Research Paper

The Cytochrome P450 Complement (CYPome) of *Leishmania* Leads to the Discovery of a Plant like Cytochrome P450 Sub-family CYP710C1 Gene

RUBY BANSAL¹, HIMANSHU BHUSAN SAMAL², V S GOWRI¹, SHIB SANKAR SEN¹, INDIRA GHOSH² and RENTALA MADHUBALA^{1,*}

¹School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067, India

²School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi 110 067, India

(Received on 14 February 2017; Revised on 04 March 2017; Accepted on 07 April 2017)

Cytochrome P450 (CYP) is a super-family of heme-containing monooxygenases and is involved in the metabolism of endogenous xenobiotic compounds. A genome-wide analysis of the complete cytochrome P450 complement (CYPome) in *Leishmania* species is presented here. Genome database search algorithms of 20 strains of *Leishmania* species and 3 strains of *Trypanosoma brucei* was employed to describe the complete CYPome. Motif search analysis and phylogenetic studies were carried out to investigate the sequence diversity and distribution of different CYPs in *Leishmania* in comparison to fungi, humans, plants and prokaryotes. *In silico* analysis predicted the presence of genes belonging to CYP51 and a plant-like CYP710C gene that encodes a plant-like sterol C-22 desaturase, a key enzyme in stigmasterol biosynthesis. This is the first report for the comparative analysis of stigma sterol biosynthetic pathway genes in *Leishmania donovani*, *Arabidopsis thaliana* and *Candida albicans*. We provide experimental evidence of stigmasterol presence in *L. donovani* promastigotes. We further demonstrate that amphotericin B-resistant *L. donovani* accumulated stigmasterol as the major sterol and ergosterol to a lesser extent. In conclusion, this study is presumably the first comprehensive report on CYPome of *Leishmania* and molecular evidence of a plant-like sterol C-22 desaturase gene in *Leishmania*.

Keywords: *Leishmania*; Cytochrome P450 Complement; CYPome; C-22 Desaturase; Stigmasterol

Introduction

Cytochrome P450s (CYPs) are present in all three domains of life and constitute a super-family of heme-containing monooxygenases. CYPs are involved in the metabolism of endogenous and xenobiotic compounds (Bernhardt, 2006; Doddapaneni *et al.*, 2005; Kelly *et al.*, 2009; Moktali *et al.*, 2012).

The CYP genome complements (CYPomes) of several species is available on the Cytochrome P450 database (<http://drnelson.uthsc.edu/CytochromeP450.html>) (Nelson, 2009). The Cytochrome P450 nomenclature committee has named all newly identified CYPs, using the standard convention for this gene super-family. According to the accepted nomenclature, a family and sub-family use a numeral

and a letter, respectively. For instance, CYP51E refers to the family 51 and sub-family E.

CYP51 is involved in sterol biosynthesis and is reported as housekeeping CYP in fungi and is an important target for anti-fungal drugs (Becher and Wirsal, 2012; Kelly *et al.*, 2009). CYP51s are found in sterol-producing animals, plants, protists, but rarely in bacteria, producing 14- α -demethylated sterols (Kelly SL, 2005). The CYP51 reaction occurs in three steps; each reaction requires one molecule of oxygen and two NADPH-derived reducing equivalents (Akhtar *et al.*, 1978; Aoyama *et al.*, 1984; Shyadehi *et al.*, 1996). Another cytochrome P450 CYP61 (sterol 22 desaturase) represents an ancient activity, as the super-family is present as CYP710 in plants

*Author for Correspondence: E-mail: rentala@outlook.com; madhubala@mail.jnu.ac.in

(Morikawa *et al.*, 2006a). CYP710 plant P450s are classified as putative C-22 desaturase that produce stigmaterol and brassicasterol/crinosterol from β -sitosterol and 24-epi-campesterol respectively (Morikawa *et al.*, 2006b).

Isoprenoid derivative sterols are biologically the most important lipids present in all eukaryotes. They play a crucial role in signalling behaviour and are critically involved in maintaining membrane integrity in eukaryotes (Edwards and Ericsson, 1999). The sterol biosynthetic pathway has been well studied in fungi, animals and land plants. Trypanosomatids (e.g. *Trypanosoma* and *Leishmania*) synthesize ergosterol, the major sterol present in fungi that are involved in growth and maintenance of viability in parasites (de Souza and Rodrigues, 2009).

Ergosterol in fungi and stigmaterols in land plants are “22-unsaturated sterols that are structurally different from sterols of animal origin (cholesterol). Sterol C-22 desaturase plays an important role as a terminal enzyme in the sterol biosynthesis in fungi and plants. Fungal cytochrome P450 monooxygenase, CYP61, is the sterol C-22 desaturase, the penultimate enzyme in the ergosterol biosynthetic pathway (Desmond and Gribaldo, 2009; Lepesheva and Waterman, 2007). In the case of plants, the first cyclization of 2,3-oxidosqualene is mediated by cycloartenol synthase and CYP710 (a C-22 desaturase) for C-22 desaturation (Morikawa *et al.*, 2006b). Interestingly, *Leishmania* and *Trypanosomes* being non-photosynthetic protozoans not only synthesize lanosterol, but also contain CYP710C (*L. major*, XP_001684965) related genes (Bach TJ, 2012; Kelly and Kelly, 2013). The presence of CYP710C related gene suggests a possible role of these C-22 desaturases in the synthesis of stigmaterol in these parasites.

Most importantly, the composition of sterols inside fungi and *Leishmania* is a major determinant of the action of anti-fungal/anti-leishmanial polyene antibiotic amphotericin B (AmB). AmB binds with its mycosamine appendage to ergosterol (Palacios *et al.*, 2011), a major sterol in fungi, *Leishmania* and *Trypanosomes*. This binding leads to the disruption of membrane integrity coupled with an extensive manipulation in redox balance resulting in the induction of cell death (Gray *et al.*, 2012).

In the present study, we have utilized the existing genomic information in combination with different bioinformatic tools to better understand the CYPome of *Leishmania*. This study includes extensive genome database searches and exploitation of a broad bioinformatics pipeline of 20 different strains of *Leishmania* species and 3 strains of *Trypanosoma brucei* in order to identify all CYP members (CYPome). The characteristic motifs, phylogeny and specific protein functions of these CYP proteins were also analyzed. We also report the presence of CYP710C1P450 family gene that encodes a plant-like sterol C-22 desaturase, leading to stigmaterol biosynthesis in *Leishmania*. We further demonstrated that AmB-resistant strain accumulated stigmaterol as the major sterol.

