

# Identification of conserved isotype-defining variable region sequences for four vertebrate $\beta$ tubulin polypeptide classes

(microtubules)

KEVIN F. SULLIVAN AND DON W. CLEVELAND

Department of Biological Chemistry, The Johns Hopkins University School of Medicine, Baltimore, MD 21205

Communicated by Bruce Alberts, March 6, 1986

**ABSTRACT** We report the determination of the complete sequences for two chicken  $\beta$  tubulin genes,  $\beta 3$  and  $\beta 5$ . Taken with the previously published efforts, we have determined the primary structures of five of the seven  $\beta$  tubulin genes in this vertebrate species. A comparison of these sequences unambiguously reveals that amino acid sequence variations among different  $\beta$  tubulin gene products are distinctly clustered within an otherwise highly conserved framework of the  $\beta$  tubulin molecule. To determine the extent to which this pattern of structural heterogeneity is conserved among vertebrates, we have isolated novel  $\beta$  tubulin sequences from human and mouse cDNA libraries and compared these and all other known vertebrate  $\beta$  tubulin sequences with the family of chicken polypeptide sequences. What emerges from such comparison is the recognition of distinct, evolutionarily conserved isotypes of  $\beta$  tubulin that are distinguished primarily by their characteristic carboxyl-terminal variable region sequence and, to a lesser extent, by sequence in an amino-terminal variable domain as well. These correlations represent a convincing demonstration that multiple  $\beta$  tubulin genes in vertebrates encode a family of closely related but structurally distinct  $\beta$  tubulin isotypes and further serve to define the sequences of four classes of polypeptide isotypes that constitute that family.

Microtubules are ubiquitous, cytoskeletal components of eukaryotic cells. Composed of filamentous, polymeric assemblies of a heterodimer of one  $\alpha$  and one  $\beta$  tubulin polypeptide, microtubules serve as the primary structural substrates for such processes as mitosis, flagellar and ciliary-based motility, and cytoplasmic transport (e.g., refs. 1-4). In addition, in concert with actin filaments and intermediate filaments, microtubules play a major role in establishing the dynamic spatial organization of the cytoplasm (e.g., ref. 5). However, it is not known how the assembly of microtubules that participate in different cellular events is specified (either spatially or temporally) nor, in fact, is it known in what ways microtubules of apparently different functions are structurally distinct.

Nonetheless, it has become increasingly clear that at least some subtle variations in the molecular properties of microtubules do exist. For example, dramatic differences in distribution of microtubule-associated proteins have been seen among and within different tissues and during development (6-9). Significant differences in the programs of expression of the small family of genes that encode  $\alpha$  and  $\beta$  tubulin are also seen during development and differentiation (10, 11). Further, microtubules even within single cells have been shown to differ in patterns of posttranslational modification (12-14). Clearly, understanding the factors that determine and regulate the molecular differentiation of microtubules

will be fundamental in understanding the role of microtubules in cell function.

The possible role of the tubulin subunit itself as a modulator that can specify microtubule function remains enigmatic. Our initial demonstration that eukaryotic genomes contain multigene families that encode  $\alpha$  and  $\beta$  tubulin (ref. 15; reviewed in ref. 16) rekindled earlier speculations that such organisms possessed functionally differentiated tubulins that could be used to construct different kinds of microtubules (17). Indeed, although initial comparative sequence analysis of  $\alpha$  and  $\beta$  tubulin genes indicated that the encoded polypeptides were among the most highly conserved during evolution (18-21), subsequent studies have demonstrated a significant degree of heterogeneity among  $\beta$  tubulins, both within and between species (11, 22-25).

In this report we have used DNA sequence analysis to determine the complete protein coding sequences of two chicken  $\beta$  tubulin genes. Taken with our previously published efforts for three other chicken  $\beta$  tubulins (23, 26), we have now determined the primary structures of five of the seven  $\beta$  tubulin genes in this species. In addition, by similar analyses of newly isolated cDNA clones, we report the predicted polypeptide sequences for one novel human and one novel mouse  $\beta$  tubulin and extend the known sequences for two other mouse  $\beta$  tubulins. Comparison of the sequences of these newly isolated genes with all other known vertebrate  $\beta$  tubulins reveals the existence of a conserved family of isotypically distinct classes of  $\beta$  tubulin polypeptides. We find that these isotypic classes differ primarily in two discrete locations within the primary sequences and, so far as it is known, each gene/isotypic class is expressed in analogous developmental programs in each species investigated. Taken together, we believe that these results unambiguously demonstrate the existence of a structurally heterogeneous family of vertebrate  $\beta$  tubulins and define the sequences of four major classes of polypeptide isotypes that constitute that family.

