Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 8771–8775, November 1990
Neurobiology

Developmentally regulated expression of an exon containing a stop codon in the gene for glutamic acid decarboxylase

(gene structure/alternative splicing)

RICHARD W. BOND*, RUSSELL J. WYBORSKI, AND DAVID I. GOTTLIEB

Department of Anatomy and Neurobiology, Washington University School of Medicine, Saint Louis, MO 63110

Communicated by W. Maxwell Cowan, August 22, 1990

ABSTRACT **In the adult rat brain, the gene for glutamic acid decarboxylase (GAD; L-glutamate 1-carboxy-lyase, EC 4.1.1.15) is expressed predominantly as a 3.7-kilobase transcript. Earlier data showed that embryonic brain expresses an RNA transcript distinct from the adult form; however, the exact structure of this form was not elucidated. Here, transcripts expressed in the embryonic but not the adult brain were cloned and analyzed. These transcripts include an exon not expressed in the adult inserted into the coding sequence. The embryonic exon contains a stop codon that is in-frame with the coding sequence. The exon is found in genomic DNA within the GAD gene where it is flanked by introns with conventional splice sites. On the basis of these structural data, we propose the hypothesis that, early in brain development, transcripts encoding a truncated form of GAD are expressed. The deduced protein cannot function as a decarboxylase because the stop codon in the embryonic exon occurs upstream of the binding site for pyridoxal phosphate, an essential cofactor. Thus, alternative splicing plays a crucial role in the pathway leading to the development of functional GABAergic neurons. The central nervous system-derived cell lines B65 and C6 express a mixture of the adult and embryonic forms of GAD mRNA. They therefore are useful clonal models of central nervous system cells in the early phases of differentiation.**

A wide body of evidence shows that morphologically and physiologically identified classes of neurons in the vertebrate central nervous system (CNS) can also be grouped according to patterns of gene expression. A clear example of this is found in the case of inhibitory neurons which use γ-aminobutyric acid (GABA) as a transmitter. The enzyme glutamic acid decarboxylase (GAD; L-glutamate 1-carboxy-lyase, EC 4.1.1.15) is expressed at high levels in these neurons in widespread regions of the CNS (reviewed in ref. 1). Neighboring excitatory neurons seem devoid of the enzyme even when sensitive immunolocalization techniques are used. The pattern of GAD protein localization is in turn dictated by the pattern of mRNA accumulation. GAD mRNA is detectable by *in situ* hybridization with recombinant DNA probes. Neurons expressing GAD protein are positive for GAD mRNA. Neurons not expressing the enzyme do not express the mRNA (2, 3). Therefore, mechanisms leading to differential mRNA accumulation must play a crucial role in the developmental pathway leading to GABAergic neurons.

There is increasing evidence that posttranscriptional mRNA processing also plays a major role in determining neuron phenotype. For instance, the choice of calcitonin or calcitonin gene-related peptide expression by neurons in the hypothalamus is determined by differential splicing of the same primary transcript (4). In the course of investigating the developmental expression of GAD mRNA in the rat, we previously detected a developmentally regulated change in the transcripts of the GAD gene (5). Protection assays were performed with a 260-base-pair (bp) probe from a cDNA encoding GAD. In the embryonic brain, two pieces were protected. Only one piece was protected by RNA from the adult brain.

Here we show that the embryonic form is associated with another type of differential splicing of RNA transcripts of the GAD gene. Early in development, transcripts are produced that contain an extra exon not found in adult mRNA. This exon contains a stop codon in-frame with the open reading frame of GAD and therefore must encode a truncated protein. We designate this exon the ES (for embryonic stop) exon. We propose that the final development of the GABAergic phenotype entails two levels of control. One is transcriptional and determines that only a subset of neurons expresses the GAD gene. The other involves regulation of the splicing system. Early in development, the ES exon is spliced into GAD transcripts; such transcripts cannot direct the synthesis of enzymatically active GAD. Next, there is a developmental transition that ensures that the ES exon is excluded from GAD mRNA, thus allowing full-length functional GAD to be synthesized.†

