Supplemental Information

Structurally unique interaction of RBD-like and PH domains is crucial for yeast pheromone signaling

Volodymyr Yerko[*,†], Traian Sulea[*], Irena Ekiel[*,‡], Doreen Harcus[*], Jason Baardsnes[*], Miroslaw Cygler[*,¶,#], Malcolm Whiteway[*,§], Cunle Wu[*,‖,@]

Supplemental Figures

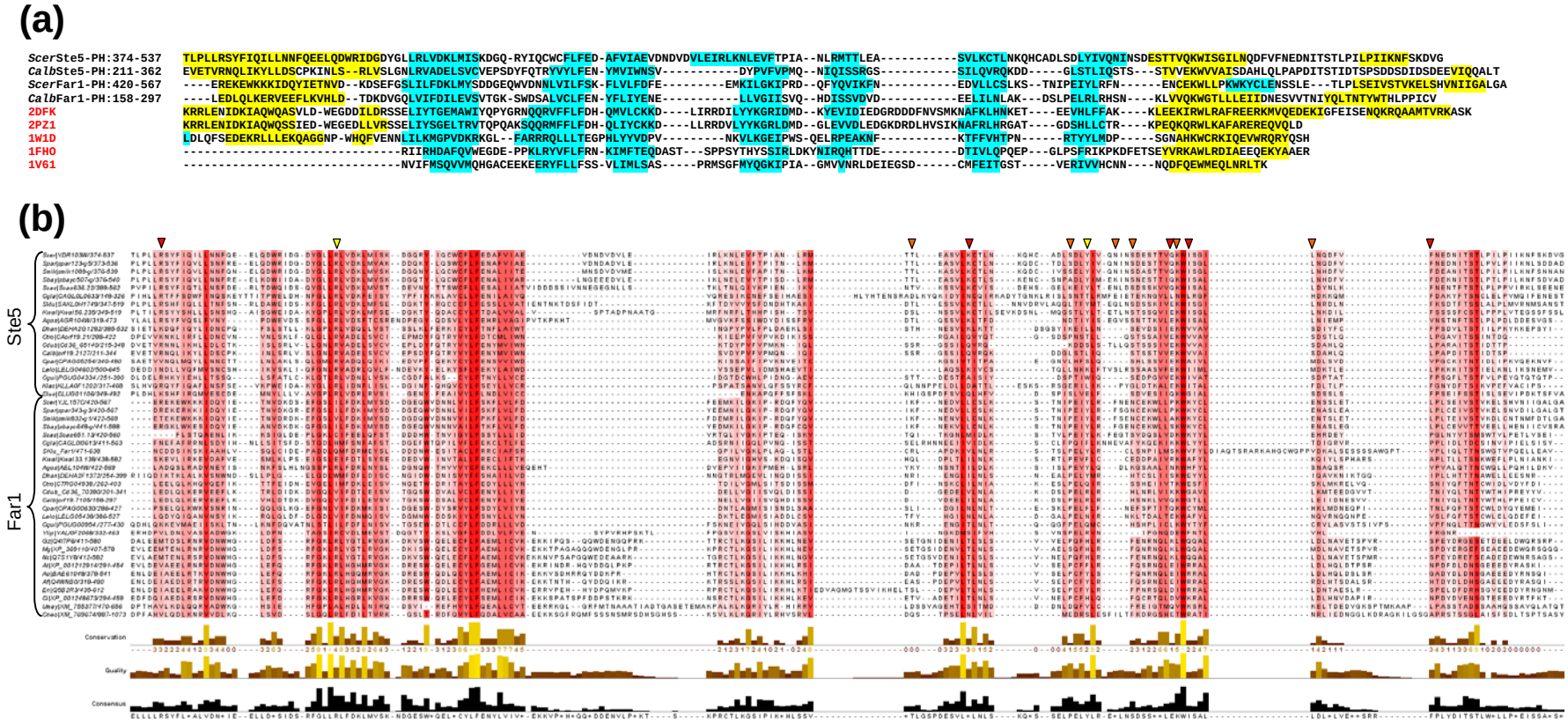


Figure S1

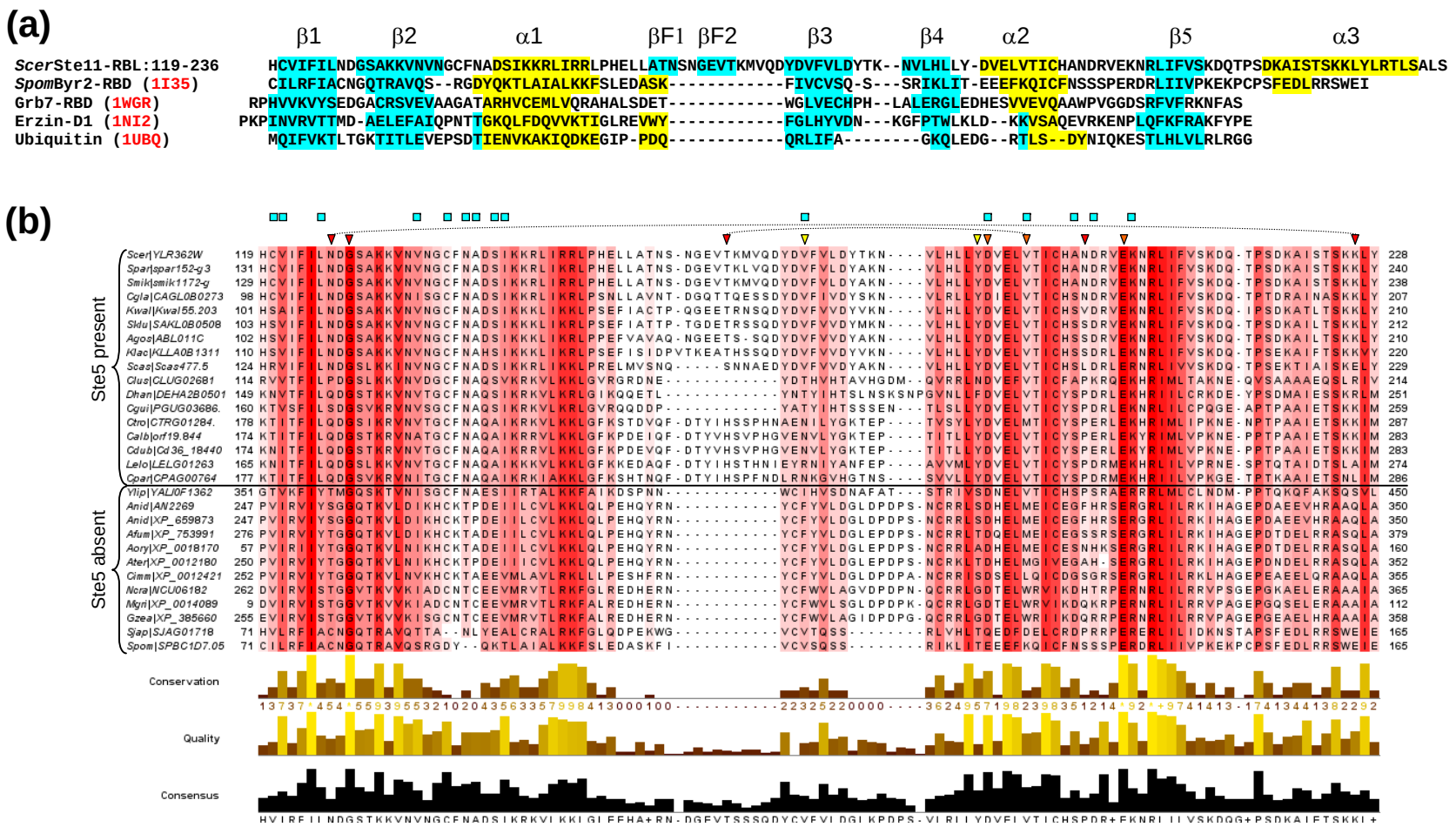

(a)
β1 β2 α1 βF1 βF2 β3 β4 α2 β5 α3
ScerSte11-RBL:119-236
SpomByr2-RBD (1I35)
Grb7-RBD (1WGR)
Erzin-D1 (1NI2)
Ubiquitin (1UBQ)
(b)
Ste5 present
Ste5 absent
Conservation
Quality
Consensus

