5.4. Shikimate kinase: a potential target for development of novel antitubercular agents. Curr Drug Targets. 2007;8(3):456-68.

Shikimate Kinase: A Potential Target for Development of Novel Antitubercular Agents

José H. Pereira[†], Igor B. Vasconcelos[§], Jaim S. Oliveira[§], Rafael A. Caceres[¤], Walter F. de Azevedo Jr.[¤], Luis A. Basso[§*] and Diógenes S. Santos[§*]

[†]*Programa de Pós-Graduação em Biofisica Molecular-Departamento de Física, UNESP, São José do Rio Preto, SP, 15054-000;* [§]*Centro de Pesquisas em Biologia Molecular e Funcional, Faculdade de Farmácia, Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul. Avenida Ipiranga 6681, Tecnopuc, Partenon 90619-900, Porto Alegre, RS, Brazil and* [¤]*Faculdade de Biociências-PUCRS. Avenida Ipiranga 6681, 90619-900, Porto Alegre, RS, Brazil*

Abstract: Tuberculosis (TB) remains the leading cause of mortality due to a bacterial pathogen, *Mycobacterium tuberculosis*. However, no new classes of drugs for TB have been developed in the past 30 years. Therefore there is an urgent need to develop faster acting and effective new antitubercular agents, preferably belonging to new structural classes, to better combat TB, including MDR-TB, to shorten the duration of current treatment to improve patient compliance, and to provide effective treatment of latent tuberculosis infection. The enzymes in the shikimate pathway are potential targets for development of a new generation of antitubercular drugs. The shikimate pathway has been shown by disruption of *aroK* gene to be essential for the *Mycobacterium tuberculosis*. The shikimate kinase (SK) catalyses the phosphorylation of the 3-hydroxyl group of shikimic acid (shikimate) using ATP as a co-substrate. SK belongs to family of nucleoside monophosphate (NMP) kinases. The enzyme is an α/β protein consisting of a central sheet of five parallel β-strands flanked by α-helices. The shikimate kinases are composed of three domains: Core domain, Lid domain and Shikimate-binding domain. The Lid and Shikimate-binding domains are responsible for large conformational changes during catalysis. More recently, the precise interactions between SK and substrate have been elucidated, showing the binding of shikimate with three charged residues conserved among the SK sequences. The elucidation of interactions between *Mt*SK and their substrates is crucial for the development of a new generation of drugs against tuberculosis through rational drug design.

Key Words: tuberculosis, malaria, shikimate kinase, shikimate pathway, drug design, drug target, *Mycobacterium tuberculosis*, *Plasmodium* sp

1. THE SHIKIMATE PATHWAY

The shikimate pathway was discovered as a biosynthetic route through the studies of Bernhard Davis and David Sprinson and their collaborators [1,2]. The shikimate pathway links metabolism of carbohydrates to biosynthesis of aromatic compounds through seven metabolic steps, where phosphoenolpyruvate and erythrose 4-phosphate are converted to chorismic acid [3,4]. Chorismic acid is a common precursor for the synthesis of aromatic compounds, such as aromatic amino acids, folate, ubiquinone and menaquinones. Amongst them, the only ones that can be synthesized by humans are tyrosine, which is synthesized from phenylalanine through a reaction catalyzed by phenylalanine hidroxylase enzyme, and ubiquinone from tyrosine through a cascade of eight aromatic precursors [5].

The molecular organization of the shikimate pathway enzymes varies between taxonomic groups [6]. Bacteria have seven individual polypeptides, which are encoded by separate genes. Plants have a molecular arrangement similar to bacteria [7], with the exception of dehydroquinase (DHQase, third enzyme) and shikimate dehydrogenase (fourth enzyme) which have been shown to be present as separate domains on a bifunctional polypeptide [8]. In fungi and apicomplexan parasites (*Toxoplasma gondii*) the shikimate pathway has been shown to include monofunctional 3-deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase and chorismate synthase (CS) enzymes and a pentafunctional polypeptide termed AROM, which accounts for the remaining five shikimate pathway reactions [9].

The shikimate pathway enzymes are attractive targets for development of non-toxic antibacterial [10] and herbicides [11], because this pathway is essential for algae, higher plants, bacteria, fungi, whereas it is absent from mammals [12]. Thus, in the case of bacterial diseases, inhibition of any of shikimate pathway enzymes is unlikely to cause toxic side effects on the host. In addition, the importance of shikimate pathway can be indicated by the finding that deletion of the *aroA* gene, which codes EPSPS, causes *Streptomyces pneumoniae* and *Bordetella bronchiseptica* strains to be attenuated for virulence [13,14].

The shikimate pathway has also been discovered in apicomplexan parasites providing several targets for the development of new antiparasite drugs [15]. *In vitro* growth of apicomplexan parasites such as *Plasmodium falciparum* (malaria), *Toxoplasma gondii* (toxoplasmosis), and *Cryptosporidium parvum* (cryptosporidiosis) was inhibited by the herbicide glyphosate, a well-characterized inhibitor [16] of the shikimate pathway enzyme 5-enolpyruvyl shikimate 3-phosphate synthase (EPSPS), at concentration of 1-6 mM [15]. In *P. falciparum* and *T. gondii*, this inhibitory effect was reversed by co-addition o p-aminobenzoate (PABA) or folate suggesting that this pathway is essential for the biosynthesis of folate precursors. Moreover, several shikimate pathway enzymes has been detected in *T. gondii* and *P. falciparum* extracts [15,17] and a gene encoding a pentafunctional polypeptide AROM has been described in *T. goondii* indicating that a complete shikimate pathway is present in this parasite [18]. Whereas the shikimate pathway gives an array of compounds in bacteria, fungi and plants, only its folate biosynthesis function has been established is in apicomplexan parasites. All folate pathway enzymes, which convert GTP to deriva-

*Address correspondence to this author at the Centro de Pesquisas em Biologia Molecular e Funcional, Faculdade de Farmácia, Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul. Avenida Ipiranga 6681, Tecnopuc, Partenon 90619-900, Porto Alegre, RS, Brazil; Tel: +55-51-33203629; Fax, +55-51-3320-3629; E-mail: diogenes@pucrs.br

1389-4501/07 $50.00+.00 © 2007 Bentham Science Publishers Ltd.

Not for Distribution

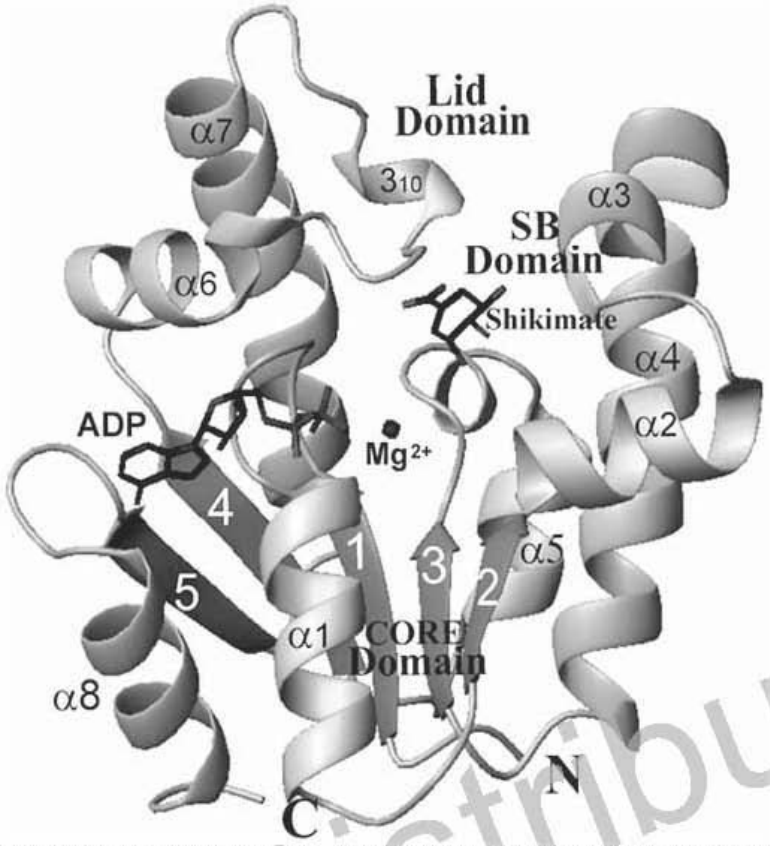


Fig. (1). Overall structure of Shikimate kinase from *M. tuberculosis* [61]. *Mt*SK displays an α/β-fold and the precise ordering of the strands 23145 in the parallel β-sheet classifies *Mt*SK as belonging to the family as the NMP kinases. The SKs are composed of three domains: CORE, LID and Shikimate-binding (SB) domains. The positions of the ADP, Mg^{2+} and shikimate are shown in the *Mt*SK structure.

tives of tetrahydrofolate, were found in Apicomplexa when the complete genome was sequenced [19].

Two shikimic acid analogs, 6-*S*-fluorshikimate and 6-*R*-fluorshikimate, have been shown to inhibit *P. falciparum* growth and inhibition shown to be specific to the shikimate pathway [20]. Despite the completion of *P.falciparum* genome sequence [21], only a single gene encoding the chorismate synthase enzyme and a potential bifunctional SK/EPSP synthase protein has been identified in the genome annotation [22]. The coding DNA sequence of *P. falciparum* chorismate synthase has been cloned and the protein has been shown to be located on cytosol by immunological studies [23]. Moreover, chorismate synthase, which catalyzes the last step of shikimate pathway, has been shown to be required for *Plasmodium falciparum* growth, as disruption of expression by RNA interference decreased parasite growth [19]. It has been proposed that the missing shikimate pathway enzymes are either substituted by non-homologous enzymes that catalyze the same reaction or that the enzymes are homologous but too divergent to be identified [22].

In mycobacteria, the chorismic acid intermediate is a precursor for the synthesis of naphthoquinones, menaquinones, and mycobactins, besides aromatic amino acids [24]. The salycilate-derived mycobactins siderophores have been shown to be essential for *M. tuberculosis* growth in macrophages [25]. Particularly, in *Mycobacterium tuberculosis*, the shikimate pathway has been shown to be essential for the bacterial viability. The disruption of *aroK* gene, which codes for the shikimate kinase enzyme (SK), was only possible when the second functional copy of *aroK* was integrated into the chromosome. Moreover, excision of the second integrated copy of *aroK* by the L5 excisionase could not be achieved in a *M. tuberculosis* strain carrying the disrupted copy of *aroK* gene, but was possible in a strain carrying a wild-type copy [26].

2. SHIKIMATE KINASE

The shikimate kinase (SK; EC 2.7.1.71), the fifth enzyme of the pathway, catalyses the regiospecific phosphorylation of the 3-hydroxyl group of shikimic acid (shikimate) using ATP as a co-substrate. In *Escherichia coli*, the SK reaction is catalyzed by two isoforms: SK I encoded by the *aroK* gene [27] and SK II encoded by the *aroL* gene [28]. The major difference between the isoenzymes is their K_m for shikimate, 20 mM for the SK I and 0.2 mM for the SK II enzyme [29]. The SK II isoform appears to play a dominant role in the shikimate pathway, its expression is controlled by the *tyr*R regulator, and it is repressed by tyrosine and tryptophan [30,31]. The physiological role of SK I in *E.coli* is not clear. Since mutations in SK I are associated with sensitivity to the antibiotic mecillinam [32] it has been suggested that SK I may have an alternative biological role that is distinct and unrelated to its shikimate kinase activity [29]. As pointed out by Parish and Stoker [26], if *M. tuberculosis aroK*-encoded SK I possess a similar activity it is possible that disruption of this activity can account for the observed inability of *M. tuberculosis* to grow in the absence of a functional copy of *aroK* gene. However, the actual nature of second activity of *aroK* gene product remains to be established. Contrary to the presence of isoenzymes in *E. coli*, complete genome sequences of a number of bacteria, for example, *Haemophilus influenzae* and *Mycobacterium tuberculosis*, have revealed the presence of only one SK-coding gene. Most of these SKs appear to be encoded by *aroK* rather than *aroL* because their amino acid sequences have higher degree of identity with *E.coli* SK I. The kinetic parameters for *aroK*-encoded *M.tuberculosis* SK (*Mt*SK) are more similar to those of *aroL*-encoded *E.coli* SK II than to those of *aroK*-encoded *E.coli* SK I. Thus, the *Mt*SK K_m value (0.41mM) for shikimate suggests that not all *aroK*-encoded SKs have high K_m values for shikimate [33].

Not For Distribution

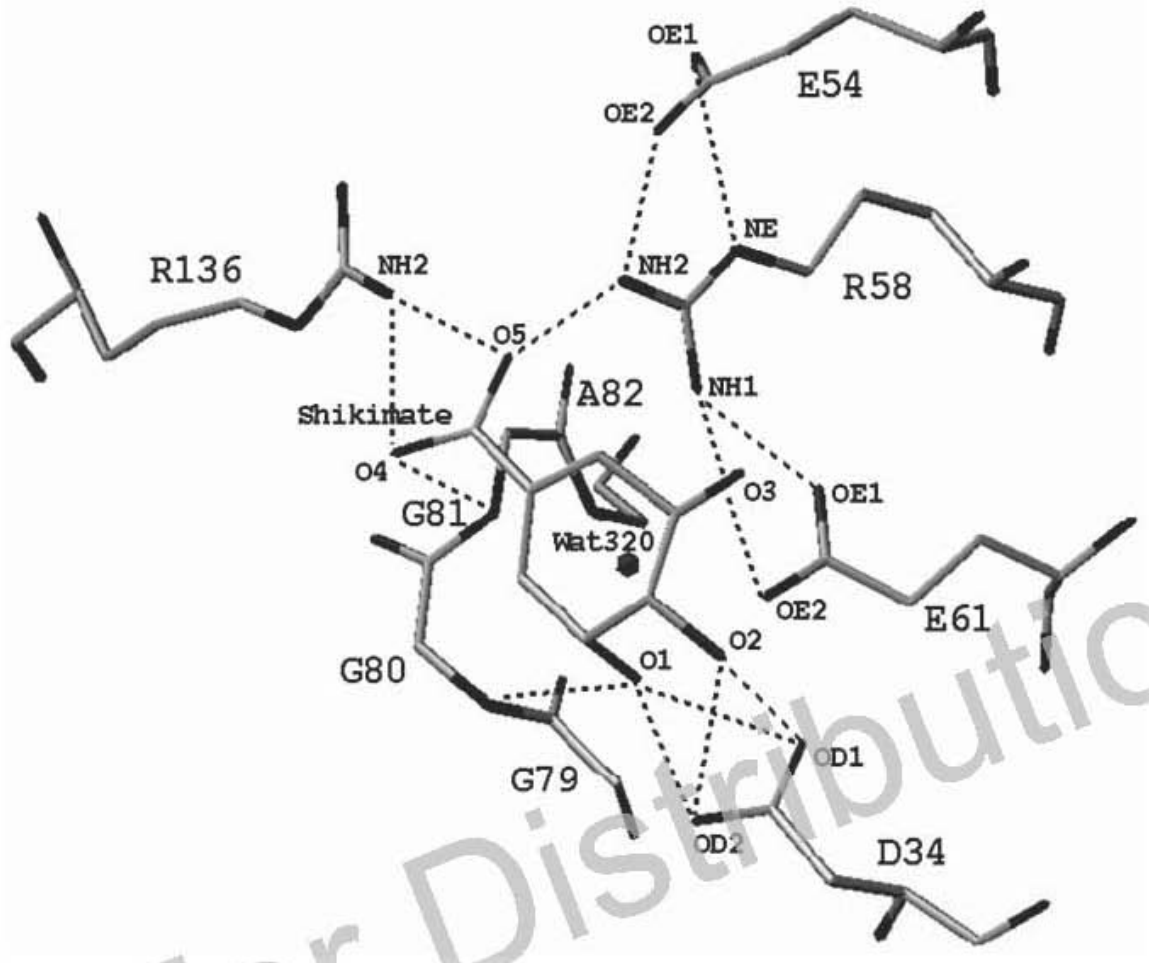


Fig. (2). Interactions involved in the binding of shikimate to the *Mt*SK active site (PDB access code: 1WE2) [61]. Hydrogen bonds are represented as broken lines. For clarity, the the water320-mediated hydrogens bonds between Gly79, Gly80, Gly81, Ala82, Arg58, and Glu61 protein residues and 3-hydroxyl group of shikimate are not shown.

Not For Distribution

2.1. The Enzyme Fold

Shikimate kinase displays an α/β-fold and consists of five central parallel β-sheet with the strand order 23145, flanked by α-helices [34] (Fig. (**1**)). The three crystal structures of SK from *Erwinia chrysanthemi* (*Erc*SK) [34,35] showed that SK belongs to the same structural family of nucleoside monophosphate (NMP) kinases for which structures are known for adenylate kinase (AK) [36,37], guanylate kinase [38], uridylate kinase [39] and thymidine kinase [40].

The NMP kinases are composed of three domains: CORE, LID and NMP-binding (NMPB) domains [41]. A characteristic feature of the NMP kinases is that they undergo large conformational changes during catalysis, for which AK is the most extensively studied [41]. There are two flexible regions of the structures that are responsible for movement: one is the NMP-binding site which is formed by a series helices between strands 2 and 3 of parallel β-sheet and the other is the LID domain, a region of varied size and structure following the fourth β-strand of the sheet [42,43]. In SK, the shikimate-binding (SB) domain corresponds to the NMPB domain of NMP kinases.

2.2. Functional Motifs

Three functional motifs of nucleotide-binding enzymes are recognizable in shikimate kinase, including a Walker A-motif (A-motif), a Walker B-motif (B-motif), and an adenine-binding loop. The Walker A-motif is located between the first β-strand (β1) and first α-helix (α1), containing the GXXXXGKT/S conserved sequence [44], where X represents any residue. This motif forms the phosphate-binding loop (P-loop), a giant anion hole which accommodates the β-phosphate of the ADP by donating hydrogen bonds from several backbone amides [45]. The side chain of P-loop lysine have a catalytic role of stabilizing the pentavalent transition state of the γ-phosphoryl group as has been shown for adenylate kinase [46] and $p21^{ras}$ [47].

In addition to the Walker A-motif, it is observed a second conserved sequence ZZDXXG called Walker B-motif [44], where Z represents a hydrophobic residue. More recently, however, sequence and structural comparisons for all P-loop-fold proteins classified Shikimate Kinase in the DxD group of enzymes [48], which has a conserved DxD motif in strand 2. The Walker B motif consensus in shikimate kinases is ZZZTGGG and the second glycine has been implicated in hydrogen bonding to the γ phosphate of ATP. This motif is located on the C-terminal segment of the third strand (β3) of the central β-sheet. The adenine-binding loop motif may be described as a sequence stretch of I/VDXXX(X)XP [33]. This motif forms a loop that wraps around the adenine moiety of ATP, connecting the β5-strand with the C-terminal α-helix.

3. TUBERCULOSIS AND *MYCOBACTERIUM TUBERCULOSIS* SHIKIMATE KINASE

Tuberculosis (TB) remains the leading cause of mortality due to a bacterial pathogen, *Mycobacterium tuberculosis*. The interruption

of centuries of decline in case rates of TB occurred, in most cases, in the late 1980s and involved the USA and some European countries due to increased poverty in urban settings and the immigration from TB high-burden countries [49]. Thus, no sustainable control of TB epidemics can be reached in any country without properly addressing the global epidemic. It is estimated that 8.2 million new TB cases occurred worldwide in the year 2000, with approximately 1.8 million deaths in the same year, and more than 95 % of those were in developing countries [50]. Approximately 2 billion individuals are believed to harbor latent TB based on tuberculin skin test surveys [51], which represents a considerable reservoir of bacilli. According to a recent report compiled by the World Health Organization (WHO), the total number of new cases of tuberculosis worldwide in 2002 had risen to approximately 9 million [52]. A key driver of the increase is the synergy with the HIV epidemic, which is having a devastating impact on some parts of the world mostly in the African Region, where 31% of new TB cases were attributable to HIV co-infection [50]. Another problem is the proliferation of multi-drug resistant (MDR) strains, defined as resistants to at least isoniazid and rifampicin, which are the most effective first-line drugs [53]. According to the 2004 Global TB Control Report of the World Health Organization, there are 300,000 new cases per year of MDR-TB worldwide, and 79 % of MDR-TB cases are now "super strains", resistant to at least three of the four main drugs used to treat TB [52]. The factors that most influence the emergence of drug-resistant strains include inappropriate treatment regimens, and patient noncompliance in completing the prescribed courses of therapy due to the lengthy standard "short-course" treatment or when the side effects become unbearable [54]. No new classes of drugs for TB have been developed in the past 30 years, reflecting the inherent difficulties in discovery and clinical testing of new agents and the lack of pharmaceutical industry in investing money and manpower for research in the area [55]. Hence, there is an urgent need to developing faster acting and effective new antitubercular agents, preferably belonging to new structural classes, to better combat TB, including MDR-TB, to shorten the duration of current treatment to improve patient compliance, and to provide effective treatment of latent tuberculosis infection [53].

