THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 39, pp. 26918–26927, September 25, 2009

Structural, Functional, and Bioinformatic Studies Demonstrate the Crucial Role of an Extended Peptide Binding Site for the SH3 Domain of Yeast Abp1p*[S]

Received for publication, June 1, 2009, and in revised form, July 6, 2009 Published, JBC Papers in Press, July 9, 2009, DOI 10.1074/jbc.M109.028431

Elliott J. Stollar[‡§], Bianca Garcia[‡], P. Andrew Chong[§], Arianna Rath[§], Hong Lin[§], Julie D. Forman-Kay[§¶1], and Alan R. Davidson[‡¶2]

From the Departments of [‡]Molecular Genetics and [¶]Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8 and the [§]Program in Molecular Structure and Function, Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada

SH3 domains, which are among the most frequently occurring protein interaction modules in nature, bind to peptide targets ranging in length from 7 to more than 25 residues. Although the bulk of studies on the peptide binding properties of SH3 domains have focused on interactions with relatively short peptides (less than 10 residues), a number of domains have been recently shown to require much longer sequences for optimal binding affinity. To gain greater insight into the binding mechanism and biological importance of interactions between an SH3 domain and extended peptide sequences, we have investigated interactions of the yeast Abp1p SH3 domain (AbpSH3) with several physiologically relevant 17-residue target peptide sequences. To obtain a molecular model for AbpSH3 interactions, we solved the structure of the AbpSH3 bound to a target peptide from the yeast actin patch kinase, Ark1p. Peptide target complexes from binding partners Scp1p and Sjl2p were also characterized, revealing that the AbpSH3 uses a common extended interface for interaction with these peptides, despite K_d values for these peptides ranging from 0.3 to 6 μM. Mutagenesis studies demonstrated that residues across the whole 17-residue binding site are important both for maximal *in vitro* binding affinity and for *in vivo* function. Sequence conservation analysis revealed that both the AbpSH3 and its extended target sequences are highly conserved across diverse fungal species as well as higher eukaryotes. Our data imply that the AbpSH3 must bind extended target sites to function efficiently inside the cell.

Many protein interactions within signaling pathways are mediated by small modular domains, which are found within larger proteins (1). SH3 domains are one of the most frequently occurring of these protein-protein interaction modules in eukaryotic cells. These domains are ~60-residue β-sheet proteins that have been generally observed to bind to short proline-rich peptides containing the core consensus sequences +*XX*P*XX*P (class I) or P*XX*P*X*+ (class II), where *X* can be a variety of residues, and + is a Lys or Arg residue (2–4). SH3 domains often bind peptides with modest affinities (5–100 μM (5)), and many SH3 domains appear to possess low specificity, binding to various P*XX*P-containing peptides with similar affinities (6–10). These observations have led to the establishment of a "promiscuous model," which postulates that the signaling specificity of pathways depends primarily on factors other than the intrinsic binding properties of isolated SH3 domains (11), and that short peptide targets are likely sufficient for SH3 domain function. Arguing against this model, it has been shown that some SH3 domains require an extended target peptide (12–30 residues) to achieve maximal binding affinity (12–15). These results imply that the intrinsic specificity of SH3 domains may indeed play a significant role. The growing realization of the importance of interactions between SH3 domains and extended peptides provides the motivation for further studies to investigate their molecular mechanisms and functional properties.

In the present study, we have investigated the SH3 domain from the yeast actin-binding protein 1 (Abp1p),[3] which is known to interact with extended peptide targets. Abp1p is an actin-associated protein that possesses an N-terminal actin depolymerizing factor/cofilin homology domain, a central Pro-rich region, and a C-terminal SH3 domain. Through its multiple protein-protein interactions, Abp1p plays important roles in coordinating the dynamic series of events leading to endocytosis (16). In particular, the Abp1p SH3 domain (AbpSH3) binds to and localizes the actin patch kinases, Ark1p and Prk1p (17), which are required for actin patch disassembly after vesicle internalization (18), and also mediates the localization of the Synaptojanin-like protein, Sjl2p, to the actin patch where it plays key roles in endocytosis (19). The AbpSH3 is essential for the biological activity of Abp1p under all genetic and growth conditions where Abp1p is required for cell viability (20, 21).

* This work was funded by operating grants from the National Cancer Institute of Canada (to J. F. K.), the Canadian Institutes of Health Research (CIHR ARD Grant MOP-13609), a Restracomp award from the Hospital for Sick Children (to E. J. S.), and a CIHR Training Program in Protein Folding award (to E. J. S.).

The atomic coordinates and structure factors (code 2rpn) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1 and S2 and Tables S1–S3.

[1] To whom correspondence may be addressed: Program in Molecular Structure and Function, Hospital for Sick Children, 555 University Ave., Toronto, Ontario M5G 1X8, Canada. Tel.: 416-813-5358; Fax: 416-813-5022; E-mail: forman@sickkids.ca.

[2] To whom correspondence may be addressed: Dept. of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada. Tel.: 416-978-0332; Fax: 416-978-6885; E-mail: alan.davidson@utoronto.ca.

[3] The abbreviations used are: Abp1p, actin-binding protein 1; AbpSH3, Abp1p SH3 domain; ITC, isothermal titration calorimetry; NOE, nuclear Overhauser effect; WT, wild type; PPII, polyproline type II; Apc, adenomatous polyposis coli protein.

Abp1p is highly conserved both in fungi and higher eukaryotes, and its homologues in mammalian cells are also involved in processes such as endocytosis that require actin rearrangement (reviewed in Ref. 22).

The SH3 domain of Abp1p (AbpSH3) has been shown to bind peptides from the yeast proteins Ark1p, Prk1p, Scp1p, Sjl2, and Srv2p (8, 17, 19, 20, 23–25). The conservation seen in an alignment of these peptide sequences and results from phage display experiments imply that a region spanning at least 10 residues is critical for AbpSH3 binding (17). In particular, the total conservation in AbpSH3 binding sites of Leu at a position three residues C-terminal to the P*XX*Px+ core consensus strongly suggests that extended peptide interactions are crucial for the function of the domain. A distinctive feature of the AbpSH3 as compared with most other yeast SH3 domains is that its interactions with target peptides in three different proteins, Ark1p, Scp1p, and Sjl2p, have been proven to be physiologically relevant. Deletion of the genes encoding these proteins, or deletion of the Abp1p SH3 domain binding sites within them, leads to a phenocopy of AbpSH3 mutations (19, 20). However, which target site residues are crucial for binding and whether the whole site is required for *in vivo* function are not known.

The well characterized nature of the AbpSH3 and its biologically relevant peptides coupled with our ability to probe the function of this domain in an authentic *in vivo* system (20) make it an excellent model for addressing the mechanism and importance of extended target peptide interactions. The goals of the work described here were to elucidate the molecular details of the AbpSH3 peptide binding and to evaluate the functional importance of the extended AbpSH3 binding interface. To this end, we used NMR spectroscopy to determine the solution structure of the AbpSH3 bound to a 17-residue peptide from the Ark1p target protein. In addition, we assessed the *in vitro* affinities and *in vivo* functional properties of three biologically relevant peptides and a series of target peptide mutants. These studies combined with an analysis of AbpSH3 binding site conservation among diverse fungal species have allowed us to clearly demonstrate the functional importance of the interaction of the AbpSH3 with extended target peptides.

