The EMBO Journal vol.14 no.10 pp.2191–2198, 1995

Affinity and specificity requirements for the first Src homology 3 domain of the Crk proteins

Beatrice S.Knudsen, Jie Zheng[1], Stephan M.Feller, John P.Mayer[2], Sarah K.Burrell[2], David Cowburn[1] and Hidesaburo Hanafusa[3]

Laboratory of Molecular Oncology and [1]Laboratory of Physical Biochemistry, The Rockefeller University, New York, NY 10021 and [2]Amgen Boulder Inc., Boulder, CO 80301, USA

[3]Corresponding author

The specificity of SH3 domain complex formation plays an important role in determining signal transduction events. We have previously identified a highly specific interaction between the first CrkSH3 domain [CrkSH3(1)] and proline-rich sequences in the guanine nucleotide exchange factor C3G. A 10 amino acid peptide derived from the first proline-rich sequence ($P^3P^4P^5A^6L^7P^8P^9K^{10}K^{11}R^{12}$) bound with a K_d of 1.89 ± 0.06 μM and fully retained the high affinity and unique selectivity for the CrkSH3(1) domain. Mutational analysis showed that P^5, P^8, L^7 and K^{10} are critical for high affinity binding. A conservative mutation, K10R, significantly decreased the affinity for the CrkSH3(1) domain while increasing the affinity for Grb2. Comparative binding studies with the K10R and K10A mutant peptides to c-Crk and v-Crk further suggested that K^{10} binds via a charge-dependent and a charge-independent interaction to the RT loop of the CrkSH3(1) domain. Besides determining important structural features necessary for high affinity and specificity binding to the CrkSH3(1) domain, our results also demonstrate that a conservative mutation in a single amino acid can significantly alter the specificity of an SH3 binding peptide.

Key words: Crk/C3G/proline-rich sequence/SH3 domain

Introduction

Src homology 3 (SH3) domains are present on many functionally diverse proteins involved in the response of cells to external stimuli (Kuriyan and Cowburn, 1993; Musacchio *et al.*, 1994b). The ability of the SH3 domains to bind to short, specific, linear, proline-rich sequences (Cicchetti *et al.*, 1992) within their binding partners has important functional consequences. SH3 domains may control the cellular localization of their binding partners (Bar-Sagi *et al.*, 1993; Rotin *et al.*, 1994), determine the substrate specificity of enzymes (Feller *et al.*, 1994a) or modulate the catalytic activity of SH3-containing (Liu *et al.*, 1993; Mayer and Baltimore, 1994) as well as SH3 binding proteins (Gout *et al.*, 1993; Pleiman *et al.*, 1994). Despite the abundance of SH3 domains within signaling proteins and their potential biological importance, only a few systems are known where SH3 domains play a direct role in physiological or pathological processes. These include genetic systems that demonstrated the crucial role of complex formation between the adaptor protein Grb2 and the guanine nucleotide exchange factor, Son of Sevenless (SOS) (Rozakis-Adcock *et al.*, 1992; Olivier *et al.*, 1993) and two human diseases that involve loss of function mutations of SH3 domains (Rawlings *et al.*, 1993) or their binding sequences (Cheng *et al.*, 1994).

Identification of two proline-rich motifs in the SH3 binding proteins 3BP1 and 3BP2 was first achieved by expression library screening with the AblSH3 domain (Cicchetti *et al.*, 1992; Ren *et al.*, 1993). Mutagenesis of the 3BP1 sequence pointed to the proline residues at positions 2, 5 and 10 as being crucial in binding to the AblSH3 domain (Ren *et al.*, 1993). Proline-rich motifs were subsequently identified in known SH3 targets (Ren *et al.*, 1994; Seedorf *et al.*, 1994; Tanaka *et al.*, 1994) and found to bind to several SH3 domains with variable affinities (Chardin *et al.*, 1993; Gout *et al.*, 1993; Liu *et al.*, 1993). Only one highly selective interaction, the binding of a proline-rich sequence in $p47^{phox}$ to the C-terminal $p67^{phox}$ SH3 domain has been reported (Finan *et al.*, 1994). More recently, new SH3 binding sequences that are not necessarily found in known proteins have been delineated by screening of combinatorial peptide (Yu *et al.*, 1994) and phage display libraries (Cheadle *et al.*, 1994; Sparks *et al.*, 1994) and provide valuable information about affinity and specificity requirements for SH3 domains.

Structural and mutational analyses of SH3 domain–peptide complexes from Sem-5, p85α, Abl and Fyn (Booker *et al.*, 1993; Lim and Richards, 1994; Musacchio *et al.*, 1994a; Yu *et al.*, 1994) revealed molecular details of interactions between SH3 domains and proline-rich sequences. The SH3 domain core is composed of two perpendicular, antiparallel, three-stranded β-pleated sheets. The hydrophobic surface contains shallow grooves formed by highly conserved aromatic amino acids which contact the proline-rich peptides. This area interacts with two co-planar proline residues from the SH3 binding motif, which forms a polyproline type II helix (Williamson, 1994). A second set of co-planar amino acids in positions next to the two proline residues contacts the hydrophobic surface, as well as the highly variable RT and n-Src loops that emerge from the core. Despite the contacts of these four amino acids, the affinities of SH3 domains for their ligands are in the micromolar range (Cussac *et al.*, 1994; Yu *et al.*, 1994).

The Crk proteins belong to a family of proteins that consist almost entirely of SH2 and SH3 domains, with little intervening sequence (Feller *et al.*, 1994b). This family presently includes v-Crk, two forms of c-Crk proteins (c-Crk-I and c-Crk-II), CRKL, Grb2/ASH from

© Oxford University Press

various species, Grb3–3 and Nck. One way to define signaling pathways that involve the Crk proteins is to identify their SH2 and SH3 domain binding partners. The first CrkSH3 domain [CrkSH3(1)] was previously shown to bind the guanine nucleotide exchange factors SOS and C3G and the tyrosine kinases Abl and Arg (Feller *et al.*, 1994a; Matsuda *et al.*, 1994; Ren *et al.*, 1994; Tanaka *et al.*, 1994). Since Crk−C3G complex formation likely plays an important role in Crk signaling, this interaction was investigated in more detail to determine structural parameters in the peptide sequence and the CrkSH3(1) domain that underlie this highly specific, high affinity interaction. The results showed that the affinity and specificity of C3G for the CrkSH3(1) domain is retained in a nine amino acid peptide. Mutagenic analysis of the first high affinity Crk binding (CB) sequence present in C3G (CB-1 peptide) demonstrated that a single basic amino acid (Lys10) plays a major role in controlling the binding specificity and affinity of the CB-1 peptide for the CrkSH3(1) domain. Differences in the RT loops of c-Crk, v-Crk and CRKL led to the elucidation of acidic amino acids in the SH3 domains that likely interact with Lys10 of the CB-1 peptide. Results from the mutational analyses are consistent with recent studies on SH3−peptide interactions that point to a role of basic amino acids in determining the orientation of SH3 ligands across an SH3 domain (Feng *et al.*, 1994; Tamemoto *et al.*, 1994) and further demonstrate that a single amino acid confers unique specificity of the CB-1 peptide for the CrkSH3(1) domain.