METHODS

Polymerase Chain Reaction. (PCR) Amplification, Cloning, and Analysis of ES Exon. Embryonic day 15 rat brain poly(A)$^+$ RNA (1 μg) was reverse transcribed with avian myeloblastosis virus reverse transcriptase using an antisense primer from a rat cDNA sequence (15) homologous to bp 926–945 (GGATATGGCTCCCCCAGGAG) in the rat or 867–886 in the feline (6). An aliquot (1 μl) was amplified by 35 rounds of PCR with primers covering bp 399–418 (CCAAGAACCTGCTTTCCTGT) and 926–945. PCR conditions were 94°C for 1 min, 61°C for 2 min, and 72°C for 1 min. The PCR products were filled in with Klenow fragments and run on a low melting point agarose gel and isolated. This band was then digested with *Bgl* II and ligated into pBS+ (Stratagene), which had been digested with *Bam*HI and *Hinc*II, to create pE15PCR; the insert is 549 bp long.

In a similar manner, a much larger region, which included the translational initiator codon and the ES exon, was amplified from embryonic RNA. An aliquot of rat embryonic day 18 whole brain RNA (1 μg) was used as the template for a reverse transcriptase reaction with a primer corresponding to bp 1182–1203 (GACATAAAGGGGAACAAATCCC) in

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: CNS, central nervous system; GABA, γ-aminobutyric acid; GAD, glutamic acid decarboxylase; ES exon, embryonic stop exon; PCR, polymerase chain reaction.

*Present address: Schering Research, Tumor Biology Division, Bloomfield, NJ 07003.

†The sequences reported in this paper have been deposited in the GenBank data base [accession nos. M38350 (cDNA) and M38351 (genomic)].

