

Structural basis of haem-iron acquisition by fungal pathogens

Lena Nasser¹, Ziva Weissman¹, Mariel Pinsky¹, Hadar Amartely², Hay Dvir^{3*} and Daniel Kornitzer^{1*}

Pathogenic microorganisms must cope with extremely low free-iron concentrations in the host's tissues. Some fungal pathogens rely on secreted haemophores that belong to the Common in Fungal Extracellular Membrane (CFEM) protein family, to extract haem from haemoglobin and to transfer it to the cell's interior, where it can serve as a source of iron. Here we report the first three-dimensional structure of a CFEM protein, the haemophore Csa2 secreted by *Candida albicans*. The CFEM domain adopts a novel helical-basket fold that consists of six α -helices, and is uniquely stabilized by four disulfide bonds formed by its eight signature cysteines. The planar haem molecule is bound between a flat hydrophobic platform located on top of the helical basket and a peripheral N-terminal 'handle' extension. Exceptionally, an aspartic residue serves as the CFEM axial ligand, and so confers coordination of Fe^{3+} haem, but not of Fe^{2+} haem. Histidine substitution mutants of this conserved Asp acquired Fe^{2+} haem binding and retained the capacity to extract haem from haemoglobin. However, His-substituted CFEM proteins were not functional *in vivo* and showed disturbed haem exchange *in vitro*, which suggests a role for the oxidation-state-specific Asp coordination in haem acquisition by CFEM proteins.

The fourth most-abundant transition metal on the earth's crust, iron, is also an essential nutrient for nearly all organisms. Its ability to accept and donate electrons readily makes it highly reactive, and toxic if unconstrained. Yet its reduction potential, mainly between the $\text{Fe}^{2+}/\text{Fe}^{3+}$ oxidation states, is valuable for many key biological processes which include, for example, oxygen transport, enzyme catalysis and cellular respiration. However, iron scarcity is the prevalent condition for most organisms that live in aerobic environments because of the extremely low solubility of its oxidized form at a physiological pH. Animals possess high-affinity iron-chelation mechanisms that additionally limit its bioavailability for invading microorganisms, and are thus considered an important element of the host nutritional immunity¹. In humans, ~70% of the iron is bound by the oxygen carrier haemoglobin as part of its haem prosthetic group². Many pathogenic microorganisms, bacterial as well as fungal, have consequently evolved mechanisms to exploit haem and haemoglobin as sources of iron^{3–5}. Some of these microorganisms utilize secreted haemophores as part of their haem-iron utilization systems, such as HasA of *Serratia marcescens*⁶, IsdX of *Bacillus anthracis*⁷ and Cig1 of *Cryptococcus neoformans*⁸.

Common in Fungal Extracellular Membrane (CFEM) proteins, characterized by a distinctive pattern of eight cysteine residues of conserved interspacing⁹, are widespread among fungi, with close to 4,000 sequences represented in the Interpro database¹⁰. CFEM proteins have been implicated in fungal haem-iron acquisition and pathogenicity^{11–14}. In the pathogenic fungus *Candida albicans*, in particular, CFEM proteins have been shown to bind haem and to play crucial roles in extracting it from the host haemoglobin, as well as in mediating its delivery to the fungal cell^{12,13}. The CFEM gene *RBT5* is among the most highly induced genes in experimental animal-infection models¹⁵, and Rbt5-specific antibodies are prominently found in the serum of patients recovering from candidaemia¹⁶.

Rbt5 and another CFEM protein, Pga7, were shown to be anchored peripherally and internally, respectively, on the fungal cell envelope and to exchange haem between them, which supports a model in which the haem is shuttled across the cell wall via a

CFEM protein relay pathway¹². Haem exchange probably involves transient interaction between the CFEM proteins¹²; however, the molecular mechanism of haem binding and transfer remained obscure, as the CFEM domain sequence bears no discernible homology to known bacterial haem-uptake systems, or to any other haem-binding protein. Here we report the first crystal structure of a member of the CFEM family, the secreted *C. albicans* haemophore Csa2, and use it to characterize functionally the haem-iron acquisition cascade in pathogenic fungi. The structure of the CFEM domain reveals a novel protein fold, as well as a unique mode of haem-iron coordination that is essential for efficient haem transfer and for growth on haemoglobin as an iron source.

Results

Csa2, secreted haemophore that participates in haem-iron acquisition. Two CFEM proteins previously identified as participating in *C. albicans* haem-iron acquisition, Rbt5 and Pga7, are linked to the cell wall and membrane, probably via glycosylphosphatidylinositol (GPI) anchors^{12,13}. Another closely related CFEM protein, Csa2, was detected previously by mass spectrometry in the culture supernatant of hyphal cells¹⁷. Using antibodies raised against recombinant Csa2, we were, indeed, able to detect it in the supernatant of a culture grown under iron restriction, but not in iron-replete cultures (Fig. 1a). Recombinant Csa2 expressed in the fungal expression system *Pichia pastoris* bound haem, as measured by isothermal calorimetry (Fig. 1b), with a dissociation constant of $0.27 \pm 0.12 \mu\text{M}$; that is, intermediate between the affinity measured for Rbt5 (5 μM) and that for Pga7 (0.05 μM)¹². Similarly to Rbt5 and Pga7 (ref. 12), Csa2 was able to extract haem efficiently from a haemoglobin column (Fig. 1c).

We next tested whether this secreted CFEM protein also participates in haemoglobin-iron utilization, like the cell-envelope-anchored CFEM proteins Pga7 and Rbt5. Compared with the wild type (WT), the *csa2*^{-/-} mutant required a twofold higher concentration of haemoglobin as the sole iron source (Fig. 1d), and this growth defect was complemented by reintroduction of the *wt* *CSA2* allele in the *csa2*^{-/-} strain

¹B. Rappaport Faculty of Medicine, Technion - Israel Institute of Technology, and the Rappaport Institute for Research in the Medical Sciences, Haifa 31096, Israel. ²Wolfson Centre for Applied Structural Biology, Hebrew University of Jerusalem, Jerusalem, Israel. ³Technion Center for Structural Biology, Technion - Israel Institute of Technology, Haifa 32000, Israel. *e-mail: hay.dvir@gmail.com; danielk@tx.technion.ac.il

