BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **238,** 126–129 (1997)
ARTICLE NO. RC977252

Cloning of a cDNA Encoding a Novel Sugar Transporter Expressed in the Neonatal Mouse Hippocampus

Naoki Matsuo, Shoko Kawamoto, Kenichi Matsubara, and Kousaku Okubo[1]
Institute for Molecular and Cellular Biology, Osaka University, Yamada-oka, Suita, Osaka 565, Japan

Received August 4, 1997

While analyzing active genes in the neonatal mouse hippocampus, we observed several novel genes that were abundantly expressed in this tissue. We report here cloning and sequencing of one of these transcripts, HiAT1 (Hippocampus Abundant Gene Transcript 1). The mRNA was 2.7 Kb in length, and the deduced amino acid sequence consisted of 490 amino acids with characteristics typical of members of the sugar transporter family. However, its overall sequence homology to known transporter cDNAs was only about 30%, suggesting strongly that it represents a novel sugar transporter gene. Northern hybridization analyses showed this transcript is detected in adult and embryonic brains, as well as in other tissues.

Expression profiling of active genes (1-3) in the neonatal mouse hippocampus showed that 32% of total mRNA from this tissue consists of 99 abundantly expressed genes, among which 43 are novel (Matsuo *et al.*, manuscript in preparation). We have initiated to identify these genes in the hope that at least some of them may have important roles either in maturation or in maintaining function of the brain. In this paper we report cloning and sequencing of one of such genes, HiAT1 (Hippocampus Abundant Gene Transcript 1).

MATERIALS AND METHODS

Detection of novel gene transcripts. Total RNA was isolated from hippocampus of 5 day old mouse, strain C57BL/6 (Japan SLC, Inc.) using acid guanidium thiocyanate-phenol-chloroform method. Poly(A) RNA was purified using QuickPrep Micro mRNA Purification kit (Pharmacia) and the 3′-directed cDNA library was constructed (4), and transformed into *E. coli.* The 3′-directed cDNA inserts were sequenced with randomly selected transformant colonies. Sequences thus obtained were compared with each other to examine abundance of gene transcripts, and then sent to GenBank for gene identification.

Cloning of a full-length cDNA. A full-length mouse neonatal hippocampus cDNA library was constructed using a ZAP-cDNA synthesis kit (Stratagene) and screened with a radiolabeled 357bp 3′-cDNA fragment of GS6410 as a probe. Positive plaques, purified through an additional round of screening were subcloned into pBluescript II (Stratagene), and inserts were sequenced for both strands using primer-walking method.

Northern blotting analysis. Total RNA samples were isolated from various regions and stages of brain, using the acid guanidium thiocyanate-phenol-chloroform method. RNA was electrophoresed in 1% agarose gels in the presence of formaldehyde, transferred onto Hybond N membranes (Amersham), and hybridized with the HiAT1 cDNA (see text) labeled by the random priming method. For analyses of tissues other than the brain, Mouse Multiple Tissue Northern Blot filter (CLONTECH) was used. Hybridization was carried out in 6× SSC, 0.5% SDS, 5 × Denhardt's solution at 65°C, and the filters were washed in 2× SSC, 0.1% SDS at 65°C.

RESULTS AND DISCUSSION

By analysis of active genes in the 5-day-old mouse hippocampus, several genes were found to be active including those encoding stathmin, ribosomal protein L38, GAP-43, beta-tubulin T beta 15, thymosin beta-4, ribosomal protein L41 and cytoskeletal gamma-actin. Each of these transcripts represented more than 0.2% of the total mRNA population. Approximately one third of the total mRNA consisted of these abundantly expressed transcripts. We observed 99 genes belonging to this category, among which 43 were novel (Matsuo *et al.*, manuscript in preparation). We initiated cloning and characterization of these active novel genes.

We initiated our work with a gene representing by the poly(A) proximal sequence, GS6410. This sequence, consisting of 357 bp, was used as a probe to screen a mouse neonatal hippocampus cDNA library in Lambda ZAP II. Five positive clones were isolated, and the clone HiAT1 that had the longest insert was sequenced by primer walking. The transcript was 2668 base pairs in

[1] To whom correspondence should be addressed. Fax: +81-06-877-1922. E-mail: kousaku@imcb.osaka-u.ac.jp.

Abbreviations: HiAT1, Hippocampus Abundant Gene Transcript 1; GS, gene signature.

The nucleotide sequence has been deposited in GenBank with the Accession No. D88315.