Figure S2

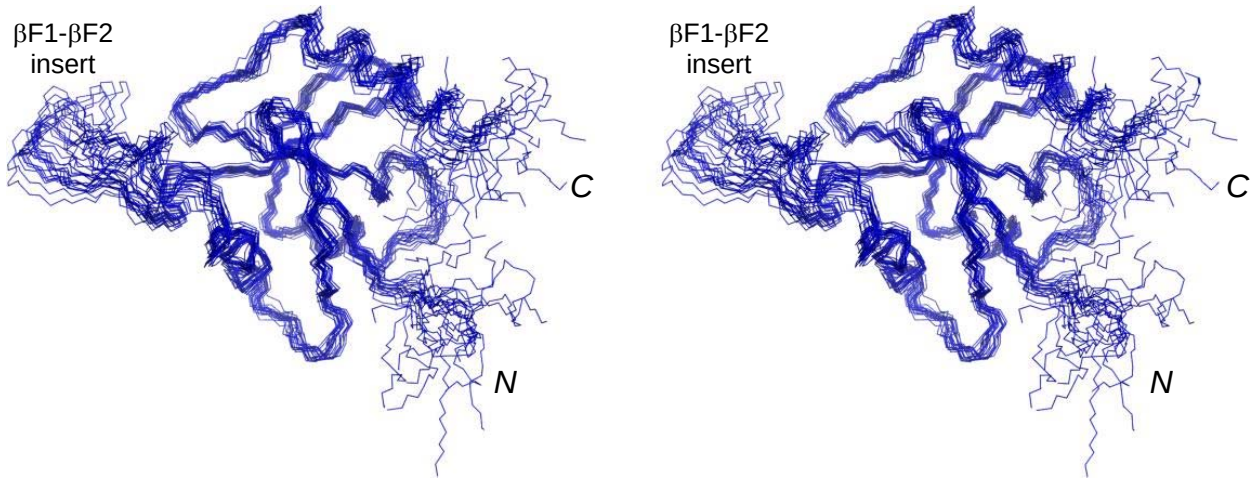


Figure S3

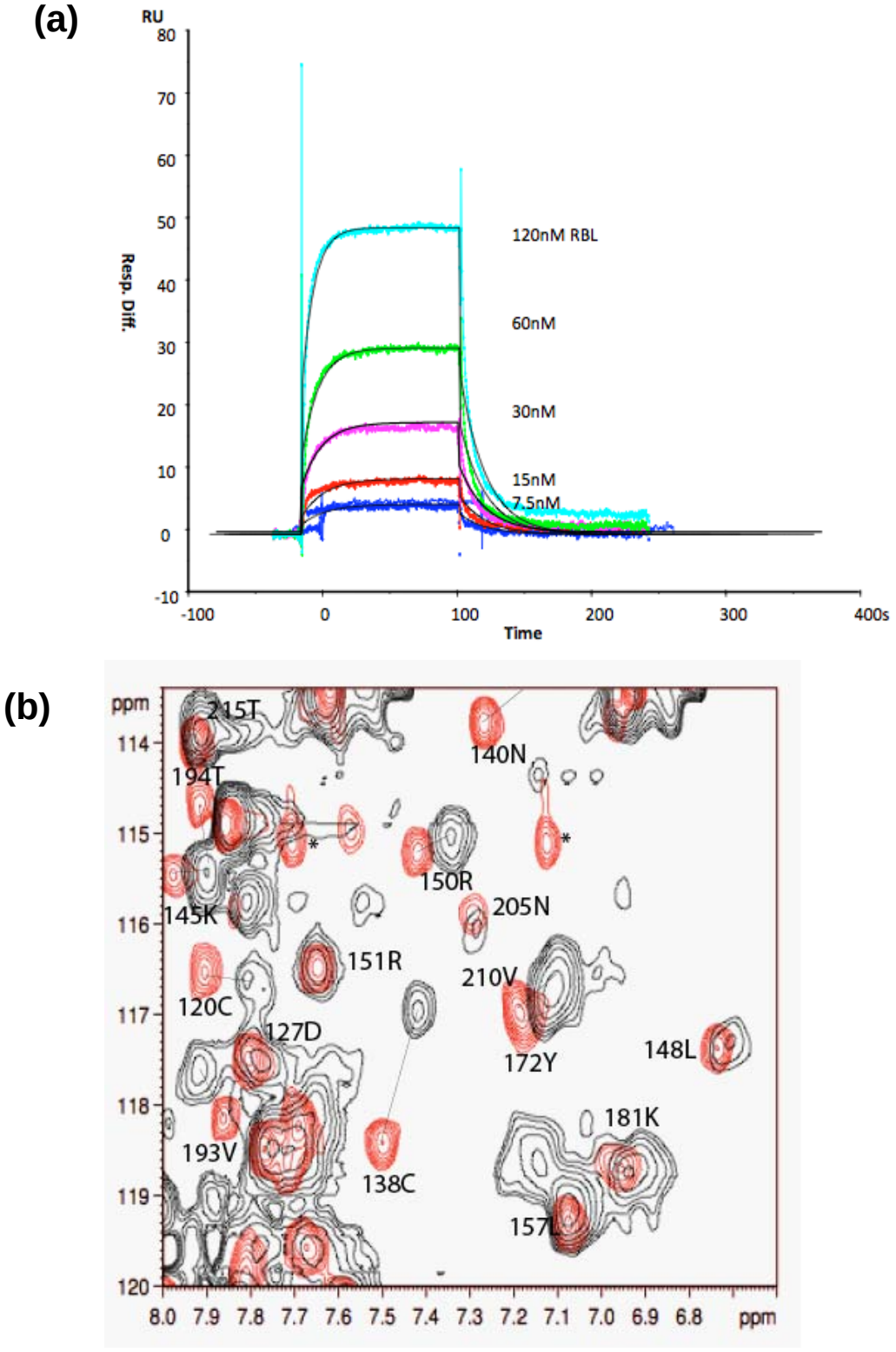
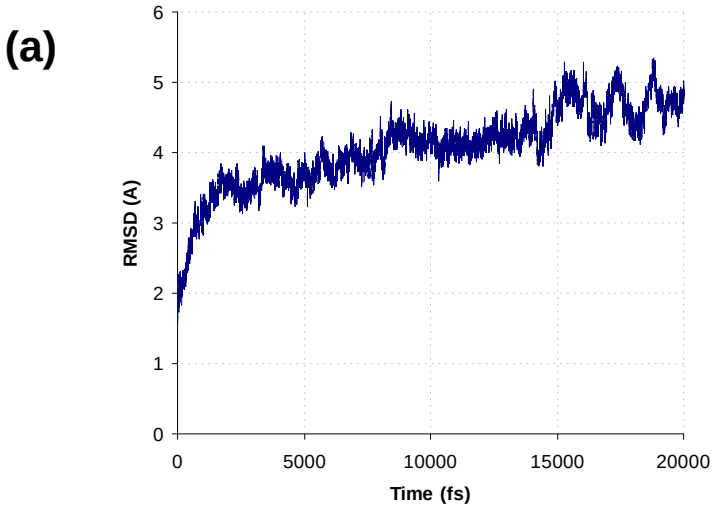
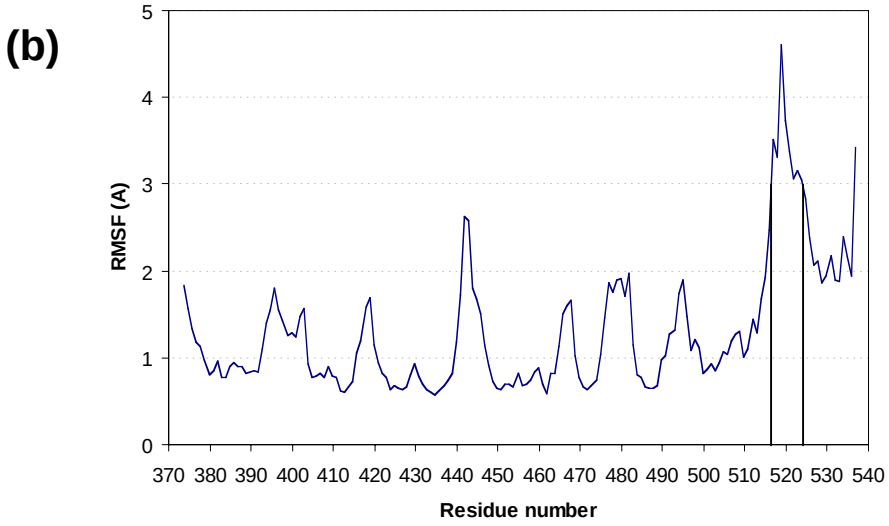

