Cloning and sequencing of a [NiFe] hydrogenase operon from Desulfovibrio vulgaris Miyazaki F

HARM M. DECKERS, FRANKIE R. WILSON and GERRIT VOORDOUW*

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

(Received 29 March 1990; revised 21 May 1990; accepted 12 June 1990)

A hydrogenase operon was cloned from chromosomal DNA isolated from Desulfovibrio vulgaris Miyazaki F with the use of probes derived from the genes encoding [NiFe] hydrogenase from Desulfovibrio vulgaris Hildenborough. The nucleic acid sequence of the cloned DNA indicates this hydrogenase to be a two-subunit enzyme: the gene for the small subunit (267 residues; molecular mass = 28763 Da) precedes that for the large subunit (566 residues; molecular mass = 62495 Da), as in other [NiFe] and [NiFeSe] hydrogenase operons. The amino acid sequences of the small and large subunits of the Miyazaki hydrogenase share 80% homology with those of the [NiFe] hydrogenase from Desulfovibrio gigas. Fourteen cysteine residues, ten in the small and four in the large subunit, which are thought to co-ordinate the iron-sulphur clusters and the active-site nickel in [NiFe] hydrogenases, are found to be conserved in the Miyazaki hydrogenase. The subunit molecular masses and amino acid composition derived from the gene sequence are very similar to the data reported for the periplasmic, membrane-bound hydrogenase isolated by Yagi and coworkers, suggesting that this hydrogenase belongs to the general class of [NiFe] hydrogenases, despite its low nickel content and apparently anomalous spectral properties.

Introduction

Two classes of hydrogenases, the iron-only and nickel-containing hydrogenases have been found in sulphate-reducing bacteria. The cloning and sequencing of the structural genes for these hydrogenases has contributed considerably to their characterization and classification. Genes for the [Fe] hydrogenase have been isolated from two species, D. vulgaris Hildenborough and D. vulgaris subsp. examicus Monticello (Voordouw et al., 1985, 1989b; Voordouw & Brenner, 1985). [Fe] hydrogenase is a periplasmic two-subunit enzyme (molecular mass = 46 kDa and molecular mass = 10 kDa for the large and small subunits, respectively). The small subunit has a 34 amino acid signal peptide for export of this hydrogenase to the periplasm (Prickril et al., 1986). The [Fe] hydrogenase subunits do not show homology with those of the nickel-containing hydrogenases, which comprise two subclasses, the [NiFe] and [NiFeSe] hydrogenases. The nucleotide sequence of the structural genes of the [NiFe] hydrogenase from D. gigas (Li et al., 1987; Voordouw et al., 1989a) indicates molecular masses of 28 and 61 kDa for the processed small and the large subunit, respectively. The former has a 50 amino acid signal sequence for export of this hydrogenase. D. gigas hydrogenase is relatively easily extracted from the Desulfovibrio periplasm (Hatchikian et al., 1978) and shares sequence homology with [NiFe] hydrogenases isolated and characterized from bacterial genera other than Desulfovibrio, i.e. Bradyrhizobium japonicum (Saya-vedra-Soto et al., 1988), Azotobacter chroococcum (Ford et al., 1990), Rhodobacter capsulatus (Leclerc et al., 1988), Rhodocyclus gelatinosus (Uffen et al., 1990) and Escherichia coli (Menon et al., 1990). The [NiFe] hydrogenases of these latter micro-organisms are generally found to be periplasmic and membrane-bound: they can be isolated only after trypsin and/or detergent treatment of cell membranes. The COOH terminus of these hydrogenases appears to have an extension of 40–50 amino acid residues, when compared with the D. gigas sequence. This extension has a central stretch of 20 hydrophobic amino acids (Menon et al., 1990) and may function in the binding of these enzymes to the membrane.