EXPERIMENTAL PROCEDURES

Sample Preparation for in Vitro Studies—Residues 535–592 of the yeast Abp1p protein (AbpSH3 construct) or the residues KKTKPTPPPKPSHLKPK from the yeast protein Ark1p peptide (ArkA construct) were expressed from a pET32b-based (Novagen) plasmid to produce proteins with thioredoxin and a 6-His tag fused at the N terminus. These fused portions were ultimately removed using tobacco etch virus protease. Protein purification was carried by nickel-affinity chromatography as previously described (20). The AbpSH3 was further purified through chromatography on a HiTrap Q column (Amersham Biosciences). Target peptides were further purified using reversed-phase high-performance liquid chromatography with a Phenomenex C18 preparatory column, using a 0 to 50% gradient of acetonitrile over 5 column volumes to obtain a purity of >95% judged by mass spectrometry. Unlabeled peptides for ITC measurements were synthesized (Biomer Technology, Hayward, CA) with C-terminal amidation and N-terminal acetylation and purified by high-performance liquid chromatography as above. The synthesized and synthetic peptides gave the same heteronuclear single quantum coherence spectra and bound with the same affinity.

NMR Spectroscopy—Experiments were measured at 10 °C on a Varian INOVA 500-MHz spectrometer equipped with a pulsed field gradient unit and a triple resonance probe. Samples were in buffer containing 50 mM Na_2HPO_4, pH 7.0, 100 mM NaCl, 1 mM EDTA, 0.05% NaN_3, 10% D_2O, unless otherwise stated. We titrated ^{15}N,^{13}C ArkA-labeled peptide into ^{15}N,^{13}C-labeled AbpSH3, to obtain a 1:1 complex as judged by minimization of peak broadening for both AbpSH3 and peptide due to exchange between free and complexed states. Standard experiments were then used to assign the backbone and side-chain resonances of the doubly labeled AbpSH3·ArkA peptide complex (26). Assignments were incomplete for atoms with complete peak broadening, which include the side chains of Tyr-10 and Phe-50 in AbpSH3 and the C_ϵ and N_Z atoms at the end of the Lys^{-3} side chain in ArkA. Structural restraints were restricted to NOEs, and HN-N, N-CO, CA-HA, CO-CA, and CA-CB residual dipolar couplings from a sample oriented in phage (27, 28). The axial component of the alignment tensor (D_a) was 9.0, and the rhombicity (D_r/D_a) was 0.37. Due to the large number of backbone restraints, we were able to exclude TALOS-based (29) dihedral angle restraints.

After initial assignment of 300 NOEs, for which symmetry-related peaks were observed, ARIA1.2 was used to facilitate the assignment of NOEs and calculate structures. The 20 lowest energy structures, of 200 calculated in the final iteration, were used for analysis. PROCHECK software (30) was used to analyze the Ramachandran plots (supplemental Table S1).

We collected backbone T_1, T_2, and ^1HN-^{15}N NOE data for the free and ArkA-bound complexes at two different concentrations (0.3 mM and 1 mM) to calculate correlation times and check for any concentration-dependent oligomerization. The correlation times stayed constant at the different concentrations (8.08 and 7.97 ns, respectively), confirming a 1:1 complex as seen from the ITC data. The combined ^1H and ^{15}N chemical shift differences between free and bound complexes were calculated using the following expression: $\Delta_{ppm} = \sqrt{((N_{ppm}\ \text{shift} * 0.1)^2 + (H_{ppm}\ \text{shift})^2)}$, where N_{ppm} shift is the difference in ppm in the nitrogen dimension and H_{ppm} is the difference in ppm in the proton dimension. The structure has been deposited into the protein data bank, accession code: 2rpn.

ITC Binding Experiments—Experiments were performed in 50 mM sodium phosphate (pH 7.0), 100 mM NaCl, following our previously described procedure (20). All experiments were repeated at least two times.

Construction of Yeast Expression Plasmid—Plasmid expression vectors to produce WT and mutant versions of Ark1p fused to green fluorescent protein were derived from pRS316. The *ARK1* gene was amplified by PCR from yeast chromosomal DNA using primers (5′-TAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTGA**GAGGACTGGCAGGGGGA**-3′ and 5′-GAATTGGGACAACTCCAGTGAAAAGTT**CTTCTCCTTTACTCTTATCCAAGGATAACTTTCG**-3′), which anneal to the *ARK1* gene (in bold)

and contain pRS316 and green fluorescent protein sequences, respectively, to flank the *ARK1* gene. All *ARK1* AbpSH3 binding site mutations were created by PCR. All generated constructs were sequenced to ensure that no unwanted mutations were introduced during plasmid construction. The plasmids used in this study can be found in supplemental Table S3.

Growth Assays in Yeast—Yeast strains and plasmids used are listed in supplemental Tables S2 and S3, respectively. Strains were manipulated, and media were prepared using standard techniques (31). Standard rich medium supplemented with glucose (YPD, 1% yeast extract, 2% peptone, 2% dextrose) was used for yeast growth without selection. Minimal medium (SD) lacking uracil was used for transformants selection and plasmid maintenance. To assay growth on solid media, cells were grown overnight to saturation in selective media supplemented with 2% dextrose and the appropriate amino acids. Serial 10-fold dilutions of an equal number of cells (0.1 $A_{600\ nm}$) were spotted onto SD (lacking uracil) plates, and incubation was for 2 days at 30 °C or 37 °C. All growth assays were repeated at least three times, and one representative experiment is shown in Fig. 4. The same levels of growth were seen in all repetitions.

Sequence Alignment Construction—We carried out bioinformatic investigations on 29 species of fungi in addition to *Saccharomyces cerevisiae* that were identified as possessing homologues of Abp1p in BLAST search (Altschul, *et al.* (45)) of the Uniprot data base (www.pir.uniprot.org/). Target protein homologues were also identified by BLAST searches. The validity of hits from these searches was confirmed by detection of regions of high sequence similarity across the length of the protein sequences. Although we analyzed sequence conservation across all 29 species, considerable sequence redundancy existed among them. We show alignments of sequences from some of the most diverse species examined. The pattern of conservation seen in these smaller alignments is fully representative of what was observed in the full alignments. The identification of AbpSH3 binding sites in higher eukaryotic genomes was carried out using the sequence motif defining and searching programs, MEME and MAST (32, 33).

RESULTS

Structural Studies of AbpSH3·peptide Complex—To determine the molecular mechanism of peptide binding by the AbpSH3, we determined its solution structure bound to a 17-residue peptide from Ark1p, which we refer to as ArkA. We previously showed that this is the tightest binding target peptide with proven biological relevance (20). The structure of AbpSH3·ArkA complex was constrained using intra-SH3 domain, intra-peptide, and intermolecular NOEs and dipolar couplings (statistics are found in supplemental Table S1). The twenty lowest energy structures overlay with backbone root mean square deviation of 0.37 ± 0.2 Å (supplemental Fig. S1). Ramachandran statistics are not as good as might be expected considering the large number of restraints on the SH3 domain backbone. This may be due to conformational fluctuations observed in the bound state SH3 domain. This was apparent for residues Lys(−3), Tyr-10, and Phe-50 based on peak broadening and for Asp-15 based on CPMG experiments performed in excess of peptide (data not shown).

The AbpSH3 displays a typical SH3 domain fold, consisting of a five-stranded β-sandwich with the long irregularly structured RT-loop between strands 1 and 2 (Fig. 1*A*). Superposition of all backbone residues to the x-ray structure of the unliganded form of AbpSH3 (pdb code 1JO8) (17) demonstrated a good fit (backbone root mean square deviation of 0.85 ± 0.04 Å). The regions with the highest root mean square deviation were the N and C termini as well as the RT-loop region with up to 3 ± 0.5 Å differences at the tip of the loop. The significant difference in the conformation of the RT-loop may be the result of a rearrangement occurring upon peptide binding, because this region is directly involved in the binding reaction. This is consistent with several other studies that have shown a role for RT-loop flexibility in binding (34, 35). The ArkA peptide, which binds in a class II orientation (2, 3), contains an N-terminal left-handed polyproline type II (PPII) helix and a C-terminal 3_{10} helix (Fig. 1*A*).

