Characterization and expression of the hydrogenase-encoding gene from Clostridium acetobutylicum P262

Joseph D. Santangelo,1 Peter Dürré1 and David R. Woods2

Author for correspondence: Joseph D. Santangelo. Tel: +49 551 39 3813. Fax: +49 551 39 3793. e-mail: jsantan@gwdg.de

1 Institut für Mikrobiologie, Georg-August-Universität Göttingen, Grisebachstraße 8, D-37077 Göttingen, Germany
2 Department of Microbiology, University of Cape Town, Rondebosch, 7700, South Africa

The hydrogenase enzyme of Clostridium acetobutylicum plays a pivotal role in controlling electron flow, and hence carbon flow, during the complex biphasic fermentation of carbohydrates to the neutral solvents acetone and butanol. We report here the cloning and molecular characterization of the hydrogenase-encoding gene (hydA) from C. acetobutylicum P262. This gene was isolated by colony hybridization, using the Clostridium pasteurianum hydrogenase-I gene as a probe. The DNA sequence encoding the hydA gene from C. acetobutylicum was determined, and revealed an ORF (1722 bp) encoding a 574 amino-acid protein. This C. acetobutylicum hydrogenase protein product has 82% similarity and 67% identity with the C. pasteurianum hydrogenase-1 protein. Northern blot analysis of RNA isolated from C. acetobutylicum indicates that the C. acetobutylicum hydrogenase protein product is translated from a monocistronic operon. RNA was isolated from the different morphological and physiological stages of a batch C. acetobutylicum fermentation, and further Northern blot analyses revealed no differences in the expression of the gene during acidogenesis as opposed to solventogenesis. Primer extension experiments confirmed these results and identified the 5’ start of the mRNA transcript. These results correlated well with the physiological need for this organism to dispose of excess reducing equivalents.

Keywords: Clostridium acetobutylicum, hydrogenase, electron flow, hydA

INTRODUCTION

Clostridium acetobutylicum is well known for its ability to produce the industrially important chemical feedstocks acetone and butanol. This strictly anaerobic, Gram-positive endospore-forming bacterium carries out a biphasic fermentation, where acetic and butyric acids are produced during the exponential growth phase, followed by solvent production occurring during the stationary phase (Jones & Woods, 1986). Since the late 1970s, a considerable amount of research has gone into studying the mechanisms behind the metabolic shift to solvent production in this organism.

The importance of electron flow, with respect to regulating carbon flow, has been the topic of a number of studies (Martin et al., 1986; Vasconcelos et al., 1994). Among the solvent-producing clostridia, the pathways for acid production alone do not provide for the disposal of excess NADH produced during glycolysis. These organisms have the ability to produce hydrogen and therefore can process both excess electrons and protons through this route. During acidogenesis, a major portion of the electron flow is directed to hydrogen production, whereas the carbon flow is directed to acid production. Studies have shown that during the exponential growth phase more hydrogen is produced than is theoretically possible from the phosphoroclastic breakdown of pyruvate alone (Martin et al., 1982; Kim & Zeikus, 1985), indicating that excess NADH is cycled through ferredoxin towards the production of hydrogen gas. In C. acetobutylicum the shift from acidogenesis to solventogenesis is accompanied by a decrease in hydrogen production and an increase in CO2 production. Less hydrogen is produced than would be expected from the oxidation of pyruvate, demonstrating that during the stationary phase both carbon and electron flow are primarily directed to solvent production.