(a)
RU
Resp. Diff.
Time
120nM RBL
60nM
30nM
15nM
7.5nM
(b)
ppm
215T
194T
140N
150R
205N
145K
151R
210V
120C
127D
172Y
148L
181K
193V
138C
157L

Figure S4


(a)
RMSD (A)
Time (fs)
0
1
2
3
4
5
6
0
5000
10000
15000
20000


(b)
RMSF (A)
Residue number
0
1
2
3
4
5
370 380 390 400 410 420 430 440 450 460 470 480 490 500 510 520 530 540

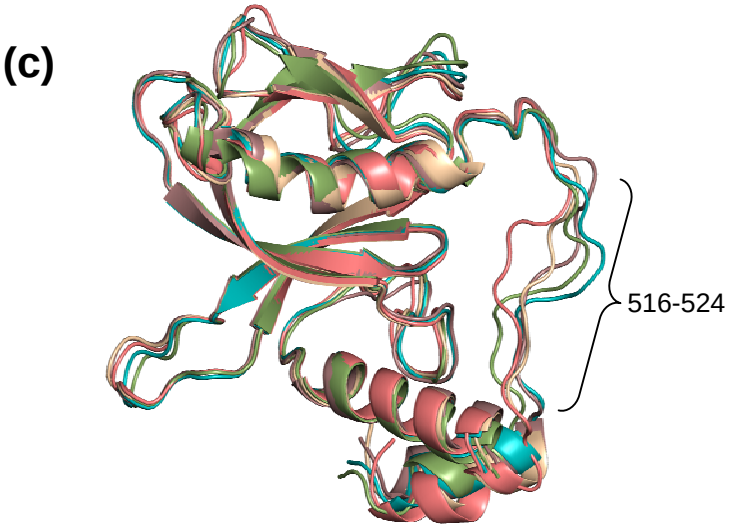
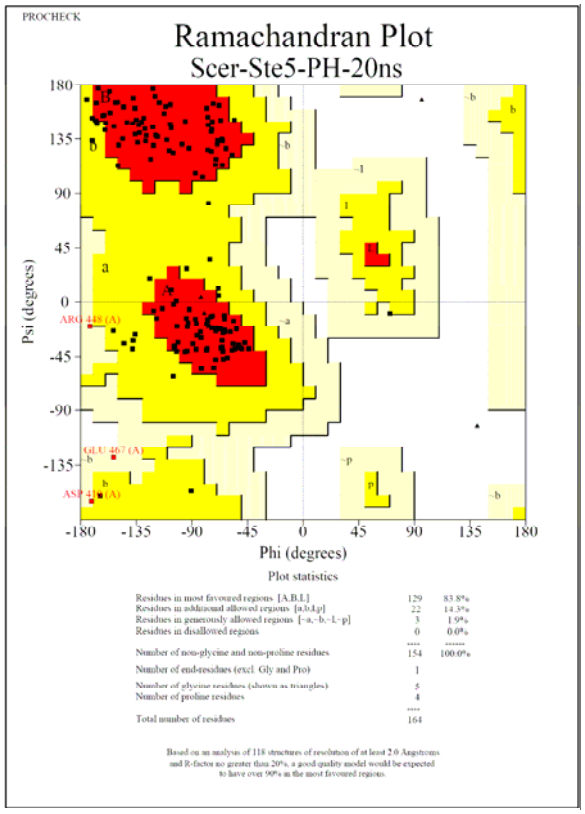
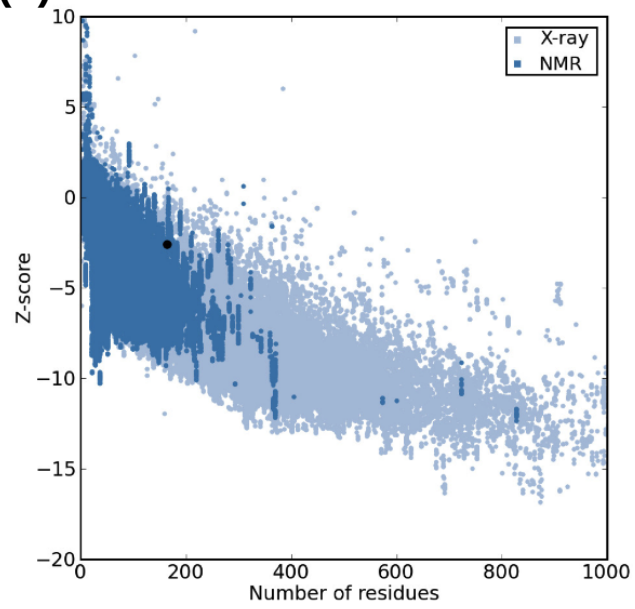

(c)
516-524

Figure S5

(a)



(b)



(c)

