

# The Microtubule Binding Domain of Microtubule-associated Protein MAP1B Contains a Repeated Sequence Motif Unrelated to that of MAP2 and Tau

Melanie Noble, Sally A. Lewis, and Nicholas J. Cowan

Department of Biochemistry, New York University Medical Center, New York 10016

**Abstract.** We report the complete sequence of the microtubule-associated protein MAP1B, deduced from a series of overlapping genomic and cDNA clones. The encoded protein has a predicted molecular mass of 255,534 D and contains two unusual sequences. The first is a highly basic region that includes multiple copies of a short motif of the form KKEE or KKE<sub>V</sub> that are repeated, but not at exact intervals. The second is a set of 12 imperfect repeats, each of 15 amino acids and each spaced by two amino acids. Subcloned fragments spanning these two distinctive regions were expressed as labeled polypeptides by translation in a cell-free system *in vitro*. These polypeptides were tested for their ability to copurify with unlabeled brain microtubules through successive cycles of polymeriza-

tion and depolymerization. The peptide corresponding to the region containing the KKEE and KKE<sub>V</sub> motifs cycled with brain microtubules, whereas the peptide corresponding to the set of 12 imperfect repeats did not. To define the microtubule binding domain *in vivo*, full-length and deletion constructs encoding MAP1B were assembled and introduced into cultured cells by transfection. The expression of transfected polypeptides was monitored by indirect immunofluorescence using anti-MAP1B-specific antisera. These experiments showed that the basic region containing the KKEE and KKE<sub>V</sub> motifs is responsible for the interaction between MAP1B and microtubules *in vivo*. This region bears no sequence relationship to the microtubule binding domains of kinesin, MAP2, or tau.

**M**ICROTUBULES prepared *in vitro* by successive cycles of assembly and disassembly consist largely of  $\alpha$ - and  $\beta$ -tubulin, together with a number of other proteins that are collectively defined as microtubule-associated proteins (MAPs)<sup>1</sup> (for reviews, see 32, 35, 52). Because neuronal cells are rich in microtubules, the best characterized MAPs are those from brain. Conventionally, these MAPs have been classified according to their molecular mass: the high molecular mass MAPs, MAP1A, MAP1B, MAP1C, and MAP2 (in the range 200–350 kD), and the low molecular mass MAPs, or tau proteins, a heterogeneous group of proteins in the range 35–40 kD (51). MAP3 and MAP4 are less abundant than these; MAP3 is primarily expressed in astroglia and some neurons (21), while MAP4 is expressed in glia, vascular, and some other tissues (37). With the exception of MAP1C, which has been identified as cytoplasmic dynein (38, 49) the function of brain MAPs is essentially unknown, though many have been shown to promote microtubule assembly upon addition to purified tubulin *in vitro*.

Among the high molecular mass MAPs, MAP1B (7) (which is referred to variously as MAP1.2 [16], MAP1(X) [9], and MAP5 [40]) differs from the others in its high abundance, its prominence both in association with microtubules

and in soluble form, and the relative inefficiency with which it cycles with tubulin, particularly during the first cycle. Nonetheless, MAP1B is generally regarded as an authentic MAP because it does cocycle with purified tubulin and promote microtubule assembly *in vitro* (40) and because antisera specific for MAP1B (both polyclonal and monoclonal) identify characteristic microtubule networks in neuronal (and other) cells (6, 12). MAP1B is expressed in the axons and dendrites of neurons, as well as in glia and other cells. Because it is especially prominent in axons during their initial outgrowth (40, 43), it has been suggested that MAP1B plays a role in neurogenesis.

Several years ago, we isolated a set of cDNA clones encoding a portion of a MAP1 subspecies (28). Subsequently, an antiserum (3d2) raised against a fusion protein derived from one of these clones was shown to uniquely recognize MAP1B on Western blots of brain microtubules and of total protein from PC12 cells (1). The same antiserum also gives a staining pattern identical to that of other anti-MAP1B antisera on brain sections (A. Matus and N. J. Cowan, unpublished observations), thereby proving that our original set of cDNA clones were derived from MAP1B mRNAs. We have now obtained two genomic clones and further overlapping cDNA clones that together encode the entirety of MAP1B. Here we report the complete sequence of MAP1B, deduced from these clones. The encoded protein has a molecular mass of 255,

1. Abbreviation used in this paper: MAP, microtubule-associated protein.

534 D and contains two unusual sequences. One is a set of 12 imperfect repeats that do not occur in any other sequenced protein, each 15 amino acids long, and spaced by 2 amino acids. The other is a highly basic region with many copies of the sequence KKEE and KKE<sub>v</sub>, repeated but not at fixed intervals. We show that it is this latter region which is responsible for the binding of MAPIB to microtubules both in vitro and in vivo.

## Materials and Methods

### cDNA and Genomic Clones

cDNA from 5-d-old mouse brain mRNA was prepared as described (26), except that half was ligated to Eco RI linkers (Boehringer Mannheim Biochemicals, Inc., Indianapolis, IN) and half to Xba I linkers, and each portion cloned into the vector λGem (Promega Biotec, Madison, WI). The resulting libraries were screened (4) as described in the text, using purified cDNA restriction fragments <sup>32</sup>P-labeled by nick translation (41). Hybridizing clones were subcloned into pUC, restriction-mapped, and, in some cases, sequenced. Sequencing of the complete set of overlapping clones on both strands was accomplished by subcloning bal31-generated fragments (27) into M13, and using the dideoxy chain terminator method (42). Four overlapping genomic clones containing the 5' end of the MAPIB gene were obtained by screening a mouse genomic library in the EMBL4 vector, kindly provided by P. D'Eustachio (New York Medical Center). The cosmid clone C (Fig. 1), including the 6-kb MAPIB exon, was isolated by screening a library (provided by S. Tonegawa, Massachusetts Institute of Technology) with clone 141A (Fig. 1).

### Constructs for Translation, Fusion Proteins, and Transfection Experiments

Two constructs encompassing repeated sequences were made for in vitro transcription, translation, and cycling experiments as follows: (a) clone 72 was digested with Xho I, blunt-ended with the Klenow fragment of DNA polymerase I, cut with Bgl II, and the coding fragment (Fig. 1, 5' arrow) cloned into the transcription vector pGEM3 (Promega Biotec) cut with Hinc II and Bam HI; (b) clone 141A was digested with Pst I and Eco RI, and the purified fragment (Fig. 1, 3' arrow) cloned into pGEM3 cut with the same restriction enzymes. For both constructs, an in-frame translational start signal was provided by the pGEM vector. The insert from the first of these pGem clones, including the initiator AUG, was cloned into a pSV vector (34) for expression in cultured cells (Fig. 1, construct X). In addition, constructs were made from the 5' and 3' end of the set of MAPIB cDNAs to place the

translational start and stop signals: (a) the insert from clone 12X was subcloned into the Xba I site of pGEM1; and (b) the Sac I/Eco RI fragment was cloned into pGEM3 cut with Acc I, blunt ended with Klenow and then digested with Eco RI; this procedure supplied an AUG codon for translational initiation in frame with the MAPIB coding sequence.

The full coding sequence of MAPIB was assembled, first by joining clones 12X and 1R at an internal Pflm I site, then by digesting the resulting construct with Xho I and Eco RI, and finally by adding both the Xho I/Ssp I fragment from cosmid clone C and the Ssp I/Eco RI fragment from clone 166 in a three-way ligation. The clones and cloning sites are depicted in Fig. 1. The full cloning region was subcloned into a pSV vector using a unique Sal I site in the 5' flanking pUC polylinker, and a unique Dra I site in the 3' untranslated region. Deleted versions were obtained by (a) joining a Msc I site to a Pvu I site blunt ended with T4 DNA polymerase (Fig. 1, 1B-H); (b) deleting an ~1-kb Msc I fragment by self-ligating a gel purified fragment from a partial Msc I digest (Fig. 1, 1B-K); and (c) by joining the Xho I and Bg II sites, by blunt ending each with the Klenow fragment of DNA polymerase I (Fig. 1, 1B-X).

To raise an antiserum specific for the second set of repeats, a Pst I/Eco RI fragment (Fig. 1, 3' arrow) from clone 141A was subcloned into a pATH vector (29) that was used to express a fusion protein whose amino terminal part is *Escherichia coli* TrpE, and whose carboxyterminal part is derived from MAPIB. Fusion protein excised from SDS polyacrylamide gels was used to immunize rabbits. The resulting serum is called YXY; its specificity is demonstrated in Fig. 7.

