the FEBS Journal

Structural and kinetic studies on adenylosuccinate lyase from *Mycobacterium smegmatis* and *Mycobacterium tuberculosis* provide new insights on the catalytic residues of the enzyme

Sanchari Banerjee[1], Monika J. Agrawal[1,*], Diptimayee Mishra[1,*], Siddharth Sharan[1], Hemalatha Balaram[2], Handanhal S. Savithri[3] and Mathur R. N. Murthy[1]

1 Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India
2 Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre of Advanced Scientific Research, Bangalore, India
3 Department of Biochemistry, Indian Institute of Science, Bangalore, India

Keywords
adenylosuccinate lyase; catalytic activity; purine nucleotide supply; slow growth rate; X-ray crystallography

Correspondence
M. R. N. Murthy, Molecular Biophysics Unit, Indian Institute of Science, Bangalore – 560012, India
Fax: +91 80 2360 0535
Tel: +91 80 2293 2458
E-mail: mrn@mbu.iisc.ernet.in

*These authors contributed equally to this work.

(Received 20 November 2013, revised 15 January 2014, accepted 23 January 2014)

doi:10.1111/febs.12730

Adenylosuccinate lyase (ASL), an enzyme involved in purine biosynthesis, has been recognized as a drug target against microbial infections. In the present study, ASL from *Mycobacterium smegmatis* (*MsASL*) and *Mycobacterium tuberculosis* (*MtbASL*) were cloned, purified and crystallized. The X-ray crystal structure of *MsASL* was determined at a resolution of 2.16 Å. It is the first report of an apo-ASL structure with a partially ordered active site C3 loop. Diffracting crystals of *MtbASL* could not be obtained and a model for its structure was derived using *MsASL* as a template. These structures suggest that His149 and either Lys285 or Ser279 of *MsASL* are the residues most likely to function as the catalytic acid and base, respectively. Most of the active site residues were found to be conserved, with the exception of Ser148 and Gly319 of *MsASL*. Ser148 is structurally equivalent to a threonine in most other ASLs. Gly319 is replaced by an arginine residue in most ASLs. The two enzymes were catalytically much less active compared to ASLs from other organisms. Arg319Gly substitution and reduced flexibility of the C3 loop might account for the low catalytic activity of mycobacterial ASLs. The low activity is consistent with the slow growth rate of *Mycobacteria* and their high GC containing genomes, as well as their dependence on other salvage pathways for the supply of purine nucleotides.

Structured digital abstract
• purB and purB bind by x-ray crystallography (View interaction)

Abbreviations
AICAR, 5-aminoimidazole-4-carboxamide ribotide; ASL, adenylosuccinate lyase; *BaASL*, *Bacillus anthracis* ASL; *BsASL*, *Bacillus subtilis* ASL; *CeASL*, *Caenorhabditis elegans* ASL; CMLE, 3-carboxy-*cis,cis*-muconate lactonizing enzyme; *EcASL*, *Escherichia coli* ASL; ESRF, European Synchrotron Radiation Facility; *FtASL*, *Francisella tularensis* ASL; *HsASL*, *Homo sapiens* ASL; *LdASL*, *Leishmania donovani* ASL; *LpASL*, *Legionella pneumophila* ASL; *MsASL*, *Mycobacterium smegmatis* ASL; MS, *Mycobacterium smegmatis*; *MtbASL*, *Mycobacterium tuberculosis* ASL; MTB, *Mycobacterium tuberculosis*; *PaASL*, *Pyrobaculum aerophilum* ASL; PDB, Protein Data Bank; *PfASL*, *Plasmodium falciparum* ASL; *PvASL*, *Plasmodium vivax* ASL; *SaASL*, *Staphylococcous aureus* ASL; SAMP, succinyl-adenosine monophosphate; *TbASL*, *Trypanosoma brucei* ASL; *TmASL*, *Thermotoga maritima* ASL.

Introduction

Purine nucleotides are synthesized in eukaryotes and prokaryotes by two distinct pathways: the low energy costing salvage pathway and the ATP dependent *de novo* pathway. Both pathways are active in most organisms. However, in obligate protozoan parasites such as *Toxoplasma gondii*, *Plasmodium falciparum* and *Leishmania donovani*, only the salvage pathway has been found [1]. The pathway utilized for obtaining purine may depend on the cellular environment, as well as the requirements of the given organism. In *Mycobacteria*, both pathways are known to exist, although they have not been extensively investigated [2–4]. Adenylosuccinate lyase (ASL; EC 4.3.2.2) is the only enzyme common to both the pathways. In the *de novo* biosynthetic pathway, it displays dual substrate specificity and catalyzes two nonsequential reactions. The same active site is involved in both the reactions. These two reactions correspond to the cleavage of 5-aminoimidazole-(*N*-succinylocarboxamide) ribotide to 5-aminoimidazole-4-carboxamide ribotide and fumarate in one step and cleavage of succinyl-adenosine monophosphate (SAMP) to AMP and fumarate in another step [5]. It has been demonstrated that the reactions involve general acid–base catalysis with β-elimination of the succinyl group from the substrates leading to the release of fumarate, which leaves the active site before the release of 5-aminoimidazole-4-carboxamide ribotide or SAMP [6,7]. A general scheme for catalysis of SAMP by ASL proposed on the basis of biochemical studies [5,8,9] is shown in Fig. 1. It proceeds by abstraction of a proton from C_β atom of the succinyl group by a catalytic base leading to the formation of a carbanion intermediate that is stabilized by resonance [10]. Protonation at N1 or N6 by a catalytic acid leads to C_α-N bond cleavage and release of AMP and fumaric acid as products.

ASL is a housekeeping enzyme that is found in many organisms. It plays a crucial role in cellular replication, purine nucleotide cycle [11] and cellular metabolism [12]. In humans, point mutations in the *purB* gene encoding ASL that lead to functional deficiency cause autosomal recessive disorders such as autism, mental retardation, epilepsy and degeneration of muscles [13–15]. It has also been observed that the nucleotide requirements of most normally dividing human cells (except liver and activated T-lymphocytes) are derived from the salvage pathway, whereas those of rapidly dividing tumor cells are derived by *de novo* biosynthesis [16], making ASL an attractive target for anticancer drugs [17]. Additionally, ASL could also be a suitable target for antibacterial drugs [18].

In the present study, we determined the X-ray crystal structure of ASL from nonpathogenic *Mycobacterium smegmatis* (*MsASL*) at a resolution of 2.16 Å. We also crystallized ASL from pathogenic *Mycobacterium tuberculosis* (*MtbASL*). However, these crystals

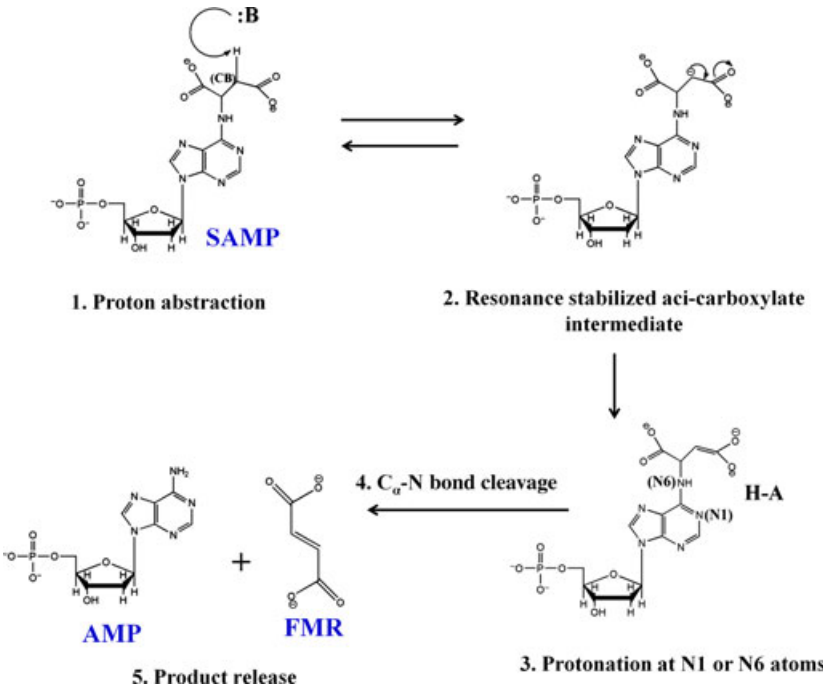

Fig. 1. General scheme showing cleavage of SAMP by ASL. The five steps involved are: (1) Proton abstraction by the catalytic base (B) from the CB atom of succinyl group; (2) stabilization of the carbanion intermediate formed; (3) protonation at N1 or N6 atom of AMP moiety by a catalytic acid; (4) cleavage of the C_α-N bond between AMP and succinyl groups; and (5) release of the products.

did not diffract X-rays. In the absence of diffracting crystals, a structural model for *MtbASL* was obtained by homology modeling using *MsASL* structure as a template. These unliganded structures reveal the most-open conformation of the C3 loop observed so far in the structures of aspartase–fumarase superfamily of enzymes. The partially ordered C3 loop in *MsASL* structure provides a basis for identification of the catalytic residues. Structural comparison and phylogenetic analysis show that the mycobacterial enzymes are more similar to the human enzyme rather than other bacterial ASLs. A comparison of the kinetic constants of ASL from several organisms demonstrates that the mycobacterial enzymes have the lowest catalytic activity. The activity of *MtbASL* is particularly low. This could be a result of the slow growth rate of *Mycobacteria* and the high GC content of their genomes, as well as their dependence on adenosine kinase (a salvage pathway enzyme) for the supply of AMP. Our structural results for *MsASL* could be used as a model system to understand the structure–function relationship in *MtbASL*. The structure of *MsASL* will also be useful for exploring possibilities with respect to the design of specific inhibitors.

Results and Discussion

MsASL structure determination

Crystals of *MsASL* diffracted X-rays to a resolution of 2.16 Å at the BM-14 beam line of the European Synchrotron Radiation Facility (ESRF). The data collection statistics for these crystals are summarized in Table 1. The structure of *MsASL* could be determined using polyalanine coordinates of human ASL as the phasing model. The model obtained was manually adjusted and the side chains were built. After refinement of positional parameters, 518 water molecules were added based on $2F_o$-F_c and F_o-F_c maps. R_{work} and R_{free} factors of the final refined structure are 20.4% and 23.9%, respectively. The final refinement statistics are summarized in Table 2.

Table 1. Data collection statistics.