0006-291X/97 $25.00

```
                                                                    60
GGACCGCACTCGCCGCTGGGGCCGCCCGCCCCCTGAGTGGCTGCTGCAGCGCCGGGAGTC
                                                                   120
GAGGATGGTAAAATGACCCAGGGGAAGAAAAAGAAACGGGCCGCGAACCGCAGTATCATG
  1                M  T  Q  G  K  K  K  K  R  A  A  N  R  S  I  M
CTGGCCAAAAAGATCATCATTAAGGACGGAGGCACGCCTCAAGGAATAGGCTCCCCTAGT     180
 17  L  A  K  K  I  I  I  K  D  G  G  T  P  Q  G  I  G  S  P  S
GTGTATCATGCGGTTATTGTCATCTTTTTGGAGTTTTTTGCTTGGGGACTACTGACAGCA     240
 37  V  Y  H  A  V  I  V  I  F  L  E  F  F  A  W  G  L  L  T  A
CCCACATTGGTGGTATTGCATGAAACCTTCCCTAAACATACATTTCTGATGAATGGCCTA     300
 57  P  T  L  V  V  L  H  E  T  F  P  K  H  T  F  L  M  N  G  L
ATTCAAGGAGTAAAGGGTTTGTTGTCATTCCTCAGTGCCCCACTTATTGGTGCTCTTTCT     360
 77  I  Q  G  V  K  G  L  L  S  F  L  S  A  P  L  I  G  A  L  S
GATGTTTGGGGCCGAAAGTCCTTCCTGCTGCTAACAGTGTTTTTCACGTGTGCCCCTATT     420
 97  D  V  W  G  R  K  S  F  L  L  L  T  V  F  F  T  C  A  P  I
CCTCTGATGAAGATCAGCCCGTGGTGGTACTTTGCTGTTATCTCTGTTTCTGGGGTTTTT     480
117  P  L  M  K  I  S  P  W  W  Y  F  A  V  I  S  V  S  G  V  F
GCAGTAACCTTCTCTGTGGTATTTGCATATGTAGCAGATATAACTCAAGAGCATGAAAGG     540
137  A  V  T  F  S  V  V  F  A  Y  V  A  D  I  T  Q  E  H  E  R
AGTATGGCTTACCGACTGGTTTCTGCTACGTTCGCTGCCAGCTTAGTTACCAGCCCTGCA     600
157  S  M  A  Y  G  L  V  S  A  T  F  A  A  S  L  V  T  S  P  A
ATTGGAGCTTACCTTGGACAAATGTATGGGGACAGCTTAGTGGTGGTCCTTGCTACAGCA     660
177  I  G  A  Y  L  G  Q  M  Y  G  D  S  L  V  V  V  L  A  T  A
ATAGCTCTGCTAGACATCTGTTTTATCCTTGTTGCTGTGCCAGAGTCATTGCCTGAGAAG     720
197  I  A  L  L  D  I  C  F  I  L  V  A  V  P  E  S  L  P  E  K
ATGCGGCCAGCATCGTGGGGAGCTCCCATTTCATGGGAACAGGCTGACCCCTTTGCATCT     780
217  M  R  P  A  S  W  G  A  P  I  S  W  E  Q  A  D  P  F  A  S
TTAAAAAAAGTGGGCCAAGACTCCATAGTGCTGCTGATCTGTATCACGGTGTTTCTCTCC     840
237  L  K  K  V  G  Q  D  S  I  V  L  L  I  C  I  T  V  F  L  S
TACCTACCTGAAGCCGGCCAGTACTCCAGCTTCTTCTTATACCTCAAACAGATAATGAAA     900
257  Y  L  P  E  A  G  Q  Y  S  S  F  F  L  Y  L  K  Q  I  M  K
TTTTCCCCAGAAAGTGTGGCAGCCTTCATAGCAGTCCTTGGCATTCTGTCCATTATTGCA     960
277  F  S  P  E  S  V  A  A  F  I  A  V  L  G  I  L  S  I  I  A
CAGACCATAGTCTTGAGTTTACTCATGAGGTCAATTGGAAATAAGAACACCATCTTACTG    1020
297  Q  T  I  V  L  S  L  L  M  R  S  I  G  N  K  N  T  I  L  L
GGTCTGGGATTTCAAATATTGCAGCTGGCATGGTACGGCTTTGGTTCCGAACCTTGGATG    1080
