The EMBO Journal vol.8 no.2 pp.393 - 399, 1989

Cloning and sequencing of the cDNA encoding an isoform of microtubule-associated protein tau containing four tandem repeats: differential expression of tau protein mRNAs in human brain

M.Goedert, M.G.Spillantini, M.C.Potier, J.Ulrich[1] and R.A.Crowther

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK and [1]Division of Neuropathology, Department of Pathology, University of Basel, Schönbeinstrasse 40, 4003 Basel, Switzerland

Communicated by R.A.Crowther

We have isolated cDNA clones encoding a 383-amino acid isoform of the human microtubule-associated protein tau. It differs from previously determined tau sequences by the presence of an additional repeat of 31 amino acids, giving four, rather than three, tandem repeats in its carboxy-terminal half. The extra repeat is encoded by a separate exon. Probes derived from cDNA clones encoding the three (type I) and four repeat (type II) tau protein isoforms detected mRNAs for both forms in all adult human brain areas examined. However, in foetal brain only type I mRNA was found. Type I and type II mRNAs were present in pyramidal cells in cerebral cortex. In the hippocampal formation, type I mRNA was found in pyramidal and granule cells; type II mRNA was detected in most, though not all, pyramidal cells but not in granule cells. These observations indicate that tau protein mRNAs are expressed in a stage- and cell-specific manner. Tau protein is found in the protease-resistant core of the paired helical filament, the major constituent of the neurofibrillary tangle in Alzheimer's disease. Taken in conjunction with previous findings, the present results indicate that both the three and four repeat-containing tau protein isoforms are present in the core of the paired helical filament.

Key words: alternative mRNA splicing/Alzheimer's disease/ microtubule-associated proteins

Introduction

Microtubule-associated tau proteins promote microtubule assembly and stabilize microtubules (Weingarten *et al.*, 1975; Cleveland *et al.*, 1977a,b; Carlier *et al.*, 1984; Horio and Hotani, 1986). Tau occurs throughout the mammalian nervous system (Binder *et al.*, 1985) and in the adult consists of a family of four to six related polypeptides with apparent mol. wts of 50 000–68 000 daltons, when examined by one-dimensional gel electrophoresis (Cleveland *et al.*, 1977a). In the foetus, however, only one tau protein band of apparent mol. wt 48 000 daltons is found (Couchie and Nunez, 1985). At present, the molecular basis for this heterogeneity is incompletely understood. In the mouse, molecular cloning has shown the existence of at least two forms of tau, 341 and 364 amino acids in length, that differ in their carboxy-terminal sequence (Lee *et al.*, 1988). We have recently determined the amino acid sequence of a 352-residue form of human tau (Goedert *et al.*, 1988), that probably represents the human equivalent of the shorter of the two mouse forms. All the tau sequences reported to date contain, in their carboxy-terminal half, a characteristic region of three tandem repeats of 31 or 32 amino acids, each displaying a distinctive Pro-Gly-Gly-Gly motif. It has been proposed that the three repeats may represent the tubulin-binding region of tau (Lee *et al.*, 1988). We describe here the cloning and sequencing of the cDNA encoding an isoform of human tau that contains four such repeats.

Although tau protein is normally associated with microtubules, it has been demonstrated by immunological techniques to be present also in the paired helical filaments that form the bulk of the pathological neurofibrillary tangle in Alzheimer's disease (Brion *et al.*, 1985; Grundke-Iqbal *et al.*, 1986; Wood *et al.*, 1986; Kosik *et al.*, 1986). We have recently shown that tau protein contributes to the protease-sensitive fuzzy coat around the paired helical filament and have also proved by direct protein sequencing that the triple repeat region of tau lies in the protease-resistant core of the filament (Goedert *et al.*, 1988; Wischik *et al.*, 1988a,b). The protein sequencing indicated the presence of two forms of tau in the core of the paired helical filament (Wischik *et al.*, 1988a) and we show here that the two sequences are likely to correspond to the three- and four-repeat forms of tau protein.

We have derived probes, which we call type I and type II, from the cDNA clones encoding the three- and four-repeat isoforms respectively. These probes distinguish between the two isoforms studied so far, but in view of the complexity of expression of tau protein we cannot be sure that they will identify unique mRNAs. Nevertheless they do serve to demonstrate different patterns of expression of the two types. We detect mRNA recognized by both types of probe in all areas of the adult human brain examined, with type I being more abundant than type II. However in foetal brain, although type I is abundant, type II could not be detected. At a cellular level, granule cells in the adult dentate gyrus showed expression of type I but not of type II mRNA. Thus, the expression of tau protein mRNAs is both stage and cell type specific.

Results

Nucleotide sequence and deduced amino acid sequence of a human tau protein isoform

