# Intramolecular interactions of the regulatory domains of the Bcr–Abl kinase reveal a novel control mechanism

Hyun-Joo Nam[1,2], Wayne G Haser[3,4], Thomas M Roberts[3,4] and Christin A Frederick[1,5]*

**Background:** The Abl nonreceptor tyrosine kinase is implicated in a range of cellular processes and its transforming variants are involved in human leukemias. The N-terminal regulatory region of the Abl protein contains Src homology domains SH2 and SH3 which have been shown to be important for the regulation of its activity *in vivo*. These domains are often found together in the same protein and biochemical data suggest that the functions of one domain can be influenced by the other.

**Results:** We have determined the crystal structure of the Abl regulatory region containing the SH3 and SH2 domains. In general, the individual domains are very similar to those of previously solved structures, although the Abl SH2 domain contains a loop which is extended so that one side of the resulting phosphotyrosine-binding pocket is open. In our structure the protein exists as a monomer with no intermolecular contacts to which a biological function may be attributed. However, there is a significant intramolecular contact between a loop of the SH3 domain and the extended loop of the SH2 domain. This contact surface includes the SH2 loop segment that is responsible for binding the phosphate moiety of phosphotyrosine-containing proteins and is therefore critical for orienting peptide interactions.

**Conclusions:** The crystal structure of the composite Abl SH3–SH2 domain provides the first indication of how SH2 and SH3 domains communicate with each other within the same molecule and why the presence of one directly influences the activity of the other. This is the first clear evidence that these two domains are in contact with each other. The results suggest that this direct interaction between the two domains may affect the ligand binding properties of the SH2 domain, thus providing an explanation for biochemical and functional data concerning the Bcr–Abl kinase.

Addresses: [1]Department of X-ray Crystallography, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA [2]Department of Cellular and Developmental Biology, Harvard Medical School, Boston, MA 02115, USA [3]Department of Cellular and Molecular Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA [4]Department of Pathology, Harvard Medical School, Harvard Boston, MA 02115, USA and [5]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.

*Corresponding author.
E-mail: CAF@Amber.DFCI.Harvard.edu

Key words: Bcr–Abl, protein kinase, SH2/SH3 domains, X-ray crystallography

Received: 21 June 1996
Revisions requested: 18 July 1996
Revisions received: 5 August 1996
Accepted: 7 August 1996

Structure 15 September 1996, 4:1105–1114



## Introduction

Protein tyrosine kinases are central components in the regulation of cellular growth, differentiation, and transformation in complex eukaryotes. The two major classes are the transmembrane receptor kinases and the predominantly cytoplasmic nonreceptor kinases. The role of the receptor kinases in transmitting an extracellular signal through the binding of a growth factor or cytokine has been well documented [1], while the role of nonreceptor kinases in signal transduction is less well understood. The Abl nonreceptor kinase is an important example of one such protein whose role in a range of cellular processes has made it the subject of extensive biochemical investigations [2]. Abl was so-called because it was originally found to be the transforming gene associated with the Abelson murine leukemia virus [3]. Subsequently it was discovered that, by a chromosomal rearrangement, N-terminal parts of the *abl* gene sequence can be substituted with a breakpoint cluster region (*bcr*) gene to result in the expression of a fusion gene, *bcr–abl*, involved in human leukemias [4–6]. In general, structural alterations in the Abl protein result in efficient cellular transformation and elevated levels of tyrosine kinase activity. In contrast to the transforming variants, overexpression of the cellular homologue of the viral gene does not result in increased levels of phosphorylation or cellular transformation [7,8].

Although Abl is not closely related to the Src family of kinases, its N-terminal regulatory region does contain an Src homology 3 (SH3) and an Src homology 2 (SH2) domain. These consensus sequences of approximately 60 and 100 amino acid residues, respectively, mediate protein–protein interactions in signal transduction pathways [9,10]. It is through the interactions of these domains that the activities of both the catalytic and non-catalytic components of the signal cascades are regulated. SH2 domains have been shown to have varying affinities for phosphotyrosine (pTyr) containing peptide sequences

whereas the SH3 domains seem to prefer proline-rich segments [11,12]. Several proteins which bind to the Abl SH3 domain have been isolated and these proline-rich binding motifs identified [13,14]. These proteins have not been shown to bind Abl *in vivo* [15]. Although the Abl SH2 domain has been shown to bind to cellular phosphoproteins [16] and experimental results using a degenerate peptide library demonstrated a high affinity for peptides with a pTyr-Glu-Asn-Pro motif [17], thus far no specific protein ligands for the Abl SH2 domain have been isolated. The Abl SH3 domain is implicated in repression of the transforming activity of the Abl kinase. This repressive activity of the Abl SH3 domain was shown to be sensitive to its relative position within the entire molecule [15]. In contrast to the SH3 domain, the SH2 domain is required for efficient cellular transformation, and its activity is relatively position independent [18].

SH2 and SH3 domains are often found together in signalling molecules, and many proteins, including Abl and Src family kinases, contain tandem SH3–SH2 modules. Biochemical data have shown that the activities of these tandem domains are coordinated in some Src family members and that the SH3 domain of Src is also involved in regulation of its internal kinase activity [19,20]. As the Abl kinase does not contain internal phosphorylation sites, its mode of regulation is expected to differ from that of members of the Src family. The crystal structure of the Abl regulatory region bearing the SH3 and SH2 domains was determined in an attempt to understand the nature of the interaction between these two domains and ultimately the complex regulation of the activities of this important kinase.

