JOURNAL OF BACTERIOLOGY, Dec. 1990, p. 7035–7042
0021-9193/90/127035-08$02.00/0

Vol. 172, No. 12

Molecular Analysis of Two Genes of the *Escherichia coli gab* Cluster: Nucleotide Sequence of the Glutamate:Succinic Semialdehyde Transaminase Gene (*gabT*) and Characterization of the Succinic Semialdehyde Dehydrogenase Gene (*gabD*)

KLAUS BARTSCH,* ASTRID VON JOHNN-MARTEVILLE, AND ARNO SCHULZ

Hoechst AG, 6230 Frankfurt 80, Federal Republic of Germany

Received 27 June 1990/Accepted 20 September 1990

We have characterized two genes of the *Escherichia coli* K-12 *gab* cluster, which encodes the enzymes of the 4-aminobutyrate degradation pathway. The nucleotide sequence of *gabT*, coding for glutamate:succinic semialdehyde transaminase (EC 2.6.1.19), alternatively known as 4-aminobutyrate transaminase, was determined. The structural gene consists of 1,281 nucleotides specifying a protein of 426 amino acids with a molecular mass of 45.76 kDa. The protein shows significant homologies to the ornithine transaminases from *Saccharomyces cerevisiae* and from rat and human mitochondria. Three functionally and structurally important amino acid residues of the transaminase were identified by sequence comparison studies, and evolutionary relationships of the aminotransferases are discussed. The *gabD* gene, encoding succinic semialdehyde dehydrogenase (EC 1.2.1.16), was cloned and shown to be located adjacent to the 5′ end of *gabT*. Expression studies with subfragments of the initially cloned DNA region revealed a maximal size of 1.7 kb for *gabD*. Both genes are cotranscribed from a promoter located upstream of *gabD*.

The *gab* cluster of *Escherichia coli* specifies the synthesis of the enzymes of the 4-aminobutyrate (GABA) degradation pathway (8, 9). The cluster, which is located at 57.6 min on the *E. coli* K-12 chromosome, was mapped genetically by Metzer et al. (22) and shown to contain four genes: *gabT*, encoding glutamate:succinic semialdehyde transaminase (GSST; GABA transaminase; EC 2.6.1.19); *gabD*, encoding succinic semialdehyde dehydrogenase (SSDH; EC 1.2.1.16); *gabP*, encoding GABA permease; and a control gene, *gabC*, coordinately regulating their expression. In a previous work, we purified and characterized a GABA transaminase from *E. coli* K-12 (26). The corresponding gene was cloned on a 1.6-kb *Dra*I-*Bam*HI fragment and overexpressed in *E. coli*, and its identity with *gabT* could be demonstrated (3). The *gabT* gene was shown to be situated on a 3.8-kb *Sal*I-*Bam*HI fragment together with the endogenous promoter (3). Moreover, we presented a restriction map of a 15-kb *Sal*I fragment containing most of the *E. coli gab* cluster (3). Metzer and Halpern recently also reported the cloning of the *E. coli* K-12 *gab* region and suggested that the *gab* genes are divergently transcribed by two different promoters (21).

In this paper, we report the complete nucleotide sequence of the GABA transaminase gene, *gabT*, and discuss the evolutionary relatedness of GABA transaminase to other aminotransferases. Furthermore, we describe the cloning and characterization of *gabD* and provide evidence that *gabT* and *gabD* are transcribed from a common promoter upstream of the *gabD* gene.

MATERIALS AND METHODS

Bacterial strains, media, enzymes, and chemicals. For the transformation and expression experiments, the *E. coli* DH1 (17) and JM103 (20) were used. Plasmid-carrying transformants were grown in Luria-Bertani medium (18) with 50 mM Tris hydrochloride (pH 7.5) and 50 μg of the appropriate antibiotic per ml. For induction with isopropylthiogalactoside (IPTG), 1 mM IPTG was added to the bacterial cultures at an optical density of 0.5, and the cells were harvested 4 h after induction. Restriction enzymes, T4 DNA ligase, T4 DNA polymerase, exonuclease III, and polynucleotide kinase were obtained from New England BioLabs, Inc., Beverly, Mass. Alkaline phosphatase, nuclease S1, the large fragment of DNA polymerase I (Klenow enzyme), and the deoxy- and dideoxyribonucleotides for DNA sequencing were obtained from Boehringer GmbH, Mannheim, Federal Republic of Germany. [α-^{35}S]dATP was purchased from New England Nuclear Research Products, Dreieich, Federal Republic of Germany.

Plasmid constructions and subcloning experiments. The plasmids used and constructed in this work are shown in Table 1. The vector pJF$_{tac}$- was derived from pJF118u by deletion of a ca. 0.3-kb *Sph*I-*Eco*RI fragment containing the *tac* promoter. Deletions of the 3.8-kb *Sal*I-*Bam*HI fragment in the vector pMLC13 were constructed from either end by digestion with exonuclease III and nuclease S1 by the method of Henikoff (12). Restriction fragments from the same DNA region were subcloned in the plasmid vectors pMLC12, pMLC13, pJF118u, and pJF$_{tac}$- by standard methods described by Maniatis et al. (18). The resulting constructs were tested for their ability to confer elevated GSST and SSDH activities.

Sequence and computer analyses. The sequence analysis of the 1.6-kb *Dra*I-*Bam*HI fragment subcloned in the vector pMLC13 was carried out by the chain termination method of Sanger et al. (25) with double-stranded DNA templates (6) and with [α-^{35}S]dATP as the radioactive label. The sequence information obtained from this fragment and a number of subfragments was extended by the use of suitable oligonucleotides as sequencing primers; these were synthesized in the Hoechst AG Pharmaceutical Department by automated phosphoramidite chemistry (28). DNA and protein homology searches were conducted with the standard alignment

* Corresponding author.

