Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 4190–4194, June 1987
Genetics

Molecular cloning and characterization of a cDNA encoding the cerebrovascular and the neuritic plaque amyloid peptides

(Alzheimer disease/Down syndrome/human chromosome 21/β protein)

NIKOLAOS K. ROBAKIS, NARAYANARAO RAMAKRISHNA, GLORIA WOLFE, AND HENRYK M. WISNIEWSKI

New York State Office of Mental Retardation and Developmental Disabilities, Institute for Basic Research, 1050 Forest Hill Road, Staten Island, NY 10314

Communicated by John J. Burns, March 16, 1987 (received for review February 10, 1987)

ABSTRACT **Deposits of amyloid fibers are found in large numbers in the walls of blood vessels and in neuritic plaques in the brains of patients with Alzheimer disease and adults with Down syndrome. We used the amino acid sequence of the amyloid peptide to synthesize oligonucleotide probes specific for the gene encoding this peptide. When a human brain cDNA library was screened with this probe, a clone was found with a 1.7-kilobase insert that contains a long open reading frame coding for 412 amino acid residues including the 28 amino acids of the amyloid peptide. RNA gel blots revealed that a 3.3-kilobase mRNA species was present in the brains of individuals with Alzheimer disease, with Down syndrome, or with no apparent neurological disorders. Southern blots showed that homologous genes are present in the genomic DNA of humans, rabbits, sheep, hamsters, and mice, suggesting that this gene has been conserved through mammalian evolution. Localization of the corresponding genomic sequences on human chromosome 21 suggests a genetic relationship between Alzheimer disease and Down syndrome, and it may explain the early appearance of large numbers of neuritic plaques in adult Down syndrome patients.**

Alzheimer disease (AD) is the most frequent cause of dementia and presently afflicts over two million elderly individuals in the United States. The pathological findings in AD brain include neurofibrillary tangles within neurons, the extracellular deposition of amyloid fibers in the neuritic (senile) plaques, and cerebrovascular amyloidosis. These lesions are used for the posthumous diagnosis of AD. Changes in the brains of Down syndrome (DS) individuals over the age of 40 closely resemble the neuropathology in AD (1). The amyloid fibers consist of β-pleated sheets and exhibit a characteristic green birefringence when stained with congo red and viewed in polarized light. This structure is observed in amyloid fibrils from a variety of types of amyloidosis and is considered to be a common characteristic of a diverse collection of diseases (2). The partial amino acid sequences of the cerebrovascular amyloid peptides present in AD and DS patients were determined and found to be essentially identical up to residue 28 (3, 4). The molecular mass of these peptides was ≈4 kDa (3). Immunocytochemical and biochemical studies showed that the cerebrovascular amyloid and the neuritic plaque amyloid share common sequences (4, 5). In the present study these peptides will be referred to as the β protein (3). Conflicting evidence has been reported on the relationship between neurofibrillary tangles and the β protein. Based on sequencing data, Masters *et al.* (6) reported that neurofibrillary tangles consist of β protein. Other groups, however, failed to detect β protein in neurofibrillary tangles (4, 7).

Although the etiology of AD is unknown, it has been demonstrated that the frequency of neuritic plaques found in the cortex of AD patients correlates with the degree of dementia. Familial cases of AD are also known and tend to follow an autosomal dominant pattern of inheritance (8). In an effort to better understand the origin and the pathogenesis of the AD amyloid plaque, we have isolated and characterized a human brain cDNA clone encoding the β protein.*

MATERIALS AND METHODS

A mixture of sixteen 72-mer oligonucleotides, 5′-GATGCT-GAGTTC$^{C}_{A}$G$^{C}_{G}$CATGAC$^{AG}_{TC}$TGGCTATCAGGTGCACCA-TCAGAAGCTGGTGTTCTTCGCTGAGGATGTG3′, was synthesized by Applied Biosystems (Foster City, CA). The nucleotide sequence was deduced by reverse translating the amino acid sequence of the amyloid peptide (3). The only degeneracies introduced into the probe are in the six codons encoding arginine and serine. The choice of the nucleotide sequence was based on the codon bias observed in human mRNAs (9), the infrequency of the dinucleotide CpG in vertebrate genes (10), the relative stability of G·T versus A·C mismatches (11), and the sequence permitting minimum secondary structure in the probe. The probe was labeled to a specific activity of 5×10^8 cpm/μg using [γ-^{32}P]ATP (6000 Ci/mmol; 1 Ci = 37 GBq; ICN) and T4 polynucleotide kinase (12). A human brain λgt11 cDNA library, constructed from poly(A)$^+$ mRNA isolated from the temporal cortex of a 50-year-old male cancer victim, was purchased from Clontech (Palo Alto, CA). The library was used to infect *Escherichia coli* strain Y1090, and duplicate nitrocellulose filters were prepared from each plate (12). Hybridization with the 72-mer probe and washing of the filters was performed as described (13). cDNA inserts were recovered from the positive clones by *Eco*RI digestion. The human genomic DNA library in λ Charon 4A phage was kindly provided to us by T. Maniatis (Harvard University, Cambridge, MA).

