Recognition of tandem PxxP motifs as a unique Src homology 3-binding mode triggers pathogen-driven actin assembly

Olli Aitio[a], Maarit Hellman[a], Arunas Kazlauskas[b], Didier F. Vingadassalom[c], John M. Leong[c], Kalle Saksela[b], and Perttu Permi[a,1]

[a]Program in Structural Biology and Biophysics, Institute of Biotechnology, University of Helsinki, FI-00014, Helsinki, Finland; [b]Department of Virology, Haartman Institute, University of Helsinki, and HUSLAB, University of Helsinki Central Hospital, FI-00014, Helsinki, Finland; and [c]Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, MA 01655

Edited by John Kuriyan, University of California, Berkeley, CA, and approved October 29, 2010 (received for review July 15, 2010)

Src homology 3 (SH3) domains are globular protein interaction modules that regulate cell behavior. The classic SH3 ligand-binding site accommodates a hydrophobic PxxP motif and a positively charged specificity-determining residue. We have determined the NMR structure of insulin receptor tyrosine kinase substrate (IRTKS) SH3 domain in complex with a repeat from *Escherichia coli*-secreted protein F-like protein encoded on prophage U ($EspF_U$), a translocated effector of enterohemorrhagic *E. coli* that commandeers the mammalian actin assembly machinery. $EspF_U$-IRTKS interaction is among the highest affinity natural SH3 ligands. Our complex structure reveals a unique type of SH3 interaction based on recognition of tandem PxxP motifs in the ligand. Strikingly, the specificity pocket of IRTKS SH3 has evolved to accommodate a polyproline type II helical peptide analogously to docking of the canonical PxxP by the conserved IRTKS SH3 proline-binding pockets. This cooperative binding explains the high-affinity SH3 interaction and is required for $EspF_U$-IRTKS interaction in mammalian cells as well as the formation of localized actin "pedestals" beneath bound bacteria. Importantly, tandem PxxP motifs are also found in mammalian ligands and have been shown to contribute to IRTKS SH3 recognition similarly.

IRSp53 | PPII Helix

Src homology 3 (SH3) domains constitute a prototypic and ubiquitous class of modular protein-binding domains that guide interactions between proteins typically involved in cell signaling (1, 2). A hydrophobic groove on the SH3 surface is adapted to bind to target peptides that adopt a left-handed polyproline type II (PPII) helical conformation (3, 4). The majority of SH3 ligands contain a consensus sequence XPxXP (wherein X is generally hydrophobic residue and x is any residue). The XP dipeptides occupy two hydrophobic pockets on the SH3 ligand-binding groove, whereas a third slot ("specificity pocket") contacts additional residues flanking the XPxXP moiety. In Src family and several other SH3 domains, this specificity pocket is negatively charged and interacts with Arg or Lys in the ligand. This basic residue may be N-terminal (+xXPxXP, class I) or C-terminal (XPxXPx+, class II) relative to the conserved proline residues, and thereby determines the orientation of ligand binding (5, 6). Conversely, SH3 domains with divergent specificity pockets show preference for ligands with other types of flanking residues (1, 2). For example, the specificity pockets of Eps8-like SH3 domains have specialized in accommodating the dipeptide DY (7).

Insulin receptor tyrosine kinase substrate (IRTKS) and IRSp53 are related proteins serving as adaptors and effectors of filamentous (F-) actin assembly (8). Recently, they were also found to provide an essential link between two bacterial proteins that regulate host cell actin reorganization (9, 10). Enterohemorrhagic *Escherichia coli* (EHEC) serotype O157:H7 is an important diarrheal pathogen that triggers F-actin assembly in the cells directly beneath bound bacteria by injecting two effector proteins, namely, translocated intimin receptor (Tir) and *E. coli*-secreted protein F-like protein encoded on prophage U ($EspF_U$) (11–14).

$EspF_U$ contains multiple 47-residue repeats (Fig. S1). The N-terminal 33-residue region (Fig. 1, "H") of the $EspF_U$ repeats can activate the Wiskott–Aldrich syndrome protein (WASP) by contacting with its GTPase-binding domain to disrupt it from an intramolecular autoinhibitory interaction (15, 16). Our recent studies showed that the proline-rich C terminus of these repeats (Fig. 1, "P") mediates recruitment of $EspF_U$ to sites of bacterial attachment by binding to the SH3 domain of IRTKS, which acts as an adapter connecting $EspF_U$ to Tir (10). We observed robust binding between the IRTKS SH3 and the complete 47-residue repeat of $EspF_U$ ($R47_5$) but no binding to a 33-residue fragment ($R33_5$) lacking the C-terminal proline residues (10) (Fig. 1). Likewise, Stradal and colleagues (9) reported that IRSp53 SH3 can recognize the sequence IPPAPNWPAP in the C-terminal portion of the $EspF_U$ repeat.

In this study, we have used NMR spectroscopy as well as peptide array and mutagenesis approaches to characterize the molecular basis of the interaction between IRTKS SH3 and $EspF_U$ $R47_5$. Our IRTKS SH3:$EspF_U$ $R47_5$ complex structure reveals a unique type of SH3 interaction that involves accommodation of two adjacent PPII helical PxxP motifs as the structural basis of IRTKS-$EspF_U$ binding, explaining the high affinity and selectivity of this interaction.

Results

High-Affinity Complex Between IRTKS SH3 and $EspF_U$ $R47_5$. NH correlations of I_{27}, A_{30}, N_{32}, W_{33}, A_{35} and T_{37}, covering the entire proline-rich region of the fifth 47-residue repeat of $EspF_U$, $R47_5$ (Fig. 1, P), experienced large changes in their chemical shift in the ^{15}N-heteronuclear single quantum coherence (HSQC) spectrum on addition of an equimolar amount of unlabeled IRTKS SH3 (Fig. S2). This is in agreement with the previous binding site mapping studies (9, 10) but indicates that the SH3-binding epitope in the $R47_5$ region extends beyond the minimal IPPAPNWPAP peptide identified earlier (9). The rest of the spectrum remained poorly dispersed, indicating that only the

Author contributions: O.A., M.H., A.K., D.F.V., J.M.L., K.S., and P.P. designed research; O.A., M.H., A.K., D.F.V., and P.P. performed research; J.M.L., K.S., and P.P. contributed new reagents/analytic tools; O.A., M.H., A.K., D.F.V., J.M.L., K.S., and P.P. analyzed data; and O.A., D.F.V., J.M.L., K.S., and P.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The resonance assignments and coordinates of the IRTKS SH3:EspFU R475 complex structure have been deposited in the BioMagResBank (accession no. 16909) and Protein Data Bank, www.pdb.org (PDB ID code 2kxc), respectively.

