J Mol Model (2008) 14:249–263
DOI 10.1007/s00894-008-0267-2

ORIGINAL PAPER

Molecular modelling and comparative structural account of aspartyl β-semialdehyde dehydrogenase of *Mycobacterium tuberculosis* (H37Rv)

Anupama Singh · Hemant R. Kushwaha · Pawan Sharma

Received: 9 July 2007 / Accepted: 3 January 2008 / Published online: 31 January 2008

Abstract Aspartyl β-semialdehyde dehydrogenase (ASADH) is an important enzyme, occupying the first branch position of the biosynthetic pathway of the aspartate family of amino acids in bacteria, fungi and higher plants. It catalyses reversible dephosphorylation of L-β-aspartyl phosphate (βAP) to L-aspartate-β-semialdehyde (ASA), a key intermediate in the biosynthesis of diaminopimelic acid (DAP)—an essential component of cross linkages in bacterial cell walls. Since the aspartate pathway is unique to plants and bacteria, and ASADH is the key enzyme in this pathway, it becomes an attractive target for antimicrobial agent development. Therefore, with the objective of deducing comparative structural models, we have described a molecular model emphasizing the uniqueness of ASADH from *Mycobacterium tuberculosis* (H37Rv) that should generate insights into the structural distinctiveness of this protein as compared to structurally resolved ASADH from other bacterial species. We find that *mt*ASADH exhibits structural features common to bacterial ASADH, while other structural motifs are not present. Structural analysis of various domains in *mt*ASADH reveals structural conservation among all bacterial ASADH proteins. The results suggest that the probable mechanism of action of the *mt*ASADH enzyme might be same as that of other bacterial ASADH. Analysis of the structure of *mt*ASADH will shed light on its mechanism of action and may help in designing suitable antagonists against this enzyme that could control the growth of *Mycobacterium tuberculosis*.

Keywords *Mycobacterium tuberculosis* · Aspartyl β-semialdehyde dehydrogenase · Protein structure · Homology modelling · Molecular dynamics simulation

Anupama Singh and Hemant R. Kushwaha contributed equally to this work.

A. Singh · H. R. Kushwaha
Centre of Computational Biology and Bioinformatics (CCBB), School of Information Technology, Jawaharlal Nehru University, New Delhi 110067, India

P. Sharma (✉)
International Centre for Genetic Engineering and Biotechnology (ICGEB), Aruna Asaf Ali Road, New Delhi 110067, India
e-mail: pawans@icgeb.res.in

Background

Throughout the world, infectious diseases caused by *Mycobacteria*, namely tuberculosis (TB) and leprosy, are a major cause of morbidity and a large number of deaths [1]. World Health Organisation (WHO) estimates suggest that about one-third of the world's population is infected and more than two million people die from TB annually. The rise of Multiple Drug Resistant (MDR) strains, evolution of various subtypes, the inefficiency of BCG (Bacillus Calmette-Guerin) as a vaccine and co-infection with HIV, make it imperative that novel antimicrobial drugs are developed that can offer acute precision, shorten the time course of treatment for better patient compliance, and more effectively treat TB. The most pertinent strategy is based on the principles of selective toxicity towards key enzymes or structural components of the pathogen. Enzymes in the lysine and leucine biosynthetic pathways have been reported as potential targets for inhibiting growth of *Mycobacterium tuberculosis* [2]. The proteins involved in these pathways are present in microbes and plants, but not in mammals; hence inhibition of these enzymes should affect only the pathogenic organism with little toxicity to the patient [3]. One of the key regulatory enzymes in the

lysine biosynthetic pathway, aspartyl β-semialdehyde dehydrogenase (ASADH), is a logical target for antimicrobial drug development [4] and that calls for characterisation of its three-dimensional structure.

Aspartate β-semialdehyde dehydrogenase (ASADH, EC-1.2.1.11) is an essential, NADP-dependent enzyme in the biosynthetic pathway of the aspartate family of amino acids in plant systems [5, 6]. The aspartate family comprises essential amino acids and constitutes almost one-quarter of all naturally occurring amino acids. ASADH catalyses a reductive dephosphorylation of aspartyl β-phosphate to aspartate β-semialdehyde in a reversible oxidation–reduction reaction in which NADPH acts as hydrogen donor to the substrate (Fig. 1).

The product aspartyl β-semialdehyde is a key branch point in the aspartyl family biosynthetic pathway and is either further reduced to homoserine, leading to biosynthesis of the amino acids methionine, threonine, isoleucine, or undergoes condensation with pyruvate, leading to dihydropicolinic acid (DAP). DAP is an integral cross-linking component in the peptidoglycan layer of Gram-negative bacterial cell walls as well as being an immediate precursor of lysine—its decarboxylation leads to lysine biosynthesis [7]. It is well established that inhibition of DAP biosynthesis leads to cell death due to peptidoglycan instability and the absence of this pathway in humans makes it a suitable target in *Mycobacterium* species [8]. The gene sequence of *asd* has previously been reported from several organisms including *Escherichia coli* [9], *Salmonella typhimurium* [10], *Streptococcus mutans* [11], *Corneybacterium glumaticum* [12], *Saccharomyces cerevisiae* [13], *Leptospira interrogans* [14], *Mycobacterium smegmatis*, *Mycobacterium avium*, *Mycobacterium leprae* and *Mycobacterium tuberculosis* [15]. It has been demonstrated that mutations in the *asd* genes of *Salmonella typhimurium* and *Streptococcus mutans* result in loss of virulence; the resulting mutants have been used to develop candidate vaccines [16].

Thus, ASADH has been the focus of sustained interest in the field of development of effective inhibitors. Consequently, ASADH is one of the most widely studied enzymes and the crystal structures of several ASADH from Gram-negative bacteria and plants have been solved. The structures of ASADH from *Escherichia coli* (*ec*ASADH) [17], *Vibrio cholerae* (*vb*ASADH) [18], *Methanococcus jannaschii* (*mj*ASADH) [19] and *Haemophilus influenzae* (*hi*ASADH) [20] have been solved experimentally. The structures of *ec*ASADH and *vb*ASADH have been solved in two states: the apoenzyme and a complex with cofactor NADP, and covalently bound to an active site inhibitor, S-methyl L-cysteine sulphoxide (SMCS) [17, 18]. Mutant studies in

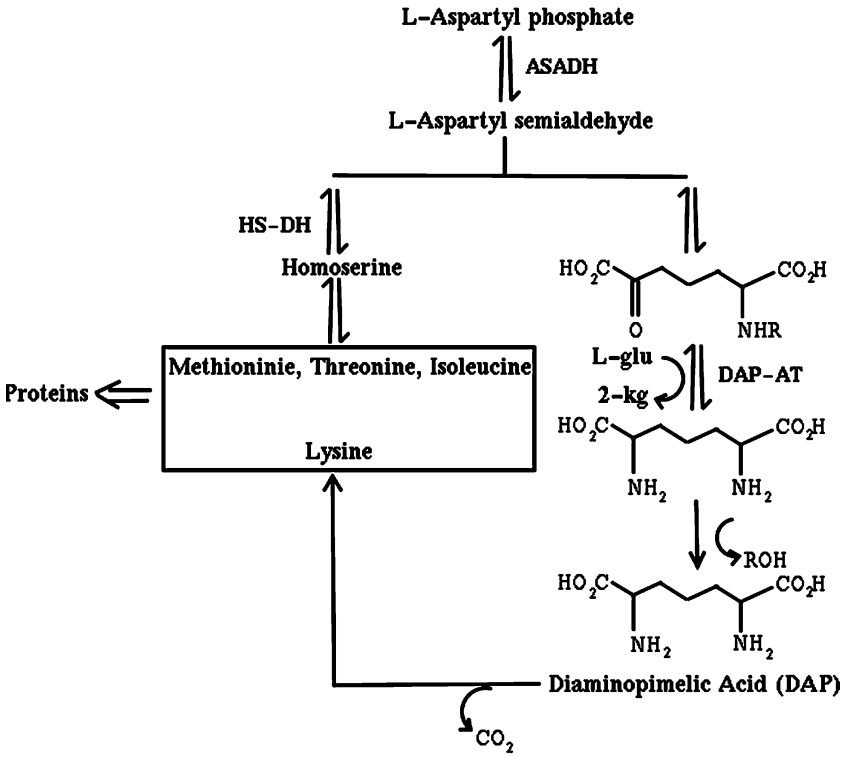


Fig. 1 The pathway leading to the biosynthesis of the essential amino acids lysine, methionine, isoleucine and threonine from aspartate, catalysed by aspartyl β-semialdehyde dehydrogenase (ASADH)

H. influenzae and *V. cholerae* have established the identity of both catalytically essential and catalytically relevant amino acid residues in ASADH [21–23]. Additional information about the catalytic mechanism of ASADH has been derived from X-ray studies and chemical-modification-based studies of glyceraldehyde 3-phosphate dehydrogenase (GADPH), since the catalytic mechanism of ASADH is similar to that of GAPDH [19]. In both enzymes, catalysis is initiated by a conserved cysteine as nucleophile, and is assisted by an adjacent histidine through formation of a thioester intermediate in the active site domain in each catalytic cycle [17].

ASADH has previously been purified from *Escherichia coli* [24]. Recently, ASADH from *Mycobacterium tuberculosis* (H37Rv) has been cloned, expressed and characterised and the protein sequence of chain A, *mt*ASADH has been submitted to GenBank (AY372113). The *mt*ASADH is a 345 amino-acid-long polypeptide with a theoretical molecular weight of 36 kDa [25].

