Three-dimensional structure of the tyrosine kinase c-Src

Wenqing Xu*, Stephen C. Harrison†‡§ & Michael J. Eck†

*Laboratory of Molecular Medicine, Children's Hospital, and Departments of * Microbiology and Molecular Genetics, † Biological Chemistry and Molecular Pharmacology, and ‡ Pediatrics, Harvard Medical School, and § The Howard Hughes Medical Institute, 300 Longwood Avenue, Boston, Massachusetts 02115, USA*

The structure of a large fragment of the c-Src tyrosine kinase, comprising the regulatory and kinase domains and the carboxy-terminal tail, has been determined at 1.7 Å resolution in a closed, inactive state. Interactions among domains, stabilized by binding of the phosphorylated tail to the SH2 domain, lock the molecule in a conformation that simultaneously disrupts the kinase active site and sequesters the binding surfaces of the SH2 and SH3 domains. The structure shows how appropriate cellular signals, or transforming mutations in v-Src, could break these interactions to produce an open, active kinase.

Src, the product of the first proto-oncogene to be characterized[1], is a protein switch. Its output is tyrosine phosphorylation, catalysed by a kinase domain; its inputs are multiple regulatory interactions, received by other parts of its polypeptide chain[2,3]. The family of Src-related protein tyrosine kinases now includes nine members (Fyn, Yes, Fgr, Lyn, Hck, Lck, Blk and Yrk, in addition to Src itself). One or more of these proteins is present in every higher-animal cell type studied to date. Many, such as Lck, are specifically expressed in haematopoietic cells; Src itself shows a more widespread cellular distribution. Although not linked covalently to extracellular receptor domains, Src family members respond to a number of receptor-mediated signals, both by changes in kinase activity and by changes in intracellular localization[2].

Proteins of the Src family have a common domain organization, with each segment designated as a Src-homology (SH) region. The N-terminal segment includes the SH4 domain, which is a myristylation and membrane-localization signal, as well as a 'unique' domain of 50–70 residues that has no particular similarity to family members. The SH3, SH2 and catalytic (SH1) domains follow in order in the polypeptide chain; there is also a short, C-terminal 'tail', which includes a critical tyrosine residue. SH2 and SH3 domains mediate protein–protein interactions in cellular signalling cascades, and are found in many proteins outside the Src family[4]. Extensive structural and functional studies of SH3 and SH2 domains have defined their specific molecular-recognition properties. SH3 domains are small, β-barrel modules that present a non-polar groove complementary to peptides in a polyproline-II conformation[5,6]. Proline-rich sequences in target molecules mediate interactions with SH3-containing proteins[7]. SH2 domains bind polypeptide segments that contain a phosphotyrosine (pY)[8], with significant context specificity, especially for the first three residues following the phosphotyrosine[9–11].

The SH3 and SH2 domains and the C-terminal tail all have roles in regulating Src kinase activity. The classically characterized v-*src* oncogene from avian retroviruses encodes a constitutively active kinase with a deleted tail[12]. Mutation of the conserved tyrosine (Tyr 527 in avian c-Src) to phenylalanine also creates an active oncogenic protein[13]. It is now clear that phosphorylation of Tyr 527 by a specific kinase, Csk (ref. 14) inhibits Src catalytic activity by creating an intramolecular binding site for the Src SH2 domain[15,16]. The interaction is believed to result in auto-inhibition by locking the molecule in an inactive state. Mutations in either the SH2 or the

Table 1 Data collection, phase determination, and refinement statistics

Crystal	Native	$KAu(CN)_2$	PIP	$KAu(CN)_2$ + PIP	Native 2
Resolution (Å)	25.0–2.0	25.0–2.5	25.0–2.45	25.0–2.0	20.0–1.5
R_{sym} (%)	5.3	4.2	3.9	4.9	4.9
Total observations	102,852	29,663	35,805	52,980	165,384
Unique reflections	25,752	14,690	15,631	24,653	48,728
Coverage (%)	90.0	89.7	90.1	80.4	66.1
R_{deri} (%)		24.9	15.3	28.0	
No. of sites		3	2	4	
Phasing power (25–2.0 Å, acent/cent)	1.79/1.35	0.63/0.55	1.50/1.19		
R_{cullis} (%, 25–2.0 Å, acent/cent)		0.68/0.72	0.95/0.92	0.73/0.80	
Refinement (with native 2 data set)					
R_{free} (%, 20–1.5 Å)	26.4				
R_{cryst} (%, 20–1.5 Å)	21.0				

			R.m.s.d.		
	Res. no.	Average *B*	Bonds (Å)	Angles (°)	*B*-values (Å^2 bonded)
Protein	436	28.3	0.010	1.16	2.16
Water	490	39.4			

$R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I is the observed intensity. $\langle I\rangle$ is the average intensity of multiple observations of symmetry-related reflections. $R_{deri} = \Sigma||F_{PH}| - |F_P||/\Sigma|F_P|$, where $|F_P|$ is the protein structure factor amplitude and $|F_{PH}|$ is the heavy-atom derivative structure factor amplitude. Phasing power = r.m.s. ($|F_H|/E$), where $|F_H|$ is the heavy-atom structure factor amplitude and E is the residual lack of closure error. $R_{cullis} = \Sigma|E|/\Sigma||F_{PH}| - |F_P||$. Figure of merit = $\langle \Sigma P(\alpha) e^{i\alpha}/\Sigma P(\alpha)\rangle$, in which α is the phase and $P(\alpha)$ is the phase probability distribution. $R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where R_{free} is calculated for a randomly chosen 5% of reflections, R_{cryst} is calculated for the remaining 95% of reflections used for structure refinement.

SH3 domain can affect the stability of this 'closed', inactive conformation[2,3,13]. Many, but not all, of these mutations are transforming in cell culture and tumorigenic in animals. In the 'open' state, phosphorylation of Tyr 416 in the kinase domain further enhances catalytic activity[17].

To analyse the mechanism of this molecular switch, we have crystallized a large fragment of human c-Src containing the SH3, SH2 and kinase domains and the C-terminal tail. We chose to prepare and use protein that is singly phosphorylated at Tyr 527, in order to select the closed, autoinhibited conformation. The structure, determined at 1.7 Å resolution, shows that there is indeed an intramolecular association of the SH2 domain with the phosphorylated Tyr 527. It also reveals how the SH3 domain contributes to the stability of the closed state, through an

