Rational Design of Specific High-Affinity Peptide Ligands for the Abl-SH3 Domain†

M. T. Pisabarro and L. Serrano*

European Molecular Biology Laboratory, Meyerhofstrasse 1, 69012 Heidelberg, Germany

Received January 29, 1996; Revised Manuscript Received May 6, 1996⊗

ABSTRACT: SH3 domains bind proline-rich peptides with affinities in the order of 0.2−50 μM. In general, these domains are quite promiscuous, and the same peptide can bind to several different SH3 domains with similar affinities (i.e., 3BP1 peptide to Abl- and Fyn-SH3). This poor affinity and specificity make it difficult to elucidate their role *in vivo* as well as the use of peptides to specifically bind to a single domain. Here, we report that by using existing biocomputing tools, as well as simple physicochemical reasoning, it is possible to design mutations in the 3BP1 peptide (Met4-Tyr, Pro5-Ser, and Leu8-Pro), so that the affinity for Abl-SH3 increases 20-fold (p40 peptide: APTYSPPPPP; K_d = 0.4 μM), while that for the closely related domain, Fyn-SH3, decreases 10-fold. Both the RT and n-Src loops are responsible for regulating the specificity for Pro-rich ligands and more specifically residues Ser15, Thr19, and Glu38 in Abl-SH3. The first six positions in the 3BP1 peptide are important for determining the specificity for SH3 domains, while the remaining four seem to be more important for the affinity. Moreover, by choosing rationally the substituents, it is possible to replace some of the Pro residues postulated to be essential for the interaction with SH3 domains and still have a significant affinity. This indicates that the sequence repertoire that could interact with a specific SH3 domain could be larger than previously thought.

The SH3[1] domains were first identified as homologous noncatalytic regions in cytoplasmic tyrosine kinases, phospholipase C, and the viral oncogene product Crk (Mayer et al., 1988). These domains have no fixed topological position within their host proteins, and they are found in a large variety of functionally different proteins (Musacchio et al., 1992a, 1994b; Pawson, 1992, 1993; Pawson & Gish, 1992; Schlessinger, 1994). SH3 are discrete protein modules which can be expressed, purified, and crystallized independently of the rest of the protein (Musacchio et al., 1992b).

The presence of SH3 domains in a wide variety of unrelated proteins, many of which involved in signal transduction, led to the speculation that this motif could play an important role in protein−protein interactions which are essential to cellular functions and/or subcellular localization, prompting many groups to search for the *in vivo* partners of SH3 domains (Pawson, 1995).

Initially, two proteins, called 3BP1 and 3BP2, were identified for their ability to interact with Abl-SH3. 3BP1 contains an homologous region to the GTPase-activating proteins of the small G-protein Rho while the function of 3BP2 remains unknown (Cicchetti et al., 1992). Mutagenic analysis of these proteins has identified short proline-rich sequences as binding sites for the Abl-SH3 domain. These regions share the consensus sequence XPXXPPPΨXP, where Ψ is a hydrophobic residue and X is any amino acid; prolines at positions 2, 7, and 10 have been shown to be essential for the binding to this domain (Ren et al., 1993). Peptides corresponding to these sequences (defined from now on as 3BP1 and 3BP2 peptides) bind to Abl-SH3 and Fyn-SH3 domains with similar affinities in the 5−40 μM range (Viguera et al., 1994). Two main approaches have been used to find putative SH3 ligands: (i) Screening of libraries using for instance the two-hybrid system has allowed the identification of several proteins which interact with different SH3 domains. Deletion experiments of these proteins have allowed identification of Pro-rich regions that bind with reasonable affinity to some SH3 domains (Pawson, 1995). (ii) The other approach has been the use of random or Pro-biased libraries of peptides (Chen, 1993; Yu et al., 1994; Cheadle et al., 1994; Rickles et al., 1994; Sparks et al., 1994, 1996; Weng et al., 1995; Combs et al., 1996). The dissociation constants found for the *in vitro* interaction of different proline-based peptides with different SH3 domains are in the 0.2−100 μM range (Viguera et al., 1994; Chen, 1993; Yu et al., 1994; Cheadle et al., 1994; Rickles et al., 1994; Cussac et al., 1994; Feng et al., 1994; Wittekind et al., 1994; Lim et al., 1994a,b; Pisabarro et al., 1994c; Alexandropoulos et al., 1995; Knudsen et al., 1995).

Although the biased combinatorial approach can be successfully employed, the sequences which are identified are clearly influenced by the biases introduced. This can be avoided by using a complete random library, but this is problematic for a 10-residue peptide due to the enormous number of combinations (10^{20}). Therefore, rational modification of the interaction of SH3 domains with proline-based peptides, when the structure of the complex is known, could be an alternative and complementary approach to the use of phage-display libraries. The complex between SH3 domains and proline-rich peptides constitute a good model system to do this, due to the fact that several different structures of SH3 domains have been solved in the presence (Yu et al., 1994; Feng et al., 1994; Wittekind et al., 1994; Lim et al., 1994a,b; Musacchio et al., 1994a; Kohda et al., 1994; Terasawa et al., 1994; Wu et al., 1995) or absence of a ligand

† M.T.P. is a Boehringer Ingelheim Fonds Foundation fellow.
* To whom correspondence should be addressed.
⊗ Abstract published in *Advance ACS Abstracts,* July 15, 1996.
[1] Abbreviations: SH3, Src homology domain 3; K_d, dissociation constant.

S0006-2960(96)00203-6 CCC: $12.00

(Musacchio et al., 1992b; Kohda et al., 1993, 1994; Yu et al., 1992, 1993; Noble et al., 1993; Koyama et al., 1993; Booker et al., 1993; Eck et al., 1994; Yang et al., 1994; Borchert et al., 1994; Farrow et al., 1994; Maignan, 1995; Liang et al., 1996; Grosser et al., 1995). We have chosen the complex between the Abl-SH3 domain and the 3BP1 peptide for several reasons:

(i) It is one of the structures solved to a higher resolution (Musacchio et al., 1994a; manuscript in preparation).