DNA Sequence Analysis. cDNA inserts or genomic DNA fragments were subcloned into M13mp18 and mp19 (14) and sequenced by the dideoxy chain-termination method (15). The universal 17-base primer and the DNA sequencing kit were purchased from Bethesda Research Laboratories. Other oligonucleotides used as primers for sequencing were synthesized on a 380B Applied Biosystems DNA synthesizer. The Cyclone subcloning system (International Biotechnologies, New Haven, CT) was used for the production of deletion mutants. ATP[^{35}S] (600 Ci/mmol) used for the sequencing was from Amersham. Each region of the cDNA was sequenced at least three times.

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: ORF, open reading frame; AD, Alzheimer disease; DS, Down syndrome.

*A preliminary report of this work was presented at the Bunbury conference on the *Neurochemistry of Aging*, Dec. 2, 1986, Cold Spring Harbor, NY.

Southern Transfers. Human brains obtained at autopsy were frozen at −70°C. High molecular weight DNA was prepared from human, hamster, mouse, sheep, and rabbit brains (12). The DNA was digested with the appropriate restriction endonucleases, electrophoresed on a 1% agarose gel, and then transferred to nitrocellulose filters (16). The cDNA insert present in clone B2.3 (Fig. 1) was labeled at a specific activity of 2×10^9 cpm/μg using [α-^{32}P]ATP (3000 Ci/mmol) purchased from New England Nuclear, and the oligonucleotide labeling kit was from Pharmacia. Filters were hybridized to the probe as described (12) in the presence of 50% (vol/vol) formamide at 42°C and then washed in 15 mM NaCl/1.5 mM sodium citrate, pH 7.0/0.1% $NaDodSO_4$ at 65°C.

RNA Blots. RNA was prepared from human, hamster, mouse, and sheep brains as described (17). All human material used was frozen no later than 6 hr postmortem. Poly(A)$^+$ RNA was isolated from total RNA (18) and stored in liquid nitrogen. RNA samples were electrophoresed on a formaldehyde gel and then transferred to nitrocellulose filters (19). The filters were hybridized to the labeled B2.3 cDNA insert and then washed as described above.

RESULTS

To identify the nucleic acid sequences encoding the β protein, we used low-stringency hybridization with long synthetic oligonucleotides. This methodology has been used (17) to isolate the cDNA encoding the hamster prion protein. The probe was synthesized on the basis of the amino acid sequence of the amyloid peptide.

In view of the uncertainty regarding the site of synthesis of the amyloid peptide and assuming that this protein is a human gene product, we first decided to search a complete human genomic DNA library contained in λ Charon 4A phage (20). We screened 6×10^5 recombinant phage from this library and isolated 26 positive clones. By increasing the washing temperature, the number of hybridizing clones was reduced to 7 clones. Three of these clones were characterized, and, although they were found to contain regions of high homology with our probe, none of them encoded the amyloid peptide (21). Therefore, we decided to screen cDNA libraries constructed from poly(A)$^+$ mRNA isolated from different human tissues. Using the 72-mer oligonucleotides as probe, we screened 2×10^5 recombinants and isolated 3 clones from a brain cDNA library. Restriction analysis showed that clone B1.1 contained a cDNA insert 450 base pairs (bp) long (Fig. 1). Sequencing of this fragment revealed a region that encodes the 26 carboxyl-terminal amino acids of the β protein. This region is part of an open reading frame (ORF) that encodes 97 amino acids and terminates with a TAG codon. The amyloid peptide sequence is located at the 5′ end of this ORF. There is perfect agreement between the amino acid sequence predicted from this clone and the sequence of the cerebrovascular amyloid isolated from a DS patient (3). The insert present in clone B2.3 shares the same 5′ end with clone B1.1 and extends another 350 bp from the 3′ end of this clone. Clone B2.1 contains a 1.7-kilobase (kb) insert that upon *Eco*RI digestion yields a 650-bp fragment that shares the same 5′ end with clones B1.1 and B2.3, plus another fragment about 1 kb long located upstream from the 650-bp fragment. The sequence of the 1-kb fragment had a long ORF that at its 3′ end contained the codons for the two amino-terminal amino acids present in the amyloid peptide. To obtain the sequence overlapping the *Eco*RI site, we subcloned and sequenced the 1.1-kb *Rsa* I fragment from clone B2.1 (Fig. 1). Fig. 2 shows the nucleotide sequence obtained from clones B1.1, B2.1, and B2.3 and the deduced amino acid sequence. All regions of the cDNA inserts were sequenced at least three times. The presence of an *Eco*RI site at position 956 of the B2.1 clone suggests that clones B1.1 and B2.3 were derived from the incomplete methylation and subsequent digestion of the cDNA by the enzyme *Eco*RI used to construct this library. At present it is not clear whether the *Eco*RI site at the 5′ end of clone B2.1 is derived from the linkers used to prepare the library or is part of a cDNA that extends upstream from this site. The ORF present in clone B2.1 extends from nucleotide 9 to nucleotide 1245 and encodes 412 amino acids. The first in-frame ATG codon is at nucleotide 151, and the TAG stop codon is at position 1247. A protein encoded by the sequence from position 151 to position 1247 would have 365 amino acids and a molecular weight of 40,200. However, since the size of the mRNA transcript detected by this cDNA is 3.3 kb (see below), it is likely that there is another in-frame ATG codon further upstream in an uncloned region. There are stop codons in all three reading frames downstream from position 1247. It is interesting to note that there are no cysteine residues encoded by this ORF. A computer search of the GenBank* revealed that this cDNA