The GenBank accession number for the sequence reported in this paper is U09760.
The hydrogenase enzyme plays a pivotal role in controlling the direction of electron flow, as well as the carbon flow appears to be directly linked to the culture vessel (Gapes et al., 1982), or by sparging the culture vessel with CO (Kim et al., 1984). Hydrogenase activity, and therefore solvent production, could be altered by changing the partial pressure of hydrogen in the culture vessel (Gapes et al., 1982), or by sparging the culture vessel with CO (Kim et al., 1984; Datta & Zeikus, 1985; Meyer et al., 1985, 1986), a known inhibitor of hydrogenase enzymes (Thauer et al., 1974; Mortensen & Chen, 1975). Under these conditions, an earlier onset of fermentation was observed. Because the hydrogenase was inhibited, the flow of electrons from reduced ferredoxin was shifted to the generation of NADH and NADPH via the actions of the respective ferredoxin-oxidoreductases. Stimulation of butanol production was therefore necessary for the disposal of electrons carried by these reduced pyridine nucleotides.

We report here the cloning, sequencing and mRNA analysis of the hydrogenase-encoding gene (bydA) from C. acetobutylicum P262, one of the latest C. acetobutylicum strains to be utilized for the industrial production of solvents (Spivey, 1978; Jones & Woods, 1986). A preliminary report of the data presented here was given at the annual meeting of the Federation of European Microbiological Societies, Hannover, Germany, 7–9 March 1994.

METHODS

Bacterial strains, plasmids and growth conditions. C. acetobutylicum P262 (Jones et al., 1982) was used as the source of DNA. Escherichia coli strains CC118 [araD139 (ara leu)7697 ΔlacX74 phoAΔ20 gatE galK thi rpiE rpsB argE (Am) recA1] (Manoil & Beckwith, 1985) and JM109 [recA1 endA1 gyrA96 thi bsdR17 supE44 relA1 Δ(lac-proAB)] (Yanisch-Perron et al., 1985) were used interchangeably throughout this study. The positive selection plasmid vector pEcoR251 was a gift from M. Zabeau, Biotechnology Business Development, Ghent, Belgium, and has been described previously (Zapf et al., 1986). Plasmid pCPh10, which contained a region of the Clostridium pasteurianum hydrogenase-1 gene (Meyer & Gagnon, 1991), was generously donated by Jacques Meyer, DBMS-Métalloprotéines, Grenoble, France.

Inoculum cultures of C. acetobutylicum P262 were grown under strictly anaerobic conditions in the Clostridium basal medium of O'Brien & Morris (1971) as described by Allcock et al. (1982). Batch fermentations of C. acetobutylicum P262 were carried out at 34 °C in TYG medium which contained per l: glucose, 60 g; yeast extract, 2 g; tryptone, 6 g; MgSO₄, 7H₂O, 0.2 g; MnSO₄.4H₂O, 10 mg; FeSO₄.7H₂O, 10 mg; p-amino benzoic acid, 2 mg; biotin, 0.1 mg; KH₂PO₄, 5 g; diammonium hydrogen phosphate, 0.6 g; adjusted to pH 6.2 with NaHCO₃. E. coli was grown aerobically in a Luria broth (LB) or on Luria agar (LA) (Sambrook et al., 1989), which was supplemented with ampicillin (100 μg ml⁻¹) when required for plasmid selection and maintenance.

Nucleic acid isolation and manipulation. Plasmid DNA was prepared by the alkaline hydrolysis method of Ish-Horowicz & Burke (1981). C. acetobutylicum P262 chromosomal DNA was prepared by the method of Marmur (1961) as modified by Zapf et al. (1986) to overcome the high nuclease activity exhibited by C. acetobutylicum (Burchardt & Dürre, 1990; Lin & Blaschek, 1984). Chromosomal DNA from other C. acetobutylicum strains and from E. coli strain JM109 was prepared as described by Bertram & Dürre (1989). Total RNA from C. acetobutylicum and E. coli was isolated according to the method of Oelmüller et al. (1990) as modified by Gerischer & Dürre (1992). DNA and RNA were manipulated by standard methods (Sambrook et al., 1989); the restriction endonucleases employed were obtained from a variety of sources and used according to the manufacturer's instructions.