## Results and discussion

### Structure of Abl SH3–SH2 domains

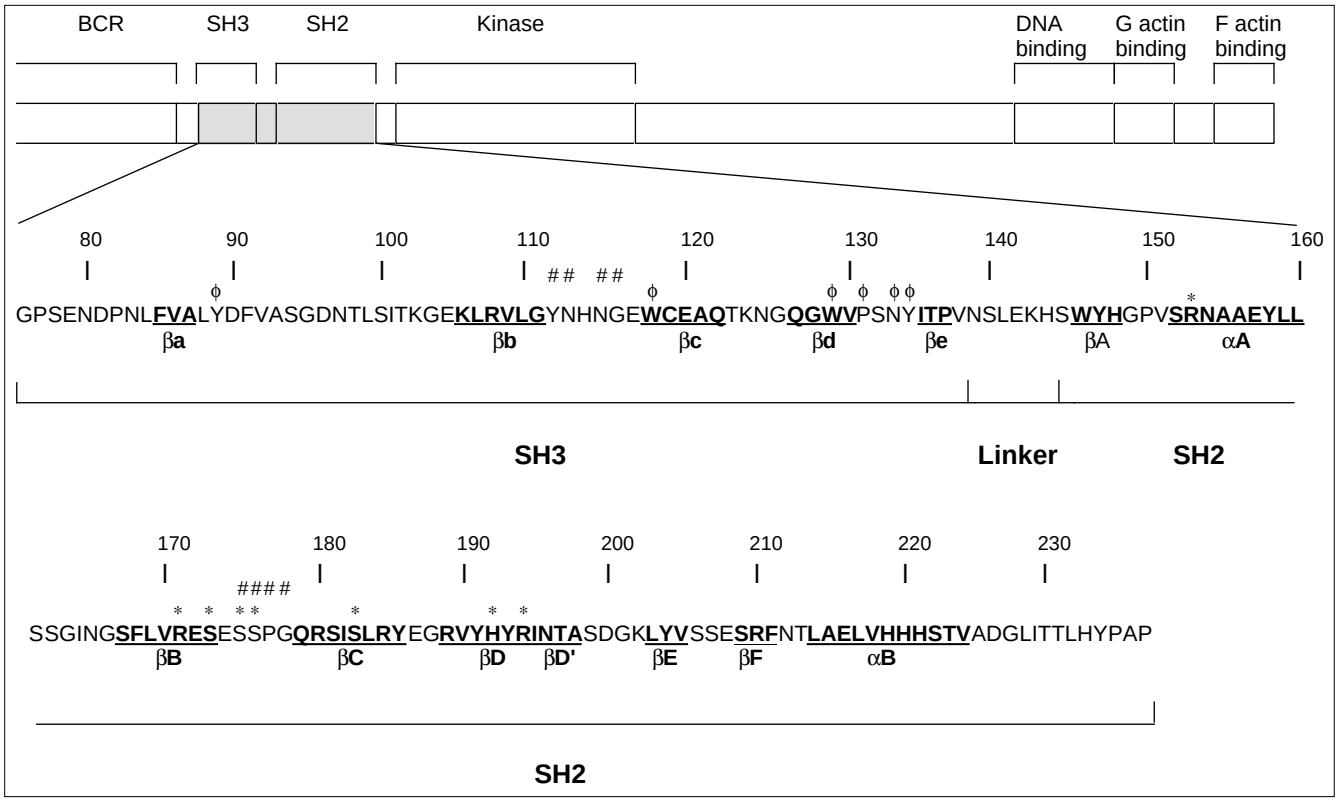
The regulatory region of the Abl kinase for which the three-dimensional (3D) structure has been determined includes the SH3 and SH2 domains connected by a seven residue linker and flanked on either end by several additional residues (Fig. 1). In this structure of the tandem domains (Fig. 2a), the SH2 and SH3 domains each still fold as individual compact units, as predicted from sequence comparisons, and as seen in the previous X-ray and NMR structures of the Abl SH3 and SH2 domains, respectively [21,22]. As expected, each domain looks very similar to other known structures in its own class. The SH3 domain is essentially a β sandwich; it consists of five β strands (βa–βe), and a long β hairpin-like turn (the ab loop). Although the precise boundaries of the secondary structure elements differ somewhat, this domain maintains essentially the same structure as that observed in the individual Abl SH3 domain crystal structure [21]. NMR structures of this domain indicate the presence of antiparallel β strands in the long ab loop. In our structure characteristic hydrogen bonding between these two neighboring strands is not observed except between the main chain amide of Phe91 and the carboxyl group of Ile101 and between the amide group of Ser100 and the carboxyl group of Phe91. The SH3 domain ends in a final β strand, βe, that begins as a single $3_{10}$ helical turn and continues into the SH2 domain through the linker region.

The structure of the SH2 domain in this tandem molecule is also not drastically different from that seen in previous structures [23]. This domain is composed of a central antiparallel β sheet that is flanked on one side by an α helix (αA), and on the other by a second α helix (αB) and a small β hairpin. In other SH2 structures the last β strand, βG, folds against the back of the central β sheet; this strand is absent in the Abl SH3–SH2 structure. Instead of forming hydrogen bonds with the βB strand, in this SH2 structure the C-terminal segment is turned away from the central β sheet. The slightly different orientation at the C terminus of the SH2 domain appears to be a result of a direct contact with a symmetry-related molecule in the unit cell. The Pro-Ala-Pro residues at the C terminus of the SH2 domain are packed against the ab loop of the SH3 domain of a symmetry-related molecule. Interestingly, stacking of these proline residues resembles the binding interactions of proline-rich peptides at the same site in SH3–ligand complexes [24].

### Sites for protein interaction

Mutational analyses with the Abl SH2 domain have shown that mutations in an arginine residue, in β strand βB, completely abolish the transforming activity of the viral gene [18]. The arginine residue is part of a conserved sequence, Phe-Leu-Val-Arg-Glu-Ser (FLVRES), which is located on the surface of the central SH2 β sheet. Side chains of the conserved residues assume positions critical for the formation of an extended binding surface, from the BC loop across the edge of the sheet to the pTyr+3 peptide binding hydrophobic pocket. In previously reported peptide ligand–SH2 structures [25,26] it was shown that ArgβB5 forms critical interactions with the phosphotyrosine moiety. This residue was also identified in the NMR structure of the Abl SH2 domain as the likely phosphotyrosine anchor [22]. In our tandem Abl SH3–SH2 structure, three arginine residues, ArgαA2, ArgβB5 and ArgβC2, extend out from the surfaces of αA, βB and βC, respectively, and point into the proposed phosphotyrosine-binding pocket. Furthermore, the structure determined in our laboratory of the Abl SH2 domain in complex with a high affinity peptide (H-JN *et al.*, unpublished data) has confirmed that these residues do indeed bind to the phosphotyrosine. The crystal structure of the Abl SH3 domain in complex with a proline-rich peptide [21], identified the ligand-binding site on this domain to be a hydrophobic patch which consists of aromatic amino acids Tyr89, Trp118, Trp128, Asn133 and Tyr134 (see Fig. 1). These residues are also shown to be clustered at one end of the three-stranded β sheet in our structure.

