Cloning and nucleotide sequence of the *Butyrivibrio fibrisolvens* gene encoding a type III glutamine synthetase

HEIDE J. K. GOODMAN and DAVID R. WOODS*

Department of Microbiology, University of Cape Town, Rondebosch 7700, South Africa

(Received 5 January 1993; revised 26 February 1993; accepted 15 March 1993)

A *Butyrivibrio fibrisolvens* glnA gene encoding glutamine synthetase (GS) was cloned on a recombinant plasmid pGS4 which enabled *Escherichia coli* glnA deletion mutants to utilize (NH₄)₂SO₄ as a sole source of nitrogen. The nucleotide sequence of a 2423 bp DNA segment containing the GS-coding region of *B. fibrisolvens* was determined and the complete amino acid sequence (701 residues) was deduced. Comparisons of the derived *B. fibrisolvens* GS protein sequence with the amino acid sequences of GS from other bacteria indicate that it is the second reported example of a type III GS, originally identified in the obligate anaerobe *Bacteroides fragilis*. The presence of GS in *B. fibrisolvens* cells and the regulation of the cloned GS in *E. coli* cells was demonstrated by Western blot analysis.

Introduction

The genus *Butyrivibrio* is composed of Gram-negative anaerobic bacteria that ferment a wide variety of carbohydrates with the production of large amounts of butyric acid. *Butyrivibrio fibrisolvens* is a ubiquitously distributed bacterial species in the gastro-intestinal tracts of ruminants and other mammals.

Ammonia is the major and often a required source of nitrogen for most rumen bacteria (Bryant & Robinson, 1962). Enzymes involved in the assimilation of ammonia are essential for the growth of most rumen organisms. In prokaryotes, glutamate dehydrogenase (GDH) and the dual enzyme system, glutamine synthetase (GS) and glutamate synthetase (GOGAT), are the two most important routes by which ammonia may be assimilated (Hespell, 1984).

There is a paucity of information concerning the pathways and regulation of ammonia assimilation in rumen organisms. Studies with ¹⁵NH₄⁺ have shown that bacterial amide is the most rapidly labelled nitrogen pool in the bacterial population of the rumen over a range of ruminal ammonia concentrations. These results imply that GS is an important route of ammonia assimilation in the bacterial population of the rumen (Smith et al., 1981).

Until recently, it appeared that prokaryotes had two forms of GS termed GSI and GSII. The majority of bacteria investigated have GSI-type enzymes but members of the *Rhizobiaceae*, *Frankiaceae* and *Streptomyces* contain both GSI- and GSII-type enzymes (Darrow & Knotts, 1977; Kumada et al., 1990; Rochefort & Benson, 1990). GSI is the typical prokaryotic GS (Streicher & Tyler, 1980) whereas GSII is similar to eukaryotic GS enzymes (Carlson & Chelm, 1986). The GS gene from the clinical isolate *Bacteroides fragilis* Bf-1 has been cloned (Southern et al., 1986), the nucleotide sequence determined (Hill et al., 1989), and the GS protein characterized (Southern et al., 1987). The GS from *B. fragilis* differs markedly from the GS enzymes of prokaryotes, eukaryotes and archaeobacteria. The holoenzyme is a single disc of six identical subunits with an *M*₉ of approximately 490000. The GSI and GSII holoenzymes are dodecamers and octamers, respectively. The *B. fragilis* GS subunit is large and has a calculated *M*₉ of 82827. It contains 729 amino acids, and is approximately 270 and 400 amino acids longer than the GSI and GSII subunits, respectively. The *B. fragilis* GS subunit is large and has a calculated *M*₉ of 82827. It contains 729 amino acids, and is approximately 270 and 400 amino acids longer than the GSI and GSII subunits, respectively. Because of these differences, Hill et al. (1989) suggested that GS enzymes with large subunits arranged as hexamers should be grouped together to form a third class of GS enzyme termed

*Author for correspondence. Tel. 21 650 2183; fax 21 650 4023; e-mail heide@micro.uct.ac.za.

Abbreviations: GS, glutamine synthetase; ORF, open reading frame.