[1]To whom correspondence should be addressed. E-mail: perttu.permi@helsinki.fi.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1010243107/-/DCSupplemental.

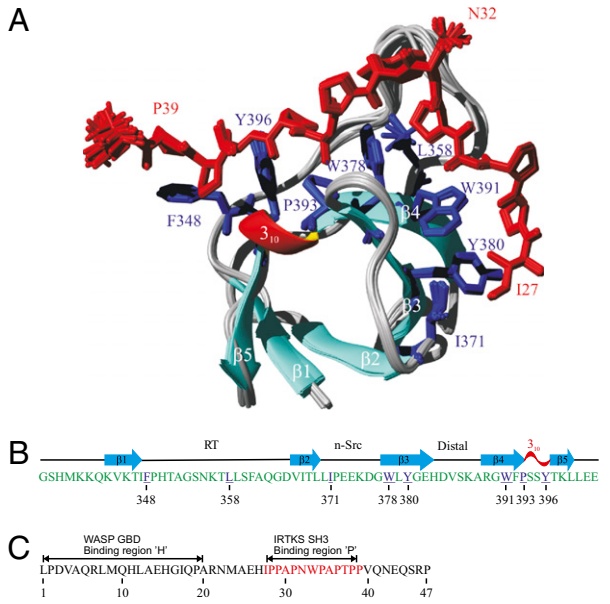


Fig. 1. (*A*) Ribbon presentation of the ensemble of 20 superimposed NMR structures of IRTKS SH3:EspF$_U$ R47$_5$ complex. The heavy atoms of R47$_5$ (residues 27–40) and IRTKS SH3 domain residues interacting with R47$_5$ are shown in red and blue, respectively. (*B*) Secondary structure elements of IRTKS SH3, with residues interacting with R47$_5$ underlined and colored in blue. (*C*) Sequence of R47$_5$, with amino acid numbering from 1 to 47, corresponding to residues 268–314 in full-length EspF$_U$. The 'H' and 'P' regions important for the EspF$_U$-WASP GTPase-binding domain (GBD) and EspF$_U$-IRTKS SH3 interactions are highlighted.

C-terminal part of R47$_5$ is involved in binding and that no conformational changes in R47$_5$ beyond the binding epitope occur on interaction.

We used both chemical shift perturbation (CSP) mapping and isothermal titration calorimetry (ITC) to determine the binding affinity of R47$_5$. The CSP mapping showed that IRTKS SH3 associates with R47$_5$ with an affinity that is unusually high for SH3 binding, because the induced ^{15}N/^{1}H chemical shift changes observed in ^{15}N-labeled IRTKS SH3 were in the regime of slow exchange in the NMR time scale (Fig. S2). This observation was confirmed more quantitatively using ITC, which indicated a dissociation constant (K_d) of 500 nM, ranking the IRTKS SH3:EspF$_U$ R47$_5$ interaction among the strongest found in naturally occurring ligands (Fig. S3).

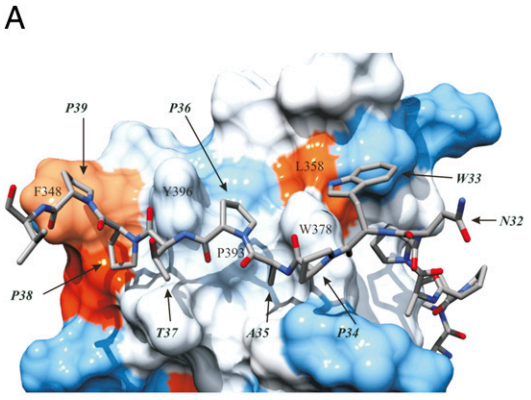
SH3 Binding Induces Folding of the EspF$_U$ Peptide. We next determined the solution structure of the human IRTKS SH3 domain in complex with EspF$_U$ R47$_5$ using triple-resonance NMR experiments (17–19). The assignment of ^{1}H/^{15}N/^{13}C resonances was performed using two samples containing a 1:1 molar ratio of IRTKS SH3:EspF$_U$ R47$_5$, wherein either the SH3 domain or R47$_5$ was uniformly ^{15}N, ^{13}C labeled. The ^{15}N-HSQC spectrum of SH3 displayed well-dispersed resonances characteristic for proteins with high β-sheet content. Unlike the SH3 domain, the ^{15}N-HSQC spectrum of free EspF$_U$ R47$_5$ exhibited poorly dispersed amide proton chemical shifts with heavily overlapping resonances, which is typical for disordered protein (20) (Fig. S2). Addition of an equimolar amount of SH3 induced chemical shift changes for few NH correlations in R47$_5$ arising from the residues interacting with SH3 (Fig. S2). Because the N-terminal segment and part of the C-terminal region of R47$_5$, which flank the SH3-binding epitope, also remained clearly disordered in the complex, we carried out the structure calculations using distance restraints (NOEs) arising from all SH3 protons and EspF$_U$ protons within residues 26–42. The structure ensemble, for residues 343–400 in IRTKS and residues 27–40 in EspF$_U$, is exquisitely well defined with a backbone and heavy atoms rmsd of 0.32 ± 0.08 Å and 0.62 ± 0.08 Å, respectively. A summary of structural statistics is shown in Table S1.

Structure of the IRTKS SH3:EspF$_U$ R47$_5$ Complex. The structure of the IRTKS SH3 in complex with EspF$_U$ R47$_5$ has a typical SH3 domain fold (Fig. 1*A*). Five β-strands (β1 to β5) and a short 3$_{10}$ helix establish the usual β-barrel fold, where the β2 strand is shared by two antiparallel β sheets. Strands β1 to β4 are connected by RT, n-Src, and distal loops, whereas the 3$_{10}$ helix connects β4 to β5. The structure of the mouse ortholog of IRTKS SH3 has been solved by NMR in free form (PDB ID code 1spk). Superimposition of secondary structure elements of 1spk and IRTKS SH3 domain gives a Cα rmsd of 0.68 Å, suggesting that the SH3 does not undergo any substantial structural rearrangements on binding to R47$_5$.