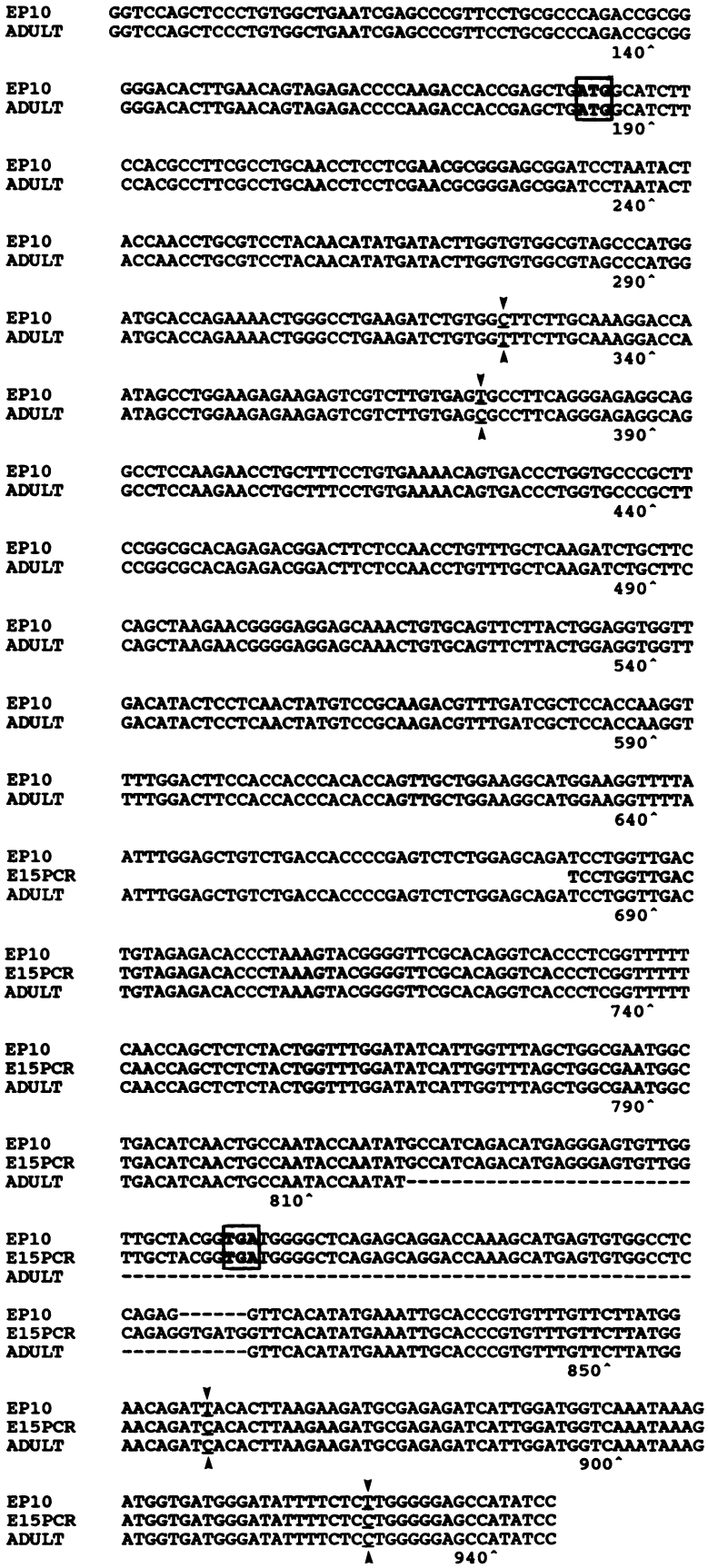

FIG. 1. Sequence of cDNAs from embryonic mRNA. Sequences were obtained from PCR amplification of embryonic mRNA. The products were subcloned as pE15PCR and pEP10. The sequences are compared with the corresponding region of the adult cDNA. Note the insertion of an 80-bp (EP10) or 86-bp (E15PCR) sequence not found in the adult. The putative initiator methionine and termination codon are indicated by boxes. The four nucleotide substitutions in the embryonic sequence are indicated by arrowheads. The adult sequence is from rat cDNA clones previously isolated (15).

```
ccctttctcttcttcttcctttatcagGCCATCAGACATGAGGGAGTGTTGGTTGCTACGG

TGATGGGGCTCAGAGCAGGACCAAAGCATGAGTGTGGCCTCCAGAGGTGATGgtaactga

acactgtt
```

FIG. 2. Genomic sequence of embryonic exon and flanking regions. The sequence of the exon exactly matched the 86-bp insert found in the pE15PCR clone and the 80-bp insert found in pEP10. Splice sites are marked by arrowheads and consensus intronic splice sequences are underlined. The alternative splice site within the ES exon is double underlined.

the adult GAD sequence. Half of the reverse transcriptase reaction was then amplified in a PCR reaction that contained a primer covering bp 926–945 (GGATATGGCTCCCCCAG-GAG) and a primer representing bp 89–106, which also contained an *Eco*RI site (GGGGAATTCGGGAGGGTC-CAGCTCCCT). PCR conditions were 94°C for 1 min, 50°C for 2 min, and 72°C for 3 min for 30 cycles. The reaction products were filled in with Klenow enzyme (Promega), the primers were removed by use of a Centricon-100, and the reaction mixture was extracted twice with phenol/$CHCl_3$/isoamyl alcohol and precipitated with ethanol. The pellet was redissolved and digested with *Eco*RI and ligated into pBS+ that had been digested with *Eco*RI and *Hin*cII. XL1-Blue (Stratagene) competent cells were transformed with the ligation mixture and plated onto NZ plates that contained ampicillin, 5-bromo-4-chloro-3-indolyl β-D-galactoside, and isopropyl β-D-thiogalactopyranoside. Clones containing an embryonic form of GAD were identified by probing lifts with an oligonucleotide (CACCTCTGGAGGCCACACTCATGC) representing the reverse complement of embryonic sequence identified in pE15PCR. A plasmid designated pEP10 was cloned and characterized.

DNA Sequencing. The two clones derived from PCR were sequenced by double-stranded DNA dideoxynucleotide sequencing (7). Reactions were carried out with Sequenase 2.0 (United States Biochemical) and the appropriate primers from the rat sequence. Complete DNA sequences were obtained in both directions.

Isolation and Analysis of Rat GAD Genomic Clones. GAD genomic clones were isolated from a rat genomic library obtained from R. Hynes (Massachusetts Institute of Technology) by screening with probes from the adult rat cDNA. Hybridization conditions were 5× SSPE (1× SSPE = 0.18 M NaCl/10 mM phosphate, pH 7.4/1 mM EDTA)/5× Denhardt's solution (1× Denhardt's solution = 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone)/salmon sperm DNA (0.5 mg/ml)/0.5% SDS/15% formamide at 65°C for 15 hr. Filters were washed in 1× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate)/0.1% SDS at 25°C for 30 min (twice) and 0.1× SSC/0.1% SDS at 65°C for 30 min (twice). A partial sequence verified that the selected clones contained coding exons for GAD. The region containing the sequences surrounding the alternative embryonic splice site was identified by Southern blots using oligonucleotide primers from within the pE15PCR sequence as probes, and the respective regions were subcloned into pBS+ (Stratagene). Double-stranded DNA sequencing utilized primers from within pE15PCR and also from the bordering intron sequences.

