Structural insights into bacterial sterol transport

Liting Zhai,[a] Amber C. Bonds,[b] Hannah Oo,[a] Paula V. Welander,[b,1] Laura M. K. Dassama[a,1]

[a]Department of Chemistry and Sarafan ChEM-H, Stanford University

[b]Department of Earth System Science, Stanford University

[1]To whom corresponding should be addressed: welander@stanford.edu, dassama@stanford.edu

Author contributions: L. Z., A. C. B., P. V, R., and L. M. K. D. designed research, L. Z., A. C. B., and H. O. performed research, L. Z., A. C. B., H. O., P. V, R., and L. M. K. D. analyzed data, L. Z., A. C. B., P. V, R., and L. M. K. D. wrote the manuscript.

Data deposition: The coordinates of structures reported herein have been deposited in the Protein Data Bank (7T1M and 7T1S).

Abstract

Sterol lipids are widely present in eukaryotes and play essential roles in signaling and modulating membrane fluidity. Although rare, some bacteria species produce sterols, but their function in the bacterial domain remains unknown. The aerobic methanotroph *Methylococcus capsulatus* was the first bacterium shown to synthesize sterols, producing a mixture of C-4 methylated sterols that are distinct from those observed in eukaryotes. Subsequent studies demonstrated that C-4 methylated sterols synthesized in the cytosol are localized to the outer membrane, suggesting that a bacterial sterol transport machinery exists. While proteins involved in eukaryotic sterol transport have been identified and characterized, their counterparts in bacteria remain enigmatic. In this study, we used bioinformatics, ligand binding analysis, and structural approaches to identify three novel bacterial sterol transporters from *Methylococcus capsulatus*. These proteins reside in the inner membrane, periplasm, and outer member of the bacterium, thereby working as a conduit to move modified sterols to the outer membrane. We report their remarkable specificity for recognizing 4-methylsterols, reveal the structural bases for substrate binding, and describe their structural divergence from eukaryotic sterol transporters. These findings provide unequivocal evidence for a distinct sterol transport system within the bacterial domain, which provide insights into our understanding of bacterial sterols and their divergence from similar lipids in higher-order organisms.

Significance

Sterol lipids, such as cholesterol, were once thought to be unique to eukaryotes. However, a few bacterial species also produce sterols, albeit with structural modifications not observed in eukaryotes. While little is known about the function of bacterial sterols, it is known that sterols in certain bacteria are localized to the outer membrane. However, how bacteria transport these lipids from the cytoplasm to the outer membrane remains unknown. Here, we identify three novel bacterial sterol transporter proteins in the bacterium *Methylococcus capsulatus* that bind sterols methylated at the C-4 position. Through structural characterization of these proteins, we reveal insights into the mechanisms of sterol transport in bacteria with implications for understanding how bacterial cells interact with sterol lipids more broadly.

Introduction

Sterol lipids are ubiquitous and essential components of eukaryotic life, playing vital roles in intra- and intercellular signaling, stress tolerance, and maintaining cell membrane integrity (1–4). Although sterol synthesis is often considered to be a strictly eukaryotic feature, several bacteria have been shown to also produce sterols (1, 5). Bacterial sterols were first discovered in *Methylococcus capsulatus* more than 40 years ago (6) and, initially, were only observed in a few isolated aerobic methanotrophs (γ-Proteobacteria) and a few myxobacteria (δ-Proteobacteria) (7–10). Subsequent comparative genomics analyses have revealed the potential to produce sterol in a variety of bacterial groups including Planctomycetes, α-Proteobacteria, and Bacteroidetes (9, 11, 12). Sterols produced by bacteria, however, tend to differ from eukaryotic sterols in both structure and in their synthesis pathways.

Sterol synthesis in both bacteria and eukaryotes requires the cyclization of the linear substrate oxidosqualene by an oxidosqualene cyclase (Osc) to generate either lanosterol or cycloartenol (13). However, fungi, vertebrates, and plants further modify these initial cyclization products to synthesize ergosterol, cholesterol, and stigmasterol, respectively (14). These biochemical transformations include demethylations, isomerizations, saturations, and unsaturations and are essential for sterols to function properly in eukaryotes (15, 16). Although bacteria have not been shown to produce fully modified sterols such as cholesterol, some bacteria, including several aerobic methanotrophs and myxobacteria, do modify their sterols. For example, *M. capsulatus* produces sterols that are demethylated once at the C-4 and C-14 positions and with a unique desaturation between C-8 and C-14 (Fig.1A) (17). In addition, studies have revealed that bacterial proteins required to modify sterols can differ from the canonical eukaryotic sterol modifying proteins. We recently identified two sterol demethylase proteins, SdmA and SdmB, in *M. capsulatus* that utilize O_2 to remove one methyl group at the C-4 position that are mechanistically distinct from the eukaryotic C-4 demethylase enzymes (18–20).

Although we have a better grasp of the taxonomic distribution of sterol synthesis in bacteria and of the distinct bacterial proteins involved in their biosynthesis, the function of these lipids in bacteria remains a mystery. It has been posited that bacterial sterols

play a role in modulating the fluidic properties of the cytoplasmic membrane, similar to what is observed for cholesterol and ergosterol in eukaryotic cells(4, 21, 22). However, studies in *M. capsulatus* have demonstrated that the distribution of sterols differs considerably from what occurs in eukaryotes. Approximately 90% of the free sterol pool in eukaryotic cells is found in the cytoplasmic membrane (23) while in *M. capsulatus*, 75% of sterols are localized to the outer membrane (24). The outer membrane is the external asymmetric membrane of Gram-negative bacteria and differs from the cytoplasmic membrane in that its outer leaflet is composed of lipopolysaccharides (LPS) in addition to phospholipids (25). Further, sterols that differ by only a single methylation in the core structure can produce significantly different effects in terms of membrane fluidity and stability (26). Given that *M. capsulatus* primarily produces sterols with one or two C-4 methyl groups, any sterol-membrane interactions in *M. capsulatus* could differ significantly from what is observed in eukaryotic membranes with cholesterol and ergosterol, which lack methyl groups at the C-4 position. Interestingly, the 4,4-dimethylsterols and 4-methylsterols found in *M. capsulatus* also exist as precursors of pathway end-products in eukaryotes and accumulation of C-4 methylated sterols has been associated with a variety of eukaryotic process. The accumulation of C-4 intermediates is considered to be the cause of human genetic diseases known as sterolosis (27–29) and 4,4-dimethylsterols such as lanosterol in the brain are implicated in Parkinson's disease (30). In a variety of yeast species, 4-methylsterol intermediates can regulate cellular processes associated with hypoxia and other conditions of cellular stress (31–33). Although the function of 4,4-dimethylsterols and 4-methylsterols in bacteria is still elusive, a functional role for these specific sterols beyond maintaining membrane fluidity and integrity seems plausible.

To better understand the functional significance of sterol production in the bacterial domain, a fuller understanding of the molecular mechanisms controlling sterol production and localization are needed. Given the observed distribution of sterols in the outer membrane of *M. capsulatus*, we hypothesized that transporters specific for C-4 methylated sterols must exist that can shuttle these substrates to the outer membrane. Trafficking of sterols to various organelles in eukaryotic cells is not fully understood (23) and impaired sterol transport has been connected with a variety of defects including

lysosomal storage diseases (34). Identifying and characterizing bacterial proteins that transport sterol could provide insight into novel sterol binding motifs and folds and could potentially provide insight into lipid trafficking and protein-sterol interaction in eukaryotes.

In this study, we reveal three new sterol transport proteins in *M. capsulatus* that exhibit remarkable specificity for 4-methylsterols over 4,4-dimethylsterols. We used bioinformatics to first identify a cytoplasmic membrane protein (MCA1013), periplasmic protein (MCA1014), and outer membrane associated protein (MCA1015). Their ability to recognize and bind sterols was confirmed using protein-lipid pull down assays, where they showed selective binding to C-4 methylated sterols when in the presence of total lipid extracts from *M. capsulatus.* Microscale thermophoresis (MST) confirmed these interactions and provided quantitative analyses of equilibrium dissociation constants. High-resolution crystallographic structures of MCA1014 and MCA1015 evince their pronounced divergence from eukaryotic sterol transporters. The structures coupled with docking studies revealed putative substrate binding sites and recognition mechanisms. Collectively, these data provide evidence for a novel sterol transport system within the bacterial domain and advances our understanding of bacterial sterol lipid physiology.

Results

Bioinformatic identification of sterol transport proteins in *M. capsulatus*

In a previous study, we used the Joint Genome Institute (JGI) Integrated Microbial Genomes (IMG) Phylogenetic Profiler to identify seventeen candidate genes in *M. capsulatus* that are unique to C-4 demethylating methanotrophs (18). Among them, a putative operon that contains five genes was found to localize next to three sterol biosynthesis genes (Fig. 1B). Two of them (*sdmA/sdmB*) have been characterized to be demethylases that are involved in sterol demethylation at the C-4 position. Homologs of the other three hypothetical genes, *MCA1013, MCA1014* and *MCA1015,* are found in all other aerobic methanotroph genomes that contain *sdmA/sdmB*. In some of these species, these three genes are also seen to be adjacent to other sterol synthesis genes such as

the oxidosqualene cyclase (*osc*) and squalene monooxygenase (*smo*), implying they may be functionally related to sterol physiology.

To better understand the function of these three unknown proteins, we performed more refined bioinformatic analyses, including protein sequence similarity networks (SSN) (35) and structure prediction with Phyre2 (36) and I-TASSER (37). The SSN of MCA1013 (pfam03176 or IPR004869) shows that the protein is grouped into a small cluster containing 5 nodes (where sequences with >50% identity are merged into a single node). Structural prediction shows MCA1013 shares a high homology with a membrane-bound hopanoid transporter, HpnN (PDB 5KHS, *Burkholderia multivorans*), which belongs to a subfamily of RND (Resistance-Nodulation-Division) transporters that transport hopanoids from the cytoplasm to the outer membrane in bacteria (38); HpnN is located in a different cluster than MCA1013 (Fig.S1A, Table S2). Similarly, MCA1014 (pfam12974) is grouped into a single cluster consisting of 16 sequences, with most of the sequences belong to *Methylococcaceae* family. Structural prediction of MCA1014 shows homology to PhnD (PDB 3QK6), a periplasmic solute-binding protein of the phosphonate uptake system in *E. coli* (39). However, the MCA1014 cluster in the SSN is separate from all structurally characterized PhnD homologs (Fig.S1B, Table.S2). Finally, the SSN of MCA1015 (pfam16811 or IPR031823) reveals that the protein is also grouped into a separate cluster consisting of 13 sequences, with most from *Methylococcaceae* and *Deltaproteobacteria*. Only one protein in the SSN has structural information: TP0956 from *Treponema pallidum* (PDB 3U64) is found in the second largest cluster of the network (Fig.S1C, Table S2). Notably, compared with SSNs of MCA1013 and MCA1014, that of MCA1015 shows the family is predominantly bacterial, with only 1 and 5 sequences from eukaryotes and archaea, respectively (Fig.S2A-C).

In summary, homology analyses reveal MCA1013, MCA1014, and MCA1015 are similar to transporters involved in distinct bacterial transport systems. However, their separation into distinct clusters in the SSNs imply their functions are not identical to characterized homologs. Because the MCA genes always co-occur with sterol synthesis genes (*osc, sdmA* and *sdmB*) in methanotrophs that produce sterols, we hypothesize that MCA1013/ MCA1014/ MCA1015 represent a novel transport system for sterols in bacteria (Fig.1B, Fig.S3).

Sterol interaction with transporters

To first determine whether MCA1013, MCA1014, and MCA1015 are putative sterol transporters, we conducted pull-down assays to assess their binding of sterols from *M. capsulatus* lipid extracts. In these studies, recombinantly-produced pure proteins (Fig. S4) were incubated in excess with either the total lipid extract (TLE) from *M. capsulatus* or a specific polar fraction enriched for native sterols (HS). Using HisPur Ni-NTA resin, the proteins were isolated, the bound lipids were extracted, and identified by GC-MS. In the presence of the lipid extracts (TLE and HS), PD-MCA1013 (the periplasmic domain of MCA1013; details in method), MCA1014, and MCA1015 bound the 4-monomethyl and 4,4-dimethyl sterols (Fig.2A, Fig.S4). In contrast, neither the 4-monomethyl nor the 4,4-dimethyl sterol were present when the sterol transporters were incubated with the DMSO negative control. These results indicate that PD-MCA1013, MCA1014, and MCA1015 preferentially bind C4-methylated sterols in a mixture of native *M. capsulatus* lipids.

Quantitative analyses of transporter-sterol interactions

The pull-down assays demonstrated that 4-methyl and 4,4-dimethyl sterols are substates for the putative sterol transporters. We next used MST to determine the equilibrium binding affinities for these substrates. Binding curves were generated upon titration of the labeled proteins with serially diluted sterol substrates, and the binding affinities were calculated by fitting the curves with different kinetic models (Fig.2B, methods). The "K_d" model applies to binding events with a single binding site or multiple independent binding sites, while the Hill model is invoked in instances where multiple binding sites that exhibit cooperativity are present, with Hill coefficients (n_H) greater than 1 signaling positive cooperativity. The average equilibrium dissociation (K_d) constant for PD-MCA1013 binding to 4-methylsterol was determined to be 4.41±0.14 μM (n_H=4.0) and ≥203±155 μM for the interaction with 4,4-dimethylsterol. For MCA1014, the K_d value of 2.43±0.13 μM (n_H of 2.4) was determined for 4-methylsterol; and ≥232±84.3 μM for 4,4-dimethylsterol. With MCA1015, the determined values were 2.50±0.39 μM and ≥80.9±14.2 μM for 4-methylsterol and 4,4-dimethylsterol, respectively. All K_d values for the 4,4-dimethylsterol could not be accurately determined due to large errors caused by

non-saturation of the substrate, even at the highest substrate concentration. In all three proteins, we measured clear binding to sterols, with marked preference for the C-4 monomethylated substrate. Additionally, MCA1013 and MCA1014 exhibit cooperativity for binding to substrate, indicating multiple binding sites are plausible.

To further define the specificity of MCA proteins for sterols, we again used MST to detect their interactions with cholesterol and lanosterol, two sterols lipids that differ in their methylation at the C-4 and C-14 positions, as well as their unsaturation patterns in the core ring structure. MST results revealed that no interactions exist between PD-MCA1013 and MCA1014 with either cholesterol or lanosterol. However, MCA1015 had detectable interaction with both sterols, hinting that the protein is more tolerant to non-native substates (Fig.S5).