Hydrogenases containing both selenium and nickel have been isolated from several species of Desulfovibrio and the cloning and sequencing of the genes from
D. baculatus has been reported (Menon et al., 1987; Voordouw et al., 1989a). The sequences of the small (molecular mass = 31 kDa) and large (molecular mass = 57 kDa) subunits are homologous to those of the D. gigas [NiFe] hydrogenase, although the degree of sequence identity is less (38 and 34%, respectively, Voordouw et al., 1989a) than that found within the [NiFe] hydrogenase family. The small subunit of [NiFeSe] hydrogenase has a 32 amino acid signal peptide, which shares sequence homology with those for the [NiFe] and [Fe] hydrogenase small subunits.

The hydrogenase from D. vulgaris Miyazaki F has been extensively studied by Yagi and co-workers (Yagi, 1970; Yagi et al., 1976, 1985; Higuchi et al., 1987). The subunit molecular masses (29 kDa and 60 kDa) and amino acid composition are similar to that of D. gigas hydrogenase (Yagi et al., 1976; Hatchikian et al., 1978). Nevertheless, the nickel content appears to be very low (0.16 mol Ni mol⁻¹) and spectroscopic characteristics have been interpreted as an indication of the absence of bound nickel (Yagi et al., 1985). In earlier studies the enzyme was shown to be inhibited by carbon monoxide, a characteristic of [Fe] hydrogenases (Yagi et al., 1976), while lack of inhibition by carbon monoxide is diagnostic for the nickel-containing hydrogenases (Gow et al., 1986; Lisso& et al., 1986). The properties of the Miyazaki F hydrogenase described by Yagi and co-workers are thus unusual and a further investigation of hydrogenases in this strain with molecular biological methods is appropriate. In this study we investigate the possible presence of [Fe] and [NiFe] hydrogenase genes in this species with currently available probes and describe the cloning and sequencing of a hydrogenase gene with high homology to the [NiFe] hydrogenase genes of D. gigas.

**Methods**

*Biochemical reagents.* Most of the enzymes used for cloning and sequencing work were obtained from Pharmacia. A sequencing kit containing T7 DNA polymerase (Sequenase) was obtained from the United States Biochemical Corporation. Radioisotopes deoxyadenosine 5'-α-[³²P]thiotriphosphate [α-³²P]dATP; (400 Ci mmol⁻¹, 148 TBq mmol⁻¹) and [α-³²P]dATP (300 Ci mmol⁻¹, 111 TBq mmol⁻¹) were purchased from Amersham, and were used for dideoxynucleotide sequencing and nick-translation, respectively. Hybond-N hybridization membranes were also obtained from Amersham.

*Bacterial strains, plasmids and DNA isolation.* DNA, isolated by the method of Marmur (1961), from D. vulgaris Miyazaki F (IAM 12604, Institute for Applied Microbiology, University of Tokyo, Japan) was kindly donated by Professor T. Yagi, Department of Chemistry, Shizuoka University, Oya, Shizuoka, Japan. DNA from D. vulgaris subsp. vulgaris Hildenborough (NCIB 8303) was similarly isolated in our own laboratory. Vectors pUC8 and M13mp8 or M13mp9 (Vieira & Messing, 1982; Messing & Vieira, 1982) were used for gene cloning and dideoxy sequencing, respectively. Plasmid DV1, containing the structural genes of the [NiFe] hydrogenase from D. vulgaris Hildenborough on a 4-8 kb Clal-Smal DNA insert in BLUESCRIPT was kindly donated by Dr A. E. Przybyla, Department of Biochemistry, School of Chemical Sciences, University of Georgia, Athens, GA, USA. A restriction map of this insert has been published (Menon et al., 1990) and is presented in Fig. 1 (a).

pHV150, which contains the hydA, B genes encoding [Fe] hydrogenase of D. vulgaris Hildenborough on a 1-9 kb BamHI-BamHI insert has been described elsewhere (Voordouw et al., 1987). A 0-7 kb hydC probe, containing the 3' end of the hydC gene (Stokkermans et al., 1989), was isolated from pHVCS (Voordouw et al., 1989b).
Cloning of Miyazaki hydrogenase genes

**Fig. 2. Southern blotting of DNA from D. vulgaris Hildenborough and Miyazaki.** Blots were incubated with: (a) a 1-9 kb probe containing the hydA,B genes for [Fe] hydrogenase from D. vulgaris Hildenborough or (b) probe P (Fig. 1a). Hildenborough DNA was digested with HindIII (lane 1), EcoRI (lane 2) or PstI (lane 3). Miyazaki DNA was digested with HindIII (lane 4), EcoRI (lane 5) or PstI (lane 6). The sizes (kb) of selected fragments derived from the migration of molecular mass markers and the nucleic acid sequence (Fig. 3) are indicated.

