Investigation of SH3 domain specificity through the study of cross-reactive domains from yeast proteins Nbp2p and Bem1p

by

Maryna Gorelik

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Molecular Biology
University of Toronto

© Copyright by Maryna Gorelik (2011)
Abstract

Investigation of SH3 domain specificity through the study of cross-reactive domains from yeast proteins Nbp2p and Bem1p

Maryna Gorelik
Graduate Department Molecular Genetics
University of Toronto

Protein interactions are often mediated through binding of protein interaction domains to short peptide motifs. However, how the specificity of interactions is regulated within domain families is not well understood. In this thesis, I investigated this issue through the study of two cross-reactive SH3 domains from the yeast adaptor proteins Nbp2p and Bem1p. Despite low sequence similarity and different cellular roles, the Nbp2p SH3 and Bem1p SH3b domains recognize targets containing the same consensus sequences, raising the question of how the specificity of these domains is controlled. The Bem1p SH3b domain is also unusual in containing a 40 amino acid long C-terminal extension required for folding and interaction with the Cdc42p GTPase, which I also investigated in this thesis.

I demonstrated that although Nbp2p SH3 and Bem1p SH3b domains recognize the same consensus sequence, each domain possesses its own unique specificity within the context of eight yeast peptides. Through structural and mutagenesis studies, I demonstrated that the Nbp2p SH3 and Bem1p SH3b domains employ different mechanism for binding, explaining how each domain can still possess its own unique specificity, while recognizing the same consensus sequence. I found that the specificity of Bem1 SH3b domain is largely regulated through inhibitory interactions, where a conserved residue in the binding interface serves to decrease binding to non-specific targets. Most significantly I observed that reducing the
specificity of Bem1 SH3b domain, without reducing its binding affinity decreases its ability to function \textit{in vivo}.

My structural and mutagenesis studies with the Bem1p SH3b domain also demonstrated that its C-terminal extension is involved in peptide binding surface and creates a separate surface required for interaction with Cdc42p. I also showed that the Bem1p SH3b domain is capable of simultaneously interacting with peptide targets and Cdc42p, suggesting a role for this domain in Cdc42p mediated activation.

Overall, my findings demonstrated the ability of SH3 domains to perform complex functions and bind their targets with finely tuned specificity. Most importantly, I demonstrated the significance of negative interactions for specificity determination and the importance of precise specificity for \textit{in vivo} function.
Acknowledgements

I would like to thank my supervisor Alan Davidson for his guidance, support, and letting me pursue interesting research angles, making my research enjoyable and exciting. I would also like to thank my supervisory committee members, Peter Roy and Brigitte Lavoie, for their insightful comments and advice over the years.

I would like to thank the people who helped me with my research: Ranjith Muhandiram, for doing NMR experiments and his explanation of basic NMR concepts; Ana Popovic, Karen Maxwell, and Lisa Pell for their help with analysis of NMR data; Summer students, Alex Tsai and Wen Chi Yin, for their contributions to the structural part of my project; Lastly, all past and present members of the Davidson Lab, for their help and for creating a positive work environment that made me happy to come to work every day.

Finally, I would like to thank my family. My parents, who raised me and instilled in me an interest in science from an early age. And my husband Terry, not only for his emotional support, but also for using his programming skills to help me with my research.

Thank you
Table of Contents

Abstract ............................................................................................................................. ii
Acknowledgements ........................................................................................................... iv
Table of Contents .............................................................................................................. v
List of Figures .................................................................................................................. viii
List of Tables ................................................................................................................... x
List of Abbreviations ....................................................................................................... xi
Chapter 1 : Introduction ..................................................................................................... 1
  1.1 Protein interaction domains ...................................................................................... 1
  1.2 SH3 domains ............................................................................................................. 2
    1.2.1 SH3 domain structure .................................................................................. 3
    1.2.2 Recognition of canonical PxxP targets by SH3 domains......................... 4
    1.2.3 Recognition of noncanonical targets by SH3 domains............................. 6
    1.2.4 Affinity and specificity .............................................................................. 8
    1.2.5 Specificity of SH3 domain mediated interactions .................................. 11
    1.2.6 Mechanisms through which SH3 domains can achieve higher specificity .. 13
    1.2.7 Yeast SH3 domains ................................................................................... 15
  1.3 Nbp2p and Bem1p ............................................................................................... 18
    1.3.1 Cdc42p .................................................................................................... 19
    1.3.2 Yeast p21 activated kinases: Ste20p, Cla4p and Skm1p ......................... 21
    1.3.3 Yeast MAPK pathways .......................................................................... 24
    1.3.4 Nbp2p ..................................................................................................... 26
    1.3.5 Bem1p .................................................................................................... 29
    1.3.6 Cross-reactivity between Nbp2p and Bem1p SH3b domains ................... 33
  1.4 Project overview ................................................................................................... 35
Chapter 2 : Structural and functional characterization of Bem1p SH3b domain ............ 37
  2.1 Overview .............................................................................................................. 37
  2.2 Introduction ......................................................................................................... 37
  2.3 Materials and Methods ...................................................................................... 39
    2.3.1 NMR spectroscopy and structure calculation ......................................... 39
    2.3.2 Protein expression and purification for in vitro binding assays ............... 40
    2.3.3 Fluorescence-based binding assays ......................................................... 40
    2.3.4 Pull-down binding assays ....................................................................... 41
  2.4 Results ................................................................................................................. 41
2.4.1 The structure of the BemSH3-CI domain from *L. elongisporus* in complex with a PXXP-containing peptide from Ste20p ................................................................. 41

2.4.2 Mutagenesis and quantitative binding studies to delineate residues important for the interaction of the BemSH3-CI domain and Cdc42p .................................................. 47

2.4.3 Cdc42p and Ste20 peptide binding sites on BemSH3-CI domain are distinct and are not allosteric ........................................................................................................... 49

2.5 Discussion .......................................................................................................................... 51

Chapter 3 : A Conserved Residue in the Yeast Bem1p SH3 Domain Maintains the High Level of Binding Specificity Required for Function ......................................................................... 58

3.1 Overview ................................................................................................................................ 58

3.2 Introduction .................................................................................................................................. 58

3.3 Experimental procedures ........................................................................................................... 59

3.3.1 Sample preparation for in vitro binding studies ..................................................................... 59

3.3.2 In vitro peptide-binding and SH3 domain stability assays .................................................. 60

3.3.3 SH3 domain binding assays in yeast extracts ....................................................................... 60

3.3.4 Yeast strains ......................................................................................................................... 61

3.3.5 Growth assays ...................................................................................................................... 61

3.4 Results ........................................................................................................................................ 63

3.4.1 The NbpSH3 and BemSH3b domains display finely tuned specificity ................................. 63

3.4.2 The conserved Lys14 residue of the BemSH3b domain is a key determinant of specificity ............................................................................................................................. 64

3.4.3 The difference in binding affinities between BemSH3b and NbpSH3 domains is biologically significant ........................................................................................................... 66

3.4.4 Altering the binding specificity of the BemSH3b domain impairs Bem1p function .......... 67

3.4.5 The K14A substitution causes an increase in non-specific binding .................................... 71

3.4.6 An NbpSH3 mutant also displays increased non-specific binding .................................... 74

3.5 Discussion ................................................................................................................................ 74

Chapter 4 : Distinct recognition of the Ste20 peptide by the Nbp2p SH3 and Bem1p SH3b domains: understanding the determinants of SH3 domain binding specificity .................................... 78

4.1 Overview ................................................................................................................................... 78

4.2 Introduction ................................................................................................................................ 78

4.3 Materials and Methods ............................................................................................................. 79

4.3.1 Protein expression, peptide binding assays and thermal stability assays ......................... 79

4.3.2 NMR spectroscopy and structure calculation .................................................................... 80

4.3 Results ...................................................................................................................................... 80
List of Figures

Figure 1.1 SH3 domain structure.................................................................4
Figure 1.2 Interaction of SH3 domains with canonical PxxP containing targets........6
Figure 1.3 Interactions of SH3 domains with noncanonical targets....................8
Figure 1.4 Mechanisms of SH3 domain specificity........................................12
Figure 1.5 An alignment of yeast SH3 domains......................................16
Figure 1.6 Cdc42p regulation and establishment of cell polarity....................21
Figure 1.7 Yeast p21 activated kinases: Ste20p, Cla4p and Skm1p....................23
Figure 1.8 Yeast MAPK pathways..............................................................25
Figure 1.9 Domain organization and function of Nbp2p. ............................27
Figure 1.10 Domain organization and function of Bem1p..............................31
Figure 1.11 Binding targets of the NbpSH3 and BemSH3b domains..............35
Figure 2.1 Heteronuclear \(^{1}H\)--\(^{15}N\) NOEs of the LeBemSH3-CI domain ....43
Figure 2.2 The structure of LeBemSH3-CI:Ste20 peptide complex .................44
Figure 2.3 An alignment of BemSH3-CI domain fungal homologues ..............45
Figure 2.4 Binding of the Ste20 peptide to the BemSH3-CI domain ...............46
Figure 2.5 Binding of Cdc42p to the BemSH3-CI domain............................48
Figure 2.6 Stability and peptide binding activity of LeBemSH3-CI mutants ......49
Figure 2.7 Simultaneous binding of the Ste20 peptide and Cdc42p by the BemSH3-CI domain........................................................................................................50
Figure 2.8 A model for BemSH3-CI domain mediated activation of Ste20p kinase. ....53
Figure 2.9 Comparison of the LeBemSH3-CI domain alternative binding surfaces with those of other SH3 domains..........................................................55
Figure 2.10 Alignment of Myo3p SH3 and Bud14p fungal homologues.............57
Figure 3.1 Tryptophan fluorescence binding assays......................................63
Figure 3.2 In vitro binding analysis of BemSH3b and NbpSH3 interactions .........65
Figure 3.3 Biological activity of Nbp2/BemSH3b hybrid proteins....................67
Figure 3.4 In vivo analysis of Bem1p function..............................................68
Figure 3.5 The effect of BemSH3b altered specificity mutants on viability .........69
Figure 3.6 BemSH3b KAD\(\rightarrow\)NDN substitution has no effect on Nbp2p function. ..70
Figure 3.7 BemSHb K14A and NbpSH3 H32A mutants display increase in non-specific binding.................................................................................................................................. 72
Figure 4.1 Mutational analysis of the Ste20 and Bck1 peptides, and identification of the N-term Cla4 and Boi1 sites displaying high selectivity for the BemSH3b domain. .................. 82
Figure 4.2 The structure of NbpSH3-Ste20 complex ................................................................. 84
Figure 4.3 Comparison of SH3-peptide interaction details between the yeast and
$L. elongisporus$ BemSH3b-Ste20 complexes.................................................................................... 86
Figure 4.4 Comparison of the NbpSH3-Ste20 and BemSH3b-Ste20 complexes................. 88
Figure 4.5 Mutational analysis of NbpSH3 and BemSH3b domains and identification of residues important for establishing NbpSH3 domain specificity ........................................... 93
List of Tables

Table 1.1 Abundance and binding preferences of the most wide-spread protein interaction domains ..................................................................................................................................2
Table 1.2 Yeast SH3 domains .............................................................................................................17
Table 1.3 Summary of NbpSH3 and BemSH3b interactions .................................................................34
Table 2.1 Structural statistics for LeBemSH3-CI:Ste20 peptide complex ..............................................42
Table 3.1 Yeast strains used in this study. ............................................................................................62
Table 3.2 Additional peptide binding affinities of the NbpSH3 and BemSH3b mutants ..............66
Table 3.3 TAP-tagged proteins tested for interaction with the BemSH3b domain ......................73
Table 4.1 Structural statistics for the NbpSH3-Ste20 peptide complex ............................................83
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAP</td>
<td>Adhesion and Degranulation promoting Adaptor Protein</td>
</tr>
<tr>
<td>BemSH3b</td>
<td>Bem1p SH3b</td>
</tr>
<tr>
<td>BemSH3-CI</td>
<td>Bem1p SH3-CI</td>
</tr>
<tr>
<td>BR</td>
<td>Basic Region</td>
</tr>
<tr>
<td>CD</td>
<td>Circular Dichroism</td>
</tr>
<tr>
<td>CI</td>
<td>Cdc42 Interacting</td>
</tr>
<tr>
<td>COSY</td>
<td>Correlation Spectroscopy</td>
</tr>
<tr>
<td>CRIB</td>
<td>Cdc42 and Rac Interactive Binding</td>
</tr>
<tr>
<td>CWI</td>
<td>Cell Wall Integrity</td>
</tr>
<tr>
<td>DI</td>
<td>Dimerization segment</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylene Diamine Tetraacetic Acid</td>
</tr>
<tr>
<td>GAP</td>
<td>GTPase Activating Protein</td>
</tr>
<tr>
<td>GEF</td>
<td>Guanine Exchange Factor</td>
</tr>
<tr>
<td>GDI</td>
<td>Guanine Dissociation Inhibitor</td>
</tr>
<tr>
<td>HOG</td>
<td>High Glycerol Osmolarity</td>
</tr>
<tr>
<td>IS</td>
<td>Inhibitory sequence</td>
</tr>
<tr>
<td>MANT-GTP</td>
<td>2'-(or-3')-O-(N-methylanthraniloyl) guanosine 5'-triphosphate</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen Activated Protein Kinase</td>
</tr>
<tr>
<td>MIA</td>
<td>Melanoma Inhibitory Activity</td>
</tr>
<tr>
<td>NbpSH3</td>
<td>Nbp2p SH3</td>
</tr>
<tr>
<td>Ni-NTA</td>
<td>Nickel Nitrilo-Triacetic Acid</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>NOE</td>
<td>Nuclear Overhauser Effect</td>
</tr>
<tr>
<td>NOESY</td>
<td>Nuclear Correlation Spectroscopy</td>
</tr>
<tr>
<td>PPII</td>
<td>Polyproline helix II</td>
</tr>
<tr>
<td>PAK</td>
<td>p21 Activated Kinase</td>
</tr>
<tr>
<td>PBD</td>
<td>p21 Binding Domain</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenyl Methane Sulfonyl Fluoride</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root Mean Square Deviation</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium Dodecyl Sulphate Polyacrylimide Electrophoresis</td>
</tr>
<tr>
<td>SD</td>
<td>Synthetic defined</td>
</tr>
<tr>
<td>SH3</td>
<td>Src Homology 3</td>
</tr>
<tr>
<td>SH2</td>
<td>Src Homology 2</td>
</tr>
<tr>
<td>TOCSY</td>
<td>Total Correlation Spectroscopy</td>
</tr>
<tr>
<td>TRIS</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>YPD</td>
<td>Yeast Peptone Dextrose</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 Protein interaction domains
Protein domains can be defined as conserved, independently folding units that perform a function using a common structural motif. An important subset of protein domains are protein interaction domains, which are non-catalytic modules that mediate interaction between proteins. Although some protein interaction domains bind folded domains, the majority recognize linear peptide motifs in their targets. Interactions between protein interaction domains and linear peptide motifs may account for 15%-40% of all interactions between proteins (Neduva et al., 2005). Since domain-peptide interactions have comparatively small interfaces, the affinity of such interactions is relatively weak. Therefore such interactions are particularly prevalent in the cellular processes that require transient interactions such as signaling pathways (Diella et al., 2008).

Src homology 3 (SH3) and Src homology 2 (SH2) domains are prototypical protein interaction domains. They were first identified as regions of homology in non-receptor tyrosine kinases that were distinct from the catalytic domain (Mayer et al., 1988; Sadowski et al., 1986). SH2 domains were later found to bind peptides containing phosphotyrosine, while SH3 domains were found to interact with proline rich ligands. Since then numerous other protein interaction domains with various binding preferences have been identified. The most common ones are listed in Table 1.1, where the SH3 domains are the most abundant, followed closely by PDZ domains and SH2 domains. Many protein interaction domains recognize peptides that bear some kind of post-translational modifications (Table 1.1), which means they are particularly well adapted for function in signaling pathways. For example the ability of the SH2 domains to recognize phosphorylated tyrosine allows them to be recruited to activated receptor tyrosine kinases and initiate downstream signaling pathways (Scott and Pawson, 2009). Another feature of the protein interaction domains is that they are often found in tandem with each other within the same protein. For example among 211 human SH3 domain containing proteins, 41 of these proteins also contain the SH2 domain and 35 also contain the PDZ domain. Such organization of protein interaction domains allows them to act in the assembly of multi-subunit protein complexes linking upstream and downstream targets within a signaling pathway. The appearance of different combinations of protein interaction
domains and the overall increase in their abundance has been linked to the increase in the phenotypic complexity of an organism (Basu et al., 2008). Although members of the same domain family share common binding preferences, there is great versatility in the modes of recognition and types of targets within a given domain family. The main focus of this thesis is examination of such complexity in the function of the protein interaction domains through the study of the SH3 domains.

Table 1.1 Abundance and binding preferences of the most wide-spread protein interaction domains

<table>
<thead>
<tr>
<th>Domain</th>
<th>Abundance (Total number/Number of proteins)</th>
<th>Yeast</th>
<th>Worm</th>
<th>Fly</th>
<th>Mouse</th>
<th>Humans</th>
<th>Binding</th>
</tr>
</thead>
<tbody>
<tr>
<td>SH3</td>
<td>27/23&lt;sup&gt;a&lt;/sup&gt;</td>
<td>80/60</td>
<td>91/71</td>
<td>279/206</td>
<td>283/211</td>
<td>Pro-rich</td>
<td></td>
</tr>
<tr>
<td>WW</td>
<td>9/6</td>
<td>30/18</td>
<td>36/21</td>
<td>82/46</td>
<td>85/47</td>
<td>Pro-rich</td>
<td></td>
</tr>
<tr>
<td>PDZ</td>
<td>3/2</td>
<td>92/64</td>
<td>102/65</td>
<td>242/141</td>
<td>256/142</td>
<td>C-terminal</td>
<td></td>
</tr>
<tr>
<td>SH2</td>
<td>1/1</td>
<td>71/67</td>
<td>37/32</td>
<td>120/111</td>
<td>120/110</td>
<td>Phospho-Tyr</td>
<td></td>
</tr>
<tr>
<td>PTB</td>
<td>0/0</td>
<td>16/15</td>
<td>9/9</td>
<td>35/31</td>
<td>38/33</td>
<td>Phospo-Tyr</td>
<td></td>
</tr>
<tr>
<td>BRCT</td>
<td>15/10</td>
<td>39/28</td>
<td>28/14</td>
<td>42/20</td>
<td>44/23</td>
<td>Phospo-Ser</td>
<td></td>
</tr>
<tr>
<td>FHA</td>
<td>13/12</td>
<td>12/12</td>
<td>16/16</td>
<td>27/27</td>
<td>29/29</td>
<td>Phospo-Thr</td>
<td></td>
</tr>
<tr>
<td>Bromo</td>
<td>14/10</td>
<td>23/15</td>
<td>26/18</td>
<td>55/42</td>
<td>58/42</td>
<td>Ac-Lys</td>
<td></td>
</tr>
<tr>
<td>Total genes&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6000</td>
<td>13000</td>
<td>19000</td>
<td>~25000</td>
<td>~25000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> These data was gathered from SMART (Simple Modular Architecture Research Tool) database (http://smart.embl-heidelberg.de/) searched in March, 2011 using Genomic mode.

<sup>b</sup> Information taken from Functional and Comparative Genomics Fact Sheet (http://www.ornl.gov/sci/techresources/Human_Genome/faq/compgen.shtml)

1.2 SH3 domains

As can be seen in Table 1.1, the SH3 domains are the most common protein interaction domains in eukaryotes. These domains are often found within signal-transduction enzymes in combination with catalytic domains including tyrosine kinase, guanylate kinase, and phospholipase modules. For example the tyrosine kinases of the Src family contain an N-terminal SH3 domain, a central SH2 domain, and a C-terminal tyrosine kinase domain. Under basal conditions, the SH3 and SH2 domains participate in an intermolecular interaction that holds the kinase domain in an inactive configuration. The binding of these domains to external ligands leads to the dissolution of this auto-inhibitory interaction and the activation of the kinase domain (Parsons and Parsons, 2004). SH3 domains are also commonly found together with other protein interaction domains within non-catalytic proteins that participate in signaling pathways. These proteins are called adaptors. Well characterized examples of these include the Crk, Grb, and Nck adaptor proteins that act as critical mediators in the activation of diverse signal transduction responses (Birge et al., 1996). Grb2, for example, uses its SH2
domain to bind activated receptor tyrosine kinase and its SH3 domain to interact with the Ras GTPase guanine exchange factor Sos. This links the activation of the receptor tyrosine kinase with the membrane recruitment of Sos, resulting in activation of the membrane-associated Ras GTPase and initiation of the mitogen activated protein kinase cascade (Lowenstein et al., 1992). In addition to being found in signaling proteins, the SH3 domains are also found in a variety of other proteins including: cytoskeletal proteins such as myosins and spectrins, endocytic proteins such as amphiphilins and endophilins, and regulatory proteins such as GTPase activating proteins and guanine nucleotide exchange factors.

1.2.1 SH3 domain structure

SH3 domains are typically 60 residues long and are composed of five β-strands linked together by the RT loop, the N-Src loop, the Distal loop and the 310 helix. A canonical SH3 domain structure is exemplified by the SH3 domain from the Fyn tyrosine kinase (Figure 1.1 A). Variations to this canonical structure usually occur in the form of sequence insertions within the N-Src or Distal loops. The amphiphysin SH3 domain contains 5 and 8 extra residues in the N-Src and Distal loops, respectively, compared to the Fyn SH3 domain (Figure 1.1B). There are also examples of SH3 domains containing extra structural elements, which contribute to the overall fold of the domain. Melanoma inhibitory activity (MIA) proteins, which are secreted by melanoma cells and function in preventing cellular adhesion, are composed of SH3 domains with ~20 residue N- and C-terminal extensions (Lougeed et al., 2001). These sequences fold into anti-parallel β-strands, extending one of the β-sheets of the MIA SH3 domains (Figure 1.1C). Another example is the SH3 domain from the T cell adapter protein ADAP (adhesion and degranulation promoting adaptor protein), which contains an N-terminal extension that folds into a helix and interacts with the rest of the SH3 domain through hydrophobic contacts ((Heuer et al., 2004), Figure 1.1D). The SH3 domain from a yeast protein Bem1p, which I investigate in this thesis, also contains a domain extension, which is one of the factors that made me interested in the functional properties of this domain (Chapter 2). As I demonstrate through my findings on Bem1p SH3 domain, such domain extensions can serve important functional roles through creation of additional and/or alternative binding surfaces.
1.2.2 Recognition of canonical PxxP targets by SH3 domains

SH3 domains typically recognize proline rich sequences containing PxxP motifs that adopt a polyproline helix II (PPII) helical conformation. These motifs are divided into class I, represented by the +xΨPxΨP consensus, and class II, represented by the ΨPxΨPx+ consensus, where Ψ is a hydrophobic amino acid and + is an Arg or Lys residue. Class I and Class II motifs bind in the opposite orientation, however the mode of the interaction with the SH3 domain surface is the same (Figure 1.2A). The ΨP or PΨ dipeptides pack into the hydrophobic
grooves on the SH3 domain surface that are composed of the Phe8/Tyr8, Phe10/Tyr10, Trp36, Pro51, and Phe54/Tyr54 residues (Figure 1.2B). These positions are conserved in most SH3 domains (Figure 1.2C; numbering according Larson and Davidson, 2000). Another conserved feature of the class I and class II complexes is the binding of the positively charged peptide residue. This residue (Lys or Arg) packs against Trp36 and forms electrostatic interactions with the negatively charged residues (Asp and Glu) in the RT loop, among which position 17 is the most highly conserved (Figure 1.2). While some SH3 domains can bind both class I and class II ligands, others are selective for one of the classes. However what determines these binding preferences is not well understood and was only directly addressed by a single study (Fernandez-Ballester et al., 2004). This study observed that certain residues at position 49 appear to lock SH3 domain in class II binding configuration, offering a possible explanation for why some SH3 domains can only bind to class II ligands. What determines selectivity of the SH3 domains for class I ligands was not addressed in this study.

The peptide sequences that flank the core PxxP motif (N-terminal in class I ligands and C-terminal in class II ligands) contact the surface of the SH3 domain located between the RT and N-Src loops (Figure 1.2B). This surface is composed of the residues that are not conserved between different SH3 domains (Figure 1.2C). Thus while the interaction with the peptide +xxPxxP or PxxPx+ core motifs, described above, serves as a common anchor for the binding of the SH3 domains, the interaction with the peptide residues flanking the core motif (extended region) contributes to the individual specificities of the interactions. Studying how a combination of the interactions with the peptide core motif and the extended region contributes to the specificity of the SH3 domain mediated interactions is one of the main focuses of this thesis. This will be discussed in greater detail through examples later in this chapter and presented through my findings in Chapters 3 and 4.
Figure 1.2 Interaction of SH3 domains with canonical PxxP containing targets

(A) An alignment of the class I and class II consensus motifs, standard peptide numbering is used (Lim et al., 1994). The interaction of the class I and class II ligands with the SH3 domain surface is depicted schematically (Mayer, 2001). The two hydrophobic pockets that interact with the ΨP dipeptides are colored green and the negatively charged pocket that interacts with the positively charged peptide residue at position (-3) is colored purple. The peptide prolines are highlighted in yellow, the adjacent hydrophobic residues (Ψ) are highlighted in orange, and the positively charged residue is highlighted in red. (B) The structure of the SH3 domain from the Src tyrosine kinase with the class I ligand VSL12 (PDB ID: 1QWF). The SH3 domain and the bound peptide are shown as cartoon representations. The key conserved residues involved in the interaction are shown as sticks. Same color scheme as in (A). The SH3 domain numbering is according to Davidson and Larson, 2000.

(C) An alignment of various SH3 domains derived from human and yeast proteins with diverse functions. The conserved aromatic residues and proline that interact with the ΨP dipeptides are highlighted in green. The negatively charged residues in the RT loop that interact with the positively charged residue at position (-3) are highlighted in purple. Other conserved residues include hydrophobic core positions and are highlighted in yellow.

1.2.3 Recognition of noncanonical targets by SH3 domains

There are numerous examples of the peptide core motifs recognized by the SH3 domains that do not belong to either class I (+xxPxxP) or class II (PxxPx+) consensus sequences. Some SH3 domains including the Abl SH3 domain (Pisabarro et al., 1998), the Irtks SH3 domain (Aitio et al., 2010) and the yeast Myo3y/Myo5p SH3 domains (Tong et al., 2002) recognize peptides defined by the ΨxxPxxP consensus, where Ψ is a hydrophobic amino acid. These
SH3 domains have evolved to possess residues in the RT loop that are capable of interacting with hydrophobic residues at peptide position (-3) instead of the positively charged Lys or Arg. For example, the Abl SH3 domain contains Thr17 instead of the negatively charged Asp or Glu. The structure of the Abl SH3 domain in complex with the high affinity APSYSPPPP ligand shows that Tyr(-3) is accommodated through formation of two hydrogen bonds between its carboxyl group and the RT loop residues Ser13 and Asp15 (Figure 1.3A, (Pisabarro and Serrano, 1996)).

There are also many examples of the SH3 domains recognizing peptides that contain core motifs which bear little resemblance to class I or class II consensus sequences. Examples of these include: the Eps8 SH3 domain which recognizes the PxxDY consensus (Aitio et al., 2008); the Fyb SH3 domain which binds the KxxYxxY containing sequences (Kang et al., 2000); and the Mona/Gads SH3 domain which interacts with the RxxK motifs (Harkiolaki et al., 2003). However, even in these cases the peptides are recognized in a way that is not too dissimilar from the canonical interactions. For example, the interaction of the Mona/Gads SH3 domain with the high affinity Slp76 APSIDRSTKPLP peptide resembles the binding mode of class II ligands. The structure of this complex reveals that the peptide residue Ile(-1) fits into the hydrophobic groove that normally accommodates the first PΨ dipeptide and the peptide residues, Pro(1) and Ala(2), fit into the adjacent groove that normally accommodates the second PΨ dipeptide. At the same time, the peptide residue Arg(-3) packs against Trp36 and interacts with negatively charged residues in the RT loop in the canonical fashion (Figure 1.3B, (Harkiolaki et al., 2003)).

In addition to binding peptides, SH3 domains have also been demonstrated to bind folded domains. For example, SH3 domains of several endocytic proteins including the yeast Sla1p and the mammalian CIN85 and amphiphysin proteins were demonstrated to interact with ubiquitin (Stamenova et al., 2007). These interactions also engage the conserved SH3 domain residues that are involved in the canonical peptide interactions. The structure of the Sla1p SH3 domain in complex with ubiquitin demonstrates that complex formation is mediated through the electrostatic interaction between ubiquitin Arg42 and the SH3 domain Glu17, and the packing of ubiquitin Val70 and Ile44 into the conserved hydrophobic grooves of the Sla1p SH3 domain (Figure 1.3C, (He et al., 2007)).
Similar to the cases described above, binding of the most atypical targets to the SH3 domains involves the same SH3 domains surfaces as those used for the recognition of the canonical targets (Kim et al., 2008). However, there are also a few examples of the SH3 domains that recognize targets through a completely different surface, which means that these domains can mediate multiple interactions through different binding surfaces (Douangamath et al., 2002; Nishida et al., 2001; Tatebayashi et al., 2006). In chapter 2 of this thesis, I describe the identification of such surface in the Bem1p SH3b domain and discuss the importance of the second binding surface for the function of this domain.

