JOURNAL OF BACTERIOLOGY, Aug. 1989, p. 4448–4456
0021-9193/89/084448-09$02.00/0

Vol. 171, No. 8

Initial Cloning and Sequencing of *hydHG*, an Operon Homologous to *ntrBC* and Regulating the Labile Hydrogenase Activity in *Escherichia coli* K-12

KAREL STOKER,* WILLEM N. M. REIJNDERS, L. FRED OLTMANN, AND ADRIAAN H. STOUTHAMER

Department of Microbiology, Biological Laboratory, Vrije Universiteit, P.O. Box 7161, 1007 MC Amsterdam, The Netherlands

Received 7 December 1988/Accepted 11 May 1989

To isolate genes from *Escherichia coli* which regulate the labile hydrogenase activity, a plasmid library was used to transform *hydL* mutants lacking the labile hydrogenase. A single type of gene, designated *hydG*, was isolated. This gene also partially restored the hydrogenase activity in *hydF* mutants (which are defective in all hydrogenase isoenzymes), although the low hydrogenase 1 and 2 levels were not induced. Therefore, *hydG* apparently regulates, specifically, the labile hydrogenase activity. Restoration of this latter activity in *hydF* mutants was accompanied by a proportional increase of the H_2 uptake activity, suggesting a functional relationship. H_2:fumarate oxidoreductase activity was not restored in complemented *hydL* mutants. These latter strains may therefore lack, in addition to the labile hydrogenase, a second component (provisionally designated component R), possibly an electron carrier coupling H_2 oxidation to the anaerobic respiratory chain. Sequence analysis showed an open reading frame of 1,314 base pairs for *hydG*. It was preceded by a ribosome-binding site but apparently lacked a promoter. Minicell experiments revealed a single polypeptide of approximately 50 kilodaltons. Comparison of the predicted amino acid sequence with a protein sequence data base revealed strong homology to NtrC from *Klebsiella pneumoniae*, a DNA-binding transcriptional activator. The 411 base pairs upstream from pHG40 contained a second open reading frame overlapping *hydG* by four bases. The deduced amino acid sequence showed considerable homology with the C-terminal part of NtrB. This sequence was therefore assumed to be part of a second gene, encoding the NtrB-like component, and was designated *hydH*. The labile hydrogenase activity in *E. coli* is apparently regulated by a multicomponent system analogous to the NtrB-NtrC system. This conclusion is in agreement with the results of Birkmann et al. (A. Birkmann, R. G. Sawers, and A. Böck, Mol. Gen. Genet. 210:535–542, 1987), who demonstrated *ntrA* dependence for the labile hydrogenase activity.

Hydrogen metabolism in *Escherichia coli* is characterized by a fermentative formate hydrogenlyase-linked H_2 production, via formate dehyrogenase and hydrogenase (29), and by an H_2 uptake route (H_2:fumarate oxidoreductase [Hup]) linking the oxidation of H_2 to an anaerobic respiratory chain, leading to fumarate (1, 4, 18, 46). The hydrogenases involved are not identical and probably not interchangeable. *E. coli* contains three hydrogenases, two electrophoretically stable isoenzymes (2), i.e., hydrogenase 1 (36) and hydrogenase 2 (3), and a considerable residual activity, which is inactivated upon neutral, nondenaturing gel electrophoresis (35). The current state of literature is not unambiguous about their physiological function. Some studies suggest that the labile species is involved in formate hydrogenlyase activity (hydrogenase 3) (5, 6, 35), whereas we provided evidence that the labile species (previously designated hydrogenase L, but probably identical to hydrogenase 3) might function in the respiratory Hup pathway (41, 42). Here we provide further evidence for the latter model.

In addition to anaerobiosis, substrates like formate, and general regulators like Fnr and sigma factor NtrA (5, 6, 40), several genes are involved in the regulation of hydrogen metabolism. The majority of these genes or loci are pleiotropic in their effect, in that they abolish hydrogenase as well as formate dehydrogenase-H activity simultaneously, e.g., *hydA* (17, 27), *hydB* (17), *hydE* (9), *hydFD12* (13, 44), *hydC* or *hydD* (45). The *hydF* mutation (42) is hydrogenase specific but pleiotropic with respect to the individual isoenzymes. Only two loci seem to control one hydrogenase isoenzyme, viz., the labile one, *hyd-17* (28, 41) and *hydL* (41).

Some (regulatory) genes such as *hydA* (16, 34), *hydB* (9, 34, 44), and *hydE* (9) have been cloned. This approach has been shown to be fruitful in the study of other gene clusters, e.g., the *nif* (14), *nar* (40), or *frd* (15) operons.

In this paper, we report the cloning, by complementation of a *hydL* mutant, of (a part of) an operon that specifically restores the labile hydrogenase activity in several *E. coli* mutants. The upstream gene (*hydH*) was partially cloned, and the second gene (*hydG*) was completely cloned. The complementation behavior in several mutants, the coding region, the product (HydG), the nucleotide sequence, and the homologies with other regulatory genes are described.

MATERIALS AND METHODS

Bacterial strains and vectors. For transformation with pUC plasmids and M13 derivatives, *E. coli* JM103 [F′ *proA*$^+$*B*$^+$ *lacI*^{q} *lacZ*Δ*M15* Δ(*lac pro*) *thi endA sbrB supE hsdR4 rpsL*] was used. For cloning purposes, *E. coli* HB101 (F$^-$ *pro leu thi lacZ gal xyl ara mtl hsdS phx supE recA13 rpsL*) was also used. For the detection of plasmid-encoded proteins, we used the minicell-producing strain DS410 (F$^-$ *thr leu minA minB rpsL*) (10). Hydrogenase-deficient strains were HB101 derivatives BO2 (*hydL*) (41), BO32 (*hydF*), BO26 (*hydC*), and BO5 (*hyd*, near 58 min) (42). (Sub)cloning was in the *lacZ′* region of plasmid pUC13 (43). For sequencing, DNA

* Corresponding author.

fragments were cloned in the M13mp7 (20) derivatives M13mp18 and M13mp19 (24).

Media and growth conditions. pUC13-transformed cells were plated on YT/X-gal agar containing (per liter) 10 g of trypton, 5 g of yeast extract, 8 g of NaCl, and 15 g of agar and supplied with 0.03 mg of X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside) per ml and 0.06 mM isopropyl-β-D-thiogalactopyranoside. For scoring M13 plaques, a YT/X-gal top layer containing 6 g of agar per liter, 0.2 mg of X-gal per ml, and 0.2 mM isopropyl-β-D-thiogalactopyranoside per liter was used. In all other cases cells were grown on plates or in (an)aerobic batch cultures in brain heart infusion broth (38 g/liter, pH 7.4) as previously described (41). Cells to be rendered transformation competent were grown in YT.

