J Mol Evol (1983) 19:397–410

Journal of Molecular Evolution

Evolution of αq- and β-Tubulin Genes as Inferred by the Nucleotide Sequences of Sea Urchin cDNA Clones

Despina Alexandraki and Joan V. Ruderman

Department of Anatomy, Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, USA

Summary. Evolutionary studies on the tubulin multigene families were initiated by nucleotide sequence analysis of cDNA clones complementary to sea urchin (*Lytechinus pictus*) testis α- and β-tubulin mRNAS. Sequence comparisons of three partial β-tubulin cDNA clones (pβ1, pβ2, pβ3) demonstrated the existence of tubulin mRNA heterogeneity. pβ2 and pβ3 contain identical tubulin-coding regions and extremely similar 3′ untranslated sequences, including a polyadenylation signal (AAUAAA). However, pβ2 contains an additional region of 3′ untranslated sequence which includes a second polyadenylation signal. These two sequences may be allelic, representing products of alternative transcription termination or processing pathways. pβ1 and pβ2 (or pβ3) cDNAs almost certainly correspond to transcripts of distinct but evolutionarily related genes. Examination of the available coding portions showed that they differ only by a few silent nucleotide substitutions and the deletion/insertion of one codon; most of the differences are clustered within the last 15 3′-end codons. In contrast, their 3′ untranslated sequences are considerably divergent. Nucleotide alignment in this region was feasible by considering specific point and segmental mutations, mainly T ↔ C transitions and small deletions/insetions associated with small direct repeats.

The sea urchin α- and β-tubulin cDNA and corresponding protein sequences were compared with previously described tubulin cDNA and protein sequences from other organisms. Both α and β tubulins are very conserved proteins, evolving with a rate comparable to that of histones. Analysis of the nucleotide divergence of the coding cDNA regions showed that replacement sites have changed with a rate 20–175 times lower than that of the silent sites. Among the 177 codons compared between the sea urchin testis and chick brain β-tubulin cDNAs, there are 7 conservative amino acid replacements and the deletion/insertion of two codons. Most of these changes are clustered near the C-terminus. The 161-amino acid portion of chick brain, rat and porcine α-tubulin sequences differs by 3 conservative amino acid replacements from the corresponding sea urchin testis α-tubulin sequence. The compared interspecies 3′ untranslated sequences are very divergent.

Key words: Tubulin cDNA sequences – Silent substitutions – Protein evolution – 3′ untranslated sequences – Multigene families

Offprint requests to: D. Alexandraki

Introduction

The rules underlying the evolution of multigene families may substantially enhance our understanding of the evolution of different developmental pathways, of differentiation processes and of the diversification of species. Ancestral genes reduplicate at differential rates, some disappear and others remain in the genome either closely linked or scattered, and these sequences diverge to different degrees, presumably depending on specific functional constraints. Detailed analysis of the anatomy of a small number of multigene families has already provided many clues to the function of particular sequences, either specific to the family or of more general occurrence.

The α- and β-tubulin multigene families are as ancient as eukaryotic organisms. αβ-tubulin heterodimers are the basic structural components of the polymeric microtubules of the cytoskeleton and several organelles involved in a variety of processes concerned with cell shape and motility. Extensive studies on the morphology and biochemistry of microtubules and tubulins have partially revealed how microtubules function in the

cell by assembling, disassembling and interacting with a number of small molecules and macromolecules (reviewed in Roberts and Hyams 1979). The conservation of their roles has been emphasized by the demonstration that tubulins are highly conserved proteins. Tubulins of several taxonomically distant species share common antigenic properties (reviewed in Ludueña 1979), can copolymerize to form microtubules in vitro (reviewed in Kirschner 1978), and give identical or closely similar peptide maps (Little et al. 1981; Ludueña and Little 1981). An extreme conservation of primary amino acid sequence was first revealed by comparison of partially sequenced α and β tubulins from chick brain and sea urchin sperm outer doublets (Ludueña and Woodward 1973) and more recently by comparison of the predicted amino acid sequences of two brain α-tubulin cDNAs, one from chicken (Valenzuela et al. 1981) and the other from rat (Lemischka et al. 1981). The cDNA sequencing analyses, as well as DNA cross-hybridization studies of both α- and β-tubulin sequences from three distantly related species (Cleveland et al. 1980), suggested that the striking interspecies conservation of tubulins partly reflects a considerable nucleotide sequence preservation.

In spite of the high interspecies conservation of tubulins there is considerable evidence in several systems for the existence of multiple tissue- and organelle-specific microheterogeneous α- and β-tubulin subtypes (reviewed in Roberts and Hyams 1979). This heterogeneity cannot be solely attributed to postranslational modifications of unique α- and β-tubulin proteins (Lefebvre et al. 1980; Brunke et al. 1982b). Proteins sequence data have shown directly the existence of several closely related forms of porcine brain α (Ponstingl et al. 1981) and β (Krauhs et al. 1981) tubulins. Functionally distinct tubulins which are products of different transcripts are known in *Naegleria* (Fulton and Simpson 1976; Lai et al. 1979) and in *Drosophila* (Raff et al. 1982). Multiple α- and β-tubulin sequences have been detected in the genomes of a variety of organisms (Cleveland et al. 1980, 1981; Sánchez et al. 1980; Kalfayan and Wensink 1981; Silflow and Rosenbaum 1981; Lemischka et al. 1981; Alexandraki and Ruderman 1981; Cowan et al. 1981; Wilde et al. 1982; Mischke and Pardue 1982; Thomashow et al. 1983). Genetic evidence argues for the stage-specific expression of different β-tubulin genes in *Drosophila* (Kemphues et al. 1979; 1980) and different α-tubulin genes in *Aspergillus* (Morris et al. 1979). Therefore, tubulin gene families provide a clear example of ubiquitous, highly conserved and yet diversified genes that are expressed with developmental specificity.

As a first step in a study of the evolution, diversification and differential expression of tubulin genes we are pursuing a detailed analysis of cloned sea urchin messenger and genomic sequences. Ten to fifteen sequences for each tubulin family are present in the sea urchin genome (Alexandraki and Ruderman 1981) at least some of which must encode organelle-specific tubulin subtypes (Stephens 1978). In this report, nucleotide sequence analysis of tubulin cDNA clones confirms our previous (indirect) evidence that heterogeneous mRNA sequences are present in the sea urchin testis cytoplasm. The same cDNA sequences are also used to make an evolutionary comparison of three sea urchin β-tubulin cDNAs and of these sea urchin sequences with a previously described chick brain β-tubulin cDNA (Valenzuela et al. 1981). Finally, the sequence of a sea urchin α-tubulin cDNA is compared with the brain α-tubulin cDNA sequences of chicken (Valenzuela et al. 1981) and rat (Lemischka et al. 1981; Ginzburg et al. 1981). The amino acid sequences of porcine brain α (Ponstingl et al. 1981) and β (Krauhs et al. 1981) tubulins are also included in the amino acid sequence comparisons. Since the sea urchin cDNA molecules were constructed from testis mRNA, whereas the chick, rat and pig sequences were derived from brain, the observed differences may represent either species- or tissue-specific variations, or both.

Materials and Methods

Recombinant plasmid DNAs were amplified in *E. coli* HB101 (Boyer and Rouland-Dussoix 1969) and purified as described by Clewell (1972). DNA end-labeling and sequencing reactions were carried out following the protocols described by Maxam and Gilbert (1980). 5'-end labeling of DNA was accomplished with [γ-^{32}P]ATP (specific activity > 3,000 Ci/mmol; ICN Radiochemicals) and T4 polynucleotide kinase (New England Biolabs) following dephosphorylation with bacterial alkaline phosphatase (Bethesda Research Laboratories Inc.).

To obtain uniquely end-labeled DNA fragments, plasmid DNA was digested with one restriction endonuclease, end-labeled and then digested with one or more additional restriction endonucleases. Alternatively, plasmid cDNA insertion sequences were isolated, cleaved with restriction endonucleases which generated only one 5'-protruding-end sequence (the other one being the 3'-protruding Pst I-cleavage site) and labeled.

Labeled and unlabeled DNA fragments were purified following electrophoresis in Tris borate EDTA polyacrylamide gels (6–8% polyacrylamide with acrylamide to bisacrylamide ratio of 30 or 20) and elution as described by Maniatis and Efstratiadis (1980). In some cases, DNA fragments labeled at both ends were sequenced following strand separation and purification by electrophoresis in polyacrylamide gels according to Maxam and Gilbert (1980). The four sequencing reaction products (G, G+A, T+C, C) were analyzed by electrophoresis in 7–12% polyacrylamide gels (40 cm long, ~0.3 mm thick) followed by autoradiography on Kodak XR5 X-ray film with DuPont Cronex Lightening Plus intensifying screens at -80°C.