TABLE 1. Plasmids used and constructed in this work

Plasmid	Description	Source or reference
pMLC12 and pMLC13	Vector, Cm^r, *lac* promoter, pUC12 polylinker in opposite orientations	16
pJF118u	Vector, Ap^r, *tac* promoter, $lacI^q$	11
pJF_{tac}^-	Derived from pJF118u by deletion of the *tac* promoter	This work
pGAB1 (pGT3)	3.8-kb *Bam*HI-*Sal*I fragment in pMLC13	3
$pGAB1_R$	3.8-kb *Bam*HI-*Sal*I fragment in pMLC12	This work
pGAB2 to pGAB5	Deletions of the 3.8-kb *Bam*HI-*Sal*I fragment from the *Sal*I end (3.5, 3.0, 2.8, and 2.5-kb, respectively) in pMLC13	This work
pGAB6	2.9-kb *Aat*II-*Bam*HI fragment in pMLC12	This work
pGAB7	2.2-kb *Sal*I-*Dra*I fragment in pMLC13	This work
pJF1	3.8-kb *Bam*HI-*Sal*I fragment in pJF118u	This work
pJF2	1.6-kb *Dra*I-*Bam*HI fragment in pJF118u	This work
$pJF1_{tac}^-$	3.8-kb *Bam*HI-*Sal*I fragment in pJF_{tac}^-	This work
$pJF2_{tac}^-$	1.6-kb *Dra*I-*Bam*HI fragment in pJF_{tac}^-	This work

programs HIBIO DNASIS and HIBIO PROSIS, distributed by Hitachi Software Engineering Co., Ltd., Yokohama, Japan. The sequence data bases used were GenBank, compiled by the National Institute of Health; EMBL, compiled by the European Molecular Biology Laboratory; and Protein Identification Resource, compiled by the National Biomedical Research Foundation.

Preparation of cell extracts. Crude protein extracts were prepared by two different methods. (i) For determination of GABA transaminase activity, the bacteria were harvested by centrifugation and washed twice with 10 mM NaCl–10 mM sodium phosphate (pH 7.0). The cells were resuspended in 20 mM sodium phosphate buffer containing 0.01 mM pyridoxalphosphate, 5 mM 2-mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride (pH 7.0) (3 ml of buffer per g of cells) and disrupted by sonication. After removal of the cell debris by centrifugation for 10 min at 6,000 × *g*, the supernatant was used for enzyme testing. (ii) For determination of SSDH activity, the bacteria were collected and washed as described above. The cells were resuspended in 100 mM sodium phosphate buffer containing 9% glycerol, 1 mM dithiothreitol, and 1 mM phenylmethylsulfonyl fluoride (pH 7.0) (1 ml of buffer per g of cells) and disrupted by sonication. The resultant slurry was centrifuged for 10 min at 6,000 × *g*. The supernatant was further purified by centrifugation for 90 min at 128,000 × *g* and subsequently used in the enzyme assay.

Protein determination. The protein content of the cell extracts was determined by the Coomassie brilliant blue-binding assay (Bio-Rad Laboratories, Munich, Federal Republic of Germany) of Bradford (5).

Assay of GABA transaminase and of SSDH. The crude protein extracts prepared by method i were tested for GABA transaminase activity as described previously (26) with test method 2, with the modification that 50 mM 2-ketoglutarate and 150 mM GABA were used as substrates. The formation of glutamate was determined with a Biotronic LC 5001 amino acid analyzer.

The SSDH assay was carried out by incubating the cell extracts prepared by method ii in a 1-ml volume at 30°C with 100 mM sodium phosphate–0.14 mM NADP–0.1 mM succinic semialdehyde (pH 7.0). The conversion of succinic semialdehyde to succinate was determined by measuring the reduction of NADP to NADPH with a Beckman Du-50 spectrophotometer at 340 nm.

Nucleotide sequence accession number. The nucleotide sequence data reported will appear in the EMBL, GenBank, and DDBJ nucleotide sequence data bases under the accession number M38417.

RESULTS

Determination of the *gabT* DNA sequence. As shown previously, the complete *gabT* gene is located on a 1.6-kb *Dra*I-*Bam*HI fragment of the *E. coli* chromosome (3). The nucleotide sequence of this fragment was determined by the chain termination method of Sanger et al. (25). Internal sequencing templates were obtained by digestion of the fragment with exonuclease III-nuclease S1 or suitable restriction enzymes. Alternatively, synthetic oligonucleotides were used as sequencing primers. The *Dra*I-*Bam*HI site region was sequenced on both strands. The sequencing strategy is depicted in Fig. 1. The nucleotide sequence is shown in Fig. 2.

GABA transaminase protein. The analysis of the DNA sequence revealed one long open reading frame of 1,287 nucleotides with two potential methionine (start) codons at positions 269 and 275. A comparison of the deduced amino acid sequence (Fig. 2) with the N-terminal amino acid sequence of the GABA transaminase protein (26) suggests that the second methionine codon at position 275 is the authentic translation start of the gene. Thus, the *gabT* gene is 1,281 nucleotides long and encodes a 426-amino-acid protein with a calculated molecular mass of 45.76 kDa. These data correlate well with the 43-kDa band observed on polyacrylamide gel electrophoresis of purified GABA transaminase (26). On the other hand, a molecular mass of 90 kDa was determined for the native protein by gel filtration experiments under nondenaturing conditions (data not shown). The results indicate that the enzymatically active GABA transaminase protein is a homodimer composed of two 45.76-kDa subunits.

Amino acid sequence homologies. A comparison of the GABA transaminase sequence with all known transaminase sequences from the GenBank, EMBL, and National Biomedical Research Foundation data bases was carried out with the standard alignment program HIBIO PROSIS. The GABA transaminase protein showed significant homologies to the ornithine transaminases (OAT) from *Saccharomyces cerevisiae* (SCOAT) and from rat and human mitochondria. The sequences were compared for identical amino acids and conservative substitutions as described by Barcer and Dayhoff (2); they exhibited identity and similarity scores of 28.2 and 46.0% between GABA transaminase and OAT from human mitochondria 27.9 and 45.3% between GABA transaminase and OAT from rat mitochondria, and 27.5 and 44.8% between GABA transaminase and SCOAT, respectively. The sequence comparison of GABA transaminase, SCOAT, and OAT from rat mitochondria is shown in Fig. 3.

