Journal of Industrial Microbiology & Biotechnology (2001) 27, 281–286
© 2001 Nature Publishing Group 1367-5435/01 $17.00

www.nature.com/jim

Nitrogen-fixation genes and nitrogenase activity in *Clostridium acetobutylicum* and *Clostridium beijerinckii*

J-S Chen, J Toth and M Kasap

Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

Several solvent-producing clostridia, including *Clostridium acetobutylicum* and *C. beijerinckii*, were previously shown to be nitrogen-fixing organisms based on the incorporation of $^{15}N_2$ into cellular material. The key nitrogen-fixation (*nif*) genes, including *nifH*, *nifD*, and *nifK* for nitrogenase component proteins as well as *nifE*, *nifN*, *nifB* and *nifV* for synthesis of the iron–molybdenum cofactor (FeMoco) of nitrogenase, have now been identified in *C. acetobutylicum* or *C. beijerinckii* or both. The organization of these genes is similar to the distinctive pattern that was first observed in *Clostridium pasteurianum*, with the *nifN* and *nifB* genes fused into the *nifN-B* gene and with the *nifV* gene split into the *nifV*ω and *nifV*α genes. The corresponding *nif* genes of these three clostridial species are highly related to each other. However, in the two solvent-producing clostridia, the *nifH* and *nifD* genes are interspersed by two *glnB*-like genes, which are absent in the corresponding region in *C. pasteurianum*. However, the *nifN-B* and *nifV*ω genes of *C. pasteurianum* are interspersed by the putative *modA* and *modB* genes (for molybdate transport), which are absent in the corresponding region in *C. acetobutylicum*. *C. acetobutylicum* and *C. beijerinckii* grew well under nitrogen-fixing conditions, and the acetylene-reducing activity of nitrogenase was measured in the two species. Acetone, butanol, and isopropanol production occurred in nitrogen-fixing cultures, but the peak of nitrogen-fixing activity preceded the active solventogenic phase. *Journal of Industrial Microbiology & Biotechnology* (2001) **27,** 281–286.

Keywords: *nif*; nitrogen fixation; solvent-producing clostridia

Introduction

Nitrogen fixation and alcohol production (the production of butanol, ethanol, and isopropanol) are two reductive metabolic processes that can conceivably compete against each other for the same source of reductant in the cell, if the two processes occur simultaneously. In the solvent-producing clostridia, active production of butanol and isopropanol normally occurs at a stage subsequent to the exponential phase of growth. If the solvent-producing clostridia could fix N_2 during the exponential phase of growth, nitrogen gas may become an inexpensive source of nitrogen for the growth of the solvent-producing clostridia. However, if the two processes overlap during solvent production, they may be utilized as an experimental system for the study of the relationship between the supply of reductants and the yield of nitrogen fixation or solvent production.

Several solvent-producing clostridia, including *Clostridium acetobutylicum* and *Clostridium beijerinckii*, were previously found to be nitrogen-fixing organisms, when the incorporation of $^{15}N_2$ was measured [10]. *Clostridium butylicum*, which has since been identified as *C. beijerinckii* was also reported as a nitrogen-fixing organism. Among the free-living nitrogen-fixing organisms, *Clostridium pasteurianum* was the first to be isolated [18], and its nitrogen-fixation (*nif*) genes have been characterized [2]. Following the sequencing of the genome of *C. acetobutylicum* ATCC 824 (http://www.cric.com/genesequences/clostridium/clospage.html), a region encompassing nucleotides 284,400 to 295,096 was found to contain open-reading frames (ORFs) that encode amino acid sequences highly similar to those encoded by the *nif* genes of *C. pasteurianum*. The identification of the *nif*-related genes in *C. acetobutylicum* is consistent with the previously observed nitrogen-fixing activity of this organism.

In order to utilize the nitrogen-fixing activity as an experimental tool for study of the regulation of solvent production in the clostridia, we examined the nitrogen-fixing ability of the currently recognized strains of *C. acetobutylicum* and *C. beijerinckii* and also searched for the presence of *nif* genes in *C. beijerinckii*. In this paper, we report the organization of the *nif* genes in *C. beijerinckii* and the growth and solvent production of the clostridia under nitrogen-fixing conditions.

Correspondence: Dr J-S Chen, Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
Received 2 September 2000; accepted 22 November 2000

Materials and methods

Bacteria, plasmids, and growth conditions

Spores of *C. acetobutylicum* ATCC 824, *C. beijerinckii* NRRL B593, and *C. pasteurianum* W5 were from laboratory stocks, and routine culture conditions were as reported [5,13]. Plasmids LITMUS 28 and 29 (New England BioLabs, Beverly, MA) and pUC19 (Life Technologies, Rockville, MD) were maintained as recommended by the manufacturers.

