Molecular and enzymological evidence for two classes of fumarase in 
*Bacillus stearothermophilus* (var. *non-diastaticus*)

**SUSAN K. REANEY,** **STEPHEN J. BUNGARD** and **JOHN R. GUEST**

**1 The Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, PO Box 594, Firth Court, Western Bank, Sheffield S10 2UH, UK**

**2 ICI Bio Products & Fine Chemicals, PO Box 1, Billingham, Cleveland TS23 1LB, UK**

(Received 13 August 1992; revised 14 October 1992; accepted 19 October 1992)

The gene (*fumA*<sub>ba</sub>) encoding an oxygen-labile fumarase of *Bacillus stearothermophilus* has been cloned and sequenced. The structural gene (1542 bp) encodes a product (*FumA<sub>ba</sub>*) of *M*<sub>r</sub> 56788 containing 514 amino acid residues. The amino acid sequence is 23% identical (37% similar) to *FumA* and *FumB*, the labile [4Fe–4S]-containing fumarases (Class I enzymes) of *Escherichia coli*. It exhibits no significant similarity to *FumC* and CitG, the stable fumarases (Class II enzymes) of *E. coli* and *Bacillus subtilis* (respectively). Enzymological studies indicated that *FumA<sub>ba</sub>* resembles the iron–sulphur-containing fumarases in being dimeric (*M*<sub>r</sub> 2 × 58500), oxygen labile and partially reactivated by Fe<sup>2+</sup> plus DTT. The *fumA*<sub>ba</sub> gene is the first gene encoding a Class I fumarase to be characterized in any organism other than *E. coli*. Enzymological and DNA-hybridization studies further indicated that *B. stearothermophilus* resembles *E. coli* in containing an oxygen-stable fumarase (Class II enzyme). Sequence comparisons revealed significant similarities between the Class I fumarases and the products of adjacent open-reading frames (*orfZ1* and *orfZ2*) located upstream of the macromolecular synthesis operon (*rpsU–dnaG–rpoD*) at 67 min in the *E. coli* linkage map. Located downstream of *fumA*<sub>ba</sub> there is an unidentified gene (*orfZ2*), which is homologous to the rhizobial *nodB* genes involved in the initiation of root nodule formation.

**Introduction**

Fumarase (fumarate hydratase: EC 4.2.1.2) catalyses the interconversion of fumarate and L-malate. It functions in the citric acid cycle during aerobic metabolism and in the reductive pathway of the branched form of the citric acid cycle which provides fumarate as an anaerobic electron acceptor during anaerobic metabolism (Guest et al., 1990).

Two distinct classes of fumarase have been discovered, primarily by cloning and sequencing the fumarase genes of *Escherichia coli* K12 (Guest & Roberts, 1983; Miles & Guest, 1984; Guest et al., 1985; Woods et al., 1986, 1988a; Bell et al., 1989). This organism contains at least three fumarase genes: *fumA* and *fumB* encoding two similar Class I enzymes, designated *FumA* and *FumB*; and *fumC*, encoding the unrelated Class II enzyme, designated *FumC*.

FumA and FumB (Class I) are oxygen-labile, iron–sulphur-containing homodimeric enzymes (subunit *M*<sub>r</sub> 60000) exhibiting 90% sequence identity (Bell et al., 1989). Aerobically inactivated FumA can be reactivated by anaerobic incubation with Fe<sup>2+</sup> and thiols (Yumoto & Tokushige, 1988; Flint et al., 1989; Ueda et al., 1991). Recent EPR spectroscopy has shown that FumA resemblesaconitase in containing a [3Fe–4S] cluster which is converted to a [4Fe–4S] cluster upon activation (Ueda et al., 1991; Flint et al., 1989, 1992). Other Class I fumarases include the immunologically cross-reacting and highly unstable enzyme of *Euglena gracilis* (Shibata et al., 1985; Woods et al., 1988a) and the oxygen-labile fumarase detected in *Bradyrhizobium japonicum* (Acuña et al., 1991). In addition, these enzymes may belong to a wider group of hydratases and other enzymes which contain Fe–S centres but do not catalyse redox-reactions e.g. aconitase, L(+)-tartrate dehydratase, maleate dehydratase, t-serine dehydratase, lactyl-CoA dehydratase, 2-hydroxyglutaryl-CoA dehydratase, dihydroxyacid dehydratase, DNA endonuclease III and glutamine PRPP amidotransferase (Switzer, 1989; Cammack, 1991; Buckel, 1992).