Six positive clones were obtained when a cDNA library prepared from normal frontal cortex was screened with a 17mer oligonucleotide probe derived from the repeat region of the published human tau sequence (Goedert *et al.*, 1988). The same oligonucleotide was used as a primer to obtain a partial nucleotide sequence of these clones. Four clones were identical to the published sequence, whereas two differed in the region sequenced. One clone with an insert

```
                                                                                  -30 TGTCGACTATCAGGTGAACTTTGAACCAGG

   1 Met Ala Glu Pro Arg Gln Glu Phe Glu Val Met Glu Asp His Ala Gly Thr Tyr Gly Leu Gly Asp Arg Lys Asp Gln Gly Gly Tyr Thr
   1 ATG GCT GAG CCC CGC CAG GAG TTC GAA GTG ATG GAA GAT CAC GCT GGG ACG TAC GGG TTG GGG GAC AGG AAA GAT CAG GGG GGC TAC ACC

  31 Met His Gln Asp Gln Glu Gly Asp Thr Asp Ala Gly Leu Lys Ala Glu Glu Ala Gly Ile Gly Asp Thr Pro Ser Leu Glu Asp Glu Ala
  91 ATG CAC CAA GAC CAA GAG GGT GAC ACG GAC GCT GGC CTG AAA GCT GAA GAA GCA GGC ATT GGA GAC ACC CCC AGC CTG GAA GAC GAA GCT

  61 Ala Gly His Val Thr Gln Ala Arg Met Val Ser Lys Ser Lys Asp Gly Thr Gly Ser Asp Asp Lys Lys Ala Lys Gly Ala Asp Gly Lys
 181 GCT GGT CAC GTG ACC CAA GCT CGC ATG GTC AGT AAA AGC AAA GAC GGG ACT GGA AGC GAT GAC AAA AAA GCC AAG GGG GCT GAT GGT AAA

  91 Thr Lys Ile Ala Thr Pro Arg Gly Ala Ala Pro Pro Gly Gln Lys Gly Gln Ala Asn Ala Thr Arg Ile Pro Ala Lys Thr Pro Pro Ala
 271 ACG AAG ATC GCC ACA CCG CGG GGA GCA GCC CCT CCA GGC CAG AAG GGC CAG GCC AAC GCC ACC AGG ATT CCA GCA AAA ACC CCG CCC GCT

 121 Pro Lys Thr Pro Pro Ser Ser Gly Glu Pro Pro Lys Ser Gly Asp Arg Ser Gly Tyr Ser Ser Pro Gly Ser Pro Gly Thr Pro Gly Ser
 361 CCA AAG ACA CCA CCC AGC TCT GGT GAA CCT CCA AAA TCA GGG GAT CGC AGC GGC TAC AGC AGC CCC GGC TCC CCA GGC ACT CCC GGC AGC

 151 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 Ser
 451 CGC TCC CGC ACC CCG TCC CTT CCA ACC CCA CCC ACC CGG GAG CCC AAG AAG GTG GCA GTG GTC CGT ACT CCA CCC AAG TCG CCG TCT TCC

 181 Ala Lys Ser Arg Leu Gln Thr Ala Pro Val Pro Met Pro Asp Leu Lys Asn Val Lys Ser Lys Ile Gly Ser Thr Glu Asn Leu Lys His
 541 GCC AAG AGC CGC CTG CAG ACA GCC CCC GTG CCC ATG CCA GAC CTG AAG AAT GTC AAG TCC AAG ATC GGC TCC ACT GAG AAC CTG AAG CAC

 211 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 Lys
 631 CAG CCG GGA GGC GGG AAG GTG CAG ATA ATT AAT AAG AAG CTG GAT CTT AGC AAC GTC CAG TCC AAG TGT GGC TCA AAG GAT AAT ATC AAA

 241 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 Ile
 721 CAC GTC CCG GGA GGC GGC AGT GTG CAA ATA GTC TAC AAA CCA GTT GAC CTG AGC AAG GTG ACC TCC AAG TGT GGC TCA TTA GGC AAC ATC

 271 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 Asp
 811 CAT CAT AAA CCA GGA GGT GGC CAG GTG GAA GTA AAA TCT GAG AAG CTT GAC TTC AAG GAC AGA GTC CAG TCG AAG ATT GGG TCC CTG GAC

 301 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 His
 901 AAT ATC ACC CAC GTC CCT GGC GGA GGA AAT AAA AAG ATT GAA ACC CAC AAG CTG ACC TTC CGC GAG AAC GCC AAA GCC AAG ACA GAC CAC

 331 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 Asp
 991 GGG GCG GAG ATC GTG TAC AAG TCG CCA GTG GTG TCT GGG GAC ACG TCT CCA CGG CAT CTC AGC AAT GTC TCC TCC ACC GGC AGC ATC GAC

 361 Met Val Asp Ser Pro Gln Leu Ala Thr Leu Ala Asp Glu Val Ser Ala Ser Leu Ala Lys Gln Gly Leu ***
1081 ATG GTA GAC TCG CCC CAG CTC GCC ACG CTA GCT GAC GAG GTG TCT GCC TCC CTG GCC AAG CAG GGT TTG TGA TCAGGCCCCTGGGGCGGT
```

Fig. 1. Nucleotide and predicted amino acid sequences of a human tau protein isoform, as deduced from clone PHF24. It differs from the previously determined human tau sequence (Goedert *et al.*, 1988) by an extra repeat of 31 amino acids (sequence underlined). Nucleotides are numbered in the 5′→3′ direction, starting with the first nucleotide of the initiating codon; the nucleotides 5′ to residue 1 are indicated by negative numbers.

of ~2.2 kb (λPHF24) was further characterized. A partial nucleotide sequence determined from PHF24, which encompasses the complete coding region, is depicted in Figure 1. It consists of 30 nucleotides of 5′-untranslated sequence, an open reading frame of 1149 nucleotides, an in-frame stop codon and 18 nucleotides of 3′-untranslated sequence. The open reading frame encodes a protein of 383 amino acids that is identical with the published human tau sequence, with the exception of an inserted stretch of 31 amino acids (amino acids 217−247). The extra sequence (underlined in Figure 1) encodes an additional repeat with a characteristic Pro-Gly-Gly-Gly motif of the type present three times in the previously reported human tau sequence. It is inserted within the first repeat in a way that preserves the periodic pattern (Figure 2). The alignment shown here, which is slightly different from that used previously (Goedert *et al.*, 1988), has been chosen to maximize the homology between the repeats. There are now 12 residues completely conserved between the four repeats and a further four residues that show conservative changes. The predicted mol. wt for the protein containing three repeats is 36 760 daltons, while that for the four-repeat form is 40 007 daltons.

Three positive clones were obtained when a human genomic library was screened with a 296-bp *Pst*I/*Hin*dIII fragment containing the region encoding the additional repeat. Synthetic oligonucleotides were used as primers to obtain a partial nucleotide sequence of these clones. This showed that the additional repeat is encoded by a separate exon, flanked by consensus splice acceptor and donor sequences (Shapiro and Semapathy, 1987). The sequence reads as follows (with intron sequences shown in lower case): 5′tctggctaccaaagGTGCAGATAA.....AGGCGGCAGT gtgagtaccttcac3′.

```
D L K N V K S K I G S T E N L K H Q P G G G K[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]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 V E V K S E K L
D F KDR V Q S K I G S L D N I T H V P G G G N K K I E T H K L
* #     *   * *   * *     * #   *   * * * *         #       #
```

Fig. 2. Four tandem repeats (31 or 32 amino acids long) identified in the human tau protein isoform sequence. They correspond to amino acids 194−224, 225−255, 256−286 and 287−318. The extra repeat sequence is bracketed. Similarity is indicated as follows: *, all four residues identical; #, three residues are identical and the fourth is a conservative replacement.