![Figure 1.3 Interactions of SH3 domains with noncanonical targets](image)

1.2.4 Affinity and specificity

One of the major focuses of this thesis is investigating the properties of the SH3 domain mediated interactions. To introduce this topic I will first describe the two important parameters of any protein interaction: affinity and specificity which I consider in this thesis. Protein interactions are often viewed in qualitative terms, where a pair of proteins is categorized as either interacting or not interacting. This is due to many protein interactions being characterized through semi-quantitative techniques, such as yeast-two-hybrid, affinity
chromatography or affinity blotting. However protein interactions are not static and a given protein can have multiple interaction partners. Therefore a more complete picture of a protein interaction network benefits from the knowledge of affinities between the interacting proteins, which can provide a more realistic view of the network that accounts for the relative amounts of different complexes. The affinity between two proteins is often represented by the equilibrium dissociation constant ($K_d$, units M), where $K_d$ is related to the free energy of complex formation ($\Delta G = RT \ln K_d$) and the ratio of association and dissociation rates ($K_d = k_{off}/k_{on}$). The affinities of protein interactions span a broad range of values ($10^{-3}$ to $10^{-15}$ M) and are influenced by the particular requirements of an interaction. Although the affinities measured between proteins in vitro are not necessarily the same as inside the cell, it is the comparison of the affinities measured under identical conditions that can provide insight into biological processes.

Since a pair of interacting proteins does not exist in isolation, another important parameter of protein interactions is specificity. A specific interaction can be defined as the binding of a protein to its biological target which contributes to a particular cellular process, while a non-specific interaction can be defined as off-target binding of this protein that does not fulfill any functional purpose. The specificity of a particular protein can in turn be defined as a fraction of this protein found in complex with its biological partner compared to the fraction involved in non-specific interactions. Considering crowded intra-cellular environment, protein interactions need to be highly specific not only to avoid miscommunication in cellular processes, but also to ensure the binding of a protein to its specific partner. The specificity of a protein interaction can be decomposed into intrinsic and contextual specificity. The intrinsic specificity is defined as an affinity of a protein for its specific interaction partner relative to the non-specific proteins ($K_d$ non-specific/$K_d$ specific). Thus the intrinsic specificity depends on the isolated proteins and is independent of other factors. On the other hand, the contextual specificity is the contribution of the environment to the interaction specificity between a pair of proteins. This can be separation of specific interacting partners from the non-specific proteins through spatial localization, temporal expression or co-localization through other protein interactions (Figure 1.4). For example, in the case of the signaling pathways that are initiated at the membrane, membrane recruitment can serve to enhance specificity by increasing the local concentration of the specific interaction partners. Certain biological
interactions clearly lack contextual specificity, necessitating the need for high intrinsic specificity. Examples of these include interactions between antibodies and antigens, hormones and receptors, and transcription factors and DNA (Greenspan, 2010; Schreiber and Keating, 2011; Szwajkajzer and Carey, 1997). Such interactions are distinguished by high intrinsic specificity, where the specific targets are bound with the affinity which is more than several orders of magnitude greater than the affinity for the non-specific targets.

The affinity of interactions with high intrinsic specificity is generally high, somewhere in the nanomolar \((10^{-9})\) to picomolar \((10^{-12})\) range. Thus, if an interaction is high affinity, this is usually taken as an indicator that this interaction also possesses high intrinsic specificity. This is partly true, since high affinity of interaction requires extensive interactions between the two molecules, which are more likely to be specific than limited interactions. On the other hand, if a contribution to the affinity of an interaction is made by a contact between the two molecules that can also be made in a non-specific complex, then the resulting higher affinity will also result in lowered specificity. For example, in the case of antigen-antibody recognition, high affinity interactions evolve through changes in the variable antibody domain that lead to higher geometric and chemical complementarity. However, one of these changes can involve a mutation that leads to increased interaction not only with the specific antigen, but also with the non-specific one leading to lower specificity (Greenspan, 2010). Indeed one study examining the capacity of different amino acids to contribute to the affinity and the specificity of synthetic antibody interactions found that while some residues contribute equally to the binding affinity, their effect on the specificity is different (Birtalan et al., 2010). For example, a Tyr residue was optimal both for the high affinity and specificity. However, an Arg residue at the same position, contributed equally to the binding affinity, but also conferred reduced specificity. The relationship between the affinity and the specificity is an important topic of investigation. It is crucial not only for understanding the basic biology, but also for the design of therapeutic or diagnostic macromolecules that need to be optimized in terms of both affinity and specificity. I will discuss this issue more extensively through my findings (Chapter 3) and in the conclusion section (Chapter 5).
1.2.5 Specificity of SH3 domain mediated interactions

As described above, many SH3 domains recognize peptides that contain class I or class II core motifs. This raises a question of how the specificity of the interactions is mediated in such a system, since domains that recognize the same class of motifs can potentially cross react with each other’s ligands. Indeed many studies on the SH3 domains demonstrated that they bind their targets with relatively low affinities (1-50 μM) and with significant cross-reactivity, binding their targets and the targets of other SH3 domains with comparable affinities (Cussac et al., 1994; Feller et al., 1995; Kaneko et al., 2003; Landgraf et al., 2004; Viguera et al., 1994). For example, in the case of the biologically relevant interaction between the Grb2 and Sos proteins (described above), the Grb2 N-terminal SH3 domain interacts with four similar Sos PxxP sites with affinities ranging from 4 μM to 120 μM (McDonald et al., 2009; Nguyen et al., 1998; Wittekind et al., 1997). At the same time, several other SH3 domains (Src, Fyn and Crk), were also demonstrated to bind the Sos-derived peptides with similar affinities (Nguyen et al., 1998; Viguera et al., 1994). The structure of the N-terminal Grb2 SH3 domain in complex with the best binding Sos peptide (VPPVPPRRR) and mutagenesis experiments demonstrated that the main contribution to peptide binding is made by residues of the conserved core motif (Arg(-3), Pro(-1) and Pro (2)) explaining the relatively low affinity and specificity of this interaction (Wittekind et al., 1997). Observations such as this led to a view that SH3 domains do not possess high intrinsic specificity and contextual specificity (defined above in section 1.2.4) must be responsible for maintaining the specificity of SH3 domain mediated interactions (Castagnoli et al., 2004; Ladbury and Arolf, 2000; Li, 2005; Mayer, 2001). For example, in the case of the Grb2-Sos interaction, the C-terminal Grb2 SH3 domain also interacts with one of the four Sos PxxP sites that are bound by the N-terminal SH3 domain (McDonald et al., 2009). Therefore, while the interactions of the individual Grb2 SH3 domains with Sos have relatively low affinity, full length Grb2 binds to Sos with a significantly higher affinity (Cussac et al., 1994). Thus it appears that in the case of Grb2-Sos interaction, the specificity is established through the co-operative binding of both Grb2 SH3 domains to Sos protein. Since many SH3 domain containing proteins contain more than one SH3 domain (Table 1.1) and SH3 domains are also often found together with other protein interaction domains, this could be a common mechanism for improving the specificities of the interactions.
An important question is whether low intrinsic specificity in SH3 domains is the result of a general inability of SH3 domains to possess high intrinsic specificity? Alternatively, the absence of evolutionary pressure on these interactions to develop high intrinsic specificity either due to the contribution of the contextual specificity or other factors could also account for low intrinsic specificity. In this thesis, I show that SH3 domains are capable of evolving high intrinsic specificity when there is a requirement for them to do so. I will introduce this subject in next section by discussing different mechanisms that contribute to the intrinsic specificity of SH3 domain mediated interactions.

![Figure 1.4 Mechanisms of SH3 domain specificity](image)

In the case of the intrinsic specificity, a specific peptide (green) can discriminate between the specific domain (blue) and the non-specific domain (pink). As outlined in the text, this can be accomplished through the interactions with an extended peptide region, recognition of a non-canonical peptide motif, or negative selection against non-specific targets. In the case of contextual specificity, the specific peptide (green) can bind both the specific (blue) and the non-specific (pink) domains equally well. In this case, other factors including multiple interaction co-operativity, complex formation, or separation in space or time must come into play to ensure the specificity of the interactions.
1.2.6 Mechanisms through which SH3 domains can achieve higher specificity

The interactions between SH3 domains and peptide residues flanking the core motif is one of the mechanisms that can contribute to the specificity of SH3 domain mediated interactions (Figure 1.4). As described earlier, peptide residues in the extended region contact the SH3 domain surface located between the N-Src and RT-loops (Figure 1.2B). The residues on this surface are not highly conserved among different SH3 domains; thus, the interactions mediated through this region are expected to confer specificity. There are numerous examples of the SH3 domain-peptide interactions where interactions between the extended region of the peptide and the SH3 domain contribute significantly to the binding energy. An extreme example of this is the interaction between the p47phox peptide and the p67phox SH3 domain that is crucial for the assembly and the activation of phagocyte NADPH oxidase (Kami et al., 2002). In this case, the interaction of the SH3 domain with the extended peptide region increases the affinity by 1000-fold from 20 μM to 0.024 μM. Another example, which is perhaps more applicable to the majority of the SH3-peptide interactions, is the interaction between the Lyn SH3 domain and the herpes virus tyrosine kinase interacting protein (Tip). In this case, the interaction of the Lyn SH3 domain with the extended region of the Tip peptide confers a more modest increase in the affinity of about 20-fold from 16.8 μM to 0.8 μM (Bauer et al., 2005). A question arises of whether such interactions with the extended peptide region, which confer higher binding affinity and specificity, are an exception or a general property of SH3 domain mediated interactions. One study examined the contribution of the residues flanking the core motif to the total binding energy for multiple families of the protein interaction domains, one of which was the SH3 domain family (Stein and Aloy, 2008). This was done through the examination of the available structures of the protein domain-peptide complexes and theoretical alanine scanning mutagenesis. They found that in the case of the SH3 domain family (represented by 95 structures), peptide residues in the extended region contribute on average 30% to the total binding energy. To put this in perspective, reducing binding energy by 30% would be equivalent to reducing the affinity of an interaction from 1 μM to 60 μM. These results suggest that the interactions between SH3 domains and extended peptide regions do in fact play an important role for many SH3-peptide interactions.
The interactions of the SH3 domains with non-canonical targets, described in section 1.2.3, would also be expected to result in higher intrinsic specificity, since they involve recognition of motifs not recognized by the majority of the SH3 domains (Figure 1.4). Indeed the interaction between the Mona/Gads SH3 domain and the noncanonical Slp76 peptide (described in section 1.2.3) was demonstrated to possess high intrinsic specificity (Harkiolaki et al., 2003). Out of the 146 human SH3 domains tested, only the Grb2 and Stam1/2 SH3 domains also bound to the Slp76 peptide as detected by an SH3 domain microarray binding assay. However, even in the case of these two SH3 domains that bound, the affinities of the interactions appeared to be considerably lower. In the case of the Grb2 SH3 domain the affinity was 1000-fold lower than the affinity of the Mona/Gads SH3 domain for the Slp76 peptide. Significantly, a version of the Slp76 peptide that carried a mutation which created a PxxP motif interacted with multiple other SH3 domains in addition to the Mona/Gads SH3 domain. This suggested that the high intrinsic specificity of the Mona/Gads SH3-Slp76 peptide interaction depends on the absence of the canonical PxxP motif in the Slp76 peptide.

The specificity of SH3 domain mediated interactions can also be improved through negative selection against promiscuous interactions (Figure 1.4). In contrast to the aforementioned mechanisms, which rely on a lack of positive interactions with non-specific targets, this mechanism implies active selection against the binding of non-specific targets accomplished through negative interactions. This was most elegantly demonstrated by the Zarrinpar et al. study (Zarrinpar et al., 2003) which examined interactions between all yeast SH3 domains and the target of the Sho1p SH3 domain, Pbs2p (VNKPLPPLP). They found that none of the 26 non-cognate yeast SH3 domains were able to bind Pbs2p. On the other hand, 6 out of the 12 non-yeast SH3 domains tested were able to bind the peptide. This suggested that the observed high intrinsic specificity of the Sho1p SH3-Pbs2p interaction does not depend on the specific interactions between the Sho1p SH3 and Pbs2p (since many non-yeast SH3 domains were able to bind Pbs2p), but rather on the ability of the Pbs2p peptide to avoid non-specific interactions with the other yeast SH3 domains. In further support of this, substitution of the prolines at positions (-2) and (2) with alanines (VNKALPALP) produced a promiscuous Pbs2p peptide. This peptide still bound Sho1p SH3 domain with a similar affinity as the wild-type peptide, but unlike wild-type peptide, also bound several other yeast SH3 domains with
the comparable affinities. Thus the overall conclusion of this study was that high intrinsic specificity in the interactions mediated by the SH3 domains or other protein interaction domains can be achieved through negative selection against non-specific interactions.

In principle, SH3 domain-peptide interactions can rely to differing extents on a combination of mechanisms to achieve high intrinsic specificity. For a given SH3-peptide pair this can be influenced by the particular requirements of that interaction. For example, if an interaction requires high affinity and specificity then combination of interaction with a canonical PxxP motif and an extended peptide region can provide that. On the other hand, if an interaction does not require high affinity binding, then recognition of a non-canonical binding motif may be a better way of achieving high specificity binding without a simultaneous increase in affinity. Understanding to what extent SH3 domains and other protein interaction domains rely on the intrinsic specificity and how this intrinsic specificity can be achieved is an important topic of investigation. My work examined intrinsic specificity in the context of the two yeast SH3 domains that recognize class I PxxP motifs. In the next section, I will therefore provide a brief overview of yeast SH3 domains and their binding properties.

1.2.7 Yeast SH3 domains

The yeast species *Saccharomyces cerevisiae* encodes 27 unique SH3 domains that are found in 23 proteins (Table 1.2). Although there are four pairs of SH3 domains found in paralogous proteins that share greater than 70% sequence identity (Lsb1p/Lsb2p, Lsb3p/Lsb4p, Boi1p/Boi2p, Myo3p/Myo5p), the rest of the SH3 domain sequences are highly diverse. Excluding these very similar SH3 domains produces a diverse collection of SH3 domains with no two displaying higher than 44% sequence identity (Figure 1.5). Most SH3 domain containing proteins in yeast have been implicated in either signal transduction or reorganization of the actin cytoskeleton (Table 1.2).

Phage display experiments involving all yeast SH3 domains proteome identified binding preferences for most of them (Tong et al., 2002; Tonikian et al., 2009). These can be roughly divided into SH3 domains that recognize class I or class II ligands, SH3 domains that recognize both classes, and SH3 domains that recognize atypical targets (Table 1.2). However, since relatively short peptides (9 or 10 amino acids long) were employed, the consensus
sequences identified in these studies do not extend far beyond the core binding motifs. As mentioned above, the interactions of the SH3 domains with extended peptide regions can make a significant contribution to the affinity and the specificity of the interactions. Therefore, the majority of the yeast SH3 domain consensus sequences identified through these phage display studies may not contain enough information for the accurate identification of their biological targets. For the same reason, the conclusion derived from these studies that many yeast SH3 domains possess a large number of putative binding partners and overlapping binding specificities (Cesareni et al., 2002; Landgraf et al., 2004; Tong et al., 2002) may be erroneous. This is because putative biological targets identified through the use of the consensus sequences with insufficient binding information may identify weak and non-specific interactors. Thus the resulting protein interaction network may not be an accurate reflection of the in vivo situation.

**Figure 1.5** An alignment of yeast SH3 domains.

The alignment was constructed using Jalview (Clamp et al., 2004), default clustalx coloring scheme was used. For SH3 domains with high sequence identity (>70%) one of the redundant sequences was eliminated.
### Table 1.2 Yeast SH3 domains

<table>
<thead>
<tr>
<th>SH3 domain</th>
<th>GO Molecular Function and Biological process</th>
<th>Structure</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sho1p SH3</td>
<td>Osmosensor Osmosensory signaling pathway</td>
<td>2vkn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>KPLPLLPLA</td>
<td></td>
</tr>
<tr>
<td>Cyk3p SH3</td>
<td>Cytokinesis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bzz1p SH3a</td>
<td>Actin filament organization</td>
<td>1zuu</td>
<td>Class I</td>
</tr>
<tr>
<td>Bzz1p SH3b</td>
<td>Actin filament organization</td>
<td>2a28</td>
<td>+xxPxxP</td>
</tr>
<tr>
<td>Bem1p SH3b</td>
<td>Protein scaffold Establishment of cell polarity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nbp2p SH3</td>
<td>Negative regulation of kinase activity</td>
<td>1yn8</td>
<td></td>
</tr>
<tr>
<td>Sla1p SH3c</td>
<td>Cytoskeleton protein binding Actin filament organization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pex13p SH3</td>
<td>Protein binding, bridging Protein import into peroxisome</td>
<td>1nz5</td>
<td></td>
</tr>
<tr>
<td>Abp1p SH3</td>
<td>Actin filament binding, Actin cortical patch assembly</td>
<td>2rpm</td>
<td>Class II</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KPKLHSPKPPPPTPKTKKA</td>
<td>PxxPx+</td>
</tr>
<tr>
<td>Lsb3p/Lsb4p SH3</td>
<td>Actin cortical patch localization</td>
<td>1ssh</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GPPPAMPARPT</td>
<td></td>
</tr>
<tr>
<td>Lsb1p/Lsb2p SH3</td>
<td>Actin cytoskeleton organization</td>
<td>1xz6</td>
<td></td>
</tr>
<tr>
<td>Bbc1p SH3</td>
<td>Myosin 1 binding Actin cytoskeleton organization</td>
<td>1zuk</td>
<td>I and II</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RGPAPPPPHR</td>
<td></td>
</tr>
<tr>
<td>Hse1p SH3</td>
<td>Ubiquitin binding Endosome to vacuole transport</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rvs167p SH3</td>
<td>Cytoskeleton protein binding Endocytosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sla1p SH3ab</td>
<td>Cytoskeleton protein binding Actin filament organization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myo3/Myo5 SH3</td>
<td>Microfilament motor activity Endocytosis, exocytosis, actin patch</td>
<td>2btt</td>
<td></td>
</tr>
<tr>
<td>Boi1/Boi2 SH3</td>
<td>Cytokinesis, establishment of cell polarity</td>
<td></td>
<td>Unusual</td>
</tr>
<tr>
<td>Hof1</td>
<td>Cytoskeleton protein binding, cytokinesis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bem1 SH3a</td>
<td>Protein scaffold Establishment of cell polarity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fus1p SH3</td>
<td>Cell fusion during mating</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cdc25p SH3</td>
<td>Guanine exchange factor Regulation of cell cycle</td>
<td></td>
<td>No peptide</td>
</tr>
<tr>
<td>Bud14p SH3</td>
<td>Cytoskeleton organization</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* a SH3 domains with higher than 70% sequence identity are grouped together
* b GO Molecular Function and Biological Process are indicated if known and were taken from the Saccharomyces Genome Database (SGD www.yeastgenome.org)
* c PDB code is given if the SH3 domain structure is available. If the SH3 domain was solved in complex with a peptide, the peptide sequence is shown
* d Class assignments are taken from phage display studies (Tong et al., 2002; Tonikian et al., 2009)
In addition to binding studies performed on the yeast SH3 domains, the structures of many these domains (13 out of 27) have been determined and are available in PDB database (Table 1.2). However, apart from the Abp1p SH3 structure (Stollar et al., 2009), most of these structures contain either no ligand or short ligands, which offer limited insight into the binding requirements of these domains.

Yeast SH3 domains serve as a good model system for investigating the principles behind protein interactions mediated by the SH3 domains. First, the existence of only 27 SH3 domains makes it feasible to characterize the interactions of each SH3 domain and thus obtain a complete SH3 domain interaction network. Second, their high sequence diversity and recognition of diverse targets suggests that yeast SH3 domain mediated interactions can be representative of the SH3 domain interactions in other organisms with larger numbers of SH3 domains. Last, the amenability of yeast to genetic manipulation and phenotypic characterization provides a comparatively easy system to access the validity and study the properties of the SH3 domain mediated interactions in the biological context. In this thesis, I explore the issues of specificity and target recognition in SH3 domain-mediated interactions through investigating the function of the SH3 domains from the yeast adaptor proteins Nbp2p and Bem1p.

1.3 Nbp2p and Bem1p

This thesis is focused on studying the specificity and function of the Nbp2 SH3 (NbpSH3) and Bem1 SH3b (BemSH3b) domains from the yeast adaptor proteins Nbp2p and Bem1p. I was intrigued by these domains because both of them appeared to recognize the same consensus binding motif $\Psi xPxRxAPxxP$, where $\Psi$ represents a hydrophobic amino acid (described below in more detail). This motif is a class I canonical motif, where the $\Psi xPx$ corresponds to the extended region and the $RxAPxxP$ corresponds to the core motif. Thus, the NbpSH3 and BemSH3b domains both recognize not only the same core motif, which is also recognized by many SH3 domains, but also the same extended region, which confers specificity to the interactions. There are other SH3 domain pairs in yeast that share binding preferences (Landgraf et al., 2004; Tong et al., 2002). However, in all of these other cases, the SH3 domains are found in paralogous proteins and are very similar in sequence. Nbp2p and Bem1p on the other hand share no similarity in function or domain composition and the NbpSH3 and
BemSH3b domains are only 36% identical (52% similar). The recognition of the same motif by two different SH3 domains with distinct functions raises the question of how the specificities of these domains are controlled with respect to each other. This is one of the main questions I investigate in this thesis, which will be addressed through my findings in chapters 3 and 4, and in the conclusion section (chapter 5).

In addition to displaying cross-reactivity with the NbpSH3 domain, the BemSH3b domain possesses another interesting property. It contains a ~40 amino acid long C-terminal extension region that is required for folding and binding to the Cdc42p Rho GTPase in a PxxP independent manner (Yamaguchi et al., 2007). There are few examples of SH3 domains that possess extra structural regions and/or use alternative binding interfaces for additional interactions. However such domains may be more common than is currently appreciated, since studies on SH3 domains are likely to be biased towards the domains with the canonical structure and binding properties. This makes the characterization of unusual domains an important objective. Thus investigation of the BemSH3b domain structure and its unusual binding properties was another goal of my project (Chapter 2).

In the following sections I will provide a brief overview of the Nbp2p and Bem1p cellular roles, with particular focus on the functions of the NbpSH3 and BemSH3b domains. Since the function of Bem1p is intimately linked with Cdc42p activity and the functions of both Nbp2p and Bem1p are associated with the yeast p21 activated kinases (PAKs) and yeast mitogen activated protein kinase (MAPK) signaling pathways, I will start by giving a brief overview of the Cdc42p, PAK, and MAPK functions in yeast.

1.3.1 Cdc42p

Cdc42p belongs to the Rho family (part of the Ras superfamily) of the GTPases, which are highly conserved in all eukaryotes, and play a key role in signaling pathways. GTPases possess intrinsic GTPase activity and alternate between the inactive GDP-bound and the active GTP-bound forms. In the active GTP-bound form, the switch regions of a Rho GTPase adopt a particular conformation that allows it to bind and regulate its effector proteins. The conversion between the active and the inactive forms of a Rho GTPase is regulated by several classes of proteins (Figure 1.6). The guanine nucleotide exchange factors (GEFs) activate Rho
GTPases by promoting dissociation of GDP. Since the cytoplasmic concentration of GTP is higher than GDP, this promotes the binding of GTP. On the other hand, GTPase-activating proteins (GAPs) inactivate Rho GTPases by promoting hydrolysis of GTP to GDP. Another class of negative regulators are guanine nucleotide dissociation inhibitors (GDIs), which inhibit dissociation of GDP and sequester Rho GTPases, which are normally targeted to the membrane, in the cytoplasm. Membrane targeting of Rho GTPases is due to the post-translational addition of geranylgeranyl lipid to their C-terminus and plays an important role in the Rho GTPase function. Studies investigating an interaction of a Rho GTPase with its effectors \textit{in vitro} require the use of a Rho GTPase in the active GTP-bound form. This poses a problem due to the intrinsic ability of GTPases to hydrolyze GTP into GDP resulting in the conversion of the GTPase into inactive conformation. This is often circumvented by the use of non-hydrolyzable GTP analogues or GTPase mutants that are unable to hydrolyze GTP, such as Rac or Cdc42 Q61L mutants. Although the Q61L mutants possess slightly higher affinity for their effector proteins (Thompson et al., 1998), they are very widely used since they offer an easier and cheaper alternative than the use of non-hydrolyzable GTP analogues. In Chapter 2, I use the Cdc42p Q61L mutant to investigate the binding between the BemSH3b domain and Cdc42p.

In yeast, Cdc42p plays a key role in the establishment of cell polarity (reviewed in (Perez and Rincon, 2010)) and the activation of several MAPK pathways (described in more detail in section 1.3.3). Cdc42p is activated by the GEF, Cdc24p, which is also targeted to the membrane through a phosphinositide-binding PH domain (Toenjes et al., 1999). During cell division Cdc24p is released from the nucleus and is recruited to the site of a new bud growth through the interaction with the Bud1p GTPase (Figure 1.6). Yeast cells initiate budding at specific locations (adjacent to the former bud in the case of the haploid cells) and Bud1p GTPase is part of the complex which serves as positional landmark for these sites (Park et al., 1993). In the absence of spatial cues provided by Bud1p GTPase the cells are still able to initiate polarization at random sites, which has been referred to as symmetry breaking (Irazoqui et al., 2003). The function of Bem1p is essential for the symmetry breaking and the deletion of both \textit{BUD1} and \textit{BEM1} is lethal (Irazoqui et al., 2003). The contribution of Bem1p to symmetry breaking behavior will be discussed in detail in section 1.3.5. At the sites of nascent buds, Cdc24p activates Cdc42p (also assisted by Bem1p), which in turn establishes
cell polarity through the activation of multiple effector proteins. Briefly, these include PAKs Ste20p and Cla4p (introduced in section 1.3.2), the formin Bni1p, the septin organizing proteins Gic1p and Gic2p, and the exocyst subunit Sec3p (Reviewed in (Perez and Rincon, 2010)). Through these interactions Cdc42p initiates the reorganization of cytoskeleton, polarized secretion, and the septin organization required for polarized cell growth.

Figure 1.6 Cdc42p regulation and establishment of cell polarity

As described in text, the activity of Cdc42p is positively regulated by the GEF, Cdc24p, and negatively by GAPs and GDIs (left panel). During cell division, Cdc24p is released from the nucleus (through G1 cyclin-Cdc28p mediated degradation of Far1, which sequesters Cdc24p within the nucleus), associates with Bud1p at the plasma membrane and activates Cdc42p. Active Cdc42p interacts with the indicated downstream effectors to initiate rearrangement of actin cytoskeleton, polarized secretion and septin formation (right panel). This Figure was adapted from Perez et al, 2010.

1.3.2 Yeast p21 activated kinases: Ste20p, Cla4p and Skm1p

An important subset of proteins regulated by Rho GTPases are PAKs, serine/threonine kinases that are activated by the binding of Rho GTPases. PAKs serve as important regulators of multiple cellular processes, including cytoskeleton organization, signal transduction in the MAPK pathways, cell cycle progression, apoptosis, and many others (Bokoch, 2003). All PAKs contain an autoregulatory region and a catalytic kinase domain (Figure 1.7A). The autoregulatory region is composed of the dimerization segment, the p21 binding domain, and the inhibitory switch domain. A Cdc42 and Rac interactive binding (CRIB) motif, which is conserved among many Cdc42 and Rac effectors, is part of the p21 binding domain. However, the complete p21 binding domain is required for maximal binding affinity to Rho GTPases. In the inactive state, a PAK kinase is kept in the autoinhibited conformation through homodimerization and the binding of the inhibitory switch domain to the kinase domain (Figure 1.7A). Since the p21 binding domain overlaps with the dimerization region and the
inhibitory switch domain, the binding of the GTPase to the p21 binding domain motif disrupts these autoinhibitory interactions, resulting in auto-phosphorylation and activation of the kinase (Lei et al., 2000; Parrini et al., 2002).

