Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The narrow active-site cleft of *O*-acetylserine sulfhydrylase from *Leishmania donovani* allows complex formation with serine acetyltransferases with a range of C-terminal sequences

Isha Raj, Sudhir Kumar and Samudrala Gourinath*

School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067, India

Correspondence e-mail: samudralag@yahoo.com

Received 1 March 2012
Accepted 16 April 2012

PDB References: *O*-acetylserine sulfhydrylase, 3spx; complex with octapeptide, 3t4p; complex with tetrapeptide, 3tbh.

Cysteine is a crucial substrate for the synthesis of glutathione and trypanothione, which in turn maintain intracellular redox homeostasis and defend against oxidative stress in the pathogen *Leishmania donovani*. Here, the identification, sequencing, characterization and crystal structure at 1.79 Å resolution of *O*-acetylserine sulfhydrylase (OASS), a cysteine-biosynthetic pathway enzyme from *L. donovani* (LdOASS), are reported. It shows binding to the serine acetyltransferase (SAT) C-terminal peptide, indicating that OASS and SAT interact with each other to form a cysteine synthase complex, further confirmed by the structure of LdOASS in complex with SAT C-terminal octapeptide at 1.68 Å resolution. Docking and fluorescence binding studies show that almost all SAT C-terminus mimicking tetrapeptides can bind to LdOASS. Some peptides had a higher binding affinity than the native peptide, indicating that SAT–OASS interactions are not sequence-specific. The structure of LdOASS with a designed peptide (DWSI) revealed that LdOASS makes more interactions with the designed peptide than with the native peptide. In almost all known SAT–OASS interactions the SAT C-terminal sequence was shown to contain amino acids with large side chains. Structural comparison with other OASSs revealed that LdOASS has a relatively less open active-site cleft, which may be responsible for its interaction with the smaller-amino-acid-containing C-terminal LdSAT peptide. Biochemical studies confirmed that LdOASS interacts with SATs from *Entamoeba histolytica* and *Brucella abortus*, further displaying its sequence-independent and versatile mode of interaction with SATs. This implicates a critical role of the size of the active-site cleft opening in OASS for SAT–OASS interaction and thus cysteine synthase complex formation.

1. Introduction

Visceral leishmaniasis (VL) is the most severe form of leishmaniasis and may be lethal if left untreated. VL is caused exclusively by species of the *Leishmania donovani* complex. *L. donovani* infections are prevalent in East Africa, India and parts of the Middle East (Lukes *et al.*, 2007). This disease is the second largest parasitic killer (after malaria) in the world and is responsible for an estimated 500 000 cases every year worldwide (Desjeux, 2001; Turrens, 2004). Chemotherapy based on pentavalent antimonials has been the mainstay of treatment of visceral and cutaneous leishmaniasis to date. However, widespread failure to respond to these drugs, as reported in the endemic region for VL in north Bihar in India (Sundar *et al.*, 2000), has raised concerns over drug resistance. This resistance has been shown to be unique to *L. donovani* (Thakur *et al.*, 2001). Analysis of the factors contributing to

resistance and the identification of new drug targets in *L. donovani* is therefore important in the current scenario.

An increase in levels of trypanothione [N^1,N^8-bis(glutathionyl)spermidine], an intracellular thiol present in pathogenic trypanosomatids, has been linked to the acquisition of resistance in several resistant leishmanial species (Croft *et al.*, 2006). Trypanothione and trypanothione-dependent enzymes protect against oxidative damage and maintain intracellular redox homeostasis in trypanosomes (*Leishmania* spp. and *Trypanosoma* spp.); the maintenance of a trypanothione pool is also essential for their survival (Dumas *et al.*, 1997; Krieger *et al.*, 2000; Comini *et al.*, 2004). The availability of cysteine is crucial for the synthesis of glutathione and therefore trypanothione (Krauth-Siegel & Comini, 2008). Moreover, increased production of cysteine also helps to replace thiols lost owing to efflux, as well as to restore thiol redox potential perturbed by the accumulation of disulfides in drug-resistant species (Wyllie *et al.*, 2004). Cysteine has also been shown to be an essential growth factor in the related trypanosomatid *T. brucei* (Duszenko *et al.*, 1992). The cysteine-biosynthetic pathway is absent in humans but is present in *L. major*, where it was observed that levels of *O*-acetylserine sulfhydrylase (OASS), a cysteine-biosynthetic enzyme, were upregulated under sulfur-deficient conditions (Williams *et al.*, 2009). In view of the above findings, the enzymes of the cysteine-biosynthetic pathway are attractive drug targets.

The cysteine-biosynthetic pathway is a two-step process that involves the conversion of serine to *O*-acetylserine (OAS) catalyzed by the enzyme SAT (serine acetyltransferase; EC 2.3.1.30) followed by the incorporation of sulfide in a β-displacement reaction to form cysteine catalyzed by the enzyme OASS (EC 2.5.1.47). In plants and bacteria, the intracellular level of Cys is strictly regulated by two mechanisms. The first mode of control involves feedback inhibition of SAT by Cys. The second mechanism involves the association of OASS and SAT to form a cysteine synthase complex that is favoured under excess sulfur concentrations. In the complex, the C-terminal end of SAT binds the active site of OASS, and OASS activity decreases while SAT activity increases (Saito *et al.*, 1995). This results in excess production of *O*-acetylserine (OAS), which induces the dissociation of the cysteine synthase complex and downregulates the SAT activity (Fig. 1). The free OASS catalyses Cys formation and reduces OAS levels, which in turn allows re-association of SAT with OASS.

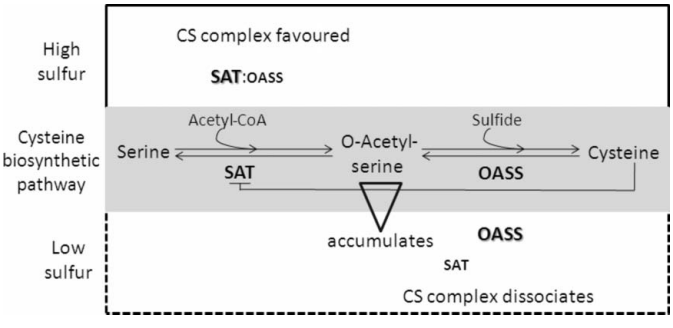

Figure 1
Schematic representation of the cysteine-biosynthetic pathway and its regulation. The grey panel outlines the general pathway. Feedback inhibition of SAT by cysteine is shown. The cysteine synthase (CS) complex forms under high-sulfur conditions and dissociates under low-sulfur conditions. Shaded and larger text indicates active forms of SAT and OASS. Smaller text indicates downregulated forms of the enzymes. Adapted and modified from Francois *et al.* (2006).

In studies performed previously, it was shown that OASS forms a complex with SAT C-terminal peptides which have large amino acids in their sequence and a crucial isoleucine at their C-terminal end, *e.g.* Asp-Tyr-Val-Ile in *Arabidopsis thaliana* (Francois *et al.*, 2006), Asn-Leu-Asn-Ile in *Haemophilus influenza* (Huang *et al.*, 2005) and Asp-Phe-Ser-Ile in *Mycobacterium tuberculosis* (Schnell *et al.*, 2007). However, it did not form a complex with OASS when the residues preceding Ile were smaller (Ser-Pro-Ser-Ile), as demonstrated in the case of *Entamoeba histolytica* (Kumar *et al.*, 2011). This trend does not seem to be dependent on any property of the amino acid involved in the interaction, but in most of these interactions there is a large hydrophobic residue at the third position from the C-terminal end (see below). These findings raise an interesting question regarding the possibility of cysteine synthase complex formation in *L. donovani*, as the putative SAT from *L. donovani* (LdSAT) has a C-terminal sequence with very small side chains: Gly-Ser-Gly-Ile.

