# Minireview

# WH2 domain: a small, versatile adapter for actin monomers

Eija Paunola, Pieta K. Mattila, Pekka Lappalainen\*

Program in Cellular Biotechnology, Institute of Biotechnology, Viikki Biocenter, P.O. Box 56, University of Helsinki, 00014 Helsinki, Finland

Received 20 October 2001; revised 30 October 2001; accepted 14 November 2001

First published online 6 December 2001

Edited by Gianni Cesareni and Mario Gimona

Abstract The actin cytoskeleton plays a central role in many cell biological processes. The structure and dynamics of the actin cvtoskeleton are regulated by numerous actin-binding proteins that usually contain one of the few known actin-binding motifs. WH2 domain (WASP homology domain-2) is a  $\sim$ 35 residue actin monomer-binding motif, that is found in many different regulators of the actin cytoskeleton, including the β-thymosins, ciboulot, WASP (Wiskott Aldrich syndrome protein), verprolin/ WIP (WASP-interacting protein), Srv2/CAP (adenylyl cyclaseassociated protein) and several uncharacterized proteins. The most highly conserved residues in the WH2 domain are important in  $\beta$ -thymosin's interactions with actin monomers, suggesting that all WH2 domains may interact with actin monomers through similar interfaces. Our sequence database searches did not reveal any WH2 domain-containing proteins in plants. However, we found three classes of these proteins: WASP, Srv2/ CAP and verprolin/WIP in yeast and animals. This suggests that the WH2 domain is an ancient actin monomer-binding motif that existed before the divergence of fungal and animal lineages. © 2002 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

*Key words:* Actin; Wiskott Aldrich syndrome protein homology domain-2; β-Thymosin; Verprolin

#### 1. Introduction

Actin is a major constituent of the cytoskeleton in all eukaryotic cells. It exists both in a monomeric form (G-actin) and in polar filamentous structures (F-actin). The actin monomer is always bound to a nucleotide, either ATP or ADP, and the nucleotide status significantly affects its biochemical properties. In most cells, the actin cytoskeleton is highly dynamic and rapid actin filament assembly and disassembly (i.e. turnover) are required for many cellular processes, such as endocytosis, motility, division and morphogenesis. The dynamic nature of actin filaments enables cells to respond quickly to extracellular signals [1].

The rapid turnover is accomplished by a plethora of actinbinding proteins. Many of these proteins interact with actin through one of the following actin-binding motifs: the calponin homology domain [2,3], the ADF-H domain [4], the gelsolin homology domain [5] or the thymosin  $\beta$ 4/WH2 (<u>W</u>ASP <u>homology domain-2</u>) domain [6,7]. The calponin homology domain appears to interact only with filamentous actin, while the ADF-H and the gelsolin homology domains are able to bind both monomeric and filamentous actin.

The WH2 domain is a small (approximately 35 amino acids) motif found in a number of different proteins. All available biochemical data indicate that the WH2 domains bind to actin monomers. Using BLAST and SMART searches, we identified WH2 domains in a total of 37 proteins from Caenorhabditis elegans, Drosophila melanogaster, Saccharomyces cerevisiae and Homo sapiens sequence databases. The molecular weights of these proteins range from 5 kDa to 150 kDa, and they contain from one to four copies of the domain per protein. Yeast has three WH2 domain-containing proteins whereas humans have 18 different WH2 domain-containing proteins. Interestingly, we did not identify any WH2 domain proteins from plants in our database searches. The WH2 domain was recently reported in a group of minor baculoviral capsid proteins, and it was suggested that these WH2 domaincontaining proteins may regulate actin dynamics during the infection process [7]. Furthermore, WH2 domain-like sequences have been identified in Listeria monocytogenes ActA protein: a protein that enables Listeria to exploit the host's actin cytoskeleton for its motility during the infection process [8]. However, these ActA sequences do not fulfill our criteria to be called a WH2 domain, because they lack some of the WH2 domain's most highly conserved and functionally critical residues.

# 2. Different classes of WH2 domain proteins

We identified 37 WH2 domain-containing proteins from the nearly completed human, fly, worm and budding yeast genomes. Many of these proteins belonged to previously identified WH2 domain-containing protein families:  $\beta$ -thymosins, ciboulots, WASPs (<u>Wiskott Aldrich syndrome proteins</u>) and verprolins. We also identified previously unrecognized WH2 domain in Srv2/CAP (adenylyl <u>cyclase-associated protein</u>) family proteins, which are known regulators of actin dynamics (Fig. 1), and we found several novel WH2 domain-containing proteins.

## 2.1. The $\beta$ -thymosin family

 $\beta$ -Thymosins are a family of highly abundant actin monomer-binding proteins. They are small proteins of 44–45 residues and are entirely composed of a single WH2 domain.  $\beta$ -Thymosins are found in vertebrates, and have not been identified in lower eukaryotes. Mammalian cells express several  $\beta$ -thymosin isoforms: thymosins  $\beta$ 4,  $\beta$ 10 and  $\beta$ 15. Their

<sup>\*</sup>Corresponding author. Fax: (358)-9-19159366.