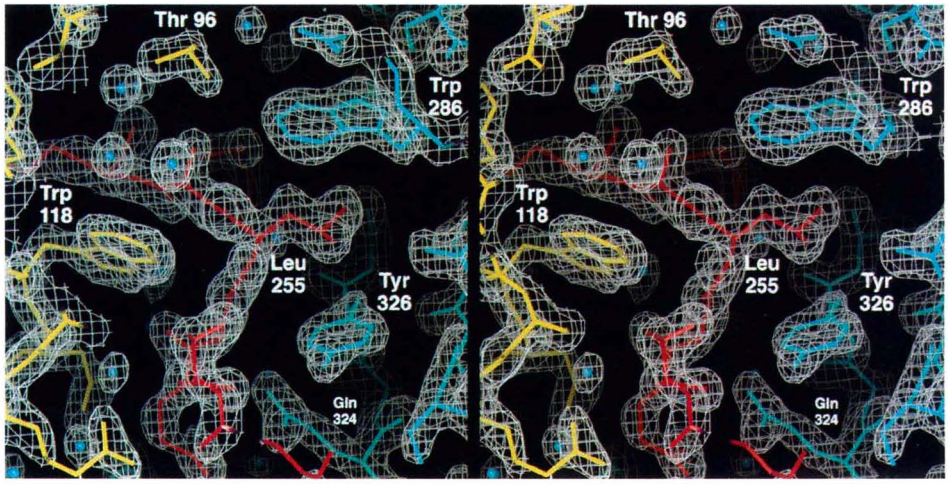


Figure 1 Stereo diagram showing the $2F_o - F_c$ electron density map and refined atomic model at the interface of the SH3, linker and kinase domains. The linker intercalates between the SH3 domain and the N-terminal lobe of the kinase domain, interacting extensively with both. The map is contoured at 1.5σ, and was calculated using data between 20 and 1.5 Å resolution. The SH3 domain is yellow, the linker red, and the kinase domain light blue. Ordered water molecules are shown as blue spheres.

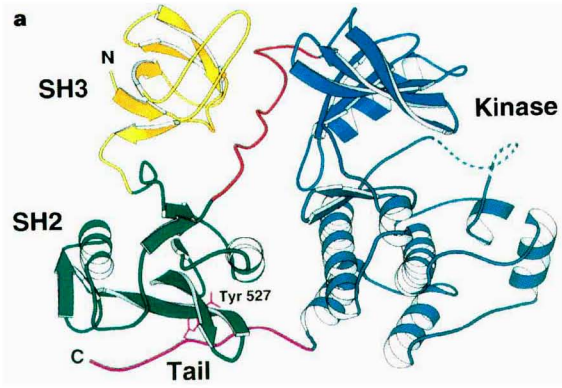
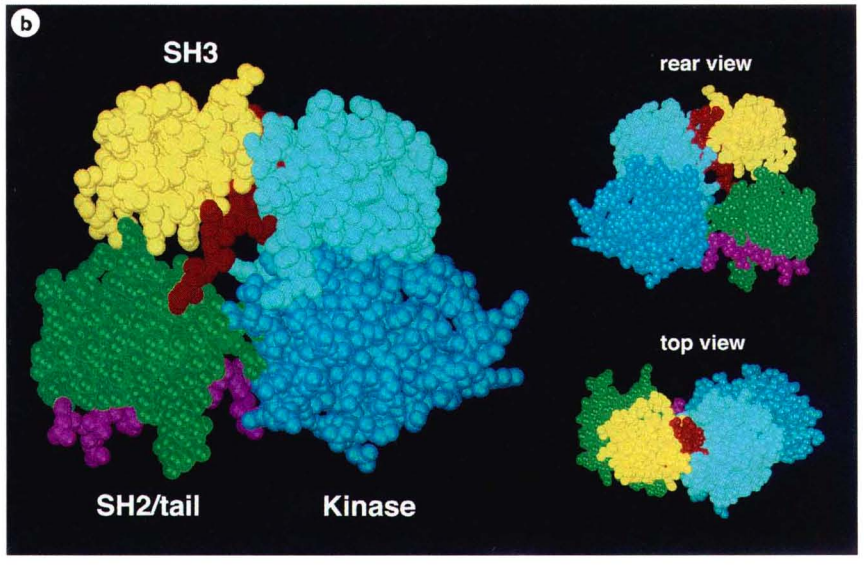


Figure 2 The 'closed' conformation of Tyr 527 phosphorylated c-Src. **a**, Ribbon diagram showing the structure and organization of domains. The SH3 and SH2 domains are coloured yellow and green, respectively. The linker connecting the SH2 and kinase domains (red) forms a polyproline II helix in complex with the SH3 domain. The N- and C-lobes of the catalytic domain are shown in cyan and blue. The disordered portion of the activation segment is shown as a dashed line. The phosphorylated tail (purple) is bound to the SH2 domain. **b**, Space-filling model. The SH2 domain makes only modest contact with the rest of the molecule. The RT and n-Src loops of the SH3 domain wrap around the linker to contact the N-lobe of the kinase in the front and rear, respectively.



