Molecular Analysis of a Novel Glutamine Synthetase of the Anaerobe
Bacteroides fragilis

By RUSSELL T. HILL, JOAN R. PARKER, HEIDE J. K. GOODMAN,
DAVID T. JONES† AND DAVID R. WOODS*

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

(Received 19 June 1989; revised 31 August 1989; accepted 18 September 1989)

The nucleotide sequence of a 2777 bp DNA segment containing the Bacteroides fragilis glnA gene was determined. The B. fragilis glnA open reading frame of 2187 bp encoded a glutamine synthetase (GS) subunit of 729 amino acid residues with a calculated M, of 82827. The apparent M, of the GS subunit determined by SDS-PAGE was approximately 75000. A single mRNA transcription start point was identified upstream of the B. fragilis glnA open reading frame. The B. fragilis GS subunit is approximately 270 and 400 amino acids longer than the GSI and GSII subunits, respectively, of other prokaryotes and eukaryotes. The GSI and GSII holoenzymes are dodecamers and octamers respectively, whereas the GS of B. fragilis is a hexamer. Although GSI and GSII subunits show amino acid similarity in five conserved regions, this similarity is not strongly conserved in the B. fragilis GS. The GS of B. fragilis is not regulated by adenylylation and lacks the adenylylation site. It also lacks the Trp residue associated with the active site in GSI and GSII enzymes from other prokaryotes and eukaryotes.

INTRODUCTION

Glutamine synthetase (GS) (EC 6.3.1.2) is an important enzyme in nitrogen metabolism and ammonia assimilation in many prokaryotes, eukaryotes and archaeobacteria. GS enzymes of eukaryotic origin are octamers (Prusiner & Stadtman, 1973) whereas the GS enzymes of eubacteria and an archaeobacterium are dodecamers of M, approximately 600000 composed of a single type of subunit, the M, of which falls in the range 44000 to 59000 (Streicher & Tyler, 1980; Bhatnagar et al., 1986).

Prokaryotes have two forms of GS, termed GSI and GSII. The majority of bacteria investigated have GS enzymes but members of the Rhizobiaceae contain both GSI and GSII enzymes. GSI is the typical prokaryotic GS whereas GSII is similar to eukaryotic GS enzymes (Carlson & Chelm, 1986). Prokaryotic GSI subunits investigated vary in length from 444 to 474 amino acids (Janssen et al., 1988), whereas the GSII subunits of eukaryotes investigated vary in length from 355 to 373 amino acids (Gebhardt et al., 1986; Hayward et al., 1986; Tischer et al., 1986). The eukaryotic GSII subunits lack the C-terminal portion of the prokaryotic GSI subunit including the adenylylation site. The GSII subunit from Bradyrhizobium japonicum is 329 amino acids long and also lacks the terminal portion of the GSI subunit and the adenylylation site (Carlson & Chelm, 1986). In a comparison of prokaryotic and eukaryotic GSI and GSII enzymes, Rawlings et al. (1987) showed that although the amino acid similarity between the enzymes was only approximately 15%, the major part of this similarity was located in five

† Present address: Department of Microbiology, University of Otago, Dunedin, New Zealand.

Abbreviations: glnA, gene encoding glutamine synthetase; GS, glutamine synthetase; ORF, open reading frame; SD, Shine–Dalgarno.

The nucleotide sequence data reported in this paper have been submitted to GenBank and have been assigned the accession number M28252.
regions corresponding to \( \beta \)-sheets in the *Salmonella typhimurium* enzyme (Almassy *et al.*, 1986). These regions are strongly conserved in all GS enzymes analysed to date (Janssen *et al.*, 1988). A feature of the five conserved amino acid regions is that they are all associated with the proposed GS active site (Almassy *et al.*, 1986; Janson *et al.*, 1986). Regions II to V are \( \beta \)-strands closely associated with two Mn\(^{2+} \) cations of one subunit, while region I contains the Trp residue which is thought to complete the active site formed between adjacent subunits.