Although the crystal structure of *mt*ASADH has not yet been resolved, the protein shows high structural identity with the ASADH proteins of *Streptococcus pneumoniae* (unpublished data), *Pseudomonas aeruginosa* (unpublished data), *Escherichia coli*, *Methanococcus jannaschii*, *Vibrio cholerae*, and *Haemophilous influenzae* [17–20], all of which have been crystallographically solved. Thus, we have used ASADH from *Escherichia coli* and *Methanococcus jannaschii* as a template to predict a potential model of *mt*ASADH, and used this modelled protein in comparative studies with the known resolved structures. We have also attempted to predict the probable mechanism of action based on the relative domain moment using molecular dynamics (MD) simulations along with sequence and structural analysis.

Methods

Fold recognition and sequence alignment

The single letter amino acid sequence of aspartate β-semialdehyde dehydrogenase (EC 1.2.1.11) (accession number AY372113, protein id AAQ75346) [23] was retrieved from the NCBI database and taken as the target sequence. For phylogenetic analysis, the target sequence was aligned against all the ASADH sequences (EC 1.2.1.11) reported in the SwissProt database. Alignment was carried out using MUSCLE (ver 3.6) [26, 27] and subsequently subjected to minor manual editing. Using the sequence alignment, bootstrap analyses were performed with 1,000 replicas. Distances were calculated using the protdist program in the Phylip (ver. 3.66) package [28]. Using neighbour joining, a consensus tree was obtained, and the rooted tree was plotted using the drawgram program from the Phylip package (http://evolution.genetics.washington.edu/phylip.html).

Secondary structure was predicted using JNET [29], SABLE [30], PREDATOR [31–34], STRIDE [33], PSIPRED [35] and SAM [36]. Fold-recognition analysis was carried out using FUGUE [37], INBGU [38], mGE NETHREADER [39], FFAS03 [40], 3DPSSM [41]. The architectural motifs and the topology of proteins with known three-dimensional structure were analysed according to SCOP [42] and CATH [43] classifications.

Homology modelling and analysis

The three-dimensional structure of ASADH of *Mycobacterium tuberculosis* (*mt*ASADH) was modelled in a stepwise procedure, starting with a template structure having PDB ID: 1GL3.pdb for *E. coli* (NADP bound); chain B, 1YS4.pdb for *Methanococcus jannaschii*; chain B, and 1BRM.pdb for *E. coli* (NADP unbound) were taken from the Protein Database (PDB) server (http://pdbbeta.rcsb.org/pdb/Welcome.do). The NADP-bound conformation was generated using 1GL3.pdb, 1YS4.pdb, and the unbound conformation was generated using 1BRM.pdb. These template structures were aligned using structure alignment software STAMP [44, 45]. These aligned structures were used as a profile for aligning the target sequence using ClustalX [46]. The automated comparative protein modelling program MODELLER8v2 [47] was then used to generate a 50 all-atom model by alignment of the target sequence with the selected template sequence in an alignment file. The best model was chosen on the basis of the stereochemistry quality report generated using PROCHECK [48], and side chains were optimised using SCWRL 3.0 [49]. The bond distance and dihedral angle restraints on the target sequence were derived from its alignment with the template three-dimensional structures. The spatial restraints and the energy minimisation steps were performed within Modeller using the CHARMM22 force field for proper stereochemistry of proteins. Further evaluation of the modelled *mt*ASADH structure was done using the PROSA II [50] tool, which compares the z-score of target and templates. The presence of conserved structural motifs was studied using STRIDE [33] on modelled *mt*ASADH in comparison to motifs present in ASADH from *Escherichia coli* and *Methanococcus jannaschii*. Structural motifs present in known templates were also studied using STRIDE. Molecular visualisation and analysis of the final model were carried out with Visual Molecular Dynamics (VMD) (http://www.ks.uiuc.edu/Research/vmd/) [51]. The electrostatic binding energy of NADP with *mt*ASADH and *mj*ASADH (1YS4.pdb) was calculated using APBS (Adaptive Poisson Boltzmann Solver) [52]. Domain movement in *mt*ASADH was studied using DynDom software (http://www.cmp.uea.ac.uk/dyndom/dyndomMain.do) [53].

The X-ray structure of the ASADH proteins of *Escherichia coli*, *Methanococcus jannaschii* and modelled *mt*ASADH were embedded in a cubical solvent box consisting of 17,244, 17,374 and 12,043 TIP3 water molecules, respectively. The water molecules were placed 5 Å around the protein molecule. Capping of the C-terminal and N-terminal ends was performed. To produce a neutral system mimicking physiological conditions, 11, 11 and 13 sodium ions were added to the water-embedded *mj*ASADH, *mt*ASADH and *ec*ASADH, respectively. The cut-off range was set to 12 Å for van der Waals interactions; the switching distance for smoothing was 10 Å, and the nonbonded pairlist distance was 14 Å. The NPT ensemble was maintained with a Langevin thermostat and isotropic Langevin Piston. Each system was minimised for 10^5 steps using conjugate gradient and the line search algorithm. This was followed by 400 ps equilibration at 300 K. A time step of 2 fs was employed in the Verlet algorithm. Production runs of 2 ns were completed with an average temperature of 300 K. The SHAKE algorithm was used to constrain hydrogen bond length [54]. MD simulations were performed on a 16 node dual processor AMD Opteron 242 machine (CCBB, JNU), taking ~28 h CPU time for each 2 ns simulation for all structures. All simulations were performed using NAMD v2.6 (http://www.ks.uiuc.edu/Research/namd/) [55] with a CHARMM22 force field.

Docking studies were performed using Flexx Software (v 2.2.0) [56]. Docking of the molecule was performed using default parameters. The small molecule database was first filtered using SMCS—a known inhibitor of the ASADH enzyme [17, 18]. Molecules with 60% structural similarity were picked for final screening. The initial screening was performed using jcsearch program of JChem software (ver. 3.2) [57].

Results and discussion

Identification of the three-dimensional folds

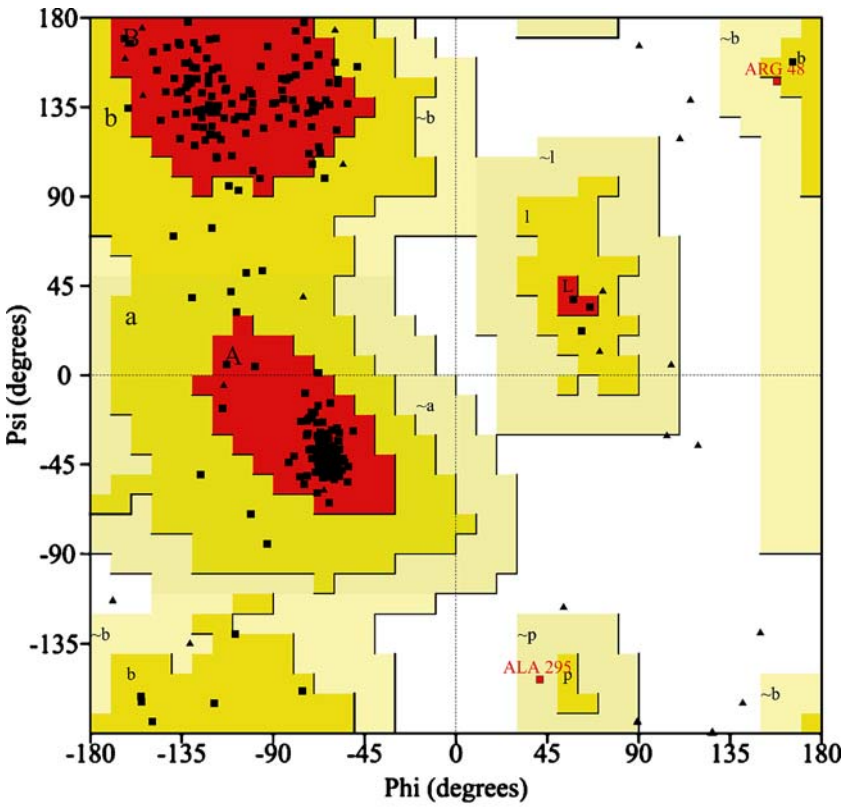
To create a model of *mt*ASADH, we first performed a BLAST search against the Protein Data Bank (PDB) for proteins with similar sequence and known 3D structures using the 345-residue-long *mt*ASADH sequence obtained from NCBI. The search result identified structures of ASADH in *Streptococcus pneumoniae*, *Pseudomonas aeruginosa*, *Methanococcus jannaschii*, *Escherichia coli*, *Haemophilus influenzae* that can be considered as possible templates for modelling *mt*ASADH using the threading approach. The threading approach helps to assess the compatibility of the target sequence with the available protein folds based not only on sequence similarities but also on structural considerations [58, 59]. The conserved domains of the sequence were identified from Pfam (Protein family database [60]). The analysis of the results obtained from Pfam showed the presence of two major domain signatures in the sequence. The first signature domain is the NAD-binding domain, which extends from residue 4 to 119, and the other domain signature corresponds to the dimerisation domain, encompassing residues 139 to 328 of *mt*ASADH. The conserved domain search against the conserved domain database, CDD [61, 62], also supported the Pfam results. For further identification of all the folds and domains, fold-recognition servers such as FUGUE and 3DPSSM were used. Results from the servers show that the models of *Methanococcus jannaschii*, *Escherichia coli* and *Pseudomonas aeruginosa* can be good templates with significant scores. As they are currently unpublished, the structures from *Streptococcus pneumoniae* and *Pseudomonas aeruginosa* were utilized only for finding conserved regions in *mt*ASADH. The structures of ASADH from *Methanococcus jannaschii*, *Escherichia coli* and *Haemophilus influenzae* were used as a template for modelling. These results show that the major secondary structure and fold region in *mt*ASADH is well conserved despite the low sequence identity within the family (~24–39%).