Nuclease Protection Assays. Purified pE15PCR was cut with *Bst*E-II and antisense RNA probe (length, 304 bp plus polylinker sequence) was transcribed with T3 RNA polymerase (Stratagene) and purified from a 4% polyacrylamide gel. This segment contained the entire ES exon and surrounding adult sequences (see Fig. 3*A*). Protection assays were performed as described (5). An excess of the ^{32}P-labeled probe was hybridized to 25 μg of the appropriate RNA in 80% formamide for 15 hr at 45°C. Unbound probe was degraded by the addition of RNases A and T1. After proteinase K treatment and phenol/chloroform extraction, the samples were precipitated with ethanol and run on a 6% polyacrylamide sequencing gel. The protected probe was visualized by autoradiography.

Cell Lines. The B65 cell line was obtained from D. Schubert (Salk Institute). C6 cells were obtained from the American Type Culture Collection. Cells were grown in Dulbecco's modified Eagle's medium with 10% fetal calf serum.

RESULTS

To determine the structure of embryonic mRNA in the region of the developmentally polymorphic site (5), cDNA was transcribed from day 15 embryonic rat brain poly(A)$^+$ RNA with reverse transcriptase by using an oligonucleotide primer downstream of the site. This cDNA was then amplified with PCR by using primers that flanked the polymorphic site. The amplified DNA was subcloned and sequenced. The sequence (designated pE15PCR) showed a continuous sequence of 86 bp inserted into the sequence of adult GAD (Fig. 1). Analysis of genomic clones (see below) reveals that the 86-bp sequence corresponds to a single exon. The 86-bp insert is referred to as the ES (embryonic stop) exon. Translation of the sequence in-frame with the adult GAD message reveals a stop codon. Since the sequence has 86 bp, the reading frame in the downstream sequences is shifted and results in multiple stop codons. Thus, the new sequence contains multiple stop codons that prevent it from encoding full-length GAD. In the sequence amplified by PCR, a total of 241 nucleotides were sequenced upstream and downstream of the 86-bp insert. These sequences are identical with the corresponding sequences in the adult cDNA. The identity of these sequences is strong evidence that the embryonic transcript is encoded by the same gene that encodes adult cDNA rather than being the product of a different but related gene.

The sequence described above strongly suggests that a specific embryonic GAD transcript exists. However, it could be argued that this sequence occurs within a highly altered transcript that would not be translated. To exclude this possibility, a longer sequence stretching from a point upstream of the initiator codon, through the ES exon, and into the flanking normal coding region was amplified. This transcript (designated pEP10) was sequenced and analyzed. pEP10 consists of the normal adult sequence beginning 100 bp 5′ to the initiator codon. It includes the ES exon and 120 bp beyond it. Outside the ES exon, 4 nucleotides differ between pEP10 and the adult sequence. Three represent silent substitutions, whereas the thymidine at position 930 changes a proline residue to a leucine. The ES exon in pEP10 is 6 nucleotides shorter than that in pE15PCR. The 3′ splice site in pEP10 is actually a stronger consensus sequence than that in pE15PCR. Most importantly, the 6-bp difference in the two forms does not affect the stop codon contained within the ES exon.

