ptl-1, a *Caenorhabditis elegans* Gene Whose Products Are Homologous to the τ Microtubule-Associated Proteins[†,‡]

Joan B. McDermott,[§] Stephanie Aamodt,[||] and Eric Aamodt*[,§]

Department of Biochemistry and Molecular Biology, Louisiana State University Medical Center, 1501 Kings Highway, Shreveport, Louisiana 71130-3932, and Department of Biology, Louisiana State University, 1 University Place, Shreveport, Louisiana 71115

Received November 7, 1995; Revised Manuscript Received May 17, 1996[⊗]

ABSTRACT: The τ microtubule-associated proteins are axonal proteins that have been implicated in axonal outgrowth, microtubule spacing, and microtubule bundling. Moreover, τ is the major structural component of the paired helical filaments present in the brains of Alzheimer's disease patients. The *Caenorhabditis elegans* Genome Sequencing Consortium identified a genomic sequence with homology to the repeat region of τ. PCR, Northern analyses, and cDNA sequencing were used here to identify transcripts containing the τ homology region. The gene that encodes these transcripts was named *ptl-1* for protein with τ-like repeats. The *ptl-1* transcript, like mammalian τ transcripts, is alternatively spliced to produce messages that encode proteins with variable numbers of repeats. The predicted *ptl-1* products have strong sequence homology to τ over the repeat region and are similar to τ in several other important respects including size, amino acid content, charge distribution, predicted secondary structure, hydrophobicity, and flexibility. Both proteins contain several potential glycosylation sites and numerous phosphorylation sites. Bacterially expressed PTL-1 bound to microtubules *in vitro*. These results show that τ-like proteins evolved early and suggests that they may be present in many different phyla. *C. elegans* is a powerful system amenable to genetic, molecular, and cellular analysis in which to study the functions of this important class of proteins.

Microtubule-associated proteins (MAPs)[1] are a group of proteins that bind to and co-purify with microtubules. A subfamily of MAPs include τ, MAP2, and MAP4. All of these proteins contain three or four imperfect repeats of 31 or 32 residues within their microtubule-binding domain (Chapin & Bulinski, 1991; Gustke et al., 1994; Lewis et al., 1988). The repeats are named R1, R2, R3, and R4. A fifth region with weak homology, R′, follows R4 (Gustke et al., 1994). All of the MAPs in the τ-MAP2-MAP4 family bind to microtubules, stimulate microtubule nucleation, stimulate microtubule assembly, stabilize microtubules, and form arm-like projections from microtubules [for reviews see Chapin and Bulinski (1992), Lee (1993), Wiche et al. (1991)].

The mammalian τ MAPs are a developmentally regulated group of phosphoproteins that are located in axons. All of the τs are encoded by a single gene that produces alternatively spliced mRNAs [Goedert et al., 1989b; Himmler, 1989; reviewed in Goedert et al. (1991)]. The N-terminal third and the final fifth of τ are highly acidic while the region in between is highly basic. The microtubule-binding domain lies within the basic region of τ. Early in development, τ variants with three repeats are made; later, variants with four repeats are preferentially produced (Kosik et al., 1988; Goedert et al., 1989a; Himmler, 1989). Interest in τ was greatly stimulated by the discovery that the paired-helical filaments of Alzheimer's disease are composed of abnormally phosphorylated τ (Kosik et al., 1988; Bramblett et al., 1993; Grunke-Iqbal et al., 1986; Lee & Trojanowski, 1992; Wischik et al., 1988; Wood et al., 1986).

Studies of τ in cell culture have suggested that τ has a role in controlling microtubule assembly and in organizing the microtubule cytoskeleton in axons. Expression of τ in Sf9 insect cells causes long axon-like processes to form (Knops et al., 1991), and expression in CHO cells results in microtubule bundling but does not increase the total amount of polymerized tubulin (Barlow et al., 1994). τ injected into non-neuronal cells in culture becomes localized to microtubules and stimulates net microtubule assembly (Drubin and Kirschner, 1986). Primary cerebellar neurons cultured in the presence of τ antisense oligonucleotides fail to elaborate axons [Caceres and Kosik, 1990; reviewed in Kosik and Caceres (1991)]. Expression of τ antisense RNA in PC12 cells causes inhibition of process formation, whereas overexpression of τ causes net microtubule assembly and more rapid neurite extension (Esmaeli-Azad et al., 1994). Finally, mice without a τ gene show no external phenotype, and axonal polarity, elongation, and stability appear normal. The microtubules of their small-caliber axons are slightly disorganized, however, and their brains have increased levels of MAP1A (Harada et al., 1994). The mouse τ knockout experiments show that τ is not necessary for regulating

[†] This work was supported by grants to E.A. from the LSU Medical Center—Shreveport Stiles Fund, The Biomedical Research Foundation of Northwest Louisiana and the LSUMC Center for Excellence in Cancer Research and by grants to S.A. from LSU—Shreveport Faculty Research Funds.

* To whom correspondence should be addressed.

[‡] Each of the authors contributed equally to this study.

[§] Louisiana State University Medical Center.

[||] Louisiana State University.

[⊗] Abstract published in *Advance ACS Abstracts,* July 1, 1996.

[1] Abbreviations: MAP, microtubule-associated protein; RT-PCR, reverse transcriptase polymerase chain reaction; RACE, rapid amplification of cDNA ends; dNTP, deoxynucleotide triphosphate; SDS, sodium dodecyl sulfate; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol; NCBI, National Center for Biotechnology Information; IPTG, isopropyl β-D-thiogalactopyranoside; PAGE, polyacrylamide gel electrophoresis; ACEDB, A *C. elegans* data base.

S0006-2960(95)02646-8 CCC: $12.00

microtubule polymerization, stability, or polarity *in vivo*, perhaps because other proteins can substitute for τ.

To date, only mammalian τ genes have been cloned and sequenced. τ-like proteins have been identified in mammalian axons, chicken erythrocytes (Lichtenberg-Kraag & Mandelkow, 1990; Murphy & Wallis, 1985), and at low levels in some other mammalian tissues (Kenner et al., 1994). Proteins antigenically similar to τ have been described in non-mammalian species, including both plants and animals (Cambiazo et al., 1995; Vantard et al., 1991, 1994). Thus, knowledge of non-mammalian τ MAPs has so far been limited to protein analyses.

The *Caenorhabditis elegans* Genome Sequencing Consortium identified sequences in the *C. elegans* genome that are predicted to encode a protein with homology to the repeat region of the τ class of MAPs. Here we use Northern blots, PCR, and cDNA sequencing to show that this region of the *C. elegans* genome does indeed code for a protein that has much in common with mammalian τ. This gene was named *ptl-1* for protein with τ-like repeats. The *ptl-1* primary transcript, like mammalian τ mRNA, is alternatively spliced to produce at least two messages. One mRNA contains four repeats while the other contains five repeats. The *ptl-1* product has sequence homology to mammalian τ only over the repeat region, but both proteins have a very similar amino acid content and a remarkably similar charge structure. Mammalian τ contains many phosphorylation sites, and *C. elegans* PTL-1 contains numerous potential phosphorylation sites. Finally, a *ptl-1* fusion gene expressed in bacteria has microtubule-binding activity. On the basis of these similarities, *ptl-1* may encode the *C. elegans* ortholog of mammalian τ. *C. elegans* will provide a convenient, easily manipulated system in which to study the function of τ-like proteins *in vivo*.