Here, we have identified, cloned and sequenced the OASS gene from *L. donovani* (LdOASS). We have also carried out fluorescence binding studies of LdOASS with native and modified C-terminal SAT peptides, enzymatic inhibition studies and high-resolution crystal structure determinations of LdOASS with and without bound SAT peptides. The results of our present study show that LdOASS not only forms a complex with a putative LdSAT peptide, but is also capable of interacting with SATs from other organisms which have either larger or smaller amino acids in their C-terminal tails.

2. Experimental

2.1. Identification, sequencing and cloning of OASS from *L. donovani*

The *L. donovani* genome has not yet been sequenced. Therefore, the genome sequence of its closest relative, *L. infantum*, was used to identify the putative gene sequence which codes for the OASS enzyme in this organism and this sequence was then used to design primers for OASS in *L. donovani*. This sequence was cross-checked with other known OASS sequences from plants, bacteria and parasitic protozoa as queries and showed high sequence similarity to all of them. Degenerate primers were then used to clone OASS from *L. donovani* (strain MHOM/IM/1983/AG83). The gene was amplified using the following forward and reverse primers: OASS Fp, 5′-ATGGCGGCACCGTTCG-3′, and OASS Rp, 5′-CTACGGGCAGGGACGACAC-3′ (Supplementary Fig. 1*a*[1]). The amplified gene was sequenced and the sequence was deposited in GenBank with accession No.

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: MV5063). Services for accessing this material are described at the back of the journal.

JF728801. The OASS gene was cloned into pET21c vector (Novagen) between *Nde*I and *Not*I sites with a C-terminal His_6 tag.

2.2. Overexpression and purification

The recombinant plasmid pET21c containing the LdOASS insert was transformed into *Escherichia coli* BL21 (DE3) cells (Novagen). Freshly transformed BL21 cells were grown in Tartoff–Hobbs HiVeg broth medium (HiMedia) supplemented with 100 µg ml^{-1} ampicillin at 310 K to an OD_{600} of 0.5. Expression of recombinant protein was induced with 0.25 m*M* IPTG (Sigma) overnight at 289 K. The cells were harvested by centrifugation at 6000*g* for 10 min at 277 K and resuspended in buffer *A* (50 m*M* Tris pH 8.0, 150 m*M* NaCl, 5 m*M* β-mercaptoethanol) containing 100 µ*M* PMSF and 0.1%(*v*/*v*) Triton X-100 (USB Corporation). The cells were lysed and the clear supernatant was passed through an Ni Sepharose column (GE Healthcare) pre-equilibrated with buffer *A*. The column was washed with buffer *A* containing 10 and 50 m*M* imidazole and LdOASS was eluted in 1 ml fractions with 300 m*M* imidazole in buffer *A*. All fractions were checked on 10% SDS–PAGE (Supplementary Fig. 1*b*); the fractions containing sufficiently pure LdOASS were concentrated and subjected to gel filtration on a HiLoad Superdex 200G 16/60 column (GE Healthcare) at a flow rate of 0.5 ml min^{-1}. The elution profile showed that LdOASS exists as a dimer (Fig. 2). The purity of the protein was assessed by SDS–PAGE and it was concentrated using Centricon tubes (Amicon) to a final concentration of 32 mg ml^{-1} as estimated by the spectroscopic method using an ε_{280} of 15 930 M^{-1} cm^{-1} for LdOASS as calculated using the *ExPASy* server.

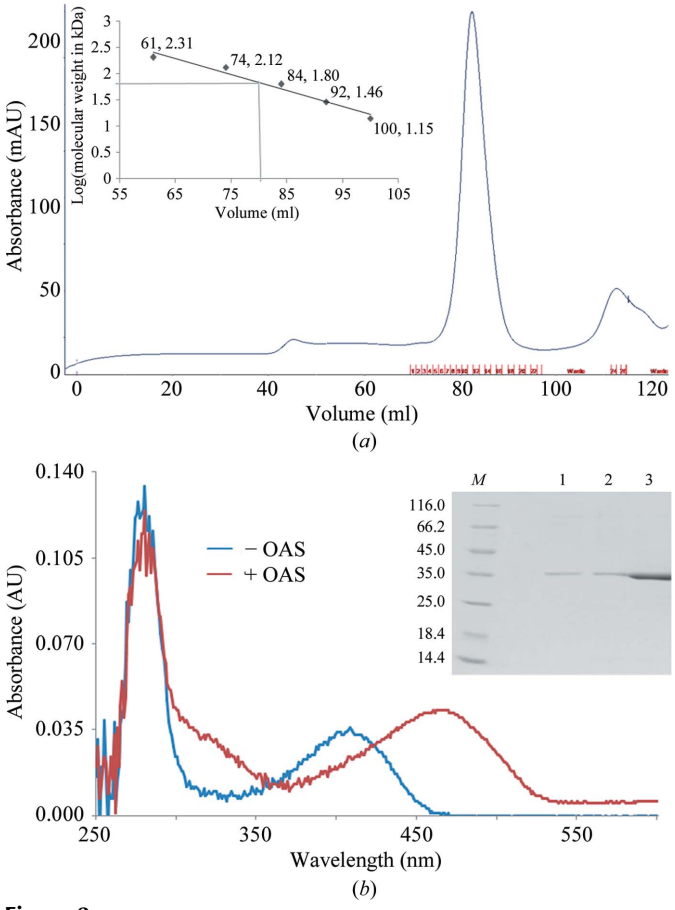

Figure 2
Characterization of LdOASS. (*a*) Gel-filtration profile of LdOASS; the protein eluted at 82 ml from a Sephadex 200 column, indicating that it exists as a dimer. The inset on the left shows a plot of log molecular weight of standards *versus* elution volume. The inset on the right shows an SDS–PAGE of fractions of the elution peak. Lane *M*, molecular-weight markers (labelled in kDa); lanes 1, 2 and 3, elution fractions. (*b*) Spectrum of LdOASS showing the peak at 412 nm that indicates the presence of PLP in the active site (blue line). When OAS is added the absorbance at 412 nm decreases and the absorbance at 470 nm increases (red line), indicating formation of the α-aminoacrylate intermediate.

2.3. Enzyme activity

UV–visible absorbance spectra of LdOASS were recorded at a concentration of 300 µg ml^{-1} (8.7 µ*M*) using a Ultrospec 21000pro spectrophotometer (GE Healthcare). Characteristic spectra of the covalent amino-acrylate complex were obtained after the addition of 150 µ*M* *O*-acetylserine to the enzyme solution. The LdOASS sulfhydrylase activity was monitored using TNB (5-thio-2-nitrobenzoate) as an alternative substrate as described previously (Kumar *et al.*, 2011). Briefly, the disappearance of TNB was monitored continuously at 412 nm. A typical assay contained 100 m*M* HEPES pH 7.0, 25 µ*M* TNB and 30 µg (2 µ*M*) LdOASS (final concentrations); this reaction mixture was incubated at room temperature for 30 min. The reaction was started by the addition of OAS and readings were taken immediately at 412 nm. To determine the K_m for OAS, its concentration was varied from 0.25 to 32 m*M* while keeping the TNB concentration constant at 25 µ*M*. Activity was derived from the linear part of the time curves (3 min) and each measurement was carried out in triplicate at room temperature.