Results and Discussion

Isolation of α- and β-Tubulin cDNA Clones

Recombinant pBR322 plasmid clones containing cDNA sequences complementary to sea urchin (*Lytechinus pictus*) testis cytoplasmic poly(A)$^+$ RNA were con-

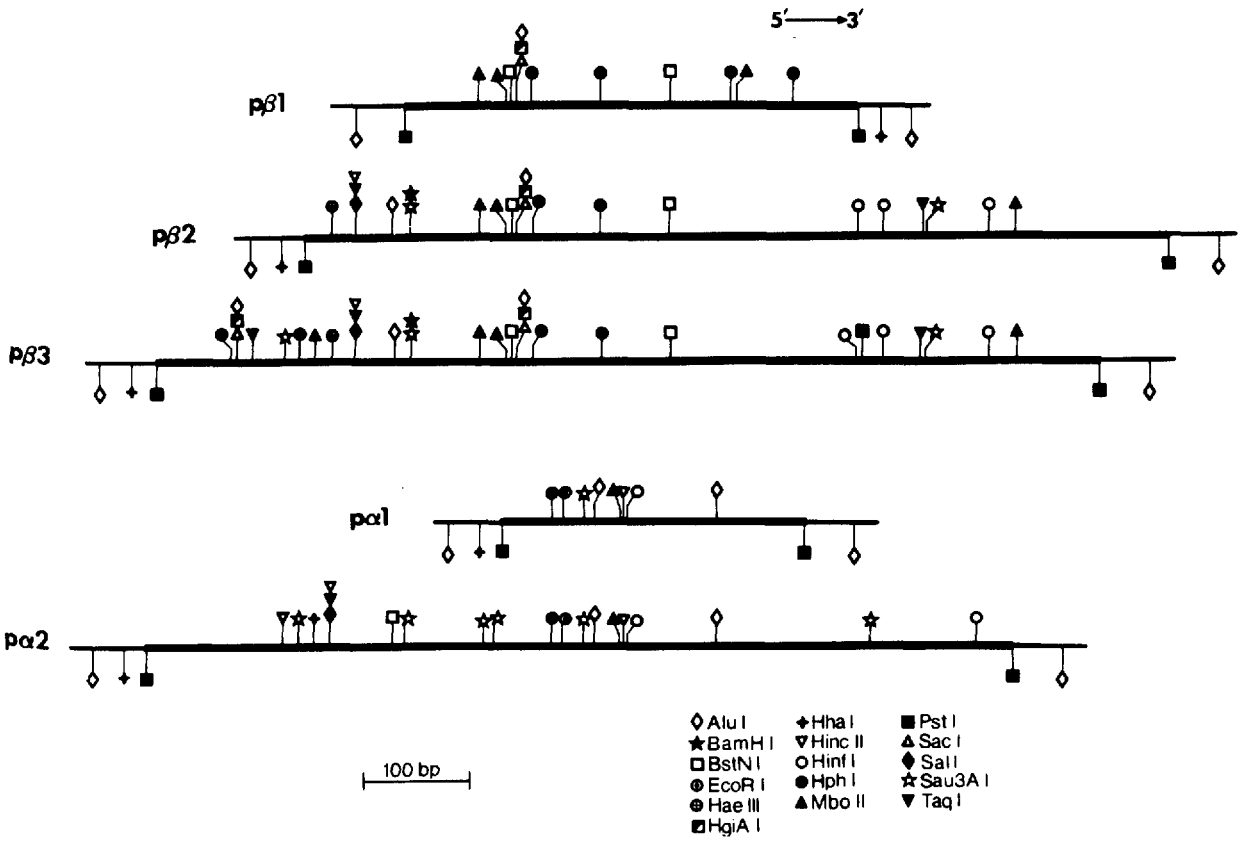

Fig. 1. Restriction endonuclease maps of tubulin cDNA clones.
The overlapping and different restriction endonuclease sites of various α- and β-tubulin cDNA clones are shown as symbols (according to given keys) on aligned insertion sequences (represented by thick lines). These restriction sites were confirmed by DNA sequencing. (An additional Hph I-recognition sequence, 250 nucleotides off the 5′ end on pα2 insertion, was predicted by the sequencing analysis, however, pα2 is not digested by Hph I at this position.)

structed as previously described (Alexandraki and Ruderman 1981). The originally identified clones pα1 and pβ1 contained partial α-tubulin and β-tubulin cDNA sequences, respectively. We have since isolated several additional tubulin cDNA clones using as probes the pα1 and pβ1 insertion DNA fragments as well as DNA from λ-*Lytechinus pictus* genomic clones containing tubulin sequences (Alexandraki and Ruderman 1981). Restriction endonuclease mapping of the longest cDNA insertions showed that most insertions contained sites that overlapped with pα1 or pβ1. However, the restriction maps of two clones (pβ2 and pβ3) suggested the existence of both common and variant (relative to pβ1) sequence regions (Fig. 1). We have chosen these clones for detailed sequence analysis in order to examine β-tubulin gene heterogeneity, without prejudging whether that heterogeneity represented allelic or non-allelic variation (the cDNA clones were constructed using RNA from testes of several animals). DNA sequence analysis of pα1 and of the longer pα2 (see Fig. 1) was also performed.

Nucleotide Sequences of α- and β-Tubulin cDNA Clones

The primary sequences of pβ1, pβ2, pβ3, pα1 and pα2 insertion DNAs were determined as described in the Methods. Figure 2 shows the strategy of sequencing overlapping restriction endonuclease fragments of each clone. Each DNA fragment was sequenced twice for confirmation of the sequences. The correct codon frames were identified by comparison to the previously described chick brain α- and β-tubulin cDNA sequences (Valenzuela et al. 1981).

The sense strand of each cDNA is shown in the 5′ to 3′ orientation (Fig. 3). The sea urchin partial cDNA sequences are presented in comparison with the corresponding portions of tubulin cDNA sequences from other organisms. pβ1, pβ2 and pβ3 sequences contain respectively 300, 392 and 532 nucleotides which code for 100, 131 and 177 C-terminal amino acids of β tubulin and, following a termination codon, 83, 390, and 323 nucleotides of the 3′ untranslated mRNA region. The pα1 and pα2 sequences contain respectively 143 and

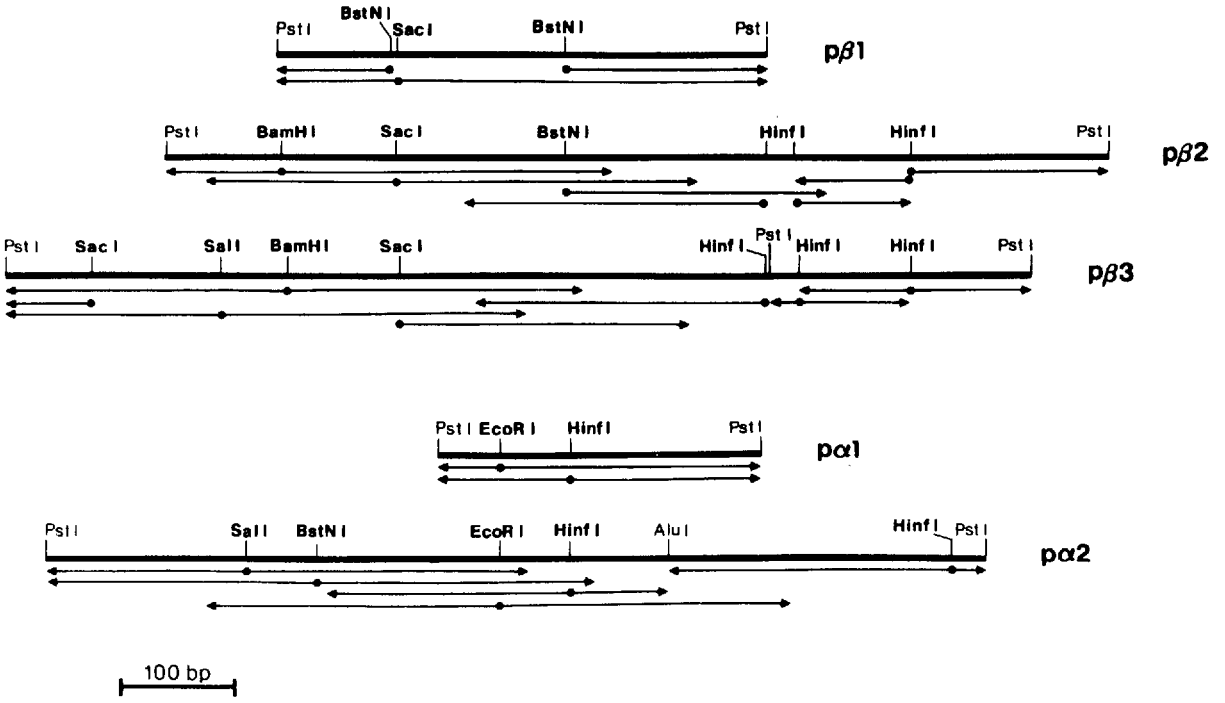

Fig. 2. Diagram illustrating the sequencing strategy for the cloned α- and β-tubulin cDNAs.