### Transcription, Translation, and Microtubule Cycling Experiments

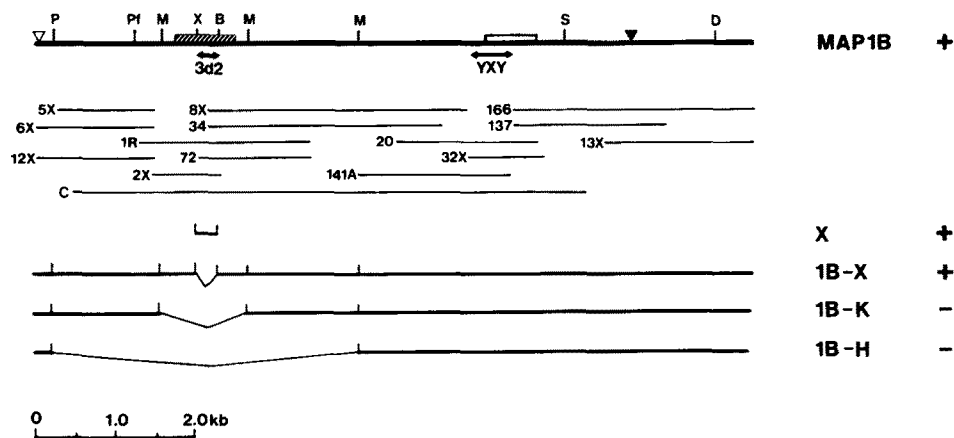
pGEM subclones were linearized at the 3' end and transcribed in vitro using SP6 polymerase as described (33). The capped mRNAs were translated in rabbit reticulocyte lysate (Promega Biotec) supplemented with either [<sup>35</sup>S]methionine or [<sup>3</sup>H]lysine. Translation products were resolved on SDS polyacrylamide gels, which were fluorographed, dried, and exposed to film. Microtubule cycling experiments were performed as described (30).

### Hybrid Selection and Primer Extension

Clone 5X (Fig. 1) was used to select MAPIB-specific mRNA from total mouse brain RNA (100 μg/2.1-cm nitrocellulose filter) by methods described previously (17). The selected RNA was used in a primer extension reaction containing an antisense oligonucleotide, 5'CCTGAGAGAAGT-GTTCCT3' (corresponding to nucleotides 17-34 of the sequence shown in Fig. 4), <sup>32</sup>P-end labeled with polynucleotide kinase. Reaction conditions and analysis of the extended product were as described (3).

### Transfection Experiments and Immunofluorescence

Hela cells were grown in DME supplemented with 5% FCS (Hyclone



**Figure 1.** MAPIB clones and constructs. The top line is a composite map of our set of MAPIB clones. The hatched and open rectangles represent regions encoding the KKEE/KKE<sub>v</sub> repeated motifs and the 15-amino acid repeats, respectively (see Fig. 4). The start codon (▽) and stop codon (▼) are indicated, as are some salient restriction sites: P, Pvu I; P<sub>1</sub>, Pflm I; M, Msc I; X, Xho I; B, Bgl II; S, Ssp I; D, Dra I. The arrows indicate fragments subcloned for in vitro cycling experiments, and for the generation of antisera 3d2 (1) and YXY (see Materials and Methods). Thin lines show the extent of cDNA and genomic clones which span the coding region of MAPIB. Clones with names ending in X came from an Xba I-linked cDNA library, while those with names ending in R came from an Eco RI-linked cDNA library (see Materials and Methods). Other clones have been previously described (28). Deletion constructs created for transfection experiments are indicated by heavy lines, and the results of these experiments are shown at the right: +, microtubule binding; -, no microtubule binding.

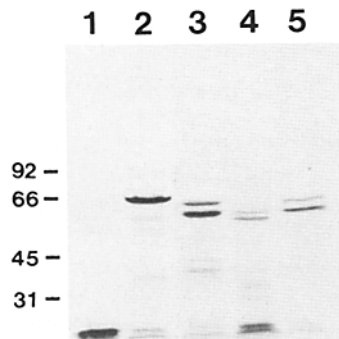
+

Laboratories, Logan UT) in a 10% CO<sub>2</sub> atmosphere; 3T3 cells were grown under the same conditions, except 10% FCS was used. Cells growing on glass coverslips were transfected with DNA by the procedure of Chen and Okayama (10), and after 72 h fixed in one of three ways: (a) with paraformaldehyde (36); (b) with cold methanol for 6 min; or (c) with cold methanol following detergent extraction with 0.2% Triton X-100 in a microtubule stabilizing buffer (47). Immunofluorescence was performed using one of the two rabbit anti-MAP1B antisera, 3d2 (1) or YXY (this paper). The regions of MAP1B to which these sera bind are shown in Fig. 1. A mouse anti- $\alpha$ -tubulin monoclonal antibody (Amersham Corp., Arlington Heights, IL) was used for double labeling; FITC-conjugated goat anti-rabbit and rhodamine conjugated goat anti-mouse second antibodies were from Boehringer Mannheim Biochemicals. All antibodies were used at a dilution of 1:500.

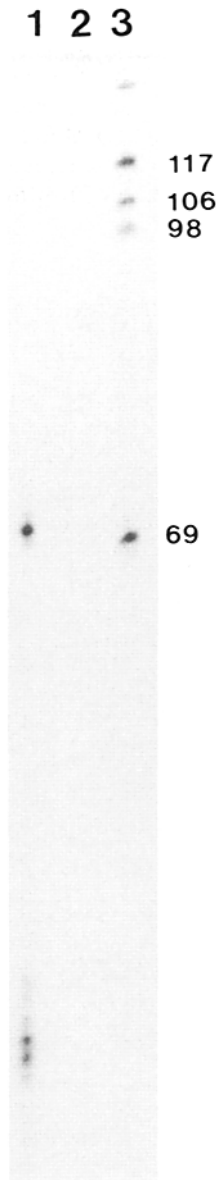
## Results and Discussion

### Isolation of Clones Encoding MAP1B

With a view to elucidating the structural and functional domains of MAP1B and its relationship to other MAPs, we first sequenced our original set of overlapping MAP1B-encoding cDNA clones. This analysis showed that the set of clones contains a long open reading frame of 5.4 kb, extending from the 5' end of the overlapping set to a location 1.5 kb from the 3' end. Because this open reading frame encodes a polypeptide of at most  $\sim$ 198,000 D, we extended our collection of overlapping cloned cDNAs in the 5' direction by further "walking" experiments. To facilitate the search for overlapping clones, we constructed two mouse brain libraries using different linkers for cloning. These were first screened with a small fragment from the 5' end of our 5' most overlapped cDNAs (clone 72; reference 28), and twice more, each time with fragments obtained from the new set of clones. Representative examples from the complete set of overlapped cDNA clones are depicted in Fig. 1.



**Figure 2.** Determining the placement of the start and stop codons within the MAP1B transcript. Fragments from the 5' and 3' end of our MAP1B cDNA set were cloned into transcription vectors; capped RNA was synthesized and translated in a cell free system containing [<sup>35</sup>S]methionine (see Materials and Methods). The resulting labeled translation products were resolved on an 8.5% SDS polyacrylamide gel. Lane 1, no RNA control; lane 2, RNA from clone 12X (nucleotides 1-1,574); lane 3, RNA from clone 137, linearized with Eco RI (Fig. 4, nucleotide 6,304-3' end of clone); lane 4, RNA from clone 137 linearized with Kpn I (Fig. 4, nucleotides 6,303-7,333); lane 5, RNA from clone 137 linearized with Nae I (Fig. 4, nucleotides 6,303-7,454). In lanes 3-5, the slowest migrating species results from initiation at an AUG triplet provided by the vector (see Materials and Methods), whereas the second prominent band presumably results from initiation at an internal AUG codon.



**Figure 3.** The MAP1B transcriptional initiation site lies 35 nucleotides upstream from sequences represented by overlapping cloned cDNAs. A synthetic antisense oligonucleotide corresponding to nucleotides 17-34 of the sequence shown in Fig. 4 was end-labeled with polynucleotide kinase and used in a primer extension reaction with RNA hybrid selected with clone 5X (Fig. 1) (see Materials and Methods). The reaction products were analyzed on a 6% polyacrylamide sequencing gel. Lane 1, selected RNA; lane 2, control (*E. coli* tRNA); lane 3, molecular size markers (shown, in nucleotides, right).

