JOURNAL OF BACTERIOLOGY, Apr. 1990, p. 1969–1977
0021-9193/90/041969-09$02.00/0

Vol. 172, No. 4

Cloning and Sequencing of a Putative *Escherichia coli* [NiFe] Hydrogenase-1 Operon Containing Six Open Reading Frames

NANDA K. MENON, JEFFREY ROBBINS, HARRY D. PECK, JR., CORINNE Y. CHATELUS, EUI-SUNG CHOI, AND ALAN E. PRZYBYLA*

Department of Biochemistry, School of Chemical Sciences, University of Georgia, Athens, Georgia 30602

Received 31 July 1989/Accepted 18 December 1989

DNA encompassing the structural genes of an *Escherichia coli* [NiFe] hydrogenase has been cloned and sequenced. The genes were identified as those encoding the large and small subunits of hydrogenase isozyme 1 based on NH_2-terminal sequences of purified subunits (kindly provided by K. Francis and K. T. Shanmugam). The structural genes formed part of a putative operon that contained four additional open reading frames. We have designated the operon *hya* and the six open reading frames *hyaA* through *F*. *hyaA* and *hyaB* encode the small and large structural subunits, respectively. The nucleotide-derived amino acid sequence of *hyaC* has a calculated molecular mass of 27.6 kilodaltons, contains 20% aromatic residues, and has four potential membrane-spanning regions. Open reading frames *hyaD* through *F* could encode polypeptides of 21.5, 14.9, and 31.5 kilodaltons, respectively. These putative peptides have no homology to other reported protein sequences, and their functions are unknown.

The anaerobic hydrogen metabolism of *Escherichia coli* and other enterobacteria is intricately regulated with regard to both hydrogen production during fermentation and hydrogen oxidation during anaerobic respiration (13, 14). This complex regulatory system responds to specific substrates as well as to global regulatory signals and results in the biosynthesis of discrete hydrogenases specific to a given metabolic pathway. Three *E. coli* hydrogenases have been described which are synthesized in response to different physiological conditions. Hydrogenases 1 and 2 have been biochemically characterized and are immunologically distinct membrane-bound nickel-containing proteins (2, 3, 31). The existence of hydrogenase 3 was originally inferred from the fact that hydrogenases 1 and 2 immunoprecipitated from cell lysates did not account for total hydrogenase activity (30). Hydrogenase 3 activity is very labile and has been only partially characterized (34). Hydrogenases 1 and 3 are induced to higher levels by growth on glucose and formate. Hydrogenase 3 has also been shown to have a role in the formate hydrogenlyase pathway and accounts for about 60 to 70% of the total hydrogenase activity in the cell. The physiological role of hydrogenase 1 has not been defined. Hydrogenase 2 has been implicated as a respiratory uptake hydrogenase coupled to fumarate reduction and is induced to a higher level by growth in the presence of glycerol and fumarate (30).

The presence of three different enzymes catalyzing the same reactions makes the biochemical, physiological, and genetic analyses of their metabolic roles technically difficult. Toward understanding the physiological roles of the different hydrogenases, a large number of *E. coli* mutants defective in hydrogenase activities have been analyzed. Most of these mutants lack all three hydrogenase activities; based on genetic analysis, the mutations appear to be exclusively at loci affecting regulation and do not encompass the hydrogenase structural genes. These genetic loci have been designated as *hydA* through *F* (8, 16, 18, 20, 27–29, 35, 42, 43). The *hydC* locus appears to be involved in nickel uptake and/or processessing (32, 42, 43), but the function of the other loci remains to be elucidated. One *E. coli hyd* mutant was unusual in that it lacked hydrogenase 3 activity but contained normal levels of hydrogenases 1 and 2 (35). It is not clear whether this locus, which maps at 59 min, contains the structural genes encoding hydrogenase 3.

In this paper we report the nucleotide sequence of a putative operon that encodes the structural genes of the two subunits of hydrogenase 1 and contains four additional open reading frames. We have designated the operon *hya* and the open reading frames *hyaA* through *F*.

MATERIALS AND METHODS

Genomic library construction. The genomic library used in these studies consisted of size fractionated 15- to 20-kilobase (kb) *Sau*3A genomic partial digests ligated into the *Bam*HI site of the lambda vector, EMBL4. Twenty genomic equivalents transfected into *E. coli* LE392 were screened for the structural genes encoding [NiFe] hydrogenase.

Probes. Hydrogenases from several bacteria contain stretches of amino acids that are highly conserved in both the large and small subunits. An M13 clone containing one of these regions from the large subunit of the [NiFe] hydrogenase of *Desulfovibrio vulgaris* (approximately amino acids 50 through 150; see Fig. 7) was used as a probe (probe 1) to screen the *E. coli* genomic library. The clone was labeled by a sequencing reaction in the presence of [α-^{32}P]dATP (26). The labeled insert was purified from M13 by *Eco*RI-*Hin*dIII digestion followed by fractionation in a 1% low-melting-point agarose gel. A second probe (probe 2), consisting of a 400-base-pair *Pst*I fragment from the carboxy-terminal end of the small subunit, was labeled by the primer extension method with [α-^{32}P]dATP (22). This probe was used to confirm the positives isolated with probe 1.

Primary library screens. About 5,000 plaques grown overnight on three petri plates were blotted for 2 min to nitrocellulose filters. The filters were processed as previously described (22) and baked for 2 h at 80°C. All heterologous hybridizations were done in 37% formamide at 42°C for 16 h, followed by washing three times for 20 min each at 60°C in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate). The filters were exposed for 16 h at −70°C to Kodak

* Corresponding author.

X-Omat film. Positive plaques identified in the primary screen were plaque purified at least twice before use in further experiments.

Genomic Southern hybridizations. Genomic Southerns were done with *Eco*RI-*Sal*I and *Eco*RV digests of *E. coli* DNA. The blots were hybridized to probe 1 in 37% formamide at 42°C for 16 h. The filters were washed in 2× SSC at 60°C and exposed to Kodak X-Omat film for 24 h.

Subclone construction and sequencing. An *Eco*RI digest of the positive lambda clone yielded a 7.5-kb fragment, which was deduced (by restriction mapping and hybridizations to the two different probes) to contain the genes encoding the hydrogenase structural subunits. This restriction fragment was subcloned into the plasmid pTZ19R. Clones containing the insert in both orientations with respect to the *lacZ* promoter were isolated and used to generate deletion libraries by the DNase 1 random nicking method (11). The deletion clones were sized on 1% agarose gels, and 50 clones in both directions representative of the entire 7.5-kb insert were isolated. Single-stranded templates were generated from the pTZ19R clones (25). Isolated colonies of *E. coli* NM522, freshly transformed with the deletion clones were picked after 12 h and grown for 3 h in 3 ml of 2× TY liquid broth (16 g of tryptone, 10 g of yeast extract, and 5 g of NaCl per liter, pH 7.25). They were then infected with the helper phage M13K07, and growth was continued for 2 h. Samples of 100 μl were transferred to 5 ml of 2× TY containing 50 μg of kanamycin per ml and grown overnight with vigorous shaking. Single-stranded templates were isolated and purified as described for M13 templates in the cloning and sequencing handbook of Amersham Corp. (Arlington Heights, Ill.).

