Architecture of a single membrane spanning cytochrome P450 suggests constraints that orient the catalytic domain relative to a bilayer

Brian C. Monka,b,1 Thomas M. Tomsiakc,1 Mikhail V. Keniyanb,1 Franziska U. Huchmannan,b,1,2 Joel D. A. Tyndalld,1 Joseph D. O’Connell IIIc, Richard D. Cannonb,2, Jeffrey G. McDonaldd, Andrew Rodriguezc, Janet S. Finer-Mooree, and Robert M. Stroudc,2

*Sir John Walsh Research Institute, aDepartment of Oral Sciences, Faculty of Dentistry, and bSchool of Pharmacy, University of Otago, Dunedin 9054, New Zealand, cDepartment of Biochemistry and Biophysics, University of California, San Francisco, CA 94158; and dDepartment of Molecular Genetics, University of Texas Southwestern Medical Center, Dallas, TX 75390

Contributed by Robert M. Stroud, January 10, 2014 (sent for review September 30, 2013)

Bitopic integral membrane proteins with a single transmembrane helix play diverse roles in catalysis, cell signaling, and morphogenesis. Complete monospanning protein structures are needed to show how interaction between the transmembrane helix and catalytic domain might influence association with the membrane and function. We report crystal structures of full-length Saccharomyces cerevisiae lanosterol 14α-demethylase, a membrane monospanning cytochrome P450 of the CYP51 family that catalyzes the first postcyclization step in ergosterol biosynthesis and is inhibited by triazole drugs. The structures reveal a well-ordered N-terminal amphipathic helix preceding a putative transmembrane helix that would constrain the catalytic domain orientation to lie partly in the lipid bilayer. The structures locate the substrate lanosterol, identify putative substrate and product channels, and reveal constrained interactions with triazole antifungal drugs that are important for drug design and understanding drug resistance.

Membrane proteins that span the lipid bilayer once constitute around 50% of all integral membrane proteins (1). Although monospanning membrane proteins carry out numerous key biological functions, including environmental sensing, organelle-specific catalysis, and the regulation of cell morphology, only individual domains or subdomains are currently represented in the Protein Data Bank, and structural information about interactions between their transmembrane domains and extramembranous components is lacking. Cytochrome P450 enzymes are prominent enzymes with orthologs found in all kingdoms of life. In eukaryotes, microsomal members of this major family of mixed-function mono-oxygenases contain a single transmembrane helix and can be grouped in two broad functional categories: biiodefense, such as the first phase of xenobiotic detoxification, and core metabolism including reactions in sterol biosynthesis and fatty acid oxidation (2).

The lanosterol 14α-demethylases or CYP51 enzymes, probably the most genetically ancient of the cytochrome P450 families, play a central role in cholesterol or ergosterol biosynthesis (3). CYP51s carries out three consecutive mono-oxygenase reaction cycles to remove the 14α-methyl group from lanosterol to yield 4,4-dimethyl-cholesta-8,14,24-trienol, a key precursor in cholesterol and ergosterol biosynthesis, releasing water and formic acid (3). Because of the key roles that CYP51s play in yeast, filamentous fungi, and some parasitic protozoa, these enzymes are therapeutic targets for antifungal and antibacterial agents, including fluconazole (FLC), voriconazole (VCZ), and itraconazole (ITC) (4). Fungal infections play an increasingly significant role in disease, impacting agriculture ecosystems and human health, especially in immuno-compromised individuals (5–7) for whom antifungal resistance continually poses a threat (8). In humans CYP51 is being tested as a target for cholesterol-lowering drugs (9) and in antiangiogenic cancer therapies (10). A limited set of cytochrome P450 isoforms (1A2, 2C8, 2C9, 2C19, 2D6, and 3A4) metabolize about 75% of the drugs currently used in medicine (11).

Membrane-bound cytochrome P450s catalyze reactions involving many hydrophobic substrates, including lipophilic drugs and steroids that partition from the lipid bilayer into the active site. Identification of enzyme orientation relative to the lipid bilayer therefore is critical for understanding substrate entry into and egress from the heme-containing active site (12). For the eukaryotic cytochrome P450 enzymes that contain transmembrane segments, most X-ray structures have been obtained after removal of the N-terminal transmembrane domain to facilitate expression and crystallization (13–18), as is the case with many other monospanning membrane protein structures. Resolved CYP51 structures include the soluble enzyme from Mycobacterium tuberculosis (19) and N-terminal truncated enzymes from Homo sapiens (20), Trypanosoma cruzi (21), Trypanosoma brucei (22), and Leishmania infantum (23). The available eukaryotic cytochrome P450 structures show the catalytic subunit has a conserved fold surrounding a central prostatic heme group. A proposed substrate channel to the membrane as well as interactions with

**Significance**

The absence in the Protein Data Bank of full-length structures of bitopic membrane proteins with one transmembrane helix, probably because of difficulties with ordered crystallization, has limited understanding of how single-transmembrane helices orient enzymes and sensors at the bilayer surface. X-ray crystal structures of full-length yeast lanosterol 14α-demethylase, a cytochrome P450, show how a helix spanning a single transmembrane domain may lead to constraints on the orientation of the putative substrate entry portal from within the bilayer. The crystal structures also locate the substrate lanosterol, identify putative substrate and product channels, and reveal constrained interactions with triazole antifungal drugs that are important for drug design and understanding the drug resistance associated with orthologs of the enzyme found in fungal pathogens.


The authors declare no conflict of interest.

Data deposition: The structures reported in this paper have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 4LXU (ScErg11p-6xHis + lanosterol) and 4K0F (ScErg11p-6xHis + itraconazole)].

See Commentary on page 3659.

1B.C.M. and T.M.T. contributed equally to this work.

2To whom correspondence should be addressed. E-mail: stroud@msg.ucsf.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1324245111/-/DCSupplemental.
several inhibitors, including ketoconazole, posaconazole, FLC, and substrate analogs such as 14c-methylenecyclopropyl-Δ7–24,25-dihydrolanosterol, have been visualized (19, 20, 22, 24).

Although bacterial and mitochondrial orthologs of cytochrome P450s and the M. tuberculosis CYP51 are soluble enzymes, bioinformatic analysis of many microsomal cytochrome P450s led to the detection of a hydrophobic surface in the catalytic domain and prediction of an N-terminal monospanning transmembrane domain that has been shown to be important for subcellular targeting and membrane insertion (25–28). Multiple in vitro approaches have suggested that a transmembrane anchor is not always required for the catalytic activity of some microsomal preparations of cytochrome P450s (29, 30). For example, M. tuberculosis CYP51 specifically converts detergent-solubilized lanosterol to its demethylated product (31). However, the in vivo role of the membrane anchor in catalysis by other CYP51s that use membrane-associated lanosterol has yet to be established. Antibody-binding accessibility assays showed that most of the surface of CYP2B4 (32) was solvent exposed with the exception of the putative transmembrane helix and the loop between the F and G helices. About 30% of the catalytic domain of CYP11A1 was inaccessible to chemical modification by fluorescent probes (33). The physical displacement of phospholipids in Langmuir–Blodgett monolayers indicated a considerably greater area of protein insertion than could be accounted for by a single N-terminal transmembrane helix (34). This finding was supported by atomic force microscopy, which showed the catalytic domain protrudes only 3.5 ± 1 nm above the membrane (35). A crystal structure of full-length aromatase at 2.9-Å resolution has been modeled, but weak electron density meant that the structures of 44 N-terminal and seven C-terminal residues could not be determined (36). Models that place the mouth of the substrate channel in direct contact with the hydrophobic environment of the lipid bilayer required the catalytic domain helices A′ and A to be imbedded and partially imbedded, respectively, in the lipid bilayer (28, 37).