Various (mainly overlapping) DNA-restriction endonuclease fragments were end labeled, purified and sequenced as described in the Methods. In this diagram, the dot at the beginning of each arrow and the length of the arrow indicate the restriction site used for 5′ end labeling and the extent of determined sequence, respectively

483 bases coding for the 47 and 161 C-terminal amino acids of α tubulin and 104 and 297 3′ untranslated bases after the termination codon (Fig. 4). The sequence of the pα1 clone is identical to a portion of the pα2 sequence (as indicated in Fig. 1), so it is not presented separately.

The Highly Conserved β-Tubulin Protein and cDNA Coding Sequences Contain Localized Differences

Comparison of the nucleotide sequences of the three sea urchin β-tubulin partial cDNA clones revealed a remarkably high degree of conservation. The pβ2 and pβ3 sequences are identical throughout their common 392-base protein-coding portion. The overlapping coding-sequence regions of pβ1 and pβ3 (or pβ2) insertions (300 bases) differ only by 5 silent nucleotide substitutions and the deletion/insertion of one codon (Fig. 3 and Table 1). It is worth noting that three of these base substitutions and the codon deletion/insertion are clustered within a region of 10 codons near the C-terminus.

The sea urchin pβ3 (longest) and the chicken β-tubulin coding sequences exhibit an 84.7% nucleotide identity over a contiguous region of 535 nucleotides. Their differences include 82 base substitutions, 74 of which are silent. The remaining eight base substitutions result in seven amino acid replacements. Just as with the sea urchin sequences, these differences are highly clustered, two near the 5′ end of the available sequences and the other six near the protein C-terminus. In the latter location three codon deletions/insertions are also present, two of which represent tandemly duplicated codons in the sea urchin sequences and the third one is a codon found only in the pβ1 and the chicken sequences (Fig. 3 and Table 1). The comparable portion of the porcine brain β tubulin is more similar to the chick brain sequence, differing only by two amino acid residues from that sequence, whereas it differs from the sea urchin pβ3 predicted tubulin chain by seven residues (Table 1).

The high amino acid conservation between the sea urchin and chicken sequences (96%) probably reflects functional constraints on the protein structure. An analysis of the nucleotide divergence in the coding regions, considering the degeneracy of the genetic code and correcting for multiple base change events, showed that although the number of silent changes exceeds the number of the available silent sites, the replacement sites have changed at a 20–30 fold lower rate (Table 2).

Most of the amino acid differences in the tubulins of the three species are rather conservative when the ionic properties (Hood et al. 1975) and the conforma-

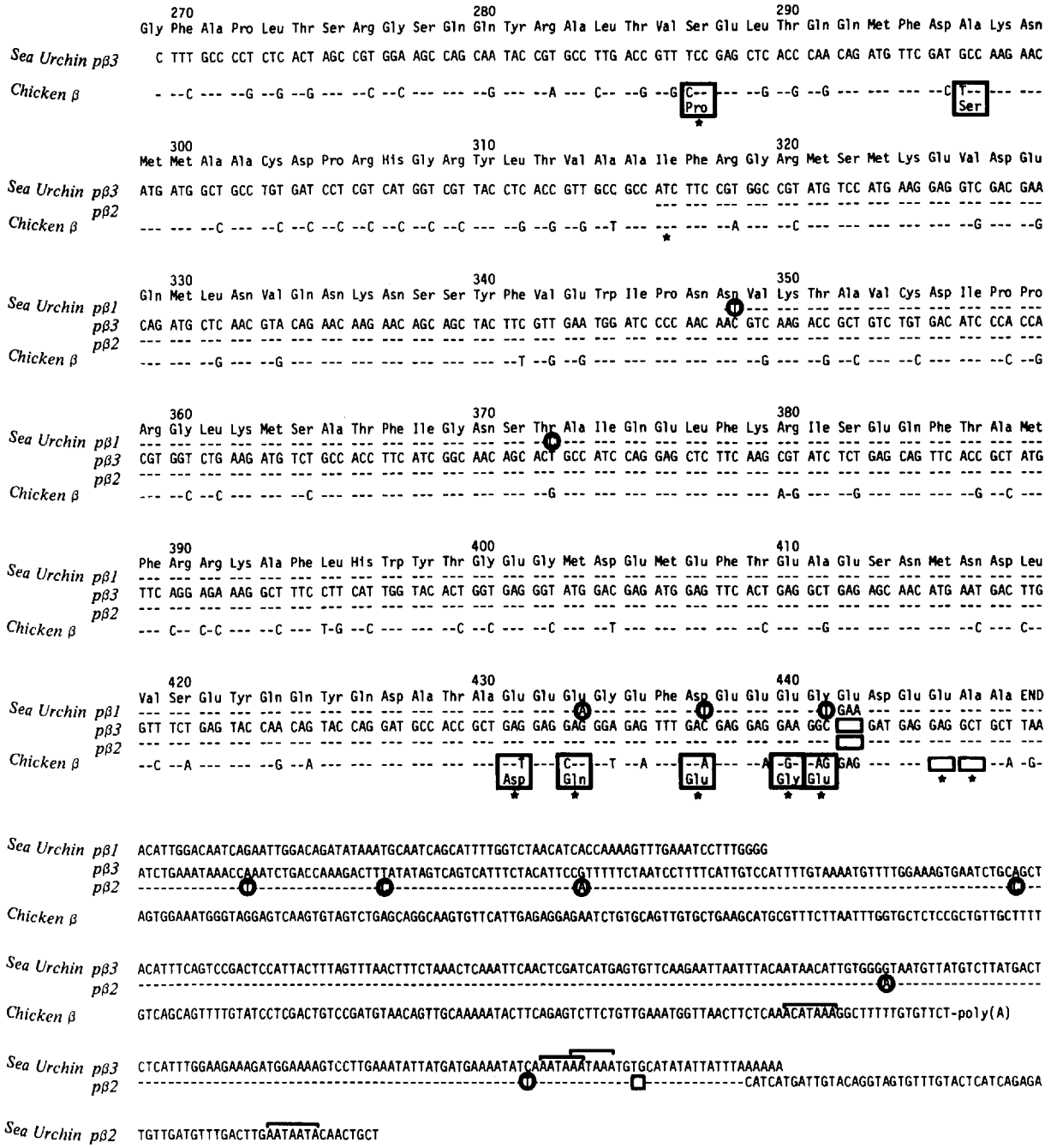

Fig. 3. Nucleotide and predicted amino acid sequence comparisons of different sea urchin testis β-tubulin partial cDNAs and a previously described chick brain β-tubulin cDNA.

