Amino Acids (2012) 42:1361–1370
DOI 10.1007/s00726-011-0831-4

ORIGINAL ARTICLE

The SH3 domain of HS1 protein recognizes lysine-rich polyproline motifs

Giuliano Siligardi · Paolo Ruzza · Rohanah Hussain · Luca Cesaro · Anna Maria Brunati · Lorenzo A. Pinna · Arianna Donella-Deana

Received: 4 August 2010 / Accepted: 4 January 2011 / Published online: 2 February 2011

Abstract HS1 is a protein involved in erythroid proliferation and apoptotic cell death, containing several structurally significant motifs including a C-terminal SH3 domain. HPK1 is a member of the Ste20-related kinase family, which contains four proline-rich sequences and is constitutively associated with HS1 in hematopoietic cells. Recombinant fusion protein GST-$SH3_{HS1}$ was expressed to assess the binding properties of 16 peptides derived from the HPK1 proline-rich regions. The binding affinities were determined by non-immobilized ligand interaction assay by circular dichroism. Our results revealed that the classical PxxPxK class II binding motif is not sufficient to induce the interaction with the GST-$SH3_{HS1}$ domain, an event dependent on the presence of additional basic residue(s) located at the C-terminus of the PxxPxK motif: Lys_{-5} in **P2** peptide and Lys_{-8} in **P4c** peptide. Lys replacement by Arg residues decreases the ligand binding affinity. The finding that both $SH3_{HS1}$ domain and full-length HS1 protein bind to **P2** peptide with similar affinity demonstrates that the whole protein sequence does not affect the interaction properties of the domain. In silico models of $SH3_{HS1}$ as a complex with **P2** or **P4c** highlight the domain residues that interact with the recognition determinants of the peptide ligand and that cooperate in the complex stabilization.

Keywords HS1 · Pro-rich peptides · SH3 domain · Protein–protein interaction

The authors G. Siligardi and P. Ruzza contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00726-011-0831-4) contains supplementary material, which is available to authorized users.

G. Siligardi · R. Hussain
Diamond Light Source Ltd., Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire OX11 0QX, UK

P. Ruzza (✉)
Institute of Biomolecular Chemistry of CNR, 35131 Padua, Italy
e-mail: paolo.ruzza@unipd.it

L. Cesaro · A. M. Brunati · L. A. Pinna · A. Donella-Deana
Department of Biological Chemistry, University of Padova, 35131 Padua, Italy

Introduction

Protein–protein interactions represent an essential mechanism in cellular signaling, which is responsible for the transmission of information inside cells. A family of protein–protein interaction modules is represented by the Src homology 3 (SH3) domain, a non-catalytic protein module of around 60 amino acids, originally discovered in the protein tyrosine kinase product of the v-Src oncogene (Cohen et al. 1995; Mayer 2001). SH3 domains have no fixed topological position within proteins and are frequently found in combination with other protein–protein interaction modules (Kaneko et al. 2008). SH3 domains have now been identified in more than 350 different proteins in organisms ranging from yeasts to humans. Whereas the yeast harbors 28 SH3 domains, the human genome encodes approximately 300 of such domains (Karkkainen et al. 2006), which show a completely conserved structure.

The majority of SH3 domains recognize proline-rich peptides that adopt a left-handed poly-proline type II (PPII) helix containing two XP dipeptides, separated by a scaffolding residue (often a proline), forming the XPxXP core element, generally flanked by a basic residue (Marchiani

et al. 2010, and references therein). The detailed requirements of ligand/SH3-domain interactions have been examined by numerous approaches. One remarkable discovery that arose from these studies is that the conserved basic residue can be either N-terminal or C-terminal to the PxxP-motif, permitting classification of Pro-rich ligands into two related, yet distinct, groups named classes I and II, respectively. Structural analysis of SH3 domains in complex with either class I or class II ligands show that these two types of ligands bind to SH3 domains in opposite orientations at the same binding site (Feng et al. 1994). The fact that ligands can bind either N-to-C or C-to-N orientations is a consequence of the twofold rotational pseudosymmetry possessed by the left-handed PPII helix.

Although most SH3 domains recognize ligands containing the PxxP core organized in class I and/or II orientations, a number of SH3 domains bind to peptide sequences that lack this motif or require additional residues outside it to interact with high affinity to SH3, which are so-called "atypical motifs" (Marchiani et al. 2010; Li 2005).

The hematopoietic lineage cell-specific protein 1 (HS1) is mostly expressed in hematopoietic cells and contains several structurally significant motifs including Pro-rich sequences and a C-terminal SH3 domain (Brunati et al. 1999). Following the hematopoietic cell antigen receptor engagement, HS1 undergoes a Tyr-phosphorylation sequentially catalyzed by Syk and Lyn tyrosine kinases (Brunati et al. 2005; Yamanashi et al. 1997; Hao et al. 2004; Hata et al. 1994; Takemoto et al. 1996). HS1 Tyr-phosphorylation has been reported to play a role in B cell antigen receptor-mediated proliferative and apoptotic responses (Yamanashi et al. 1997; Hao et al. 2004; Hata et al. 1994; Takemoto et al. 1996; Fukuda et al. 1995). In chronic lymphocytic leukemia, HS1 is differently expressed; while patients with an aggressive behavior are characterized by the prominent expression of phosphorylated HS1, those with an indolent course express mainly non-phosphorylated HS1 (Scielzo et al. 2005). Nagata et al. (1999) found that, in haematopoietic cells, HS1 is constitutively associated with haematopoietic progenitor kinase-1 (HPK1) through its SH3 domain (Kiefer et al. 1996; Hu et al. 1996).

HPK1 kinase is a member of a family of mammalian Ste20-related serine/threonine protein kinases, the expression of which is restricted to haematopoietic organs and cells in adults (Kiefer et al. 1996; Hu et al. 1996). It has been demonstrated that HPK1 initiates a kinase cascade involving several intermediates leading to SAPK activation (Kiefer et al. 1996; Hu et al. 1996; Wang et al. 1997). In contrast, less is known about the signaling events upstream of HPK1 activation. Several groups have shown that SH2/SH3 domain-containing adaptor proteins, including Crk, Crkl and Grb2, are implicated in the interaction occurring between HPK1 and the activated receptors (Liu et al. 2000).

HPK1 contains a cluster of four proline-rich sequences called **P1**, **P2**, **P3** and **P4** (Table 1). **P1**, **P2** and **P4** peptides show the canonical class II binding motif; while **P1** contains an arginine residue at the −3 position, **P2** and **P4** are characterized by the presence of a lysine residue at this position.

