MOLECULAR AND CELLULAR BIOLOGY, Apr. 1989, p. 1381–1388
0270-7306/89/041381-08$02.00/0

Vol. 9, No. 4

Tau Consists of a Set of Proteins with Repeated C-Terminal Microtubule-Binding Domains and Variable N-Terminal Domains

ADOLF HIMMLER,[1]† DAVID DRECHSEL,[2] MARC W. KIRSCHNER,[2] AND DAVID W. MARTIN, JR.[1]*

Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080,[1] and Department of Biochemistry and Biophysics, School of Medicine, University of California, San Francisco, California 94143[2]

Received 17 October 1988/Accepted 29 December 1988

Tau proteins consist of a family of proteins, heterogeneous in size, which associate with microtubules in vivo and are induced during neurite outgrowth. In humans, tau is one of the major components of the pathognomonic neurofibrillary tangles in Alzheimer's disease brain. Screening of a cDNA library prepared from bovine brain led to the isolation of several cDNA clones encoding tau proteins with different N termini and differing by insertions or deletions, suggesting differential splicing of the tau transcripts. One of the N-terminal domains and the repeated C-terminal domain of the encoded tau proteins are recognized by polyclonal antibodies to bovine tau. The bovine tau proteins are highly homologous to murine and human tau, especially within the repeated C-terminal domain. Compared with murine and human tau, bovine tau contains the insertion of three longer segments, one of which is an additional characteristic repeat. Portions of tau proteins generated by in vitro translation were used to show that these repeats represent tubulin-binding domains, two of which are sufficient to bind to microtubules assembled from purified tubulin in the presence of taxol.

The primary function of microtubules is to generate specific cell morphologies and to organize the intracellular arrangements of cytoplasmic components such as vesicles, nuclei, or mitochondria (11). It has been suggested that the novel dynamics and lability of individual microtubules in cells play an important role in establishing the initial morphologies (21). As demonstrated by in vitro experiments with pure tubulin subunits, microtubules have intrinsic dynamic properties, but these extreme dynamics may not be appropriate to cells, such as mature neurons, that have established definitive morphologies (19, 39). For this reason, one might expect such cells to contain proteins that are capable of binding to the tubulin polymer and thereby modulate the dynamic properties of this polymer. Similarly, pure tubulin polymers probably have a restricted capacity to interact directly with cellular constituents, and therefore there might also exist adapter or cross-linking proteins that interact with tubulin on one of their domains and with other cytoplasmic components on another domain, analogous to the adapter function of transfer RNA.

For these two reasons, there has been considerable interest in proteins that copurify with microtubules, that is, microtubule-associated proteins (MAPs) (32). In the vertebrate brain, the most abundant form of MAPs is represented by the high-molecular-weight MAPs, which consist of a disparate group of proteins that includes MAP2, the brain protein dynein, MAP1c, and others (33). The other major class of MAPs is called tau, a 40- to 60-kilodalton (kDa) group of related proteins that are induced (along with MAP1) during neurite extension and that stabilize microtubules when injected into fibroblast cells (8–10). Tau is present in different forms that are distinguishable by denaturing gel electrophoresis. These forms change markedly during embryonic development (7). Tau shares epitopes with paired helical filaments, a major constituent of the cytopathological neurofibrillary tangles, which are found at autopsy in brains from patients with Alzheimer's disease (14, 20, 23, 31, 42).

Although numerous cell biological and biophysical studies of MAP proteins have begun to define the possible roles of these proteins in controlling microtubule assembly and have hinted at a role in mediating interactions with other components, the lack of detailed structural information has limited our understanding. We do not understand whether there are specialized regions of these molecules that act as tubulin receptors, and in the case of tau protein we do not yet understand the basis for the similarities and heterogeneities in that group. Lee et al. (25) have shown that two cDNAs for mouse tau encode heterogeneity in the C-terminal domain, suggesting that alternative RNA splicing may contribute to the heterogeneity. They proposed that reiterated structures in the C-terminal domain might be related to tubulin-binding sites. We now show directly by gene expression and protein coassembly that the repeated domains of 31 amino acids in bovine tau are tubulin-binding domains. We also show that tau is a tripartite molecule composed of a C-terminal tubulin-binding region with different numbers of repeats, a constant middle domain, and variable N-terminal domains. The heterogeneity of the N-terminal domains has implications for the role of tau in mediating interactions with other components in the cell.

MATERIALS AND METHODS

Enzyme reactions. Enzyme reactions were carried out according to the suggestions of the manufacturer unless otherwise stated.

Isolation and characterization of bovine tau cDNA clones. A total of 750,000 clones of a bovine brain cDNA library in λgt10 (38) were screened (3). The ^{32}P-labeled hybridization probe was generated from the 0.8-kilobase (kb) *Bam*HI-*Bgl*II DNA fragment of mouse tau cDNA clone pTA2 (25) by the random priming method (12). Screening was carried out by standard procedures (28). Positive clones were plaque purified. *Eco*RI inserts of lambda clones were subcloned into Bluescript plasmids (Stratagene) and characterized by restriction endonuclease digestion, Southern blotting, and

* Corresponding author.
† Present address: Ernst Boehringer Institut, Dr. Boehringer Gasse 5-11, A-1121 Vienna, Austria.

DNA sequencing. Clone pBT43 was obtained by rescreening the library with a probe made from a 0.95-kb *Eco*RI-*Sma*I fragment from clone pBT7.

The cDNA clones were sequenced by the modified dideoxy-chain termination technique (4) and the use of wedge gels (5). A combination of M13mp18 and M13mp19 vectors was used, as well as universal or specifically synthesized oligonucleotide primers and nested deletion mutants generated in Bluescript vectors with exonuclease III and mung bean nuclease (Stratagene).

RNA extraction and Northern (RNA) blot analysis. RNA was isolated from fresh adult bovine brain by the low-temperature guanidine thiocyanate method (16) and passed through oligo(dT)-cellulose chromatography. Poly(A)$^+$ RNA was treated with glyoxal (28), separated by electrophoresis in an 0.8% agarose gel in 10 mM sodium phosphate buffer (pH 7.0), and subsequently transferred onto nitrocellulose filter in 20× SSC (1× SSC is 150 mM NaCl, 15 mM sodium citrate). End-labeled *Hin*dIII fragments of lambda DNA were glyoxalated under identical conditions and used as size markers. Hybridization was performed for 46 h at 42°C in 50% formamide, 5× SSC, 0.1% Ficoll (Pharmacia Fine Chemicals, Piscataway, N.J.), 0.1% bovine serum albumin, 0.1% polyvinylpyrrolidone, 10 mM Tris hydrochloride (pH 7.5), 0.5% sodium dodecyl sulfate (SDS), denatured salmon sperm DNA at 0.1 mg/ml. The DNA fragments were labeled by the random primer method. The oligonucleotide was end labeled with [α-^{32}P]dATP (Amersham Corp., Arlington Heights, Ill.) and terminal deoxynucleotide transferase (Bethesda Research Laboratories, Inc., Gaithersburg, Md.). The blots were washed three times for 20 min at room temperature in 2× SSC–0.1% SDS, followed by three washes over 30 min at 55°C in 0.1× SSC–0.1% SDS.

