

Lars Kammermeier · Jürg Spring · Michael Stierwald ·
Jean-Marc Burgunder · Heinrich Reichert

Identification of the *Drosophila melanogaster* homolog of the human *spastin* gene

Received: 14 April 2003 / Accepted: 5 May 2003 / Published online: 5 June 2003
© Springer-Verlag 2003

Abstract The human SPG4 locus encodes the *spastin* gene, which is responsible for the most prevalent form of autosomal dominant hereditary spastic paraplegia (AD-HSP), a neurodegenerative disorder. Here we identify the predicted gene product CG5977 as the *Drosophila* homolog of the human *spastin* gene, with much higher sequence similarities than any other related AAA domain protein in the fly. Furthermore we report a new potential transmembrane domain in the N-terminus of the two homologous proteins. During embryogenesis, the expression pattern of *Drosophila spastin* becomes restricted primarily to the central nervous system, in contrast to the ubiquitous expression of the vertebrate *spastin* genes. Given this nervous system-specific expression, it will be important to determine if *Drosophila spastin* loss-of-function mutations also lead to neurodegeneration.

Keywords *Drosophila melanogaster* · Spastin · AAA domain · Embryonic CNS

Hereditary spastic paraplegia (HSP) comprises a heterogeneous group of neurological disorders. So far several loci, known as SPG loci, have been identified (reviewed in Casari and Rugarli 2001). The autosomal dominant form of hereditary spastic paraplegia (AD-HSP) is a neurodegenerative disorder characterized by progressive spasticity, mainly of the lower limbs. The disease in 40% of families with the pure form of AD-HSP is linked to SPG4 on chromosome 2p21-p22 due to mutations of a

gene encoding *spastin*. This gene is expressed ubiquitously in fetal and adult human tissues (Hazan et al. 1999). The highest expression levels are found in the brain, with selective expression in the cortex and striatum. In the spinal cord, *spastin* is expressed exclusively in nuclei of motor neurons, suggesting that the strong neurodegenerative defects observed in patients are caused by a primary defect of *spastin* in neurons (Charvin et al. 2003). The human *spastin* gene encodes a predicted 616-amino-acid long protein and is a member of the large family of proteins with an AAA domain (ATPases Associated with diverse cellular Activities). When human *spastin* was identified, no clear invertebrate homologs were known, not even in the complete genome sequence of *Caenorhabditis elegans*; indeed there are a confusing number of similar sequences in humans, *C. elegans*, yeast and even prokaryotes.

With the availability of the first draft of the genome sequence of *Drosophila* (Adams et al. 2000) a predicted gene product, CG5977, with much higher sequence similarity to the human *spastin* than any other AAA domain protein, became evident. Our analysis revealed that the initial prediction, based on the genomic sequence contained errors, that a first version based on a partial cDNA clone (LD23843) was lacking the N-terminus, and that EST clones existed which appeared to extend the similarity with human *spastin* from the N- to the C-terminus. We sequenced one such clone, AT01057 (Fig. 1A). We also used this clone in order to prepare a probe for in situ hybridization. Subsequently AT01057 was also sequenced by the Berkeley *Drosophila* Genome Project. The two sequences are identical and code for a predicted protein of 758 amino acids with extended sequence similarity to human *spastin*, which is therefore called *Drosophila spastin* (*spas*).

Closer inspection of the *Drosophila* sequence and of the alignment of *Drosophila* and human *spastin* reveals, besides the two known highly conserved domains [namely the AAA domain (70% identity) of 300 residues, and the MIT (microtubule interacting and trafficking molecules, Ciccarelli et al. 2003) domain (55% identity)

Edited by D. Tautz

L. Kammermeier (✉) · J. Spring · M. Stierwald · H. Reichert
Institute of Zoology, Biozentrum/Pharmazentrum,
University of Basel,
Klingelbergstrasse 50, 4056 Basel, Switzerland
e-mail: Lars.Kammermeier@unibas.ch

J.-M. Burgunder
National University of Singapore
and National Neuroscience Institute,
11 Jalan Tan Tock Seng, 308433 Singapore, Singapore

