

SCIENTIFIC REPORTS



OPEN

Structural and functional characterization of the CAP domain of pathogen-related yeast 1 (Pry1) protein

Received: 03 February 2016

Accepted: 09 June 2016

Published: 27 June 2016

Rabih Darwiche^{1,*}, Alan Kelleher^{2,*}, Elissa M. Hudspeth², Roger Schneider¹ & Oluwatoyin A. Asojo²

The production, crystal structure, and functional characterization of the C-terminal cysteine-rich secretory protein/antigen 5/pathogenesis related-1 (CAP) domain of pathogen-related yeast protein-1 (Pry1) from *Saccharomyces cerevisiae* is presented. The CAP domain of Pry1 (Pry1CAP) is functional *in vivo* as its expression restores cholesterol export to yeast mutants lacking endogenous Pry1 and Pry2. Recombinant Pry1CAP forms dimers in solution, is sufficient for *in vitro* cholesterol binding, and has comparable binding properties as full-length Pry1. Two crystal structures of Pry1CAP are reported, one with Mg²⁺ coordinated to the conserved CAP tetrad (His208, Glu215, Glu233 and His250) in spacegroup *I*4₁ and the other without divalent cations in spacegroup *P*6₁22. The latter structure contains four 1,4-dioxane molecules from the crystallization solution, one of which sits in the cholesterol binding site. Both structures reveal that the divalent cation and cholesterol binding sites are connected upon dimerization, providing a structural basis for the observed Mg²⁺-dependent sterol binding by Pry1.

Members of the eukaryotic CAP (cysteine-rich secretory protein/antigen 5/pathogenesis related-1) or SCP/TAPS (Sperm-coating protein/Tpx/antigen 5/pathogenesis related-1/Sc7) superfamily have a wide range of physiological activities including sperm maturation, fertilization, fungal virulence, cellular defense, and immune evasion^{1–6}. However, the underlying mechanisms by which CAP proteins perform these seemingly unrelated functions remain poorly understood. CAP proteins contain a CAP or SCP domain, named after mammalian CAP proteins characterized in semen and seminal fluid^{7–10}. The CAP domain (NCBI domain cd00168 or Pfam PF00188) is an ~15 kDa cysteine-rich domain with limited (<38%) sequence identity^{1–4}. Several CAP protein structures have been reported and they all contain a conserved alpha-beta-alpha sandwich domain^{8,11–20}. The CAP domains differ in the lengths of their strands and helices and ~47% of the domain is composed of loops, making it difficult to accurately predict their structure^{11–13}. The prototypical CAP domain has a central cavity characterized by the cavity tetrad, formed by four key residues from the four CAP motifs: His from CAP1, Glu from CAP2, His from CAP3, and Glu from CAP4 and these tetrad residues bind divalent cations including zinc and magnesium^{8,13,21,22}. CAP1 and CAP2 are defined in the PROSITE database (<http://www.expasy.ch/prosite>) as CRISP motifs, while CAP3 and CAP4 are additional CAP motifs defined by Gibbs and colleagues⁸.

CAP proteins bind lipids; for example Tablysin-15 from a blood-feeding arthropod is an anti-inflammatory scavenger of eicosanoids with a hydrophobic channel that binds leukotrienes with sub-micromolar affinities¹⁹. Structures revealing how palmitate and other lipids bind to this hydrophobic channel have been reported¹⁹. Additionally, GLIPR2/GAPR-1 which is highly over-expressed in glioblastoma multiforme, binds to the surface of liposomes containing negatively charged lipids^{23,24}. Structures showing how GAPR-1 binds up to three phosphatidylinositol molecules strongly enough to resist denaturation or organic solvent extraction have been reported^{23,24}. Our previous studies revealed that *Saccharomyces cerevisiae* CAP proteins are required for the export of sterols *in vivo* and they bind cholesterol *in vitro*^{25,26}.

¹Division of Biochemistry, Department of Biology, University of Fribourg Chemin du Musée 10 CH 1700 Fribourg Switzerland. ²National School of Tropical Medicine, Baylor College of Medicine Houston TX 77030 USA. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to O.A.A. (email: Asojo@bcm.edu)

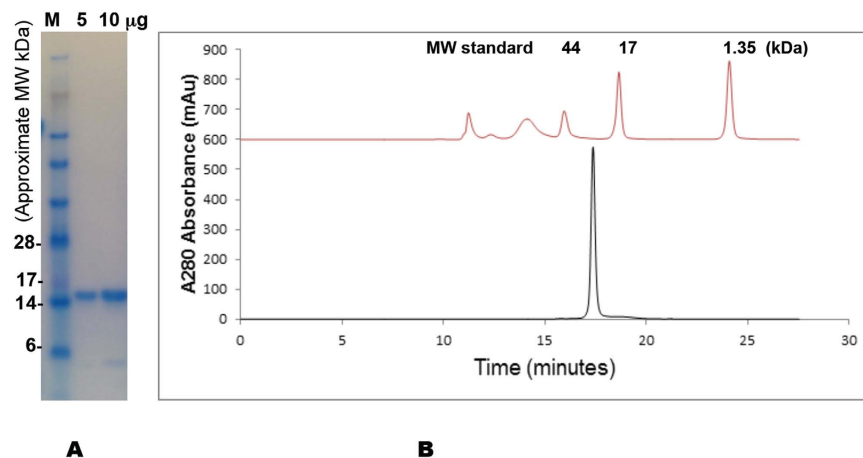


Figure 1. Pry1CAP purification and size exclusion chromatography. (A) The electrophoretic mobility of Pry1CAP on a reducing SDS Page gel reveals a monomeric weight of ~15 kDa, (B) Pry1CAP elutes as an ~31 kDa dimer from a size exclusion column.

The *Saccharomyces cerevisiae* genome encodes three CAP proteins known as Pathogen Related in Yeast (Pry1–3). Pry1 and Pry2 are secreted glycoproteins and Pry3 is associated with the yeast cell wall²⁵. The sterol binding and export properties of these proteins are localized to the CAP domain, which is sufficient to rescue sterol export properties of cells lacking endogenous Pry proteins^{25,26}. Computational modeling suggests that sterol binding to Pry1 occurs through displacement of a flexible loop, the caveolin binding motif (CBM)^{25,26}. While point mutations within the CBM abrogated sterol binding and export, mutations of residues located outside the CBM including highly conserved putative catalytic residues have minimal effect on lipid binding and sterol export²³. These studies defined the CBM as a crucial motif for lipid binding *in vitro* and sterol export *in vivo*^{22,23}. Furthermore, the expression of heterologous CAP proteins rescues the block in sterol export of yeast mutants lacking Pry function, indicating that sterol export is a conserved function of CAP proteins from different species^{26,27}. As part of our ongoing efforts to characterize the lipid binding properties of this ubiquitous superfamily of proteins^{6,11,13,25–28}, we present the structural and functional characterization of the CAP domain of Pry1 (Pry1CAP) from *Saccharomyces cerevisiae*.

