Molecular basis for the binding of SH3 ligands with non-peptide elements identified by combinatorial synthesis

Sibo Feng, Tarun M Kapoor, Fumiyuki Shirai, Andrew P Combs and Stuart L Schreiber

Background: Protein-structure-based combinatorial chemistry has recently been used to discover several ligands containing non-peptide binding elements to the Src SH3 domain. The encoded library used has the form Cap-M1-M2-M3-PLPPLP, in which the Cap and Mi's are composed of a diverse set of organic monomers. The PLPPLP portion provided a structural bias directing the non-peptide fragment Cap-M1-M2-M3 to the SH3 specificity pocket. Fifteen ligands were selected from >1.1 million distinct compounds. The structural basis for selection was unknown.

Results: The solution structures of the Src SH3 domain complexed with two ligands containing non-peptide elements selected from the library were determined by multidimensional NMR spectroscopy. The non-peptide moieties of the ligands interact with the specificity pocket of Src SH3 domain differently from peptides complexed with SH3 domains. Structural information about the ligands was used to design various homologs, whose affinities for the SH3 domain were measured. The results provide a structural basis for understanding the selection of a few optimal ligands from a large library.

Conclusions: The cycle of protein-structure-based combinatorial chemistry followed by structure determination of the few highest affinity ligands provides a powerful new tool for the field of molecular recognition.

Address: Howard Hughes Medical Institute, Department of Chemistry and Chemical Biology, Harvard University, MA 02138, USA.

Correspondence: Stuart L Schreiber

Introduction

Intracellular signaling depends critically on protein–protein interactions. The discovery of small molecule ligands that can bind to a specific target protein and block its interactions with other proteins would enable the study and control of such signaling processes. We have selected the Src homology 3 (SH3) domain as a test system to develop methods for ligand discovery.

SH3 domains are among the few small conserved protein modules that mediate the formation of protein complexes *in vivo* and participate in a diverse array of cellular events (for recent reviews, see [1,2]). The structural basis for interactions between peptide ligands and SH3 domains is now well understood, in part due to the combined use of combinatorial chemistry and structural biology [3–9]. The two classes (classes I and II) of peptide ligands bind with opposite amide backbone directionality [5]. They all have a PXXP core that anchors the ligand to the receptor in a polyproline type II (PPII) helical conformation and a nearby flanking sequence that provides binding specificity. Although the core sequences of ligands to different SH3 domains are conserved, the flanking sequences are unique to individual SH3 domains and bind to a specificity-determining pocket [8,10].

The synthesis and evaluation of libraries of millions of compounds prepared by combinatorial chemistry provide

a powerful means for ligand discovery. Structural information derived from X-ray and NMR studies can guide the selection of monomers and biasing elements to construct combinatorial libraries targeted for a specific receptor. We recently used this structure-based combinatorial chemistry strategy to identify ligands with non-peptide elements binding to the specificity pocket of the Src tyrosine kinase SH3 domain [11]. The library has the form Cap-M1-M2-M3-PLPPLP, in which the Cap and Mi's are composed of a diverse set of organic monomers. The additional PLPPLP, serving as the PXXP core, provided a structural bias that was expected to direct the non-peptide fragment Cap-M1-M2-M3 to the SH3 specificity pocket on the basis of previously determined structures of SH3–peptide complexes [8]. Using multidimensional NMR spectroscopy (Table 1), we have now determined the solution structures of the Src SH3 domain complexed with two such ligands, NL1 (dissociation constant (K_d) = 3.4 μM) and NL2 $(K_d = 11\ \mu$M) (NL stands for non-peptide ligand), discovered from the previously described library (Tables 2a, 3). As anticipated, both ligands bind in the class I orientation, and the non-peptide elements are located in the specificity pocket between two flexible loops connecting the β strands in the SH3 domain (named the RT and n-Src loops). Comparing these structures with those of the peptide ligands selected from phage-display

Table 1

Structural and energetic statistics for the complexes of NL1–Src SH3 and NL2–Src SH3.

Parameter		NL1–Src SH3			NL2–Src SH3	
		SA$_i$	‹SA›$_{ref}$		SA$_i$	‹SA›$_{ref}$
		RMS deviations from experimental distance restraints (Å)				
All	(966)	0.020 ± 0.002	0.012	(879)	0.020 ± 0.002	0.019
Interproton distances of Src SH3	(643)	0.021 ± 0.003	0.018	(642)	0.020 ± 0.002	0.019
Hydrogen-bond restraints*	(57)	0.016 ± 0.005	0.016	(57)	0.014 ± 0.003	0.015
Interproton distances of NL1	(80)	0.005 ± 0.003	0.005			
Interproton distances of NL2				(42)	0.003 ± 0.003	0.001
Intermolecular ligand–SH3 interproton distances	(186)	0.020 ± 0.003	0.021	(137)	0.020 ± 0.002	0.020
		RMS dihedral deviations (degrees)				
All	(72)	0.140 ± 0.045	0.149	(73)	0.156 ± 0.074	0.127
		Deviations from idealized geometry[†]				
Bonds (Å)	(1072)	0.003 ± 0.000	0.003	(1051)	0.003 ± 0.000	0.003
Angles (degrees)	(1938)	0.753 ± 0.018	0.741	(1901)	0.758 ± 0.008	0.748
Impropers (degrees)	(590)	0.430 ± 0.019	0.421	(579)	0.419 ± 0.013	0.403
		Energetic statistics (kcal mol^{-1})				
E_{repel}[‡]		2.6 ± 0.8	2.1		2.6 ± 0.7	2.0
E_{vdW}[§]		−143.0 ± 19.2	−143.7		−134.3 ± 24.1	−153.9
E_{NOE}[¶]		18.8 ± 4.9	9.5		16.8 ± 2.7	15.1

The SA$_i$ columns give the average and standard deviations for the indicated variables obtained from the 23 final refined simulated annealing structures. ‹SA›$_{ref}$ represents the restrained minimized mean structure obtained from individual SA structures best fitted to each other. The numbers of various restraints are given in parentheses.
*All hydrogen-bond restraints were from slowly exchanging amide protons except when noted in Materials and methods.
†Idealized geometries based on CHARMM 19 parameters [26].