An indication that the GS from the Gram-negative obligately anaerobic bacterium *Bacteroides fragilis* may differ from other GS enzymes of eukaryotes and prokaryotes was suggested by Southern *et al.* (1986, 1987). Studies of the cloned *B. fragilis* GS subunit produced in *Escherichia coli* and *B. fragilis* indicated that the *B. fragilis* GS subunit was substantially larger (\( M \), approximately 75000) than other GSI and GSII subunits (maximum \( M \), approximately 59000). However, the \( M \), of the *B. fragilis* GS holoenzyme (approximately 490000) was lower than that of the prokaryotic GSII holoenzyme (approximately 600000). It was concluded that the GS of this anaerobic prokaryote is a hexamer and not a dodecamer. To confirm the structure of the *B. fragilis* GS we sequenced the *glnA* gene and compared the nucleotide sequence and the derived amino acid sequence with other GSI and GSII enzymes. We report that the GS from *B. fragilis* differs markedly from the GSI and GSII enzymes from other prokaryotes and eukaryotes.

**METHODS**

*Bacterial strains, plasmids and growth conditions.* *B. fragilis* strain Bf-1 was described by Mossie *et al.* (1979) and was used by Southern *et al.* (1986, 1987) for the cloning of the *B. fragilis* *glnA* gene. Plasmid pJS139 [GlnA\(^* \), ampicillin resistant (Ap\(^* \)\)], is an *E. coli* recombinant plasmid which contains the *B. fragilis* Bf-1 *glnA* gene and enables *E. coli* deletion mutants to utilize \((\text{NH}_4\)\)\(_2\)\(\text{SO}_4 \) as a sole source of nitrogen (Southern *et al.*, 1986, 1987). *E. coli* strains were grown on Luria medium (Miller, 1972) or glucose minimal medium (Miller, 1972) with 1 g (\(\text{NH}_4\)\(_2\)\(\text{SO}_4 \) per litre as the sole nitrogen source. All strains were grown at 37°C with aeration.

*Preparation of DNA.* Plasmid DNA was prepared by the alkali-hydrolysis method of Ish-Horowicz & Burke (1981).

*DNA sequence analysis.* DNA sequencing of both strands of plasmids containing the cloned *B. fragilis* *glnA* gene was done by the dideoxynucleotide chain termination method of Sanger *et al.* (1977) with templates prepared from DNA subcloned in pUC18 and pUC19 and exonuclease III digestion (Henikoff, 1984). *E. coli* LK111 (Zabeau & Stanley, 1982) was used as the recipient strain. The DNA chains were radiolabelled with \([\text{\({}^{35}\text{S}\)}\text{]dCTP \ (400 Ci \ mmol}^{-1}; 14.8 \text{TBq \ mmol}^{-1} \)). The DNA and deduced amino acid sequences were analysed with an IBM XT computer using a GENEPRO (version 3.1) program (Riverside Scientific).

*Primer extension reactions.* A synthetic 17-mer oligonucleotide 5'ATAGCTCTTTGATGCG-3' was used as a primer for the synthesis of cDNA. A [\(\text{\(\alpha\)}-\text{\(^{32}\text{P}\)}\]-5' end-labelled primer (5 ng) was mixed with 10 \(\mu\)g of RNA prepared from *B. fragilis* Bf-1 or *E. coli* LK111 (pJS139) by the method of Aiba *et al.* (1981). The mixture was denatured at 95°C in 10 mm-Tris/HCl pH 8.3, 0.35 mm-EDTA for 5 min and chilled on dry ice. Primer-RNA hybridizations were incubated at 42°C for 6 h. The reverse transcription reaction was carried out at 42°C for 60 min in 20 \(\mu\)l of 50 mm-Tris/HCl, pH 8.3, 2 mm-MgCl\(_2\), 2 mm-DTT, 1 mm each of dGTP, dATP, dCTP, dTTP and 50 units of reverse transcriptase. The RNA was degraded with RNase A (20 \(\mu\)g) for 10 min at 37°C. After phenol extraction and ethanol precipitation the cDNA was resuspended in 10 \(\mu\)l of 99% (w/v) formamide, 0-1 mg bromophenol blue ml\(^{-1}\), 0-1 mg xylene cyanol ml\(^{-1}\) and electrophoresed in a sequencing gel.