Yeast contains three PAKs: Ste20p, Cla4p and Skm1p. The domain organization and the functionally important regions of these PAKs are shown in Figure 1.7B. In each PAK the autoregulatory domain is preceded by the membrane targeting domain: a conserved phospholipid binding PH domain in Cla4p and Skm1p, and a short basic region (BR) that has affinity for phospholipids in Ste20p. Since Cdc42p is localized to the membrane, the membrane targeting domains in these PAKs are important for promoting Cdc42p binding and their subsequent activation (Takahashi and Pryciak, 2007). Ste20p also contains a conserved binding site for the Gβ (Ste4p) subunit of the trimeric G-protein (Leeuw et al., 1998), which is important for its function in the mating MAPK pathway (described in more detail in section 1.3.5). All of the yeast PAKs contain the consensus ΨxPxRxAPxxP binding motif, recognized by the NbpSH3 and BemSH3b domains, located between the autoregulatory and kinase domains (Figure 1.7B). As described in more detail below, the binding of the BemSH3b domain to this motif in Ste20p is important for signal transduction in the mating pathway (Winters and Pryciak, 2005), while the interaction between the BemSH3b domain and Cla4p (most likely mediated through this motif) is important for polarized growth (Bose et al., 2001; Gulli et al., 2000).

Ste20p and Cla4p play a role in numerous cellular processes and these PAKs share at least one essential function, since deletion of both STE20 and CLA4 is lethal (Cvrckova et al., 1995). Detailed description of the multiple functions performed by the Ste20p and Cla4p kinases is not included here for the sake of brevity (reviewed in (Park and Bi, 2007)). Only the Ste20p and Cla4p cellular roles that have a proven or inferred functional connection to Bem1p or Nbp2p will be described in more detail below (section 1.3.4 and 1.3.5). Ste20p is best known for its role as an activating kinase in MAPK kinase pathways (described in more detail in section 1.3.3). Both Ste20p and Cla4p play a role in multiple processes associated with cell division including polarized growth (Goehring et al., 2003), actin cytoskeleton organization (Eby et al., 1998), exit from mitosis (Hofken and Schiebel, 2002) and vacuole inheritance (Bartholomew and Hardy, 2009). Cla4p also plays an important role in the assembly of the
septin ring that is required for proper cytokinesis (Versele and Thorner, 2004). Unlike Ste20p and Cla4p, Skm1p is poorly characterized and no clear function had been attributed to this PAK until recently. Interestingly, all three PAKs translocate into the nucleus and downregulate the transcription of genes involved in sterol uptake through interaction with the transcriptional regulator Sut1p (Lin et al., 2009). Of note, Ste20p also directly regulates chromatin state during apoptosis through histone phosphorylation, emphasizing great diversity of the cellular processes associated with PAK function (Ahn et al., 2005).

Figure 1.7 Yeast p21 activated kinases: Ste20p, Cla4p and Skm1p

(A) Mechanism of PAK activation. As described in text, the autoregulatory domain of PAKs is composed of the dimerization segment (Di), the p21 binding domain (PBD), which includes the Cdc42 and Rac interactive binding (CRIB) motif, and the inhibitory sequence (IS). A PAK kinase in the inactive conformation is homodimerized and inhibited in trans through the binding of the IS domain to the kinase domain. Cdc42p binding disrupts dimerization and unfolds the IS domain, relieving autoinhibition. Subsequent autophosphorylation (or phosphorylation by other kinases) of the IS domain and the kinase domain locks a PAK kinase in the active conformation. (B) Domain organization of yeast PAKs: Ste20p, Cla4p and Skm1p. All three kinases contain a membrane interaction domain, a regulatory/Cdc42p binding domain, an NbpSH3/BemSH3b binding PxxP motif, and a kinase domain. For clarity purposes the depiction of different segments is not to scale. The residue numbers of domain boundaries are indicated.
1.3.3 Yeast MAPK pathways
An essential property of cells is the ability to sense and respond to the external signals. The transmission of intracellular signals is often mediated by MAPK pathways - a highly conserved class of pathways found in fungi, plants and animals. All MAPK pathways share a basic architecture consisting of three kinases: MAPK kinase kinase (MAPKKK or MEKK), MAPK kinase (MAPKK or MEK) and MAPK (MAPK or ERK). As shown in Figure 1.8, an activated MAPKKK phosphorylates and activates a MAPKK, which in turn phosphorylates and activates a MAPK. Activated MAPKs phosphorylate and activate a variety of targets, such as transcription factors, thereby changing cell properties in the response to the received signal.

Yeast contains five MAPK pathways: the mating pathway, the invasive growth pathway, the high osmolarity glycerol (HOG) pathway, the cell wall integrity (CWI) pathway, and the spore wall assembly pathway (Figure 1.8). These pathways have been extensively reviewed (Chen and Thorner, 2007; Gustin et al., 1998). The mating pathway, which is one of the best characterized MAPK pathways in eukaryotes, responds to mating pheromones and prepares cells for mating through a G1 cycle arrest, polarized growth towards the mating partner and activation of genes necessary for cell fusion (yeast cells can exist as $MAT^a$ and $MAT^\alpha$ haploids, which mate to produce $MAT^a/MAT^\alpha$ diploids). The invasive growth pathway responds to limiting nutrients in the environment and changes the budding pattern, cellular morphology and cell-surface properties, which is thought to improve the ability of cells to explore the surroundings for additional nutrients. The HOG pathway responds to high osmolarity in the environment and counteracts it by raising the internal osmolarity through increased glycerol concentration. The CWI pathway is activated by cell wall damage and responds to it by synthesizing and modifying cell wall components. It is also important for the remodeling of the cell wall during polarized growth. Finally, the spore assembly pathway, which is poorly characterized, is important during sporulation of diploid cells (diploid cells undergo meiosis producing 4 haploid spores under conditions of carbon and nitrogen starvation).
Figure 1.8 Yeast MAPK pathways

Schematic representation of the five yeast MAPK pathways. The components that are unique to each pathway are colored the same color: the mating pathway-pink, the invasive growth pathway-red, the HOG pathway-blue, the CWI pathway-violet and the spore wall assembly pathway-orange. The components that are shared between the pathways are colored the same color in different pathways: Ste11p-pink, Ste7p-orange, Ste20p-purple, Cdc42p-green, and Sho1p-dark green. Additional factors and interactions, as well as connections to other pathways are omitted for clarity.

Ste20p acts as the upstream kinase in the mating, invasive growth and HOG pathways. As mentioned above, Ste20p is activated by binding of the activated (GTP-bound) Cdc42p. What mediates the activation of Cdc42p and its binding to Ste20p is relatively well understood in the case of the mating pathway and will be discussed in more detail in the section on the Bem1p function (section 1.3.5). On the other hand, it is not clear how the extracellular signal received by the membrane receptors activates Cdc42p and Ste20p in the case of the invasive growth and the HOG pathways. The CWI pathway is activated by a different upstream kinase-Pkc1p, which in turn is activated by the Rho1p GTPase. As can be seen in Figure 1.8, in addition to sharing Cdc42p-Ste20p activation module, the mating, invasive growth and HOG pathways also share other components, where Ste11p acts as the MAPKKK in all three pathways. In this case the choice of the downstream pathway is at least partly determined by the scaffold proteins: Ste5p in the case of the mating pathway and Pbs2p in the case of the HOG pathway. There are also other mechanisms including docking interactions, cross-pathway inhibition and kinetic insulation that ensure fidelity of signal transmission through these pathways (Saito, 2010).
In addition to activating MAPK pathways under desirable circumstances, the cells must also be able to inactivate these pathways once a sufficient response has been generated, and maintain the basal activity of the pathways low in the absence of stimuli. One way in which the MAPK pathways can be downregulated is through dephosphorylation of MAPKs. Since MAPKs are only fully active when phosphorylated on both the Thr and Tyr residues, their inactivation can be accomplished by either threonine/serine, tyrosine or dual specificity phosphatases. In yeast, many phosphatases have been implicated in the downregulation of the MAPK pathways: the threonine/serine phosphatases Ptc1p, Ptc2p and Ptc3p; the tyrosine phosphatases Ptp2p, Ptp3p; and the dual specificity phosphatases Sdp1p and Msg5p (reviewed in (Martin et al., 2005), Figure 1.8). As can be seen in Figure 1.8, a MAPK can be downregulated by multiple phosphatases, which contribute to different extents to the maintenance of low basal activity and the inactivation of the stimulated kinase. Unlike tyrosine phosphatases (Ptp2p and Ptp3p) and dual specificity phosphatases (Sdp1p and Msg5p) that have been demonstrated to directly bind their targets (reviewed in (Martin et al., 2005)), the mechanism by which threonine/serine phosphatases (Ptc1p, Ptc2p and Ptc3p) are targeted to their substrates is poorly understood. An exception to that is the targeting of Ptc1p phosphatase to the HOG and CWI pathways, which is accomplished through Nbp2p. The role of Nbp2p as an adaptor for the Ptc1p phosphatase in the HOG and the CWI pathways will be described in more detail in the next section focused on the Nbp2p function.

1.3.4  Nbp2p
Nbp2p is a 236 amino acid long adaptor protein that contains a central SH3 domain bounded by N-terminal and the C-terminal sequences that bear no discernable similarity to characterized protein motifs (Figure 1.9A). Nbp2p was originally identified as the interacting partner of Nap1p-a nucleosome assembly protein that is also involved in septin and microtubule organization during mitosis. The binding site of Nap1p was later mapped to the C-terminal region (Hruby et al., 2011), while the N-terminal region was demonstrated to interact with the Ptc1p phosphatase (Mapes and Ota, 2004). To date, the SH3 domain of Nbp2p (NbpSH3) was shown to interact with the ΨxPxRxAxxP motifs in four kinases: Ste20p (Winters and Pryciak, 2005), Cla4p (Hruby et al., 2011), Bck1p (Hruby et al., 2011) and Pbs2p (Mapes and Ota, 2004). While no functional roles have yet been attributed to the
Ste20p and Cla4p interactions, the function of NbpSH3 interactions with Pbs2p and Bck1p kinases is well understood (Mapes and Ota, 2004; Stanger, 2008).

Figure 1.9 Domain organization and function of Nbp2p.

(A) Schematic representation of Nbp2p functional regions. (B) The role of Nbp2p as the adaptor for Ptc1p phosphatase in the MAPK pathways. In the HOG and CWI pathways, Ptc1p acts as a negative regulator by removing the phosphate group from the activated MAPKs. Nbp2p functions to recruit Ptc1p by binding the components of these pathways through the NbpSH3 domain and the Ptc1p phosphatase via the N-terminal region. In the mating pathway, Ptc1p acts as a positive regulator through dephosphorylation of Ste5, which leads to dissociation of the activated Fus3p and induction of the shmoo response. Nbp2p may assist in the recruitment of Ptc1p to the mating pathway through interaction of NbpSH3 domain with Ste20p kinase. This role of Nbp2p is speculative and has not yet been supported by any experimental data.

Pbs2p acts as the scaffold and the MAPKK of the HOG pathway (Figure 1.9B). The binding of NbpSH3 to Pbs2p recruits the Ptc1p phosphatase to the MAPK of the pathway, Hog1p, leading to the dephosphorylation of Hog1p and downregulation of its activity. This negative regulation by the Nbp2p-Ptc1p complex is required for both lowering the basal activity and the inactivation of the stimulated pathway (Mapes and Ota, 2004). Nbp2p plays an analogous role in the CWI pathway, where it binds the MAPKK of the pathway Bck1p via its SH3 domain and recruits Ptc1p phosphatase to downregulate the activity of the MAPK-Slt2p (Figure 1.9B). This was demonstrated by Karen Stanger in our lab, which showed that mutation of either the NbpSH3 domain or its binding site within Bck1p resulted in the increased Slt2p phosphorylation levels (Stanger, 2008). A different study also showed that
deletion of either \textit{NBP2} or \textit{PTC1} increases Slt2p phosphorylation levels, however this study did not link the observed effect to the NbpSH3-Bck1p interaction (Du et al., 2006).

The role of the Nbp2p-Ptc1p complex as a negative regulator of the HOG and CWI pathways leads me to question whether the interaction between NbpSH3 and Ste20p might also function to downregulate the mating, invasive growth, and HOG pathways in which Ste20p acts as an upstream kinase. Interestingly, a recent study suggests that the opposite may be true in the case of the mating pathway (Malleshaiah et al., 2010). This study explored the molecular basis behind a switch-like mating response, where at certain pheromone concentration the cells form shmoos (a pre-fusion state characterized by polarized growth towards the mating partner) in an “all-or-none response”. They discovered that a phosphorylation state of the scaffold protein Ste5p is critical for this response, where dephosphorylation of Ste5p at a certain pheromone concentration leads to the dissociation of the activated MAPK- Fus3p and induction of the shmooring response. Significantly, they found that the Ptc1p phosphatase is required for the dephosphorylation of Ste5p and activation of the shmooring response, where deletion of \textit{PTC1} reduced shmooring and activation of Fus3p, while its overexpression enhanced both. The recruitment of Ptc1p to Ste5p was determined to be dependent on the addition of the mating factor and phosphorylation of Ste5p. Although Ptc1p was found to directly bind Ste5p through a four residue motif adjacent to the phosphorylated sites, Nbp2p might also play a role in the recruitment of Ptc1p to Ste5p, something that was not addressed in this study. In particular, this study did not explain how addition of the mating pheromone promotes association of Ptc1p and Ste5p. It is possible that Ptc1p is recruited through an association between Nbp2p and Ste20p kinase, where the recruitment of Ste20p kinase to the mating pathway is dependent on the addition of mating pheromone (described in more detail in section 1.3.5). In this case Nbp2p would act as a positive regulator of the mating pathway (Figure 1.9). Under different circumstances, the Nbp2p-Ptc1p complex may function as a negative regulator of Ste20p function, since overexpression of \textit{STE20} is toxic in cells bearing an \textit{NBP2} deletion (Hruby et al., 2011).

As mentioned earlier, no biological significance has been attributed to the NbpSH3-Cla4p interaction. However, there are a number of cellular processes where the functions of Nbp2p and Cla4p may intersect (Hruby et al., 2011). One of these may be the organization of the
septin scaffold that is regulated by Cla4p and Nap1p, which both bind different domains in Nbp2p. Indeed septin organization appeared to be abnormal in the \textit{nbp2}Δ cells (Hruby et al., 2011). Another cellular process that may involve interaction between Nbp2p and Cla4p is vacuole inheritance. Both the Nbp2p-Ptc1p complex and Cla4p were demonstrated to play important roles in this process through yet unidentified mechanisms (Bartholomew and Hardy, 2009; Jin et al., 2009). It is possible that Nbp2p also plays a role in regulating the activity of Cla4p and other PAKs in the nucleus, since it interacts with the Nap1p protein that functions in nucleosome assembly. In conclusion, in addition to the well characterized roles of Nbp2p in the HOG and CWI pathways, Nbp2p is likely to perform multiple other functions some or all of which involve association with the Ptc1p phosphatase. This is supported by a large number of synthetic lethal interactions of \textit{nbp2}Δ strain with the deletions of genes in various functional categories (Tong et al., 2001).

1.3.5 \textbf{Bem1p}

Bem1p is composed of two SH3 domains, a PX domain which binds phosphinositides and recruits Bem1p to the membrane, and a PB1 protein interaction domain at its C-terminus (Figure 1.10A). It was originally identified as one of the genes important for bud emergence, and later discovered to modulate the activity of Cdc42p, which as described above, is critical for the establishment of the cell polarity (Bose et al., 2001; Irazoqui et al., 2003; Kozubowski et al., 2008; Matsui et al., 1996; Shimada et al., 2004). Bem1p regulates Cdc42p activity by binding its GEF, Cdc24p, via the PB1 domain, its effectors, Ste20p and Cla4p, via the BemSH3b domain, and Cdc42p itself also through the BemSH3b domain. As already mentioned, the BemSH3b domain contains a C-terminal extension required for folding and interaction with the Cdc42p-GTP in a PxxP independent manner. This C-terminal extension has been referred to as the Cdc42 interacting region (CI) and the SH3 domain as the SH3-CI. For simplicity, I will use the term BemSH3b throughout this thesis and only use the BemSH3-CI term in chapter 2. In order to explain how Bem1p modulates Cdc42p activity, I will give a brief description of the characterized Bem1p roles in the establishment of cell polarity and the mating pathway.

As described in section 1.3.1, the establishment of cell polarity requires the recruitment of Cdc24p to the positional landmark Bud1p GTPase. The activity of Cdc24p appears to be
inhibited by an intra-molecular interaction involving its PB1 domain (Butty et al., 2002). While the binding of Cdc24p to Bud1p partly relieves this auto-inhibition, full release requires Bem1p, which binds to the PB1 domain of Cdc24p and stabilizes its active conformation. Activated Cdc24p then activates Cdc42p, which initiates polarized growth (Figure 1.10B). Another documented role of Bem1p in polarized growth is the recruitment of Cla4p (most likely mediated by the binding of BemSH3b to ΨxPxRxAPxxP motif) and its co-localization with Cdc24p. This was observed by two studies, which both showed that this co-localization is required for the phosphorylation of Cdc24p by Cla4p. One study suggested that phosphorylation of Cdc24p by Cla4p promotes dissociation of Cdc24p from Bem1p causing cessation of Cdc42p-GTP production and termination of polarized growth (Gulli et al., 2000), while the other study disagreed with this conclusion but offered no alternative explanation for the significance of Cdc24p phosphorylation (Bose et al., 2001). Although the significance of Bem1p interaction with Cla4p (and Ste20p) during polarized growth is not fully understood, it was shown to play a critical role in the initiation of polarization in the absence of spatial landmarks, symmetry breaking (Kozubowski et al., 2008).

As mentioned earlier, the function of Bem1p is critical for symmetry breaking in the absence of the spatial cues provided by Bud1p and deletion of both BEM1 and BUD1 genes is lethal (Irazoqui et al., 2003). A recent study showed that the role of Bem1p in symmetry breaking is to link PAKs (Ste20p or Cla4p) and Cdc24p into a single complex (Kozubowski et al., 2008). Supporting this idea, a fusion of Cla4p to Cdc24p is able to bypass the requirement for Bem1p. This offered a model for symmetry breaking, where an activated Cdc42p binds a PAK and is co-localized with Cdc24p through the PAK-Bem1p-Cdc24p interaction (Figure 1.10C). In this way Cdc24p generates more Cdc42p-GTP in the vicinity of the pre-existing Cdc42p-GTP, amplifying the local concentration of activated Cdc42p, which triggers polarization once it reaches a critical level. The ability of Bem1p itself to bind Cdc42p-GTP did not appear to be sufficient for the Cdc42p recruitment. However, it may play an accessory role since disrupting the Bem1p-Cdc42p interaction caused a mild defect in the symmetry breaking ability of Bem1p.
Figure 1.10 Domain organization and function of Bem1p.

(A) Bem1p is composed of two SH3 domains, a PX domain and a PB1 domain. The N-terminal SH3 domain does not appear to recognize canonical PxxP motifs according to phage display (Table 1.2) and with the exception of the exocyst subunit Sec15p (which binds through unknown motif) no yeast targets have been identified for this domain. The second SH3 domain (subject of this study) contains a C-terminal extension required for folding and interacts both with the canonical PxxP motifs in Ste20p, Cla4p and Boi1/2p proteins, as well as with the Cdc42-GTP in PxxP independent manner. The PX domain binds phosphoinositides and also binds the membrane non-specifically through electrostatic interactions (Stahelin et al., 2007). This domain is important for the recruitment of Bem1p to the membrane. The PB1 domain is responsible for the interaction with the Cdc24p through heterodimerization with its PB1 domain. (B) During polarity establishment, Cdc24p is recruited to the sites of polarized growth through interaction with Bud1p, where Bem1p assists in stabilizing Cdc24p in the active conformation. Activated Cdc24p activates Cdc42p, which initiates the establishment of cell polarity. Bem1p is also required for Cla4p dependent phosphorylation of Cdc24p, which plays a yet undefined role during budding. (C) In the absence of the spatial cues provided by Bud1p, Bem1p initiates budding at random sites by co-localizing the activated Cdc42p-GTP with Cdc24p through PAKs Cla4p or Ste20p. This co-localization triggers activation of Cdc42p in the vicinity of the existing active Cdc42p-GTP, thus amplifying local concentration of the active Cdc42p-GTP leading to the critical levels required for the initiation of cell polarity. (D) In the mating pathway Bem1p appears to promote signal transduction through co-localization of its multiple components including Cdc24p, Cdc42p-GTP, Ste20p and Ste5p.

The role of Bem1p in the mating pathway is not too dissimilar from its role in symmetry breaking. I will first give a brief description of how the mating pathway is activated. The binding of the mating pheromones to their G-protein coupled receptors (Ste2/Ste3) releases the membrane-associated Gβγ subunit, which interacts with Far1p, Ste20p and Ste5p. Since Far1p binds Cdc24p and Ste5p binds Ste11p, this recruits Cdc24p, Ste20p and Ste11p to the membrane in the vicinity of the activated receptors. Cdc24p then activates Cdc42p, which in turn activates Ste20p, which finally activates Ste11p, which activates the MAPK pathway.
Bem1p appears to play an accessory role in this process (in a bem1Δ strain pathway induction was reduced 5-fold (Lyons et al., 1996)) by co-localizing Cdc24p, Cdc42p-GTP and Ste20p through the aforementioned interactions. It may also help to co-localize Ste20p with Ste11p, since it was reported to interact with the scaffold Ste5p through an unidentified mechanism (Lyons et al., 1996). The interaction between BemSH3b and Ste20p plays a central role, where it links Ste20p to its upstream activator Cdc42p-GTP and downstream target Ste11p. Accordingly disrupting this interaction (through either Ste20p PxxP or BemSH3b mutations) reduced signaling 2-3 fold, which is a comparable to the decrease in signaling obtained when BEM1 is deleted (Winters and Pryciak, 2005; Yamaguchi et al., 2007). Interestingly, disrupting the interaction between Bem1p and Cdc42p-GTP did not cause any further decrease in signaling in the context of the already disrupted Bem1p-Ste20p interaction (the effect of just disrupting Bem1p-Cdc42p interaction was not tested in the same assay; (Yamaguchi et al., 2007)). To me this suggested that the role of the Bem1p-Ste20p and Bem1p-Cdc42p-GTP interactions is to co-localize Cdc42p-GTP and Ste20p, leading to the activation of Ste20p by the binding of Cdc42p-GTP. However, a different conclusion was drawn by the authors of this study, who proposed that the binding of Ste20p and Cdc42p-GTP to Bem1p functions independently to promote pathway activation (Yamaguchi et al., 2007). The reason for this conclusion was based on their results demonstrating that the BemSH3b domain can not bind Ste20p (through PxxP motif) and Cdc42p (in PxxP independent manner) at the same time. In chapter 2 of this thesis I investigate these characteristics of the BemSH3b interactions and propose a different model for the function of BemSH3b in the mating pathway.

BemSH3b was also demonstrated to bind proline rich regions of Boi1p and Boi2p proteins, which are 52% identical and are functionally redundant paralogues (Bender et al., 1996; Matsui et al., 1996). The binding site of BemSH3b in Boi1p and Boi2p proteins was mapped to proline rich regions containing FVSPRRAPKPP motif in the case of Boi2p and FESPGRAPKPP motif in the case of Boi1p. The Boi2p motif closely matches the ΨxPxRxAPxxP consensus recognized by the BemSH3b and NbpSH3 domains. Boi1p motif also resembles this consensus sequence, but lacks the critical Arg at the (-3) position.
Although several studies detected the interaction between BemSH3b and Boi1p/Boi2p proteins through yeast-two-hybrid, pull-down and phage display assays, no functional significance has yet been attributed to this interaction (Bender et al., 1996; Hertveldt et al., 2006; Matsui et al., 1996). However similar to Bem1p, Boi1p and Boi2p are also involved in the establishment of cell polarity. The growth of boi1Δboi2Δ mutants is strongly impaired with the majority of cells unable to establish cell polarity or lysing with small buds (Bender et al., 1996). Furthermore, yeast-two-hybrid interactions were detected between Boi1p and the activated form of Cdc42p (not through Bem1p) and Boi2p was identified as one of the proteins important for proper localization of Cdc24p (Bender et al., 1996; Cole et al., 2009). Thus, the interaction of BemSH3b with the Boi1p and Boi2p proteins is likely to be biologically significant and I have included the Boi1/2 PxxP motifs in my studies of the BemSH3b function.

1.3.6 Cross-reactivity between Nbp2p and Bem1p SH3b domains

The cross-reactivity between the NbpSH3 and BemSH3b domains was first demonstrated by Winters et al, who showed that both SH3 domains interact with the FIPSRPAPKPP motif in Ste20p kinase (Winters and Pryciak, 2005). Since similar motifs were also found in other, then known targets of the BemSH3b domain (Boi2p and Cla4p) and the NbpSH3 domain (Pbs2p), it appeared that this cross-reactivity should extend to these sites as well. The NbpSH3 domain consensus sequence determined by phage display (Tong et al., 2002) also matched the consensus sequence ΨxPxRxAPxxP shared by these motifs (Figure 1.11). Since then, the NbpSH3 domain was also found to interact with the ΨxPxRxAPxxP motifs in Cla4p and Bck1p and bind to Skm1p most likely through the same motif (Hruby et al., 2011; Tonikian et al., 2009). The summary of the BemSH3b and NbpSH3 domain interactions obtained in other studies is shown in Table 1.3. As mentioned earlier, recognition of the same consensus motif by the BemSH3b and NbpSH3 domains is what made me interested in investigating and comparing the binding properties of these domains, as is presented in Chapters 3 and 4 of this thesis.
Table 1.3 Summary of NbpSH3 and BemSH3b interactions

<table>
<thead>
<tr>
<th>Protein</th>
<th>NbpSH3</th>
<th>BemSH3b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ste20p - PAK kinase that acts in Mating, High</td>
<td><strong>Previously detected interactions</strong>(^a,b,c)</td>
<td><strong>Previously detected interactions</strong>(^a,b,c)</td>
</tr>
<tr>
<td>Glycerol Osmolarity (HOG) and Invasive Growth MAPK pathways</td>
<td>Affinity Capture, Y2H</td>
<td>Affinity Capture, Y2H</td>
</tr>
<tr>
<td></td>
<td>Ste20p residues 469-484</td>
<td>Ste20p residues 469-484</td>
</tr>
<tr>
<td></td>
<td>(Winters and Pryciak, 2005)</td>
<td><strong>Biological relevance: Signal transduction in the mating pathway</strong> (Winters and Pryciak, 2005)</td>
</tr>
<tr>
<td>Cla4p - PAK kinase involved in the control of</td>
<td>Interaction in split-ubiquitin assay,</td>
<td>Reconstituted Complex,</td>
</tr>
<tr>
<td>polarized growth, septin ring assembly and</td>
<td>Affinity capture</td>
<td>Affinity Capture</td>
</tr>
<tr>
<td>cytokinesis</td>
<td>Cla4p residues: 801-824</td>
<td><strong>Biological relevance: Cdc24p phosphorylation by Cla4p</strong> (Bose et al., 2001)</td>
</tr>
<tr>
<td></td>
<td>(Hruby et al., 2011)</td>
<td></td>
</tr>
<tr>
<td>Skm1p - PAK kinase involved in down-</td>
<td>Affinity Capture (high throughput)</td>
<td></td>
</tr>
<tr>
<td>regulation of sterol uptake</td>
<td>(Breitkreutz et al., 2010)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Y2H (high throughput)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Tonikian et al., 2009)</td>
<td></td>
</tr>
<tr>
<td>Bck1p - The MAPKKK in the Cell Wall Integrity</td>
<td>Interaction in split-ubiquitin assay</td>
<td></td>
</tr>
<tr>
<td>(CWI) MAPK pathway</td>
<td>Affinity capture</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bck1p residues 801-824</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Hruby et al., 2011)</td>
<td><strong>Biological relevance: Requirement of Ptc1p and Nbp2p for the down-regulation of CWI pathway</strong> (Du et al., 2006)</td>
</tr>
<tr>
<td></td>
<td><strong>Biological relevance: Requirement of Ptc1p and Nbp2p for the down-regulation of CWI pathway</strong> (Du et al., 2006)</td>
<td></td>
</tr>
<tr>
<td>Pbs2p - MAPKK and scaffold protein in the</td>
<td>Affinity Capture</td>
<td></td>
</tr>
<tr>
<td>High Glycerol Osmolarity (HOG) MAPK pathway</td>
<td>Reconstituted Complex</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pbs2p residues 187-190</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Biological relevance: Recruitment of negative regulator Ptc1 to the HOG MAPK pathway</strong> (Mapes and Ota, 2004)</td>
<td></td>
</tr>
<tr>
<td>Boi2p - Adaptor protein, redundant with</td>
<td>Reconstituted complex</td>
<td></td>
</tr>
<tr>
<td>Boi1p, functions during polar growth and in</td>
<td>(Bender at el., 1996)</td>
<td></td>
</tr>
<tr>
<td>spindle midzone checkpoint</td>
<td>Y2H , Boi2p residues 436-464</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Matsui et al., 1996)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phage display(Hertveldt et al., 2006)</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) For each interaction the methods of detection are indicated: Y2H-yeast two hybrid; affinity capture - interactions detected in cell extracts; reconstituted complex- interactions detected between purified proteins.

