UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

Tese de doutorado:

A ENZIMA 2-*trans*-ENOIL-ACP(COA) REDUTASE DE *Mycobacterium tuberculosis*: INIBIÇÃO POR UM NOVO COMPOSTO E ESTUDOS ESPECTROSCÓPICOS DO SEU MECANISMO DE RESISTÊNCIA À HIDRAZIDA DO ÁCIDO ISONICOTÍNICO

Doutorando: JAIM SIMÕES DE OLIVEIRA
Orientadores: DIÓGENES SANTIAGO SANTOS
LUIZ AUGUSTO BASSO

Banca examinadora:

Prof. Dr. **Arthur Germano Fett Neto** – Professor e Coordenador do Programa de Pós-Graduação em Biologia Molecular e Celular (PPGBCM) e Professor do Departamento de Botânica da Universidade Federal do Rio Grande do Sul (Porto Alegre).

Prof. Dr. **Luiz Juliano** – Professor do Departamento de Biofísica da Escola Paulista de Medicina, Universidade Federal de São Paulo (SãoPaulo).

Prof. Dr. **João Batista Calixto** – Professor e chefe do Departamento de Farmacologia da Universidade Federal de Santa Catarina (Florianópolis).

Prof. Dr. **Tarso Benigno Ledur Kist** (suplente) - Professor do Programa de Pós-Graduação em Biologia Molecular e Celular (PPGBCM) e Departamento de Biofísica da Universidade Federal do Rio Grande do Sul (Porto Alegre).

ÍNDICE GERAL

- Cloning, overexpression, and purification of functional human purine nuceloside phosphorylase.

- Protective immune response against methicillin resistant *Staphylococcus aureus* in a murine model using a DNA vaccine approach.

- Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*.

- One-step purification of 5-enolpyruvylshikimate-3-phosphate synthase enzyme from *Mycobacterium tuberculosis*.

- Crystal structure of human purine nucleoside phosphorylase at 2.3 Å resolution.

- Molecular models for shikimate pathway enzymes of *Xylella fastidiosa*.

- Crystallographic structure of PNP from *Mycobacterium tuberculosis* at 1.9 Å resolution.

- Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid.

- Electron transfer kinetics and mechanistic study of the thionicotinamide coordinated to the pentacyanoferrate (III)/(II) complexes: a model system for the in vitro activation of thioamides anti-tuberculosis drugs.

- DAHP synthase from *Mycobacterium tubercuosis* H37Rv: cloning, expression, and purification of functional enzyme.

ÍNDICE DE TABELAS

ÍNDICE DE FIGURAS

1. APRESENTAÇÃO DA TESE

A presente tese de doutorado foi desenvolvida no período de abril de 2001 à agosto de 2005 sob orientação dos professores Luiz Augusto Basso e Diógenes Santiago Santos. As seções de introdução, resultados e discussão serão apresentadas na forma de quatro manuscritos que descrevem os resultados experimentais mais relevantes obtidos durante o período do doutorado. A introdução (item 2) corresponde ao manuscrito intitulado “Enoyl reductase as a target for the development of anti-tubercular agents” que foi submetido a revista “Current Drug Targets” em junho do presente ano. Na introdução, são abordados a epidemiologia e quimioterapia da tuberculose, a biossíntese de ácidos micólicos em *M. tuberculosis* salientando o papel da enzima 2-*trans*-enoil-ACP(CoA) redutase (InhA) nesta rota metabólica, e também sua importância como o principal alvo da droga utilizada no tratamento da tuberculose, a hidrazida do ácido isonicotínico (isoniazida). Além disso, são descritos os mecanismos de ação e resistência a isoniazida em *M. tuberculosis*, e os recentes esforços para o desenvolvimento de novos inibidores da enzima enoil redutase de *M. tuberculosis*.

O segundo manuscrito da presente tese (item 3), intitulado “An inorganic iron complex that inhibits wild-type and an isoniazid-resistant mutant 2-*trans*-enoyl-ACP(CoA) reductase from *Mycobacterium tuberculosis*” foi publicado na revista Chemical Communications em 2004. Neste manuscrito são descritos dados de inibição da enzima InhA por um novo composto, o complexo inorgânico pentaciano(isoniazida)ferratoII. Os resultados deste artigo demonstram que este complexo inorgânico pode representar uma nova classe de compostos líderes para o desenvolvimento de agentes anti-tuberculose visando à inibição de um alvo validado.

No terceiro manuscrito são apresentados e discutidos dados que demonstram o mecanismo de inibição da enzima InhA pelo composto pentaciano(isoniazida)ferratoII como sendo do tipo “slow-binding”. No mecanismo de inibição, inicialmente ocorre a rápida formação de um complexo binário entre enzima e inibidor E-I, o qual sofre uma isomerização

unimolecular lenta para um estado E^*-I, onde o inibidor encontra-se mais fortemente ligado a enzima. Este manuscrito foi aceito para publicação na revista “Current Pharmaceutical Design” em junho de 2005.

No quarto e último manuscrito da tese (item 5) são descritos e discutidos os resultados de estudos cristalográficos e de cinética de ligação a NADH em estado pré-estacionário realizados com enzima InhA espécie selvagem e três enzimas mutantes resistentes a isoniazida (I21V, I47T, and S94A). Estas mutações foram identificadas em isolados clínicos de *M. tuberculosis* resistentes a isoniazida por mutação no gene estrutural *inha.* Neste artigo, as principais diferenças estruturais entre os complexos binários das enzimas mutantes e espécie selvagem com NADH são discutidas. Os estudos em estado pré-estacionário demonstram que a principal diferença entre as enzimas está na constante de velocidade limitante para a dissociação do NADH, a qual é maior nas enzimas InhA mutantes. Além disso, um mecanismo para ligação de NADH é proposto. Estes estudos são de extrema importância para uma melhor compreensão dos mecanismos de resistência a isoniazida em *M. tuberculosis*, e também para o desenvolvimento racional de novos inibidores para InhA, que possam vir a ser testados contra cepas sensíveis a isoniazida e também contra cepas resistentes a mesma por mutação no gene estrutural *inha*. Este manuscrito, intitulado “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*” foi submetido ao “Journal of Molecular Biology”.

Por fim, em anexo no item 6, encontram-se todos os artigos publicados durante o período de doutorado, que resultaram de atividades realizadas pelo presente autor em outros projetos.

2. INTRODUÇÃO: “Enoyl reductase as a target for the development of anti-tubercular agents” (Manuscrito submetido para a revista “Current Drug Targets”)

Enoyl reductase as a target for the development of anti-tubercular agents.

J. S. Oliveira[§], I. S. Moreira[†], D. S. Santos[§]* and L. A. Basso[‡]*

[§] Centro de Pesquisas em Biologia Molecular e Functional, Faculdade de Farmácia, Instituto de Pesquisas Biomédicas, Pontifícia Universidade do Rio Grande do Sul, Porto Alegre, RS, Brasil.

[†] Departamento de Química Inorgânica, Universidade Federal do Ceará, Fortaleza, CE, Brasil.

[‡] Centro de Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil.

[*] to whom correspondence may be addressed
LAB: phone, +55-51-33166234; fax, +55-51-33167309; e-mail, labasso@cbiot.ufrgs.br or DSS: phone, +55-51-33203629; fax, +55-51-33203629; e-mail, diogenes@pucrs.br

Key Words: tuberculosis, mycobacteria, mycolic acids, isoniazid, InhA, enoyl reductase, drugs

ABBREVIATIONS USED

ACP = acyl carrier protein

ADR = adenine ribose

BSA = bovine serum albumin

CoA = Coenzyme A

CQ = chloroquine

ENR = enoyl reductase

ETH = ethionamide

FabI = enoyl reductase

FAS-II = Type II dissociated fatty acid biosynthesis system

HPLC = High Performance Liquid Chromatography

INH = isoniazid, isonicotinic acid hydrazide

InhA = 2-*trans* enoyl-ACP (CoA) reductase

KAR = β-ketoacyl reductase

KatG = catalase-peroxidase

KAS = β-ketoacyl synthase

MabA = β-ketoacyl reductase

mAGP = mycolyl-arabinogalactan-peptidoglycan complex

MDG = Millennium Development Goals

MDR-TB = multidrug-resistant tuberculosis, defined as resistant to at least isoniazid and rifampicin

MIC = minimum inhibitory concentration

SI = selectivity index (SI = IC_{50}/MIC)

TB = tuberculosis

TCL = triclosan (2,4,4'-trichloro-2'-hydroxyphenyl ether)

WT = wild type

ABSTRACT

Tuberculosis (TB) is a neglected disease, which continue to be major cause of morbidity and mortality worldwide, killing together around 5 million people each year. Mycolic acids, the hallmark of mycobacteria, are high-molecular-weight α-alkyl, β-hydroxy fatty acids. Biochemical and genetic experimental data have shown that the product of the *M. tuberculosis inhA* structural gene (InhA) is the primary target of isoniazid mode of action, the most prescribed anti-tubercular agent. InhA was identified as an NADH-dependent enoyl-ACP(CoA) reductase specific for long-chain enoyl thioesters and is a member of the Type II fatty acid biosynthesis system, which elongates acyl fatty acid precursors of mycolic acids. *M. tuberculosis* is a target for the development of anti-tubercular agents. Here we present a brief description of the mechanism of action of, and resistance to, isoniazid. In addition, data on inhibition of mycobacterial enoyl reductase by triclosan are presented. We also describe recent efforts to develop inhibitors of *M. tuberculosis* enoyl reductase enzyme activity.

TUBERCULOSIS: EPIDEMIOLOGY AND CHEMOTHERAPY

Tuberculosis (TB) is a global health emergency that remains the leading cause of mortality due to a bacterial pathogen, *Mycobacterium tuberculosis*. The re-emergence 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 [1]. No sustainable control of TB epidemics can thus 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, which translates into more than 4,900 deaths per day, and more than 95 % of these were in developing countries [2]. In addition, approximately 12 % (226,000) of total deaths from TB was attributed to co-infection with *M. tuberculosis* and human immunodeficiency virus (TB-HIV). It has been estimated that approximately 39.4 million people are infected with HIV worldwide, and more than half of them live in sub-Saharan Africa and nearly about a fifth in South and South-East Asia [3]. Immune deficient patients with HIV are at increased risk of latent *M. tuberculosis* infections progressing to active disease, and being transmitted to others [4]. TB and HIV are intricately linked to malnutrition, unemployment, alcoholism, drug abuse, poverty and homelessness. The direct and indirect costs of illness due to TB and HIV has been estimated to be more than 30 % of the annual household income in developing countries and have thus a catastrophic impact on the economy in the developing world [5].

It has been estimated that 3.2 % of the world's new cases of TB, in 2000, were multidrug-resistant tuberculosis (MDR-TB), defined as resistant to at least isoniazid and rifampicin [6]. According to the 2004 Global TB Control Report of the World Health Organization (WHO), there are 300,000 new cases per year of MDR-TB worldwide, and 79 % of MDR-TB cases are "super strains", resistant to at least three of the four main drugs used to treat TB [7]. Localized high incidence rates of MDR-TB have been found in particular regions, which have been defined as hot zones based on either areas where the prevalence of MDR-TB cases is >5 % (that is, where >5 % of current cases are MDR-TB) or areas where the incidence of MDR-TB cases is >5 % (that is, where >5 % of new cases are MDR-TB) [6]. MDR-TB is an airborne bacterium that is spread just as easily as drug-sensitive TB. An individual who is sick with any strain of TB will

infect between 10 and 20 people each year with that same strain [8], thereby making the hot zones of particular concern to public health officials. A mathematical model showed that, paradoxically, areas with programs that successfully reduced wild-type pansensitive strains (as a result of high case detection and treatment rates) often evolved into hot zones [9]. Accordingly, it has been suggested that second-line drugs be quickly introduced to disrupt the amplification of resistance. However, the bacteriostatic second-line drugs are more toxic and less effective and are given for at least three times as long and at 100 times the cost of basic short-course chemotherapy regimens [10, 11]. A mathematical model of the impact of fitness of multidrug-resistant strains of *M. tuberculosis* on the emergence MDR-TB has shown that even when the average relative fitness of MDR strains is low and a well-functioning control program is in place, a small subpopulation of a relatively fit strain may eventually outcompete both the drug-sensitive strains and the less fit MDR strains [12]. These results indicate that current epidemiological measures and short-term trends in the burden of MDR-TB do not provide evidence that MDR-TB strains can be contained in the absence of specific efforts to limit transmission and combat MDR disease. 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 [13].

Although only 5-10 % of infected individuals develop the disease because the host immune response against *M. tuberculosis* is highly effective in controlling bacterial replication, *M. tuberculosis* is almost never eradicated due to its ability to establish and maintain latency, a period during which the infected person does not have clinically apparent TB. Post-primary TB, which is predominantly a pulmonary disease, develops later in life, and can be caused either by reactivation of bacteria remaining from the initial infection or by failure to control a subsequent infection. Post-primary TB involves extensive damage to the lungs and efficient aerosol transmission of bacilli. It has been pointed out that the success of *M. tuberculosis* as a pathogen is largely attributable to its ability to persist in host tissues, where drugs that are rapidly bactericidal *in vitro* require prolonged administration to achieve comparable effects [14]. Thus effective tuberculosis chemotherapy must include early bactericidal action against rapidly growing organisms and subsequent sterilization of the semidormant and dormant populations of

bacilli. The first-line drugs isoniazid, rifampicin, streptomycin, and ethambutol exhibit early bactericidal activity against actively metabolizing bacilli [15]. Pyrazinamide is active against the semidormant bacilli in acidic intracellular environments. The modern, standard "short-course" therapy for TB 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 an additional four months. This combination therapy must be strictly followed to prevent drug resistance and relapse, and direct observation of patient compliance is the most reliable way to ensure effective treatment and prevent the acquisition of resistance. The bacteriostatic second-line drugs (amikacin, kanamycin, capreomycin, cycloserine, *para*-aminosalicylic acid, ethionamide, and fluoroquinolones) are reserved to strengthen the treatment of drug-resistant disease or when bactericidal drugs are prohibited because of toxicity [16].

It has recently been pointed out that there is a rapidly growing capability to undertake health innovation in many developing countries that have high indigenous science and technology capacity but relatively low economic strength, including India, China, Brazil, South Africa, Thailand, Argentina, Malaysia, Mexico and Indonesia [17]. Collectively, these countries invest at least U$ 2.5 billion per year in health research and it has been proposed that they should assume a leadership position in creating health innovations that target diseases of the poor because they are closer to those that are most in need for more effective health products [18]. In addition, to achieve the Millennium Development Goals (MDG), including investments in strategies to halt and reverse the spread of TB, developing country governments should adopt bold development strategies to meet the MDG targets for 2015 [19]. In particular, more effective and less toxic anti-tubercular agents are urgently needed to shorten the duration of current treatment, improve the treatment of MDR-TB, and to provide effective treatment of latent tuberculosis infection.

Mycobacterium tuberculosis FATTY ACID BIOSYNTHESIS AND ENOYL REDUCTASE

The mycobacterial cell wall is comprised of three covalently linked macromolecules: peptidoglycan, arabinogalactan, and mycolic acid, which is often described as mycolyl-arabinogalactan-peptidoglycan complex (mAGP) [20, 21, 22]. Mycolic acids are high-molecular-

weight α-alkyl, β-hydroxy fatty acids (Fig. **1**) that appear mostly as bound esters in tetramycolylpentaarabinosyl cluster in the mycobacterial cell wall. In the pyrolytic cleavage of mycolic acids the intact fatty acid released is referred to as the α branch and the aldehyde released is referred to as the meroaldehyde, and the corresponding segment of the intact mycolate is thus referred to as the meromycolate branch. The meromycolate branch is functionalized at regular intervals by cyclopropyl, α-methyl ketone, or α-methyl methylethers groups, and, as shown more recently, by unsaturations [23].

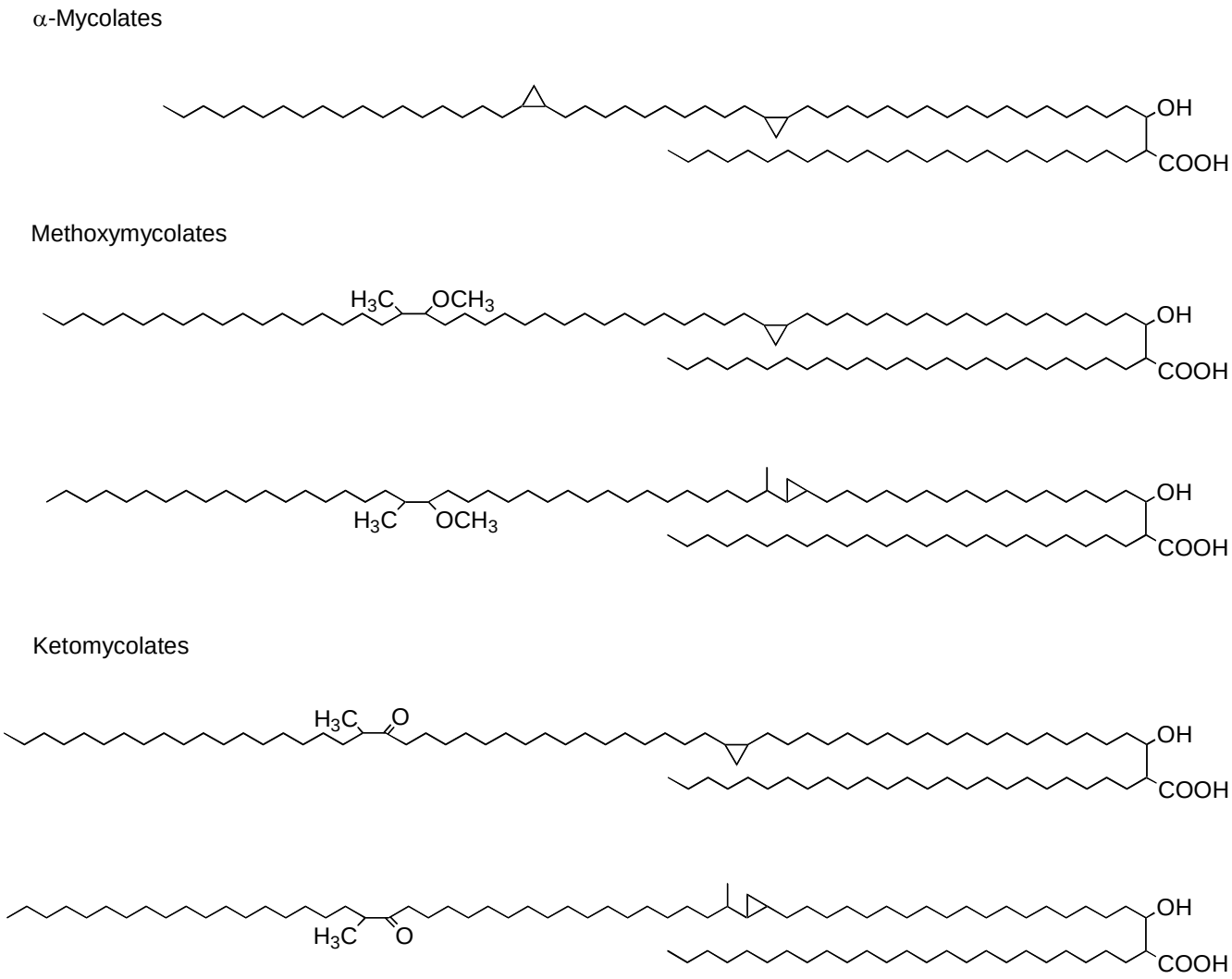


Fig. (1). The structures of mycolic acids identified in *M. tuberculosis*. α-Mycolates: its meromycolate chain contains two *cis*-cyclopropanes. Methoxymycolates: its meromycolate chain

contains an α-methyl methyl-ether moiety in the distal position and a *cis*-cyclopropane or an α-methyl *trans*-cyclopropane in the proximal position. Ketomycolates: its meromycolate chain contains an α-methyl ketone moiety in the distal position and proximal functionalities as in the methoxy series. It should be pointed out that, more recently, unsaturations have been detected in the meromycolate chain of *M. tuberculosis* [23], which are not shown here.

Fatty acid elongation occurs through repetitive cycles of condensation, β-keto reduction, dehydration, and enoyl reduction which are catalyzed by, respectively, β-ketoacyl synthase (KAS, condensing enzyme), β-ketoacyl reductase (KAR; MabA in *M. tuberculosis*), β-hydroxyacyl dehydrase (DE), and enoyl reductase (ENR; InhA in *M. tuberculosis*). These chemical reactions are catalyzed by two types of fatty acid synthase systems (FAS). The FAS-I system is a multidomain polyprotein that encodes all the enzymes necessary for fatty acid synthesis in one large polypeptide and is generally present in most eukaryotes, except in plants [24]. FAS-II systems, which are present in bacteria and plants, catalyze the individual reactions by separate proteins readily purified independently of the other enzymes of the pathway and are encoded by unique genes. Mycobacteria, unlike most organisms, have both FAS-I and FAS-II systems [25]. The mycobacterial FAS-I system catalyzes the synthesis of C_{16} and C_{18} fatty acids, the normal products of *de novo* synthesis, and elongation to produce C_{24} and C_{26} fatty acids [26]. The mycobacterial FAS-II is analogous to other bacterial FAS-II systems, with the notable exception of primer specificity. The mycobacterial FAS-II is not capable of *de novo* synthesis from acetate but instead elongates the “short-chain” acyl-ACP intermediates to fatty acids ranging from 24 to 56 carbons in length [27, 28]. In summary, the mycobacterial FAS-I produces a bimodal ($C_{14:0}$-$C_{16:0}$ to $C_{24:0}$-$C_{26:0}$) distribution of acyl-CoA fatty acids. The mycobacterial FAS-I system would provide the “shorter” acyl-CoA fatty acid precursors ($C_{14:0}$-$C_{16:0}$) for condensation with malonyl-ACP by mtFabH enzyme activity whose products, in turn, would be elongated by the FAS-II system, yielding the long carbon chain of the meromycolate branch (50-60 carbons) of mycolic acids. The longer chain acyl-CoA products ($C_{24:0}$-$C_{26:0}$) of FAS-I would be excluded from chain elongation and remain available to be utilized, presumably in the CoA form, as substrates for formation of the α-alkyl branch (20-26 carbons) of mycolic acids. The

readers are referred to recent reviews describing mycolic acid biosynthesis and their processing into the final products for a more comprehensive description of these processes [22, 29, 30, 31]. Isoniazid (INH, isonicotinic acid hydrazide; Fig. **2**) is one of the oldest synthetic antitubercular, and the most prescribed drug for active infection and prophylaxis. INH is highly active against *M. tuberculosis* with an MIC value in the range of 0.02 to 0.2 μg mL^{-1}, and is active against growing tubercle bacilli but not resting bacilli. Isoniazid uptake in *M. tuberculosis* has been proposed ro occur by passive diffusion [32]. However, the involvement of efflux pumps in isoniazid transport has recently been reported [33]. The mechanism of action of isoniazid is complex, as mutations in a number of different genes have been found to correlate with isoniazid resistance in *M. tuberculosis* [22, 34, 35].

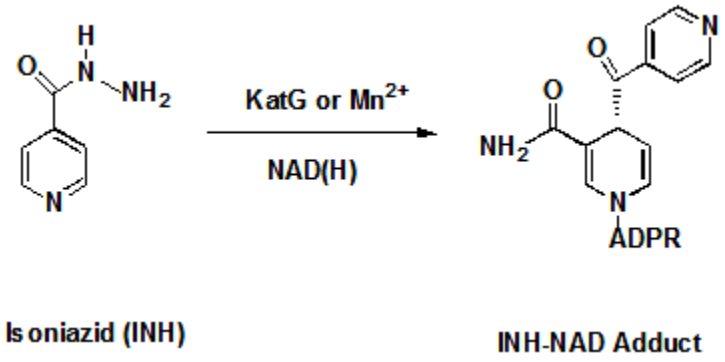


Fig. (2). Chemical structure of isoniazid and INH-NAD adduct that inhibits InhA enzyme activity.

The product of the *M. tuberculosis inhA* structural gene (InhA or ENR) has been shown to be the primary target for isoniazid [36]. InhA was identified as an NADH-dependent 2-*trans* enoyl-ACP (acyl carrier protein) reductase enzyme that exhibits specificity for long-chain enoyl thioester substrates [37]. InhA is a member of the mycobacterial FAS-II, which elongates acyl fatty acid precursors yielding the long carbon chain of the meromycolate branch of mycolic acids, the hallmark of mycobacteria [22]. Consistent with InhA as the primary target of INH mode of action, inactivation of *M. smegmatis inhA*-encoded enoyl reductase and INH treatment resulted in similar mycolic acid biosynthesis inhibition, and morphological changes to the mycobacterial cell wall leading to cell lysis [38]. Transformation of *M. smegmatis*, *M. bovis* BCG and three different strains of *M. tuberculosis* with multicopy plasmids expressing *inhA*

genes conferred a 20-fold resistance to INH and a 10-fold resistance to ethionamide (ETH) [39]. More recently, further biochemical and genetic evidence has been provided showing that InhA is the primary target of INH [40]. In agreement with these results, mutations in the *inhA* structural gene and in the *inhA* locus promoter region have been associated with isoniazid resistance [41]. Moreover, 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 values for NADH than WT InhA, whereas there were only modest differences in the steady-state parameters [41]. Consistent with these results, a comparison of the crystal structure of binary complex of WT and INH-resistant mutant InhA (S94A) in complex with NADH [42] showed that disruption of a hydrogen bond network in the mutant protein could account for higher dissociation constant value for the coenzyme. In WT InhA-NADH complex structure (2.2 Å) a well-ordered water molecule mediates two hydrogen bonds between the O2 of the P_N phosphate of NADH and protein that are lost in the S94A mutant protein [42]. This water molecule (WAT1) hydrogen bonds to the hydroxyl group of Ser94, the main-chain oxygen of Gly14, and the main-chain nitrogen atoms of Ala22 e Ile21. In the S94A-NADH complex structure (2.7 Å) the carbonyl group of residue Gly14 rotates away from the water molecule, breaking the hydrogen bond observed in wild-type structure (Fig. **3A**). A computational comparison of 102 high-resolution structures ($\leq$ 1.90 Å) of enzyme-dinucleotide complexes revealed that the WAT1 water is structurally conserved in proteins that exhibit the classic Rosmann dinucleotide-binding fold motif [43]. The conserved water molecule links, through a conserved hydrogen-bonding pattern, the glycine-rich phosphate-binding loop with the dinucleotide pyrophosphate moiety [44]. In enoyl reductase enzymes the phosphate-binding loop consensus sequence has been proposed to be GXXXXXXXG(A) (residues Gly14-Ala 22 in InhA) [43]. Typically, the WAT1 water makes four hydrogen bonds and two of them are invariant. The invariable hydrogen bonds involve the dinucleotide pyrophosphate moiety and main-chain amine of the last conserved Gly of phosphate binding loop (Ala22 in InhA). Moreover, the interaction of pyrophosphate has been shown to be stereospecific. Almost without exception, the pyrophosphate atom that interacts with the structurally conserved water is O2N in the case of NAD- or NADP-binding proteins. In structures with the sequence pattern GXXXXXXXG(A), the partners for the other two hydrogen bonds involves the carbonyl group of the first conserved Gly (Gly14 in InhA) and a C-terminal

residue of β4 strand (Ser94 in InhA) [43]. In our efforts to understand the structural basis of isoniazid resistance mechanism of InhA, we have solved the crystal structures WT InhA-NADH and S94A InhA-NADH binary complexes to a better resolution than previously reported [42], which were refined to, respectively, 1.9 Å and 2.3 Å. The NADH binding pocket analysis showed that in the S94A InhA-NADH structure there was no rotation of the carbonyl group of Gly14 and only the conserved hydrogen bond of WAT1 water molecule with hydroxyl group of Ser94 is lost owing to mutation to Ala94 (Fig **3C** and **D**). Thus, this hydrogen bond disruption probably is the main cause of a sixty-fold reduction in affinity for NADH observed for the S94A InhA protein, which corresponds to approximately 2.5 kcal mol^{-1} [41]. This is in agreement with estimates for energy hydrogen bonds which are in the range of 3 to 9 kcal mol^{-1} [45]. In addition, this WAT1 water has been proposed to be an integral characteristic of dinucleotide-binding Rossmann fold domains that both contributes significantly to dinucleotide recognition and presumably provides a favorable enthalpic contribution to the free energy of binding [43]. More recently, molecular dynamics simulations of fully solvated WT and INH-resistant clinical isolates of *M. tuberculosis* in complex with NADH showed that mutations of the glycine-rich loop residues I21V and I16T resulted in a change in the pattern of direct hydrogen bond contacts with the pyrophosphate moiety of NADH [46]. The NADH pyrophosphate moiety undergoes considerable conformational changes, reducing its interactions with InhA binding site and probably indicating the initial phase of ligand expulsion in which NADH moves apart from its binding site [46].

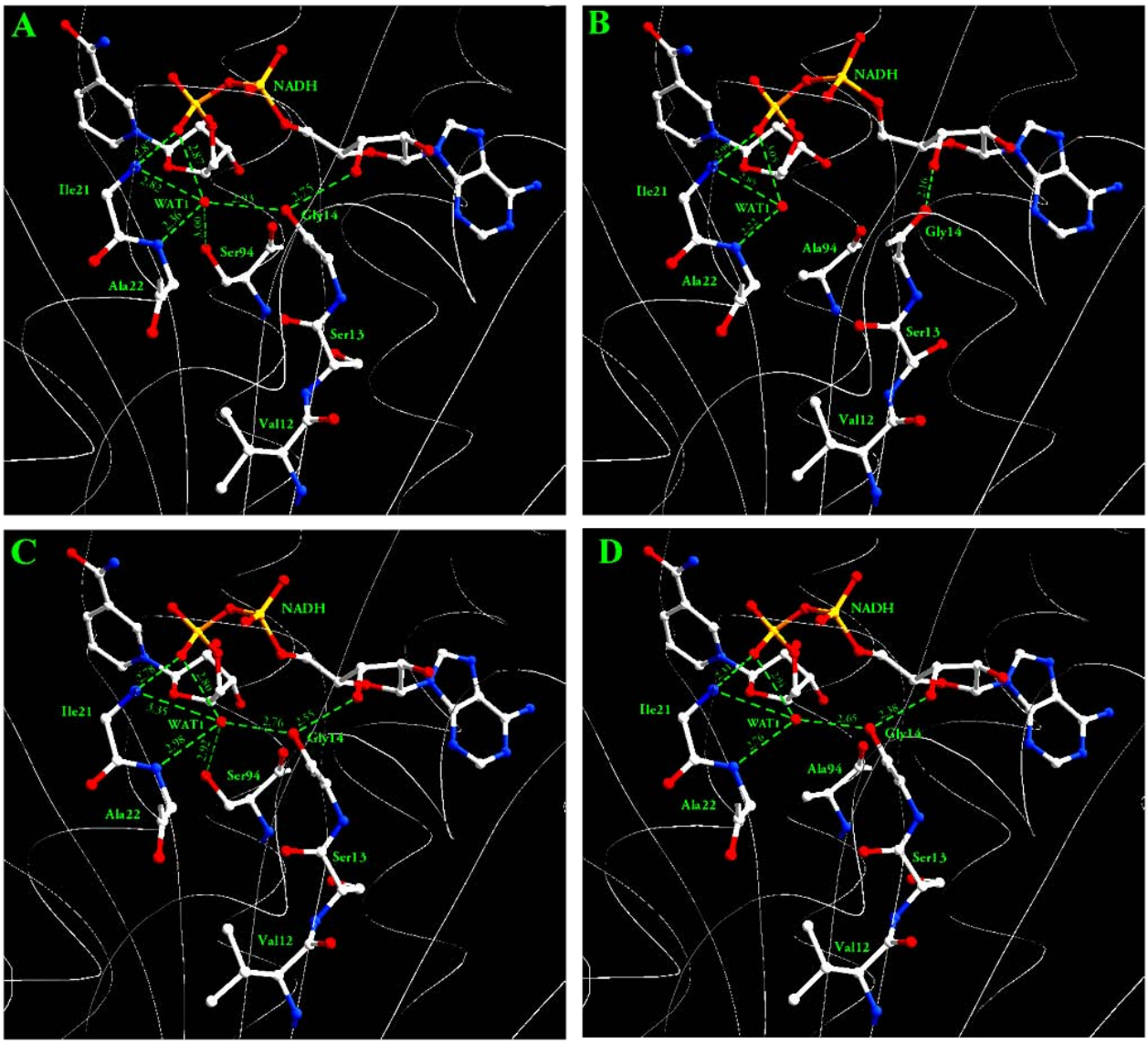


Fig. (3). Differences in hydrogen bond network between the crystal structures of WT and INH-resistant mutant InhA (S94A) complexed with NADH.

Top: Crystal structures of WT InhA-NADH (**A**) and S94A InhA-NADH (**B**) binary binary complexes refined to, respectively, 2.2 Å and 2.7 Å [42].

Bottom: Crystal structures of WT InhA-NADH (**C**) and S94A InhA-NADH (**D**) binary complexes solved to a better resolution than previously reported [42], which were refined to, respectively, 1.9 Å and 2.3 Å. Only the conserved hydrogen bond between the WAT1 water and hydroxyl group of residue S94 is lost owing to mutation to Ala94. Both structures were refined with X-PLOR-NIH program (47, 48). The completeness values for all data were 97.7% and

95.5% for, respectively, WT InhA and S94A InhA proteins. The final R_{Factor} and R_{Free} values calculated for structure of WT InhA-NADH complex were, respectively, 21.5% and 25.1%. For the structure of S94A InhA-NADH complex, the final R_{Factor} and R_{Free} values were, respectively, 20.8% and 27.4%. Crystals for both wild type and mutant were hexagonal, space group $P6_222$, with on molecule per assymetric unit. The figures were prepared using the program Swiss-PDBViewer v3.7 (www.expasy.org/spdbv) [49].
For clarity, only residues Val12, Ser13, Gly14, Ile21, Ala22, and WAT1 water and NADH molecules are shown. Direct and water-mediated hydrogen bonds between the protein and NADH pyrophosphate moiety are represented as broken green lines. The interactions distances are in Å. The atoms are colored white for carbon, red for oxygen, blue for nitrogen, and yellow for phosphorus. The InhA strucures are represented as thin-line ribbons colored gray.

The increase in the NADH dissociation constant observed for the I21V and I16T isoniazid-resistant mutants [41] may thus be attributed to a decrease in the number of H-bond interactions between NADH and aminoacids in the binding pocket and between the cofactor molecule and water molecules that mediate interactions with the enzyme [46]. In agreement, estimates of NADH free energy of binding by molecular docking experiments showed that the WT enzyme has higher affinity for NADH than the mutant enzymes [46]. The correlation between the NADH binding properties in solution, molecular dynamics simulation and docking studies and resistance to INH inactivation of *M. tuberculosis* strains harbouring *inhA*-structural gene mutations provides a mechanism of resistance at molecular level for this clinically important drug. Accordingly, current experimental data point to InhA as the primary target for INH mode of action in *M. tuberculosis*.

Approximately 50 % of INH-resistant clinical isolates of *M. tuberculosis* harbour deletions of, or missense mutations in, the *katG* gene that codes for a catalase-peroxidase enzyme [35, 50]. The S315T KatG mutant enzyme, which is one of the most commonly encountered substitutions in clinical INH-resistant strains, has been shown by isothermal titration calorimetry to have reduced affinity for INH as compared to WT KatG enzyme [51]. These authors have shown that although S315T KatG maintains reasonably good steady-state catalytic rates, poor binding of INH to the mutant enzyme would reduce INH activation and bring about drug

resistance [51]. INH is a pro-drug that is activated by the mycobacterial *katG*-encoded catalase-peroxidase enzyme in the presence of manganese ions, NAD(H) and oxygen [52, 53, 54, 55]. Data from X-ray crystallography and mass spectrometry showed that the KatG-produced acylpyridine fragment of isoniazid binds covalently to the C4 of the nicotinamide ring of NADH (Fig. **2**), forming a binary complex with InhA [56]. This isonicotinyl-NAD^+ adduct has been shown to bind to InhA with a dissociation constant value lower than 0.4 nM [57]. This adduct has recently been shown to be a slow, tight-binding competitive inhibitor of WT InhA [58]. The initial rapidly reversible weak binding of isonicotinyl-NAD^+ adduct to WT InhA (K_i = 16 nM) is followed by a slow isomerization of the enzyme-inhibitor binary complex with an overall dissociation constant (K_i*) value of 0.75 nM. 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 wild-type enzyme [58]. These results prompted the authors to suggest an alternative hypothesis to explain for INH resistance mechanism in strains harboring *inhA*-structural gene mutations. A high molecular weight FAS-II system (fatty acid synthase system II) that catalyzes the ACP-dependent synthesis of long chain fatty acids and that contains both enoyl reductase (ENR) and β-ketoacyl reductase (KAR) activities has been purified from *M. smegmatis* [59]. The *mabA* (*fabG1*)-encoded β-ketoacyl-ACP (CoA) reductase (MabA or KAR) has been shown to be a member of FAS-II in mycobacteria [60]. Based on these findings, it has been proposed that mycobacterial ENR (InhA in *M. tuberculosis*) may interact directly with other components of the FAS-II system [58]. 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 mycobacterial ENR (InhA) is tested in isolation as in *in vitro* assays. More recently, yeast two-hybrid and co-immunoprecipitation approaches have been employed to study protein-protein interactions between known components of FAS-II system [61]. These authors reported several types of FAS-II complexes 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. 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 [61]. However, it remains to be

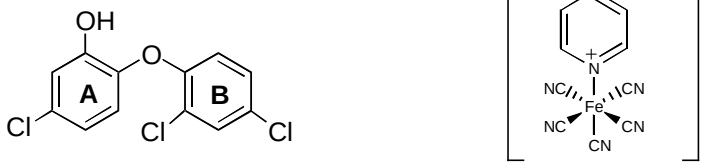
shown whether *inhA* structural gene mutations identified in INH-clinical isolates of *M. tuberculosis* will affect the inhibition of InhA by INH in the context of, for instance, InhA-KasA multienzyme complex. In agreement with InhA as the primary target for INH mode of action, recessive mutations in *M. smegmatis* and *M. bovis* BCG *ndh* gene, which codes for a type II NADH dehydrogenase (NdhII), have been found to increase intracellular NADH/NAD ratios [62]. 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 [62], 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 [41].

The INH mechanism of action requires KatG-catalyzed conversion of INH into a number of electrophilic intermediates [52]. Formation of the isonicotinyl-NAD^+ adduct has been proposed to be through addition of either an isonicotinic acyl anion to NAD^+ or an isonicotinic acyl radical to an $NAD^{\cdot}$ radical [56]. Using a combination of spectroscopic, biochemical and computational techniques, an enzyme-catalyzed mechanism for INH activation has recently been proposed, leading to isonicotinoyl radical formation (thought to be the activated form of INH) and production of the amide end product via a diazene intermediate [63]. In the proposed mechanism, the oxyferryl group of compound I of KatG, generated after reaction with peroxide, is reduced by INH in a single electron transfer to the heme. In concert, a proton is lost from the hydrazide moiety and could be accepted by His108. In the following step, the C-N bond of hydrazide is broken, yielding a diazene and the acyl radical. The diazene intermediate (that may be stabilized by Trp107, Asp137 or the oxyferryl group) is reduced to hydrazine and ammonia, which may involve deprotonation of His108, Asp137 or Arg104 of KatG, although Arg104 may also play a role. The peroxidase activity of KatG catalyzes the conversion of Mn^{2+} to Mn^{3+} [64]. Rapid freeze-quench electron paramagnetic resonance spectroscopy experiments have shown that hydrogen atoms abstractions may be initiated by the a KatG tyrosil radical [65]. It has been shown that the yield of isonicotinoyl-NAD adduct is about the same for oxidation of isoniazid by KatG as it is by Mn^{3+} [66]. Accordingly, oxidation by Mn^{3+}-pyrophosphate has been proposed as an alternative method for nonenzymatic INH activation for simple chemical synthesis of various

INH derivatives that mimic the isonicotinyl-NAD adduct [67]. X-ray data shows a single isonicotinoyl-NAD(H) adduct in the open form with a 4*S* configuration (Fig. **2**) bound to InhA [56]. However, oxidation of isoniazid by the nonenzymatic Mn^{3+}-pyrophosphate method has been shown to generate a family of isomeric INH-NAD(H) adducts [68]. Proton and carbon NMR characterization showed that there are four main species of adducts: two are open diastereoisomers consisting of the covalent attachment of the isonicotinoyl radical at C4 of the nicotinamide moiety, and the other two result from cyclization of the amide group of the nicotinamide moiety of NAD(H) and the carbonyl group of the isonicotinoyl radical yielding diastereisomeric hemiamidals. These results prompted the authors to raise the question whether there is only one active form of INH-NAD(H) adduct or there are several forms able to inhibit InhA activity. However, production of INH-NAD adduct in the active site of InhA and analysis of purified adduct by reverse-phase HPLC showed two peaks, whose production has been ascribed to the experimental conditions used for HPLC that may promote interconversions of the INH-NAD adduct into a number of different forms [58]. Moreover, molecular modeling analysis showed that the C4 (*R*)enantiomeric INH-NAD adduct is stereochemically forbidden due to severe van der Waals repulsion around Gly192, Pro193 and Ile194 residues of InhA [69]. In any case, the nonenzymatic synthesis of compounds that mimic the isonicotinyl-NAD adduct [67] should allow the synthesis of new InhA inhibitors with potential antitubercular activity. Interestingly, formation of isonicotinyl adducts by the Mn^{3+}-pyrophosphate method and InhA inhibition studies suggest that the adenylic moiety and the amide function at position 3 of the pyridinium ring may play a role in inhibitor binding [70].

The current understanding of INH mode of action has provided clues into critical and unique biosynthetic pathways in mycobacteria. Owing to the INH chemical simplicity, oral availability and favourable toxicology profile, INH analogs appear worthy of examination as antitubercular agents. The mechanism of activation for isoniazid has been proposed to occur via an electron transfer reaction [71]. More recently, a redox reversible probe metal center has been used to show the involvement of an inner sphere pathway for activation of thionamide-containing pro-drugs [72]. An approach has thus been put forward for the design of self-activating drugs for the treatment of strains of *M. tuberculosis* resistant to drugs that require activation. These drugs would be activated by electron transfer reactions before interacting with

its cellular target. Most of the INH resistance is associated with *katG* structural gene alterations resulting in catalase-peroxidase mutant enzymes with impaired ability to form activated-INH intermediates that will form the INH-NAD adduct and ensuing inhibition of InhA enzyme activity. In this context, the use of a redox reversible metal complex coordinated to the pro-drug appears as a very first system. Accordingly, we have recently shown that a pentacyano(isoniazid)ferrateII complex (Fig. **4**) inhibits enzyme activity of both wild-type InhA and I21V mutant InhA identified in isoniazid-resistant clinical isolates of *M. tuberculosis* [73]. The *in vitro* kinetics of inactivation indicate that this process requires no activation by KatG, no need for the presence of NADH, and is also effective against INH-resistant mutant InhA. An MIC value of 0.2 $\mu g\ mL^{-1}$ for this inorganic complex was determined by the radiometric BACTEC AFB system for *M. tuberculosis* H37Rv strain, and toxicity assays in HL60 leukemia and MCS-7 breast cancer cells yielded an IC_{50} value > 25 $\mu g\ mL^{-1}$; thereby indicating a good selectivity index ($SI = IC_{50}/MIC > 125$; as suggested by the Tuberculosis Antimicrobial Acquisition & Coordinating Facility of USA, for a compound to move forward through screening programs SI should be larger than 10). More recently, we have shown that the pentacyano(isoniazid)ferrateII complex is a slow-onset inhibitor of *M. tuberculosis* InhA enzyme activity, with a true overall dissociation constant value of 70 nM [74]. In this mechanism of action an initial enzyme-inhibitor complex is rapidly formed, which then undergoes a slow isomerization reaction of to an enzyme-inhibitor binary complex in which the inhibitor is more tightly bound to enzyme. The weakness in the use of classical enzyme inhibitors as drugs for clinical conditions is that inhibition results in the upstream accumulation of the substrate for the enzyme, which may overcome the inhibition. By contrast, the build up of substrate cannot have any effect on the isomerization of enzyme-inhibitor complex typical of the slow-onset mechanism and hence reversal of the inhibition [75]. In addition, a half-time value of 630 min (10.5 hours) for the limiting step for inhibitor dissociation from the binary complex is a desirable feature since it may be expected to enhance inhibitor's effectiveness [76]. This inorganic complex may represent a new class of lead compounds to the development of anti-tubercular agents aiming at inhibition of a validated target and that is effective against INH-resistant strains.



Triclosan(TCL) **Pentacyano(isoniazid)ferrateII**

Fig. (4). Chemical structures of enoyl reductase inhibitors:Triclosan(TCL) and Pentaciano(isoniazid)ferrateII.

