JOURNAL OF BACTERIOLOGY, Sept. 1992, p. 5820–5826
0021-9193/92/185820-07$02.00/0

Vol. 174, No. 18

Escherichia coli Has Two Homologous Glutamate Decarboxylase Genes That Map to Distinct Loci

DEAN K. SMITH,[1] TASLEEM KASSAM,[2] BHAGIRATH SINGH,[1] AND JOHN F. ELLIOTT[1,2]*

Departments of Medical Microbiology and Infectious Diseases[2] and Immunology,[1] University of Alberta, Edmonton, Alberta, T6G 2H7 Canada*

Received 5 May 1992/Accepted 9 July 1992

Degenerate oligonucleotides based on the published *Escherichia coli* glutamate decarboxylase (GAD) protein sequence were used in a polymerase chain reaction to generate a DNA probe for the *E. coli* GAD structural gene. Southern blots showed that there were two cross-hybridizing GAD genes, and both of these were cloned and sequenced. The two GAD structural genes, designated *gadA* and *gadB*, were found to be 98% similar at the nucleotide level. Each gene encoded a 466-residue polypeptide, named, respectively, GAD α and GAD β, and these differed by only five amino acids. Both GAD α and GAD β contain amino acid residues which are highly conserved among pyridoxal-dependent decarboxylases, but otherwise the protein sequences were not homologous to any other known proteins. By restriction mapping and hybridization to the Kohara miniset library, the two GAD genes were located on the *E. coli* chromosome. *gadA* maps at 4046 kb and *gadB* at 1588 kb. Neither of these positions is in agreement with the current map position for *gadS* as determined by genetic means. Analysis of Southern blots indicated that two GAD genes were present in all *E. coli* strains examined, including representatives from the ECOR collection. However, no significant cross-hybridizing gene was found in *Salmonella* species. Information about the DNA sequences and map positions of *gadA* and *gadB* should facilitate a genetic approach to elucidate the role of GAD in *E. coli* metabolism.

The enzyme glutamate decarboxylase (GAD; also known as glutamic acid decarboxylase; EC 4.1.1.15) catalyzes the α-decaboxylation of glutamic acid to produce γ-aminobutyric acid. Within bacteria, GAD activity seems to be relatively unique to *Escherichia coli* (37). Gale (14) proposed a general role for the inducible bacterial amino acid decarboxylases, including GAD, in the maintenance of physiological pH under acidic conditions. *E. coli* GAD has been extensively characterized with respect to its biophysical and biochemical properties (1, 14, 30, 31, 39–46), and a partial protein sequence has been available for some time (42, 43). Based on genetic linkage studies in *E. coli*, the structural gene for GAD (*gadS*) and a potential regulatory gene (*gadR*) have been mapped between *mtl* at approximately 80.7 minutes and *gltS* at approximately 82.4 minutes (3, 23, 26, 27), but these early results have not been followed up. For the purposes of our immunological studies, we were interested in obtaining the complete DNA and protein sequences of *E. coli* GAD. At the time this work was initiated, an extensive search of the literature and all available sequence data bases suggested that neither the complete protein sequence nor the gene sequence for *E. coli* GAD had been determined. To obtain this information, we generated a DNA probe for the gene based on a recent more extensive partial protein sequence (24). By using this probe on *E. coli* genomic DNA, we discovered that there were not one but two separate cross-hybridizing GAD genes. We report here the complete DNA sequences and map positions for these two genes, which encode two distinct *E. coli* GAD polypeptides. The genes are highly similar, yet they map to separate loci, neither of which correspond to the previously reported map position for *gadS*. Representatives from a diverse assemblage of *E. coli* strains were examined, and all were found to contain two GAD genes. The conservation of two homologous coding segments among diverse strains of *E. coli* suggests that both gene products play a role in *E. coli* metabolism. Although the complete protein sequence of one isoform of *E. coli* GAD was reported while this manuscript was in preparation (25), this is the first account of the dimorphic nature of the GAD genes in *E. coli* and the first time the complete DNA sequence of either isoform has been reported.

MATERIALS AND METHODS

Culture media, bacterial strains, and bacteriophage libraries. All bacterial cultures were grown in 2× YT (36) at 37°C with vigorous aeration. Recombinant plasmids were propagated in standard *E. coli* DH5α, WM1100, or TG2 as required. The Kohara λEMBL4 miniset library (version 9010) was obtained from the Japanese National Institute of Genetics (19) and plated on *E. coli* NM 621. All other strains used are described in Table 2, with the exception of *Salmonella typhimurium* LT2 (50).

General procedures. To obtain *E. coli* DNA, we lysed cells by resuspending the drained bacterial pellets in 100 mM Tris-HCl–25 mM EDTA–1% sodium dodecyl sulfate (SDS) and incubating them for 15 min at room temperature. The lysate was extracted with phenol, phenol-chloroform, and chloroform, and the DNA was precipitated with isopropanol, rinsed in 70% ethanol, and resuspended in TE (10 mM Tris-HCl [pH 8.0], 1 mM EDTA). To obtain λ phage DNA, we purified the intact phage particles on CsCl step gradients and extracted the DNA as described in reference 36. To generate the GAD gene probe, polymerase chain reaction (PCR) conditions were 94, 55, and 72°C, each for 1 min, in a total reaction volume of 100 μl. Buffer, primer, nucleotide, and *Taq* DNA polymerase concentrations were as recommended by the supplier (Perkin-Elmer Cetus, Norwalk, Conn.). For each PCR, 5 ng of purified *E. coli* DNA was amplified through 30 cycles, and the products were analyzed

* Corresponding author.