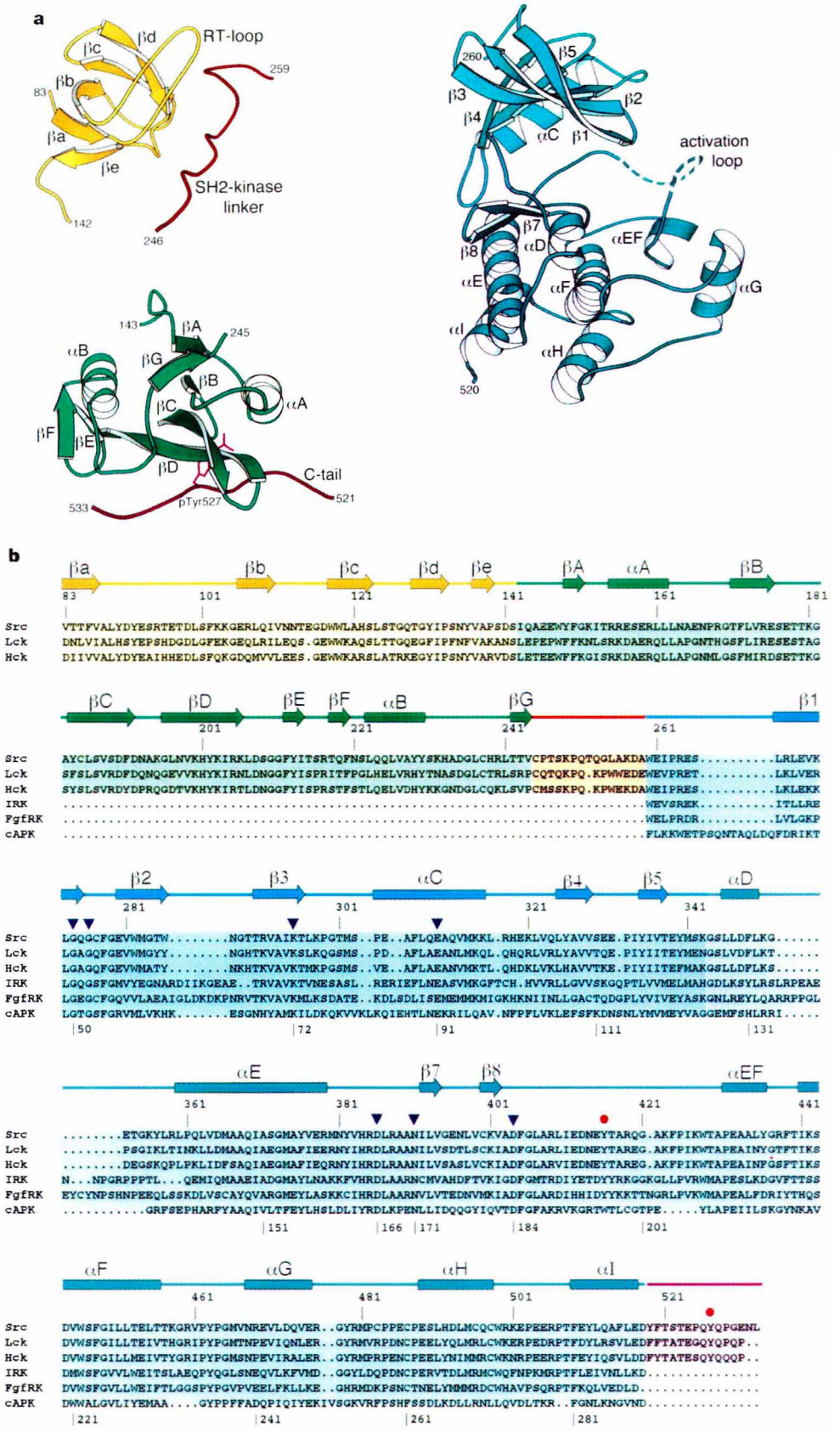


Figure 3 a, Ribbon diagrams of the SH3, SH2 and tyrosine kinase domains. Elements of secondary structure are labelled, using previously established nomenclature for the SH3, SH2 and catalytic domains. **b**, Structure-based sequence alignment, coloured by domain as in **a**. Sequences of the Src-family kinases Lck and Hck, the FGF (FgfRK) and insulin (IRK) receptor tyrosine kinases and the cAMP-dependent serine kinase (cAPK) are shown. Active-site residues are indicated by inverted triangles; major regulatory phosphorylation sites by red dots (tyrosines 416 and 527 in c-Src). Residue numbers for c-Src and cAPK are shown above and below the sequence alignment, respectively.

unexpected interaction of the SH3 domain with the 'linker' that joins the SH2 and catalytic domains. The compactly organized, highly ordered domain assembly pushes the two lobes of the catalytic domain close together and enforces a conformation in the small lobe that disables the active site. Thus, in addition to dephosphorylation by tail-directed tyrosine phosphatases, competitive interactions with SH3 or SH2 ligands could destabilize the observed conformation. The tightly coupled interdomain contacts suggest that any one of these inputs is likely to produce the open, activated structure as its output.

Structure determination

We crystallized human c-Src (residues 86–536) in a closed, inactive conformation. The structure includes the SH3, SH2, tyrosine kinase, and 'C-terminal tail' domains of the protein; it is phosphorylated only on Tyr 527 in the regulatory tail. (For historical consistency, we refer to the structure using the numbering of chicken c-Src. Residues 86–536 of the human sequence correspond to residues 83–533 of chicken c-Src.) The divergent amino-terminal domain was deleted because it is not thought to be integrated into the closed form of the kinase[18]. The recombinant protein was produced using a baculovirus vector in Sf9 insect cells[19]. Isolation of a defined phosphorylation state, with stoichiometric phosphorylation of Tyr 527, was critical for crystallization and was accomplished with anion-exchange and phosphotyrosine-affinity chromatography (W. Xu *et al.*, manuscript in preparation). Mass analysis verified that the purified protein was monophosphorylated.

The structure was determined using conventional heavy atom/multiple isomorphous replacement methods, and refined to a crystallographic R value of 21% (R_{free} = 26.4%) using data between 20.0 and 1.5 Å resolution. Phase determination and crystallographic refinement statistics are shown in Table 1. A portion of the electron density map and the refined atomic model are shown in Fig. 1. The refined model includes all residues except 410–423 in the 'activation segment' of the catalytic domain, which are not visible in our density maps and therefore must be disordered.

Overall structure and domain organization

In its closed regulatory state, c-Src is a compact ensemble of four structural domains (Figs 2 and 3). The SH2 and SH3 domains lie beside the large and small lobes, respectively, of the tyrosine kinase domain, on the side opposite the catalytic cleft. The SH2 domain binds the phosphorylated C-terminal tail (pTyr 527), which extends from the base of the adjacent catalytic domain.

A 14-residue polypeptide linker, which joins the SH2 and catalytic domains, interacts along its course with both the SH3 domain and the small lobe of the kinase domain and serves as an adaptor to fit the two together. Although it contains only one proline, the linker (residues 246–259) adopts a polyproline type II (PPII) helical conformation in complex with the recognition surface of the SH3 domain (Fig. 4a). The SH3 domain and the small lobe of the kinase also contact each other directly.

The two lobes of the catalytic domain are opposed even more tightly than they are in the 'active' form of cyclic AMP-dependent protein kinase A (cAPK)[20,21] (Figs 3a, 5a). Their close approach is accompanied by the displacement of an α-helix (helix C) in the small lobe from its position in active kinases, leading to rearrangement of residues that participate in catalysis (Fig. 5b). The conformation of the small lobe and its contacts across the catalytic cleft are strikingly similar to those seen in the uncomplexed (inactive) form of cyclin-dependent kinase 2 (Cdk2)[22,23]. The conformation of the kinase domain appears to be stabilized by interactions of its small and large lobes with the SH3 and SH2 domains, respectively. There are direct contacts between the regulatory and catalytic domains, in addition to binding of the phosphorylated C-terminal tail to the SH2 domain and of the linker to the recognition surface of the SH3 domain. The structural organization of these interactions suggests a significant degree of cooperativity.

Domain architecture and interactions

Three-dimensional structures have been determined for the isolated SH3 and SH2 domains of Src and several closely related kinases. These domains do not show any significant conformational