**RESULTS AND DISCUSSION**

*Nucleotide sequence and transcription start point of glnA gene* The nucleotide sequence of a 2777 bp region of *B. fragilis* DNA containing the *glnA* gene was determined. The DNA sequence contained an open reading frame (ORF), which from a start codon (ATG) to a stop codon contained 2187 nucleotides (Fig. 1). Primer extension experiments carried out with RNA isolated from *B. fragilis* cells grown in brain heart infusion broth.
(Holdeman & Moore, 1972) showed a single major transcription start point 18 bp from the first nucleotide of the ATG start codon (position +1) of the 2187 bp ORF (Fig. 2). Similar primer extension experiments carried out with RNA isolated from *E. coli* (pJS139) cells showed a number of weak transcription start points within 16 to 40 bp of the ATG start codon (results not shown). Although promoter consensus sequences for *B. fragilis* have not been defined, it is interesting that an *E. coli* promoter sequence containing −10 and −35 RNA-polymerase-binding consensus sequences separated by 22 bp is situated 32 bp upstream of the *B. fragilis* transcription start point (Fig. 1). Previously we reported that the apparent *M*. of the *B. fragilis* GS subunit determined by SDS-PAGE was approximately 75000 (Southern et al., 1986, 1987). The start region of the *B. fragilis* glnA ORF contains five in-phase ATG start codons which would result in proteins with calculated *M*. values of 82827, 82481, 78067, 75131 and 74345 respectively (Fig. 1). The next downstream in-phase ATG codon would produce a protein with a calculated *M*. of 55823, which is too small to be regarded as the *B. fragilis* GS. The first ATG codon is preceded by an AAGAGA sequence eight nucleotides from the presumptive start codon. This AG-rich region shows some similarity to other Shine–Dalgarno (SD) sequences (Shine & Dalgarno, 1976) and is situated close enough to the start of the ORF to be regarded as a putative SD region. The other four ATG codons were not preceded by sequences which showed similarity to SD sequences. The *B. fragilis* glnA gene was expressed in *E. coli* from a promoter site contained within the cloned *B. fragilis* DNA (Southern et al., 1986). Attempts to sequence the amino-terminal end of the GS subunit were unsuccessful as the protein appeared to be blocked. A feature of the upstream region is the presence of two near-perfect direct repeat sequences of 46 and 45 bp from positions −127 to −82 and −74 to −30 respectively (Fig. 1). This region of direct repeats overlaps the putative promoter sequence and may be involved in the regulation of GS expression. The regulation of the glnA gene in *Salmonella typhimurium* and *E. coli* involves five NR binding consensus sequences of 15 bp (Ames & Nikaido, 1985). Some of the NR binding sites overlap the glnA promoters. However, there is no homology between the repeats upstream of the *B. fragilis* glnA gene and the NR binding consensus sequences.

**Comparison of B. fragilis GS with other GS enzymes**

Recently Rawlings et al. (1987), Janssen et al. (1988), Strauch et al. (1988) and Wray & Fisher (1988) have compared the nucleotide-sequence-derived GS amino acid sequences from diverse prokaryotes including *Bradyrhizobium japonicum* (Carlson & Chelm, 1986), *E. coli* (Colombo & Villafranca, 1986), *Anabaena* sp. strain 7120 (Tumer et al., 1983), *Thiobacillus ferrooxidans* (Rawlings et al., 1987), *Salmonella typhimurium* (Janson et al., 1986), *Clostridium acetobutylicum*...
Glutamine synthetase of Bacteroides fragilis

Fig. 3. Dot-matrix comparisons of amino acid sequences of GSI enzymes from Salmonella typhimurium and Streptomyces coelicolor, of GSII enzymes from alfalfa and Bradyrhizobium japonicum, and of the Bacteroides fragilis GS with a GSI enzyme and a GSII enzyme. Dot-matrix comparisons were done using the GENEPRO (version 3.1) sequence analysis program (Riverside Scientific), checking for five matches in a window of ten residues.

(Janssen et al., 1988), Bacillus subtilis (Strauch et al., 1988), Azospirillum brasilense (Bozouklian & Elmerich, 1986) and Streptomyces coelicolor (Wray & Fisher, 1988).