## MATERIALS AND METHODS

The isolation of genomic sequences bearing the chicken  $\beta 3$  and  $\beta 5$  tubulin genes has been described (24, 27). The complete nucleotide sequences and characterization of these genes will be reported elsewhere. Mammalian cDNAs were isolated from a human placental cDNA library constructed in  $\lambda$  gt11 (courtesy of Dan Littman and Moses Chow) and from a mouse 3T3 cell library constructed in  $\lambda$  gt10 (courtesy of Lester Lau). Isolated cDNAs were then subcloned into M13 vectors (28) for DNA sequence analysis, which was performed by the dideoxy method (29). Novel cDNA clones whose sequences are reported in this work are ph $\beta 4$ , a 1.47-kilobase (kb) human cDNA derived from a human gene analogous to the chicken  $c\beta 4$  gene (23), and pm $\beta 4'$ , a 1.4-kb mouse 3T3 cell  $\beta$  tubulin cDNA that derives from a gene analogous to the human gene  $h\beta 2$  (30). In addition, a nearly full-length cDNA copy of the mouse  $m\beta 5$  tubulin gene was

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

isolated (pm $\beta$ 5), which contains the entire coding region as well as extensive 5' and 3' untranslated sequences. The complete nucleotide sequences of these cDNA clones are available upon request.

## RESULTS

**Predicted Polypeptides Encoded by Chicken Genes  $\beta$ 3 and  $\beta$ 5.** In an effort to determine whether unique polypeptide sequences that might define distinct  $\beta$  tubulin isotypes are present in vertebrate species, we and our collaborators have attempted to isolate and sequence the entire repertoire of  $\beta$  tubulin gene segments in the chicken genome. Previously, we reported the complete sequence and characterization of three such genes ( $\beta$ 1 and  $\beta$ 2, ref. 26;  $\beta$ 4, ref. 23). We have now determined the complete sequence of two additional genes,  $\beta$ 3 (isolated in ref. 27) and  $\beta$ 5 (reported in ref. 24). With the exceptions of two final genes [gene  $\beta$ 4', which encodes a ubiquitously expressed  $\beta$  tubulin (10), and gene  $\beta$ 6, which encodes a highly divergent  $\beta$  tubulin apparently expressed exclusively in erythrocyte precursors (31, 32)], these five genes appear to comprise the complete repertoire of DNA segments within the chicken genome that have high nucleotide homology to cloned  $\beta$  tubulin sequences (see discussions in refs. 10 and 24).

The predicted polypeptide sequences from these five chicken  $\beta$  tubulin genes are displayed in Fig. 1. Inspection of Fig. 1 immediately confirms our initial suggestion that within an otherwise highly conserved framework there exist several clusters of amino acid substitutions that constitute the major regions of sequence heterogeneity among the five polypeptides (23). The most obvious of these variable regions comprises the extreme carboxyl-terminal residues, whereas areas of less striking but still significant divergence are found between amino acid positions 33–57, 80–95, and 288–335.

**Identification of Carboxyl-Terminal and Amino-Terminal Variable Region Domains.** To determine which (if any) of the differences among the five chicken sequences represent evolutionarily conserved differences that serve to define different  $\beta$  tubulin polypeptide isotypes, we have sought to identify and sequence  $\beta$  tubulin cDNA clones from two other vertebrate species. As the result of this effort, we have identified from a human placental library a cDNA clone (h $\beta$ 4) that represents a previously unidentified human  $\beta$  tubulin gene. In addition, from a mouse fibroblast library, cDNA clones for three different  $\beta$  tubulin mRNAs have been obtained. Two of these represent longer clones for the partial length m $\beta$ 2 and m $\beta$ 5 clones published by Cowan and co-workers (11), whereas the third clone, m $\beta$ 4' (extending from amino acid codon 91 into the 3' untranslated region), represents a novel mouse  $\beta$  tubulin. The sequences from each of these newly isolated clones as well as those determined by Cowan and his associates for three completely sequenced human (22, 30, 33) and three partially sequenced mouse genes (11) are also compiled in Fig. 1. With the addition of known porcine  $\beta$  tubulin sequences (19) and fragmentary sequence data from three rat  $\beta$  tubulins (25) and from one bovine  $\beta$  tubulin (35), all presently available vertebrate  $\beta$  tubulin sequences are represented in Fig. 1.

