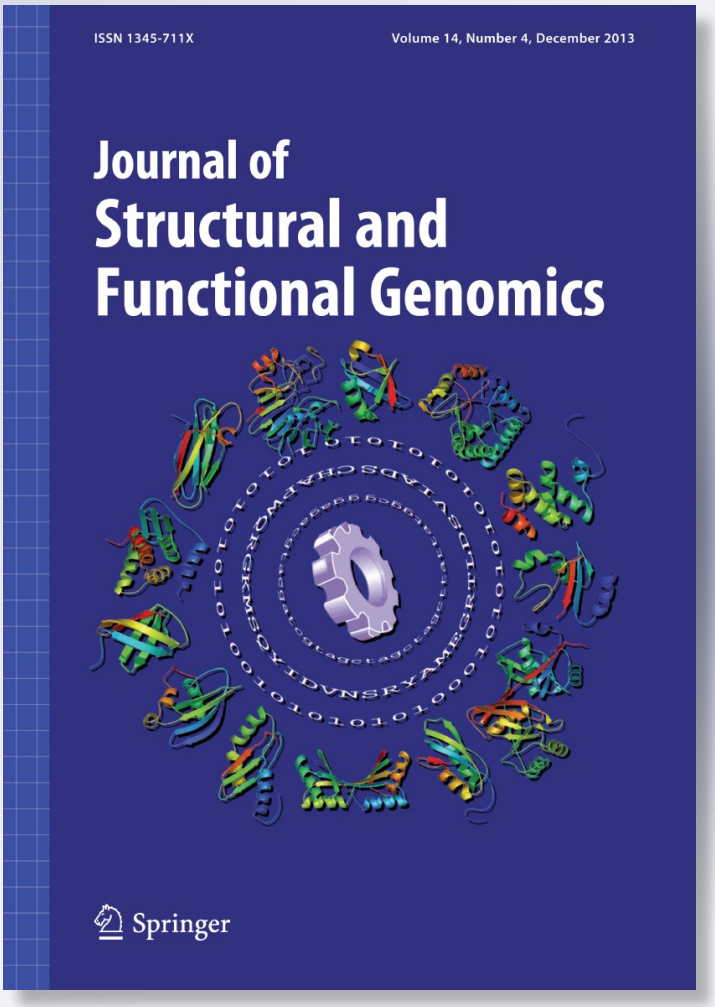
The crystal structure of sterol carrier protein 2 from Yarrowia lipolytica *and the evolutionary conservation of a large, non-specific lipid-binding cavity*

Federico Perez De Berti, Stefano Capaldi, Raúl Ferreyra, Noelia Burgardt, Juan P. Acierno, Sebastián Klinke, Hugo L. Monaco, et al.

Journal of Structural and Functional Genomics

ISSN 1345-711X
Volume 14
Number 4

J Struct Funct Genomics (2013)
14:145-153
DOI 10.1007/s10969-013-9166-6



Springer

J Struct Funct Genomics (2013) 14:145–153
DOI 10.1007/s10969-013-9166-6

The crystal structure of sterol carrier protein 2 from *Yarrowia lipolytica* and the evolutionary conservation of a large, non-specific lipid-binding cavity

Federico Perez De Berti · Stefano Capaldi · Raúl Ferreyra · Noelia Burgardt · Juan P. Acierno · Sebastián Klinke · Hugo L. Monaco · Mario R. Ermácora

Received: 7 September 2013 / Accepted: 11 November 2013 / Published online: 17 November 2013

Abstract Sterol carrier protein 2 (SCP2), a small intracellular domain present in all forms of life, binds with high affinity a broad spectrum of lipids. Due to its involvement in the metabolism of long-chain fatty acids and cholesterol uptake, it has been the focus of intense research in mammals and insects; much less characterized are SCP2 from other eukaryotic cells and microorganisms. We report here the X-ray structure of *Yarrowia lipolytica* SCP2 (YLSCP2) at 2.2 Å resolution in complex with palmitic acid. This is the first fungal SCP2 structure solved, and it consists of the canonical five-stranded β-sheet covered on the internal face by a layer of five α-helices. The overall fold is conserved among the SCP2 family, however, YLSCP2 is most similar to the SCP2 domain of human MFE-2, a bifunctional enzyme acting on peroxisomal β-oxidation. We have identified the common structural elements defining the shape and volume of the large binding cavity in all species characterized. Moreover, we found that the cavity of the SCP2 domains is distinctly formed by carbon atoms, containing neither organized water nor rigid polar interactions with the ligand. These features are in contrast with those of fatty acid binding proteins, whose internal cavities are more polar and contain bound water. The results will help to design experiments to unveil the SCP2 function in very different cellular contexts and metabolic conditions.

Keywords Sterol carrier protein · Non-specific lipid binding protein · Crystallography · *Yarrowia lipolytica* · Cholesterol · β-Oxidation · Peroxisomes · Fatty acids

Introduction

Sterol carrier protein 2 (SCP2) is present in bacteria, archaea, and eukarya, as an individual protein and as a

Electronic supplementary material The online version of this article (doi:10.1007/s10969-013-9166-6) contains supplementary material, which is available to authorized users.

F. P. De Berti · R. Ferreyra · N. Burgardt · J. P. Acierno · M. R. Ermácora (✉)
Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Sáenz Peña 352, B1876BXD Bernal, Buenos Aires, Argentina
e-mail: ermacora@unq.edu.ar

F. P. De Berti · R. Ferreyra · M. R. Ermácora
Instituto Multidisciplinario de Biología Celular, CONICET, Calle 526 y Camino General Belgrano, B1906APO La Plata, Buenos Aires, Argentina

S. Capaldi · H. L. Monaco
Biocrystallography Laboratory, Department of Biotechnology, University of Verona, Ca Vignal 1, Strada Le Grazie 15, 37134 Verona, Italy

N. Burgardt
Instituto de Química y Fisicoquímica Biológicas Prof. Alejandro C. Paladini, CONICET–Universidad de Buenos Aires, Junín 956, C1113AAD Ciudad Autónoma de Buenos Aires, Buenos Aires, Argentina

S. Klinke
Fundación Instituto Leloir, IIBBA, CONICET, Patricias Argentinas 435, C1405BWE Ciudad Autónoma de Buenos Aires, Argentina

subunit of several multidomain proteins. It owns its name to its first reported activity [1, 2]. However, SCP2 binds with high affinity a broad spectrum of lipids, including fatty acids, fatty acyl CoAs, phospholipids, sterols and bile salts, and it is better defined as a non-specific lipid transfer protein. Numerous physiological functions are ascribed to SCP2, including cholesterol uptake and secretion, intracellular lipid trafficking, signaling, and activity in a variety of metabolic processes [3, 4].