In contrast, FumC (Class II) is a stable homotetrameric
enzyme (subunit \( M_r \) 50000) which is related to the products of the \( \text{cit}G \) gene of \( \text{Bacillus subtilis} \) (Moir et al., 1984; Miles & Guest, 1985), the \( \text{FUM}1 \) gene of \( \text{Saccharomyces cerevisiae} \) (Wu & Tzagoloff, 1987) and the \( \text{fumC} \) gene of \( \text{Bradyrhizobium japonicum} \) (Acuña et al., 1991), and to the mammalian mitochondrial fumarases (Kinsella & Doonan, 1986; Suzuki et al., 1989; Sachettini et al., 1988). The Class II fumarases from diverse sources exhibit a remarkable degree of amino acid sequence conservation (50–65% identity). It is also clear that the Class II enzymes (subunit \( M_r \) 50000) which is related to the products of the \( \text{cit}G \) gene of \( \text{Bacillus subtilis} \) (Moir et al., 1984; Miles & Guest, 1985), the \( \text{FUM}1 \) gene of \( \text{Saccharomyces cerevisiae} \) (Wu & Tzagoloff, 1987) and the \( \text{fumC} \) gene of \( \text{Bradyrhizobium japonicum} \) (Acuña et al., 1991), and to the mammalian mitochondrial fumarases (Kinsella & Doonan, 1986; Suzuki et al., 1989; Sachettini et al., 1988). The Class II fumarases from diverse sources exhibit a remarkable degree of amino acid sequence conservation (50–65% identity). It is also clear that the Class II fumarases belong to a family of structurally related homotetrameric enzymes including aspartase, argininosuccinase (Woods et al., 1988b) and adenylosuccinase (He et al., 1992), which are 38%, 15% and 19% identical to \( \text{FumC} \), respectively.

In previous studies with \( \text{cit}G \) mutants of \( \text{B. subtilis} \) there was no evidence for multiple fumarases (Rutberg & Hoch, 1970), suggesting that the obligately aerobic bacilli (Gram-positive) may differ from the facultatively anaerobic enterobacteria (Gram-negative) in containing a single Class II fumarase rather than representatives of both classes. A fumarase gene of \( \text{B. stea} \)thermophilus was therefore sought for use in amplifying a thermostable and oxygen-insensitive enzyme for structural and mutagenic studies. This paper reports the characterization of a gene (\( \text{fumA}_{\text{B}} \)) encoding a Class I (oxygen-labile) fumarase, as well as preliminary evidence for the presence of a Class II fumarase in \( \text{B. stea} \)thermophilus. An adjacent gene related to the \( \text{nodB} \) genes of nodulating bacteria was also characterized, and a novel relationship between the Class I fumarases and a pair of unidentified \( \text{E. coli} \) genes was detected.

### Methods

**Bacterial strains, plasmids and phagemids.** The bacterial strains, plasmids and phagemids are listed in Table 1. \( \text{E. coli} \) JH400 lacks all three \( \text{fum} \) genes and was used in the cloning and complementation studies; other strains used for the isolation and propagation of phages and plasmids were: C600 (\( \text{iZAPII} \) derivatives), JM101 (M13 derivatives), and XL1-Blue (plasmids). The gene bank of \( \text{B. stea} \)thermophilus var. \( \text{non-diastaticus} \) (DSM 2334) contained products of complete and partial EcoRI digestion (2 to 9 kb) ligated in the Stratagene vector \( \text{iZAPII} \) (Mallinder et al., 1992). Phagemids were excised from \( \text{iZAPII} \) derivatives by the in vivo procedure provided by the supplier.

**Media and growth conditions.** The rich medium used for routine subculture was \( \text{L-broth} \). A weak peptone medium (BBL) was used to detect \( \text{fum} \) phages as described previously (Guest & Roberts, 1983). Transformants of \( \text{E. coli} \) JH400 were tested for complementation of the \( \text{Fum}^+ \) phenotype on plates of citrate-free minimal medium (Cole & Guest, 1980) containing glucose (11 mM), sodium L-malate (40 mM) or sodium fumarate (40 mM) as sole carbon source and supplemented with methionine (40 \( \mu \)g ml\(^{-1}\)) and ampicillin (100 \( \mu \)g ml\(^{-1}\)) as required. All strains were grown aerobically at either 30 °C (\( \lambda \)-infected cultures), 55 °C (thermophiles) or 37 °C (other strains).