Tissue levels of tau protein type I and type II mRNAs

RNase protection experiments were used to investigate the distribution of tau protein mRNAs using specific RNA probes derived from the cDNA clones encoding the three- and four-repeat tau proteins. The probes span the region containing the additional repeat and, when hybridized to the homologous mRNA, give rise to protected fragments of characteristic lengths. mRNAs are operationally defined as type I (hybridizing with the probe specific for the cDNA encoding the three-repeat protein) and type II (hybridizing with the probe specific for the cDNA encoding the four-repeat protein). This operational definition is taken because, based on the marked heterogeneity of tau, there are likely to be multiple transcripts that hybridize with either the type I or the type II probe. ^{32}P-Labelled anti-mRNA sense transcripts were synthesized and isolated from gels, either as a 214-nucleotide species (type I) comprising 203 nucleotides of coding region and 11 nucleotides of vector or as a 307-nucleotide species (type II) comprising 296 nucleotides of coding region and 11 nucleotides of vector.

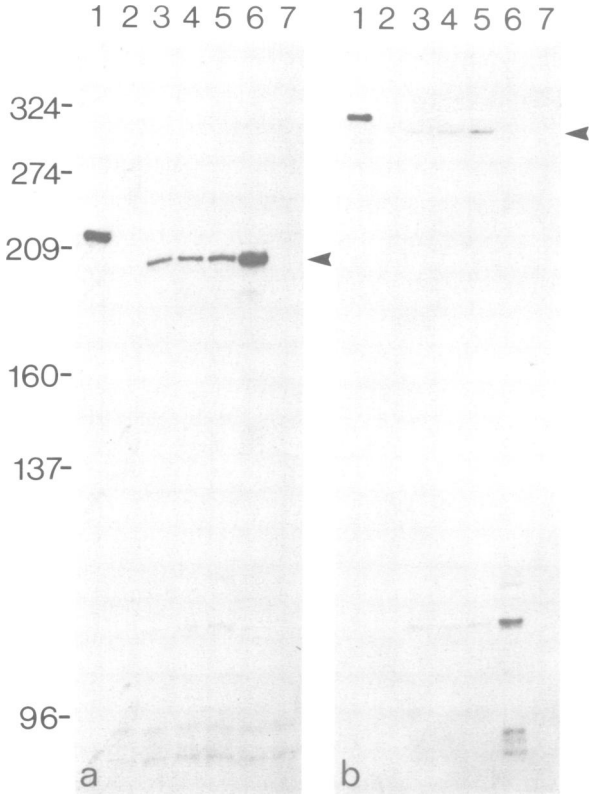

Fig. 3. Expression of tau protein type I (**a**) and type II (**b**) mRNAs in human tissues. Poly(A)$^+$ RNA (2 μg) was analysed by RNase protection. **Lanes: 1**, probe alone; **2**, probe + RNase; **3**, frontal cortex; **4**, hippocampus; **5**, cerebellum; **6**, foetal brain; **7**, liver. Mol. wt standards (*Hin*fI-cut M13mp18 labelled with [α-^{32}P]dATP) are indicated on the left. The protected fragments are indicated by arrows.

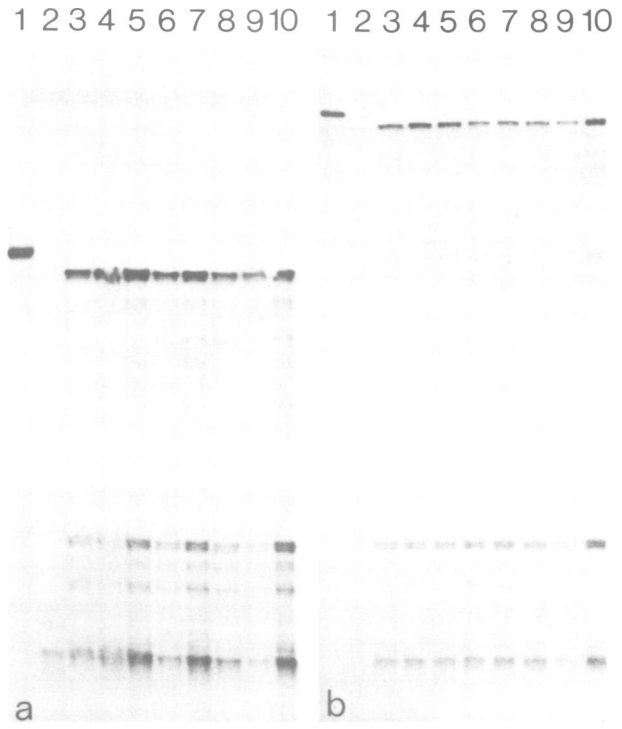

Fig. 4. Expression of tau protein type I (**a**) and type II (**b**) mRNAs in frontal cortex from control patients and from patients who had died with Alzheimer's disease. Total RNA (40 μg) was analysed by RNase protection. **Lanes: 1**, probe alone; **2**, probe + RNase; **3–6**, frontal cortex from control patients; **7–10**, frontal cortex from Alzheimer's disease patients.

Hybridizations with either total RNA or poly(A)$^+$ RNA prepared from human brain protected a 203-nucleotide species (type I) or a 296-nucleotide species (type II) against degradation by RNase (Figures 3 and 4). The two sub-fragments of 112 and 91 nucleotides which the two types have in common arise from cross-hybridization of one type of probe with mRNA of the other type. Omission of RNA during the hybridization step resulted in complete degradation of the probe by RNase (Figures 3 and 4).