Preparation of templates for in vitro transcription. Plasmids containing tau cDNA sequences downstream of the T7 promoter to generate the sense strand of RNA were linearized with restriction enzymes 3′ of the coding sequence and electroeluted from agarose gels after electrophoresis. To increase the otherwise poor translation initiation from internal ATG codons, the portions upstream of the *Pst*I site preceding the first of the tau repeats of plasmids pBT4 and pBT20 were exchanged for the synthetic oligonucleotide 5′ - GATCCTCTGTCGACTATCAGGTGGGCCTTGAACC AGGATGCTGCA, representing the authentic 34 nucleotides upstream and the initiation codon of clone pBT43, flanked by compatible ends for ligation between the *Bam*HI site of the polylinker and the *Pst*I site near the first repeat. The resulting plasmids, pBT4BSMP and pBT20BSMP, allow translation of the C-terminal domain of tau starting with repeat 1. Templates encoding exclusively all four repeats (PCR [polymerase chain reaction]-4) or only two repeats (PCR-20) were generated from plasmid pBT4BSMP or pBT20BSMP by selectively amplifying this region by the PCR (37). Approximately 10 ng of plasmid DNA, linearized with *Xho*I 3′ of the coding region of pBT4BSMP or pBT20BSMP, was mixed with 50 pmol each of the two oligonucleotide primers 5′-ACGTTGTAAAACGACGGC CAGT, binding 5′ of the T7 promoter, and 5′-CTTTTTAT TCCCTCCGCCAGGG, binding to the opposite DNA strand at the end of repeat 4. After initial denaturation of the DNA for 10 min at 95°C, the DNA sequence between these primers was amplified in 16 cycles with 2.5 U of *Thermus aquaticus* DNA polymerase (New England BioLabs, Inc., Beverly, Mass.). The cycle times were 45 s at 55°C, 3 min at 68°C, and 50 s at 95°C. After the last cycle, ending with an incubation of 15 min at 68°C, the reaction products of 0.53-kb (PCR-4) and 0.36-kb (PCR-20) length were purified by electrophoresis in an agarose gel and subsequent electroelution.

In vitro transcription and translation. Approximately 1 μg of linear DNA template was used for in vitro transcription with 10 to 20 U of T7 RNA polymerase (Stratagene) and incubated for 60 min at 37°C. The DNA template was removed subsequently by treatment with DNase (RNase free; Promega Biotech, Madison, Wis.). The RNA was purified by extraction with phenol-chloroform and ethanol precipitation. RNA generated with T7 RNA polymerase was translated in a nuclease-treated rabbit reticulocyte lysate system (Promega Biotech) containing [^{35}S]methionine (Amersham) for 60 min at 30°C.

Cosedimentation of translation products with microtubules. Microtubules were formed in vitro from purified bovine brain tubulin (29) as follows. Tubulin (1 mg/ml) in polymerizing buffer {80 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] to KOH (pH 6.8), 1 mM $MgCl_2$, 1 mM GTP, 1 mM EGTA} was supplemented with taxol (gift of M. Suffness, Natural Products Division, National Cancer Institute, Bethesda, Md.) to 0.1 μg/ml and incubated for 5 min at 37°C. Taxol was added to 1.0 μg/ml, and the mixture was incubated at 37°C for 5 minutes. Taxol was added to 10 μg/ml, and the mixture was incubated for 10 min at 37°C. Then 3 μl of reticulocyte translation mixture was added to 50 μl of microtubule suspension, and the mixture was incubated for 20 min at 25°C. A 40-μl amount of this mixture was pelleted through an 80-μl sucrose cushion (50% sucrose, 10 μg of taxol per ml in polymerizing buffer) for 15 min at 140,000 × *g* in an Airfuge (Beckman Instruments, Inc., Fullerton, Calif.). A 30-μl portion of the upper layer was saved for gel electrophoresis; the remaining sucrose cushion was aspirated to the pellet. The pellet containing microtubules and microtubule-binding fragments was suspended in SDS sample buffer containing 3% 2-mercaptoethanol, boiled, and electrophoresed on 15% SDS-polyacrylamide gels (24).

Immunoprecipitation. The translation reactions (30 μl) were treated as described elsewhere (7) except that only 12 μl of the affinity-purified antiserum to bovine tau was used.

RESULTS

Characterization of tau cDNA clones. Using a partial cDNA clone for mouse tau as an initial hybridization probe and by rescreening, we identified five different clones in a cDNA library made from bovine brain mRNA. These clones fell into two categories with different sequences at the 5′ portion (Fig. 1). Assuming that clone pBT7 contained the translation initiation codon, we isolated two cDNA clones containing complete tau-coding sequences and constructed two more complete coding sequences from partial clones. In different cDNA clones, deletions occurred in each of the three domains, generating considerable diversity of primary structure at the level of the transcripts. Nucleotide sequence analyses of these clones revealed that three clones (pBT4, pBT7, and pBT12) had identical 3′ portions including a poly(A) tail but were heterogeneous at their 5′ ends. The largest clone, pBT7, had within its 5′ region a nucleotide sequence identical to that of clones pBT4 and pBT20 but contained a deletion of 251 base pairs (bp). Surprisingly, for pBT12 the 104 bases at the very 5′ end did not show homology to the other sequences, although the remaining 1.1 kb was identical to that in pBT7 and pBT4 and therefore represented a new N-terminal sequence (Fig. 2).

