Reassignment of the gene encoding the *Escherichia coli* hydrogenase 2 small subunit

Identification of a soluble precursor of the small subunit in a hypB mutant

Frank SARGENT, Stuart P. BALLANTINE, Paul A. RUGMAN, Tracy PALMER and David H. BOXER
Department of Biochemistry, Medical Sciences Institute, University of Dundee, UK

(Received 3 February 1998) -- EJB 98 0156/4

An active tryptic fragment of hydrogenase 2 from *Escherichia coli* has been isolated from the periplasmic face of the cytoplasmic membrane, and the large and small subunits N-terminally sequenced. The large subunit is encoded by the *hybC* gene and shows no N-terminal processing, other than removal of the initiator methionine during its biosynthesis. Both N-terminal and the subsequent internal tryptic-fragment amino acid sequence indicate that the small subunit is neither encoded by *hybA*, a gene previously identified as encoding the small subunit [Menon et al. (1994) *J. Bacteriol.*, 176, 4416–4423], nor any of the remaining genes in the *hyb* operon. Genome sequence analysis revealed the presence of an open reading frame which could potentially encode the peptide sequences of the proteolysed small subunit. The gene, designated *hyb0*, lies directly upstream of, and is separated by two nucleotides from, the start of the *hybA* gene. *Hyb0*, which shares an approximate 40% identity with other hydrogenase small subunit amino acid sequences, is synthesised with an N-terminal signal sequence containing a twin-arginine motif which is probably required for export of the enzyme. In the mature enzyme the small subunit is proteolytically cleaved after Ala37. Immunological analysis of strains overproducing either recombinant *Hyb0* or *HybA* using antibodies specific for hydrogenase 2, readily identified *Hyb0* as the small subunit. In a pleiotropic *hypB* mutant, which is unable to insert nickel into the active site, both the large and small subunits accumulate as unprocessed, soluble forms, consistent with the two subunits being assembled and processed in a coordinated manner during biosynthesis.

*Keywords*: nickel; *hyb*; periplasm; overexpression; signal sequence.

Hydrogenases are found in all the major taxonomic groups in the bacterial kingdom including *Enterobacteriaceae* and Archaea, and may also be present in some lower eukaryotes such as algae [1]. *Escherichia coli* contains at least three immunologically distinct, anaerobically expressed [NiFe] hydrogenases; hydrogenase 1, hydrogenase 2 and hydrogenase 3 [2–4]. Recent genetic evidence suggests that there may be a fourth [5].

Hydrogenase 1 is a transmembrane protein which is purified as a heterodimer of a 64-kDa large subunit and a 35-kDa small subunit [2]. By analogy with the periplasmic [NiFe] hydrogenase of *Desulfovibrio gigas*, the large subunit harbours the Ni-Fe dinuclear active site [6]. The *hya* operon encoding hydrogenase 1 has been cloned, sequenced and shown to comprise six open reading frames, *hyaABCDEFG* [7]. The *hyaA* gene encodes a 40.6-kDa, Fe-S protein that possesses a large N-terminal signal sequence. This signal sequence is removed during its biosynthesis to leave the mature, membrane-bound 35-kDa small subunit. The nickel-containing large subunit is encoded by *hyaB*. The roles of the remaining gene products remain less clear. However, *hyaC* has some sequence identity with *b*-type cytochromes, most notably that of the hydrogenase of *Woollinella succinogenes* [8], suggesting that hydrogenase 1 may display a more complex structure in vivo.

Hydrogenase 3 forms a loose association with the molybdoenzyme formate dehydrogenase-H on the cytoplasmic face of the cytoplasmic membrane, where they collectively make up the formate hydrogenlyase complex [9]. Hydrogenase 3 is encoded by an operon (*hyc*) located at 58 min on the *E. coli* genome, which is composed of nine open reading frames *hycABCDEFGH/* [4, 10]. The *HycE* gene encodes the nickel-containing large subunit. A hydrogenase-3 small subunit analogue has not been identified but the *HycG* protein could be a candidate. Of the other *hyc* gene products, only *HycA* and *HycC* have been characterised. *HycA* has a role in transcriptional regulation. *HycC* is a protease involved in post-translational processing of *HycE*.