As shown in Figure 3, tau type I and type II mRNAs were found in normal cerebral cortex, hippocampus and cerebellum; similar results were obtained from striatum, midbrain, thalamus and basal forebrain (data not shown). In adult brain the mRNA levels for each type did not vary much between different regions; however, the levels of type I mRNA were always 2- to 3-fold higher than the levels of type II mRNA. No specific signal was observed in liver (Figure 3). A major difference between the distributions of tau type I and type II mRNAs was observed in foetal brain, where type I mRNA was abundant, whereas type II mRNA could not be detected (Figure 3).

The levels of tau protein type I and type II mRNAs were investigated in the frontal cortex from four control patients and from four patients who had died with a histologically confirmed diagnosis of Alzheimer's disease by using the RNase protection assay (Figure 4). The levels of tau protein type I and type II mRNAs were not significantly changed in frontal cortex in Alzheimer's disease (Figure 4).

Cellular distribution of tau protein type I and type II mRNAs

Synthetic oligonucleotide probes were used to study the distribution *in situ* of tau protein type I and type II mRNAs in the cerebral cortex and the hippocampal formation. The level of cross-hybridization of the oligonucleotides was checked by using the cDNA clones encoding the three- and four-repeat tau proteins. Under the stringency of hybridization and washing conditions used, it amounted to $<1\%$.

Tau protein type I and type II mRNAs showed a wide distribution throughout the cerebral cortex and hippocampal formation, where their localization was exclusively neuronal (Figures 5–7). No labelling of glial cells was observed. The hybridization was specific, as only background labelling was seen when a probe in the mRNA sense orientation was used as control (Figures 5b, 5d, 6b, 6d, 7b). In the cerebral cortex, hybridization-positive cells for tau protein type I and type II mRNAs were present in all layers, with the highest density in the deeper layers. The distribution and size of these cells indicate that pyramidal cells constitute the cortical cell type that predominantly, if not exclusively, expresses tau protein type I and type II mRNAs (Figure 5a,c). In the hippocampal formation the tau protein type I-specific probe in the anti-mRNA sense orientation produced strong labelling of granule cells in the dentate gyrus, pyramidal cells throughout Ammon's horn and pyramidal cells in the subiculum (Figures 6a,c, 7c). All granule cells and all pyramidal cells appeared to be hybridization positive. By contrast, tau protein type II mRNA was not detectable in granule cells in the dentate gyrus (Figure 7d), whereas it

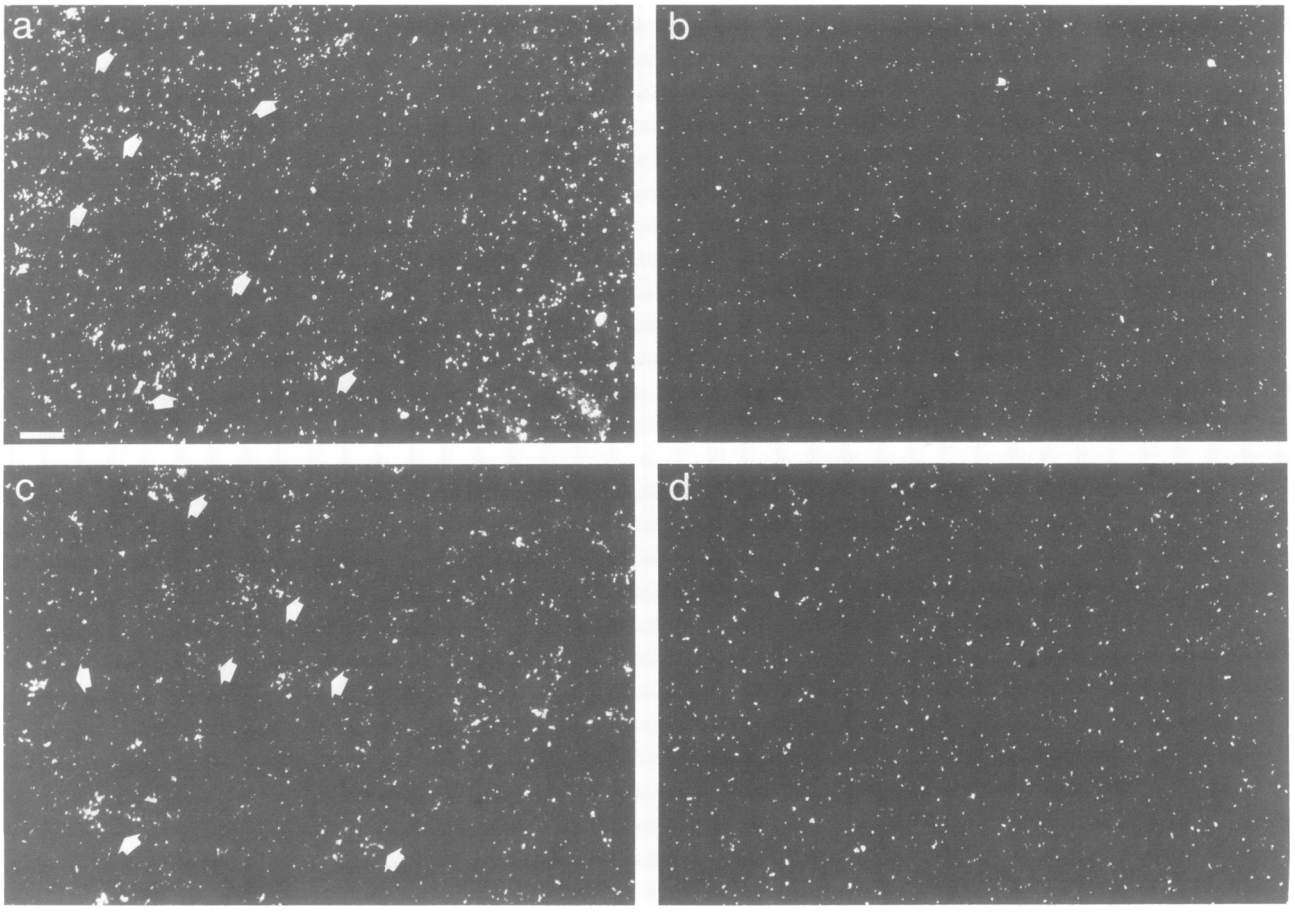

Fig. 5. Cellular localization of tau protein type I (**a,b**) and type II (**c,d**) mRNAs in frontal cerebral cortex. (**a,b**) Dark-field photomicrographs following hybridization with a type I-specific probe in the anti-mRNA sense orientation (**a**) or mRNA sense orientation (**b**). (**c,d**) Frontal cortex after hybridization with a type II-specific probe in the anti-mRNA sense orientation (**c**) or mRNA sense orientation (**d**). The arrows point to hybridization-positive cells. Scale bar, 70 μm.

was found in pyramidal cells in Ammon's horn (Figure 7a and d). However, not all pyramidal cells appeared to be hybridization positive.

