

# Supporting Information

Lohse et al. 10.1073/pnas.1410110111

## SI Results

**Minor-Groove Interactions.** To assess the relative importance of the various minor-groove protein–DNA interactions, we determined DNA-binding affinities of the WOPR domain [named for the members Wor1 (white-opaque regulator 1), Pac2, and Ryp1] mutated at the relevant position (Fig. 4, Fig. S2, and Table S3). Mutation of the hydrogen bond-forming residues Arg35 or Arg62 to alanine decreased affinity for DNA *in vitro* by ~100-fold (Table S3). The severe effects of these mutations indicate that a base-specific minor-groove interaction is critically important for YHR177w binding and that the interactions with the DNA backbone play a significant role in this process. We note that the Arg62-to-Glu charge-swap mutation also had an ~100-fold effect on YHR177w's affinity for DNA, whereas the Arg62-to-Lys mutation had a roughly 15-fold effect (Table S3). These results suggest that both the charge and the van der Waals interactions from the aliphatic portion of the arginine side chain contribute to the high affinity for DNA shown by this residue.

The YHR177w–minor-groove interactions comprise the TTWR portion of the WOPR *cis*-regulatory sequence. Using the 3DNA suite of programs (1, 2) and established criteria for a narrow minor groove (3), we determined the width of the minor groove at the R62 binding site (between bT13 and bA14) to be 3.7 Å (Table S5). After this site, the width of the minor groove expands to 8.1 Å near the loop between R62 and R73 (Table S5). More than half of the interface between the WOPR domain and DNA is formed from the interaction of the DNA with the 12 residues of the backbone-binding loop, but there are relatively few base pair-specific interactions among these residues. The preference for A or T in this region (i.e., as the nearly invariant aT7/bA14 position adjacent to the Arg62–bT13 interaction) may stem from local DNA structure considerations related to the modulation of minor-groove width, in addition to limited protein–DNA interactions of a sequence-specific nature.

**Major-Groove Interactions.** Mutations of the major-groove hydrogen bond-forming residues Arg73 and Arg185 to alanine resulted in a 10- to 30-fold decrease in affinity for DNA (Fig. 4, Fig. S2, and Table S3). Mutation of the backbone-interacting residue Ser72 produced a similar decrease in affinity for DNA. Alanine mutations of Leu79 and Tyr81, the two residues interacting with the major groove at aA10 and aG11, reduced affinity by 10- to 20-fold, but alanine mutations of Ser122 and His139, which interact with the same region, had no detectable effect on affinity. Removing the C-terminal loop ( $\Delta$ 181–201 aa from YHR177w) which passes through the major groove compromised affinity for DNA to a greater extent than any of the point mutations in that region (Table S3).

## SI Materials and Methods

**Plasmid and Strain Construction.** Lists of plasmids, strains, and oligonucleotides used in this study can be found in Table S4.

Construction of the pLIC-H3 YHR177 6–201 expression construct has been reported previously (4). Protein expression of YHR177w mutants, Wor1 truncations, and chimeras were performed using the pLIC-H3 derivative of pET28 (5) with constructs inserted between the XmaI and XhoI restriction sites. Chimeras and protein truncations were constructed using two rounds of PCR amplification with primers corresponding to the regions to be eliminated or to the points where parts of two genes were linked. Single-amino acid mutations in YHR177w were introduced using two rounds of PCR amplification with primers

corresponding to the codon to be mutated. To produce the selenomethionine-substituted protein, two additional methionine residues were introduced at leucines 93 and 135 (corresponding to L95 and L137 in wild-type YHR177w) using the techniques described above.

Wor1 ectopic expression plasmids for *Candida albicans* used the previously reported pADH33 system containing the inducible *MET3* promoter and the *SAT1* selectable marker (6) inserted between the XmaI and HindIII sites. Ectopic expression in *Saccharomyces cerevisiae* used the p413 TEF plasmid system (7) with the Wor1 linker truncations introduced between the SpeI (NheI) and XhoI restriction sites.