\(^b\) Biological role of the interactions and the SH3 target residues are indicated if known.
Figure 1.11 Binding targets of the NbpSH3 and BemSH3b domains

(A) Phage display peptides recognized by the NbpSH3 domain. All of the peptides that bound to the NbpSH3 domain as determined in Tong et al. (Tong et al., 2002) are shown. An alignment of these peptides yielded the consensus sequence shown, where Ψ is a hydrophobic residue and X is any amino acid. This consensus matches known target peptides of both the Nbp2SH3 and BemSH3b domains as shown in (B). Conserved residues are highlighted: proline residues are highlighted in yellow, basic residue in red and hydrophobic residues in green.

(B) The alignment of NbpSH3 and BemSH3b binding motifs. Coloring is the same as (A).

1.4 Project overview

Although SH3 domains initially appeared limited in their interactions binding one class of targets with low specificity, numerous examples demonstrated that these domains are capable of more sophisticated functions. They were demonstrated to achieve higher intrinsic specificity through different mechanisms, interact with a variety of different targets and even bind multiple targets through different interfaces. In this thesis, I have explored the binding properties of the yeast BemSH3b and NbpSH3 domains, providing further evidence that SH3 domains can possess high intrinsic specificity and perform complex functions.

In Chapter 2, I present the solution structure the BemSH3b domain in complex with the Ste20p peptide. Since I was unable to obtain the S.cerevisiae BemSH3b at a concentration high enough for structural studies, I instead used a better behaved BemSH3b homologue from a related fungus L.elongisporus. This structure demonstrates how the extra C-terminal region interacts with the SH3 subdomain and contributes to its functional properties. I have also used
quantitative *in vitro* binding assays to characterize the interaction between BemSH3b and its interaction partners Ste20 peptide and Cdc42p. The results of this study suggest a model for how the binding of BemSH3b to Cdc42p and the PxxP sites within the Ste20p or Cla4p kinases contributes to Cdc42p-mediated activation of these PAKs.

In Chapter 3, I used quantitative *in vitro* binding assays to characterize the binding of the NbpSH3 and BemSH3b domains to six yeast peptides (Ste20, Cla4, Skm1, Boi2, Bck1, Pbs2), some of which are proven biological targets of these domains (Table 1.3). I discovered that despite recognizing the same consensus sequence, each domain possesses finely tuned and unique specificity towards this set of peptides. Furthermore, I constructed the BemSH3b domain mutants with altered specificity, demonstrating how the specificity of this domain is controlled on molecular level. Through the use of the BemSH3b domain mutants with altered specificity, I explored the importance of specificity for the *in vivo* function of the NbpSH3 and BemSH3b domains.

In Chapter 4, I compare the binding mechanism of NbpSH3 and BemSH3b domains in order to understand how these domains are able to recognize the same consensus motif, yet still retain their own unique specificity. I solved the structure of the NbpSH3 domain in complex with the Ste20 peptide and compare it to the structure of the BemSH3b-Ste20 complex solved by another group, as well as the *L. elongisporus* structure of the BemSH3b-Ste20 complex that I solved in Chapter 1. I also performed mutagenesis analysis of the target peptides and the SH3 domains to further understand their binding mechanisms. These studies provide a model of how the NbpSH3 and BemSH3b domains are able to interact with some targets with equal affinity, and bind other targets with strikingly different affinities.
Chapter 2: Structural and functional characterization of Bem1p SH3b domain

2.1 Overview

The second SH3 domain from the yeast scaffold protein, Bem1p, interacts with a PXXP-containing peptide target in a manner similar to other SH3 domains. However, unlike most SH3 domains, it also interacts with another protein, Cdc42p, utilizing a distinct binding surface, which includes a 40 residue C-terminal Cdc42p Interacting (CI) subdomain. To characterize this unique BemSH3-CI domain and its mode of interaction with both its peptide and Cdc42p targets, I have solved the solution structure of a homologue from the fungus, *L. elongisporus*, in complex with a target peptide derived from the PAK kinase, Ste20p. I have found that the BemSH3-CI domain is comprised of a subdomain displaying a typical SH3 domain fold tightly packed against the CI subdomain, which is composed of two α-helices. Examination of the structure indicates that the interaction with the Ste20 target peptide is mediated through the canonical peptide-binding interface, and additional residues contained within the CI subdomain. The interaction with Cdc42p also requires the participation of residues within the CI and SH3 subdomains. Quantitative analyses clearly demonstrated that Cdc42p and the Ste20 peptide bind the BemSH3-CI domain independently and that a tertiary complex can be formed containing these components. I also found that the affinity of Cdc42p for the BemSH3-CI domain is considerably weaker than for Cdc42p binding motif within Ste20p, which suggests a sequential model of Cdc42p binding by these two proteins.

2.2 Introduction

Although most SH3 domains do interact with PXXP-containing peptides, there are numerous examples of SH3 domains that possess additional properties, such as binding to peptides lacking a PXXP motif, binding to folded protein domains, or even binding lipids(Chan et al., 2003; Harkiolaki et al., 2003; He et al., 2007; Heuer et al., 2005; Kang et al., 2000). In most of these cases, binding to atypical targets is mediated through the canonical peptide-binding interface (Kim et al., 2008). However, there are also a few SH3 domains that utilize a completely different binding surface (Douangamath et al., 2002; Nishida et al., 2001; Tatebayashi et al., 2006). For example, the yeast Pex13p SH3 domain can simultaneously bind
Pex14p through a canonical PXXP-based interaction and Pex5p through a distinct surface located on the opposite side of the domain (Douangamath et al., 2002). Since alternative binding surfaces on SH3 domains may occur more frequently than is currently appreciated, the identification and characterization of more examples of such surfaces is an important objective.

In this chapter, I have investigated the second SH3 domain from the fungal adaptor protein, Bem1p, which has been shown to possess a unique alternative binding surface (Yamaguchi et al., 2007). This domain is an intriguing subject for investigation because it binds both PXXP-containing peptides, and uses an alternative binding surface to interact with Cdc42p (Yamaguchi et al., 2007). Binding to Cdc42p requires an additional 40 residue region at the C-terminus of the SH3 domain, which is referred to as the Cdc42 Interacting (CI) domain. The CI domain is also required for the soluble expression of the SH3 domain, suggesting that these two regions associate to form a single domain, which I will refer to in this chapter as the BemSH3-CI domain. It was previously shown through two-hybrid and pull-down assays that the BemSH3-CI domain could not simultaneously bind to the Ste20 target peptide and to Cdc42p (Yamaguchi et al., 2007).

The goals of the work described here were to gain insight into the mechanism by which the BemSH3-CI domain interacts with both Cdc42p and PXXP-containing peptides, and to illuminate the unique interaction between the SH3 and CI components. To this end, I have determined the solution structure of a homologue of the BemSH3-CI domain from the fungus, *L. elongisporus* in complex with its target peptide from Ste20p. I have also utilized quantitative *in vitro* binding studies to investigate the binding surfaces of this domain and interactions between them. This work provides important new insight into the means by which the multiple interaction modules of Bem1p are coordinated to elicit appropriate regulation of yeast signaling pathways. In addition, by comparing the BemSH3-CI interaction surfaces to other SH3 domains, I have identified common alternative binding surfaces among diverse SH3 domains.
2.3 Materials and Methods

2.3.1 NMR spectroscopy and structure calculation

Residues 2-117 from *L. elongisporus* Bem1p (LeBemSH3-CI domain) were cloned into pET21d (Novagen) and expressed with a C-terminal 6-His tag. Purification was carried out using Ni²⁺-NTA agarose (QIAGEN) under native conditions. A peptide corresponding to Ste20p residues 468-484 was chemically synthesized (CanPeptide Inc). For NMR experiments, two samples were prepared. Sample A contained ¹⁵N ¹³C labeled LeBemSH3-CI at 0.7 mM and unlabeled Ste20 peptide at 1.5 mM in 50 mM Hepes, pH 6.8, 100 mM NaCl, 1 mM DTT, 0.5 mM EDTA, 0.5 mM PMSF, 0.05% NaN₃. Sample B contained ¹⁵N ¹³C labeled LeBemSH3-CI at 0.5 mM and unlabeled Ste20 peptide at 0.5 mM in 50 mM Phosphate buffer, pH 6.8, 100 mM NaCl, and 0.05%NaN₃. Backbone and side chain chemical shifts for the LeBemSH3-CI domain were assigned using standard triple resonance experiments (Kay, 1995) (HNCO, HN(CO)CA, CBCA(CO)NH, HNCACB, C(CO)NH, HCCH-TOCSY and HCCH-COSY) with sample A. Inter-proton distances for the LeBemSH3-CI domain were obtained with ¹⁵N-edited and ¹³C-edited NOESY spectra (sample A). Chemical shift assignments and structural restraints for the Ste20 peptide were derived from double-half-filtered 2D ¹H-¹H TOCSY, COSY and NOESY experiments (Gemmecker G., 1992; Ikura and Bax, 1992) (sample B). Intermolecular NOEs between LeBemSH3-CI domain and Ste20 peptide were obtained from 3D NOESY experiments (sample A) and ¹³C-edited ¹³C-filtered 3D NOESY (Zwahlen et al., 1997) (sample B). All NMR experiments were carried out at 25°C on a Varian INOVA 500 or 800 MHz spectrometers equipped with pulsed field gradients at the Quebec/Eastern Canada High Field NMR Facility.

All data were processed using NMRPipe (Delaglio et al., 1995) and subsequent analysis performed using Sparky program (Kneller, 1997). Structure calculations were performed using CYANA 2.1 (Guntert, 2004). The input constraints included dihedral angle restraints derived using TALOS (Cornilescu et al., 1999) and hydrogen bond upper and lower distance limits derived from examination of short range NOE data. Manually assigned NOE peaks were fixed in the first cycle of structure calculation and peak assignments. In the remaining cycles, all NOE peak assignments, with the exception of several high confidence intermolecular NOEs which remained fixed, were made by CYANA. The 20 lowest energy structures, of 100
calculated in the final iteration were used to derive the structure. Structural coordinates and NMR restraints have been deposited in the Protein Data Bank with accession number 2KYM. NMR assignment data was deposited in BioMagResBank with accession number 16970.

2.3.2 Protein expression and purification for in vitro binding assays
The LeBemSH3-CI (residues 2-117) and yBemSH3-CI (residues 155-272) domains were expressed from pET21d vector with a C-terminal 6-His tag (sequence boundaries of LeBemSH3-CI within full length protein are different from those in the yeast homologue because the LeBem1p homologue lacks the first SH3 domain found in yeast Bem1p). The Ste20 peptide (468-484) used in binding assays was expressed from pET21d-based vectors as C-terminal bacteriophage λ cI protein fusions as previously described (Maxwell and Davidson, 1998). LeCdc42p Q61L was expressed from a pET15b vector with an N-terminal 6-His tag. yCdc42p Q61L was expressed from pGEX vector with an N-terminal GST tag. The His-tagged proteins were purified using Ni²⁺-NTA agarose (QIAGEN) and GST-tagged proteins were purified using Glutathionine Sepharose 4B (GE Healthcare).

2.3.3 Fluorescence-based binding assays
For peptide-binding assays, BemSH3-CI domain at concentration of 1 μM was titrated with Ste20 peptide fusion protein. Binding was monitored by measuring tryptophan fluorescence ($\lambda_{\text{excitation}} = 295\text{nM}$, $\lambda_{\text{emission}} = 326\text{nM}$), which increased upon binding. The experiments were carried out in 50 mM phosphate pH 6.8, 100 mM NaCl. For the Cdc42p binding assay, Cdc42p proteins were pre-loaded with MANT-GTP (Invitrogen). Proteins at a concentration of 4 μM were incubated with 0.6 μM MANT-GTP at 25°C in 50 mM Tris, 100 mM NaCl, 10 mM MgCl₂, 1 mM DDT buffer. The loading of MANT-GTP was monitored by measuring the increase in MANT-GTP fluorescence ($\lambda_{\text{excitation}} = 360\text{nM}$, $\lambda_{\text{emission}} = 440\text{nM}$), which reached a maximum fluorescence after approximately 2 hours. Cdc42p:MANT-GTP complexes were then titrated with BemSH3-CI domains and binding was detected by monitoring the quenching of the MANT-GTP fluorescence signal. All titrations and fluorescence measurements were carried out on an Aviv ATF105 spectrofluorometer equipped with a Microlab 500 series automated titrator. Dissociation constants were calculated as previously described (Maxwell and Davidson, 1998).
2.3.4 Pull-down binding assays

The Ste20 peptide fusion protein was covalently linked to Affi-Gel 10 resin (Bio-Rad Laboratories) by incubating 25 mg of protein/1 ml of resin according to the manufacturer’s instructions. 40 μL of the saturated Ste20 peptide fusion resin was incubated with 4 nmol of BemSH3-CI domains, 4 nmol of Cdc42p proteins or both. The total volume of the reactions was 200 μL. Incubations were carried out at 4°C for 30 minutes in 50 mM Tris, pH 7.5, 100 mM NaCl, 10 mM MgCl2, 1 mM DTT buffer. The resin was washed 3 times with 300 μL of binding buffer. Bound proteins were eluted by re-suspending the resin in 30 μL SDS loading buffer and visualized by Coomassie staining of 15% SDS-PAGE gel.

2.4 Results

2.4.1 The structure of the BemSH3-CI domain from L. elongisporus in complex with a PXXP-containing peptide from Ste20p

To understand the mechanism by which the BemSH3-CI domain interacts with PXXP-containing peptides and Cdc42p, I attempted to determine the solution structure of this domain from yeast Bem1p. Although a variety of different C-terminally tagged constructs were tested with varying domain boundaries, I was unable to identify one that could be purified and maintained at a concentration high enough for structural studies using NMR spectroscopy. Therefore, a summer student Wen Chi Yin who was working under my supervision tested the solubility of BemSH3-CI domain from 7 different fungal species and found that the domain from Lodderomyces elongisporus was highly soluble. Since the yeast and L. elongisporus BemSH3-CI domains are 50% percent identical in sequence, this domain was expected to provide an accurate model for the yeast domain structure. In addition, it possessed very similar functional properties to the homologous yeast domain (see below).

Using standard NMR experiments (Kay, 1995), I determined the solution structure of the L. elongisporus BemSH3-CI (LeBemSH3-CI) domain in complex with the previously identified (Winters and Pryciak, 2005) PXXP-containing target peptide from Ste20p. The structure was constrained based on intra-peptide, intra-SH3 and intermolecular NOEs and dihedral angle restraints derived from TALOS (Table 2.1). Residues 1-101 of the LeBemSH3-CI domain and residues spanning Phe(-7) to Pro(3) of the peptide (numbering of the SH3 according to Larson et al (Larson and Davidson, 2000) and peptide according to Lim et al (Lim et al., 1994)) were
well defined with a root mean square deviation (RMSD) of 0.5 Å among the backbone atoms of the 20 lowest energy structures of these regions (Figure 2.2A, Table 2.1). The last 14 residues of LeBemSH3-CI domain (102-115) were determined to be unstructured based on the lack of long range NOEs and negative $^1$H-$^{15}$N heteronuclear NOEs (Figure 2.1).

### Table 2.1 Structural statistics for LeBemSH3-CI:Ste20 peptide complex

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input dihedral angles $^a$ ($\phi, \psi$)</td>
<td>RMSD from mean structure (Å)$^c$</td>
</tr>
<tr>
<td>NOE upper distance</td>
<td>Backbone atoms</td>
</tr>
<tr>
<td>Intraresidual</td>
<td>0.50</td>
</tr>
<tr>
<td>Medium-range (1≤</td>
<td>i-j</td>
</tr>
<tr>
<td>Long-range (</td>
<td>i-j</td>
</tr>
<tr>
<td>Intermolecular</td>
<td>RMS deviation from ideal geometry</td>
</tr>
<tr>
<td>Hydrogen bond $^b$</td>
<td>Distance (Å)</td>
</tr>
<tr>
<td>Lower</td>
<td>0.004</td>
</tr>
<tr>
<td>Upper</td>
<td>Angles (degrees)</td>
</tr>
<tr>
<td>Structure Calculation</td>
<td>0.7</td>
</tr>
<tr>
<td>Cyana Target Function</td>
<td>Ramachandran plot</td>
</tr>
<tr>
<td>1$^a$ Cycle</td>
<td>Favored regions</td>
</tr>
<tr>
<td>Final</td>
<td>77.1%</td>
</tr>
<tr>
<td>RMSD from mean structure (Å)$^c$</td>
<td>Additionally allowed regions</td>
</tr>
<tr>
<td>1$^a$ cycle Backbone/Heavy atoms</td>
<td>22.7%</td>
</tr>
<tr>
<td>NOEs</td>
<td>Generously allowed regions</td>
</tr>
<tr>
<td>% unassigned</td>
<td>0.1%</td>
</tr>
<tr>
<td>% long range discarded</td>
<td>Disallowed regions</td>
</tr>
<tr>
<td></td>
<td>0.0%</td>
</tr>
</tbody>
</table>

$^a$ Dihedral angles restraints were derived using TALOS(Cornilescu et al., 1999)

$^b$ Based on the predicted hydrogen bonds that were derived from the examination of secondary structure NOEs

$^c$ The RMSD is indicated for well defined regions. LeBemSH3-CI domain, residues 1-101, numbering according to Larson et al.(Larson and Davidson, 2000); Ste20 peptide, residues (-7)-(3), numbering according to Lim et al.(Lim et al., 1994)

The SH3 subdomain of LeBemSH3-CI domain displays a typical SH3 domain fold, and overlays well with other SH3 domain structures. For example, it can be superimposed onto the crystal structure of the canonical SH3 domain from the Fyn tyrosine kinase with an RMSD of 1.6 Å (Figure 2.2D), which is a similar deviation as is seen when other SH3 domains are compared (Larson and Davidson, 2000). The CI subdomain is composed of a loop followed by two $\alpha$-helices that wrap around the SH3 subdomain and an unstructured C-terminal tail (Figure 2.2B). The interface between the two subdomains reveals extensive interactions that bury 1840 Å$^2$ (Figure 2.2 E and F). The hydrophobic CI residues buried in the interface stand out as partially (Val77, Ile80, Phe84) or highly conserved (Ile81, Ile86, Val89 and Trp92) in
an alignment of Bem1p homologues (Figure 2.3). The SH3-CI interface can be divided into two major regions of hydrophobic contacts. One region is composed of CI residues Val77, Ile81, Phe84 (CI Helix 1) and Ile86 that pack against residues Leu2, His31, Phe37, Val52, Ile57 of the SH3 domain (Figure 2.2E). Helix 1 is very intimately packed against the SH3 subdomain, and this interaction creates an enlarged hydrophobic core. The second region contains CI residues Val89 and Trp92 (CI Helix 2) that contact SH3 domain residues Cys29, Ala30, Tyr32, Ile38 and Pro47 (Figure 2.2F). SH3 subdomain residues Leu2, Val52, Ile57 and Ile38 are highly conserved as hydrophobic residues among Bem1p homologues (Figure 2.3) and are completely buried by the CI subdomain. The extensive hydrophobic packing interaction between the SH3 and CI subdomains likely accounts for the insolubility of the SH3 subdomain when expressed on its own (Yamaguchi et al., 2007).

**Figure 2.1 Heteronuclear ¹H−¹⁵N NOEs of the LeBemSH3-CI domain**

Heteronuclear ¹H−¹⁵N NOEs (Farrow et al., 1994) of LeBemSH3-CI domain. The ratio of peak intensities collected with and without ¹H saturation ¹H−¹⁵N axis is plotted against position number in the LeBemSH3-CI domain. Low numbers are indicative of dynamics on the fast ns-ps timescale. Negative values obtained for C-terminal residues 102-115 demonstrate high mobility in this region.
Figure 2.2 The structure of LeBemSH3-CI:Ste20 peptide complex

(A) An overlay of the 20 lowest energy structures (RMSD 0.5 Å) is shown. (B) Ribbon diagram of the lowest energy structure. The unstructured C-terminal region (residues 102-115) is omitted. (C) An overlay of LeBemSH3-CI (blue) and yBemSH3-CI (pink, PDB ID: 2rqw) domains (RMSD 1.6 Å). (D) An overlay of LeBemSH3 subdomain (blue) and the Fyn SH3 (orange, PDB ID: 1shf) domains (RMSD 1.9 Å). (E) The intramolecular interaction between BemSH3 subdomain and CI helix 1 extends the hydrophobic core of the BemSH3 domain to include the residues at the SH3-CI interface. The hydrophobic residues involved in the intramolecular interaction are shown, the numbering of equivalent residues of the yBemSH3-CI domain is indicated in parentheses. SH3 domain hydrophobic core residues (Larson and Davidson, 2000) are colored magenta, buried CI subdomain residues are in green, and other buried SH3 subdomain residues not normally found in the hydrophobic core are in blue. (F) Intramolecular interaction of BemSH3 subdomain with CI helix 2. Labeling and residue coloring are the same as in (E).
Figure 2.3  An alignment of BemSH3-CI domain fungal homologues

An alignment of BemSH3-CI domain fungal homologues. The secondary structure of LeBemSH3-CI domain is indicated above the alignment. The standard SH3 domain numbering is used (Larson and Davidson, 2000). Residues that are involved in peptide binding (p), Cdc42p binding (c), and the intramolecular SH3-CI interaction (i) are indicated above the sequence. The peptide binding surface (Surface I or II) on which a residue is located is indicated below the alignment. Positions with greater than 67% conservation are colored according to amino acid properties: hydrophobic (light green), polar (dark green), positive (red), negative (purple), aromatic and His (light blue), Pro (yellow), Gly (orange).

Ste20 peptide residues spanning residues Phe(-7) to Pro(3) are well defined in the structure; most residues (with the exception of Ser(-4) and Pro(-2)) have multiple intermolecular NOEs indicating that within the binding interface almost every peptide residues is involved in binding. The core RxxPxxP motif of the peptide interacts in a canonical manner for SH3 domains, where Pro(0) and Pro(3) interact with the conserved aromatic residues (Tyr8, Phe10, Trp36, and Phe54 and Pro51) and the Arg(-3) side chain packs against Trp36 and interacts with the RT-loop residues Asp16 and Glu17 (Figure 2.4B). Interestingly, an additional contribution to binding is made by CI residue Tyr72, which contacts residues Lys(1) and Pro(2) of the peptide. Residues in the extended region of the peptide ((-7) to (-4)) interact with the variable surface of the SH3 subdomain (Figure 2.4B). Significantly, the Phe(-7) and Pro(-5) side chains make a large contribution to binding, as indicated by a high number of intermolecular NOEs (36 for Phe(-7) and 29 for Pro(-5)) and the complete burial of these residues upon binding. Pro(-5) packs against SH3 domain residues Tyr32, Trp36, Leu49 and
CI residue Val89, whereas Phe(-7) fits into a pocket between the SH3 and CI domains that is lined by SH3 domain residues Ile38, Pro47, Leu49 and CI residues Val89 and Lys93. The significance of SH3 domain-peptide interactions will be discussed in more detail in Chapter 4, where the binding of Ste20 peptide to the NbpSH3 and BemSH3-CI (BemSH3b) domains is compared.

Figure 2.4  Binding of the Ste20 peptide to the BemSH3-CI domain

(A) Interaction of the Ste20 peptide core region. Residues involved in the interaction are shown and the equivalent residues of the yBemSH3-CI domain are indicated in parentheses. SH3 subdomain residues are colored blue and CI subdomain residues are in green. Peptide residues are in magenta. (B) Interaction of the Ste20 peptide extended region. Same coloring scheme as in (B).

As I was completing my work on this project, another group published the solution structure of the yeast BemSH3-CI (yBemSH3-CI) in complex with the Ste20 peptide (Takaku et al., 2010). As can be seen in Figure 2.2C, the LeBemSH3-CI and yBemSH3-CI domains are very similar in structure and overlay with an RMSD of 1.9 Å over the well defined regions of their structures. A similar set of residues, most of which are conserved between the two domains (Figure 2.3) were identified as important for the formation of SH3-CI interface and peptide binding in the analysis of yBemSH3-CI structure. However the SH3-CI interface in yBemSH3-CI is not as extensive, burying approximately 1263 Å² compared to 1844 Å² in the LeBemSH3-CI domain. Consistent with this, I measured the melting temperature of LeBemSH3-CI (72 °C) to be significantly higher than the melting temperature of yBemSH3-CI domain (50 °C), which is probably due to the tighter network of hydrophobic interactions in the L. elongisporus SH3-CI interface. In the study on yBemSH3-CI, the domain was
expressed and purified as an N-terminal fusion to GST, which may explain why this group was able to obtain a sample at high enough concentration for NMR studies while I was not.

2.4.2 Mutagenesis and quantitative binding studies to delineate residues important for the interaction of the BemSH3-CI domain and Cdc42p

To quantitatively characterize the interaction between the BemSH3-CI domain and Cdc42p, I utilized a fluorescence based assay. Cdc42p was pre-loaded with a fluorescent GTP analog, MANT-GTP. Since Cdc42p:MANT-GTP complexes exhibit altered fluorescence upon binding of effectors (Nomanbhoy and Cerione, 1999), I could monitor binding between BemSH3-CI and Cdc42p:MANT-GTP by detecting changes in fluorescence (Figure 2.5A). I determined that the yBemSH3-CI domain bound to yCdc42p with a $K_d$ value of 5 µM, and LeBemSH3-CI domain bound to LeCdc42p with a $K_d$ value of 9 µM (Figure 2.5A). These values are similar to that measured for the interaction of *S. pombe* BemSH3-CI and Cdc42p ($K_d$ value = 3 µM) as determined by Isothermal Titration Calorimetry (Wheatley and Rittinger, 2005).

I used site-directed mutagenesis to delineate residues important for the interaction between BemSH3-CI domain and Cdc42p (Figure 2.5B). The LeBemSH3-CI domain was employed for these assays because I observed a greater change in fluorescence when this domain interacted with Cdc42p as compared to the yeast domain (Figure 2.5A), which allowed a more accurate calculation of $K_d$ values for weak interactions. A previous study identified a single mutation in the yBemSH3-CI domain (N253D) that disrupted binding to Cdc42p in pull down and yeast 2-hybrid assays (Yamaguchi et al., 2007). In yBemSH3-CI (Takaku et al., 2010) and LeBemSH3-CI domains this substitution is located in the unstructured region of CI subdomain. I found that an equivalent substitution in LeBemSH3-CI domain (S103D) also disrupted binding to LeCdc42p, confirming that yBemSH3-CI and LeBemSH3-CI domains interact with Cdc42p in a similar manner. However, the Ser103 side chain does not appear to play a role in binding, as the S103A substitution had no effect. Substitutions with Ala of conserved residues in this region (Y99A, I104A, and I106A) did cause significant decreases in LeCdc42p binding of 2-fold or more, confirming the involvement of this unstructured region in the binding reaction. The BemSH3-CI domain contains a highly conserved and unusual sequence at its distal loop, which suggested that this region could also be involved in Cdc42p binding. Consistent with this idea, substitution of residues in this region (K40A, I42A and
L44A) caused a considerably larger reduction of binding affinity to LeCdc42p as compared to the substitutions in the CI C-terminal tail (Figure 2.5B). For example, the L44A substitution decreased the affinity to greater than 100 μM. The substitution of residues in the folded region of LeBemSH3-CI domain (Y99A, K40A, I42A and L44A) did not have a significant effect on domain stability or affinity for Ste20 peptide, demonstrating that the detrimental effect of these substitutions on LeCdc42p binding was not due to the misfolding of the domain (Figure 2.6). Thus, the BemSH3-CI domain uses at least two distinct surfaces to interact with Cdc42p, where one surface includes the C-terminal unstructured residues of the CI subdomain and another surface is primarily composed of the SH3 subdomain Distal loop residues (Figure 2.5C).

Figure 2.5  Binding of Cdc42p to the BemSH3-CI domain

(A) A Titration of Cdc42p with the BemSH3-CI domain. Representative curves obtained for yeast and *Le. elongisporus* titrations are shown. The average $K_d$ values are indicated in parentheses. Binding affinities were calculated by titrating Cdc42p:GTP-MANT complexes with the BemSH3-CI domains and detecting binding through the change in the MANT-GTP fluorescence ($\lambda_{\text{excitation}} = 360\text{nm}$, $\lambda_{\text{emission}} = 440\text{nm}$). (B) Substitution of residues in the Distal Loop and unstructured C-terminal region of LeBemSH3-CI domain impair binding to LeCdc42p. Binding affinities of LeCdc42p to wild type and mutant LeBemSH3-CI domains are shown. (C) The structure of the LeBemSH3-CI domain showing the residues involved in LeCdc42p binding. Equivalent yeast residues are indicated in parentheses. (D) Substitution of residues in LeCdc42p switch 1 region impairs LeBemSH3-CI domain binding. Binding affinities obtained for BemSH3-CI domain interactions with wild type and mutant LeCdc42p proteins are shown. The $K_d$ values were calculated as for wild type Cdc42p.
To identify Cdc42p residues required for BemSH3-CI domain binding, I substituted residues in the effector loop of LeCdc42p (F37A, D38A and Y40C). This region is important for the binding of Cdc42p to many of its target proteins (Bishop and Hall, 2000). The Y40C substitution had a significant effect on binding, reducing the affinity at least 17-fold, while the D38A and F37A substitutions had a more moderate effect increasing the $K_d$ by several fold (Figure 2.5D). The same substitutions have been shown to disrupt the binding of Cdc42p homologues to the CRIB motif, which is conserved in many Cdc42p effectors, including Ste20p kinase (Owen et al., 2000). These data demonstrate that though the BemSH3-CI domain lacks a CRIB motif, it interacts with the same surface on Cdc42p that is bound by the CRIB motif.