MV filter assay. Colonies were screened for hydrogenase activity by transferring them to a filter paper drenched in 20 mM methylviologen (MV) (E_0' = −440 mV)–10 mM Tris hydrochloride (pH 7.4) and placing this in an atmosphere of 5% H_2–95% N_2 (41). Positive colonies turned blue within 1 to 2 min. Mutants remained colorless and only after 10 to 30 min could some coloration be observed.

Cell preparations and extracts. Whole-cell samples, solubilized extracts, and cell-free solubilized (S100) extracts were prepared as previously described (41).

Assays. Hydrogenase activity was determined by manometric (Warburg) methods or by microtitration. Hydrogenases 1 and 2 were demonstrated by electrophoresis of cell-free extracts on activity-stained neutral, nondissociating 7.5% polyacrylamide gels. Hup, formate hydrogenlyase, benzylviologen (red)-dependent fumarate reductase, formate dehydrogenase-H, and formate:nitrate oxidoreductase activities were determined by manometrical (Warburg) methods. All assays were carried out as described previously (41).

DNA techniques. General cloning techniques were essentially as previously described (19). DNA fragments were electroeluted on a LKB 2014 Extraphor (LKB Products, Sweden) by using a salt bridge. Restriction fragments with incompatible ends were rendered blunt ended by treatment with T4 DNA polymerase in the presence of deoxynucleoside triphosphates. This method was also used to introduce 1- or 2-base-pair (bp) deletions or insertions. Southern blotting was performed by capillary transfer of electrophoresed and denatured DNA to GeneScreen Plus hybridization membranes (Dupont, NEN Research Products, Boston, Mass.) (10 to 25 μg per lane). This procedure was followed by overnight hybridization to a [^{35}S]ATP-labeled probe in 1% (wt/vol) sodium dodecyl sulfate–1 M NaCl–10% (wt/vol) dextran sulfate at 65°C, exactly as described by the manufacturer.

Isolation of chromosomal DNA. Chromosomal DNA was isolated by suspending the cells from a 20-ml overnight culture in 30 ml of 25% (wt/vol) sucrose–50 mM Tris hydrochloride (pH 8)–2 mg of lysozyme per ml and incubating for 30 min at 37°C. After addition of EDTA (50 mM), sodium dodecyl sulfate (1%, wt/vol), and pronase (0.1 mg/ml), the mixture was left for another 30 min, after which the DNA was phenol extracted.

Construction of a plasmid library. Chromosomal DNA (250 μg) was partially digested with restriction endonuclease *Sau*3A, and the fragments were separated on a 4-ml sucrose gradient containing 0.1 mM NaCl–10 mM Tris hydrochloride (pH 7.5)–1 mM EDTA–0.1% (wt/vol) sarcosyl and 5 to 30% (wt/vol) sucrose. Centrifugation was at 22,000 × *g*, for 16 h at 4°C. Typically, from the sucrose fraction containing DNA fragments between 2 and 3 kilobases (kb), 10 μl was added to 0.1 μg of *Bam*HI-digested and dephosphorylated pUC13 DNA, and the DNA was ligated in a total volume of 50 μl. After transformation and plating on YT/X-gal plates with ampicillin, 10,000 white colonies were selected, separately grown, and pooled, after which the plasmid DNA content was isolated and stored in aliquots at −20°C.

Labeling of minicells. Minicells were isolated as previously described (22) and were labeled with [^{35}S]methionine (50 μCi/ml of cells).

DNA sequencing. For sequencing purposes, DNA fragments were cloned in derivatives of bacteriophage M13 (24, 33), after which the single-stranded copies were used as templates in a Klenow polymerase reaction in the presence of [^{35}S]ATP, deoxynucleotides, and dideoxynucleotides, as described by Sanger et al. (32). The latter reaction started from the M13 linker region, with the universal primer, or was primed by internally annealing 18-mer oligonucleotides synthesized on a 381A DNA synthesizer (Applied Biosystems, Foster City, Calif.).

Nucleotide and amino acid sequence analysis. DNA and protein sequence analyses were performed with the aid of the PC/Gene computer program (Genofit, Geneva, Switzerland). Comparisons were made with sequences from the Protein Identification Resource data bank, release 30 (June 1987), supported by the Division of Research of the National Institutes of Health. For the determination of functional amino acid homologies, a division into four groups was maintained by the method of Nixon et al. (23), i.e., group 1 (L, I, V, M, F, Y, W), group 2 (P, T, A, G, S, C), group 3 (Q, N, E, D), and group 4 (H, K, R).

RESULTS

Isolation of *hydL*-complementing clones. In order to isolate an *E. coli* DNA fragment that restores the electrophoretically labile hydrogenase activity, a mutant specifically impaired in this function (strain BO2, *hydL*) (41) was transformed by an *E. coli* genomic plasmid library, and the transformants were tested in the MV filter colony assay. The *hydL* mutants are lacking the labile hydrogenase activity, which accounts under these semifermentative conditions for about 80% of the total activity, and the colonies of these strains remained essentially colorless in this test. Of 10,000 transformants screened, three were complemented. The recombinant plasmids isolated from these strains were pHG10, pHG20, and pHG30.

Hydrogenase activities in pHG10-transformed mutants. The degree of complementation was determined quantitatively (Table 1). The hydrogenase activity in *hydL* strains was restored to parental level, when measured in whole-cell preparations, after anaerobic growth. Interestingly, when the cells were opened by sonication, the activity in the transformed cells was even two- to threefold higher than in the parent strain. (This might be explained by assuming that there is only a limited number of sites in the membrane where this hydrogenase can be incorporated, and it is likely that in a whole-cell assay, it is mainly the activity of this particulate fraction that is determined). The biological activity regulated by pHG10 was not influenced by the isopropyl-β-D-thiogalactopyranoside concentration, a gratuitous inducer for *lacZp*-dependent genes. No hydrogenase activity was found in transformants grown aerobically or in the presence of nitrate (40 mM).

