Structure of the regulatory domains of the Src-family tyrosine kinase Lck

Michael J. Eck*, Shane K. Atwell*, Steven E. Shoelson† & Stephen C. Harrison‡

* Howard Hughes Medical Institute and Laboratory of Molecular Medicine, Children's Hospital, 300 Longwood Avenue, Boston, Massachusetts 02115, USA
† Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA
‡ Research Division, Joslin Diabetes Center, Department of Medicine, Brigham & Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA
§ Department of Biochemistry & Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

The kinase p56Lck (Lck) is a T-lymphocyte-specific member of the Src family of non-receptor tyrosine kinases1. Members of the Src family each contain unique amino-terminal regions, followed by Src-homology domains SH3 and SH2, and a tyrosine kinase domain. SH3 and SH2 domains mediate critical protein interactions in many signal-transducing pathways2. They are small, independently folded modules of about 60 and 100 residues, respectively, and they are often but not always found together in the same molecule. Like all nine Src-family kinases (reviewed in ref. 3), Lck is regulated by phosphorylation of a tyrosine in the short C-terminal tail of its catalytic domain4. There is evidence that binding of the phosphorylated tail to the SH2 domain inhibits catalytic activity of the kinase domain5,6 and that the SH3 and SH2 domains may act together to effect this regulation7. Here we report the crystal structures for a fragment of Lck bearing its SH3 and SH2 domains, alone and in complex with a phosphotyrosyl peptide containing the sequence of the Lck C-terminal regulatory tail. The latter complex represents the regulatory apparatus of Lck. The SH3-SH2 fragment forms similar dimers in both crystals, and the tail peptide binds at the intramolecular SH3/SH2 contact. The two structures show how an SH3 domain might recognize a specific target and suggest how dimerization could play a role in regulating Src-family kinases.

The 60-residue SH3 domain is a compact antiparallel β-sandwich composed of five strands βα–βε, a helical turn, and a long β-hairpin-like loop8–10. A putative ligand-binding site has been identified on the surface of SH3 domains based on locations of conserved residues11 and on shifts in NMR spectra recorded in the presence of proline-rich ligand peptides12. The site is a hydrophobic 'pouch' on the surface of the domain. The five residues that line this pouch in the Lck SH3 are shown in green in Fig. 1 (Tyr 72, Trp 97, Phe 110, Pro 112, Phe 115).
FIG. 1 Richardson diagram produced with the program MOLSCRIPT\textsuperscript{22}, showing the secondary structure and relative orientations of the SH3 and SH2 domains within a single molecule. The SH2 domain is at the top of the figure; the SH3 domain below and slightly to the rear. The \( \beta \)-sandwich of the SH3 is seen roughly 'on edge', looking into the presumptive binding groove. Note the minimal contact between domains. On the SH2 domain, amino-acid side chains that participate in high-affinity phosphotyrosyl-peptide binding are shown in red; the phosphotyrosine (pY) and pY + 3 binding pockets are labelled. The side chain of Leu 122, the first residue in the hydrophobic core of the SH2 domain, is shown in purple. The regions of the SH2 domain that interact with the SH3 domain at a dimer contact are highlighted in green. SH3 residues that are likely to participate in ligand binding (see text) are shown in green, and the 'ab' loop, which contacts the phosphopeptide in our crystalline complex, is highlighted in blue. Lower-case letters are used to label the \( \beta \)-strands of the SH3 domain, and upper-case letters are used for the SH2 domain. Turns are named according to the secondary structure elements that they join. The N and C termini are labelled.

FIG. 2 a and b. Solvent-accessible surface\textsuperscript{23} of the dimer-related SH3 and SH2 domains of the liganded structure. The SH3 surface is coloured green, the SH2 blue. The two domains together form the groove that binds the phosphorylated tail peptide. In b, a space-filling model of the peptide is added to the surface in a. c, Space-filling model of the SH3-SH2/tail structure, showing the 'head-to-tail' dimer interactions. The orientation is similar to that in a and b. Both domains of the one molecule are coloured blue, the dimer-related molecule is white. The tail peptide is coloured by atom type. Note the extensive interaction at the dimer interface and the relatively loose interaction between covalently linked SH3 and SH2 domains. The unliganded structure reveals essentially identical head-to-tail dimeric packing interactions. (Figure prepared using Insight II, Biosym, Inc.)
FIG. 4 Electron density maps and refined atomic models for (a), the SH3-SH2 structure, and (b), the SH3-SH2/tail structure. Both maps were calculated with (2Fo-Fc) coefficients and phases from the refined atomic models and show a region at the interface between dimer-related SH3 and SH2 domains. The model for the SH3 domain is coloured green, the SH2 yellow, and in the liganded structure the peptide is coloured orange.