In eukarya, as many as six fused and unfused SCP2 variants are found in the same species, whereas in archaea and bacteria, only unfused SCP2 have been described. Typical companions of SCP2 in multidomain proteins are peroxysomal enzymes involved in the β-oxidation of lipids. These fusions are targeted to the peroxysomes by a short signal at the SCP2 C-terminus. Notably, although SCP2 domains can be found in all kinds of eukaryotic cells, there are no genes encoding SCP2 in *S. cerevisiae* or *S. pombe* [5].

Several studies addressed the structure, localization, and function of mammalian SCP2 (reviewed in [6]). In mice, abolition of SCPx—a fusion of a thiolase and SCP2—was associated with marked alterations in peroxisome proliferation, hypolipidemia, and impairment of methyl-branched fatty acylCoA catabolism. The defect causes the inefficient import of phytanoyl-CoA into peroxisomes and thiolytic cleavage of 3-ketopristanoyl-CoA [7]. Impaired catabolism of branched fatty acids was also demonstrated in mice, in which only the expression of the thiolase moiety was abolished [8].

Another multidomain peroxysomal enzyme carrying an SCP2 subunit at its C-terminus is MFE-2. This bifunctional enzyme catalyzes hydration and dehydrogenation reactions in the degradation pathway of branched-chain fatty acids and bile acid intermediates. MFE-2 deficiency in mice causes severe growth defects and the accumulation of very-long-chain fatty acid phospholipids in brain and liver, immature C27 bile acids in bile, and, after diet supplementation with phytol, pristanic and phytanic acid in liver triacylglycerols. These changes correlate with a severe impairment of peroxisomal β-oxidation of very-long-straight and branched chain fatty acids [9].

Although the structure and function of SCP2 fusions in vertebrates has been the subject of intense research, it has been difficult to establish the precise molecular function of the SCP2 domain itself, largely because its activity could not be dissociated from that of the accompanying enzymes.

In addition to fused SCP2 domains, insect genomes encode several single-domain SCP2. Deficiency in the expression of two of these single-domain SCP2 in the yellow fever mosquito, AeSCP-2 and AeSCP-2L2, has been related to abnormalities in cholesterol and fatty acid uptake, respectively [10, 11]. Interestingly, the relative binding affinities for cholesterol and fatty acids of AeSCP-2 and AeSCP-2L2 are in correspondence with the proposed role in the uptake of each kind of ligand. Even more interesting, these SCP2 variants lack peroxisomal localization and may represent an expansion of the SCP2 domain to adopt a more general role as a lipid carrier.

In plants, SCP2 domains are stand-alone proteins [12]. Deficiency in the sole SCP2 domain present in *A. thaliana* genome causes altered seed morphology, a delayed germination, and the need for an exogenous carbon source for proper seedling establishment. *A. thaliana* SCP2 localizes to peroxisomes through its C-terminal PST1 targeting signal, and it has lipid transfer activity in vitro [13].

In fungi, SCP2 is frequently found as a single-domain protein. Although several fungal SCP2 are known [5], their structure, function and biochemical and biophysical properties are poorly understood. Binding of cis-parinaric acid and palmitoyl-CoA to SCP2 from the yeast *Yarrowia lipolytica* (YLSCP2) was characterized, and the biophysical properties of that protein reported [14, 15]. The preferential localization in peroxisomal fractions and the mechanism for fatty acids transfer to membranes of YLSCP2 was also reported [16].

SCP2 domains from archaea to bacteria are single-domain proteins. Their study is just beginning and very little is known about their function [17].

There are a number of general questions regarding SCP2 that warrant further research of this protein across all forms of life. The first one is if this is a particularly successful scaffold domain for which nature evolves ancillary uses in different contexts or if it has a well-defined, specific function—traceable to the atomic details of its binding site or interactions with other proteins or membranes—that can explain its frequent association with the lipid metabolism and or uptake. To expand the current knowledge of SCP2 structure and function, we report here the first three-dimensional structure of a fungal SCP2 domain, YLSCP2 and we compare it to mammalian, insect, bacterial and archaeal SCP2.

Materials and methods

Miscellanea

YLSCP2 concentration was determined by UV absorption at 280 nm using the published extinction coefficient (6,986 M^{-1} cm^{-1}) [15]. Molecular visualization and simple calculations were performed with VMD [18]. Pockets and cavities properties were calculated using CASTp [19]. Electrostatic potentials were calculated using APBS [20] and PDB2PQR [21]. Residue pK_a were calculated with PROPKA3.0 [22].

Protein expression and purification

Recombinant YLSCP2 was expressed in *Escherichia coli* BL21 DE3 cells harboring pYLSCP2 [15]. Protein expression was induced with 1 mM IPTG, and YLSCP2 purification was performed as described previously [15], with a few medications. Briefly, after a 3-h induction, cells were harvested by centrifugation, suspended in 10 ml of lysis buffer (50 mM Tris-HCl, 100 mM sodium chloride, 1.0 mM EDTA, pH 8.0), and disrupted by pressure (1,000 psi; French Pressure Cell Press; Thermo IEC, Needham Heights, MA). The inclusion bodies isolated by centrifugation were solubilized in 8 M urea, 25 mM sodium acetate, 10 mM glycine, pH 5.5. The solution was loaded into an SP Sepharose Fast-Flow column (1.5 cm × 3.0 cm; GE Healthcare, Piscataway Township, NJ) equilibrated with 8 M urea, 25 mM sodium acetate, 10 mM glycine, pH 4.7, and elution was performed with a 200-ml linear gradient from 0 to 500 mM sodium chloride in the equilibration buffer. Fractions containing YLSCP2 were pooled and subjected to refolding by dialysis (16 h; 5 °C) against 1,000 volumes of 50 mM sodium phosphate, pH 7.0. After removing particulate matter by centrifugation, the protein was further purified by size exclusion chromatography using a Superdex G75 HR 10/30 (Amersham Bioscience, Piscataway, NJ) column and 20 mM sodium phosphate, pH 7.0 as elution buffer. Finally, the pure protein was concentrated using a filter CENTRIPREP (Millipore, Bedford, MA) with a 10 kDa cut-off.

