Microbiology (1996), **142**, 829–836 Printed in Great Britain

The [Ni–Fe] hydrogenase from the thermophilic bacterium *Acetomicrobium flavidum*

G. M. Mura, P. Pedroni, C. Pratesi, G. Galli, L. Serbolisca and G. Grandi†

Author for correspondence: G. Grandi. Tel: +39 577 243506. Fax: +39 577 243564.

ENIRICERCHE S.p.A. – Laboratories of Genetic Engineering and Microbiology, Via Maritano, 26 San Donato Milanese 20097 Milan, Italy

Biochemical analysis of the soluble hydrogenase from the thermophilic organism *Acetomicrobium flavidum* revealed that the enzyme is an $\alpha_2\beta_2$ tetramer, with the α and β subunits having a molecular mass of 50 kDa and 25 kDa, respectively. The most important biochemical properties of the enzyme are a specific activity of 77 μmol min^{-1} (mg protein)$^{-1}$, a K_m for methylviologen of 0·2 mM, a pH optimum of 7·5 and a T_{50} of about 70 °C. In addition, the enzyme is remarkably stable to oxygen inactivation, retaining full activity after 24 h exposure to air. By using oligodeoxynucleotides designed on the basis of the N-terminal sequences of the two subunits, the corresponding genes have been isolated and sequenced. When compared to the other hydrogenases so far characterized, the *A. flavidum* hydrogenase appears to be a typical [Ni–Fe] enzyme. The hydrogenase was expressed in *Escherichia coli* at high levels in a soluble form but it was not active. The analysis of the recombinant large subunit showed that it was not post-translationally processed at its C-terminus.

Keywords: hydrogenase, *Acetomicrobium flavidum*, Fe–S enzymes, heterologous expression of hydrogenases

INTRODUCTION

Many living organisms are endowed with specific metabolic pathways that allow them either to utilize or to evolve hydrogen. This latter property has stimulated an intense research activity during the last 20 years because hydrogen represents a clean alternative to the current energy sources. In particular, much attention has been focused on one of the key groups of enzymes of hydrogen metabolism, the hydrogenases (for recent reviews, see Przybyla *et al.*, 1992; Vignais & Toussaint, 1994).

A large number of hydrogenases from a variety of microorganisms have been isolated and characterized at both the biochemical and molecular level. Hydrogenases are broadly classified in two major categories on the basis of the presence or absence of Ni in their catalytic sites (Przybyla *et al.*, 1992). Indeed, the absence of Ni in hydrogenases has been demonstrated only in a few anaerobic bacteria and it is becoming evident that [Fe]-only hydrogenases are the exception to the [Ni–Fe] rule (Adams, 1990).

†**Present address:** Biocine S.p.A., Via Fiorentina, 1 53100 Siena, Italy.

The EMBL/GenBank/DDBJ accession number for the sequence data reported in this paper is X80180.

It should be pointed out, however, that the metal content of the hydrogenases as well as their mechanism of activity will probably be revised soon in the light of the recent indication that the [Ni–Fe] hydrogenase from *Desulfovibrio gigas* might contain in its active site two metal ions, one being Ni and the other, probably, Fe. This evidence comes from the crystallographic analysis of the *D. gigas* hydrogenase structure which has been recently resolved by Volbeda *et al.* (1995).

Besides their metal content, the hydrogenases also differ in the organization of their subunits, which can vary from multimeric to dimeric or monomeric.

In general, the [Ni–Fe] hydrogenases can both evolve and oxidize hydrogen, the latter being the preferred reaction, and they are relatively resistant to oxygen, which inactivates these enzymes in a reversible way. In contrast, the [Fe]-only hydrogenases are hydrogen-evolving enzymes and are irreversibly inactivated by oxygen.

In this respect, the hydrogenase from *Acetomicrobium flavidum*, the only known thermophilic bacterium which ferments 1 mol glucose to 2 mol acetate and carbon dioxide and 4 mol hydrogen (Soutschek *et al.*, 1984), is noteworthy. This hydrogenase was described by Schindler & Winter (1987) as a [Fe]-only, thermostable monomeric enzyme insensitive to oxygen with a pre-

ference for evolving rather than oxidizing hydrogen. The unusual properties of this hydrogenase prompted us to analyse the enzyme in more detail in an attempt to elucidate the structural features of its activity and stability.

Here we describe the biochemical characterization of the *A. flavidum* hydrogenase as well as the cloning and sequencing of the genes encoding it. From our work, it appears that the enzyme is a Ni-containing two-subunit enzyme which resembles the typical [Ni–Fe] hydrogenases from both thermophilic and mesophilic organisms. The enzyme is expressed at high levels and in a soluble form in *Escherichia coli* but is not active, probably because of the inability of *E. coli* to correctly process the C-terminus of the large subunit.

METHODS

Bacterial growth. *Acetomicrobium flavidum* (DSM 20664) was grown under nitrogen flow (0·3 l min^{-1}) at 58 °C in a 20 l Setric GI fermenter using the growth medium described by Schindler & Winter (1987). At the end of the exponential phase of growth (OD_{660} = 0·8 corresponding to 0·6–1 g wet wt l^{-1}), cells were harvested by centrifugation, washed once in 50 mM Tris/HCl (pH 7·5) and stored at −20 °C.

Hydrogenase activity and characterization. Hydrogenase activity was routinely determined by following hydrogen evolution at 58 °C as described previously by Pedroni *et al.* (1995) using a Varian 3400 Gas Chromatograph equipped with a 5 Å molecular sieve column (1 m, 3·175 mm). One unit of hydrogenase activity is defined as the amount of enzyme which catalyses the production of 1 μmol hydrogen min^{-1}.

Thermal stability was determined by following the residual activity of the enzyme after incubation at different temperatures. T_{50} is the temperature at which 50% of the activity is lost after a 30 min incubation.

For K_m determination, the hydrogenase activity was followed varying the methylviologen concentration from 10 to 2000 μM.

