

Minireview

SH3 – an abundant protein domain in search of a function

Andrea Musacchio^a, Toby Gibson^a, Veli-Pekka Lehto^b and Matti Saraste^a

^aEuropean Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900 Heidelberg, Germany and ^bBiocenter and Department of Pathology, University of Oulu, 90220 Oulu, Finland

Received 5 May 1992

Src-homology 3 is a small protein domain of about 60 amino acid residues. It is probably made of β -sheets. SH3 is present in a large number of eukaryotic proteins which are involved in signal transduction, cell polarization and membrane-cytoskeleton interactions. Here we review its occurrence and discuss possible functions of this domain.

SH3; Structural prediction; Occurrence; Function

1. INTRODUCTION

SH3 (Src homology region 3) is a small protein domain containing about 60 amino acid residues. It was first identified as a conserved sequence in the N-terminal non-catalytic part of the *src* protein tyrosine kinases [1]. Src-kinases participate in signal transduction pathways originating at the receptor tyrosine kinases, such as the platelet derived growth factor receptor (PDGFR). A number of other proteins which are involved in these pathways, have also been shown to contain SH3 domains. These include phospholipase C γ [2], phosphatidylinositol-3-kinase (PI-3 kinase, e.g. [3]) and the p21^{ras} GTPase activating protein (GAP, [4]). Another common non-catalytic domain of about 100 or more amino acid residues in these proteins is called SH2. One of its functions is to bind to peptides containing phosphotyrosine (see [5] and below).

SH3 is not exclusively characteristic of signalling proteins. Both in single-cell and higher eukaryotes SH3 is present in a very large group of proteins, among them families which are otherwise unrelated to each other. These include erythroid and non-erythroid α -spectrins [6,7], two unconventional myosins in *Acanthamoeba* [8], the p47 and p67 cytosolic activators of the human neutrophil respiratory burst oxidase [9,10], and putative transcriptional regulators Vav and Hs1 (e.g. [11,12]). In yeast, SH3 is found in proteins involved in morphogenesis, such as Bem-1 [13,14] and the actin-binding protein Abp1 [15,16], proteins involved in mating (Fus1 [17,18]), and in regulation of ras activity (Cdc25 and Ste6 [19,20]). Recently an SH3 domain has also been found in a 55 kDa protein which is a major palmitoyl-

ated erythrocyte membrane protein [21], and in Dgl, a homologous protein in *Drosophila* [22]. A gene (*sem5*) involved in development of nematode vulva encodes a protein that has two SH3 domains [23].

So far the function of SH3 has remained unclear. Rodaway and co-workers [24] have drawn attention to the fact that most of the SH3-containing proteins are associated with membranes. They suggested that SH3 may have a role in subcellular localization, probably through interaction with components of the cortical cytoskeleton. Several lines of evidence suggest that SH3 is involved in establishing or regulating such protein-protein interactions which may often occur close to the plasma membrane. The presence of SH3 domains in proteins involved in morphogenesis and signal transduction clearly indicates such a role. On the other hand, it has been shown that the SH2 domain of viral Src is essential for its association with detergent-insoluble cellular matrix [25]. It has also been suggested that SH3 might mediate membrane-cytoskeletal interaction by binding to actin [16] but no conclusive evidence supports this idea.

SH3 has no fixed topological position in proteins (Fig. 1). It seems likely that the position of SH3 in a protein sequence is not important for its function. A function, such as binding to a target protein, is therefore entirely confined to this domain.

2. A SEQUENCE ALIGNMENT

Figure 2 shows an alignment of SH3 sequences. The alignment is limited to about 60 amino acid residues, which seem to constitute the SH3 core. This choice of domain boundaries is justified by two facts. First, the SH3 domain in Ste6 and the first SH3 domain in Nck are at the extreme N-terminus. Second, in Hs1, Abp1,

Correspondence address: A. Musacchio and M. Saraste, EMBL, Meyerhofstr. 1, D-6900 Heidelberg, Germany. Fax: (49) (6221) 387306.

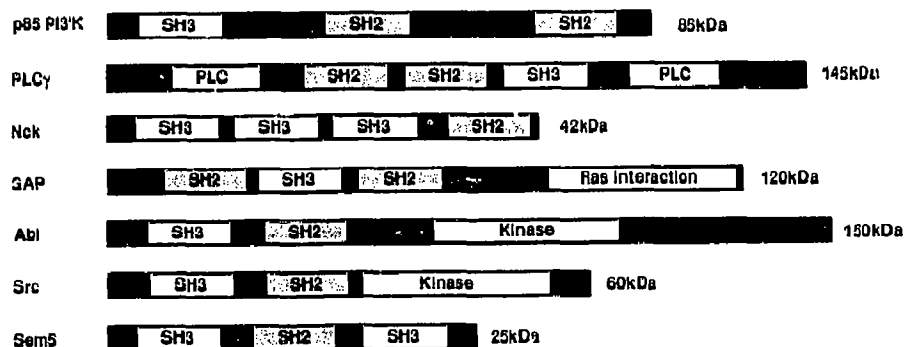


Fig. 1. SH2 and SH3 domains in selected proteins. PI3'K/p85 is a subunit of phosphatidylinositol-3'-kinase, PLC γ is phospholipase-C γ and GAP is p21^{GTP} GTPase activating protein. Nck, Abl and Src are oncogene products discussed in text. Sem5 is a protein involved in vulval development of *C. elegans*.

p80/85 (which a substrate of Src kinase [26]), and in one *Acanthamoeba* myosin-I SH3 is located precisely at the C-terminus (see Fig. 2).