The peptide-binding surface extends from the classic PxxP motif-binding pockets involving the highly conserved SH3 residues F$_{348}$, W$_{378}$, and Y$_{396}$ to an unusual and extended specificity-determining region formed by hydrophobic residues in the n-Src loop and strands β3 and β4 (Fig. 1*A*). This region in IRTKS is unique among known SH3 structures and accounts for its unusual ligand-binding properties. In contrast to the single negatively charged specificity pocket characteristic to Src family and most other SH3 domains, this region of IRTKS SH3 includes two distinct hydrophobic pockets involving residues L$_{358}$, W$_{378}$, and W$_{391}$ and residues I$_{371}$, Y$_{380}$, and W$_{391}$, respectively. L$_{358}$ occupies the position that corresponds to the conserved acidic residue in typical SH3 domains that forms a salt bridge with the basic residue in class I and II ligands.

The binding epitope in EspF$_U$ consists of 13 residues (I$_{27}$PPAPNWPAPTPP$_{39}$). Remarkably, each SH3-contacting residue in the epitope provided ≈10 intermolecular NOEs, which implies firm anchoring of R47$_5$ to the IRTKS SH3 (Fig. S4). EspF$_U$ R47$_5$ adopts a class I binding orientation and an overall V-shaped conformation with the angle at the residue N$_{32}$ (Fig. 1*A*). The residues P$_{36}$ and P$_{39}$ constitute the canonical PxxP motif-defining prolines and are part of the PPII helix that comprises residues P$_{34}$ to P$_{39}$ (Table S2). The XP dipeptide units of this motif (A$_{35}$P$_{36}$ and P$_{38}$P$_{39}$), bridged by T$_{37}$, occupy the conserved proline-binding pockets formed by F$_{348}$, Y$_{396}$ and P$_{393}$, W$_{378}$, respectively (Fig. 2*A*). The Ramachandran ϕ, ψ angles of P$_{34}$ (−71°, 153°) indicate that it is also part of the PPII helix, although its side chain does not contact with SH3. However, P$_{34}$ is likely to play a role in maintaining a suitable conformation for optimal accommodation of an A$_{35}$P$_{36}$ dipeptide into the hydrophobic slot sculpted by P$_{393}$ and W$_{378}$. This is supported by our peptide array data indicating that binding is significantly weakened if P$_{34}$ is changed to Ala (Fig. 3*A*).

W$_{33}$ of the R47$_5$ peptide occupies the position of Arg/Lys in the class I peptides (RxxPxxP) recognized by Src-type SH3 domains via an electrostatic interaction with a conserved negatively charged residue in their specificity pocket (Fig. S5). Although Trp is hydrophobic and IRTKS SH3 has a leucine (L$_{358}$) instead of conserved D/E in the specificity pocket, the complementarity between the ligand and IRTKS SH3 at this position is poor. Thus, unlike the corresponding residue (position −3) of canonical class I ligands, W$_{33}$ of R47$_5$ does not directly contribute to IRTKS binding, and inhibitory effects of residue substitutions at this position are likely to reflect its role in forming an appropriate linker between the two PPII helical regions of R47$_5$. Overall, the C-terminal PxxP-accommodating binding pocket in IRTKS is highly hydrophobic, which is further dictated by atypical hydrophobic Leu at position 358 in addition to the conserved hydrophobic residues at positions 348, 378, 393, and 396 (Fig. 2*A*).



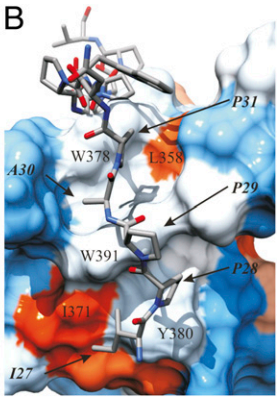


Fig. 2. Close-ups of C-terminal and N-terminal $R47_5$ binding sites on IRTKS SH3 as a hydrophobicity surface presentation. (*A*) C-terminal binding region. Residues forming the conserved proline-binding pockets are labeled on the surface. Heavy atoms of the C-terminal part of $R47_5$ residues $N_{32}APTPP_{39}$ are highlighted as a stick model: nitrogen (blue), oxygen (red), and carbon (gray). Surface coloring is according to the scale of Kyte and Doolittle (37). Blue corresponds to the most hydrophilic, white to intermediate, and red to the most hydrophobic. (*B*) N-terminal PPII helix of $EspF_U$ $R47_5$. The two hydrophobic clefts formed by L_{358}, W_{378}, W_{391}, Y_{380}, and I_{371} and the heavy atom of $R47_5$ residues $I_{27}PPAP_{31}$ occupying the clefts are shown.

The most unique feature of the IRTKS:$EspF_U$ complex is the presence of two bona fide PxxP motifs in $R47_5$, both contributing to this interaction. In addition to the previously described canonical $P_{36}xxP_{39}$ motif, the $EspF_U$ residues P_{28} and P_{31} define a PxxP motif that is part of an N-terminal PPII helix in $R47_5$ involving residues 27–31 (Fig. 2*B*). Intriguingly, this additional PxxP motif is accommodated by two hydrophobic slots in the extended specificity pocket of IRTKS SH3 in a manner that is virtually identical to the interaction of the canonical PxxP motif with the conserved SH3 proline-binding pockets of IRTKS.

P_{29} of $R47_5$ acts as the bridging residue between two XP dipeptide units formed by residues $I_{27}P_{28}$ and $A_{30}P_{31}$, which occupy the two hydrophobic pockets in the specificity region of IRTKS (Fig. 2*B*). The backbone carbonyl of P_{29} forms a hydrogen bond with the indole proton of W_{391}. The $A_{30}P_{31}$ dipeptide-binding pocket involving L_{358}, W_{378}, and W_{391} corresponds to the classic negatively charged specificity pocket, whereas W_{391}, together with Y_{380} and I_{371}, participates in forming a fourth hydrophobic cavity in IRTKS SH3, which perfectly coheres with the $I_{27}P_{28}$ moiety (Fig. 2*B*). Hence, the N-terminal binding groove in IRTKS SH3, which accommodates another PxxP motif, harbors several hydrophobic residues that render the specificity-determining region highly hydrophobic. In addition to the conserved W_{378} and W_{391}, the N-terminal binding region is occupied by uncommon L_{358}, I_{371}, and Y_{380}. Comparison of different SH3 domain sequences indicates that Y_{380} and I_{371} are very unusual residues at these positions and clearly provide unique specificity for IRTKS SH3:$EspF_U$ $R47_5$ recognition. This conclusion agrees well with our peptide array data. Deletion of the N-terminal dipeptide or replacement of either I_{27} or P_{28} with Ala drastically hindered binding to $EspF_U$ (Fig. 3 *A* and *B*).