METHODS. The SH3-SH2 protein (residues 53-226) was produced in a T7-based bacterial expression system and purified by phosphotyrosine affinity essentially as described for the Lck SH2 domain.14 The protein was maintained in a storage buffer containing 200 mM NaCl, 50 mM Tris, pH 8.5, 5 mM dithiothreitol, and 0.02% sodium azide. Unliganded crystals (space group P3_2_1; a = 82.51 Å, c = 116.22 Å) were grown in hanging drops by combining 1 μl of protein solution (15 mg ml^{-1} SH3-SH2 protein in storage buffer plus 10 mM o-phospho-tyrosine) with 1 μl reservoir solution containing 1.1 M sodium tartrate, 100 mM Tris, pH 8.5, and 0.01% 2-mercaptoethanol. Liganded crystals (space group R32, a = 72.35 Å, c = 187.36 Å) were grown in hanging drops by combining 2 μl protein/peptide solution (13 mg ml^{-1} SH3-SH2 protein in storage buffer plus 6 mM phosphopeptide TEGQpYQQQPA) with 2 μl of a well solution containing 950 mM sodium tartrate, 400 mM ammonium sulphate, and 100 mM Tris, pH 8.5. The phosphotyrosyl peptide was prepared as described.20 Data collection: Diffraction data for both structures were recorded at room temperature on a Mar Research image plate scanner mounted on an Elliot GX-13 rotating anode generator. Each data set was collected from a single crystal. One degree oscillation images were indexed, integrated and scaled using the program XDS.20 Both the liganded and unliganded structures were determined using molecular replacement. SH3-SH2 structure determination. The Lck SH2 domain,14 with phosphopeptide and solvent molecules omitted, was used as an initial search model. Rotation functions and translation searches were calculated with X-PLOR.21 With the two SH2 domains oriented and positioned in the asymmetric unit, we used the spectrum SH3 domain12 as a search model in Patterson rotation and translation functions calculated with AMORE.22 About 450 rotation peaks culled from searches calculated using a range of vector lengths were tested in translation searches, and a single rotation/translation solution had a correlation coefficient obviously above background. This solution placed the SH3 domain plausibly in the lattice, and a second SH3 domain was positioned using the non-crystallographic symmetry operator (calculated from the SH2 domains). Rigid body refinement with all four domains yielded an R-factor of 42.6% and a correlation coefficient of 0.515. The initial molecular replacement phases and electron density were improved by iterative real-space non-crystallographic symmetry averaging using the program RAVE,23 and the model was refined with multiple cycles of manual re-fitting and simulated annealing and positional refinement using X-PLOR. The program O (ref. 30) was used for all model building. Automated OMIT maps were used to relit several regions of the model, and tight non-crystallographic symmetry restraints were maintained during simulated annealing refinement steps. Restrained, individual temperature factors were refined. The current model has excellent geometry and contains residues 63-226 of Lck for both molecules in the asymmetric unit and no solvent molecules. The crystallographic R factor is 19.6% (R_{free} = 32.4%) for all data between 6.0 and 3.0 Å resolution. Although the crystallization required phosphotyrosine, no electron density was visible in the phosphotyrosine-binding pocket of the SH2 domain. SH3-SH2/tail structure determination. The SH3-SH2 structure was used as a search model in molecular replacement calculations carried out using X-PLOR.21 Rigid-body minimization of the properly rotated and translated model yielded a crystallographic R value of 42.2% for data between 12.0 and 3.2 Å resolution. The phosphopeptide model was built into the electron density map after an initial cycle of simulated annealing refinement. The model was further refined with multiple cycles of manual re-fitting and simulated annealing and positional refinement.22 Restrained individual temperature factors were refined. The current model has excellent geometry and contains residues 63-226 of Lck, residues pT to pY of the phosphopeptide, and 58 water molecules. The crystallographic R factor is 19.0% (R_{free} = 32.9%) for all data between 5.0 and 2.5 Å resolution.
