Residues Neighboring an SH3-Binding Motif Participate in the Interaction *In Vivo*

David F. Jordan[1,2,3,4,*], Alexandre K. Dubé[1,2,3,4,5], Ugo Dionne[2,3,4,6,7,8], David Bradley[1,2,3,4,5] and Christian R. Landry [1,2,3,4,5,*]

1. Département de Biochimie, microbiologie et bio-informatique, Université Laval, 1045 Avenue de la Médecine, Québec, QC, Canada, G1V 0A6
2. Institut de Biologie Intégrative et des Systèmes (IBIS), 1030, Avenue de la Médecine, Université Laval, Québec, QC, Canada, G1V 0A6
3. PROTEO-Regroupement Québécois de Recherche sur la Fonction, l'Ingénierie et les Applications des Protéines, Université Laval, 1030, Avenue de la Médecine Québec, QC, Canada, G1V 0A6
4. Centre de Recherche en Données Massives de l'Université Laval, Université Laval, 1065, Avenue de la Médecine, Québec, QC, Canada, G1V 0A6
5. Département de Biologie, Université Laval, 1045 Avenue de la Médecine, Québec, QC, Canada, G1V 0A6
6. Centre de Recherche du Centre Hospitalier Universitaire (CHU) de Québec-Université Laval, 11, côte du Palais, Québec, QC, Canada, G1R 2J6
7. Centre de Recherche sur le Cancer de l'Université Laval, 9, rue McMahon, Québec, QC, Canada, G1R 3S3
8. Present address: Lunenfeld-Tanenbaum Research Institute, Sinai Health, 600 University Avenue, Toronto, ON, Canada, M5G 1X5

* Correspondance: 1030, Avenue de la Médecine, Université Laval, Québec, QC, Canada, G1V 0A6 Email: david.jordan.2@ulaval.ca. 1030, Avenue de la Médecine, Université Laval, Québec, QC, Canada, G1V 0A6 Email: christian.landry@bio.ulaval.ca

Running head: Residues neighboring motif participate in interaction

Keywords: Short Linear motif, Src Homology 3 (SH3) domain, Deep mutational scanning, Binding affinity, Structure prediction

Abstract

In signaling networks, protein-protein interactions are often mediated by modular domains that bind short linear motifs. The motifs' sequences affect many factors, among them affinity and specificity, or the ability to bind strongly and to bind the appropriate partners. Using Deep Mutational Scanning to create a mutant library, and protein complementation assays to measure protein-protein interactions, we determined the *in vivo* binding strength of a library of mutants of a binding motif on the MAP kinase kinase Pbs2, which binds the SH3 domain of the osmosensor protein Sho1 in *Saccharomyces cerevisiae*. These measurements were made using the full-length endogenous proteins, in their native cellular environment. We find that along with residues within the canonical motif, many mutations in the residues neighboring the motif also modulate binding strength. Interestingly, all Pbs2 mutations which increase affinity are situated outside of the Pbs2 region that interacts with the canonical SH3 binding pocket, suggesting that other surfaces on Sho1 contribute to binding. We use predicted structures to hypothesize a model of binding which involves residues neighboring the canonical Pbs2 motif binding outside of the canonical SH3 binding pocket. We compared this predicted structure with known structures of SH3 domains binding peptides through residues outside of the motif, and put forth possible mechanisms through which Pbs2 can bind specifically to Sho1. We propose that for certain SH3 domain-motif pairs, affinity and specificity are determined by a broader range of sequences than what has previously been considered, potentially allowing easier differentiation between otherwise similar partners.

Summary

Protein-protein interactions are often mediated by a binding domain on one protein and a short disordered binding motif on another protein. We measured the binding strength of a mutant library of a binding motif situated in the yeast protein Pbs2 to the SH3 domain of Sho1. Many mutations in the residues neighboring the binding motif affect binding. A protein structure prediction of the interaction partners shows that residues neighboring the motif bind residues outside the known binding pocket on the SH3 domain. The Sho1-Pbs2 interaction differs enough from other known SH3-motif pairs to allow specific binding.

Introduction

Cells possess complex and robust signaling networks that detect stimuli and trigger responses. These signaling networks are composed of a series of protein-protein interactions, which are often mediated by modular interaction domains. Many classes of protein interaction domains are shared across pathways and species, yet fulfill different roles and functions (Pawson *et al.* 2002). The different constraints placed on interaction domains by their distinct roles contribute to explaining the divergence in sequence between homologs of the same domain (Ernst *et al.* 2010; Dionne *et al.* 2022). Therefore, to understand how domain sequences have evolved and continue to evolve, the phenotypic consequences of mutations must be understood. Two of the major phenotypes for protein-protein interaction domains are affinity, that is, the strength of binding to the partner protein, and specificity, that is, the ability to bind to the appropriate partner proteins, and to not form spurious interactions with other proteins in the cellular environment (Ivarsson and Jemth 2019).

Many interaction domains bind short intrinsically disordered stretches of their interaction partners, also known as short linear motifs (SLiMs) or simply binding motifs (Gouw *et al.* 2018). Binding motifs are involved in a wide swath of cell signaling pathways, and underlie many important mechanisms in human and other cells (Kumar *et al.* 2024). Uncovering the determinants of binding can elucidate the functioning of human cells and certain diseases that affect them, such as various cancers in which motif-mediated signaling plays a part (Van Roey *et al.* 2014). Additionally, viruses also use their own binding motifs to disrupt the signaling machinery of their hosts (Davey *et al.* 2011). Understanding domain-motif binding can also allow them to be used as tools, for example by attaching therapeutic proteins to domains and motifs to hydrogels, to ensure the gradual release of the therapeutic proteins in human tissues (Delplace *et al.* 2019).

Interactions mediated by domain-motif associations are relatively weak, with dissociation constants (K_D) in the micromolar range, while domain-domain binding typically results in dissociation constants values orders of magnitude smaller, indicating much stronger binding (Van Roey *et al.* 2014). While such low affinity could simply be caused by physical limitations, such as the smaller interface, adaptive hypotheses have been put forward which suggest that an overly strong affinity can have deleterious effects on the cell, for instance by compromising specificity (Haslam and Shields 2012; Karlsson *et al.* 2016). Alternatively, this low intrinsic affinity can be compensated by other factors. For example, certain protein-protein interactions depend on the simultaneous binding of many domain-motif pairs (Liao *et al.* 2020). In other cases, motifs bind outside of the canonical binding pockets of their partner domains, although this does not always lead to increased affinity (Douangamath *et al.* 2002). The regions surrounding a domain can also have an effect on motif binding preference, as inserting domains into new protein backgrounds can change the motifs that this domain binds to (Dionne *et al.* 2021, 2022).

Binding affinity and specificity have previously been explored in the context of the interaction between binding domains and motifs, by measuring the binding of mutants *in vitro* (Yu *et al.* 1994; Zarrinpar *et al.* 2003; Tonikian *et al.* 2009; Vincentelli *et al.* 2015; Kazlauskas *et al.* 2016). However, the interaction of proteins in the cellular environment is a more complex situation, and

many interactions detected *in vivo* are not detected *in vitro,* and vice versa (Kelil *et al.* 2017). Many factors can modulate binding in the cell, including colocalization of partners, expression in the same cell cycle phases, and contributions from sequences outside of the immediate binding domain and motif, as well as other proteins that can interact with one or both of the partners (Ivarsson and Jemth 2019; Dionne *et al.* 2022). Earlier studies have measured the effects of limited numbers of domain and motif mutants *in vivo* (Zarrinpar *et al.* 2003; Marles *et al.* 2004). More recently, techniques such as deep mutational scanning (DMS) have been used to study the impact of large libraries of mutations on protein stability and function *in vivo* (Fowler and Fields 2014). Methods such as protein complementation assays (Tarassov *et al.* 2008; Michnick *et al.* 2016) have also successfully been used to measure the *in vivo* binding strength of libraries of mutants in yeast (Diss and Lehner 2018; Dionne *et al.* 2021; Faure *et al.* 2022; Robles *et al.* 2023; Dibyachintan *et al.* 2024; Bendel *et al.* 2024). For example, a recent study showed the effect on binding of combining mutations in a PDZ domain and its binding motif (Zarin and Lehner 2024). These *in vivo* techniques could be used to determine how a large library of mutations can affect domain-motif interaction affinity.

One powerful model to study domain-motif interactions is SH3 domains. These globular domains are composed of around 60 residues, and bind different classes of polyproline motifs defined by the canonical forms R/KXXPXXP for class I and PXXPXR/K for class II (Kaneko *et al.* 2008). The model yeast *Saccharomyces cerevisiae*'s proteome contains 27 different SH3 domains (Dionne *et al.* 2022), including one in the High Osmolarity Glycerol (HOG) signaling pathway protein Sho1 (Saito and Posas 2012). The SH3 domain of Sho1 is known to interact with a motif on the MAPKK Pbs2, and this interaction can be strengthened by exposing cells to osmotic stress (Maeda *et al.* 1995; Posas and Saito 1997; Saito and Posas 2012) (Figure 1a). The Sho1-Pbs2 interaction is thus a powerful model to study domain-motif interaction strength, due to the potential for increasing or decreasing affinity through mutations and the simple method of inducing an increase in the interaction strength using osmotic stress.

To better understand the impact of mutations on affinity, we measured the binding strength of a nearly complete library of mutations in the Pbs2 binding motif *in vivo* using a DHFR Protein-fragment Complementation Assay (PCA) (Tarassov *et al.* 2008; Michnick *et al.* 2016). We measured the binding strength between single residue Pbs2 motif mutants and Sho1, the canonical partner of Pbs2. We find that all single mutations that increase the interaction strength are found outside the canonical binding motif, and that certain residues outside the binding motif are predicted to interact with Sho1 outside of the canonical binding pocket, and even outside of the SH3 domain. We use this finding to propose a model of Sho1-Pbs2 binding, where residues outside the canonical Pbs2 motif bind a second binding pocket on Sho1. We also compare this proposed model to other known SH3-motif interactions involving binding outside the canonical motif.

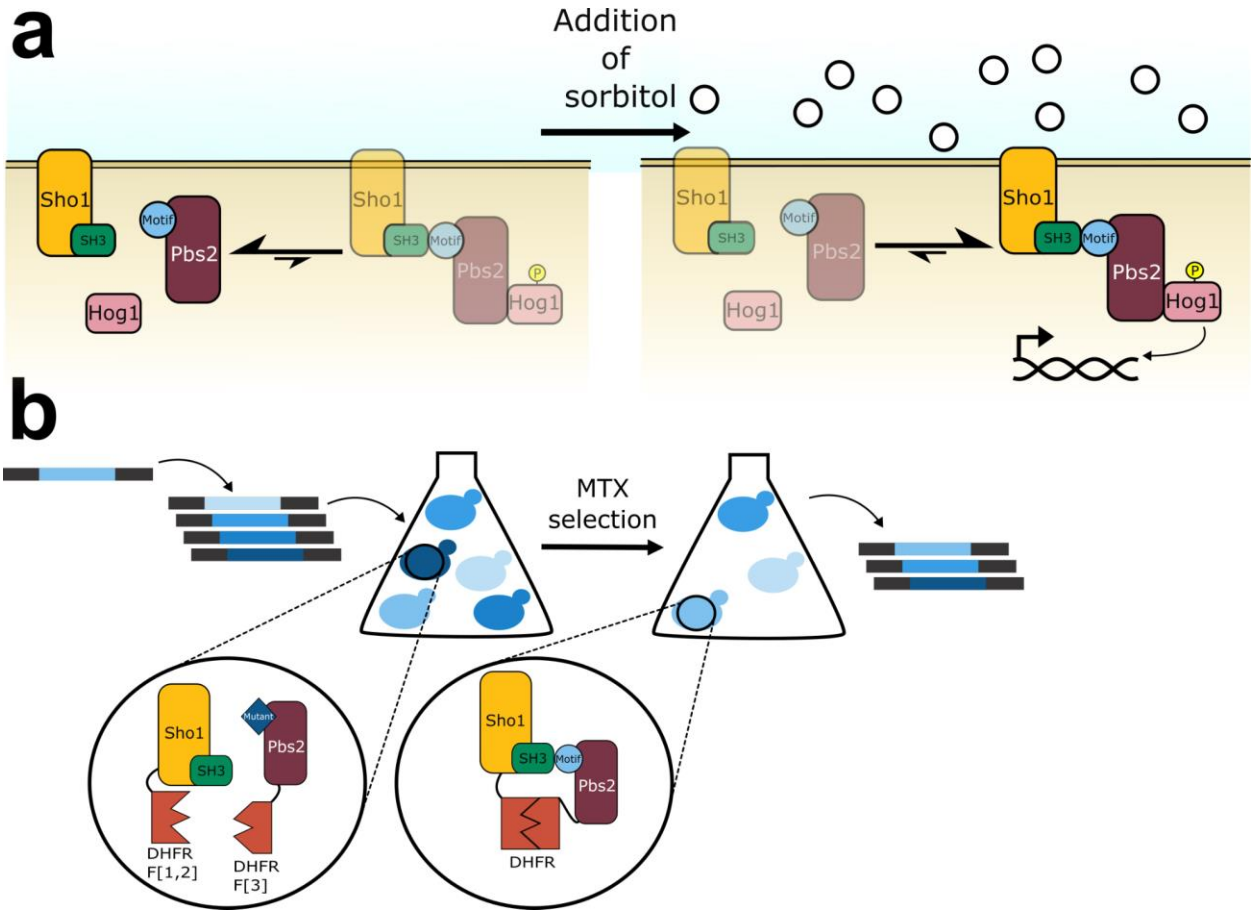


Figure 1. <u>The Sho1-Pbs2 interaction can be strengthened using sorbitol exposure, and the DHFR-PCA can be used to measure the interaction strength between Sho1 and Pbs2.</u>
a) Simplified schematic of the HOG pathway, and the interactions between Sho1, Pbs2 and Hog1. In normal conditions (left), Sho1, Pbs2 and Hog1 are mostly separated. In the presence of osmotic stress such as sorbitol (right), signaling through the pathway is induced, the proteins assemble into a complex, and Hog1 is phosphorylated, which causes it to modulate gene expression via translocation to the nucleus. b) Schematic of the DHFR-PCA screen with a DMS mutant library. Single residue variants of *PBS2* were inserted into the yeast genome at the native locus using CRISPR-Cas9 genome editing. Sho1 and Pbs2 were tagged with complementary DHFR fragments. In cases where the Pbs2 variant does not interact with Sho1, the DHFR fragments remain separated. Upon interaction between a Pbs2 variant and Sho1, the DHFR fragments combine to reconstitute the functional DHFR, and allow growth in the presence of methotrexate (MTX). More strongly interacting partners allow the cell to proliferate more rapidly (Freschi *et al.* 2013).

Results

Scan of the region surrounding the Pbs2 motif reveals that few mutations modify binding to Sho1

The interaction between Sho1 and Pbs2 is modulated by an SH3 domain on Sho1 and the canonical binding motif on Pbs2, which is a type I motif with the sequence KPLPPLP (Maeda *et al.* 1995). Previous computational work has suggested that the region surrounding binding motifs could play an important role in modulating binding, including in SH3 interactions and in particular in the Sho1-Pbs2 interaction (Stein and Aloy 2008; Kelil *et al.* 2016). To determine which Pbs2 residues play a role in binding, we undertook a Deep Mutational Scan (DMS) on a region of 56 codons comprising the binding motif and its surrounding region. A DNA library of every possible single codon mutant was created for the 56 codon stretch, and inserted into the *S. cerevisiae* genome at the PBS2 locus using a CRISPR Cas-9 based approach, replacing the wild-type sequence (Dionne *et al.* 2021). The *PBS2* gene was fused with a Dihydrofolate Reductase Protein-Fragment 3 (DHFR F[3]), and these cells were mated with cells containing a DHFR F[1,2] fusion with *SHO1*, in order to obtain diploid strains containing complementary DHFR F fusions on both interaction partners. An interaction between Sho1 and Pbs2 brings the two DHFR fragments into contact, forming a functional DHFR which allows cell division even in the presence of the drug methotrexate (MTX) (Figure 1b). Growth in the presence of MTX in a strain with both DHFR fragments is therefore a proxy for the amount of Sho1-Pbs2 complexes forming, which depends on both the binding affinity and the local abundance of the interacting proteins (Tarassov *et al.* 2008; Freschi *et al.* 2013). As the DHFR fusions are located at the C-terminus of proteins, nonsense mutations introduced in the mutant library lead to no DHFR fragment expression.

To measure the interaction strength of the DMS library, the mutants were pooled and DHFR-PCA competition assays were done in media containing both MTX and 1 M of sorbitol, in order to induce the HOG pathway (Ferrigno *et al.* 1998). The resulting interaction score was calculated as the normalization of log-2-fold change of variant frequencies before and after selection for the interaction using MTX, as detected by targeted sequencing of the *PBS2* locus. Interaction scores were normalized in each replicate, so that scores of 0 are equivalent to wild-type Pbs2, while interaction scores of -1 correspond to the median of all nonsense mutants. Mutants which possess a higher interaction strength than wild-type Pbs2 therefore have a positive interaction score, while mutants which possess a lower interaction score than wild-type Pbs2 have a negative interaction score.

Many mutations in the canonical motif change the interaction strength. However, we found that most mutations in the region surrounding the motif did not affect binding of Sho1, apart from a short section neighboring the canonical motif itself (Figure 2a, Figure S1, Figure S2, Table S1). This confirms the strong role of the motif in the interaction and validates the computational predictions that extended the role of binding to include neighboring residues (Stein and Aloy 2008; Kelil *et al.* 2016). Since the effect on binding was strongly position dependent, we categorized different sections of Pbs2 as follows: the canonical binding motif is the conserved type I binding motif sequence situated in positions 93 to 99, the extended motif is composed of the canonical motif and the neighboring residues which have a strong impact on binding from positions 85 to 99 (Figure 2b).

DHFR-PCA reports on the amount of protein complex formed, which depends on affinity and local protein abundance. To differentiate between these two parameters, we measured the effect of mutations on a Pbs2 interaction to a partner that does not depend on the SH3 binding motif. Hog1 binds Pbs2 distally to the SH3 binding motif and independently of Sho1 binding (Murakami *et al.* 2008) (Figure 1a). So, measuring Hog1-Pbs2 binding allows us to control for Pbs2 local abundance. Pbs2 mutations that only affect Sho1 binding should not have a detectable effect in the Hog1-Pbs2 DHFR-PCA, while Pbs2 mutations that affect Pbs2 stability or local abundance will have an effect in the Sho1-Pbs2 as well as the Hog1-Pbs2 DHFR-PCA.

We found that most mutations in the surrounding region of the Pbs2 motif had little effect on Hog1 binding (Figure 2a). Furthermore, mutants that had a negative effect on Hog1-Pbs2 binding also had a negative effect on Sho1-Pbs2 binding, suggesting that the reduction in signal results from a loss of local abundance of Pbs2 (Figure S3). However, most mutations that affected Sho1 binding did not impact the Hog1 interaction, and therefore did not change the local abundance of Pbs2. Strikingly, only mutations in the residues comprising the extended motif had Sho1-interaction specific effects, while mutations outside the extended motif had either little effect or destabilizing effects.

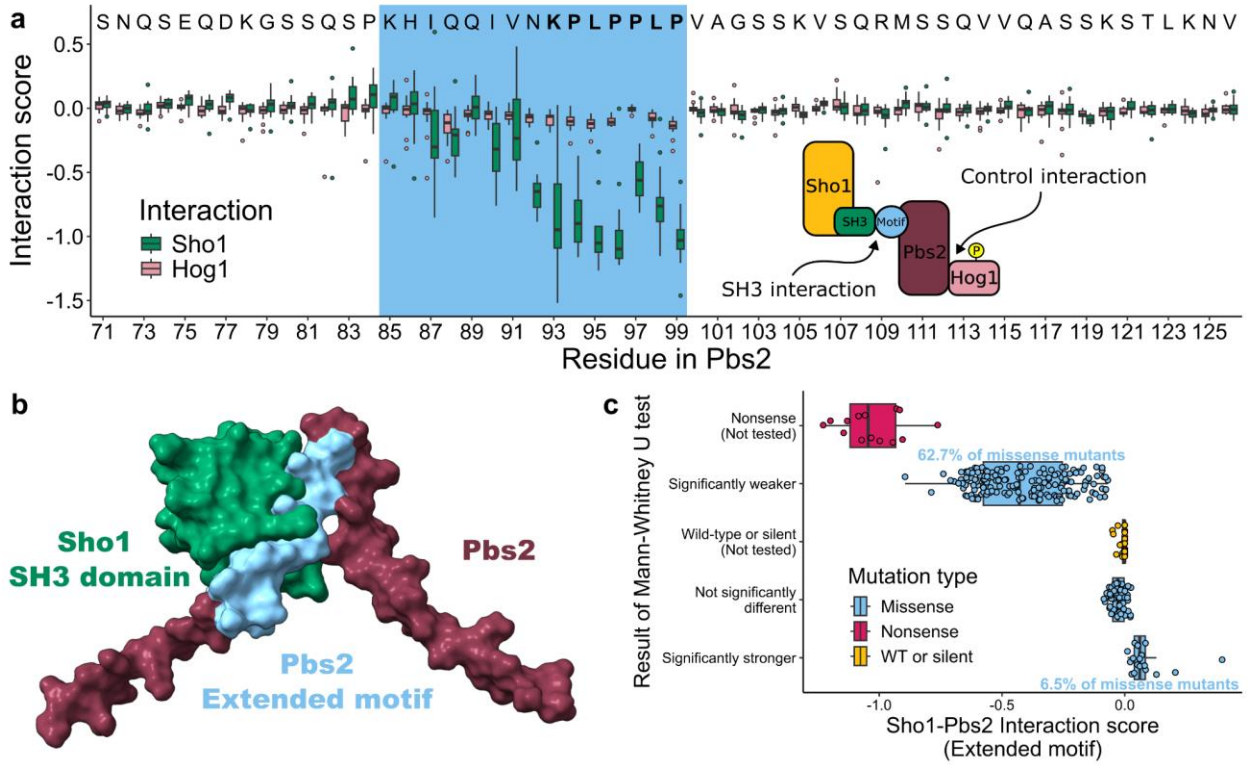


Figure 2. Mutations in Pbs2 within and around the binding motif affect binding to Sho1.
a) Effect of Pbs2 mutations in the surrounding region of the SH3 binding motif (interaction score, normalized log2-fold change before and after selection using MTX), with either Sho1 or Hog1.