In *M.tuberculosis*, the presence of the genes involved in the shikimate pathway began to be elucidated in the early 90s, with the cloning and characterization of *aro*A gene product, which codifies the EPSPS synthase [56]. Nevertheless, the complete genome sequence from *M.tuberculosis*, strain H37Rv reported by Cole *et al.* in 1998 [57] allowed the identification by sequence homology of all genes coding for shikimate pathway enzymes.

Four homologues to the shikimate pathway enzymes were located in a cluster containing the *aroD*-encoded type II DHQ dehydratase (Rv2537c), *aroB*-encoded DHQ synthase (Rv2538c), *aroK*-encoded type I shikimate kinase (Rv2539c), and *aroF*-encoded chorismate synthase (Rv2540c). The remaining homologues to shikimate pathway enzymes were annotated as follows: *aroG*-encoded class II phenylalanine-regulated DAHPS (Rv2178c), *aroE*-encoded shikimate dehydrogenase (Rv2552c), and *aroA*-encoded EPSP synthase (Rv3227).

The *aroK* structural gene is composed by 531 bp and codes a protein of 176 amino acids. The theoretical molecular mass of *Mt*SK enzyme subunit is 18.58 KDa. The first report of cloning and overexpression in soluble and functional form of *Mt*SK occurred in 2001 [58], where has been reported the PCR amplification *aroK* gene from genomic DNA of *M. tuberculosis* H37Rv strain, cloning in the plasmid pET-23a(+), and overexpression of *aroK*-encoded *Mt*SK protein in *Escherichia coli* BL21(DE3) host cells, without IPTG induction.

The crystal structure of a protein complexed with its substrates is of crucial importance for the rational design of inhibitors that target the enzyme. Although two crystal structures of SK from *M.tuberculosis* [33] have revealed the dynamic role of LID Domain in catalysis and the position of Mg^{2+} and ADP ligands in the structure, the precise positions and interactions between shikimate and *Mt*SK was not demonstrated because the shikimate-binding site was not occupied by the substrate or the electron density was not sufficient clear to position the shikimate molecule in the complex. In case of SK enzymes, a likely drawback of ATP-binding-site-based SK inhibitors would be their lack of specificity, owing to the common fold and similar ATP-binding site shared by many P-loop kinases [48]. Hence, trying to obtain a description of molecular interaction between *Mt*SK and shikimate, molecular-modeling and docking studies has been carried out and its results reported [59]. Despite these studies failed to predict all molecular interactions between *Mt*SK and shikimate [60], Asp34 and Arg136 residues and its hydrogen bonds implicated on shikimate binding are in agreement with the crystal structure of *Mt*SK-MgADP-Shikimate ternary complex reported afterwards [61]. The crystal structure of *Mt*SK-MgADP-Shikimate was the first crystallographic structure of SK with bound shikimate deposited in Protein Data Bank (PDB access code: 1WE2) [61]. Crystals were obtained by the hanging-drop vapour-diffusion method, and larger final concentration of ADP and shikimate in the drop (8.0mM) were used than those used by Gu *et al.* (2002) to obtain the crystals (4.0mM), in order to accurately determine the position of shikimate binding in the active site of *Mt*SK. The structure has been determined at 2.3 Å resolution, clearly reveling the amino-acids residues involved in shikimate binding. The molecular replacement method was used, using as a search model the structure of *Mt*SK-MgADP [33]. Almost at the same time of publishing of our article describing the structure of *Mt*SK-MgADP-Shikimate ternary complex, another article describing a similar structure of the ternary complex has been published (PDB access code: 1U8A) [60]. The crystal structures of *Mt*SK-MgADP-Shikimate [61] and *Mt*SK-ADP-Shikimate [60] ternary complexes have unequivocally revealed in detail the interactions of amino acid residues with bound shikimate and conformational changes upon substrate binding.

3.1. ADP/Mg^{2+} Interaction

Essentially all kinases require a Mg^{2+}-nucleotide complex as one of the enzyme substrates [62], an exception being the first partial reaction catalyzed by nucleoside diphosphate kinase, which can proceed independently of Mg^{2+} [63]. Nucleophilic attack on the γ-phosphate group of ATP will be most facilitated by meta-ion binding (Mg^{2+}) to the β- and γ-phosphoryl groups, whereas departure of the leaving group will be most favoured in a structure with metal binding to the α- and β-phosphoryl groups [64]. Presumably Mg^{2+} also assists in orienting the γ-phosphoryl group of 'inline' with respect to the second substrate, creating the correct geometry to complete phosphoryl transfer [62]. The binding of this cation in many P-loop proteins such as myosin [65], elongation factor EF-Tu [66], p21-Ras [67] and the heterotrimeric G-proteins [68] involves hexa-coordination of the Mg^{2+} by two oxygen atoms (from the β and γ-phosphates of the bound nucleotide), two water molecules, and two protein ligands. As in the *Mt*SK-MgADP structure (PDB code 1L4Y) [33], a typical six-coordination has been observed for Mg^{2+} in the *Mt*SK-MgADP-Shikimate structure, with some minor differences between the position of Mg^{2+} in the structures. In the *Mt*SK-MgADP-Shikimate structure Mg^{2+} interacts with a β-phosphate oxygen of ADP, Ser16 OG of the Walker A motif and four water molecules [61]. Superimposition of shikimate-free *Mt*SK-MgADP structure on shikimate-complexed *Mt*SK-MgADP structure showed that in the ternary complex water1 and water4 are in equivalent positions but that shifts of 1.32 and 2.75 Å are observed for water2 and water3, respectively. In the *Mt*SK-MgADP structure, the interaction between Asp32 and Ser16 of the Walker A motif is *via* a bridging water molecule (water6), whereas in the ternary complex structure this interaction occurs directly *via* a hydrogen bond between the two residues [61], which accounts for the exclusion of water6 from the magnesium-binding site in the ternary

Non-commercial Distribution

structure. In the ternary complex, the water1 molecule coordinated to the Mg^{2+} interacts directly with the chloride ion instead of interacting with Asp34 *via* a bridging water molecule (water5) as observed in the *Mt*SK-MgADP structure [33]. This chloride ion also interacts with the 3-hydroxyl group of shikimate and the backbone amide of Gly80 [61]. Thus, the different mode of interaction observed for residue Asp34 arises from the presence of shikimate, which leads to the exclusion of water5 from *Mt*SK active site [61]. The Mg^{2+} cation was not included in the final structure of *Mt*SK-ADP-Shikimate ternary complex reported by Dhaliwal et. al.(2004), since the best diffracting crystals grew in the absence of $MgCl_2$ [60]. Accordingly, the chloride ion observed in the active sites of both *Mt*SK binary complex [33] and *Mt*SK-MgADP-Shikimate ternary complex [61is absent in the Mg^{2+}-free structure of *Mt*SK-ADP-Shikimate ternary complex [60]. Thus, the absence of $MgCl_2$ in crystallization mixture probably have accounted for the small differences observed in conformation, position and molecular interactions of shikimate when the structure reported by Dhaliwal *et al.* is compared with *Mt*SK-MgADP-Shikimate-Cl structure [61]. In Mg^{2+}-free *Mt*SK structure the 3-hydroxyl group of shikimate is closer to β-phosphoryl group of ADP than it is in the *Mt*SK-MgADP-Shikimate structure. Furthermore, two additional water-mediated interactions between protein residues and shikimate have been shown. The NZ atom of Lys15 residue forms a 2.6 Å hydrogen-bond with oxygen atom of another water molecule, which in turn interacts with the 3-hydroxyl group of shikimate, and the main-chain nitrogen of Leu119 residue forms a 2.9 Å hydrogen bond with the oxygen atom of a water molecule, which in turn interacts with 5-hydroxyl group of substrate [60]. In the *Mt*SK-MgADP-Shikimate structure, NZ atom of Lys15 cannot form a water-mediated interaction with 3-hydroxyl group of shikimate, since the chloride ion is bound to the active site cavity between them. The distance between NZ atom of Lys15 and the chloride ion is 3.94 Å, which in turn forms a 3.36-Å hydrogen bond with 3-hydroxyl group of shikimate [61].

In the *Mt*SK-MgADP structures [33], Lys15 forms a hydrogen bond with a β-phosphate O1B atom and the chloride ion. The main-chain NH of Gly80 is hydrogen bonded to the chloride ion the binary (1L4Y) [33] and ternary complex structures [60,61]. These residues are located in vicinity of where the chemical reaction occur a may thus play a critical role in the transition state stabilization.

The molecular interactions that describe the ADP binding mode on enzyme are very similar in all available structures of *Mt*SK. The adenine moiety of ADP is sandwiched between Arg110 and Pro155 [33] and this interaction has also been observed in *Erc*SK [34], and in Adenilate kinase [69] and isoenzyme II [70]. Arg110 is located at the C-terminus of α6 where the LID domain starts. In *Mt*SK, Arg110 and Arg117 residues represent, respectively, the first and the last residue of a conserved motif of LID domain observed for P-loop kinases (typically RXX(X)R) [48]. In P-loop shikimate kinases the conserved motif of the LID domain has been proposed to be $R(X)_{6-9}R$ [61]. The Arg117 has been shown to interact with the α– and β-phosphate groups of ADP [33,60,61], it generally interacts with the γ-phosphate of ATP bound to enzymes, and it may stabilize the transition state by neutralizing the developing negative charge on the β-γ bridge O atom [71]. The Pro155 is the last residue of the adenine-binding loop motif (residues 148-155 in *Mt*SK), which was first recognized in Adenilate kinase and *Erc*SK [34] and has been described as an I/VD*XXX*(*X*)*X*P sequence stretch [33]. However, the second (aspartate) residue and the last (proline) residue are not conserved in the adenine-binding loops of *aroK*-encoded SKs from *Escherichia coli* [72] and *Campylobacter jejuni* (PDB acess code: 1VIA). However, they are recognizable from structural alignment analysis using *Mt*SK-MgADP as a reference structure [61]. In fact, it has been pointed out that the main-chain contacts with the adenine base and the presence of a structural motif independent of sequence may be the most important features for adenine binding in kinases [62]. The adenine-binding loop motif (residues 148-155 in *Mt*SK) forms a loop that wraps around the adenine moiety of ATP, connecting the β5-strand with the C-terminal α8-helix.

For catalysis, the three protein interacting residues Lys15, Arg117, and Arg136 have been proposed to be the most important [33]. Consistent with this proposal, the K15M mutant of *Erc*SK showed no detectable enzyme activity [35], although it was in fact a K15M/P115L double mutant. Contrary to the previously proposed, analysis of the *Mt*SK ternary complexes have revealed that Lys15 and Arg117 are the only positively charged residues located in the vicinity of where the reaction may occur and may therefore play critical roles in the stabilization of the transition state [33]. The Arg136 residue, instead, appears to interact with the carboxyl group of shikimate and probably, it is not involved in catalysis [60,61].

3.2. Interaction with Shikimate

The shikimate-binding domain, which follows strand β2, consists of helices α2 and α3 and the N-terminal region of helix α4 (Fig. (**1**)). In particular for *Mt*SK, the precise interactions between shikimate and SK have been elucidated, showing that the guanidinium groups of Arg58 and Arg136 and the NH backbone group of Gly81 interact with the carboxyl group of shikimate. The 3-hydroxyl group of shikimate forms hydrogen bonds with the carboxyl group of Asp34 and the main-chain NH group of Gly80 and a water molecule. The 2-hydroxyl group of shikimate hydrogen bonds to the side chain of Asp34 (Fig. (**2**)). The Glu61 is conserved in both *aro*K and *aro*L-encoded shikimate kinase enzymes. This residue is not directly involved in substrate binding, but it forms a hydrogen bond and a salt bridge with the conserved Arg58 and assists in positioning the guanidinium group of Arg58 for shikimate binding. In the ternary structure, Glu61 side chain also forms a water-mediated interaction with the 3-hydroxyl group of shikimate. Therefore, Glu61 plays an important role in the substrate-binding site [61]. The Glu54 carboxylate group also appears to anchor the guanidinium group of Arg58 for interaction with the shikimate carboxylate group; however, Glu54 is not conserved, except for *aro*K-encoded shikimate kinases [61]. In the *Mt*SK-ADP-shikimate structure [60], the main-chain nitrogen of Leu119 forms a 2.9 A H-bond with the water3 oxygen atom, which in turns interacts with the 5-hydroxyl group of shikimate, and the NZ atom of Lys15 forms a 2.6 A H-bond with the oxygen atom of water2, which also interacts with the 3-hydroxyl group of the substrate.

3.3. Conformational Changes Upon Substrate Binding

As pointed out above, NMP Kinases undergoes large conformational changes during catalysis, because their LID domain and the NMP binding-site are very flexible and can make large movements upon substrate binding. These structural changes act to position enzyme sidechains appropriately around the substrates and to sequester the substrates so as to prevent the hydrolysis of bound ATP or other phosphoryl-containing substrates prior to catalysis [41,62,64]. In addition, it has also been shown that these two domains are capable to make independently moves towards each other [41]. Previous studies have shown that hexokinases [73] and adenylate kinases [42,74], which are classified as NMP kinases, undergo a large conformational change during catalysis.

Several structural and spectroscopic studies have demonstrated that SKs undergo conformational changes on ligand binding. Circular dichroism spectra of unliganded and liganded *Erc*SK enzyme in the presence of 2mM shikimate or 2mM γN-ATP (an non-hydrolisable ATP analogue) have shown that SK undergoes conformational changes upon ligand binding [34]. Moreover, fluorescence studies were performed using a single tryptophan residue (W54) as a report group, which is positioned close to the shikimate binding site. The addition of shikimate to protein solution caused a quenching in *Erc*SK protein fluorescence and a blue shift of 3 nm

Non-commercial Distribution

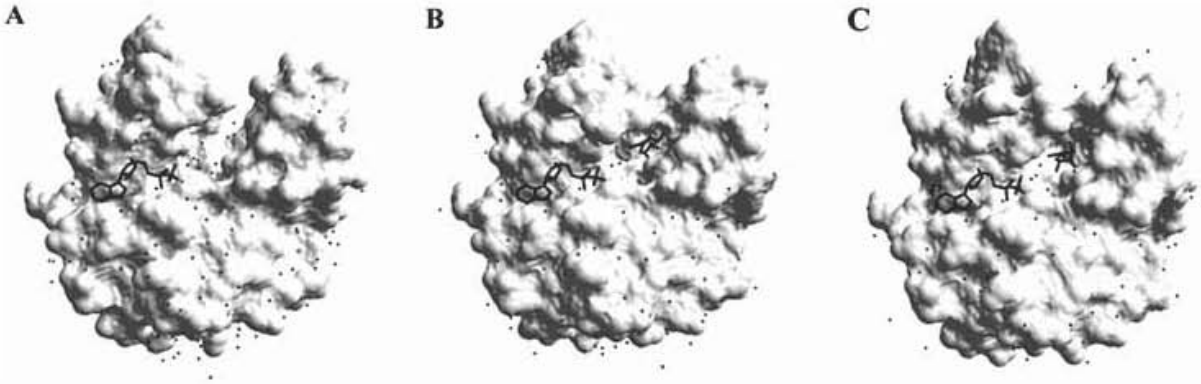


Fig. (3). Molecular surface of (A) *Mt*SK-MgADP (1L4U) [33], (B) *Mt*SK-MgADP-Shikimate (1WE2) [61], and (C) *Mt*SK-ADP-Shikimate (1U8A) [60] structures The molecular surface areas have been calculated using the program Swiss-PDBViewer v3.7 (www.expasy.org/spdbv), probe radius of 1.4 Å, and a fixed radius for all atoms. Calculated values are approximately 7246 $Å^2$ for *Mt*SK complexed with only MgADP (A), 6915 $Å^2$ for structure complexed with MgADP and Shikimate (B), and 7040 $Å^2$ for the Mg^{2+}-free structure complexed with ADP and Shikimate. Thus, approximately 330 $Å^2$ and 206 $Å^2$ of solvent-acessible surface are buried on shikimate binding to, respectively, *Mt*SK-MgADP-Shikimate and *Mt*SK-ADP-Shikimate ternary complexes. The shikimate binding leads to a counter movement of LID and SB domains and, consequently, to a partial closure of shikimate-binding pocket. Probably, the Mg^{2+} cation has a role in the closure of *Mt*SK active site. All atoms of shikimate and ADP, Mg^{+2} cation, and oxygen atoms of waters are colored grey.

in the emission maximum, consistent with the loop containing the tryptophan residue becoming more deeply buried within the protein following ligand-binding [75]. In addition, the averaged *B*-factor for all residues in crystal structure of *Erc*SK showed clear evidence of the flexibility of the molecule, where the temperature factors for both *Erc*SK molecules in the asymmetric unit (one with bound shikimate and other with unbound shikimate) indicate two regions of high mobility, corresponding to the shikimate binding-site and the LID domain and its flanking regions [34]. A comparison of the residue-averaged *B* factors between the crystal structures of *Erc*SK-MgADP, *Mt*SK-MgADP (1L4Y), and *Mt*SK-MgADP Pt-derivative (1L4U) binary complexes [33] shows that the enzyme from *M. tuberculosis* follows the same pattern of flexibility in the LID and SB domains as previously observed for *Erc*SK.

Another method used to evaluate conformational changes in proteins is the superposition of structures in order to compare different complexes of the same protein. To demonstrate conformational changes upon shikimate binding in *Mt*SK, alignment of C^{α} positions of the *Mt*SK-MgADP-Shikimate dead-end ternary complex and the *Mt*SK-MgADP binary complex structures were made [61], showing that the LID and SB domains undergoes noticeable concerted movements towards each other. In this alignment, which included all residues, were verified r.m.s. deviation values of 0.56 and 0.54 Å for 1L4U and 1L4Y, respectively. Residues 112-124, that comprises the LID domain, showed a r.m.s. deviation of 1.33 Å, and the SB doamin shift was somewhat smaller, with an r.m.s. deviation value of 0.74 Å for residues 33-61. Similar values were found when the same procedure was applied to the binary complex and the *Mt*SK-ADP-Shikimate structure [60], with a value of 0.7 Å for the overall structure (excluding residues 114-115), 1.5 Å for the LID domain and 0.4 Å for the SB domain. The residues directly involved in these movements are Val116, Pro118 and Leu119 (LID domain), Ile45, Ala46, Glu54, Phe57 and Arg58 (SB domain), where its side chains shifted upon shikimate binding [61]. A comparison of *Mt*SK-ADP-Shikimate ternary complex to *Mt*SK-MgADP binary complex showed, within of SB region, a shift of 0.9 Å of Arg58 residue towards the carboxylate group of shikimate. Phe49 residue moves approximately ~1.7 Å away from Phe57 and closer to the substrate, translating towards shikimate [60]. In addition, it has been proposed by Dhaliwal *et al.* (2004) that the observed changes in the orientation and position of Phe49 and Phe57 residues disrupt their strong ring stacking interactions and probably result from the Van Der Waals contacts made between shikimate with both phenylalanines [60]. However, in *Mt*SK-MgADP-Shikimate the strong ring stacking interaction between the phenylalanines is not disrupted in spite of the shift of Phe49 and Phe57 residues towards shikimate of, 0.49 Å and 0.79 Å, respectively. Moreover, there is a reduction in molecular surface area value the ternary complex compared to binary complex. The calculated value for *Mt*SK complexed with MgADP (1L4U) was approximately 7246 $Å^2$ and for *Mt*SK complexed with MgADP and shikimate a value of 6915 $Å^2$ was found. Based on these results, the difference between the complexes molecular surface areas is 330 $Å^2$, which are buried on shikimate binding (Fig. (**3**)) [61]. A reduction in molecular surface area value has also been observed in the Mg^{2+}-free crystal structure of *Mt*SK-ADP-Shikimate ternary complex (Fig. (**3**)). Nevertheless, in this complex a molecular surface of only 206 $Å^2$ is buried on shikimate binding, indicating that Mg^{2+} binding has a role in the closure of *Mt*SK active site. It is important to point out that both the *Mt*SK-MgADP bynary complexes [33] and *Mt*SK-MgADP-Shikimate [61] and *Mt*SK-ADP-Shikimate [60] ternary complexes have been crystallized in the same space group with similar unit-cell parameters and residues from LID and SB domains form no crystal contacts with symmetry-related *Mt*SK molecules. Thus, the conformational changes that have been described for *Mt*SK cannot merely be a reflection of the different crystal-packing arrangements.

