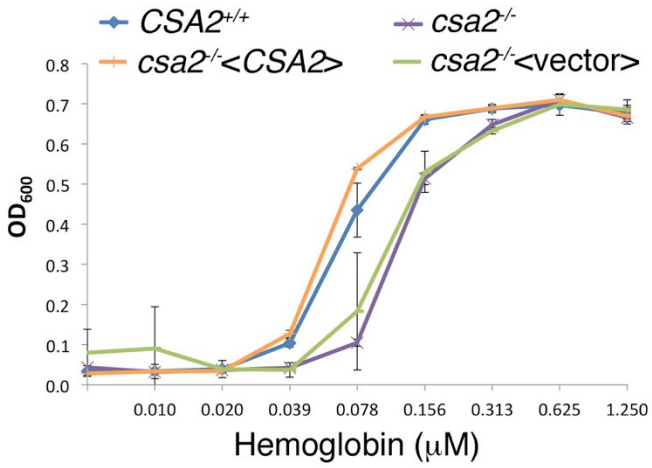
nature microbiology

SUPPLEMENTARY INFORMATION

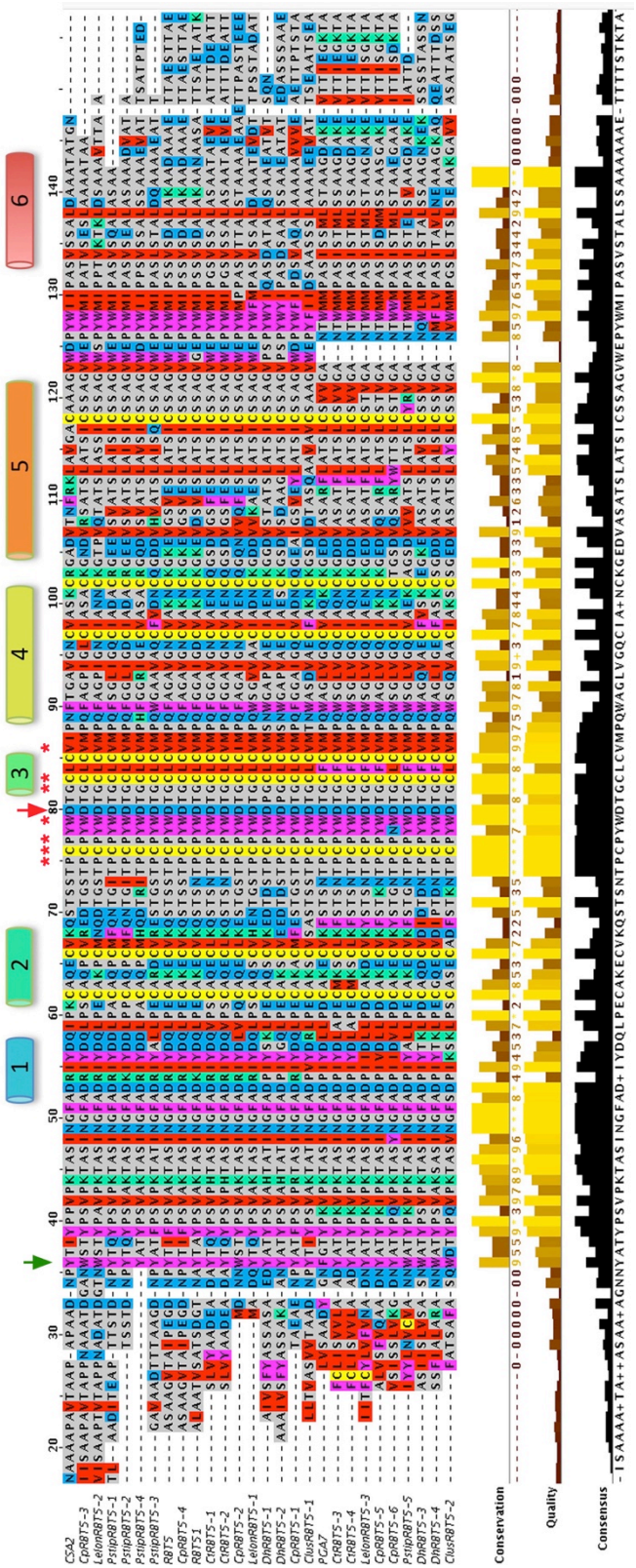
ARTICLE NUMBER: 16156 | DOI: 10.1038/NMICROBIOL.2016.156