directly on a low-melting-point (LMP) 1.2% agarose gel. The sequences of the oligonucleotide primers used in the PCR were 5′ GGNATGGCNATGAAATGG 3′, 5′ GGNATGGC NATGAA(A/G)TGG 3′, 5′ CCAGATNACCCANCCGCA 3′, and 5′ CCAGATNACCCANCC(A/G)CA 3′. For hybridization probes, DNA fragments were excised from LMP agarose gels, boiled, and labeled with [α-^{32}P]dCTP by the random hexamer primer method (12). Southern blots were transferred to Hybond-N (Amersham Corp., Arlington Heights, Ill.) and probed and washed as described previously (10). Southern blots of *S. typhimurium* DNA were probed at a lower stringency (5× SSC [1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate], 0.1% SDS, 5× Denhardt's solution at 37°C) and washed in 5× SSC–0.1% SDS at 56°C. DNA sequencing was done on double-stranded and single-stranded templates with the Sequence Kit (United States Biochemical, Cleveland, Ohio) and a series of synthetic oligonucleotide primers. Searches for homologous sequences were run for both DNA and encoded protein against the GenBank, Unique EMBL, and HIV DNA or the PIR, SWISS-PROT, and HIV protein data bases (Release 92.1.1; Intelligenetics, Mountain View, Calif.). The searches were conducted by using FastDB from the Intelligenetics Suite (Release 5.4) on a Sun SPARC Station 2 (Sun Microsystems Inc., Mountain View, Calif.).

Construction and screening of plasmid libraries. The cloning vector pHAS was constructed by modifying the vector pBluescript KS− (Stratagene, La Jolla, Calif.) in the following way: (i) digestion of KS− with *Bst*XI, blunting with T4 DNA polymerase, and religation to delete the *Bst*XI and *Sac*II sites present in the original polylinker; (ii) digestion of the resultant plasmid with *Eco*RV and ligation to the annealed oligonucleotides (5′ pCCACGTGTTTGGTGTG 3′-5′ pCCAAACACGTGG 3′); and (iii) ligation of the resultant linear plasmid to the ≈360-bp *Bst*XI fragment from the vector pCDM8 (38) to create a circular molecule, pHAS. For construction of libraries, pHAS DNA was digested with *Bst*XI, the ≈380-bp DNA stuffer fragment was removed by running the cut vector DNA twice on an LMP 0.8% agarose gel, and the DNA was purified from the final gel sliced by using GeneClean (Bio 101, La Jolla, Calif.).

Two separate *Eco*RI libraries were constructed by digesting 10 μg of *E. coli* DNA with *Eco*RI, size fractionating the digested material on an LMP 0.7% agarose gel, excising gel slices containing fragments of 4.4 to 6.6 kb and 9.4 to 23 kb, and extracting the DNA from each slice with GeneClean. The DNA fragments were ligated in separate reactions to a 50-fold molar excess of the oligonucleotides (5′ pCT CTAGGG 3′-5′ pAATTCCCTAGAGACAC 3′), purified again by electrophoresis on LMP agarose gels and GeneClean, and ligated into pHAS prepared as described above. A single *Hin*fI library was constructed in the same manner by digesting *E. coli* DNA with *Hin*fI, size fractionating the digested material on an LMP 1.0% agarose gel, excising a single gel slice containing fragments of 1.2 to 2.3 kb, and ligating to a 50-fold molar excess of the oligonucleotides (5′ pCTCTAGGG 3′-5′ pANTCCCTAGAGAC AC 3′, where N represents equal portions of G, A, T, and C). The final ligations of *E. coli* DNA into pHAS were precipitated with sodium acetate-ethanol, using 10 μg of yeast tRNA (GIBCO BRL, Gaithersburg, Md.) as carrier, transformed into *E. coli* WM1100 by electroporation (9), and plated on ampicillin plates. Colonies were transferred to nitrocellulose filters (Schliecher & Schuell, Inc., Keene, N.H.), and the filters were baked and prewashed (36) and then hybridized and washed under the same conditions as for Southern blots.

Plating and probing the Kohara miniset library. A 100-μl portion of 2× YT–10% glycerol was added to each well containing the ordered set of clones (in 96-well plates), and phage particles were allowed to elute overnight at 4°C. Fresh lawns of *E. coli* NM 621 were poured onto 10-cm plates, and after 3 h of growth, a 48-pin transfer apparatus was used to transfer an ≈2-μl drop of phage particles from each well onto the bacterial lawns. These were grown until an array of large plaques had formed with the same geometry and corresponding to the original set of clones in the 96-well plates. The plaques were transferred to nitrocellulose filters, and the filters were baked, hybridized, and washed as above.

Nucleotide sequence accession numbers. The DNA sequences reported here have been submitted to GenBank under accession numbers M84024 and M84025.

RESULTS

Four oligonucleotides were designed based on the partial protein sequence of *E. coli* GAD (24), using the regions of protein sequence underlined in Fig. 3. In four separate reactions, all possible combinations of 5′ and 3′ primers were amplified through 30 cycles of PCR. Four control reactions, each containing *E. coli* DNA plus a single primer, were done in parallel. All the PCRs which contained two primers generated a single predominant fragment of approximately 470 bp, whereas the control reactions did not generate DNA fragments of any length (data not shown). The four amplified products were excised from the gel, and the DNA sequence of each was determined on both strands by direct sequencing by using the PCR primers. Each of the DNA sequences encoded a single long open reading frame, which when translated corresponded to the known partial protein sequence of *E. coli* GAD (24). One of the DNA fragments was radiolabeled and used as a hybridization probe for the GAD gene.