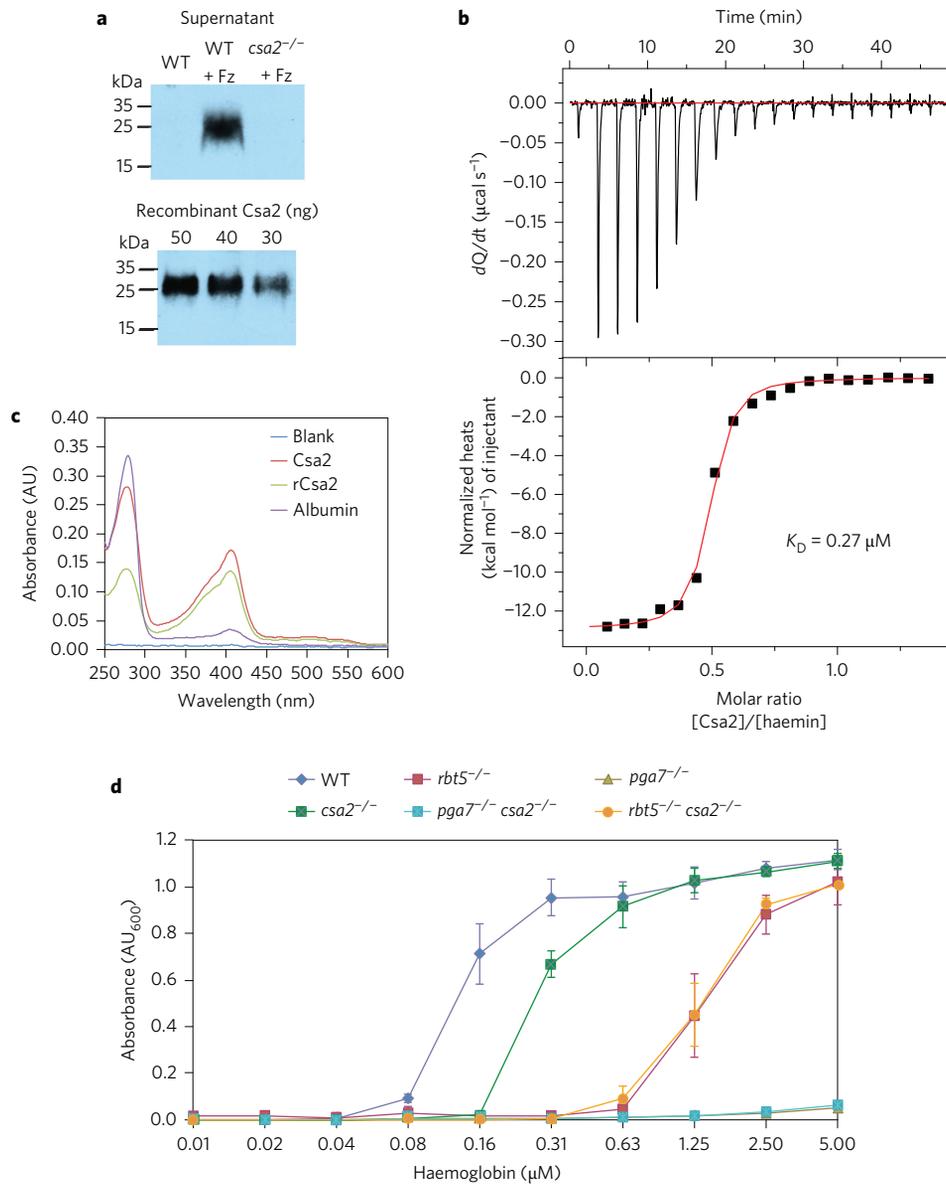


Figure 1 | *C. albicans* Csa2 is a secreted haemophore. a, Csa2 was detected by western blotting with an anti-Csa2 antibody in the supernatants (20 μl per lane, 20 \times concentrate) of cells grown for five hours in YPD medium with 1 mM Fz, an iron chelator (top panel). Known amounts of *P. pastoris*-produced recombinant Csa2 were used as a comparison (bottom panel, taken from the same gel and exposure). The secreted Csa2 reached a concentration of 100 ng ml^{-1} in the medium. Note the aberrant migration of Csa2, a 16.2 kDa protein, caused by multiple O-mannosylation. **b**, *P. pastoris*-produced recombinant Csa2 binding of haemin measured by isothermal calorimetry. A representative heat signal (upper panel) shows the exothermic peaks that correspond to 19 injections (4 μl) of 140 μM apo-Csa2 solution into 20 μM haemin performed at 150 second intervals. To derive the dissociation constant, the binding isotherm (lower panel) of the normalized heats as a function of the molar ratio was fit to a single-binding-site model using the Origin software package. The indicated affinity is the average of three such independent experiments. **c**, Haem extraction from haemoglobin covalently bound to Sepharose. The indicated proteins were incubated for 30 minutes at room temperature with the Sepharose beads. The absorbance spectrum of the supernatant is shown, in which the 406 nm peak indicates the extraction of haem from the haemoglobin Sepharose and binding to the soluble haemophores. Albumin, a haem-binding protein, was used as a negative control for the extraction. **d**, Strains deleted of *CSA2*, *RBT5* and *PGA7* and double-deletion strains were grown for two days in iron-limited medium (YPD + 1 mM Fz) with increasing concentrations of haemoglobin. Each curve represents the average of three cultures. Vertical bars indicate standard deviations. The strains are KC36, KC131, KC484, KC778, KC779 and KC781. The *pga7*^{-/-} curve is invisible because it runs exactly behind the *pga7*^{-/-} *csa2*^{-/-} curve.

(Supplementary Fig. 1). As previously observed^{12,13}, the *rbt5*^{-/-} mutant required even higher haemoglobin concentrations, and the *pga7*^{-/-} strain barely grew, even in the presence of 5 μM haemoglobin. Moreover, both of these mutations were epistatic to *csa2*^{-/-} (Fig. 1d).

Haem is transferred rapidly between Rbt5 and Pga7 *in vitro*, which suggests a cross-cell-wall haem-transfer cascade¹². To test whether Csa2 can also exchange haem with Rbt5 and Pga7, we

used their differential migration in size-exclusion chromatography (SEC) to trace the transfer of haem from one protein to the other. When holo-Csa2 was mixed with either Rbt5 or Pga7, the haem redistributed between the two proteins, regardless of which protein was pre-loaded with haemin (Fig. 2a,b), which indicates that the postulated Rbt5–Pga7 haem-transfer cascade¹² can be extended to the secreted haemophore Csa2.

Bacterial expression and crystallization of functional Csa2. The structure of the CFEM domain is unknown, and its amino acid sequence is unrelated to any protein of known structure. To establish the structural basis of haem binding by CFEM proteins and to gain insight into the mechanism of haem transfer, we expressed in bacteria a recombinant form of Csa2 that extended from position 33 to the C-terminal residue 147 (hereafter rCsa2), which includes all the sequences conserved in related CFEM proteins (Supplementary Fig. 2). rCsa2 bound haem with a dissociation constant of $0.5 \pm 0.3 \mu\text{M}$, comparable to the affinity measured for Csa2 expressed in *P. pastoris* (Supplementary Fig. 3 and Fig. 1b). Moreover, rCsa2 was able to exchange haem with Rbt5 (Fig. 2c), and to extract haem from a haemoglobin column (Fig. 1c). Finally, the quantity and homogeneity of rCsa2 permitted its crystallization in the presence of haem.

The structure of the rCsa2–haem complex was determined at a resolution of 2.0 Å, using experimental phases derived from the anomalous signals of both the haem-bound iron and the native sulfur atoms of the protein (Methods and Supplementary Table 1). rCsa2 crystallized as an asymmetric homomeric trimer that possessed no rotational symmetry but displayed extensive interaction between its monomers (Fig. 3a and Supplementary Fig. 4a). However, protein crosslinking analysis was consistent with a dimeric assembly, regardless of the presence or absence of haem (Supplementary Fig. 4b). Moreover, multiangle light scattering in line with SEC (SEC-MALS) analyses of rCsa2 (13.2 kDa) also indicated a dimer assembly for both the apo and holo protein in solution (22.7 ± 0.8 and 22.9 ± 0.4 kDa, respectively) (Supplementary Fig. 4c). The SEC-MALS analysis with fungal-expressed Csa2 (18.5–19.2 kDa) (Supplementary Fig. 4d) was consistent with an apo dimer (30 ± 4 kDa), but less conclusive for the holo assembly (39 ± 3 kDa). Thus, the prevalences of the crystallized trimer in solution and/or the geometrical assembly of the dimer measured in solution remain to be established by future structures of different lattice packing.

The CFEM domain adopts a novel ‘helical-basket’ fold. The crystal structure of *C. albicans* rCsa2 represents the first experimental three-dimensional (3D) description of the CFEM family of proteins (Fig. 3b). No similar tertiary structures could be identified in the Protein Structure Database (PDB) using the Dali server¹⁸. Global envelope similarity was detected with some of the high-scoring protein domains; however, the topology, that is, the number of helices, their connectivity and the relative orientation between them, was very distinct (Supplementary Fig. 5), which suggests that the CFEM structure displays a novel protein fold.

The structure comprises six α -helices, the third of which is perpendicular to a central bundle of five antiparallel helices (Fig. 3b,c). This bundle assembly of helices is different from the orthogonal helical fold of globins. Overall, the shape of the helical portion resembles a basket that is ‘capped’ from the top by Helix 3, whereas the long N-terminal loop (positions 34–49) appears as its ‘handle’; we therefore named it the helical-basket fold. The basket structure is kept compact by means of hydrophobic interactions that involve all the helices, with only a single water molecule in the basket core (which satisfies the indole nitrogen of Trp128 (Supplementary Fig. 6)).