Results

Analysis of the proline-rich CrkSH3(1) binding (CB) sequences in C3G

We have previously identified four similar proline-rich sequences in the guanine nucleotide exchange factor C3G (Tanaka *et al.*, 1994) that can bind individually to the CrkSH3(1) domain (Knudsen *et al.*, 1994). Furthermore, full-length C3G co-precipitates with c-Crk II and v-Crk from cell lysates, suggesting that at least one of these proline-rich sequences binds with high affinity. In addition, when testing a broad panel of SH3 domains only the first CrkSH3 domain bound to C3G in cell lysates, thus pointing to a remarkable specificity of this interaction. These results suggested that further analysis of the proline-rich sequences derived from C3G should provide better insight into affinity and specificity requirements of the first CrkSH3 domain.

Sequence alignment of CB sequences from C3G, c-Abl and SOS led to a preliminary consensus sequence for CrkSH3(1) binding spanning ~10 amino acids (Knudsen *et al.*, 1994). A sequence of 10 amino acids that corresponded to the first CB region on C3G (CB-1 peptide) and contained the CB consensus was therefore expressed as a glutathione S-transferase (GST) fusion peptide and shown to precipitate Crk from cell lysates (data not shown), indicating that a 10 amino acid sequence is sufficient for high affinity binding. Therefore, 10 amino acid peptides identical to the four CB sequences within C3G (CB-1−CB-4) were compared for binding to the first CrkSH3 domain by inhibition studies in cell lysates and fluorescence measurements (Figure 1A). All four peptides were effective in inhibiting the binding of a 12 amino acid GST fusion peptide (GST1–12, amino acids 278–291 in C3G) to c-Crk II in cell lysates (Figure 1A). For each of the peptides, the degree of inhibition in the cell lysate correlated with the K_d obtained by the fluorescence measurements. The peptide derived from the first Crk binding site within the C3G protein (CB-1 peptide) bound with the highest affinity to the CrkSH3(1) domain (K_d = 1.89 ± 0.06 μM).

Determination of essential elements for high affinity and specificity binding to the CrkSH3(1) domain

Since the CB-1 sequence bound with highest affinity, it was used for further analysis. To determine the minimal sequence that bound with high affinity to the CrkSH3(1) domain, we used a filter binding assay (Feller *et al.*, 1995). A series of truncated GST fusion peptides were Western blotted and probed for binding to [^{35}S]GST−CrkSH3(1) (Figure 1B). There was no difference in the amount of bound [^{35}S]GST−CrkSH3(1) to the 12, 10 and 9 amino acid GST fusion peptides, as long as they contained the Arg residue at the C-terminus. Deletion of the Arg or further truncation from the N-terminus diminished the binding to the [^{35}S]CrkSH3(1) probe. In addition, mutations of the two Pro from the Pro-X-X-Pro motif, the highly conserved Lys or the C-terminal Arg residue greatly reduced the binding of the mutated peptides to the CrkSH3(1) domain, demonstrating that these positions are crucial in the interaction.

Since a great preference of full-length C3G for the CrkSH3(1) domain was previously observed, binding of a ^{35}S-labeled nine amino acid GST fusion peptide (GST4–12) to a panel of SH3 domains was tested. In a filter binding assay the nine amino acid GST fusion peptide only bound strongly to the first CrkSH3 domain (Figure 2). While a weak interaction was detectable with the SH3 domains of Grb2, p85α/PI3-kinase, Fyn and α-spectrin in the filter binding assay, previous experiments showed that these SH3 domains do not form stable complexes with C3G from cell lysates (Knudsen *et al.*, 1994). In summary, these experiments are consistent with a remarkable specificity of C3G for the CrkSH3(1) domain and further demonstrate that the specificity is maintained in a nine amino acid sequence derived from CB-1.

Alanine scan through the 10 amino acid high affinity binding peptide

To comprehensively analyze the importance of each position within the CB-1 peptide with regard to binding to the CrkSH3(1) domain, an alanine scan was performed (Figure 3). Each mutant synthetic peptide was tested for its ability to inhibit binding of the wild-type GST1–12 peptide to c-Crk in cell lysates. As expected, mutations in either Pro of the Pro-X-X-Pro motif greatly reduced the ability of the mutant peptides to compete for binding of the wild-type GST1–12 peptide to the c-Crk protein. In addition, a dramatic decrease in the binding affinity was observed when Leu7 (the amino acid numbering in the peptides corresponds to the numbers shown in Figure1B for the GST1-12 peptide) was mutated. Weak but significant reductions in affinity were also observed with mutations in the conserved Lys10 and Arg12 residues. The Lys10Ala