We digested *E. coli* DNA from a K-12 strain (J53-1) with the restriction enzymes *Eco*RI, *Hin*dIII, *Hin*fI, *Pst*I, and *Sma*I and made a Southern blot. When the blot was probed with the GAD PCR gene fragment, we found two cross-hybridizing bands of different sizes but roughly equal intensities in all the digests except *Pst*I. The single band seen with the *Pst*I digest was approximately double the intensity of the bands produced in the other digests. Representative results for the *Eco*RI, *Hin*fI, and *Pst*I digests are shown in Fig. 1. With the exception of *Pst*I, we knew from the DNA sequence that our gene probe did not contain any of these restriction sites. These results indicated that there must be two cross-hybridizing genes in *E. coli*. Moreover, the double-intensity *Pst*I band suggested that this enzyme cleaves the same-sized fragment from each of the two genes. Although the probe does contain an internal *Pst*I site, it is so near one end that the PCR product apparently does not hybridize efficiently to any adjacent *Pst*I fragments under these stringency conditions. Both the 13.1- and 5.6-kb *Eco*RI fragments and the 2.2- and 1.65-kb *Hin*fI fragments shown in Fig. 1 were chosen for further analysis.

Clones *Eco*RI L (13.1-kb *Eco*RI fragment) and *Eco*RI S (5.6-kb *Eco*RI fragment) were isolated from separate libraries. A third library which contained both the longer and shorter cross-hybridizing *Hin*fI fragments was constructed and screened, and clones were isolated and determined to

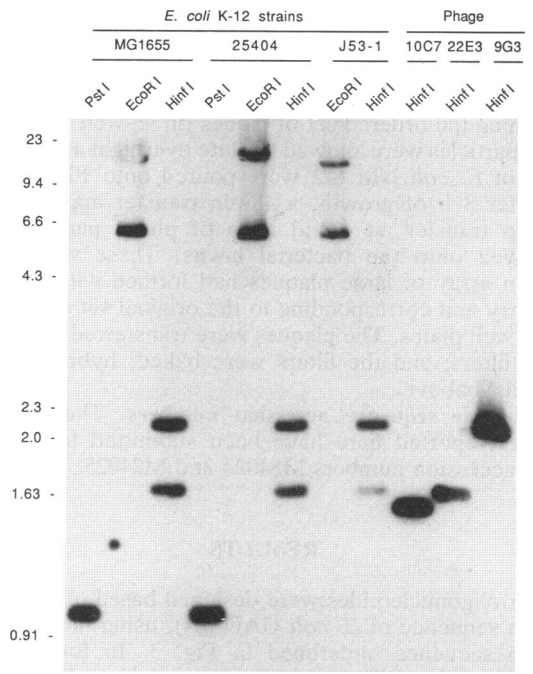

FIG. 1. Representative Southern blot made with *E. coli* DNA from three K-12 strains (left eight lanes) and with purified λ phage DNA from three of the clones in the miniset library (right three lanes), digested with enzymes as shown. The blot was probed with a ^{32}P-labeled 552-bp *Pst*I-*Eco*RI restriction fragment from *gadA*. A single band of approximately double intensity appears in the *Pst*I digests because the two genes contain identical *Pst*I sites (see Fig. 2 and 3). Numbers on left show size in kilobases.

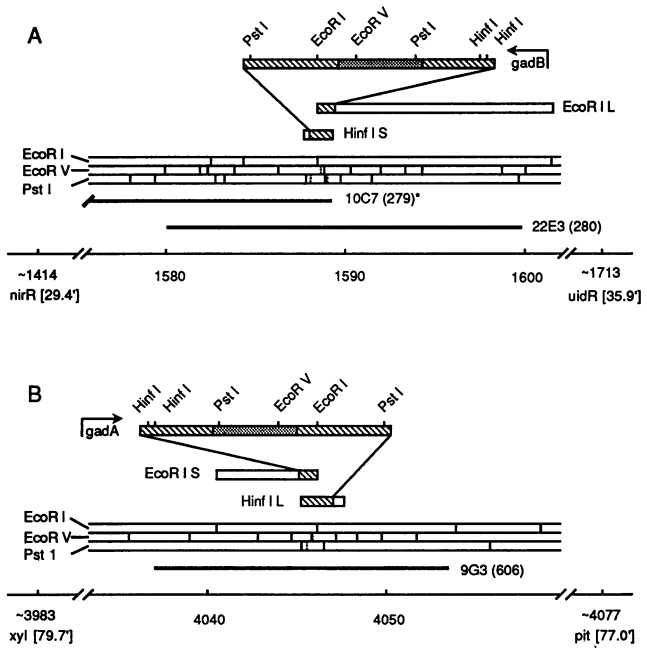

FIG. 2. Alignment of the GAD genomic clones with the *E. coli* physical map and corresponding miniset phage clones (19). Each pair of genomic clones (*Eco*RI L and *Hin*fI S or *Eco*RI S and *Hin*fI L) is shown in the center of panel A or B, and above each pair and connected by diagonal lines is an enlarged map of each *gad* open reading frame, with a number of restriction sites indicated. The arrow shows the direction of transcription. Hatched regions represent *gad* open reading frames, stippled regions overlaid on the open reading frames indicate the region encompassed by the original PCR probe, and the dotted lines within the redrawn physical map show the actual restriction sites seen in these genes. * indicates that the 10C7(279) phage clone has less overlap with 22E3(280) than is indicated in the original physical map figure (see Results).

contain either the 2.2-kb (*Hin*fI L) or 1.6-kb (*Hin*fI S) fragment on the basis of restriction mapping. Sequence data obtained by primer walking through the cloned genomic fragments revealed that although each contained an open reading frame, no single clone encoded a complete GAD gene. However, since the DNA sequence of the two GAD genes is similar but not identical (see below), by comparing sequences of the four clones, we were able determine that the *Eco*RI L clone should be matched with the *Hin*fI S clone and vice versa as shown in Fig. 2 (upper parts). By combining the DNA sequences from the matched pairs of clones, we were able to reconstruct the entire open reading frame for each gene.