Hydrogenase 2 was purified as an active tryptic fragment from the membranes of anaerobically grown, wild-type cells. The proteolysed fragment has a structure similar to that of hydrogenase 1 comprising two 61-kDa large subunits and two 31-kDa small subunits. Subsequent immunological analyses revealed that the trypsin treatment apparently left the large subunit untouched while about 5 kDa was lost from the small subunit [3]. An operon (*hyb*), believed to encode hydrogenase 2, has been cloned and sequenced [11]. The operon, present at 65 min on the *E. coli* chromosome, is composed of seven open reading frames *hybABCDEFG*. From the predicted sequences, *hybC* clearly encodes the hydrogenase-2 large subunit. *HybB*, like the *hyaC* gene product, has sequence similarity to *b*-type cytochromes. Surprisingly, none of the open reading frames encode.
a polypeptide of clear identity to any of the previously described small subunits of [NiFe] hydrogenases. It was suggested that HybA, an Fe-S protein with an N-terminal signal sequence, may represent an unusual small subunit [11]. In this communication, we show that the small subunit of hydrogenase 2 is encoded by a hitherto unsuspected gene, designated hyb0, which is directly upstream of hybA. We also show that the small subunit accumulates in a soluble precursor form in a hybB mutant which is pleiotropically unable to insert nickel into the large subunits of E. coli hydrogenases.

**MATERIALS AND METHODS**

**Bacterial strains.** The *E. coli* strains and plasmids employed in this study are listed in Table 1.

**Growth of bacteria and preparation of subcellular fractions.** Anaerobic growth of *E. coli* was in CR-fumarate medium [12], which increases the level of active hydrogenase 2 [13]. Anaerobic growth was accomplished in capped aspirator jars filled almost to the top. Strains carrying plasmids were grown anaerobically in Luria-Bertani medium at 37 °C, for 18 h. Sphaeroplast formation.

**Table 1. List of *E. coli* strains and plasmids used in this work.**

<table>
<thead>
<tr>
<th>Strain/plasmid</th>
<th>Genotype</th>
<th>Reference/source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC4100</td>
<td>F- *AlacU169 araD159 rpsL150 relA1 ptsF ribR fbdB5301</td>
<td>M. C. Casadaban</td>
</tr>
<tr>
<td>DHP-B</td>
<td>As MC4100 ΔhytB</td>
<td>Jacobi et al. [31]</td>
</tr>
<tr>
<td>M15[pREP4]</td>
<td>F- <em>lac ara mtl</em> Km*</td>
<td>Qiagen</td>
</tr>
<tr>
<td>P4X</td>
<td>His metB1</td>
<td>lab stocks</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pQE70</td>
<td>Amp*</td>
<td>Qiagen</td>
</tr>
<tr>
<td>pFS75</td>
<td>As pQE70, hybO</td>
<td>this study</td>
</tr>
<tr>
<td>pFS77</td>
<td>As pQE70, hybA</td>
<td>this study</td>
</tr>
</tbody>
</table>

**Materials and methods.** Hydrogenase and β-galactosidase activities in the soluble fraction were assayed as described below. Total enzyme activities were measured by sonication of sphaeroplasts followed by centrifugation (20000 g for 10 min at 4°C). Triton X-100 was added to the supernatant to a final concentration of 0.01%, the sample centrifuged again and the supernatant assayed for hydrogenase and β-galactosidase activities.

**Enzyme assays.** Hydrogenase-uptake activity was assayed spectrophotometrically as the H2-dependent reduction of oxidised benzyl viologen at 600 nm in anaerobic cuvettes as described by Ballantine and Boxer [14]. Samples were analysed in duplicate and the mean values displayed less than 10% deviation.