Materials and Methods

Sequence Data

Genome database analysis was done using a total of 23 strains/species (20 different *Leishmania* strains/species and 3 strains of *Trypanosoma brucei*) (Table S1).

Identification of CYP Genes

The identification process of the CYP genes in *Leishmania* was a two-step process of identification and validation (Fig. S1). The identification of CYP genes was performed using HMMER v3.1b1, <http://hmm.janelia.org/>) with Profile hidden Markov models derived from the Pfam seed alignment flatfile of PF00067 downloaded from the Pfam protein families database, (<http://pfam.xfam.org/>) against the selected *Leishmania* proteomes. The cut off of positive hits was set at E-value 10^{-3} . Supplementary File S9 represents all the identified CYP protein sequences (20 strains of different *Leishmania* strains/species and 3 strains of *Trypanosoma brucei*). The positive hits were then subjected to a validation step involving hmm search using *Leishmania*-specific profile (LeishHmmCYP).

The sterol biosynthesis pathway information of *Arabidopsis thaliana*, *Candida albicans* and *Leishmania donovani* was extracted from the KEGG metabolic pathway database (http://www.genome.jp/kegg-bin/show_pathway?map01100). A similar pathway was created by extracting the enzymes and

their corresponding product information.

Motif Search

Motif search analysis (MSA) of all the identified *Leishmania* CYPs was built by MUSCLE v3.8.31 program (Edgar, 2004) and the consensus logos of the alignment generated by WebLogo (<http://weblogo.threeplusone.com/create.cgi>) were used for visualization of conserved amino acids for each position (Crooks *et al.*, 2004; Schneider and Stephens, 1990).

Construction of Phylogenetic Tree

Alignment of all the identified *Leishmania* and *Trypanosoma brucei* CYPs was performed by MUSCLE v3.8.31 program. The phylogenetic trees from alignments of protein sequences were constructed by FastTree v2.1.7 with maximum-likelihood method (<http://www.microbesonline.org/fasttree/>) (Price *et al.*, 2010). The tree data were then submitted to iTOL (<http://itol.embl.de/upload.cgi>) for viewing phylogenetic trees and making figures (Letunic and Bork, 2007).

Parasite and Culture Conditions

L. donovani clone AG83 (MHOM/IN/80/AG83) (WT) was used in the present study. The parasites were routinely cultured at 25°C in modified M-199 medium (Sigma, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (FBS; Gibco/BRL, Life Technologies, Scotland, UK) and 0.13 mg/ml penicillin and streptomycin. Cloning of the wild-type (WT) strain of *L. donovani* was carried out on a semi-solid M199 containing 1% Bacto Agar (Difco), 10% FBS and 0.13 mg/ml penicillin and streptomycin. Colonies were picked, inoculated and propagated in liquid M199 medium.

Generation of Amphotericin B-resistant Lines

The cloned wild-type parasites were grown in M199 media (with supplements) in the presence of 40 nM of AmB. The parasite cultures were stabilized for six sub-cultures before increasing the drug concentration to one passage per week. The line propagating in 40 nM of AmB was plated on M199 agar plates in presence of the drug, and a single colony was picked for culture. The cultures were subjected to step-wise increasing drug pressure with a final concentration of

80 nM of AmB. Cells from the culture were again cloned on semi-solid medium in presence of the drug. The clone was named AmB R80 and was resistant to 80 nM of AmB. AmB-resistant line AmB R80 showed an IC₅₀ (50% inhibitory concentration) of 145 nM (making it 3.4 fold less sensitive to AmB than the wild-type strain AG83). Repeated cloning and selection process over a period of time resulted in the selection of AmB-resistant lines, AmB R200 (IC₅₀ 248 nM) and AmB R400 (IC₅₀ 407 nM) that were 5.7 fold and 9.5 fold less sensitive to AmB than the wild-type strain AG83. Exposure of the wild-type AG83 cell population to progressively increasing concentrations of AmB resulted in the selection of a strain AmB R400, which was capable of proliferation in 400 nM of AmB. Resistant parasites were propagated for more than 50 generations in presence of the requisite drug concentration to stabilize the culture and were used for subsequent studies.

Stability of the AmB-resistant Lines

Parasites were cultured for two months in the absence of drug pressure, and the IC₅₀ values were determined after 4, 6 and 8 weeks of sub-culture.

MTT Assay

The 50% inhibitory concentration (IC₅₀) was determined. Promastigotes were seeded into 96 well plates at a density of 5x10⁶ cells/ml in 200 µl of the M199 media with different drug dilutions. The culture was incubated for the next 72 h at 22°C followed by addition of 20 µl of MTT solution. The formazan crystals were solubilized, by the addition of 50 µl of solubilization buffer and optical density of the solution was measured at 570 nm. The percentage viability was calculated by comparing with untreated parasite cultures. The experiments were performed in triplicate and data are expressed as the mean ± standard deviation (SD).

Extraction of Sterols

The wild-type parasites were grown in M199 media while the AmB-resistant parasites were cultured in presence of the requisite drug concentrations. Stationary phase promastigotes (10⁹ parasites) were harvested (3000 rpm for 5 min at 4°C), and the pellets were repeatedly washed using ice-cold 1X PBS (pH

7.4). The resulting cell pellet was re-suspended in 20 ml dichloromethane: methanol solvent (2:1 vol/vol), vigorously mixed and incubated for 24 h at 4°C. The suspension was then filtered using Whatman filter paper to separate proteins. To each filtrate, 0.2 volume of 0.9% sodium chloride solution was added to facilitate the removal of non-lipid contaminants into the aqueous phase. The resulting turbid suspension was allowed to stand for 30 min at RT for phase separation, which was followed by a collection of lower dichloromethane phase in new tubes. The extracts were completely evaporated in a stream of nitrogen gas and were immediately stored at -20°C. These dried lipid residues were subsequently used for sterol extraction (BA., 1967; Galli and Maroni, 1967).