**Figure 1**



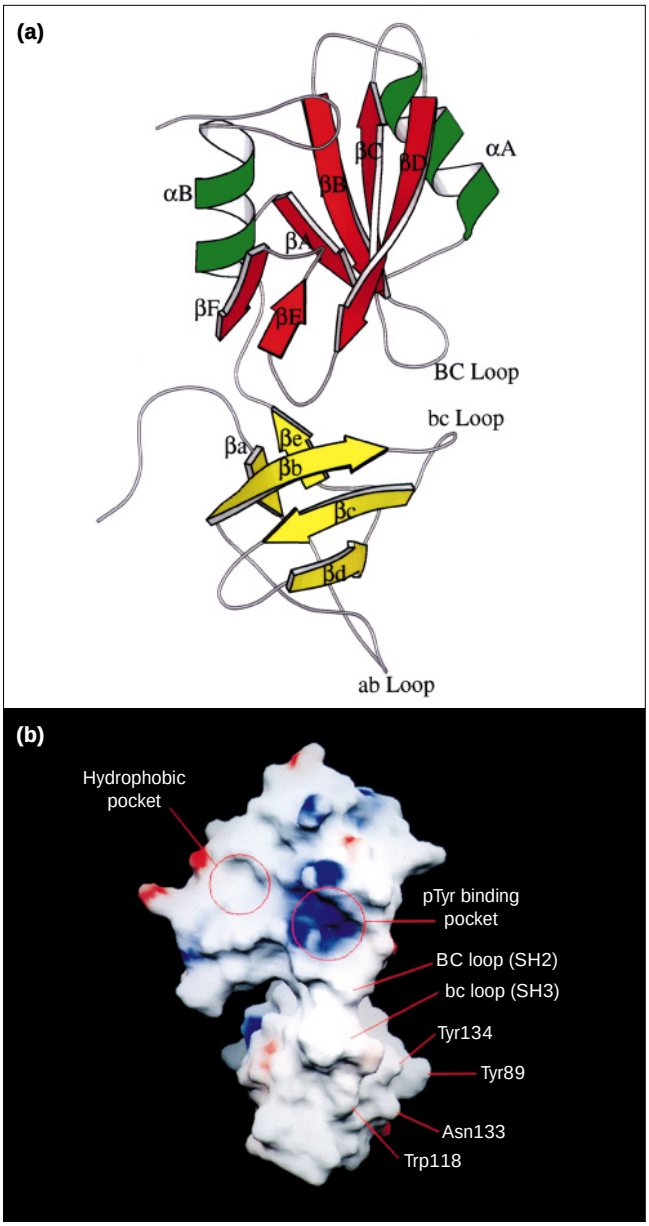

The sequence of the Abl SH3–SH2 domain. The diagram at the top is a schematic representation of the Bcr–Abl protein, showing to scale the locations of the SH2 and SH3 domains, the kinase domain, and the DNA-, G actin- and F actin-binding domains. Below is shown the sequence of the Abl SH3–SH2 domain used in this study. The boundaries of the secondary structure elements of Abl SH3–SH2 are underlined and the notations for these elements are shown below. The amino acid residues involved in the SH3 domain ligand binding (ϕ), those involved in SH2 phosphotyrosine binding (*) and those involved in the intramolecular contact between the two domains (#) are marked. Numbering of the amino acids is that of type IV cellular Abl protein [18].

In Figure 2b, the overall surface structure of the Abl SH3–SH2 protein is shown with the individual proposed peptide-binding sites marked. In the Abl SH3–SH2 structure, both of the ligand-binding sites are fully accessible, and the spacing between these sites (~30 Å) is adequate either for binding two different proteins or for binding a single protein with both the SH2- and SH3-binding motifs. These two binding sites are very different in charge and shape, possibly indicating some difference in the binding modes at these two sites. The SH2-binding surface is composed of two distinct pockets, one pocket is narrow and deep for the phosphotyrosine moiety, the second is a rounded hydrophobic pocket for the pTyr+3 residue. The surface structure shows that the phosphotyrosine-binding pocket is highly positively charged which will accommodate the negatively charged phosphate group of the phosphotyrosine. In contrast, the peptide-binding site on the SH3 domain is an extended surface with ridges and grooves, formed by the aromatic side chains, that will complement the surface character of a polyproline containing substrate [27].

### Intramolecular contact between the SH2 and SH3 domains

Although these two domains maintain distinct compact tertiary structures, the most striking feature of our tandem structure is that significant intramolecular contacts are present between the two domains. The orientation is such that one end of the SH3 β sheet contacts one edge of the central β sheet of the SH2 domain. The contact surface includes the linker segment and two loops, one from each domain, with approximately parallel backbones. These contacts, through the bc loop of the SH3 domain and BC loop of the SH2 domain, are stabilized by a series of Van der Waals interactions. As shown in Figure 3a, at the beginning of the SH2 BC loop, SerBC2 interacts with Asnbc4 in the middle of the SH3 bc loop. Continuing around these two segments, there are contacts between SerBC3 of the SH2 domain and both Glybc5 and Asnbc2 of the SH3 domain. Tyrbc1 of the SH3 domain also interacts with two residues of the BC loop, ProBC4 and GlyBC5. The BC loop of the SH2 domain, which contacts the SH3 domain, was shown to be critical for phosphotyrosine peptide binding in the

**Figure 2**




Schematic diagram and surface structure of the Abl SH3–SH2 regulatory region. **(a)** Richardson diagram of the Abl SH3–SH2 protein showing the relative positions of the SH3 and SH2 domains in one molecule. The SH2 domain is at the top of the figure and the SH3 domain at the bottom. In the SH2 domain the β strands are shown in red and the α helices in green; in the SH3 domain the β strands are in yellow. The secondary structure elements are numbered according to the convention of Eck, *et al.* [31] with individual residues identified by their position within each element or connecting loop; this numbering is used throughout. (The figure was made with the program MOLSCRIPT [51].) **(b)** The molecular surface of the Abl SH3–SH2 structure. The surface is colored according to the local electrostatic potential, ranging from blue (the most positive region) to red (the most negative). The putative phosphotyrosine-binding pocket and a hydrophobic (pTyr + 3 pocket are indicated. Important residues for the ligand binding of the SH3 domain are also indicated. (The figure was made using the program GRASP [52].)

structures of homologous domains from Src family proteins in complex with high affinity peptides [25,26].

Figure 3b shows a stereo composite omit map covering the regions of the two contact loops. The backbone connectivity is very good and the direction and position of the loops are clear in each case. In contrast to the somewhat altered conformation of the SH2 BC loop, the conformation of the bc loop of the SH3 domain is very close to that observed in the Abl SH3 crystal structure [28]. This loop shows variable length and sequences in other SH3 domains. In the SH2 domain the BC loop is flexible, as demonstrated by the different relative orientations seen in previously determined structures [25,26,29]. An extended conformation of the BC loop was observed in the unliganded Src SH2 structure and in the N- and C-terminal SH2 domains of tyrosine kinase ZAP-70 [25,30]. The conformation seen in our structure is the most extreme observed so far, in that the loop is completely extended toward the SH3 domain and one side of the resulting pocket is open.

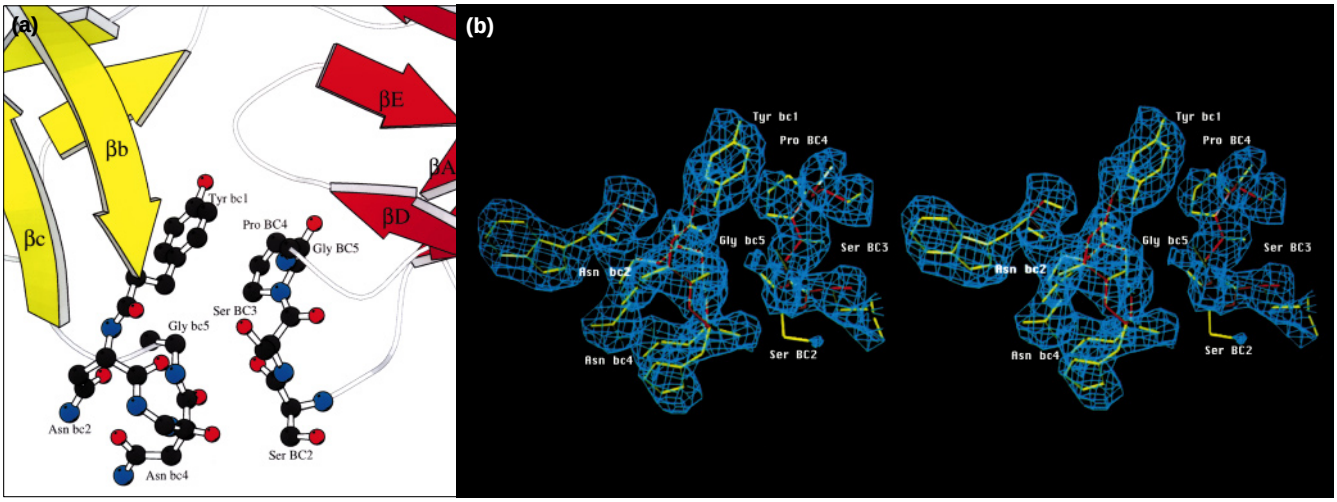
The average temperature factors for the residues in the loops can be used to give an indication of the inherent disorder in these regions of the structure. For the bc loop of the SH3 domain these values are low (~10 Å$^2$) and represent one of the very well ordered regions of the structure. This is in keeping with the high degree of structural similarity to that of the NMR solution structure and indicative of less flexibility in this segment. On the other hand, the SH2 BC loop displays higher average values (~30 Å$^2$) consistent with the fact that this loop has adopted a range of conformations in other SH2 structures, including the Abl SH2–peptide structure (H-JN *et al.*, unpublished results). This flexibility would be necessary in order to accommodate a range of functional interactions of this domain in the intact protein.