2.4. Crystallization

LdOASS was used at a concentration of 16 mg ml^{-1} in extensive crystallization screening using a Mosquito crystallization robot to dispense hanging drops for vapour diffusion. The following preliminary crystallization conditions were obtained: Molecular Dimensions PACT premier screen conditions P2-14 [0.2 *M* sodium bromide, 0.1 *M* bis-Tris propane pH 6.5, 20%(*w*/*v*) PEG 3350], P2-15 [0.2 *M* sodium iodide, 0.1 *M* bis-Tris propane pH 6.5, 20%(*w*/*v*) PEG 3350] and P2-16 [0.2 *M* potassium thiocyanate, 0.1 *M* bis-Tris propane pH 6.5, 20%(*w*/*v*) PEG 3350]. The conditions were then replicated and optimized to obtain better quality crystals in larger drops. The best crystals were obtained using 225 m*M* KSCN and 18% PEG 3350 in 100 m*M* bis-Tris propane pH 6.2 at 289 K. The C-terminal octapeptide LERDGSGI derived from the putative LdSAT sequence was cocrystallized with native LdOASS. Another tetrapeptide, DWSI, which was found to have a good binding affinity in fluorescence experiments with LdOASS was also cocrystallized in order to study the interactions that it makes with LdOASS. Cocrystallization with the peptides was achieved using the crystallization

Table 1
Crystallographic data and refinement statistics.

Values in parentheses are for the highest resolution shell. R_{free} was calculated using a randomly selected subset of 5% of the reflections.

Data set	LdOASS	LdOASS–DWSI	LdOASS–octapeptide
Crystallographic data			
X-ray source	BM14, ESRF, France		
Wavelength (Å)	0.9762	0.9762	0.9786
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit-cell parameters (Å)			
a	115.266	115.348	115.217
b	61.974	61.916	62.440
c	43.427	43.431	43.538
Resolution range	115.27–1.79 (1.84–1.79)	115.35–1.80 (1.84–1.80)	115.22–1.68 (1.72–1.68)
Completeness (%)	98.4 (87.5)	99.0 (89.1)	99.5 (98.5)
Total No. of observations	486701	439872	1166652
No. of unique observations	27999	29361	34664
Multiplicity	7.8 (5.6)	7.8 (6.0)	14.4 (12.5)
Average $I/\sigma(I)$	23.9 (2.3)	30.15 (3.2)	64.5 (6.8)
Crystal mosaicity (°)	0.29	0.45	0.57
R_{merge}	0.064 (0.386)	0.074 (0.355)	0.063 (0.349)
Refinement			
R factor (%)	17.8	17.8	18.5
R_{free} (%)	21.7	20.8	22.2
Mean *B* factor	20.66	20.25	21.91
No. of atoms			
Protein	2396	2390	2383
Water	186	169	232
Na	1	1	1
Cl	1	—	—
PEG	6	6	6
Peptide	—	36	30
R.m.s. deviations			
Bond lengths (Å)	0.02	0.03	0.02
Bond angles (°)	1.91	2.44	2.23
E.s.u. based on R_{free} (Å)	0.12	0.10	0.10

conditions established for the native enzyme. LdOASS at a concentration of 16 mg ml^{-1} was mixed with the peptides at a concentration of 5 m*M* and pre-incubated at room temperature for 60 min before setting up the crystallization drops using the conditions described above.

2.5. Data collection and processing

All of the crystals were washed and equilibrated with a cryosolution with the same composition as the mother liquor of the respective drop with an increased concentration of PEG 3350 of up to 20%(*w*/*v*) and an additional 12%(*w*/*v*) PEG 400. The crystals were then mounted in cryoloops and flash-cooled in a liquid-nitrogen stream at 100 K. Initial data were collected at the home source (Advanced Instrumentation Research Facility) using a Bruker Microstar generator and a MAR imaging-plate detector. The high-resolution diffraction data were collected on ESRF beamline BM14. The data sets were indexed, integrated and scaled using the *HKL*-2000 data-processing software (Table 1).

2.6. Structure determination and refinement of LdOASS

The structure of LdOASS was determined by the molecular-replacement method using *A. thaliana* OASS (AtOASS; PDB entry 1z7w; Bonner *et al.*, 2005), with which it has 44% sequence identity, as a search model. LdOASS crystallized in space group $P2_12_12$ with one molecule in the asymmetric unit. The output from *CHAINSAW* (Stein, 2008) of the aligned sequences of LdOASS and AtOASS was used for molecular replacement in *MOLREP* (Vagin & Teplyakov, 2010); the best solution had a correlation coefficient of 0.502 and an *R* factor of 57.2%. The default resolution range (in this case 42–2 Å) for the program *MOLREP* was used for molecular replacement. An initial cycle of restrained refinement using *REFMAC*5 (Murshudov *et al.*, 2011) with all data resulted in a decrease in the *R* factor to 33.9%. The model was then submitted to *ARP*/*wARP* (Morris *et al.*, 2002) for auto-building, which successfully built 90% of the side chains in the model with good-quality electron density and an *R* factor of 22%. The remaining parts of the structure were built manually with *Coot*, guided by σ-weighted $2F_o - F_c$ and $F_o - F_c$ electron-density maps. Rounds of manual adjustment of the model in *Coot* (Emsley & Cowtan, 2004) were alternated with refinement with *REFMAC*5. The PLP molecule, PEG molecule, water molecules, sodium ion and chloride ion were added manually guided by $F_o - F_c$ electron density at a $>3\sigma$ contour level. The water molecules were initially auto-searched and incorporated by the *Coot* graphics package and were subsequently manually checked by considering electron density and hydrogen-bond interactions with the protein. The final model consisted of 2396 protein atoms, one PLP molecule, 186 water molecules, one sodium ion, one chloride ion and one PEG molecule with good refinement statistics (Table 1). The structure has a relatively large C-terminal domain and a small N-terminal domain (Fig. 3). The coordinates of the native LdOASS structure and the structure factors have been deposited in the PDB with accession code 3spx.

2.7. SAT C-terminus mimicking peptide binding assay and inhibition studies

SAT with an Ile at its C-terminal end has previously been observed to interact with OASS in other organisms (Francois *et al.*, 2006; Salsi *et al.*, 2010). The last four amino-acid residues of SAT are the primary contributors to this interaction at the active site of OASS; we generated a computational library of tetrapeptides against LdOASS trying all permutations and combinations of the three penultimate amino acids while

keeping Ile constant at the C-terminal end. All of the peptides were docked into the active site of LdOASS using *GLIDE* (Friesner *et al.*, 2004) and the peptides with the best scores were shortlisted and obtained from GenPro Biotech. These peptides were DWSI, DFSI, DYSI, DPTI and DTTI (Supplementary Table 1). The C-terminal LdSAT-derived peptide LERDGSGI was also used for binding and inhibition studies. Since the C-terminal end of LdSAT consists of small amino acids, we also wanted to compare the interaction with the *E. histolytica* SAT1-derived C-terminal peptide SPSI. In *E. histolytica* SAT and OASS do not interact and it has been shown that the amino acids with small side chains at the end of *E. histolytica* SAT1 were responsible for the loss of interaction (Kumar *et al.*, 2011). The binding interaction was studied by fluorescence experiments based on the methods described by Campanini *et al.* (2005). Fluorescence measurements were carried out using a Cary 400 Scan spectrophotometer (Varian Inc.). Emission spectra upon excitation at 412 nm (slit excitation = 5 nm, slit emission = 10 nm) of a solution consisting of 300 µg ml^{-1} (8.7 µ*M*) LdOASS, 100 m*M* HEPES pH 7 were recorded. Fluorescence peaks were measured in the absence of the peptides and following each addition of peptide until saturation was reached (Fig. 4). The K_d value for each peptide was calculated using *SigmaPlot* as described in Kumar *et al.* (2011).

The decrease in the activity of LdOASS in the presence of the peptides was measured as follows. A standard reaction mixture consisting of 100 m*M* HEPES pH 7.0, 25 µ*M* TNB and 30 µg (2 µ*M*) LdOASS was incubated at room temperature for 30 min and the reaction was then started by adding 0.25 m*M* OAS. The decrease in absorbance at 412 nm was monitored over a fixed interval of time. To observe the effect of peptides on the activity of LdOASS, 0.25 m*M* peptide was added to the above reaction mixture and given the same treatment. All of these experiments were run in triplicate. The spectrum of the standard reaction was then compared with those with the peptides and the percentage decrease in activity was calculated using the following equation: 100 − [(decrease in absorbance for reactions with peptide/decrease in absorbance for the standard reaction) × 100].