317  G  L  G  F  Q  I  L  Q  L  A  W  Y  G  F  G  S  E  P  W  M
ATGTGGGCTGCTGGGGCAGTGGCAGCCATGTCTAGCATCACCTTTCCAGCTGTTAGCGCC    1140
337  M  W  A  A  G  A  V  A  A  M  S  S  I  T  F  P  A  V  S  A
CTTGTTTCACGGACTGCTGACGCGGATCAACAGGGTGTTGTTCAAGGGATGATAACAGGG    1200
357  L  V  S  R  T  A  D  A  D  Q  Q  G  V  V  Q  G  M  I  T  G
ATCAGAGGATTGTGCAATGGTCTGGGACCAGCCCTTTATGGGTTCATTTTCTACATATTC    1260
377  I  R  G  L  C  N  G  L  G  P  A  L  Y  G  F  I  F  Y  I  F
CATGTGGAACTTAAAGAACTGCCAATCACAGGGACAGACCTGGGGACAAACACAAGCCCT    1320
397  H  V  E  L  K  E  L  P  I  T  G  T  D  L  G  T  N  T  S  P
CAGCATCACTTTGAACAGAATTCCATCATCCCTGGTCCCCCTTTCCTCTTTGGAGCCTGT    1380
417  Q  H  H  F  E  Q  N  S  I  I  P  G  P  P  F  L  F  G  A  C
TCGGTACTGCTGGCTCTGCTCGTTGCCTTGTTTATTCCGGAACACACTAATTTAAGTTTA    1440
437  S  V  L  L  A  L  L  V  A  L  F  I  P  E  H  T  N  L  S  L
AGGTCCAGCAGTTGGAGGAAGCACTGTGGCAGTCACAGCCATCCTCACAGTACACAAGCG    1500
457  R  S  S  S  W  R  K  H  C  G  S  H  S  H  P  H  S  T  Q  A
CCAGGAGAGGCCAAAGAACCTTTACTCCAGGACACGAATGTGTGACGGAAATCAGGAAGA    1560
477  P  G  E  A  K  E  P  L  L  Q  D  T  N  V  *
CTTCTCTCTCTCGCAGCACCCAGGTCTTAGTTTTCACCTGTAGCTCTGGATGTACATTCC    1620
ATTTCCATCTGCAGTGTACTTTAAGACTGTCTTAAGAAATACTTGCATGAACTCCGTGGG    1680
AACTAAAGAAAGAACCGGACAGTTTTCCAAAGATGTTACACTTTCTTTCAAAAAGAAACC    1740
TTTTGTTTATATTAGCACCACTCTCTTGCCACTAAACTGTTTTATTACACATCTTTAATT    1800
AAAGCTCTATCTGTAACTCCGTCTCTGCCATTTCCTTAGGTGTTGAGCCTTTACCTGCTC    1860
GTGCTGACTCAGCGACACTAGGCAGGATGGTTATGGACCTGTTTGCTTTAACCTTGTAAA    1920
ACCTGAGTCATTGTAAAATCTTGGGTCAAATAATTCAAAGCCTTAATGCAGATGCACTAG    1980
AAAAGAAATGGTAAAGAATTGTTTGCATTTAAAAGAACTTGTGAAAATTGTAGAAAACCA    2040
GAATCATGTTTCAAGCTTGTTTGCCATAATTTTATTTAAAGCATTATTCACGGCTGTTCT    2100
GAAGAATGAAAAGATCAGCCACTTAGAATTCAAATTGGGGTGGTCAGAGCCTAGCCATCT    2160
GGCTCAATCTGTTTAAATCTGTTCTGCTTCTTATACCACTATTCTTTTGATATTTGCATA    2220
ATCATAAGGACCTCAACACTTGAATACATAATCTAAAAATTATATAGTAAAGCTGGTAGC    2280
CTTGAAAATGTCAATGTGATATCTATATGTAGATAAATATATATAGTGGCCTTTCAGGAC    2340
TGTCACAGTAACACTTTATTTACAGAGCTAATGTTTGTCCTAAATTTTCAGGACCCTAGA    2400
AGAGAGCTTTATACAATTACTGATGTGAATTTCTCTAAAGTGTATATTTTTGTGTCCAGT    2460
TATATTATTTAAAAAAGTGTTACTTTGTAAAAATGTACATAAAGTATTGTATAGTTTACA    2520
CTGTTTTCATCTTGTGTGTGGTTACTGCCTAATGCTTTTTAAACTTGGAGCATTCACTAT    2580
GATTAAATAAGGTCTCAAATGTAAATATTCTGTTAATAAAATTAAATATTTTAATGATTT    2640
TTTTCTTTTTAAAAAAAAAAAAAAAAA                                     2668
```