Comparison of these vertebrate  $\beta$  tubulin polypeptides (shown in Fig. 1) clearly demonstrates a significant degree of amino acid sequence heterogeneity that was not noted in the initial comparative analysis of  $\beta$  tubulin sequences (19). A total of 64 residue positions shows amino acid substitutions among the full-length sequences available for analysis. The distribution of these amino acid substitution positions within the primary sequence is not random but, as noted previously, shows clustering into more or less discrete regions of high amino acid substitution frequency (16, 22, 23). The highest degree of heterogeneity is seen in the carboxyl-terminal

sequence beyond amino acid 430, which displays variation at 85% (17 of 20) of residue positions, including insertions, deletions, and terminal addition of amino acid residues. A second site of intensive heterogeneity is seen in the amino-terminal region between residues 33 and 57, where substitutions have been accepted at 41% (10 of 24) of the residue positions. Together these two regions, which account for 9.7% of the polypeptide chain, possess 38% of the amino acid substitution sites within the vertebrate  $\beta$  tubulins.

**Identification of Isotype-Defining Consensus Sequences Within the Carboxyl-Terminal and Amino-Terminal Variable Regions.** In striking contrast to the intraspecies diversity demonstrated in Fig. 1, significant sequence homology is seen among variable region sequences from different organisms. Sequences from the carboxyl-terminal variable regions have been aligned on the basis of sequence homology in Fig. 2. Of the 17 sequences compared, 15 can be grouped into one of four classes (named class I–IV) of conserved sequence showing no more than 14% variation within a class. Two additional isotypes (the tentative testis class and an unassigned class) are defined by chicken genomic clones for which no cognate sequences have yet been isolated in another organism. Similarly, sequences in the amino-terminal variable region are conserved in different species (Fig. 1; summarized in Fig. 3), although the degree of variation among isotypes is less than that seen in the carboxyl-terminal sequences. Thus, it is clear that several distinct isotypic classes of  $\beta$  tubulin exist in vertebrates and, further, that these isotypic classes can be identified by a characteristic, highly conserved carboxyl-terminal variable region sequence.

**Identified  $\beta$  Tubulin Isotypes Share Homologous Patterns of Expression in Different Species.** Do these conserved  $\beta$  tubulin isotypes share functional properties within different vertebrate species? As a first approach to answering this question we have correlated isotypic classifications with available data on  $\beta$  tubulin gene expression to determine patterns of isotype usage in different vertebrates. Expression of  $\beta$  tubulin genes has been analyzed most extensively in the chicken (10) and in the mouse (11) and, to a lesser extent, in rat (25) and in a variety of human cell lines (22, 30).

The class I isotype is expressed in all tissues of the mouse and rat that have been surveyed as well as all human cell lines examined. We have thus designated the class I isotype as the major constitutively expressed  $\beta$  tubulin. Class II genes are also widely expressed at low levels within a variety of cell and tissue types but are detected predominantly in brain and in lung of both mouse and chicken. In addition, the class II gene encodes the major  $\beta$  tubulin transcript detected in isolated sympathetic neurons from the chicken. Thus, we have designated the class II isotype as a major neuronal  $\beta$  tubulin isotype.

The most highly divergent carboxyl-terminal sequence is present on the class III polypeptide, exemplified by the chicken c $\beta$ 4 gene. Our present isolation of a human  $\beta$  tubulin cDNA, h $\beta$ 4, which contains 85% of the coding sequence for a polypeptide virtually identical to the chicken c $\beta$ 4 gene product confirms that this gene is conserved in mammals. In addition, partial protein sequences obtained from pig (19) and bovine (35)  $\beta$  tubulins demonstrate the presence of polypeptides that share sequence homology in the amino-terminal variable region (residues 33–57 in Fig. 1), suggesting that the observed sequence conservation extends throughout the primary sequence. Expression of class III has been detected at the protein level exclusively in chordate brain (36). Similarly, mRNA corresponding to a class III gene has been detected only in neurons of the chicken (unpublished), where it is expressed at low levels relative to the class II transcript; thus, we have designated it as a minor neuronal isotype. (The