All work in *S. cerevisiae* used the previously reported  $\Delta$ mit1 $\Delta$ yrh177 strain (4, 8). This strain was transformed with previously reported Ura-tagged LacZ reporter plasmids (8) and His-tagged ectopic expression plasmids. Work in *C. albicans* used a previously reported a/a strain (6, 9). Ectopic expression strains, as well as empty control strains, were constructed by transforming with the linearized pMET3 expression constructs described above. Integration of these constructs at the proper locus was confirmed by colony PCR.

**White–Opaque Switching Assays.** Quantitative plate-based white–opaque switching assays using the pMET3 ectopic expression system were performed as previously described (10, 11). Strains were grown for 5–7 d at room temperature on repressing medium (+Met +Cys). Colonies were resuspended in water and then were plated on repressing or inducing (–Met –Cys) medium at room temperature. Plates were scored for colony phenotypes (white or opaque) after 7 d. In addition, single-cell morphology was determined for representative colonies of each phenotype. A strain transformed with an empty version of pADH33 was used as a negative control. Microscopic images were taken using a Zeiss Axiovert 200M microscope (Carl Zeiss).

**$\beta$ -Galactosidase Assays.**  $\beta$ -Galactosidase assays were performed using a standard protocol (12). Strains were grown in selective medium overnight, diluted back, and allowed to regrow to log phase before harvesting. Data in Fig. 6C are from the same day and represent the average of three replicates unless otherwise noted.

**EMSA.** Protein was purified for use in EMSA as previously described (4). Expression was for 4 h at 25 °C with 0.4 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) in the BL21 background. Pellets were lysed, and protein was purified using Ni-NTA agarose (Qiagen). Protein concentration was determined by comparison with BSA standards on a SDS/PAGE gel as well as by UV absorbance when possible. EMSAs were performed as previously described (4, 8), except that sodium chloride was reduced to 50 mM and Poly(dIdC) was omitted. In all assays, 0.4 nM of labeled DNA was used.

Quantification of binding affinity ( $k_d$ ) was determined based on EMSAs using ImageQuant 5.1 (GE Healthcare). ImageQuant was used to determine the intensities of the unshifted radiolabeled probe (unbound) as well as the entire lane (all probes). The fraction of bound probe was set as total lane minus the unbound probe. Using the known protein concentration and the fraction of probe bound, a curve was fit in GraphPad Prism using a quadratic equation for near stoichiometric binding. Unless otherwise noted in Table S3, affinity was determined based on a curve with the protein concentrations 4, 40, 200, and 1,000 nM. Although these concentrations are above the  $k_d$  for wild-type YHR177w

6–201, this approach produced  $k_d$  values for WT YHR177w 6–201 similar to those for curves with lower protein concentrations or with a larger number of points.

**Protein Purification for Crystallization.** Expression of YHR177 for crystallography was for 20 h at 25 °C with 0.4 mM IPTG in the BL21 background. Cells were lysed using an Avestin Emulsiflex homogenizer (Avestin) followed by purification on a Ni-NTA agarose column. Cells were dialyzed overnight against 20 mM Hepes (pH 7.5), 150 mM sodium chloride, 2 mM DTT. During this step, the 8xHis tag was cleaved by PreCission Protease (GE Healthcare). Cleaved protein then was purified on a HiTrap SP FF column (GE Healthcare) using a sodium chloride gradient of 300–600 mM.