For sterol extraction, dried lipid residues were heated at 80°C for 2 h for saponification using 30% ethanolic KOH. Sterol extraction was performed using hexane followed by drying the samples under a liquid nitrogen stream. The dried residues were dissolved in dichloromethane followed by addition of 2 volumes of bis (trimethylsilyl) trifluoroacetamide (BSTFA). Tubes containing the aliquots were sealed and heated to 80°C in water bath for 1 h. The trimethylsilyl ethers of sterols were subjected to GC/MS analysis.

Measurement of Sterol Levels Using Gas Liquid Chromatography Coupled to Mass Spectrometry (GC/MS)

For measurement of sterols, gas-liquid chromatography was performed using Shimadzu TD 1020 GC Mass Spectrometer QP2010 Plus that was equipped with DB5 columns (methyl/phenylsiloxane ratio, 95/5; dimension 30 m by 0.25 mm). The gas carrier was Helium (1ml/min). The column was kept at 270°C, and the injector and detectors were maintained at 300°C during analysis. The linear gradient for methyl esters was maintained from 150°C to 180°C at 10°C/min. MS conditions were 280°C, 70 eV and 2.2 kV.

Results

Identification of CYPome of *Leishmania*

The protein sequences of 20 different species/strains of *Leishmania* and 3 of *Trypanosoma brucei* were used to analyse the CYPome (Table S1). The Pfam

seed alignment file for Cytochrome P450 family (PF00067) from the Pfam database was used to build a profile using hmmbuild (a module of HMMER v3.1b1). The identification step was performed using hmmsearch (HMMER v3.1b1) to find out the possible CYP members based on the profile generated by hmmbuild (generic HMMCYP) (Fig. S1). A total of 70 Cytochrome P450 proteins were identified from 20 different *Leishmania* genomes using NCBI, UniProt and TriTrypDB database (Table S1). Most of the species had 4 CYPs. Genomes of *L. donovani* and *L. panamensis* had 3 CYPs each while *L. amazonensis* and *Leishmania donovani* had one CYP each. Interestingly, *L. turanica* strain had 5 CYPs. A total of six CYP 450s were identified in 3 different *Trypanosoma brucei* genomes. The 3 *Trypanosoma brucei* strains had 2 CYPs each (Table S1).

In addition, a *Leishmania*-specific profile (LeishHmCYP) was generated to find any new or missing CYP proteins from the *Leishmania* genome. Multiple sequence alignment (MSA) of all 70 CYPs using MUSCLE v3.8.31 program was built. Based on the alignment file the profile (LeishHmCYP) has been constructed by hmmbuild. Genomes of 20 different strains of *Leishmania* sp and 3 different strains of *Trypanosoma brucei* sp were searched using hmmsearch, for the presence of any new CYP members using LeishHmCYP profile (Fig. S1). This method again identified the same number of CYPs as reported in Table S2, thereby further validating the profile.

Cytochrome P450 database (<http://drnelson.uthsc.edu/CytochromeP450.html>) search for the reported CYPs in *Leishmania* strains/species is summarized in Table S3. The reported database indicates that some of the CYPs have incomplete sequence information.

Only two annotated cytochrome P450 proteins from *Leishmania* having accession number, DQ267494 (CYP5122A1) and UniProt ID: A2TEF2 (X-ray structure available in PDB (<http://www.rcsb.org/pdb/>) (PDB ID: 3L4D) (Hargrove et al., 2011; Verma et al., 2011) have been reported in literature.

Classification of CYPs

In the next step, we classified the above-identified

list of CYPs of *Leishmania* and *Trypanosoma brucei*. In the absence of available seed alignment files for CYP51, CYP61 and CYP710, we first generated three individual seed alignment files specific for CYP51, CYP710 and CYP61. This was done by MSA using only annotated CYP51, CYP710 and CYP61 sequences from many model organisms (parasites not included) (Table S4-S6). Based on the MSA, three individual profiles were built using hmmbuild (HMMER v3.1b1) namely, CYP51 (CYP51proHMM), CYP710 (CYP710proHMM) and CYP61 (CYP61proHMM). An extensive search was then performed using hmmsearch to find individual CYP51, CYP710 and CYP61 proteins amongst the total of 70 proteins identified in different *Leishmania* strains/species. CYP51 and CYP710 were identified using the above profile (Table S2). The CYP61 profile hits all the CYP710 class of proteins. Analysis of all the 6 proteins from the *Trypanosoma* strains led to the identification of CYP51 and none as CYP710/CYP61.

Further, using Cytochrome P450 nomenclature (<http://drnelson.uthsc.edu/CytochromeP450.html>), we report different classes of CYPs (CYP51E1, CYP710C1, CYP5123A1 and CYP5122A1) in various strains/species of *Leishmania* and *Trypanosome* (Table S2). Table S2 also summarizes the classification of the identified CYPs, chromosome location, and predicted sub-cellular localization of the proteins and probable function(s) of the CYPs

Motifs

The family diversity of CYP genes differs considerably between species except for some well-conserved regions. This family of proteins are reported to have four conserved regions (Deng *et al.*, 2007; Kelly *et al.*, 2009; Werck-Reichhart and Feyereisen, 2000). A unique motif FXXGXRXCXG (located at position d in Fig. 1) corresponds to the heme-binding domain. The second conserved motif EXXR (located at position b in Fig. 1) and the third consensus PER (located at position c in Fig. 1) form E–R–R triad that is necessary for locking the heme pocket into position and assured stabilization of the core structure. The fourth relatively conserved motif AGXDTT (located at position a in Fig. 1) contributes to oxygen binding and activation.

In the present study, sequence logo of all the four conserved motifs was generated from the

identified list of CYPs in *Leishmania*. A comparison of conserved motifs of *Leishmania* with the motifs of fungi, human, plant and prokaryotes (Chen *et al.*, 2014) was made. Our results show that in spite of high conservation of residues, there were some noticeable differences amongst the motifs. Comparison of the motif “a” (oxygen binding) of *Leishmania* with fungi, human, plant and bacterial counterparts was made. It was observed that *Leishmania* motif “a” is similar to that of plant and fungi. Ala, Gly, Asp and Thr (AGXDTX) are the most conserved residues in this motif. In the case of *Leishmania*, the third position of this motif contains His and is very similar to archaea and bacteria. A noticeable difference was the presence of Ser in place of Thr at the sixth position, showing similarity with plants. Motif “b” of *Leishmania* remains well conserved (EXXR) and shows similarity with all the four kingdoms. Motif “c” (Heme locking region) of *Leishmania* is similar to those of plants and fungi. Motif “c” (FXPERF/W) shows conservation of Phe, Pro, Glu and Arg in *Leishmania*. The second position is not conserved, while in the sixth position Phe is replaced by Trp as in the case of fungi. Comparison of *Leishmania* motif “d” showed that the sequence FXXGXRXCXG is conserved and shows similarity to archaea and bacteria. The only noticeable difference was the frequent presence of His in place of Arg in FGXGXR/HXCXG. The second position (G/S) is not highly conserved whereas the first, fourth, eighth and tenth positions (Phe, Gly, Cys, Gly) are highly conserved.