In this study, several peptide analogs derived from the four proline-rich regions of HPK1 were synthesized and tested for their ability to interact with the SH3 domain of HS1 ($SH3_{HS1}$) by non-immobilized ligand interaction assay by CD spectroscopy (NILIA-CD). Our results demonstrated that the class II motif present in the peptides is not sufficient to promote the interaction with the $SH3_{HS1}$ domain. The presence of an additive lysine, located two or five residues after the basal PxxPxK sequence, is required for the high-affinity binding of the peptide ligand to the domain, allowing the identification of two new "atypical" SH3 binding motifs, PxxPxKxKx and PxxPxKxxxxK, respectively.

Experimental procedures

Peptide synthesis

Fmoc-protected amino acids and preloaded Wang or 2-chlorotrityl resins were purchased from Calbiochem–Novabiochem (Läufelfingen, Switzerland). HBTU, HOBt,

Table 1 Proline-rich sequences contained in HPK1 protein and their values of binding affinity to GST-$SH3_{HS1}$ domain

Name	HPK1 residues	Sequence															K_d (μM)
		6	5	4	3	2	1	0	−1	−2	−3	−4	−5	−6	−7	−8	
P1	308–319	P	E	L	P	P	A	I	P	R	R	I	R				12.5 ± 1.2
P2	394–403				P	P	P	L	P	P	K	P	K	F			1.4 ± 0.2
P3	432–443			P	P	P	N	S	P	R	P	G	P	P	P		nbd
P4	468–477			K	P	P	L	L	P	P	K	K	E				>130.0

nbd no binding detectable

Table 2 Proline-rich derivatives of P2 peptide and their values of binding affinity to GST-$SH3_{HS1}$ domain

Name	Sequence										K_d (μM)
	3	2	1	0	−1	−2	−3	−4	−5	−6	
P2	P	P	P	L	P	P	K	P	K	F	1.4 ± 0.2
P2a	P	P	P	L	P	P	K	P	K		12.2 ± 1.0
P2b	P	P	P	L	P	P	K	P			nbd
P2c	P	P	P	L	P	P	K				nbd
P2d		P	P	L	P	P	K	P	K	F	20.0 ± 1.5
P2e			P	L	P	P	K	P	K	F	25.0 ± 1.8
P2f	P	P	P	L	P	P	K	P	R	F	10.0 ± 0.9
P2g	P	P	P	L	P	P	R	P	K	F	14.8 ± 1.4
P2h	P	P	P	L	P	P	R	P	R	F	11.1 ± 0.9
P2i	F	K	P	K	P	P	L	P	P	P	nbd

nbd no binding detectable

DIEA and NMP were obtained from Applied Biosystems-Perkin Elmer (Foster City, CA). Peptides were assembled by manual solid-phase synthesis and the Fmoc/HBTU chemistry in 0.15-mM scale. **P2a**, **P2b**, **P2c** and **P2i** peptides (Table 2) were synthesized using 2-Cl-Trt resins to preclude undesired side reactions (Steinauer and White 1995). Coupling yields were monitored on aliquots of peptide resin either by Kaiser test for the amino groups or by evaluation of Fmoc displacement (Wellings and Atherton 1997). Peptides were side chain deprotected and removed from the resin by TFA treatment in the presence of 2.5% TIS, 2.0% anisole and 0.5% water, and then precipitated by addition of diethyl ether.

Crude peptides were purified by preparative reversed-phase HPLC using a Shimadzu LC-8 (Shimazdu, Kyoto, Japan) system with a Vydac 218TP1022, 10 μ, 250 × 22 mm column (Grace Davison Discovery Sciences, Deerfield, IL). The column was perfused at a flow rate of 12 mL min^{-1} with a mobile phase containing solvent A (0.05% TFA in water) and a linear gradient from 10 to 30% of solvent B (0.05% TFA in acetonitrile/water, 9:1 by volume) in 40 min. The fractions containing the desired product were collected and lyophilized to constant weight in the presence of 0.01 N HCl. Analytical HPLC analyses were performed on a Shimadzu LC-10 instrument fitted with a Jupiter C18, 10 μ, 250 × 4.6 mm column (Phenomenex, Torrance, CA) using the described solvent system (solvents A and B), with a flow rate of 1 mL min^{-1} and detection at 216 nm. All peptides showed less than 1% impurities. Molecular weights of compounds were determined by ESI-MS on a Mariner (PerSeptive Biosystem, Foster City, CA) mass spectrometer instrument. The mass was assigned, using a mixture of neurotensin, angiotensin and bradykinin at a concentration of 1 pmol μL^{-1}, as external standard. The amino acid compositions of the peptide acid hydrolysates (6 M HCl, 22 h at 110°C in sealed evacuated vials) were determined with a Carlo Erba 3A30 (Milan, Italy) amino acid analyzer.

Expression and purification of GST-SH3, free SH3 and full-length HS1

The oligonucleotides 5′- CGT GGG ATC CCC GGG ATC TCA TCA GCT-3′ and 5′-GAC CCG GGA ATT CTC CAG AAG CTT GAC-3′ were used to amplify by PCR the region encoding the sequence 430–486 containing the SH3 domain (amino acids 434–482) from pTcrHis HS1 (Brunati et al. 1999). The introduced BamHI and EcoRI restriction sites were exploited to directionally clone the insert in the expression vector pGEX-5X-3 (Pharmacia Biosciences,). GST and GST-$SH3_{HS1}$ fusion protein were expressed in *Escherichia coli* BL21 and affinity purified by glutathione–agarose column (Sigma) according to the manufacturer's instructions. GST-$SH3_{HS1}$ domain concentration was determined by absorption spectroscopy (ε = 58,580 M^{-1} cm^{-1} at 280 nm). The occurrence of GST-$SH3_{HS1}$ dimerization in solution was analyzed by FPLC gel filtration on a Superdex G75 (Amersham Pharmacia Biotech). GST-$SH3_{HS1}$ was eluted in 5 mM Tris–HCl, pH 7.0, buffer at a flow rate of 0.4 mL min^{-1}. Globular marker proteins (Sigma) were used for molecular mass determination.

The $SH3_{HS1}$ domain was cleaved from GST-$SH3_{HS1}$ on glutathione–Sepharose beads using thrombin (0.1 U/1 mg of protein) at 20°C for 14 h, and the thrombin excess was removed using benzamidine Sepharose 6B (Pharmacia Biotech, Piscataway, NJ). The eluate was collected and concentrated with an Amicon Ultra-10 (Millipore). The $SH3_{HS1}$ domain concentration was determined by absorption spectroscopy (ε = 15,470 M^{-1} cm^{-1} at 280 nm).

Full-length HS1 was prepared as described elsewhere (Brunati et al. 1999).