The nucleotide sequence data reported in this paper have been submitted to GenBank and have been assigned the accession no. L08256.
GSII. We report the cloning and characterization of the glnA gene from the rumen anaerobe B. jibrisolvens. This glnA gene encodes a GS which is similar to the B. fragilis GS and is the second example of a GS which belongs to the GSII class.

**Methods**

**Bacterial strains, vectors, and growth conditions.** Chromosomal DNA from B. jibrisolvens H17c (Dehority, 1966) was used for the cloning of the B. jibrisolvens glnA gene. The E. coli host strain used for the original cloning experiment was the YMC11 glnA deletion mutant (Backman et al., 1981). Subcloning and other genetic manipulations were carried out in E. coli LK111 (Zabeau & Stanley, 1982). The E. coli Bacillus subtilis shuttle vector pEB1, used for the construction of the B. jibrisolvens genomic library, has been described previously (Lin et al., 1990). Subcloning for sequencing was done using the high copy number pUCBM20 and -21 vectors (Boehringer Mannheim). B. fragilis strain Bf-1 was used by Southern et al. (1986, 1987) and Hill et al. (1989) for the cloning and molecular analysis of the B. fragilis glnA gene. B. jibrisolvens H17c was grown in M10 medium as described by Strydom et al. (1986). E. coli strains were grown in Luria-Bertani (LB) medium or glucose minimal medium (Miller, 1972) with 1 g (NH4)2SO4 l−1 (GMMN) as the sole nitrogen source. Glucose minimal medium for nitrogen regulation studies contained 15 mm-glutamine (GMML) or 15 mm-glutamate and 13 mm-(NH4)2SO4 (GMMH). Ampicillin (100 μg ml−1) was added to media for plasmid selection and maintenance. B. fragilis Bf-1 was grown on supplemented brain heart infusion (BHI) agar (Holdeman & Moore, 1972) as described by Hill et al. (1989).

**Preparation of DNA and recombinant DNA techniques.** Construction of the B. jibrisolvens H17c genomic library in pEB1 has been described (Lin et al., 1990). Plasmid DNA prepared from pools of clones containing B. jibrisolvens DNA was used to transform the E. coli YMC11 glnA deletion strain. An E. coli YMC11 GlnA* ampicillin resistant transformant was isolated on GMMN medium. DNA fragments for subcloning were purified from 0.8% (w/v) agarose Tris/acetate gel slices by adsorption onto glass beads (Gene Clean). The preparation of plasmid DNA and restriction endonuclease mapping of the clones was done using standard techniques (Sambrook et al., 1989). B. jibrisolvens chromosomal DNA was prepared as described by Berger et al. (1989) and Southern hybridization was used to confirm that the insert DNA originated from B. jibrisolvens.

**Nucleotide sequence determination and analysis.** The ExoIII Mung bean nuclease technique was used to create a deletion series for sequencing (Henikoff, 1984). The nucleotide sequence of the glnA gene was determined by the dyeoxy-chain termination method using the Sequenase 2.0 kit (US Biochemical Corp.) according to the manufacturer's specifications. The nucleotide and deduced amino acid sequences were analysed using the Genetics Computer Group Inc. software package (version 7.0). All the current databases accompanying the GCG package were screened for related nucleotide and amino acid sequences.

Western immunoblotting. Western blotting from SDS-PAGE gels onto nitrocellulose membranes was done by the method of Towbin et al. (1979). B. fragilis GS antibody binding and the development of bands using a goat anti-rabbit serum conjugated to horse-radish peroxidase were done as described previously (Southern et al., 1987).

**Results and Discussion**

Cloning and expression of a GS gene from B. jibrisolvens in E. coli

The E. coli glnA deletion strain YMC11 was transformed with the genomic library of B. jibrisolvens. An E. coli YMC11 GlnA* transformant was isolated on GMMN agar plates containing ampicillin. The resident plasmid was designated pGS4 and was shown to harbour a 6 kb insert of B. jibrisolvens DNA. A 5-3 kb fragment of pGS4 insert DNA was subcloned in both orientations using the pUCBM20 and -21 sequencing vectors (pGS20 and pGS21, respectively). Both subclones retained GS activity. Hybridization of HpaI and BglII fragments internal to the pGS20 insert to B. jibrisolvens H17c chromosomal fragments of the same size confirmed the origin of the cloned DNA fragment (data not shown).