E-mail address: pekka.lappalainen@helsinki.fi (P. Lappalainen).

<sup>0014-5793/02/</sup> $22.00 \otimes 2002$  Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved. PII: S 0 0 1 4 - 5793(0 1) 0 3242 - 2



Fig. 1. Domain structures of WH2 motif-containing proteins. Thymosin  $\beta$ 4 is composed of a single WH2 domain. *Drosophila* ciboulot has three tandem WH2 domains. Yeast verprolin and human WIP are proline-rich proteins that contain two WH2 domains in their N-terminal regions and a WASP-binding domain near their C-terminus. The N-terminal region of N-WASP contains a WH1 domain and a CRIB region that are important for its regulation, its two WH2 domains and its C-terminal acidic region are essential for interactions with actin and Arp2/3. Similarly, the N-terminal regions of WAVE proteins have a regulatory role, while the WH2 domains and acidic regions are involved in the Arp2/3-induced actin filament assembly. Srv2/CAP contains the cyclase-binding domain followed by a proline-rich region and a WH2 domain.

amino acid sequences are approximately 70% homologous (Fig. 2). Thymosin  $\beta4$  is the most extensively studied member of the group. It is ubiquitously expressed [9], whereas thymosins  $\beta10$  and  $\beta15$  appear to be expressed at certain stages of embryonic development [10] and in cancer cells, respectively [11,12].

Thymosin  $\beta$ 4 forms a 1:1 complex with actin monomers and it binds ATP-actin monomers with a 100-fold higher affinity than ADP-actin monomers. Thymosin  $\beta$ 4 prevents actin polymerization and its high cellular concentration ( $\leq 600 \ \mu M$ ) allows it to maintain the large unpolymerized actin monomer pool in the cytoplasm. Although actin bound to thymosin  $\beta$ 4 does not polymerize, profilin can compete with thymosin  $\beta$ 4 for actin-binding and may release the actin monomer from thymosin  $\beta 4$  onto the barbed end of the actin filament [13]. At high concentrations (100-200 µM) thymosin β4 was reported to interact with F-actin creating defects in the structure of the polymer. Thus, thymosin  $\beta$ 4 may have a more complex role in actin dynamics, and this may be especially true in motile blood cells, where its concentration can be up to 500  $\mu$ M [14]. A more complex role for  $\beta$ -thymosins in actin dynamics is also supported by the observation that overexpression of one member of the family, thymosin  $\beta 10$ , decreases the cytoplasmic actin monomer pool in cultured NIH3T3 cells [15].

#### 2.2. Ciboulot-like proteins and actobindin

Recently WH2 domain-containing proteins with strong sequence similarity to  $\beta$ -thymosins were identified in *D. melanogaster* and in *C. elegans*. These proteins were named ciboulot [16] and tetrathymosin  $\beta$  ([6]; Van Troys et al., unpublished). Yeast and mammalian ciboulot homologues have not been identified to date. D. melanogaster ciboulot has a molecular weight of 14.4 kDa and is composed of three WH2 domains, whereas the C. elegans tetrathymosin  $\beta$  contains four WH2 domains [9]. Although Drosophila ciboulot contains three potential actin-binding sites (WH2 domains) it forms a 1:1 complex with actin monomers. Ciboulot shares strong sequence similarity with  $\beta$ -thymosins, but the biochemical properties of these two classes of WH2 domain proteins differ from each other. Ciboulot binds G-actin like β-thymosin, but it promotes the assembly of actin monomers at the filament's barbed end while β-thymosin inhibits actin polymerization [16]. Ciboulot's activity is similar to profilin's, a structurally unrelated actin monomer-binding protein [17], and overexpressing profilin compensates for the lack of ciboulot during the development of adult brain [16]. The C. elegans tetrathymosin displays similar profilin-like properties in vitro, which most probably arise from a differential, though cooperative activity of the WH2 repeats which are all able to bind actin. Interestingly, one WH2 repeat functions as a pure sequesterer whereas another preferentially interacts with F-actin (Van Troys et al., unpublished).

Actobindin is a small actin-binding protein of *Acanthamoe*ba castellanii. It consists of two WH2 domains and binds one or two actin molecules depending on the proteins' concentrations [18]. The heterotrimeric actobindin–actin complex is incompetent for nucleation, self-association and elongation; however, when actobindin dissociates from this complex, the actin dimer can join to a pre-existing filament. Thus actobindin can accelerate the elongation at the ends of uncapped filaments while blocking formation of new filaments [19]. Further biochemical studies are required to determine whether actobindin's function is more closely related to that of ciboulot or the  $\beta$ -thymosins.

#### 2.3. WASP, N-WASP (neural WASP) and WAVE

WASP was identified as a gene that is mutated in the hematopoietic cells of patients with Wiskott Aldrich syndrome. Its overexpression caused actin clustering, suggesting a role in actin polymerization [20]. N-WASP was initially isolated from the brain [21] but is ubiquitously expressed and its sequence is approximately 50% similar to WASP. WASP, N-WASP, and their yeast homologue, Las17p, stimulate the Arp2/3 complex to nucleate actin filaments [22–24].