Comparisons were performed by using the pβ3 sequence as reference, and only the different nucleotides or amino acids of the other sequences are shown. The entire 3′ untranslated sequences of pβ1 and chicken β are shown due to their high divergence from pβ3 (or pβ2). The amino acids were numbered according to a known complete primary sequence of porcine brain β tubulin. Amino acid replacements are boxed. Insertions/deletions are shown by open rectangles. Asterisks indicate the amino acids which are different in the porcine β-tubulin sequence. Nucleotide substitutions among the sea urchin sequences are encircled. Overlined sequences indicate the putative polyadenylation signals

tional parameters (Chou and Fasman 1978) of the amino acids are considered. The two rather radical replacements, both in terms of charge and conformational coefficient, of glutamic acid with glycine at position 440 and of glycine with glutamic acid at position 441, could be viewed as a simple interchange having a neutral overall effect. Also, the deletion/insertion of the glutamic acid at positon 442 and the presence of glutamic

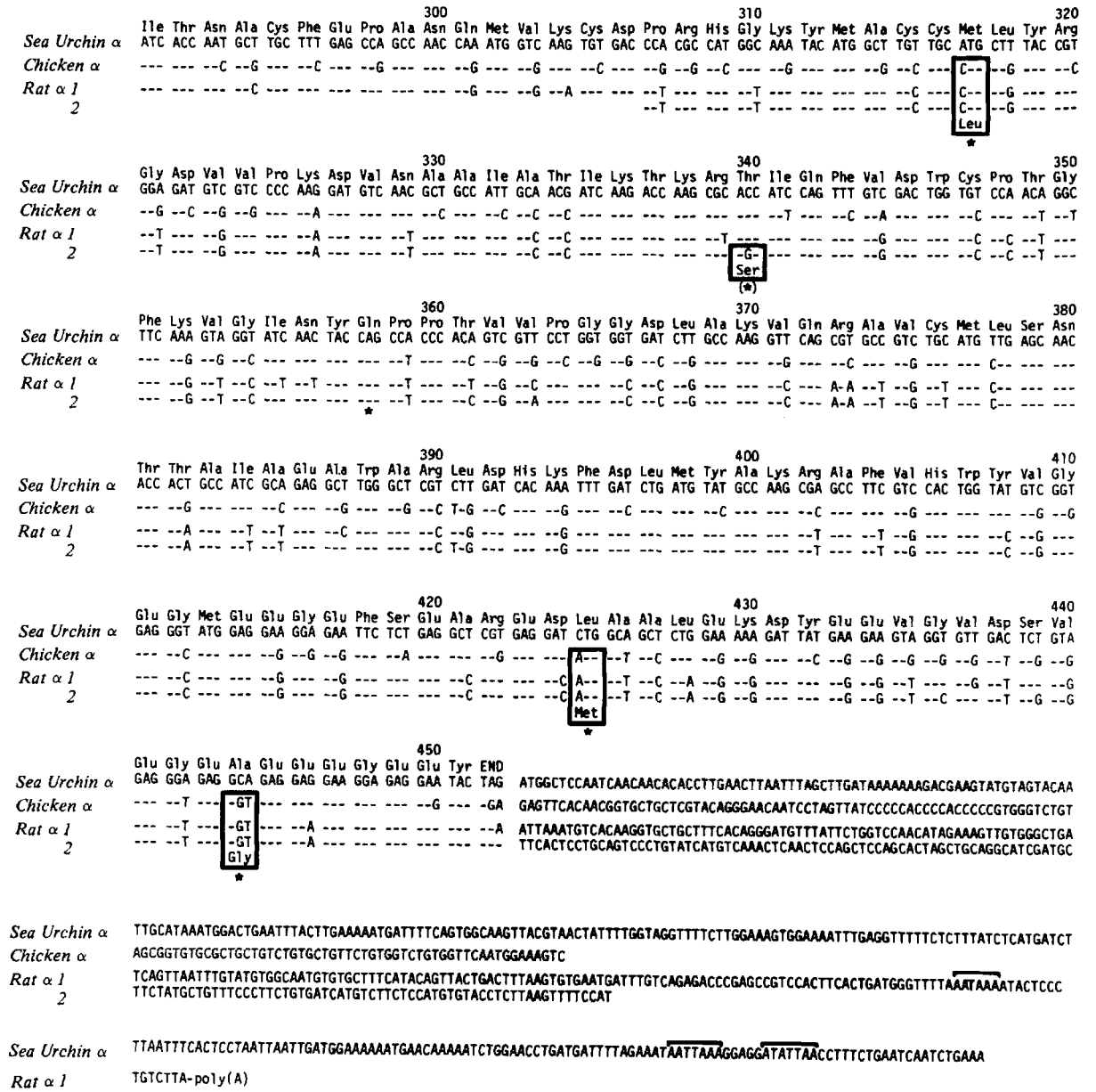

Fig. 4. Nucleotide and predicted amino acid sequence comparisons of a sea urchin testis α-tubulin partial cDNA (pα2) with known chick and rat brain α-tubulin cDNAs.

The sea urchin sequence was used as reference. As rat 1 is indicated the sequence presented by Lemischka et al. (1981), and as rat 2 the sequence presented by Ginzburg et al. (1981). The amino acids were numbered according to a published complete primary sequence of porcine brain α tubulin. All other conventions are as in Fig. 3. The asterisk at position 340 is in parentheses because the porcine sequence is heterogeneous at this position, containing either threonine or serine

acid and alanine at positions 445 and 446 in some of the sequences may not disturb the properties of this highly acidic, glutamic acid-rich region.

All of these interspecies amino acid replacements, except two at positions 287 and 316, lie within two of the regions (288–302 and 426–445) of the molecule that were predicted to be highly helical by analyses of the porcine β tubulin (Krauhs et al. 1981). Most of them are clustered in the main helical portion at the C-terminus, which is highly acidic and is thought to assume an α-helical or randomly-coiled configuration depending on the pH conditions (Lu and Elzinga 1978). This portion of β tubulin is probably exposed on the exterior of the molecule, since it was found to be phosphorylated in vivo in the rat brain (Eipper 1974). Also, this is possibly a region for the binding of Ca^{++} or of cationic microtubule-associated proteins (Ludueña 1979). It is likely that different functional constraints and therefore different flexibilities govern the evolution of specific parts of the molecule. Sites involved

Table 1. Amino acid replacements in tubulin sequences

Position	Sea urchin			Chicken	Pig	Position	Sea urchin	Chicken	Rat		Pig
	pβ1	pβ2	pβ3	β	β		pα2	α	α1	α2	α
287	←- - - - - - - -	Ser	- - -→	Pro	Pro	317	Met	Leu - - - - - - - - - - - - →			
296	←- - - - - - - -	Ala	- - -→	Ser	Ala	340	Thr - - - - - - - - - - - - - - →			Ser	Thr, Ser
316	←- - - - - - - -	Ile	- - -→	Ile	Val	358	Gln - - - - - - - - - - - - - - - - - →				Glu
431	←- - - - - - - -	Glu	- - -→	Asp	Asp	425	Leu	Met - - - - - - - - - - - - →			
433	←- - - - - - - -	Glu	- - -→	Gln	Gln	444	Ala	Gly - - - - - - - - - - - - →			
437	←- - - - - - - -	Asp	- - -→	Glu	Glu						
440	←- - - - - - - -	Glu	- - -→	Gly	Gly						
441	←- - - - - - - -	Gly	- - -→	Glu	Glu						
442	Glu (GAA)	–	–	Glu (GAG)	Glu						
444,445	←- Glu-Glu (GAG-GAG) - →			Glu (GAG)	Glu						
446,447	←- Ala-Ala (GCT-GCT) - →			Ala (GCA)	Ala						

Amino acid replacements and codon deletions/insertions of the sequences presented in Figs. 3 and 4 are tabulated

Table 2. Multiple-hit corrected divergences of α- and β-tubulin coding cDNA sequences

β-Tubulin cDNA sequences	Replacement sites (%)	Silent sites (%)
sea urchin pβ2 – chicken	4.25	138.8
sea urchin pβ1 – chicken	6.88	128.4
pβ1 – pβ2	0	11.3
α-Tubulin cDNA sequences		
sea urchin pα2 – chicken	1.84	186.7
sea urchin pα2 – rat 1	1.85	118.7
2	2.33	119.1
chicken – rat 1	0.55	96.7
2	0.78	78.8
rat 1 – rat 2	0.34	8.9

The percent divergences of silent and replacement substitution sites were calculated according to the method presented in Perler et al. (1980), considering the number of potential silent or replacement sites and including corrections for multiple base change events at each site. Codons corresponding to gaps were not included in the calculation

in the assembly process, in αβ heterodimer formation or in other universally required interactions with other molecules may be highly conserved. Regions that are post-translationally modified or interact with distinct molecules in different organelles and tissues may be allowed to diverge more freely in order to acquire these specific functions. Alternatively, the acidic region of β tubulin may be the only region of the molecule in which selectively neutral changes are possible.

The conservation of β tubulin sequences is extremely high, at least within the compared portion of the molecule. The uniform rate of protein sequence evolution is conventionally expressed as the unit evolutionary period (UEP), the average time required for a 1% amino acid divergence to occur in the same protein of two evolutionary lineages (Wilson et al. 1977). An estimate of the UEP of β tubulin suggests a rate of evolution comparable to that of histones H4 and H3 (Table 3). The UEP values obtained by considering the three examined species are not uniform. This finding is not in agreement with the evolutionary clock hypothesis which states that the amino acid sequence of a protein changes at a constant rate in evolution (reviewed in Wilson et al. 1977). However, it is quite possible that the compared sequences are not orthologous (descendants of the same gene), and therefore not eligible for such a comparison, since the corresponding mRNAs derive from different tissues (testis and brain) that are enriched in different microtubular organelles (flagellar and cytoplasmic microtubules). Nevertheless, the order of magnitude of the estimated UEP values is informative about the overall rate of β tubulin evolution. It should be added that although this analysis deals with only a portion of the tubulin sequence (177 C-terminal residues), comparison of the partial N-terminal sequences in the same two species (Ludueña and Woodward 1973) resulted in similar UEP values (Table 3).