The colorimetric determination of the hydrogenase activity on 10% (w/v) native polyacrylamide gels was performed at 58 °C as already described by Pedroni *et al.* (1995). Proteins were determined spectrophotometrically with the Bio-Rad micromethod based on the Bradford assay (Bradford, 1976) using BSA as the standard.

SDS-PAGE and native PAGE were carried out according to the method of Laemmli (1970).

Hydrogenase purification

Native enzyme. For the purification of the *A. flavidum* hydrogenase, all buffers were degassed and flushed with helium and then supplemented with 5 mM DTT. Purification was preferentially carried out at 4 °C but we found that operating at room temperature did not cause a significant loss of hydrogenase activity. Frozen cells (1·8 g) were resuspended in 50 mM Tris/HCl (pH 7·5)/5 mM DTT and broken by three passages through a French Pressure Cell apparatus set at 18000 p.s.i (124·2 MPa). After removing cell debris by centrifugation at 10000 ***g*** for 30 min, the soluble fraction was obtained by ultracentrifugation at 130000 ***g*** for 90 min. The soluble extract was loaded onto a Sephacryl S-300 HR column (2·6 × 65 cm; Pharmacia) previously equilibrated with 50 mM Tris/HCl (pH 7·5)/5 mM DTT/100 mM NaCl. The column was washed with the same buffer at a flow rate of 0·5 ml min^{-1}. Fractions of 10 ml each were collected and those containing the hydrogenase activity were pooled and loaded onto a MonoQ HR 16/10 column (FPLC; Pharmacia) equilibrated with 50 mM Tris/HCl (pH 7·5)/100 mM NaCl/5 mM DTT. The column was washed with 50 ml of the same buffer and the adsorbed proteins were eluted with a 200 ml linear gradient of 0–1 M NaCl at a flow rate of 2·5 ml min^{-1}. Fractions (5 ml) were collected.

Sequencing grade hydrogenase was obtained by running the partially purified hydrogenase on a 10% native polyacrylamide gel. After activity staining using sodium dithionite and methylviologen, the activity bands were excised from the gel and electroeluted with a 422 Electro-eluter (Bio-Rad) using a Tris/glycine buffer in 0·1% SDS (Legendre & Matsudaira, 1989). The elution products were analysed by SDS-PAGE and blotted onto a PVDF membrane (Millipore). The N-terminal sequencing was performed with a Beckman LF Protein Sequencer, using the proteic material adsorbed on the PVDF membrane.

Periplasmic proteins. For the analysis of periplasmic proteins, the cells (1 g wet wt) were resuspended in 10 ml 100 mM Tris/HCl (pH 8·0)/0·5 mM EDTA/0·5 M sucrose. An equal volume of cold water containing 2 mg lysozyme ml^{-1} was added and incubated on ice for 30 min. Spheroplasts were finally separated from the supernatant by centrifugation at 5000 ***g*** for 20 min at 4 °C.

Recombinant enzyme. For the purification of the recombinant hydrogenase, *E. coli* cells were grown aerobically at 37 °C in LB medium (2 l) supplemented with 0·5% glucose, 10 μM $NiCl_2$, 10 μM $FeSO_4$ and 100 μg ampicillin ml^{-1}. When the culture reached an OD_{600} value of 0·5, the expression of hydrogenase was induced with 1 mM IPTG for 3 h and thereafter the cells (7·7 g) were collected by centrifugation and resuspended in 20 ml 50 mM Tris/HCl (pH 7·5), 1 mM PMSF, 1 mM benzamidinium chloride and 0·1 mg each of DNase and RNase ml^{-1}. The cells were broken by three passages through a French Press apparatus at 18000 p.s.i. (124·2 MPa) and the soluble fraction was obtained by centrifugation at 40000 ***g*** for 60 min. The soluble extract (1077 mg protein) was loaded onto a Q-Sepharose Fast Flow column (1·6 × 20 cm; Pharmacia) previously equilibrated with 50 mM Tris/HCl (pH 7·5)/5 mM DTT. The column was washed with 60 ml of the same buffer and the adsorbed proteins were eluted with a 500 ml linear gradient of 0–0·4 M NaCl. Fractions (10 ml) were collected at a flow rate of 0·5 ml min^{-1}. The hydrogenase-containing fractions were pooled (110 mg protein) and loaded onto a Chelating Sepharose 6B Fast Flow column (1 × 16 cm; Pharmacia) previously activated with $NiCl_2$ and equilibrated with 50 mM Tris/HCl (pH 7·5)/5 mM DTT. The column was washed with 30 ml 50 mM Tris/HCl (pH 7·5)/5 mM DTT/0·2 M NaCl and the adsorbed proteins were eluted with 100 ml of a linear gradient of 0–0·2 M imidazole. Fractions (2 ml) were collected at a flow rate of 0·4 ml min^{-1}. The fractions containing the hydrogenase large subunit were collected (46 mg protein) and loaded onto a Sephacryl S-200 HR column (5 × 88 cm; Pharmacia) previously equilibrated in 50 mM Tris/HCl (pH 7·5)/5 mM DTT/0·2 M NaCl and proteins were eluted at a flow rate of 1·3 ml min^{-1}. The fractions containing the small subunit (27 mg protein) were concentrated and loaded onto the same Sephacryl S-200 column and proteins were eluted as described above.

Preparation of antibody to the N-terminus of the hydrogenase large subunit. The peptide having the sequence TEVFKLEINPVTRIC and corresponding to the N-terminal sequence of the 50 kDa large subunit of the *A. flavidum* hydrogenase from amino acids 1–14 was synthesized using the fluoren-9-ylmethoxycarbonyl (Fmoc)-polyamide chemistry

(Atherton & Sheppard, 1989) with an LKB Biolynx 4170 automatic peptide synthesizer. Fmoc-pentafluorophenyl and 3,4-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazin-3-yl esters (Novobiochem, Switzerland) were used throughout, except for the first coupling reaction, which was carried out by the highly reactive Fmoc-amino acid symmetrical anhydride with the addition of 0·1 equivalents 4-dimethylaminopyridine as catalyst for the esterification reaction. The C-terminal cysteine was added to facilitate peptide crosslinking to the carrier protein (see below).