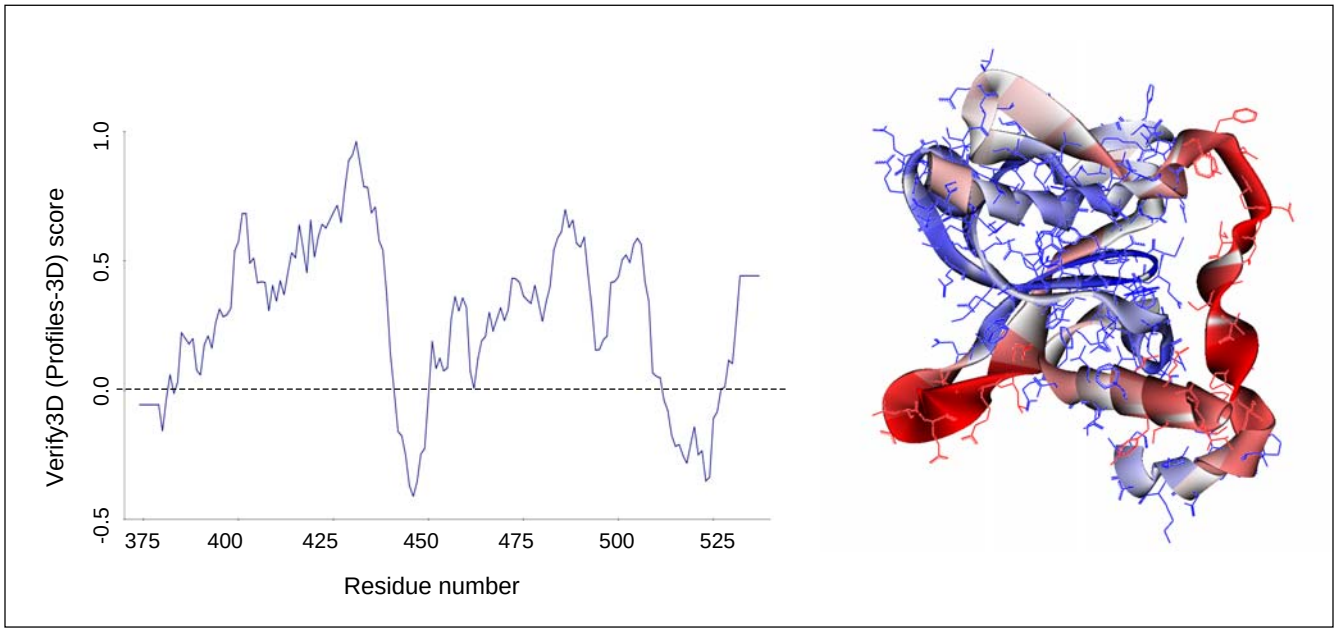


Figure S6

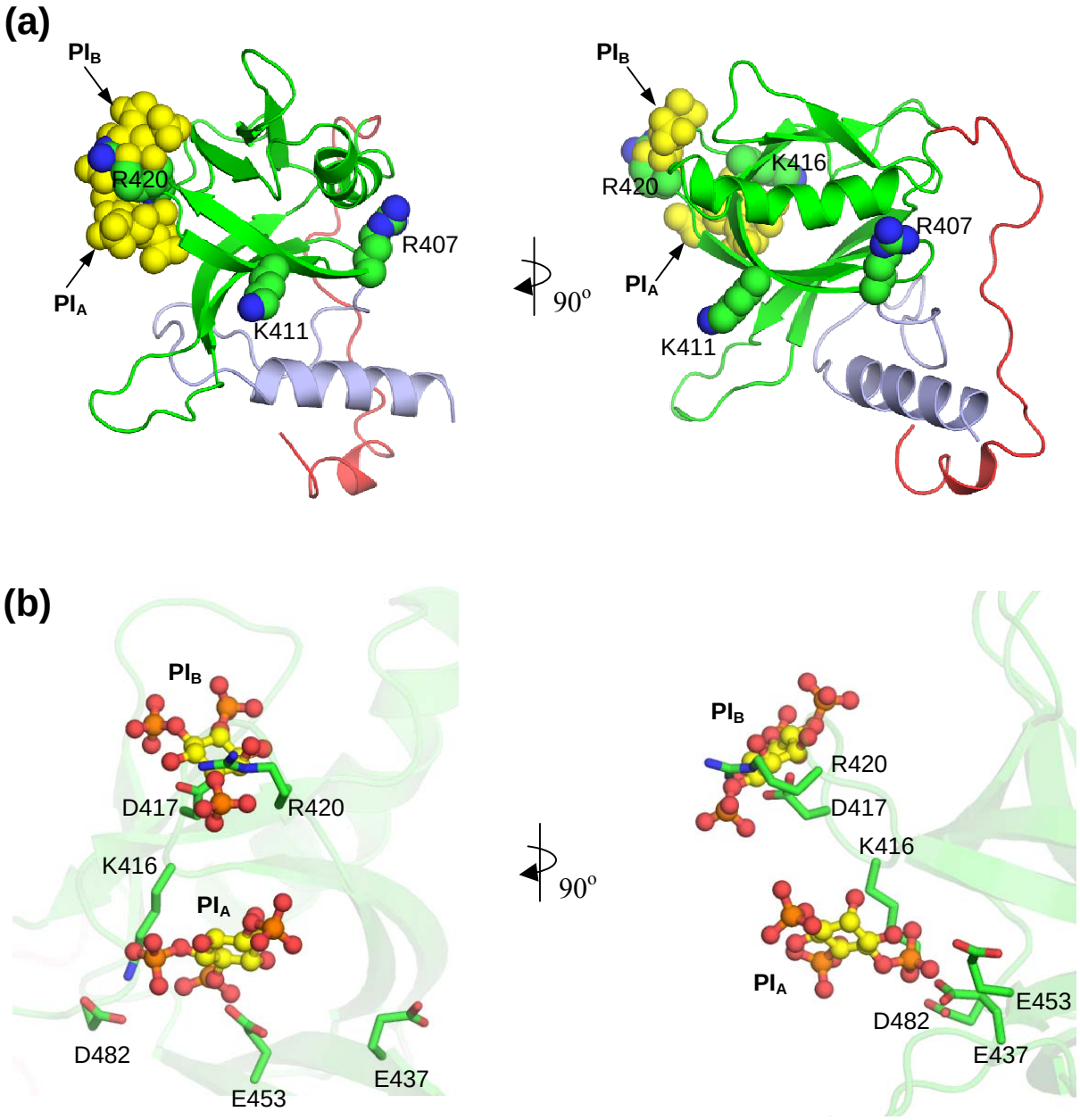

(a)
PIB
R420
PIA
R407
K411
90o
PIB
K416
R420
R407
PIA
K411
(b)
PIB
D417
R420
K416
PIA
D482
E453
E437
90o
PIB
R420
D417
K416
PIA
E453
D482
E437

Figure S7

TABLE S1. Structural statistics for the RBL^{Ste11} domain.

Restraints for structure calculation	
Total	1696
Total NOE	1492
Intra-residual	358
Sequential (|i - j| = 1)	431
Medium range (1<|i - j| < 5)	256
Long range (|i - j | ≥ 5)	447
Hydrogen bonds	12
Dihedral angles	172
RMSD of the 20 lowest energy structures from the mean coordinates (Å)	
Average backbone RMSD to mean:	1.02 +/- 0.35
Average heavy atom RMSD to mean:	1.49 +/- 0.29
Ramachandran plot statistics	
Most favoured (%)	77.2
Additionally allowed	20.6
Generously allowed	1.1
Disallowed	1.0

Supplemental figure legends

Figure S1. Multiple sequence alignments of predicted PH domains of fungal Ste5 and Far1 proteins. (a) Multiple-query/multiple-template sequence alignment of PH domains from *S. cerevisiae* and *C. albicans* Ste5 and Far1 proteins with similar PH domain template structures. Color shading represents secondary structure elements, β-strands (cyan) and α-helices (yellow), predicted for queries and observed experimentally for templates (PDB codes in red). (b) Extensive multiple sequence alignment of PH domains from fungal Ste5 and Far1 proteins. Marked residues correspond to those mutated in this study, with colors corresponding to those in Figure 5A.

Figure S2. Multiple sequence alignments of predicted RBD-like domains of fungal Ste11 proteins. (a) Multiple sequence alignment of the RBD-like domain from *S. cerevisiae* with similar structures adopting ubiquitin fold (including various RBD domains and ubiquitin). Color shading represents secondary structure elements, β-strands (cyan) and α-helices (yellow), observed experimentally for the RBL^{Ste11} domain from *S. cerevisiae* (this work; labeled on top), and for other structures (PDB codes in red). (b) Extensive multiple sequence alignment of RBD-like domains from fungal Ste11 proteins. Residues marked by triangles are functionally important as determined by random mutagenesis in this study, with symbol colors according to residue colors in Figure 5B (dotted connecting lines indicate double mutants). Residues marked by cyan squares interact with the PH^{Ste5} domain according to the NMR experiments in this study (also shown in Figure 5C).

Figure S3. Stereoview showing the backbone superposition of 20 lowest-energy NMR solution conformations of the RBL^{Ste11} domain structure. Flexible regions are labeled.

Figure S4. *In vitro* interaction of the RBL^{Ste11} domain with the PH^{Ste5} domain. (a) Surface plasmon resonance (SPR) analysis of bacterially expressed and purified RBL^{Ste11} and PH^{Ste5} domain. (b) Expanded region of NMR HSQC spectra of the RBL^{Ste11}:PH^{Ste5} complex. ^{15}N-labeled RBL^{Ste11} domain alone (red spectrum) and in complex with the unlabeled PH^{Ste5} domain (black spectrum).

Figure S5. Structural convergence of the MD simulation of the PH^{Ste5} domain. (a) Backbone root-mean-square deviations (RMSD) from the initial structure. (b) Backbone root-mean-square fluctuations (RMSF) averaged per each residue over the last 5 ns of MD trajectory (16^{th} ns to 20^{th} ns). The residues 517-527 show backbone RMSF values larger than 3 Å. (c) Overlay of 5 average 1-ns structures from the last 5 ns of MD trajectory (16^{th} ns to 20^{th} ns). The residues with backbone RMSF values larger than 3 Å are indicated.