Sequencing was done by the dideoxy-chain termination method (26). T7 DNA polymerase (Sequenase; U.S. Biochemical Corp., Cleveland, Ohio) was used for the reactions, and DNA fragments were fractionated in 6% wedge gels (0.25 to 0.75 mm thick). Compressions were resolved in duplicate reactions by using dITP in place of dGTP.

Sequence analyses and comparisons were done with Beckman Microgenie, Staden (33), PC/Gene, and IntelliGenetics programs.

Materials. All restriction and DNA-modifying enzymes were purchased from American Allied Biochemicals, Boehringer Mannheim Biochemicals, and New England BioLabs, Inc., and were used per manufacturer specifications. Sequencing reagents were purchased from U.S. Biochemical Corp. Radioactive isotopes were purchased from Amersham.

RESULTS AND DISCUSSION

Isolation and identification of hydrogenase clones. A restriction map of a 4.8-kb *Cla*I-*Sma*I fragment containing the structural genes encoding the two subunits of the membrane-bound [NiFe] hydrogenase from *D. vulgaris* is shown in Fig. 1. The regions of the *D. vulgaris* genes used as probes to screen an *E. coli* genomic library and probe *E. coli* genomic digests have been indicated. The region of the large subunit gene comprising probe 1 is conserved in both [NiFe] and [NiFeSe] hydrogenases.

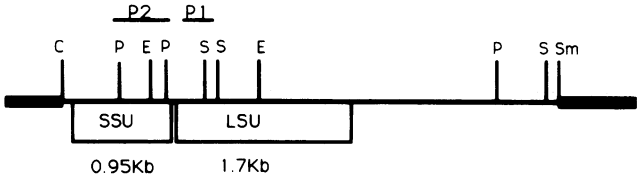


FIG. 1. Restriction map of a 4.8-kb *Cla*I-*Sma*I restriction fragment from *D. vulgaris* containing the genes encoding the large and small subunits of a membrane-bound [NiFe] hydrogenase. Probe 1 represents an M13 clone containing the marked region, and probe 2 represents a *Pst*I fragment from the small subunit. Restriction sites: C, *Cla*I; E, *Eco*RI; P, *Pst*I; S, *Sal*I; Sm, *Sma*I.

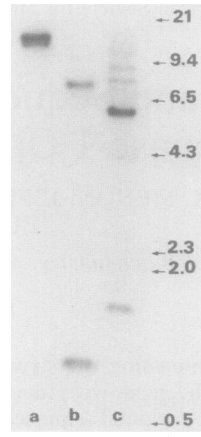


FIG. 2. Southern blot of *E. coli* genomic DNA hybridized to probe 1 for 16 h at 42°C in 37% formamide. Digests (lanes): a, *Pst*I; b, *Eco*RI-*Sal*I; c, *Eco*RV. Standard markers used were from *Hin*dIII-digested lambda DNA.

E. coli DNA digested with *Eco*RV (Fig. 2, lane c) or with a combination of *Eco*RI and *Sal*I (Fig. 2, lane b) showed two regions of homology to probe 1. This indicated that either there were restriction sites for both enzymes in the same region of the *E. coli* hydrogenase large subunit gene or that the two signals represented two different hydrogenase genes. The following results suggest the presence of two sets of hydrogenase genes.

Eighteen putative hydrogenase clones were isolated with probe 1, and four were further characterized. Three of these clones were identical, and the data for one of these clones, EC1, are presented. Probe 1 hybridized to a 1-kb *Eco*RI-*Sal*I fragment (Fig. 3, lane a) and a 5.8-kb *Eco*RV fragment (data not shown) of EC1. These bands corresponded to the 1- and 5.8-kb signals obtained on genomic digests with the same enzymes in 37% formamide (Fig. 2, lanes b and c). Probe 2 hybridized to a 540-base-pair *Eco*RI-*Sal*I fragment (Fig. 3, lane b) and a 5.8-kb *Eco*RV fragment (data not shown) of EC1. This indicated that the *Eco*RV fragment contained

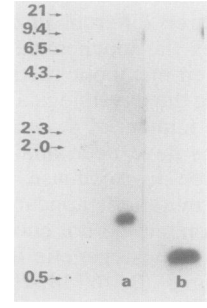


FIG. 3. Southern blot of a positive lambda EMBL4 clone, EC1, restricted with *Eco*RI-*Sal*I and hybridized to probe 1 (lane a) and probe 2 (lane b) for 16 h at 42°C in 42% formamide. Standard markers used were from *Hin*dIII-digested lambda DNA.

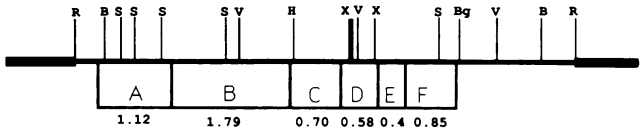


FIG. 4. Restriction map of a 7.5-kb *Eco*RI fragment isolated from EC1 containing the *hya* operon. The *Eco*RI site at the 5′ end of the operon is derived from the lambda vector. Restriction sites: B, *Bam*HI; Bg, *Bgl*I; H, *Hpa*I; R, *Eco*RI; V, *Eco*RV; S, *Sal*I; X, *Xmn*I. ORFs are labeled A through F, and their prospective sizes in kilobase pairs are indicated below.

genes encoding both the large and small subunits of a [NiFe] hydrogenase, and that there were internal *Eco*RI and/or *Sal*I sites within the operon. These results also implied that the 7-kb *Eco*RI-*Sal*I and 1.8-kb *Eco*RV signals obtained with probe 1 on genomic digests (Fig. 2, lanes b and c) represented a second set of hydrogenase genes.

It has been reported that all three hydrogenases in *E. coli* are immunologically distinct (2). Similarly, the antibodies to the *Desulfovibrio* [NiFe] and [NiFeSe] hydrogenases do not cross-react (10). However, the nucleotide probes we used for screening represented conserved domains found in all [NiFe] hydrogenases sequenced to date, including the [NiFeSe] hydrogenase of *Desulfovibrio baculatus*. Hence, it is to be expected that our screening procedure would identify more than one set of hydrogenase genes.

Sequence analysis of the *hya* operon. A 6.0-kb *Eco*RI-*Bgl*II region of the 7.5-kb *Eco*RI fragment isolated from EC1 was found to contain the entire *hya* operon (Fig. 4). Six putative open reading frames (ORFs) were identified which could encode polypeptides of 40.6, 66.2, 27.6, 21.5, 14.9, and 31.5 kilodaltons, respectively. The polypeptides coded by ORFs 1 and 2 showed homology to the amino acid sequences of the small and large subunits of [NiFe] hydrogenases from other organisms (Fig. 5).