**β-Galactosidase activity was measured by the method described by Miller [16]. Each assay was performed in triplicate and the average result calculated; values rarely differed from the mean by more than 15%.**

**Purification of hydrogenase 2 tryptic derivative.** Two purification protocols were employed in the isolation of hydrogenase 2: the purification procedure exactly as described by Ballantine and Boxer [3] and a modified version described below. Washed cell paste of anaerobically grown *E. coli* (25 g) was resuspended in 25 ml 50 mM Tris/HCl, pH 7.5, 100 mM KCl, 1 mM benzamide, 10 μg/ml deoxyribonuclease I, 0.1% (mass/vol.) 2-mercaptoethanol, broken, and cell membranes prepared as described. Hydrogenase 2 activity was released from the membrane by the addition of trypsin (Sigma) to a final concentration of 2.5% (mass/vol.) membrane protein followed by incubation for 1 h at 30 °C. The reaction was stopped by the addition of soybean trypsin inhibitor (Sigma) to a final concentration of 5% (mass/vol.). Membrane residue was removed by centrifugation at 4°C for 90 min at 15000 g and the resultant supernatant was loaded onto a DEAE-Sepharose CL 6B column (11.5 cm x 4.5 cm) equilibrated with 50 mM Tris/HCl, pH 7.5, 0.1% (mass/vol.) 2-mercaptoethanol. After extensive washing, bound hydrogenase 2 activity was eluted in a single peak by application of a linear 0–200 mM KCl gradient (1000 ml). Hydrogenase-containing fractions were pooled and (NH₄)₂SO₄ was added to a final concentration of 0.5 M. The sample was then applied to a 5-ml FPLC-linked phenyl-Superose column, which had been equilibrated with 50 mM Tris/HCl, pH 7.5, 0.5 M (NH₄)₂SO₄, 5 mM benzamide, 0.1% (mass/vol.) 2-mercaptoethanol, and proteins were eluted with a linear gradient (30 ml) of 0.5–0M (NH₄)₂SO₄. Hydrogenase activity eluted as a single peak at 0M (NH₄)₂SO₄. The active fractions were pooled and concentrated to a volume of 0.5 ml using a Centricon-10 concentrator (Amicon). The concentrate was applied to an FPLC-linked Superose 12 gel-filtration column which had been previously equilibrated with 50 mM Tris/HCl, pH 7.5, 250 mM NaCl, 5 mM benzamide, 0.1% (mass/vol.) 2-mercaptoethanol, and proteins were eluted in the same buffer. The hydrogenase-containing fractions were pooled and rapidly frozen as 50-μl beads in liquid nitrogen. Preparations typically yielded 250 μg hydrogenase 2 tryptic derivative with a specific H₂/benzyl viologen oxidoreductase activity of 220 μmol benzyl viologen reduced · min⁻¹ · mg⁻¹.

**Peptide isolation and amino acid sequencing.** Hydrogenase 2 was purified by the method of Ballantine and Boxer [3]. 100 μg enzyme was subjected to SDS/PAGE, the small subunit excised from the gel and digested with trypsin by the method described by Knopf et al. [17]. The resultant tryptic peptides were separated by reverse-phase HPLC using a Waters C₁₈ column. Amino acid sequencing of peptides and N-terminal amino acid sequencing of proteins was carried out as described by Rivers et al. [18] using a Beckman 470A gas-phase automatic peptide sequencer.
Polyacrylamide gel electrophoresis and Western immunoblotting. Polyacrylamide gel electrophoresis in the presence of SDS was performed as described by Laemmli [19] using 10% or 12.5% (mass/vol.) acrylamide, 0.4% (mass/vol.) bisacrylamide. Immunoblotting of SDS/polyacrylamide gels was performed essentially as described by Towbin et al. [20]. After SDS/polyacrylamide gel electrophoresis, proteins were transferred to Hybond-C nitrocellulose membranes (Amersham) using a Bio-Rad Trans-blot cell. Immunoreactive proteins were detected with a rabbit anti-hydrogenase 2 serum (diluted 1:5000) coupled to the Promega ProtoBlot detection system (Promega).

Cloning and overexpression of hyb0 and hybA. The hyb0 gene was amplified from E. coli cells by PCR. The oligonucleotide primer targeted to the 5’ end of hyb0 had the sequence 5’-GGAAATAACTATGCACTGCCGAGAT-3’ and incorporated an engineered SphI recognition sequence. The primer targeted to the 3’ end of hyb0 had the sequence 5’-GGCCAGATCTTCTCCGCCGTAGTCAGCGT1A-3’ and incorporated an engineered BglII recognition sequence. The PCR product was digested with SphI and BglII and integrated into pQE70 to give plasmid pFS75.