### Table 1. **Bacterial strains, phagemids and plasmids**

<table>
<thead>
<tr>
<th>Bacterium and plasmids</th>
<th>Relevant genotype or phenotype</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bacterium</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{E. coli} )</td>
<td>Wild-type</td>
<td></td>
</tr>
<tr>
<td>W3110</td>
<td>( \Delta (\text{lac-proAB}) ) ( \text{F}^+ ) (( \text{traD36 proA}^+ \text{B}^+ \text{lacI}^+ \Delta \text{M15} ))</td>
<td>Messing (1983)</td>
</tr>
<tr>
<td>JM101</td>
<td>( \Delta (\text{dgs-fumCA-mamA}) \Delta (\text{mel-fumB}) \text{zdg-232-Tn10} \text{metB1 spoT1 relA} )</td>
<td>Henson et al. (1987)</td>
</tr>
<tr>
<td>JH400</td>
<td>( \text{thr leu6 thi1 supE44 tonA21 lacY1} )</td>
<td></td>
</tr>
<tr>
<td>C600</td>
<td>( \text{recA1 lac endA1 gvrA96 thi hsrd17 supE44 relA1} )</td>
<td>Guest &amp; Roberts (1983)</td>
</tr>
<tr>
<td>XL1-Blue</td>
<td>( \text{F}^+ ) (( \text{proA}^+ \text{B}^+ \text{lacI}^+ \Delta \text{M15 Tn10} ))</td>
<td>Bullock et al. (1987)</td>
</tr>
<tr>
<td>( \text{B. stea} )thermophilus var. ( \text{non-diastaticus} )</td>
<td>Undefine auxotrophy</td>
<td>Liao et al. (1986)</td>
</tr>
<tr>
<td>DSM 2334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{B. stea} )thermophilus ( \text{var.} ) ( \text{non-diastaticus} )</td>
<td>Undefine auxotrophy</td>
<td></td>
</tr>
<tr>
<td>NRRL 1174</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{B. subtilis} ) 1604</td>
<td>( \text{trpC2} )</td>
<td>Moir et al. (1979)</td>
</tr>
<tr>
<td>( \text{B. caldotenax} ) DSM 406</td>
<td></td>
<td>Sharp et al. (1986)</td>
</tr>
<tr>
<td><strong>Phagemid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{pBluescript} )</td>
<td>( \text{Ap}^R )</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pGS507</td>
<td>( \text{pBluescript derivative, Ap}^R )</td>
<td></td>
</tr>
<tr>
<td>pGS509</td>
<td></td>
<td>This study</td>
</tr>
<tr>
<td>pGS511</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Plasmid</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pGS96</td>
<td>( \text{citG}_{\text{B}} ) in ( \text{pHV33}, \text{Ap}^R )</td>
<td>Moir et al. (1984)</td>
</tr>
<tr>
<td>pGS54</td>
<td>( \text{fumAC}_{\text{Es}} ) in ( \text{pBR322}, \text{Ap}^R )</td>
<td>Guest &amp; Roberts (1983)</td>
</tr>
<tr>
<td>pGS210</td>
<td>( \text{fumC}_{\text{Es}} ) in ( \text{pUC12}, \text{Ap}^R )</td>
<td>Woods &amp; Guest (1987)</td>
</tr>
</tbody>
</table>
Enzymology. Cultures were grown in L-broth to early stationary phase, harvested, and disrupted by ultrasonic treatment (6 periods of 30 s, frequency 10 MHz, at 0 °C with 30 s intervals) in 40 mm-potassium phosphate buffer, pH 7.8. Cell debris was removed by centrifuging at 11 600 g for 30 min and supernatants (cell-free extracts) were assayed the same day for fumarase by following the formation of fumarate from L-malate at 240 nm and at 55 °C (unless otherwise stated); activities are expressed in units (nmol fumarate formed min⁻¹) per mg protein using \( \text{B. stearothermophilus} \) (1.2 units) and plasmid (pGS509)-amplified JH400 (29000); cytochrome \( c \) (200 000) determined by gel-filtration with crude extracts (Amersham) was used with intact phagemid DNA and karyotic DNA-directed translation kit containing an E. coli s-30 extract (Amersham) was used with intact phagemid DNA and transformation (Sambrook et al., 1989). The sequencing strategy involved cloning the 0.9 kb, 0.6 kb and 1.4 kb EcoRI fragments of pGS509 into the corresponding site of M13mp18 and sequencing with universal primer (Fig. 1). Deleting \( \text{HindIII} \) and \( \text{SalI} \) fragments from appropriate restriction fragment derivatives allowed further sequence to be obtained. Synthetic primers were used to complete the sequence with pGS509 as the double-stranded template. Single-stranded and double-stranded templates were sequenced by the dyeoxy chain termination method, using Sequenase (Sequenase protocols, United States Biologicals). Oligonucleotides were synthesized using an Applied Biosystems 381 DNA synthesizer.

Nucleotide sequences were compiled and analysed using the Staden computer programs (Staden, 1980, 1982, 1984; Staden & McLachlan, 1982). Multiple sequence alignments were generated initially by the CLUSTAL microcomputer programs (Higgins & Sharp, 1988, 1989) and refined by DIAGON and visual inspection. For other computer analyses, including database searching with TFASTA and FASTA, the University of Wisconsin Genetics Computer Group (UWCGC) package (Devereux et al., 1984) was accessed through the SEQUEN VAX 3600 on node DLVH (Daresbury Laboratory, Warrington, UK).