All attempts to extend this cloned set further in the 5' direction resulted in the isolation of numerous unrelated and nonoverlapping sequences that were presumed to be cloning artifacts generated by the unavoidable adventitious ligation of unrelated cDNA sequences during construction of the cDNA library, or by "hairpinning" of cDNA generated during the synthesis of the second strand. Sequence analysis of the overlapping 5' cDNA clones extended the long open reading frame to 7.5 kb. However, because this open reading frame extended to the extreme 5' end of our overlapped cDNAs, we cannot be certain whether the first AUG triplet indeed encodes the initiator methionine residue; it is possible that further upstream sequences encoding additional aminoterminal MAP1B residues might be unrepresented in our cDNA libraries. However, we did three further experiments which strongly suggest that translation of MAP1B does begin at the first AUG triplet found in our cDNA clones. Firstly, a fragment extending from the extreme 5' end of the overlapped cDNAs (Fig. 1) was cloned into a vector

AGAGAGCGCGGAGAGAACTTCTCTCAGCGTTGACAGAGCGCGGAGGAGG 60  
H A T 3  
CGTGGTGTGGAAGACCAACCGGAGCCATCGGACGATCGCAACCGCGCGCT 120  
V V V E A T E P E P B S B I G N P A A S 23  
CCACTCGCCAGCCTTCCGACCGCTTCCGACGCAAGTCTACTTCTGCTGGTGG 180  
T B P S L S N R F L D S K F Y L L V V V 43  
TCGCGAGACCGTAACCGAAGACCTGAGCGCTGCATCGCAACATCGAGCTGGGA 240  
G E T V T E E H L R R A I G N I E L G I 923  
TCGACTCGGCAACAACTGATGAGTCAACTTGAACGAGAACTCAAACTTTCC 300  
R B W D T N L I E D Q E L K L F V 83  
TATCTCGACTCTGCAAGATCTCTGCAAGTTCAGGACAAAGATCTTCAATCCC 340  
S R H B A R F S P E V P G Q K I L M H R 103  
GAAGTACCTTGAAGAACTGTTCTGATCAACCTTCAGATGAAGCGCTCAGCACTE 420  
S D V L E T V V L I N P B D E A V B T E 123  
AGTGGCTTGTGATGATCACTGACCTGCCGCCATAAATGCTGGTCTCAGCGGACST 480  
V R L M I T D A A R M K L L V L T G Q C 143  
GCTTTGAGAACTCGGAGCTCATCTGCAAGTCTGCTCTTCTGCTTCCAGAACTCA 540  
F E N T G E L I L Q G S B F B F Q N F I 163  
TAGAGATTTTCAAGATCAAGAGATGTTGATGCTTCAAGCAACCCCACTGCAACA 600  
E I F T D Q E I G E L L S T H P A N K 103  
AAGCCAGCTCACCTCTCTGCTCCGAGGAGGAGCTGGAAGAACTCCAACTTGACA 660  
A S L T L F C P E E G D M K N S N L D R 703  
GACACAATCTCAAGACTTCACTCAAGATTAAGCTCAAGCTGCTTCTTCCGAGAAA 720  
H N L Q D F I M I K L N B A S I L P E M 223  
TGGAGGAGCTTCTGAGTCAAGGATATCTCTGAGAGCTGTAAGAGTCCCCTCTCC 780  
E G L B E F T C C E S E V E P B P F 243  
TTGACATCTGGAACCCCAACTCGGCGGCTTTGGAAGCTCTCAAGCTTCTGCTTT 840  
D I L E P P T B G G F L K L B K P C C Y 263  
ACATTTCCAGGCGCTGGGAGCTCGGCTGTTTCAAGTGAAGGATCAAGTGC 900  
I F P G G R G D S A L F A V N G F N H L 283  
TCATCAAGGAGGATCAGAGAAAGTCTGCTTCCGAGCACTCATCGGCACTTGACC 960  
I N G G S E R K B C T F M K L I R H L D R 303  
GTGTGACTCCATCTGCTCACACATTTGGGATGACAATTTGCTGGATCAATAGCA 1020  
V D S I L L T H I G D D N L P G I N S M 323  
TGTGAGCGCAAGATTCAGAGCTGGAAGGAGCGGCTCCAGGAGTCCAGCAGCA 1080  
L Q R K I A E L E E R B Q S T S N S 343  
GTGACTGATGAAACCTCATCTCCCTGAGCTGGGCTTGTGTTTCAAGTACTCE 1140  
D M K N L I S P D L G V V F N V P E 363  
AAAACCTGAAAGACCCAGAACCAACATCAAGATGAAGAGAGCAGATAAGAGCATGCT 1200  
N L K D P E P N I K M K R S I E E A C F 383  
TCACTCCAGTACTTAAACAACTGCTGATGAAACAGGCTTATTTAGAAATGTAG 1260  
T L Q Y L N K L S M K P L F R B S V G 403  
GCAACCACTTGAAGCTGATCTGCTTCCAGAGATGGAGTGGGAGGCTTGAAGAT 1320  
N T I E P V I L F Q K H G V G K L E H Y 423  
AGTGTCTAAACCAAGTCAAGAGCAGCAAGGAAATGCAATTTTATGCAAGTGGACTE 1380  
V L N P V K B S K E M Q Y F M Q Q M T G 443  
GCACCAAGAGCAAGGCAACTAATCTGCAAGTGGTCAAGAGTAGACATCCGGA 1440  
T N K D K A E L I L P N G Q E V D I P I 1640  
TTTCTACTTAACTCAGTCTCATCTTGTGTTGTTGCTCCGCAACCCGCTGAGA 1500  
S Y L T S V B B L I V W H P A N P A E K 483  
AAATCTTGGGTTCTGTTCCGGAACAGTACCACTCAACATCTAGAGGACTGG 1560  
I I R V L F P P G N S T Q V N I L E G L E 503  
AGAACTCAACATAGATTTCTAAAGCAGTCCGCAACCCAAAAGATCTCACTE 1620  
K L K H L D F L K Q P L A T Q K D L T G 523  
GCCAGGTCGCCACCCCGCAGTGAACAGGCTCAACTGAAACAGCGGCTGACAGCGAG 1680  
G V P T P P V K Q V K L K Q R A D S R E 543  
AGATCTGAAAGCAGCAAGCAAGTCAAGTAAATCTGTGAGGAGGAGTCCAAAG 1740  
S L K P A P A T K P V A B K S V R K E S K E 563  
AGGAAACCCCTGAAGTCAACAAGTCAAGTGAAGGAAACCAAGATTTGAAGAGCA 1800  
E T P E V T K T B V T K E K T P K V E S G 583  
AAGAGAAAGTGTAGTGAAGAAAGCAAGGCTGTAAGAGCAGAAAGTAACTTCAAGTGA 1860  
E K V L V K K D K P V K T E S K P S V T 603  
CAGAAAGAGGTTCCAGCAAGGAGGAGTCCCAAGTGAAGGAGGAGTCCCTGAGA 1920  
E K E V S S K E E Q S P V K A E V A E K 623  
AGCAAGCCTGAGAGCAACCCAGTCAAGGAGCAAGTGGTGAAGGAGGAAATAA 1980  
Q A T E S K P K V T K D K V V K K E I K 643  
AGCAAACTCGAGAAAGAAAGAGGAGGAGCCCAAGAAAGAGTGGTTAAGAGGAG 2040  
T K L E E K K E E K P K K E V V K K E D 663  
ACAAAACCCCTCAAGAAAGCAGAGGCTCAAGAAAGGAGGAGTGAAGAGGAGTCA 2100  
K T P L K K D E K P R K E E V K K E I K 683  
AGAAAGAAATCAAAAGAGAGAGAGGAGTGAAGAGGAGTGAAGAGGAGAAAC 2160  
K E I K K K E I K K E E V K K E I K 703  
CCCTGAGGATGCAAGAGGAGTGAAGAAAGGAGGAGGAGGAGTAAAGAGGAG 2220  
L K D A K K E E V K K E E K K E V K K E E 723  
AGAGGAGCCCAAGAAAGTCAAGAGATCTCAAGGACATAAGAAATCAACTCTC 2280  
K E P K K E I K K I S K D I K K S T P Q 743  
AGTCAAGCAAGAGGAGTCTGCTTAAACCAAAGTGAAGAGGAGGAGGAGTCCA 2340  
S D T K K P A L K P K A V K K E E S T 763  
CTAAGAGGAGCTTCTGCTGAGAACTCAAGGACAGGAGGAGGAGGAGTCAATTA 2400  
K K E P L A A A G K L K D K G K V K V I K 783  
AAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460  
K E G K T T E A A T A V G T A G T A T A 803  
CTGCGTGTGGCTGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2520  
A V V A A A G I A A S G P V K E L E A E 823