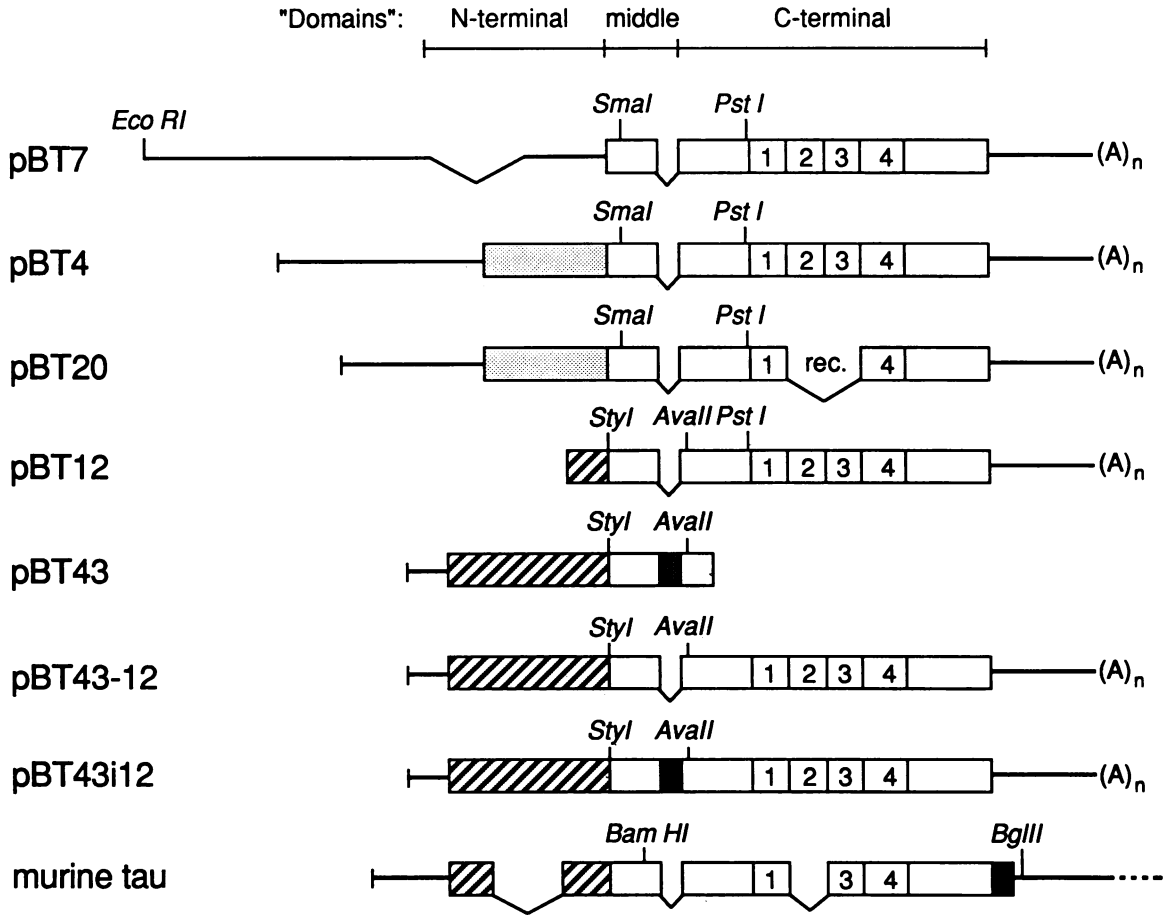

FIG. 1. Schematic map of tau cDNA clones, aligned to show their relationship. Regions that are predicted to encode proteins are drawn as bars. The divergent 5′ regions in a subset of clones are shaded to indicate the identical sequences within the identically shaded aligned regions. Locations of the four putative tubulin-binding repeats are indicated by integers 1 to 4. The restriction sites used for the generation of hybridization probes (Fig. 4) and for the production of truncated tau proteins (Fig. 5) are also indicated. The two types of murine tau cDNAs (25), differing only at the 3′ ends (solid bars), are aligned for best homology. The deletion marked rec. in clone pBT20 is assumed to be a recombination artifact, as described in the Discussion. Clone pBT43-12 was constructed by ligating the 5′ portion of clone pBT43 with the 3′ part of pBT12 at the *Sty*I site. pBT43i12, differing only by an in-frame insertion of 54 bp, was constructed by using the indicated *Ava*II site 3′ of the insertion (shaded) to ligate the fragments of pBT43 and pBT12 isolated after partial digestion. With the exception of pBT12, all tau cDNAs were inserted into the Bluescript plasmid in an orientation so that transcription starting at the T7 promoter generated the sense strand of RNA.

Rescreening the cDNA library with a probe derived from clone pBT7 led to the isolation of one more clone, pBT43, which did not contain a poly(A) tail but showed a nucleotide sequence identical to that of the distinct 5′ end of clone pBT12 and homology to the proposed translation initiation site of a mouse tau cDNA clone (25). However, pBT43 contained an insertion in the middle domain.

Clones pBT43 and pBT12, apparently derived from similar messages, were recombined to construct two variants of presumably complete tau-coding sequences. The resulting open reading frame in pBT43-12 predicted a polypeptide of 430 amino acids (44.4 kDa), and the other open reading frame in pBT43i12, including the in-frame insertion of 54 bp (Fig. 2), encoded 448 amino acids (46.3 kDa).

The open reading frames of clones pBT4, pBT20, and pBT7 encoded proteins of 402, 340, and 302 amino acids, respectively, having C-terminal portions identical to those of pBT12 but different N termini. Within the C-terminal portion were four contiguous repetitive stretches of 31 amino acids, denoted to 1 to 4, that were highly homologous (Fig. 3). The deletion in the 3′ portion of clone pBT20 maintains the reading frame, eliminating precisely the middle two of the four repeated peptide sequences.

Messenger RNA analysis. It was reported that two mouse tau mRNAs with different 3′ untranslated regions have similar lengths of approximately 6 kb (7, 25). In bovine tau, we found different-size mRNAs corresponding to the different transcripts (Fig. 4). The reconstructed clone pBT43i12, encoding the complete tau protein, produced a strong hybridization signal by Northern blots, with a message of 5 kb and a very weak signal at 9 kb. The fragment of clone pBT20 containing 5′ untranslated and coding regions for the alternative N-terminal domain hybridized to RNAs of 5, 8, and 9 kb and showed a higher relative ratio of the 9-kb to the 5-kb signal than did the previous probe. The similar probe derived from pBT7 containing a deletion of 251 bp hybridized to the 5- and 9-kb RNAs, whereas the 8-kb RNA was not detected.