There are significant numbers of intermolecular NOEs throughout the binding interface (17 peptide residues spanning from Lys^{6} to Lys^{-8}) indicating that the complete peptide is involved in binding (we have used a standard numbering system (3) for peptide positions) (Fig. 1*B*). An overlay of the ArkA peptide conformers from multiple low energy structures shows that the backbone and side-chain atoms overlay well across 13 residues from position 3 to −9 (Fig. 1*E*). Even at peptide position Lys^{6} where the peptide structure is more variable, parts of this residue are still packed against the domain in every structure (Fig. 1*C*). The P*XX*P region of the peptide interacts in a typical manner for SH3 domains with P(2), P(0), and P(−1) well packed into two grooves lined primarily by the conserved residues Tyr^{8}, Tyr^{10}, Trp^{36}, Tyr^{54}, Pro^{51}, and Asn^{53} (36) (Fig. 1, *A* and *B*). We refer to this conserved PPII helix-interaction surface as surface I (37). Peptide residue Lys^{3} is also packed tightly against the domain, and its side chain lies near both Asp^{9} and Asp^{11} where electrostatic interactions may play a stabilizing role. Although it lies in a less well defined region of peptide structure, the side chain of peptide residue Lys^{6} invariably points toward the domain positioned near residues Asp^{9} and Glu^{22} (Fig. 1*A*). Thus, most N-terminal residues in the peptide interact with the AbpSH3.

Residues in the C-terminal extended region of the ArkA peptide (positions −4 to −9) make multiple contacts with residues in a surface that is distinct from the P*XX*P binding interface, which we refer to as surface II (37). Most notably, Leu^{-7} is highly buried, packing against Val^{32}, Asp^{33}, Trp^{36}, and Leu^{49} on the domain, and Pro^{-4} and His^{-6} on the peptide (Fig. 1*D*). His^{-6} is also able to form an intermolecular salt bridge with Asp^{33}. Interestingly, peptide residues extending out as far as Lys^{-8} and Pro^{-9} make close contacts with the domain at RT-loop residues Glu^{14} and Asp^{15}, respectively. Finally, as seen in other structures (37), peptide residue Lys^{-3} interacts at the dividing line between surface I and II, packing against Trp^{36} and making electrostatic interactions with the RT-loop residues, Glu^{14} and Asp^{15}. Surprisingly, the C_{ϵ} and N_{Z} atoms at the end of the Lys^{-3} side chain are completely broadened, indicating substantial dynamics even though this residue is crucial for the

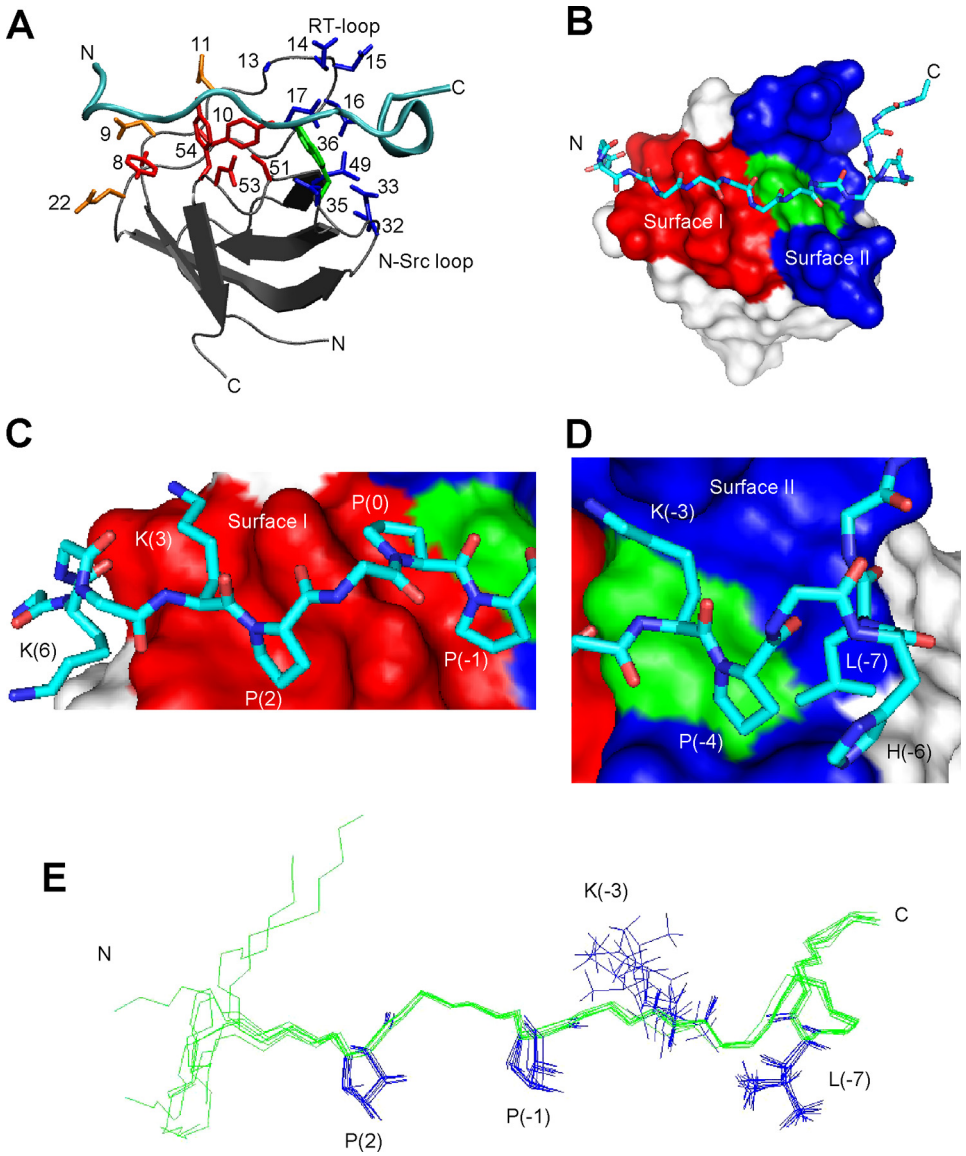


FIGURE 1. **The structure of the AbpSH3·ArkA peptide complex.** *A*, schematic of the AbpSH3·ArkA complex. The SH3 domain is in *silver* with the peptide in *cyan*. Side chains of conserved surface I residues within yeast SH3 domains (Y8, Y10, P51, N53, Y54) are shown in *red*, key surface II residues for AbpSH3 (Ala^{13}, Glu^{14}, Asp^{15}, Asn^{16}, Glu^{17}, Val^{32}, Asp^{33}, Asp^{35}, and Leu^{49}) are shown in *blue*, Trp^{36}, which is at the border of surfaces I and II is shown in *green*. Asp^{9}, Asp^{11}, and Glu^{22}, which interact with the N terminus of the peptide, are in *orange*. *B*, surface representation of the AbpSH3 highlighting surface I (*red*) and surface II (*blue*). The surface of Trp^{36} is indicated in *green*. The ArkA peptide is represented in *stick*, with nitrogen in *dark blue*, carbon in *light blue*, and oxygen in *red*. *C*, interactions of the PPII-helical region of ArkA with binding surface I. The positions of key peptide side chains are indicated. *D*, interactions of the extended region of the ArkA peptide with surface II. *E*, overlay of eight low energy ArkA peptide conformers.

binding reaction, explaining the undefined position of its side chain (Fig. 1*E*).

