe–Se] hydrogenases from *D. gigas* and *D. baculatus* did thus establish one of the ligands to Ni (*5*). The (seleno)cysteine is the first of a pair of cysteine residues in the strictly conserved sequence DPCXXC, present in all α subunits of nickel-containing hydrogenases sequenced to date (Table III, consensus 1).

Group III comprises seven [NiFe] hydrogenase sequences from *B. japonicum, Rhizobium leguminosarum, Azotobacter chroococcum, Azotobacter vinelandii, Rhodobacter capsulatus, Rhodocyclus gelatinosus,* and *E. coli*. The sequences for elements 1–4 in this group are again highly homologous (Table III) and strongly resemble those for the [Ni–Fe] hydrogenases in group II. At the DNA level the degree of sequence identity is sufficient to allow the *hya* operon of *E. coli* to be cloned with the use of a group II DNA probe, derived from the *hynB,A* operon of *D. vulgaris* (*18*). Despite these strong similarities there are

two major, related differences between group II and group III hydrogenases. First, the group III [Ni–Fe] hydrogenases are encoded by a polycystronic operon. The *hya* operon of *E. coli* comprises six genes, of which the first two (*hyaA* and *hyaB*) encode the electron-transferring and the nickel-binding subunits, respectively. Reading frame 3 (*hyaC*) encodes a hydrophobic integral membrane protein of 27.6 kDa, which may interact with the *hyaA* and *hyaB* gene products. Evidence that the other six hydrogenases in group III are encoded by a similar polycystronic operon is provided by the observation that a gene homogolous to *hyaC* has been found immediately downstream from the hydrogenase structural genes in *B. japonicum* (*orf3*), *A. chroococcum* (*orf3*), *A. vinelandii* (*orf3*), and *R. capsulatus* (*orfX*). The presence of additional genes downstream from *hupS* and *hupL* has been reported for *R. leguminosarum* (*12*). At the time of writing of this article, the *hya* operon is the only one for which the complete sequence has been published, and it remains to be established whether these other operons also comprise six genes and whether genes 4–6 share homology with *hyaD–hyaF*. This structural difference with group II [Ni–Fe] hydrogenases is precipitated by a different mode of action: group III hydrogenases deliver the electrons derived from hydrogen directly to a membrane-bound electron transport chain, whereas group II hydrogenases donate their electrons to a soluble, nonmembrane-bound, periplasmic cytochrome.

The second difference between group II and III hydrogenases is that the COOH-terminus of the electron-transferring β subunit of group III hydrogenases has an extension of ~50 amino acids, when compared with the group II β subunit sequences. This extension is hydrophobic and serves to anchor the $\alpha\beta$ hydrogenase dimer to the membrane. Isolation of group III [Ni–Fe] hydrogenases therefore requires detergent (to disrupt the membrane) or protease (to cleave the hydrophobic extension) treatment. This structural feature is presumably related to the first and helps in the interaction of the $\alpha\beta$ dimer with the other, membrane-bound components encoded by the operon.

The group IV hydrogenases are isolated as three- or four-subunit enzymes. This group includes the methyl viologen-reducing hydrogenase from the archaebacteria *Methanobacterium thermoautotrophicum* and *Methanothermus fervidus*, a three-subunit enzyme encoded by the *mvhD*, *mvhG*, and *mvhA* genes. The operon encoding this enzyme includes a fourth gene (*mvhB*) that encodes a polyferredoxin (*21, 23*), which does not copurify with the methyl viologen-reducing hydrogenase. The F_{420}-reducing hydrogenase from *M. thermoautotrophicum* also belongs to this group and is encoded by three structural genes (*frhA*, *frhB*, and *frhG*). Finally, the NAD^+-reducing hydrogenase of

the eubacterium *Alcaligenes eutrophus* belongs to group IV. Its four subunits are encoded by the *hoxF, hoxU, hoxY,* and *hoxH* genes, which are organized in the *hoxS* operon (*20*). The nickel-binding subunits in these operons, encoded by the *hoxH, mvhA,* and *frhA* genes, share extensive sequence homology as indicated by the sequences for elements 1–4 in Table III. In element 4, for instance, the sequence IRAY is unique to this group.

Finally, group V is represented by only a single sequence for the nickel-binding subunit of hydrogenase-3 from *E. coli* encoded by *orf5* (*19*). This hydrogenase is, like the hydrogenases in group IV, a cytoplasmic enzyme. Hydrogenase 3 functions in hydrogen production in the formate hydrogenlyase reaction. The functional difference is reflected in the sequence of the nickel-binding subunit of his hydrogenase, which differs appreciably from those in groups I–IV. Highest homologies are observed with the enzymes from group IV (Table III).

Comparison of the sequences for elements 1–4 for all five groups leads to the definition of 18 strictly conserved residues, identical in all 16 sequences (Table III, consensus 1). These will be indicated with Roman numerals (R-I to C-XVIII) in the discussion below. Although it may seem an oversimplification to consider only 18 residues in a chain of 500–600, it must be realized that the actual degree of homology is much higher, since (1) the 18 conserved residues are grouped in four sequences (elements 1–4) that are similarly spaced in all 16 chains, and (2) the other residues in these elements often show only limited variation, as can easily be verified by comparing the sequences in Table III.