*National Institutes of Health (1986) Genetic Sequence Data Bank: GenBank (Research Systems Div., Bolt, Beranek, and Newman, Cambridge, MA), Tape Release 46.

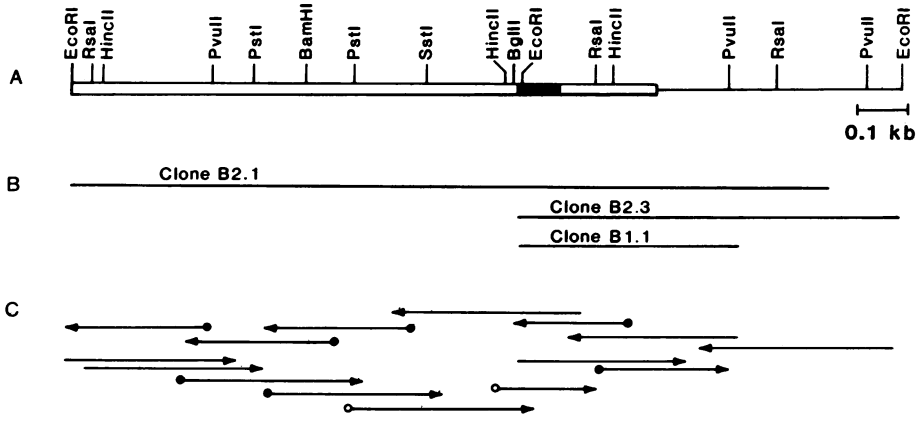

FIG. 1. Physical map of the cDNAs in clones B2.1, B2.3, and B1.1. (*A*) Restriction map of the cDNA inserts. Coding region, long rectangle. Single line, 3′-noncoding region. Filled box, region encoding the amyloid peptide. (*B*) Schematic representation of the three λgt11 recombinant clones isolated from the human brain cDNA library. (*C*) Nucleotide sequencing strategy. The direction and length of each sequence determination is indicated by an arrow. To sequence past the internal *Eco*RI site the 1.1-kb *Rsa* I fragment was subcloned into M13. Arrows with open-circle ends, sequencing of deletion clones produced by means of the Cyclone subcloning system (International Biotechnologies). Arrows with solid-circle ends, sequencing with synthetic oligonucleotides as primers. All other sequencing reactions were performed with the universal primer (Bethesda Research Laboratories).

```
                Glu Glu Val Val Arg Val Pro Thr Thr Ala Ala Ser Thr Pro Asp Ala Val Asp Lys Tyr   20
      GAATTCCGG GAA GAG GTG GTT CGA GTT CCT ACA ACA GCA GCC AGT ACC CCT GAT GCC GTT GAC AAG TAT   69

   Leu Glu Thr Pro Gly Asp Glu Asn Glu His Ala His Phe Gln Lys Ala Lys Glu Arg Leu Glu Ala Lys   43
   CTC GAG ACA CCT GGG GAT GAG AAT GAA CAT GCC CAT TTC CAG AAA GCC AAA GAG AGG CTT GAG GCC AAG  138

   His Arg Glu Arg Met Ser Gln Val Met Arg Glu Trp Glu Glu Ala Glu Arg Gln Ala Lys Asn Leu Pro   66
   CAC CGA GAG AGA ATG TCC CAG GTC ATG AGA GAA TGG GAA GAG GCA GAA CGT CAA GCA AAG AAC TTG CCT  207

   Lys Ala Asp Lys Lys Ala Val Ile Gln His Phe Gln Glu Lys Val Glu Ser Leu Glu Gln Glu Ala Ala   89
   AAA GCT GAT AAG AAG GCA GTT ATC CAG CAT TTC CAG GAG AAA GTG GAA TCT TTG GAA CAG GAA GCA GCC  276

   Asn Glu Arg Gln Gln Leu Val Glu Thr His Met Ala Arg Val Glu Ala Met Leu Asn Asp Arg Arg Arg  112
   AAC GAG AGA CAG CAG CTG GTG GAG ACA CAC ATG GCC AGA GTG GAA GCC ATG CTC AAT GAC CGC CGC CGC  345

   Leu Ala Leu Glu Asn Tyr Ile Thr Ala Leu Gln Ala Val Pro Pro Arg Pro Arg His Val Phe Asn Met  135
   CTG GCC CTG GAG AAC TAC ATC ACC GCT CTG CAG GCT GTT CCT CCT CGG CCT CGT CAC GTG TTC AAT ATG  414

   Leu Lys Lys Tyr Val Arg Ala Glu Gln Lys Asp Arg Gln His Thr Leu Lys His Phe Glu His Val Arg  158
   CTA AAG AAG TAT GTC CGC GCA GAA CAG AAG GAC AGA CAG CAC ACC CTA AAG CAT TTC GAG CAT GTG CGC  483

   Met Val Asp Pro Lys Lys Ala Ala Gln Ile Arg Ser Gln Val Met Thr His Leu Arg Val Ile Tyr Glu  181
   ATG GTG GAT CCC AAG AAA GCC GCT CAG ATC CGG TCC CAG GTT ATG ACA CAC CTC CGT GTG ATT TAT GAG  552