Table 3. Rates of α and β tubulin divergence

		No. of amino acid replacements/total no. of amino acids	Time of divergence[c] in million years	UEP[d] in million years
β Tubulin[a]				
sea urchin – chicken		7/177	700	177
		1/ 25[b]		175
sea urchin – pig		7/177	700	177
chicken – pig		2/177	300	265
α Tubulin				
sea urchin – chicken		3/161	700	376
		0/ 25[b]		
sea urchin – rat	1	3/161	700	376
	2	4/145		254
sea urchin – pig		4 or 5/161	700	282 or 225
chicken – rat	1	0/161	300	
	2	1/145		435
chicken – pig		1 or 2/161	300	483 or 241
pig – rat	1	2/161	60	48
	2	1/145		87
Histones				
H4				400
H3				330
H2A				60
H2B				60
H1				8
Insulin				14
Insulin C peptide				1.9
Globin α				3.7
β				3.3
Fibrinopeptide	A			1.7
	B			1.1

a Rates of divergence were established using the sequences presented in Figs. 3 and 4

b Data taken from the partial N-terminal protein sequences published by Ludueña and Woodward (1973)

c The times of lineage divergence are taken from Lewin (1980; Table 11.2)

d The UEP values were corrected for multiple amino acid replacements at the same position by using the Table 29, in Dayhoff (1976). (The correction is insignificant for the very small number of amino acid replacements in tubulins). The UEP values given for other proteins shown below the tubulins are taken from Wilson (1977; Table 1)

Conservation of the α-Tubulin Protein and cDNA Coding Sequences

As shown in Fig. 4 the sea urchin α-tubulin cDNA sequence differs in 86 out of a total of 483 nucleotides from the chicken cDNA sequence (82.2% homology), in 72 out of 483 nucleotides (85.1% homology) from the rat 1 sequence and in 68 out of 438 nucleotides (84.5%) from the rat 2 sequence. Three base substitutions resulted in three amino acid replacements in the chicken sequence which have also been retained in both rat sequences (positions 317, 425, 444; see also Table 1). These three amino acid changes have also been conserved in the pig α-tubulin sequence. An additional base substitution caused the replacement of threonine with serine at position 340 in the rat 2 sequence. This is one of the heterogeneous positions found in the porcine protein sequence and contains either one of those two amino acids. The porcine sequence differs from all the other sequences at position 358.

Similar to the case of β tubulin, the α-tubulin protein sequence appears selectively conserved in spite of the high mutational pressure observed in the divergence of the silent nucleotide sites (Table 2). The replacement sites were estimated to have changed 50–175 times slower than the silent sites (Table 2).

Two of the observed amino acid replacements in the compared protein sequences, those at positions 425 and 444, are located within two α-helical regions of the C-terminus (413–435 and 440–450), according to the prediction in the analysis of the porcine α tubulin (Ponstingl et al. 1981). This is a very acidic region and is probably exposed on the exterior of the molecule (Ponstingl et al. 1979), being subject to post-translational modifications. One such modification

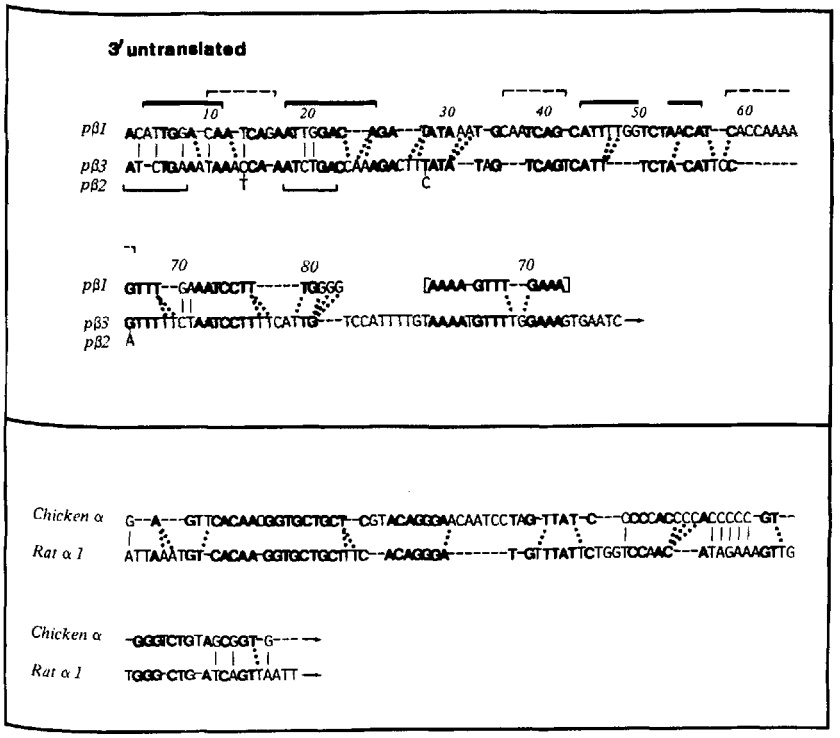

Fig. 5. Alignment of 3′ untranslated sequences.
The entire available pβ1 3′ untranslated portion has been aligned with the initial part of the corresponding sequence of pβ3 (or pβ2). Alignment was also performed in the initial 3′ untranslated portions of the chicken and rat 1 α-tubulin cDNAs. The aligned homologous bases are shown in bold face and base substitutions are indicated by vertical lines. Base duplications are shown by dotted lines and deletions by dashes. Thick, thin, or dashed horizontal lines indicate some gross sequence repetition

is tyrosylation (Raybin and Flavin 1977; Thompson et al. 1979). All of the α-tubulin cDNAs sequenced so far (including the sea urchin sequences) contain a tyrosine codon immediately before the termination codon. However, Ponstingl et al. (1981) found that only 15% of the porcine α-tubulin protein preparation contained this C-terminal tyrosine. Apparently, this tyrosine residue can be removed and possibly added post-translationally.

Both amino acid replacements found in the C-terminus of α tubulin (positions 425 and 444) are very conservative. They do not cause any charge change and the alanine in the sea urchin molecule has an even higher α-helical coefficient than the glycine present in the other species. The replacement of methionine with leucine at position 317 as well as the alternative presence of threonine and serine at position 340 are very conservative. On the other hand, the presence of glutamine at position 358 in the sea urchin, chicken and rat sequences reinforces the regional conformation of β-sheet predicted in the porcine sequence (Ponstingl et al. 1981) although the change of glutamic acid to glutamine is not a conservative one.

The comparison of the available α-tubulin sequences within the region of the C-terminal 161 residues revealed a high sequence conservation, just as in the case of β tubulins. An estimation of the rate of α tubulin divergence gave UEP values comparable to those of histones or even higher (Table 3). Similarly, comparison of the entire chicken and rat α-tubulin predicted amino acid sequences by Lemischka et al. (1981) also showed extreme sequence conservation. The observed variation in our estimated UEP values may be due to the comparison of sequences which do not derive from the same gene, as discussed above for β tubulins. A more accurate rate of tubulin divergence requires the comparison of orthologous full-length sequences since restricted parts of the molecules may be under different functional constraints as indicated, for instance, by the clustering of changes in β tubulins.

Divergence of the 3′ Untranslated Regions in Both α- and β-Tubulin cDNAs

Comparisons of the 3′ noncoding sequences of the β- or α-tubulin cDNAs, shown in Fig. 3 and 4, revealed a considerably lower degree of conservation than that observed in the respective coding sequences. The sea urchin pβ3 clone probably contains the entire 3′ untranslated portion of the corresponding mRNA, since it ends with a stretch of six adenines which probably represent a portion of the 3′ poly(A) tail. In addition, 26 bases upstream of the 3′ end there is a sequence AAUAAAUAAA which includes the hexanucleotide AAUAAA. This latter sequence is commonly observed at this position in eukaryotic mRNAs and considered to be involved in the polyadenylation of transcripts (Proudfoot and Brownlee 1976; Brethnach and Chambon 1981). The pβ2 3′ untranslated sequence is very similar to the pβ3 sequence for 316 bases immediately following the termination codon; in this portion it differs from the pβ3 sequence by six base substitutions and one dinucleotide deletion. By contrast, the pβ2

sequence lacks the adenine stretch present at the pβ3 3′ end and continues for an additional 71 bases. Since no poly(A) tail is observed at the 3′ end of pβ2, the untranslated sequence of the corresponding mRNA must extend even further. However, the pβ2 3′ sequence may be almost complete since a variation of the typical hexanucleotide AAUAAA, the sequence AAUAAUA (Setzer et al. 1980; Tosi et al. 1981), precedes the last eight bases.