It is known that the PCR amplification procedure can generate sequence variants (8). Therefore, it was crucial to show that the ES sequence amplified from the cDNA is represented in genomic DNA. Genomic clones for GAD were isolated from a rat library in EMBL3 by using GAD cDNA probes at high stringency. Partial sequences of the genomic clones were obtained with primers from the GAD cDNA sequence. These partial sequences confirm that the cloned genomic DNA encodes the GAD cDNA. The subset of clones

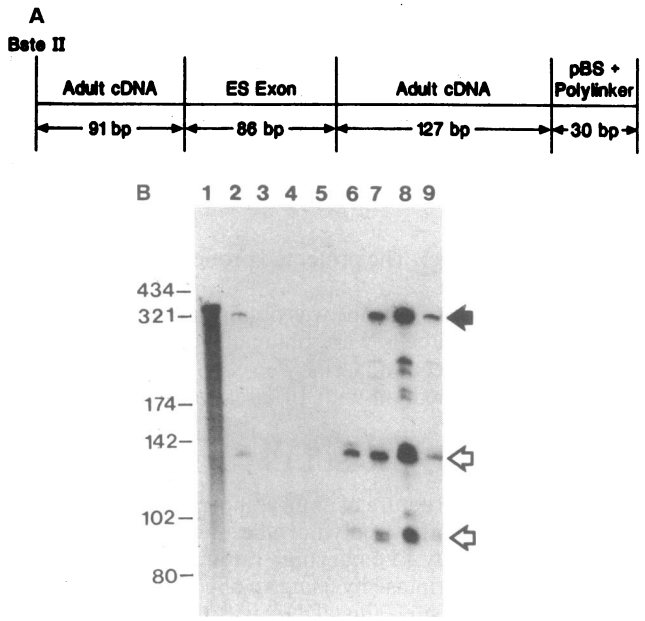

FIG. 3. Nuclease protection assay with GAD cDNA probe, which includes ES exon. (*A*) Diagrammatic representation of pE15PCR T3 RNA antisense probe. (*B*) Total RNA samples (25 μg) were analyzed by nuclease protection assay. Lanes: 1, undigested probe (334 bp); 2, B65; 3, C6; 4, no RNA; 5, liver; 6, adult brain; 7, embryonic day 18 brain; 8, embryonic day 18 striatum; 9, embryonic day 18 cortex. Solid arrow indicates full-length protected probe. Open arrows are fragments corresponding in size to those surrounding the ES exon.

with sequences corresponding to the embryonic sequence was isolated and analyzed by Southern blot analysis. Fragments recognized by the probe from the embryonic sequence were subcloned and sequenced. The 86-bp sequence from pE15PCR is found as an exon in the GAD genomic clone (Fig. 2). The exon is surrounded by intron sequences with consensus splice sites. This exon is designated the ES (embryonic stop) exon.

To analyze the expression of the ES exon in the developing brain, protection assays were done with a probe that includes the ES exon and flanking cDNA sequences (Fig. 3). In RNA from whole embryonic brain, embryonic cortex, and embryonic striatum the full-length piece is protected. Smaller bands that correspond in size to the adult sequences flanking the ES exon are also protected. This implies that by embryonic day 18, the adult as well as the embryonic form is being expressed. There are no protected RNA bands in the liver RNA samples. The intensity of the full-length signal shows that RNA containing the ES exon is fairly abundant in the embryonic brain. In adult brain, the full-length protected piece is absent. However, two fragments of lower molecular weight are protected. Together these fragments add up to the length of the probe minus the ES exon and thus correspond to the regions of the cDNA flanking the ES exon. This shows that the embryonic exon is not expressed in the adult. Three additional bands appear in the protection patterns of embryonic but not adult RNA. These may represent fragments due to intron-containing primary transcripts in the RNA samples.

It has been reported that C6, a clonal cell line of CNS origin, expresses GAD mRNA (9). B65, another CNS-derived clonal cell line expresses GAD enzymatic activity (10). RNA from these cell lines was probed with pE15PCR in protection assays. Full-length pieces were protected with RNA from both lines; the protected band from B65 RNA is visible in Fig. 3. The level of expression is very low in C6 and the bands do not show up in Fig. 3 but are present upon longer exposure. Therefore, both lines express the ES exon in GAD transcripts. There are also bands corresponding to the size of the two fragments from adult brain. This suggests that these lines also express the adult form of the GAD transcript.

DISCUSSION

The availability of cloned gene probes offers the opportunity to analyze the development of CNS neurons in terms of regulatory events affecting cell-specific genes and their products. It is already clear that adult GABAergic neurons accumulate high levels of GAD protein and mRNA compared to neighboring excitatory neurons (1–3). The relative contributions of differential transcription, mRNA stability, and mRNA translation efficiency to this final result have not been determined. In this report, we show that another basic mechanism, differential splicing, must play an important role in the final emergence of functional GABAergic neurons. An exon that has a stop codon has been detected within the GAD gene. This exon is spliced into embryonic RNA transcripts but not into those from the adult.