FIG. 3 a, Diagram showing non-covalent interactions at the SH3/SH2 dimer interface. Open circles show residues in the SH3 domain, filled circles the SH2 domain. Solid lines indicate non-polar contacts across the interface; dotted lines indicate hydrogen bonds or salt bridges. The interface of each SH3/SH2 dimer contact burial $\sim 500$ $\AA^2$ on each domain, or a total of $\sim 2,000$ $\AA^2$ (refs 23, 24). This is comparable to the buried surface in relatively strong protein/protein interactions$^a$. By contrast, the interface between covalently linked SH3 and SH2 domains in our structure buries only $\sim 225$ $\AA^2$ per domain. b, Aligned sequences of the SH3, SH2 and tail regions of the nine Src-family tyrosine kinases. Backbone of Abk, Atk and Csk. The last three kinases have consecutive SH3 and SH2 domains, but not the Src-family members lack a regulatory C-terminal tail. Regions of $\beta$-sheet and $\alpha$-helix, shaded and labelled within each domain. Residue numbering is indicated for Lck. Residues that participate in hydrophobic and polar interactions at the dimer interface are indicated by a $\phi$ and an asterisk, respectively. Polar interactions stabilizing the interface include salt bridges (Glu 73-Arg 184, Arg 186-Glu 96, Asp 187-His 70 or Lys 84) and a hydrogen bond (Asn 114- Pro 195 carbonyl). Note that these interactions are conserved within the Src family, but not in more distantly related kinases that lack regulatory tails. Residues in the 'ab' loop of the SH3 domain and the C-terminal tail that are likely to form an ion-pair are boxed (see text). Histidine 76 in the 'ab' loop forms a salt bridge with the terminal carboxyl group of the peptide in our structure. This peptide is one residue longer than the actual Lck tail, but the histidine side chain could equally well interact with the terminal carboxyl of the authentic peptide. The length of the tail is the same in other members of the Lck subfamily, and the histidine is present in all but Btk (where the Arg residue later could substitute). In the Src subfamily, the tail is two residues longer, but a negatively charged residue is present at the position corresponding to the C-terminal tail in Lck, and an arginine in the SH3 domain (at the position of Ser 75 in Lck) could salt-bridge to it. This arginine is the site of mutations that affect Src regulation (see text). c, Diagram showing how dimerization could regulate Src-family kinases. The SH3, SH2 and catalytic domains are shown schematically. The tail peptides are shown with a tyrosine side chain. A red star in the SH3 domain marks the location of its ligand binding surface and of the adjacent 'ab' loop, which contacts the tail peptide in our crystals. The N-terminal domains, which are myristoylated and bound to the cytoplasmic tails of CD4 or CD8, are not shown. The view of the dimer is from the membrane. The N termini of the SH3 domains turn out of the plane of the page (towards the membrane) and therefore could connect to the membrane-associated N-terminal domains. The structure suggests that phosphorylation of Tyr 505 could stabilize a dimer through SH3/tail interactions. The tail peptide interacts with both the SH2 and SH3 domains, increasing the total buried surface area of the interaction. In addition to tethering and perhaps inhibiting the kinase domain, the closed structure would leave the SH2 domain occupied and unavailable to bind a heterologous target, and the SH3 domain sequestered with its ligand-binding surface buried by the SH2 domain and unable to recognize a target proline-rich ligand. A number of events might trigger a transition to an 'open' state: dephosphorylation of Tyr 505, which would unblock the SH2 domain and probably weaken the dimer contact; apposition of a 'high-affinity' SH2 ligand, which could compete for binding with the tail peptide and repel the SH3 domain; or apposition of a ligand for the SH3 domain, which might dissociate the closed structure independently of any phosphorylation or dephosphorylation event. In an open state, the SH3 and SH2 domains would be free to bind their respective ligands. Auto phosphorylation of the catalytic domain (Tyr 394 in Lck) is also important for full activation of kinase activity. The SH3/SH2 dimer interactions seen here might orient the catalytic domains in a manner that would prevent cross-auto phosphorylation. We note that an inactive monomeric state can readily exist, if the C-terminal tail binds to the SH2 domain of the same polypeptide.
The 100-residue SH2 domain contains a central β-pleated sheet flanked by α-helices A and B (refs 11–14). The proximal face of the central sheet, the BC loop, and the amino-terminal residues of the A helix form a phosphotyrosine binding pocket. The distal face of the central sheet and the opposing ‘jaws’ made by the EF and BG loops create a second pocket capable of accepting a hydrophobic side chain (Fig. 1). The SH2 domains of Lck and Src bind particularly strongly to peptides containing the sequence pYEEI (ref. 15), and the structures of these domains with a bound ‘YEEI phosphopeptide show the peptide in an extended conformation and the phosphotyrosine (pY) and isoleucine (pY+3) inserted into the pockets just described16-16.