The wild-type residues of Pbs2 are indicated for each position at the top of the figure. The positions comprising the extended motif are highlighted in blue, while the positions in the canonical binding motif have the wild-type residue indicated in bold. Boxplots show the distribution effects for each residue mutated to, and each residue is measured in between 3 and 18 replicates. Nonsense mutants omitted. b) A predicted structure (AlphaFold2-Multimer) of the SH3 domain of Sho1 (green) in complex with the Pbs2 region surrounding the motif (burgundy). The extended motif in positions 85 to 99 is colored in light blue. c) Distribution of interaction scores between Sho1 and Pbs2 single mutants in the extended motif DMS library (positions 85 to 100). Mutations are colored and placed on the y-axis by the nature of the mutation, either a change to a stop codon (nonsense), a change to the wild-type codon or a synonymous codon (WT or silent) or a change to a codon coding for a different residue (missense). The missense mutants were compared to the WT and silent mutations, using a Mann-Whitney U test with false discovery rate correction. They were grouped into 3 categories: Significantly weaker (p-value < 0.05 and negative interaction score), Not significantly different (p-value > 0.05) or Significantly stronger (p-value < 0.05 and positive interaction score).

<u>Extended motif DMS library identifies mutations modulating interaction strength</u>

In light of these results, we decided to focus solely on the extended motif of Pbs2, comprising codons 85 to 99, and adding position 100 to verify no effects were taking place in the residue immediately following the canonical motif. The DHFR-PCA on the initial DMS library of the entire region surrounding the motif took place in a diploid strain. Consequently, a large part of Pbs2 binding occurred with the Sho1 copy which did not possess the DHFR F[1,2] fusion, and these interactions did not contribute to the DHFR-PCA signal. Since this limited our ability to precisely measure the Sho1-Pbs2 interaction, we built a new haploid strain containing both DHFR F[1,2] on Sho1 and DHFR F[3] on Pbs2. We also constructed a Hog1-DHFR F[1,2]/Pbs2-DHFR F[3] strain to once more measure the local abundance of Pbs2 mutants. These new strains capture all Pbs2 binding to Sho1 or Hog1, therefore allowing for more sensitive measurements. We inserted a second Pbs2 DMS library, comprising all possible single mutations in Pbs2 positions 85 to 100, into the two haploid DHFR strains, using the same CRISPR-Cas9 approach as the first library. The interaction strength of the variants was measured using the DHFR-PCA as previously (Figure S4, Figure S5). At this point, we filtered out the 12 mutants which had an interaction score with Hog1 significantly different from wild type and synonymous Pbs2 mutants (Two-sided Mann-Whitney U test with false discovery rate corrected p-value < 0.05) from our analysis (Figure S6). As before, the goal was to disregard mutants which modulated the interaction through change of local abundance or stability, which would affect both the Sho1 and Hog1 interaction.

The impacts on the Sho1-Pbs2 interaction were varied (Figure 2c, Figure S7, Table S2). Few mutants had interaction scores as low as the nonsense mutants, suggesting that few single mutations are able to completely prevent Sho1-Pbs2 binding. By comparing missense mutant interaction scores to the wild type and synonymous mutants scores, we found that many missense mutants had significantly stronger or weaker scores than the wild-type Pbs2 sequence (Two sided Mann-Whitney U test with false discovery rate corrected p-value < 0.05). 18 mutants (6.5% of all missense mutants) interacted significantly more strongly with Sho1 than wild-type Pbs2, while

173 mutants (62.7% of all missense mutants) interacted significantly more weakly than wild-type Pbs2. The impact of mutants that reduced the interaction scores was generally greater.

As the Sho1-Pbs2 interaction is strengthened under osmotic stress (Maeda *et al.* 1995; Posas and Saito 1997; Saito and Posas 2012), we also compared the interaction strength between these two proteins in the presence and absence of 1 M of sorbitol (Figure S8). We found that interaction scores in the presence and absence of sorbitol correlate strongly (Spearman's rho 0.98, p-value $< 2.2X10^{-16}$). The effects on binding were stronger in the presence of sorbitol, as a result of the induction of the HOG signaling pathway. When sorbitol is present, strongly interacting Pbs2 mutants are even more often in contact with Sho1, increasing the number of reconstituted DHFR complexes, while weakly interacting mutants are even further outcompeted by the strongly interacting mutants. Still, the high correlation in interaction scores between these two conditions indicates that Pbs2 mutants behave similarly in the presence and absence of osmotic stress, and that the general strengthening or weakening effect of mutations does not depend on the activation of the signaling pathway.

To validate the pooled measurements, we individually reconstructed 24 mutants spanning the range of measured interaction scores in Pbs2-DHFR F[3] strains which were mated with the Sho1-DHFR F[1,2] strain. We then measured the growth rate of these mated strains in growth assays in DHFR-PCA medium. As with the competition assay, the growth rate is a proxy for interaction strength, which reflects affinity in the absence of a change in abundance. The growth curve results correlate strongly with the interaction scores (Spearman's rho 0.88, p-value $7.3x10^{-5}$ for interactions in presence of 1 M sorbitol, spearman's rho of 0.73, p-value $2.9x10^{-6}$ for interactions in absence of sorbitol), thus upholding the results of the competition assay (Figure S9a). To validate further that the effects on interaction strength were not the result of noise or the result of another source, we also performed a smaller scale competition assay DHFR-PCA, individually reconstructing 67 mutants with varying interaction scores (Figure S10, Table S3). The measurements of this validation assay correlated strongly with the assay on the DMS library of the extended motif (Spearman's rho 0.98, $p < 2.2X10^{-16}$ for interactions in presence of 1 M sorbitol, spearman's rho of 0.98, $p < 2.2X10^{-16}$ for interactions in absence of sorbitol), thus validating the results of the measurement of the DMS library (Figure S9b). We also measured the effect of mutations on cell proliferation by constructing strains with the same 67 mutations, but with no DHFR tags. Cell growth and division should thus only depend on the effect of the mutations, and not interaction strength. We undertook a pooled competitive growth assay in SC synthetic media with 1 M of sorbitol, measuring how much each mutant proliferated. We found that all mutations except nonsense mutations had no significant effect on cell proliferation (Welch's t-test false discovery rate corrected p-value < 0.05) (Figure S11, Table S3). Interestingly, mutant I87W, which is the strongest interacting mutant, seems to have a detectable, though not statistically significant, deleterious effect on cell proliferation.

Structure prediction suggests a secondary contact between extended motif and Sho1

A crystal structure of the SH3 domain of Sho1 in complex with a 9 residue-long segment of Pbs2 reveals that the canonical motif occupies the entirety of the canonical binding pocket of Sho1 (Kursula *et al.* 2008). However, many mutations outside the canonical motif have a strong effect on binding. These residues must affect the interaction of Pbs2 with Sho1, but it is unclear how. One possibility is that certain residues in the extended motif are interacting with Sho1 outside of the canonical binding pocket of the SH3 domain. This would not be unprecedented, as NMR structures of different SH3 domains interacting with long peptides reveal interactions involving residues outside of the canonical binding pocket, which affected the strength and specificity of the interaction (Rickles *et al.* 1995; Feng *et al.* 1995; Stollar *et al.* 2009; Takaku *et al.* 2010; Gorelik and Davidson 2012). No experimental structure has been determined for the entirety of Sho1 or Pbs2, so we used structure predictions from AlphaFold-Multimer (Jumper *et al.* 2021; Evans *et al.* 2022) to determine if binding outside the canonical binding pocket was likely to occur (Figure S12). By modeling the entire Sho1 protein along with the surrounding region of the Pbs2 motif (positions 71-126), we found potential interactions outside of the canonical binding pocket of Sho1, mediated by residues neighboring the canonical binding motif of Pbs2 (Figure 3a). In the structural model, a hydrophobic pocket is formed between the SH3 domain and a loop of Sho1, which interacts with an alpha helix composed of Pbs2 positions 84 to 92 (Figure 3b). Interestingly, all but one of the mutants in the extended motif which significantly increase interaction strength are found in the positions comprising the helix. Certain Pbs2 residues towards the C-terminus relative to the canonical motif are also predicted to be in proximity to Sho1, with 2 residues within 2 Å of Sho1, but as mutations in these positions did not have substantial effects on binding (Figure 2a), we decided to concentrate on residues in and near the extended motif.

To help explain how different mutations could modify binding, we predicted the structures of the 5 Pbs2 variants possessing the strongest interaction strength with Sho1, using the same method as with wild-type Pbs2 (Figure S12). In particular, I87W, which is the strongest interacting mutant measured, positions its side-chain in close contact (1.77 Å) with the non-canonical hydrophobic pocket, and particularly in a sub-pocket formed of both SH3 and non-SH3 residues. Additionally, it is predicted to form a hydrogen bond with the aspartic acid side chain in position 333 of the Sho1 sequence (Figure 3c). These two factors potentially explain its strong interaction strength relative to all other Pbs2 variants. Of the other strengthening variants tested, H86W, V91L and V91M all place a hydrophobic side chain in the non-canonical hydrophobic pocket, potentially strengthening the interaction. For the last mutant tested, Q89D, the hydrophilic side-chain faces away from the hydrophobic pocket, thus not disrupting the hydrophobic interaction, although it is unclear how the interaction is strengthened (Figure S13).

We further reasoned that substituting a non-hydrophobic residue in one of the Pbs2 positions which is predicted to contact the non-canonical hydrophobic pocket could interrupt this interaction, and weaken the affinity. Indeed, when classifying the mutations in the alpha helix by their hydrophobicity (Monera *et al.* 1995), we find that replacing the initially hydrophobic residues in positions I87, I90, and V91 with a non-hydrophobic residue reduces the binding strength. In

general, a mutation to another hydrophobic residue in these positions has a weaker negative impact on binding, and in the case of V91, many hydrophobic substitutions increase the interaction strength (Figure 3d). Based on the results of the hydrophobic or non-hydrophobic substitutions, we reasoned that these three positions situate their side chains in the non-canonical hydrophobic binding pocket. This finding helps explain the strong interactions strength of Pbs2 mutant I87W, by showing the important role of hydrophobic interactions in position 87. The predicted models of interactions between Sho1 and different Pbs2 variants, along with our interaction score measurements suggest that residues outside the canonical binding motif of Pbs2 can interact directly with Sho1, and thus they have the capacity to modulate affinity.

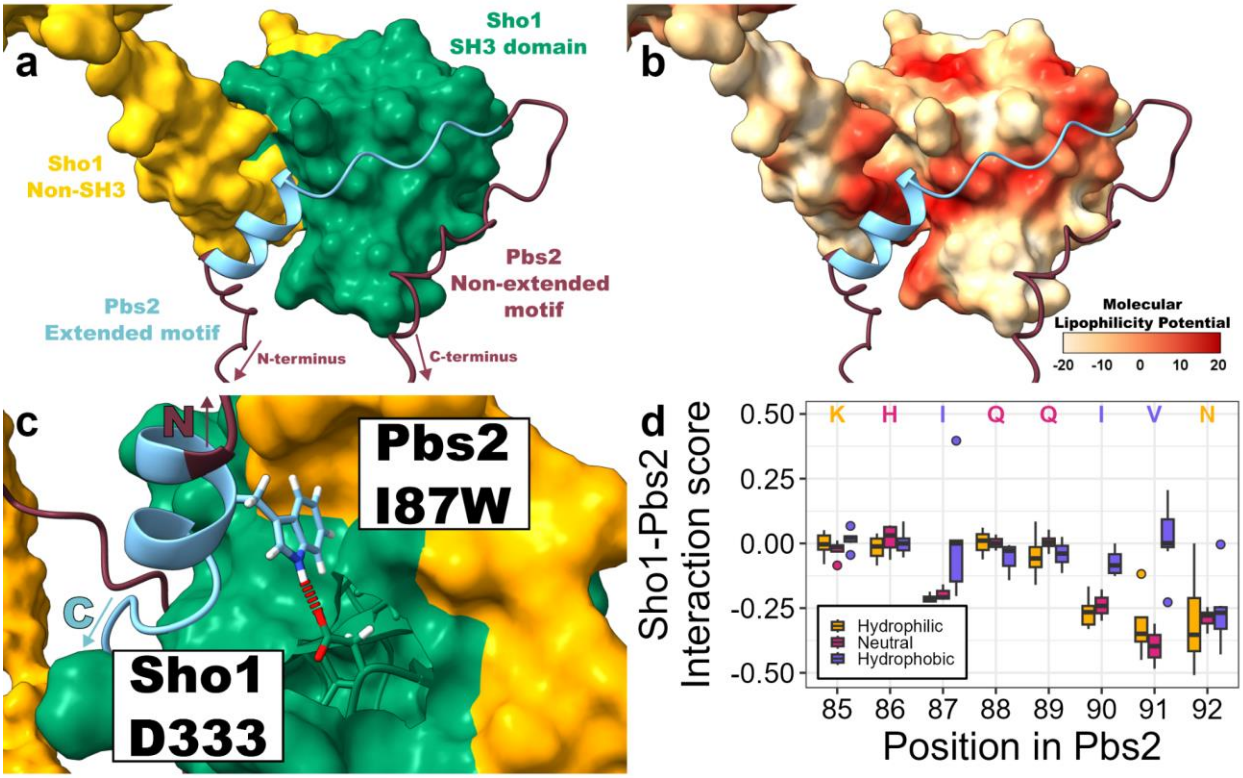


Figure 3. <u>Contacts mediated by positions outside the canonical binding motif of Pbs2 could modulate binding</u>

a) Structure of Sho1 in complex with residues 71 to 126 of Pbs2 as predicted by AlphaFold2-Multimer (Jumper *et al.* 2021; Evans *et al.* 2022). The structure is positioned to show the helix on Pbs2 interacting with a hydrophobic pocket on Sho1, while the canonical motif interacts with the canonical binding pocket. The structure is colored by the different regions of Sho1 and Pbs2, as indicated on the panel. The orientation of Pbs2 is shown by annotations indicating which end is nearer the N-terminus and which end is nearer the C-terminus. b) Hydrophobicity map of the predicted structure of Sho1, with high molecular lipophilicity potential (MLP) indicating a hydrophobic surface, and low MLP indicating a hydrophilic surface. MLP calculated using

Chimera X 1.8. Positioning of both proteins, and coloring of Pbs2 as in panel a). c) Close-up of the predicted structure of Sho1 with residues 71 to 126 of Pbs2 with mutation I87W. The side-chain of W87 is shown, and a hydrogen bond (red dotted line) is predicted with D333 of Sho1. Coloring of the proteins is as described in panel a). The orientation of Pbs2 is shown by annotations indicating which end is nearer the N-terminus (N) and which end is nearer the C-terminus (C). d) Boxplot of interaction scores of Pbs2 mutants for selected positions from the DHFR-PCA of the DMS library of the extended motif, based on the hydrophobicity of the substituted residue. The wild-type residue of Pbs2 for every position is indicated and colored according to its hydrophobicity. Nonsense mutants omitted. Hydrophobicity for the different amino acids was obtained from (Monera *et al.* 1995), and F, I, W, L, V, M, Y, C, and A were classified as hydrophobic, T, H, E, S and Q were classified as neutral and R, K, N, G, P and D were classified as hydrophilic.

Comparison of Sho1-Pbs2 binding to other SH3-motif pairs

As we propose a role in modulating interaction strength for the extended motif and the additional binding pocket of Sho1, we wondered if these structures could play a part in modulating the binding specificity of Pbs2 to Sho1. If this additional pocket were a unique structure among yeast SH3 domains, that could increase the specificity of binding, as the Pbs2 extended motif could be uniquely tuned to bind this unique feature. First, we generated a multiple sequence alignment of yeast SH3 domain sequences with 25 adjacent residues on each side, to verify the conservation of the amino acids forming the additional binding pocket. This window includes all residues predicted to be in proximity to Pbs2 in Sho1. Apart from a few key residues in the SH3 domain, we found no significant conservation of the positions which in Sho1 are predicted to be in proximity (< 5 Å) to Pbs2 positions situated in the extended motif, but outside of the canonical motif (Figure S14a). We reasoned that the structure of the additional binding pocket might be conserved, even if the particular residues were not, so we used the MUSTANG (Konagurthu *et al.* 2006) tool to carry out a multiple structural alignment on structure predictions of yeast SH3 domains with 25 neighboring residues, obtained from the AlphaFold Protein Structure Database (Varadi *et al.* 2022). We found that only the positions in the SH3 domains were structurally conserved, with no significant structural conservation for the positions outside the SH3 domains (Figure S14b), including those forming the additional binding pocket. We therefore conclude that the additional binding pocket on Sho1 is a unique structural feature.

Even without a conserved structure of non-SH3 sequences, there are reports of three other yeast SH3 domains binding extended motifs: Abp1, the second SH3 domain of Bem1 (referred to as Bem1-2) and Nbp2 (Stollar *et al.* 2009; Takaku *et al.* 2010; Gorelik and Davidson 2012). We tested whether there was an *in vivo* interaction between these SH3 domains tuned to bind extended motifs, and the extended motif of Pbs2. To do so, we once again used a DHFR-PCA assay, though this time measuring colony growth on solid medium as is routinely done for this (Tarassov *et al.* 2008), with the SH3-containing proteins tagged with DHFR F[1,2] and Pbs2 tagged with DHFR F[3]. We used two variants of Pbs2: a wild-type, and a control where the

extended motif and the region surrounding it (positions 71 to 126, the same region covered in the initial DMS library) were replaced by a neutral, flexible stuffer sequence.

We found that Abp1 and Nbp2 interact with Pbs2 with approximately the same strength as Sho1 does, while Bem1-2 does not detectably interact with Pbs2 (Figure 4a). However, Abp1 and Nbp2 also interact with the stuffed Pbs2 variant at the same level as the wild-type variant, indicating that the interaction is not mediated by the Sho1-binding motif. Nbp2 is known to bind a second, distal motif on Pbs2 through its SH3 domain (Mapes and Ota 2004), while Abp1 must bind Pbs2 through an unknown mechanism involving a different site than the one bound by Sho1. We additionally measured the interaction with all other known yeast SH3-containing proteins, and found that apart from Abp1 and Nbp2, all proteins interacted significantly less strongly with either Pbs2 variant than Sho1 with wild-type Pbs2 (One sided Mann-Whitney U test with false discovery rate corrected p-value < 0.05) (Figure S15). Additionally, measuring the interaction of Pbs2 with a Sho1 variant where the SH3 domain was replaced by the neutral, flexible stuffer sequence showed an interaction strength above the level of stuffed Pbs2 with the same Sho1 variant, confirming the role of the non-SH3 portion of Sho1 in binding the extended motif of Pbs2 (Figure S15, Sho1_stuffed). Together, these findings show that the extended motif of Pbs2 specifically binds to Sho1 both inside and outside the SH3 domain, and doesn't interact with other SH3-containing proteins *in vivo*.

To determine whether we could explain the specificity of Pbs2 binding, we compared the predicted structure of the Sho1-Pbs2 interaction to NMR structures of the Abp1-Ark1, Bem1-2-Ste20, and Nbp2-Ste20 interactions, all of which include long peptides with lengths comparable to the extended motif of Pbs2. The structures of the four SH3 domains and peptides are very similar, with all peptides binding in the canonical binding pocket of their respective SH3 domains, and extending towards the space corresponding to what we identified as the additional binding pocket on Sho1 (Figure 4b, Figure S16). However, Abp1 and Nbp2 only bind the extended motif through the SH3 surface, while Bem1-2 has a second, non-SH3 interface that forms a pocket with the SH3 domain into which the extended motif fits, much like Sho1. However, this surface has no structural or sequence similarity with Sho1 and has been previously identified as the Cdc42 interacting (CI) region (Takaku *et al.* 2010). Contrary to Sho1, this additional binding region is located C-terminal to the SH3 domain.

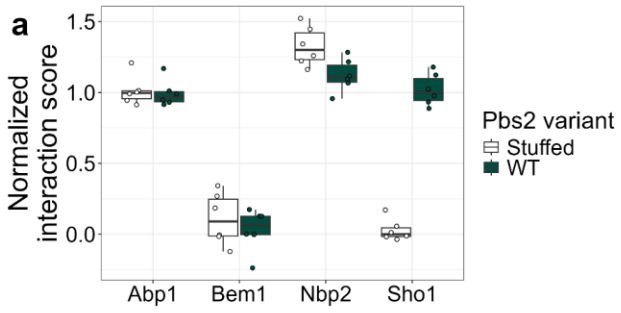
There is little sequence conservation between the extended motifs of Pbs2, Ark1, and Ste20, apart from the conserved positions of the canonical motifs. These extended motifs can be aligned using a previously developed standard numbering, where the first conserved proline is designated 0, residues preceding the motif are numbered with decreasing negative numbers, and residues following the motif are numbered with increasing numbers (Lim *et al.* 1994) (Figure 4c). The binding specificity exhibited by Pbs2 may depend on the possible contacts with the different SH3 domains, and the non-SH3 surfaces in the cases of Sho1 and Bem1-2. Examining the different motifs and their contacts with their respective interaction partners, we can compare Pbs2 to the two other extended motifs, and in particular the positions which are in the extended motif but outside the canonical binding motif, to establish determinants of specificity.

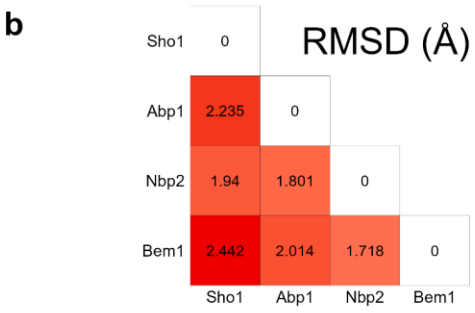
Foremost, Pbs2 binds using residues more distal from the canonical motif than either Ste20 or Ark1. Contacts are predicted for Q(-14) as well as H(-10) on the predicted Sho1-Pbs2 structure, which is not the case for any other interaction. In Pbs2, I(-9), which can have a large impact on binding when mutated, contacts the hydrophobic residues L290 and Y293 on the non-SH3 surface of Sho1. The equivalent position on Ste20, G(-9) contacts the polar residues Q240 and K243 in the non-SH3 surface of Bem1, which would be unfavorable for the isoleucine in Pbs2. In Ark1 binding to Abp1, the residue at position (-9) contacts the SH3 domain (Figure 5a).

Position (-7) is important for binding in both Ste20 (F(-7)) and Ark1(L(-7)). When binding to their respective partners, including both Bem1-2 and Nbp2 for Ste20, mutation of position (-7) to an alanine reduced the strength of binding by at least an order of magnitude. Sequentially mutating all positions in the extended motif to an alanine showed that position (-7) had the most impact of any position outside the canonical motif in all cases (Stollar *et al.* 2009; Gorelik and Davidson 2012). In Pbs2, position (-7) is occupied by a polar glutamine, as opposed to the hydrophobic residues found in the same position in the other two peptides. Also, in the predicted Sho1-Pbs2 structure, the side chain of Q(-7) is pointed away from Sho1 and is not predicted to have any contacts, while in the other structures, the equivalent side chains are pointed towards the SH3 domains (Figure S17a).

