Vol. 159, No. 1, 1989 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
February 28, 1989 Pages 297-304

CHARACTERIZATION OF THE 5'-END REGION AND THE FIRST TWO EXONS OF THE β-PROTEIN PRECURSOR GENE

Giuseppe La Fauci, Debomoy K. Lahiri, Stephen R.J. Salton
and Nikolaos K. Robakis

Mount Sinai School of Medicine, Department of Psychiatry and Fishberg Center of Neurobiology, One Gustave L. Levy Place, Box 1230, New York, NY 10029

Received January 2, 1989

Summary: Human genomic clones encoding the promoter region and the first two exons of the β-amyloid protein precursor (β-APP) gene were isolated. The first exon is 205 base pairs (bp) long and encodes 19 amino acids. The second exon is 168 bp long and encodes 56 amino acids. The 5' -flanking sequence of the β-APP gene was found to display promoter activity in several cell lines including PC12 cells where the highest activity was detected. The promoter region of this gene lacks the typical "TATAA"and "CAAT" boxes usually associated with eukaryotic promoters. Five copies of the GGGCGC sequence are located between positions -107 and -188 and one copy is located within the first exon of the β-APP gene. Consensus sequences recognized by the transcription factors Sp1 and AP-1 are located upstream from the RNA start site. Palindromic sequences capable of forming stable hairpin-like structures are found around the main transcription initiation site. The structural characteristics of the β-APP promoter indicate that multiple elements participate in the regulation of the expression of this gene.

The histopathological findings in the brain of the Alzheimer disease (AD) and adult Down Syndrome (DS) patients include the extracellular deposition of amyloid fibers in the neuritic (senile) plaques and vascular amyloidosis. The amyloid fibers consist of a 4kDa peptide, termed β-protein (1) or A_4 protein (2). Molecular cloning experiments showed that β-protein derives from the proteolytic cleavage of a precursor protein encoded by a gene located on chromosome 21 (3-7). Recently it was shown that at least three distinct forms of the of β-amyloid protein precursor (β-APP) mRNA exist: β-APP_{695} mRNA encodes a protein 695 residues long. β-APP_{751} mRNA is identical to β-APP_{695} mRNA except for a 168 bp insert which encodes 56 amino acids. β-APP_{770} mRNA is identical to β-APP_{751} mRNA except for a 57 bp insert encoding 19 amino acids (8-10). All β-APP mRNAs are probably derived from a single gene by alternative splicing. The amyloid precursor mRNAs are expressed in both neuronal and non-neuronal tissues with the form encoding 695 residues being found predominantly in the brain and the forms encoding 751 and 770 amino acids found mainly in the peripheral tissues (9). In the brain, β-APP mRNA is expressed in both neuronal and glial cells(11), although different neuronal subpopulations may contain different levels of this mRNA (12). In AD the expression of certain β-APP mRNA species may be altered (13,14). An understanding of the regulation of the expression of the different β-APP forms will require the structural characterization of the genomic locus encoding these proteins. A partial characterization of the amyloid gene promoter was reported recently (15). The results presented here extend the characterization of the β-APP promoter and describe the structure of the first two exons of the β-APP gene.

0006-291X/89 $1.50

MATERIALS AND METHODS

Isolation and Sequencing of Human Genomic Clones. A flow sorted human chromosome 21 library was kindly provided to us by Dr. Young (16). Oligonucleotides were prepared in a 380B Applied Biosystems DNA synthesizer and labelled using ^{32}P ATP and polynucleotide kinase (17). Screening of the library was performed as described (17). Hybridizing phages were subcloned into M13mp19 and sequenced by the dideoxynucleotide chain termination procedure (18). To avoid artifacts due to regions with high GC content both strands were sequenced several times using either the Klenow fragment of DNA polymerase I or reverse transcriptase.