Comparative modelling of mtASADH

X-ray resolved structures of ASADH from *Methanococcus jannaschii*, *Haemophilus influenzae* and *Escherichia coli* are available from Protein Data Bank (PDB). Hence, we modelled ASADH from *Mycobacterium tuberculosis* using a comparative modelling strategy [63]. Since the protein structures used for the alignment and modelling have well conserved structural motifs and regions, and functional information is also available, the problem of low-sequence identity could be overcome, i.e. a multiple sequence alignment obtained from the known sequences can provide a reasonable approach to comparative structure modelling. In order to verify the quality of the sequence alignment and optimise the position of gaps, corresponding positions from secondary structures were used. Experimentally determined structures were aligned and used as a profile for aligning the target sequence and to obtain an alignment, which was then used for modelling *mt*ASADH.

To generate a 3D model of *mt*ASADH, two sets of 50 structural models each were generated using MODELLER8v2, one set using 1BRM.pdb (*ec*ASADH, unbound conformation) and the other using 1GL3.pdb (*ec*ASADH NADP-bound conformation). Ramachandran plots were generated for two *mt*ASADH structures without NADP, and with NADP to determine deviations from normal bond lengths, dihedrals and nonbonded atom–atom distances, and with a view to comparing the *mt*ASADH model against the ASADH structure solved by X-ray crystallography. The plot shows that the modelled *mt*ASADH (Fig. 2) has 91%

residues in most favourable regions with the remaining 9% of residues occurring in allowed regions. No residue was found in the disallowed region. This is expected for crystallographic models with at least 2.2 Å resolution. X-ray resolved structures give comparable stereochemical data: that of the *ec*ASADH is 86% favourable, with 13.7% in the additional allowed region, and in *mj*ASADH 90.4% is favourable with 9.6% in the additional allowed region. The PROCHECK result summary showed 10 out of 346 residues labelled while the torsion angles of the side chain designated by $\chi1-\chi2$ plots showed only 7 labelled residues out of 171. All main-chain and side-chain parameters were found to be in the 'better' region. G-factor is essentially a log odds score based on the observed distribution of stereochemical parameters such as main chain bond angles, bond length and phi–psi torsion angles. The score for G-factors should be above −0.50 for a reliable model. The observed G-factor scores of the present model were −0.09 for dihedral bonds, −0.30 for covalent bonds and −0.17 overall. The distribution of the main chain bond lengths and bond angles were 98.4% and 90.6% within limits. This distribution in the bond angles is attributed to many residues that are not conserved in *mt*ASADH in comparison to the template structures. For example, from Phe193 to Val215 the conservation is low and *mt*ASADH has more proline residues in this region, resisting the formation of any secondary structure normally present in the template structures. In the modelled structure, only two bond angles, at Gly73 and Ala330, show deviation from their respective standard values. Conclusively, the modelled structure of *mt*ASADH is comparable to the structurally resolved ASADHs as structural motifs are conserved. The



Plot statistics

Residues in most favoured regions [A,B,L]	263	91.0%
Residues in additional allowed regions [a,b,l,p]	24	8.3%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.7%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	289	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	33	
Number of proline residues	21	
Total number of residues	345	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Fig. 2 Ramachandran plot of ψ/Φ [distribution of ASADH from *Mycobacterium tuberculosis* (*mt*ASADH) model produced by PROCHECK]

PROSA II energy plots for *mt*ASADH with its template show all residues to have negative interaction energy (Fig. 3). With the exception of a few residues in *ec*ASADH, the template structures display positive interaction energy. The z-score of pair, surface and combined energy for modelled *mt*ASADH are −10.05, −7.87 and −6.30, respectively, while the corresponding values for *mj*ASADH and *ec*ASADH are −9.33, −6.14, and −7.86, and −12.25, −8.78 and −8.08, respectively. The comparable z-score values and the interaction energy results further confirm the quality of the modelled structure. Additionally, electrostatic binding energy calculated using APBS [65] for ligand (NADP) and protein is also found to be comparable for NADP-*mt*ASADH and NADP-*mj*ASADH at −2.23 kJ mol^{-1} and −1.53 kJ mol^{-1}, respectively. Regions of the secondary structure were also verified using PREDATOR and STRIDE software. Time-dependent Cα rms deviations (RMSD) were used to provide a picture of the global drift of the homology model in comparison to the template structures. Thus, during simulation, the RMSD drift of Cα atoms from the initial protein structure was determined (Fig. 4).

The modelled *mt*ASADH was used to screen a database of ligands that could act as potent inhibitors of *mt*ASADH using Flexx Software (ver 2.2.0). Flexx predicts the geometry as well as the free energy of the binding of the protein and ligand molecule [56]. A list of the ten best-scoring ligands along with their respective IUPAC names, as well as a natural known inhibitor of the ASADH enzyme is shown in Table 1.

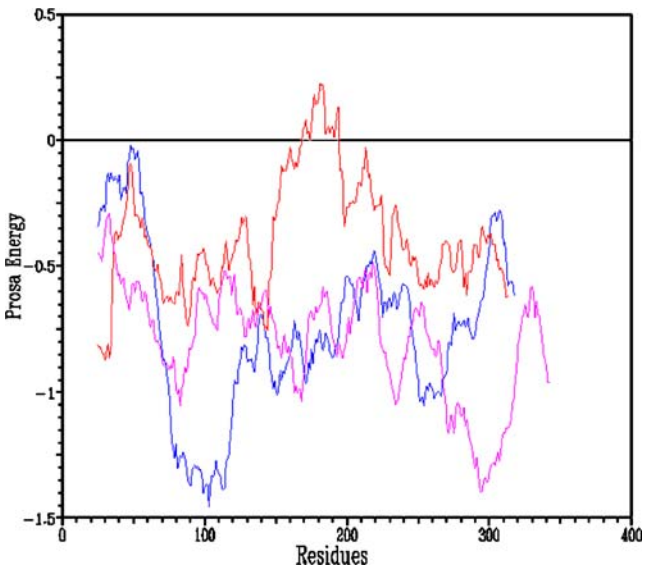


Fig. 3 PROSA energy profiles calculated for the modelled *mt*ASADH, template ASADH from *Escherichia coli* (*ec*ASADH) and *Methanococcus jannaschii* (*mj*ASADH). *x-axis* Residue numbers, *y-axis* Prosa energy. *Blue line mt*ASADH, *red line mj*ASADH, *magenta line ecASADH*

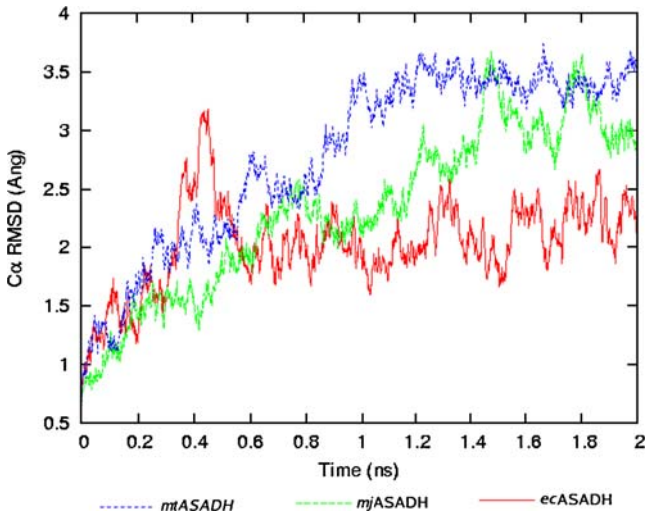


Fig. 4 Analysis of the molecular dynamics (MD) simulation. Drift of protein structure from the initial model. The time-dependent Cα rms deviations (RMSD) of all Cα atoms from the starting structure are shown as a function of time

The three-dimensional structure of mtASADH

A comparison of the amino acid sequence of *mt*ASADH with that of other enzyme members across the family reveals more than 40 conserved amino acid residues. Despite the very low level of sequence identity, *mt*ASADH retains the same overall fold and, more importantly, appears to have the same catalytic mechanism as other bacterial ASADHs. Conserved residues include those directly involved in substrate binding and catalysis, an N-terminal GxxGxxG Rossmann fold motif, and a C-terminal GAA sequence. As shown in Fig. 5b, *mt*ASADH might be present as a functional dimer. Alignment of the *mt*ASADH sequence shows a number of insertions and deletions with respect to other ASADH members that have structural, and thus probable functional, significance (Fig. 6).

The first β-α unit (α1-β1) contains the highly conserved GxxGxxG NAD-binding motif (residues 1–10)—a conserved motif commonly present among all ASADHs. It is well known that GADPH performs very similar chemistry and both enzymes have been shown to react via a stable acyl intermediate [24, 64]. Previously homology studies between ASADH and GADPH were used to inform mutational studies in ASADH [65]. Both binding domains as well as the dimerisation domain show similar folds with slightly different orientations [17].