The GSII subunit from Bradyrhizobium japonicum contained 329 amino acids (M, 36865), whereas the GSI subunits of the other bacteria contained 444 to 474 amino acids (M, 49000 to 59000). In contrast, the Bacteroides fragilis glnA ORF of 2187 nucleotides encodes a GS subunit of 729 amino acid residues with a calculated M, of 82827. Thus the B. fragilis GS subunit is approximately 270 and 400 amino acids longer than the GSI and GSII subunits respectively. The calculated M, of the B. fragilis GS confirms our previous results (Southern et al., 1987) involving PAGE analysis of the GS which indicated that the B. fragilis GS subunit was substantially larger than other prokaryote GS subunits.

The GSI and GSII enzymes can be separated into two groups on the basis of amino acid similarity. Although there is only limited amino acid similarity between these two groups, a high degree of similarity exists between GS enzymes within each group. This is illustrated by the dot-matrix comparisons of GS subunits from distantly related organisms within each group (Fig. 3). The GS from B. fragilis did not show any overall amino acid similarity with either GSI or GSII enzymes (Fig. 3).

The major regions of amino acid sequence similarity between the prokaryote GS enzymes involve five β-sheets which are all associated with the proposed GS active site (Almassy et al., 1986; Janson et al., 1986; Rawlings et al., 1987). These regions are also conserved in GS from eukaryotes (Janssen et al., 1988). A comparison of these five regions from representative groups of eukaryotes, prokaryotes and B. fragilis is shown in Fig. 4. In the GS enzyme from B. fragilis...
Fig. 4. Comparison of amino acid sequences of five regions of homology of GS enzymes from *Homo sapiens* (Hs) (Gibbs *et al.*, 1987), alfalfa (Af) (Tischer *et al.*, 1986), *Bradyrhizobium japonicum* (Bj) (Carlson & Chelm, 1986), *Anabaena* sp. strain 7120 (An) (Tumer *et al.*, 1983), *Salmonella typhimurium* (St) (Janson *et al.*, 1986), *Thiobacillus ferrooxidans* (Tf) (Rawlings *et al.*, 1987), *Clostridium acetobutylicum* (Ca) (Janssen *et al.*, 1988), *Streptomyces coelicolor* (Sc) (Wray & Fisher, 1988), *Bacillus subtilis* (Bs) (Strauch *et al.*, 1988) and *Bacteroides fragilis* (Bf). The amino acids are identified by the single-letter code and the positions of the first amino acids in each region are indicated. Numbering of residues begins with the start methionine at the N-terminal end. Identical residues are boxed. No region IV was found in *B. fragilis*.

these five regions show far less amino acid similarity although certain conserved amino acids can be identified in four of these five regions.

Region I of the *B. fragilis* GS shows recognizable but relatively poorly conserved amino acid similarity with the other GS enzymes (Fig. 4). The amino acid sequence DGSS at the start of region I is conserved in all the eukaryote and prokaryote GS enzymes. A Ser residue (the homologue of Ser-64 in *Salmonella typhimurium*) is conserved in all GS enzymes except in the *B. fragilis* GS, where it is replaced by a Cys residue (Cys-164 of *B. fragilis*).

In region I the *S. typhimurium* GS has a Trp residue at position 58, which is thought to complete the active site formed between adjacent subunits (Almassy *et al.*, 1986). In the GS from three Gram-positive bacteria, *Bacillus subtilis*, *Clostridium acetobutylicum* and *Streptomyces coelicolor*, the Trp residue is replaced by the functionally similar Phe residue. The *Bacteroides*
Glutamine synthetase of *Bacteroides fragilis*

*B. fragilis* GS differs from all the other prokaryotic GSs in that the active site Trp in region I is replaced by Val, which shows no functional similarity with Trp (Fig. 4).

In the *Salmonella typhimurium* GS holoenzyme the Trp-58 residue is situated within 0.6 nm of the Tyr residue at position 398 in the carboxy-terminal portion of the adjacent subunit. Tyr-398 has been shown to be the site of adenylylation (Almassy *et al.*, 1986; Shapiro & Stadtman, 1968). Since the Trp-58 of one subunit is situated close to the Tyr-398 of another subunit in the *S. typhimurium* GS, it is suggested that the three-dimensional structure of this region may control the ability of an enzyme to be regulated by adenylylation. The GS enzymes which are subject to adenylylation showed strong similarity among the 18 amino acids adjacent to this Tyr residue (Janssen *et al.*, 1988). *B. fragilis* GS is the first example of a GS from a Gram-negative bacterium which is not regulated by adenylylation (Southern *et al.*, 1987). It is perhaps not surprising, therefore, that the *B. fragilis* GS does not contain either a Trp residue at the relevant position in region I, or an 18-amino-acid sequence containing a Tyr residue which even remotely resembles the *S. typhimurium* adenylylation region.