Figure S6. Quality validation tests for the PH^{Ste5} structural model. (a) Ramachandran plot from PROCHECK analysis (Laskowski et al., 1993), showing that the backbone phi and psi torsion angles for 98% of modeled residues occupy the most favoured and additionally allowed regions of the conformational space. (b) Global quality from ProSA analysis (Wiederstein and Sippl, 2007). The global Z-score is -2.6 for 164 residues (black dot), which lies within the space occupied by native protein structures determined by X-ray and NMR methods. (c) Local quality from Verify3D analysis of Profiles-3D scores (Luthy et al., 1992) in Discovery Studio v2.5 (Accelrys, Inc., San Diego, CA). Averaging over a 10-residue sliding window was applied. The backbone ribbon rendering of the model in the panel on right has the color and width proportional with the Verify3D score. Negative values (full red, widest ribbon) may represent regions of lower quality, which in the present model structure are the flexible, solvent exposed

segments with weak intramolecular packing: the flexible linker between the C-terminal αC and αC1 helices, and the tip of the longer β3-β4 hairpin loop.

Figure S7. Assessing potential PI-binding sites on the modeled PH^{Ste5} domain. (a) Global views. PI binding sites are identified by PI molecules (yellow CPK models) extracted from PH domains superimposed onto the modeled PH^{Ste5} domain. No structural refinement was attempted in order to optimize the interactions between PH^{Ste5} domain and each PI. The positions of the canonical (PI_A) and non-canonical (PI_B) binding sites are obtained from the PH domains of DAPP1 (PDB code 1FAO) and β-spectrin (PDB code 1BTN), respectively. The PI shown is PI(4,5)P2 at both the PI_A site (obtained after removal of the 3-phosphatefrom the PI(3,4,5)P3 bound in the template structure 1FAO) and PI_B site (as complexed in the template structure 1BTN). The side-chains of four basic residues of Ste5 mutagenized previously in order to study PI-binding capacity of the PH^{Ste5} domain (Garrenton et al., 2006) are represented as CPK models and labelled. (b) Close-up views of PH^{Ste5} domain interactions at the canonical and non-canonical PI-binding sites. PI models are shown as ball-and-stick models, and select side-chains of charged residues in the vicinity of these sites are shown as sticks and labeled.

Supplemental Experimental Procedures

Reagents

Restriction endonucleases and DNA-modifying enzymes were obtained from New England Biolabs (Beverly, MA) and GE Healthcare (Quebec, Canada). High fidelity Expand thermostable DNA polymerase, and tablet protease inhibitors were purchased from Roche Molecular Biochemicals (Quebec, Canada). Acid-washed glass beads (450-600 □m), synthetic α-mating factor, protease inhibitors, and bovine serum albumin were purchased from Sigma (Ontario, Canada). Geneticin was purchased from Life Technologies (Ontario, Canada), and nourseothricin (clonNAT) from Werner BioAgents (Jena-Cospeda, Germany). Plasmids pGEX-4T-3 and pGEX-2TK, glutathione-Sepharose beads, glutathione, NiNTA resin, and protein A/G Sepharose beads were obtained from GE Healthcare. Vector pET15b was from EMD4Biosciences. The antibody against GST was described previously (Wu et al., 1999), and horseradish peroxidase conjugated secondary antibodies were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Nitrocellulose and FVDF membranes were purchased from Millipore. The enhanced chemiluminescence (ECL) assay system was purchased from Roche Molecular Biochemicals.

Yeast strains and manipulations

Yeast media, culture conditions and manipulations of yeast strains were as described (Rose et al., 1990). Yeast transformations with circular or linearized plasmid DNA were carried out after treatment of yeast cells with lithium acetate (Gietz et al., 1992; Rose et al., 1990). The yeast GST-ORF library was purchased from Invitrogen (Carlsbad, CA), and the yeast deletion strain collection was purchased from ATCC. The yeast strains used in this study are listed below:

Yeast strains used in this study.

Strain	Relevant genotype	Source
W303-1A	*MAT***a** *ade2 ura3 his3 leu2 trp1 can1*	R. Rothstein
W303-1B	*MAT*□ *ade2 ura3 his3 leu2 trp1 can1*	R. Rothstein
BY4741	*MAT***a** *ura3 his3 leu2 met15*	ATCC

BY4742	*MAT□ ura3 his3 leu2 lys2*	ATCC
YCW338	*MATa ura3 his3 leu2 sst1::hisG FUS1-LacZ::LEU2*	Wu et al, 1999
YCW340	*MAT***a** *ste11Δ::Kan^R ssk2Δ::LEU2 ssk22Δ::LEU2 ura3 leu2 his3*	Wu et al, 1999
YCW1172	*MAT***a** *ste11Δ::Kan^R sst1::hisG FUS1-LacZ::LEU2 ura3 leu2 his3*	This study
YCW1216	*MAT***a***ste5Δ::TRP1 FUS1-LacZ::LEU2 sst1::hisG ura3 leu2 his3*	This study
YCW757	*MAT***a***ura3 leu2 his3 ssk2Δ::LEU2 ssk22Δ::LEU2 ste11Δ::Kan^R ste50Δ::TRP1 ste5Δ::hisG*	Wu et al, 2006
YCW1476	*MAT***a** *ste50Δ::TRP1 ssk2Δ::LEU2 ssk22Δ::LEU2 ura3 leu2 his3*	This study
YCW1477	*MAT□ ste50Δ::TRP1 ssk2Δ::LEU2 ssk22Δ::LEU2 ura3 leu2 his3*	This study
DC17	*MAT□ his1*	J. Hicks
DC16	*MAT***a** *his1*	J. Hicks

Plasmid construction and mutagenesis

To map protein-interaction boundary of PH^{Ste5} domain with Ste11, we used the yeast two-hybrid system (Y2H) developed recently in our lab for the detection of protein-protein interactions in cytoplasm (Cote et al., 2011); detailed information of this Y2H system will be published elsewhere. Briefly, this yeast two-hybrid system is based on the interaction of Ste11 (MAPKKK) and Ste50 that is required for the HOG pathway activation and osmoadptation, which is critical for the survival of yeast cells under hyperosmotic stress. The interaction of Ste11 and Ste50 through their respective SAM domains that is required to activate the HOG pathway can be replaced by interaction of other protein interacting modules (Wu et al., 2006). We used this property to analyze the interaction of the PH^{Ste5} domain with Ste11. The PH^{Ste5} domain fragments were cloned into the SmaI site of plasmid pYL45 containing the fragment of *STE50* encoding Ste50 protein without its SAM domain (aa 115-346) at the SalI site of pGREG503 (Jansen et al., 2005) through *in vivo* recombination (IVR) in yeast strain YCW1476. All the primers used for the PCR amplification reactions contain gene specific sequences and common sequences used for IVR in a layout as follows: 5'-ATTCTAGAGCGGCCGCACTAGTGGATCCCCCGGG-gene specific sequence (starting with ATG)-3' for the forward orientation, and 5'-TCGATAAGCTTGATATCGAATTCCTGCAGCCCGGG-gene specific sequence (delete the stop codon)-3' for the reverse orientation. To query the bait-prey interaction, IVR positive clones were selected and their ability to activate the HOG pathway measured as the ability to grow on hyperosmolarity media was scored.

For random mutagenesis of the PH^{Ste5} domain and to effectively select against nonsense mutations, we first generated the IFM-Ste50ΔSAM construct in which the *NAT1* (nourseothricin resistance marker) was cloned into pYL45 into the AatII site in frame with Ste50ΔSAM, and the resulting plasmid was digested with SmaI to add a stuffer *URA3* marker as SmaI fragment, which

replaces the N-terminal part (1-16aa) of Nat1 to create pVY0017. The error prone PCR of the PH^{Ste5} domain (corresponding to aa 373 to 537) was performed with the primers OCW357 and OVY13 (the low case letters corresponding to the sequence of PH^{Ste5} domain). Error prone PCR was carried out with Taq DNA polymerase with 5 mM $MgCl_2$, 0.2 mM dNTP (each) mix, BSA (0.8 ug/μl) for 30 cycles. The PCR products were cloned into SmaI digested pVY17 through *in vivo* recombination in yeast strain YCW1476. Recombinants were first selected on SD-his plates, then on YP/Gal-nourseothricin plates for nourseothricin resistance (IFM function), and finally on SD/Gal –his +sorbitol hyperosmolarity plates. PH^{Ste5} domain mutants that were unable to grow on high osmolarity media but were nourseothricin resistant indicative of defective binding to the RBL^{Ste11} domain were picked. The PH^{Ste5} domain mutants were transferred through IVR with pCW789 to make a CEN plasmid-born *STE5* under its own promoter for further analysis. GFP-tagged Ste5 alleles were constructed by cloning the Ste5 mutants into pGREG576 according to the procedure as described (Jansen et al., 2005).