(ii) Although the homology between Abl-SH3 and Fyn-SH3 is not high, both proteins bind the 3BP1 peptide with the same affinity (Viguera et al., 1994), and the structures of both Abl- and Fyn-SH3 domains, in complex with the 3BP1 and 3BP2 peptides, respectively, are known (Musacchio et al., 1994a). This makes the design project more interesting because we can simultaneously attempt to increase the affinity for Abl-SH3 and decrease it for Fyn-SH3.

(iii) Using existing biocomputing tools, we modeled the Abl-SH3 domain (Pisabarro et al., 1994b) and predicted the region involved in the interaction with the proline-rich peptide, prior to the crystal structure being known (Pisabarro et al., 1994c). The modeling exercise gives good background knowledge for further design.

On the basis of above information, we decided to design specific mutations that should increase the specificity and affinity of proline-rich peptides for the Abl-SH3 domain. The final goal has been to obtain a ligand that would have better affinity for Abl-SH3, discriminating between Abl- and Fyn-SH3 domains.

MATERIALS AND METHODS

Cloning and Expression of the Abl-SH3 Domain. The pET3d plasmid coding for the Abl-SH3 domain was a generous gift of A. Musacchio. The Abl-SH3 domain was cloned, expressed, and purified as previously indicated (Viguera et al., 1994; Musacchio et al., 1994a).

Peptide Synthesis. Peptides were provided by the EMBL peptide service. They were synthesized in the solid phase in an MPS column. The molecular weight of the resulting product was confirmed by mass spectrometry. Peptide purity was assessed by analytical HPLC.

Peptide Concentration Determination. Peptide concentration was determined by amino acid analysis. The typical error in this procedure is around 10%.

Circular Dichroic Spectroscopy. CD spectra were recorded in a Jasco-710 instrument at 25 °C. Far-UV spectra were recorded in the range 180−250 nm with a 0.01 cm path-length cuvette. The near-UV spectra were recorded in the range 250−320 nm with a 1 cm path-length cuvette. Peptide concentrations were 200 μM in 5 mM phosphate buffer at pH 7.0.

Fluorescence Spectroscopy. The fluorescence emission spectra of the tryptophan residues in the Abl-SH3 domain were used to monitor any changes in their environment upon peptide binding. Fluorescence was measured in an AMINCO Bowman Series 2 luminescence spectrometer. Excitation was at 298 nm with a 2 nm slit. Fluorescence was detected through a 4 nm slit. In these experiments the final protein concentration was kept at 10 μM, except in the case of those peptides with very high affinity in which 1 μM protein was used. All experiments were performed at 25 °C.

Determination of the K_d by Interpolation. A fast screening of the mutated peptides was done by determining the change in fluorescence of Abl- and Fyn-SH3 domains resulting from the addition of two different peptide concentrations (22 and 111μM) to a final concentration of 10 μM Abl- or Fyn-SH3 domains. At the lowest concentration, the changes in fluorescence of the Abl- or Fyn-SH3 domains are linear with the concentration of 3BP2 and 3BP1 peptides. At the highest concentration the domain is almost saturated (Viguera et al., 1994). Interpolation of the fluorescence change in the titration curves of the Abl- or Fyn-SH3 domains, with 3BP1 and/or 3BP2 (Viguera et al., 1994), allows estimation of the K_d value for the peptides. Those peptides containing a Tyr were interpolated in the titration curves done with 3BP2, while the rest were interpolated on those done with 3BP1. When the fluorescence change was close to that found for the 3BP1 or 3BP2 peptides, we estimated the K_d from the effect produced by the lower concentration of the peptide. If very little fluorescent change took place at the lowest peptide concentration, then we estimated the K_d from the effect produced by the higher peptide concentration. The estimated average error committed with this procedure is around 50% of the K_d value when it is between 2 and 10 μM and up to 100% of the value when it is higher than 100 μM. In those cases in which the extrapolated K_d was close to that of the 3BP2 peptide with Abl-SH3 ($\sim$5 μM), we measured it more accurately as described below (Viguera et al., 1994).

Determination of the K_d by Titration. Assuming a one−one complex between the SH3 domains and the peptide, it is possible to determine the K_d for the interaction of the different peptides, by monitoring the changes in the fluorescence emission, using the equation:

$$F = F_f + (F_b - F_f)[pep_f]/(K_d + [pep_f]) \quad (1)$$

where F_f is the fluorescence of the free domain (Dom_f), F_b is the fluorescence of the complex (Comp), [pep_f] is the concentration of the free peptide in solution, and K_d is the dissociation constant in the equilibrium.

$$K_d = [Dom_f][pep_f]/[Comp] \quad (2)$$

The concentration of the free peptide can be calculated by subtracting the fluorescence estimated concentration of the complex from the concentration of added peptide.

Search for Substituents at Different Positions of the 3BP1 Peptide. The rational modifications of the 3BP1 peptide to enhance its affinity for the Abl-SH3 domain were based on the combined use of biochemical information, molecular graphics analysis (INSIGHTII, 1993), and the GRID (Goodford, 1985) and LUDI (Bohm, 1992a,b) programs. Parameters and protocols similar to the ones used by Pisabarro *et al.*, for the successful modification of human synovial fluid phospholipase A_2 inhibitors were used (Pisabarro et al., 1994a).

Modeling of the Complex of the Abl-SH3 Domain and the High-Affinity Mutant Peptide. The coordinates of the X-ray structure of the Abl-SH3 domain complexed with the 3BP1 peptide (Musacchio et al., 1994a) were used to model the complex of Abl-SH3 with the P40 mutant form of the 3BP1 peptide. First, Met4 was replaced by Tyr, Pro5 by Ser, and Leu8 by Pro in the 3BP1 peptide, and the complex was

energy minimized with the AMBER force field (Pearlman, 1991), as implemented in DISCOVER (INSIGHTII, 1993), by using steepest descent (500 steps) and conjugate gradient energy minimization (500 steps) while keeping the protein fixed. A 10 Å cutoff for the nonbonded interactions and a dielectric constant $e = 4rij$ were used. Then, a 5 Å shell of TIP3P water molecules was added to the complex, which was optimized with a 10 Å cutoff for the nonbonded interactions and a distance-dependent dielectric constant ($1rij$) as follows: First, all the hydrogen atoms were reoriented by using steepest descent (200 steps) and conjugate gradient (500 steps) energy minimization. Then, 500 more steps with each of these minimizers in turn were applied to hydrogen atoms and water molecules and finally to the whole systems.