Nitrogen-fixing cultures

Nitrogen-fixing cultures were grown in a medium previously described [15], except that sucrose was replaced by glucose. When needed, yeast extract was included at 0.5 or 1.0 g/l. Cultures were grown at 35°C under N_2 in rubber-stoppered 160-ml serum bottles containing 20 ml of medium. *C. beijerinckii* was also grown in a defined medium containing sucrose, 60 g/l; dibasic, potassium phosphate, 3.5 g/l; mineral solution [15], 1 ml/l; cysteine, 0.5 g/l; alanine, 0.027 g/l; isoleucine, 0.016 g/l; leucine, 0.023 g/l;

proline, 0.013 g/l; valine, 0.019 g/l, and a mixture of B vitamins (45 mg/l each of vitamins B-1, B-2, B-6, niacinamide, pantothenic acid, choline bitartrate, inositol, *p*-aminobenzoic acid; 45 μg/l each of vitamin B-12 and biotin; 180 μg/l of folic acid).

Acetylene reduction assay

The reduction of acetylene to ethylene by nitrogenase was measured by gas chromatography using a Poropak N column. The assay was performed in a 9.2-ml vial that contained 0.95 ml of bacterial culture and 0.5 ml of a 20% (w/v) solution of glucose in 1 M potassium phosphate buffer at pH 7.2. One milliliter of argon in the gas phase was replaced by acetylene. The vials were incubated at 30°C with shaking, and 200-μl samples were periodically taken for the measurement of ethylene formation. When cultures were grown in the defined medium with 60 g/l of sucrose, the culture samples were assayed directly without addition of glucose.

DNA cloning, amplification, and sequencing

Protocols for the isolation of DNA from *C. beijerinckii* NRRL B593 and Southern analysis were previously reported [13]. Depending on the properties of the primers, PCR was performed within the following range of conditions: denaturation temperature, 92–95°C; annealing temperature, 42–60°C; elongation temperature, 70–72°C. Either *Taq* DNA polymerase (Qiagen, Valencia, CA) or herculase enhanced polymerase (Stratagene, La Jolla, CA) was used. The probe for the *nifH* gene was generated using PCR to amplify a conserved region of the *C. beijerinckii nifH* gene. The primer pair was 5′-GGWTGTGAYCCWAAGGCWG (corresponding to the amino acid sequence GCDPKAD) and 5′-AKWGCCATCATYTCWCC (complementry to the sequence GEMMAL). The PCR product was sequenced to confirm its identity. A 2.4-kb *Hin*dIII fragment was identified by the 350-bp probe and was cloned into LITMUS 28 to give pMK40. The insert in pMK40 was sequenced, and it contained the entire *nifH* gene and two downstream genes.

To amplify a longer fragment containing the downstream region of the *nifH* gene, *Pst*I-digested genomic DNA was ligated to pUC19, and the ligation mixture was used as template in the PCR. The primer pair was 5′-CATATACACCTGATTTAGACTATG (corresponding to the NifH sequence YTPDLDY) and an M13 universal primer. The PCR product was sequenced.

Attempts to amplify a fragment containing the *nifE* to *nifV*ω region yielded multiple products. However, a fragment containing the *nifE* and *nifN-B* genes was amplified from the mixture, using the primer pair 5′-CTTAATATTGTTCAATGTGC (corresponding to the sequence LNIVQCA) and 5′-GCTWGCCATTCTTCTTC-CATC (complementary to DGRRMAS). The PCR product was sequenced.

DNA sequencing was performed by either the DNA Sequencing Facility, Virginia-Maryland Regional College of Veterinary Medicine, Blacksburg, VA, or Davis Sequencing of Davis, CA.

Analysis of DNA sequences

Both the MacVector (Genetics Computer Group, Madison, WI) and the Lasergene (DNASTAR, Madison, WI) software packages were used for the management of DNA sequences. BLAST programs, provided by the National Center for Biotechnology Information, were used in searching databases for related amino acid sequences.

Nucleotide sequence accession number

The sequence of the 2391-bp insert in pMK40 has been deposited in the GenBank nucleotide sequence database under accession number AF266462.