Arginine and cysteine residues were protected with 4-methoxy-2,3,6-trimethylbenzenesulfonyl and acetamidomethyl (Acm) groups, respectively. The S-protected peptide was obtained by treating the resin with a trifluoroacetic acid/phenol mixture (95:5, v/v) for 6 h at room temperature. The crude peptide (160 mg) was loaded onto an anion-exchange column (Q-Sepharose Fast Flow, 1·6 × 16 cm) and eluted with a linear gradient of 20–500 mM NH_4HCO_3 (200 ml, pH 7·7). Fractions were collected and freeze-dried. Edman sequencing and analytical reverse-phase HPLC confirmed the identity and purity of the synthesized peptide. The Acm group was cleaved by treating the peptide with mercuric acetate (Veber *et al.*, 1972). Mercuric ions were precipitated by adding a 15-fold excess of DTT. After preparative reverse-phase HPLC to remove DTT, the thiol-containing peptide was freeze-dried and resuspended in 50 mM potassium phosphate buffer (pH 7·5). The antigen was prepared by crosslinking the thiol peptide to purified Keyhole Limpet haemocyanin activated with thiol-reactive maleimide groups (Pierce).

Antigen equivalent to 300 μg peptide was injected subcutaneously into two New Zealand White rabbits together with 300 μl complete Freund's adjuvant (Sigma). Immunization was repeated at 2–3 week intervals using the same amount of antigen supplemented with incomplete Freund's adjuvant. After 2 months the rabbits were bled and the serum stored at −20 °C.

Recombinant DNA techniques. The molecular biology methods were according to standard protocols (Sambrook *et al.*, 1989). When necessary, the subcloning of the PCR fragments was done using the pUC18 plasmid from Boehringer Mannheim or the pCRII TA cloning kit from Invitrogen. The oligodeoxynucleotides used for sequencing, PCR and screening of the genomic library were synthesized with a Beckman DNA synthesizer (OLIGO 1000). The following primers were used (from 5′ to 3′; deg = degenerate): S1deg, GC(TCAG)-AA(AG)GT(TCAG)GC(TCAG)AC(TCAG)TT(CT)TGG; S2deg, TC (TCAG) A (GA) (GA) AA (TCAG) (CG) (TA) CAT(GA)TG(GA)CA; L1deg, GA(AG)GT(TCAG)TT(CT)-AA(GA)(CT)T(TCAG)GA(AG)AT; L2deg, TC(AG)TC-(TCAG)A(GA)CAT(TCAG)AC(TCAG)GT; p1, TGGTTA-GAAGCTTGTGCGGGATG; p2, ATGTCCTTCTATCCT-AGTTACAGGGTTTAT.

Cloning of *A. flavidum* hydrogenase genes. The isolation of the hydrogenase genes was carried out following the PCR strategy already described by Pedroni *et al.* (1995). PCR (30 cycles of denaturation at 94 °C for 1 min, annealing at the appropriate temperature for 1 min and elongation at 72 °C for 2 min) was performed using a Perkin Elmer Cetus thermal cycler in 100 μl mixtures covered with 100 μl mineral oil. Reaction samples included template genomic DNA (500 ng), 200 μM dNTPs, 150 pmol degenerate primers or 100 pmol gene-specific primers and 2·5 units of either *Taq* or cloned *Pyrococcus furiosus* DNA polymerase.

The chromosomal region encoding the two hydrogenase subunits was isolated by screening a genomic library constructed after digestion of the *A. flavidum* chromosomal DNA (100 μg) with *Xba*I. Fragments in the 3·6–3·8 kb range were purified by sucrose gradient centrifugation (5–20 %, w/v) and ligated into *Xba*I-digested pUC18. The ligase mixture was used to transform *E. coli* TG1 competent cells (Dower *et al.*, 1988) and about 2000 clones were screened by colony hybridization using the DIG System (Boehringer Mannheim) and the p1 and p2 oligonucleotides as probes.

Nucleotide sequencing and analysis. DNA sequencing was performed on both strands by the dideoxy chain-termination method (Sanger *et al.*, 1977) using the CircumVent (New England Biolabs) sequencing kit. The sequence was compiled and analysed by the PC/GENE software package (IntelliGenetics). Multiple amino acid sequence alignments were performed using the MACAW program (Schuler *et al.*, 1991).

RESULTS AND DISCUSSION

Purification and characterization of the *A. flavidum* hydrogenase

For the isolation of the *A. flavidum* hydrogenase we followed the same strategy successfully applied for the cloning of the *P. furiosus* hydrogenase genes (Pedroni *et al.*, 1995). Briefly, the approach is based on the use of degenerate oligonucleotides to amplify the target genes from the chromosomal DNA by PCR and inverse PCR. The oligonucleotides are designed on the basis of the N-terminal sequence of the protein whose gene is to be cloned.

In order to purify the *A. flavidum* hydrogenase and to determine its N-terminal sequence, we took advantage of the information available from the previously published work of Schindler & Winter (1987), who isolated and characterized this enzyme for the first time. Since in the course of our purification we found a few differences in the chromatographic behaviour of the enzyme and in its physical and chemical properties, a brief description of the hydrogenase isolation and characterization is given.

Upon cell disruption, approximately 98 % of the hydrogenase activity was found in the soluble protein fraction. When soluble proteins were run on a 10 % polyacrylamide gel under native conditions and the gel was stained for hydrogenase activity, two major activity bands having R_F values of 0·01 and 0·31 were detected (data not shown). The hydrogenase was partially purified following the procedure summarized in Table 1. The SDS-PAGE analysis of the purified enzyme indicated that it consists of two subunits having molecular masses of 50 and 25 kDa. However, on gel-filtration chromatography most of the hydrogenase activity eluted with an apparent molecular mass of approximately 150 kDa and the native gel analysis of the hydrogenase-containing fractions from both gel-filtration chromatography and ion-exchange chromatography indicated that the slow and fast migrating activity bands co-eluted. These data suggested that the *A. flavidum* enzyme is organized in an $\alpha_2\beta_2$ structure with the tendency to aggregate in more complex forms.