**In vitro transcription-translation and Southern blotting.** The prokaryotic DNA-directed translation kit containing an E. coli S-30 extract (Amersham) was used with intact phagemid DNA and L-[\( ^{35} \text{S} \)]-

![Restriction maps of phagemids](image)

**Fig. 1.** Restriction maps of phagemids (pGS507, pGS509 and pGS511) containing the \( \text{fumA}_{92} \) gene and summary of sequencing strategy. The unlabelled arrows denote the polarity of \( \text{lacZ} \) transcription in \( \text{pBluescript} \). The enlarged region shows the 3 kb fragment cloned in pGS509, the subfragments cloned in M13mp18 and the \( \text{HindIII} \) or \( \text{SalI} \) deletion derivatives. The positions and polarities of coding regions are indicated by filled arrows and the hybridization probes by broken arrows. In the sequencing strategy, thin arrows denote directions and extents of sequence obtained from M13 derivatives or double stranded pGS509 using universal primer (I) or specifically designed primers (>). Relevant restriction sites are as follows: E, EcoRI; H, HindIII; and S, SalI.
methionine, according to the manufacturer's instructions. For Southern blotting, bacterial DNA was digested with EcoRI, fractionated by electrophoresis (0.8% agarose, Tris/acetate/EDTA gel) and transferred to nitrocellulose membrane (Sambrook et al., 1989). The blots were hybridized for 18 h at 60 °C with 6x SSC (low stringency) and washed (2× 15 min) with 2× SSC plus 0.1% SDS (w/v) at 65°C [2× SSC contains NaCl (0.3M) and sodium citrate (0.03 M, pH 7)]. Probes containing specific regions of four fumarase genes were derived from M13 templates by primer-extension labelling with digoxygenin (Boehringer Mannheim): citG, 0.8 kb BclI-HindIII fragment of pGS96 in M13mp10 (Miles & Guest, 1985); fumC, 0.8 kb XhoI-BclI fragment in M13mp10 (Woods et al., 1986); fumA, 0.7 kb PstI-BclI fragment in M13mp8 (Miles & Guest, 1984); fumABst, M13mp18 derivatives F14LH and F6R, see Fig. 1. The corresponding restriction fragments were isolated by electrophoresis in a low melting point agarose gel for use in hybridization and detection was according to the manufacturer's instructions.

Materials. Restriction endonucleases and DNA ligase were purchased from Northumbria Biologicals, the Sequenase kit was from Cambridge Bioscience, the non-radioactive (digoxygenin) DNA-labelling and detection kit was from Boehringer Mannheim and the prokaryotic DNA-directed translation kit, [35S]methionine (>37 TBq mmol⁻¹) and [α-35S]thio-dATP (>37 TBq mmol⁻¹) were purchased from Amersham. Mₖ markers for SDS-PAGE and gel-filtration were obtained from BDH and Sigma, respectively.

Results and Discussion

Cloning a fumarase gene of *B. stearothermophilus*

A λZAPII gene bank of *B. stearothermophilus* was screened for λ*fum* phages by their ability to produce transduction plaques on *E. coli* JH400 (Δ*fumAC* Δ*fumB*). Two distinct and stably inherited plaque morphologies were detected. Those designated smoke rings were turbid rings of 3–6 mm diameter having clear centres that eventually filled with diffuse micro-colonies, whereas the galaxies consisted of compact clusters of micro-colonies.

![Graph](image-url)

Fig. 2 Aerobic and ammonium persulphate inactivation and molecular properties of FumABst. (a) Extracts were incubated aerobically and anaerobically at 50 °C for different periods prior to assaying for residual fumarase activity (%). Strains and initial specific activities are: JH400(pGS509), aerobic (■), anaerobic (□), 42.3 units (mg protein)⁻¹; *B. stearothermophilus*, aerobic (▲), anaerobic (△), 2.2 units (mg protein)⁻¹. (b) Extracts were treated with ammonium persulphate (1 mM) at 0 °C in air and assayed for residual fumarase activity (%). Strains and initial specific activities are: JH400(pGS509), ammonium-persulphate-treated (■), untreated (□), 14.2 units (mg protein)⁻¹; *B. stearothermophilus*, ammonium-persulphate-treated (▲), untreated (△), 1.2 units (mg protein)⁻¹. (c) Gel-filtration profiles for fumarase activities (units ml⁻¹) with extracts of JH400(pGS509) (■), and *B. stearothermophilus* (△). The vertical arrows denote Mₖ deduced from the elution of standard proteins. (d) Coomassie-blue-stained SDS-PAGE gel (15%) of cell-free extracts of JH400 (lane 1) and JH400(pGS509) (lane 2). FumABst (arrowed) and the positions of Mₖ markers are indicated.
Bacillus stearothermophilus fumarases

The aetiology of these morphologies is unknown, but neither was exactly like the dense transduction plaques obtained with ΔfumA and ΔfumB or the lysogen-filled plaques obtained with the B. subtilis counterpart, ΔcitG (Guest & Roberts, 1983; Moir, 1983). This is probably because ΔZAPII contains a ColEI replicon which would prevent stable prophage integration. Studies with the corresponding lysogens were frustrated by instability and irreproducibility, possibly for the same reason.