The uniqueness of the helical-basket fold appears to stem primarily from the formation of four distinctive disulfide bonds that probably contribute to its rigidity and compactness, as well as to the distinctive orientation between its helices (Fig. 3c). These S–S bridges are formed between the eight canonical cysteine residues of the CFEM domain⁹ (Supplementary Fig. 2). Topologically, helices 2 and 4 are held together via two disulfide bridges, one formed by the first and the seventh cysteine residues (C1–C7) and the second by

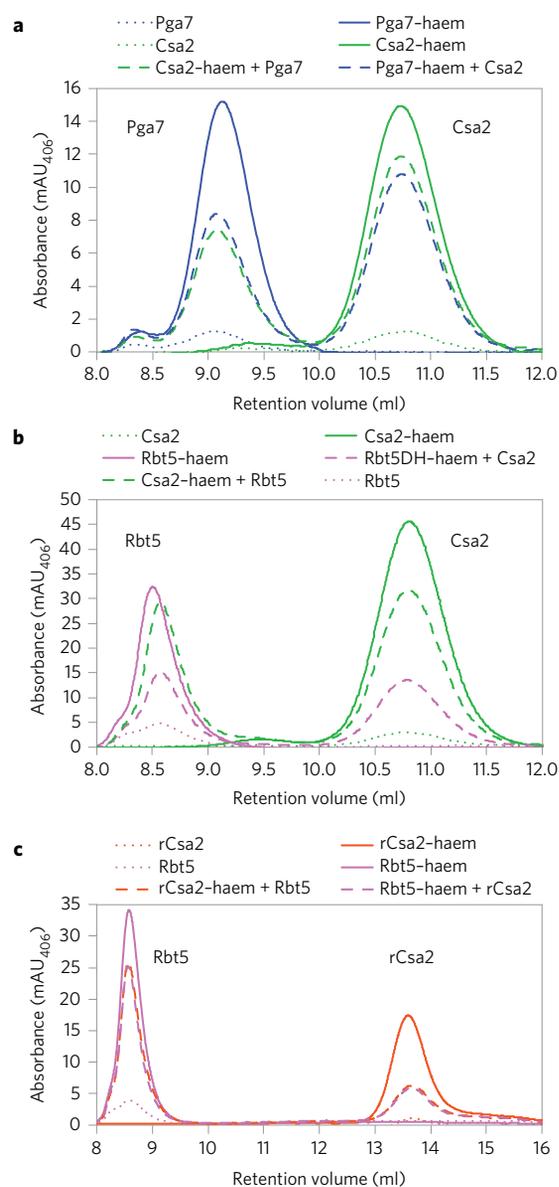


Figure 2 | Csa2 and rCsa2 exchange haem with Rbt5 and Pga7. **a**, Pga7 or Csa2 (approximately 50 μM each) were pre-loaded separately with 10 μM haemin, that is, ferric haem coordinated to chloride (which becomes displaced by the specific haem-iron coordinating residues of the protein) and run on a SEC column. Where indicated, the protein–haem complex was mixed with the second protein in apo-form, and incubated for five minutes at room temperature before loading on the SEC column. The protein-bound haem was detected by monitoring absorbance at 406 nm, the Soret peak of haem bound to the CFEM protein. **b**, Same as for **a**, but with Rbt5 instead of Csa2, and Csa2 was pre-loaded with 20 μM haemin. **c**, Same as for **a**, but with rCsa2 and Rbt5. The ‘rCsa2–haem + Rbt5’ and ‘Rbt5–haem + rCsa2’ curves overlap.

the second and sixth (C2–C6). Helices 3 and 5 are linked via a single disulfide bond (C5–C8), and Helix 3 is linked to an important haem-interacting loop via the C3–C4 disulfide bridge.

The CFEM domain has a peripheral binding site for haem.

Crystals of rCsa2 in the apo form did not diffract, and only the holo-rCsa2, that is, rCsa2 complexed with haem B, diffracted well. Unlike haemoproteins that bind haem as a prosthetic group buried at the core of a protein domain, in the case of CFEM, the haem is bound outside the helical core, between the helical basket

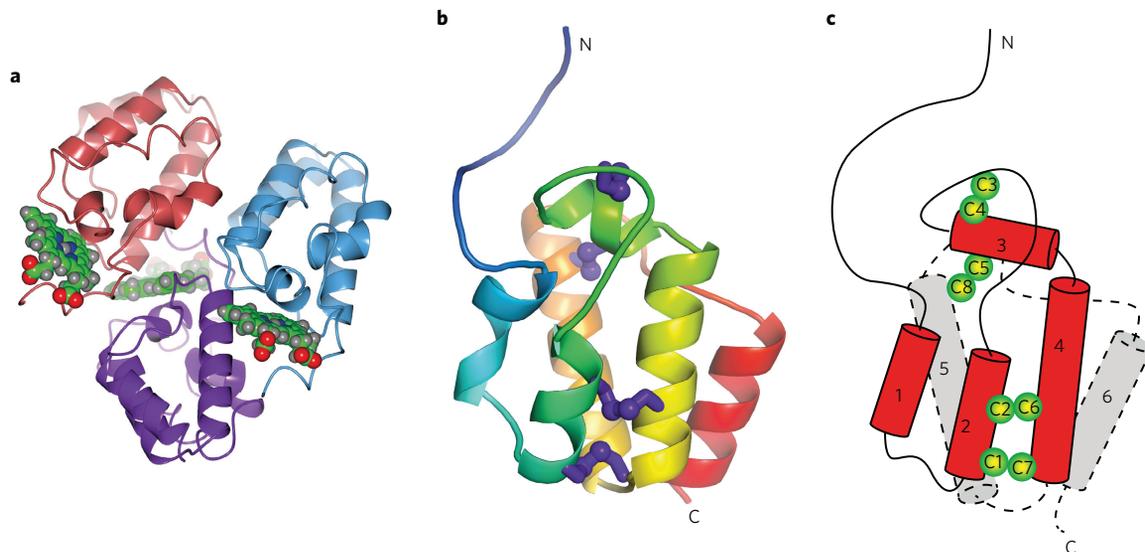


Figure 3 | The crystal structure of the Csa2 CFEM protein reveals a novel helical-basket fold. **a**, The holo-Csa2 crystal structure has a trimer in the ASU with no apparent rotational symmetry between homomers, which are coloured by chain. The haem groups are shown as green sphere models. **b**, The 3D structure of the CFEM domain is made up of six α -helices arranged in a compact helical-basket shape with an elongated N-terminal loop as its handle. The structure is colour coded with a gradient from blue at the N-terminus to red at the C-terminus, and the four unique disulfide bridges are coloured purple. For clarity, the bound haem is not displayed in this panel. **c**, The 2D topology diagram with a similar orientation shows four helices in the front (red) and two in the back (grey). Here the cysteine residues, numbered by their sequential position (C1 through C8), are drawn as green spheres.

and the N-terminal handle (Fig. 4a), and consequently has a large solvent-accessible area (~33% (Supplementary Table 2)). This is consistent with a reversible ligand mode of binding suitable for the relatively rapid association and dissociation of the haem, as is common to haemophores.

Limited proteolysis of rCsa2 revealed that the haem-bound protein is significantly more resistant to cleavage by trypsin than the apo form as it requires at least an 80-fold higher trypsin concentration to attain full cleavage (Fig. 4b). Mass spectrometric analysis of the tryptic digestion products (Supplementary Fig. 7) located the cleavage site protected by haem binding to the N-terminal handle, between residues K44 and T45 (Fig. 4a). This observation suggests that the N-terminal handle—extending from the N-terminal residue 33 to position 49—is flexible in the apo form. Haem binding appears to stabilize the N-terminal handle to its well-resolved conformation seen in the crystal structure.

CFEM proteins have unique haem-iron coordination chemistry.