The highest degree of conservation within the SH3 domain is localised at the N- and C-termini and in some parts of the central region. Some conserved residues form patterns which may help to define the secondary structure of the SH3 domain. This is especially evident in the central region of the sequence, where the patterns of hydrophobic and hydrophilic amino acids suggest formation of β -strands. Three β -strands may correspond to residues 17–20, 25–30 and 35–42 (Fig. 2). These seem to be connected to loop regions, corresponding to amino acids 13–16, 21–24, and 31–34. The first and the latter show a low degree of conservation, while the second putative loop region is rather conserved and shows a preference for charged residues. The N-terminal residues 4, 6, 8 and 10 are all more highly conserved than the alternating residues 3, 5, 7 and 9, implying the presence of a long β -strand. Part of this stretch is especially conserved and forms the characteristic ALYDY motif of SH3. The C-terminal region has also a regular pattern of hydrophilic and hydrophobic amino acids, which may be compatible with formation of one or two further strands.

Prediction of secondary structure can be assisted by the distribution of glycine and proline residues. Figure 2 shows regions where glycine and proline are poorly tolerated and which may correspond to β -strands. Glycine residues are most often present in the putative loop regions. There are several conserved glycine positions (23, 34, and 45 following the numbering in Fig. 2). G48 is almost invariant, together with P51.

We suggest that the SH3 domain folds into an all β -structure, with 5 or 6 strands connected by loops. Circular dichroic spectrum determined for the isolated SH3 motif of a non-erythroid α -spectrin purified after expression in *E.coli* indicates a high content of β -structure (around 70%) in agreement with this suggestion (A. M. and M. S., unpublished).

Most of the conserved amino acids seem to be important for the secondary structure. Very few residues in the predicted loops are strictly conserved. The surface loops could be important in binding to other proteins. A number of mutations in the SH3 domain of Src, which convert it into a transforming protein, have been described [27,28]. These mutations include substitutions of both Y8 and Y10 with phenylalanines and substitutions which affect R13 and T14. They fall into a loop region and could perturb a protein binding site. How-

Fig. 2. An alignment of SH3 sequences. This alignment was generated with the Genetic Data Environment colour alignment editor (S. Smith, Harvard University) and the colour masking program COLORMASK (J. Thompson, EMBL) emphasizing conserved features. Symbols used for the consensus (bottom lines) are: %, >50% hydrophobic residues; #, >85% hydrophobic residues; +, >60% K+R; -, >60% D+E. Capitals indicate residues with >85% conservation. The sequences are coloured to illustrate structurally important features. All G (orange) and P (yellow) residues are coloured. Columns are coloured by most conserved feature. For >50% occurrence of a residue property, the matching residues are coloured according to the following scheme: hydrophobic = blue, hydrophobic tendency = light blue, basic = red, acidic = purple, Q or N and S or T = green. Columns which are left white show poor conservation of residue or property. The single letter code in lower key is used for amino acids. The sequences were taken from the SwissProt Database (release #20). SPC-A denotes α -spectrin from chicken (C) and fruit fly (Dm); MIL-B and -C are cytoplasmic myosins from *Acanthamoeba* (Ae). Human (H), mouse (M), rabbit (R), chicken, bull frog (Xl) or fruit fly (proto)oncogene products FYN, VAV, FGR, SRC, YES, NCK, GAG-CRK, LYN, HCK, BLK, LSK, ABL, DLG, TEC, CSK and TKL are of viral (v-) or cellular (c- or no prefix) origin. ASV, RSV and FSF are avian, Rous and feline sarcoma viruses, respectively. HSI is a nuclear protein; NCF1 and NCF2 are neutrophil cytoplasmic factors; human PLC and PLC1 as well as rabbit and bovine (B) PLC-II are phospholipase C's; GAP is ras-GTPase activating protein; P55 is a major palmitoylated erythrocyte membrane protein; P80/85 is a substrate of Src kinase; P85A and B are subunits of PI-3-kinase. These, and the yeast (Y, Sp) proteins ABP1, BEM1, CDC25, FUS1 and STE6 are discussed in the text where the references are found. Multiple copies of SH3 are indicated with #1,2,3. Residue numbering is shown on the top. Stars (*) indicate N- and C-termini.