Discussion

The present results provide the complete amino acid sequence of a second form of human tau protein, which differs from the previously described form (Goedert *et al.*, 1988) by the presence of an extra repeat of 31 amino acids. It contains four tandem repeats located in the carboxy-terminal half, each containing a characteristic Pro-Gly-Gly-Gly motif. A different number of repeats is thus one way in which multiple forms of tau protein can be generated. Another way has already been described in the mouse, where a cDNA clone encoding a tau protein with a carboxy-terminal extension has been found (Lee *et al.*, 1988). Current evidence indicates the existence of only one tau gene (Drubin *et al.*, 1984; Neve *et al.*, 1986; Goedert *et al.*, 1988); it thus appears likely that tau protein mRNA is subject to alternative splicing. This is confirmed by our finding that one exon completely and exclusively encodes the extra repeat described here. The functional implications of different numbers of repeats in human tau proteins are unknown at present. It has been suggested that the repeat region of tau may represent the tubulin-binding domain (Lee *et al.*, 1988) and current evidence is consistent with this (Aizawa *et al.*, 1988). Interestingly, the existence of four binding sites on tau for a synthetic peptide derived from the carboxy terminus of β-tubulin has been described (Maccioni *et al.*, 1988). It has also been shown that the high mol. wt forms of tau protein possess a higher affinity for tubulin than the lower mol. wt forms (Carlier *et al.*, 1984). It is tempting to speculate that this may result, at least in part, from the presence of an extra repeat in the higher mol. wt forms.

Transcripts for tau protein type I and type II mRNAs showed a wide distribution throughout the adult human nervous system, as both forms could be detected in all brain regions examined. By *in situ* hybridization they were found exclusively in neurones in cerebral cortex and hippocampal formation. Type I mRNA was present in pyramidal cells in the cerebral cortex, hippocampus and subiculum, as well as in hippocampal granule cells. By contrast, type II mRNA, although present in most pyramidal cells, could not be detected in granule cells in the hippocampus. This difference in distribution between the two types suggests a functional difference between the tau protein isoforms encoded by the two types of mRNA.

The distribution of tau protein mRNAs in the hippocampal formation is identical to that of the amyloid β protein precursor mRNAs (Bahmanyar *et al.*, 1987; Goedert, 1987) and it is therefore likely that both mRNAs are present in the same cells. Thus an insult to these cells in the course of the development of Alzheimer's disease could give rise