Structural basis of haem-iron acquisition by fungal pathogens

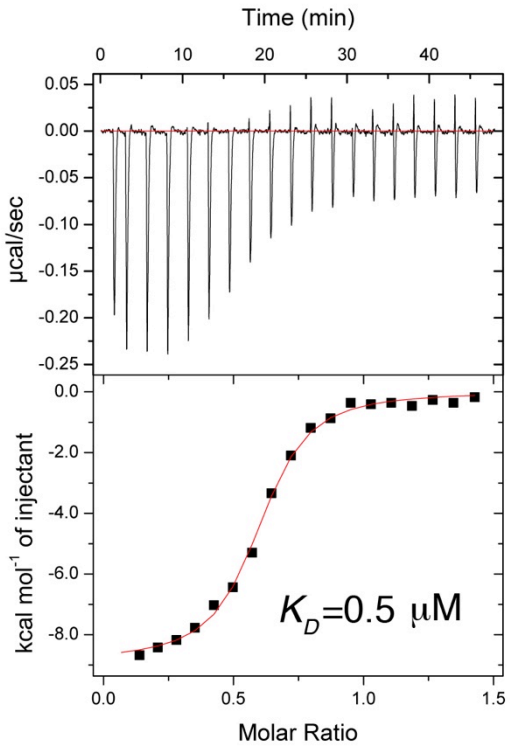
L. Nasser *et al.*



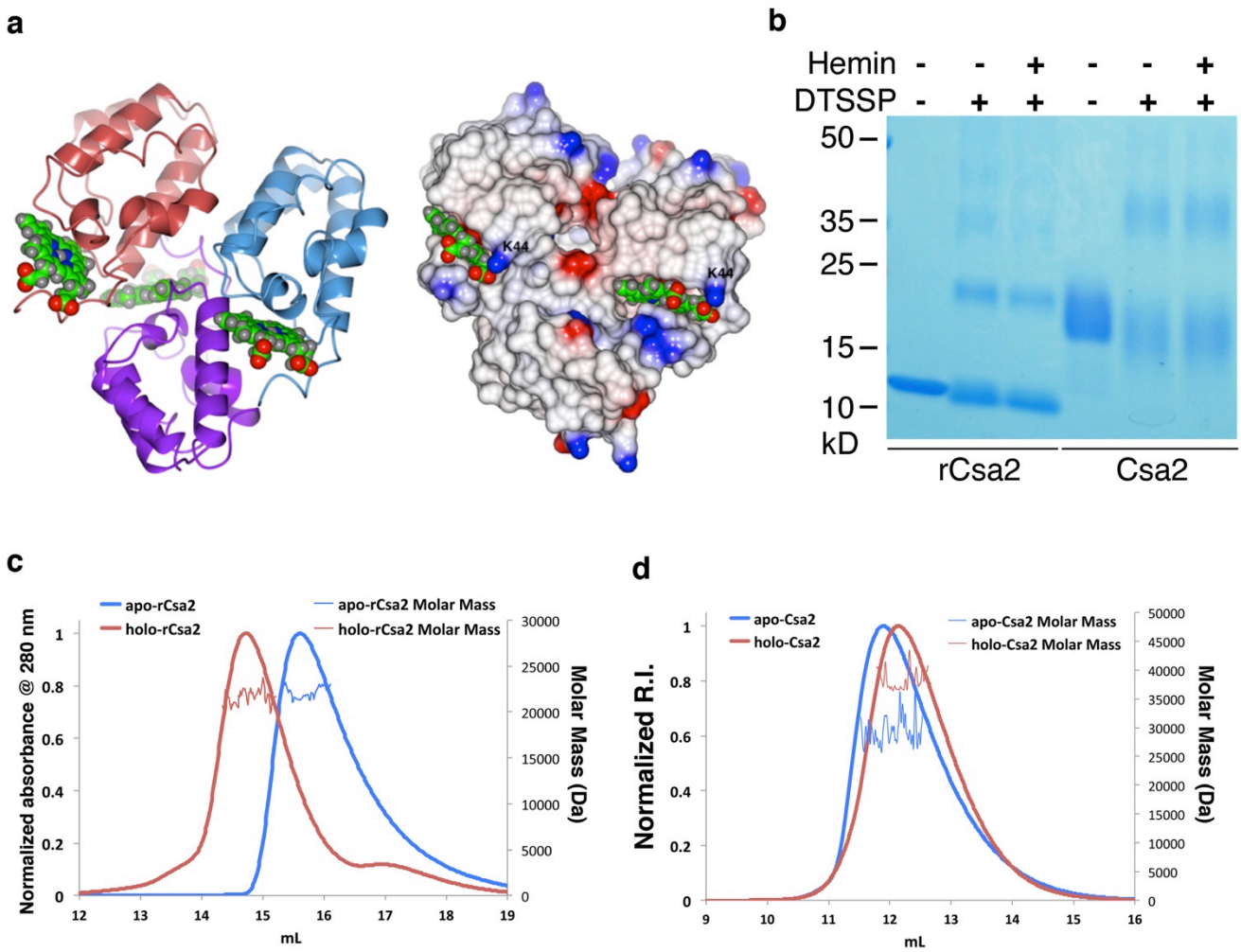
Supplementary Figure 1. Reintegration of *CSA2* into the knockout strains restores wild-type growth of the *C. albicans* strain. Growth conditions were as described for Figure 1d. Each curve represents the average of 3 cultures. Vertical bars indicate standard deviations. The strains are KC36, KC778, KC787, and KC812.



Supplementary Figure 2. Sequence alignment of Csa2 with additional *Candidaceae* CFEM proteins. Numbering corresponds to the Csa2 sequence. The canonical CFEM cysteine residues are marked in yellow. The positions of the 6 α-helices of the Csa2 CFEM domain structure are indicated above the sequences. The red arrow marks Asp80 that coordinates the heme iron, whereas the green arrow marks Tyr36 which interacts with the opposite side of heme plane. The residues forming the hydrophobic platform around Asp80 (Fig. 4d) are indicated by red asterisks.