The aligned DNA and encoded protein sequences of *gadA* and *gadB* are shown in Fig. 3. Each gene contains a long open reading frame which encodes a protein of 466 amino acids, with calculated molecular masses of 52,651 and 52,634 Da, respectively. A putative ribosome binding sequence is found immediately 5′ of the start codon (underlined in Fig. 3). Preliminary sequence data (not shown) suggest that there are potential promoter sequences (16) 5′ to each of the GAD genes, and these are under investigation. The two DNA coding sequences are 98% similar, and the deduced protein sequences show 99% similarity. The 5 amino acids which differ between the two proteins all occur within the N-terminal 22 residues, and some of these represent nonconservative substitutions (Fig. 3 and Table 1, upper part). Our deduced protein sequence for GAD α is in agreement with the published protein and deduced protein sequences for *E. coli* GAD (25), with the differences listed in the lower part of Table 1. The *E. coli* GAD gene sequences reported here were not obtained from PCR amplification, which can generate mutations (22), and for each GAD gene, we independently sequenced the upper and lower strands and found the two strands to be in complete agreement with each other. We suggest therefore that our DNA and deduced protein sequences are correct and that those differences which exist between our sequences and the protein and partial DNA sequences published by Maras et al. (25) are due either to polymorphisms between the different strains of *E. coli* studied or to errors in the previously published sequences, which in the case of the DNA data may have resulted from artifacts introduced during PCR amplification.

To map the positions of the two *E. coli* GAD genes, we radiolabeled the *Hin*fI L and *Hin*fI S DNA fragments and in separate reactions hybridized these to duplicate sets of filters containing DNA from the Kohara miniset library (19). With both probes, three of the λ EMBL 4 clones gave a strong hybridization signal [clones 10C7(279), 22E3(280), and 9G3(606)], with the stronger signal for the *Hin*fI L probe corresponding to clone 9G3 and the stronger signal for the *Hin*fI S probe corresponding to clones 10C7 and 22E3. At least five additional λ clones gave an intermediate cross-hybridizing signal that was clearly below that of the 9G3, 10C7, and 22E3 clones but above background (discussed below). Knowledge of the map positions of the strongly cross-hybridizing λ clones together with the lengths of the

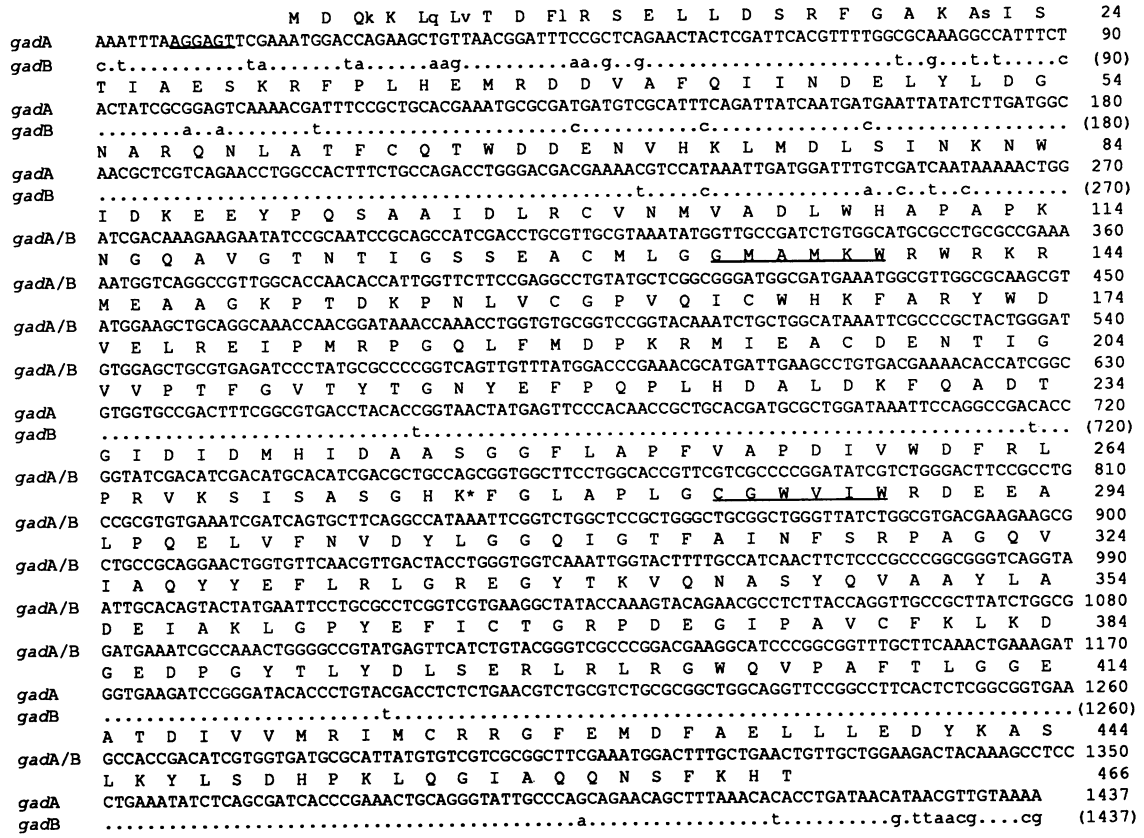

FIG. 3. Nucleotide and deduced amino acid sequences of *gadA* and *gadB*. All nucleotides and amino acids which are unique to *gadB* are shown in lowercase; otherwise, they are identical with those found in *gadA*. Regions of complete nucleotide identity are indicated by *gadA/B* at the start of the line. The putative Shine-Dalgarno sequence and the regions of protein sequence corresponding to the degenerate primers used in the PCR are underlined. * indicates the lysine residue which binds pyridoxal phosphate.