The gene coding for the small subunit (*hyaA*) was located upstream of the large subunit gene. A consensus ribosome-binding site occurs 8 bp upstream from the start of the small subunit. Possible −10 and −35 promoter regions as identified by a Staden analysis program have been indicated (Fig. 5). The putative promoters do not share a strong homology to the consensus *E. coli* −10 and −35 promoter elements (25). In the 150 nucleotides of sequence 5′ to the small subunit gene, neither a consensus anaerobox (6) indicative of *fnr*-mediated regulation nor a formate box (5) characteristic of formate-inducible systems was identified. Both have been implicated in the regulation of hydrogenase 1 (5, 6, 30). Further sequence analysis of this region and transcript mapping are required to characterize the promoter and regulatory elements. The stop codon of the small subunit gene was found to overlap with the initiator methionine codon of the large subunit gene. A putative ribosome-binding site was identified 9 base pairs upstream from the start of the large subunit, within the coding region of the small subunit.

Protein sequencing of the amino-terminal end of the small subunit of *E. coli* hydrogenase 1 revealed that the first 15 residues are identical to the nucleotide-derived amino acid sequence of the small subunit gene (*hyaA*) starting from residue 45 (K. Francis and K. T. Shanmugam, personal communication). This indicates that *hyaA* is the small subunit gene and codes for a putative 45-amino-acid signal peptide. The first 10 residues of the large subunit (Francis and Shanmugam, personal communication) are colinear with the start of the nucleotide-derived sequence of *hyaB*; thus the large subunit is not synthesized with a signal peptide. The molecular weights of the two subunits of hydrogenase 1 as determined by electrophoresis are also in good agreement with the data obtained from the nucleotide derived sequences of *hyaA* and *B* (31).

Comparison of [NiFe] hydrogenases. The small subunits of hydrogenases from various organisms, including periplasmic [NiFeSe] and [NiFe] hydrogenases from the anaerobic sulfate reducers *D. baculatus* (23) and *Desulfovibrio gigas* (21, 39) as well as membrane-bound [NiFe] hydrogenases from the facultative anaerobe *E. coli*, the nitrogen fixer *Bradyrhizobium japonicum* (32), and the phototroph *Rhodobacter capsulatus* (19) are compared in Fig. 6. These hydrogenases show different degrees of homology and fall into two groups when classified by overall sequence homology. There is 60 to 70% homology between the nucleotide-derived amino acid sequences of the small and large subunits of the [NiFe] hydrogenases from *B. japonicum*, *R. capsulatus*, and *E. coli*. The enzymes from these systems, however, share only a 30 to 35% overall homology with the *D. gigas* [NiFe] and *D. baculatus* [NiFeSe] hydrogenases. A third group of nickel-containing hydrogenases represented by the [NiFe] hydrogenase of *Methanobacterium thermoautotrophicum* (24) has less than 20% homology with the other two groups of hydrogenases. Nevertheless, there are interesting conserved domains and features shared by all nickel-containing hydrogenases sequenced to date.

Except for the small subunit of the [NiFe] hydrogenase from *M. thermoautotrophicum*, the small subunits of the hydrogenases are characterized by unusually long signal peptides (40 to 45 residues in [NiFe] hydrogenases and about 35 residues in [NiFeSe] hydrogenases that do not conform to the consensus structure of procaryotic signal peptides (37). These putative signal peptides, which share limited homologies (Fig. 6), appear to resemble some mitochondrial and chloroplast transit peptides and have the structural characteristics necessary to form amphipathic helices.

The gene for the small subunit of the [NiFe] hydrogenase 1 of *E. coli* encodes about 50 more amino acids than do the [NiFe] hydrogenases of sulfate-reducing bacteria. This carboxy-terminal extension contains a central stretch of 20 hydrophobic amino acids followed by a shorter stretch of 10 to 12 charged residues. A homologous extension is found in the derived amino acid sequence of the [NiFe] hydrogenases from *B. japonicum* (32), *R. capsulatus* (19), and *A. chroococcum* (36; C. M. Ford, K. H. Tibelius, M. G. Yates, D. J. Arp, and L. G. Seefeldt, 7th Int. Congr. N_2 Fixation, p. 275, 1988). The carboxy terminus of the small subunit of the *M. thermoautotrophicum* [NiFe] hydrogenase, a cytoplasmic enzyme (41), lacks the extension (24). Hence it is possible that this stretch of amino acid residues plays a role in the membrane attachment of the protein. However, this motif is not found in the membrane-bound [NiFe] and [NiFeSe] hydrogenases of *D. vulgaris*, suggesting that the mode of membrane attachment may be different in these anaerobic bacteria.

The genes for the large subunits of the [NiFe] hydrogenases encode a variable number of cysteinyl residues, but only four of these are conserved in all nickel-containing hydrogenases; they are present as Cys-X-X-Cys motifs at the amino and carboxyl ends of the protein (Fig. 7, regions 2 and 4). One of these conserved cysteines at the carboxyl end is substituted by a TGA codon (Fig. 7, region 4) in the [NiFeSe] hydrogenase from *D. baculatus* (39) and a membrane-bound [NiFeSe] hydrogenase from *D. vulgaris* (E. S. Choi, unpublished results). The TGA codon has been shown