A strategy identical to that designed for the cloning of hyb0 was employed in the cloning of hybA. The hybA gene was amplified from E. coli cells by PCR. The synthetic oligonucleotide primer targeted to the 5’ end of hybA had the sequence 5’-GGAGAATAACGACACGACGAATTT-3’ and incorporated an engineered SphI recognition sequence. The primer targeted to the 3’ end of hybA had the sequence 5’-GGCCAGATCTTCTCCGCCGTAGTCAGCGT1A-3’ and incorporated an engineered BglII site. The PCR product was digested with SphI and BglII before being integrated into the expression vector pQE70 to yield plasmid pFS77.

Both the hyb0 gene in pFS75 and the hybA gene in pFS77 were engineered such as to incorporate a polyhistidine tail onto the C-termini of the gene-products. DNA sequencing of the plasmids confirmed that the coding regions were correct and that the altered bases specified by the primers had been incorporated.

pFS75 and pFS77 were transformed separately into the E. coli host strain M15 carrying plasmid pREP4 which expresses high levels of lac repressor. Overexpression of the cloned genes was obtained by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to exponentially growing cultures at a final concentration of 2 mM.

Recombinant DNA techniques and DNA sequencing. Isolation of plasmid DNA, restriction enzyme digests, modification, ligation, transformation and electrophoresis were essentially as described [21]. Nucleotide sequencing reactions of the cloned DNA in plasmids pFS75 and pFS77 were performed using the ABI PRISM Dye Terminator Cycle Sequencing kit and products were analysed using an ABI 373 DNA sequencer. The nucleotide sequences of both DNA strands were determined. Synthetic oligonucleotides were supplied by Gibco BRL.

DNA sequences were collated, compiled and analysed using the GCG package, Daresbury Laboratory Seqnet Facility. Protein sequences were analysed using the Pileup and Prettybox programmes, on the GCG package, Daresbury Laboratory Seqnet Facility.

RESULTS

Tryptsin-mediated release of a soluble active fragment of hydrogenase 2 from membranes. The active preparation of hydrogenase 2 that has been previously described is a proteolytically derived of the enzyme obtained from material solubilised from a membrane preparation following exposure to trypsin [3]. This preparation contains a large subunit of 61 kDa which is electrophoretically indistinguishable from the large subunit present in the membrane-bound mature enzyme as revealed by Western immunoblot analysis. The second polypeptide of the trypsin-solubilised preparation, the small subunit, has an apparent molecular mass of 30 kDa and is derived from a subunit of apparent molecular mass 35 kDa, which is part of the membrane-bound enzyme [3]. It appears that the active fragment is released from the membrane following trypsin cleavage of the small subunit of the enzyme. No information is available regarding other polypeptides that may be associated with the mature membrane-bound enzyme. Advantage can be taken of the trypsin-mediated release of the hydrogenase 2 derivative to establish the location of this portion of the enzyme at the cytoplasmic membrane. Treatment of sphaeroplasts with trypsin over 60 min was sufficient to release approximately 40% of the total hydrogenase (H2: benzyl viologen oxidoreductase) activity (Fig. 1). That this activity was that of the hydrogenase 2 trypsin derivative was established by the immunoprecipitation of all of this activity with antibodies specific for hydrogenase 2. The immunoprecipitated enzyme, as anticipated, comprised the apparently mature large subunit and the smaller 31-kDa derivative of the small subunit. A control measurement indicated that there was little breakage of sphaeroplasts over the incubation period and that the hydrogenase release was dependent on the presence of trypsin. Hydrogenase 2 not released by trypsin during the experiment remained membrane-associated with the larger mature form of the small subunit (not shown). It is clear, therefore, that the large subunit and the major part of the small subunit of hydrogenase 2 is located at the periplasmic face of the cytoplasmic membrane. The small subunit could, therefore, have an important role in the anchoring of the enzyme to the periplasmic face of the membrane. Furthermore, this subunit, which is thought to be in extensive intimate contact with the large subunit [6], may play a role in the translocation of the Ni-centre-containing and Fe-S-centre-containing enzyme to its effective periplasmic location.