Colony hybridization. Colony hybridization techniques were employed for the isolation of the bydA gene from a C. acetobutylicum P262 gene library using the 1·23 kb EcoRI–AccI fragment of the C. pasteurianum hydrogenase-1 gene (corresponding to positions 352–1589 in the GenBank sequence CLOHDGI, accession number M62754; Meyer & Gagnon, 1991) as a probe. This probe DNA was radiolabelled with [α-³²P]dATP using a nick translation kit (Amersham) according to the manufacturer's instructions. Colonies harbouring recombinant plasmids with C. acetobutylicum insert DNA were duplicated onto Hybond-N nylon membranes (Amersham) which were placed on LA medium. After overnight growth at 37 °C the membranes were placed on agar plates containing chloramphenicol (170 μg ml⁻¹) and incubated at room temperature for approximately 16 h in order to amplify the plasmid DNA. The membranes were then processed and probed using standard techniques (Sambrook et al., 1989).

DNA: DNA and DNA: RNA hybridization. Plasmid DNA and C. acetobutylicum chromosomal DNA were digested to completion with the appropriate endonucleases and the resulting fragments were fractionated by electrophoresis in 0·8 % (w/v) agarose gels in Tris/acetate buffer. The DNA was transferred monodirectionally and fixed to GeneScreen-Plus nylon membranes (DuPont de Nemours) according to the manufacturer's instructions. Membranes were prehybridized for 1–3 h at 62 °C in a buffer (75 μl cm⁻²) containing 0·2 % (w/v) polyvinylpyrrolidone, 0·2 % (w/v) BSA, 0·2 % (w/v) Ficoll, 1 M NaCl, 1 % (w/v) SDS, 0·1 % (w/v) sodium pyrophosphate, 10 % (w/v) dextran sulfate, 50 mM Tris/HCl (pH 7·5), and 100 μg debranched salmon sperm DNA ml⁻¹. Hybridization with radiolabelled DNA fragments (approximately 1·×10⁶ c.p.m.) was carried out for approximately 16 h at 62 °C (Southern, 1975). Total RNA was separated in 1 % (w/v) denaturing formaldehyde agarose gels (Sambrook et al., 1989). The fractionated RNA was then transferred monodirectionally to Hybond-N nylon membranes (Amersham) as specified by the manufacturer. The RNA was fixed to the membrane by exposure to UV light (254 nm) for 2·5 min. Hybridization with radiolabelled DNA fragments was performed as described above for Southern hybridization, except that salmon sperm DNA was omitted from the hybridization solution and the hybridization temperature was reduced to 35 °C.

Probes used for the Southern and Northern hybridization experiments were labelled with [α-³²P]dATP using the random primers DNA labelling system of Gibco/BRL. Life Technologies according to the manufacturer's instructions.

Nucleotide sequencing. A 1026 bp SpI DNA fragment, containing the majority of the C. acetobutylicum bydA gene, was subcloned into the EcoRV site of Bluescript plasmid SK (Stratagene). Progressive deletions of this subclone from both the 5' and 3' ends of the insert were generated by unidirectionally.
digested KpnI–AclI and SstI–XbaI fragments with exonuclease III (Henikoff, 1984). The deletions were transformed into E. coli CC118 and transformants were selected on LA containing ampicillin (100 μg ml⁻¹).

Nucleotide sequencing was carried out by the dideoxynucleotide triphosphate chain-termination method developed by Sanger et al. (1977) according to the protocol outlined by Tabor & Richardson (1987), using the T7 sequencing kit (Pharmacia Biotech). The nucleotide and deduced amino acid sequences were analysed by a VAX 6000-330 computer using the Genetics Computer Group suite of sequence analysis programs (Devereux et al., 1984).