Results

Pry1CAP forms dimers in solution. Pry1CAP was overexpressed in *Pichia pastoris* with a yield of approximately 500 mg of ~99% pure protein from a 1 L shake flask culture. Pry1CAP migrates on a reducing Coomassie stained SDS PAGE gel at ~16 kDa (Fig. 1A). As was observed for some CAP proteins including Glipr-1, GAPR-1 and Na-ASP-2, Pry1CAP crystallizes with a monomer in the asymmetric unit while dimerizing in solution^{8,11,12}, eluting from a size exclusion column as a single sharp peak, with a molecular mass of 30.8 kDa, which is twice the theoretical molecular weight of 15.9 kDa (Fig. 1B).

Pry1CAP binds and exports cholesterol. Pry1CAP used in these studies is functional for cholesteryl acetate export *in vivo* because a plasmid encoding Pry1CAP complemented the cholesteryl acetate export defect of mutant yeast cells that lacked endogenous Pry1 and Pry2 (Fig. 2A). The efficiency of cholesteryl acetate export by Pry1CAP was comparable to that of full-length Pry1 as indicated by the similarity of the export indices (Fig. 2B). Addition of an increasing amount of [³H]-cholesterol resulted in a concentration dependent and saturable binding of cholesterol to Pry1CAP protein. *In vitro*, Pry1CAP displayed saturation binding kinetics with an apparent K_d of $2.08 \pm 0.07 \mu\text{M}$, which is comparable to cholesterol binding by full-length Pry1 (K_d of $1.25 \pm 0.42 \mu\text{M}$) (Fig. 2C). The saturable binding observed by titration of the ligand is not due to limited solubility of cholesterol, as indicated by the control experiment in which the *in vitro* affinity of Pry1 to cholesterol was measured by increasing the concentrations of the purified protein (0–500 pmol) rather than increasing the concentration of the radioligand. This results in a saturable binding curve with an apparent K_d of 0.87 ± 0.18 , which is very similar to the K_d obtained by titrating the ligand (K_d of $1.25 \pm 0.42 \mu\text{M}$). Furthermore, cholesterol binding by Pry1CAP is inhibited by EDTA and adding magnesium ions restores sterol binding, indicating that magnesium is important for sterol binding by Pry1CAP (Fig. 2D).

Overall Structure of Pry1CAP. Two Pry1CAP structures solved by molecular replacement with a monomer in the asymmetric unit are reported and their atomic coordinate and structure factors have been deposited in the protein data bank under accession number 5ETE and 5JYS, for the 1,4-dioxane and Mg^{2+} complex, respectively. Both Pry1CAP structures are very similar with an rmsd of all atoms of 0.318 Å. In both structures amino-terminal amino acids residues Ser151 through Ser157, which contains two predicted O-glycosylated sites, are disordered. The overall monomer surface area for Pry1CAP is 7125 \AA^2 and its topology is an alpha-beta-alpha sandwich made up of 6 helices, 2 disulfide bonds, 12 beta turns and 1 beta bulge (Fig. 3B,C). The overall structure of Pry1CAP is 16.9% strand, 33.1% alpha-helix, 2.1% 3–10 helix and 47.9% loop with longer loops than predicted