AGAGTCCCIATGTCTCCCTCCGAGGATCAACCAAGGACTTTGAGGAGCTAAAGGCTE 2580  
R S L M B S P E D L T K D F E E L K A E 843  
AGAGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640  
E D V A K D I K P Q L E L I E D E E K 863  
AACTGAGGAAACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2700  
L K E T Q P G E A Y V I Q K E T E V B K 883  
AAGTTCGCTGAGTCAACCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2760  
G S A E S P D E G I T T T E G E G E 903  
AGCAACCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820  
Q T P E E L E P V E K Q G V D D I E K F 923  
TTGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2880  
E D E G A G F E E S S E T G D Y E E K A 943  
CGGAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940  
E T E A E E P E E D G E S S A S 963  
CCTTAAGCAGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3000  
S K H S P E D D E S A K A E A D V H L 983  
TCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3060  
K E K R E S V V S G D D R A E E D M D D 1003  
ATGTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3120  
V L E K G E A E Q S E E E G E D K A 1023  
CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3180  
E D A R E E S Y E P D K T E A E D Y V H 1043  
TGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3240  
A V A D K A A E A G V T E E Q Y G L 1063  
GCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3300  
T S A K Q P G I Q S P S R E P A S I H 1083  
ATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3360  
D E T L P G G S E S E A T A S D E E N R 1103  
GGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3420  
E D Q P E E F T A T S G Y T C A G T I E 1123  
TATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3480  
S S E P T P M D E M S T P R D V M S D E 1143  
AGACCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3540  
T M N E E T E S P S Q E F V N I T K Y E 1163  
AGTCTGCTGATTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3600  
S L Y S Q E Y S K P A V A B F N G L S 1183  
CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3660  
E G S K T D A T D G K D Y N A S A S T I 1203  
TATCACCACTTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3720  
S P P S S M E E D K F S K S A L R D A Y 1223  
ACTGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3780  
C S E E K E L K A S A E L D I K D V S G 1243  
ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3840  
E R L S P A K S P S L S P S P S P I E 1263  
AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3900  
K T P L G E R S V N F S L T P N E I K V 1283  
TGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3960  
S A E G E A R S V S P G V T Q A G V T G 1303  
AACATTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4020  
H C A S P E E K T L E V V S P S Q S V T 1323  
CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4080  
G S A G H T P Y Y Q S P T D E K S S H L 1343  
TACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4140  
P T E V S E M A Q A V P V S F E F S E A 1363  
CCAAAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4200  
K D E N E R A S L S P H D E P V P D S E 1383  
AGTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4260  
S P V E K V L S P L R S P P L L G S E 1403  
CCCCATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4320  
P Y E D F L S A D S K V L G R R S A E S 1423  
CCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4380  
F E G K N G K Q G F P D R E S P V S D L 1443  
TGACTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4440  
T S T G L Y Q D K Q E E K S I S E G T V S 1463  
TAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4500  
K E D F G P E K K T S D V E T M S S H E 1483  
CAGCTAGCTTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4560  
A L A L D E R K L G G D V S P T Q I D V 1503  
TCAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4620  
S Q F G S F K E D T K M S I S E G T V S 1523  
CAGCAAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4680  
K S A T P V D E G V A E D T Y S H M E 1543  
AGGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4740  
G V A C S V S T A S V A T S S F P E P T T 1563  
CAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4800  
D D V S P S L H A E V G S P H S T E V D 1583  
ATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4860  
D S L S V S V V Q T P T T F Q E T E M S 1603  
CTCATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4920  
P S K E E C P R P M S I S P P D F S P K 1623  
AGACAGCCAACTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4980  
T A K S R T P V Q D H R S E S M S I 1643  
TTGAGTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5040  
E F G Q E S P E H S F A M D F S R Q S P 1663

CAGATCACCTACTCTG66TGCACGTGCTTCATATCACTGAAAACGG66CAACTGAA6 5100  
 D H P T L G A S V L H I T E N G P T E V 1683  
 T6GACTATAGTCCCTGTGATTCAG6GACTCTAGT7TTC6GATAGATCCCGCAAC66 5160  
 D Y S P C D I Q O S S L S H K I P P T E 1703  
 AAG6GACCTCTCACCCAGGATAAGTGTCTGAGCTCATCTCTG6TCTCAG6T6 5220  
 E P S Y T Q D N D L S E L I S V S Q V E 1723  
 A66CTTCC66TCCACCTCTTCA6CTCACACCCCTTCTCAGAT6ACCTCTCTCTCAG6 5280  
 A S P S T S S A H T P S I A S P L Q E 1743  
 AAG6ACTCTCTGATGTGTCTCCAG6AGAT6TCTTATATGCTCAGTCTG6T 5340  
 D T L S D V V P P R E H S L Y A S L A S 1763  
 CT6AAAAGT6CAG6GCT6G6A6G6A6GCTCTCTC66AATCT6A1ATTCTCCAC 5400  
 E K V Q S L E G E K L S P K S D I S P L 1783  
 TCACCCAG6AG6TCTCACCCTTATATTCACCTG6CTTTCAGATTTCACTCTG6AG 5460  
 T P R E S S P L Y S A F S D S T S A A 1803  
 CTAAG6ACT6CAG6AGCTCAG6AGCTCTCTCTCCACCAATAGAT6CAG6CAAC6 5520  
 K E T A A A H G A S S S P P I D A A T A 1823  
 CAG6G6CCTAT6GCTCTC6GCTCTCAATGCTT7T6ATACAA6CAG6CAATCATCTG6CCT 5580  
 E P Y G F R S S M L F D T H Q M H L A L 1843  
 T6AATAG6ATTT6ACTACATAGTGT6G6A6A6CAG6G6A6G6A6CAG6C6G6 5640  
 N R D L T T S S V E K T P G S G K T P G D 1863  
 ACTTTA6CTAT6GCTCAAAAAGCT6G6AT6G6C6G6G6ATCTCAG6T6A6A6AGAT 5700  
 F N Y A Y Q K P E M A A G S P D E E D Y 1883  
 AT6ACT6AATCTCA6G6A6A6A6GCTCTCC6G6C6C6C6G6CT6G6T6G6T6ACT6ATG 5760  
 D Y E S Q E K T I R T H O V V R Y Y Y E 1903  
 A6A6G6CAG6A6A6C6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6 5820  
 K T E R T I K S P C D S G Y S Y E I I E 1923  
 A6A6G6CAG6A6A6C6C6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6 5880  
 K T I K T P E D G G Y T C E I T E K T I 1943  
 CTC6G6C6C6C6A6A6G6G6T6G6T6T6CATAT6G6AT6CAG6G6A6A6C6G6A6C6G6 5940  
 R T P E E G G Y S Y E I S E K T I T R T P 1963  
 CT6A6AT6A6G6G6CTACACCTAT6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6 6000  
 E V B G Y T R B P D I S A Y C V E T H E K 1983  
 TT6G6A6T6G6T6AT6G6A6C6T6G6G6C6C6C6C6A6C6T6G6G6A6T6G6A6C6T6 6060  
 S N S Y D D T I F D S G M Y L S D C S Y S 2003  
 CCTAT6A6 6120  
 Y E T E K I T S F P E S E S Y S Y E I 2023  
 CA6GTCA6 6180  
 S T K T R B P D I S A Y C V E T H E K 2043  
 AGATC6A6 6240  
 I T K T P G A S T Y S Y E T S D R C Y T 2063  
 CTC6G6A6 6300  
 T E K K S P S E A R Q D V D L C L V S S 2083  
 CCT6T6A6ATCA6G6ATCCCA6G6C6G6A6G6CTCTG6CCTTCTCTTAA6TCCA6ATCCT 6360  
 C E F K H P K T E L S P S F I N P N P L 2103  
 TT6AGT6T7T6G6T6G6A6A6G6C6C6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6 6420  
 E W F A G E E P T E E S E K P L T Q S G 2123  
 G6G6A6C6 6480  
 G A P P P S G G K Q Q G R Q C D E T P P 2143  
 CCACCTCAGTCA6T6A6T6C6G6C6C6C6C6C6C6C6C6C6C6C6C6C6C6C6C6 6540  
 T S V S E S A P S I Q T D S D V P P A G E 2163  
 A6A6T6C6CCTCCATCAG6C6G6C6C6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6 6600  
 E C P S I T A D A A M I D B E D E S E T I 2183  
 TCCCA6CAG6A6 6660  
 P T D K T V T Y K H M D P P P A P M Q D 2203  
 ACC6A6G6CCTTCTCTC6C6C6C6C6G6A6T6T6C6A6T6G6G6A6C6G6CCTT6G6 6720  
 R S P S P R H P D V B H V D P D A L A V 2223  
 TT6ATCAG6A6CTAG6CA6G6CT6T6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6 6780  
 D Q N L G K A V K K D L K E K T K T K K 2243  
 A6C6A6G6C6A6 6840  
 P G T K T K S S S P V K K G D G K S K P 2263  
 CTTT6AC6G6CTTCC6C6A6A6C6A6G6CTT6A6G6A6T6C6G6A6A6G6T6C6A6 6900  
 L A A S P K P G A L K E B S D K V S R V 2283  
 T6GCTTCC6A6 6960  
 A B P K K K E S V E K A T K T T T T P E 2303  
 A6T6A6A6C6A6C6G6G6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6 7020  
 V K A T R G E E K D K E T K N A A N A S 2323  
 CTC6ATCAAATCAG6A6A6C6C6A6C6A6G6A6G6A6A6A6A6A6A6A6A6A6A6 7080  
 A S S A K T A T T G P G T T K T A K S 2343  
 C6TCC6C6G6CTC6A6G6CCTCCT6T6T6T6G6A6T6G6A6C6T6T6C6A6A6C6A 7140  
 S T V P P G L P V Y L D L C Y I P N H S 2363  
 GCAAC6A6A6A6T6G6A6T6G6T6T6C6A6G6A6G6A6G6A6G6A6G6A6G6A6 7200  
 N S K N V D V E F F K R V R S S Y Y V V 2383  
 T6A6T6G6A6A6A6C6C6T6C6C6G6A6G6A6C6C6G6G6C6T6G6T6G6A6T6T6 7260  
 S G M D P A A E E P S R A V L D A L L E 2403  
 A6G6A6G6CTCAGT6G6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6 7320  
 G K A Q M G S N H Q V T L I P T H D S E 2423  
 A6T6AT6A6G6A6T6G6T6C6A6G6A6A6A6A6A6A6A6A6A6A6A6A6A6A6A6 7380  
 V M R E W Y Q E T H E K Q Q D L N I M V 2443  
 TCTT6A6 7440  
 L A S B S T V V H Q D E S C F T P A C K I E 2463  
 AACT6TAG6A6C6C6G6C6C6C6C6C6C6C6C6C6C6C6C6C6C6C6C6C6C6 7500  
 L 2464  
 GAATT6A6 7560  
 TACCTG6CAG6T6CTAT6T6T6T6T6T6T6T6T6T6T6T6T6T6T6T6T6T6T6 7620