Two embryonic transcripts have been sequenced. The shorter of these, pE15PCR, contains the ES exon plus 241 bp of flanking sequence upstream that is identical to the adult sequence. The stop codon is in-frame with the open reading frame of the adult sequence. This suggests that pE15PCR is a segment of a transcript that encodes a truncated protein that

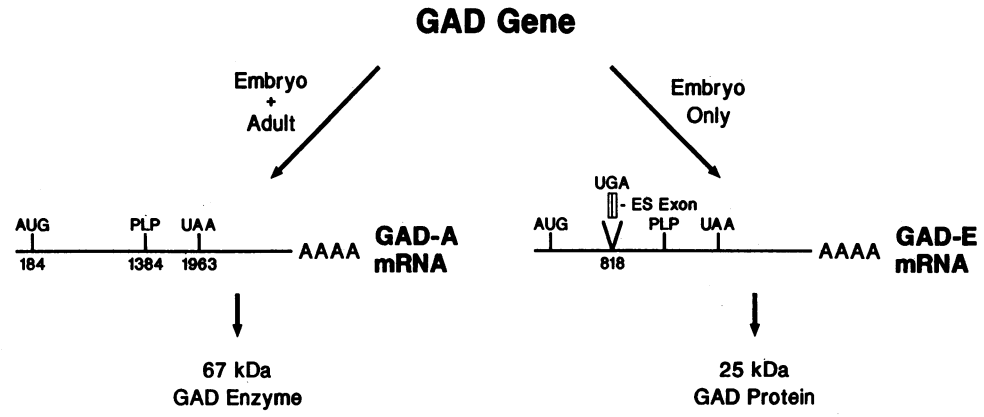

FIG. 4. Proposed model for developmentally regulated alternative splicing of GAD transcripts. During development, differential splicing of the GAD mRNA produces two forms: GAD-A mRNA, which is found both in the adult and embryo, and GAD-E mRNA, which is found only in the embryo. GAD-A encodes fully functional, enzymatically active, 67-kDa GAD. GAD-E encodes a truncated, 25-kDa, GAD-like protein, which lacks the essential pyridoxal phosphate binding site (PLP). AUG is the initiator methionine, UAA is the adult termination codon, and UGA is the embryonic termination codon. Numbers refer to the location of sites according to the adult rat sequence (15).

begins at the initiator codon of the adult sequence and ends at the stop codon of the ES exon. To verify this hypothesis, a longer segment of embryonic RNA was amplified. A sequence including the translation initiation codon and the ES exon was determined (pEP10). It contains an open reading frame of 672 bp that terminates in the stop codon of the ES exon. No other long open reading frame is found in the sequence. Conceptual translation of this sequence predicts a protein of 25 kDa beginning at the initiator methionine of the adult sequence and terminating at the UGA of the ES exon. Ultimate proof that transcripts containing the ES exon are translated must await analysis of the GAD polypeptides found in embryonic brain. The sequence of pEP10 outside the ES exon differs from the adult GAD sequence at four positions. Of these, three are silent substitutions and one changes an amino acid. These differences could be allelic, since different rats were used to amplify the adult and embryonic sequences. Alternatively, some of the changes could be due to substitutions that occurred during PCR amplification. Alternative splicing within an exon such as seen within the ES exon has been demonstrated in another neuronal gene, tyrosine hydroxylase (11).

Nuclease protection assays show that transcripts containing the ES exon are present in embryonic day 18 embryonic brain but not in the adult. However, transcripts without the ES exon are present in embryonic day 18 and adult brains (this scheme is summarized in Fig. 4). The data show that some regions of the brain, including the striatum and cortex, pass through a stage where the embryonic form of the GAD gene is expressed. It has not been determined whether expression of the embryonic form and adult form occurs in the same cells. One possibility is that individual cells first express the embryonic form and then switch to the adult form. Alternatively, the embryonic form may be restricted to a subset of cells that never go on to express the adult form.