	MsASL
Wavelength (nm)	0.9537
Resolution limits (Å)	50.0–2.16
Highest resolution shell (Å)	2.28–2.16
Exposure time (s)	5
Crystal to detector distance (mm)	240
Space group	$P2_12_12$
Cell parameters (Å)	a = 74.6, b = 176.83, c = 73.3
Protomers in the asymmetric unit	2
R_{merge} (%)	8.3 (27.4)[a]
Total observations	253 669
Unique observations	52 292
Mean redundancy	4.85
Completeness (%)	98.9 (96.4)[a]
Mean ($I/\sigma I$)	11.3 (4.2)[a]

[a]Values given in parentheses are for reflections of the highest resolution shell extending from 2.28 to 2.16 Å.

Table 2. Refinement statistics.

	MsASL
Resolution (Å)	50.0–2.16
Reflections used in refinement[a]	49 552
R_{work} (%)[a]	20.4
R_{free} (%)[b]	23.9
rms bond lengths (Å)[c]	0.0044
rms bond angles (°)[c]	0.76
Ramachandran statistics	
1 Most favored (%)	92.6
2 Additionally allowed (%)	6.4
3 Generously allowed (%)	1.0
4 Outliers (%)	0.0
Mean B-factor for protein atoms (Å^2)	20.8
Mean B-factor for water atoms (Å^2)	34.0

[a]R_{work} defined as $\Sigma(F_o-F_c)/\Sigma F_o$, where F_o and F_c refer to observed and calculated structure factor amplitudes. R_{free} is defined similarly for the reflections that were excluded from the refinement.
[b]5% of these reflections were not included in the refinement and were used for calculation of *R*-free.
[c]rmsd of the observed bond lengths and bond angles from the standard values.

Model quality of *MsASL*

The final electron density map was of good quality throughout most of the polypeptide main chain (Fig. 2A). Two protomers, A and B are present in the asymmetric unit of the crystal. A crystallographic two-fold axis generates the other two protomers (C and D) of the tetrameric *MsASL*. The observed electron density for the two protomers in the asymmetric unit is largely similar, although with some minor differences. Electron density is not observed in both the protomers for the first three residues at the N-terminus and the last three at the C-terminus. Density is absent for residue 76 in protomer A and residue 49 in protomer B, as well as for a short stretch of three residues (90–92) in both A and B. A long stretch of residues from 272–287 that corresponds to the C3 loop (close to the active site and harbouring the signature sequence '^{278}GSSxxPxKxN^{287}') is disordered in all the native ASL structures determined so far [12,19,20]. Density

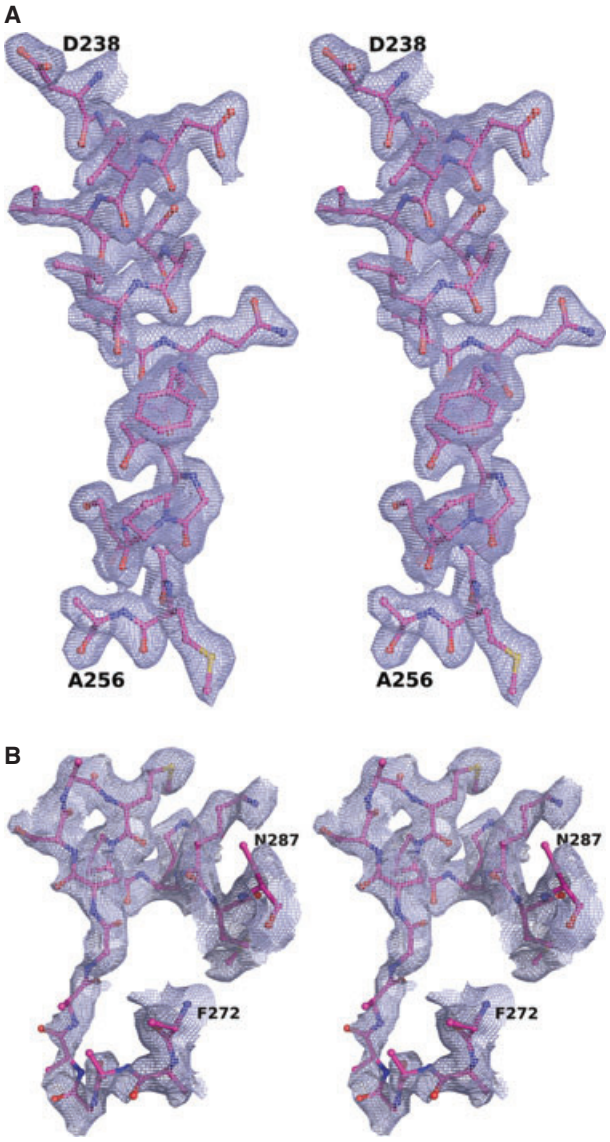

Fig. 2. Quality of the final electron density map. (A) Stereo view of a representative section (residues 238–256) of the map contoured at 1.5 σ along with the fitted polypeptide chain (pink). (B) Stereo view of the $2F_o$-F_c electron density map for the partially ordered residues 271–287 of the C3 loop (red), which is disordered in most of the apo-ASL structures. The map has been contoured at 0.5 σ.

was observed for residues 272–279 in protomer A, although at a lower contour level of 0.5 σ (Fig. 2B). The density for C3 loop residues 272–284 is absent in protomer B. Residues Arg17, Glu115 and Gln360 have been built in two alternate conformations, each with half occupancy. The side chains of 11 residues in protomer A and 18 residues in protomer B have been truncated according to the observed electron density and seven of these residues are common to both the protomers.

Validation of the structure using PROCHECK showed that 92.6%, 6.4% and 1.0% of the residues are in the most favored, additionally allowed and generously allowed regions, respectively, of the Ramachandran map [21]. There were no outliers in the final structure. No additional density that might correspond to bound ligands apart from water molecules was observed.

X-ray crystal structure of *MsASL*

ASL belongs to the aspartase/fumarase superfamily of enzymes, all members of which share characteristic tertiary and quaternary structures. This superfamily also includes argininosuccinate lyase, δ2-crystallin, class II fumarase, L-aspartase and 3-carboxy-*cis,cis*-muconate lactonizing enzyme (CMLE). The sequence identity across the members of the superfamily is ~ 10–20%, yet they are uniquely characterized by a common fold and their ability to catalyze β-elimination reactions of succinyl containing substrates leading to the formation of fumarate, except in the CMLE catalyzed reaction, which results in the formation of lactone [22].

MsASL structure adheres to the fold of the enzymes of the aspartase/fumarase superfamily. Each protomer of *MsASL* is composed of three distinct domains: N-terminal D1 (residues 1–100), central D2 (residues 101–366) and C-terminal D3 (367–474) domains arranged in the shape of a dumbbell (Fig. 3A). The structure is predominantly helical, with each protomer consisting of 22 helical segments and four β-strands. The four strands that occur in domain D2 are organized as two anti-parallel β-sheets, each consisting of two strands (β-sheet I: residues 142–148 and 151–157; β-sheet II: residues 267–269 and 353–354).

ASL is biologically active only as a homotetramer as the active site residues are from three highly conserved regions, C1 (residues 95–110), C2 (residues 146–157) and C3 (residues 272–288) contributed by three independent protomers [23]. *MsASL* forms a tight tetramer (Fig. 3B) with three independent stable interfaces (A/B, A/D, A/C) associated with buried surface areas of 5890 Å^2, 3620 Å^2 and 3270 Å^2 respectively. There are 58, 12 and 15 hydrogen bonds and 12, eight and three salt bridges in the three interfaces, respectively.

The C3 loop, which is disordered in most apo-ASL structures, could be built in protomers A and D related to the crystal by two-fold. The mean B-factor for the main chain atoms of residues 272–287 is 71.1 Å^2, which is substantially higher than those of all main chain atoms of protomer A (32.9 Å^2) and

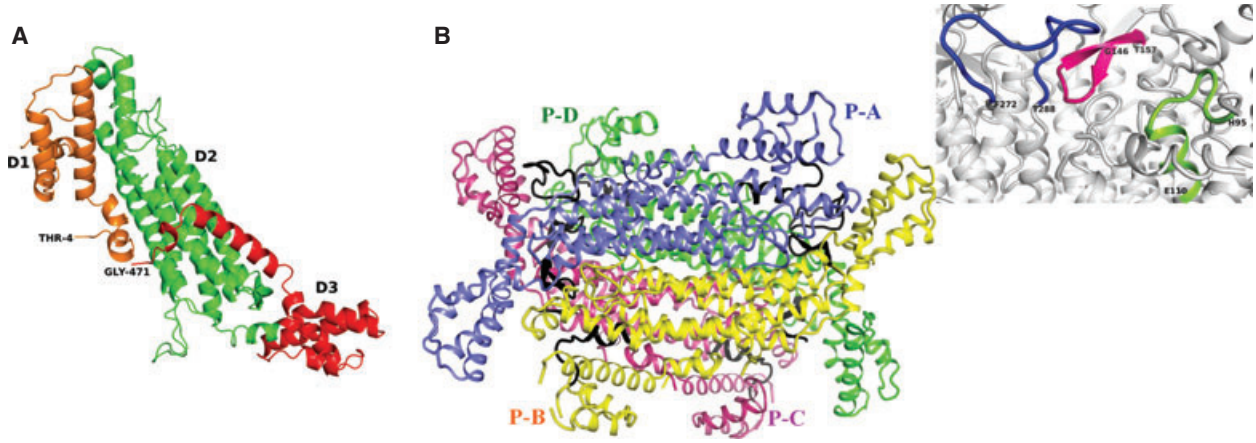

Fig. 3. *MsASL* X-ray crystal structure. Cartoon representation of the (A) protomer and (B) tetramer. The three domains in *MsASL* protomer are labeled D1–D3. The N- and C-terminal residues are marked. The four protomers in *MsASL* tetramer in (B) are labeled P–A to P–D and colored in blue, yellow, pink and green, respectively. The residues forming the active sites are shown in black. The view in (B) is related to that of (A) by a rotation of 90°. The inset is an enlarged view of one of the active sites formed by the conserved regions C1, C2 and C3, contributed by three separate protomers.

protomer B (31.7 Å^2), reflecting their flexibility and partial disorder. The C3 loop was mostly disordered in subunits B and C. *MsASL* is the first reported unliganded ASL structure in which the C3 loop is ordered at least partially. Upon ligand binding, this loop becomes ordered in ASLs from *Escherichia coli* (*EcASL*; Protein Data Bank (PDB) code: 2PTQ) [20], *Staphylococcous aureus* (*SaASL*; PDB code: 2X75) [18], *Bacillus anthracis* (*BaASL*; PDB code: 2PFM) and *Trypanosoma brucei* (*TbASL*; PDB code: 4EFC), whereas it remains disordered in ASLs from *Francisella tularensis* (*FtASL*; PDB code: 4EEI) and *Homo sapiens* (*HsASL*; PDB code: 2VD6 and 2J91). A superposition of the ordered C3 loop of *MsASL* with those of *EcASL* and *SaASL* (C3 loop conformation in *BaASL* and *TbASL* is similar to that of *EcASL*) shows that the loop has the most-open conformation in *MsASL* (Fig. 4A). In the native and L-aspartate bound forms of aspartase B, an enzyme of the aspartase/fumarase superfamily, the C3 loop is in open and closed conformations, respectively [24]. Superposition of these structures on *MsASL* confirms that the C3 loop of *MsASL* has the most-open conformation (Fig. 4B).