The pβ2 and pβ3 cDNAs are copies of two different mRNA molecules. These two mRNAs could be transcribed form two different genes. Alternatively, since the pβ2 sequence appears to contain the pβ3 sequence (including its polyadenylation signal) the two mRNAs may be transcribed from a single gene which either has two different transcription termination sites or has one termination site and the long mRNA is processed in either of two alternative ways. In the second case, the few base substitutions by which the two sequences differ would be attributed to allelic polymorphism. Overlapping cellular mRNAs differing in the length of their 3′ untranslated region occur in other systems as well. For example, two functionally distinct mouse μ immunoglobulin heavy chain mRNAs are produced from a single gene by alternative processing pathways (Early et al. 1980). Four overlapping mouse dihydrofolate reductase mRNAs appear to be multiple products of a single gene (Setzer et al. 1980). Two overlapping α-amylase mRNAs are transcribed from a single gene in both mouse liver and salivary gland tissues (Tosi et al. 1981). Multiple mRNAs appear to be produced from a chicken vimentin single copy gene (Zehner and Paterson 1983). Analysis of the corresponding genomic sequences is necessary to resolve this issue for the β-tubulin transcripts.

The pβ1 insertion contains only a portion (83 bases) of the 3′ untranslated mRNA region. Comparison of the untranslated sequence in pβ1 with that in pβ2 (or pβ3) shows that these sequences are related but extensively divergent, differing by a number of segmental mutations as well as base substitutions. pβ1 and pβ2 sequences are different enough to postulate that they are products of two different genes. An optimal matching of these sequences is shown in Fig. 5. The alignment was performed manually, according to the following set of accepted rules. Blocks of ≥4 consecutive identical bases in both sequences were first identified, and then the regions between those blocks were aligned by introducing not more than one gap for each 2-bp matching (Jones and Kafatos 1980). The segmental mutations assumed in this analysis frequently are duplications/multiplications of a base; 11 out of 16 gaps can be accounted for this way. In addition, two deletions/insertions of longer sequences and three one-base gaps were found. Among base substitutions, transitions and especially T↔C substitutions predominate. Six out of nine base substitutions are transitions; five of these are T↔C. This holds also for the base substitutions observed in the coding regions of pβ1 and pβ3 (or pβ2) (all five substitutions are transitions, four of which are (T↔C), and in the noncoding regions of pβ3 and pβ2 (four out of six substitutions are transitions, two of which are T↔C).

While alignment of these sequences without the positioning of any gaps showed a number of regions of considerable matching (two to seven contiguous bases), it also showed long stretches of unmatched bases. Evidence from other investigations suggests that the introduction of gaps may not be arbitrary if associated with regions of small direct repeats. In the case of pβ1 and pβ3, both the multiplication of single bases and the repetition of the CA dinucleotide within positions 35–40 and 55–62 can be considered gaps related to small direct repeats. It is worth noting that the single segmental mutation found in the coding region of the pβ sequences is associated with a repeated codon (GAA) at positions 440 and 442. Similarly, the only deleted TG dinucleotide from the pβ2 3′ untranslated sequence is repeated twice at that position in the pβ3 sequence. Also, the interspecies deletion/insertion of two codons at position 445, 446 is related with repeated codons (444, 447) (see Fig. 3). Deletions between or at short repeated sequences have been demonstrated in many cases in *E. coli* (see, for example, Farabough et al. 1978), in phages (Pribnow et al. 1981) and in the human β-globin gene family (Efstratiadis et al. 1980). A more recent analysis of a large number of silkmoth chorion sequences provides many examples of segmental mutations associated with direct small tandem or nontandem repeats (Jones and Kafatos 1982). Repeats in general and tandem repeats in particular have been implicated as hot-spots for unequal crossing-over events. A mechanism that has been proposed as being responsible for the observed deletions/insertions between or at small repeated sequences relies on the "slipped mispairing" and excision of a looped out DNA strand during replication or repair (Farabough et al. 1978; Efstratiadis et al. 1980; Albertini et al. 1982; Jones and Kafatos 1982).

Comparison of the interspecies α- or β-tubulin 3′ untranslated sequences also revealed an extreme divergence of this region compared to the coding portion of the cDNAs. An effort to align these sequences gave very ambiguous results. Homologies of small base stretches were found at only a very few sites, and a considerable number of deletions/insertions had to be assumed. Significant homology was only found between the initial portions of the rat 1 and chicken α-tubulin noncoding sequences (Fig. 5). The similarities observed between the other interspecies sequences may be artifactual. In general, the noncoding regions have a biased base composition, usually AT enriched (McReynolds et al. 1978; Nunberg et al. 1980). Such composition tends to form simple sequences which could be mistaken for real

homologies. In the case of the tubulin untranslated sequences biases also exist: pβ1-64% AT; pβ2 and pβ3-71% AT; chick β-60% TG; pα2-71.5% AT; chick α-28% each of T, C, G; rat 1 α-62% AT; rat 2 α-66% AT. The existence of simple sequences may also artifactually generate the repetition of small DNA segments which can be seen in all of the examined tubulin non-coding sequences. However, some "real" gross repetition of DNA blocks was detected, as indicated in the pβ1 and pβ3 sequences in Fig. 5. Whether repetition has any functional significance or whether it simply represents a random evolutionary process is not known. Theoretical studies based on characteristics of sequences such as the 5S spacer DNA and satellite DNA suggest that simple repetition can be generated in any randomly mutated sequence which is not selectively conserved, by unequal crossing-over events. The repeats remain identical or they diverge depending on the mutational rate relative to the frequency of crossovers (Smith 1976).

Further Discussion

The cDNA and protein sequences for α and β tubulins of several organisms have established directly that tubulin amino acid sequences are highly conserved, even between species as distantly related as echinoderms and mammals. Our analysis also shows that this conservation is reflected by the high preservation of the replacement sites relative to the significant divergence of the silent sites at the nucleotide level. However, the available sea urchin β-tubulin cDNAs may represent an example of selective conservation of the protein coding nucleotide sequences among different gene copies within the same species.

Comparison of the pβ1 and pβ3 (or pβ2) sequences showed that, although their 3′ untranslated regions are very divergent, the coding regions are conserved far beyond what would be required for conservation of the amino acid sequence. The corresponding divergences are 50% (or 81%, when corrected for multiple hits[1]) in the 3′untranslated region and 1.7% (or 11.3% corrected - Table 2) in the coding region. Most importantly, the majority of the differences in the coding sequences appear clustered at the 3′ terminus, heighboring the 3′ noncoding sequences. Ths leaves a region of 252 bases (84 codons) that is virtually identical, differing only by one base substitution. Apparently, some mechanism has maintained almost identical certain parts of the coding region in the two corresponding β-tubulin genes. The alternative explanation that the two different untranslated sequences (along with an ~45-base portion of coding region) derive from alternative splicing of the same coding sequence is made less likely by the following results. Copy number measurements indicated that while the coding sequences are repeated 10–15 times in the sea urchin genome, the noncoding sequences of pβ1 and pβ2 are each repeated 2–5 times (Alexandraki and Ruderman, in preparation). It has been suggested that species-specific homogeneity of the multiple members of eukaryotic gene families is maintained both by stochastic and directional processes involving unequal exchange, gene conversion and transposition (Dover and Coen 1981; Baltimore 1981; Selker et al. 1981; Dover 1982; Roberts and Axel 1982). The homogeneity of tandemly repeated genes can be generated by unequal crossing-over (Smith 1973; Petes 1980; Szostak and Wu 1980). Such gene copies, as the highly repeated early histone gene clusters in the sea urchins (Hentschel and Birnstiel 1981) and the distinct clusters of chorion genes in the silkmoths (Jones and Kafatos 1981), contain almost identical or closely homologous coding sequences and noncoding sequences which evolve in parallel. β-tubulin genes are mostly dispersed in the sea urchin genome and, although some clustering also exists (Alexandraki and Ruderman 1981), the corresponding genes of the two cDNAs were not found closely linked in total genomic DNA or cloned genomic DNA fragments (unpublished observations). A process that can generate homogeneity of genes which are dispersed in the same or in different chromosomes is gene conversion. This correcting mechanism is probably responsible for the presence of a virtually identical region within two closely linked $^{G}\gamma$- and $^{A}\gamma$-globin genes adjacent to sequences that differ by 10–20% (Slightom et al. 1980; Efstratiadis et al. 1980). It may also explain the concerted evolution of other nonallelic α-globin genes (Liebhaber et al. 1981) and families of dispersed genes in yeasts (Jackson and Fink 1981; Klein and Petes 1981). The case of pβ1 and pβ2 (or pβ3) sequences resembles that of genes corrected by gene conversion. However, analysis of the corresponding genomic sequences is necessary before we can substantiate any conclusions.