In Vitro Analysis of the AbpSH3 Interaction with Target Peptides—To identify the thermodynamically critical interactions in the AbpSH3-peptide interface, and evaluate the affinity requirements for *in vivo* function, isothermal titration calorimetry (ITC) was used to quantitate the binding affinities of wild-type (WT) and mutant AbpSH3-binding peptides. Initially, affinities for 17-residue long peptides from the three different biologically relevant Abp1p-interacting proteins, Ark1p, Scp1p, and Sjl2, were measured. Because Ark1p possesses tandem binding sites for the AbpSH3, which we call ArkA and ArkB, we tested these sites separately. Although the AbpSH3 bound the ArkA and Scp1 sites very tightly with K_d values of 0.27 and 0.7 μM, respectively, the ArkB and Sjl2 sites bound much more weakly with K_d values of ~6 μM (Fig. 2*A*). Thus, the biologically relevant interactions of this domain are seen to vary over a 20-fold range in affinity. Interestingly, combining the ArkA and ArkB sites into one site as they are found in Ark1p resulted in no significant increase in affinity over ArkA. It is also notable that the Scp1 and Sjl2 sites lack a P*XX*P motif, and yet were still able to bind with high affinities compared with the majority of SH3 domain-peptide interactions. To ensure that the Sjl2, Scp1, and ArkB peptides bound in the same manner as the ArkA peptide investigated in our NMR studies, we assigned the NMR resonances of AbpSH3 in complex with these peptides, and calculated the difference in the chemical shifts of each backbone amide resonance in the free and complexed state. As seen in Fig. 3, the overall pattern of chemical shift changes was very similar for the Sjl2 and ArkA peptides despite their very different amino acid sequences (only 6 identical residues out of 17), indicating that the Sjl2 peptide binds the same surface and utilizes the same binding mechanism as the ArkA peptide. A similar result was obtained when the other target peptides were tested in this manner (supplemental Fig. S2).

The importance of residues at the N and C termini of each target peptide was demonstrated by the 3- to 6-fold reductions in binding affinity that were observed when shortened (12 residue) peptides were tested (Fig. 2*B*). Furthermore, the significant contribution of interactions at both extremities of the ArkA peptide was shown by the 3-fold reduction in affinity resulting from deletion of Pro^{-9} and Lys^{-10} (Fig. 2*B*, ArkA 15-mer), and the further 2-fold reduction occurring when the three N-terminal residues were removed (ArkA, 12-mer). The contributions of these regions are consistent with our structure, in which well defined interactions were seen involving the Lys^{6} and Pro^{-9} positions. Versions of the ArkA peptide longer than 17 residues did not bind more tightly than the 17-mer peptide, as is demonstrated in reactions with a 20-mer ArkA site and with the ArkAB site (Fig. 2*A*).

A

Peptide	Kd (μM)	6	5	4	3	2	1	0	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10
	Class II core consensus :					P	x	x	P	x	+	"Extended Region"						
Scp	0.70 ± 0.15	K	K	P	R	P	P	V	K	S	K	P	K	H	L	Q	D	G
SjI	6.10 ± 1.30	K	P	E	K	P	P	V	V	K	K	P	H	Y	L	S	V	A
Prk	0.21 ± 0.20	K	S	R	P	P	R	P	P	P	K	P	L	H	L	R	T	E
ArkB	6.10 ± 1.05	S	H	L	K	P	K	P	P	P	K	P	L	L	L	A	G	R
ArkA	0.27 ± 0.08	K	K	T	K	P	T	P	P	P	K	P	S	H	L	K	P	K
ArkAB	0.23 ± 0.10	K	K	T	K	P	T	P	P	P	K	P	S	H	L	K	P	K
ArkA(20mer)	0.31 ± 0.02	K	K	T	K	P	T	P	P	P	K	P	S	H	L	K	P	K

P P P K P L L L A G R

P P P

B

						P	x	x	P	x	+	"Extended Region"						
Scp(12mer)	3.80 ± 1.14				R	P	P	V	K	S	K	P	K	H	L	Q		
SjI(12mer)	27.05 ± 4.02				K	P	P	V	V	K	K	P	H	Y	L	S		
ArkA(12mer)	1.70 ± 0.61				K	P	T	P	P	P	K	P	S	H	L	K		
ArkA(15mer)	0.77 ± 0.20	K	K	T	K	P	T	P	P	P	K	P	S	H	L	K		

C

						P	x	x	P	x	+	"Extended Region"						
K(3)A	0.39 ± 0.11	K	K	T	A	P	T	P	P	P	K	P	S	H	L	K	P	K
P(2)A	7.50 ± 1.02	K	K	T	K	A	T	P	P	P	K	P	S	H	L	K	P	K
P(2)V	2.95 ± 0.51	K	K	T	K	V	T	P	P	P	K	P	S	H	L	K	P	K
P(0)A	0.96 ± 0.17	K	K	T	K	P	T	A	P	P	K	P	S	H	L	K	P	K
P(-1)A	0.81 ± 0.18	K	K	T	K	P	T	P	A	P	K	P	S	H	L	K	P	K
K(-3)A	31.10 ± 2.56	K	K	T	K	P	T	P	P	P	A	P	S	H	L	K	P	K
K(3)R	6.20 ± 1.0	K	K	T	K	P	T	P	P	P	R	P	S	H	L	K	P	K
P(-4)A	1.75 ± 0.30	K	K	T	K	P	T	P	P	P	K	A	S	H	L	K	P	K
H(-6)A	2.00 ± 0.63	K	K	T	K	P	T	P	P	P	K	P	S	A	L	K	P	K
L(-7)A	6.00± 1.82	K	K	T	K	P	T	P	P	P	K	P	S	H	A	K	P	K

FIGURE 2. ***In vitro* binding analysis of AbpSH3 interactions.** *A*, binding of the AbpSH3 to biologically relevant extended peptides. Conserved positions in AbpSH3 binding sites are highlighted in *gray*. Above the sequences the standard nomenclature for class II SH3 domain target sequences is defined as well as the extended region of the peptide. The residues *boxed* in the ArkAB sequence indicate the sequence deleted for the ΔArkB sequence used in Fig. 4. *B*, binding of the AbpSH3 to truncated peptide targets. *C*, binding of the AbpSH3 to mutant ArkA sites. The mutated sites are *boxed*.