**Results**

**Architecture of the Cytochrome P450 Saccharomyces cerevisiae Lanosterol 14α-Demethylase.** We report X-ray crystal structures of full-length Saccharomyces cerevisiae lanosterol 14α-demethy lase (ScErg11p-6xHis) with the membrane domain structurally discernable and two ligands cocrystallized: the substrate lanosterol and the triazole inhibitor ITC. Clear electron density for the 50 N-terminal amino acids revealed two helices oriented at ∼60° to each other (Fig. 1A and B). The N-terminal helix [membrane helix 1 (MH1), residues 6–23] is amphipathic and makes extensive crystal contacts with symmetry-related molecules (Fig. 2A and B). The distribution of hydrophobic and hydrophilic side chains suggests that the hydrophobic face may contact the inner leaflet of the endoplasmic reticulum (ER) membrane (Fig. 1B). MH1 connects to TMH1 via a short proline-containing turn (residues 24–26), a slightly kinked helix 37.5 Å long (residues 27–51) that is of sufficient length to traverse the lipid bilayer. The side chains of amino acids R30 and Q29, adjacent to the turn, form a C-terminal cap for helix MH1 and an N-terminal cap for TMH1, respectively.

Seventeen of the 25 residues in TMH1 are hydrophobic, with a high concentration in the N-terminal portion, as is consistent with immersion of TMH1 in the lipid bilayer. Beyond the proline kink at P28, TMH1 makes extensive polar contacts with the catalytic domain (Fig. 1C). The helix then consists of opposing hydrophilic and hydrophobic faces that interact with the catalytic domain and lipid bilayer, respectively. An ionic interaction between the guanidinium group of R52 and the carboxylate of D54 enables a sharp turn between TMH1 and the first loop of the catalytic domain.

Paired polar residues constrain the orientation between TMH1 and the catalytic domain. The ε-amide group of Q46, which hydrogen bonds with the main-chain carbonyl and amide of Y61, is absolutely conserved in the fungal CYP51 family. The hydroxyl of Y61 also makes a hydrogen bond with the δ-amide group of N42. Water molecules bridge the ε-amide group of Q46 with the main-chain amide oxygen of L58 and the ε-hydroxyl of Y49 with the hydroxyl of S92. Other polar contacts in this region
include the guanidinium group of R55 and the carbonyl of A399; the main-chain amide of K53 and the hydroxyl of S397; and the main-chain carbonyl of D54 and the amide of A399. These extensive contacts bury a surface area of 51.8 Å² on TMH1 and appear to form a rigid architecture that is important for establishing the pose of the enzyme in the bilayer (Figs. 1A and 2A and C). The energetic stabilization conferred by the extensive contacts between TMH1 and the catalytic domain make it unlikely that the crystal contacts between the amphipathic MH1 (residues S6, E10, E13, and H21) and residues K270 and S267 of one symmetry-related molecule and G532 and E354 of the other (Fig. 2B) would have distorted the overall structure of ScErg11p significantly. The ScErg11p-6xHis structures show few significant differences between structures, with the most extensive changes (0.36 Å rmsd on Cα) being in the F/G loop (Fig. 2C) of the lanosterol and ITC costructures.

**Inferred Orientation of ScErg11p Relative to the Membrane.** To establish the membrane orientation, ScErg11p-6xHis was inserted into preformed lipid vesicles and labeled from the outside at free carboxylates with glycinamide in the presence of N,N′-Dicyclohexylcarbodiimide (DCCD) at pH 6 (Fig. 3A) (38). Tandem mass spectrometry of tryptic digests identified labeled residues only on the catalytic domain, with no labeling on the N-terminal MH1 (residues 1–26), as is consistent with topological labeling of other cytochrome P450s and with MH1 and the catalytic domain associating with opposite sides of the bilayer.

The orientation of ScErg11p also was explored using two algorithms for membrane insertion: the all-atomistic positioning of protein in the membrane (PPD) procedure (39) and the knowledge-based Ex3d procedure (40). Both calculated poses were very similar to those inferred from our crystallographic analysis, i.e., TMH1 residues 27–50 link the N-terminal amphipathic helix MH1 and the C-terminal catalytic domain that are partially submerged in opposite leaflets of the lipid bilayer (Fig. 4A). The OPM analysis, which calculated a partitioning free energy of −31.3 kcal/mol into the lipid bilayer, was the best fit with our topological labeling, including the labeling of residue E249 on the F/G loop, and was used in this analysis (Fig. 3).

The pose we infer for ScErg11p relative to the membrane buries portions of the loop connecting helices F–F′ and G in the lipid bilayer (Figs. 1D, 2C, and 3). It is consistent with mutagenesis experiments showing this region is important for membrane binding (33, 41, 42), fluorescence quenching analysis (5), the predicted pose of aromatase (36), and computational simulations using Delphi software (44) reveal an electrostatic potential surface showing a distinct hydrophobic/hydrophilic boundary (Fig. 3B). Ordered water molecules are present only in proposed solvent-accessible areas of the model and are absent from the bilayer-accessible surface of the structure (Fig. 1B and C). As a corollary to the positioning of the lipid bilayer, the protein crystallized with alternating layers of proposed membrane and soluble segments (Fig. 2A) corresponding to the placement of hydrophobic and hydrophilic elements in the crystal structure, including the surface-associated ligands.

We hypothesize that a cluster of charged residues found at the N-terminal end of the G helix forms a major membrane-binding contact with the negatively charged phosphate head groups of the lipid bilayer (Fig. 1D). This orientation places the end of a substrate channel in direct contact with the cytoplasmic surface of the lipid bilayer (Figs. 1A, 2C, 3B, and 4C). This substrate channel is bifurcated near its mouth, yielding a secondary vestibule (Fig. 3C) not previously identified in CYP51s. Although an N-decyl-D-maltoside (DM) detergent micelle differs from an intact lipid bilayer, the proposed orientation supports a model in which the substrate first is abstracted directly from the surface of the membrane via the juxtaposed channel mouth. After heme reduction and complex formation with oxygen, an oxygen-bound mechanism converts the C14 methyl to the first of three products that result in 4,4-dimethyl-cholesta-8,14,24-trienol (2).

**Substrate and Product Binding.** The enzyme was purified and cocrystallized with the substrate lanosterol or with four other ligands (itraconazole, estriol, FLC, or VCZ). The lanosterol (Fig. 4A) and estriol (Fig. S1A) cocystal structures showed density consistent with the common four-ring sterol in the active-site cavity. A mixed state probably involving both endogenous lanosterol and added ligand complicated interpretation of electron density in the active site. This mixed state did not occur when ScErg11p was purified in the presence of ITC and crystallized. As a result, only the lanosterol and ITC structures are detailed here with the four others reported as models in SI Discussion.

The lanosterol (Fig. 4A) and estriol (Fig. S1A) cocystal structures showed density consistent with the common four-ring sterol in the active-site cavity. Electron density consistent with the tail of lanosterol is located in the primary vestibule that leads up to the active site (Fig. 5B). Similar electron density was observed in the same locations when no ligand was supplied (Fig. S1D), indicating that significant amounts of substrate copurify with the enzyme. HPLC-MS in positive ion mode (45) showed that lanosterol copurifies with ScErg11p-6xHis in the absence of added ligand, supporting this interpretation (Fig. S1E).

Electron density proximal to the heme iron, consistent with the presence of an oxygen molecule coordinated to the heme iron, fits with the ferrous dioxo intermediate predicted for the CYP51 reaction cycle. This feature also was seen in the crystal
structure without added ligand, which also had bound lanosterol (Fig. S1D). To interpret this peak, the electron density near the heme was assessed using crystallography difference-map refinement. Fo–Fc maps calculated after several rounds of refinement showed that a single water (10 electrons) would generate insufficient positive density, whereas imidazole (32 electrons) would give a strong negative density peak. Placing O2 (16 electrons) gave a flat Fo–Fc electron density map consistent with O2 being at the heme site. Although unusual for P450s, this result is not without precedent (46) and may be caused by reduction after X-ray radiolysis (47) or purification of ScErg11p-6xHis trapped in the oxygen-bound state. The distance of ∼8 Å between oxygen and C14 methyl is consistent with a lanosterol binding in a pre-catalytic rather than a catalytically active state. Visual inspection of the electron density maps and a high real-space R-factor indicated the need for considerable care in placement of the ligand in the lanosterol costructure. Minimally biased Fo–Fc electron density omit maps, calculated with lanosterol omitted to place the lanosterol correctly, resulted in a ligand placement with reasonable correlation coefficient (0.813). Bound lanosterol is oriented differently from other lanosterol-like molecules in related CYP51s (24), and the electron density offers no evidence of these alternate orientations.