		10	20	30	40	50	60	70	80	90	100	110	120
CHICKEN $\beta 2$		MREIVHIQAGQCGNQIGAKFW	EVISDERGIDPTG	SYHGDS	DLQLERIN	VYNEATG	NGKYVP	RAILVD	LEPGTMD	SVRSGPFGQIF	RPDNFV	FGQSGAGN	NWAKGHYTEGAELVDSVLDVV
CHICKEN $\beta 1$							A						
CHICKEN $\beta 3$		L			T		G	V		R			
CHICKEN $\beta 4$				S N V		S	SSH		A H	I			
CHICKEN $\beta 5$			T	A G V	A		SSSQ	V		L	I T		
HUMAN M40				T		D S	G						
HUMAN $\beta 2$				T			G	V					
HUMAN $5\beta$		L		T			G	V					A
HUMAN $\beta 4$								[H $\beta 4$ :		A HL	I		
MOUSE M $\beta 4$								[M $\beta 4$ :					A
MOUSE M $\beta 4'$										[M $\beta 4'$ :			
MOUSE M $\beta 5$				T		D S	G						
PIG (A)							A						
PIG (B)				[V		S	SSH]						
BOVINE			[	S N V		S	SSH]						
		130	140	150	160	170	180	190	200	210	220	230	240
CHICKEN $\beta 2$		RKESESCDCLQGFLTHSLGGT	GSGMGTLLISKIRE	EYDPDRIM	TFVMPSPK	VSDTV	VEPYNATLS	SVRQL	VENTDETY	CIDNEALYD	ICFRITL	KLTTPTYG	DLNHLVSATMSGVITCL
CHICKEN $\beta 1$						V							
CHICKEN $\beta 3$		A				V							
CHICKEN $\beta 4$		C N			V	V		I			A		S
CHICKEN $\beta 5$		C H											S
HUMAN M40		A				V							
HUMAN $\beta 2$		A				V							
HUMAN $5\beta$		A			F	V							
HUMAN $\beta 4$		C N			V	V		I			A		S
MOUSE M $\beta 2$													[M $\beta 2$ :
MOUSE M $\beta 4$		A			F	V							
MOUSE M $\beta 4'$		A				V							
MOUSE M $\beta 5$		A				V							
PIG $\beta$						V							
		250	260	270	280	290	300	310	320	330	340	350	360
CHICKEN $\beta 1$		RFPGLNADLRKLA	VNMVFPRLHFF	MPGFAPLTS	RSGSQYRAL	TVPELTQ	QMFDSKN	MAACDPR	BGRYLTV	AAIFRGR	MSMKEVDE	QMLNVGN	KNSSYFVEV
CHICKEN $\beta 2$													
CHICKEN $\beta 3$													
CHICKEN $\beta 4$					R		A	E G	V		AI S		V
CHICKEN $\beta 5$					A		A		TV		AI		V
HUMAN M40													
HUMAN $\beta 2$						D	V A		V				
HUMAN $5\beta$							A		V				
HUMAN $\beta 4$				A		G	A		V		S S		
MOUSE M $\beta 2$													
MOUSE M $\beta 4$							V A		V		L		
MOUSE M $\beta 4'$							I A		V		S S		A
MOUSE M $\beta 5$							V NA		V				
PIG $\beta$					(A)				V				
		370	380	390	400	410	420	430	440				
CHICKEN $\beta 1$		LKMSATFIGNSTA	IGELFKRI	SEQFTAMF	RRKAFLHW	YTTGEGMD	EMEFTEA	ESNMNDL	VSEYQOYQ	QDATAD	EQGE*F	EEFG	EDEAX
CHICKEN $\beta 2$													
CHICKEN $\beta 3$													
CHICKEN $\beta 4$		S											
CHICKEN $\beta 5$		AS		S		F							
HUMAN M40		AV											
HUMAN $\beta 2$													
HUMAN $5\beta$		AV											
HUMAN $\beta 4$		S											
MOUSE M $\beta 2$													
MOUSE M $\beta 4$		A											
MOUSE M $\beta 4'$													
MOUSE M $\beta 5$		AV											
PIG $\beta$													
RAT RBT.1													
RAT RBT.2													
RAT RBT.3													