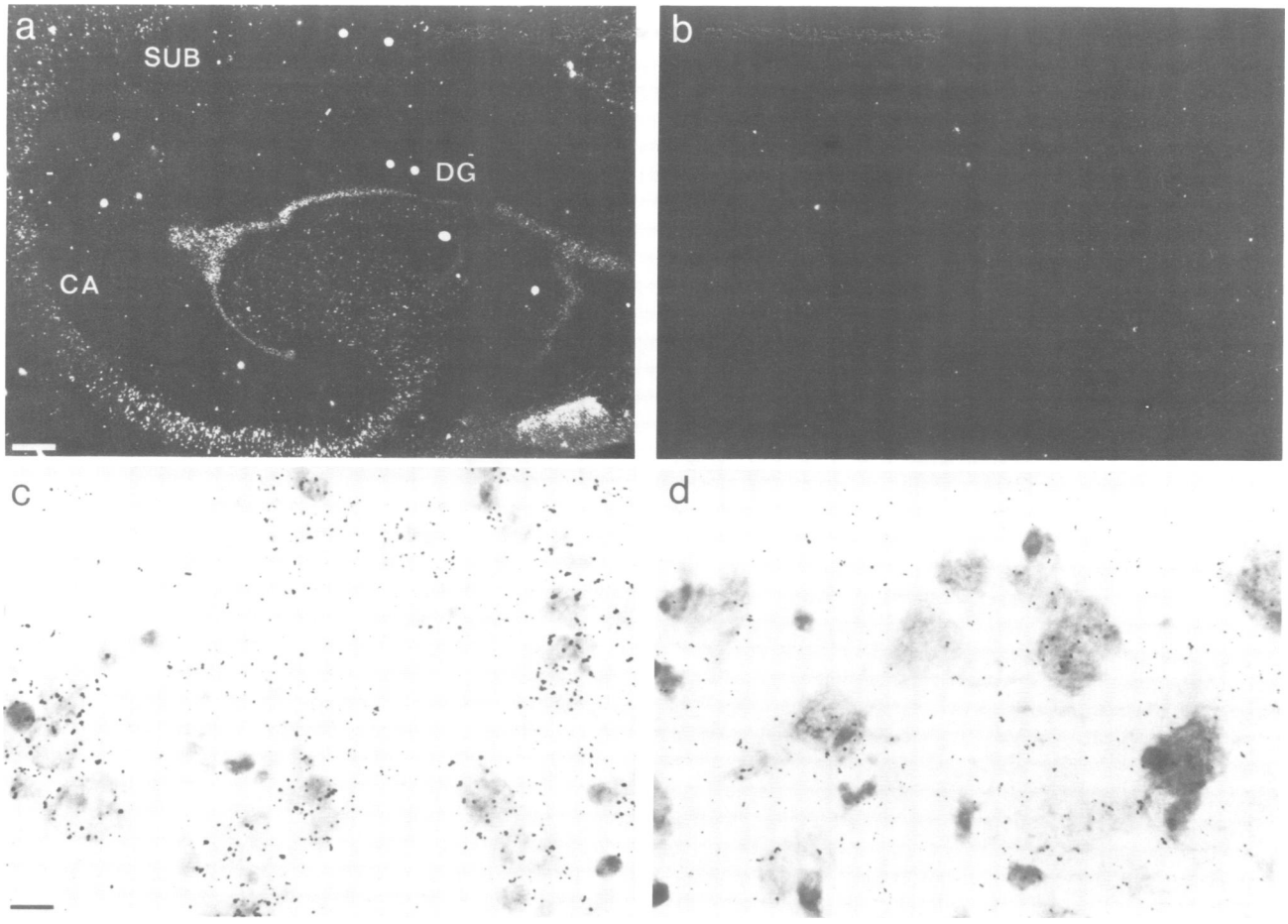

Fig. 6. Cellular localization of tau protein type I mRNA in the hippocampal formation. (**a,b**) Dark-field photomicrographs of hippocampal formation after hybridization with a probe in the anti-mRNA sense orientation (**a**) or mRNA sense orientation (**b**). (**c,d**) Light-field photomicrographs of pyramidal cells in the CA3 region of the hippocampus following hybridization with a probe in the anti-mRNA sense orientation (**c**) or mRNA sense orientation (**d**). DG, dentate gyrus; CA, cornu ammonis; SUB, subiculum. Scar bars: in (**a**) for (**a,b**), 350 μm; in (**c**) for (**c,d**), 18 μm.

to both amyloid plaques and neurofibrillary tangles. Protein sequencing has indicated that the three-repeat version of human tau protein forms a component of the protease-resistant core of the paired helical filament of Alzheimer's disease (Wischik *et al.*, 1988a). In addition, the sequence -Ile-Lys-X-Val-, which is not present in the three-repeat form, was obtained (Wischik *et al.*, 1988a). We show here that this sequence is present in the four-repeat form of human tau, with -X- corresponding to -His- (amino acids 239–242). It follows that the core of the paired helical filament of Alzheimer's disease contains at least two isoforms of tau protein. We have previously investigated, by Northern blotting, the levels of tau protein mRNA in the frontal cortex from control patients and from patients who died with a histologically confirmed diagnosis of Alzheimer's disease; no significant difference between the two groups was found (Goedert *et al.*, 1988). However that study did not distinguish between the mRNAs encoding different forms of tau protein. We therefore used an RNase protection assay to study the levels of tau protein type I and type II mRNAs in Alzheimer's disease. No difference in levels of either type I or type II mRNAs was found. Thus simple over-expression of type I or type II tau protein mRNAs is unlikely to be the cause of the formation of paired helical filaments.

Type I and type II tau protein mRNAs appear to be developmentally regulated, since type I mRNA was found to be abundant in foetal brain, where type II mRNA could not be detected. This indicates that tau is a nervous system-specific protein whose mRNA is subject to developmentally regulated alternative splicing. A developmental change in mRNA levels for the microtubule-associated protein MAP2c in rat brain has also been reported recently (Garner and Matus, 1988). Changes in the composition of tau proteins during brain development have been observed in the rat (Mareck *et al.*, 1980; Ginzburg *et al.*, 1982). Juvenile tau consists of a single major component with an apparent mol. wt of 48 000 daltons (Couchie and Nunez, 1985). The transition from this simple pattern to the more complex adult pattern of 4–6 bands starts around post-natal day 15 and is complete by post-natal day 35 (Couchie and Nunez, 1985). It appears unlikely that the tau protein type I mRNA encodes the human juvenile form, as it is expressed at high levels both at early and late stages of development, though it must be remembered that type I mRNA could contain multiple species. To date, developmental studies on tau proteins in human brain have not been performed, so it is possible that the pattern of expression in embryonic human brain may be different from that in the rat. Alternatively, type I mRNA may not be translated at early developmental stages: the existence of a translational control system regulating the