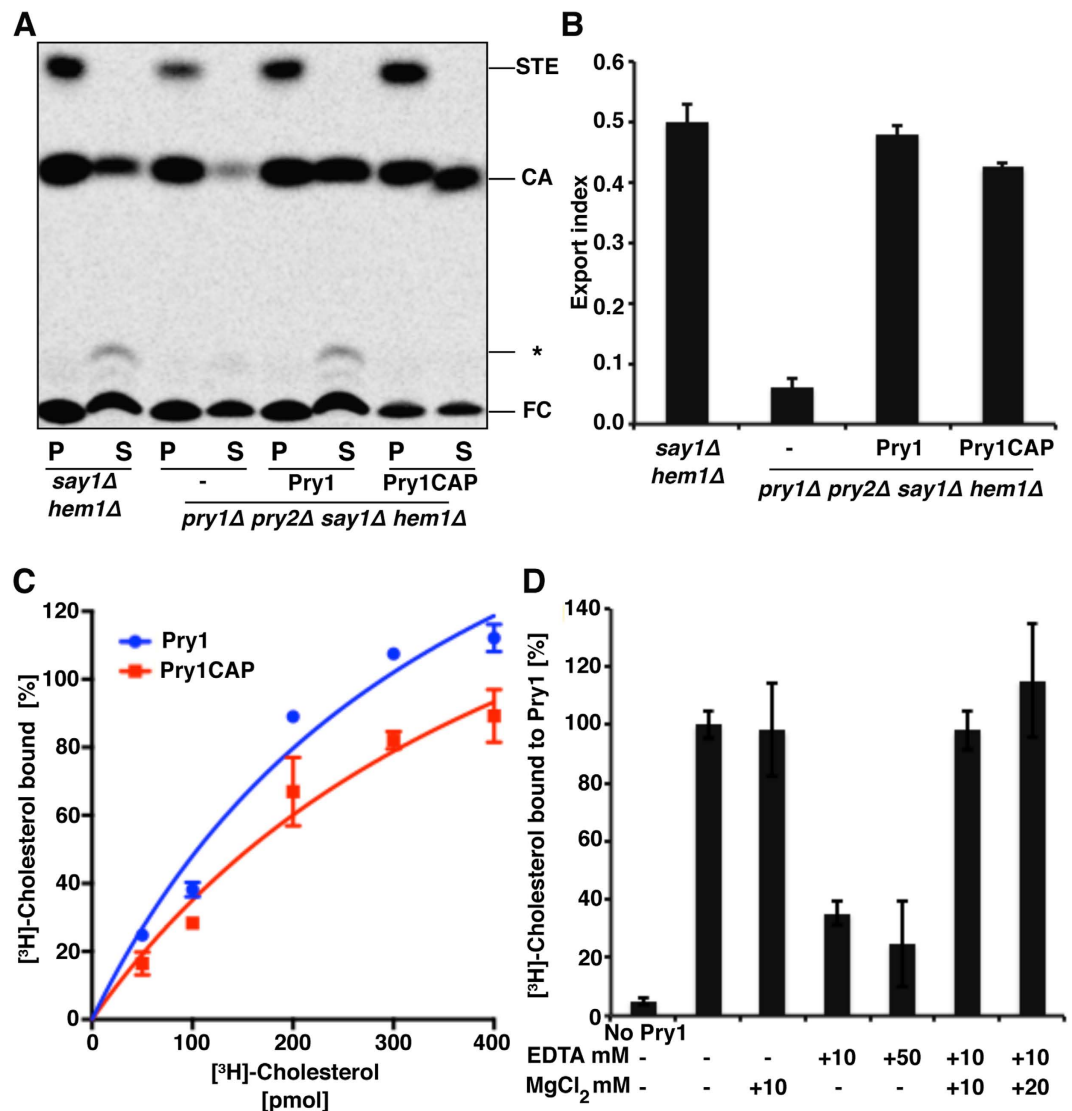


Figure 2. The CAP domain of Pry1 rescues the defect in cholesterol acetate export of the yeast *pry1Δ pry2Δ* double mutant *in vivo* and binds cholesterol *in vitro*. (A) Expression of Pry1CAP complements the sterol-export defect of yeast cells lacking their endogenous CAP proteins. Levels of lipids from Say1 and Heme deficient cells (*say1Δ hem1Δ*) containing either an empty plasmid (–) or a plasmid with Pry1 or Pry1CAP in the cell pellet (P) and the culture supernatant (S) after separation by thin-layer chromatography. The positions of free cholesterol (FC), cholesteryl acetate (CA), steryl esters (STE), and other cholesterol derivatives (*) are indicated. (B) Quantification of the levels of exported CA from yeast cells lacking their endogenous CAP proteins compared with levels from cells complemented with Pry1 or Pry1CAP. (C) Pry1CAP binds cholesterol *in vitro* similarly to full-length Pry1. (D) Sterol binding is Mg²⁺-dependent.

based on other CAP structures (Fig. 4A). The 3–10 helix is involved in the alpha-beta-alpha sandwich and was previously observed in SmVAL4 but not in any of the other representative CAP protein structures.

Divalent cation binding. Pry1CAP has a large central (1638 Å³ volume) cavity, which contains the tetrad that binds divalent cation other CAP protein^{11,14,19,21,23,29}, and this cavity is distinct from the CBM cavity (Figs 3C and 4A). The CAP cavity tetrad residues of Pry1CAP (His208, Glu215, Glu233 and His247) superpose well with the corresponding residues from representative CAP proteins^{8,11,13,29} (Fig. 4B). In one Pry1CAP structure, a magnesium ion from the crystallization solution is coordinated by the tetrad which results in a slight shift in the histidines when compared to the structure without the divalent cation (Fig. 4C).

Discussion

1,4-Dioxane binds to the cholesterol binding site. Electron density for four 1,4-dioxane molecules were observed in the initial molecular replacement $2F_o - F_c$ maps when dioxane was present in the crystallization solution; this density is absent in the Mg²⁺ structure, which does not have 1,4-dioxane in the crystallization

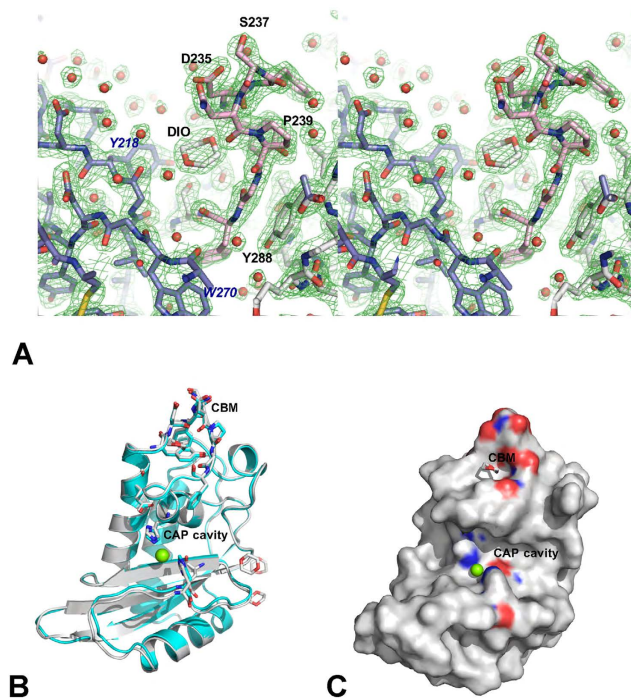


Figure 3. The Pry1CAP structure. (A) Stereoview of the 1,4-dioxane molecule in a $2F_o - F_c$ electron-density map (green) contoured at 1.5σ bound in the caveolin-binding motif (CBM) cavity of Pry1CAP. Main chain residues are colored gray, the CBM residues are shown in pink stick, residues from a symmetry related monomer are colored in blue, and water molecules are shown as red spheres. (B) Superposed structures of the complexes Pry1CAP with Mg^{2+} ion (green) and 1,4-dioxane (gray). (C) Location of the CBM and CAP tetrad on Pry1CAP monomer surface representation.

solution. Density corresponding to one 1,4-dioxane molecule was observed in the 194 \AA^3 volume cavity created by the CBM, confirming for the first time the ability of the cholesterol binding site of Pry1 to bind a ligand.

Since a 1,4-dioxane molecule binds to the CBM cavity, and the CBM is verified to be important for cholesterol binding²⁶, the effect of 1,4-dioxane on cholesterol binding was tested. 1,4-dioxane inhibits *in vitro* binding of [³H]-cholesterol to Pry1CAP in a dose dependent manner (Figure S.1). However, the inhibition was only observed at relatively high concentrations of 1,4-dioxane, and is possibly due to the increased hydrophobicity of the 1,4-dioxane containing solvent compared to the aqueous buffer. Interestingly, 1,4-dioxane does not bind to other lipid binding sites notably the palmitate binding site of Tablysin-15¹⁹ or phosphatidylinositol binding sites of GAPR-1^{23,24}. The presence of dioxane in the crystallization solution was incompatible with the formation of complexes of Pry1CAP with cholesterol or palmitate and all co-crystallization and soaking experiments yielded crystals with dioxane only. While cholesterol is virtually insoluble in water, it is soluble in 1,4-dioxane, and 1,4-dioxane is a major component of the CryoSol kit that is used for co-crystallizing proteins with hydrophobic ligands^{30–32}. Efforts are underway to identify either a sterol solubilizing agent that is suitable for co-crystallization or a crystallization condition that is compatible with sterol binding.