Density consistent with a second lanosterol-like molecule projects from the secondary vestibule (Fig. 5 A and B). In all ScErg11p structures examined, a hydrophobic tail fills the narrow channel from the secondary vestibule to the surface of Erg11p, whereas a multiring head that projects into solution appears to lack the C14 methyl group. Lack of density for the C14 methyl group suggested that the lanosterol-like molecule was the reaction product 4,4-dimethyl-cholesta-8,14,24-trienol or a subsequent metabolite that likely copurified with the enzyme. A bond links the heme–oxo complex. (B) ITC in a coordination complex with active-site heme. The color scheme is as in A, with ITC carbons colored magenta. Electron density is depicted as an Fo–Fc map contoured to 2.5σ with the ligand density omitted from the calculation. (C) Conserved amino acids in fungi that are commonly mutated in antifungal resistance are depicted with carbon atoms colored orange. Amino acids conserved in S. cerevisiae and in the main fungal pathogens of humans but not in human CYP51, and not mutated, are depicted with carbon atoms colored green. Amino acids that are similar only in the pathogenic fungi are shown in cyan. Other atoms are colored as in A.

Discussion

Azole Binding and Antifungal Resistance. Our structures identify important ligand-binding constraints. The antifungal triazole drug ITC extends from the active site to just beyond the mouth of the entry channel, similar to posaconazole in the T. brucei CYP51 structure (49). The triazole group of the ITC makes a coordination bond with the heme iron. The space occupied by ITC (Fig. 4B) fits closely with that occupied by lanosterol and bound O2 (Fig. 4A), with the triazole head group displacing the O2 and the di-halogenated headgroup replacing the first sterol ring. The expected chlorine atoms on the dichlorophenyl group consistently showed strongly negative Fo–Fc density, suggesting that the chloro groups had been damaged through photoreduction upon X-ray data collection. As a result, the chloro groups have been refined to zero occupancy in the model. In confirmation of our findings, the triazole head groups in the smaller triazole drugs FLC and VCZ coordinate with the heme iron and displace the presumed O2, and their unmodified difluorinated head groups fill the space occupied by the first sterol ring of lanosterol and the sterol analog estriol (Fig. 4 A and B and Fig. S1A–C). The long tail of ITC fills the entry channel and appears to exclude all but one water molecule. The tail of ITC lies within a hydrophobic pocket lined by helices F and F′ (F236, P238, and F241), consistent with previous structures of truncated CYP51 family members (50), and by the residues L380HSL383 and F506TSMVS10 (Fig. 4B and see Fig. S4A).

**Fig. 4.** Lanosterol and ITC binding in ScErg11p. (A) Lanosterol is depicted with carbon atoms colored cyan and heme with carbon atoms colored yellow. Selected oxygen atoms are colored red, nitrogen blue, sulfur gold, and iron brown. Electron density is depicted as a simulated annealing 2Fo–Fc omit map contoured to 0.7σ with the ligand, heme, and oxygen omitted from the electron density calculation. A bond links the heme–oxo complex. (B) ITC in a coordination complex with active-site heme. The color scheme is as in A, with ITC carbons colored magenta. Electron density is depicted as an Fo–Fc map contoured to 2.5σ with the ligand density omitted from the calculation. (C) Conserved amino acids in fungi that are commonly mutated in antifungal resistance are depicted with carbon atoms colored orange. Amino acids conserved in S. cerevisiae and in the main fungal pathogens of humans but not in human CYP51, and not mutated, are depicted with carbon atoms colored green. Amino acids that are similar only in the pathogenic fungi are shown in cyan. Other atoms are colored as in A.
Although less than 24% of the amino acids in ScErg11p are conserved among fungi (50), there is an especially high concentration of conserved residues near the heme (Fig. S4A) and active-site cavity. These conserved residues include the absolutely conserved coordination of the C470 sulfur to the heme iron and the ionized K151, R385, and H468 side chains. The sequence G310LVMG314 in helix I is a consensus sequence (GXXGXL) for sterol (cholesterol) binding, with the carbonyl of G310 making the only polar interaction with lanosterol and a slightly kinked face toward ligand.

The structure of the active site and substrate channel may explain the susceptibility of some azole-resistant clinical isolates, including *Candida albicans* mutants F126L, Y132F/H, P230L, G307S, and F380S, *Aspergillus fumigatus* G345E/F and I301, *Cryptococcus neoformans* Neof145, and *A. fumigatus* Cyp51A, are in *S. cerevisiae* (Fig. S3C) is equivalent to Y132 in C. albicans, Y145 in C. neoformans, and Y136 in A. capsulatus. The absence of the heme-binding hydrogen bond in *S. cerevisiae* changes the orientation of the heme and reduce susceptibility to both FLC and ITC. In contrast, G73E/F mutations in ScErg11p, equivalent to G54E/F in *A. fumigatus* Cyp51A, are in the mouth of the substrate channel and might be expected to modify stability to the long-tailed triazoles ITC and posaconazole but not short-tailed VCZ.

Numerous azole-resistance mutations occur outside the active site. Significantly, the N22D azole-resistance mutation in TMH1 of *A. fumigatus* Cyp51A (ScErg11p N42D) supports the proposition that interactions between TMH1 and the catalytic domain can affect active-site function. CYP51s from *S. cerevisiae* and the fungal pathogens of humans (Fig. S4) contain a group of conserved/similar aromatic residues (F60, Y61, W62, Y77, Y77, F79, F80, Y87, F90, H381) bordered by TMH1 that form a compact, buried structure. In concert with neighboring conserved/similar hydrophobic residues (V59, L95, L96 L383, L407) and the polypeptide backbone (A69, G73, S382), this grouping supports a narrow entry channel into the active site and egress vestibules. The tight vestibules orient the lanosterol and may provide gated entry of demethylated reaction product into the egress vestibules. Finally, mutation conferring azole resistance at residues homologous to ScErg11p L95, F241, H381, and S382 has yet to be reported. These residues along with the chemically similar residue I139 could provide fungal-specific side-chain contacts for the design of novel antifungals directed against the major fungal pathogens of humans (Fig. 4C and Fig. S3F).

**Bitopic Membrane Proteins.** Full-length structures of bitopic membrane proteins with one transmembrane helix are absent in the Protein Data Bank, probably because of the difficulties with ordered crystallization. This deficiency has limited understanding of how single-transmembrane helices orient enzymes/sensors at the bilayer surface (1). The X-ray structure of ScErg11p-6xHis, a full-length fungal cytochrome P450 enzyme, shows that the transmembrane domain not only tethers the protein to the ER but also may orient it relative to the bilayer. The surprisingly rigid interaction between the transmembrane domain and catalytic domain may establish the pose required for catalysis and interaction with other protein partners such as Erg24p and Erg25p-Erg27p. With monospanning membrane proteins constituting about half of all membrane proteins, structural information supporting properties beyond tethering and localization in ScErg11p provides clues about constraining polar residues in other monospanning enzymes. It also may define a motif that may be a feature of other membrane-bound enzymes.