Subsequently, pHG10 was transferred to the *hydF*, *hydC*, and *hyd* mutants which were recently described (42). As a result, an identical restoration of hydrogenase activity was observed in the *hydF* strains, which are genetically closely

TABLE 1. Some physiological data on HB101, hydrogenase mutants, and pHG10-transformed mutants[a]

Enzyme	Activity in strain:				
	BO32 (*hydF*)		BO2 (*hydL*)		HB101
	−pHG10	+pHG10	−pHG10	+pHG10	
Hydrogenase[b]					
Whole cells	0.9	5.7	0.9	5.8	5.6
Open cells	2.1	35.2	2.3	37.0	12.5
Polyacrylamide gel electrophoresis	0.1	0.1	100	100	100
Hup[c]	2.1	11	0.01	0.01	13
Formate hydrogenlyase[d]	6.5	6.6	6.6	6.4	6.5

[a] Cells were plate grown. Determinations were as previously described (41).

[b] H_2:methylviologen oxidoreductase activity was determined manometrically. Units represent nanomoles of H_2 consumed per minute per milligram of protein. Hydrogenase (1 and 2) activity is given relative to that of parent HB101; read off from activity-stained neutral gels (polyacrylamide gel electrophoresis, %).

[c] Determined manometrically on whole cells. Units represent nanomoles of H_2 consumed per minute per milligram of protein.

[d] Determined manometrically on whole cells. Units represent nanomoles of H_2 plus CO_2 produced per minute per milligram of protein.

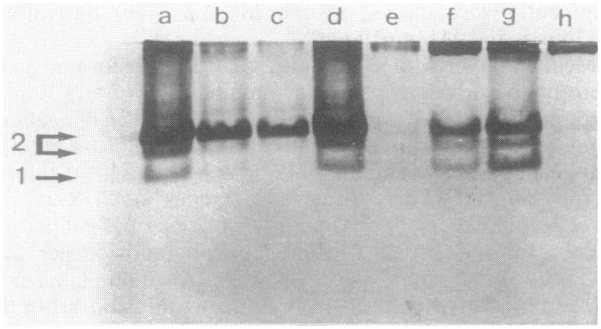

FIG. 1. Complementation of a *hydF* strain (BO32) by F′ 116 or pHG10. A neutral, nondenaturing polyacrylamide gel, stained for hydrogenase activity, on which cell-free extracts (S100) were run is shown. Samples were from HB101 (lanes a, b, and c; 300, 150, and 30 μg of protein, respectively), a *hydL* mutant (BO2, lane d; 300 μg), a *hydF* mutant (BO32, lane e; 300 μg), an F′ 116-transformed *hydF* (300 μg), plate grown (lane f) or grown in anaerobic batch culture (lane g), and a pHG10-transformed *hydF* mutant (lane h; 300 μg). Numbers 1 and 2 indicate hydrogenase isoenzymes according to Ballantine and Boxer (2). Experimental procedures were as described in Materials and Methods.

linked to *hydL* (Table 1). Partial complementation (25%) was found in *hydC* mutants carrying a lesion 10 min distant from the *hydLF* cluster (not shown). (This mutant is impaired in the uptake of nickel [42], a cofactor in probably all three hydrogenase isoenzymes. It may also have regulatory effects on several other genes involved in H_2 metabolism [45]. In the pHG10-transformed *hydC* strains, the hydrogenase core enzymes may be produced at wild-type levels, while the activity remains quite low, due to an insufficient nickel influx.) The *hyd* mutants, roughly mapped in the 58 min region, remained unaffected.

Isoenzyme specificity. The *hydF* mutants are impaired not only in the labile part of their hydrogenase content, but also in hydrogenase 1 and 2 activities. Therefore, these mutants could be used for the assessment of the specificity of the complementation with respect to the individual isoenzymes. A *hydF* mutant was transformed by pHG10 and also, as a control, by F-prime factor F′ 116, which had previously been shown to restore the hydrogenase activity in this mutant (42). The extracts were analyzed on activity-stained neutral gels. No increase of the hydrogenase 1 and 2 activities in the pHG10-transformed *hydF* mutant could be observed (Fig. 1, lanes e and h), whereas F-prime factor F′ 116 restored all three hydrogenase activities (lanes e, f, and g). Apparently, pHG10 controls only the electrophoretically labile part of the hydrogenase activity.

Restoration of Hup function. In addition to the hydrogenase activity, the influence of pHG10 expression on several related physiological activities was determined in strains BO32 (*hydF*) and BO2 (*hydL*) (Table 1). In neither strain was any effect observed on formate dehydrogenase-H, benzylviologen (red)-dependent fumarate reductase, formate hydrogenlyase, or formate:nitrate oxidoreductase activity levels. However, in strain BO32(pHG10), the Hup activity was elevated from 16 to 80% of the wild-type level. In BO2(pHG10) this activity remained essentially zero. Apparently, a correlation exists between respiratory H_2 uptake and electrophoretically labile hydrogenase activity (see also the Discussion section).

Hybridization of pHG10 with mutant chromosomal DNA. Since pHG10 (partially) complemented three different Tn*5* insertion mutants (*hydL* [41], *hydF* [42], and *hydC* [42, 45]), the question arose as to in which one the transposition had taken place into the gene carried on pHG10. Chromosomal DNA from BO2 (*hydL*), BO32 (*hydF*), BO26 (*hydC*), and parent HB101 was digested with several restriction endonucleases, for which Tn*5* carried no or one recognition site, electrophoresed, and filter hybridized with pHG10 as probe. If a gene was interrupted by transposon Tn*5*, this would cause a change in the electrophoretic mobility of fragments from this gene. In our case, however, the mutant patterns of hybridizing bands were identical to that of the parent strain (data not shown). Therefore, it was concluded that the gene carried on pHG10 is not identical to *hydL*, *F*, or *C*. It is probably also not identical to *hydA*, *B*, or *E*, since these genes have been cloned and their physical maps do not show homology (9, 16, 34, 44). The gene carried on pHG10 was apparently a new one and was designated *hydG*.

Assuming that the genomic library used contained all genes (about one gene per insert), these results also demonstrated that the labile hydrogenase activity in *hydL* mutants cannot not be restored by the homologous *hydL* gene. A possible explanation will be discussed below (see Discussion).