A.

```
                                                                                                    1
                                                                                                   met ala glu pro
CGGCGCCCACCTTCTGCCGCCGCCACCACAGCCACTTTCTCCTCTGCCTCCCTCTACTGTCCTCGGCCCTCTGTCGACTATCAGGTGGGCCTTGAACCAGG   ATG GCT GAG CCC

                    10                                      20                                      30
arg gln glu phe asp val met glu asp his ala gln gly asp tyr thr leu gln asp gln glu gly asp met asp pro gly leu lys glu
CGC CAG GAG TTC GAC GTG ATG GAA GAT CAT GCT CAG GGG GAC TAC ACC CTG CAA GAC CAG GAG GGT GAC ATG GAC CCC GGC CTG AAA GAG

                    40                                      50                                      60
ser pro leu gln thr pro ala asp asp gly ser glu glu pro gly ser glu thr ser asp ala lys ser thr pro thr ala glu asp ala
TCT CCC CTG CAG ACC CCG GCC GAT GAT GGA TCT GAG GAA CCA GGC TCT GAA ACC TCT GAT GCT AAG AGC ACT CCG ACG GCG GAA GAT GCG

                    70                                      80                                      90
thr ala pro leu val asp glu gly ala pro gly glu gln ala ala ala gln ala pro ala glu ile pro glu gly thr ala ala glu glu
ACA GCA CCC TTA GTG GAT GAG GGA GCC CCC GGT GAG CAG GCG GCC GCT CAG GCC CCC GCG GAG ATC CCA GAA GGA ACC GCA GCT GAA GAA

                    100                                     110                                     120
ala gly ile gly asp thr ser asn leu glu asp gln ala ala gly his val thr gln ala arg met val ser lys gly lys asp gly thr
GCA GGC ATT GGC GAC ACG TCC AAC CTG GAA GAC CAA GCT GCC GGA CAC GTG ACC CAA GCT CGC ATG GTC AGT AAA GGC AAA GAT GGG ACT
             ■

                    130                                     140                                     150
gly pro asp asp lys lys thr lys gly ala asp gly lys pro gly thr lys ile ala thr pro arg gly ala ala pro pro gly gln lys
GGA CCC GAT GAC AAA AAA ACC AAG↑GGG GCG GAT GGT AAG CCT GGA ACG AAG ATT GCC ACA CCC CGG GGA GCA GCC CCT CCA GGC CAG AAA
                        StyI                                            SmaI
                    160                                     170                                     180
gly gln ala asn ala thr arg ile pro ala lys thr thr pro thr pro lys thr ser pro ala thr met gln val gln lys lys pro pro
GGC CAG GCC AAC GCC ACC CGG ATT CCA GCA AAA ACC ACT CCC ACC CCG AAG ACC TCG CCA GCA ACC ATG CAA GTG CAG AAA AAA CCA CCC
                                                                                ...................................

                    190                                     200                                     210
pro ala gly ala lys ser glu arg gly glu ser gly lys ser gly asp arg ser gly tyr ser ser pro gly ser pro gly thr pro gly
CCT GCA GGG GCA AAA TCT GAG AGA GGT GAA TCT GGG AAA TCC GGG GAC CGC AGC GGC TAC AGC AGC CCC GGC TCC CCA GGC ACT CCG GGC
................................

                    220                                     230                                     240
ser arg ser arg thr pro ser leu pro thr pro pro thr arg glu pro lys lys val ala val val arg thr pro pro lys ser pro ser
AGC CGC TCC CGC ACA CCC TCC CTG CCG ACC CCG CCC ACC CGG GAG CCC AAG AAG GTG GCG GTG GTC CGC ACT CCC CCC AAG TCG CCG TCT
     ●                                                                                                                PstI
                    250                                     260                                     270
ala ala lys ser arg leu gln ala ala pro gly pro met pro asp leu lys asn val lys ser lys ile gly ser thr glu asn leu lys
GCA GCC AAG AGC CGC CTG CAG GCC GCT CCC GGG CCC ATG CCA GAC CTG AAG AAC GTC AAG TCC AAA ATC GGC TCC ACG GAA AAC CTG AAG
                    PstI
                    280                                     290                                     300
his gln pro gly gly gly lys val gln ile ile asn lys lys leu asp leu ser asn val gln ser lys cys gly ser lys asp asn ile
CAC CAG CCA GGA GGT GGC AAG GTG CAG ATA ATT AAT AAG AAG CTG GAT CTT AGC AAC GTC CAG TCC AAG TGT GGC TCA AAG GAT AAT ATC
                        _______________________________________________________________________________________________

                    310                                     320                                     330
lys his val pro gly gly gly ser val gln ile val tyr lys pro val asp leu ser lys val thr ser lys cys gly ser leu gly asn
AAA CAC GTG CCA GGA GGC GGC AGT GTG CAA ATA GTC TAC AAA CCA GTG GAT CTG AGC AAG GTG ACC TCC AAG TGT GGC TCA TTA GGC AAC
_______________________________________________________________________________________________________________________

                    340                                     350                                     360
ile his his lys pro gly gly gly gln val glu val lys ser glu lys leu asp phe lys asp arg val gln ser lys ile gly ser leu
ATC CAT CAT AAG CCA GGA GGT GGC CAA GTG GAA GTG AAA TCT GAG AAG CTG GAC TTC AAG GAT AGA GTC CAG TCG AAG ATT GGG TCC CTG
_______________________________

                    370                                     380                                     390
asp asn ile thr his val pro gly gly gly asn lys lys ile glu thr his lys leu thr phe arg glu asn ala lys ala lys thr asp
GAC AAC ATC ACA CAC GTC CCT GGC GGA GGG AAT AAA AAG ATC GAA ACC CAC AAG CTG ACC TTC CGC GAG AAC GCC AAA GCC AAG ACC GAC

                    400                                     410                                     420
his gly ala glu ile val tyr lys ser pro val val ser gly asp thr ser pro arg his leu ser asn val ser ser thr gly ser ile
CAC GGG GCG GAG ATC GTG TAC AAG TCA CCG GTG GTG TCG GGG GAC ACT TCC CCC CGG CAC CTC AGC AAC GTG TCC TCC ACC GGC AGC ATC

                    430                                     440                             448
asp met val asp ser pro gln leu ala thr leu ala asp glu val ser ala ser leu ala lys gln gly leu OP
GAC ATG GTG GAC TCG CCG CAG CTC GCC ACC CTC GCT GAC GAG GTG TCC GCC TCC CTG GCC AAG CAG GGT TTG TGA TCAGGCCCCGGGGCGGTCAAT

AATCGTGGAGAGAAGAGAGTGAGAGTGTGGGAAAAAAAAAGAATAATGCTCTGGCCCTTCGCCCTCTGCCCTCCCCCAGCTGCTCCTCACAGATCGGTTCCTCGGTTAATCACTAACCTG

CTTGTCGCTCAGCTCTGGCTCGGGGCTTCCAAATCAGTGACCGGAACAGAGCCAAGTTCAGCTTCCCAAATTGATGGGTGGGCCAGTCCTAATAAAGTATTTTTAAAAACACTCC-poly(A)
                                                                                           ______
```

FIG. 2. Nucleotide and deduced amino acid sequences of bovine tau cDNAs. (A) Nucleotide sequence of pBT43i12 (Fig. 1) with the encoded tau protein. The beginning of clone pBT12 and the end of pBT43 are indicated with a square and a dot, respectively. The polyadenylation signal of the mRNA (35) is underlined. The position where the common sequence of all clones starts is marked with an arrow. The 54-bp insertion of pBT43 is indicated with dots; the deletion of repeats 2 and 3 in pBT20 is underlined. (B) Sequence of alternative 5′ end of bovine tau, starting with the internal *Eco*RI site of pBT7. The beginnings of clones pBT4 and pBT20 are indicated. The deletion in pBT7 is underlined. The deleted fragment starts with the bases GTAAGT and ends with a T-rich stretch of 20 bases followed by AAG, matching the consensus sequence for intron borders compiled from many genes (30). Because of this deletion, the first potential translation initiation site of pBT7 (boxed) is further downstream. The sequence ends where it becomes identical to the one in panel A (arrow).