NADP binding domain

The N-terminal Rossmann fold domain has the same overall architecture as that of *vc*ASADH and *mj*ASADH. This domain is conventionally built around three β-α-β motifs that form six parallel β-strands sandwiched by three α-

Table 1 Flexx docking results of aspartyl β-semialdehyde dehydrogenase (ASADH) inhibitors

S.No.	IUPAC name	Energy (kJ mol^{-1})
1	S-Methyl-L-cysteine sulfoxide	−26.88
2	4-Aminiumyl-3,3-dimethylbutanoate	−24.72
3	(2R)-4-Aminiumyl-2-methylbutanoate	−22.84
4	(3S)-5-Oxothiomorpholine-3-carboxylate	−20.94
5	(2R)-4-Aminiumyl-2-hydroxy-3,3-dimethylbutanoate	−19.37
6	4-Amino-2-hydroxy-3,3-dimethyl-butyric acid	−19.37
7	3-Mercapto-adamantane-1-carboxylic acid	−18.17
8	(3-Mercapto-adamantan-1-yl)-acetic acid	−17.33
9	3-(Phenylsulfanyl)propanoate	−15.67
10	4-Amino-3-hydroxy-4-methyl-pentanoic acid amide	−15.47
11	2-[(2S)-2-Sulfanylpropanamido]acetate	−9.94

helices (Fig. 5a). The N-terminus of *mt*ASADH forms the first β-α-β motif starting from the first β-strand (β1) (residues 3–9) and the first helix (α1) (residues 13–24), followed by two short strands β2 (residues 29–37)–β3 (residues 51–54). The motif is connected to a short alpha helix α2 (residues 38–43). The end of the surface loop leads to the second β-α-β motif [β4 (residues 67–71)-α4 (residues 74–88)-β5 (residues 90–94)], which is similar to other

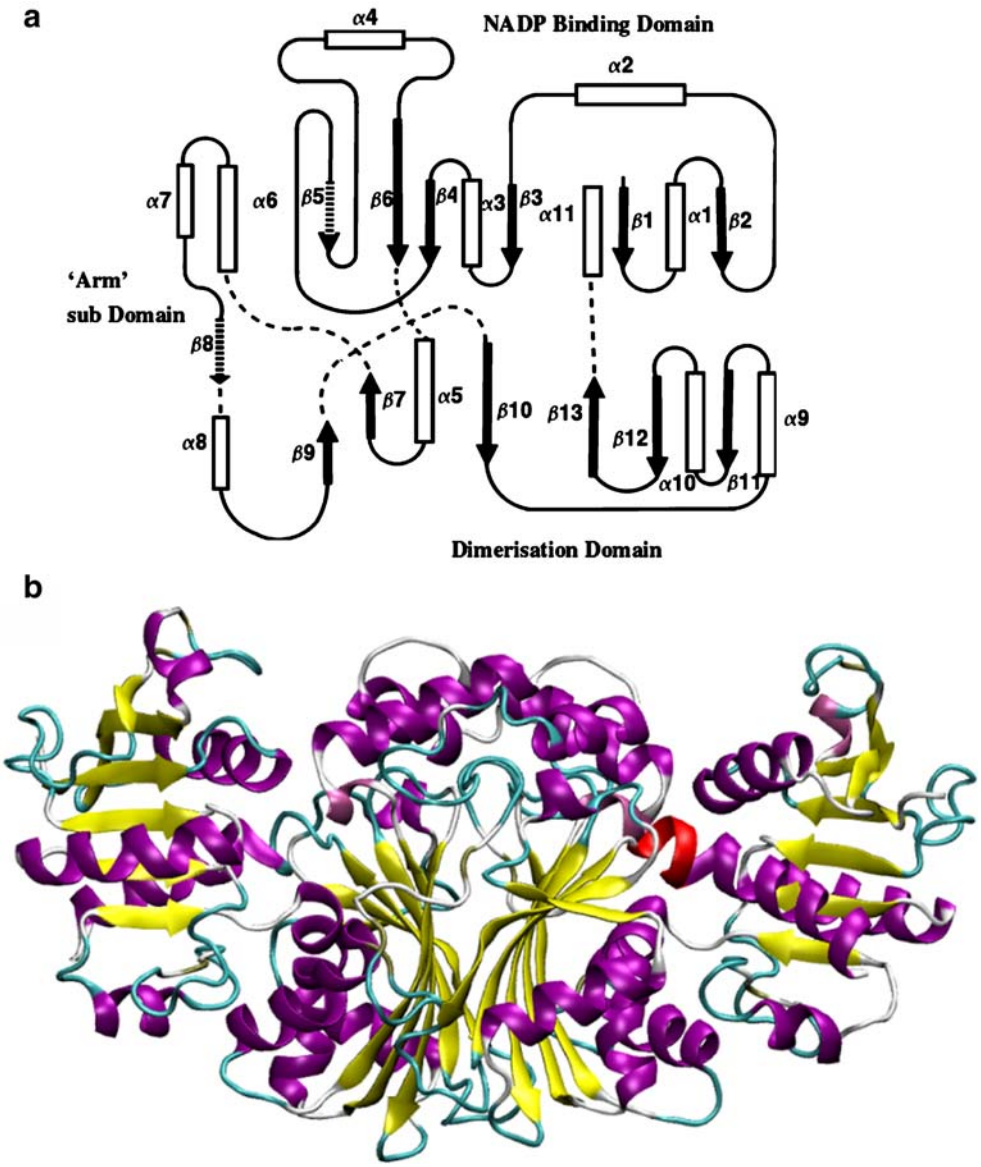


Fig. 5 **a** Secondary structure topology of ASADH showing the NADP binding domain, the "Arm" sub domain and the dimerisation domain. The *dotted arrow* shows the absence of a β8 secondary structure motif in *mt*ASADH, normally present in the template structures. **b** New cartoon view diagram of a dimer of the modelled *mt*ASADH. The figure was prepared using Visual Molecular Dynamics (VMD) (http://www.ks.uiuc.edu/Research/vmd/)

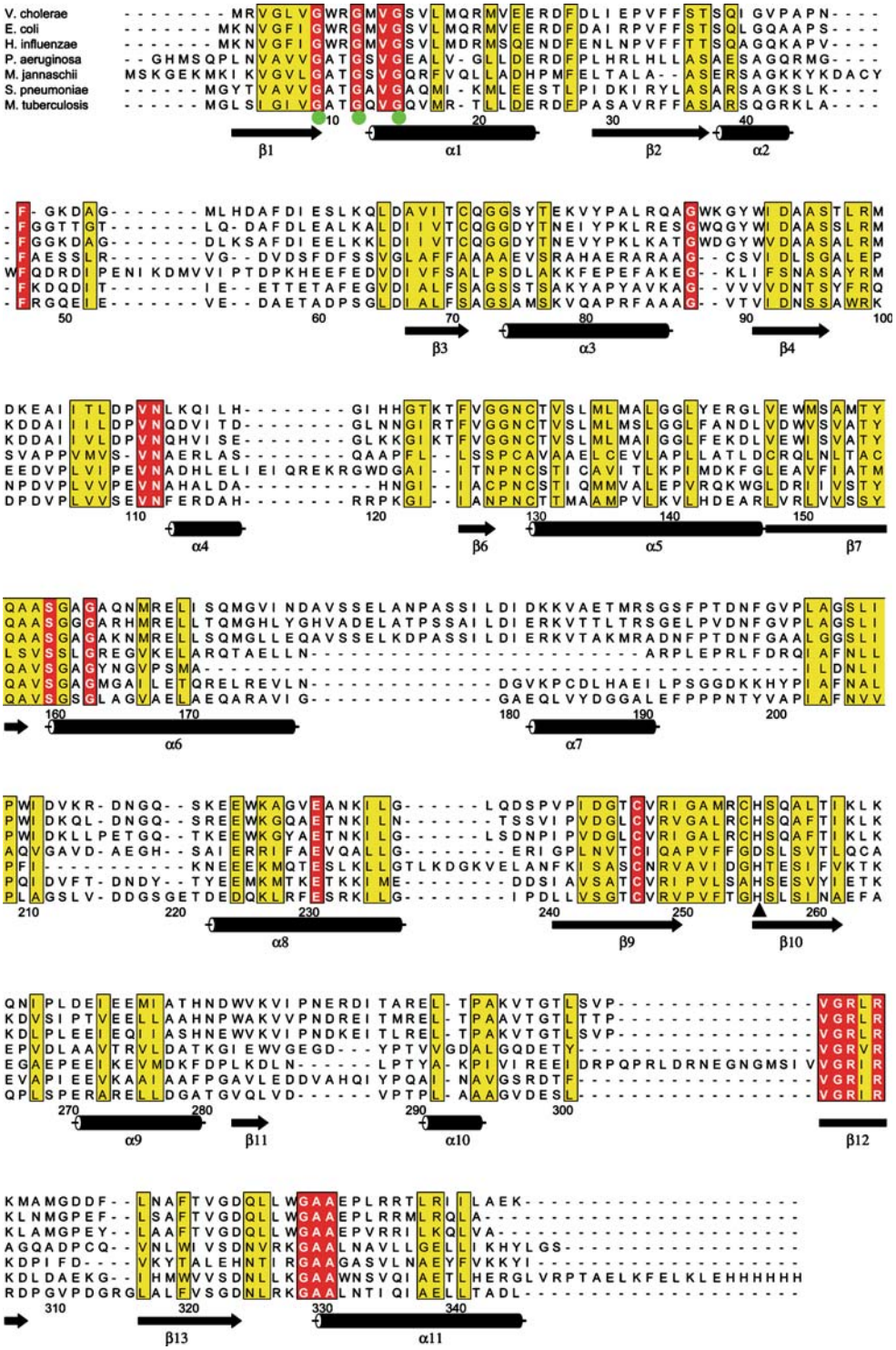


Fig. 6 Alignment of ASADH sequences, showing structural motifs in the modelled *mt*ASADH. The *blue triangles* show the presence of conserved Cys130 and His256. The *green dots* highlight the conserved GxxGxxG region. Secondary structure elements are indicated (α-helices as *cylinders* and β-strands as *arrows*). The absence of β5, β8 in the alignment shows the absence of motifs in modelled *mt*ASADH, which are otherwise present in the template structures. Each motif is marked by its representative secondary structure. The sequenced is numbered with respect to *mt*ASADH. The figure was prepared using Alscript program [72]