** Primer extension.** Primer extension studies were performed as previously described (Gerischer & Dürre, 1992), except that SuperScript-plus reverse transcriptase (Gibco/BRL Life Technologies) was employed. A 17-mer oligonucleotide (5'-TTTGCTAGGGCAGCTTG-3'), which is complementary to the 5' end of the hydA transcript, was radiolabelled at the 5' end with [α-³²P]ATP with the aid of T4 polynucleotide kinase (Gerischer & Dürre, 1992). In order to determine the exact start point (5'-terminus) of the mRNA, the cDNA was electrophoresed next to a standard sequencing reaction which was prepared using the same 17-mer oligonucleotide.

**RESULTS**

**Gene identification and Southern hybridization**

The cloned *C. acetobutylicum* hydA gene was isolated from a gene library via colony hybridization. Southern hybridization experiments were performed in order to locate the position of this gene within the 5-4 kb insert DNA of the clone (p9B8) using the 1-2 kb *EcoRI–AclI* *C. pasteurianum* hydrogenase-1-encoding fragment as a probe. A 1026 bp SspI fragment and a 2492 bp HindIII fragment yielded positive hybridization signals and these fragments were subcloned into the Bluescript vector SK, resulting in plasmids pS12 and pH12, respectively. Plasmid pH12 was employed as a probe. Equal sized hybridization signals (2128 bp) were obtained with the pH12 and SstI-HindIII (double digest) fragment, it can be concluded that the sequenced hydA-containing region represents a single chromosomal fragment.

Additional DNA hybridization experiments were performed in order to obtain information about the hydA genes from other *C. acetobutylicum* strains. Apart from *C. acetobutylicum* P262, chromosomal DNA from strains DSM 792, DSM 1731, ATCC 824 and NCIMB 8052 was used for these hybridization experiments. Three representative blots are presented here, where the chromosomal DNAs were digested with *SspI*, *HindIII* and *SstI–HindIII* (double digest), respectively. In all cases, the 1026 bp radiolabelled SspI fragment described above was employed as a probe. Equal sized hybridization signals (1026 bp) were obtained from *SspI*-digested pH12 and chromosomal DNA from *C. acetobutylicum* strains P262, DSM 792, DSM 1731 and ATCC 824. No positive hybridization signal was obtained with NCIMB 8052 chromosomal DNA. Similarly, these chromosomal DNAs were digested with *HindIII* and analysed via Southern hybridization (Fig. 2). In this case, plasmid pH12 was not included as there is only one *HindIII* recognition site in the insert DNA of this plasmid and no similarly sized fragment would be expected in the chromosomal digests. A single fragment in *HindIII*-digested P262 chromosomal DNA, with an approximate size of 2.6 kb, hybridized with the *SspI* probe. HindIII-digested chromosomal DNA from the three *C. acetobutylicum* strains DSM 792, DSM 1731 and ATCC 824 yielded identical hybridization patterns with two DNA fragments (approx. 1.7 and 0.9 kbp) having positive hybridization signals. Again, no hybridization signal was obtained with chromosomal DNA isolated from the *C. acetobutylicum* strain NCIMB 8052. In the third blot (Fig. 2), pH12 plasmid DNA along with the five different chromosomal DNAs, were digested with *HindIII* and *SstI*. As stated previously, identical sized signals (2128 bp) were obtained with the pH12 and P262 DNA. Chromosomal DNA from the three strains

![Diagrammatic representation of the hydA-containing clone p9B8. The thick shaded bar represents the 2267 bp region for which the nucleotide sequence data was determined. The thick open bar represents the approximately 3.5-kbp of unsequenced *C. acetobutylicum* insert DNA in clone p9B8. The thin lines represent pEcoR251 vector DNA. The hydrogenase-encoding region between the insert and vector HindIII sites was subcloned into Bluescript SK to yield plasmid pH12 (see text). The location and direction of transcription for the hydA gene is indicated by the thick black line below the diagram.](image-url)
Fig. 2. Hybridization of the hydA-containing, 1026 bp SspI fragment to chromosomal DNA from five different C. acetobutylicum strains. DNA was digested with SspI (a), HindIII (b) and HindIII–StuI (c). Lanes: 1 (a and c only), pH12 plasmid DNA; 2–6, C. acetobutylicum chromosomal DNA from strains P262, DSM 792, DSM 1731, ATCC 824 and NCIMB 8052, respectively. The autoradiogram signal sizes (kbp) are indicated to the left of each panel. Weaker signals were presumed to be due to incomplete DNA digestion.