Homology modeling of *MtbASL* structure

In the absence of diffracting crystals, *MtbASL* tetrameric structure was modeled using *MsASL* structure because the amino-acid sequence identity between the two enzymes is 84.5%. The modeled structure of *MtbASL* shows a similar conserved fold of the aspartase/fumarase superfamily with striking resemblance to *MsASL* and *HsASL*. The structural alignment of the two mycobacterial ASLs and *HsASL* is shown in Fig. 5. Validation of the tetrameric structure using PROCHECK showed that 93.2%, 6.1% and 0.6% of the residues are in the most favored, additionally allowed and generously allowed regions, respectively, of the Ramachandran map. Ser277 of chains A and C are present in the disallowed region of the Ramachandran map. This serine is in the partially disordered C3 loop. Superposition of C_α atoms of *MtbASL* and *MsASL* tetramers yielded an rmsd of 0.42 Å.

The conformation of the C3 loop in *MtbASL* was identical to that observed in the template when *MsASL*, *EcASL* or *SaASL* with ordered C3 loops was used for modeling. On the other hand, modeling *MtbASL* based on structures from which the C3 loop was omitted showed variable conformations for the loop (Fig. 4C), suggesting that the loop is flexible and might assume different conformations depending on the environment.

Structural analysis

A sequence-based phylogenetic analysis of ASLs, whose structures have been determined, unexpectedly showed that the mycobacterial ASLs are more closely related to the higher eukaryotic ASLs such as *HsASL* and *Caenorhabditis elegans* ASL (*CeASL*) than to lower eukaryotic ASLs such as *TbASL* and *Plasmodium vivax* ASL (*PvASL*) or other bacterial ASLs (Fig. 6). Although amino acid sequence identity

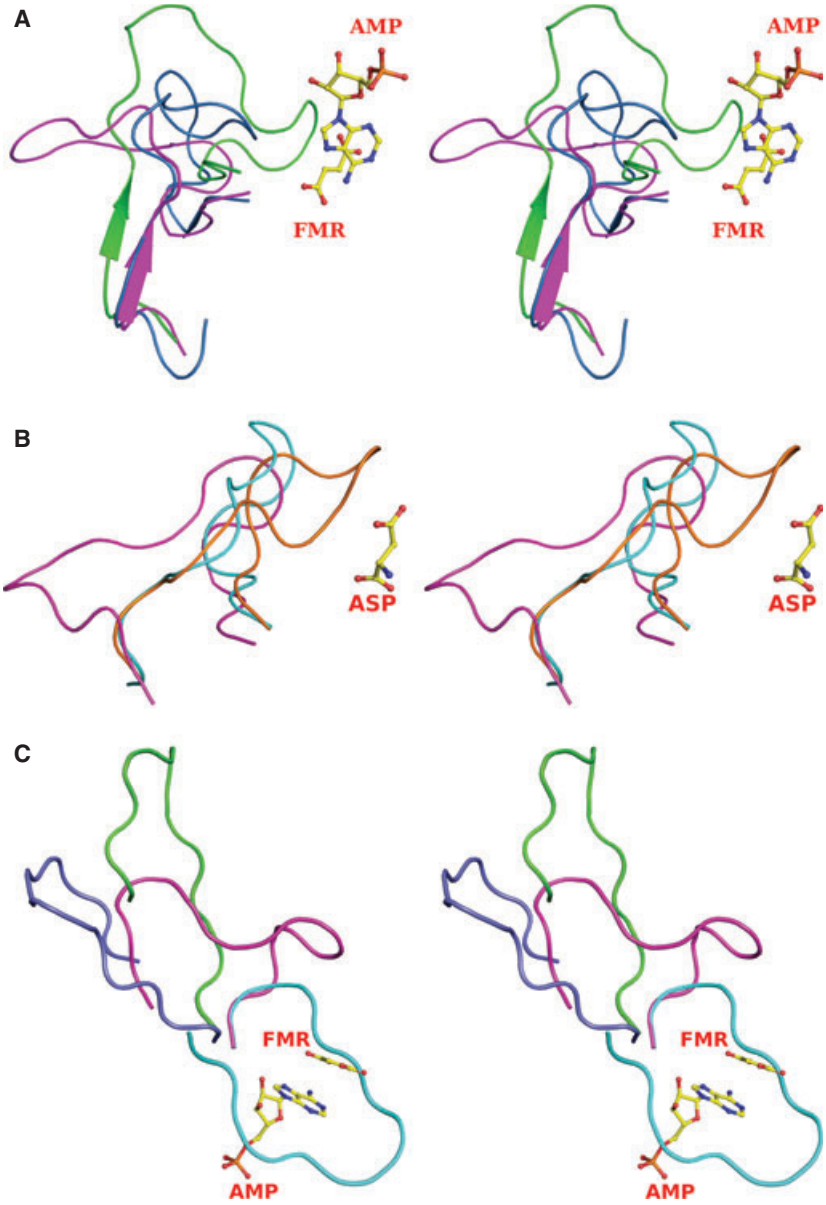

Fig. 4. C3 loop conformation in *MsASL* and related proteins. (A) Comparison of the conformations of the C3 loop in apo-*MsASL* (pink), *EcASL* (green, PDB code: 2PTQ) and *SaASL* (blue, PDB code: 2X75). The loop is most open in *MsASL*. (B) Comparison of the C3 loop conformation of *MsASL* (pink) with unliganded (cyan, PDB code: 3R6Q) and L-aspartate bound (orange, PDB code: 3R6V) aspartase B. (C) Comparison of the conformation C3 loop modeled in *MtbASL*. The conformation of the loop obtained by using the templates *MsASL* with C3 loop and *MsASL*, *EcASL* and *SaASL* without C3 loop are colored in pink, cyan, green and blue, respectively. The location of the active site is represented by bound ligands, AMP and fumaric acid in (A) and (C) and L-aspartate in (B).

between *MsASL* and human ASL is 33%, it is only 20% between the *E. coli* and mycobacterial enzymes. Structure based comparisons also reveal a similar phylogeny (results not shown). The organization of the three structural domains of mycobacterial enzymes is similar to that of human ASL. The rmsd values observed upon structural superposition of *MsASL* with *HsASL*, *CeASL*, *PvASL* and *EcASL* are 1.7, 1.5, 2.3 and 2.4 Å, respectively. The rmsd values upon independent structural superposition of the three domains of *MsASL* and human ASL protomers are 1.1, 1.1 and 1.4 Å for D1, D2 and D3, respectively. Similar structural superposition of the domains of *MsASL* and *EcASL* results in rmsd values of 2.4, 1.5 and 3.2 Å, respectively. D2 appears to be the most conserved domain, whereas D3 is the most variable (Fig. 7A). D3 has fewer interactions with the major domain D2 compared with domain D1.

To compare the relative disposition of D3 with respect to D2 in different ASL structures, ASL protomers were transformed using the matrix obtained by superposing domain 2 of all other ASLs on *MsASL* D2. D3 was found to occupy two extreme positions in *EcASL* and *MsASL* (Fig. 7A). The disposition of D3 found in *EcASL* was also observed in most of the other enzymes, including *TbASL*, *PvASL*,

```
MsASL-A.pdb   1  -------TIPNVLANRYASDEMVAIWSPEAKIIAERRLWLAVLRAQAELG 43
MtbASL-A.pdb  1  ------MSIPNVLATRYASAEMVAIWSPEAKVVSERRLWLAVLRAQAELG 44
2vd6-A.pdb    1  GDHGSPDSYRSPLASRYASPEMCFVFSDRYKFRTWRQLWLWLAEAEQTLG 50
                        sipnvLA RYAS EMvaiwSpeaK   eRrLWLavlrAqaeLG

MsASL-A.pdb  44 VAVPDGVVEDYERVLENVDLESIAARERVRHDVKA-RIE-EFNA-----H 86
MtbASL-A.pdb 45 VAVADSVLADYERVVDDVDLASISARERVLRHDVK-ARI-EEFNALA-GH 91
2vd6-A.pdb   51 LPITDEQIQEMKSNLENIDFKMAAEEEKRLRHDVMAHVHTFGHCCPKA-A 99
                   vav D v  dyervlenvDl siaarErvlrhdv      e        h

MsASL-A.pdb  87 EHVHKGMTSRDLTENVEQLQIRQSLELVFSHGVAVVARLAERAVVYRDLV 136
MtbASL-A.pdb 92 EHVHKGMTSRDLTENVEQLQIRRSLEVIFAHGVAAVARLAERAVSYRDLI 141
2vd6-A.pdb  100GIIHLGATSCYVGDNTDLIILRNALDLLLPKLARVISRLADFAKERASLP 149
                   ehvHkGmTSrdlteNveqlqiR sLel f hgvavvaRLAerAv yrdL

MsASL-A.pdb  137MAGRSHNVAAQATTLGKRFASAAEETLVALTRLRELIDRYPLRGVKGPMG 186
MtbASL-A.pdb 142MAGRSHNVAAQATTLGKRFASAAQEMMIALRRLRELIDRYPLRGIKGPMG 191
2vd6-A.pdb   150TLGFTHFQPAQLTTVGKRCCLWIQDLCMDLQNLKRVRDDLRFRGVKGTTG 199
                   maGrsHnvaAQaTTlGKRfasaaqe   aL rLreliDryplRGvKGpmG

MsASL-A.pdb  187TAQDMLDLFGGDVGKLADLERRVAEFLGFTEVFTSVGQVYPRSLDHDVLS 236
MtbASL-A.pdb 192TGQDMLDLLGGDRAALADLERRVADFLGFATVFNSVGQVYPRSLDHDVVS 241
2vd6-A.pdb   200TQASFLQLFEGDDHKVEQLDKMVTEKAGFKRAFIITGQTYTRKVDIEVLS 249
                   T qdmLdLfgGD  kladLerrVaeflGF  vF svGQvYpRslDhdVlS

MsASL-A.pdb  237ALVQFGAGPSSMAHTIRLMAGHELVTEGFAPGAVGSSAMPHKMNTRSCER 286
MtbASL-A.pdb 242ALVQLGAGPSSLAHTIRLMAGHELATEGFAPGQVGSSAMPHKMNTRSCER 291
2vd6-A.pdb   250VLASLGASVHKICTDIRLLANLKEMEEPF--------E-PYKRNPMRSER 290
                   aLvqlGAgpss ahtIRLmAghel tEgF        a PhKmNtrscER

MsASL-A.pdb  287VNGLQVVLRGYASMAAELAGAQWNEGDVFCSVVRRVALPDAFFAIDGQTE 336
MtbASL-A.pdb 292VNGLQVVLRGYASMVAELAGAQWNEGDVFCSVVRRVALPDSFFAVDGQIE 341
2vd6-A.pdb   291CCSLARHLMTLVMDPLQTASVQWFERTLDDSANRRICLAEAFLTADTILN 340
                   vngLqvvLrgyasm aelAgaQWnEgdvfcSvvRRvaLpdaFfa Dgq e

MsASL-A.pdb  337TFLTVLDEFGAYPAVIQRELDRYLPFLATTRILMAAVRAGVGREAAHEVI 386
MtbASL-A.pdb 342TFLTVLDEFGAYPAVIGRELDRYLPFLATTKVLMAAVRAGMGRESAHRLI 391
2vd6-A.pdb   341TLQNISEGLVVYPKVIERRIRQELPFMATENIIMAMVKAGGSRQDCHEKI 390
                   TfltvldefgaYPaVI ReldryLPFlATt ilMAaVrAG gRe aHe I

MsASL-A.pdb  387KEHAVAVALAMREQGREPDLIDRLAGDPRLP-LDKVALEAALEDKQAFTG 435
MtbASL-A.pdb 392SEHAVATALAMREHGAEPDLLDRLAADPRLT-LGRDALEAALADKKAFAG 440
2vd6-A.pdb   391RVLSQQAASVVKQEGGDNDLIERIQVDAYFSP-IHSQLDHLLD-PSSFTG 438
                    ehava Alamre G epDLidRla Dprl      aLeaaL  k aFtG

MsASL-A.pdb  436AAGDQVDGVVAA-VGELVSRYPEAAKYTSG--- 464
MtbASL-A.pdb 441AAGDQVDDVVAM-VDALVSRYPDAAKYTPGAIL 472
2vd6-A.pdb   439RASQQVQRFLEEEVYPLLKPYES---------- 461
                   aAgdQVd vva  V  LvsrYp
```