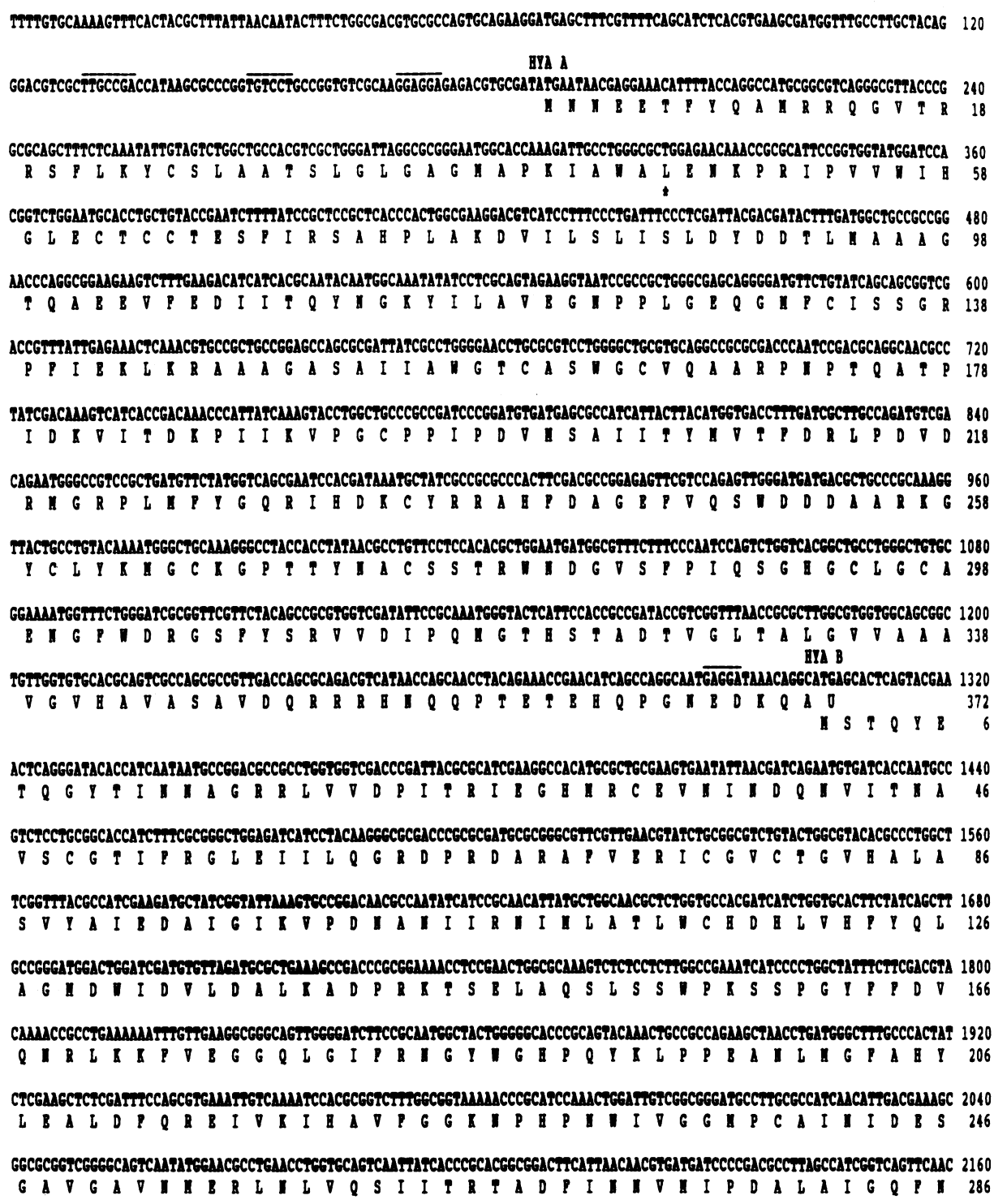
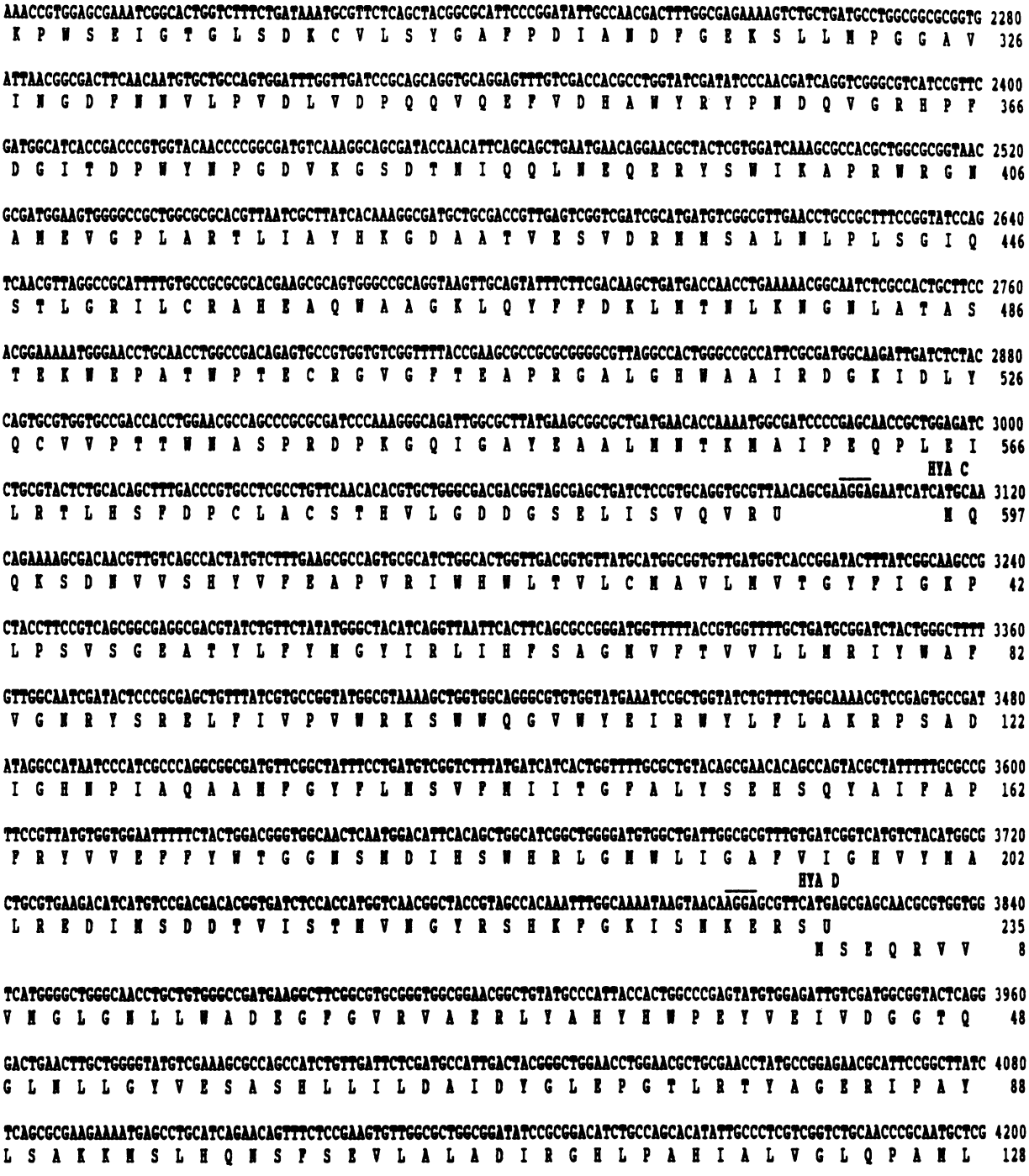


FIG. 5. Nucleotide and derived amino acid sequences of the *hya* operon. Possible promoter sites (−10 and −35) as identified by a computer search have been overlined. The asterisk indicates first residue of the mature small subunit of hydrogenase 1 (HyaA) as identified by amino-terminal sequencing; the dotted line signifies an inverted repeat that could function as a transcription terminator. The six polypeptides derived from ORFs 1 through 6 are designated HYA A through F, respectively.