DSM 792, DSM 1731 and ATCC 824 showed a major hybridization signal with a size of approximately 1.7 kbp. The logical conclusion is that this 1.7 kbp signal is analogous to the 1.7 kbp signal obtained with the HindIII-digested chromosomal DNAs. As seen in the first two blots, no hybridization signal was observed with chromosomal DNA from the NCIMB 8052 strain.

Nucleotide sequence

The nucleotide sequences of both strands of the insert DNA contained in plasmid pS12 were determined using overlapping DNA fragments generated by exonuclease III digestion. Unfortunately, this 1026 bp SspI fragment did not contain the entire hydA gene (bases 409–1435 in Fig. 3), and the remaining sequence data were therefore obtained from plasmid pH12 using ‘primer walking’ techniques. These sequence data revealed the presence of one complete ORF (574 amino acids), which encodes the hydA gene, and is represented by bp 456–2177 (Fig. 3). This coding region is preceded by a 145 bp intergenic region, and a putative ribosome-binding site (AGGAGC) was located 10 bp upstream of the ATG start codon. An invert repeat, characteristic of factor-independent terminators (Brendel & Trifonov, 1984), was located downstream of the hydA gene (bp 2207–2246).

Amino acid similarity

The deduced amino acid sequence encoded by the C. acetobutylicum hydA gene was used to search the GenBank, EMBL, SWISS-protein, NBRF-nucleic, and NBRF-protein databases with the BLASTP and TBLASTN computer programs (Altschul et al., 1990). As expected, there was considerable sequence similarity between the C. acetobutylicum HydA protein sequence and the hydrogenase-1 protein sequence from C. pasteurianum (Meyer & Gagnon, 1991). The amino acid sequences of other [Fe] hydrogenases also showed a high degree of similarity to the C. acetobutylicum HydA protein sequence. These included the large subunits (α-subunit) of the [Fe] hydrogenases from both Desulfovibrio vulgaris subsp. vulgaris (Voordouw & Brenner, 1985) and Desulfovibrio vulgaris subsp. oxamicus (Voordouw et al., 1989), the γ-hydrogenase (HydC) from D. vulgaris subsp. vulgaris (Stokkermans et al., 1989) and the fourth subunit of the potential NAD-reducing hydrogenase from Desulfovibrio fructosovorans (S. Malki and others, unpublished, GenBank accession number U07229). GAP, the alignment program based on the method of Needleman & Wunsch (1970), was used to compare the deduced amino acid sequence from the C. acetobutylicum hydA gene to the amino acid sequences of the above mentioned hydrogenases. The amino acid sequence of the C. pasteurianum hydrogenase-1 contained 67% identical residues and 82% similar residues to the C. acetobutylicum HydA protein. The amino acid sequences of the D. vulgaris subsp. vulgaris hydrogenase α-subunit, the D. vulgaris subsp. oxamicus hydrogenase α-subunit, the D. vulgaris γ-hydrogenase, and the fourth subunit of the D. fructosovorans NAD-reducing hydrogenase showed 43, 42, 38 and 43% identity and 61, 63, 59 and 64% similarity to the HydA protein sequence, respectively.