**Southern blotting, cloning and dideoxy sequencing.** Southern blotting was done with the formamide procedure for Hybond-N membranes as described previously (Voordouw et al., 1989b). Probes P1 and P2 (Fig. 1a) appeared to hybridize with 3-8 kb and 3-1 kb EcoRI fragments of D. vulgaris Miyazaki DNA, respectively (see Fig. 1b and Results). These hybridizing fragments were cloned in EcoRI-cleaved pUC8 and pHCal2 and pMZSS1-56, containing the 3.8 and 3.1 kb EcoRI inserts, were isolated using techniques described previously (Voordouw et al., 1989b). The inserts of these plasmids were sequenced with the dideoxy chain termination method of Sanger et al. (1977), using the procedures for generating M13 clones with random inserts obtained by sonication as outlined by Bankier & Barrell (1983). E. coli TG2 (from Toby Gibson, Laboratory of Molecular Biology, Cambridge, UK) was used as the host for these molecular biological manipulations. The sequence data obtained were compiled and analysed in a DEC-VAX computer using the programs of Staden (Staden, 1982, 1984a, b; Staden & McLachlan, 1982) and of the University of Wisconsin Genetics Computer Group (Deverewux et al., 1984).

**Results**

Using Southern blotting it can readily be shown that D. vulgaris Miyazaki F lacks the hydA,B genes, encoding a periplasmic [Fe] hydrogenase as found in the Hildenborough strain. Digests of chromosomal DNA from the two strains with restriction enzymes HindIII, EcoRI and PstI were electrophoresed through agarose and Southern blotted onto Hybond-N. Following hybridization of the blots with the hydA,B probe the results shown in Fig. 2(a) were obtained. In agreement with earlier data (Voordouw et al., 1985), hybridizing fragments of 14 kb (HindIII), 4-8 kb (EcoRI) and 4-8 and 6-6 kb (PstI) were observed for Hildenborough DNA (Fig. 2a, lanes 1–3), while Miyazaki DNA was entirely negative (Fig. 2a, lanes 4–6). A second gene in D. vulgaris Hildenborough, which could encode a [Fe] hydrogenase, was recently described by Stokkermans et al. (1989). This gene, hydC, is located immediately downstream from the Hildenborough hydA,B genes and has 30–50% sequence homology with these genes. Hybridization of Miyazaki DNA with a D. vulgaris Hildenborough hydC gene probe was tested but was also found to be negative (not shown).

Although D. vulgaris Miyazaki appears to lack genes for [Fe] hydrogenase with homology to those of the Hildenborough strain, its DNA does hybridize with probes derived from the [NiFe] hydrogenase genes of D. vulgaris Hildenborough. Probe P1 hybridizes with a 3-8 kb EcoRI fragment and PstI fragments of 1-3 and 0-58 kb (Fig. 2b, lanes 5 and 6) in agreement with the restriction map for this region of the Miyazaki chromosome, that was established by these experiments and nucleic acid sequencing (Fig. 1b). One can derive from the distribution of EcoRI sites in Fig. 1(b), that the 3-8 kb
EcoRI fragment arises from partial digestion: a 1 kb fragment of Miyazaki DNA is expected to hybridize with P1 upon complete digestion. This 1 kb fragment is present in EcoRI digests of Miyazaki DNA at low intensity (Fig. 2b, lane 5). Complete digestion of the D. vulgaris Miyazaki chromosome with EcoRI proved hard to achieve. In contrast complete PstI digests, showing more complex EcoRI band pattern is due to partial restriction fragments expected on the basis of the map (Fig. 2b), were achieved consistently, indicating that the more complex EcoRI band pattern is due to partial digestion rather than to the presence of multiple genes.