Position (-6) also marks a difference, specifically between Pbs2 and Ark1. In Pbs2, I(-6) contacts Y42 on the SH3 domain. In Ark1, H(-6) forms a salt bridge with an aspartic acid situated on a different loop of the Abp1 SH3 domain (Figure 5b). The structural alignment of SH3 domains indicates that Sho1 has a gap at the position of the salt bridge forming aspartic acid (Figure S14b, position 112). The last major difference is in position (-4), where the residues in both Ste20 and Ark1 make contact with an SH3 domain, while N(-4) in Pbs2 is not predicted to have any contacts. As with position (-7), the polar side-chain is pointed away from the SH3 domain (Figure S17b).

Overall, we can see that the Pbs2 extended motif is very different from the other extended motifs known to bind SH3 domains in yeast. The additional contacts in positions (-14) and (-10), the differing contacts in positions (-9) and (-6), and the missing contact in position (-4) and especially the important missing contact in position (-7) may preclude Pbs2 from binding to SH3 domains other than Sho1. If this is truly the case, the presence of a non-conserved portion in the extended motif allows much more differentiation in binding than solely the canonical motif.



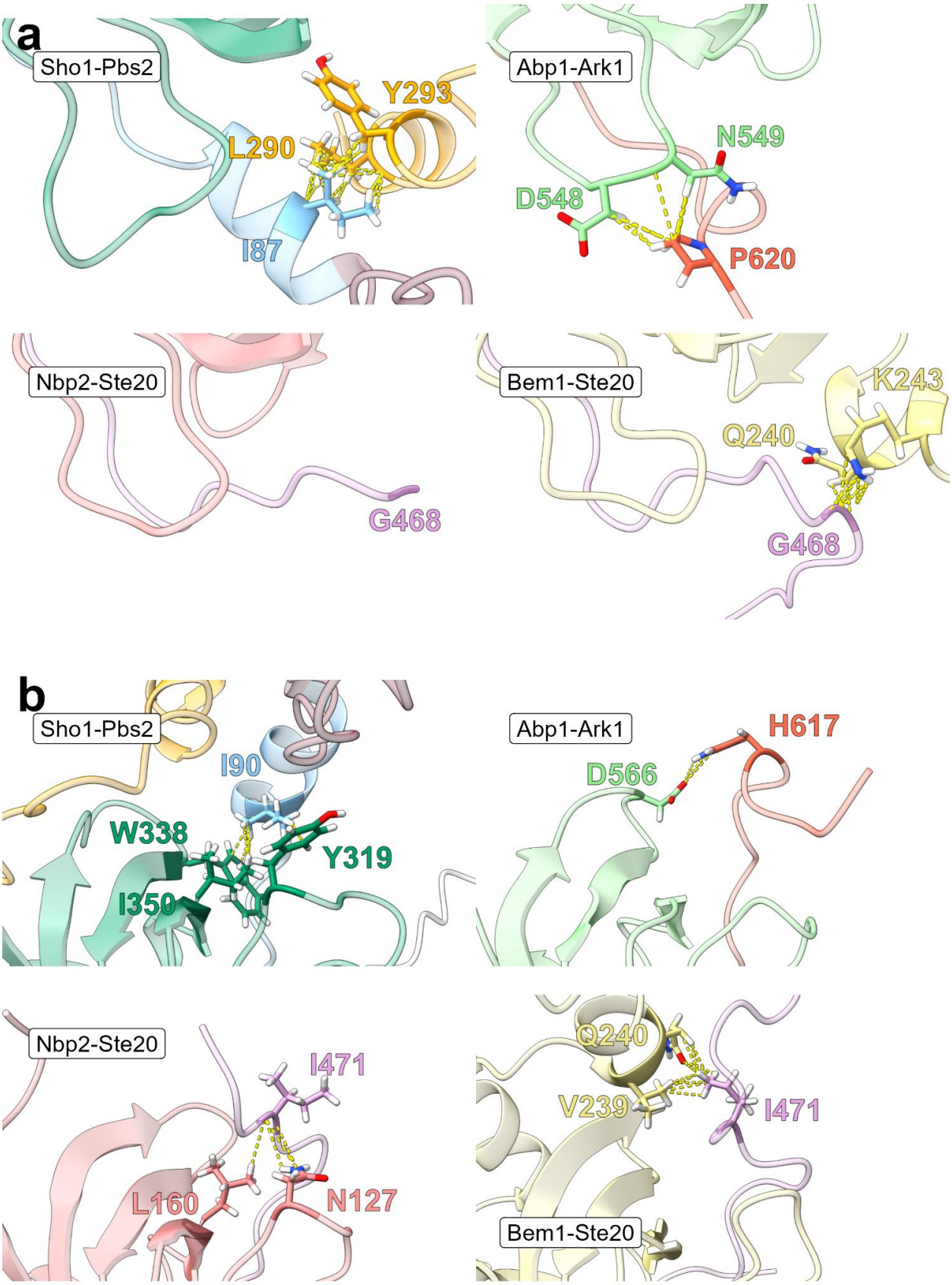


c

												Canonical motif										
Standard numbering	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7
Pbs2	Q	S	P	K	H	I	Q	Q	I	V	N	K	P	L	P	P	L	P	V	A	G	S
Ark1					K	P	K	L	H	S	P	K	P	P	P	T	P	K	T	K	K	A
Ste20	S	S	S	A	N	G	K	F	I	P	S	R	P	A	P	K	P	P	S	S	A	
Pbs2 position	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103

Figure 4. The Pbs2 motif binds specifically to Sho1, despite similarities of other SH3 domains and motifs

a) Interaction score of wild-type Pbs2 (green) or Pbs2 where positions 76 to 126 have been replaced with a flexible stuffer sequence (Stuffed, white) with select yeast SH3-containing proteins. All replicates shown. Scores were normalized, with 1 being the median of Sho1 - wild-type Pbs2 interactions and 0 being the median of Sho1 - stuffed Pbs2 interactions. b) Root mean square deviation of atom locations between matching backbone atoms in each pair of structures, as calculated using the R package bio3d (Grant *et al.* 2006). c) Sequences of Ste20, Ark1 and Pbs2, with standard numbering according to the motif, to compare equivalent positions (Lim *et al.* 1994). For Ste20 and Ark1, the positions included are those present in the structures in panel b, with Pbs2 positions chosen to overlap all represented positions.


a
Sho1-Pbs2
Y293
L290
I87
Abp1-Ark1
N549
D548
P620
Nbp2-Ste20
G468
Bem1-Ste20
K243
Q240
G468
b
Sho1-Pbs2
I90
W338
Y319
I350
Abp1-Ark1
H617
D566
Nbp2-Ste20
I471
L160
N127
Q240
V239
I471
Bem1-Ste20

537

Figure 5. Differing contacts between SH3-extended motif pairs

a) Visualization of the side-chains and contacts of position (-9) in the Sho1-Pbs2 predicted structure, Abp1-Ark1 structure (PDB accession 2RPN), Nbp2-Ste20 structure (PDB accession 2LCS) and Bem1-2-Ste20 (PDB accession 2RQW). Sho1-Pbs2 predicted structure colored as in figure 3a, with the SH3 domain of Sho1 in green, the non-SH3 portions of Sho1 in orange, the Pbs2 extended motif in light blue and the rest of Pbs2 in burgundy. All four structures were superimposed using the Chimera X tool "Matchmaker" (Pettersen *et al.* 2021) to minimize root mean square deviance of backbone atom distances of the SH3 domains with Sho1, then pictures were captured separately. Position (-9) and its contacts are kept opaque while the rest of the structures are semi-transparent. b) Visualization of the side-chains and contacts of position (-6) of the same structures as in panel a). All four structures were superimposed using the Chimera X tool "Matchmaker" (Pettersen *et al.* 2021) to minimize root mean square deviance of backbone atom distances of the SH3 domains with Sho1, then pictures were captured separately. Position (-6) and its contacts are kept opaque while the rest of the structures are semi-transparent.

Discussion

Binding motifs face the double challenge of binding their partner with appropriate strength while avoiding spurious interactions with non-partner proteins. Using deep mutational scanning libraries of the Pbs2 binding motif and its surrounding sequences, we measured the impact of single residue mutations on interaction strength with Sho1. We also predicted the structure of the Sho1-Pbs2 interaction, and compared it to known structures of SH3-mediated interactions. We use these comparisons to propose a number of ways in which the Pbs2 extended motif may bind specifically to Sho1, and not other SH3 domains.

One important consideration in domain-peptide interactions is the potential trade-off between interaction affinity and specificity. Previous studies on SH2 and PDZ domains, both domains that bind short motifs similarly to SH3 domains, have put forth that mutants which bind with higher affinity also possess lower specificity (Ernst *et al.* 2010; Haslam and Shields 2012; Kaneko *et al.* 2012; Karlsson *et al.* 2016), suggesting that high affinity and high specificity may be incompatible. This represents an intuitive idea: a protein capable of having strong interactions with a certain motif will also have more frequent and stronger spurious interactions with motifs of the same family, which are physically and chemically similar. On the other hand, a computational assay studying binding motifs, including SH3 binding motifs, found that an increase in affinity can lead

to an increase in specificity (Kelil *et al.* 2016). Other works have emphasized the importance of the residues surrounding the canonical binding motif — in this case defined as K/RXXPXXP, corresponding to positions 93 to 99 of Pbs2 — in determining both affinity and specificity (Li 2005; Ivarsson and Jemth 2019). In fact, another study computed that nearly 30% of binding energy from the binding motif was contributed by residues outside the canonical motif, and suggests that the residues outside the canonical motif mostly play a role in determining specificity (Stein and Aloy 2008). In our experimental work, we found that all mutations which significantly increase the interaction strength between Sho1 and Pbs2 are situated outside of the canonical motif. Consequently, the wild-type canonical motif could be already optimized for the strongest possible interaction strength. This suggests different roles for the canonical motif and its neighboring residues in determining affinity and specificity, with canonical motif residues optimized and therefore more constrained, while the neighboring residues are less constrained and able to increase affinity and specificity when mutated.

One interesting possibility that may explain the higher binding strength conferred by mutations to certain residues outside the Pbs2 canonical binding motif is that the mutated residues may interact with either residues outside the canonical binding pocket in the SH3 domain or even with non-SH3 surfaces on Sho1. Non-binding pocket interactions have previously been observed in multiple SH3-mediated interactions (Lee *et al.* 1996; Dalgarno *et al.* 1997; Li 2005; Stollar *et al.* 2009; Takaku *et al.* 2010; Gorelik and Davidson 2012; Gaussmann *et al.* 2024). Using predicted protein complex structures generated by AlphaFold-Multimer, we were able to hypothesize that a hydrophobic pocket is formed between the SH3 domain and a non-SH3 loop of Sho1, which residues of the extended motif can fill. The presence of such a pocket can explain the increased affinity of many Pbs2 mutants in the extended motif, as predicted structures of these models show that substituted hydrophobic side chains can insert themselves into the predicted pocket. Perhaps the disordered region encompassing the Sho1-binding motif can place itself in such a way to maximize the hydrophobic interaction between any mutated side chain and the hydrophobic pocket. These predicted structures may represent only one possible form of the Sho1-Pbs2 interaction, which may in reality be dynamic due to the disordered nature of the region surrounding the Pbs2 motif. Nevertheless, they offer a potential example of binding mediated by positions outside the canonical motif. By removing the need for affinity to be dependent only on canonical binding motifs, the residues surrounding motifs could bind to protein surfaces which are less conserved among SH3 domains, and thus easier to differentiate. The presence of unique components, such as extended motifs or extended binding surfaces could therefore contribute to specificity among domains of the same family. The hydrophobic pocket formed by both SH3 and non-SH3 parts of Sho1 may represent an example of a unique structural feature which may help differentiate Sho1 from other SH3 domains, and allow increased affinity to develop in Pbs2 without increasing the affinity to other SH3 containing proteins. This proposed model of binding also involves a larger number of residues in tuning affinity and specificity, potentially uncoupling affinity and specificity and therefore avoiding a trade-off between these two phenotypes. Affinity and specificity may only be linked for highly conserved residues and structures, such as the canonical binding motif or the canonical SH3 binding pocket, but for less constrained residues, may develop independently.

Studies on interaction domains have shown that SH3 domain mediated interactions are context specific, meaning that the properties of the interactions depend not only on the domains themselves, but also on the protein context. This context can be composed of many factors, including the non-SH3 sequence of the protein, other proteins interacting with the SH3 containing protein, and cellular localization (Dionne *et al.* 2022). SH3 mediated interactions therefore do not simply depend on the SH3 sequence. For example, when inserting an SH3 domain into a different protein or expressing an SH3 domain without its host protein, the SH3 domain's interaction profile substantially changes (Dionne *et al.* 2021; Lemieux *et al.* 2024; Dibyachintan *et al.* 2024). Thus, Pbs2 mutations may increase or decrease binding strength not only with the Sho1 SH3 domain, but also within the Sho1 context. It is not entirely clear how Pbs2 mutations modulate binding within the Sho1 context, but the binding of the Pbs2 extended motif to non-SH3 residues could explain a large part of this phenomenon. Other causes could include stereochemical adjustments of SH3 binding loops or changes which contribute to stabilizing or destabilizing interactions with other Sho1 binding partners. There are observations of other proteins, namely Ste11 and Ste50, interacting with the Sho1 SH3 domain outside of the canonical binding pocket (Zarrinpar *et al.* 2004; Tatebayashi *et al.* 2006). Changes to the Pbs2 peptide may affect the stability of the interaction of these proteins, which are also involved in the osmotic stress response pathway.

Previous work has already focused on the binding of Pbs2 motif variants (Zarrinpar *et al.* 2003). We replicate some of the findings of this previous work, for example, showing that mutations near the canonical Pbs2 motif can cause both gain and loss of binding strength. This study also finds that a promiscuous double mutant of the Pbs2 motif has lower fitness (Zarrinpar *et al.* 2003). We find that, apart from nonsense variants, no mutants confer a significant loss of fitness. However, Pbs2 mutant I87W has a noticeable, though not statistically significant, negative effect on cell fitness (Figure S11). I87W is also the mutant which most increases the interaction strength, which may be a cause of the slightly reduced fitness. Mutations which increase affinity have been suggested to have deleterious effects, by reducing the dynamicity of the pathway, and impeding dissociation of the partners (Wang *et al.* 2018). For the Sho1-Pbs2 interaction, previous work has found that a mutation in the Sho1 SH3 domain that increased affinity also led to a small reduction in cell growth (Marles *et al.* 2004), indicating that the wild-type affinity may be already optimized, and that stronger binding may disrupt signaling.

In conclusion, we show that mutations in positions neighboring the canonical binding motif of Pbs2 can both increase and decrease binding strength with Sho1. We propose a role for residues outside the canonical Pbs2 binding motif in increasing binding affinity independently of binding in the canonical SH3 binding pocket. We hypothesize that Sho1 is similar to certain other yeast SH3 domains which bind longer motifs through additional binding pockets (Stollar *et al.* 2009; Takaku *et al.* 2010; Gorelik and Davidson 2012), and propose mechanisms through which an extended binding motif could better differentiate binding to different SH3 domains. Going forward, studies of SH3-motif binding should consider a wider range of residues, both on the SH3 domain and surrounding the motif. Our work, along with previous work by other groups, suggests that *in vitro* work using isolated domains and shortened peptides does not capture the full picture of domain-motif interaction. By understanding and harnessing the potential of extended motif binding,

researchers may be better able to understand the impact of motifs in cell signaling and disease, and better able to use motif binding in technical and medical applications.

Methods

Growth conditions

Escherichia coli cells were grown in 2YT medium (1% yeast extract, 1.6% tryptone, 0.2% glucose, 0.5% NaCl, and 2% agar for solid plates) with shaking at 37 °C for liquid cultures. When specified, 100 µg/mL of ampicillin (Amp, Bioshop Canada) were added as a selection agent. *Saccharomyces cerevisiae* cells were grown in YPD medium (1% yeast extract, 2% tryptone, 2% glucose, and 2% agar for solid plates), synthetic complete medium buffered to a pH of 6.0 (0.174% yeast nitrogen base without amino acids, without ammonium sulfate, 2% glucose, 0.134% amino acid dropout complete, 0.1% monosodium glutamate, with 1% succinic acid and 0.6% NaOH as buffer) or PCA medium (0.67% yeast nitrogen base without amino acids and without ammonium sulfate, 2% glucose, 10% liquid drop-out without adenine, methionine, and lysine, and 200 µg/ml methotrexate (MTX, Bioshop Canada) diluted in dimethyl sulfoxide (DMSO, Bioshop Canada), and 2.5% noble agar for solid plates), as specified. When specified, 200 µg/mL of G418 (Bioshop Canada), 100 µg/mL of nourseothricin (Nat, Jena Bioscience) or 250 µg/mL of hygromycin B (Hyg, Bioshop Canada) were used as selection agents. Also, when specified, 1 M of sorbitol (Bioshop Canada) was added to either liquid or solid PCA medium. Yeast cultures were grown at 30 °C with agitation at 250 rpm. See Table S4 for details of all growth media used.

Cloning

All plasmids were constructed by Gibson assembly. As template for the *PBS2* DMS library construction (pUC19-Pbs2, see below), a pUC19 plasmid backbone was amplified (Addgene #50005, primers 1F and 1R, all primers detailed in Table S5), then digested for 1 hour at 37 °C using SacI (New England Biolabs). This linearized plasmid was used for Gibson assembly with an insert composed of a 248 base pair region of *PBS2* amplified from genomic DNA, using primers that add homology arms (Primers 2F and 2R). For the CRISPR-Cas9 based insertion of sequences, the pCAS plasmid (Addgene #60847) was modified as previously described (Ryan *et al.* 2016) to change the sgRNA to target either the Sho1 binding motif of Pbs2 (pCAS-Pbs2, primers 3F and 3R), the SH3 domain of Sho1 (pCAS-Sho1, primers 4F and 4R), or the stuffer sequence which was inserted into *PBS2* (pCAS-stuffer, primer 5F and 5R) (Dionne *et al.* 2021). Proper plasmid construction was verified using Sanger sequencing (Primers 6F and 6R) (Messing 1983).

Strain construction

All strains used are summarized in Table S6. Yeast strains were constructed with either the F[1,2] or the F[3] fragments of an engineered murine DHFR protein fused to the C-terminus of different

proteins, to measure protein interaction strength as previously described (Tarassov *et al.* 2008; Michnick *et al.* 2016). The transformed sequence contains the DHFR F[1,2] fusion sequence as well as a nourseothricin resistance cassette, or the DHFR F[3] fusion sequence with a hygromycin B resistance cassette. This allows selection of the strains with DHFR fragment fusions. The DHFR F[1,2] tagged strains were derived from MATa strain BY4741 (*his3Δ leu2Δ met15Δ ura3Δ*), while the DHFR F[3] strains were derived from MATα strain BY4742 (*his3Δ leu2Δ lys2Δ ura3Δ*). The Hog1-DHFR F[1,2] strain and the Pbs2-DHFR F[3] strain used in the paper were taken from the Yeast Protein Interactome Collection (Horizon Discovery) (Tarassov *et al.* 2008). The Sho1-DHFR F[1,2] strain additionally had a 1X FLAG tag at the end of the fragment construct, and its construction was described in a previous publication (Dionne *et al.* 2021). Briefly, the DHFR F[1,2], the Nat cassette, a flexible linker and the FLAG tag sequence were amplified from the pAG25-DHFR F[1,2]-linker-FLAG plasmid (Dionne *et al.* 2021) using primers with homology arms matching the regions either side of the *SHO1* termination codon. The cassette was transformed into BY4741 cells and integrated by homologous recombination. Further deletions were done in the DHFR F[1,2] tagged strains in order to knock out *PBS2*, using the *LEU2* cassette from pUG37 (Gueldener *et al.* 2002), through homologous recombination (Primers 7F and 7R). Proper deletion was verified by PCR (Primers 8 to 11) and sequencing (Gueldener *et al.* 2002). After the mating of the DHFR F[1,2] MATa strains with the Pbs2-DHFR F[3] MATα, only one copy of *PBS2* is present. This prevents the additional wild-type allele from influencing the measurements of interaction strength of the variants.

The Pbs2 mutants of interest for the DHFR-PCA growth curves were constructed from the Pbs2-DHFR F[3] fusion strain. First, the region surrounding the Sho1 binding motif of Pbs2 (codons 71 to 126) was replaced with a stuffer sequence (GGCGGAAGTTCTGGAGGTGGTGGT) that translates into a flexible linker sequence (GGSSGGGG), by co-transformation of the pCAS-Pbs2 plasmid, and of a repair template consisting of the stuffer sequence flanked by homology arms matching the sequences around codons 71 to 126 of *PBS2* (produced by amplifying the stuffer sequence with primers 12F and 12R), following a previously published protocol (Ryan *et al.* 2016). This produced the Pbs2-stuffed-DHFR F[3] strain, which was used for the construction of the individually reconstructed mutants, for the construction of the DMS libraries (see below), and also as a control to measure Pbs2 binding to SH3-containing yeast proteins. The individually reconstructed mutants contain all mutants used in a previous paper exploring Pbs2 mutants (Zarrinpar *et al.* 2003), including a P94A+P97A double mutant. The mutants were individually constructed by co-transforming the Pbs2-stuffed-DHFR F[3] strain with the pCAS-stuffer plasmid targeting the aforementioned stuffer, and synthesized oligonucleotide sequences corresponding to the desired mutation (Integrated DNA Technologies) (Table S7). The stuffed strain as well as all reconstructed strains were verified by PCR and Sanger sequencing (Primers 13F and 13R). The pCAS plasmids were purged by growth in liquid YPD without G418, and loss of the plasmid was verified by lack of growth on YPD+G418 plates.

For the extended motif DMS library and the individually reconstructed mutants used in the validation competition assays, three additional haploid strains were built, into which Pbs2 DHFR mutations were inserted. The first two strains were Hog1-DHFR F[1,2]/Pbs2-stuffed-DHFR F[3] and Sho1-DHFR F[1,2]/Pbs2-stuffed-DHFR F[3]. These were built by adding the DHFR F[1,2] tag

to either *HOG1* or *SHO1* in the MATα Pbs2-stuffed-DHFR F[3] strain previously constructed. The DHFR F[1,2] moieties were amplified from plasmid pAG25-DHFR F[1,2]-linker-FLAG plasmid using primers with homology arms to either the *HOG1* or *SHO1* locus (Primers 14 F to 15R) (Dionne *et al.* 2021). Proper tagging was verified using PCR and Sanger sequencing (Primers 16 to 18) (Freschi *et al.* 2013; Dionne *et al.* 2021). The third strain was Pbs2-stuffed, which was used to measure the proliferation of mutants without the effect of the DHFR fragment. This was constructed in the haploid strain BY4742, using the same strategy to insert the stuffer into *PBS2* as detailed above.