Analysis of the 5' -End of the β-APP mRNA. Total human brain RNA was prepared by the guanidine thiocyanate procedure (19). For S1 nuclease mapping the 1.0kb MaeIII-BamHI fragment (Fig. 1), was end-labeled with ^{32}P at the BamHI site. This probe was added to 20μg of total human brain RNA in 85% formamide, 400mM NaCl, 1mM EDTA and 40mM Pipes pH 6.4. The samples were heated at 80°C for 10 minutes and then incubated at 55 or 60°C for 16 hours. S1 digestion was performed at 37°C (20) and the products were analyzed on a 6% acrylamide sequencing gel. For primer extension analysis, a 20mer oligonucleotide corresponding to residues 155-174 of the full length cDNA clone (5) was end-labelled with ^{32}P. Twenty micrograms of total human RNA and 5pmol of the primer were denatured and then allowed to hybridize. Primer extension was then performed as described (20) and the reaction products were analyzed on a 6% polyacrylamide gel.

Plasmid Construction and Chloramphenicol Acetyltrasferase (CAT) Assays. Plasmid pAmyl-CAT was constructed by inserting the 1.2kb PstI-BamHI fragment of the amyloid promoter (Fig. 1) into plasmid pBLACT3 in front of the reporter cat gene (21). This fragment contains 1100bp located immediately upstream from the main transcription start site and the first 105bp of the transcribed sequence.

Cell Culture and CAT assays. All cell lines were grown in Dulbeco's modified Eagle's medium (DMEM) except rat PC12 cells which were grown in RPMI 1640 medium. All media were supplemented with 10% horse serum and 5% fetal calf serum (J.R. Scientific). Twenty four hours before transfection the cells were seeded on poly-L-Lysine coated dishes at a density of 10^6 cells per dish. Transfection was carried out by the calcium phosphate precipitate procedure as described (22) and cell extracts were prepared forty eight hours after glycerol shock. CAT activity was then assayed (22) using 25μl of cell extracts that were normalized for total protein and co-transfected β-galactosidase activity (22). After 2 hours of incubation at 37°C, chloramphenicol and its acetylated derivates were separated by silica gel thin layer chromatography in chloroform: methanol (19:1) and visualized by autoradiography.

RESULTS AND DISCUSSION

To isolate the promoter of the β-APP gene we screened a human chromosome 21 library with an oligonucleotide corresponding to the first 41 nucleotides of the coding sequence of the β-APP cDNA (5). A clone (G70) was identified which contained a 7.8kb insert. A partial restriction map of this clone is presented in Fig. 1. The 1.2kb PstI-BamHI DNA fragment which contains the promoter region of the β-APP gene was used to probe Southern blots of total human genomic DNA digested with the enzymes EcoRI, BamHI or PstI. The number and length of the hybridizing fragments obtained, were those predicted from Fig. 1 (data not shown). These results suggest that only one copy of the the β-APP gene exists in the human haploid genome and that the various β-APP mRNAs derive from a single genomic locus through alternative splicing.

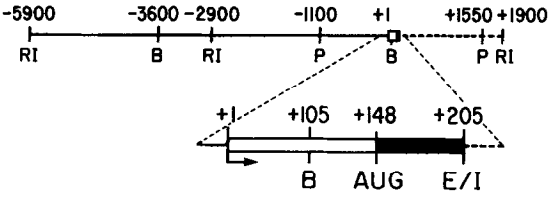

Fig. 1 Structure and partial restriction enzyme map of clone G-70. Untranscribed regions are represented by a solid line. The open box represents the 5'-untranslated region. The filled box represents coding region. The main transcriptional start site is at +1. Arrow indicates the direction of the transcription. Intron sequences are represented by a broken line. AUG is the translation initiation Codon. E/I is the first Exon-Intron junction. RI, B, P represent recognition sites for the enzymes EcoRI, BamHI and PstI respectively. The numbers represent base pairs.