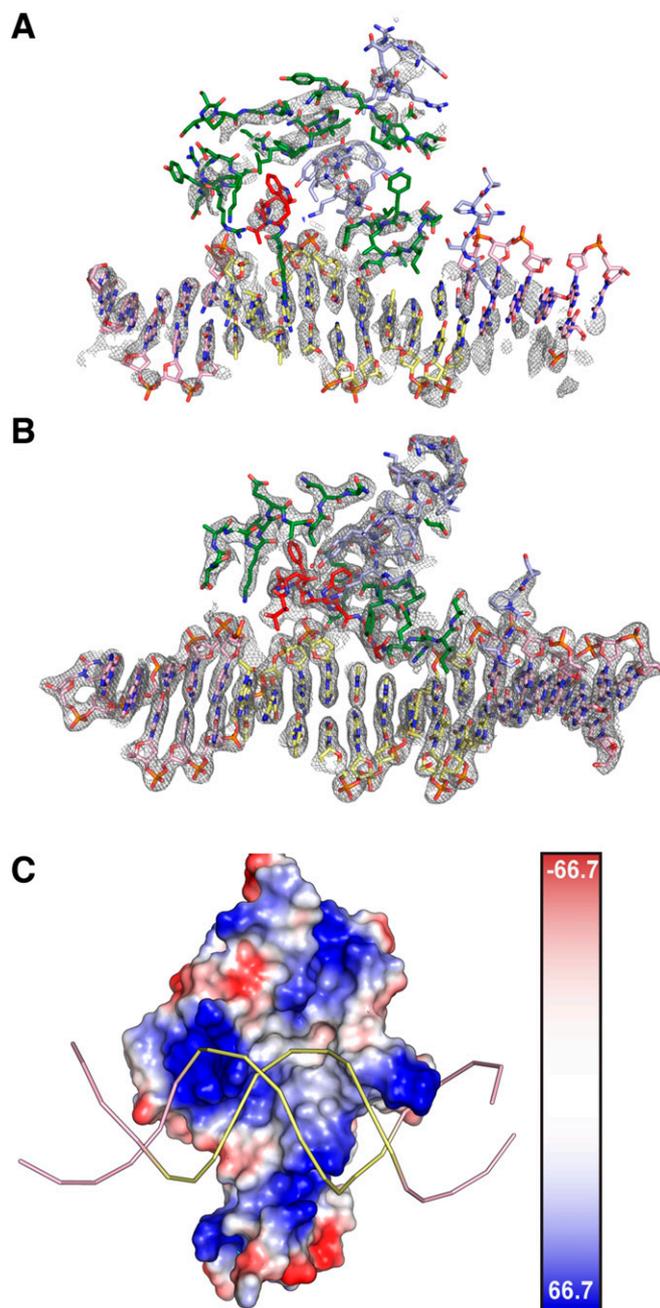
Oligonucleotides for crystallization were resuspended at ~2 mM in 20 mM Tris (pH 7.4), 50 mM sodium chloride, 1 mM Tris (2-carboxyethyl)phosphine (TCEP). Exact concentrations were determined based on UV absorbance. Corresponding nucleotides were mixed at a 1:1 ratio, heated for 5 min at 95 °C, and then allowed to cool to room temperature to anneal the strands together.

For the DNA1 fragment, protein and annealed dsDNA were mixed at a 1:1.1 molar ratio and then purified over a Superdex 200 10/300 GL column (GE Healthcare) into 20 mM Tris (pH 7.4), 50 mM sodium chloride, 1 mM TCEP. After concentration, the protein concentration was determined using the Bradford Colometric Assay. For the DNA2 fragment, the protein was purified over the Superdex 200 10/300 GL column by itself, requantified, and then mixed at a 1:1.05 molar ratio with dsDNA.

Expression of the seleno-Met YHR177w mutant [L93M, L135M] was conducted in M9 minimal medium supplemented with glucose, MgSO<sub>4</sub>, FeSO<sub>4</sub>, and vitamins. Fifteen minutes before induction, an amino acid mixture was added to stop methionine biosynthesis. Expression was for 16 h at 25 °C, and purification followed the methods described above.