The selected Pbs2 mutants for the validation assay were constructed in these three same strains by transformation with repair templates containing the mutations of interest, formed through fusion PCR. Briefly, forward and reverse primers were designed for each mutant to be constructed, which contained the mutated sequence instead of the wild-type sequence (Primers 19F to 110R). These primers were used for separate PCR amplification of the *PBS2* sequence along with either common forward or reverse primers (Primers 111F and 111R), to create two overlapping fragments. These were combined in a second PCR using only the two common primers, which fused the overlapping sequences together, to create a single fragment of 338 base pairs, which was identical to *PBS2* except for the mutated codon of interest. These fragments were transformed into DHFR F[1,2]/Pbs2-stuffed-DHFR F[3] strains using the CRISPR-Cas9 strategy described above. Mutants were confirmed by PCR and sequencing (Primers 112F and 112R). Most mutations were constructed in all three DHFR strains, but a certain number could only be built in one or two of the strains. The successfully constructed mutants are listed in Table S6.

For the DHFR-PCA screen on solid media, the previously constructed wild-type Pbs2-DHFR F[3] and Pbs2-stuffed-DHFR F[3] were used. The SH3-containing proteins DHFR F[1,2] strains were obtained from the Yeast Protein Interactome Collection (Horizon Discovery). A version of Sho1 was also built where the entire SH3 domain was replaced by the same flexible stuffer used in the Pbs2-stuffed strain. As with Pbs2, the stuffer was amplified using primers adding homology to the sequences either side of the Sho1 SH3 domain locus (Primers 113F and 113R). The stuffer was inserted into the Sho1 sequence by cotransformation of the pCas-Sho1 plasmid and the repair template of the stuffer with homology arms. Proper stuffing was verified by PCR and Sanger sequencing (Primers 114F and 114R).

<u>Construction of the surrounding region DMS library</u>

The initial DMS library covering the surrounding region of Pbs2 was constructed as previously reported (Dionne *et al.* 2021). 56 codons (168 nucleotides) were targeted, which correspond to codons 71 to 126 of the *YJL128C*/*PBS2* ORF.

A degenerate oligonucleotide was designed for each codon to be mutated, with the three nucleotides of the mutated codon replaced by an NNN codon (Primers 115 to 170). This series of 56 oligonucleotides was used to amplify pUC19-Pbs2, in a two-step PCR procedure (Miyazaki 2011). At all steps, the different codons were kept in separate reactions and the mutants for each codon were created in separate reactions. In the first amplification, a megaprimer was created

which includes the *PBS2* coding sequence using the degenerate oligonucleotide and a common reverse primer (Primer 171R). The megaprimer thus carries the DMS library for one codon. In the second PCR step, this megaprimer was used to amplify the entire plasmid. The PCR products were digested with DpnI (New England Biolabs) at 37°C for 1 hour to remove the original methylated template, and keep only the mutated plasmids amplified by the megaprimer. The remaining mutated amplified but unligated plasmids were transformed into chemocompetent *E. coli* cells strain MC1061, and plated on solid 2YT+Amp. The amplified plasmids were ligated and thus recircularized in the cells. All transformations resulted in hundreds of colonies and these colonies were suspended in 5 mL of 2YT. Their plasmids were extracted using a miniprep plasmid extraction kit (Presto Mini Plasmid Kit, Geneaid) to obtain a plasmid library containing all the DMS mutants. Prepared libraries were sequenced using 300-PE MiSeq technology (see DNA sequencing section), to verify that the desired codon diversity was present. Once diversity was confirmed, the libraries were amplified from the plasmids (Primers 172F and 172R), and transformed into the Pbs2-stuffed-DHFR F[3] yeast strain using the CRISPR-Cas9 strategy described in the strain construction section.

Transformed cells were plated on solid YPD+G418+Hyg for selection, and then resuspended in liquid YPD. The optical density of each suspension (each codon) was measured and all suspensions were pooled into a masterpool, with 5 optical density units (OD) added for each codon. Frozen glycerol stocks were prepared for the individual suspensions of each codon, and multiple stocks were prepared for the masterpool. Genomic DNA was extracted from the masterpool using phenol/chloroform DNA extraction (Amberg *et al.* 2005), and as before, the library was sequenced to verify diversity. At this point, three codons had unsatisfactory diversity. A new template was created for codons 103 and 119 using fusion PCR, creating a sequence of 967 base pairs around the mutated codon, using the original degenerate primers as well as new reverse degenerate primers to create an NNN codon at either codon 103 or 119 (Primers 173 to 175). For codon 88, a new template was ordered as an oligonucleotide of 503 base pairs (Twist Bioscience), around codon 88, which is replaced by an NNN codon. These templates were transformed into Pbs2-stuffed-DHFR F[3] as described above, and diversity at the desired codon was verified with Sanger sequencing. The three supplementary cultures were pooled into the masterpool, with an OD corresponding to 1/56th of the OD of the master pool.

<u>DHFR-PCA competition assay for surrounding region DMS library</u>

The PCA selection followed a previously published protocol (Dubé *et al.* 2022). Three liquid cultures were started, each from 100 µL of the masterpool containing all mutated codon positions, in SC complete medium. These cultures were incubated at 30 °C for 16 hours. 300 µL of this preculture was mixed with 150 µL of either Sho1-DHFR F[1,2] or Hog1-DHFR F[1,2], and 1.2 mL of fresh liquid YPD. These mixes of strains were left to mate for 8 hours at 30 °C with agitation. The mated diploid cells were then selected in two successive cycles of diploid selection in liquid YPD+Nat+Hyg at 30 °C with agitation. The first cycle consisted of a 5/12 dilution of the mating culture in 3 mL of liquid YPD for 16 hours, and the second cycle consisted of a 1/10 dilution in 3 mL of liquid SC media for 24 hours. A volume equivalent to 5 OD was spun down and the supernatant was removed, to form a cellular pellet, which was stored at -80 °C. This is the initial

time point used for sequencing. The selected diploid cultures were then diluted to 0.1 OD/mL in 15 mL of liquid PCA medium with either 200 µg/mL of methotrexate with 1 M of sorbitol, the same concentration of methotrexate without sorbitol, or a control without methotrexate but with 1 M sorbitol. This PCA selection was grown without light for 96 hours, except for the control cultures which were saturated after 24 hours. After this first growth cycle, the cultures were diluted 0.1 OD/mL into 15 mL of fresh liquid PCA medium, and a 5 OD pellet was spun down and stored at -80 °C. The second PCA cycle lasted the same amount of time, and at the end, as many 5 OD pellets as possible were prepared from each culture. Genomic DNA was extracted from the pellets of the second diploid selection cultures and the second PCA selection culture. The Pbs2 locus was amplified and sequenced, as described in the DNA sequencing section.

Construction of the extended motif DMS library

A second DMS library was constructed of the extended motif of *PBS2*, comprising codons 85 to 100. An oligonucleotide pool (Integrated Data Technologies, Table S8) was synthesized containing all possible NNK codons (K indicating either a G or T nucleotide) for the 16 codon region of interest, with homology arms for a total length of 248 base pairs. This pool was integrated into 2 different yeast strains: Sho1-DHFR F[1,2]/Pbs2-stuffed-DHFR F[3] and Hog1-DHFR F[1,2]/Pbs2-stuffed-DHFR F[3], to allow respectively for PCA measurements of Sho1-Pbs2 interaction in a haploid strain, and PCA measurements of Hog1-Pbs2 interaction in a haploid strain.

The mutant pool was integrated into these strains by replacing the Pbs2 stuffer using the CRISPR-Cas9 strategy described in the strain construction section. As before, the integrated libraries were sequenced to verify diversity. Transformed yeast libraries were stored at -80 °C.

DHFR-PCA competition assay for extended motif DMS library

The liquid PCA selection for the haploid double DHFR tagged strains was done following the same protocol as the first DHFR-PCA selection, only ignoring the mating step. However, a fourth control condition of PCA medium without methotrexate and without sorbitol was added. Three replicate pools were done for each growth condition. Frozen pellets were prepared as in the first DHFR-PCA selection.

Validation DHFR-PCA competition assay of individually reconstructed mutants

The individually reconstructed validation strains were pooled by combining 100 µL of saturated overnight culture of each mutant in YPD. 6 pools were made for each condition, three pools each from two individual reconstructed colonies of the mutant. The same conditions were used as the second liquid PCA selection: PCA medium both with and without 1 M of sorbitol and with and without methotrexate for the DHFR tagged strains. For the cell proliferation measurements, SC medium with and without 1 M of sorbitol was used. Two cycles of selection were done, with 96-hour cycles for the conditions with methotrexate and 24 hour cycles for the conditions without methotrexate, and the SC medium. Frozen pellets were prepared as for the previous liquid PCA selections.

DNA sequencing

DNA sequencing was done in generally the same manner for all experiments, with minor changes in certain cases. DNA was extracted from the frozen pellets using a standard phenol/chloroform genomic DNA extraction protocol (Amberg *et al.* 2005). The extracted DNA was amplified and barcoded using a Row-Column DNA barcoding strategy as previously described (Dubé *et al.* 2022). From the extracted genomic DNA, the *PBS2* region of interest was amplified using primers containing 3' overhangs which allow a second amplification to add barcodes (Primers 176F and 176R). PCR products were diluted 1/2500, and used as a template to add row and column barcode primers onto the sequence, with a unique combination of 5' and 3' barcodes for each replicate in each condition. The row-column barcode primers were previously described (Dubé *et al.* 2022). PCR products were migrated on an agarose gel, and DNA concentrations were estimated based on band size using the Image Lab software (BioRad Laboratories). Samples were then pooled with different volumes to obtain the same quantity of DNA from each sample. The pools were purified using magnetic beads, measured using a NanoDrop (ThermoFisher), and diluted to 0.1 ng/µL. The diluted pools were then amplified with primers adding a second set of barcodes, called plate barcodes (Dubé *et al.* 2022). These double-barcoded pools were purified on magnetic beads, and sent for sequencing on Illumina NGS machines. Unique combinations of barcodes were used to identify the DNA extracted from each pellet in each replicate of each condition. The sequencing of the first DHFR-PCA, on the larger region surrounding the Pbs2 motif was done using Illumina NovaSeq paired-end 250 base pair technology (CHUL sequencing platform, Quebec, Canada), while all other sequencing was done using Illumina MiSeq paired-end 300 bp technology (IBIS Genomic Analysis Platform, Quebec, Canada).

After preliminary analysis of the sequencing of the DNA from the pellets of the DHFR-PCA on the extended motif DMS library, it was determined that more reads would be needed. Additional sequencing libraries were prepared fresh from the original DNA extractions. Instead of using a row-column barcoding approach, the *PBS2* region was amplified using primers (Primers 177F and 177R) with homology arms allowing the addition of Illumina Nextera barcodes (Illumina). These allow automated demultiplexing by the Illumina MiSeq instrument. All other steps were as described above.

Analysis of DNA sequencing results

The variant frequency in DMS libraries during construction and in the DHFR-PCA and proliferation screens was evaluated using custom scripts based on already published work (Després *et al.* 2022). Python libraries pandas 1.5.2 (The pandas development team 2023), matplotlib 3.6.2 (Hunter 2007) and numpy 1.24.1 (Harris *et al.* 2020) were used for data manipulation and visualization. Quality was assessed using FastQC 0.12.1 (Andrews 2010). Reads were trimmed using Trimmomatic 0.39 (Bolger *et al.* 2014), then demultiplexed using bowtie 1.3.1 for the sequencing of the surrounding region DMS library (Langmead *et al.* 2009), Interstellar 1.0 following the RCP-PCR configuration for the first sequencing run of the extended motif DMS library (Kijima *et al.* 2023) and cutadapt 4.7 for the validation competition assay sequencing

(Martin 2011). The second sequencing run for the extended motif DMS library was barcoded using Illumina Nextera primers, and was thus automatically demultiplexed by the sequencing instrument. Forward and reverse reads were then merged using the PANDAseq software (Masella *et al.* 2012). Next, identical reads were grouped using vsearch (Rognes *et al.* 2016), and aligned to the wild-type *PBS2* sequence using the needle function from the EMBOSS software (Rice *et al.* 2000). From this, the frequency of each variant was obtained.

From the frequency of each variant, the interaction scores and selection coefficients of each variant were calculated using custom R 4.3.1 (R Core Team 2023) scripts. The R package collection tidyverse (Wickham *et al.* 2019), as well as the packages rstatix 0.7.0 (Kassambara 2021), ggpubr 0.4.0 (Kassambara 2020) and bioseq 0.1.4 (Keck 2020) were used for data analysis. R packages ggExtra 0.10.1 (Attali and Baker 2023), ggrepel 0.9.4 (Slowikowski 2023), GGally 2.2.0 (Schloerke *et al.* 2023) and cowplot 1.1.1 (Wilke 2020) were used for data visualization. Briefly, variant counts were imported, identified by the mutation type (wild type, nonsense, missense or silent) then transformed into frequencies by dividing each variant count by the total number of counts present in that sequencing library. The interaction score was calculated as the log-2-fold-change of the frequency of each variant at the end of the second cycle of methotrexate selection and the frequency of each variant after the selection for diploids (so immediately preceding the methotrexate selection). These log-2-fold-change scores were rescaled for each library, so that the median of nonsense codons was equivalent to -1 and the value of wild-type variants was 0. Variants were then filtered to keep only those that had 20 or more reads detected after diploid selection, in order to remove bias caused by low read counts leading to inflated scores. Next, a unique score was calculated for each variant by taking the median score of all codons coding for the same amino acid variant in all replicates of the same condition. Scores were only kept for amino acid variants that had at least 3 replicates from any combination of codons not filtered out. For the competition assay of cell proliferation, the number of generations elapsed during the selection for each replicate was calculated, as the log-2-fold-change the optical density of the cultures at the end and at the beginning of the growth cycles. The selection coefficient for each variant was calculated as follows:

$$s = \frac{ln\left(\frac{variant\ frequency\ after\ PCA\ selection}{variant\ frequency\ before\ PCA\ selection}\right) - ln\left(\frac{wt\ frequency\ after\ PCA\ selection}{wt\ frequency\ before\ PCA\ selection}\right)}{number\ of\ generations}$$

The frequency of the wild type used is the frequency of the wild type in the same replicate as the variant. The same filtering criteria were used for the proliferation assay as for the DHFR-PCA.

When measuring the interaction scores of the preliminary and extended motif DMS libraries, we also included control conditions without methotrexate, the selective agent allowing measurement of DHFR fragment reconstitution. The measurements in this condition should not give any information about interactions between Pbs2 and Sho1, but can inform us about any other defects in the mutant strains. We calculated DMSO scores for growth in the control condition in the same way as the interaction scores. We noticed that some nonsense mutants had growth scores

comparable to wild-type Pbs2, and that some silent mutants had growth scores comparable to nonsense mutants. These abnormal scores could indicate defects in the strains, possibly resulting from the Cas9 mediated mutation process. We statistically verified which nonsense and silent mutants were abnormal by comparing every nonsense or silent mutation to all other nonsense or silent mutations (Two-sided Mann-Whitney U test, $p < 0.05$). The mutants which were significantly different were removed from the interaction score dataset. Similarly, we wanted to verify that no missense mutants had any defects caused by other sources than the interaction. We statistically verified which mutants had defects in the control condition by comparing each missense mutant to the nonsense mutants (Two-sided Mann-Whitney U test with false discovery rate corrected p-value > 0.05). The missense mutants which were not significantly different from the nonsense mutants were also removed from the interaction score dataset.

We also removed from the interaction score dataset the mutants for which an abundance change was detected through gain or loss of interaction strength between Pbs2 and Hog1, as explained in the results section. Pbs2 missense mutants that had an interaction strength with Hog1 which was significantly different from wild-type Pbs2 and silent mutations (Two-sided Mann-Whitney U test with false discovery rate corrected p-value < 0.05), either in the presence or absence of sorbitol, were removed from the datasets.

DHFR-PCA growth curves of individually reconstructed mutants

Growth curves were measured for individually reconstructed Pbs2 mutants in DHFR-PCA medium, to validate Pbs2 interaction strength measurements. The DHFR F[3] tagged strains with the mutations were mated with a Sho1-DHFR F[1,2]/*pbs2::LEU2* strain, in order to produce diploids with complementary DHFR fusions. To mate these strains, 50 µL of precultures of each strain were mixed into 900 µL of YPD and incubated at 30 °C overnight without agitation. 2 µL of each mated culture was spotted on solid YPD+Nat+Hyg, to select for diploids. Exponential phase precultures of diploids were diluted to 0.1 OD/mL in 80 µL of PCA medium, in a 384 well plate, in 4 replicates, at 30 °C. Cultures were grown either with and without 200 µg/mL of methotrexate, and with and without 1 M sorbitol. The optical density of each well was measured every 15 minutes in a Tecan Spark plate reader (Tecan). Growth curves were analyzed using a custom script written in R (R Core Team 2023), using the tidyverse package collection (Wickham *et al.* 2019) and the R package ggpubr 0.6.0 (Kassambara 2020). A function was written to find the maximal instantaneous growth rate for each well. Mutant P94R was disregarded, as growth of all replicates in the control condition in the absence of methotrexate was much lower than all other samples, suggesting a problem with strain construction.

Structure prediction

Structure predictions were done for Pbs2 mutants in complex with Sho1 using AlphaFold 2.3.2 and the AlphaFold-Multimer implementation (Jumper *et al.* 2021; Evans *et al.* 2022). Default options were used, and only the top scoring model for each mutant was relaxed. For Sho1, the

sequence of the entire protein was used for prediction (Uniprot entry P40073). For Pbs2, residues 71 to 126 were used for the prediction (Uniprot entry P08018). Predicted structures were visualized using ChimeraX 1.8 (Pettersen *et al.* 2021).

Sequence and structural alignments

The protein sequences of 24 yeast SH3-containing sequences were obtained from the alliancemine server (Bult and Sternberg 2023). The sequence of the SH3 domains plus 25 residues on each side was extracted from the complete protein sequences using the SH3 location information from Interpro (Paysan-Lafosse *et al.* 2023), and a custom R script. When an SH3 domain was within less than 25 residues of the N- or C-terminus of a protein, the sequence up to the terminus was used. The sequences of these SH3 domains and adjoining regions were aligned using mafft 7.526, using the L-INS_i (local alignment) algorithm with the command line : "/usr/bin/mafft" --localpair --maxiterate 16 --inputorder “<input_file>” > “<output_file>” (Katoh and Standley 2013). For the structural alignment, we used previously predicted structures from the AlphaFold Protein Structure Database (Varadi *et al.* 2022). Using the R packages AlphaMissenseR (Morgan and Nguyen) and bio3d (Grant *et al.* 2006), we downloaded the predicted structures of the 24 same yeast SH3-containing proteins from the AlphaFold Protein Structure Database, and trimmed the pdb files to only contains residues composing the SH3 domains as well as 25 residues on each side of the domain. To align the predicted sequences, we used the MUSTANG 3.2.4 tool, with the command line ./bin/mustang-3.2.3 -f ./path_to_files.txt -F fasta -o SH3_flanking -s off (Konagurthu *et al.* 2006). The ./path_to_files.txt file contains the paths of all trimmed pdb files.

DHFR-PCA on solid media against SH3 domain-containing proteins

The DHFR-PCA on solid media was based on previous work (Tarassov *et al.* 2008; Rochette *et al.* 2015). All colony manipulation was done using a robotic pin tool platform (BM5-SC1, S&P Robotics Inc.). The strains containing DHFR F[1,2]-tagged SH3-containing proteins and Pbs2-DHFR F[3] were taken from the Yeast Protein Interactome Collection (Horizon Discovery). The Pbs2-stuffed-DHFR F[3] strain is the same used for DMS library construction, and is described in the strain construction section. Cultures were started from frozen stocks in YPD with nourseothricin for DHFR F[1,2] strains and hygromycin B for DHFR F[3] strains, in deep welled 96-well plates.

Colonies were cherry picked from the 96-well plates, and arrayed onto YPD+Nat plates for the DHFR F[1,2] strains or YPD+Hyg plates for the DHFR F[3] strains, in a 384 colony array. The DHFR F[1,2] strains were organized into a randomized pattern to avoid any effects caused by neighboring colonies. Each 384 colony plate of DHFR F[3] strains contained only copies of the same variant. The 384 colony arrays were condensed into 1536 colony arrays on YPD+Nat or YPD+Hyg, with the randomized DHFR F[1,2] pattern repeated four times. The outer two rows and columns of the 1536 arrays were composed of LSM8-DHFR F[1,2] and CDC39-DHFR F[3], which grow well on PCA medium and serve to avoid any measurement bias from being on the edge of the array for the interactions of interest. The 1536 colony arrays were mated together by pinning the repeated DHFR F[1,2] array onto YPD plates, and then pinning one of the DHFR F[3] arrays

onto each plate. In this collection of arrays, each interaction between an SH3 containing protein and a Pbs2 variant was measured in 5 or 6 replicates. After 48 hours of mating at 30 °C, the mated colonies were pinned onto YPD+Nat+Hyg plates to select for diploid cells. Two growth cycles of 48 hours at 30 °C on YPD+Nat+Hyg were done. After diploid selection, the plates were photographed using an EOS Rebel T5i camera (Canon), to verify that all colonies grew correctly. The mated arrays were then pinned onto solid PCA medium plates, either with or without methotrexate, and with 1 M of sorbitol. The arrays were grown for two cycles of 48 hours at 30 °C in a custom growth and imaging platform (S&P Robotics Inc.), which incubated the plates and took a picture of each plate every 2 hours.

Analysis of DHFR-PCA of Pbs2 mutants binding to SH3 domain-containing proteins

First, the ImageMagick command line tool (ImageMagick Studio LLC 2023) was used to crop the images of the selection plates as well as change them to grayscale and invert the colors, using the commands: convert -colorspace LinearGray, convert -crop 4200X2800+530+320 and convert -negate. Colony sizes were then quantified using the Python package Pyphe (Kamrad *et al.* 2020), using the command pyphe-quantify timecourse --grid auto_1536 --t 1 --d 3 --s 0.05. The colony sizes were analyzed and visualized using a custom script written in R 4.3.1 (R Core Team 2023), based on a previous analysis (Dionne *et al.* 2021), and using R packages from the tidyverse collection (Wickham *et al.* 2019). Colony sizes were filtered, and colonies which had not grown at the end of the diploid selection were removed from the PCA results. Colony sizes were log2 transformed and then normalized to avoid bias from a position in the array or from the plate they grew on by subtracting the log2-transformed median colony size for each row, column and 1536 array, including the border colonies. Scores were rescaled, so that the median of the scores of the Sho1 and wild-type Pbs2 interaction was equivalent to 1 and the median of scores of the Sho1 and stuffed Pbs2 interaction was equivalent to 0.