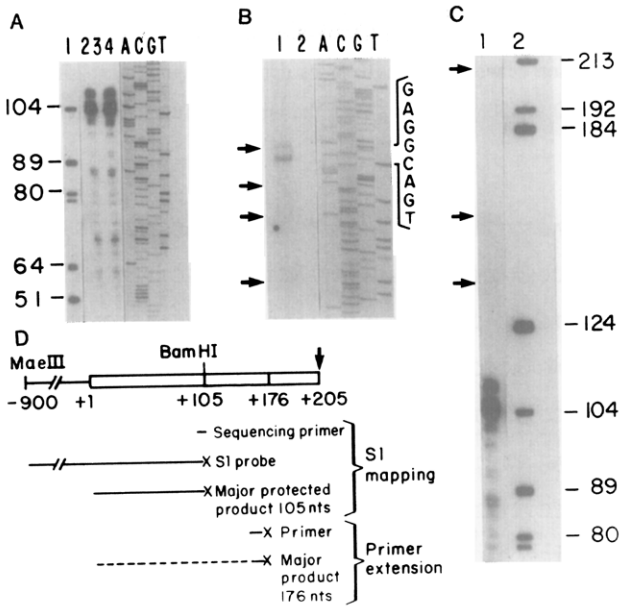

Fig. 2 Determination of the 5'-end of the β-AAP RNA by S1 nuclease mapping (A) and primer extension (B). (A), the 1Kb MaeIII-BamHI fragment (-900 to +105) labeled with ^{32}P at the 5'-end of the BamHI site was used as a probe. Hybridization and S1 digestion were performed as described under "Materials and Methods". Sequencing was carried out using as a primer a 20-mer whose 5'-end coincides with the 5'-end of the S1 probe. Lane 1; DNA size markers. Lanes 2 and 4; S1 digestion products after the probe was hybridized to total brain RNA at 55°C or 60°C respectivity. Lane 3; as in lane 2 in the presence of 20μg of yeast tRNA instead of brain RNA. Sequencing and S1 digestion products were analyzed in a 6% sequencing gel. (B),the primer for the extension and sequencing experiments is described in the "Materials and Methods" section. Arrows indicate the position of additional minor extension products. Lane 1, extension products in the presence of total brain RNA; lane 2, extension products in the presence of tRNA. The sequence of the antisense strand around the major transcription initiation site is indicated. (C), lane 1 shows the detection of low abundance long DNA fragments in the S1 nuclease assay. Lane 2, DNA size markers. (D), schematic representation of the probes and primers used in the S1 nuclease and primer extension assays. Numbers represent base pairs, x indicates the 5'-labelled end of the probes and the arrow indicates the exon/intron junction.

A combination of the S1 nuclease protection assay and primer extension reaction was used to identify the transcription initiation site of the β-APP gene. The 1kb MaeIII-BamHI DNA fragment extending from nucleotide +105 to nucleotide -900 (Fig.2) was hybridized to human brain RNA. S1 nuclease treatment of the DNA-RNA hybrids produced many protected fragments suggesting that there are multiple transcription initiation sites in this gene (Fig.2). The two most prominent fragments were about 105 and 109 nucleotides long. Comparison of the S1 products to the fragments obtained from the sequencing of the MaeIII-BamHI probe indicated that the initiation sites are located 148 and 152 nucleotides upstream from the translation initiation codon (Fig.2). Primer extension and sequencing experiments using a 20-mer primer located 176 nucleotides downstream from the putative transcription start site (Fig. 2) also indicated that the main RNA start site is located 148 nucleotides upstream from the translation initiation codon. Therefore, this site was assigned as the transcription initiation site (+1). A less prominent extension product 179 nucleotides long was also detected. This product corresponds to the second major initiation site detected by the S1 assay.These results differ from a recent report where only one major initiation site was detected (15). Minor extension products corresponding to transcripts initiated at positions +9, +19 and +36 were also found (Fig. 2B). Three more S1 products were detected which indicated minor initiation sites further upstream around positions -40, -65 and -130 (Fig 2C).

Inspection of the DNA sequences around the transcription initiation site (Fig. 3) shows that the "TATA" box usually present in a region between 20 and 30 bases upstream from the cap site is missing here. Similarly, the "CAAT" box located 100 to 150bp upstream from the initiation site is absent. Several other genes with mul-