Physical mapping and determination of the *hydG* coding region. After establishing a physical map of pHG10 (see Fig. 2 and 4), deleted clones were constructed by cutting with *Sac*I (pHG11), *Bss*HII plus *Sac*I (pHG12), *Tth*III I plus *Sac*I (pHG13), or with *Pst*I (cutting in the linker and insert, pHG14), after which they were tested for their ability to complement *hydL* strains (as measured by the MV filter assay). It was found (Fig. 2) that the essential region extended from the left junction side, just after the promoter and the initial coding triplets of *lacZ*, to a region of 160 bp between the *Tth*III I and *Bss*HII recognition sites, about 1.3 kb in length. However, when the insert of pHG10 was recloned in an inverted orientation with respect to the *lacZ* promoter region (pHG15) or in the *Pst*I site of pBR322 (not shown), no complementation was found. It was concluded that the cloned gene lacked a functional promoter and that the transcription of *hydG* was apparently dependent on the *lacZ* promoter. In addition, when a frameshift mutation was introduced between the *lacZ* ATG start codon and the vector

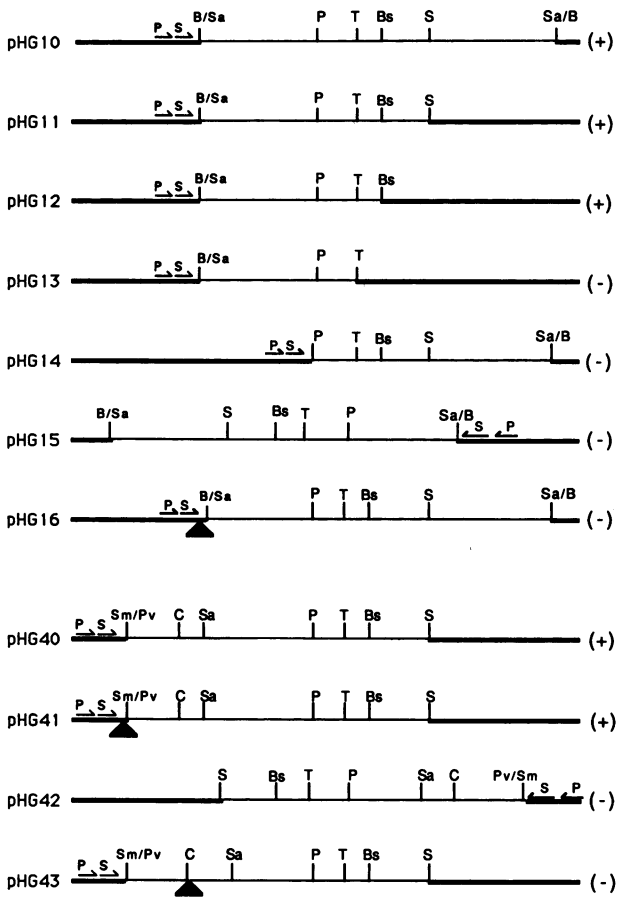

FIG. 2. Delimitation of the *hydG* coding region. In the original clones, mutations were introduced and the effect on the ability to complement *hydL* mutants with respect to the H_2:MV oxidoreductase activity was determined (MV filter assay), as indicated by (+) or (−). The bold line represents vector DNA (pUC13). B, *Bam*HI; Sa, *Sau*3A; P, *Pst*I; T, *Tth*III I; Bs, *Bss*HII; S, *Sac*I; Sm, *Sma*I; Pv, *Pvu*II; C, *Cla*I. For *Sau*3A more sites were present but were not indicated. P→, lacZ promoter region; S→, *lacZ* ribosome-binding site and ATG start codon; ▲, a 1- or 2-bp insertion or deletion. Procedures were as described in Materials and Methods.

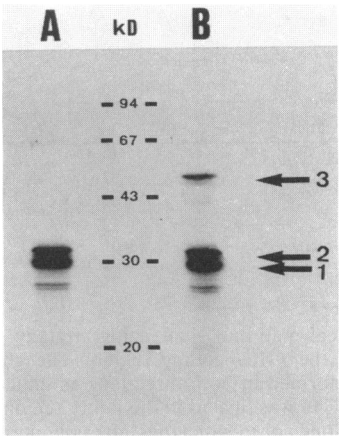

FIG. 3. Autoradiograph of a sodium dodecyl sulfate-polyacrylamide gel on which [^{35}S]methionine-labeled polypeptides encoded by pUC13 (lane A) and pHG40 (lane B) were run. Arrows: 1 and 2, β-lactamase proteins; 3, the polypeptide encoded by *hydG*, showing an apparent molecular mass of approximately 50 kDa. Molecular masses of marker proteins are indicated. Minicell procedures and electrophoresis were carried out as described in Materials and Methods.

insert junction (pHG16), the activity was also abolished, demonstrating the absence of a ribosome-binding (SD) sequence as well, the translation being dependent on the vector-derived start signals.

Therefore, a new partial genomic library of *E. coli* was constructed with DNA digested with *Pvu*II and *Sac*I. This digestion mixture had been shown by Southern blotting experiments to contain a 2.1-kb fragment that hybridized with the 1.4-kb *Bam*HI-*Sac*I insert from pHG11. This 2.1-kb fragment thus carries a 5′ extension of 0.7 kb, in comparison with the pHG11 insert. The plasmid carrying this 2.1-kb insert, pHG40, was isolated, and its complementation behavior was found to be identical with that of the smaller pHG11 clone. Analogous delineation experiments (Fig. 2) showed that this time a frameshift mutation between the SD sequence of *lacZ* and the beginning of the insert (pHG41) had no influence on the complementing behavior of this clone but that a 1-bp insertion at the *Cla*I site (pHG43) abolished this activity. Apparently, the larger pHG40 insert did contain the *hydG* translation start signals located upstream from the *Cla*I recognition site. The inverted clone (pHG42) was again no longer active. Therefore, the 0.7-kb 5′-extended clone also did not contain the *hydG* promoter.

The physical maps of pHG20 and pHG30 were identical to that of pHG10.

Taking all data together, it was concluded that pHG40 contained the complete *hydG* coding sequence, whereas pHG10 lacks about 200 bp from the 5′ terminus of this gene (as derived from the later-obtained sequence data, given below). Remarkably, this deletion did not affect the biological activity, suggesting that the approximately 70 N-terminal amino acids encoded by this region are not functionally essential, at least not in these complementation experiments (see also Discussion).

Molecular size of the *hydG*-encoded polypeptide. Minicells transformed with pUC13 vector DNA produced two labeled polypeptides, which showed apparent molecular masses of 28 and 33 kilodaltons (kDa), encoded by the β-lactamase gene (31). The pHG40-transformed cells showed an additional band, corresponding to approximately 50 kDa (Fig. 3). Therefore, the insert of pHG40 codes for one polypeptide, the *hydG* gene product (HydG).