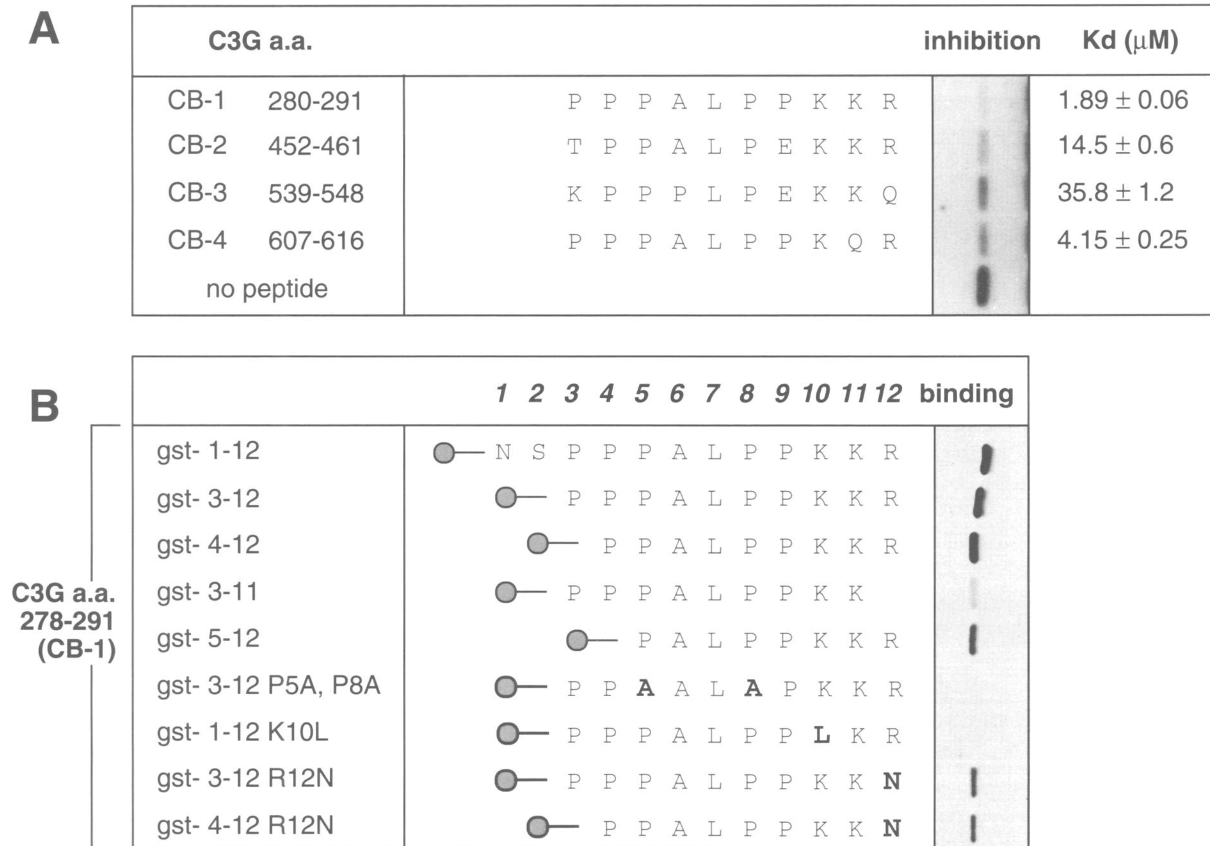

Fig. 1. (**A**) Affinities of the four proline-rich sequences in C3G for the CrkSH3(1) domain. C3G-derived peptides (CB-1–CB-4) were used to inhibit binding of the GST-1–12 peptide to c-Crk from lysates of chick embryo fibroblasts (CEF)s over-expressing c-Crk at a 1:500 molar ratio of GST1–12 to peptide. GST1–12–c-Crk complexes were precipitated with glutathione beads, blotted and probed with Crk antiserum. The bottom lane in the autoradiogram from a representative peptide inhibition study shows the amount of Crk precipitated in the absence of competitor peptide. Binding affinities of the peptides to GST–CrkSH3(1) were determined by tryptophan fluorescence, as described in Materials and methods. (**B**) Binding of CB-1–GST fusion peptides to the CrkSH3(1) domain. Wild-type or mutant sequences derived from the CB-1 region of C3G (CB-1 peptide and derivatives) were expressed as GST fusion peptides, as described under Materials and methods. The numbering of amino acid positions indicated is used throughout the paper. For the SH3 binding assay, 1 μg GST fusion peptide was blotted and the blot incubated with [^{35}S]GST–CrkSH3(1) (35 000 d.p.m./μg) at a concentration of 2 μg/ml in blocking buffer. After washing, binding of [^{35}S]GST–CrkSH3(1) was visualized by autoradiography for 12 h.

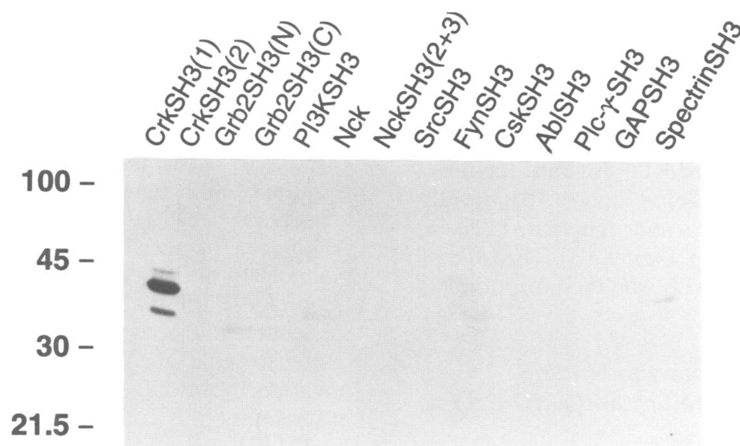

Fig. 2. Binding analysis of a nine amino acid GST fusion peptide derived from CB-1 to a panel of SH3 domains. The SH3 domains indicated above each lane were expressed as GST fusion proteins. GST–CrkSH3(1) (1 μg) or equimolar amounts of the other GST fusion proteins were blotted and probed with [^{35}S]GST4–12 (shown in Figure 1B). The 3 day exposure of the autoradiogram in this experiment allowed detection of additional SH3 domains that weakly bound to the nine amino acid GST fusion peptide.

mutant showed a considerable loss of binding affinity, but less than expected from the Lys10Leu mutation in the GST fusion peptide (Figure 1B), where a large reduction in binding was obtained. The K_d values for both mutant peptides were therefore measured fluorometrically. While the K_d was increased 11.7-fold with Lys10Ala, the K_d

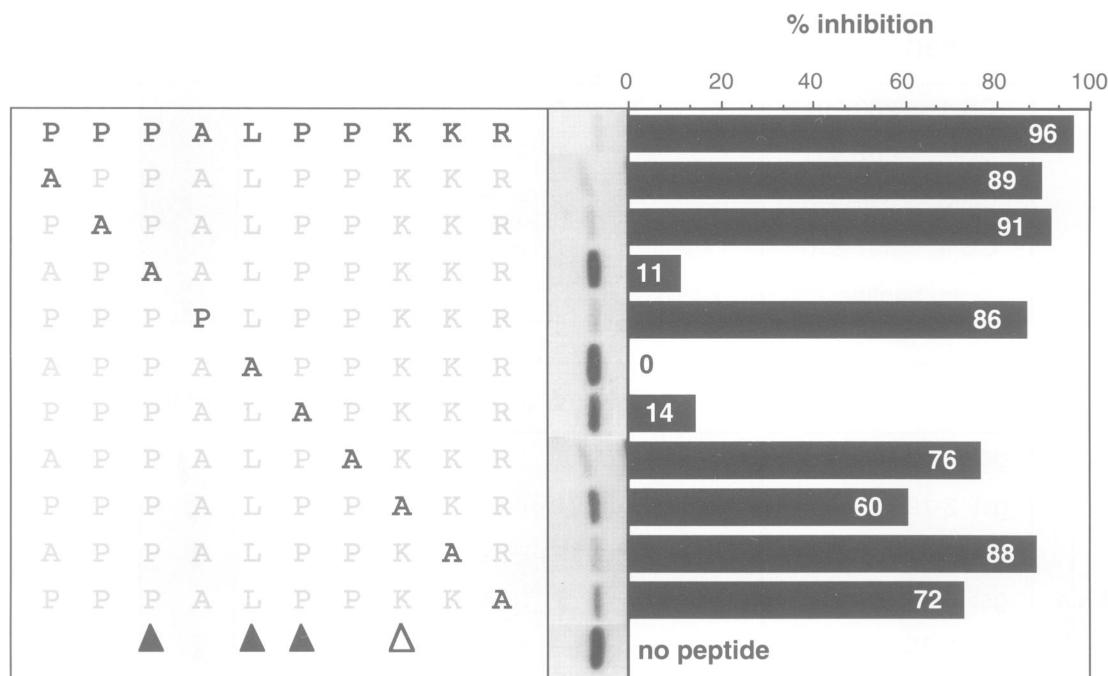

Fig. 3. Alanine scan through the CB-1 peptide. The synthetic peptides shown were used in competition studies, as described under Materials and methods. The filled arrowheads indicate the positions where a mutation to Ala almost entirely abolished binding. The open arrowhead points to the Lys10Ala mutation that significantly reduced binding. A representative blot is shown. Each synthetic peptide was used in a 500-fold molar excess to inhibit the binding of 1 μM GST1–12 to c-Crk from lysates of CEFs over-expressing c-Crk. Gst1–12–c-Crk complexes were precipitated with glutathione beads, blotted and detected with the Crk antibody. The last lane represents c-Crk precipitated in the absence of competitor peptide. The autoradiogram of the blot was scanned and the competition for each peptide is indicated as percent inhibition of GST1–12–c-Crk complex formation in the bar graph. Data show one out of four representative experiments.

increased 200-fold with the Lys10Leu mutant (see Table II), which is consistent with the results obtained from the peptide inhibition study.