We also generated the sequence logo for CYP51 (CYP51proHMM), CYP710 (CYP710proHMM) and CYP61 (CYP61proHMM) profiles that were used for classification of the total CYPome (Fig. 1). CYP51 profile showed that the motif “a” was similar to *Leishmania* except for the presence of His in place of Asp in AGQD/HTSS. In the third position, Gln is dominant over His. In motif “b” the sequence EXXR is highly conserved. Motif “c” shows similarity to *Leishmania* motif “c” except for the presence of Asp in place of Glu i.e., FXPE/DR. Motif “d” showed greater similarity to *Leishmania* except at the second and sixth positions where the Gly and His were present as invariant residues i.e., FGXGXHXCXG. Comparison of motifs of the CYP710 profile to *Leishmania* motifs identified many differences. In motif “a”, the Gly is replaced with Ala and Thr is

Fig. 1: Sequence logos of the conserved CYP motifs of *Leishmania* and their comparison with fungi, human, plant and prokaryotes. The figure also illustrates a comparison with three individual profiles, i.e. CYP51 profile (CYP51proHMM), CYP710 profile (CYP710proHMM) and CYP61 profile (CYP61proHMM))

replaced with Ala i.e. AG/AQDT/AST. Motif “b” was highly conserved. However, when compared with the other two profiles, Thr in the second position is replaced with Val i.e., ET/VLR. Motif “c” had an insertion of an amino acid between Arg and Phe and also showed the frequent presence of Asp in place of Glu i.e. FXPE/DRF. The sequence FXXGXHCXG of motif “d” remains conserved when compared to the other two profiles, but comparison with *Leishmania* showed greater similarity except at the second and the sixth position where Gly and His were present as invariant residues. Comparison of the CYP61 profile with a CYP710 profile of *Leishmania* showed that as far as motif “a” is concerned, the only difference observed was the replacement of amino acid Gly with

Ser i.e. AG/SXDT/AST. The motif “b” and “c” remain well conserved and were highly similar to CYP710 of *Leishmania*. Motif “d” is also well conserved and shows greater similarity with the other two profiles.

Phylogenetic Tree

Alignment of all the identified CYPs was performed by MUSCLE v3.8.31 program and the likelihood phylogenetic tree for *Leishmania* and *Trypanosoma brucei* is shown in Fig. 2. Four clusters of CYPs belonging to CYP51E1, CYP710C1, CYP5123A1 and CYP5122A1 were observed (Fig. 2). *Leishmania* and *Trypanosoma* species were represented in both CYP51E1, CYP5122A1 cluster, whereas CYP710C1

The un-rooted phylogenetic tree shows the

distribution of CYP710, CYP51 of *Leishmania* and CYP51 of *Trypanosoma* and their comparison with fungi, plant, bacteria and animal counterparts (Fig. 3). *Leishmania* CYP710 and plant CYP710 reside in one branch, further indicating that they are highly similar. They were also found to be similar to CYP61 of fungi indicating their common ancestry. *Leishmania* CYP51 and *Trypanosoma* CYP51, on the other hand, were present in one branch and were

Fig. 3: Phylogeny of *Leishmania* CYP51s, CYP710s and *Trypanosoma brucei* CYP51 and their counterparts from animal, plant, bacteria and fungi. This is based on 19 animal CYP51s, 10 bacteria CYP51s, 42 plant CYP51s, 23 fungal CYP51, 27 fungal CYP710 and 30 plant CYP710s (extracted from the CYP database (<http://drnelson.uthsc.edu/CytochromeP450.html>))

found to be closely related to bacteria and plant CYP51. *Leishmania* and *Trypanosoma* CYP51 are grouped in one cluster along with bacteria, plant, animal and fungi CYP51 indicating that they may have a common ancestor.

Comparative Computational Analysis of Sterol/Stigmasterol Biosynthetic Pathway Genes of *Leishmania donovani*

A comparison of stigmasterol biosynthesis pathway of *Arabidopsis thaliana*, *Candida albicans* and *Leishmania donovani* was made using information extracted from KEGG metabolic pathway database (http://www.genome.jp/kegg-bin/show_pathway?map01100) (Fig. 4, Supplementary Table S7). Sterol synthesis starts with the formation of farnesyl-PP from the condensation reaction of isoprenoids, catalyzed by the enzyme farnesyl diphosphate synthase.

Farnesyltransferase (LDBPK_313050) promotes a head-to-head condensation of two molecules of farnesyl diphosphate to produce squalene (Fig. 4). Production of squalene then drives the synthesis of squalene-2,3-epoxide, a reaction catalyzed by the enzyme squalene monooxygenase (LDBPK_131360). Squalene-2,3-epoxide (30-carbon linear chain) is then cyclized in the presence of lanosterol synthase (LDBPK_060670) and sterol-24-c methyltransferase (LDBPK_362510) to form lanosterol and 24-methylene cycloartenol respectively. In the case of plants, cycloartenol is formed from squalene-2,3-epoxide with the help of cycloartenol synthase (AT2G07050), and then cycloartenol undergoes demethylation at the C-24 position to form 24-methylene cycloartenol in the presence of cycloartenol-C-24-methyltransferase (AT5G13710). In the case of *Leishmania donovani* and *Candida*

Fig. 4: Comparison of sterol biosynthesis pathway of *Leishmania donovani*, *Arabidopsis thaliana* and *Candida albicans*. Light green colour indicates enzyme involved in the plant (*Arabidopsis thaliana*), light yellow colour indicates enzyme involved in fungi (*Candida albicans*) and light purple colour indicates *Leishmania donovani*. Information was taken from KEGG Metabolic Pathway Database (www.genome.jp/kegg-bin/show_pathway?map01100). The enzyme nomenclature is reported in Table S7

albicans, squalene-2,3-epoxide is directly converted to 24-methylene cycloartenol through the action of sterol 24-c-methyltransferase and delta-24 methyltransferase respectively (Fig. 4).