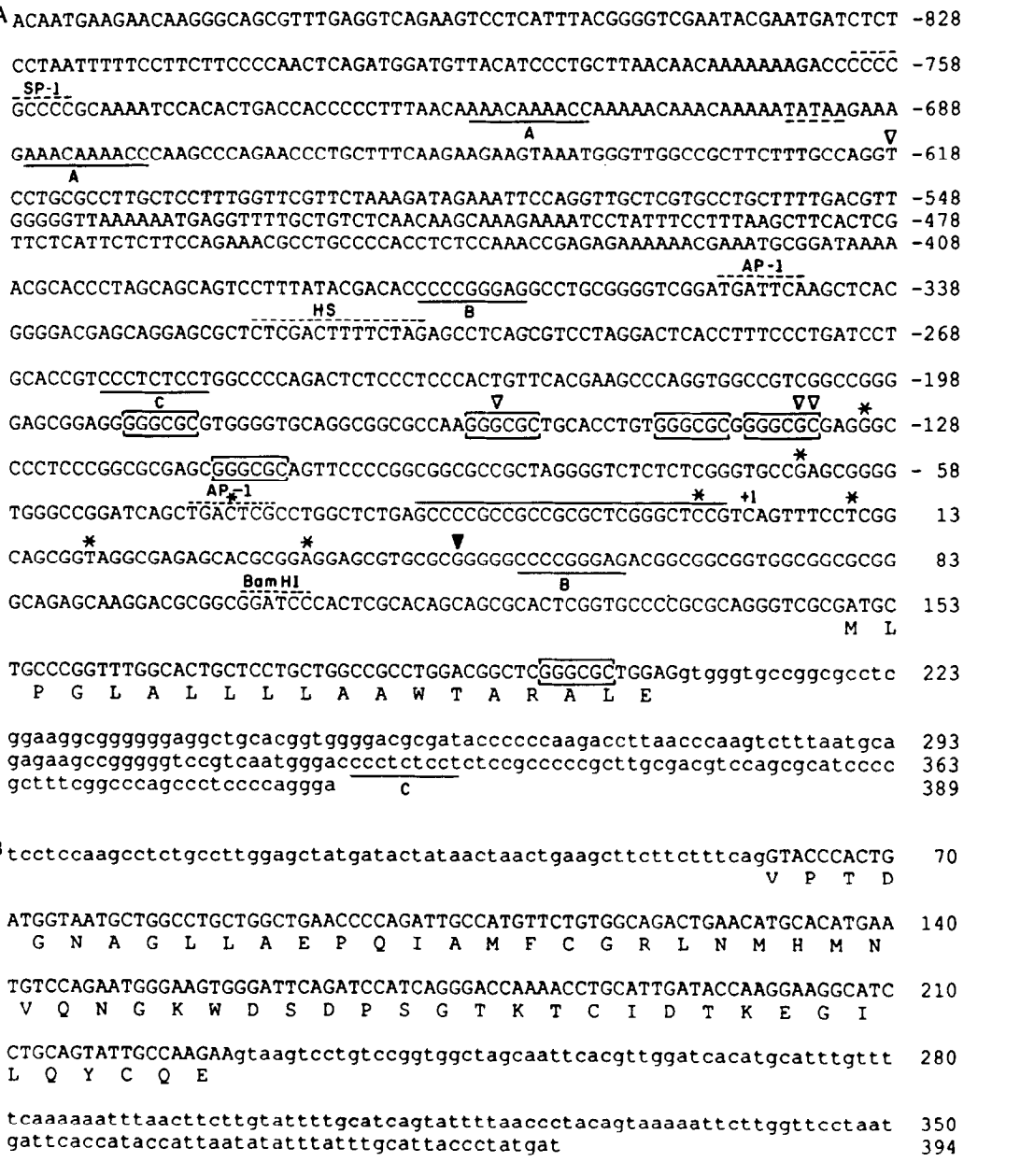

Fig. 3 (A), Nucleotide sequence of the 5′ region and first exon of theβ- AAP gene. Nucleotide +1 corresponds to the cytidine of the major transcription initiation site. * Represents additional transcription initiation sites. The putative SP-1, AP-1 and heat shock (HS) control element binding sites are overlined with a broken line. The six GGGCGC elements are boxed. A, B, and C, represent three pairs of direct repeat elements. The sequence capable of forming hairpin-like structures is overlined with a solid line. The TATAA box downstream from the SP-1 site is underlined with a broken line. Open arrowheads mark the sites of difference with previously published β-APP promoter sequence (15) and the filled arrowhead indicates the difference with the cDNA clones (5,8). (B), The sequence of the second exon of the AAP gene. Amino acids are represented by the single letters code and small letter represents intron sequences.