Results

The nif cluster of C. acetobutylicum

From the nucleotide sequence of the chromosome of *C. acetobutylicum* ATCC 824, a cluster of ORFs was identified as putative *nif* genes because of their similarity to the *nif* genes of *C. pasteurianum* and other organisms (GenBank accession number AE001437). This cluster is located between nucleotide (nt) number 284,400 and 295,096 (Figure 1). The assigned ORF number and the position (from–to) number for the ORFs as well as the length of the deduced polypeptides are listed below. It should be noted that the position numbers for some ORFs encompass a length somewhat greater than the needed coding capacity for the deduced polypeptides. For the nitrogenase component proteins, *nifH* (CAC317; nt 284,400–285,230; 272 amino acid residues) encodes the iron protein, *nifD* (CAC320; nt 286,090–287,706; 531 amino acid residues) encodes the α subunit of the molybdenum–iron protein, and *nifK* (CAC321; nt 287,700–289,070; 454 amino acid residues) encodes the β subunit of the molybdenum protein. Downstream from the nitrogenase genes are the genes encoding proteins for the synthesis of the iron–molybdenum cofactor of nitrogenase: *nifE* (CAC322; nt 289,161–290,567; 466 amino acid residues), *nifN-B* (CAC323; nt 290,568–293,231; 867 amino acid residues), *nifV*ω (CAC324; nt 293,245–294,318; 356 amino acid residues) and *nifV*α (CAC325; nt 294,212–295,096; 265 amino acid residues). The structure and organization of *nif* genes are strikingly similar in *C. acetobutylicum* and *C. pasteurianum*, including the characteristically fused *nifN-B* gene [Ref. [2]; S.-Z. Wang, S. Xiong, and J.-S. Chen, unpublished results] and the split *nifV* genes [17]. In *C. acetobutylicum*, the *nifV*ω and the *nifV*α genes, which encode homocitrate synthase, overlap by 20 bases.

There are also significant differences in the *nif* clusters of *C. acetobutylicum* and *C. pasteurianum*. Between the *nifH* and *nifD* genes of *C. acetobutylicum*, two ORFs, which are designated the *glnB*-like gene 1 (CAC318; nt 285,350–285,700; 106 amino acid residues) and *glnB*-like gene 2 (CAC319; nt 285,634–286,089; 125 amino acid residues), are present, whereas the *C. pasteurianum nif* cluster does not contain these genes. The *glnB*-like genes encode amino acid sequences highly similar to the *glnB*-encoded nitrogen-regulatory protein P_{II} of *Escherichia coli* [12] and the ORFs present between *nifH* and *nifD* of several methanogens [3,11] and *Clostridium cellobioparum* (GenBank accession number AAB63258). However, there are two ORFs, now designated *modA* (261 amino acid residues; J. Toth, S.-Z. Wang, and J.-S. Chen, unpublished results) and *modB* (formerly *nifC*; 286 amino acid residues) [16], that are located between *nifN-B* and *nifV*ω genes in *C. pasteurianum* but not in the corresponding region in *C. acetobutylicum*. However, a *modA*-like ORF (CAC357; 322,061–321,210; 284 amino acid residues) and a *modB*-like ORF (CAC356; 321,189–320,380; 270 amino acid residues) are located about 25 kb downstream from the *nif* cluster and in the opposite orientation on the *C. acetobutylicum* chromosome. The *modA* and *modB* genes are members of a