Effects of other domains on CrkSH3(1) binding

To rule out the possibility that sequences outside the CrkSH3(1) domain have a significant impact on CrkSH3(1)–peptide interactions, the affinities of several murine constructs, all containing the CrkSH3(1) domain for the CB-1-derived peptide, were compared (Figure 4B). The presence of CrkSH2 or the second CrkSH3 domain had little effect on the affinity of CrkSH3(1) for the CB-1 peptide, nor did the GST tag significantly influence the affinity of the isolated CrkSH3(1) domain [compare GST–CrkSH3(1) with CrkSH3(1), Figure 4B] or full-length murine Crk for the CB-1 peptide [compare GST–c-Crk (mouse) with c-Crk (mouse), Figure 4B]. The isolated CrkSH3(1) domain fused to GST was therefore used in subsequent binding studies.

A single amino acid influences the binding specificity of a peptide to the Crk and Grb2 SH3 domains

Despite the large number of published SH3 domain NMR and crystal structures, essential features of physiologically important, highly specific SH3 ligands are still poorly understood. Interaction of a Pro-X-X-Pro motif with the hydrophobic SH3 domain surface occurs in the majority of the SH3 domains currently analyzed and cannot by itself account for the binding specificity. Consequently, amino acids outside the Pro-X-X-Pro motif likely control specificity by their unique interactions. Comparison of binding sequences from C3G, which uniquely bind the CrkSH3(1) domain, with sequences from the more promiscuous SOS protein, which binds to the SH3 domains of Grb2 (Olivier *et al.*, 1993), Crk, Nck and Src (S.M.Feller *et al.*, in preparation), pointed to several potentially important differences in the binding peptides. Although all C3G and SOS sequences in Table I contain a Pro-X-ψ-Pro-Pro-β (ψ hydrophobic, β basic) motif, the binding affinities vary (Table I). The only non-conservative difference between the CB-1 wild-type peptide and the SOS-1 peptide, an Ala3 (CB-1 wild-type) to Pro (SOS-1) change, did not effect the binding specificity, as demonstrated by the affinities of the C3Gmut(1) peptide for the Crk and Grb2 SH3 domains (Table I), suggesting that another position controls the binding specificity. Next we analyzed conservative changes between the CB-1 and SOS peptides. Alignment of the four proline-rich SOS-1 sequences pointed to a conserved Arg (Table I). Consequently, fluorometric affinity measurements showed that these SOS sequences bound to Grb2 better than to the CrkSH3(1) domain (Table I). We therefore instituted a Lys10Arg mutation in the CB-1 sequence to generate C3Gmut(2). Surprisingly, this conservative mutation increased the K_d for CrkSH3 ~9-fold, whereas the K_d for Grb2 decreased from 142 ± 3 to 23.5 ± 0.8 μM. These data suggest that the Lys10 residue of the CB-1 sequence plays a crucial role in controlling the binding specificity of the CB-1 peptide for the CrkSH3(1) domain.

Interaction of acidic residues in the CrkSH3 domains with basic peptide residues

The interaction of Lys10 in the CB-1 sequence with the CrkSH3(1) domain was further investigated. The NMR