two *Eco*RI fragments enabled us to assign positions for *Eco*RI L and *Eco*RI S and to deduce the orientation of the open reading frames for the two GAD genes (Fig. 3). To confirm the assignments for the *Hin*fI L and *Hin*fI S fragments, we first purified λ DNA from the 10C7, 22E3, and 9G3 phages, digested it with *Hin*fI, made Southern blots, and then probed with the original GAD gene probe (Fig. 1, right lanes). These results indicate clearly that clone 9G3 contains the *Hin*fI L fragment and clone 22E3 contains the *Hin*fI S fragment. Clone 10C7 appears to contain a slightly truncated *Hin*fI S fragment, suggesting that it is missing a portion of the GAD gene, a result which is in keeping with our understanding of the map position of this λ clone relative to the position we assigned to *gadB* (Fig. 2). The miniset library was constructed from *Sau*3AI partial digests of *E. coli* W3110 DNA, and examination of the DNA sequence for *gadB* reveals several *Sau*3AI sites downstream of the two *Hin*fI sites (Fig. 2 and 3). Therefore, the insert in clone 10C7(279) was apparently generated by digestion of the *E. coli* chromosome at one of these *Sau*3AI sites, and the truncated *Hin*fI fragment results from cutting at the usual *Hin*fI site on the one side and a *Hin*fI site present in the λ vector on the other. With respect to the *E. coli* physical map, there is an inversion in the chromosome of *E. coli* W3110 in the region that encompasses GAD α (17). Thus, although the two GAD genes are shown to be transcribed in opposite directions in Fig. 2, this would not be the case in other *E. coli* K-12 strains.

All of our initial results for the GAD gene were based on *E. coli* K-12 strains. We were interested in knowing whether other strains of *E. coli* also had two GAD genes, or whether this was a property which was unique to the K-12 strains. The ECOR collection of *E. coli* strains provides a set of bacteria with diverse genetic backgrounds, originating from a wide range of hosts (33). We extracted DNAs from a number of representative members in this collection (29), digested each with a series of restriction enzymes, and subjected these fragments to Southern blot analysis using a GAD gene probe. The results of these experiments are summarized in Table 2. Although the size of the cross-hybridizing bands varied somewhat between strains, all *E. coli* strains examined contained two bands for most digests. It would appear, therefore, that most if not all strains of *E. coli* contain two GAD genes.

Since *S. typhimurium* is closely related to *E. coli*, we wondered whether we could detect a cross-hybridizing GAD gene in this species. Even at the lowest stringency of hybridization, we were unable to detect a clear signal on Southern blots made from *S. typhimurium* DNA using the *E. coli* gene as a probe (data not shown). Very long autoradiographic exposures revealed multiple faint cross-hybridizing bands, but these were more consistent with the less specific signals seen with long exposures of the *E. coli* Southern blots. If in fact *S. typhimurium* does make GAD, we suggest that the gene sequence is significantly different from that found in *E. coli*.

TABLE 1. Comparison of amino acids at substituted positions for GAD α, GAD β, and GAD* (25)[a]

Position	GAD α DNA-encoded sequence	GAD* protein sequence	GAD* DNA-encoded sequence	GAD β DNA-encoded sequence
3	Q	Q	NA[b]	K
5	L	L	NA	Q
6	L	L	NA	V
9	F	F	NA	L
22	A	A	NA	S
64	C	S	NA	C
73	H	R	NA	H
153	D	N	D	D
165	C	S	S	C
208	T	T	N	T
295	L	L	V	L
355	D	N	D	D

[a] Ambiguities exist between the protein and partial DNA-encoded sequences for GAD*, so both have been listed. The partial DNA sequence of GAD* spans the region from amino acids 148 to 431 of the 466-residue protein. Residues which are unique to GAD β are shown above; residues which differ between GAD α-GAD β and GAD* are shown below.
[b] NA, not available.

TABLE 2. Sizes of *Eco*RI, *Hin*fI, and *Pst*I restriction fragments from various strains of *E. coli* which hybridize to a *gadA* probe[a]

Strain (reference)	Restriction fragment(s) (kb)		
	*Eco*RI	*Hin*fI	*Pst*I
K-12			
J53-1 (4)	13.1, 5.6	2.2, 1.65	**1.02**
Wild type (4)	12.3, 5.5	2.2, 1.64	**0.96**
MG1655 (15)	13.2, 5.6	2.2, 1.65	**0.97**
25404 (ATCC)	15.5, 5.6	2.2, 1.66	**0.98**
ECOR[b]			
20 (33)	13.0, 5.5	2.1, 1.60	ND[c]
26 (33)	16.5, 5.6	2.1, 1.54	ND
29 (33)	16.5, 5.6	2.2, 1.65	ND
37 (33)	6.3, 5.6	**2.2**	ND
46 (33)	14.0, 5.7	2.2, 1.64	ND
Other			
W (7)	16.5, 5.9	2.2, 1.68	**1.02**
11246 (32)	15.5, 6.0	2.2, 1.67	1.20, 1.10

[a] The probe is described in the legend to Fig. 1. Double-intensity bands are shown in boldface. ATCC, American Type Culture Collection.
[b] ECOR strain hosts: 20, steer (Bali); 26, human infant (United States); 29, kangaroo rat (United States); 37, marmoset (U.S. zoo); 46, Celebese ape (U.S. zoo).
[c] ND, not determined.