**Nucleotide sequence of the B. jibrisolvens glnA gene**

The nucleotide sequence of a 2423 bp region of B. jibrisolvens DNA containing the glnA gene was determined (Fig. 1). The DNA sequence contained an open reading frame (ORF) of 2103 nucleotides, which began with an ATG codon at position 201 and ended with a TAA codon at position 2303. The predicted size of the polypeptide encoded by this ORF was 701 amino acids, which has a calculated Mr of 79645. A potential ribosomal binding sequence (GGAGAGAG) was situated 6 bp upstream of the ATG initiation codon. No significant sequence homology was found upstream of the initiation codon with the canonical E. coli σ70 promoter sequence or with the promoter consensus sequence for genes under ntr control (GAGAGAG) (Dixon, 1984; Reitzer & Magasanik, 1985). The B. jibrisolvens glnA regulatory region was scanned for the presence of a consensus sequence similar to the consensus established for the binding of the ntrC gene product to the glnA promoter region in enteric bacteria (Dixon, 1984). Such a sequence could not be found. A 13 bp inverted repeat sequence capable of forming a stem-loop structure was located 33 bp downstream of the putative glnA stop codon.

Comparison of the translated amino acid sequence with sequences of GS enzymes from other organisms

The degree of amino acid sequence homology between the B. jibrisolvens GS and those of other prokaryotes was determined on the basis of identical amino acid sequence.

---

Fig. 1. (facing page). Nucleotide and deduced amino acid sequences of the B. jibrisolvens glnA gene and flanking regions. The initiation codon ATG is at nucleotide 200 and the termination codon TAA is at nucleotide 2301. The derived amino acid sequence of 701 amino acids is given in the one letter code. The putative ribosome binding site and the inverted repeat sequence are underlined.
comparisons (Table 1). The highest amino acid sequence identity was obtained with the GS subunit sequence from *B. fragilis* (40% identity). An analysis of Table 1 places the GS enzymes from Gram-negative and Gram-positive organisms in separate clusters based on homology, with the GS enzymes from *B. fibrisolvens* and *B. fragilis* showing a greater divergence in sequence similarity.

Alignment of the predicted GS polypeptide sequences from *B. fibrisolvens* and *B. fragilis* revealed considerable homology in a number of highly conserved blocks, particularly in the central regions of the proteins (Fig. 2). The *B. fibrisolvens* GS is 28 amino acid residues shorter than the *B. fragilis* GS enzyme, with the majority of the missing amino acid residues being located at the N-terminal end (residues 1–20) of the *B. fragilis* GS protein. Nevertheless, the two sequences share 40% amino acid identity when optimally aligned.

Five major regions of amino acid similarity between the prokaryotic type I and the eukaryotic type II GS enzymes have been identified (Rawlings *et al.*, 1987). The active site of the type I GS is formed between enzyme subunits, and it has been shown that the five conserved amino acid regions are associated with the proposed GS active site (Almassy *et al.*, 1986). Regions II to V are β-strands associated with two Mn$^{2+}$ cations of one subunit while region I contains the tryptophan residue which is thought to complete the active site formed between adjacent subunits.

Hill *et al.* (1989) reported that these five regions were not strongly conserved in the *B. fragilis* GS protein, but they were able to identify regions I, II, III and V which showed some similarity to the corresponding regions in the GSI and GSII proteins (Fig. 2).

In a comparison with the GSI and GSII enzymes, only two of the five major regions of homology (regions II and III) could be identified in the GS enzyme from *B. fibrisolvens* (Fig. 2). Hill *et al.* (1989) reported that region II of all GSI and GSII enzymes contained four highly conserved amino acids which are not contiguous [EV(X10)Q(X2)E]. These four amino acids are also conserved in region II of the *B. fibrisolvens* GS. Region III is considered to be the ATP-binding site because the sequence resembles the sequence XXXXXXXXXGXXGXXGKT found in several ATP-binding proteins (Walker *et al.*, 1982; Tischer *et al.*, 1986). In region III of the *B. fibrisolvens* GS enzyme, an ATP-binding sequence can be identified. Furthermore, region III contains seven amino acid residues (KXXXXXNGXXHXXXX) which are conserved in the ten eukaryote and prokaryote GS enzymes investigated (Hill *et al.*, 1989), as well as in the GS enzymes from *Azospirillum brasilense* (Bozouklian & Elmerich, 1986), *Azotobacter vinelandii* (Toukdarian *et al.*, 1990), *Bacillus cereus* (Nakano *et al.*, 1989), *E. coli* (Colombo & Villafranca, 1986), *Methylococcus capsulatus* (Cardy & Murrell, 1990) and *Rhizobium leguminosarum* (Colonna-Romano *et al.*, 1987). In all GS enzymes investigated, the spacing between regions II and III is also highly conserved and is approximately 42 amino acid residues.