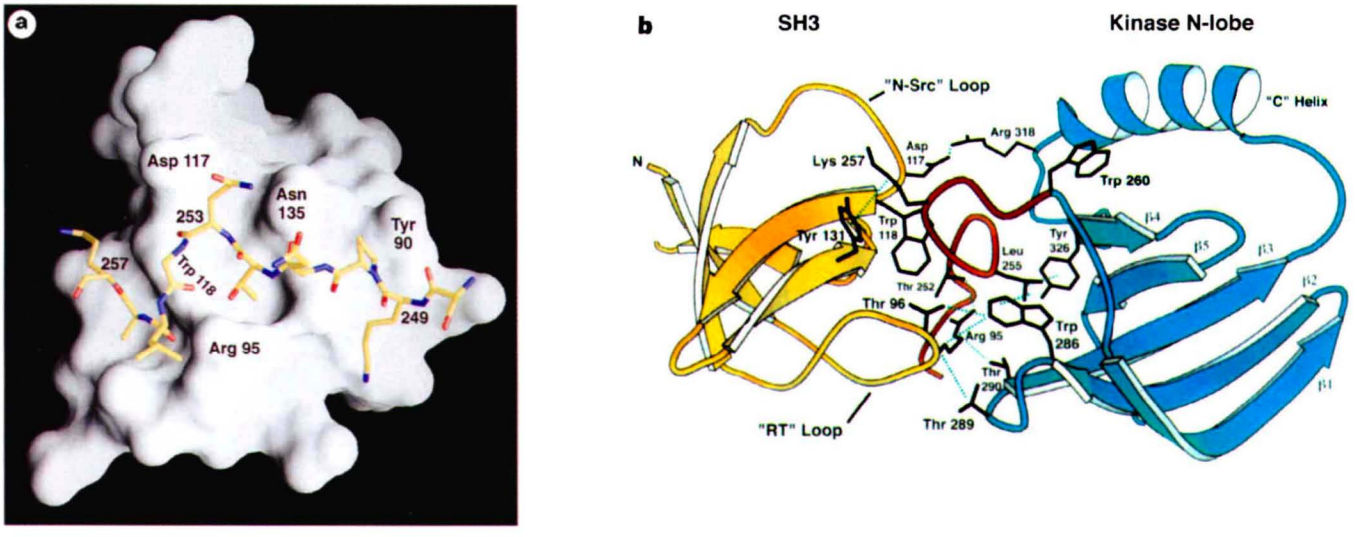


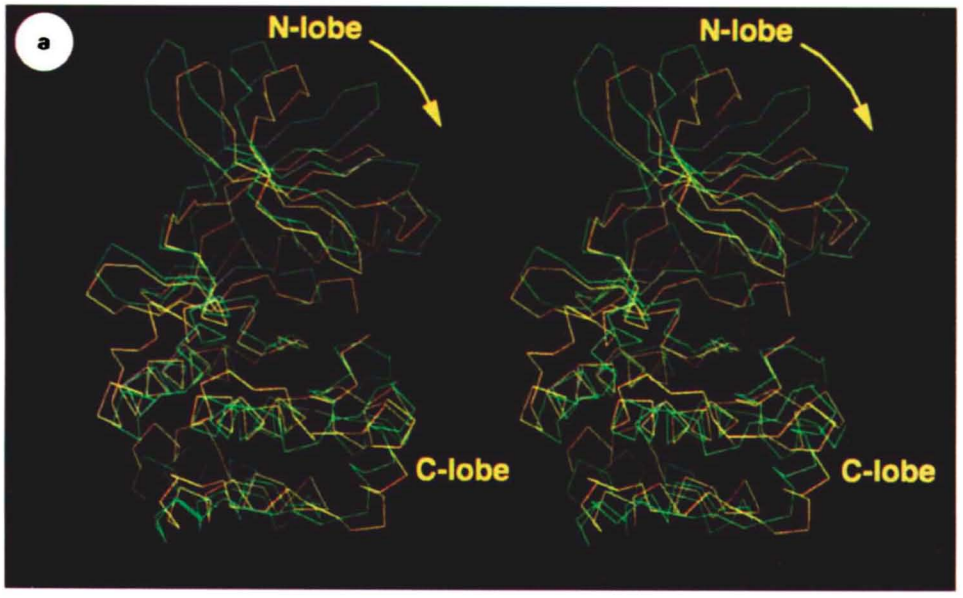
Figure 4 a, Molecular surface of the SH3 domain in complex with part of the linker segment. Residues in the domain are designated by name and number; those in the linker, by number only. The linker interacts in a class II orientation, with residues 249–253 in a polyproline type II (helical) conformation. C-terminal to Gln 253, the linker is not in a PPII conformation, but it continues to make extensive contact with the SH3 domain. Gly 254 introduces a kink, which allows the chain to arch over Trp 118. Leu 255 extends away from the surface of the SH3 domain and intercalates between Tyr 326 and Trp 286 in the kinase domain (see Fig. 1, and panel **b**). The C-terminal residues of the linker (256–259) form a β-turn, which also packs between the SH3 and kinase domains. **b**, 'Top' view of the interactions between the SH3, linker and kinase domains. The linker (red) 'glues' the SH3 domain to the N-terminal lobe of the kinase. In addition, the RT and N-Src loops of the SH3 domain contact the kinase directly. Hydrogen-bond and van der Waals contacts are indicated by dashed green lines. Trp 260, the first residue in the kinase domain, packs against the 'C' helix and anchors the C-terminal portion of the linker.

rearrangements in our structure. The large and small lobes of the tyrosine kinase domain are also very similar to the known structures of the insulin[24] and fibroblast growth factor (FGF) receptor tyrosine kinases[25] and to several protein serine kinases[26].

The SH3 domain. The SH3 domain (residues 83–142) is a compact five-stranded β-sandwich (Fig. 3a). Its ligand-binding surface is formed by a cluster of hydrophobic residues and is flanked by the 'RT' and 'n-Src' loops, which connect strands βa and βb, and strands βb and βc, respectively. The RT loop is a site of activating mutaions in v-Src[27–29]; the n-Src loop contains a six-residue insertion in the more active neuronal isoform of Src[30]. In our structure, the linker binds in the hydrophobic binding surface, and the RT and n-Src loops extend on either side of it to contact the catalytic domain (Figs 2b and 4). Residues 249–253 of the linker form a left-handed PPII helix and bind in a characteristic class II orientation, as seen with peptide ligands having PXXPXR consensus[31,32]. Proline 250, the only proline in this segment, packs between Tyr 90 and Tyr 136, the position usually assumed by the first proline in the class II binding sequence[31,32]. Gln 253 occupies the other site that normally requires proline. Its long polar side chain cannot intercalate into the binding cleft, and so the course of the linker deviates from that of proline-rich peptides at this point (Fig. 4a).

Interaction of the n-Src loop with the catalytic domain is modest; Asp 117 forms a salt bridge with Arg 318 in the kinase domain. The RT loop makes a more extensive contact. In particular, Arg 95 and Thr 96 make van der Waals and hydrogen bond contacts with the β2–β3 loop in the catalytic domain, and their side chains are in hydrophobic contact with Trp 286, the last residue in strand β2. The extended side chain of Arg 95 forms hydrogen bonds with Thr 252 and with the carbonyl of Leu 255 in the linker, and interacts through well-ordered water molecules with the hydroxyl groups of tyrosines 136 and 326 in the SH3 and catalytic domains, respectively (Fig. 4b).

Interaction between the SH3 and SH2 domains is minimal. Residues 142–146 at the SH2 amino terminus form a 3_{10} helical turn, which lies between small clusters of hydrophobic residues on each of the domains. The SH3/SH2 connection must be quite flexible; in a structure of an SH3/SH2 fragment of Lck, we observed a very different relative orientation of the two domains, even though the corresponding residues in Lck form an identical 3_{10} helical turn[33]. The difference in observed orientation is due entirely to backbone rotation about the two residues just N-terminal to this turn (Asp 141 and Ser 142). We suggested that dimeric interactions in the SH3/SH2 crystal structure might serve as a model for the cooperative role of the SH3, SH2 and tail domains in regulation[33].