Crystallographic structure of MCA1014

To determine the structural basis for sterol recognition and binding, we crystallized MCA1014 (Table S1). The structure of MCA1014 was determined by experimental phasing to the Bragg diffraction limit of 1.6 Å. The electron density allows the unambiguous building of a model of MCA1014, except for one disordered loop spanning residues 202-204. The MCA1014 structure comprises two globular α/β domains (Fig.3A, domains A and B) that form a horizontal cleft in the middle. Domain A contains six β-strands, while domain B contains five. All the β-strands are flanked by α-helices. The domains are connected by two loops: one 9-residue loop that runs from β4 in domain A to β5 in domain B, and a second 8-residue loop that connects β9 in domain B and β10 in domain A. These two long loops could work as a hinge of the structure to allow a bending motion of two domains. In addition, an α-helix formed by the C-terminal residues lies behind the cleft of and may also allow flexibility of the two domains. The flexible hinge presents the potential for conformation change upon substrate binding. Several extend loops are found on the domain-domain interface. The loops (residues 34-38 and residues 171-177) exhibit poor densities, suggesting conformational sampling of more than one state in preparation to undergo a rearrangement upon binding of the substrate (Fig.S6). The cleft between domains A and B forms a bulky cavity with an area of ~ 1142 Å and a volume of ~ 953 Å that is only accessible to solvent from one side of the structure. The

cavity is predominately hydrophobic, suggesting large and hydrophobic substrates are more favorable (Fig.3B). The residues whose side chains form the cavity are shown in Table S3.

A DALI structural homology search (40) suggests the structure of MCA1014 closely resembles that of periplasmic binding proteins (PBPs) which share a bilobed architecture and undergo conformational change upon substrate binding (41). A comparison of the MCA1014 structure with the phosphonate-binding protein PhnD from *E. coli* with and without its substrate (2-aminoethyl phosphonate, 2AEP, PDB 3P7I and 3S4U) was performed (39). The comparison revealed that MCA1014 adopts a conformation more similar to the unliganded PhnD (3S4U) over the liganded one (3P7I, Fig.3C). However, the cleft in MCA1014 is less open (~10 Å) than that in apo-PhnD (~22 Å). There are two possible explanations for this: 1) compared with apo-PhnD structure, the cleft of MCA1014 is deeper and more hydrophobic. Given that the substrates of MCA1014 are sterols, which have a nearly planar architecture and are more hydrophobic than phosphonates, the different architecture of the cleft in MCA1014 may be related to its functional role; 2) it is known that binding of non-substrate ligands can at times trigger conformational change in PBPs (42). Although no obvious ligand density is observed in the MCA1014 structure, weak unmodeled densities were observed inside the cavity. Most of these are small patches that cannot be satisfactorily modeled with water and do not closely correspond to any reagent used for crystallization. It is therefore possible that the binding of ligands co-purified with the protein leads it to adopt a “partially-open” conformation in our structure.

To obtain insights into the mechanism of substrate binding, we used the Auto Dock Vina (43) to generate a docking model of the MCA1014 structure and 4-methylsterol. The docked ligand with the lowest free energy (-10.4 kcal/mol) was selected. The docking result suggests the sterol is positioned in the hydrophobic cavity in domain A, with its polar head pointing toward domain B. Hydrogen bonds could be formed between the hydroxyl group of the sterol and residue Glu 118 and Tyr 120 at the bottom of the cleft (Fig.3D). The aliphatic tail resides in the hydrophobic cavity lined by the side chains of Tyr 33, Pro 34, Met 47, Met 50, Ser 91, Leu 92, Pro 110, Glu 118, Tyr 120, Phe 150, Leu 219, Met 220, Leu 253, Cys 254, Phe 257, Ile 259 and Phe 262 in domain A. Sequence

alignment shows these residues are highly conserved between proteins in the same MCA1014 cluster of the SSN, implying their significance in forming a substrate binding site (Fig.S7). Mutagenesis to remove or reduce the interactions with the polar head and the aliphatic tail did not show obvious influences (<2-fold) on the sterol affinity, suggesting that the interaction between the protein and sterol substrate is extensive and principally depends on the hydrophobic interaction (Fig.S8). Of additional note in the structure is the presence of positive difference density that could not be fit with water in proximity to the position of the sterol polar head. One of them could be modeled with a chloride ion, which can also form polar interactions with the surrounding Glu 118, Tyr 120 and Phe 150. Additionally, unmodeled positive density is also observed in the cavity where the aliphatic tail is positioned (Fig.S9).

A second pocket (area=172 Å; volume=117 Å) is found in domain B adjacent to the central cleft. This pocket is localized behind the unmodeled loop (residues 202-204). The extremely poor density in this part implies of high flexibility. The high hydrophobicity of this pocket suggests it could be an additional binding site for the sterol substrate (Fig.3B; Fig.S10), in support of the binding data that suggests cooperativity between multiple substrate binding sites. However, the sterol substrate failed to dock into this pocket with reasonably low Gibbs free energy, raising a possibility that the site may be a subsidiary binding site or that cooperative binding can only occur after some conformational change of the protein. Other separate hydrophobic areas are also observed on the protein surface (Fig.S10), and a new hydrophobic interface could be formed by two hydrophobic areas across the cleft upon conformational change induced by substrate binding. These areas may be helpful for mediating protein-protein interactions (e.g., with MCA1013 or MCA1015 to facilitate substrate transfer).

Crystallographic structure of MCA1015

The structure of MCA1015 was determined by experimental phasing to a Bragg diffraction limit of 1.9 Å. Two monomers (RMSD of 0.656 Å for all Cαs) in the asymmetric unit form a dimer with 180° rotational symmetry. Each monomer adopts an all-α-helical structure with 12 α-helices in total, designated with H1 to H12; 9 antiparallel arranged helices (H4-H12) form an irregular ring with a pore running through the center. H4, H6,

H8, H10, H12 constitute the inside wall of the pore, while the others form the outside wall. Two helices (H1-H2) at the N-terminus sit on the side-top of the ring with a short helical turn (H3) connecting the two (Fig. 4A). It is notable that, unlike the other helices which all have a straight structure, the H7 helix is bent, with the lower part bent towards the C-terminal of the protein at an angle of ~54 °. This bent part contributes to the formation of the pore (Fig.S11).

Hydrophobicity analysis shows that MCA1015 has a predominantly hydrophilic exterior surface (Fig.S12). Three cavities are found in the center of the structure. We annotated them as Tunnel 1 (TN1), Cavity 1 (CA1) and Cavity 2 (CA2), with volumes of 19.2 Å×1.8 Å, 12.2 Å×2 Å and 16 Å×1.8 Å, respectively. TN1 is closer to the N-terminal face of the structure, while CA1 and CA2 are proximal to the C-terminal face and form two open hydrophobic pockets. TN1 exhibits dichotomous hydrophobicity, with the half side homolateral to CA1 being hydrophobic while the opposite side is more hydrophilic (Fig. 4B). The highly hydrophobic environment inside the MCA1015 structure suggests it prefers hydrophobic substrates. Residues that line the cavities are listed in Table.S4. Notably, electrostatic calculation shows that the C-terminal face, including the entrance to the cavities, exhibits strong negative charge compared to the N-terminal of the protein (Fig.S13). This charged surface may mediate protein-protein interactions, including with MCA1014.

To assess the substrate-binding mode, we performed the docking with MCA1015 and 4-methysterol using Auto Dock Vina. The final docking model was obtained by choosing the one with the least minimum free energy (-9.7 kcal/mol). In this model, the sterol is positioned with its head-down in hydrophobic cavity (CA1) lined by the side chains of Ala 132, Met 135, Ala 138, Leu 139, Leu 142, Leu 145, Arg 166, Leu 167, Met 170, Leu 171, Lys 174, Leu 204, and Phe 208, distributed in H6, H7 (bent part), H8 and H10. All these residues show high conservation in the sequence alignment with proteins in the same SSN cluster (Fig. S14). No polar interaction is observed between the sterol with surrounding residues. Notably, the sterol is positioned near the entrance of the tunnel from the C-terminal side. Next to the entrance is an extended loop (residues 176-183) that consists of hydrophobic and neutral residues. It appears that this loop could serve as

a “lid” to the tunnel. The lid is open in the apo structure and might become closed upon substrate binding to restrict access to the tunnel (Fig.4C).

A structure similarity search using the DALI server revealed high homology between MCA1015 and TP0956 (PDB 3U64, *Treponema pallidum strain Nichols*). TP0956 is a T-component protein of a subfamily of TRAP (tripartite ATP-independent periplasmic) transporters known as TPATs (tetratricopeptide repeat-protein associated TRAP transporters), which are implicated in transport of hydrophobic ligands in bacteria (44, 45). Tp0956 contains four helical hairpins that are similar to tetratricopeptide repeat (TPR) motifs, which are hallmarks of proteins that are involved in protein-protein interactions (46). Compared to Tp0956, two antiparallel α-helices exist in Tp0956 are missing in the structure of MCA1015 (between H5-H6), as well as a shorter H4 and H5. But the TPR-like substructure is also observed in MCA1015 and forms part of the pore and lateral surface of the protein, hints that MCA1015 might engage with itself or other proteins (Fig. S15). Size-exclusion chromatography with multiangle light scattering (SEC-MALS) data shows that the purified MCA1015 exists both as a monomer and dimer in solution (Fig. S16), with the dimer being a small fraction of the total protein. However, a structure of the dimer form reveals little interaction between the two monomers in the asymmetric unit. It suggests that no functional dimer exists in solution. However, we cannot rule that dimerization occurs upon binding of the substrate.

Comparison to eukaryotic sterol transporters

In most eukaryotic cells, lipid transfer proteins (LTPs) are used to transfer sterols between subcellular membranes by non-vesicular transport to maintain sterol homeostasis in different sub-cellular organelles (47). The most-studied family in mammals is the steroidogenic acute regulatory protein (StAR)-related lipid transfer (START) domain (STARD) proteins. All fifteen STARD proteins identified to date possess lipid-harboring START domain and can be divided into six subfamilies according to their sequence similarity and ligand specificity. In addition to STARD proteins, a novel yet evolutionarily conserved family that contain the START-like domains called the GRAMD1s/Lam/Ltc family (GRAMD1s in mammals and Lam/Ltc in yeast) were shown to mediate non-vesicular sterol transport from the plasma membrane to the endoplasmic

reticulum to maintain sterol homeostasis (48). A third family consists of the Osh/ORP/OSBP proteins (Oxysterol-binding homology in yeast and Oxysterol-binding Protein/OSBP-Related Protein in mammals) are considered to be either sterol-sensing and/or sterol-transfer proteins (49). Besides these, two Niemann-Pick type C proteins, NPC1 and NPC2, were also reported to directly bind and transport cholesterols and various oxysterols; defective NPC proteins causes NPC disease, a fetal neurodegenerative disorder where sterols aberrantly accumulate in the lysosome (50).

To better understand the differences between eukaryotic and bacterial sterol transport, we compared the structures of MCA1014 and MCA1015 with well-studied eukaryotic sterol transporters from different families. Rabbit SCP2(PDB 1C44), human STARD4 (PDB 6L1D), yeast Lam4p (PDB 6BYM), yeast Osh4 (PDB 1ZI7 and 1ZHY) and yeast NPC2 (PDB 6R4M and 6R4N) were chosen for the comparison. The crystal structures of STARD4, Lam4p and Osh4 show that they share a similar fold in that they are dominated by β-sheet motifs, which wrap around the longest α-helix and fold into a half-barrel/groove structure with the other side (or two) packed with several α-helices (Fig.S17A-D). A central hydrophobic cavity can accommodate sterols. In most cases, the sterols are positioned proximal to the tunnel entrance with the hydroxyl end oriented towards the inside of the tunnel. Also, a “lid” comprising a flexible “Ω1” loop is observed near the entrance of tunnel in STARD4 and sterol-bound Lam4. This loop is replaced by an N-terminal α-helix in Osh4, which serves a similar function. The SCP2 structure also exhibits an α/β-fold with a “lid”-like α-helix observed close to the N-terminus (Fig.S17E). The sterol-bond NPC2 adopts a different immunoglobulin-like β-sandwich fold with seven strands arranged into two β-sheets with a loosely-packed hydrophobic cavity formed between the two sheets. The sterol is positioned at the entrance of the cavity with the hydroxyl end pointing out; no “lid”-like structure is observed in NPC2 structure (Fig. S17F-G).

Our comparative analysis shows that the eukaryotic sterol transporters share very little resemblance with MCA1014 and MCA1015. The bilobed architecture of MCA1014 is pronounced in its divergence from eukaryotic sterol transporters, suggesting that this protein evolved separately and may work in a distinct manner. For the all-helical MCA1015, the one similarity is the presence of a hydrophobic tunnel formed in the center

of the protein. Additionally, MCA1015 has a loop next to the tunnel entrance; it is possible that similar to the eukaryotic proteins, the loop is used as a “lid” to restrict access to the substrate binding site (Fig. 4C).

Discussion

Since their discovery more than 40 years ago, the function of bacterial sterols has remained a mystery. Recent identification of bacterial sterol demethylase enzymes that diverge from their eukaryotic counterparts provided renewed enthusiasm for deciphering the molecular machinery involved in bacterial sterol synthesis and trafficking (18). The aerobic methanotroph *M. capsulatus* is known to synthesize 4-methylsterol from 4,4-dimethylsterol and to accumulate these sterols in its outer membrane. Transport of the C4-methylsterol products to the outer membrane is crucial for maintaining the sterol homeostasis in the cytoplasm. Until now, the transport process remained elusive owing to lack of knowledge of transporters that facilitate the trafficking. We now fill this gap with the identification of three new sterol transporters in *M. capsulatus*: MCA1013, MCA1014 and MCA1015. Pull-down assays and quantitative in vitro binding analyses reveal their specificity for 4-methylsteros. High resolution crystal structures show that MCA1014 shares an architecture similar to periplasmic binding proteins that undergo a conformation change induced by substrate binding. For MCA1015, the “all-helical” structure has cavities of striking hydrophobicity. Docking of 4-methylsterol into both structures reveal the sterol binds to bulky hydrophobic cavities. The polar head of the sterol could be stabilized through hydrogen-bonding to several amino acids in MCA1014. For MCA1015, a continuous hydrophobic environment is observed, leaving few hydrophilic areas that might stabilize the substrate through polar interactions. The combination of structures and docking studies suggest that hydrophobic interactions dominate the binding of sterols in both MCA1014 and MCA1015.