Comparison of Pry1CAP with other CAP proteins. Using PDBFold, GAPR-1 with bound inositol hexakisphosphate (PDB entry 4aiw)²³, and the apo structure of the same protein (PDB entry 1smb)¹⁴ were identified as the most similar structures to that of Pry1CAP. The second best score was that of VAL4 from *Schistosoma mansoni* (SmVAL4), which lacks the prototypical CAP cavity (PDB entry 4p27)²⁷. This is followed by the NMR structure of a plant CAP protein (P14a, PDB entry 1cfe)¹⁵; crystal structures of the hookworm CAP protein *Na-ASP-2*, (PDB entries 4ly5, 4nuj, 4nuo, 4nuk)^{13,22}; the structures of human glioma pathogenesis related protein (sGLIPR1 PDB entry 3q2r)¹³; and structures of snake venom CRISPs notably pseudochetoxin (PDB entry 2dda)²⁹. All these proteins share under 35% sequence identity with Pry1CAP and the greatest differences between representative CAP protein structures are in loop regions as well as in the length of helices and strands (Fig. 4A). The flexible regions in CAP proteins are important for ligand binding and make up ~47% of the structure. Therefore, each new CAP protein structure offers information that cannot be generated simply by homology modeling.

The CAP tetrad and CBM cavities are connected in the dimer. The observation that EDTA inhibits cholesterol binding and that addition of Mg^{2+} restores cholesterol binding by Pry1 is surprising since the cholesterol binding CBM and the CAP tetrad that binds Mg^{2+} are located at distinct sites in the Pry1CAP monomer. Furthermore, our previous studies identified SmVAL4, a CAP protein lacking the CAP tetrad as an effective sterol binder and exporter²⁷. Mg^{2+} dependent sterol binding by Pry1CAP requires the interaction of both binding sites, likely occurring *via* dimerization, because Pry1CAP forms dimers in solution, and both crystal structures have

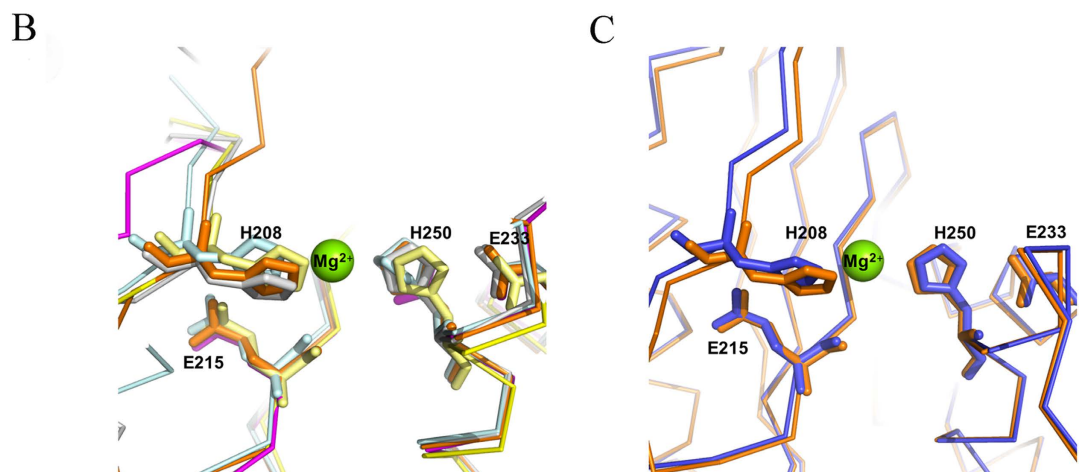
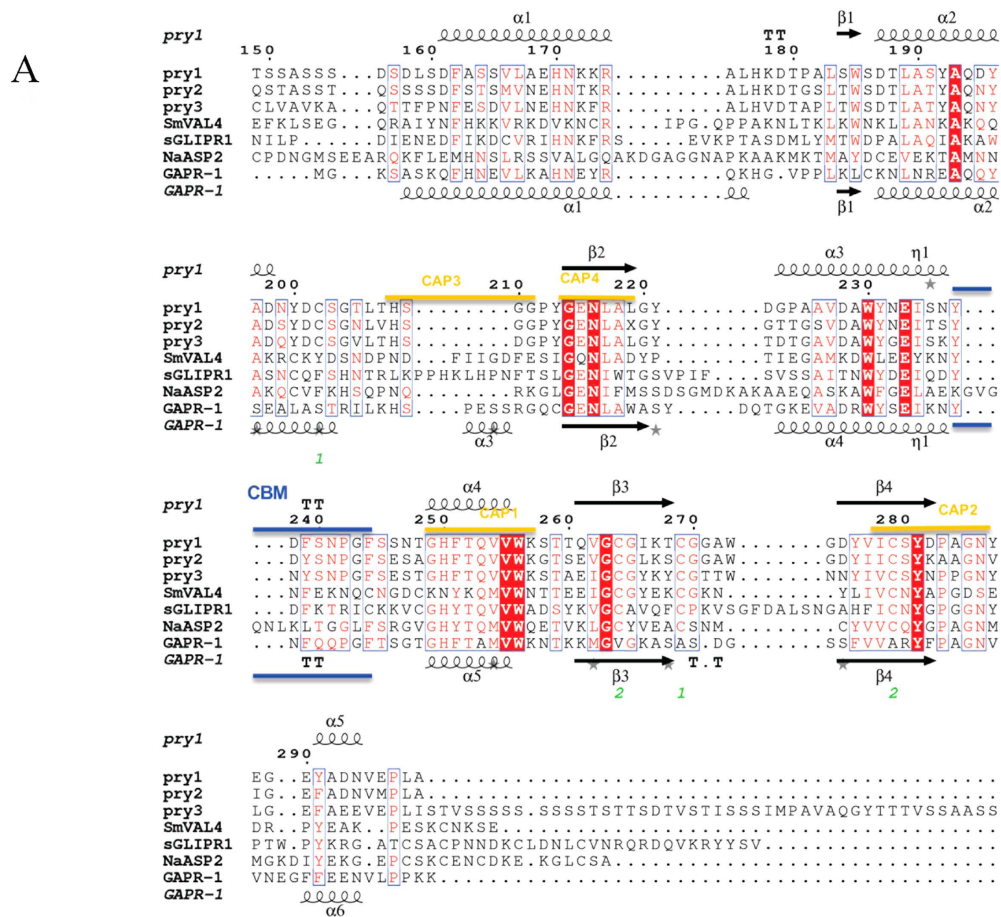


Figure 4. CAP protein motifs. (A) Structural features of Pry1CAP and primary sequence alignment with selected representative CAP proteins. This figure was generated with ESPrift⁴⁸. The different secondary structure elements shown are alpha helices as large squiggles labelled (α), 3_{10} -helices as small squiggles labelled (η), beta strands as arrows (β), and beta turns (TT). Identical residues are shown in white on red background, and conserved residues in red. The locations of the cysteine residues involved in disulfide bonds are numbered in green. CAP motifs are highlighted in gold. The representative CAP structures are *Na-ASP-2* (PDB entry 1u53), *SmVAL4* (PDB entry 4p27), *GAPR-1* (PDB entry 1smb), and *sGLIPR1* (PDB entry 3q2r). (B) The CAP tetrad residues of Pry1CAP (PDB entry 5jys) superimpose with other CAP structures. CAP structures are colored as follows Pry1CAP (orange), *SmVAL4* (magenta), *GAPR-1* (cyan), *sGLIPR1* (gray) and *Na-ASP-2* (yellow). (C) The superposed divalent cation binding site of Pry1CAP with Mg^{2+} (PDB entry 5jys, orange) and without Mg^{2+} (PDB entry 5ete, blue). The numbers correspond to those of Pry1CAP and the magnesium ion is shown as green sphere.