RESULTS

Design of the Mutations. Interaction of proline-rich peptides with SH3 domains takes place through formation of a polyproline helix II (PPII), for at least half of the peptide length and the stacking of the prolines with aromatic rings in the domain, modulated by specific interactions of non-proline residues with the n-Src and RT loops of SH3 domains (Yu et al., 1994; Feng et al., 1994; Wittekind et al., 1994; Lim et al., 1994a,b; Knudsen et al., 1995; Musacchio et al., 1994a; Kohda et al., 1994). We are interested, first, in increasing the affinity for Abl-SH3 and, second, in designing a ligand that will discriminate between Abl- and Fyn-SH3 domains. To do so we have used two different strategies:

(i) Increasing the tendency of the free peptide to adopt a PPII conformation. This can be done by mutating non-proline residues with a PPII conformation to proline so that the entropic cost upon formation of the complex will be smaller (Rabanal et al., 1993; Williamson, 1994) and/or eliminating alternative conformations in the free peptide incompatible with the structure in the complex (i.e., β-turns etc.). This strategy could increase the affinity for Abl-SH3 but not the specificity, since the 3BP1 and 3BP2 peptides bind with similar conformation to Abl- and Fyn-SH3 domains, respectively (Musacchio et al., 1994a).

In the complex between 3BP1 and the Abl-SH3, there are some peptide residues which are partly or completely buried (Pro2, Met4, Pro6, Pro7, and Pro9), while others are quite solvent exposed (Musacchio et al., 1994a) (Thr3, Pro5, Leu8, and Pro10). Residues Pro2 and Pro5 to Pro10 are adopting a PPII conformation, with Pro7 and Pro10 occupying external positions P0 and P3, respectively, in the consensus nomenclature (Lim et al., 1994b). Leu8 is the only non-proline residue which adopts PPII angles and consequently constitutes a prime target for the first strategy outlined above. Mutation of this residue for a Pro should decrease the entropic cost of forming the complex. Also, since the side chain of Leu8 is quite exposed to the solvent, we reasoned that in solution the peptide could adopt other conformations different from the one in the complex, in which the side chain of Leu8 is partly buried. Mutation of this residue to a smaller and less apolar residue (i.e., Ala) could eliminate some of these alternative conformations and increase the affinity for the protein.

(ii) Creating new favorable interactions specific to Abl-SH3. To identify reasonable substitutions at each one of the 10-positions of the 3BP1 peptide, we scanned the Abl-SH3, with the programs GRID (Goodford, 1985) and LUDI (Bohm, 1992a,b), as well as performed a molecular graphics visual inspection of the 3BP1−Abl-SH3 and 3BP1−Fyn-SH3 complexes (see Materials and Methods). GRID calculates interaction energies of different chemical probes with molecules of known structures. LUDI is a automatic method for *de novo* ligand design that searches in a fragment database and can append a substituent, or modify a chemical group, onto an already existing ligand. The most favorable substitutions that introduced new interactions (i.e., hydrogen bonding, aromatic−aromatic interaction, histidine−aromatic interaction, etc.) or that replaced the Pro residues by another hydrophobic amino acid were selected (Table 1). Although at some positions Leu, Ile, or Val could have fitted, we decided to choose Leu to prevent the known conformational preferences of Val and Ile for the β-strand region of the Ramachandran plot (Muñoz & Serrano, 1994; Swindells et al., 1995). LUDI detected the possibility of an aromatic ring making a very favorable interaction with the same aromatic residues of Abl-SH3 that interact with Pro2 of 3BP1 (Swindells et al., 1995). However, due to the orientation of Pro2, mutation by Phe results in the aromatic ring pointing away from the Abl-SH3 aromatic residues (data not shown).

Table 1: Determination of the K_d Values for the Different Peptides for the Abl- and Fyn-SH3 Domains

	name	sequence[a]	Fyn int[b] (μM)	Fyn cal[c] (μM)	Abl int[b] (μM)	Abl cal[c] (μM)
	3bp2	PPAYPPPPVP		34 ± 3		5 ± 0.2
	3bp1	**R**APTMPPPLPP		34 ± 3		34 ± 3
	p0R	**R**APT**Y**PPPLPP				7 ± 2
	p0	APT**Y**PPPLPP	37	54 ± 5	8	6 ± 2
position 10	p1	APT**Y**PPPLP**H**	114		27	
	p2	APT**Y**PPPLP**L**	156		30	
	p3	APT**Y**PPPLP**Y**	174		30	
position 9	p4	APT**Y**PPPL**H**P	112		27	
	p5	APT**Y**PPPL**N**P	161		95	
	p6	APT**Y**PPPL**L**P	114		20	
position 8	p7	APT**Y**PPP**P**PP	17	27 ± 3	4	2 ± 0.2
	p8	APT**Y**PPP**A**PP	24	65 ± 3	6	5 ± 0.6
	p9	APT**Y**PPP**S**PP	60		7	
position 7	p10	APT**Y**PP**Y**LPP	111		30	
	p11	APT**Y**PP**L**LPP	141		41	
	p12	APT**Y**PP**T**LPP	149		41	
position 6	p13	APT**Y**P**H**PLPP	165		41	
	p14	APT**Y**P**T**PLPP	66		40	
	p15	APT**Y**P**L**PLPP	42		41	
	p16	APT**Y**P**Y**PLPP	117		39	
position 5	p17	APT**YS**PPLPP	136		5	2 ± 0.3
	p18	APT**YH**PPLPP	134		130	
position 4	p19	APT**Q**PPPLPP	102		27	
	p20	APT**R**PPPLPP	76		18	
	p30	APT**K**PPPLPP	106		17	
	p31	Ac-APT**K**PPPLPP	155		4	16 ± 1
position 3	p21	AP**SY**PPPLPP	50		6	
	p22	AP**HY**PPPLPP	54		6	
	p23	AP**GY**PPPLPP	73		8	
position 2	p24	A**LPY**PPPLPP	122		45	
	p25	A**FTY**PPPLPP	161		41	
position 1	p26	**G**PT**Y**PPPLPP	93		20	
	p27	**S**PT**Y**PPPLPP	83		7	

