ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

GENOMICS

Genomics 81 (2003) 437–441

www.elsevier.com/locate/ygeno

Short Communication

The identification of a conserved domain in both spartin and spastin, mutated in hereditary spastic paraplegia

Francesca D. Ciccarelli,[a,b] Christos Proukakis,[c,d] Heema Patel,[c] Harold Cross,[e] Shakil Azam,[c] Michael A. Patton,[c] Peer Bork,[a,b] and Andrew H. Crosby[c,*]

[a] *European Molecular Biology Laboratory, Meyerhofstr. 1, 69012 Heidelberg, Germany*

[b] *Max-Delbrueck-Centrum, PO Box 740238, D-13092 Berlin, Germany*

[c] *Department of Medical Genetics, St. George's Hospital Medical School, University of London, Cranmer Terrace, London SW17 ORE, UK*

[d] *Department of Clinical Neurosciences, Royal Free & University College Medical School, London NW3 2PF, UK*

[e] *Department of Ophthalmology, University of Arizona School of Medicine, 655 N. Alveron Way, Tuscon, USA*

Received 24 September 2002; accepted 10 December 2002

Abstract

Multiple sequence alignment has revealed the presence of a sequence domain of ~80 amino acids in two molecules, spartin and spastin, mutated in hereditary spastic paraplegia. The domain, which corresponds to a slightly extended version of the recently described ESP domain of unknown function, was also identified in VPS4, SKD1, RPK118, and SNX15, all of which have a well established and consistent role in endosomal trafficking. Recent functional information indicates that spastin is likely to be involved in microtubule interaction. With this new information relating to its likely function, we propose the more descriptive name 'MIT' (contained within microtubule-interacting and trafficking molecules) for the domain and predict endosomal trafficking as the principal functionality of all molecules in which it is present.

Troyer syndrome, a form of complicated hereditary spastic paraplegia found in the Old Order Amish, was recently found to be due to a frameshift mutation in a novel gene designated *spartin* (*SPG20,* spastic paraplegia autosomal recessive Troyer syndrome) [1]. Screening the non-redundant protein database using the entire amino acid sequence of human spartin, we identified the likely orthologues from *F. rubipens*, *D. melanogaster*, and *C. elegans.* Multiple sequence alignment reveals that the frameshift mutation identified leads to the loss of the C-terminal region, which corresponds to the most highly conserved region of these proteins (data not shown). Whilst no functional information is available for any of the spartin orthologues, restricted regions of similarity with several annotated proteins are clearly identifiable in both the C- and N-termini of the protein. In particular, in the C-terminal portion we detected a strong sequence similarity to a number of uncharacterized plant proteins (data not shown). The only functional aspect reported so far for these sequences is their expression under stress conditions such as dehydration and senescence [2]. Conversely, significantly more is known about the function of the proteins sharing conserved features with the N-terminal region of spartin. We scanned the protein non-redundant database using PSI-BLAST [3] starting with a profile derived from the alignment of the N-terminal regions of human spartin (residues 1–145) and its fugu orthologues. The iterative search converged at the sixth round retrieving several members of different protein families. At the first iteration, the N-terminal part of spastin was identified, with an E value of $1e^{-8}$. Spastin is commonly mutated in autosomal dominant forms of HSP and has recently been implicated in microtubule interaction [4]. Further iterations recovered SNX15 ($E = 3e^{-4}$, second round), proteins similar to ribosomal S6 kinase RSK such as the recently described RPK118 ($E = 3e^{-3}$, third round), SKD1 ($E = 3e^{-3}$, fourth round) and VPS4 ($E = 2e^{-8}$, fifth round). The sorting nexins, including SNX15, are a family of mammalian pro-

* Corresponding author. Fax: +44-0-20-8725-1039.

E-mail address: acrosby@sghms.ac.uk (A.H. Crosby).