Determination of the pHG40 nucleotide sequence. The insert of pHG40 was cloned in M13mp18/19 (Fig. 4), and the nucleotide sequence was determined (Fig. 5). A 1,314-bp open reading frame was deduced between residues 411 and 1726. It was preceded by a possible ribosome-binding site but not by a recognizable promoter sequence. The polypeptide predicted for this open reading frame contained 437 amino acid residues, with a calculated molecular mass of 48 kDa, which is in reasonable agreement with the value determined in minicells (Fig. 3). It was assumed to represent the *hydG* gene product.

Comparison with other protein sequences. A computer search was used to compare the predicted HydG amino acid sequence with a compilation of sequences in the Protein Identification Resource data base. Unexpectedly, a high degree of homology was found (Fig. 6) between our gene product and the NtrC protein from *Klebsiella pneumoniae* (8, 11) and its equivalent in *E. coli*, GlnG (21), both global regulators of assimilatory nitrogen metabolism. An absolute

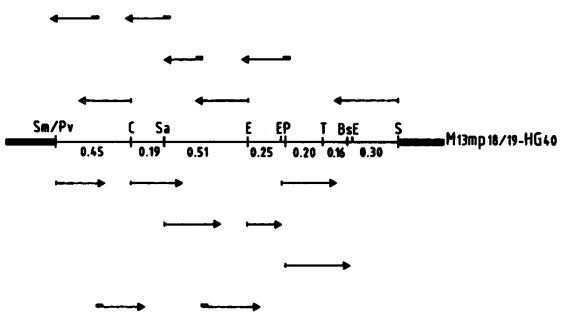

FIG. 4. Physical map and sequencing strategy of a DNA fragment containing the *hydG* coding region. The chain termination reaction was performed in both directions as indicated by arrows. The insert of pHG40 was ligated in the linker region of M13mp18 or mp19. In these clones or in subclones, the polymerase reaction was primed from M13 by using the universal primer (⟶) or from internally annealing 18-mer synthetic primers (▬⟶). Restriction sites are indicated as in Fig. 2. Procedures were as described in Materials and Methods.

homology of 41% was observed, whereas the functional homology was found to be 64%. This homology was equally spread over all functional domains on NtrC and GlnG, as delineated by Drummond et al. (11), although in the case of HydG, the interdomain region C was not particularly hydrophilic, as had been found to be the case for NtrC and related proteins such as OmpR and SfrA. One region of nonhomology was very conspicuous, an apparent deletion of 25 residues starting from HydG residue 399 (Fig. 6). (By checking different *hydG*-containing clones independently isolated from different libraries, it was determined that this deletion was not a cloning artifact.)

The homology within the C-terminal domain included the conserved DNA-binding motive, which is found in many DNA-binding proteins analyzed thus far (12, 23, 26, 38). When the amino acids are represented by their functional group number, this sequence can be written as 31A–1G12–21, giving rise to a typical α helix-turn-α helix structure (39). The functional homology of this 9-bp consensus with the analogous stretch in HydG (between residues 418 and 431) was 8 out of 9 (32A–1G12–21).

As a consequence of the homology with NtrC, there was also homology with other regulatory proteins which respond to other environmental stimuli such as phosphate limitation (PhoB), altered osmolarity (OmpR), presence of female cells (SfrA), presence of plant exudate (*Agrobacterium tumefaciens* VirG), the presence of C_4-dicarboxylates (*Rhizobium leguminosarum* DctD) or nutrient depletion, giving rise to chemotaxis (CheB, CheY) or sporulation in *Bacillus subtilis* (SpoOA, SpoOF) (11, 14, 23). There is also homology with the NifA protein from different species (*K. pneumoniae*, *Rhizobium meliloti*), although limited to the internal and C-terminal parts (11).

Significant homologies with Fnr (37), the product of *fnr* (a gene involved in the regulation of several aspects of anaerobic respiration in *E. coli*), were not found. These two proteins share only the presumed DNA-binding consensus mentioned above.

The overall homology on the DNA level between *hydG* and *ntrC* was 46%, although regions with a considerably higher degree of conservation occurred, e.g., between *hydG* residues 900 to 921 (85%), 927 to 943 (93%), 1041 to 1069 (85%), 1089 to 1141 (76%), 1335 to 1386 (76%), and 1458 to 1489 (83%) (Fig. 5).

Analysis of the *hydG* upstream sequence. When the *hydG* upstream sequence was analyzed, a part of a second open reading frame overlapping *hydG* by four bases was found (Fig. 5, upper part). Interestingly, when the amino acid sequence of 137 residues deduced from this frame was compared with the Protein Identification Resource data base, again significant homology was observed, this time with the C-terminal part of NtrB (Fig. 7), as well as with the carboxyl termini of other analogous proteins such as EnvZ (23). The absolute homology with NtrB was 34%, and the functional homology was 64%. This, in combination with the absence of a functional promoter for *hydG*, was considered to be strong evidence that *hydG*, like the other *ntrC*-like genes, is probably also part of an operon preceded by a *ntrB*-like gene, designated *hydH*.

DISCUSSION

In a previous paper (41) a mutant was described (*hydL*) that had lost the labile hydrogenase activity (hydrogenase 3 [35]) and, concomitantly, the respiratory Hup activity. Here we report the cloning of a DNA fragment (pHG40) which carries a new hydrogenase regulatory gene, *hydG*, that restores, specifically, the labile hydrogenase activity in *hydL* mutants, without restoring the Hup pathway.

The first conclusion was, therefore, that *hydL* mutants are lacking not only the labile hydrogenase but also a second component, possibly an electron transport protein, linking the oxidation of hydrogen to the anaerobic respiratory chain leading to fumarate. (This is provisionally designated component R. This model resembles that described by Bonnefoy-Orth and co-workers, who studied nitrate reductase regulation [7, 25]). Both the labile hydrogenase and component R seem to be required for a proper Hup function.

Furthermore, we had to account for three other observations: (i) hydrogenase 3 and component R disappear simultaneously upon transposon insertion in *hydL* (41); (ii) hydrogenase 3 activity can be induced in *hydL* mutants by the *hydG* gene product (HydG), without inducing the expression of component R (since the Hup pathway remains impaired); and (iii) hydrogenase 3 activity cannot be restored in *hydL* mutants by introducing in *trans* the homologous *hydL* gene (since this gene could not be picked up from the plasmid library).