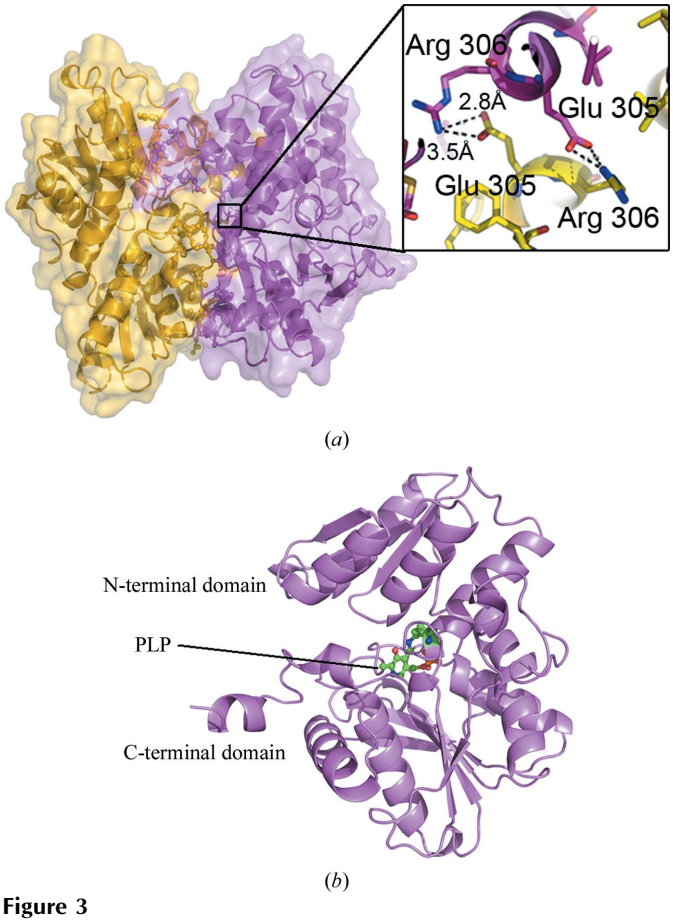

Figure 3
The dimeric arrangement of LdOASS. (*a*) Two molecules from two different asymmetric units interact with each other to form dimers, which are displayed as a cartoon generated by *PyMOL* (DeLano, 2002). The dimeric interface makes symmetrical interactions, forming several hydrophilic and hydrophobic interactions. Salt bridges between Arg306 of one monomer and Glu305 of the other are observed in the dimer interface, as seen in the inset. (*b*). The monomer, with a large C-terminal domain and a small N-terminal domain separated by active-site cleft. PLP cofactor is observed at the active-site cleft.

2.8. Structure determination and refinement of the LdOASS–peptide complexes

The structures of the LdOASS–peptide complexes were solved by molecular replacement using the native LdOASS structure as a search model. The electron density calculated after the initial refinement rounds indicated the presence of the peptide bound at the enzyme active site. Refinement of the complexes was carried out using a similar protocol to that outlined above. The final model with DWSI comprised 319 LdOASS residues, the four-residue peptide, one sodium ion, one PEG molecule and 169 water molecules. The model with the natural LdSAT peptide consisted of 318 LdOASS residues, only five residues of the C-terminal octapeptide, 232 water molecules, one PEG molecule and one sodium ion. The protein models were analyzed with *PROCHECK* (Laskowski *et al.*, 1993) to monitor their stereochemistry. Details of the refinement and protein models are given in Table 1. The coordinates and structure factors of the structures of LdOASS in complex with the octapeptide (LERDGSGI) and with the tetrapeptide (DWSI) have been deposited in the PDB with PDB codes 3t4p and 3tbh, respectively.

2.9. Inhibition studies of LdOASS with SATs from other organisms

SAT1 from *E. histolytica* (EhSAT1) was expressed and purified as described in Kumar *et al.* (2011). The cloning, expression and purification of SAT from *Brucella abortus* (BaSAT) was performed in our laboratory and will be reported later. BaSAT also has smaller amino acids at its C-terminal end with the sequence GDGI. The putative LdSAT has the sequence GSGI at its C-terminal end, which is similar to the BaSAT sequence; therefore, it would be interesting to determine its effect on LdOASS activity. Concentrated EhSAT1 and BaSAT in a buffer consisting of 150 m*M* NaCl, 50 m*M* HEPES pH 7 were used in inhibition studies with

LdOASS. Inhibition was studied at equimolar concentrations (25 μ*M*) of OAS and EhSAT1/BaSAT in a 400 µl reaction mixture; the LdOASS concentration was kept at 60 µg (4 μ*M*). The reaction was run in duplicate. The decrease in enzyme activity was monitored and the percentage inhibition was calculated as in the case of the peptides.

3. Results and discussion

3.1. Identification of the gene encoding OASS from *L. donovani*

Based on the phylogenetic proximity of *L. donovani* and *L. infantum*, we hypothesized the presence of an orthologous gene encoding a putative OASS in *L. donovani*, which was demonstrated when this gene was subsequently cloned from the *L. donovani* genome. This gene of 975 base pairs was cloned into pET21c between *Nde*I and *Not*I restriction sites. The sequence was deposited in the NCBI GenBank and given the accession No. JF728801. The gene, which encodes a protein of 324 amino acids (34.4 kDa), has a high sequence identity to other protozoan OASSs and type A isoforms of bacterial and plant cysteine synthases (Supplementary Fig. 2). The catalytic residues and the cofactor PLP-binding motif sequence are also conserved.

3.2. Characterization of OASS from *L. donovani*

The 6×His-tagged recombinant LdOASS was purified in two steps, metal-affinity chromatography and size-exclusion chromatography (SEC), as described in §2. The SEC profile revealed LdOASS to be dimeric in solution (Fig. 2*a*), as observed in many other organisms. Purified LdOASS has the yellow colour characteristic of all OASSs. The absorbance spectra for LdOASS have maxima at 280 and 410 nm, with an A_{280}:A_{410} ratio of about 4. The absorbance peak at 410 nm is characteristic of the internal aldimine formed between the cofactor PLP and the active-site lysine residue (Lys51), which is responsible for the yellow colour of the protein. Addition of OAS results in a decrease in absorbance at 410 nm with a concomitant increase in absorbance at 320 and 470 nm, indicating the formation of the α-aminoacrylate intermediate (Fig. 2*b*). The activity of the enzyme was characterized using OAS and TNB as substrates. It has a K_m of 18.2 ± 0.2 m*M* for OAS, indicating very weak binding affinity for its substrate, which is comparable to the affinity of the OASS enzymes from *L. major* and *Salmonella typhi* (Williams *et al.*, 2009; Tai *et al.*, 1993).