FIG. 1. Primary sequences of vertebrate  $\beta$  tubulins: identification of variable regions within a conserved framework. The polypeptide sequences for all known vertebrate  $\beta$  tubulins are shown in the one-letter amino acid code. The predicted polypeptide sequence of the chicken  $c\beta 2$  gene (20, 26) is shown on the top line. Sequences of other vertebrate  $\beta$  tubulin genes are shown in comparison, explicitly displaying only those residues that differ from the chicken  $c\beta 2$  sequence. The sequences of the chicken  $c\beta 3$  and  $c\beta 5$  polypeptides as well as the human  $h\beta 4$  and mouse  $m\beta 4'$  sequences and portions of the  $m\beta 2$  (residues 238–345) and  $m\beta 5$  (residues 1–324) sequences are from our sequencing efforts. The remaining mouse sequences were determined by Lewis *et al.* (11). Complete sequences for three human  $\beta$  tubulins were taken from the work of Cowan and coworkers (22, 30, 33, 34) and carboxyl-terminal sequences for three rat  $\beta$  tubulins were from Farmer *et al.* (25). Sequences for two porcine polypeptides are from Krauhs *et al.* (19), whereas fragmentary sequences for one bovine  $\beta$  tubulin are from Little and Luduena (35). The brackets delineate the positions at which sequence data begin or end. [Sequences have been corrected at residue 400 ( $m\beta 2$ ,  $m\beta 4$ , and  $m\beta 5$ ), residue 111 ( $m\beta 4$ ), and residue 433 ( $h5\beta$ ) (N. J. Cowan, personal communication).]

## CLASS I ISOTYPE - CONSTITUTIVE

```

human hβ1      E E E E D F G E E A E E E A
rat rbt.3      E E E E D F G E E A E E E A
mouse mβ5      E E E E D F G E E A E E E A

```

## CLASS II ISOTYPE - MAJOR NEURONAL

```

chick cβ1/cβ2  D E Q G E F E E E G E E D E A
pig           D E Q G E F E E E G E E D E A
rat rbt.1     D E Q G E F E E E G E E D E A
mouse mβ2     D E Q G E F E E E G E E D E A

```

## CLASS III ISOTYPE - MINOR NEURONAL POLYPEPTIDE

```

chick cβ4      E E E G E M Y E D D E E E S E - Q G A K
human hβ4      E E E G E M Y E D D E E E S E S Q G P K

```

## CLASS IV ISOTYPE

## variant IIIa - BRAIN SPECIFIC

```

rat rbt.2      E E G E F E E E A E E E V A
human 5β       E E G E F E E E A E E E V A
mouse mβ4      E E G E F E E E A E E E V A

```

## variant IIIb - CONSTITUTIVE

```

human hβ2      E E E G E F E E E A E E E V A
mouse mβ4'     E E E G E F E E E A E E E V A

```

## TENTATIVE DOMINANT TESTIS ISOTYPE

```

chick cβ3      E E E G E F E E E A E E E A E

```

## UNASSIGNED ISOTYPE - UBIQUITOUS, EXCEPT BRAIN

```

chick cβ5      N D G E E A F E D D E E E I N E

```

FIG. 2. Alignment of the carboxyl termini of known vertebrate  $\beta$  tubulins into distinct isotypic classes. Each of the 17 known carboxyl-terminal sequences of vertebrate  $\beta$  tubulins shown in Fig. 1 has been grouped according to sequence homology and, where known, to dominant pattern of gene expression. All but 2 sequences (those of chicken  $c\beta 3$  and  $c\beta 5$ ) fall into one of four classes of conserved carboxyl-terminal sequence. The human class I isotype sequence, here labeled  $h\beta 1$  for clarity, is taken from the human cDNA  $D\beta 1$  (22), which is derived from the human  $\beta$  tubulin gene  $M40$  (34).

pattern of expression of the human  $h\beta 4$  gene has not yet been determined.)