Crystallization, X-ray diffraction and data analysis

Crystals of YLSCP2 were obtained after 4 weeks at 20 °C with the hanging drop method. The reservoir solution (500 µl) was 1.4 M sodium citrate, 100 mM HEPES, pH 6.5. The drop (2 µl) was a 1:1 blend of reservoir and protein solutions (10 mg/ml in 20 mM sodium phosphate, 10 mM palmitic acid, 10 % ethanol, pH 7.0). Diffraction data were collected at 100 K with an ADSC Quantum 270 CCD detector at the X6A beam line at the National Synchrotron Light Source, Brookhaven National Laboratory. Data processing and reduction were carried out with the programs MOSFLM and Scala (CCP4 suite [23, 24]). Relevant data collection parameters are given in Table 1.

The structure was solved by molecular replacement with MOLREP [25], using the MFE–2 SCP2 domain (1 ikt.pdb; [26]) coordinates as the search probe. The model was rebuilt using the program Coot [27] and finally subjected to a round of TLS refinement using REFMAC [28]. The final R_{factor} was 18.7 % and the R_{free} 21.7 %. Two palmitate, two citrate, and 75 water molecules were included in the

Table 1 X-ray statistics

Data collection[a]	
Wavelength (Å)	0.9795
Space group	$P3_121$
Unit cell parameters (Å)	
$a = b$	67.35
c	117.02
Resolution (Å)	30.0–2.20 (2.32–2.20)[b]
Unique reflections	16,186 (586)
R^{c}_{merge}	0.068 (0.453)
$\langle I \rangle / \langle \sigma_I \rangle$	23.1 (5.4)
Completeness (%)	99.9 (100.0)
Redundancy	10.0 (9.8)
Wilson B factor (Å^2)	35.2
Matthews coeff. (Å^3/Da)	2.73
Solvent fraction (%)	54.9
Copies/ASU	2
Refinement	
Resolution (Å)	29.19–2.20
Reflections	15,538
R^{d}_{work}	0.187
R^{e}_{free}	0.217
Protein atoms	1,883
Ligand atoms	62
Water molecules	75
Geometry	
Average isotropic *B*-factors (Å^2)	
All protein atoms	40.7
Main chain atoms	39.5
Palmitic acid atoms	68.8
Water atoms	41.7
Bond *RMSD*	
Lengths (Å)	0.008
Angles (°)	1.277
Ramachandran plot[f] (%)	
Most favoured	91.7
Additionally allowed	7.8
Generously allowed	0.5
Outliers	0.0
RCSB deposition	
PDB code	4JGX

[a] Data were collected on a single crystal

[b] Values in parentheses are for the highest-resolution shell

[c] $R_{\text{merge}} = \sum_{\text{hkl}} \sum_{\text{i}} |I_{\text{hkl,i}} - \langle I_{\text{hkl}} \rangle| / \sum_{\text{hkl}} \sum_{\text{i}} I_{\text{hkl,i}}$

[d] $R_{\text{work}} = \sum |F_{\text{obs}} - F_{\text{calc}}| \sum |F_{\text{obs}}|$, where F_{calc} and F_{obs} are the calculated and observed structure factor amplitudes, respectively

[e] R_{free} is the same as R_{work}, but calculated for 5.0 % of the total reflections (chosen at random and omitted during refinement)

[f] Corresponding to a single residue, Asn 60, in both chains and as defined by PROCHECK [29]

final model, validated using PROCHECK [29], to account for the observed electron density. Further statistics on the refinement process can be found in Table 1.

Results

Structure determination

The structure of YLSCP2 in complex with palmitic acid was solved at 2.2 Å resolution by molecular replacement techniques. YLSCP2 crystals belong to the space group $P3_1$ 2 1, and the asymmetric unit contains a homodimer (Fig. 1). Each monomer is an spheroid with overall dimensions 40 × 30 × 30 $Å^3$. The two monomers exhibit nearly identical backbone structure for residues 1–109 (RMSD = 0.54 Å). Residues 110–119 are α-helical in both monomers but differ in orientation, and the C-terminal residues 120–128 have well defined electron density only in chain B.

According to the PISA program [30], the asymmetric unit homodimer is stable in solution. The molecular surface buried upon formation of the complex is 704 $Å^2$ per subunit (10 % of the total area of the complex) with an estimated free energy of dissociation of 1.4 kcal mol^{-1}. However, the homodimer is uncommon because it lacks the typically observed C2 symmetry (i.e., the pattern of interacting residues is different in each subunit). For identical interacting monomers, lack of C2 symmetry enables the formation of open chains. It is not the case for YLSCP2 because the conformational differences mentioned above between chains A and B result in the formation of only one stable interface per homodimer. Instead, a second kind of interface, predicted as unstable by PISA, is established between homodimers along a 3_1 screw axis of the unit cell (see Fig. 1 in online resource).

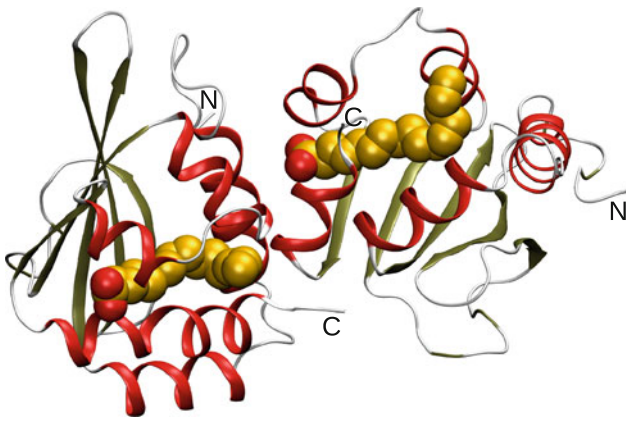


Fig. 1 The asymmetric unit of YLSCP2. The two protein chains are colored according to the secondary structure, and the two liganded molecules of palmitate are shown as space-filling models. N- and C-termini are indicated

YLSCP2 fold

YLSCP2 domain belongs to the α+β class and to the SCP-like fold (SCOP 1.75 release; [31]). The fold consists of a five-stranded β-sheet covered on the internal face by a layer of five α-helices. The external, solvent-exposed face is traversed by a crossover loop (Fig. 2). The β-sheet exhibits strand order 32,145, with all strands antiparallel except 1 and 4. The α-helices and the internal face of the β-sheet form a large cavity where the ligand binds.