The ScErg11p structures identify lanosterol binding in the active site of a CYP51, a substrate channel linking to the lipid bilayer, and a secondary vestibule with a product exit channel. The model also more fully maps how triazole antifungals interact with a fungal cytochrome P450 to block catalysis and identifies possible interactions that confer azole resistance applicable to pathogenic fungal species. Like other CYP51 structures, the intact enzyme displays less conformational heterogeneity between structures than seen in many truncated ectodomain eukaryotic cytochrome P450s, suggesting that the transmembrane domain itself or associated hydrophobic molecules such as lipids or reaction product may decrease conformational heterogeneity and protect the active site from bulk water. The structures shown here may enable the development of new molecular models to facilitate drug design that targets fungal CYP51s or other eukaryotic cytochrome P450s and may provide a practical basis for the design of therapeutics with minimized off-target effects.

**Methods**

*S. cerevisiae* ScErg11p was overexpressed homogeneously as a 6xHis derivative (ScErg11p-6xHis) from the *PDR5* locus in the host yeast strain ADΔ (Table S1) (55). ScErg11p-GFP expressed in ADΔ was found in nuclear-associated ER (Fig. S4A). ScErg11p-6xHis was purified from strain MMLY941 by extraction of a crude membrane fraction with the detergent DM, Ni-NTA-agarose affinity chromatography, and size-exclusion chromatography (SEC) in the presence of 60 mM NaCl. During purification, lanosterol behaved as a detergent micelle-protein monomer during SEC, as is consistent with a molecular mass for the protein of ~62 kDa (Fig. S2C). It bound ergosterol and showed type II binding of known antifungals (Fig. S5D). The purified protein was authenticated by mass spectrometry of tryptic fingerprints at the Stanford University Mass Spectrometry facility (Fig. S5E).

Red crystals were obtained initially by hanging-drop vapor diffusion at 18 °C in 200-nL drops with a MemGold (Molecular Dimensions Limited) screen set using a Mosquito robot (TTP Labtech). Crystallization conditions were optimized, and large, red, boat-shaped crystals (100 μm) were obtained in 2- to 6-μL drops with the purified protein at ~20 μg/mL in 10% (wt/vol) glycerol, 150 mM NaCl, 4x critical micellar concentration DM, and 20 mM Hapes (pH 7.5) mixed 1:1 with 43-45% (v/v) PEG 400 in 100 mM glycine (pH 9.3-9.5) (Fig. S5F).

Complete datasets were obtained for crystals at the Berkeley Advanced Light Source BL 8.3.1 with a beam at wavelength 1.18Å (Table S2). Data were processed, integrated, and scaled using the HKL2000 (53) package. An initial model was determined by molecular replacement using the soluble domain of the HsCYP51 model 3LD6 (20) in Phaser (54). Refinement was performed using both Refmac5 in the CCP4 (55) suite of programs and Phenix (56), and model building was performed using Coot (57). Model validation was performed in Phenix and with Molprobity (58) and PROCHECK (59). Display of electrostatic maps was performed in Chimera (59).

**ACKNOWLEDGMENTS.** We thank James Holton and George Meigs at beamline 8.3.1 of the Advanced Light Source at Lawrence Berkeley National Laboratory and James Holton at the University of California, Berkeley for helpful discussions.
Laboratory; Bill Harries, Rebecca Robbins, and Erwin Lamping for support; Paul Ortiz de Montellano and John Cutfield for critical reading; and Avner Schlesinger for helpful computational suggestions. This research was supported by National Institutes of Health Grants U54GM094625, GM24485, and GM073710, the Marsden Fund of the Royal Society of New Zealand, and the Otago Medical Research Foundation.

Supporting Information

Monk et al. 10.1073/pnas.1324245111

SI Methods

Expression of Saccharomyces cerevisiae Lanosterol 14α-Demethylase-6xHis in Saccharomyces cerevisiae. Saccharomyces cerevisiae lanosterol 14α-demethylase (ScErg11p) was overexpressed as a C-terminal 6xHis derivative from the PDR5 locus in a host yeast strain (ADA) (1, 2) that is hypersensitive to xenobiotics because of the deletion of seven ATP-binding cassette (ABC) transporters, including pumps responsible for the efflux of azole drugs (Table S1). The strain used to express ScErg11p-6xHis (MMLY941) is multiply deleted of the Pleiotropic Drug Resistance (PDR) class of drug efflux pumps, the YOR1 drug efflux pump, and the PDR3 transcriptional regulator but retains a single copy of its endogenous ERI11 gene. It also contains the gain-of-function pdr1-3 mutation in the Pdr1p transcriptional regulator that renders transcription from the PDR5 promoter constitutive (3). Placement of the PDK1 transcription terminator and the URA3 gene immediately downstream of the ORF inserted at the PDR5 locus (4) enables termination of transcripts from the PDR5 locus and provides a selection marker, respectively. Control EGFP derivatives of the S. cerevisiae plasma membrane ABC transporter ScPdr3p (ScPdr3p-GFP) and the cytoplasmic Candida albicans orotidine-5'-phosphate decarboxylase (CaUra3p-GFP) expressed in ADA were visualized by confocal microscopy almost exclusively in the plasma membrane and the cytoplasm, respectively (Fig. S5A). The ScErg11p-GFP derivative appeared enriched in the endoplasmic reticulum (ER), particularly in a region surrounding or capping the DAPI-stained nucleus (Fig. S5A). Overexpressed proteins that normally localize to the ER often are concentrated in layers of the ER around much of the nucleus (5). The overexpressed ScErg11p-6xHis protein from strain MMLY941 was recovered in a crude membrane fraction, readily solubilized using the detergent N-decyl-β-D-maltoside (DM), and purified by Ni-NTA affinity chromatography and size-exclusion chromatography (SEC) (Fig. S5 B and C).

Preparation of Purified ScErg11p-6xHis. Twelve liters of MMLY941 cells were grown in 1-L volumes of YPD medium [1% yeast extract, 2% (wt/vol) peptone, and 2% (wt/vol) d-glucose] in baffled 3-L flasks at 30 °C with shaking at 200 rpm to ensure vigorous aeration. The cells were harvested by centrifugation when OD600nm = 18-20, were washed once with distilled water, and were resuspended to a final volume of about 250 mL in ice-cold 2% (wt/vol) glycerol, 250 mM NaCl, 20 mM Tris, 0.5 mM PMSF containing Roche minus EDTA protease inhibitors and adjusted to pH 8 at 4 °C by centrifugation at 92,400 × g in a Beckman Ti70 rotor for 70 min. The membranes (2.5 mg/mL protein) were extracted using the same buffer containing 16 mM (~10 times critical micelle concentration, cmc) DM for 1 h with stirring on ice. The DM-soluble fraction was obtained by centrifugation at 92,400 × g for 70 min at 6 °C. Two milliliters (1/50th volume) of 1-M imidazole (20 mM final), pH 8, was added to the supernatant, which then was mixed with a 2-mL packed volume of preequilibrated Ni-NTA-agarose resin and incubated overnight with mixing at 6 °C. The affinity matrix was preequilibrated by washing with 20 volumes of solubilizing buffer containing 16 mM DM and 20 mM imidazole. After overnight incubation with the detergent extract, the Ni-NTA-agarose affinity matrix was placed in a 20-mL Bio-Rad Econo-Column, was washed four times with 10 volumes (20 mL) of solubilizing buffer containing 16 mM DM and 20 mM imidazole, and then was eluted using eight 1-mL samples of the same buffer containing 200 mM imidazole. The affinity-purified fraction was enriched for a 62-kDa band detected by SDS/PAGE with the size predicted for ScErg11p-6xHis (Fig. S5B). The preparation gave an absorbance peak at 424 nm consistent with the type II binding of imidazole (6). In some cases 250 nmol of itraconazole (ITC) (125 μL, 2 mM in DMSO) was added to 5 mL of the pooled eluted protein. After 5 min on ice this preparation gave an absorbance peak at 422 nm, which was indicative of the displacement of the imidazole in the active site by the ITC. The bulk of non-heme-containing ScErg11p-6xHis was precipitated by standing overnight at 6 °C. The soluble supernatant recovered after centrifugation at 16,000 × g for 10 min was concentrated by centrifugal filtration to about 1.0 mL using a 50-kDa cutoff membrane. Centrifugation at 16,000 × g for 10 min or filtration through a 0.22-μm filter was used to remove residual insoluble material. Samples (0.5 mL) of the concentrated affinity-purified protein were separated by SEC using a Superdex 200 10/300 column (GE Healthcare) equilibrated and eluted with filtered and degassed 10% (wt/vol) glycerol, 150 mM NaCl, 6.4 mM DM (approximately 4× cmc), 0.5 mM PMSF, 1 Roche minus EDTA protease inhibitor tablet per 400 mL, 20 mM Hepes (pH 7.5) ± 2 mM ITC. The 62-kDa ScErg11p-6xHis protein coeluted with the heme as a single peak at about 14 mL (Fig. S5C) and then was concentrated by centrifugal filtration to >20 mg/mL in ~200 μL for spectral analysis (Fig. SSD) and crystallization.