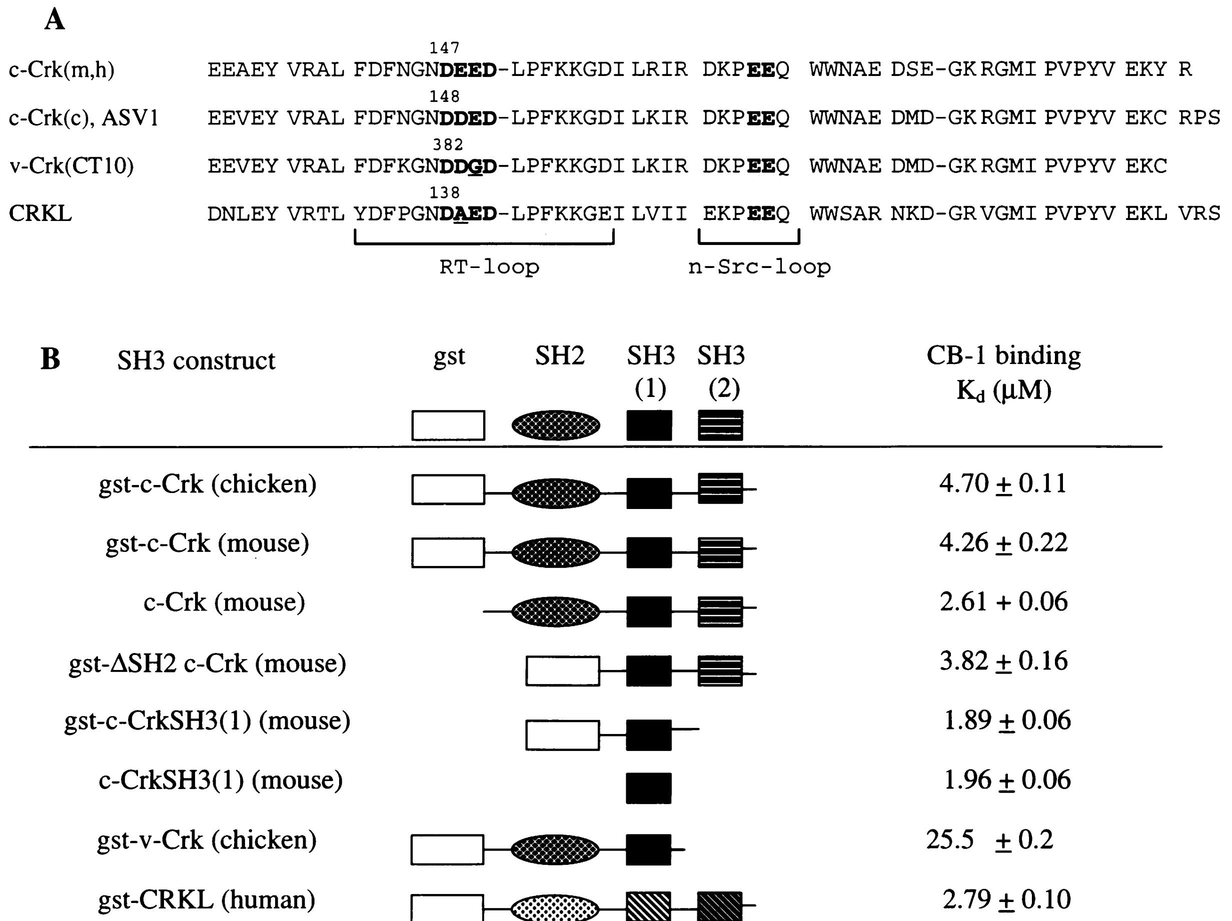

Fig. 4. (**A**) Alignment of Crk/CRKL SH3 domains. Sequences of the first SH3 domains of mouse [m] (Ogawa *et al.*, 1994), human [h] (Margolis *et al.*, 1992) and chicken [c] (Reichman *et al.*, 1992) c-Crk-II, v-Crk (CT10 virus; Mayer *et al.*, 1988) and CRKL (ten Hoeve *et al.*, 1993) were aligned by following the likely boundaries of structural elements (Cowburn, 1994). The murine and human CrkSH3(1) domains have identical sequences. ASV-1 (Tsuchie *et al.*, 1989), another independently isolated viral v-Crk, is identical to chicken Crk-II and did not show the point mutations found in the CT10 virus. The bold amino acids represent the acidic clusters in the RT loop and the n-Src loop. The numbers above the RT loop indicate the position of the first Asp of the acidic cluster. (**B**) Binding of CB-1 peptide to c-Crk, v-Crk and CRKL SH3 domains. The K_d values were calculated from fluorescence measurements, as described under Materials and methods.

structure of p85αSH3 with a high affinity binding peptide suggested that basic peptide residues interact with acidic residues on the loops of the SH3 domain (Yu *et al.*, 1994). The RT loops of the Crk/CRKL SH3(1) domains contain a three (v-Crk, CRKL) or four (c-Crk) amino acid acidic cluster, whereas the n-Src loop contains two adjacent Glu residues in all cases (Figure 4). In the RT loop the two Asp residues, Asp148 and Asp151 in chicken c-Crk (Reichman *et al.*, 1992) are conserved amongst the aligned Crk/CRKL SH3 domains. The second position shows a conservative change between chicken and mouse and an Ala in CRKL (Figure 4A) without a significant loss of affinity for the CB-1 peptide (Figure 4B). In v-Crk, the third acidic residue is mutated from Glu to Gly. This mutation reduces the affinity of the v-Crk SH3 domain to the CB-1 peptide by ~14-fold, suggesting that Glu150 in chicken c-Crk and the corresponding amino acids in mouse c-Crk and CRKL are involved in the binding to basic peptide residues.

The complexity of interaction of this acidic cluster on the RT loop with basic peptide residues became apparent in binding studies with mutant peptides (Table II). The Lys10 residue was mutated to Arg, Ala or Leu and the affinity of these mutant peptides for c-CrkSH3(1), v-Crk and CRKL was determined. A Lys10Ala mutation eliminated the positive charge and reduced the binding affinity 11.6-fold for c-CrkSH3(1) and 9.7-fold for v-Crk. In contrast, a conservative mutation from Lys10 to Arg only decreased the binding of the mutant peptide to c-Crk and not to v-Crk. Binding to CRKL of this mutant peptide was 4-fold reduced. The SOS-1 peptide resembles the mutant peptide, since it also contains an Arg at the position

Table I. Basic amino acids of a proline-rich motif determine the binding specificity to the first Crk and to the Grb2 SH3 domains

Peptide		c-CrkSH3(1) K_d (μM)	Grb2SH3 K_d (μM)
Origin	Sequence		
CB-1/wt	PPPALPP**K**KR	1.89 ± 0.06	142 ± 3
SOS-1	PPPVPPP**R**RRR	5.24 ± 0.16	3.54 ± 0.16
SOS-2	PPAIPP**R**QPT	127 ± 5	42 ± 4
SOS-3	PPLLPP**R**EPV	206 ± 9	88 ± 3
SOS-4	GPPVPP**R**QST	663 ± 3	94 ± 5
C3Gmut(1)	PPPPLPP**K**KR	1.45 ± 0.03	203 ± 3
C3Gmut(2)	PPPALPP**R**KR	17.2 ± 0.04	23.5 ± 0.8

Table II. Point mutations in the v-Crk and CRKL SH3 domains affect the binding to the basic peptide residues

Peptide	gst-c-CrkSH3(1) K_d (μM)	gst-v-Crk K_d (μM)	gst-CRKL K_d (μM)
PPPALPPKKR	1.89 ± 0.06	25.5 ± 0.2	2.79 ± 0.1
PPPALPP**R**KR	17.2 ± 0.4	23.5 ± 0.8	11.2 ± 0.5
PPPALPP**A**KR	22.1 ± 0.5	248 ± 10	
PPPALPP**L**KR	379 ± 1	437 ± 6	
PPPALPPK**A**R	5.74 ± 0.14		
PPPALPPKK**A**	9.3 ± 0.2	71.8 ± 1.1	
PPPALPPKK**K**	5.0 ± 0.3	36.2 ± 0.8	
PPPVPPRRRR	5.24 ± 0.16	23.7 ± 0.6	10.5 ± 0.8

corresponding to Lys10. Accordingly, binding of the SOS-1 peptide to v-Crk and CRKL occurred with affinities similar to the CB-1 Lys10Arg mutant [C3G mut(2), Table II]. However, the SOS-1 peptide bound to c-CrkSH3(1) with a K_d of 5.24 ± 0.16 μM, which is a lower K_d than expected from the binding affinity of the Lys10Arg mutant, suggesting that the loss of affinity caused by the Arg is compensated for by binding of other peptide residues.

Lys11 in the CB-1 peptide is poorly conserved amongst all CrkSH3(1) binding sequences. Accordingly, a mutation to Ala did not significantly reduce the binding affinity (Table II and Figure 3). The role of the conserved Arg residue at the C-terminus of the peptide remains poorly understood. An Arg is present in all high affinity Crk binding sequences derived from various proteins, but mutations of this residue do not greatly affect the affinities of the mutated peptides (Table II and Figure 3). However, a deletion of Arg12 in the GST fusion peptide abolished binding almost entirely (Figure 1). Our results from binding studies with GST fusion peptides are consistent with the K_d measurements obtained with synthetic peptides and point to important features in Crk binding sequences that are necessary for a specific, high affinity interaction with the CrkSH3(1) domain.