Figure 2.6 Stability and peptide binding activity of LeBemSH3-CI mutants

(A) Temperature melts of the wild type LeBemSH3-CI domain and LeBemSH3-CI domains carrying substitutions that affect LeCdc42p binding. Although all LeBemSH3-CI mutants show slightly decreased stability relative to wild type, they all remain folded until at least 50 °C, which is much higher than the temperature of 25°C at which Cdc42p binding assays were performed. The loss of secondary structure was monitored using circular dichroism spectroscopy as previously described (Maxwell and Davidson, 1998). Denaturation of the LeBemSH3-CI domain was not reversible; thus, the data from these curves were not fit to determine precise melting temperatures. (B) Substitutions in LeBemSH3-CI that impair LeCdc42p binding have no effect on Ste20 peptide binding. The affinities of wild type and mutant LeBemSH3-CI domains towards Ste20 peptide are indicated.

2.4.3 Cdc42p and Ste20 peptide binding sites on BemSH3-CI domain are distinct and are not allosteric

The Ste20 peptide and Cdc42p binding sites are located on two different sides of the BemSH3-CI domain, suggesting that a ternary complex of BemSH3-CI domain, Cdc42p and
the Ste20 peptide should be able to form. Supporting this supposition, I showed by a pull-
down assay that the Ste20 peptide was able to bind a complex of BemSH3-CI domain and
Cdc42p (Figure 2.7 A and B). This was true of both the yeast and L. elongisporus domains.
These results disagree with a previous study that used both pull-down and yeast-two-hybrid
assays to show that the BemSH3-CI domain could not simultaneously bind Cdc42p and the
Ste20 peptide (Yamaguchi et al., 2007).

Figure 2.7 Simultaneous binding of the Ste20 peptide and Cdc42p by the BemSH3-CI domain.
(A) Pull-down assays with purified L. elongisporus proteins. The proteins indicated were mixed with Ste20
peptide fusion protein that had been covalently bound to Affi-Gel resin. Ste20 peptide resin was incubated with
BemSH3-CI domain, Cdc42p or both. The Ste20 peptide fusion migrates as a monomer and a dimer on the 15%
SDS-PAGE gel, thus displaying 2 bands in the lane with no other proteins added. (B) The same experiments as
shown in (A) were carried out with yeast proteins. Yeast Cdc42p was expressed as a fusion to GST. The upper
band seen in the gel corresponds to the Cdc42p-GST fusion, and the lower band corresponds to the break down
products, GST and Cdc42p. (C) Presence of the Ste20 peptide does not affect the binding affinity of LeBemSH3-
CI domain for LeCdc42p. LeCdc42p was titrated with LeBemSH3-CI domain alone or in the presence of excess
amount of Ste20 peptide. (D) Binding of the LeBemSH3-CI domain to Ste20 peptide is not affected by the
presence of LeCdc42p. The LeBemSH3-CI domain was titrated with Ste20 peptide alone or in the presence of
excess LeCdc42p.

To accurately assess whether binding of the Ste20 peptide to BemSH3-CI domain reduces its
affinity for Cdc42p, I determined the affinity of the LeBemSH3-CI for LeCdc42p in the
presence of an excess amount of the Ste20 peptide. As shown in Figure 2.7C, no change in affinity was observed under these conditions. Similarly, the affinity of the Ste20 peptide for the LeBemSH3-CI domain was not altered by the presence of excess levels of Cdc42p (Figure 2.7D). These data conclusively demonstrate that the two binding sites on the BemSH3-CI domain function independently.

**2.5 Discussion**

In this study, I have shown that the BemSH3-CI domain is comprised of a subdomain displaying a typical SH3 domain fold coupled in a novel manner to the helical CI subdomain. These subdomains are tightly associated through hydrophobic packing interactions, particularly where Helix I of the CI subdomain forms a common hydrophobic core with the SH3 subdomain (Figure 2.2E). Although the peptide interaction surface of the BemSH3-CI domain is similar to other SH3 domains, peptide-binding is enhanced by interactions between the N-terminal residues of the target peptide and residues in the CI domain. In particular Phe(-7) of the bound peptide is tightly packed between the SH3 and CI domains (Figure 2.4B) and is likely to make significant contribution to binding. I have also clearly shown that optimal binding of Cdc42p requires both the C-terminal unstructured region of the CI subdomain and residues within the highly conserved Distal loop of the SH3 subdomain. Notably, substitutions in the Distal loop cause a much greater decrease in Cdc42p-binding than those in the C-terminal region. Since previous studies investigating the *in vivo* importance of the BemSH3-CI domain interaction with Cdc42p (Kozubowski et al., 2008; Yamaguchi et al., 2007) both utilized an amino acid substitution (N253D in yeast Bem1p) that only partially reduced Cdc42p-binding in our assays, further studies using the more severely perturbing substitutions tested here (e.g. L44A) could provide additional insight. My structural determination of the LeBemSH3-CI and delineation of the peptide and Cdc42p binding surfaces are consistent with a recent NMR study on the yeast BemSH3-CI domain (Takaku et al., 2010). However, my quantitative analyses of Cdc42p binding have provided important new insight into the energetic contributions of individual residues on the binding surfaces of this domain and its interaction partners.
In contrast to a previous study (Yamaguchi et al., 2007), here I found that a ternary complex could be formed comprised of BemSH3-CI domain, the Ste20 peptide, and Cdc42p (Figure 2.7). The formation of this ternary complex was also demonstrated by Takaku et al. (Takaku et al., 2010). The failure of Yamaguchi et al. to observe this complex may be due to their use of slightly different constructs (e.g. their BemSH3-CI domain construct was 15 residues longer at its N-terminus and was utilized in the form of a GST-fusion). Yamaguchi et al. also reported that Cdc42p is capable of binding both the Ste20p CRIB motif and BemSH3-CI domain simultaneously, and that a Cdc42p substitution (Y40C) that impairs binding to the CRIB motif does not affect the binding to the BemSH3-CI domain. However, I found that the same substitution in LeCdc42 and others in the switch 1 region (F37A and D38A), which are known to impair binding to the CRIB motif, also markedly reduced binding to LeBemSH3-CI domain (Figure 2.5D). These results imply that the CRIB motif and BemSH3-CI domain interact with the same surface on Cdc42p. In agreement with this conclusion, Takaku et al. showed that addition of the Ste20p CRIB motif displaces Cdc42p from BemSH3-CI domain (Takaku et al., 2010) and the same result was obtained in a separate study focused on the Bem1p homologue of S. pombe, Scd2 (Endo et al., 2003). Furthermore, the latter study also showed that, consistent with our results, the Cdc42 D38A substitution impaired binding to the Scd2 SH3-CI domain. In another report, substitution of several switch I residues in yeast Cdc42p (including F37G, D38A and Y40C) impaired binding to full length Bem1p and the same substitutions affected Ste20p binding in an almost identical manner (Gladfelter et al., 2001). Given our results and the balance of the other published data, I conclude that the BemSH3-CI domain can bind to the Ste20 peptide and Cdc42p simultaneously. On the other hand, Cdc42p employs the same surface to bind both the BemSH3-CI domain and Ste20p CRIB motif, thus, Cdc42p can bind only one of these proteins at a time.

In this study I measured the affinities between BemSH3-CI domains and Cdc42p from yeast and L. elongisporus to be 5 μM and 9 μM respectively (Figure 2.5A), which are similar to the affinity of 3 μM obtained for this interaction with S. pombe proteins. These affinities are significantly weaker than those between Cdc42 and its other binding partners. In particular, the interaction between Cdc42 and CRIB motifs from a variety of Cdc42 effectors display $K_d$ values in the low nM range (Owen et al., 2000). For example, the Ste20p CRIB motif of C. albicans binds its cognate Cdc42p with a $K_d$ value of 43 nM (Su et al., 2005). Since it is
conserved across several fungal species, the lower affinity of BemSH3-CI:Cdc42p interaction as compared to the CRIB:Cdc42p interaction is likely to be functionally important. This difference in affinities may allow Cdc42p to be transferred from the BemSH3-CI domain to Ste20p CRIB, once the latter becomes available for binding (Figure 2.8).

Figure 2.8  A model for BemSH3-CI domain mediated activation of Ste20p kinase.

Cdc42p-GTP binds to the BemSH3-CI domain despite a relatively low affinity because it is locally present at a high concentration due to the interaction of Bem1p with Cdc24p (GEF of Cdc42p). Ste20p kinase is recruited in close proximity to Cdc42p owing to the high affinity of its PXXP motif (0.04 μM) for the BemSH3-CI domain. Upon the release of the Ste20p CRIB motif from an intramolecular interaction, Cdc42p is transferred from the low affinity BemSH3-CI site (4.6 μM) to the high affinity Ste20p CRIB motif (0.04 μM), leading to the activation of Ste20p kinase. The affinities of yeast BemSH3-CI domain interaction with Ste20 peptide (Chapter 3) and Cdc42p obtained in this thesis are shown. The affinity between \textit{C. albicans} Ste20p CRIB and Cdc42p is indicated for the Ste20p-Cdc42p interaction (Su et al., 2005) as this affinity has not been measured for the yeast proteins.

Although the affinity of Cdc42 for the isolated PAK kinase CRIB domain is high, the structure of the PAK1 kinase shows that the CRIB domain is occluded from Cdc42-binding by inhibitory interactions when it is within context of the full length kinase (Figure 1.7B, (Lei et al., 2000)). Consistent with this structural picture, full length human αPAK binds to Cdc42 with 40-fold lower affinity than the isolated CRIB domain (Buchwald et al., 2001), presumably due to the requirement for the CRIB domain to be released from the inhibitory interactions before Cdc42-binding can occur. I hypothesize that the role of the BemSH3-CI domain is to co-localize Ste20p and Cdc42p, so that release of Ste20p inhibitory interactions can be directly coupled to Cdc42p binding and subsequent activation of Ste20p. The high affinity of the Ste20 peptide for the BemSH3-CI domain ensures that Ste20p is co-localized
with Bem1p-bound Cdc42p at high frequency. On the other hand, the relatively low affinity of 
the BemSH3-CI domain for Cdc42p facilitates its transfer to the tighter-binding CRIB motif 
of Ste20p when this motif becomes exposed (Figure 2.8). This model would also apply to 
Cla4p kinase that is activated by Cdc42p and is a known target of BemSH3-CI domain.

In addition to the canonical PXXP-containing peptide binding surfaces, the BemSH3 
subdomain also possesses distinct surfaces interacting with the CI subdomain and with 
Cdc42p. Although few SH3 domains have been demonstrated to utilize such alternative 
binding surfaces, there is a striking overlap between the alternative binding surfaces identified 
on the BemSH3 subdomain and those found on other SH3 domains. For example, the 
BemSH3 subdomain surface bound by CI helix 1 (Figure 2.2) is in the same position as the 
Vav2 SH3 domain surface that participates in an atypical interaction with the Grb2 SH3 
domain (Figure 2.9A). It can be seen that residues at equivalent positions in the primary 
sequence of these two domains comprise their alternative binding surfaces (Figure 2.9 A and 
D). In a similar manner, the BemSH3 subdomain interacts with helix II of the CI subdomain 
(Figure 2.2F) utilizing the same surface as that employed by the ADAP SH3 domain to 
engage an α-helix positioned at its N-terminus (Figure 2.9 B and D). Another similarity 
between the BemSH3-CI and ADAP SH3 domains is that the extra helical regions extend the 
canonical binding surface. In the case of ADAP SH3, a lipid is bound by an interface of the 
canonical binding surface and residues in the N-terminal helix residues (Heuer et al., 2005; 
Heuer et al., 2006). Our study and previous work (Takaku et al., 2010) has defined the Cdc42p 
interaction surface to include the Distal loop and residues on strands 1 and 2 of the SH3 
subdomain. This binding surface coincides remarkably well with the alternative surface of the 
yeast Pex13p SH3 domain involved in Pex5p binding (Figure 2.9 C and D). Residues in the 
Distal loop of the yeast Sho1p SH3 domain (Figure 2.9) have also been shown to play an 
important functional role in binding to the Ste11p/Ste50p complex (Tatebayashi et al., 2006).
Figure 2.9  Comparison of the LeBemSH3-CI domain alternative binding surfaces with those of other SH3 domains

(A) An overlay of the LeBemSH3-CI domain (SH3-blue, CI-green) with the Vav2 SH3 (magenta)-Grb2 SH3 (orange) complex (PDB ID: 1gcq). Residues at the BemSH3:BemCI (helix I) subdomain interface and the Vav2 SH3:Grb2 SH3 domain interface that occupy equivalent positions in the BemSH3 and Vav2 SH3 domains are indicated. (B) An overlay of the LeBemSH3-CI (SH3-blue, CI-green) and ADAP SH3 (SH3-magenta, N-terminal sequence-orange, PDB ID:1ri9) domains. Residues involved in the intramolecular binding surfaces of BemSH3-CI (helix 2) and ADAP SH3 are shown, residues with equivalent positions in the binding interfaces of BemSH3 and ADAP SH3 domains are labeled. (C) An overlay of the LeBemSH3 (blue) and the Pex13p SH3 (magenta, PDB ID:1n5z) domains. The LeBemSH3 residues involved in the interaction with Cdc42p (demonstrated in this work and Takaku et al. (Takaku et al., 2010) to affect Cdc42p binding when mutated) are shown in orange and labeled. Equivalent yeast residues are indicated in parentheses. Pex13p residues involved in the noncanonical interaction with Pex5p (demonstrated to affect Pex5p binding upon substitution (Douangamath et al., 2002)) are shown in red. (D) Alignment of the SH3 domains. Residues that function in alternative surfaces are highlighted in the following manner: blue, residues in the BemSH3 subdomain that interact with CI helix 1 and in the Vav2 SH3 domain interact with Grb2 SH3 domain; red, residues in the BemSH3 subdomain that interact with CI helix2 and in the ADAP hSH3 subdomain interact with the extra N-terminal helix; green, residues in the BemSH3 subdomain that are involved in Cdc42p binding, in Pex13p SH3 domain are involved in Pex5p binding and in the Sho1p SH3 domain are involved in the Ste11p/Ste50p interaction.

Since atypical binding surfaces and domain extensions, such as those found in the BemSH3-CI domain, can play important functional roles, it is important to ask whether these unusual features could be predicted through bioinformatics analysis. In the alignment of BemSH3 sequences from diverse fungal homologues, a very high degree of conservation is observed at positions outside of those comprising the typical peptide-binding surfaces (Figure 2.3). For example, the sequence of the Distal loop is almost completely conserved. This situation
contrasts sharply with alignments of other fungal SH3 domains, such as that from Myo3p (Figure 2.10A), which display high conservation at positions involved in peptide-binding and at hydrophobic core positions, but show wide diversity in other regions. Thus, an SH3 domain that possesses an unusual degree of conservation at positions outside of the typical peptide-interaction and hydrophobic core residues may be suspected to utilize alternative binding surfaces. Similarly, SH3 domain alignments which display additional regions of high conservation immediately adjacent to the typical domain boundaries may possess a relevant domain extension. One example of such a domain is that from yeast Bud14p, which displays an additional conserved region at its C-terminus that is required for soluble expression (Tonikian et al., 2009), and additional conserved positions within its SH3 region (Figure 2.10B). Interestingly, a typical PXXP-containing target has never been found for this domain (Tong et al., 2002; Tonikian et al., 2009). Future studies will determine whether this domain and others utilize alternative surfaces in a manner similar to the BemSH3-C1 domain.
Figure 2.10 Alignment of Myo3p SH3 and Bud14p fungal homologues

(A) Alignment of Myo3p SH3 domain fungal homologues. Positions conserved across most SH3 domains (Larson and Davidson, 2000) are indicated above alignment positions involved in peptide binding are labeled ‘p’, and hydrophobic core positions are labeled ‘h’. Positions with >67% conservation are colored. Residues are divided into the following groups: hydrophobic (light green), polar (dark green), positive (red), negative (purple), aromatic and His (light blue), Pro (yellow), Gly (orange). (B) Alignment of Bud14p SH3 domain fungal homologues. The extra conserved sequence shown is at the C-terminus of the SH3 domain. Highly conserved positions in Bud14p that could be involved in additional binding interactions (i.e. they are not positioned in the canonical peptide-binding interface) are indicated by arrows.
Chapter 3: A Conserved Residue in the Yeast Bem1p SH3 Domain Maintains the High Level of Binding Specificity Required for Function

This work has been published: Gorelik M, Stanger K, Davidson A. A conserved residue in the yeast Bem1p SH3 domain maintains the high level of binding specificity required for function. Journal of Biological Chemistry. 2011 Jun 3;286(22):19470-7.

3.1 Overview
The yeast BemSH3b and NbpSH3 domains are unusual because they bind to peptides containing the same consensus sequence, yet they perform different functions and display low sequence similarity. In this work, by analyzing the interactions of these domains with six biologically relevant peptides containing the consensus sequence, they are shown to possess finely tuned and distinct binding specificities. I also identify a residue in the Bem1p SH3b domain that inhibits binding, yet is highly conserved for the purpose of preventing non-specific interactions. Substitution of this residue results in a marked reduction of in vivo function that is caused by titration of the domain away from its proper targets through non-specific interactions with other proteins. This work provides a clear illustration of the importance of intrinsic binding specificity for the function of protein-protein interaction modules, and the key role of “negative” interactions in determining the specificity of a domain.

3.2 Introduction
Signals are transmitted through cellular pathways by relays of protein-protein interactions resulting in specific outputs, such as cell growth, differentiation, or apoptosis. To generate appropriate responses from signaling pathways, the protein-protein interactions involved must be specific, and not lead to misactivation of other pathways. This requisite precision can be readily achieved by proteins that possess “high intrinsic specificity”, directly binding their intended targets much more tightly than any other protein. For protein-DNA interactions, this can involve differences of greater than three orders of magnitude or more in $K_d$ value between target and non-target binding (Jen-Jacobson, 1997). However, protein-protein interactions are often mediated by small conserved modular domains that recognize short sequence motifs in
their target proteins and may not possess intrinsically high specificity (Bhattacharyya et al., 2006; Castagnoli et al., 2004). Since members of the same domain family often bind to similar peptide sequences, and individual domains have been found to bind many different peptides with similar affinities, many protein-protein interaction modules have been described as “promiscuous”, meaning that they are unable on their own to distinguish correct from incorrect binding sites. In these cases, the interactions of proteins containing these modules may still achieve high specificity through alternative mechanisms such as coordinated temporal and spatial localization within the cell, or participation in cooperative multi-protein complexes. Currently, the mechanisms by which protein interaction specificity within signaling pathways is achieved are not well understood (Bhattacharyya et al., 2006; Pawson, 2007). In this work, I address three key questions on this topic: How much intrinsic binding specificity is encoded in small protein-protein interaction modules? What are the mechanisms for encoding this specificity? Are there biological consequences to alteration of the, intrinsic specificity of a domain?

To address the role of binding specificity in SH3 domain function, I have examined an unusual pair of SH3 domains from the yeast adaptor proteins, Nbp2p and Bem1p. As explained in Chapter 1, the NbpSH3 and BemSH3b domains appeared to possess identical binding specificity despite their distinct biological roles and relatively low amino acid sequence identity. To investigate the specificity determinants of the NbpSH3 and BemSH3b, I measured the binding affinities of these two domains for six different target sequences. Although both domains were able to bind all six sequences, several sequences were bound with surprisingly different affinities by each of the two domains, demonstrating a finely tuned level of specificity. Further studies on this unique pair of cross-reactive SH3 domains have allowed me to identify residues that are responsible for determining specificity, address the in vivo consequences of altering specificity, and establish a mechanism by which specificity is maintained.

3.3 Experimental procedures

3.3.1 Sample preparation for in vitro binding studies
The NbpSH3 domain (residues 110-172) and BemSH3b domain (residues 155-252) were expressed with a C-terminal 6-His tag from the pET21d (Novagen) vector. Target peptides
were expressed as C-terminal fusions to bacteriophage λ cI repressor carrying C-terminal 6-His tag, as previously described (Maxwell and Davidson, 1998). Proteins were purified using nickel-affinity chromatography and dialyzed against 50 mM sodium phosphate (pH 7.0), 100mM NaCl buffer. All subsequent assays were carried out in this buffer.

3.3.2 In vitro peptide-binding and SH3 domain stability assays
The degree of binding of the SH3 domain to target peptides was assessed by monitoring the increase in SH3 domain tryptophan fluorescence at 326 nM using an Aviv ATF105 spectrofluorometer. To calculate binding affinities, SH3 domains (0.5 μM or 1 μM) were titrated with cI-peptide fusion proteins using a Microlab 500 series automated titrator. After mixing, the samples were equilibrated for 1 min and excited at 295 nM with a 5-s averaging time. The dissociation constant ($K_d$) values were calculated as previously described (Maxwell and Davidson, 1998). The $K_d$ values calculated are not affected by changes to the fluorescence difference between the bound and unbound states that could be caused by some amino acid substitutions. The thermal stability of wild type and mutant SH3 domains were measured using an Aviv Circular Dichroism Model 202 Spectrometer. Proteins were heated from 20 to 90 ºC and protein unfolding was assessed by monitoring the change in ellipticity at 212 nm for the BemSH3b domain and at 230 nm for the NbpSH3 domain, respectively.

3.3.3 SH3 domain binding assays in yeast extracts
Yeast BY4741 and BY4741 strains encoding TAP-tagged proteins were grown in YPD to an OD$_{600}$ of ~1.0 and cells from 150 ml of culture were collected by centrifugation. Protein extracts were prepared by re-suspending the cell pellet in 1.8 ml of lysis buffer (100 mM Tris-HCl pH 7.9, 250 mM NaCl, 5mM EDTA, 50mM NaF, 0.1% NP-40, 1 mM DTT, 10% Glycerol) and lysing the cells with glass beads. Ni$^{2+}$-NTA agarose resin (50 μL) was pre-loaded with 100 nmol of purified BemSH3b or NbpSH3 domains (1 ml of protein at a concentration of 100 μM) and then incubated with 200 μL of the prepared protein extract. The resin was washed 3 times with 300 μL lysis buffer and bound proteins were eluted by re-suspending the resin in 40 μL of 2X SDS sample buffer. 20 μL eluates were loaded onto 15% SDS-polyacrylamide gel and visualized by Coomassie staining or immunoblotting with α-TAP antibody (Open Biosystems, CAB1001). Each experiment was performed at least 3 times and one representative gel was chosen for the figures shown.
3.3.4 Yeast strains
Yeast strains used in this study are listed in Table 3.1. All gene modifications were introduced at chromosomal loci using PCR-based homologous recombination following standard procedures (Longtine et al., 1998). To construct strains carrying Nbp2/BemSH3b hybrid proteins, the chromosomal region encoding the NbpSH3 domain was first replaced with a URA3 cassette. This cassette was subsequently replaced by homologous recombination with a DNA fragment encoding the BemSH3b flanked by DNA homologous to the NBP2 gene. The desired recombinants lacking URA3 were selected using 5-FOA. The same method was applied to obtain strains encoding single amino acid substitutions in the NbpSH3 and BemSH3b domains. All double mutants were constructed through mating, sporulation of the diploids, and tetrad dissection, where the resulting double mutants were identified by the associated markers. To construct a plasmid expressing wild type Bem1p, a DNA fragment including the 900 base pair BEM1 promoter region and encoding Bem1p fused to a C-terminal 3HA tag was amplified from genomic DNA of strain MG40 constructed in this study (Table 3.1). The amplified DNA was inserted into a centromeric URA3- marked pRS316 plasmid.

3.3.5 Growth assays
Standard rich medium (YPD) or minimal medium (SD) supplemented with the appropriate amino acids were used for growth assays. The cells were grown overnight to saturation, normalized by O.D.600 and spotted in five-fold serial dilutions. Growth was monitored for 2 days at standard growing temperature (30 °C) or high temperatures (36-39 °C). It should be noted that changes to the incubation temperature as small as 0.5°C, caused a significant effect on cell fitness.
Table 3.1. Yeast strains used in this study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>YCW362</td>
<td>MATa ura3 leu2 his3 trp1 ssk2Δ::LEU2 ssk22Δ::LEU2</td>
<td>(Wu et al., 1999)</td>
</tr>
<tr>
<td>SY28a</td>
<td>NBP2::3HA[TRP1]</td>
<td>This study</td>
</tr>
<tr>
<td>SY29a</td>
<td>nbp2Δ::URA3</td>
<td>This study</td>
</tr>
<tr>
<td>SY56a</td>
<td>NBP2::nbp2 sh3Y3A::3HA[TRP1]</td>
<td>This study</td>
</tr>
<tr>
<td>SY58a</td>
<td>NBP2::nbp2 sh3F54A::3HA[TRP1]</td>
<td>This study</td>
</tr>
<tr>
<td>MG52a</td>
<td>NBP2::BEM1 SH3b-3HA[TRP1]</td>
<td>This study</td>
</tr>
<tr>
<td>MG54a</td>
<td>NBP2::bem1 sh3bKAD-NDN-3HA[TRP1]</td>
<td>This study</td>
</tr>
<tr>
<td>Y7092</td>
<td>can1Δ::STE2pr-Sp_his5 lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0</td>
<td>Charlie Boone Laboratory</td>
</tr>
<tr>
<td>bem1Δb</td>
<td>bem1Δ::natR</td>
<td>(Tong et al., 2004)</td>
</tr>
<tr>
<td>nbp2Δb</td>
<td>nbp2Δ::natR</td>
<td>(Tong et al., 2004)</td>
</tr>
<tr>
<td>MG40b</td>
<td>BEM1::3HA[natMX4]</td>
<td>This study</td>
</tr>
<tr>
<td>MG46b</td>
<td>BEM1::bem sh3bKAD-NDN-3HA[natMX4]</td>
<td>This study</td>
</tr>
<tr>
<td>MG134b</td>
<td>BEM1::bem sh3bK14A-3HA[natMX4]</td>
<td>This study</td>
</tr>
<tr>
<td>MG43b</td>
<td>BEM1::bem sh3bF54A-3HA[natMX4]</td>
<td>This study</td>
</tr>
<tr>
<td>MG127b</td>
<td>NBP2::3HA[natMX4]</td>
<td>This study</td>
</tr>
<tr>
<td>MG128b</td>
<td>NBP2::nbp2 sh3F54A-3HA[natMX4]</td>
<td>This study</td>
</tr>
<tr>
<td>MG130b</td>
<td>NBP2::nbp2 sh3H2A-3HA[natMX4]</td>
<td>This study</td>
</tr>
<tr>
<td>BY4741</td>
<td>MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0</td>
<td>(Brachmann et al., 1998)</td>
</tr>
<tr>
<td>smi1Δc</td>
<td>smi1Δ::kanMX4</td>
<td>(Winzeler et al., 1999)</td>
</tr>
<tr>
<td>bni1Δc</td>
<td>bni1Δ::kanMX4</td>
<td>(Winzeler et al., 1999)</td>
</tr>
<tr>
<td>bem2Δc</td>
<td>bem2Δ::kanMX4</td>
<td>(Winzeler et al., 1999)</td>
</tr>
<tr>
<td>bud1Δc</td>
<td>bud1Δ::kanMX4</td>
<td>(Winzeler et al., 1999)</td>
</tr>
<tr>
<td>cnb1Δc</td>
<td>cnb1Δ::kanMX4</td>
<td>(Winzeler et al., 1999)</td>
</tr>
<tr>
<td>lte1Δc</td>
<td>lte1Δ::kanMX4</td>
<td>(Winzeler et al., 1999)</td>
</tr>
<tr>
<td>msb1Δc</td>
<td>msb1Δ::kanMX4</td>
<td>(Winzeler et al., 1999)</td>
</tr>
<tr>
<td>msg5Δc</td>
<td>msg5Δ::kanMX4</td>
<td>(Winzeler et al., 1999)</td>
</tr>
<tr>
<td>ptp2Δc</td>
<td>ptp2Δ::kanMX4</td>
<td>(Winzeler et al., 1999)</td>
</tr>
<tr>
<td>kre1Δc</td>
<td>kre1Δ::kanMX4</td>
<td>(Winzeler et al., 1999)</td>
</tr>
<tr>
<td>Las17p-TAPc</td>
<td>LAS17::TAP[his3MX]</td>
<td>(Ghaemmaghami et al., 2003)</td>
</tr>
<tr>
<td>TiF32p-TAPc</td>
<td>TIF32::TAP[his3MX]</td>
<td>(Ghaemmaghami et al., 2003)</td>
</tr>
<tr>
<td>Cdc48p-TAPc</td>
<td>CDC48::TAP[his3MX]</td>
<td>(Ghaemmaghami et al., 2003)</td>
</tr>
<tr>
<td>Uba2p-TAPc</td>
<td>UBA2::TAP[his3MX]</td>
<td>(Ghaemmaghami et al., 2003)</td>
</tr>
<tr>
<td>Gln4p-TAPc</td>
<td>GLN4::TAP[his3MX]</td>
<td>(Ghaemmaghami et al., 2003)</td>
</tr>
<tr>
<td>Mas5p-TAPc</td>
<td>MASP::TAP[his3MX]</td>
<td>(Ghaemmaghami et al., 2003)</td>
</tr>
</tbody>
</table>

a-YCW362 derivative; b-BY4742 derivative; c-BY4743 derivative; d-ATCC 201388 derivative
3.4 Results

3.4.1 The NbpSH3 and BemSH3b domains display finely tuned specificity
To determine the binding specificities of the BemSH3b and Nbp2SH3 domains, I tested the
binding of these SH3 domains to six peptides derived from yeast proteins implicated in the
function of either Bem1p or Nbp2p (Table 1.3). The BemSH3b domain construct (155-252)
used in the binding assays included a C-terminal extension that is required for folding of the
domain and contributes to peptide binding by extending the peptide binding surface (Chapter
2, (Takaku et al., 2010)). The interaction of the SH3 domains to target peptides was monitored
by measuring the increase in intrinsic Trp fluorescence resulting from peptide-binding (Figure
3.1).