ruler	1	10	20	30	40	50	60
ASV v-SRC	..ggvttfvaly	yesrtetdl	sf..kkgarlqivnntgwwlahslttgqgt	gyipsnyvapsds..
RSV v-SRC	..ggvttfvaly	yeswtetdl	sf..kkgarlqivnntgwwlahslttgqgt	gyipsnyvapsds..
H c-SRC-1	..ggvttfvaly	yesrtetdl	sf..kkgarlqivnntgwwlahslstgqgt	gyipsnyvapsds..
X1 c-SRC-1	..ggvttfvaly	yesrtetdl	sf..kkgarlqivnntgwwlahslssgqgt	gyipsnyvapsds..
C c-SRC	..ggvttfvaly	yesrtetdl	sf..kkgarlqivnntgwwlahslttgqgt	gyipsnyvapsds..
M n-SRC	..ggvttfvaly	yesrtetdl	sf..kkgarlqivnnt	rkvrregwwlahslstgqgt	gyipsnyvapsds..
X1 c-SRC-2	..ggvttfvaly	yesrtetdl	sf..kkgarlqivnntgwwlahslssgqgt	gyipsnyvapsds..
ASV v-YES	..ggvttfvaly	yearttddl	sf..kkgarfqiinntgwwearslatgkt	gyipsnyvapads..
C c-YES	..ggvttfvaly	yearttddl	sf..kkgarfqiinntgwwearslatgkt	gyipsnyvapads..
H c-YES-1	..ggvttfvaly	yearttddl	sf..kkgarfqiinntgwwearslatgkt	gyipsnyvapads..
X1 c-YES	..ggvttfvaly	yearttddl	sf..kkgarfqiinntgwwearslatgkt	gyipsnyvapads..
X1 c-FYN	..tgvtlfvaly	yearteddl	sf..kkgkfqiilnssgwwearslttgqgt	gyipsnyvapads..
H c-FYN	..tgvtlfvaly	yearteddl	sf..kkgkfqiilnssgwwearslttgqgt	gyipsnyvapads..
M c-FGR	..tgvtlfvaly	yearteddl	sf..kkgkfqiilnntgwwearslssghr	gyipsnyvapads..
H c-FGR	..igvtlfialy	yearteddl	sf..kkgkfqiilnntgwwearslssgkt	gyipsnyvapads..
Ha STK	..pgvttfvaly	yeariseal	sf..kkgarlqiintagwwyarslitnse	gyipstyvapeks..
H HCK	..sedilvvaly	yeaihhedl	sf..kkgqmvvleesgwwkarslatrke	gyipsnyvarvds..
M HCK	..sedilvvaly	yeaihhedl	sf..kkgqmvvleesgwwkarslatkke	gyipsnyvarvds..
H LYN	..eqgdilvvaly	pydgihpdal	sf..kkgkmlvleehgwwkarslltkke	gyipsnyvakint..
M BLK	..eeerfvvalf	yaavndrdl	qv..lkglqlvirstgwwlarslvtg	gryvpsnfavpet..
M LSK-T	..lqdnlvialhsyepshdgdl	gf..ekgqlrileqsgwwkarslttgqgt	gyipfnfvakans..	
H LCK	..lqdnlvialhsyepshdgdl	gf..ekgqlrileqsgwwkarslttgqgt	gyipfnfvakans..	
FSV v-ABL	..ndpnlfvaly	ifvasgdntls	l..tkgklrvlgynhngwceagtkngggwvpsnyitpvn..	
H c-ABL	..ndpnlfvaly	ifvasgdntls	l..tkgklrvlgynhngwceagtkngggwvpsnyitpvn..	
M c-ABL	..ndpnlfvaly	ifvasgdntls	l..tkgklrvlgynhngwceagtkngggwvpsnyitpvn..	
Dm ABL-1	..ddpqlfvaly	ifgaggengls	l..kkgqvrilsynksgwwceahssngv	gryvpsnyvtpplns..
C c-TKL	..lqdklvvaly	yeepthdgdl	gf..kkgklrvleesgwwraqlttgqgt	gyipfnfvamvns..
Dm SRC-1	..vlkrvvsly	fyksrdesdl	sf..mkggmveiddtswwrvnltrq	gqglplnfvaers..
ASV GAG-CR	..eeveyvralf	fykgndddgl	fpf..kkgdkilrkdkseqwmaedmdgr	krmpvpyvekfrp..
C Spc-a	..tgkclvialy	fygeksprevtm	..kkgdiltllnsnkwwkvevndrq	gfpaayvkkldp..
Dm Spc-a	..tgkclvialy	fyteksprevtm	..kkgdiltllnsnkwwkvevndrq	gfpaayvkkldp..
H PLC	..mpqrtvkalysy	fykakrsdels	sf..rcgalihnyvskpgwwkcdytri	qgyipsnyvedist..