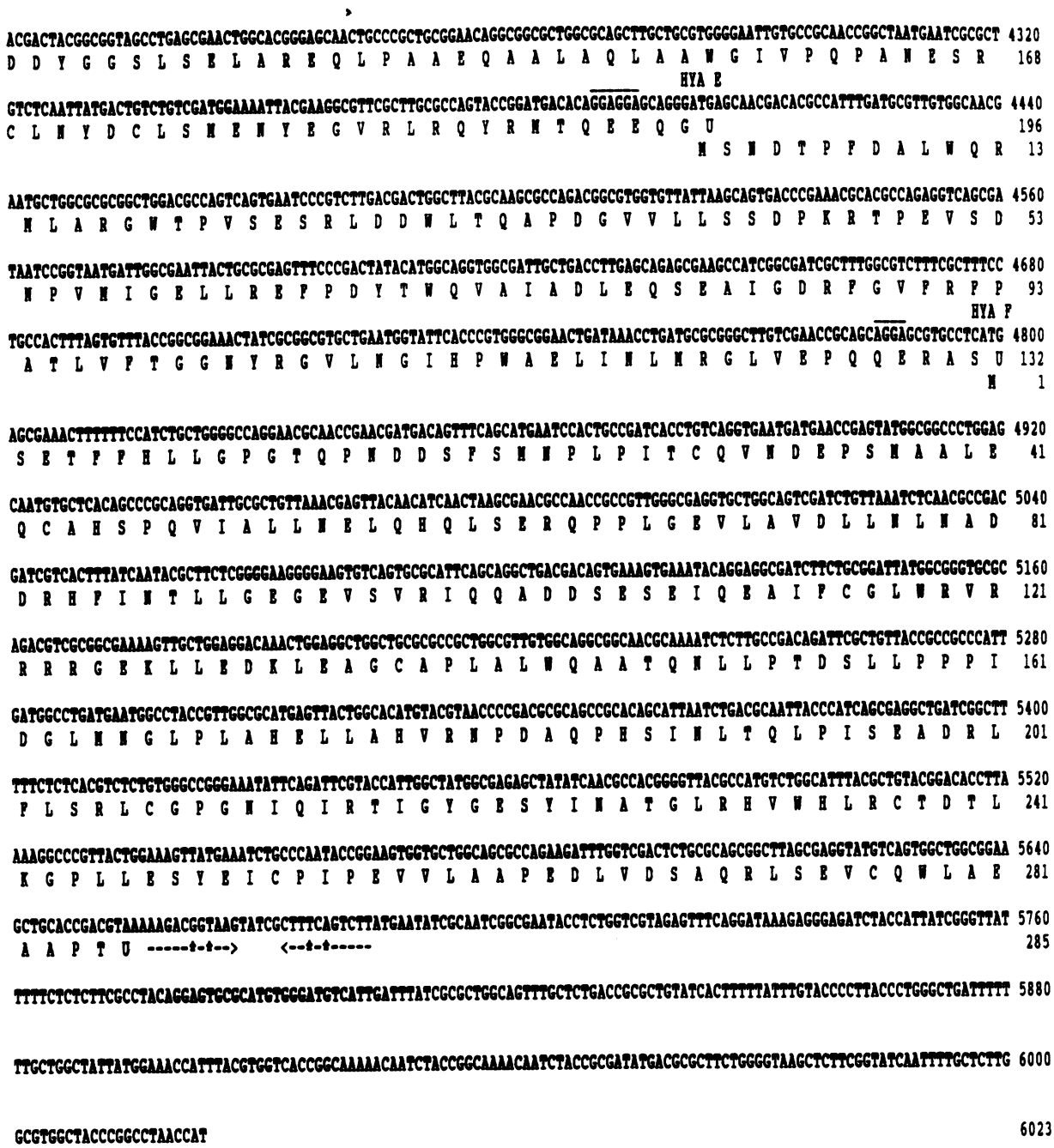
to encode the incorporation of selenocysteine into proteins, as in the case of *E. coli* formate dehydrogenase and mouse glutathione peroxidase (7, 44). Extended X-ray absorption fine structure (EXAFS) and electron paramagnetic resonance data obtained with ^{77}Se-enriched [NiFeSe] hydrogenase of *D. baculatus* indicate that selenium acts as a ligand to nickel (9, 12). This result might imply that analogous cysteinyl residues in [NiFe] hydrogenases (Fig. 7, region 4) serve as ligands to nickel.

EXAFS data on the [NiFeSe] hydrogenase from *D. bac-*



ulatus and studies on the incorporation of ^{33}S into the [NiFe] hydrogenase of *Wolinella succinogenes* indicate that there are two sulfur ligands to nickel, which raises the possibility of the involvement of the second of the paired cysteinyl residues (Fig. 7, region 4) in nickel liganding (1, 42). EXAFS data have also indicated that, in addition to the sulfur ligands, there are three or four ligands involving either nitrogen or oxygen at the nickel-binding site of the [NiFe] hydrogenase of *D. gigas* and the [NiFeSe] hydrogenase of *D. baculatus* (9). The presence of these additional ligands suggests that there might be other highly conserved domains involved in the formation of nickel ligands. Two additional conserved domains which could be the source of oxygen and/or nitrogen ligands have been identified in the large subunit of the [NiFe] and [NiFeSe] hydrogenases (Fig. 7, regions 1 and 3).

Additional ORFs in the *hya* operon. The gene encoding the large subunit is immediately followed by a ribosome-binding site and an ATG initiator codon, indicating the presence of a third ORF in the putative operon. This arrangement is repeated for three additional ORFs. A codon usage analysis supports the presence of four potential ORFs in the *hya* operon in addition to the genes encoding the two structural subunits. Similarly, the presence of additional ORFs in the



hydrogenase operon of *B. japonicum* is indicated by the presence of a ribosome-binding site and an ATG codon immediately downstream of the large subunit gene (32). Consistent with this possibility, the partial nucleotide sequence of the putative third ORF of the *B. japonicum* hydrogenase operon would encode a polypeptide with 50% homology to the polypeptide encoded by the third ORF in the *E. coli hya* operon (Fig. 8). This level of homology would be highly unlikely in a nongenic region.

The derived amino acid sequence of the third ORF from *E. coli* indicates that its product, HyaC, is very hydrophobic, is rich in aromatic residues, and has four putative hydrophobic membrane-spanning regions (Fig. 9). Its general topological homology to bacteriorhodopsin (17) is striking (data not shown) and suggests there may be as many as six membrane-spanning helices. Recently, the dimethyl sulfoxide reductase (*dms*) operon from *E. coli* was fully sequenced (4); and one of the ORFs in the operon (*dmsC*) codes for a protein with eight hydrophobic helices. It has been speculated that this protein may play a role in anchoring of the structural subunits to the membrane, proton translocation, or electron transfer. In *Alcaligenes eutrophus* the regulatory and structural genes for the hydrogenase system are found on a megaplasmid. One of the loci necessary for hydrogenase activity (HoxM), is required for anchoring of the [NiFe] hydrogenase to the membrane. This region lies immediately downstream of the genes encoding the structural subunits of this membrane-bound nickel-containing hydrogenase and might be analogous to HyaC of *E. coli* (13).