The surface involved in the intramolecular contact, which includes the linker segment and two loops, spans roughly 700 Å$^2$. The extent of the contact surface is consistent with what might be expected for complex intermolecular interactions and is considerably greater than has been previously observed for SH3–SH2 intramolecular contacts [31,32]. Figure 2 further emphasizes the close intramolecular contacts between the SH3 and SH2 domains as the molecular surface is clearly continuous between the SH3 and SH2 domains. In addition, this SH3–SH2 intramolecular domain interaction is totally different from the intermolecular packing contacts described below. However, the relative orientation observed in our crystal structure and the significant intradomain contacts are in sharp contrast to a recently reported NMR based solution model for these two domains [33].

### Molecular packing

Unlike the previously reported tandem SH3–SH2 structures of Lck and Grb2, in the Abl SH3–SH2 crystal only

**Figure 3**




Schematic diagram and electron density of the Abl SH2 and SH3 contact region. (a) Schematic diagram showing the side chains in the Abl SH2 and SH3 contact region. The color scheme is the same as that used for Figure 2a; atoms are shown in standard colors. Residues which are not involved in the intramolecular contact are omitted for simplicity. (The figure was made with the program MOLSCRIPT [51].) (b) Stereo view of a composite SA omit map (contour level 0.7σ) of the Abl SH2 and SH3 contact region overlaid on the refined atomic model. The electron-density maps around the BC and bc loops of the SH2 and SH3 domains, respectively, were calculated with coefficients $2F_o–F_c$ and phases calculated from the current model less the atoms of residues 172–179 (βB7–βC1) and residues 110–117 (βb6–βc1), respectively. The simulated-annealing omit map option of X-PLOR [46] was used for the calculation. The BC loop region of the SH2 domain is shown on the right, and the bc loop of the SH3 domain on the left. The polypeptide backbone is shown in red, and the side chains, carbonyl groups, and amide groups are shown in yellow.

one molecule is present in an asymmetric unit and the crystal packing does not introduce any true dimer formation. In the complete unit cell, two pairs of molecules are present which are related by a translation along the xy direction. The molecules within each pair are related by twofold crystallographic symmetry along the z axis. This symmetrical arrangement results in intermolecular packing contacts at three regions that are unrelated to the intramolecular surface described above.

One such contact point is between the N terminus and βc–βd loop of the SH3 domain and the αA helix of the SH2 domain in an adjacent molecule. Here only two pairs of amino acid residues are in contact through Van der Waals interactions. The two molecules related by twofold symmetry are packed against each other around the rotation axis resulting in contact mainly through the linker region of one molecule and β strands βa and βb of its partner's SH3 domain. Finally, the C-terminal segment of an SH2 domain is packed against the βa–βb loop of an adjacent SH3 domain. As mentioned previously, the stacking of proline residues at the end of the SH2 sequence resembles the binding interactions of proline-rich peptides at this site in SH3–ligand complexes and is the only intermolecular contact for which it is possible to link a functional significance. As is typical of protein crystals, there are large solvent channels within the unit cell and all of these intermolecular contact points involve very limited regions of the protein surface. While it is probably true that the linker segment in the tandem molecule is flexible and able to adopt a range of conformations, it does not appear that the intramolecular contact observed here is constrained by the packing arrangement.

### Functional implications of the SH3–SH2 interaction

The significance of the intramolecular contact is emphasized by the fact that it is the BC loop of the SH2 domain that has been shown to be directly involved in binding phosphotyrosine residues. The exact region of the BC loop involved in the contact of the SH3 domain is part of a conserved sequence, Ser-Ser-Pro-Gly, which was shown to be critical for phosphotyrosine peptide binding and is conserved in Src family SH2 domains. In particular, the amino acid residues at this position were shown to be important for binding of the phosphotyrosine moiety itself. The BC loop folds over the phosphate group of the phosphotyrosine and forms the top portion of the binding pocket [25,26]. The contact loop from the SH3 domain was shown in the SH3–ligand complex structure to be located near the ligand-binding site [21], however, this loop is not yet directly implicated in any functional role.

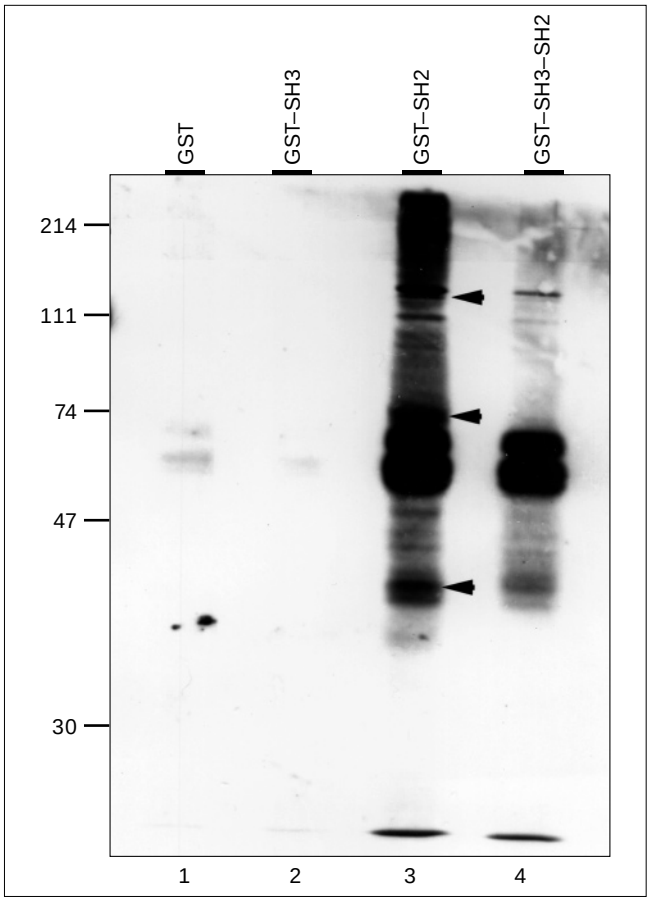
The extended conformation of the SH2 BC loop observed in our structure can be analyzed with respect to its functional significance. This is possible because the overall tertiary structure of the SH2 peptide-binding site is directly

influenced by this interaction. The close contact between the SH3 and SH2 domains implies not only that these domains interact to stabilize the entire molecule, but also that the function of one could be modified by the presence of the other. This also agrees with previous mutational analyses regarding the effects of changes in relative domain positions and linker size on Abl kinase activity. While the SH2 domain is required for efficient cellular transformation, the presence of the SH3 domain has been shown to have a strong negative effect and is the major cause of the normally non-transforming phenotype of the cellular Abl protein. However, mutants with only slight alterations in the relative position of the SH3 domain within the Abl sequence have full transformation efficiency [15]. These properties are consistent with there being direct physical contact between these two domains as observed in our structure. However, the presence of the SH3 domain has no effect on the kinase activity of Abl protein *in vitro*, suggesting that overall cellular regulation is more complex [15].