Both WASP and N-WASP contain several different protein domains, including a WH1/EVH1 domain, a GBD/CRIB domain, a basic region, a proline-rich domain, a cofilin-like domain and an acidic region. WASP also contains one WH2 domain and N-WASP contains tandem WH2 domains (Fig. 1). The activity and localization of WASP and N-WASP are regulated through the N-terminal WH1/EVH1 and GBD/ CRIB domains [23,25,26], and their effector domains are the WH2 and cofilin-like domains and the acidic region [27]. The acidic region, together with its neighboring elements, binds the Arp2/3 complex causing it to initiate the assembly of new filaments. The WH2 domain binds actin monomers and is believed to facilitate the assembly of these monomers into newly forming actin filaments [28]. A synthesized WH2 domain of WASP inhibits the spontaneous nucleation of actin filaments and prevents the assembly of monomers to the filaments' pointed ends but not to their barbed ends [29].

WAVE1/Scar proteins were identified by virtue of their sim-

		* *		**						
Hs THYMB10	MADKPD	MGEIA	SFDKA	-KLK	TET	-OEKNTI	PTKET	IEOEKR	SEIS-	44
HS THYMB4	MSDKPD	MAEIR	KFDKS	-KLK	TET	OEKNPI	PSKET	IEOEKO	AGES-	44
Hs THYMB15	MSDKPD	LSEVE	TFDKS	-KLK	TNT	-EEKNTI	PSKET	IOOEKE	YNORS	45
Dm CibD2	KTOOSI	FEGIT	AFNON	-NLK	TET	-NEKNPI	PDKEA	IEOEKE	K	41
Dm CibD3	KEKNOF	IAGIE	ENFDAK	-KLK	TET	-NEKNVI	PTKEV.	IEAEKO	A	41
Ce CibD3	LEL	TDKIN	INFPSE	-NLK	TET	-IEKNVI	PSPTD	VAREKT	L	38
Ce CibD2	KOHVER	IHEIE	HFDST	-KLH	STPV	-KEKIVI	PSADD	IKQEKQ	H	41
Dm CibD1	KVAENL	KSQLE	GFNQD	-KLK	NAST	-QEKIII	PTAED	VAAEKT	Q	41
Ce_CibD1	LPKMNQEL	AGAVE	REGL	-ELK	<b>VET</b>	-TEKNVI	PTKED	VAEEKQ	H	41
Ce CibD4	QM	AASFI	K-SA	LHI	IVET	-IVST	DV	RVTEAO		28
Hs_Q9Y544	PTGDNSEL	LAEIF	A-GK	-SLK	PTPQ	-SKGLT-	TVFSG	IGQPAF	QP	41
Hs_CAP	ESASRSAL	FAQIN	Q-GESIT	HALK	IVSD	-DMKTH-	KNPAL	KAQSGP	V	44
Ce_Vrp2	SSNARNAL	LGDIH	IK-GL	-KLK	TVT	-NDRSA-	PSVGK	VVGSSG	ss	41
Sc_Vrp1D1	VMQGRDAL	LGDIE	RK-GM	-KLK	KAET	-NDRSA-	PIVGG	GVVSSA	SG	41
Hs_WIPD1	EQAGRNAL	LSDIS	SK-GK	-KLK	KTVT-	-NDRSA-	PILDK	PKGAGA	GG	41
Hs_043312	DTPQGEDM	LNAIR	RR-GV	-KLK	<b>TTT</b>	-NDRSA-	PRFS-			32
Hs_095763	DPKGRSAL	LADIÇ	QQ-GT	-RLR	KVTQ:	INDRSA-	PQIES	SKGTN-		39
Ce_Vrp1D2	GTVDRGEF	LKGIÇ	QG-GF	-KLK	TTT.	-NDKSG-	-LFVDE	EMRERS	VK	41