Based on sequence and structure similarity, YLSCP2 belongs to the SCP super family, one of the four super families with the SCP-like fold (Fig. 3). The structural superposition of characterized members of this superfamily indicates that the closest relative of YLSCP2 is the human MFE-2 SCP2 domain, the C-terminal domain of a complex bifunctional enzyme acting on the peroxisomal β-oxidation pathway for fatty acids. The backbones of YLSCP2 and MFE-2 SCP2 domains can be superposed with 0.94 Å RMSD (87 residues out of 128). However, and despite a low sequential identity, the fold of all members of the superfamily is well preserved from bacteria to mammals (Table 2).

The binding site

YLSCP2 crystallization milieu contained 5 mM palmitic acid, and a molecule of it was found bound to each chain in

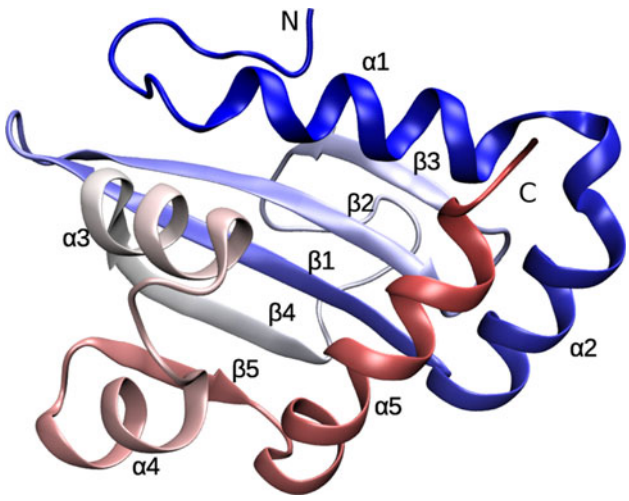


Fig. 2 Overall fold cartoon of YLSCP2. The color scale indicates the N- to C-terminal direction (*blue* to *red*, respectively). The chain represented is B, for which all residues could be modeled. In chain A, the last ten residues exhibit no defined electron density. Secondary structure elements are labeled

```
                10        20        30        40
           ----:----|----:----|----:----|----:----
                 αααααααααααα    αααααααααα 888888888
SCP2_YARLI   SLKVDGFTSSIIFDVIRDGLNDPSQAKQKAESIKKANAIIVFNLKNKAG-
DHB4_HUMAN   ------LQSTFVFEEIGRLKDI-----GPEVVKKVNAVFEWHITKGGN-
NLTP_HUMAN   -SASDGFKANLVFKEIEKKLEEE-----GEQFVKKIGGIFAFKVKDGPG-
NLTP_RABIT   SSAGDGFKANLVFKEIEKKLEEE-----GEQFVKKIGGIFAFKVKDGPGG
Q86PR3_AEDAE -GIRMSLKSDEVFAKIAKRLESI------DPANRQVEHVYKFRITQGGK-
Q0GY13_AEDAE -SPGIRMSVETIIERIKARVGAV-----DPNGPRKVLGVFQLNIKTASG-
Q16LC3_AEDAE -----ALKTDQILDKLNEKLAQV-----DRSKRSF-TVILFVHLRQEGK-
O28738_ARCFU -----MSEAKELIKKMCDLQNS---NEEIQKEMAGWSGVVQYKLD-----
Q5SL92_THET8 -----ELFTEAWAQAYCRKLNESE---AYRKAASTWEGSLALAVRPDPKA

                50        60                70        80
                 |----:----|-  ---:--     --|----:----|----:----
                 888888888     88888              888888 αααααα
SCP2_YARLI   ---KTESWYLDLKND--GDVGKG-----NKSPKGDADIQLTLSDDHFQQL
DHB4_HUMAN   ---IGAKWTIDLKSG-SGKVYQGP-------AKGAADTTIILSDEDFMEV
NLTP_HUMAN   --GKEATWVVDVKNG-KGSVLPN--------SDKKADCTITMADSDFLAL
NLTP_RABIT   ---KEATWVVDVKNG-KGSVLPN--------SDKKADCTITMADSDLLAL
Q86PR3_AEDAE ---VVKNWVMDLKN---VKLVE---------SDDAAEATLTMEDDIMFAI
Q0GY13_AEDAE ----VEQWIVDLKQ---LKVDQG--------VFASPDVTVTVGLEDMLAI
Q16LC3_AEDAE ---VVRSVVLDFNDL-KISEIELAVTSTADYPAERIDASITIDDNDFYLV
O28738_ARCFU ----GEEFYVEYKSDGTCEFKEG--------VHSSPTFTVVAPPDFWLAV
Q5SL92_THET8 GFPKGVAVVLDLWHG-ACRGAK------AVEGEAEADFVIEADLATWQEV

                      100       110       120
                |----: ----|----:----|----:----|----:----
                αα      αααααααα  88888 αααααααααααααα
SCP2_YARLI   VEGKAN-AQRLFMTGKLKVKGNVMKAAAIEGILKNAQ-NNL---
DHB4_HUMAN   VLGKLD-PQKAFFSGRLKARGNIMLSQKLQMILKDY--AKL---
NLTP_HUMAN   MTGKMN-PQSAFFQGKLKITGNMGLAMKLQNLQLQPGNAKL---
NLTP_RABIT   MTGKMN-PQSAFFQGKLKITGNMGLAMKLQNLQLQPGKAKL---
Q86PR3_AEDAE GTGALP-AKEAMAQDKMEVDGQVELIFLLEPFIASLK-------
Q0GY13_AEDAE SGKTLT-VGDALKQGKIELSGDADLAAKLAEVI-----------
Q16LC3_AEDAE ATKETS-FAALIEQGKVDITGNKQAFLTLDEKFRNK--------
O28738_ARCFU LKGQED-PVSGFMMGKYRIEGNIMEAQRLAGVIKKFQGKFEL--
Q5SL92_THET8 LEGRLEPLSALMRGLLELKKGTIAALAPYAQAAQELVKVAREVA
```