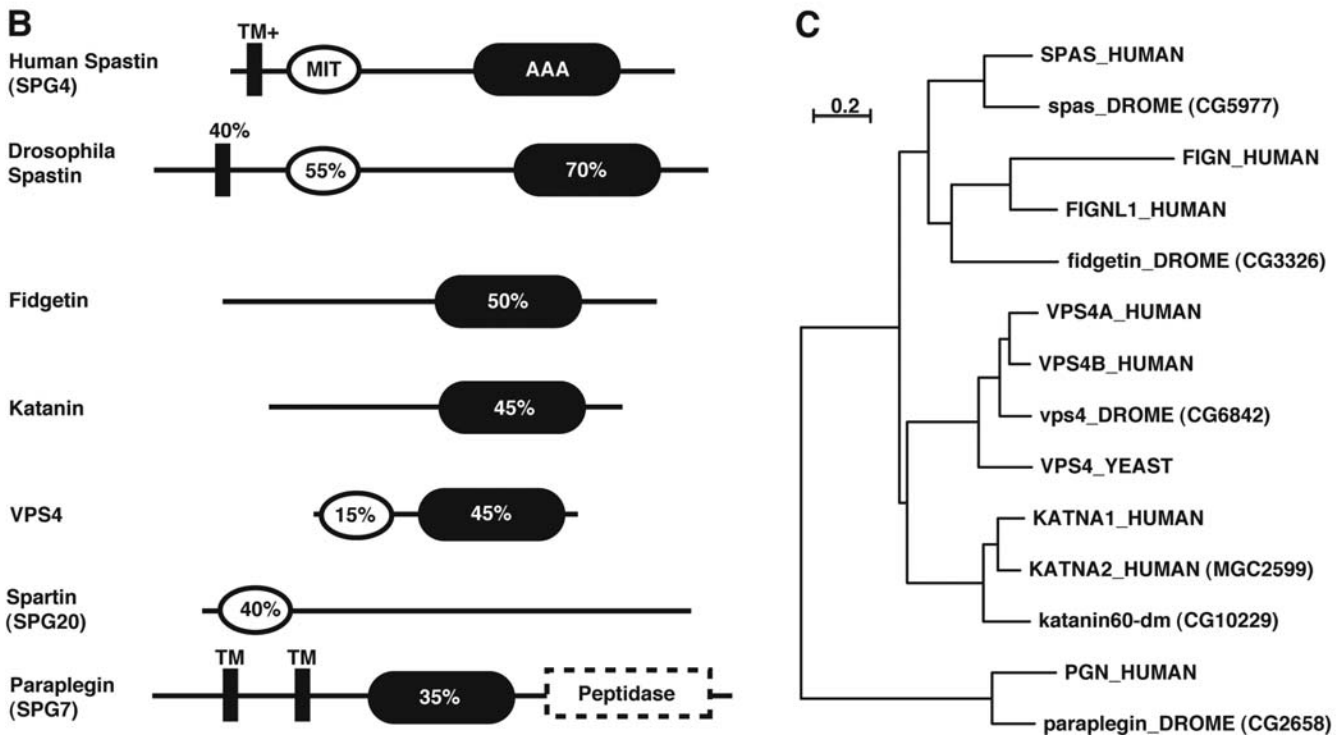
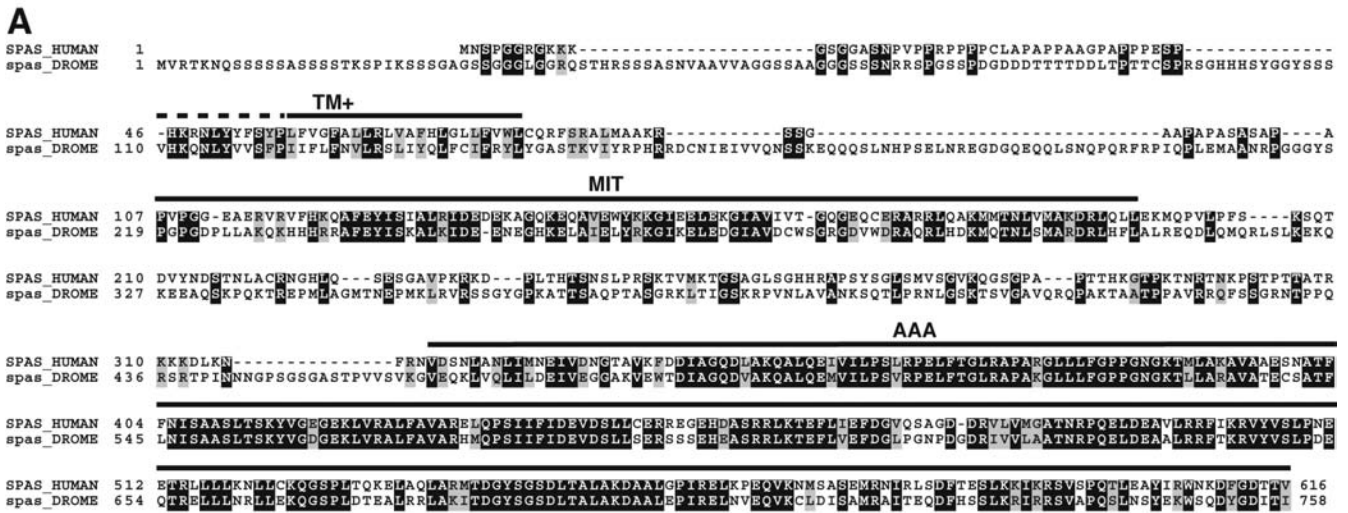