To confirm that the two hydrogenase activities are different aggregation forms of the same enzyme, an aliquot of the purified hydrogenase was loaded onto an acrylamide

Table 1. Purification steps of *A. flavidum* hydrogenase

Fraction	Volume (ml)	Protein (mg)	Enzyme (U)	Specific activity (U mg^{-1})	Purification (-fold)	Recovery (%)
Soluble extract	7·5	168·0	511	3	1	100
Sepharose S-300	30·0	23·0	474	21	7	93
Mono-Q	15·0	4·5	346	77	26	68

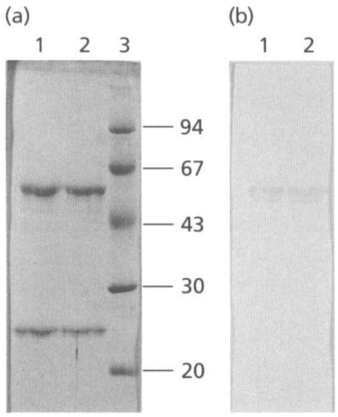

Fig. 1. Coomassie-Blue-stained 12·5% SDS-polyacrylamide gel (a) and Western blot analysis (b) of the slow and fast migrating hydrogenase activities electroeluted from a 10% native polyacrylamide gel. Lanes: 1, fast migrating activity; 2, slow migrating activity; 3, molecular mass standards (94, 67, 43, 30 and 20 kDa).

gel and electrophoretically separated under native conditions. After identification of the two hydrogenase activities by methylviologen gel staining, the two activity bands were electroeluted. The electrophoretic analysis on an SDS-PAGE gel of both samples is shown in Fig. 1(a). The fast migrating activity was resolved in two protein species having apparent molecular masses of 50 kDa and 25 kDa. These two protein species also appeared to be the components of the slow migrating hydrogenase activity. The identity between the protein species was confirmed by both Western blotting, using antibodies prepared against a 15 amino acid peptide having a sequence identical to the N-terminus of the large subunit (see below) (Fig. 1b), and by N-terminal sequence analysis (data not shown).

The biochemical properties of the partially purified enzyme, including pH and temperature optima, specific activity and K_m for methylviologen were also calculated. In general, the values are in agreement with the data previously published by Schindler & Winter (1987), the pH and temperature optima being 7·5 and 72 °C, respectively, and the K_m being 0·2 mM (data not shown). Besides the structural organization of the enzyme in two, rather than one, subunits, the most important difference between our data and the data previously published was found in the specific activity of the enzyme, which we found, using a partially purified enzyme, to be approximately 77 μmol min^{-1} (mg protein)$^{-1}$ (Table 1) against the 2 μmol min^{-1} (mg protein)$^{-1}$ value of the enzyme purified by Schindler & Winter (1987). In addition to that, we found that the enzyme represents less than 0·1% of the total soluble protein, much less than the figure reported by Schindler & Winter (1987) (4% of the total soluble protein).

Cloning and sequence analysis of the *A. flavidum* hydrogenase (*hyd*) genes

On the basis of the N-terminal sequencing of the small and the large subunits, degenerate forward and reverse primers (S1deg/S2deg and L1deg/L2deg) were used to amplify the 5′ regions of the coding genes. The deduced amino acid sequences of the PCR products perfectly matched the N-terminal amino acid sequences of the purified small and large subunits (data not shown). From the nucleotide sequences of the PCR products, the p1 and p2 primers were designed and used in a second PCR reaction, which gave a fragment of about 600 bp. From the analysis of this fragment, the entire sequence of the small subunit gene was obtained, together with the sequence of the 5′ region of the large subunit. Finally, the p1 and p2 oligonucleotides were used as probes to screen an *A. flavidum* genomic library. One clone harbouring the hybrid plasmid pSM696 and hybridizing to both probes was isolated. The plasmid contained a 3·7 kb *Xba*I chromosomal fragment which was completely sequenced. The restriction map and the physical organization of this fragment carrying the *A. flavidum hyd* genes is shown in Fig. 2, and its nucleotide sequence with the derived amino acid sequence is given in Fig. 3.

The genes corresponding to the small (*hydS*) and the large (*hydL*) subunits encode polypeptides of 179 and 475 amino acids, respectively. The calculated molecular masses of the two subunits (20·4 kDa and 54·1 kDa, respectively) are in agreement with those estimated by SDS-PAGE analysis.

The comparison of the N-terminal protein sequences with the corresponding deduced amino acid sequences indicates that both subunits are post-translationally processed at their N-termini, lacking the initial methionine. No typical signal-peptide-like encoding sequence is found