MATERIALS AND METHODS

Growth and Maintenance of C. elegans. Basic methods for culturing *C. elegans* were according to Brenner (1974). Wild-type refers to strain N2 var. Bristol. Genetic nomenclature follows the guidelines of Hodgkin (1995). Nematode growth medium agar (NGM) was used to grow *C. elegans* on Petri plates (Sulston & Hodgkin, 1988). To enhance the mass of *C. elegans* on each plate, the peptone was increased 5-fold. These are referred to as 5× peptone plates. Approximately 0.1 g of *C. elegans* could be harvested from each 100 mm diameter 5× peptone plate.

To harvest larger quantities of *C. elegans*, 2−3 mL of chicken egg yolk was added to a 5× peptone plate and the plate was seeded with approximately 0.1 g of *C. elegans*. After 3−4 days of growth, the egg-yolk plates contained over a gram of *C. elegans*. To make the egg-yolk plates, an egg was washed thoroughly with 70% ethanol and the yolk was separated from the white by standard culinary methods. The intact egg yolk was placed in a sterile Petri plate, and the liquid center of the yolk was carefully removed with a 10 cc syringe tipped with a 16 gauge needle. The yolk was then ejected directly onto 5× peptone plates. To harvest and clean *C. elegans* grown on egg-yolk plates, the plates were washed with 0.1 M NaCl. The wash was cooled on ice and the animals were collected by centrifugation for 30 s at 1000*g* in a swinging bucket clinical centrifuge. *C. elegans* were purified from the yolk plate debris by sucrose flotation (Sulston & Hodgkin, 1988).

RNA Isolation. A slight modification of the procedure of Maes and Messens (1992) was used to isolate whole RNA from *C. elegans*. Frozen *C. elegans* (0.2−2.0 grams) were ground with 3.0 mL of acid phenol (pH 5.1) in a mortar and pestle cooled with liquid nitrogen. The powder was transferred to a 15 mL centrifuge tube and 3.0 mL of SDS-buffer (1% SDS, 50 mM sodium acetate, 10 mM EDTA, pH 5.1) was added. The mixture was heated to 60 °C for 10 min and then cooled on ice. Chloroform (0.6 mL) was added, and the phases were separated by centrifugation. The top phase was recovered and re-extracted with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) and then extracted with chloroform:isoamyl alcohol (24:1) twice. The RNA was precipitated with 0.1 volumes of 3 M sodium acetate and 2.5 volumes of ethanol. The RNA was further purified by LiCl precipitation followed by a second ethanol precipitation. The final pellet was dried under vacuum and resuspended in DEPC-treated distilled water. Poly-A-enriched RNA was isolated from whole RNA with the PolyATract system from Promega.

RT-PCR. cDNA for amplification was prepared from 20 μg of total RNA. The RNA was incubated with 45 pmol of oligo $d(T)_{18}$ at 65 °C for 3 min and then chilled. To this mixture was added 100 units M-MuLV reverse transcriptase, dNTPs (final concentration 1.0 mM each), buffer (final concentration 50.0 mM Tris-HCl, 8.0 mM $MgCl_2$, 3.0 mM KCl, 10.0 mM DTT, pH 8.3), 50 units of human placental RNase inhibitor and 5.0 μg of bovine serum albumin. The final volume was brought to 50 μL, and the mixture was incubated at 37 °C for 1 h. The reaction was terminated by the addition of 450 μL of TE (10 mM Tris, 1.0 mM EDTA, pH. 7.5). For each 100 μL PCR reaction, 5.0 μL of the cDNA mixture was used. The primers used were SL1 (GGTTTAATTACCCAAGTTTGAG), 1U (ATGAGCCCACCTAACCTGTCCA), 1L (CAACCCGGGAAAGAGTAAGTAT), 2L (CTACAACTTCTTCTTCCTTCTCG), 3U (ACTCCATCACCACCCCAACATA), 3L (CATAATTCTGTCGGCTTGTTGAT), 4U (AACTGCTACTCCCTCGTCTCAA), 5U (AACGGTCGGATGTTCAAAAATC), 5L (TGACACTTCCCACTTTGCTCTT), 6U (AGACTCTACAACGCCCAATCCA), 6L (TTATGCGCTGCGTTGTCCATTG), 7L (CAACCTTACTCTCGGCTTTCCA), and 8L (CTGATGTCCAGCGTAGAATGAT).

Southern and Northern Hybridization. *C. elegans* genomic DNA was isolated by a modification of a standard protocol (Sulston & Hodgkin, 1988). 2 g of *C. elegans* was grown on egg-yolk plates made with agarose rather than agar. They were cleaned by sucrose flotation and frozen in liquid nitrogen. The frozen worms were ground to a fine powder in a mortar cooled with liquid nitrogen, and the powder was transferred to two Oakridge centrifuge tubes each containing 20 mL of 100 mM EDTA (pH 8.0), 0.5% SDS, 50 μg of proteinase K/mL, and 1% β-mercaptoethanol. The mixture was incubated at 50 °C overnight, cooled on ice, and extracted with 20 mL of phenol three times by slow rolling at 4 °C for 15 min. The DNA was precipitated with ammonium acetate and ethanol, washed, and resuspended in TE buffer (Ausubel et al., 1993).

For Southern analysis, genomic DNA was digested with restriction enzymes, electrophoresed through 1.0% agarose in 0.5× TBE buffer (1× TBE is 89 mM Tris, 89 mM boric acid, 2 mM EDTA), blotted onto an uncharged nylon membrane, denatured, and cross-linked by UV irradiation

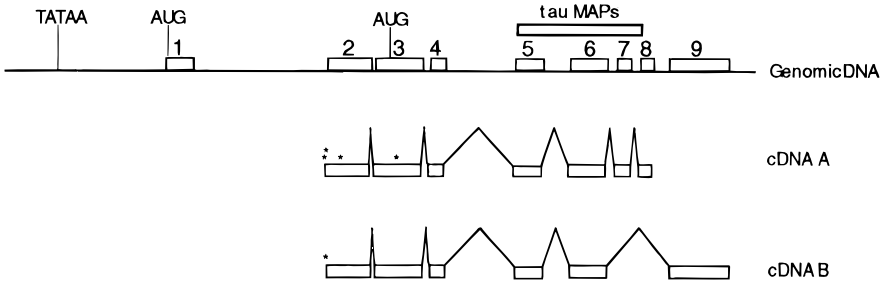

FIGURE 1: Intron−exon structure of the *ptl-1* genomic region, cDNA A, and cDNA B. Exons 1−8 were previously identified by the program Genefinder. Exon 9 was identified from the sequence of cDNA B. A potential TATA box and two potential translational start sites are indicated by TATAA and AUG, respectively. The position of sequences with homology to mammalian τ MAPs is indicated with a bar. Both exons 8 and 9 contain standard polyadenylation sites. The 5′ endpoints of the four *ptl-1a* and single *ptl-1b* cDNA clones are indicated by asterisks; each cDNA contained a poly-A tail. The GenBank accession numbers for the *ptl-1a* cDNA and *ptl-1b* cDNA sequences are U38982 and U38983, respectively.