The incomplete LID-domain closure observed in crystal structures of both *Mt*SK-MgADP-Shikimate [61] and *Mt*SK-ADP-Shikimate [60] ternary complexes suggests that γ-phosphate of ATP is necessary for the completion of the domain movement [34]. This find is not unexpected, since total active-site closure upon dead-end ternary complex formation would result in locking the enzyme active site in an inactive form in which shikimate substrate binding to *Mt*SK enzyme prior to MgADP dissociation from its active site would result in an inactive abortive complex. Several previously reported results are consistent with these structural data. Measurements of *Erc*SK intrinsic tryptophan fluorescence (Trp54) on shikimate binding to either *Erc*SK or the *Erc*SK-MgADP binary complex showed a modest synergism of binding between these substrates, since the dissociation constant value for shikimate (K_d = 0.72 mM) decreased to 0.3 mM in the presence of 1.5 mM ADP [75]. In addition, measurements of the quenching of protein fluorescence of the *aroL*-encoded *Erc*SK upon nucleotide binding demonstrated the dissociation-constant values for ADP and ATP to be 1.7 and 2.6 mM, respectively [35]. The *Km* for ATP (620 μM) was found to be approximately four times lower than the dissociation constant in the absence of shikimate [35]. These results prompted the proposal that the conformational changes in the *Erc*SK enzyme associated with the binding of the first substrate led to an increase

Not for Distribution

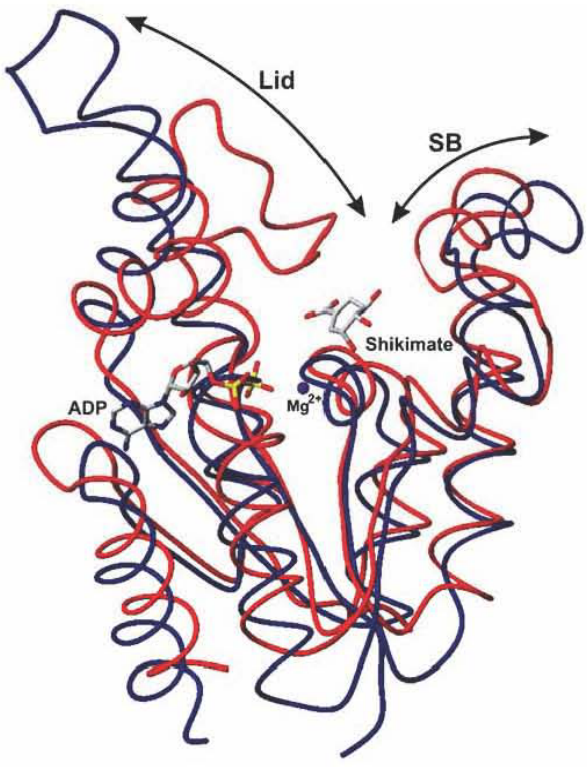


Fig. (4). Superposition of SK-MgADP-Shikimate ternary complex from *Mycobacterium tuberculosis* [61] (red line) and apo-SK (K15M/P115L) from *Erwinia chrysanthemi* [34] (blue line).

in the affinity for the second substrate. However, this synergism on substrate binding does not appear to hold for *Mt*SK since the apparent dissociation-constant values for ATP (89μM) and shikimate (440 μM) are similar to their K_m values, 83 μM for ATP and 410 μM for shikimate considering either a rapid equilibrium random-order bi-bi enzyme mechanism or a steady-state-ordered bi-bi enzyme mechanism [33]. Moreover, no evidence for synergism between shikimate and ATP could be observed in substrate binding to *Erc*SK in a chloride buffer system [76].

The K15M *Erc*SK mutant has been crystallized in an open conformation that is proposed to presumably be equivalent to an apo-enzyme structure in which neither ADP (or ATP) nor shikimate would be bound [35]. This mutant was produced to evaluate the role of the conserved Lys15 of the Walker A motif. However, an unwanted mutation point mutation in the LID domain (Pro115Leu) was detected during the refinement of the model. Furthermore, the enzyme was crystallized with two independent molecules in the asymmetric unit and extensive contacts of neighboring LID domains lead to a stabilization of this part of the molecule that is not visible in the native crystal structure [34,35]. It therefore appears unwarranted to consider the double K15M/P115L *Erc*SK mutant a model for the apo enzyme. Since a better model for the apo-enzyme is not currently available, here we considered the double K15M/P115L *Erc*SK mutant as a model for the apo-SK. The superposition of *Mt*SK-MgADP-Shikimate and apo *Erc*SK (K15M+P115L) show the large conformational changes in the LID and SB domains associated with the binding of ADP and shikimate on *Mt*SK (Fig. (**4**)).

4. MOLECULAR DYNAMICS SIMULATIONS OF THE SHIKIMATE KINASE STRUCTURE

Molecular dynamics (MD) simulations are playing an important role in a wide area of research, which includes structural biology, condensed matter physics, material science, and bioinformatics. The use of molecular dynamics simulations allow monitoring the dynamics of individual atoms, thereby giving a unique insight into the molecule behavior that cannot be easily extracted from laboratory experiments. The application of molecular dynamics simulation allows prediction of structural flexibility features of complexes of enzyme and ligands. We performed molecular dynamics simulations of the complex MtSK-ADP-shikimate (system 1) and apoenzyme MtSK (system 2) in order to elucidate the influence of the shikimate on the overall structure of the MtSK. The root-mean square deviation (RMSD) of the positions for all backbone Cα atoms from their initial configuration as a function of simulation time for both systems are shown in Fig. (5). Analysis of the molecular dynamics simulations indicates the LID domain undergoes confor-

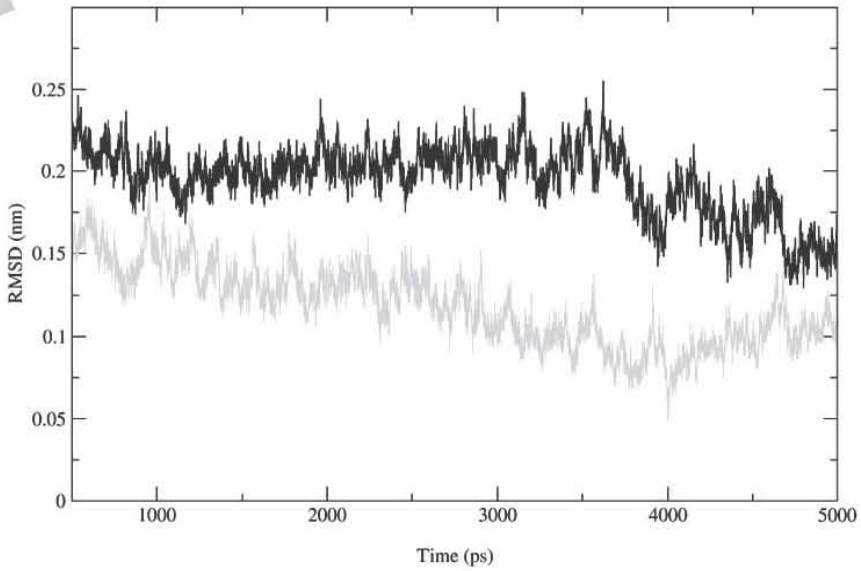


Fig. (5). Time-dependent Cα RMSD of both complexes with respect to the initial structures for MtSK-ADP-shikimate (system 1)(grey line) (a), MtSK, apoenzyme (system 2) (black line) (b). Molecular dynamics simulations were performed using the program GROMACS [97]. The starting geometries for the simulations of the systems 1 were generated from the X-ray structure obtained from the Protein Data Bank (PDB access code: 1WE2). For the simulations of the system 2 the atomic coordinates for the shikimate, ADP and ions were deleted from the PDB file of system 1.

Not for Distribution

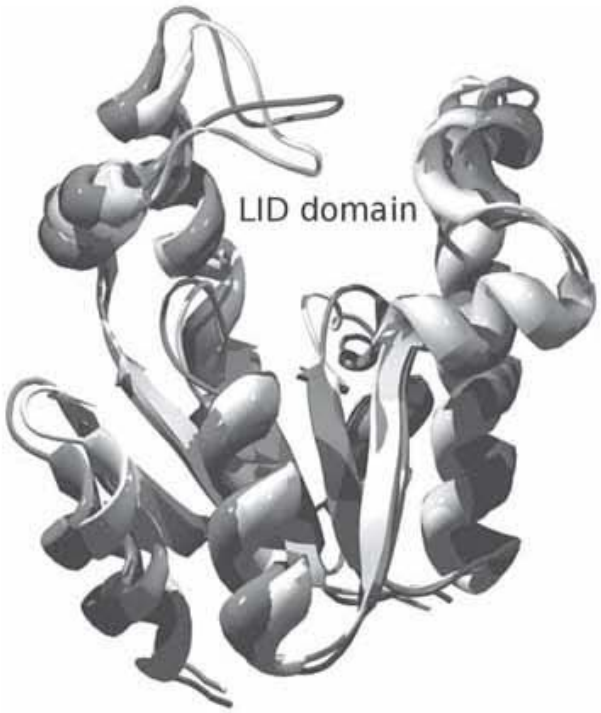


Fig. (6). Superposition of the atomic coordinates of the initial (light grey) and final (grey) structures for system 2. The atomic coordinates for the final structure were those at the end of 5ns molecular dynamics simulation.

mation change in the system 2 (without shikimate). The structure without the molecule of MtSK presents high flexibility for the residues involved in the LID domain in the system 2. Analysis of the residues participating in the LID domain in the system 1 indicates the presence of a hydrogen bond network involving residues 112-124 and the shikimate, which promotes the stabilization of the LID domain in the closed conformation. Molecular dynamics simulations of the system 2 for 5ns show the movement of the LID domain in the MtSK structure. Superposition of the atomic coordinates of the initial and final structures for system 2 after 5ns of molecular dynamics simulation is shown in Fig. (**6**). The LID domain is moved towards the solvent in the structure at the final of the 5ns molecular dynamics simulation, demonstrating that the molecular dynamics simulation results are in agreement with experimental X-ray crystallographic data (Fig. (**4**)), confirming that the binding of the shikimate to the SK structure is enough to promote the closure of the LID domain. On the other hand, the lacking of the shikimate in the MtSK promotes the opening of the LID domain. This structural feature should be considered in any structure-based design of SK inhibitors, which are competitive against shikimate. A detailed analysis of the molecular dynamics simulations for MtSK with different ligand is underway and will be published elsewhere.

5. CONCLUDING REMARKS

The disruption of gene *arok*, which codes for the shikimate kinase, has been shown to be essential for the viability of *M. tuberculosis*. The evidence that shikimate pathway is essential for *M. tuberculosis* even in the presence of exogenous supplements as p-aminobenzoate, p-hydroxibenzoate and aromatic amino acids reinforces its attractiveness as a drug target [26]. It is interesting to note that shikimic acid can be used to supply the aromatic amino acid and aromatic vitamin requirements of *E. coli* blocked in any of the first three steps of the shikimate pathway [77]. This ability has been related to the presence of the shikimic acid transport system encoded by the gene *shiA* [78]. Thus, there is the possibility that shikimic acid could be used as a supplement to allow the growth *in vitro* of mutant *M. tuberculosis* strains disrupted on *aroG*, *aroB*, *aroD* or *aroE* genes, which code for first four shikimate pathway enzymes in *M. tuberculosis*. These *aro*-disrupted *M. tuberculosis* strains could be used in experiments to determine if shikimate pathway is essential to *in vivo growth* of *M. tuberculosis*.

A comprehensive structural picture of the interactions between *M. tuberculosis* SK enzyme and Mg^{2+}, ADP and shikimate substrates have been obtained from crystallographic studies [33,60,61]. In addition, the LID and SB domains conformational changes upon ADP and shikimate substrates binding, which result in a partial closure of and solvent expulsion from *Mt*SK active site, has been described.

There has been considerable debate within the literature as whether enzyme-catalyzed phosphoryl transfer reactions operate primarily with an associative (S_N2-like) or dissociative (S_N1-like) transition state [62]. A criterion that has been used for distinguishing between these mechanisms is based on the geometry and reaction coordinate distances between the terminal phosphoryl group of ATP and the acceptor substrate in ground state enzyme-substrate or enzyme inhibitor complexes [79]. For the transition state to have some degree of associative character, these distances are expected to be less than the sum of a P-O van der Waals contacts and a P-O single bond (i.e in the order of approximately 4.9 Å) [62]. In dissociative mechanisms these distances would be expected to be longer than this to allow space for the intermediate monomeric metaphosphate to exist between the leaving group (ADP) and the entering group [79]. A crystal structure of *Mt*SK in complex with MgADP, Shikimate, and AlF_3 (a structural analog of γ-phosphoryl group transfered in the reaction) could permit a snapshot of the transition-state like structure of *Mt*SK. As it has been shown to other kinases [80-93], this structural study could distinguish between associative and dissociative transition states in the reaction catalyzed by *Mt*SK.

It has been recently described the structural alterations caused by binding of the shikimate, magnesium and chloride ions in the SK from *Mycobacterium tuberculosis* (MtSK). These new findings indicate that both ions and the shikimate influence the structural conformation of MtSK. The magnesium ion also appears to influence both the hydroxyl groups of shikimate molecule and some residues from active site. The chloride ion can influence the affinity of shikimate for MtSK, in the binding position of ADP in MtSK active site and also in the opening length of LID domain of the MtSK structure. The shikimate binding causes a closing of LID domain, which appears to influence drastically the crystallographic packing of the protein [94-95].

Structures of SKs deposited in the protein data bank (PDB) were superimposed and positions of residues involved in binding (Lys15, Asp34, Arg58, Glu61, Gly79, Gly80, Gly81, Arg136) between shikimate and SK are highly conserved [61]. The precise interactions between *Mt*SK and shikimate have been elucidated, showing the residues involved in the binding of substrate and the conformational changes upon substrate binding. Accordingly, the availability of the *M. tuberculosis* shikimate kinase structures complexed with shikimate provide crucial information for the design of non-promiscuous SK inhibitors that target both the shikimate- and ATP-binding pockets or uniquely, the shikimate-binding site. Moreover, the knowledge of functional factors that lead to active site closure could be used for designing inhibitors that force *Mt*SK to a closed conformation that would be unable to catalyze the phosphoryl transfer to shikimate. These inhibitors could block the biosynthesis of aromatic aminoacids [96] and other compounds (folate, micobactins, etc), which are essential for the growth and viability of the microorganism.

ACKNOWLEDGEMENTS

Financial support for this work was provided by FAPESP (Proc. 02/05347-4, 01/07532-0, 04/00217-0) and Millennium Initiative

Nao Comercializavel

Program MCT-CNPq, Ministry of Health - Secretary of Health Policy (Brazil) to DSS and LAB. DSS (304051/1975-06) WFA (CNPq, 300851/98-7), and LAB (520182/99-5) are research career awarded from the National Research Council of Brazil (CNPq).

ABBREVIATIONS

TB	=	Tuberculosis
Mt	=	*Mycobacterium tuberculosis*
SK	=	Shikimate kinase
DAHPS	=	3-deoxy-D-arabino-heptulosonate-7-phosphate synthase
CS	=	Chorismate synthase
EPSPS	=	5-enolpyruvylshikimate-3-phosphate synthase
PABA	=	p-aminobenzoate
GTP	=	Guanosine triphosphate
ATP	=	Adenosine triphosphate
ADP	=	Adenosine diphosphate
Km	=	Michaellis-Menten constant
NMP	=	Nucleoside monophosphate
AK	=	Adenylate kinase
WHO	=	World Health Organization
HIV	=	Human immunodeficiency virus
MDR	=	Multidrug-resistant
DHQ	=	Dehydroquinase
PCR	=	Polymerase chain reaction
IPTG	=	Isopropyl-beta-D-thiogalactoside
Mg	=	Magnesium
r.m.s.	=	Root mean square
Erc	=	*Erwinia chrysantemi*
AMP-PCP	=	Adenosine -β,γ-methyleneadenosine 5'-triphosphate
AMP-PCP	=	β,γ-imidoadenosine 5'-triphosphate