Properties of ScErg11p-6xHis. SEC in the presence of DM ± ITC (2 μM) gave a heme-containing protein monomer/detent mutant micelle complex (Fig. S5C) with a relative migration consistent with a molecular mass of 96 kDa by comparison with molecular mass standards (BenchMark, Invitrogen). Excess DM was detected by refractive index measurements as a micelle with an apparent molecular size of 35 kDa during SEC. This result (96 – 35 kDa = 61 kDa) is consistent with the size expected for ScErg11p-6xHis (62 kDa). Before SEC or the addition of ITC, the affinity-purified protein had an absorbance peak at 424 nm. Preparations of the SEC-purified apo-enzyme obtained in the absence of ITC had an absorbance peak at 418 nm, whereas preparations obtained by SEC in the presence of ITC had an absorbance peak at 421 nm (Fig. 1D), as expected for type II binding of the azole drug (7). After centrifugal concentration to >20 mg/mL, precipitation of a colorless 62-kDa protein, which appeared to be ScErg11p-6xHis depleted of the heme, ceased once the 418 nm/280 nm or 421 nm/280 nm absorbance ratio reached ~0.65. The
purified soluble protein then could be stored for at least 1 mo at 4 °C. Mass spectrometry of tryptic fragments of the purified protein obtained from an SDS/PAGE gel slice gave ~70% coverage of the protein primary sequence (Fig. S5E).

Preparation of Proteoliposomes and Analysis of Amino Acid Accessibility. Azolectin (25 mg l-α-Phosphatidylcholine; Sigma) dissolved in 2 mL of chloroform was dried onto the surface of a Corex tube using a flow of nitrogen gas. The lipid film was emulsified in 5 mL of filtered and degassed 20-mM Hepes (pH 7.5) buffer by vortexing and was transferred to a 15-mL Falcon tube. The emulsion (2.5 mL) was sonicated on ice until optically clear. Samples of 0.5 mL were extruded 19 times using a minieextruder (Avanti Polar Lipids, Inc.) equipped with a polycarbonate membrane containing 100-nm pores. Samples (1.2 mL) of the resultant lipid vesicles were recovered by centrifugation at 100,000×g for 45 min using a Beckman mini-ultracentrifuge, were washed twice with 1 mL of the Hepes buffer, and were resuspended in the same buffer. A 20-μL sample of 23 mg/mL ScErg11p-6xHis containing 6 mM DM was added to 0.9 mL of the lipid vesicle preparation and incubated overnight at 6 °C to enable spontaneous proteoliposome formation. The preparation was centrifuged at 5,000 × g for 5 min to remove a small amount of precipitated ScErg11p-6xHis. The proteoliposomes then were pelleted at 100,000 × g for 45 min and washed twice by centrifugation in 1 mL of the Hepes buffer. The slightly colored pellet was resuspended in 1 mL 20-mM Hepes, pH 7.5, and stored on ice. Accessible acidic residues were glycinated at pH 5.0, and accessible lysine residues were acetylated at pH 9.0 according to the method of Izumi et al. (8) after washing in the appropriate buffer. The modified ScErg11p-6xHis residues in the treated, washed, and acetone-precipitated proteoliposomes were identified by mass spectrometry of tryptic fingerprints of SDS/PAGE-separated protein bands at Stanford University Mass Spectrometry facility.

SI Discussion

Structural Analysis. Use of the Constraint-based Multiple Alignment Tool (Cobalt; National Center for Biotechnology Information), together with some minor adjustment of the alignments by eye, allowed structural alignment of the primary sequences of ScErg11p with the Erg1pS from 10 major fungal pathogens of humans (C. albicans, Candida glabrata, Candida dubliniensis, Candida tropicalis, Cryptococcus neoformans, Aspergillus fumigatus, Coccidioides immitis, Pneumocystis carinii, Ajellomyces capsulatus, Malassezia globosa), including the Cyp51A and Cyp51B sequences from Aspergillus fumigatus. This alignment showed that only 24% of the amino acid residues in the structural alignment with ScErg11p were absolutely conserved, and a further 25% of the amino acid residues were similar i.e., fell into the groups STA, NEOK, NHOK, NDEQ, QHRK, MILV, HY, FWY, CSÅ, ATÅ, SAG, STNK, SPTA, SGND, SNDEQK, NDEQHK, NEQHRK, MILVF, HFHY Three separate views of the ScErg11p-6xHis structure (Fig. S4) show the distribution of conserved and similar amino acid residues for S. cerevisiae and the major fungal pathogens as red and white sticks, respectively. The most highly conserved regions, which contain the bulk of the conserved residues, surround and stabilize the heme, appear within the active-site cavity in support of enzyme function, contribute to the walls of the channel to the active site, occur in a strip running from the proximal surface of the heme beside helix C and toward the active vertex site, and are found in a buried region of aromatic groups that links between the C-terminal half of the transmembrane segment TMH1 and the active-site channel. The similar amino residues are found in regions enriched for absolutely conserved residues (and therefore are likely to be important for function) and also tend to be more widely dispersed throughout the structure and appear likely to contribute to the conservation of overall enzyme architecture. In contrast, the highest concentrations of dissimilar residues are found on the front and side surfaces of the catalytic domain and in membrane-embedded portions of the molecule that interact with the lipid bilayer. We propose that the strip of conserved residues beside helix C, which reaches toward the product egress channel and which includes a number of charged residues, is the site that enables productive interaction with NADPH-cytochrome P450 reductase. Previous studies have suggested a role for helix C in interactions with the reductase (9). We also propose that the region of conserved and similar aromatic groups (F59, F60, Y61, W62, Y72, Y77, F79, F80, Y87, F91, H381) surrounded by two concave β-sheet surfaces, the A and A’ helices, and one of the bifurcating faces of the substrate entry and product egress channel provides a compact, buried structure when it is bounded by TMH1. With the assistance of neighboring conserved and similar hydrophobic residues (V59, L95, L96, L383, L407) and the polypeptide backbone (A69, G73, S382), this structure contributes to a narrow entry channel into the active site and egress pathway. We hypothesize that this structure could help ensure an efficient oxidation reaction by orienting lanosterol in the active site, allowing entry of only the demethylated reaction product into the egress channel and blocking external water from traveling through the occupied entrance and exit channels to the active-site heme.