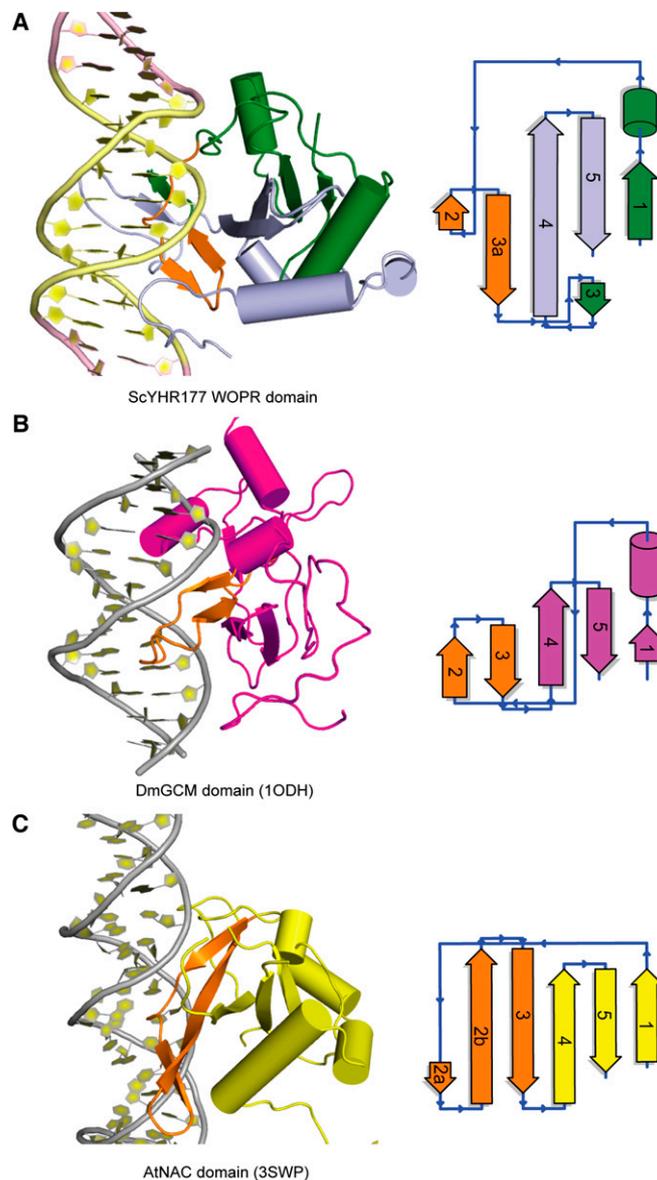
**Structural Analysis and Bioinformatics.** Primary sequence-based searches for WOPR homologs were conducted using the BLAST (13) and PSI-BLAST (14) utilities against the sequenced genomes available on the National Center for Biotechnology Information website. Primary sequence-based searches for structurally related proteins to the WOPR family were conducted using the PHYRE2 server (15).