FIG. 1. Nucleotide sequence of HiAT1 cDNA and its deduced amino acid sequence. Nucleotide numbers are shown on the right, and amino acid numbers are shown on the left. Sugar transport protein signatures (PS00216) and gram-positive cocci surface proteins anchoring hexapeptide (PS00343) searched in the PROSITE data base are underlined. The complete nucleotide sequence has been deposited in GenBank with the Accession No. D88315.

length and had one major open reading frame covering a region from position 73 to 1545, encoding a protein of 490 amino acids with a calculated molecular weight of 52,973 Da. The nucleotide and deduced amino acid sequences of the HiAT1 are shown in Fig 1. A homology search with the deduced amino acid sequence showed that it shares 31% homology with *E. coli* tetracycline resistance protein class C (5, 6), and 29% homology with mouse glucose transporter type2 (7, 8) and type4 (9). The *E. coli* tetracycline transporters and facilitative glucose transporters are members of the sugar transporter family which transport antibiotics, carboxylates and sugars in bacteria and mammals (10-12). Several characteristic structural motifs in the sugar transporters have been reported, among which the twelve transmembrane helices (10-14) are the most prominent. HiAT1 showed a similar structure (Fig. 2). The size of the product (490 a.a.) is also in accordance with the size of sugar transporters. The sugar transporter-specific D-R/K-X-G-R-R/K motif is located between the second and third putative transmembrane domains (Figs. 1 and 3A), and a cluster similar to the P-E-S-P-R motif is present at the end of the sixth putative membrane spanning region (Fig. 3B). This motif is conserved in facilitative glucose transporters (10-12), strengthening the idea that the HiAT1 represents a new member of the sugar transporter gene family.

The HiAT1 also has a gram-positive cocci surface proteins anchoring hexapeptide (L-P-X-T-G-X) (see Fig. 1) in the predicted long loop between the eleventh and twelfth transmembrane domains. This is interesting since this hexapeptide is also homologous to the proposed site of cleavage and attachment of glycosyl phosphatidylinositol in eukaryotic proteins anchored with compounds involved in cell adhesion (17, 18). No other sugar transporters carry this motif, suggesting that the protein is a new type of transporter.

To analyze the tissue distribution of the HiAT1 transcript, the cDNA probe was used for northern analysis with RNA prepared from various mouse tissues. In addition to the brain, the 2.7 Kb transcript was widely

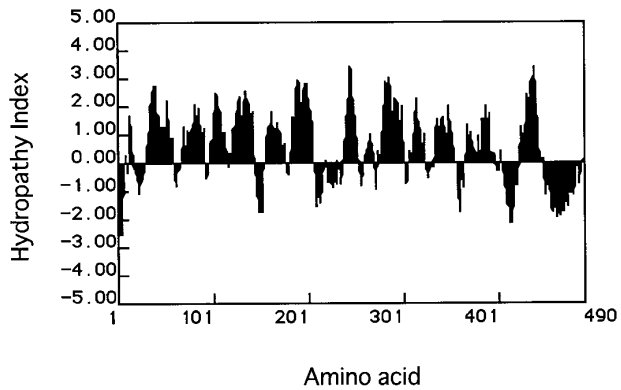

FIG. 2. Hydropathy profile of the deduced amino acid sequence of HiAT1 according to the algorithm of Kyte and Doolittle (15), showing the twelve membrane spanning structures.