Thus, the overall binding epitope can be considered as a double-PxxP motif, wherein the flanking N- and C-terminal XPxXP motifs are linked by the tripeptide $N_{32}WP_{34}$. The role of this linker was further investigated by peptide array mutagenesis studies, which indicated that although divergent tripeptide combinations could be used to replace it, there were strict steric requirements for the linker to coordinate the binding of the two PxxP motifs of $R47_5$ correctly (Fig. 3). Elongating or shortening the linker by duplication or deletion of either N_{32} or W_{33} residue seriously impaired binding. Ala was well tolerated in place of N_{32} but not at position 33 or 34 (Fig. 3*A*). Conversely, replacement of the whole linker with a trialanine sequence still allowed significant binding, whereas triglycine or triproline linkers did not. Finally, tripeptides RAP, TTS, and QSV, found in the corresponding position of reported cellular ligands of IRTKS, could be used to replace the NWP linker of $R47_5$ without loss of binding (Fig. 3*B*).

To examine the individual roles of the PxxP motifs of $R47_5$ further, we also carried out NMR titration experiments with N- and C-terminal halves of the IPPAPNWPAPTPP peptide. The individual binding of N-terminal EHIPPAPN and C-terminal PAPTPPVQ peptides to IRTKS SH3 was monitored using ^{15}N-HSQC spectra. In good agreement with our peptide array studies, binding was either very weak, with an estimated K_d in the millimolar range for EHIPPAPN, or completely abolished for PAPTPPVQ, indicating that both PxxP motifs are equally recognized and pivotal to the high-affinity binding. However, on addition of the EHIPPAPN peptide, the observed CSPs were determined in

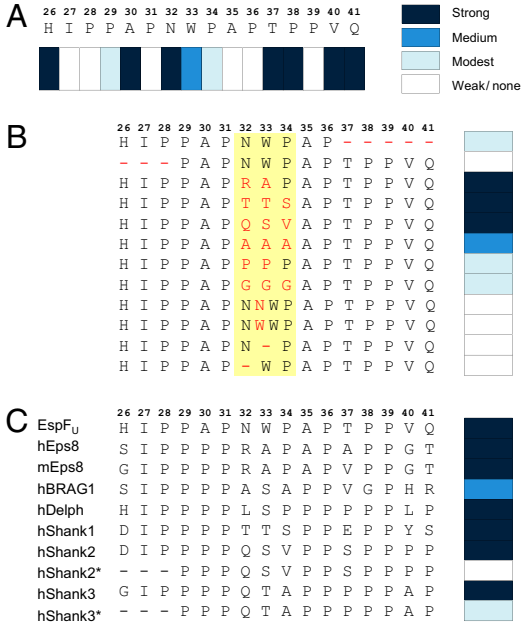


Fig. 3. Fine mapping of IRTKS SH3-binding preferences. Peptide array technology was used to define the residues critical for IRTKS SH3 binding in the $EspF_U$ $R47_5$ and in a panel of cellular proteins reported or suspected to interact with IRTKS/IRSp53. (*A*) Sixteen residues encompassing the region of $R47_5$ observed to contact IRTKS SH3 were systematically replaced with Ala. At positions naturally occupied by Ala, an uncharged polar amino acid (Asn) was used instead. Binding signals were quantified and classified into four categories. Strong indicates >85%, medium indicates 30–85%, modest indicates 5–29%, and weak/none indicates <5% of the average signal from triplicate dots printed with the corresponding unmodified peptide. (*B*) N- and C-terminal ends and the linker region (shaded in yellow) of 16-mer IRTKS SH3-binding peptide were altered as indicated (red font) and examined as above. (*C*) IRTKS binding of the 16-mer $EspF_U$ peptide was compared with that of similar peptides from cellular proteins. Human (h) and mouse (m) Eps8 protein sequence differ in this region and have both been included. N-terminal truncated versions of Shank2/3 peptides (*) were included because an earlier study suggested the minimal IRSp53 SH3-binding site to reside within this sequence (24).

MICROBIOLOGY

CHEMISTRY

W_{391}, Y_{380}, and I_{371}, indicating that the N-terminal fragment of the peptide could independently bind to the specificity pocket formed by these hydrophobic residues.

Binding Sites in Cellular Partners of IRTKS SH3. Based on our complex structure, we carried out a search (http://au.expasy.org/prosite/) for human proteins containing the amino acid string IPxZPxxxZPxZP (wherein Z is P, A, I, L, or V). Among the genes found in this search were the previously reported cellular partners of IRSp53/IRTKS SH3 domains, namely, Shank1–3, Eps8, and BRAG1 (21–23), as well as potential previously undescribed partners, such as Delphilin, a postsynaptic scaffolding protein involved in actin cytoskeleton regulation (22). Testing 16-mer peptide sequences from these proteins corresponding to the $EspF_U$ characterized by this approach confirmed strong binding to IRTKS in all cases, including the Delphilin peptide (Fig. 3*C*), suggesting that tandem PxxP motifs are functional in mammalian ligands and their recognition is evolutionary conserved.

Based on mapping of the IRSp53 SH3-binding sites in the proteins, a related but less well-defined consensus sequence, PpPxxxppxPP, has been proposed (24). We therefore also included shorter Shank2/3 peptides in our array that contained this motif but lacked three N-terminal residues, including the Ile and Pro residues critical for $EspF_U$ binding. These 13-mer Shank peptides showed dramatically reduced binding to IRTKS, supporting the idea that the complete IPpΦPxxxΦPxΦP consensus sequence and dual recognition of the tandem PxxP motifs are also critical for interaction with cellular partners of IRTKS/IRSp53 SH3.