‡Repulsive potential. The value of the quartic repulsive force constant used in the structure calculations was 1 kcal mol^{-1} Å^{-4}.
§Lennard–Jones van der Waals energy calculated with standard CHARMM 19 parameters.
¶Total distance restraint energy. The values of the square-well NOE and torsion angle force constants were 50 kcal mol^{-1}Å^{-2} and 200 kcal mol^{-1} rad^{-2}, respectively.

libraries shows that NL1 and NL2 use a common set of binding interactions not present in the peptide–SH3 complexes. We also report the study of various NL1 and NL2 homologs to understand the different contributions to the ligand–SH3 complexation. The results demonstrate that the new ligands would have been impossible to design *a priori* based on known peptidic ligands, providing a dramatic illustration of the power of combinatorial chemistry in ligand discovery.
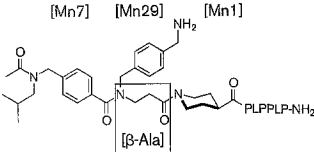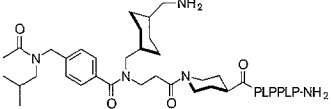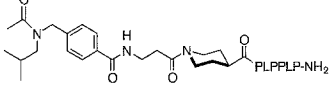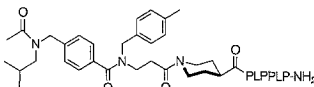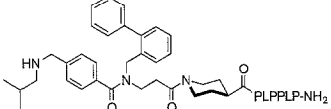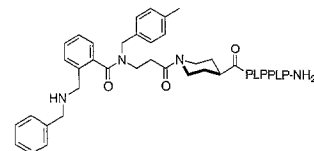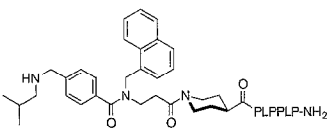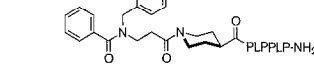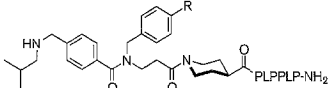
Results

Structure the of Src SH3–NL1 complex

The average root-mean-square deviations (rmsds) of the 23 calculated structures of the Src SH3–NL1 complex versus their mean coordinate are 0.55 Å for the backbone and 0.97 Å for the heavy atoms, respectively (Fig. 1). There are a total of 186 intermolecular nuclear Overhauser

effects (NOEs). The heavy atom rmsd for the non-peptide Mn7-Mn29-Mn1 portion is 0.84 Å (Mn#'s = monomers identified in the library screen), and is 0.66 Å for the entire ligand. As noted befo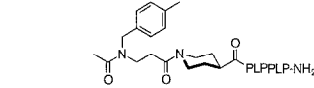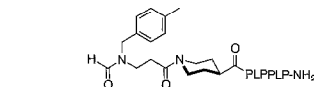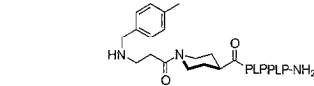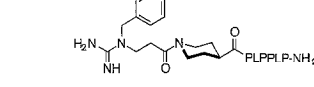re with peptide ligands, ligand binding results in few structural changes in the SH3 domain. The protein structure is essentially the same as the free SH3 domain and the SH3 domain complexed with VSL12, a class I peptide ligand VSLARRPLPPLP selected from a phage-display library (K$_d$ for Src SH3 = 0.45 μM) [8,10]. The accessible surface area on the SH3 domain buried by NL1 is 615 Å^2, comparable to the 622 Å^2 buried by VSL12 (calculated in GRASP [12] using a probe of radius 1.4 Å). The binding mode of the PLPPLP portion of NL1 is the same as the PLPPLP sequence in VSL12. It adopts a left-handed PPII helix conformation, and the two Leu–Pro dipeptides occupy pockets 1 and 2, respectively.

Table 2

(a) Binding affinity of NL1-related ligands: variations on Mn7.		(b) Binding affinity of NL1 -related ligands: variations on Mn29.	
Ligand	K_d, μM^*	Ligand	K_d, μM^*
[Mn7] [Mn29] [Mn1] ... [β-Ala] NL1	3.4 / 32 (to Src Y131A SH3 mutant) / 170 (to PI3K SH3 domain)	... NL1-9	12
... NL1-2	6.2	... NL1-10	80
... NL1-3	5.0	... NL1-11	6.4
... NL1-4	11	... NL1-12	8.6†
... NL1-5	15	R = H NL1-13	5.4
... NL1-6	21	CH₃ NL1-14	4.0
... NL1-7	44†	OMe NL1-15	11
... NL1-8	10	CO₂⁻ NL1-16	5.5
		CO₂Me NL1-17	7.9
		CF₃ NL1-18	6.4
		F NL1-19	7.7
		NO₂ NL1-20	12†
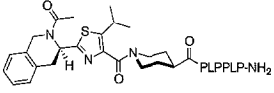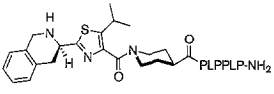
*All K_d values were measured by fluorescence perturbation assay against the wild type Src SH3 domain at pH 7.4 except when noted.
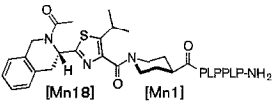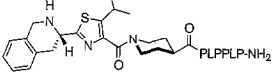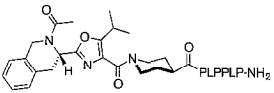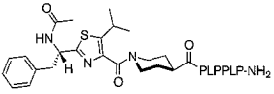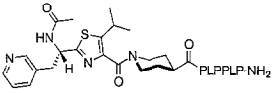†The K_d was measured by fluorescence polarization assay.

The non-peptide part of the ligand is located in the specificity pocket bordered by W118, Y131, and the n-Src and RT loops (Fig. 1). The carbonyl group of Mn1 makes a hydrogen bond to the indolic NH of W118 of the SH3 domain; this hydrogen bond is conserved in the peptide–SH3 complex between Arg6 of VSL12 and W118. The NOE and coupling constant patterns indicate that the piperidine ring predominantly adopts a chair conformation, although some conformational averaging exists as indicated by the linewidth broadening and intensity attenuation for its proton resonances. Part of the piperidine ring leans on top of W118, as indicated by the NOEs between W118 and protons on Mn1. The β-alanine portion of Mn29 (Table 2a) adopts the lowest energy *anti* conformation with an average dihedral angle of 164° for

the backbone, and packs against the sidechain of W118. Because of chemical shift degeneracy in the proton dimension, the assignment of the four protons on β alanine was facilitated by isotope-editing NMR experiments using a partially ¹³C-labeled sample (Fig. 2). The aminomethyl benzyl moiety of Mn29 lies on the floor of the pocket and makes an intermolecular aromatic–aromatic contact with Y131, adopting a geometry not commonly observed in protein structures [13]. For the average minimized coordinate, the two phenyl ring centroids are separated by 3.6 Å, and the dihedral angle between the two rings is 30°. In addition to the NOEs to Y131, the aromatic protons of Mn29 have NOEs to E115 on the n-Src loop and to T98 on the RT loop. An edge-to-face 'herringbone' aromatic–aromatic packing most often

Table 3

Binding affinity of NL2-related ligands.