[a] The affinities for the 3BP1 and 3BP2 peptides have been previously described (Viguera et al., 1994). The sequences of the peptides in the one letter code are shown. In bold we show the residues mutated with respect to the 3BP1 peptide. Ac- means that the N-terminus of the peptide was acetylated. [b] K_d values determined by interpolation of the fluorescent change produced when adding two different concentrations of the peptides to the Abl or Fyn-SH3 domains (see Materials and Methods). [c] K_d values obtained using the classical fluorescence titration method (Viguera et al., 1994).

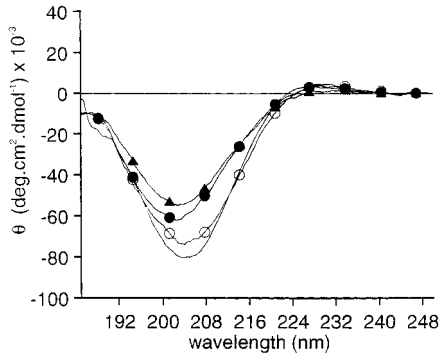

FIGURE 1: Far-UV circular dichroism spectra of a 10-residue polyproline peptide (continuous line) and of peptides p0 (APTYPPPLPP, filled triangles), p7 (APTYPPPPPP, open circles), and p8 (APTYPPPAPP, filled circles). The conditions are those indicated in Materials and Methods.

To allow for the interaction of an aromatic ring with these residues, we mutated two residues, Ala1 and Pro2, to Phe and Gly, respectively (peptide p28, Table 1). Modeling of the p28 peptide indicates that Gly at position 2 could allow the aromatic ring at position 1 to pack favorably with the Abl-SH3 aromatic residues (data not shown). As a control we designed a peptide in which Pro2 was mutated by a Phe (p25).

Far-UV Circular Dichroism Analysis of the Peptides in Solution. Substitution of Leu8 by a Pro should result in an increase in the PPII conformation of the peptide. To see that this is the case, we have analyzed by far-UV circular dichroism the modified version of the 3BP1 peptide that we use here as a reference (p0; see Table 1), the two peptides in which Leu8 has been mutated by a Pro (p7) or Ala (p8), and that of a 10-residue polyproline that adopts a PPII conformation (Viguera et al., 1994). As expected, the CD spectrum of p7 is closer than that of p0 to that of the 10-residue polyproline peptide (Figure 1), thus indicating a significant increase in the PPII conformation. Interestingly enough, the spectrum for the p8 peptide indicates a small increase in the PPII conformation (Figure 1), thus supporting our hypothesis that the long hydrophobic side chain of the Leu could favor alternative non-PPII conformations in aqueous solution.

Determination of the Dissociation Constants of the Mutant Peptides for Abl-SH3 and Fyn-SH3 Domains. A fast screening of the mutated peptides was done by interpolation of the fluorescence change produced by adding two different concentrations of the peptides (22 and 111 μM) to Abl- or Fyn-SH3 domains into the titration curves of these domains with 3BP1 and/or 3BP2 (Viguera et al., 1994); those peptides with a Tyr were interpolated on the 3BP2 titration curves, while the rest were interpolated on the 3BP1 titration curves. The average error in the estimation of the K_d is high for those peptides having low affinity (>100 μM) but smaller for those with high affinity (2 to 10 μM; around 50% of the real value on average; see Materials and Methods). Therefore, interpolation of the fluorescence change in the titration curves of the Abl- or Fyn-SH3 domains, with 3BP1 and 3BP2 (Viguera et al., 1994), allows us to obtain a reasonable estimation of the K_d for the high-affinity mutant peptides. The results of this analysis are shown in Table 1. Whenever, at a certain position, we found that the estimated K_d values of the mutated peptides for Abl-SH3 were lower or similar to that of the wild-type peptide, and significantly higher for the Fyn-SH3 domain, we measured the K_d with more precision using the standard titration procedure (Viguera et al., 1994) (Table 1). The only exceptions were the point mutations at positions 8 of the peptide (peptides p7 and p8). In this case we measured the K_d with more precision since in both Abl- and Fyn-SH3 domains there was a significant increase in affinity.

Table 2: Determination of K_d Values for Those Peptides Containing Some of the Mutations That Increased the Affinity for the Abl-SH3 Domain

name	sequence[a]	Fyn int[b] (μM)	Fyn cal[c] (μM)	Abl int[b] (μM)	Abl cal[c] (μM)
p28	**FGTY**PPPLPP	206	97 ± 10	7	8 ± 0.4
p32	**SPSYSPPP**PP				1.8 ± 0.2
p33	Ac-**SPSYSPPP**PP		496 ± 60		1.9 ± 0.6
p40	APT**YSPPP**PP		472 ± 55		0.4 ± 0.1
p41	AP**SYSPPP**PP		273 ± 30		1.5 ± 0.1
p50	APTMPPP**P**PP				6.2 ± 1.3
p51	APTM**S**PPLPP				22 ± 2

[a] The sequences of the peptides in the one letter code are shown. Ac- means that the N-terminus of the peptide was acetylated. [b] K_d values determined by interpolation of the fluorescent change produced when adding two different concentrations of the peptides to the Abl- or Fyn-SH3 domains (see Materials and Methods). [c] K_d values obtained using the classical fluorescence titration method (Viguera et al., 1994).

When we first analyzed the affinity of the 3BP1 and 3BP2 peptides for different SH3 domains, we placed an extra Arg residue at the 3BP1 N-terminus (Viguera et al., 1994; Pisabarro et al., 1994c). In Table 1 we show that this extra Arg residue does not influence the affinity for the Abl-SH3 domain (see p0 and p0R). Therefore, the increase in affinity that we observe when mutating Met4 to Tyr in 3BP1 (Pisabarro et al., 1994c) is solely due to the Tyr residue.