Fig. 5. Structural alignment of *MsASL*, *MtbASL* and *HsASL*.

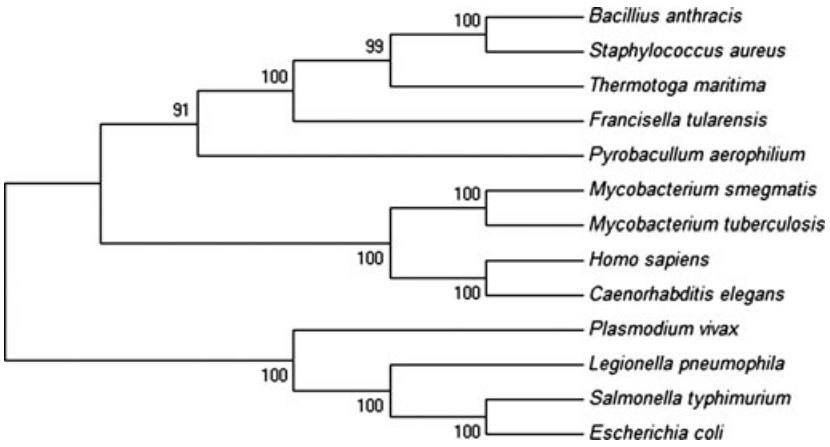

Fig. 6. A sequence based phylogenetic tree of all ASLs for which structures are available in the PDB. The mycobacterial ASLs are evolutionarily closer to the higher eukaryotes (human and *C. elegans*) compared to other bacterial ASLs.

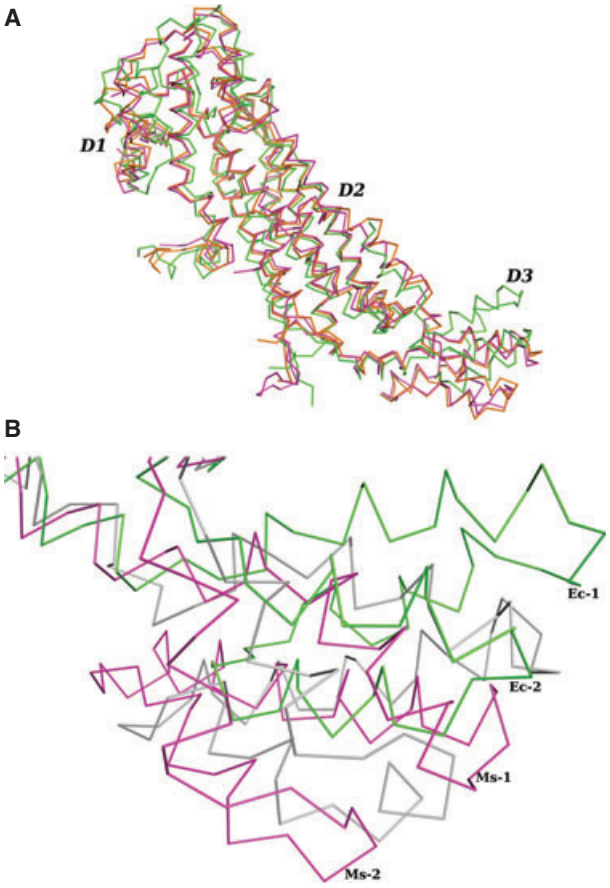

Fig. 7. Analysis of domain 3 disposition. (A) Ribbon diagram representing superposition of protomers in ASLs from *M. smegmatis* (pink), human (PDB code: 2VD6, orange) and *E. coli* (PDB code: 2PTS, green). Domain organization in *MsASL* resembles that of eukaryotic human ASL more than that of bacterial *EcASL*. The structure of domain 2 is mostly invariant in the three ASLs. Largest differences are observed for domain 3. The three domains are labeled as D1, D2 and D3. (B) The two extreme conformations of D3 in *MsASL* and *EcASL*. The two helices appearing immediately after D2 are labeled as 1 and the next two helices as 2. *TmASL* (grey) represents the intermediate conformation.

Legionella pneumophila ASL (*LpASL*) and *Pyrobaculum aerophilum* ASL (*PaASL*). By contrast, in *HsASL*, the position of D3 with respect to D2 was similar to that in *MsASL*. Disposition of D3 in *BaASL*, *Thermotoga maritima* ASL (*TmASL*) and *SaASL* is intermediate to that in *EcASL* and *MsASL* (Fig. 7B). The length and conformation of the loop following the helical residues 389–410 of *MsASL* is variable in ASL structures and might contribute to the relative disposition of domain D3 with respect to D2. The length of this loop is shortest in ASLs with a D3 conformation similar to *EcASL* and longest in *MsASL* and *HsASL*.

A comparison of the disposition of D3 with respect to D2 across the members of aspartase/fumarase superfamily shows that the positional heterogeneity of D3 observed in *MsASL*, *EcASL* and *TmASL* is also found in this superfamily (data not shown). All the structures of class II fumarase and L-aspartase deposited in the PDB have a D3 conformation similar to that of *MsASL*. All argininosuccinate lyases and δ-2 crystallin structures in the PDB have a D3 conformation resembling that of *EcASL*. The two intact structures of CMLE found in the PDB have a D3 conformation similar to that of *TmASL*. It has been demonstrated that ASL interacts with glutamyl-tRNA synthetase [25], leading to linkage between protein biosynthesis and nucleotide metabolism. It has been proposed that D3 might be involved in this interaction [19]. Therefore, the disposition of D3 with respect to D2 as observed in ASL structures might influence this interaction and be of physiological importance.

Kinetic parameters

The kinetic parameters of the two enzymes for the second reaction of *de novo* purine biosynthetic pathway in both directions were determined at a pH of 7.6 and temperature of 37 °C. The variation of their activities as a function of pH was studied over the range 5.2–9.6. The optimum pH values for the catalytic activities of *MsASL* and *MtbASL* were 8.0 and 7.6, respectively (data not shown). The Michaelis–Menten curves of the two mycobacterial enzymes for SAMP, AMP and fumarate are shown in Fig. 8. The kinetic parameters, k_{cat}, K_m and k_{cat}/K_m of the two mycobacterial ASLs are shown in Table 3. The catalytic efficiencies (k_{cat}/K_m) of *MsASL* for SAMP, AMP and fumarate are ~ 40-, 42- and ten-fold greater than the corresponding values for *MtbASL*.

A comparison of the specific activities of all the biochemically characterized ASLs is provided in Table 4. Apart from the mycobacterial enzymes, these include *Bacillus subtilis* ASL (*BsASL*) [23], *EcASL* [20], *Plasmodium falciparum* ASL (*PfASL*) [26], *HsASL* [7] and *Leishmania donovani* ASL (*LdASL*) [27]. Most activity measurements on ASLs have been performed at 25 °C. However, the catalytic activity of the mycobacterial enzymes was very low at this temperature. The activity increased linearly with temperature until 50–55 °C and started reducing when the temperature was raised further (data not shown). Therefore, physiological conditions (temperature of 37 °C and pH 7.6), where moderate activity of mycobacterial ASLs was observed, were selected for activity measurements. Mycobacterial ASLs, particularly

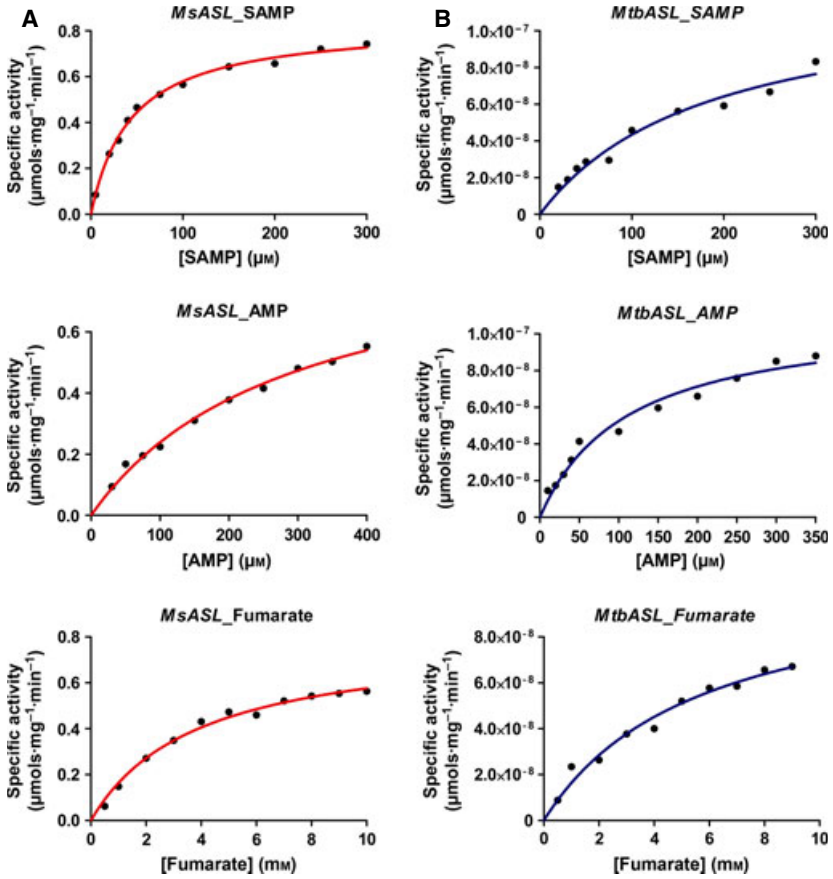

Fig. 8. Plots of specific activities in (μmol·mg^{-1}·min^{-1}) of (A) MsASL (in red) and (B) MtbASL (in blue) as a function of substrate concentration fitted to the Michaelis–Menten equation.