(Ausubel et al., 1993). Hybridization was performed in a Hyb-aid rotating incubator at 65 °C in 6× SSC (1× SSC is 0.15 M NaCl, 15 mM sodium citrate, pH 7.0), 1% SDS, 5× Denhardt's solution, and 100 μg of denatured salmon sperm DNA/mL. An *Xho*I genomic DNA fragment derived from cosmid F42G9 extending from exon 5 to exon 8 was radiolabeled with the Prime-it kit (Stratagene) and used to probe the membrane.

For Northern analysis, whole RNA was treated with glyoxal and electrophoresed through 1.2% agarose in 10 mM phosphate buffer, pH 6.5 (Thomas, 1983). The RNA was transferred to an uncharged nylon membrane and hybridized as described for the Southern analysis except that the hybridization buffer contained 50% formamide and hybridization was done at 42 °C. The probe was made from the cloned RT-PCR product of primers 3U and 8L as described above. Both Northern and Southern blots were washed at room temperature in 2× SSC, 0.1% SDS, and at 50−60 °C in 0.1× SSC, 0.1% SDS.

cDNAs. *ptl-1* cDNAs were isolated from a commercial mixed stage library made with the Uni-Zap XR vector (Stratagene). Filters representing 5×10^5 phage were hybridized as described for Southern analysis with the probe used for Northern analysis. Phage were purified through secondary platings. Plasmids containing the cDNAs were excised from the isolated phage by the ExAssist protocol in SOLR cells (Stratagene).

5′-RACE. The 5′-end of the *ptl-1* transcript was identified by 5′-RACE using the Gibco BRL 5′-RACE System for Rapid Amplification of cDNA Ends. Following the manufacturer's protocol, the cDNA strand was prepared using primer 5L, followed by a first round of amplification with the kit anchor primer and primer 3L. A second round of amplification used the kit universal amplification primer and primer 2L. The products were isolated from agarose gels and ligated into the pGEM-T vector (Promega) for sequencing.

Computer Analysis. The *ptl-1* genomic sequence was determined by the *C. elegans* Genome Sequencing Consortium (Sulston, 1992) and made available through ACEDB (Durbin & Mieg, 1991). The program Genefinder in ACEDB identified a possible gene that included eight exons (Figure 1). Homologies were identified with the Blast program on a computer at NCBI (Altschul et al., 1990). The Wisconsin package of programs (Genetics Computer Group, 1994) was used for assembling sequences, for identifying potential modification sites, and for predicting secondary structure and protein flexibility. Design of the oligonucleotides was aided by the program Oligo (National Biosciences), and the predicted amino acid content and isoelectric points were calculated using DNAsis (Hitati Software Engineering Co.).

Microtubule Isolation. Microtubules were isolated from rabbit brains by the method of Vallee (1982) as modified by Aamodt and Culotti (1986). MAPs were eluted from the microtubules with 0.4 M NaCl, and the microtubules were stored in liquid nitrogen.

Expression of a PTL-1a fusion protein in Escherichia coli. A fragment of *ptl-1a* extending from exon 3 through exon 8 with *Bam*HI and *Hin*dIII restriction sites was generated by PCR using oligonucleotides 5′-AGATATGGATCCGAGAGAGCTCGAGTCTCTAA and 5′-CCGATCATAACAAGCTTATGTCCAGCGTAG. The PCR product was cloned into the pET21b vector (Novagen) between the *Bam*HI and *Hin*dIII sites with the PTL-1a sequence fused in-frame to the N-terminal T7 epitope tag. Sequence analysis confirmed that the insert was *ptl-1a* cDNA. This plasmid was named pTAUe. pTAUe was transformed into bacterial strain BL21-(pLys S), and expression was induced by the addition of 1.0 mM IPTG to cells growing in M9 media.

Microtubule-Binding Assay and Western Analysis. Bacteria containing recombinant PTL-1a were resuspended in microtubule assembly buffer (50 mM Pipes, 1.0 mM EGTA, 1.0 mM $MgSO_4$, 1.0 mM GTP, and 10 μM taxol), lysed, and centrifuged at 30 000*g* for 15 min. The supernatant is referred to as bacterial extract. Taxol-stabilized rabbit brain microtubules (200 μg) were mixed with 25, 50, or 100 μL of bacterial extract or 100 μL of microtubule assembly buffer and incubated at 30 °C for 5 min. The mixtures were centrifuged through 10% sucrose cushions at 30 000*g* for 15 min. The pellets were rinsed and resuspended in 40 μL of microtubule assembly buffer. Half of the samples were reserved for analysis, and NaCl was added to a concentration of 0.6 M to the remainder of the samples. The samples were centrifuged at 30 000*g* for 15 min, and the supernatants were recovered. The samples were electrophoresed on 10% acrylamide SDS−PAGE gels by standard methods. Duplicate gels were stained with Coomassie Blue or transferred to an Immobilon-P polyvinylidene difluoride membrane (Millipore) for Western analysis. Western analysis of the bacterially expressed PTL-1a was done as described (Harlow & Lane, 1988) with a primary antibody (Novagen) against the T7 epitope (1:2000 dilution) and an alkaline phosphatase-

conjugated goat anti-rabbit secondary antibody (1:2000 dilution).

RESULTS

ptl-1 Predicted Gene Structure. The *C. elegans* Genome Sequencing Consortium identified genomic sequences with strong sequence homology to the repeat region of the microtubule-binding domain of the τ-class of MAPs. The exons identified by the Consortium with the program Genefinder (Durbin & Mieg, 1991) correspond to exons 1−8 in Figure 1. Exon 1 begins with AACAAAAATGA, which matches the *C. elegans* consensus for translation initiation. Upstream from this ATG by 116 bp is a TATAA sequence. Exons 2 and 3 contain homologies to mammalian neurofilament H1 protein (and several other proline-rich proteins), which is thought to be an extended arm-like protein (Lees et al., 1988). Exon 2 contains a distinctive PE repeat sequence that occurs in several proteins including proline-rich protein from wheat (Raines et al., 1991), procyclin from *Trypanosoma brucei* (Jackson et al., 1993) and several other proteins. The structure and role of this sequence are not known. Exons 5−8 show homology to the repeat region of mammalian MAPs (see Figure 4).

ptl-1 cDNAs. We isolated five *ptl-1* cDNA clones representing two RNA isoforms using an RT-PCR-generated probe extending from exon 3 to exon 8, including the putative microtubule-binding domain (Figure 1). The two isoforms, which we refer to as *ptl-1a* and *ptl-1b*, differed at their 3′-ends. The four *ptl-1a* clones contained exons 2 through 8, including the predicted 3′-untranslated region (3′-UTR) of the gene and a poly-A tail. Exon 7 was 14 bp longer than predicted by Genefinder, and the reading frame of exon 8 was changed by the addition of this extra sequence. The *ptl-1b* cDNA did not contain exons 7 or 8 but joined exon 6 to an exon (exon 9 in Figure 1) farther downstream, which was not predicted by Genefinder. Exon 9 lies 81 bp downstream of the polyadenylation sequence of exon 8; the resulting intron between exons 6 and 9 is 474 bp. Exon 9 encodes 30 amino acids and also contains a polyadenylation site. Thus, the predicted carboxyl-termini of the PTL-1a and PTL-1b proteins differ from one another and from the predicted protein.