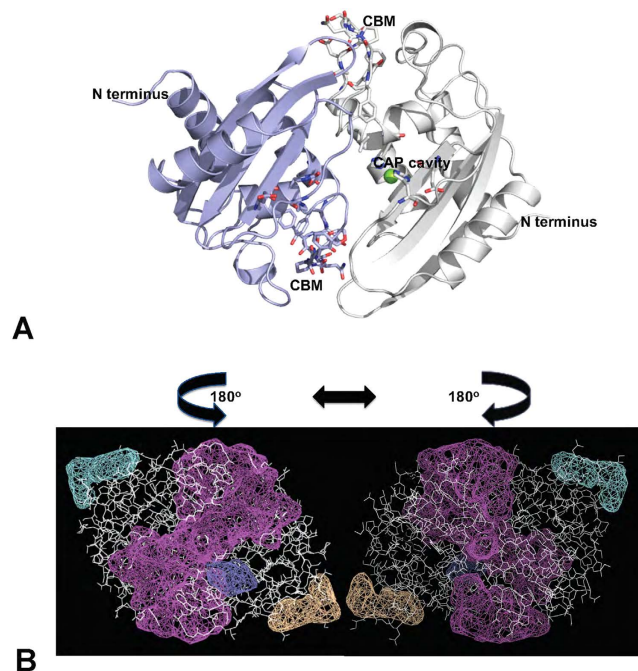


Figure 5. Pry1CAP dimer. (A) Ribbon diagram of the crystallographic dimer with one monomer colored blue and the other gray. (B) A large central cleft (purple) wraps around the protein (white sticks) connecting the CBM and CAP tetrad sites of both monomers, the left panel is the same view as Fig. 5A while the right panel represents a 180° rotation around the 2-fold axis. Additional clefts in the structure are shown as mesh in other colors.

a crystallographic dimer in which the CBM and CAP tetrad are connected within a large 7063 Å³ volume cleft (Fig. 5). The same dimer was observed in two different structures of Pry1CAP, that were obtained from different crystallization conditions with different spacegroups and crystal morphologies. Incidentally, the cavities of the two covalently linked CAP domains of *Na-ASP-1*¹² are connected by a large cleft and prior to this study a dimer that allowed the connection of the CAP tetrad and lipid binding cavities of single CAP domain proteins had not been identified^{8,11,12}. Beyond affecting sterol binding, the implications of having such a large interconnected cleft are undefined; however, the cavity volume is large enough to bind other ligands, such as peptides, which may be relevant for some observed functions of other eukaryotic CAP proteins, for example, the serine protease activity of the cone snail CAP protein *Tex31*³³ and the ability of the hookworm NIF to inhibit neutrophils³⁴.

Conclusions

Two crystal structures of the CAP domain of pathogen-related yeast protein from *Saccharomyces cerevisiae*, Pry1CAP, are presented. Both structures reveal that the cholesterol binding CBM is large enough for cholesterol binding. In one structure a 1,4-dioxane molecule from the crystallization mixture occupies the CBM cavity confirming its ability to bind hydrophobic ligands. In the second structure Mg²⁺ is coordinated by the CAP cavity tetrad residues. Pry1CAP is functional in cholesterol export *in vivo* and binds cholesterol *in vitro* with comparable affinity to full length Pry1. Interestingly, cholesterol binding by Pry1CAP is inhibited by EDTA and restored by the addition of Mg²⁺ indicating that presence of the divalent cation is important for sterol binding. The cholesterol and Mg²⁺ binding sites are distinct and unconnected in the Pry1CAP monomer. Pry1CAP is a dimer in solution and the cholesterol and Mg²⁺ binding sites are connected by a large cleft in a crystallographic dimer, providing a structural basis for Mg²⁺-dependent sterol binding by Pry1CAP.

Methods

Recombinant protein expression and purification of Pry1CAP. The carboxyl terminal CAP domain (corresponding to amino acid residues Ser151 through Ala296) of Pry1 from *Saccharomyces cerevisiae* was amplified by PCR and ligated into pPICZαA vector using the XhoI and XbaI sites. After linearization, the vector was transformed into *Pichia pastoris* strain X33. The transformants were selected on zeocin-resistant YPD plates and verified by PCR amplification using pPICZαA vector flanking primers (α-factor and 3'AOX1). Ten colonies with the right insert were picked and screened for induction of recombinant Pry1CAP protein with 0.5% methanol at 30 °C for 72 hours. The highest expressing colony was chosen for recombinant protein expression and purification as previously described for SmVAL4²⁷. Full length Pry1 was produced similarly to Pry1CAP.

Size-exclusion chromatography (SEC). For SEC experiments, 20 μg of Pry1CAP was injected onto a Yarra SEC-2000 column (Phenomenex, Torrance, CA) at flow-rate of 0.5 ml/min with a Shimadzu Prominence series HPLC (Kyoto, Japan) using PBS pH 7.4 as the mobile phase. The elution was monitored with a photo diode array detector (Shimadzu). The system was calibrated using Bio-Rad gel filtration standard (Hercules, CA) consisting