REFERENCES

[1] Davis, B.D. and Mingioli, E. S. (1953) *J. Bacteriol.* **66**(2), 129-136.
[2] Sprinson, D.B. (1960) *Adv. Carbohydrate Chem.* **15**, 235-270.
[3] Pittard, A.J. (1987) *In* F.C. Neidhardt, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter, and H.E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol.**1**, pp. 368-394. American Society for Microbiology, Washington, D.C.
[4] Haslam, E. (1993) Shikimic acid: metabolism and metabolites. John Wiley & Sons, Chichester, United Kingdom.
[5] Folkers, K.(1996). *Biochem. Biophys. Res. Commun.* **224**(2), 358-361.
[6] Coggins, J.R., Duncan, K., Anton, I.A., Boocock, M.R., Chaudhuri, S., Lambert, J.M., Lewendon, A., Millar, G., Mousdale, D.M., Smith, D.D., (1987) *Biochem. Soc. Trans.* **15**(4), 754-759.
[7] Butler, J.R., Alworth, W.L., Nugent, M.J., (1974) *J. Am. Chem. Soc.* 96, 1617-1618.
[8] Mousdale, D.M., Campbell, M.S., Coggins, J.R. (1987) *Phytochemistry* **26**(10), 2665-2670.
[9] Duncan, K., Edwards, R.M., Coggins, J.R. (1987) *Biochem. J.* **246**(2), 375-386.
[10] Davies, G.M., Barret-Bee, K.J., Jude, D.A., Lehan, M., Nichols, W.W., Pinder, P.E., Thain, J.L., Watkins, W.J., Wilson, R.G. (1994) *Antimicrob. Agents Chemother*. **38**(2), 403-406.
[11] Coggins, J.R. (1989) The shikimate pathway as a target for herbicides. In Herbicides and Plant Metabolism (Dodge, A., ed.), Cambridge University Press, Cambridge, UK, pp. 97-112.
[12] Bentley, R. (1990). *Crit. Rev. Biochem. Mol. Biol.* **25**(5), 307-384.
[13] McDevitt, D., Payne, D.J., Holmes, D.J., and Rosenberg, M. (2002) *J. Appl. Microbiol.* **92**, 28S-34S.
[14] McArthur, J.D., West, N.P., Cole, J.N., Jungnitz, H., Guzman,C.A., Chin, J., Lehrbach, P.R., Djordjevic, S.P., and Walker, M.J. (2003) *FEMS Microbiol. Lett.* **221**(1), 7-16.
[15] Roberts, F., Roberts, C.W., Johnson, J.J., Kyle, D.E., Krell, T., Coggins, J.R., Coombs, G.H., Milhous, W.K., Tzipori, S., Fergunson, D.J.P., Chakrebarti, D., and McLeod, R. (1998). *Nature* **393**, 801-805.
[16] Kishore, G.M. & Shah, D.M. (1988) *Annu. Rev. Biochem.* **57**, 627-663.
[17] Dieckmann, A., and Jung, A. (1986) *Mol. Biochem. Parasitol.* **19**(2), 143-147.
[18] Campbell, S.A., Richards, T.A., Mui, E.J., Samuel, B.U., Coggins, J.R., McLeod, R., Roberts, C.W. (2004) *Int. J. Parasitol.* **34**(1), 5-13.
[19] McRobert, L., McConkey, G.A. (2002) *Mol. Biochem. Parasitol.* **119**(2), 273-278.
[20] McConkey G.A. (1999) *Antimicrob. Agents Chemother*. **43**(1), 175-177.
[21] Gardner, M.J., Shallom, S.J., Carlton, J.M., Salzberg, S.L., Nene, V., Shoaibi, A., Ciecko, A., Lynn, J., Rizzo, M., Weaver, B., Jarrahi, B., Brenner, M., Parvizi, B., Tallon, L., Moazzez, A., Granger, D., Fujii, C., Hansen, C., Pederson, J., Feldblyum, T., Peterson, J., Suh, B., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., White, O., Cummings, L.M., Smith, H.O., Adams, M.D., Venter, J.C., Carucci, D.J., Hoffman, S.L., Fraser, C.M. (2002) *Nature* **419**, 498-511.
[22] McConkey, G.A., Pinney, J.W., Westhead, D.R. (2004) *Trends Parasitol.* **20**(2), 60-65.
[23] Fitzpatrick T, Ricken S, Lanzer M, Amrhein N, Macheroux P, Kappes B. (2001) *Mol. Microbiol.* 40(1), 65-75.
[24] Ratledge, C. (1982). The Biology of the Mycobacteria. *Academic Press, London* **1**, 185-271.
[25] Voos, J.J., Rutter, K., Schroder, B.G., Su, H., Zhu, Y., and Barry, C.B. III. (2000) *Proc. Natl. Acad. Sci. USA*, **97**, 1252-1257.
[26] Parish, T., and Stoker, N. G. (2002) *Microbiology* **148**, 3069-3077.
[27] Whipp, M.J. & Pittard, A.J. (1995) *J. Bacterial.* **177**(6), 1627-1629.
[28] Millar, G., Lewendon, A., Hunter, M.G. & Coggins, J.R. (1986). *Biochem. J.* **237**(2), 427-437.
[29] De Feyter, R.C. & Pittard, J. (1986) *J. Bacteriol.* **165**(1), 331-333.
[30] Ely, B. and Pittard, J. (1979) *J. Bacteriol.* **138**(3), 933-943.
[31] De Feyter, R.C., Davidson, B.E., and Pittard, J. (1986) *J. Bacteriol.* **165**(1), 233-239.
[32] Vinella, D., Gagny, B., Joseleau-Petit, D., D'Ardi, R. & Cashel, M. (1996) *J. Bacteriol.* **178**(13), 3818-3828.
[33] Gu, Y., Reshetnikova, L., Li, Y., Wu, Y., Yan, H., Singh, S., Ji, X. (2002) *J. Mol. Biol.* **319**(3), 779-789.
[34] Krell, T., Coggins, J.R. & Lapthorn, A.J. (1998) *J. Mol. Biol.* **278**(5), 983-997.
[35] Krell, T., Maclean, J., Boam, D.J., Cooper, A., Resmini, M., Brocklehurst, K. *et al*. (2001) *Protein Sci*. **10**(6), 1137-1149.
[36] Dreusicke, D., Karplus, A. & Schulz, G.E. (1988) *J. Mol. Biol.* **199**(2), 359-371.
[37] Schlauderer, G.J. & Schulz, G.E. (1996) *Protein Sci*. **5**(3), 434-441.
[38] Stehle, T. & Schulz, G.E. (1990) *J. Mol. Biol*. **211**(1), 249-254.
[39] Müller-Dieckmann, H.-J. & Schulz, G.E. (1994) *J. Mol. Biol.* **236**(1), 361-367.
[40] Wild, K., Bohner, T., Aubry, A., Folkers, G. & Schulz, G.E. (1995) *FEBS Lett.* **368**(2), 289-292.
[41] Vonrhein, C., Schlauderer, G. J. & Schulz, G. E. (1995) *Structure*, **3**(5), 483-490.
[42] Müller, C.W., Schlauderer, G.J., Reinstein, J. & Schulz, G.E. (1996) *Structure*, **4**(2), 147-156.
[43] Gerstein, M., Schulz, G.E. & Chothia, C. (1993) *J. Mol. Biol.* **229**(2), 494-501.
[44] Walker, J.E., Saraste, M., Runswick, M.J. & Gay, N.J. (1982). *EMBO J.* **1**(8), 945-951.
[45] Smith, C.A. & Rayment, I. (1996) *Biophys. J.* **70**(4), 1590-1602.
[46] Reinstein, J., Schlichting, I., and Wittinghofer, A. (1990) *Biochemistry* **29**(32), 7451-7459.
[47] Sigal, I.S., Gibbs, J.B., D'Alonzo, J.S., Temeles, G.L., Wolanski, B.S., Socher, S.H., and Scolnick, E.M. (1986) *Proc. Natl. Acad. Sci. USA* **83**(4), 952-956.
[48] Leipe, D.D., Koonin, E.V., and Aravind, L. (2003) *J. Mol. Biol.* **333**(4), 781-815.
[49] Raviglione, M.C. (2003) *Tuberculosis*, **83**(1-3), 4-14.

Not For Distribution

[50] Corbett, E.L., Watt, C.J., Walker, N., Maher, D., Willians, B.G., Raviglione, M.C., Dye, C. (2003). *Arch. Intern. Med.* **163**(9), 1009-1021.
[51] Dye C, Scheele S, Dolin P, Pathania V, Raviglione, M.C. (1999) *JAMA* **282**(7), 677-686.
[52] World Health Organization: Global Tuberculosis Control. WHO Report 2004. ISBN 92 4 156264 1.
[53] Basso, L.A., and Santos, D.S. (2005) *Med. Chem. Rev.* Online **2**, 393-413.
[54] Duncan, K. (2003) *Tuberculosis*, **83**(1-3), 201-207.
[55] O'Brien, R.J., Nunn, P.P. (2001) *Am. J. Resp. Crit. Care Med.* **163**(5), 1055-1058.
[56] Garbe, T., Joens, C., Charles, I., Dougan, G., Young, D. (1990) *J. Bacteriol.* **172**(12), 6774-6782.
[57] Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M.A.; Rajandream, M.-A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J.E.; Taylor, K.; Whitehead, S. and Barrell, B.G. (1998) *Nature* **393**, 537-544.
[58] Oliveira, J.S., Pinto, C.A., Basso, L.A., Santos, D.S. (2001) *Protein Expr. Purif.* **22**(3), 430-435.
[59] De Azevedo, W. F. Jr., de Oliveira, J. S., Basso, L. A., Palma, M.S., Pereira, J. H., Canduri, F., and Santos, D. S. (2002) *Biochem. Biophys. Res. Commun*, **295**(1), 142-148.
[60] Dhaliwal, B., Nichols, C.E., Ren, J., Lockyer, M., Charles, I., Hawkins, A.R., Stammers, D.K. (2004) *FEBS Lett.* **574**(1-3), 49-54.
[61] Pereira, J.H., Oliveira, J.S, Canduri, F., Dias, M.V.B., Palma, M.S., Basso, L.A., Santos, D.S., and Azevedo Jr., W.F. (2004) *Acta Cryst. Section D* **60**, 2310-2319.
[62] Matte, A., Tari, L.W., and Delbaere, T.J. (1998) *Structure*, **6**(4), 413-419.
[63] Williams, R.L., Oren, D.A., Muñoz-Dorado, J., Inouye, S., Inouye, M., and Arnold, M. (1993) *J. Mol. Biol.* **234**(4), 1230-1247.
[64] Jencks, W.P. (1975) *Adv. Enzymol.* **43**, 219-410.
[65] Smith, C.A., Reyment, I. (1995) *Biochemistry*, **34**(28), 8973-8981.
[66] Berchthold, H., Reshetnikovca, L., Reiser, C.O.A., Schrimer, N.K., Sprinzl, M., Hilgenfeld, R. *Nature* (1993) **365**(6442), 126-132.
[67] Pai, E.F., Krengel, U., Petski, G.A., Goody, R.S., Kabsch, W., Wittinghofer, A. . (1990) *EMBO J.* **9**(8), 2351-2359.
[68] Coleman, D.E., Berghuis A.M., Lee, E., Linder, M.E., Gilman, A.G., Sprang, S.R. (1994) *Science* **265**(5177), 1405-1412.
[69] Abele, U. and Schulz, G. E. (1995) *Protein Sci.* **4**(7), 1262-1271.
[70] Schlauderer, G.J. & Schulz, G.E. (1996) *Protein Sci.* **5**(3), 434-441.
[71] Hasemann, C. A., Istvan, E. S., Uyeda, K. & Deisenhofer, J. (1996) *Structure* **4**(9), 1017-1029.
[72] Romanowski, M.J., Burley, S.K. (2002) *Proteins: Structure, Function Genet.* **47**(4), 558-562.
[73] Bennett, W.S. and Steitz, T.A. (1980) *J. Mol.Biol.* **140**(2), 211-230.
[74] Schulz, G.E., Müller, C.W. and Diederichs, K. (1990) *J. Mol. Biol.* **213**(4), 627-630.
[75] Idziak, C., Price, N.C., Kelly, S.M., Krell, T., Boam, D.J., Lapthorn, A.J., and Coggins, J.R. (1997) *Biochem. Soc. Trans.* **25**(4), S627.
[76] Cerasoli, E., Kelly, S.M., Coggins, J.R., Lapthorn, A.J., Clarke D.T. (2003) *Biochim. Biophys.Acta*, **1648**(1-2), 43-54.
[77] Pittard, A.J., and Wallace, B.J. (1966) *J. Bacteriol.* **92**(4), 1070-1075.
[78] Whipp, M.J., Camakaris, H., and Pittard, A.J. (1998) *Gene* **209**(1-2), 185-192.
[79] Mildvan, A.S. (1979) *Adv. Enzymol. Relat. Areas Mol. Biol.* **49**, 103-126.
[80] Schlichting, I., and Reinstein, J. (1997) *Biochemistry* **36**(31), 9290-9296.
[81] Xu, Y.-W., Moréra, S., Janin, J., and Cherfils, J. (1997) *Proc. Natl. Acad. Sci. USA*, **94**(8), 3579-3583.
[82] Scheffzek, K., and Wittinghofer, A. (1997) *Science*, **277**, 333-338.
[83] Basso, L. A., Pereira da Silva, L. H., Fett-Neto, A. G., de Azevedo Jr., W. F., Moreira, I. S., Palma, M. S., Calixto, J. B., Astolfi Filho, S., dos Santos, R. R., Soares, M. B. P., Santos, D. S. (2005) *Mem. Inst. Oswaldo Cruz*, **100**(6): 475-506.
[84] da Silveira, N. J. F., Uchoa, H. B., Canduri, F.,Pereira, J. H., Camera Jr., J. C., Basso, L. A., Palma, M. S., Santos, D. S., de Azevedo Jr., W. F. (2004) *Biochem Biophys. Res. Commun.* **322**(1), 100-104.
[85] Uchoa, H. B., Jorge, G. E., da Silveira, N. J., Camera J. C., Canduri, F., De Azevedo, W.F. (2004) *Biochem. Biophys. Res. Commun.* **325**(4), 1481-1486.
[86] da Silveira, N. J. F., Bonalumi, C. A., Uchoa, H. B., Pereira, J. H., Canduri, F.,Pereira, J. H., de Azevedo Jr., W. F. (2006) *Cell Biochem. Biophys.* **44**(3), 366-374.
[87] De Azevedo, W.F. Jr.; Mueller-Dieckmann, H.J.; Schulze-Gahmen, U.; Worland, P.J.; Sausville, E.; Kim S.-H. (1996) *Proc. Natl. Acad. Sci. USA*. **93**(7), 2735-2740.
[88] Kim, S.-H.; Schulze-Gahmen, U.; Brandsen, J.; De Azevedo, W. F. Jr. (1996) *Prog. Cell Cycle Res.* **2**, 137-145.
[89] Canduri, F.; Uchoa, H.B.; de Azevedo, W.F.Jr. (2004) *Biochem. Biophys. Res. Commun.* **324**(2), 661-666.
[90] De Azevedo, W.F.Jr.; Canduri, F.; Silveira, N.J.F. (2002) *Biochem Biophys. Res. Commun.* **293**(1), 566-571.
[91] De Azevedo, W.F.Jr.; Leclerc, S.; Meijer, L.; Havlicek, L.; Strnad, M.; Kim, S.-H. (1997) *Eur. J. Biochem.* **243**, 518-526.
[92] De Azevedo, W.F.Jr., Gaspar, R.T., Canduri, F.; Camera, J.C.Jr., Silveira, N.J.F. (2002) *Biochem Biophys. Res. Commun.* **297**(5), 1154-1158.
[93] De Azevedo, W. F. Jr., Canduri, F. (2005) *Curr. Comput. Aided Drug Design* **1**, 53-64.
[94] Gan, J., Gu, Y., Li, Y., Yan, H., Ji, X. (2006) *Biochemistry* ***45,*** 8539-8545.
[95] Dias, M. V. B., Faim, L. M., Vasconcelos, I. B., Oliveira, J. S., Basso, L. A., Santos, D. S., De Azevedo, W. F. *Acta Crystallographica F.* (In press).
[96] Ducati, R. D., Basso, L. A., Santos, D. S. (2007) *Curr. Drug Targets* **8**(1) In press.
[97] van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. Berendsen, H. J. C. (2005) *J. Comp. Chem.* **26**(16),1701-1718.

Received: December 12, 2006 Accepted: May 26, 2006 Updated: September 11, 2006

Non-commercial Distribution

5.5. Crystallographic studies on the binding of isonicotinyl-NAD adduct to wild-type and isoniazid resistant 2-*trans*-enoyl-ACP (CoA) reductase from *Mycobacterium tuberculosis*. J Struct Biol. 2007;159(3):369-80.

Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 159 (2007) 369–380
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

Crystallographic studies on the binding of isonicotinyl-NAD adduct to wild-type and isoniazid resistant 2-*trans*-enoyl-ACP (CoA) reductase from *Mycobacterium tuberculosis*

Marcio Vinicius Bertacine Dias [a], Igor Bordin Vasconcelos [b], Adriane Michele Xavier Prado [a], Valmir Fadel [a], Luiz Augusto Basso [b], Walter Filgueira de Azevedo Jr. [c,*], Diógenes Santiago Santos [b,*]

[a] *Programa de Pós-Graduação em Biofísica Molecular—Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil*
[b] *Centro de Pesquisa em Biologia Molecular e Funcional, Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil*
[c] *Faculdade de Biociências—Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681, Porto Alegre-RS CEP 90619-900, Brazil*

Received 16 January 2007; received in revised form 27 March 2007; accepted 11 April 2007
Available online 3 May 2007

Abstract

The resumption of tuberculosis led to an increased need to understand the molecular mechanisms of drug action and drug resistance, which should provide significant insight into the development of newer compounds. Isoniazid (INH), the most prescribed drug to treat TB, inhibits an NADH-dependent enoyl-acyl carrier protein reductase (InhA) that provides precursors of mycolic acids, which are components of the mycobacterial cell wall. InhA is the major target of the mode of action of isoniazid. INH is a pro-drug that needs activation to form the inhibitory INH–NAD adduct. Missense mutations in the *inhA* structural gene have been identified in clinical isolates of *Mycobacterium tuberculosis* resistant to INH. To understand the mechanism of resistance to INH, we have solved the structure of two InhA mutants (I21V and S94A), identified in INH-resistant clinical isolates, and compare them to INH-sensitive WT InhA structure in complex with the INH–NAD adduct. We also solved the structure of unliganded INH-resistant S94A protein, which is the first report on apo form of InhA. The salient features of these structures are discussed and should provide structural information to improve our understanding of the mechanism of action of, and resistance to, INH in *M. tuberculosis*. The unliganded structure of InhA allows identification of conformational changes upon ligand binding and should help structure-based drug design of more potent antimycobacterial agents.
© 2007 Elsevier Inc. All rights reserved.

Keywords: *Mycobacterium tuberculosis*; InhA; Crystal structure; Isoniazid; Drug resistance; Enoyl-ACP reductase

1. Introduction

Tuberculosis (TB), which is caused mainly by *Mycobacterium tuberculosis*, is a global human health emergency that remains the leading cause of mortality among infectious diseases. It has been estimated that 8.2 million new TB cases occurred worldwide in the year 2000, with approximately 1.8 million deaths in the same year, which translates into more than 200 deaths per hour, and more than 95% of these were in developing countries (Corbett et al., 2003). In the same year, 3.2% of the world's new cases of TB were multidrug-resistant tuberculosis (MDR-TB), defined as strains resistant to at least isoniazid and rifampicin (Espinal, 2003; Ormerod, 2005). Treatment of MDR-TB strains requires the administration of second-line drugs that are more toxic and less effective, and are given

* Corresponding authors. Fax: +55 17 51 3220 3629.
E-mail addresses: walter.junior@pucrs.br (W.F. de Azevedo Jr.), diogenes@pucrs.br (D.S. Santos).

1047-8477/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2007.04.009

for at least three times as long as, and 100 times as expensive as basic chemotherapeutic regimens (Pablos-Mendez et al., 2002). More recently, a survey of the frequency and distribution of extensively drug-resistant (XDR) TB cases, which are defined as cases in persons with TB whose isolates were resistant to isoniazid and rifampicin and at least three of the six main classes of second-line drugs, showed that during 2000–2004, of 17,690 TB isolates, 20% were MDR and 10% of these were XDR (CDC, 2006). XDR-TB has a wide geographic distribution, poses a public health threat, is an impediment to TB control, and opens up the possibility that epidemics of virtually untreatable TB may develop (Anon, 2006). New antimycobacterial agents are thus needed to improve the treatment of MDR- and XDR-TB, as well as to provide more effective treatment of drug-sensitive TB infection. An understanding of drug resistance mechanisms in this pathogen should contribute to the rational design of new chemotherapeutic agents to treat TB.

The modern, standard "short-course" therapy for tuberculosis is based on a four-drug regimen of isoniazid, rifampicin, pyrazinamide, and ethambutol or streptomycin for two months, followed by treatment with a combination of isoniazid and rifampicin for four months (Mitchison, 1985). Isoniazid (INH, isonicotinic acid hydrazide) was first reported to be effective in the treatment of TB in 1952 (Bernstein et al., 1952) and, soon after, the first INH-resistant *M. tuberculosis* strains were isolated (Middlebrook and Cohn, 1953). Genetic and biochemical studies have shown that the *inhA*-encoded protein is the primary target for isoniazid (Banerjee et al., 1994; Quémard et al., 1995; Larsen et al., 2002; Kremer et al., 2003). InhA was identified as an NADH-dependent 2-*trans*-enoyl-ACP (acyl carrier protein) reductase enzyme that exhibits specificity for long-chain thioester substrates. InhA is a member of the mycobacterial type II fatty acid synthase system (FAS-II), which elongates acyl fatty acid precursors yielding the long carbon chain of the meromycolate branch of mycolic acids, the hallmark of mycobacteria (Schroeder et al., 2002). INH is a pro-drug that is activated by the mycobacterial catalase-peroxidase enzyme KatG in the presence of manganese ions, NAD(H) and oxygen (Johnsson and Schultz, 1994; Johnsson et al., 1995; Basso et al., 1996; Zabinski and Blanchard, 1997). The KatG-produced acylpyridine fragment of isoniazid is covalently attached to the C4 position of NADH forming an INH–NAD adduct, which, in turn, forms an inhibitory binary complex with the wild-type (WT) enoyl reductase of *M. tuberculosis* (Rozwarski et al., 1998) with an equilibrium dissociation constant value lower than 0.4 nM (Lei et al., 2000). The isonicotinyl-NAD adduct has been shown to be a slow, tight-binding competitive inhibitor of WT InhA with an overall inhibition constant value of 0.75 nM (Rawat et al., 2003).