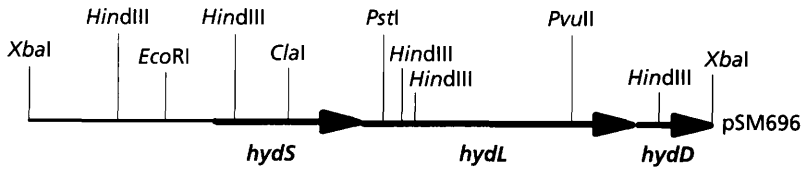

Fig. 2. Restriction map and physical organization of the 3·7 kb *Xba*I genomic fragment carrying the *A. flavidum hydSLD* genes. The thick arrows indicate the coding regions and the orientation of the three genes.

```
GAATTCGCGAAGAAGTTGGGAGTTAAGTTAGAGAGGCCTGAGGTATCTCCTTGCGAGGAAAT      62
CTTTAAATCGCAACTACGGGCCAGCGAAGCGCGTATTGTTACTTTGTAAACAGTCTCACTGG     124
                                                              *
                    → hydS
GAGGTGAAATGTTCCATGGCCAAGGCAAAGGTGGCTACATTTTGGTTAGAAGCTTGTGCGGG     186
****          (M) A  K  A  K  V  A  T  F  W  L  E  A  C  A  G
ATGTCATATGTCCTTTCTTGATTTAGATGAAAGGCTGATAGATTTGTTTCAAAATGTGGAAA     248
  C  H  M  S  F  L  D  L  D  E  R  L  I  D  L  F  Q  N  V  E
TACTCTTTTCACCCATCGTAGATGCAAAGGATATTCCGAATATAGATGTAGGAGTACTGTCC     310
  I  L  F  S  P  I  V  D  A  K  D  I  P  N  I  D  V  G  V  L  S
GGCGGGCTGGGCAATGTGGAAGAGGTAGAGCTGGCCAAAAAGATGAGGGAACGTTGCAAGTA     372
  G  G  L  G  N  V  E  E  V  E  L  A  K  K  M  R  E  R  C  K  Y
TTTGGTAGCTTGGGGTGACTGCGCAGTGTTTGGCGGTATTAACTGCATGCGCAACTTCATAC     434
  L  V  A  W  G  D  C  A  V  F  G  G  I  N  C  M  R  N  F  I
CCAAGGATGTCGTCTTAAGAGAGGGATACATTGAAACTGCAAGCACAGTTAATCCTCAAGGC     496
  P  K  D  V  V  L  R  E  G  Y  I  E  T  A  S  T  V  N  P  Q  G
ATTGTTCCAAGCGAAGACATTCCAGAATTGCTTCCGAGGGCATTGCCGATCGATTACGAAGT     558
  I  V  P  S  E  D  I  P  E  L  L  P  R  A  L  P  I  D  Y  E  V
GAAAGTCGATGTCTATGTGCCTGGATGTCCCCCTGATGCTGATACGATTTACTACGTTTTCA     620
  K  V  D  V  Y  V  P  G  C  P  P  D  A  D  T  I  Y  Y  V  F
AGGAGTTATTGGCTGGACGAGTTCCTAAGGTTCCCTCAGAAATGATGCGCTATGACTAGAAT     682
  K  E  L  L  A  G  R  V  P  K  V  P  S  E  M  M  R  Y  D  -
                    → hydL
AAGGGGAGGAAATGGACATGACCGAGGTCTTTAAGCTGGAGATAAACCCTGTAACTAGGATA     744
*********     (M) T  E  V  F  K  L  E  I  N  P  V  T  R  I
GAAGGACATGGAAAAATAACGGTTATGTTGGATGAGTCCGGGCATGTACGCGAGACGCGCTT     806
  E  G  H  G  K  I  T  V  M  L  D  E  S  G  H  V  R  E  T  R  F
TCACGTGACTCAATACCGTGGCTTTGAAGTCTTTACGCATGGAAGAGATTTCAGGGAGATGC     868
  H  V  T  Q  Y  R  G  F  E  V  F  T  H  G  R  D  F  R  E  M
CCGTAATCACCCCTAGGATTTGTGGGATTTGCCCTGTGTCTCACCACCTTGCCTCAGCAAAG     930
P  V  I  T  P  R  I  C  G  I  C  P  V  S  H  H  L  A  S  A  K
GCGTGCGACGAGATTCTGGGGGTTACGATAACACCTGCAGCCCACAAGCTCCGCGAATTGAT     992
  A  C  D  E  I  L  G  V  T  I  T  P  A  A  H  K  L  R  E  L  M
GCATATGGGACAAATTGTACAATCTCACGCCCTAAGCTTCTTTCATCTGTCCAGCCCGGATA    1054
  H  M  G  Q  I  V  Q  S  H  A  L  S  F  F  H  L  S  S  P  D
TTCTATTGGGTTTCGATGCGCCCGTTAAGATAAGAAATGTTGCTGGGTTGGTAGATCGTTAC    1116
  I  L  L  G  F  D  A  P  V  K  I  R  N  V  A  G  L  V  D  R  Y
CCTGAGCTTGCAAAGAAGGGCATCATGCTCCGCAAGTTTGGGCAGGAAATAATAAAAACCCT    1178
  P  E  L  A  K  K  G  I  M  L  R  K  F  G  Q  E  I  I  K  T  L
GGGTGGCAAGAAGATACATCCTTGGCATAGCATTCCAGGAGGGGTCAACAGAAGCTTGACTC    1240
  G  G  K  K  I  H  P  W  H  S  I  P  G  G  V  N  R  S  L  T
CCCAAGAGCGGGATGCCATAGCTGCTCAGCTTCCCGAAATGAAATCCATCGCTATGGAAGCC    1302
  P  Q  E  R  D  A  I  A  A  Q  L  P  E  M  K  S  I  A  M  E  A
ATCAAGTTGATCAAGGATTATCTGCAAGAAGGTGGCGAAGAATTAAAGGAATTTCCTACGCT    1364
  I  K  L  I  K  D  Y  L  Q  E  G  G  E  E  L  K  E  F  P  T  L
AGATACGGCCTATATGGGGTTAGTCCGTGACGGATACCTGGAACTTTATGACGGTGAAGTAC    1426
  D  T  A  Y  M  G  L  V  R  D  G  Y  L  E  L  Y  D  G  E  V
GCATTAAGGCTCCAAGGGGCAGGATTTTGGATCAATTTGATCCGAAGGATTATCTGGACCAT    1488
  R  I  K  A  P  R  G  R  I  L  D  Q  F  D  P  K  D  Y  L  D  H
ATCGGCGAACATGTAGAACCTTGGAGCTATTTGAAGTTTCCCTTCTATAAGGCTCTGGGCTT    1550
  I  G  E  H  V  E  P  W  S  Y  L  K  F  P  F  Y  K  A  L  G  F
CCCTCATGGGTCTTACAGGGTTGGGCCGCTTGCCCGGTTGAACGCGGCTGATGCCGTGAGCA    1612
  P  H  G  S  Y  R  V  G  P  L  A  R  L  N  A  A  D  A  V  S
CTCCGGAAGCGTCAAAGGAGTTCGCCCTTTACAAGGAGATGGGTGAGGACGGCATAGTGCCT    1672
  T  P  E  A  S  K  E  F  A  L  Y  K  E  M  G  E  D  G  I  V  P
TATACCTTGTATTATCACTATGCTAGATTAATAGAAGCACTTTACGGTTTGGAGCGCATAGA    1736
  Y  T  L  Y  Y  H  Y  A  R  L  I  E  A  L  Y  G  L  E  R  I  E
GCAGCTGCTGGCCGATCCTGATATCACCTCAAGCGACCTTCGCGTTACATCTAAAGAAATAA    1798
  Q  L  L  A  D  P  D  I  T  S  S  D  L  R  V  T  S  K  E  I
ATCCCGAAGGCATCGGCGTGATAGAAGCTCCAAGGGGAACATTGATTCACCATTATCAGGTC    1860
  N  P  E  G  I  G  V  I  E  A  P  R  G  T  L  I  H  H  Y  Q  V
AACGAAAGCGGCGTCATTACAAAGGTGAACCTCATTGTAGCTACCGGTCATAATAATTTTGC    1922
  N  E  S  G  V  I  T  K  V  N  L  I  V  A  T  G  H  N  N  F  A
CATGAATAAAGGCGTAGAGATGGTCGCAAAGAAATACATTACAGGCACAAATGTCCCGGAGG    1984
  M  N  K  G  V  E  M  V  A  K  K  Y  I  T  G  T  N  V  P  E
GCGTTTTTAACAGACTGGAGCACGTCATTCGTGCTTATGATCCCTGTTTATCGTGTTCTACC    2046
  G  V  F  N  R  L  E  H  V  I  R  A  Y  D  P  C  L  S  C  S  T
CATGCCGTAGGGAAGATGCCTTTGAAATTGGAGTTAGTCGGACCCACGGGCGAAATATTAAA    2108
  H  A  V  G  K  M  P  L  K  L  E  L  V  G  P  T  G  E  I  L  K
                *****     → hydD
GGAAGTCACGAGGGACTAAAGCCTAATGAACGTGATCGTGTTGGGTTACGGGAATCCCTTTA    2170
  E  V  T  R  D  -        M  N  V  I  V  L  G  Y  G  N  P  F
GAACGGACGACGGGGTAGGGCCTGTCGGAGCAGAAGCGATAGTACAGTGGCTTGAAGATGAA    2232
  R  T  D  D  G  V  G  P  V  G  A  E  A  I  V  Q  W  L  E  D  E
GGACATGAAGTGACGACCTGGATAGGCCAACAACTTCTCCCAGAAGTCGTGTTGGAGCTGGA    2294
  G  H  E  V  T  T  W  I  G  Q  Q  L  L  P  E  V  V  L  E  L  E
GGGAAAAGATCTGGCCATCTTCGTCGATGCTTCTGTGGAACAATATCCCGAAAGCTTTCGTA    2356
  G  K  D  L  A  I  F  V  D  A  S  V  E  Q  Y  P  E  S  F  R
TATGCGAAGTCCTGCCCTCTGGCGATCTAGA                                    2387
I  C  E  V  L  P  S  G  D  L
```