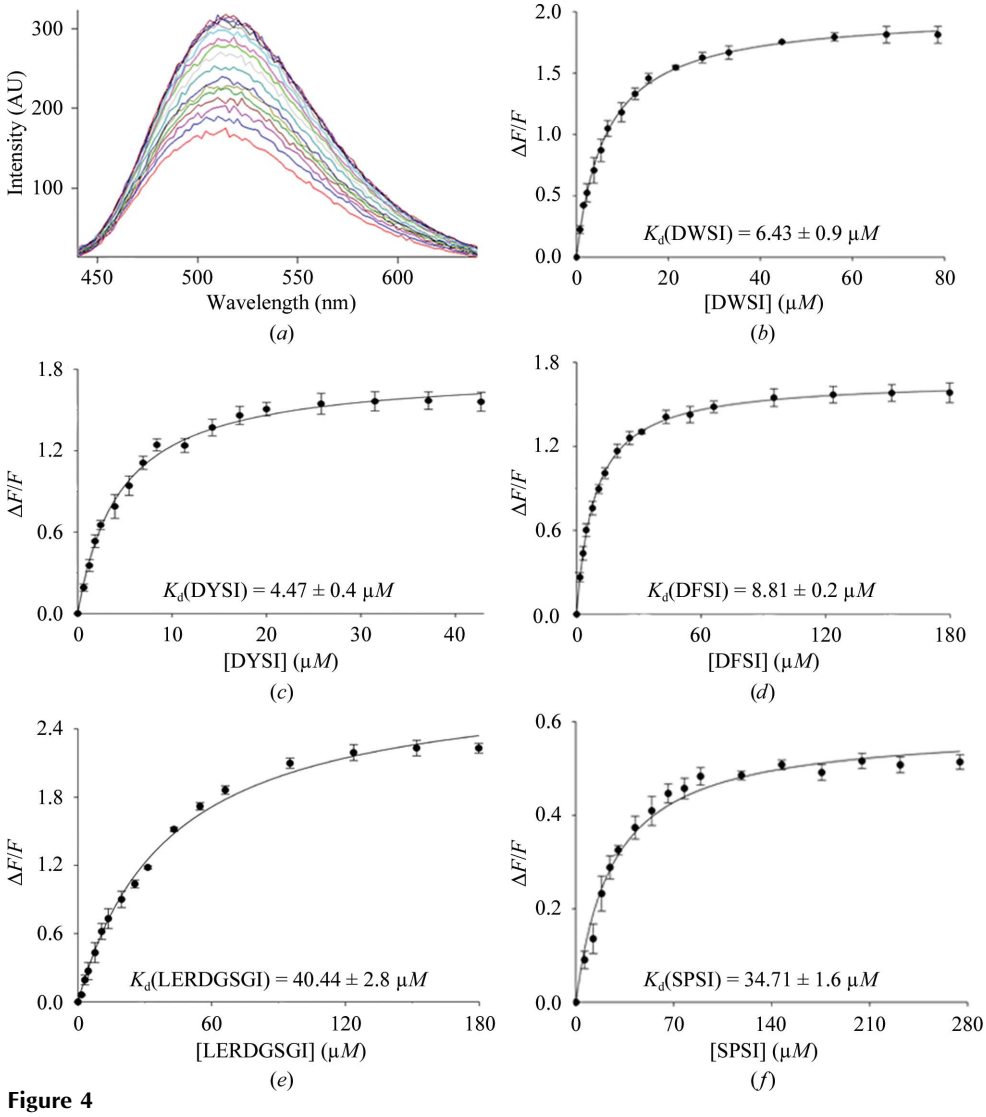

Figure 4
Changes in the fluorescence spectrum of LdOASS used to study the binding affinity of peptides. (*a*) Typical fluorescence spectrum of LdOASS with titration of peptide. The fluorescence emission of PLP in the active site was monitored by changes in the fluorescence at 508 nm *versus* (*b*) DWSI concentration, (*c*) DFSI concentration, (*d*) DYSI concentration, (*e*) LERDGSGI concentration and (*f*) SPSI concentration.

3.3. The C-terminal peptide of SAT binds to LdOASS

As described above, OASS associates with SAT to form a bi-enzyme cysteine synthase complex in which the OASS activity is downregulated. Crystallographic studies of OASSs from *H. influenzae* and *A. thaliana* have shown that an essential part of the interactions between the two enzymes involves the binding of the four C-terminal residues of SAT to the active site of OASS (Francois *et al.*, 2006; Salsi *et al.*, 2010). A conserved Ile at the

C-terminal end of SAT is essential for binding OASS to form a regulatory cysteine synthase complex (Francois *et al.*, 2006; Supplementary Fig. 3).

The C-terminal octapeptide of *L. donovani* SAT deduced from the putative gene coding for SAT in *L. major* has the sequence LERDGSGI with an Ile at the C-terminal end. SAT–OASS interactions were nevertheless assumed to be absent in *L. donovani* because of previous studies on the *E. histolytica* OASS enzyme (EhOASS; Kumar *et al.*, 2011), in which it was shown that EhOASS did not interact with EhSAT1 despite the presence of Ile at the C-terminus. It appeared that the small amino-acid side chains just prior to the Ile in EhSAT1 caused weakened binding since the modification of these residues to DWSI, DYSI or DFSI yielded improved binding to EhOASS. The C-terminal end of LdSAT (GSGI) also contains small amino-acid side chains.

To verify this prediction, we obtained the LdSAT octapeptide LERDGSGI and measured its binding to LdOASS by fluorescence. Based on a plot of the monitored change in the fluorescence signal of the active-site PLP against the peptide concentration (Fig. 4), the binding affinity of this octapeptide was measured to be 40 μ*M*. Surprisingly, SAT binds relatively strongly to OASS in *L. donovani*.

We therefore aimed to determine what allows these two proteins to interact in *L. donovani* but not in *E. histolytica* despite the presence of small amino-acid side chains in both of their SATs. To answer this question, we have carried out additional binding studies and have determined and analyzed high-resolution crystal structures, as described in the remainder of this manuscript.

3.4. Large side chains in SAT-mimicking peptides strengthen binding to LdOASS

A library of peptides mimicking the C-terminal end of SAT was generated to study their binding affinity to LdOASS. Different combinations of tetrapeptides were generated for docking studies, keeping Ile at the C-terminal end. The tetrapeptides were docked into the active site of LdOASS based on the MtbOASS–tetrapeptide complex as a model (Schnell *et al.*, 2007) following similar methods as reported in Kumar *et al.* (2011). On the basis of energy parameters, five peptides were shortlisted (Supplementary Table 1).

The binding affinities of the three best SAT-mimicking tetrapeptides with the highest binding energies (DWSI, DYSI and DFSI) were determined. The *E. histolytica* SAT-derived tetrapeptide (SPSI) was included in the study as it has small amino-acid side chains that are comparable to those in the *Leishmania* SAT C-terminal peptide. The results showed that the peptides with large side chains (DWSI/DFSI/DYSI) bind LdOASS with K_d values in the range 4–8 μ*M*, whereas the native *E. histolytica* SAT-derived and putative *L. donovani* SAT-derived peptides had K_d values of approximately 40 μ*M* (Fig. 4). The LdOASS enzyme-inhibition studies also indicated that the peptides DWSI and DYSI inhibited the enzyme activity more strongly than the native LdSAT-derived peptide (Supplementary Table 2), which is consistent with the measured differences in binding affinity. Taken together, the affinity measurements indicate that LdOASS can bind relatively strongly to almost any SAT with a C-terminal sequence ending with Ile, but that the strength of the binding affinity correlates with the size of the penultimate residues, specifically a large hydrophobic residue at the third position from the C-terminal end.

Table 2
Comparison of cleft opening at the active site.

The distance between the N-terminal upper domain and the C-terminal lower domain near the cleft opening of the OASS active site in the native state.