An important difference between the prokaryotic GSI

**Table 1. Comparison of the identity of amino acid sequences of GS enzymes from various bacteria**

<table>
<thead>
<tr>
<th></th>
<th>BFIB</th>
<th>BFRA</th>
<th>ABRA</th>
<th>RLEG</th>
<th>ECOL</th>
<th>STYP</th>
<th>AVIN</th>
<th>MCAP</th>
<th>TFER</th>
<th>NGON</th>
<th>ANAB</th>
<th>BCER</th>
<th>BSUB</th>
<th>CACE</th>
<th>MVOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>BFIB</td>
<td>40</td>
<td>22</td>
<td>25</td>
<td>24</td>
<td>25</td>
<td>21</td>
<td>22</td>
<td>21</td>
<td>22</td>
<td>24</td>
<td>26</td>
<td>24</td>
<td>24</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>BFRA</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>21</td>
<td>24</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>24</td>
<td>25</td>
<td>23</td>
<td>24</td>
<td>20</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>ABRA</td>
<td>68</td>
<td>62</td>
<td>63</td>
<td>64</td>
<td>63</td>
<td>62</td>
<td>62</td>
<td>58</td>
<td>43</td>
<td>44</td>
<td>44</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>41</td>
</tr>
<tr>
<td>RLEG</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>63</td>
<td>63</td>
<td>58</td>
<td>57</td>
<td>56</td>
<td>41</td>
<td>41</td>
<td>41</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>ECOL</td>
<td>97</td>
<td>69</td>
<td>68</td>
<td>65</td>
<td>63</td>
<td>55</td>
<td>54</td>
<td>44</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>42</td>
<td>39</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>STYP</td>
<td>69</td>
<td>69</td>
<td>65</td>
<td>63</td>
<td>55</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>AVIN</td>
<td>71</td>
<td>67</td>
<td>65</td>
<td>53</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>MCAP</td>
<td>70</td>
<td>63</td>
<td>54</td>
<td>41</td>
<td>40</td>
<td>37</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>TFER</td>
<td>64</td>
<td>55</td>
<td>54</td>
<td>41</td>
<td>41</td>
<td>39</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>ANAB</td>
<td>35</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>BCER</td>
<td>86</td>
<td>62</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>BSUB</td>
<td>64</td>
<td>64</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
<td>52</td>
</tr>
</tbody>
</table>

The results are given as the percentage identity of the amino acid residues in the GS enzymes. Sequence data of the proteins were taken from Hill *et al.* (1989) for *B. fragilis* (BFRA), Bozouklian & Elmerich (1986) for *Azospirillum brasilense* (ABRA), Colonna-Romano *et al.* (1987) for *Rhizobium leguminosarum* (RLEG), Colombo & Villafranca (1986) for *E. coli* (ECOL), Janson *et al.* (1986) for *Salmonella typhimurium* (STYP), Toukdarian *et al.* (1989) for *Azotobacter vinelandii* (AVIN), Cardy & Murrell (1990) for *Methylococcus capsulatus* (MCAP), Rawlings *et al.* (1987) for *Thiobacillus ferrooxidans* (TFER), Tumer *et al.* (1983) for *Anaabaena* sp. (ANAB), Nakano *et al.* (1989) for *Bacillus cereus* (BCER), Strauch *et al.* (1988) for *B. subtilis* (BSUB), Janssen *et al.* (1988) for *Clostridium acetobutylicum* (CACE), Possot *et al.* (1989) for *Methanococcus voltae* (MVOL) and *B. fibrisolvens* (BFIB) (this paper). The sequence data for *Neisseria gonorrhoeae* (NGON) were obtained from the GenBank database (accession no. M84113).
and the prokaryotic GSIII enzymes is that the former contains a tryptophan residue, or the functionally similar phenylalanine residue, at the position corresponding to position 139 in region I of the B. fragilis GS. The prokaryotic GSIII enzyme from B. jibrisolvens, which is not regulated by adenylylation, is replaced by either valine, in the region IV of the B. fragilis GS, or isoleucine in the region IV of the B. fragilis GS. The B. fragilis GS protein, which is not regulated by adenylylation, is not identified in the B. fragilis GS (Hill et al., 1989). The underlined regions are the regions identified in the B. fragilis GS protein, showing homology with type I and type II GS subunits.