Determination of transcript size

RNA was isolated from acid-producing and solvent-producing C. acetobutylicum P262 cells which were grown in TYG medium. When cultivated under these conditions,
Fig. 3. Nucleotide sequence of the hydA gene and flanking regions from C. acetobutylicum. The deduced amino acid sequence is provided in single-letter code from nucleotide positions 456-2177 (574 amino acid residues). The mRNA start point (position 416), determined from primer extension experiments, is indicated above the nucleotide sequence. The −10 and −35 promoter regions and the proposed ribosome-binding site (SD) are double underlined and labelled above the sequence. The identification of an inverted repeat sequence, capable of forming a stem-loop structure, is shown by dashed lines below the sequence (position 2207-2246).
\textbf{C. acetobutylicum} P262 follows a sufficiently synchronous developmental cycle to allow for correlative morphological, physiological and biochemical studies (Woods \\& Jones, 1986). Results from Northern hybridization analysis suggest that the \textit{hydA} gene is expressed at the same level during both acidogenesis and solventogenesis, and the transcript size was estimated to be approximately 1.95 kbp (Fig. 4). Identical results were obtained with RNA isolated from \textit{E. coli} cells which contained either plasmid p9B8 or plasmid pH12 (data not shown). The same \textit{hydA}-containing \textit{SspI} fragment (bases 409–1435 in Fig. 3) used for the Southern hybridization studies was also employed as a probe for these Northern hybridization studies. A second hybridization signal, with an approximate size of 1.65 kbp, was also observed. Since 1.65 kbp is not large enough to encode for the entire \textit{hydA} gene, it was assumed that this signal was a result of non-specific binding to 16S rRNA (approximate size 1.7 kbp).

**Determination of 5' mRNA start point**

Total RNA isolated from both acid-producing and solvent-producing cells, as well as \textit{E. coli} containing plasmid pH12, was utilized for primer extension analysis. Total RNA isolated from \textit{E. coli} cells which harboured the Bluescript vector pSK (no insert) was used as a negative control, and in this case no primer extension signal was observed (data not shown). The results of the primer extension experiments (Fig. 5) revealed one major signal which indicates that the 5' end of the \textit{hydA} transcript lies 40 bp upstream of the ATG initiation codon, with G being the first transcribed nucleotide (see also Fig. 3). No difference was observed, with respect to the start point or signal intensity, between RNA isolated from acid-producing and solvent-producing \textit{C. acetobutylicum} cells. The same signal was seen with RNA isolated from \textit{E. coli} cells which contained plasmid pH12. A second, slightly stronger signal was also observed with RNA isolated from this \textit{E. coli} clone, which corresponds to a transcript that is one nucleotide shorter than that observed with the \textit{C. acetobutylicum} RNA. The observation that the strongest signal was obtained with RNA isolated from the \textit{E. coli} pH12 clone is most likely a function of the high copy number of the Bluescript plasmid vector.

**DISCUSSION**

The low G+C content of \textit{C. acetobutylicum} and \textit{C. pasteurianum}, as well as a similarly restrictive codon bias within this genus, allowed the isolation of the \textit{C. acetobutylicum} P262 \textit{hydA} gene via colony hybridization.
using the *C. pasteurianum* hydrogenase-1 gene as a probe. In retrospect, we can now calculate that the DNA sequence of the *C. pasteurianum* probe used for colony hybridization shares a 72% identity with the analogous region in the cloned *C. acetobutylicum* hydA gene. Obviously, this level of homology was sufficient to allow cross hybridization under the stringency conditions used.

Enough sequence data were obtained to reveal the presence of an additional ORF upstream of the *hydA* gene, and a 102 amino-acid-length protein sequence could be deduced. Searches of the various protein and nucleic acid databases showed that this 102 amino-acid sequence was similar to the C-terminal end of a protein sequence coded by a DNA region upstream of the *C. pasteurianum* hydrogenase-1 gene. This coding region from the *C. pasteurianum* chromosome (complement of bp 4608–5180; GenBank accession number Z28533) would be transcribed in the opposite direction to the hydrogenase-1 gene from this organism, which is not the case in *C. acetobutylicum*. Although these two organisms may contain analogous genes upstream of their respective hydrogenases, the configurations are somewhat different, and further sequence data and biochemical studies are required to determine their roles and organization.