Table 3. Kinetic parameters determined at pH 7.6 and 37°C for the forward and reverse reactions with SAMP catalyzed by *MsASL* and *MtbASL*.

	MsASL	*MtbASL*
1 SAMP		
a k_{cat} (s^{-1})	0.7 ± 0.01	0.1 ± 0.02
b K_m (μM)[a]	43.7 ± 2.6	204.2 ± 48.2
c k_{cat}/K_m ($\mu M^{-1}\cdot s^{-1}$)	0.02	0.0005
2 AMP		
a k_{cat} (s^{-1})	0.8 ± 0.07	0.1 ± 0.01
b K_m (μM)[a]	277.8 ± 45.4	134.0 ± 27.0
c k_{cat}/K_m ($\mu M^{-1}\cdot s^{-1}$)	0.03	0.0007
3 Fumarate		
a k_{cat} (s^{-1})	0.7 ± 0.03	0.1 ± 0.01
b K_m (mM)[5]	3.9 ± 0.5	5.7 ± 1.2
c k_{cat}/K_m ($\mu M^{-1}\cdot s^{-1}$)	0.2	0.02

[a] As a result of high K_m values, *MsASL* did not reach saturation when titrating with AMP, whereas *MtbASL* did not reach saturation with SAMP and fumarate. Thus, the apparent K_m values for the substrates are reported.

MtbASL, were observed to have low specific activities compared to other ASLs despite using near optimal conditions. A comparison of the catalytic efficiency

Table 4. Comparison of kinetic parameters for SAMP catalysis across ASL homologs.

	Assay conditions	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1}\cdot s^{-1}$)
MsASL	37 °C, pH 7.6	0.7	43.7	0.02
MtbASL	37 °C, pH 7.6	0.1	204.2	0.0005
EcASL	25 °C, pH 7.0	4.1	–	–
HsASL	25 °C, pH 7.5	287.0	10.6	26.3
PfASL	25 °C, pH 7.4	7.5	32.0	0.2
LdASL	25 °C, pH 7.0	28.0	24.0	1.2
BsASL	25 °C, pH 7.0	1.3	3.5	0.4

k_{cat}/K_m of the enzymes shows that *HsASL* is the most active enzyme. Activities of other ASLs are comparable, although lower than *HsASL*. Both *MsASL* and *MtbASL* showed very low activity, with *MsASL* being one order of magnitude less, and *MtbASL* being three orders of magnitude less, than those of other bacterial ASLs. The human enzyme is ~ 40 000 times more efficient compared to mycobacterial ASLs.

Catalytic residues of mycobacterial ASLs

ASL catalysis (Fig. 1) follows a general acid–base reaction involving two separate groups to function as the catalytic acid and base [26]. The catalytic base is involved in proton abstraction (Fig. 1, step 1), followed by protonation by the catalytic acid (Fig. 1, step 3). In earlier studies, residues equivalent to His78 and His149 of *MsASL* were proposed to function as the acid and base, respectively [28]. In later studies, His149 or an activated water molecule and Ser279 were proposed to function as the acid and base, respectively [20]. In a more recent study on *SaASL*, it was proposed that residue equivalent to His149 and Lys285 of *MsASL* might function as the catalytic acid and base, respectively [18]. It is therefore of interest to examine the position and conformation of these proposed catalytic residues with respect to the putative ligand binding site in *MsASL* structure and their potential to function in catalysis.

Figure 9A shows the distance of His78 and His149 from the N1 and N6 atoms of AMP in apo-*MsASL* and holo-*HsASL* structures. His78 does not show a large variation in its conformation and C_α position between the apo- and holo-structures. Its $N_{\varepsilon 2}$ atom is at a distance > 5.0 Å from N1 of the adenine ring of SAMP in all the structures, meaning that His78 is unlikely to function as a catalytic acid. The conformation of His149 in apo-*MsASL* resembles that observed in the holo-structures. The $N_{\varepsilon 2}$ atom of *MsASL* His149 is close to the N6 atom of the adenine ring of SAMP with distances of 4.1 and 3.7 Å in *MsASL* and *HsASL*, respectively. By contrast, it is more distant from the C_β atom of the fumarate moiety with distances of 4.7 and 4.5 Å in *MsASL* and *HsASL*, respectively, suggesting that His149 is unlikely to function as a base extracting proton from fumarate but might function as the catalytic acid protonating N6 of SAMP.

The identity of the catalytic base of ASLs has been extensively debated [18,20]. Figure 10A shows the position of the residues Lys285 and Ser279 with respect to fumarate. Both of these residues, present on the highly flexible C3 loop, have been proposed to function as base in one or other of ASL structures [18,20]. Based on the available ASL crystal structures conformation of the C3 loop in *MsASL*, *EcASL* and *SaASL* structures could be considered as most-open, most-closed and intermediate, respectively. The conformation of Lys285 is highly variable in the apo- and holo-structures of ASLs. It is observed that substrate-binding brings this lysine within interacting distance of His149 and the carboxylate oxygens of the succinyl

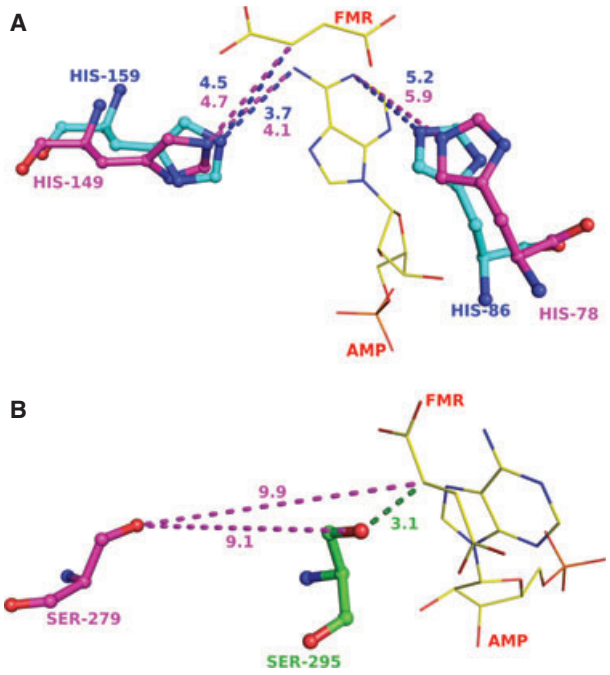

Fig. 9. Positions of the proposed catalytic residues in apo-*MsASL* (pink) with respect to the bound ligands (yellow), AMP and fumaric acid (FMR). (A) Structural superposition of apo-*MsASL* and holo-*HsASL* (cyan, PDB code: 2VD6) showing the position of $N_{\varepsilon 2}$ of His78 and His149 with respect to the N1 and N6 atoms of the adenine ring of AMP. The position of $N_{\varepsilon 2}$ of His149 with respect to C_β of FMR is also shown. The equivalent histidine residues in holo-*HsASL* are His86 and His159. (B) Structural superposition of apo-*MsASL* and holo-*EcASL* (green, PDB code: 2PTQ) showing the position of *MsASL* Ser279 and equivalent *EcASL* Ser295 with respect to C_β of FMR. Closure of the C3 loop might bring Ser279 close to the substrate. The color of the labels of the distances shown indicates the color of the corresponding structure.

group of SAMP in all ASLs except *SaASL*. The conformation of Lys285 in *MsASL* is similar to that observed in the holo-structure of *EcASL*. The equivalent lysine in *EcASL* shows strong hydrogen-bonding interactions with the carboxylate oxygens of SAMP. Mutation of this lysine has been shown to increase K_m of the enzyme for SAMP by several folds in *BsASL* [23], suggesting that Lys285 has a role in substrate binding. Lys285 might contribute to substrate binding by its interaction with the carboxylate group. It might also act as a base as the mutant is devoid of catalytic activity.

In *EcASL*, the distance between the O_γ atom of the serine residue (Ser295) equivalent to *MsASL* Ser279 and C_β atom of the fumarate moiety is 3.1 Å (Fig. 9B), suggesting that it could function as the catalytic base. In *MsASL*, the C_α of Ser279 is displaced

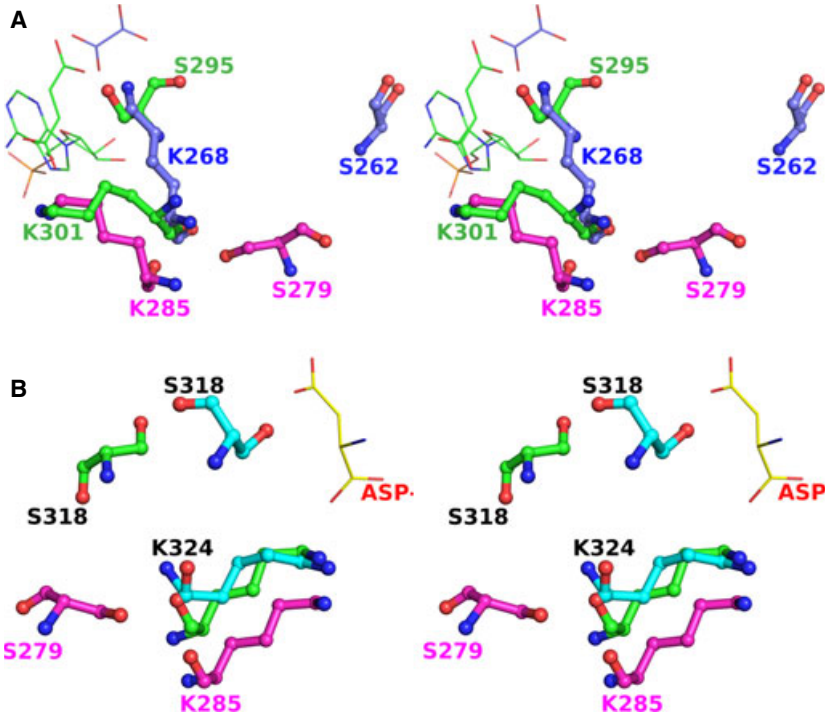

Fig. 10. Conformation of Ser279 in apo-*MsASL*. (A) Stereo view of the orientations of the proposed catalytic bases, Ser279 and Lys285 in apo-*MsASL* (pink) after superposition on holo-*EcASL* (green) and holo-*SaASL* (blue). (B) Stereo view of the orientations of Ser279 and Lys285 in apo-*MsASL* (pink), apo-AspB (cyan, PDB code: 36RQ) and holo-AspB (orange, PDB code: 36RV) after structural superposition. The equivalent Ser318 in apo-AspB is initially far away from the active site and is brought closer to the substrate in the holo-structure. The bound ligands at the active site are shown in line representation.