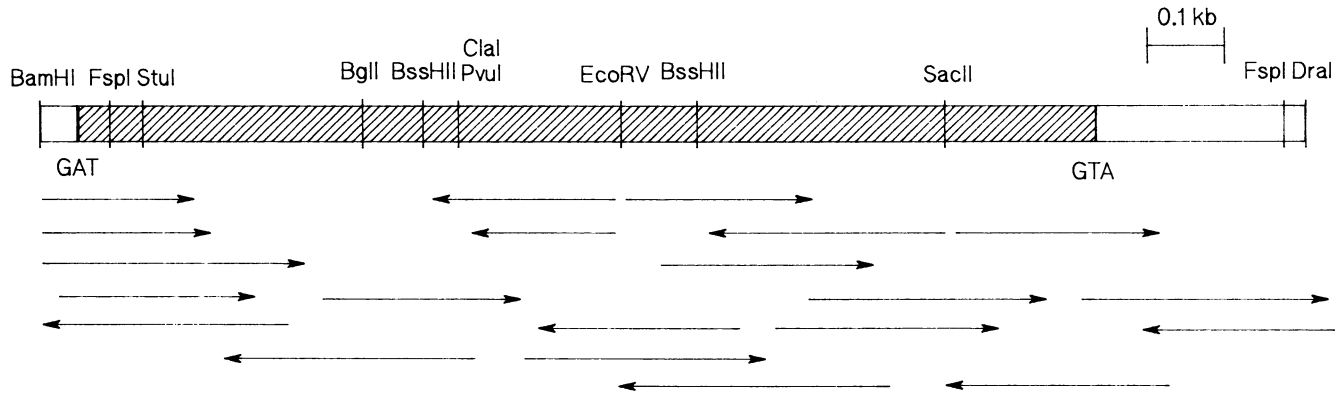

FIG. 1. Sequencing strategy for the 1.6-kb *Dra*I-*Bam*HI fragment. The position and orientation of the *gabT* gene are indicated by the translational start and stop codons (ATG, TAG). Arrows represent the direction and extent of sequences.

No significant homology to other aminotransferases was observed. A more rigorous search for conserved sequence segments was conducted with the information on amino acids with a specific structural or functional role in aspartate, tyrosine, and histidinol-phosphate transaminases and OAT, as discussed by Mehta et al. (19) and Mueckler and Pitot (23). By comparing a number of highly homologous peptides which are conserved among 16 aminotransferases (19) to the GABA transaminase, we found that 3 of the 12 invariant amino acid residues could be identified in the GABA transaminase protein (Fig. 4). Lys-268 in GABA transaminase corresponds to Lys-258 in cytosolic chicken aspartate aminotransferase (AAT); this residue is known to form a Schiff base with the cofactor pyridoxalphosphate. Arg-398 corresponds to Arg-386 in cytosolic chicken AAT; this residue binds the α-carboxylate group of the substrate. Ala-304 corresponds to Ala-303 in cytosolic chicken AAT; the function of this residue is unknown. Ala-304 seems not to be conserved in the OAT sequences.

Localization of the *gabD* gene. It is known that the *gabD* gene, which encodes SSDH, is located adjacent to *gabT* in the *E. coli gab* cluster (22) (Fig. 5). As shown previously (3), the 1.6-kb *Dra*I-*Bam*HI fragment contains the *gabT* gene and is situated directly at the *Bam*HI end of a 3.8-kb *Sal*I-*Bam*HI fragment. To elucidate whether the region upstream of *gabT* could encode the SSDH protein, we tested recombinant bacteria harboring the entire 3.8-kb fragment as well as subcloned parts of it for expression of elevated SSDH activity (Table 2). *E. coli* transformants carrying the complete 3.8-kb fragment in pMLC13, i.e., in the same orientation as that which causes a high overexpression of the *gabT* gene (3), exhibited a ca. 30-fold elevation of SSDH activity. When the same fragment was cloned in the wrong orientation with respect to the *lac* promoter into the vector pMLC12, SSDH activity was still threefold higher than the background level. These results indicate that the complete structural *gabD* gene, together with the endogenous promoter, is located on the 3.8-kb *Sal*I-*Bam*HI fragment and that it is transcribed in the same direction as the *gabT* gene. The position of the gene was further mapped by constructing a series of exonuclease III-nuclease S1 deletions from the *Sal*I end of the fragment in the vector pMLC13. High levels of SSDH activity were observed in bacteria carrying either a 3.5-kb or a 3.0-kb digest, while the enzyme activity dropped to the background level in transformants carrying shorter parts of the 3.8-kb fragment (Fig. 5). Furthermore, we subcloned a 2.9-kb *Aat*II-*Bam*HI fragment into the *Sma*I-*Bam*HI-digested vector pMLC12 and a 2.2-kb *Sal*I-*Dra*I fragment into the *Sal*I-*Sma*I-digested vector pMLC13, both in the correct orientation with respect to the *lac* promoter. None of these constructs conferred elevated SSDH activity. From these results, we concluded that the 5′ end of the *gabD* structural gene is located between the *Aat*II site at 2.9 kb and the 3.0-kb deletion (Fig. 5). The 3′ end of the gene is situated between the *Dra*I site at 1.6 kb and the start of the *gabT* structural gene at 1.3 kb. The maximal possible size of the *gabD* gene is therefore 1.7 kb, leading to a gene product with a maximum of 566 amino acids and a molecular mass of ca. 60 kDa.

***gabT/gabD* promoter.** When the 3.8-kb *Bam*HI-*Sal*I fragment was cloned in the wrong orientation with respect to the *lac* promoter into the vector pMLC12, a certain level of expression of both genes (*gabT* and *gabD*) was still maintained (Table 2), indicating the presence of endogenous promoter activities on the fragment. These promoter activities were investigated more thoroughly in the vector derived from pJF118u by deletion of the *tac* promoter (Table 1).

Transformants carrying the 3.8-kb *Sal*I-*Bam*HI fragment subcloned in the *Sma*I-*Bam*HI-digested vector pJF_{tac^-} (plasmid $pJF1_{tac^-}$) expressed both GABA transaminase and SSDH at levels about fivefold higher than background levels (Table 3). On the other hand, transformants harboring the 1.6-kb *Dra*I-*Bam*HI fragment subcloned in the same vector (plasmid $pJF2_{tac^-}$) exhibited no elevated expression levels for either of the two genes. The 1.6-kb *Dra*I-*Bam*HI fragment contains the complete *gabT* structural gene as well as the 3′ end of the *gabD* gene (Fig. 5). Therefore, it is obvious that *gabT* is not preceded by its own promoter but that both genes are transcribed coordinately from a common promoter located upstream of *gabD*.

To compare the strength of the *gabT/gabD* promoter with that of a well-characterized *E. coli* promoter, we cloned the 3.8-kb *Sal*I-*Bam*HI fragment and the 1.6-kb *Dra*I-*Bam*HI fragment downstream of the *tac* promoter of the *Sma*I-*Bam*HI-digested vector pJF118u (Table 3). The expression of the genes in these constructs was inducible by IPTG, and expression levels were ca. 40-fold higher than the background level. These results suggest that *gab* promoter activity is roughly 1 order of magnitude lower than *tac* promoter activity.