bacterial ASADHs. The third β-α-β motif of the Rossmann fold is less conserved in *mt*ASADH but the region is most similar to that of *mj*ASADH compared to the other bacterial ASADH. The first beta sheet (β5) is missing from *mt*ASADH, which instead contains an unstructured loop leading into a helix (α4) (residues 113–120) and then to the final sheet of the fold (β6) (residues 124–127). As in the case of *mj*ASADH, the absence of the beta sheet in this region might contribute to the altered co-enzyme properties of this enzyme [19].

NADP binding to the apoenzyme induces conformational change in *ec*ASADH [66]. Structural analysis of NADP-bound and unbound ASADH can provide some information on the catalytic features of the enzyme. In *ec*ASADH,

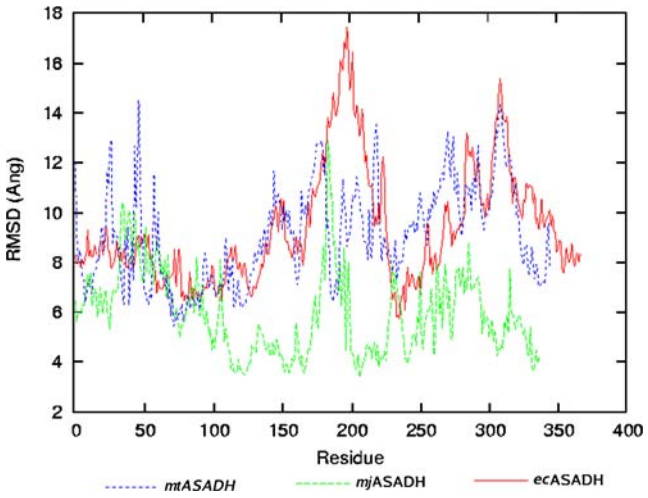


Fig. 7 Fluctuation of protein coordinates. The rms fluctuations of the Cα coordinates from their time-averaged values are shown as a function of residue number for the structures

NADP binding induces domain closure, i.e. changes in the relative position of domains due to significant rotation of the NADP-bound N-domain towards the dimerisation domains [66]. Modelling of *mt*ASADH using NADP-bound and unbound enzyme as template reveals a significant change in the conformation of the protein. To monitor and assess the NADP-induced conformational change in the two modelled *mt*ASADH structures, we relied on Cα trace overlap showing a bend in the Cα chain. Domain movement in *mt*ASADH is observed from the analysis using DynDom, with modelled ASADH in the bound and unbound state. The DynDom program determines the protein domain, hinge axes and amino acid residues involved in hinge bending. Movement in the domain was found to occur only along one axis that divides the protein into a fixed domain (309 residues) and a movable domain (21 residues). The residues involved in bending are residues from Glu172 to Arg175 and from Pro195 to Tyr199. The residues Val177, Ile178, Gly180, Phe193, Pro190 and Pro195 act as a mechanical hinge along which the overall movement in the loop region between α6–α7 and α7– α8 occurs. The absence of the β6 secondary structure in *mt*ASADH, which is otherwise present in ASADH from *Escherichia coli*, lead to the formation of a long loop region, which may allow movement of this region. Significant movement in the region has been observed in MD simulations, possibly due to the absence of any stable secondary structure (Fig. 7). The corresponding changes in the open and closed structures of ASADHs from *V. cholerae* and *E. coli* have been used to deduce the probable mechanism of catalysis [18, 34].

In multiple sequence alignment (Fig. 6), the essential conserved residues for NADP binding are Gln/Asn at position 326 of *mt*ASADH. The Gln/Asn pair is implicated to have an important role in hydrogen bonding with the carboxylic oxygen of the nicotinamide ring of NADP, and it has been found conserved across species in all known ASADH proteins. The other important residues involved in interaction with NADP, Gly161 and Ser162, are also conserved in *mt*ASADH with respect to the known template structures. These residues collectively play an important role in determining the stereospecificity of NADP binding in this enzyme in *Escherichia coli* [66], and their conservation in *mt*ASADH supports a similar role. The other important region, the internal loop from Ala158–Gly163 (in *mt*A-SADH) located in β7-α6, in the cross-over region between the C- and N-domains, is highly conserved in all known species. The consensus sequence S/AxSGS/A/G, where x is a small hydrophobic residue, plays a similar role as the GxxGxxG motif of the Rossmann fold (Fig. 6). These consensus residues have small side chains and offer the least stereo constraints to main chain hydrogen bonding with the nucleotide and pyrophosphate of NADP (Table 2). Experimental studies in *E. coli* identified Arg31 and Glu39 as being involved in binding with the 2′ PO_4 group of NADP [66], and substitution of Gln39 results in a ten-fold reduction in K_m value for NADP [65]. The chemical and kinetic mechanisms of ASADH in *E. coli* are well studied [35]. From various K_{cat}/ K_m values for ASA, PO_4 and NADP of *mt*ASADH, it is proposed that *mt*ASADH is not an efficient enzyme in comparison with the ASADHs of *E. coli*, *V. cholerae* and *H. influenzae* (Table 3) [25, 66]. K_{cat} of *mt*ASADH is comparable to that of *mj*ASADH, suggesting some similarities between the Gram-positive ASADH and *archaeal* ASADH [19]. It should be noted here that the K_m value for NADP binding is significantly lower than this value in other bacteria (Table 3). The residues observed to be involved in major hydrogen bond interaction with the adenine end of NADP in *mt*ASADH are Thr11 and Gly12. Gln13 forms a hydrogen bond with the pyrophosphate oxygen (Fig. 8). In *E. coli* and *V. cholerae*, adenine exhibits a hydrogen bond interaction with Pro193 in *ec*ASADH [19, 66] in an extended internal loop. However, in *mt*ASADH

Table 2 Hydrogen bonds between NADP and modelled ASADH from *Mycobacterium tuberculosis* (*mt*ASADH). Hydrogen bond distances were calculated using HBPLUS [73]

NADP	*mt*ASADH		Distance (Å)
AOP3	SER40	OG	2.67
AOP3	SER37	OG	2.66
AOP3	SER37	N	3.04
NO2	GLN13	N	3.15
AO3*	GLY12	N	2.95
AOP1	THR11	OG1	2.63
AO3*	THR11	OG1	2.90

Table 3 Comparison of the enzymatic efficiency and K_m values of ASADH of different species including *Mycobacterium tuberculosis* (*mt*). The enzymatic efficiency of *mt*ASADH is lowest among the reported ASADH [19, 21, 25, 66]

Organism	ASA(K_m)	ASA(K_{cat}/ K_m)	PO_4(K_m)	PO_4(K_{cat}/ K_m)	NADP(K_m)	NADP(K_{cat}/ K_m)
Escherichia coli	0.17	3.60×10^6	4.8	1.30×10^5	0.19	3.60×10^6
*mt*ASADH	0.96	8.89×10^4	11.39	7.45×10^2	0.07	1.31×10^5
Vibrio cholerae I	0.19	6.30×10^5	1.1	1.10×10^5	0.32	3.80×10^5
V. cholerae II	0.16	3.60×10^5	2.2	3.00×10^4	0.36	1.60×10^5
Haemophilus influenzae	0.24	1.40×10^6	1.6	2.10×10^5	0.2	2.20×10^6
Pseudomonas aeruginosa	0.12	1.30×10^6	1.35	1.10×10^5	0.13	1.60×10^6

and *mj*ASADH none of these residues interact with the adenine ring of NAD. Structural analysis shows that some conserved residues form a stable hydrogen bond interaction with NADP in *mt*ASADH (Table 2). The hydrogen bond interaction between adenine ring N6 with the O of Pro193, which is more prominent in *E. coli* and in *V. cholerae* due to π charge distribution [19, 66], is not present in *mt*ASADH. The absence of the Pro–nicotinamide interaction in *mt*ASADH may also influence the relative orientation of the nicotinamide ring in NADP in the N-domain. NADP binding induces the formation of several new hydrogen bonds such as Ser99 with Arg240, Asn134 with Gly241 and with Lys244 within the side chains of Cα, which stabilises the loop structures around the active site in the *E. coli* enzyme [66]. The corresponding residues in *mt*ASADH, Ser96, Asp223, Asn129, Glu224 and Lys227, are conserved. Since Arg240 in *ec*ASADH is replaced by Asp223 in *mt*ASADH, there is no hydrogen bond interaction with Ser96, while a new hydrogen bond forms between Asn129 and Lys227 as in *E. coli*. However, in *V. cholerae*, the active site loop is already structured and the active site is preformed [19]. A per residue RMSD plot between the bound and unbound conformation of *mt*ASADH shows three regions of large deviation in residues from Ala35 to Asp56, Val185 to Asp225 and Phe320 to Leu345, which corresponds to the N-terminal loop, the surface loop in the "arm" domain and helix α11 of the N-domain, respectively (Fig. 7).