Fig. 3 Structural alignment of representative SCP superfamily members. SCP (Pfam ID: PF02036) is one of the four superfamilies sharing the SCP-like fold as defined in SCOP [31]. Equivalent backbone atoms were superimposed (Table 2). Residues that contribute atoms to the corresponding binding cavity are shown in *bold*. The *boxed* region highlights an invariant glycine and the following residue that might have a crucial structural role (see the text). The *numbering* corresponds to the YLSCP2 sequence. The secondary structure of YLSCP2 is indicated with *greek letters*. Sequences are identified with the corresponding UniProtKB ID (SCP2_YARLI: YLSCP2, 4jgx.pdb; DHB4_HUMAN: MFE-2 SCP2 domain, 1ikt.pdb; NLTP_HUMAN: NSL-TP SCP2 domain, 2c0 l.pdb; NLTP_RABIT: NSL-TP SCP2 domain, 1c44.pdb; Q86PR3_AEDAE: SCP2 from *Aedes aegypti*, 1pz4.pdb; Q0GY13_AEDAE: SCP2-L2 from *Aedes aegypti*, 2qzt.pdb; Q16LC3_AEDAE: SCP2–L3 from *Aedes aeg*ypti, 3bkr.pdb; O28738_ARCFU: SCP2 from *A. fulgidus*, 3cnu.pdb; Q5SL92_THET8: SCP2 from *Thermus thermophilus*, 2cx7.pdb)

Table 2 Structural and sequential similarity of SCP superfamily members

	RMSD[a] (Å)	Aligned (%)	Identity (%)	Similarity (%)
YLSCP2_YARLI	0.00	100	100	100
DHB4_HUMAN	0.94	69	31	59
NLTP_HUMAN	1.35	62	30	46
NLTP_RABBIT	1.27	64	30	46
Q86PR3_AEDAE	1.19	62	17	33
Q0GY13_AEDAE	1.39	51	16	33
Q16LC3_AEDAE	2.02	60	14	28
O28738_ARCFU	1.28	52	16	41
Q5SL92_THET8	2.61	51	12	27

[a] The SCP2 domains were superposed to YLSCP2 to minimize the RMSD of structurally equivalent backbone atoms. The percentage of structurally aligned atoms is given. Protein names corresponding to each Uniprot code are listed in the legend to Fig. 3. Percentages of identity and similarity are for the structural alignment shown in Fig. 3. Similarity was calculated using the BLOSUM62 matrix [37]

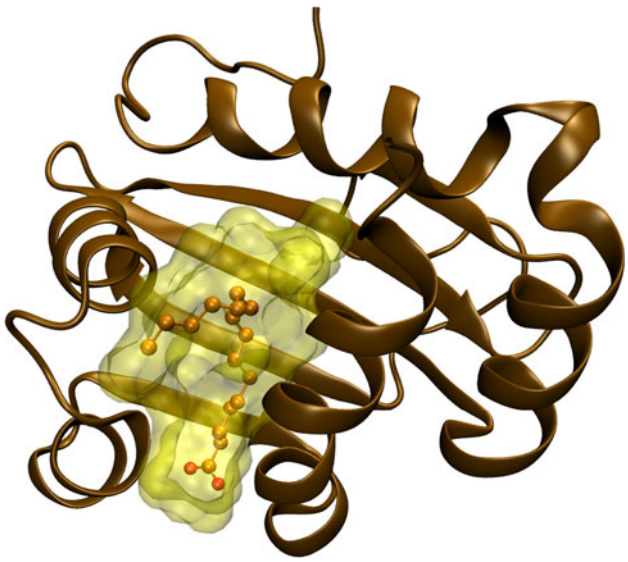


Fig. 4 The binding site of YLSCP2. The liganded palmitate is shown as a ball-and-stick model, and the van der Waals surface of the binding cavity is depicted in *yellow*

the asymmetric unit (Fig. 1). The binding site of YLSCP2 is a 19 × 18 × 10 $Å^3$, T-shaped cavity with a volume of 796 $Å^3$ and at least three putative entry–exit surface openings (Fig. 4). The internal face of the β-sheet acts as a flat hydrophobic surface on which the cavity is defined by the N- and C-terminal helices, and a helix-loop-helix motif connecting strands 4 and 5. A large fraction of the domain residues participates directly in the formation of the binding cavity (Fig. 3).

The putative portal of the binding site that accommodates the ligand carboxylate has a tunnel-like shape and fits closely the first three to four methylene groups. Farther away from the entry, the tunnel turns into a central pocket delimited by the five helices and with a connection to the surface. Thus the last portion of the ligand occupies the cavity loosely and is partially exposed to the solvent.

The binding cavity is lined almost exclusively by carbon atoms which establish only van der Waals interactions with the ligand (Fig. 2 in online resource). The electron density map of the ligand and nearby side chains is shown in Fig. 3 of the online resource. Neither organized water nor polar interactions were identified in the binding site. The oxygen

atoms of the ligand are partially buried at the bottom of an open pocket and presumably interacting with water. However, they show no convincing evidence of hydrogen bonding or ionic contacts neither with the solvent nor with other protein atoms. The ligands bound to chain A and B show a similar overall orientation in the binding site. However, the ligand in chain A is somewhat displaced toward the center of the cavity (Fig. 3 of the online resource), and in both chains, the electron density is much weaker for carbons 12–16. These relative differences and local disorder may provide entropic stabilization to a binding mode mediated mostly by hydrophobic contacts.

The cavity is only partially occupied by the fatty acid, and it can accommodate larger and/or wider ligands. To illustrate this, a molecule of phytanic acid was modeled into the binding site in the same position and orientation as the palmitic acid. The quadruply branched fatty acid, with extra methyl groups at positions 3, 7, 11 and 15, fits the binding site with no steric clashes (see Fig. 4 in online resource). In addition, the accessibility of the palmitate carboxyl group to the protein surface and the overall orientation of the ligand allow to envisage how the acyl CoA compounds might bind to YLSCP2. Indeed, a CoA moiety can be accommodated readily on the protein surface with the proper geometry to form the thioester bond (Fig. 5 in online resource). Interestingly, the phosphate groups of CoA can be located in the model in approximately the same position as a citrate anion observed in the solved structure of YLSCP2.

To compare the binding cavities of YLSCP2 and other SCP2 domains, the available X-ray structures were examined and the results are summarized in Table 3. A general trend among these structures is that water molecules and polar contacts within the binding cavity are almost completely absent.

The ligand orientation and bending in Triton X-100 liganded human MFE-2 (1ikt.pdb; [26]) are similar than in YLSCP2, and the binding site size is 40 % larger. The volume increase is caused mainly by a rigid–body, hinge-like movement of helix α5. In fact, the movement of this helix is likely to be an adjustment of the molecule to the bulkier nature of the (1,1,3,3-tetramethyl) buthyl head of Triton X-100 compared with a regular fatty acid chain. In human NSL-TP SCP2 (2c0 l.pdb; [32]), the last ten residues of the C-terminal helix α5 unwind and extend away from the body of the protein to form the interaction domain with the Pex5p receptor responsible for the translocation to the peroxisome. Concomitantly with the unwinding, the α5 helix moves as a hinge toward the center of the SCP2 domain resulting in an almost complete collapse of the binding site. In rabbit NSL-TP SCP2 (1c44.pdb; [33]) the unoccupied binding site is segmented and significantly smaller that in YLSCP2. The effect is caused by the partial unwinding of the C-terminus of helix α5, which in addition moves toward the binding site obliterating one of its ends.