Analysis of the Catalytic Site and Entry and Egress Channels for Antifungal Discovery. A detailed knowledge of the active site, the substrate entry channel, and the product egress channel for ScErg11p, together with reactants and antifungal inhibitors, should expedite structure-directed drug design. The unambiguous placement of both the head groups and the detection of the conformation of the long tail of ITC in the crystal structure identifies putative anchor points in ScErg11p for a pharmacophore that could enable the design of antifungals with broad-spectrum activity against the dominant fungal pathogens and minimal activity against the human homolog HsCYP51. Amino acid residues within 6 Å of ITC are illustrated in Fig. S5A. These residues include the previously identified fungal substrate recognition sequences SRS-1, SRS-2, SRS-4, SRS-5, and SRS-6 (7). The helix between G67 and P76 and the loop between F92 and R96 also contribute amino acid residues within 6 Å of the ITC. The amino acid residues visualized in Fig. S5A can be categorized as conserved only in the S. cerevisiae and the major fungal pathogens of humans (red sticks), conserved in the fungal Erg11ps and HsCYP51 (pink sticks), similar only in fungi (white sticks), similar in fungal Erg11ps and HsCYP51 (gray sticks), dissimilar in fungi (blue sticks), or dissimilar in fungi but with polypeptide backbone carbonyl groups within 4 Å of ITC (sticks with atoms colored by element), i.e., within possible hydrogen-bonding distance of the antifungal. The ScErg11p X-ray structure shows that the amino acid residues L95, F241, H381, and S382 conserved in fungi and not in humans have side-chain atoms within 4.1 Å of ITC. In addition the backbone carbonyl but not the side chain of the similarly conserved M313 is within 4.5 Å of ITC. Other side chains within 6 Å of ITC are likely to confer lower specificity because they are conserved in the fungal Erg11ps and HsCYP51. These include Y126, L129, T130, F133, Y140, and F236, whereas G73 and G315 offer polypeptide backbone carbonyl contacts. I136 and F506 are similar in fungal Erg11ps but not in HsCYP51. I136 provides a side-chain contact, and F506 provides a backbone carbonyl contact. A69, Y72, L307, V311, T318, L380, and V510 are similar amino acid residues in fungal Erg11ps and HsCYP51. These amino acid residues provide possible side-chain contacts within 6 Å of the ITC, whereas the similar residues A69, G310, and G314 provide carbonyl contacts within 4 Å of the drug. Of these residues, A69 provides the only direct hydrogen bond between the polypeptide and the tail of ITC. The backbone carbonyl of H381 and the side-chain
hydroxyl of S382 form a hydrogen bond network with water molecules that are hydrogen bonded to ITC. Finally, the dissimilar amino acid residues T507, S308, and M509 have poly-peptide backbone carbonyls within 4 Å of the ITC. We conclude that the four fungal-specific amino acid side-chain contacts (L95, F241, H281, and S282) around ITC are augmented by a large number of lower-specificity contacts that could enable the development of either azole or nonazole antifungals with broad specificity for fungal pathogens and minimal opportunity for effective interaction with HsCYP51. We also note that azole resistance has yet to be conferred at any of the four amino acids that contribute a fungal-specific side-chain contact despite a significant number of single-amino acid changes that confer azole resistance in C. albicans, A. fumigatus, C. neoformans, or A. capsulatus mapping to residues within the active site and substrate channel (Fig. S5C). These include ScErg11p A69 (C. albicans A61V), Y70 (A. fumigatus Cyp51A S52T), G73 (A. fumigatus Cyp51A G54E/K/V/R/W), A127 (A. albicans K134E/R), F134 (A. albicans F126L), V138 (A. albicans V130I), Y140 (A. albicans Y132F/H, C. neoformans Y145F, A. capsulatus Y136F), T237 (A. albicans T229A), P239 (A. albicans P230L, A. fumigatus P216L), V242 (A. fumigatus M220K/I/T/V), G314 (A. albicans G307S), T318 (A. fumigatus S297T), and P504 (C. albicans P503L). Consideration of such mutations will be important in pharmacophore definition and antifungal design.

The detection of the putative reaction product in all ScErg11p-6xHis crystal structures adds to pharmacophore definition. Lanosterol is within 6 Å of the bulk of the amino acid residues that are within a similar distance of ITC but has similar proximity with less of SRS-4 (helix I) and G7–G67 (Fig. S3 A and B).

The putative reaction product detected in the egress channel, however, has amino acid residues within 6 Å that extend SRS-1 (backbone carbonyls of A124 and A125), SRS-2 (V242 side chain similar in fungi and human CYP51s), SRS-5 (backbone carbonyl of R385 and the dissimilar K386 sidechain), F92–R98 to V102 (similar in fungi and HsCYP51 R98 side chain and the dissimilar M100 side chain) and adds the G403–L407 β-sheet (H405 is aromatic in fungi but dissimilar in HsCYP51). The backbone carbonyl and the aromatic face of the conserved in fungi F241 are adjacent to the tail of the putative reaction product, and the edge of its aromatic side chain abuts the tail of lanosterol. These results indicate additional features that could contribute to the design of a ScErg11p-based pharmacophore.

Conformational Homogeneity. Cytochrome P450s involved in sterol biosynthesis show remarkable structural conservation and rigidity, especially as compared with microsomal xenobiotic modifying cytochrome P450s (10–12). The ScErg11p-6xHis structure is consistent with a more restrained model of substrate binding obtained using fast X-ray crystallography (13), NMR, and fluorescent spectroscopy (5). For example, we find single conformations for the substrates lanosterol and the triazole drugs ITC (Fig. 2 C) and FLC (Fig. S1B), unlike the multiple conformations found with truncated CYP51s from eukaryotic parasites (14). Furthermore, molecular dynamics simulations on a CYP3A4 without a transmembrane helix gave an open-state structure that may not be physiologically relevant when bound to the lipid bilayer (11). This property may have important consequences for the design of inhibitors that specifically target CYP51 family members or other membrane-bound cytochrome P450s and has been considered in the design of novel inhibitory scaffolds specific for Trypanosoma and Leishmania CYP51s (14).

3. Balzi E, Wang M, Leterme S, Van Dyck L, Goffeau A (1994) PDR5, a novel yeast gene and a putative reaction product detected in the egress channel, however,
Fig. S1. The ScErg11p active-site models in the presence or absence of estriol, FLC, and VCZ. (A) Plus estriol. (B) Plus FLC. (C) Plus VCZ. (D) Ligand free. The estriol, triazoles, and heme are depicted with oxygen atoms colored red, nitrogen atoms blue, and iron brown. The carbon atoms of the heme are colored yellow, the carbon atoms of the estriol are colored salmon, the carbon atoms of fluconazole are colored orange, and the carbon atoms of voriconazole are colored white. Electron densities for the estriol and the triazoles were calculated as Fo–Fc maps contoured to 2.5σ for estriol and FLC and to 3σ for the empty structure and VCZ with the ligand density omitted from the calculation. A covalent bond links the heme–oxo complex. (E) Mass spectrometry profile showing the lanosterol and zymosterol identified in ScErg1p-6xHis preparations. The standards including fluid follicular meiosis activating sterol (FMAS) and testicular meiosis activating sterol (TMAS) are shown.