Structural comparison shows that neither MCA1014 nor MCA1015 share resemblance with known eukaryotic sterol transporters. The marked divergence between the structures of sterol transporters in bacteria and eukaryotes suggests distinct evolutionary paths for the sterol trafficking. The bioinformatic analyses show that MCA1015 and its homologs rarely exist in eukaryotes, implying their unique function in

bacteria. Also, no member of the MCA1014 family has been identified as a putative sterol transporter. Instead, these proteins share structural similarities with bacterial transporters implicated in the trafficking of essential nutrients such as sugars, amino acids, vitamins, solutes, and metal ions (51, 52). The implication of MCA1014 and MCA1015 in the transport of sterols expands the substrate repertoire for their respective families while hinting that there are more bacterial sterol transporters to be discovered.

Acknowledgments

This work was supported in part by NSF grant MCB1919153 (to L.M.K.D. and P.V.W.). Crystallographic data were acquired at the Stanford Synchrotron Radiation Lightsource. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (P30GM133894). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

Figures and legends

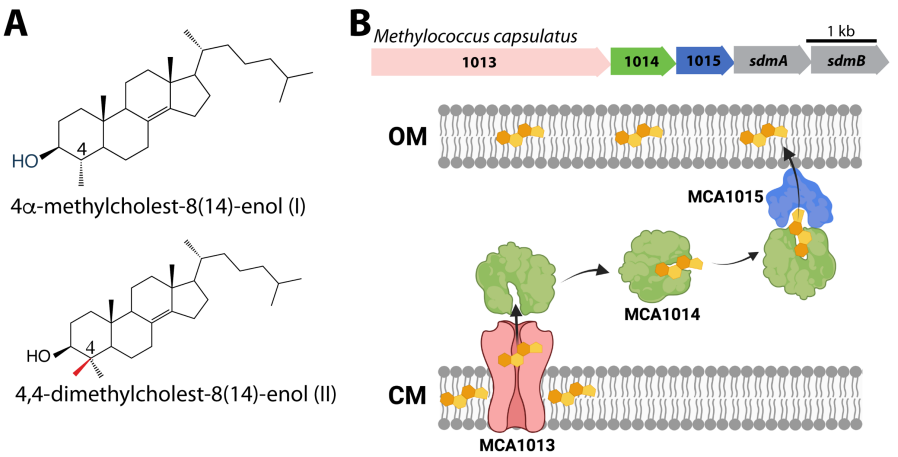


Figure 1. (A) Structures of C-4 methylsterols synthesized by *Methylococcus capsulatus*. (B) Operon of sterol biosynthesis and transport genes and (top) and schematic description of their localization.

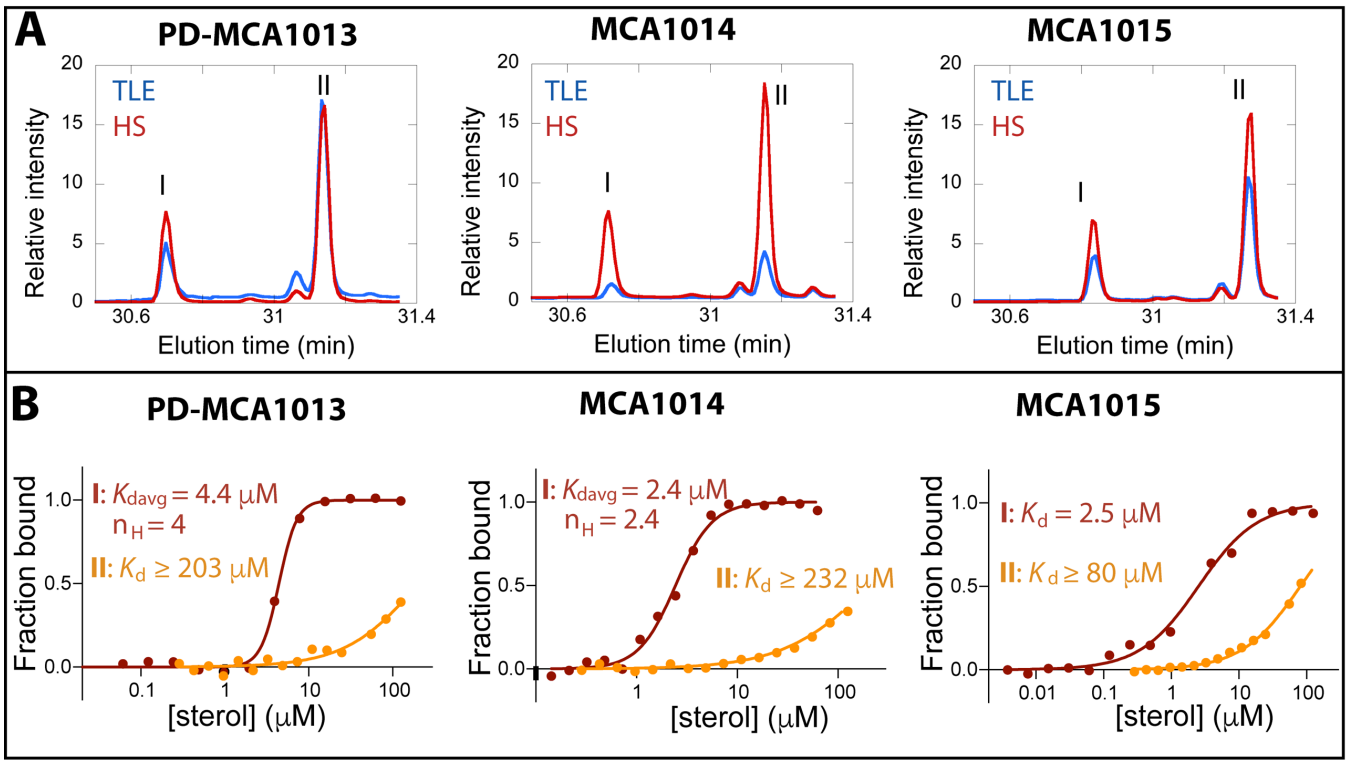


Figure 2. Binding of C-4 methylsterols to transporter proteins. (A) Pull-down of sterols in the total lipid extract (TLE) and hydroxy lipids (HS) of *M. capsulatus* using 40 μM purified PD-MCA103, MCA1014, and MCA1015. The proteins preferentially bind 4-monomethyl (I) and 4,4-dimethyl (II) sterols. (B) MST analysis of the proteins (50 nM) binding to C-4 methylsterols (3.8 nM -1125 mM). Equilibrium dissociation constants (K_d) and Hill coefficients (n_H) are reported on the plots.

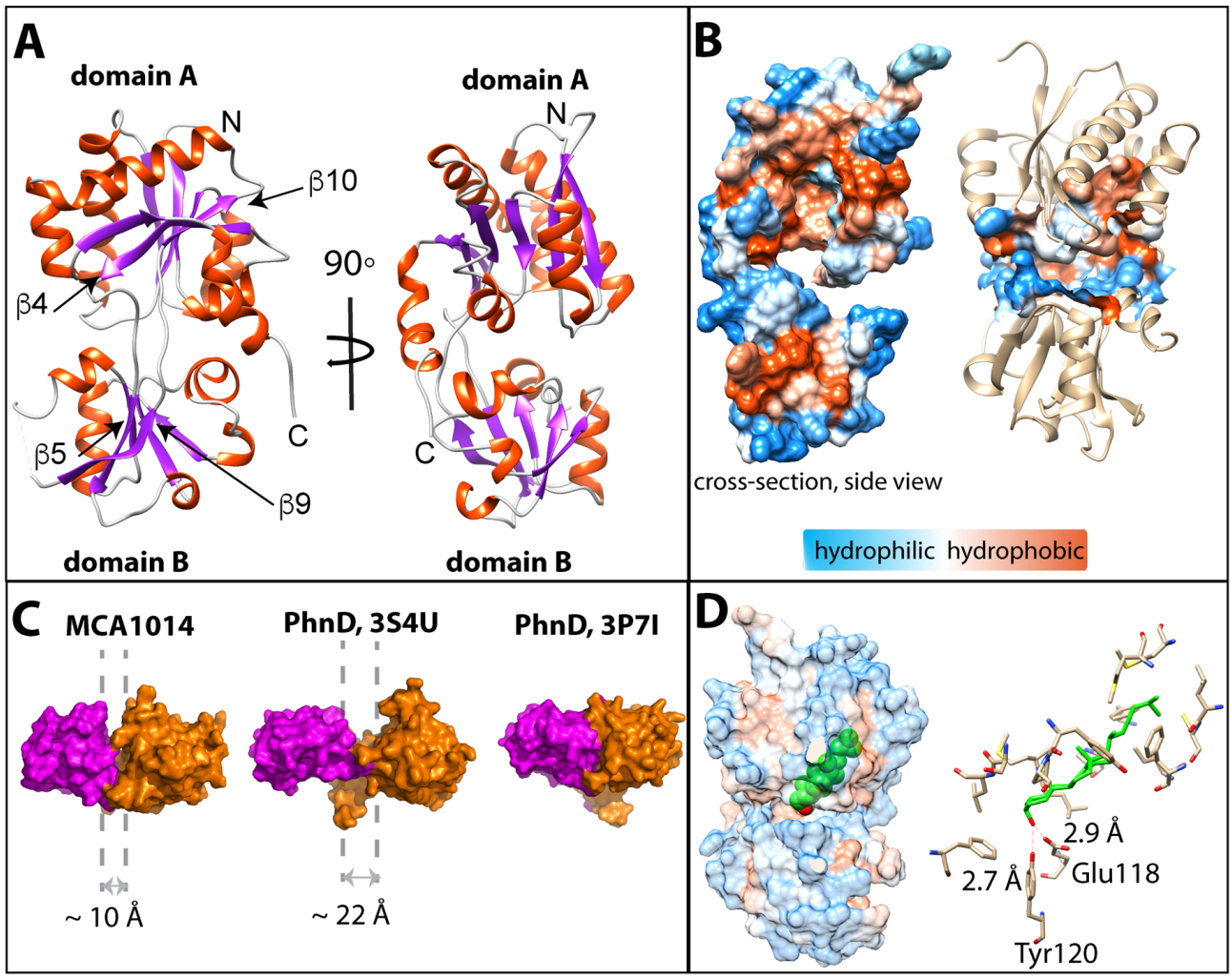


Figure 3. Structure of MCA1014. (**A**) Cartoon representation of MCA1014 with the α-helixes and β-strands colored in red and magenta, respectively. (**B**) Hydrophobicity representations of MCA1014 show a cleft in the middle and a cavity in domain A. The cavity is dominatingly hydrophobic and open only to one side. (**C**) Comparison of the cleft between the two domains of MCA1014, apo-PhnD (3S4U), and liganded PhnD (3P7I). MCA1014 displays a similar but narrower open (~10 Å) compared to non-ligand PhnD (~22 Å). (**D**) A docking model generated with the structures of MCA1014 with 4-methylsterol. The sterol is positioned in the hydrophobic cavity formed by the cleft and domain A. The sterol is shown in sphere and colored in green with the hydroxyl group is colored in red; the surrounding residues are shown in sticks and colored in tan. Hydrogen bonds between the hydroxyl group of the sterol and Glu118 and Tyr120 are shown as red dashed lines.

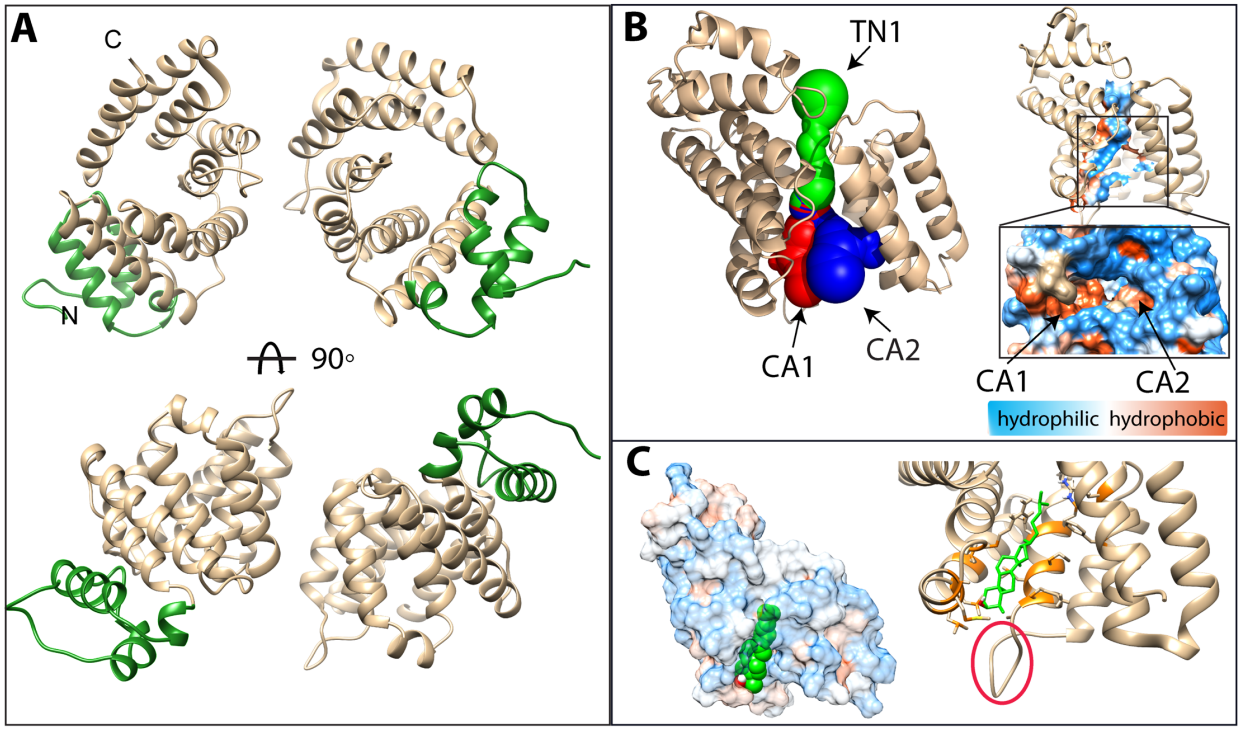


Figure 4. Structure of MCA1015. (**A**) Cartoon representation of MCA1015. The two α-helixes at the N-terminus are colored in green. Two monomers (RMSD of 0.656 Å for all Cαs) with 180◦ rotational symmetry exist in the asymmetric unit. (**B**) A representation of cavities MCA1015 (left, with Tunnel 1 colored in green, Cavity 1 in red, and Cavity 2 in blue) alongside hydrophobicity representation of the cavities (right). (**C**) A docking model generated with MCA1015 with 4-methylsterol. The sterol is colored in green with the hydroxyl group is colored in red; the surrounding residues are shown in stick and colored in orange. No polar interaction is observed between the sterol and surrounding residues.