Identification of the gene encoding the small subunit of hydrogenase 2. In order to assess whether N-terminal proteolytic
processing, which is characteristic of periplasmically located proteins, had occurred for either the small or large subunits of the enzyme, and to attempt to locate the trypsin-sensitive sites of the membrane-bound small subunit, N-terminal amino acid sequencing was carried out on both the subunits present in the isolated hydrogenase 2-active derivative. The N-terminal amino acid sequence of the large subunit was exactly as predicted from the DNA sequence of the hybc gene, after removal of the N-terminal methionine, namely SQRITI. This confirmed that exposure to trypsin employed for the release of the enzyme from the membrane in the purification did not cleave the N-terminus of the large subunit. Despite its effective periplasmic location, it is clear that no N-terminal processing of the large subunit has occurred during its biosynthesis and assembly. This agrees with the conclusions of others [11] that the large subunit is encoded by the hybc gene and that the proteolytic processing of this subunit associated with nickel acquisition is confined to the C-terminus [22].

The N-terminal amino acid sequence obtained for the small subunit derivative of the hydrogenase 2 preparation was EM-AESVTNPQ. This sequence is not present in the putative gene hybC as was previously suggested. Furthermore, it cannot be encoded by any other gene in the operon. The two sequenced internal tryptic peptides, which were derived from internal parts of the subunit sequence, were isolated by reverse-phase HPLC. The sequences of two such peptides were NRPTFYAGR and LIHEHXER. Neither of these sequences correspond to any potential translated product of the hybc gene. This eliminates the possibility that the N-terminus of the small subunit is blocked and that the first sequence obtained was derived from a contaminant. It is clear, therefore, that the small subunit of hydrogenase 2 is not the product of the hybc gene as was previously suggested. Furthermore, it cannot be encoded by any other gene in the hybcABCDEFG operon.

Reference to the E. coli genome sequence database [23] lead to the solution of the problem. A putative further, previously undetected, gene is present in the database immediately upstream of hybc. Comparison of the amino acid sequences of the possible products of this additional gene with the amino acid sequences determined for the small subunit of hydrogenase 2, confirmed that one of the reading frames for this gene encodes the small subunit of the enzyme. We designate this gene hyba. The translated product of the hyba gene is a protein of 372 residues and molecular mass of 39652 Da (Fig. 2).

Proteolytic processing and trypptic cleavage of the hydrogenase 2 small subunit. The two sequenced internal tryptic peptides are derived from adjacent regions located near the middle of the hyba primary translated gene product (Fig. 2). The N-terminal sequence of the small subunit, however, corresponded to the sequence starting at Glu38, Tryptic cleavage cannot be responsible for this result since the amino acid residue predicted for the N-terminal side of Glu38 is not Lys or Arg but Ala. It has previously been suggested that the small subunits of Ni-Fe hydrogenases may be proteolytically processed at the N-terminus during biosynthesis and assembly [1]. The sequence analysis for hydrogenase 2 indicates that this is indeed the case.

A post-translational processing cleavage at Glu38 during biosynthesis would yield a mature, membrane-bound small subunit with a predicted molecular mass of 35794 Da. This is very close to that estimated for the mature membrane-bound small subunit of 35 kDa [3]. A signal peptide of 37 residues is removed during biosynthesis. A motif RXRXFK has been identified for several bacterial species, including E. coli, as a signal sequence that leads to the N-terminal cleavage of the signal peptides for periplasmically located proteins that contain prokaryotic groups [24]). The sequence RRDFMK occurs at residues 15–20 in the deduced small subunit amino acid sequence (Fig. 2). This is consistent with the periplasmic orientation of the small subunit and the presence of Fe-S centres within this subunit [6].