Triclosan (TCL; 2,4,4'-trichloro-2'-hydroxyphenyl ether; Fig. **4**) is a chlorinated bisphenol that possesses broad spectrum antimicrobial action and has a favorable satefy profile [77]. Accordingly, TCL has been used in many contemporary consumer and professional health care products, including hand soaps, shower gels, deodorant soaps, health care personnel hand washes, hand lotions and creams, toothpastes, mouth washes, underarm deodorants, surgical scrubs and drapes, and hospital over-the-bed table tops.

Strains of *Escherichia coli* that were selected for resistance to TCL harbored mutation in the enoyl-ACP reductase-encoding *fabI* gene provided evidence that TCL inhibits lipid biosynthesis [78], and the G93S enoyl reductase mutant was shown to have higher IC_{50} values than the wild-type enzyme for TLC [79]. Triclosan increases the affinity of *E. coli* FabI for NAD^+, leading to formation of a stable FabI-NAD^+-TCL ternary complex through non-covalent interactions with amino acid residues in the enzyme active site [80, 81]. TCL has been shown to act as a slow-onset inhibitor of *E. coli* FabI and that TCL binds to FabI-NAD^+ complex with a dissociation constant value of 20 - 40 pM [82]. The inhibition of *E. coli* enoyl reductase prompted studies on its mycobacterial counterpart. Genetic studies have thus shown that the enoyl-ACP reductase-encoding *inhA* gene is the target for triclosan in *M. smegmatis* [83] and *M. tuberculosis* [40]. TCL has been shown to be a submicromolar uncompetitive inhibitor of *M. tuberculosis* InhA, and that the M161V and A124V mutants have significantly higher inhibition constants for triclosan [84]. *In vitro* whole-cell assays demonstrated TCL activity against INH-

resistant and -sensitive *M. tuberculosis* strains with MIC values ranging from 20 to 60 μM, and a value of 8.5 μM for the inhibition constant of *M. tuberculosis* InhA [85]. The 2.6 Å-resolution three-dimensional structure of *M. tuberculosis* InhA, NAD^+ and triclosan complex shows that the hydroxyl-substituted ring of triclosan (Fig. **4**; the "A" ring) stacked with the nicotinamide ring of NAD^+ and formed a conserved hydrogen-bonding pattern with the 2'-OH group of NAD^+ and with Tyr158 in the catalytic active site [85]. The dichlorophenyl ring (Fig. **4**; the "B" ring) was oriented orthogonally to the "A" ring, and the chlorine atoms were projected toward the solvent. Interestingly, two molecules of triclosan were found by X-ray crystallography to bind to *M. tuberculosis* enoyl reductase [85], such mode of binding has not been observed for any TCL-enoyl reductase complex solved to date. The first TCL molecule occupies the same site as the singly-bound enzyme, whereas the second molecule is in an inverted orientation relative to the first TCL molecule, residing in an almost entirely hydrophobic site within the binding cavity and making van der Waals interactions between themselves and the protein. The presence of two molecules of triclosan in the active site of *M. tuberculosis* enoyl reductase has been attributed to the specificity for long chain fatty acyl substrates. To accommodate long fatty acids, *M. tuberculosis* enoyl reductase possesses a hydrophobic substrate-binding loop (residues 197-226) that is approximately 10 residues longer (residues 203-212) than the corresponding loop in *E. coli* and *B. napus* FabI. These data suggest that selective targeting of *M. tuberculosis* enoyl reductase may be achieved by a dual complex between triclosan molecules. Synthesis of triclosan analogues and high-throughput screening efforts to target enoyl-ACP reductase activity may lead to the development of new anti-tubercular agents.

CONCLUDING REMARKS

M. tuberculosis and enoyl-ACP reductase is a validated target for anti-tubercular agents development and to be used in drug screening efforts. In addition, owing the TCL and INH chemical simplicity and favourable toxicology profile, analogs of TCL and INH appear worthy of examination as antitubercular agents. An analog of INH, 1-isonicotinyl-2-nonaoyl hydrazine, has recently been shown to exhibit a two-fold increase in the degree of susceptibility against *M. tuberculosis* H37Rv [86]. Enzymatic acetylation of INH by *N*-acetyltransferase reduces the

therapeutic activity of this drug, resulting in underdosing, decreased bioavailability and the consequent possibility of acquired INH resistance. The lipophilic Schiff base N^2-cyclohexylidenyl isonicotinic acid hydrazide, in which the hydrazine moiety is blocked toward acetylation, has shown activity against both intracellular and extracellular *M. tuberculosis* (Erdman strain) *in vitro* and excelent bioavailability [87]. However, a 25-fold increase in the MIC value of this compound for INH-suceptible *M. tuberculosis* strain (0.03 μg mL^{-1}) as compared to an INH-resistant strain (> 0.75 μg mL^{-1}) casts doubt on its usefulness for the treatment of patients infected with resistant strains. A completely new approach, which was mentioned above, to INH analog design based on inorganic atoms attached to the nitrogen atom of the heterocyclic ring of isoniazid has provided promising results [73]. Two novel inhibitors, named Genz-10850 and Genz-8575, have been identified in a high-throughput screen against *M. tuberculosis* enoyl reductase with IC_{50} values in the submicromolar range [85]. The compound Genz-8575 showed good activity in whole-cell asssays against several drug-susceptible and drug-resistant strains of *M. tuberculosis*, with MIC values ranging from 1.25 to 30 μM (MIC = 1.5 μm for isoniazid against *M. tuberculosis* H37Rv strain). The X-ray crystal structure of the *M. tuberculosis* enoyl reductase and Genz-10850 binary complex showed that Genz-10850 has hydrogen-bonding interactions with NAD^+ and the catalytic residue Tyr158 similar to the binary complex formed between the enzyme and triclosan. The crystal structure should help optimization of these compounds. However, whether or not Genz-10850 and Genz-8575 target(s) the *M. tuberculosis* enoyl reductase *in vivo* as well as any effect they may have on mycolic acid biosynthesis will have to await for further studies.

Aminopyridine-based inhibitors of *Staphylococcus aureus* and *Haemophilus influenzae* enoyl reductase (FabI) enzymes have been developed through a combination of iterative medicinal chemistry and X-ray crystal structure based design [88]. One of these compounds, 3-(6-aminopyridin-3-yl)-*N*-methyl-*N*-[(1-methyl-1*H*-indol-2-yl)methyl]acrylamide, was shown to be a low micromolar inhibitor of FabI from *S. aureus* (IC_{50} = 2.4 μM) and *H. influenzae* (IC_{50} = 4.2 μM). In addition, this compound has shown good *in vitro* antibacterial activity against *S. aureus* (MIC = 0.5 μg mL^{-1}). Even though the inhibitory activity of this compound, if any, against *M. tuberculosis* enoyl reductase still needs to be shown, it may represent a new lead for antitubercular agent development. A series of 1,4-disubstituted imidazoles have shown *in vitro*

inhibitory activity against *S. aureus* and *E. coli* enoyl reductase enzyme, and whole-cell antimicrobial activity against *S. aureus* determined by a broth microdilution assay [89]. Compounds based on 2,9-disubstituted 1,2,3,4-tetrahydropyrido[3,4-*b*]indole have been synthesized and their inhibitory activity against *S. aureus* and *E. coli* enoyl reductase assessed *in vitro* [90]. The compounds showing submicromolar IC_{50} against enoyl reductase activity were tested for *in vitro* antibacterial activity with MIC values ranging from 0.5 to 4.0 $\mu g\ mL^{-1}$ against *S. aureus*. More recently, a naphthyridinone-based series of compounds were found to have inhibitory activity against *S. aureus* enoyl reductase with IC_{50} and MIC values lower than those for triclosan [90]. The activity of these compounds could be determined against *M. tuberculosis* to evaluate their potential as mycobacterial enoyl reductase inhibitors and as lead compounds to the development of antitubercular agents. Moreover, *M. tuberculosis* enoyl reductase may represent a molecular target for immobilization on solid support in natural-product screening efforts. The determination of the complete genome sequence of *M. tuberculosis* H37Rv strain has had a striking impact on researchers in the TB field [91] and represent a valuable tool for discovery of new drug targets.

ACKNOWLEDGMENTS

Financial support for this work was provided by Millennium Initiative Program MCT-CNPq, Ministry of Health - Secretary of Health Policy (Brazil) to DSS and LAB. DSS (304051/1975-06) and LAB (520182/99-5) are research career awardeed from the National Research Council of Brazil (CNPq).

REFERENCES

[1] Raviglione, M.C. (2003) *Tuberculosis*, **83**, 4-14.

[2] Corbett, E.L.; Watt, C.J.; Walker, N.; Maher, D.; Williams, B.G.; Raviglione, M.C. and Dye, C. (2003) *Arch. Intern. Med*., **163**, 1009-1021.

[3] Sharma, S.K.; Mohan, A. and Kadhiravan, T. (2005) *Indian J. Med. Res*., **121**, 550-657.

[4] Morens, D.M.; Folkers, G.K. and Fauci, A.S. (2004) *Nature*, **430**, 242-248.

[5] Russell, S. (2004) *Am. J. Trop. Med. Hyg*., **71**, 147-155.

[6] Espinal, M.A. (2003) *Tuberculosis*, **83**, 44-51.

[7] World Health Organization. Anti-tuberculosis drug resistance in the world. Third Global Report. 2004.

[8] Nunn P. WHO Global TB Programme, Tuberculosis Research and Surveillance Unit of the WHO Global TB Programme; Press Release WHO/74 (www.who.int/gtb/press/who74.htm), 1997.

[9] Blower, S.M. and Chou, T. (2004) *Nature*, **10**, 1111-1116.

[10] McKinney, J.D.; Jacobs, W.R.Jr. and Bloom, B.R. (1998) in: *Emerging Infections* (Krause, R.M., Ed.), Academic Press, New York, pp. 51-146.

[11] Pablos-Méndez, A.; Gowda, D.K. and Frieden, T.R. (2002) *Bull. World Health Organ*., **80**, 489-495.

[12] Cohen, T. and Murray, M. (2004) *Nat. Med*., **10**: 1117-1121.

[13] Duncan, K. (2003) *Tuberculosis*, **83**, 201-207.

[14] Gomez, J.E. and McKinney, J.D. (2004) *Tuberculosis*, **84**, 29-44.

[15] Heifets, L.B. (1994) *Semin. Respir. Infect.*, **9**, 84-103.

[16] Dutt, A.K. and Stead, W. (1994) *Dis. Month*, **43**, 247-274.

[17] Mashelkar, R.A. (2005) *Innovation Strategy Today*, **1**, 16-32.

[18] Morel, C.; Broun, D.; Dangi, A.; Elias, C.; Gardner, C.; Gupta, R.K.; Haycock, J.; Heher, T.; Hotez, P.; Kettler, H.; Keusch, G.; Krattiger, A.; Kreutz, F.; Lee, K.; Mahoney, R.; Mashelkar, R.A.; Min, H.-K.; Matlin, S.; Mzimba, M.; Oehler, J.; Ridley, R.; Senanayake, P.; Thorsteinsdóttir, H.; Singer, P.A. and Yun, M. (2005) *Innovation Strategy Today*, **1**, 1-15.

[19] UN Millennium Project (2005) www.unmillenniumproject.org/documents.

[20] Brennan, P.J. and Nikaido, H. (1995) *Annu. Rev. Biochem.*, **64**, 29-63.

[21] Brennan, P.J. (2003) *Tuberculosis*, **83**, 91-97.

[22] Schroeder, E.K.; de Souza, O.N.; Santos, D.S.; Blanchard, J.S. and Basso, L.A. (2002) *Curr. Pharm. Biotechnol*., **3**, 197-225.

[23] Watanabe, M.; Aoyagi, Y.; Mitome, H.; Fujita, T.; Naoki, H.; Ridell, M. and Minnikin, D.E. (2002) *Microbiology*, **148**, 1881-.902.

[24] Cronan, J.E. and Rock, C.O. (1996) in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, (Neidhardt, F.C., Ed.), ASM Press, Washington, pp. 612-636.

[25] Brindley, D.N., Matsumura, S. and Bloch, K. (1969) *Nature*, **224**, 666-669.

[26] Bloch, K. (1975) *Methods. Enzymol*. **35**, 84-90.

[27] Mdluli, K., Slayden, R.A., Zhu, Y., Ramaswamy, S., Pan, X., Mead, D., Crane, D.D., Musser, J.M. and Barry, C.E.III (1998) *Science* **280**, 1607-1610.

[28] Slayden, R.A.; Barry, C.E.III *Tuberculosis*, (2002), **82**, 149-160.

[29] Barry III, C.E.; Lee, R.R.; Mdluli, K.; Sampson, A.E.; Schroeder, B.G.; Slayden, R.A. and Yuan, Y. (1998) *Prog. Lipid Res*. **37**, 143-179.

[30] Basso, L.A. and Santos, D.S. (2005) Med. Chem. Rev. - Online. (accepted for publication; Ms. MCRO/03/INV).

[31] Takayama, K.; Wang, C. And Besra, G.S. (2005) *Clin. Microbiol. Rev.* **18**, 81-101.

[32] Bardou, F.; Raynaud, C.; Ramos, C.; Laneelle, M.A. and Laneelle, G. (1998) *Microbiology*, **144**, 2539-2544.

[33] Viveiros, M.; Portugal, I.; Bettencourt, R.; Victor, T.C.; Jordaan, A.M.; Leandro, C.; Ordway, D. and Amaral, L. (2002) *Antimicrob. Agents Chemother*., **46**, 2804-2810.

[34] Basso, L.A. and Blanchard, J.S. (1998) *Adv. Exp. Med. Biol*., **456**, 115-144.

[35] Ramaswamy, S.V.; Reich, R., Dou, S.-J., Jasperse, L., Pan, X., Wanger, A., Quitugua, T. and Graviss, E.A. (2003) *Antimicrob. Agents Chemother*., **47**, 1241-1250.

[36] Banerjee, A.; Dubnau, E.; Quémard, A.; Balasubramanian, V.; Um, K.S.; Wilson, T.; Collins, D.; de Lisle, G. and Jacobs, W.R.Jr. (1994) *Science*, **263**, 227-230.

[37] Quémard, A.; Sacchettini, J.C.; Dessen, A.; Vilchèze, C.; Bittman, R.; Jacobs, W.R.Jr. and Blanchard, J.S. (1995) *Biochemistry*, **34**, 8235-8241.

[38] Vilchèze, C.; Morbidoni, H.R.; Weisbrod, T.R.; Iwamoto, H.; Kuo, M.; Sacchettini, J.C. and Jacobs, W.R.Jr. (2000) *J. Bacteriol.*, **182**, 4059-4067.

[39] Larsen, M.H.; Vilchèze, C.; Kremer, L.; Besra, G.S.; Parsons, L.; Salfinger, M.; Heifets, L.; Hazbon, M.H.; Alland, D.; Sacchettini, J.C. and Jacobs, W.R.Jr. (2002) *Mol. Microbiol*., **46**, 453-466.

[40] Kremer, L.; Dover, L.G.; Morbidoni, H.R.; Vilchèze, C.; Maughan, W.N.; Baulard, A.; Tu, S.-C.; Honoré, N.; Deretic, V.; Sacchettini, J.C.; Locht, C.; Jacobs, W.R.Jr. and Besra, G.S. (2003) *J. Biol. Chem.*, **278**, 20547-20554.

[41] Basso, L.A.; Zheng, R.; Musser, J.M.; Jacobs, W.R.Jr. and Blanchard, J.S. (1998) *J. Infect. Dis*., **178**, 769-775.

[42] Dessen, A.; Quémard, A.; Blanchard, J.S.; Jacobs, W.R.Jr. and Sacchettini, J.C. (1995) *Science*, **267**, 1638-1641.

[43] Bottoms, C.A.; Smith, P.E. and Tanner, J.J. (2002) *Protein Sci.*, **11**, 2125-2137.

[44] Wierenga, R.K., De Maeyer, M.C.H., and Hol, W.G.J. (1985) *Biochemistry*, **24**, 1346-1357.

[45] Fersht, A. (1999) in *Structure and mechanism in protein science*, W.H. Freeman and Company, New York, pp. 324-348.

[46] Schroeder, E.K.; Basso, L.A.; Santos, D.S. and de Souza, O.N. (2005) *Biophys. J.*, (Epud ahead of print; **PMID: 15908576**].

[47] Brünger, A.T. (1992) X-Plor version 3.1 - A system for X-ray Crystallography and NMR, Yale University Press, New Haven, CT.

[48] Schwieters, C.D., Kuszewski, J.J., Tjandra, N. and. Clore, G.M. (2003*) J. Magn. Res.*, **160**, 66-74.

[49] Guex, N., and Peitsch, M.C. (1997) *Electrophoresis*, **18**, 2714-2723.

[50] Heym, B.; Alzari, P.M.; Honoré, N. and Cole, S.T. (1995) *Mol. Microbiol*., **15**, 235-245.

[51] Yu, S.; Girotto, S.; Lee, C. and Magliozzo, R.S. (2003) *J. Biol. Chem*., **278**, 14769-14775.

[52] Johnsson, K. and Schultz, P.G. (1994) *J. Am. Chem. Soc*., **116**, 7425-7426.

[53] Johnsson, K.; King, D.S. and Schultz, P.G. (1995) *J. Am. Chem. Soc*., **117**, 5009-5010.

[54] Basso, L.A.; Zheng, R. and Blanchard, J.S. (1996) *J. Am. Chem. Soc.*, **118**, 11301-11302.

[55] Zabinski, R.F. and Blanchard, J.S. (1997) *J. Am. Chem. Soc.*, **119**, 2331-2332.

[56] Rozwarski, D.A.; Grant, G.A.; Barton, D.H.R.; Jacobs, W.R.Jr. and Sacchettini, J.C. (1998) *Science*, **279**, 98-102.

[57] Lei, B.; Wei, C.-J. and Tu, S.-C. (2000) *J. Biol. Chem*., **275**, 2520-2526.

[58] Rawat, R.; Whitty, A. and Tonge, P.J. (2003) *Proc. Natl. Acad. Sci. USA*, **100**, 13881-13886.

[59] Odriozola, J.M. and Block, K. (1977) *Biochim. Biophys. Acta*, **488**, 198-206.

[60] Marrakchi, H.; Ducasse, S.; Labesse, G.; Montrozier, H.; Margeat, E.; Emorine, L.; Charpentier, X.; Daffé, M. and Quémard, A. (2002) *Microbiology*, **148**, 951-960.

[61] Veyron-Churlet, R.; Guerrini; O.; Mourey, L.; Daffé, M. and Zerbid D. (2004) *Mol. Microbriol*., **54**, 1161-1172.

[62] Vilchèze, C.; Weisbrod, T.R.; Chen, B.; Kremer, L.; Hazbón, M.H.; Wang, F.; Alland, D.; Sacchettini, J.C. and Jacobs, W.R.Jr. (2005) *Antimicrob. Agents Chemother.*, **49**, 708-720.

[63] Pierattelli. R.; Banci, L.; Eady, N.A.J.; Bodiguel, J.; Jones, J.N.; Moody, P.C.E.; Raven, E.L.; Jamart-Grégoire, B. and Brown, K.A. (2004) *J. Biol. Chem*. **279**, 39000-39009.

[64] Magliozzo, R.S. and Marcinkeviciene, J.A. (1997) *J. Biol. Chem*. **272**, 8867-8870.

[65] Chouchane, S.; Girotto, S.; Yu, S. and Magliozzo, R.S. (2002) *J. Biol. Chem*., **277**, 42633-42638.

[66] Wilming, M. and Johnsson, K. (1999) *Angew. Chem. Int. Ed*. **38**, 2588-2590.

[67] Nguyen, M.; Quemard, A.; Marrakchi, H.; Bernadou, J. and Meunier, B. (2001) *C. R. Acad. Sci. Paris, Série Iic, Chimie*, **4**, 35-40.

[68] Broussy, S.; Coppel, Y.; Nguyen, M.; Bernadou, J. and Meunier, B. (2003) *Chem. Eur. J*., **9**, 2034-2038.

[69] Scior, T.; Morales, I.M.; Eisele, S.J.G.; Domeyer, D. And Laufer, S. (2002) *Arch. Pharm. Pharm. Med. Chem*., **11**, 511-525.

[70] Nguyen, M.; Quémard, A.; Broussy, S.; Bernadou, J. and Meunier, B. (2002) *Antimicrob. Agents Chemother*. **46**, 2137-2144.

[71] DeBarber, A.; Mdluli, K.; Bosman, M.; Bekker, L.-G. and Barry, C.E.III. (2000) *Proc. Natl. Acad. Sci. USA.*, **97**, 9677-9682.

[72] Sousa, E.H.S.; Pontes, D.L.; Diógenes, I.C.N.; Lopes, L.G.F.; Oliveira, J.S.; Basso, L.A.; Santos, D.S. and Moreira, I.S. (2005) *J. Inorg. Biochem*., **99**, 368-375.

[73] Oliveira, J.S.; Sousa, E.H.S; Basso, L.A.; Palaci, M.; Dietze, R.; Santos, D.S. and Moreira, I.S. (2004) *Chem. Commun*., **3**, 312-313.

[74] Oliveira, J.S.; Sousa, E.H.S.; de Souza, O.N.; Moreira, I.S.; Santos, D.S. and Basso, L.A. (2005) Current Pharmaceutical Design (accepted for publication; Ms. P138313).

[75] Morrison, J.F. and Walsh C.T. (1988) *Adv. Enzymol. Relat. Areas Mol. Biol*., **61**, 201-301.

[76] Schloss, J.V. (1988) *Acc. Chem. Res*., **21**, 348-353.

[77] Tierno, P.M.Jr. (1999) *Am. J. Infect. Control*, **27**, 71-72.

[78] McMurry, L.M.; Oethinger, M. and Levy, S.B. (1998) *Nature*, **394**, 531-532.

[79] Heath, R.J.; Yu, Y.-T.; Shapiro, M.A.; Olson, E. and Rock, C.O. (1998) *J. Biol. Chem*. **273**, 30316-30321.

[80] Heath, R.J.; Rubin, J.R.; Holland, D.R.; Zhang, E.; Snow, M.E. and Rock, C.O. (1999) *J. Biol. Chem*. **274**, 11110-11114.

[81] Levy, C.W.; Roujeinikova, A.; Sedelnikova, S.; Baker, P.J.; Stuitje, A.R.; Slabas, A.R.; Rice, D.W. and Rafferty, J.B. (1999) *Nature* **398**, 383-384.

[82] Heath, R.J. and Rock, C.O. (2000) *Nature* **406**, 145-146.

[83] Heath, R.J.; Su, N.; Murphy, C.K. and Rock, C.O. (2000) *J. Biol. Chem*. **275**, 40128-40133.

[84] Parikh, S.L.; Xiao, G. and Tonge, P.J. (2000) *Biochemistry* **39**, 7645-7650.

[85] 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, W.R.Jr.; Sacchettini, J.C.; Fidock, D.A. (2003) *J. Biol. Chem*., **278**, 20851-20859.

[86] Mohamad, S.; Ibrahim, P.; Sadikun, A. (2004) *Tuberculosis*, **84**, 56-62.

[87] Hearn, M.J.; Cynamon, M.H. (2004) *J. Antimicrob. Chemother*., **53**, 185-191.

[88] Miller, W.H.; Seefeld, M.A.; Newlander, K.A.; Uzinskas, I.N.; Burgess, W.J.; Heerding, D.A.; Yuan, C.C.K.; Head, M.S.; Payne, D.J.; Rittenhouse, S.F.; Moore, T.D.; Pearson, S.C.; Berry, V.; DeWolf, W.E.; Keller, P.M.; Polizzi, B.J.; Qiiu, X.; Janson, C.A. and Huffman, W.F. (2002) *J. Med. Chem*., **45**, 3246-3256.

[89] Heerding, D.A.; Chan, G.; DeWolf, W.E.; Fosberry, A.P.; Janson, C.A.; Jaworski, D.D.; McManus, E.; Miller, W.H.; Moore, T.D.; Payne, D.J.; Qiu, X.; Rittenhouse, S.F.; Slater-Radosti, C.; Smith, W.; Takata, D.T.; Vaidya, K.S.; Yuan, C.C.K. and Huffman, W.F. (2001) *Bioorg. Med. Chem. Lett*., 11, 2061-2065.

[90] Seefeld, M.A.; Miller, W.H.; Newlander, K.A.; Burgess, W.J.; Payne, D.J.; Rittenhouse, S.F.; Moore, T.D.; DeWolf, W.E.; Keller, P.M.; Qiu, X.; Janson, C.A.; Vaidya, K.; Fosberry, A.P.; Smyth, M.G.; Jaworski, D.D.; Slater-Radosti, C. and Huffman, W.F. (2001) *Bioorg. Med. Chem. Lett*., **11**, 2241-2244.

[91] 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.

3. An inorganic iron complex that inhibits wild-type and an isoniazid-resistant mutant 2-*trans*-enoyl-ACP(CoA) reductase from *Mycobacterium tuberculosis* (Artigo publicado na revista "Chemical Communications" em 2004)

An inorganic iron complex that inhibits wild-type and an isoniazid-resistant mutant 2-*trans*-enoyl-ACP (CoA) reductase from *Mycobacterium tuberculosis*

Jaim S. Oliveira[†], Eduardo H. S. Souza[‡], Luiz A. Basso[†*], Moisés Palaci[¶], Reynaldo Dietze[¶], Ícaro S. Moreira[‡], and Diógenes S. Santos[&*].

[†] Universidade Federal do Rio Grande do Sul, Departamento de Biologia Molecular e Biotecnologia, Av. Bento Gonçalves 9500, Porto Alegre, RS 91501-970, Brazil; [‡] Universidade Federal do Ceará, Departamento de Química Orgânica e Inorgânica, Campus Pici, Fortaleza, CE 60455-760, Brazil; [¶] Universidade Federal do Espírito Santo, Núcleo de Doenças Infecciosas, Av. Marechal Campos 1468, Vitória, ES 29040-091, Brazil; [&] Faculdade de Farmácia, Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, CEP 90619-900, Brazil. E-mails: labasso@dna.cbiot.ufrgs.br or diogenes@dna.cbiot.ufrgs.br)

Received (in Cambridge, UK) 28th October, 2003, Accepted 24th Novenber 2003

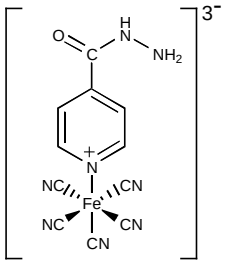
First published as an Advanced Article on the web 7th January 2004

The *in vitro* kinetics of inactivation of both wild-type and I21V InhA enzymes by $[Fe^{II}(CN)_5(INH)]^{3-}$ indicate that this process requires no activation by KatG, and no need for the presence of NADH. This inorganic complex may represent a new class of lead compounds to the development of anti-tubercular agents aiming at inhibition of a validated target.

In 1952 isoniazid (INH, isonicotinic acid hydrazide) was first reported to be effective in the treatment of tuberculosis.[1] However, strains of *Mycobacterium tuberculosis*, the causative agent of the disease, resistant to INH were reported shortly after its introduction.[2] The primary mechanism of multiple drug resistance in *M. tuberculosis* is the accumulation of mutations in individual drug target genes.[3] The mechanism of action of INH is complex, as mutations in at least five different genes (*katG*, *inhA*, *ahpC*, *kasA*, and *ndh*) have been found to correlate with isoniazid resistance.[4] The primary target of INH has been shown to be the product of the *inhA* structural gene.[5] The *inhA* gene codes for an NADH-dependent 2-*trans*-enoyl-ACP (CoA) reductase that exhibits specificity for long chain ($C_{18} > C_{16}$) enoyl thioester substrates,[6] consistent with its role in mycobacterial cell wall biosynthesis.[7] Missense mutations in the *inhA* structural gene, but lacking mutations in the *inhA* promoter region, *katG* gene and *oxyR-ahpC* region, were identified in INH-resistant clinical isolates of *M. tuberculosis*[8] and shown to correlate with changes in the NADH binding properties of enoyl reductase.[9] Moreover, deletions

of, or missense mutations in, the *katG* gene have been associated with decreased susceptibility to INH in approximately 50 % of clinical isolates of *M. tuberculosis*.[10] The isoniazid mechanism of action involves the conversion of INH by the mycobacterial *katG*-encoded catalase-peroxidase into a number of electrophilic intermediates.[11] Although isoniazid does not bind to the *inhA*-encoded enoyl reductase[12], the KatG-activated drug intermediate binds to, and inhibits the enoyl reductase activity in the presence of NAD^+ or NADH.[13] The three-dimensional structure determination of WT InhA, NADH, and activated isoniazid intermediate ternary complex has shown that the acylpyridine fragment of isoniazid is covalently attached to the C4 position of NADH.[14] This isonicotinyl-NAD^+ adduct binds to WT InhA with a dissociation constant value lower than 0.4 nM.[15] Isoniazid is, therefore, a KatG-activated pro-drug, that upon formation of an isonicotinyl-NAD^+ adduct inhibits the *M. tuberculosis* enoyl reductase, resulting in reduction of mycolic acid synthesis.

In trying to find better alternatives to INH, we have investigated an INH analog that contains a cyanoferrate moiety (**1**) and tested its ability to inhibit both WT and isoniazid-resistant I21V mutant enoyl reductases from *M. tuberculosis*.



1

Incubation of WT InhA with $Na_3[Fe^{II}(CN)_5(INH)]\cdot 4H_2O$[16] in the absence of NADH resulted in the time-dependent inactivation of the enzyme with an apparent first-order rate constant value of 327 (± 34) × 10^{-3} min^{-1} (Figure 1A, l inset; $t_{1/2}$ = 2.1 ± 0.2 min). The rate constant values were 65 (± 4) × 10^{-3} min^{-1} (Figure 1A, n inset; $t_{1/2}$ = 10.7 ± 0.7 min) in the presence of 10 μM NADH and 15.7 (± 0.7) × 10^{-3} min^{-1} (Figure 1A, ▲ inset; $t_{1/2}$ = 44 ± 2 min) in the presence of 100 μM NADH. Inactivation of WT InhA by oxidized INH derivatives produced by KatG in the absence of NADH has been reported not to occur at detectable levels in the time range tested here (25 – 30 min).[9, 21] Moreover, a value of 8.9 × 10^{-3} min^{-1} has been

reported for the rate constant for WT InhA enzyme inactivation by KatG-activated isoniazid in the presence of 100 μM NADH.[9]

The results presented here clearly demonstrate that WT InhA inactivation by $[Fe^{II}(CN)_5(INH)]^{3-}$ requires no activation by KatG, no need for the presence of NADH, and its mechanism of action probably involves interaction with the NADH binding site of the enzyme. The *in vitro* kinetics of inactivation of the I21V mutant enzyme, under identical conditions, proceeded with apparent first-order rate constant values of 315 (± 38) × 10^{-3} min^{-1} (Figure 1B, l inset; $t_{1/2}$ = 2.2 ± 0.3 min) in the absence of NADH, 99 (± 4) × 10^{-3} min^{-1} (Figure 1B, n inset; $t_{1/2}$ = 7.0 ± 0.3 min) in the presence of 10 μM NADH, and 14 (± 1) × 10^{-3} min^{-1} (Figure 1B, ▲ inset; $t_{1/2}$ = 50 ± 3 min) in the presence of 100 μM NADH. As for the WT InhA enzyme, these results demonstrate that inactivation of I21V mutant enzyme from a INH-resistant clinical isolate of *M. tuberculosis* by $[Fe^{II}(CN)_5(INH)]^{3-}$ requires no activation by KatG, no need for NADH, and probably the same site of interaction. Interestingly, since the *in vivo* NADH concentration in *M. tuberculosis* H37Rv has been estimated to be < 10 μM[22], the inorganic complex would probably display a better efficacy against isoniazid-resistant *M. tuberculosis* strains harboring *inhA* structural gene mutations than WT InhA strains. In order to verify if slow formation of a covalent binary compound between $[Fe^{II}(CN)_5(INH)]^{3-}$ and NADH could result in inactivation of WT and I21V enzymes, pre-incubation experiments[23] were performed and the kinetics of inactivation followed for WT and I21V enzymes. Apparent first-order rate constant of inactivation values of 15 (± 1) × 10^{-3} min^{-1} and 13 (± 2) × 10^{-3} min^{-1} were obtained for, respectively, WT and I21V InhA enzymes (data not shown). These values are within standard error of the ones determined with no pre-incubation suggesting that there is no slow formation of an intermediate compound capable of inactivating WT and I21V enzymatic activities.

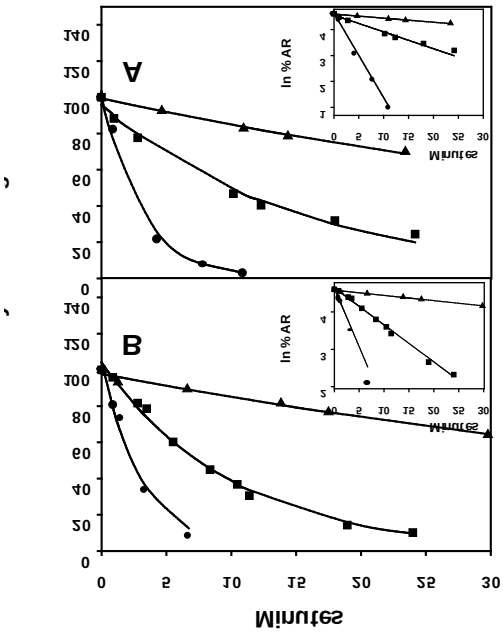


Fig. 1. (A) Inactivation of WT InhA (3 μM) by $[Fe^{II}(CN)_5(INH)]^{3-}$ (100 μM):[18] l, no NADH; n, 10 μM NADH; ▲, 100 μM NADH. The inset shows a plot of the natural log of the percentage of WT InhA activity remaining (%AR) versus time in the absence (l), presence of 10 μM NADH (n), and presence of 100 μM NADH (▲). (B) Inactivation of I21V InhA under the same experimental conditions described in part A (except that WT InhA was replaced for I21V InhA): no NADH (l), 10 μM NADH (n); and 100 μM NADH (▲). The inset shows a plot of the natural log of the percentage of I21V InhA activity remaining (%AR) versus time in the absence (l), presence of 10 μM NADH (n), and presence of 100 μM NADH (▲).

An MIC value of 1.0 μg mL^{-1} for the $[Fe^{II}(CN)_5(INH)]^{3-}$ compound was determined by the radiometric BACTEC AFB system.[24]. This value is in agreement with an MIC value of 0.2 μg mL^{-1} provided by the Tuberculosis Antimicrobial Acquisition & Coordinating Facility (TAACF) through a research and development contract with the U.S. National Institute of Allergy and Infectious Diseases. *M. tuberculosis* is susceptible to isoniazid in the range of 0.02 – 0.2 μg mL^{-1}.[26] Accordingly, the $[Fe^{II}(CN)_5(INH)]^{3-}$ compound appears to be a promising candidate for further antitubercular drug development and may represent a new class of lead compounds. Efforts to obtain crystals of the binary complex formed between this inorganic

compound and either WT or I21V InhA are currently underway and should assist in the design of a new class of antimycobacterial agents.

ACKNOWLEDGMENT. Financial support for this work was provided by QuatroG Pesquisa e Desenvolvimento to D.S.S., L.A.B., and I.S.M. Part of this work was also supported by Millennium Initiative Program MCT-CNPq, Ministry of Health, Secretary of Health Policy (Brazil) grant awarded to D.S.S., R.D., L.A.B. and M.P.

REFERENCES.

(1) Bernstein, J.; Lott, W.A.; Steinberg, B.S.; Yale, H.L. *Am. Rev. Tuberc*. **1952**, *65*, 357. Middlebrook, G. *Am. Rev. Tuberc*. **1952**, *65*, 765.

(2) Middlebrook, G. *Am. Rev. Tuberc*. **1952**, *65*, 765.

(3) Blanchard, J.S. *Annu. Rev. Biochem*. **1996**, *65*, 215. Basso, L.A.; Blanchard, J.S. *Adv. Exp. Med. Biol*. **1998**, 456, 115. Somoskovi, A.; Parsons, L.M.; Salfinger, M. *Respir. Res*. **2002**, *2*, 164.

(4) Schroeder, E.K.; de Souza, O.N.; Santos, D.S.; Blanchard, J.S.; Basso, L.A. *Curr. Pharm. Biotechnol*. **2002**, *3*, 197.

(5) Banerjee, A.; Dubnau, E.; Quémard, A.; Balasubramanian, V.; Um, K.S.; Wilson, T.; Collins, D.; Lisle, G.; Jacobs, W.R.Jr. *Science* **1994**, *263*, 227. 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, W.R.Jr. *Mol. Microbiol*. **2002**, *46*, 453. Kremer, L.; Dover, L.G.; Morbidoni, H.R.; Vilchèze, C.; Maughan, W.N.; Baulard, A.; Tu, S.-C.; Honoré, N.; Deretic, V.; Sacchettini, J.C.; Locht, C.; Jacobs, W.R.Jr.; Besra, G.S. *J. Biol. Chem*. **2003**, *278*, 20547.

(6) Quémard, A.; Sacchettini, J.C.; Dessen, A.; Vilchèze, C.; Bittman, R.; Jacobs, W.R.Jr.; Blanchard, J.S. *Biochemistry* **1995**, *34*, 8235.

(7) Vilchèze, C.; Morbidoni, H.R.; Weisbrod, T.R.; Iwamoto, H.; Kuo, M.; Sacchettini, J.C.; Jacobs, W.R.Jr. *J. Bacteriol*. **2000**, *182*, 4059.

(8) Ramaswamy, S.; Musser, J.M. *Tuber. Lung Dis*. **1998**, *79*, 3.

(9) Basso, L.A.; Zheng, R.; Musser, J.M.; Jacobs, W.R.Jr.; Blanchard, J.S. *J. Infect. Dis*. **1998**, *178*, 769.

(10) Heym, B.; Cole, S.T. *Res. Microbiol*.**1992**, *143*, 721. Cockeril, F.R.; Uhl, J.R.; Temesgen, Z.; Zhang, Y.; Stockman, L.; Roberts, G.D.; Williams, D.L.; Kline, B.C. *J. Infect. Dis*. **1995**, *171*, 240. Heym, B.; Alzari, P.M.; Honore, N.; Cole, S.T. *Mol. Microbiol*. **1995**, *15*, 235. Musser, J.M.; Kapur, V.; Williams, D.L.; Kreiswirth, B.N.; Soolingen, D.V.; van Embden, J.D.A. *J. Infect. Dis*. **1996**, *173*, 196. Ramaswamy, S.V.; Reich, R.; Dou, S.J.; Jasperse, L.; Pan, S.; Wanger, A.; Quituga, T.; Graviss, E.A. *Antimicrob. Agents Chemother*. **2003**, *47*, 1241.

(11) Johnson, K.; Schultz, P.G. *J. Am. Chem. Soc*. **1994**, *116*, 7425.

(12) Quémard, A.; Dessen, A.; Sugantino, M.; Jacobs, W.R.Jr.; Sacchettini, J.C.; Blanchard, J.S. *J. Am. Chem. Soc*. **1996**, *118*, 1561.

(13) Johnsson, K.; King, D.S.; Schultz, P.G. *J. Am. Chem. Soc*. **1995**, *117*, 5009.

(14) Rozwarski, D.A.; Grant, G.A.; Barton, D.H.R.; Jacobs, J.W.Jr.; Sacchettini, J.C. *Science* **1998**, *279*, 98.

(15) Lei, B.; Wei, C.-J.; Tu, S.-C. *J. Biol. Chem*. **2000**, *275*, 2520.

(16) The complex was prepared by the direct reaction of $Na_3[Fe(CN_5)(NH_3)]\cdot 3H_2O$ with INH, in aqueous solution, at room temperature, following the same procedure as reported elsewhere for similar iron cyanoferrate complexes.[17]

(17) Sousa, J.R.; Diogenes, I.C.N.; Carvalho, I.M.M.; Temperini, M.L.; Tanaka, A.A.; Moreira, I.S. *Dalton Transactions*, **2003**, *11*, 2231.

(18) Reactions of time-dependent inactivation of the enoyl reductases were carried out in 100 mM Na_2HPO_4, pH 7.5 at 25 °C, either 3 μM WT InhA or 3 μM I21V. Aliquots were taken at times specified on the x-axis of Figure 1, and steady-state enzyme activities were determined from rates of decrease in absorbance at 340 nm using 2-*trans*-dodecenoyl-CoA (100 μM) and NADH (200 μM). The substrate 2-*trans*-dodecenoyl-CoA was synthesized from 2-*trans*-dodecenoic acid and CoA by the mixed anhydride method[19] and purified by reverse-phase high-performance liquid chromatography using a 19 X 300 mm C_{18} μBondapak column (Waters Associates, Milford, MA) as previously described.[6] The ratio of absorbance of purified 2-*trans*-dodecenoyl-CoA at 232 nm and 260 nm was 0.62, a value that meets the established criterion for pure thioesters ($A_{232}/A_{260} \geq 0.52$.[20] WT InhA and I21V mutant enzymes were expressed and purified to homogeneity as described elsewhere.[9] The specific activity of WT InhA and I21V mutant with 2-*trans*-dodecenoyl-CoA (100 μM) and NADH

(200 μM) were, respectively, 13.5 ± 0.4 and 8.1 ± 0.3 U mg^{-1}. Activity measurements for both WT and I21V InhA proteins were performed up to 2 hours in 100 mM Na_2HPO_4, pH 7.5 at 25 °C, showing that no loss of enzyme activity could be observed in the time range studied.

(19) Goldman, P.; Vagelos, P.R. *J. Biol. Chem*. **1961**, *236*, 2620.

(20) Constantinides, P.P.; Steim, J.M. *J. Biol. Chem*. **1985**, *260*, 7573.

(21) Basso, L.A.; Zheng, R.; Blanchard, J.S. *J. Am. Chem. Soc*. **1996**, *118*, 11301.

(22) Gopinathan, K.P.; Sirsi, M.; Ramakrishnan, T. *Biochem. J*. **1963**, *87*, 444.

(23) $[Fe^{II}(CN)_5(INH)]^{3-}$ (200 μM) was pre-incubated with NADH (100 μM) for 2 hours in 100 mM Na_2HPO_4, pH 7.5 at 25 °C (solution A). An equal volume of this solution was added to a solution containing 6 μM of either WT InhA or I21V InhA mutant and 100 μM NADH, pre-incubated under the same conditions, and the kinetics of inactivation followed as described previously.[18]

(24) The Minimum Inhibitory Concentration (MIC) of $[Fe^{II}(CN)_5(INH)]^{3-}$ was determined radiometrically for *M. tuberculosis* H_{37}Rv ATCC 27294 strain (American Type Culture Collection, Rockville, Md.) by the method described by Heifets[25] in a BACTEC 460 instrument (Becton Dickinson).

(25) Heifets, L.B. *Am. Rev. Resp. Dis*. **1988**, *137*, 1217.

(26) Heifets, L.B. *Semin. Respir. Infect*. **1994**, *9*, 84.

4. Slow-onset inhibition of 2-*trans*-Enoyl-ACP(CoA) reductase from *Mycobacterium tuberculosis* (Artigo aceito para publicação na revista "Current Pharmaceutical Design" em Junho de 2005)

SLOW-ONSET INHIBITION OF 2-*trans*-ENOYL-ACP (CoA) REDUCTASE FROM *Mycobacterium tuberculosis* BY AN INORGANIC COMPLEX

Jaim S. Oliveira[§†], Eduardo H. S. de Sousa[¥], Osmar N. de Souza[§π], Ícaro S. Moreira[¥], Diógenes S. Santos[§]*, Luiz A. Basso[§†]*

[§] Centro de Pesquisas em Biologia Molecular e Funcional, Instituto de Pesquisas Biomédicas, Pontifícia Universidade do Rio Grande do Sul, Porto Alegre, RS, Brazil

[†] Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

[¥] Departamento de Química Orgânica e Inorgânica, Universidade Federal do Ceará, Campus Pici, Fortaleza, Brazil.

[π] Laboratório de Bioinformática, Modelagem e Simulação de Biossistemas – LABIO, Faculdade de Informática, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil.