Sc_Vrp1D2	IPGMGAPQ	LGDII	LAGGIP	-KLK	IINN	-NASTK-	- <mark>PSP</mark> SA	SAPPIP	GA	43
Dm_Vrp1D1	GADARSAL	LSSI	QK-GT	-KLK	<b>TTT</b>	-VDKSG-	-PALSG	KVCGGD	GG	41
Dm_Vrp1D2	GLGNGTPK	LGGLI	E-GLSQM	PKLK	PVNG	-IRAT	-PSAGS	AATTTS	KS	44
Hs_WIPD2	FGGGGPPG	LGGLI	QAGMP	-KLR	STAN	-RDNDS	GSRPP	LLPPGG	R	43
Ce_Vrp1D1	HRNSNENA	QLEI	KK-GF	-KLR	PTKT-	-VDKSKI	VITAE	NEDESE	v	41
Hs_CAP2	SSPSRSAL	FAQLI	IQ-GEAIT	KGLR	IVTD	-DQKTY-	KNPSL	RAQGGQ	TQ	45
Dm_CAP	AGDDRSAL	FAQIN	Q-GADIT	KGLK	<b>KVTG</b>	-DMQTH-	-KNPSL	RTGPAP	F	44
Ce_CAP1	DKASRDAL	FASL	IQ- <mark>GEGVT</mark>	SRLK	KVTA-	-DMQTH-	-KN <mark>P</mark> NL	RGTAVV	PA	45
Sc_Srv2	NKGGIGAV	FAEL	Q-GENIT	KGLK	KVDK-	-SQQTH-	-KNPEL	RQSSTV	SS	45
Ce_CAP2	KPSGVGAL	LESL	T-GLSAT	SRLK	KVTP-	-EMQTH-	-KNPVL	REVNGQ	MN	45
Hs_075128D1	DTSLHSAL	MEAIH	ISAGGKD-	-RLR	<b>TAE</b>	-HTGEGI	RPAKLS	YTEAEG	E	44
Hs_075128D2	AEGERSAL	LAAIH	RGHSGTC-	-SLR	KVAS-	-SASEEI	QSFRD.	AALSAQ	G	44
Sc_Las17	GDAGRDAL	LASI	RGAGGI	GALR	KVDK-	-SQLDK-	-PSVLL	QEARGE	SA	44
Dm_SpireD2	EPSPREQL	MESI	RK- <mark>G</mark> K	-ELK	2ITP.	- <mark>PEAP</mark> TI	RERVL	PSANST	L	41
Hs_AA64	EKNSRDQL	LAAII	RS-SN	LK	2LKK	-VEV <mark>P</mark> KI	LL	GTGH		34
Hs_Q9UDY7	ISDAHSDL	LSAI	CQ-GF	-QLR	RVEE	-QQEQE-	-KWDVV	GNDVAT	IL	41
Hs_Wave2	VSDARSDL	LSAI	RQ- <mark>G</mark> F	-QLR	RVEE	-QREQE-	-KRDVV	GNDVAT	'I	40
Hs_Wave1	ISDARSVL	LEAIH	RK-GI	-QLR	KVEE-	-QREQE	KHERI	ENDVAT	'I	41
Hs_Wave3	ISDARSDL	LAAIH	RM-GI	-QLK	KVQE-	-QREQE	AKREPV	GNDVAT	'I	41
Dm_Scar	FHDPRNDL	KMAIH	RD-GI	-TLR	KVEK-	-SEQKE-	-IEMAA	PLDVAS	I	40
Ce_WaspD2	GGDARGDV	MAQIE	RQ-GA	-QLK	HVDA	-AAEQE-	-RRKST	TSGAAG	MG	41
Dm_N-WaspD1	GGDPRDAF	LESIE	RQ-GY	-KLK	KVDQ-	-KAATI-	-SGIKP	RPERKP	VT	41
Ce_WaspD1	AQDGRSNL	LAEI	QA-GK	-QLR	SVQQ	-TADSP-	-KSAGG	DARGDV	MA	41
Dm_Wasp	APDPRNAL	MDAIN	RK-GT	-QLK	KVDT-	-TALST	3S	GDSRSD	LMDI-	40
Ce_Scar	QPDARSDL	LAQIO	2S-GI	-KLK	KVQR-	-AEAEA	AENAAL	EANNVA	A	41
Hs_N-WaspD2	SCSGRDAL	LDQII	Q-GI	-QLK	SVAD	-GQESTI	PTPAP	TSGIVG	A	41
Hs_Wasp	PGGGRGAL	LDQII	Q-GI	-QLN	KTPG	-APESSA	ALQPPP	QSSEGI	V	41
HS_N-WaspD1	TAGNKAAL	LDQII	KE-GA	-QLK	VEQ.	-NSRPV-	SCSGR	DALLDQ	IR	41
Dm_CG6771	RERRSGRE	LKILI	KS-KLT	-KLK	VKEE	AAKKEI	DALKQ.	AMKKNQ	5	42
Dm_SpireD1	ELTPYEIL	MGDIE	CAKKY	-QLR	KVMV-	-NGDIP-	PRVKK	DAHAMI	<b>L</b> E	42
Dm_N-waspD2	RKPVTTDF	LSEL	CL-GI	-TLR	VKN	-PADN-I	ISEES	ESQA		37
HB_095559	PMGDNSEL	LAETI	A-GK	-SLK	LIL60	-SKGLT-	-TVFSG	SKQPAF	QP	41