0888-7543/03/$ – see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0888-7543(03)00011-9

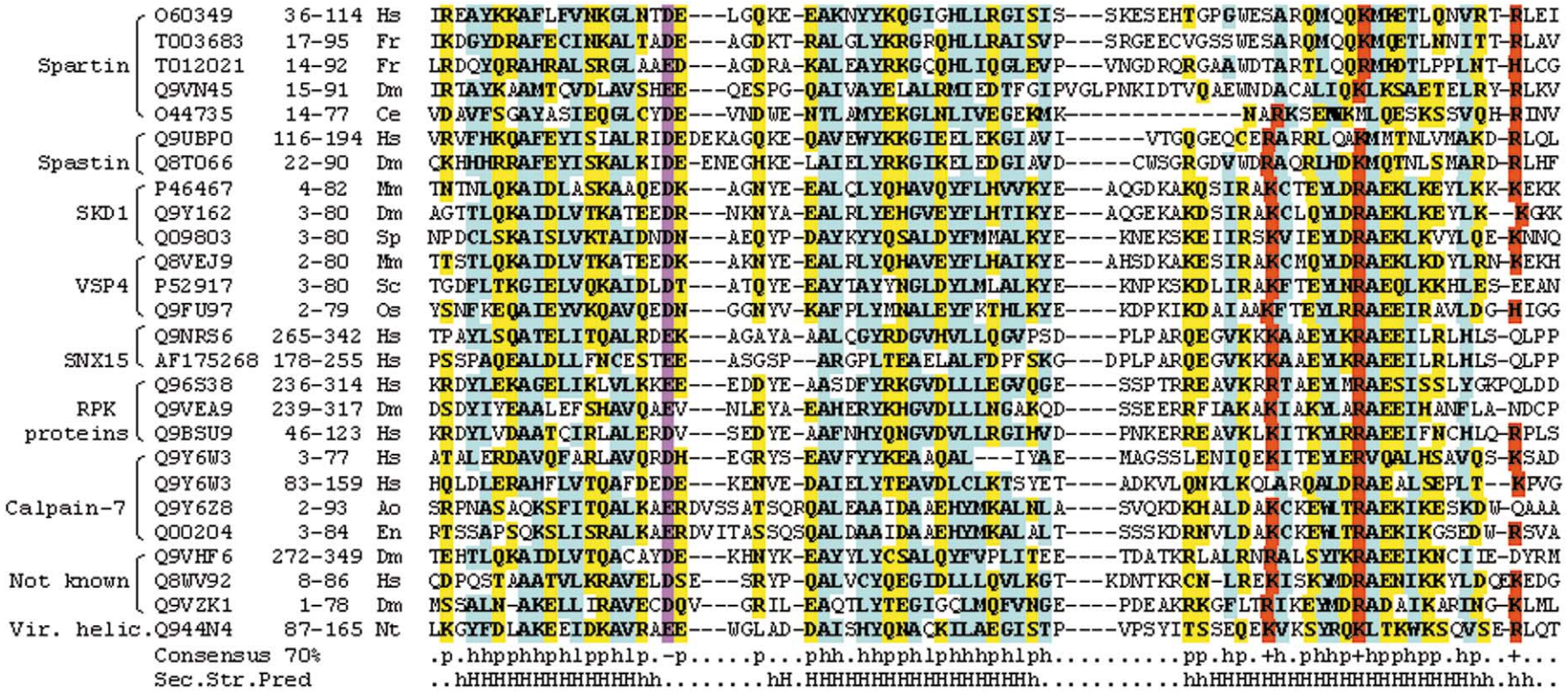

Fig. 1. Multiple sequence alignment of selected MIT domain containing proteins. Sequences are indicated using their database accession number followed by the starting and the ending residues of the domain and by the species. The consensus in at least 70% of the sequences is reported below the alignment: h, l, K, p, R, − and + indicate hydrophobic, aliphatic, lysine, polar, arginine, negatively and positively charged residues, respectively. Hydrophobic and aliphatic residues are highlighted in blue, polar residues in yellow, conserved lysine, arginine, and positive residues in red, negative residues in pink. The secondary structure prediction (Sec.Str.Pred.) at the bottom of the alignment is derived from the alignment (H, helix predicted with expected average accuracy >82%; h, helix predicted with expected average accuracy <82%) [17]. Abbreviations: Ao, *Aspergillus oryzae*; Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; En, *Emericella nidulans*; Fr, *Fugu rubripes*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Nt, *Nicotiana tabacum*; Os, *Oryza sativa*; Sc, *Saccharomyces cerevisiae*; Sp, *Schizosaccharomyces pombe.*