Table 3 Comparison of SCP2 domain binding sites

	Ligand	Stoichiometry[a]	Volume[b] (Å^3)	PDB[c] (ID)
YLSCP2_YARLI[d]	Palmitate	1:1	796	4jgx
DHB4_HUMAN	Triton-X100	1:1	1,071	1ikt
NLTP_HUMAN	–	–	90	2c0 l
NLTP_RABBIT	–	–	575	1c44
Q86PR3_AEDAE	Palmitate	1:1	750	1pz4
Q0GY13_AEDAE	Palmitate	1–2:1	1,127	2qzt
Q16LC3_AEDAE	Palmitate	1:1	1,070	3bkr
O28738_ARCFU	–	–	271	3cnu
Q5SL92_THET8	–	–	671	2cx7

[a] Molecules of ligand per SCP2 domain

[b] Volumes of the binding cavity calculated with CASTp [19] considering only protein atoms

[c] References [10, 17, 26, 32–35]

[d] Protein names corresponding to each UniProtKD code are given in the legend to Fig. 3

The three insect SCP2 domains described so far show a characteristic unwinding of helix α2, which forms a long connecting loop to strand β1 [10, 34]. This conformational feature results in the remodeling of the binding site: one new binding chamber perpendicular to the main tunnel is created, and, as a result, the liganded fatty acid is inserted backwards, with the carboxylate pointing toward the new loop. The first insect domain, insect SCP2 (1pz4.pdb, [35]) possesses a single bound palmitate and a binding cavity with nearly the same volume as that of YLSCP2. The increase in volume produced by the unwinding of helix α2 is compensated by a movement of helix α5 toward the center of the cavity, which obliterates part of the space available in YLSCP2. The second insect domain, SCP2-L2 (2qzt.pdb, [34]), in addition to the unfolded helix α2, possesses a much shorter helix α5 (Figs. 2, 4), which results in a large volume increase. The extra space is used to accommodate a second ligand, perpendicularly oriented to the first. Moreover, since SCP2-L2 is a dimer whose interacting surface connects the binding cavities in each monomer, it might act as a larger container with up to four ligands carried simultaneously in a physically continuous deposit. The third insect SCP2 domain, SCP2-L3 (3bkr.pdb, [34]), is similar to the first, with only one palmitate moiety occupying the binding site.

Finally, the archaeal SCP2 domain from *Archaeoglobus fulgidus* displays a remarkable similarity to mammalian SCP2 domains and to YLSCP2. However, it crystallizes as a dimer with swapped α5 helices and the unoccupied binding site is reduced in volume by a movement of the latter toward the center of the molecule.

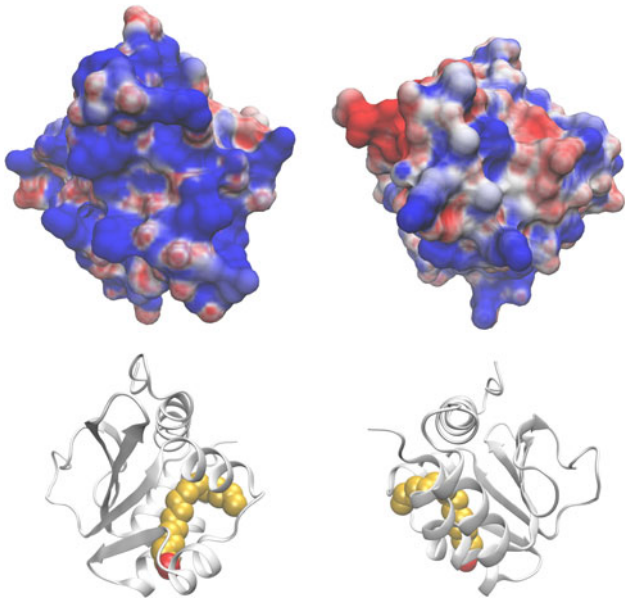


Fig. 5 A highly positive surface of YLSCP2. Above, two opposite faces (a 180° rotation about a vertical axis in the plane of the paper) of YLSCP2 colored according to the electrostatic potential (from *blue* +5 kT/e to *red*-5kT/e). Below, a cartoon representation of the above molecular orientations. The liganded palmitate is shown as van der Waals spheres

Electrostatic potential

YLSCP2 is a highly basic protein. The pI calculated using PROPKA3.0 [22] for the folded structure is 9.9. This is a consequence of the high content of basic residues since the pI calculated for the unfolded structure is nearly the same. The electrostatic potential at the protein surface calculated using APBS [20] is shown in Fig. 5. The calculation evidences that YLSCP2 possesses an unusually extended, strongly positive surface. The basic nature and asymmetry in the positive charge distribution was also observed in rabbit and human NSL-TP SCP2, and in MFE-2 SCP2 (not shown). Interestingly, both features are absent in the insect, archaeal and bacterial SCP members listed in Fig. 2, which are all acidic proteins.

Discussion

The X-ray structure reported herein significantly expands the set of characterized eukaryotic SCP2 domains by including a fungal representative, and it also adds to those from bacteria and archaea previously characterized [10, 17, 32–35]. Thus, the amount of structural information available offers an opportunity for a comparative examination of this ubiquitous protein fold and its properties in very different cellular contexts.

Regarding the quaternary structure, several of the SCP2 domains described so far showed no evidence of homomeric association. Two exceptions are *Aedes aegypti* SCP2-L2 and the archaeal SCP2 from *A. fulgidus*: the crystals of the insect protein contain a dimer with twofold symmetry in the asymmetric unit and the protein behaves as a dimer in solution [10]; the archaeal protein also associates with two fold symmetry in a very tight complex by swapping the C-terminal helices. Here, the X-ray structure of YLSCP2 reveals a stable dimer without two fold symmetry in the asymmetric unit. This confirms our previous report on the concentration-dependent tendency of YLSCP2 to form dimers and higher order aggregates in solution, as shown by SAXS and chemical cross-linking [14]. Surprisingly, unlike SCP2-L2 and the archaeal protein, the YLSCP2 homodimer lacks two fold symmetry and buries different surfaces in each subunit, which is a very uncommon way of homodimerization. Thus, the aggregation state of SCP2 domains is not a conserved feature across species and moreover a distinct behavior in terms of the particular dimer formed is observed. Since the oligomerization state certainly has an impact in the mechanisms by which the SCP2 exerts its function, it must be considered on a case-by-case basis.