[*] to whom correspondence may be addressed:
LAB: phone, +55-51-33166234; fax, +55-51-33167309; e-mail, labasso@cbiot.ufrgs.br; or DSS: phone, +55-51-33203513; fax, +55-51-33202515; e-mail, diogenes@pucrs.br

Key Words: tuberculosis, mycolic acids, enoyl reductase, InhA, isoniazid, inhibitors, inorganic complex, molecular modeling

Abstract

Tuberculosis (TB) remains the leading cause of mortality due to a bacterial pathogen, *Mycobacterium tuberculosis*. The reemergence of tuberculosis as a potential public health threat, the high susceptibility of human immunodeficiency virus-infected persons to the disease, and the proliferation of multi-drug-resistant strains have created a need for the development of new antimycobacterial agents. Mycolic acids, the hallmark of mycobacteria, are high-molecular-weight α-alkyl, β-hydroxy fatty acids, which appear mostly as bound esters in the mycobacterial cell wall. The product of the *M. tuberculosis inhA* structural gene (InhA) has been shown to be the primary target for isoniazid (INH), the most prescribed drug for active TB and prophylaxis. InhA was identified as an NADH-dependent enoyl-ACP reductase specific for long-chain enoyl thioesters. InhA is a member of the mycobacterial Type II fatty acid biosynthesis system, which elongates acyl fatty acid precursors of mycolic acids. Although the history of chemotherapeutic agent development demonstrates the remarkably successful tinkering of a few structural scaffolds, it also emphasizes the ongoing, cyclical need for innovation. The main focus of our contribution is on new data describing the rationale for the design of a pentacyano(isoniazid)ferrateII compound that requires no KatG-activation, its chemical characterization, *in vitro* activity studies against WT and INH-resistant I21V *M. tuberculosis* enoyl reductases, the slow-onset inhibition mechanism of WT InhA by the inorganic complex, and molecular modeling of its interaction with WT InhA. This inorganic complex represents a new class of lead compounds to the development of anti-tubercular agents aiming at inhibition of a validated target.

1. Introduction
2. Chemotherapy
3. Isoniazid (INH) mechanism of action
4. INH analogs and/or INH-adducts as antimycobacterial agents
5. A pentacyano(isoniazid)ferrateII inorganic complex
   5.1. Rationale for the design of $[Fe^{II}(CN)_5(INH)]^{3-}$
   5.2. Chemical characterizaton of $Na_3[Fe^{II}(CN)_5(INH)]^{3-}\bullet 3H_2O$
   5.3. *In vitro* activity of $[Fe^{II}(CN)_5(INH)]^{3-}$ against WT and I21V InhA enzymes
   5.4. Slow-binding inhibition kinetics of WT InhA by $[Fe^{II}(CN)_5(INH)]^{3-}$
   5.5. Dissociation of $[Fe^{II}(CN)_5(INH)]^{3-}$ from WT InhA and overall inhibition constant (K_i*)
   5.6. Two-step mechanism of WT InhA inhibition by $Fe^{II}(CN)_5(INH)]^{3-}$
6. Molecular modeling of mode of interaction between $[Fe^{II}(CN)_5(INH)^{3-}]$ and WT InhA
   6.1. Docking of $[Fe^{II}(CN)_5(INH)]^{3-}$ to NADH bound enzyme (control experiment)
   6.2. Docking of $[Fe^{II}(CN)_5(INH)]^{3-}$ to the 1ENY and 1BVR crystal structures.
7. Concluding remarks

Abbreviations

ACP = acyl carrier protein

ADR = adenine ribose

BSA = bovine serum albumin

CQ = chloroquine

DD-CoA = 2-*trans*-dodecenoyl-CoA

ENR = enoyl reductase

ETH = ethionamide

FASII = Type II dissociated fatty acid biosynthesis system

INH = isoniazid, isonicotinic acid hydrazide

InhA = 2-*trans* enoyl-ACP (CoA) reductase

LGA = Lamarckian Genetic Algorithm

MabA = β-ketoacyl reductase

MDR-TB = multidrug-resistant tuberculosis, defined as resistant to at least isoniazid and rifampicin

MLCT = metal-to-ligand charge transfer transition

NA = nicotinamide

PP = pyrophosphate

RMSD = root-mean-square deviation

TB = tuberculosis

WT = wild type

1. Introduction

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 [1]. 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 these were in developing countries [2]. In addition, approximately 12 % (226,000) of total deaths from TB was attributed to co-infection with *M. tuberculosis* and human immunodeficienty virus (TB-HIV). Immune deficient patients with HIV are at increased risk of latent *M. tuberculosis* infections progressing to active disease, and being transmitted to others [3]. Approximately 2 billion individuals are believed to harbor latent TB based on tuberculin skin test surveys [4], which represents a considerable reservoir of bacilli. It has been estimated that 3.2 % of the world's new cases of TB, in 2000, were multidrug-resistant tuberculosis (MDR-TB), defined as resistant to at least isoniazid and rifampicin [5]. According to the 2004 Global TB Control Report of the World Health Organization (WHO), 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 [6]. Localized high incidence rates of MDR-TB have been found in particular regions, which have been defined as hot zones based on either areas where the prevalence of MDR-TB cases is >5 % (that is, where >5 % of current cases are MDR-TB) or areas where the incidence of MDR-TB cases is >5 % (that is, where >5 % of new cases are MDR-TB) [5]. MDR-TB is an airborne bacterium that is spread just as easily as drug-sensitive TB. An individual who is sick with any strain of TB will infect between 10 and 20 people each year with that same strain [7], thereby making the hot zones of particular concern to public health officials. More recently, a mathematical model using Monte Carlo sampling methods has been employed to determine the key causal factors on generating the hot zones [8]. Paradoxically, areas with programs that successfully reduced wild-type pansensitive strains (as a result of high case detection and treatment rates) often evolved into hot zones. It has therefore been suggested that second-line drugs be quickly introduced to disrupt the amplification of

resistance. However, the bacteriostatic second-line drugs (amikacin, kanamycin, capreomycin, cycloserine, *para*-aminosalicylic acid, ethionamide, and fluoroquinolones) are more toxic and less effective and are given for at least three times as long and at 100 times the cost of basic short-course chemotherapy regimens [9, 10]. 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 [11]. Hence more effective and less toxic anti-tubercular agents are urgently needed to shorten the duration of current treatment, improve the treatment of MDR-TB, and to provide effective treatment of latent tuberculosis infection.

2. Chemotherapy

M. tuberculosis has been considered the world's most successful pathogen and this is largely due to the ability of the bacillum to persist in host tissues, where drugs that are rapidly bactericidal *in vitro* require prolonged administration to achieve comparable effects [12, 13]. TB is transmitted by *M. tuberculosis*-containing aerosolized droplets generated by the cough of a person infected with bacilli. The primary infection involves replication of *M. tuberculosis* within alveolar macrophages of the terminal air spaces of the lung, spread to local lymph nodes within the lung, and eventual dissemination of infection to remote sites in the body within 1 or 2 years after an initial infection. Only 5-10 % of infected individuals develop the disease because the host immune response against *M. tuberculosis* is highly effective in controlling bacterial replication. However, *M. tuberculosis* is almost never eradicated due to its ability to establish and maintain latency, a period during which the infected person does not have clinically apparent TB. Post-primary TB, which is predominantly a pulmonary disease, develops later in life, and can be caused either by reactivation of bacteria remaining from the initial infection or by failure to control a subsequent infection. Post-primary TB involves extensive damage to the lungs and efficient aerosol transmission of bacilli.

Effective chemotherapies that decreased TB mortality rate worldwide was led by the discovery of the antibacterial and antitubercular properties of streptomycin in 1944 [14], and

both isoniazid and pyrazinamide in 1952 [15, 16]. The later introduction of ethionamide, rifampicin, ethambutol and ciprofloxacin to the arsenal used to treat tuberculosis seemed to provide an adequate number of effective antimicrobial agents [17]. Effective tuberculosis chemotherapy must include early bactericidal action against rapidly growing organisms and subsequent sterilization of the semidormant and dormant populations of bacilli [18]. The first-line drugs isoniazid, rifampicin, streptomycin, and ethambutol exhibit early bactericidal activity against actively metabolizing bacilli [19]. Pyrazinamide is active against the semidormant bacilli in acidic intracellular environments. The modern, standard "short-course" therapy [18] for TB 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 an additional four months. This combination therapy must be strictly followed to prevent drug resistance and relapse, and direct observation of patient compliance is the most reliable way to ensure effective treatment and prevent the acquisition of resistance. The DNA mutations identified in drug-resistant strains of *M. tuberculosis* have been reviewed [17, 20, 21, 22]. The bacteriostatic second-line drugs ethionamide, cycloserine and *para*-aminosalicylic acid (PAS) are reserved to strengthen the treatment of drug-resistant disease or when bactericidal drugs are prohibited because of toxicity [23].

3. Isoniazid (INH) mechanism of action

Isoniazid (INH, isonicotinic acid hydrazide; Fig. **1**) is one of the oldest synthetic antitubercular, and the most prescribed drug for active infection and prophylaxis [24]. The product of the *M. tuberculosis inhA* structural gene (InhA or ENR), which is located downstream of an ORF coding for a β-ketoacyl reductase (MabA) in the *inhA* operon, has been shown to be a major target for isoniazid [25]. InhA was identified as an NADH-dependent enoyl-ACP (acyl carrier protein) reductase enzyme, which exhibits specificity for long-chain ($C_{18} > C_{16}$) enoyl thioester substrates [26]. InhA is a member of the mycobacterial Type II dissociated fatty acid biosynthesis system (FASII), which elongates acyl fatty acid precursors yielding the long carbon chain of the meromycolate branch of mycolic acids, the hallmark of mycobacteria [22]. The mycobacterial cell wall is comprised of three covalently linked macromolecules: peptidoglycan,

arabinogalactan, and mycolic acid, which is often described as mycolyl-arabinogalactan-peptidoglycan complex (mAGP). Mycolic acids have become one of the defining taxonomic characteristics of many species in genera such as *Mycobacterium, Corynebacterium, Dietzia, Nocardia, Rhodococcus,* and *Tsukamurella*. Mycolic acids are high-molecular-weight α-alkyl, β-hydroxy fatty acids, which appear mostly as bound esters in tetramycolylpentaarabinosyl cluster in the mycobacterial cell wall. Consistent with InhA as the major target of INH mode of action, inactivation of the *M. smegmatis inhA*-encoded enoyl reductase and INH treatment resulted in similar morphological changes to the mycobacterial cell wall leading to cell lysis [27]. Overexpression of *inhA* has been shown to confer resistance to INH and ethionamide (ETH) in *M. smegmatis*, *M. bovis* BCG, and *M. tuberculosis* [28]. Further biochemical and genetic evidence has been given likewise showing that InhA is the primary target of INH [29]. Consistent with these results, mutations in the *inhA* structural gene and in the *inhA* locus promoter region have been associated with isoniazid resistance [30]. Moreover, 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 values for NADH than wild-type (WT) InhA, whereas there were only modest differences in the steady-state kinetic parameters [30]. Consistent with these results, a comparison of the crystal structure of binary complex of WT and INH-resistant mutant enoyl reductases with NADH [31] showed that disruption of a hydrogen bond network in the mutant protein could account for higher dissociation constant value for the coenzyme.

More recently, molecular dynamics simulations of WT and INH-resistant clinical isolates of *M. tuberculosis* showed that mutations of the glycine-rich loop residues I21V and I16T resulted in a change in the pattern of direct hydrogen bond contacts with the pyrophosphate moiety of NADH, in which the pyrophosphate moiety of NADH moves apart from its binding site [32]. The increase in the NADH dissociation constant observed for the I21V and I16T isoniazid-resistant mutants [30] can be attributed to a decrease in the number of direct hydrogen bond interactions between the cofactor and aminoacids in the binding pocket of the enzyme and to a decrease in water-mediated interactions between cofactor and enzyme [32]. The correlation between the NADH binding properties in solution, molecular dynamics simulation studies and resistance to INH inactivation of *M. tuberculosis* InhA mutants provides a mechanism of

resistance at molecular level for this clinically important drug. 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 [22]. However, a drug target for a bactericidal drug has been suggested to be an enzyme of the bacterium (i) that binds the drug, (ii) that is inhibited by the drug, and (iii) whose inhibition induces the death of the bacterium [27]. Accordingly, based on these criteria, genetic and biochemical results point to InhA as the bona fide target for INH in *M. tuberculosis*.

Deletions of, or missense mutations in, the *katG* gene have been associated with decreased susceptibility to INH in approximately 50 % of clinical isolates of *M. tuberculosis* [33, 34]. It has recently been shown that the S315T KatG mutant, which is one of the most commonly encountered substitutions in clinical INH-resistant strains, has reduced affinity for INH, which would result in decreased drug activation and ensuing INH resistance [35]. INH is a pro-drug that is activated by the mycobacterial *katG*-encoded catalase-peroxidase enzyme in the presence of manganese ions, NAD(H) and oxygen [36-39]. The KatG-produced acylpyridine fragment of isoniazid is covalently attached to the C4 position of NADH (Fig. **1**), and forms a binary complex with the wild-type enoyl reductase of *M. tuberculosis* [40]. This isonicotinyl-NAD^+ adduct has been characterized spectroscopically and shown to bind to InhA with a dissociation constant value lower than 0.4 nM [41]. The isonicotinyl-NAD^+ adduct has recently been shown to be a slow, tight-binding competitive inhibitor of WT InhA [42]. The initial rapidly reversible weak binding (K_i = 16 nM) is followed by a slow isomerization leading to a tighter enzyme-inhibitor complex with an overall dissociation constant (K_i*) value of 0.75 nM.

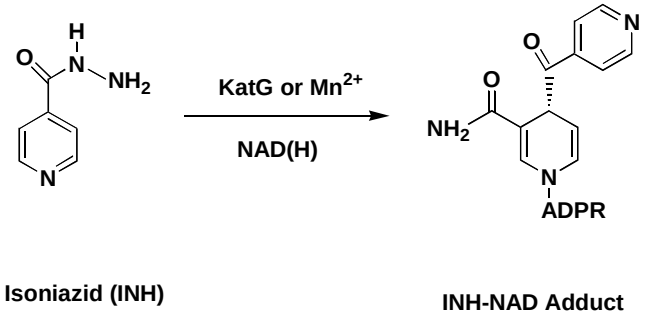


Fig. (1). Chemical structure of isoniazid and INH-NAD adduct that inhibits InhA enzyme activity.

The isoniazid mechanism of action requires the conversion of INH by mycobacterial KatG into a number of electrophilic intermediates [36]. It has been proposed that formation of the isonicotinyl-NAD^+ adduct may be through addition of either an isonicotinic acyl anion to NAD^+ or an isonicotinic acyl radical to an $NAD^{\cdot}$ radical [40]. The mycobacterial heme-containing KatG has been shown to function as a highly active catalase as well as a broad specificity peroxidase [43]. The catalase-peroxidase KatG enzyme is classified as Class I peroxidase whose catalytic cycle includes a classical Compound I (oxyferryl [Fe(IV)=O] porphyrin π-cation radical) that is catalytically competent, and that is reduced by a number of substrates including INH without evidence for stabilization of Compound II [44]. Rapid freeze-quench electron paramagnetic resonance spectroscopy experiments have shown that a KatG tyrosyl radical doublet signal was rapidly quenched by addition of isoniazid, suggesting that oxidations or, more likely, hydrogen atoms abstractions may be initiated by the tyrosyl radical [45]. Tyrosine-353 residue has been proposed to be the principal site for tyrosyl radical formation in mycobacterial KatG [46]. However, tyrosine-353 appears to be an unlikely candidate for the tyrosyl radical based on the crystal structure of *M. tuberculosis* KatG refined to 2.4 Å [47]. The KatG crystal structure and computational studies allow to predict that the binding site for isoniazid is located near the *δ-meso* heme edge rather than in a surface loop structure as previously proposed, and offer a rational interpretation of the association between a number of mutations in the *katG* gene and INH resistance [47]. An enzyme-catalyzed mechanism for INH activation has recently been proposed, leading to isonicotinoyl radical formation (thought to be the activated form of INH) and production of the amide end product via a diazene intermediate [48]. In the proposed mechanism, the oxyferryl group of compound I of KatG, generated after reaction with peroxide, is reduced by INH in a single electron transfer to the heme with concomitant transfer of a proton from the hydrazide moiety to His-108. The C-N bond of hydrazide is broken, yielding a diazene and the acyl radical. The diazene intermediate (that may be stabilized by Trp107, Asp137 or the oxyferryl group) is reduced to hydrazine and ammonia, which may involve deprotonation of His108, Asp137 or Arg104 of KatG.

4. INH analogs and/or INH-adducts as potential antimycobacterial agents

The serendipitous discovery of isoniazid has now provided clues into critical and unique biosynthetic pathways in mycobacteria. The chemical simplicity of this molecule, and decades of synthetic endeavors after the initial discovery of INH, suggest that more potent analogs, prepared by classic organic synthesis or novel combinatorial synthetic methods, would appear to be hard to find. However, given the oral availability and favorable toxicology profile, INH analogs appear to be worthy of examination as antitubercular agents.

The peroxidase activity of mycobacterial KatG catalyzes the conversion of Mn^{2+} to Mn^{3+} [49]. It has been shown that the yield of isonicotinoyl-NAD adduct is about the same after oxidation of isoniazid by KatG or Mn^{3+} [50]. Accordingly, Mn^{3+}-pyrophosphate has been proposed as an alternative oxidant that mimics the activity of KatG, thus providing a nonenzymatic isoniazid activation method for simple chemical synthesis of various INH derivatives that mimic the isonicotinyl-NAD adduct [51]. A single isonicotinoyl-NAD(H) adduct in the open form with a 4*S* configuration (Fig. **1**) has been detected in the active site of *M. tuberculosis* InhA by X-ray crystallography [40]. In agreement with this configuration for the adduct, a molecular model of the C4 (*R*)enantiomeric INH-NAD adduct indicates that this form is stereochemically forbidden due to van der Waals contacts [52]. The nonenzymatic synthesis of compounds that mimic the isonicotinyl-NAD adduct [51] should allow the synthesis of new InhA inhibitors with potential antimycobacterial activity.

An analog of INH, 1-isonicotinyl-2-nonaoyl hydrazine, has recently been shown to exhibit a two-fold increase in the degree of susceptibility against *M. tuberculosis* H37Rv [53]. However, no data have been presented on INH-resistant strains of *M. tuberculosis*, thereby making this analog of limited therapeutic value if not proven to be effective against INH-resistant strains of *M. tuberculosis*. Enzymatic acetylation of INH by *N*-acetyltransferase reduces the therapeutic activity of this drug, resulting in underdosing, decreased bioavailability and the consequent possibility of acquired INH resistance. The lipophilic Schiff base N^2-cyclohexylidenyl isonicotinic acid hydrazide, in which the hydrazine moiety is blocked toward acetylation, has shown activity against both intracellular and extracellular *M. tuberculosis*

(Erdman strain) *in vitro* and excelent bioavailability [54]. However, this INH Schiff base analog showed no antitubercular activity against INH-resistant strain (>0.75 μg mL^{-1}) that obviously casts doubt on its usefulness for the treatment of patients infected with resistant strains.

5. A pentacyano(isoniazid)ferrateII inorganic complex

Although the history of chemotherapeutic agent development demonstrates the remarkably successful tinkering of a few structural scaffolds by generations of medicinal chemists to meet the challenges of evolution of drug resistance, it also emphasizes the ongoing, cyclical need for innovation [55]. A brief account of a completely new approach to INH analog design based on inorganic atoms attached to the nitrogen atom of the heterocyclic ring of isoniazid has recently been given [56]. Here we describe the rationale for the design of a pentacyano(isoniazid)ferrateII compound $[Fe^{II}(CN)_5(INH)]^{3-}$ (Fig. **2**), its chemical characterization, *in vitro* activity studies of this inorganic complex against WT and INH-resistant I21V *M. tuberculosis* enoyl reductase enzymes (InhA), the slow-onset inhibition mechanism of WT InhA by the inorganic complex, and molecular modeling of its interaction with WT InhA. This inorganic complex represents a new class of lead compounds to the development of anti-tubercular agents aiming at inhibition of a validated target.

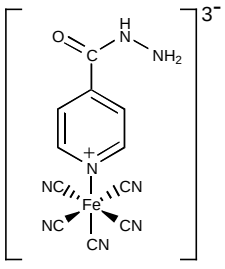


Fig. (2). Structural formula of pentacyano(isoniazid)ferrateII complex: $[Fe^{II}(CN)_5(INH)]^{3-}$

5.1. Rationale for the design of $[Fe^{II}(CN)_5(INH)]^{3-}$

It has been proposed that free radicals generated during the oxidation of INH by catalase-peroxidase of *M. tuberculosis* are responsible for the toxicity of this drug [36]. Various oxidizing products including isonicotinic acid, isonicotinamide, and isonicotinaldehyde are formed from the *in vitro* activation of the INH by KatG enzyme. A widely accepted pathway for the *in vivo* INH activation is that it is metabolized by oxidation to acyl diimide, then to diazonium ion or isonicotinyl radical that binds to enoyl reductase, and inhibits its activity in the presence of NAD^+ or NADH [37]. Isoniazid is, therefore, a KatG-activated pro-drug, that upon formation of an isonicotinyl-NAD^+ adduct inhibits the *M. tuberculosis* enoyl reductase, resulting in reduction of mycolic acid synthesis. Based on the mechanism of activation propose for isoniazid, via electron transfer reaction [57], an alternative self-activation route can be proposed for designing new drugs for the treatment of WT and INH-resistant tuberculosis. These drugs would be activated by electron transfer reactions before interacting with its cellular target. Most of the INH resistance is associated with *katG* structural gene alterations resulting in catalase-peroxidase mutant enzymes with impaired ability to form activated-INH intermediates. One approach to overcome this drug resistance is based on the synthesis of a new molecule capable to promote an inner-sphere electron transfer reaction. In this context, the use of a redox reversible metal complex coordinated to the prodrugs appears as a very first system. For this purpose, we have found that the pentacyanoferrate(II) metal center and isoniazid ligand is a suitable model system on the perspective of a new approach to drug design for the treatment of WT and INH-resistant tuberculosis by proposing a self-activation mechanism as illustrated in Fig. **3**. Upon oxidation of $[Fe^{II}(CN)_5(INH)]^{3-}$ complex ($E_{1/2}$ = 549 mV *vs*. NHE) by oxygen and/or cellular oxidants, it is postulated that the thermodynamically unstable iron(III) complex would undergo a rapid intramolecular electron transfer reaction forming the $[Fe^{II}(CN)_5(isonicotinyl)]^{3-}$ intermediate species which could decay to the $[Fe^{II}(CN)_5(L)]^{3-}$ complex (L= isonicotinic acid, isonicotinamide, or isonicotinaldehyde) and bind to enoyl reductase resulting in inhibition of mycolic acid synthesis without the interference of KatG enzyme**.**

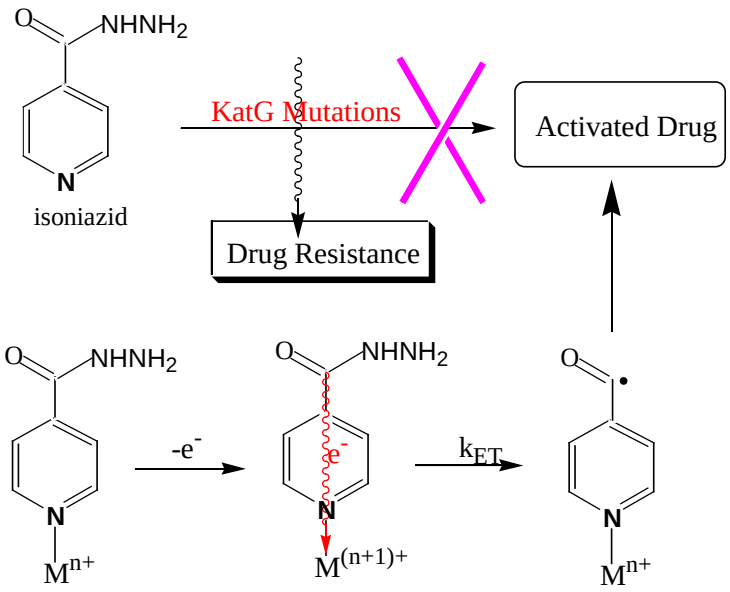


Fig. (3). A proposal for a self-activating mechanism of isoniazid.

5.2. Chemical characterizaton of $Na_3[Fe^{II}(CN)_5(INH)^{3-}]\bullet 3H_2O$

The $Na_3[Fe^{II}(CN)_5(INH)]\bullet 3H_2O$ complex was prepared by the direct reaction of $Na_3[Fe(CN_5)(NH_3)]\bullet 3H_2O$ with isoniazid, in aqueous solution, at room temperature, following the same procedure reported elsewhere for similar pentacyanoferrate(II) complexes [58]. Results of microanalysis of $Na_3[Fe^{II}(CN)_5(INH)]\bullet 3H_2O$ compound yielded the following results: %C(exp)=28.2 (theor=28.4), %H(exp)=3.3 (theor=3.2), %N(exp)=23.8 (theor=24.0). Chemical shifts (δ) of 500 MHz ^{1}H NMR spectrum are consistent with INH coordinated to the $[Fe^{II}(CN)_5]^{3-}$ ion complex. The ^{1}H NMR spectrum for isoniazid shows a downfield chemical shift of the doublet at δ 8.66 ppm with coupling constants (J) values of 1.65 Hz and 4.58 Hz, which were assigned to the deshielded *ortho* hydrogens to nitrogen 1 of the pyridine ring. The upfield chemical shifts of the doublet at δ 7.67 ppm with J values of 1.67 Hz and 4.58 Hz were attributed to the shielded *meta* hydrogens of isoniazid. Interestingly, no signal could be observed for the hydrogen atoms of the hydrazine substituent of the pyridine ring due to rapid exchange with solvent. The ^{1}H NMR spectrum of $Na_3[Fe^{II}(CN)_5(INH)]\bullet 3H_2O$ showed the signals characteristic of the hydrogens of the pyridine ring (δ 9.11 ppm for the *ortho* and δ 7.41 ppm for the *meta*

hydrogens). These results are consistent with coordination of the metal center with the pyridine ring nitrogen resulting in deshielding of the *ortho* hydrogens and shielding of *meta* hydrogens due to the cyanide anisotropy and the (INH)pπ^* ← dπ(Fe^{II}) back-bonding. The Mössbauer parameters values (δ = 0.219 ± 0.001 mms^{-1}, Δ = 0.219 ± 0.002 mms^{-1}, referred to SNP) for the $Na_3[Fe^{II}(CN)_5(INH)]\bullet 3H_2O$ complex strongly suggest the presence of an unique ^{57}Fe(II) Mössbauer nucleus, consistent with the proposed chemical composition [59]. The most prominent feature of the electronic spectrum of the $[Fe^{II}(CN)_5(INH)]^{3-}$ complex is the band at 436nm (ε = 4.0 x 10^3 $M^{-1}cm^{-1}$) assigned to the [(INH)pπ^* ← dπ(Fe^{II})] metal-to-ligand charge transfer (MLCT) transition, suggesting a strong interaction between the metal ion and the nitrogen atom of the pyridine ring of INH. The infrared (IR) vibrational spectrum of $[Fe^{II}(CN)_5(INH)]^{3-}$ displays bands at wavenumbers (cm^{-1}) associated with stretching frequencies of νC≡N groups (2044 cm^{-1}), νFe^{II}–C (563 cm^{-1}), C−H bonds involving sp^2-hybridized carbons of the pyridine ring (3300 cm^{-1}), and C=O stretching (1668 cm^{-1}). The pentacyano(isoniazid)ferrateII compound is stable at pH 2.0 (acid-base equilibrium of CN groups with a pK_a = 2.3). The stability of the inorganic complex at acidic pH indicates that it has favourable characteristics to be administered orally, since it would not be degraded by the acidic pH of human digestive system.

5.3. *In vitro* activity of $[Fe^{II}(CN)_5(INH)]^{3-}$ against WT and I21V InhA enzymes

The I21V enoyl reductase (InhA) mutant enzyme, which was identified in INH-resistant clinical isolates of *M. tuberculosis* harboring *inhA*-structural gene mutations, has been shown by binding studies to exhibit increased dissociation constant value for NADH (13.9 μM) as compared to WT InhA enzyme (0.57 μM) [30]. A comparison of the *in vitro* kinetics of inactivation by KatG-activated INH, which was carried out in the presence of NADH (10μM), KatG (2μM) and INH (200μM), showed a decreased value for the apparent first-order rate constant (k) of inactivation for the I21V mutant enzyme (3.3 x 10^{-3} min^{-1}; $t_{1/2}$ = 209 min) as compared to the WT enzyme (4.87 x 10^{-3} min^{-1}; $t_{1/2}$ = 142min) [30]. The k values of inactivation by $[Fe^{II}(CN)_5INH]^{3-}$ (100 μM), in the absence of NADH, for WT InhA (327 x 10^{-3} min^{-1}; $t_{1/2}$ = 2.1 min) and I21V mutant (315 x 10^{-3} min^{-1}; $t_{1/2}$ = 2.2 min) were within experimental error [56].

The values for inhibition in the presence of NADH (10 μM), and absence of both KatG and INH, were determined to be 99 x 10^{-3} min^{-1} ($t_{1/2}$ = 7.0min) and 65 x 10^{-3} min^{-1} ($t_{1/2}$ = 10.7 min) for, respectively, I21V and WT enzymes [56]; whereas no inactivation could be detected for KatG-activated INH in the same time range [30]. In addition, pre-incubation studies showed that no slow formation of a covalent binary compound between $[Fe^{II}(CN)_5(INH)]^{3-}$ and NADH capable of inactivating WT and mutant enzyme activities was detected [56]. These studies indicate that the mode of action of $[Fe^{II}(CN)_5(INH)]^{3-}$ compound requires no activation by KatG, no need for the presence of NADH, and is also effective against INH-resistant mutant InhA [56].

5.4. Slow-binding inhibition kinetics of WT InhA by $[Fe^{II}(CN)_5(INH)]^{3-}$

In order to try to elucidate the kinetic mechanism of time-dependent inhibition observed for InhA, typical slow-binding inhibition experiments [60] were carried out for WT InhA in the presence of various concentrations of $[Fe^{II}(CN)_5(INH)]^{3-}$. The progress curves for enzyme activity indicate that the inorganic iron complex is a slow-binding inhibitor of WT InhA (Fig. **4**). In this mechanism of action an initial complex EI is rapidly formed, which then undergoes a slow isomerization reaction to EI^*, where the inhibitor is more tightly bound to enzyme (Fig. **5**, mechanism B).

For slow-binding inhibitors the binding, which is the establishment of the equilibrium between enzyme, inhibitor, and enzyme-inhibitor complexes, occurs slowly on the steady-state time scale of seconds to minutes [60]. The slow-binding experiments were performed under conditions where the concentration of inhibitor was significantly greater than that of the enzyme ($I_t >> E_t$) and measurements made within a time range over which there was no significant substrate depletion (<10%). Thus, apparent pseudo first-order rate constants could be calculated by fitting the progress curves for time-dependent inactivation of WT InhA (Fig. **4**) to the following integrated equation for slow binding inhibition [42, 60-61],

$$A_t = A_0 - v_s t - (v_i - v_s)(1 - \exp(-k_{obs} t))/k_{obs} \quad \textbf{[1]}$$

where A_t and A_0 are the 340-nm absorbance at time t and time 0, v_i, v_s, and k_{obs} represent, respectively, the initial velocity, the final steady-state velocity and the apparent first-order rate constant for the establishment of the equilibrium between EI and EI^* complexes (the observed rate constant for formation of EI* or the apparent isomerization rate constant).

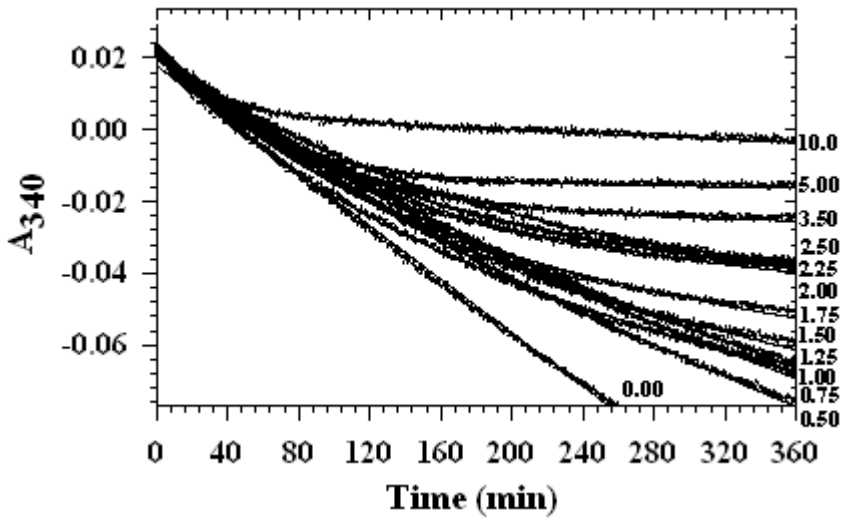


Fig. (4). Time-dependent inactivation of WT InhA by 0-10 μM $[Fe^{II}(CN)_5(INH)]^{3-}$. The solid curves represent the best fit of the data to equation **1** for slow-binding inhibition. The assay mixtures were kept at 25°C and contained 100 mM phosphate (pH 7.5), 85 μM 2-*trans*-dodecenoyl-CoA (DD-CoA), 250 μM NADH, 8% (vol/vol) glycerol, 0.1 mg mL^{-1} BSA, and 0-10 μM $[Fe^{II}(CN)_5(INH)]^{3-}$. The reactions were initiated by adding enzyme (0.1-0.25 nM) to assay mixture and were monitored by following the decrease in NADH absorbance at 340 nm for 6 hours.

Glycerol and BSA were added to stabilize WT InhA activity since the enzyme was inactive in the absence of these reagents when used in concentrations of 3 nM or lower. The same assay conditions have been used previously in experiments of slow, tight-binding inhibition of WT InhA by the isoniazid-NAD adduct [42], thereby allowing a comparison to be made between the results described here and for the INH-NAD adduct.

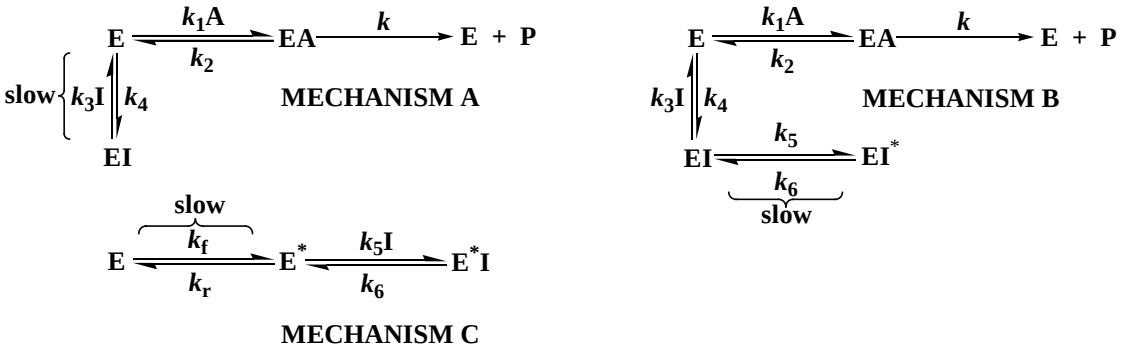


Fig. (5). Mechanisms that describe reversible slow-binding inhibition of enzyme-catalyzed reactions.

Higher concentrations of inhibitor caused the steady-state to be reached more quickly but yielded lower v_s values. Accordingly, values for v_s could only be estimated from progress curves for assay mixtures containing up to 10 μM of $[Fe^{II}(CN)_5(INH)]^{3-}$. However, this inhibitor concentration was not sufficient to differentiate between the two basic slow-binding mechanisms that have been described by Morrison and Walsh as mechanism A and mechanism B [Fig. **5**, 60]. In mechanism A, it is assumed that the time-dependent enzyme inactivation observed is due to a slow bimolecular interaction between inhibitor and enzyme, because the inhibitor concentration is low and/or because of barriers that the inhibitor encounters in its binding at the active site of the enzyme. For mechanism B, it is assumed that there is an initial rapid interaction between enzyme and inhibitor to form EI which then undergoes a slow isomerization (conformational change) to EI^* [62]. In an inhibition that conforms to this mechanism , the overall dissociation constant (K_i^*) must be lower than the dissociation constant for initially formed complex EI (K_i), a condition that can be satisfied only when k_6 is < k_5 ($K_i^* = K_i\, k_6 /(k_5 + k_6)$ and when k_5 and k_6 are slower than all other steps [60].

According to mechanism B, in the presence of a fixed concentration of substrate (A), the value of k_{obs} increases hyperbolically as a function of inhibitor concentration with limiting values of k_6 and $k_5 + k_6$ at, respectively, zero and infinite concentrations of inhibitor. In mechanism A, k_{obs} varies as a linear function of the inhibitor concentration [60, 62]. The k_{obs} values obtained from progress curves of inhibitor concentrations up to 10 μM displayed a linear dependence on inhibitor concentration. However, when the K_i is much higher than K_i* and inhibitor

concentrations used in the assays are lower than the K_i (1 + A/K_a), k_{obs} values display a linear dependence on inhibitor concentration. Under these circumstances, the steady-state concentration of EI would be insignificant and it can be considered that E is converted directly to EI^* [60]. Accordingly, inactivation assays to differentiate between the two mechanisms were carried out in the presence of higher concentrations of $[Fe^{II}(CN)_5(INH)]^{3-}$. When the concentration of inhibitor is much larger than its apparent overall inhibition constant (K_i^*), v_s will be negligible and equation **1** simplifies to [63]:

$$A_t = A_0 - (v_i/k_{obs})(1 - \exp(-k_{obs}t)) \quad \textbf{[2]}$$

Thus, data from progress curves obtained for $[Fe^{II}(CN)_5(INH)]^{3-}$ in the concentration range of 15-1500 µM were fitted by nonlinear regression to equation **2**. The k_{obs} values obtained from equations **1** and **2** at, respectively, low and high inhibitor concentrations are consistent with mechanism B, since the k_{obs} values display, at fixed concentration of substrates, a hyperbolic dependence on $[Fe^{II}(CN)_5(INH)]^{3-}$ concentration (Fig. **6**), which can be described by the following equation [42]:

$$k_{obs} = k_6 + k_5(I/(K_i(1 + A/K_a + B/K_b) + I)) \quad \textbf{[3]}$$

where K_i is the dissociation constant for the initial enzyme-inhibitor binary complex (EI), k_5 is the forward isomerization rate constant for conversion of EI to EI^*, k_6 is the reverse isomerization rate constant (conversion of EI^* to EI), I is the concentration of inhibitor ($[Fe^{II}(CN)_5(INH)]^{3-}$), A and B are the concentration of the two competitive substrates (NADH and 2-*trans*-dodecenoyl-CoA), and K_a and K_b are the Michaelis-Menten constants for A and B, i.e. K_m for NADH and K_m for DD-CoA, respectively. In a previous report, the mechanism of action of WT InhA inactivation by $[Fe^{II}(CN)_5(INH)]^{3-}$ has been suggested to involve an interaction with the NADH binding site of the enzyme, since it has been demonstrated that both high concentrations of NADH protects WT InhA from inactivation by $[Fe^{II}(CN)_5(INH)]^{3-}$ and the apparent first-order rate constant obtained for the inactivation process is dependent on NADH concentration [56]. Thus, these results indicate that InhA enzyme inhibition by $[Fe^{II}(CN)_5(INH)]^{3-}$ appears to be competitive with NADH. To test if the DD-CoA substrate

protects against enzyme inactivation by $[Fe^{II}(CN)_5(INH)]^{3-}$, time-dependent inactivation experiments were performed, where WT InhA was preincubated with the inorganic complex either in the absence or in the presence of DD-CoA. All reactions were carried out at 25 °C in mixtures containing 100 mM Na_2HPO_4, pH 7.5, 3 μM WT InhA, 100 μM $[Fe^{II}(CN)_5(INH)]^{3-}$, in the absence of DD-CoA or in the presence of either 85 μM or 150 μM of DD-CoA. Higher concentrations of DD-CoA could not be used because of a decrease in the enzyme activity at concentrations of enoyl-CoA substrate above ~150 μM, possibly due to substrate inhibition or the formation of micelles at high concentrations of DD-CoA [30]. Control experiments contained no inhibitor. The residual enzyme activity in preincubation mixtures was determined by diluting an aliquot (10 μL) of the reaction mixture into the assay solution (1000 μL), which contained 100 mM Na_2HPO_4 (pH 7.5), 200 μM NADH, and 100 μM DD-CoA. Thus, the residual enzyme activity in preincubation studies could be assessed at appropriate intervals by measurements of initial velocity at 340 nm for 1 min. The apparent first-order rate constants for WT InhA inactivation (k_{inact}) could be calculated from a plot of residual enzyme activity against time (data not shown). Values of 0.25 ± 0.02 min^{-1}, 0.129 ± 0.009 min^{-1}, and 0.031 ± 0.001 min^{-1}, for DD-CoA concentrations of 0, 85, and 150 μM, respectively, were determined for k_{inact}. Since the apparent first-order rate constant values obtained for the enzyme inactivation decreased as a function of increasing DD-CoA concentration in the preincubation mixtures, the protection afforded by DD-CoA against enzyme inactivation suggests that the inhibition mechanism probably involves interaction with both NADH and DD-CoA binding sites. Accordingly, the WT InhA inhibition by $[Fe^{II}(CN)_5(INH)]^{3-}$ appears to be competitive with respect to both DD-CoA and NADH substrates. Michaelis-Menten constant values (K_m) of 66 μM and 46 μM for, respectively, NADH and DD-CoA have previously been determined for WT InhA in the same experimental conditions used in the slow-binding inhibition studies described here (42). Accordingly, these K_m values were used in the present work for K_i and K_i^* calculations. Values of 32 ± 3 μM and 0.41 ± 0.01 min^{-1} for, respectively, the rapidly reversible dissociation constant for the WT InhA-$[Fe^{II}(CN)_5(INH)]^{3-}$ binary complex (K_i) and for the forward rate constant (k_5) were determined from fitting the k_{obs} values to equation **3** (Fig. **6**).

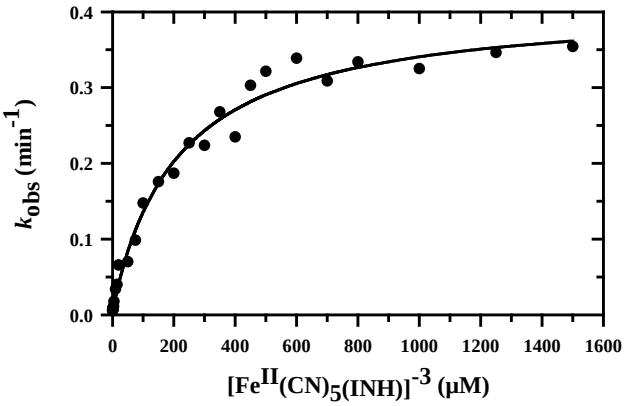


Fig. (6). Determination of dissociation constant (K_i) for the rapidly reversible EI complex and forward rate constant (k_5). The values of K_i and k_5 were obtained from plotting apparent pseudo-first order rate constants (k_{obs}) as a function of $[Fe^{II}(CN)_5(INH)]^{3-}$ and fitting the data to equation **3**. The experimental values for k_{obs} were determined from fits of progress curves to equations **1** and **2**.

A third mechanism (mechanism C) has also been described to account for slow-binding inhibition of enzyme catalyzed reactions (Fig. **5**). The mechanism C involves a slow initial isomerization of two forms of free enzyme in solution (E and E^*), followed by a rapid binding of inhibitor to form E^*I complex. The k_f and k_r stand for the forward and reverse rate constants for the conversion of free enzyme (E) to a form of free enzyme (E^*) to which the inhibitor binds. The three slow-binding mechanisms presented in Fig. (**5**) can be differentiated by the relationship between k_{obs}, and inhibitor concentration [64-65]. The k_{obs} predicted by mechanism C is given by equation 4, where $K_i = k_6/k_5$.

$$k_{obs} = k_f/(1 + [A]/K_a) + k_r K_i/([I] + K_i) \qquad \textbf{[4]}$$

The equation **4** predicts a decrease in k_{obs} with increasing inhibitor concentration. Therefore, in the present case, mechanism C was immediately ruled out since k_{obs} increases, rather than decreases, as a function of increasing $[Fe^{II}(CN)_5(INH)]^{3-}$ concentration.