Expression of the class IV isotype is controlled in mammals by two genes, each of which encodes a highly homologous protein. Expression of the gene sequence designated IVa has been detected uniquely in brain tissue in mouse and rat. We have reported here the isolation of a cDNA from mouse 3T3 cells that encodes a protein with a carboxyl-terminal sequence homologous to that of the mouse class IVa (gene  $m\beta 4$ ) differing by only a single amino acid insertion at position 433. However, it differs from  $m\beta 4$  in nucleotide sequence within the coding region and is completely divergent in the 3' untranslated region and thus clearly derives from a different gene (data not shown). Similarly, two human genes,  $h\beta 2$  and  $h\beta 5$ , encode proteins with class IV carboxyl-terminal sequences. Although  $h\beta 2$  is expressed in numerous human cell lines, significant expression of  $h\beta 5$  has been detected only in cells of neural origin (30, 33). Thus, in mammals the class IV isotype is encoded by two genetic variants differing in expression properties: variant IVa (e.g.,  $h\beta 5$  and  $m\beta 4$ ) has been detected only in neural tissues, whereas variant IVb ( $h\beta 2$  and  $m\beta 4'$ ) is apparently constitu-

tively expressed. Curiously, neither class IVa nor IVb isotypes have been found in the chicken.

A potential additional isotype class is represented at present only by the chicken  $c\beta 3$  sequence. This gene is the predominant  $\beta$  tubulin expressed in chicken testis and shows only low levels of expression in other chicken cell types. Although it shares a high degree of homology with class IV isotypes, the  $\beta 3$  sequence terminates with an acidic residue and also differs by nonconservative substitutions at positions 15, 83, and 306. Although it may in fact represent the chicken class IVb gene, we have tentatively classified  $c\beta 3$  as an independent isotype based on these sequence features as well as its pattern of expression. Confirmation of this designation as an independent isotype awaits the characterization of  $\beta$  tubulin sequences expressed in mammalian testes.

The chicken  $c\beta 5$  gene potentially defines a sixth isotypic species, which appears most closely related to the class III polypeptide. Since a cognate sequence has not yet been detected in another organism and it lacks the striking specificity of expression of the chicken  $c\beta 3$  gene, we have not designated it as a specific isotypic class.

The overall degree of conservation of sequences within the

ISOTYPE CLASS	EXPRESSION	C-TERMINAL VARIABLE REGION	N-TERMINAL VARIABLE REGION
I	MAJOR CONSTITUTIVE	E E E E D * F G E E A E E E A	T G T Y H G D S D L Q D E R S N V Y Y N E A T G G
II	MAJOR NEURONAL	D E Q G E * F E E E G E E D E A	T G S Y H G D S D L Q L E R I N V Y Y N E A T G N
III	MINOR NEURONAL	E E E G E M Y E D D E E E S E S Q G P K	S G N Y V G D S D L Q L Q R I S V Y Y N E A S S N
IV A	NEURAL SPECIFIC	E E * G E * F E E E A F E E V A	T G T Y H G D S D L Q L E R I N V Y Y N E A T G G
IV B	CONSTITUTIVE	E E E G E * F E E E A E E E V A	T G T Y H G D S D L Q L E K I N V Y Y N E A T G G

FIG. 3. Proposed carboxyl-terminal and amino-terminal isotype-defining consensus sequences for five  $\beta$  tubulins. The sequences compiled in Fig. 1 have been classified according to variable region sequence homology. The carboxyl- and amino-terminal variable region sequences are shown. The carboxyl-terminal sequences shown are the consensus of individual sequences within each class, except for class III, for which the human variant sequence is shown. For each class the dominant pattern of gene expression is also noted.



different isotypic classes may be estimated from the data of Fig. 1. Divergence between two class II polypeptides, from pig and chicken, was reported previously to be 0.9%. The two available class I sequences show only 0.7% divergence, whereas class III peptides show 0.7% divergence within the 86% of the polypeptide available for comparison. A comparable degree of conservation is seen in the other isotypic classes, although the present data do not allow extensive comparison. Thus, although different isotypes of  $\beta$  tubulin show significant heterogeneity, amino acid sequences within each isotypic class of  $\beta$  tubulin have been very highly conserved during evolution.