Fig. 3. Nucleotide sequence of the 2387 bp *Eco*RI–*Xba*I fragment encoding HydS, HydL and part of HydD. The amino acid sequence of each ORF is given in one-letter code under the DNA sequence. The putative ribosome-binding sites (asterisks), the start codons of the three polypeptides (short heavy arrows) and the amino acid residues determined by N-terminal sequencing (underlined) are indicated. The initiator methionines of HydS and HydL are in parentheses since they are absent in the protein sequences.

at the 5′ end of the small subunit gene, in agreement with our experimental evidence, indicating that this hydrogenase is a cytoplasmic enzyme.

Downstream from the gene encoding the hydrogenase large subunit, the 5′ region of a third ORF, designated *hydD* because of its sequence homology with other hydrogenase-linked ORFs (see below), is present.

The *hydSLD* genes are closely linked: *hydS* and *hydL* are separated by 20 nucleotides and 6 nucleotides separate *hydL* from *hydD*. Each start codon is preceded by a putative ribosome-binding site: the sequence GGGAGG is located 11 bp upstream from *hydS*, AAGGGGAGG is found 8 bp before *hydL* and 3 bp downstream from the TAG stop codon of *hydS* and the sequence GAGGG is located within the *hydL* sequence, 2 bp upstream from its stop codon and 11 bp upstream from the start codon of *hydD*.

Table 2. Comparison of *A. flavidum* HydS and HydL with the [Ni–Fe] hydrogenase subunits from several species

Organism	HydS		HydL	
	Subunit	Identity (%)	Subunit	Identity (%)
Alcaligenes eutrophus	HoxY	43·6	HoxH	45·0
Methanobacterium thermoautotrophicum	MvhG	44·7	MvhA	44·3
	FrhG	28·5	FrhA	34·0
Methanococcus voltae	VhuG	39·6	VhuA	45·2
	VhcG	39·1	VhcA	42·0
	FruG	30·7	FruA	33·6
	FrcG	24·0	FrcA	37·4
Pyrococcus furiosus	HydD	35·2	HydA	32·5
Desulfovibrio fructosovorans	HydS	31·8	HydL	30·7
Desulfovibrio gigas	HydS	22·3	HydL	30·9
Rhodobacter capsulatus	HupS	26·2	HupL	31·1
Bradyrhizobium japonicum	HupS	20·1	HupL	33·0
Azotobacter vinelandii	HoxK	27·3	HoxG	28·2
Escherichia coli hydrogenase-1 operon	HyaA	36·3	HyaB	29·6

Table 2 shows the results of a sequence homology study which suggests that the *A. flavidum* enzyme can be included in the class IV [Ni–Fe] hydrogenases described by Wu & Mandrand (1993), the highest homology being found with the *Alcaligenes eutrophus* HoxY small subunit (Tran-Betcke *et al.*, 1990). Interestingly, the small subunits of these two enzymes share the unique structural characteristic of being shorter than the small subunits of other class IV hydrogenases, with the C-terminal end truncated by approximately 50 amino acid residues. Considering that two out of the three CxxC Fe–S-cluster-binding motifs normally present in the small subunits of the [Ni–Fe] hydrogenases are located in the missing region, the presence of a single Fe–S-binding site would suggest an uncommon mechanism of electron transfer operating in both *Alc. eutrophus* and *A. flavidum* hydrogenases.