The complementing phages were purified, converted to pBluescript phagemids, and studies with double-stranded phagemid DNA from nine independent phages showed that they contain a nested set of three to five EcoRI fragments inserted with the same polarity. Four phagemids (represented by pGS509) contained the common 3 kb insert consisting of three EcoRI fragments (0.92, 0.64 and 1.40 kb), whereas three (represented by pGS511) contained an extra 1.16 kb EcoRI fragment, and two (represented by pGS507) contained a further 1.50 kb EcoRI fragment (Fig. 1). There was no correlation between plaque morphology and inserted DNA.

Characterization of the phagemid-encoded fumarase

The fumarase-deficient strain JH400 grows on malate minimal medium but not with fumarate as the sole carbon source. This nutritional lesion was complemented by transformation with pGS507, pGS509 and pGS511, as with pGS54 (fumAC), pGS96 (citG) and pGS210 (fumC), but not with pBluescript. The fumarase activity [units (mg protein)] in cell-free extracts of JH400 (0.006) was likewise increased in transformants containing pGS507 (45.3), pGS509 (31.8), pGS511 (8.4), pGS54 (91.0), pGS96 (7.4) and pGS210 (40.3). Further studies with JH400(pGS509) indicated that a polypeptide of M_r 58000 is amplified to approximately 40% of soluble cell protein in freshly transformed cells (Fig. 2). In gel-filtration, the fumarase activity eluted with an apparent M_r of 117000 indicating that the enzyme is homodimeric. The fumarase activity was unstable in air and rapidly inactivated by ammonium persulphate (Fig. 2) but the rate of inactivation could be lowered by fumarate (2 mM) and partial reactivation was obtained by incubating with 0.5 mM-Fe^{2+} plus 50 mM-dithiothreitol. These results indicate that the common 3 kb inserts of pGS507, pGS509 and pGS511 express an oxygen-labile homodimeric (Class I) fumarase, FumA_{Bst}. When amplified in E. coli, the enzyme had a thermostability consistent with its source. The temperature optimum was 55–60 °C and, when incubated under anaerobic conditions, the half-lives (min) at different temperatures were: > 100 (50 °C), 20 (60 °C) and < 5

![Fig. 3. In vitro transcription-translation analysis. Autoradiogram of 35S-labelled polypeptides expressed from pGS507 (lane 1), pGS509 (lane 2), pGS511 (lane 3), pBluescript (lane 4), no DNA (lane 5) and pAT153 (lane 6). Relevant labelled products and the positions of M_r markers are indicated.](image-url)
Fig. 4. For legend see facing page.
Bacillus stearothermophilus fumarases

The complete nucleotide sequence of the common 3 kb fragment was determined by the strategy summarized in Fig. 1. The sequence (Fig. 4) was fully overlapped and 98% was derived from both strands. Two complete coding regions were detected using the FRAMESCAN and POSITIONAL BASE PREFERENCES options of ANALYSEQ.

(70 ºC). In contrast, only a fraction (30–40% of the fumarase activity in cell-free extracts of B. stearothermophilus was oxygen-labile (Fig. 2). Indeed, most of the activity was insensitive to ammonium persulphate and eluted with an apparent Mr of 160 000 in gel-filtration, strongly indicating that, like E. coli, B. stearothermophilus contains an oxygen-stable (Class II) fumarase as well as an oxygen-labile (Class I) enzyme.

An in vitro transcription-translation analysis (Fig. 3) showed that the three phagemids express a major labelled product (M, 380 000), presumed to be the fumarase subunit, two minor products (M, 43 000 and 27 000) and β-lactamase (M, 30 000). The novel products exceed the coding capacity of the common insert (M, 100 000 approximately), which could mean that one or both of the minor products are in vitro artefacts or degraded forms of the major product.

Sequence analysis and location of the coding regions

The complete nucleotide sequence of the common 3 kb fragment was determined by the strategy summarized in Fig. 1. The sequence (Fig. 4) was fully overlapped and 98% was derived from both strands. Two complete coding regions were detected using the FRAMESCAN and POSITIONAL BASE PREFERENCES options of ANALYSEQ.
In conjunction with a codon usage table compiled from the open reading frames detected by Mallinder et al. (1992). The first (coordinates 174 to 1715) encodes a product of $M_r$ 56788 (514 amino acid residues including formyl-methionine), which corresponds to the major in-vitro-synthesized product. It could be identified as the Class I fumarase of B. steaotermophilus. $FumA_{Bst}$, by virtue of its sequence similarity with FumA and FumB. The second coding region (orf2, coordinates 1842 to 2636) encodes a product of $M_r$ 30563 (Orf2, 265 amino acids), which probably corresponds to the in-vitro-synthesized product of $M_r$ 27000 (Fig. 3) and resembles the NodB proteins (see below). There is also an incomplete coding region starting at coordinate 2699 and extending for 263 bp to the end sequenced region, but this could not be identified by database searching.