tiple RNA start sites have been reported to lack the TATA and CAAT boxes. For most of these genes multiple RNA start sites have been described (23-26). The region immediately upstream from the main initiation site is GC rich and contains five copies of the sequence GGGCGC in an area extending from -107 to -188. These hexamers may constitute signals for the binding of a transcription factor. We have noticed that the same hexamer occurs at positions -95, -340, -480, -495, -540, -800 and -827 in the promoter of the human insulin receptor gene (26). These observations suggest that the β-APP gene and the insulin receptor gene may share common regulatory elements. There are several differences between the sequences reported here and those reported previously (Fig. 3). In addition we find only one copy of the 9-base pair concensus sequence GGGCGC$^{G}_{A}$AGG reported to occur six times in this region (15). Since we sequenced both strands of this region several times with two different enzymes (see methods), the observed differences could be due to either the isolation of two different allelles of the amyloid gene or to errors in the sequence reported previously (15). Nevertheless, since some of the differences occur within two GC boxes, they could affect the activity of the promoter.

A nine-nucleotide sequence found at position -740 of the amyloid promoter shows a perfect homology to the inverted sequence recognized by transcription factor SP1 (Fig. 3). Transcription signals for this factor have been shown to function in either direction (27) and therefore this sequence might represent an SP1 binding site. This sequence is followed by the sequence TATAA located at -684. The relative position of these two promoter elements suggests that they may be involved in the minor transcription initiation events occurring upstream of the main initiation site. At -310 we also find a sequence homologous to the consensus binding site of the heat shock control element (15). There are two pairs of direct repeats in the β-APP promoter each at least nine nucleotides long. Two heptamer sequences similar to the consensus sequence of the enhancer element TGACTCA recognized by the transcription factor AP1 are located at -41 and at -348 (Fig. 3). Binding sites for this factor have been found in the enhancer region of the collagenase, human metallothionein IIA and SV40 genes (28). These genes are transcriptionally induced by the tumor promoting agent phorbol ester which seems to modulate the activity of the factor AP1. It would be interesting to see whether the amyloid promoter also responds to treatment by Phorbol esters.

Using the Intelligenetics Search Algorithm (29) we observed that the region encompassing nucleotides -2 to -26 can form stable hairpin like structures by pairing with any one of the following regions: -46 to -68, +7 to +29; +16 to +41, +33 to +53 and +49 to +75. The free energy of these structures are -30,-36,-26,-40 and -49 kcal/mole respectively (Fig. 4).This suggests that transcripts initiated upstream from the +1 site (Fig 3) will have a high probability of forming hairpin like structures which may act as transcription termination signals. It has been suggested that similar structures function as transcription termination signals in both prokaryotic and eukaryotic cells (30). Efficient termination at the hairpin structures would explain the very low abundnance of transcripts initiated upstream from the two main RNA start sites. Alternatively, these dyad summetries may function as signals for the binding of factors which modulate the regulation of the expression of this gene.

We mapped the first intron-exon junction by partially sequencing the 4.8kb EcoRI genomic fragment (Figs 1 and 3). This junction occurs 57 nucleotides after the ATG codon and the sequences around the splicing point obey the AG/GT rule (31). Therefore, the first exon of the β-APP gene is 205 nucleotides long and encodes 19 amino acids. Comparing the sequence of the first exon with those of the cDNA (5,8) we observed that the genomic sequence contained an additional G in position 48. Interestingly a GC box is found within the first exon and a nine-baise pairs direct repeat element is common between the promoter region and the first intron of the β-APP gene (Fig. 3) suggesting that these regions may participate in the regulation of the expression of this