R PLC-II	..tfkcalvkalysy	fykaqredelt	sf..tkksailqnvekggwwrgdygkq	qlwfpnyveemin..
B PLC-II	..tfkcalvkalysy	fykaqredelt	sf..tkksailqnvekggwwrgdygkq	qlwfpnyveemin..
H PLC1	..tfkcalvkalysy	fykaqredelt	sf..tkksailqnvekggwwrgdygkq	qlwfpnyveemin..
RASA/GAP	..edrrrvrallpytkvptdte	dl..lkggmivhneldgmwvtnlrtde	gqlive	lveevgr..
B GAP	..edrrrvrallpytkvptdte	dl..lkggmivhneldgmwvtnlrtde	gqlive	lveevgr..
Ac MIL-B	..pakpqvkalysy	fydaqtgdelt	sf..kegdtlihvqkdpagwweelngkrgwvpsnyvqdi*	
Ac MIL-C	..pgpegaraly	fyasenpdelt	sf..negavtvinksnpwwegelngqrgvpsnyvelipr..	
H HS1	..algsavaly	fyqgegsdels	sf..dpdvatdiemvegwrgchghf	glfpanyvklle*
H VAV	..kyfgtakary	fycardrsels	sf..kegdilklknkqqwwrgciygrv	gwpfanyveedys..
M TEC	..nteeivvamy	fygateahdlr	l..lrgqeyilleknlhwrrardkygy	grntnrskae..
Dm SRC-2	..hfvklvvaly	fykaieggdl	sf..geknaeyeviddsqehwkvkdlgn	vgypnyvqaeal..
R CSK	..psgterialkyn	fyhgtaeqdl	fpf..ckgqvitavtkdnpnykankv	regilpanyvqkre
H NCK/1	*maeevvvvaly	fyvaqqeqald	l..kknerrwllddskswrvnsmnkt	gfpsnyverks..
H NCK/2	..dlmnpayvkinymeredels	sf..ikgtkvivmekcsdwwrgsyngqrgwvpsnyvteed..		
H NCK/3	..qvlhvvaly	fyfsssndeeln	sf..ekgqvmviekpndpwwkcklngm	vlpknyvvtmqn..
H NCF1/1	..ilqttyralanyektsqsemal	..stgavveevkssgwffcqmakrgwvpsnyvteed..		
H NCF1/2	..yagcpvvaly	fykavegdeval	..legeavevikhlldwwvirkddvt	gyfpsnyvteedcat..
H NCF2/1	..legeahrvl	fyfvpetkeelq	v..mpgnivfvlkkndnwatvmfngkq	glvpcnyvteedcat..
H NCF2/2	..kkgqsvealfsytqpedel	sf..qegailvlskvnewlegeckgkvgifpkvteedcat..		
Y ABP1	..kenpwatvey	fydaaednelt	sf..vendklniefvdwwlgelkdsk	glfpsnyvteedcat*
Y BEM1/1	..spekvikakysyqtskals	sf..megaffvysgdekwykaspstke	gvpvptyfevdr..	
Y BEM1/2	..mgslyalvly	fykaekadelt	sf..yvgnelficahncwfiaaplg	rlgppdlvpyvtsidi
Y FUS1	..qlgktytviq	fyeprltdelr	l..ylgskvilathtdgwclvekntq	gslhvsyddkryl..
Y CDC25	..rpigivvaay	fyfnpikksasqllsvqqggt	fyilnknssgwddglviddsngkvnrgwvpsnyvteedcat..		
OC CAC-b	..afavrtvayvpspdevpve	gvaltf..epkafihikekynnwwigrlvkeg	cevgfipsnyvteedcat..	
Dm DLG	..krslyvralfydpnddglpsrclpf	..khgdilhvtnasdwwqrrylgdnede	gippskrvdkrm..	
H P55	..alqmfnraqfydpkkdnlipckasglkatgdlilqinkdsnwqgrvegsskesaglipspelpqerv..		
C P80/85	..elcditaly	fyqgagddels	sf..dpddiltntiemdwwrgvckgry	glfpanyvteedcat..
B P85A	..pegfayraly	fyfrrerpedel	l..lpgdvlvvraalqalgvaegnercpqsgvmp	glnerttrq..
B P85B	..aegyqyraly	fykkereedid	l..hlodiltvknkslvalgfsdqqekpee	gln.gynettge..
Sp STE6	*mrfqttalysyssnpsflkfsagdtli	fyevldwwgdcicsekrgwvpsnyvteedcat..	
CONS.%Alydy.a-lftf	..kg-#.#.#.g-Ww.a..%g.g.#Psnv..#.g.g.#Psnv..#.g.g.#Psnv..#.
pro	ppppppp.p	pppppp.p	ppppppp.p	ppppp
gly	gggggg.gg.g	gggggg.g	gggggg.g	ggggg