Which of these conserved residues coordinate to redox prosthetic groups and which types of redox prosthetic groups are present in the nickel-binding subunits? As indicated above, sequence comparison and spectroscopic measurements have established C-XVI as a nickel ligand. Extended X-ray absorption fine structure (EXAFS) studies have established that the coordination sphere of [Ni–Fe] and [Ni–Fe–Se] hydrogenases is very similarly occupied by 3 ± 1 N,O donors and 2 ± 1 S donors (*47, 49*). As discussed elsewhere, another S (in addition to C-XVI) and one N (e.g., imidazole) are likely to coordinate to Ni (*50*). These could be provided by one of the three strictly conserved cysteine and histidine residues of the consensus 1 sequence. If a histidine residue coordinates to nickel, it is unlikely to be one of the two histidines present in element 4 in group II and group III hydrogenases, since these are not conserved in groups IV and V (Table III). The nickel-binding subunit may also coordinate one Fe–S cluster, in close proximity to Ni. The interpretation of recent EXAFS studies (*49*) indicated that the

active site Ni of [Ni–Fe] and [Ni–Fe–Se] hydrogenase may be present in a Ni, Fe, and S cluster in which the Ni shares sulfur ligands with the Fe–S cluster. The [Ni–Fe] hydrogenase from *D. gigas* has been shown to contain two 4Fe–4S clusters, one 3Fe–4S (or 3Fe–*x*S) cluster, and a nickel atom (*51, 54*). As reviewed elsewhere (*50*), the small β subunit of this two-subunit nickel-containing hydrogenase is likely to coordinate the two electron-transferring 4Fe–4S clusters, whereas the 3Fe–4S cluster is coordinated by the α subunit. Taken together, the data indicate the existence of a Ni, Fe, and S cluster at the active site in the nickel-binding subunit of nickel-containing hydrogenases that is coordinated by the four cysteine residues and some of the other conserved residues indicated in the consensus 1 sequence of Table III. The remaining conserved residues could function as proton conductors in the hydrogen evolution/consumption reaction as discussed in Section II,A for [Fe] hydrogenase. The exact arrangement will only become clear from X-ray crystallographic studies (*55, 56*) and/or from spectroscopic work (e.g., EXAFS) on site-directed mutants.

2. *Electron-Transferring Subunit Sequences*

The sequences of conserved elements 5–10 of the electron-transferring subunit of the nickel-containing hydrogenases are compared in Table IV. It appears that a single comparison for all hydrogenase groups I–V is not meaningful: the electron-transferring subunits of two of the enzymes in group IV lack some conserved elements, whereas hydrogenase 3 (group V) lacks all conserved elements 5–10. Focusing first on groups I–III and the *mvhG* gene products of group IV, it appears that the 13 sequences have 17 strictly conserved positions as indicated in the consensus 2 sequence. These will again be labeled with Roman numerals, C-I to C-XVII, and include 10 cysteine, 4 glycine, and 3 proline residues. Interestingly, there are with the possible exception of the cysteines no conserved proton-conducting residues (e.g., histidines), confirming that the active site of hydrogenase and its proton-conducting channels are located on the nickel binding α subunit: the β subunit has a strictly electron-transferring function. As discussed in Section II,B,1, the 10 conserved cysteine residues are likely to coordinate two 4Fe–4S clusters. Assignment of specific cysteines to these two clusters is difficult because they are spread out over the β subunit amino acid sequence. Although two pairs and a triplet are present (Table IV, elements 5, 9, and 10), an F-cluster binding motif C-X-X-C-X-X-C-X-X-X-C as in [Fe] hydrogenase (Section II,A) is not found. Nevertheless, an assignment can be made when the sequence of the *A. eutrophus hoxY* gene product is considered. This protein shares elements 5, 6, and

7 with the consensus 2 sequence but lacks elements 8, 9, and 10. The *hoxY* gene product (22.9 kDa) is considerably smaller than the 13 consensus 2 electron-transferring subunits (28–35 kDa). Its polypeptide chain terminates immediately following element 7. The *hoxY* gene product binds only a single 4Fe–4S cluster (*20*), which is likely coordinated by the four cysteine residues of the consensus 3 sequence (Table IV). This cluster, referred to as F1 from hereon, is therefore also likely to be coordinated by the four corresponding cysteine residues of the consensus 2 sequence: C-I, C-II, C-IV, and C-VII. The second 4Fe–4S cluster, F2, of the electron-transferring subunit of consensus 2 hydrogenases, could then be coordinated by four of the six remaining conserved cysteine residues of the consensus 2 sequence: C-X, C-XI, C-XII, C-XV, C-XVI, and C-XVII. The variability in sequence of electron-transferring subunits is further demonstrated by considering the 25.7-kDa protein of the F_{420}-reducing hydrogenase of *M. thermoautotrophicum*. It shares the two cysteines (C-IV and C-VII) of elements 6 and 7 with the consensus 3 sequence. However, C-I and C-II are lacking, although two other cysteine residues are present in a sequence that is not homologous to element 5. Nevertheless, these four cysteines could coordinate the F1-cluster. Interestingly, although the COOH-terminal portion of this protein lacks elements 8–10, it does contain eight cysteine residues in an 8Fe–8S ferredoxin motif, which are likely to coordinate two additional F-clusters. The *frhG* gene product may thus coordinate three rather than two electron-transferring clusters. The culmination of these changes is provided by hydrogenase-3, which does not share any of the elements 5–10 in the sequences of the *orf2* and *orf6* gene products, which clearly encode electron transfer proteins since both contain 8Fe–8S ferredoxin motifs (Table IV). It thus appears that though the sequence of the nickel-binding subunit is relatively conserved in nickel-containing hydrogenases, the electron transfer function can be accommodated by a variety of Fe–S cluster-containing redox proteins.

III. Evolution of Hydrogenase Genes: "Redon Shuffling" and Hydrogenase Export

Comparison of the 19 hydrogenase sequences in Section II clearly confirms the existence of two different families, the iron-only and the nickel-containing hydrogenases. There are no significant homologies between the polypeptides encoding the active sites of these enzymes, respectively, the COOH-terminal portion (residues 106–420) of the α and the β subunits (89 residues) of periplasmic iron-only hydrogenases

and the nickel-containing subunit (500–650 residues) of the nickel-containing hydrogenases. Members within a family are related, but the two families must have evolved independently. The hydrogenase from *Anabaena cylindrica* (*24*) is again completely different and forms a third family that will not be considered here.