The wealth of available haemoprotein structures indicates that the predominant protein axial ligand for the coordination of haem-bound iron is a His residue (~80%), but other side chains, such as Met (~7%), Cys (~6%), Tyr (~4%) and Lys (~2%), have also been observed¹⁹. Surprisingly, the CFEM domain has an aspartic acid side chain (D80 in Csa2) as its axial ligand (Fig. 4c). To our knowledge, this is the first documentation of a negatively charged residue as an axial protein ligand for haem-iron coordination. The proximal residues that surround the axial Asp form a flat hydrophobic platform (Fig. 4d) onto which haem intimately docks at van der Waals distances. Consistent with its role in haem binding, this almost sequential region of Csa2 is the most conserved by comparison with other *Candidaceae* CFEM sequences (Supplementary Figs 2 and 8b). At the opposite side of the haem plane, the sixth iron-coordination position typical of the octahedral coordination geometry is vacant. However, a Tyr residue (Y36 of Csa2) located on the N-terminal handle occupies this position and, apparently, interacts with the haem through an uncharacterized haem- π -stacking interaction (Fig. 4e).

Role of the aspartic acid residue as an axial iron-coordination ligand.

In addition to the eight defining CFEM cysteines, the aspartic residue that coordinates the haem iron is one of the most-conserved amino acids in the CFEM proteins. As such, previously it had been mutated to alanine in Rbt5 and Pga7, and found to be essential for haem binding¹². We therefore studied the effect of replacing this Asp with His, the prevalent haem-iron coordinating residue, on haem binding *in vitro* and on the function of the CFEM proteins *in vivo*. Such His-for-Asp substitutions in the *C. albicans* CFEM proteins Csa2, Rbt5 and Pga7 (referred to as DH mutants) yielded proteins that bound haem efficiently *in vitro* based on Soret bandshifts (Fig. 5a). Furthermore, these DH mutant proteins were capable of extracting haem from haemoglobin as effectively as the WT proteins (Fig. 5b).

The exceptional observation of a hard-base axial ligand—the carboxyl moiety of Asp—rather than the soft Lewis bases commonly found as haem-iron ligands prompted us to test the preference of Csa2 for interaction with Fe^{3+} (a strong Lewis acid) over interaction with Fe^{2+} (a soft Lewis acid known to favour soft base ligands²⁰). As the haemin used in our binding assays contains ferrihaem (Fe^{3+}), we assumed that the CFEM-bound haem is oxidized (Fe^{3+}) as well. We thus tested for the ability of haem-bound Csa2 to retain haem binding on its reduction to ferrohaem (Fe^{2+}) by migration through a dithionite layer in a gel-filtration column²¹. Whereas the WT Csa2 lost its bound haem on reduction, the histidine-substituted mutant (D80H) retained its haem (Fig. 5c). Together, these results strongly suggest that the loss of haem binding under reduced conditions is attributable to the inability of the Asp80 to coordinate Fe^{2+} -haem, rather than to an alteration of protein conformation by dithionite reduction.

Surprisingly, despite the ability of the histidine-substituted mutant proteins to bind haem *in vitro*, the DH mutants of RBT5, PGA7 and CSA2 were totally inactive *in vivo* (Fig. 5d). Therefore, we tested further the ability of the DH mutant proteins to exchange haem between themselves and with WT proteins *in vitro*. Mixing WT and DH mutant proteins resulted in limited or no haem exchange, or a transfer in one direction only (Supplementary

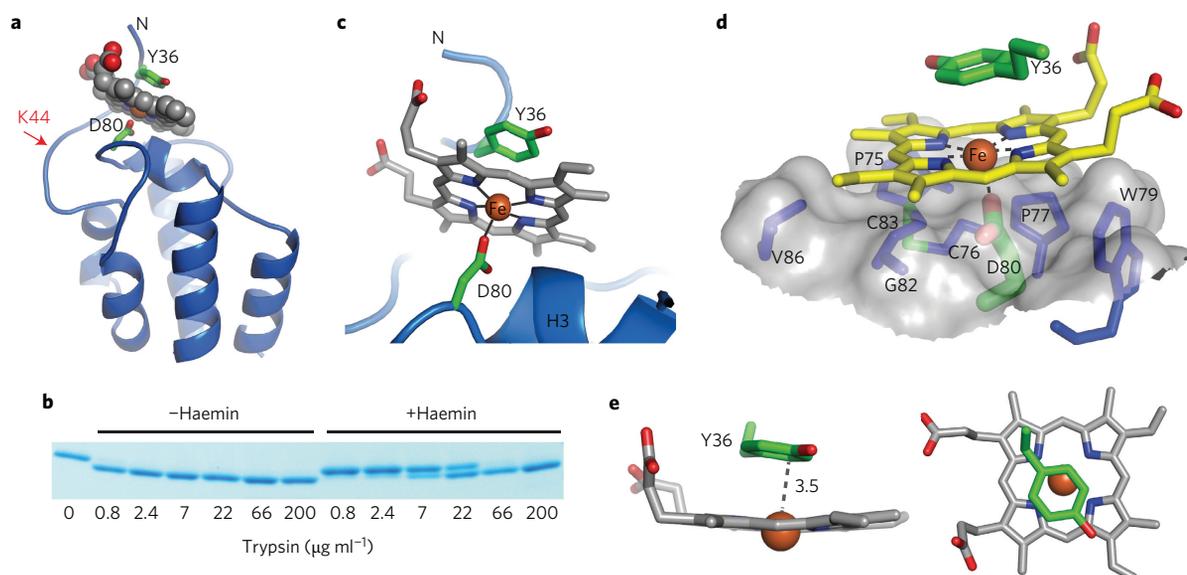


Figure 4 | CFEM has a unique binding site for haem and an exceptional iron coordination. **a**, Ribbon representation of the Csa2 structure complexed with the haem represented in a spheres format. The view has a similar orientation to that in Fig. 3b, but rotated clockwise $\sim 25^\circ$ around the vertical axis. The red arrow indicates the haem-sensitive trypsin cleavage site near the N-terminus of the protein. **b**, Trypsin digestion of rCsa2 (0.2 mg ml⁻¹) with increasing concentrations of the protease, with or without 30 μM haemin. **c**, Asp80 is the axial ligand of the haem-iron. **d**, Csa2 utilizes a conserved flat-platform determinant for binding haem that consists of proximal residues (P75, C76, P77, W79, G82, C83 and V86) surrounding the axial Asp (D80). The platform is shown as a grey surface, and the underlying residues forming it are drawn in blue. The haem is drawn in yellow, whereas the Asp axial ligand and the haem-stacking Tyr are coloured green. The numbering corresponds to the Csa2 sequence. **e**, The second canonical octahedral coordination position of the haem lacks a discrete coordinating atom, but is occupied by a Tyr residue (Y36) which stacks against haem from the top, with its aromatic plane approximately parallel to that of the haem molecule, which gives rise to unusual haem- π -stacking interactions. Although Phe, Tyr and His residues have been observed in the vicinity of protein-bound haem⁴⁰, they mostly stack against the haem pyrrole rings, probably via π - π interactions. Here Tyr36 is aligned more centrally above the haem, and is thus less likely to participate in π - π interactions, but apparently forms an uncharacterized haem- π -stacking interaction (although the observed distance agrees with cation- π interactions, the latter interactions commonly involve quaternary amines and/or cations of alkali metals⁴¹, and we are unaware of examples in which iron is involved in cation- π interactions).

Fig. 9). When Csa2DH was mixed with either Rbt5DH or Pga7DH, some transfer did occur. However, unlike the exchange between WT proteins, which showed the same final distribution regardless of the initial location of the haem (Fig. 2a,b), with the mutant proteins, the distribution depended on the initial location of the haem, which indicates that the redistribution had not reached equilibrium (Fig. 5e). Mixing of holo-Csa2DH with apo-Rbt5DH or apo-Pga7DH yielded intermediate migration peaks in addition to the holo-Csa2DH peak, consistent with the formation of stable heterologous complexes, which were not detectable with the WT proteins. We conclude that haem transfer is much less efficient between histidine-substituted CFEM mutants.