To create a Ste11 mutant with an internal in-frame deletion of the region encompassing amino acid residues 117-240 corresponding to the predicted RBL domain, primer OCW331 was used with pCW199 (Wu et al., 1999) as template using a Quickchange kit (Strategene). The mutagenesis also introduced unique restriction enzyme sites for XhoI and HpaI, and all the modifications were confirmed by DNA sequencing. This gave plasmid pCW850. A SmaI-site ended *URA3*-stuffer was then put into the HpaI site of pCW850 to create pCW719 to be used for the mutagenesis analysis of the RBL^{Ste11} domain. Random mutagenesis of RBL^{Ste11} was performed by IVR of XhoI-digested pCW719 with error-prone PCR product of Ste11 fragment encompassing the RBL region. The loss of the *URA3*-stuffer marker was used as the selection of IVR product at the Ste11 BRL region.

To make Ste5 constructs carrying the SAM domain (corresponding to aa 29 - 131) of Ste50, we amplified the SAM domain using PCR with the primers OCW541 and OCW542R, and inserted it into *STE5* at the XhoI site through *in vivo* recombination with XhoI digested Ste5 plasmids. The manipulation resulted in the Ste50-SAM domain fused to Ste5 C-terminally at the XhoI site.

Site-directed mutagenesis of *STE5* and *STE11* were performed either with either mutagenic sewing PCR (Ho et al., 1989) followed by *in vivo* recombination (IVR) in yeast, or with a Quick-change mutagenesis kit according to manufacturer's protocol. All desired mutations were confirmed by sequencing. Oligonucleotides and plasmids used in this study are listed below:

Oligonucleotides used in this work.

Name	**Sequence**
OCW331	CGATGAATTCAGAATTGATTCCTctcgaggttaacTCCTCGTCAAATTTGTTGGCA
OCW357	ATTCTAGAGCGGCCGCACTAGTGGATCCCCCGGGATGacaacgttgccgctgtta
OCW358R	TCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGattgaatacaaaatcctgatt

OCW359R	TCGATAAGCTTGATATCGAATTCCTGCAGCCCGGGcagggtcgaagtgatattgtc
OVY0013	TGTCGGTGGTGAAGGACCCATCCAGTGCCTCGATGGCCTCGGCGTCacca acatcttttgaaaagtt
OVY0007	tata gacgtc GGTACCACTCTTGACGACAC
OVY0008	agaa gacgtc ACC GCT CGC GGGGCAGGGCATGCTCATGT
OVY0011	tata gacgtc tctgttattaatttcacagg
OVY0012	agaa gacgtc ACC GCT CGC tttcttagcatttttgacga
OCW411	AAGAGCAAAGGATGACGAAACCAAAcccggggttaacGGTAGGCACGAGAC GAGT
OCW411R	ACTCGTCTCGTGCCTACCgttaaccccgggTTTGGTTTCGTCATCCTTTGC
OCW450	GCTAATGATTCTATTTCTGCTGTTTCCAATTCGG TAAGAGCAAA GGATGACGAAACCAAAACAACGTTGCCGCT GTTA
OCW451R	AACTTCAACAACTTTGTTAGGATTGATTAAACCTAGAAAGGTACTCGTC TCGTGCCTACCATTACCAACATCTTTTGAAAAGTT
OCW541	GGAATCGATGGCATAACCAGACGCAGTTCATTCTCGAGTaatgaagacttttccc agtg
OCW542R	GAGGGGACAGTTGTTATTACCGCTCTCTATAAGACTCGAGcc ttgcaatttcgcagatgtagt

Plasmids used in this work.

Name	
pCW199	pRS313-*STE11::HIS3*, own promoter, CEN/ARS, Wu et al, 1999.
pCW51	pRS316-*STE5::URA3*, own promoter, CEN/ARS.
pCW185	YCP50-*STE11-1*, CEN/ARS, (Stevenson et al., 1992).
pVY2	pFO1a-Ste5(373-537), N-terminally His-tagged
pVY3	pFO1a-Ste5(373-523), N-terminally His-tagged
pVY4	pFO1a-Ste11(116-236), N-terminally His-tagged
pVY6	pGEX-Ste5(373-537), N-terminal GST-fusion
pVY7	pGEX-Ste5(373-523), N-terminal GST-fusion
pVY8	pGEX-Ste11(116-236), N-terminal GST-fusion
pVY-1-18	pGAL1-GFP-Ste5^{I504T}::*URA3*, CEN/ARS
pVY31-11-11	pGAL1-GFP-Ste5^{Q501R}::*URA3*, CEN/ARS
pVY31-11-17	pGAL1-GFP-Ste5^{F514L}::*URA3*, CEN/ARS
pVY31-8-5	pGAL1-GFP-Ste5^{wt}::*URA3*, CEN/ARS
pVY13	Derivative of pGAL1- PH^{Ste5}-Ste50ΔSAM::*HIS3* where *Nat1* was

	inserted into AatII site in frame with upstream PH^{Ste5} domain and downstream Ste50ΔSAM, CEN/ARS and *HIS3* as plasmid selection marker.
pVY17	Derivative of pVY0013 where PH^{Ste5} domain and the part of *Nat1* were replaced by *URA3*-stuffer marker for IVR selection at SmaI sites.
pCW719:	p*STE11*ΔRBL(117-240)-*URA3::HIS3*, own promoter, CEN/ARS, *URA3*-stuffer marker for IVR selection, *HIS3* for plasmid selection, XhoI digestion for IVR.
pCW789	p*STE5*ΔPHD(373-538)-*HIS3::URA3*, own promoter, CEN/ARS, *HIS3*-stuffer marker for IVR selection, *URA3* for plasmid selection, SmaI digestion for IVR.
pCW790	p*STE5*ΔPHD(373-538)-*URA3::HIS3*, own promoter, CEN/ARS, *URA3*-stuffer marker for IVR selection, *HIS3* for plasmid selection, SmaI digestion for IVR.
pCW850:	p*STE11*ΔRBL(117-240)*::HIS3*, CEN/ARS, *HIS3* for plasmid selection, Hpa1/XhoI sites introduced at the deletion.

Yeast HOG, pheromone response and other assays

Halo assays to test cell growth inhibition in response to α-mating factor, assays for the ability of cells to grow on hyperosmotic media to test the function of the HOG pathway, and yeast extract preparation and Western blot analyses were performed as described previously (Wu et al., 1999). Quantitative β-galactosidase reporter assays for the pheromone response and HOG pathways were performed as described (Tatebayashi et al., 2006; Wu et al., 1999). GFP fluorescence photomicroscopy was performed as previously described (Wu et al., 1999).

Protein expression and purification

Various Ste5 and Ste11 fragments were cloned into pGEX or pFO1a (a derivative of pET15b) vectors, and expressed in the BL21 or Rosetta pLys strains of *E. coli*. Protein was expressed in *E. coli* grown in 2YT medium supplemented with 200 mg/l ampicillin. Cultures were grown at 25°C to an OD_{600} of 0.5-0.6. The temperature was then decreased to 15°C and protein expression was induced with 1 mM IPTG. After incubation at 15°C overnight, cells were harvested by centrifugation. Cells were re-suspended in PBS supplemented with 300 mM NaCl, 10 mM imidazole and 1 mM DTT, and lysed by sonication. The lysate was cleared of cell debris by centrifugation at 37,000 x *g* for 1 hour at 4°C. All His-tagged fusion proteins were purified by affinity chromatography using NiNTA resin (GE Healthcare) using standard protocols. The GST fusion proteins were purified by affinity chromatography on glutathione Sepharose according to a modification of the manufacturer's recommendation. Cleared lysates were incubated with resin on ice for 30 minutes. Resin was then collected in a column and washed with 50 ml of 20 mM HEPES, pH 7.5, 300 mM NaCl, 5 mM DTT. For resin binding assay, the beads were blocked for 30 min with the same buffer supplemented with 1% BSA and 0.1% Triton X-100 before incubation with binding partners to be tested. For further purification of proteins for biophysical studies, the bound fusion protein was cleaved with thrombin for 4 hours at room temperature and the protein was eluted in the HEPES buffer. Size exclusion

chromatography was done on a Superdex 75 10/30 column (GE Healthcare) equilibrated with 20 mM HEPES, pH 7.5, 300 mM NaCl, 5 mM DTT. The buffer was sparged with argon for 30 minutes before use. The flow rate was 0.5 ml/min and the sample volume was 200 µl. Protein concentrations were estimated by the Bradford method and the BCA (Pierce) method following the protocols supplied by the manufacturers.