We performed 5′-RACE and RT-PCR experiments to determine the transcriptional initiation site and to assay for transcription of predicted exon 1. The 5′-ends of the cDNAs are marked in Figure 1 with asterisks. The longest *ptl-1a* cDNA and the *ptl-1b* cDNA began with the first base of predicted exon 2. The major products of two 5′-RACE reactions were cloned and sequenced to determine the sequence of full-length transcripts. In each case, the trans-spliced leader sequence SL1 (Krause & Hirsh, 1987) was joined to the 5′ splice acceptor site of exon 2. These results show that at least some of the transcripts are trans-spliced and suggests that most *ptl-1* mRNAs in our RNA preparation were trans-spliced to SL1 at the 5′ splice site of exon 2. RT-PCR amplification of whole RNA from a mixed stage population of *C. elegans* failed to produce evidence for mRNAs containing predicted exon 1 (Table 1). In each experiment in which primer 1U was used with a downstream antisense primer, no band was detected that corresponded in size to the predicted cDNA. Control RT-PCR experiments done with primers to the other exons had major bands that corresponded to the predicted cDNAs.

Table 1: RT-PCR Products[a]

	1U	3U	4U	5U	6U
1L	*****	*****	*****	*****	*****
2L	G	*****	*****	*****	*****
3L	G	*****	*****	*****	*****
5L	NT	+	+	*****	*****
6L	G	+	+	+	*****
7L	G	+	+	+	+
8L	NT	+	+	+	NT

[a] "+" indicates that the major band was the size predicted for the cDNA. "G" indicates that the major product is the size predicted for genomic DNA. "NT" indicates a primer combination that was not tested.

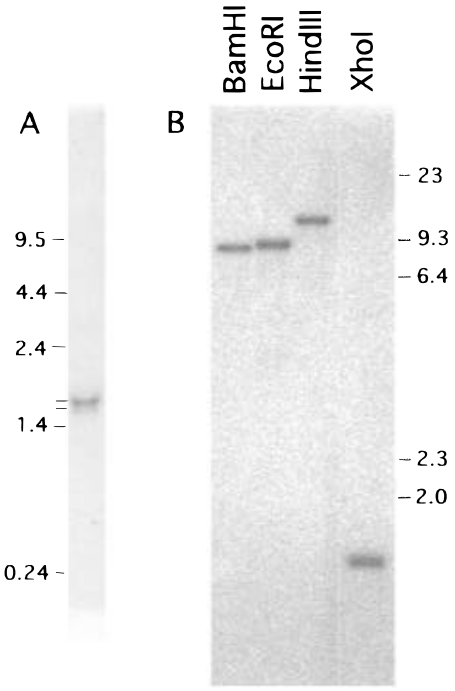

FIGURE 2: (A) Northern analysis of whole *C. elegans* RNA detected two *ptl-1* transcripts (lines). RNA markers (kb) are indicated on the left. (B) Southern analysis of *C. elegans* genomic DNA digested with *Bam*HI, *Eco*RI, *Hin*dIII, and *Xho*I. Molecular weight markers (kbp) are indicated on the right. The blot shown here was probed as described in Materials and Methods, but the same bands were seen after a low-stringency hybridization (2× SSC, 0.1% SDS, 37 °C).

Northern and Southern Analysis. Northern and Southern analyses indicated that *ptl-1* is a single-copy gene that produces at least two transcripts. Northern hybridization of RNA prepared from mixed stage animals revealed the presence of two RNA species of approximately 1.5 kb (Figure 2A), in close agreement with the lengths of the *ptl-1a* and *ptl-1b* cDNAs, which are 1403 and 1526 bp, respectively. *C. elegans* genomic DNA was digested with several restriction enzymes and probed with a genomic *Xho*I fragment that contained the entire repeat region of *ptl-1a*. For each restriction enzyme used, the probe hybridized a single fragment, even at low stringency (Figure 2B). Thus, the *C. elegans* genome apparently has only one gene with homology to mammalian τ.

Predicted PTL-1 Proteins. The predicted PTL-1a and PTL-1b proteins can be divided into domains that are similar to the domains of mammalian τs (Gustke et al., 1994). Figure 3 shows that exons 2 and 3 encode an acidic proline-rich region similar to the amino-terminus of mammalian τ.