Circular dichroism and GST-SH3 titration

All measurements were obtained at room temperature using nitrogen-flushed Jasco J-600 and J-720 spectropolarimeters (Tokyo, Japan) with either a 0.5 or a 1.0 cm (near-UV) and a 0.02 cm (far-UV) quartz cells, respectively. Each titration was performed in aqueous buffer (5 mM Tris–HCl, pH 7.0, 1 mM dithiothreitol) at room temperature by addition of small aliquots of peptide stock solution in the same buffer as described in Siligardi et al. (2002b). The contribution of the peptide alone, evaluated by control CD titrations, was subtracted from the CD spectra of the domain in complex with each peptide containing the aromatic Phe residue (**P2** peptide and its derivatives). Peptide concentration was determined by weight using a Mettler Toledo microbalance (Columbus, OH) model AT21 Comparator (sensitivity ± 1 μg). Protein concentration was determined by absorption spectroscopy (GST-$SH3_{HS1}$ = 43.0 μM; $SH3_{HS1}$ = 74.2 μM; HS1 = 50.3 μM). The dissociation constants K_d of the different complexes were determined by analyzing the CD data at a single wavelength by a nonlinear least-squares computer fitting to the equation based on 1:1 binding stoichiometry:

$$\Delta A = \frac{(\Delta\varepsilon_l - \Delta\varepsilon_0)((1 + K_a \times P + K_a \times L) - \sqrt{((1 + K_a \times P + K_a \times L)^2 - 4 \times K_a^2 \times P \times L)})}{2K_a} + (\Delta\varepsilon_0 \times P)$$

where $\Delta\varepsilon_0$ and $\Delta\varepsilon_l$ are the differential molar extinction coefficients (M^{-1} cm^{-1}) either of protein alone or bound to ligand, respectively. L and P represent the total concentration (M) of peptide and protein, respectively, and K_a is the association constant (M^{-1}) for the peptide–protein interaction (Siligardi and Hussain 1998).

In silico studies

The design of the peptides **P2** and **P4c** and their binding to $SH3_{HS1}$ were based on the crystal structure of the human cortactin SH3 domain in complex with a peptide ligand (PDB code: 2D1X). The energy minimization of the SH3–peptide complexes were performed with the program DiscoveryStudio (Accelrys Software, Inc.) until the energy reached a minimum. The images were drawn by using the program PYMOL (W. L. DeLano, http://www.pymol.org).

Results

The interaction occurring between the synthetic Pro-rich peptides derived from HPK1 and GST-$SH3_{HS1}$ fusion protein was analyzed using the near-UV NILIA-CD technique. The region of the SH3 domain involved in the interaction with the proline-rich core of the peptide ligands is a surface patch formed by the side chains of a few well-conserved Trp and Tyr residues. Consequently, aromatic residues are excellent in situ molecular probes of protein–peptide interaction, and CD associated with the aromatic residues can be successfully used to screen ligand binding, both qualitatively and quantitatively. The binding of the peptides to GST-$SH3_{HS1}$ domain can be clearly determined by the spectral changes of the two strong features at approximately 290 and 295 nm. These changes are indicative of a reduction in the mobility and/or of local environmental changes of the Trp residues strictly related to the binding process (Fig. 1a, b). The values of the apparent dissociation constant K_d can be determined by the nonlinear regression analysis of the ΔA values of the CD spectra at 295 nm with increasing SH3/peptide molar ratios (Fig. 2 and Tables 1, 2, 3).

Table 1 shows that **P2** peptide interacts with GST-$SH3_{HS1}$ fusion protein with a very favorable K_d value (1.4 ± 0.2 μM), while **P1** binds with an affinity about ninefold lower (K_d = 12.5 ± 1.2 μM) than that of **P2**. As expected, **P3** peptide, which does not include a basic residue at the −3 position, is unable to interact with GST-$SH3_{HS1}$ fusion protein in agreement with the data reported by Anafi et al. (1997) and for this reason used as scrambled peptide to evaluate the presence of non-specific peptide effects on the environment of the Trp and Tyr residues of the $SH3_{HS1}$ domain. Surprisingly **P4**, which contains the canonical class II binding motif, is a poor peptide ligand (K_d = > 130 μM).

Control experiments performed with recombinant GST protein and increasing **P2** concentration demonstrated that

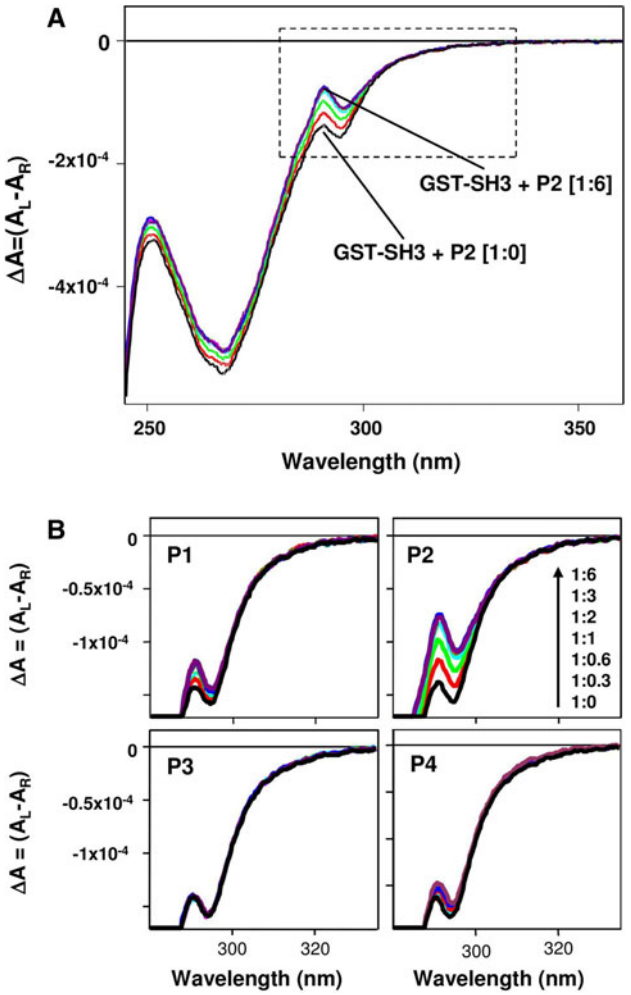

Fig. 1 Near-UV CD titration of GST-$SH3_{HS1}$ domain with HPK1 proline-rich peptides. The analysis was performed in Tris–HCl buffer solution, pH 7.5, at room temperature. **a** The CD changes in the near-UV region (245–360 nm) on addition of increasing amounts of **P2** peptide to the GST-$SH3_{HS1}$ domain (43.0 μM), due to the perturbation of the local tertiary structure of the SH3 Tyr (about 270 nm) and Trp (about 295 nm) aromatic side chain residues induced by the peptide interaction. **b** The details of CD changes observed in the near-UV region (280–340 nm) on the addition of increasing amounts of **P1**, **P2**, **P3** or **P4** peptide to GST-$SH3_{HS1}$ domain. The Δ*A* values were measured as a function of the increasing SH3/peptide molar ratios