Synthesis of tau proteins in vitro. To examine the relationship between transcript heterogeneity and tau protein size heterogeneity, an in vitro transcription-translation system was used. Transcripts of the different cDNAs, generated in vitro with T7 RNA polymerase, were translated in rabbit reticulocyte lysates to obtain radioactively labeled tau proteins. This procedure revealed that the translation initiation site of clone pBT43 was much more efficient than were those of clones pBT4 and pBT7, consistent with the rules of Kozak (22). The translation product of clone pBT43-12 migrated in SDS-polyacrylamide gels with an apparent size of 61 kDa (Fig. 5), although the molecular mass deduced from the cDNA sequence was only 44.4 kDa. Nevertheless, this migration rate corresponded well to that of natural bovine tau proteins, estimated by electrophoresis to be about 54 to 65 kDa (2, 7), and suggested that plasmids pBT43-12 and pBT43i12 contained the complete coding sequence for two bovine tau isoforms.

Immunoprecipitation of the various translation products with affinity-purified polyclonal antiserum to bovine tau proteins (34) confirmed that the isolated clones represented tau messages. Both the N-terminal domain of clone pBT43

B.

```
GAATTCCGTGGAGAGAGGAGCCTGGAAGGTACAGTCCATGGGGTCGCAAAGGGTCGGACACGACTGAGCAACTAACAGTTTCACTGTGTGGCTCAGGGATCTCTACTCAGTGCTCTGTG

GTGACCTAAATGGGAAGGAAATCCAAAAAAGAGGGCATATATATATACATATACCTGATTCACTGTGCTGTACAGTAGAAACTAACAACAGTGCAAAGCAACTACACTCCAAGAAAACT
                                                                          ┌→BT4
TAATTTTAAAAATATTGGGTGGGGGGAGATTTACCAGACAGTTTAGTTTTGAGACTTTAATCAGAAAGTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTCTAAGTGGTATCAGA
                                                                                                          ┌→BT20
CAGAATCCTTAAGACCTGCCTAATCAGGTACAGTCCTGATGTTGGCAAGGACGATTCTATCAGTGCAGTGGAAGAGGGAGAGGGGTTAGGGGGCCAGCATGAGCTCCCCACCTGGAATT

GAGCAGAACCAGGGGCACTGCTGACCCATGACACACTGGAGATGTTTTCTGGAGAAACATCCTAAGAATCATGTACTTCTGGCGATCTCCATTGACTCTCTGGGTGCGAGCCAGCTTTT

AACCAGTGGTCTCCAAAGTGCGGAGCTGAAATAGATCCAGTGGGCTGGAGAAAGAACATACCAACACGTCTGTCATTTTCTCCAGCTCTGCCATATGTAAGTTTCATAATGTACGTCAT
                                                                                                                      1
                                                                                                                    met
ATGCTAGCATCGTAGTATATGCATTAGCTGATGAATGACTGCAAGTTCATGGTCACTTTCTGGTCTTCCTGATAAGGACTGTTAAGGGTGAATGATCCACATGGTTTGGAGATC ATG

                                 10                                      20                                      30
pro leu asn his tyr leu pro tyr leu phe leu val ser val leu phe gln phe val pro phe ser his val leu thr phe ile leu ile
CCT TTA AAC CAT TAC CTG CCT TAT TTG TTT TTA GTT AGT GTC CTT TTT CAG TTT GTA CCT TTC TCC CAT GTG CTG ACT TTT ATT TTG ATT

                                 40                                      50                                      60
leu phe met phe met phe lys pro ser thr pro ser ser ala lys thr leu lys asn arg pro cys leu ser pro lys arg pro thr pro
TTA TTT ATG TTT ATG TTT AAG CCA TCC ACA CCT TCT TCC GCT AAA ACC CTG AAA AAT AGG CCT TGC CTT AGC CCC AAA CGC CCC ACT CCT

                                 70                                      80                                      90
gly ser ser asp pro leu ile lys pro ser ser pro ala val cys pro glu pro ser ser ser pro lys his val ser ser val thr pro
GGT AGC TCA GAC CCT TTG ATC AAA CCC TCC AGC CCT GCC GTG TGC CCA GAG CCA TCT TCC TCT CCT AAA CAC GTC TCT TCT GTC ACA CCC

                                 100
arg thr gly asn ser gly ala lys glu [met] lys val lys
CGA ACT GGC AAT TCT GGA GCA AAG GAG [ATG] AAA GTC AAG
```

and the C-terminal domain of tau were recognized by the antibodies (Fig. 5).

Characterization of the microtubule-binding domain. The most distinctive feature of the bovine tau proteins, as deduced from the cDNA sequences, was the occurrence of four consecutive repeats of 31 or 32 amino acids, each containing a Pro-Gly-Gly-Gly motif (Fig. 3). To evaluate the previous suggestion (25) that these repeats, three of which are present in murine and human tau, might be the tubulin-binding sites, we generated radioactively labeled polypeptides representing various parts of tau in vitro. To increase the otherwise poor translation initiation at internal ATG codons, a short region upstream of the translation initiation site of clone pBT43 was inserted between the T7 promoter and the beginning of the first repeat in plasmids pBT4 and pBT20, replacing the 5′ portion of the cDNAs. To produce proteins ending exactly with repeat 4, we required precisely trimmed transcripts. Since there is no restriction site at this position, appropriate DNA fragments were generated in vitro by amplifying the region from the T7 promoter to the end of repeat 4, using the PCR. With these different constructions, it was possible to synthesize in vitro polypeptides corresponding to the N-terminal 132 amino acids of tau from clone pBT43, the entire C-terminal domain of tau starting with repeat 1, all four repeats alone, and repeats 1 and 4 alone. These fragments of tau were tested for their abilities to bind to and cosediment with microtubules (Fig. 5). This experiment revealed that the tubulin-binding domain of tau is localized in the C-terminal domain containing the repeats and that two of these repeats alone are sufficient to coprecipitate with microtubules.