In order to determine whether the presence of the Abl SH3 domain affects the phosphoprotein binding activity of the corresponding SH2 domain, substrate selectivity of the Abl SH2 domain in the presence and absence of the corresponding SH3 domain was examined. In Figure 4, the binding of proteins containing either or both of these domains to phosphotyrosine containing cellular proteins is compared. Similar methods have been used to examine potential phosphotyrosine binding proteins in other systems [34]. As expected certain proteins do indeed bind to the SH2 domain in both contexts. However, there are several significant differences in the presence of the SH3 domain. The important obvious result is that the overall level of protein binding is considerably less in the presence of the SH3 domain. In addition, proteins with approximate molecular weights of 160 kDa, 60–70 kDa and 40 kDa, which bind to the Abl SH2 domain in the absence of the SH3 domain, do not bind to the Abl SH3–SH2 fragment. These corresponding bands are not detectable, even at much longer exposures of the blot, indicating that some specific binding interactions are disrupted. This initial biochemical evidence taken together with the Abl SH3–SH2 structure strongly suggests that the binding activity of the Abl SH2 domain is modulated by the SH3 domain through the intramolecular contact of the BC and bc loops of the SH2 and SH3 domains, respectively. The absence of such SH3 modulation of the SH2 domain may partly explain the transforming activities observed in SH3 deleted mutants.

It has been proposed that the negative regulatory effect of the SH3 domain in the Abl system is the result of the interaction of a *trans* inhibitor protein that exerts its effects by interacting with the Abl SH3 domain as well as another region of the protein [15]. Our results do not exclude this possibility in that general steric hindrance or conformational adjustments, resulting from one domain being so close to the other, might affect the potential interaction of yet another protein or domain. However, the simple model we are proposing does account for the differential activity. In addition, a protein has been identified in humans [35], and an analogous protein in mice [36], whose activity suggests a *cis* mode of Abl kinase inhibition. The protein isolated from lymphocytes, Abi-2, binds to both the SH3 domain as well as an additional C-terminal sequence of the cellular Abl protein; mutants of both regions will activate Abl transforming potential [35]. The authors suggest that the SH3 domain therefore functions in *cis* by binding to this or another region of the Abl protein thus locking it in an inactive form. They also noted that Abi-2 contains tyrosine phosphorylation sites that might function as a binding site for the Abl SH2 domain as well. The orientation of the two

**Figure 4**




Binding profiles of Abl SH3, Abl SH2 and Abl SH3–SH2 domains to phosphotyrosine-containing proteins. Western blot analysis of phosphotyrosine-containing proteins precipitated by various Abl SH3/SH2 constructs (see Materials and methods). The domain or domains used for precipitation are indicated above each lane of the gel; the gel was loaded keeping the amount of total protein constant in each lane. The arrows indicate proteins which show clear differences in binding between the GST–SH2 and GST–SH3–SH2 constructs. Molecular masses are given on the left in kDa.

domains seen in our structure provides a *cis* SH3–SH2 interaction that represents a less active form of the protein, with respect to phosphoprotein binding.

In contrast to our findings, a recent NMR model of the Abl SH3–SH2 region gave a completely different relative orientation of the two domains. These authors state that the only significant intramolecular contacts are near the linker region. This NMR model was based on comparing assignments from the individual domains and attempting to identify potential differences in the tandem molecule. As not all resonances could be assigned there may still be some room for adjustment of this model. However, regardless of that fact what is clear is that we have captured one real conformational arrangement that could have very important functional implications.

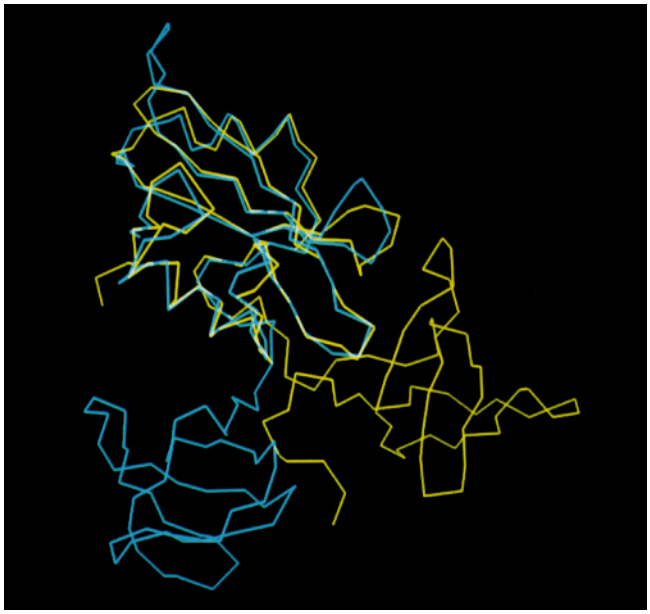
#### Comparison with other SH3–SH2 structures

While the Abl SH3–SH2 structure is consistent with the biochemical and genetic data for the Abl system, the relative orientation of the SH2 and SH3 domains is different from that observed in the two existing examples of SH3–SH2 domain structures, in Grb2 and Lck SH3–SH2 proteins. In the composite SH3–SH2 domain structure of Src family member kinase, Lck, solved by Eck *et al.* [31], there is very little intramolecular contact between the two domains. In their structure the molecule appears as a dimer with significant head-to-tail intermolecular interactions, in contrast to the monomer seen in the Abl SH3–SH2 structure. The residues that are involved in the dimer interaction are conserved in other Src family members but not in the Abl family kinases. In addition, the catalytic activity of the Src family kinases is thought to be regulated by binding of their SH2 domain to a tyrosine phosphorylated sequence, near the C terminus of the catalytic domain [19,37], which is absent in the Abl kinase. The SH3 and peptide interaction in the Lck structure may resemble the *in vivo* intermolecular regulatory mechanism. In contrast, experiments with various Bcr–Abl constructs have suggested that a serine/threonine rich region of the Bcr segment may interact with the Abl SH2 domain to influence kinase activity [38].

Figure 5 emphasizes the striking difference in the relative domain orientation for Abl and Lck. It is clear, from this superposition of the Cα backbones for the two SH2 domains, that the SH3 domain of Abl is rotated about the linker segment and then angled up toward the other side of the SH2 domain as compared to the Lck structure. The closest regions for intramolecular contact are in the vicinity of the αB helix and the βD and βE strands of the Lck SH2 domain, as opposed to the BC loop region in Abl.

Recently the crystal structure of the mediator protein Grb2 was also reported [32]. This protein consists of two SH3 domains flanking a lone SH2 domain. The structure of Grb2 showed there is essentially no interaction between domains within one molecule however, there are extensive intermolecular contacts between SH3 domains. This lack of interaction is consistent with the independent binding function of the SH2 and SH3 domains in such modular proteins [39,40]. It may result from the need for this protein to bind to more than one cellular partner thus bringing these proteins together to interact. The striking differences in the domain contacts between Abl and these other proteins reflect the different modes of interaction and cellular regulation.

**Figure 5**



Superposition of SH3–SH2 proteins from Abl and Lck. A least squares fit was performed using the backbone coordinates of the SH2 domain from the current structure and those of the SH2 domain from the tandem structure of Eck *et al.* [31]. The resulting figure represents the entire Cα traces from both molecules. The Abl protein is shown in yellow and the Lck protein in cyan. The figure is oriented with the SH2 domains at the top and the SH3 domains pointing downward; an approximately –45° in plane rotation with respect to the view shown in Figure 2a. The close intramolecular loop contact can be seen in the upper right of the Abl contact region.