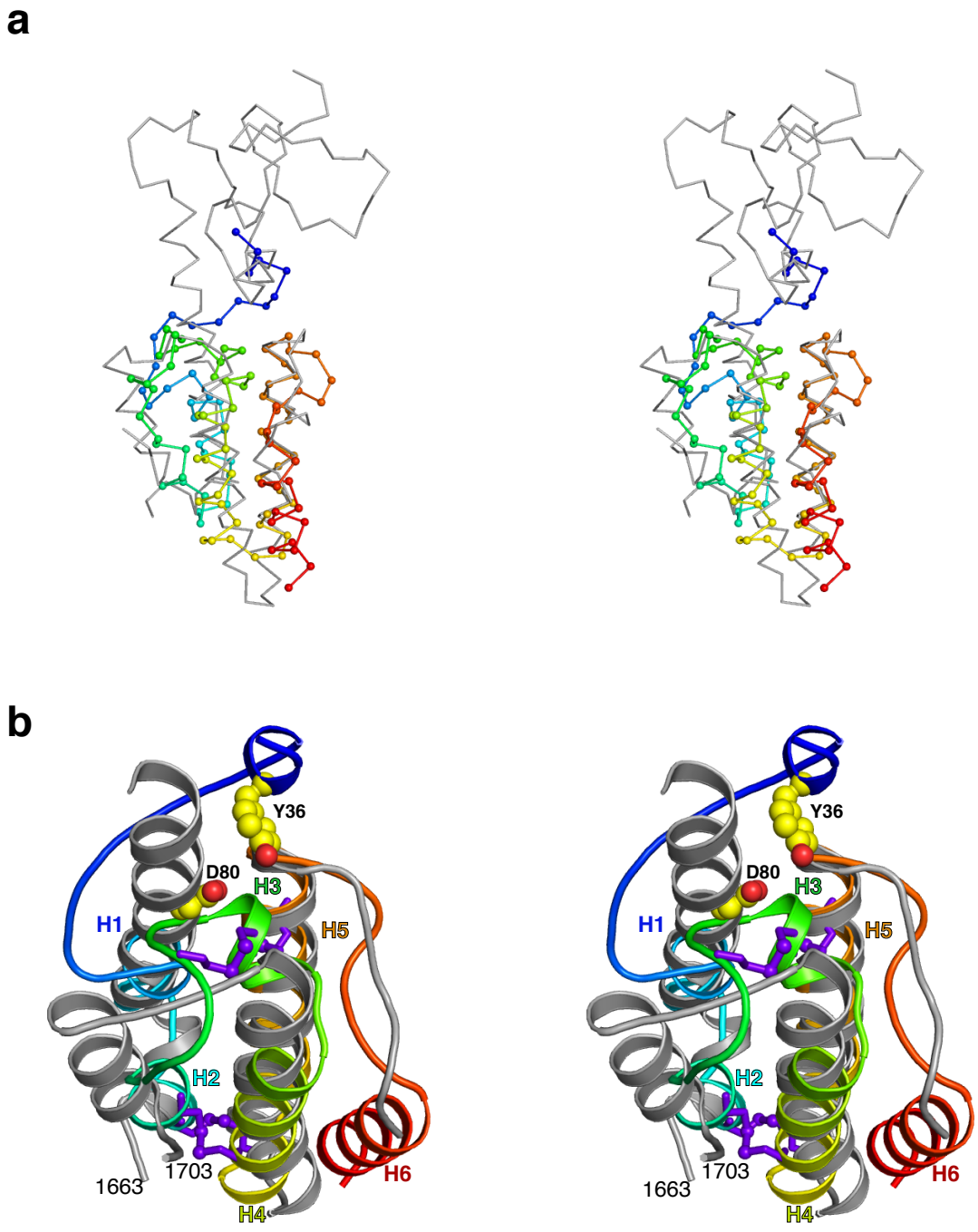


Supplementary Figure 3. Isothermal calorimetry of rCsa2 with hemin. The experiment was carried out as described in the Methods. Representative heat signal thermogram (upper panel) and a binding isotherm plot of normalized heats as a function of the molar ratio (lower panel) corresponding to nineteen 4-µL injections of 140 µM apo-rCsa2 solution into 20 µM hemin at 150 s intervals. The binding isotherm profile was fit to a single binding site model and the binding affinity was derived using the "Origin" software package. The dissociation constant value, 0.5 ±0.3 µM, is the average of 3 such independent experiments.