Discussion

The interaction between the guanine nucleotide exchange factor C3G and Crk proteins likely plays an important role in signaling pathways that involve the c-Crk and v-Crk proteins (Knudsen *et al.*, 1994; Tanaka *et al.*, 1994). We have previously identified four similar proline-rich sequences in the middle of C3G that can bind independently and with high affinities to the first SH3 domain in c- and v-Crk (Knudsen *et al.*, 1994). In addition to the high affinity of this region in C3G for the CrkSH3(1) domain, a unique specificity of these sequences for the CrkSH3(1) domain was noticed. This highly specific interaction was therefore analyzed in more detail to determine its underlying structural features. A peptide derived from the first Crk binding sequence within C3G (CB-1 peptide) bound with the highest affinity (1.89 μM) to the first c-Crk SH3 domain [CrkSH3(1)] (Figure 1A). The affinity for the CrkSH3(1) domain remained unchanged upon truncation of the binding sequence to a nine amino acid GST fusion peptide. Further truncation to an eight amino acid peptide from the N-terminus (Figure 1B, GST5–12) reduced the affinity considerably, whereas deletion of the C-terminal Arg12 (Figure 1B, GST3–11) abolished all binding. The loss of affinity with Arg12 mutants (Figure 1B, GST3–12R12N, GST4–12R12N) also points to the importance of a positively charged amino acid at this position, which is highly conserved amongst Crk binding sequences in proteins (Knudsen *et al.*, 1994). However, we cannot predict whether this Arg contacts the CrkSH3(1) domain or whether it influences the conformation of the binding peptides. In general, the number of amino acids that are necessary for high affinity complex formation with SH3 domains lies between seven and nine (Finan *et al.*, 1994; Yu *et al.*, 1994). Since the design of peptides and structural analogs that inhibit SH3 domain function is one of the goals of this analysis, it is important to demonstrate that the unique specificity of C3G for the CrkSH3(1) domain was retained in the nine amino acid peptide sequence (Figure 2). In contrast to the less specific SOS-1-derived peptide, which bound with similar affinities to the SH3 domains of Grb2 and Crk, the C3G-derived CB-1 peptide bound 75-fold better to the CrkSH3(1) domain than to Grb2 (Table I). Consequently, when introduced into cells, the CB-1 peptide should result in a selective inhibition of Crk signaling pathways, without significantly interfering with Grb2 signaling.

We demonstrated with an alanine scan through the CB-1 sequence the critical role of Leu7 within the Pro-X-Leu-Pro motif in binding to the CrkSH3(1) domain (Figure 3). In addition to the Pro-X-Leu-Pro sequence, a motif that interacts with many SH3 surfaces, we have shown the importance of a lysine residue (Lys10) in controlling binding specificity and affinity of the CB-1 sequence for the CrkSH3(1) domain. Most recently, two groups (Feng *et al.*, 1994; Lim *et al.*, 1994) demonstrated that non-proline residues in SH3 binding motifs determine the orientation of these motifs along the SH3 domain surface. Orientation of the SOS-1 peptide N- to C-terminal across the surface of the Grb2(C)SH3 domain, defined as the 'minus' orientation, occurs through interactions of a Val and an Arg residue (Lim *et al.*, 1994). By sequence analogy, the four C3G-derived sequences should be positioned similarly to the SOS-1 peptide in the 'minus' orientation through the interaction of their corresponding Leu7 and Lys10 residues with the CrkSH3(1) domain, which is not consistent with the orientation suggested by Feng *et al.* (1994) for the C3G-derived sequence. In addition to controlling the orientation of the peptide, Lys10 strongly influences the preference of the CB-1 sequence for the CrkSH3(1) domain. A Lys→Arg mutation at this position decreases the affinity for the CrkSH3(1) domain, but increases the affinity of the peptide for Grb2 (Table I), while the orientation of the peptide is maintained. This observation demonstrates that a single conservative amino acid change in an SH3 binding sequence can effect the binding specificity. The strong preference of the CrkSH3 domain for a Lys at this position is a unique property of the CrkSH3 domain. In contrast, an Arg at this position allows for binding to several SH3 domains, as illustrated by the interaction of the SOS-1 peptide with the SH3 domains of Grb2 and Crk (Table I) and binding of the SOS protein to the SH3 domains of Grb2 (Olivier *et al.*, 1993), Crk, Src and Nck (S.M.Feller *et al.*, in preparation).

The difference in the binding spectrum of the CB-1 and SOS-1 peptides demonstrates that high affinity interactions may be of variable specificity.

Structural studies of p85α and Abl SH3 domains with high affinity binding peptides (Musacchio *et al.*, 1994a; Yu *et al.*, 1994) pointed to an interaction of the SH3 ligand with an Asp residue in the RT loop of the p85αSH3 domain or a Thr residue in the RT loop of the AblSH3 domain, both corresponding to Asp151 in chicken Crk. A Thr at this position is unique to the AblSH3 domain and explains the preference of AblSH3 for an interacting Pro in the 3BP1 peptide. However, the Asp residue is conserved amongst most SH3 domains and by itself does not explain the specificity of p85α binding peptides for the p85αSH3 domain. An Asp is also present at a corresponding position in the SH3 domains of chicken c-Crk (Asp151), v-Crk (Asp386) and CRKL (Asp141) (Figure 4). Surprisingly, our binding studies show a 15-fold difference in binding of the c-Crk and v-Crk SH3 domains to the CB-1 and SOS-1 peptides (Table II). Therefore, additional acidic residues surrounding Asp151 in chicken c-Crk may influence the interaction with basic peptide residues. v-Crk has a Glu→Gly change at position 384, which causes the decrease in binding affinity to the SOS-1 and CB-1 peptides. In agreement with the peptide binding studies, we have previously noticed less binding of C3G to GST−v-Crk than to GST−c-Crk in cell lysates (B.S.Knudsen, unpublished observation). Since Lys10 in the CB-1 peptide corresponds to Arg in the SOS peptide, which forms a salt bridge with Glu172 in the RT loop of the Grb2SH3 domain (Feng *et al.*, 1994), acidic residues in the RT loop and not the n-Src-loop likely contact Lys10. This hypothesis has recently been confirmed upon determination of the crystal structure of the CrkSH3(1)−CB-1 peptide complex (Wu *et al.*, in preparation). Interestingly, a Lys10Arg mutation decreased binding to the c-CrkSH3(1) domain by 10-fold, but had no effect on binding of the mutant peptide to the v-CrkSH3 domain. In contrast, a Lys10Ala mutation decreased binding by 10-fold to both c-Crk and v-Crk, suggesting that the charge of Lys10 plays an important role in binding of the CB-1 peptide to the c-Crk and v-Crk SH3 domains. The structural parameters that explain the similar binding affinities of the Lys10Arg mutant peptide for c-Crk and v-Crk have been elucidated by analyzing the crystal structures of the CrkSH3(1) domain complexed with the CB-1 or SOS-1 peptides and are discussed in detail elsewhere (Wu *et al.*, in preparation).