Fig. 1A–C Sequence analysis of *Drosophila* spastin. **A** Comparison of human and *Drosophila* spastin homologs reveals three regions of increased sequence conservation, the extended transmembrane domain *TM+*, the *MIT* (microtubule interacting and trafficking) and the *AAA* (ATPases associated with diverse cellular activities) domains, respectively. **B** Human, mouse and *Drosophila* spastin are the only proteins in current databases with precisely this domain structure. The fidgetin and katanin subfamilies belong to the *AAA* family, but lack transmembrane or *MIT* domains. The *VPS4* subfamily lacks transmembrane domains. *Spartin* (*SPG20*) contains a *MIT* domain with 40% of sequence identity to spastin, but does not belong to the *AAA* family. *Paraplegin* (*SPG7*) belongs

to a very different subgroup of the numerous *AAA* subfamilies with only 35% of sequence identity in the *AAA* domain with spastin and contains unrelated transmembrane domains and a peptidase domain. **C** A phylogenetic analysis based on the *AAA* domain of representative proteins reveals that spastins and fidgetins are most similar in the *AAA* domain, but not in the overall structure, while *VPS4* proteins are less similar in the *AAA* domain although they contain a *MIT* domain and are highly conserved from yeast to humans. *Spartin* and *paraplegin* are single genes in their subfamily in *Drosophila* and humans, while *fidgetin*, *katanin* and *VPS4* are represented by two paralogs in humans

of about 100 residues], an additional conserved region of 33 residues near the N-terminus (Fig. 1A). All three domains are also recognized with SMART domain analysis (Letunic et al. 2002) as *AAA*, *MIT* and as a

putative transmembrane domain, respectively. Since the similarity in the N-terminal domain extends beyond the hydrophobic core that could be a transmembrane domain, we call this a *TM+* domain. This domain appears to be