The significant homology among the nickel-binding subunits of 16 nickel-containing enzymes (Table III, consensus 1), representing different bacterial genera and classified here in five distinct enzyme groups, suggests evolution from a common ancestor. This polypeptide coordinates a Ni, Fe, and S cluster and specifies possibly conserved proton conduction pathways, which are essential for the conversion of hydrogen into protons and electrons or vice versa. It associates generally with one electron-transferring subunit, which is not as strongly conserved and shows a variety of Fe–S cluster coordination patterns. The present data should not be interpreted as indicating that the consensus 2 sequence for this subunit is the most common, since the sample of sequences currently available may not be representative. With this reservation in mind, the data do indicate that the suggested F1-cluster, coordinated by the four cysteines of the consensus 3 sequence (Table III), is more conserved than the F2-cluster, coordinated by cysteine residues in the COOH-terminal domain of the consensus 2 sequence. This observation suggests the following path for the electrons during hydrogen uptake: (Ni, Fe, S) → F1 → F2. The functional equivalent of cluster F2 is likely to reside on one of the other two subunits of the NAD^+-reducing hydrogenase of *A. eutrophus*. The γ subunit (30 kDa) is thought to contain two 4Fe–4S clusters, whereas the δ subunit (63 kDa) contains a 2Fe–2S cluster and a bound FAD. Electrons may thus flow from F1 in the β subunit to clusters in γ and δ subunits and FAD, which then reduces NAD^+. (Note: since α and β denote hydrogenase subunits in this review, all subunit symbols used here have a different meaning from those in Ref. *20*.)

The different forms of the electron-transferring subunits found in group IV and V hydrogenases encoded by *frhG*, *orf2*, and *orf6* may have resulted by shuffling and combining genes encoding smaller redox proteins. Several of these genes have now been cloned and sequenced, e.g., the *rub* gene (156 nucleotides) encoding rubredoxin from *D. vulgaris* Hildenborough encodes a protein of only 52 amino acids, with two pairs of cysteines C-X-X-C present at the NH_2 and COOH terminus (*57*). The *dsr* gene of *D. gigas* (108 nt) encodes desulforedoxin, a protein of only 36 amino acids, which also coordinates one Fe per polypeptide (58b). The 8Fe–8S ferredoxins (50–60 amino acids) are widespread (see Ref. *58* for a review) and their genes (*frd*, 150–160 nt) have been

analyzed from a variety of sources. The mode of coordination of eight cysteine residues to two 4Fe–4S clusters in these small redox proteins has been elucidated by X-ray crystallography (*34*). These small units of DNA (100–200 nt), encoding the smallest possible polypeptide domain coordinating a defined redox prosthetic group, will be referred to as redons. Genes for larger redox proteins can be assembled from these basic units by "redon shuffling." For instance, the *rbo* gene encoding a novel redox protein recently discovered in *D. vulgaris* Hildenborough has the *dsr* redon at its 5′ end (*59*). The *rbr* gene, encoding rubrerythrin from the same organism, has the *rub* redon at its 3′ end (*60*). Fusion of six *frd* redons has given rise to the polyferredoxin gene (*mvhB*) of archaebacteria. The sequence variability of the electron-transferring subunit genes of nickel-containing hydrogenases thus originates from redon shuffling, e.g., the *frhG* gene may have formed by replacing DNA encoding the F2 cluster region by a *frd* redon. The variability found so far is large when one considers the limited number of sequences examined and indicates that other sequences may yet be found.

The electron-transferring subunit of nickel-containing hydrogenases has one other sequence variability option of great consequence: the presence of a complex 30- to 50-amino acid residue signal sequence at its NH_2 terminus destines the enzyme for export to the periplasm. Lack of the signal sequence causes a cytoplasmic location. The enzymes of groups I–III all have signal sequences, which are compared in Table V, and are periplasmic with the possible exception of the [NiFeSe] hydrogenase (*61*). The enzymes of groups IV and V do not have signal sequences and are cytoplasmic. There are two intriguing aspects to the signal sequences in Table V:

1. All sequences contain a strictly conserved element (consensus 5). This is most unusual for signal sequences and suggests that all of these hydrogenases are exported via a unique, conserved mechanism.
2. The nickel-binding subunit (generally referred to in the literature as the large subunit, since all periplasmic hydrogenases are two-subunit enzymes) lacks a signal sequence.

The mechanism of export of hydrogenase has been investigated in some detail (*6, 29, 62*). It appears that a pro-β,α complex may be formed in the cytoplasm or at the cytoplasmic face of the inner membrane, which is then exported, resulting in cleavage of the signal peptide. Thus a single signal peptide operates in the export of both subunits. Recent studies by Niviere *et al.* (*63*), in which DNA encoding the signal peptide for [NiFe] hydrogenase of *D. vulgaris* Hildenborough was fused to the gene for β-lactamase, lacking its natural signal peptide, provided fur-

TABLE V

HYDROGENASE SIGNAL PEPTIDES

Group	N^a	Gene	Signal sequence[b]
			++ + + + +
[Fe]	1	*hydB*	MQIASIT RR G F L K VACVTTGAALIGIRMTGKAVA↓AVK
			++ + + + +
[Fe]	2	*hydB*	MQIVNLT RR G F L K AACVVTAAALISIRMTGKAVA↓AAK
			++ - + -
I	7	*hysB*	MSLS RR E F V K LCSAGVAGLGISQIYHPGIVHA↓MTE
			+ + + --+ -++ ++ - + + -
II	4	*hynB*	MKCYIGRGKNQVEERLERRGVS RR D F M K FCTAVAVAMGMGPAFAPKVAEA↓LTA
			+ +- --+ -+ ++ - + - +
	5	*hynB*	MKISIGLGKEGVEERLAERGVS RR D F L K FCTAIAVTMGMGPAFAPEVARA↓LMG
			+ -++ ++ - + + -
	6	*hynB*	MNFSVGLGRMNAEKRLVQNGVS RR D F M K FCATVAAAMGMGPAFAPKVAEA↓LTA
			- ++ ++ + + -
III	8	*orf1*	MGAATETFYSVIRRQGIT RR S F H K FCSLTATSLGLGPLAASRIANA↓LET
			- ++ ++ + _ _
	9	*hupS*	MATAETFYDVIRRQGIT RR S F T K FCSLTAASLGFGPGAATAMAEA↓LET
			++ ++ + +
	10	*hupS*	MRRQGIT RR S F L K YCSLTGRPCLGPTFAPQIAHA↓MET
			+ - - ++ ++ + + -
	11	*hoxK*	MSRLETFYDVMRRQGIT RR S F L K YCSLTAAALGLGPAFAPRIAHA↓MET
			- - - ++ ++ + + + - -
	12	*hupS*	MMSDIETFYDVMRRQGIT RR S F M K SVRSPQHVLGLGPSFVPKIGEA↓MET
			- - ++ ++ + -
	13	*hupS*	METFYEVMRRQGIS RR S F L K YCSLTATSLGLAPSFVPQIAHA↓MET
			-- ++ ++ + + -
	14	*hyaA*	MNNEETFYQAMRRQGVT RR S F L K YCSLAATSLGLGAGMAPKIAWA↓LEN
			++ +
		CONSENSUS 5	RR F K