The 3-8 kb EcoRI fragment was next cloned by ligating a size-fractionated EcoRI digest (3-5 kb) of Miyazaki DNA to EcoRI-cut pUC8 and transforming the ligation mixtures into E. coli TG2. Recombinant clones were probed with radiolabelled P2, pMZSS1-56 containing the 3-1 kb EcoRI insert as indicated in Fig. 1(b) was isolated.

Both the 3-1 and 3-8 kb EcoRI fragments were sequenced with the shotgun procedure outlined by Bankier & Barrell (1983). This led to complete determination of the sequence of the 3-8 kb fragment, while that of the 3-1 kb fragment was partially determined. Together the two sequences form a contig of 6190 nucleotides (nt), as indicated in Fig. 1(c). The sequence of 3100 nt of this contig is shown in Fig. 3, with the proposed reading frames for the small and large subunits of Miyazaki hydrogenase indicated by translation into protein. Every nt of the sequence shown in Fig. 3 was determined at least once on each strand. Several severe compressions could only be read by using T7 DNA polymerase ('Sequenase') rather than Klenow polymerase and by substituting dGTP by dITP in the sequencing reactions.

The coding regions for the hydrogenase operon from D. vulgaris Miyazaki F. The coding regions for the small (nt 219-1169) and large (nt 1231-2931) subunits have been translated with the one letter amino acid code. Both genes are preceded by a purine-rich (−), ribosome-binding site (rbs). Potential promoter sequences (prl and pr2), discussed in the text, are indicated, as well as a hairpin loop (−−), that could function in transcription termination.

Fig. 3. Nucleotide sequence for the hydrogenase operon from D. vulgaris Miyazaki F. The coding regions for the small (nt 219-1169) and large (nt 1231-2931) subunits have been translated with the one letter amino acid code. Both genes are preceded by a purine-rich ribosome-binding site (rbs). Potential promoter sequences (pr1 and pr2), discussed in the text, are indicated, as well as a hairpin loop (−−), that could function in transcription termination.

Cloning of Miyazaki hydrogenase genes

Gene (Fig. 1 b). When Southern blots of EcoRI-digested D. vulgaris Miyazaki DNA were incubated with radiolabelled P3, two hybridizing fragments of 3-8 and 3-1 kb were observed, while just the 3-1 kb fragment was found to hybridize with probe P2 (not shown). These results indicated that the 3-1 kb fragment contains the 3' end of the small subunit gene. It was cloned by ligating 2-4 kb EcoRI fragments and probing the resulting recombinant clones with P2. pMZSS1-56 containing the 3-1 kb EcoRI insert as indicated in Fig. 1(b) was isolated.

Both the 3-1 and 3-8 kb EcoRI fragments were sequenced with the shotgun procedure outlined by Bankier & Barrell (1983). This led to complete determination of the sequence of the 3-8 kb fragment, while that of the 3-1 kb fragment was partially determined. Together the two sequences form a contig of 6190 nucleotides (nt), as indicated in Fig. 1(c). The sequence of 3100 nt of this contig is shown in Fig. 3, with the proposed reading frames for the small and large subunits of Miyazaki hydrogenase indicated by translation into protein. Every nt of the sequence shown in Fig. 3 was determined at least once on each strand. Several severe compressions could only be read by using T7 DNA polymerase ('Sequenase') rather than Klenow polymerase and by substituting dGTP by dITP in the sequencing reactions.