ever, mutations in conserved residues may also cause gross structural alterations, which indirectly abolish function.

Y8 and Y10 are not normally phosphorylated *in vivo*. However, an activated form of Src has probably been found to be phosphorylated on tyrosine(s) in its N-terminal part [29], most likely on one or both of these

residues. Although this suggests that Y8 and Y10 may be potential sites for regulation of the SH3 function in Src [30], these tyrosines are not strictly invariant in SH3 and are often substituted with phenylalanines (Fig. 2).

The low similarity between SH3 sequences has also created confusion in the identification of putative SH3 domains. For example, an SH3 domain has been re-

ported in the IRA1 protein of *Saccharomyces cerevisiae* (see [5]) but this sequence does not fit into our alignment. The consensus sequence of Fig. 2, in conjunction with the pattern of glycines and prolines in it, may be a good guide for identification and future classification of SH3 sequences.

3. SH3 IN YEAST

Four SH3-containing proteins have been described in the budding yeast *Saccharomyces cerevisiae*. Two of these, Bem-1 and Abp-1, have a role in budding [13–16]. Fus-1 is an essential factor for cell fusion during mating [17,18]. Cdc25 is a guanine nucleotide exchange factor, which is thought to regulate the Ras/adenylate cyclase pathway [19]. So far, no SH2 sequences have been detected in yeast.

In budding yeast, both cell growth and mating are polarised phenomena, which imply definition of cell axis. During cell growth, microtubules and actin cables define an axis along which cellular materials are transported towards the growing area of the membrane [31,32]. In the bud, actin forms patches in the cortical region. Abp1, an actin-binding protein, is exclusively localised to these structures. Its function is related to formation of actin patches: overexpression of Abp1 leads to formation of patches in abnormal areas such as the cortical region of the mother cell [15]. The same phenotype is observed upon disruption of other genes, which have been classified as 'bud site assembly' genes. These include *cdc24*, *cdc42*, and *cdc43* and encode a putative calcium-binding protein, a ras-like GTP-binding protein and a factor involved in C-terminal prenylation of small GTP-binding proteins, respectively [31].

Small GTP-binding proteins appear also to be important for the localization of the bud site. The localization depends on the orientation of the axis in the previous cell division. Mutations in five genes (*bud1–5*) lead to a random budding pattern [31,33]. Bud1 (also known as Rsr1) is a small GTP-binding protein [34]. *bud5* encodes a guanine nucleotide exchange factor whose catalytic domain is related to the catalytic domain of Cdc25.

Genetic studies have established that Bud5 interacts with Bem1 [13], an SH3-containing protein, which is encoded by one of the 'bud site assembly' genes. The localization of budding site is a signal for a class of proteins produced by these genes. Their function is to limit surface growth to this area. Bud site assembly involves a specific organization of actin and tubulin both in the mother cell and in the bud. This organization is thought to drive organelle transport to the site of assembly, and in this way promote surface growth [31]. Interplay between bud-site-assembly gene products and cytoskeletal elements is necessary for polarising yeast cells.

It would be interesting to know if any SH3-containing proteins have a comparable function in development of

cell polarity in higher eukaryotes. Woods and Bryant [22] have recently shown that the product of the *dgl* tumor suppressor gene in *Drosophila*, an SH3-containing protein which is localized at septate junctions, is involved in development of apical–basal polarity in cells of imaginal discs. Dgl is homologous to a 55 kDa SH3-containing protein of the erythrocyte membrane. Both proteins have a domain homologous to the yeast guanylate kinase suggesting that Dgl and p55 may have guanylate kinase activity [35].

4. SH3 IN SIGNAL TRANSDUCTION

PDGFR and epidermal growth factor receptor (EGFR) belong to a family of transmembrane receptors, whose cytoplasmic moiety has intrinsic tyrosine kinase activity [36,37]. Many details of the molecular mechanisms by which signal transduction through these receptors takes place, have been elucidated in recent years. Upon induction with PDGF, for example, a number of proteins which include Src (or a src-related kinase), PLC γ , PI-3-kinase and GAP, become physically associated with the activated receptor (see [5]).

SH2 is a noncatalytic domain common to many substrates of the activated receptors, and may play an essential role in the formation of the transducing complex. Upon induction with the appropriate growth factor, receptors of the tyrosine kinase family are thought to dimerise and autophosphorylate on tyrosine residues. SH2 domains recognize, and bind to, phosphorylated tyrosine residues which are within a specific amino acid sequence. Phosphotyrosine residues constitute the signal which promotes association of a number of signalling proteins via their SH2 domains to the receptor [38–40]. These proteins are then phosphorylated by tyrosine kinase activity of the receptor. The association to the receptors is important for activation of the intracellular steps in signal transduction, which lead to changes in gene expression, cellular metabolism, and cytoskeletal architecture [41].

In contrast, the role of SH3 in signal transduction remains unclear. Most of the available information comes from the analysis of mutations in the SH3 domains of different signalling proteins. Several mutations of Src, which cause morphological alterations and cellular transformation due to uncontrolled kinase activity, have been isolated (e.g. [28,42]). Some of these mutations are in the catalytic domain, whereas others map into the non-catalytic regions, including the SH3 domain. Deletions of the SH3 domain in Src also cause cellular transformation. Similarly, Jackson and Baltimore [43] have reported that deletions in the SH3 domain of the c-Abl cytoplasmic tyrosine kinase induce a transformed phenotype when they are introduced into fibroblasts.