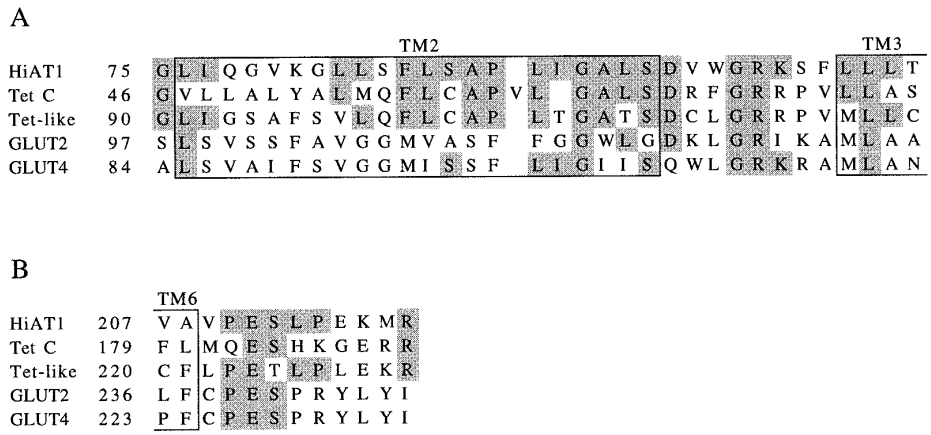

FIG. 3. Multiple sequence alignments around the D-R/K-X-G-R-R/K motif (A) and the P-E-S-P-R motif (B) of HiAT1. Tet C, the *E. coli* tetracycline transporter class C (5, 6); Tet-like, the human tetracycline transporter-like protein found in the Huntington's disease region of 4p16.3 (16); GLUT2, the mouse glucose transporter type2 (7,8); and GLUT4, the mouse glucose transporter type4 (9). Homologous regions are shaded, and putative transmembrane domains (TM) are boxed. Amino acid numbers are shown on the left.

expressed in various tissues, with the exception of the testis in which an additional hybridization-positive 2.2 Kb transcript was highly expressed (Fig. 4A). Northern blotting analyses of RNA extracted from various regions of the mouse brain and at different stages of development showed that the HiAT1 transcripts were evenly distributed throughout the neonatal brain, including olfactory bulb, cerebral cortex, brain stem, hippocampus and cerebellum (Fig. 4B). HiAT1 was highly expressed in the brain at least from embryonic day 13 (E13) to adulthood (Fig. 4C). Thus, the HiAT1 gene is likely to be expressed in almost all cells throughout

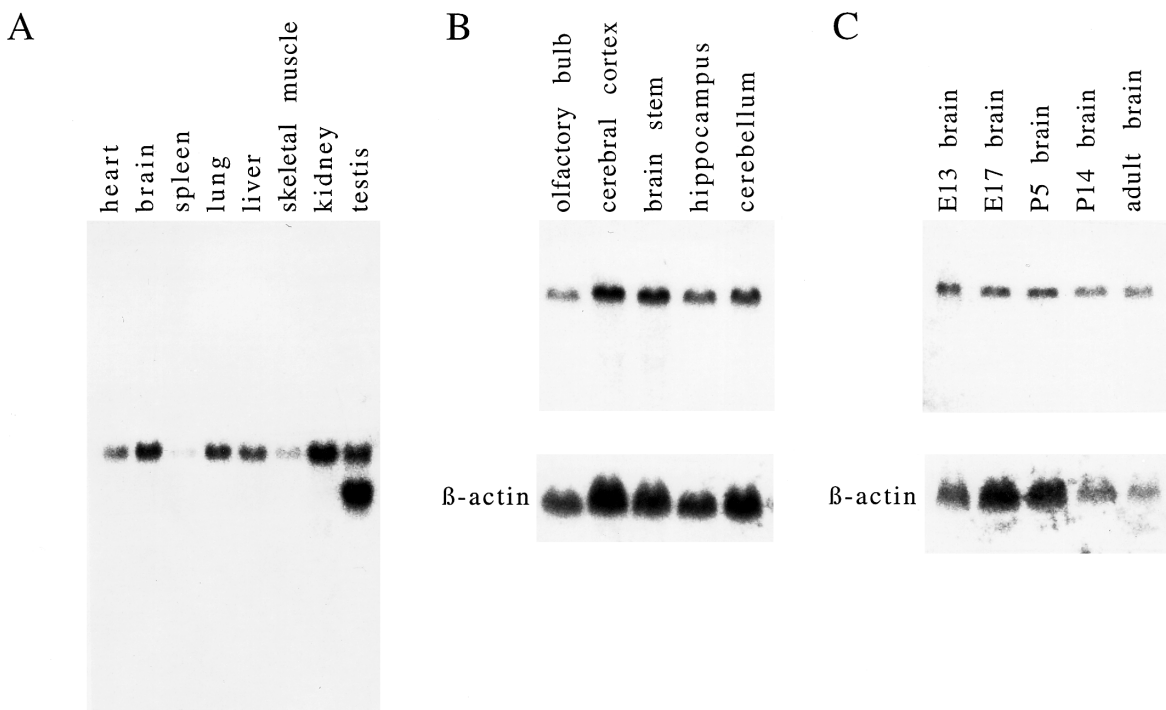

FIG. 4. Northern blotting analyses of the HiAT1 mRNA in various adult mouse tissues (A), in various regions of 5-day-old neonatal mouse brain (B) and in the developing brain (C). Total RNA (15μg) was electrophoresed in a 1% agarose gel. Commercial poly(A) RNA (2 μg) was used for analyses of mRNA from adult tissues (CLONTECH). The same filters were stripped and rehybridized with a β-actin cDNA probe (lower panel).