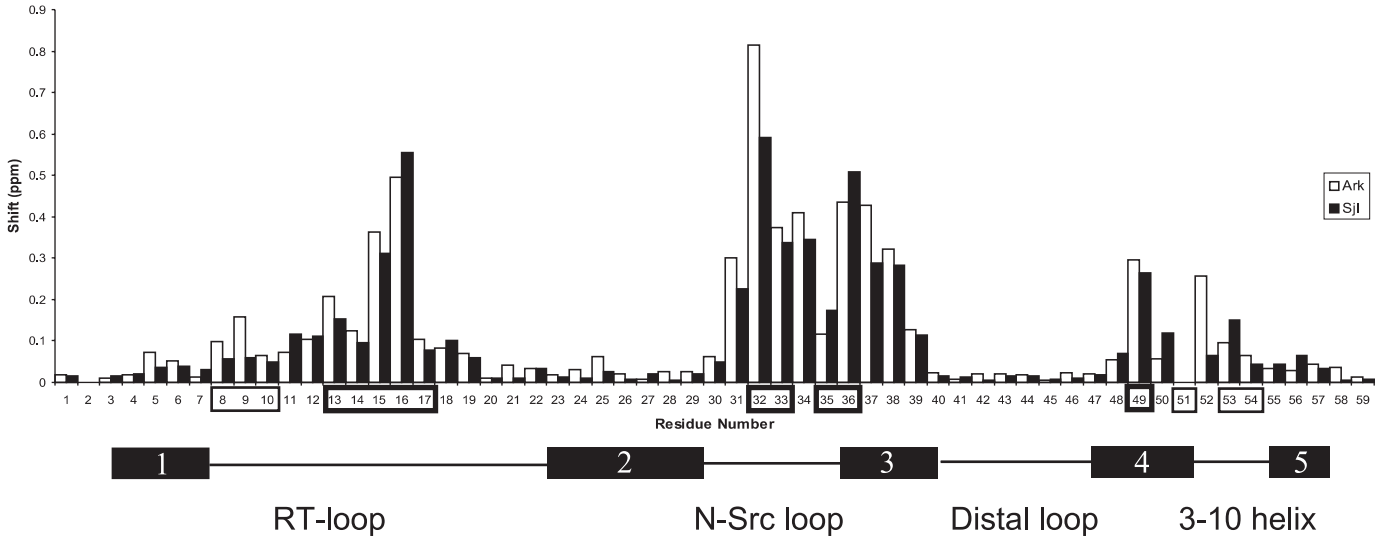


FIGURE 3. **Backbone amide chemical shift differences between the free and bound AbpSH3 complexes.** The ArkA target complex is in *white* and SjI2 target complex is in *black*. Surface I and surface II residues are *boxed* in *thin* and *thick lines*, respectively. A schematic of the secondary structure of the SH3 domain is shown at the *bottom*. Residues 2 and 51 are prolines and have no values in this plot.

To assess the individual binding contributions of ArkA peptide residues, single amino acid substitutions were constructed at eight positions that contacted the AbpSH3 in our structure. The three substitutions causing the most dramatic reductions in binding affinity (*i.e.* 20- to 100-fold decreases in affinity), $Pro^{2} \rightarrow Ala$, $Lys^{-3} \rightarrow Ala$, and $Leu^{-7} \rightarrow Ala$, were distributed across the whole binding surface, which emphasizes the importance of the extended target peptides in stabilizing interactions with the AbpSH3 (Fig. 2*C*). On the other hand, the only 3-fold decrease in affinity observed for the $Pro^{-1} \rightarrow Ala$ substitution confirmed that the presence of a P*XX*P motif in AbpSH3 binding sites is not critical for strong binding. While substitutions

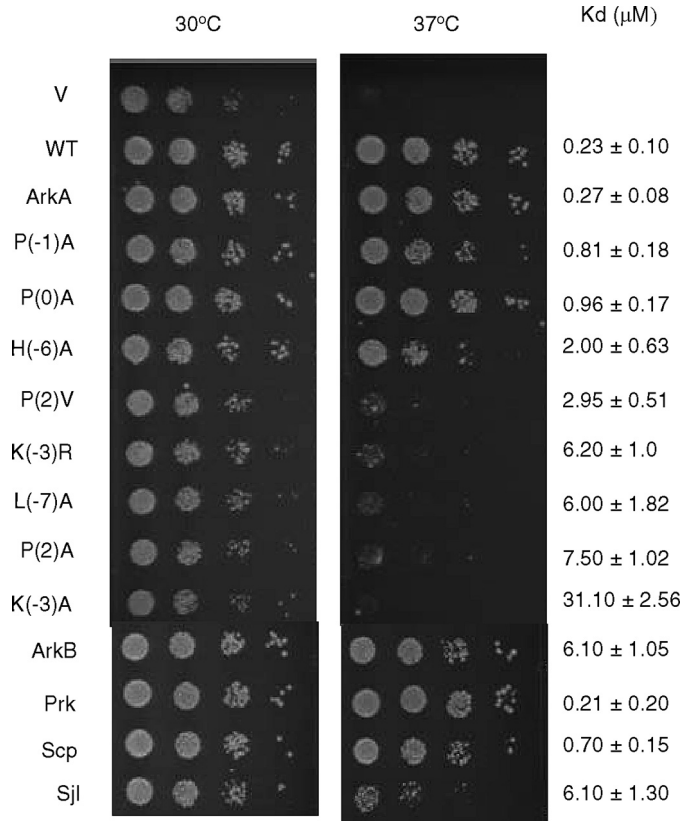


FIGURE 4. **Growth assays of cells carrying Ark1p with WT or mutant AbpSH3 binding sites.** The test strain was deleted for Prk1p and possessed an AbpSH3 binding site deletion in its chromosomal copy of Ark1p. Ark1p bearing the indicated AbpSH3 binding sites was expressed from a plasmid. Growth at 37 °C was dependent upon the interaction of the AbpSH3 with the plasmid-expressed Ark1p. Serial 10-fold dilutions of each culture were spotted on the plate. *ArkA* refers to the 17-mer ArkA.

that had the largest effect on binding were at positions conserved in all the targets studied here, substitution of the non-conserved His^{-6} position also produced a marked 7-fold reduction in affinity, which was similar in magnitude to the reduction resulting from substitution of the conserved residue, Pro^{-4}. Surprisingly, substitution of Lys^{3}, which appears to make significant packing and electrostatic interactions with the domain, produced only a small change in affinity.

In Vivo Activity of AbpSH3-interacting Peptides—To evaluate the relevance of the extended peptide binding interface of the AbpSH3 to *in vivo* function, we developed a system to test the biological activity of AbpSH3-interacting peptides. DNA encoding WT and mutant AbpSH3 target peptides was incorporated into a low copy number plasmid that expresses full-length Ark1p under the control of its native promoter. The target peptides were incorporated into Ark1p in such a way that its tandem AbpSH3 binding site (*i.e.* ArkA plus ArkB) was precisely replaced by a 17-residue site to be tested. These plasmids were transformed into a mutant yeast strain (*prk1*Δ*/ark1-*Δ*PP*) in which Ark1p binding to Abp1p is required for cell growth at 37 °C (20). It can be seen in Fig. 4 that a version of Ark1p possessing only the 17-mer ArkA binding site is able to complement the mutant strain as well as WT Ark1p. On the other hand, Ark1p carrying the very weak binding $Lys^{-3} \rightarrow$ Ala mutant ArkA site mediates cell growth that is as poor as the empty vector negative control (Fig. 4). These data show that, under these conditions, cell growth requires a robustly functioning AbpSH3 binding site within Ark1p, and also that the tandem binding site found within WT Ark1p is not necessary.

Assays of the single residue substitutions in the ArkA site demonstrated a clear correlation between *in vitro* binding affinity and *in vivo* function. Mutants displaying K_d values of 3 μM (*e.g.* the $Pro^{2} \rightarrow$ Val mutant) or greater displayed little or no growth at 37 °C, whereas those with K_d values of 1 or less grew as well as WT. The H6A mutant with a K_d value of 2 μM showed a marked reduction in growth at 37 °C, but grew considerably better that weaker binding mutants (Fig. 4). These results demonstrate that relatively tight binding (*i.e.* 1 μM) is required for optimal function of the ArkA site and that the biological function of this site is sensitive to small changes in affinity. These data also emphasize that residues at both extremes of the ArkA site (*i.e.* from Pro^{2} to Leu^{-7}) are crucial for *in vivo* function. It should be noted that we previously found that an AbpSH3 mutant with a K_d value of only 15 μM for the ArkA peptide was still able to mediate growth (20), but here target sites with affinities of between 2 and 3 μM did not mediate robust growth. This discrepancy is most likely due to the different growth conditions used in the two studies.