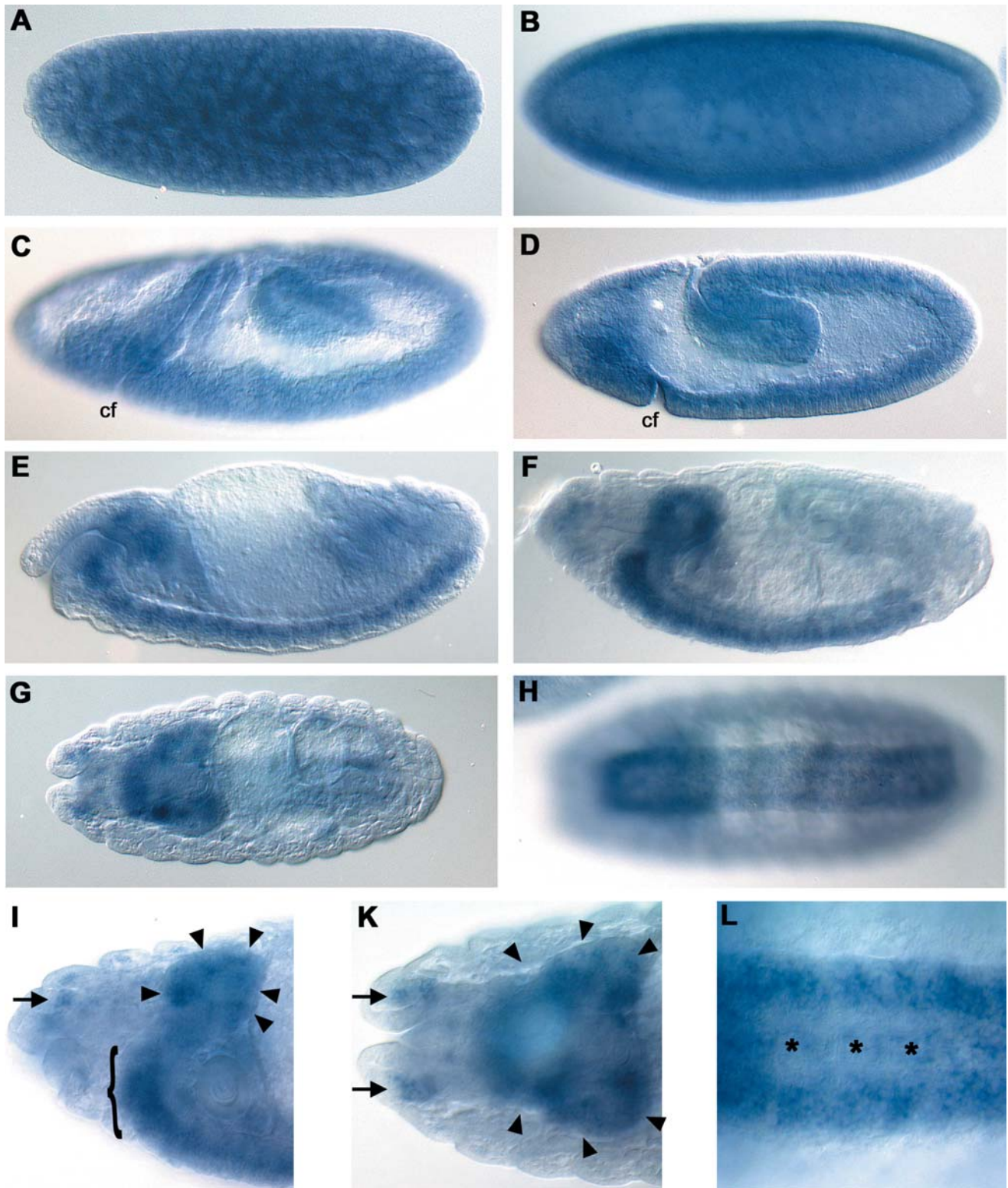


Fig. 2A-L Embryonic expression of *Drosophila spastin*. Whole-mount in situ hybridization of *spastin* transcripts in wild-type embryos. **A-F, I** Lateral views; **G, K** dorsal views; **H, L** ventral views, anterior is to the *left*. **A** Preblastoderm stage, **B** blastoderm stage, **C, D** stage 9, different focal planes, **E** stage 13, **F-L** stage 16–17 embryos; stages according to Campos-Ortega and Harten-

stein (1997). **I, K, L** Enlarged views. *Arrowheads* indicate the expression in cell clusters of the supraesophageal ganglion; *bracket* expression in the subesophageal ganglion; *asterisks* expression in some midline cells of the VNC; *arrows* expression in some sensory head organs of the PNS

specific for spastin proteins. MIT domains, and to a comparable degree AAA domains, also appear in unrelated proteins; one was discovered in *spartin*, the gene responsible for SPG20 (Patel et al. 2002). The association of MIT and AAA domains, however, appears to be evolutionarily old since even the highly conserved vacuolar protein sorting 4 (VPS4) subfamily shows this domain structure. The members of the AAA family most closely related to spastin, however, are not VPS4 but members of the fidgetin subfamily (Fig. 1C), which may be of neurological interest because of the shaking phenotype in the fidget mouse (Cox et al. 2000). *Paraplegin*, the gene responsible for SPG7, however, is only distantly related to *spastin*, and the paraplegin protein is localized to the mitochondria and contains a peptidase domain (Casari et al. 1998; Fig. 1B, C). Since *spastin* was already shown not to be targeted to the mitochondria, but appears to be localized to the nucleus (Charvin et al. 2003), it will be important to test whether the potential transmembrane domain with the highly conserved neighbouring residues could be involved in localization of spastin to the nuclear membrane system. Other studies using expression systems of wild-type and mutated spastin with reporter sequences in different cell lines have suggested an involvement of spastin in microtubule dynamics (Errico et al. 2002).