The $fumA_{Bst}$ coding region is preceded by a putative promoter sequence resembling the $\sigma^70$ and $\sigma^43$ promoters of E. coli and B. subtilis (respectively) and a potential ribosome-binding site (Fig. 4). It is followed by a potential stem–loop structure ($\Delta G = -89.9$ kJ mol$^{-1}$) which could serve as a rho-independent terminator. The start codon (TTG) predicted for orf2 is not uncommon in Gram-positive bacilli – e.g. B. subtilis argC (Smith et al., 1986) and gerA (Feavers et al., 1990), and B. pumilis cat86 (Harwood et al., 1983) and it is preceded by a well-placed ribosome-binding site. Neighbouring ATG codons at coordinates 1777, 1822, 1972, 1990 and 2184 (Fig. 4) are not associated with a ribosome-binding site, and in the latter case are situated well downstream of a region of similarity with the NodB proteins (see below). No potential terminator sequences for orf2, or ribosome-binding sites for orf3, were detected.

Multiple sequence alignment and detection of two $fumA$-related genes in E. coli

In addition to the observed similarities between $FumA_{Bst}$ and the FumA and FumB proteins, a database search revealed that OrfZ1 and OrfZ2, the products of adjacent open reading frames located upstream of the rpsU–dnaG–rpoD operon at 67 min in the E. coli linkage map (Nesin et al., 1987) are similar to the respective N- and C-terminal moieties of the FumA-related proteins. This suggests that OrfZ1 and OrfZ2 combine in forming an iron–sulphur-containing hydratase–dehydratase with an $\alpha_2\beta_m$ structure like the L-tartrate dehydratase of Pseudomonas putida (Kelly & Scopes, 1986) or the ($\alpha\beta$), structure like 2-hydroxyglutaryl-CoA dehydratase of Acetaminococcus fermentans (Schweiger et al., 1987), rather than the homodimeric ($\alpha_2$) arrangement in the Class I fumarases.

A multiple sequence alignment for the FumA family is shown in Fig. 5 and an analysis by distances showed that $FumA_{Bst}$ is slightly closer to the combined OrfZ1/Z2 sequence (25% identity, 45% similarity) than to FumA (23%, 38%) and FumB (23%, 37%), which are themselves remarkably close (90%, 98%; Bell et al., 1989). The $FumA_{Bst}$ and OrfZ1/Z2 sequences widen the family of FumA-related proteins and draw attention to potentially important residues and motifs which could perform common structural and catalytic roles. Previous comparisons could only be made between the Class I fumarases of E. coli (FumA/B) and the potentially analogous iron–sulphur proteins (aconitase, ACN; the iron-responsive-element binding protein, IREBP; and isopropylmalate isomerase, IPMI), and the Class II fumarases and their relatives [aspartase, ASP; argininosuccinase, ARS; and adenylosuccinase, ASL – see Prodromou et al. (1992); Woods et al. (1988a, b); Aimi et al. (1990)]. The alignment shows that 58 (12.7%) of 457 equivalent residues are conserved in all of the sequences, including 3 of the 9 or 10 cysteine residues, only 1 of the 11 to 14 histidine residues, and as many as 16 glycine plus proline residues (possibly because the latter residues flank regions of conserved secondary structure).

Cysteine residues generally provide ligands for iron atoms in iron–sulphur proteins via motifs such as CxxCxxC in bacterial ferredoxins or CxxC in ACN, where the [4Fe–4S] cluster has three cysteine ligands (Robbins & Stout, 1989a, b). The three conserved cysteine residues, $C_{62}$, $C_{186}$ and $C_{374}$ at positions 108, 238 and 334 in the alignment (Fig. 5), could presumably perform this function in FumA. Previously, residue $C_{317}$ (position 334) and two residues in the CxxC motif near the C-terminus of FumA were proposed as the three iron ligands (Prodromou et al., 1992) but it is now clear that the CxxC motif is not conserved in the other FumA-related proteins, $FumA_{Bst}$ and OrfZ1/Z2 (Fig. 5).

Only one histidine residue (H$_{396}$ in FumA, position 420) is conserved in the alignment (Fig. 5). This residue was identified previously as a potential active-site residue because its context resembles that of an active-site histidine residue in ACN (Prodromou et al., 1992). The similarity extends to aligned sequences in related enzymes, ACN, IREBP and IPMI (Fig. 6a). However, these sequences show no similarity to those containing the sole conserved histidine in the Class II fumarases, ARS and ASL, where there is a GxTH/XXA consensus (Woods et al., 1988b; Aimi et al., 1990).