**Figure 4.** Complete sequence of MAPIB. The nucleotide sequence of our set of overlapping cDNA clones is shown together with the encoded amino acid sequence, in the one letter code. (The sequence of 1.4 kb of 3' untranslated region is not shown, but the entire sequence has been submitted to GenBank.) A repeated basic motif encoded toward the 5' end is underlined; a second set of repeats near the 3' end is indicated with dashed underlines. Possible phosphorylation sites for casein kinase II are indicated in two ways: predicted (●) and strongly predicted (▲) (see text).

designed for in vitro transcription. Capped mRNA synthesized from this construct was translated in a cell-free system containing [<sup>35</sup>S]methionine, and the reaction products were analyzed by SDS-PAGE. The data from this experiment (Fig. 2) show an abundant translation product whose molecular size is consistent with initiation at the first AUG within the long open reading frame. Therefore, the 5' proximal AUG functions as an efficient translational initiator in vitro. The same AUG is also capable of efficient translational initiation in vivo (see below). Secondly, to examine the possible existence of MAPIB-encoding aminoterminal sequences upstream from the region covered by our overlapping cDNAs, we isolated a corresponding cloned genomic fragment by screening a mouse genomic library with a MAPIB cDNA probe. An antisense synthetic oligonucleotide corresponding to sequences close to the 5' end of overlapped MAPIB-encoding cDNAs was then used as a sequencing primer to determine upstream genomic sequences. This experiment showed that while the reading frame specified by the MAPIB amino acid sequence remains open, there are no potential upstream AUG codons; however, there is a canonical TATA sequence that could potentially form part of the promoter for the MAPIB gene (data not shown). Thirdly, we performed a primer extension experiment using an antisense oligonucleotide corresponding to nucleotides 17-34 of our MAPIB

cDNA sequence (see below) and RNA isolated (from total mouse brain RNA) by hybrid selection with clone 5X (Fig. 1). The result of this experiment is shown in Fig. 3: an extended product 69 nucleotides in length appears in the primed reaction containing selected RNA that is not present in a control reaction containing an equivalent mass of tRNA. Because the 5' end of the primer used in this experiment is located at nucleotide 34 in the sequence determined from overlapping cloned cDNAs, this experiment locates the MAPIB cap site only 35 nucleotides 5' to the beginning of the sequence shown in Fig. 4. Given the absence of upstream AUG codons in upstream genomic DNA and the very short 5' sequence not represented in our cloned cDNAs, the likelihood of translational initiation occurring at a site 5' to the first AUG present in the long open reading frame of our cloned cDNAs seems very remote.

To confirm the position of the translational termination codon, which our sequence analysis placed 1.5 kb from the 3' end of our cDNA clones, a restriction fragment spanning the region containing the 3' end of the long open reading frame was cloned into a vector designed for the transcription of RNA sequences. An initiator methionine codon was supplied in frame with the cloned insert via the AUG contained in the polylinker of the pGEM vector. This construct was linearized at restriction sites that we predicted would be either

**A**

```

Y A Y Q K P E N A A G S P D E
Y D Y E S Q E K T I R T H D V
Y Y Y E K T E R T I K S P C D
Y S Y E T I E K T T R T P E D
Y T C E I T E K T T R T P E E
Y S Y E I S E K T T R T P E V
Y T Y E K T E R S R R R L L D D
N G Y D D T E D G G H T L G D
Y S Y E T T E K I T S F P E S
Y S Y E T S T K T T R S P D T
Y C Y E T M E K I T K T P Q A
Y S Y E T S D R A T K S P S E
      C Y E K

```

**B**

```

Y S Y E T S E R T T R S P E E
      T E K T K T P D D

```

**Figure 5.** 12 imperfect repeats in the sequence of MAP1B. The repeats indicated by dashed underlines in Fig. 4 are aligned (A), and a consensus sequence is shown (B).

before or after the putative stop codon. These linearized DNAs were then used as templates for *in vitro* mRNA synthesis, and the mRNA was translated *in vitro* in a rabbit reticulocyte lysate containing [<sup>35</sup>S]methionine. As shown in Fig. 2, only the template cut 5' to the predicted stop codon (at Kpn I) yields a truncated product, thus confirming the existence of the stop codon at the position indicated in our MAP1B cDNA clones.

Because of the large size of MAP1B and the total absence of any protein sequence data, we checked the putative open reading frame in two ways, in addition to verifying the start and stop codons as described above. First, we verified that the reading frame in each of our original λgt11 expression clones matched the predicted long open reading frame. Second, we made seven *E. coli*-trpE fusion proteins with various fragments derived from the set of cDNAs in the reading frames predicted by our sequence data. In each case, the size of the fusion protein was in good agreement with that calculated from our sequence data (data not shown). Two of these fusion proteins were used to raise antisera: 3d2 (reported previously [1]) and YXY (see Fig. 1). These two fusion proteins were chosen because of their unusual amino acid sequences (see below). The sera generated using these fusion proteins are especially useful because they uniquely recognize two widely spaced and well defined regions of MAP1B. In particular, they are in all probability not against phosphorylated epitopes; antibodies to phosphorylated epitopes can produce confusing results, in part because the phosphorylation of MAP1B is dependent on developmental stage (43), and also because some of these antibodies cross react with phosphorylated epitopes on the neurofilament and other proteins (2, 31).

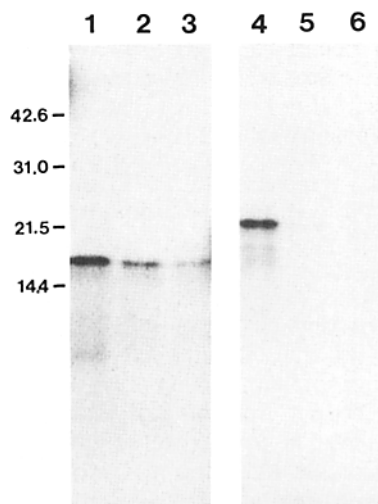
### Deduced Sequence of MAP1B

The sequence of our overlapping MAP1B cDNAs together with the deduced amino acid sequence of MAP1B is shown in Fig. 4. The encoded protein is 2,463 amino acids long, with a molecular mass of 255,534 D. While this is smaller than its apparent size on SDS polyacrylamide gels, MAP1B (like tau, MAP2, and NF-M [18, 24, 25]) is known to run anomalously large in such gels (12). The predicted size of MAP1B is, however, ~25% larger than the calculated size of MAP2, which is consistent with their relative mobilities.

Two regions of repetitive amino acid sequence, each ~200 residues long, are evident in Fig. 4. One is a highly basic region containing many sequences of the form KKEX, where X is almost always an acidic or hydrophobic residue. These repeated motifs do not fall into any strict pattern, although they are clustered in the central part of a region containing

many basic amino acids. Indeed, the region from amino acids 589–787 has a calculated net charge of +30, and both the methods of Chou-Fasman (11) and Garnier et al. (15) predict that it folds into short, accessible hydrophilic α-helices, although these predictions are notoriously unreliable (50). The second repetitive region of MAP1B spans amino acids 1865–2070, and consists of 12 imperfect 15-amino acid-long repeats (Fig. 5); each repeat is separated by two nonconserved amino acids. This region is also quite hydrophilic, and is predicted to fall on a series of turns in the MAP1B molecule.