Ligand		K_d (μM)*
[Mn18] [Mn1] (structure, PLPPLP-NH$_2$)	NL2	11
(structure, PLPPLP-NH$_2$)	NL2-2	30
(structure, PLPPLP-NH$_2$)	NL2-3	62
(structure, PLPPLP-NH$_2$)	NL2-4	5.4 7.5 (pH = 5.9)
(structure, PLPPLP-NH$_2$)	NL2-5	6.2 3.3 (pH = 5.9)
(structure, PLPPLP-NH$_2$)	NL2R	5.4
(structure, PLPPLP-NH$_2$)	NL2R-2	33

*All K_d values were measured by fluorescence perturbation assay against the wild type Src SH3 domain at pH 7.4, except when noted.

observed in proteins (mean centroid separation 5.1 Å and mean dihedral angle 77°) [13] does not satisfy the NOE restraints. The sidechain of E115 has NOEs to the methylene of the aminomethyl group on Mn29, suggesting a polar interaction between the Mn29 amino group and the E115 carboxylate may exist. After descending to the bottom of the pocket, the ligand makes an abrupt turn at Mn7, which lies in a hydrophobic groove between Mn1 and the namesake R95 and T96 of the RT loop. The Mn7 phenyl ring packs directly against the γMe of T96, whose chemical shift changes from 1.19 ppm to 0.25 ppm upon binding. The isobutyl group of Mn7 makes hydrophobic contacts with the Arg95 sidechain.

Except for the amide bond between the acetyl cap and Mn7, all other amide bonds in the ligand adopt a single configuration in the bound state, as defined by the corresponding NOEs. The two amide rotamers resulting from rotation around the Ac-Mn7 amide bond exist in about a 1:1 ratio. The two sets of NMR resonances from the rotamers exhibit the same NOE patterns except for the acetyl methyl, thus the rotation around this bond has little effect on the overall binding of the ligand.

Structure of the Src SH3–NL2 complex

The average rmsds of the 23 calculated structures of the Src SH3–NL2 complex versus their mean coordinate are 0.67 Å for the backbone and 1.13 Å for the heavy atoms, respectively (Fig. 3). There are a total of 137 intermolecular NOEs. The heavy atom rmsd for the non-peptide Mn18-Mn1 portion is 0.22 Å, and is 0.39 Å for the entire ligand. The protein structure is essentially the same as the free SH3 domain. The accessible surface area on the SH3 domain buried by NL2 is 482 Å^2, substantially smaller than the 615 Å^2 buried by NL1, and thus correlates with their relative affinities (NL1 K_d = 3.4 μM; NL2 K_d = 11 μM). The PLPPLP portion of NL2 binds to the receptor as the corresponding region from NL1 does, and it directs the Ac-Mn18-Mn1 fragment into the specificity pocket.

The carbonyl of Mn1 in NL2 makes a hydrogen bond to the indolic NH of W118 similar to the one between NL1 and W118, and its piperidine ring also adopts a chair conformation. However, the relative orientation of the piperidine rings in NL1 and NL2 differs by almost 90°. The ring in NL1 binds the receptor with a 'face-on' geometry, and protons on both carbons β to the carbonyl group of Mn1 have NOEs to W118. In contrast, the piperidine in NL2 adopts an 'edge-on' geometry, and only protons on one edge of the ring show NOEs to W118. The thiazolyl group of Mn18 serves as a rigid linker that has minimal contact with the protein. The rotational freedom of the thiazole is largely restricted by its tendency to avoid 1,3-allylic strain with Mn1. The tetrahydroisoquinoline (Tic) bicyclic moiety of Mn18 binds to a region between Y131 and the RT loop as defined by its NOEs to Mn1 and to T96, T98 and Y131. Interestingly, comparison with NL1 shows that the phenyl ring of Tic is located approximately at the mid point of the two phenyls belonging to Mn29 and Mn7 in NL1 (Fig. 4b,c).

Similar to NL1, all the amide bonds in the bound NL2, except the acetyl-Mn18 amide, adopt a single configuration. The ratio of *trans* to *cis* amide rotamers in this group is about 3:1. The *trans* rotamer should bind the protein better than the *cis* rotamer, since only the *trans* acetyl can make hydrophobic contact with Y131. Based on the NOE patterns, the rotation around this amide bond does not affect the binding mode of the rest of the ligand. Only the *trans* configuration is used in generating the structures.

Binding of NL1-related ligands

In addition to the various truncated ligands reported earlier [11], we synthesized two groups of new variants of NL1 and evaluated their binding to Src SH3. The first group contains ligands with various functional groups replacing Mn7 (Table 2a). Relative to ligands NL1 and NL1-2, replacement of the isobutyl with benzyl (ligand NL1-3) leads to little change in affinity. However, substituting