Data Collection	PDB entry 5ete	PDB entry 5jys
X-ray Source	Rigaku FR-E+	Rigaku FR-E+
Detector	Rigaku HTC	Rigaku HTC
Wavelength	0.15418 nm	0.15418 nm
Space group	$P6_22$	$I4_1$
Cell dimensions	a = b = 124.74 Å, c = 59.06 Å $\alpha = \beta = 90.00^\circ$, $\gamma = 120.00^\circ$	a = b = 69.11 Å, c = 97.73 Å $\alpha = \beta = \gamma = 90.00^\circ$
Resolution (Å)	30.8–2.1 (2.16–2.10)	34.5–1.89 (1.99–1.90)
Number of total reflections	230544 (17841)	65957 (4185)
Number of unique reflections	16255 (1304)	18355 (1200)
$^{\dagger}R_{merge}$	9.7 (62.4)	8.2 (42.8)
$I/\sigma(I)$	17.5 (3.8)	10.4 (2.3)
Completeness (%)	99.8 (100)	100 (100)
† Redundancy	14.2 (13.7)	3.6 (3.5)
Mn(I) half-set correlation CC(1/2)	0.998 (0.963)	0.995 (0.902)
Average Mosaicity	1.8	1.01
Resolution (Å)	30.8–2.01 (2.10–2.01)	34.5–1.79 (1.83–1.79)
Number of total reflections		77208 (4323)
Number of unique reflections	18511 (1324)	21607 (1254)
$^{\dagger}R_{merge}$	0.105 (1.296)	0.073 (0.629)
$I/\sigma(I)$	15.4 (1.7)	9.4 (1.4)
Completeness (%)	99.7 (98.9)	100 (100)
Redundancy	14.1 (13.0)	3.6 (3.4)
Refinement (PHENIX)		
Resolution (Å)	28.46–2.10 (2.23–2.10)	34.5–1.9 (2.0–1.9)
Percentage Data completeness	99.4 (99.4)	100 (100)
R_{free} test set	831 reflections (5.40%)	961 reflections (5.34%)
Wilson B-factor (Å ²)	29.5	27.3
Anisotropy	0.171	0.027
No. of non-H protein atoms	1119	1072
No. of water molecules	128	196
Ions and ligands	4 DIO (1,4-dioxane)	1 Mg ²⁺
$^aR_{work}$	0.167 (0.188)	0.198
$^bR_{free}$	0.190 (0.222)	0.208
Correlation coefficient $F_o - F_c$	0.954 (0.939)	0.94
Average B-factors (Å ²)	31.3	30.2
Protein (Å ²)	29.2	29.6
Water and other small molecules (Å ²)	40.5	42.6
r.m.s. deviations		
Bond lengths (Å)	0.011	0.003
Bond angles (°)	0.935	0.533
MolProbity analysis		
Ramachandran outliers	0%	0
Ramachandran favored	100%	100
Rotamer outliers	3%	0
C-beta deviations	0	0
Clashscore	0.96	0.99

Table 1. Statistics for data collection and model refinement.

of proteins with molecular weights of 670, 158, 44, 17 and 1.35 kDa. Data analysis was performed on the 280 nm wavelength extracted chromatograph using Shimadzu LCsolution version 1.25. Experiments were conducted in triplicate.

Crystallization. Crystallization conditions were identified and optimized after screening for crystals using the following screens, Qiagen Cryos, PEGs and Hampton Crystal Screen I. The largest crystals were obtained at 298 K by vapor diffusion in sitting drops by mixing 4 μ L of protein solution (15 mg/ml in PBS 7.4.) with 1.5 μ L of the reservoir solution (1.6 M ammonium sulfate, 10% (v/v) 1,4-dioxane, and 0.1 M MES pH 6.5) and equilibrating against 350 μ L of reservoir solution. Crystals of dimension 0.2 mm X 0.3 mm X 0.5 mm grew after 3 months. The crystallization process was shortened to 5 days by increasing the protein concentration to 50 mg/ml and the

temperature to 37 °C, without compromising the diffraction quality of the crystals. A second crystal form was obtained with 100 mM Tris pH 8.5 and 200 mM magnesium chloride as the crystallization buffer. These crystals required higher protein concentrations (150 mg/ml Pry1CAP in 0.1 M Tris pH 8.0) and took up to 3 weeks to grow. No crystallization hits were observed when Pry1CAP was screened in 50 mM sodium acetate pH 4.5, 50 mM sodium HEPES pH 7.5, or sodium citrate pH 6.5, whereas weakly diffracting plate-like crystals were obtained with 50 mM sodium cacodylate as protein storage buffer.

Data collection, structure determination and analysis. Crystals were transferred into a cryo-protecting solution containing 75% precipitant solution and 25% glycerol for ~30 seconds, and flash-cooled directly in a stream of N₂ gas at 113 K prior to collecting diffraction data. X-ray diffraction data were collected at the Baylor College of Medicine core facility (Rigaku HTC detector, Rigaku FR-E+ SuperBright microfocussing anode generator, with VariMax HF optics) using the Crystal Clear (d*trek) package³⁵. Data was integrated using MosFLM and scaled with SCALA³⁶. Pry1CAP structures were solved after multiple molecular replacement attempts (MR) using different search models^{11–13} with PHASER^{37,38}. The optimal model that resulted in the lowest R_{free} was a polyalanine model of GAPR-1 (PDB entry 1smb)¹⁴. The MR solution indicated a monomer per asymmetric unit giving a Matthews' coefficient of 4.3 Å³/Da with 72% solvent content³⁹. Molecular replacement was followed by automatic model building using ARP/wARP^{40,41}. The final models were obtained by iterative manual model building cycles using the program Coot⁴² followed by structure refinement with REFMAC5^{43,44} and PHENIX⁴⁵. Structural figures were generated using PyMOL⁴⁶. Structures were analyzed and Fig. 5B was generated using PDBSumm (<https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html>). Details of the quality of the structure as well as data collection are shown in Table 1. Protein structures most similar to Pry1CAP were identified using PDBFold's structure similarity option (<http://www.ebi.ac.uk/msd-srv/ssm/>), which allows a 3-D structural alignment taking both the alignment length and rmsd into account.

In vitro lipid binding. The radioligand binding assay was performed as described previously^{25,47}. 100 pmol of purified protein in binding buffer (20 mM Tris, pH 7.5, 30 mM NaCl, 0.05% Triton X-100) was incubated with 0–400 pmol of [³H]-cholesterol (American Radiolabeled Chemicals Inc., St Louis, Missouri, USA) for 1 h at 30 °C. The protein was then separated from the unbound ligand by adsorption to Q-sepharose beads (GE healthcare, USA), beads were washed, and the radioligand was quantified by scintillation counting. The effect of 1,4-dioxane was determined by performing the *in vitro* assay in the presence of 1,4-dioxane (0–4% v/v). The effect of divalent cations on cholesterol binding was measured by performing the *in vitro* binding reaction in the presence of different concentrations of EDTA and magnesium chloride. At least two independent experiments were performed under each experimental condition and data is reported as the mean ± standard deviation. Calculation of the K_d value and curve fitting was performed using the statistical software GraphPad Prism, La Jolla, CA.