Both PxxP Motifs Are Required for $EspF_U$ Recruitment and Pedestal Formation by EHEC. $EspF_U$ alleles vary in the number of C-terminal 47-residue repeats, but all contain at least two repeats (25). As noted above, each $EspF_U$ repeat harbors two functional elements, a helical portion that binds and activates WASP/N-WASP (15, 16) and a proline-rich region that binds to IRTKS SH3 (9, 10) (Fig. 1*C*). Although a single proline-rich sequence is capable of recognition by the SH3 domain, two proline-rich sequences flanking a helical region, which we term PHP^{WT}, are required for recruitment to sites of bacterial attachment and actin pedestal formation in mammalian cells. To investigate the relationship between SH3 recognition and $EspF_U$ recruitment to sites of bacterial attachment, myc-tagged PHP^{WT} was expressed in KC12, an *E. coli* derivative that expresses EHEC Tir and is capable of type III translocation. When fibroblast-like cells (FLCs) were infected with this strain, PHP^{WT} was readily recruited to sites of bacterial attachment and promoted robust pedestal formation (Fig. 4, second row). In contrast, a variant of $EspF_U$ consisting of a single repeat (HP) was neither recruited nor mediated in pedestal formation (Fig. 4, first row).

To test whether mutations in the tandem PxxP motifs that interact with IRTKS SH3 affect $EspF_U$ function in vivo, we constructed mutant alleles of $EspF_U$ PHP carrying Ala substitutions in either PxxP motif (referred to as $PHP^{P28/31A}$ and $PHP^{P36/39A}$, respectively). Immunoblotting revealed that these mutants were expressed in *E. coli* KC12 at levels equivalent to HP and PHP^{WT}. Strikingly, neither $PHP^{P28/31A}$ nor $PHP^{P36/39A}$ localized to sites of KC12 attachment or promoted the formation of pedestals (Fig. 4, third and fourth rows). These data indicate that both PxxP motifs in the $EspF_U$ proline-rich region are required for recruitment by IRTKS and subsequent localized actin assembly.

Discussion

We have determined the structure of the human IRTKS SH3 domain in complex with the fifth 47-residue repeat of $EspF_U$. The most important finding of our study is that IRTKS contains an SH3 domain that establishes a high-affinity complex, which is

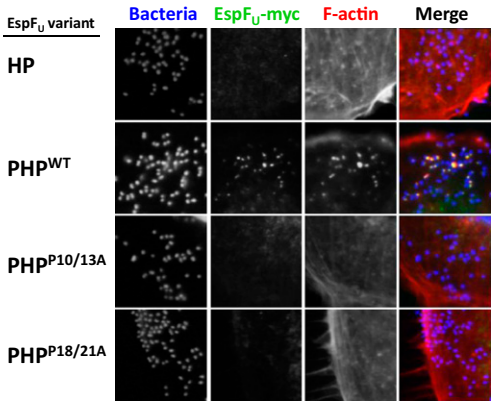


Fig. 4. Tandem PxxP motifs are required for $EspF_U$ recruitment and pedestal formation by EHEC. FLCs were infected for 3.5 h with strain EPEC KC12 expressing the $EspF_U$ HP ($R47_5$), PHP, $PHP^{P28/31A}$, and $PHP^{P36/39A}$ variants and were visualized using fluorescent microscopy after staining with DAPI to localize bacteria (blue), FITC–anti-myc antibody was used to detect $EspF_U$ derivatives (green), and Alexa568-phalloidin was used to detect F-actin (red).

mediated solely by hydrophobic interactions using two consecutive PxxP motifs (i.e., without any polar contacts to the specificity pocket). IRTKS SH3 harbors atypical hydrophobic residues in the n-Src loop and β3 to β4 strands as well as in its specificity pocket. Compared with a negatively charged residue typically found in this position of the SH3 specificity pocket, the L_{358} residue of IRTKS creates a repulsive force toward typical SH3 ligands. The highly hydrophobic binding interface of IRTKS SH3 is unusually large, containing four hydrophobic slots, with each capable of accommodating an XP dipeptide moiety. Consequently, two adjacent PxxP motifs in $EspF_U$ $R47_5$ establish two conjugated PPII helices spanning from I_{27} to P_{31} and via P_{34} to P_{39} (Table S2). The interaction between IRTKS and $EspF_U$ $R47_5$ is one of the strongest observed in natural ligands bound to an SH3 domain, and the tandem PxxP motif is essential for the ability of EHEC to localize $EspF_U$ beneath bound bacteria and trigger the formation of an actin pedestal.

An epitope of 13 residues in $R47_5$ was found to be required for the high-affinity binding. Our NMR and peptide array data indicated the amino acid sequence IPxΦPxxxΦPxΦP as the optimal consensus target site for IRTKS. This consensus motif can be found in the previously identified cellular ligand for IRTKS/IRSp53 SH3 domains and could be used to identify potential unique binding partners, such as Delphilin.

Specific peptide recognition by IRTKS SH3 relies on the presence and correct positioning of the two consecutive PxxP motifs. Removal of an XP moiety from either the N or C terminus drastically impairs binding, indicating that the whole 13-aa epitope is indispensable (Fig. 3*B*). The molecular basis of proline recognition is analogous at both PPII helix docking sites. In a left-handed PPII helix conformation, the bond between N and Cδ of the critical PxxP prolines points toward the hydrophobic surface of the SH3 domain. Any other natural amino acid would have a backbone NH proton at this position, resulting in repulsive interaction, and N-substitution of proline is thus required to prevent this unfavorable interaction (26). The high proline content of the tandem PxxP peptide is also likely to contribute to the affinity of binding through constrained conformational flexibility, resulting in reduced entropic cost on binding.