The mechanism of action of isoniazid is complex, as mutations in at least five different genes (*katG*, *inhA*, *ahpC*, *kasA*, and *ndh*) have been found to correlate with isoniazid resistance (Schroeder et al., 2002; Blanchard, 1996; Basso and Blanchard, 1998; Glickman and Jacobs, 2001; Basso and Santos, 2005; Oliveira et al., 2007). Consistent with InhA as the primary target of INH mode of action, INH-resistant clinical isolates of *M. tuberculosis* harboring *inhA*-structural gene missense mutations, but lacking mutations in the *inhA* promoter region, *katG* gene and *oxyR-ahpC* region, were shown to have higher dissociation constant (K_d) values for NADH than INH-sensitive WT InhA, whereas there were only modest differences in the steady-state parameters (Blanchard, 1996). We have recently reported the crystal structures of binary complexes formed between NADH and INH-sensitive WT InhA, and INH-resistant S94A, I47T, and I21V InhA mutant enzymes (Oliveira et al., 2006). Even more recently, both specialized linkage transduction has been used to introduce S94A single point mutation within the *inhA* structural gene and X-ray crystallographic on INH-resistant S94A InhA protein has been reported (Vilchèze et al., 2006). However, even though there are several crystal structures of InhA in complex with a variety of ligands, there has been no report on unliganded InhA structure and, thus, no high resolution information on the InhA structure before ligand binding. In our efforts to understand the molecular basis for the reduced inhibition of the INH–NAD adduct to InhA mutants, here we report co-crystallization of INH-resistant I21V and S94A InhA mutant enzymes, which were identified in INH-resistant clinical isolates of *M. tuberculosis* (Blanchard, 1996; Morlock et al., 2003), with the INH–NAD adduct, and compare them to the INH-sensitive WT InhA structure. This is the first report on the crystal structure of the complex formed between INH-resistant I21V InhA and INH–NAD adduct refined to 2.2 Å. We also report the crystal structure of INH-resistant S94A InhA and INH-sensitive WT InhA both in complex with INH–NAD adduct to 2.0 and 2.2 Å of resolution, respectively. Moreover, we report, for the first time, the crystal structure of apo INH-resistant S94A InhA refined to 2.15 Å, which shows the protein conformational changes upon ligand binding. It is hoped that the data presented here will provide structural insight into an understanding of the drug resistance mechanism, which, in turn, should aid the rational design of chemical compounds to efficiently inhibit both INH-resistant and -sensitive InhA enzymes with potential antimycobacterial activity.

2. Materials and methods

2.1. Crystallization

WT, I21V, and S94A InhA enzymes were expressed and purified to homogeneity as described elsewhere (Quémard et al., 1995; Basso et al., 1998). INH–NAD synthesis was carried out as described elsewhere (Rozwarski et al., 1998). Crystals of binary complex InhA:INH–NAD were obtained by the hanging-drop vapor-diffusion method under similar conditions as described by Dessen et al.

(1995) and Rozwarski et al. (1998). InhA enzymes were dialyzed against 50 mM Hepes, pH 7.5 and concentrated to 5–10 mg/mL. Enzyme–inhibitor complexes were obtained at room temperature by incubating 5–10 mg/mL InhA proteins with NADH, $MnCl_2$ and isoniazid for 1 h using molar ratios of 1:50, 1:10, and 1:100, respectively. The complexes InhA:INH–NAD were crystallized in hanging droplets containing 1 μL of inhibited-protein solution and 1 μL of crystallization solution containing 50 mM Hepes, pH 7.2; Sodium citrate buffer and 5–10% MPD.

The apoenzyme S94A InhA mutant was dialyzed against 50 mM Hepes, pH 7.5, 10% glycerol and 300 mM of KCl. The protein was concentrated to 10 mg/mL and crystallized in hanging droplets containing 1 μL of protein solution and 0.5 μL of crystallization solution containing 100 mM sodium citrate, pH 5.6, 200 mM ammonium acetate and 20–30% of PEG 4000.

2.2. Data collection and processing

All data sets were collected at a wavelength of 1.427 Å using Synchrotron Radiation Source (Station PCr, LNLS, Campinas, Brazil; and a CCD detector (MARCCD) (Polikarpov et al., 1998). Data collection were performed in cryogenic conditions at 100 K in a cold nitrogen stream generated and maintained with an Oxford Cryosystem. Prior to flash-cooling, glycerol was added to the crystallization drop up to 20% (v/v). The data sets were processed using the program MOSFLM (CCP4, 1994) and scaled with SCALA (CCP4, 1994).

2.3. Structure determination

The crystal structures were determined by standard molecular replacement methods using the program AMoRe (Navaza, 1994). Initially, atomic coordinates for binary complex WTInhA:INH–NAD to 2.7 Å resolution were used as search model for the structure of WT InhA:INH–NAD (PDB Access code: 1ZID) (Rozwarski et al., 1998). Our atomic coordinates for WT InhA:INH–NAD at 2.2 Å resolution were then used as search model to solve the other structures. The structure of apoenzyme S94A InhA mutant was solved using as search model the structure of binary complex InhA:Genz-10850 at 2.6 Å resolution (PDB Access code: 1P44) (Kuo et al., 2003). The atomic positions, which generated the higher correlation coefficient magnitude obtained from molecular replacement method, were used for the crystallographic refinement. The refinements of structures were performed using REFMAC 5.2 program (Murshudov et al., 1997). The XtalView/Xfit (McRee, 1999) was used for visual inspection and addition of water molecules. The water molecules were also checked based on *B*-factor values. The stereochemistry correctness of the models was checked using PROCHECK program (Laskowski et al., 1993) and PARMODEL (Uchoa et al., 2004). The final atomic models were superposed using the program LSQKAB program from CCP4i package (CCP4, 1994).

3. Results and discussion

The crystals of apo S94A InhA are triclinic. These crystals diffracted to 2.15 Å resolution. The asymmetric unit presents four monomers forming the characteristic tetramer of InhA. This structure presents final *R*-factor and *R*-free values of 16.2% and 25.5%, respectively. The crystals of WT, I21V, and S94A InhA in complex with the INH–NAD adduct are hexagonal and crystallized in the space group $P6_222$, having one molecule in the asymmetric unit. Table 1 summarizes the data processing, crystallographic refinement statistics and structural quality for the four structures presented here. Analysis of the crystallographic refinement for the structures here described indicates that WtInhA:NADH:INH, I21V InhaA:NADH–INH and S94Inha:NADH–INH present difference between *R*-factor and *R*-free ranging from 3.9% to 5.5%, which indicates fairly good overall refinement statistics, and the structure S94A InhA presents a difference of 9.3%, below 10% , which makes this structure also acceptable for structural comparisons. These structures present good geometry, although some residues are located in regions not permitted in Ramachandran plots.

InhA belongs to the short chain dehydrogenase/reductase (SDR) family of enzymes. The main characteristic of this family is a polypeptide backbone topology in which each subunit consists of a single domain with a central core that contains a Rossmann fold supporting an NADH binding site. The structure displays a α/β folding consisting of a central β-sheet composed of parallel strands and flanked by α-helices (Fig. 1a). The structure presented here is in accordance with the homotetrameric quaternary structure in solution determined by analytical size-exclusion chromatography, and possesses an internal 222 symmetry (Fig. 1b).

3.1. Influence of INH on the structure of MtInhA

The activated form of INH consists of an isonicotinic-acyl group attached through its carbonyl group to the C4 of the nicotinamide ring, replacing the 4*S* hydrogen of NADH, which is the same position involved in the hydride transfer that occurs during reduction of enoyl-ACP substrates (Quémard et al., 1995).

We compared the structure of INH-sensitive WT InhA presented here and the structure previously solved by Rozwarski et al. (1998) at 2.7 Å of resolution. Although Vilchèze et al. (2006) have solved two structures of InhA in complex with NAD–INH (WT InhA and S94A mutant at 2.0 and 1.9 Å at resolution, respectively) it was only observed one structure deposited in the Protein Data Bank (S94A InhA mutant). This way, the S94A InhA solved by Vilchèze et al. (2006) is argued to be related to effect of this

Table 1
Crystallographic data, refinement statistics and analysis of the quality of InhA structures

Crystallographic data	WtInhA:NADH:INH	I21V InhA:NADH INH	S94A InhA:NADH INH	S94A InhA
a (Å)	97.09	96.59	96.83	54.63
b (Å)	97.09	96.59	96.83	63.52
c (Å)	136.94	136.35	136.42	65.18
α	90	90	90	97.21
β	90	90	90	85.81
γ	120	120	120	102.87
Space group	*P*6222	*P*6222	*P*6222	*P*1
Number of image	80	60	80	160
Number of measurements with $I > 2\ \sigma\ (I)$	245,823	134,890	185,407	77,775
Number of independent reflections	26,409	19,563	19,782	44,054
Completeness (%) (outermost shell)	99.9 (100)	98.7 (91.6)	99.8 (99.5)	98.6 (85.2)
R_{sym}[a] (%) (outermost shell)	9.8 (83.0)	7.2 (59.1)	9.5 (77.4)	5.8 (29)
Redundancy	9.3	6.9	9.4	1.6
Refinement statistics				
Resolution range (Å)	40.19 2.0	39.41 2.2	45.64 2.2	42.26 2.14
Reflections used for refinement	25,045	18,527	18,193	41,818
Number of atoms	2283	2261	2257	8682
Number of residues	267	267	267	1068
Number of water molecules	239	216	212	701
Final *R*-factor [b] (%)	18.7	18.4	17.9	16.2
Final *R*-factor free[c] (%)	22.6	23.6	23.4	25.5
Correlation coefficient (%)	96.0	95.8	95.8	95.8
B values (A^2)				
Main chain	30.56	36.69	34.08	27.52
Side chain	32.19	37.70	35.33	28.59
NADH INH	32.19	31.93	29.85	
Waters	40.93	43.39	42.13	33.23
Ramachandran plot				
Favorable	86.2	87.1	88.9	88
Additional allowed	9.8	9.8	9.3	10.3
Generously allowed	0.0	0.0	0.0	0.7
Disallowed	4.0	3.1	1.8	1.0

[a] $R_{sym} = 100\ PI(h)\ \_\ ÆI(h)æ/PI(h)$ with $I(h)$, observed intensity and $ÆI(h)æ$, mean intensity of reflection h overall measurement of $I(h)$.
[b] *R*-factor $= 100 \times \sum(F_{obs} - F_{calc})/\sum(F_{obs})$, the sums being taken over all reflections with $F/\sigma(F) > 2\sigma(F)$.
[c] *R*-free = *R*-factor for 10% of the data that were not included during crystallographic refinement.

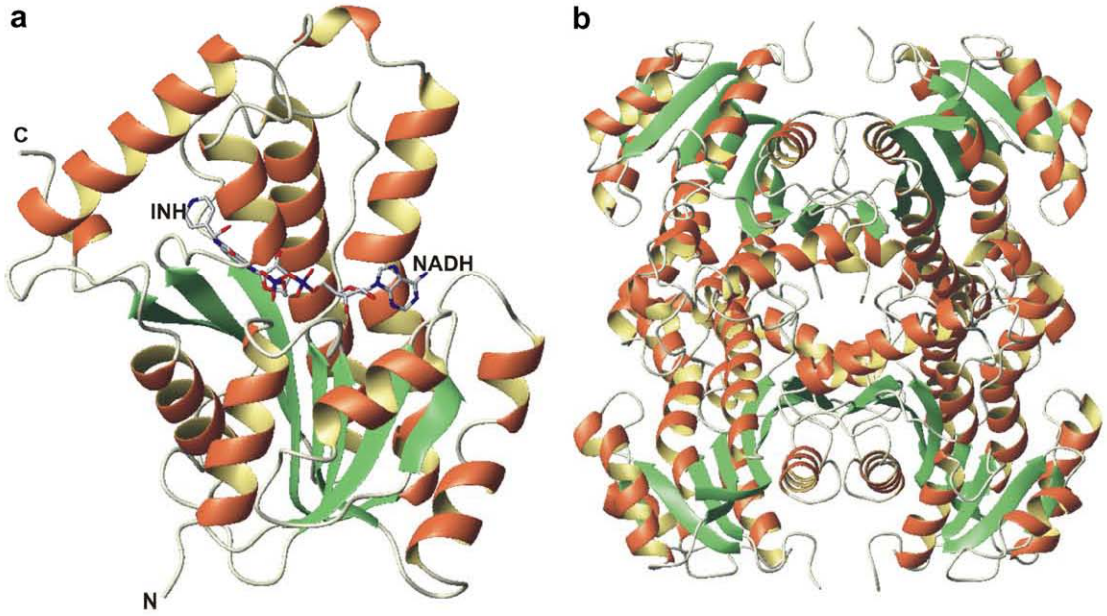


Fig. 1. Ribbon diagram of InhA. (a) Monomeric structure of the wild-type InhA:INH NAD binary complex, and (b) Tetrameric structure of the unliganded InhA. The figures were drawn using program MolMol (Koradi et al., 1996).

mutation on the binding process of the INH–NAD at a later stage.

The structure of the INH-sensitive WT InhA presented here shows differences when compared with the structure determined by Rozwarski et al. (1998), mainly in important residues located in the binding site of INH–NAD (Tyr158, Phe149, Trp222, Leu218, and Phe41) and in the INH–NAD molecule. There is a difference of approximately 15° in the position of Tyr158 side chain of WT InhA in comparison with the structure solved by Rozwarski et al. (1998), and the Tyr158 side chain moves closer to INH. Interestingly, a conformational change involving rotation of Tyr158 side chain upon binding of the enoyl substrate to InhA has been invoked to account for the observed inverse solvent isotope effect (Parikh et al., 1999). These authors have proposed that Tyr158 functions as an electrophilic catalyst, stabilizing the transition state for hydride transfer by hydrogen bonding to the substrate carbonyl. The Phe41 side chain of WT InhA:INH–NAD structure presented here underwent 30° torsion as compared to the structure solved by Rozwarski et al. (1998). This residue appears to be important in anchoring the adenine moiety of NADH. The side chain of Phe149 and Leu218 residues move closer to INH molecule (Fig. 2a). Although it could be argued that these differences are due to the lower resolution of the previously published structure, these residues appear to play an importance role in the enzyme–ligand interaction and should be considered in structure-based drug design.

The main effect of INH moiety of INH–NAD adduct on WT InhA is a 90° rotation of Phe149 side chain (Fig. 2b), which provides room to accommodate the INH moiety. Owing to this rotation, a water molecule, which has been identified in the InhA structures without the isonicotinic-acyl group (Oliveira et al., 2006), was removed from the InhA active site (not shown in Fig. 2). Furthermore, it is observed alterations in the Met155, Leu218 and Ile215, Met199, probably due to alteration in the side chain position of Phe149. The Ile215 moves away from the active site due to rotation of the side chain of the Tyr158 (Fig. 2b). The side chain of Met199 underwent a rotation of approximately 30° away from the active site. Two water molecules are present in the structure of InhA:INH–NAD complex, but absent from the structures between WT and S94A, I21V, and I47T mutants InhA proteins in complex with NADH (Oliveira et al., 2006). The isonicotinic group also causes the expulsion of one water molecule next to C4 of NADH molecule. This causes a rotation of approximately 20° of nicotinamide group of NADH molecule, when compared with other InhA:NADH structures, but the phosphates, ribose and adenine are highly superposed among the analyzed structures. This change results in a more

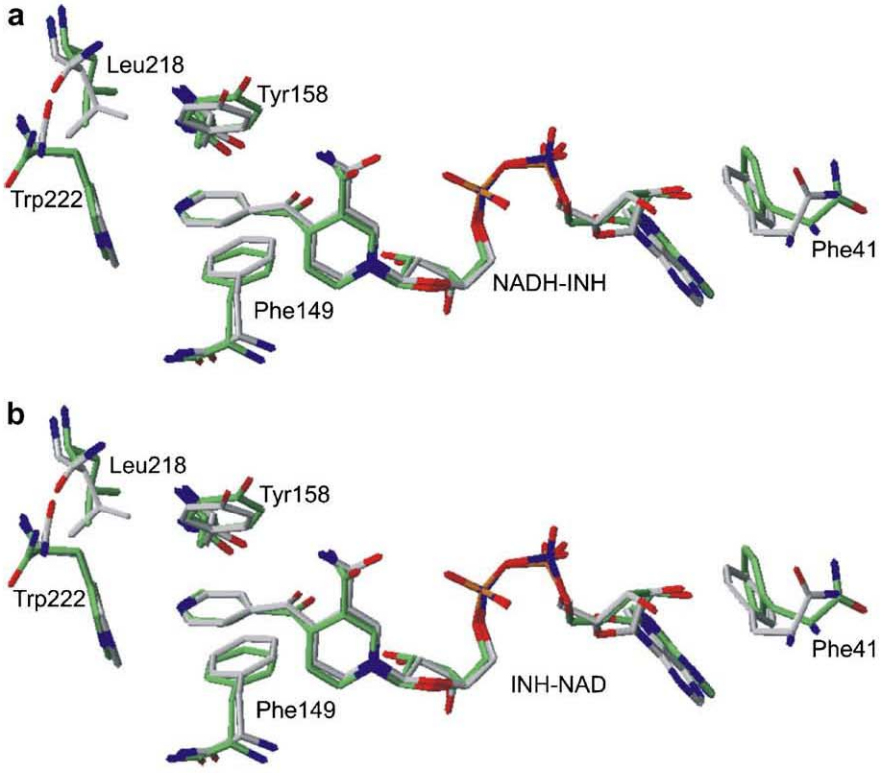


Fig. 2. Superposition of active site residues of InhA. (a) It is shown the superposition of the active site structure of wild-type InhA:INH NAD solved by Rozwarski et al. (1998) at 2.7 Å resolution (in green) and wild-type InhA:INH NAD at 2.0 Å resolution presented here (in grey). (b) It is shown the superposition of the structure of wild-type InhA:NADH complex binary solved by Oliveira et al. (2006) (in blue) and the structure of wild-type InhA: INH NAD presented here (in grey). The figures were generated with MolMol (Koradi et al., 1996). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

snugly fit of the INH moiety of INH–NAD in the active site of InhA.

There is one intermolecular hydrogen bond that is conserved in the present structures and in the structure reported by Rozwarski et al. (1998) involving oxygen atom from Gly14 and O2 from the pentose of NADH, which strongly indicates the importance of this hydrogen bond for the interaction of NADH and the InhA.