Region IV is thought to be the glutamate-binding site, but either the B. fragilis or the B. fragilis GS enzymes show convincing similarity with the other GS enzymes in region IV and region II cannot be identified in the B. fragilis GS enzymes. Region V was identified in the B. fragilis GS by relatively high similarity with region V of the Streptomyces coelicolor GS (Hill et al., 1989). However, the B. fragilis GS shows no convincing similarity with the B. fragilis or other GS enzymes in region V (Fig. 2).

The GSIII enzyme from B. fragilis is not regulated by adenylylation (Southern et al., 1987) and does not contain an 18-amino-acid sequence containing a tyrosine residue which has been shown to be the site of adenylylation in GSI enzymes (Shapiro & Stadtman, 1968; Almassy et al., 1986; Hill et al., 1989). An 18-amino-acid sequence containing tyrosine which even remotely resembles the GSI adenylylation region could not be identified in the B. fragilis GS. The B. fragilis GS appears to be the second example of a GS from a Gram-negative bacterium which is not regulated by adenylylation.

Detection of GS in crude cell extracts by Western blotting

Since the B. fragilis GS showed homology to the B. fragilis GS, an antiserum raised against the purified B. fragilis GS was tested to determine whether it reacted with the B. fragilis GS. The B. fragilis GS antiserum detected a single polypeptide with an apparent Mr of 75,000 in crude cell extracts of B. fragilis cells, B. fragilis cells and E. coli YMC11(pGS4) cells (Fig. 3a,b). Previously, we showed that the B. fragilis

Fig. 2. Alignment of amino acid sequences of GS proteins from B. fibrisolvens (BFIB; this paper) and B. fragilis (BFRA; Hill et al., 1989). The underlined regions are the regions identified in the B. fragilis polypeptide sequence by Hill et al. (1989) showing homology with type I and type II GS subunits.
antiserum did not react with any SDS-PAGE-fractionated proteins from E. coli YMC11 cells (Southern et al., 1987). It appears that the production of the cloned B. fibrisolvens GS is regulated in E. coli YMC11 cells since the production of the B. fibrisolvens GS was induced by growth of E. coli YMC11(pGS4) cells in nitrogen-limiting GMML medium (Fig. 3b). Although E. coli YMC11 is a heterologous host, the fact that the production of the B. fibrisolvens GS appears to be regulated in E. coli YMC11 will be helpful in studying the regulation of expression of the B. fibrisolvens glnA gene. Studies on the regulation of GS activity in B. fibrisolvens are not feasible at present, since a suitable nitrogen-limiting medium which supports the growth of B. fibrisolvens is not available.

Conclusion

Although the establishment of the novel type-III class of GS is justified on the characterization of a single GS from B. fragilis, the demonstration that this GS type is present in other bacteria is important in confirming and extending the relevance and importance of GSIII in prokaryotes. The GS from B. fibrisolvens can be assigned to the GSIII class because of its very large subunit size, and amino acid sequence and antigen homology with the B. fragilis GS. Size and amino acid sequence homology differentiate the B. fragilis and B. fibrisolvens GSIII enzymes from enzymes belonging to the GSI and GSII classes. The GSIII enzyme has been demonstrated in two diverse members of the family Bacteroidaceae which inhabit similar environments. It will be interesting to determine whether other members of the Bacteroidaceae have the GSIII enzyme and whether this different GS is a distinguishing characteristic of this bacterial family.

References


Butyrivibrio fibrisolvens glutamine synthetase