## Biological implications

Protein tyrosine kinases are central components in the regulation of cellular growth, differentiation, and transformation in complex eukaryotes. Although not closely related to the Src family of kinases, the N-terminal domain of Abl kinase contains two Src homology domains, SH2 and SH3. In the Abl kinase system transforming activity can be acquired by several known mechanisms. These include the conversion from the cellular Abl protein, c-Abl, to a viral form, v-Abl, resulting from the deletion of the SH3 domain or by activating point mutations in either the kinase or SH3 domains.

The SH3 domain supresses transformation whereas the SH2 domain is known to be required for efficient transformation. The roles of these two domains are thought to be somewhat different from those of the homologous domains from other kinases. Bcr–Abl proteins, generated by the upstream fusion of segments of the breakpoint cluster region to the c-*abl* coding sequence, lack the functions associated with the oncogenic activation of v-Abl. Bcr–Abl proteins contain an intact SH3 domain and have no additional mutations in the corresponding Abl sequences. In addition, the Bcr segments of these fusion proteins are required for *in vivo* autophosphorylation so it has been proposed that they might in some way alter the protein function in the same manner as loss of an SH3 domain does in other viral Abl proteins. It has been suggested that the suppression of transforming activity imposed by the SH3 domain may be caused by binding a cellular inhibitory factor that also binds to the kinase domain.

Our results provide an alternative explanation for the cellular and biochemical data observed for the Abl system. The intramolecular contact seen in the Abl SH3–SH2 structure gives the first clear indication of direct interaction between these two domains. If the SH2 and SH3 domains communicate with each other in this fashion, it offers a straightforward reason as to why the presence of one directly influences the activity of the other. The intramolecular interaction could either lock the protein in an inactive form for phosphotyrosine binding or in an inaccessible conformation for binding of other substrates. In an analogous manner an NMR study has detected in the murine Crk-family adaptor protein, Crk II, an intramolecular SH2–phosphotyrosine interaction that in turn effects the ability of its SH3 domain to bind cellular targets [41]. The novel regulatory mechanism of the SH2 and SH3 domains of the Abl tyrosine kinase, as compared to other systems, is reflected in the structural differences of the Abl and Lck SH3–SH2 tandem proteins. The examination of intramolecular versus intermolecular interactions may provide new information for understanding the biological roles of these oncogenic proteins.

## Materials and methods

### *Cloning of the Abl SH3–SH2 domains and protein purification*

Gene segments encoding SH3–SH2 domains of *bcr–abl* were amplified using the polymerase chain reaction. The amplified segments were inserted into the pRSET (InvitroGen) expression vector using *Nde*I and *Hin*dIII sites. Several constructs were made with varying additional residues upstream of the canonical SH3 start site and downstream of the corresponding SH2 stop site, to increase the chance of obtaining crystals. Cultures of *Escherichia coli* strain BL21(DE3) (10 L) transformed with this expression vector were grown at 37° C. After induction with 200 mg $l^{-1}$ isopropyl-β-thiogalactoside (IPTG) for 4 h, cells were harvested by centrifugation at 4000 g for 20 min and resuspended in a lysis buffer containing 100 ml of 50 mM Tris pH 7.5, 0.2 M NaCl, 2 mM DTT and 1 mM phenylmethyl sulfonyl fluoride (PMSF). The cells were lysed by sonication, and lysates were centrifuged for 40 min at 15 000 g. The supernatant was loaded onto a 15 ml phosphotyrosine-affinity column pre-equilibrated in a buffer containing 50 mM Tris pH 7.5, 0.2 M NaCl and 2 mM DTT. The proteins were eluted with a solution containing 25 mM phosphotyrosine. Gel filtration using a Superdex 75 FPLC column in a buffer containing 20 mM Tris pH 8.0, 0.2 M NaCl and 2 mM DTT removed high molecular weight contaminants. The purified proteins were concentrated using an ultrafiltration technique.

### *Crystallization and data collection*

Crystals were grown by vapor diffusion in hanging drops at room temperature. A 60 mg $ml^{-1}$ solution of the protein was mixed with an equal volume of well solution containing 2 M NaCl, 5 mM $MnCl_2$, 100 mM HEPES pH 7.5. Thin rod-shaped crystals of monoclinic space group C2 ($a = 65.62$ Å, $b = 45.85$ Å, $c = 54.45$ Å, $\beta = 95.18°$) were observed in approximately three weeks. A complete data set was collected to 2.5 Å resolution from a single crystal at 20° C. A Mar Research image plate scanner mounted on a Rigaku RU300 rotating-anode X-ray generator was used for data collection. Data were processed and reduced using the programs DENZO and SCALEPAK [42]. A total of 19 460 observations were processed to a unique data set of 5.585 intensities. Data are 99.3 % complete to 2.5 Å with an $R_{sym}$ equal to 0.048 % (Table 1).

### *Structure determination and refinement*

Structure-factor phases were determined by the molecular replacement technique. A series of SH2 and SH3 crystal or solution structures were tried as search models. Ultimately, it was the Src SH2 domain [43] and the Abl SH3 domain [28] crystal structures that gave identifiable solutions. The cross-rotation function option of the program AMoRe [44] was used to determine the rotation angles for each domain. The correct rotation solutions obtained for the SH2 and the SH3 domain (8 Å to 3.5 Å data) had correlation factors of 16.3 and 23.6, compared with background values of 10.3 and 11.9, respectively. These solutions were further analyzed by Patterson correlation refinement [45] and the appropriately transformed coordinates were input to X-PLOR [46]; subsequent rotation searches gave angles of essentially 0° for the top

**Table 1**

**X-ray data collection and refinement summary.**

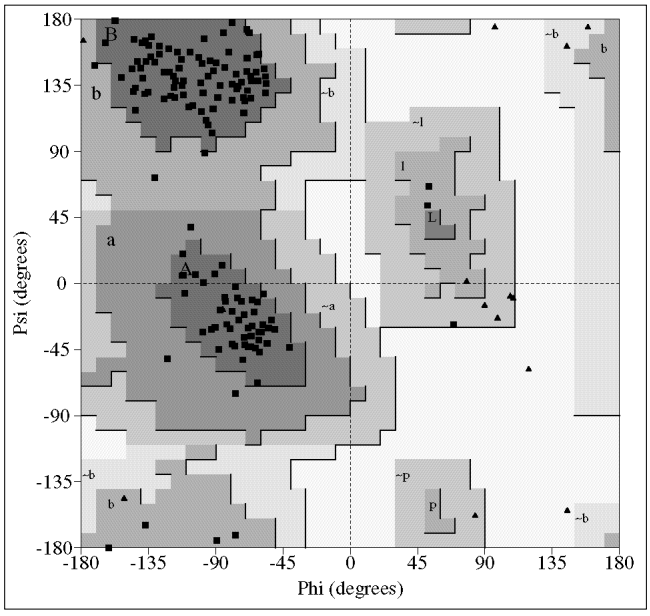
| | |
|---|---|
| Resolution range (Å) | 12–2.5 |
| Space group | C2 |
| Number of total reflections | 19 460 |
| Number of unique reflections | 5 585 |
| Completeness of data (%) | |
| 2.81–2.69 (Å) | 99.5 |
| 2.69–2.59 (Å) | 98.6 |
| 2.59–2.50 (Å) | 98.0 |
| all data | 99.3 |
| $R_{sym}$ (%)* | 4.8 |
| $R_{cryst}$ ($F > 2\sigma(F)$) | 18.3% (10.0–2.5Å) |
| $R_{free}$ (%)† | 26 |
| Number of non-hydrogen atoms | 1281 |
| Number of solvent molecules | 40 |
| Rms deviation from ideality | |
| bond lengths (Å) | 0.018 |
| bond angles (°) | 3.5 |
| dihedral angles (°) | 26.5 |
| Average B factors (Å$^2$) | |
| main chain | 12.3 |
| all protein atoms | 13.1 |