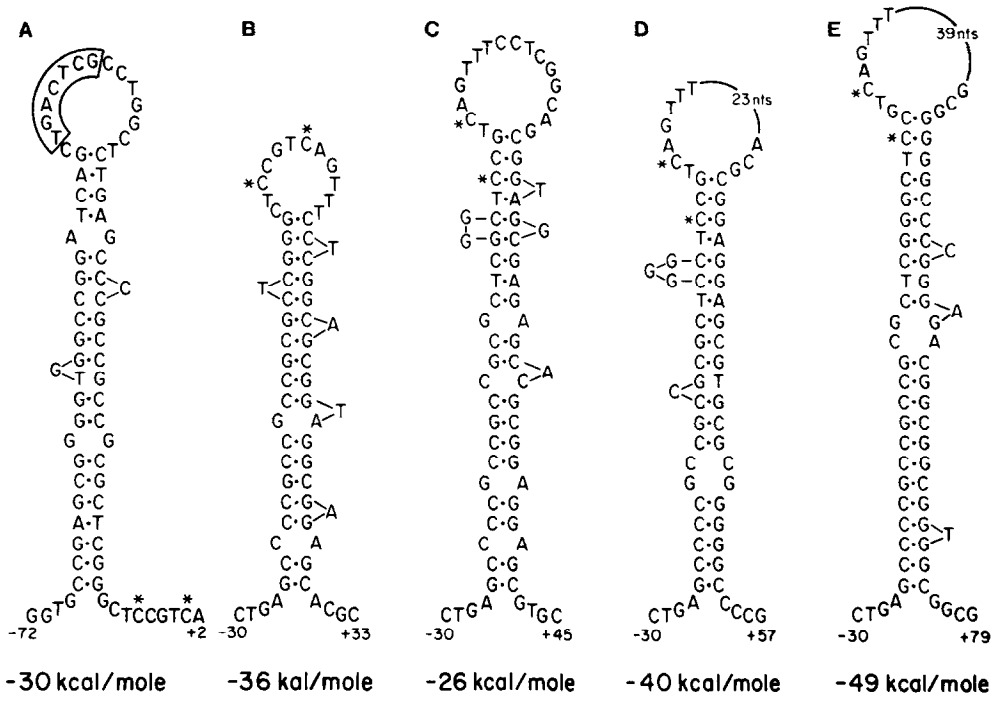

Fig. 4 Proposed secondary structure of the palindromic sequences around the two major transcription initiation sites. The two major transcriptional start sites are indicated by *. The putative AP-1 binding site is boxed. The estimated thermal stability of each structure is given at the bottom of the figure. For more details see Text.

gene. Screening of the library with an oligonucleotide probe corresponding to nucleotides 62 to 101 (5) a clone (E91) containing a 1.8kb insert was isolated. Sequencing revealed that this clone contained the second exon (Fig. 3B). Clone E91 did not hybridize to clone G-70 indicating that the first intron of the β-APP gene is at least 1.7kb long. Clone E91 failed to hybridize to cDNA sequences immediately 3′ to the second exon-intron junction indicating that no other coding sequences are present in this clone. The sequence around both the 5′ and 3′ intron-exon junctions were in agreement with the AG/GT rule (31). We have isolated 13 additional genomic clones each enccoding a distinct exon unit of the amyloid locus (G. La Fauci, R. Ramakrishna, N. Robakis, unpublished results). These observations suggest that this gene contains a large number of exons and therefore additional β-APP mRNAs may exist as a result of alternative splicing.

To demonstrate that the PstI-BamHI 1.2kb fragment encodes for a functional promoter we constructed plasmid pAmyl-CAT where the gene encoding the enzyme chloramphenicol acetyltransferase is under the control of the amyloid promoter (Fig 5). This construct was used to transfect HeLa, CV-1 and PC12 cells. The transient expression of the CAT gene was monitored by measuring CAT activity (22). As positive control we used plasmids pRSV-CAT and pTK-CAT, in which the Rous sarcoma virus LTR region and the Herpes simplex virus (HSV) thymidine kinase promoter respectively, have been fused to the CAT gene. To compare the strength of the β-APP promoter among different cell lines, differences in transfection efficiency were normalized by cotransfecting plasmid pRSV-β-galactosidase, which encodes for a β-galactosidase fusion protein, and the various CAT-plasmid constructs. β-Galactosidase activity was then assayed as described (22). In addition cell extracts were normalized for total protein. It can be seen in Fig 5 that β-APP gene promoter activity was