Supporting Information

Structural insights into bacterial sterol transport

Liting Zhai,[a] Amber C. Bonds,[b] Hannah Oo,[a] Paula V. Welander,[b,1] Laura M. K. Dassama[a,1]

[a]Department of Chemistry and Sarafan ChEM-H, Stanford University

[b]Department of Earth System Science, Stanford University

[1]To whom corresponding should be addressed: welander@stanford.edu, dassama@stanford.edu

Materials and Methods

Bioinformatics

The SSNs were generated using EFI-EST sever (http://efi.igb.illinois.edu/efi-est/stepa.php). Proteins are represented by nodes. Only the nodes that share pairwise sequence alignments greater than the user-specified value will be connected by edges to create a network that reveals potential functional relationships among the proteins (35). HMMER webserver (https://www.ebi.ac.uk/Tools/hmmer/) was used to predict the superfamilies of MCA proteins. MCA1013 is predicted to be a member of MMPL family (pfam03176 or IPR004869). The SSN for this family was generated using the EFI-EST online tool (53) with an E-value threshold of 10^{-60} (>~35% sequence identity). UniRef50 (sharing >50% sequence identities will be merged into a single node) was used to facilitate the calculation and save time during the SSN creation. MCA1014 belongs to a family of ABC transporter phosphonate periplasmic substrate-binding proteins (pfam12974), which was selected to create an SSN at E-value threshold of 10^{-50} (>~40% sequence identity). MCA1015 is a member of TRAP transporter T-component family (pfam16811 or IPR031823), which facilitates the import of small molecules into cytoplasm (54). An SSN was constructed with this dataset at E-value threshold of 10^{-55} (>~40% sequence identity). UniRef90 standard (sequences that share ≥90% sequence identity are merged into a single node) was used in SSN creation for MCA1014 and MCA1015. Cystoscope software(55) was used to visualize the SSNs with the “organic” layout. The genome neighborhood network and diagram were generated using the Enzyme Function Initiative-Genome Neighborhood Tool (EFI-GNT; https://efi.igb.illinois.edu/efi-gnt/) with the SSNs generated above. The neighborhood reading frame was set to 10 frames and the minimal co-occurrence was set to 20%.

Protein expression and purification

The genomic DNA of *M. capsulatus* was subjected to several rounds of PCR to isolate the genes encoding MCA1014 and MCA1015 without their respective signal peptides (cleavage sites determined by the SignalP 5.0 server (56)). The final gene products were assembled into pET28a vectors containing TEV-cleavable N-terminal hexa-histidine (His-) tags or pET20b vectors containing C-terminal His-tags. DNA

sequencing of the plasmids confirmed the correct insertion of the genes. The MCA1013 gene was synthesized and inserted into pET28a after codon optimization for expression in *E. coli*. The soluble periplasmic domain of MCA1013 (PD-MCA1013) was designed based on the predicted structure of MCA1013 by AlphaFold (Fig. S18) (57). Two fragments spanning residues 54-277 and 510-732 were connected with a GSGSGSGS linker to replace the membrane spanning connecting sequence. The DNA encoding PD-MCA1013 was cloned into pET28a. All the protein sequences and primers are listed in Table S5.

PD-MAC1013, MCA1014 (residues 24-281) and MCA1015 (residues 16-264) proteins were expressed in *E. coli* BL21(DE3) cells grown at 37 °C until OD_{600nm} of 0.6. Protein expression was induced upon the addition of 0.25 mM isopropyl β-D-1thiogalactopyranoside (IPTG) while shaking at 18 °C for 12h. Cells were harvested through centrifugation at 8000 × g for 15 min and resuspended in lysis buffer containing 25 mM Tris-HCl (pH 8.0), 100 mM NaCl, 2 mM 2-mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride (PMSF). Cells were lysed by microfluidization and centrifuged at 11,000 × g for 40 min to remove cell debris. The supernatant was applied to a Ni-NTA affinity column (HiTrap IMAC FF, Cytiva product # 17092104), washed with lysis buffer supplemented with 15 mM imidazole to remove weakly bound proteins, and eluted with lysis buffer containing 300 mM imidazole. The eluate was concentrated and applied to a size-exclusion column (HiLoad 16/600 Superdex 200 pg, Cytiva product # 28989335) to remove additional contaminants, and purity was assessed through gel electrophoresis.

Selenium-methionine (Se-Met) labeled MCA1014 and MCA1015 were produced by supplementing the endogenous methionine synthesis of the expression host with Se-Met. The cells were cultured in M9 minimal media until OD_{600nm} of 0.4. L-lysine (100 mg/L), L-phenylalanine (100 mg/L), L-threonine (100 mg/L), L-isoleucine (50 mg/L), L-leucine (50 mg/L) and L-valine (50 mg/L) and L-Se-Met (60 mg/L) were added to the media. 0.25 mM IPTG was added after 1h to commence induction and cells were cultured for an additional 12h at 18°C before harvesting. The proteins were purified by the protocol described above.

The N-terminal His-tags of proteins were removed during overnight incubation at room temperature with TEV protease at a molar ratio of 1:100 (TEV: protein). Proteins without tags were further purified by size-exclusion chromatography. The molecular weight of proteins was determined by SEC-MALS (Wyatt Technology). The purity of all purified proteins was assessed using SDS-PAGE stained with Coomassie Brilliant Blue, after which aliquots of the purest fractions were flash-cooled in liquid nitrogen and stored at -80 °C.

Size-exclusion chromatography (SEC)-coupled multi-angle light scattering (MALS)

SEC-MALS was performed using an in-line Superdex 200 Increase 3.2/300 GL SEC column (GE Healthcare) combined with a miniDawn Multi-Angle Light Scattering (MALS) detector coupled with an Optilab refractive index detector (Wyatt Technology, Santa Barbara, CA, USA). 15 μL (~ 200 ng) protein was centrifuged at 13000 × g for 10 min before injected into the pre-equilibrated SEC column with buffer 25 mM Tris-HCl (pH 8.0), 100 mM NaCl and 2 mM 2-mercaptoethanol. Proteins were separated at a flow rate of 0.15 mL/min at room temperature. Molecular masses were calculated using the Astra6.1 software (Wyatt Technology).

Sterols extraction from *M. capsulatus*

Lipids were extracted from cell pellets of 2 L *M. capsulatus* cultures using a modified Bligh-Dyer extraction method as previously described (18). Cells were re-suspended in 10:5:4 (vol:vol:vol) methanol: dichloromethane (DCM):water and sonicated for 1 hour. The organic phase was separated with 1:1 (vol:vol) DCM:water, later washed three times with DCM and then transferred and dried with N_2 to yield the total lipids extraction (TLE). The TLE underwent Si column chromatography to isolate the alcohol soluble lipids. The 4-methylsterol and 4,4-dimethylsterol were purified from the alcohol soluble lipids by liquid-chromatograph (LC) using an Agilent 1260 Infinity II LC System. Briefly, lipids were run through a InfinityLab Poroshell120 EC-C18 column (4.6 x 150 mm, 2.7 Micron). The solvent system consisted of Solvent A: MeOH (with .04% formic acid and 0.1% ammonia), Solvent B: acetonitrile, and Solvent C: Water (with .04% formic acid and 0.1% ammonia). The solvent gradient started with 84% Solvent A, 1% Solvent B, and 15% Solvent C and

ended with 100% Solvent A over the course of 50 minutes. Isolated sterols were derivatized to trimethylsilyl ethers with 1:1 (vol:vol) Bis(trimethylsilyl)trifluoroacetamide: pyridine and confirmed using GC-MS analysis as described in Lee et al, PNAS 2018 (18). Lipids were separated with an Agilent 7890B Series GC equipped with 2 Agilent DB-17HT columns in series, each (30 m x 0.25 mm i.d. x 0.15 µm film thickness) with helium as the carrier gas at a constant flow of 1.1 mL/min and programmed as follows: 100°C for 2 min, then 12°C/min to 250°C and held for 10 min, then 10°C/min to 330°C and held for 17.5 min. 2 uL of each sample was injected in splitless mode at 250°C. The GC was coupled to an Agilent 5977A Series MSD with the ion source at 230°C and operated at 70 eV in EI mode scanning from 50 to 850 Da in 0.5 s. Lipids were identified based on their retention time and comparison to previously published spectra.

Protein-Lipid Pull Down Assay

Proteins (PD-MCA1013, MCA1014, MCA1015) were incubated with two different *M. capsulatus* lipid extracts: total lipid extractions (TLE) or the enriched sterol fraction (HS), both dissolved in DMSO. The TLE (139.5 µg total) contained 65.7 ng of 4-monomethyl sterol and 338.9 ng of 4,4-dimethyl sterol. The HS fraction (99.5 µg total) contained 313.8 ng and 967 ng of 4-monomethyl sterol and 4,4-dimethyl sterol, respectively. The reaction conditions are as follows: 40 µM protein (PD-MCA1013, MCA1014, MCA1015) was incubated with 7.5 µL of either the TLE or HS lipids. The reactions were carried out in reaction buffer of 200 mM HEPES pH 8.0, 100mM NaCl and 0.05% Triton X-100. As a negative control, the proteins were incubated with DMSO that did not contain any lipids. The reactions (both control and experiment) were incubated at 25 °C for 20 hours. The proteins were purified using HisPur Ni-NTA his resin (Thermos Fisher), lipids from the eluted protein fractions were extracted using a modified Bligh Dyer protocol (18, 58), and lipid samples were analyzed by GC-MS as previously described. HisPur Ni-NTA wash buffer: 25 mM Tris pH 8.0,100 mM NaCl and elution buffer: 25 mM Tris pH 8.0,100 mM NaCl, 300mM imidazole. SDS PAGE electrophoresis was conducted to demonstrate that the fractions analyzed did in fact contain protein.

Microscale Thermophoresis

Microscale thermophoresis (MST) experiments were performed with NanoTemper Monolith NT.115 instrument according to the manufacturer's instructions (NanoTemper Technologies GmbH). In brief, His-tagged MCA1013, MCA1014, and MCA1015 were fluorescently labeled with Monolith His-Tag Labeling Kit RED-tris-NTA 2nd Generation kit according to the manufacturer's instructions. The 4-methylsterol and 4,4-dimethylsterol were extracted from *M. capsulatus*, while lanosterol (CAS:79630, Sigma-Aldrich) and cholesterol (CAS: AAA1147018, fisher Scientific) were prepared in DMSO with a final concentration of 3.9 mM and 7.76mM, respectively.

50 nM of labeled proteins were diluted in assay buffer (20 mM HEPES pH 8.0, 100 mM NaCl, and 0.05% Tween 20), in the presence of sixteen-step serial dilution of 4-methylsterols or 4,4-dimethylsterols (ligand solubilized in DMSO at a final constant concentration of 5%). After incubation, samples were loaded into Monolith NT.115 Premium Capillaries (NanoTemper Technologies GmbH) and measurements were taken at a constant temperature of 23 °C. MST traces were collected with an LED excitation power of 40% and an MST laser power of 40%. The MO. Control Analysis software (NanoTemper) was used to analyze the interaction affinity and the equilibrium dissociation constant (K_d) for each ligand using the K_d fit model or Hill fit model. The data was exported and plotted in Prism.

Equation of K_d -model fitting:

$$f(c) = \frac{c + c_T + K_d - \sqrt{(c+c_T+K_d)^2 + 4\,c\,c_T}}{2\,c_T}$$

f(c) is the fraction bound at a given ligand concentration c. K_d is the equilibrium dissociation constant and c_T is the final concentration of target in the assay.

Equation for Hill-model fits:

$$f(c) = Unbound + \frac{Bound - Unbound}{1 + \left(\frac{EC_{50}}{c}\right)^{n_{Hill}}}$$

f(c) is the fraction bound at a given ligand concentration, c. "Unbound" is the F_{norm} signal of the target. "Bound" is the F_{norm} signal of the Complex. EC_{50} is the half-maximal effective concentration and n_{Hill} is the Hill coefficient. The Hill coefficient n_{Hill} describes

the degree of cooperativity of an interaction: $n_{Hill}>1$ indicates positive cooperativity, while $n_{Hill}<1$ indicates negative cooperativity.

Crystallization of MCA1014 and MCA1015

Sparse-matrix crystallization screens were performed with purified proteins at concentrations of about 5 mg/mL for each. All crystals were obtained by the sitting-drop vapor diffusion method by mixing 0.5 μL protein and 0.5 μL precipitant solution at 20°C. Native and Se-Met MCA1015 crystals appeared in the MCSG2-B3 condition containing 20% (v/v) PEG3500 and 0.2 M ammonium citrate dibasic after 2 weeks of incubation at 20 °C. Crystals were cryoprotected in the same conditions with the addition of 33% (v/v) PEG400 before being flash-cooled and stored in liquid nitrogen. Native and Se-met MCA1014 crystals appeared in the MCSG1-C12 condition containing 25% (v/v) PEG3500 and 0.1M Bis-tris pH 6.5 after 1 week of incubation at 20 °C. Crystals were cryoprotected in the same conditions with the addition of 20% (v/v) PEG400 before being flash-cooled and stored in liquid nitrogen.

Data collection, processing, and structure determination

All diffraction datasets were collected at 100 K at the Stanford Synchrotron Radiation Lightsource (Menlo Park, CA, USA). MCA1014 datasets were collected used a Pilatus 6M detector, at beamline BL9-2 while MCA1015 datasets were collected at beamline BL14-1 with MARCCD325 detector. Native datasets were collected at a fixed wavelength at 0.97946 Å. MAD datasets were collected on ligand-free Se-Met-substituted crystals, at three wavelengths near the Se K-edge (0.97896 Å, 0.91837 Å, and 0.97930 Å). All diffraction datasets were indexed and processed by XDS (59, 60) and HKL3000 packages (61). The structures of MCA1014 and MCA1015 were phased by the MAD method with Phenix AutoSol using the Se-anomalous diffraction (62). Subsequent density modification gave excellent electron-density maps, which allowed the building of the models. Iterations of refinement were carried out with Phenix Refine, and model building was performed in Coot. The statistics of data collection and refinement are listed in Table S1.

All structural figures were prepared using PyMOL (63) or Chimera (64). Cavities in MCA1014 were assessed using CASTp (65) with a probe radius of 1.4 Å, equivalent to the radius of water. The tunnel through MCA1015 was identified with MOLE2.0 (66) using default settings and a probe radius of 1.4 Å.