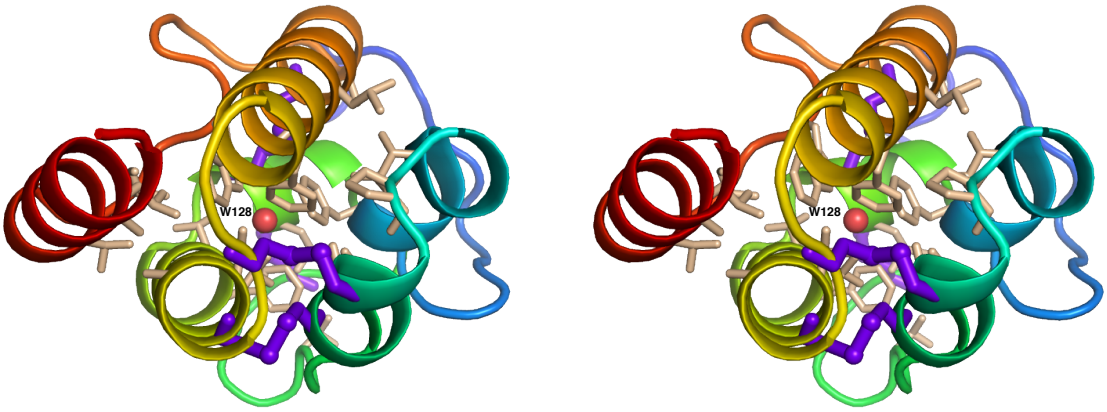
Supplementary Figure 4. Quaternary structure of rCsa2. (**a**) Holo Csa2 crystallized as an asymmetric homomeric trimer, colored by chain on the left. Extensive interaction interfaces exist between its monomers (5520 Å^2 buried surface) compared to the relatively small overall trimer surface area (14520 Å^2), suggesting that despite its asymmetry it is a rather stable assembly, with a calculated ΔG of -73.4 kcal/mol (pdbe.org/pisa). In this assembly each of the heme groups (green spheres models) interacts primarily with a single Csa2 protein chain, but also makes additional non-specific hydrophobic contacts with neighboring Csa2 homomers, *i.e.* the heme on the right interacts with the blue and purple chains while the heme in the middle-back interacts with the purple and maroon chains. The heme drawn on the left interacts primarily with the maroon chain but also mildly interfaces with a symmetry related molecule of the blue chain (not shown). The electrostatic surface representation on the right shows how the propionate groups of the heme protrude out toward the solvent where they can also interact with nearby lysine side chains (K44). (**b**) Crosslinking of rCsa2 or Csa2 with DTSSP yields a ~50% conversion to dimers, regardless of the presence of hemin. (**c**) Multi Angle Light Scattering in line with SEC (Superdex75 10/300, GE) (SEC-MALS) with the bacterially expressed rCsa2 (MW = 13.2 kDa) indicates a molar mass of 22.7 ± 0.8 kDa for apo-rCsa2 and 22.9 ± 0.4 kDa for holo-rCsa2, consistent with a dimeric assembly. Absorbance at 280 nm normalized to 1 (= peak absorbance) is indicated on the left Y-axis and molar mass values are indicated on the right axis. (**d**) SEC (Superdex75 10/300, GE) -MALS with the *P. pastoris*-expressed Csa2 (MW = 16.2 kDa + 15-18

mannosyl adducts, i.e. 18.6-19.2 kDa total). Refractive index (R.I.) was used to measure protein concentration, with dn/dc = 0.179 based on 84% protein, 16% sugars. Molecular masses were 30 ± 4 kDa for apo-rCsa2 and 39 ± 3 kDa for holo-Csa2, consistent with a dimeric assembly. R.I. normalized to 1 (=peak R.I.) is indicated on the left Y-axis. Molar mass is indicated on the right axis.