The spatial and temporal expression patterns of *Drosophila spastin* were analysed by in situ hybridization. For this, the EST clone AT01057 that covers *spastin* was used to generate digoxigenin-labeled sense and antisense RNA probes (Tautz and Pfeifle 1989). These were hybridized to wild-type whole-mount embryos. Maternally contributed *spastin* transcripts characterized early embryogenesis with high expression levels seen until blastoderm stage (Fig. 2A). At the cell formation stage, expression was strongest near the basal part of the cell layer underlying the surface (Fig. 2B). During germband extension and stomodeal plate formation, expression was seen in the ventral head and trunk ectoderm, as well as in cells near the cephalic furrow and in the invaginating hindgut and midgut primordia (Fig. 2C, D). After germband retraction and delamination of neuroblasts at stage 13, transcripts were observed in subsets of cells in all neuromeres of the CNS including those of the supraesophageal and subesophageal ganglia (Fig. 2E). In later embryonic stages, marked expression of *spastin* was observed in cell clusters throughout the supraesophageal ganglion (Fig. 2F, G, I, K, arrowheads), with pronounced expression also seen in the subesophageal ganglion (Fig. 2I, bracket). In the ventral nerve cord (VNC), transcripts were seen in two broad longitudinal stripes located laterally, and weaker expression was observed in some midline cells (Fig. 2H, L, asterisks). In addition to expression in the CNS, *spastin* transcripts were also

observed in some sensory head organs of the peripheral nervous system (PNS), most probably the Bolwig's organs and/or the dorsal organs (Fig. 2F, G, I, K, arrows).

Thus, during embryogenesis, expression of *Drosophila spastin* is mainly restricted to the central nervous system, in contrast to the ubiquitous expression of the vertebrate *spastin* gene (Hazan et al. 1999).

Acknowledgements We thank Boris Egger, Simon Sprecher, and Urs Stiefel for technical assistance, and Bruno Bello, Frank Hirth and Ronny Leemans for comments on the manuscript. This work was supported by the Swiss NSF.

References

- Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amantides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science* 287:2185–2195
- Campos-Ortega J, Hartenstein V (1997) The embryonic development of *Drosophila melanogaster*, 2nd edn. Springer, Heidelberg Berlin New York
- Casari G, Rugarli E (2001) Molecular basis of inherited spastic paraplegias. *Curr Opin Genet Dev* 11:336–342
- Casari G, De Fusco M, Ciarmatori S, Zeviani M, Mora M, Fernandez P, De Michele G, Filla A, Cocozza S, Marconi R, et al. (1998) Spastic paraplegia and OXPHOS impairment caused by mutations in paraplegin, a nuclear-encoded mitochondrial metalloprotease. *Cell* 93:973–983
- Charvin D, Cifuentes-Diaz C, Fonknechten N, Joshi V, Hazan J, Melki J, Betuing S (2003) Mutations of SPG4 are responsible for a loss of function of spastin, an abundant neuronal protein localized in the nucleus. *Hum Mol Genet* 12:71–78
- Ciccarelli FD, Proukakis C, Patel H, Cross H, Azam S, Patton MA, Bork P, Crosby H (2003) The identification of a conserved domain in both spartin and spastin, mutated in hereditary spastic paraplegia. *Genomics* 81:437–441
- Cox GA, Mahaffey CL, Nystuen A, Letts VA, Frankel WN (2000) The mouse *fidgetin* gene defines a new role for AAA family proteins in mammalian development. *Nat Genet* 26:198–202
- Errico A, Ballabio A, Rugarli EI (2002) Spastin, the protein mutated in autosomal dominant hereditary spastic paraplegia, is involved in microtubule dynamics. *Hum Mol Genet* 11:153–163
- Hazan J, Fonknechten N, Mavel D, Paternotte C, Samson D, Artiguenave F, Davoine CS, Cruaud C, Durr A, Wincker P, et al. (1999) Spastin, a new AAA protein, is altered in the most frequent form of autosomal dominant spastic paraplegia. *Nat Genet* 23:296–303
- Letunic I, Goodstadt L, Dickens NJ, Doerks T, Schultz J, Mott R, Ciccarelli F, Copley RR, Ponting CP, Bork P (2002) Recent improvements to the SMART domain-based sequence annotation resource. *Nucleic Acids Res* 30:242–244
- Patel H, Cross H, Proukakis C, Hershberger R, Bork P, Ciccarelli FD, Patton MA, McKusick VA, Crosby AH (2002) SPG20 is mutated in Troyer syndrome, a hereditary spastic paraplegia. *Nat Genet* 31:347–348
- Tautz D, Pfeifle C (1989) A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* 98:81–85