*$R_{sym} = \Sigma(|<I> - I|)/\Sigma<I>$, where I is the observed intensity and $<I>$ is the averaged intensity obtained from multiple measurements of symmetry-related relections. †$R_{free} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$ for all measured relections within the test set, with $F_{calc}$ scaled to $F_{obs}$. The calculated structure factors $F_{calc}$ are obtained from a model which was constructed and refined against the working set. ‡Rcryst for values in the resolution shell 10.0–2.5 Å.

solutions. Translation functions were then calculated using AMoRe and the top solutions subjected to rigid-body refinement. These yielded correlation and R factor values of 35.2 (background of 26.4) and 50.0 (background of 51.7) for the SH3 domain and 38.2 (background of 24.7) and 50.2 (background of 54.2) for the SH2 domain.

After rotation and translation solutions were obtained for the individual SH2 and SH3 domains, the y translation between the two domains was determined. This was achieved by searching for the position of the SH3 domain relative to that of the fixed SH2 domain that gave the lowest R factor after rigid-body refinement (X-PLOR). The y axis was sampled first in 4 Å increments and once the general position was identified a fine sampling was performed in that region. The solution obtained in this manner had no significant packing violations. After rigid-body refinement of the model followed by refinement of all atomic positions by conjugate-gradient optimization, initial electron-density maps were calculated. Based on these maps, the model determined by molecular replacement was rebuilt to reflect the amino acid sequence of the Abl SH3–SH2 domain. The rebuilt model was refined through 38 cycles of manual rebuilding using the program O [47] and by using the simulated-annealing and positional refinement options of X-PLOR. Simulated-annealing omit maps were used to refit several regions of the model including the segment linking the SH3 and SH2 domains and the N and C termini. At the final stage of the refinement, a simulated omit map was used for the whole structure to minimize the model bias. Individual B factors were refined and water molecules were added at the final stage of refinement. The final model includes 163 of 165 amino acids, with two residues missing at the C terminus. The R factor of the final model is 18.3% at a resolution of 2.5 Å with 40 solvent molecules. This refined model has a root mean square (rms) deviation of the bond lengths of 0.018 Å and backbone angles of 3.5°. Analysis of backbone torsion angles identified 89% of the residues within the most favored regions and the remaining 11% in additional allowed regions with only one residue whose phi/psi angles are outside the accepted range [48] (Fig. 6). This single residue in the generously allowed region is an asparagine residue in the SH3 long βa–βb loop.

**Figure 6**




Ramachandram plot of the refined structure. Glycine and non-glycine ψ, ϕ pairs are designated by triangles and squares, respectively. Disallowed, generously allowed, favorable and most favorable regions are indicated by progressively darker shading. Of all the residues, 89% are in the most favorable regions and none of the residues are in disallowed regions.

### *Precipitation and western blot analysis of phosphotyrosine-binding proteins*

The human leukemia cell line M07, which was transformed with p210 *bcr–abl* [34] was lysed in NP40 lysis buffer. Total protein from the lysate (200 mg) were incubated with glutathione-s-transferase (GST) protein (lane 1), or GST fusion proteins Abl SH3 (lane 2), Abl SH2 (lane 3) and Abl SH3–SH2 (lane 4). These proteins were covalently cross-linked to glutathione beads (Pharmacia) using dimethyl pimelimidate [49]. Precipitates were washed with ten volumes of a buffer containing 20 mM Tris-HCl pH 8.0 and 0.5 M $Li_2SO_4$ and ten volumes of phosphate buffered saline (PBS).The samples were subjected to 10% SDS PAGE followed by Western blotting. The filter was probed for phosphotyrosine containing protein using the antiphosphotyrosine monoclonal anti+body 4G10 [50].

### *Accession numbers*

The coordinates for the current structure are being deposited in the Brookhaven Prote+in Databank.

## Acknowledgements

We thank A McCarthy for help with molecular replacement and D Kim for technical assistance. This work was supported by grants from Sandoz Pharmaceutical Company (91032) (CAF) and from the National Institutes of Health (CA43803) (TMR).

## References

1. Lemmon, M.A. & Schlessinger, J. (1994). Regulation of signal transduction and signal diversity by receptor oligomerization. *Trends Biochem. Sci.19, 459-493.*
2. Wang, J.Y.J. (1993). Abl tyrosine kinase in signal transduction and cell-cycle regulation. *Curr. Opin. Genet. Devel.* **3**, 35–43.
3. Goff, S.P., Gilba, E., Witte, O.N. & Baltimore, D. (1980). Structure of the Abelson murine leukemia virus genome and the homologous cellular gene: studies with cloned viral DNA. *Cell* **22**, 777–785.
4. Shtivelman, E., Lifshitz, B., Gale, R.P. & Canaani, E. (1985). Fused transcript of the *abl* and *bcr* genes in chronic myelogenous leukemia. *Nature* **315**, 550–554.
5. Mes-Masson, A.-M., McLaughlin, J., Daley, G.Q., Paskind, M. & Witte, O.N. (1986). Overlapping cDNA clones define the complete coding region for the p210c-*abl* gene product associated with chronic myelogenous leukemia cells containing the Philadelphia chromosome. *Proc. Natl. Acad. Sci. USA* **83**, 9768–9772.
6. Daley, G.Q., Van Etten, R.A. & Baltimore, D. (1990). Induction of chronic myelogenous leukemia in mice by the p210*bcr*/*abl* gene of the Philadelphia chromosome. *Science* **247**, 824–830.
7. Jackson, P. & Baltimore, D. (1989). N-terminal mutations activate the leukemogenic potential of the myristoylated form of c-abl. *EMBO J.* **8**, 449–456.
8. Franz, W.M., Berger, P. & Wang, J.Y.J. (1989). Deletion of an N-terminal regulatory domain of the c-Abl tyrosine kinase activates its oncogenic potential. *EMBO J.* **8**, 137–147.
9. Pawson, T. (1992). SH2 and SH3 domains. *Curr. Opin. Struct. Biol.* **2**, 432–437.
10. Schlessinger, J. (1994). SH2/SH3 signaling proteins. *Curr. Opin. Genet. Dev.* **4**, 25–30.
11. Birge, R.B. & Hanafusa, H. (1993). Closing in on SH2 specificity. *Science* **262**, 1522–1524.
12. Viguera, A.R., Arrondo, J.L.R., Musacchio, A., Saraste, M. & Serrano, L. (1994). Characterization of the interaction of natural proline-rich peptides with five different SH3 domains. *Biochemistry* **33**, 10925–10933.
13. Cicchetti, P., Mayer, B.J., Thiel, G. & Baltimore, D. (1992). Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. *Science* **257**, 803–806.
14. Ren, R., Mayer, B.J., Ciccheti, P. & Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. *Science* **259**, 1157–1161.
15. Mayer, B. & Baltimore, D. (1994). Mutagenic analysis of the roles of SH2 and SH3 domains in regulation of the Abl tyrosine kinase. *Mol. Cell Biol.* **14**, 2883–2894.