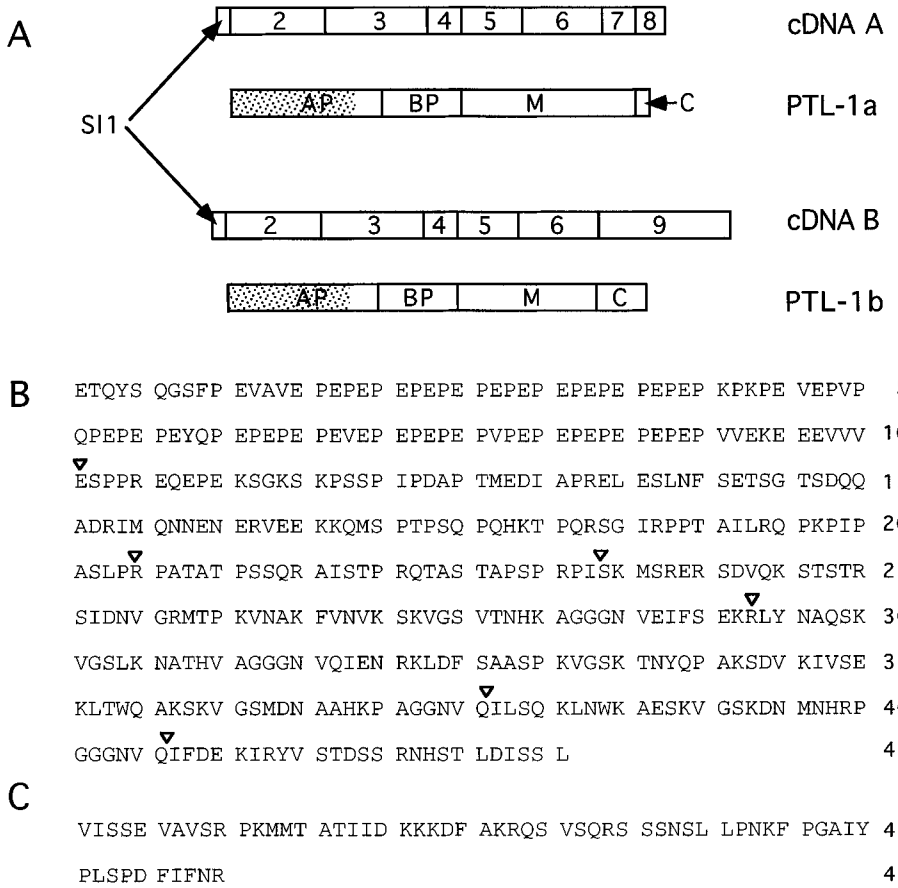

FIGURE 3: (A) Domain structure of the predicted products of cDNA A (PTL-1a) and cDNA B (PTL-1b). Both *ptl-1* products are predicted to contain an acidic proline-rich region at their N-terminus (AP) followed by a basic proline-rich region (BP). M indicates the region with homology to the repeats of the τ microtubule-binding domain and C indicates the carboxyl-terminus. The stippled region of AP is to indicate that the translation start site is unknown. (B) The predicted coding sequence of PTL-1a from the 5′-end of exon 2. Arrowheads mark the positions of introns. (C) The C-terminal end of PTL-1b is encoded in exon 9.

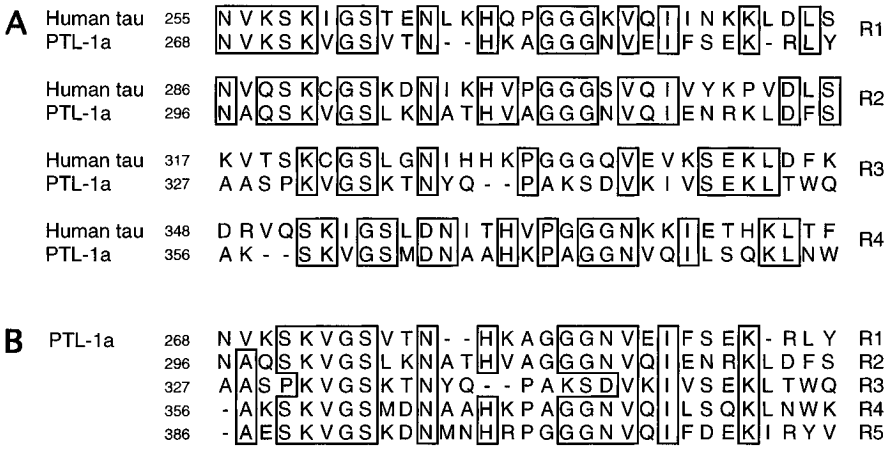

FIGURE 4: (A) Comparison of the conserved repeat region of PTL-1a and the human τ isoform hτ40 [accession number 88154 (Goedert et al., 1989a)]. Identities are indicated by boxes. The amino acid residue numbers are indicated on the left. The alignments were generated with the Gap program of the Wisconsin Package (Genetics Computer Group, 1994), and spaces were inserted into the PTL-1a sequence by the program to maintain the alignment. (B) PTL-1a repeats aligned with each other.

Exons 3 and 4 encode a basic proline-rich region. Exons 5−8 in PTL-1a and exons 5, 6, and 9 in PTL-1b encode a sequence with strong homology to the repeat region of the mammalian τ microtubule-binding domain (Figure 4A). There was no convincing sequence homology to mammalian τs outside the repeats.

In mammals, alternatively spliced τ variants can differ in the number of repeats they contain. *C. elegans* τ variants also differ in the number of encoded repeats. PTL-1a encodes five repeats, and PTL-1b encodes four repeats (Figure 3B,C). PTL-1a is 48% identical to mammalian τ over the repeat region, allowing the gaps shown in Figure 4A. The PTL-1 repeats are more similar to each other than to the repeats of mammalian MAPs, as shown in Figure 4B. Mammalian τs show a maximum of four repeats (R1−R4) and a fifth weakly conserved repeat referred to as R′ (Gustke et al., 1994). PTL-1a contains five repeats. The homology is weakest in the third repeat of PTL-1a, but this repeat still has stronger homology to the other repeats of the *C. elegans* protein than to the R′ of mammalian τs.

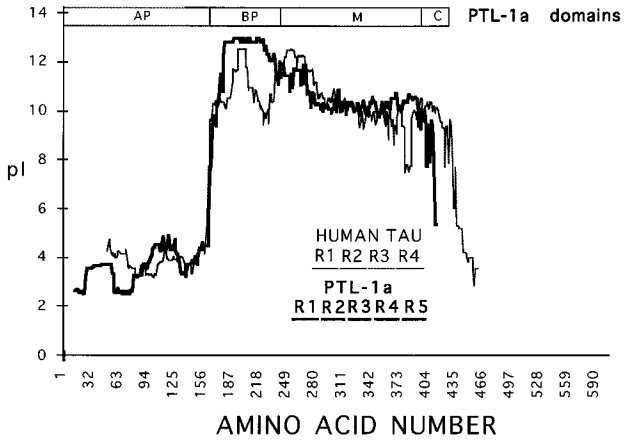

FIGURE 5: Charge distribution of PTL-1a (black line) and human τ (gray line). Each point represents the calculated p*I* for a 30-amino acid stretch. The sequences were aligned so that the sharp transitions between the acidic and the basic proline-rich sequences were aligned. The position of the repeats are indicated for both human τ and PTL-1a.

PTL-1a is strikingly similar to human τ in amino acid composition, charge distribution (Figure 5), predicted structure, and modification sites (Figure 6). Both PTL-1 and human τ are predicted to be highly hydrophilic, and most of both sequences are predicted to be at the surface (Figure 6), probably because these are extended hydrophilic proteins with little secondary or tertiary structure. Both are predicted to be highly flexible, and both contain many potential phosphorylation sites and several potential glycosylation sites. In addition to the predicted phosphorylation sites for casein kinase II and protein kinase C indicated in Figure 6, five potential sites for proline-directed phosphorylation (Lew & Wang, 1995) are found between Ser102 and Thr180. The site at Thr180 is a consensus site for cdc2 kinase phosphorylation. Two potential phosphorylation sites are found in the final exon of PTL-1b.

Bacterially expressed PTL-1a fusion protein bound to microtubules and could be eluted from the microtubules with high salt. The fusion protein contains an N-terminal T7 epitope tag followed by amino acids 132−429 of PTL-1a (from exon 3 to the C-terminus). SDS−PAGE analysis of bacterial lysates revealed that two polypeptides were induced with IPTG and that the supernatant fraction (lane S in Figure 7) contained much of the expressed proteins (data not shown). The predicted molecular weight of the fusion protein was 37 kDa, and the apparent molecular weights of the induced proteins were near 45 kDa (Figure 7). To test for microtubule binding, bacterial extract was mixed with taxol-stabilized rabbit brain microtubules. The microtubules and associated proteins were pelleted through sucrose. A salt extraction (0.6 M NaCl) was performed to elute MAPs, and the samples were run on gels. Figure 7 shows duplicate gels that were either stained with Coomassie Blue or transferred to a membrane for Western analysis with an antibody to the T7 epitope tag. Although the microtubules co-migrate with the induced proteins, comparison of the stained gel and Western blot shows that the epitope-tagged proteins are pelleted with microtubules. Both of the induced proteins bound microtubules and are eluted by salt. The antibody bound to both of the expressed proteins, but the induced protein of higher mobility was much more darkly stained than the other.