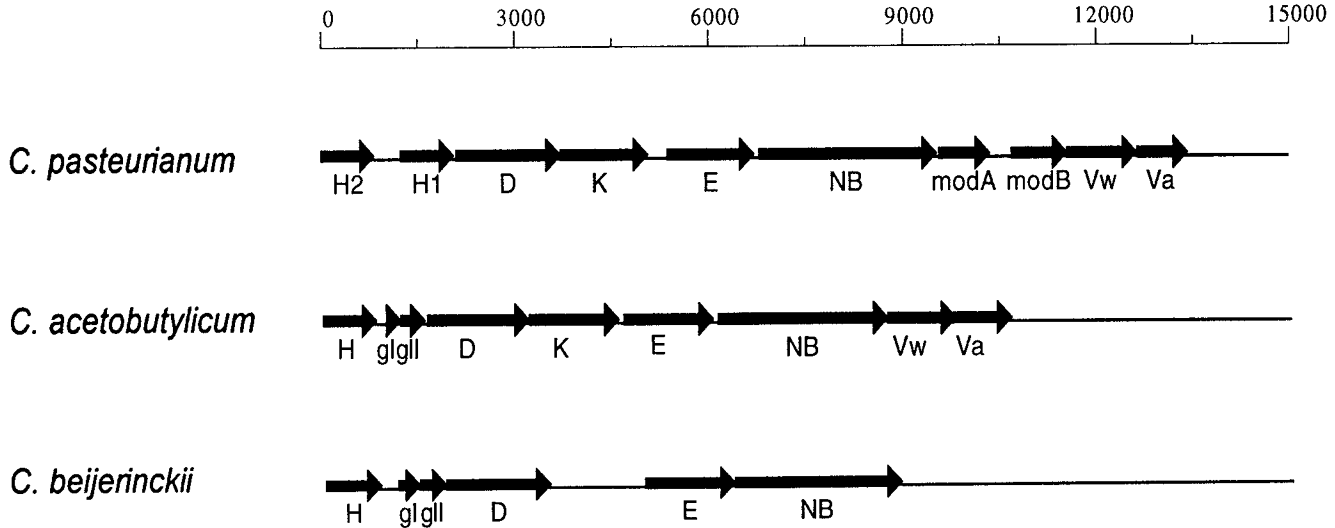

Figure 1 Physical maps of the *nif* clusters of *C. pasteurianum* W5 [Ref. [2], and this study], *C. acetobutylicum* ATCC 824 [GenBank accession number AE001437, and this study], and *C. beijerinckii* NRRL B593 [this study]. Symbols designate the following genes: H2, *nifH2*; H1, *nifH1*; H, *nifH*; gI, *glnB*-like gene 1; gII, *glnB*-like gene 2; D, *nifD*; K, *nifK*; E, *nifE*; NB, *nifN-B*; Vw, *nifVω*, Va, *nifVα*. The scale bar indicates length in base pairs.

proposed molybdate transport system [8]. In addition, the multiple copies of *nifH* genes of *C. pasteurianum* [15] are not present in *C. acetobutylicum*.

The nif cluster of C. beijereinckii

The organization of the *nif* genes in *C. beijerinckii* is similar to that found in *C. acetobutylicum*, as deduced from the nucleotide sequences and the size of the PCR products. The sequenced portion of the *nif* cluster of *C. beijerinckii* NRRL B593 is shown in Figure 1. These sequences were determined from a cloned restriction fragment as well as from overlapping DNA fragments generated by PCR.

The *nifH* gene, the complete *glnB*-like gene 1, and part of the *glnB*-like gene 2 are located on a 2391-bp *Hin*dIII fragment (GenBank accession number AF266462), which was initially detected by Southern hybridization using an amplified 350-bp fragment of the *C. beijerinckii nifH* gene as a probe (data not shown).

Attempts to generate a fragment encompassing the *nifE* and *nifVω* genes by PCR always gave multiple products. Using the PCR-generated multiple products as template, it was possible, however, to amplify the *nifE–nifN* region for direct sequencing. Using a primer based on the *nifH* sequence and a universal M13 primer, a 1.5-kb *Pst*I fragment was amplified from a ligation mixture containing *Pst*I-digested fragments of the genomic DNA ligated to pUC19. This 1.5-kb fragment was flanked by partial *nifH* and *nifD* genes, between which the two *glnB*-like genes are located.

Once the nucleotide sequences for portions or entire *nifH*, *glnB*-like genes 1 and 2, *nifD*, and *nifN-B* genes of *C. beijerinckii* were available, PCR could be used to generate additional DNA fragments that encompass the following regions for sequence analysis: *glnB*-like gene 2 to *nifD*, *nifD* to *nifN-B*, and *nifN-B* to *nifVα*. Except for the *nifK* gene and the *nifV* region, which are yet to be sequenced, the organization of the other *nif* genes of *C. beijerinckii* are proposed as shown in Figure 1.

Comparison of GlnB-like proteins of C. acetobutylicum and C. beijerinckii

The GlnB-like protein 1 of *C. acetobutylicum* ATCC 824 (106 amino acid residues, pI=8.345) and the GlnB-like protein 1 of *C. beijerinckii* NRRL B593 (108 amino acids; pI=6.714) share a 62.6% positional identity at the amino acid level. They are 56.9% and 60.9%, respectively, identical to the GlnB of *C. cellobioparum* (accession number AAB 63258). An examination of the eight amino acid sequences that are encoded by the *glnB*-like gene 1 occurring between the *nifH* and *nifD* genes in the clostridia and the methanogens revealed a conserved region flanking the proposed active-site tyrosine residue (residue 47 and 49, respectively, of *C. acetobutylicum* and *C. beijerinckii* GlnB-like protein 1), which resembles the corresponding region in the GlnB protein of *E. coli* (Figure 2).