It is widely accepted that abnormal tyrosine phosphorylation caused by these mutations leads to cellular

transformation [27,28,42], but mechanisms by which this could happen have not been fully elucidated. In any case, these observations suggest that SH3 may be involved in down-regulation of the Src tyrosine kinase activity [27,44]. One can imagine at least three ways by which SH3 deletions or mutations may lead to a failure to regulate the kinase activity. First, SH3 may inhibit the kinase activity through an intramolecular mechanism, for example by binding to the kinase domain. In this case, deletion of SH3 may cause increased phosphorylation of normal substrates as well as phosphorylation of new, abnormal substrates. The current model is, however, that SH2 is the main intramolecular regulator of the Src kinase activity [45]. And it is evident that the sole function of SH3 cannot be the intramolecular regulation of kinase activity, because it is present in proteins that have no catalytic kinase domain (Fig. 1).

Secondly, SH3 may bind to another molecule which is able to switch the kinase activity on and off in response to appropriate stimuli. Small GTP-binding proteins could have such a function. They are apparently able to regulate other proteins' activity through conformational rearrangements [46].

The third possibility is that an SH3 deletion may impair the ability of kinase to choose right substrates, which would lead to phosphorylation of wrong ones. A consequence of such a model is that it should be possible to see changes in the pattern of proteins phosphorylated by Src after deleting SH3. Brugge and coworkers [42] have recently shown that deletions in the SH3 domain of Src lead to an increase in tyrosine phosphorylation. However, the phosphorylation pattern of cellular proteins by the transforming forms of Src with an intact SH3 domain is similar to the one observed after deletion of SH3. This suggests that SH3 does not regulate catalytic activity of Src via substrate recognition but via a more direct effect on the kinase activity itself. As pointed out, this may occur either through an intramolecular mechanism or through interaction with another regulatory protein. In the latter case, mutations in a gene encoding such a regulatory protein would also cause cellular transformation [43].

Development of experimental approaches which combine genetics and biochemical techniques give some hope in unravelling the impressively complex network of interactions between signalling, cytoskeletal, and regulatory proteins in signalling pathways. Organisms like *Drosophila* and *C. elegans* seem to be ideal systems for such experiments.

Recently, Clark and coworkers [23] have shown that mutations in both SH3-coding regions of the *sem5* gene can affect vulval development and sex myoblast migration during development of *C. elegans*. The product of *sem5* is a 228 amino acid protein, which contains only two SH3 and one SH2 domain with no evident catalytic domain (Fig. 1). In the latter respect mammalian proteins such as the p85 subunit of PI-3-kinase, Nck, v-Crk,

and c-Crk are comparable to Sem5, and also involved in signal transduction [5]. The lack of catalytic activity suggests that these proteins may act as adaptors, creating a link between phosphotyrosine containing proteins (via SH2) and specific classes of substrates or regulatory proteins (via SH3 [47]). Two other genes are known to control signal transduction in vulval cells of *C. elegans*. The product of the *let23* gene is a transmembrane tyrosine kinase homologous to the EGFR, and Let60 is a ras-like GTP-binding protein. Genetic experiments indicate that Sem5 and Let23 act before Let60 in vulval induction. It is possible that the SH2 domain of Sem5 promotes binding to phosphotyrosine residues on the Let23 protein, which leads to activation of the *let60* gene product [23]. In this case, SH3 could have the important function of selecting a target – maybe a GDP exchange factor – which is able to influence the activity of Let60 protein. Mutations in both SH3 domains of Sem5 have a deleterious effect on its function. This suggests that these two domains are functionally not redundant. Hopefully, further genetic analysis will lead to identification of the functions involved in this signalling process downstream of Sem5 activity (see [23,47]).

5. CONCLUSIONS AND HYPOTHESES

The long evolutionary history of SH3 and its presence in diverse proteins must reflect the fact that SH3 carries out a vital function in different cellular processes. The molecular mechanisms through which this domain acts, have not yet been elucidated. The most urgent problem is to identify proteins that bind to SH3 with high affinity. The fact that SH3 can regulate the activity of different protein tyrosine kinases makes this objective particularly interesting. Such studies might lead to identification of other proteins involved in the control of cell differentiation and proliferation.

Sequence analysis (Fig. 2) shows that SH3 has probably a β -sheet structure. Localization of possible binding sites for protein targets is complicated by the lack of extensive homology among different SH3 sequences. This evolutionary divergence becomes meaningful, if the proteins that bind to SH3 are also divergent but evolutionarily related. Some hypotheses have been made on the SH3 target. None of them seems to be conclusive. As most of the SH3-containing proteins are localised to the cortical region of the cell, actin has been indicated as a potential target. Moreover, some proteins which are known to bind actin, contain an SH3 domain. These include Abp1, cytoplasmic myosins in amoeban cells, and spectrin. The structures of the Abp1 protein and unconventional amoeban myosins are quite similar [46]. Their SH3 domains are located at the C-terminus, and similarity in their amino acid sequences extends towards the N-termini [16]. A 30 kDa C-terminal fragment of *Acanthamoeba* myosin-I contains an ATP-inde-

pendent actin-binding site [49]. The SH3 domain in spectrin is, in contrast, in the middle of the α -chain while its (main) actin-binding domain is at the N-terminus of the β -chain [7,50]. Other proteins, like ezrin and calpactin, have been indicated as possible ligands of the spectrin SH3 [16,51]. Calpactin is a potential substrate of Src, and moreover colocalised with spectrin [16,52].