DISCUSSION

In this study, the isolation and characterization of several cDNA clones for bovine tau proteins showed that part of the protein size heterogeneity is reflected in the primary structure of the transcripts, which differ by insertions or deletions. The sizes of translation products, recognition by affinity-purified antiserum, and the ability of the products to bind to microtubules suggested that the isolated cDNA clones account for complete tau-coding sequences. The difference in size between the cDNA clones which apparently contain the natural 3′ end and the tau mRNAs as determined by Northern blot analysis suggested an unusually long stretch of 5′ noncoding region, even for the smallest tau message. The strength of the hybridization signal indicated that the most abundant tau-related transcript in cow brain is 5 kb. The additional weak signals of mRNAs of 8 and 9 kb may represent rare forms of different tau transcripts. Alternatively, the 9-kb band may have been generated by cross-hybridization to the mRNA of the considerably larger MAP, MAP2. MAP2 has been shown to cross-react immunologically with certain tau antibodies (44) and in humans is generated from a transcript of 9 kb (6). Similar functions and characteristics of MAP2 and tau proteins, such as tubulin binding, could be due to conserved stretches in their primary structures, allowing cross-hybridization even under the stringent conditions applied.

Alignment and comparison of the nucleotide and its derived amino acid sequences with those of the murine and human tau cDNAs indicate that tau is a highly conserved protein (13, 25). Clone pBT43i12 shows 82% sequence

```
repeat1  Q A A P G P M P D L K N - V K S K I G S T E N L K H Q P G G G K
repeat2  V Q I I N K K L D L S N - V Q S K C G S K D N I K H V P G G G S
repeat3  V Q I V Y K P V D L S K - V T S K C G S L G N I H H K P G G G Q
repeat4  V E V K S E K L D F K D R V Q S K I G S L D N I T H V P G G G N
```

FIG. 3. Alignment of amino acids 251 to 375 encoded in clone pBT43i12 (Fig. 2), containing four repeats and representing the microtubule-binding domain of tau. The amino acids are given in one-letter notation, and conserved residues are boxed. Dashes represent gaps introduced for optimal alignment of the sequences. The corresponding three repeats (1, 3, and 4) in the murine and human tau sequences deduced from analyses of cDNA clones (13, 25) are identical to those of bovine except that repeats 1 of murine and human each contain QTAPVP as their first six amino acids.

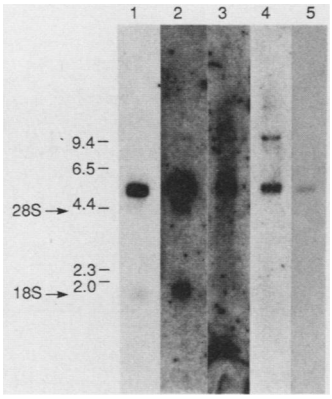

FIG. 4. Northern blot hybridization of bovine brain mRNA, using different ^{32}P-labeled bovine tau DNAs as probes. Each lane contained 2 μg of bovine brain poly(A)$^+$ RNA. Hybridization probes: lanes 1 and 2, total cDNA insert of pBT43i12; lane 3, 0.9-kb *Eco*RI-*Sma*I fragment of pBT7; lane 4, 0.6-kb *Eco*RI-*Sma*I fragment of pBT20 encoding the alternative N terminus; lane 5, synthetic antisense oligonucleotide (56 nucleotides) against repeat 2. Exposure times were 17 h for lane 1 and 5 days for the others. Sizes of markers are in kilobases. rRNAs are marked with arrows.

identity at the nucleotide level and 85% identity at the amino acid level with the shorter of the two mouse tau clones. The most striking difference between the isolated clones is the occurrence of several insertions and deletions within the coding region, which always maintain the reading frame; this observation therefore suggests differential splicing of the tau mRNAs. After determining the exon-intron structure of the bovine tau gene (17), we concluded that the deletion in clone pBT20 which removed repeats 2 and 3 was generated by a recombination event in the lambda bacteriophage within a perfect direct repeat of the 14 bases AGCCAGGAGGTGGC at the ends of the deletion and was not represented in an actual transcript. Despite this fact, clone pBT20 was useful for studying the interaction of tau with microtubules. Compared with the mouse sequence, the sequence of bovine tau protein has four insertions of 2 to 58 amino acids (Fig. 2). Beyond the difference of the presence of repeat 2, the C-terminal half of the tau protein is highly conserved among cows, mice, and humans, being identical for the last 136 amino acids and showing only one nonconservative and five conservative changes in the region from amino acids 197 to 281. Interestingly, the point of divergence between clones pBT43-pBT12 and pBT4 divides the tau proteins into common middle and C-terminal domains, which are predominantly positively charged (46 basic versus 23 acidic amino acids), and a rather acidic N-terminal domain (32 acidic versus 9 basic amino acids) in the case of pBT43 or a hydrophobic 37-amino-acid-long stretch at the N terminus in the case of pBT4.

Cosedimentation of in vitro translated tau proteins and tau fragments with microtubules revealed that the microtubule-binding domain of tau resides in the C-terminal domain containing the repeats. The N-terminal 132 amino acids of pBT43 did not bind to microtubules. Truncating the polypeptide containing the C-terminal domain and reducing the number of repeats revealed that two repeats alone are sufficient to bind to microtubules. This is in good agreement with the observation that a larger chymotryptic digestion product of bovine tau of M_r 14,000 contains the microtubule-binding domain of tau (1). The reported amino-terminal sequence of this peptide was identical with the sequence Ser-205 to Pro-223 in pBT43i12. Judging from the cleavage