The active-site tyrosine in the *E. coli* GlnB (Tyr-51 in GlnB or P_{II} protein) is reversibly uridylylated to inactivate and activate this protein for the regulation of glutamine synthetase and the NtrB (NR_{II}) protein in nitrogen metabolism [12]. This region contains the GxGxxxG motif, which precedes the tyrosine residue and may be involved in nucleotide binding. It may be proposed that the product of the *glnB*-like gene 1, which is located between the *nifH* and *nifD* genes of some clostridia and methanogens, may play a role similar to that of the *E. coli* protein. However, the spacing (seven versus nine residues) between the GxxGxxxG motif and the tyrosine residue as well as the hydrophobicity of the intervening amino acid residues differ between the *E. coli* protein and the putative protein from the anaerobes. It is thus possible that the GlnB-like protein 1 of the solvent-producing clostridia has a somewhat different function than the *E. coli* protein. Because the *nif* cluster of *C. pasteurianum* does not contain a *glnB*-like gene, it should be useful to search the *C. pasteurianum* genome for a similar gene. The presence of a similar gene(s) in *C. pasteurianum* would suggest a shared regulatory function among these clostridia.

The GlnB-like protein 2 of *C. acetobutylicum* ATCC 824 (125 amino acids; pI=5.212) and the GlnB-like protein 2 of *C. beijerinckii* NRRL B593 (127 amino acids; pI=6.665) share a

Proposed active site of the GlnB-like protein 1 encoded by a gene occurring between the *nifH* and *nifD* genes

```
* *       *                   ◆
GRGKQ(KR)G(ILV)x(ILV)xx(IV)xYDE(IL)
```

Proposed active site of the *E. coli* GlnB (P_{II}) protein

```
* *   *       ◆
GFGRQKGHTELYRGAEYMVD
```

Figure 2 A comparison of the proposed active site of the GlnB-like protein 1, which is encoded by a gene occurring between the *nifH* and *nifD* genes, and the proposed active site of the *E. coli* GlnB (P_{II}) protein. The consensus sequence encompasses the GlnB-like proteins of the following organisms (with the accession number given in parenthesis): GlnB-like protein 1 of *M. barkeri*, *nifHD1* region (P54808); GlnB-like protein 1 of *M. sarcina*, *nifHD2* region (P54807); GlnB-like protein 1 of *Methanobacterium ivanovii*, *nifHD2* region (P51603); GlnB-like protein 1 of *Methanococcus maripaludis*, *nifHD* region (P71524); GlnB-like protein 1 of *Methanococcus thermolithotrophicus*, *nifHD* region (P25771); GlnB of *C. cellobioparum* (AAB63258); GlnB-like protein 1 of *C. acetobutylicum* (AE001437), and GlnB-like protein 1 of *C. beijerinckii* (this study). The *E. coli* GlnB sequence is from Son and Rhee [12]. The active-site tyrosine residue is marked by ◆. The glycine residues within the GxGxxxG motif are marked by *.

50% positional identity at the amino acid level. They are 30.4% and 31.0%, respectively, identical to the GlnB protein of *C. cellobioparum*. The GlnB-like protein 2 of *C. acetobutylicum* and *C. beijerinckii* differ from the protein from the methanogens: the protein of these two solvent-producing clostridia contains the proposed active-site tyrosine, located 11 amino acids after a GRG motif, whereas the protein of the methanogens (with the exception of one of the two GlnB-like proteins 2 of *Methanosarcina barkeri*; Swissport accession number P54807) does not contain this tyrosine residue. It is conceivable that the GlnB-like protein 2 has a function different from that of the GlnB-like protein 1.

Comparison of nif genes of C. acetobutylicum and C. beijerinckii

The sequence of the *nifH* gene of *C. beijerinckii* NRRL B593 has been determined (accession number AF 266462). The *nifH* gene encodes the nitrogenase iron protein, which is one of the most conserved proteins [19]. It is hence not surprising that the amino acid sequences of the *C. acetobutylicum* ATCC 824, *C. beijerinckii* NRRL B593, and *C. pasteurianum* W5 NifH display a positional identity between 79.8% and 83.2%. The *nifE* gene encodes a polypeptide homologous to the α-subunit (NifD) of the nitrogenase MoFe protein. Together with NifN, the $NifE_2$ $NifN_2$ tetramer serves as the scaffold for biosynthesis of the iron–molybdenum cofactor of nitrogenase [4]. The NifE protein (about 460 amino acids in length) is also highly conserved. For the region of 436 amino acid residues so far sequenced (lacking the N-terminal region), the NifE of *C. beijerinckii* NRRL B593 is 60.8% and 65.7%, respectively, identical to the NifE of *C. acetobutylicum* ATCC 824 and *C. pasteurianum* W5.