Recently, it has been discovered that MAP2 and tau have a highly similar carboxyterminal domain that contains three (or four [19, 20]) 18-amino acid-long imperfect repeats (24, 30). This domain has been shown to constitute the microtubule binding site of MAP2 and tau both *in vitro* (24, 30) and *in vivo* (S. A. Lewis, I. E. Ivanov, G. H. Lee, and N. J. Cowan, manuscript submitted for publication). However, the repeats in MAP2 and tau have no similarity to either the basic repeated motif of the 15-amino acid-long repeats in MAP1B, nor are there any regions of significant sequence homology between MAP1B and either tau or MAP2. Indeed, when the deduced amino acid sequence of MAP1B is compared with the PIR protein data base, no significant homol-



**Figure 6.** Translated protein corresponding to the first repeated motif in MAP1B cycles with microtubules *in vitro*. Capped RNAs transcribed from two regions (Fig. 1, arrows) were translated in a cell free system containing [<sup>3</sup>H]lysine or [<sup>35</sup>S]methionine (see Materials and Methods). Translation products were cycled with microtubules, and at each stage, aliquots containing the same amount of protein (40 μg) as judged by absorption at A<sub>280</sub>/A<sub>260</sub> were analyzed on a 14% SDS-polyacrylamide gel. Lane 1, <sup>3</sup>H-labeled protein derived from 5' subclone; lanes 2 and 3, 40 μg of microtubule protein after 1 and 2 additional cycles of assembly and disassembly in the presence of added <sup>3</sup>H-labeled translation product. Lane 4, <sup>35</sup>S-labeled protein derived from the 3' subclone; lanes 5 and 6, 40 μg of microtubule protein after 1 and 2 additional cycles of assembly and disassembly in the presence of added <sup>35</sup>S-labeled translation product. The times of exposure have been adjusted to compensate for the difference in specific activity between [<sup>3</sup>H]lysine and [<sup>35</sup>S]methionine in the cell-free system, the relative abundance of lysine and methionine residues in the respective peptides, and the relative fluorographic detection efficiency of the two isotopes.

ogies to other proteins are uncovered. However, the basic domain of MAP1B encompassing the KKEE and KKE<sub>v</sub> repeated motifs shows some (probably fortuitous) sequence similarity to basic regions in other proteins, primarily to histones and other nuclear proteins but also to the highly charged tails of the neurofilament triplet proteins. This may account for the cross-reactivity of several anti-MAP1B antibodies with nuclear (12) or neurofilament (31) antigens.

The phosphorylation of MAP1B increases during neurite outgrowth, both in PC12 cells induced to differentiate with nerve growth factor (1, 5, 14) and in neuroblastoma cells that generate neurites when serum is removed from the growth medium (13). MAP1B is also found in phosphorylated form in developing axons *in vivo* (43). These facts, together with the elevated levels of MAP1B protein found in developing axons (or neurites) *in vivo* and *in vitro* (14, 16, 43), and the increased affinity of phosphorylated MAP1B for microtubules (13), suggest that phosphorylated MAP1B plays a role in the cytoskeletal changes that accompany neurite extension. A casein kinase II type activity is implicated in MAP1B and  $\beta$ -tubulin phosphorylation both in brain and in cultured differentiating neuroblastoma cells (13, 46). Casein kinase II phosphorylates serine and threonine residues in sequences of the form  $\frac{X}{Y} \times \frac{R}{S}$  where  $X \neq K, R$ ; however, only those sequences that fall on  $\beta$  turns in a protein appear to be phosphorylated (39). The primary recognition sequence for casein kinase II occurs 81 times in the MAP1B sequence; in 30 cases these sequences are predicted to form  $\beta$  turns by the Chou-Fasman and Garnier algorithms, and, in seven of these cases, turns are strongly predicted. These 30 potential phosphorylation sites are indicated in Fig. 4; nine are clustered in the region of MAP1B containing the 12 imperfect repeats.

### The Microtubule Binding Domain of MAP1B

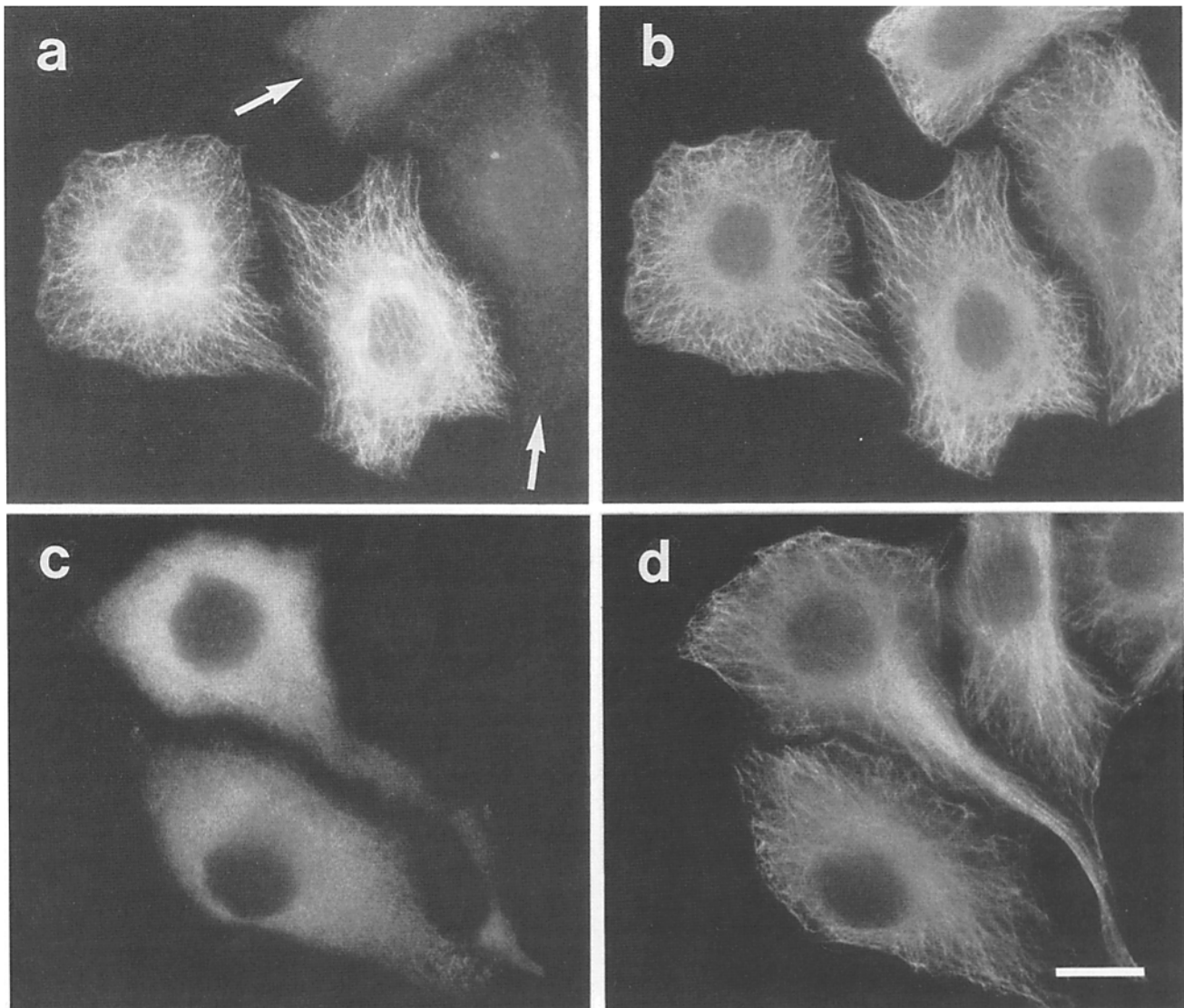
There exists within the MAP1B amino acid sequence two regions containing repeated motifs (discussed above). The fact that the microtubule binding regions of both MAP2 and tau contain repeated sequences (24, 30), and that microtubules are themselves composed of repeating subunits suggested that one of these regions might constitute the microtubule binding domain. To test the ability of these domains to bind to microtubules, restriction fragments encoding regions spanning each set of repeats (Fig. 1, *arrows*) were subcloned into RNA transcription vectors; a translational start signal in the correct reading frame was in each case provided by the vectors. The resulting constructs were used as templates for the transcription of capped mRNA, and the transcription products used to direct protein synthesis *in vitro* in a rabbit reticulocyte lysate containing either [<sup>3</sup>H]lysine (in the case of the first repeat region, which contains no methionine or cysteine residues) or [<sup>35</sup>S]methionine (in the case of the second repeat region). The translation products were added to two-cycle-purified, depolymerized bovine brain microtubules, which were then taken through two further cycles of polymerization and depolymerization. After each cycle, aliquots containing the same amount (40  $\mu$ gm) of protein were removed and analyzed by SDS-PAGE, to see whether either labeled polypeptide could coassemble with the unlabeled carrier microtubules. The result of this experiment is shown in Fig. 6. A labeled polypeptide derived from the first repeat region was observed to cycle with brain microtubules, though

there is a significant loss (of  $\sim 60\%$ ) of labeled polypeptide with each successive cycle; in contrast, a labeled polypeptide derived from the second repeat region completely failed to cycle. The relative inefficiency with which the labeled polypeptides from the first repeat region cycles with microtubules *in vitro* could result from competition with endogenous unlabeled MAP1B present in the microtubule preparations used in these experiments, or from the inherently poor cycling efficiency of MAP1B with microtubules *in vitro* (6), or a combination of both these factors. Nonetheless, these data suggest that the region containing the basic repeated sequence motifs is responsible for the cycling of MAP1B with microtubules during their purification.