   Arg Met Asn Gln Ser Leu Ser Leu Leu Tyr Asn Val Pro Ala Val Ala Glu Glu Ile Gln Asp Glu Val  204
   CGC ATG AAT CAG TCT CTC TCC CTG CTC TAC AAC GTG CCT GCA GTG GCC GAG GAG ATT CAG GAT GAA GTT  621

   Asp Glu Leu Leu Gln Lys Glu Gln Asn Tyr Ser Asp Asp Val Leu Ala Asn Met Ile Ser Glu Pro Arg  227
   GAT GAG CTG CTT CAG AAA GAG CAA AAC TAT TCA GAT GAC GTC TTG GCC AAC ATG ATT AGT GAA CCA AGG  690

   Ile Ser Tyr Gly Asn Asp Ala Leu Met Pro Ser Leu Thr Glu Thr Lys Thr Thr Val Glu Leu Leu Pro  250
   ATC AGT TAC GGA AAC GAT GCT CTC ATG CCA TCT TTG ACC GAA ACG AAA ACC ACC GTG GAG CTC CTT CCC  759

   Val Asn Gly Glu Phe Ser Leu Asp Asp Leu Gln Pro Trp His Ser Phe Gly Ala Asp Ser Val Pro Ala  273
   GTG AAT GGA GAG TTC AGC CTG GAC GAT CTC CAG CCG TGG CAT TCT TTT GGG GCT GAC TCT GTG CCA GCC  828

   Asn Thr Glu Asn Glu Val Glu Pro Val Asp Ala Arg Pro Ala Ala Asp Arg Gly Leu Thr Thr Arg Pro  296
   AAC ACA GAA AAC GAA GTT GAG CCT GTT GAT GCC CGC CCT GCT GCC GAG CGA GGA CTG ACC ACT CGA CCA  897

   Gly Ser Gly Leu Thr Asn Ile Lys Thr Glu Glu Ile Ser Glu Val Lys Met Asp Ala Glu Phe Arg His  319
   GGT TCT GGG TTG ACA AAT ATC AAG ACG GAG GAG ATC TCT GAA GTG AAG ATG GAT GCA▲GAA TTC CGA CAT  966

   Asp Ser Gly Tyr Glu Val His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly  342
   GAC TCA GGA TAT GAA GTT CAT CAT CAA AAA TTG GTG TTC TTT GCA GAA GAT GTG GGT TCA AAC AAA GGT 1035

   Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val Ile Val Ile Thr Leu Val Met Leu  365
   GCA ATC ATT GGA CTC ATG GTG GGC GGT GTT GTC ATA GCG ACA GTG ATC GTC ATC ACC TTG GTG ATG CTG 1104

   Lys Lys Lys Gln Tyr Thr Ser Ile His His Gly Val Val Glu Val Asp Ala Ala Ala Val Thr Pro Glu Glu  388
   AAG AAG AAA CAG TAC ACA TCC ATT CAT CAT GGT GTG GTG GAG GTT GAC GCC GCT GTC ACC CCA GAG GAG 1173

   Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu Gln Met Gln  411
   CGC CAC CTG TCC AAG ATG CAG CAG AAC GGC TAC GAA AAT CCA ACC TAC AAG TTC TTT GAG CAG ATG CAG 1242

   Asn End                                                                                      412
   AAC TAG ACCCCCGCCACAGCAGCCTCTGAAGTTGGACAGCAAAACCATTGCTTCACTACCCATCGGTGTCCATTTATAGAATAATGTGG 1331
   GAAGAAACAAACCCGTTTTATGATTTACTCATTATCGCCTTTTGACAGCTGTGCTGTAACACAAGTAGATGCCTGAACTTGAATTAATCCA 1422
   CACATCAGTAATGTATTCTATCTCTCTTTACATTTTGGTCTCTATACTACATTATTAATGGGTTTTGTGTACTGTAAAGAATTTAGCTGTA 1513
   TCAAACTAGTGCATGAATAGATTCTCTCCTGATTATTTATCACATAGCCCCTTATGCCAGTTGTATATTATTCTTGTGGTTTGTGACCCAA 1604
   TTAAGTCCTACTTTACATATGTTTAAGAATCGATGGGGGATGCTTCATGTGAACGTGGGAGTTCAGCTGCTTCTCTTGCCTAAGTATTCCT 1695
   TTCCTGATCACTATGCATTTTAAAGTTAAACATTTTCG                                                       1733
```