away from the C_α of Ser295 of *EcASL* by 8.8 Å. However, its side chain has a similar orientation as in *EcASL* and points towards the active site with a distance of 9.9 Å between its O_γ atom and C_β of the fumarate moiety. It is possible that substrate binding induces closure of the C3 loop reducing the distance between Ser279 and C_β of the succinyl group, as seen in *EcASL*, enabling proton abstraction by Ser279. Similar closure of the C3 loop bringing the equivalent serine residue closer to the substrate has also been observed in L-aspartase apo- and holo-structures (Fig. 10B). Hence, it is possible that Ser279 functions as the catalytic base. In the light of these results, it is likely that His149 functions as the catalytic acid and Lys285 or Ser279 functions as the catalytic base in mycobacterial ASLs, as well as in the aspartase/fumarase superfamily. Further mutational experiments followed by kinetic studies might establish the identity of the catalytic base.

Low catalytic activity of mycobacterial ASLs

Variation in the catalytic activity of ASLs from different species could be a result of sequence alterations or change in the geometry of the active site or change in the accessibility of the active site to substrates and products. Alignment of sequences of ASLs from various organisms (~ 150 ASL sequences representing different phyla) shows that the proposed active site residues are mostly conserved. In *Mycobacteria*, two residues near the active site, proposed to be important for substrate binding, were found to be different [20,30]. The residue corresponding to *MsASL*-Ser148 and *MtbASL*-Ser146 is either retained as serine or replaced by threonine in most other ASLs. This variation may not be the cause of drastic reduction in the activity of mycobacterial ASLs because the substitution does not lead to changes in the hydrogen bonding interactions near the active site. The second residue is a conserved arginine in all ASLs except in actinobacteria (a phylum that includes the genus *Mycobacteria*), where it is replaced by a glycine (Gly319 in *MsASL*). The drastic mutation from arginine to glycine will lead to a large cavity. In *MsASL*, this cavity is partially filled by the guanidino group of Arg103 of *MsASL* (Arg101 of *MtbASL*), although the orientation of the guanidino group is opposite to that of other ASLs (Fig. 11). Therefore, Arg319Gly substitution might be partially responsible for the reduced efficiency of mycobacterial enzymes. It would be of interest to examine the catalytic properties of Gly319Arg mutant of *MsASL*.

The low value of k_{cat} of mycobacterial ASLs suggests that the active site might not be easily accessible and the rate-limiting step of mycobacterial ASL catalysis could be slow substrate binding or product release. This might be a result of sluggish loop or subdomain movements of the enzyme and slow re-organization of the residues required for catalysis. The sequence at the

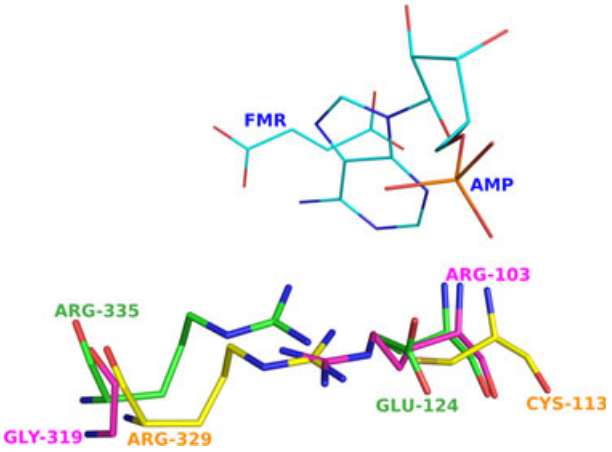

Fig. 11. Arg319Gly substitution in *MsASL* (pink). This residue is a strictly conserved Arg except in Actinobacteria. The guanidine group of Arg103 of *MsASL*, although oriented differently, partially fills this cavity. The residues equivalent to 319 and 103 of *MsASL* are also shown for *EcASL* (green) and *HsASL* (yellow). The active site is shown by bound ligands, AMP and fumaric acid (FMR), colored in cyan.

N-terminal segment of the flexible C3 loop is (ELVTEGFAP) in *MsASL*, (FKQKTIAGE) in *EcASL* and (KEIEEPFEK) in *HsASL*. The loop appears to undergo open–close conformational transition in the catalytic cycle. The human enzyme has six charged residues in this segment compared to three in *EcASL* and two in *MsASL*. The flexibility of the C3 loop may depend on these charged residues. Thus, the human enzyme with six charged residues may be most flexible. The loop is found to be disordered in the ligand bound structures (PDB code: 2VD6, 2J91) of *HsASL*. With an intermediate charge, the loop is disordered in the ligand free *E. coli* enzyme but gets ordered upon ligand binding [20]. By contrast, with minimal charges, the loop in *MsASL* is partially ordered even in the unliganded enzyme. The dynamic nature of this loop may contribute to the rate of catalysis leading to an efficient human enzyme and sluggish mycobacterial enzymes.

Many enzymes of purine salvage and biosynthetic pathways have been shown to be up- or down-regulated allosterically by the pathway intermediates or purine nucleotides [31,32]. Whether mycobacterial ASLs are also subject to allosteric regulation remains to be investigated.

Importance of ASL in *Mycobacteria*

The replication cycle of mycobacterial species is much longer than those of other bacteria. Thus, although *E. coli* divides once every 20 min, *Mycobacteria* divide once in several hours. Between the two mycobacterial species of the present study, *Mycobacteria smegmatis* (MS) is the faster growing (6-h cycle) compared to *Mycobacteria tuberculosis* (MTB), which divides once in 24 h (Goodfellow and Magee, 1998). A comparison of the k_{cat} for the cleavage of SAMP in all the kinetically characterized bacterial ASLs shows an inverse relationship with the corresponding doubling time (Table 5). Similar differences in the efficiency of the enzymes from *Mycobacteria* and other organisms as a result of differences in growth rates have been reported previously [33,34]. Furthermore, the mycobacterial genomes have a high GC content of ~ 65% and hence their requirement of adenine nucleotides required for DNA or RNA synthesis is correspondingly low. Therefore, the low activity of ASL might suffice for the slow growth rate and low AT content of these bacteria [33,34].

The second possible reason for the lower activity of mycobacterial ASLs could be the preference in these bacteria for the salvage pathway over the *de novo* biosynthetic pathway for deriving purine nucleotides. MTB genome codes for several purine salvage enzymes [35]. *Mycobacteria microti* and *Mycobacteria avium* have very low *de novo* purine biosynthetic activity and the pathway appears to be absent in *Mycobacteria leprae* [36]. Adenosine kinase, a salvage pathway enzyme supplying adenine nucleotides, is present in all eukaryotes as well as in *Mycobacteria*, whereas it is absent in most of the intracellular bacteria [37,38]. Furthermore, adenosine kinase has been shown to play a key role in supplementing adenine nucleotides in *M. leprae* [36]. Therefore, low catalytic activity of mycobacterial ASLs might be compensated by adenosine kinase. However, the function of ASL cannot be entirely substituted by adenosine kinase because it has been demonstrated that, although adenosine kinase deficient strains of MTB can grow *in vitro* [37], ASL deficient strains fail to survive. This suggests that the *purB* gene is essential for mycobacterial growth [39–41] and could be a target for developing specific inhibitors. Mycobacterial enzymes structurally resemble the human enzyme more than bacterial enzymes although they are almost 40 000 times slower than the human

Table 5. Relationship between the doubling time of the bacteria and k_{cat} for SAMP catalysis.

	Doubling time	k_{cat}·(s^{-1})
MtbASL	24 h	0.1
MsASL	6 h	0.7
BsASL	120 min	1.3
EcASL	20 min	4.1

enzyme. It would be of great interest to design an inhibitor for mycobacterial enzymes that does not affect the human enzyme.

Conclusions

ASLs from MS and MTB were cloned, over-expressed, purified and crystallized. Crystals of *MsASL* diffracted X-rays to high resolution and its structure could be determined at a resolution of 2.16 Å. This is the first X-ray crystal of an unliganded ASL to be reported with a partially ordered C3 loop in its most-open conformation. Crystals of *MtbASL* diffracting X-rays to useful resolution could not be obtained and hence its tetrameric structure was modeled using *MsASL* co-ordinates as a template. The X-ray crystal structure of *MsASL* and the modeled structure of *MtbASL* adhere to the fold that is common to all members of the aspartase/fumarase superfamily. The structures of mycobacterial ASLs were found to be closer to *HsASL* compared to other bacterial ASLs. A comparison of *MsASL* structure with other ASL structures suggests that His149 and Lys285 or Ser279 are most likely to function as the catalytic acid and catalytic base, respectively.

MsASL was found to be catalytically more active than *MtbASL*. However, both the enzymes were much slower compared to ASLs of other organisms. Most of the active site residues of mycobacterial enzymes were found to be conserved, with only two exceptions: Ser148 and Gly319 of *MsASL*, which correspond to threonine and arginine residues, respectively, in most of the other organisms. We speculate that the C3 loop at the active site is most dynamic in the human enzyme and least dynamic in mycobacterial enzymes. The substitution of Gly in *Mycobacteria* for Arg of human ASL and the nature of dynamics of the C3 loop might account for the low activity of mycobacterial ASLs. It is likely that the low catalytic activity of *MsASL* and *MtbASL* suffices for their slow growth rates, dependence on other salvage pathways for adenine nucleotides and their GC-rich genomes. However, basal expression of ASL is important for the sustained multiplication and infection of *Mycobacteria*. The structures of these mycobacterial ASLs might therefore be useful for the development of specific inhibitors as drug candidates for mycobacterial infections.