[a] See Table I.

[b] Positively (+) and negatively (−) charged residues as well as the signal peptidase cleavage site (↓) are indicated.

ther evidence for a highly specialized export mechanism. In *E. coli* [Ni–Fe], hydrogenases 1 and 3 are only expressed under anaerobic conditions. It appeared that the constructed fusion was only efficiently exported and processed under anaerobic conditions. The results supported the theory that under these conditions at least one protein was expressed that specifically facilitates hydrogenase export and processing.

Thus, assuming that a cytoplasmic, nickel-containing hydrogenase existed first, the evolutionary path toward a periplasmic enzyme left the nickel-binding subunit and its assembly locus relatively unchanged. Instead, the malleable electron-transferring subunit was equipped with

a unique signal peptide and a specific export mechanism evolved, accommodating its own export and that of the nickel-binding subunit following cytoplasmic assembly of the (Ni, Fe, S) cluster in the latter.

The limited number of [Fe] hydrogenase sequences makes it harder to speculate on their evolution. A gene for an H-cluster binding protein (encoding residues 200–606 of the putative *hydC* gene product; Ref. *3*) may have fused with a *frd* redon to create a hydrogenase II-like gene. This gene encodes a simple cytoplasmic, single-subunit hydrogenase, which accommodates both the H-cluster and two F-clusters. The electron transfer function may also be less strictly conserved in cytoplasmic [Fe] hydrogenases, as observed for the nickel-containing enzymes (Section II,B and above). Fusion with a second *frd* redon at the 5′ end of the hydrogenase II gene may have created the hydrogenase I gene, which encodes a cytoplasmic hydrogenase with four F-clusters (see Section V). Periplasmic localization was achieved, by fusing DNA encoding the specific signal peptide (Table V) near the 3′ end of a hydrogenase II-like gene, such that a two-subunit enzyme was created. Assembly and export of this enzyme is through a similar mechanism as described for the group I–III nickel-containing hydrogenases: assembly of F- and H-clusters in the α subunit, binding of pro-β, and export and processing. It is indeed remarkable that the two classes of periplasmic hydrogenases do not share sequence homology, except the consensus 5 sequence (Table V).

IV. Functions of Hydrogenases

The main purpose of the speculations in Section III was to draw together, in an evolutionary model, the structures of the various hydrogenases that exist today. Such models, e.g., the evolution of genes for a periplasmic hydrogenase from those for a cytoplasmic hydrogenase, are generally hard to prove, and Section III will therefore remain largely speculative. Although it would appear that it must be easier to define the function of hydrogenases presently occurring in prokaryotes (Table I), this is actually difficult in cases in which multiple enzymes are present and/or when the organism lives both fermentatively (favoring hydrogen production) and respiratively (favoring hydrogen consumption).

The most straightforward case is that of the group III nickel-containing hydrogenases. These are membrane-bound, periplasmic-uptake hydrogenases that deliver the electrons from hydrogen to the membrane-bound electron transport chain, where they eventually reduce a

higher potential electron acceptor A (e.g., fumarate, nitrate, and oxygen). The overall reaction catalyzed by hydrogenase and the electron transport chain is thus:

$$H_2(\text{out}) + nH^+(\text{in}) + A \rightarrow (n + 2)H^+(\text{out}) + A^{2-} \quad (1)$$

Equation (1) indicates that the flow of two electrons from hydrogen to the electron acceptor A leads to free-energy conservation by coupled export of n protons to the periplasm, where n will increase with an increased difference in reduction potentials of the $2H^+/H_2$ and the A/A^{2-} couples. If these potentials are similar, then $n \approx 0$ and there may be no need for tight structural coupling of hydrogenase with the membrane-bound electron transport chain. This could be one of the main reasons for the structural differences of group II nickel-containing hydrogenases of sulfate-reducing bacteria and those of group III (Section II,B,1).

Since the hydrogenase uptake reaction releases two protons into the periplasm, the total number of protons released is $(n + 2)$, and assuming that m protons flow back via ATP synthase per ATP synthesized from ADP and P_i, a total of $(n + 2)/m$ moles of ATP can be formed per mole of hydrogen oxidized. From this discussion it would appear energetically advantageous to locate uptake hydrogenases in the periplasm. However, this may be an oversimplification: periplasmic and cytoplasmic compartments are kept at a different pH and potential, such that conversion of 1 mol of hydrogen to 2 mol of protons in the cytoplasm is accompanied by a larger decrease in free energy of the system (ΔG_c) than when this reaction is carried out in the periplasm (ΔG_p). A thermodynamic "cycle" in which cytoplasmic conversion is followed by export of 2 mol of protons (positive free energy, ΔG_e) could be energetically equivalent ($\Delta G_c + \Delta G_e = \Delta G_p$), but such a mechanism does of course require a proton pump. In conclusion, there may be no thermodynamic arguments against cytoplasmic uptake hydrogenases and these do indeed occur, e.g., those of archaebacteria (*21–23, 64*) and the enzyme of *Desulfotomaculum orientis* (*65*).