DISCUSSION

As is evident from the genetic map established by Metzer et al. (22), the *E. coli gab* cluster comprises three structural

```
   1    T TTA AAG ATG AAG CTG ATG TGA TTG CGC AAG CCA ATG ACA CCG AGT
  47  TTG GCC TTG CCG CCT ATT TCT ACG CCC GTG ATT TAA GCC GCG TCT TCC
  95  GCG TGG GCG AAG CGC TGG AGT ACG GCA TCG TCG GCA TCA ATA CCG GCA
 143  TTA TTT CCA ATG AAG TGG CCC CGT TCG GCG GCA TCA AAG CCT CGG GTC
 191  TGG GTC GTG AAG GTT CGA AGT ATG GCA TCG AAG ATT ACT TAG AAA TCA
 239  AAT ATA TGT GCA TCG GTC TTT AAC TGG AGA ATG CGA ATG AAC AGC AAT
   1                                              M   N   S   N
 287  AAA GAG TTA ATG CAG CGC CGC AGT CAG GCG ATT CCC CGT GGC GTT GGG
   5   K   E   L   M   Q   R   R   S   Q   A   I   P   R   G   V   G
 335  CAA ATT CAC CCG ATT TTC GCT GAC CGC GCG GAA AAC TGC CGG GTG TGG
  21   Q   I   H   P   I   F   A   D   R   A   E   N   C   R   V   W
 383  GAC GTT GAA GGC CGT GAG TAT CTT GAT TTC GCG GGC GGG ATT GCG GTG
  37   D   V   E   G   R   E   Y   L   D   F   A   G   G   I   A   V
 431  CTC AAT ACC GGG CAC CTG CAT CCG AAG GTG GTG GCC GCG GTG GAA GCG
  53   L   N   T   G   H   L   H   P   K   V   V   A   A   V   E   A
 479  CAG TTG AAA AAA CTG TCG CAC ACC TGC TTC CAG GTG CTG GCT TAC GAG
  69   Q   L   K   K   L   S   H   T   C   F   Q   V   L   A   Y   E
 527  CCG TAT CTG GAG CTG TGC GAG ATT ATG AAT CAG AAG GTG CCG GGC GAT
  85   P   Y   L   E   L   C   E   I   M   N   Q   K   V   P   G   D
 575  TTC GCC AAG AAA ACG CTG CTG GTT ACG ACC GGT TCC GAA GCG GTG GAA
 101   F   A   K   K   T   L   L   V   T   T   G   S   E   A   V   E
 623  AAC GCG GTA AAA ATC GCC CGC GCC GCC ACC AAA CGT AGC GGC ACC ATC
 117   N   A   V   K   I   A   R   A   A   T   K   R   S   G   T   I
 671  GCT TTT AGC GGC GCG TAT CAC GGG CGC ACG CAT TAC ACG CTG GCG CTG
 133   A   F   S   G   A   Y   H   G   R   T   H   Y   T   L   A   L
 719  ACC GGC AAG GTG AAT CCG TAC TCT GCG GGC ATG GGG CTG ATG CCG GGT
 149   T   G   K   V   N   P   Y   S   A   G   M   G   L   M   P   G
 767  CAT GTT TAT CGC GCG CTT TAT CCT TGC CCG CTG CAC GGC ATA AGC GAG
 165   H   V   Y   R   A   L   Y   P   C   P   L   H   G   I   S   E
 815  GAT GAC GCT ATC GCC AGC ATC CAC CGG ATC TTC AAA AAT GAT GCC GCG
 181   D   D   A   I   A   S   I   H   R   I   F   K   N   D   A   A
 863  CCG GAA GAT ATC GCC GCC ATC GTG ATT GAG CCG GTT CAG GGC GAA GGC
 197   P   E   D   I   A   A   I   V   I   E   P   V   Q   G   E   G
 911  GGT TTC TAC GCC TCG TCG CCA GCC TTT ATG CAG CGT TTA CGC GCT CTG
 213   G   F   Y   A   S   S   P   A   F   M   Q   R   L   R   A   L
 959  TGT GAC GAG CAC GGG ATC ATG CTG ATT GCC GAT GAA GTG CAG AGC GGC
 229   C   D   E   H   G   I   M   L   I   A   D   E   V   Q   S   G
1007  GCG GGG CGT ACC GGC ACG CTG TTT GCG ATG GAG CAG ATG GGC GTT GCG
 245   A   G   R   T   G   T   L   F   A   M   E   Q   M   G   V   A
1055  CCG GAT CTT ACC ACC TTT GCG AAA TCG ATC GCG GGC GGC TTC CCG CTG
 261   P   D   L   T   T   F   A   K   S   I   A   G   G   F   P   L
1103  GCG GGC GTC ACC GGG CGC GCG GAA GTA ATG GAT GCC GTC GCT CCA GGC
 277   A   G   V   T   G   R   A   E   V   M   D   A   V   A   P   G
1151  GGT CTG GGC GGC ACC TAT GCG GGT AAC CCG ATT GCC TGC GTG GCT GCG
 293   G   L   G   G   T   Y   A   G   N   P   I   A   C   V   A   A
1199  CTG GAA GTG TTG AAG GTG TTT GAG CAG GAA AAT CTG CTG CAA AAA GCC
 309   L   E   V   L   K   V   F   E   Q   E   N   L   L   Q   K   A
1247  AAC GAT CTG GGG CAG AAG TTG AAA GAC GGA TTG CTG GCG ATA GCC GAA
 325   N   D   L   G   Q   K   L   K   D   G   L   L   A   I   A   E
1295  AAA CAC CCG GAG ATC GGC GAC GTA CGC GGG CTG GGG GCG ATG ATC GCC
 341   K   H   P   E   I   G   D   V   R   G   L   G   A   M   I   A
1343  ATT GAG CTG TTT GAA GAC GGC GAT CAC AAC AAG CCG GAC GCC AAA CTC
 357   I   E   L   F   E   D   G   D   H   N   K   P   D   A   K   L
1391  ACC GCC GAG ATC GTG GCT CGC GCC CGC GAT AAA GGC CTG ATT CTT CTC
 373   T   A   E   I   V   A   R   A   R   D   K   G   L   I   L   L
1439  TCC TGC GGC CCG TAT TAC AAC GTG CTG CGC ATC CTT GTA CCG CTC ACC
 389   S   C   G   P   Y   Y   N   V   L   R   I   L   V   P   L   T
1487  ATT GAA GAC GCT CAG ATC CGT CAG GGT CTG GAG ATC ATC AGC CAG TGT
 405   I   E   D   A   Q   I   R   Q   G   L   E   I   I   S   Q   C
1535  TTT GAT GAG GCG AAG CAG TAG CGC CGC TCC TAT GCC GGA GAG CAC TGC
 421   F   D   E   A   K   Q   *
1583  GCG TCT TGT CCG GCC TAC GGG GAT CC
```