Protein (PDB code)	Equivalent residues at the opening of the cleft	Distance (Å)
LdOASS (3spx)	Gly129–Pro225	11.1
MtbOASS (2q3b)	Gly121–Pro217	12.0
EhOASS (2pqm)	Gly135–Pro231	13.1
StOASS (1oas)	Gly118–Pro223	12.5
AtOASS (1z7w)	Gly124–Pro220	12.2
TmOASS (1o58)	Glu108–Lys206	7.9

3.5. The crystal structure of *L. donovani* OASS reveals a narrow cleft

The overall structure of LdOASS is very similar to those of OASSs from other organisms (Fig. 3). One molecule of LdOASS is present in one asymmetric unit; however, the protein is a dimer in the crystal and its molecular axis coincides with the crystallographic axis (Fig. 3). The dimer is stabilized by the formation of multiple hydrogen bonds and hydrophobic interactions, primarily between the N-terminal region and α-helices 9 and 10 of the C-terminal domain. Among the many interactions, salt bridges between Glu305 and Arg306 of each monomer also stabilize the dimer (Fig. 3*a*). The buried surface area at the dimer interface is 2488.9 Å^2, which corresponds to 17.2% of the monomer surface area. Each monomer has the typical fold of type II PLP enzymes and consists of two domains, each with an α/β-fold. The cofactor PLP is bound to Lys51 by a Schiff linkage and rests in the active-site groove lined by α-helix 2 and loops from the C-terminal domain. PLP is present at the intersection of the two domains (Fig. 3*b*). The 3′-hydroxyl group of PLP forms a hydrogen bond to the side chain of Asn74 from a highly conserved sequence stretch: ^{76}ESSSGNTG^{83}. These interactions are identical to those determined in other OASS enzymes (Schnell *et al.*, 2007; Burkhard *et al.*, 1998; Huang *et al.*, 2005). LdOASS differs from most other structures by having a short helix formed by residues 125–128 lining the active-site entrance (Fig. 5); a similar helix is also present in OASS from *T. maritima* (Heine *et al.*, 2004).

Although the sequence and overall structures of OASSs are very similar among the different organisms, a detailed comparison reveals critical differences in the opening of the active-site cleft that explain the different SAT peptide-binding properties of LdOASS and EhOASS. TmOASS has the most closed cleft, at just 8 Å wide, but LdOASS has the second narrowest cleft at 11 Å wide. The clefts of all other OASS structures are at least 12 Å wide, with that of EhOASS being

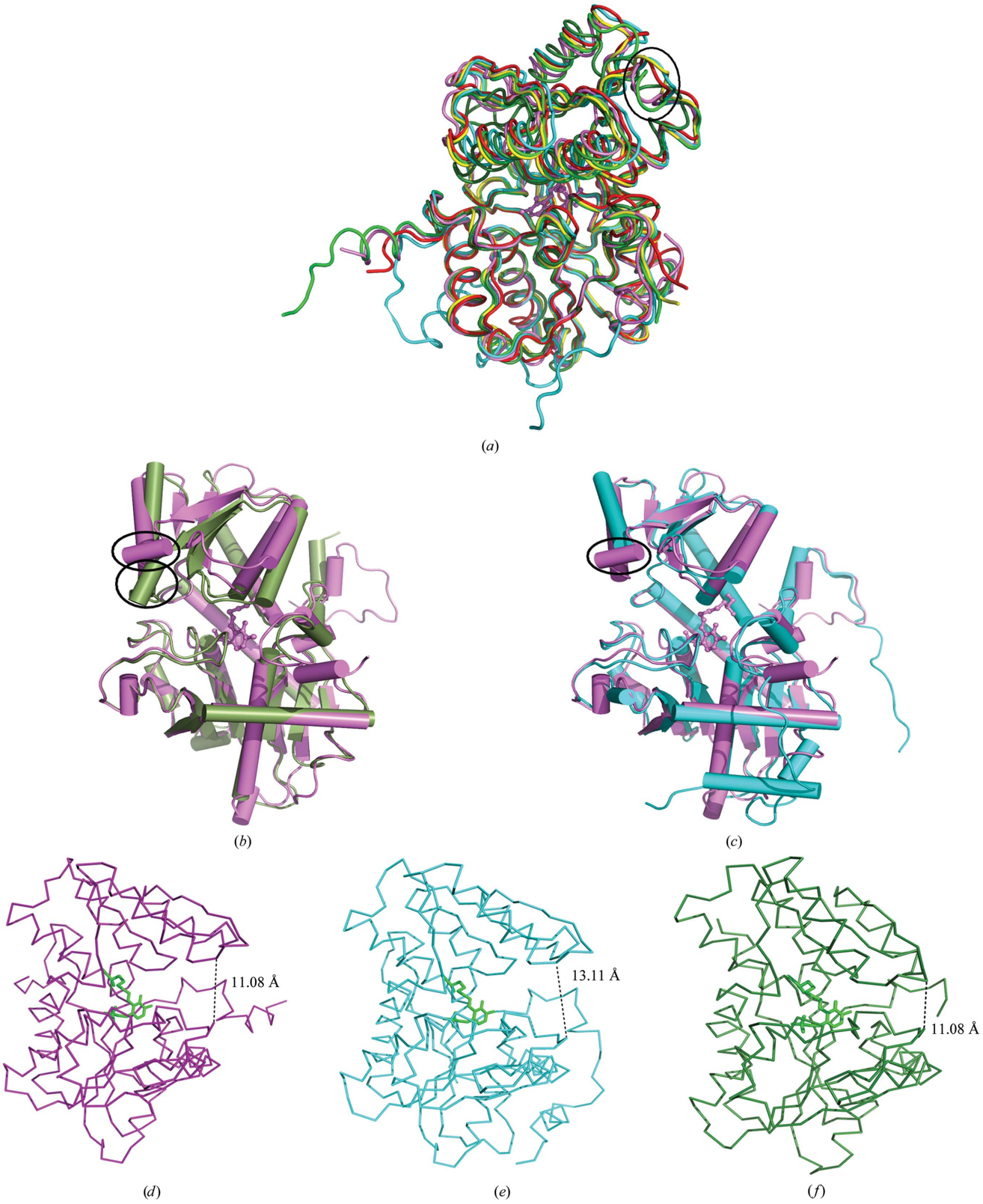

Figure 5
Superimposition and comparison of various OASS structures. (*a*) All of the OASS structures were superposed on each other with respect to the C-terminal domain; deviation is observed at a loop (66–74) in the N-terminal domain indicated by a circle. The structures of EhOASS (cyan), AtOASS (green), TmOASS (forest green), MtbOASS (yellow) and HiOASS (red) were superimposed on LdOASS (violet). Superimposed structures of (*b*) TmOASS (sap green) and (*c*) EhOASS (cyan) on LdOASS (violet) showing the encircled extra helix corresponding to residues 125–128 of LdOASS. This helix is also observed in TmOASS but is absent in EhOASS and all other OASS structures. The widths of the opening of the active-site cleft in (*d*) LdOASS, (*e*) EhOASS and (*f*) TmOASS are indicated. [The orientation of the molecule in (*b*) and (*c*) is rotated by 180° around the *y* axis compared with those in (*d*), (*e*) and (*f*).]

the widest at 13.1 Å (Table 2). The difference in the cleft opening arises from a small variable N-terminal domain loop at residues 66–74 (Supplementary Fig. 2) that connects to the active-site region α/β motif (containing β-strand residues 73–80 and α-helical residues 83–94; Fig. 5). An analysis of the surface area of various SAT C-terminal tetrapeptides indicates that the peptides which have higher binding affinity (DWSI/DFSI/DYSI) for OASS have a surface area of around 450 Å^2 or more (Supplementary Fig. 3). However, the EhSAT1 C-terminal tetrapeptide has a surface area of only 382 Å^2; the surface area of the LdSAT and BaSAT tetrapeptides are 328 and 341 Å^2, respectively, but they still interact with LdOASS, albeit with a lower binding affinity.

The inability of EhOASS to bind to peptides with small side chains may therefore be a consequence of its unusually wide active-site cleft. Similarly, the moderately narrow cleft of LdOASS appears to explain its ability to bind peptides with small as well as large side chains (see below). This property may be unique to *L. donovani*: several structural and functional studies reported for HiOASS (Huang *et al.*, 2005), MtbOASS (Schnell *et al.*, 2007) and AtOASS (Francois *et al.*, 2006) show that all of these OASS enzymes only bind strongly to suitably large SAT side chains.