DISCUSSION

To obtain a gene probe for *E. coli* GAD, we designed oligonucleotides based on two independent segments of protein sequence whose relative positions were known and then used PCR to generate the unknown gene sequence lying between them. This represents a powerful application of the PCR technology which is far superior to screening strategies based on degenerate oligonucleotides. To clone the *E. coli* GAD genes quickly and efficiently, we constructed a series of sublibraries by digesting *E. coli* DNA with *Eco*RI or with *Hin*fI and selecting fragments within a certain size range as described in the Materials and Methods. To facilitate the construction of the sublibraries, we created a new plasmid vector, pHAS. When pHAS DNA is digested to completion with *Bst*XI, a DNA fragment of approximately 380 bp is cleaved from the center of the polylinker, and the resulting ends of the vector are sticky but not self-cohesive (38). Ligation of foreign DNA into *Bst*XI-cut pHAS prepared as described in the Materials and Methods is highly efficient and results in very low numbers of nonrecombinant (i.e., empty) clones in the final libraries, typically <0.1%.

Since our deduced GAD protein sequences are virtually identical to the recently published protein sequence of Maras et al. (25), we are confident that the *gadA* and *gadB* sequences shown here do indeed encode *E. coli* GAD. Further, all three protein sequences are in agreement with the early partial protein sequence and total amino acid composition data (42, 43). Predicted molecular masses for both isoforms of GAD are also in agreement with estimates based on sedimentation rates for monomeric GAD protein (42). Protein sequence homology searches conducted by using the FastDB algorithm produced results similar to those obtained by Maras and colleagues (25). Our optimized searches identified a plant tryptophan decarboxylase (8) as the most similar protein in the data bases. However, when our *E. coli* GAD protein sequences were aligned with other protein sequences from the family of pyridoxal-dependent decarboxylases (aligning with respect to the active-site lysine which binds pyridoxal phosphate), the *E. coli* protein sequences contained many of the conserved residues which were originally noted by Jackson (18) for this group of enzymes. For the GAD α and GAD β proteins reported here, the region containing these conserved residues, which likely contributes to the active site, extends from the methionine at position 240 to the cysteine at position 284 (data not shown). Very little protein sequence similarity exists outside of this region when comparing the *E. coli* protein sequences with the known eucaryotic GADs, and this observation also extends to other members of the group of pyridoxal-dependent decarboxylases, both eucaryotic and procaryotic. It is interesting that in those eucaryotic organisms in which the GAD genes have been cloned, two related but non-cross-hybridizing genes have virtually always been found (6). Given the differences between the *E. coli* and eucaryotic GAD gene sequences and the distinct biochemical properties of the procaryotic and eucaryotic enzymes (11, 49), the fact that two separate GAD genes occur in such diverse species most likely represents an example of convergent evolution.

As noted in the Results, longer exposures of both the *E. coli* and *S. typhimurium* Southern blots allowed us to detect numerous additional weakly cross-hybridizing bands. These likely represent genes encoding other pyridoxal-dependent decarboxylases, since our probe included the highly conserved region bracketing the pyridoxal-binding lysine. Such an explanation would also account for the intermediate hybridization signals observed for some of the clones when the minset library was screened.

Our hybridization results with the miniset library, in combination with the restriction mapping data and Southern blot analysis, argue strongly that we have correctly assigned the positions of the two GAD structural genes (*gadA* and *gadB*) on the *E. coli* physical map (Fig. 2). The fact that neither GAD gene mapped to the 82-minute region, assigned to *gadRS* (3), suggests that this assignment should be revised, although it is possible that a regulatory gene (*gadR*) accounted for the earlier mapping results. It is also possible that a single-copy structural or regulatory gene responsible for the production of the pyridoxal phosphate cofactor could

map to the 82-minute position. A mutation at such a locus could result in a loss of GAD activity, as well as a loss of enzyme activity for all other pyridoxal-dependent enzymes. This type of mutation, however, would not necessarily affect the production of pyridoxal-dependent enzymes. GAD immunoreactive protein with little or no enzymatic activity was observed by Lupo and Halpern (23) in a Gad$^-$ *E. coli* strain, but this was interpreted to represent a lesion in the *gad* structural gene and considered supportive of the 82-minute map position with a single *gad* locus.

The similarity between the two GAD genes and the encoded proteins is quite striking. Other examples of duplication of structural genes in which the two homologs map to distinct loci in *E. coli* include *argI* and *argF* (5, 47) and *tufA* and *tufB* (2, 48). The *arg* genes have diverged to the point that the two enzymes are biochemically distinguishable (20, 21). In contrast, the *tuf* genes are nearly identical, and the only difference in the 339-residue structural proteins is a single substitution at the C terminus. The *tufA* and *tufB* gene products are essentially indistinguishable functionally and structurally (13, 28), but the TufA protein appears to be produced at higher levels (34, 35). Comparison of the *tufA* and *tufB* gene sequences shows that nucleotide differences tend to be clustered toward the 5′ and 3′ ends of the genes, with very few differences occurring in the central region. This same clustering of differences is also observed in *gadA* and *gadB* (Fig. 3). An and Friensen (2) suggested that in *tufA* and *tufB* this clustering of differences is consistent with the notion that recombination events are more likely to occur within the central portion of the genes. Alternatively, it can be argued that physiological constraints have prevented divergence of these homologous gene pairs. Since the gene products of the *tuf* loci were functionally indistinguishable, it was suggested that the duplication provided an emergency backup when demand for translation elongation factor EF-Tu was high (2). A similar argument may apply to the two *gad* loci, since the two *E. coli* GAD proteins are virtually identical. Availability of the cloned GAD genes together with accurate information about the locations of *gadA* and *gadB* on the *E. coli* chromosome should facilitate a genetic approach to elucidate the role of GAD in *E. coli* metabolism.