Analysis of contacts in SH3-motif structures

Contacts were analyzed for 4 structures. Sho1-Pbs2 was generated by AlphaFold-Multimer, as described above. Abp1-Ark (PDB accession 2RPN), Nbp2-Ste20 (PDB accession 2LCS), and Bem1-2-Ste20 (PDB accession 2RQW) were all previously measured NMR structure ensembles, composed of the 20 individual structures. All 20 structures were used for determining the contacts, although only the designated representative structure was used for visualization. The structures were imported into ChimeraX 1.8 (Pettersen *et al.* 2021), and the ChimeraX “Contacts” tool was used to detect all pairs of atoms from different proteins with a Van der Waals radius overlap ≥ -0.40 Å. This was done for all atoms in the motif-bearing protein preceding position -3 in the standard motif numbering. So, in Pbs2, this was done for all atoms in positions 71 to 92 (SNQSEQDKGSSQSPKHIQQIVN), for Ark1 this was done for positions -10 to -4 (KPKLHSP), and for Ste20 this was done for positions -14 to -4 (SSSANGKFIPS). To identify equivalent positions in the four SH3 domains which form contacts, the structures of the 4 domains were aligned using the MUSTANG tool (Konagurthu *et al.* 2006), and a standard numbering was determined based on this alignment (Table S9). The contacts between atoms were filtered to simply indicate all contacts between residues, and are summarized in Table S10, using the

aforementioned standard numbering. The minimum distance between residues or proteins was calculated in a custom R script, using the bio3d package (Grant *et al.* 2006).

Data availability

Strains and plasmids are available upon request. All analysis scripts and visualization scripts are available on Github at the following address: https://github.com/Landrylab/Jordan_et_al_2024. All data, including demultiplexed sequencing data, results from the DHFR-PCA experiments, growth curve data, predicted structures, protein alignments, and all files and information necessary to run the scripts are available in the following Dryad repository: https://doi.org/10.5061/dryad.79cnp5j3z.

Acknowledgments

The authors thank Philippe C. Després, Moïra Dion and Dan Evans-Yamamoto for help with sequencing data analysis. We also thank all members of the LandryLab team for thoughtful comments on the experiments and on figure conception. All molecular graphics were performed using UCSF ChimeraX 1.8. The protein structure predictions were enabled by computing resources provided by the Digital Research Alliance of Canada.

Funding

This work was supported by Canadian Institutes of Health Research Foundation grant number 387697 and Human Frontier Science Program research grant RGP34/2018 to CRL. C.R.L. holds the Canada Research Chair in Cellular Systems and Synthetic Biology. D.J. was supported by the Canada Graduate Scholarship - Master's program from NSERC, a graduate scholarship from PROTEO, a graduate scholarship from the NSERC CREATE program EvoFunPath, a Postgraduate Scholarship - Doctoral from NSERC and a Doctoral scholarship from FRQS. D.B. was funded by the EMBO Long-Term Fellowship (LTF) (ALTF 1069-2019). U.D. was funded by an NSERC graduate fellowship.

Author contributions

All authors were involved in the design of the study. C.R.L. supervised this work and obtained funding. D.J. and A.K.D. performed the experiments. D.J. performed the computational analyses.

D.J. wrote the first draft of the manuscript. All authors contributed to revising and writing the final version of the manuscript.

Supplementary Figures

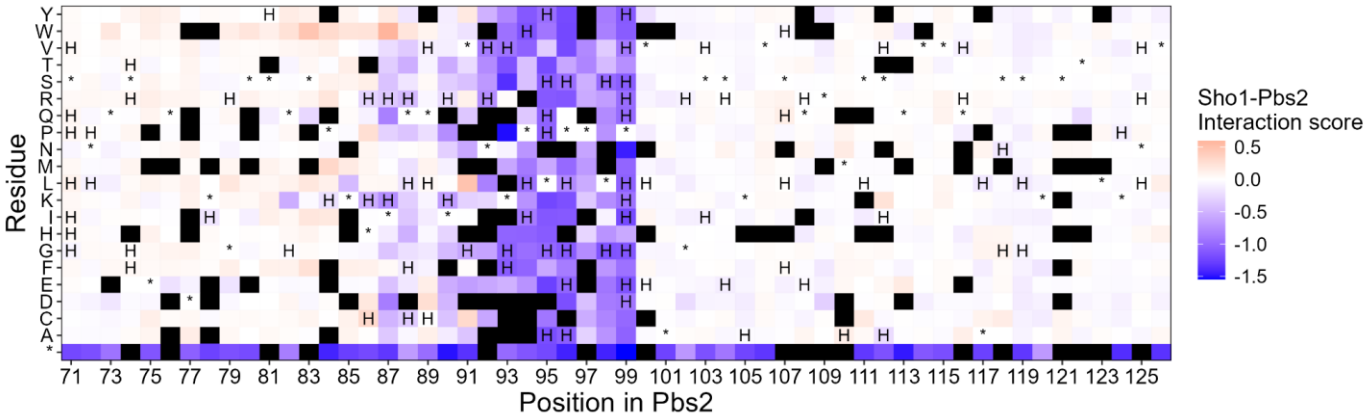


Figure S1.

Sho1-Pbs2 interaction scores for Pbs2 mutants in the DHFR-PCA on the DMS library of the region surrounding the Pbs2 motif. Scores are indicated for each mutant according to their position in Pbs2 and the residue that it was mutated to. The wild-type residue for each position is indicated by an asterisk (*). Mutants which significantly affect the Hog1-Pbs2 interaction are marked with the letter H (Two-sided Mann-Whitney U test with false discovery rate corrected p-value < 0.05). Mutants which were removed from the dataset for having too few reads in the initial timepoint are marked in black. Each measurement is the median of all codons in each replicate coding for the same residue, with 3 to 18 replicates for each mutant.

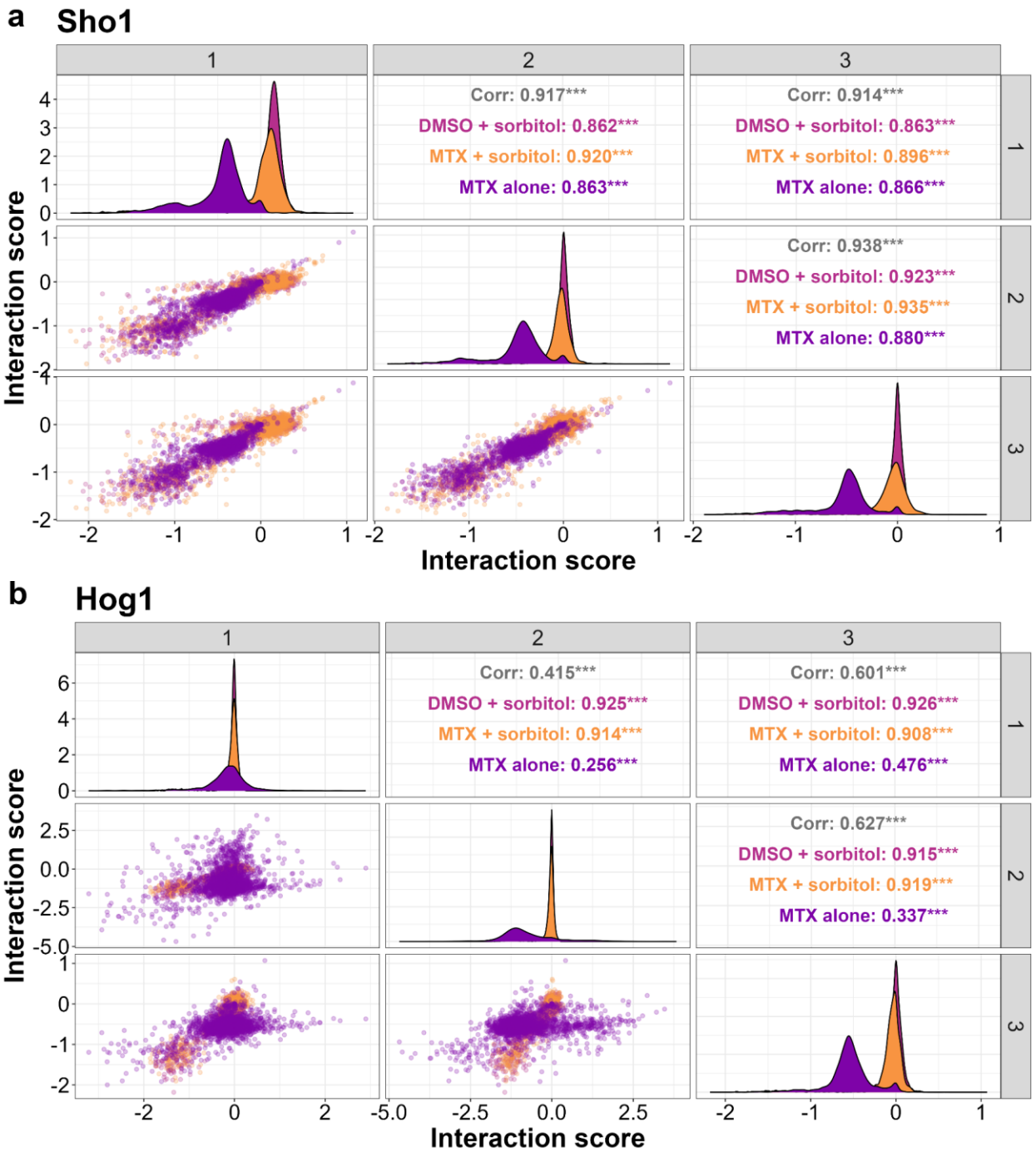


Figure S2.

Correlation between replicates in the DHFR-PCA experiments on DMS libraries surrounding region of the Pbs2 binding motif a) Scatterplot between the 3 replicates of the Sho1-Pbs2 DHFR-PCA interaction scores for mutants in the DMS library of the region surrounding the Pbs2 motif. Each point represents one Pbs2 codon variant in one condition (with or without sorbitol, with MTX or with only DMSO), as indicated by the color in the corresponding panel on the top right. Pearson correlations also indicated in the top right panels. In the panels on the diagonal, the distribution

of scores for a given condition and replicate is shown. The replicate number is indicated in the gray rectangle. b) Scatterplot between the 3 replicates of the Hog1-Pbs2 DHFR-PCA interaction scores for mutants in the DMS library of the region surrounding the Pbs2 motif. Each point represents one Pbs2 codon variant in one condition (with or without sorbitol, with MTX or with only DMSO), as indicated by the color in the corresponding panel on the top right. Pearson correlations also indicated in the top right panels. In the panels on the diagonal, the distribution of scores for a given condition and replicate is shown. The replicate number is indicated in the gray rectangle.

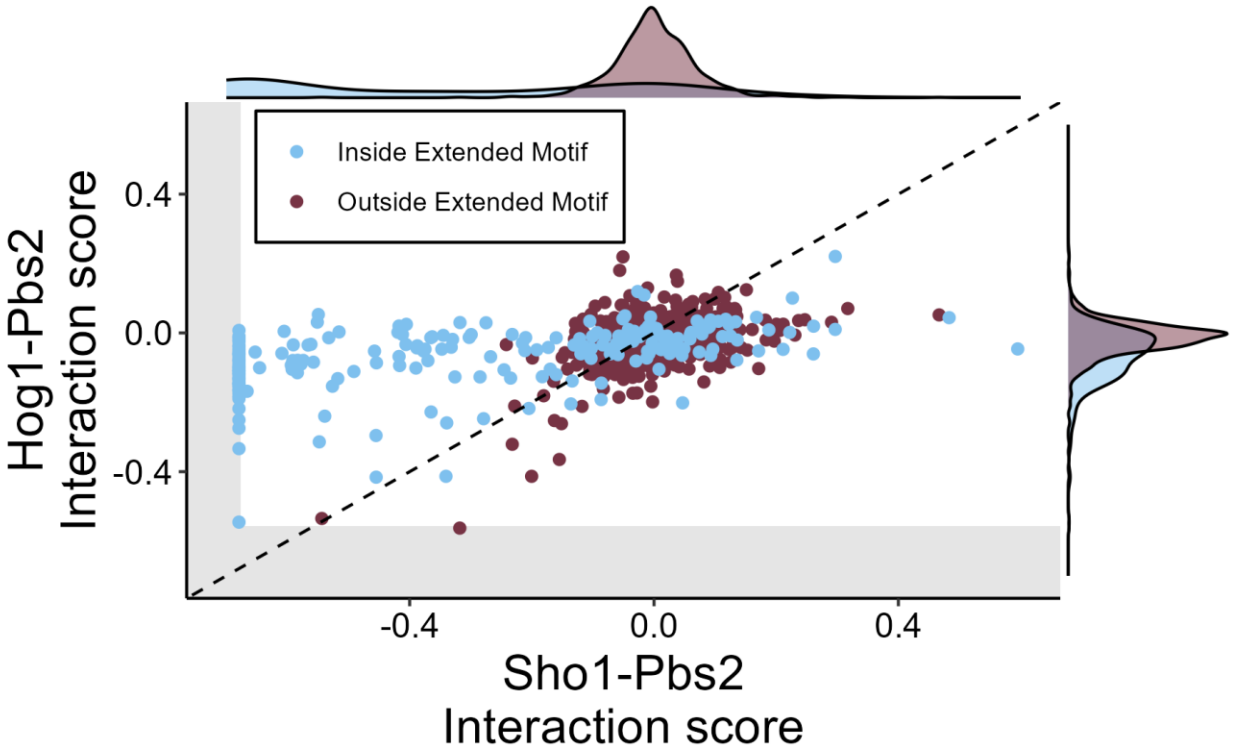


Figure S3.

Scatterplot of interaction scores for missense and silent mutants in the DMS library of the region surrounding the motif, for the interaction of Pbs2 with either Sho1 or Hog1. Points are colored according to whether the mutants are situated inside or outside the extended motif. The dotted line represents the diagonal. The gray rectangles on the left and bottom of the panel represent the 97.5 percentile of scores for the nonsense variants of Pbs2, which do not express the DHFR F[3] fragment. As such, they represent the limit of detectable signal in the assay. The points which have scores comparable to the nonsense mutations, and therefore no detectable interaction, were placed at the limit of the gray rectangle to indicate that no interaction was detected. In the margins, density plots show the distribution of scores for mutants situated either inside or outside the extended motif. Each point represents the median of all codons in each replicate coding for the same mutation, with 3 to 18 replicates for each mutant.

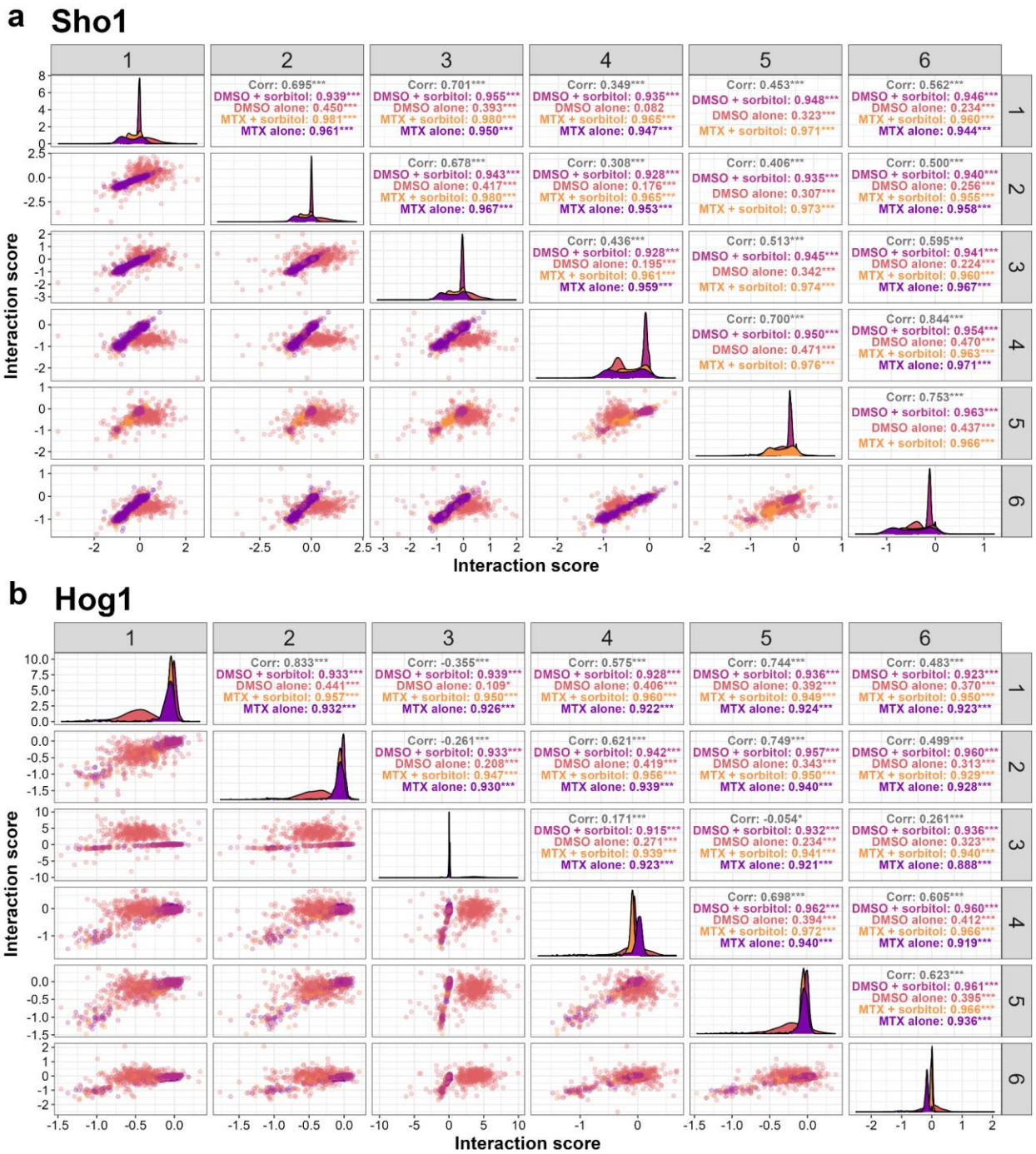


Figure S4.

<u>Correlation between replicates in the DHFR-PCA experiments on DMS libraries of the extended Pbs2 binding motif</u> a) Scatterplot between the 6 replicates of the Sho1-Pbs2 DHFR-PCA interaction scores for mutants in the DMS library of the extended Pbs2 motif. Each point represents one Pbs2 codon variant in one condition (with or without sorbitol, with MTX or with only DMSO), as indicated by the color in the corresponding panel on the top right. Pearson correlations also indicated in the top right panels. In the panels on the diagonal, the distribution of scores for a given condition and replicate is shown. The replicate number is indicated in the

gray rectangle. b) Scatterplot between the 6 replicates of the Hog1-Pbs2 DHFR-PCA interaction scores for mutants in the DMS library of the extended Pbs2 motif. Each point represents one Pbs2 codon variant in one condition (with or without sorbitol, with MTX or with only DMSO), as indicated by the color in the corresponding panel on the top right. Pearson correlations also indicated in the top right panels. In the panels on the diagonal, the distribution of scores for a given condition and replicate is shown. The replicate number is indicated in the gray rectangle.

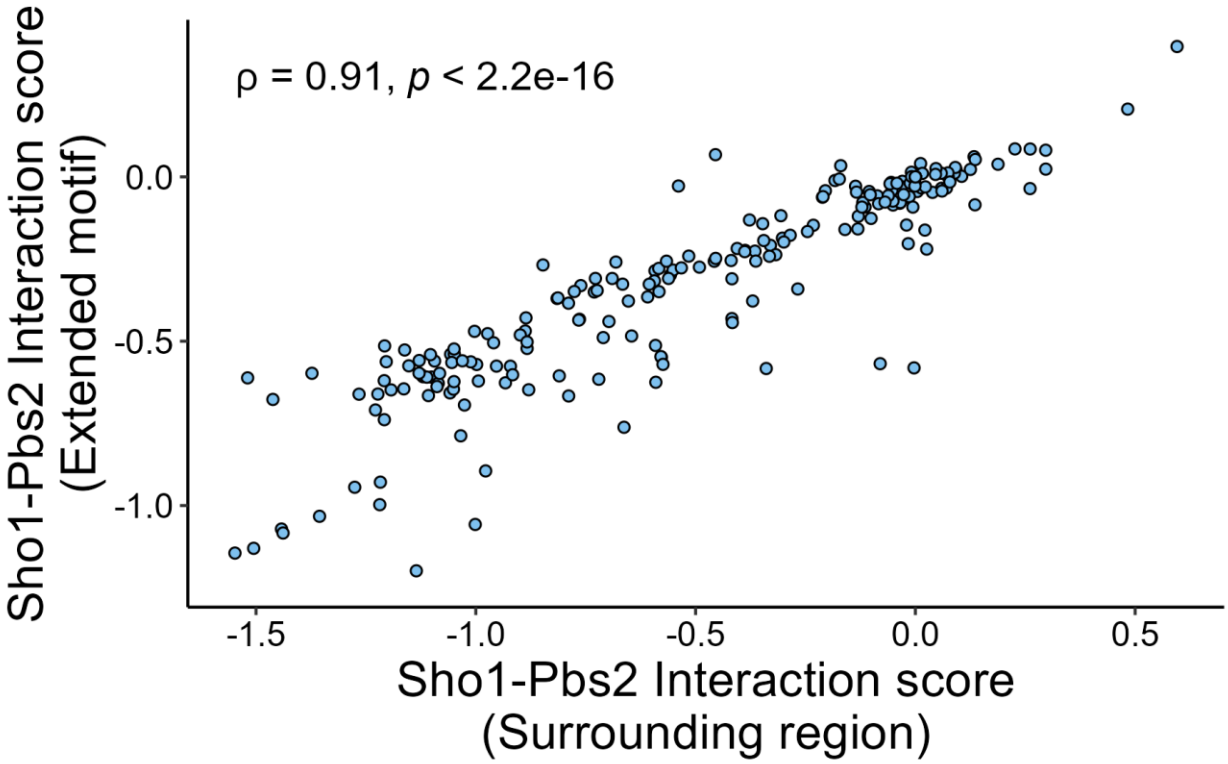


Figure S5.

Scatterplot of interaction scores of Pbs2 mutants in the initial DHFR-PCA screen on the DMS library of the surrounding region of the motif, and the subsequent DHFR-PCA screen on the DMS library of only the extended motif (positions 85 to 100). Scores shown were measured in the presence of methotrexate and 1 M of sorbitol. Each point represents the median of all codons in each replicate coding for the same mutation, with 3 to 18 replicates for each mutant. Spearman's rho 0.91, $p < 2.2X10^{-16}$.