The polypeptide segment linking the two domains is quite short. Only two amino acids intervene between the last residue of strand β of the SH3 domain and Leu 122, which forms part of the SH2 hydrophobic core. The corresponding residue is either a leucine or isoleucine in all nine Src family members, which also have identical spacings between their SH3 and SH2 domains.

The consecutive SH3 and SH2 domains make very little contact with each other, aside from the polypeptide chain linking them (Fig. 1). This limited contact includes a hydrogen bond from the main-chain amide of Lys 118 in the SH3 domain to the side chain of Ghu 204 in the SH2 domain, and in the liganded molecule a hydrogen bond from Arg 207 to the carbonyl of Val 116. In addition, Pro 200 in the SH2 domain is in contact with the side chains of Leu 69 and Ala 117 in the SH3 domain. The ω-carbons of Tyr 181 (βD5) and Pro 112 (δ1), which mark the centres of the ligand-binding regions of the SH2 and SH3 domains, respectively, are 31 Å apart. The two domains could in principle bind to sites on the surface of a single larger ligand molecule or to successive proline and phosphotyrosine motifs on an extended peptide loop. But given the paucity of interactions between the two domains, we expect the link between them to be flexible, allowing interaction of the two domains with separate ligands.

The unit cell contains two crystallographically independent SH3–SH2 molecules. These two molecules have essentially identical conformations, and they pack together in the crystal related by a non-crystallographic 2-fold symmetry axis. This arrangement creates a 'head-to-tail' interaction between the two molecules, in which the SH3 of one molecule binds the SH2 of the other and vice versa (Fig. 2). In contrast to the minimal contact region seen between covalently linked SH3 and SH2 domains, the intermolecular SH3/SH2 contacts in the dimer are extensive, creating a large solvent-excluded area. The contact contains a large central hydrophobic core surrounded by a number of hydrogen bonds and salt bridges (Fig. 3a). On the SH3 domain, the interface includes residues from the 'α' loop, the helical 'de' turn, and the 'be' loop. On the SH2 domain, it includes portions of the βD, βE and βF strands and the EF turn (one of the 'jaws' of the pY+3 pocket).

The structure with the bound 'tail' peptide has the same overall arrangement of domains as in the unliganded SH3–SH2 fragment. There is only one molecule in the R32 asymmetric unit, but the crystallographic 2-fold axis generates a dimer that is nearly identical to the non-crystallographic dimer. Other crystal lattice contacts are very different. The contact between dimer-related SH3 and SH2 domains is as well conserved as the one between covalently linked domains. The phosphotyrosyl peptide (sequence TEGOpYQPOP) binds at the intermolecular SH3/SH2 contact (Fig. 2), and interacts with both domains. The phosphotyrosine side chain inserts into the pY pocket of the SH2 domain exactly as it does in the SH2/peptide complexes, and most of the interactions previously described are present. Arginine α2, which in the high-affinity YEEI complex participates in a number of hydrogen bonds, including an amino-aromatic bond to the phosphotyrosine ring, here donates only one hydrogen bond to the phosphate from its Ne; the guanidinium group turns away from the ligand. The conformation of the rest of the peptide differs substantially from that of the YEEI peptide bound to the Lck SH2 (ref. 14). Instead of forming an extended chain across the D-strand with anchors at pY and pY+3, the peptide appears to follow the seam between the SH2 domain and the apposed SH3 domain, with contacts to both domains along its length. The first four residues of the peptide are poorly ordered and are not included in the current model.

The glutamate at pY+1 adopts a conformation very similar to that of the glutamic acid at the corresponding position in the YEEI peptide complexes of Lck and Src, but the backbone of the peptide between pY+1 and pY+3 is kinked by Pro pY+2, and it is not in close contact with the surface of the SH2 domain.

The side chain of Pro pY+2 packs into the top of the pY+3 binding pocket. Histidine 76 (αB8) on the SH3 domain is positioned to bind the C-terminal carboxylate of the peptide.