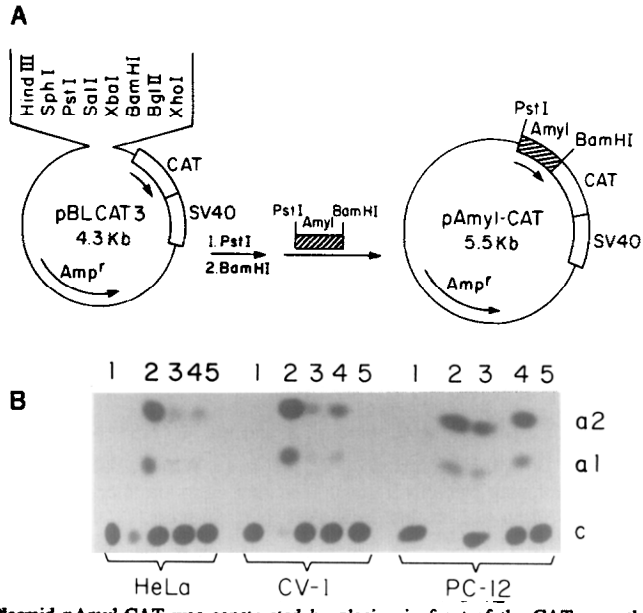

Fig. 5 (A), Plasmid pAmyl-CAT was constructed by placing in front of the CAT gene the 1.2 Kb PstI-BamHI DNA fragment of clone G70. (B), The CAT enzymatic activity in different cell lysates was measured 48 hours after glycerol shock with the following plasmids: Lane 1, no plasmid; Lane 2,pRSV-CAT; Lane 3, pTK-CAT; lane 4, pAmyl-CAT; lane 5, promoterless plasmid pBLCAT3. The cell lines transfected are indicated. C, Chloramphenicol; al and a2, are the two forms of monoacetylated chloramphenicol. For more details see text.

detected in each of the three cell lines transfected. The strength of the amyloid promoter was comparable to the strength of the TK promoter. In HeLa cells a 3 to 4% conversion to acetylated chloramphenicol products was detected for both promoters. In CV-1 cells the percent conversion for the TK and amyloid promoters was 6 and 11 respectively, and in PC12 cells the respective numbers were 30% and 35%. The highest conversion was detected with the RSV promoter constructs in all three cell lines. Amyloid promoter activity was also detected in GH3 cells (rat pituitary), AtT20 (mouse pituitary) and L cells (mouse fibroblasts) (data not shown). Transfection with the promoterless plasmid pBLCAT3 (21), showed very little activity (Fig 5). The detection of the β-APP gene promoter activity in six different cell lines accords with the almost ubiquitous expression of this gene(9). The high biological activity of this promoter in PC12 cells is consistent with recent reports demonstrating the presence of high levels of the amyloid precursor peptides in these cells (32). Further experiments will allow a detailed ayalysis of the structure and molecular mechanisms underlying the regulation of the expression of this gene.

ACKNOWLEDGMENTS

We thank J.R. Lundblad for providing pTK-CAT pRSV-β-galactosidase and pRSV-CAT and his help with the computer program, A.K. Jaiswal for providing plasmid pBLACAT3, J.L. Roberts in whose laboratory the transfection studies were performed, G. Wolfe for fine technical assistance, and D. Carman for typing the manuscript. S.R.J.S. was supported by a Pfizer post-doctoral fellowship. This work was supported by NIH grant AG08200 and by a grant from the McKnight Foundation to N.K.R.

REFERENCES

(1) Glenner, G.G., and Wong, C.W. (1984) Biochem. Biophys. Res. Commun. 122, 1131-1135.

(2) Masters, C., Simms, G., Weinman, N.A., Nulthaup, G., McDonald, B.L. and Beyreutler, K. (1985) Proc. Natl. Acad. Sci USA 82, 4245-4249.

(3) Robakis, N.K., Wisniewski, H.M., Jenkins, E.C., Devine-Gaye, E.A., Houck, G.H., YXO, X.L., Ramakrishma, N.; Wolfe, G.; Silverman, W.P.; and Brown, W.T. (1987) Lancet 1, 384-385.

(4) Robakis, N.K., Ramakrishna, N., Wolfe, G. and Wisniewski, H.M. (1987) Proc. Natl. Acad. Sci. USA 84, 4190-4194.

(5) Kang, J., Lemaire, H.G., Unterbeck, A; Salbaum, J.M.; Masters, CL.L., Greschik, K.H.; Multhaup, G.; Beyreuther, K., and Muller-Hill, B., (1987) Nature 325, 733-736.

(6) Goldgaber, D., Lerman, M.L., McBride, O.W., Saffiotti, U. and Gujdusek, D.C. (1987) Science 235, 877-880.

(7) Tanzi, R.E., Gusella, J.F., Watkins, P.L., Bruns, G.A.P., George-Hyslop, P.S. , Vankeuren, M.L. Patterson, D., Pagan, S., Kurnit, D.M. and Never, R.L. (1987) Science 235, 880-884.