- Zheng G, Lu XJ, Olson WK (2009) Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Res* 37(Web Server Issue):W240–246.
- Lu XJ, Olson WK (2008) 3DNA: A versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat Protoc* 3(7):1213–1227.
- Rohs R, et al. (2009) The role of DNA shape in protein-DNA recognition. *Nature* 461(7268):1248–1253.
- Cain CW, Lohse MB, Homann OR, Sil A, Johnson AD (2012) A conserved transcriptional regulator governs fungal morphology in widely diverged species. *Genetics* 190(2): 511–521.
- Hammon J, Palanivelu DV, Chen J, Patel C, Minor DL, Jr. (2009) A green fluorescent protein screen for identification of well-expressed membrane proteins from a cohort of extremophilic organisms. *Protein Sci* 18(1):121–133.
- Lohse MB, et al. (2013) Identification and characterization of a previously undescribed family of sequence-specific DNA-binding domains. *Proc Natl Acad Sci USA* 110(19): 7660–7665.
- Mumberg D, Müller R, Funk M (1995) Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 156(1):119–122.
- Lohse MB, Zordan RE, Cain CW, Johnson AD (2010) Distinct class of DNA-binding domains is exemplified by a master regulator of phenotypic switching in *Candida albicans*. *Proc Natl Acad Sci USA* 107(32):14105–14110.
- Lohse MB, Johnson AD (2010) Temporal anatomy of an epigenetic switch in cell programming: The white-opaque transition of *C. albicans*. *Mol Microbiol* 78(2):331–343.
- Zordan RE, Miller MG, Galgoczy DJ, Tuch BB, Johnson AD (2007) Interlocking transcriptional feedback loops control white-opaque switching in *Candida albicans*. *PLoS Biol* 5(10):e256.
- Zordan RE, Galgoczy DJ, Johnson AD (2006) Epigenetic properties of white-opaque switching in *Candida albicans* are based on a self-sustaining transcriptional feedback loop. *Proc Natl Acad Sci USA* 103(34):12807–12812.
- Guthrie C (1991) *Guide to Yeast Genetics and Molecular Biology*, ed Fink GR (Elsevier Academic, San Diego, CA), Vol 194.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410.
- Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25(17):3389–3402.
- Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: A case study using the Phyre server. *Nat Protoc* 4(3):363–371.



**Fig. S1.** Examples of the experimental and calculated density maps. (A) Density-modified experimental map (contoured at  $1.0 \sigma$ ) (gray) generated from a multiwavelength anomalous diffraction experiment done on a selenomethionine-derived crystal. (B) View of a 2Fo-Fc map (contoured at  $1.0 \sigma$ ) calculated with data from the P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> crystal form (gray). (C) Color plot of the vacuum electrostatics of the WOPR domain, as generated in Pymol. Blue represents a positive charge.





**Fig. S3.** Comparison with other transcription factors that bind DNA using a  $\beta$ -hairpin motif. Several transcription factors with weak matches to YHR177w were identified using the DALI algorithm. Structural alignments (*Left*) and simplified topology diagrams (*Right*) are shown for (A) YHR177w, (B) DmGCM (z-score 4.7) [Protein Data Bank (PDB) ID code 1ODH (1)], and (C) AtNAC (z-score 3.9) [PDB ID code 3SWP (2)]. These three proteins all use a  $\beta$ -hairpin to bind to the major groove of DNA and share a common  $\beta$ -sheet topology. The  $\beta$ -hairpins used to bind DNA are indicated in orange. The GCM and NAC proteins bind exclusively to the major groove and do not share the key minor-groove interactions seen in the WOPR domain.

1. Cohen SX, et al. (2003) Structure of the GCM domain-DNA complex: A DNA-binding domain with a novel fold and mode of target site recognition. *EMBO J* 22(8):1835–1845.
2. Welner DH, et al. (2012) DNA binding by the plant-specific NAC transcription factors in crystal and solution: A firm link to WRKY and GCM transcription factors. *Biochem J* 444(3): 395–404.

#### Table S1. Data collection and refinement statistics

##### [Table S1](#)

Statistics for the highest-resolution shell are shown in parentheses.

**Table S2. WOPR protein–DNA interactions, as determined by the PISA server (1)**[Table S2](#)

Interactions are separated into hydrogen bonds and residues with buried surface area.

1. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797.

**Table S3. Effects of WOPR domain single-point mutants, truncations, and other mutations on DNA-binding affinity**[Table S3](#)**Table S4. Plasmids, strains, and oligonucleotides used in this study**[Table S4](#)**Table S5. Calculated major- and minor-groove widths**[Table S5](#)

This table lists the direct P–P distances in the DNA backbone steps, as calculated in the program 3DNA. The distances are transformed to be equivalent to those calculated by the program curves (1). The critical minor-groove-binding residue, R62, binds at the seventh position in the chart, at a narrowing of the groove that is followed by an abrupt widening.

1. Rohs R, et al. (2009) The role of DNA shape in protein-DNA recognition. *Nature* 461(7268):1248–1253.