16. Mayer, B.J., Jackson, P.K. & Baltimore, D. (1991). The noncatalytic Src homology region 2 segment of Abl tyrosine kinase binds to tyrosine-phosphorylated cellular proteins with high affinity. *Proc. Natl. Acad. Sci. USA* 88, 627–631.
17. Songyang, Z., *et al*., & Cantley, L.C. (1993). SH2 domains recognize specific phosphopeptide sequences. *Cell* 72, 767–778.
18. Mayer, B., Jackson, P., Van Etten, R. & Baltimore, D. (1992). Point mutations in the Abl SH2 domain coordinately impair phosphotyrosine binding *in vitro* and transforming activity *in vivo*. *Mol. Cell Biol.* 12, 609–618.
19. Superti-Furga, G., Fumagalli, S., Koegl, M., Coutneidge, S.A. & Draetta, G. (1993). Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src. *EMBO J.* 12, 2625–2634.
20. Panchamoorthy, G., *et al*., & Band, H. (1994). Physical and functional interactions between SH2 and SH3 domains of the Src family protein tyrosine kinase p59*fyn*. *Mol. Cell Biol.* 14, 6372–6385.
21. Musacchio, A., Saraste, M. & Wilmanns, M. (1994). High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline rich peptides. *Nat. Struct. Biol.* 1, 546–551.
22. Overduin, M., Rios, C., Mayer, B., Baltimore, D. & Cowburn, D. (1992). Three-dimensional solution structure of the Src homology 2 domain of c-Abl. *Cell* 70, 697–704.
23. Kuriyan, J. & Cowburn, D. (1993). Structures of SH2 and SH3 domains. *Curr. Opin. Struct. Biol.* 3, 828–837.
24. Musacchio, A., Wilmanns, M. & Saraste, M. (1994). Structure and function of the SH3 domain. *Prog. Biophys. Mol. Biol.* 61, 283–297.
25. Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D. & Kuriyan, J. (1993). Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. *Cell* 72, 779–790.
26. Eck, M.J., Shoelson, S.E. & Harrison, S.C. (1993). Recognition of a high-affinity phosphotyrosyl peptide by the Src-homology-2 domain of p56 lck. *Nature* 362, 87–91.
27. Yu, H., Chen, J.K., Feng, S., Dalgarno, D.C., Brauer, A.W. & Schreiber, S.L. (1994). Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* 76, 933–945.
28. Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. & Saraste, M. (1992). Crystal structure of a Src homology 3 (SH3) domain. *Nature* 359, 851–855.
29. Lee, C.H., *et al*., & Kuriyan, J. (1994). Crystal structures of peptide complexes of the amino-terminal SH2 domain of the Syp tyrosine phosphatase. *Structure* 2, 423–38.
30. Hatada, M.H., *et al*., & Karas, J.L. (1995). Molecular basis for interaction of the protein tyrosine kinase ZAP-70 with the T-cell receptor. *Nature* 377, 32–38.
31. Eck, M., Atwell, S.K., Shoelson, S.E. & Harrison, S.C. (1994). Structure of the regulatory domains of the Src-family tyrosine kinase Lck. *Nature* 368, 764–769.
32. Maignan, S., Guilloteau, J.-P., Fromage, N., Arnoux, B., Becquart, J. & Ducruix, A. (1995). Crystal structure of the mammalian Grb2 adaptor. *Science* 268, 291–293.
33. Gosser, Y.Q., Zheng, J., Overduin, M., Mayer, B. & Cowburn, D. (1995). The solution structure of Abl SH3 and its relationship to SH2 in the SH(32) construct. *Structure* 3, 1075–1086.
34. Ernst, T.J., Slattery, K.E. & Griffin, J.D. (1994). p210Bcr/Abl and p160v-abl induce an increase in the tyrosine phosphorylation of p93c-fes. *J. Biol. Chem.* 269, 5764–5769.
35. Dai, Z. & Pendergast, A.M. (1995). Abi-2, a novel SH3-containing protein interacts with the Abl tyrosine kinase and modulates c-Abl transforming activity. *Genes Dev.* 9, 2569–2582.
36. Shi, Y., Alin, K. & Goff, S.P. (1995). Abl-interactor-1, a novel SH3 protein binding to the carboxy-terminal portion of the Abl protein, suppresses v-abl transforming activity. *Genes Dev.* 9, 2583–2597.
37. Margolis, B., *et al*., & Schlessinger, J. (1990). The tyrosine phosphorylated carboxy terminus of the EGF receptor is a binding site for GAP and PLC-γ. *EMBO J.* 9, 4375–4380.
38. Pendergast, A.M., Muller, A., Havlik, M., Maru, Y. & Witte, O.N. (1991). Bcr sequences essential for transformation by the Bcr-Abl oncogene bind to the Abl SH2 regulatory domain in a non-phosphotyrosine-dependent manner. *Cell* 66, 161–171.
39. Lemmon, M.A., Ladbury, J.E., Mandiyan, V., Zhou, M. & Schlessinger, J. (1994). Independent binding of peptide ligands to the SH2 and SH3 domains of Grb2. *J. Biol. Chem.* 269, 31653–31658.
40. Cussac, D., Frech, M. & Chardin, P. (1994). Binding of the Grb2 SH2 domain to phosphotyrosine motifs does not change the affinity of its SH3 domains for Sos proline-rich motifs. *EMBO J.* 13, 4011–4021.
41. Rosen, M.K., Yamazaki, T., Gish, G.D., Kay, C.M., Pawson, T. & Kay, L.E. (1995). Direct demonstration of an intramolecular SH2-phosphotyrosine interaction in the Crk protein. *Nature* 374, 477–479.
42. Otwinowski, Z. (1993). Oscillation data reduction program. In *Data Collection and Processing, Proceedings of the CCP4 Study Weekend*. (Sawyer, L., Isaacs, N. & Bailey, S., eds), pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
43. Waksman, G., *et al*., & Kuriyan, J. (1992). Crystal structure of the phosphotyrosine recognition domain SH2 of v-*src* complexed with tyrosine-phosphorylated peptides. *Nature* 358, 646–653.
44. Navaza, J. (1994). AMoRe: An automated package for molecular replacement. *Acta. Cryst. A* 50, 157–163.
45. Brünger, A.T. (1990). Extension of molecular replacement: a new search strategy based on Patterson correlation refinement. *Acta. Cryst. A* 46, 46–57.
46. Brünger, A.T. (1993) X-PLOR: Version 3.1. A system for Crystallography and NMR. Yale University Press, New Haven, CT, USA.
47. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* 47, 110–119.
48. Ramachandran, G.M. & Sasisekharan, V. (1968). Conformations of polypeptides and proteins. *Adv. Protein Chem.* 23, 283–437.
49. Simanis, V. & Lane, D.P. (1985). An immunoaffinity purification procedure for SV40 large T antigen. *Virology* 144, 88–100.
50. Druker, B.J., Mamon, H.J. & Roberts, T.M. (1989). Oncogenes, proteins, and signal transduction. *N. Engl. J. Med.* 321, 1383–1391.
51. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* 24, 946–950.
52. Nicolls, A., Bharadwaj, R. & Honig, B. (1993). GRASP: graphical representation and analysis of surface properties. *Biophys. J.* 64, 166–170.