Dimerisation domain

The active site residue Cys-130 is located at the top of helix α6 (residues 130–147), sandwiched between the co-enzyme binding domain and the dimerisation domain. α6 leads to β7, which can be considered as the first beta sheet of the dimerisation domain. The dimerisation domain is composed of six beta sheets and is hydrophobic in nature (Fig. 5a). The end of β7 connects to the surface exposed loop known as the "arm" sub domain of ASADH.

In all the bacterial ASADHs, the arm sub domain begins with a conserved SGxG sequence; in *mt*ASADH it begins with the same conserved sequence. The conserved residues contribute to the interaction with the nicotinamide ring of NADP (Fig. 8). In *mt*ASADH, this domain consists of helix α6 (residues 161–179), which is followed by another helix structure, α7 (residues 182–192). In *mt*ASADH, beta sheet that is otherwise present in other bacterial ASADHs is missing from this domain. The absence of the beta sheet in this region is due to the existence of multiple proline residues present in the sequence, which does not support the formation of stable secondary structure. This region in *mt*ASADH is also not conserved when compared to the other bacterial ASADHs. This domain plays a major role in communication between the two subunits of the enzyme [66], which is reported to be absent in *mj*ASADH [19].

In *ec*ASADH, substrate binding induces an asymmetric rotational twist in the central β-sheet core of the dimer, thereby creating a geometrically favourable catalytic site in one subunit, while distorting it in the other, which gives a suitable explanation of the observed HOSR (Half of Site Reactivity) [66]. The signal of the active state of one monomer is transferred through the linker sequences to the sister subunit via this conformational movement of the catalytically active subunit. A number of electrostatic interactions between the subunits also serve to maintain the conformational integrity of the dimer [20]. The tetrad in *mt*ASADH has conserved residues except for the presence of Ser in place of Thr with respect to the templates (Fig. 6). From the multiple sequence alignment it is observed that replacement of Thr by Ser in the tetrad region is common among ASADH in different species.

As in all the other bacterial ASADHs, a conserved tyrosine is present in *mt*ASADH that plays a major role in forming the hydrogen-bonding network bridging the two active sites. These tyrosine residues are ð-stacked to maintain their hydrogen bonding orientation, and this is further stabilised by edge-on stacking with phenylalanine residues from each subunit. This inter-domain communication results in the half-site reactivity observed in all the bacterial ASADHs, but which is absent in the archaeal enzyme [66].

The dimerisation domain consists of six beta sheets, the middle four of which are arranged in antiparallel fashion. Dimerisation occurs mainly through the formation of a β-sandwich by the β-sheets from the A and B subunits in the dimer [23].

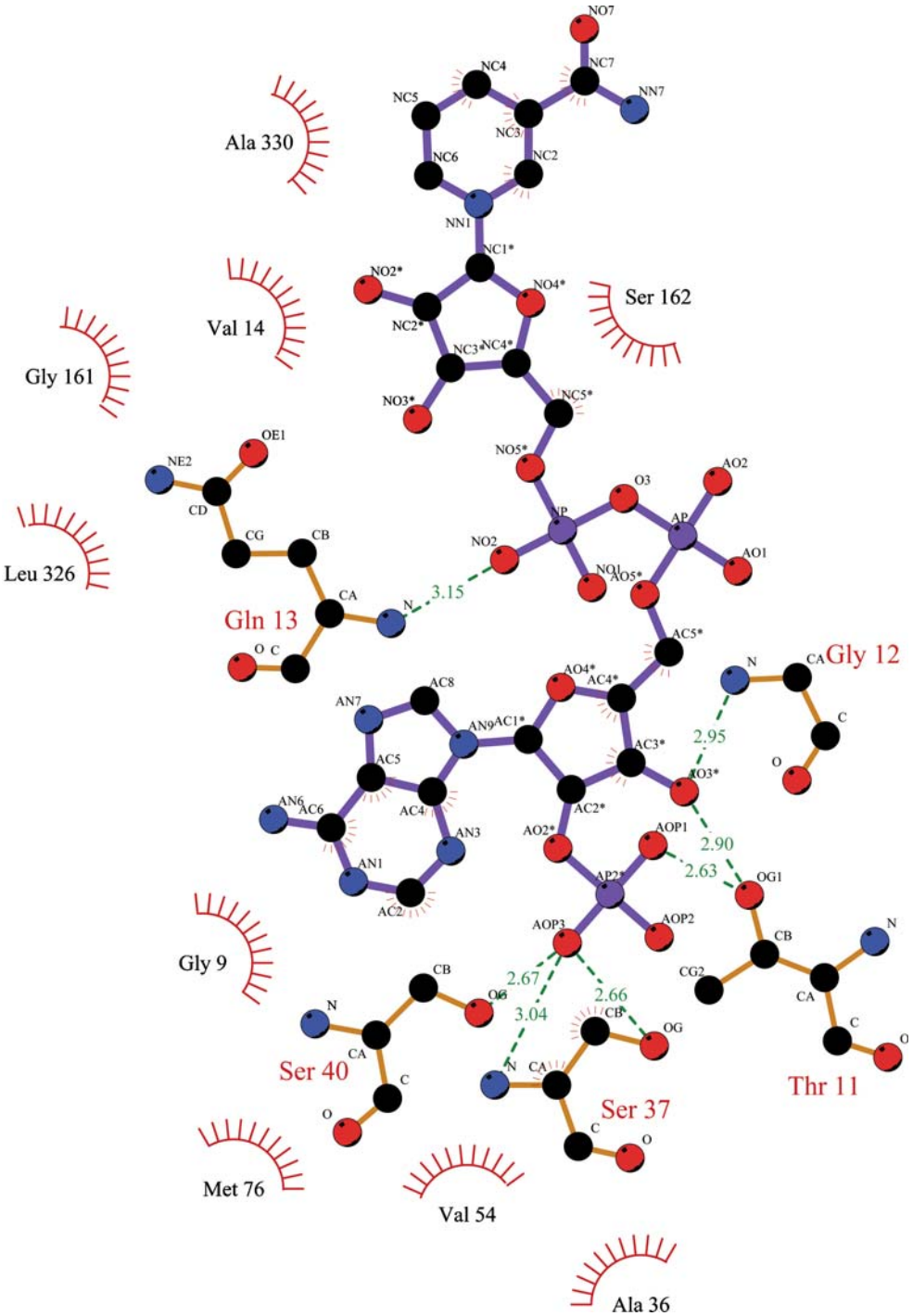


Fig. 8 Interaction of NADP with *mt*ASADH. Ser37, Ser40, Thr11, Gln13 and Gly12 show hydrogen bond interactions with NADP, while other non-ligand residues shown are involved in forming hydrophobic contacts. The figure was prepared using LIGPLOT [74]

Analysis of active site

The active site residue Cys 130 in *mt*ASADH is conserved in all the bacterial ASADHs, and has been identified through biochemical studies as the catalytic cysteine residue that forms an acyl intermediate in the conversion of β-aspartyl phosphate to L-β-aspartate semialdehyde [23]. The importance of cysteine in the catalytic cycle of ASADH has been well established by mutational analysis performed on *hi*ASADH [22]. Two other residues, Gln157 and His256, which are located in the active site cleft, and which are well conserved in all the bacterial ASADHs, play an essential role in catalysis and have been studied by site-directed mutagenesis [65]. The histidine residue has previously been identified to play the role of base catalyst in the proposed mechanism of ASADH catalysis. Mutation studies have shown that

Table 4 Hydrogen bonds between S-methyl L-cysteine sulphoxide (SMCS) and *mt*ASADH. Hydrogen bond distances were calculated using HBPLUS [73]

SMCS	*mt*ASADH		Distance (Å)
O	HIS256	NE2	2.98
SG	HIS256	NE2	2.87
O	ARG249	NH2	2.66
N	GLU22	OE2	2.98

removal of this histidine residue would reduce the activity by 10^3-fold in comparison to the native enzyme [67]. In *E. coli*, it was shown that His274, Gln162 and Arg267 are involved in hydrogen bonds to the single water molecule, which appears to approximate the site of substrate binding. These residues are also conserved in *mt*ASADH and therefore might play the same role as in other bacterial ASADHs. The conserved residue His256 seems to form a stable hydrogen bond with SMCS (Table 4). Hence, conserved residues such as Cys130 and His256 can be good targets for selective inhibitor development against *mt*ASADH.

In *ec*ASADH, the tetrad Thr160-Tyr161-Gln162/Phe345 has been proposed as the linker system allowing signalling between the two active sites; this tetrad is also conserved in other bacterial ASADHs [66]. In *mt*ASADH, Thr is replaced by Ser in the tetrad, which is otherwise well conserved (Ser155-Tyr156-Gln157/Phe320), and thus can be assumed to have same signalling function between the two active sites.