5.5. Dissociation of $[Fe^{II}(CN)_5(INH)]^{3-}$ from WT InhA and overall inhibition constant (K_i*).

To demonstrate reversible binding and to obtain a reliable estimate for k_6 value, the reverse rate constant for the conversion of EI* to EI, WT InhA*-inhibitor complex (EI*) was formed by incubation of 25 μM WT InhA (concentration of active sites) with 25, 75 and 125 μM of $[Fe^{II}(CN)_5(INH)]^{3-}$ in 100 mM phosphate, pH 7.5 for 5 hours at 25°C. We have observed that, in the absence of NADH, this preincubation time is sufficient to allow the system reach equilibrium. After preincubation, samples were diluted by factors of 1:100 000 or 1:250 000 into assay mixtures containing 100 mM phosphate buffer (pH 7.5), 85 μM DD-CoA, 250 μM NADH, 8 % (vol/vol) glycerol and 0.1 mg mL^{-1} BSA, and absorbance at 340 nm monitored for 6 hours. The rate of activity regain of WT InhA was calculated by fitting the values of absorbance at 340 nm versus time to the equation **1** [60]. Control experiments contained no inhibitor. Dilution into a relatively large volume of assay mix containing near-saturating concentrations of one of the competitive substrate (NADH) causes inhibitor dissociation and regain of activity. Under these conditions the final inhibitor concentration in the assay mixtures and/or its effect are negligible and the rate of activity regain will provide k_6 [60]. Values of 1.1 (± 0.1) x 10^{-3} min^{-1}, 1.18 (± 0.09) x 10^{-3} min^{-1}, and 1.01 (± 0.03) x 10^{-3} min^{-1}, for inhibitor concentrations of 25, 75, and 125 μM, respectively, were determined for k_6. The similar values calculated for k_6 corroborate that both the inhibitor rebinding was insignificant and that the true reverse isomerization rate constant was estimated. Furthermore, as required, the k_6 value is independent of WT InhA concentration since enzyme dilutions of 1:100 000 or 1:250 000 yielded similar estimates. A typical activity regain experiment is presented in Fig. (**7**). A value of 630 ± 28 min was estimated for the half-time of the limiting step for inhibitor dissociation ($t_{1/2} = \ln 2/k_6$) from E*-$[Fe^{II}(CN)_5(INH)]^{3-}$ binary complex. A slow rate of dissociation is desirable, as it may be expected to enhance an inhibitor's effectiveness [66]. In addition, a value of 373 can be estimated for the equilibrium constant for the E-$[Fe^{II}(CN)_5(INH)]^{3-}$ to E*-$[Fe^{II}(CN)_5(INH)]^{3-}$ isomerization process (k_5/k_6 = 0.41 min^{-1}/1.1 x 10^{-3} min^{-1}). The ratio, k_5/k_6, can be taken as an index of the accumulation of EI* and the energetics of its formation. The higher the value of the k_5/k_6 ratio, the longer lived is EI* and more likely the inhibitor is to have a useful *in vivo* lifetime. The weakness in the use of classical enzyme inhibitors as drugs for clinical conditions is that inhibition results in the upstream accumulation of the substrate for the enzyme, which may overcome the inhibition. By

contrast, the buildup of substrate cannot have any effect on the isomerization of EI^* to EI and hence reversal of the inhibition [60].

The overall dissociation constant (K_i^*) is a function of the relative magnitudes of the forward (k_5) and reverse (k_6) rates for WT InhA*-$[Fe^{II}(CN)_5(INH)]^{3-}$ isomerization process (Fig. **5**). A K_i^* value of 86 ± 11 nM was estimated from the relationship $K_i^* = K_i k_6/(k_5+k_6)$ using an average k_6 value (1.1×10^{-3} min^{-1}) determined from the activity regain experiments. Alternatively, the K_i^* was determined by analysis of final steady-state velocities (v_s) versus inhibitor concentration (Fig. **8**) using the relationship $v_s = v_0/(1+I/K_i^{app})$, where v_0 is the initial rate in absence of inhibitor, I is the inhibitor concentration, and K_i^{*app} is the apparent overall dissociation constant value obtained in the presence of 250 μM NADH and 85 μM DD-CoA [62-63]. A true overall dissociation constant (K_i^*) value of 70 ± 10 nM was calculated by using the relationship $K_i^* = K_i^{*app}/(1+A/K_a+ B/K_b)$, where A is the NADH concentration, B is the DD-CoA concentration, and K_a and K_b are the Michaelis-Menten constants previously determined for, respectively, NADH (66 μM) and DD-CoA (46 μM) [42] in the same experimental conditions used in the assays described here. The K_i^* value estimated from regain of activity experiment (86 ± 11 nM) is within the standard error of the value determined from v_s versus inhibitor concentration plot (70 ± 10 nM), hence supporting the robustness of the data and the validity of the method used for analysis. A comparison between the inhibitor dissociation constant value for the rapidly reversible EI complex formation (K_i = 32 μM) and the overall dissociation constant value (K_i^* = 70 nM) indicates that the E^*-$[Fe^{II}(CN)_5(INH)]^{3-}$ is –3.6 kcal mol^{-1} more stable than the initially formed E-$[Fe^{II}(CN)_5(INH)]^{3-}$complex.

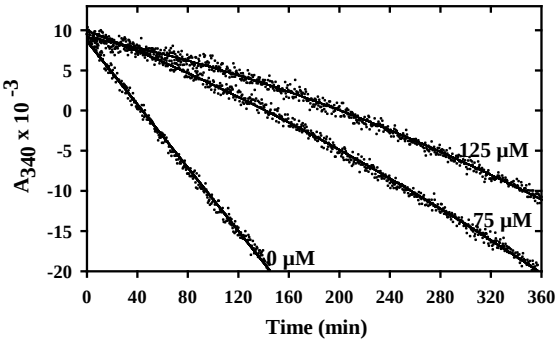


Fig. (7). Dissociation of $[Fe^{II}(CN)_5(INH)]^{3-}$ from WT InhA*. Enzyme (25 µM) was incubated with 75 µM and 125 µM of $[Fe^{II}(CN)_5(INH)]^{3-}$ for 5 h at 25 °C. Control experiment contained no inhibitor. Samples were diluted by a factor of 1:250 000, as exemplified here, into assay mixtures containing 100 mM phosphate buffer (pH 7.5), 85 µM DD-CoA, 250 µM NADH, 8% (vol/vol) glycerol, and 0.1 mg ml^{-1} BSA. WT InhA activity regain was continuously monitored at 340 nm for 6 h.

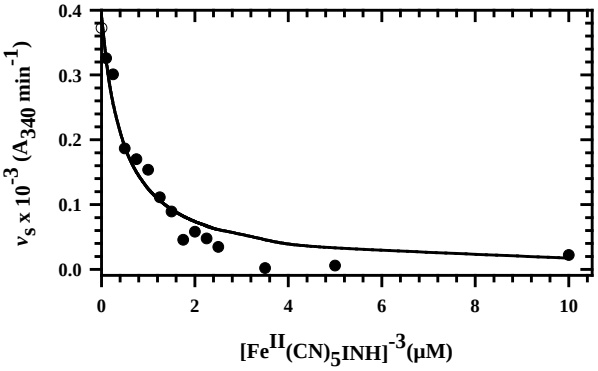


Fig. (8). Determination of the overall dissociation constant. K_i^* was determined from a replot of the final steady-state velocity (v_s) as a function of $[Fe^{II}(CN)_5(INH)]^{3-}$ concentration.

5.6. Two-step mechanism of WT InhA inhibition by $Fe^{II}(CN)_5(INH)]^{3-}$

The two-step mechanism of WT InhA inhibition by $[Fe^{II}(CN)_5(INH)]^{3-}$was confirmed in an experiment where the enzyme was preincubated in the presence of various concentrations of

inhibitor and the decrease of WT InhA activity was assessed by withdrawing aliquots of enzyme from preincubation mixtures at appropriate intervals, followed by rapid measurement of initial velocity. Reactions of time-dependent inactivation of WT InhA were carried out at 25 °C in mixtures containing 100 mM Na_2HPO_4, pH 7.5, 3 μM WT InhA, and various concentrations of $[Fe^{II}(CN)_5(INH)]^{3-}$. Control experiments contained no inhibitor. At various times, the residual enzyme activity was determined by diluting an aliquot (5 μL) of the reaction mixture into the assay solution (500 μL), which contained 100 mM Na_2HPO_4 buffer (pH 7.5), 200 μM NADH, and 100 μM DD-CoA. Since the value of K_i is much larger than K_i^* and the inhibitor concentration in the assay solution is lower than the K_i value ($[I]<K_i$) and there is a large excess of the competitive substrates NADH and DD-CoA over $[Fe^{II}(CN)_5(INH)]^{3-}$, the steady-sate concentration of EI in the assay solution is kinetically insignificant. In addition, the low reverse isomerization rate constant (k_6) value ($t_{1/2}$ = 630 ± 28 min) would result in a slow formation of EI from EI^* in a time range that is far removed from the time taken to make enzyme activity measurements, so that the EI^* concentration could be regarded as constant in the activity assay time range. Thus, the residual enzyme activity in preincubation studies could be assessed by measurements of initial velocity at 340 nm for 1 min. The apparent first-order rate constants for WT InhA inactivation (k_{inact}) could be calculated from a plot of residual enzyme activity against time for each inhibitor concentration. A plot of k_{inact} values versus inhibitor concentration shows that the k_{inact} values increase hyperbolically as a function of inhibitor concentration (Fig. **9**). These data were fitted to equation **5**, which is derived from equation **3** setting the concentration of substrates to zero ([NADH] = 0 and [DD-CoA] = 0). Equation **5** describes the slow-binding inhibition for mechanism B. Values of 0.51 ± 0.03 min^{-1} and 73 ± 14 μM were calculated for, respectively, the forward isomerization rate constant (k_5), and dissociation constant for formation of EI (K_i) These values are in reasonably good agreement with the ones obtained in slow-binding inhibition experiments, reinforcing a two-step mechanism for WT InhA inactivation by $[Fe^{II}(CN)_5(INH)]^{3-}$.

$$k_{obs} = k_6 + k_5(I/(K_i+I)) \qquad \textbf{[5]}$$

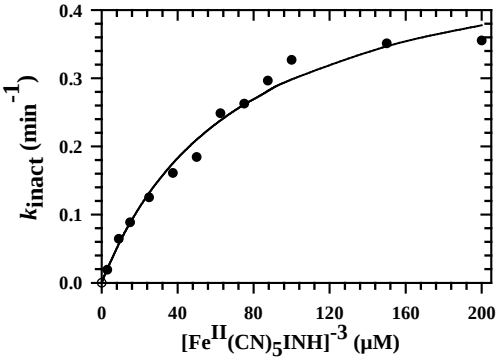


Fig. (9). Two-step inactivation of WT InhA by $[Fe^{II}(CN)_5(INH)]^{3-}$. The values for apparent first-order rate constant of inactivation (k_{inact}) were obtained from a plot of residual enzyme activity versus time (not shown). The k_{inact} values were plotted against $[Fe^{II}(CN)_5(INH)]^{3-}$ concentration in the preincubation mixture. The data were fitted to equation **5**.

The two-step mechanism followed by $[Fe^{II}(CN)_5(INH)]^{3-}$ inactivation of WT InhA involves an enzyme-inhibitor isomerization step. The enzyme isomerization could involve, for instance, a protein conformational change or alteration of water structure at the active site. For the inhibitor the change could be hydration/dehydration of a carbonyl, change in ionization state, or a conformational change. These alterations could take place in the enzyme inhibitor complex or in solution [66]. The slow-binding inhibition of WT InhA described here cannot be due to a slow isomerization of the inhibitor to a more potent species in solution, since we have previously shown that apparent first-order rate constant values obtained when $[Fe^{II}(CN)_5(INH)]^{3-}$ was pre-incubated for 2 hours before adding to a solution containing WT InhA was within the standard error of the one determined with no pre-incubation [56]. However, the results described above cannot distinguish between the various possible conformational changes of enzyme and/or inhibitor, changes in ionization states, or hydration/dehydration steps. Molecular modeling studies described below suggests though that $[Fe^{II}(CN)_5(INH)]^{3-}$ inhibitor molecule may undergo small conformational changes after initial docking in the active site of WT InhA enzyme.

6. Molecular modeling of mode of interaction between $[Fe^{II}(CN)_5(INH)]^{3-}$ and WT InhA

Complementary to the experimental studies on slow-binding inhibition kinetics presented and discussed above, the interaction of $[Fe^{II}(CN)_5(INH)]^{3-}$ ligand with WT InhA was modeled using computational docking, since no crystal structure of the inorganic compound-enzyme binary complex is available. The aim of computational docking is to evaluate the interaction energy between the ligand and a protein, giving an estimate of the free energy of binding, and predict the binding mode of the interacting molecules [67]. With the recent advances in search algorithms [68] and scoring functions [69], fast automated docking methods have become a valuable tool to explore the interaction between an enzyme and its inhibitors [70]. In the lack of experimental data, it contributes significantly to the understanding of the structural and energetic basis of protein-ligand interactions [71].

All docking simulations were performed using a force-field based free energy scoring function with the Lamarckian Genetic Algorithm (LGA) in the version 3.05 of the software AutoDock [72]. AutoDock rapidly evaluates the interaction energy by precalculating atomic affinity potentials for each atom type in the ligand, combined with fast search algorithms, to find possible binding positions on the protein. The structures were generated as follows:

Ligand: The structure of $[Fe^{II}(CN)_5(INH)]^{3-}$ was generated with the Molecular Builder module of Insight II (version 2000.1, Accelrys Inc.). In order to calculate AMBER force field charges, the geometry of $[Fe^{II}(CN)_5(INH)]^{3-}$ was optimized at the ab initio HF/6-31G* level. The optimized geometry was used to calculate the electrostatic potentials at the HF/6-31G* level. Gaussian98 program [73] was used for these calculations. The calculated electrostatic potentials were used to determine the charges by using the RESP [74] module of AMBER5 [75]. Then, the inorganic complex was processed with AutoTors to incorporate the apolar hydrogens to their respective heavy atoms and define the rotatable bonds when applicable. In this study $[Fe^{II}(CN)_5(INH)]^{3-}$ was kept rigid.

Proteins: The coordinates of WT InhA in complex with NADH (PDB 1ENY) (31) and with NAD^+ and a substrate analog (PDB 1BVR) [76] were taken from the Protein Data Bank [77].

Polar hydrogen atoms and partial charges were added in the EDIT module AMBER4.1 [78]. After adding solvation parameters with AddSol, the active site was defined by AutoGrid using default parameters. The number of grid points, spaced at 0.375 Å, was set to 61 x 61 x 61 generating a grid with approximate dimensions 23 x 23 x 23 $Å^3$, centered at the initial position of the ligand in the active site of the enzyme. This grid is large enough to include the NAD(H), as well as the C16 fatty acyl substrate binding pocket in 1BVR.

Each docking simulation was composed of 100 independent runs, for which a maximum number of 27,000 LGA was generated on the initial population of 50 individuals, a maximum number of 1.5 x 10^6 energy evaluations, with an elitism value of 1, a mutation rate of 0.02, and a cross-over rate of 0.8. For the local search the pseudo-Solis and Wets method was applied using default parameters. Each run provides one predicted binding mode. At the end of the docking experiment binding modes with root-mean-square deviation (RMSD) of 3.0 Å within each other were placed in the same cluster. The docking results are described separately for each crystal structure.

6.1. Docking of $[Fe^{II}(CN)_5(INH)]^{3-}$ to NADH bound enzyme (control experiment)

The 100 independent docking simulations for $[Fe^{II}(CN)_5(INH)]^{3-}$ in the NADH binding pocket, in the crystal structure (PDB code: 1ENY), resulted in 31 clusters with estimated average free energies of binding (ΔG_{bind}) ranging from -9.6 to -3.4 kcal/mol. Despite these differences, all 31 clusters showed a similar binding mode as depicted for the top ranked cluster in Fig. **10**. This docking simulation was performed as a control experiment to show, as expected, that $[Fe^{II}(CN)_5(INH)]^{3-}$ and the extended conformation of NADH cannot be simultaneously located in the crystallographic structure of the cofactor binding pocket.

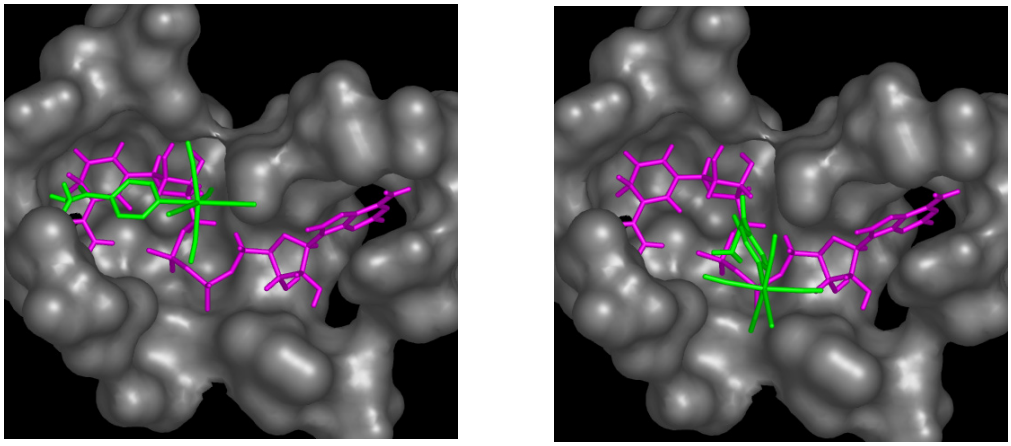


Fig. (10). A molecular surface representation of the NADH binding pocket in the WT InhA enzyme (PDB 1ENY) for all residues at least 5 Å from NADH (magenta). Left - the initial position of $[Fe^{II}(CN)_5(INH)]^{3-}$ (green) for all docking experiments in this work. Right - the binding mode of $[Fe^{II}(CN)_5(INH)]^{3-}$ for the lowest free energy of binding run in the top ranked cluster. The pentaferrocyanate portion of $[Fe^{II}(CN)_5(INH)]^{3-}$ is strongly repelled by the pyrophosphate moiety of NADH, which drives it outwards the crevice entrance.

6.2. Docking of $[Fe^{II}(CN)_5(INH)]^{3-}$ to 1ENY and 1BVR crystal structures.

The NAD(H) molecules make 25 residues contacts in 1ENY whilst 27 contacts are observed in 1BVR. The presence of the C16 fatty acyl substrate in the latter structure has resulted in a small conformational change that accounts for the difference in the contact numbers. A structural consequence of this change is a more compact NAD^+ binding pocket in 1BVR. Hence, by using these structures in the docking of $[Fe^{II}(CN)_5(INH)]^{3-}$ we are also implicitly considering the flexibility of the macromolecule. The docking simulation for $[Fe^{II}(CN)_5(INH)]^{3-}$ in the free cofactor binding pocket resulted in 11 and 10 clusters for the 1ENY and 1BVR crystal structures, respectively. Six clusters from each experiment were discarded for their binding mode did not take place in the active site and had higher energies. Only five distinct clusters were identified, each corresponding to one particular binding mode and ranked according to the lowest ΔG_{bind} as summarized in Table **1** and depicted in Figures **10** and **11**. It is important to note, however, that because there is yet no structural information about the inhibitor binding to InhA, a direct structural comparison is not possible. Instead, we focus on

analyzing the possible binding modes of $[Fe^{II}(CN)_5(INH)]^{3-}$ in InhA. In both structures the $[Fe^{II}(CN)_5(INH)]^{3-}$ inhibitor preferentially occupies the pyrophosphate and nicotinamide sites in the NAD(H) binding pockets (Table **1**). The $[Fe^{II}(CN)_5(INH)]^{3-}$ complex is bulky and the kinetic studies indicate that it can also occupy the NADH binding pocket, it is though not surprising that $[Fe^{II}(CN)_5(INH)]^{3-}$ makes approximately half of the number of contacts between protein and NADH. Furthermore, for the top 2 clusters in both targets the observed difference in pocket geometry and dimensions between 1ENY and 1BVR is reflected in the number of $[Fe^{II}(CN)_5(INH)]^{3-}$ -protein contacts.

1ENY: The pentacyanoferrate portion of cluster 1 and cluster 2 occupies the nicotinamide (NA) binding site, with the isoniazid moiety pointing towards the pyrophosphate (PP) in cluster 1 and upwards in cluster 2 (Figures **11b** and **c**), reducing cluster 2 contacts with the enzyme (Table **1**). Clusters 3, 4, and 7 (Figures **11d**, **e** and **f**) all have their pentacyanoferrate placed on top of PP. While in cluster 3 the isoniazid moiety points upwards, in cluster 4 and 7 it points in opposite directions, towards adenine ribose (ADR) and NA, respectively. Cluster 1 and 2 represent the best docking results, or the preferred binding modes of $[Fe^{II}(CN)_5(INH)]^{3-}$, as shown by their estimated free energy of binding and corresponding dissociation constant K_d (Table **1**).

1BVR: This structure is interesting for it comprises the conformations of the NAD^+ binding pocket and the active site of a complexed state. As mentioned above, it represents a distinct conformational state of InhA, thus implicitly accounting for some flexibility in the docking model. Consequently, different docking results for $[Fe^{II}(CN)_5(INH)]^{3-}$ were expected. In fact, because the entrance to the active site is slightly tighter than in 1ENY, the pentacyanoferrate portion cannot bind to the NA site as observed for clusters 1 and 2 of 1ENY. Instead, all binding modes are shifted away from the NA site resulting in an overall reduction of the inhibitor contacts with the protein, and a corresponding decrease in the fee energy of binding and affinity for the enzyme (Table **1**). A ubiquitous feature of these docking results is that the top ranking binding mode in 1ENY is similar to the bottom ranking binding mode in 1BVR, with the latter shifted away from the NA site. Conversely, the top ranking binding mode of 1BVR is similar to the bottom ranking binding mode in 1ENY. Clearly, this indicates that the flexibility of both protein and inhibitor must be considered for proper account of their interactions. Thus, the exact

position of $[Fe^{II}(CN)_5(INH)]^{3-}$ in the NAD(H) binding pocket of InhA cannot be assigned unambiguously. Nonetheless, these studies provide a rough description of its position and contacts. The position of $[Fe^{II}(CN)_5(INH)]^{3-}$ observed in the clusters of 1BVR docking overlap with the position occupied by the *trans* double bond of C16 fatty acyl substrate in the 1BVR crystal structure, which is adjacent to the nicotinamide ring of NAD^+ [76]. These results are consistent with the lower apparent first-order rate constant values for InhA inactivation observed in the presence of either NADH or DD-CoA.

Table 1. Results of 100 independent docking runs for $[Fe^{II}(CN)_5(INH)]^{3-}$ in the free NAD(H) binding pocket of the crystallographic structures of WT InhA, 1ENY and 1BVR. For 1BVR the C16 fatty acyl substrate also was removed.

Target	N_{tot}	Cluster	N	(ΔG_{bind}) kcal/mol	K_d (nM)	Number of contacts	NAD(H) Sites		
							NA	PP	ADR
1ENY	11	1	6	-13.4	0.2	14	$[Fe^{II}(CN)_5]^{3-}$	(INH)	-
		2	5	-12.0	1.5	12	$[Fe^{II}(CN)_5]^{3-}$	-	-
		3	2	-10.4	24.3	11	-	$[Fe^{II}(CN)_5]^{3-}$	-
		4	1	-9.7	75.8	11	-	$[Fe^{II}(CN)_5]^{3-}$	(INH)
		7	5	-8.7	455.0	10	(INH)	$[Fe^{II}(CN)_5]^{3-}$	-
1BVR	10	1	12	-10.0	46.7	13	(INH)	$[Fe^{II}(CN)_5]^{3-}$	-
		2	14	-9.9	57.2	10	-	$[Fe^{II}(CN)_5]^{3-}$	(INH)
		3	1	-9.4	129.0	11	-	(INH)	-
		4	5	-9.3	144.0	11	-	(INH)	-
		7	1	-8.8	371.0	12	$[Fe^{II}(CN)_5]^{3-}$	(INH)	-

[a] N_{tot} is the total number of clusters; N is the number of results for each cluster; ΔG_{bind} is the estimated lowest free energy of binding for each cluster. K_d is the estimated inhibition or dissociation constant. The seventh column shows the number of contacting residues for each binding mode. Columns 8, 9, and 10 show where the pentaferrocyanate and isoniazid parts of the $[Fe^{II}(CN)_5(INH)]^{3-}$ ligand sit in the free NADH (NAD^+) binding pocket (NA = nicotinamide; PP = pyrophosphate; ADR = adenine ribose).

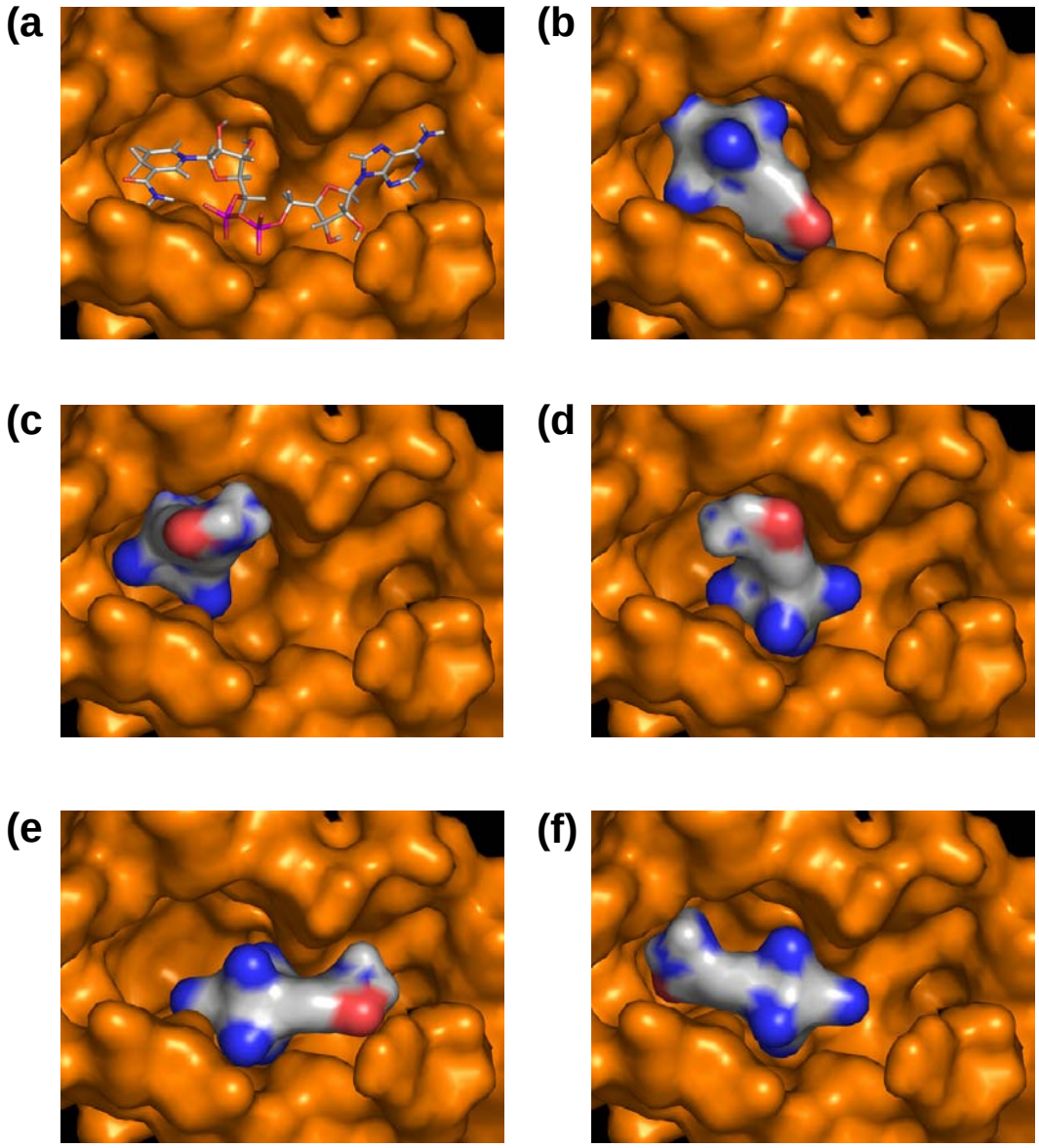


Fig. (11). The 5 top-ranked binding modes for $[Fe^{II}(CN)_5(INH)]^{3-}$ in the NADH binding pocket, which are represented as molecular surfaces, of the 1ENY crystal structure. The atoms are colored as follows: carbon and hydrogen (grey), nitrogen (blue), oxygen (red), phosphorus (magenta). (a) NADH in its pocket from the 1ENY crystal structure, (b) cluster 1, (c) cluster 2, (d) cluster 3, (e) cluster 4, and (f) cluster 5.

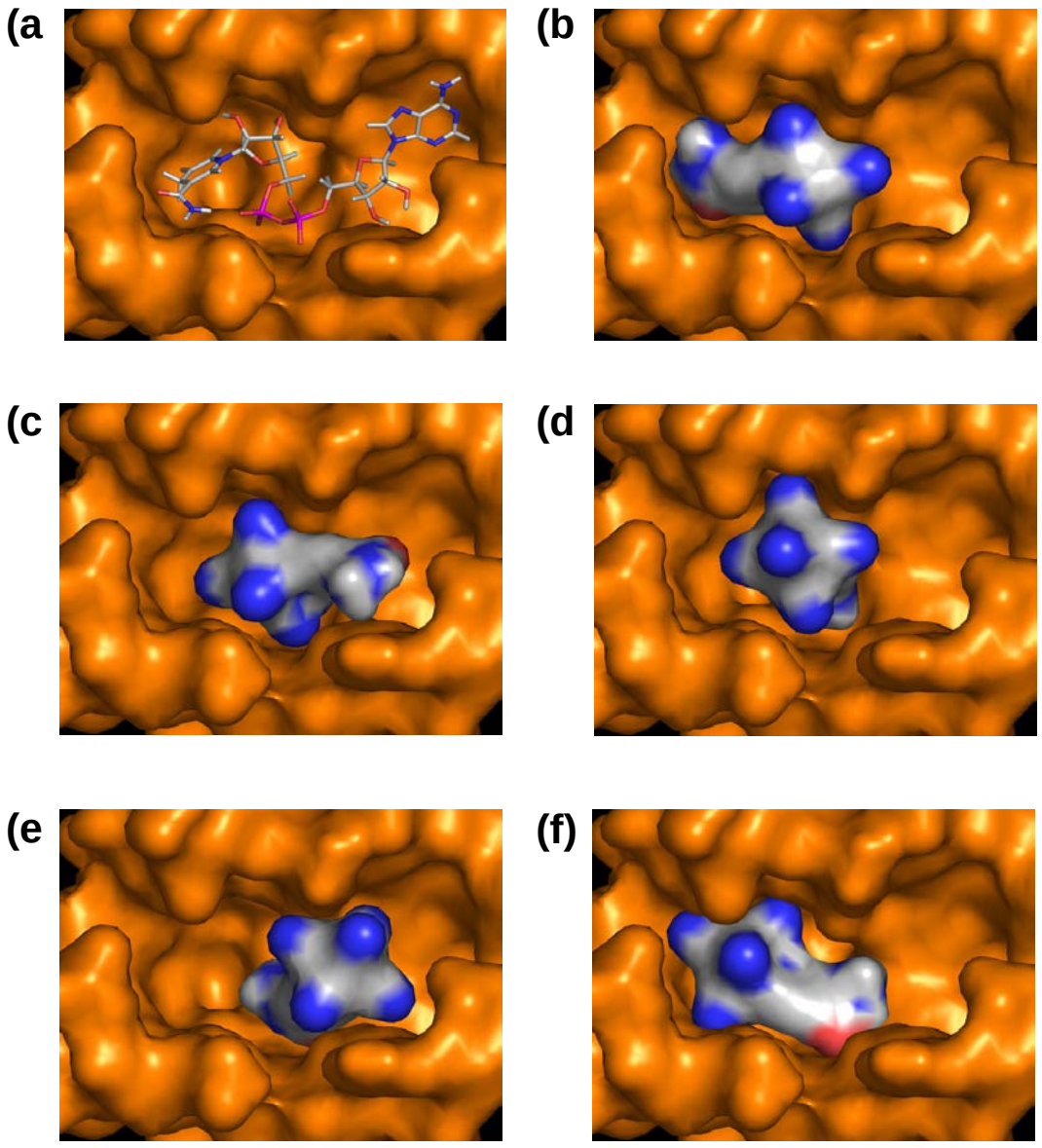


Fig. (12). The 5 top-ranked binding modes for $[Fe^{II}(CN)_5(INH)]^{3-}$ in the NAD^+ binding pocket, which are represented as molecular surfaces, of the 1BVR crystal structure. The atoms are colored as in Fig. **10**. (a) NAD^+ in its pocket from the 1BVR crystal structure, (b) cluster 1, (c) cluster 2, (d) cluster 3, (e) cluster 4, and (f) cluster 5.

7. Concluding remarks

The determination of the complete genome sequence of *M. tuberculosis* [79] has had a striking impact on researchers in the TB field. Comparison of the complete genome sequence of *M. tuberculosis* H37Rv and the clinical isolate CDC1551, which is highly infectious in humans, revealed a number of polymorphisms with potential relevance to disease pathogenesis, immunity and evolution [80]. The *M. leprae* genome represents the minimal mycobacterial genome as it contains a large number of pseudo-genes, the remnants of genes which were once functional [81]. Intact gene family members found in the *M. leprae* genome may point to the genes that are essential for survival of other mycobacteria. More recently, the complete genome sequence of *M. bovis* has shown that it is >99.95 % identical to *M. tuberculosis* and that there are no genes unique to *M. bovis* [82], thereby implying that differential gene expression may be the key to the host tropisms of human and bovine bacilli. One of the expected benefits of genome analysis of pathogenic bacteria is in the area of human health, particularly in the design of more rapid diagnostic reagents and development of new vaccines and antimicrobial agents. New approaches are being sought including comparative genomics using bioinformatics, functional genomics, proteomics, transcriptomics, and structural genomics. It is hoped that a better understanding of the complex biology of this pathogen will yield new drug targets for the drug development process. In addition, studies being conducted by the TB Structural Genomics consortium will make available the structures of many potential target proteins and should provide important information such as the biological function of proteins harboring no homology to any known protein [83]. A number of these proteins must be selected for further analysis and computational chemistry techniques applied to identify leads.

As pointed out above, there is an urgent need for more effective and less toxic anti-tubercular agents to shorten the duration of current treatment, improve the treatment of MDR-TB, and to provide effective treatment of latent tuberculosis infection. It thus appears reasonable to examine the potential antitubercular activity of INH analogs due to oral availability and favorable toxicology profile of INH. Moreover, InhA is a validated target and inhibitors of its activity could provide a faster route towards the development of anti-TB agents, while new targets are yet to be unveiled and validated. In addition, there is a cyclical need for innovation to

provide new structural scaffolds (lead compounds) to undergo chemical iteration aiming at optimization before reaching candidate drug status. Here we describe a new approach to the rational design of an INH analog based on an inorganic atom attached to the nitrogen atom of the heterocyclic ring of isoniazid that inhibits a validated target (InhA). The pentacyano(isoniazid)ferrateII compound appears to be a promising candidate for further antitubercular drug development and may represent a new class of lead compounds. Efforts to determined the three-dimensional structure of $[Fe^{II}(CN)_5INH]$•WT InhA binary complex are currently underway to confirm the binding mode and to assess the molecular interactions made with the protein active site residues. In addition, evaluation of the ability of this inorganic complex to kill *M. tuberculosis* in mouse bone marrow macrophages and to inhibit mycobacterial growth in animal models are essential steps currently underway to establish the pentacyano(isoniazid)ferrateII complex as a potential drug candidate.

Screening threshold is typically set in the range of 1-10 μM for ligands to produce hits to undergo chemical modifications to construct a synthetic library to provide a population of molecules that have a chance of moderate- to high-affinity recognition by specific regions of target macromolecules [55]. Architectural diversity or functional group density are desirable features to foster an increase in new structural leads to provide quality not quantity. In this respect, the inorganic complex here described represents a step towards this goal, since different metal centers and functional groups may be introduced to improve its efficacy, if needed. In addition the overall dissociation constant value of 86 nM for inhibition of InhA by the pentacyano(isoniazid)ferrateII compound represents an excelent starting point from which to improve efficacy.

Both *M. tuberculosis* and *M. bovis* BCG are susceptible to isoniazid in the range of 0.02-0.2 μg mL^{-1} [Heifets, 1994]. An MIC value of 0.2 μg mL^{-1} for the pentacyano(isoniazid)ferrate was determined by the radiometric BACTEC system [56]. Preliminary *in vitro* cytotoxicity results in HL60 leukemia and MCS-7 breast cancer cells indicate an $IC_{50} > 25$ μg mL^{-1}, which yields a favorable selectivity index ($SI = IC_{50}/MIC > 125$). The molecular mass of most of the orally active drugs is less than 500 daltons, and biologically relevant structural features would have to be achieved within this size constraint if oral activity of a new chemotherapeutic agent is

to be routinely achieved [84]. The $Na_3[Fe^{II}(CN)_5INH]\bullet 3H_2O$ compound is stable at acidic pH and its molecular mass is 446 daltons, thus imparting favorable characteristics to this compound as a potential oral drug.

Metal-based chemotherapeutic compounds have been investigated for potential medicinal application; for instance, the platinum-based cisplatin is widely used in the treatment of cancer [85]. Metal complexes have also been described as alternative chemoterapheutic agents to treat tropical diseases as malaria, trypanosomiasis, and leishmaniasis [86]. The most successful examples of metal-based anitmalarials are derived from modifying the most widely used drug chloroquine (CQ) through coordination to metal-containing fragments. The *in vitro* and *in vivo* tests of CQ complexes of transition metals (Rh, Ru) against *Plasmodium berghei* showed that the incorporation of the metal fragments generally produce an enhancement of the efficacy of chloroquine [87]. More recently, gold-chloroquine complexes were shown to be active against CQ-resistant and CQ-sensitive strains of *Plasmodium falciparum* [88-89]. The highest activity for this series was obtained for $[(CQ)Au(PPh_3)][PF_6]$, which was found to be 9 times more active than chloroquine diphosphate against the CQ-resistant FcB1 strain of *P. falciparum*, 5 times more active for the more aggressive strain FcB2 strain and 22 times more active against the rodent malaria parasite *P. berghei* [86]. In particular, metal complexes of carboxamidrazone analogs resulted in an enhancement in their *in vitro* antitubercular activity against *M. tuberculosis* H37Rv strain [90]. However, the target(s) of these metal complexes are unknown and this approach could hardly be considered as a rational design of a drug.

As pointed out by Gerard Jaouen (chemistry professor at Ecole Nationale Superieure de Chimie de Paris), the bio-inorganic chemistry discipline “needs to prove it can create something useful for society” [91]. There is little doubt that other new- metal-contaning drugs will be developed in a variety of therapeutic areas, including the treatment of parasitic diseases, as the field of inorganic medicinal chemistry continues to grow. We hope that the approach described here will bear fruit and help the development of new antitubercular agents.

Acknowledgments

Financial support for this work was provided by The Department of Science and Technology (DECIT)/ Ministry of Health and UNESCO and Millennium Initiative Program MCT-CNPq to DSS and LAB. DSS (304051/1975-06) and LAB (520182/99-5) are research career awardees from the National Research Council of Brazil (CNPq). ONS acknowledges a PROADI grant from FAPERGS (01/0554.9) and CAPES (PI 18-23/2001). We, the authors, would like to acknowledge Chang-Guo Zhan for providing charge parameters of the pentacyano(isoniazid)ferrateII inorganic complex needed for the molecular modeling studies.

References

[1] Raviglione MC. The TB epidemic from 1992 to 2002. Tuberculosis 2003; 83: 4-14.

[2] Corbett EL, Watt CJ, Walker N, Maher D, Williams BG, Raviglione MC, Dye C. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. Arch Intern Med 2003; 163: 1009-21.

[3] Morens DM, Folkers, GK, Fauci, AS. The challenge of emerging and re-emerging infectious diseases. Nature 2004; 430: 242-248.

[4] Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. JAMA 1999; 282: 677-86.

[5] Espinal MA. The global situation of MDR-TB. Tuberculosis 2003; 83: 44-51.

[6] World Health Organization. Anti-tuberculosis drug resistance in the world: Third Global Report. 2004.

[7] Nunn P. WHO Global TB Programme, Tuberculosis Research and Surveillance Unit of the WHO Global TB Programme; Press Release WHO/74 (www.who.int/gtb/press/who74.htm), 1997.

[8] Blower SM, Chou T. Modeling the emergence of the 'hot zones': tuberculosis and the amplification dynamics of drug resistance. Nat Med 2004; 10: 1111-16.

[9] McKinney JD, Jacobs WRJr., Bloom BR. In: Krause RM Ed, Emerging Infections. New York, Academic Press. 1998; 51-146.

[10] Pablos-Méndez A, Gowda DK, Frieden TR. Controlling multidrug-resistant tuberculosis and access to expensive drugs: a rational framework. Bull. World Health Organ 2002; 80: 489-95.

[11] Duncan K. Progress in TB drug development and what is still needed. Tuberculosis 2003; 83: 201-7.

[12] Hingley-Wilson SM, Sambandamurthy VK, Jacobs WRJr. Survival perspectives from the world's most successful pathogen, Mycobacterium tuberculosis. Nat Immunol 2003; 4: 949-55.

[13] Gomez JE, McKinney JD. M. tuberculosis persistence, latency, and drug tolerance. Tuberculosis 2004; 84: 29-44.

[14] Schatz A, Bugie E, Waksman SA Streptomycin, a substrate exhibiting antibiotic activity against gram-positive and gram-negative bacteria. Proc Soc Exp Biol Med 1944; 55: 66.

[15] Middlebrook G. Sterilization of tubercle bacilli by isonicotinic acid hydrazide and the incidence of variants resistant to the drug *in vitro*. Am Rev Tuberc 1952; 65: 765-7.

[16] Kushner S, Dalalian H, Sanjurjo JL, Bach FL, Safir SR, Smith VKJr, Williams JH Experimental chemotherapy of tuberculosis. The synthesis of pyrazinamides and related compounds. J Am Chem Soc 1952; 74: 3617-21.

[17] Basso LA, Blanchard JS. Resistance to antitubercular drugs. Adv Exp Med Biol 1998; 456: 115-44.

[18] Mitchison DA The action of antituberculosis drugs in short-course chemotherapy. Tubercle 1985; 66: 219-25.

[19] Heifets LB. Antimycobacterial drugs. Semin Respir Infect 1994; 9: 84-103

[20] Ramaswamy S, Musser JM. Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. Tuber Lung Dis. 1998, 79: 3-29.

[21] Glickman MS, Jacobs WRJr. Microbial pathogenesis of Mycobacterium tuberculosis: dawn of a discipline. Cell 2001; 104: 477-85.

[22] Schroeder EK, de Souza ON, Santos DS, Blanchard JS, Basso LA. Drugs that inhibit mycolic acid biosynthesis in *Mycobacterium tuberculosis*. Curr Pharm Biotechnol 2002; 3: 197-225.

[23] Dutt AK, Stead W. The treatment of tuberculosis. Dis. Month 1994; 43: 247-74.

[24] Blanchard JS. Molecular mechanisms of drug resistance in Mycobacterium tuberculosis. Annu Rev Biochem 1996; 65: 215-39.

[25] Banerjee A, Dubnau E, Quémard A, Balasubramanian V, Um KS, Wilson T, Collins D, de Lisle G, Jacobs WRJr. inhA, a Gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis. Science 1994; 263: 227-30.

[26] Quémard A, Sacchettini JC, Dessen A, Vilchèze C, Bittman R, Jacobs WRJr, Blanchard JS. Enzymatic characterization of the target for isoniazid in Mycobacterium tuberculosis. Biochemistry 1995; 34: 8235-41.

[27] Vilchèze C, Morbidoni HR, Weisbrod TR, Iwamoto H, Kuo M, Sacchettini JC, Jacobs WRJr. Inactivation of the inhA-encoded fatty acid synthase II (FASII) enoyl-acyl carrier protein reductase induces accumulation of the FASI end products and cell lysis of Mycobacterium smegmatis. J Bacteriol 2000; 182: 4059-67.

[28] Larsen MH, Vilchèze C, Kremer L, Besra GS, Parsons L, Salfinger M, Heifets L, Hazbon MH, Alland D, Sacchettini JC, Jacobs WRJr. Overexpression of inhA, but not kasA, confers

resistance to isoniazid and ethionamide in Mycobacterium smegmatis, M. bovis BCG and M. tuberculosis. Mol Microbiol 2002; 46: 453-66.

[29] Kremer L, Dover LG, Morbidoni HR, Vilchèze C, Maughan WN, Baulard A, Tu S-C, Honoré N, Deretic V, Sacchettini JC, Locht C, Jacobs WRJr, Besra GS. Inhibition of InhA activity, but not KasA activity, induces formation of a KasA-containing complex in mycobacteria. J Biol Chem 2003; 278: 20547-54.

[30] Basso LA, Zheng R, Musser JM, Jacobs WRJr, Blanchard JS. Mechanism of isoniazid resistance in Mycobacterium tuberculosis: enzymatic characterization of enoyl reductase mutants identified in isoniazid-resistant clinical isolates. J Infect Dis 1998; 178: 769-75.

[31] Dessen A, Quémard A, Blanchard JS, Jacobs WRJr, Sacchettini JC. Crystal structure and function of the isoniazid target of Mycobacterium tuberculosis. Science 1995; 267: 1638-41.