Materials and methods

Cloning and protein purification

The *purB* gene encoding ASL of *Mycobacterium smegmatis* strain MC2 155 was amplified by PCR using gene-specific primers: *MsASL*-sense primer 5′-GCTAGCCATATG TTC GTG ACG ATT CCG AAT-3′ and *MsASL*-antisense primer 5′-GGATCCTTAAAGCTT CAA GAT CGC GCC CGA-3′ (the underlined nucleotides represent restriction sites). The *purB* gene of *Mycobacterium tuberculosis* strain H37rv was amplified by PCR using gene-specific primers: *MtbASL*-sense primer 5′-GCTAGCCATATG AGC ATT CCC AAC GTG CTG-3′ and *MtbASL*-antisense primer 5′-GGATCCTTAAAGCTT AAG AAT TGC ACC CGG CGT-3′. The sense primers had *Nhe*I and *Nde*I restriction enzyme sites, whereas the antisense primers had *Bam*HI and *Hind*III restriction enzyme sites. Both *MsASL* and *MtbASL* genes were cloned in pRSET C vector between *Nhe*I and *Bam*HI sites. The strategy used for cloning resulted in the addition of fourteen amino acids from the vector sequence at the N-terminus of the protein including a hexa-histidine tag useful for Ni^{2+}-NTA (nitrilotriacetic acid) affinity purification. *MsASL* and *MtbASL* proteins were over-expressed in *E. coli* BL21 (DE3) pLysS cells by the addition of 0.3 mM isopropyl thio-β-D-galactoside. The induced cultures of *MsASL* and *MtbASL* were grown for further 6 h at 30 °C and 12 h at 20 °C, respectively. The proteins were purified in a two-step process starting with Ni^{2+}-NTA affinity chromatography, followed by gel filtration using a Superdex S-200 preparative column (GE Healthcare, Little Chalfont, UK). The purified proteins were in the buffer containing 50 mM Tris (pH 7.5) and 200 mM NaCl. The proteins were concentrated using an Amicon centrifugal filter device (molecular weight cut-off of 30 000; Millipore, Billerica, MA, USA). The molecular masses were confirmed by MALDI-TOF and corresponded to 53 and 52 kDa for *MsASL* and *MtbASL*, respectively. Protein concentrations were estimated using a Jasco V-630 spectrophotometer (Jasco Inc., Easton, MD, USA) and an extinction coefficient of 31 400 $M^{-1}\cdot cm^{-1}$ as calculated by EXPASY PROTPARAM [42]. The concentration obtained was verified by the Bradford assay [43].

Structural studies

Crystallization

Crystals for *MsASL* were obtained using a microbatch method [44]. Crystallization was set up with 2 μL of 10 $mg\cdot mL^{-1}$ protein and 2 μL of the crystallization condition containing 0.1 M sodium Hepes (pH 7.5), 1.4 M sodium citrate and 0.1 M betaine hydrochloride. The crystals diffracted X-rays to a resolution of 2.5 Å at home-source [Microstar Ultra II Rotating Anode X-ray Generator (Bruker Instruments, Inc., Bellerica, MA, USA) equipped with a MAR imaging plate detector (Marresearch GmbH, Norderstedt, Germany)] and a resolution of 2.16 Å on the BM14 beam line at the ESRF synchrotron facility (Grenoble, France).

Crystals for *MtbASL* were obtained using microbatch and hanging drop methods [45]. Crystallization was set up using 2 μL of 5 $mg \cdot mL^{-1}$ protein and 2 μL of different crystallization conditions. Crystallization was set up for native protein, as well as for the protein incubated with various concentrations of SAMP, AMP and fumarate. Attempts were made to enhance the crystal quality by adjusting crystallization conditions and including additives to the crystallization droplet. The crystals were also soaked for varying time periods in different cryoprotectants such as polyethylene glycol and glycerol. However, crystals of quality useful for structural studies could not be obtained.

X-ray diffraction data collection and processing of *MsASL*

Crystals of *MsASL* were transferred to the crystallization condition containing 5% glycerol as the cryoprotectant for a few seconds and then mounted in cryo-loops. They were tested by exposing them to Cu-K_α radiation at 100 K maintained using a stream of nitrogen gas. After confirming the quality, the crystals were frozen in liquid nitrogen and sent to BM-14, ESRF (Grenoble, France) for data collection. The frames were processed using MOSFLM [46] in the orthorhombic space group P222 with a = 74.6 Å, b = 176.83 Å and c = 73.3 Å. The resulting intensities were scaled using SCALA of the CCP4 software suite [47,48].

Structure determination of *MsASL*

Systematic absences in the diffraction data indicated the presence of two 2_1 screw axes along the a- and b-directions. Hence, the space group of the crystal was set to $P2_12_12$. A BLAST search [49] revealed that *MsASL* shares 33% sequence identity with *HsASL*. The 2.0-Å resolution structure of *HsASL* (PDB code: 2VD6) was retrieved from the PDB. It has an average B-factor of 28.8 $Å^2$. The structure of *MsASL* was determined by molecular replacement using the co-ordinates of the polyalanine chains of the AB dimer of 2VD6 as the phasing model in PHASER [50] of the CCP4 software suite. The quality of the model obtained from phaser was improved using the automatic model building software BUCCANEER [51] of the CCP4 software suite and AUTOBUILD [52] of PHENIX, version 1.7-650 [53] followed by manual fitting using COOT [54]. The individual atomic positions and the isotropic temperature factors were refined using REFMAC 5 [55] of the CCP4 software suite. During the final stages of refinement, TLS refinement [56] was used, where each of the protomer in the asymmetric unit consisted of one TLS group. This resulted in a decrease of R_{work} and R_{free} by 0.4% and 0.5%, respectively. The average B-factor of the final structure is 32.3 $Å^2$. The B-factors specified refer to the temperature factors after adding the TLS contribution. No noncrystallographic symmetry restraints were applied during the course of refinement. The progress of refinement was monitored using 5% (2615) of the 52 292 independent reflections, which were used for the calculation of free R-factor. The structure of *MsASL* was validated using PROCHECK [57] of the CCP4 software suite. The coordinates and the structure factors of *MsASL* have been deposited in the Protein Data Bank under code 4NLE.

Homology modeling of *MtbASL*

The tetrameric structure of *MtbASL* was modeled using MODELLER, version 9v7 [58,59], based on the co-ordinates of *MsASL* as the template. The output with the lowest molecular probability density function was selected and its stereochemical quality was verified using PROCHECK.

Structural analysis

Sequence alignment was carried out using either CLUSTAL W [60] or MULTALIN [61]. Structure based sequence alignment was carried out using the MUSTANG-MR structural sieving server [62]. Superposition of all structures was achieved using SSM SUPERPOSE [63] of COOT and ALIGN [64]. The interactions in the oligomer were identified using PDBEPISA [65]. The secondary structure was evaluated using PDBSUM [66]. Phylogenetic analyses were carried out using MEGA, version 4.0 [67]. All figures were prepared using PYMOL [68].

Biochemical studies

Steady-state kinetics of both *MsASL* and *MtbASL* for the forward (cleavage of SAMP) and reverse (formation of SAMP) reactions of the second reaction catalyzed by ASL in *de novo* purine biosynthesis were followed by monitoring the decrease or increase, respectively, in the absorbance of SAMP at 290 nm as a function of time using a double beam JASCO V-630 spectrophotometer at pH 7.6 and temperature 37 °C. The K_m and V_{max} for the forward reaction were determined by titrating SAMP over a range of concentrations. The K_m and V_{max} for the reverse reaction with the two substrates (AMP and fumarate) were determined independently by keeping one substrate at saturating concentration and titrating the other substrate over a range of concentrations. A mixed buffer (pH range 4.0–10.0) containing succinic acid, sodium dihydrogen phosphate and glycine in the ratio of 2 : 7 : 7 was used for the reactions in order to minimize the effect of buffer components on enzyme activities.

For determining the kinetic parameters for SAMP in the forward reaction, 35 μg of *MsASL* and 125 μg of *MtbASL* were used in a final assay volume of 500 μL and titrating SAMP concentrations over the range of 5–300 μM.

For determining the kinetic parameters for AMP in the reverse reaction, 35 μg of *MsASL* and 125 μg of *MtbASL* were used in a final assay volume of 500 μL with an almost saturating fumarate concentration of 8 mM and titrating AMP concentrations of 5–350 μM.

For determining the kinetic parameters for fumarate in the reverse reaction, 35 μg of *MsASL* and 125 μg of *MtbASL* were used in a final assay volume of 500 μL with almost saturating AMP concentration of 350 μM and titrating fumarate concentrations of 500 μM to 10 mM.

The reactions were initiated by the addition of substrate to the sample cuvette after setting the absorbance reading with respect to the reference cuvette to zero. The amount of substrate consumed was estimated using the difference extinction coefficient of 4.05 $mM^{-1} \cdot cm^{-1}$ [69]. Specific activity was calculated as μmoles of substrate converted to product in 1 min by a milligram of enzyme. Kinetic parameters were obtained by fitting the data to the Michaelis–Menten equation using GRAPHPAD PRISM, version 5.02 (GraphPad Software, Inc., San Diego, CA, USA).

Acknowledgements

We thank the staff at MBU X-ray laboratory and the beam line scientists at BM14, ESRF (Grenoble, France) for help during the X-ray diffraction data collection. Sourav Roy (JNCASR) kindly assisted in the analysis of the kinetic experiments. MRNM, HSS and HB thank the Department of Biotechnology and the Department of Science and Technology, India, for providing financial support. SB thanks the Council for Scientific and Industrial Research (CSIR), India for Junior & Senior Research Fellowships. We thank Fathima Benazir for critically reading the manuscript.