DISCUSSION

ptl-1 contains sequence homologous to the repeat regions of the mammalian MAPs τ, MAP2, and MAP4. We have demonstrated that *ptl-1* is transcribed and that the primary transcript is alternatively spliced to produce at least two mRNAs, *ptl-1a* and *ptl-1b*. The PTL-1a and PTL-1b predicted proteins have sequence homology to the τ-class of MAPs (Figure 4) and a domain structure similar to mammalian τs (Figure 3). Both proteins contain an acidic proline-rich amino-terminus, both have a basic proline-rich region, and the transition between acidic and basic domains is very sharp. Both proteins have very similar amino acid compositions and remarkably similar charge distributions (Figure 5). Mammalian τs are very hydrophilic and highly flexible proteins, and the *ptl-1* products are also predicted to be very hydrophilic and highly flexible (Figure 6). Mammalian τs are highly phosphorylated, and PTL-1a contains up to 20 predicted phosphorylation sites (Figure 6). Both the *ptl-1* and the mammalian τ gene transcripts are alternatively spliced so as to produce products with different numbers of repeats (Figure 1). Mammalian τs contain either four repeats plus R′ or three repeats plus R′. PTL-1a contains five repeats, and PTL-1b contains four repeats. A bacterially expressed fusion protein containing the PTL-1a microtubule-binding region bound to microtubules and was eluted from the microtubules by high salt (Figure 7). On the basis of these similarities, we predict that the products of *ptl-1* perform many of the same functions as τ in mammals.

MAPs isolated from *C. elegans* were previously shown to contain proteins with apparent molecular weights similar to mammalian τ (Aamodt & Culotti, 1986). Subsequently, Western analysis of *C. elegans* MAPs and whole *C. elegans* proteins, probed with an antibody against mammalian MAPs, revealed four proteins with apparent molecular weights similar to mammalian τs and the predicted *ptl-1* gene products (S. Aamodt, Y. Jia, and E. Aamodt, unpublished results). It is not yet clear whether these proteins include PTL-1a or PTL-1b.

The cDNAs sequences and the Northern analysis showed that at least two transcripts of approximately 1.5 kb are produced. We have not determined which species on the Northern blot corresponds to *ptl-1a* and *ptl-1b*. In RNA from mixed stages of *C. elegans*, the larger transcript is more abundant. However, four of the five isolated cDNAs represent *ptl-1a*, which is 120 bases shorter than *ptl-1b*. Polyadenylation may account for this apparent discrepancy. In addition, it is possible that other transcripts of similar sizes are produced.

The sequence of the *ptl-1a* cDNA corresponded to the sequence of the transcript predicted by the *C. elegans* Sequencing Consortium with two exceptions. The cDNA clones do not include the first predicted exon, and the 3′ splice site of exon 7 was downstream of the predicted site. Examination of the splice sites reveals that the predicted 5′ and 3′ splice sites match the *C. elegans* consensus splicing sequences well, with the exception of the 3′ splice donor site of exon 1 and the 3′ splice donor site of exon 7. Because the *ptl-1* mRNAs are trans-spliced, the transcriptional start site of *ptl-1* is not known. The failure of Genefinder to predict the 3′ splice site of exon 7 is probably due to the small size of the intron between exons 7 and 8. The intron is 49 bp, one base pair smaller than the intron size allowed

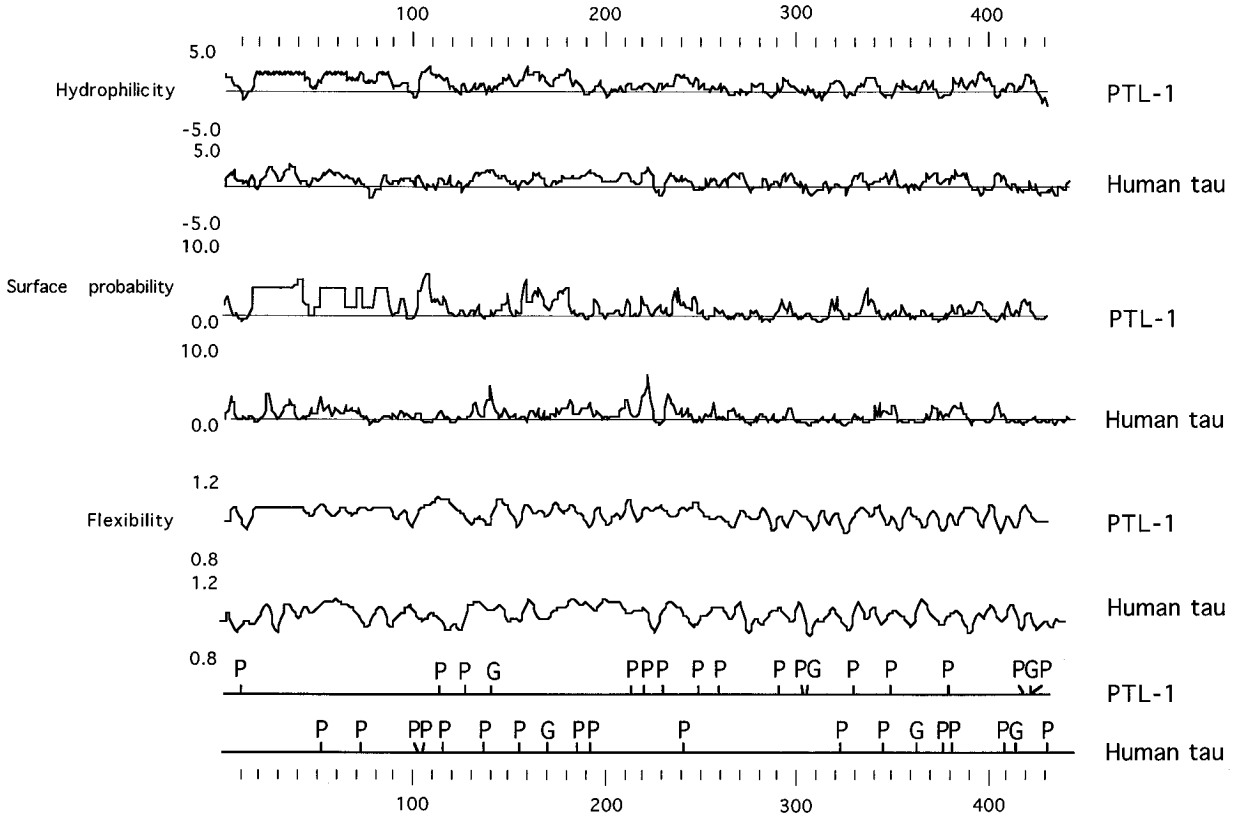

FIGURE 6: Comparison of PTL-1a and human τ predicted hydrophilicity, surface probability, flexibility, and predicted modification sites. The position of predicted casein kinase II and protein kinase C phosphorylation sites (P) and N-glycosylation sites (G) are marked. The data were generated with the Wisconsin Package (Genetics Computer Group, 1994) using the Prosite database.