To reconcile these data, it is hypothesized that *hydL* might be the structural gene for component R, located adjacent to and upstream from the structural gene for hydrogenase 3, both genes normally being transcribed from the *hydL* promoter (e.g., when fermenting on complex medium). However, the downstream hydrogenase 3 gene might be preceded by a second promoter, which is dependent on HydG. Under the growth condition mentioned above, HydG is apparently not expressed from its locus in the genome. (Otherwise *hydL* mutants would only lack component R, and not hydrogenase 3, and would be positive in the MV filter assay). However, when cloned and stripped from any regulatory sequences, as on plasmid pHG10 or pHG40, HydG is expressed and leads in *hydL* mutants to the transcriptional activation of the structural gene for hydrogenase 3, leading to a MV^+ Hup^- phenotype (since component R is still lacking).

Under which growth conditions the genomic counterpart of *hydG* (and *hydH*) will be transcribed is still a question. Also unknown is whether and to what extent the *hydH* product is required for HydG function.

Finally, the model also makes clear that the effect of *hydL* interruption on hydrogenase 3 expression is *cis* dominant

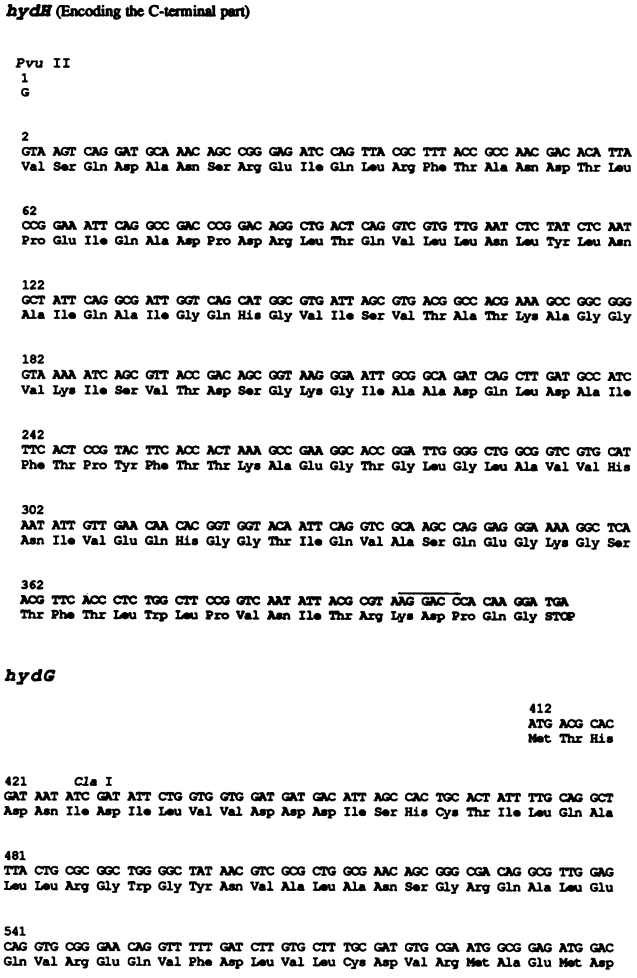

FIG. 5. Nucleotide sequence of *hydG* and the C-terminal part of *hydH*. The two frames are overlapping by four bases. Shown are 1,740 nucleotides from the pHG40 insert (Fig. 2) starting from the *Sma*I-*Pvu*II junction site. Predicted amino acid sequences of 437 residues (*hydG*) and 137 residues (*hydH*) are given. Straight line, tentative Shine-Dalgarno ribosome-binding site. Sequence determination was as described in Materials and Methods.

and cannot be relieved by introducing in *trans* a *hydL*-containing DNA fragment.

More direct evidence for the physiological role of HydG was derived from its homology with the *K. pneumoniae* NtrC product, a regulatory protein. NtrC is a transcriptional activator (and in some cases a repressor) of several operons involved in nitrogen assimilation (*nif*, *hut*, and *put*) and can be present in an active or inactive state (14). It is activated by NtrB, a phosphokinase. NtrB might be considered a sensor protein which responds to a physical stimulus (glutamine/α-ketoglutarate ratio). The encoding genes, *ntrB* and *ntrC*, are arranged in an operon. The NtrB-NtrC couple is the prototype of a distinct set of two-component regulatory systems. Most NtrC-like regulators function in pairs (PhoR-PhoB, EnvZ-OmpR, CpxA-SfrA, VirA-VirG, DctB-DctD, CheA-CheB, CheY), including NifA, which is modified by the NifL product. NtrC binds to DNA, and at least for NtrC and NifA, it has been demonstrated that they perform their function in a coordinate action with a specific sigma factor (NtrA).

HydG showed considerable homology with respect to the presumed DNA-binding consensus (90%), as well with the other functional domains that can be discerned on NtrC. However, whether all these structural similarities also imply a similar physiological function is still a question.

Also, the sequence data, although still incomplete, suggest that *hydG* may be arranged in an operon as well, being preceded by a *ntrB*-like gene, *hydH*, although this has to be demonstrated in a more direct way.

Nevertheless, (a part of) a new two-component regulation system controlling the labile hydrogenase 3 activity and possibly related functions was revealed.

The conclusion that the labile hydrogenase activity in *E. coli* is controlled by a NtrA/B/C-like system is in agreement