Figure 1

The structure of the Src SH3–NL1 complex.
(a) Stereoview of the superposition of 23
final refined structures of the Src SH3–NL1
complex. The backbone traces of the SH3
domain are blue, and the sidechains of the
SH3 residues at the binding sites are red.
The non-peptide portion of NL1 is yellow and
the PLPPLP fragment is purple. **(b)** The
molecular surface of Src SH3 with NL1. The
three binding pockets and the monomers in
the non-peptide fragment are labeled in red.
The two Leu-Pro dipeptides of PLPPLP bind
to pockets 1 and 2, respectively. Pocket 3
(the specificity pocket) accommodates the
non-peptide sequence Ac-Mn7-Mn29-Mn1.
(c) Binding site of the Src SH3–NL1
complex. The Cα trace of the protein is red.
The sidechains for the SH3 domain (labeled
in yellow) are colored yellow for carbon, red
for oxygen, and blue for nitrogen. The ligand
is colored white for carbon, red for oxygen,
and blue for nitrogen. The amino and carboxyl
termini of the ligand are indicated by N and C
in white, respectively. Both (b) and (c) were
drawn using the program GRASP [12].
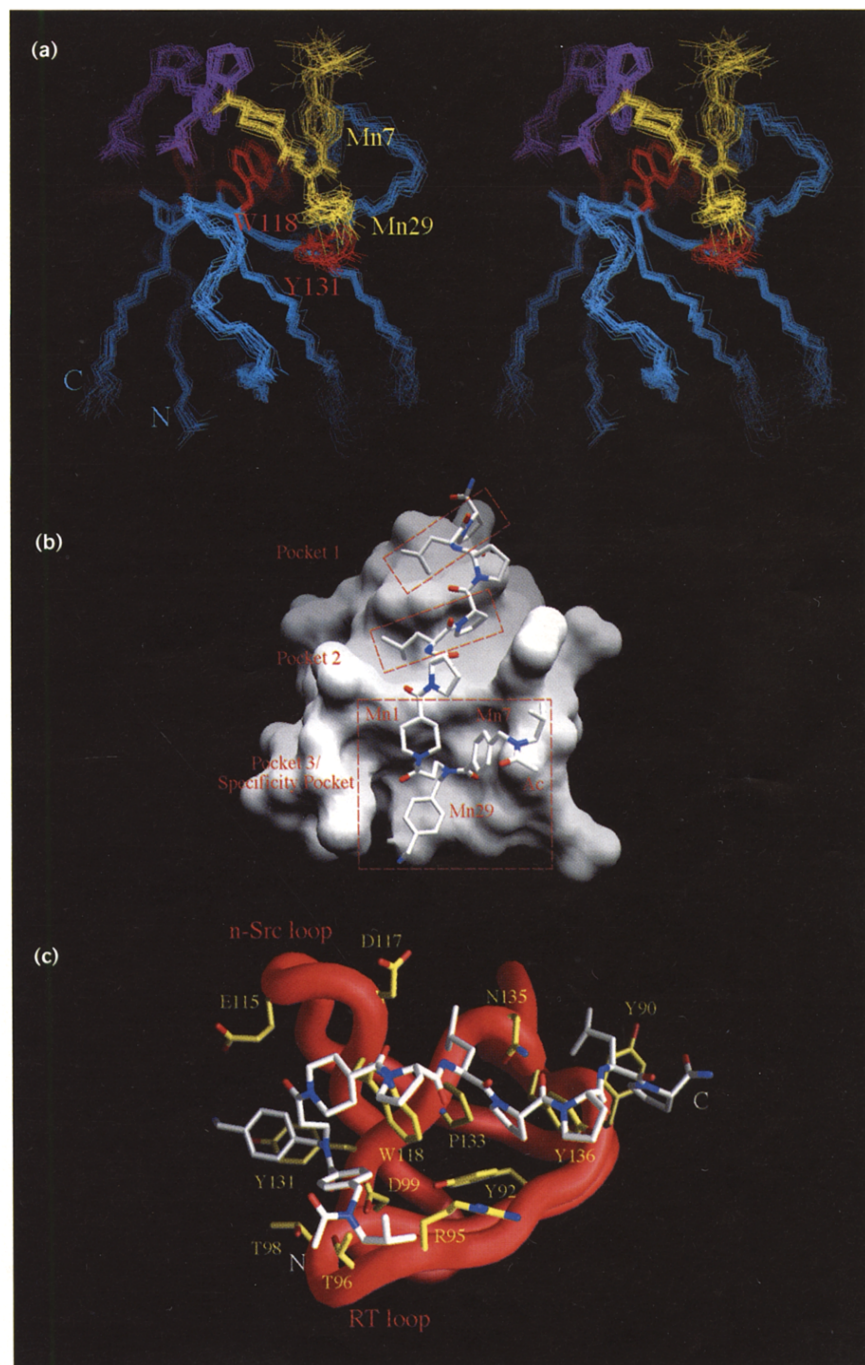
Mn7 with smaller hydrophobic groups including benzoyl,
acetyl, formyl and hydrogen results in successive decreases
in binding affinity (ligands NL1-4 to NL1-7). On the other
hand, replacing the formyl with a guanidinium group
(ligand NL1-8) results in a two-fold increase in affinity,
possibly through a guanidinium–D99 electrostatic interaction.

The second group of ligands contain compounds with Mn29
variants (Table 2b). The importance of the aromatic–aromatic

interaction between Mn29 and Y131 is demonstrated by
the observation that replacing the aminomethyl benzyl
with aminomethyl cyclohexyl group leads to a 4-fold drop
in affinity (ligand NL1-9), and that NL1 binds to the
Y131A mutant of Src SH3 10-fold more weakly than does
the wild-type protein (Table 2a). Removing the amino-
methyl benzyl of Mn29 (ligand NL1-10) leads to a 23-fold
drop in affinity. On the other hand, enlarging the amino
benzyl moiety of Mn29 to a biphenyl (NL1-11) or a

Figure 2
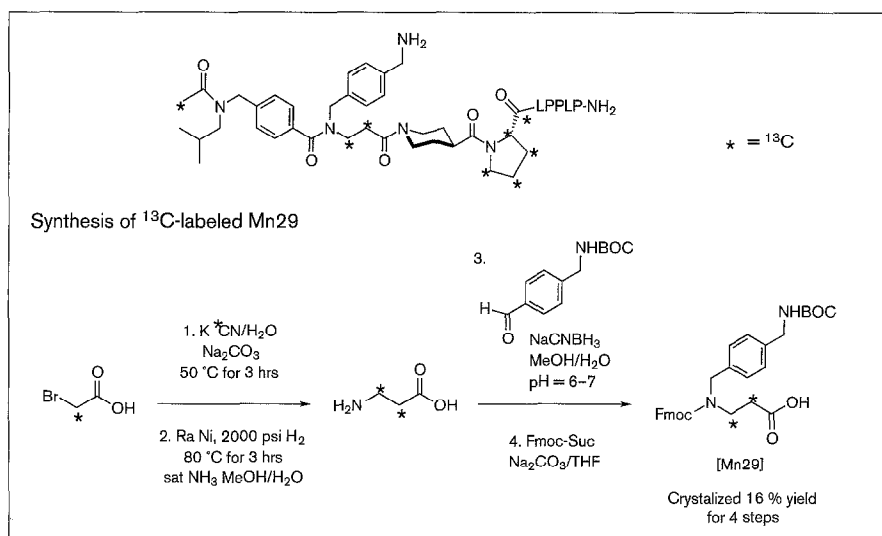
Synthesis of partially ^{13}C-labeled NL1 for NMR structure determination.

---

naphthyl (NL1-12), to increase the hydrophobicity, does not improve the K_d; the bulky substitutions probably give rise to unfavorable steric interactions with the nearby W118 and the n-Src loop.

To investigate the nature of the π–π interaction further, ligands NL1-13 to NL1-20 were synthesized, which have various *para* substituents on the phenyl of Mn29, including the electron-donating groups –Me, –OMe and the electron-withdrawing groups $-CO_2^-$, $-CO_2Me$, $-CF_3$, –F and $-NO_2$. The results suggest that varying the electronic properties of Mn29 has only a small effect on binding affinity.

Binding of NL2-related ligands

Replacing the thiazolyl in Mn18 by the homologous oxazolyl group (Table 3, ligand NL2-3) results in a substantial decrease in affinity, suggesting that even a subtle alteration of this rigid linker is unfavorable. For the oxazolyl group, the C–O bond length is 1.37 Å and the C–O–C angle is 104°, whereas for the thiazole the C–S bond length is 1.74 Å and the C–S–C angle is 89°. The smaller oxazole probably makes it difficult for the Tic group to make favorable contacts with the RT loop.