[32] Schroeder EK, Basso LA, Santos DS, de Souza ON. Molecular dynamics simulation studies of the wild-type, I21V and I16T mutants of isoniazid resistant Mycobacterium tuberculosis enoyl reductase (InhA) in complex with NADH: towards the understanding of NADH-InhA different affinities. Biophys J 2005 (accepted for publication, Epud ahead of print, PMID:15908576).

[33] Heym B, Alzari PM, Honoré N, Cole ST. Missense mutations in the catalase-peroxidase gene, katG, are associated with isoniazid resistance in Mycobacterium tuberculosis. Mol Microbiol 1995; 15: 235-45.

[34] Ramaswamy SV, Reich R, Dou S-J, Jasperse L, Pan X, Wanger A, Quitugua T, Graviss EA. Single nucleotide polymorphisms in genes associated with isoniazid resistance in Mycobacterium tuberculosis. Antimicrob Agents Chemother 2003; 47: 1241-50.

[35] Yu S, Girotto S, Lee C, Magliozzo RS. Reduced affinity for isoniazid in the S315T mutant of Mycobacterium tuberculosis KatG is a key factor in antibiotic resistance. J Biol Chem 278: 14769-75.

[36] Johnsson K, Schultz PG. Mechanistic studies of the oxidation of isoniazid by the catalase peroxidase from Mycobacterium tuberculosis. J AmChem Soc 1994; 116: 7425-6.

[37] Johnsson K, King DS, Schultz PG. Studies on the mechanism of action of isoniazid and ethionamide in the chemotherapy of tuberculosis. J Am Chem Soc 1995; 117: 5009-10.

[38] Basso LA, Zheng R, Blanchard JS. Kinetics of inactivation of WT and C243S mutant of Mycobacterium tuberculosis enoyl reductase by activated isoniazid. J Am Chem Soc 1996; 118: 11301-2.

[39] Zabinski RF, Blanchard JS. The requirement for manganese and oxygen in the isoniazid-dependent inactgivation of Mycobacterium tuberculosis enoyl reductase. J Am Chem Soc 1997; 119: 2331-2.

[40] Rozwarski DA, Grant GA, Barton DHR, Jacobs WRJr, Sacchettini JC. Modification of the NADH of the isoniazid target (InhA) from Mycobacterium tuberculosis. Science 1998; 279: 98-102.

[41] Lei B, Wei C-J, Tu S-C. Action mechanism of antitubercular isoniazid. J Biol Chem 2000; 275: 2520-6.

[42] Rawat R, Whitty A, Tonge PJ. The isonizid-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 2003; 100: 13881-6.

[43] Johnsson K, Froland W.A., Schultz PG. Overexpression, purification, and characterization of the catalase-peroxidase KatG from Mycobacterium tuberculosis. J Biol Chem 1997; 272: 2834-40.

[44] Chouchane S, Lippai I, Magliozzo RS. Catalase-peroxidase (Mycobacterium tuberculosis KatG) catalysis and isoniazid activation. Biochemistry 2000; 39: 9975-83.

[45] Chouchane S, Girotto S, Yu S, Magliozzo RS. Identification and characterization of tyrosyl radical formation in Mycobacterium tuberculosis catalase-peroxidase (KatG). J Biol Chem 2002; 277: 42633-8.

[46] Zhao X, Girotto S, Yu S, Magliozzo RS. Evidence for radical formation at Tyr-353 in Mycobacterium tuberculosis catalase-peroxidase (KatG). J Biol Chem 2004; 279: 7606-12.

[47] Bertrand T, Eady NAJ, Jones JN, Jesmin, Nagy JM, Jamart-Grégoire B, Raven EL, Brown KA. Crystal structure of Mycobacterium tuberculosis catalase-peroxidase. J Biol Chem 2004; 279: 38991-9.

[48] Pierattelli R, Banci L, Eady NAJ, Bodiguel J, Jones JN, Moody PCE, Raven EL, Jamart-Grégoire B, Brown KA. Enzyme-catalyzed mechanism of isoniazid activation in Class I and Class III peroxidases. J Biol Chem 2004; 279: 39000-9.

[49] Magliozzo RS, Marcinkeviciene JA. The role of Mn(II)-peroxidase activity of mycobacterial catalase-peroxidase in activation of the antibiotic isoniazid. J Biol Chem 1997; 272: 8867-70.

[50] Wilming M, Johnsson K. Spontaneous formation of the bioactive form of the tuberculosis drug isoniazid. Angew Chem Int Ed 1999; 38: 2588-90.

[51] Nguyen M, Quémard A, Marrakchi H, Bernadou J, Meunier B. The nonenzymatic activation of isoniazid by Mn^{III}-pyrophosphate in the presence of NADH produces the inhibition of the enoyl-ACP reductase InhA from Mycobacterium tuberculosis. C R Acad Sci Paris, Série IIc, Chimie 2001; 4: 35-40.

[52] Scior T, Morales IM, Eisele SJG, Domeyer D, Laufer S. Antitubercular isoniazid and drug resistance of Mycobacterium tuberculosis – a review. Arch Pharm Pharm Med Chem 2002; 11: 511-25.

[53] Mohamad S, Ibrahim P, Sadikun A. Susceptibility of Mycobacterium tuberculosis to isoniazid and its derivative, 1-isonicotinyl-2-nonanoyl hydrazine: investigation at cellular level. Tuberculosis 2004; 84: 56-62.

[54] Hearn MJ, Cynamon MH. Design and synthesis of antituberculars: preparation and evaluation against Mycobacterium tuberculosis of an isoniazid Schiff base. J Antimicrob Chemother 2004; 53: 185-191.

[55] Walsh C. Where will new antibiotics come from? Nature Rev Microbiol 2003; 1: 65-70.

[56] Oliveira JS, Sousa EHS, Basso LA, Palaci M, Dietze R, Santos DS, Moreira IS. An inorganic iron complex that inhibits wild-type and an isoniazid-resistant mutant 2-trans-enoyl-ACP (CoA) reductase from Mycobacterium tuberculosis. Chem Commun 2004; 3: 312-3.

[57] DeBarber A, Mdluli K, Bosman M, Bekker L-G, Barry CEIII. Ethionamide activation and sensitivity in multidrug-resistant Mycobacterium tuberculosis. Proc Natl Acad Sci USA. 2000; 97: 9677-82.

[58] Moreira IS, Franco DG. Ruthenium(II) and iron(II) complexes of 4,4-dithiodipyridine. Synthesis, characterization, and reactivity studies. Inorg. Chem. 1994; 33: 1607-13.

[59] Moreira IS, Lima JB, Franco DW. Manifestation of sulfur-to-sulfur interaction in complexes of iron, ruthenium and osmium. Coord Chem Rev. 2000; 196: 197-217.

[60] Morrison JF, Walsh CT. The behaviour and significance of slow-binding enzyme inhibitors. Adv Enzymol Relat Areas Mol Biol 1988; 61: 201-301.

[61] Liu W, Rogers CJ, Fisher AJ, Toney MD. Aminophosphonate inhibitors of dialkylglycine decarboxylase: structural basis for slow binding inhibition. Biochemistry 2002; 41:12320-8.

[62] Morrison JF, The slow-binding and slow, tight-binding inhibition on enzyme-catalysed reactions. Trends Biochem Sci 1982, 9: 102-105.

[63] Stone SR, Hermans JM. Inhibitory mechanism of serpins. Interaction of thrombin with antithrombin and protease nexin 1. Biochemistry 1995; 34: 5164-72.

[64] Erion MD, Walsh CT. 1-Aminocyclopropanephosphonate: Time-dependent inactivation of 1-aminocyclopropanecarboxylate deaminase and *Bacillus stearothermophilus* alanine racemase by slow dissociation behaviour. Biochemistry 1987; 26: 3417-25.

[65] Shapiro R, Riordan JF. Inhibition of angiotensin converting enzyme: Dependence on chloride. Biochemistry 1984; 23: 5334-40.

[66] Schoss JV. Significance of slow-binding enzyme inhibition and its relationship to reaction-intermediate analogues. Acc Chem Res 1988; 21: 348-53.

[67] Sotriffer CA, Flader W, Winger RH, Rode BM, Liedl KR, Varga JM. (2000), Automated docking of ligands to antibodies: methods and applications. Methods: A Companion to Methods in Enzymology 2000a; 20: 280-91.

[68] Thomsen R. Flexible ligand docking using evolutionary algorithms: investigating the effects of variation operators and local search hybrids. Biosystems 2003; 72: 57-73.

[69] Stahl M, Rarey M. Detailed analysis of scoring functions for virtual screening. J Med Chem 2001; 29: 1035-42.

[70] Sotriffer CA, Ni HH, McCammon JA. Active site binding modes of HIV-1 integrase inhibitors. J. Med. Chem 2000; 43: 4109-17.

[71] Brooijmans N, Kuntz ID. Molecular recognition and docking algorithms. Annu Rev Biophys Biomol Struct 2003; 32: 335-73.

[72] Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian Genetic Algorithm and and empirical binding free energy function. J Computational Chemistry 1998; 19:1639-62.

[73] Frisch M J, Trucks G W, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Stratmann RE, Burant JC, Dapprich S, Millam J M, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez AC, Head-Gordon M, Replogle ES, Pople JA. Gaussian 98, Revision A.6, Gaussian, Inc.: Pittsburgh, PA, 1998.

[74] Bayly CI, Cieplak P, Cornell WD, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges – the RESP model. J Phys Chem 1993; 97: 10269-80.

[75] Case DA, Pearlman DA, Caldwell JW, Cheatham I, Ross WS, Simmerling C, Darden T, Merz KM, Stanton RV, Cheng A, Vincent JJ, Crowley M, Ferguson DM, Radmer R, Seibel GL, Singh US, Weiner PK, Kollman PA. AMBER, 5.0; University of California: San Francisco, 1997.

[76]Rozwarski DA, Vilchèze C, Sugantino M, Bittman R, Sacchettini JC. Crystal structure of the Mycobacterium tuberculosis enoyl-ACP reductase, InhA, in complex with NAD^+ and a C16 fatty acyl substrate. J Biol Chem 1999; 274: 15582-89.

[77] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000; 28: 235-42.

[78] Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, Ferguson DM, Seibel GL, Singh UC, Weiner PK, Kollman PA. AMBER 4.1, University of California, San Francisco, 1995.

[79] Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, 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 MA, Rajandream M-A, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 1998; 393: 537-44.

[80] Fleischmann RD, Alland D, Eisen JA, Carpenter L, White O, Peterson WJ, DeBoy R, Dodson R, Gwinn M, Haft D, Hickey E, Kolonay JF, Nelson WC, Umayam LA, Ermolaeva M, Salzberg SL, Delcher A, Utterback T, Weidman J, Khouri H, Gill J, Mikula A, Bishai W, Jacobs WRJr, Venter JC, Fraser CM. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. J Bacteriol 2002; 184: 5479-90.

[81] Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holroyd S, Hornsby T, Jagels K, Lacroix C, MaClean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream MA, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JB, Barrell BG. Massive gene decay in the leprosy bacillus. Nature 2001; 409: 1007-11.

[82] Garnier T, Eiglmeier K, Camus J-C, Medina N, Mansoor H, Pryor M, Duthoy S, Grondin S, Lacroix C, Monsempe C, Simon S, Harris B, Atkin R, Doggett J, Mayes R, Keating L, Wheeler

PR, Parkhill J, Barrell BG, Cole ST, Gordon SV, Hewinson,RG. The complete genome sequence of *Mycobacterium bovis*. Proc Natl Acad Sci USA 2003; 100: 7877-82.

[83] Terwilliger TC, Park MS, Waldo GS, Berendzen J, Hung L-W, Kim C-Y, Smith CV, Sacchettini JC, Bellinzoni M, Bossi R, De Rossi E, Mattevi A, Milano A, Riccardi G, Rizzi M, Roberts MM, Coker AR, Fossati G, Mascagni P, Coates ARM, Wood SP, Goulding CW, Apostol MI, Anderson DH, Gill HS, Eisenberg DS, Taneja B, Mande S, Pohl E, Lamzin V, Tucker P, Wilmanns M, Colovos C, Meyer-Klaucke W, Munro AW, McLean KJ, Marshall KR, Leys D, Yang JK, Yoon H-J, Lee BI, Kwak JE, Han BW, Lee JY, Baek S-H, Suh SW, Komen MM, Arcus VL, Baker EN, Lott JS, Jacobs WRJr, Alber T, Rupp B. The TB structural genomics consortium: a resource for *Mycobacterium tuberculosis* biology. Tuberculosis 2003; 83: 223-49.

[84] Teague SJ, Davis, AM, Leeson PD, Oprea T. The design of leadlike combinatorial libraries. Angew Chem Int Ed Engl 1999; 38: 3743-3748.

[85] Zhang CX, Lippard SJ. New metal complexes as potential therapeutics. Curr Opin Chem Biol 2003; 7:481-9.

[86] Sanchez-Delgado RA, Anzellotti A. Metal complexes as chemotherapeutic agents against tropical diseases: trypanosomiasis, malaria and leishmaniasis. Mini Rev Med Chem. 2004; 4: 23-30.

[87] Sanchez-Delgado RA, Navarro M, Perez H, Urbina JA. Toward a novel metal-based chemotherapy against tropical diseases. 2. Synthesis and antimalarial activity in vitro and in vivo of new ruthenium- and rhodium-chloroquine complexes. J Med Chem. 1996; 39:1095-99.

[88] Navarro M, Perez H, Sanchez-Delgado RA. Toward a novel metal-based chemotherapy against tropical diseases. 3. Synthesis and antimalarial activity in vitro and in vivo of the new gold-chloroquine complex [Au(PPh3)(CQ)]PF6. J Med Chem. 1997; 40: 1937-9.

[89] Navarro M, Vasquez F, Sanchez-Delgado RA, Perez H, Sinou V, Schrevel J. Toward a novel metal-based chemotherapy against tropical diseases. 7. Synthesis and in vitro antimalarial activity of new gold-chloroquine complexes. J Med Chem. 2004; 47:5204-9.

[90] Sandbhor U, Padhye S, Billington D, Rathbone D, Franzblau S, Anson CE, Powell AK. Metal complexes of carboxamidrazone analogs as antitubercular agents. 1. Synthesis, X-ray crystal-structures, spectroscopic properties and antimycobacterial activity against *Mycobacterium tuberculosis* H37Rv.J Inorg Biochem. 2002; 90:127-36.

[91] Dagani R. The bio side of organometallics. C&EN 2002; 80: 23-29.

5 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* (Manunscrito publicado no "Journal of Molecular Biology" em Junho de 2006)

Running title: WT and isoniazid-resistant enoyl reductases

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*.

Jaim S. Oliveira[§], José H. Pereira[†], Fernanda Canduri[†], Nathália C. Rodrigues[†], Osmar N. de Souza[¶], Walter F. de Azevedo Jr[§], Luiz. A. Basso[§]*, and Diógenes. S. Santos[§]*

From the [§]Centro de Pesquisas em Biologia Molecular e Funcional, Faculdade de Farmácia e Faculdade de Biociências, Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre, RS, Brasil., [†] Departamento de Física, UNESP, São José do Rio Preto, Brazil, [¶] Laboratório de Bioinformática, Modelagem e Simulação de Biossistemas, Faculdade de Informática, PUCRS, Porto Alegre, RS, Brasil,

[*] to whom correspondence may be addressed
LAB: phone, +55-51-33203629; fax, +55-51-33203629; e-mail, luiz.basso@pucrs.br or DSS: phone, +55-51-33203629; fax, +55-51-33203629; e-mail, diogenes@pucrs.br

An understanding of isoniazid (INH) drug resistance mechanism in *Mycobacterium tuberculosis* should provide significant insight for the development of newer anti-tubercular agents able to control the INH-resistant tuberculosis (TB). The *inhA*-encoded 2-*trans* enoyl-acyl carrier protein reductase enzyme (InhA) has been shown through biochemical and genetic studies to be the primary target for INH. In agreement with these results, mutations in the *inhA* structural gene have been found in INH-resistant clinical isolates of *M. tuberculosis*, the causative agent of TB. In addition, the InhA mutants were shown to have higher dissociation constant values for NADH and lower values for the apparent first-order rate constant for INH inactivation as compared to wild-type InhA. Here, in trying to identify structural changes between wild-type and INH-resistant InhA enzymes, we have solved the crystal structures of wild-type and of S94A, I47T and I21V InhA proteins in complex with NADH to resolutions of, respectively, 2.3 Å, 2.2 Å, 2.0 Å, and 1.9 Å. The more prominent structural differences are located in, and appear to indirectly affect, the dinucleotide binding loop structure. Moreover, studies on pre-steady-state kinetics of NADH binding have been carried out. The results showed that the limiting rate constant values for NADH dissociation from the InhA-NADH binary complexes (k_{off}) were 11-, 5-, and 10-fold higher for, respectively, I21V, I47T, and S94A INH-resistant mutants of InhA as compared to INH-sensitive wild-type InhA. Accordingly, these results are proposed to be able to account for the reduction in affinity for NADH for the INH-resistant InhA enzymes.

Keywords: enoyl-ACP reductase, isoniazid resistance, *Mycobacterium tuberculosis*, pre-steady-state kinetics, NADH binding

Tuberculosis (TB) is a curable, yet widespread disease that infects one-third of the world's population. 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, which translates into more than 4,900 deaths per day, and more than 95 % of these were in developing countries.[1] In addition, it has been estimated that 3.2 % of the world's new cases of TB, in 2000, were multidrug-resistant tuberculosis (MDR-TB), defined as resistant to at least isoniazid and rifampicin.[2] According to the World Health Organization (WHO), there are 300,000 new cases per year of MDR-TB worldwide, and 79 % of MDR-TB cases are "super strains", resistant to at least three of the four main drugs used to treat TB.[3] Thus, despite the availability of effective chemotherapy and a moderately protective vaccine, newer agents are urgently needed to decrease the global incidence of TB. Accordingly, an understanding of drug resistance mechanisms in this pathogen should provide significant insight into the development of newer anti-tubercular agents able to combat the MDR-TB strains. Isoniazid (INH) was introduced as an antituberculosis drug in 1952[4,5] and, soon after, the first INH-resistant *M. tuberculosis* strains were isolated.[6] The *inhA*-encoded protein has been shown through a number of biochemical and genetic studies to be the primary target for isoniazid.[7-11] InhA was identified as an NADH-dependent 2-*trans* enoyl-ACP (acyl carrier protein) reductase enzyme that exhibits specificity for long-chain enoyl thioester substrates.[8] 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.[12] To inhibit InhA, INH needs to be activated by the catalase/peroxidase KatG and the isoniazid-activated intermediate forms an isonicotinyl-NAD adduct (INH-NAD), through addition of either an isonicotinic acyl anion to NAD^+ or an isonicotinic acyl radical to an $NAD^{\cdot}$ radical.[13] The INH-NAD adduct has been shown to bind to InhA active site[13] with an equilibrium dissociation constant value of 0.4 nM.[14]. More recently, this adduct has been shown to be a slow, tight binding competitive inhibitor, which inhibits the InhA activity with an overall inhibition constant (K_i*) value of 0.75 ± 0.08 nM.[15] 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 wild-type (WT) InhA, whereas there were only modest differences in the steady-state parameters.[16] In this study,

K_d values of 0.57, 13.9, 85.1, and 36.4 μM were determined for NADH binding to, respectively, WT InhA, I21V, I47T and S94A mutant InhA enzymes. The S94A mutant enzyme has been initially identified in an isoniazid resistant strain of *Mycobacterium smegmatis* generated in laboratory.[7] More recently, however, the Ser94→Ala mutation in the *inha* structural gene was identified in clinical isolates of *M. tuberculosis* strains with coresistance to INH and ethionamide.[17]

In order to gain insight into the structural basis that could account for the lower NADH affinity observed for INH-resistant InhA mutants identified in clinical isolates of *M. tuberculosis* as compared to WT enzyme, we have purified, crystallized, and solved for the first time the crystal structures of INH-resistant I47T and I21V InhA mutant proteins in complex with NADH to resolutions of, respectively, 2.2 Å, and 2.0 Å. Furthermore, we also report the crystal structures of INH-sensitive WT InhA and INH-resistant S94A mutant enzymes to a better resolution than previously described. Altogether, these structures provide a description of the enzyme-NADH interactions that are lost due to the mutations in the NADH binding site of InhA. In addition, we have carried out pre-steady-state kinetics to assess the NADH binding kinetic mechanism and to investigate which rate constant(s) is(are), if any, changed in INH-resistant InhA mutant enzymes as compared to WT InhA. The data presented here provide structural and mechanistic insights into the differences between INH-sensitive WT InhA and INH-resistant mutants that can account for the INH resistance in clinical isolates of *M. tuberculosis* strains harboring inhA structural gene mutations. These data should aid in the rational design of compounds able to efficiently inhibit both INH-resistant and -sensitive InhA enzymes with potential antitubercular activity.

RESULTS

Differences between crystal structures of WT, I21V, I47T and S94A InhA enzymes in complex with NADH.

The InhA structures display an α/β folding with a Rossmann-fold pattern, consisting of a central β-sheet composed of parallel strands and flanked by α-helices (Fig. 1), typical and common to other enzymes from short-chain dehydrogenase/reductase (SDR) superfamily. The SDR enzymes have been classified in five SDR families, which were designated classical, extended, intermediate, divergent, and complex.[18] The InhA enzyme is part of the divergent family, which includes NADH-dependent enoyl reductases from bacteria and plants.

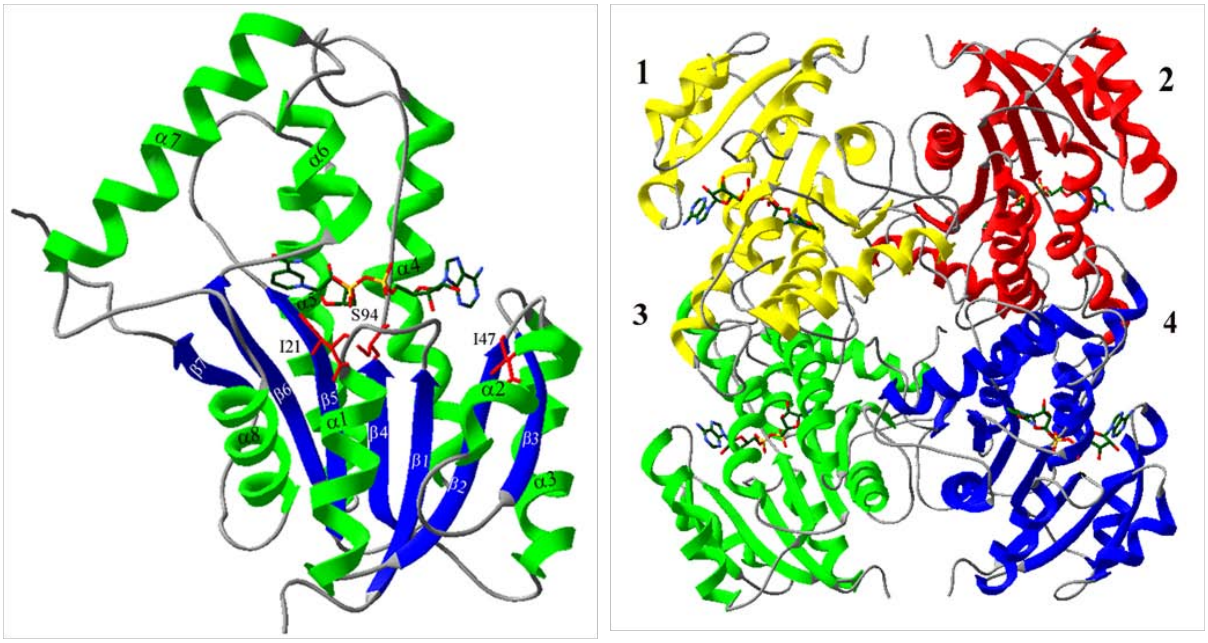


Fig. 1. WT InhA structure solved to 1.92 Å. **Left**: Ribbon representation of one InhA subunit. The strands are colored blue and the helices are colored green. The loops and turns are colored gray. The strands and helices are labeled α1 to α8 and β1 to β7, respectively. The NADH molecule bound to active site is colored by atom (green for carbon, blue for nytrogen, red for oxygen, and yellow for phosphorus). The residues Ile21, Ile47, and Ser94 that are mutated in the isoniazid-resistant InhA enzymes of the present study are shown in red. **Right**: InhA tetramer viewed along one of three molecular twofold axes (the C axis). Subunits are numbered 1 (blue), 2 (green), 3 (yellow), and 4 (red) with secondary structure elements represented as ribbons. The NADH bound to active site of the fours subunits is shown in all-atom representation and is

colored by atom type (green for carbon, blue for nytrogen, red for oxygen, and yellow for phosphorus). The figure has been produced using the program DeepView/Swiss-PdbViewer 3.7 (SP4).[61]

In trying to identify differences between the structures of WT and INH-resistant InhA enzymes reported here, the backbone C-α carbons of I21V, I47T, and S94A mutant InhAs were superimposed on WT InhA. No significant differences in secondary structure and backbone conformations have been observed between these structures. Furthermore, the pattern of distribution for B-factor values per protein residues is very similar in the four crystal structures of InhA (Supplementary Material). One of the more flexible regions is the stretch of residues 42-46 located in the loop connecting strand β2 to helix α2 and in the N-terminal of helix α2. The main-chain carbonyl group of Arg43 hydrogen bonds to the backbone NH of Ile47 and appears to help stabilize the helix α2 in the InhA structures. However, in the crystal structure of I47T InhA the first six residues (Leu44-Arg49) of helix α2 appear to be more flexible than in the other three InhA structures, probably owing to two hydrogen bonds between main-chain carbonyl group of Arg43 with water321 and OG1 atom of Thr47 that are absent in the WT structure (Fig. 2). Thus, we suggest that in the I47T structure the acceptor O atom of Arg43 has less occupancy in hydrogen bonding to the helix residue Thr47 than in the WT structure. The electron density for NADH is very clear in WT and in the three isoniazid-resistant InhA proteins. The mean B-factor values for bound NADH molecule are very similar in all structures of the binary complexes (Table 1).

A summary of all direct and water-mediated hydrogen bond interactions between NADH and protein residues in the crystal structures of binary complexes of WT-NADH, I21V-NADH, I47T-NADH, and S94A-NADH shows that the interactions are conserved between the four InhA enzymes and 54% of hydrogen bonds are mediated by 10 water molecules in WT InhA (Table 5 of the Supplementary Material). It should be pointed out that a water-mediated interaction was regarded as a hydrogen bond between an NADH atom and the protein that is mediated by only one water molecule. In addition, two interactions between O3 and AO1 atoms of pyrophosphate moiety of NADH and backbone oxygen atom of Gly96 are mediated by two water molecules in the binary complex WT InhA-NADH. These interactions are also conserved in the three isoniazid-resistant InhA proteins. Bottoms *et al*. have pointed out the importance of bridging

water molecules in the dinucleotide-binding interface in proteins with classic Rossmann fold structures.[19]

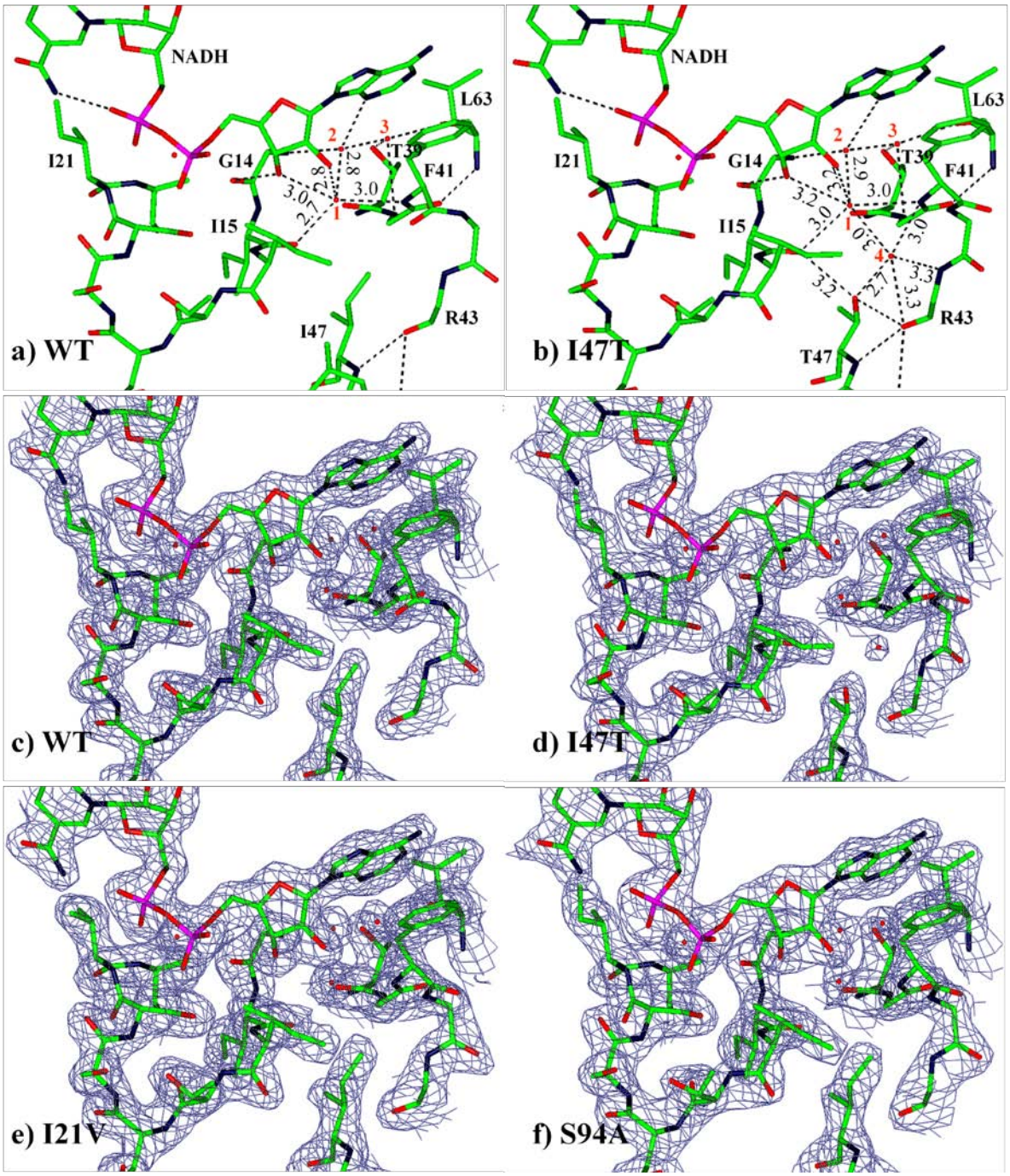


Fig. 2. Surroundings of position 47 in the crystal structures of WT (a) and (c), I47T (b) and (d), I21V(e), and S94A(f). For clarity only NADH, dinucleotide binding loop (Gly14-Ala22)

residues, strand β2 residues Thr39 and Gly40, residues from the turn between strand β2 and helix α2 (Phe41-Arg43), Leu63, Ile47 (a) or Thr47 (b), and water molecules near the position 47 that are involved in hydrogen bond network. The atoms are colored green for carbon, blue for nytrogen, red for oxygen, and magenta for phosphorus. The water molecules are numbered 1, 2, 3 and 4 and corresponds to waters 272, 282, and 281 in both WT InhA and I47T InhA structures and to water 321 in I47T InhA structure, respectively. The most relevant hydrogen bonding differences between the structures of WT-NADH and I47T-NADH binary complexes are shown in the parts (a) and (b), respectively. The ($2F_{obs} - F_{calc}$) electron density maps encompassing the residue 47 and the surrounding atoms of the structures of WT-NADH, I47T-NADH, I21V-NADH, and S94A-NADH binary complexes are shown in the parts (c), (d), (e), and (f), respectively. The ($2F_{obs} - F_{calc}$), α_{calc} maps were calculated in the REFMAC 5.2 program[48] using all unique reflections and are countoured at a level of 1.2σ above the mean. The figure was prepared with CCP4 Molecular Graphics 0.12 program.[62]

They found that on average, FAD-, NADP-, and NAD-binding proteins accommodate about, respectively, 12, 11, and 9 interfacial water molecules per dinucleotide binding site (within 3.75Å of the protein and dinucleotide). Actually, the major differences on NADH interaction between the WT and INH-resistant InhA proteins, which could account for the lower NADH affinity observed for the mutant proteins, have been identified in the surroundings of each mutated residue and will be described below. Notwithstanding, a water-mediated hydrogen bond between NADH and the Thr196 OG1 atom of WT InhA protein, which don`t make any contact with residues from the mutation points in the NADH binding site, is likely weaker in the mutant proteins reported here. The analysis of simulating annealing omit electron density maps has shown that in the isoniazid-resistant InhA mutants the Thr196 side-chain has rotated ~180° in relation to WT InhA structure, resulting in hydrogen bond distances between the OG1 atom of Thr196 and the bridging water (Water380) that are significantly larger (3.3 – 3.4 Å) than the 2.5 Å distance observed in WT enzyme (supplementary material). Although this finding could be a consequence of the lower NADH affinity of INH-resistant InhA proteins since some polar interactions with NADH are expected to be less stable in I21V, I47T, and S94A InhA mutant proteins than in WT enzyme, there are no structural constraints or interactions that could restrict the rotation of the Thr196 side-chain in both wild-type or isoniazid-resistant InhA structures

presented here. Accordingly, the Thr196 conformation observed in I21V, I47T, and S94A InhA mutant proteins has also been observed in the previously solved structure of WT InhA-NADH binary complex **(1ENY)**[27]. This finding sugest that this interaction cannot account for the lower NADH affinity observed for the isoniazid-resistant InhA proteins.

Table 1. Summary for data collection, structural and refinement statistics.

	WT InhA	S94A InhA	I21V InhA	I47T InhA
Unit cell dimensions a, c (Å)	96.02, 138.16	96.25, 138.05	95.95, 138.24	95.440, 138.71
Space group	$P6_222$	$P6_222$	$P6_222$	$P6_222$
Number of unique reflections	27,062	15,559	23,889	17,898
Refinement resolution range (Å)	35.62 ~ 1.92 (1.971~1.921)[a]	40.29 ~ 2.30 (2.360~2.300)	41.56 ~ 2.01 (2.062~2.010)	40.36 ~ 2.20 (2.257~2.200)
Completeness (%)	97.2 (90.1)	94.1 (90.5)	97.8 (89.4)	96.2 (98.5)
$I/\sigma(I)$	11.1 (1.2)	10.7 (3.8)	10.1 (3.9)	6.2 (3.2)
R_{sym}	0.046 (0.391)	0.049 (0.177)	0.059 (0.193)	0.076 (0.210)
R_{full}	0.035 (0.363)	0.042 (0.166)	0.050 (0.138)	0.062 (0.157)
Multiplicity	2.8 (2.7)	2.4 (2.5)	9.1 (9.3)	4.9 (4.6)
R-factor[b]	0.187 (0.306)	0.189 (0.226)	0.194 (0.262)	0.186 (0.216)
R-free	0.213 (0.326)	0.242 (0.316)	0.227 (0.314)	0.213 (0.219)
Estimated Overall coordinate error based on Maximum Likelihood (Å)	0.08	0.15	0.10	0.11
RMSD from ideality				
Bond lengths (Å)	0.015	0.022	0.017	0.020
Bond angles (°)	1.750	2.136	1.957	1.956
Total protein residues	267 (Gly3-Leu269)	268 (Ala2-Leu269)	268 (Ala2-Leu269)	268 (Ala2-Leu269)
Heteroatoms	1 (NADH)	1 (NADH)	1 (NADH)	1 (NADH)
Water molecules	176	134	152	147
Average B-factors (Å^2)				
Protein residues	23.9	27.4	22.9	26.7
Protein main chain	23.1	26.7	22.0	26.1
Protein side chain	24.9	28.1	24.1	27.5
NADH	20.3	22.9	19.4	21.7
Water	31.3	30.5	29.5	32.1
Ramachandran Plot in PROCHECK[50]				
Most favorable region	91.6%	90.7%	91.6%	90.2%
Additional allowed region	7.1%	8.4%	7.6%	8.9%
Generously allowed region	0	0	0	0

Disallowed region	1.3%	0.9%	0.9%	0.9%
Ramachandran Plot in MOLPROBITY[62]				
Favorable regions	96.6%	97.0%	96.2%	95.5%
Allowed regions	100%	100%	100%	100%
Disallowed regions	0%	0%	0%	0%

[a]In parenthesis are the data obtained for the highest resolution shell.
[b]*R*-factor = $\sum|F_o - F_c|/\sum F_o$, where F_o and F_c are the observed and calculated structure-factor amplitudes for the reflection with Miller indices *h = (h,k,l)*. The free *R*-factor is calculated for a test set of reflections (5 %) which are not included in atomic refinement.

I21V-NADH structure — Alpha-carbon superposition between the crystal structures of WT InhA (1.92 Å) and I21V InhA enzymes showed an RMSD of 0.09 Å. Moreover, the direct hydrogen bonds and most water-mediated hydrogen bonds between NADH and protein residues are conserved in both crystal structures. A study reporting a systematic analysis of 21 crystal structures of NAD(H)-bound SDR enzymes has identified conserved interactions between a protein fragment of six to seven residues (positions S1 to S7) nearly parallel to diphosphate moiety of NAD(H) and the NAD(H) molecule.[20] The amino acid in position S6 was recognized to be always hydrophobic probably because its side chain is directed toward the nicotinamide ring of NAD with its conserved apolar environment. In addition, the S6 residue always makes a hydrogen bond with a phosphate oxygen atom through its main-chain N atom.[20] In fact, this protein fragment corresponds to the classical glycine-rich diphosphate binding loop (Gly14-Ala22 in InhA) connecting the first β-strand with the first α-helix in dinucleotide binding proteins that have the Rossmann fold motif. Accordingly, the apolar side-chain of the amino acid in position S6 corresponds to Ile21 residue in InhA. In the previously solved WT InhA structures (**1ENY**, **1BVR**) and in our WT and I21V InhA structures the main-chain NH atom of residue 21 makes a hydrogen bond with the phosphate oxygen NO2, and two interactions mediated by the conserved water271 with NO2 and AO5* oxygen atoms of NADH (Fig. 3). In WT InhA, the CD1 atom of Ile21 apolar side-chain makes van der Waals contacts with the nicotinamide ring, and nicotinamide ribose, and phosphate oxygen atoms. The distances between the atoms of nicotinamide, nicotinamide ribose, and phosphate oxygens and CD1 atom of Ile121 are given in Table 2. In the INH-resistant I21V InhA enzyme these van der Walls contacts are lost owing to the mutation of Ile→Val, since the Val side -chain does not have the CD1 atom and its carbon atoms superpose very well with the equivalent carbon atoms of Ile21 residue in the WT InhA

crystal structure (Fig. 3). Moreover, the van der Waals contacts between Ile21 side-chain and nicotinamide moiety of NADH are structurally conserved in the crystal structures of enoyl-ACP reductases in complex with NAD^+ from other sources. In *Escherichia coli*[21] and *Helicobacter pylori*[22] structures (PDB codes: **1QSG** and **1JW7**, respectively) the van der Waals contacts are made by Ile20 residue, whereas in *Brassica napus*[23] (PDB code: **1D7O**) and *Plasmodium falciparum*[24] (PDB codes: **1V35** and **1UH5**) structures these contacts are made by, respectively, Tyr32 and Tyr111. Thus, since the interaction between Ile21 side-chain and nicotinamide appears to be conserved in enoyl-ACP reductases and in other SDR enzymes, we propose that this residue makes a contribution in stabilizing the bound NAD(H) in the InhA active site. Accordingly, the loss of contacts between the side-chain of amino acid 21 with the nicotinamide could account for the higher equilibrium dissociation constant value for Ile21V mutant as compared to the WT InhA enzyme.[16] A lower stabilization of NADH bound to I21V InhA compared to WT InhA is consistent with the recently reported results of molecular dynamic simulation with I21V and WT InhA enzymes in complex with NADH.[25] It was observed a higher flexibility of pyrophosphate moiety of NADH and a lower occupancy of conserved direct bonds between the coenzyme and protein residues in I21V than in WT InhA, during the last 2 ns of the molecular dynamic simulation. In addition, the in silico-mutated enzyme showed a decreased total number of diret and water-mediated hydrogen bonds between cofactor and protein residues as compared to WT InhA.[25] It thus appears that the loss of van der Waals contacts between the nicotinamide moiety and the apolar side chain of Ile21 observed in the crystal structure of I21V InhA would lead to a destabilization of NADH bound to I21V InhA active site, thereby accounting for the reduced number and occupancy of hydrogen bonds and the higher flexibility of NADH observed in the molecular dynamic simulation with I21V InhA.[25]

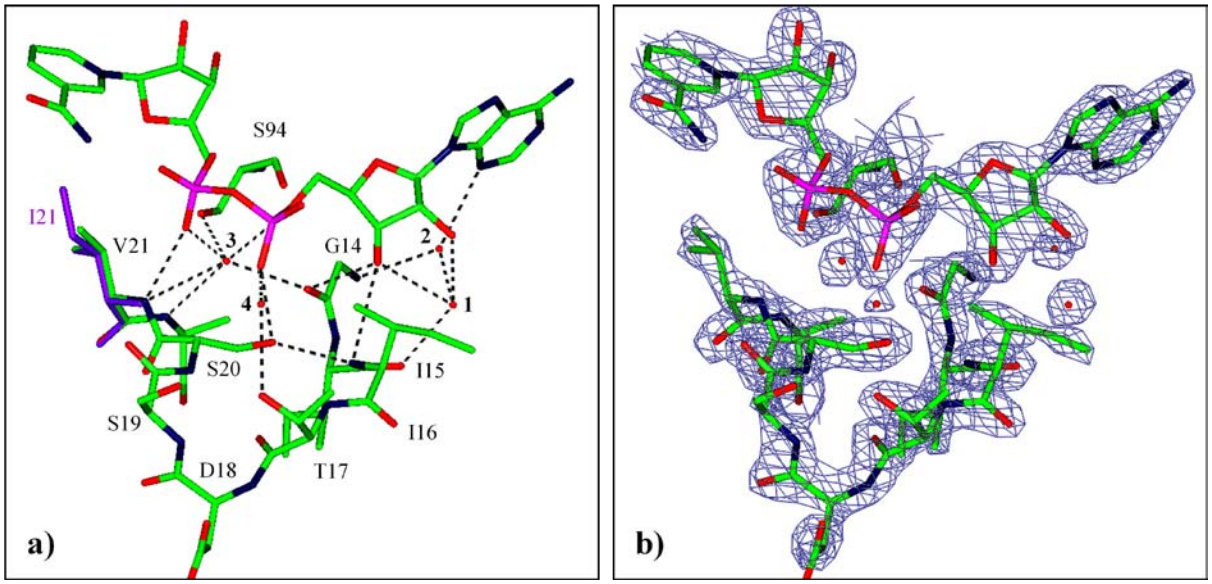


Fig. 3. a) Interactions of dinucleotide binding loop in the crystal structure of the binary complex I21V InhA-NADH. For clarity only the NADH molecule, dinucleotide binding loop residues (Gly14 to Ala22), and the conserved bridging waters are shown. The direct and water-mediated hydrogen bond interactions between NADH and the loop are represented by black dashed lines. The waters that are numbered 1, 2, 3, and 4 correspond to waters 272, 282, 271, 285 in both I21V InhA-NADH and WT InhA-NADH crystal structures. The atoms are colored green for carbon, blue for nitrogen, red for oxygen, and magenta for phosphorus. The atoms of Ile21 side-chain of WT InhA-NADH structure superimposed on the structure I21V InHA-NADH are shown in purple. The CD1 atom of Ile21 side-chain (WT InhA) makes several contacts with the nicotinamide moiety that are lost due to the mutation to valine.

b) $(2F_{obs} - F_{calc})$, α_{calc} electron density map encompassing the NADH molecule, the conserved 272, 282, 271, 285 water molecules and the dincleotide binding loop of the crystal structure of I21V InhA-NADH binary complex. The $(2F_{obs} - F_{calc})$, α_{calc} electron density map was calculated in the REFMAC 5.2 program[48] using all unique reflections and is countoured at a level of 2.3σ above the mean. The figure was prepared with CCP4 Molecular Graphics 0.12 program.[62]

Table 2. Van der Walls contacts between the CD1 atom of Ile21 side-chain and NADH in the binary complex WT-NADH (1.92 Å).