Jackson and Baltimore [43] have included G-proteins as possible targets of the SH3 domain in the Abl tyrosine kinase. This hypothesis is intriguing, because it postulates the existence of a network of interactions among apparently separate signal transduction pathways. The localization of many G-proteins at the plasma membrane also makes this hypothesis coherent from a topological point of view.

The possibility that SH3 domains in diverse proteins bind to different small GTP-binding proteins is interesting. Small GTP-binding proteins take part in different cellular processes such as protein transport and sorting, cell differentiation and proliferation, organization of cytoskeleton and protein synthesis [46]. This functional diversity would match with the presence of SH3 domains in cytoskeletal and signalling proteins as well as in proteins involved in cell polarization. The p47 and p67 cytosolic components of the human neutrophil respiratory burst oxidase have been partially purified using GTP-agarose. This observation suggests that complexes are formed between these proteins, which both contain a SH3 domain, and (unidentified) GTP-binding proteins [53,54].

Acknowledgement: We thank Sara Courtneidge for critical reading of the manuscript.

REFERENCES

- [1] Meyer, B.J., Hamaguchi, M. and Hanafusa, H. (1988) *Nature* 332, 272-275.
- [2] Stahl, M.L., Ferenz, C.R., Kelleher, K.L., Kriz, R.W. and Knopf, J.L. (1988) *Nature* 332, 269-272.
- [3] Otsu, M., Hiles, I., Gout, I., Fry, M.J., Ruiz-Larrea, F., Panayotou, G., Thompson, A., Dhand, R., Hsuan, J., Totty, N., Smith, A.D., Morgan, S.J., Courtneidge, S.A., Parker, P.J. and Waterfield, M.D. (1991) *Cell* 65, 91-104.
- [4] Trahey, M., Wong, G., Halenbeck, R., Rubinfeld, B., Martin, G.A., Ladner, M., Long, C.M., Crosier, W.J., Watt, K., Kothe, K. and McCormick, F. (1988) *Science* 242, 1697-1700.
- [5] Koch, C.A., Anderson, D., Moran, M.F., Ellis, C. and Pawson, T. (1991) *Science* 252, 668-674.
- [6] Sahr, K.E., Laurila, P., Kotula, L., Scarpa, A.L., Coupal, E., Leto, T.L., Linnenbach, A.J., Winkelmann, J.C., Speicher, D.W., Marchesi, V.T., Curtis, P.J. and Forget, B.G. (1990) *J. Biol. Chem.* 265, 4434-4443.
- [7] Wasenius, V.-M., Saraste, M., Salven, P., Erämaa, M., Holm, L. and Lehto, V.-P. (1989) *J. Cell Biol.* 108, 79-93.
- [8] Jung, G., Korn, E.D. and Hammer III, J.A. (1987) *Proc. Natl. Acad. Sci. USA* 84, 6720-6724.
- [9] Volpp, B.D., Nauseef, W.M., Donelson, J.E., Moser, D.R. and Clark, R.A. (1989) *Proc. Natl. Acad. Sci. USA* 86, 7195-7199.
- [10] Leto, T.L., Lomax, K.J., Volpp, B.D., Nunoi, H., Sechler, J.M.G., Nauseef, W.M., Clark, R.A., Gallin, J.I. and Malech, H.L. (1990) *Science* 248, 727-730.
- [11] Bustelo, X.R., Ledbetter, J.A. and Barbacid, M. (1992) *Nature* 356, 68-71.
- [12] Kitamura, D., Kaneko, H., Miyagoe, Y., Ariyasu, T. and Watanabe, T. (1989) *Nucleic Acid Res.* 17, 9367-9379.
- [13] Chant, J., Corrado, K., Pringle, J.R. and Herskowitz, I. (1991) *Cell* 65, 1213-1224.
- [14] Chenevert, J., Corrado, K., Bender, A., Pringle, J. and Herskowitz, I. (1992) *Nature* 356, 77-79.
- [15] Drubin, D.G., Miller, K.G. and Botstein, D.J. (1988) *Cell Biol.* 107, 2551-2561.
- [16] Drubin, D.G., Mulholland, Z. and Botstein, D. (1990) *Nature* 343, 288-290.