FIG. 2. Nucleotide sequence of the 1.6-kb *Dra*I-*Bam*HI fragment containing the *gabT* gene. The DNA sequence is numbered beginning at the *Dra*I site. The deduced amino acid sequence of the GABA transaminase (GSST) protein is given in the bottom line. *, Termination codon.

```
ECGSST   1  MNSNKELMQRRSQAIPRGVGQIHPIFADRAENCRVWDVEGREYLDFAGGIAVLN
SCOAT    4  ATLSKQTIEWENKYSAHNYHPLPVVF-HKAKGAHLWDPEGKLYLDFLSAYSAVN
mROAT   37  PPSSEYIFERESKYGAHNYHPLPVAL-ERGKGIYMWDVEGRQYFDFLSAYGAVS
                * *   .* *   ..    .   *  **    .****** **** .. . .*

55  TGHLHPKVVAAVEAQLKKLSHTCFQVLAYEPYLELCEIMNQKVPGDFAKKTLLVTTGSEAVE
57  QGHCHPHIIKALTEQAQTLTLSS-RAFHNDVYAQFAKFVTE--FSGF-ETVLPMNTGAEAVE
90  QGHCHPKIIEAMKSQVDKLTLTS-RAFYNNVLGEYEEYITK--LFNY-NKVLPMNTGVEAGE
     ** ***.. *.  *  **. *       . * *  * .       *  * * . ** ****

117 NAVKIARA--------ATKRSGTIAFSGAYHGRTHYTLALTGKVNPYSAGMGLMPGHVYRAL
115 TALKLARRWGYMKKNIPQDKAIILGAEGNFHGRTFGAISLSTDYEDSKLHFGPFVPNVASG-
148 TACKLARRWGYTVKGIQKYKAKIVFAVGNFWGRTLSAVSSSTD-PTSYDGFGPFMP------
     * *.**             ..  .   * .****  ...*..       * *.  . *

171 YPCPLHGISEDDAIASIHRIFKNDAAPEDIAAIVIEPVQGEGGFYASSPAFMQRLRALCDEH
176 --HSVHKIRYGHA-EDFVPIL-ESPEGKNVAAIILEPIQREAGIVVPPADYFPKVSALCRKH
203 ---GFETIPYNDL-PALERAL-QDP---NVAAFMVEPIQGEAGVIVPDPGYLTGVRELCTRH
       . * *   **      **  .*.   ..***..**.***.*   . * .   .****  *

233 GIMLIADEVQSGAGRTGTLFAMEQ--MGVAPDLTTFAKSIAGG-FPLAGVTGRAEVMDAVAP
234 NVLLIVDEIQTGIGRTGELLCYDHYKAEAKPDIVLLGKALSGGVLPVSCVLSSHDIMSCFTP
257 QVLFIADEIQTGLARTGRWLAVDH--ENVRPDIVLLGKALSGGLYPVSAVLCDDDIMLTIKP
     ..*****.*.* **** * * .     * **.   .*...**  *.. *    ..*    *

292 GGLGGTYAGNPIACVAALEVLKVFEQENLLQKANDLGQKLKDGLLAIAEK-HPEIGDVRGLG
296 GSHGSTFGGNPLASRVAIAALEVIRDEKLCQRAAQLGSSFIAQLKALQAKSNGIISEVRGMG
317 GEHGSTYGGNPLGCRIAIAALEVLEEEHLAENADKMGAILRKELMKLPSDVVTA---VRGKG
    *  *.**.***.**  *.  * * *.* * * *  **  *   * *.  *  . *  *** *

353 AMIAIELFEDGDHNKPDAKLTAEIVARARDKGLILLSCGPYYNVLRILVPLTIEDAQIRQGL
358 LLTAI----VIDPSKANGKTAWDLCLLMKDQG--LLAKPTHDHIIRLAPPLVISEEDLQTGV
376 LLNAI----VIRETKDCD--AWKVCLRLRDNG--LLAKPTHGDIIRLAPPLVIKEDEIRESV
     . **      * *   * .  .  * **.*  **. ..   ..*.  ** * . .** *.

415 EIISQCFDEAKQ
414 ETIAKCIDLL
430 EIINKTILSF
    ***  * *
```

FIG. 3. Amino acid sequence comparison of the GABA transaminase from *E. coli* (ECGSST), the SCOAT, and the OAT from rat mitochondria (mROAT). Invariant amino acids are shown in boldface type. Identical residues in ECGSST and at least one of the two OAT are marked by asterisks. Conservative substitutions (2) are marked by dots.

genes, *gabT*, *gabD*, and *gabP*, as well as the control gene *gabC*, which coordinately regulates their expression (Fig. 5). In the present paper, we have determined the nucleotide sequence of *gabT* encoding GABA transaminase and characterized the *gabD* gene encoding SSDH. Both genes are located on a 3.8-kb *Sal*I-*Bam*HI fragment which is part of an initially cloned 15-kb *Sal*I fragment. In a previous study (3), we aligned the 15-kb *Sal*I fragment to the physical map of the *E. coli* K-12 genome constructed by Kohara et al. (15) at positions ca. 2800 to 2815. The *Sal*I-*Bam*HI fragment matches the physical map at positions 2802 to 2806. The restriction pattern of the map of Kohara et al. predicts a *Bgl*I site close to the *Bam*HI end of the *gabT* gene. This restriction site was identified in the *gabT* sequence (Fig. 1) and comprises nucleotides 1184 to 1194 (Fig. 2).

```
              268          304     398
               *            *       *
ECGSST     TFAKSIA      PIACVAA     LRILV-PLTIE
ECHPAT     TLSKAFA      PVADIAA     LRITV-G-TRE
SCHPAT     TLSKSFG      LASEYAL     LRITV-G-THE
ECTAT      SFSKIFS      PPNFGAQ      RMCVAGLNTA
ECAAT      SYSKNFG      PPAHGAS      RVNVAGMTPD
cCAAT      SFSKNFG      PPSQGAR      RINMCGLTTK
               *            *       *
              258          303     386
SCOAT      LLGKALS      PLASRVA      RLAP-PLVIS
mROAT      LLGKALS      PLGCRIA      RLAP-PLVIK
```

FIG. 4. Comparison of conserved regions in the amino acid sequences of the GABA transaminase from *E. coli* (ECGSST) and several other aminotransferases: ECHPAT, histidinol-phosphate transaminase from *E. coli*; SCHPAT, histidinol-phosphate transaminase from *S. cerevisiae*; ECTAT, aromatic aminotransferase from *E. coli*; ECAAT, AAT from *E. coli*; cCAAT, cytosolic chicken AAT; SCOAT; mROAT, OAT from rat mitochondria. Identical amino acids in ECGSST and other transaminases are shown in boldface type. The residues with a structural or functional role are marked by asterisks. The positions of these residues in the amino acid sequences of ECGSST and cytosolic chicken AAT are marked by numbers.