Tryptsinolysis, which liberates the active fragment of hydrogenase 2 from the membrane, cannot take place at the N-terminal part of the small subunit. The small subunit derivative present in the isolated trypsin-released enzyme has an apparent molecular mass of about 30 kDa. The site of trypsin cleavage which leads to the release of hydrogenase 2 from the membrane must be located at about residue 325 in order for a 5-kDa C-terminal fragment to be removed. Three candidate trypsin cleav-

Fig. 2. DNA sequence of the region covering hyba. This shows part of the E. coli chromosome as sequenced by Blattner et al. [23], which lies upstream of the hyba gene. The nucleotides are numbered from the translational start of hyba. The predicted amino acid sequence of Hyba is given in single-letter code below the DNA sequence. Possible ribosome-binding sites for hyba and hybc are indicated by a single line and the letters RBS above the sequence. Amino acid sequences which are underlined are those for which the sequence has been obtained experimentally from the tryptic small subunit fragment. The N-terminus of the mature Hyba is denoted in bold. Consensus signal sequence motifs [24] for Hyba and HybA are indicated. The arrows indicate putative sites for trypsin cleavage of the small subunit from the membrane.
The coding sequence for hybA was also amplified in a similar manner to that for the small subunit to give plasmid pFS77. A high level of expression of the C-terminally His-tagged HybA protein was obtained (Fig. 3 A).

The His-tagged small subunit (Hyb0) and the His-tagged HybA fusion proteins were visualised as protein-staining bands on SDS/PAGE of 39 kDa and 35 kDa, respectively, which are consistent with the predicted molecular masses for the recombinant proteins of 40.7-kDa and 34.8-kDa. N-terminal sequence analyses of the recombinant proteins confirmed that neither was proteolytically processed in the expression system. The overexpressed proteins were each recovered as inclusion bodies in the pellet fraction following cell breakage and centrifugation. Western immunoblot analysis using antibodies specific for hydrogenase 2, of SDS/PAGE gels of crude extracts of cells overexpressing either Hyb0 or HybA, readily identified the hyb0 gene product as the small subunit of the enzyme (Fig. 3 B). The antibodies did not cross react with hybA gene product even though large amounts were present on the gel. This independently confirms the identity of hyb0 as the small subunit gene and shows that the HybA polypeptide is not recognised by the anti-hydrogenase 2 serum.

The hyb operon. The result of this reassessment of the hyb operon is shown (Fig. 4). The hyb0 gene is located directly upstream of the previously identified hybA gene within the hyb operon. The stop codon of hyb0 is separated from the putative start codon of hybA by 2 bp (Fig. 2). A potential Shine-Dalgarno sequence is located 7 bp upstream of the hyb0 start within hyb0. It appears that hyb0 is an integral part of the hyb transcriptional unit under control of a promoter region upstream of hyb0. There is no further potential gene in the immediate upstream vicinity of hyb0.

The complete primary structure of Hyb0 demonstrates excellent similarity with previously identified hydrogenase small subunits (Figs. 5). Hyb0 has 42.5% identity with Rhodobacter capsulatus hydrogenase small subunit, HupS, 41.1% sequence similarity with Azotobacter vinelandii HoxK, 37.0% sequence identity with E. coli HyaA (hydrogenase 1 small subunit), and 36.8% identity with Desulphovibrio gigas HynB. The sequence comparison demonstrates convincingly that the small subunit contains three Fe-S centres, as has been shown by X-ray crystallography for the D. gigas hydrogenase [6].

The hybA gene product shows some sequence similarity to hydrogenase small subunits to which it has been previously compared [11]. A database search revealed that HybA has 36% sequence identity with HmcB from D. vulgaris (Fig. 6). HmcB, which has an N-terminal ‘twin-arginine’ signal sequence, is located at the periplasmic face of the cytoplasmic membrane to

---

**Fig. 3.** Overproduction and immunological analysis of Hyb0 and HybA. (A) Overproduction of the Hyb0 and HybA proteins as histidine-tagged C-terminal fusions. Protein samples in lanes 1–4 are 3 mg whole-cell protein of the strain M15 [pREP4] carrying the following plasmids: lane 1, pFS75 (hyb0), before IPTG induction; lane 2, pFS75 after induction of expression of plasmid-encoded proteins with IPTG; lane 3, pFS77 (hybA) after induction of expression of plasmid-encoded proteins with IPTG; lane 4, pFS77 before IPTG induction. The His6-tagged small subunit of molecular mass 30 765 Da, and Lys327 (3 1 µ300 Da) are present in the polypeptide. Cleavage could occur at one of these sites specifically, or indeed at all three (Fig. 2). The exact site of trypsin cleavage remains to be established.