One of the functions of group III nickel-containing hydrogenases has been to trap hydrogen produced in the cytoplasm by another reaction, e.g., the fixation of nitrogen in *Azotobacter, Rhizobium,* and other species (*66*):

$$N_2 + 8H^+ + 8e + 16\text{MgATP} \rightarrow 2NH_3 + H_2 + 16\text{MgADP} + 16P_i \quad (2)$$

Azotobacter eutrophus appears to have a group III nickel-containing uptake hydrogenase, in addition to its NAD^+-reducing hydrogenase (*20, 67*). The latter is not present in all members of the autotrophic

hydrogen-oxidizing bacteria (*67*). Of course, both NADH and ATP are needed in the fixation of CO_2 into carbohydrate and the combined action of these two hydrogenases may well provide a correct ratio of these two prerequisites, e.g., as provided by the combined action of photosystems I and II in plants.

The function of hydrogenase 3 has already been briefly mentioned in Section II,B,1. It is the only enzyme of Table I that has a firmly established function in hydrogen production [see Eq. (3)]. In addition to hydrogenase 1, *E. coli* has a second, nickel-containing uptake hydrogenase, hydrogenase 2. No structural information is available for this enzyme at present. The sequence of the hydrogenase 2 operon has not yet been reported. The enzyme is immunologically distinct from hydrogenase 1 (*26, 68*). Hydrogenase 1 is expressed under the same conditions that lead to expression of hydrogenase 3 and is, therefore, suggested to function in the uptake of hydrogen produced by hydrogenase 3 during the formate hydrogenlyase reaction:

$$HCOOH \rightarrow H_2 + CO_2 \qquad (3)$$

Hydrogenase 2 is thought to couple hydrogen oxidation to, e.g., fumarate reduction under anaerobic conditions. Hydrogenase 2 is present but is believed to be inactive during fermentative growth in the presence of formate (*69*), and it has been suggested that, under these conditions, hydrogenase 1 recycles the fermentatively produced hydrogen by reducing endogenously generated electron acceptors (i.e., fumarate) (*69*). The rationale for having these two enzymes, which seem to share the same terminal electron acceptor(s), is thus not clear at present.

It is similarly difficult to prove definitively the roles of the various hydrogenases in the sulfate-reducing bacteria *Desulfovibrio*. As indicated in Table VI, the genes for periplasmic [Fe], periplasmic [Ni–Fe], and [Ni–Fe–Se] hydrogenase are distributed such that four classes arise (*28*). All contain the group II nickel-containing hydrogenase, which is the exclusive enzyme in the class 4 *Desulfovibrio* species. It has been proposed elsewhere that both the periplasmic [NiFe] and [Fe] hydrogenase of *Desulfovibrio* serve as hydrogen uptake enzymes (*6*), but that the [Fe] enzyme acts only at high hydrogen concentrations in view of its low H_2 affinity ($K_m \approx 100\ \mu M$) compared to the [Ni–Fe] enzyme ($K_m \approx 1\ \mu M$). The advantage of expressing periplasmic [Fe] hydrogenase for class 1 and 2 strains (Table VI) would then be that at high hydrogen concentration hydrogen can be taken up at a high rate, since the turnover number of [Fe] hydrogenase is ≈30-fold higher than that of [Ni–Fe] hydrogenase. This proposal may apply to *D. vulgaris* Hildenborough, in which 95% of the periplasmic hydrogenase activity

TABLE VI

DISTRIBUTION OF HYDROGENASE GENES IN *Desulfovibrio*

Class 1:	[Fe], [Ni–Fe–Se], and [Ni–Fe] hydrogenase (9 species)
	Desulfovibrio vulgaris Hildenborough
	Desulfovibrio vulgaris Wandle
	Desulfovibrio vulgaris Brockhurst Hill
	Desulfovibrio vulgaris ssp. *oxamicus* Monticello 2
	Desulfovibrio desulfuricans Berre Sol
	Desulfovibrio desulfuricans Canet 41
	Desulfovibrio desulfuricans G200
	Desulfovibrio africanus Walvis Bay
	Desulfovibrio africanus Bhengazi
Class 2:	[Fe] and [Ni–Fe] hydrogenase (4 species)
	Desulfovibrio desulfuricans El Agheila Z
	Desulfovibrio desulfuricans NCIMB 8307
	Desulfovibrio multispirans
	Desulfovibrio fructosovorans
Class 3:	[Ni–Fe–Se] and [Ni–Fe] hydrogenase (6 species)
	Desulfovibrio vulgaris Miyazaki F
	Desulfovibrio vulgaris ssp. *oxamicus* UofA
	Desulfovibrio desulfuricans Norway 4
	Desulfovibrio salexigens British Guiana
	Desulfovibrio salexigens California
	Desulfovibrio salexigens NCIMB 8365
Class 4:	[Ni–Fe] hydrogenase (3 species)
	Desulfovibrio vulgaris Groningen
	Desulfovibrio desulfuricans Teddington R
	Desulfovibrio gigas

(measured at high hydrogen concentration) is due to [Fe] hydrogenase (*35*) and the remaining 5% is due to the two nickel-containing hydrogenases. However, in *Desulfovibrio fructosovorans,* which contains only the [Fe] and [Ni–Fe] hydrogenase, the latter was found responsible for 90% of the total uptake hydrogenase activity (*70*). Inactivation of the *hynB,A* genes by marker exchange mutagenesis (Ref. *70;* the first time directed gene inactivation has been achieved in this genus) caused a lag phase in the growth on hydrogen sulfate medium. However, the cells grew to the same density, indicating that hydrogen uptake through [Fe] hydrogenase was equally efficient thermodynamically. Repression of [Fe] hydrogenase in *D. vulgaris* Hildenborough by expression of *hydA,B* antisense mRNA slowed bacterial growth on lactate–sulfate medium and caused a reduced accumulation of H_2 in the

medium head space (*35*). *Desulfovibrio vulgaris* Hildenborough is a net hydrogen producer when grown in this medium. These results were interpreted as indicating that the function of [Fe] hydrogenase in *D. vulgaris* Hildenborough is hydrogen production, a function attributed to [Ni–Fe–Se] hydrogenase by others (*61*). It would be worthwhile to test growth of the [Fe] hydrogenase-repressed *D. vulgaris* Hildenborough on hydrogen–sulfate medium (*71*), on which *D. vulgaris* Hildenborough can also grow (R. K. Thauer, personal communication, 1991).