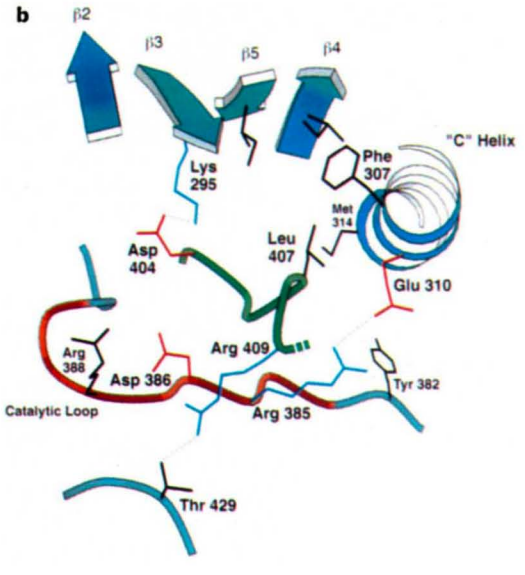


Figure 5 a, Alpha-carbon traces of the catalytic domains of c-Src (yellow) and cAPK (green), superimposed using the C-terminal lobes. In c-Src, interactions of the SH3 and SH2 domains rotate the N lobe by ~9° relative to its position in cAPK, closing the catalytic cleft. The break in the Src chain, seen at the right-hand edge of the central cleft, is the disordered part of the activation segment. **b**, The active site is partially disassembled. By analogy with cAPK, in an active conformation Glu 310 should contact Lys 295, positioning it to coordinate the α- and β-phosphates of ATP. In our structure, Glu 310 on helix C is rotated out of the catalytic conformation and is separated from Lys 295 by over 12 Å; its position in this inactive state is stabilized by an interaction with Arg 385. Leu 407, fixed in position by the anchoring of Arg 409, also enforces displacement of helix C. Phosphorylation of Tyr 416 may re-establish the catalytic configuration by coordinating Arg 385 and Arg 409, thereby relieving the electrostatic and steric barriers preventing helix C from assuming its active position. Other critical active-site residues (and in parentheses, their equivalents in cAPK) include the catalytic base, Asp 386 (Asp 166); Asp 404 (Asp 184), which coordinates a catalytic magnesium ion; and Arg 388 (Lys 168), which coordinates the γ-phosphate.

Our structure shows that this model is not relevant to the inactive state.

The SH2 domain. The conserved SH2 domain fold includes a central four-stranded β-sheet and two α-helices, which pack on either side of it (Fig. 3a). Numerous X-ray and NMR structures of SH2 domains in complex with phosphotyrosine-containing peptides show similar modes of recognition—phosphopeptides bind in an extended conformation across the surface of the domain, roughly perpendicular to the β-D edge of the central sheet[34]. Phosphotyrosine is bound in a pocket on one side of the central sheet, and in high-affinity complexes, residues C-terminal to phosphotyrosine bind in a pocket or groove on the opposite side. In Src-family SH2 domains, a hydrophobic pocket recognizes a leucine or isoleucine residue at position pY + 3 in high-affinity peptides[10,11]. The phosphorylated tail in our structure binds with the SH2 domain in a manner suggestive of a low-affinity interaction[34], consistent with experimental measurements of the affinity of the isolated Src SH2 domain for phosphopeptides modelled on the C-terminal tail[35]. The conformation and interactions of residues pY − 1 to pY + 2 (Gln-pTyr-Gln-Pro) in the tail are the same as those made by the corresponding residues in high-affinity polyoma middle-T sequence in previous Src and Lck SH2 structures[10,11], but outside this region the tail is poorly ordered and does not appear to make specific interactions. In particular, there is no side chain occupying the pY + 3 pocket.

The phosphorylated tail is a short but flexible tether that anchors the SH2 domain. Apart from the tail, the SH2 domain makes contact with the catalytic domain only along its A-helix, which runs roughly antiparallel to the E-helix of the kinase domain. The helices are not closely packed. Their juxtaposed surfaces are electrostatically complementary, however, and several charged and polar side chains interact across their interface, which also contains a number of bound water molecules.

Tyrosine kinase. The c-Src catalytic domain has an unembellished protein kinase fold, comprising a smaller, N-terminal lobe connected by a flexible 'hinge' to a larger, C-terminal lobe[20]. The N-terminal lobe is a five stranded antiparallel β-sheet, with a single helix (helix C) connecting strands β3 and β4; the C-terminal lobe is mostly α-helical (Fig. 3a). With the exception of helix C, the small and large lobes of the Src kinase domain superimpose well on the corresponding parts of cAPK; for example, 134 α-carbon atoms in the core of the large domain superimpose with an r.m.s.d. of 1.7 Å.

The two lobes of the catalytic domain approach each other closely. Superposition on the active conformation of cAPK, using only residues of the large lobe, shows that the small lobe in Src is rotated towards the large lobe by 9° relative to cAPK, creating an even more occluded interface (Fig. 5a). Superposition of the small lobes show that helix C in Src is displaced from this interface by ~5 Å and rotated away from it relative to its position in cAPK (Figs 3b, 5b). A precisely similar rearrangement is seen in the inactive form of Cdk2 (ref. 22). Helix C contains the conserved residue Glu 310 (Fig. 5b). Structures of kinases in their active conformations show that the side chain of this glutamate projects into the catalytic cleft to form a salt bridge with the equivalent of Lys 295, an important ligand for the α- and β-phosphates of ATP. In our structure (and in Cdk2), the glutamate faces outwards. It forms an alternative salt bridge with Arg 385, while Lys 295 interacts instead with Asp 404. The displaced helix C bears against Leu 407 near the beginning of the so-called 'activation segment'; this interaction prevents the helix from assuming its catalytic position (Fig. 5b). Leu 407 is the second residue in a β-turn, which is firmly anchored by insertion of the side chain of Arg 409 into a pocket near Thr 429. Residues following Arg 409 are disordered. In Cdk2, it is also a structured region near the beginning of the activation segment that displaces helix C (the PSTAIRE helix)[22]. Thus the same local inactivation mechanism is used in both Src and Cdk2. However, the molecular interactions controlling this local switch are quite different. Cdk2 is activated by interaction with cyclin A, which binds helix C and pushes it into its active conformation[23]. In Src, restoration of helix C to its active position is probably accomplished by phosphorylation of Tyr 416, and perhaps by disruption of the intramolecular SH3 and SH2 interactions, as discussed below.

Activation of the Src kinase

Reorientation of helix C to create a functional kinase would require a conformational change in the activation segment to remove interference from Leu 407. In a recent structure of the kinase domain of Lck, activated by phosphorylation at the position equivalent to Tyr 416, the phosphotyrosine forms salt bridges with the equivalents of both Arg 385 and Arg 409 (ref. 36). The homologous interactions in activated Src would necessarily expel Arg 409 from the pocket that it occupies in our structure, leading to rearrangement and ordering of the entire activation segment (Fig. 5b). Moreover, the side chain of Arg 385 would move away from Glu 310. Helix C would then be free to shift back to the position required for an active catalytic site.

How is this local switch, mediated by a concerted set of salt-bridge rearrangements, coupled to the global regulatory switch, mediated by the SH3 and SH2 domains? The structure described here suggests that the observed interactions of the regulatory domains with the linker, the C-terminal tail, and the kinase domain itself, all serve to push the lobes of the kinase together, reducing access to the catalytic cleft and forcing helix C to shift outwards. The Hck structure, reported elsewhere in this issue[37], shows that very similar SH3 and SH2 contacts are also compatible with some opening up of the lobes of the catalytic domain, perhaps due to the presence of quercitin in those crystals, which binds at the adenine site. Nonetheless, helix C is still in its displaced (inactive) position in the tail-phosphorylated Hck. Src phosphorylated on Tyr 527 in addition to Tyr 416 retains about 20% of its kinase activity[38]. Catalysis must require the 'active' position of helix C. It is not yet clear, however, whether catalysis can occur with the regulatory domains clamped in place, or if it requires release of the C-terminal tail from the SH2 domain and the linker from the SH3. The structures of other phosphorylated states of Src will be useful for answering these questions.