Disconnecting the cyclic structure in the Tic moiety leads to a ligand (NL2-4) with good affinity. Since the Tic aromatic ring is near D99, we synthesized a pyridyl-containing ligand (NL2-5) to see whether the basicity of the pyridyl nitrogen could improve the binding through an electrostatic interaction with D99. The pKa of free pyridinium is 5.2. At pH 7.4, no affinity improvement is observed presumably because a minimal amount of the pyridyl is protonated. At pH 5.9, the ratio of protonated to unprotonated pyridyl increases by 32-fold relative to pH 7.4, and we did observe a small (2-fold) increase in

affinity. Unfortunately, the pI of the protein is 4.9, and we cannot measure the K_d at significantly lower pH without denaturing the protein. As a control, the homologous phenyl-containing ligand (NL2-4) did not show an increase in affinity at pH 5.9 versus pH 7.4. Assuming that the pKa of the pyridinium moiety is ≥ 5.2 (due to the nearby D99 carboxylate), extrapolation of the pyridyl ligand data to a pH at which all the pyridyl would be protonated gives a $K_d \geq 1$ μM, representing a gain of ≤ 1 kcal mol^{-1} free energy relative to the unprotonated state. It is possible that a hydrogen bond forms between the pyridinium nitrogen in NL2-5 and the carboxylate of D99 at low pH, but this needs to be confirmed by further structural studies.

A surprising result is that when the chirality of Mn18 is inverted from (*S*) to (*R*), the resulting ligand (NL2R) binds as well as does NL2. Removing the acetyl group of either NL2 or NL2R (ligands NL2-2 and NL2R-2, respectively) has a similar effect in decreasing the affinity, probably as a consequence of reduction in hydrophobicity of the ligands. NMR data show that both NL2 and NL2R are optically pure, and no racemization occurred during the synthesis as detected by NMR. Studies are under way to determine how NL2R interacts with the SH3 domain.

Discussion

**Comparison of peptide and non-peptide ligands**

Despite the divergence in their flanking sequences, the binding modes for the PLPPLP cores in NL1, NL2 and VSL12 are essentially the same, as evidenced by the fact that the inter- and intra-molecular NOE patterns are the same. Thus, the PLPPLP core directs the neighboring sequences of both the peptide and non-peptide elements to the specificity pocket.

Figure 3

Structures of the Src SH3–NL2 complex. **(a)** Stereoview of the superposition of 23 final refined structures of the Src SH3–NL2 complex. The backbone traces of the SH3 domain are blue, and the sidechains of the SH3 residues at the binding sites are red. The non-peptide portion of NL2 is yellow and the PLPPLP fragment is purple. **(b)** The molecular surface of Src SH3 with NL2 shown. The three binding pockets and the monomers in the non-peptide fragment are labeled red. The two Leu-Pro dipeptides of PLPPLP bind to pockets 1 and 2, respectively. Pocket 3 (the specificity pocket) accommodates the non-peptide sequence Ac-Mn18-Mn1. **(c)** Binding site of the Src SH3–NL2 complex. The Cα trace of the protein is red. The sidechains for the SH3 domain (labeled in yellow) are colored yellow for carbon, red for oxygen, and blue for nitrogen. The ligand is colored white for carbon, red for oxygen, blue for nitrogen and green for sulfur. The amino and carboxyl termini of the ligand are indicated by N and C, respectively.
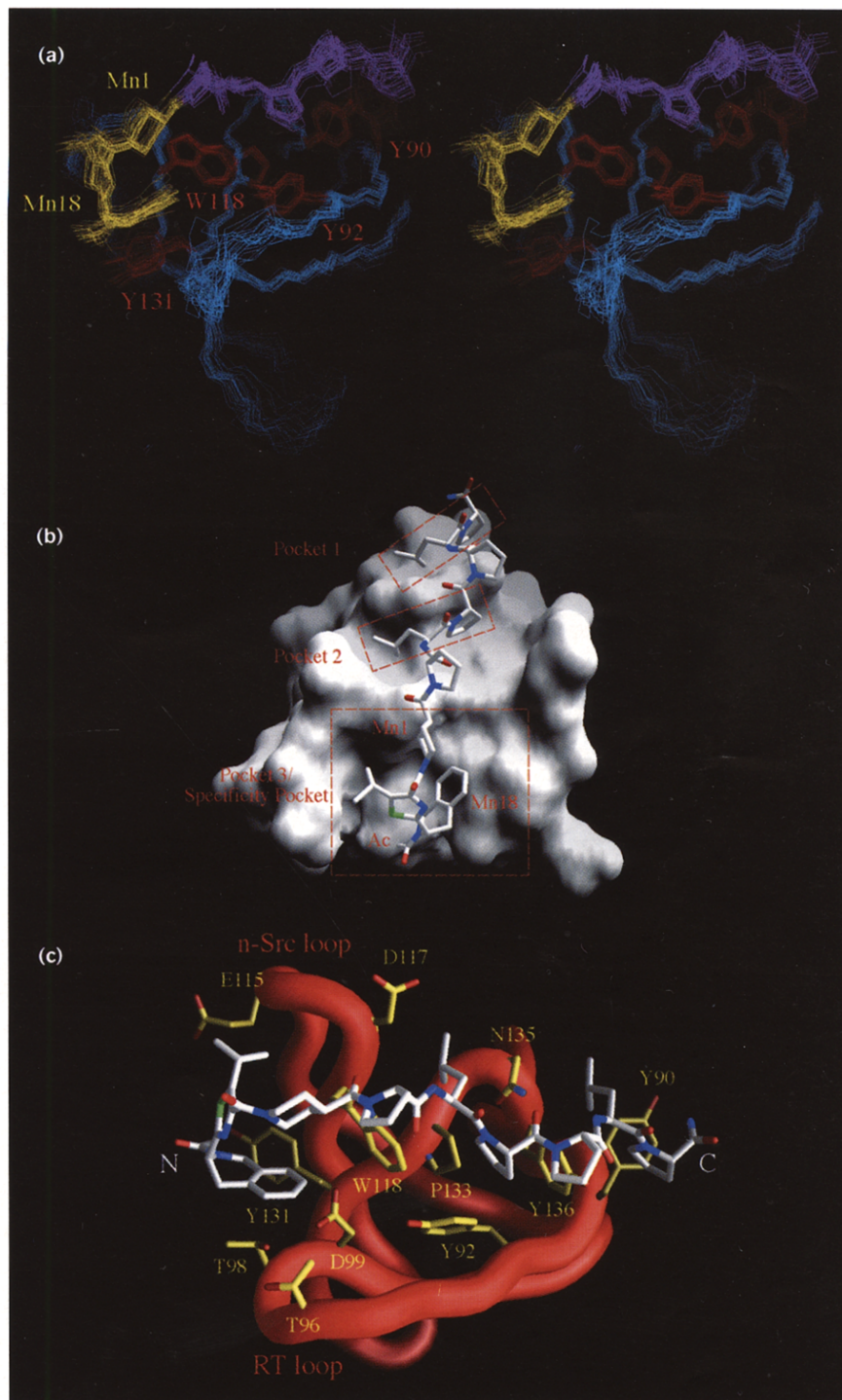
There are, however, significant differences among the binding modes of VSL12, NL1 and NL2 in the specificity pocket (Fig. 4a–c). No aromatic interactions are observed in the peptide ligand VSL12. The peptide fragment VSLARR of VSL12 makes a salt bridge to D99 of Src SH3 that contributes about 2.1 kcal mol^{-1} free energy of binding. The majority of the hydrophobic interactions in VSL12 involve W118, Y131 and the n-Src loop, whereas for the RT loop only a small area of hydrophobic contact exists between Leu3 of VSL12 and T98 of Src SH3 [8]. In contrast, no salt bridge to D99 is present with NL1. In addition to the Mn29–Y131 π–π packing, other intermolecular interactions are primarily hydrophobic in nature. Whereas only limited contact between