The overall backbone fold of YLSCP2 is nearly identical to that of human MFE-2 SCP2. However, all characterized SCP2 domains have very similar folds and differ only in specific details. YLSCP2, as well as the other SCP2 proteins, makes use of nearly one-third of the residues distributed along all its secondary structure elements to build a functional hydrophobic internal cavity (see Figs. 3, 4). This functional requirement is the more likely explanation for the low conformational variation of the fold: whereas for function most proteins need to preserve strictly only the 3D structure of specific backbone segments, SCP2 need to do so with nearly all of them.

The comparison of known SCP2 domains reveals subtle differences that shed light on the principles on which the lipid binding site of these proteins is built and on its function. At the backbone level, the entire β-sheet and the motif α3-loop-α4 (see Fig. 2) are almost identical in all the structures compared. On the other hand, helices α1, α2 and α5 concentrate the variant–dependent conformational fluctuations that shape and define the characteristics of the binding site. Particularly interesting are the changes in α2, which in insects is distinctly unwound to create a new internal chamber for the ligand, and helix α5 whose position and length determines the volume and compartmentalization of the binding site. Moreover, helix α5 seems to have multiple and crucial roles in SCP2: (a) it is the interaction domain in the complex between NS-LTP SCP2 and the import receptor peroxin Pex5p, responsible for the

translocation of SCP2 to the peroxisomal matrix [32]; and (b), as mentioned above, the C-terminal helix participates in the homodimerization of the archaeal SCP2 domain. All these roles of the C-terminal helix involve hinge movements about the loop connecting β4 and α5. Interestingly, this loop is formed by highly conserved residues, specially a glycine that is invariant in all SCP2 domains (see Fig. 3).

The volume of the central cavity of the SCP2 domain is its most important structural feature. From a protein folding point of view, this poses a fundamental question: how to conciliate the need of well packed hydrophobic core—the hallmark of folded proteins—with the functional need of a large internal void? In this regard, SCP2 adopts a radically different structural logic compared with other lipid binding proteins. Fatty acid binding proteins for instance, with cavities of similar size to those in SCP2, have an internal surface that is partially hydrophilic and filled with a large number of organized water molecules [36]. Internal water is almost absent in all SCP2 structures reported so far and the atoms lining the cavities are overwhelmingly carbons. Thus, whereas fatty acid binding proteins seem to stabilize its internal surface by mimicking as much as possible a regular protein surface, SCP2 domains do it differently. The common mechanisms for the stabilization is not obvious from the comparison of the apo forms of SCP2. However, the apo forms tend to have a significantly reduced cavity. Moreover, the cavity is almost eliminated in the human SCP2 bound to Pex5p. This cavity contractions are mediated by hinge-like rigid-body movements of helices α1, α2 and α5 and likely suffice to lower the energy of the unliganded structures to acceptable levels for proper folding.

Unlike most protein-ligand interactions that rely heavily on elaborated directional and polar contacts, binding of hydrocarbon chains to SCP2 is mediated mostly by carbon–carbon van der Waals contacts. These kind of interactions are much less restrictive in geometrical terms, and thus, the hydrophobic ligand lies loosely in the cavity and can even adopt different orientations. Rather than a lock for a key, the binding sites of SCP2 seem more like hydrophobic containers, ample enough to accommodate the ligand in a variety of positions and orientations. In this regard, the information provided by the X-ray structure of YLSCP2 is of central importance. The only previous structures with a physiological ligand were those of insect SCP2. And in these, at least two different orientations of the ligand were observed. One of these orientations was found in YLSCP2 as well, and it was in agreement with the position adopted by a Triton molecule bound to mammalian MFE-2. Thus, all the different orientations were confirmed in more than one structure and must be considered relevant.

The lack of geometric constraints in the binding site is consistent with the low binding specificity of SCP2. In turn, lack of restrictive geometry in the binding site means fewer restrictions at the sequence level. Indeed, the only restriction for recruiting the residues that conform the binding site seems to be the hydrophobic nature of the side chain and compensated volume changes. No specific residue is conserved in the binding site of these domains to which a defined structural role could be ascribed.

In summary, we report here the X-ray structure of the first fungal SCP2 domain. Our results confirm that the SCP2 fold is very well preserved in all forms of life, and that its main structural feature is a very large internal cavity built by a significant fraction of the residues. The size and shape of the cavity are modulated by specific backbone fluctuations that can expand or obliterate the void, likely to permit binding of hydrophobic ligands with low selectivity and the stability of the unliganded form. Distinctly, the binding cavity of the SCP2 domains is lined mostly by carbon atoms, with neither polar interactions nor crystallographically-detectable water, which allows considerable latitude in the orientation of the bound molecules. Moreover, no common functional residues could be identified examining all available liganded structures, which highlights the lack of strong constraints in the geometry of the protein-ligand interaction. The general characterization presented here will help to design future investigations aimed to address the structure-function relationship of this important metabolic factor in all living organisms.

Acknowledgments This work was supported by grants from CONICET, UNQ, ANPCyT. We acknowledge the access to the X6A beam line, funded by the National Institute of General Medical Sciences, National Institute of Health under agreement GM-0080. The National Synchrotron Light Source, Brookhaven National Laboratory is supported by the U.S. Department of Energy under contract DE-AC02-98CH10886. We are grateful to Edwin Lazo for X-ray data collection.