In both GS I and GSII enzymes, regions II to V are β-strands closely associated with two Mn²⁺ cations of one subunit (Janson *et al.*, 1986). Region II contains four highly conserved amino acids which are not contiguous [EV(X₃₋₄)Q(X₁)E] (Fig. 4). Region III is considered to be the ATP-binding site because the sequence resembles the sequence KXXXXGXXGXGKT found in several ATP-binding proteins (Walker *et al.*, 1982; Tischer *et al.*, 1986). In the *B. fragilis* GS, region III shows the highest degree of conservation; it contains eight amino acid residues which are conserved in all the 10 eukaryote and prokaryote GS enzymes investigated, and an ATP-binding sequence can be identified (Fig. 4). Region IV is thought to be the glutamate-binding site because it is similar to the sequence DRGASIV (amino acid positions 18 to 24) of bovine and chicken glutamine dehydrogenase substrate-binding sites (Moon & Smith, 1972; Rasched *et al.*, 1974; Tischer *et al.*, 1986; Rawlings *et al.*, 1987). The *B. fragilis* GS shows no convincing similarity with the other GS enzymes in region IV. Region V, between positions 471 and 484 of the *B. fragilis* GS, shows relatively high similarity with region V of the *Streptomyces coelicolor* GS between positions 357 and 373. Seven amino acids are conserved and three positions contain amino acid residues which occur in one or more other GS enzymes at the equivalent position (Glu-472, Arg-474 and Asn-481 of *B. fragilis*). These Glu and Arg residues are conserved in all the GS enzymes (Fig. 4).

In Fig. 5 the relative spacings of the GS regions I to V are compared. It is apparent that the relative positions of the best-conserved regions in the *B. fragilis* GS (regions I, II, III and V) are...
similar to the positioning of these regions in the other GS enzymes. The start of region I of the GSI subunits is located approximately 50 amino acids from the amino-terminal end, and the end of region V is approximately 100 amino acids from the carboxy-terminal end. In comparison, region I of the B. fragilis GS is preceded by 149 amino acids and there is a long carboxy-terminal region of 391 amino acids extending beyond region V. From Fig. 5 it is apparent that in all the GS enzymes the spacing between regions II and III is highly conserved and is approximately 42 amino acid residues.

Conclusions

The GS from the Gram-negative obligate anaerobe B. fragilis differs markedly from the GS I and GS II enzymes from other prokaryotes and eukaryotes. Our results demonstrate the very large size of the B. fragilis GS subunit, and confirm our previous suggestion that the GS is a hexamer (Southern et al., 1987). Other prokaryote GS enzymes normally consist of 12 subunits (Streicher & Tyler, 1980; Bhatnagar et al., 1986). The GS from Clostridium pasteurianum has been reported to consist of 20 subunits (Krishnan et al., 1986).

Classification of Bacteroides species by ribosomal RNA sequence analysis indicated that the genus Bacteroides belongs to a distinct assemblage of genera including Bacteroides, Flavobacterium and Cytophaga (Woese, 1987). The taxonomic separation of Bacteroides is supported by the structure of the B. fragilis GS. It will be interesting to determine whether the GS enzymes of other members of this group are similar to the GS of B. fragilis. If this is the case then GS enzymes with large subunits arranged as hexamers should be grouped together to form a third class of GS enzymes (GSIII). Genetic studies in Agrobacterium and Rhizobium have suggested the possible occurrence of a third GS by the cloning of a gene glnT which complements an E. coli glnA deletion, allowing growth in the absence of glutamine (Rossbach et al., 1988). However, no enzyme activity has yet been ascribed to the glnT gene product.

We are grateful to D. E. Rawlings for supplying GS protein sequences and for useful discussions.

REFERENCES


Glutamine synthetase of Bacteroides fragilis 3279

Clostridium acetobutylicum. Journal of Bacteriology 170, 400–408.