NADH moiety	NADH atom	Distance (Å)
Pyrophosphate	NO1	3.8
	NO2	4.0
Nicotimamide ribose	NO4*	3.8
Nicotinamide	NN1	4.0
	NC2	3.7
	NC3	3.9
	NC4	4.4
	NC7	4.3
	NN7	4.1

I47T-NADH structure — In the I47T mutant enzyme, the apolar residue isoleucine is replaced by the slightly polar threonine residue. A structural alignment with all atoms between I47T InhA and WT InhA structures reported here resulted in RMSD deviations of 0.20 Å and 0.64 Å for, respectively, backbone atoms and all atoms. Hence, there are no significant differences in I47T backbone conformation in relation to WT enzyme. In fact, the structural differences are limited to the surroundings of the mutation site (Fig. 2). In the present WT InhA crystal structure, a water molecule (Water272) mediates hydrogen-bonding interactions between the carbonyl group of Ile15 with 2'- and 3'-OH groups of adenine ribose moiety of NADH, and makes hydrogen bonds with the main-chain N atom of Phe41 residue, and with Water282 molecule. The Water282, in turn, mediates an interaction between the backbone nitrogen atom of Gly14 with N3 atom of adenine and hydrogen bonds to Water281, which is hydrogen bonded to OG atom of Thr39 (C-end of strand β2), main-chain N atom of Gly40, and main-chain oxygen atom of Leu63 (C-terminal of strand β3). This network of water-mediated interactions causes the formation of a chair-type conformation on which the adenine sits. A superimposition of WT InhA structure with crystal structures of complexes between NAD^+ and enoyl reductases from other sources showed that the a water-mediated hydrogen bond between the second residue of the dinucleotide binding loop (Ile15 in InhA) with the first residue from the turn connecting strand β2 with helix α2 (Phe41 in InhA) is conserved in enoyl reductase enzymes from *Helycobacter pylori*[22] (PDB ID: **1JW7**, subunit 2), *Plasmodium falciparum*[24] (**IUH5**, subunit 1), *Brassica napus*[23] (**1D7O**), and *Escherichia coli*[26] (**1QG6**, subunit 1). Furthermore, as observed in InhA, the conserved bridging

water (Water272 in InhA) makes two additional hydrogen bonds with 2'-OH and 3'-OH groups of adenine ribose moiety in all of these crystal structures. Thus, we propose that these water interactions are important to the correct geometry of the dinucleotide-binding loop for the optimal nucleotide binding in InhA and in other enoyl reductases.

In I47T InhA structure some new interactions are observed close to the waters that mediate the interactions discussed above. The slightly polar side-chain of Thr47 allowed the entrance of an additional water molecule (Water321) to the NADH binding site. The Water321 makes hydrogen-bond contacts with Thr47 (OG atom), Phe41 (main-chain N atom), Arg43 (main-chain O and N atoms), and the conserved Water272. In addition, the OG atom of Thr47 makes a hydrogen bond with the carbonyl group of Ile15 residue from the dinucleotide binding loop, which, in turn, interacts with 2'-OH and 3'-OH groups of adenine ribose moiety through Water272. Hence, the new polar interactions promoted by the presence of Water321 and the direct hydrogen bond of Thr47 side-chain with dinucleotide binding loop residue Ile15 led to a new hydrogen bond pattern in the NADH binding site of InhA that probably affects the binding of NADH. This is borne out by the significantly larger hydrogen bond distances observed in I47T as compared to WT InhA structure for the interactions between the carbonyl group of Ile15 with Water 272, and Water272 with 2'-OH atom of adenine ribose (Fig. 2a and 2b), considering the estimated overall coordinate error values of 0.08 and 0.11 Å that were calculated by REFMAC 5.2 program for, respectively, the structures of WT-NADH and I47T-NADH binary complexes (Table 1). The Water272 also appears to be less ordered in the I47T-NADH than in the WT-NADH structure as indicated by a higher B-Factor value for the mutant enzyme (33.2 $Å^2$) as compared to the WT enzyme (22.9 $Å^2$). Thus, the Ile15 carbonyl group- and Water272-mediated hydrogen bonds are expected to be weaker and of lower occupancy than the equivalent interactions in WT enzyme, since in the structure of I47T-NADH binary complex, Ile15 and Water272 are partially occupied in hydrogen bonding to Water321 (Fig. 2b). These features could account for the higher equilibrium dissociation constant value for NADH of this mutant enzyme. The Water272 that mediates interactions between protein residues with hydroxyl groups of adenine ribose appears to be stabilized by the hydrophobic Ile47 side-chain in the WT-NADH binary complex. However, in the structure of I47T-NADH binary complex the presence of a second water molecule (Water321) appears to be necessary to maintain the Water272 interacting with the 2'- and 3'-OH groups of adenine ribose, since in the absence of Watert321, the Water272

could instead interact with OG atom of slightly polar Thr47 side-chain. It is thus tempting to suggest that the presence of Water321 in the active site of I47T-NADH would maintain a weak interaction between Water272 and the 2`- and 3`-OH groups of adenine ribose, thereby avoiding a complete lack of interaction between these groups that could be deleterious to the enzyme activity and ensuing survival of the bacilli. Electron density maps calculated for the WT and isoniazid-resistant InhA structures (shown in Fig. 2b, 2c, 2d, and 2e) give evidence that the Water 321 is absent in the structures of WT-NADH, I21V-NADH, and S94A-NADH binary complexes. Actually, in the I47T InhA-NADH structure superimposed on WT InhA-NADH structure the Water 321 oxygen atom is at a distance of only 1.4 Å from the Ile47 CD1 atom. Thus, in these structures, the steric hindrance promoved by the CD1 atom of Ile47 residue side-chain and the hydrophobic character of its side-chain probably impair the presence of a additional water molecule in this region.

S94A-NADH structure — A previous study compared the crystal structure of binary complex of WT and INH-resistant S94A mutant InhA in complex with NADH and proposed that disruption of a hydrogen bond network in the mutant protein due to rotation of Gly14 residue could account for the higher coenzyme dissociation constant value.[27] In WT-NADH crystal structure (2.2 Å) a well-ordered water molecule was shown to mediate two hydrogen bonds between the O2 of the P_N phosphate of NADH and protein (shown in Fig. 4a) that are lost in the S94A mutant protein.[27] This water molecule (WAT1) hydrogen bonded to the hydroxyl group of Ser94, the main-chain oxygen of Gly14, and the main-chain nitrogen atoms of Ala22 and Ile21. In the S94A-NADH complex structure (2.7 Å), the carbonyl group of residue Gly14 was shown to rotate away from the water molecule, breaking the hydrogen bond observed in the WT protein structure (shown in Fig. 4b). However, in the S94A[27] the main-chain oxygen atom is at a distance of only 2.16 Å from the 3'-OH oxygen of adenine ribose that is shorter than the lower limit distance (2.29 Å) expected to a hydrogen bond.[28] In fact, this interaction appears to be unfavorable, since the van der Waals radii for a hydroxyl oxygen atom and for a carbonyl oxygen is 1.72 Å and 1.66 Å, respectively.[29] Thus, it should be expected a high van der Waals repulsion energy between two oxygens separated by 2.16 Å. Accordingly, the rotation of Gly14 residue in the S94A mutant appears to play no role in reducing NADH affinity and ensuing INH resistance mechanism as previously suggested.

A computational comparison of 102 high-resolution structures (≤ 1.90 Å) of enzyme - dinucleotide complexes revealed that the WAT1 water molecule is structurally conserved in proteins that exhibit the classic Rosmann dinucleotide-binding fold motif.[19] This conserved water molecule links, through a conserved hydrogen-bonding pattern, the glycine-rich phosphate-binding loop with the dinucleotide pyrophosphate moiety.[30]. In enoyl reductase enzymes the phosphate-binding loop consensus sequence has been proposed to be GXXXXXXXG(A) (residues Gly14-Ala22 in InhA).[19] Typically, the WAT1 molecule makes four hydrogen bonds and two of them are invariant. The invariable hydrogen bonds involve the dinucleotide pyrophosphate moiety and main-chain amine of the last conserved Gly of phosphate binding loop (Ala22 in InhA). Moreover, the interaction of pyrophosphate has been shown to be stereospecific. Almost without exception, the pyrophosphate atom that interacts with the structurally conserved water is NO2 (O2N) in the case of NAD- or NADP-binding proteins. In structures with the sequence pattern GXXXXXXXG(A), the partners for the other two hydrogen bonds are the carbonyl group of the first conserved Gly (Gly14 in InhA) and a C-terminal residue of β4 strand (Ser94 in InhA).[19] In our efforts to understand the structural basis of isoniazid resistance mechanism of InhA, we have solved the crystal structures WT-NADH and S94A-NADH binary complexes to a better resolution (refined to, respectively, 1.92 Å and 2.3 Å) than previously reported[27] (refined to, respectively, 2.2 Å and 2.7 Å). Analysis of the NADH binding pocket showed that there was no rotation of the carbonyl group of Gly14 in the S94A-NADH structure reported here (Fig 4c, 4d, 4e, and 4f). In addition, only a hydrogen bond between the conserved WAT1 molecule (Water 271 in the present structures) and the hydroxyl group of Ser94 is lost owing to Ser94→Ala mutation (Fig 4c and 4d). It thus appears that this hydrogen bond disruption is likely the main reason for a sixty-fold reduction in affinity for NADH observed for the S94A InhA protein, which corresponds to approximately 2.5 kcal mol^{-1}.[16] This is in good agreement with estimates for the energies of hydrogen bonds which are in the range of 3 to 9 kcal mol^{-1}.[31] In addition, this WAT1 molecule has been proposed to be an integral characteristic of dinucleotide-binding Rossmann fold domains that both contributes significantly to dinucleotide recognition and presumably provides a favorable enthalpic contribution to the free energy of binding.[19]

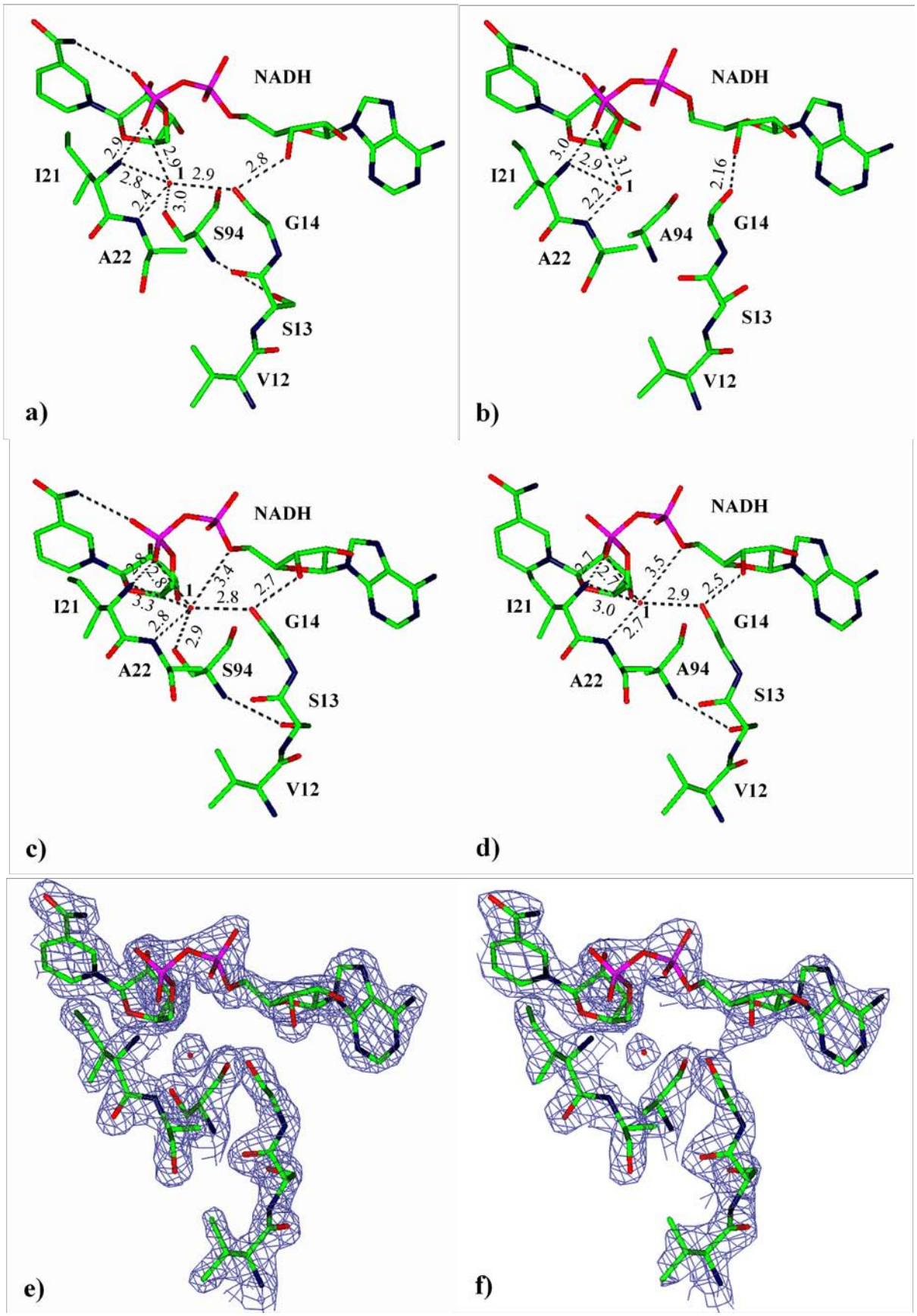

NADH
I21
A22
S94
G14
S13
V12
a)
NADH
I21
A94
A22
G14
S13
V12
b)
NADH
I21
A22
S94
G14
S13
V12
c)
NADH
I21
A22
A94
G14
S13
V12
d)
e)
f)

Fig. 4. Differences in hydrogen bond network between the crystal structures of WT and INH-resistant mutant InhA (S94A) complexed with NADH. **Top**: Crystal structures of WT InhA-NADH (a) and S94A InhA-NADH (b) binary binary complexes refined to, respectively, 2.2 Å and 2.7 Å.[27] **Middle**: Crystal structures of WT InhA-NADH (C) and S94A InhA-NADH (D) binary complexes refined to, respectively, 1.9 Å and 2.3 Å (present work). Only the conserved hydrogen bond between the conserved Water271water (**1)** and hydroxyl group of residue S94 is lost owing to mutation to Ala94. **Bottom:** $(2F_{obs} - F_{calc})$, α_{calc} electron density map encompassing the active site residues from crystal structures of WT InhA-NADH (e) and S94A InhA-NADH (f) binary complexes. The map clearly shows the position of Gly14carbonyl group in both structures. For clarity, only residues Val12, Ser13, Gly14, Ile21, Ala22, and **1** (WAT1 water in 1ENY and 1ENZ and Water271 in the present InhA structures) and NADH molecules are shown. Direct and water-mediated hydrogen bonds between the protein and NADH pyrophosphate moiety are represented as dashed black lines. The interactions distances are in Å. The atoms are colored green for carbon, red for oxygen, blue for nitrogen, and magenta for phosphorus. . The $(2F_{obs} - F_{calc})$, α_{calc} electron density map were calculated in the REFMAC 5.2 program[48] using all reflections and are countoured at a level of 2.1σ above the mean.The figures were prepared using the program CCP4 Molecular Graphics 0.12.[62]

Analysis of oligomeric state and intersubunit contacts.

Molecular weight values of 115,930 Da, 104,890 Da, 118,080 Da, and 108,720 Da were estimated by gel-permeation chromatography for, respectively, WT, I21V, I47T, and S94A homogeneous InhA proteins (Supplementary material). The subunit molecular weight (MW) of WT InhA is 28,368 Da, so a MW value of 113,472 Da would be expected for an InhA homotetramer. Thus, these results suggest that WT InhA and the three INH-resistant mutants are homotetramers in solution, and that the mutations in NADH binding site do not alter the oligomeric state of InhA. The gel filtration results are supported by the packing in $P6_222$ crystals of WT and INH-resistant mutant InhA enzymes in which a tetramer occurs with approximate dimensions 80 Å x 70 Å x 60 Å (Fig. 1). The tetramer has 222 symmetry and consequently there are three distinguishable intersubunit interfaces about three orthogonal molecular twofold axes, which are designated here as A, B, and C. In this case, each monomer in the tetramer makes contacts with all three symmetry-related partners (Fig. 1). Contacts between subunits 1 and 2 (3

and 4) are about the A axis, 1 and 3 (2 and 4) about the B axis, and 1 and 4 (2 and 3) about the C axis (Fig. 1).

It has been reported that mutations in the *inhA*-structural gene identified in INH-resistant strains of *M. tuberculosis* not only decreased the affinity of NADH for the enzyme but also resulted in appearance of cooperativity in cofactor binding.[16] Moreover, it has been suggested that homotypic interactions within the InhA tetramer could modulate, or be modulated by, ligand binding and that this allosteric response to ligand binding would be affected by InhA mutations that occur in INH-resistant mycobacterial strains.[15] Indeed, our analysis of the crystal structures of three INH-resistant InhA complexed with NADH showed that although almost all the direct intersubunit contacts are conserved in all the four InhA structures, some water mediated interactions are lost in the three orthogonal twofold axes (A, B, and C) in INH-resistant InhA enzymes. Although the lack of waters could be due to the poorer resolution of the structures of InhA mutant proteins (2.0 - 2.3 Å) compared to wild-type enzyme (1.9), these differences can be correlated with the slight cooperativity observed for isoniazid-resistant mutants.[16]. In addition, it can be reconciled with our stopped-flow studies on the slow phase of NADH binding (described below), since the association between protein monomers in the homotetramers are expected to be less stable in the mutant enzymes than in the WT enzyme. It should be pointed out that as the InhA homotetramer has three twofold axes of symmetry, the association between the protein monomers is isologous, i.e. any group that contributes to the binding in one subunit furnishes precisely the same contribution in the other subunit.[32] Hence, any bonded group-pair is represented four times within the InhA homotetramer, and in the same way any interaction that is lost in InhA mutant enzymes is lost four times in the context of the tetramer. A summary of all direct and water-mediated hydrogen bonding ocurring in the intersubunit interfaces about the three orthogonal molecular twofold axes (A, B, and C) of WT InhA homotetramer is given in the Supplementary Material (Table 6).

Differences between intersubunit contacts in the A axis — In WT InhA, the OG1 atom of $Thr241^{1}$ (subunit 1) residue interacts with $Leu250^{2}$, $Pro251^{2}$, $Ala262^{2}$, $Thr253^{2}$ residues from subunit 2 through a well-ordered water molecule (WAT385, B-Factor = 21.3 $Å^{2}$), and the main-chain oxygen of $Arg177^{1}$ makes a water-mediated interaction with main-chain oxygen of $Pro227^{2}$. These interactions are lost in I47T and S94A mutant enzymes, since those water

molecules in equivalent positions are not present in these structures. In I21V InhA, only the Water385-mediated interaction between OG1 atoms from Thr241[1] and Thr253[2] residues is lost due to a rotation of side-chain dihedral angle χ_1 of Thr253[1].

Differences between intersubunit contacts in the B axis — In the I21V InhA homotetramer, the water356 molecule, which in WT InhA mediates the interaction between Glu68[1] OE2 and Tyr125[1] OH atoms with Tyr113[3] N atom, is not present and thus these interactions are lost. In the I47T InhA, only the water-mediated-interaction between the side-chains of Ser166[1] and Arg173[3] residues is lost since the Water283 molecule is not present. Finally, in the S94A InhA structure, there are no water molecules in equivalent positions to Water283 (B-Factor = 21.27) and water304 (B-Factor = 26.87) molecules, which in WT InhA mediate, respectively, the interactions of nitrogen backbone atoms of Phe108[1] and Phe109[1] residues with OE1 atom of Glu178[3] residue, and of Ser166[1] side-chain with Arg173[3] side-chain. In addition, the NZ atom of Lys132[3] has moved away from the backbone oxygen of Phe109[1], breaking the hydrogen bond between these residues.

Differences between intersubunit contacts in the C axis — In the C axis, most of the contacts involve Arg153 residue that makes one direct hydrogen bond with His265 residue and water-mediated hydrogen bonds with Asp150, Ser152, Arg153, Arg225, Ala264, and Gln267 residues from the symmetry equivalent subunit. In I21V structure, the Arg153 interacts with the same residues as observed in WT InhA structure, but the hydrogen bonding pattern is somewhat different owing to small conformational changes of Arg153 side-chain occurring in this mutant enzyme. More noticeable differences in intersubunit interactions involving Arg153 are observed for the S94A enzyme. In this enzyme, a conformational change of Arg153 side-chain led to a loss of all water-mediated hydrogen bonds that have been observed in WT InhA.

Kinetics of NADH binding

The kinetics of NADH binding was determined by stopped-flow spectroscopy, monitoring the nucleotide fluorescence. The binding process was characterized by an initial fast phase followed by a second slow phase in WT and INH-resistant InhA mutant enzymes. Moreover, the two phases remain well separated from each other over the whole range of NADH

concentrations tested. Accordingly, all data sets were best fitted to a double exponential function (Fig. 5).

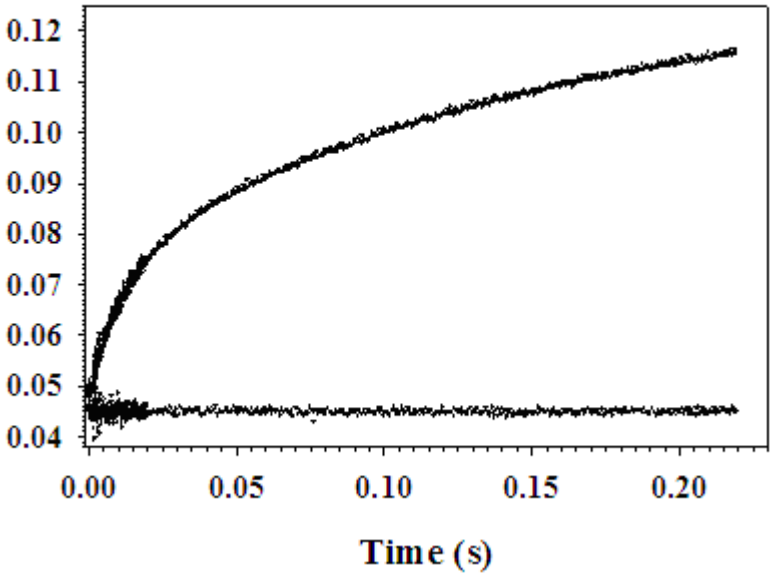


Fig. 5. Representative stopped-flow trace of NADH binding to WT InhA enzyme. This typical trace shows the biphasic enhancement in nucleotide fluorescence (λ_{ex} = 370 nm, band pass =4.65 nm), λem ≥ 420 nm) upon the binding of NADH (5 µM) to WT InhA (2 µM) enzyme (concentrations after mixing). The curve through the experimental trace shows the best fit to a double exponential function, which yielded values of 80.5 ± 2.3 and 6.9 ± 0.3 s^{-1} for the apparent rate constants of the fast phase (k_{obs1}) and slow phase (k_{obs2}), respectively. The trace above is an average of twelve individual traces. The control trace (at the bottom) was obtained by mixing 10 µM NADH solution with an equal volume of Pipes 100 mM, pH 7.0 buffer. This trace shows that there is no loss of nucleotide fluorescence enhancement signal owing to any process ocurring in the dead time (1.36 ms) of the stopped-flow instrument.

The values obtained for the apparent rate constant of the first phase (k_{obs1}) were plotted against the final NADH concentrations and appeared to increase linearly for WT, I21V and I47T mutant enzymes, whereas for S94A enzyme the rate constant values appeared to have a hyperbolic dependence on NADH concentration (Fig. 6). Hence, the fast phase rate constants obtained for WT, I21V and I47T InhA enzymes were tentatively fitted to the following equation for a single-step reversible bimolecular association,

$k_{obs1} = k_2[L] + k_{-2}$ **[1]**

where k_2 is the bimolecular association rate constant, k_{-2} is the rate constant for NADH dissociation from InhA-NADH binary complex, and [L] is the final NADH concentration (see scheme in Fig 7). The calculated rate constants values are given in Table 3. From these analysis equilibrium dissociation constant values (K_{d2}) of 0.3 ± 0.3, 5.9 ± 0.3, and 6.5 ± 0.3 μM could be estimated for WT-NADH, I21V-NADH, and I47T-NADH binary complexes by the following relationship: $K_{d2} = k_{-2}/k_2$.

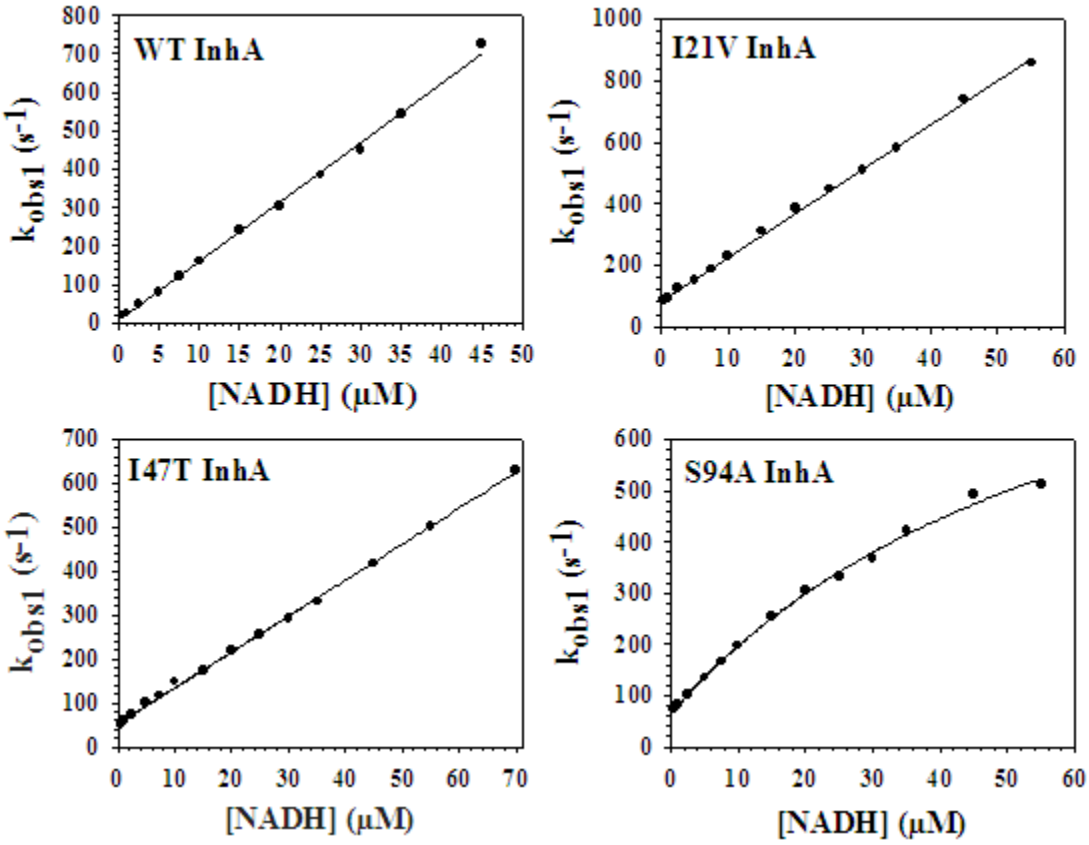


Fig. 6. Analysis of the NADH concentration dependence of the k_{obs1} values calculated from the fast phase (first) of NADH binding to WT, I21V, I47T, and S94A InhA enzymes. For WT, I21V, and I47T InhA enzymes the data were fitted to the equation **1**, which describes a single bimolecular association, whereas for S94A InhA the curve was fitted to equation **2**, which describes a rapid bimolecular association process followed for a unimolecular isomerization process.

Table 3. Kinetic constants estimated from the fast phase of NADH-binding to wild-type and isoniazid-resistant mutant InhA enzymes.

Kinetic constant	WT InhA	I21V InhA	I47T InhA	S94A InhA
k_2 (M^{-1} s^{-1})	15.5 ± 0.2 (x 10^6)	14.3 ± 0.1 (x10^6)	8.17 ± 0.08 (x10^6)	≥ 15 (x10^6)
k_{-2} (s^{-1})	5 ± 5	85 ± 4	53 ± 2	-
K_{d2} (μM)	0.3 ± 0.1[a] 0.39 ± 0.03[b] 0.3 ± 0.3[c]	7.6 ± 0.4[a] 4.76 ± 0.08[b] 5.9 ± 0.3[c]	6.9 ± 0.8[a] 3.7 ± 0.3[b] 6.5 ± 0.3[c]	67 ± 9[c]
mean K_{d2} (μM)[f]	0.33 ± 0.1	6.1 ± 0.3	5.7 ± 0.7	-
k_3 (s^{-1})	-	-	-	1014 ± 84
k_{-3} (s^{-1})	-	-	-	65 ± 6
K_{d2}* (μM)	-	-	-	2.7 ± 0.2[a] 3.9 ± 0.5[b] 4.0 ± 0.5[d]
mean K_{d2}* (μM)[f]	-	-	-	3.5 ± 0.4
k_{off}[e] (s^{-1})	6.1 ± 0.4	68 ± 1	30 ± 2	62 ± 1

K_{d2}* = K_{d2} overall, calculate with equation **3.**
[a] Calculated from analysis of the fast-phase amplitude values with equation **1.**
[b] Calculated with the k_{off} value obtained from the displacement experiment by using either the relationship $K_{d2}=k_{-2}/k_2$ or, particularly to S94A InhA, the equation **5.**
[c] Calculated from the relation $K_{d2} = k_{-2}/k_2$, or particularly to S94A InhA, by fitting the data to equation **4**.
[d] Calculated for S94A InhA with K_{d2}, k_3, and k_{-3} values by using equation **5.**
[e] Limiting rate constant for NADH dissociation estimated from the displacement of NADH from InhA-NADH complexes by high NAD^+ concentrations.
[f] Mean value calculated from the three K_{d2}(WT, I21V, and I47T) or K_{d2}^* (S94A) values yielded both from analysis of fast phase of NADH-binding and by using the k_{off} estimated in the displacement experiment.

The fast phase rate constants obtained for S94A InhA were best interpreted as a two-step mechanism, described by equation **2**, where a rapid bimolecular association between NADH and the mutant enzyme would be followed by a unimolecular isomerization of the binary complex,

$$k_{obs1} = (k_3[NADH]/([NADH]+K_{d2})) + k_{-3} \quad \textbf{[2]}$$

where K_{d2} is the NADH dissociation constant, k_{-2}/k_2, and k_3 and k_{-3} are, respectively, the forward and reverse rate constants for binary complex isomerization (see scheme in Fig. 7). This analysis yielded values of 67 ± 9 μM, 1014 ± 84 s^{-1}, and 65 ± 6 s^{-1} for K_{d2}, k_3, and k_{-3}, respectively. Hence, a minimum value for the second-order rate constant for the binding of NADH to S94A InhA enzyme was calculated to be 15 x 10^6 M^{-1} s^{-1} (k_3/K_{d2}), which is of similar magnitude to the k_2 values calculated for WT, I21V and I47T enzymes (see Table 3). A value of 4.0 ± 0.5 μM for the overall dissociation constant (K_{d2}*), which takes into account the bimolecular association between NADH and the E form of InhA enzyme and the following isomerization of E-NADH complex to E-NADH* complex, can be estimated from equation **3.**

$$K_{d2}^* = K_{d2}k_{-3}/(k_3+k_{-3}) \qquad \textbf{[3]}$$

The K_{d2} values estimated for WT, I21V, and I47T enzymes and K_{d2}* value estimated for S94A are in good agreement with the respective K_{d2} values calculated by using k_{off} values obtained independently in the displacement experiment with high NAD^+ concentrations described below (see Table 3).

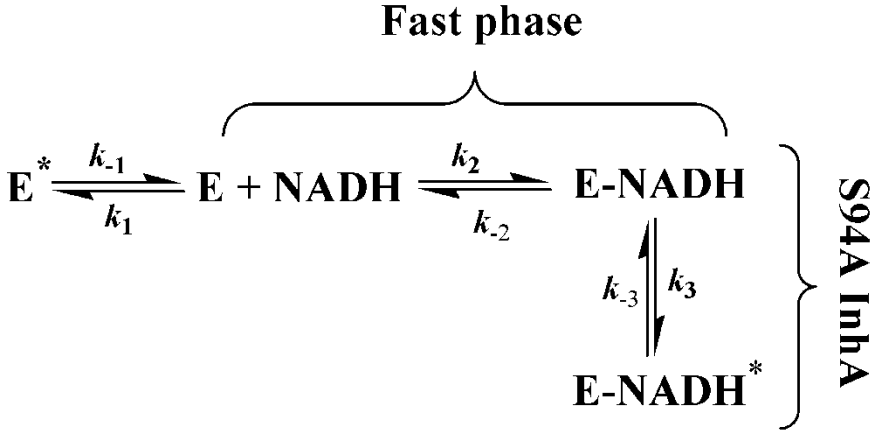


Fig. 7. Scheme showing the minimum kinetic mechanism proposed for NADH binding to WT and isoniazid-resistant InhA enzymes. The fast phase corresponds to a bimolecular association between InhA (E form) and NADH. In case of NADH binding to S94A the fast phase also includes a further isomerization of E-NADH complex to E-NADH*. The rate of slow phase is limited by the slow isomerization of E* form of InhA to the E form, where only the E form binds NADH.

The amplitude values for the fast phase, which were corrected for the dead time of the instrument by the method of Hiromi[33], were plotted versus NADH concentration and fitted to a quadratic equation (Eq. **7**) for a second-order binding process[34,35] as described in the Material and Methods section. This analysis yielded values of 0.3 ± 0.1, 7.6 ± 0.4, and 6.9 ± 0.8 μM for the equilibrium dissociation constant of WT InhA, I21V, and I47T, and also a value of 2.7 ± 0.2 μM for the overall K_d (K_d^*) for S94A InhA. These values are consistent with the K_{d2} values estimated for the WT, I21V, and to I47T InhA enzymes and also to the K_{d2}* (overall) estimated for the two-step NADH binding to S94A InhA from analysis of the fast phase rate constant values, supporting the validity of the method used for data analysis of the fast phase.

For the slow phase of NADH-binding to WT enzyme, the apparent first-order constant (k_{obs2}) values appeared to be independent of NADH concentration, whereas the I21V, I47T, and S94A mutant enzymes showed a striking concentration dependence (Fig. 8). There is a decrease in the slow phase rate constant values with increasing NADH concentration for the InhA mutant enzymes. This inverse dependence is characteristic of mechanisms in which the enzyme itself, not the enzyme-substrate complex, undergoes rate-limiting isomerization (most probably a conformational change of the free enzyme).[33]

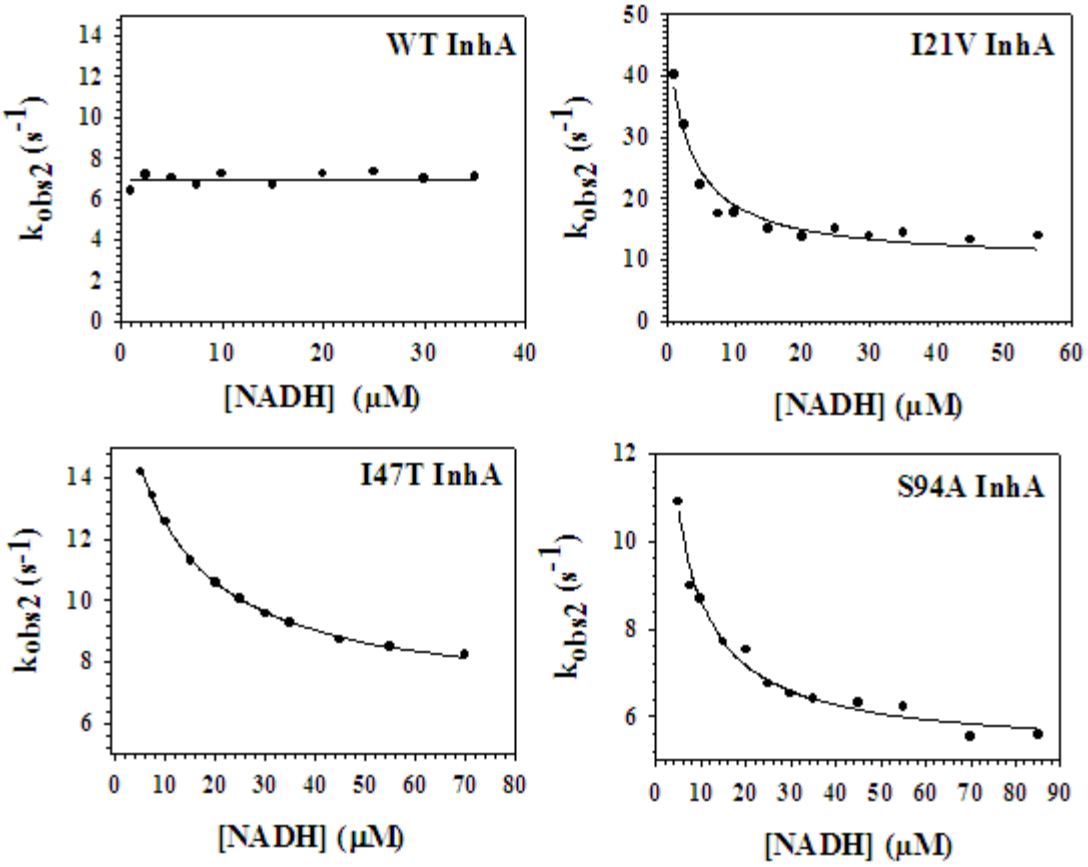


Fig. 8. Analysis of NADH concentration dependence of the k_{obs2} values calculated from the slow phase (second phase) of NADH-binding to WT, I21V, I47T, and S94A InhA enzymes. The I21V, I47T, and S94A InhA data were fitted to equation **4**, which describes a bimolecular reaction including two interconvertible forms of the enzyme (InhA), E and E^*, of which only one form (E) can bind the ligand (NADH) (see scheme depicted in Fig. 7).

Accordingly, the slow phase data obtained for the three INH-resistant InhA enzymes were tentatively fitted to an equation that describes a system in which the enzyme exists in two interconvertible forms, E and E*, and the substrate S can only bind to E (see Fig. 7):

$$k_{obs2} = (k_1 K_{d2}/([NADH] + K_{d2})) + k_{-1} \quad \textbf{[4]}$$

where K_{d2} is the equilibrium dissociation constant of InhA-NADH to InhA and NADH, and k_1 and k_{-1} are, respectively, the forward and reverse rate constants for the isomerization of free

enzyme form E (InhA that binds NADH) and E* (InhA* that does not bind NADH). Analysis of the data yielded values for k_1, k_{-1}, and K_{d2} constants that are given in Table 4.

Table 4. Kinetic constants estimated from the slow phase of NADH-binding to wild-type and isoniazid-resistant mutant InhA enzymes.

Kinetic constant	WT InhA	I21V InhA	I47T InhA	S94A InhA
$k_1(s^{-1})$[a]	n.d.	37 ± 4	11.2 ± 0.2	14 ± 4
k_{-1} (s^{-1})[a]	7.0 ± 0.2	9.7 ± 1.5	6.6 ± 0.2	5.2 ± 0.2
K_1[b]		3.8 ± 0.6	1.70 ± 0.05	2.7 ± 0.8
K_{d2}[c] (μM)		3 ± 1	11 ± 1	3.3 ± 1.6[c]
K_D(overall) (μM)		14 ±5[d]	30 ± 3[d]	12 ± 6[d]
		29 ± 5[e]	15 ± 2[e]	13 ±4[e]
	2.5 ± 0.3[f]	14.9 ± 0.7[f]	32 ± 2[f]	23 ± 3[f]

[a] Calculated by fitting the slow phase data (K_{obs2} versus [NADH]) to equation **6**
[b] Equilibrium constant for isomerization of E to E* calculated with the relation, $K1 = k_1/k_{-1} = [E^*]/[E]$.
[c] Calculated by fitting the slow phase data to equation **6**, however in specific to S94A InhA enzyme the K_{d2} term in equation **6** corresponds to the K_{d2}* (K_{d2} overall).
[d] Calculated by applying the K_{d2} and K_1 values, which were estimated by fitting the slow phase data to equation **6**, to the relationship $K_D(overall) = (K_1 + 1)K_{d2}$ (eq. **7**), where Kd2 = *k*-2/*k*2 for I21V and I47T enzymes. In particular to S94A InhA, the K_{d2} term corresponds to the K_{d2}^* value.
[e] calculated by using the equation **7** and the meam values for K_{d2} (I21V, and I47T) and K_{d2}^* (S94A) yielded from analysis of fast phase data.
f Overall dissociation constant obtained from the NADH concentration dependence of the algebraic sum of the amplitudes of the fast and slow phases (Amp total = Amp1 + Amp2). The K_D (overall) value was estimated by fitting the data to equation **1**.

It is noteworthy that the K_{d2} values of 3 ± 1, 11 ± 1, and 3.3 ± 1.6 (Table 4) estimated from the slow phase analysis for, respectively, I21V, I47T, and S94A InhA enzymes are consistent with the K_{d2} and K_{d2}^* values estimated from the fast phase data analysis (Table 3). This suggests that the equilibrium constants (K_{d2} and K_{d2}^*) estimated from both phases are probably describing the same process, a single bimolecular association between InhA and NADH for WT, I21V and I47T enzymes, and a bimolecular association followed by a unimolecular isomerization of the binary complex (E-NADH) to a more tightly bound complex (E-NADH*) for the S94A mutant enzyme. In a situation where the bimolecular process (E+S↔ES) is sufficiently faster than the unimolecular isomerization process (E↔E*), or, more

strictly, when the inequality k_2[NADH] + k_{-2} >> k_1 + k_{-1} holds, two phases, one fast and other slow can be observed in the stopped-flow. In this case, the dependence of the fast phase k_{obs1} on ligand concentration can be described by equation **1**, whereas the dependence of the slow phase k_{obs2} on ligand binding is described by equation **4**.

The pre-steady-state kinetics data on NADH binding to WT and INH-resistant InhA enzymes are consistent with the whole mechanism depicted in Fig. 7, where the observed fast phase corresponds to a rapid association between NADH and the E form of InhA enzyme that can bind NADH. The slow phase is rate-limited by the slow conversion of E* form of InhA into the E form that then binds NADH. For the S94A enzyme, in addition to a bimolecular association, the fast phase includes a unimolecular isomerization of E-NADH complex to E-NADH* (Fig. 7).

The hyperbolic decreasing value of k_{obs2} (slow) as a function of increasing NADH concentration that has been observed for INH-resistant I21V, I47T and S94A mutants could not be observed for INH-sensitive WT InhA. It may be because the equilibrium between E* and E is shifted to the latter or because the NADH K_{d2} (k_{-2}/k_2) value calculated for WT InhA (0.3 µM) is lower than the K_{d2} values for I21V (5.9 µM) and I47T (6.5 µM), and the K_{d2}^* value for S94A (4.0 µM). The k_{obs2} values (slow phase) could not be accurately measured when NADH concentrations lower than either 5 µM (I47T, S94A) or 1 µM (I21V) were used in the assays, probably because the amplitude values for the second phase were very small for these low NADH concentrations to confidently obtain estimates for k_{obs2} values. Accordingly, the k_{obs2} values obtained for WT InhA enzyme showed no dependence on the NADH concentration range used in the assays. Moreover, the average k_{obs2} value obtained for WT InhA corresponds to the isomerization rate constant for conversion of E to E* form of InhA enzyme (k_{-1}), as depicted in Fig. 7.

The overall K_D for NADH-binding mechanism depicted in Fig. 7 is given by the following relationship:

$$K_D(\text{overall}) = (K_1 + 1)K_{d2}, \qquad \textbf{[5]}$$

where K_1 is the equilibrium isomerization constant to the E^* from E form of InhA ($K_1 = E^*/E = k_1/k_{-1}$), and K_{d2} is the equilibrium dissociation constant of NADH from E-NADH complex given

by $K_{d2} = k_{-2}/k_2$ for WT, I21V and I47T enzymes, whereas. the K_{d2} (equation **5**) for S94A mutant corresponds to K_{d2}^* calculated from equation **3** ($K_{d2}^* = K_{d2}k_{-3}/(k_3+k_{-3})$). Thus, using the K_1 and K_{d2} values (Table 4) estimated from an analysis of the slow phase data, K_D(overall) values of 14 ± 5, 30 ± 3, and 12 ± 6 μM for, respectively, I21V, I47T, and S94A mutant enzymes could be readily determined from equation **5**. These estimates are also consistent with the K_D(overall) values calculated by using mean values for K_{d2} (I21V and I47T) and K_{d2}* (S94A) constants obtained from the analysis of both fast phase data and NADH displacement data (Tables 3 and 4).