Docking of substrate to MCA1014 and MCA1015

All molecular docking studies were performed using the Autodock Vina package (43). Briefly, chain A of MCA1014 and MCA1015 crystal structure was used as the rigid molecule. The molecule was added with non-polar hydrogens and assigned partial atomic charges in Chimera (64). The coordinates of 4-methylsterols were generated using ChemDraw (PerkinElmer Informatics). A grid box with 43 × 47 × 47 (for MCA1014) grid points and 55.8 × 29.3 × 41(for MCA1015) was used as the searching space. The results were clustered and ranked. Result analyses and figure rendering were performed using PyMOL (63) and Chimera.

Supporting figures and legends

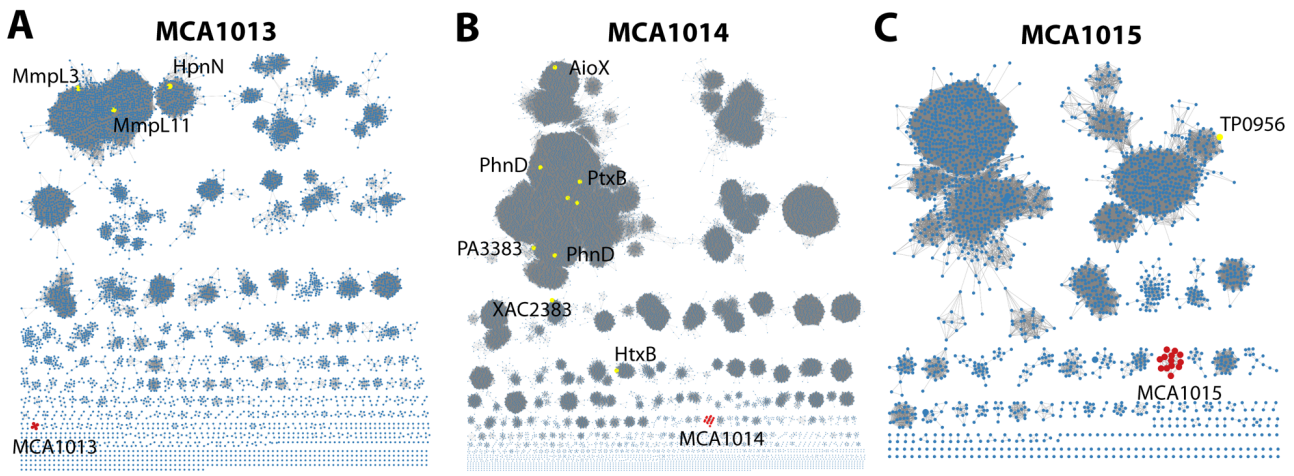


Figure S1. Bioinformatic analysis of MCA proteins. (A-C): Protein sequence similarity networks (SSN) of MCA1013, MCA1014 and MCA1015. Colored in red are clusters containing MCA proteins.

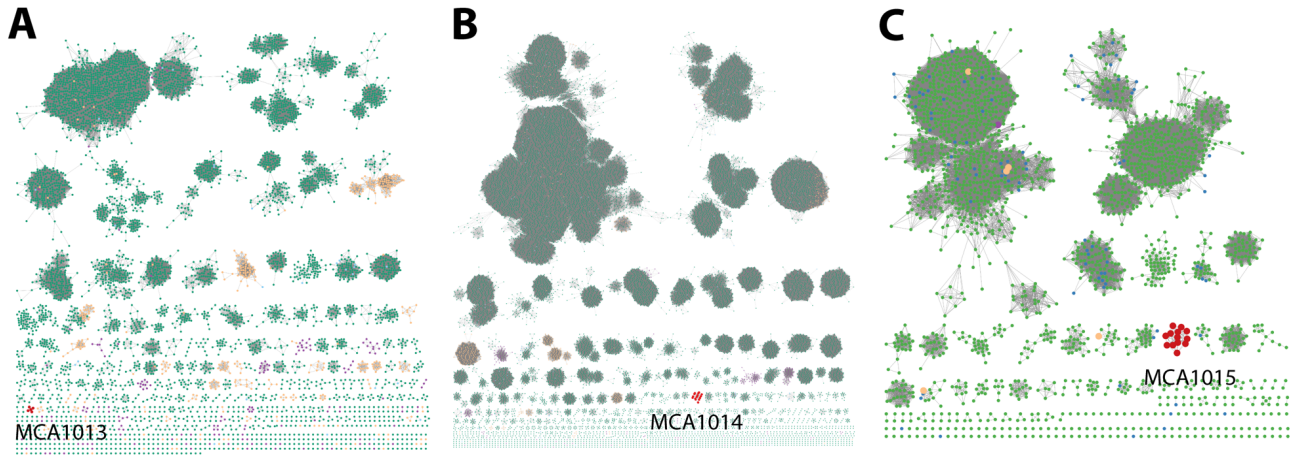


Figure S2. Taxonomic distribution of MCA protein families showing their presence in bacterial (green), eukaryotic (purple), and archaeal (orange) genomes.

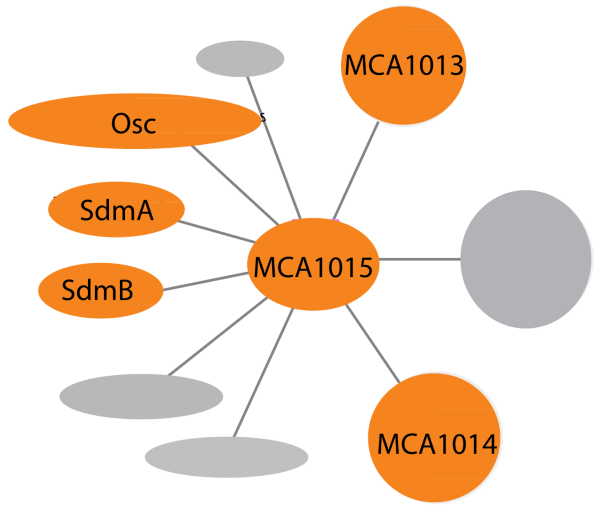


Figure S3. Genome neighborhood network of MCA proteins. (**A**) SSN cluster hub- and Pfam family spoke-nodes. Shown in circles and ovals are genes within 10 genes of MCA1015; colored in orange are those with identified roles in sterol synthesis and trafficking.

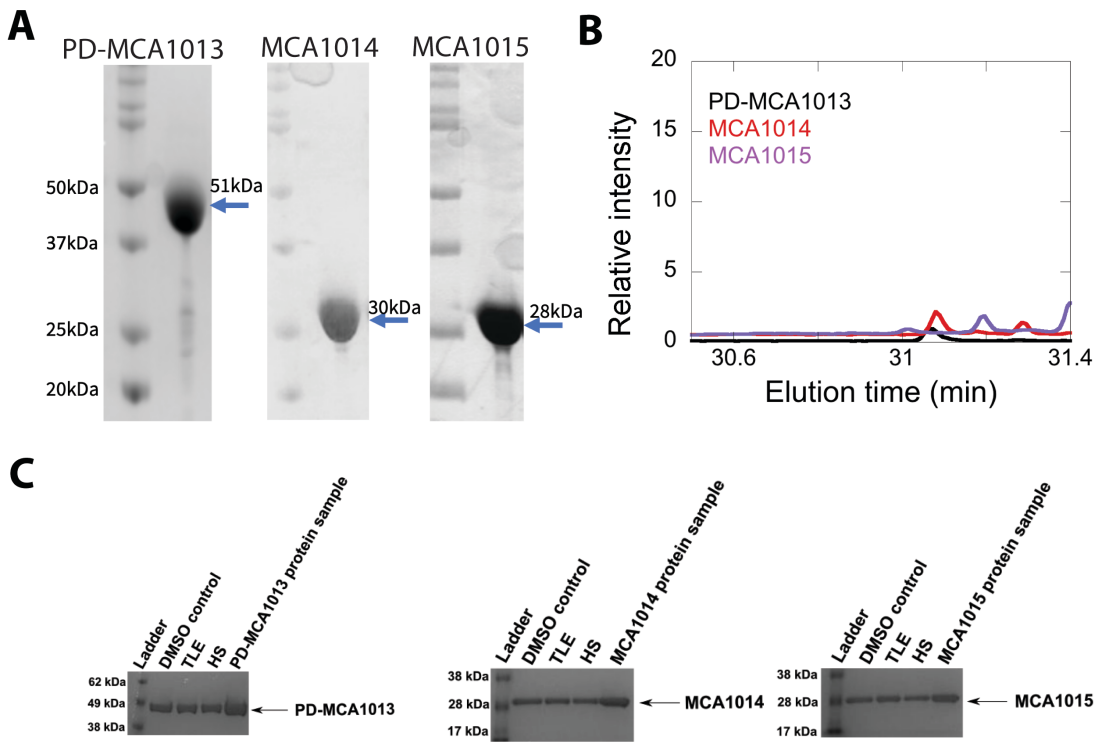


Figure S4. (**A**) SDS-PAGE gel of pure proteins. PD-MCA1013 (lane 2), MCA1014 (lane 4) and MCA1015 (lane 6). Molecular weight markers are shown in lanes 1, 3, and 5. All proteins are estimated to be more than 90% pure. **(B)** SDS PAGE gels demonstrating that the samples analyzed in the protein-lipid pull down assays did contain PD-MCA1013, MCA1014, and MCA1015. (**C**) Control experiments with DMSO-treated samples reveal that the proteins do not contain C4-methylsterols.

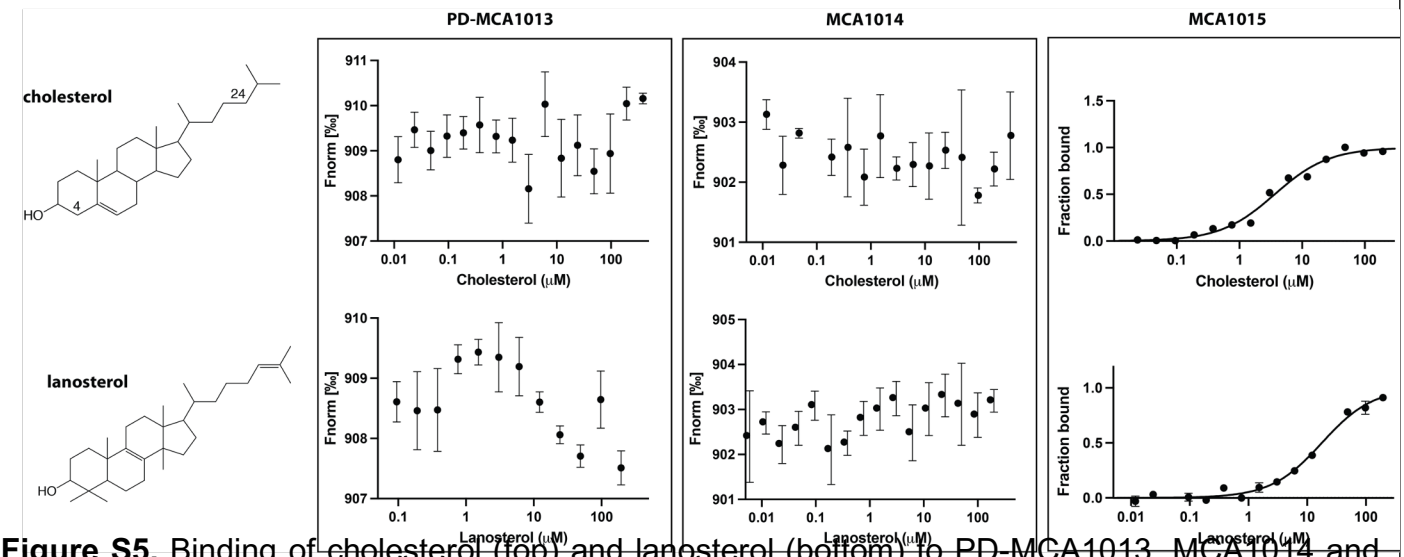


Figure S5. Binding of cholesterol (top) and lanosterol (bottom) to PD-MCA1013, MCA1014 and MCA1015. The data were acquired using the same conditions as reported for **Figure 2.**

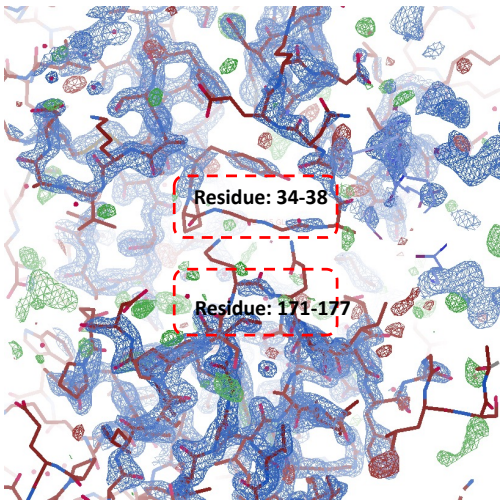


Figure S6. A portion of the 2Fo − Fc electron density omit map for MCA1014 contoured at 2 σ (blue mesh). Loops (residues 34-38 and 171-177) are shown in red boxes. The poor density of these loops could be due to protein flexibility in the solution.