As indicated in Fig. 3, the gene for the small subunit (nt 219-1169) precedes that for the large subunit (nt 1231-2931) as in other operons encoding [NiFe] or [NiFeSe] hydrogenases (Li et al., 1987; Menon et al., 1987, 1990; Voordouw et al., 1989; Leclerc et al., 1988;
Sayavedra-Soto et al., 1988; Uffen et al., 1990; Ford et al., 1990). Identical putative ribosome-binding sites (AAGGAG) are located at position -12 to -7 from the initiating ATG codon of both genes, which are separated by an intergenic region of 58 nt (Fig. 3). When the nt sequences upstream from the D. vulgaris Miyazaki and D. gigas hydrogenase operons are compared, no homology is observed except for two elements of 10 and 12 nt, CATT(T/A)TCAAT and TG(T/G)ACCT(T/G)CCCG. These are, respectively, at position -190 to -181 and -159 to -148 from the translational start of the Miyazaki small subunit gene (Fig. 3) and at position -196 to -187 and -162 to -151 from the translational start of the D. gigas small subunit gene (Voordouw et al., 1989). Although these elements do not resemble the two E. coli consensus promoter sequences (TTGACA and TATAAT), they are separated by 21 nt, a number similar to the 19 nt found between the two elements in the E. coli consensus promoter and are, as indicated above, located in conserved positions relative to the translational start of the small subunit gene in D. vulgaris Miyazaki and D. gigas. They could therefore be involved in transcription initiation. Since a strong hairpin sequence, which could serve as a transcription terminator, is located immediately downstream from the large subunit gene (nt 2987–3013), it follows that the two structural genes could be transcribed as a single mRNA as a single mRNA of approximately 3000 nt.

The coding regions for the small (nt 219–1169) and large (nt 1231–2931) subunit of Miyazaki hydrogenase are homologous to those for the small and large subunits of D. gigas [NiFe] hydrogenase. A comparison of the amino acid sequences, derived from the sequences of these genes, is shown in Fig. 4 and is discussed below. As expected from the Southern blotting results (Fig. 2a), there is no homology with the [Fe] hydrogenase genes of D. gigas hydrogenase sequences of D. gigas. (Fig. 3)

**Discussion**

The cloning and sequencing of [NiFe] hydrogenase genes from a variety of bacterial sources has allowed an extensive comparison of homologies in the small and large subunit polypeptide chains, which, in turn, has led to the definition of conserved amino acid, particularly cysteine, residues. Conserved cysteine side-chains are likely to co-ordinate iron–sulphur clusters and the active-site nickel in [NiFe] hydrogenases. There are three iron–sulphur clusters in the [NiFe] hydrogenase of D. gigas, which has probably been best characterized with spectroscopic methods of all enzymes in this class (Li et al., 1987; Moura et al., 1982). Two of these are of the Fe₃S₄ type, requiring four cysteine residues for coordination. The third is a Fe₅S₆ cluster, which is
presumably co-ordinated by three cysteines (Kissing er al., 1989). It thus appears that 11 cysteines are required for the co-ordination of iron-sulphur clusters in [NiFe] hydrogenases.

Comparison of primary structures of [NiFe] hydrogenase small and large subunits has shown that there are 14 strictly conserved cysteine residues. Perhaps surprisingly, it appears that 10 of these are present in the small, while only four are found in the large subunit. The small subunit appears, therefore, to be the main site for iron-sulphur cluster co-ordination in [NiFe] hydrogenases. Using the numbering of the D. gigas sequence in Fig. 4, the 10 strictly conserved cysteines of the small subunit are C-67, C-70, C-162, C-198, C-238, C-263, C-269, C-278, C-296 and C-299, while in the large subunit two groups of two residues are conserved at the NH$_2$-terminus, C-65 and C-68, and at the COOH terminus, C-530 and C-533. Cysteine residue C-530 serves a special function, since its TGC codon is replaced by a TGA codon, thought to encode selenocysteine, in the [NiFeSe] hydrogenase gene of D. baculatus (Voordouw et al., 1989a). Since it has been shown by spectroscopic studies that this selenocysteine co-ordinates to the nickel (Eidness et al., 1989) it is now thought that C-530 is also a nickel ligand in [NiFe] hydrogenase.

Inspection of Fig. 4 reveals that all 14 cysteine residues listed above have been conserved in the sequence of the Miyazaki hydrogenase. In particular the conservation of C-530 (C-546 in the Miyazaki sequence, Fig. 4b) indicates this hydrogenase to belong to the [NiFe] hydrogenase family.