As can be seen from Table 1, with the exception of two mutations, Asn instead of residue Pro9 [P_2 in Lim nomenclature (Lim et al., 1994b)] and His instead of residue Pro5 [P_{-2} in Lim nomenclature (Lim et al., 1994b)], in all cases the affinity of the mutated peptides for Abl-SH3 is lower or similar, within the experimental error, to that for the wild-type peptide 3BP1 peptide (Viguera et al., 1994).

Analysis of Peptides with Multiple Favorable Mutations. When some of the peptides that in the first analysis exhibited similar or lower K_d values to that of the original peptide [p0 (Pisabarro et al., 1994c)] were analyzed more precisely by the classical titration procedure, it was found that, in some cases (p7, p8, and p17), the affinity for Abl-SH3 was increased (Table 1).

Assuming that the effect of mutations at different positions is additive, we should expect an increase in affinity and specificity for Abl-SH3 by combining the individual mutations. We analyzed different peptides bearing several mutations to find if this is the case and also to analyze the effect of other mutations that by the interpolation method seemed to result in K_d values similar to that of the original peptide p0 (Table 2). Acetylation does not change significantly the affinity of the peptides for the Abl-SH3 domain (see p32 and p33), thus indicating that the positive charge at the N-terminus does not play an important role. The best combination we found is the one containing the Met4 by Tyr, the Pro5 by Ser, and the Leu8 by Pro mutations (p40). This peptide has a K_d which is around 15 times lower (~400 nM) than that of peptide p0 and 100 times lower than that of the wild-type 3BP1 peptide (Table 2). What is more

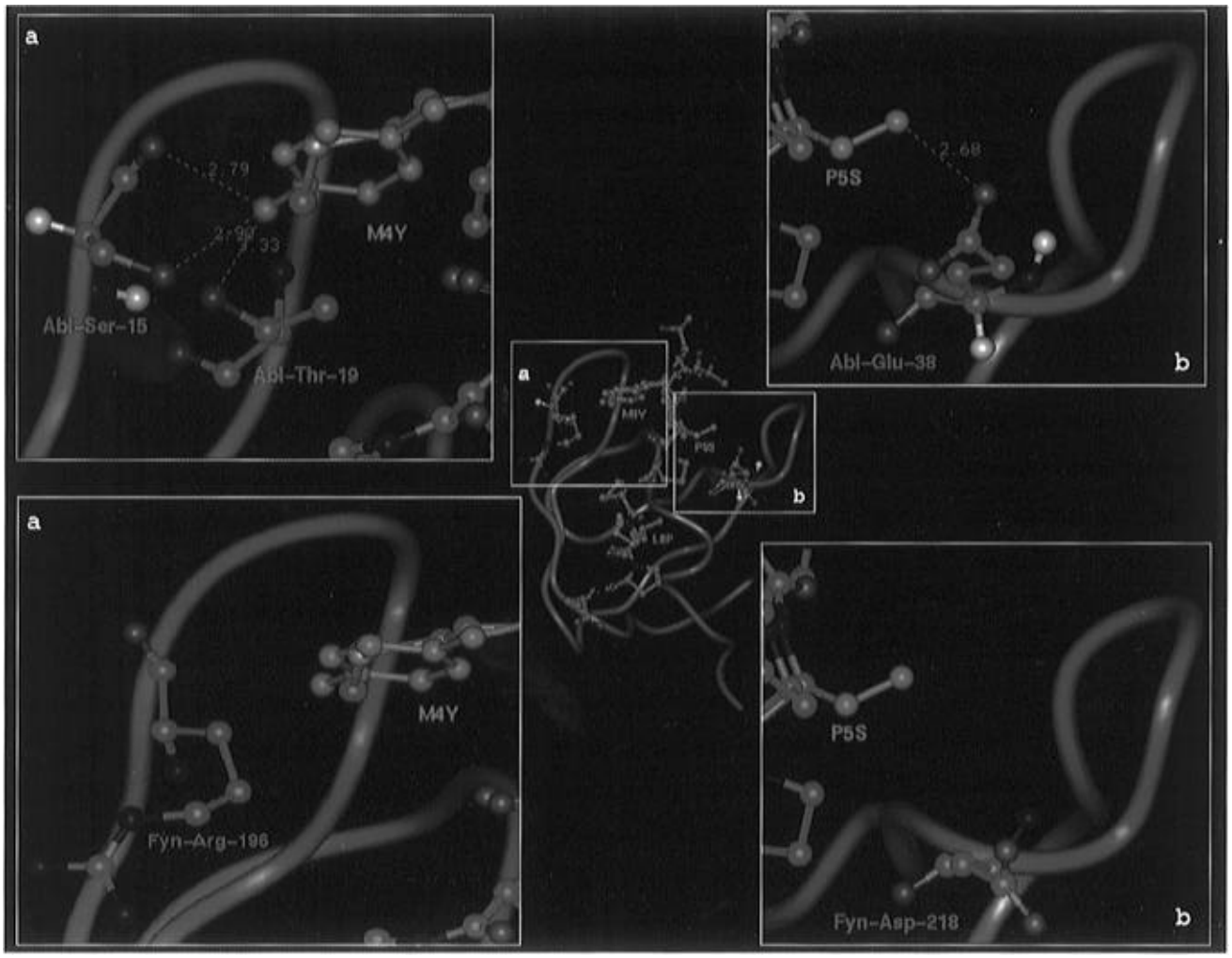

FIGURE 2: The central part of the picture shows the modeled complex between the Abl-SH3 domain (shown as a magenta ribbon) and the p40 peptide modeled on 3BP1. The mutated residues in the p40 peptide (M4Y, P5S, L8P) are shown in magenta, while the corresponding residues in 3BP1 are shown in amino acid color code. Only the main and side chains of the residues in the Abl- and Fyn-SH3 domains involved in specific interactions with the p40 peptide are shown (Abl-Ser-15, Abl-Thr-19, Abl-Glu-38; Fyn-Arg-196, Fyn-Asp-218). Regions corresponding to boxes a and b in the central picture are shown in more detail at the right and left corners of the picture. Abl-SH3 is shown in the upper boxes a and b, while Fyn-SH3 is shown in the lower ones.

important is that the affinity for Fyn-SH3 is now 10 times lower than that of the wild-type peptide, so we have increased the specificity by around 1000 times.