Fig. S2. Nonmodeled regions of electron density associated with ScErg11p. (A) Electron density extending from near M313 in the catalytic site toward the proposed membrane boundary. (B) Electron density detected in the neighborhood between K220 and helix G.
Fig. S3. Features in the active site and entry and egress channels of ScErg11p. (A) Amino acid residues within 6 Å of ITC in ScErg11p. Amino acid residues within 6 Å of ITC are shown as colored sticks on a cartoon polypeptide backbone that extends (in cyan) the primary sequence two amino acids N-terminal and C-terminal of these residues. Structural features previously identified as possibly contributing to the active site are marked as SRS-1, SRS-2, SRS-4, SRS-5, and SRS-6 (7). The G67-P76 α-helix and the F92-R96 loop also contribute residues within 6 Å of the ITC. The heme is shown in brown, and the molecular dimensions of ITC are shown with yellow lines and black dots. Amino acid residues conserved in ScErg11p and Erg11ps of the major pathogenic fungi of humans and not in HsCYP51 (L95, F241, M313, H381, S382) are shown in red. Residues that are conserved in fungal Erg11ps and HsCYP51 (G73, Y126, L129, T130, F133, Y140, F236, G315) are shown in pink. Amino acid residues that are similar in fungi but not in humans are shown as white sticks (I139, F506); residues that are similar in humans and fungi are in gray. Amino acid residues that are dissimilar in humans and fungi are shown as blue sticks. Dissimilar residues that project the polypeptide backbone carbonyl to within 4 Å of the ITC (T507, S508, and M509) are shown by element (oxygen, red; nitrogen, blue; carbon, green; sulfur, yellow). (B) Amino acid residues within 6 Å of lanosterol and the lanosterol-like product in ScErg11p. Amino acid residues within 6 Å of lanosterol and the lanosterol-like product are shown as colored sticks on a cartoon polypeptide backbone that extends the primary sequence (in cyan) two amino acids N-terminal and C-terminal of these residues. Structural features previously identified as possibly contributing to the active site are marked as SRS-1, SRS-2, SRS-4, SRS-5, and SRS-6 (7). The G67-P76 α-helix and the F92-R96 loop also contribute residues within 6 Å of the ITC. The heme is shown in brown, and the molecular dimensions of ITC are shown with yellow lines and black dots. Amino acid residues conserved in ScErg11p and Erg11ps of the major pathogenic fungi of humans and not in HsCYP51 (L95, F241, M313, H381, S382) are shown in red. Residues that are conserved in fungal Erg11ps and HsCYP51 (G73, Y126, L129, T130, F133, Y140, F236, G315) are shown in pink. Amino acid residues that are similar in fungi but not in humans are shown as white sticks (I139, F506); residues that are similar in humans and fungi are in gray. Amino acid residues that are dissimilar in humans and fungi are shown as blue sticks. Dissimilar residues that project the polypeptide backbone carbonyl to within 4 Å of the ITC (T507, S508, and M509) are shown by element (oxygen, red; nitrogen, blue; carbon, green; sulfur, yellow). (C) Amino acid residues within 6 Å of lanosterol and the lanosterol-like product in ScErg11p. Amino acid residues within 6 Å of lanosterol and the lanosterol-like product are shown as colored sticks on a cartoon polypeptide backbone that extends the primary sequence (in cyan) two amino acids N-terminal and C-terminal of these residues. Structural features previously identified as possibly contributing to the active site are marked as SRS-1, SRS-2, SRS-4, SRS-5, and SRS-6 (7). The G67-P76 α-helix and the F92-R96 loop also contribute residues within 6 Å of the ITC. The heme is shown in brown, and the molecular dimensions of ITC are shown with yellow lines and black dots. Amino acid residues conserved in ScErg11p and Erg11ps of the major pathogenic fungi of humans and not in HsCYP51 (L95, F241, M313, H381, S382) are shown in red. Residues that are conserved in fungal Erg11ps and HsCYP51 (G73, Y126, L129, T130, F133, Y140, F236, G315) are shown in pink. Amino acid residues that are similar in fungi but not in humans are shown as white sticks (I139, F506); residues that are similar in humans and fungi are in gray. Amino acid residues that are dissimilar in humans and fungi are shown as blue sticks. Dissimilar residues that project the polypeptide backbone carbonyl to within 4 Å of the ITC (T507, S508, and M509) are shown by element (oxygen, red; nitrogen, blue; carbon, green; sulfur, yellow). (D) Amino acid residues within 6 Å of lanosterol and the lanosterol-like product in ScErg11p. Amino acid residues within 6 Å of lanosterol and the lanosterol-like product are shown as colored sticks on a cartoon polypeptide backbone that extends the primary sequence (in cyan) two amino acids N-terminal and C-terminal of these residues. Structural features previously identified as possibly contributing to the active site are marked as SRS-1, SRS-2, SRS-4, SRS-5, and SRS-6 (7). The G67-P76 α-helix and the F92-R96 loop also contribute residues within 6 Å of the ITC. The heme is shown in brown, and the molecular dimensions of ITC are shown with yellow lines and black dots. Amino acid residues conserved in ScErg11p and Erg11ps of the major pathogenic fungi of humans and not in HsCYP51 (L95, F241, M313, H381, S382) are shown in red. Residues that are conserved in fungal Erg11ps and HsCYP51 (G73, Y126, L129, T130, F133, Y140, F236, G315) are shown in pink. Amino acid residues that are similar in fungi but not in humans are shown as white sticks (I139, F506); residues that are similar in humans and fungi are in gray. Amino acid residues that are dissimilar in humans and fungi are shown as blue sticks. Dissimilar residues that project the polypeptide backbone carbonyl to within 4 Å of the ITC (T507, S508, and M509) are shown by element (oxygen, red; nitrogen, blue; carbon, green; sulfur, yellow).
and C-terminal of these residues. The V70-K75 α-helix, the F92-V102 loop, and the G403-L407 β-sheet also contribute residues within 6 Å of the lanosterol and its reaction product. Amino acid residues that are conserved in the Erg11ps of \textit{S. cerevisiae} and the major pathogenic fungi of humans and not in \textit{HsCYP51} (L95, F241, M313, H381, S382) are shown in red. Amino acid residues that are conserved in fungal Erg11ps and \textit{HsCYP51} are shown in pink (G73, Y126, L129, T130, F133, Y140, F236, G315, R385). Amino acid residues that are similar in fungi but not in humans are shown in white (I139, F507, H405); residues that are similar in humans and fungi are in gray. Amino acid residues that are dissimilar in humans and fungi are shown as blue sticks. Dissimilar residues that project the polypeptide backbone carbonyl to within 4 Å of the lanosterol or the reaction product (A124, A125, S508, and M509) are shown by element (oxygen, red; nitrogen, blue; carbon, green; sulfur, yellow). (C) Mutations in amino acid residues neighboring the active site and substrate entry channel that confer FLC resistance to fungal pathogens. The annotation is as in A, with mutated residues shown in yellow and annotated in black.