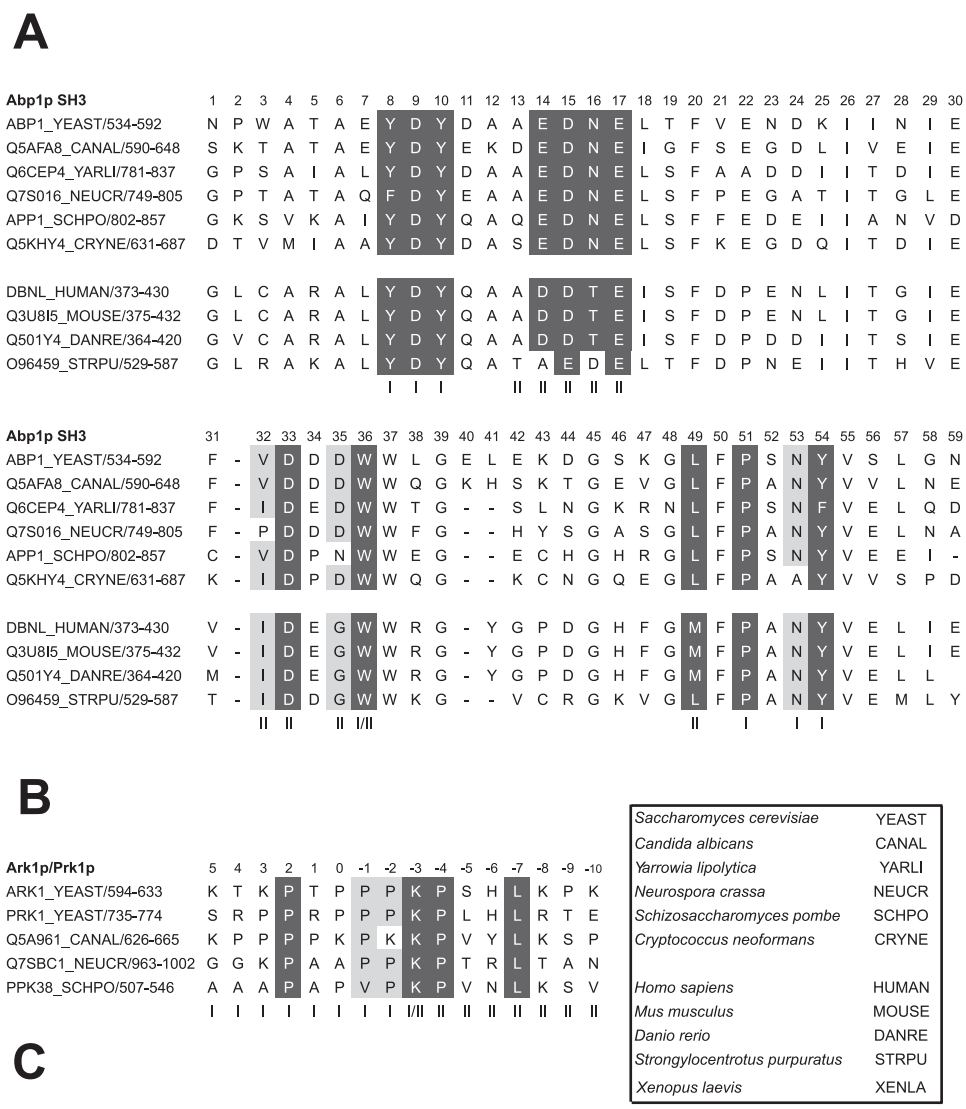
We also investigated the biological activity of the 17-residue Scp1, Sjl2, and ArkB binding sites. The Scp1 site mediated a WT level of growth, which was expected, because this site binds with a K_d of <1 (Fig. 4). Surprisingly, however, the ArkB site with a K_d value of 6 μM was able mediate a WT level of cell growth. Although cells carrying the Sjl2 site (K_d = of 6.1 μM) grew poorly compared with WT, their growth was more robust than ArkA mutants with the same affinity (*e.g.* Lys^{-3} Arg and Leu^{-7} Ala) (Fig. 4). Thus, the level of growth observed for strains carrying the ArkB and Sjl2 sites cannot be predicted using the correlations found for the ArkA mutant series, implying that other factors must be at play for these sites.

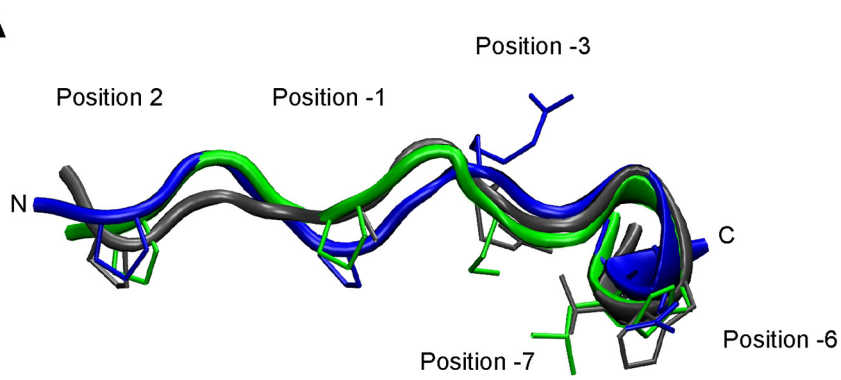
Conservation of the AbpSH3 Binding Interface and Binding Sites—To further investigate the importance of the extended AbpSH3:peptide interface, we analyzed its evolutionary conservation. In Fig. 5*A* where AbpSH3 sequences from diverse fungal species were aligned, it can be seen that 15 of the 16 positions corresponding to residues comprising the peptide binding interface are highly conserved (Fig. 5*A*). This high degree of conservation across both binding surfaces I and II implies that these homologues should all recognize similar target sequences. To test this idea, we aligned fungal homologues of AbpSH3 target proteins and searched for conserved binding sites. In most Sjl2p homologues, we found a conserved AbpSH3 binding site embedded within a highly diverged region of these proteins (Fig. 5*C*). Remarkably, homologous binding sites were found even in the distantly related species, *Cryptococcus neoformans* and *Schizosaccharomyces pombe*, which are estimated to have diverged from yeast 500 million to 1 billion years ago (38). Homologues of Ark1p also clearly possessed AbpSH3 binding sites (Fig. 5*B*). Conserved sites were also found in Scp1p homologues, but the species distribution of these sites was not as wide (data not shown). The binding specificity of the AbpSH3 is conserved not only in fungi, but also in Abp1p homologues in higher eukaryotes. Fig. 5*A* shows that AbpSH3

sequences from several organisms, including humans, display a high degree of sequence similarity to the fungal sequences, particularly in binding surface residues. Furthermore, we identified conserved sites closely resembling AbpSH3 binding sites in two higher eukaryotic proteins, the adenomatous polyposis coli protein (Apc) and the faciogenital dysplasia protein (Fgd1) (Fig. 5, *D* and *E*). It has been experimentally demonstrated that mouse Fgd1 does directly interact with mouse Abp1 both *in vitro* and *in vivo* (23), and this interaction is mediated through the binding site identified by our computational search. Although Apc has not yet been demonstrated to interact with Abp1, it does bind directly to actin and localizes to actin structures within cells (39); thus, a direct interaction with Abp1, which localizes to the same regions, could be expected. In summary, our sequence conservation data indicate that the target-specific binding interfaces of the AbpSH3 are conserved from yeast to humans.