development from the early stages of embryogenesis similarly to other sugar transporters (19-21), although the expression of the HiAT1 transcripts before E13 has not been examined. Studies to localize of the HiAT1 gene products in cells and their characterization by transport assays are currently in progress.

Many transporters have been isolated by low-stringency hybridization strategy using previously characterized transporter cDNA probes. Cloning of abundantly expressed novel genes in the neonatal mouse brain fortuitously revealed the existence of a new transporter.

ACKNOWLEDGMENTS

This work was supported in part by a Grant-in-Aid for Creative Research (Human Genome Program) from the Ministry of Education, Science and Culture, Japan.

REFERENCES

1. Okubo, K., Hori, N., Matoba, R., Niiyama, T., Fukushima, A., Kojima, Y., and Matsubara, K. (1992) *Nature Genet.* **2,** 173–179.
2. Okubo, K., and Matsubara, K. (1997) *FEBS Letters* **403,** 225–229.
3. Matsubara, K., and Okubo, K. (1993) *Curr. Opin. Biothech.* **4,** 672–677.
4. Okubo, K., Hori, N., Matoba, R., Niiyama, T., and Matsubara, K. (1991) *DNA Seq.* **2,** 137–144.
5. Sutcliffe, J. G. (1979) *Cold Spring Harb. Symp. Quant. Biol.* **43**(Pt. 1), 77–90.
6. Peden, K. W. (1983) *Gene* **22,** 277–280.
7. Suzue, K., Lodish, H. F., and Thorens, B. (1989) *Nucleic Acids Res.* **17,** 10099.
8. Asano, T., Shibasaki, Y., Lin, J. L., Akanuma, Y., Takaku, F., and Oka, Y. (1989) *Nucleic Acids Res.* **17,** 6386.
9. Kaestner, K. H., Christy, R. J., McLenithan, J. C., Braiterman, L. T., Cornelius, P., Pekala, P. H., and Lane, M. D. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 3150–3154.
10. Henderson, P. J. F. (1991) *Curr. Opin. Struct. Biol.* **1,** 590–601.
11. Griffith, J. K., Baker, M. E., Rouch, D. A., Page, M. G. P., Skurray, R. A., Paulsen, I., Chater, K. F., Baldwin, S. A., and Henderson, P. J. F. (1992) *Curr. Opin. Cell Biol.* **4,** 684–695.
12. Baldwin, S. A. (1993) *Biochimi. Biophys. Acta.* **1154,** 17–49.
13. Maiden, M. C. J., Davis, E. O., Baldwin, S. A., Moore, D. C. M., and Henderson, P. J. F. (1987) *Nature* **325,** 641–643.
14. Kilty, J. E., and Amara, S. G. (1992) *Curr. Opin. Biothech.* **3,** 675–682.
15. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
16. Duyao, M. P., Taylor, S. M., Buckler, A. J., Ambrose, C. M., Lin, C., Groot, N., Chirch, D., Barnes, G., Wasmuth, J. J., Housman, D. E., MacDonald, M. E., and Gusella, J. F. (1993) *Hum. Mol. Genet.* **2,** 673–676.
17. Ferguson, M. A. J., and Williams, A. F. (1988) *Ann. Rev. Biochem.* **57,** 285–320.
18. Schneewind, O., Jones, K. F., and Fischetti, V. A. (1990) *J. Bacteriol.* **172,** 3310–3317.
19. Hogan, A., Heyner, S., Charron, M. J., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Thorens, B., and Schultz, G. A. (1991) *Development* **113,** 363–372.
20. Aghayan, M., Rao, L. V., Smith, R. M., Jarett, L., Charron, M. J., Thorens, B., and Heyner, S. (1992) *Development* **115,** 305–312.
21. Bauer, H., Sonnleitner, U., Lametschwandtner, A., Steiner, M., Adam, H., and Bauer, H. C. (1995) *Brain Res. Dev. Brain Res.* **86,** 317–325.