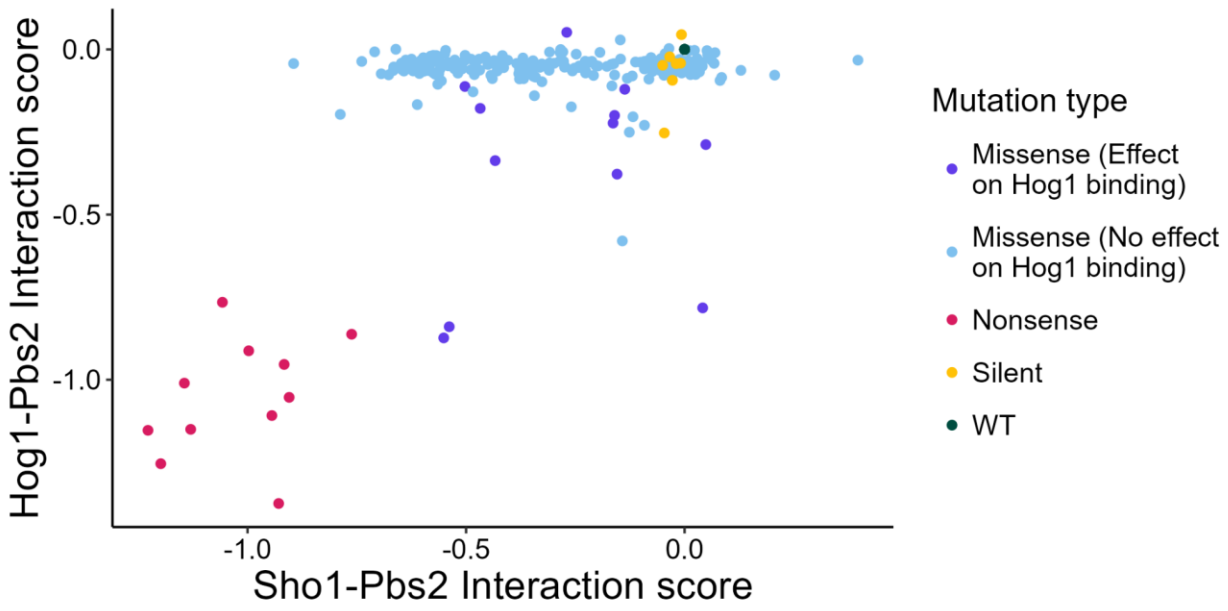


Figure S6.
Scatterplot of interaction scores of the Pbs2 extended motif DMS library mutants with Sho1 and Hog1, in the presence of 1 M of sorbitol. Purple missense points represent Pbs2 mutants which have an interaction score with Hog1 which is significantly different to the combination of silent mutants and wild-type Pbs2 (Two-sided Mann-Whitney U test with false discovery rate corrected p-value < 0.05), while blue missense points represent mutants for which the Hog1 interaction score is not significantly different to the combination of silent mutants and wild-type Pbs2. The combination of silent mutations and wild type represents all mutants expressing the wild-type protein sequence. Since the Hog1 interaction is used to measure abundance and stability of the Pbs2 mutants, the significantly different mutants are considered to have an effect on the abundance or the stability of Pbs2 and were not kept for the remainder of the analysis. The only exceptions are the nonsense mutants, which were kept because they provide a useful reference. Each point represents the median of all codons in each replicate coding for the same mutation, with 3 to 18 replicates for each mutant.

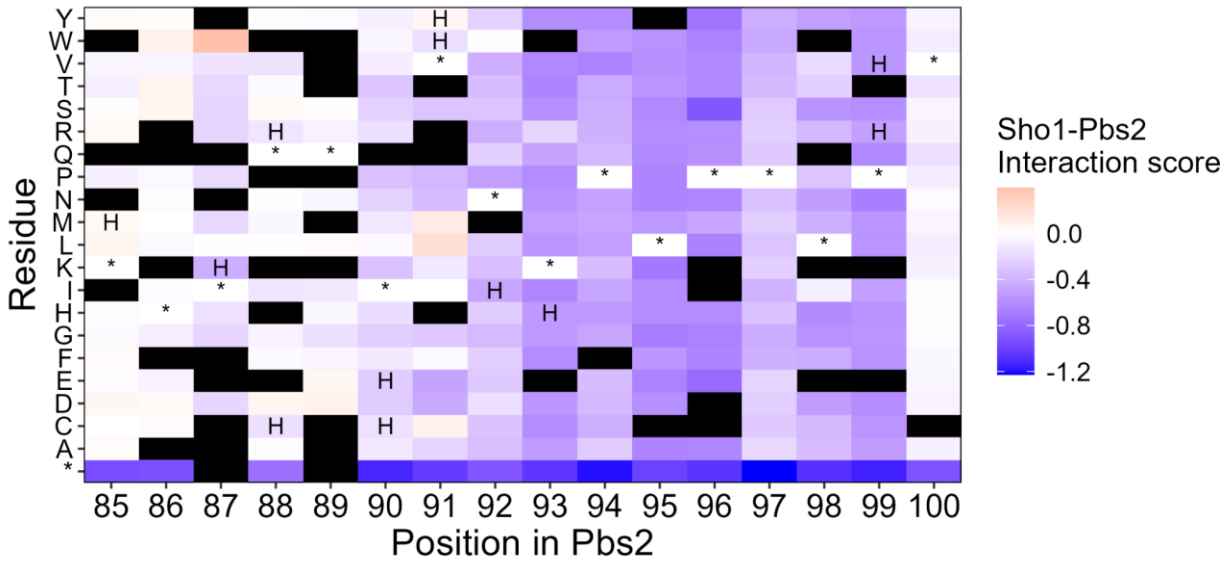


Figure S7.

Sho1-Pbs2 interaction scores for Pbs2 mutants in the DHFR-PCA on the DMS library of the extended Pbs2 motif. Scores are indicated for each mutant according to their position in Pbs2 and the residue that it was mutated to. The wild-type residue for each position is indicated by an asterisk (*). Mutants which significantly affect the Hog1-Pbs2 interaction are marked with the letter H (Two-sided Mann-Whitney U test with false discovery rate corrected p-value < 0.05). Mutants which were removed from the dataset for having too few reads in the initial timepoint are marked in black. Each measurement represents the median of all codons in each replicate coding for the same residue, with 3 to 18 replicates for each mutant.

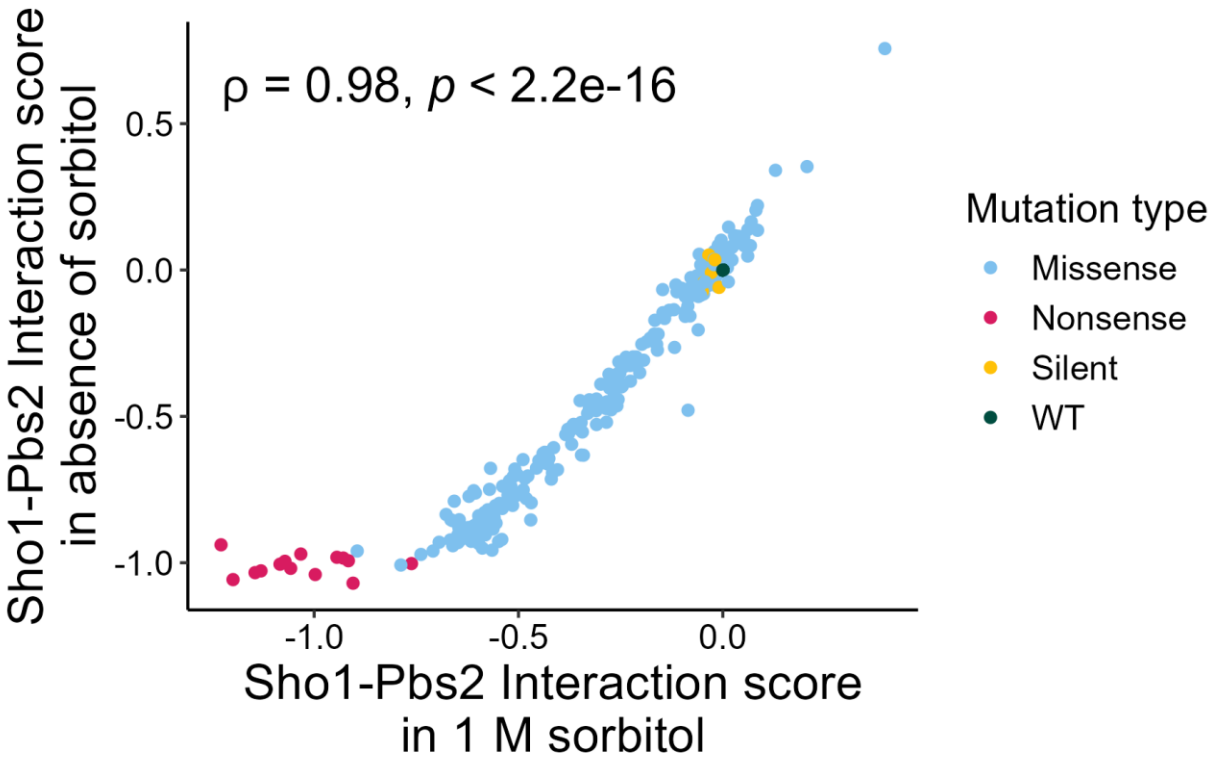


Figure S8.

Scatterplot of the Sho1 interaction scores for all mutants in the Pbs2 extended motif DMS library, in the presence and absence of 1 M of sorbitol. Measurements in DHFR-PCA media with methotrexate. The correlation has a spearman's rho of 0.98, and a p-value < $2.2X10^{-16}$. The measurements for the interaction in 1 M sorbitol have a greater range due to the greater activation of the HOG pathway. Each point represents the median of all codons in each replicate coding for the same mutation, with 3 to 18 replicates for each mutant.

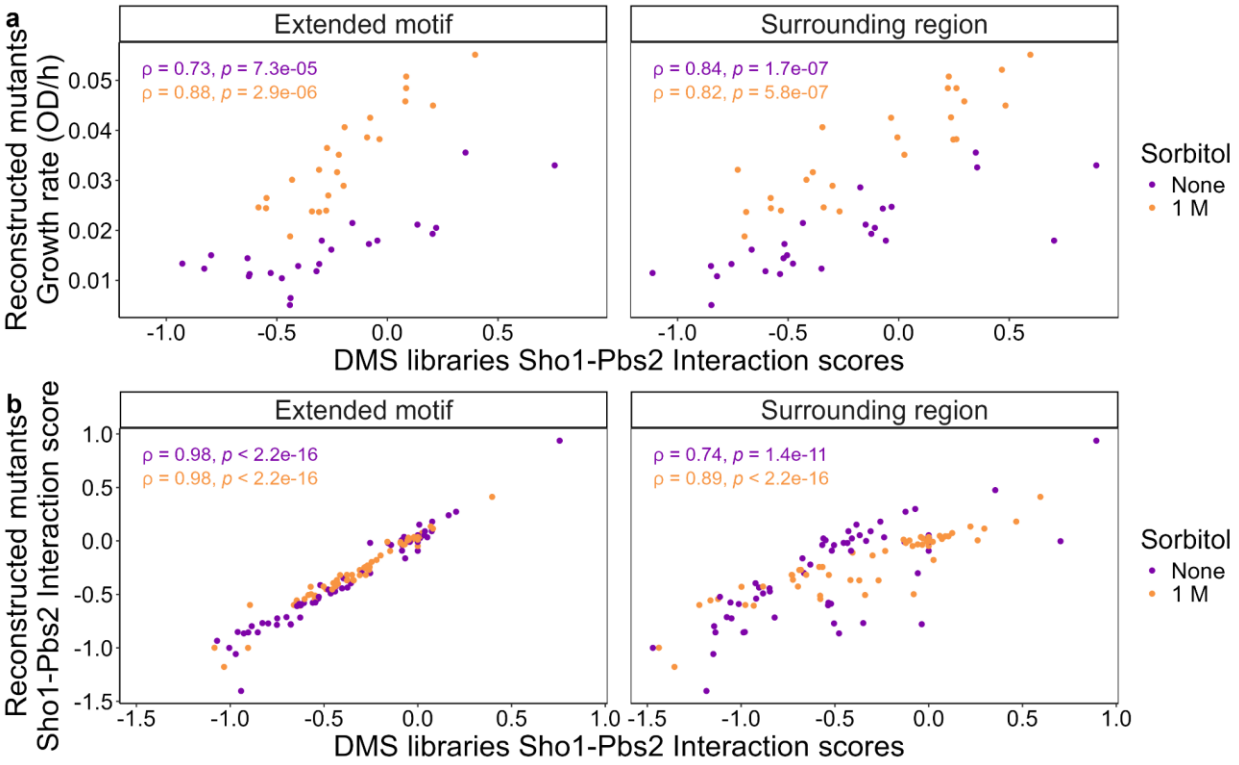


Figure S9.

<u>Validation of pooled DHFR-PCA effects by individually reconstructed mutants</u>

a) Scatterplot of Sho1-Pbs2 interaction scores as measured in pooled competition DHFR-PCA of the DMS libraries (x-axis) and growth rate of the same mutations individually reconstructed, in DHFR-PCA growth curves (y-axis). Results from the DMS library of the extended motif (left) and the DMS library of the surrounding region (right), in the presence (orange) or absence (purple) of 1 M of sorbitol. The spearman correlations between the two methods of measuring the interaction strength are indicated on the plots. The values for the DMS library interaction scores represent the median of all codons in each replicate coding for the same mutation, with 3 to 18 replicates for each mutant. The values for the growth rate represent the median of 3 replicates. b) Scatterplot of Sho1-Pbs2 interaction scores as measured in pooled competition DHFR-PCA of the DMS libraries (x-axis) and measured in pooled competition DHFR-PCA of individually reconstructed mutants (y-axis). Results from the DMS library of the extended motif (left) and the DMS library of the surrounding region (right), in the presence (orange) or absence (purple) of 1 M of sorbitol. The spearman correlations between the two methods of measuring the interaction strength are indicated on the plots. The values for the DMS library interaction scores represent the median of all codons in each replicate coding for the same mutation, with 3 to 18 replicates for each mutant. The values for the reconstructed mutants represent the median of 3 to 6 replicates.

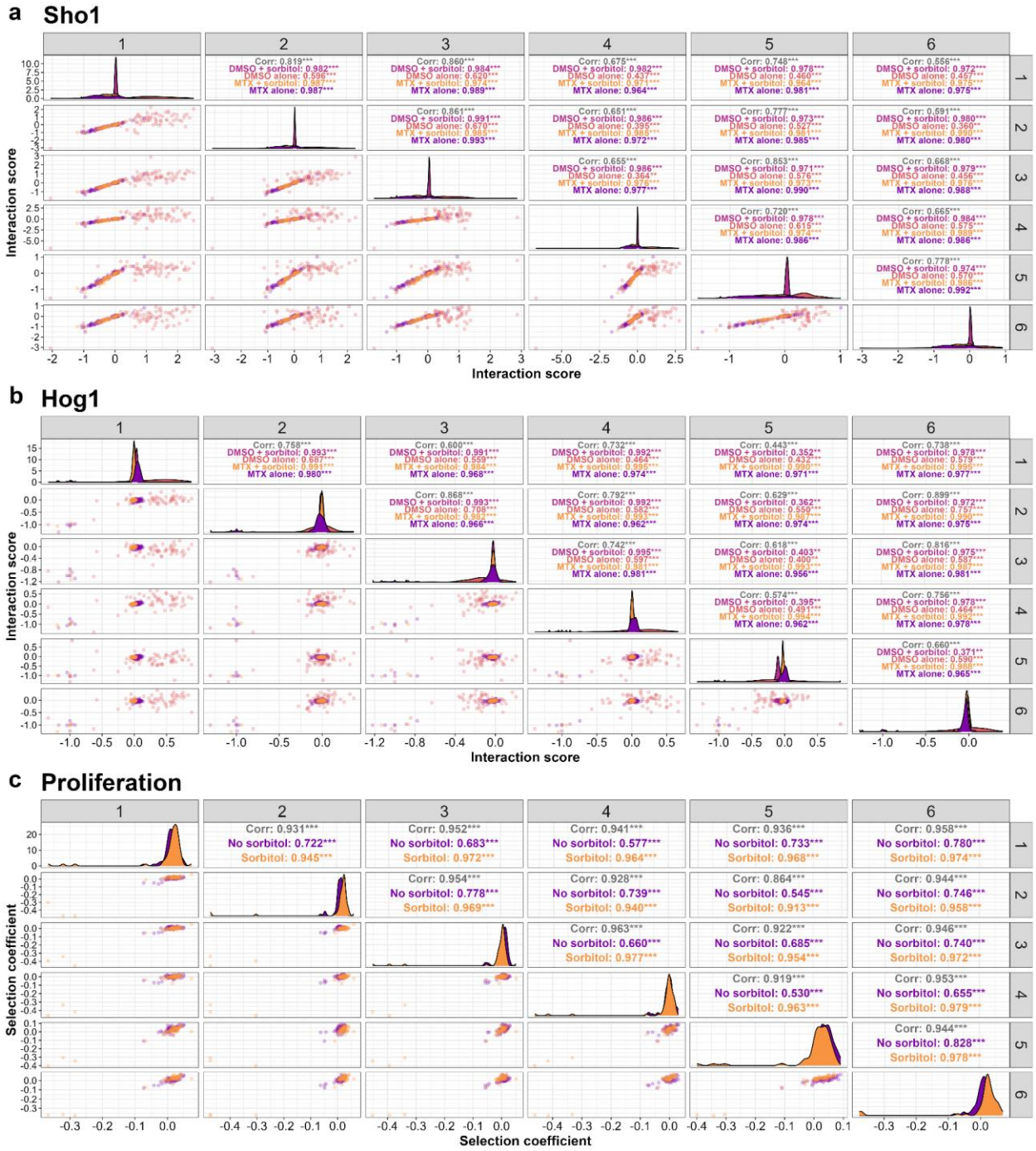


Figure S10.

Correlation between replicates in the DHFR-PCA experiments on DMS libraries of the extended Pbs2 binding motif

a) Scatterplot between the 6 replicates of the Sho1-Pbs2 DHFR-PCA interaction scores for mutants in the Pbs2 validation DMS library. Each point represents one Pbs2 codon variant in one condition (with or without sorbitol, with MTX or with only DMSO), as indicated by the color in the corresponding panel on the top right. Pearson correlations also indicated in the top right panels.

In the panels on the diagonal, the distribution of scores for a given condition and replicate is shown. The replicate number is indicated in the gray rectangle. b) Scatterplot between the 6 replicates of the Hog1-Pbs2 DHFR-PCA interaction scores for mutants in the Pbs2 validation DMS library. Each point represents one Pbs2 codon variant in one condition (with or without sorbitol, with MTX or with only DMSO), as indicated by the color in the corresponding panel on the top right. Pearson correlations also indicated in the top right panels. In the panels on the diagonal, the distribution of scores for a given condition and replicate is shown. The replicate number is indicated in the gray rectangle. c) Scatterplot between the 6 replicates of selection coefficients for mutants in the Pbs2 validation DMS library measuring cell proliferation. Each point represents one Pbs2 codon variant in one condition (with or without sorbitol), as indicated by the color in the corresponding panel on the top right. Pearson correlations also indicated in the top right panels. In the panels on the diagonal, the distribution of scores for a given condition and replicate is shown. The replicate number is indicated in the gray rectangle.

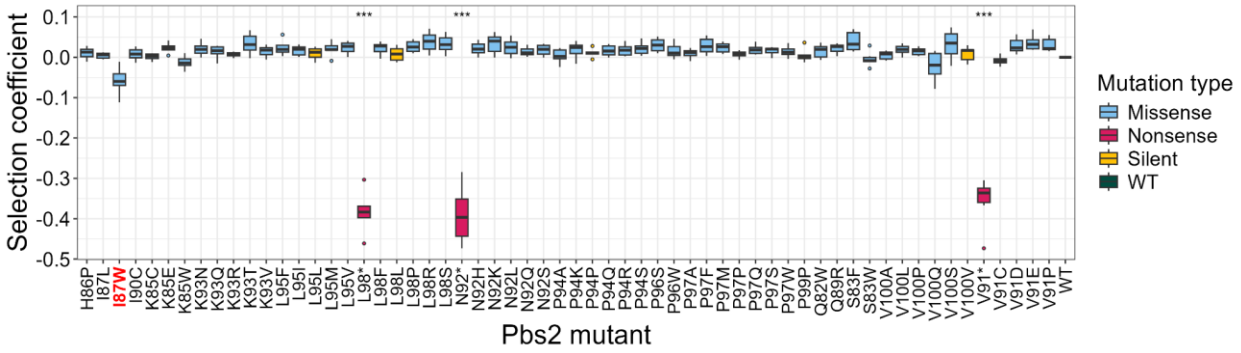


Figure S11.

Impact on growth rate measured as selection coefficients for individually reconstructed Pbs2 mutants in SC synthetic media with 1 M sorbitol. The strains contain no DHFR fragments on any proteins. Boxplots show the distribution of 6 replicates per mutant. The selection coefficient is normalized by the growth of the wild-type sequence, making the wild-type selection coefficient zero. Only the 3 nonsense mutants grew significantly less than wild-type Pbs2 and the silent mutants (Welch's t-test with false discovery rate correction p-value < 0.05), although I87W (highlighted in red on axis), which is the mutant with the highest Sho1-Pbs2 interaction score shows a marked decrease in selection coefficient (p-value = 0.066).