Despite several examples of atypical SH3 interactions (27–29) the majority of SH3 domains rely on idiotypic recognition of PPII helical PxxP motifs in their partner selection. The binding energy in such Src-like SH3 interactions is governed by hydrophobic, nonspecific van der Waals forces and by an electrostatic

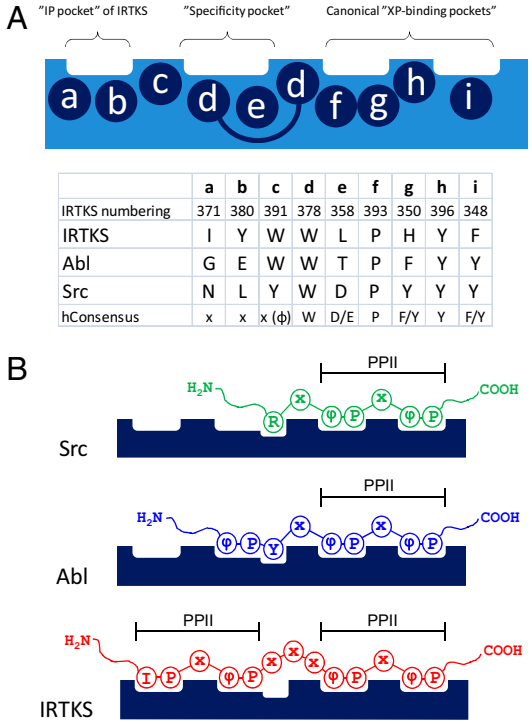
interaction between positively charged R/K of a peptide and negatively charged D/E in the specificity pocket of SH3 (Fig. 5).

A relevant SH3 domain to be discussed here is that of the Abl tyrosine kinase. Early studies on Abl provided paradigm-forming examples of structural basis and specificity in SH3 binding (4, 30). However, it later became evident that ligand binding by Abl is anomalous rather than a typical SH3 interaction. Like IRTKS/IRSp53, Abl SH3 lacks the highly conserved D/E residue at position 358 (numbering according to IRTKS in Fig. 1*B*) and has Thr instead (Fig. 5*A*). Accordingly, Abl SH3 does not bind to typical class I or II peptides; instead, it prefers ligands containing a class I-like ΦxxPxxP consensus sequence, wherein Φ is a hydrophobic residue, most often Tyr (Fig. 5*B*). The structure of Abl SH3 in complex with a high-affinity (K_d = 1.5 μM) ligand APSYSPPPPP shows that the Y_4 residue of the peptide provided significant affinity for the interaction via hydrogen bonds with side chains of Ser and Asp residues in the RT loop of Abl SH3 (31). IRTKS has Gly and Asn at the corresponding positions, indicating that IRTKS cannot engage in a similar hydrogen-bonding network with ligands selected by Abl SH3. Indeed, W_{33} of $R47_5$, which occupies a position analogous to Y_4 of Abl ligands, is part of the linker connecting the two PxxP motifs in $R47_5$ and has no direct role in IRTKS SH3 binding. Comparison of IRTKS SH3-bound $EspF_U$ $R47_5$ with the Abl SH3-bound APSYSPPPPP ligand and a canonical RxxPxxP class I ligand is illustrated in Fig. S5.

Apart from this difference at IRTKS position 358, the rest of the specificity pocket of Abl is formed by two Trp residues in a manner that is very similar to W_{378} and W_{391} in IRTKS (Fig. 5*A*). In this regard, it is of interest that the first consensus-binding motif reported for Abl SH3 was PxxxxPxxP (30). Notably, the first Pro of such consensus peptides is accommodated by the specificity pocket of Abl SH3 analogously to the interaction of P_{31} of $R47_5$ by W_{378} and W_{391} of IRTKS SH3. A similar interaction can be seen in an unrelated SH3/ligand interaction, namely, a typical intermediate affinity (7.4 μM) complex of βPIX SH3 with a class I ligand from atropin-interacting protein 4 (AIP4), which also involves a proline-directed contact between βPIX SH3 and an extended PPII helical scaffold in the AIP4 peptide (32). However, in contrast to IRTKS:$R47_5$, the βPIX:AIP4 complex involves a canonical polar interaction between an arginine residue of the AIP4 peptide and negatively charged residue in the specificity pocket of βPIX SH3, which accounts for much of the binding affinity.

Despite these similarities, the mode of ligand binding in which canonical recognition of tandem PxxP motifs gives rise to high affinity of binding in the absence of any polar contacts with the ligand peptide is unique. This unique binding specificity of IRTKS SH3 can be attributed most prominently to the presence of the very unusual hydrophobic residues I_{371} and Y_{380} found in the n-Src loop and strand β3, respectively (Figs. 1 and 5*A*). In Abl and βPIX, as well as in most other SH3 domains, these sites are occupied by nonhydrophobic residues. As is evident from our structural and mutational data, this "IP" pocket plays a pivotal role in specificity, because the recognition of the dipeptide is highly exclusive. Conversely, some SH3 domains have occupied positions 371 and 380 with hydrophobic residues but are lacking a hydrophobic residue at position 358. Indeed, among the ≈300 *Homo sapiens* SH3 domains, IRTKS/IRSp53 are the only ones that harbor hydrophobic residues in the critical positions F_{348}, L_{358}, I_{371}, W_{378}, Y_{380}, W_{391}, P_{393}, and Y_{396} (Fig. 5*A*).

SH3 domains have addressed the promiscuity of binding in a cellular context by using non-PxxP motifs, additional residues flanking the PxxP region, or extended binding surfaces to enhance the affinity and specificity of diverse interactions (31, 33, 34). A comparison of IRTKS SH3-bound $EspF_U$ $R47_5$ with three other SH3 complexes involving unusually high binding affinity is shown in Fig. S5. Our studies highlight the evolutionary diversity of the SH3 specificity-determining region. In contrast to the relatively homotypic recognition of PPII helical proline-rich peptides by the conserved PxxP motif-binding surface of SH3 domains, the region contributed in part to the n-Src- and RT loops has specialized in accommodating a wide range of structures linked to PxxP core peptides, thus providing selectivity for SH3 ligand binding. The recognition of a second PPII helical PxxP motif by the specificity-determining region observed in the IRTKS SH3:$EspF_U$ complex is an extreme example of such adaptation, which plays a decisive role in mediating host-pathogen interaction, resulting in seizure of host's actin assembly machinery.



	a	b	c	d	e	f	g	h	i
IRTKS numbering	371	380	391	378	358	393	350	396	348
IRTKS	I	Y	W	W	L	P	H	Y	F
Abl	G	E	W	W	T	P	F	Y	Y
Src	N	L	Y	W	D	P	Y	Y	Y
hConsensus	x	x	x (ϕ)	W	D/E	P	F/Y	Y	F/Y

Fig. 5. Schematic presentation of class I peptide binding to Src-, Abl-, and IRTKS-type SH3 domains. (*A*) Amino acids (a to i) forming the individual peptide-binding pockets of SH3 domains (numbering according to IRTKS SH3 in Fig. 1). (*B*) Peptide recognition is similar for all SH3 domains at the canonical "XP-binding pockets," wherein each peptide adopts a left-handed PPII helical conformation. Selectivity is determined at the "specificity pocket," wherein peptide conformation and peptide/SH3 domain interactions are different in each case. In IRTKS SH3, the specificity and "IP pockets" also accommodate PPII helical moiety.