**Overexpression and immunological analysis of the small subunit and the hybA gene product.** The coding sequence for hyb0 was amplified by PCR and the resulting fragment cloned into the C-terminal His-Tag vector pQE70 to produce plasmid pFS75 (see Materials and Methods). The His6-tagged small subunit (Hyb0) with the His6-tag at the C-terminus of the protein was overproduced in the E. coli strain M15/pREP4 transformed with pFS75. Overexpressed His6-tagged Hyb0 was produced to a high level following induction with IPTG (Fig. 3 A).
**Fig. 5. Comparison of hydrogenase small subunit sequences.** Alignments were made using the Pileup and Prettybox programmes (Daresbury).

| hoxk | hyaa | hyb0 |*
|------|------|------|
| hoxk | hyaa | hyb0 | 75
| hoxk | hyaa | hyb0 | 1

**Fig. 6. Proteins with identity to HybA.** Alignments were made using the Pileup and Prettybox programmes (Daresbury).

<table>
<thead>
<tr>
<th>hmcb</th>
<th>hyba</th>
<th>nrfc</th>
<th>fdoh</th>
</tr>
</thead>
</table>
| hmcb | hyba | nrfc | fdoh | 715
| hmcb | hyba | nrfc | fdoh | 52
The hyp operon contains several genes that appear to catalyze cofactor acquisition by the hydrogenases during biosynthesis and are required for the active expression of all E. coli hydrogenases. Mutations in the hypB gene lead to the pleiotropic loss of all hydrogenase activities due to the inability of such mutants to insert nickel into the large subunit of the hydrogenases during biosynthesis [31, 34]. Western immunoblot analysis of the hydrogenase 2 small subunit region of an SDS/PAGE gel (Fig. 7A) allows direct comparison with the mature (N-terminally processed), trypsin exposed (N-terminally processed and cleaved near the C-terminus by trypsin) and recombinant, his-tagged but unprocessed small subunit. It can be seen that the small subunit of hydrogenase 2 in this hypB deletion strain is unprocessed and is located exclusively in the soluble fraction.

Mutants defective in hypB can be restored to active hydrogenase sufficiency by growth in media supplemented with nickel (II) salts [31, 34]. Although the primary lesion in a hypB strain is the inability to insert nickel into the large subunit, the biosynthesis of the small subunit is disrupted in an apparently similar manner to the large subunit (Fig. 7B). Consistent with what has been shown previously for other hydrogenases, the large subunit is found in a non-proteolytically processed form in the soluble fraction of the mutant. This subunit becomes proteolytically processed and associated with the membrane when the mutant is grown in media supplemented with NiCl2. Consistent with the very large amount of inter-subunit contact anticipated between the large and small subunits in the mature active hydrogenase as suggested by X-ray crystallography [6], it appears that the two subunits are assembled and processed in a coordinated manner during biosynthesis.

**DISCUSSION**

The DNA sequence of the structural operon hybABCDEFG for E. coli hydrogenase 2 was established by Menon et al. [33]. The purified enzyme, characterised as an H2:benzyl viologen oxidoreductase, was obtained following trypsin-dependent release from the membrane fraction and was characterised as a soluble tetramer comprising two large subunits of 61 kDa and two small subunits of 31 kDa [3]. That the hybc gene encoded the large subunit of the enzyme was established following the determination of some amino acid sequence for the large subunit [33]. The authors were less certain about the gene encoding the small subunit but suggested, following predicted sequence comparison, that the small subunit was encoded by hyba, noting, however, that the small subunit sequence appeared somewhat unusual. It is clearly established here that the small subunit of the enzyme is not encoded by hyba but by a previously overlooked gene, termed hyb0, within the operon but located immediately upstream hyb0.