References

1. Gallegos AM, Schoer JK, Starodub O, Kier AB, Billheimer JT, Schroeder F (2000) A potential role for sterol carrier protein-2 in cholesterol transfer to mitochondria. Chem Phys Lipids 105:9–29
2. Puglielli L, Rigotti A, Greco AV, Santos MJ, Nervi F (1995) Sterol carrier protein-2 is involved in cholesterol transfer from the endoplasmic reticulum to the plasma membrane in human fibroblasts. J Biol Chem 270:18723–18726
3. Gallegos AM, Atshaves BP, Storey SM, Starodub O, Petrescu AD, Huang H, McIntosh AL, Martin GG, Chao H, Kier AB, Schroeder F (2001) Gene structure, intracellular localization, and functional roles of sterol carrier protein-2. Prog Lipid Res 40:498–563
4. Stolowich NJ, Petrescu AD, Huang H, Martin GG, Scott AI, Schroeder F (2002) Sterol carrier protein-2: structure reveals function. Cell Mol Life Sci 59:193–212
5. Edqvist J, Blomqvist K (2006) Fusion and fission, the evolution of sterol carrier protein-2. J Mol Evol 62:292–306

6. Schroeder F, Atshaves BP, McIntosh AL, Gallegos AM, Storey SM, Parr RD, Jefferson JR, Ball JM, Kier AB (2007) Sterol carrier protein-2: new roles in regulating lipid rafts and signaling. Biochim Biophys Acta 1771:700–718
7. Seedorf U, Raabe M, Ellinghaus P, Kannenberg F, Fobker M, Engel T, Denis S, Wouters F, Wirtz KW, Wanders RJ, Maeda N, Assmann G (1998) Defective peroxisomal catabolism of branched fatty acyl coenzyme A in mice lacking the sterol carrier protein-2/sterol carrier protein-x gene function. Genes Dev 12:1189–1201
8. Atshaves BP, McIntosh AL, Landrock D, Payne HR, Mackie JT, Maeda N, Ball J, Schroeder F, Kier AB (2007) Effect of SCP-x gene ablation on branched-chain fatty acid metabolism. Am J Physiol Gastrointest Liver Physiol 292:G939–G951
9. Baes M, Huyghe S, Carmeliet P, Declercq PE, Collen D, Mannaerts GP, Van Veldhoven PP (2000) Inactivation of the peroxisomal multifunctional protein-2 in mice impedes the degradation of not only 2-methyl-branched fatty acids and bile acid intermediates but also of very long chain fatty acids. J Biol Chem 275:16329–16336
10. Dyer DH, Wessely V, Forest KT, Lan Q (2008) Three-dimensional structure/function analysis of SCP-2-like2 reveals differences among SCP-2 family members. J Lipid Res 49:644–653
11. Blitzer EJ, Vyazunova I, Lan Q (2005) Functional analysis of AeSCP-2 using gene expression knockdown in the yellow fever mosquito, *Aedes aegypti*. Insect Mol Biol 14:301–307
12. Zheng BS, Rönnberg E, Viitanen L, Salminen TA, Lundgren K, Moritz T, Edqvist J (2008) Arabidopsis sterol carrier protein-2 is required for normal development of seeds and seedlings. J Exp Bot 59:3485–3499
13. Edqvist J, Rönnberg E, Rosenquist S, Blomqvist K, Viitanen L, Salminen TA, Nylund M, Tuuf J, Mattjus P (2004) Plants express a lipid transfer protein with high similarity to mammalian sterol carrier protein-2. J Biol Chem 279:53544–53553
14. Burgardt NI, Ferreyra RG, Falomir-Lockhart L, Córsico B, Ermácora MR, Ceolín M (2009) Biophysical characterisation and urea-induced unfolding of recombinant *Yarrowia lipolytica* sterol carrier protein-2. Biochim Biophys Acta 1794:1115–1122
15. Ferreyra RG, Burgardt NI, Milikowski D, Melen G, Kornblihtt AR, Angelica ECD, Santomé JA, Ermácora MR (2006) A yeast sterol carrier protein with fatty-acid and fatty-acyl-CoA binding activity. Arch Biochem Biophys 453:197–206
16. Lockhart LJF, Burgardt NI, Ferreyra RG, Ceolin M, Ermácora MR, Córsico B (2009) Fatty acid transfer from *Yarrowia lipolytica* sterol carrier protein 2 to phospholipid membranes. Biophys J 97:248–256
17. Goroncy AK, Murayama K, Shirouzu M, Kuramitsu S, Kigawa T, Yokoyama S (2010) NMR and X-ray structures of the putative sterol carrier protein 2 from *Thermus thermophilus* HB8 show conformational changes. J Struct Funct Genomics 11:247–256
18. Humprey W, Dalke A, Schulten K (1996) VMD-Visual Molecular dynamics. J Mol Graph 14:33–38
19. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res 34:W116–W118
20. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci USA 98:10037–10041
21. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res 35:W522–W525
22. Li H, Robertson AD, Jensen JH (2005) Very fast empirical prediction and rationalization of protein p*K*a values. Proteins 61:704–721
23. Leslie AGW, Powell HR (2007) Processing diffraction data with mosflm, in evolving methods for macromolecular crystallography. Springer, Berlin, p 41
24. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr D67:235–242
25. Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. Acta Crystallogr D Biol Crystallogr 66:22–25
26. Haapalainen AM, van Aalten DM, Meriläinen G, Jalonen JE, Pirilä P, Wierenga RK, Hiltunen JK, Glumoff T (2001) Crystal structure of the liganded SCP-2-like domain of human peroxisomal multifunctional enzyme type 2 at 1.75 A resolution. J Mol Biol 313:1127–1138
27. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126–2132
28. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol Crystallogr 67:355–367
29. Laskowski RA, Moss DS, Thornton JM (1993) Main-chain bond lengths and bond angles in protein structures. J Mol Biol 231:1049–1067
30. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774–797
31. Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247:536–540
32. Stanley WA, Filipp FV, Kursula P, Schüller N, Erdmann R, Schliebs W, Sattler M, Wilmanns M (2006) Recognition of a functional peroxisome type 1 target by the dynamic import receptor pex5p. Mol Cell 24:653–663
33. Choinowski T, Hauser H, Piontek K (2000) Structure of sterol carrier protein 2 at 1.8 A resolution reveals a hydrophobic tunnel suitable for lipid binding. Biochemistry 39:1897–1902
34. Dyer DH, Vyazunova I, Lorch JM, Forest KT, Lan Q (2009) Characterization of the yellow fever mosquito sterol carrier protein-2 like 3 gene and ligand-bound protein structure. Mol Cell Biochem 326:67–77
35. Dyer DH, Lovell S, Thoden JB, Holden HM, Rayment I, Lan Q (2003) The structural determination of an insect sterol carrier protein-2 with a ligand-bound C16 fatty acid at 1.35-A resolution. J Biol Chem 278:39085–39091
36. Lücke C, Huang S, Rademacher M, Rüterjans H (2002) New insights into intracellular lipid binding proteins: the role of buried water. Protein Sci 11:2382–2392
37. Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 89:10915–10919