Fig. 2. Multiple sequence alignment of WH2 motifs found in human, worm, fly, and budding yeast proteins. Residues that are essential for actin-binding are indicated with asterisks. The position of the N-terminal α-helix is shown above the sequences. Protein names, database, and accession numbers are listed below. *H. sapiens*: CAP SwissProt Q01518, CAP2 SwissProt P40123, ID Q9Y544 TrEMBL AL031848, ID 095550 TrEMBL AL035288, ID 043312 TrEMBL AB007889, ID 095763 TrEMBL AC004912, ID O75128 TrEMBL AB014533, N-WASP TrEMBL D88460, ID 043312 TrEMBL AB007889, WIP TrEMBL AF031588, ID 060794 TrEMBL AL022578, ID 075128 TrEMBL AB014533, WAVE3 TrEMBL AB020707, ID 095559 TrEMBL AL035288, ID 095763 TrEMBL AC004912, ID Q9UDY7 TrEMBL AF134304, WASP TrEMBL AF115549, WAVE3 TrEMBL AB026543, ID Q9Y544 TrEMBL AL031848, WAVE2 TrEMBL AB026542, ID AA64 PIR S26815, WAVE1 EMBL D87459, thymosin β4 EMBL M17733, thymosin β10 SwissProt P13472, thymosin β15 NCBI AAH000183. *S. cerevisiae:* Srv2 SGD YNL138W, Las17 SGD YOR181W, Vrp1 SGD YLR337W. *C. elegans*: Cib F08F1.8 NCBI AAB71308.1, Vrp2 PIR T16755, Vrp1 PIR T25220, Scar PIR T23959, CAP1 NCBI AAK85482, WASP PIR T15446. *D. melanogaster*: Scar NCBI AAF53042, Cib NCBI CAA21832, Vrp1 NCBI AAF46800.1, N-WASP NCBI AAF5448.1, WASP NCBI AAF56819, CG6771 gene product NCBI AAF54641, CAP NCBI AAF51408, SPIRE TrEMBL AF184975.

ilarity to WASP and N-WASP [30]. Mammalian WAVE1 and WAVE3 are expressed only in brain, whereas WAVE2 is ubiquitously expressed except in skeletal muscle [31]. The WAVEs consist of an N-terminal SHD or <u>Scar homology domain</u>, a basic region, a long proline-rich region, a WH2 domain, a cofilin-like region and a highly acidic region at the C-terminus (Fig. 1). However, they lack the N-terminal WH1 domain and GBD domain, and are assumed to be regulated differently. Like WASP and N-WASP, WAVE1/Scar's acidic region activates the Arp2/3 complex, and the WH2 domain is believed to facilitate the assembly of actin monomers to the newly formed filaments [32].

The WASP proteins activate the Arp2/3 complex to different degrees. The strongest activator is N-WASP. The studies with chimeric proteins revealed that neither the actin-binding activity of the WH2 domain nor the Arp2/3 complex-binding affinity of the CA domain is related directly to a potency to activate Arp2/3 complex. Instead, the number of WH2 domains is important, the tandem WH2 domain arrangement being the most effective [33].



# 0.1

Fig. 3. An unrooted phylogenetic tree of WH2 domains. This tree was produced from the alignment in Fig. 2 by the Clustal-X software. Circles indicate the six classes of WH2 domain-containing proteins described in this article. A bar showing 10% divergence is included.

#### 2.4. Verprolin and WIP

Verprolin is an 817-amino acid proline-rich (22%) WH2 domain-containing protein in the yeast *S. cerevisiae*. Cells containing mutant verprolins have an aberrant cytoskeleton and defects in endocytosis, distribution of mitochondrial proteins and bud-site selection [34,35]. The carboxyl-terminus of verprolin interacts with Las17p, the yeast homologue of WASP [36]. The WH2 domain-containing N-terminal 70 amino acids of verprolin bind actin in a yeast two-hybrid assay, but deleting these residues does not totally prevent actin-binding [35]. Our BLAST searches reveal that verprolin has a second WH2 domain, perhaps explaining the verprolin's N-terminal deletant's residual actin-binding activity (Figs. 1 and 2).

WIP (<u>WASP-interacting protein</u>) is a human 503-amino acid long protein, that is homologous to yeast verprolin. Like its yeast homologue, WIP contains two WH2 domains and has a high proline content. WIP is a functional homologue of verprolin because it suppresses the defects observed in  $\Delta vrp1$  and vrp1-1 cells; this verprolin-compensatory ability requires its WH2 and profilin-binding domains [37]. Overexpression of WIP in mammalian cells increases F-actin content and induces the appearance of actin-containing projections on the cell surface; this also requires the presence of its WH2 domain [38]. Purified WIP directly interacts with N-WASP, G-actin and F-actin. WIP retards N-WASP/Cdc42-activated polymerization of actin mediated by the Arp2/3 complex and stabilizes actin filaments [39].

#### 2.5. Srv2 and CAP

The yeast Srv2p/CAP was identified as a component of an adenylyl cyclase-containing complex and as a gene required for Ras signaling [40,41]. Srv2p is a regulator of the actin cytoskeleton, as illustrated by the abnormal morphology

and actin cytoskeleton in budding yeast SRV2 knockouts. Biochemical studies showed that Srv2p and its mammalian homologues bind and sequester monomeric actin in vitro [42,43]. A proline-rich region in yeast Srv2p contains a consensus SH3-binding motif (PXXP), that is recognized by the SH3 domains of several proteins in vitro and is required for the localization of Srv2p to cortical actin patches [44]. A good candidate for a targeting protein is Abp1, which binds the proline-rich region of Srv2p through its SH3 domain [45]. Srv2p's actin-binding activity maps to its C-terminal region, but the protein motif responsible for actin monomer-binding was not identified. Our domain searches revealed a WH2 domain in Srv2p's C-terminal region (Figs. 1 and 2). This WH2 domain is highly conserved in C. elegans, D. melanogaster and mammalian homologues of Srv2p, suggesting that this region is important for the biological function of Srv2p/CAP.