Discussion

We report here the first crystal structure of a CFEM protein, the secreted *C. albicans* haemophore Csa2. Like the two previously characterized CFEM proteins Rbt5 and Pga7, Csa2 is able to extract haem from haemoglobin and to exchange it with its paralogues. As the three CFEM proteins are differentially located, that is, Csa2 is secreted, Rbt5 is peripherally located on the cell wall and Pga7 is localized more internally in the cell envelope (Fig. 1 and Kuznets *et al.* model in which haem is extracted from haemoglobin in the extracellular medium or at the periphery of the cell wall, and is then transferred from the medium and across the cell wall towards the cell membrane by sequential binding to soluble and anchored CFEM proteins (Fig. 6). Csa2, Rbt5 and Pga7 are closely related proteins (Supplementary Figs 2 and 8). A phylogenetic analysis indicates that many Candidaceae have at least one homologue of each of the three CFEM proteins (Supplementary Fig. 8), which suggests that the Csa2-Rbt5-Pga7

haem-transfer cascade is conserved in many fungal species, including saprophytic species. This probably reflects the fact that haem represents an available iron source in many environments, not only in the parasitic context.

The effect of the deletion of CSA2 on haemoglobin-iron utilization was relatively small compared with that of *RBT5* and *PGA7*. However, the observation that a secreted haemophore is required at all for haem-iron utilization in a pure culture was in itself surprising, given that haemoglobin can diffuse freely to the cell surface, where it could be utilized directly by Rbt5 and Pga7. It is possible that Csa2 plays a role in haem transfer between cells in a biofilm, such as may be formed in our microtitre-well-based growth assays, or, alternatively, that it contributes to the transfer of haem between the CFEM proteins embedded in the cell-wall matrix. Furthermore, in natural environments with mixed microbial populations, Csa2 could help *C. albicans* compete with other haem-scavenging organisms.

To gain insight into the CFEM haem extraction and transfer mechanisms, we determined the crystal structure of Csa2 expressed in bacteria, rCsa2. The structure of Csa2 is unrelated to that of any known protein, and in particular to any known bacterial haemophores (reviewed in Contreras *et al.*²²). Gram-negative bacteria secrete a well-characterized haemophore, HasA, that coordinates haem iron via a histidine and a tyrosine residue²³. In Gram-positive bacteria, a haem-utilization system, first identified in *Staphylococcus aureus*²⁴, consists of a cascade of cell-envelope-attached proteins, all of which contain one or more near-iron transporter (NEAT) domains (reviewed in Grigg *et al.*²⁵), whereas *B. anthracis* also secretes a soluble NEAT-domain-containing haemophore, IsdX (ref. 7). The crystal structure of the NEAT domains reveals iron coordination via tyrosine as an axial ligand^{7,25,26}. Although Csa2 also utilizes a tyrosine

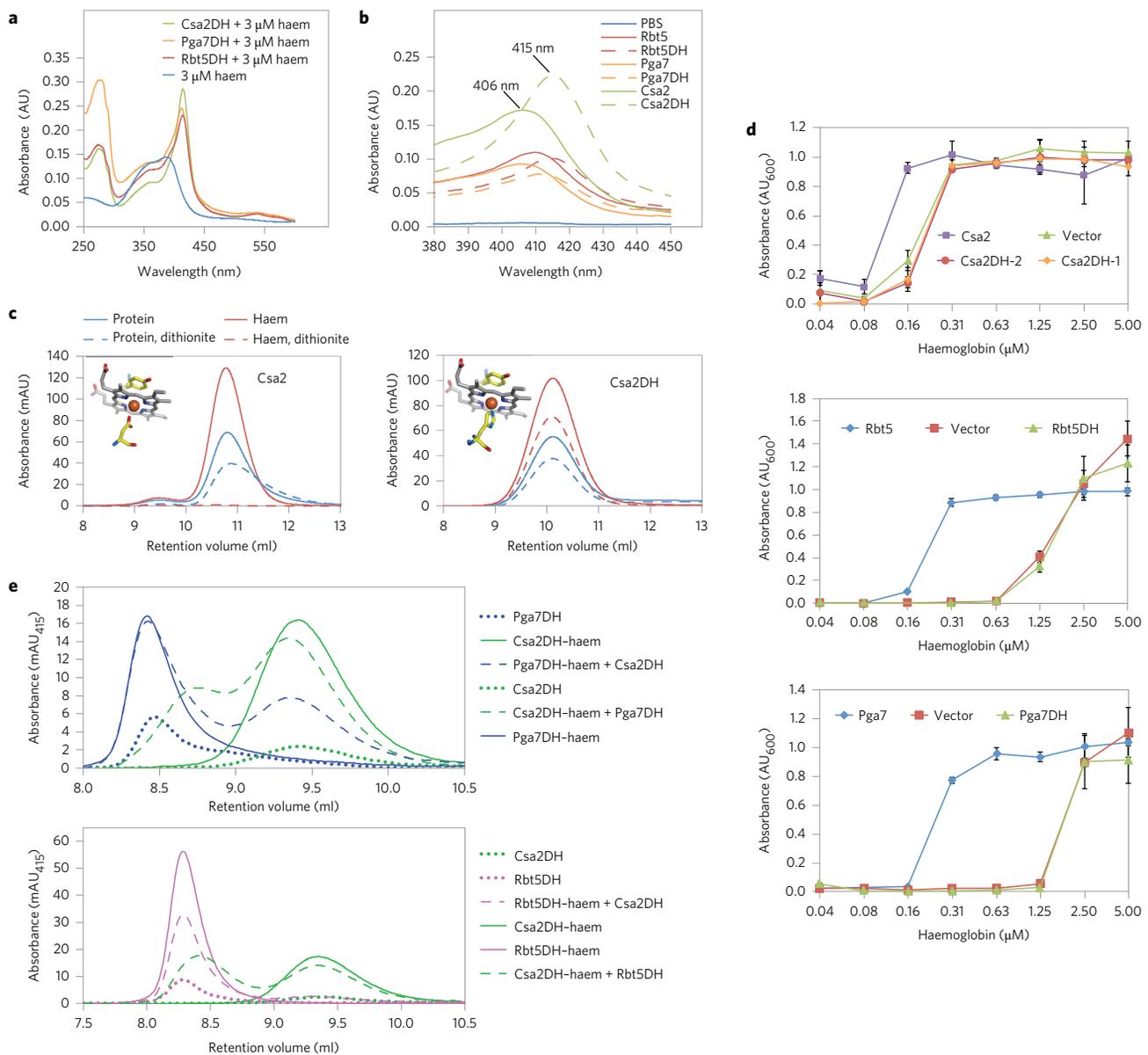


Figure 5 | Histidine-substituted CFEM mutants bind haem and extract haem from haemoglobin, but are inactive and defective in haem exchange.

a, UV-vis spectrum of the indicated histidine-substituted mutant proteins (DH) loaded with $3 \mu\text{M}$ haemin versus haemin alone. The mutant proteins have absorbance maxima of 414 nm (Csa2D80H, Rbt5D72H) and 412 nm (Pga7D63H) versus 406 nm for the WT proteins. **b**, Extraction of haem from a haemoglobin column by WT versus histidine-substituted mutants. **c**, Dithionite treatment of Csa2 or Csa2DH. Holo proteins (1 mg ml^{-1} , $50 \mu\text{l}$) were migrated through a SEC column with (dashed lines) or without (continuous lines) prior injection of $100 \mu\text{l}$ of 10% dithionite. Protein was detected by absorbance at 280 nm (blue curves), and haem was detected at 406 nm (WT protein) or 415 nm (mutant protein) or 415 nm (mutant protein) (red curves). The Csa2DH model was obtained by manual mutation of Asp80, and by selecting a probable His rotamer that best supports Fe coordination geometry. **d**, To test the effect of the histidine substitution on the function of the CFEM proteins *in vivo*, *C. albicans* deletion mutant strains of each of the three CFEM genes, *RBT5*, *PGA7* and *CSA2*, were complemented with the corresponding WT or histidine-substituted mutant alleles. The *rbt5*^{-/-}, *pga7*^{-/-} and *csa2*^{-/-} strains were transformed with their WT genes or with the *RBT5* D72H, *PGA7* D63H or *CSA2* D80H alleles, and grown as described for Fig. 1d. Each curve represents the average of three cultures. Vertical bars indicate standard deviations. The strains used are KC787, KC812 and KC940 (two independent clones) (top), KC922, KC955 and KC956 (middle), and KC536, KC904 and KC921 (bottom). **e**, Haemin-exchange experiment with the mutant proteins Csa2^{D80H} and Pga7^{D63H} (top) or Csa2^{D80H} and Rbt5^{D72H} (bottom) loaded with $10 \mu\text{M}$ haemin. The experiment was performed as described for Fig. 2.

residue (Tyr36) for haem binding, it does not serve as a coordination ligand for the iron but rather stacks parallel to the haem (Fig. 4e). The utilization of aromatic residues for such haem-stacking appears to be common to CFEM proteins, which possess a Tyr, Trp or Phe residue at this position (Supplementary Fig. 2). Therefore, the mechanism of haem-iron coordination by CFEM proteins is distinct from that of the bacterial haemophores and, as discussed further below, it is also unique among haemoproteins in general.