In a recent publication, Metzer and Halpern (21) also reported the cloning of the *E. coli* K-12 *gab* region. They localized *gabT* and *gabD* on a 4-kb *Sal*I-*Bam*HI fragment which is most probably identical to the 3.8-kb *Sal*I-*Bam*HI fragment described in this work, even though the restriction patterns of the two fragments do not match precisely. We determined the position and orientation of *gabT* and *gabD* on the 3.8-kb *Sal*I-*Bam*HI fragment. In agreement with the genetic map, *gabD* was found to be located adjacent to the 5′ end of *gabT*, and both genes are transcribed in the *Sal*I→ *Bam*HI direction (Fig. 5).

From early experiments with *E. coli gab* cluster regulatory mutants, it was concluded that the *gab* genes do not constitute a single operon (9). This hypothesis was supported recently by expression studies with fragments of the cloned *gab* region (21). It was reported that the *gab* genes were divergently transcribed from two promoters located in the vicinity of the unique *Bam*HI site, whereas the 4-kb *Sal*I-*Bam*HI fragment exhibited no promoter activity in either orientation.

In contrast to these findings, we provide evidence that the corresponding *Sal*I-*Bam*HI fragment described in this paper (Fig. 5) contains an endogenous promoter. In a previous communication (3), we demonstrated that this promoter is most probably the primary *gabC*-regulated promoter, since it is under the dual control of catabolite repression and nitrogen availability. Gene expression from this promoter is

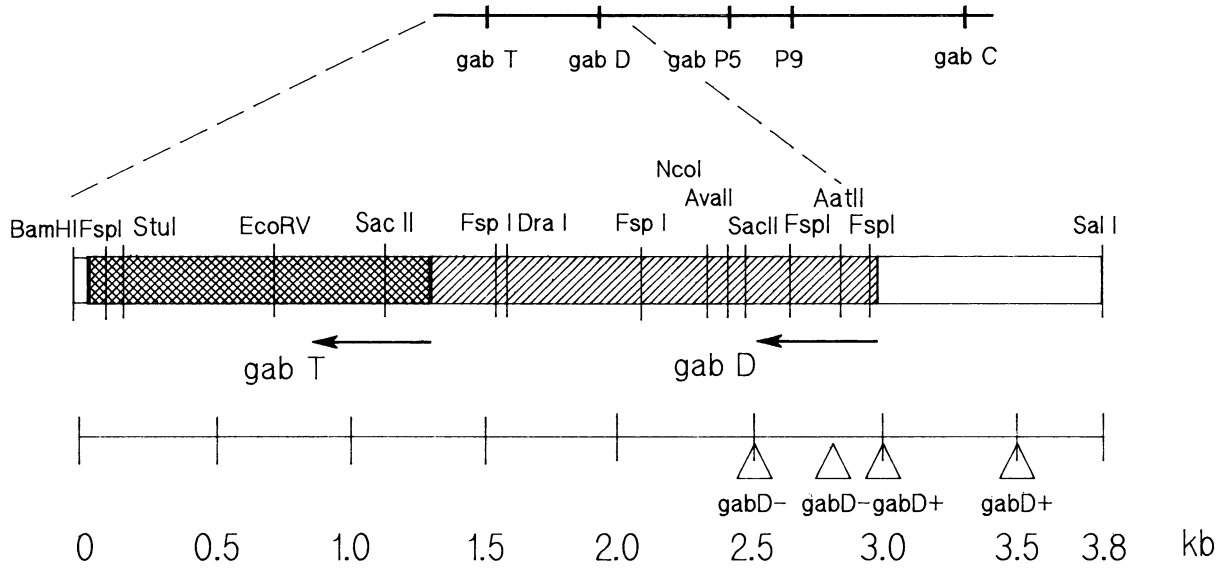

FIG. 5. Genetic map of the *E. coli* K-12 *gab* cluster at 57.5 min (22) and restriction map of the 3.8-kb *Sal*I-*Bam*HI fragment carrying the *gabT* and *gabD* genes. The locations of *gabT* and of the fragment containing the complete *gabD* structural gene are indicated. The direction of transcription of the two genes is marked by arrows. △, Positions of deletions (from the *Sal*I end of the fragment) used for mapping of the *gabD* gene. + and − indicate the ability of the fragment in question to confer elevated SSDH activity.

repressed by glucose but can be restored specifically when GABA is added as the sole nitrogen source. In the present study, we localized this promoter 5′ to the *gabD* gene and showed that it drives the transcription of both *gabD* and *gabT*.

Given the orientation of the two genes within the *gab* cluster, as deduced from Fig. 5, the two divergently transcribing promoters defined by Metzer and Halpern (21) are located at the 3′ end of *gabT* and can therefore not be responsible for the expression of *gabT* and *gabD*. However, it is possible that the promoter transcribing in the *Bam*HI⟶*Hin*dIII direction controls the expression of *gabP* which, in contrast to previous results (22) (Fig. 5), seems to be located downstream of *gabT* (21).

The maximal size of the *gabD* structural gene was estimated to be 1.7 kb, sufficient to encode a protein with a maximal molecular mass of ca. 60 kDa. The expression studies with transformants of *E. coli* carrying the *gabD* gene revealed that SSDH is specific for NADPH as a cofactor (data not shown). Two forms of bacterial SSDH with different cofactor requirements have previously been isolated from *Klebsiella pneumoniae* and characterized (24). One enzyme was specific for NADPH, whereas the other could use either NADH or NADPH. Their apparent molecular

TABLE 2. Expression of *gabT* and *gabD* in *E. coli* DH-1[a]

Plasmid	Insertion	Relative activity[b] of:	
		GSST	SSDH
pGAB1	3.8-kb *Sal*I-*Bam*HI	2,900	3,100
$pGAB1_R$	3.8-kb *Bam*HI-*Sal*I	350	300
pGAB2	3.5-kb *Sal*I-*Bam*HI (0.3 kb)[c]	4,100	2,400
pGAB3	3.0-kb *Sal*I-*Bam*HI (0.8 kb)[c]	3,900	4,600
pGAB4	2.8-kb *Sal*I-*Bam*HI (1.0 kb)[c]	3,200	70
pGAB5	2.5-kb *Sal*I-*Bam*HI (1.3 kb)[c]	3,800	80
pGAB6	2.9-kb *Aat*II-*Bam*HI	3,000	110
pGAB7	2.2-kb *Sal*I-*Dra*I	90	80
pMLC13		100[d]	100[d]

[a] Expression was measured as the GABA transaminase (GSST) and SSDH activities of transformants carrying the 3.8-kb *Bam*HI-*Sal*I fragment or parts of it in the plasmid vectors pMLC12 and pMLC13.