## DISCUSSION

The recent accumulation of  $\beta$  tubulin sequence data has produced somewhat paradoxical conclusions drawn from different studies. Initial biochemical studies of tubulin as well as comparison of the first complete sequences for vertebrate  $\beta$  tubulins suggested that tubulin had been very highly conserved throughout metazoan evolution (19, 20). Subsequently, the demonstration that tubulins are encoded by multigene families in multicellular eukaryotes (ref. 15; reviewed in ref. 16) and the determination of DNA sequences that encode divergent polypeptides within and between species (11, 22, 23, 30) suggested that, in contrast to the initial inferences,  $\beta$  tubulin is relatively variable and has undergone significant divergence during evolution.

Our present comparative sequence analysis of 19 vertebrate  $\beta$  tubulin polypeptides has, at least in part, resolved this situation by revealing the existence of a family of isotypically distinct vertebrate  $\beta$  tubulins that differ in amino acid sequence within two discrete variable domains. Within each proposed isotypic class of  $\beta$  tubulin both variable region sequences and developmentally regulated programs of expression are conserved among several different vertebrate species. The salient features of these data have been summarized in Fig. 3, in which are presented consensus sequences derived for the carboxyl-terminal variable regions, the corresponding amino-terminal variable regions, and the major patterns of expression for each of the four isotypic classes identified.

These findings rekindle previous suggestions that microtubule differentiation may be achieved (in part) by variations in the tubulin subunits themselves. Although the conservation of variable region sequences demonstrated here does not, in itself, prove functional divergence among the different  $\beta$  tubulin isotypes, the stringent conservation of variable region sequences strongly suggests that each isotype-specific sequence has been positively selected during evolution. Such positive selection would imply a functional role for the structural differentiation of  $\beta$  tubulin isotypes. Direct evidence that  $\beta$  tubulin isotypes are biochemically distinguished within cells has been presented by Gard and Kirschner (37) and Edde *et al.* (38), who have documented the differentiation-dependent phosphorylation of an electrophoretically distinct isoform of  $\beta$  tubulin in cultured neuroblastoma cells. Clearly, the presence of cellular mechanisms that distinguish among variant tubulins mandates that such variants are in fact utilized for unique roles in cells.

Consideration of all of the presently available data makes it seem likely that, at least in some instances,  $\beta$  tubulin isotypes do play a role in modulating the molecular properties of microtubules. However, in view of the failure to detect such properties in *in vitro* assays (39) or in gene transfection experiments (40), we suggest that such functional distinctions are probably subtle and manifest only in conjunction with other specific microtubule components.

We thank Drs. Moses Chow and Dan Littman for generously providing the human cDNA library cloned in  $\lambda$  gt11. Similarly, we thank Dr. Lester Lau for providing the mouse fibroblast cDNA library. A big thank you is also extended to Mr. Floyd Benko and Ms. Paula Machlin for their enthusiastic and expert technical assistance in DNA cloning and sequencing. K.F.S. has been supported by a National Institutes of Health Postdoctoral Fellowship. This work has been supported by grants from the National Institutes of Health and March of Dimes to D.W.C., who is the recipient of a National Institutes of Health Research Career Development Award.