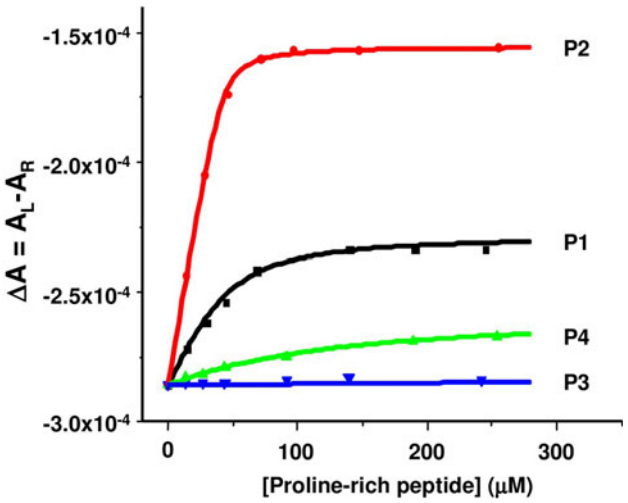

Fig. 2 Determination of the K_d values by CD spectroscopy analyses. The Δ*A* values, measured as a function of different SH3/peptide molar ratios, are plotted versus the peptide concentration

GST does not interact with the Pro-rich peptide (Fig. 1S, supplementary material). Moreover, the occurrence in solution of a potential GST-SH3 dimerization was ruled out by FPLC gel filtration analysis.

Cleaved $SH3_{HS1}$ domain and the recombinant full-length HS1 protein were also titrated with **P2** peptide and the obtained CD data were analyzed as previously detailed. CD studies revealed that free $SH3_{HS1}$ domain and the full-length HS1 protein interact with **P2** peptide with a similar affinity binding (K_d values: 6.1 ± 1.0 and 5.9 ± 1.0 μM, respectively), which is in the same range but slightly lower than that of GST-$SH3_{HS1}$ (Fig. 3). Because of the low yield of both $SH3_{HS1}$ domain and full-length HS1 protein expression, CD studies were performed using the stable GST-$SH3_{HS1}$ recombinant protein.

As **P2** displayed the highest $SH3_{HS1}$ binding affinity, this peptide was selected to determine the minimum sequence requirement and to identify the key residues involved in the SH3 domain recognition. Therefore, several peptide-analogs of **P2** were synthesized and analyzed for their binding affinity to the GST-$SH3_{HS1}$ fusion protein.

Surprisingly, the results obtained demonstrated that the typical class II PxxPxK motif is not sufficient to promote the peptide–$SH3_{HS1}$ interaction and that the whole **P2** sequence is required for a high-affinity binding (Table 2). In particular, the shortening of the **P2** sequence (PPPLPPKPKF) from the C-terminus showed that the Phe-removed **P2a** analog displays an SH3 affinity of about one order of magnitude lower (K_d = 12.2 ± 1.0 μM) than that by **P2**. The further deletion of the Lys_{-5} residue (**P2b** peptide) abrogates the interaction with $SH3_{HS1}$ domain, indicating the crucial role played by the Lys residue at the −5 position (Table 2). Consistently, the peptide **P2c**, which lacks also the Pro_{-4} residue and shows a class II consensus motif, does not bind to GST-$SH3_{HS1}$ (Table 2). On the other hand, the deletion of the N-terminal Pro_3 and Pro_2 residues of the **P2** sequence does not abrogate, but highly reduces the GST-$SH3_{HS1}$ binding affinity of **P2d** (K_d = 20 ± 1.5 μM) and **P2e** (K_d = 25 ± 1.8 μM) peptides, respectively (Table 2). This might be consistent with the disruption and perturbation of the required PPII helical conformation that, together with the cationic residues in the C-terminal region, contributes to promote the interaction with the $SH3_{HS1}$ domain.

Table 3 Proline-rich derivatives of P4 peptide and their values of binding affinity to GST-$SH3_{HS1}$ domain

Name	Sequence													K_d (μM)
	4	3	2	1	0	−1	−2	−3	−4	−5	−6	−7	−8	
P4	K	P	P	L	L	P	P	K	K	E				>130.0
P4a	K	P	P	L	L	P	P	K	K	E	K			>130.0
P4b	K	P	P	L	L	P	P	K	K	E	K	nL		>130.0
P4c	K	P	P	L	L	P	P	K	K	E	K	nL	K	1.4 ± 0.2

nL norleucine

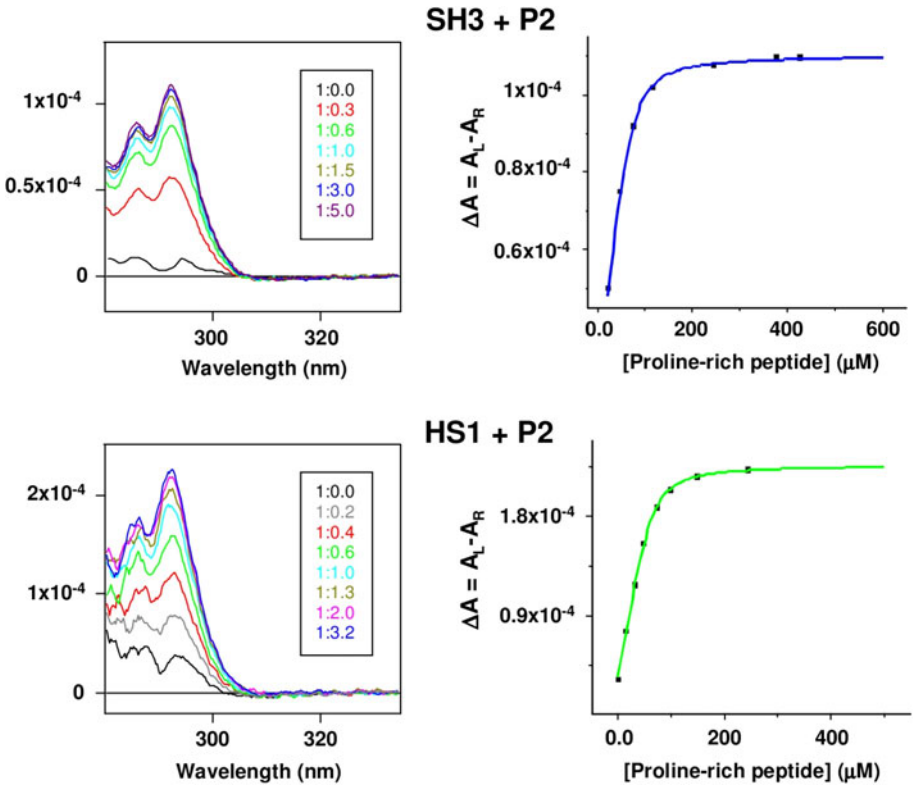

Fig. 3 Near-UV CD titration of isolated $SH3_{HS1}$ domain (74.2 μM) and whole HS1 protein (50.3 μM) with **P2** peptide. CD analysis was performed in Tris–HCl buffer solution, pH 7.5, at room temperature. The figure shows the details of CD changes observed in the near-UV region on addition of increasing amounts of **P2** peptide to either isolated $SH3_{HS1}$ domain (*upper panel*) or whole HS1 protein (*lower panel*). The Δ*A* values, measured as a function of different SH3/peptide molar ratios, are plotted versus the peptide concentration

The retro peptide **P2i** (FKPKPPLPPP), which includes the canonical KxxPxxP class I motif, also does not bind to GST-$SH3_{HS1}$.