After the cyclization of squalene-2,3-epoxide to form lanosterol, several sequential transformations occur to form ergosterol and 7-dehydrocholesterol. One of the earliest steps in the lanosterol pathway is the demethylation of the ring system at the C14 position in a two-stage oxidative reaction catalyzed by a cytochrome P-450-containing monooxygenase enzyme known as lanosterol-14 α -methyl demethylase (LDBPK_111100/CYP51E) to form 4,4-dimethyl cholesta-8,14,24-trienol, which then undergoes reduction reaction catalyzed by the enzyme C-14 sterol reductase (LDBPK_322470) with NADPH as cofactor to form 14-demethyl lanosterol. It is then converted to fecosterol by the enzyme sterol 24-c-methyltransferase (LDBPK_362510, adds methyl group at the C-24 position). This is followed by isomerisation of the double bonds in fecosterol, in a reaction catalyzed by C-8 sterol isomerase (LDBPK_292250) to form episterol. Episterol further undergoes reduction reaction to form ergosterol. Here, 5,7,24(28)-ergostatrienol is formed from episterol by the enzyme lathosterol oxidase (LDBPK_231560). 5,7,24(28)-ergostatrienol is directly converted to ergosterol by sterol C-24 reductase (LDBPK_330730). Ergosterol is absent in *Arabidopsis thaliana* whereas, in *Candida albicans*, ergosterol is the main sterol. In the case of *Candida albicans*, 5,7,24(28)-ergostatrienol forms 5,7,22,24(28)-ergostetraenol by the enzyme C-22 sterol desaturase (CaO19.12645/CYP61), which then forms ergosterol. However, in the case of *Leishmania donovani*, this step is absent.

In comparison to *L. donovani*, delta (14) sterol reductase catalyzes the reaction to form 14-demethyl lanosterol in plants and fungi as well. 14-demethyl lanosterol is converted to 4-methyl zymosterol carboxylate in the presence of C-4 methyl sterol oxidase in *Candida albicans*. However, this reaction step is absent in *L. donovani* as well as in *A. thaliana*. 4-methyl zymosterol carboxylate is converted to 3-keto-4-methyl zymosterol by the enzyme 3 β -hydroxysteroid-dehydrogenase in *A. thaliana* and C-3 sterol dehydrogenase in *Candida albicans*. 3-keto-4-methyl zymosterol formation drives the synthesis of 4-methyl zymosterol to zymosterol in

the presence of 3-keto sterol reductase, which causes demethylation at C-4 in *C. albicans* which is also absent in *L. donovani*.

In *Arabidopsis thaliana*, zymosterol is converted to lathosterol in two consecutive steps. The first step involves reduction at C-24 by the enzyme delta-24-sterol reductase to form cholesta-8-en-3 β -ol which then undergoes isomerisation reaction to form lathosterol through the action of delta-7-isomerase. In *Candida albicans* and *L. donovani*, lathosterol is not formed. 7-dehydrocholesterol which functions as a precursor of cholesterol is synthesized in the presence of the enzyme lathosterol oxidase (LDBPK_231560) in the case of *L. donovani* and is also present in plants and fungi and is catalyzed by delta(7)-sterol-C5-(6)-desaturase-1 and sterol C5, 6 desaturases. In *arabidopsis*, cholesterol formation is driven by 7-dehydrocholesterol reductase, which can be converted to cholesterol ester in the presence of steryl ester hydrolase (in the case of both *Arabidopsis* and *Candida*). In *L. donovani* cholesterol is absent, as the enzyme 7-dehydrocholesterol reductase responsible for the conversion is not present. Desmosterol is a molecule similar to cholesterol. Desmosterol is an immediate precursor of cholesterol (Vainio et al., 2006). Delta(24)-sterol reductase catalyzes the reduction of desmosterol to cholesterol and vice versa. However, desmosterol is absent in both *L. donovani* and *Candida albicans*.

24-methylene cycloartenol undergoes C-14 demethylation reaction to form delta-8,14-sterol, catalyzed by the enzyme lanosterol 14-alpha-demethylase (LDBPK_111100/CYP51E). In contrast to *L. donovani*, *arabidopsis* contains enzymes sterol-4alpha-methyl oxidase and cyclopropyl isomerase that catalyzes the formation of cycloeucalenol and obtusifolol. Obtusifolol is then converted to delta-8,14-sterol in the presence of sterol-14-demethylase. Delta-8,14-sterol drives the formation of 4alpha-methylfecosterol in the presence of C-14 sterol reductase (LDBPK_322470). 4alpha-methylfecosterol leads to desaturation reaction at the C-5 position to form 5-dehydroavenasterol in the presence of lathosterol oxidase (LDBPK_231560) in *L. donovani*. However, in the case of *Arabidopsis thaliana*, at least three products are formed from 4alpha-methylfecosterol prior to the formation of 5-

dehydroavenasterol, these are 24-methylenelophenol, 24-ethylidene lophenol and delta-7-avenasterol catalyzed by the enzymes delta-7-isomerase, 24-methylenesterol C-methyltransferase and sterol 4-alpha-methyl-oxidase. Sequential transformation of 24-methylenelophenol leads to the formation of campesterol. Episterol formation is catalyzed by the sterol 4-alpha-methyl oxidase, which is present only in the plants. Episterol is converted to 5-dehydroepisterol by the enzyme delta(7)-sterol-C5(6)-desaturase. This is then converted to 24-methylene cholesterol and campesterol in the presence of the enzyme 7-dehydrocholesterol reductase and delta-24-sterol reductase. Campesterol formation is absent in *Leishmania donovani* and *Candida albicans*. However, in both these cases, 5-dehydroepisterol is directly formed from 4alpha-methylfecosterol by the enzyme lathosterol oxidase (LDBPK_231560) and sterol C5, 6 desaturase.

Lathosterol oxidase also catalyzes the formation of 5-dehydroavenasterol. Cytochrome P450 (CYP710C), a hemeprotein catalyzes the formation of stigmasterol from 5-dehydroavenasterol in *L. donovani* (Fig. 4). However, in *Arabidopsis thaliana*, a three-step reaction leads to the formation of stigmasterol, from 5-dehydroavenasterol. Enzymes involved in the conversion are 7-dehydrocholesterol reductase, delta-24-sterol reductase and cytochrome P450 (CYP710A). Stigmasterol is absent in *Candida albicans*. Instead, 5-dehydroavenasterol is present.