Materials and Methods

Cloning, Expression, and Purification of IRTKS SH3 and $EspF_U$. A detailed description of cloning, expression, and purification of IRTKS SH3 and $EspF_U$ $R47_5$ and construction of plasmids expressing $EspF_U$ derivatives is given in *SI Materials and Methods*.

Mammalian Cell Infection. Mouse embryo FLCs were infected with strain KC12 expressing $EspF_U$ derivatives as previously described (11). Cells were visualized after staining with DAPI to reveal bacteria, FITC–anti-myc antibody (Invitrogen) to detect $EspF_U$-myc fusions, and Alexa568-phalloidin to detect F-actin.

ITC. ITC experiments were performed at 25 °C using a VP-ITC microcalorimeter (Microcal, Inc.). Gel-filtrated IRTKS SH3 and $R47_5$ were both added to the NMR buffer. $R47_5$ (0.2 mM) was titrated into IRTKS SH3 (10 μM). The experiment was repeated twice. To measure heats of dilution, control experiments were performed by titrating $R47_5$ to buffer and subtracting the results from raw titration data. The thermodynamic profile of the IRTKS SH3 and $R47_5$ interaction was obtained by nonlinear least-square fitting of ex-

perimental data using a single-site binding model of Origin 7 software (Microcal, GE Healthcare).

NMR Spectroscopy, Structure Calculations, and Analysis. All NMR experiments were carried out at 25 °C on a Varian INOVA 800- or 600-MHz spectrometer, equipped with a $^{15}N/^{13}C/^{1}H$ triple-resonance cold probe and a z-axis gradient system. NMR spectra for structure determination of IRTKS SH3:$EspF_U$ $R47_5$ complex were recorded on two samples: 0.48-mM uniformly ^{15}N, ^{13}C-labeled IRTKS SH3 in 1:1 complex with unlabeled $EspF_U$ $R47_5$ and 0.90-mM uniformly ^{15}N, ^{13}C-labeled $EspF_U$ $R47_5$ in 1:1 complex with unlabeled IRTKS SH3. A set of triple-resonance experiments [e.g., 3D iHNCACB, CBCA(CO)NH, HBHA(CO)NH, CC(CO)NH, HCC(CO)NH, HCCH-COSY] was used for the assignment of backbone and aliphatic side chain resonances of ^{15}N,^{13}C-labeled IRTKS SH3 (17, 18). Aromatic resonances were assigned using (HB)CB(CGCD)HD, (HB)CB(CGCDCE)HE, and aromatic HCCH-COSY experiments (17). For the assignment of ^{15}N, ^{13}C-labeled $EspF_U$ $R47_5$, a similar set of experiments as in the case of IRTKS SH3 was used, but the assignment of main chain resonances was supplemented with a set of HA-detected experiments (19).

For the structure calculation of SH3:$R47_5$ complex, the $R47_5$ sequence was connected to the C terminus of the SH3 domain sequence through a set of weightless noninteracting dummy atoms. NOE peaks were picked manually from ^{15}N- and ^{13}C-edited NOESY spectra recorded with 100 ms of mixing time. The peak lists, together with the chemical shift assignments, were used as input for the iterative NOE assignment and structure calculations in Cyana (35). We generated 200 conformers in each of the seven cycles of the combined automated NOESY and structure calculation algorithm. A more comprehensive description of sample conditions, structure calculation procedure, and structure validation is given in *SI Materials and Methods*.

Accession Codes. The resonance assignments and coordinates of the IRTKS SH3:$EspF_U$ $R47_5$ complex structure have been deposited in the BioMagResBank (accession number 16909) and Protein Data Bank (PDB ID code 2kxc), respectively.

Peptide Array. Peptides of interest were synthesized by the Peptide and Protein Laboratory (Haartman Institute, University of Helsinki) as peptide-cellulose conjugates using Multipep (INTAVIS Bioanalytical Instruments) according to the manufacturer's protocol for SPOT synthesis (36) and printed in parallel arrays on glass slides using SlideSpotter equipment (INTAVIS Bioanalytical Instruments). Experimental details are provided in *SI Materials and Methods*.

Construction of Plasmids Expressing $EspF_U$ Derivatives. The C-terminal proline-rich repeats of $EspF_U$ were deleted from plasmid pKC471 (13) by inverse PCR using primers flanked by KpnI and BamHI restriction sites, creating plasmid pDV48. Plasmid pDV48 is designed to receive DNA inserts between sequences encoding the N-terminal translocation sequence of $EspF_U$ and a C-terminal 5-myc tag, and thus allows for expression of proteins that can be translocated into mammalian cells by the EHEC/enteropathogenic *Escherichia coli* (EPEC) type III secretion system. $EspF_U$ fifth repeat ($R47_5$, also referred to as HP) was amplified by PCR and cloned into pDV48 to create plasmid pCL1. $EspF_U$ PHP WT and mutant fragments were synthesized (Integrated DNA Technologies) and cloned into pDV48, creating plasmids pCL2 (PHP^{WT}), pDV168 ($PHP^{P28/31A}$), and pDV169 ($PHP^{P36/39A}$). All constructs were transformed into EPEC strain KC12, and their expression level was analyzed by Western blot using an anti-myc 9E10 monoclonal antibody (Santa Cruz Biotechnology).

ACKNOWLEDGMENTS. We thank Elina Ahovuo for excellent technical assistance, Cindy Lai (University of Massachusetts Medical School, Worcester, MA) for plasmid pCL1, Brian Skehan for communication of unpublished results, and Jussi Hepojoki for help and advice with peptide arrays. Financial support for this study was provided by Academy of Finland Grants 122170 and 131144 (to P.P.) and 127197 (to K.S.), National Institutes of Health Grant R01AI46454 (to J.M.L.), a Sigrid Juselius Foundation grant (to K.S.), and Research Council of Helsinki University Hospital Grant TYH2008307 (to K.S.).