Table 2
Effect of the mutations I21V and S94A on the binding process of the NADH on the InhA

Mutation	Mutation effect	Consequence
2I21V	Lose of Van der Waals interaction between NADH and the CD1 atom present in the valine residue	Decrease of the stability of the binding of the NADH in the active site of the protein
S94A	Alteration in the binding network involving a conserved water molecule and O9 atom of molecule of NADH	Increase of the flexibility of the conserved water molecule and decrease of the affinity of NADH by protein

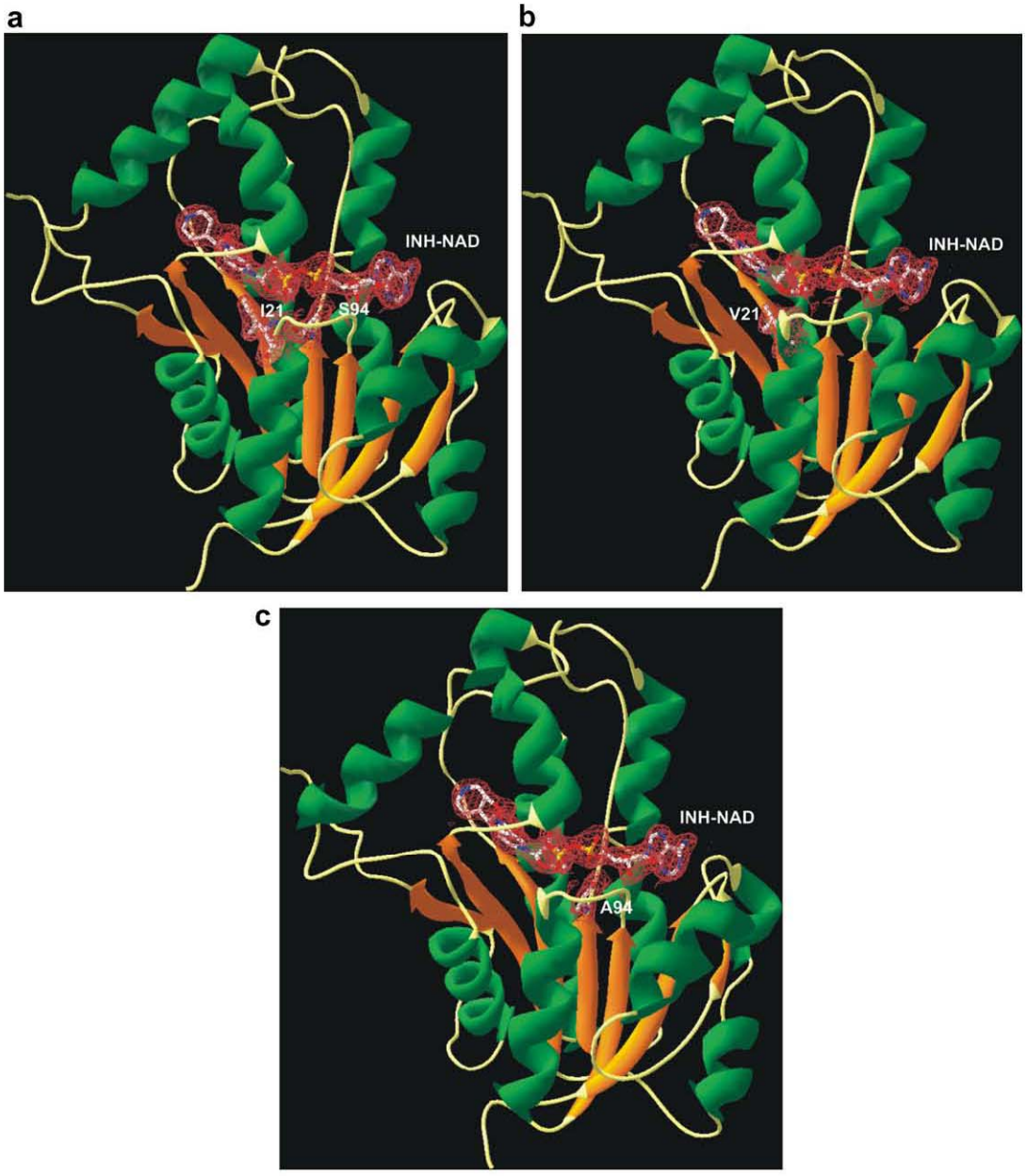


Fig. 3. Ribbon diagrams for InhA structures, showing the electronic density for INH NAD and mutated residue. (a) Wild-type InhA, (b) I21V mutant, and (c) S94A mutant. The figures were generated using program SpdbView (Guex and Peitsch, 1997).

3.2. Influence of mutations I21V and S94A on INH–NAD binding to InhAs

Van der Waals contacts between Val21 side chain and the nicotinamide ring, nicotinamine ribose, and phosphate oxygen atoms are missing in the INH-resistant I21V InhA structure as compared to WT-sensitive InhA. These interactions do not occur due to the absence of the CD1 atom in the mutated residue Val21 (Oliveira et al., 2006), which appear to play an important role in stabilizing bound NADH or INH–NAD in the InhA active site (Oliveira et al., 2006; Basso et al., 1998).

The crystal structure of S94A InhA showed that a conserved hydrogen bond to a water molecule is lost owing to mutation to Ala94 (Oliveira et al., 2006), which has been proposed to account for the reduction in affinity for NADH observed for the INH-resistant S94A InhA enzyme (Basso et al., 1998). Vilchèze et al. (2006) have solved the structure of INH-resistant S94A InhA enzyme (PDB code 2NV6) and INH sensitive InhA in complex with INH–NAD (data do not deposited in the PDB). These authors propose the loss of the serine causes a shift in position of water molecule that promotes a disruption in hydrogen binding network. Analysis of the structure of S94A InhA enzyme present here does not present the disruption between this water molecule and the atom O9 of the molecule of INH–NAD. Analysis of other structures of S94A InhA, such a ones solved by Oliveira et al. (2006), shows a binding between this conserved water molecule and the O9 of the molecule of INH–NAD. Thus, we believed that the mutation S94A can cause increase flexibility or decrease affinity of this molecule due the loss binding between this water molecule and the OG atom of the Ser94, which is observed in the structure wild-type. The influence of the mutation on this water molecule could cause a reduction in the affinity of INH–NAD or NADH by InhA S94A. Thus, the movement observed in the water molecule by Vilchèze et al. (2006) can be due an increase of flexibility of this molecule carrying a false impression of an ordered movement.

The effect of the mutations I21V and S94A on binding process of the NADH on protein is summarized in the Table 2. Analysis of contact area between the InhAs and INH–NAD reveals that the INH-sensitive WT enzyme presents larger contact area than the INH-resistant mutants. The value for WT enzyme is 467.4 $Å^2$ while for S94A and I21V mutants the values are 464.2 and 461.9 $Å^2$, respectively. The smaller value observed for I21V mutant can be due the absence of the contact between the CD1 of Ile21 that is missing in the I21V mutant. The I21V and S94A mutations do not seem to alter the position of the isonicotinic-acyl group because there are no significant changes when the mutant structures are compared with WT structure (Fig. 3a–c). These results are in agreement with the proposal that INH resistance is due to reduction in NADH affinity for mutant InhAs (S94A and I21V) thereby hampering binding of INH–NAD adduct.

3.3. Effect of INH–NAD on the binding of substrate in the InhA

The isonicotinic-acyl group of INH–NAD adduct replaces the 4*S* hydrogen of C4 of NADH, which is the same position involved in the hydride transfer catalyzed by InhA. As a member of the family of SDR enzymes, the catalytic triad formed by Phe149, Tyr158, and Lys165 (Rozwarski et al., 1999) is present in InhA (Fig. 4). The Phe149 may play a role in anchoring the nicotinamide moiety of NADH for hydride transfer to the fatty acyl substrate. Lys165 interacts with the 3′-hydroxyl oxygen of the nicotinamide ribose of NADH, in agreement with a previous proposal that Lys165 plays a role in cofactor binding (Parikh et al., 1999).

The proposed mechanism of how InhA catalyzes the reduction of the 2-*trans* double bond of the fatty acyl substrate consists of the formation of an enolate intermediate through the direct transfer of a hydride ion from NADH to position C3 of the substrate, followed by protonation of position C2 (Rozwarski et al., 1999). It is noteworthy that the Tyr158 occupies approximately the same position for S94A:INH–NAD and apo S94A, whereas there is a conformational change upon binding of NADH to S94A, and NAD^+ and a C16 enoyl substrate to WT InhA resulting in position of Tyr158 to similar positions in these latter structures (Fig. 4). A comparison between the WT InhA–NADH structure and an inactive ternary complex formed by WT InhA, NAD^+, and a C16 enoyl substrate revealed

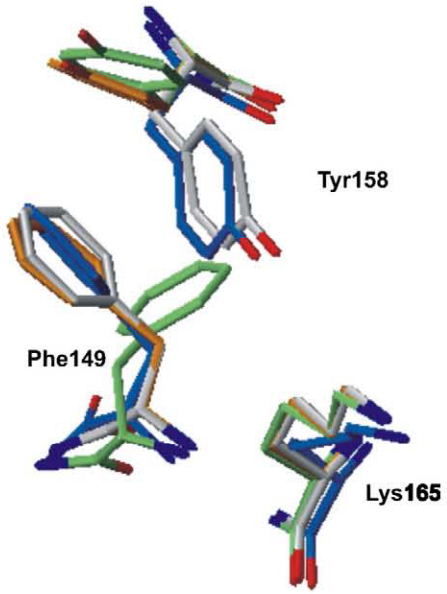


Fig. 4. Catalytic triad for InhA. In green is shown the residues of S94A InhA in complex with INH NAD, in grey is shown the residues of S94A InhA unliganded, in orange is shown the residues of S94A in complex with NADH, and in blue is shown the residues of WT InhA in complex with C16 enoyl substrate and NAD^+ (Rozwarski et al., 1999). The figure shows conformational changes in these residues due to different ligands. The figure was generated using program MolMol (Koradi et al., 1996). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

that upon binding the enoyl substrate there is a 60° rotation about the Tyr158 Cα—Cβ bond that enables the Tyr158 to hydrogen bond to the substrate carbonyl group (Parikh et al., 1999). It has thus been suggested that Tyr158 provides electrophilic stabilization of the transition state(s) for the reaction by hydrogen bonding to the carbonyl of the substrate. However, the InhA:NAD:C16 enoyl substrate is a non-productive ternary complex and is likely not part of the reaction course, which is in agreement with the structural results presented here for apo S94A InhA. The isonicotyl-acyl group causes a large change in the position of the Phe149 side chain (Fig. 4), and appears to be involved in stacking interactions with the INH–NAD adduct.

3.4. INH-resistant S94A mutant in an uncomplexed form

The structure of unliganded S94A InhA reveals that the protein undergoes three noticeable conformational changes in its main chain upon NADH binding. The main alterations occur between the residues 99 and 112, between the residues 197 and 213 (Fig. 5a), and in the Phe41 (Fig. 5b). It appears that NADH binding is sufficient to cause a closure of the active site of InhA. Analysis of *B*-factor for three different structures of InhA (Fig. 6) shows that in the absence of both substrate and cofactor or in the presence of ones, the protein presents the substrate binding loop (residues 196–219) more disordered than the structure in the presence of cofactor. By analysis of these three structure, we can observe that in the uncomplexed structure, part of the substrate binding loop is closed (197–203) and part is opened (205–214) for InhA:NAD^+:substrate structure (Fig. 5a). In the S94A:NADH structure, part of the substrate binding loop is more open (197–203) whereas part is more closed (205–214). The conformational change of InhA:NAD^+:substrate ternary complex is due to the geometry of the α-helix of the longer substrate binding loops that could sterically hinder the loop from folding downward into the substrate binding cavity, leaving a larger opening for the fatty acyl substrate between residues 205–214 (Rozwarski et al., 1999). The residues 205–214 of the substrate binding loop moves approximately 6 Å away from the active site in the unliganded InhA, while the substrate of InhA:NAD^+:substrate presents the same loops away approximately 4 Å. The opening of the substrate binding loop for these structures are consistent with apoenzyme providing access to substrates to enter the enzyme active site and non-productive ternary complex providing access to solvent for product release. Analysis of molecular surface shows the effect of the substrate binding loop on the active site of InhA (Fig. 7).

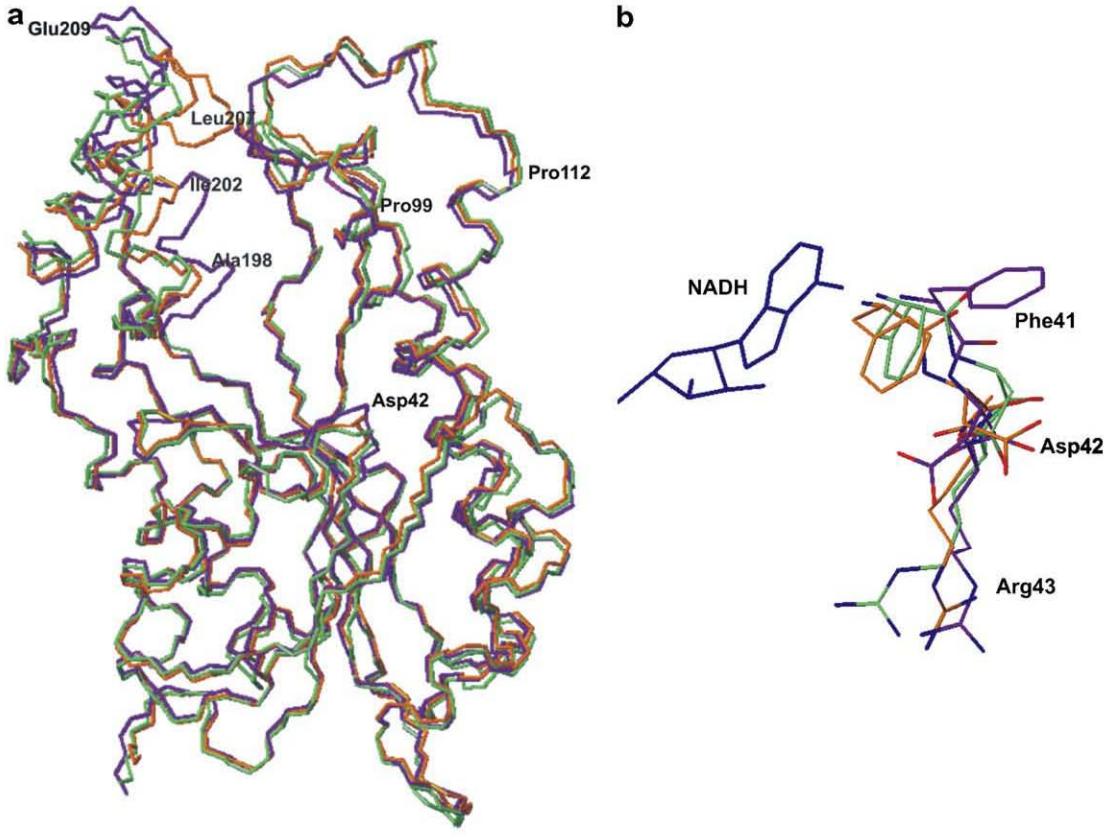


Fig. 5. Conformational changes observed in the structure of S94A InhA upon binding of NADH or both NADH and substrate. (a) Cα traces of S94A InhA unliganded (purple), S94AInhA in complex with NADH (orange), and WT InhA in complex with both NAD^+ and substrate (green). (b) Conformational changes of residues 41 43 (coloring was the same as in a). The figures were generated using program MolMol (Koradi et al., 1996). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

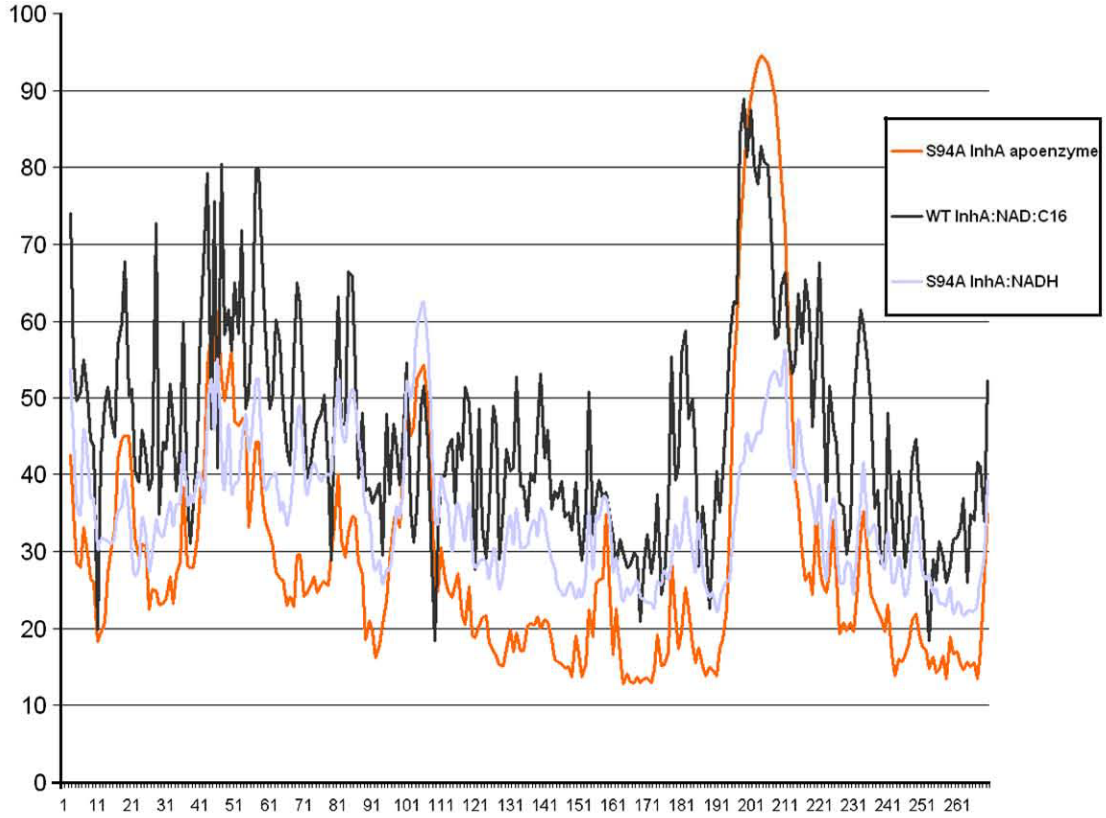


Fig. 6. Residue-averaged *B*-factor for S94A resistant mutant apoenzyme (red line), WT InhA in complex with NAD^+ and C16 substrate (black line), and S94A resistant mutant in complex with NADH (purple line). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

In this manner, the binding of NADH in the active site of the protein can cause the partial closure of substrate binding loop, however, arranging a conformation for the binding of the substrate. In the presence of the cofactor and substrate the protein presents the active site newly opened for liberation of the product of the reaction. Thus, in according with this hypothesis the molecule of NADH is the main responsible by closure of the active site of InhA.

The Van der Waals interactions between the Phe41 and NADH molecule may play in NADH binding to the enzyme active site, since Phe41 moves closer to the adenine moiety upon cofactor binding (Fig. 5b). Asp42 and Arg43 side chains also undergo conformational changes upon binding of NADH or both NADH and substrate (Fig. 5b). However, the absence of substrate does not seem to alter the oligomeric state of InhA because the tetrameric form in the asymmetric unit and the contact area among the monomers in the unliganded structure are similar to the other structures (InhA:NAD^+:substrate ternary complex and InhA:NADH binary complex) (data not show). Thus, the movement of substrate binding loop appears not to interfere with the protein oligomeric state.