Stigmasterol is the Major Sterol in Amphotericin B-resistant *L. donovani*

AmB was found to inhibit the wild-type AG83 in a dose-dependent manner (IC_{50} : 43 ± 2.3 nM). (Table S8). AmB-resistant promastigotes of *L. donovani* were obtained as discussed in the methods section. Three lines resistant to AmB designated as AmB R80, AmB R200 and AmB R400 displayed approximately 3.4-fold (IC_{50} : 147 ± 3.0 nM), ~ 5.5-fold (IC_{50} : 248 ± 2.5 nM), ~ 9.5-fold (IC_{50} : 407 ± 6.0 nM) resistance respectively to AmB. (Table S8). The growth pattern of the wild-type AG83 and AmB-resistant parasites was determined over a period of 5 days. AmB-resistant parasites showed same doubling time as compared to the wild-type AG83 (data not shown). The drug-resistance phenotype was stable as parasites when

grown in the absence of drug for several generations retained the resistance phenotype. We investigated the cross-resistance profile of the AmB-resistant strain (AmB R200) to different anti-leishmanial drugs. There was no significant change in the sensitivity profile of AmB R200 strain to sodium antimony gluconate (SAG), pentamidine, paromomycin and miltefosine, when compared to the wild-type strain (data not shown).

GC/MS analysis of the wild-type AG83 and AmB R400 strain showed that ergosterol (21.4%) and its isomer ergosta-5, 7, 22-trien-3-beta-ol (70%) constitute more than 90% of total ergosterol in the wild-type *Leishmania* while ergosterol and its isoforms were largely undetected in the AmB-resistant parasites (Table 1). The amount of stigmasterol (4.9%) and its isoform stigmasta-5, 22-dien-3-ol (~ 56.5%) were significantly higher in the AmB R400 strains, as compared to the wild-type AG83 (Table 1). There was no significant difference in the composition in cholesterol and its isoforms in both wild-type and AmB-resistant parasites.

Discussion

Leishmaniasis is caused by the protozoan parasites of genus *Leishmania* and affects nearly 12 million people worldwide with almost 2 million new cases per year (<http://www.who.int/tdr/diseases/leish/diseaseinfo.html>). The disease treatment completely relies on a limited repertoire of antileishmanial chemotherapy. Due to the emergence of widespread drug resistance against a limited pool of anti-leishmanial drugs, identification and characterization of novel cellular drug targets in *Leishmania* is one of the major thrust areas in protozoan research (Ouellette *et al.*, 2004; Singh *et al.*, 2012). The availability of whole-genome sequences for a number of *Leishmania* species has enabled the research community to identify novel drug targets and the presence of pathways unique to the parasite.

We report different classes of CYPs (CYP51E1, CYP710C1, CYP5123A1 and CYP5122A1) in various strains/species of *Leishmania* and *Trypanosomes*. *Leishmania* CYPs showed highly conserved characteristic motifs. The characteristic motifs of *Leishmania* CYPs were found to be similar to those of plants and fungi.

Table 1: Composition of different sterols identified using GC/MS approach in wild-type AG83 and AmB R400 strain of *L. donovani*. Results corresponding to other components were detected in at least three biological replicates and are mean concentration $\pm$ S.D. of at least two independent experiments. ND: not detected

Name of sterol and isoforms	% Composition	
	Wild-type AG83	AmB R400
Ergosterol and isoforms		
Ergosterol	21.2 $\pm$ 4.89	ND
Ergosta-5,7,22-trien-3-beta-ol	70.81 $\pm$ 2.6	ND
Ergosta-5,24(28)-dien-3-ol	0.59 $\pm$ 0.33	ND
Ergosta-tetraen-ol	1.77	ND
Ergosta-5,22-dien-3-ol	ND	0.7
Total ergosterol	94.3	0.7
Cholesterol and isoforms		
Cholesterol	1.375 $\pm$ 0.39	1.245 $\pm$ 0.46
Cholest-5-en-3-beta-ol	1.01	0.525 $\pm$ 0.16
(3-beta-5-alpha)-cholesta-8, 24-dien-3-ol	0.21	0.3 $\pm$ 0.18
Cholesta-5,24-dien-3-ol	ND	0.62
Total cholesterol	2.595	2.69
Stigmasterol and isoforms		
Stigmasterol	2.7 $\pm$ 0.79	4.29
Stigmasta-5, 22-dien-3-ol	29.67	56.47
Stigmast-5-en-3-ol	ND	0.38
Total stigmasterol	32.37	61.14

The crystal structure of CYP51 from *T. cruzi* and *T. brucei* bound to the anti-fungal drugs fluconazole and posaconazole (Chen *et al.*, 2010; Hargrove *et al.*, 2012b; Lepesheva and Waterman, 2011) and that of *L. infantum* bound to fluconazole is available (Chen *et al.*, 2010; Lepesheva and Waterman, 2011). The structure would help in designing newer drugs with improved efficacy and reduced toxicity (Chen *et al.*, 2010; Hargrove *et al.*, 2012a; Hargrove *et al.*, 2012b; Lepesheva and Waterman, 2011). *Leishmania* CYP51s and CYP710s may provide a useful model for CYP structure/function studies (Chen *et al.*, 2010; Hargrove *et al.*, 2012a; Hargrove *et al.*, 2012b; Lepesheva and Waterman, 2011).

In mammalian cells, cholesterol is the main sterol found in various membranes. However, ergosterol plays a key role in eukaryotic organisms such as fungi and protozoa. Plants usually possess more complex sterol compositions. Sitosterol and stigmasterol are the major sterols in vascular plants. Cholesterol and

ergosterol synthesis is driven by lanosterol as the key intermediate. On the other hand, cycloartenol is the key intermediate in the synthesis of stigmasterol and campesterol. In *Arabidopsis thaliana*, synthesis of stigmasterol and brassicasterol is catalyzed by two separate sterol C-22 desaturases, encoded by the genes *CYP710A1* and *CYP710A2*, respectively (Morikawa *et al.*, 2006a). These proteins belong to a small cytochrome P450 sub-family having four members, denoted by CYP710A1-A4, and are related to the yeast sterol C-22 desaturase Erg5p acting in ergosterol synthesis. Our *in silico* analysis showed the presence of a single CYP710C1 gene in *Leishmania* species. In order to identify the biological relevance of plant-like CYP710 protein in *Leishmania*, we employed a GC/MS based sterol (stigmasterol) profiling approach using wild-type and AmB-resistant strains of *Leishmania donovani*. The rationale behind including the AmB-resistance parasites in this metabolic screen served two purposes. Firstly, we wanted to check the presence of stigmasterol and its

isoforms in both the wild-type and AmB-resistant parasites. Secondly, this approach served the purpose of identification and differential abundance of different sterols (critical targets of AmB), including stigmasterol in both wild-type and AmB-resistant parasites.