**Fig. S4.** Conservation of amino acid residues in \textit{ScErg11p} and the Erg11ps of 10 prominent fungal pathogens. The cartoon of \textit{ScErg11p} is shown with the heme as blue sticks, ITC in yellow sticks, and with either conserved amino acid residues as red sticks or similar amino acid residues as white sticks. The figure compares the primary sequence of \textit{ScErg11p} used in the present study with its homologs in 10 major fungal pathogens affecting humans: \textit{C. albicans} (accession no. EMBL P10613.2), \textit{C. glabrata} (EMBL P50859.1), \textit{C. dubliniensis} (EMBL AAK57519.1), \textit{C. tropicalis} (EMBL P14263.2), \textit{Cryptococcus neoformans} (EMBL AAP5366.1), \textit{Aspergillus fumigatus} (EMBL Q4WNTS and EMBL Q96W81), \textit{Coccidioides immitis} (RefSeqXP_001240306.1), \textit{Pneumocystis carinii} (EMBL ABG91757), \textit{Ajellomyces capsulatus} or \textit{Histoplasma capsulatum} (EMBL AAAU01158.1), and \textit{Malassezia globosa} (RefSeqXP_001730619.1). The primary sequences were aligned using the Constraint-based Multiple Alignment Tool (Cobalt; \textit{www.ncbi.nlm.nih.gov/tools/cobalt/cobalt.cgi?link_loc=BlashHomeLink}) plus some minor modifications by eye. (A) Side view of \textit{ScErg11p}. (B) Opposite side view of \textit{ScErg11p}. (C) Top view of \textit{ScErg11p}.
Fig. S5. Expression of ScErg11p-GFP and purification of ScErg11p-6xHis. (A) Confocal microscopy of GFP conjugates of the plasma membrane ATP-binding cassette transporter ScPdr5p, the cytosolic marker CaUra3p, and ScErg11p. Yeast cells were grown in YPD medium to OD$_{600nm}$ = 3.0, harvested, and washed with fresh medium, and suspensions at OD$_{600nm}$ = 1.0 were treated with 0.5 μg/mL DAPI for 1–20 min at room temperature before analysis by confocal microscopy (Zeiss 710 confocal laser scanning microscope) using white light illumination and laser illumination at 405 and 488 nm. The fluorescence emission was detected at ~455 nm (DAPI, false color blue) and at ~507 nm (GFP, false color green) for slices of 0.4 μm. White light and fluorescent images are combined for a single slice. The images show that ScPdr5p-GFP is associated almost exclusively with the plasma membrane, Ura3p-GFP is associated with the cytosol and is excluded from organelles, and ScErg11p-GFP surrounds the nucleus. (B) SDS/PAGE analysis of samples from the purification of the ScErg11p-6xHis used for crystallization. Lanes 1 and 9: protein molecular weight markers (BenchMark, Invitrogen); lane 2: 50 μg crude yeast membranes; lane 3: DM extract from 50 mg crude yeast membranes; lane 4: flow through of Ni-NTA column from 50 μg crude yeast membranes; lane 5: 10 μL nonconcentrated Ni-NTA affinity-purified ScErg11p-6xHis; lane 6: 10 μL 5× concentrated Ni-NTA affinity-purified ScErg11p + 50 μM ITC; lane 7: 23 μg Ni-NTA affinity- and SEC-purified ScErg11p-6xHis; lane 8: 124 μg Ni-NTA affinity- and SEC-purified and concentrated ScErg11p-6xHis. (C) Preparation of purified ScErg11p-6xHis by SEC. Samples of 0.5 mL of concentrated Ni-NTA affinity-purified ScErg11p-6xHis in the presence of 50 μM ITC were separated using a Superdex 20 10/300 column (GE Healthcare) equilibrated with 20 mM Heps (pH 7.5), 0.15 M NaCl, 10% glycerol (v/vol), 2 μM ITC, 6.4 mM DM, 0.5 mM PMSF, and 1 Roche minus EDTA protease minipill per 400 mL at a flow rate of 0.4 mL/min. Protein was detected at 280 nm, and the heme detected at 420 nm was normalized relative to the peak maximum. The protein and heme peak coincided, and their relative shapes indicated that the bulk of the protein in the peak was heme-containing ScErg11p-6xHis, which ran with a mobility expected of the 61-kDa protein inserted in a DM micelle of 35 kDa. The enzyme purified in the absence of ITC showed the same behavior during SEC. The molecular weight standards used were thyroglobulin, gamma globulin, ovalbumin, myoglobin, and vitamin B$_{12}$. The DM micelle was detected using refractive index measurements. (D) Spectra of concentrated Ni-NTA affinity- and SEC-purified ScErg11p-6xHis ± 2 μM ITC. The spectra of 1-μL samples were determined using a Pearl Nanodrop with a 0.2-mm path length. The alpha and beta peaks were detected in these preparations at 570 and 540 nm, respectively. Legend continued on following page
Mass spectrometry analysis of tryptic fingerprints of ScErg11p. The protein sequence of ScErg11p-6xHis used in the present study. A sample of the protein used for crystallography was separated by SDS/PAGE as a 62-kDa protein band and analyzed by mass spectrometry of tryptic peptides at the Stanford University Mass Spectrometry facility. Peptides detected by tandem mass spectrometry are shown in bold. Coverage of 70% of the ScErg11p-6xHis primary sequence was obtained.

Crystals of ScErg11p-6xHis. The crystals were formed by vapor diffusion in 6-μL drops by mixing the purified protein 1:1 with 45% (wt/vol) PEG4000 in 100 mM glycine at pH 9.3. The crystals are ~100 μm in length.

Table S1. Yeast strains used in the present study

<table>
<thead>
<tr>
<th>Strain number</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMLY663 (ADΔ)</td>
<td>MATα his1Δ ura3Δ pdr1-3Δ pdr3::hisG Δyor1::hisG Δsnq2::hisG Δycf1::hisG Δpdr10::hisG Δpdr11::hisG Δpdr15::hisG Δpdr5::hisG Δpdr11::hisG Δpdr15::hisG Δpdr5::hisG</td>
</tr>
<tr>
<td>MMLY941</td>
<td>ADΔ Δpdr5::pABC3ScERG11-6xHis</td>
</tr>
<tr>
<td>MMLY942</td>
<td>ADΔ Δpdr5::pABC3ScERG-GFP</td>
</tr>
<tr>
<td>MMLY939</td>
<td>ADΔ Δpdr5::pABC3ScPDR5-GFP</td>
</tr>
<tr>
<td>MMLY1834</td>
<td>ADΔ Δpdr5::pABC3CaURA3-GFP</td>
</tr>
</tbody>
</table>

The MMLY941 strain was constructed by transforming the yeast host strain ADΔ with a linear transformation cassette comprising the AscI fragment from the pABC3ScERG11-6xHis at the PDR5 locus and selecting ura+ progeny (1). The strains MMLY942, MMLY939, and MMLY1834 were obtained using recombinant PCR to prepare constructs that replaced a C-terminal 6xHistidine tag with EGFP as previously described (1). The PDR5 locus in each yeast transformant was recovered by PCR of genomic DNA, and its DNA sequence was determined.

Table S2. X-ray crystal structure dataset statistics

<table>
<thead>
<tr>
<th>Erg11 datasets</th>
<th>Empty pocket</th>
<th>+ ITC</th>
<th>+ FLZ</th>
<th>+ VCZ</th>
<th>+ Estriol</th>
<th>+ Lanosterol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P21</td>
<td>P21</td>
<td>P21</td>
<td>P21</td>
<td>P21</td>
<td>P21</td>
</tr>
<tr>
<td>a</td>
<td>80.1 Å</td>
<td>79.6 Å</td>
<td>80.3 Å</td>
<td>79.4 Å</td>
<td>79.4 Å</td>
<td>78.2 Å</td>
</tr>
<tr>
<td>b</td>
<td>67.7 Å</td>
<td>67.6 Å</td>
<td>67.3 Å</td>
<td>67.7 Å</td>
<td>67.5 Å</td>
<td>67.0 Å</td>
</tr>
<tr>
<td>c</td>
<td>81.0 Å</td>
<td>80.7 Å</td>
<td>81.3 Å</td>
<td>81.2 Å</td>
<td>80.8 Å</td>
<td>80.4 Å</td>
</tr>
<tr>
<td>α</td>
<td>90°</td>
<td>90°</td>
<td>90°</td>
<td>90°</td>
<td>90°</td>
<td>90°</td>
</tr>
<tr>
<td>β</td>
<td>100.14°</td>
<td>99.75°</td>
<td>100.34°</td>
<td>100.14°</td>
<td>99.58°</td>
<td>99.54°</td>
</tr>
<tr>
<td>γ</td>
<td>90°</td>
<td>90°</td>
<td>90°</td>
<td>90°</td>
<td>90°</td>
<td>90°</td>
</tr>
<tr>
<td>Completeness</td>
<td>95.8% (96.7%)</td>
<td>97.0% (99.0%)</td>
<td>91.7% (86.7%)</td>
<td>98.9% (99.7%)</td>
<td>97.0% (94.7%)</td>
<td>96.5% (72.8%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* R_sym = \( \sum_i l_i(hkl) - \bar{l}(hkl) / \sum_i l_i(hkl) \), where \( l_i(hkl) \) and \( \bar{l}(hkl) \) are the \( i \)th and mean measurements of the intensity of reflection \( hkl \).

\[ R_{cryst} = \sum_i |F_{obs}(hkl) - F_{calc}(hkl)| / \sum_i |F_{obs}(hkl)|, \]

where \( F_{obs} \) and \( F_{calc} \) are the observed and calculated structure factors for the reflection \( hkl \) and \( k \) is a weighing factor.

\[ R_{free} = \sum_{i \in T} |F_{obs}(hkl) - kF_{calc}(hkl)| / \sum_{i \in T} |F_{obs}(hkl)|, \]

where \( F_{obs} \) and \( F_{calc} \) are the observed and calculated structure factors for the reflection \( hkl \) and \( k \) is a weighing factor. \( T \) is the test set of reflections.

\[ \text{Rmsd.} \]

Monk et al. www.pnas.org/cgi/content/short/1324245111