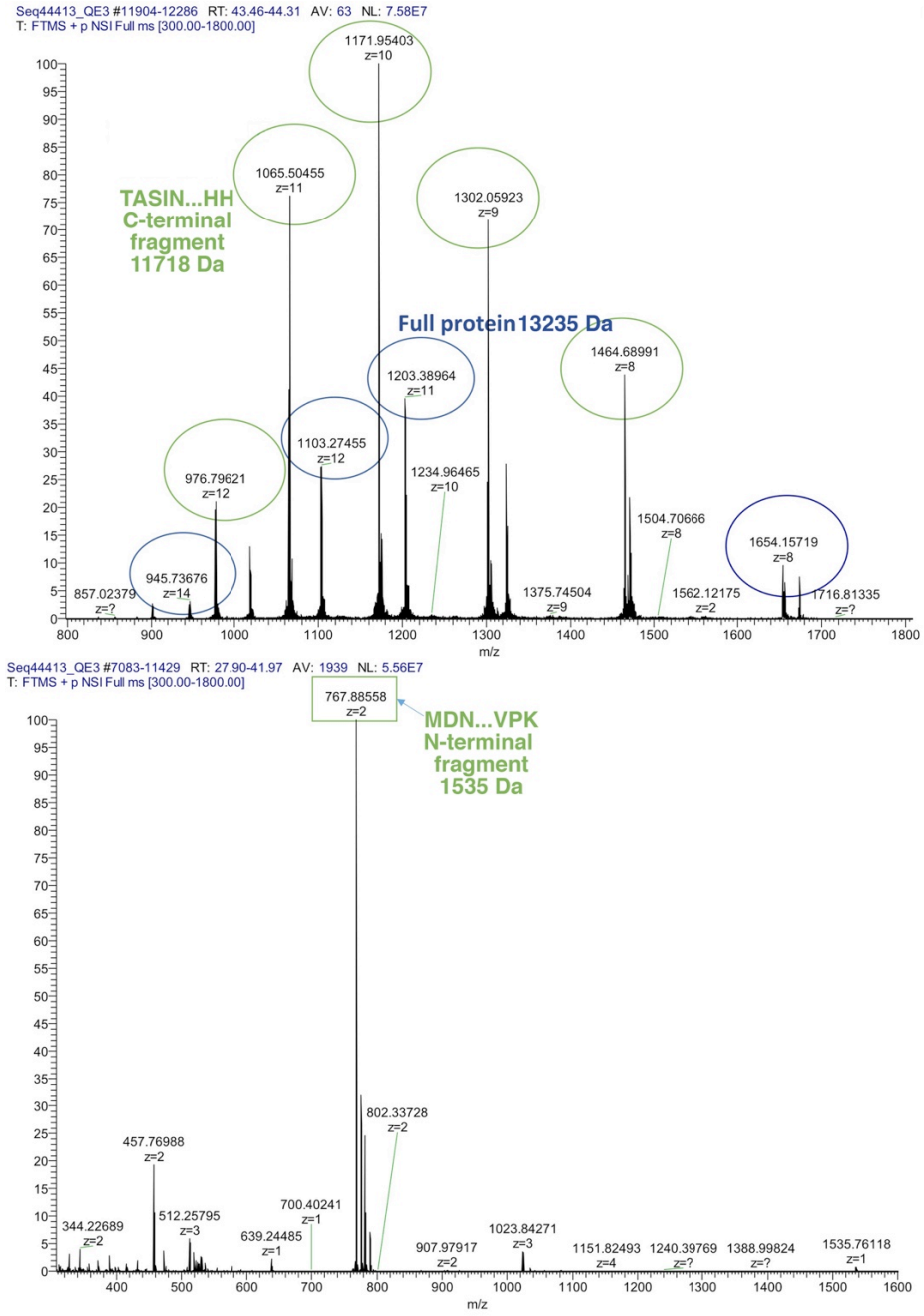


Supplementary Figure 5. Structural alignment of the bromodomain of MLL1 and rCsa2. The highest scoring Dali structural alignment was found with a portion of the bromodomain of the MLL1 protein (PDB ID 3LQJ), yielding a Z score of 5.2 indicating a similar structure. However, the Dali cutoff Z score for a strong match for a protein of this size is 10.5 [1], suggesting that the similarity is superficial. **(a)** Stereo view of the overall structural alignment of the Cα traces of MLL1 (grey sticks) and rCsa2 (balls and sticks colored with gradient from

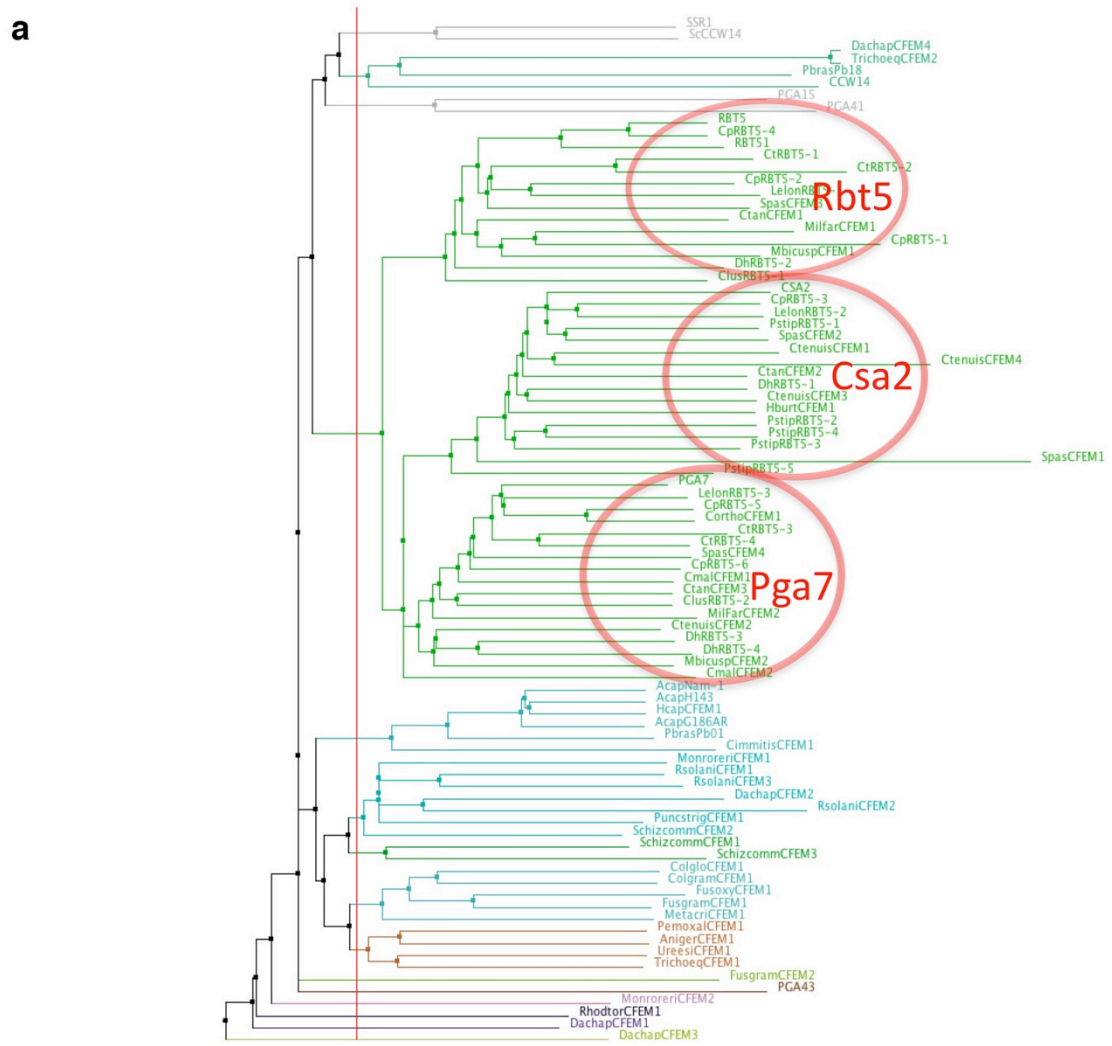
blue at the N-terminus to red at the C-terminus). **(b)** A zoom-in stereo view on the aligned regions colored as above. It can be seen that compared to the 6-helical-basket fold of Csa2, the aligned portion of the bromodomain has a pseudo four-helix bundle fold with its helix 2 interrupted by a missing 40 residue protein stretch (see flanking tails marked 1663 and 1703). Although helices 4 and 5 of Csa2 align nicely with helices 3 and 4 of the bromodomain, helix 1 and its pseudo helix 2 only partially superimpose on helices 1 and 2 of Csa2, while helices 3 and 6 of Csa2 are completely missing in the bromodomain structure. Furthermore, this portion of the bromodomain lacks cysteine residues. Thus, despite the global structural alignment between these structures they represent completely different folds. For orientation, Asp 80 and Tyr 36 are displayed as yellow spheres and the cysteines in purple ball-and-stick format.