Figure 4
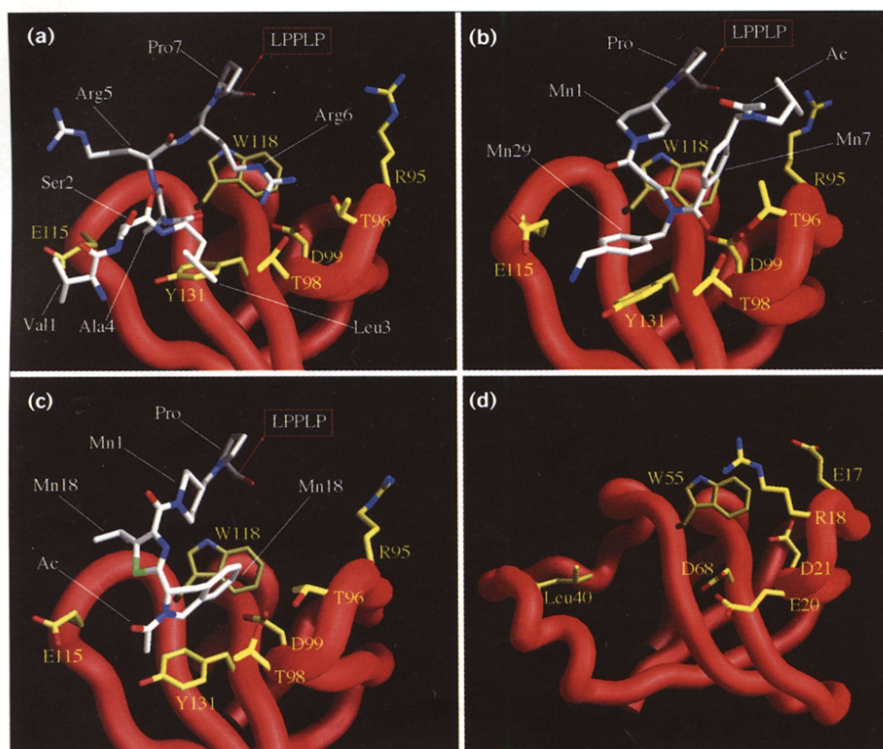
Comparison of peptide and non-peptide fragments binding to the specificity pockets of SH3 domains. The Cα trace of the SH3 domain is shown as a red worm. The SH3 residues are labeled in yellow, and the ligand residues in white. The position where the LPPLP portion in the ligand is attached is indicated. **(a)** Peptide ligand VSL12 at the specificity pocket of Src SH3. The ligand sequence Val-Ser-Leu-Ala-Arg-Arg-Pro is shown. **(b)** Ligand NL1 at the specificity pocket of Src SH3. The sequence Ac-Mn7-Mn29-Mn1-Pro is shown. **(c)** Ligand NL2 at the specificity pocket of Src SH3. The sequence Ac-Mn18-Mn1-Pro is shown. **(d)** Specificity pocket of PI3K SH3. Except for the conserved Trp55 (Trp118 in Src), the other residues in Src SH3 that have hydrophobic interactions with NL1 are substituted by hydrophilic charged residues in PI3K (see Discussion).

Mn29 of NL1 and the n-Src loop (E115) of Src SH3 is observed, a large area of hydrophobic contact is present between Mn7 and Mn29 of NL1 and Y131 and the RT loop, including residues R95, T96 and T98. The binding affinity decreases substantially as Mn7 or Mn29 are replaced with groups of reduced hydrophobicity (Table 2a,b). In the specificity pocket, NL2 also makes hydrophobic contacts with the RT loop and Y131, with virtually no contact to the n-Src loop. Such dramatic differences in the interactions between VSL12 and NL1 and NL2 with the receptor suggest that the non-peptide fragments would have been impossible to envision by traditional structure-based modeling.

The intermolecular aromatic–aromatic interaction

Although π–π interactions are common in proteins, the face-to-face geometry between the phenyl rings of Mn29 and Y131 is not commonly observed in protein structures. In the average minimized structure, the dihedral angle between the two phenyl rings is 30° and the distance between the two centroids is 3.6 Å. In contrast, the majority of the aromatic–aromatic pairs in proteins adopt the characteristic edge-to-face 'herringbone' packing, with a mean dihedral angle of 77° and mean centroid separation of 5.1 Å [13]. Both types of packing geometry are observed in small organic molecules (e.g., [14,15]).

This packing between Mn29 and Y131 is important for NL1 binding. The binding geometry predicts an attractive

aromatic–aromatic interaction based on theoretical calculations [14]. From the K_d data of NL1 binding to the wild-type Src SH3 and its Y131A mutant (Table 2a), we estimate that the interaction is worth 1.3 kcal mol^{-1} free energy of association, consistent with theoretical data that a typical aromatic–aromatic interaction has an energy of between –1 and –2 kcal mol^{-1} [13]. The π–π interaction has primarily two components, electrostatic and van der Waals forces. The interaction between the cyclohexyl group in NL1-9 (Table 2b) and Y131 in Src SH3 primarily involves van der Waals contact and lacks the electrostatic component; NL1-9 binding is four-fold lower in affinity compared to NL1.

We have shown that varying the electronic property of *para* substituents on the Mn29 phenyl ring has in general only a small effect on K_d (Table 2b). The *para* substituent on Mn29 has no direct contact with Y131. This result is consistent with the proposal that it is the properties of atoms at the points of intermolecular contact rather than the overall redox properties of the molecules that determine the aromatic–aromatic interaction [14].

Structural basis of ligand specificity and biological implications

Ligand NL1 binds to the SH3 domain of Src 52-fold more strongly than to the SH3 domain of the phosphatidyl-inositol 3-kinase (PI3K) (Table 2a) [11]. This selectivity can be explained by examining the structures of the two SH3 domains, which have different specificity pockets

[16]. In this pocket, many of the residues in Src SH3 that have substantial hydrophobic interactions with NL1, including R95, T96, T98 and Y131, are replaced by E17, R18, E20, and D68 in PI3K, respectively (Fig. 4d). In comparison, the peptide ligand VSL12 binds only three-fold better to Src SH3 than to PI3K SH3; the single largest contribution to binding, a salt bridge between R6 of VSL12 and D99 of Src SH3, is conserved in VSL12 binding to PI3K SH3 [8].