A final problem to be considered here in the function of hydrogenases in *Desulfovibrio* is how the electrons make their way to the cytoplasm, where the sulfate is reduced. As indicated in Section II,B,1, all three hydrogenase types are encoded by simple bicistronic operons that do not contain genes whose products could conduct the electrons through the membrane. The physiological electron carrier for hydrogenases in *Desulfovibrio* is thought to be cytochrome c_3, a 13-kDa *c*-type cytochrome that binds four *c*-type hemes and is exclusively present in the periplasm. The problem is that the next electron carrier in the chain (hydrogenase → cytochrome c_3 → ?) is not known. The structure of the high-molecular-weight cytochrome (*72*), which is also a periplasmic protein, was recently determined by cloning and sequencing its gene (*73*). It appeared that the sequence of this protein, which binds 16 hemes covalently to a polypeptide chain of 58.9 kDa, can be described in terms of four cytochrome c_3-like domains. Three complete domains, coordinating four *c*-type hemes similarly as cytochrome c_3, and one incomplete domain, which may have a high-potential histidine–methionine coordinate heme. Although this more complex structure does not by itself solve the above problem, it was subsequently found (W. B. R. Pollock, 1992, unpublished) that Hmc is encoded by the first gene (*hmc*) in an operon that contains at least six open reading frames, encoding proteins Orf1–Orf6. Orf1 (Hmc) is the periplasmic high-molecular-weight cytochrome; Orf2 is a largely periplasmic iron–sulfur protein anchored to the membrane; Orf3, Orf4, and Orf5 are integral membrane proteins; and Orf6 is a cytoplasmic iron–sulfur protein. Thus the operon structure suggests that all proteins, Orf1–Orf6, may be physically interacting and that the *hmc* operon encodes the link between the periplasmic hydrogenases and the cytoplasmic redox chain in *D. vulgaris*.

V. Perspectives

Sections I–IV of this review were completed just prior to the Third International Conference on Molecular Biology of Hydrogenases in

Troia, Portugal, from July 29 to August 1, 1991. At this meeting, Thauer and co-workers reported on a novel family of hydrogenases in archaebacteria in which the presence of a metal ion or other redox prosthetic group has yet to be demonstrated. Extensive nucleic acid sequencing of the [Ni–Fe] hydrogenase operons of *R. capsulatus* (Vignais and co-workers), *A. vinelandii* (Mortenson and co-workers), and *R. leguminosarum* (Ruiz-Argueso and co-workers) revealed that these comprise ~15 genes. In addition to the two hydrogenase structural genes and the equivalent of the *E. coli hyaC* gene, these include genes for nickel-processing proteins and genes homologous to open reading frames of the *E. coli hya, hyp,* and hydrogenase 3 operons. Possibly, these large operons specify all necessary functions to form an active periplasmic [NiFe] hydrogenase in these organisms. In *E. coli,* which likely requires additional genes to specify its three hydrogenases, these genes have been scattered over different genomic loci. The observation of widespread homology among the multitude of genes required for hydrogenase formation in different organisms makes a more uniform gene nomenclature desirable and a proposal in this direction is being prepared by P. Vignais. Kroger and co-workers reported the sequence of the [Ni–Fe] hydrogenase genes from *Wolinella succinogenes,* belonging to group III (Table III), and showed that the *E. coli hyaC* equivalent gene expresses a membrane-bound, *b*-type cytochrome. The sequence of hydrogenase I from *C. pasteurianum* was reported by Meyer and Gagnon (*74*). It showed striking homologies with the *hydC* and *hydA,B* gene products and confirmed the conserved elements listed in Table II. Hydrogenase I shares an additional seven conserved cysteine residues with the *hydC* gene product in a sequence located on the N-terminal side of elements 1 and 2 (Table II). These residues are not present in an 8Fe–8S ferredoxin motif, but must, in view of their conservation, contribute to the coordination of the two additional F-clusters of hydrogenase I. Site-directed mutagenesis of the *E. coli hyaB* gene (Przybyla and co-workers) confirmed several of the consensus-1 residues (Table III: R-IV, C-VII, D-XV, C-XVII, and C-XVIII) as essential. These and other presentations confirmed the great progress that has been made in the molecular biology of hydrogenases. What is still lacking is a three-dimensional structure and it must be hoped that current efforts (*55, 56*) to solve the structure of group II [Ni–Fe] hydrogenases will be successful soon. Lacking also is a firm understanding of the role of individual enzymes in organisms expressing multiple hydrogenases (e.g., *Desulfovibrio*). This understanding may be achieved by progress in directed mutagenesis studies and the achievement of a more thorough thermodynamic description of the metabolic energy transformations in these bacteria.

ACKNOWLEDGMENTS

Work in the author's laboratory was supported financially by the National Science and Engineering Research Council of Canada and intellectually by Michael Brumlik, Harm Deckers, Vincent Niviere, Brent Pollock, Hansje Voordouw, and Frankie Wilson.