If indeed the coding sequences of dispersed β-tubulin genes are preferentially highly conserved, that would indicate that the differences in the primary structure of the various β-tubulin subtypes may be very subtle. (For example, pβ1 and pβ2 differ by one glutamic acid residue in a highly acidic, glutamic acid-rich region). In contrast, each gene could be easily identified by its surrounding divergent noncoding sequences. This situation would be similar to that of the actin gene families in the sea urchin and other systems (Cleveland et al. 1980; McKeown and Firtel 1981a, b; Fyrberg et al. 1981; Scheller et al. 1981; Davidson et al. 1982). These findings have led to the hypothesis, as stated by

[1] The corrected divergence of the untranslated region was found 82.3% when the sequences were compared directly without any prior special alignment and 79.5% when the alignment shown in Fig. 5 was first performed and each nonhomologous base pair as well as each nucleotide gap were scored as mismatches

Davidson et al. (1982), that the primary sequence of morphogenetic (structural) proteins encoded by small multigene families may not be the determining factor for the regulated expression of each gene and the assignment of each different protein to specific structures. Instead, member genes may have been duplicated and inserted into different regions of the genome, possibly into specific developmental regulatory units, in order to be coordinately expressed with other genes as part of a particular tissue-specific program of gene expression. This hypothesis is consistent with the finding of a testis-specific β-tubulin subtype in *Drosophila melanogaster*, which is involved in multiple structures in late spermatogenesis but is absent from similar structures in early spermatogenesis and oogenesis (Kemphues et al. 1979, 1980, 1982). Investigation of the genomic β-tubulin sequences along with their noncoding and flanking regions will be necessary to reveal the evolution and function of this gene family in the sea urchins.

Acknowledgments. We are grateful to F. C. Kafatos for invaluable advice and his criticisms of the manuscript. We thank P. Kennedy, N. Rosenthal and G. Thireos for advice on the DNA sequencing procedures and C. Fuller, S. Tabor and J. Pustell for help with the computer analysis of the sequencing data. This work was supported by NIH grant HD 15351 to J.V.R.

References

Albertini AM, Hofer M, Calos MP, Miller JH (1982) On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. Cell 29:319–328

Alexandraki D, Ruderman JV (1981) Sequence heterogeneity, multiplicity, and genomic organization of α- and β-tubulin genes in sea urchins. Mol Cell Biol 1:1125–1137

Baltimore D (1981) Gene conversion: some implications for immunoglobin genes. Cell 24: 592–594

Boyer HW, Rouland-Dussoix D (1969) A complementation analysis of the restriction and modification of DNA in *E. coli*. J Mol Biol 41:459–472

Breathnach R, Chambon P (1981) Organization and expression of eucaryotic split genes ccding for proteins. Ann Rev Biochem 50:349–383

Brunke KJ, Young EE, Buchbinder BU, Weeks DP (1982a) Coordinate regulation of the four tubulin genes of *Chlamydomonas reinhardii*. Nucleic Acids Res 10:1295–1310

Brunke KJ, Collis PS, Weeks DP (1982b) Post-translational modification of tubulin dependent on organelle assembly. Nature 297:516–518

Chou PY, Fasman GD (1978) Empirical predictions of protein conformation. Ann Rev Biochem 47:251–276

Cleveland DW, Hughes SH, Stubblefield E, Kirschner MW, Varmus HE (1981) Multiple α and β tubulin genes represent unlinked, dispersed gene families. J Biol Chem 256:3130–3134

Cleveland DW, Lopata MA, MacDonald RJ, Cowan NJ, Rutter WJ, Kirshner MW (1980) Number and evolutionary conservation of α- and β-tubulin, cytoplasmic β- and α-actin genes using specific cloned probes. Cell 20:95–105

Clewell DB (1972) Nature of ColE1 plasmid replication of *Escherichia coli* in the presence of chloramphenicol. J Bacteriol 110:667–676

Cowan NJ, Wilde CD, Chow LT, Wefald FC (1981) Structural variation among human β-tubulin genes. Proc Natl Acad Sci USA 78:4877–4881

Davidson EH, Thomas TL, Scheller RH, Britten RJ (1982) The sea urchin actin genes and a speculation on the evolutionary significance of small gene families. In: Dover GA, Flavell RB (eds) Genome Evolution. Systematics Assoc. special vol 20. Academic Press, New York, pp 177–191

Dayhoff MO (ed) (1976) Atlas of protein sequence and structure. Vol 5, Suppl 2, Silver Spring, Natl Biomed Res Found

Dover G (1982) Molecular drive and molecular drift: a novel mode of species evolution. Nature 299:111–117

Dover GA, Coen ES (1981) Springcleaning ribosomal DNA: a model for multigene evolution? Nature (London) 290: 731–732

Early P, Rogers J, Davis M, Calame K, Bond M, Wall R, Hood L (1980) Two mRNAs can be produced from a single immunoglobin gene by alternative RNA processing pathways. Cell 20:313–319

Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA, DeRiel JK, Forget BG, Weissman SM, Slightom JL, Blechl AE, Smithies O, Baralle FE, Shoulders CC, Proudfoot NJ (1980) The structure and evolution of the human β-globin gene family. Cell 21:653–668

Eipper BA (1974) Properties of rat brain tubulin. J Biol Chem 249:1407–1416

Farabaugh PJ. Schmeissner U, Hofer M, Miller JH (1978) Genetic studies of the lac repressor. VII. On the molecular nature of spontaneous hotspots on the *lacI* gene of *Escherichia coli*. J Mol Biol 126: 847–857

Fulton C, Simpson PA (1976) Selective synthesis and utilization of flagellar tubulin. The multi-tubulin hypothesis. In: Goldman R, Pollard T, Rosenbaum JL (eds) Cell Motility, Book C. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 987–1005

Fyrberg EA, Bond BJ, Hershey ND, Mixter KS, Davidson N (1981) The actin genes of *Drosophila:* protein coding regions are highly conserved but intron positions are not. Cell 24: 107–116

Ginzburg I, Behar L, Givol D, Littauer UZ (1981) The nucleotide sequence of rat α-tubulin: 3'-end characteristics and evolutionary conservation. Nucleic Acid Res 9: 2691–2697

Hentschel CC, Birnstiel ML (1981) The organization and expression of histone gene families. Cell 25:301–313

Hood LE, Wilson JH, Wood WB (1975) Molecular biology of eucaryotic cells, Vol 1. Benjamin WA, Inc, Menlo Park, California

Jackson JA, Fink GR (1981) Gene conversion between duplicated genetic elements in yeast. Nature 292:306–311

Jones CW, Kafatos FC (1982) Accepted mutations in a gene family: evolutionary diversification of duplicated DNA. J Mol Evol 19:87–103

Jones CW, Kafatos FC (1981) Linkage and evolutionary diversification of developmentally regulated multigene families: tandem arrays of the 401/18 chorion gene pair in a silkmoth. Mol Cell Biol 1:814–828

Jones CW, Kafatos FC (1980) Structure, organization and evolution of developmentally regulated chorion genes in a silkmoth. Cell 22:855–867

Kalfayan L, Wensink PC (1981) α-tubulin genes of *Drosophila*. Cell 24: 97–106

Kemphues KJ, Kaufman TC, Raff RA, Raff EC (1982) The testis-specific β-tubulin subunit in *Drosophila melanogaster* has multiple functions in spermatogenesis. Cell 31:655–670

Kemphues KJ, Raff EC, Raff RA, Kaufman TC (1980) Mutation in a testis-specific β-tubulin in *Drosophila*: analysis of its effects on meiosis and map location of the gene. Cell 21: 445–451

Kemphues KJ, Raff RA, Kaufman TC, Raff EC (1979) Mutation in a structural gene for a β tubulin specific to testis in *Drosophila melanogaster.* Proc Natl Acad Sci USA 76:3991–3995