Yeast GST fusions of the small GTPases of Ras/Rho family were cloned from the GST-ORF library collection. The expression and purification of the GST fusion onto glutathione Sepharose beads, and subsequent resin binding assay with bacterially expressed His-tagged RBL^{Ste11} were carried out in conditions as previously described (Annan et al., 2008). The presence His-tagged protein and GST fusions were analyzed by Western-blotting with anti-His and anti-GST antibodies respectively.

Surface plasmon resonance analysis of protein-protein interaction

Surface plasmon resonance analysis of the RBL^{Ste11}:PH^{Ste5} domain interaction was performed using a Biacore 3000 (GE Healthcare Bio-Sciences Corp., Piscataway, NJ). The PH^{Ste5} domain was covalently immobilized to a CM-5 sensorchip using standard amine coupling methods. Briefly, a 35 µL injection of a mixture of 0.05 M N-hydroxysuccinimide (NHS) and 0.2 M N-ethyl-N'-(3-diethylaminopropyl) carbodiimide hydrochloride (EDC) was followed by the manual injection of His-tagged PH^{Ste5} domain diluted 100-fold in 10 mM sodium acetate pH 5 until approximately 700 RU immobilization level was reached. The remaining activated surface groups were inactivated by a 35 µL injection of 1 M ethanolamine. Similarly, a blank control surface was created by NHS / EDC treatment followed immediately with ethanolamine.

Using the KINJECT command, various concentrations of RBL^{Ste11} domain in running buffer containing 10 mM HEPES, 0.02% Tween 20, 50 µM EDTA, pH 7.4 were injected over the PH^{Ste5} surface for 2 minutes at a flow rate of 25 µL/min and a dissociation of 300 s. The PH^{Ste5} surface was regenerated for subsequent RBL^{Ste11} domain injections using a 15 s pulse of 20 mM HCl. Sensorgrams were aligned and double referenced to the control surface using buffer injections, and analyzed by both global fitting to a 1:1 interaction and steady state analysis with BiaEvaluation (v3.2) software.

Structural bioinformatics

Homologous protein sequences were retrieved by either BLASTP or TBLASTN search (Altschul and Lipman, 1990) in the Fungal genome database (http://seq.yeastgenome.org/) and by protein domain architecture search in the SMART database (http://smart.embl.de/; (Letunic et al., 2006)). Comparative sequence analysis of the assembled datasets (29 Ste11-like, 27 Far1-like and 18 Ste5-like sequences) was carried out with MUSCLE (Edgar, 2004a, b) and MAFFT v6 (Katoh et al., 2005) for deriving multiple sequence alignments with the L-INS-i iterative refinement algorithm and to generate phylogenetic trees, and visualized using Jalview v2.4.0B2 (Waterhouse et al., 2009).

Structural fold detection for regions of representative Ste5, Far1 and Ste11 fungal proteins was carried out at the Structure Prediction Meta Server (http://bioinfo.pl/meta/), which assembles state-of-the-art fold recognition methods, and provides a consensus sequence-to-structure scoring using the 3D-Jury meta-predictor (Ginalski and Rychlewski, 2003). Multiple-queries/multiple-templates sequence alignments were assembled using a combination of: (i) multiple sequence alignment of the queries using MAFFT v6 (Katoh et al., 2005), (ii) structure-

based sequence alignment of the templates using 3D-Coffee (Armougom et al., 2006), (iii) 3D-Jury consensus fold recognition based alignment between queries and templates, (iv) manual minor local improvements in the overall alignment of secondary structure elements predicted using SAM-T02 with DSSP and STRIDE alphabets (Karplus et al., 2005), and (v) 3D-structural alignment by alternate domain fit in Swiss-PdbViewer v4.0.1 (Guex and Peitsch, 1997).

Model building and refinement

Homology modeling of the PH domain of *S. cerevisiae* Ste5 was done in MODELLER v9.1 (Fiser and Sali, 2003a, b; Marti-Renom et al., 2000). Using a multiple-queries/multiple-templates sequence alignment (Figure S1) several hundreds of initial models were built and ranked based on global and residue-based DOPE scores (Shen and Sali, 2006). Based on this analysis, the structure of PH domain of the guanine nucleotide exchange factor collybistin (PDB code 2DFK) (Xiang et al., 2006) was selected as the most suitable template structure, providing a good overall sequence and secondary structure alignment together with structural homology in the functionally important N- and C-terminal extensions beyond the canonical PH fold. A stepwise loop refinement protocol was devised, consisting of sequential refinement of structurally adjacent loop pairs, using the DOPE score to select best the model at each stage.

The structure resulted from the loop refinement protocol was used as input for a classical MD simulation carried out under the AMBER force field with the FF03 parameters (Duan et al., 2003; Lee and Duan, 2004) using the AMBER 9 software (Case et al., 2005). The structure was solvated in a truncated octahedron TIP3P water box (Jorgensen et al., 1983), and electroneutrality was achieved by adding Na^+ counterions. Applying harmonic restraints with force constants of 10 kcal/(mol $Å^2$) to all solute atoms, the system was energy-minimized first, followed by heating from 100 to 300 K over 25 ps in the canonical ensemble (constant number of particles, volume, and temperature, NVT) and by equilibrating to adjust the solvent density under 1 atm pressure over 25 ps in the isothermal-isobaric ensemble (constant number of particles, pressure, and temperature, NPT) simulation. The harmonic restraints were then gradually reduced to 0 with four rounds of 25-ps simulations. A 20-ns production NPT run was obtained with snapshots collected every 1 ps, using a 2-fs time step and 9-Å nonbonded cutoff. The Particle Mesh Ewald method (Darden and Pedersen, 1993) was used to treat long range electrostatic interactions, and bond lengths involving bonds to hydrogen atoms were constrained by SHAKE (Ryckaert et al., 1977). Standard analyses of MD trajectories were carried out with PTRAJ in AMBER 9. Model validation was carried out with PROCHECK (Laskowski et al., 1993), ProSA (Wiederstein and Sippl, 2007), and Verify3D (Luthy et al., 1992).

Protein production for NMR spectroscopy

Constructs expressing the RBL^{Ste11} domain (aa 116-236) were prepared using the expression vector pFO1, a derivative of pET15b (Novagen), to obtain an N-terminal His_8-tagged thrombin-cleavable construct. After verification by DNA sequencing, the constructs were transformed into *E. coli* Rosetta pLysS (Novagen) for protein expression. The RBL^{Ste11} domain was produced in minimal medium (M9) enriched with ^{15}N-ammonium chloride or ^{15}N-ammonium chloride /^{13}C-glucose and expression was induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside at 20 ºC for 18 hrs. The recombinant protein was purified by standard immobilized metal affinity chromatography using Ni^{2+}-NTA resin (Qiagen). The protein was eluted with buffer containing 350 mM imidazole, and the His-tag was cleaved using thrombin.

For studies of RBL/PH domain interactions, His-tagged PH^{Ste5} domain (aa 373-537) was used, produced and purified as described above.

NMR spectroscopy

Samples for NMR measurements were prepared in 50 mM sodium phosphate buffer pH 6.8, 200 mM NaCl, 5 mM DTT, 0.02% (w/v) NaN_3 at protein concentrations of 1.2 – 1.5 mM. All NMR data were collected at 300 K on a Bruker Avance500 spectrometer equipped with a triple-resonance cryoprobe and with z-gradient pulse field gradient accessories.