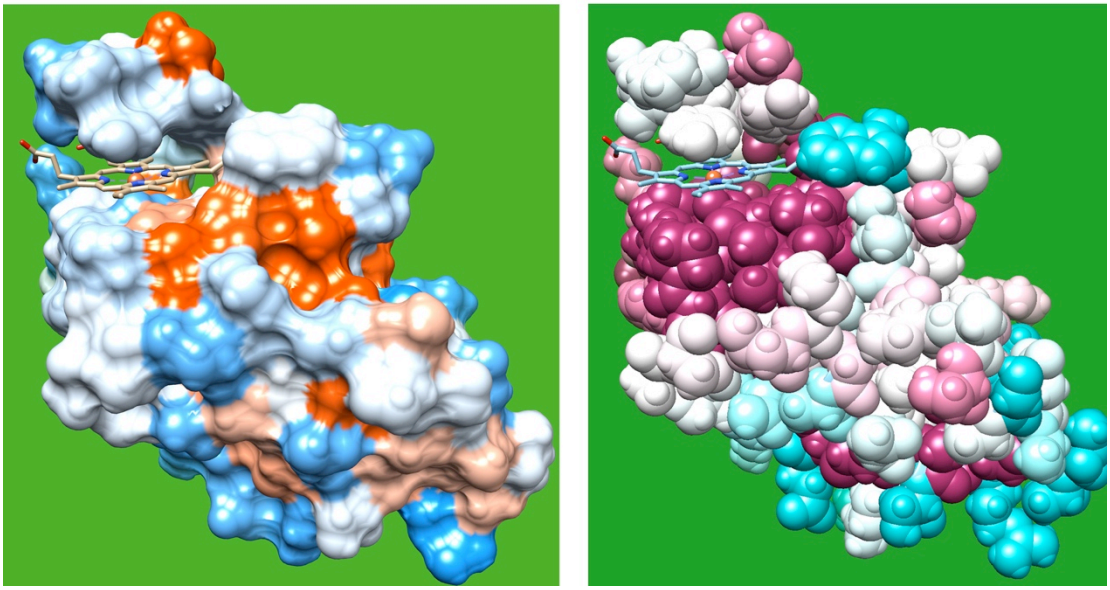
Supplementary Figure 6. The CFEM helical basket fold is stabilized by hydrophobic interactions in addition to four unique disulfide bonds. A stereo view is shown of the structure of Csa2 in cartoons format, color-coded by gradient from blue at the N-terminus to red at the C-terminus. Disulfide bonds are shown as purple balls-and-sticks, and the hydrophobic side-chains pointing to the core of the structure, and thus stabilizing the helical packing, are rendered as beige sticks. A single water molecule is bound at the core of the basket, forming a hydrogen bond with Trp128.



Supplementary Figure 7. MS analysis of rCsa2 digested by trypsin. The digestion reactions were split in two; one half was analyzed by SDS-PAGE shown in Fig. 4b, and the other half of the reaction (only of rCsa2 supplemented with hemin) was analyzed at the Technion Smoler proteomics center by LC-MS/MS on Q Exactive plus (Thermo) and identified by Discoverer software version 1.4 using the Sequest search engine vs. the specific sequence. The following main species are detected: the undigested full protein of 13235 Da (with multiple charged species circled blue, top chart), the fragment extending from T45 to the C-terminus of 11718 Da (top chart circled green), and the N-terminal fragment extending to position K44 of 1535 Da (bottom chart, green rectangle).