**Figure 3.1 Tryptophan fluorescence binding assays.**

(A) Fluorescence spectra of the NbpSH3 domain alone and in the presence of 20 μM Ste20 peptide are shown. Similar changes in emission fluorescence were observed upon incubation of NbpSH3 domain with other peptides and in the same experiments with the BemSH3b domain. (B) A representative binding curve from a single titration experiment is shown where 0.5 μM NbpSH3 domain was titrated with Ste20 peptide. Binding was detected by monitoring the change in emission fluorescence. The line joining the points in the graph is the theoretical fit of the data based on previously published binding equations (Maxwell and Davidson, 1998).

Although both SH3 domains bound all of the tested peptides, the affinities of each domain for the set of peptides span a broad range of values. The NbpSH3 domain bound with $K_d$ values ranging from 0.2 to 11 μM, while the BemSH3b domain bound with an even greater range of affinities varying from 0.05 to 26 μM (Figure 3.2). While some target peptides, including Ste20, Cla4, and Pbs2 were bound with relatively similar affinities by the two domains (<4 fold difference), the Boi2, Bck1, and Skm1 sites showed a much greater difference in their affinities for the two domains ranging from 20- to 30-fold. In the case of the Boi2 site, the
BemSH3 bound more tightly, while the Nbp2SH3 bound the Bck1 and Skm1 sites with higher affinity. It is also notable that each domain bound to most sites with relatively high affinity compared to typical SH3 domain target peptide interactions. For example the $K_d$ value of the BemSH3b-Ste20 interaction (0.05 µM) is among the lowest observed for SH3 domains. In summary, although these domains recognize the same consensus sequence, their specificities are seen to be finely tuned when assayed against a panel of yeast peptides, some of which are proven biological targets of NbpSH3 or BemSH3b domains (Chapter 1, Table 1.3).

3.4.2 The conserved Lys14 residue of the BemSH3b domain is a key determinant of specificity

To identify key residues that determine the distinct specificities of the BemSH3b and NbpSH3 domains, I exchanged residues 14-16 in the RT-loops of these domains (Figure 3.2C). I chose these positions because they were occupied by residues with different properties in the two domains (KAD in BemSH3b versus NDN in NbpSH3) and this region of SH3 domains had been implicated in determining specificity in several different systems (Hiipakka and Saksela, 2007; Panni et al., 2002). Remarkably, the BemSH3b$_{KAD\rightarrowNDN}$ mutant bound most sites more tightly than the wild-type domain (Figure 3.2B). In particular, it bound the Bck1 and Skm1 sites 50- and 20-fold more tightly, respectively. Overall, the specificity profile of the BemSH3b$_{KAD\rightarrowNDN}$ mutant more closely resembles the specificity of the wild-type NbpSH3 except that it is still able to bind the Boi2 site with a much higher affinity than the Nbp2SH3. In contrast to the BemSH3b$_{KAD\rightarrowNDN}$ mutant, the NbpSH3$_{NDN\rightarrowKAD}$ mutant displayed significantly reduced binding affinity to all the peptides.

To investigate the mechanism of the changes in specificity induced by the RT-Loop substitutions of the BemSH3b domain, I tested the effects of substitutions with Ala in this region. Strikingly, the K14A substitution caused large increases in affinity for the Cla4, Skm1, and Bck1 sites, similar to the KAD$\rightarrow$NDN substitution, and also bound more tightly to the Ste20 and Boi2 sites (Figure 3.2B). A K14A/D16A double mutant showed no increase in affinity for any sites beyond that observed for the K14A mutant (Table 3.2). Thus, the altered binding properties the BemSH3b$_{KAD\rightarrowNDN}$ mutant can be mostly attributed to the elimination of Lys14 side chain, suggesting that this residue interferes with the binding mediated by the
Figure 3.2 In vitro binding analysis of BemSH3b and NbpSH3 interactions

(A) The alignment of yeast peptides used in in vitro binding assays. The consensus sequence for peptides bound by the BemSH3b and NbpSH3 domains is indicated, where Ψ is a hydrophobic residue and X is any amino acid. Peptide numbering is according to Lim et al. (Lim et al., 1994). (B) Binding affinities of wild-type and mutant BemSH3b and NbpSH3 domains. The $K_d$ values were determined as described in Experimental Procedures. Mean values with standard errors are indicated. The measurements with the wild-type domains were performed at least 3 times and the measurements with the mutant domains were performed at least twice. The thermodynamic stability of each domain was measured by temperature-induced unfolding experiments and the temperature midpoints ($T_m$) of these transitions are shown. (C) Alignments of BemSH3b and NbpSH3 domains from diverse fungal species. The numbering of SH3 domain residues is according to a standard convention (Larson and Davidson, 2000). BemSH3b secondary structure and residues involved in peptide binding (indicated by ‘P’) obtained from the structure of BemSH3b-Ste20 complex (Takaku et al., 2010) are shown. The RT-loop residues exchanged between the BemSH3b and NbpSH3 domains are boxed. The BemSH3b K14 and NbpSH3 H32 residues that interfere with binding are indicated by the arrows. Positions with >70% conservation are colored: hydrophobic (light green), polar (dark green), positive(red), negative(purple), aromatic and His(light blue), Pro(yellow), Gly(orange).
BemSH3b domain. This result is surprising because the alignment of BemSH3b domains from widely diverse fungal species (Figure 3.2C) shows that a positively charged residue has been conserved at this position over at least 500 million years of evolution. Substitutions at such conserved positions in protein-protein interaction interfaces usually cause large decreases in binding affinity.

Table 3.2  Additional peptide binding affinities of the NbpSH3 and BemSH3b mutants

<table>
<thead>
<tr>
<th>SH3 domains</th>
<th>$K_d$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ste20</td>
</tr>
<tr>
<td>BemSH3b (Figure 3.2)</td>
<td>0.05±0.01</td>
</tr>
<tr>
<td>BemSH3b KAD→AAA</td>
<td>0.16±0.01</td>
</tr>
<tr>
<td>BemSH3b H32A</td>
<td>0.37±0.06</td>
</tr>
<tr>
<td>BemSH3b F54A</td>
<td>9.1±2</td>
</tr>
</tbody>
</table>

3.4.3 The difference in binding affinities between BemSH3b and NbpSH3 domains is biologically significant

To determine whether the differences in binding specificity between NbpSH3 and BemSH3b domains observed in vitro have an effect on in vivo function, I replaced the chromosomally encoded NBP2 gene with one encoding a hybrid protein in which the NbpSH3 domain was replaced with the BemSH3b domain. As shown in Figure 3.3, substitution of NbpSH3 domain with the BemSH3b domain resulted in growth defects at 39 °C comparable to those seen with an nbp2Δ mutant (Ohkuni et al., 2003). To assess the defect in NbpSH3 domain function conferred by replacing it with the BemSH3b domain, I compared the growth of cells expressing the Nbp2/BemSH3b hybrid protein to cells expressing Nbp2p bearing Y8A and F54A SH3 domain substitutions, which cause large reductions in peptide-binding but are not expected to affect specificity because they interact with the PXXP region (Larson and Davidson, 2000). The level of growth mediated by the Nbp2/BemSH3b hybrid protein was intermediate between the NbpSH3 Y8A and F54A mutants which display 45- and 280-fold reductions, respectively, in binding affinity to the Ste20 peptide (Table 3.2). This suggested that in Nbp2/BemSH3b hybrid protein, SH3 domain function is significantly affected, but not completely lost. In contrast to cells expressing the Nbp2/BemSH3b hybrid protein, those
expressing the Nbp2/BemSH3b$^{KAD\rightarrow NDN}$ mutant grew as well as cells expressing wild-type Nbp2p (Figure 3.3A). All of these mutant proteins were expressed at levels similar to wild-type (Figure 3.3B). These data show that the specificity differences between the NbpSH3 and BemSH3b observed in vitro result in an inability of the BemSH3b to replace the NbpSH3 in vivo. However, the BemSH3b$^{KAD\rightarrow NDN}$ mutant was able to replace the NbpSH3 because its specificity much more closely resembles that of the NbpSH3 domain.

**Figure 3.3 Biological activity of Nbp2/BemSH3b hybrid proteins**

(A) An npb2Δ strain and strains expressing Nbp2p bearing the indicated SH3 domain amino acid substitutions, which impair peptide binding (Table 3.2), were included for comparison. Strains were grown overnight, normalized to an O.D.$_{600}$ of 1.0 and spotted onto YPD plates in five-fold serial dilutions. (B) The expression level of Nbp2p mutants. Wild type and mutant versions of Nbp2p were tagged at the C-terminus with the 3HA tag and the accumulated levels of protein were detected by immunoblotting with α-HA antibody (Santa Cruz Biotechnology, sc-7392). To control for gel loading and membrane transfer, the same membrane was probed with an α-tubulin antibody (Sigma, T-5168).

3.4.4 Altering the binding specificity of the BemSH3b domain impairs Bem1p function

To investigate the effects of altered binding specificity on the in vivo function of the Bem1p protein, I constructed yeast strains in which the BEM1 gene was replaced by genes encoding the BemSH3b KAD$\rightarrow$NDN or K14A mutants. Although strains bearing these mutations displayed no reduction in viability in a wild-type strain background (Figure 3.4A), they displayed marked reductions in cell viability at high temperature (Figure 3.5) when combined with smi1Δ, bni1Δ or bem2Δ gene deletions, which are synthetically lethal with bem1Δ (Figure 3.4B, (Tong et al., 2004)). The in vivo effects of the altered specificity mutants were not the result of altered Bem1p expression levels (Figure 3.4D) or BemSH3b domain instability (Figure 3.2B). To gauge the importance of BemSH3b domain function in these backgrounds, I also tested the effect of the BemSH3b F54A substitution, which causes a large
reduction in binding affinity (Table 3.2). Introduction of this substitution resulted in complete lack of growth at high temperature (Figure 3.5) or synthetic lethality in the case of bem2Δ (Figure 3.4C) emphasizing a critical role of SH3 domain activity in these backgrounds.

**Figure 3.4 In vivo analysis of Bem1p function**

(A) Cell fitness effects of BemSH3b amino acid substitutions analyzed in this study. While the BemSH3b KAD-NDN and K14A mutations that affect specificity have no effect on cell fitness in the wild-type background, the F54A substitution, which severely impairs binding, causes growth impairment at 39 °C. The K14A substitution also caused no loss in viability when tested in several other backgrounds that were previously shown to have negative genetic interactions with bem1Δ. These backgrounds included bud1Δ (Irazoqui et al., 2003), msb1Δ (Bender and Pringle, 1991), cnb1Δ (Parsons et al., 2004), lte1Δ (Costanzo et al., 2010). By contrast, the F54A caused severe temperature-sensitivity in all of these backgrounds. (B) Synthetic lethality between bem1Δ and bem2Δ, smi1Δ and bni1Δ gene deletions. bem1Δ, bem1Δ bem2Δ, bem1Δ smi1Δ, bem1Δ bni1Δ strains carrying a URA3-marked BEM1-expressing plasmid were grown on SD medium lacking uracil and medium containing 5-fluoro-orotic acid (5-FOA) at 30 °C as indicated. In the presence of FOA, only cells that have lost the URA3-marked BEM1 expressing plasmid are able to grow. Cells carrying bem1Δ bem2Δ, bem1Δ smi1Δ, bem1Δ bni1Δ double deletions were inviable on 5-FOA medium, whereas cells carrying bem1Δ alone were viable on both media. (C) Synthetic lethality between bemSH3bΔ and bem2Δ. Tetrad dissection of asci derived from the indicated heterozygous diploid strains are shown. The 3:1 and 2:2 segregation for viability indicates the inability of bem2Δ bemSH3bΔΔ spores to form colonies. None of the colonies shown were able to grow on medium that selected for both bem2Δ and bemSH3bΔΔ markers (KAN and NAT). (D) The level of Bem1p expression from the pRS316-derived CEN-based plasmid. The plasmid and endogenously expressed versions of Bem1p were tagged with the 3HA tag at the C-terminus and the level of expression was detected by immunoblotting with the α-HA antibody (Santa Cruz Biotechnology, sc-7392). (E) Analysis of Bem1p wild-type and SH3b mutant proteins showed that they were all expressed at similar levels. Proteins were detected by immunoblotting as described in (D). (F) Cell viability in a diploid background is decreased by reducing the amount of functional Bem1p in half through heterozygosity. The viability of a +/-bem1Δ heterozygote was assayed a bni1Δ/bni1Δ homozygous background. The cells were grown overnight, normalized to O.D.600 of 1.0 and spotted in five-fold serial dilutions.
The phenotypes caused by the BemSH3b KAD→NDN and K14A mutations could have been the result of a gain-of-function whereby the altered binding specificity led to deleterious misactivation of a pathway. To explore this possibility, I tested the effect of plasmid-mediated expression of wild-type Bem1p within these mutant cells. As shown in Figure 3.5, expression of wild-type Bem1p at endogenous levels (Figure 3.4) led to normal levels of growth for the BemSH3b mutants in all three backgrounds. This recessive-like behavior indicated that altering the binding specificity of the BemSH3b domain caused a loss of Bem1p function that could be overcome by expressing wild-type Bem1p. If these mutations had caused a gain-of-function, I would have expected their effect to be dominant.

Figure 3.5  The effect of BemSH3b altered specificity mutants on viability.

Chromosomally expressed mutants of Bem1p bearing the indicated amino acid substitutions in the SH3b domain were assayed in smi1Δ, bmi1Δ, bem2Δ genetic backgrounds. A BemSH3bF54A mutant that significantly impairs peptide binding (Table 3.2) was included for comparison. All strains also carried a centromeric plasmid. The plasmid in strains shown in the right panels expressed wild-type Bem1p under the control of its own promoter, while the plasmid in strains shown in the left panels contained no insert. Strains were grown overnight, normalized to O.D.600 of 0.2 and spotted in 5-fold serial dilutions on SD plates lacking uracil.

Since the BemSH3bKAD→NDN mutant displayed specificity that was similar to that of the NbpSH3 domain, I postulated that Bem1p carrying this substitution might interfere with
Nbp2p function by competing for NbpSH3 binding sites within the cell. However, in strains carrying this mutation, I failed to observe any phenotypes indicative of a defect in the function of Nbp2p. These might have included hyperactivation of the Cell Wall Integrity pathway (Du et al., 2006) and growth defects in strains sensitized to Nbp2p SH3 function by deletion of additional genes (Figure 3.6).

**Figure 3.6** BemSH3b KAD→NDN substitution has no effect on Nbp2p function.

(A) Previously demonstrated synthetic sick/lethal interactions between nbp2Δ and the ptp2Δ (Mapes and Ota, 2004), msg5Δ (Costanzo et al., 2010), kre1Δ (Tong et al., 2004) gene deletions are recapitulated in crosses with a mutant nbp2 producing the NbpSH3F54A substituted protein, which is impaired in peptide binding (Table 3.2). Crosses with a mutant bem1 expressing the BemSH3bKAD→NDN substitution, which binds with high affinity to Nbp2 specific sites, causes no fitness defects in the context of the same gene deletions. The tested gene deletions are specifically sensitive to NbpSH3 function, but not BemSH3b, as evidenced by the absence of any genetic interactions with BemSH3bV36K mutant that impairs peptide binding (Endo et al., 2003). The cells were grown overnight, normalized to O.D._{600} of 1 and spotted in 5-fold dilutions. (B) CWI pathway hyper-activation is
observed in \( nbp2\Delta \) and \( npbsh3^{\text{F54A}} \) mutants, while the \( \text{BemSH3b}^{\text{KAD-NDN}} \) mutant displays wild-type levels of pathway activity. Cells were grown to mid-log phase, normalized by \( \text{O.D.}_{600} \), quick extracts were prepared and run on an SDS-PAGE gel. Phosphorylated Slt2 was detected using anti-phospho-p44/p42 antibody (Cell Signaling, 9101). To control for gel loading and membrane transfer the same membrane was probed with \( \alpha\)-tubulin antibody (Sigma, T-5168).

3.4.5. The K14A substitution causes an increase in non-specific binding

The recessive behavior of the BemSH3b altered specificity mutants led me to hypothesize that the impairment of Bem1p function caused by these mutations might be due to increased non-specific binding, which might divert Bem1p from its proper target proteins. To investigate this idea, I compared the abilities of wild-type and mutant BemSH3b domains coupled to Ni-affinity resin to bind proteins from yeast extracts. These assays were performed with high concentrations of SH3 domains such that non-specific binding could be detected. Strikingly, the K14A mutant bound to many more proteins than the wild-type domain (Figure 3.7A). The large number of different proteins bound by the K14A mutant and the apparently high concentration of these proteins (they were easily visualized in a Coomassie stained SDS-PAGE gel) suggested that these interactions were non-specific. The wild-type domain also bound many of these same proteins, but the level of binding was much lower. Most of the bound proteins were absent in the eluate of the BemSH3b domain containing F54A substitution which abrogates PXXP binding interface. These indicated that these non-specific interactions still required Pro-rich sequences.

With this in mind, I used the same assay to test binding of the BemSH3b domain to six different TAP-tagged proteins containing RXXPXXP motifs, which have no functional connection to Bem1p (Table 3.3). Among these proteins, Tif32p, Cdc48p, and Las17p showed increased binding to the K14A mutant when compared to the wild-type domain but absence of binding to the F54A mutant (Figure 3.7). These results mirrored those seen in the binding assay to total yeast lysates.

To confirm increased non-specific peptide binding activity of the K14A mutant I assessed its binding to an RXXPXXP-containing peptide that possessed none of the other conserved features of peptides bound tightly by the BemSH3b domain (Figure 3.7C). The wild-type BemSH3b domain displayed little detectable binding to this peptide as detected by examination of Trp fluorescence spectra even when the peptide was present in 100-fold excess. By contrast, addition of this peptide to the K14A mutant caused approximately 5-fold
increase in Trp fluorescence intensity at 330 nm compared to the WT domain and a blue-shift in the peak wavelength of emission, indicating that the mutant bound significantly more strongly to the non-specific target. In summary, these results demonstrate an increased level of non-specific binding by the BemSH3b\textsuperscript{K14A} mutant, which could reduce its ability to bind sufficiently to its required partners in vivo.

Figure 3.7 BemSH3b K14A and NbpSH3 H32A mutants display increase in non-specific binding

(A) Binding of SH3 domains to proteins in yeast extracts. Purified SH3 domains bound to Ni-affinity resin were incubated with yeast extracts. Bound proteins were eluted and visualized by Coomassie staining of SDS-PAGE gels. Protein bands with higher intensity in the NbpSH3 H32A and BemSH3b K14A lanes are indicated by arrows. The bands corresponding to the eluted SH3 domains (intense bands at the bottom) are indicated.

(B) Binding of the BemSH3K14A mutant to TAP-tagged proteins. The same procedure was used as in part (A), except proteins bound by the SH3 domain were detected by immunoblotting with and antibody directed against the TAP tag. The tested TAP-tagged proteins included Tif32p (130,343 Da), Cdc48p (111,995 Da) and Las17p (87,571 Da). See also Table 3.3.

(C,D) Binding of BemSH3b and NbpSH3 domains to a non-specific RXXPXXP containing target. Representative Trp fluorescence spectra of SH3 domains at a concentration of 1 μM alone or in the presence of 100 μM target peptide are shown. \(K_d\) values were not obtained for non-specific binding reactions because the low affinity of these reactions made it impossible to reach binding saturation with the quantities of peptide-fusion protein available.
Table 3.3 TAP-tagged proteins tested for interaction with the BemSH3b domain.

<table>
<thead>
<tr>
<th>Name</th>
<th>Function</th>
<th>RxxPxxP</th>
<th>Abundance&lt;sup&gt;b&lt;/sup&gt; (molecules/cell)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Las17p</td>
<td>Actin assembly factor, activates the Arp2/3 protein complex that nucleates branched actin filaments; localizes with the Arp2/3 complex to actin patches; homolog of the human Wiskott-Aldrich syndrome protein</td>
<td>PQQNRPLPQLP, NRNNRPVPPPP, PPRRGAPP, QATGRRGPAPP</td>
<td>8580</td>
</tr>
<tr>
<td>Cdc48p</td>
<td>ATPase in ER, nuclear membrane and cytosol with homology to mammalian p97; in a complex with Npl4p and Ufd1p participates in retrotranslocation of ubiquitinated proteins from the ER into the cytosol for degradation by the proteasome</td>
<td>NAQLRKTPLEP</td>
<td>78400</td>
</tr>
<tr>
<td>Tif32p</td>
<td>eIF3a subunit of the core complex of translation initiation factor 3 (eIF3), essential for translation; part of a subcomplex (Prt1p-Rpg1p-Nip1p) that stimulates binding of mRNA and tRNA(i)Met to ribosomes</td>
<td>RAGNREPPSTP</td>
<td>52700</td>
</tr>
<tr>
<td>Uba2</td>
<td>Nuclear protein that acts as a heterodimer with Aos1p to activate Smt3p (SUMO) before its conjugation to proteins (sumoylation), which may play a role in protein targeting; essential for viability</td>
<td>VCTIRSTPSQP</td>
<td>18800</td>
</tr>
<tr>
<td>Mas5</td>
<td>Protein chaperone involved in regulation of the HSP90 and HSP70 functions; involved in protein translocation across membranes; member of the DnaJ family</td>
<td>ILPPRIVPAIP</td>
<td>44670</td>
</tr>
<tr>
<td>Gln4</td>
<td>Glutamine tRNA synthetase, monomeric class I tRNA synthetase that catalyzes the specific glutaminylation of tRNA(Glu); N-terminal domain proposed to be involved in enzyme-tRNA interactions</td>
<td>KEFFRLTPNQP</td>
<td>37500</td>
</tr>
</tbody>
</table>

<sup>a</sup>-Description of protein function was copied from Saccharomyces Genome Database (<http://www.yeastgenome.org>)

<sup>b</sup>-The protein abundance data shown here is taken from the global expression study of yeast proteins (Ghaemmaghami et al., 2003)
3.4.6 An NbpSH3 mutant also displays increased non-specific binding

My findings regarding the Lys14 position of the BemSH3b domain led me to seek similar specificity-regulating residues within the NbpSH3 domain. Through investigation of several conserved positions within the putative peptide-interaction interface of this domain, I discovered that the H32A substitution led to an approximately 10-fold increase in affinity for most of the tested target peptides (Figure 3.2B). Interestingly, the BemSH3b also possesses a His residue at this position, but substitution of this residue had little effect on binding affinity (Table 3.2). Like the BemSH3bK14A mutant, the NbpSH3H32A mutant showed an increased level of non-specific binding to proteins in a yeast cell extract as compared to the wild-type domain (Figure 3.7A). This increased non-specific binding was still observable even though the wild-type NbpSH3 displayed an overall higher level of non-specific binding as compared to the BemSH3b domain. Also similar to the BemSH3bK14A mutant, the NbpSH3H32A mutant displayed increased non-specific binding in vitro (Figure 3.7D).

3.5 Discussion

To understand how cells are able to elicit specific responses to changes in environmental conditions, it is essential to delineate the mechanisms by which protein-protein interaction specificity is generated. A key question is how much of the information governing pathway specificity is encoded within individual protein-protein interaction modules. In this work, I have addressed this question using the BemSH3b and NbpSH3 domains, and demonstrated that these domains possess finely tuned intrinsic binding specificities, despite recognizing a common consensus sequence. Furthermore, I have shown that altering this intrinsic binding specificity through amino acid substitutions markedly impairs function in vivo, and that a highly conserved residue in the BemSH3b domain plays a key functional role in reducing non-specific interactions. This work provides a clear illustration of the importance of intrinsic binding specificity for the function of protein-protein interaction modules, and the key role of “negative” interactions in determining the specificity of a domain.

Although the BemSH3b and NbpSH3 domains both bound the same six biologically relevant peptides sequences, they did so with distinct specificities. They bound some sites with very similar affinities, while affinity differences of up to 30-fold were seen for other sites (Figure 3.2B). Remarkably, the observed binding affinities of these domains for the six sites varied
over a 500-fold range even though these sites all conform to the same consensus sequence. These results illustrate that the intrinsic binding specificity of SH3 domains can be highly nuanced, and that residues outside of the conserved positions of the consensus play important roles. The distinct binding specificities of the NbpSH3 and BemSH3b domains are the result of different binding mechanisms, which will be demonstrated in the next chapter (Chapter 4).

At the same time, my observation that the BemSH3b and NbpSH3 domains can interact with the same site with similar high affinity (e.g. Cla4) suggests that these domains may at times function by competing for binding to the same protein. There is a biological rationale for such a competition because as explained in the introduction these proteins appear to generate opposite responses: Nbp2p acts as a negative regulator by recruiting Ptc1p phosphatase, while Bem1p acts as an activator by assembling components of various pathways.

The changes in binding specificity resulting from amino acid substitutions in the BemSH3b domain caused significant in vivo effects. Although the function of the NbpSH3 could not be assumed by the wild-type version of the BemSH3b, a variant of Nbp2p bearing the mutant BemSH3b\(^{\text{KAD} \rightarrow \text{NDN}}\) domain was able to function normally (Figure 3.3). The ability of this mutant BemSH3b domain to replace the NbpSH3 domain is likely due to its increased affinity for sites that are bound strongly by the NbpSH3. For example, the BemSH3b\(^{\text{KAD} \rightarrow \text{NDN}}\) mutant bound 50-fold more tightly to the site from the Bck1p MAP kinase, which is a required component of the CWI pathway. I also found that replacement of the genomic copy of \textit{BEM1} with a gene encoding either of the mutations affecting specificity in vitro (i.e. KAD\rightarrowNDN or K14A) caused a marked decrease in viability in several different genetic backgrounds (Figure 3.5). It is remarkable that despite the increased binding affinity of these mutants for all of the sites tested, they were still debilitated in their function. This implies that the altered binding specificity of these mutants is the cause of their in vivo deficiency. Although a number of studies have investigated amino acid substitutions in protein-protein interaction modules that alter binding specificity (Fazi et al., 2002; Hiipakka and Saksela, 2007; Panni et al., 2002), this study provides the first easily demonstrable viability phenotype for such a mutant. One comparable result was obtained in a study of the yeast Sho1p SH3 domain in which amino acid substitutions that caused reduced specificity were generated in the target site for this domain. The presence of these promiscuous sites within a yeast strain decreased viability, but this phenotype only became evident after many generations of growth under stressful conditions.
conditions in competition with a wild-type strain (Zarrinpar et al., 2003). In another study, similar promiscuous binding sites for the Gads SH3 domain were isolated, but these sites showed no alteration in \textit{in vivo} activity (Seet et al., 2007).