# 2.6. Others

*D. melanogaster* p150-Spir protein contains an acidic region followed by two WH2 domains, and it has been proposed that it promotes actin filament assembly in a fashion similar to WASP/WAVE proteins [46]. The WH2 domain mediates p150-Spir's interaction with actin monomers, but its biochemical activities have not been characterized. *D. melanogaster* p150-Spir seems to function with actin and rho family GTPases in patterning the *Drosophila* oocyte. Defects found in *spir* mutants resemble those obtained with treatment with the actin polymerization inhibitor, cytochalasin D [47].

The BLAST and SMART searches revealed several unidentified proteins. Their WH2 domain sequences are most similar to the WASPs and WIPs (Figs. 2 and 3). They contain up to three WH2 domains and have molecular weights ranging from 40 to 150 kDa. Some of these proteins also contain ankyrin (ANK) repeats, which are known to mediate protein–protein interactions.

## 3. Structure and actin interfaces of the WH2 domain

Our knowledge of the structure and actin interfaces of the WH2 domains is from studies on thymosin  $\beta$ 4. The structure of thymosin  $\beta$ 4 has been analyzed by nuclear magnetic resonance and circular dichroism spectroscopy. Under physiological conditions, the N-terminal region (residues 5-16) forms an  $\alpha$ -helix, and the rest of the molecule is unstructured [48,49]. The binding of thymosin  $\beta$ 4 to an actin monomer results in an increase in non-random structure of the molecule [49]. In the preliminary model of the thymosin  $\beta$ 4/actin monomer complex, the N-terminal  $\alpha$ -helical region and the following loop of thymosin  $\beta$ 4 interact with subdomain-1 of the actin monomer, and the C-terminal half of thymosin B4 extends towards actin subdomain-2 [49]. The N-terminal helix is relatively well conserved in all WH2 domains (Fig. 2), suggesting that they all have similar  $\alpha$ -helical structures. In contrast, the C-terminal halves of WH2 domains are more variable and may therefore adopt different structures in solution or in contact with actin.

The actin interfaces of thymosin  $\beta$ 4 have been mapped by using chemically synthesized full-length thymosin  $\beta$ 4 variants. These studies showed that the N-terminal part has to adopt an  $\alpha$ -helical conformation for actin-binding, and that several residues play key roles in this binding [50,51]. Four of the critical actin-binding residues, Met7, Ile10, Leu18 and Lys19, are conserved in all WH2 domains (Fig. 2), suggesting that WH2 domains interact with actin through a similar interface. Two other critical actin-binding residues in thymosin  $\beta$ 4, Phe13 and Lys15, are conserved only in the  $\beta$ -thymosin family: these residues may account for  $\beta$ -thymosin's specific actin-binding features.

It is important to note that while  $\beta$ -thymosins inhibit the assembly of actin filaments, other WH2 domain-containing proteins characterized to date (WASP, WIP and ciboulot) promote the assembly of actin filaments at the barbed end. Boquet et al. [16] speculated that  $\beta$ -thymosin-specific residues, Glu9 and Lys12, may be important for the actin monomer sequestering function of thymosin  $\beta$ 4. However, mutations in these residues do not result in dramatic defects in the actin monomer sequestering activity of thymosin  $\beta$ 4 [50]. Interestingly, in  $\beta$ -thymosins there is another highly conserved residue, Lys15, which is not found in any other WH2 domain proteins (Fig. 2). Biochemical experiments showed that the replacement of this residue by alanine has a relatively small effect on actin monomer-binding, but results in a strong defect in actin monomer sequestering activity of thymosin  $\beta$ 4 [50]. Therefore, this residue is an attractive candidate for being responsible for the biochemical differences between β-thymosins and other WH2 domain proteins. However, other studies have demonstrated that also the C-terminal region of β-thymosins is involved in actin-binding, and that functional differences between  $\beta$ -thymosin family members are specified by the C-terminal variability [52]. It is possible that the biochemical differences between various WH2 domains result from differences in the N-terminal  $\alpha$ -helix as well as in the C-terminal variable region.

## 4. Evolution of WH2 domain

Three WH2 domain proteins, WASP, verprolin/WIP and Srv2/CAP, are present from yeast to mammals, and their domain organizations and positions are well conserved: these three classes were probably present in the common ancestor of yeast and animals. Interestingly, we did not find these three proteins in plants, and our BLAST searches from the nearly complete *Arabidopsis* genome did not reveal any WH2 domains. Therefore, the WH2 domain and WH2 domain proteins may have evolved after the divergence of the plant lineage.