The predicted polypeptides from ORFs 4, 5, and 6 were

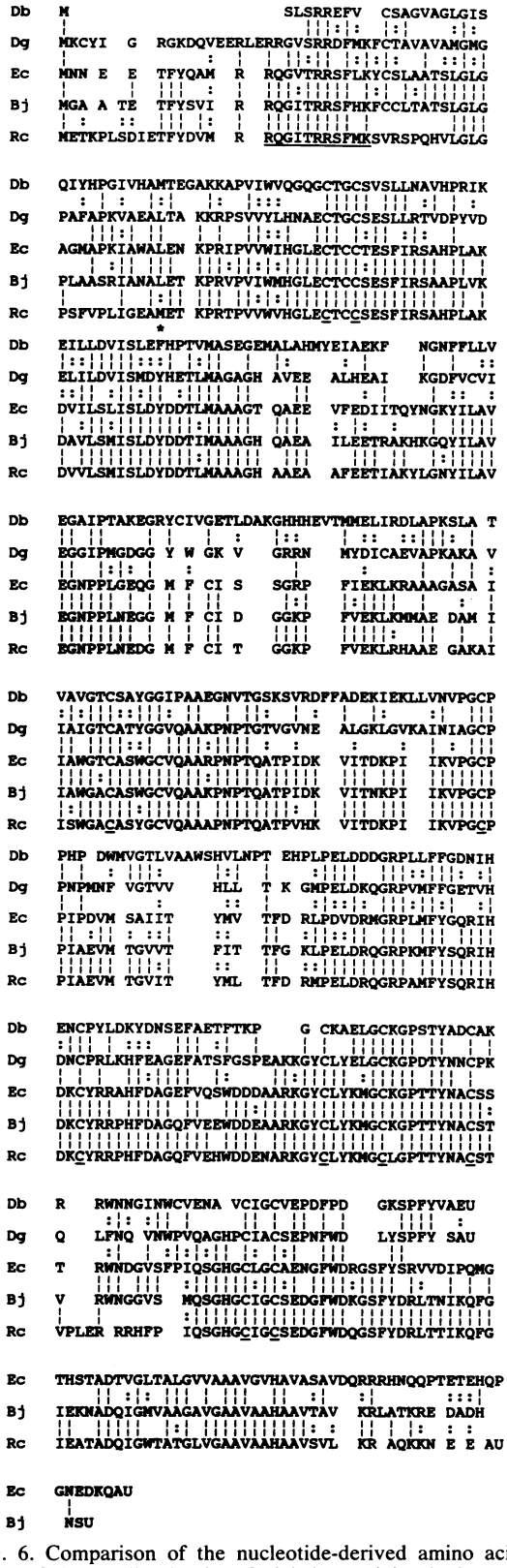


FIG. 6. Comparison of the nucleotide-derived amino acid sequences of the small subunits of nickel-containing hydrogenases. Abbreviations: Db, *D. baculatus*; Dg, *D. gigas*; Ec, *E. coli*; Bj, *B. japonicum*; Rc, *R. capsulatus*. The asterisk marks the first residue of the mature small subunits as determined by amino-terminal protein sequencing. A conserved region found in all hydrogenase signal peptides is underlined. Conserved cysteine residues are underlined. Vertical lines ( ) indicate identical amino acids, whereas colons (:) indicate neutral changes.

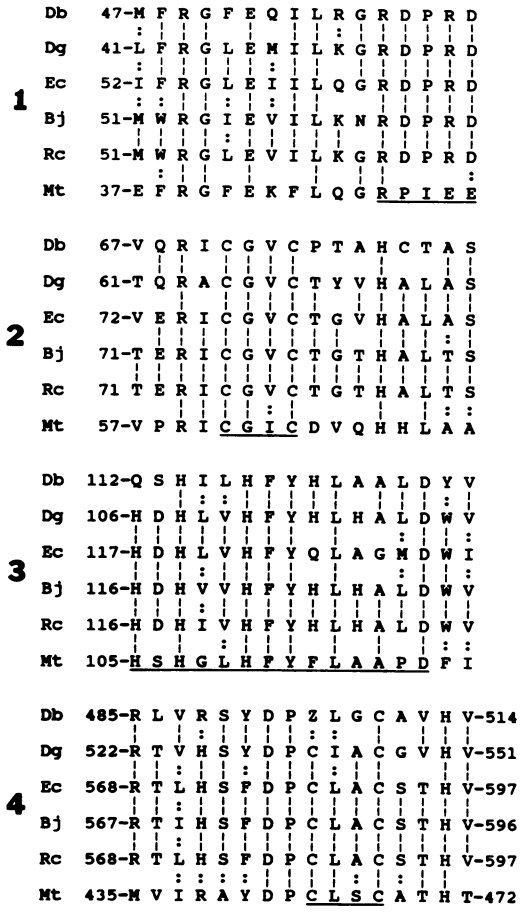


FIG. 7. Comparison of four highly conserved regions (numbered 1 through 4) of the large subunits of nickel-containing hydrogenases. Abbreviations indicate organisms as in Fig. 6 with the following addition: Mt, *M. thermoautotrophicum*. Numbers at the beginning of each fragment indicate their positions in the complete polypeptide. The most conserved portion of each region is underlined, and each is discussed in the text. The residue marked Z at position 495 in the *D. baculatus* sequence is selenocysteine, which is coded for by the TGA triplet.

compared with existing proteins in the Protein Identification Resource Sequence data bank, The University of Geneva Sequence data base, and the N.B.R.F. protein data bank. The search did not reveal homologous proteins, and at present we do not know the role played by these gene products in the processing and activation of hydrogenase. Although the operons for the [Fe], [NiFe], and [NiFeSe] hydrogenases (21, 23, 38, 40) of the sulfate-reducing bacteria contain only the genes encoding the structural subunits, hydrogenase operons from other systems may include genes that code for additional proteins. The operon for the [NiFe] hydrogenase of *M. thermoautotrophicum* has been shown to encode two additional proteins, one of which appears to be a unique polyferredoxin (24). The regulatory system for the [NiFe] hydrogenases in *A. eutrophus* has also been shown to involve more than one operon (B. Friedrich, Proc. Int.

```
  1   M M D A V A P A S D A R P D L A S R D A
      |           | |                   |
  1   M         Q Q K S D       N V V S

 21   S G E R A V G R P T V Y V Y E A P V R I
                  :             | | : | | | | |
 11                H           Y V F E A P V R I

 41   C H W V N A F S I I V L M V T G Y L I G
        | | :               :   | | | | | | |   | |
 21   W H W L T V L C M A V L M V T G Y F I G

 61   T P L P T V A G E A S D N F V M G Y I R
        | | | : | : | | | :       |   | | | | |
 41   K P L P S V S G E A T Y L F Y M G Y I R

 81   F A H F A A G Q V L A V F F L T R I L W
        | | : | |   |   | :   |     |   | |   |
 61   L I H F S A G M V F T V V L L M R I Y W

101   A F V G N H H S R Q I
      | | | | | : | | :
 81   A F V G N R Y S R E L
```

FIG. 8. Comparison of the nucleotide-derived amino acid sequence of *E. coli* HyaC to the amino acid sequence derived from the nucleotide sequence 3′ to the hydrogenase large subunit gene of *B. japonicum* (see Results and Discussion).