Sequence-specific backbone and aliphatic side chain chemical shift assignments for the RBL^{Ste11} domain were obtained using combined HNCA, CBCA(CO)NH, HNCACA, H(CCCO)NH, HBHA(CO)NH and (H)CC(CO)NH experiments. The ^1H chemical shift assignments for aromatic side chains were based primarily on 2D DQF COSY, TOCSY and NOESY experiments. The ^1H chemical shifts were referenced directly to internal DSS at 0 ppm and the ^1C and the ^{15}N chemical shifts were referenced indirectly to DSS. NMR spectra were processed using XWINNMR (Bruker Biospin) and analysed with programs CARA, XEASY (Bartels et al., 1995), Sparky (Goddard and Kneller), SPARTA and MARS (Jung and Zweckstetter, 2004). The interaction between the ^{15}N-labeled RBL^{Ste11} and the PH^{Ste5} domains was studied by recording the HSQC spectra of ^{15}N-labeled RBL domain, titrated with unlabeled PH domain. To assist in assigning shifted NMR signals in ^1H-^{15}N HSQC spectra, amino acid specific ^{15}N-enriched samples were prepared using defined media and ^{15}N-labeled valine and leucine.

NMR structure calculation

The 3D ^1H-^{15}N NOESY-HSQC, 3D ^1H-^{13}C NOESY-HSQC (in ^2H_2O) and 2D homonuclear NOESY experiments were used to collect NOE-restraints for structure calculation. A mixing time of 100 ms was used in all experiments. Dihedral restraints were derived from $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C'$, $^1H^{\alpha}$ and ^{15}N chemical shifts using the program TALOS (Cornilescu et al., 1999). Structure calculations were performed using CYANA 2.1(Güntert, 2004). NOE upper limit distances were used together with dihedral angles in the standard CYANA protocol of seven interactive cycles of calculations. Automatic NOE assignments from CYANA were manually verified/corrected. The 20 final structures with lowest energy or violations were retained for refinement. Hydrogen bond constraints derived from ^1H, ^{15}N HSQC exchange experiment were added at the last stage of calculations. A summary of the results is given in Table S4, each structure was refined by conjugate-gradient energy minimization using the AMBER force field (Hornak et al., 2006) and a distance-dependent dielectric constant in a stepwise protocol in which the flexible termini were relaxed first, followed by minimization of the side-chains, and finally full relaxation of the structure.

References:

Altschul, S.F., and Lipman, D.J. (1990). Protein database searches for multiple alignments. Proceedings of the National Academy of Sciences of the United States of America *87*, 5509-5513.

Annan, R.B., Wu, C., Waller, D.D., Whiteway, M., and Thomas, D.Y. (2008). Rho5p Is Involved in Mediating the Osmotic Stress Response in Saccharomyces cerevisiae, and Its Activity Is Regulated via Msi1p and Npr1p by Phosphorylation and Ubiquitination. Eukaryot Cell *7*, 1441-1449.

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., and Notredame, C. (2006). Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. Nucleic acids research *34*, W604-608.

Bartels, C., Xia, T., Billeter, M., Güntert, P., and Wüthrich, K. (1995). The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. J Biomol NMR *6*, 1-10.

Case, D.A., Cheatham, T.E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K.M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R.J. (2005). The Amber biomolecular simulation programs. J Comput Chem *26*, 1668-1688.

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

Cote, P., Sulea, T., Dignard, D., Wu, C., and Whiteway, M. (2011). Evolutionary reshaping of fungal mating pathway scaffold proteins. MBio *2*, e00230-00210.

Darden, T.A., and Pedersen, L.G. (1993). Molecular modeling: an experimental tool. Environ Health Perspect *101*, 410-412.

Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T.*, et al.* (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem *24*, 1999-2012.

Edgar, R.C. (2004a). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics *5*, 113.

Edgar, R.C. (2004b). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research *32*, 1792-1797.

Fiser, A., and Sali, A. (2003a). Modeller: generation and refinement of homology-based protein structure models. Methods Enzymol *374*, 461-491.

Fiser, A., and Sali, A. (2003b). ModLoop: automated modeling of loops in protein structures. Bioinformatics *19*, 2500-2501.

Garrenton, L.S., Young, S.L., and Thorner, J. (2006). Function of the MAPK scaffold protein, Ste5, requires a cryptic PH domain. Genes Dev *20*, 1946-1958.

Gietz, D., St Jean, A., Woods, R.A., and Schiestl, R.H. (1992). Improved method for high efficiency transformation of intact yeast cells. Nucleic acids research *20*, 1425.

Ginalski, K., and Rychlewski, L. (2003). Detection of reliable and unexpected protein fold predictions using 3D-Jury. Nucleic acids research *31*, 3291-3292.

Goddard, T.D., and Kneller, D.G. SPARKY 3, University of California, San Francisco
Guex, N., and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis *18*, 2714-2723.

Güntert, P. (2004). Automated NMR structure calculation with CYANA. In Methods Mol Biol, A.K. Downing, ed. (Totowa, NJ, Humana Press Inc.).

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene *77*, 51-59.

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins *65*, 712-725.

Jansen, G., Wu, C., Schade, B., Thomas, D.Y., and Whiteway, M. (2005). Drag&Drop cloning in yeast. Gene *344*, 43-51.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. J Chem Phys *79*, 926-935.
Jung, Y.S., and Zweckstetter, M. (2004). Mars -- robust automatic backbone assignment of proteins. J Biomol NMR *30*, 11-23.

Karplus, K., Katzman, S., Shackleford, G., Koeva, M., Draper, J., Barnes, B., Soriano, M., and Hughey, R. (2005). SAM-T04: what is new in protein-structure prediction for CASP6. Proteins *61 Suppl 7*, 135-142.

Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic acids research *33*, 511-518.

Laskowski, R.A., Moss, D.S., and Thornton, J.M. (1993). Main-chain bond lengths and bond angles in protein structures. Journal of molecular biology *231*, 1049-1067.

Lee, M.C., and Duan, Y. (2004). Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model. Proteins *55*, 620-634.

Letunic, I., Copley, R.R., Pils, B., Pinkert, S., Schultz, J., and Bork, P. (2006). SMART 5: domains in the context of genomes and networks. Nucleic acids research *34*, D257-260.

Luthy, R., Bowie, J.U., and Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. Nature *356*, 83-85.

Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F., and Sali, A. (2000). Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct *29*, 291-325.

Rose, M.D., Winston, F., and Hieter, P. (1990). Methods in yeast genetics. A laboratory manual (Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press).

Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys *23*, 327-341.

Shen, M.Y., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. Protein Sci *15*, 2507-2524.

Stevenson, B.J., Rhodes, N., Errede, B., and Sprague, G.F., Jr. (1992). Constitutive mutants of the protein kinase STE11 activate the yeast pheromone response pathway in the absence of the G protein. Genes Dev *6*, 1293-1304.

Tatebayashi, K., Yamamoto, K., Tanaka, K., Tomida, T., Maruoka, T., Kasukawa, E., and Saito, H. (2006). Adaptor functions of Cdc42, Ste50, and Sho1 in the yeast osmoregulatory HOG MAPK pathway. Embo J *25*, 3033-3044.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191.

Wiederstein, M., and Sippl, M.J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic acids research *35*, W407-410.

Wu, C., Jansen, G., Zhang, J., Thomas, D.Y., and Whiteway, M. (2006). Adaptor protein Ste50p links the Ste11p MEKK to the HOG pathway through plasma membrane association. Genes Dev *20*, 734-746.

Wu, C., Leberer, E., Thomas, D.Y., and Whiteway, M. (1999). Functional characterization of the interaction of Ste50p with Ste11p MAPKKK in Saccharomyces cerevisiae. Mol Biol Cell *10*, 2425-2440.

Xiang, S., Kim, E.Y., Connelly, J.J., Nassar, N., Kirsch, J., Winking, J., Schwarz, G., and Schindelin, H. (2006). The crystal structure of Cdc42 in complex with collybistin II, a gephyrin-interacting guanine nucleotide exchange factor. Journal of molecular biology *359*, 35-46.