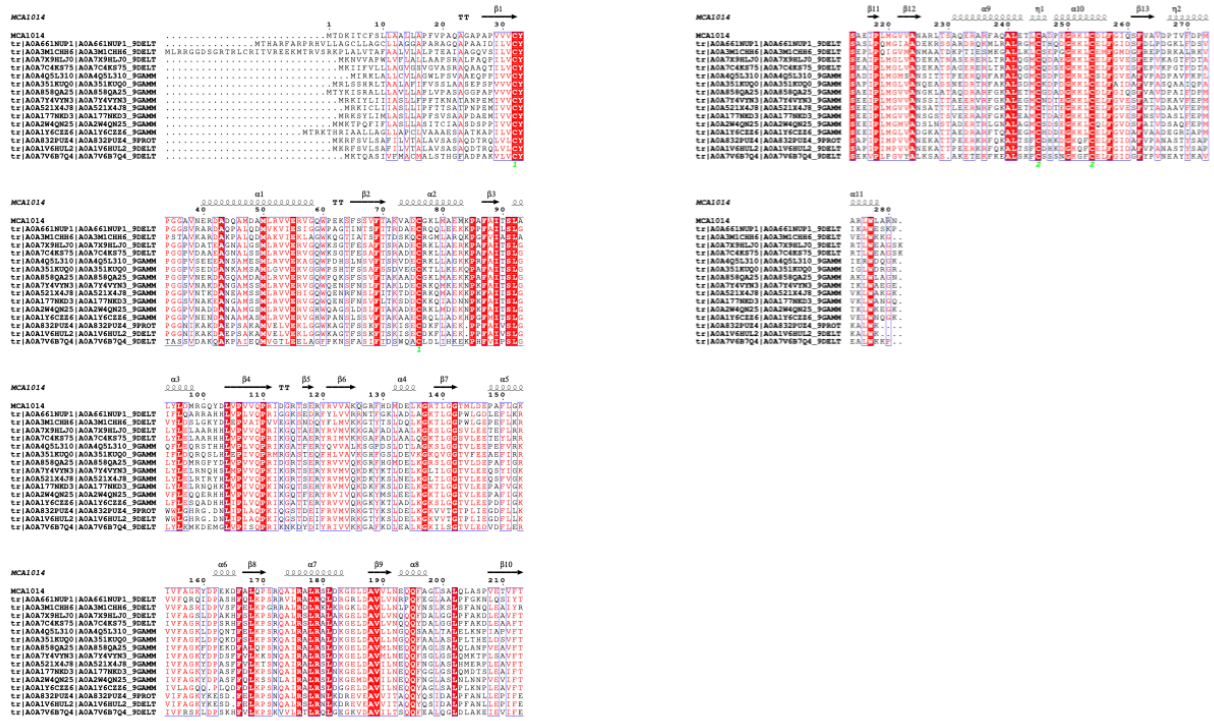
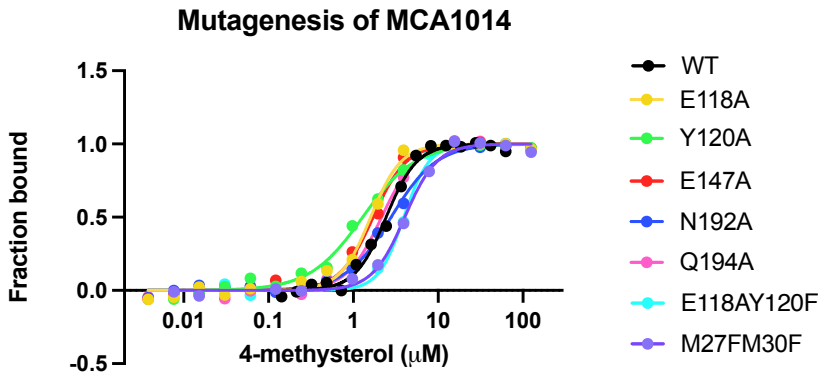


Figure S7. Amino acid sequence alignment of MCA1014. Alignment was performed using the CLUSTALW server (https://www.genome.jp/tools-bin/clustalw) and displayed using Esprit (http://espript.ibcp.fr). Secondary structure elements are shown based on the MCA1014 crystal structure. The conserved residues are highlighted in red.



	WT	E118A	Y120A	E147A	N192A	Q194A	E118A Y120F	M27F M30F
n_H	2.375	2.406	1.393	2.194	1.624	2.099	2.9	2.282
Kdave	2.435	1.601	1.287	1.715	2.789	2.157	3.934	4.041

Figure S8. MST binding curves of MCA1014 variants with 4-methylsterols.

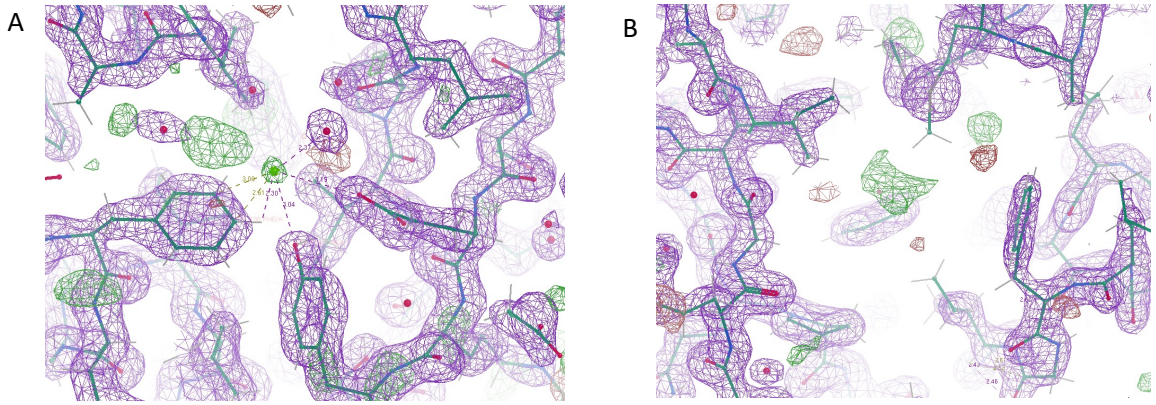


Figure S9. A portion of the 2Fo−Fc electron density omit map (purple mesh, contoured at 2s) the Fo−Fc electron density omit map (green mesh, contoured at 3σ) of MCA 1015. **(A)** The built-in chloride forms polar interactions with the surrounding Glu118, Tyr120 and Phe150. **(B)** Unmodeled density found in the cavity where the aliphatic tail of the sterol is positioned in the docking model.

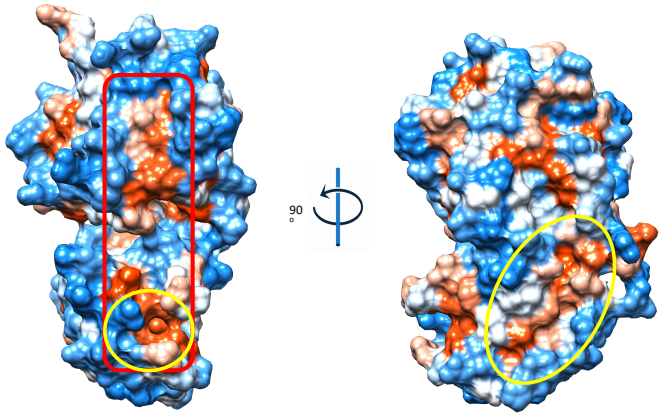


Figure S10. Hydrophobicity representation of the MCA1014 structure shows a second hydrophobic pocket located in domain B below the missing loop (202-204). The pocket (circled) has an area of 172 Å and a volume of 117 Å and may function as a second substrate binding site. The surface in red box is the potential interface after substrate binding.

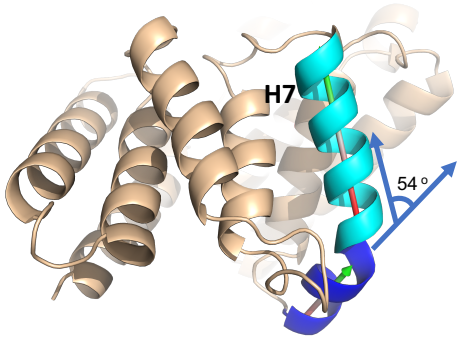


Figure S11. The H7 of MCA1015 exhibits a bent helical structure, with the lower part bent towards the C-terminal (bottom) of the protein at an angle of 54 °**.**

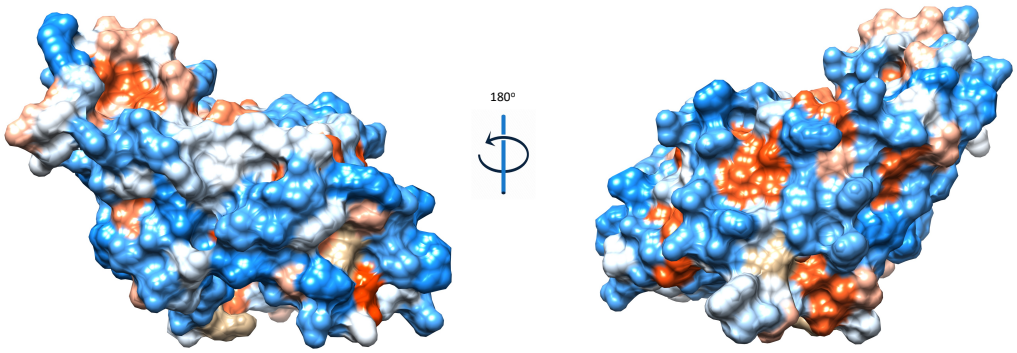


Figure S12. Hydrophobicity representation of the MCA1015 structure shows it is predominantly hydrophilic on the exterior surface.

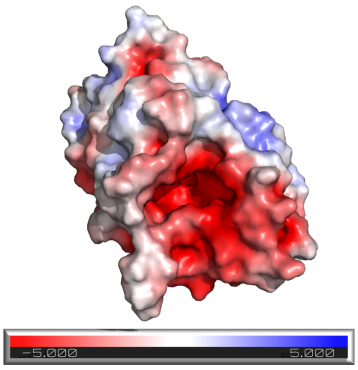


Figure S13. Electrostatic calculation shows that the C-terminal face, including the entrance to the cavities of MCA1015, exhibits strong negative charge.

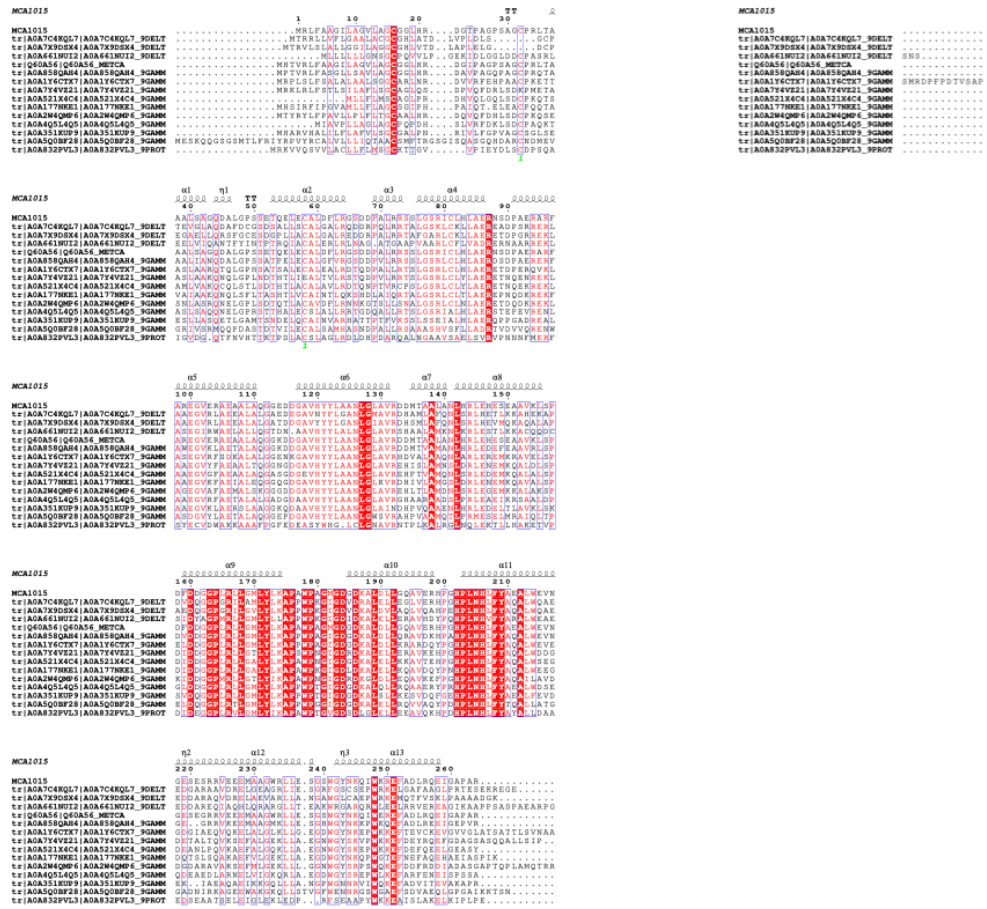


Figure S14 Amino acid sequence alignment of MCA1015. Alignment was performed using the CLUSTALW server (https://www.genome.jp/tools-bin/clustalw) and displayed using Esprit (http://espript.ibcp.fr). Secondary structure elements are shown based on the MCA1015 crystal structure. The conserved residues are highlighted in red.

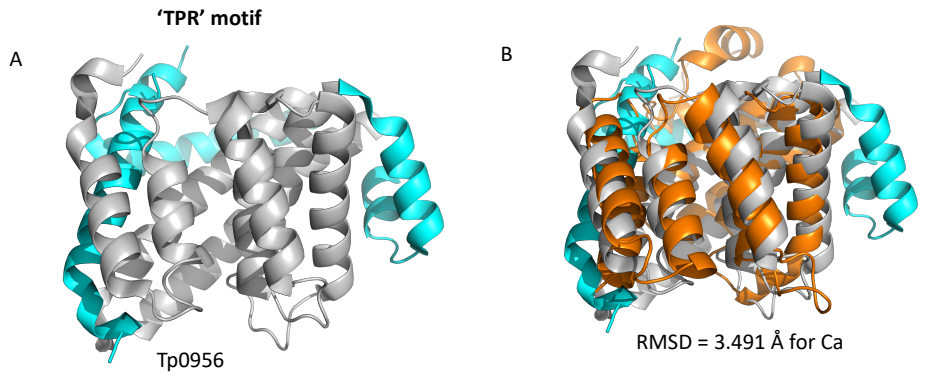


Figure S15. (**A**) Cartoon depiction of Tp0956 (PDB 3U64), the structural homolog of MCA1015. The protein contains four helical hairpins that are similar to tetratricopeptide repeat (TPR) motifs, which are hallmarks of proteins that engage in protein–protein interactions. The TPR motif was colored in grey. (**B**) Superposition of the MCA1015 structure with Tp0956 (RMSD of 3.491 Å for all Cαs). MCA1015 colored in orange.

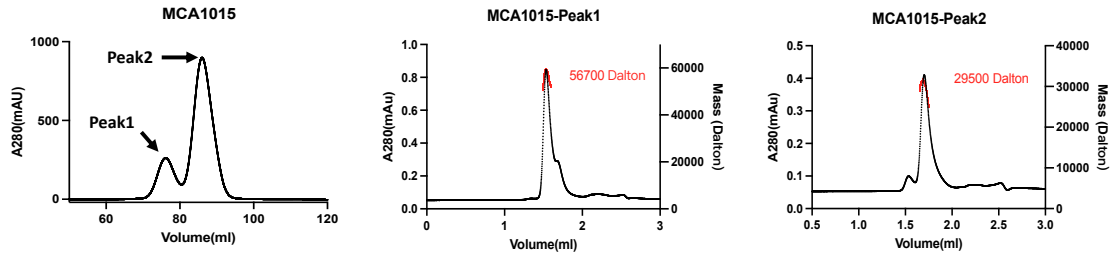


Figure S16. Size exclusion chromatography (SEC, left) and SEC-multiangle light scattering (MALS, middle and right) analysis of MCA1015. Two peaks are observed in the SEC chromatogram, with SEC-MALS revealing that they correspond to monomeric and dimeric MCA1015.