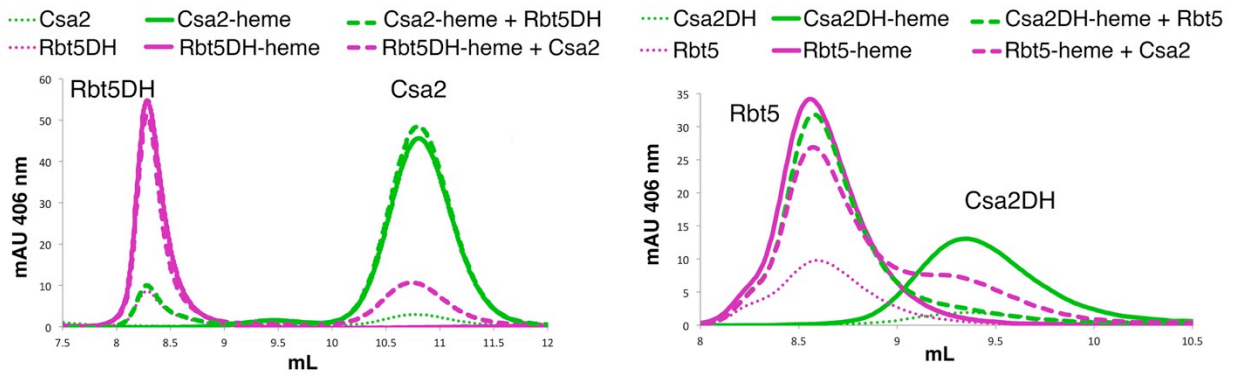


b



Supplementary Figure 8

Supplementary Figure 8. (a) Proximity tree of Ascomycete CFEM protein sequences. The sequences were gathered by homology search among fungal genomes, using the Rbt5 sequence as query. The 46 Candiaceae CFEM proteins (green lines) cluster in three groups, with homology to Rbt5, Csa2 and Pga7, respectively (see also Fig. S2). Represented are 46 sequences from 15 species: *C. albicans, C. parapsilosis, C. orthopsilosis, C. lusitaniae, C. tropicalis, C. maltosa, C. tenuis, C. tanzawaensis, Debaryomyces hansenii, Pichia stipitis, Millerozyma farinosa, Lodderomyces elongisporus, Spathaspora passalidarum, Hyphopichia burtonii, Metschnikowia bicuspidata*. Alignment of these 46 sequences was used to map the conservation of each residue onto the Csa2 structure. **(b)** Left: hydrophobicity surface representation of Csa2 (red = hydrophobic, blue = charged) together with a stick representation of heme. Right: conservation of Csa2 residues among Candidaceae CFEM proteins (from red = most conserved, to blue = least conserved). The sequences shown in the top panel were aligned and mapped onto the Csa2 sequence using ConSurf (http://consurf.tau.ac.il).



Supplementary Figure 9. Defective heme exchange between wild-type and DH mutant proteins. Rbt5DH and Csa2 (approximately 50 µM each) were pre-loaded separately with 20 µM hemin (left) and Csa2DH and Rbt5 (approximately 50 µM each) were pre-loaded separately with 10 µM hemin (right), before loading on a SEC column (Superdex 75 10/300, GE). Where indicated, the protein-heme complex was mixed with the second protein in apo-form, and incubated 5 min at room temperature before loading onto the SEC column. The protein-bound heme was detected by monitoring absorbance at 406 nm, the Soret peak of heme bound to the CFEM protein. Whereas no transfer of heme was detected between Rbt5DH and Csa2 (left panel), heme was efficiently transferred from Csa2DH to Rbt5, but not in the opposite direction (right panel).

Supplementary Table 1. Data collection and refinement statistics for the structure of rCsa2.

Space group	**$P2_12_12_1$**
Cell dimensions *a*, *b*, *c* [Å]	61.34, 61.21, 97.30
Resolution [Å]	32.4 – 2.0 (2.03- 2.0)
Wavelength [Å]	1.54178
Total # of Reflections	369858
Unique reflections	25528
R_{merge}	0.061 (0.625)
R_{pim}	0.017 (0.279)
CC1/2	0.999 (0.832)
CC (estimate of true CC)*	1.000 (0.953)
Mean $I/\sigma(I)$	43.5 (2.04)
Completeness [%]	96.0 (67.5)
Multiplicity	15.1 (5.4)
Wilson B-factor [Å^2]	19.1
Phasing (Res range 14-2.5 Å)	
Method	SAD
Anomalous scatters used in phasing	3 Fe, 4 S atoms
FOM	0.7849
R_{cullis}	0.72
Phasing power	0.89
Refinement	
R_{work} (cryst)	0.1526
R_{free} (5%)	0.1916
Number of non-hydrogen atoms	2791
# Water atoms	217
# Protein atoms	2441
# Ligand atoms	133
# Protein residues	331
Estimated overall coordinate error based on maximum likelihood	0.19 Å
R.M.S. deviations	
Bond lengths [Å]	0.005
Bond angles [°]	1.31
Ramachandran analysis (MOLPROBITY [2])	Preferred – 97.2% Allowed – 2.8%, Outliers - 0.0% , Clashscore 3.59
Average B-factor [Å^2]	32.1
Protein	31.6
Ligand	30.8
Solvent	38.0

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

$$R_{merge} = \frac{\sum_h \sum_i |I_i - < I >|}{\sum_h |I_i|}, \quad R_{cryst} = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

$$R_{cullis\ ano} = \frac{\sum_h \sum_\phi \mathrm{P}(\phi)|\Delta_{ano}(obs) - \Delta_{ano}(calc)|}{\sum_h |\Delta_{ano}(obs)|}, \quad (\mathrm{P}(\phi)\ porbability\ weight)$$