Binding of the common active site inhibitor SMCS has been previously studied in the bound state in *vc*ASADH and *ec*ASADH. In *mt*ASADH, the active site residue Cys130 (corresponding to the Cys135 in *E. coli* and Cys134 in *V. cholerae* ASADHs) plays an essential role in binding, as it forms a disulfide linkage with the inhibitor (Fig. 9). Other active site residues that play an essential role in binding are His256, Arg249, Glu224, which form hydrogen bonds with the inhibitor.

Evolutionary analysis

The phylogenetic tree (Fig. 10) derived from the protein sequences reveals several interesting facts about the phylogenetic relationship amongst the ASADH enzymes of different species. Bootstrap values indicate that the topology of end branches of the phylogeny having higher values can be considered as consistent and correct. However, the topology at higher levels but with low bootstrap values are inconsistent in some cases, e.g. exterior (root) branch of *Aquifex*, *Prochlorococcus*, *Rickettesia* etc. The *Mycobacterium* family and closely related *Corynebacterium* members can be seen to be two closely related family groups. The ASADH enzymes from *Mycobacterium tuberculosis* and

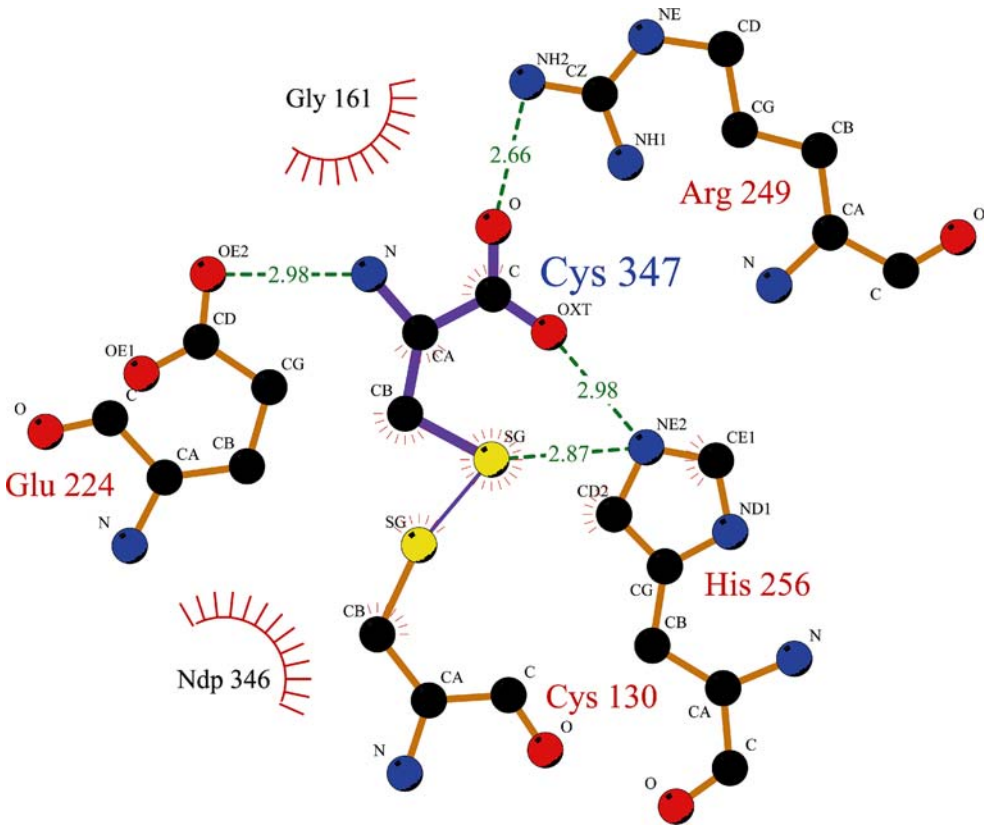


Fig. 9 Interaction of S-methyl L-cysteine sulphoxide (SMCS) with *mt*ASADH conserved catalytic residues, Cys130 and His256. Other non-ligand residues shown are involved in forming hydrophobic contacts. The figure was prepared using LIGPLOT [74]

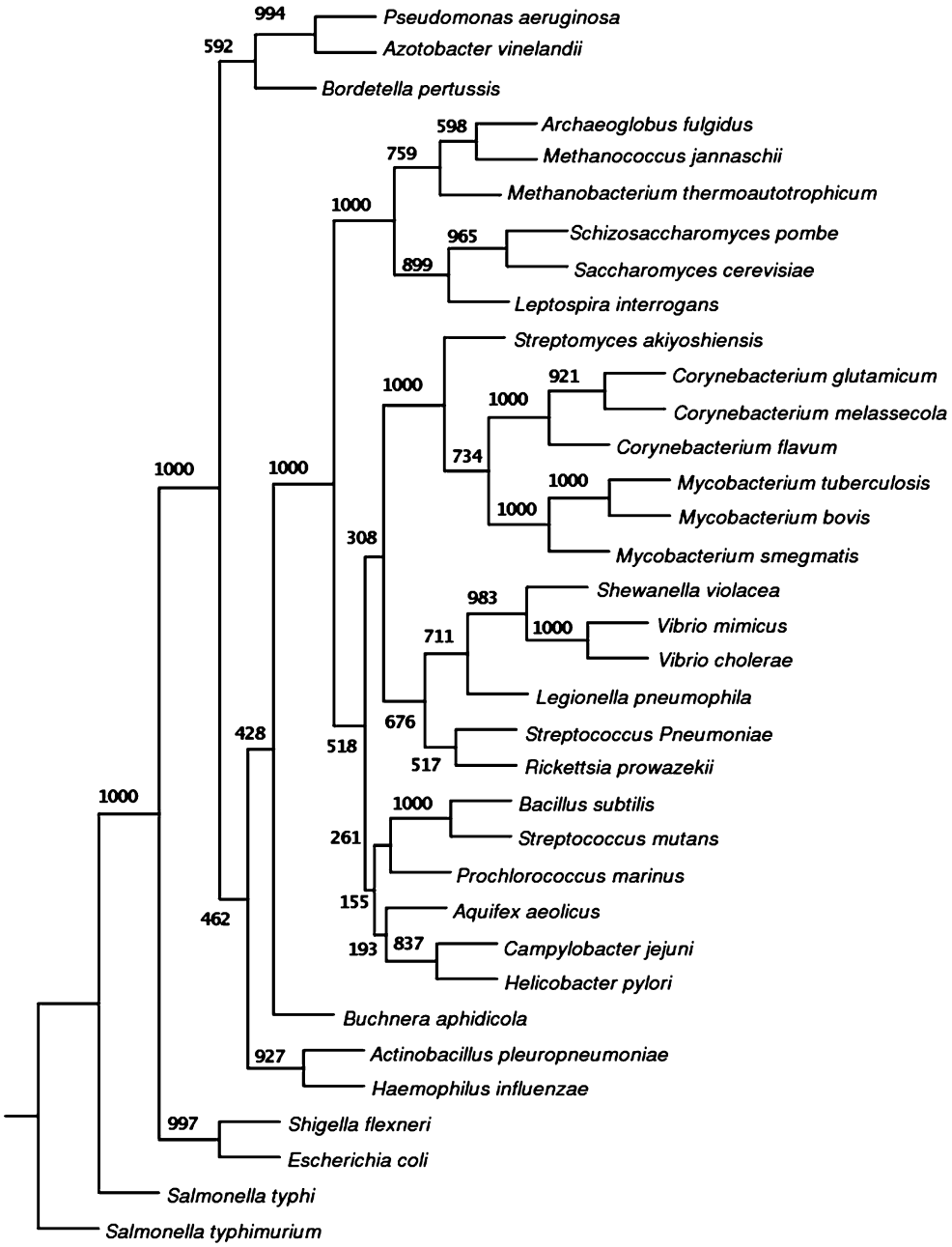


Fig. 10 Phylogenetic relationship of ASADH enzyme among bacterial species. This is a bootstrap consensus tree based on 1,000 replicates. The numbers on the nodes are bootstrap values

Vibrio cholera show relatively correct and consistent close proximity in their enzymology, although from the tree relationship the phylogenetic proximity between *Vibrio cholerae* and *Mycobacterium tuberculosis* can be inferred only with a low level of confidence.

Conclusions

The major obstacle in the application of structure-based drug design is the lack of structural data for the validated targets. This is where homology or comparative modelling comes to the rescue, by which it is easier to predict the structure of the target protein with the help of a structurally resolved template protein. The methodology to predict the 3D structure of a protein starting from its sequence has improved in accuracy and statistical robustness [68]. In fact, a protein sequence with over 30% identity to a known structure can often be predicted with an accuracy equivalent to a low-resolution X-ray structure [69]. In the absence of a crystal structure of aspartyl β-semialdehyde dehydrogenase in *Mycobacterium tuberculosis*, which could be used as a potential target against this microbe, an attempt has been made to model its 3D structure. A detailed study of the

stereochemistry using PROCHECK and structural conservation of motifs using STAMP and STRIDE software provided insights into the structural conservation of the folds in the 3D model of *mt*ASADH. An evaluation of the stereochemical quality of the model shows the reliability of the modelled protein. The presence of conserved catalytically active residues, Cys130 and His256, provides insight into the catalytic mechanism of this enzymes and indicates that it may be similar to that of *Escherichia coli* and *V. cholerae* ASADH. Thus, these catalytically active residues can be targeted to develop antagonists against the enzyme. The absence of any well defined secondary structure in the arm sub domain in *mt*ASADH can be attributed to the presence of residues that do not allow the formation of stable secondary structure. In the C-domain of the modelled *mt*ASADH, a non-conserved proline and three consecutive occurrences of alanine in α10, divide the molecule to form a coil and a helix rather than two helices as observed in the template structures. Molecular dynamics simulation analysis with modelled *mt*ASADH and the structurally resolved *Escherichia coli* and *M. jannaschii* enzymes provides a comparative view of the region that binds to NADP. Comparison of the kinetic behaviour of *mt*ASADH with the enzymes considered for template yielded a low binding affinity of the mtASADH enzyme for NADP, as observed in experimental outcomes [19, 21, 25, 65]. Phylogenetic analysis of all the sequences of ASADH from different species indicate that the enzyme has evolved in distinct groups from a common ancestor.