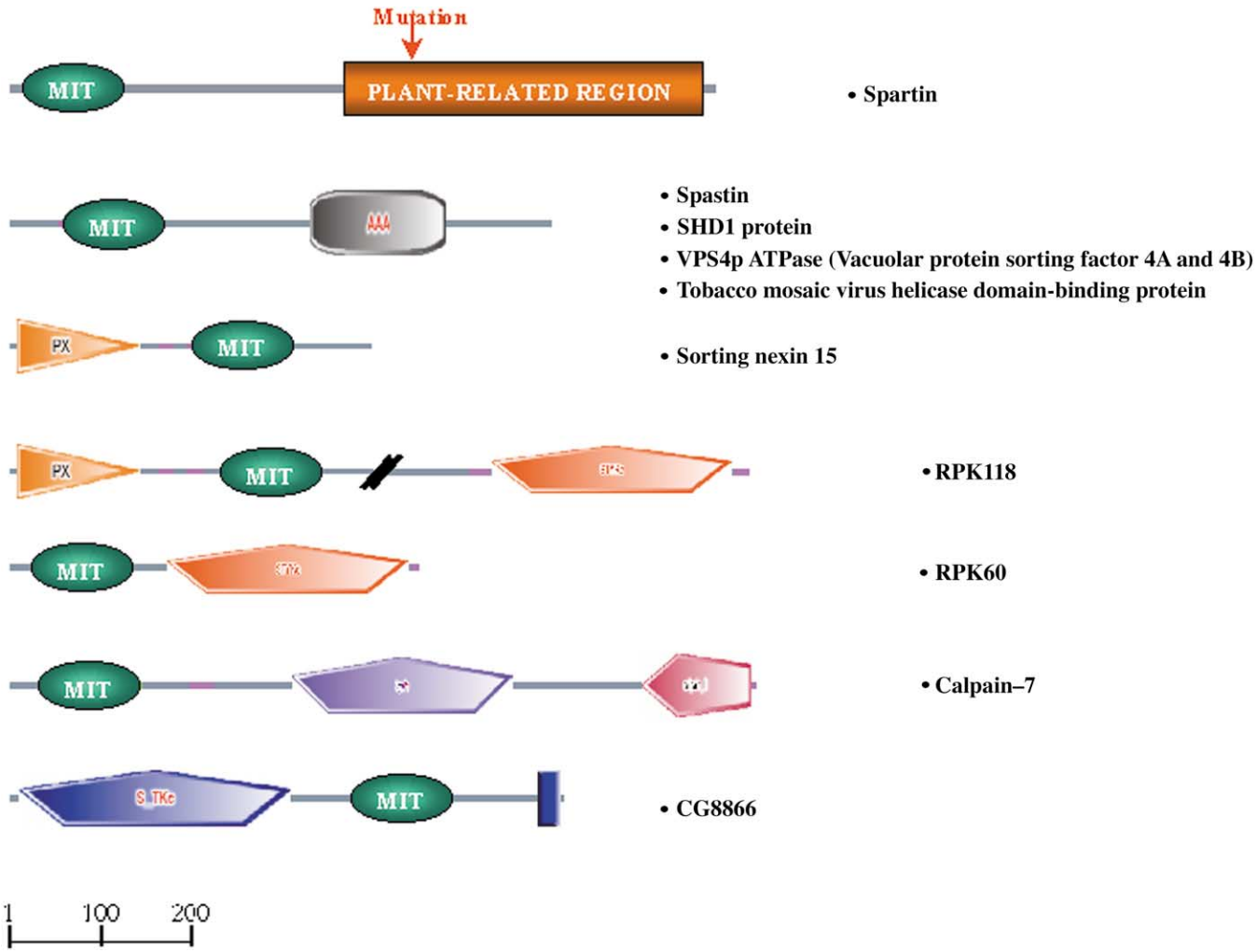

Fig. 2. Domain architectures of proteins containing the MIT domain. The domains are named according to the SMART database [18,19] (http://smart.embl-heidelberg.de). Pink regions represent the segments with low complexity in their amino acidic composition and blue boxes the transmembrane domains. Abbreviations: AAA, ATPase domain associated with different cellular activities; Calpain_III, 3rd calpain protease conserved domain; CysPc, calpain-like thiol protease family; PX, phoX homologous domain present in p47phoX and p40phoX; S_TKc, serine/threonine protein kinases catalytic domain; STYKc, protein kinase domain without a known specificity.