The total amplitude changes derived from the biphasic enhancement in nucleotide fluorescence upon NADH binding to InhA enzymes are given as the algebraic sum of the amplitudes of the fast (Amp1) and slow (Amp2) phases (Total amplitude = Amp1 + Amp2). The NADH concentration dependence of Amp1 (fast phase) appears to corresponds to a bimolecular association between E form of InhA and NADH (followed by an isomerization in S94A InhA), which would be determined either by K_{d2} for WT, I21V and I47T enzymes or by K_{d2}^* for S94A InhA enzyme. In this case, the amplitude of the observed slow phase (Amp2) could include to some extent a contribution of the fast process, in addition to the characteristic change of the slow process.[33] Accordingly, the overall dissociation constant, K_D, could be determined from the NADH concentration dependence of total signal amplitude.[33] Indeed, as shown in Table 4, the K_D values obtained by fitting the total amplitude values as a function of NADH concentration to equation **7** (Material and Methods section) are in excellent agreement with the K_D (overall) values obtained from the analysis of rate constants, supporting the mechanism proposed here for NADH binding to WT and INH-resistant enoyl reductase enzymes from *M. tuberculosis* (Fig. 7). In addition, from this analysis we verified a 6-, 13-, and 9-fold reduction in NADH affinity for I21V, I47T, and S94A mutant enzymes as compared to WT enzyme (Table 4). Contrary to what was observed for WT InhA, the slow-phase amplitude values plotted against NADH concentration were best fitted to the Hill equation (eq. **8**) for isoniazid-resistant InhA mutant enzymes, yielding Hill coefficients of and 1.24 ± 0.05, 1.31 ± 0.03, and 1.26 ± 0.03 for I21V, I47T, and S94A InhA, respectively (data not shown). A slight cooperativity has also been observed in static fluorescence measurements of NADH binding to I21V, I47T, and S94A mutant enzymes, where the Hill coefficient values obtained were 1.4 ± 0.1, 1.50 ± 0.04, and 1.64 ± 0.06, respectively.[16] These data indicate that the mutations in the NADH binding site of InhA

affected the intersubunit contacts within the homotetramer, leading to the slight cooperativity observed in the slow phase of NADH binding to the enzyme, which is rate limited by the interconversion of E* form (that does not bind NADH) to the E form (that binds NADH) of InhA. This hypothesis is supported by the crystal structures of WT, I21V, I47T, and S94A InhA in complex with NADH, in which several water-mediated intersubunit contacts are lost in the mutant enzymes. The loss of these intersubunit contacts could be an explanation to the weak cooperativity observed in the here described stopped-flow data on NADH binding to I21V, I47T, and S94A InhA mutant enzymes. In accordance with the mechanism depicted in Fig. 7, the InhA subunit conformation observed in the InhA crystal structures presented here are equivalent to the E form of InhA enzyme, since the four InhA structures are in complex with NADH and only this enzyme form is able to bind NADH. Thus, the loss of the water-bridged hydrogen bonds between the protein subunits of the homotetramer that was observed in the crystal structures of InhA mutant enzymes should lead to destabilization of the observed conformation of the InhA with NADH bound, and appearance of a weak NADH-binding dependent cooperativity in the isomerization of the E* to E form of InhA. The conformation of each enzyme subunit in an oligomer is constrained and stabilized by the quaternary bonds, i.e. its conformation is dependent upon its quaternary structure.[32]

The conformation equivalent to the E* form described above may be observed in a substrate-free structure of InhA. In fact, we have crystallized and collected X-ray diffraction data of the S94A InhA enzyme in its apo-form in two distinct space groups, P1 and C2 (unpublished data), which can correspond to two different forms of InhA in solution (E and E^*), in agreement with the pre-steady-state data reported here. The refinement of the WT and S94A crystal structures in their apo-form is currently underway in our laboratory.

NADH dissociation rate constant from WT, I21V, I47T, and S94A binary complexes.

The dissociation rate constants of NADH from its complex with the InhA enzymes were studied in a stopped-flow apparatus by observing the decrease in fluorescence intensity of the InhA-NADH complexes due to displacement of NADH by NAD^+. The loss in nucleotide fluorescence accompanying dissociation of NADH from the InhA-NADH binary complexes was monophasic for all NAD^+ concentrations tested and the k_{obs} values were thus fitted to a single-exponential equation. The experimental conditions of our experiments and data analysis were as

described by Fatania *et al.*[36] and Chock and Gutfreund.[37] The NAD^+ concentration range (0.75-20 mM) is larger than the NADH and InhA concentrations, which were present at very much lower concentrations (5 µM and 10 µM, respectively, mixing chamber concentrations). Under these conditions, there is no general solution to the rate equations describing the competition reaction (Fig. 9), which can only be solved analytically when steady-state assumptions can be made.[37]

$$\text{InhA-NADH} \underset{k_{on\,1}}{\overset{k_{off\,1}}{\rightleftharpoons}} \text{NADH} + \text{InhA} + \text{NAD}^+ \underset{k_{off\,2}}{\overset{k_{on\,2}}{\rightleftharpoons}} \text{InhA-NAD}^+$$

Fig. 9. Representation of the reactions that occur in the displacement experiment after mixing the pre-incubated mix of InhA (20 µM) and NADH (10µM) with excess of NAD^+.

As the concentration of NAD^+ approaches infinity, the concentration of unliganded enzyme can be assumed to be in a steady-state and the relationship between the observed first-order rate constants (k_{obs}) and the NAD^+ concentration can be interpreted by the following equation[37] with the assumption that d[free enzyme]/dt = 0,

$$k_{obs} = k_{on\,1}[\text{InhA}] + k_{off\,1} \qquad \textbf{[6]}$$

where [InhA] is the concentration of free InhA enzyme, $k_{on\,1}$ is the association rate constant for formation of [InhA-NADH] binary complex, and $k_{off\,1}$ is the dissociation rate constant of NADH from the binary complex. The concentration of the free InhA enzyme is dependent on the relative affinity and concentration of the two ligands, NADH and NAD^+. As previously observed[37] for displacement of NADH from the [GPDH-NADH] complex by NAD^+, our k_{obs} values decreased hyperbolically with increasing concentration of NAD^+, probably due to a decrease in free InhA concentration. Thus, the experimental values of k_{obs} obtained for the displacement of NADH from InhA-NADH complex, could be fitted to a rectangular hyperbola that extrapolates to the true value for the rate constant ($k_{off\,1}$) describing the dissociation of NADH (Fig. 10). Computer simulation studies on displacement reactions have demonstrated that this behavior can be observed when the displacing ligand (NAD^+) has a lower affinity than the one being displaced

(NADH) if the magnitude of the rate constant for bimolecular association between enzyme and NAD^+ ($k_{on\,2}$) in relation to the other constants is high enough (see scheme in Fig. 9).

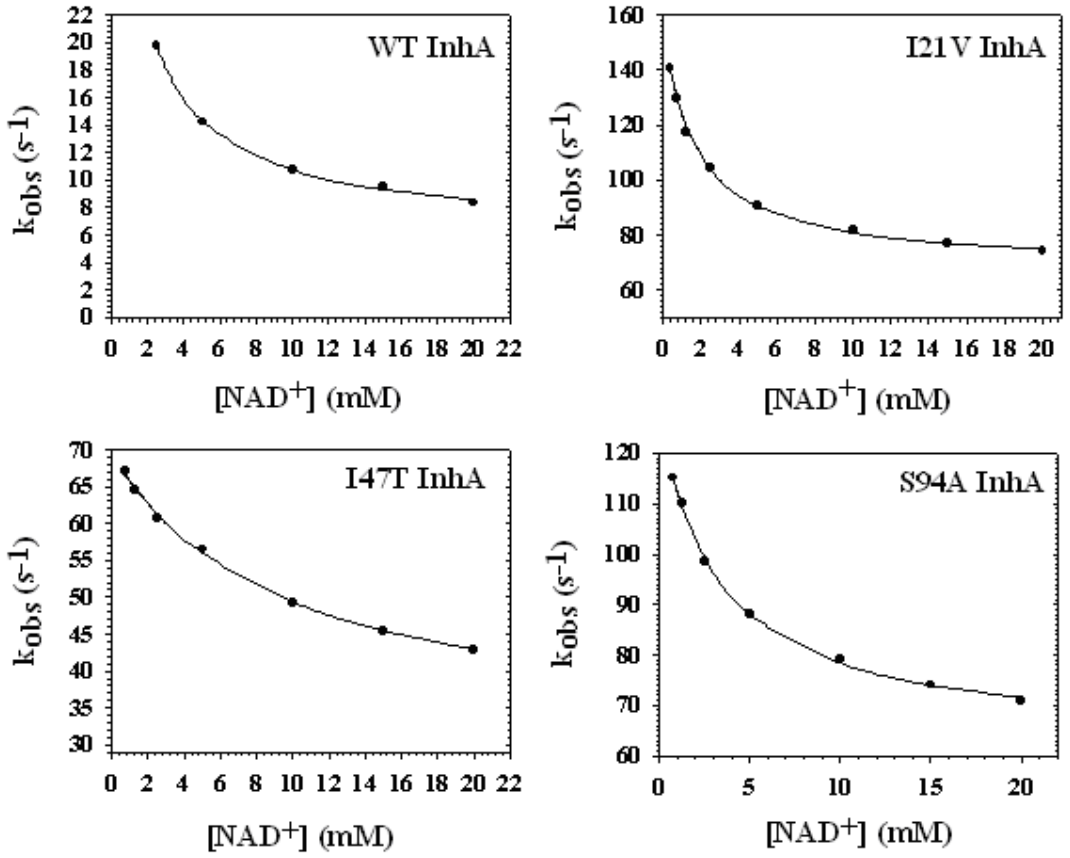


Fig. 10. Relationship between the observed first-order rate constants, determined with the stopped-flow apparatus, for displacement of NADH from its complex with WT (A), I21V (B), I47T (C), and S94A InhA enzymes, and the concentration of NAD^+. The data were obtained by mixing NAD^+ solutions (of twice the indicated concentration) with a mixture of 20 µM of each InhA enzyme and 10 µM NADH at 25°C. The curves were fitted to retangular hyperbolas that extrapolate to the true value for the rate constant that describes the dissociation of NADH from InhA. Values of 6.1 (± 0.4), 68 (± 1), 30 (± 2) s^{-1}, and 62 (± 1) were calculated for WT, I21V, I47T, and S94A InhA, respectively.

It would thus imply that the combination of NAD^+ with free enzyme is fast enough to decrease the concentration of the latter to its equilibrium value in the very early stages of the reaction. It has been shown for displacement of NADPH by $NADP^+$ from isocitrate dehydrogenase-NADPH binary complex that only a fraction of NADPH is displaced at equilibrium for relatively low $NADP^+$ concentrations and therefore the half-time for approaching this equilibrium is smaller than that for complete ligand displacement.[36] In our experiments, the k_{obs} values for the displacement reaction are greater than the intrinsic rate constant for NADH dissociation. True rate constant values of 6.1 ± 0.4, 68 ± 1, 30 ± 2, 62 ± 1 s^{-1} (Fig. 10) were estimated for the dissociation of NADH from, respectively, WT-NADH, I21V-NADH, I47T-NADH, and S94A binary complexes (Table 3). These values are in good agreement with the values of 5 ± 5, 85 ± 4, 53 ± 2, and 65 ± 6 s^{-1} that were estimated for either k_{-2} (WT, I21V, and I47T) or k_{-3} (S94A) from the plot of k_{obs1} values against NADH concentration (Fig. 6), suggesting that the rate constant values estimated from the displacement experiment corresponds to either the true dissociation rate constant value k_{-2} or to the reverse isomerization rate constant k_{-3}, which appears to limit the rate of NADH dissociation from the E^*-NADH in S94A InhA (see scheme in Fig. 7). It should be pointed out that NADH displacement was monophasic for both WT and INH-resistant mutant enzymes with all NAD^+ concentrations tested. Furthermore, the fluorescence amplitude values estimated from the NADH displacement curves obtained when 20 μM of each InhA enzyme pre-incubated with 10 μM NADH was mixed with an equal volume of 40 mM NAD^+ are approximately equal to the total fluorescence amplitude (i.e. the sum of amplitudes of fast and slow phases) estimated from double exponential NADH-binding curves obtained when 20 μM of each InhA enzyme was mixed with a equal volume 10 μM NADH in the stopped flow apparatus (data not shown). These results are consistent with the mechanism depicted in Fig. 7, since the fluorescence amplitude of the slow phase would correspond to equivalent InhA-NADH complex formed in the fast phase of NADH-binding to the InhA enzymes. Accordingly, the displacement of NADH from InhA by NAD^+ appears to be monophasic and the fluorescence amplitude value of the process is equivalent to the amplitude value of the fast and slow phases of NADH-binding to InhA.

DISCUSSION

The more prominent difference observed between WT InhA and the INH-resistant InhA mutant enzymes in our pre-steady-state studies on NADH binding is in the value of the limiting rate constant to the dissociation of NADH from the InhA-NADH binary complexes (k_{off}). The k_{off} values found for I21V, I47T, and S94A are, respectively, 11-, 5-, and 10-fold larger than the k_{off} value estimated for WT InhA. There are nearly no differences between the values for bimolecular association rate constants (k_2) of WT enzyme, and I21V and S94A mutant enzymes. The k_2 value of 8.17 ± 0.08 ($x10^6$) M^{-1} s^{-1} for I47T enzyme is only 1.8-fold lower than the k_2 value of 15.5 ± 0.2 ($x\ 10^6$) M^{-1} s^{-1} for WT enzyme. Hence, the reduction in affinity for NADH previously observed[16] and thoroughly described here for I21V, I47T, and S94A proteins can be attributed to the larger k_{off} values observed for the INH-resistant mutant InhA enzymes. This rate constant appears to be equivalent to the k_{-2} constant for WT, I21V, and I47T InhA enzymes, whereas for S94A InhA the NADH dissociation appears to be limited by the reverse isomerization rate constant k_{-3}. In addition, the stopped-flow data on NADH binding suggest a mechanism where there are two forms of InhA in equilibrium (E and E*), and only the E form can bind NADH. The INH-NAD adduct has been proposed to be formed either in solution[15] or on the InhA active site through the binding of KatG-activated INH to the NADH-bound enzyme.[13] Based on our results, *M. tuberculosis* strains harbouring the mutant InhA enzymes described here would have a higher E*/E intracellular ratio than would the WT InhA enzyme, which would imply that a larger amount of mutant enzyme should be present in the E* form as compared to WT InhA. Since NADH molecule is part of the INH-NAD adduct that inhibits the InhA enzyme, it is tempting to suggest that this adduct can only bind to the E form of InhA. Under physiological NADH concentration (<10 µM),[38] the larger E*/E ratio for the INH-resistant mutant enzymes would result in an increase in the concentration of the enzyme form that does not binds NADH (E*) and ensuing reduction in enzyme inhibition by INH-NAD.

The three mutations studied here either are located on or appear to indirectly affect the dinucleotide binding loop structure. The S94A mutation result in a break of hydrogen bond of Ser94 residue with a conserved water molecule, which is also present in the crystal structures of apo-SDR enzymes and that helps maintain the Gly-rich loop conformation.[19] In the conservative substitution of Val for Ile in the penultimate position of the dinucleotide binding loop only van

der Waals interactions between nicotinamide moiety and CD1 atom of Ile21 are lost. Finally, in the I47T mutant protein only a perturbation in the hydrogen bond network involving a conserved water, the second residue of dinucleotide binding loop (Ile15), the 2'- and 3'-hydroxyl groups of adenine ribose, and Phe41 residue are observed owing to the presence of the slightly polar Thr15 side chain and of a new water molecule which hydrogen bonds to OG1 atom of Thr47 residue in the crystal structure of the binary complex I47T-NADH. It should be pointed out that mutations in conserved residues of the loop have been correlated with attenuation or elimination of enzyme activity[39-41] and it is thus not surprising that the mutations reported here result in only very small changes in the backbone conformation of the dinucleotide binding loop, since this could result in an inactive enzyme that would be lethal to *M. tuberculosis*. These small structural differences probably result in a decreased stability of the NADH bound to InhA, and thus are implicated in the higher k_{off} values observed for isoniazid-resistant enzymes in the pre-steady-state kinetics studies on NADH binding reported here. As NADH molecule is part of the INH-NAD adduct that inhibits the InhA enzyme, in which only the 4*S* hydrogen of niconinamide ring is replaced by the isonicotinic acyl moiety of isoniazid, it appears plausible that the lower affinity for the NAD moiety of INH-NAD adduct should contribute to INH resistance. Consistent with this proposal, the *in vitro* kinetics inactivation of the I21V and I95P mutant enzymes has been compared with that of the WT enzyme, and both mutations resulted in increased values for the NADH equilibrium dissociation constant and decreased values for the apparent first-order rate constant for enzyme inactivation.[16] Contrary to our proposal, the kinetic and thermodynamic parameters for the interaction of isonicotinyl-NAD with INH-resistant I21V, I47T, and S94A InhA mutant enzymes were found to be similar to those of the WT enzyme in slow-onset inhibition experiments.[15] These results prompted the authors to suggest that INH resistance of *M. tuberculosis* strains harboring *inhA*-structural gene mutations may be explained by an alteration of protein-protein interactions between InhA and other components of the FAS-II system. However, it remains to be shown whether *inhA*-structural gene mutations identified in INH-resistant clinical isolates of *M. tuberculosis* will affect the sensitivity of InhA to INH in the context of multienzyme complex. Moreover, increasing NADH levels have been shown to protect InhA against inhibition by the INH-NAD adduct formed upon KatG activation of INH, and mutations in mycobacterial *ndh* gene have been proposed to result in increased intracellular NADH concentrations, which competitively inhibits the binding of INH-NAD adduct to InhA.[42]

The crystal structures of INH-resistant I21V, I47T, and S94A InhA mutant enzymes complexed with NADH presented here showed that a few water-mediated interactions are lost in the three orthogonal twofold axes. These differences can be correlated with the slight cooperativity observed for isoniazid-resistant mutants.[16] In addition, it can be reconciled with our stopped-flow studies on the slow phase of NADH binding since the INH-resistant mutant InhA homotetramers are likely less stable the WT enzyme.

Since the results of pre-steady-state kinetics on NADH binding reported here suggest the presence of two forms of InhA enzyme in solution, it would be interesting to carry out similar pre-steady-state kinetic experiments on binding of INH-NAD adduct to WT and INH-resistant InhA enzymes to assess the binding mechanism and affinity of INH-NAD adduct and compare with reported data on slow-onset inhibition experiments.[15] The results presented here provide both a mechanism for NADH binding to *M. tuberculosis* enoyl-ACP reductase and compelling evidence for the effects of *inhA*-structural gene mutations, identified in clinical isolates of *M. tuberculosis* strains resistant to isoniazid, on structural and kinetic features of these enzymes. It is hoped that the results reported here will assist the development of new anti-tubercular agents that can overcome isoniazid-resistance. Moreover, our results highlight the importance of regarding water molecules when analyzing protein crystal structures. This issue is critical for future docking calculations and structure-based drug design and optimization studies of lead inhibitors against both wild-type and isoniazid resistant enoyl-ACP reductases from *M. tuberculosis*.

MATERIAL AND METHODS

Determination of oligomeric state of WT, I21V, I47T, and S94A InhA enzymes — The oligomeric state of wild-type and isoniazid-resistant enoyl-ACP reductase homogenous proteins was estimated by gel-permeation chromatography on a Superdex™ 200 HR column (1.0 cm × 30 cm) (Amersham Biosciences). The column was eluted with 20 mM Pipes containing 0.2 M NaCl, pH 7.3 at a flow rate of 0.4 mL min^{-1}. The eluate was monitored at 215 and 280 nm and the column was calibrated with the following protein standards (Amersham Biosciences): Ribonuclease A (13,700 Da) from bovine pancreas, Chymotrypsinogen (25,000 Da) from bovine pancreas, Ovalbumin (43,000 Da) from hen egg, Albumin (67,000 Da) from bovine serum, Purine nucleoside phosphorylase from *M. tuberculosis* (82,720 Da), Aldolase (158, 000 Da), Catalase (232,000 Da), and Ferritin (440,000 Da). Blue Dextran 2000 was used to determine the void volume (V_0). The K_{av} value was calculated for each protein using the equation V_e-V_0/V_t-V_0, where V_e is the elution volume for the protein and V_t is the total bed volume, and K_{av} was plotted against the logarithm of standard molecular weights.

Crystallization, refinement and analysis — WT, I21V, I47T and S94A InhA enzymes were expressed and purified to homogeneity as described elsewhere.[8,16] The binary complexes of InhA-NADH were crystallized by hanging drop vapor diffusion method under similar conditions as described by Dessen *et al*.[27] A 3 µL drop of reservoir solution containing 50 mM Hepes pH 7.2, 8 to 12% methylpentanediol (MPD), and 50 mM sodium citrate (pH 6.5) was mixed with an equal volume of 13 $mg.mL^{-1}$ protein solution (containing a 1:2 ratio of enzyme and NADH) on a silanized cover slit that was inverted and sealed above the precipitant solution. Crystals were cryoprotected by transfer to crystallization solution with 20% glicerol and flash-frozen at 104K in a cold nitrogen stream generated and maintained with an Oxford Cryosystem. All data sets were collected at 1.4310 Å wavelength on a CCD detector (MAR CCD) using synchrotron radiation source (Station PCr, LNLS, Campinas, Brazil).[43] The data sets were processed using MOSFLM program and scaled with SCALA program.[44] All protein-NADH complexes have crystallized in the hexagonal space group $P6_222$, with one subunit per asymmetric unit. Initially, a crystal structure of the binary complex WT InhA-NADH has been solved at 2.2 Å by molecular replacement with the coordinates of the previously solved WT InhA structure without

NADH and water molecules[27] (PDB code: 1ENY) as a search model, using data collected by cryocrystallography. This WT InhA structure without NADH and water molecules has been used as starting model in standard molecular-replacement methods using the program AMoRe[45] to solve the crystal structures of WT InhA, I21V, I47T, and S94A InhA mutant enzymes in complex with NADH, which have been refined to, respectively, 1.92 Å, 2.0 Å, 2.2 Å, and 2.3 Å (Table 1) with X-PLOR-NIH[46,47] and REFMAC 5.2[48] programs. The structures have initially been refined with the X-PLOR-NIH[46,47] program, where only reflections between 6.0 Å and the high resolution cutoff were used in refinement. Subsequently, the structures were further refined with the maximum likelihood based program REFMAC 5.2, using a solvent mask option, which allow inclusion of low resolution data in refinement. No sigma cutoff has been applied to select the unique reflections that have been used in the refinement process. A test set of 5.1% reflections, which were used for free R value calculations through the cross-validation method, have been randomly selected and omitted from the refinement process. Since both all diffraction data sets are isomorphous to and the structures have similar unit cell dimensions to the WT InhA starting model, there will be correlations between the different data sets of test reflections and the working data set used in the refinement of the WT InhA starting model. Thus, the test data sets could not provide an independent check of the quality of the building models.[49] To minimize this problem, high temperature simulated annealing (starting temperature between 2000-4000K) has been performed in the begining of the refinement for WT InhA, I21V, I47T, and S94A InhA mutant proteins to uncouple the working and free R values, i.e., to remove remove bias toward the test set.[49,50] The simulating annealing steps have been carried out throughout of refinement process with X-PLOR-NIH program.[46,47] In order to minimize the model bias,[51] the models were checked and rebuilt when necessary in the program XTALVIEW (Xfit tool)[52] against simulating annealing electron density omit maps, in which either 10 contiguous residues or spheres encompassing no more than 10% of the total residues were omitted, or against omit maps calculated *on the fly* of XTALVIEW program (Xfit tool)[52] together with the SHAKE option. At the final stages of the refinements, water molecules were manually added to electronic density peaks $\geq 3\sigma$ in ($|F_{obs}|$ - $|F_{calc}|$) density maps, and the model subject to cycles of isotropic B-factor refinement. The solvent molecules were subsequently rejected from the models if their B-factors were above 60 $Å^2$ after refinement or if they were not located within 2.2-3.5 Å from one or more possible hydrogen bonding donor(s) or acceptor(s) atom(s). The stereochemistries of the final

models were verified with the program PROCHECK[53] and MOLPROBITY[54] (http://molprobity.biochem.duke.edu/). In the PROCHECK evaluation, 90.2-91.6% of residues are found in the most favorable region of the classical Ramachandran plot (Table 1). Only three residues from crystal structure of WT InhA (Asp42, Ala157, and Asn159) and two residues from crystal structure of I21V, I47T, and S94A InhA mutant proteins (Ala157 and Asn159), which are located in turns, are found in the disallowed region of the Ramachandran plot from PROCHECK[53]. However, the electron density for Asp42, Ala157, and Asn159 residues is sufficiently clear to unequivocally assign their conformations in the structures of InhA proteins (suplementary material). Actually, Asp42, Ala157, and Asn159 are not ouliers when the InhA structures were evaluated with the program MOLPROBITY[54] (Table 1), which use an updated φ,ψ plot (Ramachandran plot) for over 100,000 well-ordered non-Gly, non-Pro, non pre-Pro residues in 500 proteins at 1.8 Å resolution or better.[54] Thus, the conformations oserved for these residuwes are probably genuine. Final statistics for all the refined structures are summarized in Table 1. Atomic coordinates have been deposited in the RCSB Protein Data Bank[55] with the following PDB ID codes: **2AQ8** (WT InhA), **2AQK** (S94A), **2AQI** (I47T), and **2AQH** (I21V).

Clear electronic density was found to unambiguously position all moieties of NADH molecule in the models of the four structures at the beginning of each refinement. No clear electronic density was found for the first two residues of WT InhA structure and for the first residue of the InhA mutant enzymes. Three water molecules in special positions have been added to the final model of WT InhA-NADH structure. In the structures of I21V InhA-NADH, I47T InhA-NADH, and S94A InhA-NADH, one water, three water, and two water molecules, have, respectively, been included in special positions in the final model. The waters in special positions have been refined with an occupancy value of 0.5. The analysis of the interactions between NADH and protein residues has been made with a cut off distance of 3.5 Å for hydrogen bonding and of 4.5 Å for van der Waals and hydrophobic interactions. The molecular-surface areas were calculated using the program Swiss PDB Viewer v.3.7 (http://www.expasy.org/spdv), a probe radius of 1.4 Å and a fixed radius for all atoms. The three symmetry-related partners, which make part of a tetramer in the crystal packing, were generated in the CNS program using the input file neighbours.inp.[56]

Pre-steady-state kinetics of NADH binding — Rapid reactions were followed using an Apllied-Photophysics SX-18MV-R (Leatherhead, UK) stopped-flow spectrofluorimeter, operated at 25°C, in fluorescence mode (dead time ≤ 1.5 ms). The excitation wavelength was selected by focusing a 150-W Xenon arc lamp onto a monochromator fitted with a 250 nm holographic grating. The emitted light was selected with a WG420 Scott filter, positioned between the photomultiplier and the sample cell as previously described.[57,58] Data were stored on an Acorn A5000 computer, and analyzed by non-linear regression. Data acquisition was carried out using a split time base (20 and 200 ms) with the first half of data acquired over 10% of the time period monitored for the second half of the split time base. This procedure allows for more accurate determination of the rate and amplitude of an exponentially changing signal.[59] The enzyme and coenzyme concentrations are the final concentrations after mixing unless stated otherwise.

The pre-steady state kinetics of NADH binding to either 2 μM (WT InhA) or 4 μM (I21V, I47T, S94A mutant InhAs) enzyme concentrations was determined by monitoring the nucleotide fluorescence using an emission cut-off filter ≥ 420 nm. In order to minimize inner filter effects at high NADH concentrations, an excitation wavelength of 370 nm (band pass = 4.65 nm) was chosen, which is on the red edge of the longest wavelength absorption band of NADH. All measurements were performed in 100 mM PIPES pH 7.0 using a 0.5-85 μM NADH concentration range. At least five stopped flow traces were averaged prior to non-linear regression analysis. All trace curves for NADH binding were characterized by a biphasic enhancement in nucleotide fluorescence and the data were best fitted to a double exponential function. The amplitude data from the fast phase were fitted to a quadratic equation (eq. **7**) relating fluorescence change to NADH concentration for a second-order binding process.[34,35]

$$\frac{\Delta F}{\Delta F_{max.}} = \frac{1}{2[E]_0}\left\{(K_d + [E]_0 + [S]_0) \pm \sqrt{(K_d + [E]_0 + [S]_0)^2 - 4[E]_0[S]_0}\ \right\} \qquad \textbf{[7]}$$

In equation **7**, $[E]_0$ and $[S]_0$ represent the initial enzyme and substrate concentrations, respectively. ΔF is the fluorescence amplitude estimated for each NADH concentration, ΔF_{max} is the maximum fluorescence amplitude at saturating NADH concentration.

The fluorescence amplitude values estimated from the slow phase of NADH-binding to WT InhA were fitted to equation **7**. However, the amplitude values calculated for the slow phase of NADH-binding to I21V InhA, I47T InhA, and S94A InhA mutant enzymes were best fitted to the Hill equation (eq. **8**),

$$\Delta F = \Delta F_{max} [S]^h / K_{0.5} + [S]^h \quad \textbf{[8]}$$

Where $K_{0.5}$ is the NADH concentration in which 50% active sites are occupied, [S] is the total NADH concentration, and h is the Hill coefficient.[60]

Displacement of NADH by NAD^+ from WT InhA-NADH, I21V InhA-NADH, I47T InhA-NADH, and S94A InhA-NADH binary complexes — The apparent NADH dissociation rate values were determined independently by measuring the quenching of nucleotide fluorescence that resulted from the displacement of NADH from the InhA-NADH binary complexes by NAD^+. In order to obtain a reliable estimate to the actual limiting value for the dissociation rate constant of NADH, NAD^+ was used at a series of increasing concentrations. In these experiments, 20 μM of enoyl reductase pre-incubated with 10 μM NADH was rapidly mixed with 0.75, 1.5, 2.5, 5, 10, 20, 30, and 40 mM NAD^+ (syringe concentrations), and NADH release was measured by monitoring the fluorescence (λ_{exc} = 370 nm, band pass = 4.65 nm, $\lambda_{em} \geq 420$ nm). All reactions were measured at 25°C in 100 mM Pipes, pH 7.0.

ACKNOWLEDMENTS

Financial support for this work was provided by Millennium Initiative Program MCT-CNPq, Ministry of Health/DECIT – Secretary of Health Policy and PRONEX/FAPERGS/CNPq (Brazil) to DSS and LAB. DSS (304051/1975-06), LAB (520182/99-5), and WFAJr (300851/98-7) are research career awardees from the National Research Council of Brazil (CNPq).

REFERENCES

1. 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.

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

3. World Health Organization. Anti-tuberculosis drug resistance in the world. Third Global Report. 2004.

4. 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.

5. Fox, H. H. (1952). The chemical approach to the control of tuberculosis. *Science* **116,** 129-134.

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

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

8. Quémard, A.; Sacchettini, J.C.; Dessen, A.; Vilchèze, C.; Bittman, R.; Jacobs, W.R.Jr. & Blanchard, J.S. (1995). Enzymatic characterization of the target for isoniazid in Mycobacterium tuberculosis. *Biochemistry*, **34**, 8235-8241.

9. Vilchèze, C.; Morbidoni, H.R.; Weisbrod, T.R.; Iwamoto, H.; Kuo, M.; Sacchettini, J.C. & Jacobs, W.R.Jr. (2000). Inactivation of the inhA-encoded fatty acid synthase II (FASII) enoyl-

acyl carrier protein reductase induces accumulation of the FASI end products and cell lysis of Mycobacterium smegmatis. *J. Bacteriol.* **182**, 4059-4067.

10. 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, W.R.Jr. (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.

11. Kremer, L.; Dover, L.G.; Morbidoni, H.R.; Vilchèze, C.; Maughan, W.N.; Baulard, A.; Tu, S.-C.; Honoré, N.; Deretic, V.; Sacchettini, J.C.; Locht, C.; Jacobs, W.R.Jr. & 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.

12. 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.

13. 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.

14. Lei, B., Wei, C.-J. & Tu, S.-C. (2000). Action mechanism of antitubercular isoniazid. Activation by Mycobacterium tuberculosis KatG, isolation, and characterization of inha inhibitor. *J. Biol. Chem.* **275,** 2520-2526.

15. 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.

16. Basso, L. A.; Zheng, R.; Musser, J. M.; Jacobs, W. R. Jr. & 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.

17. 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.

18. Persson, B., Kallberg, Y., Oppermann, U. & Jörnvall, H. (2003). Coenzyme-based functional assignments of short-chain dehydrogenases/reductases (SDRs). *Chem. Biol. Interact*. **143-144,** 271-278.

19. Bottoms, C.A.; Smith, P.E. & Tanner, J.J. (2002). A structurally conserved water molecule in Rossmann dinucleotide-binding domains. *Protein Sci.*, **11**, 2125-2137.

20. Carugos, O. & Argos, P. (1997). NADP-dependent enzymes. I: Conserved stereochemistry of cofactor binding. *Proteins* **28,** 10-28.

21. Stewart, M. J., Parikh, S., Xiao, G., Tonge, P. J. & Kisker, C. (1999). Structural basis and mechanism of enoyl reductase inhibition by triclosan. *J. Mol. Biol.* **290,** 859-865.

22. Lee, H. H., Yun, J., Moon, J., Han, B. W., Lee, B. I., Lee, J. Y., Suh, S. W. (2002). Crystallization and preliminary X-ray crystallographic analysis of enoyl-acyl carrier protein reductase from Helicobacter pylori.
Acta Crystallogr. D Biol. Crystallogr. **58**, 1071-1073.

23. Roujeinikova, A., Levy, C. W., Rowsell, S., Sedelnikova, S., Baker, P. J., Minshull, C. A., Mistry, A., Colls, J. G., Camble, R., Stuitje, A. R., Slabas, A. R., Rafferty, J. B., Pauptit, R. A., R Viner, Rice, D. W. (1999). Crystallographic analysis of triclosan bound to enoyl reductase. *J. Mol. Biol.* **294**, 527-535.

24. Pidugu, L. S., Kapoor, M., Surolia, N., Surolia, A. & Suguna, K. (2004). Structural basis for the variation in triclosan affinity to enoyl reductases. *J. Mol. Biol.* **343**, 147-155.

25. Schroeder, E. K., Basso, L. A., Santos, D. S. & de Souza, O. N. (2005). Molecular dynamics simulation studies of the wild-type, I21V, and I16T mutants of isoniazid-resistant Mycobacterium tuberculosis enoyl reductase (InhA) in complex with NADH: toward the understanding of NADH-InhA different affinities. *Biophys. J.*, **89,** 876-884.

26. Ward, W. H. Holdgate, G. A. Rowsell, S. McLean, E. G. Pauptit, R. A. Clayton, E. Nichols, W. W. Colls, J. G. Minshull, C. A. Jude, D. . Mistry, A. Timms, D. Camble, R. Hales, N. J. Britton, C. J. & Taylor, I.W. (1999). Kinetic and structural characteristics of the inhibition of enoyl (acyl carrier protein) reductase by triclosan. *Biochemistry* **38,** 12514-12525.

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

28. Cleland, W. W., Frey, P. A. & Gerlt, J. A. (1998). The low barrier hydrogen bond in enzymatic catalysis. *J. Biol. Chem.* **273,** 25529-25532.

29. Cornell, W. D., Cieplack, C. P., Bayly, C. I., Gould, I. R., Merz, K. M., Fergunson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (**1995).** A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules**.** *J. Am. Chem. Soc.* **117,** 5179-5197.

30. Wierenga, R.K., De Maeyer, M.C.H. & Hol, W.G.J. (1985). Interaction of pyrophosphate moieties with alpha.-helixes in dinucleotide-binding proteins. *Biochemistry,* **24**, 1346-1357.

31. Fersht, A. (1999) Forces between molecules, and binding energies. In *Structure and mechanism in protein science* (Fersht, A.), pp. 324-348, W.H. Freeman and Company, New York.

32. Monod, J., Wyman, J. & Changeux, J-P. (1965). On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* **12,** 88-118.

33. Hiromi, K. (1979) In *Kinetics of Fast Enzyme Reactions*. Halsted Press, New York, U.S.A.

34. Birdsall, B., King, R. W., Wheeler, M. R., Lewis, C. A., Goode, S. R., Dunlap, R. B. & Roberts, G. C. K. (1983). Correction for light absorption in fluorescence studies of protein-ligand interactions. *Anal. Biochem.* **132,** 353-361.

35. Nakatani, H. & Hiromi, K. (1980). Analysis of signal amplitude in stopped-flow method for enzyme-ligand systems. *J. Biochem.* **87,** 1805-1810.

36. Fatania, H. R., Matthews, B. & Dalziel, K. (1982). On the mechanism of NADP+-linked isocitrate dehydrogenase from heart mitochondria. I. The kinetics of dissociation of NADPH from its enzyme complex. *Proc. Soc. London Ser. B* **214,** 369-387.

37. Chock, P. B. & Gutfreund, H. (1988). Reexamination of the kinetics of the transfer of NADH between its complexes with glycerol-3-phosphate dehydrogenase and with lactate dehydrogenase. *Proc. Natl. Acad. Sci. USA* **85,** 8870-8874.

38. Gopinathan, K. P., Sirsi, M. & Ramakrishnan, T. (1963). Nicotinamide-adenine nucleotides of *Mycobacterium tuberculosis* H37Rv. *Biochem. J.* **87,** 444-448.

39. Rescigno, M. & Perham, R. N. (1994). Structure of the NADPH-binding motif of glutathione reductase: efficiency determined by evolution. *Biochemistry* **33,** 5721-5727.

40. Nishiya, Y. & Imanaka, T. (1996). Analysis of interaction between the Arthrobacter sarcosine oxidase and the coenzyme flavin adenine dinucleotide by site-directed mutagenesis. *Appl. Environ. Microbiol.* **62,** 2405-2410.

41. Eschenbrenner, M., Chlumsky, L. J., Khanna, P., Strasser, F. & Jorns, M. S. (2001). Organization of the multiple coenzymes and subunits and role of the covalent flavin link in the complex heterotetrameric sarcosine oxidase. *Biochemistry* **40,** 5352-5367.

42. Vilchéze, C., Weisbrod, T. R., Chen, B., Kremer, L., Hasbón, M. H., Wang, F., Alland, D., Sacchettini, J. C. & Jacobs, W. R. Jr. (2005). Altered NADH/NAD^+ ratio mediates coresistance to isoniazid and ethionamide in mycobacteria. *Antimicrob. Agents Chemother.* **49,** 708-720.

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

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

45. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst.* A **50,** 157-163.

46. Brünger, A.T. (1992). X-Plor version 3.1 - A system for X-ray Crystallography and NMR, Yale University Press, New Haven, CT.

47. Schwieters, C.D., Kuszewski, J.J., Tjandra, N. &. Clore, G.M. (2003). The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.*, **160**, 66-74.

48. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst.* **D53,** 240-255.

49. Kleywegt, G. J. & Brünger, A. T. (1996) Checking your imagination: applications of the free R value. *Structure* **4,** 897-904.

50. Brünger, A. T. (1993) Assessment of phase accuracy by cross validation: the free R value. Methods and applications. *Acta Cryst.* **D49,** 24-36.

51. Hodel, A., Kim, S.-H. & Brunger, A. T. (1992). Model bias in macromolecular crystal structures. *Acta Cryst. A* **48**, 851-858.

52. McRee, D. E. (1999). XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125,** 156-165.

53. 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.

54. Lovell, S. C., Davis, I. W., Arendall, W. B. 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., Richardson, D. C. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*. **50**, 437-450.

55. Berman, H. M., Westbrook, J., Feng , Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acid Res.* **28,** 235-242.

56. Brünger, A. T., Paul D. Adams, P. D., Clore, G. M., Warren L. DeLano, W. l., Grose, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilgesg, M., Pannuh, N. S., Readi, R. J., Riceb, L. M., Simonsonj, T. & Warren, G. L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst.*.D**54**, 905-921.

57. Basso, L. A., Engel, P. C. & Walmsley, A. R. (1993). The mechanism of substrate and coenzyme binding to clostridial glutamate dehydrogenase during oxidative deamination. *Eur. J. Biochem*., **213**, 935-945.

58. Basso, L. A., Engel, P. C. & Walmsley, A. R. (1995). The mechanism of substrate and coenzyme binding to clostridial glutamate dehydrogenase during reductive amination. *Eur. J. Biochem*., **234**, 603-615.

59. Walmsley, A. R. & Bagshaw, C. R. (1989). Logarithmic timebase for stopped-flow data acquisition and analysis. *Anal. Biochem.* **176,** 313-318.

60. Hill, A. V. (1913). The combinations of haemoglobin with oxygen and with carbon monoxide. *Biochem. J.* **7,** 471-480.

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

62. Potterton, L.,McNicholas, S.,Krissinel, E., Gruber, J., Cowtan, K., Emsley, P. & Murshudov, G. N., Cohen, S., Perrakis, A., Noble, M. (2004) Developments in the CCP4 molecular-graphics project. *Acta Crystallogr. D Biol. Crystallogr*. **60(Pt 12 Pt 1),** 2288-9224.

FOOTNOTES

The abbreviations used are: ETH, ethionamide; FAD, Flavin adenine dinucleotide; FAS, fatty acid synthase system; GPDH, glycerol-3-phosphate dehydrogenase; INH, isoniazid; INH-NAD, isonicotinoyl-NAD; InhA=ENR, enoyl acyl carrier protein reductase; KatG, catalase/peroxidase from *Mycobacterium tuberculosis*; MDR-TB, multidrug-resistant tuberculosis; NAD, nicotinamide adenine dinucleotide; RMSD, root-mean-squared deviation; SDR, short chain dehydrogenase/reductase; 17β-HSD, 17β-hydroxysteroid dehydrogenase; WT, wild-type.

SUPPLEMENTAL MATERIAL

Analysis of intersubunit contacts

The solvent accessible surface areas of an isolated monomer and a tetramer (WT InhA) are approximately 11078 $Å^2$ and 33991 $Å^2$, respectively. Thus, around 2670 $Å^2$ of the molecular surface is buried per monomer on formation of tetramer. Interestingly, a cavity is delimited by the intersubunit contacts in the center of the tetramer with a molecular surface area of approximately 1639 $Å^2$. The molecular surface area values calculated for the crystal structures of isoniazid-resistant InhA enzymes were very similar to the values calculated for WT enzyme. The values calculated for monomer and tetramer of crystal structures of I21V, I47T, and S94A mutants were, respectively, 11053 $Å^2$ and 33805 $Å^2$, 11123 $Å^2$ and 34004 $Å^2$, and 11147 $Å^2$ and 34102 $Å^2$. A summary of all direct and water-mediated hydrogen bonding ocurring in the intersubunit interfaces about the three orthogonal molecular twofold axes (A, B, and C) of WT InhA homotetramer is given in Table 6.

Intersubunit contacts in the A axis — In the A axis hydrophobic interactions and direct and water-mediated hydrogen occur between side chains of helices $\alpha8^1$ and $\alpha8^2$ with residues from the loops connecting helix α8 and strand β7 (Ser247 - Gly255) in the subunits 1 an 2, respectively. There are hydrophobic interactions between Pro237, and Thr241 and Ala244 from helix α8 with Ala252, and Leu250, from the loop, respectively. Furthermore, a series of direct and water-mediated hydrogen bonds are observed for residues that lie in strands $\beta7^1$ and $\beta7^2$, which run antiparalell to each other, and which are located on the edge of the paralell β-sheets. Additional interface contacts are made by residues from the loop between helix α8 and strand β7, the turn between helices α7 and α8, the C-terminus of helix $\alpha5^2$, and the C-terminal stretch of residues Asp261-Leu269. Approximately 908 $Å^2$ of molecular surface area is buried per monomer in this interface.

Intersubunit contacts in the B axis — A four helix bundle is formed about B axis by helices $\alpha4^1$ (Tyr113-Leu135) and $\alpha5^1$ (Met161-Lys181) and their respective symmetry equivalents, $\alpha4^3$ and $\alpha5^3$. Helix $\alpha4^1$ forms extensive hydrogen-bonding contacts with its symmetry equivalent $\alpha4^3$,

which is supplemented by some hydrophobic contacts between $\alpha4^1$ and helices $\alpha4^3$, and $\alpha5^3$. Furthermore, several hydrophobic contacts are formed between Phe108 and Phe109 residues, which are located between strand β4 and helix α4, with non-polar residues from helices α4, and α5, and Trp160 residue of the symmetry-related InhA subunit. The phenylalanine residues (108 and 109) also make two water-mediated hydrogen bonds with Glu178 residue (helix α5), and the Phe109 forms a direct hydrogen bond with NZ atom of Lys132 residue (helix α4) from the symmetry equivalent subunit. A hydrophobic-packing interface is also observed between helices $\alpha5^1$ and $\alpha5^3$ with additional hydrogen-bonding contacts. In the WT InhA structure, approximately 947 $Å^2$ of molecular surface is buried per monomer in this interface. In addition, Pro151 and Ala154 residues, which are located in the turn between strand β5 and helix α5, make hydrogen bond interactions with, respectively, Arg173 and Ser170 residues of helix α5 from symmetry equivalent subunit.