The regulatory switch

The structure and regulation of Src and its family members have been analysed extensively by site-specific mutagenesis and by comparison of c-Src and v-Src[2]. Introduction of isolated mutations into the SH3, SH2, kinase and tail regions of c-Src can be sufficient to create a constitutively active, and potentially transforming protein. In v-Src from Rous sarcoma virus, an unrelated 12-residue fragment replaces the last 19 residues of c-Src[12], removing the final turn of the 'I' helix and the entire regulatory tail. Various strains of the virus bear additional point mutations in the v-Src SH3 and catalytic domains[33]. The tail mutation alone activates the kinase and leads to a transforming protein, suggesting that the conformation we observe is destabilized in the absence of the SH2/tail interaction. Some of the point mutations are in solvent-exposed residues, and these would not be expected to affect the structure or activity of the kinase, but others, including tryptophan swapped for arginine at position 95 (Arg95Trp), Thr97Ile, Asp117Asn and Thr338Ile, are found at domain interfaces and could further destablize the closed conformation. Introduced on their own, substitutions for Arg 95 and Thr 96 in the RT loop are activating and partially transforming[27,28], and we would indeed expect them to disrupt the interaction of the SH3 domain with the linker and with the N-terminal lobe of the catalytic domain, thereby destabilizing the closed conformation. Likewise, we would expect mutation of Asp 117 to Asn to disrupt the interaction of the n-Src loop with the catalytic domain. Transformation by the isolated Thr 338 → Ile mutation is intriguing[27]. This residue is at the back of the

nucleotide-binding pocket, near the interdomain hinge. Introduction of isoleucine at this position would disrupt local hydrogen-bond interactions and might also require a small domain rotation, which could destabilize the closed conformation.

Spontaneous single-site mutations of Glu 378 to glycine, and of Ile 441 to phenylalanine, confer transforming activity on c-Src[39]. These residues pack together in the three-dimensional structure, just under Tyr 382, which helps to stabilize Glu 310 in its inactive position (Fig. 5b). Thus, disruption of this interaction may trigger activation of the kinase.

The interdependence of Tyr 527 phosphorylation and the SH3 and SH2 domains in regulation has been explicitly examined in a yeast system by coexpressing various Src mutants and Csk. These studies demonstrate that down-regulation of the kinase by phosphorylation of the tail requires both the SH3 and SH2 domains and, in particular, an intact ligand-binding surface and RT loop in the SH3 domain[29,40–42]. The present structure shows that the SH3 domain binds the linker and the N-lobe of the kinase with this surface, and in turn positions the SH2 domain to interact with the kinase and phosphorylated tail.

A number of cellular inputs, including tail dephosphorylation or apposition of a high-affinity ligand for the SH2 or the SH3 domain[52], could produce a transition to the open, active conformation. The flexibility of the SH3/SH2 interaction and the long linker between the SH2 and kinase domains suggest that the open conformation might be relatively floppy, with little interaction among the component domains. The regulatory domains would be free in such a state to bind their cognate cellular targets and to direct activated c-Src to its appropriate substrate and subcellular location. □

Methods

Protein purification. Insect cells bearing c-Src(ΔM85)[19] were lysed in 150 mM NaCl, 25 mM HEPES, pH 7.6, and 5 mM DTT. The lysate was cleared by ultracentrifugation, and c-Src(ΔN85) was isolated by sequential column chromatography on DEAE–Sepharose CL-6b, γ-aminophenyl ATP–Sepharose[43], and Superdex-200 (Pharmacia). Phosphorylation states were defined by mass spectroscopy, routinely identified by native PAGE electrophoresis (Phastegel, Pharmacia), and separated by phosphotyrosine and Q-Sepharose HP chromatography. Purified protein was maintained in a storage buffer containing 20 mM HEPES, pH 7.6, 0.1 M NaCl and 5 mM DTT. The yield of Tyr 527-phosphorylated material was increased by incubating non-phosphorylated protein (1 mg ml^{-1}) in storage buffer with purified recombinant Csk (1 mg ml^{-1}), in the presence of 10 mM $MnCl_2$ and 1.0 mM ATP.

Crystallization and data collection. Src crystals were grown in hanging drops by combining 1 μl protein solution (15 mg ml^{-1} protein in storage buffer) with 1 μl reservoir solution (50 mM PIPES, pH 6.5, 0.8 M sodium tartrate, 20 mM DTT). Crystals grow to maximum size ($0.25 \times 0.25 \times 0.6$ mm^3) in about a week at room temperature. Different crystal forms can grow in the same drop; the form we used belongs to space group $P2_12_12_1$, and has cell dimensions $a = 51.76$ Å, $b = 87.38$ Å, $c = 101.30$ Å, with one molecule per asymmetric unit. For derivatization and cryogenic data collection, crystals were transferred stepwise into 50 mM PIPES, pH 6.5, 1.15 M sodium tartrate, 15% glycerol and 0.1 M NaCl, and allowed to stabilize for at least 24 h before transfer to a final stabilizing solution containing 20% glycerol, 20 mM PIPES, pH 6.5, 1.15 M sodium tartrate and 0.1 M NaCl.

Diffraction data from native and derivative crystals were recorded with a Mar Research image plate scanner mounted on an Elliot GX-13 rotating anode source with mirror optics, and integrated and scaled with the programs DENZO and SCALEPACK[44] or XDS and XSCALE[45] (Table 1). A high-resolution native data set, complete to ~ 2.1 Å but ~ 55% complete between 2.1 and 1.5 Å, was recorded (from the same native crystal used for phasing) on an ADSC 1K CCD detector at the CHESS F-1 beamline (Cornell University). The synchotron data set was merged with the native 1 data set to create native 2, the data set used in refinement. We estimate an 'effective' resolution of 1.69 Å, based on the number of reflections observed with $I/\sigma > 2$.

Structure determination and refinement. The structure was determined by conventional multiple isomorphous replacement (MIR) using $KAu(CN)_2$ (1mM, 2 days), di-μ-iodobis (ethylene diamine) diplatinum (II) nitrate (PIP) (50 μM, 18 h) and a double-soak ($KAu(CN)_2$ and PIP) as derivatives (Table 1). Heavy-atom positions were located by Patterson and difference Fourier methods with the CCP4 program package[46]. Heavy-atom parameters were refined and phases were calculated with MLPHARE (overall figure of merit was 0.483 at 25.0–2.0 Å resolution). The initial MIR map was improved with real-space density modification using the program DM[46]. Skeletonization of the improved map with BONES[47] allowed us to fit models of the Lck SH3 domain[11], v-Src SH2 domain[10], and N and C domains of the insulin-receptor kinase[24] as rigid bodies. This model was rebuilt to reflect the sequence of human c-Src and refitted to the map using program O[47]. After an initial cycle of crystallographic refinement using XPLOR[48], the linker and tail regions were built. The model was refined with simulated annealing and positional refinement and manual rebuilding using the programs XPLOR[48] and O[47]. Simulated-annealing omit maps were computed to check the conformation of C-terminal tail. Water molecules were added with the aid of the program ARP[49]. Refinement statistics are given in Table 1. There are no residues with disallowed main-chain torsion angles. The refined model includes residues 83–409, 424–533, the N-terminal methionine, and 490 water molecules. Phe 424 is modelled as an alanine. The main-chain density is well-defined except for residues 521–525 and 530–533 in the 'tail' region which are poorly ordered and probably adopt multiple conformations.