FIG. 2. Nucleotide sequence obtained from clones B2.1, B2.3, and B1.1 and the deduced amino acid sequence. Upper number, amino acids; lower number, nucleotides. β Protein sequence, overlined. Three potential N-linked glycosylation sites, boxed. Hydrophobic region, overlined with a broken line. Arrowhead, 5′ end of clones B2.3 and B1.1.

clone contains no significant homologies with any known sequences including the gene encoding the prion protein that forms the amyloid fibers observed in the unconventional slow virus infections (22, 23). There are three potential N-linked glycosylation sites (Asn-Xaa-Thr/Ser) at positions 184, 213, and 401 (Fig. 2). The presence of proline at position 402 makes it unlikely that Asn-401 is glycosylated. A hydropathy plot of the deduced amino acid sequence obtained by the method of Kyte and Doolittle (24) is presented in Fig. 3. This protein is largely hydrophilic with a highly hydrophobic region near the carboxyl terminus. Regions of proteins with 19 or more amino acids and with hydropathy averages >1.6 are usually tightly associated with lipid and often span the membrane in an α-helical configuration (24). The stretch of 23 hydrophobic residues near the carboxyl terminus (Figs. 2 and 3) has an average hydropathy value of 2.8 and is, therefore, most probably the transmembrane region of this protein. The predicted amino acid sequence from residue 37 to residue 147 contains six pairs of amino acids Xaa-Arg where Xaa is an acidic amino acid. Amino acids 366–412 probably represent the cytoplasmic part of this protein. This region contains the same number of acidic and basic residues.

Southern Blot Analysis. We used the cDNA insert present in clone B2.3 as a probe to analyze human genomic DNA isolated postmortem from the brains of a DS patient, an AD patient, and a person with no history of neurological disorders. When these DNA samples were digested with restriction endonucleases *Eco*RI and *Pst* I and analyzed by Southern blot hybridization, all samples yielded identical hybrid-

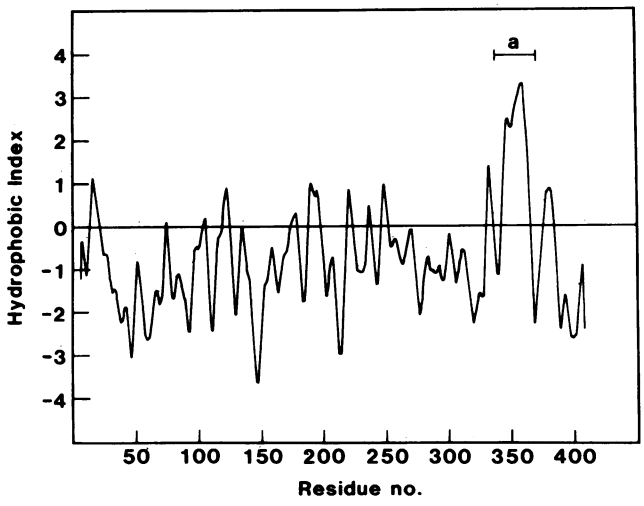

FIG. 3. Hydropathy plot of the β-protein precursor amino acid sequence. Hydropathy values (24) for a window of 11 amino acid residues were averaged, assigned to the middle residue of the span, and plotted with respect to position along the amino acid sequence. Bar, proposed transmembrane region. Positive numbers, hydrophobicity; negative numbers, hydrophilicity.