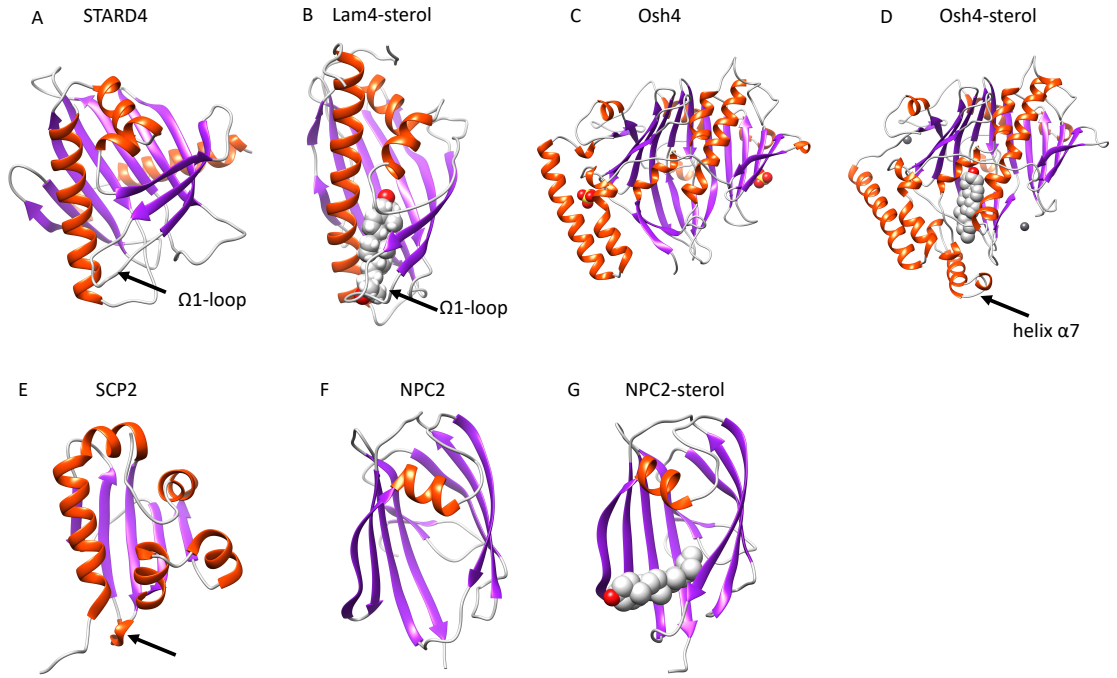


Figure S17. Structural comparison of eukaryotic sterol-binding proteins.
(A) Apo human STARD4 (PDB 6L1D). **(B)** Yeast Lam4p bind with 25-hydroxycholesterol (PDB 6BYM). The Ω1-loop is depicted in structures that possess it. **(C-D)** Yeast apo- (**C**, PDB 1ZI7) and cholesterol-bound (**D**, PDB 1ZHY) Osh4. The N-terminal helix α7 in Osh4, which is missing in the apo-Osh4 structure, may work similar to the Ω1-loop. **(E)** The non-specific LTP, rabbit SCP2 (PDB 1C44); **(F)** apo- (PDB 6R4M) and (**G**) ergosterol-bound (PDB 6R4N) NPC2. All structures are shown in cartoon representation with the α-helixes and β-strands colored in red and magenta, respectively. The sterol substrates are depicted in grey spheres with the hydroxyl group colored in red.

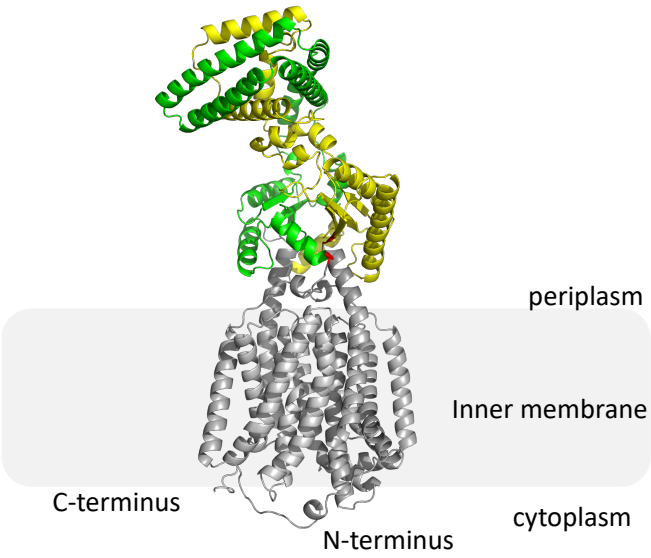


Figure S18. AlphaFold-predicted structure of MCA1013. To create PD-MCA1013, the two fragments of the periplasmic portion (colored in yellow and green) were connected by a "GSGSGSGS" linker rather than the transmembrane helix of the full-length protein.

Table S1. Data collection and refinement statistics.

	apo-MCA1014(SeMet)	apo-MCA1015(SeMet)*
Resolution range	38.4 -1.6 (1.657-1.6)	28.63 -1.91 (1.978-1.91)
Space group	P 21 21 21	P 1 21 1
Unit cell (a, b, c; α, β, γ)	39.882, 40.355, 142.263; 90, 90, 90	55.083, 57.257, 85.825; 90, 97.26, 90
Total reflections	606476 (60185)	200826 (19731)
Unique reflections	31071 (3061)	41338 (4084)
Multiplicity	19.5 (19.7)	4.9 (4.8)
Completeness (%)	97.81 (93.74)	99.93 (99.98)
Mean I/sigma(I)	30.96 (2.07)	11.09 (2.78)
Wilson B-factor	22.86	15.65
R-merge	0.07677 (1.462)	0.1267 (0.6131)
R-meas	0.07881 (1.501)	0.1423 (0.6891)
R-pim	0.01755 (0.3345)	0.06419 (0.3117)
CC1/2	1 (0.745)	0.994 (0.762)
CC*	1 (0.924)	0.999 (0.93)
Reflections used in refinement	30554 (2873)	41326 (4084)
Reflections used for R-free	1964 (184)	1996 (198)
R-work	0.2038 (0.2892)	0.1850 (0.2523)
R-free	0.2234 (0.2749)	0.2123 (0.2638)
CC(work)	0.955 (0.786)	0.948 (0.879)
CC(free)	0.948 (0.725)	0.939 (0.834)
Number of non-hydrogen atoms	2117	4058
macromolecules	1974	3546
ligands	1	298
solvent	142	388
Protein residues	257	464
RMS(bonds)	0.010	0.005
RMS(angles)	1.03	0.67
Ramachandran favored (%)	99.21	100.00
Ramachandran allowed (%)	0.79	0.00
Ramachandran outliers (%)	0.00	0.00
Rotamer outliers (%)	0.97	0.00
Clashscore	2.28	1.66
Average B-factor	33.82	19.26
macromolecules	33.84	17.79
ligands	30.00	33.53
solvent	33.49	28.15
Number of TLS groups	3	1

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

Table S2. Available structural information in the SSNs.

Superfamily	Pfam	Protein	PDB	Orgnasim
MMPL	PF03176	HpnN	5khs	*Burkholderia multivorans*
MMPL	PF03176	MmpL3	6ajf	*Mycobacterium smegmatis*
MMPL	PF03176	MmpL11	4y0l	*Mycobacterium tuberculosis*
Phosphonate-bd	PF12974	AioX	6esk	*Pseudorhizobium banfieldiae*
Phosphonate-bd	PF12974	PhnD	5lq5	*Prochlorococcus marinus*
Phosphonate-bd	PF12974	PtxB	5lv1	*Prochlorococcus marinus*
Phosphonate-bd	PF12974	PtxB	5o2k	*Pseudomonas stutzeri*
Phosphonate-bd	PF12974	PtxB	5jvb	*Trichodesmium erythraeum*
Phosphonate-bd	PF12974	PA3383	3n5l	*Pseudomonas aeruginosa*
Phosphonate-bd	PF12974	PhnD	3qk6	*Escherichia coli*
Phosphonate-bd	PF12974	XAC2383	5ub3	*Xanthomonas axonopodis pv. citri*
Phosphonate-bd	PF12974	HtxB	5me4	*Pseudomonas stutzeri*
TAtT	PF16811	Tp0956	3u64	*Treponema pallidum (strain Nichols)*

Table S3. Residues involved in forming cavity in cleft of MCA1014.

Domain	SeqID	AA		Domain	SeqID	AA
A	29	Val		A	115	Arg
A	31	Val		A	117	Ser
A	33	Tyr		A	118	Glu
A	34	Pro		B	120	Tyr
A	35	Gly		B	143	Thr
A	36	Gly		B	144	Met
A	37	Ala		B	146	Asp
A	38	Val		B	147	Glu
A	39	Asn		B	150	Phe
A	40	Glu		B	171	Ser
A	43	Ala		B	172	Arg
A	44	Asp		B	173	Gln
A	46	Ala		B	174	Ala
A	47	Met		B	175	Ile
A	48	Asp		B	192	Asn
A	49	Ala		B	193	Glu
A	50	Met		B	194	Gln
A	51	Leu		B	195	Gln
A	53	Val		B	217	Ile
A	54	Val		A	218	Pro
A	67	Ser		A	219	Leu
A	69	Phe		A	221	Gly
A	71	Ala		A	222	Val
A	73	Val		A	223	Val
A	89	Ile		A	237	phe
A	90	Thr		A	241	Leu
A	91	Ser		A	244	Leu
A	92	Leu		A	253	Leu
A	93	Ala		A	254	Cys
A	94	Leu		A	256	Leu
A	95	Tyr		A	257	Phe

A	106	Pro		A	258	Gly
A	108	Val		A	259	Ile
A	109	Gln		A	262	Phe
A	110	Pro		A	270	Phe
A	112	Ile				

Table S4. Residues involved in forming cavity in MCA1015.

Tunnel1	Chamber2	Chamber3
ASP 45	GLY 128	LEU 129
PRO 49	LEU 129	ARG 132
LEU 82	ARG 132	LEU 167
HIS 83	MET 135	MSE 170
GLU 86	ALA 138	LYS 174
TYR 121	LEU 139	LEU 207
TYR 122	LEU 142	PHE 208
LEU 129	LEU 145	GLU 211
ARG 132	LEU 167	GLU 251
ASP 160	MSE 170	ASP 254
ARG 166	LEU 171	LEU 255
LEU 167	LYS 174	
LEU 204	PHE 208	
TRP 241		
ASN 244		
TRP 248		

Table. S5 Protein sequences and primers.

NAME	SEQUENCE	PRIMERS
MCA1013	MNIPHLAALAAERFARRPWRVLALAMALSALSLWAVSRL PVHTSRQALLPHDNAVAQRFDAFLDKFGAASDLIVVLEG APPDELKPFADELATALAAEPEIAQATARLDLRFVLEHAY LAVTPERLGTLAGVLEKFGAGAIPEDSSQVDATLGRLLQW LEGAPAMPAAGIDLPTVEVGLKLLGASLDEWHRWLSAGE VPAALDWTRLLAGLGGSEIANDGYFVSRDGRMYFLFVHP ASASEDFTAIGPFVEKVRTVAADRAARARAAGRTAPKVG LTGLPAIEYEEHVSIRHDIALVVGSAAGLIVLLILVVVRSW RWALVIFVPMGLGVLWSLGLALVTIGHLTLITASFIAVLFG LGADYGIFTSARIAEERRRGKPLTEAIGAGMGASFQAVFT AGGASVVIFGALATVDFPGFSELGLVAAKGVMLILVSTWL VQPALYALLPPKLAPLPAAASAGAIEPGRMPFRGSVAVIL VAGALATAAFGIGSGYELPFDYDVLSLLPKDSESAYYQNR MVAESDYQAEVVIFTAPDLEEARRIAAEAGRLGSVAKVQ SLMDLFPPDADARALEARRIGELADDGYAVRLARLAAIGL PEGTFGRVRTILEKGGDFIDQSQELAFSAGHSGLVAALED VRGRLDAVRSAIEADPVQARERSERFFRMLLSAAERGVAL LAEWRQARPITPAQLPPALRDRFFAADGTVAVYAFPAKT VYDPANLDRLMQEIYGVSPDATGFPATHQVFSKSVVESFT HGTREAVTVCLLWLALVLRNWRGFVLASMPLLIGGGWM LGLMALCGIRYNYANIIALPLVIALAVDYGVWFSQRWFDL	pET20b-Forward: GAAGGAGATATACATATGAACATCCCC CATTTGGCCGCTC pET20b-Reverse: GTGGTGGTGGTGCTCGAGTGGTTTCTTT CTTGAATCGAGAAGC pET28a-Forward: AGAATCTTTATTTTCAGGGCCATAACAT TCCGCACCTGGCGGC pET28a-Reverse: CTCAGTGGTGGTGGTGGTGGTGCTCGAG CTTAGGTTTCTTACGGCTGTCCAG