Modeling of the SH3-Abl Domain Complexed to the p40 Peptide. On the basis of the three-dimensional structure of the complex between the SH3-Abl domain and the 3BP2 peptide, we have modeled the structure of the complex with the mutated p40 peptide. After substitution of the side chains of the mutated residues (Met4 by Tyr, Pro5 by Ser, and Leu8 by Pro) and minimization of the complex in the presence of water (see methods), the reasons for the affinity and specificity increase are evident (Figure 2). The side chain of p40-Tyr4 occupies the hydrophobic cavity occupied by Met2 in 3BP1, but its OH group makes two short hydrogen bonds with the main chain carbonyl and the side chain hydroxyl of Abl-Ser15 and a long one with that of Abl-Thr19. In Fyn-SH3 the equivalent residue to Ser15 is Fyn-Arg196, whose side chain cannot make a hydrogen bond to the hydroxyl group of p40-Tyr4. Also, instead of Thr19, Fyn has an Asp218 residue which will be partially buried by p40-Tyr4. The side chain OH group of p40-Ser5 makes a hydrogen bond to the side chain of Abl-Glu38 that overcompensates the entropic cost paid by the mutation of a Pro to a Ser. In Fyn-SH3 this residue is an Asp, whose side chain is too short to make the same hydrogen bond. This explains why mutation of Pro5 by Ser increases the affinity for Abl-SH3 but significantly decreases it for Fyn-SH3. Finally, in the case of the Leu8 to Pro mutation, the increase in affinity should be due to the smaller loss of conformational freedom in the peptide upon binding to the protein (see above), since the position is solvent exposed and has PPII dihedral angles. The fact that there is also an increase in affinity for Fyn-SH3 supports this hypothesis (from ~54 to ~27 μM; Table 1).

DISCUSSION

Many of the so far identified ligands for SH3 domains using the two-hybrid system or phage-display random biased libraries are rich in Pro residues. This could be due to the fact that they fit better into the aromatic pockets of SH3 domains and readily adopt a PPII conformation in solution without a significant loss in conformational entropy. This PPII conformation seems to fit best the steric and hydrogen-bonding pattern of the SH3 ligand surface (Lim et al., 1994b). However, the possibility that other sequences with fewer Pro residues could interact with SH3 domains is still open, being

the entropic cost of fixing them in a PPII conformation paid by the surrounding protein (Grzesiek et al., 1996). Mutation of the solvent-exposed Leu8 by Pro supports the idea that the entropic cost is one of the reasons why Pro-rich ligands are selected. Also, as indicated by the higher affinity of the Leu8 to Ala mutant, the existence of conformations in the free peptide different from that of PPII, which could be stabilized by certain amino acids, favors Pro residues.

All the mutations we have done on the p0 peptide, with the exception of Pro5 to His and Pro9 to Asn, did not decrease the affinity from 6μM to more than $\sim$30$-$40 μM. This is opposite to what has been suggested about the crytical role of some Pro residues in SH3 ligands (Ren et al., 1993; Lim et al., 1994b). In principle, all the Pro residues in the p0 peptide can be substituted for different amino acids as long as a good replacement and/or a compensating mutation is introduced elsewhere, and consequently, there seems to be no essential Pro residues in the interaction of a peptide with a SH3 domain. A very good example of this is the p28 peptide. In this peptide, mutation of Ala1 by Phe and Pro2, which has been suggested to be crucial for binding (Musacchio et al., 1994a), by Gly does not produce a significant change in affinity with respect to the p0 peptide. This indicates that the repertoire of SH3 binding peptides could be much larger than previously thought. However, we cannot eliminate, in all cases, the unlikely possibility that mutation of one of the so-called critical Pro residues in the 3BP1 and 3BP2 peptides results in the mutant peptide adopting a different orientation (Yu et al., 1994; Feng et al., 1994; Wittekind et al., 1994; Lim et al., 1994a,b; Pisabarro et al., 1994c; Musacchio et al., 1994a; Kohda et al., 1994). If this occurs, then some of the other Pro residues could play the role of the mutated ones. Although it seems possible to substitute a Pro residue by another amino acid that interacts favorably with the SH3 domains, this does not mean that we can substitute several Pro residues at the same time. Mutation of a Pro residue by other amino acids introduces a certain degree of freedom in the polypeptide chain, but this effect is not additive and grows exponentially.