Apart from cysteine residues the two small subunits show 69% overall sequence identity and 81% sequence homology, as calculated with the GAP program (Devereux et al., 1984). Amino acids 1–50 of the D. gigas small subunit (Fig. 4a) constitute a signal sequence for protein export to the periplasm and the NH$_2$-terminus of the mature small subunit starts with L-51 (Li et al., 1987, Niviere et al., 1988). The sequence of residues 1–50 of the Miyazaki small subunit is highly conserved. In particular a box of six amino acid residues (RRXFXK), which is present in the small subunit signal sequence of every hydrogenase determined to date (Voordouw et al., 1989b) is also found in the Miyazaki sequence, while residues 50–51 comprise a signal peptidase processing site identical to that found in the D. gigas small subunit. Residues 1–50 of the Miyazaki small subunit are therefore also expected to function in the export of this hydrogenase to the periplasm. The amino acid sequence of the mature small subunit is 267 residues long and has a calculated molecular mass of 28763 Da.

The D. vulgaris Miyazaki and D. gigas large subunits show 68% sequence identity and 80% sequence homology (Fig. 4b). The two sequences are highly homologous at the COOH terminus and end with the same sequence.

### Table 1. Amino acid composition of D. vulgaris Miyazaki hydrogenase as determined from the sequence of the gene (n$_D$) and from the data (n$_Y$) of Yagi et al. (1976), recalculated for a total of 833 residues

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>n$_D$</th>
<th>n$_Y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>48</td>
<td>45</td>
</tr>
<tr>
<td>H</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>R</td>
<td>34</td>
<td>27</td>
</tr>
<tr>
<td>D + N</td>
<td>77</td>
<td>78</td>
</tr>
<tr>
<td>T</td>
<td>52</td>
<td>50</td>
</tr>
<tr>
<td>S</td>
<td>31</td>
<td>38</td>
</tr>
<tr>
<td>E + Q</td>
<td>70</td>
<td>71</td>
</tr>
<tr>
<td>P</td>
<td>56</td>
<td>59</td>
</tr>
<tr>
<td>G</td>
<td>67</td>
<td>73</td>
</tr>
<tr>
<td>A</td>
<td>89</td>
<td>88</td>
</tr>
<tr>
<td>C</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>V</td>
<td>67</td>
<td>64</td>
</tr>
<tr>
<td>M</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>I</td>
<td>43</td>
<td>38</td>
</tr>
<tr>
<td>L</td>
<td>58</td>
<td>56</td>
</tr>
<tr>
<td>Y</td>
<td>32</td>
<td>28</td>
</tr>
<tr>
<td>F</td>
<td>35</td>
<td>34</td>
</tr>
<tr>
<td>W</td>
<td>11</td>
<td>28</td>
</tr>
<tr>
<td>Total</td>
<td>833</td>
<td>833</td>
</tr>
</tbody>
</table>

The Miyazaki large subunit thus also lacks the COOH terminal extension found in the membrane-bound [NiFe] hydrogenases of E. coli, B. japonicum and R. capsulatus (see Introduction). Apart from an NH$_2$-terminal extension of 16 residues in the Miyazaki large subunit, there are no major differences with the D. gigas sequence. The calculated large subunit molecular mass is 62495 Da (566 residues, excluding the initiating methionine).

The Miyazaki hydrogenase genes, that were cloned and sequenced in the present study, are thus highly homologous to those encoding the periplasmic [NiFe] hydrogenase of D. gigas. We consider it very likely that these genes encode the periplasmic (Tamura et al., 1988), membrane-bound enzyme purified and characterized by Yagi et al. (1976, 1985). The subunit molecular masses derived from the gene sequence are very similar to those determined by Yagi et al. (1985: 29 and 60 kDa) and the amino acid composition determined for purified hydrogenase (Yagi et al., 1976) is in good agreement with that derived from the sequence of the gene (Table 1). It appears, therefore, that D. vulgaris Miyazaki harbours a [NiFe] hydrogenase and there are no obvious features in the amino acid sequence of the enzyme that can explain its unusual enzymic and spectroscopic properties (Yagi et al., 1976, 1985).

This research was supported by an operating grant from the National Science and Engineering Research Council of Canada (NSERC) to G.V. The authors are indebted to Drs T. Yagi and A. E. Prybyla for gifts of chromosomal and plasmid DNA, to Drs P. M. Vignais and A. E. Prybyla for sending preprints of manuscripts and to J. K. Voordouw for technical support.
References