DNA hybridization studies showed that the *C. acetobutylicum* P262 *hydA* gene did not hybridize to chromosomal DNA isolated from strain NCIMB 8052 under the stringency conditions used. Hybridization signals were observed with chromosomal DNA isolated from *C. acetobutylicum* strains DSM 792, DSM 1731 and ATCC 824 (Fig. 2). In all cases the hybridization patterns obtained using chromosomal DNA from these strains were identical. It appears that the organization of the hydrogenase-encoding regions in these three *C. acetobutylicum* strains is essentially identical. This is in good agreement with the results reported by Wilkinson & Young (1993), where it was shown that *C. acetobutylicum* strains DSM 1731 and ATCC 824 have an almost identical genome organization, and supports the results of Sauer & Dürrè (1993), who showed that the cloned *C. acetobutylicum* DSM 792 hsp18 gene hybridized to identically sized chromosomal DNA fragments from strains DSM 792 and ATCC 824. The hydrogenase-encoding region of the P262 strain is significantly different to yield a hybridization pattern that is distinguishable from the other strains tested, and it is also important to note that the hybridization signal was always weaker with chromosomal DNA from the other strains as compared to chromosomal DNA from the source organism.

The *C. acetobutylicum* P262 *hydA* coding region is 1722 bp in length and encodes for a 574 amino-acid-protein with a predicted molecular mass of approximately 63 kDa. The amino acid sequence of this protein showed considerable similarity only to other [Fe] hydrogenases, particularly the *C. pasteurianum* protein, and supports the observation that [Fe] hydrogenases are unrelated to [NiFe(Se)] hydrogenases (Voordeck et al., 1989). An alignment of the *C. acetobutylicum* HydA amino acid sequence with the *C. pasteurianum* hydrogenase-1 sequence and with the sequence of the putative fourth subunit of the *D. fructosovorans* NAD-reducing hydrogenase is shown in Fig. 6. Overall, these three amino acid sequences are remarkably similar, with the two clostridial sequences being slightly shorter than that from *D. fructosovorans*. The other [Fe] hydrogenase amino acid sequences referred to previously were not included in the alignment because their difference in sequence length would necessitate the introduction of large gaps into the aligned sequences.

Meyer & Gagnon (1991) noted the presence of three domains in the *C. pasteurianum* hydrogenase-1 protein, based on the location of 22 cysteine residues. All 22 residues are conserved in the *C. acetobutylicum* HydA protein sequence, with one additional cysteine residue present at position 175 (Fig. 6). The N-terminal domain, amino acids 1–120 (Fig. 6), contains eight cysteine residues in the two clostridial sequences, and seven cysteine residues in the *D. fructosovorans* sequence. Despite their atypical arrangement, these eight cysteine residues function as ligands for two of four [4Fe–4S] clusters (Adams, 1990; Meyer & Gagnon, 1991). An unusual distribution of cysteine residues in [4Fe–4S] clusters has been noted elsewhere (Hausinger & Howard, 1983; Cunningham et al., 1989; Beinert, 1990). The eight cysteine residues in the central domain (amino acids 140–220, Fig. 6) are distributed in a typical ferredoxin-like pattern and function as ligands for the remaining two [4Fe–4S] clusters. Compared to the *D. fructosovorans* protein and the HydC protein from *D. vulgaris* (not shown), the *C. pasteurianum* and *C. acetobutylicum* proteins contain one and two additional cysteine residues in this domain, respectively. The C-terminal domain, which contains the necessary ligands for the hydrogen activating H cluster, is characterized by five conserved cysteine residues (Fig. 6). Notably, four methionine residues are also conserved between the *C. acetobutylicum* and *C. pasteurianum* proteins. It has been suggested that some or all of these methionine residues could also function as ligands in the H cluster (Meyer & Gagnon, 1991).