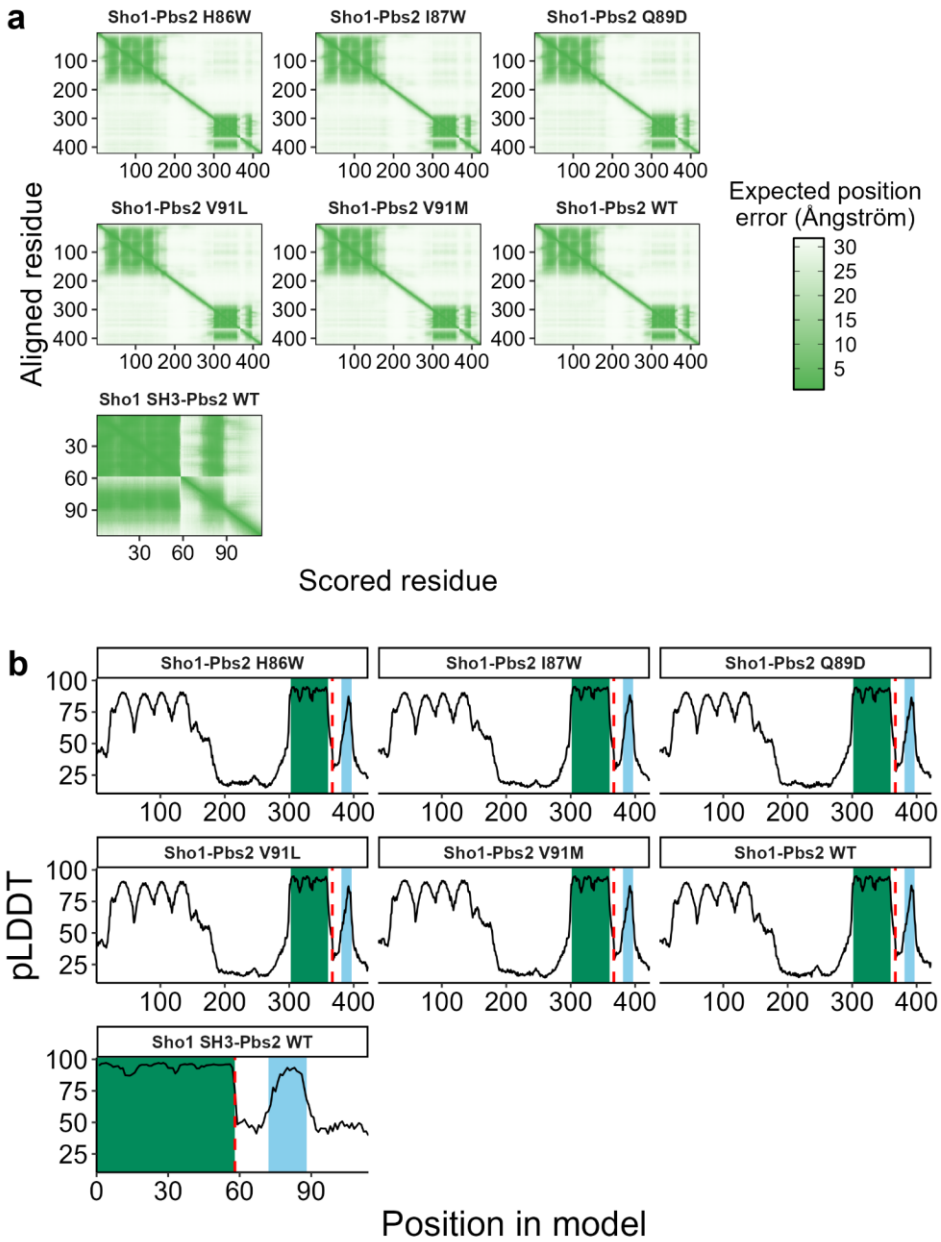


Figure S12.

Confidence metrics for the seven structures predicted using AlphaFold2-Multimer. Sho1 SH3-Pbs2 WT is the structure in figure 2b, Sho1-Pbs2 WT is the structure in figure 3a and 3b, Sho1-Pbs2 I87W is the structure in figure 3c, and all others are the structures with the same Pbs2 mutants in figure S13. a) Predicted aligned error (PAE) values for the 7 predicted structures, output from AlphaFold-Multimer. b) pLDDT value for each position in the 7 predicted structures, output from AlphaFold-Multimer. Positions corresponding to the SH3 domain of Sho1 are indicated by a green background, and positions corresponding to the extended motif of Pbs2 are indicated by a blue background. Since AlphaFold-Multimer combines the sequences of both proteins into a single model, a red dashed line was used to mark the end of the Sho1 portion of the model, and the beginning of the Pbs2 portion of the model.

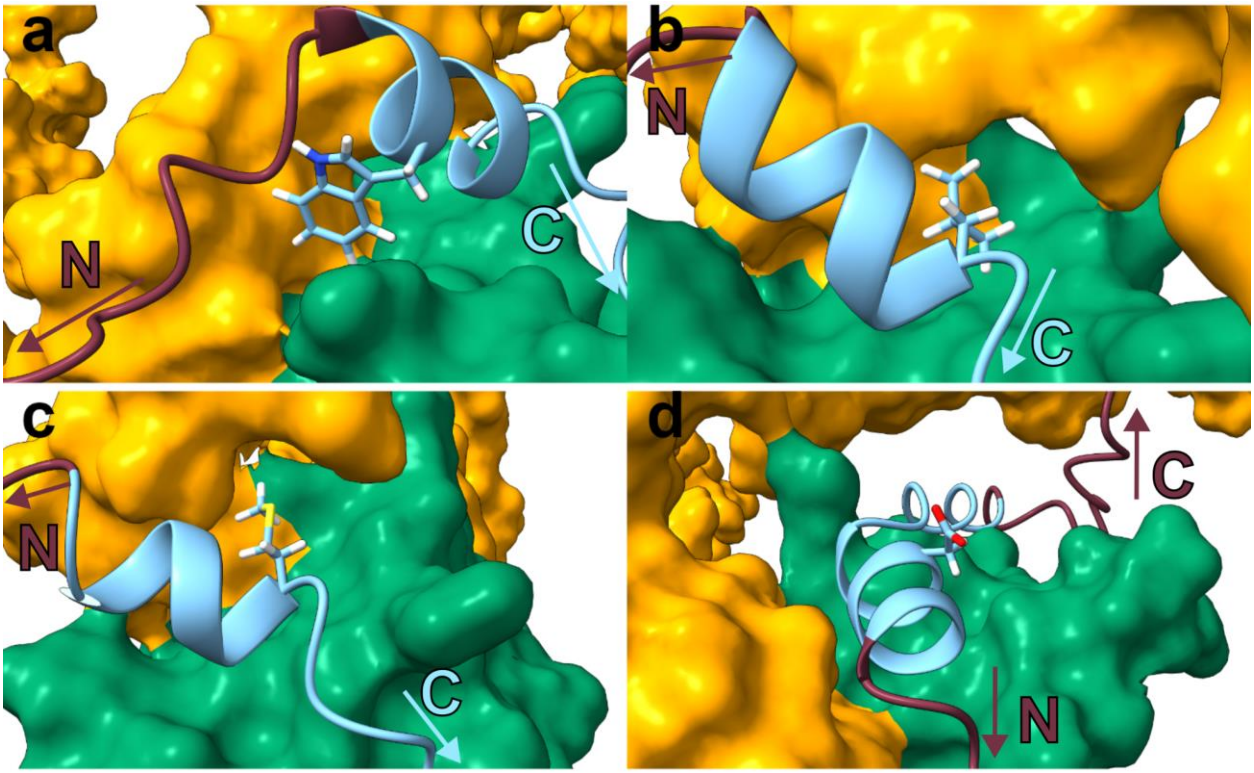
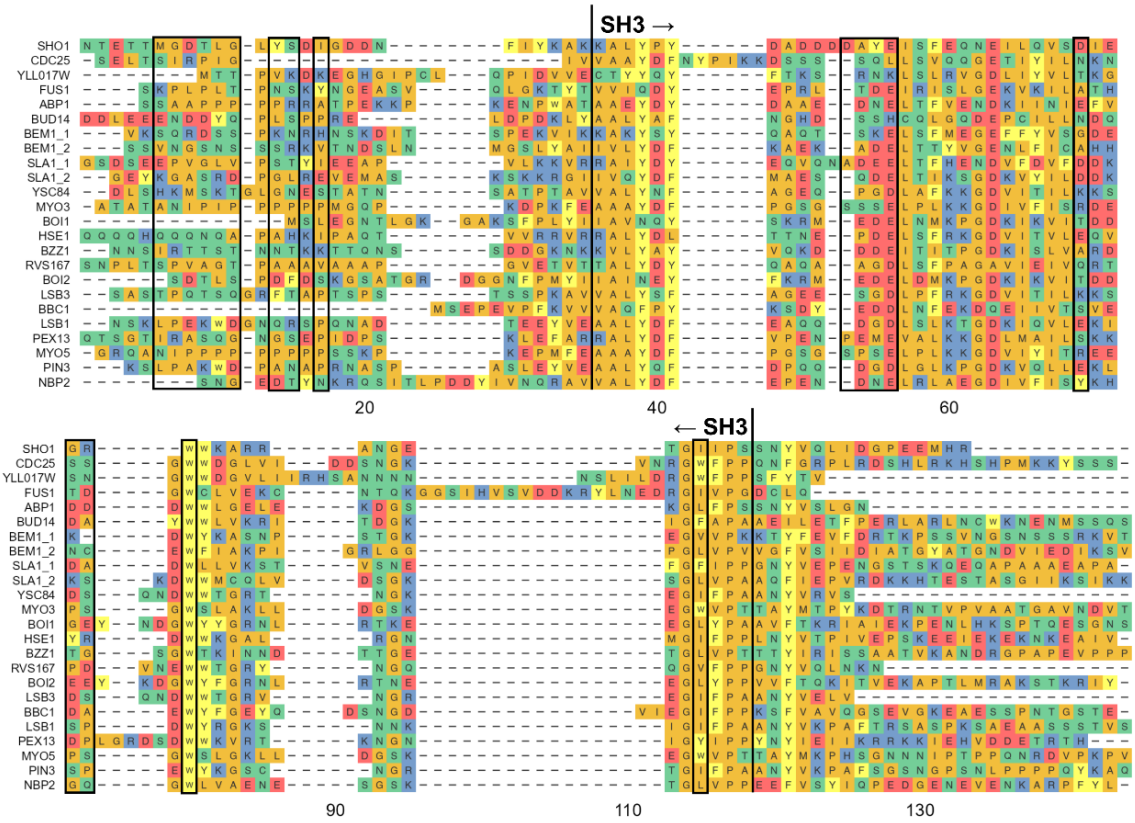


Figure S13.

Close-up details of AlphaFold-Multimer predicted complexes of Sho1 with Pbs2 mutants a) H86W, b) V91L, c) V91M and d) Q89D. The predicted structures are colored as in panel a of figure 3, with the SH3 domain of Sho1 in green, the non-SH3 portions of Sho1 in orange, the extended motif of Pbs2 in light blue and the rest of Pbs2 in burgundy. The side chain of the mutated residue is shown with oxygen atoms in red, nitrogen atoms in dark blue, sulfur atoms in yellow, hydrogen atoms in white and carbon atoms the same color as the extended motif, while all other side chains in Pbs2 are hidden. The orientation of Pbs2 is shown by annotations indicating which end is nearer the N-terminus (N) and which end is nearer the C-terminus (C).

a


SH3 →
SHO1
CDC25
YLL017W
FUS1
ABP1
BUD14
BEM1_1
BEM1_2
SLA1_1
SLA1_2
YSC84
MYO3
BOI1
HSE1
BZZ1
RVS167
BOI2
LSB3
BBC1
LSB1
PEX13
MYO5
PIN3
NBP2
20
40
60
← SH3
90
110
130

b

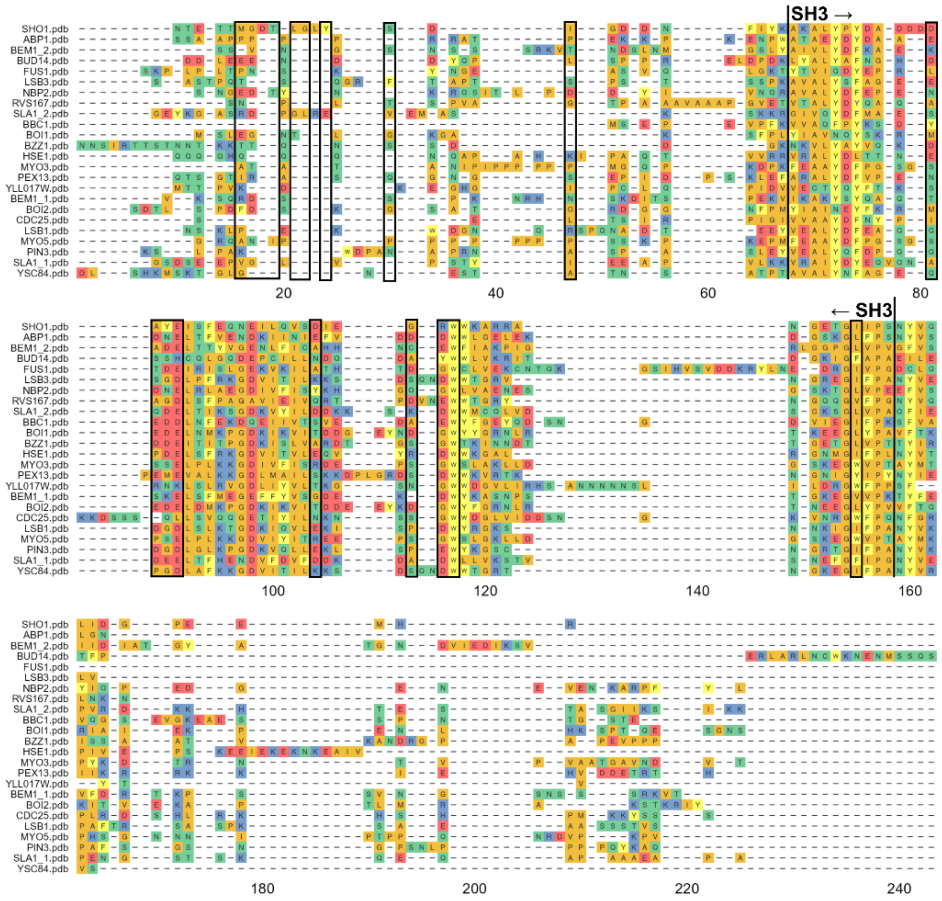

SH3 →
SHO1.pdb
ABP1.pdb
BEM1_2.pdb
BUD14.pdb
FUS1.pdb
LSB3.pdb
NBP2.pdb
RVS167.pdb
SLA1_2.pdb
BBC1.pdb
BOI1.pdb
BZZ1.pdb
HSE1.pdb
MYO3.pdb
PEX13.pdb
YLL017W.pdb
BEM1_1.pdb
BOI2.pdb
CDC25.pdb
LSB1.pdb
MYO5.pdb
PIN3.pdb
SLA1_1.pdb
YSC84.pdb
20
40
60
80
← SH3
100
120
140
160
180
200
220
240

Figure S14.

Sequence and structural alignments of yeast SH3 domains and adjacent sequences a) Multiple sequence alignment of yeast SH3 domains with 25 residues on each side. Residues which in the predicted Sho1-Pbs2 structure are within 5 Å of Pbs2 residues situated in the extended motif but not the canonical motif are surrounded by a black rectangle. Beginning and end of the SH3 domain of Sho1 marked above the alignment. b) Multiple structural alignment of predicted yeast SH3 domains with 25 residues on each side, from MUSTANG. Residues which in the predicted Sho1-Pbs2 structure are within 5 Å of Pbs2 residues situated in the extended motif but not the canonical motif are surrounded by a black rectangle. Beginning and end of the SH3 domain of Sho1 marked above the alignment.

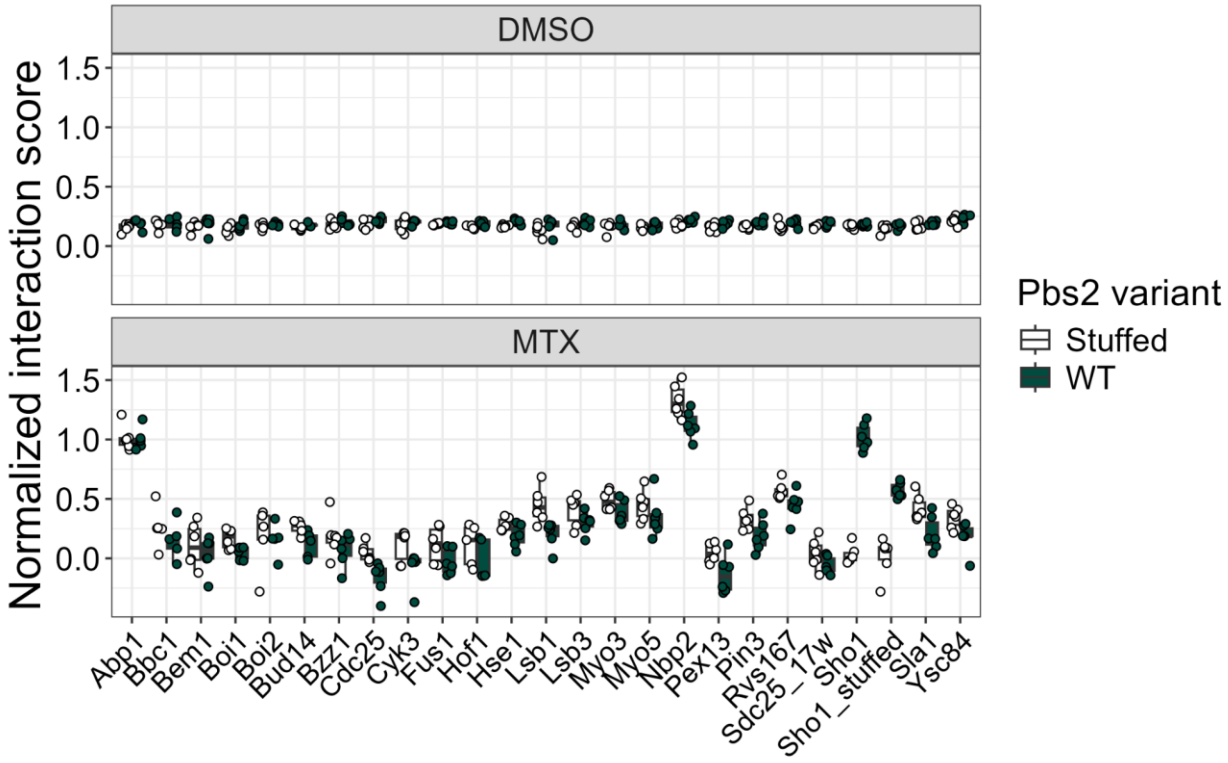


Figure S15.

Interaction score of wild-type Pbs2 (green) or Pbs2 where positions 76 to 126 have been replaced with a flexible stuffer sequence (Stuffed, white) with all yeast SH3-containing proteins. Controls grown on DMSO and not MTX, and therefore not measuring the interaction, are shown in the top panel. All replicates shown as points. Scores normalized with 1 being the median of Sho1 - wild type Pbs2 interactions and 0 being the median of Sho1 - stuffed Pbs2 interactions. A control where the SH3 domain of Sho1 was replaced by the same neutral, flexible stuffer sequence used for Pbs2-stuffed was also included, with the name Sho1_stuffed.

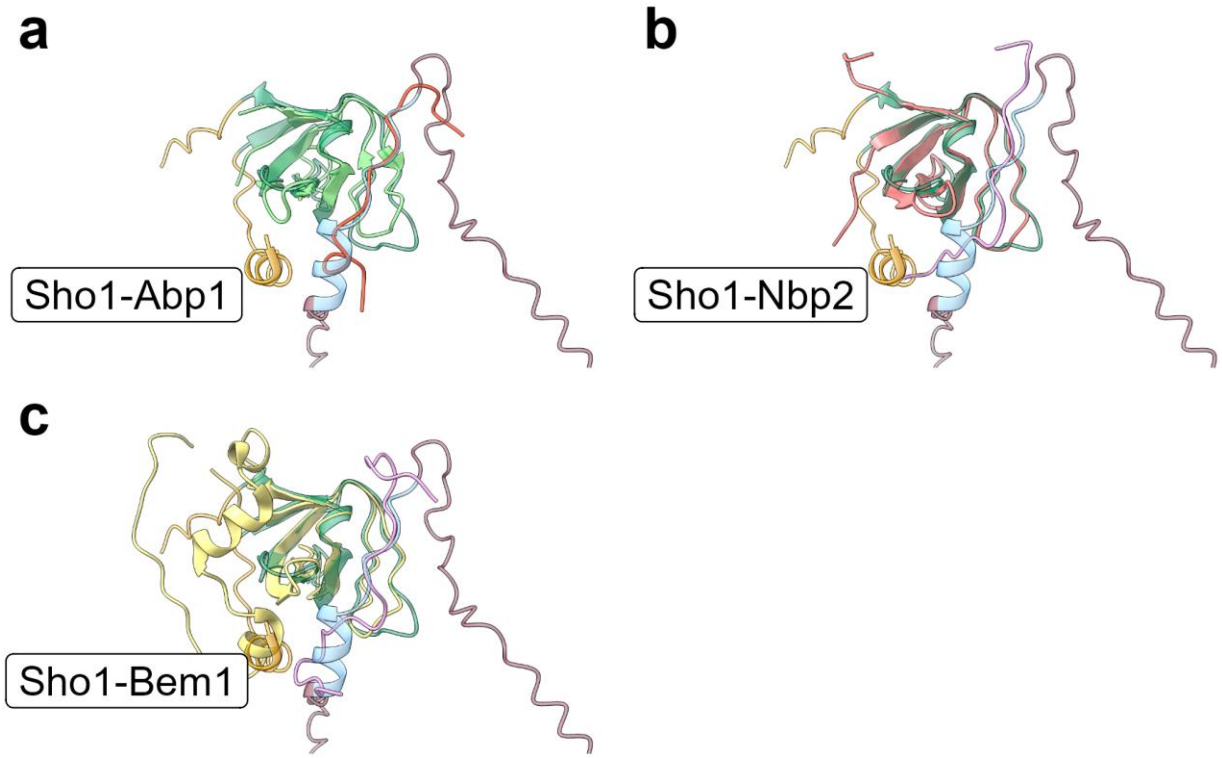
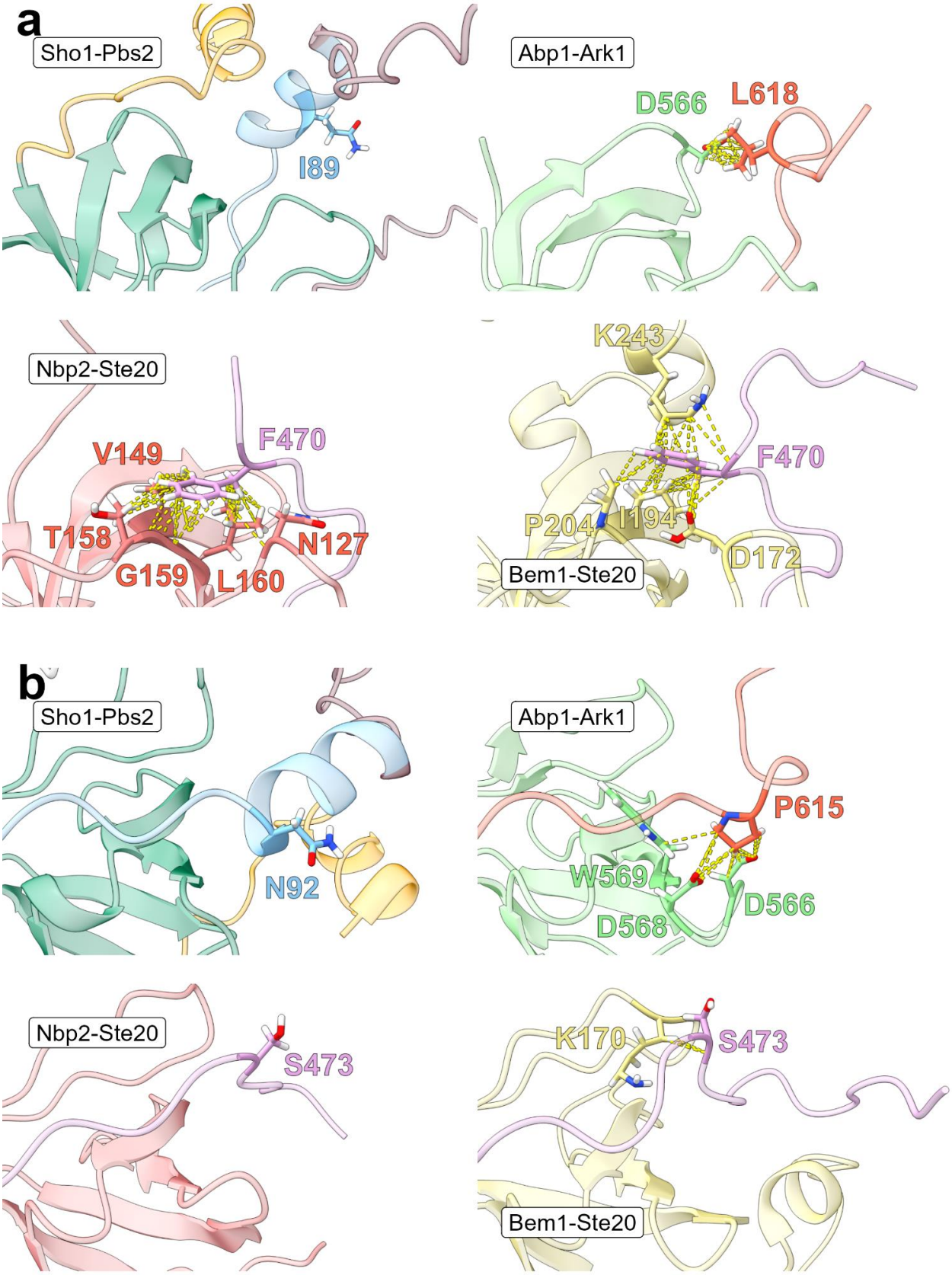


Figure S16.