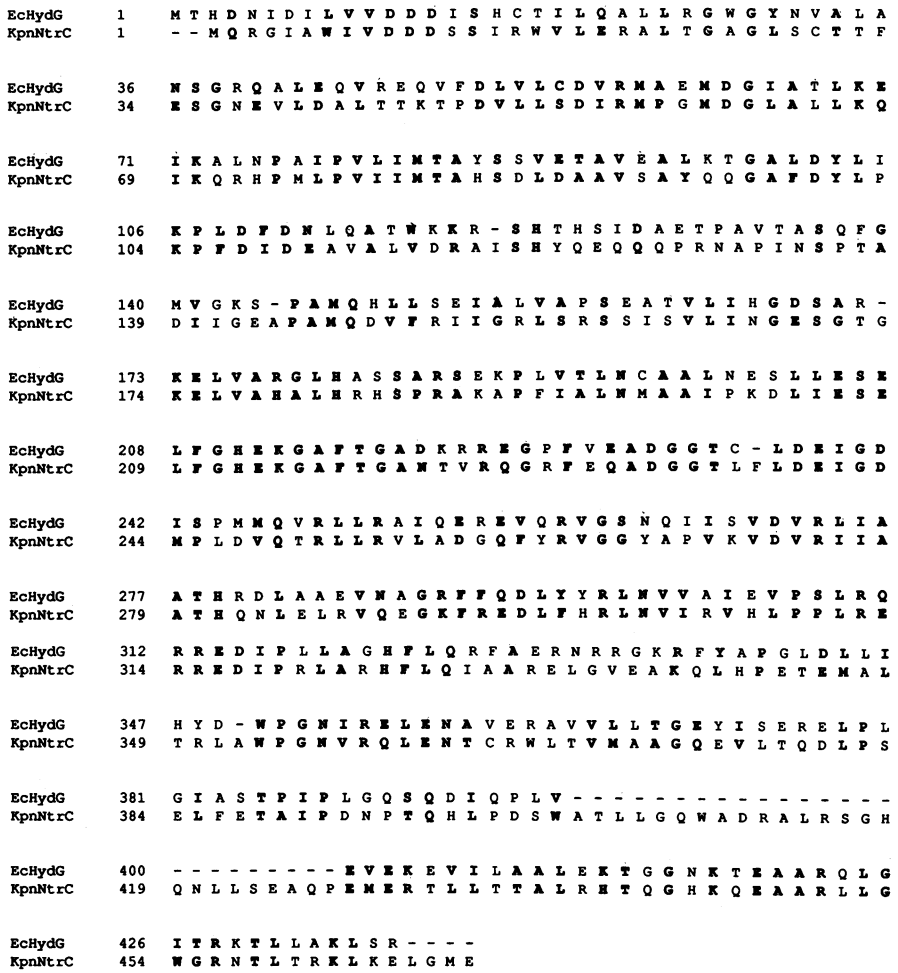

FIG. 6. Homologies between the amino acid sequences of HydG from *E. coli* (EcHydG) and NtrC from *K. pneumoniae* (KpnNtrC). Identical residues and functionally conservative substitutions are represented by bold characters. Search program and functional grouping were as described in Materials and Methods.

with and nicely complementary to the results of Birkmann et al. (5). These authors demonstrated that the labile hydrogenase activity (and that of formate dehydrogenase-H) was *ntrA* dependent, whereas we provided evidence for the involvement of genes homologous to *ntrB* and *ntrC*.

Also, the hydrogenase activity in *Alcaligenes eutrophus* might be ntrA dependent (30). Finally, our results, as well as those of Birkmann et al. (5), clearly demonstrate that the hydrogenase isoenzymes 1 and 2 are not under the control of *ntrA* or *hydG*.

Several of our own observations can be explained if it is assumed that *hydG* encodes a transcriptional activator. The fact that elevation of the expression of HydG by isopropyl-β-D-thiogalactopyranoside did not elevate the hydrogenase

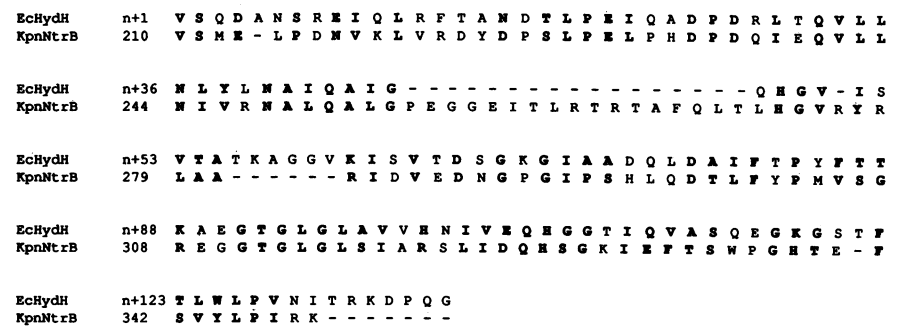

FIG. 7. Homologies between the C-terminal amino acid sequences of HydH from *E. coli* (EcHydH) and NtrB from *K. pneumoniae* (KpnNtrB). Identical residues and functionally conservative substitutions are represented by bold characters. Search program and functional grouping were as described in Materials and Methods.

activity in complemented mutants demonstrated that even a very small amount of this product suffices to reach a maximal effect, which is typical for a regulatory protein. Furthermore, the frameshift mutation, introduced at the *Tth*III I site (Fig. 5), completely eliminated the biological activity. From the amino acid sequence, it could be derived that this mutation destroyed the presumed DNA-binding structure, suggesting that association with DNA might be a requirement for biological function. Apparently not essential are the first 70 residues from the N terminus, which are lacking in pHG10, which is nevertheless equally active as pHG40, containing the complete gene. Indeed, when comparing HydG with NtrC (Fig. 6), it was found that the absolute homology between residues 1 and 60 was only 21%, whereas the homology between residues 61 and 122 was 50%, another indication that the essential region in the B domain is located somewhat more distal from the N terminus. (An interesting alternative explanation is that the N-terminal part of HydG is not required for activation but is required for inactivation).

ACKNOWLEDGMENT

This research was supported by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organization for the Advancement of Science (N.W.O.).