	KDRSLTQINRVAGGVIGLAAGTELAGLGAITLANYRGVSS LGVNITVGLLCCLAATLWVAPAIGQLLDSRKKP	
PD-MCA1013	MAVAQRFDAFLDKFGAASDLIVVLEGAPPDELKPFADEL ATALAAEPEIAQATARLDLRFVLEHAYLAVTPERLGTLAG VLEKFGAGAIPEDSSQVDATLGRLLQWLEGAPAMPAAGI DLPTVEVGLKLLGASLDEWHRWLSAGEVPAALDWTRLL AGLGGSEIANDGYFVSRDGRMYFLFVHPASASEDFTAIGP FVEKVRTVAADRAARARAAGRTAPKVGLT**GSGSGSGS**ES AYYQNRMVAESDYQAEVVIFTAPDLEEARRIAAEAGRLG SVAKVQSLMDLFPPDADARALEARRIGELADDGYAVRLA RLAAIGLPEGTFGRVRTILEKGGDFIDQSQELAFSAGHSGL VAALEDVRGRLDAVRSAIEADPVQARERSERFFRMLLSA AERGVALLAEWRQARPITPAQLPPALRDRFFAADGTVAV YAFPAKTVYDPANLDRLMQEIYGV	pET28a-Forward1: CTTTATTTTCAGGGCCATATGGCGGTGG CGCAACGTTTCGAC pET28a-Reverse1: AGTACGCGCTCTCGCTACCGCTACCGCT GCCGCTACCGGTCAGACCAACTTTCGGC pET28a-Forward2: GTTGGTCTGACCGGTAGCGGCAGCGGT AGCGGTAGCGAGAGCGCGTACTATCAG AAC pET28a-Reverse2: CTCAGTGGTGGTGGTGGTGGTGCTCGAG TCACACACCATAAATTTCCTG
MCA1014	M**TDKITCFSLLAALLAPFVPAQA**GAPAPVVVCYPGGAV NERDADQAMDAMLRVVERVGQWPEKSFSSVFTAKVADC GKLMAEMKPAFAITSLALYLDMRGQYDLVPVVQPRIDGR TSERYRVVAKQGRFHDMDELKGRTLGGTMLDEPAFLGKI VFAGKYDPEKDFALQPSRQAIRALRSLDKGELDAVVLNE QQFAGLSALQLASPVETVFTSAEIPLMGVVANARLTSAQE RARFAQALETLCADPEGRKLCDLFGIQSFVAVDPTVFDPM ARLWLARN	pET20b-Forward: GAAGGAGATATACATATGGGCGCCCCC GCACCG pET20b-Reverse: GTGGTGGTGGTGCTCGAGGTTCCGCGCC AGCCA pET28a-Forward: CGAGAATCTTTATTTTCAGGGCCATGGC GCCCCCGCACCGGTGGTC pET28a-Reverse: CTCAGTGGTGGTGGTGGTGGTGCTCGAG TTAGTTCCGCGCCAGCCAGAG
MCA1015	M**RLFAAGILAGVLAG**CGGLHRDGTPAGPSAGCPRLTAA ALSAGQDALGPSSETQELECALDFLRGSDDPALRRSSLGS RICLHLAERNSDPAERARFAREGVERAEAALAQGGEDDG AVHYYLAANLGLAVRDDMTAALANLHRLEHESEAAVKL SPDFDDGGPLRLLGMLYLKAPAWPAGMGDGDKALDLLG QAVERHPGHPLNHLFYAEALWEVNGESESRRVEEEMAAG WRLLESGSWGYNKQIWKREFADLRQEIGAPAR	pET20b-Forward: GAAGGAGATATACATATGTGCGGCGGT CTTCAC pET20b-Reverse: GTGGTGGTGGTGCTCGAGCCTAGCGGG CGCCCC pET28a-Forward: ATCTTTATTTTCAGGGCCATTGCGGCGG TCTTCACCGCGATG pET28a-Reverse: CTCAGTGGTGGTGGTGGTGGTGCTCGAG TTACCTAGCGGGCGCCCCGAT

* The signal peptide in MCA1014 and MCA1015 is in bold.

References

1. K. Bloch, Chapter 12 Cholesterol: evolution of structure and function. *New Compr Biochem* 20, 363–381 (1991).

2. X. Bi, G. Liao, Cholesterol in Niemann-Pick Type C disease. *Sub-cell Biochem* 51, 319–35 (2010).

3. P. Huang, *et al.*, Cellular Cholesterol Directly Activates Smoothened in Hedgehog Signaling. *Cell* 166, 1176-1187.e14 (2016).

4. L. Miao, *et al.*, From Lanosterol to Cholesterol: Structural Evolution and Differential Effects on Lipid Bilayers. *Biophys J* 82, 1429–1444 (2002).

5. G. Ourisson, M. Rohmer, K. Poralla, Prokaryotic Hopanoids and other Polyterpenoid Sterol Surrogates. *Annu Rev Microbiol* 41, 301–333 (1987).

6. C. W. BIRD, *et al.*, Steroids and Squalene in Methylococcus capsulatus grown on Methane. *Nature* 230, 473–474 (1971).

7. W. Kohl, A. Gloe, H. Reichenbach, Steroids from the Myxobacterium Nannocystis exedens. *Microbiology*+ 129, 1629–1635 (1983).

8. T. E. Patt, R. S. Hanson, Intracytoplasmic membrane, phospholipid, and sterol content of Methylobacterium organophilum cells grown under different conditions. *J Bacteriol* 134, 636–644 (1978).

9. H. B. Bode, *et al.*, Steroid biosynthesis in prokaryotes: identification of myxobacterial steroids and cloning of the first bacterial 2,3(S)-oxidosqualene cyclase from the myxobacterium Stigmatella aurantiaca. *Mol Microbiol* 47, 471–481 (2003).

10. S. Schouten, J. P. Bowman, W. I. C. Rijpstra, J. S. S. Damsté, Sterols in a psychrophilic methanotroph, Methylosphaera hansonii. *Fems Microbiol Lett* 186, 193–195 (2000).

11. A. Pearson, M. Budin, J. J. Brocks, Phylogenetic and biochemical evidence for sterol synthesis in the bacterium Gemmata obscuriglobus. *Proc National Acad Sci* 100, 15352–15357 (2003).

12. J. H. Wei, X. Yin, P. V. Welander, Sterol Synthesis in Diverse Bacteria. *Front Microbiol* 7, 990 (2016).

13. I. Abe, Natural Products: Discourse, Diversity, and Design. 293–316 (2014).

14. E. Desmond, S. Gribaldo, Phylogenomics of Sterol Synthesis: Insights into the Origin, Evolution, and Diversity of a Key Eukaryotic Feature. *Genome Biol Evol* 1, 364–381 (2009).

15. W. D. Nes, G. G. Janssen, F. G. Crumley, M. Kalinowska, T. Akihisa, The Structural Requirements of Sterols for Membrane Function in Saccharomyces cerevisiae. *Arch Biochem Biophys* 300, 724–733 (1993).

16. F. Xu, *et al.*, Dual roles for cholesterol in mammalian cells. *Proc National Acad Sci* 102, 14551–14556 (2005).

17. P. Bouvier, M. Rohmer, P. Benveniste, G. Ourisson, Δ8(14)-steroids in the bacterium Methylococcus capsulatus. *Biochem J* 159, 267–271 (1976).

18. A. K. Lee, *et al.*, C-4 sterol demethylation enzymes distinguish bacterial and eukaryotic sterol synthesis. *Proc National Acad Sci* 115, 201802930 (2018).

19. A. Rahier, Dissecting the sterol C-4 demethylation process in higher plants. From structures and genes to catalytic mechanism. *Steroids* 76, 340–352 (2011).

20. D. Gachotte, R. Barbuch, J. Gaylor, E. Nickel, M. Bard, Characterization of the Saccharomyces cerevisiae ERG26 gene encoding the C-3 sterol dehydrogenase (C-4 decarboxylase) involved in sterol biosynthesis. *Proc National Acad Sci* 95, 13794–13799 (1998).

21. R. E. Summons, A. S. Bradley, L. L. Jahnke, J. R. Waldbauer, Steroids, triterpenoids and molecular oxygen. *Philosophical Transactions Royal Soc B Biological Sci* 361, 951–968 (2006).

22. L. W. Parks, S. J. Smith, J. H. Crowley, Biochemical and physiological effects of sterol alterations in yeast—A review. *Lipids* 30, 227–230 (1995).

23. N. Jacquier, R. Schneiter, Mechanisms of sterol uptake and transport in yeast. *J Steroid Biochem Mol Biology* 129, 70–78 (2012).

24. L. L. JAHNKE, H. STAN-LOTTER, K. KATO, L. I. HOCHSTEIN, Presence of methyl sterol and bacteriohopanepolyol in an outer-membrane preparation from Methylococcus capsulatus (Bath). *Microbiology*+ 138, 1759–1766 (1992).

25. N. Ruiz, T. Wu, D. Kahne, T. J. Silhavy, Probing the Barrier Function of the Outer Membrane with Chemical Conditionality. *Acs Chem Biol* 1, 385–395 (2006).

26. K. Bacia, P. Schwille, T. Kurzchalia, Sterol structure determines the separation of phases and the curvature of the liquid-ordered phase in model membranes. *P Natl Acad Sci Usa* 102, 3272–3277 (2005).

27. A. König, R. Happle, D. Bornholdt, H. Engel, K. Grzeschik, Mutations in the NSDHL gene, encoding a 3β-hydroxysteroid dehydrogenase, cause CHILD syndrome. *Am J Med Genet* 90, 339–346 (2000).

28. M. He, *et al.*, Mutations in the human SC4MOL gene encoding a methyl sterol oxidase cause psoriasiform dermatitis, microcephaly, and developmental delay. *J Clin Invest* 121, 976–984 (2011).

29. K. W. McLarren, *et al.*, Hypomorphic Temperature-Sensitive Alleles of NSDHL Cause CK Syndrome. *Am J Hum Genetics* 87, 905–914 (2010).

30. L. Lim, *et al.*, Lanosterol induces mitochondrial uncoupling and protects dopaminergic neurons from cell death in a model for Parkinson's disease. *Cell Death Differ* 19, 416–427 (2012).

31. A. L. Hughes, C.-Y. S. Lee, C. M. Bien, P. J. Espenshade, 4-Methyl Sterols Regulate Fission Yeast SREBP-Scap under Low Oxygen and Cell Stress*. *J Biol Chem* 282, 24388–24396 (2007).

32. N. D. Serratore, *et al.*, A Novel Sterol-Signaling Pathway Governs Azole Antifungal Drug Resistance and Hypoxic Gene Repression in Saccharomyces cerevisiae. *Genetics* 208, genetics.300554.2017 (2017).

33. B. L. Todd, E. V. Stewart, J. S. Burg, A. L. Hughes, P. J. Espenshade, Sterol Regulatory Element Binding Protein Is a Principal Regulator of Anaerobic Gene Expression in Fission Yeast†. *Mol Cell Biol* 26, 2817–2831 (2006).

34. J. E. Vance, K. B. Peake, Function of the Niemann–Pick type C proteins and their bypass by cyclodextrin. *Curr Opin Lipidol* 22, 204–209 (2011).

35. H. J. Atkinson, J. H. Morris, T. E. Ferrin, P. C. Babbitt, Using Sequence Similarity Networks for Visualization of Relationships Across Diverse Protein Superfamilies. *Plos One* 4, e4345 (2009).

36. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10, 845–858 (2015).

37. J. Yang, *et al.*, The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 12, 7–8 (2015).

38. D. M. Doughty, *et al.*, The RND-family transporter, HpnN, is required for hopanoid localization to the outer membrane of Rhodopseudomonas palustris TIE-1. *P Natl Acad Sci Usa* 108, E1045-51 (2011).

39. I. Alicea, *et al.*, Structure of the Escherichia coli Phosphonate Binding Protein PhnD and Rationally Optimized Phosphonate Biosensors. *J Mol Biol* 414, 356–369 (2011).

40. L. Holm, P. Rosenstrï¿½m, Dali server: conservation mapping in 3D. *Nucleic Acids Res* 38, W545–W549 (2010).

41. F. A. Quiocho, P. S. Ledvina, Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol Microbiol* 20, 17–25 (1996).

42. M. de Boer, *et al.*, Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. *Elife* 8, e44652 (2019).

43. O. Trott, A. J. Olson, AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31, 455–461 (2010).

44. R. K. Deka, *et al.*, Structural, bioinformatic, and in vivo analyses of two Treponema pallidum lipoproteins reveal a unique TRAP transporter. *J Mol Biol* 416, 678–96 (2012).

45. C. A. Brautigam, R. K. Deka, P. Schuck, D. R. Tomchick, M. V. Norgard, Structural and Thermodynamic Characterization of the Interaction between Two Periplasmic Treponema pallidum Lipoproteins that are Components of a TPR-Protein-Associated TRAP Transporter (TPAT). *J Mol Biol* 420, 70–86 (2012).

46. L. D. D'Andrea, L. Regan, TPR proteins: the versatile helix. *Trends Biochem Sci* 28, 655–662 (2003).

47. N. A. Baumann, *et al.*, Transport of Newly Synthesized Sterol to the Sterol-Enriched Plasma Membrane Occurs via Nonvesicular Equilibration †. *Biochemistry-us* 44, 5816–5826 (2005).

48. A. T. Gatta, *et al.*, A new family of StART domain proteins at membrane contact sites has a role in ER-PM sterol transport. *Elife* 4, e07253 (2015).

49. S. Raychaudhuri, W. A. Prinz, The Diverse Functions of Oxysterol-Binding Proteins. *Annu Rev Cell Dev Bi* 26, 157–177 (2010).

50. E. D. Carstea, *et al.*, Niemann-Pick C1 Disease Gene: Homology to Mediators of Cholesterol Homeostasis. *Science* 277, 228–231 (1997).

51. C. F. Higgins, ABC Transporters: From Microorganisms to Man. *Annu Rev Cell Biol* 8, 67–113 (1992).

52. A. L. Davidson, E. Dassa, C. Orelle, J. Chen, Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiol Mol Biol R* 72, 317–364 (2008).

53. J. A. Gerlt, *et al.*, Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochimica Et Biophysica Acta Bba - Proteins Proteom* 1854, 1019–1037 (2015).

54. J. D. Radolf, *et al.*, Treponema pallidum, the syphilis spirochete: making a living as a stealth pathogen. *Nat Rev Microbiol* 14, 744–759 (2016).

55. P. Shannon, *et al.*, Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* 13, 2498–2504 (2003).

56. J. J. A. Armenteros, *et al.*, SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37, 420–423 (2019).

57. J. Jumper, *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589 (2021).

58. D. C. Ekiert, *et al.*, Architectures of Lipid Transport Systems for the Bacterial Outer Membrane. *Cell* 169, 273-285.e17 (2017).

59. W. Kabsch, Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr Sect D Biological Crystallogr* 66, 133–144 (2010).

60. W. Kabsch, XDS. *Acta Crystallogr Sect D Biological Crystallogr* 66, 125–132 (2010).

61. W. Minor, M. Cymborowski, Z. Otwinowski, M. Chruszcz, HKL-3000: the integration of data reduction and structure solution – from diffraction images to an initial model in minutes. *Acta Crystallogr Sect D Biological Crystallogr* 62, 859–866 (2006).

62. D. Liebschner, *et al.*, Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr Sect D Struct Biology* 75, 861–877 (2019).

63. S. L, D. W., PyMOL (2020).

64. E. F. Pettersen, *et al.*, UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem* 25, 1605–1612 (2004).

65. W. Tian, C. Chen, X. Lei, J. Zhao, J. Liang, CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res* 46, gky473- (2018).

66. L. Pravda, *et al.*, MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update). *Nucleic Acids Res* 46, W368–W373 (2018).