ACKNOWLEDGMENTS

This research was supported by research grants from the Alberta Heritage Foundation for Medical Research, the Canadian Diabetes Association, the Juvenile Diabetes Foundation International, and the Muttart Diabetes Research and Training Center. D.K.S. is supported by an Alberta Heritage Foundation studentship.

We thank W. L. Albritton, B. J. Bachmann, and D. E. Taylor for supplying *E. coli* strains and advice, Y. Kohara for supplying the miniset library, A. E. Hudson for proofreading the manuscript, and K. I. Jensen for moral support.

REFERENCES

1. **Almasov, V. P., Y. V. Morozov, F. A. Savin, and B. S. Sukhareva.** 1985. Model of enzymic decarboxylation of glutamic acid. Mol. Biol. (Moscow) **19**:359–370.
2. **An, G., and J. D. Friesen.** 1980. The nucleotide sequence of *tuf*B and four nearby tRNA structural genes of *Escherichia coli*. Gene **12**:33–39.
3. **Bachmann, B. J.** 1990. Linkage map of *Escherichia coli* K-12, edition 8. Microbiol. Rev. **54**:130–197.
4. **Bachmann, B. J.** 1972. Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol. Rev. **36**:525–557.
5. **Bencini, D. A., J. E. Houghton, T. A. Hoover, K. F. Foltermann, J. R. Wild, and G. A. O'Donovan.** 1983. The DNA sequence of *arg*I from *Escherichia coli* K-12. Nucleic Acids Res. **11**:8509–8519.
6. **Bu, D.-F., M. G. Erlander, B. C. Hitz, N. J. K. Tillakaratne, D. L. Kaufman, C. B. Wagner-McPherson, G. A. Evans, and A. J. Tobin.** 1992. Two human glutamate decarboxylases, 65-kDa GAD and 67-kDa GAD, are each encoded by a single gene. Proc. Natl. Acad. Sci. USA **89**:2115–2119.
7. **Davis, B. D., and E. S. Mingioli.** 1950. Mutants of *Escherichia coli* requiring methionine or vitamin B_{12}. J. Bacteriol. **60**:17–28.
8. **De Luca, V., C. Marineau, and N. Brisson.** 1989. Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: comparison with animal dopa decarboxylases. Proc. Natl. Acad. Sci. USA **86**:2582–2586.
9. **Dower, W. J., J. F. Miller, and C. W. Ragsdale.** 1988. High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. **16**:6127–6145.
10. **Elliott, J. F., G. R. Albrecht, A. Gilladoga, S. M. Handunnetti, J. Neequayes, G. Lallinger, J. N. Minjas, and R. J. Howard.** 1990. Genes for *Plasmodium falciparum* surface antigens cloned by expression in COS cells. Proc. Natl. Acad. Sci. USA **87**:6363–6367.
11. **Erlander, M. G., and A. J. Tobin.** 1991. The structure and functional heterogeneity of glutamic acid decarboxylase: a review. Neurochem. Res. **16**:215–226.
12. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
13. **Furano, A.** 1977. The elongation factor Tu coded by the *tufA* gene of *Escherichia coli* K-12 is almost identical to that coded by the *tufB* gene. J. Biol. Chem. **252**:2154–2157.
14. **Gale, E. F.** 1946. The bacterial amino acid decarboxylases. Adv. Enzymol. **6**:1–32.
15. **Guyer, M. S., R. R. Reed, F. A. Steitz, and K. B. Low.** 1980. Identification of a sex-factor-affinity site in *E. coli* as γδ. Cold Spring Harbor Symp. Quant. Biol. **45**:135–140.
16. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
17. **Hill, C. W., and B. W. Harnish.** 1981. Inversions between ribosomal RNA genes of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **78**:7069–7072.
18. **Jackson, F. R.** 1990. Prokaryotic and eukaryotic pyridoxal-dependent decarboxylases are homologous. J. Mol. Evol. **31**: 325–329.
19. **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.
20. **Legrain, C., P. Halleux, V. Stalon, and N. Glansdorff.** 1972. The dual genetic control of ornithine carbamoyltransferase in *Escherichia coli*. A case of bacterial hybrid enzymes. Eur. J. Biochem. **27**:93–102.
21. **Legrain, C., V. Stalon, and N. Glansdorff.** 1976. *Escherichia coli* ornithine carbamoyltransferase isoenzymes: evolutionary significance and the isolation of λ *argF* and λ *argI* transducing bacteriophages. J. Bacteriol. **128**:35–38.
22. **Lundberg, K. S., D. D. Shoemaker, M. W. W. Adams, J. M. Short, J. A. Sorge, and E. J. Mathur.** 1991. High-fidelity amplification using a thermostable DNA polymerase isolated from *Pyrococcus furiosus*. Gene **108**:1–6.
23. **Lupo, M., and Y. S. Halpern.** 1970. Gene controlling L-glutamic acid decarboxylase synthesis in *Escherichia coli* K-12. J. Bacteriol. **103**:382–386.
24. **Maras, B., E. De Luca, D. Barra, R. A. John, and F. Bossa.** 1990. Primary structure of glutamate decarboxylase from *E. coli*, p. 71–74. *In* T. Fukui et al. (ed.), Enzymes dependent on pyridoxal phosphate and other carbonyl compounds as cofactors. 8th International Symposium on Vitamin B_6 and Carbonyl Catalysis. International Union of Biochemistry Symposium 199, Osaka, Japan. Pergamon Press, Oxford, United Kingdom.
25. **Maras, B., G. Sweeney, D. Barra, F. Bossa, and R. A. John.** 1992. The amino acid sequence of glutamate decarboxylase from *Escherichia coli*; evolutionary relationship between mam-