Illustrations. Figures 1 and 2b were prepared with O[47]; Figs 2a, 3a, 4b and 5b with MOLSCRIPT[50]. The program GRASP[51] was used to create Figs 4a and 5a.

Received 13 December 1996; accepted 22 January 1997

1. Bishop, J. Viral Oncogenes. *Cell* **42,** 23–28 (1985).
2. Brown, M. T. & Cooper, J. A. Regulation, substrates and functions of sec. *Biochim. Biophys. Acta* **1287,** 121–149 (1996).
3. Superti-Furga, G. & Courtneidge, S. A. Structure–function relationships in Src family and related protein tyrosine kinases. *Bioessays* **17,** 321–330 (1995).
4. Pawson, T. Protein modules and signalling networks. *Nature* **373,** 573–580 (1995).
5. Yu, H. *et al.* Solution structure of the SH3 domain and Src and identification of its ligand-binding site. *Science* **258,** 1665–1668 (1992).
6. Musacchio, A., Saraste, M. & Willmanns, M. High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nature Struct. Biol.* **1,** 546–551 (1994).
7. Ren, R., Mayer, B. J., Cicchetti, P. & Baltimore, D. Identification of a ten amino acid proline-rich SH3 binding site. *Science* **259,** 1157–1161 (1993).
8. Mayer, B. J., Jackson, P. K. & Baltimore, D. 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 (1991).
9. Songyang, Z *et al.* SH2 domains recognize specific phosphopeptide sequences. *Cell* **72,** 767–778 (1993).
10. Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D. & Kuriyan, J. 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 (1993).
11. Eck, M. J., Shoelson, S. E. & Harrison, S. C. Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of $p56^{lck}$. *Nature* **362,** 87–91 (1993).
12. Takeya, T. & Hanafusa, H. Structure and sequence of the cellular gene homologous to the RSV *src* gene and the mechanism for generating the transforming virus. *Cell* **32,** 881–890 (1983).
13. Hunter, T. A tail of two src's: mutatis mutandis. *Cell* **49,** 1–4 (1987).
14. Nada, S., Okada, M., MacAuley, A., Cooper, J. A. & Nakagawa, H. Cloning of a complementary DNA for a protein-tyrosine kinase that specifically phosphorylates a negative regulatory site of p60c-src. *Nature* **351,** 69–72 (1991).
15. Matsuda, M., Mayer, B. J., Fukui, Y. & Hanafusa, H. Binding of transforming protein, P47gag-crk, to a broad range of phosphotyrosine-containing proteins. *Science* **248,** 1537–1539 (1990).
16. Roussel, R. R., Brodeur, S. R., Shalloway, D. & Laudano, A. P. Selective binding of activated pp60c-src by an immobilized synthetic phosphopeptide modeled on the carboxyl terminus of pp60c-src. *Proc. Natl Acad. Sci. USA* **88,** 10696–10700 (1991).
17. Cooper, J. A. & Howell, B. The when and how of Src regulation. *Cell* **73,** 1051–1054 (1993).
18. Koegl, M., Courtneidge, S. A. & Superti-Furga, G. Structural requirements for the efficient regulation of the Src protein tyrosine kinase by Csk. *Oncogene* **11,** 2317–2329 (1995).
19. Ellis, B. *et al.* Purification and characterization of deletional mutations of pp60c-src tyrosine kinase. *J. Cell. Biochem.* (suppl.) **18B,** 276 (1994).
20. Knighton, D. R. *et al.* Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253,** 407–414 (1991).
21. Madhusudan *et al.* cAMP-dependent protein kinase: crystallographic insights into substrate recognition and phosphotransfer. *Protein Sci.* **3,** 176–187 (1994).
22. De Bondt, H. L. *et al.* Crystal structure of cyclin-dependent kinase 2. *Nature* **363,** 595–602 (1993).
23. Jeffrey, P. D. *et al.* Mechanism of CDK activation revealed by the structure of a cyclinA–CDK2 complex. *Nature* **376,** 313–320 (1995).
24. Hubbard, S. R., Wei, L., Ellis, L. & Hendrickson, W. A. Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature* **372,** 746–754 (1994).
25. Mohammadi, M., Schlessinger, J. & Hubbard, S. R. Structure of the FGF receptor tyrosine kinase domain reveals a novel autoinhibitory mechanism. *Cell* **86,** 577–587 (1996).
26. Johnson, L. N., Noble, M. E. & Owen, D. J. Active and inactive protein kinases: structural basis for regulation. *Cell* **85,** 149–158 (1996).
27. Kato, J. Y. *et al.* Amino acid substitutions sufficient to convert the nontransforming p60c-src protein to a transforming protein. *Mol. Cell. Biol.* **6,** 4155–4160 (1986).
28. Potts, W. M., Reynolds, A. B., Lansing, T. J. & Parsons, J. T. Activation of pp60c-src transforming potential by mutations altering the structure of an amino terminal domain containing residues 90–95. *Oncogene Res.* **3,** 343–355 (1988).
29. Superti-Furga, G., Fumagalli, S., Koegl, M., Courtneidge, S. A. & Draetta, G. Csk inhibition of c-Src activity requires both the SH2 and SH3 domains of Src. *EMBO J.* **12,** 2625–2634 (1993).