An interesting property of the **P2** sequence is the presence of a lysine instead of the more common arginine as a basic residue at the −3 position. The substitution of either one or both arginine for lysine residues leads to a reduction in the affinity of **P2** analogs (**P2f**, **P2g** and **P2h** peptides) for the GST-$SH3_{HS1}$ domain (see Table 2). In particular, replacing both lysine residues, the cationic region of **P2h** becomes very similar to that of **P1**, which shares similar affinity for GST-$SH3_{HS1}$ (K_d = 11.1 ± 0.9 and 12.5 ± 1.2 μM, respectively).

With the aim of improving the binding affinity of **P4** peptide for the $SH3_{HS1}$ domain, we synthesized a series of peptides containing an increased numbers of residues at the C-terminus (Table 3). Indeed, in a recent paper Rubini et al. (2010) reported that the elongation of the **P4** peptide with three additional C-terminal residues is sufficient to increase its affinity toward the murine cortactin SH3 domain, consistent with previous results reported by Lewitzky et al. (2004). In **P4b** and **P4c** peptides, the native methionine residue was replaced by the isosteric Nle to avoid its oxidation during the CD assays and thus a decrease of bioactivity (Borin et al. 1987; Hruby et al. 1980). The peptide analogs elongated with either one (Lys, **P4a**) or two (Lys-Nle, **P4b**) residues do not exhibit an affinity toward GST-$SH3_{HS1}$ domain higher than that of **P4**. On the other hand, the elongation of the **P4b** sequence by a further Lys residue (**P4c**) increases the ligand binding by almost two orders of magnitude (Table 3), reaching the same K_d value of the **P2** peptide (1.4 ± 0.2 μM).

Discussion

Different mechanisms are necessary to regulate the functions of the over 800 SH3 domains found in the mammalian genomes (Karkkainen et al. 2006). Besides the binding specificity for proline-rich sequences, the SH3 domain may achieve interaction specificity using additional residues located outside the class I or class II binding motifs.

In the case of the HS1 protein, previous studies have demonstrated that its interaction with HPK1 kinase is mediated by the SH3 domain of HS1 and the proline-rich region of the kinase (Nagata et al. 1999; Siligardi et al. 2002a). Here, the binding properties of 16 peptides derived from the HPK1 Pro-rich sequences were investigated using near-UV CD spectroscopy. The reported K_d values could be assessed by different methods of analysis; however, it is noteworthy that the affinity constants obtained in similar studies by CD and ITC approaches were comparable (Salvatella et al. 2000; Rubini et al. 2010). Among the peptides examined, the GST-$SH3_{HS1}$ fusion protein displays the highest affinity toward **P2** ($K_d = 1.4 \pm 0.2$ μM) and **P4c** ($K_d = 1.4 \pm 0.2$ μM) peptides, which reproduce the regions 394–403 and 468–480 of HPK1, respectively (Tables 1 and 3). A peculiar feature of these peptides, which contain a class II binding motif, is the presence of a lysine instead of the more common arginine at the −3 position, and of an additive Lys residue outside this motif at either −5 or −8 position, respectively (Tables 1 and 3).

As CD analysis demonstrated that recombinant GST protein does not interact with **P2** peptide, the outcome that cleaved $SH3_{HS1}$ shows a **P2** binding affinity of about fourfold lower than that displayed with GST-$SH3_{HS1}$ suggesting that the GST fusion induces a "stabilization" of the domain structure. The finding that cleaved $SH3_{HS1}$ and full-length HS1 protein display similar K_d values (6.1 ± 1.0 and 5.9 ± 1.0 μM, respectively) demonstrates that the whole protein sequence does not affect the affinity of the peptide interaction and that the SH3 domain contains all the binding recognition determinants.

The replacement in the **P2** peptide of Lys at the −3 position by an arginine residue (peptide **P2g**) leads to a tenfold reduction in the affinity for GST-$SH3_{HS1}$, suggesting the presence of Lys-specific binding site(s) in the domain, as confirmed by further binding studies performed with the **P2f** and **P2h** peptides (Table 2). The occurrence of Lys-specific interactions with the SH3 domain, a rare property that confers further selectivity to the peptide binding, was first observed in the Crk N-terminal SH3 domain. The crystal structure of the complex between $SH3_{Crk}$ and a ligand peptide revealed that in the RT loop of the domain, three out of four acidic residues were involved in an interaction with a specific Lys residue. The substitution of Arg for Lys disrupted some of these specific interactions as confirmed by crystal data, which showed that the side chain of the Arg residue interacted with only two of the four acidic residues of the RT loop (Knudsen et al. 1995).

The crucial role of the charged residues located outside the classical class II binding motif was highlighted by the affinity values obtained with both **P2b** and **P2c** peptides, which lack the Lys residue at the −5 position. The deletion of this residue abrogates the binding capability of these peptides (Table 2). On the contrary, the deletion of the N-terminal Pro–Pro dipeptide (peptide **P2e**), which removes the PxxP binding motif, unexpectedly does not abrogate the peptide ligand binding.

The influence of positive charges outside the classical class II binding motif was also confirmed by the affinity data obtained with the **P4** derivatives (Table 3). The introduction of an additional Lys_{-8} residue (peptide **P4c**) strongly decreases the K_d value of the interaction with GST-$SH3_{HS1}$, which reaches a binding affinity similar to that of the **P2** peptide.

The HS1 SH3 domain is highly similar (about 86%) to cortactin SH3 domain (Kitamura et al. 1989). Consistently, we have recently demonstrated that the murine cortactin SH3 domain displays peptide binding properties similar to those of the HS1 domain. In particular, also $SH3_{m\text{-}cort}$ prefers a lysine instead of the more common arginine as interacting basic residue and recognizes with high affinity Pro-rich sequences containing additional basic residues located at either the N- or the C-terminal of the PxxPxK basal motif (Rubini et al. 2010).

Considering the high similarity between the SH3 domains of HS1 and cortactin, in silico models of $SH3_{HS1}$ in complex with either **P2** or **P4c** peptide ligands (Fig. 4a, b, respectively) were constructed using the crystal structure of human cortactin SH3 domain bound to AMAP1 proline-rich peptide (Hashimoto et al. 2006). Figure 4a shows that the PPII helical conformation of the class II binding motif (PxxPxK) provides a "semi-rigid" template, enabling hydrogen bonding between the accessible peptide amide carbonyl groups and the SH3 Tyr54 phenolic hydroxy group and the Trp38 indolic amino group. Other interactions involve the long cationic side chain of Lys_{-3} and the anionic side chain residues of the SH3 domain, Glu15, Asp18 and Glu19. In addition, Lys_{-5} side chain forms a salt bridge with the side chain of the SH3 domain Asp35 that is required to assist and maintain the binding.