Three features stand out in the rCsa2 structure. First, the CFEM domain adopts a novel fold, the helical-basket fold, which consists of six α -helices (Fig. 3). The eight characteristic cysteines of the CFEM domain form four disulfide bonds that appear to play a primarily structural role. The top of the helical-basket structure consists of a flat hydrophobic platform made by residues of Helix 3 and of the loop connecting it to Helix 2, which presents a complementary structural determinant for the docking of the planar

overcome host nutritional immunity, may be used to block it for therapeutic benefits.

Methods

Media and growth conditions. *C. albicans* strains were grown in rich medium (YPD (1% yeast extract, 2% Bacto peptone, 2% glucose, 25 $\mu\text{g ml}^{-1}$ uridine)).

Transformants were selected on SC-URA plates (0.17% yeast nitrogen base, 0.5% ammonium sulfate, 2% glucose, 2% agar). For iron starvation, cells were grown in iron-poor medium (IPM) prepared by supplementing YPD with 1 mM ferrozine (Fz (Sigma))^{27,28}.

P. pastoris was grown in YPD. Transformants were selected on YPD supplemented with 100 or 1,000 $\mu\text{g ml}^{-1}$ Zeocin (Invitrogen). For protein expression, buffered glycerol-complex medium (BMGY (1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% YNB, $4 \times 10^{-5}\%$ biotin, 1% glycerol)) and buffered methanol-complex medium (BMMY (1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% YNB, $4 \times 10^{-5}\%$ biotin, 0.5% methanol)) were used.

Haemin and haemoglobin preparation. All stocks were freshly prepared before each assay. For growth assays, stock solutions of 0.1 mM human haemoglobin (Sigma) in PBS, 0.5 mM bovine haemoglobin (Sigma) in PBS or 2 mM haemin (haem B containing a ferric iron coordinated to Cl (Frontier Scientific)) in 0.2 N NaOH were chelated with 5% w/v Chellex100 (Sigma) for two hours before use. For all the other assays, haemin was prepared in $1 \times$ PBS in the presence of 10 mM NaOH. As haemoglobin (ferrohaem, Fe^{2+}) is rapidly oxidized in solution to methaemoglobin (haem, Fe^{3+}), in all our assays the predominant haemoglobin species was methaemoglobin.

Plasmids and strains. Strain genotypes are summarized in Supplementary Table 3. The CSA2 knockout plasmids KB2221/2 were constructed by cloning the CSA2 upstream (−626 to −1, HindIII–PstI) and downstream (+447 to +900; SpeI–SacI) regions into plasmids KB985/6 (ref. 29). For CSA2 reintegration, the CSA2 region was cloned as a HindIII–KpnI fragment (−929 to +771) into pBES116 (ref. 30) to generate KB2353. PCR mutagenesis was used to introduce the D80H allele (G238C) into KB2353, which generated KB2454. For RBT5 reintegration, the RBT5 region was cloned as a NotI–KpnI fragment (−1,170 to +1,673) into pBES116 (ref. 30) to generate KB1664. PCR mutagenesis was used to introduce the D72H allele (G214C) into KB1664, to generate KB2436. Plasmid KB2435 was generated by PCR mutagenesis of PGA7 of KB2111 (ref. 12) to $\text{PGA7}^{\text{D63H}}$ (G187C). The CSA2 coding sequence (residues 33–147) was cloned between NdeI–XhoI into pET22b+ (Novagen) to generate the *E. coli* expression plasmid KB2386. The CSA2 and CSA2^{D80H} coding sequences (residues 19–147) were cloned into picZalpha (Invitrogen) to generate the *Pichia* expression plasmids KB2366 and KB2412. The PGA7 and RBT5 expression plasmids KB2259 and KB2258 (ref. 12) were mutagenized by PCR to the $\text{PGA7}^{\text{D63H}}$ and $\text{RBT5}^{\text{D72H}}$ alleles to generate KB2437 and KB2449.

Growth assays. Overnight cultures grown in YPD were diluted in the morning into a series of twofold dilutions of haemoglobin in IPM. The various strains were inoculated, in triplicate, at a density corresponding to 0.005 MAU (600 nm) in IPM in flat-bottomed 96-well plates. Plates were incubated at 30 °C with 60 revolutions per minute (r.p.m.) shaking. Growth was measured as OD₆₀₀ with an enzyme-linked immunosorbent-assay reader after three days.

Recombinant protein expression and purification from *P. pastoris* supernatant. Csa2, Rbt5 and Pga7 were expressed as C-terminal Myc–6xHis tagged proteins in the *P. pastoris* expression system as described in the *Pichia* expression manual (Invitrogen). Briefly, cells from an overnight culture in BMGY were washed in two volumes of water and diluted in BMMY to OD₆₀₀ 0.5–1. Cultures were then incubated at 28 °C with vigorous shaking for 96 hours, with the addition of 0.5% methanol every 24 hours. The cultures were centrifuged for 30 minutes at 1,500 r.p.m., 10 mM imidazole was added to the supernatant and the pH was adjusted to 8.0. The medium was centrifuged at 13,000 r.p.m. for 20 minutes and the supernatant was mixed with 6 ml of 50% Ni-NTA (nitrilotriacetic acid) resin (Qiagen), incubated for one hour at room temperature, washed with wash/elution buffer (50 mM Tris, 50 mM NaH_2PO_4 , 300 mM NaCl, pH 8.0) that contained 10 mM imidazole (1 \times), 20 mM imidazole (2 \times) and 50 mM imidazole (2 \times). Finally, Ni-NTA-bound proteins were eluted by incubating the beads three times with two bed volumes of 250 mM imidazole. The eluate was concentrated with a Centricon 3k cutoff filter, and then dialysed twice against two litres of PBS. To remove bound haem and obtain pure apo-protein, the proteins were treated with 3 M imidazole by tilting for two hours at 4 °C, centrifuged for 30 minutes at 4 °C and then the supernatant was re-dialysed.

Production in *E. coli* and crystallization of rCsa2. Rbt5, Pga7 and Csa2 expressed in fungi appear to be heavily glycosylated, based on their aberrant migration in SDS-PAGE^{12,13} (Fig. 1a) and on mass spectrometry analysis, which indicated up to 18 hexose adducts per Csa2 molecule. To obtain non-glycosylated proteins that are