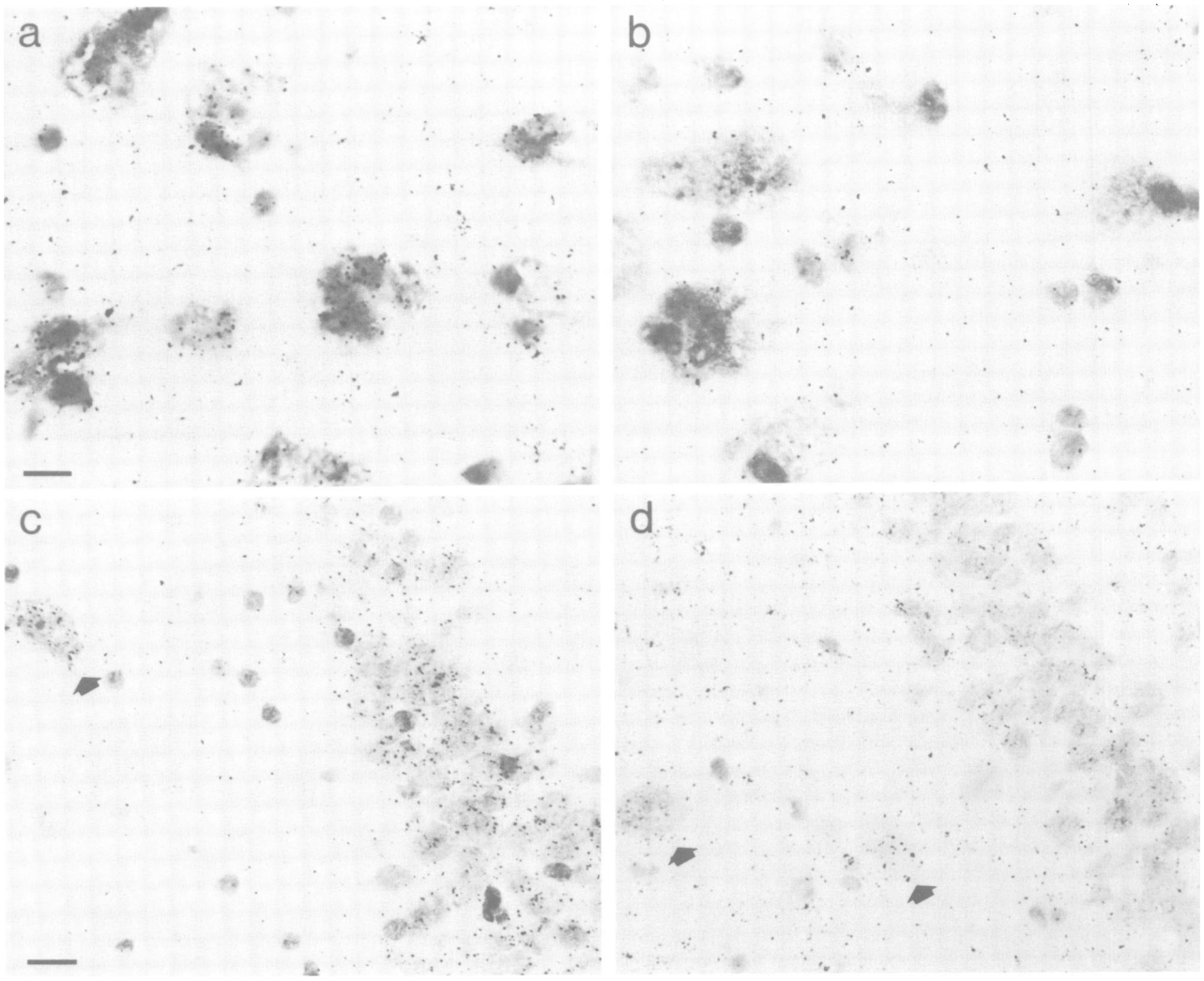

Fig. 7. Differential labelling of hippocampal neurones by tau protein type I- and type II-specific probes. (**a,b**) Light-field photomicrographs of the CA2 region of the hippocampus following hybridization with a type II-specific probe in the anti-mRNA sense orientation (**a**) or mRNA sense orientation (**b**). (**c,d**) Dentate gyrus sections hybridized with type I- (**c**) or type II- (**d**) specific probes in the anti-mRNA sense orientation. Note that the granule cells are preferentially labelled in (**c**) but not (**d**), while pyramidal cells (arrows) are labelled in both cases. Scale bars: in (**a**) for (**a,b**), 18 μm; in (**c**) for (**c,d**), 25 μm.

expression of stage-specific tau proteins has been suggested in the rat (Ginzburg *et al.*, 1982). The work reported here does demonstrate the stage- and cell-specific expression of at least some types of human tau protein mRNA.

Materials and methods

cDNA library construction and screening

A cDNA library was constructed in the insertion vector λgt10 as previously described (Goedert *et al.*, 1988) by using poly(A)$^+$ RNA prepared from the frontal cortex of a 68-year-old patient who had died without neurological or psychiatric disorders. The tissue was obtained 4 h after death. 10 μg poly(A)$^+$ RNA yielded a library of 7.4 × 10^6 clones. Replica filters of a portion of this library (100 000 plaques) were screened at high stringency with an oligonucleotide probe (5′CAAATAGTCTACAAACC3′) derived from the published human tau sequence (nucleotides 652–668) (Goedert *et al.*, 1988) and labelled with [γ-^{32}P]ATP using T4 polynucleotide kinase.

Genomic library construction and screening

A genomic library was constructed in the replacement vector λEMBL3 (Kaiser and Murray, 1985), using human cerebellar DNA. 10 μg genomic DNA yielded a library of 2 × 10^6 clones. Replica filters of a portion of this library (400 000 plaques) were screened at high stringency with a 296 bp *Pst*I/*Hin*dIII fragment derived from a cDNA clone encoding the four-repeat tau protein (nucleotides 558–853, Figure 1), labelled with [α-^{32}P]dCTP using random hexanucleotides as primers (Feinberg and Vogelstein, 1983).

DNA sequencing

Sequencing of DNA was performed by a modification (Biggin *et al.*, 1983) of the dideoxy chain termination method (Sanger *et al.*, 1977) with overlapping restriction fragments subcloned into M13mp18 or M13mp19 and synthetic oligonucleotide primers.

RNase protection assays

Human tissues were obtained <6 h after death and stored at −70°C until use. Several different brain regions and peripheral tissues were dissected from a 63-year-old patient, who had died with no neurological or psychiatric disorders. The frontal cortex was used from the same patient and from three others who had also died without neurological or psychiatric diseases (ages 72, 79 and 83 years) and from four patients who had died with a histologically confirmed diagnosis of Alzheimer's disease (ages 65, 70, 74 and 76 years). Whole brain was used from a 15-week-old human foetus. RNA extractions and poly(A) selections were performed as described previously (Goedert, 1987).

*Pst*I/*Hin*dIII fragments spanning the region where the additional repeat described here is inserted were taken from cDNA clones encoding the three- and four-repeat tau proteins. For the three-repeat form this corresponds to nucleotides 558–760 (Goedert *et al.*, 1988) and for the four-repeat form to nucleotides 558–853 (this paper, Figure 1). The fragments were subcloned into the vector pGEM2 (Promega Inc.) and transcribed using T7 RNA polymerase (Melton *et al.*, 1984). The two fragments, which are 203 (type I) and 296 (type II) nucleotides in length, differ only by the extra repeat sequence present in type II. Total RNA (40 μg) or poly(A)$^+$ RNA (2 μg) was annealed to a molar excess of ^{32}P-labelled anti-sense RNA for 5 min

at 85°C in 80% formamide, 40 mM Pipes buffer (pH 6.7), 0.4 M NaCl, 1 mM EDTA, and the hybridization was continued for at least 10 h at 45°C. Single-stranded RNA was digested by a 30-min incubation at 37°C with 40 μg/ml RNase A and 2 μg/ml RNase T1. The RNases were then inactivated by treatment with SDS and proteinase K, and the protected RNA extracted with phenol. Following ethanol precipitation the protected RNA was dissolved in 80% formamide and electrophoresed on a 6% acrylamide/8 M urea gel. The gel was dried and exposed at −70°C using a Du Pont Cronex intensifying screen.