Kirschner MW (1978) Microtubule assembly and nucleation. In: Bourne GH, Danielli GF (eds) Int Rev of Cytology 54. Academic Press: New York, pp 1–71

Klein HL, Petes T (1981) Intrachromosomal gene conversion in yeast: a new type of genetic exchange. Nature 289: 144–148

Krauhs E, Little M, Kempf T, Hofer-Warbinek R, Ade W, Ponstingl H (1981) Complete amino acid sequence of β-tubulin from porcine brain. Proc Natl Acad Sci USA 78:4156–4160

Lai EY, Walsh C, Wardell D, Fulton C (1979) Programmed appearance of translatable flagellar tubulin mRNA during cell differentiation in *Naegleria.* Cell 17: 867–878

Lefebvre PA, Silflow CD, Wieben CD, Rosenbaum JL (1980) Increased levels of mRNAs for tubulin and other flagellar proteins after amputation or shortening of *Chlamydomonas* flagella. Cell 20:469–477

Lemischka IR, Farmer S, Racaniello VR, Sharp PA (1981) Nucleotide sequence and evolution of a mammalian α-tubulin messenger RNA. J Mol Biol 151:101–120

Lewin B (1980) Gene Expression, Vol 2 Eucaryotic Chromosomes, 2nd Ed, T Wiley and Sons, New York

Liebhaber SA, Goossens M, Kan YW (1981) Homology and concerted evolution at the α1 and α2 loci of human α-globin. Nature 290:26–29

Little M, Ludueña RF, Langford GM, Asnes CF, Farrell K (1981) Comparison of proteolytic cleavage patterns of α-tubulins and β-tubulins from taxonomically distant species. J Mol Biol 149:95–107

Lu RC, Elzinga M (1978) The primary structure of tubulin. Sequences of the carboxyl terminus and seven other cyanogen bromide peptides from the α-chain. Biochem Bioph Acta 537:320–328

Ludueña RF (1979) Biochemistry of tubulin. In: Roberts K, Hyams JS (eds) Microtubules. Academic Press, London, pp 65–116

Ludueña RF, Little M (1981) Comparative structure and chemistry of tubulins from different eukaryotes. Biosystems 14:231–238

Ludueña FR, Woodward DO (1973) Isolation and partial characterization of α- and β-tubulin from outer doublets of sea urchin sperm and microtubules of chick embryo brain. Proc Natl Acad Sci USA 70:3594–3598

Maniatis T, Efstratiadis A (1980) Fractionation of low molecular weight DNA or RNA in polyacrylamide gels containing 98% formamide and 7 M urea. Methods Enzymol 65:299–305

Maxam A, Gilbert W (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol 65: 499–559

McKeown M, Firtel RA (1981a) Differential expression and 5'end mapping of actin genes in *Dictyostelium.* Cell 24: 799–807

McKeown M, Firtel RA (1981b) Evidence for sub-families of actin genes in *Dictyostelium* as determined by comparisons of 3'end sequences. J Mol Biol 151:593–606

McReynolds L, O'Malley BW, Nisbet AD, Fothergill JE, Girol D, Fields S, Robertson M, Brownlee GG (1978) Sequence of ovalbumin mRNA. Nature 273:723–728

Mischke D, Pardue ML (1982) Organization and expression of α-tubulin genes in *Drosophila melanogaster.* One member of the α-tubulin multigene family is transcribed in both oogenesis and later embryonic development. J Mol Biol 156:449–466

Morris NR, Lai MH, Oakley CE (1979) Identification of a gene for α-tubulin in *Aspergillus nidulans.* Cell 16:437–442

Nunberg JH, Kaufman RJ, Chang ACY, Cohen SN, Schimke RT (1980) Structure and genomic organization of the mouse dihydrofolate reductase gene . Cell 19:355–364

Perler F, Efstratiadis A, Lomedico P, Gilbert W, Kolodner R, Dodgson J (1980) The evolution of genes: The chicken preproinsulin gene. Cell 20:555–566

Petes TD (1980) Unequal meiotic recombination within tandem arrays of yeast ribosomal DNA genes. Cell 19:765–774

Ponstingl H, Little M, Krauhs E, Kempf T (1979) Carboxy-terminal amino acid sequence of α-tubulin from porcine brain. Nature 282:423–424

Ponstingl H, Krauhs E. Little M, Kempf T (1981) Complete amino acid sequence of α-tubulin from porcine brain. Proc Natl Acad Sci USA 78:2757–2761

Pribnow D, Sigurdson DC, Gold L, Singer BS, Napoli C, Borsius J, Dull TJ, Noller HF (1981) rII cistrons of bacteriophage T4 DNA sequence around the intercistronic divide and positions of genetic landmarks. J Mol Biol 149:337–376

Proudfoot NJ, Brownlee GG (1976) 3'non-coding region sequences in eukaryotic messenger RNA. Nature 263:211–214

Raff EC, Fuler MT, Kaufman TC, Kemphues KJ, Rudolph JE, Raff RA (1982) Regulation of tubulin gene expression during embryogenesis in *Drosophila melanogaster.* Cell 28: 33–40

Raybin D, Flavin M (1977) Enzyme which specifically adds tyrosine to the α chain of tubulin. Biochemistry 16:2189–2194

Roberts JM, Axel R (1982) Gene amplification and gene correction in somatic cells. Cell 29:109–119

Roberts K, Hyams JS (eds) (1979) Microtubules. Academic Press, London

Sanchez F, Natzle JE, Cleveland DW, Kirschner MW, McCarthy BJ (1980) A dispersed multigene family encoding tubulin in *Drosophila melanogaster.* Cell 22:845–854

Scheller RH, McAllister LB, Crain Jr WR, Durica DS, Posakony JW, Britten RJ, Davidson EH (1981) Organization and expression of multiple actin genes in the sea urchin. Mol Cell Biol 1:609–628

Selker EU, Yanofsky C, Driftmier K, Metzenberg RL, Alzner-DeWeerd B, RajBhandary UL (1981) Dispersed 5S RNA genes in *N. crassa:* Structure, expression, and evolution. Cell 24:819–828

Setzer DR, McGrogan M, Nunberg JH, Schmike RT (1980) Size heterogeneity in the 3'end of dihydrofolate reductase messenger RNAs in mouse cells. Cell 22:361–370

Silflow CD, Rosenbaum JL (1981) Multiple α- and β-tubulin genes in *Chlamydomonas* and regulation of tubulin mRNA levels after deflagellation. Cell 24:81–88

Slightom JL, Blechl AE, Smithies O (1980) Human fetal $^{G}\gamma$- and $^{A}\gamma$-globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell 21:627–638

Smith GP (1973) Unequal crossover and the evolution of multigene families. Cold Spring Harbor Symp Quant Biol 38: 507–513

Smith GP (1976) Evolution of repeated DNA sequences by unequal crossover. Science 191:528–535

Stephens RE (1978) Primary structural differences among tubulin subunits from flagella, cilia, and the cytoplasm. Biochemistry 17:2882–2891

Szostak JW, Wu R (1980) Unequal crossing over in the ribosomal DNA of *Saccharomyces cerevisiae.* Nature 284: 426–430

Thomashow LS, Milhausen M, Rutter WJ, Agabian N (1983) Tubulin genes are tandemly linked and clustered in the genome of *Trypanosoma brucei.* Cell 32:35–43

Thompson WC, Deanin GG, Gordon MW (1979) Intact microtubules are required for rapid turnover of carboxyl-terminal

tyrosine of α-tubulin in cell cultures. Proc Natl Acad Sci USA 76:1318–1322

Tosi M, Young RA, Hagenbuchle O, Schibler U (1981) Multiple polyadenylation sites in a mouse α-amylase gene. Nucleic Acids Res 9:2313–2323

Valenzuela P, Quiroga M, Zaldivar J, Rutter WJ, Kirschner MW, Cleveland DW (1981) Nucleotide and corresponding amino acid structure encoded by α- and β-tubulin mRNA. Nature 289:650–655

Wilde CD, Crowther CE, Cowan NJ (1982) Isolation of a multigene family containing human α-tubulin sequences. J Mol Biol 155:533–538

Wilson AC, Carlson SS, White TJ (1977) Biochemical evolution. Ann Rev Biochem 46:573–639

Zehner ZE, Paterson BM (1983) Characterization of the chicken vimentin gene: single copy gene producing multiple mRNAs. Proc Natl Acad Sci USA 80:911–915

Received March 21, 1983. Revised July 18, 1983