References

1 Berens RL, Krug EC & Marr JJ (1995) Purine and pyrimidine metabolism. In Biochemistry and Molecular Biology of Parasites (Marr JJ & Muller M, eds), pp. 89–117. Academic Press Ltd, London.
2 Ducati RG, Breda A, Basso LA & Santos DS (2011) Purine salvage pathway in *Mycobacterium tuberculosis*. *Curr Med Chem* **18**, 1258–1275.
3 Malathi V & Ramakrishnan T (1966) Biosynthesis of nucleic acid purines in *Mycobacterium tuberculosis* H37Rv. *Biochem J* **98**, 594.
4 Le Nours J, Bulloch EM, Zhang Z, Greenwood DR, Middleditch MJ, Dickson JM & Baker EN (2011) Structural analyses of a purine biosynthetic enzyme from *Mycobacterium tuberculosis* reveal a novel bound nucleotide. *J Biol Chem* **286**, 40706–40716.
5 Ratner S (1972) Argininosuccinases and Adenylosuccinases. In The Enzymes (Boyer PD, ed), pp. 167–197. Academic Press, New York.
6 Bridger W & Cohen LH (1968) The kinetics of adenylosuccinate lyase. *J Biol Chem* **243**, 644–650.
7 Stone R, Zalkin H & Dixon J (1993) Expression, purification, and kinetic characterization of recombinant human adenylosuccinate lyase. *J Biol Chem* **268**, 19710–19716.
8 Hanson KR & Havir EA (1972) The enzymic elimination of ammonia. In The Enzymes (Boyer PD, ed), pp. 75–166. Academic Press, New York.
9 Casey PJ & Lowenstein JM (1987) Purification of adenylosuccinate lyase from rat skeletal muscle by a novel affinity column. Stabilization of the enzyme, and effects of anions and fluoro analogues of the substrate. *Biochem J* **246**, 263–269.
10 Porter DJ, Rudie NG & Bright HJ (1983) Nitro analogs of substrates for adenylosuccinate synthetase and adenylosuccinate lyase. *Arch Biochem Biophys* **225**, 157–163.
11 Van den Berghe G, Bontemps F, Vincent M & Van den Bergh F (1992) The purine nucleotide cycle and its molecular defects. *Prog Neurobiol* **39**, 547–561.
12 Toth EA & Yeates TO (2000) The structure of adenylosuccinate lyase, an enzyme with dual activity in the *de novo* purine biosynthetic pathway. *Structure* **8**, 163–174.
13 Jaeken J, Wadman S, Duran M, Van Sprang F, Beemer F, Holl R, Theunissen P, De Cock P, Van den Bergh F & Vincent M (1988) Adenylosuccinase deficiency: an inborn error of purine nucleotide synthesis. *Eur J Pediatr* **148**, 126–131.
14 Maaswinkel-Mooij P, Laan L, Onkenhout W, Brouwer O, Jaeken J & Poorthuis B (1997) Adenylosuccinase deficiency presenting with epilepsy in early infancy. *J Inherit Metab Dis* **20**, 606–607.
15 Verginelli D, Luckow B, Crifò C, Salerno C & Gross M (1998) Identification of new mutations in the adenylosuccinate lyase gene associated with impaired enzyme activity in lymphocytes and red blood cells. *Biochim Biophys Acta* **1406**, 81–84.
16 Jackson RC & Harkrader RJ (1981) The contributions of *de novo* and salvage pathways of nucleotide biosynthesis in normal and malignant cells. In Nucleosides and Cancer Treatment (Tattersall MHN & Fox RM, eds), pp. 18–31. Academic Press, Sydney.
17 McLeod HL, Cassidy J, Powrie RH, Priest DG, Zorbas MA, Synold TW, Shibata S, Spicer D, Bissett D & Pithavala YK (2000) Pharmacokinetic and pharmacodynamic evaluation of the glycinamide ribonucleotide formyltransferase inhibitor AG2034. *Clin Cancer Res* **6**, 2677–2684.
18 Fyfe PK, Dawson A, Hutchison M-T, Cameron S & Hunter WN (2010) Structure of *Staphylococcus aureus*

adenylosuccinate lyase (PurB) and assessment of its potential as a target for structure-based inhibitor discovery. *Acta Crystallogr Sect D: Biol Crystallogr* **66**, 881–888.
19 Toth EA, Worby C, Dixon JE, Goedken ER, Marqusee S & Yeates TO (2000) The crystal structure of adenylosuccinate lyase from *Pyrobaculum aerophilum* reveals an intracellular protein with three disulfide bonds. *J Mol Biol* **301**, 433–450.
20 Tsai M, Koo J, Yip P, Colman RF, Segall ML & Howell PL (2007) Substrate and product complexes of *Escherichia coli* adenylosuccinate lyase provide new insights into the enzymatic mechanism. *J Mol Biol* **370**, 541–554.
21 Ramachandran G, Ramakrishnan CT & Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* **7**, 95–99.
22 Puthan Veetil V, Fibriansah G, Raj H, Thunnissen A-MW & Poelarends GJ (2012) Aspartase/fumarase superfamily: a common catalytic strategy involving general base-catalyzed formation of a highly stabilized aci-carboxylate intermediate. *Biochemistry* **51**, 4237–4243.
23 Brosius JL & Colman RF (2002) Three subunits contribute amino acids to the active site of tetrameric adenylosuccinate lyase: Lys268 and Glu275 are required. *Biochemistry* **41**, 2217–2226.
24 Fibriansah G, Veetil VP, Poelarends GJ & Thunnissen A-MW (2011) Structural basis for the catalytic mechanism of aspartate ammonia lyase. *Biochemistry* **50**, 6053–6062.
25 Gendron N, Breton R, Champagne N & Lapointe J (1992) Adenylosuccinate lyase of *Bacillus subtilis* regulates the activity of the glutamyl-tRNA synthetase. *Proc Natl Acad Sci* **89**, 5389–5392.
26 Bulusu V, Srinivasan B, Bopanna MP & Balaram H (2009) Elucidation of the substrate specificity, kinetic and catalytic mechanism of adenylosuccinate lyase from *Plasmodium falciparum*. *Biochim Biophys Acta* **1794**, 642–654.
27 Boitz JM, Strasser R, Yates PA, Jardim A & Ullman B (2013) Adenylosuccinate synthetase and adenylosuccinate lyase deficiencies trigger growth and infectivity deficits in *Leishmania donovani*. *J Biol Chem* **288**, 8977–8990.
28 Lee TT, Worby C, Bao Z-Q, Dixon JE & Colman RF (1999) His68 and His141 are critical contributors to the intersubunit catalytic site of adenylosuccinate lyase of *Bacillus subtilis*. *Biochemistry* **38**, 22–32.
29 Segall ML & Colman RF (2004) Gln212, Asn270, and Arg301 are critical for catalysis by adenylosuccinate lyase from *Bacillus subtilis*. *Biochemistry* **43**, 7391–7402.
30 Macol CP, Tsuruta H, Stec B & Kantrowitz ER (2001) Direct structural evidence for a concerted allosteric transition in *Escherichia coli* aspartate transcarbamoylase. *Nat Struct Mol Biol* **8**, 423–426.
31 Smith JL, Zaluzec EJ, Wery J-P, Niu L, Switzer RL, Zalkin H & Satow Y (1994) Structure of the allosteric regulatory enzyme of purine biosynthesis. *Science* **264**, 1427–1433.
32 Goodfellow M & Magee JG (1998) Taxonomy of mycobacteria. In *Mycobacteria* (Gangadharam PRJ & Jenkins PA, eds), pp. 1–75. Chapman & Hall, London.
33 Sang PB & Varshney U (2013) Biochemical properties of MutT2 proteins from *Mycobacterium tuberculosis* and *M. smegmatis* and their contrasting antimutator roles in *Escherichia coli*. *J Bacteriol* **195**, 1552–1560.
34 Wheeler PR & Ratledge C (1994) Metabolism of *Mycobacterium tuberculosis*. In Tuberculosis pathogenesis, protection, and control (Bloom BR, ed), pp. 353–385. ASM Press, Washington DC.
35 Cole S, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon S, Eiglmeier K, Gas S & Barry CR (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544.
36 Wheeler PR (1987) Enzymes for purine synthesis and scavenging in pathogenic mycobacteria and their distribution in *Mycobacterium leprae*. *J Gen Microbiol* **133**, 3013–3018.
37 Long MC, Escuyer V & Parker WB (2003) Identification and characterization of a unique adenosine kinase from *Mycobacterium tuberculosis*. *J Bacteriol* **185**, 6548–6555.
38 McNally T, Helfrich RJ, Cowart M, Dorwin SA, Meuth JL, Idler KB, Klute KA, Simmer RL, Kowaluk EA & Halbert DN (1997) Cloning and expression of the adenosine kinase gene from rat and human tissues. *Biochem Biophys Res Commun* **231**, 645–650.
39 Xu G, Liu B, Wang F, Wei C, Zhang Y, Sheng J, Wang G & Li F (2013) High-throughput screen of essential gene modules in *Mycobacterium tuberculosis*: a bibliometric approach. *BMC Infect Dis* **13**, 227.
40 Sassetti CM, Boyd DH & Rubin EJ (2003) Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**, 77–84.
41 Sassetti CM & Rubin EJ (2003) Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci* **100**, 12989–12994.
42 Gasteiger EHC, Gattiker A, Duvaud S, Wilkins MR, Appel RD & Bairoch A (2005) Protein Identification and Analysis Tools on the ExPASy Server. In The Proteomics Protocol Handbook (Walker JM, ed), pp. 571–607. Humana Press, New Jersey.
43 Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.

44 Chayen NE, Shaw Stewart P, Maeder D & Blow D (1990) An automated system for micro-batch protein crystallization and screening. *J Appl Crystallogr* **23**, 297–302.
45 McPherson A (1985) Crystallization of macromolecules: general principles. *Methods in Enzymology*. **114**, 112–120.
46 Leslie AG & Powell HR (2007) Processing diffraction data with Mosflm. In Evolving methods for macromolecular crystallography, pp. 41–51. Springer, Netherlands.
47 CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760–763.
48 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG & McCoy A (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr Sect D: Biol Crystallogr* **67**, 235–242.
49 Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
50 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
51 Cowtan K (2006) The Buccaneer software for automated model building. *Acta Crystallogr Sect D: Biol Crystallogr* **62**, 1002–1011.
52 Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung L-W, Read RJ & Adams PD (2007) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr Sect D: Biol Crystallogr* **64**, 61–69.
53 Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ & Grosse-Kunstleve RW (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr Sect D: Biol Crystallogr* **66**, 213–221.
54 Emsley P, Lohkamp B, Scott W & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr Sect D: Biol Crystallogr* **66**, 486–501.
55 Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr Sect D: Biol Crystallogr* **53**, 240–255.
56 Winn M, Isupov M & Murshudov G (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr Sect D: Biol Crystallogr* **57**, 122–133.
57 Laskowski RA, MacArthur MW, Moss DS & Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26**, 283–291.
58 Šali A & Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779–815.
59 Fiser A & Šali A (2003) Modeller: generation and refinement of homology-based protein structure models. *Methods Enzymol* **374**, 461–491.
60 Higgins DG & Sharp PM (1988) CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**, 237–244.
61 Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucl Acids Res* **16**, 10881–10890.
62 Konagurthu ASRC, Schmidberger JW, Irving JA, Lesk AM, Stuckey PJ, Whisstock JC & Buckle AM (April 2010) MUSTANG-MR structural sieving server: applications in protein structural analysis and crystallography. *PLoS ONE* **5**, 1–7.
63 Krissinel E & Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr Sect D: Biol Crystallogr* **60**, 2256–2268.
64 Cohen G (1997) ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. *J Appl Crystallogr* **30**, 1160–1161.
65 Krissinel E & Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774–797.
66 Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML & Thornton JM (1997) PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends Biochem Sci* **22**, 488–490.
67 Tamura K, Dudley J, Nei M & Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596–1599.
68 DeLano WL (2002) The PyMOL molecular graphics system.
69 Tornheim K & Lowenstein JM (1972) The purine nucleotide cycle The production of ammonia from aspartate by extracts of rat skeletal muscle. *J Biol Chem* **247**, 162–169.