Figure 4b shows that interactions very similar to those described for **P2** occur between the $SH3_{HS1}$ domain and **P4c** peptide backbone. Moreover, the model highlights additional contacts occurring between the side chain of **P4c** Lys residues and the side chains of the $SH3_{HS1}$ Glu and Asp residues. In particular, contacts are present between the N-terminal Lys_4 residue of **P4c** and Asp11 of the domain, as

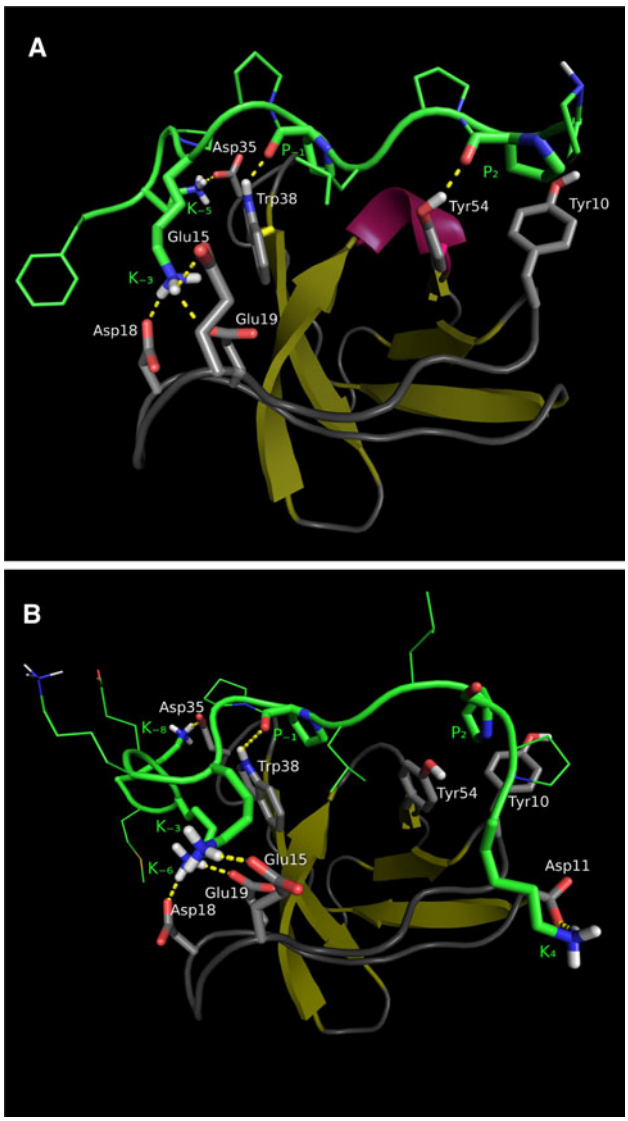

Fig. 4 In silico modeling of SH3–peptide complexes. The interactions occurring between the $SH3_{HS1}$ domain (*cartoon*) and either P2 (**a**) or P4c (**b**) peptide ligand (*green backbone*) are shown. The peptide design and its binding to $SH3_{HS1}$ were performed as described in the section "Experimental procedures". In (**a**) the interactions occurring between the residues of the ligand PxxPxK binding motif and the amino acids belonging to the SH3 domain are: Pro_2–Tyr54, Pro_{-1}–Trp38 and Lys_{-3}–Glu15/Asp18/Glu19; another essential interaction is present between the side chain of Lys_{-5} of **P2** peptide and the side chain of Asp35 of $SH3_{HS1}$ domain. In (**b**) the interactions involving the residues of the ligand PxxPxK binding motif and the amino acids belonging to the SH3 domain are: Pro_2–Tyr54, Pro_{-1}–Trp38 and Lys_{-3}–Glu15; other essential interactions are between the side chains of **P4c** cationic residues and SH3 domain anionic residues, Lys_4–Asp11, Lys_{-6}–Asp18/Glu19 and Lys_{-8}–Asp35 (color figure online)

well as between the Lys_{-3} and Glu15, Lys_{-6} and both Asp18 and Glu19, and between the side chain of Lys_{-8} and Asp35. This last interaction supports the CD binding data, which indicate the positive contribution of the C-terminus Lys residue in the binding of **P2** and **P4c** peptides.

To understand the mechanism of the peptide binding recruitment peculiar to the $SH3_{HS1}$ domain, we analyzed some available structures of other SH3 domains complexed with extended PxxP-containing ligands (Fig. 5). According to the recent work of Kim et al. (2008), it is possible to define two distinct regions of the SH3 structure, referred as Surface I and Surface II, which are in contact with the peptide ligands. Surface I has been extensively described and comprises a fairly narrow region that is well suited for binding a Pro-rich sequence in PPII helical conformation (Li 2005; Zarrinpar et al. 2003). On the other hand, Surface II, referred to as the "specificity pocket" in some studies (Li 2005; Zarrinpar et al. 2003), is much broader and contains residues less conserved (Larson and Davidson 2000). Some residues in this surface, which are located in the loops RT or N-Src and in the strands *c* and *d* interact with the −3 position and other amino acids flanking the core consensus motif of the peptide ligand.

The importance of the $SH3_{HS1}$ Surface II in mediating the domain binding to its ligand is strongly supported by the behavior of **P2e** peptide, which, in spite of the lack of the PxxP binding motif, interacts with the $SH3_{HS1}$ domain suggesting a minor role of Surface I in the peptide recruitment. This hypothesis is also supported by our previous work on **P2** analogs in which proline residues located at different *i*, $i + 3$ positions were separately replaced by 4-fluoroproline residues (Ruzza et al. 2006). The induction of a stable PPII helical conformation in the peptide ligand is not correlated with an increased affinity toward the $SH3_{HS1}$ domain, suggesting that a rigid PPII scaffold is not suited for the ligand binding surface (Ruzza et al. 2006), more likely due to a decrease in the peptide flexibility, which is necessary for an optimal interaction with the Surface II of the $SH3_{HS1}$ domain. Therefore, the ability of the domain Surfaces I and II to cooperate in the binding to the peptide PxxP-motif and flanking residues, respectively, is an $SH3_{HS1}$ requirement for a high binding affinity to its peptide ligands.

Conclusions

Summarizing, the results of the present study suggest that the $SH3_{HS1}$ domain recognizes extended proline-rich motifs characterized by an extra Lys residue outside the minimal PxxPxK class II motif. The binding sequence of **P2** (xPxxPxKxKx) and **P4c** (xPxxPxKxxxxK) are two more examples of the growing family of extended and non-canonical SH3 domain binding motifs.