In vivo yeast sterol export assay. Acetylation and export of sterols into the culture supernatant was examined as described elsewhere²⁶. In this sterol export assay, the export of radiolabeled cholesterol acetate by *hem1Δ say1Δ* deficient *Saccharomyces cerevisiae* cells upon complementation with plasmid containing genes of interest is monitored. Yeast cells (*say1Δ hem1Δ*) were cultivated in the presence of cholesterol/Tween 80 and were labeled with 0.025 μCi/ml [¹⁴C]-cholesterol (American Radiolabeled Chemicals Inc., St Louis, Missouri, USA). Cells were harvested by centrifugation, washed twice with synthetic complete (SC) media, and grown overnight in fresh medium containing non-radiolabeled cholesterol. Cells were centrifuged and lipids were extracted from the cell pellet and the culture supernatant using chloroform/methanol [1:1, (v/v)]. Samples were dried and separated by thin-layer chromatography on silica gel 60 plates (TLC; Merck, Darmstadt, Germany) using the solvent system petroleum ether/diethyl ether/acetic acid [70:30:2, (v/v)]. TLCs were exposed to phosphorimager screens and radiolabeled lipids were visualized and quantified using a phosphorimager (Bio-Rad Laboratories, Hercules, California, USA). The export index is the relative levels of CA that is exported by the cells, and calculated as the ratio of extracellular CA to the sum of intracellular and extracellular CA. Export experiments were performed in triplicate and reported as the mean ± standard deviation of three independent experiments.

References

- Ding, X., Shields, J., Allen, R. & Hussey, R. S. Molecular cloning and characterisation of a venom allergen AG5-like cDNA from *Meloidogyne incognita*. *Int J Parasitol* **30**, 77–81 (2000).
- Hawdon, J. M., Narasimhan, S. & Hotez, P. J. Ancylostoma secreted protein 2: cloning and characterization of a second member of a family of nematode secreted proteins from *Ancylostoma caninum*. *Mol Biochem Parasitol* **99**, 149–165 (1999).
- Zhan, B. *et al.* Molecular characterisation of the Ancylostoma-secreted protein family from the adult stage of *Ancylostoma caninum*. *Int J Parasitol* **33**, 897–907 (2003).
- Gao, B. *et al.* Molecular characterisation and expression of two venom allergen-like protein genes in *Heterodera glycines*. *Int J Parasitol* **31**, 1617–1625 (2001).
- Wilhelm, S. Parasitism and pathogenesis of root-disease fungi. *Plant pathology: problems and progress* **1958**, 356–366 (1908).
- Schneider, R. & Di Pietro, A. The CAP protein superfamily: function in sterol export and fungal virulence. *Biomolecular concepts* **4**, 519–525, doi: 10.1515/bmc-2013-0021 (2013).
- Gibbs, G. M. *et al.* The cysteine-rich secretory protein domain of Tpx-1 is related to ion channel toxins and regulates ryanodine receptor Ca²⁺ signaling. *J Biol Chem* **281**, 4156–4163, doi: M506849200/jbc.M506849200 (2006).
- Gibbs, G. M., Roelants, K. & O'Bryan, M. K. The CAP superfamily: cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins—roles in reproduction, cancer, and immune defense. *Endocr Rev* **29**, 865–897, doi: er.2008-0032/er.2008-0032 (2008).
- Gibbs, G. M. & O'Bryan, M. K. Cysteine rich secretory proteins in reproduction and venom. *Soc Reprod Fertil Suppl* **65**, 261–267 (2007).