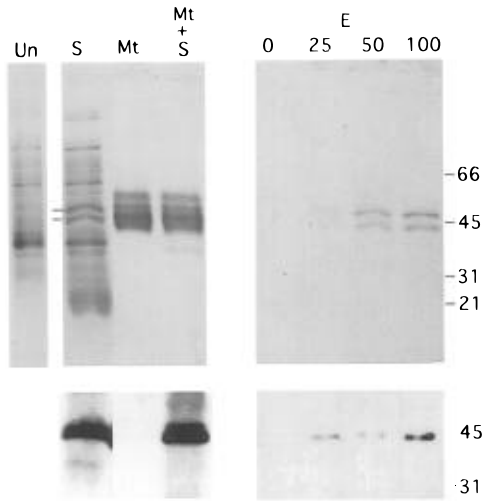

FIGURE 7: PTL-1a fusion protein binding to microtubules. Top: SDS−PAGE analysis of samples from the binding assay. Un, uninduced bacterial cells; S, supernatant from bacterial cells expressing PTL-1a fusion protein (the induced bands are marked with lines); Mt, microtubule pellet without bacterial extract; Mt + S, pellet from microtubules incubated with 100 μL of bacterial extract; E, proteins eluted from the microtubules with 0.6 M NaCl. The lanes show elution from microtubule pellets incubated with 0, 25, 50, and 100 μL of bacterial extract. Bottom: Western blot of a duplicate gel showing that the T7-tagged PTL-1 fusion proteins in the bacterial extract were precipitated with microtubules and were eluted from the microtubules with 0.6 M NaCl.

by Genefinder's default parameters. The true 3′ splice donor site of exon 7 matches the consensus better than the predicted site.

The most likely site of transcription initiation is between predicted exon 1 and the TATA box 116 bp upstream. There are no obvious TATA boxes in the sequences between exons 1 and 2. Thus, the *ptl-1* primary transcript may initiate upstream of predicted exon 1 and extend to beyond exon 9. Most, if not all, of the *ptl-1* primary transcripts undergo trans-splicing, a common phenomenon in nematode RNA processing. 5′-RACE experiments revealed that splice leader 1 (SL1) is joined to exon 2, excising the sequences upstream of exon 2. These experiments do not eliminate the possibility that *ptl-1* transcripts vary at the 5′-end. Standard methods of mapping transcriptional start sites cannot be applied to trans-spliced transcripts.

The translational start codon for the *ptl-1* transcripts is unknown. The first AUG is in exon 3, which is 400 bases from the 5′-end of the transcripts. The sequence at this site (CCCAACAATGG) matches the consensus for translation initiation in *C. elegans* (Epstein & Shakes, 1995), but translation of most trans-spliced messages begins within a few bases of the splice leader (Epstein & Shakes, 1995). Although little is known about translation initiation from non-AUG codons in *C. elegans*, translation may initiate from some alternative codon, as has been observed in other systems (Boeck & Kolakofsky, 1994). That exon 2 may be translated at times is suggested by its highly favorable codon bias, the presence of the SL1 leader, and the homology of exon 2 to other proteins. Alternatively, exon 2 could be translated from another mRNA species that includes exon 1.

Analysis of the predicted amino acid sequence of exon 1 indicates this region could be a transmembrane domain. Electron micrographs often show cross-links between microtubules and membranes (including the plasma membrane), but no protein has yet been isolated and shown to form this

type of connection. Such a protein would represent an important new class of MAPs. Exon 1 contains an AUG in a context typical of translational initiation sites in *C. elegans*. We attempted, therefore, to determine whether a mRNA was made using exon 1 and exon 2, 3, 6, or 7. RT-PCR on RNA isolated from a developmentally mixed population of animals generated products of the expected size for other exons but not for exon 1 (Table 1). Although these experiments confirmed that the cDNA structures shown by sequencing cloned cDNAs correspond to the major products of the *ptl-1* gene, we obtained no evidence for transcription of exon 1. While PCR is a sensitive method, we cannot conclude that exon 1 is never present in a message together with the downstream *ptl-1* exons. Alternatively spliced variants containing exon 1 could exist at very low levels or in very few cells or for a very short period of development or under conditions not reproduced in the laboratory.

Alternative splicing of mRNAs is characteristic of mammalian τ genes. Many spliced isoforms of human τ, for example, have been isolated (Sawa et al., 1994). Splicing of multiple isoforms gives rise to heterogeneity of mammalian τs both at the N-terminus and within the repeat region (Himmler, 1989). In addition to the *ptl-1a* and *ptl-1b* mRNAs, other *C. elegans ptl-1* RNA isoforms could be generated using the observed splice sites. *ptl-1a* exons 5−8 encode repeats R1 through R5; *ptl-1b* exons 5, 6, and 9 encode repeats R1 through R4. If variants encoding exons 5, 6, and 8 or exons 5, 6, 7, and 9 are produced, then the corresponding proteins would include four and five repeats, respectively. All of the splice junctions within the repeat region fall between codons, so that alternative splicing of transcripts would not change the reading frame and could generate multiple proteins containing various repeats in a modular fashion. Heterogeneity at the N-terminus could be generated by the inclusion of exon 1 or the excision of exon 3.

If *ptl-1* encodes *C. elegans* τ MAPs, as suggested by these results, then the τ MAPs must have evolved very early and τs may be present in most animal phyla. Furthermore, the fact that the repeat region is so well conserved over such a long period of evolution shows that the repeats are very important for the function of the τ class of MAPs.

Whereas mammalian genomes contain at least three genes of the τ family (the genes for τ, MAP2, and MAP4), Southern analysis suggests that in *C. elegans*, *ptl-1* is a single-copy gene. Therefore, radiation of the τ-class of MAPs may have occurred after the vertebrate and invertebrate lines separated during evolution, and the *ptl-1* products may serve all of the functions that τ, MAP2, and MAP4 serve in mammals. In this respect, it probably is not possible to know whether the *ptl-1* gene products are more like mammalian τ or MAP2 or MAP4. The homology of the repeats is slightly closer to τ, and PTL-1a and PTL-1b are predicted to be the size of τs; however, like MAP4, the *ptl-1* products can have five repeats.

While the τ-class of MAPs has been extensively studied and while many interesting and provocative experiments have suggested important roles for these proteins, it is still not clear what they do *in vivo*. *C. elegans* is a system in which the powerful tools of genetics can be brought to bear on the role of τ-like proteins. The location of PTL-1a and PTL-1b at each stage of development can be determined. *ptl-1* can be eliminated or over-expressed in particular cells at particular points in development, and once *ptl-1* mutants are obtained, finding proteins that interact with the PTL-1 proteins should be possible by standard genetic procedures.