REFERENCES

1. Voordouw, G., and Brenner, S., *Eur. J. Biochem.* **148,** 515 (1985).
2. Voordouw, G., Strang, J. D., and Wilson, F. R., *J. Bacteriol.* **171,** 3881 (1989).
3. Stokkermans, J., van Dongen, W., Kaan, A., van den Berg, W., and Veeger, C., *FEMS Microbiol. Lett.* **58,** 217 (1989).
4. Li, C., Peck, H. D., Jr., LeGall, J., and Przybyla, A. E., *DNA* **6,** 539 (1987).
5. Voordouw, G., Menon, N. K., LeGall, J., Choi, E.-S., Peck, H. D., Jr., and Przybyla, A. E., *J. Bacteriol.* **171,** 2894 (1989).
6. Voordouw, G., *in* "Microbiology and Biochemistry of Strict Anaerobes Involved in Interspecies Hydrogen Transfer" (J. P. Belaich, M. Bruschi, and J. L. Garcia, eds.), p. 37. Plenum, New York, 1990.
7. Deckers, H. M., Wilson, F. R., and Voordouw, G., *J. Gen. Microbiol.* **136,** 2021 (1990).
8. Rousset, M., Dermoun, Z., Hatchikian, C. E., and Belaich, J. P., *Gene* **94,** 95 (1990).
9. Menon, N. K., Peck, H. D., Jr., LeGall, J., and Przybyla, A. E., *J. Bacteriol.* **169,** 5401 (1987).
10. Sayavedra-Soto, L. A., Powell, G. K., Evans, H. J., and Morris, R. O., *Proc. Natl. Acad. Sci. U.S.A.* **85,** 8395 (1988).
11. Schneider, C. G., Schmitt, H. J., Schild, Ch., Tichy, H. V., and Lotz, W., *Nucleic Acids Res.* **18,** 5285 (1990).
12. Hidalgo, E., Levya, A., and Ruiz-Argueso, T., *Plant. Mol. Biol.* **15,** 371 (1990).
13. Ford, C. M., Garg, N., Garg, R. P., Tibelius, K. H., Yates, M. G., Arp, D. J., and Seefeldt, L. C., *Mol. Microbiol.* **4,** 999 (1990).
14. Menon, A. L., Stults, L. W., Robson, R. L., and Mortenson, L. E., *Gene* **96,** 67 (1990).
15. Leclerc, M., Colbeau, A., Cauvin, B., and Vignais, P. M., *Mol. Gen. Genet.* **214,** 97 (1988).
16. Richaud, P., Vignais, P. M., Colbeau, A., Uffen, R. L., and Cauvin, B., *FEMS Microbiol. Rev.* **87,** 413 (1990).
17. Uffen, R. L., Colbeau, A., Richaud, P., and Vignais, P. M., *Mol. Gen. Genet.* **221,** 49 (1990).
18. Menon, N. K., Robbins, J., Peck, H. D., Jr., Chatelus, C. Y., Choi, E.-S., and Przybyla, A. E., *J. Bacteriol.* **172,** 1969 (1990).
19. Bohm, R., Sauter, M., and Bock, A., *Mol. Microbiol.* **4,** 231 (1990).
20. Tran-Betcke, A., Warnecke, U., Bocker, C., Zaborosch, C., and Friedrich, B., *J. Bacteriol.* **172,** 2920 (1990).
21. Reeve, J. N., Beckler, G. S., Cram, D. S., Hamilton, P. T., Brown, J. W., Krzycki, J. A., Kolodziej, A. F., Alex, L., Orme-Johnson, W. H., and Walsh, C. T., *Proc. Natl. Acad. Sci. U.S.A.* **86,** 3031 (1989).
22. Alex, L. A., Reeve, J. N., Orme-Johnson, W. H., and Walsh, C. T., *Biochemistry* **29,** 7237 (1990).
23. Steigerwald, V. J., Beckler, G. S., and Reeve, J. N., *J. Bacteriol.* **172,** 4715 (1990).

24. Ewart, G. D., Reed, K. C., and Smith, G. D., *Eur. J. Biochem.* **187,** 215 (1990).
25. Prickril, B. C., He, S.-H., Li, C., Menon, N., Choi, E. S., Przybyla, A. E., DerVartanian, D. V., Peck, H. D., Jr., Fauque, G., LeGall, J., Texeira, M., Moura, I., Moura, J. J. G., Patil, D., and Huyn, B. J., *Biochem. Biophys. Res. Commun.* **149,** 369 (1987).
26. Sawers, R. G., and Boxer, D. H., *Eur. J. Biochem.* **156,** 265 (1986).
27. Stoker, K., Oltmann, L. F., and Stouthamer, A. H., *J. Bacteriol.* **170,** 1220 (1988).
28. Voordouw, G., Niviere, V., Ferris, F. G., Fedorak, P. M., and Westlake, D. W. S., *Appl. Environ. Microbiol.* **56,** 3748 (1990).
29. Voordouw, G., *in* "The Nitrogen and Sulphur Cycles" (J. A. Cole and S. Ferguson, eds.), p. 147. Soc. Gen. Microbiol. Symp. 42. Cambridge Univ. Press, Cambridge, 1988.
30. Voordouw, G., and Wall, J. D., *in* "Genetics and Molecular Biology of Anaerobic Bacteria" (M. Sebald, ed.), in press. Springer-Verlag, New York, 1991.
31. Fauque, G., Peck, H. D., Jr., Moura, J. G., Huynh, B. H., Berlier, Y., DerVartanian, D. V., Texeira, M., Przybyla, A. E., Lespinat, P. A., Moura, I., and LeGall, J., *FEMS Microbiol. Rev.* **54,** 299 (1988).
32. Reeve, J. N., and Beckler, G. S., *FEMS Microbiol. Rev.* **87,** 419 (1990).
33. Hagen, W. R., van Berkel-Arts, A., Kruse-Wolters, K. M., Voordouw, G., and Veeger, C., *FEBS Lett.* **203,** 59 (1986).
34. Adman, E. T., Sieker, L. C., and Jensen, L. H., *J. Biol. Chem.* **259,** 7045 (1973).
35. van den Berg, W. A. M., van Dongen, W. M. A. M., and Veeger, C., *J. Bacteriol.* **173,** 3688 (1991).
36. Adams, M. W. W., Eccleston, E., and Howard, J. B., *Proc. Natl. Acad. Sci. U.S.A.* **86,** 4932 (1989).
37. van Dijk, C., Mayhew, S. G., Grande, H. J., and Veeger, C., *Eur. J. Biochem.* **102,** 317 (1979).
38. Voordouw, G., and Brenner, S., *Eur. J. Biochem.* **159,** 347 (1986).
39. Pierrot, M., Haser, R., Frey, M., Payan, F., and Astier, J. P., *J. Biol. Chem.* **257,** 14341 (1982).
40. Higuchi, Y., Kusunoki, M., Matsuura, Y., Yasuoka, W., and Kakudo, M., *J. Mol. Biol.* **172,** 109 (1984).
41. Voordouw, G., Hagen, W. R., Kruse-Wolters, M., van Berkel-Arts, and Veeger, C., *Eur. J. Biochem.* **162,** 31 (1987).
42. Van den Berg, W. A. M., Stokkermans, J. P. W. G., and van Dongen, W. M. A. M., *J. Biotechnol.* **12,** 173 (1989).
43. Powell, B., Mergeay, M., and Christofi, N., *FEMS Microbiol. Lett.* **59,** 269 (1989).
44. Voordouw, G., Pollock, W. B. R., Bruschi, M., Guerlesquin, F., Rapp-Giles, B. J., and Wall, J. D., *J. Bacteriol.* **172,** 6122 (1990).
45. Zinoni, F., Birkman, A., Stadtman, T. C., and Bock, A., *Proc. Natl. Acad. Sci. U.S.A.* **83,** 4650 (1986).
46. Chambers, I., Frampton, J., Goldfarb, P., Affara, N., McBain, W., and Harrison, P. R., *EMBO J.* **5,** 1221 (1986).
47. Eidsness, M. K., Scott, R. A., Prickril, B., DerVartanian, D. V., LeGall, J., Moura, I., Moura, J. J. G., and Peck, H. D., Jr., *Proc. Natl. Acad. Sci. U.S.A.* **86,** 147 (1989).
48. He, S.-H., Texeira, M., LeGall, J., Patil, D. S., DerVartanian, D. V., Huyn, B. H., and Peck, H. D., Jr., *J. Biol. Chem.* **264,** 2678 (1989).
49. Maroney, M. J., Colpas, G. J., Bagyinka, N., and Mascharak, P. K., *J. Am. Chem. Soc.* **113,** 3962 (1991).
50. Hatchikian, E. C., Fernandez, V. M., and Cammack, R., *in* "Microbiology and