10. Kasahara, M., Gutknecht, J., Brew, K., Spurr, N. & Goodfellow, P. N. Cloning and mapping of a testis-specific gene with sequence similarity to a sperm-coating glycoprotein gene. *Genomics* **5**, 527–534 (1989).
11. Asojo, O. A. *et al.* X-ray structure of Na-ASP-2, a pathogenesis-related-1 protein from the nematode parasite, *Necator americanus*, and a vaccine antigen for human hookworm infection. *J Mol Biol* **346**, 801–814, doi: 10.1016/j.jmb.2004.12.023 (2005).
12. Asojo, O. A. Structure of a two-CAP-domain protein from the human hookworm parasite *Necator americanus*. *Acta crystallographica. Section D, Biological crystallography* **67**, 455–462, doi: 10.1107/S0907444911008560 (2011).
13. Asojo, O. A., Koski, R. A. & Bonafe, N. Structural studies of human glioma pathogenesis-related protein 1. *Acta crystallographica. Section D, Biological crystallography* **67**, 847–855, doi: 10.1107/S0907444911028198 (2011).
14. Serrano, R. L. *et al.* Structural analysis of the human Golgi-associated plant pathogenesis related protein GAPR-1 implicates dimerization as a regulatory mechanism. *J Mol Biol* **339**, 173–183, doi: 10.1016/j.jmb.2004.03.015 (2004).
15. Fernandez, C. *et al.* NMR solution structure of the pathogenesis-related protein P14a. *J Mol Biol* **266**, 576–593, doi: 10.1006/jmbi.1996.0772 (1997).
16. Wang, J. *et al.* Blocking effect and crystal structure of natrin toxin, a cysteine-rich secretory protein from *Naja atra* venom that targets the BKCa channel. *Biochemistry* **44**, 10145–10152, doi: 10.1021/bi050614m (2005).
17. Shikamoto, Y., Suto, K., Yamazaki, Y., Morita, T. & Mizuno, H. Crystal structure of a CRISP family Ca²⁺-channel blocker derived from snake venom. *J Mol Biol* **350**, 735–743, doi: S0022-2836(05)00551-6/j.jmb.2005.05.020 (2005).
18. Guo, M. *et al.* Crystal structure of the cysteine-rich secretory protein stecrip reveals that the cysteine-rich domain has a K⁺ channel inhibitor-like fold. *J Biol Chem* **280**, 12405–12412, doi: M413566200/jbc.M413566200 (2005).
19. Xu, X., Francischetti, I. M., Lai, R., Ribeiro, J. M. & Andersen, J. F. Structure of protein having inhibitory disintegrin and leukotriene scavenging functions contained in single domain. *J Biol Chem* **287**, 10967–10976, doi: 10.1074/jbc.M112.340471 (2012).
20. Borloo, J. *et al.* Structure of *Ostertagia ostertagi* ASP-1: insights into disulfide-mediated cyclization and dimerization. *Acta crystallographica. Section D, Biological crystallography* **69**, 493–503, doi: 10.1107/S0907444912050019 (2013).
21. Wang, Y. L. *et al.* Cobra CRISP Functions as an Inflammatory Modulator via a Novel Zn²⁺- and Heparan Sulfate-dependent Transcriptional Regulation of Endothelial Cell Adhesion Molecules. *J Biol Chem* **285**, 37872–37883, doi: M110.146290/jbc.M110.146290 (2010).
22. Mason, L. *et al.* Probing the equatorial groove of the hookworm protein and vaccine candidate antigen, Na-ASP-2. *The international journal of biochemistry & cell biology* **50**, 146–155, doi: 10.1016/j.biocel.2014.03.003 (2014).
23. van Galen, J. *et al.* Interaction of GAPR-1 with lipid bilayers is regulated by alternative homodimerization. *Biochim Biophys Acta* **1818**, 2175–2183, doi: 10.1016/j.bbame.2012.04.016 (2012).
24. Van Galen, J. *et al.* Binding of GAPR-1 to negatively charged phospholipid membranes: unusual binding characteristics to phosphatidylinositol. *Mol Membr Biol* **27**, 81–91, doi: 10.3109/09687680903507080 (2010).
25. Choudhary, V. & Schneider, R. Pathogen-Related Yeast (PRY) proteins and members of the CAP superfamily are secreted sterol-binding proteins. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 16882–16887, doi: 10.1073/pnas.1209086109 (2012).
26. Choudhary, V. *et al.* The caveolin-binding motif of the pathogen-related yeast protein Pry1, a member of the CAP protein superfamily, is required for *in vivo* export of cholesteryl acetate. *Journal of lipid research* **55**, 883–894, doi: 10.1194/jlr.M047126 (2014).
27. Kelleher, A. *et al.* *Schistosoma mansoni* venom allergen-like protein 4 (SmVAL4) is a novel lipid-binding SCP/TAPS protein that lacks the prototypical CAP motifs. *Acta crystallographica. Section D, Biological crystallography* **70**, 2186–2196, doi: 10.1107/S1399004714013315 (2014).
28. Asojo, O. A. *et al.* Crystallization and preliminary X-ray analysis of Na-ASP-1, a multi-domain pathogenesis-related-1 protein from the human hookworm parasite *Necator americanus*. *Acta Crystallograph Sect F Struct Biol Cryst Commun* **61**, 391–394 (2005).
29. Suzuki, N. *et al.* Structures of pseudochetoxin and pseudocin, two snake-venom cysteine-rich secretory proteins that target cyclic nucleotide-gated ion channels: implications for movement of the C-terminal cysteine-rich domain. *Acta crystallographica. Section D, Biological crystallography* **64**, 1034–1042, doi: /S0907444908023512 (2008).
30. Polsinelli, I. *et al.* A new crystal form of human transthyretin obtained with a curcumin derived ligand. *J Struct Biol* **194**, 8–17, doi: 10.1016/j.jsb.2016.01.007 (2016).
31. Ciccone, L. *et al.* X-ray crystal structure and activity of fluorenyl-based compounds as transthyretin fibrillogenesis inhibitors. *J Enzyme Inhib Med Chem*, 1–10 (2015).
32. Ciccone, L., Tepshi, L., Nencetti, S. & Stura, E. A. Transthyretin complexes with curcumin and bromo-estradiol: evaluation of solubilizing multicomponent mixtures. *N Biotechnol* **32**, 54–64, doi: 10.1016/j.nbt.2014.09.002 (2015).
33. Milne, T. J., Abbenante, G., Tyndall, J. D., Halliday, J. & Lewis, R. J. Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily. *J Biol Chem* **278**, 31105–31110, doi: 10.1074/jbc.M304843200 (2003).
34. Rieu, P., Ueda, T., Haruta, I., Sharma, C. P. & Arnaout, M. A. The A-domain of beta 2 integrin CR3 (CD11b/CD18) is a receptor for the hookworm-derived neutrophil adhesion inhibitor NIF. *J Cell Biol* **127**, 2081–2091 (1994).
35. Pflugrath, J. W. The finer things in X-ray diffraction data collection. *Acta crystallographica. Section D, Biological crystallography* **55** (Pt 10), 1718–1725 (1999).
36. Leslie, A. G. The integration of macromolecular diffraction data. *Acta crystallographica. Section D, Biological crystallography* **62**, 48–57 (2006).
37. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta crystallographica. Section D, Biological crystallography* **61**, 458–464 (2005).
38. Storoni, L. C., McCoy, A. J. & Read, R. J. Likelihood-enhanced fast rotation functions. *Acta crystallographica. Section D, Biological crystallography* **60**, 432–438 (2004).
39. Matthews, B. W. Solvent content of protein crystals. *J Mol Biol* **33**, 491–497 (1968).
40. Morris, R. J. *et al.* Breaking good resolutions with ARP/wARP. *J Synchrotron Radiat* **11**, 56–59 (2004).
41. Morris, R. J., Perrakis, A. & Lamzin, V. S. ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol* **374**, 229–244 (2003).
42. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486–501, doi: 10.1107/S0907444910007493 (2010).
43. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta crystallographica. Section D, Biological crystallography* **67**, 355–367, doi: 10.1107/S0907444911001314 (2011).
44. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**, 235–242, doi: 10.1107/S0907444910045749 (2011).
45. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213–221, doi: 10.1107/S0907444909052925 (2010).
46. Delano, W. L. *The PyMOL Molecular Graphics System*. DeLano Scientific LLC Palo Alto, California, USA. URL <http://www.pymol.org> (2002).
47. Im, Y. J., Raychaudhuri, S., Prinz, W. A. & Hurley, J. H. Structural mechanism for sterol sensing and transport by OSBP-related proteins. *Nature* **437**, 154–158, doi: 10.1038/nature03923 (2005).
48. Gouet, P., Robert, X. & Courcelle, E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* **31**, 3320–3323 (2003).

Acknowledgements

RS thanks the Swiss National Science Foundation for financial support [31003A_153416] OAA thanks the National School of Tropical Medicine at BCM for startup funds. Thanks to Gonteria Robinson, Auja Smith, Ebe Ewere and Zuriel Hodges for performing preliminary crystallization experiments and Nate Wolf for proof-reading the manuscript. Thanks to Drs. Peter Hotez, Maria-Elena Bottazzi, and Bin Zhan for continued support of our CAP protein studies.

Author Contributions

O.A.A. and R.S. conceived the studies and interpreted the results, A.K. produced and crystallized Pry1CAP, R.D. performed *in vivo* sterol export and *in vitro* binding studies, E.M.H. performed SEC analysis, O.A.A. collected crystallographic data, and solved crystal structure. All authors contributed to the writing and editing of the final manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Darwiche, R. *et al.* Structural and functional characterization of the CAP domain of pathogen-related yeast 1 (Pry1) protein. *Sci. Rep.* **6**, 28838; doi: 10.1038/srep28838 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>