Following the reasoning for our design explained in results, we have found that by introducing favorable interactions it is possible to increase the specificity and affinity of Pro-rich ligands for Abl-SH3. The positions we have found which determine the specificity for Abl- and Fyn-SH3 domains are the first five in the 3BP1 peptide. These positions mainly interact with the RT and n-Src loops, confirming what has been previously indicated (Musacchio et al., 1994a), that these are the main regions involved in SH3 specificity. Especially important in our case seems to be residues 4 and 5. Introduction of a Tyr at position 4 of the peptide increases significantly the affinity with respect to Met (Pisabarro et al., 1994c) and also improves the specificity with respect to Fyn-SH3. The reason for it seems to be the presence of a Ser residue at position 15 and a Thr at position 19 of Abl-SH3 that can make a hydrogen bond to the Tyr4 side chain of the peptide. Mutation of Thr19 of Abl-SH3 by Asp results in a 12-fold decrease in its affinity for a peptide containing a Tyr at the position equivalent to residue 4 of 3BP1 (Weng et al., 1995). The authors concluded that Thr19 was important for conferring selectivity toward Tyr. In our modeling the OH group of Thr19 is close to the hydroxyl group of Tyr4 in the peptide and can make a long hydrogen bond, explaining the experimental results in the T19D mutation. In Fyn-SH3 there is an Asp residue at position 19 and an Arg at position 15; therefore, a Tyr residue should not be favorable, and this is what we have found. At position 5 of our peptides we have found that a Ser residue is better than a Pro, although Pro5 adopts a PPII conformation (Mussachio et al., 1992a). The reason for it seems to be because its side chain can make a hydrogen bond to the side chain of Glu38 in Abl-SH3. In Fyn-SH3 there is an Asp residue whose side chain is too short to make the equivalent hydrogen bond. The effect of the individual mutations at positions 4 and 5 seems to be approximately additive. This can be seen when mutating Leu8 by Pro or Pro5 by Ser in the 3BP1 peptide (p50 and p51, respectively) or in the p0 peptide (p7 and p17, respectively). Addition of an Arg at the N-terminus does not seem to modify the affinity for Abl-SH3 of peptide p0 (peptide p0R) nor does an acetylated N-terminus (see Tables 1 and 2). However, we cannot discard the fact that other residues placed before position 1 of 3BP1 could still improve the affinity and specificity for Abl-SH3. At positions 6$-$10 we have not found any mutation which significantly improves the specificity for Abl- over Fyn-SH3. However, although the data are not very reliable because it is obtained through interpolation, we find that substitution of Pro6 by Leu or Thr in the p0 peptide results in a more specific binding for Fyn-SH3. In Src-SH3, which is closer in sequence to Fyn-SH3 than to Abl-SH3, it has been found also that an aliphatic residue is generally preferred at the equivalent position to residue 6 of 3BP1 (Sparks et al., 1996).

Interestingly enough, three different screening studies of phage-displayed libraries (Rickles et al., 1994; Weng et al., 1995; Sparks et al., 1996) also found that an aromatic residue is favored at the equivalent position to residue 4 of 3BP1. At the equivelent position to residue 5 of 3BP1 phage-displayed studies found mainly Pro, and only in one case was a Ser detected (Weng et al., 1995). However, the same group also found a peptide with a His at this position, and we have seen that this is quite unfavorable (p18, Table 1).

In conclusion, by using rational design we have found a peptide with the sequence APTYSPPPPP that binds to Abl-SH3 with an affinity 100 times higher than the original 3BP1 peptide and has 1000 times more affinity for Abl-SH3 than for Fyn-SH3 domains. The increase in affinity can be explained by the introduction of more favorable specific interactions and also the reduction of the entropy cost in fixing the peptide in a PPII conformation. The increase in specificity is mainly due to interactions made by residues Tyr4 and Ser5 with the RT and *n-Src* loops. We suggest that residues Ser15, Thr19, and Glu38 in the Abl-SH3 domain play an important role in determining ligand specificity. Also, we have found that, in principle, it is possible to substitute every Pro residue in 3BP1 without producing a large decrease in affinity as long as a good substituent is chosen. This is, for example, the case of the sequence FGTYPPPLPP, which also has a very good K_d value (8 $\pm$ 0.4 μM). Therefore, we propose in the line of recent evidence that SH3 ligands *in vivo* do not necessarily need to be very rich in Pro residues as long as the entropic cost in the ligand binding region is paid by the rest of the protein. Our work confirms that, by using existing software tools and simple chemical reasoning, it is possible to rationally modify the affinity and specificity of a peptide to an SH3 domain when the structure of the complex is known.

ACKNOWLEDGMENT

We are grateful to Dr. M. Saraste for his helpful suggestions.