Supplementary Table 2. Areas of surfaces and interfaces in the rCsa2 trimer [Å^2]

Calculated using PDBePISA:

Mol		**SAA** (SAA/ASA)	**MolA**	**MolB**	**MolC**
	ASA		5858	5942	5777
HemD	833	267.7 (32%)	493.6	71.7	
HemE	823	203.5 (24%)		533.5	86.0
HemF	808	272 (33%)			536.7

ASA - is the Accessible Surface Area, which is an estimate of the overall surface area of a molecule calculated by rolling a ball-probe of radius of 1.4 Angstrom around it.
SAA - is the Solvent Accessible Area, which represents the portion of ASA that is not in contact with a macromolecule.
SAA/ASA is the percent solvent exposed area.
Highlighted in yellow are the interaction surface areas of each heme molecule with its main binding Csa2 monomer in the trimer. For instance, HemF which does not interact with MolB or MolA has ~66% (536/808*100) of it surface buried in monomer C of Csa2.
The CFEM-bound heme has a large solvent accessible area (~33%), much higher than in catalases (1.3%), globin-like proteins (~16%) or cytochrome b5 (~22%) [3], but comparable to the heme solvent accessibility in the bacterial hemophore IsdC (34%; [4]).

Supplementary Table 3. List of *C. albicans* and *P. pastoris* strains.

Name	Genotype or description	Plasmid	Origin
	***Candida albicans* strains**		
KC2 = CAF3-1	*ura3D::imm434/ura3D::imm434*		[5]
KC36	KC2 *ccc2Δ::hisG/ccc2Δ::hisG-URA3-hisG*		[6]
KC68	KC2 *ccc2Δ::hisG/ccc2Δ::hisG*		[6]
KC131	CAI-4 *rbt5Δ::hisG /rbt5Δ::hisG ccc2Δ::hisG /ccc2Δ::hisG-URA3-hisG*		[7]
KC139	KC131 *ura3-*		[8]
KC484	KC68 *pga7Δ::hisG /pga7Δ:: hisG-URA3-hisG*		[8]
KC485	KC68 *pga7Δ::hisG /pga7Δ:: hisG*		[8]
KC778	KC68 *csa2Δ::hisG /csa2Δ:: hisG-URA3-hisG*		This study
KC779	KC485 *csa2Δ::hisG /csa2Δ:: hisG-URA3-hisG*		This study
KC781	KC139 *csa2Δ::hisG /csa2Δ:: hisG-URA3-hisG*		This study
KC782	KC68 *csa2Δ::hisG /csa2Δ:: hisG*		This study
KC787	KC782 *ade2::URA3 CSA2*	KB2353	This study
KC812	KC782 *ade2::URA3*	pBES116	This study
KC940	KC782 *ade2::URA3 CSA2 D80H*	KB2454	This study
KC536	KC485 *ade2::URA3 PGA7*	KB2111	[8]
KC904	KC485 *ade2::URA3*	pBES116	This study
KC921	KC485 *ade2::URA3 PGA7 D63H*	KB2435	This study
KC922	KC139 *ade2::URA3 RBT5 D72H*	KB2436	This study
KC955	KC139 *ade2::URA3 RBT5*	KB1664	This study
KC956	KC139 *ade2::URA3*	pBES116	This study
	***Pichia pastoris* X33 protein expression strains**		
KC755	Pga7(18-195a.a.)- Myc-6xHis	KB2259	[8]
KC925	Pga7(18-195a.a.) D63H-Myc-6xHis	KB2437	This study
KC758	Rbt5(23-219a.a.)-Myc-6xHis	KB2258	[8]
KC843	Rbt5(23-219a.a)D72H-Myc-6xHis	KB2449	This study
KC842	Csa2(19-147a.a.)-Myc-6xHis	KB2366	This study
KC889	Csa2(19-147a.a.) D80H-Myc-6xHis	KB2412	This study

Supplementary References

1. Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24**, 2780-1 (2008).
2. Chen, V.B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**, 12-21 (2010).
3. Schneider, S., Marles-Wright, J., Sharp, K.H. & Paoli, M. Diversity and conservation of interactions for binding heme in b-type heme proteins. *Nat Prod Rep* **24**, 621-30 (2007).
4. Sharp, K.H., Schneider, S., Cockayne, A. & Paoli, M. Crystal structure of the heme-IsdC complex, the central conduit of the Isd iron/heme uptake system in Staphylococcus aureus. *J Biol Chem* **282**, 10625-31 (2007).
5. Fonzi, W.A. & Irwin, M.Y. Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* **134**, 717-728 (1993).
6. Weissman, Z., Shemer, R. & Kornitzer, D. Deletion of the copper transporter CaCCC2 reveals two distinct pathways for iron acquisition in Candida albicans. *Mol Microbiol* **44**, 1551-60 (2002).
7. Weissman, Z. & Kornitzer, D. A family of Candida cell surface haem-binding proteins involved in haemin and haemoglobin-iron utilization. *Mol Microbiol* **53**, 1209-20 (2004).
8. Kuznets, G. et al. A Relay Network of Extracellular Heme-Binding Proteins Drives C. albicans Iron Acquisition from Hemoglobin. *PLoS Pathog* **10**, e1004407 (2014).