Fig. 3 shows an unrooted phylogenetic tree of all known human, fly, worm and budding yeast WH2 domains. It is important to note that the WH2 domain is a rather small ( $\sim$ 35 amino acid) protein motif, so reliable phylogenetic analysis is difficult. However, WH2 domains from each of the known WH2 domain-containing protein families form independent branches, providing support for the relevance of this analysis. The WH2 domains of Srv2/CAP, verprolin/ WIP and WASP proteins (with three exceptions) from yeast to mammals are in the same branches of the tree, suggesting that a WH2 domain was present in the ancestral WASP, WIP and Srv2 proteins before the yeast and animal lineages diverged (Fig. 3). β-Thymosins and ciboulot-like proteins are not found in yeast, and they also form a separate branch in the phylogenetic tree: these proteins either evolved in the animal lineage or were lost in fungi. Ciboulot and B-thymosins are most homologous to the WH2 domains in verprolin/WIP. Perhaps these proteins evolved from one of verprolin's WH2 domains. Finally, it is important to note that although  $\beta$ -thymosins and ciboulot are in the same branch and have highly similar sequences, they are biochemically different [16].

#### 5. Conclusions

The analysis presented here demonstrates that WH2 domain is a highly exploited and evolutionarily conserved actin monomer-binding motif. Interestingly, different WH2 domains can either promote or prevent the assembly of actin monomers into filament. In the future, it will be important to elucidate the molecular details of WH2 domain–actin interactions and to identify the structural variations responsible for the biochemical differences between various members of this family. There are also several uncharacterized WH2 domain-containing proteins whose biological and biochemical properties have to be evaluated.

Acknowledgements: We thank Drs. Marleen Van Troys and Christophe Ampe for critical reading of the manuscript and for sharing their unpublished results, and Pauli Ojala for help with sequence analysis.

#### References

- Sheterline, P., Clayton, J. and Sparrow, J. (1998) Actin Protein Profile 2, 1–103.
- [2] Castresana, J. and Saraste, M. (1995) FEBS Lett. 374, 149-151.
- [3] Stradal, T., Kranewitter, W., Winder, S.J. and Gimona, M. (1998) FEBS Lett. 431, 134–137.
- [4] Lappalainen, P., Kessels, M.M., Cope, J.T.V. and Drubin, D.G. (1998) Mol. Biol. Cell 9, 1951–1959.
- [5] Weeds, A. and Maciver, S. (1993) Curr. Opin. Cell Biol. 5, 63-69.
- [6] Van Troys, M., Vandekerckhove, J. and Ampe, C. (1999) Biochim. Biophys. Acta 1448, 323–348.
- [7] Machesky, L.M., Insall, R.H. and Volkman, L.E. (2001) Trends Cell Biol. 11, 286–287.
- [8] Zalevsky, J., Grigorova, I. and Mulllins, D. (2001) J. Biol. Chem. 276, 3468–3475.
- [9] Nachmias, V.T. (1993) Curr. Opin. Cell Biol. 5, 56-62.
- [10] Carpintero, P., del Amo, F., Anadon, R. and Gomez-Marquez, J. (1996) FEBS Lett. 394, 103–106.
- [11] Bao, L., Loda, M., Janmey, P.A., Stewart, R., Anand-Apte, B. and Zetter, B.R. (1996) Nat. Med. 2, 1322–1328.
- [12] Gold, J.S., Bao, L., Ghoussoub, R., Zetter, B.R. and Rimm, M.D. (1997) Mod. Pathol. 10, 1106–1112.
- [13] Pantaloni, D. and Carlier, M.-F. (1993) Cell 75, 1007-1014.
- [14] Carlier, M.-F., Didry, D., Erk, I., Lepault, J., Van Troys, M., Vandekerckhove, J., Perelroizen, I., Yin, H., Doi, Y. and Pantaloni, D. (1996) J. Biol. Chem. 271, 9231–9239.
- [15] Sun, H.-Q., Kwiatkowska, K. and Yin, H.L. (1996) J. Biol. Chem. 271, 9223–9230.
- [16] Boquet, I., Boujemaa, R., Carlier, M.-F. and Préat, T. (2000) Cell 102, 797–808.
- [17] Pollard, T.D. and Cooper, J.A. (1984) Biochemistry 23, 6631-6641.
- [18] Vancompernolle, K., Vandekerchove, J., Bubb, M.R. and Korn, E.D. (1991) Biochemistry 266, 15427–15431.
- [19] Bubb, M.R. and Korn, E.D. (1995) Biochemistry 34, 3921-3926.
- [20] Symons, M., Derry, J.M., Karlak, B.J., Jiang, S., Lemahieu, V., McCormick, F., Francke, U. and Abo, A. (1996) Cell 84, 723– 734.