ization patterns (Fig. 4). Southern blots of DNA isolated from WA17, a mouse–human hybrid cell line with three copies of human chromosome 21 per cell, as the only human chromosomes present (W. T. Brown and R. Dutkowski, personal communication), were probed with our cDNA clones and revealed the presence of hybridizing fragments produced by two different restriction enzymes identical in mobility to the hybridizing fragments produced from similar digests of the human genomic DNA (Fig. 4). Similar results were obtained from the analysis of DNA isolated from 2FU^{r}1, a Chinese hamster–human hybrid cell line that contains the long arm of chromosome 21 as the only human DNA (25). These data are in agreement with *in situ* hybridization experiments where it was shown that the amyloid genomic sequences are located on human chromosome 21q21 (26). Since there is neither an *Eco*RI nor a *Pst* I site in the B2.3 cDNA insert used as a probe, the presence of two hybridizing fragments suggests that either there are intervening sequences in the corresponding genomic locus that encodes this cDNA or that there are two copies of a sequence homologous to the probe on this chromosome. Using stringent conditions of hybridization, we detected the presence of homologous sequences in the genomic DNA of rabbit, sheep, hamster, and mouse suggesting that this gene has been highly conserved among mammalian species (Fig. 4). Hybridization of the filters under reduced stringency [20% (vol/vol) formamide, 45°C] and washing at 45°C in 15 mM NaCl showed additional hybridizing fragments in the human, mouse, hamster, sheep, and rabbit DNA indicating the presence of an amyloid multigene family in these species (data not shown).

RNA Blot Analysis. RNA blot transfer hybridization experiments revealed a 3.3-kb species present in RNA samples prepared from postmortem human brain. This mRNA seems to be present at similar levels throughout the human cortex (Fig. 5). No difference in the amount of the amyloid mRNA was detected between preparations from a 70-year-old AD patient and from an age-matched patient with no apparent neurological disorders (Fig. 5). A similar transcript was present in a sample obtained postmortem from a DS patient.

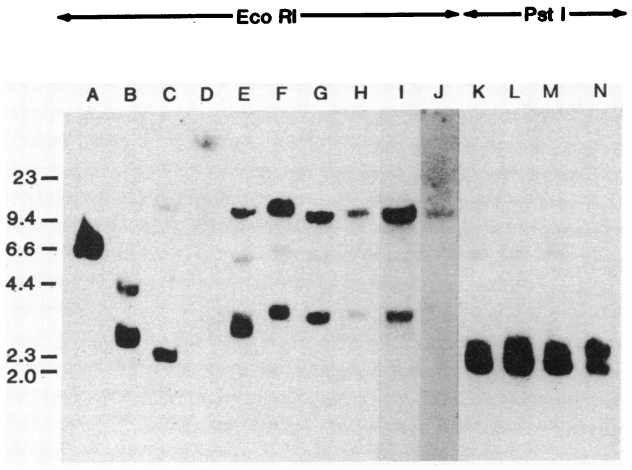

FIG. 4. Southern blot analysis of genomic DNA. DNA isolated from rabbit, sheep, Syrian hamster, and mouse brains, respectively (lanes A, B, C, and D). DNA isolated from the mouse–human cell hybrid WA17 (lanes E and N). DNA prepared postmortem from the brain of a DS patient (lanes F and K). DNA prepared from a patient with no apparent neurological disorders (lanes G and L). DNA isolated from the brain of an AD patient (lanes H and M). DNA from the Chinese hamster–human cell line 2 FU^{r}1 (lane I). Chinese hamster DNA (lane J). EcoRI; DNA digested with restriction endonuclease *Eco*RI. Pst I; DNA digested with restriction endonuclease *Pst* I. DNA (8 μg) was loaded on each lane. Incomplete digestion is observed in lane D.

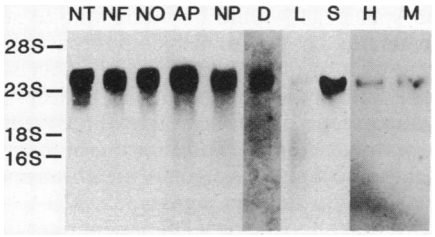

FIG. 5. RNA blot analysis. RNA isolated postmortem from the temporal, frontal, and occipital cortex, respectively, from the brain of a patient with no apparent neurological disorders (lanes NT, NF, and NO). RNA isolated from the parietal cortex of an AD victim and an age-matched victim of heart attack, respectively (lanes AP and NP). RNA isolated from the frontal cortex of a DS patient (lane D). RNA isolated from the liver of a patient with no apparent neurological disorders (lane L). For an explanation for the weak signal in this lane see text. RNA isolated from sheep, hamster, and mouse brains, respectively (lanes S, H, and M). The weak signal in lanes H and M is probably due to partial degradation of the poly(A)$^+$ RNA used. Total RNA (30 μg) was used for lanes NT, NF, NO, AP, NP, L, and S. Poly(A)$^+$ RNA (5 μg) was used in lanes D, H, and M.

A homologous RNA species was detected at low levels in two different preparations from a human liver specimen (Fig. 5). However, due to difficulties in obtaining human tissue, we were unable to determine whether this reflects a low level of expression of this mRNA in the liver or is the result of the condition of the tissue. The expression of the amyloid peptide gene in the liver is consistent with the immunological detection of this antigen in the blood (J. Currie and K. S. Kim, personal communication). In agreement with the genomic DNA hybridization experiments homologous RNA transcripts were detected in poly(A)$^+$ RNA isolated from the brains of sheep, hamster, and mouse (Fig. 5).