Docking with the known inhibitor SMCS illustrates the role played by the conserved active site residues in the *mt*ASADH. Furthermore, we have presented some probable ligand molecules, which have a similar structure to that of the known inhibitor, obtained from large scale screening of the drug molecule library using Flexx docking software.

Earlier efforts have been made to design inhibitors for ASASDH [70, 71]. The results of the present work, the modelled 3D-structure of ASADH from *Mycobacterium tuberculosis*, may play a useful role in the design and development of specific inhibitors, leading to the formation of suitable compounds that can kill *Mycobacterium tuberculosis*, the causative agent of tuberculosis in humans. The structural distinctiveness observed in the theoretical model can assist in resolving the exact three dimensional model of the SASDH from *Mycobacterium tuberculosis*.

Acknowledgements The authors would like to thank Prof. Alok Bhattacharya, for giving opportunity to work at CCBB, JNU and also like to thank Dr. Naidu Subba Rao for critical reading of the manuscript. Authors would like to thank Department of Biotechnology, Government of India for providing funds to Center of Computational Biology and Bioinformatics (CCBB). High performance computational facility at CCBB, JNU has been established under UGC-UPOE scheme. Anupama Singh would like to acknowledge the SRF received from CSIR.

References

1. Bloom BR, Murray CJL (1992) Science 257:1055–1064
2. Singh R, Kefala G, Janowski R (2004) EMBL-Hamburg Research Report 289–292
3. Grandoni JA, Marta PT, Schloss JV (1998) J Antimicrob Chemother 42:475–482
4. Moore RA, Bocik WE, Viola RE (2002) Prot Exp Purif 25:189–194
5. Cohen GN (1983) In: Herman KM, Somerville RL (eds), Addison–Wesley, Reading MA, pp 147–171
6. Viola RE (2001) Acc Chem Res 34:339–349
7. Patte JC (1983) In: Herman KM, Somerville RL (eds), Addison–Wesley, Reading MA, pp 213–228
8. Cirillo JD, Weisbrod TR, Pascopella L, Bloom BR, Jacobs WR Jr (1994) Mol Microbiol 11:629–639
9. Haziza C, Stragier P, Patte JC (1982) EMBO J 1:379–384
10. Galan JE, Nakayama K, Curtiss R (1990) Gene 94:29–35
11. Jagusztyn-Krynicka EK, Smorawinska M, Curtiss R 3rd (1982) J Gen Microbiol 128:1135–1145
12. Kalinowski J, Bachmann B, Thierbach G, Puhler A (1990) Mol Gen Genet 224:317–324
13. Thomas D, Surdin-Kerjan Y (1989) Mol Gen Genet 217:149–154
14. Baril C, Richaud C, Fournie E, Baranton G, Saint Girons I (1992) J Gen Microbiol 138:47–53
15. Cirillo JD, Weisbrod TR, Pascopella L, Bloom BR, Jacobs WR Jr (1994) Mol Microbiol 11:629–639
16. Cirillo JD, Barletta RG, Bloom BR, Jacobs WR Jr (1991) J Bacteriol 173:7772–7780
17. Hadfield A, Kryger G, Ouyang J, Petsko GA, Ringe D, Viola R (1999) J Mol Biol 289:991–1002
18. Blanco J, Moore RA, Kabaleeswaran V, Viola RE (2003) Protein Sci 12:27–33
19. Faehnle CR, Ohren JF, Viola RE (2005) J Mol Biol 353:1055–1068
20. Blanco J, Moore RA, Viola RE (2003) Proc Natl Acad Sci USA 100:12613–12617
21. Blanco J, Moore RA, Faehnle CR, Viola RE (2004) Acta Crystallogr D Biol Crystallogr 60:1808–1815
22. Blanco J, Moore RA, Faehnle CR, Coe DM, Viola RE (2004) Acta Crystallogr D Biol Crystallogr 60:1388–1395
23. Hadfield A, Shammas C, Kryger G, Ringe D, Petsko GA, Ouyang J, Viola RE (2001) Biochemistry 40:14475–14483
24. Biellmann JF, Eid P, Hirth C, Jornvall H (1980) Eur J Biochem 104:53–58
25. Shafiani S, Sharma P, Vohra RM, Tewari R (2005) J Appl Microbiol 98:832–838
26. Edgar Robert C (2004) Nucleic Acids Res 32:1792–1797
27. Edgar RC (2004) BMC Bioinformatics 5:113
28. Felsenstein J (1989) Cladistics 5:164–166
29. Cuff JA, Barton GJ (2000) Proteins 40:502–511
30. Adamczak R, Porollo A, Miller J (2004) Proteins 56:753–767
31. Frishman D, Argos P (1996) Protein Eng 9:133–142
32. Frishman D, Argos P (1997) Proteins 27:329–335
33. Frishman D, Argos P (1995) Proteins 23:566–579
34. Kabsch W, Sander C (1983) Biopolymers 22:2577–2637
35. McGuffin LJ, Bryson K, Jones DT (2000) Bioinformatics 16:404–405
36. Karplus K, Barrett C, Cline M, Diekhans M, Grate L, Hughey R (1999) Proteins Suppl 3:121–125
37. Shi J, Blundell TL, Mizuguchi K (2001) J Mol Biol 310:243–257
38. Fischer D (2000) Pac Symp Biocomput 110–130
39. Jones DT (1999) J Mol Biol 287:797–815
40. Rychlewski L, Jaroszewski L, Li W, Godzik A (2000) Protein Sci 9:232–241
41. Kelley LA, MacCallum RM, Sternberg MJ (2000) J Mol Biol 299:499–520

42. Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) J Mol Biol 247:536–540
43. Orengo CA, Michie AD, Jones S, Jones DT, Swindells MB, Thornton JM (1997) Structure 5:1093–1108
44. Russell RB, Barton GJ (1992) Proteins 14:309–323
45. Robert BR Version 4.2, User Guide [http://www.compbio.dundee.ac.uk/manuals/stamp.4.2/stamp.html]
46. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) Nucleic Acids Res 24:4876–4882
47. Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M (1995) Proteins 23:318–326
48. Laskowski RA, MacArthur MW, Smith DK, Jones DT, Hutchinson EG, Morris AL, Naylor D, Moss DS, Thorton JM (1993) J Appl Crystallogr 26:283–291
49. Canutescu AA, Shelenkov AA, Dunbrack RL (2003) Protein Sci 12:2001–2014
50. Sippl MJ (1993) Proteins 17:355–362
51. Humphrey W, Dalke A, Schulten K (1996) J Mol Graphics 14:33–38
52. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Proc Natl Acad Sci USA 98:10037–10041
53. Hayward S, Berendsen HJC (1998) Proteins 30:144–154
54. Ryckaert JP, Ciccotti G, Berendsen HJ (1977) J Comp Phys 23:327–341
55. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) J Comput Chem 26:1781–1802
56. Rarey M, Kramer B, Lengauer T, Klebe G (1996) J Mol Biol 261:470–489
57. Csizmadia F (2000) J Chem Inf Comput Sci 40:323–324
58. Bujnicki JM (2003) Curr Protein Pept Sci 4:327–337
59. Godzik A (2003) Methods Biochem Anal 4:525–546
60. Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer EL (2002) Nucleic Acids Res 30:276–280
61. Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Yamashita RA, Yin JJ, Zhang D, Bryant SH (2005) Nucleic Acids Res 33:D192–D196
62. Marchler-Bauer A, Anderson JB, DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko AR, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, Bryant SH (2003) Nucleic Acids Res 31:383–387
63. Tramontano A (1998) Methods 14:293–300
64. Holland MJ, Holland JP, Thill GP, Jackson KA (1981) J Biol Chem 256:1385–1395
65. Ouyang J, Viola RE (1995) Biochemistry 34:6394–6399
66. Nicholas CE, Dhaliwal B, Lockyer M, Hawkins AR, Stammers DK (2004) J Mol Biol 341:797–806
67. Karsten WE, Viola RE (1991) Biochem Biophys Acta 1077:209–219
68. Thornton JM (2001) Science 292:2095–2097
69. Xiang Z (2006) Curr Protein Pept Sci 7:217–227
70. Cox RJ, Gibson JS, Mayo Martin MB (2002) Chembiochem 3:874–886
71. Cox RJ, Hadfield AT, Mayo Martin MB (2001) Chem Commun 21:1710–1711
72. Barton GJ (1993a) Prot Eng 6:37–40
73. McDonald IK, Thornton JM (1994) J Mol Biol 238:777–793
74. Wallace AC, Laskowski RA, Thornton JM (1995) Protein Eng 8:127–134