The presence of the sequences GxCAVxGG (in the central part of the protein) and VDVYxPGCPP in the *A. flavidum* small subunit are worth noting. The first sequence has been identified in all methylviologen-reducing hydrogenases (Wu & Mandrand, 1993), whereas the second one has been found in specific NADH:ubiquinone oxidoreductase subunits from different organisms (Xu *et al.*, 1992; Böhm *et al.*, 1990; Weidner *et al.*, 1993; Pilkington *et al.*, 1991).

As far as the *A. flavidum* large subunit is concerned, the highest degree of homology has been found with the large subunits of the hydrogenases belonging to class IV of the Wu & Mandrand (1993) classification, whereas the homologies with the subunits of the hydrogenases of class I and class II are less pronounced.

Among the most important structural motifs found in the *A. flavidum* large subunit are the two Ni-binding sites RxCxxCxxxH and DPCxxCxxH located at the N-terminus and C-terminus of the protein, respectively, and the sequence HHLxxxKxxD conserved among the methylviologen-reducing hydrogenases.

Although the *hydD* gene of the *A. flavidum* hydrogenase operon has been sequenced only partially, the homology study with the subunits of other hydrogenases deserves comment. In fact, on the basis of the similarity shared with the *E. coli* HycI protein, the protease which activates the hydrogenase large subunit HycE by cleaving its C-terminus (Rossmann *et al.*, 1995), it is reasonable to assume that the function of HydD is to cleave the *hydL* gene product. If this should be the case, the HydD protein might become crucial for the functional expression of the *A. flavidum* hydrogenase in heterologous systems (see below).

Expression of the *A. flavidum* hydrogenase in *E. coli*

A 2400 bp *Eco*RI–*Bam*HI DNA fragment carrying the genes encoding the small and large subunits of the *A. flavidum* hydrogenase was subcloned into the pBtac1 expression vector (De Boer *et al.*, 1983). The resulting plasmid was designated pSM707. The *E. coli* TG1 cells transformed with pSM707 were grown aerobically and total soluble proteins were analysed by SDS-PAGE. As shown in Fig. 4(a), both the small and the large subunits appeared to be expressed to a level which allowed their easy identification after Coomassie Blue staining.

Following the procedure described in Methods, the two subunits were purified (Fig. 4b). Interestingly, the two subunits appear to co-elute during the ion-exchange chromatography whereas they behave differently afterwards. This might indicate the existence of a weak interaction between the two proteins.

When the recombinant large subunit was electrophoresed together with the hydrogenase extracted from *A. flavidum*,

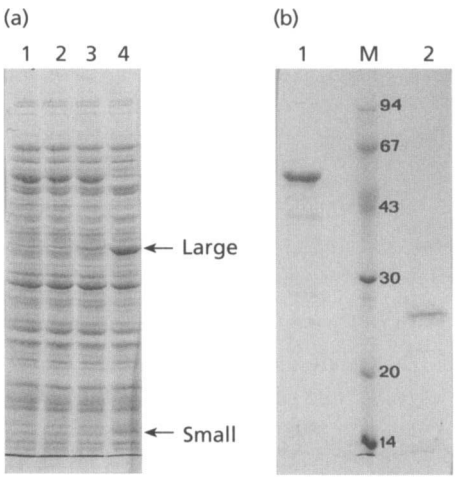

Fig. 4. Electrophoretic analysis of the *A. flavidum* hydrogenase expressed in *E. coli*. (a) SDS-PAGE (10% gel) of the total soluble proteins from *E. coli* TG1(pBtac1) (lanes 1 and 2) and *E. coli* TG1(pSM707) (lanes 3 and 4) cultures under induced (lanes 2 and 4) and uninduced (lanes 1 and 3) conditions. For details see text. (b) SDS-PAGE (12·5% gel) of the purified small (lane 2) and large (lane 1) subunits of the *A. flavidum* hydrogenase expressed in *E. coli*. M, Molecular mass standards of 94, 67, 43, 30, 20 and 14 kDa.

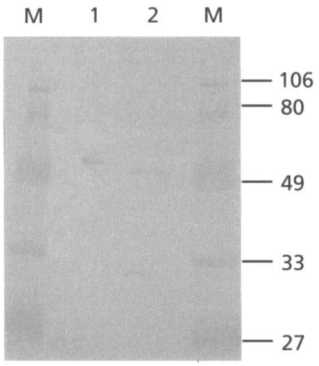

Fig. 5. Western blot analysis of the total soluble proteins from *A. flavidum* (lane 2) and the purified recombinant hydrogenase large subunit (lane 1). After electroblotting, the membrane was incubated in the presence of antibodies prepared by immunizing rabbits with a synthetic peptide having the sequence of the first 15 amino acids of the hydrogenase large subunit. M, Prestained molecular mass standards of 106, 80, 49, 33 and 27 kDa.

it became evident that the *E. coli*-produced protein migrated with an apparent molecular mass higher than the native subunit (Fig. 5). Since the N-terminal sequence revealed that the N-terminus of the recombinant protein is identical to the N-terminus of the native *A. flavidum* large subunit (data not shown), it is reasonable to assume that the different molecular mass of the two large subunits has to be ascribable to a difference at their C-termini.

When the recombinant enzyme was tested for hydrogen evolution capacity using the standard assay the enzyme turned out to be completely inactive.