Another important observation in this study is that the function of Lys14 within the BemSH3b domain is to prevent interactions with non-target proteins. My experiments showed that the K14A mutant bound more tightly than the wild-type to at least 9 different proteins (Figure 3.7A). Since the specific targets of NbpSH3 and BemSH3b domains investigated in this study are expressed at a level that would not be detected on a Coomassie stained gel (<2000 molecules per cell, (Ghaemmaghami et al., 2003)), the proteins detected in our experiments are likely to be abundant proteins that bound non-specifically and with low affinity. While both the wild-type and K14A mutant domains bind to these proteins, the K14A mutant clearly displays increased non-specific binding. This conclusion is further supported by our observation that the K14A mutant bound more tightly in pull-down experiments to Las17p, Cdc48p, and Tif32p, which possess only sites conforming to the RXXPXXP consensus (Figure 3.7B), and the K14A mutant bound more tightly \textit{in vitro} to a non-specific site (Figure 3.7C). We conclude that despite the increased affinity of the K14A mutant for biologically relevant target proteins such as Ste20p and Cla4p (Bose et al., 2001; Winters and Pryciak, 2005), its increased affinity for numerous abundant non-target proteins, causes it to be titrated away from its correct biological targets, resulting in a relative deficiency of Bem1p activity. This suggestion is supported by our finding that the loss in viability resulting from chromosomally introduced KAD → NDN and K14A mutations could be rescued by wild-type Bem1p expressed from a plasmid. This recessive behavior would not have been observed if loss of viability were caused by activation of an inappropriate pathway. Furthermore, we also observed that in the \textit{bni1Δ} diploid background the \textit{BEM1/bem1Δ} heterozygote displayed some loss in viability at high temperature (Figure 3.4F), which implies that cells are sensitive to decreased levels of Bem1p as would result from the non-specific binding activity of the K14A mutant. In summary, our data indicate that the primary function of Lys14 is not to form stabilizing interactions with biologically relevant target peptides, but to prevent non-specific interactions with other cellular proteins.
Consistent with my findings, a recent study of SH2 domains showed that residues preventing binding to certain sequences play an important role in determining binding specificity \textit{in vitro}, and that both positive and negative interactions allow for a high degree of nuance in the specificities of these domains. Similarly, the promiscuous target of the Sho1p SH3 domain, mentioned above, contained alanine substitutions that did not reduce its affinity for Sho1p SH3 domain, yet caused it to bind other SH3 domains with higher affinity (Zarrinpar et al., 2003). Similar to these previous studies, my work emphasizes the importance of high intrinsic specificity for the function of protein-protein modules, and the key role of negative selection in determining specificity. However, I have also demonstrated a pronounced \textit{in vivo} phenotype resulting from loss of binding specificity, and identified a single residue within a binding domain that appears to been highly conserved for the sole purpose of preventing non-specific interactions.

Although to my knowledge Lys14 of the BemSH3b domain is the first residue to be identified in a protein-protein module that is likely conserved for the purpose of preventing non-specific interactions, I expect that many other such residues exist. A hallmark of such residues may be similar to that seen for Lys14 where elimination of the side chain increases affinity to biologically relevant sites. Interestingly, the Ala substitution of the conserved His32 position of the NbpSH3 also increased binding affinity (Figure 3.2) and non-specific binding (Figure 3.7 A and D). My ability to identify residues in two different domains that may be conserved for specificity determination while other studies have failed to identify such residues could be a result of a dearth of systematic mutagenesis studies of residues comprising the binding interface of protein-protein interaction modules. Most studies involving mutagenesis of such domains have focused on positions that are conserved across the whole domain family and thus are not expected to be specificity determinants. Future studies, aimed at locating such determinants, should focus on positions that are only conserved among othologous domains, such as Lys14 of BemSH3b and His32 of NbpSH3. In this way, further insight into the importance of intrinsic binding specificity for the function of these domains will be obtained.
Chapter 4: Distinct recognition of the Ste20 peptide by the Nbp2p SH3 and Bem1p SH3b domains: understanding the determinants of SH3 domain binding specificity

4.1 Overview
In chapter 3, I demonstrated that although NbpSH3 and BemSH3b domains recognize the same consensus sequence and bind some targets with a similar affinity, they are able to discriminate against other targets, binding them with 20 to 30-fold differences in affinity. In this chapter, I delineate binding mechanisms of these SH3 domains and determine what regulates their specificity with respect to each other and other yeast SH3 domains. I have solved the structure of NbpSH3 domain in complex with the Ste20 peptide and compared it to the available structure of the BemSH3b-Ste20 complex, as well as the structure of the *L. elongisporus* BemSH3b-Ste20 complex that I solved in Chapter 2. Combined with the mutagenesis of Ste20 peptide and the SH3 domains, my structural analysis demonstrates that NbpSH3 and BemSH3b use a different mechanism for binding, explaining how they are able to bind some sites with a similar affinity, while discriminating against other sites. While NbpSH3 forms stronger interactions with the core peptide motif, BemSH3b binding is more dependent on interactions with the extended peptide region, demonstrating that SH3 domain specificity can be fine tuned through modulation of interactions with the core and extended peptide regions. My results also suggest that in the case of BemSH3b its unique C-terminal extension contributes significantly to specificity, while in the case of the NbpSH3 domain unique residues on the variable surface are important. This study provides further evidence for the ability of SH3 domains to possess high intrinsic specificity and provides an example of how distinct specificities can be produced even when two SH3 domains recognize the same consensus sequence.

4.2 Introduction
Yeast contains 27 SH3 domains with high sequence diversity, making it a good model system for investigating specificity of SH3 domain mediated interactions. Although several large scale studies examined binding preferences of yeast SH3 domains (Cesareni et al., 2002; Landgraf et al., 2004; Tong et al., 2002; Tonikian et al., 2009), how the specificity is
established is not well understood for most SH3 domains. In Chapter 3, I demonstrated that although the NbpSH3 and BemSH3b domains share binding preferences, each domain still possesses its own unique specificity. Furthermore, I have identified a residue within BemSH3b domain (Lys14), which serves to prevent high affinity binding to “NbpSH3 specific sites” (Bck1 and Skm1) and plays a general role in maintaining BemSH3b domain specificity required for its optimal function.

Several studies have compared binding mechanisms of different SH3 domains with the same target (Ababou et al., 2009; Casares et al., 2007; McDonald et al., 2009). However many of these studies employed relatively short peptides that bound targets with modest affinities and may thus not be as informative for understanding SH3 domain specificity. The ability of the NbpSH3 and BemSH3b domains to bind extended peptides with high affinity, while simultaneously discriminating against other targets provides a unique model system to examine the determinants of SH3 domain specificity. In this chapter, I used a combination of structural analysis and \textit{in vitro} binding assays to dissect the binding mechanism of these two SH3 domains. The main questions I address are whether the NbpSH3 and BemSH3b domains use the same or different mechanisms for binding and how their binding mechanisms allow them to share preferences for some binding sites, while simultaneously discriminating against other sites. In addition, I examined what contributes to the specificity of these domains not only with respect to each other, but also when considered in the context of the whole yeast SH3 proteome.

4.3 Materials and Methods

4.3.1 Protein expression, peptide binding assays and thermal stability assays.
The NbpSH3 and BemSH3b domains (Nbp2p residues 110-172 and Bem1p residues 155-252), and peptides (Ste20p residues 468-483 and Bck1p residues 800-815) used in the binding assays were expressed and purified as described in Chapter 3, section 3.3.1. Peptide binding assays and thermal stability assays were performed as described in Chapter 3, section 3.3.2. It should be noted that the stability of mutant SH3 domains is reported as the change in the melting temperature value ($\Delta T_m$) with respect to the wild-type. Although each measurement was done only once, the wild-type domain was always included as a control.
4.3.2 NMR spectroscopy and structure calculation

A peptide corresponding to Ste20p residues 468-484 was chemically synthesized (CanPeptide Inc). For NMR experiments, two samples were prepared. Sample A contained $^{15}$N $^{13}$C labeled NbpSH3 at 0.7 mM and unlabeled Ste20p peptide at 1.5 mM in 50 mM phosphate, pH 6.8, 100 mM NaCl, 0.05% NaN3. Sample B contained $^{15}$N $^{13}$C labeled NbpSH3 at 0.5 mM and unlabeled Ste20p peptide at 0.5 mM in 50 mM Phosphate buffer, pH 6.8, 100 mM NaCl, and 0.05% NaN3. The NMR experiments with the corresponding samples were the same as described for the LeBemSH3b-Ste20 complex in chapter 2, section 2.3.1. The same NMR experiments as described for the *L. elongisporus* complex (Chapter 2, section 2.3.1) were performed to assign chemical shifts and intramolecular/intermolecular NOEs. The same strategy was also used to process NMR data, derive input constraints, and perform structure calculation. The only difference was that while manually assigned NOE peaks were fixed in the first cycle of structure calculation in both cases, in the case of the NbpSH3-Ste20 complex calculation all NOE peak assignments in the remaining cycles were made by CYANA. Structural coordinates and NMR restraints have been deposited in the Protein Data Bank with accession number 2LCS.

4.3 Results

4.4.1 Mutational analysis of the Ste20 and Bck1 peptides

To compare the binding requirements of the NbpSH3 and BemSH3b domains, I used Trp fluorescence spectroscopy to quantitate the binding affinities of both SH3 domains to wild-type and mutant Ste20 peptides. This Ste20 site is a proven biological target of BemSH3b (Winters and Pryciak, 2005) and is also the peptide which bound with the highest affinity to both BemSH3 ($K_d=0.05 \mu M$) and NbpSH3 ($K_d=0.2 \mu M$) domains (Chapter 3). As shown in Figure 4.1A, substitution of residues comprising the $\Psi$PxRxAPxxP motif within Ste20 site significantly decreased the binding affinity of both the NbpSH3 and BemSH3b domains, confirming that both domains rely on the recognition of this motif for binding. As expected substitutions of arginine and first proline of the core RxxPxxP motif ((R(-3)A and P(0)A) caused greater than 100-fold reduction in the binding affinity of both the NbpSH3 and BemSH3b domains. However, substitution of the second proline (P(3)A) had a relatively small effect, decreasing the NbpSH3 and BemSH3b domain binding 10-fold and 6-fold
respectively. Significantly, substitution of the consensus residues in the extended region of the Ste20 peptide (F(-7)A and P(-5)A) caused a much larger reduction in the binding affinity of the BemSH3b domain than NbpSH3 domain. In particular, the F(-7)A substitution reduced the BemSH3b domain binding affinity 700-fold, while decreasing the NbpSH3 domain binding only 20-fold (Figure 4.1A). In general, substitutions of peptide residues in the extended region had a significantly larger effect on the binding of the BemSH3b domain as compared to the NbpSH3 domain. In contrast, substitution of Ste20 peptide core motif residues had a larger effect on the binding of the NbpSH3 domain. However, in the later case the difference was much less pronounced. These data demonstrate that while the NbpSH3 and BemSH3b domains both recognize the same consensus motif in Ste20 peptide, the BemSH3b domain is more dependent on the interactions with the extended region of the peptide than the NbpSH3 domain.

While the Ste20 peptide interacts tightly with both the NbpSH3 and BemSH3b domains, the Bck1 peptide binds the NbpSH3 and BemSH3b domains with quite different affinities of 0.8 μM and 26 μM respectively (Chapter 3). Mutational analysis of the Ste20 site (Figure 4.1A) suggests that the inability of BemSH3b to interact tightly with Bck1 site is due to weak interactions with the extended region of the peptide. To test this, I substituted the residues in the extended region of the Bck1 site for the corresponding residues in the Ste20 site. As shown in Figure 4.1B, all of the substitutions tested increased the binding of BemSH3b several fold, while a quadruple mutant containing all of the substitutions increased the binding affinity 81-fold to 0.31 μM. The affinity of the BemSH3b domain towards the Bck1 site was increased incrementally by each substitution, suggesting that the low affinity of the BemSH3b domain towards the Bck1 peptide is due to multiple suboptimal interactions with the extended region.
Figure 4.1 Mutational analysis of the Ste20 and Bck1 peptides, and identification of the N-term Cla4 and Boi1 sites displaying high selectivity for the BemSH3b domain.

(A) Binding of the NbpSH3 and BemSH3b domains to the wild-type and mutant Ste20 peptides. Ste20 residues that are part of the ΨxPxRxAPxxP consensus recognized by the NbpSH3 and BemSH3b domains are shown in bold. Peptide residue numbering is according to Lim et al. (Lim et al., 1994). The affinities of interactions and the fold change in binding relative to the wild-type Ste20 peptide are indicated. The values are shown as measurement mean value ± standard error. All measurements were repeated at least twice. The values marked by asterisk were taken from Chapter 3. (B) Binding of the NbpSH3 and BemSH3b domains to the wild type and mutant Bck1 peptides. (C) Interactions of the BemSH3b and NbpSH3 domains with the sites derived from the Boi1p/Boi2p homologues and Cla4p. The data in (B) and (C) is represented in the same way as in (A).
4.4.2 Structure of the NbpSH3-Ste20 complex

To determine the molecular mechanism of the NbpSH3 domain binding and to compare the interactions of the NbpSH3 and BemSH3 domains with the Ste20 peptide in molecular detail, I determined the solution structure of the NbpSH3 domain in complex with the Ste20 peptide. Standard NMR experiments were used to assign chemical shifts and collect NOE data (Kay, 1995). The structure was constrained based on intra-peptide, intra-SH3 and intermolecular NOEs, hydrogen bond restraints and dihedral angle restraints derived from TALOS (Table 4.1). The 20 lowest energy structures overlay extremely well with a backbone RMSD of 0.22Å and 0.15Å for well defined regions of the SH3 domain and Ste20 peptide respectively (Table 4.1, Figure 4.2A). The NbpSH3 domain displays a typical fold and overlays closely (backbone RMSD of 0.93Å over residues 1-59) with the previously determined crystal structure of the free NbpSH3 domain (PDB ID: 1YN8 chain B, Figure 4.2B). The major differences map to the RT-loop and N-Src-loop regions, which are directly involved in peptide binding. The Ste20 region involved in the interaction with the NbpSH3 domain includes positions (-7) to (3), while the rest of the peptide residues lack any NOEs to the SH3 domain and are disordered (Figure 4.2A).

Table 4.1. Structural statistics for the NbpSH3-Ste20 peptide complex

<table>
<thead>
<tr>
<th>Constraints</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input dihedral angles a (φ, ψ)</td>
<td>RMSD from mean structure (Å) c</td>
</tr>
<tr>
<td>NOE upper distance</td>
<td>Backbone atoms (SH3, -1 to 61) 0.22</td>
</tr>
<tr>
<td>Intraresidual</td>
<td>Heavy atoms (SH3 domain, -1 to 61) 0.67</td>
</tr>
<tr>
<td>Medium-range (1&lt;</td>
<td>i-j</td>
</tr>
<tr>
<td>Long-range (</td>
<td>i-j</td>
</tr>
<tr>
<td>Intermolecular</td>
<td>RMS deviation from ideal geometry</td>
</tr>
<tr>
<td>Hydrogen bond b</td>
<td>Distance (Å) 0.004</td>
</tr>
<tr>
<td>Lower</td>
<td>Angles (degrees) 0.8</td>
</tr>
<tr>
<td>Upper</td>
<td>Ramachandran plot</td>
</tr>
<tr>
<td>Structure Calculation</td>
<td>Favored regions 82.3%</td>
</tr>
<tr>
<td>Cyana Target Function (Å²)</td>
<td>Additionally allowed regions 17.7%</td>
</tr>
<tr>
<td>1st Cycle</td>
<td>Generously allowed regions 0.1%</td>
</tr>
<tr>
<td>Final</td>
<td>Disallowed regions 0.0%</td>
</tr>
<tr>
<td>RMSD from mean structure (Å) c</td>
<td></td>
</tr>
<tr>
<td>1st cycle Backbone/Heavy atoms</td>
<td>1.01/1.42</td>
</tr>
<tr>
<td>NOEs (%)</td>
<td></td>
</tr>
<tr>
<td>Unassigned</td>
<td></td>
</tr>
<tr>
<td>Discarded long-range</td>
<td></td>
</tr>
</tbody>
</table>

a Dihedral angles restraints were derived using TALOS (Cornilescu et al., 1999)
b Based on the predicted hydrogen bonds that were derived from the examination of secondary structure NOEs
c The RMSD is indicated for well defined regions. The NbpSH3 domain, residues -1 to 61 and the Ste20p peptide, residues (-7) to (3).
The core RXXPXXP motif of the peptide (positions (-3) to (3)) interacts with the NbpSH3 domain in a canonical manner (Figure 4.2D). The two XP dipeptides of the PPII helix (Ala(-1)Pro(0) and Pro(2)Pro(3)) fit into the conserved hydrophobic grooves. The Arg(-3) residue is packed against Trp36, and while its guanidinium group is disordered due to the lack of NOE data (Figure 4.2C), it lies near and is likely to make electrostatic interactions with the carboxyl group of Glu13 and Glu17. The side chains of other core motif residues including Pro(-2) and Lys(1) face away from the SH3 domain surface. The interaction of the Ste20 peptide extended region (positions -7 to -4) with the NbpSH3 domain is primarily mediated by Phe(-7) and Pro(-5), while the Ile(-6) and Ser(-4) side chains point away from the SH3 domain (Figure 4.2D). The Phe(-7) and Pro(-5) residues are packed against the hydrophobic surface of the NbpSH3 domain located between RT and N-Src loops.

**Figure 4.2 The structure of NbpSH3-Ste20 complex**

(A) An overlay of 20 lowest energy structures. The NbpSH3 domain is in blue and the Ste20 peptide is in magenta. Well defined peptide positions are numbered. (B) An overlay of the NbpSH3 domain crystal structure (green, 1YN8 chain B) with the NMR structure of the NbpSH3 domain (blue) in complex with the Ste20 peptide (magenta). (C) An overlay of five low energy Ste20 peptide conformers corresponding to peptide positions (-7)
to (3). (D) Surface representation of the NbpSH3 domain. The residues involved in the interaction with the Ste20 peptide are highlighted, hydrophobic residues are highlighted in green and negatively charged residues in red. The Ste20 peptide is shown in magenta.

4.4.3 Comparison of the Nbp2SH3-Ste20 and BemSH3b-Ste20 complexes

The NMR structure of the yeast BemSH3b domain in complex with the Ste20 peptide was determined by another group (Takaku et al., 2010), allowing for comparison of the Ste20 peptide binding mode between the NbpSH3-Ste20 and BemSH3b-Ste20 complexes. In addition to this structure, I have solved the NMR structure of the Ste20 peptide in complex with the BemSH3b homologue from *L. elongisporus* (described in Chapter 2). In order to strengthen my comparison of the Ste20 interaction details between the NbpSH3-Ste20 and BemSH3b-Ste20 complexes, I have also compared SH3-peptide interaction details observed in the yeast BemSH3b-Ste20 complex to those in the LeBemSH3b-Ste20 complex (see below and Figure 4.3) with an expectation that peptide interaction details conserved between both complexes are likely to be correct. Such comparison is important, since NMR structures are sometimes unable to provide accurate positions for all the residues due to the lack of NOE data or other factors.

Comparison of the NbpSH3-Ste20 and BemSH3b-Ste20 complexes shows that the Ste20 peptide binds the NbpSH3 and BemSH3b domains in a similar conformation and follows the same trajectory across the surfaces of both SH3 domains (Figure 4A). The complexes overlay with a backbone RMSD of 1.58Å (between the lowest energy models) over the well defined regions in both structures. These correspond to the SH3 domain residues 2-40 and 48-58, and the peptide residues (-7) to (-3). In both structures the rest of the peptide sequence is disordered (Takaku et al., 2010). Similar to the interactions observed in NbpSH3-Ste20 complex, the Ste20 peptide core motif interacts with the BemSH3b domain in a canonical manner, while the interaction of the Ste20 peptide extended region with the BemSH3b domain is primarily mediated through Phe(-7) and Pro(-5) residues (Figure 4.4A). Thus, overall the Ste20 peptide interacts with the NbpSH3 and BemSH3b domains in a comparable manner, however a more detailed analysis of these interactions presented below reveals a number of differences between the two complexes.
Figure 4.3 Comparison of SH3-peptide interaction details between the yeast and *L. elongisporus* BemSH3b-Ste20 complexes

(A) Interactions with the Ste20 peptide core motif observed in the yeast BemSH3b-Ste20 (PDB ID: 2RQW) and LeBemSH3b-Ste20 (PDB ID: 2KYM) complexes. The BemSH3b domain residues are shown in green and the bound Ste20 peptide in orange. The LeBemSH3b domain residues are in cyan and the Ste20 peptide in salmon. SH3 domains residues are numbered according to Larson et al (Larson and Davidson, 2000). (B) Interactions with the extended region of the Ste20 peptide. Colors and labels are the same as in (A).

### 4.4.4 The NpbSH3 domain forms stronger interactions with the core motif of the Ste20 peptide

Although the core RXXPXXP motif of the Ste20 peptide interacts with both SH3 domains in a similar manner (Figure 4.4C), it appears to form stronger interactions with the NbpSH3 domain than the BemSH3b domain. The most noticeable difference is the binding of the Arg(-3) residue (Figure 4.4B). While Arg(-3) packs against Trp36 and contacts Glu13 in both the NbpSH3-Ste20 and BemSH3b-Ste20 complexes, the examination of multiple low energy structures suggests that this association is closer in the NbpSH3-Ste20 complex. There is a potential for a hydrogen bond between the Arg(-3) backbone carboxyl and the Trp36 side...
chain in the NbpSH3-Ste20 complex, which is absent in the BemSH3b-Ste20 complex. In the 20 lowest energy structures of the NbpSH3-Ste20 complex the average distance between Arg(-3) CO and Trp36 Hε is 1.79 Å, with the donor-hydrogen-acceptor angle of 170°, which suggests a formation of a strong hydrogen bond. In contrast the average distance between Arg(-3) CO and Trp36 Hε in the 20 lowest energy structures of BemSH3b-Ste20 complex is 2.77 Å, which is greater than the distance of 2.5 Å required to satisfy hydrogen bonding potential (McDonald and Thornton, 1994). The same is true for LeBemSH3b-Ste20 complex with the average distance between Arg(-3) CO and Trp36 Hε of 3.15 Å.

The Arg(-3) side chain is also more disordered in the BemSH3b-Ste20 complex relative to the NbpSH3-Ste20 complex (Figure 4.4B), suggesting that the interaction with the Arg(-3) side chain may be stronger in the NbpSH3-Ste20 complex. Consistent with this prediction, the R(-3)A substitution caused a 3-fold greater decrease in the binding of the NbpSH3 domain than the BemSH3b domain (Figure 4.1A). It should be noted that Glu17, conserved between the NbpSH3 and BemSH3b domains, appears to interact with the Arg(-3) side chain in the NbpSH3-Ste20 complex, but points away from the peptide in the published BemSH3b-Ste20 complex (Figure 4.4B). In contrast, I find that Glu17 residue is oriented towards Ste20 peptide in LeBemSH3b complex (Figure 4.3) and residues at this position are similarly oriented in other SH3 domain complexes (Larson and Davidson, 2000) suggesting that the interaction between Arg(-3) and Glu17 is also present in the BemSH3b-Ste20 complex.

The weaker interaction between the BemSH3b domain and the Arg(-3) residue is likely due to the Lys14 residue, which I have previously demonstrated to inhibit the binding of the BemSH3b domain (Chapter 3). In the BemSH3b-Ste20 complex the Lys14 side chain crosses the binding interface and thus appears to interfere with binding of Arg(-3) both through steric and electrostatic repulsion (Figure 4.4B). Consistent with this, the K14A substitution does not alter the BemSH3b affinity for the Ste20 peptide bearing R(-3)A substitution even though it increases the affinity for the wild-type peptide by more than 5-fold (Figure 4.4B). In the NbpSH3 domain position 14 is occupied by an Asn residue and its substitution with alanine causes little change in the binding affinity of the NbpSH3 domain (Figure 4.5A).
Figure 4.4 Comparison of the NbpSH3-Ste20 and BemSH3b-Ste20 complexes

(A) Overlay of the NbpSH3-Ste20 and BemSH3b (PDB ID: 2RQW) complexes (backbone RMSD 1.58Å). The Nbp2SH3 is in blue with the bound Ste20 peptide in magenta. BemSH3b is in green with the bound Ste20 peptide in orange. (B) Interaction of Ste20 peptide Arg(-3) with the NbpSH3 and BemSH3b domains. Residues involved in the interactions are shown as sticks and are represented by five low energy models. The hydrogen
bond between Trp36 and Arg(-3) present in the NbpSH3-Ste20 complex is indicated by a dashed line. The
binding affinities of the wild-type BemSH3b domain or K14A mutant for the wild-type Ste20 peptide or R(-3)A
mutant are shown. Comparison of these affinities demonstrates that the inhibitory effect of the BemSH3b Lys14
residue on peptide binding depends on the interaction with the Arg(-3) of the peptide. (C) and (D) Detailed
comparison of the NbpSH3 and BemSH3b domain interactions with the core and extended regions of the Ste20
peptide. Schematic diagram of the interactions is depicted below the structures. Coloring is the same as in (A).

The stronger interaction of the NbpSH3 domain with the Ste20 peptide core motif is also
reflected in closer interactions with the PPII helix. In both complexes the first dipeptide of the
PPII helix (Ala(-1) and Pro(0)) fits into the hydrophobic groove lined by the same residues
(Trp36, Phe12, Pro51 and Phe54) in the NbpSH3 and BemSH3b domains. However this
interaction is more extensive in the NbpSH3-Ste20 complex, where Ala(-1) is intimately
packed against Trp36 (Figure 4.4C). By contrast, in the BemSH3-Ste20 complex, the Ala(-1)
side chain points away from Trp36 and the same positioning of Ala(-1) is observed in the
LeBemSH3b-Ste20 complex (Figure 4.3). The A(-1)T substitution also caused a 3-fold
greater decrease in the binding of NbpSH3 relative to the BemSH3b domain (Figure 4.1A).

4.4.5 BemSH3b domain interacts more extensively with the extended region of Ste20
peptide
In contrast to the NbpSH3 domain, which interacts more extensively with the core motif of the
Ste20 peptide, BemSH3b appears to form stronger interactions with the extended region. This
is largely due to the additional peptide binding surface provided by the CI subdomain of the
BemSH3b (Figure 4.4D). In the BemSH3b-Ste20 complex peptide Phe(-7) is packed at the
hydrophobic interface between the SH3 and CI subdomains, composed of SH3 residues
Asp16, Ile38, Pro47 and CI residues Lys87 and Val83 (Figure 4.4D). Similar interactions are
also observed in the LeBemSH3b-Ste20 complex (Figure 4.3). Although in the NbpSH3-Ste20
complex the Phe(-7) side chain interacts with a similar surface of the SH3 domain composed
of residues Asn16, Tyr30, Val38, Thr47 and Leu49, it lacks the additional interactions
provided by the CI subdomain of the BemSH3b domain. Interestingly, the Tyr30 residue in
the NbpSH3 domain (in the BemSH3b domain this position is occupied by Ala residue that is
completely buried at the interface of the SH3 and CI subdomains) spatially overlaps with the
Val83 residue in the BemSH3b CI subdomain and thus appears to at least partially recapitulate
the additional hydrophobic surface provided by the BemSH3b CI subdomain. The CI
subdomain of BemSH3b also interacts with peptide residue Ile(-6) through the Val83 and
Gln84 residues, while the NbpSH3 domain does not form any contacts with this residue (Figure 4.4D). Consistent with the importance of the BemSH3b CI subdomain for the interaction with peptide Phe(-7) and Ile(-6), alanine substitutions of these residues had a much larger effect on the binding of BemSH3b domain than on the NbpSH3 domain (Figure 4.1A).

Finally, in both the NbpSH3-Ste20 and BemSH3b-Ste20 complexes, peptide residue Pro(-5) is accommodated in the pocket composed of His32, Trp36 and Leu49, which are conserved between the two SH3 domains. However the orientation of His32 differs significantly between the two complexes. In the BemSH3b domain, His32 is oriented towards the peptide and is tightly packed against Pro(-5), while in the NbpSH3 domain His32 points away from the peptide. I previously demonstrated that while the H32A substitution decreases the binding affinity of the BemSH3b domain, it actually increases the binding affinity of the NbpSH3 domain (Chapter 3, Figure 4.5A). The partial burial of His32 in the hydrophobic interface between the SH3 and CI subdomains may account for the different orientation and binding contributions of this residue in the NbpSH3 and BemSH3b domains (Figure 4.4D). Consistent with the positive contribution of His32 toward peptide binding in the case of BemSH3b and negative contribution in the case of NbpSH3, the binding of BemSH3b domain was significantly more affected by the Ste20 peptide P(-5)A substitution (Figure 4.1A).