DISCUSSION

Although several other SH3 domains have been shown to bind to extended peptides in a manner similar to the AbpSH3, our combined structural, functional, and bioinformatic analysis has provided a compelling argument for the biological importance of an extended target peptide interaction. Our NMR structure showed that the AbpSH3 is in contact with a region spanning 17 residues of its target peptide, and deletion of residues at either end of this extended region affected binding affinity (Fig. 2*B*). Furthermore, two of the three key residues for mediating binding (Pro^{2} and Leu^{-7}) are near the N and C termini of the target peptide, separated by eight residues, and these same two residues are also critical for *in vivo* function (Fig. 4). Finally, we have shown that the sequence of AbpSH3 binding sites is conserved in fungal Abp1p binding partner homologues across hundreds of millions of years





B

SH3 domain	Peptide	Kd (μM)	Kd (μM) (ΔC-term)	Surface I 3	2	1	0	-1	-2	-3	-4	Surface II -5	-6	-7	-8	-9	-10	-11
Abp1p	ARK	0.27	15	K	P	T	P	P	P	K	P	S	H	L	K			
BetaPix	PAK2	1.05	ND	P	P	V	I	A	P	R	P	E	H	T	K			
Csk	PEP	0.8	no binding	P	P	P	L	P	E	R	T	P	E	S	F			
Lyn	TIP	0.8	15.8	T	P	P	L	P	P	R	P	A	N	L	G			
Gads	HPK1	2.4	no binding	P	P	L	V	P	P	R	K	E	K	M	R			

FIGURE 6. **Similarity of the ArkA peptide conformation with other extended peptides bound to SH3 domains.** *A*, structural alignment of the central 12 residues from of bound ArkA (*green*, this study), with bound PAK2 (*gray*, root mean square deviation = 1.4 Å, pdb: 2DF6) and PEP (*blue*, root mean square deviation = 1.9 Å, pdb: 1JEG). The side chains at key positions, 2, −1, −3, −6, and −7 are shown. *B*, alignment of the peptide sequences shown in *A*, including two extra peptides found in other complexes. The helical element is *boxed*, the surface I P*XX*P motif is highlighted in *red*, and the key surface II contact residues within the helical element are highlighted in *blue*. K_d values are shown for the WT peptides and peptides where the *boxed* residues are mutated or deleted (ΔC-term). *ND*, not determined.

of evolution, and that even the human Abp1 SH3 domain recognizes a similar binding sequence (Fig. 5). The extreme conservation of the AbpSH3 binding can only be explained by concluding that the biological function of this domain requires the high degree of affinity and/or specificity that is supplied by its extended site. Our work implies that the "promiscuous model," in which SH3 domains bind many different targets with relatively low affinity and specificity (6–10), does not apply to the AbpSH3. Consistent with this conclusion, a previous *in vitro* study showed that the ArkA peptide is highly specific for the AbpSH3 and does not interact at a significant level with any other yeast SH3 domain (25).

The structure of the AbpSH3·ArkA peptide complex bears significant similarities to other structures of SH3 domains bound to extended peptides (12–15, 40). In particular, the extended region of the ArkA peptide (*i.e.* those residues C-terminal to position −3), contacts the same residues in the RT-loop, N-Src loop, and β-strand *d* as were observed to contact extended peptide regions in these other structures. The surface formed by these residues, which we refer to as surface II (Fig. 1*D*), is much broader and flatter than the P*XX*P binding surface (surface I), experiences more conformational change on peptide binding, and can, thus, accommodate a variety of peptide conformations. This feature and the non-conserved nature of residues lying on this surface provide an explanation for the highly divergent specificities of SH3 domains for peptide regions extending beyond the P*XX*P region (37). Interestingly, the conformation of the ArkA peptide in our structure is similar to that seen for two other SH3-bound extended peptides in that each peptide displays a short α- or 3_{10}-helical segment following the region in PPII conformation (Fig. 6*A*). Furthermore, each of these peptides (along with two other similar peptides) possesses a Pro at the N terminus of the helical region that may help to stabilize this conformation. Each peptide also relies on its C-terminal extension for high affinity binding (Fig. 6*B*). The helical elements observed in these peptide structures provide an efficient means to produce the sharp backbone turn that is required for simultaneous interaction with both surface I and II on the SH3 domain, and we expect that this structural motif will occur in many other complexes of SH3 domains with extended peptides.

In agreement with our previous studies of amino acid substitutions in the AbpSH3 on the yeast Sho1p SH3 domain (20, 41), our experiments with ArkA peptide mutants clearly demonstrated a good correlation between the measured binding affinity of peptides *in vitro* and their ability to function *in vivo* with all mutants with K_d values greater than 3 μM showing little growth above background (Fig. 4). Unexpectedly, however, the ArkB peptide, with a K_d value of 6 μM, functioned as well as the WT ArkA peptide *in vivo*, and the Sjl2 site also functioned somewhat better than mutant ArkA sites with the same affinity. It appears that peptide-based K_d values may not always capture all the relevant properties that determine functional ability *in vivo*. Although we are currently unable to explain these data, our results suggest a high sensitivity of the AbpSH3 to the precise sequence of its targets, and indicate that the information content of these sequences may be greater than previously

FIGURE 5. **Sequence alignments of the AbpSH3 and its conserved target sequences.** *A*, an alignment of the SH3 domains from Abp1p homologues found in a divergent group of fungal species. Conserved residues in the surface I and surface II peptide binding interfaces are shaded in *gray*. In this alignment, positions with >83% conservation are *shaded* in *dark gray*, and positions with 75–83% conservation are *shaded light gray*. Residues are considered equivalent for the following groups: (A, I, L, V, M), (F, Y), (D, E), (N, Q), (R, K, H), and (S, T). A standard SH3 domain numbering system is used (36). Below the sequences from fungal homologues, the sequences of homologues of the AbpSH3 domain from higher eukaryotes are shown. Surface I positions (indicated by "*I*") were defined as those residues with side chains close (within 5 Å or less) to the PPII-helical region of ArkA (residues Pro^{2} to Pro^{-2}), surface II residues (indicated by "*II*") are close to the ArkA C-terminal extension (residues Pro^{-4} to Lys^{-10}). Residues on the dividing line between surface I and surface II (*i.e.* they contact both the PPII and C-terminal extended region of the peptide) are indicated by "*I/II*." *B*, an alignment of conserved Abp1p binding sites in Ark1/Prk1 homologues. The standard peptide numbering is shown (3). *C*, an alignment of conserved Abp1p binding sites in Sjl2 homologues. Sequences flanking the sites on either side are shown to emphasize the general lack of conservation observed in this region of these proteins. Also shown are alignments of conserved AbpSH3 binding sites in homologues of the higher eukaryotic proteins Apc (*D*), and Fgd1 (*E*).

appreciated. It is also interesting that the Sjl2 site, which is biologically relevant and highly conserved, binds to the AbpSH3 20-fold more weakly than the ArkA peptide, and does not function well *in vivo* within the context of Ark1p. The variation in affinity requirements for sites in different AbpSH3 target proteins may be due to different target protein concentrations within the cell, or to differences in the functions of the complexes in which various target proteins reside.

It is not currently known whether most SH3 domains bind promiscuously to short peptides, or require higher specificity interactions as does the AbpSH3. However, examples of SH3 domains binding to extended peptides or to other folded domains are steadily growing in number (37), and we suspect that SH3 domains and other protein-protein interaction modules may generally possess more binding specificity than is currently appreciated. A deficiency in many studies addressing SH3 domain specificity is that they focus primarily on P*XX*P binding and have investigated peptides that are too short to encompass the extended binding target for a given domain, thus missing the important binding contributions of surface II. For example, the large scale phage display study carried out by Tong *et al.* (24) on yeast SH3 domains used only 9-mer target sequences, and the consensus binding sites from this study provided the basis for other global yeast SH3 domain studies (8, 42). In one case, this process led to the characterization of AbpSH3 binding peptides lacking residues beyond the −4 peptide position, which resulted in measured K_d values for target peptides from Ark1p and Scp1 of 24 and 14 μM, respectively, and binding to the biologically relevant peptide from Sjl2p was undetectable (8). Because we have shown that the Leu at the −7 peptide position is crucial for high affinity binding and biological activity, the characterization of Abp1p peptides that do not extend to this position would necessarily be misleading. This situation demonstrates the confusion that can be caused by the bipartite nature of the SH3 domain binding interface, and the difficulty in predicting the interactions mediated by surface II (37). Furthermore, it is difficult to make conclusions about the importance of affinity and *in vivo* function in SH3 domains unless interactions with biologically relevant target peptides of sufficient length are being investigated.