The absence of enzymic activity in the recombinant enzyme can be attributed to (i) the inability of *E. coli* to process the large subunit (indeed the SDS-PAGE analysis clearly demonstrates the difference in size between the recombinant and the native enzyme large subunits) and (ii) the absence in the recombinant enzyme of the proper metal centres (Ni and/or Fe–S clusters). Preliminary data indicate that less than 1 mol Fe, Ni and S is present per mol enzyme, suggesting that the metal centres required for electron transfer might not be properly assembled in the new host. The inability of *E. coli* to assemble the metal centres in other Fe–S-containing enzymes when expressed at high levels has already been reported (Armengaud & Jouanneau, 1995). In addition to that, no difference in hydrogenase activity was detected between *E. coli* TG1 and the same strain carrying the hydrogenase genes when grown both aerobically and anaerobically (data not shown).

A large body of information now available clearly indicates that the [Ni–Fe] hydrogenases go through a complex pathway of maturation for which a number of ancillary proteins are required (Vignais & Toussaint, 1994). Generally, the genes encoding these proteins are clustered with the hydrogenase genes and many of them are organized in highly regulated transcriptional units. From the analysis of the chromosomal DNA so far sequenced, it appears that the *A. flavidum* hydrogenase does not represent an exception to this rule of complexity. In fact at least one other gene (*hydD*) is co-transcribed with the hydrogenase genes, and preliminary data obtained in our laboratories would indicate that other hydrogenase-associated genes might be present in the same region.

When these putative ancillary proteins are characterized, those strictly required for hydrogenase activation might be identified by their sequential co-cloning with the hydrogenase genes. The fact that *E. coli* expresses the *A. flavidum* hydrogenase in a soluble form is particularly promising in this context.

ACKNOWLEDGEMENTS

We are grateful to Dr J. Winter for the information about the *A. flavidum* strain and to Dr M. W. W. Adams for helpful discussion. We also wish to thank Dr Patrizia Polverino de Laureto for her collaboration in the N-terminal sequencing of hydrogenase. This work was partially supported by ENIRICERCHE S.p.A. and was performed under the management of the Research Institute of Innovative Technology for the Earth (RITE) as a part of the Research and Development Project on Environmentally Friendly Technology for the Production of Hydrogen supported by the New Energy and Industrial Technology Development Organization (NEDO), Japan.

REFERENCES

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

Armengaud, J. & Jouanneau, Y. (1995). Overexpression in *Escherichia coli* of the *fdxA* gene encoding *Rhodobacter capsulatus* ferredoxin II. *Protein Expr Purif* **6**, 176–184.

Atherton, E. & Sheppard, R. C. (1989). *Solid Phase Peptide Synthesis: a Practical Approach.* Oxford: IRL Press.

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

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal Biochem* **72**, 248–254.

De Boer, H., Comstock, H. L. & Vasser, V. (1983). The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoter. *Proc Natl Acad Sci USA* **80**, 21–25.

Dower, W. J., Miller, J. F. & Ragsdale, C. W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res* **16**, 6127–6145.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Legendre, N. & Matsudaira, P. (1989). Gel electrophoresis. In *A Practical Guide to Protein and Peptide Purification for Microsequencing*, pp. 52–69. Edited by P. Matsudaira. San Diego: Academic Press.

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

Pilkington, S. J., Skehel, J. M., Gennis, R. B. & Walker, J. E. (1991). Relationship between mitochondrial NADH-ubiquinone reductase and a bacterial NAD-reducing hydrogenase. *Biochemistry* **30**, 2166–2175.

Przybyla, A. E., Robbins, J., Menon, N. & Peck, H. D., Jr (1992). Structure-function relationship among the nickel-containing hydrogenases. *FEMS Microbiol Rev* **88**, 109–136.

Rossmann, R., Maier, T., Lottspeich, F. & Böck, A. (1995). Characterisation of a protease from *Escherichia coli* involved in hydrogenase maturation. *Eur J Biochem* **227**, 545–550.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.

Schindler, F. & Winter, J. (1987). Purification and properties of a soluble, thermophilic hydrogenase of *Acetomicrobium flavidum*. *Biochim Biophys Acta* **913**, 81–88.

Schuler, G. D., Altschul, S. F. & Lipman, D. J. (1991). A workbench for multiple alignment construction and analysis. *Proteins* **9**, 180–190.

Soutschek, E., Winter, J., Schindler, F. & Kandler, O. (1984). *Acetomicrobium flavidum*, gen. nov., sp. nov., a thermophilic, anaerobic bacterium from sewage sludge, forming acetate, CO_2 and H_2 from glucose. *Syst Appl Microbiol* **5**, 377–390.

Tran-Betcke, A., Warnecke, U., Bocker, C., Zaborosch, C. & Friedrich, B. (1990). Cloning and nucleotide sequences of the genes for the subunits of NAD-reducing hydrogenase of *Alcaligenes eutrophus* H16. *J Bacteriol* **172**, 2920–2929.

Veber, D. F., Milkowsky, J. D., Varga, S. L., Denkewalter, R. G. & Hirschmann, R. (1972). Acetamidomethyl. A novel thiol protecting group for cysteine. *J Am Chem Soc* **94**, 5456–5461.

Vignais, P. M. & Toussaint, B. (1994). Molecular biology of membrane-bound H_2 uptake hydrogenases. *Arch Microbiol* **161**, 1–10.

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

Weidner, U., Geier, S., Ptock, A., Friedrich, T., Leif, H. & Weiss, H. (1993). The gene locus of the proton-translocating NADH-ubiquinone oxidoreductase in *Escherichia coli*. *J Mol Biol* **233**, 109–122.

Wu, L.-F. & Mandrand, M. A. (1993). Microbial hydrogenases: primary structure, classification signatures and phylogeny. *FEMS Microbiol Rev* **104**, 243–270.

Xu, X., Matsuno-Yagi, A. & Yagi, T. (1992). Gene cluster of the energy-transducing NADH-quinone oxidoreductase of *Paracoccus denitrificans*: characterization of four structural gene products. *Biochemistry* **31**, 6925–6932.

Received 6 December 1995; accepted 13 December 1995.