Hydrogenase Symp., 1988). It has been shown that the NAD^+-reducing [NiFe] hydrogenase system encodes the structural proteins plus one additional protein and that the membrane-bound [NiFe] hydrogenase system encodes five new peptides of unknown function (15).

The role of *E. coli* hydrogenase 1 is perhaps the least understood among the three hydrogenases in *E. coli*. Studies on its expression and regulation are complicated by the facts that hydrogenase 1 accounts for less than 10% of the total hydrogenase activity and that expression of the different hydrogenases is in many cases coordinately regulated, making physiological studies difficult. To better understand the role of this enzyme in *E. coli* anaerobic metabolism, we constructed deletion mutants encompassing the complete *hya* operon as well as individual structural genes. Complementation of these mutants with sequential deletions in the *hya* operon showed that at least five of the ORFs are essential to produce active hydrogenase 1 (about 30% of the wild type activity was recovered when mutants were complemented with *hyaA* through *E*). Restoration of wild-type hydrogenase levels occurred only when mutants were complemented with all six ORFs. Preliminary mapping studies revealed that the *hya* operon is located at 21 min, hence defining a new locus involved in hydrogenase metabolism in *E. coli*. All of these results will be reported in detail in the future (N. K. Menon, K. T. Shanmugam, J. Wendl, J. Robbins, H. D. Peck, Jr., and A. E. Przybyla, manuscript in preparation).

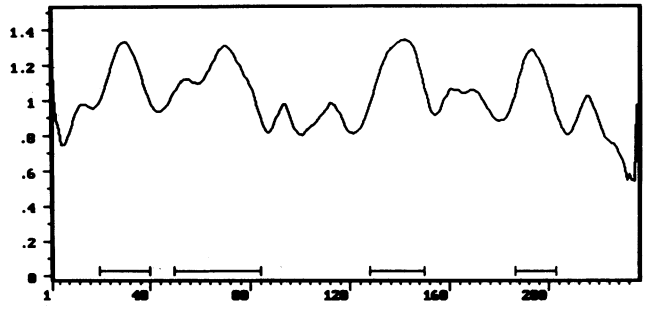


FIG. 9. Plot of the buried helix profile for HyaC. Probable membrane-spanning regions have been marked. The analysis was done on the PC/Gene Raoargos program. The *x* axis represents amino acid residue number, and the *y* axis represents the hydrophobicity index.

ACKNOWLEDGMENTS

We thank Sidney Kushner for an EMBL4 library of *E. coli* genomic DNA, K. T. Shanmugam for sequences of the amino termini of the large and small subunits of *E. coli* hydrogenase 1, H. Evans for allowing us to analyze the nucleotide sequence 3′ to the *B. japonicum* hydrogenase genes, and R. Robson for helpful discussions and advice.

This work was supported in part by Public Health Service grant GM 34903 from the National Institutes of Health and by National Science Foundation grant DMB-87-06205 to A.E.P. and H.D.P. and by Department of Energy grant ER10489 to H.D.P.

LITERATURE CITED

1. **Albracht, S. P. J., A. Kröger, J. W. van der Zwaan, G. Unden, R. Böcher, H. Mell, and R. D. Fontijn.** 1986. Direct evidence for sulphur as a ligand to nickel in hydrogenase: an EPR study of the enzyme from *Wolinella succinogenes* enriched in ^{33}S. Biochim. Biophys. Acta **874**:116–127.
2. **Ballantine, S. P., and D. H. Boxer.** 1985. Nickel-containing hydrogenase isoenzymes from anaerobically grown *Escherichia coli* K-12. J. Bacteriol. **163**:454–459.
3. **Ballantine, S. P., and D. H. Boxer.** 1986. Isolation and characterization of a soluble active fragment of hydrogenase isoenzyme 2 from membranes of anaerobically grown *Escherichia coli*. Eur. J. Biochem. **156**:277–284.
4. **Bilous, P. T., S. T. Cole, W. F. Anderson, and J. H. Weiner.** 1988. Nucleotide sequence of the Dms ABC operon encoding anaerobic dimethyl sulphoxide reductase of *E. coli*. Mol. Microbiol. **2**:785–795.
5. **Birkmann, A., and A. Böck.** 1989. Characterization of a cis regulatory DNA element necessary for formate induction of the formate dehydrogenase gene (*fdhF*) of *E. coli*. Mol. Microbiol. **3**:187–195.
6. **Birkmann, A., R. G. Sawers, and A. Böck.** 1987. Involvement of the *ntrA* gene product in the anaerobic metabolism of *E. coli*. Mol. Gen. Genet. **210**:535–542.
7. **Chambers, I., J. Frampton, P. Goldfarb, N. Affara, W. McBain, and P. R. Harrison.** 1986. The structure of the mouse glutathione peroxidase gene: the selenocysteine in the active site is encoded by the termination codon, TGA. EMBO J. **5**:1221–1227.
8. **Chaudhuri, A., and A. I. Krasna.** 1987. Isolation of genes required for hydrogenase synthesis in *Escherichia coli*. J. Gen. Microbiol. **5**:1221–1227.
9. **Eidsness, M. K., R. A. Scott, B. Prickrill, D. V. DerVartanian, J. LeGall, I. Moura, J. J. G. Moura, and H. D. Peck, Jr.** 1989. Evidence for selenocysteine coordination to the active site nickel in the (NiFeSe) hydrogenase from *Desulfovibrio baculatus*. Proc. Natl. Acad. Sci. USA **86**:147–151.
10. **Fauque, G., H. D. Peck, Jr., J. J. G. Moura, B. H. Huynh, Y. Berlier, M. Teixeira, A. E. Przybyla, P. A. Lespinat, D. V. DerVartanian, I. Moura, and J. LeGall.** 1989. The three classes of hydrogenases from sulfate reducing bacteria of the genus *Desulfovibrio*. FEMS Microbiol. Rev. **54**:299–344.
11. **Frischamp, A. M., H. Garoff, and H. Lehrach.** 1980. A subcloning strategy for DNA sequence analysis. Nucleic Acids Res. **8**:5541.
12. **He, S. H., M. Teixeira, J. LeGall, D. S. Patil, D. V. DerVartanian, B. H. Huynh, and H. D. Peck, Jr.** 1989. Studies with ^{77}Se enriched (NiFeSe) hydrogenase of *Desulfovibrio baculatus*. Evidence for a selenium ligand to the active-site nickel. J. Biol. Chem. **86**:147–157.
13. **Jamieson, D. J., and C. F. Higgins.** 1986. Two genetically distinct pathways for transcriptional regulation of anaerobic

gene expression in *Salmonella typhimurium*. J. Bacteriol. **168:** 405–411.