The specificity of SH3 domain interactions is believed to be important in controlling the formation of protein−protein complexes that allow the spatially and temporally correct propagation of a signal. Our mutational analysis, together with the recently obtained crystal structure, define molecular parameters that underlie the uniquely specific complex formation between the guanine nucleotide exchange factor C3G and Crk proteins. In contrast to C3G, the guanine nucleotide exchange factor SOS binds with high affinity to several SH3 domains. This seemingly low degree of specificity of the SOS protein, which has also been observed with other SH3 binding motifs (Gout *et al.*, 1993; Liu *et al.*, 1993), may either be a functionally important parameter of SH3 domain−ligand interactions in the cells or may be an artifact of the *in vivo* or *in vitro* systems that utilize high concentrations of SH3 domains and binding proteins to demonstrate these interactions. The promiscuity of the binding of proline-rich sequences in SOS to multiple SH3 domains could generate a point of convergence of several receptor pathways to Ras. The biological importance of the unique specificity between Crk and C3G compared with the lesser specificity of the SOS protein needs to be further investigated.

Materials and methods

Proteins expressed in bacteria

GST fusion peptides derived from the C3G CB-1 sequence (amino acids 278–291) were constructed with double-stranded oligonucleotides that possessed a 5′ *Bam*HI and a 3′ *Eco*RI overhang. The oligonucleotides were cloned into *Bam*HI/*Eco*RI cut pGEX-2T. The correct reading frame and sequence of the oligonucleotides was confirmed by sequencing using a UBI sequencing kit. Most GST−SH3 fusion proteins have previously been described (Knudsen *et al.*, 1994). The GST−CRKL construct was a generous gift of J.ten Hoeve and J.Groffen. FynSH3 was obtained from L.Cantley. All SH3 domain GST fusion proteins and ^{35}S-labeled GST fusion proteins were produced as previously described (Feller *et al.*, 1995). Briefly, GST fusion proteins were metabolically labeled in bacteria with 1 mCi [^{35}S]methionine (from a 200 ml culture) in 20 ml methionine-free medium upon induction with 0.1 mM IPTG. GST fusion proteins were isolated with glutathione beads, eluted and dialyzed against phosphate-buffered saline (PBS; 10 mM phosphate, pH 7.0, 150 mM NaCl), containing 0.05% Tween 20 and 5% glycerol. The specific activity was ~35 000 d.p.m./μg GST fusion protein.

Peptide synthesis

Peptides were synthesized by the automated solid phase methodology utilizing an Applied Biosystems Inc. model 431A synthesizer programed with the manufacturer's standard fluorenylmethoxycarbonyl (FMOC) single coupling protocol. Pre-loaded resins, FMOC-protected amino acids and other pre-packaged reagents were purchased from Applied Biosystems Inc. Cleavage from the resin support and simultaneous side chain deprotection was accomplished with a 4 h treatment with 90% trifluoroacetic acid, 2.5% thioanisole, 2.5% 2-mercaptoethanol, 5% phenol at room temperature. The precipitated, crude peptides were purified to >95% homogeneity using preparative HPLC chromatography. Characterization consisted of analytical HPLC, amino acid analysis and electrospray mass spectrometry. Peptides were resuspended in Tris-buffered saline (TBS; 10 mM Tris−HCl, pH 7.5, 150 mM NaCl) containing 0.05% Tween 20 at 5 mg peptide/ml for all measurements.

Fluorescence measurements

The fluorescence measurements are based predominantly on interactions of the peptides with the aromatic residues, predominantly Trp (Yu *et al.*, 1994), in the CrkSH3 domain. Measurements were made with a Perkin-Elmer 760-40 fluorescence spectrophotometer at a excitation wavelength of 290 nm (slit width 2 nm) and an emission wavelength of 345 nm (slit width 17 nm). A mini magnetic stirrer (CUV-O-STIR model 333; HELLMA, Long Island, NY) was used to mix the solution in a 1 cm^2 quartz fluorescence cell. A circulating water bath was used to maintain the sample temperature at 18°C. To obtain the titration curves for calculation of the binding constants, peptides from a stock solution of 5 mg/ml in TBS, 0.05% Tween 20 were added in small increments to 1 ml 0.5 μM SH3 domain in PBS, 1 mM dithiothreitol (DTT). Upon addition of the peptide solutions, maximal changes in fluorescence of between 25 and 40% were observed.

Calculation of the binding constant

Since the concentration of the SH3 domain-containing protein was low, the experimental data were fitted to the following equation:

$$F = F_{max} \times [\text{peptide}]/(K_d + [\text{peptide}]) \quad (1)$$

where [peptide] is the final peptide concentration at each data point, F the measured protein fluorescence intensity at the particular peptide concentration and F_{max} the observed maximal fluorescence intensity of the protein when saturated with the peptide. Non-linear regression curve fitting using SigmaPlot (Jandel Scientific, San Rafael, CA) was carried out to fit the experimental data to Equation 1, with F_{max} and K_d as fitted

parameters. The change in protein concentration as a result of addition of the peptide was properly corrected.

SH3 filter binding assay

The SH3 filter binding assay was carried out as described previously (Feller *et al.*, 1995). Briefly, ~30 pmol of each GST fusion peptide or GST fusion SH3 domain were blotted and incubated at room temperature overnight in a buffer (TBS, 0.05% Tween 20, 1 mM DTT, 2% bovine serum albumin, 1% ovalbumin) to allow renaturation. The blot with the GST fusion peptides was probed with 1–2 μg/ml ^{35}S-labeled GST–CrkSH3(1), whereas the blot containing the SH3 domains was probed with ^{35}S-labeled CB-1–GST4–12 fusion peptide at 1–2 μg/ml. The blots were washed three times with TBS, 0.05% Tween 20, 1 mM DTT, 0.2% bovine serum albumin, 0.1% ovalbumin and bound probe was visualized by autoradiography.

Peptide competition assay

Competition of peptides for binding to the CrkSH3 domain was carried out in total cell lysates from chicken embryo fibroblasts over-expressing c-Crk upon infection with c-Crk virus (Reichman, 1993). Cell lysate (1 mg) in RIPA buffer (TBS, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS) was diluted 1:1 with a buffer (TBS, 0.05% Tween 20) containing 2 mM DTT, 2% bovine serum albumin, 1% ovalbumin and protease inhibitors (20 μg/ml aprotinin, 10 mM phenylmethylsulfonyl fluoride, 10 μM leupeptin, 10 μM antipain, 10 μM pepstatin). CB-1–GST1-12 (1 μM) was added to each tube and competing peptides were added at a concentration of 500 μM. CB-1–GST1-12 peptide was precipitated with glutathione beads, washed four times with RIPA buffer and subjected to SDS–PAGE. After blotting, c-Crk bound to the CB-1–GST1-12 fusion peptide was detected with Crk antiserum (Mayer and Hanafusa, 1990). To quantitate the amount of c-Crk bound to CB-1–GST1-12, each lane was scanned with a gel scanner. Inhibition of c-Crk CB-1–GST1-12 binding by the competing peptide was then calculated as follows:

$$\text{percent binding} = 100\% - [\text{peak(competitor peptide)} \times 100] / [\text{peak(no peptide)}].$$

Acknowledgements

We thank Rosemary Williams for excellent technical assistance. We also thank Zoe Gerard and Joshua Derman for their help in the expression and purification of GST fusion peptides. We wish to thank Xiaodong Wu and John Kuriyan for many helpful discussions related to the CrkSH3(1) crystal structure. The generous gift of the GST–CRKL construct by Johanna ten Hoeve and John Groffen was greatly appreciated. We thank Lewis Cantley for the GST–FynSH3 construct and Ann Marie Pendergast for the GST–Grb2 constructs. This study was supported by grant VM2 from the American Cancer Society and research grant CA44356 and training grant CA09673 from the National Cancer Institute. S.M.F. is supported by a fellowship from the Revson foundation.