teins whose yeast orthologues are also known to be essential for intracellular protein trafficking [5]. Vps4p is an essential component of the intracellular protein transport machinery and regulates membrane association of several proteins [6]. It is required for normal endosome function, enabling protein transport out of the prevacuolar endosomal compartment [7]. Numerous lines of evidence point to a similar role in protein trafficking of late endosomes and the regulation of endosomal morphology for the recently demonstrated human paralogues VPS4-A and VPS4-B and mouse orthologue SKD1, which also contain this region of similarity [8,9]. Consistent with an endosomal role, human RPK118 has been shown to colocalise with early endosomes [10].

Calpain 7 (retrieved at the second iteration, $E = 2e^{-3}$), an intracellular protease of unknown function [11], is one of the few proteins containing the conserved region that have not yet been shown to be involved in endosomal trafficking. The calpains, calcium-dependent thiol proteases, are widely expressed molecules with both ubiquitous and tissue-specific isoforms involved in cell proliferation, apoptosis, and differentiation. Calpain 7 is the only member of this family shown to contain this region. Intriguingly, other studies indicate that calpain 7 is a highly divergent protease with a novel C-terminal domain [11]. This may indicate that calpain 7 has roles distinct from those of the other members of the group. A number of other proteins were also shown to contain the conserved region: the tobacco mosaic virus helicase domain binding protein (Q944N4) which interestingly also has an AAA cassette like spastin ($E = 2e^{-3}$, third iteration), and the uncharacterized Q8WV92 ($E = 4e^{-3}$, fifth iteration), Q9VHF6 (CG8866, a putative Ser/Thr kinase), and Q9VZK1 ($E = 5e^{-3}$ and $E = 2e^{-4}$, respectively, sixth iteration).

This region partially resembles the previously reported ‘ESP’ domain, so named because it was originally identified in only three molecules and their orthologues: End13/

