Molecular Evolution of a Group of Microtubule-Associated Proteins
Sharing Partial Similarities in Their Primary Structures

Yurika Hashi, Susumu Kotani, and Kiyotaka Tokuraku

Abstract: Several microtubule-associated proteins (MAPs), including mammalian MAP2, MAP4, and τ, share a repeated motif called the assembly-promoting (AP) sequence, and the evolutionary relationship among the MAPs is claiming increased attention. In this report, we first searched the protein database for AP sequence-bearing proteins using the MAP4 repeat sequences as a query, and found some 600 proteins from 170 different species. We assessed the authenticity of the detected sequences by manually aligning them along the consensus sequence, v-sk-gs-nikh-ppgg. Interestingly, the AP sequence-bearing proteins were found only in Bilateria, despite the importance of microtubules for all eukaryotes. The differentiation of the AP sequence-dependent microtubule regulation likely followed the emergence of Bilateria in biological evolution. Secondly, we closely inspected the resemblance of the AP sequences in the MAPs from three mammals and one nematode. By carefully comparing the AP sequences of the four MAPs, we discovered that the nematode MAP, PTL-1, is most similar to MAP4 among the three, although it was originally identified as a τ-like protein. From an evolutionary point of view, the closer relationship of PTL-1 to the ubiquitous MAP4 is more reasonable than to a protein restricted to mammalian brains.

Keywords: microtubule, microtubule-associated proteins (MAPs), Assembly-Promoting (AP) sequence

Introduction

Microtubules are major cytoskeletal components that play important roles in multiple cellular events, such as cell division, intra-cellular transport, and cell movement. These microtubule-mediated functions are regulated by a variety of accessory proteins, referred to as microtubule-associated proteins (MAPs), which include motor proteins, severing factors, depolymerizing factors, stabilizing factors, and others.

The mammalian MAPs, MAP1, MAP2, MAP4, and τ, are stabilizing factors: they bind to the surface of microtubules, stimulate microtubule assembly, and stabilize preformed microtubules. Among them, MAP2, MAP4, and τ share structural similarities, with distinct N-terminal projection and C-terminal microtubule-binding domains. The microtubule-binding domain is further divided into three subdomains, the Pro-rich region, the Repeat region, and the Tail region (Fig. 1), based on the characteristics of their primary structures. We and others have studied the functions of the three subdomains and concluded that their coordination accomplishes the overall MAP activity.

The Repeat region has attracted much attention, since the region contains a trio, quartet or quintet of 18-amino-acid-residue repeated motifs, with sequences that are more than 70% identical among MAP2, MAP4, and τ. Synthetic polypeptides with the motif sequence exhibited microtubule-binding activity, and we named the motif the assembly-promoting (AP) sequence.
promoting (AP) sequence\(^{15}\). Since the Repeat region is essential for a MAP to bind properly to microtubules, we hypothesized that MAP2, MAP4, and \(\tau\) share a common microtubule-binding mechanism mediated by the AP sequences. We are presently performing comparative biochemical analyses, using the microtubule-binding domains of the three mammalian MAPs (Hashi et al., manuscript in preparation).

Meanwhile, various organisms besides those in Mammalia reportedly express proteins containing repeat sequences that are very similar to those of the mammalian AP sequences\(^{16}\). Since the AP sequence itself is capable of binding to microtubules\(^{13-15}\), it is highly probable that those AP-like sequence-bearing proteins function as MAPs in their own organisms, and interact with microtubules via a mechanism common to that of the mammalian MAPs. Considering the biological significance of microtubules and their regulatory systems in eukaryotic cells, it is reasonable to expect that a certain primitive MAP molecule emerged at an early stage of eukaryote evolution. Currently, an AP sequence-bearing protein is the most promising candidate for the ancestral MAP, and the present-day AP sequence-bearing proteins may be its descendants. To investigate this idea, we are now analyzing an AP sequence-bearing nematode MAP, PTL-1, and examining its evolutionary relationship with the mammalian MAPs by comparing their biochemical characteristics (Hashi et al., manuscript in preparation).

In this study, we tackled the evolutionary problem by taking another approach, i.e., bioinformatics. We performed a thorough investigation of the sequences homologous to the AP sequence, using the protein database, and discussed their relationship from a phylogenetic point of view. We also made detailed inter-molecular and intra-molecular comparisons among the primary structures of the AP sequences in MAP2, MAP4, \(\tau\), and PTL-1.

Materials and Methods
All of the protein sequences were retrieved from Uniprot (http://www.uniprot.org) and were analyzed by the BLAST search program\(^{17}\), using the repeat sequence of Bos taurus MAP4 (amino acid residues 881-1007, GenBank: D90149.1) as the query sequence. We also compared the sequences manually, to judge the authenticity of the BLAST search results. The multiple alignment was performed using ClustalW, version 2.1\(^{18}\).