There are three conserved lysine residues and two conserved arginine residues in the multiple alignment (Fig. 5) which could provide carboxyl-binding sites. One of the lysine residues (K$_{465}$ in FumA at position 489) is in the only significant sequence motif that is shared by both the FumA (Class I) and FumC (Class II) ASP, ARS and
Fig. 5. A multiple amino acid sequence alignment for three Class I fumarases and the unidentified OrfZ1 and OrfZ2 proteins. Conserved residues are shown below the alignment and coordinates for the alignment and for individual sequences are included. The aligned sequences are: fumarases A (FumA.Eco) and B (FumB.Eco) of E. coli; fumarase A of B. steaothermophilus (FumA.Bst); and OrfZ1 (OrfZ1.Eco) and OrfZ2 (OrfZ2.Eco) of E. coli. Conserved cysteine residues (I), a putative active-site residue (V), and insertion-deletions (.), are indicated.
Fig. 6. Sequence similarities between the Fe-S-containing fumarases and related enzymes. (a) Sequences resembling an active-site histidinyl peptide of porcine aconitase. (b) Sequences containing a motif which is conserved between Class I and Class II fumarases and related enzymes that catalyse \( \beta \)-elimination reactions yielding fumarate. The sequence sources and abbreviations are as follows: FumA\textsubscript{nat}, FumA.Bst; FumA and FumB of \textit{E. coli}, FumAB.Eco (Miles & Guest, 1984; Bell \textit{et al.,} 1989); OrfZ2, OrfZ2.Eco (Nesin \textit{et al.,} 1987); FumC, FumC.Eco (Woods \textit{et al.,} 1986); aspartase, AspA. Eco (Woods \textit{et al.,} 1986); adenylsuccinase, ASL. Eco (He \textit{et al.,} 1992); aconitase, ACN. Eco (Prodromou \textit{et al.,} 1992); porcine heart aconitase, ACN. Por (Zheng \textit{et al.,} 1990); isopropylmalate isomerase, IPMI.Sty (Rosenthal & Calvo, 1990); human iron-responsive-element binding protein, IREBP.Hum (Rouault \textit{et al.,} 1990); \textit{B. subtilis} fumarase, CitG.Bsu (Miles & Guest, 1985); human argininosuccinase, ARS.Hum (Aimi \textit{et al.,} 1990).

ASL protein families (Fig. 6b). This motif was originally discovered by Woods \textit{et al.,} (1988a, b) and later designated as the `signature sequence' for enzymes which catalyse \( \beta \)-elimination reactions (Aimi \textit{et al.,} 1990). However, with the inclusion of further sequences, the original G\textit{X}\textit{X}Mxx\textit{Kx}N consensus has degenerated to G\textit{XX}\textit{X}xx\textit{Kxx}, but it still draws attention to a potentially important lysine residue (Fig. 6b). It may also be significant that the leucine and glutamate dehydrogenases contain a conserved lysine residue in a KGGM motif (McPherson \textit{et al.,} 1988; Nagata \textit{et al.,} 1988; Teller \textit{et al.,} 1992) and such a motif occurs at positions 192-195 in the three Class I fumarase sequences (Fig. 5). This motif is not retained in OrfZ1/Z2, where the substrate is unknown, but there are other lysine residues within conserved KG and K\textit{X}G\textit{X}/p motifs at positions 489-490 and 119-121, respectively.

Distribution of the Class I and Class II fumarase genes

Hybridization probes containing segments of the \textit{fumA}, \textit{fumA}\textsubscript{nat}, \textit{fumC} and \textit{citG} genes were used with \textit{EcoRI} digests of bacterial DNA to assess the distribution of the two classes of fumarase in different species by Southern blotting. Identical results were obtained with the two \textit{fumA}\textsubscript{nat} probes shown in Fig. 1, specific hybridization being detected with \textit{B. stearothermophilus} NRRL1174 (23 kb) and \textit{B. caldotenax} (9.5 kb) but not with \textit{B. subtilis} 1604 or \textit{E. coli} (Fig. 7a). No hybridization was detected between the \textit{E. coli} \textit{fumA} probe and any of the \textit{Bacillus} species. Using the cross-hybridizing \textit{citG} or \textit{fumC} probes (Guest \textit{et al.,} 1985), hybridization was detected with \textit{B. stearothermophilus} NRRL1174 (8.8 kb), \textit{B. caldotenax} (10 kb) and \textit{B. stearothermophilus} var. non-diastaticus (4.1 kb) (see Fig. 7b for results with the \textit{citG} probe).

These results indicate that the thermophilic bacilli contain a Class I fumarase gene (analogous to \textit{fumA}\textsubscript{nat}) and a Class II fumarase gene (analogous to the \textit{fumC} and \textit{citG} genes of \textit{E. coli} and \textit{B. subtilis}). The \textit{B. stearothermophilus} gene bank is now being screened by hybridization for phages containing \textit{fumC}\textsubscript{nat} DNA since none was recovered by complementation. Despite the absence of hybridization with \textit{B. subtilis} DNA and the lack of fumarase activity in \textit{citG} mutants (Rutberg & Hoch, 1970), it cannot be concluded that \textit{B. subtilis} has no Class I fumarase, because the corresponding \textit{B. subtilis} gene may not hybridize with the \textit{fumA} probes and an oxygen-labile product may have escaped detection in previous work.