4. Conclusion

Here, we present four structures of InhA from *M. tuberculosis*: three structures for complexes of InhA:INH–NAD and one in the unliganded form. The structures in complex with INH–NAD are: INH-sensitive WT and two INH-resistant mutants (I21V and S94A). Comparison between our WT InhA:INH–NAD structure and the structure previously determined by Rozwarski et al. (1998) reveals that there are changes in important residues in the active site that were not previously observed. Moreover, the comparison of our structures in complex with INH–NAD shows that there is no large influence of mutation in the binding of INH–NAD. The INH–NAD adduct has been shown to be a slow, tight-binding competitive inhibitor of WT and INH-resistant InhA mutant enzymes (Rawat et al., 2003). Interestingly, the kinetic and thermodynamic parameters for the interaction of isonicotinyl-NAD^+ adduct with INH-resistant I21V, I47T, and S94A InhA mutant enzymes were found to be similar to those of the WT enzyme (Rawat et al., 2003). These results prompted the authors to suggest an alternative hypothesis to explain for INH resistance mechanism in strains harboring *inhA*-structural gene mutations, in which InhA may interact directly with other components of the FAS-II system. Accordingly, the resistance-associated mutations in the *inhA*-structural gene would affect the susceptibility of InhA to INH inhibition only in the context of the multienzyme complex, and not when InhA is tested in isolation as in *in vitro* assays. Several protein–protein interactions between FAS-II enzymes have been detected by yeast

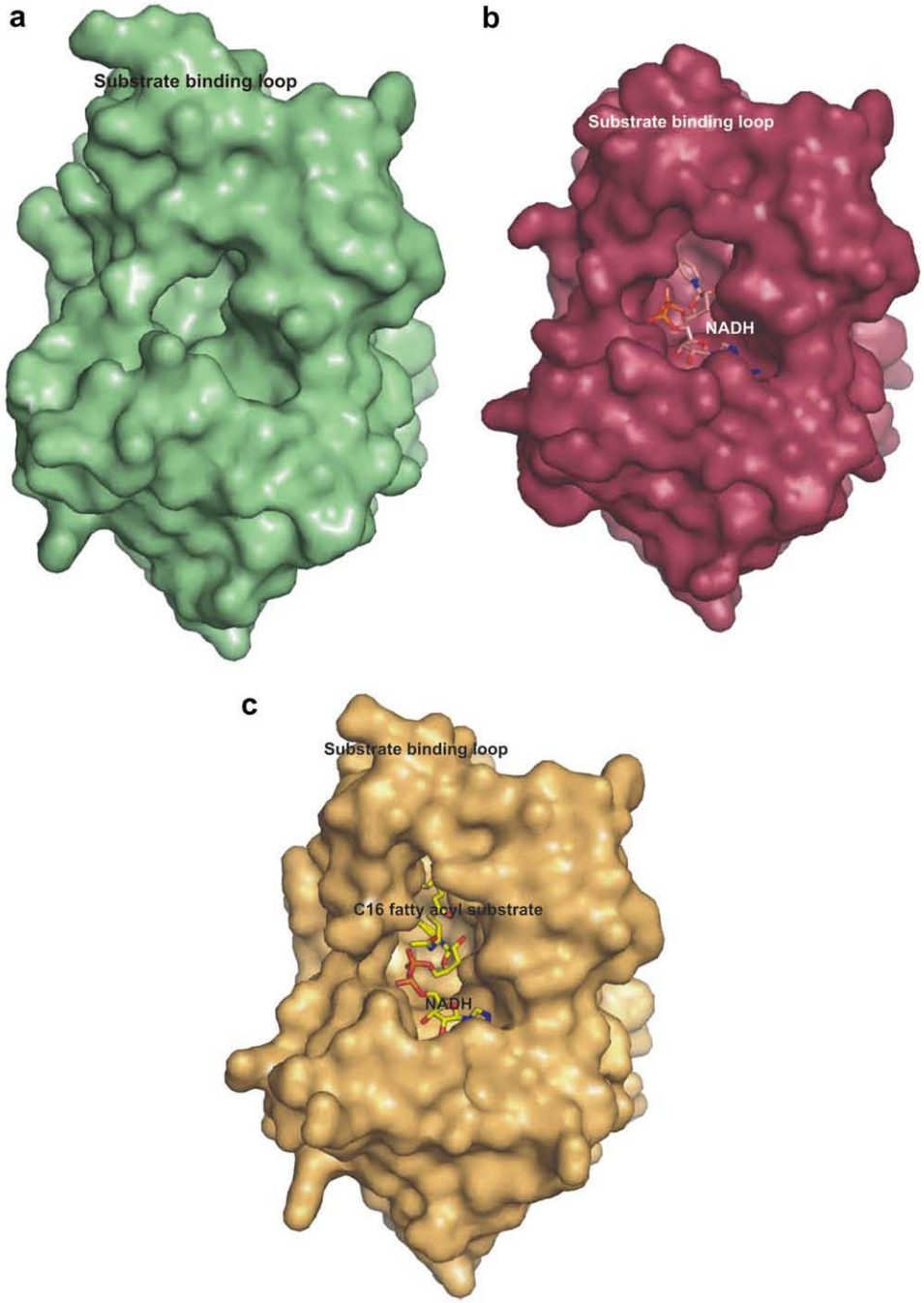


Fig. 7. Molecular surface for (a) Unliganded S94A InhA, (b) S94A InhA in complex with NADH, and (c) WT InhA in complex with both NADH and substrate. The figures were generated using program Pymol (DeLano, 2004).

two-hybrid and co-immunoprecipitation studies and proposed that either these complexes might coexist or the quaternary structure of a "unique" FAS-II might change from one composition to another during the time and according to the degree of elongation of the substrate (Veyron-Churlet et al., 2004). In particular, *M. tuberculosis* InhA was shown to interact with KasA (β-ketoacyl synthase A) and this protein–protein interaction has been suggested as a probable explanation to occurrence of INH-resistant mutant in KasA, even if InhA is indeed the only primary target of INH. However, it remains to be shown whether *inhA* structural gene mutations identified in INH-clinical isolates of *M. tuberculosis* will affect the inhibition of InhA by INH in the context of, for instance, InhA–KasA multienzyme complex. In agreement with InhA as the primary target for INH mode of action, recessive mutations in *Mycobacterium smegmatis* and *Mycobacterium bovis* BCG *ndh* gene, which codes for a type II NADH dehydrogenase (NdhII), have been found to increase intracellular NADH/NAD ratios (Vilchèze et al., 2005). Increasing NADH levels protected InhA against inhibition by the INH–NAD adduct formed upon KatG activation of INH. Hence,

mutations in mycobacterial *ndh* gene resulted in increased intracellular NADH concentrations, which competitively inhibits the binding of INH–NAD adduct to InhA, in agreement with the higher dissociation constant values for NADH found for INH-resistant clinical isolates harbouring *inhA*-structural gene mutations as compared to WT InhA (Basso et al., 1998). Moreover, it has been shown that subtle structural changes result in increased values for the limiting dissociation rate constant for NADH from INH-resistant mutants (Oliveira et al., 2006). These observations are in agreement with the results described here.

The unliganded structure presented here shows conformational changes that occur upon binding of substrates (NADH and/or fatty acyl substrate). The structure presents movements in the substrate binding loop and a striking change in the position of Phe41, which may play an important role in NADH stabilization in the enzyme active site. NADH or substrate binding causes no apparent change in the oligomeric state of InhA.

The results presented here provide structural information that should better our understanding of the mechanism of action of, and resistance to, isoniazid in *M. tuberculosis*. The unliganded structure of InhA allows identification of conformational changes upon ligand binding that are important for substrate anchoring. We hope that these results will help structure-based drug design of more potent antimycobacterial agents.

5. PDB Accession codes

Protein Data Bank: Atomic coordinates and structure factors have been deposited with Accession codes: 2idz, 2ie0, 2ieb, and 2ied.

Acknowledgments

Financial support for this work was provided by Millennium Initiative Program MCT-CNPq, Ministry of Health-Department of Science and Technology (DECIT)-UNESCO (Brazil), and PRONEX/CNPq/FAPERGS (Brazil) to D.S.S. and L.A.B. This work was also supported by grants from FAPESP (SMOLBNet, proc. 01/07532-0, 03/12472-2, 04/00217-0) to W.F.A. D.S.S. (304051/1975-06), L.A.B. (520182/99-5), and WFA (CNPq, 300851/98-7) are research career awardees from National Research Council of Brazil (CNPq).

References

Banerjee, A., Dubnau, E., Quémard, A., Balasubramanian, V., Um, K.S., Wilson, T., Collins, D., de Lisle, G., Jacobs Jr., W.R., 1994. *inhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. Science 263, 227 230.

Basso, L.A., Blanchard, J.S., 1998. Resistance to antitubercular drugs. Adv. Exp. Med. Biol. 456, 115 144.

Basso, L.A., Santos, D.S., 2005. Drugs that inhibit mycolic acid biosynthesis in *Mycobacterium tuberculosis* an update. Med. Chem. Ver. online 2, 393 413.

Basso, L.A., Zheng, R., Blanchard, J.S., 1996. Kinetics of inactivation of WT and C243S mutant of *Mycobacterium tuberculosis* enoyl reductase by activated isoniazid. J. Am. Chem. Soc. 118, 11301 11302.

Basso, L.A., Zheng, R., Musser, J.M., Jacobs Jr., W.R., Blanchard, J.S., 1998. Mechanisms of isoniazid resistance in *Mycobacterium tuberculosis*: enzymatic characterization of enoyl reductase mutants identified in isoniazid-resistant clinical isolates. J. Infect. Dis. 178, 769 775.

Bernstein, J.W., Lott, A., Steinberg, B.A., Yale, H.L., 1952. Chemotherapy of experimental tuberculosis. V. Isonicotinic acid hydrazide (nydrazid) and related compounds. Am. Rev. Tuberc. 65, 357 374.

Blanchard, J.S., 1996. Molecular mechanisms of drug resistance in *Mycobacterium tuberculosis*. Ann. Rev. Biochem. 65, 215 239.

CDC (Centers for Disease Control and Prevention), 2006. Emergence of *Mycobacterium tuberculosis* with extensive resistance to second-line drugs worldwide, 2000 2004. Morb. Mortal. Wkly. Rep. 55, 301 305.

Collaborative Computational Project, Number 4., 1994. The CCP4 suite: programs for protein. Acta Crystallogr. D 50, 760 763.

Corbett, E.L., Watt, C.J., Walker, N., Maher, D., Williams, B.G., Raviglione, M.C., Dye, C., 2003. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. Arch. Intern. Med. 163, 1009 1021.

DeLano, W.L., 2004. The PyMOL Molecular Graphics System. CA DeLano Scientific.

Dessen, A., Quémard, A., Blanchard, J.S., Jacobs Jr., W.R., Sacchettini, J.C., 1995. Crystal structure and function of the isoniazid target of *Mycobacterium tuberculosis*. Science 267, 1638 1641.

Espinal, M.A., 2003. The global situation of MDR-TB. Tuberculosis 83, 44 51.

Glickman, M.S., Jacobs Jr., W.R., 2001. Microbial pathogenesis of *Mycobacterium tuberculosis*: dawn of a discipline. Cell 104, 477 485.

Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18, 2714 2723.

Johnsson, K., King, D.S., Schultz, P.G., 1995. Studies on the mechanism of action of isoniazid and ethionamide in the chemotherapy of tuberculosis. J. Am. Chem. Soc. 117, 5009 5010.

Johnsson, K., Schultz, P.G., 1994. Mechanistic studies of the oxidation of isoniazid by the catalase peroxidase from *Mycobacterium tuberculosis*. J. Am. Chem. Soc. 116, 7425 7426.

Koradi, R., Billeter, M., Wüthrich, K., 1996. MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graph. 14 (1), 51 55.

Kremer, L., Dover, L.G., Morbidoni, H.R., Vilchèze, C., Maughan, W.N., Baulard, A., Tu, S.C., Honore, N., Deretic, V., Sacchettini, J.C., Locht, C., Jacobs Jr., W.R., Besra, G.S., 2003. Inhibition of InhA activity, but not KasA activity, induces formation of a KasA-containing complex in mycobacteria. J. Biol. Chem. 278, 20547 20554.

Kuo, M.R., Morbidoni, H.R., Alland, D., Sneddon, S.F., Gourlie, B.B., Staveski, M.M., Leonard, M., Gregory, J.S., Janjigian, A.D., Yee, C., Musser, J.M., Kreiswirth, B., Iwamoto, H., Perozzo, R., Jacobs Jr., W.R., Sacchettini, J.C., Fidock, D.A., 2003. Targeting tuberculosis and malaria through inhibition of enoyl reductase: compound activity and structural data. J. Biol. Chem. 278 (23), 20851 20859.

Larsen, M.H., Vilchèze, C., Kremer, L., Besra, G.S., Parsons, L., Salfinger, M., Heifets, L., Hazbon, M.H., Alland, D., Sacchettini, J.C., Jacobs Jr., W.R., 2002. Overexpression of *inhA*, but not kasA, confers resistance to isoniazid and ethionamide in *Mycobacterium smegmatis*, *M. bovis* BCG and *M. tuberculosis*. Mol. Microbiol. 46, 453 466.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283 291.

Lei, B., Wei, C.-J., Tu, S.-C., 2000. Action mechanism of antitubercular isoniazid. J. Biol. Chem. 275, 2520 2526.

McRee, D.E., 1999. XtalView/Xfit a versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. 125, 156 165.

Middlebrook, G., Cohn, M.L., 1953. Some observations on the pathogenicity of isoniazid-resistant variants of tubercle bacilli. Science 118, 297 299.

Mitchison, D.A., 1985. The action of antituberculosis drugs in short-course chemotherapy. Tubercle 66, 219 225.

Morlock, G.P., Metchock, B., Sikes, D., Crawford, J.T., Cooksey, R.C., 2003. ethA, inhA, and katG loci of ethionamide-resistant clinical *Mycobacterium tuberculosis* isolates. Antimicrob. Agents Chemother 47, 3799 3805.

Murshudov, G.N., Vagin, A.A., Dodson, E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D 53, 240 255.

Navaza, J., 1994. AMoRe: an automated package for molecular replacement. Acta Crystallogr. A 50, 157 163.

Anon, 2006. Addressing the threat of tuberculosis caused by extensively drug-resistant *Mycobacterium tuberculosis*. Wkly. Epidemiol. Rec. 21, 386 390.

Oliveira, J.S., Pereira, J.H., Canduri, F., Rodrigues, N.C., de Souza, O.N., de Azevedo Jr., W.F., Basso, L.A., Santos, D.S., 2006. Crystallographic and pre-steady-state kinetics studies on binding of NADH to wild-type and isoniazid-resistant enoyl-ACP(CoA) reductase enzymes from *Mycobacterium tuberculosis*. J. Mol. Biol. 359 (3), 646 666.

Oliveira, J.S., Vasconcelos, I.B., Moreira, I.S., Santos, D.S., Basso, L.A., 2007. Enoyl reductases as targets for the development of anti-tubercular and antimalarial agents. Curr. Drug Targets 8 (3), 399 411.

Ormerod, L.P., 2005. Multidrug-resistant tuberculosis (MDR-TB): epidemiology, prevention and treatment. Br. Med. Bull. 73 74 (1), 17 24.

Pablos-Mendez, A., Gowda, D.K., Frieden, T.R., 2002. Controlling multidrug-resistant tuberculosis and access to expensive drugs: a rational framework. Bull. World Health Organ. 80, 489 495.

Parikh, S., Moynihan, D.P., Xiao, G., Tonge, P.J., 1999. Roles of tyrosine 158 and lysine 165 in the catalytic mechanism of InhA, the enoyl-ACP reductase from *Mycobacterium tuberculosis*. Biochemistry 38, 13623 13634.

Polikarpov, I., Perles, L.A., de Oliveira, R.T., Oliva, G., Castellano, E.E., Garratt, R.C., Craievich, A., 1998. Set-up and experimental parameters of the protein crystallography beam line at the Brazilian National Synchrotron Laboratory. J. Synchrotron Radiat. 5, 72 76.

Quémard, A., Sacchettini, J.C., Dessen, A., Vilcheze, C., Bittman, R., Jacobs Jr., W.R., Blanchard, J.S., 1995. Enzymatic characterization of the target for isoniazid in *Mycobacterium tuberculosis*. Biochemistry 34, 8235 8241.

Rawat, R., Whitty, A., Tonge, P., 2003. The isoniazid NAD adduct is a slow, tight-binding inhibitor of InhA, the *Mycobacterium tuberculosis* enoyl reductase: adduct affinity and drug resistance. Proc. Natl. Acad. Sci. USA 100, 13881 13886.

Rozwarski, D.A., Grant, G.A., Barton, D.H.R., Jacobs, W.R., Sacchettini, J.C., 1998. Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*. Science 279, 98 102.

Rozwarski, D.A., Vilchèze, C., Sugantino, M., Bittman, R., Sacchettini, J.C., 1999. Crystal structure of the *Mycobacterium tuberculosis* enoyl-ACP reductase, InhA, in complex with NAD^+ and a C16 fatty acyl substrate. J. Biol. Chem. 274 (22), 15582 15589.

Schroeder, E.K., de Souza, O.N., Santos, D.S., Blanchard, J.S., Basso, L.A., 2002. Drugs that inhibit mycolic acid biosynthesis in *Mycobacterium tuberculosis*. Curr. Pharm. Biotechnol. 3, 197 225.

Uchoa, H.B., Jorge, G.E., Freitas Da Silveira, N.J., Camera Jr., J.C., Canduri, F., De Azevedo Jr., W.F., 2004. Biochem. Biophys. Res. Commun. 325, 1481 1486.

Veyron-Churlet, R., Guerrini, O., Mourey, L., Daffe, M., Zerbib, D., 2004. Protein-protein interactions within the fatty acid synthase-II system of *Mycobacterium tuberculosis* are essential for mycobacterial viability. Mol. Microbiol. 54, 1161 1172.

Vilchèze, C., Wang, F., Arai, M., Hazbon, M.H., Colangeli, R., Kremer, L., Weisbrod, T.R., Alland, D., Sacchettini, J.C., Jacobs Jr., W.R., 2006. Transfer of a point mutation in *Mycobacterium tuberculosis inhA* resolves the target of isoniazid. Nat. Med. 12, 1027 1029.

Vilchèze, C., Weisbrod, T.R., Chen, B., Kremer, L., Hazbon, M.H., Wang, F., Alland, D., Sacchettini, J.C., Jacobs Jr., W.R., 2005. Altered $NADH/NAD^+$ ratio mediates coresistance to isoniazid and ethionamide in mycobacteria. Antimicrob. Agents Chemother. 49, 708 720.

Zabinski, R.F., Blanchard, J.S., 1997. The requirement for manganese and oxygen in the isoniazid-dependent inactivation of *Mycobacterium tuberculosis* enoyl reductase. J. Am. Soc. Chem. 119, 2331 2332.

5.6. New sensitive fluorophores for selective DNA detection. Org Lett. 2007;9(20):4001-04.

ORGANIC LETTERS

2007
Vol. 9, No. 20
4001−4004

New Sensitive Fluorophores for Selective DNA Detection

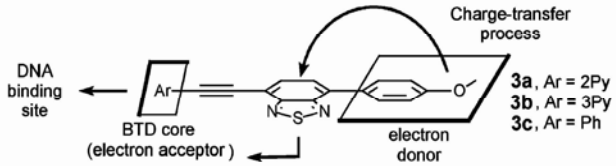
Brenno A. D. Neto,*,† Alexandre A. M. Lapis,‡ Fabiana S. Mancilha,‡ Igor B. Vasconcelos,† Caroline Thum,† Luiz A. Basso,† Diógenes S. Santos,† and Jaïrton Dupont*,‡

Centro de Pesquisas em Biologia Molecular e Funcional (CPBMF), Tecnopuc, PUC-RS, Brazil, and Institute of Chemistry, UFRGS Porto Alegre, RS 90501-970, Brazil

brenno.ipi@gmail.com; dupont@iq.ufrgs.br

Received July 18, 2007

ABSTRACT



4,7-Disubstituted benzothiadiazoles containing 1-arylethynyl and 4-methoxyphenyl groups are selective photoluminescent "light up" probes to duplex DNA with unprecedented sensibility in both spectrophotometric and spectrofluorimetric measurements.

Biosensor technologies that focus on the direct detection of nucleic acids are currently an area of tremendous interest as they play a major role in clinical, forensic, and pharmaceutical applications.[1,2] The molecular probes that cause an increase in both absorbance and emission intensity ("light up" probes) by association with the host biomacromolecules (e.g., DNA, RNA, and proteins) are very useful photoluminescent markers in genomics and proteomics.[3] These simple and straightforward spectroscopic methods are especially advantageous because small organic dyes absorb and emit at wavelengths that do not interfere with the absorption of the DNA bases (λ max $\approx$ 260 nm). Indeed, spectrophotometric and spectrofluorimetric titrations are direct methodologies that indicate the association of a specific dye with DNA.[4]

The ability of planar polycyclic aromatic molecules to intercalate, i.e., to be inserted between two consecutive base pairs of DNA, is the basic molecular geometry of many intercalators that are used in antitumor chemotherapy,[5] as DNA cleavage agents[6] and fluorescent DNA intercalators,[7] and for various other purposes.[8] However, in the case of small molecular organic fluorophores, this association is not very clear, mainly because of the great diversity of the possible resulting structures.[9] Cationic organic dyes normally enhance the propensity of a small probe to bind to DNA, most via an interaction of the positive charge with the phosphate backbone in the double-strand DNA macromolecules.[10,11] However, these charged dyes have some draw-

* To whom correspondence should be addressed. Phone: 55 51 3320 3629 and 55 51 3308 6321.
† CPBMF.
‡ UFRGS Porto Alegre.