The binding of NL1 and NL2 to the RT loop region around the namesake R95 and T96 through hydrophobic interactions is particularly intriguing. A crystal structure shows that this region is involved in the binding of the phosphorylated tail peptide to the Src family kinase Lck's regulatory domains, SH2 and SH3 [17]. Mutating R95 to Trp diminishes the ability of Src to be regulated by phosphorylation of Y527 at its carboxy-terminal tail [18]. In addition, the oncogenic form of viral Src kinase has a Trp and an Ile at these positions, respectively. On the other hand, it was recently shown that the HIV-1 Nef protein binds to the SH3 domains of a subset of Src family kinases including Hck and Lyn (but not Fyn or Lck) [19]. This interaction is necessary for optimal spread of HIV-1 infection in primary cell cultures. An Ile at the RT loop unique to Hck and Lyn (corresponding to R95 in Src or R96 in Fyn) is a crucial specificity-determinant for binding to Nef, and an R96I mutant of Fyn SH3 can bind to Nef with an affinity comparable to that of Hck SH3 [20]. The crystal structure of Nef complexed with the R96I Fyn SH3 mutant shows that, in addition to the contacts between a contiguous proline-rich sequence of Nef to the SH3 domain, a tertiary hydrophobic interaction is observed between Nef and the Ile on the RT loop — a residue that has no direct contact with any known SH3-binding peptide [21].

NL1 and NL2 are the first contiguous SH3 ligands that can interact with this part of the RT loop on the SH3 domain. Therefore, through their unique interactions, the non-peptide ligands may be used as specific molecular probes to study the activation of Src kinases. They may also be used as inhibitors to block the Nef–SH3 association more effectively than their peptidic counterparts.

Conclusion

The studies of the ligands NL1 and NL2 reported herein reveal the molecular basis of the ligand–receptor recognition. The results suggest that extending the contact with the n-Src loop region and introducing electrostatic interaction with D99 may be important to improve the ligand affinity and specificity further. The synthesis of new libraries incorporating this and other features is in progress. Moreover, new libraries using Mn7-Mn29-Mn1 or Mn18-Mn1 as the biasing elements can now be constructed to direct the non-amino-acid building blocks to the Leu–Pro binding pockets. Because of the practical limitation on the

library size and diversity, using structural information to bias the library design is likely to be productive. Iterative cycles of library synthesis and structure determination may lead to a solution to the ligand–receptor binding problem in general.

Significance

The present study provides the first structural insights into the non-peptide binding elements discovered by a structure-based library design strategy. When compared to peptide ligands bound to the same Src SH3 receptor, the non-peptide ligands are found to interact with the receptor in a novel way. Peptide elements tend to contact the n-Src loop of the SH3 specificity pocket, whereas non-peptide elements tend to contact the RT loop of the same pocket. Both types of ligands make critical contacts with the tyrosine-based floor of the pocket. The combined information provides a clear direction for future ligand discovery involving ligands that contact both loops. Comparing the two types of structures makes it clear that the non-peptide binding elements could not have been designed by traditional modeling techniques, and thus reveals the power of structure-based combinatorial chemistry to identify ligands that might not otherwise be found.

Materials and methods

Generation of wild-type and mutant SH3 domains
The expression and purification of the wild-type Src SH3 domain and the Y131A Src SH3 mutant have been described previously [8,22]. The ^{15}N-labeled and ^{13}C and ^{15}N doubly labeled Src SH3 proteins for NMR study were generated by growing transformed *E. coli* BL21 cells in M9 minimal medium with ^{15}NH$_4$Cl and [^{13}C]-glucose as the sole nitrogen and carbon sources, respectively. The c-Src SH3 contains residues 85–140, and the numbering system used in the text is that of full-length c-Src from chicken.

NMR methods and structure determination
The NMR samples contained 2–3 mM of a 1:1 complex of the c-Src SH3 domain and ligand, in a D$_2$O or 90% H$_2$O/10% D$_2$O buffer containing 50 mM potassium phosphate (pH 6.0) and 150 mM KCl. All spectra were recorded at 27 °C using a Bruker DMX500 spectrometer.

The assignment of the SH3 domain was accomplished by standard two- and three-dimensional ^{15}N/^{13}C edited experiments using a 1:1 mixture of labeled SH3 domain and unlabeled ligands as described previously [8]. The ^{1}H, ^{13}C and ^{15}N assignments of Src SH3 in the SH3–VSL12 and SH3–APP12 complexes [8] (S.F. and S.L.S., unpublished results) were particularly valuable, because many of the chemical shift changes upon the binding of NL1 and NL2 were localized around the binding site. Thus, the residues that are remote from the binding site could be assigned by direct comparison with the previous assignments.

The assignment of the ligand resonance was done by a series of two-dimensional correlation spectroscopy, total correlation spectroscopy, and nuclear Overhauser enhancement spectroscopy (NOESY) filter experiments in which the signals from the ^{15}N- and ^{13}C-attached protons of the SH3 domain were purged and only the signals from the unlabeled ligand were observed [23,24]. To overcome the resonance degeneracy, we also synthesized a partially ^{13}C-labeled NL1 in which the methyl of the acetyl cap, the β-Ala moiety of Mn29, and the first

proline of the PLPPLP were labeled with ^{13}C (Fig. 2). A ^{13}C-edited, three-dimensional NOESY and a ^{1}H–^{13}C heteronuclear multiple quantum coherence study of this partially labeled ligand complexed with an unlabeled SH3 domain gave the unambiguous assignment of the ^{13}C atoms and their attached protons.