To determine whether this domain is also responsible for MAP1B-microtubule interactions *in vivo*, a construct designed to express the entirety of MAP1B was made and transfected into cultured cells. To facilitate the assembly of a full-length MAP1B coding region, we isolated a mouse genomic clone from a cosmid library containing a 6-kb exon (Fig. 1, *clone C*). After joining this exon to cDNAs derived from the 5' and 3' ends of the MAP1B transcript, the full coding sequence was inserted into a vector containing promoter and terminator sequences derived from SV40. The final construct was transfected into HeLa and 3T3 cells and the disposition of MAP1B in transfected cells was assayed by double label immunofluorescence, using either of the anti-MAP1B antisera 3d2 and YXY, described above, and a mouse anti- $\alpha$ -tubulin monoclonal antibody. The observed patterns of MAP1B staining were dependent on the fixation conditions: no MAP1B was observed bound to detergent extracted cytoskeletal preparations, but in cells fixed directly in methanol (in which case proteins are fixed and extracted simultaneously) MAP1B-decorated microtubules were observed (Fig. 7, *a* and *b*). Notably, in cells fixed directly with paraformaldehyde, MAP1B filled the cytoplasm, obscuring the microtubule-bound MAP1B (data not shown). This implies that much or most MAP1B in transfected cells is in soluble form. Thus, in contrast to MAP2 (S. A. Lewis, I. E. Ivanov, G.-H. Lee, and N. J. Cowan, manuscript submitted for publication), MAP1B binds to microtubules weakly, and the equilibrium favors the unbound state since none remains bound to microtubules when the unbound portion is removed by detergent extraction. This is consistent with the behavior of MAP1B in brain (32), and the poor cycling efficiency of MAP1B with microtubules *in vitro* (7).

This assay for the binding of MAP1B to microtubules *in vivo* allowed us to define the microtubule binding domain by making constructs expressing deleted forms of MAP1B, and transfecting them into cultured cells. The results of these experiments are summarized in Fig. 1; representative immunofluorescence data are shown in Fig. 7. Deletion of the 5' half of the MAP1B coding region (construct 1B-H) or of those sequences encoding amino acids 524-848 (construct 1B-K) results in the expression of truncated forms of MAP1B that have completely lost the capacity to bind to microtubules (e.g., Fig. 7, *c* and *d*). Thus, the region of MAP1B containing the repeated KKEE and KKE<sub>v</sub> motifs is alone responsible for microtubule binding both *in vitro* and *in vivo*.

Two further constructs, X and 1B-X (Fig. 1), show that the microtubule binding domain is divisible. The protein expressed from X includes only amino acids 646-732, whereas that expressed from 1B-X contains all of MAP1B with the ex-



**Figure 7.** Characterization of the MAPIB microtubule binding domain *in vivo*. Constructs MAPIB and 1B-K (Fig. 1) were transfected into HeLa cells. After 72 h, cells were fixed with methanol, and analyzed by indirect double label immunofluorescence using the YXY antibody (*a* and *c*) directed against a segment of MAPIB (see Fig. 1), and an anti- $\alpha$ -tubulin monoclonal antibody (*b* and *d*). *a* and *b*, Cells transfected with MAPIB; *c* and *d*, cells transfected with the 1B-K construct (Fig. 1); untransfected cells in *a* are indicated by arrows. Bar, 10  $\mu$ m.

ception of this region. Surprisingly, the protein expressed from each of these complementary constructs binds to microtubules with approximately the same efficiency (within the limits of our assay) as intact MAPIB. This implies that at least two subregions from the KKEE and KKE<sub>v</sub>-containing region bind independently to microtubules. It therefore seems likely that MAPIB binds to at least two tubulin subunits in the polymer, and that this bridging of subunits is involved in nucleating microtubule polymerization and in stabilizing microtubules.

### Conclusions

We have cloned and sequenced cDNAs spanning the entire coding region of the MAPIB gene, and shown that a basic region of the MAPIB protein, containing many repeats of the form KKEE or KKE<sub>v</sub>, is the MAPIB-microtubule binding domain. This was accomplished experimentally *in vitro* by

cycling translated fragments with microtubules, and *in vivo*, by transfecting constructs expressing full-length and deleted forms of MAPIB into cultured cells. The MAPIB binding domain is completely unrelated to the highly similar microtubule-binding domains of MAP2 and tau, although like those domains, it has a net positive charge and contains repeated elements. Thus, in common with MAP2 and tau (44, 45) MAPIB probably interacts ionically with the negatively charged carboxyterminal portion of  $\alpha$ - and/or  $\beta$ -tubulin, though it has not been shown whether these molecules compete for the identical site on microtubules. The binding of MAPIB to microtubules is weak compared with MAP2, and unlike MAP2 (S. A. Lewis, I. E. Ivanov, G.-H. Lee, and N. J. Cowan, manuscript submitted for publication) expression of high levels of MAPIB in cultured cells does not result in any obvious cytoskeletal changes (Fig. 7, *a* and *b*). However, MAPIB has two light chains that also associate with



MAP1A (22, 23, 43); these and/or other coexpressed proteins may alter the properties of MAP1B.

The data presented here for MAP1B, together with that reported previously for MAP2 and tau (19, 20, 24, 30) and for kinesin (53), define three structurally unrelated kinds of microtubule binding domains. These differences must in some way reflect the different functions served by the interaction of each MAP with microtubules. For example, kinesin, which acts as a motor for anterograde transport along microtubules (8, 48), would be expected to have completely different binding properties from MAP2, which crosslinks and stabilizes microtubules in dendrites (S. A. Lewis, I. E. Ivanov, G.-H. Lee, and N. J. Cowan, manuscript submitted for publication). Therefore, it is not surprising that their binding domains are completely dissimilar. It seems probable that each of these three binding domains is representative of a separate family of functionally related microtubule-associated proteins.

We thank Dashou Wang for help with DNA sequencing.

This work was supported by grants from the National Institutes of Health and the Alzheimer's Disease Association.

Received for publication 11 August 1989 and in revised form 18 September 1989.