The location at the periplasmic face of the cytoplasmic membrane of that portion of hydrogenase 2 released from the membrane by trypsin exposure comprising the large subunit and a large portion of the small subunit has been established, confirming the results reported by Rodrigue et al. [35]. The large subunit is not subjected to N-terminal processing during biosynthesis but the small subunit is processed during assembly to release a 37-residue leader sequence from its N-terminus. The small subunit is clearly substantially exposed at the periplasmic surface since trypsin cleaves the subunit at one of three adjacent sites situated at about 50 amino acid residues from the C-terminus of the 335-residue mature polypeptide. The amino acid sequences of both the large and small subunits are closely similar to those of the D. gigas hydrogenase for which a three-dimensional structure has been obtained by crystallography [6]. In that structure, the
three Fe-S-centre-containing small subunit makes extensive protein-protein contact with the Ni-Fe-containing large subunit. Molecular modelling (results not shown) indicates that a similar situation pertains for these subunits of *E. coli* hydrogenase 2. However, only the small subunit has the necessary leader sequence for membrane translocation. It is likely that the large subunit associates with the small subunit after acquiring their cofactors, prior to its translocation as a complex to the periplasmic face of the membrane. The membrane-bound hydrogenase of *Alicaligenes eutrophus* resembles *E. coli* hydrogenase 2 in that both the large and small subunits are exposed to the periplasm [36]. Removal of just the small subunit signal sequence prevented membrane translocation and activation of the enzyme complex in this organism [37]. The *E. coli* trimethylamine-N-oxide reductase, a molybdoenzyme located in the periplasm, is also thought to fold and acquire its cofactor prior to membrane translocation via a Sec-independent export pathway [38].

The clear identification of the hydrogenase 2 small subunit has allowed its study in a *hypB* mutant, which is defective in the ability to insert nickel into the large subunit during biosynthesis [31, 34]. The small subunit, in its precursor form, is found exclusively in the soluble fraction in this mutant, as is the large subunit. Recovery of hydrogenase activity by growth with nickel chloride, leads to the mature forms of both subunits being located in the membrane fraction. This strongly supports a coordinated assembly and translocation process for these subunits. Brito et al. [39] also reported that the small subunit of the *Rhizobium leguminosarum* [NiFe] hydrogenase remained unprocessed in *hypB* mutant.

The precursor form of the small subunit incorporates a 37-residue N-terminal signal peptide. The signal peptide has an RRXFXK motif which has been implicated in effecting localisation of proteins containing prosthetic groups, to the periplasm or to the periplasmic face of the cytoplasmic membrane [24, 40]. The *hypB* gene product also possesses the RRXFXK motif and appears to associate with four Fe-S centres. It seems highly likely, therefore, that this protein is also located at the periplasmic face of the cytoplasmic membrane. The nature of the *HybA* protein places it as a strong candidate for being part of the membrane-bound hydrogenase 2 enzyme, probably serving as the electron acceptor from the small subunit. Further experimentation is clearly required to assess these suggestions. It seems clear, however, that hydrogenase 2 possesses two polypeptides with the twin-arginine motif. The possibility that *HybA* protein translocation and assembly is coordinated with that of the small and large subunits of the enzyme also remains to be addressed. The *hyp* operon is an example of an operon encoding more than one polypeptide with the twin-arginine motif, others are the *E. coli* and *H. influenzae* *hypB* operons which encode periplasmic nitrate reductases [30, 41]. The cleavage site for the release of the signal peptide of the small subunit occurs at Ala37-Glu38 to the C-terminal side of a triple alanine sequence for membrane translocation. It is likely that the physiological complex includes HybA, HybB, a polypeptide with identity to the cytochrome-b-g subunit of *W. succinogenes* hydrogenase [8], along with the large and small subunits. It is less clear whether any of the other *hyp* gene products are also part of the active enzyme, indeed the *hybD* gene product appears to be a protease required for releasing a C-terminal peptide from the large subunit following nickel acquisition [10]. HybA appears to have no homologues in the gene clusters for the well characterised *A. eutrophus* or *Bradyrhizobium japonicum* hydrogenases [37, 43]. The *hya* operon encoding the closely related hydrogenase 1 of *E. coli* also lacks a HybA analogue. The present work offers a route to the further definition of the membrane-bound hydrogenase 2.

This work is supported by a BBfSRC grant to DHB. Derek Richard and Ben Berks are thanked for useful discussion.

**REFERENCES**