(1) (a) Wang, J. *Nucleic Acids Res.* **2000**, *28*, 3011−3016. (b) Niemeyer, C. M.; Blohm, D. *Angew. Chem., Int. Ed.* **1999**, *38*, 2865−2869.
(2) (a) Pu, L. *Chem. Rev.* **2004**, *104*, 1687−1716. (b) Valeur, B. *Molecular Fluorescence: Principles and Applications*; Wiley-VCH: Weinheim, 2002.
(3) Prentø, P. *Biotech. Histochem.* **2001**, *76*, 137−161.
(4) Cantor, C. R.; Schimmel, P. R. *In Biophysical Chemistry*; W. H. Freeman and Co.: San Francisco, 1980; Part II, pp 392−463.
(5) (a) Denny, W. A.; Baguley, B. In *Molecular Aspects of Anti-Cancer Drug-DNA Interaction*; Neidle, S., Warning, M., Eds.; Macmillan: London 1994. (b) Wang, L.; Price, H. L.; Juusola, J.; Kline, M.; Phanstiel, O., IV *J. Med. Chem.* **2001**, *44*, 3682−3691. (c) Lee, Y.-A.; Lee, S.; Cho, T.-S.; Kim, C.; Han, S. W.; Kin, S. K. *J. Phys. Chem. B* **2002**, *106*, 11351−11355.
(6) Viola, G.; Acqua, F. D.; Gabellini, N.; Moro, S.; Vedaldi, D.; Ihmels, H. *ChemBioChem* **2002**, *3*, 550−558.
(7) For a review, see: Erkkila, K. E.; Odom, D. T.; Barton, J. K. *Chem. Rev.* **1999**, *99*, 2777−2795.
(8) Cholody, W. M.; Kosakowka-Cholody, T.; Hollingshead, M. G.; Hairprakasha, H. K.; Michejda, C. J. *J. Med. Chem.* **2005**, *48*, 4474−4481.
(9) Granzhan, A.; Ihmels, H.; Viola, G. *J. Am. Chem. Soc.* **2007**, *129*, 1254−1267.

10.1021/ol701708y CCC: $37.00 © 2007 American Chemical Society
Published on Web 09/05/2007

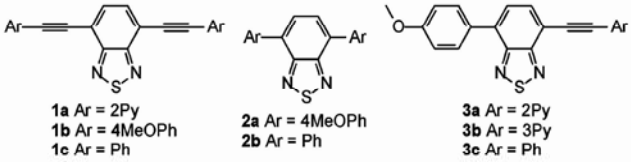
backs especially in electrophoresis purification, due to the concomitant migration of the cation and the anion along the gel. Therefore, the DNA bands in the agarose gel sometimes get poorly resolved as in the case of the commercially available ethidium bromide, which is widely used for DNA visualization in agarose gel electrophoresis.[12] Furthermore, the design of new small molecular organic fluorophores is not a simple task.[9] Higher selectivity, sensitivity, shorter assay times, and greater simplicity in performing the assay are trends that must be taken into account in the design of new intercalators that may be commercially viable.[13]

Neutral and highly polar dyes as DNA intercalators are rare examples,[14] but they can be also used for visualization of biomacromolecules. In this respect, fluorophores such as quinoxalines,[15] benzimidazoles,[16] and 2,1,3-benzothiadiazoles[17] have widespread use in scientific and technological areas,[18] mainly in light technology applications.[19] In this respect, we have recently shown that the photophysical properties of 2,1,3-benzothiadiazoles (BTD) can be tuned by proper choice of the nature of the π-extension in positions 4 and 7.[20−22]

These molecules possess the geometry and electronic properties desirable for DNA duplex intercalators. Here, we demonstrate that new neutral and highly polar BTD are outstanding light up probes for human DNA.

In order to check both electronic and geometrical parameters on the properties of BTD as light up probes to DNA, we have changed the groups attached to the 4,7-BTD core (Chart 1).

Chart 1. Tested BTD Dyes **1a−c**, **2a**,**b**, and **3a−c**



The dyes **1a−c** and **2a,b** were synthesized using a methodology previously described.[20,22] The BTD **3a−c** were synthesized combining both Sonogashira and Suzuki coupling reactions (see the Supporting Information for details).

The photophysical properties of compounds **1a−c**, **2a**,**b**, and **3a−c** were performed in phosphate buffer solutions (100 mM, pH = 7.0), and the results obtained are summarized in Table 1.

Table 1. UV−vis and Fluorescence of the BTD Dyes **1a−c**, **2a**,**b**, and **3a−c** in Phosphate Buffer 100 mM (pH = 7.0)

dye	log ϵ (ϵ)	λ_{abs}^{max} (nm)	λ_{em}^{max} (nm)	Stokes shift (nm)	Φ_f [a]
1a	4.35 (22342)	365	471	106	0.86
1b	3.76 (5790)	429	563	134	0.29
1c	3.81 (6482)	438	552	114	0.37
2a	3.22 (1644)	411	535	124	0.51
2b	3.33 (2150)	367	506	139	0.80
3a	3.97 (9152)	444	544	100	0.40
3b	3.78 (6076)	401	547	146	0.44
3c	4.01 (10208)	426	525	99	0.47

[a] Quantum yield of fluorescence [quinine sulfate in 1 M H_2SO_4, f = 0.55, as standard].

In phosphate buffer (100 mM, pH = 7.0), the lowest energy absorption bands for compounds **1a−c**, **2a**,**b**, and **3a−c** (see the Supporting Information) are assigned to $\pi-\pi^*$ transitions by virtue of their large molar extinction coefficients (log ϵ values in the range of 3.22−4.35). The absorption (λ_{abs}^{max}) and emission (λ_{em}^{max}) maxima lie between 365 and 444 nm and 471−563 nm, respectively. It is worth noting that all dyes (**1a−c**, **2a**,**b**, and **3a−c**) have large Stokes shifts in solution, 99−146 nm, allowing unambiguous detection without reabsorption effects and not interfering with the background fluorescence of biomolecules. These high values also indicate a very efficient intramolecular charge transfer (ICT) in the excited state between the terminal aromatic group (phenyl ring or a methoxyphenyl group) and the BTD moiety.

Spectrophotometric Titrations. All synthesized dyes showed λ_{abs}^{max} in the near-UV region of the spectrum (~390 nm), well separated from the one of the nucleic bases (~260 nm). Negative results were obtained during titration using the systems **2a,b** since it is necessary to use a high concentration of the intercalating agent and high concentrations of DNA. However, all compounds having a C≡C π spacer **1a−c** and especially **3a−c**, were successfully tested as sensitive probes for selective DNA detection (Figure 1(A) and Figures S4 and S6, Supporting Information). This fact indicated the necessity of the triple bound spacer C≡C in order to facilitate the intercalation binding between the dyes and DNA duplex. All compounds (**1a−c**, **2a**,**b**, and **3a−c**) were also tested against human and *Mycobacterium tuberculosis* purine nucleoside phosphorylase (PNP) enzymes with nega-

(10) Medhi, C.; Mitchell, J. B. O.; Price, S. L.; Tabor, A. B. *Biopolymers* **1999**, *52*, 84−93.
(11) Ihmels, H.; Faulhaber, K.; Vedaldi, D.; Dall'Acqua, F.; Viola, G. *Photochem. Photobiol.* **2005**, *81*, 1107−1115.
(12) (a) Luedtke, N.; Liu, Q.; Tor, Y. *Chem. Eur. J.* **2005**, *11*, 495−508. (b) Heller, D. P.; Greenstock, C. L. *Biophys. Chem.* **1994**, *50*, 305.
(13) Wong, E. L. S.; Gooding, J. *J. Anal. Chem.* **2006**, *78*, 2138.
(14) For example, see: Valis, L.; Enthart, E. M.; Wagenknecht, H-A. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3184−3187.
(15) Cui, Y.; Zhang, X.; Jenekhe, S. A. *Macromolecules* **1999**, *32*, 3824−3826.
(16) Gao, Z. Q.; Lee, C. S.; Bello, I.; Lee, S. T.; Wu, S. K.; Yan, Z. L.; Zhang, X. H. *Synth. Met.* **1999**, *105*, 141−144.
(17) (a) Akhtaruzzaman, M.; Tomura, M.; Nishida, J.; Yamashita, Y. *J. Org. Chem.* **2004**, *69*, 2953−2958. (b) Akhtaruzzaman, M.; Tomura, M.; Zaman, M. B.; Nishida, J.; Yamashita, Y. *J. Org. Chem.* **2002**, *67*, 7813−7818.
(18) Lee, M. T.; Yen, C. K.; Yang, W. P.; Chen, H. H.; Liao, C. H.; Tsai, C. H.; Chen, C. H. *Org. Lett.* **2004**, *6*, 1241−1244.
(19) Kraft, A.; Grimsdale, C. A.; Homes, B. A. *Angew. Chem., Int. Ed.* **1998**, *37*, 402−428.
(20) Mancilha, F. S.; Neto, B. A. D.; Lopes, A. S.; Moreira, P. F., Jr.; Quina, F. H.; Gonçalves, R. S.; Dupont, J. *Eur. J. Org. Chem.* **2006**, 4924−4933.
(21) Neto, B. A. D.; Lopes, A. S.; Wust, M.; Costa, V. E. U.; Ebeling, G.; Dupont, J. *Tetrahedron Lett.* **2005**, *46*, 6843−6846.
(22) Neto, B. A. D.; Lopes, A. S.; Ebeling, G.; Goncalves, R. S.; Costa, V. E. U.; Quina, F. H.; Dupont, J. *Tetrahedron* **2005**, *61*, 10975−10982.

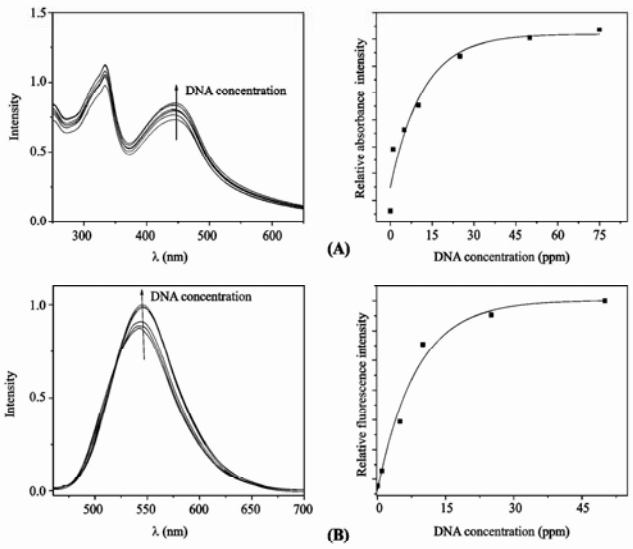


Figure 1. Spectrophotometric (**A**) and Spectrofluorimetric (**B**) titrations of DNA to compound **3a** (10 μM in a phosphate buffer (100 mM) in all measurements) (ppm = ng/mL).

tive results, indicating their sensitivity and selectivity toward DNA.

The addition of DNA (10−100 ppm) to the buffered solutions of the BTD **1a** (50 μM) causes a significant decrease of the absorbance (hypochromic effect) and a blue shift (3−5 nm) of the long wavelength absorption maxima. The exponential decay of the relative absorbance intensity furnishes a quantitative detection of the DNA using BTD **1a**. Note that using spectrophotometric titration, this compound is the only one that acts as a "light off" probe, and this clearly indicates the importance of a donor group (PhOMe) attached to the BTD core.

Compound **1b** (50 μM) is not a good candidate due to the pronounced rotational bands that result in a poorly resolved spectra. Moreover, it indicates that the interaction occurs, increasing the absorbance intensity exponentially upon binding with DNA, and the correlation with the biomacromolecules is almost perfect. Among compounds **1a**−**c**, BTD **1c** showed the best results using spectrophotometric titration. Upon addition of DNA (5−75 ppm) to the buffered solutions of the BTD **1c** (50 μM), a significant increase of the absorbance (hyperchromic effect) and a red shift (5−12 nm) of the long wavelength absorption maxima was observed.

The dyes **3a**−**c** gave the best results among all compounds (Figure 1(A) and Figures S4 and S6). The molecular architecture designed for these sensitive dyes indicates the efficiency of our strategy of combining in the organic systems a donating group (4-MeOPh) directly attached to the BTD nucleus on one side and the presence of C≡C π spacer on the other.

Compounds **3a**−**c** possess high sensitivity to DNA detection, i.e., down to 1 ppm of DNA (the hyperchromic effect observed were very pronounced, especially observed for compound **3c**). A red shift between 2 and 7 nm of the long wavelength absorption maxima was observed for compound **3a**, 2−5 nm for **3b**, and impressive 5−19 nm for BTD **3c**. To the best of our knowledge, these series (**3a**−**c**) are among the most sensitive light up probes for DNA detection using spectrophotometric titrations reported to date.

Spectrofluorimetric Titrations. Spectrofluorimetric titrations of DNA to the BTD derivatives **1a**−**c**, **2a**,**b**, and **3a**−**c** were performed in an aqueous buffer solution at a ligand concentration of 10 μM. Under these conditions, no precipitation was observed. No significant results were obtained using the BTD systems **2a**,**b**, which indicates that even in a more sensitive technical analysis (fluorescence), the presence of a C≡C π spacer is fundamental to allow the binding of the dye to the DNA duplex. All compounds (**1a**−**c**, **2a**,**b**, and **3a**−**c**) were equally tested against human and *M. tuberculosis* purine nucleoside phosphorylase (PNP) enzymes with negative results indicating once more their sensitivity and selectivity to DNA.

Compounds **1a**−**c** showed interesting results (see Figure S6 in the Supporting Information). In particular, **1b** presents a very different behavior than all other BTD. This compound decomposes upon irradiation at 429 nm. Compounds **1a** and **1c**, however, apparently could be useful as light up probes to DNA detection using fluorimetric titration. In both cases, the dyes used in very low concentration (10 μM) could detect even 1 ppm of DNA in phosphate buffer solutions. Nevertheless, after a few weeks in the stock solution, we observe degradation of compound **1c.** Compound **1a** presented its long wavelength emission maximum below 500 nm (473 nm for **1a**, see Table 1) and therefore not suitable for the detection of fluorescence without distortion by autofluorescence of the cell matrix.[23] However, dyes **3a**−**c** present the best results for fluorescence detection and quantification (Figure 1 (B) and Figures S4 and S6). The insertion of one 4-MeOPh group on the molecular architecture of compounds **3a**−**c** increases the thermal, electrochemical, and excited-state stability of the small fluorescent organic systems.

In all cases, the dyes used in very low concentration (10 μM) could detect even 1 ppm of DNA in phosphate buffer solutions. BTD **3a** showed a significant increase of the fluorescence intensity (hyperchromic effect) and a red shift (2−5 nm) of the long-wavelength emission maximum. The exponential increase of the relative fluorescence intensity enabled us to carry out a quantitative detection of the DNA using BTD **3a** in an almost perfect match as seen in Figure 1.

Compounds **3b** and **3c** presented similar results with a significant increase in the fluorescence intensity and a slight red shift (1−4 nm and 2−3 nm, respectively) of the long wavelength emission maxima. Both fluorescent systems also presented an exponential increase of fluorescence upon increasing the DNA concentration. It is interesting for a comparison that commonly used florescent probes normally possess a detection limit that is equal to or above 10 ppm[24] (see Table T1 in the Supporting Information) such as the widely used ethidium bromide.[24a] In fact, BTD **3a**−**c**

(23) Patonay, G.; Antoine, M. D. *Anal. Chem.* **1991**, *63*, 321A-327A.

fluorescent light up probes, especially, are among the most sensitive probes ever described.

Upon binding with DNA, compounds **3a**−**c** maintain their high Stokes shift, with a red shift after the association with DNA, indicating a more efficient ICT process. If the association proceeded through the electron-donating group, the ICT process would not be expected to be so efficient. All results indicate that the intercalation occurs at the C≡C side of the BTD, and as a consequence, the PhOMe portion would be free to perform the ICT process with the BTD core, as rationalized in Figure 2. Preliminary stopped-flow kinetic

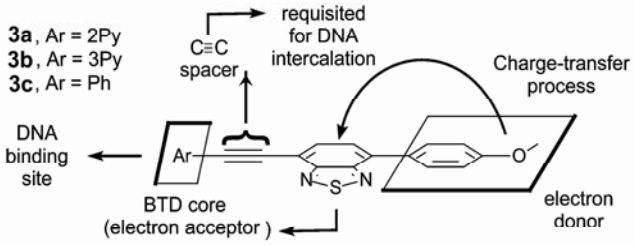


Figure 2. Proposed architecture upon DNA binding.

experiments indicate high binding constants (immediately binding to DNA).

A comparative and qualitative experiment of resolution and sensitivity to DNA was performed with dye **3a**, comparing with the commercially available ethidium bromide (Figure 3). The commercially available DNA plasmid (pCINeo) at different concentrations (200, 100, 50, 20, 10, 5 ng) was submitted in two different agarose gels (1%) to electrophoresis purification using commercial 1 kb plus as a control. It is clear from Figure 3 that the agarose gel A (used with compound **3a**) reveals DNA in the six different concentrations, while agarose gel B (used with ethidium bromide) reveals only the most concentrated DNA channel (200 ng). It is also worth noting that there is no migration of compound **3a**, as commonly occurs using charged dyes such as ethidium bromide. We can also note that using compound **3a**, we were able to reveal the bands with small mass (see the control band in Figure 3), while it is not possible with ethidium bromide (see Supporting Information for **3b**).

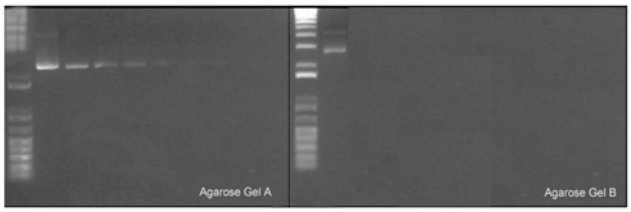


Figure 3. UV irradiation at 360 nm: (A) dye **3a** (1 μL of 1 mM solution) and (B) ethidium bromide (1 μL of 1 mM solution). Gel channels, left to right: control, 1 Kb plus; pCINeo, 200 ng; pCINeo, 100 ng; pCINeo, 50 ng; pCINeo, 20 ng; pCINeo, 10 ng; pCINeo, 5 ng.

In summary, BTDs containing a PhOMe donating group on one side associated with an Ar−C≡C group is a suitable molecular architecture for the spectroscopic selective detection and quantification of DNA. The new dyes are among the most sensitive probes for the DNA detection in solution by spectrophotometric and spectrofluorimetric titrations. A full intercalating model study with completely kinetic and binding model with stopped-flow, surface plasmon resonance, real time PCR, electrochemical, and theoretical approach will be published in due course.

Acknowledgment. We thank CAPES and CNPq for financial support.

Supporting Information Available: Synthetic and titration experimental details. This material is available free of charge via the Internet at http://pubs.acs.org.

OL701708Y

(24) (a) Paigen, K. *Anal. Biochem.* **1980**, *102*, 344−352. (b) Li, W. Y.; Xu, J. G.; Guo, X. Q.; Zhu, Q. Z.; Zhao, Y. B. *Anal. Lett.* **1997**, *30*, 527−536. (c) Li, W. Y.; Xu, J. G., Zhu, Q.; Zhao, Y. *Anal. Lett.* **1997**, *30*, 245−257. (d) Sueda, S.; Ihara, T.; Juskowiak, B.; Takagi, M. *Anal. Chim. Acta* **1998**, *365*, 27−34. (e) Yang, J.; Tong, C.; Jie, N.; Gong, H. *Chin. Biochem. J.* **1996**, *12*, 143. (f) Ci, Y.; Li, Y. Z.; Chang, W. B. *Anal. Chim. Acta* **1991**, *248*, 589−594. (g) Ci, Y. X.; Li, Y. Z.; Liu, X. *J. Anal. Chem.* **1995**, *67*, 1785−1788. (h) Huang, C. Z.; Li, K. A.; Tong, S. Y. *Anal. Lett.* **1996**, *29*, 1705−1717. (i) Huang, C. Z.; Li, K. F.; Tong, S. Y. *Anal. Lett.* **1997**, *30*, 1305−1319. (j) Wu, X.; Yang, J.; Huang, F. *Anal. Lett.* **1999**, *32*, 2417−2425. (l) Liu, R. T.; Yang, J. H.; Wu, X. *J. Lumin.* **2002**, *96*, 201−209. (m) Skripinets, Y. V.; Egorova, A. V.; Ukrainets, I. V.; Antonovich, V. P. *J. Anal. Chem.* **2006**, *61*, 44−51. (n) LePecq, J. B.; Paoletti, C. *Anal. Biochem.* **1966**, *17*, 100.