14. **Jamieson, D. J., R. G. Sawers, P. A. Rugman, D. H. Boxer, and C. F. Higgins.** 1986. Effects of anaerobic regulatory mutations and catabolite repression on regulation of hydrogen metabolism and hydrogen isoenzyme composition in *Salmonella typhimurium*. J. Bacteriol. **168:**405–411.
15. **Kärst, V., and C. G. Friedrich.** 1987. Identification of new peptides synthesized under the control system of *Alcaligenes eutrophus*. Arch. Microbiol. **147:**346–353.
16. **Karube, I., M. Tomiyama, and A. Kikuchi.** 1984. Molecular cloning and physical mapping of the *hyd* gene of *Escherichia coli* K12. FEMS Microbiol. Lett. **25:**165–168.
17. **Khorana, H. G.** 1988. Bacteriorhodopsin, a membrane protein that uses light to translocate protons. J. Biol. Chem. **263:** 7439–7442.
18. **Krasna, A. I.** 1984. Mutants of *Escherichia coli* with altered hydrogenase activity. J. Gen. Microbiol. **130:**770–787.
19. **Leclerc, M., A. Colbeau, B. Cauvin, and P. M. Vignais.** 1988. Cloning and sequencing of the genes encoding the large and small subunits of the H_2 uptake hydrogenase (hup) of *Rhodobacter capsulatus*. Mol. Gen. Genet. **214:**97–107.
20. **Lee, J. H., P. Patel, P. Sankar, and K. T. Shanmugam.** 1985. Isolation and characterization of mutant strains of *Escherichia coli* altered in hydrogen metabolism. J. Bacteriol. **162:**344–352.
21. **Li, C., H. D. Peck, Jr., J. LeGall, and A. E. Przybyla.** 1987. Cloning characterization and sequencing of the genes encoding the large and small subunits of the periplasmic (NiFe) hydrogenase of *Desulfovibrio gigas*. DNA **6:**539–551.
22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Menon, N. K., H. D. Peck, Jr., J. LeGall, and A. E. Przybyla.** 1987. The cloning and sequencing of the genes encoding the large and small subunits of the periplasmic (NiFeSe) hydrogenase of *Desulfovibrio baculatus*. J. Bacteriol. **170:**5401–5407. (Author's correction **170:**4429, 1988.)
24. **Reeve, J. N., G. S. Beckler, D. S. Cram, P. T. Hamilton, J. W. Brown, J. A. Krzycki, A. F. Kolodziej, L. Alex, W. H. Orme-Johnson, and C. T. Walsh.** 1989. A hydrogenase linked gene in *Methanobacterium thermoautotrophicum* (ΔH) encodes a polyferredoxin. Proc. Natl. Acad. Sci. USA **86:**3031–3035.
25. **Reznikoff, W., and W. R. McClure.** 1986. *E. coli* promoters, p. 2–29. *In* W. Reznikoff and L. Gold (ed.), Maximizing gene expression. Butterworth Publishers, Inc., Woburn, Mass.
26. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain termination inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
27. **Sankar, P., J. H. Lee, and K. T. Shanmugam.** 1985. Cloning of hydrogenase genes and fine structure analysis of an operon essential for H_2 metabolism in *Escherichia coli*. J. Bacteriol. **162:**353–360.
28. **Sankar, P., and K. T. Shanmugam.** 1988. Hydrogen metabolism in *Escherichia coli*: biochemical and genetic evidence for a *hydF* gene. J. Bacteriol. **170:**5446–5451.
29. **Sankar, P., and K. T. Shanmugam.** 1988. Biochemical and genetic analysis of hydrogen metabolism in *Escherichia coli*: the *hydB* gene. J. Bacteriol. **170:**5433–5439.
30. **Sawers, R. C., S. P. Ballantine, and D. H. Boxer.** 1985. Differential expression of hydrogenase isoenzymes in *Escherichia coli* K-12. Evidence for a third isoenzyme. J. Bacteriol. **164:**1324–1331.
31. **Sawers, R. G., and D. H. Boxer.** 1986. Purification and properties of membrane-bound hydrogenase isoenzyme 1 from anaerobically grown *Escherichia coli* K12. Eur. J. Biochem. **156:** 265–275.
32. **Sayavedra-Soto, L. A., G. K. Powell, H. J. Evans, and R. O. Morris.** 1988. Nucleotide sequence of the genetic loci encoding the subunits of *Bradyrhizobium japonicum* uptake hydrogenase. Proc. Natl. Acad. Sci. USA **85:**8395–8399.
33. **Staden, R., and A. D. McLachlan.** 1982. Codon preference and its use in identifying protein coding regions in long DNA sequences. Nucleic Acids Res. **10:**141–156.
34. **Stoker, K., L. F. Oltmann, and A. H. Stouthammer.** 1988. Partial characterization of an electrophoretically labile hydrogenase activity of *Escherichia coli* K-12. J. Bacteriol. **170:**1220–1226.
35. **Stoker, K., L. F. Oltmann, and A. H. Stouthammer.** 1989. Randomly induced *Escherichia coli* K-12 Tn*5* insertion mutants defective in hydrogenase activity. J. Bacteriol. **171:**831–836.
36. **Tibelius, K. H., R. L. Robson, and M. G. Yates.** 1987. Cloning and characterization of hydrogenase genes from *Azotobacter chroococcum*. Mol. Gen. Genet. **206:**285–290.
37. **Von Heijne, G.** 1985. Signal sequences: the limits of variation. J. Mol. Biol. **184:**99–105.
38. **Voordouw, G., and S. Brenner.** 1985. Nucleotide sequence of the gene encoding the hydrogenase from *Desulfovibrio vulgaris* (Hildenborough). Eur. J. Biochem. **159:**347–351.
39. **Voordouw, G., N. K. Menon, J. LeGall, E.-S. Choi, H. D. Peck, Jr., and A. E. Przybyla.** 1989. Analysis and comparison of the nucleotide sequences encoding the (NiFe) and (NiFeSe) hydrogenases from *Desulfovibrio gigas* and *Desulfovibrio baculatus*. J. Bacteriol. **171:**2894–2899.
40. **Voordouw, G., J. D. Strang, and F. R. Wilson.** 1989. Organization of the genes encoding [Fe] hydrogenases in *Desulfovibrio vulgaris* subsp. *oxamicus* Monticello. J. Bacteriol. **171:**3881–3889.
41. **Wackett, L. P., E. A. Hartwieg, J. A. King, W. H. Orme-Johnson, and C. T. Walsh.** 1987. Electron microscopy of nickel-containing methanogenic enzymes: methyl reductase and F420 reducing hydrogenase. J. Bacteriol. **169:**718–727.
42. **Waugh, R., and D. H. Boxer.** 1986. Pleiotropic hydrogenase mutants of *Escherichia coli* K12: growth in the presence of nickel can restore hydrogenase activity. Biochimie **68:**157–166.
43. **Wu, L. F., and M. A. Mandrand-Berthelot.** 1986. Genetic and physiological characterization of new *Escherichia coli* mutants impaired in hydrogenase activity. Biochimie **68:**167–179.
44. **Zinoni, F., A. Birkman, T. C. Stadtman, and A. Böck.** 1986. Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase linked) from *Escherichia coli*. Proc. Natl. Acad. Sci. USA **83:**4650–4654.