## References

- Aletta, J. M., S. A. Lewis, N. J. Cowan, and L. A. Greene. 1988. Nerve growth factor regulates both the phosphorylation and steady-state levels of microtubule-associated protein 1.2 (MAP 1.2). *J. Cell Biol.* 106: 1573-1581.
- Asai, D. J., W. C. Thompson, L. Wilson, C. F. Dresden, H. Schulman, and D. L. Purich. 1985. Microtubule-associated proteins (MAPs): a monoclonal antibody to MAP-1 decorates microtubules *in vitro* but stains stress fibers and not microtubules *in vivo*. *Proc. Natl. Acad. Sci. USA.* 82:1434-1438.
- Balcarek, J. M., and N. J. Cowan. 1985. Structure of the mouse glial fibrillary acidic protein gene: implications for the evolution of the intermediate filament multigene family. *Nucleic Acids Res.* 13:5527-5543.
- Benton, W. D., and R. W. Davis. 1977. Screening  $\lambda$ gt recombinant clones by hybridization of single plaques *in situ*. *Science (Wash. D.C.)* 196: 180-181.
- Black, M. M., J. M. Aletta, and L. A. Greene. 1986. Regulation of microtubule composition and stability during nerve growth factor-promoted neurite outgrowth. *J. Cell Biol.* 103:545-57.
- Bloom, G. S., F. C. Luca, and R. B. Vallee. 1985. Identification of high molecular weight microtubule associated proteins in anterior pituitary tissue and cells using taxol-dependent purification combined with microtubule-associated protein specific antibodies. *Biochemistry.* 24:4185-4191.
- Bloom, G. S., F. C. Luca, and R. B. Vallee. 1985. Microtubule associated protein 1B: identification of a major component of the neuronal cytoskeleton. *Proc. Natl. Acad. Sci. USA.* 82:5404-5408.
- Brady, S. T. 1985. A novel brain ATPase with properties expected for the fast axonal transport motor. *Nature (Lond.)* 317:73-75.
- Calvert, R., and B. H. Anderson. 1985. A microtubule-associated protein (MAP1) which is expressed at elevated levels during development of the rat cerebellum. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:1171-1176.
- Chen, C., and H. Okayama. 1987. High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell Biol.* 7:2745-2752.
- Chou, P. Y., and G. T. Fasman. 1974. Prediction of protein conformation. *Biochemistry.* 13:222-245.
- Diaz-Nido, J., and J. Avila. 1989. Characterization of proteins immunologically related to brain microtubule-associated protein MAP1B in non-neuronal cell. *J. Cell Sci.* 92:607-620.
- Diaz-Nido J., L. Serrano, E. Mendez, and J. Avila. 1988. A casein kinase II-related activity is involved in phosphorylation of microtubule-associated protein MAP1B during neuroblastoma cell differentiation. *J. Cell Biol.* 106:2057-2065.
- Drubin, D. G., S. Feinstein, E. Shooter, and M. Kirschner. 1985. Nerve growth factor-induced outgrowth in PC12 cells involves the coordinate induction of microtubule assembly and assembly promoting factors. *J. Cell Biol.* 101:1799-1807.
- Garnier, J., D. J. Osguthorpe, and B. Robson. 1978. Analysis of the accuracy and implication of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.* 120:97-120.
- Greene, L. A., R. K. H. Liem, and M. L. Shelanski. 1983. Regulation of a high molecular weight microtubule-associated protein in PC12 cells by nerve growth factor. *J. Cell Biol.* 96:76-83.
- Hall, J. L., L. Dudley, P. R. Dobner, S. A. Lewis, and N. J. Cowan. 1983. Identification of two human  $\beta$ -tubulin isotypes. *Mol. Cell Biol.* 3:854-862.
- Hernandez, M. A., J. Avila, and J. M. Andreu. 1986. Physicochemical characterization of microtubule-associated protein MAP-2. *Eur. J. Biochem.* 154:41-48.
- Himmler, A. 1989. Structure of the bovine tau gene: alternatively spliced transcripts generate a protein family. *Mol. Cell Biol.* 9:1389-1396.
- Himmler, A., D. Drechsel, M. Kirschner, and D. W. Martin. 1988. Tau consists of a set of proteins with repeated C-terminal microtubule-binding domains and variable N-terminal domains. *Mol. Cell Biol.* 9:1381-1388.
- Huber, G., D. Alaimo-Beuret, and A. Matus. 1985. MAP3: characterization of a novel microtubule-associated protein. *J. Cell Biol.* 100:496-507.
- Kuznetsov, S. A., and V. I. Gelfand. 1987. 18kDa microtubule-associated protein: identification as a new light chain (LC-3) of microtubule-associated protein 1 (MAP-1) *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 212: 145-148.
- Kuznetsov, S. A., V. I. Rodionov, E. S. Nadezhkina, D. B. Murphy, and V. I. Gelfand. 1986. Identification of a 34-kD polypeptide as a light chain of microtubule-associated protein-1 (MAP-1) and its association with a MAP-1 peptide that binds to microtubules. *J. Cell Biol.* 102:1060-1066.
- Lee, G., N. Cowan, and M. Kirschner. 1988. The primary structure and heterogeneity of tau protein from mouse brain. *Science (Wash. DC)* 239:285-288.
- Levy, E., R. K. H. Liem, P. D'Eustachio, and N. J. Cowan. 1987. Structure and evolutionary origin of the gene encoding mouse NF-M, the middle-molecular-mass neurofilament protein. *Eur. J. Biochem.* 166: 71-77.
- Lewis, S. A., and N. J. Cowan. 1985. Genetics, evolution, and expression of the 68,000-mol-wt neurofilament protein: isolation of a cloned cDNA probe. *J. Cell Biol.* 100:843-850.
- Lewis, S. A., M. E. Gilmartin, J. L. Hall, and N. J. Cowan. 1985. Three expressed sequences within the human  $\beta$ -tubulin multigene family each define a distinct isotype. *J. Mol. Biol.* 182:11-20.
- Lewis, S. A., P. Sherline, and N. J. Cowan. 1986. A cloned cDNA encoding MAP 1 detects a single copy gene in mouse and brain-abundant RNA whose level decreases during development. *J. Cell Biol.* 102:2106-2114.
- Lewis, S. A., A. Villasante, P. Sherline, and N. J. Cowan. 1986. Brain specific expression of MAP2 detected using a cloned cDNA probe. *J. Cell Biol.* 102:2098-2105.
- Lewis, S. A., D. Wang, and N. J. Cowan. 1988. Microtubule-associated protein MAP2 shares a microtubule binding motif with tau protein. *Science (Wash. DC)* 242:936-939.
- Luca, F. C., G. S. Bloom, and R. B. Vallee. 1986. A monoclonal antibody that cross-reacts with phosphorylated epitopes on two microtubule-associated proteins and two neurofilament polypeptides. *Proc. Natl. Acad. Sci. USA.* 83:1006-1010.
- Matus, A. 1988. Microtubule-associated proteins: their potential role in determining neuronal morphology. *Annu. Rev. Neurosci.* 11:29-44.
- Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* 12:7035-7056.
- Mulligan, R. C., and P. Berg. 1981. Selection for animal cells that express the *Escherichia coli* gene coding for xanthine-guanine phosphoribosyltransferase. *Proc Natl. Acad. Sci. USA.* 78:2072-2076.
- Olmstead, J. B. 1986. Microtubule-associated proteins. *Annu. Rev. Cell Biol.* 2: 421-457.
- Osborn, M., K. Weber. 1982. Immunofluorescence and immunocytochemical procedures with affinity purified antibodies: tubulin containing structures. *Methods Cell Biol.* 244:98-129.
- Parysek, L. M., J. J. Woloszewicz, and J. Olmsted. 1984. MAP4: a novel microtubule-associated protein specific for a subset of tissue microtubules. *J. Cell Biol.* 99:2287-2296.
- Paschal, B. M., and R. B. Vallee. 1987. Retrograde transport by the microtubule associated protein MAP1C. *Nature (Lond.)* 330:181-183.
- Pinna, L. A., V. A. Deano, and F. Meggri. 1979. Structural features determining the site specificity of a rat liver cAMP independent protein kinase. *Biochem. Biophys. Res. Commun.* 87:114-120.
- Riederer, B., R. Cohen and A. Matus. 1986. MAP5: a novel brain microtubule-associated protein under strong developmental regulation. *J. Neurocytol.* 15:763-775.
- Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labelling DNA to high specific activity *in vitro* by nick-translation with DNA polymerase I. *J. Mol. Biol.* 113:237-251.
- Sanger, F. A., A. R. Coulson, B. G. Barrell, A. J. H. Smith, and B. Roe. 1980. Cloning single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* 143:161-178.
- Schoenfeld, T. A., L. McKerracher, R. Ober, and R. B. Vallee. 1989. MAP1A and MAP1B are structurally related microtubule associated proteins with distinct developmental patterns in the CNS. *J. Neurosci.*

- 9:1712-1730.
44. Serrano, L., J. Avila, and R. B. Maccioni. 1984. Controlled proteolysis of tubulin by subtilisin: localization of the site for MAP-2 interaction. *Biochemistry*. 23:4675-4681.
  45. Serrano, L., E. Montejo, M. A. Hernandez, and J. Avila. 1985. Localization of the tubulin binding site for tau protein. *Eur. J. Biochem.* 153:595-600.
  46. Serrano, L., J. Diaz-Nido, F. Wandosell, and J. Avila. 1987. Tubulin phosphorylation by casein kinase II is similar to that found in vivo. *J. Cell Biol.* 105:1731-1739.
  47. Solomon, F., M. Magendantz, and A. Salzman. 1979. Identification with cellular microtubules of one of the co-assembling microtubule-associated proteins. *Cell*. 18:431-438.
  48. Vale, R. P., T. S. Reese, and M. P. Sheetz. 1985. Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility. *Cell*. 42:39-50.
  49. Vallee, R. B., B. M. Paschal, H. S. Shpetner, and J. S. Wall. 1988. Microtubule associated protein MAP1C from brain is a two-headed cytosolic dynein. *Nature (Lond.)*. 332:561-563.
  50. Von Heijne, G. 1987. Sequence analysis in molecular biology: treasure trove or trivial pursuit. Academic Press Inc., NY.
  51. Weingarten, M. D., M. H. Lockwood, S. Y. Hwo, and M. W. Kirschner. 1975. A protein factor essential for microtubule assembly. *Proc Natl. Acad. Sci. USA*. 72:1858-1862.
  52. Wiche, G. 1985. High-molecular-weight microtubule-associated proteins: an ubiquitous family of cytoskeletal connecting lines. *Trends Biochem. Sci.* 10:67-70.
  53. Yang, J. T., R. A. Layman, and L. S. B. Goldstein. 1989. A three domain structure of kinesin heavy chain revealed by DNA sequence and microtubule binding analyses. *Cell*. 56:879-889.