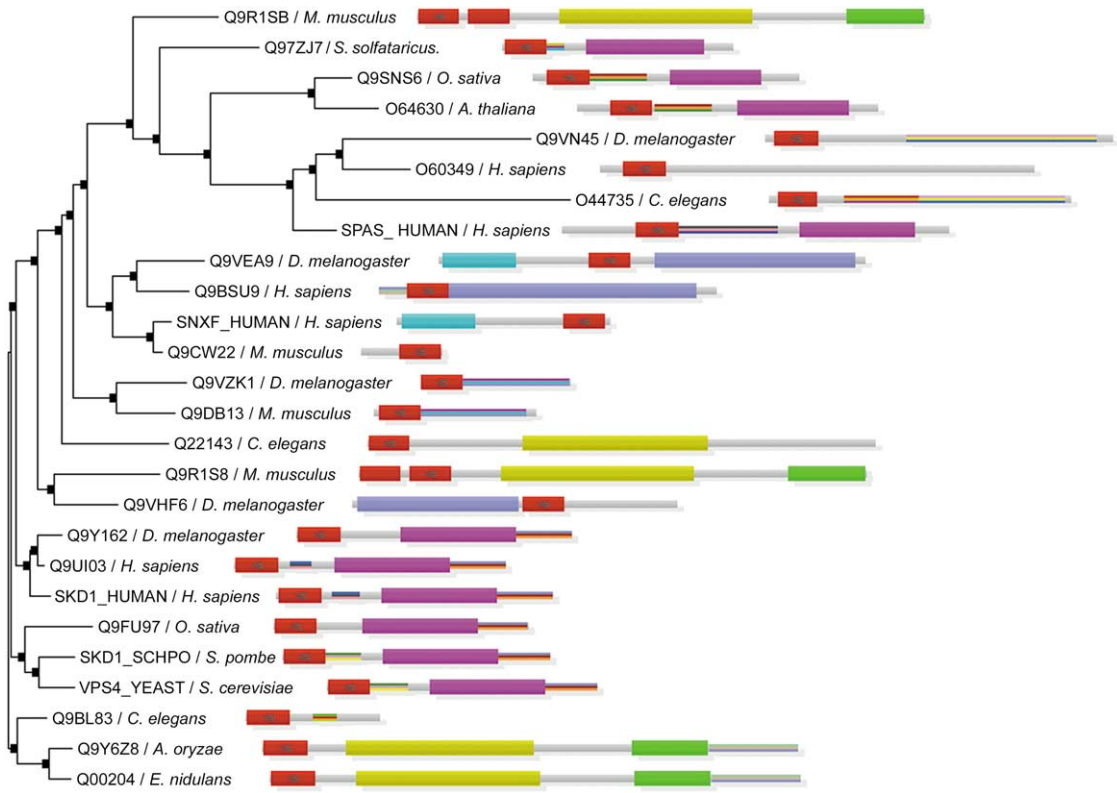

Fig. 3. NIFAS tree [20] showing the similarity of the MIT domains (red box) in the various molecules in which it is contained, indicating that the domains of spartin (O60349) and spastin (SPAS) are closely related being positioned within the same lineage (results from http://www.sanger.ac.uk/Software/Pfam).

Vps4p, SNX15 (sorting nexin 15), and PalB (5). Due to the limited information available at the time, no functional implications were formulated and the domain was named 'ESP'. Multiple sequence alignment of the conserved N-terminal spartin region shared with the molecules discussed above reveals a longer domain comprising approximately 80 amino acids, with alternating conserved hydrophobic and polar areas and a predicted helical secondary structure (Fig. 1). As we have identified this domain in proteins involved in microtubule binding and in intracellular transport, we propose the alternative more descriptive term 'MIT' (contained within microtubule-interacting and trafficking molecules). A schematic representation of proteins containing the MIT domain is shown in Fig. 2. Our analysis reveals that both the full-length and the alternatively spliced form of SNX15 possess the MIT domain, although it was originally said to be present only in the full-length molecule (5). NIFAS, a publicly available tool for the visual analysis of domain evolution in proteins, shows that the MIT domains of spastin and spartin, both of which may be mutated in hereditary spastic paraplegia, are closely related (Fig. 3).

The precise function of the MIT domain is unclear at present. It may be required for interaction with membranes, as some preliminary data from deletion of a part of it in Vps4p suggests [6]. Studies utilizing spastin deletion mutants have indicated that the capability to interact with microtubules resides within the N-terminal portion of the molecule, but no domains were identified in this region [4]. As the MIT domain is in this region, it is possible that it conveys the capability to interact with microtubules. No interaction of spastin with endosomes was detected.

Membrane-bound organelles such as endosomes are known to associate with microtubules. To date, a number of unrelated protein families have been implicated in membrane-microtubule interactions. For example, the cytoplasmic linker proteins (CLIPs) bear an amino terminal domain (MTB) that is able to bind to microtubules only at their plus extremity [12,13]. In common with some of the MIT containing proteins, an overexpression of some CLIP-containing molecules (for example, CLIPR-59) results in a strong perturbation of endosomal trafficking [14]. The Golgi membrane protein GMAP-210 has also been shown to bind microtubules and to perturb the microtubule network when overexpressed [15]. Finally, the LisH (LIS1 homology) motif has been found in proteins associated with neurological diseases, and its presence in p60 katanin, involved in microtubule severing, suggests that this motif binds microtubules [16]. Interestingly spastin is highly homologous to katanin over the AAA cassette, and spastin overexpression caused microtubule disassembly [4], leading to the suggestion that it may have a similar function to katanin. Spastin is, however, not homologous to katanin in its N-terminal half, where it contains a MIT domain and not a LisH motif. As we were unable to find a clear sequence similarity

between MIT and these other microtubule binding domains, it is possible that they may play different roles in protein interactions with microtubules. Given the presence of the MIT domain in the microtubule-interacting spastin and in endosomal molecules, it is tempting to speculate that the MIT domain may be involved in membrane-microtubule interactions. At the very least, the identification of a closely related MIT domain in spastin and spartin suggests that a common pathway underlies the pathogenesis of some forms of hereditary spastic paraplegia.