An increasing body of evidence highlights the importance of SH3-mediated protein interactions in critical signaling processes that result in cancers and other pathological conditions. Therefore, several attempts have been

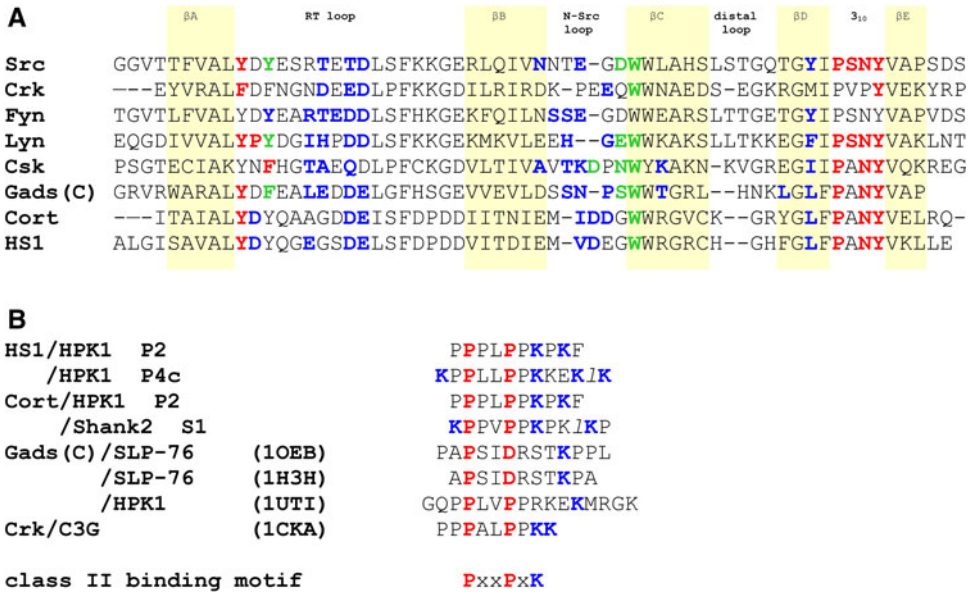

Fig. 5 Sequence alignment of the SH3 domains described to interact with extended peptide sequences and their related peptide ligands. **a** The SH3 sequence alignment was constructed in ClustalW (http://www.uniprot.org). The specific interactions occurring between the SH3 binding pockets and their related ligands are *highlighted* as determined by the available PDB domain/peptide structures. The amino acids of Surface I, which contact the peptide PxxP region, are colored *red*, while the residues of Surface II, which contact Lys_{-3} and its flanking residues, are colored *blue*. The residues contacting both the PxxP regions and other parts of the peptides are colored *green*. **b** The *panel* shows the peptide ligands of the SH3 domains aligned in (**a**). The peptide residues and their interacting amino acids located in the cognate SH3 domain are similarly *colored*. The canonical class II consensus motif is shown *below* the peptide sequences. The PDB files of ligands in complex with the SH3 domains are reported in *parentheses*; the data of $SH3_{cort}$ are from Rubini et al. (2010) (color figure online)

performed, aimed at generating drugs that can interfere with the SH3-mediated processes. However, the promiscuity and restrictions characterizing the SH3/ligand interactions are challenges for rational drugs design. In this respect, the finding that $SH3_{HS1}$ domain specifically interacts with proteins containing new Pro-rich binding motifs might facilitate the design of specific inhibitors to be analyzed in experimental models, where HS1-signaling has been demonstrated to be involved in malignancy.

Acknowledgments The work was supported by Italian National Research Council (Short-Term Mobility 2001), Italian Association for Cancer Research (grant to L.A.P.) and "Italian Cystic Fibrosis Research Foundation (Grant #4/2007)" with the contribution of Banca Popolare di Verona e Novara and "Fondazione Giorgio Zanotto".

References

Anafi M, Kiefer F, Gish GD, Mbamalu G, Iscove NN, Pawson T (1997) SH2/SH3 adaptor proteins can link tyrosine kinases to a Ste20-related protein kinase, HPK1. J Biol Chem 272: 27804–27811

Borin G, Calderan A, Ruzza P, Moroder L, Goehring W, Bovermann G, Wuensch E (1987) Synthesis of human [15-norleucine]little-gastrin-II and des-1-tryptophan-[12-norleucine]minigastrin-II. Biol Chem Hoppe-Seyler 368:1363–1372

Brunati AM, Donella-Deana A, James P, Quadroni M, Contri A, Marin O, Pinna LA (1999) Molecular features underlying the sequential phosphorylation of HS1 protein and its association with c-Fgr protein-tyrosine kinase. J Biol Chem 274:7557–7564

Brunati AM, Deana R, Folda A, Massimino ML, Marin O, Ledro S, Pinna LA, Donella-Deana AD (2005) Thrombin-induced tyrosine phosphorylation of HS1 in human platelets is sequentially catalyzed by Syk and Lyn tyrosine kinases and associated with the cellular migration of the protein. J Biol Chem 280:21029–21035

Cohen GB, Ren RB, Baltimore D (1995) Modular binding domains in signal-transduction proteins. Cell 80:237–248

Feng SB, Chen JK, Yu HT, Simon JA, Schreiber SL (1994) 2 binding orientations for peptides to the Src SH3 domain—development of a general-model for SH3–ligand interactions. Science 266:1241–1247

Fukuda T, Kitamura D, Taniuchi I, Maekawa Y, Benhamou LE, Sarthou P, Watanabe T (1995) Restoration of surface Igm-mediated apoptosis in an anti-Igm-resistant variant of Wehi-231 lymphoma-cells by HS1, a protein-tyrosine kinase substrate. Proc Natl Acad Sci USA 92:7302–7306

Hao JJ, Carey GB, Zhan X (2004) Syk-mediated tyrosine phosphorylation is required for the association of hematopoietic lineage cell-specific protein 1 with lipid rafts and B cell antigen receptor signalosome complex. J Biol Chem 279:33413–33420

Hashimoto S, Hirose M, Hashimoto A, Morishige M, Yamada A, Hosaka H, Akagi KI, Ogawa E, Oneyama C, Agatsuma T, Okada M, Kobayashi H, Wada H, Nakano H, Ikegami T, Nakagawa A, Sabe H (2006) Targeting AMAP1 and cortactin binding bearing an atypical src horology 3/proline interface for prevention of breast cancer invasion and metastasis. Proc Natl Acad Sci USA 103:7036–7041

Hata D, Nakamura T, Kawakami T, Kawakami Y, Herren B, Mayumi M (1994) Tyrosine phosphorylation of MB-1, B29, and HS1 proteins in human B-cells following receptor cross-linking. Immunol Lett 40:65–71