3.6. Structures of LdOASS–peptide complexes

Since crystallographic studies of HiOASS and AtOASS have shown that the essential interactions between OASS and SAT involve binding of the C-terminal four residues of SAT to the active site of OASS (Francois *et al.*, 2006; Salsi *et al.*, 2010), we set out to examine whether similar interactions are formed in *L. donovani*. We have been able to obtain crystals and determine X-ray structures of LdOASS in complex with the SAT-derived C-terminal octapeptide LERDGSGI and with the tetrapeptide DWSI obtained from docking and binding

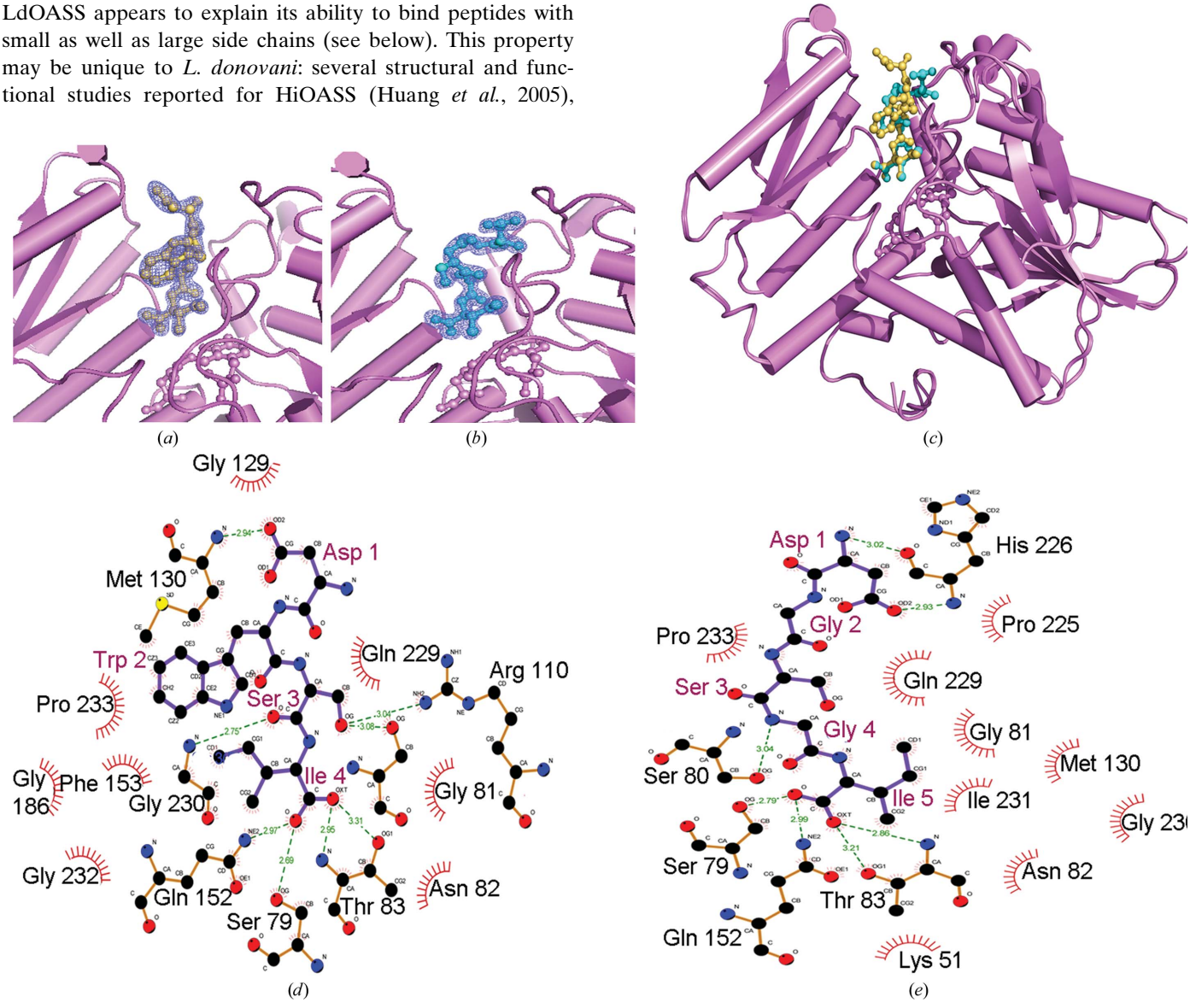

Figure 6
Peptide-bound structures of LdOASS at the active site. The difference Fourier electron-density map ($2F_o - F_c$) of peptides at 2σ was superposed on the peptides (*a*) DWSI and (*b*) LERDGSGI with the protein molecule around the active site. (*c*) Superposition of the active site shows that the peptides bind to the same location in the active site. The peptide interactions with LdOASS are shown with *LIGPLOT* for (*d*) DWSI and (*e*) DGSGI. The Ile at the C-terminal end forms the same set of interactions with the protein in both peptides.

studies. In each case the peptide extends from the enzyme surface into the interior of the LdOASS active site (Fig. 6), locking the enzyme into an open conformation. The peptides are better ordered at their C-termini. Only five (DGSGI) of the eight amino acids could be modelled in the case of the LdSAT-derived octapeptide, whereas all four amino acids in the DWSI complex could easily be traced.

The relevant interactions formed by the two different peptides at the LdOASS active-site binding pocket are shown in Figs. 6(*a*) and 6(*b*). Some of these interactions are similar to those reported previously for *A. thaliana* (Francois *et al.*, 2006), *H. influenzae* (Huang *et al.*, 2005) and *M. tuberculosis* (Schnell *et al.*, 2007). A list of contacts between the peptides and LdOASS generated by *get_near_res* (*UCSF Dock* v.5.1.1; http://dock.compbio.ucsf.edu; Supplementary Table 3) shows that the same set of amino-acid residues in LdOASS contribute to binding interactions with the peptide, even though the residues in the peptides differ. The terminal Ile forms the most contacts, stabilizing the peptide in the active site by hydrogen bonds to the conserved amino-acid stretch ^{78}SSSGNTG^{84} and by hydrophobic interactions with Gly186, Thr187 and Ile231. The amino-acid residue at the penultimate position makes the usual interaction with Ser80, Gln229 and Gly230 in both cases; the serine in DWSI forms an extra hydrogen bond to Arg110 of the main chain compared with the native peptide. The major difference between the two peptides occurs at the third residue from the C-terminal end. Trp in this position forms extensive interactions with LdOASS, forming apolar contacts with the hydrophobic loop ^{232}GPGF^{235}, Val134, Phe153 and His226. Trp at this position contributes to a total of 11 interactions, whereas serine at this position in the native peptide (DGSGI) only interacts with the main chain at Met130. The Asp at the fourth residue from the C-terminus in DWSI forms hydrogen bonds to the backbone amide N atoms of Met130 and Lys131, while Gly at the same position only participates in one interaction with Met106. The overall number of interactions made by the four terminal amino acids in DWSI and DGSGI are 27 and 19, respectively (Fig. 6; Supplementary Table 3). The Asp at the fourth residue from the C-terminus in DWSI forms hydrogen bonds to the backbone amide N atoms of Met130 and Lys131, while the Asp at the fifth residue from the C-terminus in DGSGI interacts at other side of the cleft with the backbone N atom of His223 and the O atom of Gly224. From a previous peptide-binding study with MtbOASS it was inferred that peptide recognition is mediated *via* side-chain interactions in a sequence-specific manner and that the only enzyme–peptide interaction common to all OASSs was the binding mode of the conserved C-terminal isoleucine residue *via* specific hydrogen bonds to its carboxyl group and hydrophobic interactions with the side chain. However, in LdOASS–peptide complexes even the penultimate amino-acid serine in DWSI and glycine in DGSGI formed the same set of interactions with the conserved Ser80 of the main chain. A similar study with designed peptides showed that Asn or Gln at the penultimate position would interact better with the conserved Ser79; hence, these side-chain interactions are not sequence-specific. Even though the specific interactions are contributed by the two C-terminal residues of the peptides, it appears that the binding strength and affinity are contributed by the third and fourth amino acids.