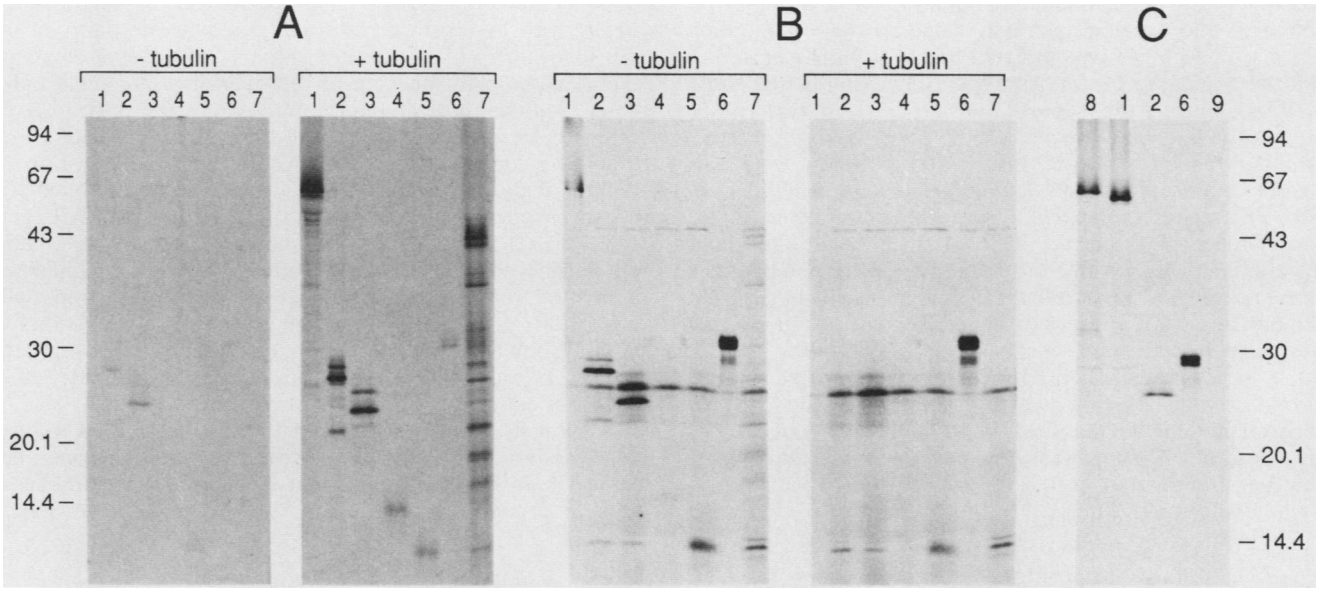

FIG. 5. Cosedimentation with microtubules and immunoprecipitation of in vitro translated tau domains. Transcripts generated with T7 RNA polymerase from linearized plasmids were translated in rabbit reticulocyte lysates and incubated with microtubules. Microtubules were pelleted through a sucrose cushion to separate them from unpolymerized material. (A) Pelleted translation products in the absence or presence of tubulin. (B) Translation products in the supernatant in the absence or presence of tubulin. (C) Immunoprecipitation of in vitro translation products with affinity-purified antiserum to bovine tau protein. Lanes: 1, translation products from clone pBT43-12 (complete tau protein); 2, the C-terminal domain of tau, starting at amino acid 250, just before repeat 1 (pBT4BSMP); 3, C-terminal half of pBT20, containing only repeats 1 and 4 (pBT20BSMP); 4, four repeats alone (PCR-4); 5, repeats 1 and 4 (PCR-20); 6, N-terminal 132 amino acids of pBT43 (pBT43 truncated with *Sty*I); 7, tau with alternative N terminus (pBT4); 8, complete tau protein from clone pBT43i12; 9, the same translation products of pBT43i12 but using rabbit preimmune serum. Molecular weights of marker proteins are in thousands.

specificity of chymotrypsin and the slower than expected migration of C-terminal tau fragments in SDS-gels, we concluded that the sequence Ser-205 to Tyr-317 (molecular weight, 11,723) in pBT43i12, containing two repeats, corresponds to the M_r-14,000 peptide. The basic nature of the tau repeats and the whole C-terminal domain and the reported interaction of synthetic peptides from the acidic C-terminal domain of tubulin with tau and MAP2 (27) indicate that these proteins interact electrostatically by their C-terminal domains. Recently described anti-idiotypic antibodies found in the sera of rabbits immunized with these tubulin domains bound both to tau and MAP2 proteins (36). This result suggests that the microtubule-binding domains of tau and MAP2 have a conserved structure which might also be present in the much larger group of microtubule-associated proteins.

Tau protein identified as a component of paired helical filaments in neurofibrillary tangles of Alzheimer's disease patients shows as extensive size heterogeneity as does nonhuman tau (15). It was reported that the tau immunoreactivity of neurofibrillary tangles can be removed by treatment with sulfhydryl reducing agents (40). The only cysteine residues within the constant region of tau that could generate disulfide bonds are located within the repeats. Recent reports provide evidence that the repeat region is present in the core of the paired helical filament of Alzheimer's disease (42). This finding indicates that the repeat region of tau may be involved not only in tubulin binding but also in binding to or the assembly of paired helical filaments.

The overall structure-function relationship of tau proteins resembles that of other protein families having constant effector domains linked to variable receptor domains. The constant effector (C-terminal microtubule-binding) domain of tau would be similar to the constant regions of the immunoglobulin superfamily that includes immunoglobulins, T-cell antigen receptors, and cell adhesion molecules (18, 41) or the similar tyrosine kinase domains of the family of growth factor receptors (e.g., epidermal growth factor, platelet-derived growth factor, and insulin) and their related oncogene products (43). The variable N-terminal domains of tau with still unknown function may be compared with the variable receptor domains of the above-mentioned protein families involved in binding to different antigens or ligands. It remains to be determined whether these variations between the tau isoforms can affect the kinetics of tubulin polymerization or binding properties or may be responsible for different tubulin functions in the context of different cellular compartments and physiological conditions.

ACKNOWLEDGMENTS

We thank Peter Schofield for providing the cDNA library, the Genentech, Inc., DNA synthesis group for providing the oligonucleotides, Kerrie Andow and Carol Morita for preparing the figures, and Karen Butner and Justus Cohen for helpful discussions.

ADDENDUM

While this manuscript was under review, Lewis et al. (26) published a description of the repeat domains in MAP2 which bind microtubules and are homologous to the domains of tau protein that are shown herein to coprecipitate with tubulin.