The dimer interface between the two domains (Figs 3 and 4) suggests how an SH3 domain might recognize a proline-rich ligand. Proline 195 (EF2) on the SH2 domain projects into the hydrophobic pouch of the apposing SH3 domain. The proline lies in a van der Waals pocket defined by the side chains of Try 72, Try 97, Pro 112, and Phe 115. Additional hydrophobic contacts with the SH3 are made by SH2 residues Leu 186, Tyr 192, and Thr 198. It is possible that similar interactions might be made by the proline-rich sites believed to be the targets for SH3 domains17. As a motif for protein–protein interactions, proline-rich peptides might be rationalized as being a source of soluble hydrophobic residues. Interspersed non-proline residues might account for specificity. One target for the Lck SH3 domain is phosphatidylinositol-3-OH kinase18: two sequences within the p85 chain of this kinase (PFTPKPRPRPLP and APALPPKPKP) are known to bind the Fyn SH3 domain and have been proposed as the likely binding site for the Lck SH3 domain19. Both contain prolines interspersed with proline residues. Model building suggests that the PPRPPK sequence could make salt-bridge and van der Waals interactions with the Lck SH3 domain that are similar to those made by Arg 184, Tyr 192, Pro 195 and Arg 196 of the apposed Lck SH2 domain.

Does the dimeric association of SH3–SH2 fragments reveal an interaction that actually occurs at some stage of activation or inactivation in the Lck signal transduction pathway? We have not detected dimerization in solution of the SH3–SH2 protein in the presence or absence of phosphopeptide, and Src-family kinases are monomeric in detergent lysates19. But dimer formation will be favoured in the intact cell because the Src-family kinases are tethered to the membrane. Moreover, the Fyn SH3 domain binds directly to bead-immobilized Fyn SH2 or SH3–SH2 (H. Band, personal communication). In addition, the extent and intimacy of the intermolecular SH3/SH2 contact and its conservation in two otherwise different crystal forms suggest that it might be more than an adventitious packing interaction: (1) it is significantly more complementary than typical lattice interactions (Figs 2, 3 and 4); (2) most of the residues at the interface, including those involved in salt bridges and hydrogen bonds, are conserved within the Src family but not in other molecules that contain adjacent SH3–SH2 domains, such as kinases with no regulatory tail (Fig. 3); (3) SH3 residues that contact the tail peptide are conserved in the Lck subfamily (which has one type of tail sequence), and they are characteristically different in the Src subfamily (which has a different type of tail sequence) (Fig. 3b); (4) the dimeric interaction of the ab loop in the SH3 domain with the tail peptide provides a structural explanation for the transforming potential of mutations in the corresponding loop in the Src SH3 domain20,21. Introduction of even single amino-acid changes in the SH3 ab loop is sufficient to activate the catalytic activity of c-Src22. Moreover, one of these mutations, Arg 95 to Trp, diminishes the ability of Src to be regulated by phosphorylation of its C-terminal tail (Y527) (ref. 7). Our dimeric structure suggests that this mutation would destroy a salt bridge between Arg 95 and pY+4 in the tail and that it might also sterically hinder interaction of the SH3 with...
the SH2 domain. In contrast, within a monomeric SH3-SH2/tail complex, the αβ loop on the SH3 domain is remote from the phosphopeptide (Fig. 1), and any contact between the two seems unlikely, given the short hinge region between the covalently adjacent domains. In particular, the intermolecular SH3/SH2 contact in our crystals could not form intramolecularly without substantial unfolding of one or both domains.

These observations and correlations lead us to propose that the dimer structure may represent a model for the regulatory regions in a 'closed' state of Src-family kinases. In this state, the SH3 and SH2 domains and the C-terminal tails would all be sequestered and the kinase domains presumably oriented to prevent cross phosphorylation (Fig. 3c). Transition to an open, probably monomeric, state would free the SH3 and SH2 domains to bind heterologous ligands and to direct, together with the unique N-terminal segment, the specificity and cellular localization of the active kinase.

Received 21 January; accepted 7 March 1994.


ACKNOWLEDGEMENTS: We thank H. Band, L. Cantley, B. Mayer, M. Milburn and D. Winkler for discussion; H. Band for information before publication; and M. Musacchio, M. Nobile, M. Saraste and R. Wiernaga for atomic coordinates of the spectrin and Fyn SH3 domains. S.S. thanks the NSF for support. Coordinates will be deposited with the Brookhaven Protein Databank.