[b] Enzyme activities were determined in crude protein extracts of transformants carrying either a recombinant plasmid or the plasmid vector as a control.

[c] Numbers in parentheses indicate deletions of nucleotides in kilobases from the *Sal*I end of the fragment.

[d] 100% relative activity corresponds to specific activities of 3.6 nkat/mg of protein (GSST) and 5.0 nkat/mg of protein (SSDH).

TABLE 3. Expression in transformants of *E. coli* JM103 of *gabT* and *gabD* under the control of the endogenous promoter or the *tac* promoter

Plasmid	Insertion	IPTG induction[a]	Relative activity[b] of:	
			GSST	SSDH
$pJF1_{tac}$-	3.8-kb *Sal*I-*Bam*HI	−	650	500
		+	500	320
$pJF2_{tac}$-	1.6-kb *Dra*I-*Bam*HI	−	100	80
		+	120	95
pJF_{tac}-		−	100	100
		+	110	70
pJF1	3.8-kb *Sal*I-*Bam*HI	−	1,180	950
		+	5,610	4,600
pJF2	1.6-kb *Dra*I-*Bam*HI	−	900	120
		+	3,600	90
pJF118u		−	105	65
		+	100	70

[a] Cultures were induced by the addition of 1 mM IPTG at an optical density at 650 nm of 0.5 and harvested 4 h after induction.

[b] Enzyme activities were determined in crude protein extracts of transformants carrying either a recombinant plasmid or the plasmid vector as a control.

masses were determined to be 300 and 275 kDa, respectively. With respect to the estimated maximal molecular mass of the *gabD* gene product, only 60 kDa, it is likely that the native SSDH from *E. coli* is an oligomeric protein.

The nucleotide sequence of *gabT* encoding GABA transaminase revealed a structural gene of 1,281 nucleotides specifying a protein of 426 amino acids with a molecular mass of 45.76 kDa. Even though GABA transaminases play an important role in the regulation of the nervous system and the enzymes have been purified from several sources (4, 14, 30, 31), the first complete nucleotide sequence of a GABA transaminase became available only when the present paper was in preparation (1). The amino acid sequences of more than 20 transaminases are known; among them are 12 AAT (EC 2.6.1.1) (19). The three-dimensional structural data for several AAT and mutated forms of the enzymes have been reported (for a review, see reference 7), and several amino acids with a specific structural or functional role have been identified (19). A high degree of homology is conserved between all AAT and *E. coli* tyrosine aminotransferase (EC 2.6.1.5) (10). On the basis of the knowledge of the structurally and functionally important amino acid residues, two histidinol-phosphate transaminases (EC 2.6.1.9) could also be included in this homology group, whereas there is no similarity to all other known aminotransferases (19).

Our sequence comparison studies revealed a significant homology between the *E. coli* GABA transaminase and SCOAT and OAT from rat and human mitochondria. These findings are in accordance with the data for *S. cerevisiae* GABA transaminase obtained by André and Jauniaux (1). In this context, it is interesting that ornithine is accepted as a substrate by several GABA transaminases (26, 30, 31). Thus, the structural relatedness of these aminotransferases may reflect functional similarities of the proteins. Both enzymes are able to catalyze the transfer of ω-amino groups, whereas all other sequenced aminotransferases are specific for the transamination of α-amino acids.

To identify amino acid residues which might be involved in the active-site formation of the GABA transaminase protein, we compared conserved sequence segments containing important invariant residues from several aspartate, tyrosine, and histidinol-phosphate transaminases (19) with the corresponding segments of the GABA transaminase sequence. By this approach three homologous sequence stretches were identified (Fig. 4). The sequence of the Lys-268 fragment from GABA transaminase (Thr-Phe-Ala-Lys-Ser-Ile-Ala) correlates well with the consensus sequence of the pyridoxalphosphate-binding peptide of AAT and histidinol-phosphate transaminases (Ser [or Thr]-X-X-Lys followed by a cluster of nonpolar amino acids), which was given by Hsu et al. (13). de Biase et al. (4) have localized the pyridoxalphosphate-binding site of human GABA transaminase by affinity labeling on a tryptic peptide of 21 amino acids and found the sequence Thr-Phe-Ser-Lys, which is highly homologous to the corresponding sequence of the *E. coli* enzyme. However, for the remainder of the tryptic peptides, no homology to the GABA transaminase from *E. coli* was detectable. It is striking that the Lys-268 fragment and the Arg-398 fragment of GABA transaminase share the highest homologies not with the ornithine aminotransferases but with the histidinol-phosphate transaminases (Fig. 4). The sequence segment Leu-Arg-Ile-Leu-Val of GABA transaminase containing the Arg-398 residue is almost identical to the homologous peptides of the histidinol-phosphate transaminases from *E. coli* and *S. cerevisiae*. In contrast, the homology of the corresponding regions from the aspartate and tyrosine transaminases and OAT is essentially limited to the invariant Arg residue itself.

Even though conserved stretches of structurally or catalytically important amino acid residues can be found to various degrees in a number of transaminase sequences, homology of all aminotransferases is still not conclusively evident (19). The structural relationship between α- and ω-transaminases remained unclear up to now, since only one complete amino acid sequence from the latter group became recently available (1). The results presented in our work suggest that the GABA transaminase from *E. coli* shares several conserved regions with the histidinol-phosphate, aspartate, and tyrosine transaminases. Moreover, the folding pattern of the large domain of ω-amino acid:pyruvate aminotransferase from *Pseudomonas* sp. is very similar to that of aspartate transaminase (29). These findings may in fact promote the view that all aminotransferases constitute a family of proteins. However, the high degree of homology between GABA transaminase and OAT and the close relationship between aspartate and tyrosine transaminases (10) as well as between branched-chain amino acid and D-amino acid transaminases (27) may indicate the existence of distinct aminotransferase subfamilies. Further sequence information on other ω-transaminases will be necessary to clarify their evolutionary position within the aminotransferase proteins.

ACKNOWLEDGMENTS

We thank E. Ulhmann (Hoechst AG) for the synthesis of oligonucleotides and P. Taggeselle (Hoechst AG) for excellent technical assistance.