Biochemistry of Strict Anaerobes Involved in Interspecies Hydrogen Transfer" (J. P. Belaich, M. Bruschi, and J. L. Garcia, eds.), p. 53. Plenum, New York, 1990.
51. Cammack, R., Patil, D. S., Hatchikian, E. C., and Fernandez, V. M., *Biochim. Biophys. Acta* **912,** 98 (1987).
52. Fernandez, V. M., Hatchikian, E. C., Patil, D. S., and Cammack, R., *Biochim. Biophys. Acta* **883,** 145 (1986).
53. Cammack, R., *Adv. Inorg. Chem.* **32,** 297 (1988).
54. Huynh, B. H., Patil, D. S., Moura, I., Texeira, M., Moura, J. J. G., Dervartanian, D. V., Czechowski, M. H., Peck, H. D., Jr., and LeGall, J., *J. Biol. Chem.* **262,** 795 (1987).
55. Higuchi, Y., Yasuoka, N., Kakudo, M., Katsube, Y., Yagi, T., and Inokuchi, H., *J. Biol. Chem.* **262,** 2823 (1987).
56. Niviere, V., Hatchikian, C., Cambilleau, C., and Frey, M., *J. Mol. Biol.* **195,** 969 (1987).
57. Voordouw, G., *Gene* **69,** 75 (1988).
58. Bruschi, M., and Guerlesquin, F., *FEMS Microbiol. Rev.* **54,** 155 (1988).
58b. Brumlik, M. J., LeRoy, G., Bruschi, M., and Voordouw, G., *J. Bacteriol.* **172,** 7289 (1990).
59. Brumlik, M. J., and Voordouw, G., *J. Bacteriol.* **171,** 4996 (1989).
60. Prickril, B. C., Kurtz, D. M., Jr., LeGall, J., and Voordouw, G., *Biochemistry* **30,** 11118 (1991).
61. Rohde, M., Furstenau, U., Mayer, F., Przybyla, A. E., Peck, H. D., Jr., LeGall, J., Choi, E.-S., and Menon, E. K., *Eur. J. Biochem.* **191,** 389 (1990).
62. Van Dongen, W., Hagen, W., van den Berg, W., and Veeger, C., *FEMS Microbiol. Lett.* **50,** 5 (1988).
63. Niviere, V., Wong, S.-L., and Voordouw, G., submitted for publication (1992).
64. Thauer, R. K., Jungermann, K., and Decker, K., *Bacteriol. Rev.* **41,** 100 (1977).
65. Cypionka, H., and Dilling, W., *FEMS Microbiol. Lett.* **36,** 257 (1986).
66. Evans, H. J., Harker, A. R., Papen, H., Russell, S. A., Hanus, F. J., and Zuber, M., *Annu. Rev. Microbiol.* **41,** 335 (1987).
67. Bowien, B., and Schlegel, H. G., *Annu. Rev. Microbiol.* **35,** 405 (1981).
68. Ballantine, S. P., and Boxer, D. H., *Eur. J. Biochem.* **156,** 277 (1986).
69. Sawers, R. G., Jamieson, D. J., Higgins, C. F., and Boxer, D. H., *J. Bacteriol.* **168,** 398 (1986).
70. Rousset, M., Dermoun, Z., Chippaux, M., and Belaich, J. P., *Mol. Microbiol.* **5,** 1735 (1991).
71. Badziong, W., Thauer, R. K., and Zeikus, J. G., *Arch. Microbiol.* **116,** 41 (1978).
72. Higuchi, Y., Inaka, K., Yasuoka, N., and Yagi, T., *Biochim. Biophys. Acta* **911,** 341 (1987).
73. Pollock, W. B. R., Loutfi, M., Bruschi, M., Rapp-Giles, B. J., Wall, J. D., and Voordouw, G., *J. Bacteriol.* **173,** 220 (1991).
74. Meyer, J., and Gagnon, *Biochemistry* **30,** 9697 (1991).