Results and Discussion
Tubulin, the principal constituent of microtubules, is a highly conserved protein, and its MAP binding site is reportedly localized to a restricted region of the C-terminal \(\alpha\)-helices\(^{19}\). Consequently, the binding site on the MAP side should also share some structural similarity, and the AP sequence-like structure is the most promising candidate for the similarity, as described in the introduction. Some 600 proteins from 170 different species were found to contain sequences homologous to the Repeat region of MAP4, by the BLAST search. To exclude pseudo-homologues from the BLAST results, we manually checked the detected sequences to confirm that they contain the consensus sequence of the MAP2/MAP4/\(\tau\) AP sequences (v-sk-gs-nikh-pggg). Among the 170 candidates, 165 biological species were verified to contain AP sequence-bearing proteins, while the remaining five did not. Incidentally, all five belonged to fungi. AP sequence-bearing proteins were found to be present in a variety of species, as

\[\text{Projection domain} \quad \text{Microtubule-binding domain} \]

MAP4

MAP2

\(\tau\)

200 a.a

Fig. 1  Schematic primary structures of MAP4, MAP2, and \(\tau\). AP sequences are indicated by black boxes in the Repeat region.
Interestingly, the expression of AP sequence-bearing proteins was restricted to Bilateria (Fig. 3). Although microtubules are essential in all eukaryotic cells, it is highly probable that the emergence of Bilateria was accompanied by new microtubule-dependent cellular events that required the AP sequence-based regulation. Poriferans are unigerminal, and cnidarians and ctenophores are diploblastic, whereas bilaterians are triploblastic. Consequently, bilaterians are expected to be equipped with a highly sophisticated mechanism of division plane determination in mitosis. Since the plane is under the control of microtubule-based mitotic apparatuses, new accessory proteins may have emerged to regulate the quite complicated microtubule dynamics at this stage of evolution. In addition, bilaterians are known to have developed a central nervous system consisting of neuronal cells, in which microtubules are involved in a variety of essential events, including neurite formation and axonal transport. It is also possible that new roles, which required novel regulatory factors, were assigned to microtubules in neural cells.

The Repeat regions of mammalian MAP2, MAP4, and \(\tau\) contain three to five AP sequences, which share very similar, but not identical primary structures. We reported elsewhere the detailed comparison of the AP sequences in the three MAPs, and discussed their evolutionary relationship. In this study, we expanded the analysis to include the PTL-1 AP sequences. Our previous results indicated that the primary structure of each AP sequence is more similar to the corresponding AP sequence in other MAPs, rather than to the AP sequences within the same molecule. Therefore, we first aligned the entire Repeat regions of the four MAPs, starting from the first AP sequence...
sequences, and compared them (Fig. 4A). We found that the first AP sequences were very similar: the four proteins have 7 amino acid residues in common. Nevertheless, the similarity was considerably lower in the second (2 identical residues) and third (4 identical residues) AP sequences. We then modified our approach, and compared each AP sequence to every AP sequence in the other MAPs. This approach revealed that the first AP sequence of MAP4 was similar to the first AP sequences of τ and PTL-1, but not to that of MAP2. The three AP sequences of MAP2 were similar to the third to fifth AP sequences of MAP4, in this order (Fig. 4B). The primary structures of the second to fourth AP sequences of τ, and the second to fourth AP sequences of PTL-1, also resembled the same region of MAP4. The second AP sequence of MAP4 is quite unique; no counterparts were found in MAP2, τ, and PTL-1. MAP4 may have acquired the sequence in the course of its evolution as a ubiquitous MAP in Mammalia. The fifth AP sequence of PTL-1 is also unique. Since its sequence is most similar to the fourth AP sequence in the same molecule, it is possible that the AP sequence was specifically duplicated during the evolution of the nematode MAP. The similarities in the primary structures of the AP sequences of the four MAPs are much more pronounced, as shown in the new alignment based on our findings (Fig. 4C). The inter-repeat sequences of the four MAPs also showed greater similarities in the new alignment (data not shown).

The nematode MAP, PTL-1, was originally identified as a τ-like protein, and hence it was named Protein with τ-like repeats (PTL-1). The characteristics of PTL-1, such as its primary structure, molecular mass, and localization, were considered to be similar to those of τ. However, our thorough analysis of the primary structure similarity clarified that PTL-1 is more similar to MAP4 (53% identity) than τ (48% identity), on a structural basis (Fig. 4C). As shown in Fig. 3, AP sequence-bearing proteins emerged at a rather early stage of eukaryotic development, and all of these proteins may be evolutionarily related. Considering the fact that PTL-1 is the only AP sequence-bearing protein in the organism, the closer relationship between PTL-1 and MAP4 is more reasonable, since MAP4 is a ubiquitous protein in mammalian cells and tissues, while τ is restricted to a specific type of cell.

Many questions still remain ahead of us, in our efforts to elucidate the molecular evolution of AP sequence-bearing proteins. First of all, the homology search technique adopted in the present...
study should be pursued further. We should also investigate the biochemical characteristics of the proteins, as well as their tertiary structures from an evolutionary viewpoint. By addressing these issues, we seek to clarify the phylogeny of AP sequence-bearing proteins, and the evolution of microtubule regulating systems in eukaryotic cells in the future.

References

Fig. 4. Alignment of the Repeat regions of MAP2, MAP4, τ, and PTL-1. A, The four sequences are aligned, starting from the first AP sequence. Residues identical in at least two proteins are boxed, and identical residues in all four proteins are marked by asterisks. B, Schematic representation of the new alignment rule. The four sequences are aligned so that the most similar AP sequences are vertically lined up. The AP sequences are depicted as white boxes. The grey boxes are inter-repeat sequences. AP sequences are numbered from the N-terminus. The horizontal lines indicate the absence of homologous sequences. C, Alignment of the four MAP sequences based on the rule in b. Residues identical in at least 2 proteins are boxed, and identical residues in all 4 proteins are marked by asterisks.