The presence of stigmasterol, ergosterol and their corresponding isoforms in *L. donovani* clearly indicates that *Leishmania* has acquired both the fungus and plant pathways for sterol biosynthesis. Furthermore, AmB has been reported to have a higher affinity for ergosterol and significantly lower affinity for stigmasterol (Patterson *et al.*, 1979). Characterization of an AmB-resistant clinical isolate from Bihar (India) shows that the absence of ergosterol in the resistant parasite's membranes and the up-regulated AmB efflux and ROS scavenging machinery have a cumulative effect in conferring resistance against AmB to the *Leishmania* parasite (Purkait *et al.*, 2012). Our observation showing lack of ergosterol and higher abundance of stigmasterol and its isoforms in AmB-resistance as compared to the wild-type strain could explain the possible mechanism of resistance to AmB.

In conclusion, this study is presumably the first comprehensive report on CYPome of *Leishmania* and molecular evidence of plant like CYP710 gene in *Leishmania*. A comparison of the sterol biosynthesis pathway of *Arabidopsis thaliana*, *Candida albicans* and *Leishmania donovani* is reported, indicating the key enzymes and their corresponding product information in *L. donovani*.

This study further shows that stigmasterol biosynthetic pathway could act as a novel drug target in Leishmaniasis. CYP710C1 represents an important drug target in *Leishmania* because of its similarity to

plant-like genes. The presence of high levels of stigmasterol and its isoforms in AmB-resistant parasites in comparison to the wild-type suggests a dynamic change in the membrane properties in resistant parasites. However, the biological significance of this phenotype remains to be validated in the case of AmB-resistance. Further studies are underway to characterize the CYP710C1 gene in *Leishmania donovani* using different molecular approaches.

Acknowledgements

Rentala Madhubala is a JC Bose National Fellow supported by Department of Science and Technology, Government of India. This work was supported by the Department of Science and Technology grant and DBT-BUILDER program, Government of India. Ruby Bansal is a Senior Research Fellow of the Council for Scientific and Industrial Research, Government of India. Himanshu Bhusan Samal is supported by DBT-BUILDER program. VSG and SS are funded by UGC DS Kothari Post-doctoral Fellowship. We would like to acknowledge the usage of the Advanced Instrumentation Research Facility (AIRF) at JNU. The funding was from DST PURSE, UGC_RNW.

Abbreviations

CYPomes, CYP genome complements; HMMER, Sequence analysis using Hidden Markov Models; Pfam, Protein families' database; Leish HMM CYP, *Leishmania*-specific profile; MSA, multiple sequence alignment; AmB, amphotericin B; AmB R200, AmB-resistant strain; SAG, sodium antimony gluconate; farnesyl-PP, farnesyl diphosphate; GC-MS, Gas chromatography coupled with mass spectrometry; BSTFA, bis (trimethylsilyl) trifluoroacetamide.

References

- Akhtar M, Alexander K, Boar R B, McGhie J F and Barton D H (1978) Chemical and enzymic studies on the characterization of intermediates during the removal of the 14 α -methyl group in cholesterol biosynthesis. The use of 32-functionalized lanostane derivatives *The Biochemical journal* **169** 449-463
- Aoyama Y, Yoshida Y and Sato R (1984) Yeast cytochrome P-

450catalyzing lanosterol 14 α -demethylation. II. Lanosterol metabolism by purified P-450(14)DM and by intact microsomes *The Journal of biological chemistry* **259** 1661-1666

- BA K (1967) Identification of Plant Sterols Using Combined GLC/Mass Spectrometry *J Chromatogr Sci* **5** 273-282
- Bach T J, R M (2012) Isoprenoid Synthesis in Plants and Microorganisms: New Concepts and Experimental