DISCUSSION

We have isolated and characterized a human brain cDNA encoding the β protein. This clone contains a long ORF that defines a 412-residue polypeptide. The 28-amino acid sequence of the β protein is located near the carboxyl terminus of this ORF, and it is closely followed by a region encoding a sequence of highly hydrophobic residues. These results suggest that β protein is cleaved posttranslationally from a larger precursor whose physiological function is unknown. RNA gel blot hybridization experiments using this cDNA as a probe indicate that the 4-kDa β protein is encoded by a ≈3.3-kb mRNA.

There are several similarities between the β-protein precursor and the precursor of the prion proteins associated with the unconventional slow virus infections (27). Both proteins are largely hydrophilic with a highly hydrophobic region at or near the carboxyl terminus (17). Each contains a sequence that after cleavage from its precursor has the capacity of forming amyloid fibers in the brain (22, 28). These amyloid-forming sequences are closely followed by hydrophobic regions that are most likely used as transmembrane-spanning domains (24). Finally, although the amyloid peptides can accumulate in the brain at high levels, the amounts of the corresponding mRNAs vary much less (17). It should be noted, however, that the human prion protein gene and the AAP gene are located on different chromosomes (29).

It is interesting that at least two other proteins, human transforming growth factor and epidermal growth factor, are cleaved from their precursor peptides at regions closely followed by transmembrane sequences (30). In unconventional slow virus infections (22) the human and the mouse prion proteins form amyloid fibers. Like the prion protein gene, the gene encoding β protein is conserved among

mammalian species. It would be interesting to see whether the protein encoded by the mouse counterpart of the human amyloid gene is also amyloidogenic and under what conditions this amyloid would be formed. The hydrophobic region of the amyloid precursor is followed by three lysines (Fig. 3). Positively charged sequences often occur at the cytoplasmic domain-membrane junctions of many membrane proteins and probably act as stop-transfer signals (31, 32).

It is known that almost every DS patient over the age of 40 develops neuritic plaques and deposits of cerebrovascular amyloid (1). The presence of the β-protein gene on chromosome 21 suggests a causal relationship between the triple gene-dosage effect seen in trisomy 21 and AD neuropathology. The factor(s) leading to the formation of the amyloidogenic protein and its aggregation into amyloid fibers in the brains of AD and DS patients are not known. We also do not know whether the structure of the amyloid peptide precursor is exactly the same in all people. Point mutations in the amyloid gene could result in changes in the primary structure of the amyloid precursor protein and this change could in turn interfere with the normal processing mechanism of this protein. Alternatively it is conceivable that in AD and DS patients the enzyme(s) involved in the processing of the amyloid precursor is defective. Both situations could lead to the accumulation of abnormally high levels of the amyloid peptide in the brain of these patients and its subsequent aggregation into amyloid fibers. This hypothesis is in agreement with the observation that cells of the brain reticuloendothelial system (pericytes, microglia, and macrophages) are associated with the formation of amyloid fibers (33). Amyloidosis indicates a failure of this system to fully degrade the amyloidogenic protein. A defect in the reticuloendothelial system would also be consistent with the observed clustering of some AD cases in certain families (8).

At this point it is not clear what the relationship is between the amyloid deposition and the neuritic dystrophy. Ultrastructural studies (33) suggest that the first changes appear to be related to the hypersecretory activity of macrophages, microglia, and pericytes that release a proteinaceous material into the extracellular space where the assembly of the amyloid fibers takes place. This supports the theory that neuritic abnormalities are a consequence of the amyloid plaque formation. The availability of a cDNA clone encoding the β-protein precursor should help to elucidate the mechanism of the formation of the amyloid plaque in the brains of AD and DS patients.

Note. After the submission of this paper, two other groups (34, 35) independently reported the isolation and sequencing of human cDNA clones encoding the amyloid peptide precursors. A third group (36) reported the detection of the amyloid precursor mRNA in 11 different human tissues.

We thank Yu-Wen Hwang and Giuseppe LaFauci for their help with the synthesis of the oligonucleotides and sequencing gels, D. Soifer for his help with the computer search, D. Bolton for his help with the hydrophobic plot, T. W. Brown for providing the WA17 DNA, S. Antonarakis for providing the 2Fu^{r}1 DNA, and D. Miller for helpful discussions. This work was supported in part by Grant AG04220 from the National Institutes of Health.