more amenable to crystallization, we expressed Rbt5, Pga7 and Csa2 in bacteria. Although all three proteins could be expressed in *E. coli*, they were largely inactive based on spectroscopic measurements of haem binding. Only Csa2, when expressed in the *trxB-gor* strain that permits disulfide-bond formation within the bacterial cytoplasm¹⁸, yielded an active protein species. This bacterial recombinant form, positions 33–147 (rCsa2), lacks the first 14 residues (following the predicted signal peptide cleavage site¹⁹ at position 19) that are present in Csa2 orthologues but absent in Rbt5 and Pga7 orthologues (Supplementary Fig. 2). rCsa2 was expressed in *E. coli* Origami B(DE3) transformed with the plasmid KB2386, a T7-expression vector that contains a sequence that extends from codon 33 of the CSA2 coding sequence to the last CSA2 codon (codon 147), followed by a hexahistidine tag. Briefly, bacterial culture was grown at OD₆₀₀ = 0.6 in 1.5 litres of 2 \times YT (yeast extract and triptone) medium that contained 0.05 mg ml^{−1} of ampicillin and 0.025 mg ml^{−1} of chloramphenicol, followed by four hours of induction with 1 mM isopropyl β -D-1-thiogalactopyranoside at 37 °C. Cells were then centrifuged (5,000 r.p.m. for ten minutes), resuspended in PBS and centrifuged again, and the bacterial pellets were frozen at −20 °C overnight. The pellets were then thawed, resuspended in 8 ml lysis buffer (50 mM Tris, 50 mM NaH_2PO_4 , 300 mM NaCl, 8 M urea, 10 mM imidazole, pH 8.0) and incubated for one hour at room temperature. The lysate was then centrifuged in a microcentrifuge for 30 minutes at 4 °C. The supernatant was mixed with 6 ml of 50% Ni-NTA resin (Qiagen), incubated for one hour at room temperature, washed with wash/elution buffer (50 mM Tris, 50 mM NaH_2PO_4 , 300 mM NaCl, pH 8.0) that contained 10 mM imidazole (1 \times), 20 mM imidazole (2 \times) and 50 mM imidazole (2 \times). Finally, rCsa2 was eluted by incubating the beads sequentially three times with two volumes of 250 mM imidazole. The eluate was dialysed twice against two litres of 10 mM NaCl, 20 mM Tris pH 8.5 and then loaded on a Hitrap Q HP column (GE Healthcare) at pH 8.5 and eluted with an NaCl gradient. The gradient fractions that contained rCsa2 (20–60 mM NaCl) were concentrated and repurified on a Sephacryl S-200 HR 16/60 SEC column (GE Healthcare) to yield 5–10 mg of 99% pure protein as determined by ultraviolet (280 nm) absorption and an extinction coefficient of 23,500 M^{−1} cm^{−1}. To obtain holo protein, before crystallization rCsa2 was mixed with an equimolar amount of haemin using a 5 mM haemin stock (Sigma-Aldrich) solubilized in 50 mM NaOH. The NaOH was neutralized with HCl, followed by ten minutes of centrifugation at 20,000g to remove aggregates.

Crystals grew by the hanging-drop vapour-diffusion method at 20 °C by mixing the protein sample (10 mg ml^{−1} supplemented with 5% *n*-octyl- β -D-glucoside) at a 2:1 ratio with the reservoir solution (0.1 M HEPES pH 7.5, 8% ethylene glycol, 13.75% poly(ethylene glycol) 10,000). Prior to the diffraction experiments crystals were manually fished out of their mother liquor using nylon CryoLoop (Hampton Research) and flash-frozen in liquid nitrogen.

Crystal-structure determination. X-ray data were collected at 100 K in-house on a Rigaku FR-X rotating anode generator coupled to an R-Axis HTC image-plate detector. X-ray data that extend to a resolution of 2 Å were indexed, integrated, merged and scaled using HKL3000³¹ in space group $P2_12_12_1$ (Supplementary Table 1). The highly redundant data set collected for phasing based on small anomalous scatterers resulted in an R_{merge} of 0.625 at the highest resolution shell, but the data were highly significant and extended safely to 2 Å resolution, as indicated by the high $CC_{1/2}$ (Supplementary Table 1). The structure was solved experimentally by the single anomalous dispersion method. Briefly, a substructure comprising five anomalously scattering sites that corresponded to three Fe atoms and two of the total of 12 disulfide S–S bonds of the asymmetric unit (ASU) was located using SHELXD³². Refinement of those sites and phasing was carried out using MLPHARE³³ followed by density modification using DM³⁴. Automated model building using Buccaneer³⁵ docked 85% of the protein sequence in the experimental electron-density maps, and the rest of the atomic model, including three haem B ligands, was fitted manually using Coot³⁶. The structure, containing three holo-Csa2–haem B complexes in the ASU, was refined using the PHENIX interface³⁷. This included several cycles of restrained refinement with simulated annealing, torsion-angle non-crystallographic symmetry restraints between the three Csa2 molecules and individual atomic displacement parameters (ADPs) refinement, each followed by a manual model fitting and the addition of structured water molecules. As the structure converged, a combined TLS (translation/libration/screw) and individual ADPs refinement protocol was used until the *R*-values converged to *R* = 15.26% and a free-*R* = 19.16% (Supplementary Table 1). The coordinates and structure factors have been deposited in the PDB (4Y7S). Structural illustrations were made using PyMOL³⁸.

Haemin extraction from haemoglobin. Bovine haemoglobin was covalently bound to CnBr-activated Sepharose-4B beads (Sigma). The haemoglobin beads were washed extensively before incubation with the proteins to remove any spontaneously released haem. The protein (6 μl) was mixed with 2 μl of haemoglobin beads, then the samples were tilted for 30 minutes at room temperature and an ultraviolet-visible (UV–vis) spectrum of the supernatant was taken using a Thermo Scientific NanoDrop2000C (BD bioscience). The average spectrum of a triplicate reading was plotted for each sample.

Isothermal titration calorimetry (ITC). Titrations were performed at 25 °C using a MicroCal iTC200 system (GE Healthcare). The protein stocks and a freshly prepared haemin stock solution were diluted in PBS. All the injections were carried out at 150 second intervals. To prevent/minimize haem adsorption, the calorimeter cell and the microsyringe used for the injections were washed extensively with 10 N NaOH after each experiment. Protein concentration was determined using a NanoDrop spectrophotometer, and haemin stock concentration was determined by absorbance at 398 nm in an organic solution³⁹. For the titration experiments, the concentrations of protein were 140 μM in the syringe and 20 μM haemin in the cell. The resulting titration data were analysed and fitted using the Origin for ITC software package supplied by MicroCal to obtain the stoichiometry (*n*), the dissociation constants (*K_D*) and the enthalpy (ΔH) and entropy (ΔS) changes of binding. For the fit, any constraints on the stoichiometry and ΔH were not fixed.

Absorption spectroscopy. UV–vis absorption spectra were collected using a Cary 60 scanning spectrophotometer (GE Healthcare) and a 1 cm pathlength 1 ml quartz cuvette. The haemin stock concentration was measured as described above. The protein concentrations were determined by the absorbance at 280 nm using predicted absorption coefficients based on sequence ($\epsilon = 23,470 \text{ mM}^{-1}$ for rCsa2).

Analytical SEC and SEC-MALS. Samples (50 μl) were injected on a Superdex 75 10/300 GL column (GE Healthcare) equilibrated with PBS, and monitored on the ÄKTApurifier system. Absorbance was recorded at 280 nm, 406 nm and/or 415 nm for the DH mutants. For the haem-exchange experiments, 25 μl of each protein, at 1 mg ml⁻¹ for Rbt5 and Pga7, or 0.7 mg ml⁻¹ for Csa2 (~50 μM), were mixed with each other or with PBS and incubated for five minutes at room temperature before loading on the column. For mass calculation, a miniDawn TREOS MALS detector together with an Optilab T-rEX dRI detector (Wyatt technology) were used in line with an ÄKTA Explorer Superdex 75 10/300 column (Wolfson Center for Structural Biology, Hebrew University of Jerusalem).

Trypsin digestion. Trypsin (TPCK-treated (Sigma)) was dissolved in PBS to 1 mg ml⁻¹ and diluted sequentially 1:2 in PBS. An aliquot of 20 μl of 0.2 mg ml⁻¹ protein stock in PBS (rCsa2 or Csa2, with or without equimolar haemin) was mixed with 5 μl of each trypsin concentration, incubated for 30 minutes at 22 °C and then the reaction was stopped by boiling in protein SDS loading buffer that contained 2 mM phenylmethanesulfonyl fluoride. The extracts were separated by SDS–PAGE, the gels were stained by Coomassie and the unprocessed and fully processed rCsa2 bands were cut out and submitted to mass spectroscopy analysis (Smoler Proteomics Center).

Csa2 antibodies. Rabbit polyclonal antibodies (Adar Biotech) were raised against *E. coli*-produced rCsa2, and used at 1:1,000 titre on western blot membranes.

Data availability. The crystal structure of the CFEM protein Csa2 been deposited in the PDB under accession number 4Y7S.