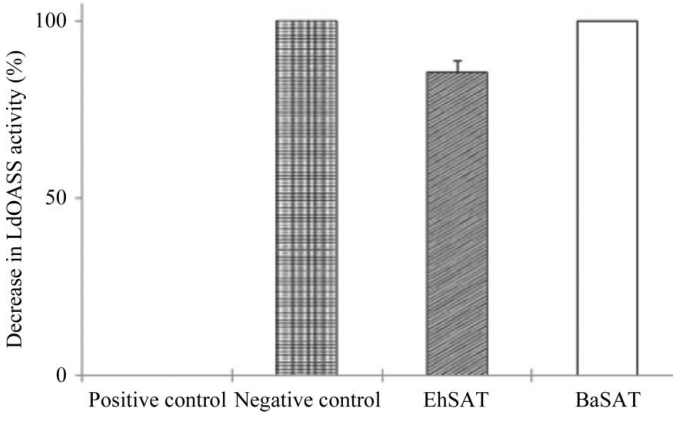

Figure 7
Effects of SATs from various organisms on the activity of EhOASS in the presence of an equimolar concentration of OAS (25 μ*M*). EhSAT inhibits LdOASS activity by around 85%, whereas BaSAT causes a 100% decrease in activity. The decrease in EhOASS activity is shown as a percentage relative to the positive control measured in the absence of any inhibitor. The negative control is shown in the absence of substrate.

Our biochemical results show that LdOASS is almost completely inhibited (100%) by SAT from *B. abortus* and is partially inhibited (85%) by SAT1 from *E. histolytica* (Fig.7). These results indicate that LdOASS can indeed bind to SATs from various organisms to form a cysteine synthase complex.

4. Conclusion

OASS from *L. donovani* turns out to be an unusually indiscriminate binder of SATs. As expected, LdOASS shares numerous features with OASSs from other organisms. For example, it is dimeric (both in solution and in the crystal structure) and has an overall structure very similar to those of other OASSs.

However, in contrast to other well characterized OASSs, especially EhOASS, which bind only SAT C-terminal peptides containing large side chains, LdOASS can bind a variety of such peptides including those with small side chains (albeit not quite as strongly as those with large side chains). LdOASS appears to accomplish this feat by forming an active-site cleft of just the right size: narrower than that of OASS from *E. histolytica* and most other organisms. (The size of the cleft seems to result from the structure of a small variable loop comprising residues 66–74 in the N-terminal domain.) Taking these results together with inhibition studies that show that LdOASS interacts with SATs from *E. histolytica* and *B. abortus*, we conclude that LdOASS can interact with SAT from any organism.

We thank the staff of beamline 14 at the European Synchrotron Radiation Facility and the Department of

Biotechnology, Government of India for obtaining access to the beamline. We thank Professor Chinmay K. Mukhopadhyay, SCMM, JNU for genomic DNA and Isha Nagpal for helping in docking studies. We thank the Department of Biotechnology and Department of Science and Technology, Government of India for funding. IR thanks the CSIR for a fellowship and SK thanks the UGC for a fellowship.

References

Bonner, E. R., Cahoon, R. E., Knapke, S. M. & Jez, J. M. (2005). *J. Biol. Chem.* **280**, 38803–38813.
Burkhard, P., Rao, G. S., Hohenester, E., Schnackerz, K. D., Cook, P. F. & Jansonius, J. N. (1998). *J. Mol. Biol.* **283**, 121–133.
Campanini, B., Speroni, F., Salsi, E., Cook, P. F., Roderick, S. L., Huang, B., Bettati, S. & Mozzarelli, A. (2005). *Protein Sci.* **14**, 2115–2124.
Comini, M. A., Guerrero, S. A., Haile, S., Menge, U., Lünsdorf, H. & Flohé, L. (2004). *Free Radic. Biol. Med.* **36**, 1289–1302.
Croft, S. L., Sundar, S. & Fairlamb, A. H. (2006). *Clin. Microbiol. Rev.* **19**, 111–126.
DeLano, W. L. (2002). *PyMOL.* http://www.pymol.org.
Desjeux, P. (2001). *Trans. R. Soc. Trop. Med. Hyg.* **95**, 239–243.
Dumas, C., Ouellette, M., Tovar, J., Cunningham, M. L., Fairlamb, A. H., Tamar, S., Olivier, M. & Papadopoulou, B. (1997). *EMBO J.* **16**, 2590–2598.
Duszenko, M., Mühlstädt, K. & Broder, A. (1992). *Mol. Biochem. Parasitol.* **50**, 269–273.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Francois, J. A., Kumaran, S. & Jez, J. M. (2006). *Plant Cell*, **18**, 3647–3655.
Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P. & Shenkin, P. S. (2004). *J. Med. Chem.* **47**, 1739–1749.
Heine, A. *et al.* (2004). *Proteins*, **56**, 387–391.
Huang, B., Vetting, M. W. & Roderick, S. L. (2005). *J. Bacteriol.* **187**, 3201–3205.
Krauth-Siegel, R. L. & Comini, M. A. (2008). *Biochim. Biophys. Acta*, **1780**, 1236–1248.
Krieger, S., Schwarz, W., Ariyanayagam, M. R., Fairlamb, A. H., Krauth-Siegel, R. L. & Clayton, C. (2000). *Mol. Microbiol.* **35**, 542–552.
Kumar, S., Raj, I., Nagpal, I., Subbarao, N. & Gourinath, S. (2011). *J. Biol. Chem.* **286**, 12533–12541.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lukes, J. *et al.* (2007). *Proc. Natl Acad. Sci. USA*, **104**, 9375–9380.
Morris, R. J., Perrakis, A. & Lamzin, V. S. (2002). *Acta Cryst.* D**58**, 968–975.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Saito, K., Yokoyama, H., Noji, M. & Murakoshi, I. (1995). *J. Biol. Chem.* **270**, 16321–16326.
Salsi, E., Bayden, A. S., Spyrakis, F., Amadasi, A., Campanini, B., Bettati, S., Dodatko, T., Cozzini, P., Kellogg, G. E., Cook, P. F., Roderick, S. L. & Mozzarelli, A. (2010). *J. Med. Chem.* **53**, 345–356.
Schnell, R., Oehlmann, W., Singh, M. & Schneider, G. (2007). *J. Biol. Chem.* **282**, 23473–23481.
Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.
Sundar, S., More, D. K., Singh, M. K., Singh, V. P., Sharma, S., Makharia, A., Kumar, P. C. & Murray, H. W. (2000). *Clin. Infect. Dis.* **31**, 1104–1107.
Tai, C.-H., Nalabolu, S. R., Jacobson, T. M., Minter, D. E. & Cook, P. F. (1993). *Biochemistry*, **32**, 6433–6442.
Thakur, C. P., Dedet, J. P., Narain, S. & Pratlong, F. (2001). *Trans. R. Soc. Trop. Med. Hyg.* **95**, 187–189.
Turrens, J. F. (2004). *Mol. Aspects Med.* **25**, 211–220.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Williams, R. A., Westrop, G. D. & Coombs, G. H. (2009). *Biochem. J.* **420**, 451–462.
Wyllie, S., Cunningham, M. L. & Fairlamb, A. H. (2004). *J. Biol. Chem.* **279**, 39925–39932.