Acknowledgment

This work was supported by the Birth Defects Foundation (UK).

References

[1] H. Patel, et al., *SPARTIN* is mutated in Troyer syndrome, an hereditary spastic paraplegia, Nat. Genet. 31 (2002) 347–348.
[2] T. Panavas, A. Pikula, P.D. Reid, B. Rubinstein, E.L. Walker, Identification of senescence-associated genes from daylily petals, Plant Mol. Biol. 40 (1999) 237–248.
[3] S.F. Altschul, et al., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[4] A. Errico, A. Ballabio, E.I. Rugarli, Spastin, the protein mutated in autosomal dominant hereditary spastic paraplegia, is involved in microtubule dynamics, Hum. Mol. Genet. 11 (2002) 153–163.
[5] S.A. Phillips, V.A. Barr, D.H. Haft, S.I. Taylor, C.R. Haft, Identification and characterization of SNX15, a novel sorting nexin involved in protein trafficking, J. Biol. Chem. 276 (2001) 5074–5084.
[6] M. Babst, T.K. Sato, L.M. Banta, S.D. Emr, Endosomal transport function in yeast requires a novel AAA-type ATPase, Vps4p, EMBO J. 16 (1997) 1820–1831.
[7] M. Finken-Eigen, R.A. Rohricht, K. Kohrer, The VPS4 gene is involved in protein transport out of a yeast pre-vacuolar endosome-like compartment, Curr. Genet. 31 (1997) 469–480.
[8] S. Scheuring, et al., Mammalian cells express two VPS4 proteins both of which are involved in intracellular protein trafficking, J. Mol. Biol. 312 (2001) 469–480.
[9] T. Yoshimori, et al., The mouse SKD1, a homologue of yeast Vps4p, is required for normal endosomal trafficking and morphology in mammalian cells, Mol. Biol. Cell 11 (2000) 747–763.
[10] S. Hayashi, et al., Identification and characterization of RPK118, a novel sphingosine kinase-1-binding protein, J. Biol. Chem. 277 (2002) 33319–33324.
[11] T. Franz, M. Vingron, T. Boehm, T.N. Dear, Capn7: a highly divergent vertebrate calpain with a novel C-terminal domain, Mamm. Genome 10 (1999) 318–321.
[12] P. Pierre, J. Scheel, J.E. Rickard, T.E. Kreis, CLIP-170 links endocytic vesicles to microtubules, Cell. 70 (1992) 887–900.
[13] F. Perez, G.S. Diamantopoulos, R. Stalder, T.E. Kreis, CLIP-170 highlights growing microtubule ends in vivo, Cell 96 (1999) 517–527.
[14] F. Perez, et al., CLIPR-59, a new trans-Golgi/TGN cytoplasmic linker protein belonging to the CLIP-170 family, J. Cell Biol. 156 (2002) 631–642.
[15] C. Infante, F. Ramos-Morales, C. Fedriani, M. Bornens, R.M. Rios, GMAP-210, A cis-Golgi network-associated protein, is a minus end microtubule-binding protein, J. Cell Biol. 145 (1999) 83–98.
[16] Emes, R.D., and Ponting C.P. A new sequence motif linking lissencephaly, Treacher Collins and oral-facial-digital type 1 syndromes, microtubule dynamics and cell migration. Hum. Mol. Genet. 10 2813–2820.
[17] Rost, B., Sander, C., and Schneider, R. (1994) Redefining the goals of protein secondary structure prediction. J. Mol. Biol. 7;235 13–26. Review.
[18] J. Schultz, F. Milpetz, P. Bork, C.P. Ponting, SMART, a simple modular architecture research tool: identification of signaling domains, Proc. Natl. Acad. Sci. USA 95 (1998) 5857–64.
[19] I. Letunic, et al., Recent improvements to the SMART domain-based sequence annotation resource, Nucleic Acids Res. 30 (2002) 242–244.
[20] C.E. Storm, E.L. Sonnhammer, NIFAS: visual analysis of domain evolution in proteins1, Bioinformatics 17 (2001) 343–348.