ACKNOWLEDGMENT

We thank Keith Simpson and Travis Harrington for technical assistance, Dr. Alan Coulson for cosmid F42G9, and Dr. G. R. Vandenbark for the cDNA library. We are especially grateful to Dr. Peter Good and Dina Herring for their assistance with expression and analysis of the PTL-1 fusion protein.

REFERENCES

Aamodt, E. J., & Culotti, J. G. (1986) *J. Cell Biol. 103*, 23−31.
Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403−410.
Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., & Struhl, K. (1993) *Current Protocols in Molecular Biology*, Current Protocols, New York.
Barlow, S., Gonzalez-Garay, M. L., West, R. R., Olmsted, J. B., & Cabral, F. (1994) *J. Cell Biol. 126*, 1017−1029.
Boeck, R., & Kolakofsky, D. (1994) *EMBO J. 13*, 3608−3617.
Bramblett, G. T., Goedert, M., Jakes, R., Merrick, S. E., Trojanowski, J. Q., & Lee, V. M.-Y. (1993) *Neuron 10*, 1089−1099.
Brenner, S. (1974) *Genetics 77*, 71−94.
Caceres, A., & Kosik, K. S. (1990) *Nature 342*, 461−463.
Cambiazo, V., Gonzalez, M., & Maccioni, R. B. (1995) *J. Neurochem. 64*, 1288−1297.
Chapin, S. J., & Bulinski, J. *C.* (1991) *J. Cell Science 98*, 27−36.
Chapin, S. J., & Bulinski, J. C. (1992) *Cell Motil. Cytoskeleton 23*, 236−243.
Drubin, D. G., & Kirschner, M. W. (1986) *J. Cell Biol. 103*, 2739−2746.
Durbin, R., & Mieg, J. T. (1991) *A C. elegans Data Base.* Documentation and code data available for anonymous FTP servers at lirmm.tilmm.fr, cele.mrc-lmb.cam.ac.uk, and ncbi.nlm.nih.gov.
Epstein, H. F., & Shakes, D. C. (1995) in *Methods in Cell Biology* (Wilson, L., & Matsudaira, P., Eds.), Academic Press, San Diego.
Esmaeli-Azad, B., McCarty, J. H., & Feinstein, S. C. (1994) *J. Cell Sci. 107*, 869−879.
Genetics Computer Group (1994) *Program Manual for the Wisconsin Package, Version 8*, Genetics Computer Group, Madison, WI.
Goedert, M., Spillantini, M. G., Jakes, R., Rutherford, D., & Crowther, R. A. (1989a) *Neuron 3*, 519−526.
Goedert, M., Spillantini, M. G., Potier, M. C., Ulrich, J., & Crowther, R. A. (1989b) *EMBO J. 8*, 393−399.
Goedert, M., Crother, R. A., & Garner, C. C. (1991) *Trends Neurosci. 14*, 193−199.
Grunke-Iqbal, I., Iqbal, K., Quinlan, M., Tung, Y.-C., Zaidi, M. S., & Wisniewski, H. M. (1986) *J. Biol. Chem. 261*, 6084−6089.
Gustke, N., Trinczek, B., Biernat, J., Mandelkow, E.-M., & Mandelkow, E. (1994) *Biochemistry 33*, 9511−9522.
Harada, A., Oguchi, K., Okabe, S., Kuno, J., Terada, S., Ohshima, T., Sato-Yoshitake, R., Takel, Y., Noda, T., & Hirokawa, N. (1994) *Nature 369*, 488−491.
Harlow, E., & Lane, D. (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY.
Himmler, A. (1989) *Mol. Cell Biol. 9*, 1389−1396.
Hodgkin, J. (1995) *Trends in Genetics: Genetic Nomenclature Guide*, pp 24−25, Elsevier Trends Journals, Tarrytown, NY.
Jackson, D. G., Smith, D. K., Luo, C., & Elliott, J. F. (1993) *J. Biol. Chem. 268*, 1894−1900.
Kenner, L., el-Shabrawi, Y., Hutter, H., Forstner, M., Zatloukal, K., Hoefler, G., Preisegger, K. H., Kurzbauer, R., & Denk, H. (1994) *Hepatology 20*, 1086−1089.
Knops, J., Kosik, K. S., Lee, G., Pardee, J. D., Cohen-Gould, L., & McConlogue, L. (1991) *J. Cell Biol. 114*, 725−733.
Kosik, K. S., & Caceres, A. (1991) *J. Cell Sci.* (Suppl. 15), 69−74.

Kosik, K. S., Orecchio, L. D., Binder, L., Trojanowski, J. Q., Lee, V. M., & Lee, G. (1988) *Neuron 1*, 817−825.
Krause, M., & Hirsh, D. (1987) *Cell 49*, 753−761.
Lee, G. (1993) *Curr. Opin. Cell Biol. 5*, 88−94.
Lee, V. M.-Y., & Trojanowski, J. Q. (1992) *Curr. Opin. Neurobiol. 2*, 653−656.
Lees, J. F., Shneidman, P. S., Skuntz, S. F., Carden, M. J., & Lazzarini, R. A. (1988) *EMBO J. 7*, 1947−1955.
Lew, J., & Wang, J. H. (1995) *Trends Biochem. Sci. 20*, 33−37.
Lewis, S. A., Wang, D., & Cowan, N. J. (1988) *Science 242*, 936−939.
Lichtenberg-Kraag, B., & Mandelkow, E.-M. (1990) *J. Struct. Biol. 105*, 46−53.
Maes, M., & Messens, E. (1992) *Nucleic Acids Res. 20*, 4374.
Murphy, D. B., & Wallis, K. T. (1985) *J. Biol. Chem. 260*, 12293−12301.
Raines, C. A., Lloyd, J. C., Chao, S. M., John, U. P., & Murphy, G. J. (1991) *Plant Mol. Biol. 16*, 663−670.
Sawa, A., Oyama, F., Matsushita, M., & Ihara, I. (1994) *Mol. Brain Res. 27*, 111−117.
Sulston, J. (1992) *Nature 356*, 37−41.
Sulston, J., & Hodgkin, J. (1988) in *The Nematode Caenorhabditis elegans* (Wood, W. B., Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
Thomas, P. S. (1983) *Methods Enzymol. 100*, 255−266.
Vallee, R. B. (1982) *J. Cell Biol. 92*, 435−442.
Vantard, M., Schellenbaum, P., Fellous, A., & Lambert, A. M. (1991) *Biochemistry 30*, 9334−9340.
Vantard, M., Peter, C., Fellous, A., Schellenbaum, P., & Lambert, A.-M. (1994) *Eur. J. Biochem. 220*, 847−853.
Wiche, G., Oberkanins, C., & Himmler, A. (1991) *Int. Rev. Cytol. 124*, 217−273.
Wischik, C. M., Novak, M., Edwards, P. C., Klug, A., Tichelaar, W., & Crowther, R. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4884−4888.
Wood, J. G., Mirra, S. S., Pollock, N. J., & Binder, L. I. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 4040−4043.

BI952646N