Regulation of hydrogen production in *C. acetobutylicum* has been discussed in a number of studies (Gapes et al., 1982; Andersch et al., 1983; Kim & Zeikus, 1985; Yerushalmi & Volesky, 1985; Vasconcelos et al., 1994). Kim & Zeikus (1985) observed that the specific rate of hydrogen production decreased in stages during the course of a batch fermentation, where the highest rate of hydrogen production occurred during the initial exponential growth phase. These authors noted a significant decrease in the rate of hydrogen production which correlated with the shift from acid production to solvent production. The specific hydrogenase activity in whole cells from acid-producing cultures was shown to be approximately 22-fold higher than that measured from solvent-producing cultures (Kim & Zeikus, 1985). It was concluded that because hydrogenase activity was not affected by either pH or acid concentration, the decrease in hydrogen production during solventogenesis is a result of regulation of hydrogenase production as opposed to an inhibition of enzyme activity. In contrast to these conclusions, Andersch et al. (1983) reported that hydro-
Hydrogenase activity was similar in both acid- and solvent-producing cells. Hydrogenase activity in solvent-producing cells was present in an inactive form and the conditions included that the hydrogenase from the solvent-producing cells could only be detected in this assay after an initial lag period. The results presented here in Figs. 4 and 5 support this hypothesis, as no difference was noted in the transcript size and expression level of bydA gene from C. acetobutylicum and C. pasteurianum sequences; +, additional cysteine residues in the C. acetobutylicum and C. pasteurianum sequences; =, conserved methionine residues. The two regions containing cysteine residues; =, conserved methionine residues. The two regions containing cysteine residues in a [4Fe-4S] ferredoxin-like pattern are underscored. In all cases the transcript size and expression level was equivalent to the results in Fig. 4. It can therefore be concluded that the bydA gene from C. acetobutylicum is constitutively expressed from a monocistronic operon throughout the course of a typical batch fermentation.

The 5' end of the bydA transcript was identified via primer extension, and these results indicated that only one transcription initiation site is utilized in both acid-producing and solvent-producing C. acetobutylicum cells (Fig. 5). Based on these primer extension results, a putative promoter sequence from positions 381 to 409 could be deduced (TATATT-17 bp-TAAAT; see Fig. 3); the proposed Gram-positive promoter consensus sequence is TTGACA-16 to 18 bp-TATAAT (Graves & Rabinowitz, 1986). Conservation of nucleotide sequences starting 3 bp upstream of the transcription initiation site is utilized in both acid-producing and solvent-producing C. acetobutylicum cells. Apart from the traditional promoter core regions, it is apparent that in Gram-positive organisms, upstream and downstream regions also display conserved sequences (Graves & Rabinowitz, 1986). Conservation of nucleotide sequences in the extended regions can compensate for a lack of conservation in the core regions. For example, a TG pair starting 3 bp upstream of the −10 region has been shown to be sufficient to compensate for the absence of a specific −35 region (Ponnambalam et al., 1986; Keilty & Rosenberg, 1987). The proposed −35 region for the bydA gene is indeed very poorly conserved.
there is a TG pair located 3 bp upstream of the well conserved — 10 region (see Fig. 3). The similarity of this putative promoter region to consensus sequences for vegetative sigma factors in Gram-positive bacteria and E. coli adds further support to the conclusion that the hydA mRNA is constitutively produced by C. acetobutylicum.

ACKNOWLEDGEMENTS

J. Santangelo acknowledges financial support obtained from the Alexander von Humboldt Stiftung, which allowed him to continue this work in Germany. Thanks are also owed to Jaques Meyer who provided the C. pasteurianum hydrogenase-1 clone which was used to screen the C. acetobutylicum gene library.

REFERENCES


Received 20 June 1994; revised 30 August 1994; accepted 28 September 1994.