malian and bacterial enzymes. Eur. J. Biochem. **204:**93–98.
26. **Marcus, M., and Y. S. Halpern.** 1969. Genetic and physiological analysis of glutamate decarboxylase in *Escherichia coli* K-12. J. Bacteriol. **97:**1509–1510.
27. **Marcus, M., and Y. S. Halpern.** 1967. Genetic analysis of glutamate transport and glutamate decarboxylase in *Escherichia coli*. J. Bacteriol. **93:**1409–1415.
28. **Miller, D. L., S. Nagarkatti, R. A. Laursen, J. Parker, and J. D. Friesen.** 1978. A comparison of the activities of the products of the two genes for elongation factor Tu. Mol. Gen. Genet. **159:**57–62.
29. **Miller, R. D., and D. L. Hartl.** 1986. Biotyping confirms a nearly clonal population structure in *Escherichia coli*. Evolution **40:**1–12.
30. **Mishin, A. A., and B. S. Sukhareva.** 1986. Glutamate decarboxylase from *Escherichia coli*: catalytic role of a histidine residue. Dokl. Akad. Nauk SSSR **290:**1268–1271.
31. **Morosov, Y. V., V. P. Almasov, F. A. Savin, and B. S. Sukhareva.** 1982. Experimental and theoretic investigation of the reaction mechanism of glutamate decarboxylase from *E. coli*. Stud. Biophys. **89:**115–122.
32. **Najjar, V. A., and J. Fisher.** 1952. Glutamic acid decarboxylase from *E. coli*. Fed. Proc. **11:**264. (Letter.)
33. **Ochman, H., and R. K. Selander.** 1984. Standard reference strains of *Escherichia coli* from natural populations. J. Bacteriol. **157:**690–693.
34. **Pedersen, S., R. M. Blumenthal, S. V. Reeh, J. Parker, P. Lemaux, R. A. Laursen, S. Nagarkatti, and S. Friesen.** 1976. A mutant of *Escherichia coli* with an altered elongation factor Tu. Proc. Natl. Acad. Sci. USA **73:**1698–1701.
35. **Pedersen, S., S. V. Reeh, J. Parker, R. J. Watson, R. D. Friesen, and N. P. Hill.** 1976. Analysis of the proteins synthesized in ultraviolet light-irradiated λ d*rif*^{d}18 and λ d*fus*-3. Mol. Gen. Genet. **144:**339–343.
36. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. **Schubert, R., J. G. Esanu, and V. Schafer.** 1988. The glutamic acid decarboxylase disc test: an approach to a faster and more simple detection of *E. coli*. Zentralbl. Bakteriol. Mikrobiol. Hyg. Ser. B **187:**107–111.
38. **Seed, B.** 1987. An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. Nature (London) **329:**840–842.
39. **Shukuya, R., and G. W. Schwert.** 1960. Glutamic acid decarboxylase. I. Isolation procedures and properties of the enzyme. J. Biol. Chem. **235:**1649–1652.
40. **Shukuya, R., and G. W. Schwert.** 1960. Glutamic acid decarboxylase. II. The spectrum of the enzyme. J. Biol. Chem. **235:**1653–1657.
41. **Shukuya, R., and G. W. Schwert.** 1960. Glutamic acid decarboxylase. III. The inactivation of the enzyme at low temperatures. J. Biol. Chem. **235:**1658–1661.
42. **Stausbauch, P. H., and E. H. Fisher.** 1970. Chemical and physical properties of *Escherichia coli* glutamate decarboxylase. Biochemistry **9:**226–233.
43. **Stausbauch, P. H., and E. H. Fisher.** 1970. Structure of the binding site of pyridoxal 5′-phosphate to *Escherichia coli* glutamate decarboxylase. Biochemistry **9:**233–238.
44. **Sukhareva, B. S., E. I. Darij, A. I. Goryunov, N. A. Tyul'kova, and L. G. Malikova.** 1989. Study on the properties of bacterial glutamate decarboxylase. Biokhimiya **54:**740–744.
45. **Sukhareva, B. S., A. S. Tikhonenko, and A. I. Goryunov.** 1979. Glutamate decarboxylase of *E. coli*: reversible dissociation of the oligomeric structure, p. 304–310. *In* B. F. Poglazov (ed.), Sborka predbiologicheskikh i biologicheskikh struktur. Materiali Osnovnih Dokladov Mezhdunarodnogo Simpoziuma, Izdatelstvo Nauka, Moscow, 1982.
46. **Tikhonenko, A. S., B. S. Sukhareva, and A. E. Braunstein.** 1968. Electron-microscopic investigation of *Escherichia coli* glutamate decarboxylase. Biochim. Biophys. Acta **167:**476–479.
47. **Van Vliet, F., R. Cunin, A. Jacobs, J. Piette, D. Gigot, M. Lauwereys, A. Pierard, and N. Glansdorff.** 1984. Evolutionary divergence of genes for ornithine and aspartate carbamoyltransferases; complete sequence and mode of regulation of the *Escherichia coli arg*F gene: comparison of *arg*F with *arg*I and *pyr*B. Nucleic Acids Res. **12:**6277–6289.
48. **Yokota, T., H. Sugisaki, M. Takanami, and Y. Kaziro.** 1980. The nucleotide sequence of the cloned *tuf*A gene of *Escherichia coli*. Gene **12:**25–31.
49. **Youngs, T. L., and G. Tunnicliff.** 1991. Substrate analogues and divalent cations as inhibitors of glutamate decarboxylase from *Escherichia coli*. Biochem. Int. **23:**915–922.
50. **Zinder, N. D., and J. Lederberg.** 1952. Genetic exchange in *Salmonella*. J. Bacteriol. **64:**679–699.