Approaches

- Becher R and Wirsal S G (2012) Fungal cytochrome P450 sterol 14 α -demethylase (CYP51) and azole resistance in plant and human pathogens *Applied microbiology and biotechnology* **95** 825-840
- Bernhardt R (2006) Cytochromes P450 as versatile biocatalysts *J Biotechnol* **124** 128-145
- Chen C K, Leung S S, Guilbert C, Jacobson M P, McKerrow J H and Podust L M (2010) Structural characterization of CYP51 from *Trypanosoma cruzi* and *Trypanosoma brucei* bound to the antifungal drugs posaconazole and fluconazole *PLoS Negl Trop Dis* **4** e651
- Chen W, Lee M K, Jefcoate C, Kim S C, Chen F and Yu J H (2014) Fungal cytochrome p450 monooxygenases: their distribution, structure, functions, family expansion, and evolutionary origin *Genome Biol Evol* **6** 1620-1634
- Crooks G E, Hon G, Chandonia J M and Brenner S E (2004) WebLogo: A sequence logo generator *Genome research* **14** 1188-1190
- de Souza W and Rodrigues J C (2009) Sterol Biosynthesis Pathway as Target for Anti-trypanosomatid Drugs *Interdisciplinary perspectives on infectious diseases* **2009** 642502
- Deng J, Carbone I and Dean R A (2007) The evolutionary history of cytochrome P450 genes in four filamentous Ascomycetes *BMC evolutionary biology* **7** 30
- Desmond E and Gribaldo S (2009) Phylogenomics of sterol synthesis: insights into the origin, evolution, and diversity of a key eukaryotic feature *Genome biology and evolution* **1** 364-381
- Doddapaneni H, Chakraborty R and Yadav J S (2005) Genome-wide structural and evolutionary analysis of the P450 monooxygenase genes (P450ome) in the white rot fungus *Phanerochaete chrysosporium*: evidence for gene duplications and extensive gene clustering *BMC Genomics* **6** 92
- Edgar R C (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput *Nucleic Acids Res* **32** 1792-1797
- Edwards P A and Ericsson J (1999) Sterols and isoprenoids: signaling molecules derived from the cholesterol biosynthetic pathway *Annual review of biochemistry* **68** 157-185
- Galli G and Maroni S (1967) Mass spectrometric investigations of some unsaturated sterols biosynthetically related to cholesterol *Steroids* **10** 189-197
- Gray K C, Palacios D S, Dailey I, Endo M M, Uno B E, Wilcock B C and Burke M D (2012) Amphotericin primarily kills yeast by simply binding ergosterol *Proceedings of the National Academy of Sciences of the United States of America* **109** 2234-2239
- Hargrove T Y, Kim K, de Nazare Correia Soeiro M, da Silva C F, Batista D D, Batista M M, Yazlovitskaya E M, Waterman M R, Sulikowski G A and Lepesheva G I (2012a) CYP51 structures and structure-based development of novel, pathogen-specific inhibitory scaffolds *Int J Parasitol Drugs Drug Resist* **2** 178-186
- Hargrove T Y, Wawrzak Z, Liu J, Nes W D, Waterman M R and Lepesheva G I (2011) Substrate preferences and catalytic parameters determined by structural characteristics of sterol 14 α -demethylase (CYP51) from *Leishmania infantum* *J Biol Chem* **286** 26838-26848
- Hargrove T Y, Wawrzak Z, Liu J, Waterman M R, Nes W D and Lepesheva G I (2012b) Structural complex of sterol 14 α -demethylase (CYP51) with 14 α -methylenecyclopropyl-Delta7-24, 25-dihydrolanosterol *J Lipid Res* **53** 311-320
- Kelly D E, Krasevec N, Mullins J and Nelson D R (2009) The CYPome (Cytochrome P450 complement) of *Aspergillus nidulans* *Fungal Genet Biol* **46** S53-61
- Kelly S L, K D, Jackson C J, Warrilow A G S and Lamb D C (2005) The Diversity and Importance of Microbial Cytochromes P450. In: Montellano PRO de, ed. Cytochrome P450. Springer US, 585-617
- Kelly S L and Kelly D E (2013) Microbial cytochromes P450: Biodiversity and biotechnology. Where do cytochromes P450 come from, what do they do and what can they do for us? *Philosophical transactions of the Royal Society of London Series B, Biological sciences* **368** 20120476
- Lepesheva G I and Waterman M R (2007) Sterol 14 α -demethylase cytochrome P450 (CYP51), a P450 in all biological kingdoms *Biochimica et biophysica acta* **1770** 467-477
- Lepesheva G I and Waterman M R (2011) Sterol 14 α -demethylase (CYP51) as a therapeutic target for human trypanosomiasis and leishmaniasis *Curr Top Med Chem* **11** 2060-2071
- Letunic I and Bork P (2007) Interactive Tree of Life (iTOL): An online tool for phylogenetic tree display and annotation *Bioinformatics* **23** 127-128
- Moktali V, Park J, Fedorova-Abrams N D, Park B, Choi J, Lee Y H and Kang S (2012) Systematic and searchable classification of cytochrome P450 proteins encoded by fungal and oomycete genomes *BMC Genomics* **13** 525
- Morikawa T, Mizutani M, Aoki N, Watanabe B, Saga H, Saito S, Oikawa A, Suzuki H, Sakurai N, Shibata D et al. (2006a) Cytochrome P450 CYP710A encodes the sterol C-22

- desaturase in Arabidopsis and tomato *The Plant cell* **18** 1008-1022
- Morikawa T, Mizutani M and Ohta D (2006b) Cytochrome P450 subfamily CYP710A genes encode sterol C-22 desaturase in plants *Biochemical Society transactions* **34** 1202-1205
- Nelson D R (2009) The cytochrome p450 homepage *Hum Genomics* **4** 59-65
- Ouellette M, Drummelsmith J and Papadopoulou B (2004) Leishmaniasis: Drugs in the clinic, resistance and new developments *Drug resistance updates : Reviews and commentaries in antimicrobial and anticancer chemotherapy* **7** 257-266
- Palacios D S, Dailey I, Siebert D M, Wilcock B C and Burke M D (2011) Synthesis-enabled functional group deletions reveal key underpinnings of amphotericin B ion channel and antifungal activities *Proceedings of the National Academy of Sciences of the United States of America* **108** 6733-6738
- Patterson J, Holland J and Bieber L L (1979) Studies on the competition of polyene antibiotics for sterols *The Journal of antibiotics* **32** 646-653
- Price M N, Dehal P S and Arkin A P (2010) FastTree 2—approximately maximum-likelihood trees for large alignments *PloS one* **5** e9490
- Purkait B, Kumar A, Nandi N, Sardar A H, Das S, Kumar S, Pandey K, Ravidas V, Kumar M, De T, Singh D and Das P (2012) Mechanism of Amphotericin B Resistance in Clinical Isolates of *Leishmania donovani* **56** 1031-1041
- Schneider T D and Stephens R M (1990) Sequence logos: a new way to display consensus sequences *Nucleic acids research* **18** 6097-6100
- Shyadehi A Z, Lamb D C, Kelly S L, Kelly D E, Schunck W H, Wright J N, Corina D and Akhtar M (1996) The mechanism of the acyl-carbon bond cleavage reaction catalyzed by recombinant sterol 14 alpha-demethylase of *Candida albicans* (other names are: lanosterol 14 alpha-demethylase, P-45014DM, and CYP51) *The Journal of biological chemistry* **271** 12445-12450
- Singh N, Kumar M and Singh R K (2012) Leishmaniasis: current status of available drugs and new potential drug targets *Asian Pacific journal of tropical medicine* **5** 485-497
- Vainio S, Jansen M, Koivusalo M, Rog T, Karttunen M, Vattulainen I and Ikonen E (2006) Significance of sterol structural specificity. Desmosterol cannot replace cholesterol in lipid rafts *The Journal of biological chemistry* **281** 348-355
- Verma S, Mehta A and Shaha C (2011) CYP5122A1, a novel cytochrome P450 is essential for survival of *Leishmania donovani* *PLoS One* **6** e25273
- Werck-Reichhart D and Feyereisen R (2000) Cytochromes P450: A success story *Genome biology* **1** REVIEWS3003.