In *C. pasteurianum*, the *nifN* and *nifB* genes are fused to have a coding capacity for a polypeptide of 929 amino acids [2]. This unusual feature is also present in *C. acetobutylicum* ATCC 824 and *C. beijerinckii* NRRL B593. Between the conserved domains of NifN and NifB, a stretch of amino acids (the linker) connects these two domains to form the fused NifN-B protein. An examination of the linker region in the NifN-B of *C. acetobutylicum* ATCC 824 (26 amino acids), *C. beijerinckii* NRRL B593 (42 amino acids) and *C. pasturianum* W5 (49 amino acids) reveal an interesting relationship (Figure 3). The linker in the NifN-B of *C. beijerinckii* and *C. pasteurianum* is significantly more similar to each other than to the linker of *C. acetobutylicum* NifN-B. However, there are also amino acid residues that are shared only between the NifN-B of *C. acetobutylicum* and *C. beijerinckii* or between the NifN-B of *C. acetobutylicum* and *C. pasteurianum*. The linker region of these three *nifN-B* genes may be useful for deduction of the evolutionary path for *nif* genes in the clostridia, which can be compared with that deduced from the 16S rRNA gene sequences [7].

Nitrogen-fixing growth and nitrogenase activity of C. acetobutylicum and C. beijerinckii

Both *C. acetobutylicum* ATCC 824 and *C. pasteurianum* W5 (used as control) grew well after serial transfers in a defined medium that contained no yeast extract or other fixed nitrogen. *C. beijerinckii* NRRL B593 grew well in the defined medium supplemented with 0.5 or 1 g/l of yeast extract. Twenty-four-hour cultures of these three species reached a similar level of optical density when grown in the same medium. Acetylene-reducing activity was detected in *C. acetobutylicum* ATCC 824 and *C. beijerinckii* NRRL B593 cells after 24 h of growth; the level of acetylene-reducing activity was variable, however, when cells from different sets of growth experiments were tested (data not shown). The measured acetylene-reducing activity in 24-h cultures of *C. acetobutylicum*

```
450       460       470       480       490       500
 *         *         *         *         *         *
FRENLYNEYY------NEETEINNEVTVIEGEKKMEARVISKEEV-EEKTKTHPC     C. beijerinckii
||   |              || |   |   |   |  | ||      |  ||||
FRKDSYDRFFIPKKEGEDNMEIANKTNVTIEEPLVEEKVLSKKELNAKKSCTHPC     C. pasteurianum
||||                |   |                       | |||||
FRKDAFNNYY---------LESVEK--------------MKADNMKEEKTCTHPC     C. acetobutylicum
      ^ ^^                                    ^^ ^
```

Figure 3 The linker region within the fused NifN-B polypeptides of *C. acetobutylicum*, *C. beijerinckii*, and *C. pasteurianum*. Numbering is based on the amino acid sequence of the *C. pasteurianum* NifN-B polypeptide. The NifN domain occupies the N-terminal half of the protein. Amino acid residues that are shared only between the *C. acetobutylicum* and *C. beijerinckii* sequences are marked by ^.

ATCC 824 and *C. beijerinckii* NRRL B593 ranged from about 1% to 50% of that of *C. pasteurianum* W5 cells, with *C. beijerinckii* cells usually exhibiting a higher activity than *C. acetobutylicum* cells. The presence of up to 1 g/l of yeast extract in the growth medium did not repress nitrogenase activity in *C. beijerinckii* or *C. pasteurianum*. In three sets of growth studies with *C. acetobutylicum*, high levels of acetylene-reducing activity were measured only in cells grown in the absence of yeast extract.

When *C. beijerinckii* NRRL B593 was grown in a defined medium in which yeast extract was replaced by vitamins and amino acids, the acetylene-reducing activity decreased 10-fold when the growth rate slowed down and the culture entered into the solventogenic phase (data not shown). This preliminary observation may explain the fluctuation in acetylene-reducing activities in 24-h-old cultures, as the small batch cultures could have different lag periods and 24-h-old cultures were normally in the solventogenic phase already. Further experiments are under way to measure the expression of *nif* and *sol* genes and the corresponding enzymic activities in cells at different stages of growth.