The current data are compatible with two alternative interpretations of the role of the ES exon in the ontogeny of GABAergic cells. In the first, the truncated protein predicted by the RNA sequence, whose molecular mass is 25 kDa, would have a biological function. This function clearly cannot be that of a decarboxylase. The binding site for pyridoxal phosphate, the cofactor for GAD, and other decarboxylases is at nucleotide positions 1384–1396 in the rat sequence (15), well downstream from the site of insertion of the ES exon. Thus, the resulting protein would certainly be enzymatically inactive as a decarboxylase. A search of sequence data bases for sequences homologous to GAD fails to turn up homologous sequences. Therefore, it is impossible to propose a specific function for the truncated protein. In the second interpretation, the truncated protein itself is without function. The role of the ES exon is to suppress the synthesis of enzymatically active GAD during early stages of brain development. Perhaps GAD activity is particularly deleterious to immature brain cells.

Two other examples of stop codon-containing exons have recently been found. In *Drosophila*, flies of both sexes transcribe the sex lethal gene. A stop codon-containing exon is spliced into the mRNA of the gene in male but not in female flies. Consequently, the biological activity of this gene is restricted to the female fly (12). Mammalian muscle transcribes a variant form of neural cell adhesion molecule mRNA (13) that also contains a stop codon-containing exon. The truncated protein encoded by this form lacks a membrane attachment site and is hence secreted rather than membrane bound. In a general way, these examples suggest that stop codon-containing exons play a role in increasing the diversity of proteins encoded by the genome and in restricting expression of their ultimate function. The data presented here make it very likely that this mechanism plays a role in the emergence of the GABAergic phenotype during CNS development.

B65 and C6 cells have been extensively used as models for neurons and glia, respectively. The finding that these clonal cells express transcripts with the ES exon implies that they express a repertoire of genes characteristic of early developing cells. Additional support for this idea comes from the fact that B65 cells express the GD3 ganglioside, a marker restricted to embryonic stem cells and immature neurons (14).

We thank Drs. K. O'Malley, E. Sadler, D. Schlessinger, and R. Waterston for helpful discussions, and Dr. R. Hynes for the rat genomic library. This work was supported by National Institutes of Health Grant RO1-NS12867 and by a grant from the Monsanto Corporation.

1. Mugnaini, E. & Oertel, W. H. (1985) in *Handbook of Chemical Neuroanatomy*, eds. Bjorklund, A. & Hokfelt, T. (Elsevier, New York), Vol. 4, pp. 436–608.
2. Wuenschell, C. W., Fisher, R. S., Kaufman, D. L. & Tobin, A. J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6193–6197.
3. Chesselet, M. R. L., Weiss, C., Wuenschell, C., Tobin, A. J. & Affolter, A. J. (1987) *J. Comp. Neurol.* **262**, 125–140.
4. Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S. & Evans, R. M. (1982) *Nature (London)* **298**, 240–249.
5. Bond, R. W., Jansen, K. R. & Gottlieb, D. I. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3231–3234.
6. Kobayashi, Y., Kaufman, D. L. & Tobin, A. J. (1987) *J. Neurosci.* **7**, 2768–2772.
7. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
8. Tindall, K. R. & Kunkel, T. A. (1988) *Biochemistry* **27**, 6008–6013.
9. Tillakaratne, N. J. K. & Tobin, A. J. (1986) *Soc. Neurosci. Abstr.* **12**, 1458.
10. Schubert, D., Heinemann, S., Carlisle, W., Tarikas, H., Kimes, B., Patrick, J., Steinbach, J. H. & Culp, W. (1974) *Nature (London)* **249**, 224–227.
11. O'Malley, K. L., Anhalt, M. J., Martin, B. M., Kelsoe, J. R., Winfield, S. L. & Ginns, E. I. (1987) *Biochemistry* **26**, 6910–6914.
12. Bell, L. R., Maine, E. M., Schedl, P. & Cline, T. W. (1988) *Cell* **55**, 1037–1046.
13. Gower, H. J., Barton, C. H., Elsom, V. L., Thompson, J., Moore, S. E., Dickson, G. & Walsh, F. S. (1988) *Cell* **55**, 955–964.
14. Levine, J. M., Beasely, L. & Stallcup, W. B. (1984) *J. Neurosci.* **4**, 820–831.
15. Wyborski, R. J., Bond, R. W. & Gottlieb, D. I. (1990) *Mol. Brain Res.* **8**, 193–198.