(8) Ponte, P., Gonzalez-DeWhitt, Schilling, J., Miller, J., Hsu, D., Greenberg, B., Davis, K., Wallace, W. Lieberburg, I., Fuller, F. and Cordell, B. (1988) Nature 331, 525-527.

(9) Tanzi, R.E., McClatchey, A.I., Lomberti, E.D., Villa-Komaroff, K., Gusella, J.F. and Neve, R.L. (1988) K., Gusella, J.F. and Neve, R.L. (1988) Nature 331, 526-530.

(10) Kitayuchi, N., Takahashi, Y., Tokushima, Y., Shiojiri, S. and Ito, H. (1988) Nature 331 530-532.

(11) Robakis, N.K., Lahiri, D., Brown, H., Rubenstein, R., Mehta, p., Wisniewski, H.M., and Goller, N. (1988) In: "Disorders of the Developing Nervous System: Changing views on thier origins, Diagnoses and Treatments." (Swann, J.W. ed), Alan R. Liss, Inc. New York, NY. In press.

(12) Bahmanyar, J., Higgins, G.A., Goldgaber, D. Lewis, D.A., Morrison, J.H. Wilson, M.L., Shankar, S.K. and Gajdusek, D.C. (1987), Science 237, 77-80.

(13) Higgins, G.A., Lewis, D.A., Bahmanyar, S. Goldbaber, D. GAjdusek, D.C. Young, W.G., Morrison, J.H. and Wilson, M.L. (1988) Proc. Natl. Acad. Sci. USA 85, 1297-1301.

(14) Cohen, M.L., Golde, T.E., Vsiak, M.F. and Younkin, S.G. (1988) Proc. Natl. Acad. Sci. USA 85, 1227-1231.

(15) Salbaum, J.M., Weidermann, A., Lemaire, H-G., Masters, C.L. and Beyreuther, K. (1988) EMBO J. 7, 2807-2813.

(16) Krumlauf, R., Jeanspierre, M., Young, B.D. (1982) Proc. Natl. Acad. sci. USA 79, 2971-2975.

(17) Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

(18) Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.

(19) Chirgwin, J.M., Przybyla, A.E., MacDonald. R.J., and Rutter, W.J. (1979) Biochemistry 18, 5294-5299.

(20) Calzone,F.J., Britter, R.J. and Davidson, E.H. (1987) Methods in Enzymol. 152, 611-632.

(21) Luckow, B. and Schutz, G. (1987) Nucleic Acid Res. 15, 5490.

(22) Gorman, C.M.(1987) In DNA cloning: A practical approach (D.M. Glover ed) vol. 2 pp. 143 - 190. IRL press, Oxford- New York.

(23) Valerio, D., Duyvesteyn, M.G.C. Dekker, B.M.M., Weeda, G. Berkvens, T.M., van der Voorn, L., van Ormandt, H. and Van der Eb, A.J. (1985) EMBO J. 4, 437-443.

(24) Ishii S., Xu, Y-H, Stratton, R.H., Roe, B.A., Merlino, G.T., and Pastan, I. (1985) Proc. Natl. Acad. Sci. USA 82, 4920-4924.

(25) Melton, D.W. Konecki, D.S., Brennand, J. and Laskey, C.T. (1984) Proc. Natl. Acad. Sci. USA 81, 2147-2151.

(26) Araki, E., Shimade, F., Uzawa, H., Mori, M., and Ebina, Y., (1987) J. Biol. Chem. 262, 16186-16191.

(27) Mitchell, P.J., Carothers , A.M., Han, J.H. Harding, J.D.,Kas, E., Venolia, L. and Chasin, L.A. (1986) Mol. Cell biol. 6, 425-440.

(28) Lee, W., Mitchell, P., and Tjian, R. (1987) Cell 49, 741-752.

(29) Korn, L.J., Queen, C.L., and Wegman, M.N. (1977) Proc. Natl. Acad. Sci. USA 74, 4401-4405.

(30) Platt, T. (1986) Ann. Rev. Biochem. 55, 339-372.

(31) Mount, S.M. (1982) Nucleic Acids Res. 10, 459-472.

(32) Schubert, D., Schroeder, R., LaCorbiere, M., Tsunao, S. and Cole, G. (1988) Science 241, 223-226.