Features of Orf2

Database searches revealed a significant sequence relationship (21–39% identity, 39–52% similarity) between the 211-residue C-terminal segment of Orf2 and the NodB proteins of \textit{Rhizobium} and \textit{Bradyrhizobium} (Fig. 8). The precise function of the \textit{nodB} gene has not been defined but it forms part of the \textit{nodABC} operon,
Bacillus stearothermophilus fumarases

which is involved in the initiation of root nodulation, possibly by generating signalling molecules that stimulate plant cell mitosis (Downie, 1991; Schmidt et al., 1991; Fisher & Long, 1992). Compared with the NodB proteins, Orf2 has a 54-residue N-terminal extension possessing some features of a signal peptide over the first 23 residues. It is not clear why *B. stearothermophilus* might express such a protein but further studies might help to define a common role for Orf2 and the NodB proteins in *Bacillus* and *Rhizobium*.

**Conclusions**

Molecular genetic and biochemical studies have shown that *B. stearothermophilus* joins *E. coli* and *Bradyrhizobium japonicum* (Acuña et al., 1991) in containing two types of fumarase. The *fumA* gene encoding a Class I fumarase of *B. stearothermophilus* was cloned by nutritional complementation of a fumarase-deficient mutant of *E. coli*, sequenced, and identified by its ability to express an oxygen-labile fumarase and by the sequence similarity between FumA and the FumA and FumB enzymes of *E. coli*. Evidence for a Class II fumarase (*FumC*) was obtained from enzyme studies and DNA hybridization. The reason why these bacteria possess two classes of fumarase is not clear but recent studies with *E. coli* indicate that *fumC* belongs to the *soxR* modulon (Liochev & Fridowich, 1992). The *soxR* regulatory system operates during oxidative stress to maintain cellular redox balance and normal metabolic flux, and this would be favoured by replacing the unstable FumA by FumC. In contrast, *fumA* is a member of the *arcA* and *crp* modulons, i.e. subject to aerobic respiratory control and catabolite repression, whereas *fumB* is subject to anaerobic repression and belongs to the *fnr* modulon (Woods & Guest, 1987; Guest & Russell, 1992).

Database searches revealed significant similarity between the Class I fumarases and the products, OrfZ1 and OrfZ2, of two adjacent *E. coli* genes, indicating that the latter combine to form an oxygen-labile, iron–sulphur-containing hydratase–dehydratase. Work is now in progress to isolate the *fumC* gene and to characterize the OrfZ1/Z2 enzyme. Putative active-site residues and structural motifs were detected by sequence comparisons with aconitase, isopropylmalate isomerase and other related enzymes. *B. stearothermophilus* was also shown to contain a gene that resembles the *nodB* genes of *Rhizobium* and *Bradyrhizobium* where it is essential for the initiation of root nodulation.

Fig. 7. Southern hybridization of *fumA* and *citG* DNA probes to EcoRI fragments of total DNA from other bacteria. The digoxygenin-labelled probes were (a) *fumA*, F14LH (Fig. 1) and (b) *citG* from *B. subtilis* (see Methods), the sizes (kb) of the hybridizing fragments are shown for digests of the following bacteria: 1, *B. subtilis* 1604; 2, *E. coli* W3110; 3, *B. stearothermophilus* NRL1174; 4, *B. caldotenax* DSM406; and 5, *B. stearothermophilus* var. *non-diastaticus*.
Fig. 8. Multiple amino acid sequence alignment for NodB proteins and Orf2. Conserved residues are shown below the alignment and coordinates are given for the alignment and for the C-terminal residues of individual segments. The sequences and sources are: NodB. Rm4, *Rhizobium meliloti* 41 (Torok et al., 1984); NodB. Rm1, *Rhizobium meliloti* 1021 (Egelhoff et al., 1985); NodB. Rle, *Rhizobium leguminosarum* plasmid pRL1J1 (Rossen et al., 1984); NodB. Rlu, *Rhizobium leguminosarum* (sequence unverified, submitted to EMBL Data Library, accession number S10232, by J. A. Downie); NodB. Rtr, *Rhizobium leguminosarum* bv. *trifolii* (Schofield & Watson, 1986); NodB. Brs, *Bradyrhizobium* sp. strain ANU289 (Scott, 1986); NodB. Aca, *Azorhizobium caulinodans* (Goethals et al., 1989).
We are very grateful to P. R. Mallinder, A. Moir, S. C. Andrews and C. Prodromou for helpful advice and materials. The work was supported by a SERC-CASE studentship (S. K. R.) with ICI Bio-Products and Fine Chemicals.

References


Prodromou, C., Artymiuk, P. J. & Guest, J. R. (1992). The aconitase of Escherichia coli: nucleotide sequence of the aconitase gene and amino acid sequence similarity with mitochondrial aconitases, the