4.4.6 Identification of the N-terminal Cla4 and Boi1 sites that are highly selective for the BemSH3b domain

The larger dependence of the BemSH3b domain on the interaction with the extended region of the Ste20 peptide suggested to me that peptides which contain suboptimal core motif would be selective for the BemSH3b domain, since it is less dependent on the interactions with the core motif relative to the NbpSH3 domain. In Chapter 3, I have identified a single peptide derived from the Boi2p, which showed high selectivity for the BemSH3b domain, binding it with the 25-fold higher affinity compared to the NbpSH3 domain (Figure 4.1). However, in the case of this site, it is not clear how the specificity is established since the core motif consensus (RxAPxxP) is conserved. As mentioned in the introduction, the Boi2p homologue-Boi1p, contains a similar site, but one in which the Arg(-3) residue is replaced with the Gly residue creating a suboptimal PxxP motif which lacks the critical negative residue. According to my analysis, such site should be selective for the BemSH3b domain, which is less dependent on
the interactions with the core motif and in particular the Arg (-3) residue due to the presence of the Lys14 residue which weakens Arg(-3) mediated interactions. Indeed, the Boi1 peptide bound the BemSH3b domain with an affinity of 0.94 μM and the NbpSH3 domain with such low affinity (>200 μM) that it could not be accurately quantified, confirming my prediction (Figure 4.1C). Interestingly, a similar site which contains Gly residue instead of the Arg residue at the (-3) position, is also present at the N-terminus of the Cla4p (a proven biological target of the BemSH3b domain). It should be stressed, that this Cla4 site (N-term Cla4) is different from the Cla4 site characterized in Chapter 3 (Cla4), which conforms to the ΨxPxRxAPxxP consensus and binds both the NbpSH3 and BemSH3b domains with high affinity (Figure 4.1C). Similar to the Boi1 peptide, the N-term Cla4 peptide also bound the BemSH3b domain with significantly higher affinity compared to the NbpSH3 domain (Figure 4.1C). Thus I have identified two yeast sites, both of which lack the Arg(-3) of the RxxPxxP motif and bind the BemSH3b domain with >100-fold higher affinity than the NbpSH3 domain, confirming my prediction that sites with the suboptimal core motif should display strong preference for the BemSH3b domain.

4.4.7 Mutational analysis of the NbpSH3 and BemSH3b domains

In order to determine which NbpSH3 and BemSH3b residues are critical for peptide binding and compare their contribution to binding between the two SH3 domains, I have tested the effect of amino acid substitutions at several positions in the NbpSH3 and BemSH3b domains (Figure 4.5A). Strikingly, the alanine substitution of the Tyr8 residue had a significantly larger effect on the NbpSH3 domain binding, decreasing its binding to Ste20 peptide 46-fold, while only affecting the binding of the BemSH3b domain 6-fold. Since the Tyr8 residue interacts with the core peptide motif (Figure 4.4), this result provides further support for the existence of stronger interactions with the core peptide motif in the NbpSH3-Ste20 complex, relative to the BemSH3b-Ste20 complex.

It is notable that E13A substitution had a relatively minor effect on the binding of both SH3 domains (Figure 4.5A), even though this residue appears to play a main role in the interaction with the positively charged Arg(-3) of the peptide in both complexes (Figure 4.4B). Another negatively charged residue which is poised to interact with peptide Arg(-3) is Glu17. While E17A caused significant destabilization of the NbpSH3 domain and prevented isolation of the
BemSH3b domain due to poor solubility, a more conservative substitution at this position (E17Q) produced proteins with stability similar to the wild-type domains (Figure 4.5A). However, both the NbpSH3 and BemSH3b domains bearing E17Q substitution bound to the Ste20 peptide with such low affinity that it could not be accurately measured. Therefore, Glu17 plays a major role in the binding of both NbpSH3 and BemSH3b domains. It should be noted that this position contributes to the structure of the RT loop (Larson and Davidson, 2000) and the Ste20 mutant containing R(-3)A substitution still bound to the NbpSH3 and BemSH3b domains. Thus the disruptive effect of the E17Q substitution on binding is not solely due to the loss of interaction with the Arg (-3) of the peptide.

4.4.8 Identification of NbpSH3 residues Tyr30 and Val38 as important mediators of its specificity

Structural analysis of the BemSH3b domain demonstrates that its unique CI subdomain is largely involved in the interaction with the extended region of Ste20 peptide and likely makes a major contribution to the specific recognition of the ΨxPxRxAPxxP consensus in the Ste20 peptide by the BemSH3b domain. However, it is not clear which residues are responsible for the recognition of the ΨxPxRxAPxxP consensus in the case of the NbpSH3 domain. In order to understand this, I have examined the alignment of all yeast SH3 domains that like the NbpSH3 domain, recognize class I binding motifs (Chapter 1, Table 1.2) and thus have potential to compete with for binding (Figure 4.5B). As shown in the Figure 4.5B, the NbpSH3 domain residues Ty30 and Val38 stand out as highly unusual at these positions. The NbpSH3 domain is the only SH3 domain that contains large hydrophobic residue at position 30, while the only other SH3 domain that has β-branched aliphatic residue at position 38 is the BemSH3b domain. In the NbpSH3-Ste20 complex these residues interact with the peptide Phe(-7) side chain, which is part of the ΨxPx extended region consensus recognized by the NbpSH3 domain. These residues are thus good candidates for establishing the specificity of the NbpSH3 domain for the sites containing ΨxPxRxAPxxP consensus relative to the other RxxPxxP sites. To examine energetic contribution of these residues towards peptide binding I have tested the effect of alanine substitutions of these positions. Significantly, V38A and Y30A substitutions decreased the NbpSH3 domain binding 19-fold and 13-fold respectively (Figure 4.5B), which is comparable to the 21-fold decrease in binding observed with the Ste20
F(-7)A substitution (Figure 4.1A). Additionally substitution of Val38 with Lys, which is the most common residue found at this position in other SH3 domains (Figure 4.5B), also caused a large decrease in binding, reducing the NbpSH3 domain affinity towards the Ste20 peptide 22-fold. Thus Tyr30 and Val38 residues are unique to the NbpSH3 domain and are also critical for the interaction with the Ste20 peptide Phe(-7), which identifies these residues as important mediators of the NbpSH3 domain binding specificity.

<table>
<thead>
<tr>
<th></th>
<th>NbpSH3+Ste20</th>
<th>BemSH3+Ste20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SH3 Kd (µM)</td>
<td>Fold change</td>
</tr>
<tr>
<td>WT</td>
<td>0.2 ± 0.01</td>
<td>1</td>
</tr>
<tr>
<td>Y8A</td>
<td>9.1 ± 2.0</td>
<td>46</td>
</tr>
<tr>
<td>E13A</td>
<td>2.5 ± 0.2</td>
<td>13</td>
</tr>
<tr>
<td>N14A</td>
<td>0.13 ± 0.0</td>
<td>0.65</td>
</tr>
<tr>
<td>E17A no binding</td>
<td>-11.0</td>
<td>280</td>
</tr>
<tr>
<td>E17Q &gt;200</td>
<td>1000</td>
<td>48</td>
</tr>
</tbody>
</table>

Figure 4.5 Mutational analysis of NbpSH3 and BemSH3b domains and identification of residues important for establishing NbpSH3 domain specificity

(A) Effect of alanine substitutions in the NbpSH3 and BemSH3b domains on the binding affinity to the Ste20 peptide. The affinities of interactions and the fold change in binding relative to the wild-type SH3 domains are shown. Each measurement was repeated at least twice. The Kd values marked by asterisk were taken from Chapter 3. ΔTm values represent the change in temperature induced denaturation midpoint with respect to the wild-type domain, where negative values indicate the decreased stability of mutant domains. The Tm values for the wild-type NbpSH3 and BemSH3 domains are 53.7°C and 54.6°C respectively. (B) The alignment of the yeast SH3 domains that similar to the NbpSH3 domain recognize class I motifs (Table 1.2) and are thus potential competitors of the NbpSH3 domain. Numbering is according to Larson et al (Larson and Davidson, 2000). Positions that in NbpSH3 domain are involved in peptide binding, according to the structure of the NbpSH3-Ste20 peptide complex, are indicated by ‘P’. Positions 30 and 38 that are occupied by the residues that are unique to the NbpSH3 domain (Tyr30) or to the NbpSH3 and BemSH3b domains (Val38 and Ile38) are boxed. Binding preferences of the yeast SH3 domains were taken from the large scale phage display studies (Tong et al., 2004; Tonikian et al., 2009).
4.4 Discussion

I have demonstrated that although the NbpSH3 and BemSH3b domains recognize the same consensus binding sequence, they use a different mechanism for binding. Relative to the BemSH3b domain, the NbpSH3 domain forms stronger interactions with the core motif of Ste20 peptide, while the BemSH3b domain compensates for weaker interactions with the core motif by forming stronger interactions with the extended region of Ste20 peptide. This observation is supported both by the structural and mutagenesis analysis of the NbpSH3 and BemSH3b domain interactions with the Ste20 peptide. Substitution of each position in the Ste20 peptide demonstrated that the binding of the BemSH3b domain is significantly more affected by the substitutions in the extended region of the peptide. The combined effect of alanine substitutions in this region on binding is 4000-fold greater in the case of the BemSH3b domain relative to the NbpSH3 domain (Figure 4.1A). On the other hand, substitutions to the core motif in the Ste20 peptide had a larger effect on the NbpSH3 domain binding. Although in this case the difference was less pronounced, where the combined effect of all substitutions on binding was approximately 60-fold greater in the case of the NbpSH3 domain, it is nonetheless significant (Figure 4.1A). Comparison of the NbpSH3-Ste20 and BemSH3b-Ste20 complexes demonstrated that the BemSH3b domain interacts more extensively with the peptide residues Phe(-7), Ile(-6) and Pro(-5) due to the presence of unique C-terminal extension that contributes to the peptide binding interface (Figure 4.4D). On the other hand, the NbpSH3 domain forms more intimate contacts with the Arg(-3) and Ala(-1) peptide residues of the core motif (Figure 4.4C). The stronger interactions of the NbpSH3 domain in this region can be at least partially explained by the presence of the inhibitory Lys14 residue in the BemSH3b domain. Finally, the significantly larger effect of the Y8A substitution on the binding of the NbpSH3 domain relative to the BemSH3b domain, also demonstrated that the core motif interactions are stronger in the NbpSH3 domain (Figure 4.5A).

The different mechanism used by the NbpSH3 and BemSH3b domains to interact with the Ste20 peptide explains how these domains are able to simultaneously display the same affinity for some yeast peptides, while having quite different affinities for other peptides (Gorelik et al., 2011). In the case of sites including Ste20 and Cla4, which both SH3 domains bind with high affinity, although the strength of core motif and extended motif interactions are different
between the two SH3 domains, the sum of these interactions amounts to a similar strength of binding. On the other hand, there are sites such as Bck1 and Skm1, which the NbpSH3 domain binds with higher affinity. In this case, the weaker interaction of the BemSH3b domain with the extended region of these peptides is not sufficient to compensate for suboptimal interactions with the core motif that according to my analysis are at least partly caused by the Lys14 residue of the BemSH3b domain. This is consistent with the data described in Chapter 3, showing that a BemSH3b K14A substitution substantially increases the affinity of the BemSH3b domain towards Skm1 and Bck1 sites. In addition, substitution of residues in the extended region of the Bck1 site with the corresponding residues from the Ste20 site, resulted in the incremental increase in the affinity of the BemSH3b domain. It is noteworthy that each of the four substitutions increased the affinity in an almost additive manner, where multiplying the effect of individual substitutions produces a theoretical 180-fold increase in binding that is similar to the 80-fold increase obtained with the quadruple mutant (Figure 4B). These data demonstrates that much lower affinity of the BemSH3b domain towards the Bck1 site relative to the Ste20 site is due to multiple suboptimal interactions with the extended region of the peptide.

The opposite is true in the case of the Boi1, Boi2 and N-term Cla4 sites, which bind the BemSH3b domain with significantly higher affinity than the NbpSH3 domain (Figure 4.1C). In this case, the interaction of the NbpSH3 domain with the extended region of these peptides is not sufficient to overcome suboptimal interactions with the core motif. Although it is not clear, why the binding of the core motif is compromised in the case of the Boi2 site, it is probably the lack of Arg at the (-3) position in the Boi1 and the N-term Cla4 sites which prevents them from binding the NbpSH3 domain with high affinity. However, the BemSH3b domain which is less dependent on the interaction with the peptide Arg(-3) due to the presence of the inhibitory Lys14 residue and forms stronger interactions with the extended regions of these peptides can still bind with high affinity. It was previously suggested that SH3 domains may evolve to be highly specific by weakening the interactions with the peptide core motif, which is recognized by many SH3 domains, while strengthening interactions with the extended peptide motif (Kim et al., 2008). The binding mechanism employed by the BemSH3b domain relative to the NbpSH3 domain exemplifies exactly such situation. The
interactions of the BemSH3b domain with the core motif are weakened by the Lys14 residue, while the interactions with the extended region are strengthened by the contribution of the CI subdomain. Consequentially, the BemSH3b domain is capable of recognizing sites such as N-term Cla4 and Boi1, which lack the canonical core RxxPxxP motif and hence displays higher level of specificity relative to the NbpSH3 domain. Consistent with this, the BemSH3b K14A substitution which strengthens the interaction of the BemSH3b domain with the peptide core motif also reduces its specificity (Gorelik et al., 2011)

Although SH3 domain mediated interactions are thought to possess low affinity and specificity (Cussac et al., 1994; Feller et al., 1995; Kaneko et al., 2003; Landgraf et al., 2004; Viguera et al., 1994) numerous studies have demonstrated that SH3 domains are capable of achieving high affinity and specificity through interactions with the extended region of target peptides (Bauer et al., 2005; Ghose et al., 2001; Kami et al., 2002; Lewitzky et al., 2004; Stollar et al., 2009). The interactions of the NbpSH3 and BemSH3b domains with the extended region of the Ste20 peptide provide another such example. In the case of the BemSH3b domain, its unique CI subdomain provides additional strength to the interaction with the extended region of Ste20 peptide producing a high affinity and specificity interaction. In the case of the NbpSH3 domain, unique residues Tyr30 and Val38, contribute to the affinity and specificity through the interaction with the Phe(-7) residue of the Ste20 peptide. The increase in the affinity resulting from these interactions is likely to be biologically significant. Reducing the binding affinity of the NbpSH3 domain for the Ste20 peptide 45-fold through the Y8A substitution (comparable to the 20-fold decrease obtained with the Ste20 peptide F(-7)A and the NbpSH3 domain V38A substitutions) produced a significant defect in the biological function of the NbpSH3 domain (Chapter 3, Figure 3.3). Characterization of SH3 domain interactions with the peptide extended regions is an important objective, since their comparison between different SH3-peptide complexes would allow generalization of different mechanisms used by the SH3 domains to achieve high affinity and specificity binding. For example, where the NbpSH3 domain uses Tyr30, Val38 and Leu49 to interact with the Phe(-7) of the Ste20 peptide, the mammalian IRTKS SH3 domain uses Ile30, Tyr38 and Trp49 to accommodate Pro(-7) and Ile(-8) of its target peptide (Aitio et al., 2010). Considering a relatively small number of well characterized interactions between SH3
domains and extended peptides, the similarity of interactions observed between NbpSH3 and IRTKS SH3 domains is likely to extend to other, yet uncharacterized, SH3-peptide complexes.

In summary, my structural and binding studies on the NbpSH3 and BemSH3b domains demonstrate that these domains use different mechanism for binding, while recognizing the same consensus sequence. This is achieved by modulating the strength of interactions with different regions of the peptide and allows these domains to share specificity for some targets, while maintaining different specificity for other targets (Chapter 3). My findings highlight the ability of SH3 domains to possess finely tuned specificity and reinforce the importance of peptide residues flanking the core motif for achieving high affinity and specificity in SH3-peptide interactions.
Chapter 5 : Conclusions and Future directions

5.1 Conclusions
One of the main arguments against the ability of the SH3 domains and other protein interaction domains to possess high intrinsic specificity is that the domains within a single family all rely on the recognition of the same core motif. Yet my work demonstrates that even in the case of the NbpSH3 and BemSH3b domains that recognize not only the same core motif, but also the same extended sequence, each domain possesses its own unique binding specificity (Chapter 3). First, I showed that although NbpSH3 and BemSH3b domains recognize some yeast sites with a similar affinity, they display greater than >20-fold differences in affinities towards the other yeast sites including Bck1, Skm1, Boi1/Boi2 and N-term Cla4 (Chapter 3, Figure 3.2; Chapter 4, Figure 4.1). Second, I confirmed that this difference in specificities observed in vitro is also biologically relevant. I showed that the BemSH3b domain is not able to substitute for the NbpSH3 domain in vivo, while the BemSH3b domain KAD→NDN mutant that binds “NbpSH3 specific sites” with high affinity is able to do so (Chapter 3, Figure 3.3). Most importantly, I showed that reducing the specificity of the BemSH3b domain through KAD→NDN or K14A mutations impairs its ability to function in vivo, despite the increased affinity of these mutants towards all the yeast sites tested (Chapter 3, Figure 3.5). These results demonstrate that the specificity of the SH3 domains and other protein interaction modules can be finely nuanced and that the intrinsic specificity plays an important role in the function of these domains.

In Chapter 4, I have explored the molecular basis for the observed difference in specificity between the NbpSH3 and BemSH3b domains. My results demonstrate that this difference in specificities is achieved through modulation of interactions with the core and the extended regions of the target peptides. Relative to the NbpSH3 domain, the binding of the BemSH3b domain is more dependent on the interactions with the extended region of a target peptide and less on the interactions with the core motif. This is due to the unique C-terminal extension in the BemSH3b domain that interacts with the extended region of the target peptides and the inhibitory Lys14 residue, which interferes with the binding to the core motif. Thus the presence of the suboptimal residues in the extended regions of the target peptides produces
sites such as Bck1 and Skm1, which bind the BemSH3b domain with low affinity. I have confirmed this for the Bck1 site, showing that substitution of the extended region residues from the Ste20 site (BemSH3b binds Ste20 with the highest affinity) produced an incremental increase in the binding affinity of the BemSH3b domain (Chapter 4, Figure 4.1). It is noteworthy that no single substitution produced a significant increase in binding, suggesting that a strong interaction between the BemSH3b domain and the extended region of a peptide depends on multiple favourable interactions rather than a single critical interaction. On the other hand, the binding of the NbpSH3 domain is more dependent on the interactions with the peptide core motif relative to the BemSH3b domain. This is because the NbpSH3 domain lacks interactions formed between the BemSH3b CI subdomain and the extended region of the peptides. Thus sites such as Boi1 and N-term Cla4, which lack the Arg(-3) residue of the core RxxPxxP motif, bind the NbpSH3 domain with very low affinity, while still binding the BemSH3b domain with high affinity. Overall the comparison of the NbpSH3 and BemSH3b binding mechanisms demonstrates that the specificity of these domains is finely tuned by modulating the strength of multiple interactions. This highlights the underlying complexity in the binding behaviour of the SH3 domains, which should be accounted for in the future studies aimed at predicting and exploring SH3 domain specificity.

An important finding in my work is that SH3 domain mediated interactions can be regulated by negative interactions that function to maintain specificity (Chapter 3). I found that both BemSH3b (Lys14) and NbpSH3 (His32) domains contain residues that inhibit binding, where removal of these residues by alanine substitutions increased the binding of the domains to non-specific targets (Figure 3.7). In addition to general insight into the mechanisms of SH3 domain specificity, this also has important practical implications. For example, this should be taken into account in the design of macromolecules intended for the inhibition of SH3 domain-peptide interactions. Such inhibition may be desirable since SH3 domains mediate interactions in signaling pathways that are misregulated in tumor cells and thus present attractive targets for anti-cancer therapies (Vidal et al., 2001). Many studies have pursued this goal by selecting small molecules or peptide based analogs that display the highest affinity towards the targeted SH3 domain (Atatreh et al., 2008; Ferguson et al., 2004; Jacquot et al., 2007; Santiveri et al., 2009). However, based on my results, higher affinity can come at the
expense of lower specificity if it is achieved through neutralizing the interactions which inhibit binding and function to maintain specificity. This may pose a problem for the development of the SH3 domain inhibitors based on tight binding since selection for high affinity may produce molecules with reduced specificity.

My work with the NbpSH3 and BemSH3b domains will guide future studies focused on the function of these proteins. First, my characterization of the BemSH3b binding properties (Chapter 2) suggested a model for the BemSH3b domain role in Cdc42p mediated activation of the PAK kinases. As described below (Section 5.2.1), this model can be tested in the future studies aimed at exploring the in vivo function of the BemSH3b domain. Second, the affinities of the interactions observed between the NbpSH3 and BemSH3b domains and the tested yeast sites can serve to identify biologically relevant interactions. For example the Ste20 and Cla4 sites, which bind the BemSH3b domain with high affinity, are proven biological targets of this domain (Chapter 1, Table 1.3). In contrast the Bck1 site, which binds the NbpSH3 domain with high affinity and the BemSH3b domain with low affinity, is a biological target of the NbpSH3 domain. Therefore the Skm1 site, which mirrors the binding behavior of the Bck1 site, is also likely to be the biological target of the NbpSH3 domain, but not the BemSH3b domain. Similarly the Boi2 site, which binds the BemSH3b domain with a high affinity, is likely to interact with the BemSH3b domain in vivo, but not the NbpSH3b domain. The Ste20 and Cla4 sites, which bind both the NbpSH3 and BemSH3b domains with high affinity, probably interact with both domains in vivo and it is possible that under some conditions the competition of the NbpSH3 and BemSH3b domains for these sites plays a functional role.

5.2 Future directions
Further studies aimed at providing additional insights into BemSH3b function and the SH3 domain specificity are described in the following sections.

5.2.1 Further investigation of BemSH3b function through in vivo experiments
In chapter 2 of this thesis, I proposed a model for the BemSH3b role in the activation of a PAK kinase. According to this model, the function of the BemSH3b domain is to co-localize the activated Cdc42p-GTP with PAK, in order to promote Cdc42p-GTP binding to PAK and its subsequent activation. The validity of this model can be examined by disrupting the
BemSH3b interactions with Cdc42p and PAKs and testing the effect of this on PAK activation. For example, the activation of the Ste20p kinase in the mating pathway can be measured by observing the induction of the mating pathway. If this model is correct, then disrupting either the BemSH3b-PAK, BemSH3b-Cdc42p interaction, or both should produce the same defect in PAK kinase activation. Although similar experiments were performed in the Yamaguchi et al. study (Yamaguchi et al., 2007), their results can be improved through the use of different mutations and further experiments. As mentioned in the introduction (Chapter 1, section 1.3.5), while they tested the effect of disrupting the BemSH3b-Ste20p interaction on its own and the effect of disrupting the BemSH3b interaction with both Ste20p and Cdc42p (which produced the same defect), the effect of disrupting the BemSH3b-Cdc42p interaction on its own was never tested. In addition, the Bem1p N253D mutation that was used to disrupt BemSH3b-Cdc42p interaction in this study may not completely abolish binding. The equivalent substitution in the LeBemSH3b homologue (S103D) disrupted binding only 5-fold (Chapter 2, Figure 2.5B). According to my results, combining the Bem1p N253D mutation with the L201A mutation (equivalent to the BemSH3b L45A mutation which caused the biggest decrease in binding) should ensure the complete disruption of the BemSH3b-Cdc42p interaction. The Bem1p W192K substitution, which was used to disrupt the Bem1p-Ste20p interaction (equivalent to BemSH3b W36K substitution), may also affect the stability of the BemSH3b domain. Despite several attempts, I was not able to purify BemSH3b containing W36K substitution, making results obtained with this mutation difficult to interpret.

According to my results obtained in Chapter 4, the BemSH3b E17Q mutation, which had the strongest effect on BemSH3b binding without affecting its stability (Chapter 4, Figure 4.5A), would be more appropriate to specifically disrupt the BemSH3b-Ste20p interaction.

It is likely that the interaction between the BemSH3b and Cdc42p functions independently of PAK activation in some cellular processes, such as establishment of cell polarity. In this case, disrupting this interaction in the context of the already disrupted Bem1p-PAK interaction should produce a worse phenotype. Indeed, Yamaguchi et al. showed that disrupting both interactions through Bem1p W192K and N253D mutations produced a bigger effect on cell growth than either mutation on its own (Yamaguchi et al., 2007). Disrupting BemSH3b mediated interactions with PAKs and Cdc42p individually and in combination, and then
testing the effect of this on Bem1p function under different conditions, would provide insight into how often the binding of both Cdc42p and PAK is required.

5.2.2 Expanding specificity studies to other yeast SH3 domains

In Chapter 3, I used a pull down assay with purified SH3 domains and yeast extract to assess the specificity changes associated with the BemSH3b K14A and NbpSH3 H32A mutations. My main goal was to show that the BemSH3b K14A mutation decreases the specificity of the BemSH3b domain, but this assay also demonstrated that the specificity of the BemSH3b domain is higher than of the NbpSH3 domain (Chapter 3, Figure 3.7). This was shown by a significantly higher intensity of the protein bands pulled down by the NbpSH3 domain relative to the BemSH3b domain. As I demonstrated in Chapter 4, the NbpSH3 domain is more dependent on the interactions with the core peptide motif than the BemSH3b domain, which is consistent with the NbpSH3 domain having lower specificity. Thus the pull down assay with purified SH3 domains and proteins in yeast extract may provide a relatively straightforward method to assess the specificity of a yeast SH3 domain.

This assay can be applied to every yeast SH3 domain in order to assess different levels of SH3 domain specificity in yeast proteome. To improve the resolution, the proteins pulled down by the SH3 domains can be separated by running 2D protein gels, instead of the conventional 1D gel, that I used in my experiments. The proteins pulled down by multiple SH3 domains would likely represent non-specific interactions. At the same time, the amount of these proteins obtained with a given SH3 domain would provide with a relative measure of the intrinsic specificity possessed by this SH3 domain. The results obtained for different SH3 domains can be compared to the results obtained with the SH3 domains that are known to be highly specific (for example Sho1p SH3 (Zarrinpar et al., 2003) or Abp1p SH3 domains (Landgraf et al., 2004)). These experiments would allow a ranking of the yeast SH3 domains from the least to the most specific and produce an overall view of different specificity levels displayed by the yeast SH3 domains. The results would also be valuable for the identification of the biological targets for a given SH3 domain. The peptides bound by the SH3 domains with high specificity are more likely to be true biological targets than the peptides bound by the SH3 domains with low specificity.
5.2.3 Investigation of the determinants of the NbpSH3 domain specificity with respect to other yeast SH3 domains

Based on the results in Chapter 4, I postulated that the specificity of NbpSH3 domain towards the ΨxPxRxAxxP consensus is largely determined by the Tyr30 and Val38 residues. I based this conclusion on the fact that Tyr30 and Val38 are unique to the NbpSH3 domain (among yeast SH3 domains that recognize class I motif) and that these residues contact the extended region of the Ste20 peptide. However, in order to confirm this prediction further experiments with other yeast SH3 domains are required. For example, other yeast SH3 domains that recognize class I motifs (Table 1.2) can be mutated to contain Tyr30/Val38 residues to test whether this increases their binding to Ste20 peptide and other peptides bearing the ΨxPxRxAxxP consensus. The ultimate goal would be to take a yeast SH3 domain and completely switch its specificity towards the NbpSH3 domain. It would be interesting to see how many amino acid changes would be required to accomplish this (i.e. would Tyr30 and Val38 substitutions be sufficient for some yeast SH3 domains?). This in turn will give an idea of the number of residues involved in controlling the specificity of the NbpSH3 domain. The NbpSH3 domain is a simple domain (unlike BemSH3b it does not have any extra sequences) that recognizes a canonical class I motif. Therefore, the results obtained with the NbpSH3 domain would be applicable to other SH3 domains and would provide further insight into the evolution of SH3 domain specificity.
Chapter 6: References


Chan, B., Lanyi, A., Song, H.K., Griesbach, J., Simarro-Grande, M., Poy, F., Howie, D.,

lessons learned from the yeast Saccharomyces cerevisiae. Biochim Biophys Acta 1773, 1311-
1340.

Bioinformatics 20, 426-427.

proteins and phosphorylations regulate Saccharomyces cerevisiae Cdc24p localization. FEBS
Lett 583, 3339-3343.

Cornilescu, G., Delaglio, F., and Bax, A. (1999). Protein backbone angle restraints from
searching a database for chemical shift and sequence homology. J Biomol NMR 13, 289-302.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh,
425-431.

phosphotyrosine motifs does not change the affinity of its SH3 domains for Sos proline-rich
motifs. EMBO J 13, 4011-4021.

protein kinases are required for normal localization of cell growth and for cytokinesis in
budding yeast. Genes Dev 9, 1817-1830.

Multidimensional Spectral Processing System Based on Unix Pipes. Journal of Biomolecular
Nmr 6, 277-293.

Diella, F., Haslam, N., Chica, C., Budd, A., Michael, S., Brown, N.P., Trave, G., and Gibson,
T.J. (2008). Understanding eukaryotic linear motifs and their role in cell signaling and
regulation. Front Biosci 13, 6580-6603.

Douangamath, A., Filipp, F.V., Klein, A.T., Barnett, P., Zou, P., Voorn-Brouwer, T., Vega,
M.C., Mayans, O.M., Sattler, M., Distel, B., et al. (2002). Topography for independent
binding of alpha-helical and PPII-helical ligands to a peroxisomal SH3 domain. Mol Cell 10,
1007-1017.

inheritance via Slt2p. EMBO J 25, 4413-4422.


Parrini, M.C., Lei, M., Harrison, S.C., and Mayer, B.J. (2002). Pak1 kinase homodimers are autoinhibited in trans and dissociated upon activation by Cdc42 and Rac1. Mol Cell 9, 73-83.