1. Mayer BJ (2001) SH3 domains: Complexity in moderation. *J Cell Sci* 114:1253–1263.
2. Kaneko T, Li L, Li SS (2008) The SH3 domain—A family of versatile peptide- and protein-recognition module. *Front Biosci* 13:4938–4952.
3. Yu H, et al. (1994) Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* 76:933–945.
4. 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.
5. Feng S, Chen JK, Yu H, Simon JA, Schreiber SL (1994) Two binding orientations for peptides to the Src SH3 domain: Development of a general model for SH3-ligand interactions. *Science* 266:1241–1247.
6. Lim WA, Richards FM, Fox RO (1994) Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* 372:375–379.
7. Aitio O, et al. (2008) Structural basis of PxxDY motif recognition in SH3 binding. *J Mol Biol* 382:167–178.
8. Scita G, Confalonieri S, Lappalainen P, Suetsugu S (2008) IRSp53: Crossing the road of membrane and actin dynamics in the formation of membrane protrusions. *Trends Cell Biol* 18:52–60.
9. Weiss SM, et al. (2009) IRSp53 links the enterohemorrhagic E. coli effector Tir and $EspF_U$ for actin pedestal formation. *Cell Host Microbe* 19:244–258.
10. Vingadassalom D, et al. (2009) Insulin receptor tyrosine kinase substrate links the E. coli O157:H7 actin assembly effectors Tir and EspF(U) during pedestal formation. *Proc Natl Acad Sci USA* 106:6754–6759.
11. Hayward RD, Leong JM, Koronakis V, Campellone KG (2006) Exploiting pathogenic Escherichia coli to model transmembrane receptor signalling. *Nat Rev Microbiol* 4: 358–370.
12. Frankel G, Phillips AD (2008) Attaching effacing Escherichia coli and paradigms of Tir-triggered actin polymerization: Getting off the pedestal. *Cell Microbiol* 10:549–556.
13. Campellone KG, Robbins D, Leong JM (2004) EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly. *Dev Cell* 7:217–228.
14. Garmendia J, et al. (2004) TccP is an enterohaemorrhagic Escherichia coli O157:H7 type III effector protein that couples Tir to the actin-cytoskeleton. *Cell Microbiol* 6: 1167–1183.
15. Cheng HC, Skehan BM, Campellone KG, Leong JM, Rosen MK (2008) Structural mechanism of WASP activation by the enterohaemorrhagic E. coli effector EspF(U). *Nature* 454:1009–1013.
16. Sallee NA, et al. (2008) The pathogen protein EspF(U) hijacks actin polymerization using mimicry and multivalency. *Nature* 454:1005–1008.
17. Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Magn Reson Spectrosc* 34:93–158.
18. Tossavainen H, Permi P (2004) Optimized pathway selection in intraresidual triple-resonance experiments. *J Magn Reson* 170:244–251.
19. Mäntylahti S, Aitio O, Hellman M, Permi P (2010) HA-detected experiments for the backbone assignment of intrinsically disordered proteins. *J Biomol NMR* 47:171–181.
20. Dyson HJ, Wright PE (2001) Nuclear magnetic resonance methods for elucidation of structure and dynamics in disordered states. *Methods Enzymol* 339:258–270.
21. Funato Y, et al. (2004) IRSp53/Eps8 complex is important for positive regulation of Rac and cancer cell motility/invasiveness. *Cancer Res* 64:5237–5244.
22. Miyagi Y, et al. (2002) Delphilin: A novel PDZ and formin homology domain-containing protein that synaptically colocalizes and interacts with glutamate receptor delta 2 subunit. *J Neurosci* 22:803–814.
23. Sanda M, et al. (2009) The postsynaptic density protein, IQ-ArfGEF/BRAG1, can interact with IRSp53 through its proline-rich sequence. *Brain Res* 1251:7–15.
24. Bockmann J, Kreutz MR, Gundelfinger ED, Böckers TM (2002) ProSAP/Shank postsynaptic density proteins interact with insulin receptor tyrosine kinase substrate IRSp53. *J Neurochem* 83:1013–1017.
25. Garmendia J, et al. (2005) Distribution of tccP in clinical enterohemorrhagic and enteropathogenic Escherichia coli isolates. *J Clin Microbiol* 43:5715–5720.
26. Nguyen JT, Turck CW, Cohen FE, Zuckermann RN, Lim WA (1998) Exploiting the basis of proline recognition by SH3 and WW domains: Design of N-substituted inhibitors. *Science* 282:2088–2092.
27. Mongioví AM, et al. (1999) A novel peptide-SH3 interaction. *EMBO J* 18:5300–5309.
28. Kami K, Takeya R, Sumimoto H, Kohda D (2002) Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67(phox), Grb2 and Pex13p. *EMBO J* 21: 4268–4276.
29. Hoelz A, et al. (2006) Crystal structure of the SH3 domain of βPIX in complex with a high affinity peptide from PAK2. *J Mol Biol* 358:509–522.
30. Ren R, Mayer BJ, Cicchetti P, Baltimore D (1993) Identification of a ten-amino acid proline-rich SH3 binding site. *Science* 259:1157–1161.
31. Pisabarro MT, Serrano L, Wilmanns M (1998) Crystal structure of the abl-SH3 domain complexed with a designed high-affinity peptide ligand: Implications for SH3-ligand interactions. *J Mol Biol* 281:513–521.
32. Janz JM, Sakmar TP, Min KC (2007) A novel interaction between atrophin-interacting protein 4 and beta-p21-activated kinase-interactive exchange factor is mediated by an SH3 domain. *J Biol Chem* 282:28893–28903.
33. Harkiolaki M, et al. (2003) Structural basis for SH3 domain-mediated high-affinity binding between Mona/Gads and SLP-76. *EMBO J* 22:2571–2582.
34. Ghose R, Shekhtman A, Goger MJ, Ji H, Cowburn D (2001) A novel, specific interaction involving the Csk SH3 domain and its natural ligand. *Nat Struct Biol* 8:998–1004.
35. Herrmann T, Güntert P, Wüthrich K (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* 319:209–227.
36. Frank R (2002) The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports—Principles and applications. *J Immunol Methods* 267:13–26.
37. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157:105–132.