30. Levy, J. B. & Brugge, J. S. Biological and biochemical properties of the c-src+ gene product overexpressed in chicken embryo fibroblasts. *Mol. Cell. Biol.* **9,** 3332–3341 (1989).
31. Feng, S., Chen, J. K., Yu, H., Simon, J. A. & Schreiber, S. L. Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science* **266,** 1241–1247 (1994).
32. Lim, W. A., Richards, F. M. & Fox, R. O. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* **372,** 375–379 (1994).
33. Eck, M. J., Atwell, S. K., Shoelson, S. E. & Harrison, S. C. Structure of the regulatory domains of the Src-family tyrosine kinase Lck. *Nature* **268,** 764–769 (1994).
34. Kuriyan, J. & Cowburn, D. Modular peptide recognition domains in eukaryotic signaling. *Annu. Rev. Biophys. Biomol. Struct.* (in the press).
35. Payne, G., Shoelson, S. E., Gish, G. D., Pawson, T. & Walsh, C. T. Kinetics of p56lck and p60src Src homology 2 domain binding to tyrosine-phosphorylated peptides determined by a competition assay or surface plasmon resonance. *Proc. Natl Acad. Sci. USA* **90,** 4902–4906 (1993).
36. Yamagushi, H. & Hendrickson, W. A. Structural basis for activation of the human lymphocyte kinase Lck upon tyrosine phosphorylation. *Nature* **384,** 484–489 (1996).
37. Sicheri, F., Moarefi, I. & Kuriyan, J. *Nature* **385,** 602–609 (1997).
38. Boerner, R. J. *et al.* Correlation of the phosphorylation states of pp60 c-src with tyrosine kinase activity: the intramolecular pY530-SH2 complex retains significant activity if Y419 is phosphorylated. *Biochemistry* **35,** 9519–9525 (1996).
39. Levy, J. B., Iba, H. & Hanafusa, H. Activation of the transforming potential of p60c-src by a single amino acid change. *Proc. Natl Acad. Sci. USA* **83,** 4228–4232 (1986).
40. Murphy, S. M., Bergman, M. & Morgan, D. O. Suppression of c-Src activity by C-terminal Src kinase involves the c-Src SH2 and SH3 domains: analysis with *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **13,** 5290–5300 (1993).
41. Okada, M., Howell, B. W., Broome, M. A. & Cooper, J. A. Deletion of the SH3 domain of Src interferes with regulation by the phosphorylated carboxyl-terminal tyrosine. *J. Biol. Chem.* **268,** 18070–18075 (1993).
42. Erpel, T., Superti-Furga, G. & Courtneidge, S. A. Mutational analysis of the Src SH3 domain: the same residues of the ligand binding surface are important for intra- and intermolecular interactions. *EMBO J.* **14,** 963–975 (1995).
43. Haystead, C. M., Gregory, P., Sturgill, T. W. & Haystead, T. A. Gamma-phosphate-linked ATP-sepharose for the affinity purification of protein kinases. Rapid purification to homogeneity of skeletal muscle mitogen-activated protein kinase kinase. *Eur. J. Biochem.* **214,** 459–467 (1993).
44. Otwinowski, Z. in *Proceedings of the CCP4 Study Weekend* (eds Sawyer, L., Isaacs, N. & Burley, S.) 56–62 (SERC Daresbury Laboratory, Daresbury, UK, 1993).
45. Kabsch, W. Evaluation of single crystal diffraction data from a position sensitive detector. *J. Appl. Crystallogr.* **21,** 916–924 (1988).
46. Collaborative Computational Project Number 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–776 (1994).
47. Jones, T. A., Bergdoll, M. & Kjeldgaard, M. in *Crystallographic Computing and Modeling Methods in Molecular Design* (eds Bugg, C. & Ealick, S.) (Springer, New York, 1989).
48. Brunger, A. T. *X-PLOR Version 3.0: A System for Crystallography and NMR* (Yale University Press, New Haven, CT, 1992).
49. Lamzin, V. S. & Wilson, K. S. Automated refinement of protein models. *Acta Crystallogr. D* **49,** 129–147 (1993).
50. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24,** 946–950 (1991).
51. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11,** 281–296 (1991).
52. Alexandropoulos, K. & Baltimore, D. Coordinate activation of c-Src by SH3- and SH2-binding sites on a novel p130Cas-related protein, Sin. *Genes Dev.* **10,** 1341–1355 (1996).

Acknowledgements. We thank B. Ellis, L. Overton, C. Hoffman, R. Boerner, K. Blackburn, W. B. Knight, M. Milburn and M. Luther of Glaxo-Wellcome for the c-Src(ΔN85) insect cell pellet, for mass spectroscopic analysis, and for discussions; M. Lawrence and the staff at CHESS for help with synchrotron data collection; P. Cole for purified recombinant Csk; K. Svenson for technical assistance; R. Nolte for help with crystallographic computation; J. Kuriyan and co-workers, for communicating results prior to publication; and T. Sweeney for help in preparing the manuscript. W.X. is supported by the Irvington Institute for Medical Research. This work was funded in part by a grank from BASF Bioresearch Corporation to M.J.E. S.C.H. is an investigator in the Howard Hughes Medical Institute. M..J.E. is the recipient of a Burroughs-Wellcome Fund Career award in the Biomedical Sciences.

Correspondence and requests for materials should be addressed to M.J.E. Coordinates have been deposited at the Brookhaven Protein Data Bank, accession code 1FMK, and are also available by e-mail from eck@rascal.med.harvard.edu.

Crystal structure of the Src family tyrosine kinase Hck

Frank Sicheri†, Ismail Moarefi*† & John Kuriyan*

*Laboratories of Molecular Biophysics, and * Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA*
† These authors contributed equally to this work.

The crystal structure of the haematopoietic cell kinase Hck has been determined at 2.6/2.9 Å resolution. Inhibition of enzymatic activity is a consequence of intramolecular interactions of the enzyme's Src-homology domains SH2 and SH3, with concomitant displacement of elements of the catalytic domain. The conformation of the active site has similarities with that of inactive cyclin-dependent protein kinases.

The Src-family kinases, named after the *src* oncogene of Rous sarcoma virus, are a closely related group of non-receptor tyrosine kinases that play critical roles in eukaryotic signal transduction. The viral and cellular forms of the *src* gene, v-*src* and c-*src*, were the first oncogene–proto-oncogene pair to be identified (reviewed in ref. 1). A critical difference between v-Src and c-Src is the loss in the former of a regulatory tyrosine residue that is located in the carboxy-terminal tail of c-Src (Tyr 527). Phosphorylation of Tyr 527 reduces the tyrosine kinase activity of c-Src, and without this inhibitory control the viral form of the protein greatly elevates cytoplasmic levels of tyrosine phosphorylation and is a potent cellular transforming agent[1]. With the aim of understanding the molecular basis for this critical signalling mechanism we have pursued a crystallographic investigation of the Src-family member Hck (haematopoietic cell kinase) in the auto-inhibited form.

The Src-family members share a common regulatory mechanism, but differ in cellular expression and localization. Nine Src-family tyrosine kinases have been identified (Src, Lck, Hck, Fyn, Fgr, Yes, Blk, Lyn and Yrk)[1]. c-Src is widely expressed and phosphorylates a wide range of substrates, whereas Lck plays a more restricted but critical role in T-cell signalling. Hck, the subject of this study, is expressed in lymphoid and myeloid cells[2,3], and is bound to B-cell receptors in unstimulated B cells. Knockout studies in mice have shown that simultaneous deletion of the genes for *hck* and *src*, or *hck* and *fgr*, leads to severe developmental anomalies and impaired immunity[4,5].

The highly conserved regulatory apparatus of the Src family members consists of two peptide-binding modules, the Src-homology domains SH2 and SH3 (refs 6, 7). These modules bind to targets containing phosphotyrosines and polyproline type II helices, respectively, and mediate the formation of protein–protein complexes during signalling[8]. Interactions between the SH2 domain and a C-terminal phosphotyrosine residue in the Src kinases (Tyr 527 in c-Src) results in repression of catalytic activity, with additional inhibitory interactions provided by the SH3 domain. The inhibitory phosphorylation at Tyr 527 is mediated by a distinct tyrosine kinase, Csk (c-Src kinase), whereas autophosphorylation at another tyrosine residue (Tyr 416 in c-Src), located within the 'activation segment' of the catalytic domain, is required for catalytic activity[1].

The catalytic domain alone is functional as a tyrosine kinase, but the SH2 and SH3 domains are required for full biological activity. Certain mutations in these domains lead to host cell-dependent phenotypes, indicating that they play a dual role in the Src kinases[1]. The SH2 and SH3 domains are required for inhibition of the enzyme, but once released from that role they function to target the kinase to specific substrates.

We now describe the crystal structure, determined at 2.6/2.9 Å resolution, of the downregulated form of the haematopoietic cell kinase, Hck. The SH2 domain, long implicated in regulating enzyme activity, is bound to the C-terminal phosphorylated tail, but is distant from the active site. Unexpectedly, the linker con-