- [17] McCaffrey, G., Clay, F.J., Kelsay, K. and Sprague Jr., G.F. (1987) *Mol. Cell. Biol.* 7, 2680-2690.
- [18] Trueheart, J., Boeke, J.D. and Fink, G.R. (1987) *Mol. Cell. Biol.* 7, 2316-2328.
- [19] Broek, D., Toda, T., Michaeli, T., Levin, L., Birchmeier, C., Zoller, M., Powers, S. and Wigler, M. (1987) *Cell* 48, 789-799.
- [20] Hughes, D.A., Fukui, Y. and Yamamoto, M. (1990) *Nature* 344, 355-357.
- [21] Ruff, P., Speicher, D.W. and Husain-Chishti, A. (1991) *Proc. Natl. Acad. Sci. USA* 88, 6595-6599.
- [22] Woods, D.F. and Bryant, P.J. (1991) *Cell* 66, 451-464.
- [23] Clark, S.G., Stern, M.J. and Horvitz, H.R. (1992) *Nature* 356, 340-344.
- [24] Rodaway, A.R.F., Sternberg, M.J.E. and Bentley, D.L. (1989) *Nature* 342, 624.
- [25] Fukui, Y., O'Brien, M.C. and Hanafusa, H. (1991) *Mol. Cell. Biol.* 11, 1207-1213.
- [26] Wu, H., Reynolds, H.B., Kanner, S.B., Vines, R.R. and Parsons, J.T. (1991) *Mol. Cell. Biol.* 11, 5113-5124.
- [27] Potts, W.M., Reynolds, A.B., Lansing, T.J. and Parson, J.T. (1988) *Oncogene Res.* 3, 343-355.
- [28] Kato, J., Takeya, T., Grandori, C., Iba, H., Levy, J.B. and Hanafusa, H. (1986) *Mol. Cell. Biol.* 6, 4155-4160.
- [29] Yonemoto, W., Jarvis-Morar, M., Brugge, J.S., Bolen, J.S. and Israel, M.A. (1985) *Proc. Natl. Acad. Sci. USA* 82, 4568-4572.
- [30] Grandori, C. (1989) *Nature* 338, p. 467.
- [31] Drubin, D.G. (1991) *Cell* 65, 1093-1096.
- [32] Adams, A.E.M. and Pringle, J.R. (1984) *J. Cell Biol.* 98, 934-945.
- [33] Chant, J. and Herskowitz, I. (1991) *Cell* 65, 1203-1212.
- [34] Ruggieri, R., Bender, A., Matsui, Y., Powers, S., Takai, Y., Pringle, J.R. and Matsumoto, K. (1992) *Mol. Cell. Biol.* 12, 758-766.
- [35] Bryant, P.J. and Woods, D.F. (1992) *Cell* 68, 621-622.
- [36] Williams, L.T. (1989) *Science* 243, 1564-1570.
- [37] Carpenter, G. and Cohen, S. (1990) *J. Biol. Chem.* 265, 7709-7712.
- [38] Anderson, D., Koch, C.A., Grey, L., Ellis, C., Moran, M.F. and Pawson, T. (1990) *Science* 250, 979-982.
- [39] Koch, C.A., Moran, M.F., Anderson, D., Liu, X., Mbamalu, G. and Pawson, T. (1992) *Mol. Cell. Biol.* 12, 1365-1374.
- [40] Klippel, A., Escobedo, J.A., Fantl, W.J. and Williams, L.T. (1992) *Mol. Cell. Biol.* 12, 1451-1459.
- [41] Pawson, T. (1992) *Curr. Op. Gen. Dev.* 2, 4-12.
- [42] Seidel-Dugan, C., Meyer, B.E., Thomas, S.M. and Brugge, J.S. (1992) *Mol. Cell. Biol.* 12, 1835-1845.
- [43] Jackson, P. and Baltimore, D. (1989) *EMBO J.* 8, 449-456.
- [44] Pawson, T. (1988) *Oncogene* 3, 491-495.
- [45] Cantley, L.C., Auger, K.R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R. and Soltoff, S. (1991) *Cell* 64, 281-302.
- [46] Hall, A. (1990) *Science* 249, 635-640.
- [47] Pawson, T. (1992) *Nature* 356, 285-286.
- [48] Cheney, R.E. and Mooseker, M.S. (1992) *Curr. Op. Cell Biol.* 4, 27-35.
- [49] Adams, R.J. and Pollard, T.D. (1989) *Nature* 340, 365-368.
- [50] Matsudaira, P. (1991) *Trends Biochem. Sci.* 16, 87-92.
- [51] Lehto, V.-P., Wasenius, V.-M., Salven, P. and Saraste, M. (1988) *Nature* 334, 388.

- [52] Lehto, V.-P., Virtanen, I., Paasivuo, R., Ralston, R. and Alitalo, K. (1983) *EMBO J.*, 2, 1701-1705.
- [53] Volpp, B.D., Nauseef, W.M. and Clark, R.A. (1988) *Science* 242, 1295-1297.
- [54] Quinn, M.T., Parkos, C.A., Walker L., Orkin, S.H., Dinauer, M.C. and Jesaitis, A.J. (1989) *Nature* 342, 198-200.