The intermolecular NOEs between the SH3 molecule and ligand were obtained by two- and three-dimensional half-filtered NOESY experiments in which the ^{13}C and their attached protons on the SH3 are correlated with protons attached to ^{12}C on the ligand. Distance (NOE and hydrogen bond) and dihedral (ϕ and χ1) restraints derived from NMR experiments were used to calculate the structures. The mixing times used for NOESY experiments were 75 ms, 150 ms and 200 ms. The NOE restraints were grouped into four ranges: 1.8–2.7 Å, 1.8–3.3 Å (1.8–3.5 for NOEs involving NH protons), 1.8–5.0 Å and 1.8–6.0 Å for strong, medium, weak, and very weak NOEs, respectively. Hydrogen bond restraints ($1.5 \leq d_{NH-O} \leq 2.3$ Å; $2.4 \leq d_{N-O} \leq 3.3$ Å) were derived from slowly exchanging amide protons identified by two-dimensional ^{15}N–^{1}H heteronuclear single quantum coherence spectra recorded about 2 h after dissolving the lyophilized protein in a D$_2$O buffer. Additional hydrogen-bond restraints identified in the VSL12–SH3 complex involving W118 and Asn135 [8] were also used at the final stage of the calculation because the convergence of the initial structures defined by NOEs confirmed the presence of the hydrogen bonds. The values for backbone ϕ torsional angle restraints were restricted to $-120°\pm40°$ for large $^{3}J_{NH-C\alpha H}$ coupling constants and to $-55°\pm35°$ for small $^{3}J_{NH-C\alpha H}$ coupling constants. The sidechain χ1 angle restraints were restricted to a range of $\pm60°$. The three-dimensional structures were calculated from the experimental restraints using the program X-PLOR [25] with minor modifications. The structure files for the non-peptide elements were built in X-PLOR based on homologous structures in the Cambridge Structure Database and incorporated to the CHARMM force field [26]. All amide bonds were defined in the structure files, and were not allowed free rotation during the calculation. At the end of the simulated annealing procedure, an additional energy minimization was performed with the various energy terms turned on successively rather than simultaneously. The coordinates have been deposited in the Brookhaven Protein Data Bank.

Ligand synthesis and affinity measurement
Ligands were synthesized on 4-methyl benzhydrylamine resin (MBHA) (Nova Biochem) by first synthesizing the common core (Mn1-PLPPLP) using Applied Biosystems Peptide Synthesizer (Model 431A). The remaining sequences were completed by manual solid phase synthesis using standard FMOC chemistry. All ligands were cleaved from the supports using 5% H$_2$O/CF$_3$CO$_2$H (5 ml x 2 for 30 min), purified by high pressure liquid chromatography, and characterized by fast atom bombardment mass spectrometry. The affinities of the ligands were measured by fluorescence perturbation and fluorescence polarization assays at 20 °C as described previously [3,8]. The partially ^{13}C-labeled Mn29 was synthesized according to Figure 2.

Acknowledgements

References
1. Cohen, G.B., Ren, R. & Baltimore, D. (1995). Modular binding domains in signal transduction proteins. *Cell* **80**, 237–248.
2. Pawson, T. (1995). Protein modules and signalling networks. *Nature* **373**, 573–580.
3. Chen, J.K., Lane, W.S., Brauer, A.W., Tanaka, A. & Schreiber, S.L. (1993). Biased combinatorial libraries: novel ligands for the SH3 domain of phosphatidylinositol 3-kinase. *J. Am. Chem. Soc.* **115**, 12591–12592.
4. Yu, H., Chen, J.K., Feng, S., Dalgarno, D.C., Brauer, A.W. & Schreiber, S.L. (1994). Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* **76**, 933–945.
5. Feng, S., Chen, J.K., Yu, H., Simon, J.A. & Schreiber, S.L. (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, W.A., Richards, F.M. & Fox, R.O. (1994). Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* **372**, 375–379.
7. 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.
8. Feng, S., Kasahara, C., Rickles, R.J. & Schreiber, S.L. (1995). Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. *Proc. Natl. Acad. Sci. USA* **92**, 12408–12415.
9. Chen, J.K. & Schreiber, S.L. (1995). Combinatorial synthesis and multidimensional NMR spectroscopy: an approach to understanding protein–ligand interactions. *Angew. Chem. Int. Ed. Engl.* **34**, 953–969.
10. Rickles, R.J., Botfield, M.C., Zhou, X.-M., Henry, P.A., Brugge, J.S. & Zoller, M.J. (1995). Phage display selection of ligand residues important for SH3 domain binding specificity. *Proc. Natl. Acad. Sci. USA* **92**, 10909–10913.
11. Combs, A.P., Kapoor, T.M., Feng, S., Chen, J.K., Daude-Snow, L.F. & Schreiber, S.L. (1996). Protein structure-based combinatorial chemistry: discovery of non-peptide binding elements to Src SH3 domain. *J. Am. Chem. Soc.* **118**, 287–288.
12. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11**, 281–296.
13. Burley, S.K. & Petsko, G.A. (1985). Aromatic–aromatic interaction: a mechanism of protein structure stabilization. *Science* **229**, 23–28.
14. Hunter, C.A. & Sanders, J.K.M. (1990). The nature of π–π interactions. *J. Am. Chem. Soc.* **112**, 5525–5534.
15. Paliwal, S., Geib, S. & Wilcox, C.S. (1994). Molecular torsional balance for weak molecular recognition forces. Effect of 'Tilted-T' edge-to-face aromatic interactions on conformational selection and solid-state structure. *J. Am. Chem. Soc.* **116**, 4497–4498.
16. Koyama, S., Yu, H., Dalgarno, D.C., Shin, T.B., Zydowsky, L.D. & Schreiber, S.L. (1993). Structure of the PI3K SH3 domain and analysis of the SH3 family. *Cell* **72**, 945–952.
17. Eck, M.J., Atwell, S.K., Shoelson, S.E. & Harrison, S.C. (1994). Structure of the regulatory domains of the Src-family tyrosine kinase Lck. *Nature* **368**, 764–769.
18. Superti-Furga, G., Fumagalli, S., Koegl, M., Courtneidge, S.A. & Draetta, G. (1993). Csk inhibition of c-Src activity requires both the SH2 and the SH3 domains of Src. *EMBO J.* **12**, 2625–2634.
19. Saksela, K., Cheng, G. & Baltimore, D. (1995). Proline-rich (PXXP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef$^+$ viruses but not for downregulation of CD4. *EMBO J.* **14**, 484–491.
20. Lee, C.-H., Leung, B., Lemmon, M.A., Zheng, J., Cowburn, D., Kuriyan, J. & Saksela, K. (1995). A single amino acid in the SH3 domain of Hck determines its high affinity and specificity in binding to HIV-1 Nef protein. *EMBO J.* **14**, 5006–5015.
21. Lee, C.-H., Saksela, K., Mirza, U.A., Chait, B.T. & Kuriyan, J. (1996). Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. *Cell* **85**, 931–942.
22. Yu, H., Rosen, M.K., Shin, T.B., Seidel, D.C., Brugge, J.S. & Schreiber, S.L. (1992). Solution structure of the SH3 domain of Src and identification of its ligand-binding site. *Science* **258**, 1665–1668.
23. Petros, A.M., Kawai, M., Luly, J.R. & Fesik, S.W. (1992). Conformation of two non-immunosuppressive FK506 analogs when bound to FKBP by isotope-filtered NMR. *FEBS Lett.* **308**, 309–314.
24. Ikura, M. & Bax, A. (1992). Isotope-filtered 2D NMR of a protein-peptide complex: study of a skeletal muscle myosin light chain kinase fragment bound to calmodulin. *J. Am. Chem. Soc.* **114**, 2433–2440.
25. Brünger, A.T. (1992). *X-PLOR Version 3.0 Manual.* Yale University, New Haven, CT.
26. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217.