In situ hybridization

In situ hybridization was performed by using 42mer oligonucleotide probes specific for tau type I and type II in the anti-mRNA sense orientation. The probe sequences were as follows: type I-specific, complementary to nucleotides 628–669 (Goedert *et al.*, 1988), 5′TGGTTTGTAGACTATTTGCACCTTCCCGCCTCCCGGCTGGTG3′; type II-specific, complementary to nucleotides 652–693 (this paper, Figure 1), 5′GGACTGGACGTTGCTAAGATCCAGCTTCTTATTAATTATCTG3′. Sense type I and type II-specific oligonucleotides were used as controls. Probes were 3′-end-labelled to the same specific activity (6×10^9 c.p.m./μg) using [α-^{35}S]dATP (1350 Ci/mmol) and terminal deoxynucleotidyl transferase (Deng and Wu, 1981). Prior to use, non-incorporated label was removed by centrifugation through a Sephadex G-25 column.

Brain tissues from three adult patients were dissected <4 h after death and processed as previously described (Goedert, 1987). Hybridizations were performed overnight at 42°C in 50% formamide, 4 × SSC (1 × SSC = 150 mM sodium chloride, 15 mM sodium citrate), 50 mM sodium phosphate buffer, pH 7.0, 120 μg/ml heparin, 100 μg/ml acid/base cleaved salmon sperm DNA and 100 μg/ml poly(A). The probe concentration was 10 000 c.p.m./μl. The sections were washed for 1 h in 2 × SSC at room temperature, followed by a 1-h wash in 1 × SSC at 55°C. The slides were dehydrated in 70% ethanol, dried and dipped in Ilford K-5 emulsion and kept at −20°C for 4 weeks. They were developed in safe-light conditions using Kodak D19 developer, stained with cresyl violet, dehydrated, cleaned in Histoclear and mounted using Depex.

Acknowledgements

We thank Drs A.Klug, C.M.Wischik and W.Wisden for helpful discussions. Some of the tissues used came from the collection of the Medical Research Council, Cambridge Brain Bank Laboratory. This work was supported in part by a fellowship from Fidia Research Laboratories (M.G.S.), a fellowship from the Commission of the European Community (M.C.P.) and by Imperial Chemical Industries PLC and ICI Americas, Inc.

References

Aizawa,H., Kawasaki,H., Murofushi,H., Kotani,S., Suzuki,K. and Sakai,H. (1988) *J. Biol. Chem.*, **263**, 7703–7707.

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

Biggin,M.D., Gibson,T.J. and Hong,C.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963–3965.

Binder,L.I., Frankfurter,A. and Rebhun,L.I. (1985) *J. Cell Biol.*, **101**, 1371–1378.

Brion,J.P., Passareiro,H., Nunez,J. and Flament-Durand,J. (1985) *Arch. Biol.*, **95**, 229–235.

Carlier,M.F., Simon,D. and Pantaloni,D. (1984) *Biochemistry*, **23**, 1582–1590.

Cleveland,D.W., Hwo,S.Y. and Kirschner,M.W. (1977a) *J. Mol. Biol.*, **116**, 207–225.

Cleveland,D.W., Hwo,S.Y. and Kirschner,M.W. (1977b) *J. Mol. Biol.*, **116**, 227–247.

Couchie,D. and Nunez,J. (1985) *FEBS Lett.*, **188**, 331–335.

Deng,G. and Wu,R. (1981) *Nucleic Acids Res.*, **9**, 4173–4188.

Drubin,D.G., Caput,D. and Kirschner,M.W. (1984) *J. Cell Biol.*, **98**, 1090–1097.

Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.

Garner,C.C. and Matus,A. (1988) *J. Cell Biol.*, **106**, 779–783.

Ginzburg,I., Scherson,T., Giveon,D., Behar,L. and Littauer,U.Z. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4892–4896.

Goedert,M. (1987) *EMBO J.*, **6**, 3627–3632.

Goedert,M., Wischik,C.M., Crowther,R.A., Walker,J.E. and Klug,A. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4051–4055.

Grundke-Iqbal,I., Iqbal,K., Quinlan,M., Tung,Y.C., Zaidi,M.S. and Wisniewski,H.M. (1986) *J. Biol. Chem.*, **261**, 6084–6089.

Horio,T. and Hotani,H. (1986) *Nature*, **321**, 605–607.

Kaiser,K. and Murray,N.E. (1985) In Glover,D.M. (ed.), *DNA Cloning–A Practical Approach*. IRL Press, Oxford, Vol. 1, pp. 1–47.

Kosik,K.S., Joachim,C.L. and Selkoe,D.J. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 4044–4048.

Lee,G., Cowan,N. and Kirschner,M. (1988) *Science*, **239**, 285–288.

Maccioni,R.B., Rivas,C.I. and Vera,J.C. (1988) *EMBO J.*, **7**, 1957–1963.

Mareck,A., Fellous,A., Francon,J. and Nunez,J. (1980) *Nature*, **284**, 353–355.

Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.

Neve,R.L., Harris,P., Kosik,K.S., Kurnit,D.M. and Donlon,T.A. (1986) *Mol. Brain Res.*, **1**, 271–280.

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

Shapiro,M.B. and Semapathy,P. (1987) *Nucleic Acids Res.*, **15**, 7155–7175.

Weingarten,M.D., Lockwood,A.H., Hwo,S.Y. and Kirschner,M.W. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 1858–1862.

Wischik,C.M., Novak,M., Thøgersen,H.C., Edwards,P.C., Runswick,M.J., Jakes,R., Walker,J.E., Milstein,C., Roth,M. and Klug,A. (1988a) *Proc. Natl. Acad. Sci. USA*, **85**, 4506–4510.

Wischik,C.M., Novak,M., Edwards,P.C., Klug,A., Tichelaar,W. and Crowther,R.A. (1988b) *Proc. Natl. Acad. Sci. USA*, **85**, 4884–4888.

Wood,J.G., Mirra,S.S., Pollock,N.J. and Binder,L.I. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 4040–4043.

Received on October 17, 1988; revised on November 21, 1988