1. Wisniewski, K. E., Dalton, A. J., Crapper McLachlan, D. R., Wen, G. Y. & Wisniewski, H. M. (1985) *Neurology* **35,** 957–961.
2. Glenner, G. G. (1980) *New Engl. J. Med.* **302,** 1283–1292.
3. Glenner, G. G. & Wong, C. W. (1984) *Biochem. Biophys. Res. Commun.* **120,** 885–890.
4. Wong, C. W., Quaranta, V. & Glenner, G. G. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 8729–8732.
5. Masters, C. L., Jimms, G., Weinman, N. A., Beyreuther, K., Multhaup, G. & McDonald, L. B. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 4245–4249.
6. Masters, C. L., Multhaup, G., Simms, G., Jutton, P., Martins, R. N. & Beyreuther, K. (1985) *EMBO J.* **4,** 2757–2763.
7. Grundke-Iqbal, I., Iqbal, K., Quinlan, M., Tung, Y. C., Zaidi, M. S. & Wisniewski, H. M. (1985) *J. Biol. Chem.* **261,** 6084–6089.
8. Heston, L. L. & White, J. (1978) *Behav. Genet.* **8,** 315–331.
9. Grantham, C., Gautier, M. J. & Mercier, R. (1981) *Nucleic Acids Res.* **9,** r43–r74.
10. Bird, A. P. (1980) *Nucleic Acids Res.* **1,** 1499–1504.
11. Aganwal, K. K., Brunsjedt, J. & Noyes, B. E. (1981) *J. Biol. Chem.* **256,** 1023–1028.
12. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
13. Ullrich, A., Berman, C. H., Dull, T. J., Gray, A. & Lee, J. M. (1984) *EMBO J.* **3,** 361–364.
14. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9,** 309–321.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
16. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
17. Robakis, N. K., Sawh, P. R., Wolfe, G. C., Rubenstein, R., Carp, R. I. & Innis, M. A. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 6377–6381.
18. Aviv, H. & Leder, P. (1982) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
19. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
20. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. (1978) *Cell* **15,** 1157–1174.
21. Robakis, N. K., Wolfe, G. Ramakrishna, N. & Wisniewski, H. M. (1986) in *Neurochemistry of Aging*, Banbury Conference Proceedings, eds. Finch, C. & Davies, P. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), in press.
22. Merz, P. A., Rohwer, R. G., Kascsak, R., Wisniewski, H. M., Somerville, R. A., Gibbs, C. J., Jr., & Gajdusek, P. C. (1984) *Science* **225,** 437–440.
23. Prusiner, S. B., McKinley, M. P., Bowman, K. A., Bolton, D. C., Bendheim, P. E., Groth, D. P. & Glenner, G. G. (1983) *Cell* **35,** 349–358.
24. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157,** 105–132.
25. Vankeuren, M. L., Watkins, P. C., Drabkin, H. A., Jabs, E. W., Gusella, J. F. & Patterson, D. (1986) *Am. J. Hum. Genet.* **38,** 793–804.
26. Robakis, N. K., Wisniewski, H. M., Jenkins, E. C., Devine-Gage, E. A., Houck, G. E., Yao, X. L., Ramakrishna, N., Wolfe, G., Silverman, W. P. & Brown, T. W. (1987) *Lancet* **i,** 384–385.
27. Manuelidis, L. & Manuelidis, E. E. (1986) *Prog. Med. Virol.* **33,** 78–79.
28. Oesch, B., Westaway, D., Wulchli, M., McKinley, M. P., Kent, S. B. H., Aebersold, R., Barry, R. A., Tempst, P., Teplow, D. B., Hood, L. E., Prusiner, S. B. & Weismann, C. (1985) *Cell* **40,** 735–746.
29. Robakis, N. K., Devine-Gage, E. A., Jenkins, E. C., Kascsak, R. J., Brown, T. W., Krawczun, M. S. & Silverman, W. P. (1986) *Biochem. Biophys. Res. Commun.* **140,** 758–765.
30. Derynck, R., Roberts, A. B., Winkler, M. E., Chen, E. Y. & Goeddel, D. V. (1984) *Cell* **38,** 287–297.
31. Blobel, G. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 1496–1500.
32. Von Heijne, G. (1980) *Eur. J. Biochem.* **120,** 275–278.
33. Wisniewski, H. M., Moretz, R. C. & Lossinsky, A. J. (1981) *Ann. Neurol.* **10,** 517–522.
34. Goldgaber, D., Lerman, M. I., McBride, O. W., Saffioti, U. & Gajdusek, D. C. (1987) *Science* **235,** 877–880.
35. Kang, J., Lemaire, H.-G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, G., Multhaup, G., Beyreuther, K. & Müller-Hill, B. (1987) *Nature (London)* **325,** 733–736.
36. Tanzi, R. E., Gusella, J. F., Watkins, P. C., Bruns, G. A. P., St. George-Hyslop, P., Van Keuren, M. L., Patterson, D., Pagan, S., Kurnit, D. M. & Neve, R. L. (1987) *Science* **235,** 880–884.