References

Bar-Sagi,D., Rotin,D., Batzer,A., Mandiyan,V. and Schlessinger,J. (1993) *Cell*, **74**, 83–91.

Booker,G.W., Gout,I., Downing,A.K., Driscoll,P.C., Boyd,J., Waterfield,M.D. and Campbell,I.D. (1993) *Cell*, **73**, 813–822.

Chardin,P., Camonis,J.H., Gale,N.W., Van Aelst,L., Schlessinger,J., Wigler,M.H. and Bar-Sagi,D. (1993) *Science*, **260**, 1338–1343.

Cheadle,C., Ivashchenko,Y., South,V., Searfoss,G.H., French,S., Howk,R., Ricca,G.A. and Jaye,M. (1994) *J. Biol. Chem.*, **269**, 24034–24039.

Cheng,G., Ye,Z.-S. and Baltimore,D. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 8152–8155.

Cicchetti,P., Mayer,B.J., Thiel,G. and Baltimore,D. (1992) *Science*, **257**, 803–806.

Cowburn,D. (1994) *Nature Struct. Biol.*, **1**, 489–491.

Cussac,D., Frech,M. and Chardin,P. (1994) *EMBO J.*, **13**, 4011–4021.

Feller,S.M., Knudsen,B. and Hanafusa,H. (1994a) *EMBO J.*, **13**, 2341–2351.

Feller,S.M., Ren,R., Hanafusa,H. and Baltimore,D. (1994b) *Trends Biochem. Sci.*, **19**, 453–459.

Feller,S.M., Knudsen,B., Wong,T.W. and Hanafusa,H. (1995) *Methods Enzymol.*, in press.

Feng,S., Chen,J.K., Yu,H., Simon,J.A. and Schreiber,S.L. (1994) *Science*, **266**, 1241–1247.

Finan,P., Shimizu,Y., Gout,I., Hsuan,J., Truong,O., Butcher,C., Bennett,P., Waterfield,M.D. and Kellie,S. (1994) *J. Biol. Chem.*, **269**, 13752–13755.

Gout,I. *et al.* (1993) *Cell*, **75**, 25–36.

Knudsen,B.S., Feller,S.M. and Hanafusa,H. (1994) *J. Biol. Chem.*, **269**, 32781–32787.

Kuriyan,J. and Cowburn,D. (1993) *Curr. Opin. Struct. Biol.*, **3**, 828–837.

Lim,W.A. and Richards,F.M. (1994) *Nature Struct. Biol.*, **1**, 221–225.

Lim,W.A., Richards,F.M. and Fox,R.O. (1994) *Nature*, **372**, 375–379.

Liu,X., Marengere,L.E.M., Koch,C.A. and Pawson,T. (1993) *Mol. Cell. Biol.*, **13**, 5225–5232.

Margolis,B., Silvennoinen,O., Comoglio,F., Roonprapunt,C., Skolnik,E., Ullrich,A. and Schlessinger,J. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 8894–8898.

Matsuda,M., Hashimoto,Y., Muroya,K., Hasegawa,H., Kurata,T., Tanaka,S., Nakamura,S. and Hattori,S. (1994) *Mol. Cell. Biol.*, **14**, 5495–5500.

Mayer,B.J. and Baltimore,D. (1994) *Mol. Cell. Biol.*, **14**, 2883–2894.

Mayer,B.J. and Hanafusa,H. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 2638–2642.

Mayer,B.J., Hamaguchi,M. and Hanafusa,H. (1988) *Cold Spring Harbor Symp. Quant. Biol.*, **LIII**, 907–914.

Musacchio,A., Sarastre,M. and Wilmanns,M. (1994a) *Nature Struct. Biol.*, **1**, 546–551.

Musacchio,A., Wilmanns,M. and Sarastre,M. (1994b) *Prog. Biophys. Mol. Biol.*, **61**, 283–297.

Ogawa,S. *et al.* (1994) *Oncogene*, **9**, 1669–1678.

Olivier,J.P., Raabe,T., Henkemeyer,M., Dickson,B., Mbamalu,G., Margolis,B., Schlessinger,J., Hafen,E. and Pawson,T. (1993) *Cell*, **73**, 179–191.

Pleiman,C.M., Hertz,W.M. and Cambier,J.C. (1994) *Science*, **263**, 1609–1612.

Rawlings,D.J. *et al.* (1993) *Science*, **261**, 358–361.

Reichman,C.T. (1993) PhD Thesis, Rockefeller University, New York, NY.

Reichman,C.T., Mayer,B.J., Keshav,S. and Hanafusa,H. (1992) *Cell Growth Differentiat.*, **3**, 451–460.

Ren,R., Mayer,B.J., Cicchetti,P. and Baltimore,D. (1993) *Science*, **259**, 1157–1161.

Ren,R., Ye,Z.-S. and Baltimore,D. (1994) *Genes Dev.*, **8**, 783–795.

Rotin,D., Bar-Sagi,D., O'Brodovich,H., Merilainen,J., Lehto,V.P., Canessa,C.M., Rossier,B.C. and Downey,G.P. (1994) *EMBO J.*, **13**, 4440–4450.

Rozakis-Adcock, M. *et al.* (1992) *Nature*, **360**, 689–692.

Seedorf,K., Kostka,G., Lammers,R., Bashkin,P., Daly,R., Burgess,W.H., Van der Bliek,A.M., Schlessinger,J. and Ullrich,A. (1994) *J. Biol. Chem.*, **269**, 16009–16014.

Sparks,A.B., Quilliam,L.A., Thorn,J.M., Der,C.J. and Kay,B.K. (1994) *J. Biol. Chem.*, **269**, 23853–23856.

Tamemoto,H. *et al.* (1994) *Nature*, **372**, 182–186.

Tanaka,S. *et al.* (1994) *Proc. Natl Acad. Sci. USA*, **91**, 3443–3447.

ten Hoeve,J., Morris,C., Heisterkamp,N. and Groffen,J. (1993) *Oncogene*, **8**, 2469–2474.

Tsuchie,H., Chang,C.H.W., Yoshida,M. and Vogt,P.K. (1989) *Oncogene*, **4**, 1281–1284.

Williamson,M.P. (1994) *Biochem. J.*, **297**, 249–260.

Yu,H., Chen,J.K., Feng,S., Dalgarno,D.C., Brauer,A.W. and Schreiber, S.L. (1994) *Cell*, **76**, 933–945.

Received on December 21, 1994; revised on February 14, 1995