In summary, our experimental and bioinformatic results demonstrate that the AbpSH3 interacts with functionally important extended conserved target binding sites, implying a high degree of specificity. Given that the extended interface characterized in this study is structurally similar to several other SH3 domain-peptide complexes found in higher eukaryotes (12, 14, 15, 43, 44), we believe these conclusions apply to human SH3 domain function as well. Future studies such as ours, incorporating detailed structural, functional, and evolutionary analyses, may illuminate the importance of extended interactions in other protein interaction modules.

Acknowledgments—We thank Ranjith Muhandiram, Lewis Kay, and the members of the Kay group for assistance in setting up NMR experiments, Tony Mittermaier for initial NMR measurements on the system, and Frank Sicheri and Jun Lui for use of their ITC machines.

REFERENCES

1. Pawson, T., and Nash, P. (2003) *Science* **300,** 445–452
2. Feng, S., Chen, J. K., Yu, H., Simon, J. A., and Schreiber, S. L. (1994) *Science* **266,** 1241–1247
3. Lim, W. A., Richards, F. M., and Fox, R. O. (1994) *Nature* **372,** 375–379
4. Pawson, T. (1995) *Nature* **373,** 573–580
5. Jia, C. Y., Nie, J., Wu, C., Li, C., and Li, S. S. (2005) *Mol. Cell Proteomics* **4,** 1155–1166
6. Castagnoli, L., Costantini, A., Dall'Armi, C., Gonfloni, S., Montecchi-Palazzi, L., Panni, S., Paoluzi, S., Santonico, E., and Cesareni, G. (2004) *FEBS Lett.* **567,** 74–79
7. Ladbury, J. E., and Arold, S. (2000) *Chem. Biol.* **7,** R3–R8
8. Landgraf, C., Panni, S., Montecchi-Palazzi, L., Castagnoli, L., Schneider-Mergener, J., Volkmer-Engert, R., and Cesareni, G. (2004) *PLoS. Biol.* **2,** E14
9. Li, S. S. (2005) *Biochem. J.* **390,** 641–653
10. Mayer, B. J. (2001) *J. Cell Sci.* **114,** 1253–1263
11. Agrawal, V., and Kishan, K. V. (2002) *Protein Pept. Lett.* **9,** 185–193
12. Bauer, F., Schweimer, K., Meiselbach, H., Hoffmann, S., Rösch, P., and Sticht, H. (2005) *Protein Sci.* **14,** 2487–2498
13. Berry, D. M., Nash, P., Liu, S. K., Pawson, T., and McGlade, C. J. (2002) *Curr. Biol.* **12,** 1336–1341
14. Ghose, R., Shekhtman, A., Goger, M. J., Ji, H., and Cowburn, D. (2001) *Nat. Struct. Biol.* **8,** 998–1004
15. Kami, K., Takeya, R., Sumimoto, H., and Kohda, D. (2002) *EMBO J.* **21,** 4268–4276
16. Kaksonen, M., Toret, C. P., and Drubin, D. G. (2005) *Cell* **123,** 305–320
17. Fazi, B., Cope, M. J., Douangamath, A., Ferracuti, S., Schirwitz, K., Zucconi, A., Drubin, D. G., Wilmanns, M., Cesareni, G., and Castagnoli, L. (2002) *J. Biol. Chem.* **277,** 5290–5298
18. Sekiya-Kawasaki, M., Groen, A. C., Cope, M. J., Kaksonen, M., Watson, H. A., Zhang, C., Shokat, K. M., Wendland, B., McDonald, K. L., McCaffery, J. M., and Drubin, D. G. (2003) *J. Cell Biol.* **162,** 765–772
19. Stefan, C. J., Padilla, S. M., Audhya, A., and Emr, S. D. (2005) *Mol. Cell. Biol.* **25,** 2910–2923
20. Haynes, J., Garcia, B., Stollar, E. J., Rath, A., Andrews, B. J., and Davidson, A. R. (2007) *Genetics* **176,** 193–208
21. Lila, T., and Drubin, D. G. (1997) *Mol. Biol. Cell* **8,** 367–385
22. Engqvist-Goldstein, A. E., and Drubin, D. G. (2003) *Annu. Rev. Cell Dev. Biol.* **19,** 287–332
23. Hou, P., Estrada, L., Kinley, A. W., Parsons, J. T., Vojtek, A. B., and Gorski, J. L. (2003) *Hum. Mol. Genet.* **12,** 1981–1993
24. Tong, A. H., Drees, B., Nardelli, G., Bader, G. D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S., Quondam, M., Zucconi, A., Hogue, C. W., Fields, S., Boone, C., and Cesareni, G. (2002) *Science* **295,** 321–324
25. Zarrinpar, A., Park, S. H., and Lim, W. A. (2003) *Nature* **426,** 676–680
26. Kanelis, V., Forman-Kay, J. D., and Kay, L. E. (2001) *IUBMB Life* **52,** 291–302
27. Mittermaier, A., and Kay, L. E. (2001) *J. Am. Chem. Soc.* **123,** 6892–6903
28. Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson.* **131,** 373–378
29. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
30. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
31. Guthrie, C., and Fink, G. R. (1991) *Guide to Yeast Genetics and Molecular Biology*, pp. 12–18, Academic Press, Inc., San Diego, CA
32. Bailey, T. L., and Gribskov, M. (1998) *Bioinformatics* **14,** 48–54
33. Bailey, T. L., Williams, N., Misleh, C., and Li, W. W. (2006) *Nucleic Acids Res.* **34,** W369–W373
34. Arold, S., O'Brien, R., Franken, P., Strub, M. P., Hoh, F., Dumas, C., and Ladbury, J. E. (1998) *Biochemistry* **37,** 14683–14691
35. Lee, C. H., Saksela, K., Mirza, U. A., Chait, B. T., and Kuriyan, J. (1996) *Cell* **85,** 931–942
36. Larson, S. M., Di Nardo, A. A., and Davidson, A. R. (2000) *J. Mol. Biol.* **303,** 433–446
37. Kim, J., Lee, C. D., Rath, A., and Davidson, A. R. (2008) *J. Mol. Biol.* **377,** 889–901

38. Taylor, J. W., and Berbee, M. L. (2006) *Mycologia* **98,** 838–849
39. Moseley, J. B., Bartolini, F., Okada, K., Wen, Y., Gundersen, G. G., and Goode, B. L. (2007) *J. Biol. Chem.* **282,** 12661–12668
40. Ogura, K., Nobuhisa, I., Yuzawa, S., Takeya, R., Torikai, S., Saikawa, K., Sumimoto, H., and Inagaki, F. (2006) *J. Biol. Chem.* **281,** 3660–3668
41. Marles, J. A., Dahesh, S., Haynes, J., Andrews, B. J., and Davidson, A. R. (2004) *Mol. Cell* **14,** 813–823
42. Beltrao, P., and Serrano, L. (2005) *PLoS. Comput. Biol.* **1,** e26
43. Lewitzky, M., Harkiolaki, M., Domart, M. C., Jones, E. Y., and Feller, S. M. (2004) *J. Biol. Chem.* **279,** 28724–28732
44. Mott, H. R., Nietlispach, D., Evetts, K. A., and Owen, D. (2005) *Biochemistry* **44,** 10977–10983
45. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410