Hruby VJ, Sawyer TK, Yang YCS, Bregman MD, Hadley ME, Heward CB (1980) Synthesis and structure–function studies of melanocyte stimulating hormone analogs modified in the 2 and 4(7) positions: comparison of activities on frog skin melanophores and melanoma adenylate cyclase. J Med Chem 23:1432–1437

Hu MCT, Qiu WR, Wang XP, Meyer CF, Tan TH (1996) Human HPK1, a novel human hematopoietic progenitor kinase that activates the JNK/SAPK kinase cascade. Genes Dev 10: 2251–2264

Kaneko T, Li L, Li SSC (2008) The SH3 domain—a family of versatile peptide- and protein-recognition module. Front Biosci 13:4938–4952

Karkkainen S, Hiipakka M, Wang JH, Kleino I, Vaha-Jaakkola M, Renkema GH, Liss M, Wagner R, Saksela K (2006) Identification of preferred protein interactions by phage-display of the human Src homology-3 proteome. EMBO Rep 7:186–191

Kiefer F, Tibbles LA, Anafi M, Janssen A, Zanke BW, Lassam N, Pawson T, Woodgett JR, Iscove NN (1996) HPK1, a hematopoietic protein kinase activating the SAPK/JNK pathway. EMBO J 15:7013–7025

Kim JM, Lee CD, Rath A, Davidson AR (2008) Recognition of non-canonical peptides by the yeast Fus1p SH3 domain: elucidation of a common mechanism for diverse SH3 domain specificities. J Mol Biol 377:889–901

Kitamura D, Kaneko H, Miyagoe Y, Ariyasu T, Watanabe T (1989) Isolation and characterization of a novel human gene expressed specifically in the cells of hematopoietic lineage. Nucleic Acid Res 17:9367–9379

Knudsen BS, Zheng J, Feller SM, Mayer JP, Burrell SK, Cowburn D, Hanafusa H (1995) Affinity and specificity requirements for the first Src homology-3 domain of the crk proteins. EMBO J 14:2191–2198

Larson SM, Davidson AR (2000) The identification of conserved interactions within the SH3 domain by alignment of sequences and structures. Protein Sci 9:2170–2180

Lewitzky M, Harkiolaki M, Domart MC, Jones EY, Feller SM (2004) Mona/Gads SH3C binding to hematopoietic progenitor kinase 1 (HPK1) combines an atypical SH3 binding motif, R/KXXK, with a classical PXXP motif embedded in a polyproline type II (PPII) helix. J Biol Chem 279:28724–28732

Li SSC (2005) Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. Biochem J 390:641–653

Liu SK, Smith CA, Arnold R, Kiefer F, McGlade CJ (2000) The adaptor protein Gads (Grb2-related adaptor downstream of Shc) is implicated in coupling hemopoietic progenitor kinase-1 to the activated TCR. J Immunol 165:1417–1426

Marchiani A, Antolini A, Calderan A, Ruzza P (2010) Atypical SH3 domain binding motifs: features and properties. Curr Top Pept Protein Res (in press)

Mayer BJ (2001) SH3 domains: complexity in moderation. J Cell Sci 114:1253–1263

Nagata Y, Kiefer F, Watanabe T, Todokoro K (1999) Activation of hematopoietic progenitor kinase-1 by erythropoietin. Blood 93:3347–3354

Rubini C, Ruzza P, Spaller MR, Siligardi G, Hussain R, Udugamasooriya DG, Bellanda M, Mammi S, Borgogno A, Calderan A, Cesaro L, Brunati AM, Donella-Deana A (2010) Recognition of lysine-rich peptide ligands by murine cortactin SH3 domain: CD, ITC and NMR studies. Biopolymers 94:298–306

Ruzza P, Siligardi G, Donella-Deana A, Calderan A, Hussain R, Rubini C, Cesaro L, Osler A, Guiotto A, Pinna LA, Borin G (2006) 4-Fluoroproline derivative peptides: effect on PPII conformation and SH3 affinity. J Pept Sci 12:462–471

Salvatella X, Peczuh MW, Gairi M, Jain RK, Sanchez-Quesada J, de Mendoza J, Hamilton AD, Giralt E (2000) Side chain elongation causes a change from enthalpy driven to entropy driven binding in the molecular recognition of tetraanionic peptides. Chem Commun: 1399–1400

Scielzo C, Ghia P, Conti A, Bachi A, Guida G, Geuna M, Alessio M, Caligaris-Cappio F (2005) HS1 protein is differentially expressed in chronic lymphocytic leukemia patient subsets with good or poor prognoses. J Clin Invest 115:1644–1650

Siligardi G, Hussain R (1998) Biomolecules interactions and competitions by non-immobilised ligand interaction assay by circular dichroism. Enantiomer 3:77–86

Siligardi G, Hussain R, Donella-Deana A, Brunati A, Pinna LA, Ruzza P, Calderan A, Borin G (2002a) Binding interactions of derived HPK1 Pro-rich peptides to HS1-SH3 domain suggest a possible adapter function of HS1 protein. In: Benedetti E, Pedone C (eds) Peptides 2002. Ziino, Napoli, pp 886–887

Siligardi G, Panaretou B, Meyer P, Singh S, Woolfson DN, Piper PW, Pearl LH, Prodromou C (2002b) Regulation of Hsp90 ATPase activity by the co-chaperone Cdc37p/p50(cdc97). J Biol Chem 277:20151–20159

Steinauer R, White P (1995) Multi-gram synthesis of peptides containing C-terminal proline. In: Epton R (ed) Innovation and perspectives in solid phase synthesis 1994. Mayflower Worldwide Ltd, Kingswinford, pp 689–692

Takemoto Y, Sato M, Furuta M, Hashimoto Y (1996) Distinct binding patterns of HS1 to the Src SH2 and SH3 domains reflect possible mechanisms of recruitment and activation of downstream molecules. Int Immunol 8:1699–1705

Wang WF, Zhou GS, Hu MCT, Yao ZB, Tan TH (1997) Activation of the hematopoietic progenitor kinase-1 (HPK1)-dependent, stress-activated c-Jun N-terminal kinase (JNK) pathway by transforming growth factor beta (TGF-beta)-activated kinase (TAK1), a kinase mediator of TGF beta signal transduction. J Biol Chem 272:22771–22775

Wellings DA, Atherton E (1997) Standard Fmoc protocols. In: Fields GB (ed) Solid-phase peptide synthesis. Academic Press, New York, pp 44–67

Yamanashi Y, Fukuda T, Nishizumi H, Inazu T, Higashi K, Kitamura D, Ishida T, Yamamura H, Watanabe T, Yamamoto T (1997) Role of tyrosine phosphorylation of HS1 in B cell antigen receptor-mediated apoptosis. J Exp Med 185:1387–1392

Zarrinpar A, Bhattacharyya RP, Lim WA (2003) Sci STKE 179:RE8