LITERATURE CITED

1. **Aizawa, H., H. Kawasaki, H. Murofushi, S. Kotani, K. Suzuki, and H. Sakai.** 1988. Microtubule-binding domain of tau proteins. J. Biol. Chem. **263**:7703–7707.
2. **Baudier, J., and R. D. Cole.** 1987. Phosphorylation of tau proteins to a state that in Alzheimer's brain is catalyzed by a calcium/calmodulin-dependent kinase and modulated by phospholipids. J. Biol. Chem. **262**:17577–17583.
3. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. Science **196**:180–182.
4. **Biggin, M. D., T. J. Gibson, and G. F. Hong.** 1983. Buffer gradient gels and ^{35}S label as an aid to rapid sequence determination. Proc. Natl. Acad. Sci. USA **80**:3963–3965.
5. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.
6. **Dammerman, M., M. Goldstein, S.-H. C. Yen, and B. Shafit-Zagardo.** 1988. Isolation and characterization of cDNA clones encoding epitopes shared with Alzheimer neurofibrillary tangles. J. Neurosci. Res. **19**:43–51.
7. **Drubin, D. G., D. Caput, and M. W. Kirschner.** 1984. Studies on the expression of the microtubule-associated protein, tau, during mouse brain development, with newly isolated complementary DNA probes. J. Cell Biol. **98**:1090–1097.
8. **Drubin, D. G., S. C. Feinstein, E. M. Shooter, and M. W. Kirschner.** 1985. Nerve growth factor-induced neurite outgrowth in PC12 cells involves the coordinate induction of microtubule assembly and assembly-promoting factors. J. Cell Biol. **101**:1799–1807.
9. **Drubin, D. G., and M. W. Kirschner.** 1986. Tau protein function in living cells. J. Cell Biol. **103**:2739–2746.
10. **Drubin, D., S. Kobayashi, and M. Kirschner.** 1986. Association of tau protein with microtubules in living cells. Ann. N.Y. Acad. Sci. **466**:257–268.
11. **Dustin, P.** 1984. Microtubules. Springer-Verlag, New York.
12. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
13. **Goedert, M., C. M. Wischik, R. A. Crowther, J. E. Walker, and A. Klug.** 1988. Cloning and sequencing of the cDNA encoding a core protein of the paired helical filament of Alzheimer disease: identification as the microtubule-associated protein tau. Proc. Natl. Acad. Sci. USA **85**:4051–4055.
14. **Grundke-Iqbal, I., K. Iqbal, M. Quinlan, Y.-C. Tung, M. S. Zaidi, and H. M. Wisniewski.** 1986. Microtubule-associated protein tau: a component of Alzheimer paired helical filaments. J. Biol. Chem. **261**:6084–6089.
15. **Grundke-Iqbal, I., K. Iqbal, Y.-C. Tung, M. Quinlan, H. M. Wisniewski, and L. I. Binder.** 1986. Abnormal phosphorylation of the microtubule-associated protein τ (tau) in Alzheimer cytoskeletal pathology. Proc. Natl. Acad. Sci. USA **83**:4913–4917.
16. **Han, J. H., C. Stratowa, and W. J. Rutter.** 1987. Isolation of full-length putative rat lysophospholipase cDNA using improved methods for mRNA isolation and cDNA cloning. Biochemistry **26**:1617–1625.
17. **Himmler, A.** 1989. Structure of the bovine tau gene: alternatively spliced transcripts generate a protein family. Mol. Cell. Biol. **9**:1389–1396.
18. **Hood, L., M. Kronenberg, and T. Hunkapiller.** 1985. T cell antigen receptors and the immunoglobulin supergene family. Cell **40**:225–229.
19. **Horio, T., and H. Hotani.** 1986. Visualization of the dynamic instability of individual microtubules by dark-field microscopy. Nature (London) **321**:605–607.
20. **Ihara, Y., N. Nukina, R. Miura, and M. Ogawara.** 1986. Phosphorylated tau protein is integrated into paired helical filaments in Alzheimer's disease. J. Biochem. **99**:1807–1810.
21. **Kirschner, M., and T. Mitchison.** 1986. Beyond self-assembly: From microtubules to morphogenesis. Cell **45**:329–342.
22. **Kozak, M.** 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. **15**:8125–8148.
23. **Ksiezak-Reding, H., D. W. Dickson, P. Davies, and S.-H. Yen.** 1987. Recognition of tau epitopes by anti-neurofilament antibod-

ies that bind to Alzheimer neurofibrillary tangles. Proc. Natl. Acad. Sci. USA **84**:3410–3414.

24. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:580–585.
25. **Lee, G., N. Cowan, and M. Kirschner.** 1988. The primary structure and heterogeneity of tau protein from mouse brain. Science **239**:285–288.
26. **Lewis, S. A., D. Wang, and N. J. Cowan.** 1988. Microtubule-associated protein MAP2 shares a microtubule binding motif with tau protein. Science **242**:936–939.
27. **Maccioni, R. B., C. I. Rivas, and J. C. Vera.** 1988. Differential interaction of synthetic peptides from the carboxy-terminal regulatory domain of tubulin with microtubule-associated proteins. EMBO J. **7**:1957–1963.
28. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Mitchison, T., and M. Kirschner.** 1984. Microtubule assembly nucleated by isolated centrosomes. Nature (London) **312**:232–237.
30. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
31. **Nukina, N., K. S. Kosik, and D. S. Selkoe.** 1987. Recognition of Alzheimer paired helical filaments by monoclonal neurofilament antibodies is due to crossreaction with tau protein. Proc. Natl. Acad. Sci. USA **84**:3415–3419.
32. **Olmsted, J. B.** 1986. Microtubule-associated proteins. Annu. Rev. Cell Biol. **2**:421–457.
33. **Paschal, B. M., H. S. Shpetner, and R. B. Vallee.** 1987. MAP 1C is a microtubule-activated ATPase which translocates microtubules in vitro and has dynein-like properties. J. Cell Biol. **105**:1273–1282.
34. **Pfeffer, S. R., D. G. Drubin, and R. J. Kelly.** 1983. Identification of three coated vesicle components as α and β tubulin linked to a phosphorylated 50,000 dalton polypeptide. J. Cell Biol. **97**: 40–47.
35. **Proudfoot, N. J., and G. G. Brownlee.** 1976. 3′ Non-coding region sequences in eukaryotic messenger RNA. Nature (London) **263**:211–214.
36. **Rivas, C. I., J. C. Vera, and R. B. Maccioni.** 1988. Anti-idiotypic antibodies that react with microtubule-associated proteins are present in the sera of rabbits immunized with synthetic peptides from tubulin's regulatory domain. Proc. Natl. Acad. Sci. USA **85**:6092–6096.
37. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Ehrlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.
38. **Schofield, P. R., M. G. Darlison, N. Fujita, D. R. Burt, F. A. Stephenson, H. Rodriguez, L. M. Rhee, J. Ramachandran, V. Reale, T. A. Glencorse, P. H. Seeburg, and E. A. Barnard.** 1987. Sequence and functional expression of the GABA receptor shows a ligand-gated receptor super-family. Nature (London) **328**:221–227.
39. **Schulze, E., and M. Kirschner.** 1987. Dynamic and stable populations of microtubules in cells. J. Cell Biol. **104**:277–288.
40. **Shaw, G., and V. Chau.** 1988. Ubiquitin and microtubule-associated protein tau immunoreactivity each define distinct structures with differing distributions and solubility properties in Alzheimer brain. Proc. Natl. Acad. Sci. USA **85**:2854–2858.
41. **Staunton, D. E., S. D. Martin, C. Stratowa, M. L. Dustin, and T. A. Springer.** 1988. Primary structure of ICAM-1 demonstrates interaction between members of the immunoglobulin and integrin supergene families. Cell **52**:925–933.
42. **Wischik, C. M., M. Novak, H. C. Thogersen, P. C. Edwards, M. J. Runswick, R. Jakes, J. E. Walker, C. Milstein, M. Roth, and A. Klug.** 1988. Isolation of a fragment of tau derived from the core of the paired helical filament of Alzheimer disease. Proc. Natl. Acad. Sci. USA **85**:4506–4510.
43. **Yarden, Y., and A. Ullrich.** 1988. Growth factor receptor tyrosine kinases. Annu. Rev. Biochem. **57**:443–478.
44. **Yen, S.-H., D. W. Dickson, A. Crowe, M. Butler, and M. L. Shelanski.** 1987. Alzheimer's neurofibrillary tangles contain unique epitopes and epitopes in common with the heat-stable microtubule associated proteins tau and MAP2. Am. J. Pathol. **126**:81–91.