- Inoue, S. (1981) *J. Cell Biol.* **91**, 131s-147s.
- Gibbons, I. R. (1975) in *Molecules and Cell Movement*, eds. Inoue, S. & Stephens, R. E. (Raven, New York), pp. 207-232.
- Hayden, J. H. & Allen, R. D. (1984) *J. Cell Biol.* **99**, 1785-1793.
- Vale, R. D., Schnapp, B. J., Reese, T. S. & Sheetz, M. P. (1985) *Cell* **40**, 449-454.
- Joshi, H. C., Chu, D., Bauxbaum, R. E. & Heideman, S. R. (1985) *J. Cell Biol.* **101**, 696-705.
- Bloom, G. S., Luca, F. C. & Vallee, R. B. (1984) *J. Cell Biol.* **98**, 331-340.
- Huber, G. & Matus, A. (1984) *J. Neurosci.* **4**, 151-160.
- Parysek, L. M., Asnes, C. F. & Olmsted, J. B. (1984) *J. Cell Biol.* **99**, 1309-1315.
- Binder, L. I., Frankfurter, A. & Rebhun, L. I. (1985) *J. Cell Biol.* **101**, 1371-1378.
- Havercroft, J. C. & Cleveland, D. W. (1984) *J. Cell Biol.* **99**, 1927-1935.
- Lewis, S. A., Gwo-Shu Lee, M. & Cowan, N. J. (1985) *J. Cell Biol.* **101**, 852-861.
- Gundersen, G. G., Kalnoski, M. H. & Bulinski, J. C. (1984) *Cell* **38**, 779-789.
- L'Hernault, S. & Rosenbaum, J. L. (1985) *Biochemistry* **24**, 473-479.
- Cumming, R., Burgoyne, R. D. & Lytton, N. A. (1984) *J. Cell Biol.* **98**, 347-351.
- Cleveland, D. W., Lopata, M. A., MacDonald, R. J., Cowan, N. J., Rutter, W. J. & Kirschner, M. W. (1980) *Cell* **20**, 95-105.
- Cleveland, D. W. & Sullivan, K. F. (1985) *Annu. Rev. Biochem.* **54**, 331-365.
- Fulton, C. & Simpson, P. A. (1976) in *Cell Motility*, eds. Goldman, R., Pollard, T. & Rosenbaum, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), p. 987-1005.
- Ponstingl, H., Krauhs, E., Little, M. & Kempf, T. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2757-2761.
- Krauhs, E., Little, M., Kempf, T., Hofer-Warbinek, R., Ade, W. & Ponstingl, H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4156-4160.
- Valenzuela, P., Quiroga, M., Zaldivar, J., Rutter, W. J., Kirschner, M. W. & Cleveland, D. W. (1981) *Nature (London)* **289**, 650-655.
- Lemischka, I. R., Farmer, S. R., Racaniello, V. R. & Sharp, P. A. (1981) *J. Mol. Biol.* **150**, 101-120.
- Hall, J. L., Dudley, L., Dobner, P. R., Lewis, S. A. & Cowan, N. J. (1983) *Mol. Cell Biol.* **3**, 854-862.
- Sullivan, K. F. & Cleveland, D. W. (1984) *J. Cell Biol.* **99**, 1754-1760.
- Sullivan, K. F., Havercroft, J. C. & Cleveland, D. W. (1984) in *Molecular Biology of the Cytoskeleton*, eds. Borisy, G. G., Cleveland, D. W. & Murphy, D. B. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 321-332.
- Farmer, S. R., Bond, J. F., Robinson, G. S., Mbangkollo, D., Fenton, D. & Berkowitz, E. M. (1984) in *Molecular Biology of the Cytoskeleton*, eds. Borisy, G. G., Cleveland, D. W. & Murphy, D. B. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 333-342.
- Sullivan, K. F., Lau, J. T. Y. & Cleveland, D. W. (1985) *Mol. Cell Biol.* **5**, 2454-2465.
- Lopata, M. A., Havercroft, J. C., Chow, L. T. & Cleveland, D. W. (1983) *Cell* **32**, 713-724.
- Messing, J. & Vieira, J. (1982) *Gene* **19**, 269-276.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
- Lewis, S. A., Gilmartin, M. E., Hall, J. L. & Cowan, N. J. (1985) *J. Mol. Biol.* **182**, 11-20.
- Murphy, D. B. & Wallis, K. T. (1983) *Biol. Chem.* **258**, 7840-7845.
- Murphy, D. B. & Wallis, K. T. (1983) *Biol. Chem.* **258**, 8357-8364.
- Gwo-Shu Lee, M., Loomis, M. C. & Cowan, N. J. (1984) *Nucleic Acids Res.* **12**, 5823-5836.
- Gwo-Shu Lee, M., Lewis, S. A., Wilde, C. D. & Cowan, N. J. (1983) *Cell* **33**, 477-487.
- Little, M. & Luduena, R. F. (1985) *EMBO J.* **4**, 51-56.
- Luduena, R. F., Roach, M. C., Trcka, P. P., Little, M., Palanivelu, P., Binkley, P. & Prasad, V. (1982) *Biochemistry* **21**, 4787-4794.
- Gard, D. L. & Kirschner, M. W. (1985) *J. Cell Biol.* **100**, 764-774.
- Edde, B., Jeantet, C. & Gros, F. (1981) *Biochem. Biophys. Res. Commun.* **3**, 1035-1043.
- Swan, J. A. & Solomon, F. (1984) *J. Cell Biol.* **99**, 2108-2113.
- Bond, J. F., Fridovich-Keil, J. L., Pillus, L., Mulligan, R. C. & Solomon, F. (1985) *Cell* **44**, 461-468.