LITERATURE CITED

1. **André, B., and J. C. Jauniaux.** 1990. Nucleotide sequence of the yeast *UGA1* gene encoding GABA transaminase. Nucleic Acids Res. **18**:3049.
2. **Barcer, W. C., and M. O. Dayhoff.** 1972. Detecting distant relationships: computer methods and results, p. 101–110. *In* M. O. Dayhoff (ed.), Atlas of protein sequence and structure, vol. 5. National Biomedical Research Foundation, Washington, D.C.
3. **Bartsch, K., R. Dichmann, P. Schmitt, E. Uhlmann, and A. Schulz.** 1990. Stereospecific production of the herbicide phosphinothricin (glufosinate) by transamination: cloning, characterization, and overexpression of the gene encoding a phosphinothricin-specific transaminase from *Escherichia coli*. Appl. Environ. Microbiol. **56**:7–12.
4. **de Biase, D., J. B. Bolton, D. Barra, F. Bossa, and R. A. John.** 1989. Stoichiometry and stability of the adduct formed between human 4-aminobutyrate aminotransferase and 4-aminohex-5-enoate: sequence of a labelled peptide. Biochimie **71**:491–495.
5. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
6. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.
7. **Danishefsky, A., D. L. Smith, M. D. Toney, J. F. Kirsch, W. L. Finlayson, S. C. Almo, and D. Ringe.** 1989. Use of mutagenesis to probe the structure and function of AATase. Br. Crop Prot. Counc. Monogr. **42**:43–56.
8. **Dover, S., and Y. S. Halpern.** 1972. Control of the pathway of γ-amino-butyrate breakdown in *Escherichia coli* K-12. J. Bacteriol. **110**:165–170.
9. **Dover, S., and Y. S. Halpern.** 1974. Genetic analysis of the γ-aminobutyrate utilization pathway in *Escherichia coli* K-12. J. Bacteriol. **117**:494–501.
10. **Fotheringham, I. G., S. A. Dacey, P. P. Taylor, T. J. Smith, M. G. Hunter, M. E. Finlay, S. B. Primrose, D. M. Parker, and R. M. Edwards.** 1986. The cloning and sequence analysis of the

aspC and tyrB genes from Escherichia coli K-12—comparison of the primary structures of the aspartate aminotransferase and aromatic aminotransferase of E. coli with those of the pig aspartate aminotransferase isoenzymes. Biochem. J. **234**:593–604.

11. **Fürste, J. P., W. Pansegrau, R. Frank, H. Blöcker, P. Scholz, M. Bagdasarian, and E. Lanka.** 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tacP* expression vector. Gene **48**:119–131.
12. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**:351–359.
13. **Hsu, L. C., M. Okamoto, and E. E. Snell.** 1989. L-Histidinol phosphate aminotransferase from *Salmonella typhimurium*. Kinetic behavior and sequence at the pyridoxal-P binding site. Biochimie **71**:477–489.
14. **Kim, D. S., and J. E. Churchich.** 1987. Active site modification of 4-aminobutyrate aminotransferase with ATP analogs. Biochim. Biophys. Acta **916**:265–270.
15. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495–508.
16. **Levinson, A., D. Silver, and B. Seed.** 1984. Minimal size plasmids containing an M13 origin for production of single-strand transducing particles. J. Mol. Appl. Genet. **2**:507–517.
17. **Low, B.** 1968. Formation of meroploids in matings with a class of Rec$^-$ recipient strains of *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA **60**:160–167.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Mehta, P. K., T. I. Hale, and P. Christen.** 1989. Evolutionary relationships among aminotransferases. Tyrosine aminotransferase, histidinol-phosphate aminotransferase, and aspartate aminotransferase are homologous proteins. Eur. J. Biochem. **186**:249–253.
20. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–313.
21. **Metzer, E., and Y. Halpern.** 1990. In vivo cloning and characterization of the *gabCTDP* gene cluster of *Escherichia coli* K-12. J. Bacteriol. **172**:3250–3256.
22. **Metzer, E., R. Levitz, and Y. S. Halpern.** 1979. Isolation and properties of *Escherichia coli* K-12 mutants impaired in the utilization of γ-aminobutyrate. J. Bacteriol. **137**:1111–1118.
23. **Mueckler, M. M., and H. C. Pitot.** 1985. Sequence of the precursor to rat ornithine aminotransferase deduced from a cDNA clone. J. Biol. Chem. **260**:12993–12997.
24. **Sanchez, M., M. M. Fernández, A. Gibello, and A. Garrido-Pertierra.** 1989. Purification and properties of two succinic semialdehyde dehydrogenases from *Klebsiella pneumoniae*. Biochim. Biophys. Acta **990**:225–231.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
26. **Schulz, A., P. Taggeselle, D. Tripier, and K. Bartsch.** 1990. Stereospecific production of the herbicide phosphinothricin (glufosinate) by transamination: isolation and characterization of a phosphinothricin-specific transaminase from *Escherichia coli*. Appl. Environ. Microbiol. **56**:1–6.
27. **Tanizawa, K., S. Asano, Y. Masu, S. Kuramitsu, H. Kagamiyama, H. Tanaka, and K. Soda.** 1989. The primary structure of thermostable D-amino acid aminotransferase from a thermophilic *Bacillus* species and its correlation with L-amino acid aminotransferases. J. Biol. Chem. **264**:2450–2454.
28. **Urdea, M. S., J. P. Merryweather, G. T. Mullenbach, D. Coid, U. Heberlein, P. Valenzuela, and J. Barr.** 1983. Chemical synthesis of a gene for human epidermal growth factor urogastrone and its expression in yeast. Proc. Natl. Acad. Sci. USA **80**:7461–7465.
29. **Watanabe, N., K. Sakabe, N. Sakabe, T. Higashi, K. Sasaki, S. Aibara, Y. Morita, K. Yonaha, S. Toyama, and H. Fukutani.** 1989. Crystal structure analysis of ω-amino acid:pyruvate aminotransferase with a newly developed Weissenberg camera and an imaging plate using synchrotron radiation. J. Biochem. **105**:1–3.
30. **Yonaha, K., K. Suzuki, and S. Toyama.** 1985. 4-Aminobutyrate: 2-oxoglutarate aminotransferase of *Streptomyces griseus*: purification and properties. Eur. J. Biochem. **146**:101–106.
31. **Yonaha, K., and S. Toyama.** 1980. γ-Aminobutyrate:α-ketoglutarate aminotransferase from *Pseudomonas sp.* F-126: purification, crystallization, and enzymologic properties. Arch. Biochem. Biophys. **200**:156–164.