REFERENCES

Alexandropoulos, K., Cheng, G., & Baltimore, D. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3110−3114.
Bohm, H. J. (1992a) *J. Comput.-Aided Mol. Des. 6*, 61−78.
Bohm, H. J. (1992b) *J. Comput.-Aided Mol. Des. 6*, 593−606.
Booker, G. W., Gout, I., Downing, A. K., Driscoll, P. C., Boyd, J., Waterfield, M. D., & Campbell, I. D. (1993) *Cell 73*, 813−822.
Borchert, T. V., Mathieu, M., Zeelen, J. P., Courtneidge, S. A., & Wierenga, R. K. (1994) *FEBS Lett. 341*, 79−85.
Cheadle, C., Ivashchenko, Y., South, V., Searfoss, G. H., French, S., Howk, R., Ricca, G. A., & Jaye, M. (1994) *J. Biol. Chem. 269*, 24034−24039.
Chen, J. K., Lane, W. S., Brauer, A. W., Tanaka, A., & Schreiber, S. L. (1993) *J. Am. Chem. Soc. 115*, 12591−12592.
Cicchetti, P., Mayer, B. J., Thiel, G., & Baltimore, D. (1992) *Science 257*, 803−806.
Combs, A. P., Kapoor, T. M., Feng, S., Chen, J. K., Daude-Snow, L. F., & Schreiber, S. L. (1996) *J. Am. Chem. Soc. 118*, 287−288.
Cussac, D., Frech, M., & Chardin, P. (1994) *EMBO J. 13*, 4011−4021.
Eck, M. J., Atwell, S. K., Shoelson, S. E., & Harrison, S. C. (1994) *Nature 368*, 764−769.
Farrow, N. A., Zhang, O., Forman, K. J., & Kay, L. E. (1994) *J. Biomol. NMR 4*, 727−734.
Feng, S., Chen, J. K., Yu, H., Simon, J. A., & Schreiber, S. L. (1994) *Science 266*, 1241−1247.
Goodford, P. J. (1985) *J. Med. Chem. 28*, 849−857.
Gosser, Q. Y., Zheng, J., Overduin, M., Mayer, B. J., & Cowburn, D. (1995) *Structure 3*, 1075−1086.
Grzesiek, S., Bax, A., Clore, M., Gronenborn, A. M., Hu, J., Haufman, J., Palmer, I., Stahl, S. J., & Wingfield P. T. (1996) *Nat. Struct. Biol. 3*, 340−345.
INSIGHTII, (1993) Biosym Technologies, 9685 Scranton Road, San Diego, CA 92121−2777.
Knudsen, B. S., Zheng, J., Feller, S. M., Mayer, J. P., Burrell, S. K., Cowburn, D., & Hanafusa, H. (1995) *EMBO J. 14*, 2191−2198.
Kohda, D., Hatanaka, H., Odaka, M., Mandiyan, V., Ullrich, A., Schlessinger, J., & Inagaki, F. (1993) *Cell 72*, 953−960.
Kohda, D., Terasawa, H., Ichikawa, S., Ogura, K., Hatanaka, H., Mandiyan, V., Ullrich, A., Schlessinger, J., & Inagaki, F. (1994) *Structure 2*, 1029−1040.
Koyama, S., Yu, H., Dalgarno, D. C., Shin, T. B., Zydowsky, L. D., & Schreiber, S. L. (1993) *FEBS Lett. 324*, 93−98.
Liang, J., Chen, J. K., Schreiber, S. L., & Clardy, J. (1996) *J. Mol. Biol. 257*, 632−643.
Lim, W. A., Fox, R. O., & Richards, F. M. (1994a) *Protein Sci. 3*, 1261−1266.
Lim, W. A., Richards, F. M., & Fox, R. O. (1994b) *Nature 372*, 375−379.
Maignan, S. G., Guilloteau, J.-P., Fromage, N., Arnoux, B., Becquart, J., & Ducruix, A. (1995) *Science 268*, 291−293.
Mayer, B. J., Hamaguchi, M., & Hanafusa, H. (1988) *Nature 332*, 272−275.
Muñoz, V., & Serrano, L. (1994) *Proteins 20*, 301−311.
Musacchio, A., Gibson, T., Lehto, V. P., & Saraste, M. (1992a) *FEBS Lett. 307*, 55−61.
Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., & Saraste, M. (1992b) *Nature 359*, 851−855.
Musacchio, A., Saraste, M., & Wilmanns, M. (1994a) *Nat. Struct. Biol. 1*, 546−551.
Musacchio, A., Wilmanns, M., & Saraste, M. (1994b) *Prog. Biophys. Mol. Biol. 61*, 283−297.
Noble, M. E., Musacchio, A., Saraste, M., Courtneidge, S. A., & Wierenga, R. K. (1993) *EMBO J. 12*, 2617−2624.
Pawson, T. (1992) *Curr. Opin. Struct. Biol. 2*, 432−437.
Pawson, T. (1995) *Nature 373*, 573−580.
Pawson, T., & Gish, G. D. (1992) *Cell 71*, 359−362.
Pawson, T. S. J. (1993) *Curr. Biol. 3*, 434−445.
Pearlman, D. A., Case, D., Caldwell, J. C., Seibel, G. L., Singh, U. C., Weiner, P., & Kollman, P. A. (1991) University of California, San Francisco.
Pisabarro, M. T., Ortiz, A. R., Palomer, A., Cabre, F., Garcia, L., Wade, R. C., Gago, F., Mauleon, D., & Carganico, G. (1994a) *J. Med. Chem. 37*, 337−41.
Pisabarro, M. T., Ortiz, A. R., Serrano, L., & Wade, R. C. (1994b) *Proteins 20*, 203−215.
Pisabarro, M. T., Ortiz, A. R., Viguera, A. R., Gago, F., & Serrano, L. (1994c) *Protein Eng. 7*, 1455−1462.
Rabanal, F., Ludevid, M. D., Pons, M., & Giralt, E. (1993) *Biopolymers 33*, 1019−1028.
Ren, R., Mayer, B. J., Cicchetti, P., & Baltimore, D. (1993) *Science 259*, 1157−1161.
Rickles, R. J., Botfield, M. C., Weng, Z., Taylor, J. A., Green, O. M., Brugge, J. S., & Zoller, M. J. (1994) *EMBO J. 13*, 5598−5604.
Schlessinger, J. (1994) *Curr. Opin. Genet. Dev. 4*, 25−30.
Sparks, A. B., Quilliam, L. A., Thorn, J. M., Der, C. J., & Kay, B. K. (1994) *J. Biol. Chem. 269*, 23853−23856.
Sparks, A. B., Rider, J. E., Hoffman, N. G., Fowlkes, D. M., Quilliam, L. A., & Kay, B. K. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 1540−1544.
Swindells, M. B., MacArthur, M. W., & Thornton, J. M. (1995) *Nat. Struct. Biol. 2*, 596−603.
Terasawa, H., Kohda, D., Hatanaka, H., Tsuchiya, S., Ogura, K., Nagata, K., Ishii, S., Mandiyan, V., Ullrich, A., Schlessinger, J., & Inagaki, F. (1994) *Nat. Struct. Biol. 1*, 891−897.
Viguera, A. R., Arrondo, J. L., Musacchio, A., Saraste, M., & Serrano, L. (1994) *Biochemistry 33*, 10925−10933.
Weng, Z., Rickles, R. J., Feng, S., Richard, S., Shaw, A. S., Schreiber, S. T., & Brugge, J. S. (1995) *Mol. Cell. Biol.* 5627−5634.
Williamson, M. P. (1994) *Biochem J. 297*, *249*−260.
Wittekind, M., Mapelli, C., Farmer, B. N., Suen, K. L., Goldfarb, V., Tsao, J., Lavoie, T., Barbacid, M., Meyers, C. A., & Mueller, L. (1994) *Biochemistry 33*, 13531−13539.
Wu, X., Knudsen, B., Feller, M. S., Zheng, J., Sali, A., Cowburn, D., Hanafusa, H., & Kuriyan, J. (1995) *Structure 3*, 215−226.
Yang, Y. S., Garbay, C., Duchesne, M., Cornille, F., Jullian, N., Fromage, N., Tocque, B., & Roques, B. P. (1994) *EMBO J. 13*, 1270−1279.
Yu, H., Rosen, M. K., Shin, T. B., Seidel, D. C., Brugge, J. S., & Schreiber, S. L. (1992) *Science 258*, 1665−1668.
Yu, H., Rosen, M. K., & Schreiber, S. L. (1993) *FEBS Lett. 324*, 87−92.
Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W., & Schreiber, S. L. (1994) *Cell 76*, 933−945.

BI960203T