Discussion

The *nif* genes represent a highly conserved group of genes that are present in diverse bacterial taxons. Although the *nifH* gene is the most conserved among the *nif* genes, phylogenetic diversity in the *nifH* gene is present in the free-living *Bacillaceae* [1], the symbiotic rhizobia [19], and the symbiotic microbial community in the gut of termites [9]. The organization of the *nif* genes is, however, much more diverse than the diversity of *nif* genes themselves [4]. A study of the structure and organization of the *nif* genes may therefore help shed light on the evolution of the genome within a bacterial group.

Among the nitrogen-fixing organisms, the clostridia appear to have the most concise cluster of *nif* genes [4]. The fused *nifN-B* gene and the split *nifV* gene are landmarks of the clostridial nitrogen-fixation system. An examination of the *nif* cluster of *C. acetobutylicum*, *C. beijerinckii*, and *C. pasteurianum* revealed two interesting differences: (1) the two solvent-producing species, like some methanogens, have two *glnB*-like genes occurring in tandem between the *nifH* and *nifD* genes, whereas *C. pasteurianum* does not have these genes in its *nif* cluster, and (2) in *C. pasteurianum*, the *nifN-B* gene is followed by the *modA* and *modB* genes, whereas these genes for molybdate transport occur elsewhere in the genome of *C. acetobutylicum*. Because the *glnB*-like genes may serve an important regulatory function in nitrogen metabolism, it should be useful to search for the presence of similar genes in the genome of *C. pasteurianum* and to conduct genetic analysis to determine the function of the *glnB*-like genes in the clostridia.

For the two solvent-producing species *C. acetobutylicum* and *C. beijerinckii*, information on the structure and organization of three important groups of genes is either available now or will be available in the near future. These three groups are (1) the SOL or solvent-production genes, (2) the BCS or butyric acid-production genes, and (3) the NIF or nitrogen-fixation genes. Among the three groups, the BCS genes are required for the production of butyric acid, which is a primary metabolic function for the butyric clostridia and the genes could be considered as housekeeping genes for these organisms. However, the SOL and NIF activities are involved in secondary metabolism. The three groups of genes could have been acquired by these organisms at different points of time and subjected to different selection pressures. The difference in the linker region of the fused *nifN-B* gene (Figure 3), for example, may be a marker for an analysis of the evolution of the structure and function of the nitrogen-fixation proteins. The genes for solvent-production are located on a plasmid in *C. acetobutylicum*, whereas the genes for butyric acid production and nitrogen fixation are located on the chromosome in this species. The different locations may also affect the rate of evolution and may be the subject of further analysis.

It is well known that growth conditions can have a profound effect on the outcome of solvent fermentation [6]. Because active solvent production occurs late during growth, it is intriguing how the earlier nutritional status of a cell affects its solvent-producing capacity. That the solvent-producing clostridia are also nitrogen-fixing organisms may provide a useful approach for probing this relationship. The nitrogen-fixation reaction catalyzed by nitrogenase is an energy- and reductant-intensive process, and under nitrogen-fixing conditions an organism grows at a slower rate than when fixed nitrogen is available. The expression of nitrogen-fixing activity is regulated at both the transcriptional and the enzymic activity levels, and it responds rapidly to the presence of fixed nitrogenous compounds [14]. When a nitrogen-fixing culture of *C. acetobutylicum* or *C. beijerinckii* is given a sudden supply of fixed nitrogen, such as ammonium ion or amino acids, its nitrogen-fixing activity can be expected to fall quickly. This decrease in nitrogen-fixing activity should be followed by a change in the energy charge and redox states as well as the growth rate of the cell. Using this approach, it may be possible to differentiate the effect of energy charge and redox states and growth rate on the onset or progression of solventogenesis. It may also allow an analysis on how the different nitrogenous compounds affect solvent productivity. Such information should contribute to the development of a better solvent fermentation.

Acknowledgements

This study was supported by the U.S. Department of Energy grant DE-FG02-97-ER20276 and by the CSREES, U.S. Department of Agriculture, under project 6122000. We thank Farzana Ahmed and Kanit Vichitphan for participation and assistance in the measurement of nitrogen-fixing growth of the solvent-producing clostridia. We also thank Jork Nolling and Douglas Smith of the Genome Therapeutics for the timely release of the nucleotide sequence and the annotation table for the genome of *Clostridium acetobutylicum* ATCC 824.