LITERATURE CITED

1. **Adams, M. W. W., L. E. Mortenson, and J.-S. Chen.** 1981. Hydrogenase. Biochim. Biophys. Acta **594**:105–176.
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 fragment of hydrogenase isoenzyme 2 from the membranes of anaerobically grown *Escherichia coli*. Eur. J. Biochem. **156**:277–284.
4. **Bernhard, T., and G. Gottschalk.** 1978. Cell yields of *Escherichia coli* during anaerobic growth on fumarate and molecular hydrogen. Arch. Microbiol. **116**:235–238.
5. **Birkmann, A., R. G. Sawers, and A. Böck.** 1987. Involvement of the *ntrA* gene product in the anaerobic metabolism of *Escherichia coli*. Mol. Gen. Genet. **210**:535–542.
6. **Birkmann, A., F. Zinoni, G. Sawers, and A. Böck.** 1987. Factors affecting transcriptional regulation of the formate-hydrogenlyase pathway of *Escherichia coli*. Arch. Microbiol. **148**:44–51.
7. **Bonnefoy-Orth, V., M. Lepelletier, M.-C. Pascal, and M. Chippaux.** 1981. Nitrate reductase and cytochrome *b* nitrate reductase structural genes as parts of the nitrate reductase operon. Mol. Gen. Genet. **181**:535–540.
8. **Buikema, W. J., W. W. Szeto, P. V. Lemley, W. H. Orne-Johnson, and F. M. Ausubel.** 1985. Nitrogen specific regulatory genes of *Klebsiella pneumoniae* and *Rhizobium meliloti* share homology with the general nitrogen regulatory gene of *K. pneumoniae*. Nucleic Acids Res. **13**:4539–4555.
9. **Chaudhuri, A., and A. I. Krasna.** 1987. Isolation of genes required for hydrogenase synthesis in *Escherichia coli*. J. Gen. Microbiol. **133**:3289–3298.
10. **Dougan, G., and D. Sherratt.** 1977. The transposon Tn1 as a probe for studying Col E1 structure and function. Mol. Gen. Genet. **151**:151–160.
11. **Drummond, M., P. Whitty, and J. Wootton.** 1986. Sequence and domain relationships of *ntrC* and *nifA* from *Klebsiella pneumoniae*: homologies to other regulatory proteins. EMBO J. **5**: 441–447.
12. **Friedrich, M. J., and R. J. Kadner.** 1987. Nucleotide sequence of the *uhp* region of *Escherichia coli*. J. Bacteriol. **169**:3556–3563.
13. **Graham, A., D. H. Boxer, B. A. Haddock. M.-A. Mandrand-Berthelot, and R. W. Jones.** 1980. Immunochemical analysis of the membrane-bound hydrogenase of *Escherichia coli*. FEBS Lett. **113**:167–172.
14. **Gussin, G. N., C. W. Ronson, and F. M. Ausubel.** 1986. Regulation of nitrogen fixation genes. Annu. Rev. Genet. **20**: 567–591.
15. **Jones, H. M., and R. P. Gunsalus.** 1985. Transcription of the *Escherichia coli* fumarate reductase genes (*frdABCD*) and their coordinate regulation by oxygen, nitrate, and fumarate. J. Bacteriol. **164**:1100–1109.
16. **Karube, I., M. Tomiyama, and A. Kikuchi.** 1984. Molecular cloning and physical mapping of the *hyd* gene of *Escherichia coli* K-12. FEMS Microbiol. Lett. **25**:165–168.
17. **Lee, J. H., P. Patel, P. Sankar, and K. T. Shanmugam.** 1985. Isolation and characterization of mutant strains of *Escherichia coli* altered in H_2 metabolism. J. Bacteriol. **162**:344–352.
18. **Macy, J., H. Kulla, and G. Gottschalk.** 1976. H_2-dependent anaerobic growth of *Escherichia coli* on L-malate: succinate formation. J. Bacteriol. **125**:423–428.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
21. **Miranda-Ríos, J., R. Sanchez-Pescador, M. Urdea, and A. Covarrubias.** 1987. Complete nucleotide sequence of the *glnALG* operon of *Escherichia coli* K12. Nucleic Acids Res. **15**:2757–2770.
22. **Mooi, F. R., N. Harms, D. Bakker, and F. K. De Graaf.** 1981. Organization and expression of genes involved in the production of the K88ab antigen. Infect. Immun. **32**:1155–1163.
23. **Nixon, B. T., C. W. Ronson, and F. M. Ausubel.** 1986. Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes *ntrB* and *ntrC*. Proc. Natl. Acad. Sci. USA **83**:7850–7854.
24. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene **26**:101–106.
25. **Orth, V., M. Chippaux, and M.-C. Pascal.** 1980. A mutant defective in electron transfer to nitrate in *Escherichia coli* K-12. J. Gen. Microbiol. **117**:257–262.
26. **Pabo, C. O., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53**:293–321.
27. **Pascal, M.-C., F. Casse, M. Chippaux, and M. Lepelletier.** 1975. Genetic analysis of mutants of *Escherichia coli* K-12 and *Salmonella typhimurium* LT2 deficient in hydrogenase activity. Mol. Gen. Genet. **141**:173–179.
28. **Pecher, A., F. Zinoni, C. Jatisatienr, R. Wirth, H. Hennecke, and A. Böck.** 1983. On the redox control of synthesis of anaerobically induced enzymes in enterobacteriaceae. Arch. Microbiol. **136**:131–136.
29. **Peck, H. D., Jr., and H. Gest.** 1957. Formic dehydrogenase and the hydrogenlyase enzyme complex in coli-aerogenes bacteria. J. Bacteriol. **73**:706–721.
30. **Römermann, D., M. Lohmeyer, C. G. Friedrich, and B. Friedrich.** 1988. Pleiotropic mutants from *Alcaligenes eutrophus* defective in the metabolism of hydrogen, nitrate, urea and fumarate. Arch. Microbiol. **149**:471–475.
31. **Sancar, A., A. M. Hack, and W. D. Rupp.** 1979. Simple method for identification of plasmid-coded proteins. J. Bacteriol. **137**: 692–693.
32. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
33. **Sanger, F., R. Coulson, B. G. Barrell, J. H. Smith, and B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143**:161–178.
34. **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.
35. **Sawers, R. G., 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.

36. **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.
37. **Shaw, D. J., and J. R. Guest.** 1982. Nucleotide sequence of the *fnr* gene and primary structure of the Fnr protein of *Escherichia coli*. Nucleic Acids Res. **19:**6119–6130.
38. **Shaw, D. J., D. W. Rice, and J. R. Guest.** 1983. Homology between CAP and Fnr, a regulator of anaerobic respiration in *Escherichia coli*. J. Mol. Biol. **166:**241–247.
39. **Steitz, T. A., D. H. Ohlendorf, D. B. McKay, W. F. Anderson, and B. W. Matthews.** 1982. Structural similarity in the DNA-binding domains of catabolite gene activator and Cro repressor proteins. Proc. Natl. Acad. Sci. USA **79:**3097–3100.
40. **Stewart, V.** 1988. Nitrate respiration in relation to facultative metabolism in Enterobacteria. Microbiol. Rev. **52:**190–232.
41. **Stoker, K., L. F. Oltmann, and A. H. Stouthamer.** 1988. Partial characterization of an electrophoretically labile hydrogenase activity of *Escherichia coli* K-12. J. Bacteriol. **170:**1220–1226.
42. **Stoker, K., L. F. Oltmann, and A. H. Stouthamer.** 1989. Randomly induced *Escherichia coli* K-12 Tn*5* insertion mutants defective in hydrogenase activity. J. Bacteriol. **171:**831–836.
43. **Vieira, J., and J. Messing.** 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19:**259–268.
44. **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.
45. **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.
46. **Yamamoto, I., and M. Ishimoto.** 1978. Hydrogen-dependent growth of *Escherichia coli* in anaerobic respiration and the presence of hydrogenases with different functions. J. Biochem. **84:**673–679.