Predicted structure of the Sho1-Pbs2 interaction superimposed with other SH3-motif pairs. The Sho1-Pbs2 structure is semi-transparent and colored as in figure 3a, with the SH3 domain of Sho1 in green, the non-SH3 portions of Sho1 in orange, the Pbs2 extended motif in light blue and the rest of Pbs2 in burgundy. Sho1-Pbs2 structure was superimposed with the a) Abp1-Ark structure (Abp1 - light green, Ark1 - pink), b) Nbp2-Ste20 structure (Nbp2 - orange-pink, Ste20 - light pink), and c) Bem1-2-Ste20 structure (Bem1 - yellow, Ste20 - light pink). Structures were superimposed using the Chimera X tool “Matchmaker” (Pettersen *et al.* 2021) to minimize root mean square deviance of backbone atom distances of the SH3 domains with Sho1.


a
Sho1-Pbs2
I89
Abp1-Ark1
D566
L618
Nbp2-Ste20
V149
F470
T158
G159
L160
N127
K243
F470
P204
I194
D172
Bem1-Ste20
b
Sho1-Pbs2
N92
Abp1-Ark1
P615
W569
D566
D568
Nbp2-Ste20
S473
K170
S473
Bem1-Ste20

1382

Figure S17.
Additional differing contacts between SH3-extended motif pairs
a) Visualization of the side-chains and contacts of position (-7) in the Sho1-Pbs2 predicted structure , Abp1-Ark1 structure (PDB accession 2RPN), Nbp2-Ste20 structure (PDB accession 2LCS) and Bem1-Ste20 (PDB accession 2RQW). Sho1-Pbs2 predicted structure colored as in figure 3a, with the SH3 domain of Sho1 in green, the non-SH3 portions of Sho1 in orange, the Pbs2 extended motif in light blue and the rest of Pbs2 in burgundy. All four structures were superimposed using the Chimera X tool “Matchmaker” (Pettersen *et al.* 2021) to minimize root mean square deviance of backbone atom distances of the SH3 domains with Sho1, then pictures were captured separately. b) Visualization of the side-chains and contacts of position (-4) of the same structures as in panel a). All four structures were superimposed using the Chimera X tool “Matchmaker” (Pettersen *et al.* 2021) to minimize root mean square deviance of backbone atom distances of the SH3 domains with Sho1, then pictures were captured separately.

Literature cited

Amberg D., D. C. Burke, and J. N. Strathern, 2005 *Methods in yeast genetics: a Cold Spring Harbor Laboratory course manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Andrews S., 2010 FastQC: a quality control tool for high throughput sequence data

Attali D., and C. Baker, 2023 *ggExtra: Add Marginal Histograms to “ggplot2”, and More “ggplot2” Enhancements*.

Bendel A. M., A. J. Faure, D. Klein, K. Shimada, R. Lyautey, *et al.*, 2024 The genetic architecture of protein interaction affinity and specificity. Nat. Commun. 15: 8868.

Bolger A. M., M. Lohse, and B. Usadel, 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114–2120.

Bult C. J., and P. W. Sternberg, 2023 The alliance of genome resources: transforming comparative genomics. Mamm. Genome 34: 531–544.

Dalgarno D. C., M. C. Botfield, and R. J. Rickles, 1997 SH3 domains and drug design: Ligands, structure, and biological function. Pept. Sci. 43: 383–400.

Davey N. E., G. Travé, and T. J. Gibson, 2011 How viruses hijack cell regulation. Trends Biochem. Sci. 36: 159–169.

Delplace V., A. Ortin-Martinez, E. L. S. Tsai, A. N. Amin, V. Wallace, *et al.*, 2019 Controlled release strategy designed for intravitreal protein delivery to the retina. J. Control. Release 293: 10–20.

Després P. C., A. F. Cisneros, E. M. M. Alexander, R. Sonigara, C. Gagné-Thivierge, *et al.*, 2022 Asymmetrical dose responses shape the evolutionary trade-off between antifungal resistance and nutrient use. Nat Ecol Evol 6: 1501–1515.

Dibyachintan S., A. K. Dube, D. Bradley, P. Lemieux, U. Dionne, *et al.*, 2024 Cryptic genetic variation shapes the fate of gene duplicates in a protein interaction network. bioRxiv 2024.02.23.581840.

Dionne U., É. Bourgault, A. K. Dubé, D. Bradley, F. J. M. Chartier, *et al.*, 2021 Protein context shapes the specificity of SH3 domain-mediated interactions in vivo. Nat. Commun. 12: 1597.

Dionne U., L. J. Percival, F. J. M. Chartier, C. R. Landry, and N. Bisson, 2022 SRC homology 3 domains: multifaceted binding modules. Trends Biochem. Sci. 47: 772–784.

Diss G., and B. Lehner, 2018 The genetic landscape of a physical interaction. Elife 7. https://doi.org/10.7554/eLife.32472

Douangamath A., F. V. Filipp, A. T. J. Klein, P. Barnett, P. Zou, *et al.*, 2002 Topography for Independent Binding of α-Helical and PPII-Helical Ligands to a Peroxisomal SH3 Domain.

Mol. Cell 10: 1007–1017.

Dubé A. K., R. Dandage, S. Dibyachintan, U. Dionne, P. C. Després, *et al.*, 2022 Deep Mutational Scanning of Protein–Protein Interactions Between Partners Expressed from Their Endogenous Loci In Vivo, pp. 237–259 in *Yeast Functional Genomics: Methods and Protocols*, edited by Devaux F. Springer US, New York, NY.

Ernst A., D. Gfeller, Z. Kan, S. Seshagiri, P. M. Kim, *et al.*, 2010 Coevolution of PDZ domain-ligand interactions analyzed by high-throughput phage display and deep sequencing. Mol. Biosyst. 6: 1782–1790.

Evans R., M. O'Neill, A. Pritzel, N. Antropova, A. Senior, *et al.*, 2022 Protein complex prediction with AlphaFold-Multimer. bioRxiv 2021.10.04.463034.

Faure A. J., J. Domingo, J. M. Schmiedel, C. Hidalgo-Carcedo, G. Diss, *et al.*, 2022 Mapping the energetic and allosteric landscapes of protein binding domains. Nature 604: 175–183.

Feng S., C. Kasahara, R. J. Rickles, and S. L. Schreiber, 1995 Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. Proc. Natl. Acad. Sci. U. S. A. 92: 12408–12415.

Ferrigno P., F. Posas, D. Koepp, H. Saito, and P. A. Silver, 1998 Regulated nucleo/cytoplasmic exchange of HOG1 MAPK requires the importin beta homologs NMD5 and XPO1. EMBO J. 17: 5606–5614.

Fowler D. M., and S. Fields, 2014 Deep mutational scanning: a new style of protein science. Nat. Methods 11: 801–807.

Freschi L., F. Torres-Quiroz, A. K. Dubé, and C. R. Landry, 2013 qPCA: a scalable assay to measure the perturbation of protein-protein interactions in living cells. Mol. Biosyst. 9: 36–

43.

Gaussmann S., R. Peschel, J. Ott, K. M. Zak, J. Sastre, *et al.*, 2024 Modulation of peroxisomal import by the PEX13 SH3 domain and a proximal FxxxF binding motif. Nat. Commun. 15: 3317.

Gorelik M., and A. R. Davidson, 2012 Distinct peptide binding specificities of Src homology 3 (SH3) protein domains can be determined by modulation of local energetics across the binding interface. J. Biol. Chem. 287: 9168–9177.

Gouw M., S. Michael, H. Sámano-Sánchez, M. Kumar, A. Zeke, *et al.*, 2018 The eukaryotic linear motif resource – 2018 update. Nucleic Acids Res. 46: D428–D434.

Grant B. J., A. P. C. Rodrigues, K. M. ElSawy, J. A. McCammon, and L. S. D. Caves, 2006 Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics 22: 2695–2696.

Gueldener U., J. Heinisch, G. J. Koehler, D. Voss, and J. H. Hegemann, 2002 A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. Nucleic Acids Res. 30: e23.

Harris C. R., K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, *et al.*, 2020 Array programming with NumPy. Nature 585: 357–362.

Haslam N. J., and D. C. Shields, 2012 Peptide-Binding Domains: Are Limp Handshakes Safest? Sci. Signal. 5: e40–pe40.

Hunter J. D., 2007 Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. 9: 90–95.

Illumina, Nextera DNA Indexes. Illumina Support Docs.

ImageMagick Studio LLC, 2023 *ImageMagick*.

Ivarsson Y., and P. Jemth, 2019 Affinity and specificity of motif-based protein-protein interactions. Curr. Opin. Struct. Biol. 54: 26–33.

Jumper J., R. Evans, A. Pritzel, T. Green, M. Figurnov, *et al.*, 2021 Highly accurate protein structure prediction with AlphaFold. Nature 596: 583–589.

Kamrad S., M. Rodríguez-López, C. Cotobal, C. Correia-Melo, M. Ralser, *et al.*, 2020 Pyphe, a python toolbox for assessing microbial growth and cell viability in high-throughput colony screens. Elife 9. https://doi.org/10.7554/eLife.55160

Kaneko T., L. Li, and S. S.-C. Li, 2008 The SH3 domain--a family of versatile peptide- and protein-recognition module. Front. Biosci. 13: 4938–4952.

Kaneko T., H. Huang, X. Cao, X. Li, C. Li, *et al.*, 2012 Superbinder SH2 Domains Act as Antagonists of Cell Signaling. Sci. Signal. 5: ra68–ra68.

Karlsson O. A., G. N. Sundell, E. Andersson, Y. Ivarsson, and P. Jemth, 2016 Improved affinity at the cost of decreased specificity: a recurring theme in PDZ-peptide interactions. Sci. Rep. 6: 34269.

Kassambara A., 2020 ggpubr: “ggplot2” Based Publication Ready Plots

Kassambara A., 2021 rstatix: Pipe-Friendly Framework for Basic Statistical Tests

Katoh K., and D. M. Standley, 2013 MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30: 772–780.

Kazlauskas A., C. Schmotz, T. Kesti, J. Hepojoki, I. Kleino, *et al.*, 2016 Large-Scale Screening of Preferred Interactions of Human Src Homology-3 (SH3) Domains Using Native Target

Proteins as Affinity Ligands. Mol. Cell. Proteomics 15: 3270–3281.

Keck F., 2020 Handling biological sequences in R with the bioseq package. Methods Ecol. Evol. 11: 1728–1732.

Kelil A., E. D. Levy, and S. W. Michnick, 2016 Evolution of domain–peptide interactions to coadapt specificity and affinity to functional diversity. Proc. Natl. Acad. Sci. U. S. A. 113: E3862–E3871.

Kelil A., B. Dubreuil, E. D. Levy, and S. W. Michnick, 2017 Exhaustive search of linear information encoding protein-peptide recognition. PLoS Comput. Biol. 13: e1005499.

Kijima Y., D. Evans-Yamamoto, H. Toyoshima, and N. Yachie, 2023 A universal sequencing read interpreter. Sci Adv 9: eadd2793.

Konagurthu A. S., J. C. Whisstock, P. J. Stuckey, and A. M. Lesk, 2006 MUSTANG: a multiple structural alignment algorithm. Proteins 64: 559–574.

Kumar M., S. Michael, J. Alvarado-Valverde, A. Zeke, T. Lazar, *et al.*, 2024 ELM-the Eukaryotic Linear Motif resource-2024 update. Nucleic Acids Res. 52: D442–D455.

Kursula P., I. Kursula, Y. H. Song, K. Paraskevopoulos, and M. Wilmanns, 2008 YEAST SHO1 SH3 DOMAIN COMPLEXED WITH A PEPTIDE FROM PBS2

Langmead B., C. Trapnell, M. Pop, and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10: R25.

Lee C. H., K. Saksela, U. A. Mirza, B. T. Chait, and J. Kuriyan, 1996 Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain. Cell 85: 931–942.

Lemieux P., D. Bradley, A. K. Dubé, U. Dionne, and C. R. Landry, 2024 Dissection of the role of

a Src homology 3 domain in the evolution of binding preference of paralogous proteins. Genetics 226. https://doi.org/10.1093/genetics/iyad175

Li S. S.-C., 2005 Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. Biochem. J 390: 641–653.

Liao T.-J., H. Jang, R. Nussinov, and D. Fushman, 2020 High-affinity interactions of the nSH3/cSH3 domains of Grb2 with the C-terminal proline-rich domain of SOS1. J. Am. Chem. Soc. 142: 3401–3411.

Lim W. A., F. M. Richards, and R. O. Fox, 1994 Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. Nature 372: 375–379.

Maeda T., M. Takekawa, and H. Saito, 1995 Activation of Yeast PBS2 MAPKK by MAPKKKs or by Binding of an SH3-Containing Osmosensor. Science 269: 554–558.

Mapes J., and I. M. Ota, 2004 Nbp2 targets the Ptc1-type 2C Ser/Thr phosphatase to the HOG MAPK pathway. EMBO J. 23: 302–311.

Marles J. A., S. Dahesh, J. Haynes, B. J. Andrews, and A. R. Davidson, 2004 Protein-Protein Interaction Affinity Plays a Crucial Role in Controlling the Sho1p-Mediated Signal Transduction Pathway in Yeast. Mol. Cell 14: 813–823.

Martin M., 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17: 10–12.

Masella A. P., A. K. Bartram, J. M. Truszkowski, D. G. Brown, and J. D. Neufeld, 2012 PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics 13: 31.

Messing J., 1983 [2] New M13 vectors for cloning, pp. 20–78 in *Methods in Enzymology*, Academic Press.

Michnick S. W., E. D. Levy, C. R. Landry, J. Kowarzyk, and V. Messier, 2016 The Dihydrofolate Reductase Protein-Fragment Complementation Assay: A Survival-Selection Assay for Large-Scale Analysis of Protein–Protein Interactions. Cold Spring Harb. Protoc. 2016: db.prot090027.

Miyazaki K., 2011 Chapter seventeen - MEGAWHOP Cloning: A Method of Creating Random Mutagenesis Libraries via Megaprimer PCR of Whole Plasmids, pp. 399–406 in *Methods in Enzymology*, Synthetic Biology, Part B. edited by Voigt C. Academic Press.

Monera O. D., T. J. Sereda, N. E. Zhou, C. M. Kay, and R. S. Hodges, 1995 Relationship of sidechain hydrophobicity and alpha-helical propensity on the stability of the single-stranded amphipathic alpha-helix. J. Pept. Sci. 1: 319–329.

Morgan M., and T. Nguyen, *AlphaMissenseR*.

Murakami Y., K. Tatebayashi, and H. Saito, 2008 Two adjacent docking sites in the yeast Hog1 mitogen-activated protein (MAP) kinase differentially interact with the Pbs2 MAP kinase kinase and the Ptp2 protein tyrosine phosphatase. Mol. Cell. Biol. 28: 2481–2494.

Pawson T., M. Raina, and P. Nash, 2002 Interaction domains: from simple binding events to complex cellular behavior. FEBS Lett. 513: 2–10.

Paysan-Lafosse T., M. Blum, S. Chuguransky, T. Grego, B. L. Pinto, *et al.*, 2023 InterPro in 2022. Nucleic Acids Res. 51: D418–D427.

Pettersen E. F., T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, *et al.*, 2021 UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. 30: 70–82.

Posas F., and H. Saito, 1997 Osmotic Activation of the HOG MAPK Pathway via Ste11p

MAPKKK: Scaffold Role of Pbs2p MAPKK. Science 276: 1702–1705.

R Core Team, 2023 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rice P., I. Longden, and A. Bleasby, 2000 EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 16: 276–277.

Rickles R. J., M. C. Botfield, X. M. Zhou, P. A. Henry, J. S. Brugge, *et al.*, 1995 Phage display selection of ligand residues important for Src homology 3 domain binding specificity. Proc. Natl. Acad. Sci. U. S. A. 92: 10909–10913.

Robles J. T., H. J. Lou, G. Shi, P. L. Pan, and B. E. Turk, 2023 Linear motif specificity in signaling through p38α and ERK2 mitogen–activated protein kinases. Proceedings of the National Academy of Sciences 120: e2316599120.

Rochette S., G. Diss, M. Filteau, J.-B. Leducq, A. K. Dubé, *et al.*, 2015 Genome-wide protein-protein interaction screening by protein-fragment complementation assay (PCA) in living cells. J. Vis. Exp. https://doi.org/10.3791/52255

Rognes T., T. Flouri, B. Nichols, C. Quince, and F. Mahé, 2016 VSEARCH: a versatile open source tool for metagenomics. PeerJ 4: e2584.

Ryan O. W., S. Poddar, and J. H. D. Cate, 2016 CRISPR-Cas9 Genome Engineering in Saccharomyces cerevisiae Cells. Cold Spring Harb. Protoc. 2016. https://doi.org/10.1101/pdb.prot086827

Saito H., and F. Posas, 2012 Response to Hyperosmotic Stress. Genetics 192: 289–318.

Schloerke B., D. Cook, J. Larmarange, F. Briatte, M. Marbach, *et al.*, 2023 *GGally: Extension to "ggplot2."*

Slowikowski K., 2023 *ggrepel: Automatically Position Non-Overlapping Text Labels with "ggplot2."*

Stein A., and P. Aloy, 2008 Contextual specificity in peptide-mediated protein interactions. PLoS One 3: e2524.

Stollar E. J., B. Garcia, P. A. Chong, A. Rath, H. Lin, *et al.*, 2009 Structural, functional, and bioinformatic studies demonstrate the crucial role of an extended peptide binding site for the SH3 domain of yeast Abp1p. J. Biol. Chem. 284: 26918–26927.

Takaku T., K. Ogura, H. Kumeta, N. Yoshida, and F. Inagaki, 2010 Solution structure of a novel Cdc42 binding module of Bem1 and its interaction with Ste20 and Cdc42. J. Biol. Chem. 285: 19346–19353.

Tarassov K., V. Messier, C. R. Landry, S. Radinovic, M. M. S. Molina, *et al.*, 2008 An in Vivo Map of the Yeast Protein Interactome. Science 320: 1465–1470.

Tatebayashi K., K. Yamamoto, K. Tanaka, T. Tomida, T. Maruoka, *et al.*, 2006 Adaptor functions of Cdc42, Ste50, and Sho1 in the yeast osmoregulatory HOG MAPK pathway. EMBO J. 25: 3033–3044.

The pandas development team, 2023 *pandas-dev/pandas: Pandas*.

Tonikian R., X. Xin, C. P. Toret, D. Gfeller, C. Landgraf, *et al.*, 2009 Bayesian Modeling of the Yeast SH3 Domain Interactome Predicts Spatiotemporal Dynamics of Endocytosis Proteins. PLoS Biol. 7: e1000218.

Van Roey K., B. Uyar, R. J. Weatheritt, H. Dinkel, M. Seiler, *et al.*, 2014 Short Linear Motifs: Ubiquitous and Functionally Diverse Protein Interaction Modules Directing Cell Regulation. Chem. Rev. 114: 6733–6778.

Varadi M., S. Anyango, M. Deshpande, S. Nair, C. Natassia, *et al.*, 2022 AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 50: D439–D444.

Vincentelli R., K. Luck, J. Poirson, J. Polanowska, J. Abdat, *et al.*, 2015 Quantifying domain-ligand affinities and specificities by high-throughput holdup assay. Nat. Methods 12: 787–793.

Wang C.-H., P. Mehta, and C. J. Bashor, 2018 The strength of protein-protein interactions controls the information capacity and dynamical response of signaling networks. bioRxiv 469197.

Wickham H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, *et al.*, 2019 Welcome to the tidyverse. Journal of Open Source Software 4: 1686.

Wilke C. O., 2020 cowplot: Streamlined Plot Theme and Plot Annotations for “ggplot2”

Yu H., J. K. Chen, S. Feng, D. C. Dalgarno, A. W. Brauer, *et al.*, 1994 Structural basis for the binding of proline-rich peptides to SH3 domains. Cell 76: 933–945.

Zarin T., and B. Lehner, 2024 A complete map of specificity encoding for a partially fuzzy protein interaction. bioRxiv 2024.04.25.591103.

Zarrinpar A., S.-H. Park, and W. A. Lim, 2003 Optimization of specificity in a cellular protein interaction network by negative selection. Nature 426: 676–680.

Zarrinpar A., R. P. Bhattacharyya, M. P. Nittler, and W. A. Lim, 2004 Sho1 and Pbs2 Act as Coscaffolds Linking Components in the Yeast High Osmolarity MAP Kinase Pathway. Mol. Cell 14: 825–832.

1629