References

1 Achouak W, P Normand and T Heulin. 1999. Comparative phylogeny of *rrs* and *nifH* genes in the *Bacillaceae*. *Int J Syst Bacteriol* 49: 961–967.
2 Chen JS and JL Johnson. 1993. Molecular biology of nitrogen fixation in the clostridia. In: Woods DR (Ed), The Clostridia and Biotechnology. Butterworth-Heinemann, Boston, pp. 371–392.
3 Chien YT and SH Zinder. 1994. Cloning, DNA sequencing and characterization of a *nifD*-homologous gene from the archaeon *Methanosarcina barkeri* 227 which resembles *nifD1* from the eubacterium *Clostridium pasteurianum*. *J Bacteriol* 176: 6590–6598.
4 Dean DR and MR Jacobson. 1992. Biochemical genetics of nitrogenase. In: Stacy G, RH Burris and HJ Evans (Eds), Biological Nitrogen Fixation, Chapman & Hall, New York, pp. 763–834.

5 Johnson JL, J Toth, S Santiwanakul and JS Chen. 1997. Cultures of "*Clostridium acetobutylicum*" from various collections comprise *Clostridium acetobutylicum*, *Clostridium beijerinckii* and two other distinct types based on DNA–DNA reassociation. *Int J Syst Bacteriol* 47: 420–424.
6 Jones DT and DR Woods. 1986. Acetone–butanol fermentation revisited. *Microbiol Rev* 50: 484–524.
7 Keis S, CF Bennett, VK Ward and DT Jones. 1995. Taxonomy and phylogeny of industrial solvent-producing clostridia. *Int J Syst Bacteriol* 45: 693–705.
8 Maupin-Furlow JA, JK Rosentel, JH Lee, U Deppenmeier, RP Gunsalus and KT Shanmugam. 1995. Genetic analysis of the *modABCD* (molybdate transport) operon of *Escherichia coli*. *J Bacteriol* 177: 4851–4856.
9 Ohkuma M, S Noda and T Kudo. 1999. Phylogenetic diversity of nitrogen fixation genes in the symbiotic microbial community in the gut of diverse termites. *Appl Environ Microbiol* 65: 4926–4934.
10 Rosenblum ED and PW Wilson. 1949. Fixation of isotopic nitrogen by *Clostridium*. *J Bacteriol* 57: 413–414.
11 Sibold L, M Henriquet, O Possot and J-P Aubert. 1991. Nucleotide sequence of *nifH* regions from *Methanobacterium ivanovii* and *Methanosarcina barkeri* 227 and characterization of *glnB*-like genes. *Res Microbiol* 142: 5–12.
12 Son HS and SG Rhee. 1987. Cascade control of *Escherichia coli* glutamine synthetase — purification and properties of P_{II} protein and nucleotide sequence of its structural gene. *J Biol Chem* 262: 8690–8695.
13 Toth J, AA Ismaiel and JS Chen. 1999. The *ald* gene, encoding a coenzyme A-acylating aldehyde dehydrogenase, distinguishes *Clostridium beijerinckii* and two other solvent-producing clostridia from *Clostridium acetobutylicum*. *Appl Environ Microbiol* 65: 4973–4980.
14 Upschurch RG and LE Mortenson. 1980. In vivo energetics and control of nitrogen fixation: changes in the adenylate energy charge and adenosine 5′-diphosphate/adenosine 5′-triphosphate ratio of cells during growth on dinitrogen versus growth on ammonia. *J Bacteriol* 143: 274–284.
15 Wang SZ, JS Chen and JL Johnson. 1988. The presence of five *nifH*-like sequences in *Clostridium pasteurianum*: sequence divergence and transcription properties. *Nucleic Acids Res* 16: 439–454.
16 Wang S, JS Chen and JL Johnson. 1990. A nitrogen-fixation (*nifC*) gene in *Clostridium pasteurianum* with sequence similarity to *chlJ* of *Escherichia coli*. *Biochem Biophys Res Commun* 169: 1122–1128.
17 Wang SZ, DR Dean, JS Chen and JL Johnson. 1991. The N-terminal and C-terminal portions of NifV are encoded by two different genes in *Clostridium pasteurianum*. *J Bacteriol* 173: 3041–3046.
18 Winogradsky MS. 1895. Recherches sur l'assimilation de l'azote libre de l'atmosphere par les microbes. *Arch Sci Biol (St Petersburg)* 3: 297–352.
19 Young JPW. 1993. Molecular phylogeny of rhizobia and their relatives. In: Palacios R, J Mora and WE Newton (Eds), New Horizons in Nitrogen Fixation. Kluwer Academic Publishers, Dordrecht, pp. 587–592.