Received 12 April 2016; accepted 26 July 2016;
published 12 September 2016

References

- Weinberg, E. D. Nutritional immunity. Host's attempt to withhold iron from microbial invaders. *J. Am. Med. Assoc.* **231**, 39–41 (1975).
- Andrews, N. C. Iron homeostasis: insights from genetics and animal models. *Nat. Rev. Genet.* **1**, 208–217 (2000).
- Almeida, R. S., Wilson, D. & Hube, B. *Candida albicans* iron acquisition within the host. *FEMS Yeast Res.* **9**, 1000–1012 (2009).
- Cassat, J. E. & Skaar, E. P. Iron in infection and immunity. *Cell Host Microbe* **13**, 509–519 (2013).
- Caza, M. & Kronstad, J. W. Shared and distinct mechanisms of iron acquisition by bacterial and fungal pathogens of humans. *Front. Cell Infect. Microbiol.* **3**, 80 (2013).
- Letoffe, S., Ghigo, J. M. & Wandersman, C. Iron acquisition from heme and hemoglobin by a *Serratia marcescens* extracellular protein. *Proc. Natl Acad. Sci. USA* **91**, 9876–9880 (1994).
- Ekworomadu, M. T. *et al.* Differential function of lip residues in the mechanism and biology of an anthrax hemophore. *PLoS Pathog.* **8**, e1002559 (2012).
- Cadioux, B. *et al.* The mannoprotein Cig1 supports iron acquisition from heme and virulence in the pathogenic fungus *Cryptococcus neoformans*. *J. Infect. Dis.* **207**, 1339–1347 (2013).
- Kulkarni, R. D., Kelkar, H. S. & Dean, R. A. An eight-cysteine-containing CFEM domain unique to a group of fungal membrane proteins. *Trends Biochem. Sci.* **28**, 118–121 (2003).
- Mitchell, A. *et al.* The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* **43**, D213–D221 (2015).
- Ding, C. *et al.* Conserved and divergent roles of Bcr1 and CFEM proteins in *Candida parapsilosis* and *Candida albicans*. *PLoS One* **6**, e28151 (2011).
- Kuznets, G. *et al.* A relay network of extracellular heme-binding proteins drives *C. albicans* iron acquisition from hemoglobin. *PLoS Pathog.* **10**, e1004407 (2014).
- Weissman, Z. & Kornitzer, D. A family of *Candida* cell surface haem-binding proteins involved in haemin and haemoglobin-iron utilization. *Mol. Microbiol.* **53**, 1209–1220 (2004).
- Bailao, E. F. *et al.* Hemoglobin uptake by *Paracoccidioides* spp. is receptor-mediated. *PLoS Negl. Trop. Dis.* **8**, e2856 (2014).
- Amorim-Vaz, S. *et al.* RNA enrichment method for quantitative transcriptional analysis of pathogens *in vivo* applied to the fungus *Candida albicans*. *mBio* **6**, e00942–15 (2015).
- Mochon, A. B. *et al.* Serological profiling of a *Candida albicans* protein microarray reveals permanent host–pathogen interplay and stage-specific responses during candidemia. *PLoS Pathog.* **6**, e1000827 (2010).
- Sorgo, A. G. *et al.* Mass spectrometric analysis of the secretome of *Candida albicans*. *Yeast* **27**, 661–672 (2010).
- Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
- Li, T., Bonkovsky, H. L. & Guo, J. T. Structural analysis of heme proteins: implications for design and prediction. *BMC Struct. Biol.* **11**, 13 (2011).
- Outten, F. W. & Theil, E. C. Iron-based redox switches in biology. *Antioxid. Redox Signal* **11**, 1029–1046 (2009).
- Dixon, H. B. & McIntosh, R. Reduction of methaemoglobin in haemoglobin samples using gel filtration for continuous removal of reaction products. *Nature* **213**, 399–400 (1967).
- Contreras, H., Chim, N., Credali, A. & Goulding, C. W. Heme uptake in bacterial pathogens. *Curr. Opin. Chem. Biol.* **19**, 34–41 (2014).
- Arnoux, P. *et al.* The crystal structure of HasA, a hemophore secreted by *Serratia marcescens*. *Nat. Struct. Biol.* **6**, 516–520 (1999).
- Mazmanian, S. K. *et al.* Passage of heme-iron across the envelope of *Staphylococcus aureus*. *Science* **299**, 906–909 (2003).
- Grigg, J. C., Ukpabi, G., Gaudin, C. F. & Murphy, M. E. Structural biology of heme binding in the *Staphylococcus aureus* Isd system. *J. Inorg. Biochem.* **104**, 341–348 (2010).
- Grigg, J. C., Vermeiren, C. L., Heinrichs, D. E. & Murphy, M. E. Haem recognition by a *Staphylococcus aureus* NEAT domain. *Mol. Microbiol.* **63**, 139–149 (2007).
- Bodanszky, M., Klausner, Y. S. & Said, S. I. Biological activities of synthetic peptides corresponding to fragments of and to the entire sequence of the vasoactive intestinal peptide. *Proc. Natl Acad. Sci. USA* **70**, 382–384 (1973).
- 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**, 1551–1560 (2002).
- Atir-Lande, A., Gildor, T. & Kornitzer, D. Role for the SCF^{CD₄} ubiquitin ligase in *Candida albicans* morphogenesis. *Mol. Biol. Cell* **16**, 2772–2785 (2005).
- Feng, Q., Summers, E., Guo, B. & Fink, G. Ras signaling is required for serum-induced hyphal differentiation in *Candida albicans*. *J. Bacteriol.* **181**, 6339–6346 (1999).
- Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000 the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr. D* **62**, 859–866 (2006).
- Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D* **58**, 1772–1779 (2002).
- Otwinowski, Z. in *Proc. CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering* (eds Wolf, W., Evans, P. R. & Leslie, A. G. W.) 80–85 (Daresbury Laboratory, 1991).
- Cowtan, K. D. & Zhang, K. Y. Density modification for macromolecular phase improvement. *Progr. Biophys. Mol. Biol.* **72**, 245–270 (1999).
- Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D* **62**, 1002–1011 (2006).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
- Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
- DeLano, W. L. *The PyMOL Molecular Graphics System* (DeLano Scientific, 2002).
- Weinstein, J. D. & Beale, S. I. Separate physiological roles and subcellular compartments for two tetrapyrrole biosynthetic pathways in *Euglena gracilis*. *J. Biol. Chem.* **258**, 6799–6807 (1983).
- 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**, 621–630 (2007).
- Zacharias, N. & Dougherty, D. A. Cation–π interactions in ligand recognition and catalysis. *Trends Pharmacol. Sci.* **23**, 281–287 (2002).
- Hargrove, M. S., Whitaker, T., Olson, J. S., Vali, R. J. & Mathews, A. J. Quaternary structure regulates heme dissociation from human hemoglobin. *J. Biol. Chem.* **272**, 17385–9 (1997).

Acknowledgements

We are grateful to D. Hiya (Technion Center for Structural Biology) for help with protein crystallization, T. Ziv (Smoler Proteomics Center) for mass spectroscopy analysis, G. Kuznets and Y. Gorelik for constructing plasmids KB2366 and KB2221/2, U. Roy for producing Csa2 cysteine mutants, M. Lebendiker (Hebrew University of Jerusalem) for

advice with SEC-MALS, R. Zarivach (Ben-Gurion University) for the Origami B strain and O. Lewinson, I. Silman, N. Adir, N. Levanon, A. Haber and S. Selig for discussions and critical reading of the manuscript. This research was supported by grants from the Israel Science Foundation and the Ministry of Health's Chief Scientist Office to D.K. H.D. thanks the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 330879-MC-CHOLESTRUCTURE for financial support.

Author contributions

L.N., Z.W., M.P. and D.K. constructed the plasmids and strains, purified the proteins and performed experiments, H.A. performed the SEC-MALS analysis, H.D. crystallized the

protein and determined its structure, and H.D. and D.K. developed the project, interpreted the data and wrote the paper.

Additional information

Supplementary information is [available for this paper](#). Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to H.D. and D.K.

Competing interests

The authors declare no competing interests.