- [21] Miki, H., Miura, K. and Takenawa, T. (1996) EMBO J. 15, 5326–5335.
- [22] Rohatgi, R., Ma, L., Miki, H., Lopez, M., Kirchhausen, T., Takenawa, T. and Kirschner, M.W. (1999) Cell 97, 221–231.
- [23] Higgs, H. and Pollard, T.D. (2000) J. Cell Biol. 150, 1311–1320.
- [24] Li, R. (1997) J. Cell Biol. 136, 649-658.
- [25] Rohatgi, R., Ho, H.-Y.H. and Kirschner, M.W. (2000) J. Cell Biol. 150, 1299–1309.
- [26] Fukuoka, M., Suetsugu, S., Miki, H., Fukami, K., Endo, T. and Takenawa, T. (2001) J. Cell Biol. 152, 471–482.
- [27] Machesky, L.M., Mullins, R.D., Higgs, H.N., Kaiser, D.A., Blanchoin, L., May, R.C., Hall, M.E. and Pollard, T.D. (1999) Proc. Natl. Acad. Sci. USA 96, 3739–3744.
- [28] Machesky, L.M. and Insall, R.H. (1999) J. Cell Biol. 146, 267– 272.
- [29] Higgs, H., Blanchoin, L. and Pollard, T. (1999) Biochemistry 38, 15212–15222.
- [30] Miki, H., Suetsugu, S. and Takenawa, T. (1998) EMBO J. 17, 6932–6941.
- [31] Suetsugu, S., Miki, H. and Takenawa, T. (1999) Biochem. Biophys. Res. Commun. 260, 296–302.
- [32] Takenawa, T. and Miki, H. (2000) J. Cell Sci. 114, 1801–1809.
- [33] Yamaguchi, H., Miki, H., Suetsugu, S., Ma, L., Kirschner, M.W. and Takenawa, T. (2000) Proc. Natl. Acad. Sci. USA 97, 12631– 12636.
- [34] Donnelly, S.F.H., Pocklington, M.J., Pallotta, D. and Orr, E. (1993) Mol. Microbiol. 10, 585–596.
- [35] Vaduva, G., Martin, N.C. and Hopper, A.K. (1997) J. Cell Biol. 139, 1821–1833.
- [36] Naqvi, S.N., Zahn, R., Mitchell, D.A., Stevenson, B.J. and Munn, A.L. (1998) Curr. Biol. 8, 959–962.
- [37] Vaduva, G., Martinez-Quiles, N., Anton, I.M., Martin, N.C., Geha, R.F., Hopper, A.K. and Ramesh, N. (1999) J. Biol. Chem. 274, 17103–17108.
- [38] Ramesh, N., Anton, I.M., Hartwig, J.H. and Geha, R.F. (1997) Proc. Natl. Acad. Sci. USA 94, 14671–14676.
- [39] Martines-Quiles, N., Rohatgi, R., Anton, I.M., Medina, M., Saville, S.P., Miki, H., Yamaguchi, H., Takenawa, T., Hartwig, J.H., Geha, R.S. and Ramesh, N. (2001) Nat. Cell Biol. 3, 484–491.
- [40] Field, J., Vojtek, A., Ballester, R., Bolger, G., Colicelli, J., Ferguson, K., Gerst, J., Kataoka, T., Michaeli, T., Powers, S., Riggs, M., Rodgers, L., Wieland, I. and Wigler, M. (1990) Cell 61, 319–327.
- [41] Fedor-Chaiken, M., Deschenes, R.J. and Broach, J.R. (1990) Cell 61, 329–340.
- [42] Freeman, N.L., Chen, Z., Horenstein, J., Weber, A. and Field, J. (1995) J. Biol. Chem. 270, 5680–5685.
- [43] Zelicof, A., Protopopov, V., David, D., Lin, X.-Y., Lustgartens, V. and Gerst, J.E. (1996) J. Biol. Chem. 271, 18246–18252.
- [44] Freeman, N.L., Lila, T., Mintzer, K.A., Chen, Z., Pahk, A.J., Ren, R., Drubin, D.G. and Field, J. (1996) Mol. Cell. Biol. 16, 548–556.
- [45] Lila, T. and Drubin, D.G. (1997) Mol. Biol. Cell 8, 367-385.
- [46] Otto, I.M., Raabe, T., Rennefahrt, U.E.E., Bork, P., Rapp, U.R. and Kerkhoff, E. (2000) Curr. Biol. 10, 345–348.
- [47] Wellington, A., Emmons, S., James, B., Calley, J., Grover, M., Tolias, P. and Manseau, L. (1999) Development 126, 5267–5274.
- [48] Czisch, M., Schleigher, M., Horger, S., Voelter, W. and Holak, T.A. (1993) Eur. J. Biochem. 218, 335–344.
- [49] Safer, D., Sosnick, T.R. and Elzinga, M. (1997) Biochemistry 36, 5806–5816.
- [50] Van Troys, M., Dewitte, D., Goethals, M., Carlier, M.-F., Vandekerckhove, J. and Ampe, C. (1996) EMBO J. 15, 201–210.
- [51] Simenel, C., Van Troys, M., Vandekerckhove, J., Ampe, C. and Delepierre, M. (2000) Eur. J. Biochem. 267, 3530–3538.
- [52] Eadie, J.S., Kin, S.W., Allen, P.G., Hutchison, L.M., Kantor, J.D. and Zetter, B.R. (2000) J. Cell. Biochem. 77, 277–287.