Intersubunit contacts in the C axis — Contacts around the C axis are dominated by direct and water-mediated hydrogen bonds between residues from the turn connecting strand β5 and helix α5 from subunit 1 with residues from the symmetry equivalent turn and C-terminal residues (Asp264-Leu268) in subunit 4. Additional hydrogen bonds are formed between C-terminal residues (Asp264-Leu268) from subunit 1 and residues in the turn between strand $\beta5^4$ and helix $\alpha5^4$. Approximately 271 $Å^2$ of molecular surface is buried per monomer in this interface.

Table 5. Direct and water-mediated hydrogen bonds ocurring between NADH and protein residues in WT InHA-NADH, I21V InhA-NADH, I47T InhA-NADH, and S94A InHA-NADH binary complexes.

NADH moiety	NADH atom	WT InhA atom	Distance (Å)	Atom from protein bonded to water molecule	Distance (Å)
Nicotinamide Ribose	NO7	Ile194 N	2.8		
	NO7	Water 352 O	2.7	Tyr158 OH	2.7
	NN7	Ile194 O	3.2		
	NN7	Water380	3.4	Thr196 OG1	2.5
	NO2*	Lys165 NZ	2.9		
	NO3*	Lys165 NZ	3.0		
	NO3*	Ile95 O	3.1		
	NO3*	Water287 O	3.2	Lys165 NZ	2.7
				Ile95 O	2.7
				Ser123 O	2.7
	NO3*	Water288 O	2.8	His93 ND1	2.9
				His93 O	3.2
				Ile95 N	3.3
				Ile95 O	3.4
				Met147 N	2.9
				Met147 O	2.9
	NO1	Water380 O	2.7	The196 OG1	2.5
	NO2	Ile21 N	2.8		
	NO2	Water271 O	2.8	Gly14 O	2.8
				Ile21 N	3.3
				Ser94 OG	2.9
				Ala22 N	2.8
Pyrophosphate	O3	Water380 O	3.1	Thr196 OG1	2.5
	O3	Water381 O	3.1	Water400 O – Gly96 O	2.7
	AO1	Water381 O	2.6	Water400 O – Gly96 O	2.7
	AO1	Water378 O	2.6	Water337 O – Gly96 O	2.6
	AO2	Ser20 OG1	2.6		
	AO2	Water285	2.7	Thr17 OG1	2.9
				Ser20 OG	3.4
	AO5*	Water271	2.7	Gly14 O	2.8
				Ile21 N	3.3
				Ala22 N	2.8
				Ser94 OG	2.9
Adenine Ribose	AO4*	Gly96 N	3.4		
	AO3*	Gly14 O	2.7		
	AO3*	Water272 O	3.0		
	AO2*	Water272 O	2.8	Ile15 O	2.7
				Phe41 N	3.0
Adenine	AN3	Water282 O	3.0	Gly14 N	2.9
	AN6	Asp64 OD1	2.9		
	AN6	Water349	2.8	Asp64 OD1	3.3
				Gln66 OE1	2.7

Table 5 (continuation)

NADH moiety	NADH atom	I21V InhA atom	Distance (Å)	Atom from protein bonded to water molecule	Distance (Å)
Nicotinamide Ribose	NO7	Ile194 N	2.6		
	NO7	Water 352 O	2.4	Tyr158 OH	2.7
	NN7	Ile194 O	3.2		
	NN7	Water380	3.4	Thr196 OG1	3.4
	NO2*	Lys165 NZ	3.0		
	NO3*	Lys165 NZ	3.0		
	NO3*	Ile95 O	3.1		
	NO3*	Water287 O	3.2	Lys165 NZ	2.7
				Ile95 O	2.6
				Ser123 O	2.7
	NO3*	Water288 O	2.8	His93 ND1	2.9
				His93 O	3.2
				Ile95 N	3.4
				Ile95 O	3.3
				Met147 N	2.9
				Met147 O	2.9
	NO1	Water380 O	2.7	The196 OG1	3.4
	NO2	Ile21 N	2.8		
	NO2	Water271 O	2.7	Gly14 O	2.8
				Ile21 N	3.3
				Ser94 OG	2.8
				Ala22 N	2.9
Pyrophosphate	O3	Water380 O	3.1	Thr196 OG1	3.4
	O3	Water381 O	3.1	Water400 O – Gly96 O	2.4
	AO1	Water381 O	2.6	Water400 O – Gly96 O	2.4
	AO1	Water378 O	2.5	Water337 O – Gly96 O	3.0
	AO2	Ser20 OG1	2.6		
	AO2	Water285	2.7	Thr17 OG1	3.0
				Ser20 OG	3.6
	AO5*	Water271	3.4	Gly14 O	2.8
				Ile21 N	3.3
				Ala22 N	2.9
				Ser94 OG	2.8
Adenine Ribose	AO4*	Gly96 N	3.5		
	AO3*	Gly14 O	2.7		
	AO3*	Water272 O	3.1		
	AO2*	Water272 O	3.0	Ile15 O	2.6
				Phe41 N	2.9
Adenine	AN3	Water282 O	3.0	Gly14 N	3.00
	AN6	Asp64 OD1	3.0		
	AN6	Water349	2.8	Asp64 OD1	3.5
				Gln66 OE1	2.7

Table 5 (Continuation)

NADH moiety	NADH atom	I47T InhA atom	Distance (Å)	Atom from protein bonded to water molecule	Distance (Å)
Nicotinamide Ribose	NO7	Ile194 N	2.6		
	NO7	Water 352 O	2.5	Tyr158 OH	2.6
	NN7	Ile194 O	3.1		
	NN7	Water380	3.3	Thr196 OG1	3.4
	NO2*	Lys165 NZ	3.0		
	NO3*	Lys165 NZ	3.0		
	NO3*	Ile95 O	3.1		
	NO3*	Water287 O	3.3	Lys165 NZ	2.7
				Ile95 O	2.8
				Ser123 O	2.6
	NO3*	Water288 O	2.9	His93 NE2	3.4
				His93 O	3.2
				Ile95 N	3.4
				Ile95 O	3.3
				Met147 N	2.9
				Met147 O	3.0
	NO1	Water380 O	2.6	The196 OG1	3.4
	NO2	Ile21 N	2.8		
	NO2	Water271 O	2.7	Gly14 O	2.9
				Ile21 N	3.3
				Ser94 OG	2.9
				Ala22 N	2.9
Pyrophosphate	O3	Water380 O	3.1	Thr196 OG1	3.4
	O3	Water381 O	3.1	Water400 O – Gly96 O	3.1
	AO1	Water381 O	2.6	Water400 O – Gly96 O	3.1
	AO1	Water378 O	2.6	Water337 O – Gly96 O	2.8
	AO2	Ser20 OG1	2.6		
	AO2	Water285	2.7	Thr17 OG1	2.9
				Ser20 OG	3.5
	AO5*	Water271	3.3	Gly14 O	2.9
				Ile21 N	3.3
				Ala22 N	2.9
				Ser94 OG	2.9
Adenine Ribose	AO4*	Gly96 N	3.4		
	AO3*	Gly14 O	2.8		
	AO3*	Water272 O	3.2		
	AO2*	Water272 O	3.2	Ile15 O	3.0
				Phe41 N	3.0
Adenine	AN3	Water282 O	3.3	Gly14 N	2.8
	AN6	Asp64 OD1	3.0		
	AN6	Water349	2.8	Asp64 OD1	3.3
				Gln66 OE1	2.6

Table 5 (Continuation)

NADH moiety	NADH atom	S94A InhA atom	Distance (Å)	Atom from protein bonded to water molecule	Distance (Å)
Nicotinamide Ribose	NO7	Ile194 N	2.5		
	NO7	Water 352 O	2.5	Tyr158 OH	2.8
	NN7	Ile194 O	3.2		
	NN7	Water380	3.1	Thr196 OG1	3.3
	NO2*	Lys165 N Z	2.9		
	NO3*	Lys165 NZ	3.1		
	NO3*	Ile95 O	3.2		
	NO3*	Water287 O	3.4	Lys165 NZ	2.8
				Ile95 O	2.7
				Ser123 O	2.6
	NO3*	Water288 O	2.8	His93 ND1	2.8
				His93 O	3.3
				Ile95 N	3.3
				Ile95 O	3.2
				Met147 N	3.0
				Met147 O	3.0
	NO1	Water380 O	2.5	The196 OG1	3.3
	NO2	Ile21 N	2.7		
	NO2	Water271 O	2.7	Gly14 O	2.9
				Ile21 N	3.0
				Ala94	-
				Ala22 N	2.7
Pyrophosphate	O3	Water380 O	3.1	Thr196 OG1	3.3
	O3	Water381 O	2.8	Water400 O – Gly96 O	3.3
	AO1	Water381 O	2.5	Water400 O – Gly96 O	3.3
	AO1	Water378 O	2.7	Water337 O – Gly96 O	2.6
	AO2	Ser20 OG1	2.6		
	AO2	Water285	2.7	Thr17 OG1	3.2
				Ser20 OG	3.4
	AO5*	Water271	3.4	Gly14 O	2.9
				Ile21 N	3.0
				Ala22 N	2.7
				Ala94	-
Adenine Ribose	AO4*	Gly96 N	3.4		
	AO3*	Gly14 O	2.5		
	AO3*	Water272 O	3.1		
	AO2*	Water272 O	2.6	Ile15 O	2.8
				Phe41 N	2.9
Adenine	AN3	Water282 O	3.0	Gly14 N	3.0
	AN6	Asp64 OD2	3.0		
	AN6	Water349	2.8	Asp64 OD1	3.3
				Gln66 OE1	2.8

Table 6. Summary of direct and water-mediated hydrogen-bonding occurring in intersubunit interfaces about the three orthogonal molecular twofold axes (A, B, and C) of WT InhA homotetramer.

Twofold Axis	Atom from subunit 1	Atom from symmetry-related subunit	Distance (Å)	Atom from symmetry-related subunit bonded to water molecule	Distance (Å)
A	Arg173 O	Gln267 NE2	3.0		
A	Arg177 NH1	Gln267 OE1	2.8		
A	Arg177 O	Wat386[1] O	2.6	Pro227 O	2.9
A	Pro227 O	Wat386[2] O	2.9	Arg177 O	2.6
A	Lys240 NZ	Water325[1] O	2.5	Asp248 O	2.9
A	Thr241 OG1	Water385[2] O	2.6	Pro251 N	3.4
A				Ala252 N	2.9
A				Thr253 N	3.2
A				Thr253 OG1	2.7
A	Asp248 O	Water325[2] O	2.9	Lys240 NZ	2.5
A	Ala252 O	Asp261 N	3.1		
A	Ala252 O	Gly262 N	2.7		
A	Pro251 N	Water385[1] O	3.4		
A	Ala252 N	Water385[1] O	2.9		
A	Thr253 N	Water385[1] O	3.2		
A	Thr253 OG1	Water385[1] O	2.7	Thr241 OG1	2.6
A	Thr254 N	Water280[2] O	3.2		
A	Thr254 O	Water280[2] O	3.2		
A	Asp256 OD2	Water280[2] O	2.7	Tyr259 O	3.3
A				Asp261 OD1	3.0
A				Gly263 N	3.5
A	Asp256 OD1	His265 NE2	3.5		
A	Asp256 OD2	Water317 O	2.9	Tyr259 N	2.9
A	Tyr259 N	Water317[1] O	2.9	Asp256 OD1	2.9
A	Tyr259 O	Water280[1] O	3.3		
A	Asp261 OD1	Water280[1] O	3.0		
A	Gly263 N	Water280[1] O	3.5	Thr254 N	3.2
A				Thr254 O	3.2
A				Asp256 OD2	2.7
A	Ile257 O	Wat402 O	2.8	Ile257 O	2.8
A	Asp261 N	Ala52 O	3.1		
A	Gly262 N	Ala252 O	2.7		
A	His265 NE2	Asp256 OD1	3.5		
A	Gln267 NE2	Arg173 O	3.0		
A	Gln267 OE1	Arg177 NH1	2.8		

[1] Water molecule from subunit1, [2] Water molecule from symmetry-related subunit 2, [3] Water molecule from symmetry-related subunit 3.

Table 6 (continuation)

Twofold Axis	Atom from subunit 1	Atom from symmetry-related subunit	Distance (Å)	Atom from symmetry-related subunit bonded to water molecule	Distance (Å)
B	Glu68 OE2	Water356[3] O	2.9	Tyr113 N	2.8
B	Phe108 N	Water303[1] O	3.0	Glu178 OE2	2.7
B	Phe108 N	Water304[1] O	3.0		
B	Phe109 N	Water304[1] O	2.9	Glu178 OE1	3.3
B	Phe109 O	Lys132 NZ	3.4		
B	Asp110 OD1	Lys132 NZ	3.5		
B	Ala111 O	Tyr125 OH	2.5		
B	Tyr113 N	Tyr125 OH	3.4		
B	Tyr113 N	Water356[1] O	2.8	Glu68 OE2	2.9
B				Tyr125 OH	3.4
B	Tyr113 OH	Water294 O	2.7	Tyr113 OH	2.5
B	Tyr113 OH	Ser117 OG	3.1		
B	Tyr113 OH	Ser117 O	2.6		
B	Tyr113 OH	His121 N	3.3		
B	Ser117 OG	Tyr113 OH	3.1		
B	Ser117 OG	Ser117 OG	2.7		
B	Ser117 O	Tyr113 OH	2.6		
B	His121 N	Tyr113 OH	3.3		
B	Tyr125 OH	Ala111 O	2.5		
B	Tyr125 OH	Tyr113 N	3.4		
B	Tyr125 OH	Water356[3] O	3.4	Tyr113 N	2.8
B	Lys132 NZ	Phe109 O	3.4		
B	Lys132 NZ	Asp110 OD1	3.5		
B	Pro151 O	Arg173 NH1	3.1		
B	Pro151 O	Water335[3] O	2.8		
B	Ala154 N	Water335[3] O	3.4	Arg173 NH1	3.5
B				Arg173 NE	2.8
B				Ser170 OG	3.0
B	Ser166 OG	Ser170 OG	2.7		
B	Ser166 OG	Water283[1] O	2.6	Arg173 NH1	2.9
B	Ser170 OG	Ser166 OG	2.7		
B	Ser170 OG	Water335[1] O	3.0	Pro151 O	2.84
B				Ala154 N	3.4
B	Arg173 NH1	Pro151 O	3.1		
B	Arg173 NH1	Water283[3] O	2.9	Ser166 OG	2.6
B	Arg173 NH1	Water335[1] O	3.5		
B	Arg173 NE	Water335[1] O	2.8	Pro151 O	2.8
B				Ala154 N	3.4
B	Glu178 OE2	Water303[3] O	2.7	Phe108 N	3.0
B	Glu178 OE1	Water304[3] O	3.3	Phe108 N	3.0
B				Phe109 N	2.9

[1] Water molecule from subunit 1, [3] Water molecule from symmetry-related subunit 3.

Table 6 (continuation)

Twofold Axis	Atom from subunit 1	Atom from symmetry-related subunit	Distance (Å)	Atom from symmetry-related subunit bonded to water molecule	Distance (Å)
C	Asp150 OD2	Water308[1] O	2.6	Arg153 NH2	3.0
C	Ser152 OG	Water308[1] O	3.4		
C	Arg153 NH2	Water308[4] O	3.0	Asp150 OD2	2.6
C				Ser152 OG	3.4
C	Arg153 NE	His265 O	2.7		
C	Arg153 NE	Water 362[4] O	3.5	Ala264 O	3.4
C				Arg225 NE	3.1
C				Gln267 O	2.7
C	Ala154 N	Thr266 O	2.96		
C	Ala154 O	Leu268 N	2.7		
C	Arg225 NE	Water362[1] O	3.1	Arg153 NE	3.5
C	Ala264 O	Water362[1] O	3.4		
C	His265 O	Arg153 NE	2. 7		
C	Thr266 O	Ala154 N	2.9		
C	Gln 267 O	Wat362[1] O	2.7	Arg153 NE	3.5
C	Leu268 N	Ala154 O	2.7		

[1] Water molecule from subunit 1, [4] Water molecule from symmetry-related subunit 4.

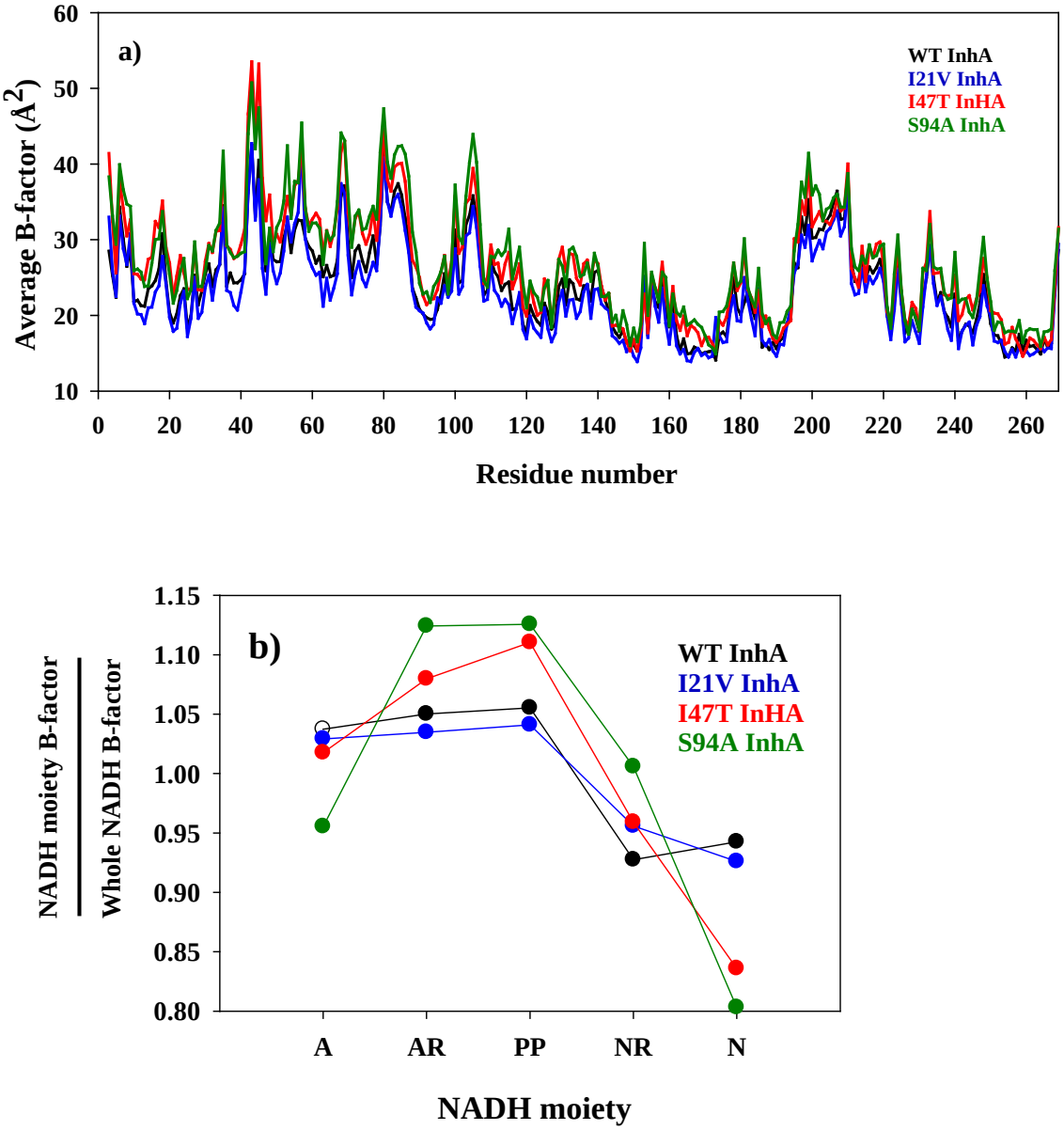


Fig.11**. (a)** B-factor plot per protein residues. The pattern of distribution for B-factor values per protein residues is very similar in the four crystal structures of InhA. The more flexible regions are the A loop (residues 100 -112) that connects strand β4 with helix α4, the substrate binding loop (residues 196-219) that includes α6 and α7 helices, the stretch of residues 79-87 that

corresponds to the C-terminal part of α3 helix and the the loop between α3 and β4, the stretch of residues 42-46 located in the loop connecting strand β2 to helix α2 and in the N-terminal part of helix α2, the residues 56-58 from the loop between helix α2 and strand β3, and the stretch of residues 67-70 that corresponds to N-terminal part of helix α3.

(b) Plot of the ratio between mean B-factor value for each NADH moiety and mean B-factor value for the whole NADH molecule per NADH moiety for WT-NADH, I21V-NADH, I47T-NADH, and S94A-NADH structures. The mean B-factor values for bound NADH molecule are very similar in all structures of the binary complexes. Furthermore, the plot above showed a pattern very similar for the four InhA structures, in which the pyrophosphate moiety appears to have the highest mean B-factor value, suggesting this moiety as the most flexible region of the NADH molecule.

Abbreviations: N, nicotinamide, NR, nicotinamide ribose, PP, pyrophosphate, AR, adenine ribose, and A, adenine.

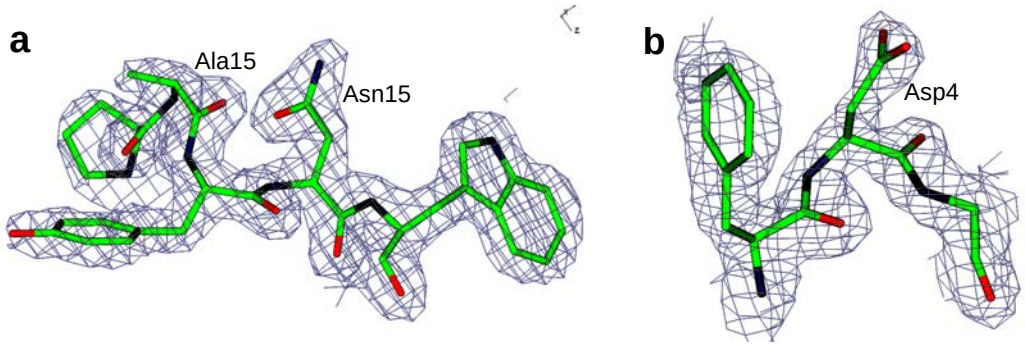


Fig. 12. Amino acid residues that are found in the disallowed region of the Ramachandram Plot from the program PROCHECK[53]. The Ala157, Asn169 (**a**), and Asp 42 (**b)** residues are well fitted to a ($2F_{obs} - F_{calc}$), α_{calc} eletron density map encompassing Ala157, Asn169 (**a**), and Asp 42 (**b)** residues from the crystal structure of WT InhA-NADH binary complex is shown. The electron density map was calculated in the REFMAC 5.2 program[48] using all unique reflections and is countoured at a level of 1.3σ above the mean. The figure was prepared with CCP4 Molecular Graphics 0.12 program.[62] The residues Ala157 and Asn159 are located in a turn connecting strand β5 to helix α5, and Asp42 is located in a turn connecting strand β2 to helix α2.

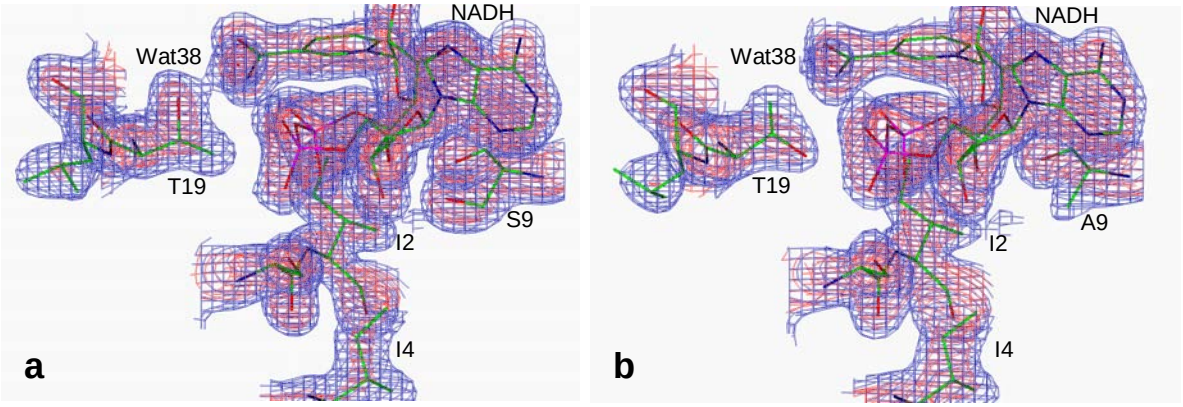


Fig. 13. ($2F_{obs} - F_{calc}$) Simulating annealing omit electron density maps showing the different conformations of Thr196 side-chain in the crystal structures of WT InhA-NADH (a) and S94A InhA-NADH (b) binary complexes. The simulating annealing omit map was calculated in the CNS program[56] and the omit map was analysed in the XTALVIEW program[52]. The residues situated within 3.5Å from Thr196 residue were omitted from simulating annealing refinement and structure factor calculations, and the residues within 2 Å from the omitted region were harmonically restrained to their former positions throughout the refinement. The slow-cooling protocol was performed with a starting temperature of 1000K using all refletions. The maps are countoured at a level of 1.0σ (blue) and 2.0σ (red) above the mean. For clarity only S20, I21, I47, S94(a) or Ala94(b), Leu197, and Thr196 residues,Water380, and NADH molecule are shown. The electron density maps give evidence that in the structure of S94A InhA the Thr196 side-chain has rotated approximately 180 in relation to its conformation in the structure of WT InhA, resulting in hydrogen bond distances between the OG1 atom of Thr196 and the bridging water (Water380) that are significantly larger (3.3 – 3.4 Å) than the 2.5 Å distance observed in WT enzyme.

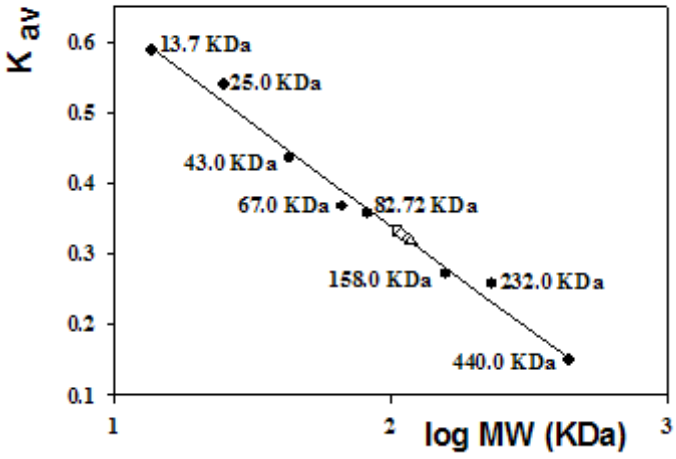


Fig 14. Determination of molecular weight of wild-type and isoniazid-resistant enoyl-ACP reductase enzymes in solution. The sample (100 μL) was loaded on an FPLC Superdex 200 HR 10/30 column, and run at a flow rate of 0.4 mL min^{-1}. K_{av} values were calculated as described in Material and Methods. Protein molecular weight markers (●) were Ribonuclease A (13,700 Da), Chymotrypsinogen (25,000 Da), Ovalbumin (43,000 Da), Albumin (67,000 Da), mtPNP (82,720 Da), Aldolase (158, 000 Da), Catalase (232,000 Da), and Ferritin (440,000 Da). Molecular mass values of 115,930 Da, 104,890 Da, 118,080 Da, and 108, 720 Da were estimated from the plot of K_{av} against the logarithm of standard molecular weights for WT InHA (○), I21V InhA (□), I47T InhA (Δ) , and S94A InhA (◊) enzymes, suggesting that the four proteins are homotetramers in solution.

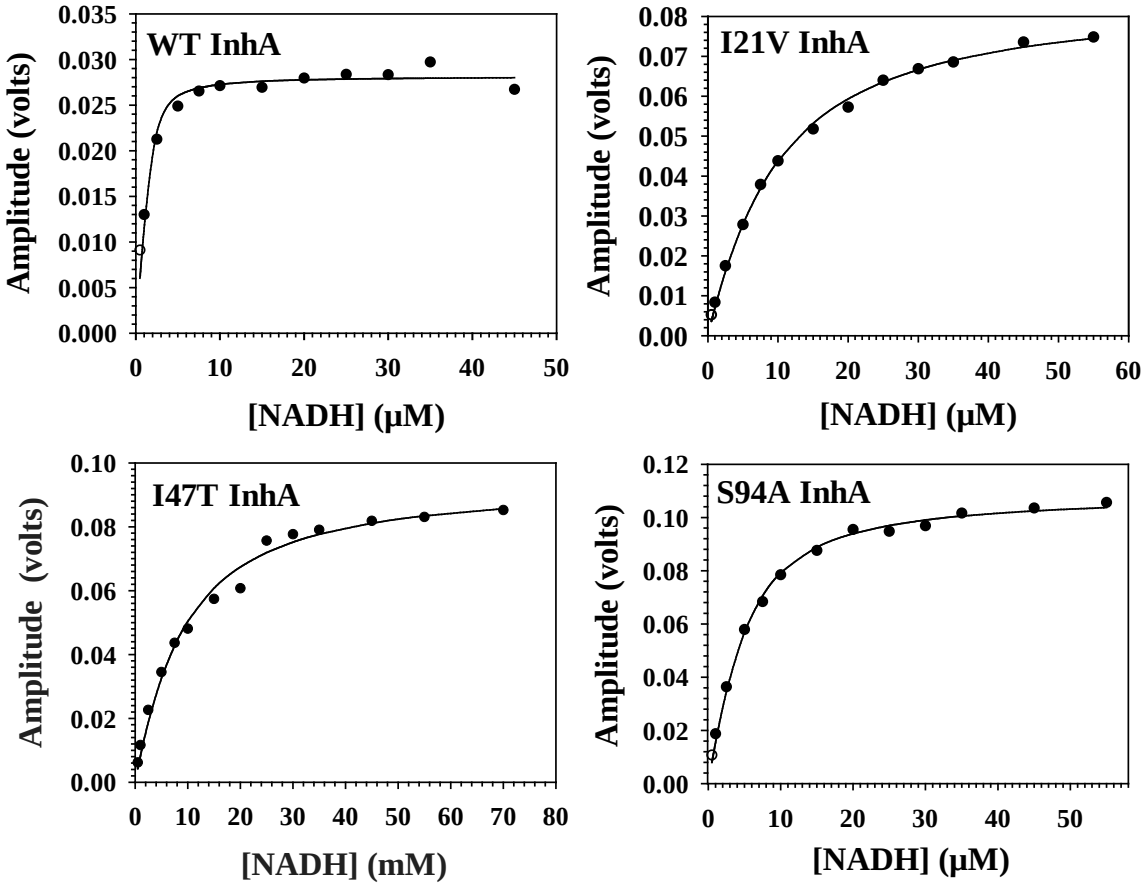


Fig. 15. Analysis of the NADH concentration dependence of the fluorescence amplitude values obtained in the fast phase of NADH binding to WT, I21V, I47T, and S94A InhA enzymes. Values of K_{d2} and K_{d2}^{*} (S94A InhA) for NADH binding were determined by fitting the four data sets to a quadratic equation (eq. **1**), which describes a second-order binding process.

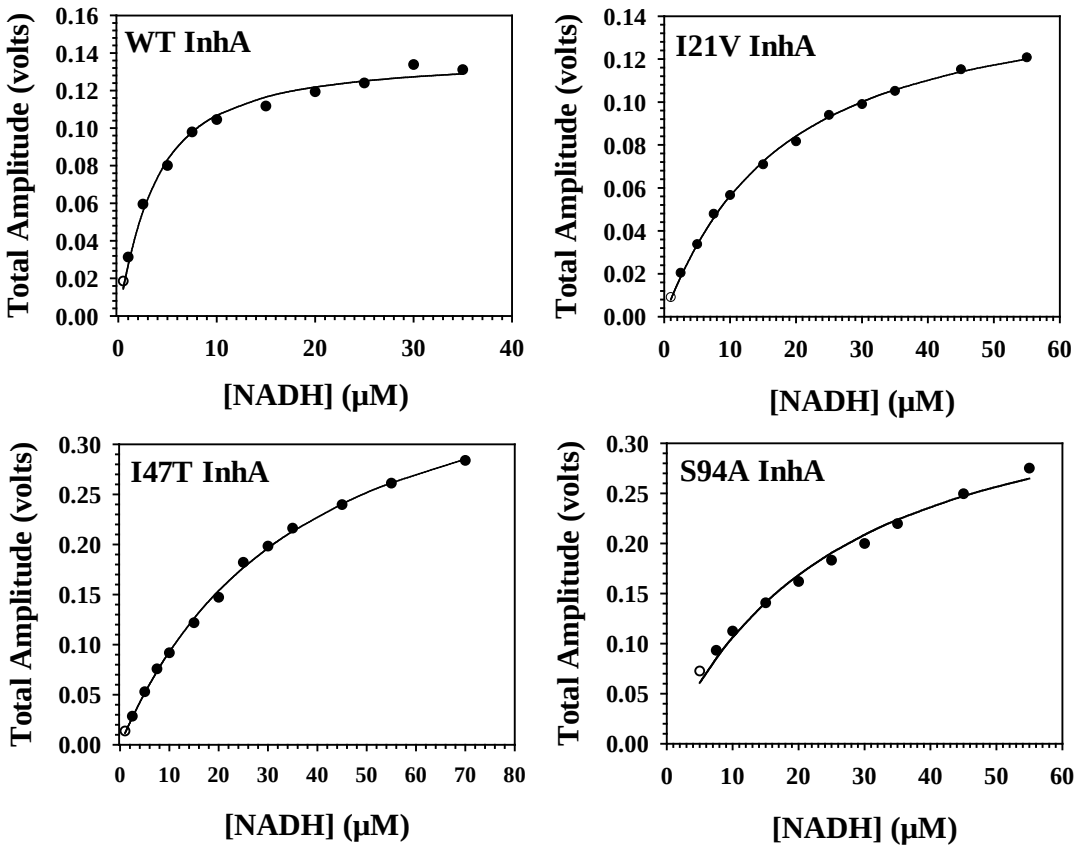


Fig. 16**.** NADH concentration dependence of the algebraic sum of the signal amplitudes of the fast and slow phases (Total amplitude) of NADH binding to WT, I21V, I47T, and S94A InhA enzymes. Estimates for the dissociation constant of the overall reaction were obtained by fitting the four data sets to the quadratic equation **1.**

6. CONCLUSÃO FINAL

A Tuberculose (TB) é um problema de saúde mundial prioritário, que continua sendo a principal causa de mortalidade devido à um patógeno bacteriano, o *Mycobacterium tuberculosis*. A estimativa é de que no ano 2000, 3,2 % dos novos casos que ocorreram no mundo foram de Tuberculose multiresistente (MDR-TB), definida como Tuberculose resistente a pelo menos isoniazida e rifampicina. De acordo com o relatório de Controle Global da TB, publicado pela Organização Mundial da Saúde em 2004 (WHO), atualmente ocorrem 300 mil novos casos por ano de MDR-TB no mundo, e 79 % destes são causados por "super cepas", que são resistentes à pelo menos três das quatro principais drogas utlizadas para tratar TB. A cepas de MDR-TB são transmitidas pelo ar tao fácil quanto as cepas de TB sensiveis às drogas. Os casos de MDR-TB são de difícil tratamento, sendo recomendado a introdução de drogas de segunda linha para romper a disseminação da resistência. Entretanto, estas são mais tóxicas e menos eficientes. Além disso, o tratamento é pelo menos três vezes mais longo e 100 vezes mais caro que os regimes de curta-duração utilizados para o tramento de TB sensível. Assim, é imprescindivel que novos agentes quimioterápicos anti-TB, mais efetivos e de baixa toxicidade sejam desenvolvidos o quanto antes para encurtar o tratamento da MDR-TB e evitar a disseminação das cepas multiresistentes.

No Capitulo 3 da presente tese, um novo composto anti-*Mycobacterium tuberculosis* foi apresentado, o pentaciano(isoniazida)ferrato II ($[Fe^{II}(CN)_5(INH)]^{3-}$). Este composto é um complexo inorgânico análogo à isoniazida, que contém um parte pentacianoferrato ligada ao átomo de nitrogênio do anel piridínico da isoniazida através do átomo de ferro. O composto inibe a enzima 2-*trans*-enoil-ACP(CoA) redutase (InhA) espécie selvagem com a mesma eficiência que inibe a enzima mutante e resistente à isoniazida I21V InhA, cuja mutação foi previamente identificada em isolados clínicos resistentes à isoniazida. Além disso, ao contrário da isoniazida, o pentaciano(isoniazida)ferrato II não necessita ativação pela catalase/peroxidase KatG, nem a

presença de NADH para inibir a InhA. Outros dados que demonstram a importância deste novo composto são os valores de concentracão inibitória mínima (MIC) obtidos frente a uma cepa virulenta de *Mycobcterium tuberculosis*. O novo composto apresentou um MIC de 0.2 μg mL^{-1}, enquanto que o *M. tuberculosis* é suscetivel à isonizada na faixa de 0.02 – 0.2 μg mL^{-1}. Esses dados apontam o composto $[Fe^{II}(CN)_5(INH)]^{3-}$ como um candidato promissor à composto líder, no qual basear o desenvolvimento de novos agents anti-tuberculose. No entanto, futuros estudos estruturais do complexo binário InhA-$[Fe^{II}(CN)_5(INH)]^{3-}$ deverão ser realizados para demonstrar as interações químicas específicas responsáveis pela inibição da InhA. Atualmente, esforços em nosso laboratório para a obtenção de cristais do complexo binário formado entre o pentaciano(isoniazida)ferrato II e a InhA espécie selvagem e também a I21V InhA resistente à isoniazida estão em andamento. Uma estrutura tridimensional destes complexos binários será crucial para auxiliar no desenho de uma nova classe de agentes antimicobacterianos que possam ser usados no tratamento da Tuberculose causada tanto por cepas sensíveis quanto resistentes à isoniazida.

Dando continuidade aos nossos estudos de caracterização do novo inibidor e também do processo de inibição do principal alvo da isoniazida em *Mycobacterium tuberculosis*, a enzima InhA, no Capítulo 4 foram descritos os fundamentos que levaram ao desenho do novo inibidor e a caracterização química deste. Além disso, foram apresentados dados cinéticos detalhados da inibição da InhA pelo pentaciano(isoniazida)ferrato II e dados de modelagem molecular e docking para avaliar a energia livre de ligação para a formação do complexo enzima-inibidor e predizer o modo de ligação entre as moléculas. A análise cinética permitiu a elucidação do mecanismo de inibição como sendo do tipo de ligação lenta (“slow binding”). No processo de iniibição, inicialmente ocorre uma ligação rápida do inibidor à InhA com afinidade moderada (K_i = 32 ± 3 μM) seguido por uma lento passo de isomerização do complexo inicial enzima-inibidor (complexo EI), resultando em um complexo final de mais alta afinidade (complexo EI*), onde o inibidor encontra-se mais fortemente ligado à enzima (K_i^* = 70 nM). O processo de dissociação do complexo EI* também é bastante lento, com uma constante de velocidade de dissociação de 1.1 (± 0.1) x 10^{-3} min^{-1} e um tempo de meia vida de 630 ± 28 min. Os dados também demosntraram que a dissociação do inibidor é limitada pelo passo de isomerização reversa do complexo EI* para o complexo EI. Sendo assim, a constante de velocidade de dissociação que foi estimada representa a veradeira constante de velocidade de isomerização reversa do

complexo EI* para o complexo EI (ver Figura 5 do Capítulo 4, mecanismo B). Estes dados indicam uma longa vida para o complexo EI* e, quanto maior a estabilidade do complexo EI* maior a probabilidade de que o inibidor apresente um longo tempo de atividade *in vivo*. Uma das fraquezas do uso de inibidores clássicos como drogas em ensaios clínicos é que a inibição resulta na acumulação do substrate da enzima, que eventualmente pode superar a inibição. No presente mecanismo de inibição, o aumento da concentração do substrato (NADH ou dodecenoil-ACP) nao terá qualquer efeito sobre a isomerização reversa do complexo EI* para o complexo EI e assim, não afetará a inibição.

Na ausência de uma estrutura cristalina do complexo Inha-($[Fe^{II}(CN)_5(INH)]^{3-}$, os resultados de modelagem molecular e docking foram muito importantes para predizer a posição da ligação do inibidor no sitio ativo da InhA, e para um entendimento inicial das bases energéticas e estruturais das interacções entre proteína e inibidor. Embora a posição exata de ligação do $[Fe^{II}(CN)_5(INH)]^{3-}$ no bolsãode ligação a NAD(H) da InhA não pode ser determinada sem ambiguidade, o inibidor $[Fe^{II}(CN)_5(INH)]^{3-}$ preferencialmente ocupa os sítios de ligação às partes pirofosfato e nicotinamida do bolsãode ligação do NADH na enzima InhA. Além disso, a posição de ligação observada no docking do inibidor $[Fe^{II}(CN)_5(INH)]^{3-}$ na InhA sobrepõe-se à posição ocupada pela ligação dupla trans do substrato acil C16 na estrutura cristalina da enoil redutase de *Brassica napus* (1BVR) (1). Estes resultados são consistentes com os valores menores das constantes de velocidade de primeira ordem aparentes, que foram estimados experimentalmente na presença de NADH ou dodecenoil-CoA.

A determinação do genoma completo do *M. tuberculosiss* (2) teve um grande impacto no avanço das pesquisas no campo da TB. Um dos benefícios esperados das análises genômicas de bactérias patogênicas é na área da saúde, principalmente no desenvolvimento de métodos de diagnóstico mais rápidos, novas vacinas e novos agentes antimicrobianos. As novas abordagens que estão sendo exploradas incluem a comparação de genomas usando bioinformática, análise funcional de genes, proteoma, transcriptoma, and genoma estrutural. Espera-se que um melhor entendimento da biologia complexa do *M. tuberculosis* revele novos alvos para o processo de desenvolvimento de drogas. Além disso, os estudos estruturais sendo conduzidos pelo “TB Structural Genomics Consortium” disponibilizarão estruturas de muitos alvos protéicos potenciais e devem prover informações importantes, tais como, a função biológica de proteínas que não apresentam homologia à nenhuma proteína conhecida (3). Estas proteínas deve m ser

selecionadas para análises posteriores e para técnicas de química computacional para identificação de compostos líderes.

Como enfatisado anteriormente na presente tese, é urgente a necessidade de agents anti-TB mais efetivos e menos tóxicos. Só assim, será possivel encurtar o tratamento atual para TB, melhorar o tratamento da MDR-TB, e obter um tratamento efetivo para a Tunerdulose latente. Assim, nos parece razoável examinar a potencial atividade anti-TB de análogos à isoniazida (INH) devido a sua disponibilidade oral e ao favorável perfil de toxicidade da INH. Além disso, a InhA é um alvo validado e inibidores da sua atividade podem representar um atalho no desenvolvimento de agents anti-TB. Enquanto que os alvos novos primeiro necessitam ser identificados e validados. Existe um necessidade cíclica de inovação para prover novos arcabouços estruturais (compostos líderes), que devem passar por uma iteração química com o objetivo de otimização antes de atingir o status de candidato à droga. No Capítulo 4 da presente tese n nós descrevemos uma nova abordagem para o desenho racional de um análogo à INH, que é baseada em umátomo inorgânico atachado ao átomo de nitrogênio do do anel heterocíclico da INH. Os resultados apresentados nos Capítulos 3 e 4, levaram-nos a conclusão de que o pentaciano(isoniazida)ferrato II é um candidato promissor para o futuros desenvolvimento de uma nova droga anti-TB e pode representar uma nova classe de compostos líderes. A avaliação da capacidade deste complexo inorgânico de matar o *M. tuberculosis* dentro de macrófagos de medula óssea de camundongo, e também de inibir o crescimento micobacteriano em modelos animais, são etapas essencias, que estão atualmente em desenvolvimento para estabelecer o pentaciano(isoniazida)ferrato II como um potencial candidato à droga.

Compostos quimioterapêuticos baseados em metais tem sido avaliados para possível aplicação medicinal, po exemplo, a cisplatina, que é baseada em platina, é amplamente usada no tratamento do cancer (4). Complexos metálicos também tem sido considerados como agentes quimioterapêuticos alternativos para tratar doenças tropicais como Malária, Tripanossomíase e Leishmaniose (5). O exemplo de maior sucesso de antimláricos baseados em metais são compostos derivados da droga Cloroquina (CQ) por meio da coordenação à fragmentos metálicos. Os testes *in vitro* e *in vivo* de complexos de CQ com metais de transição (Rh, Ru) contra *Plasmodium berghei* mostraram que a incorporação de fragmentos metálicos produz geralmente uma eficácia aumentada da CQ (6). Mais recentemente, complexos de ouro-chloroquina foram ativos contra cepas de *Plasmodium falciparum* resistentes à CQ e também

contra cepas sensíveis à CQ (7, 8). A atividade mais alta para esta série de compostos foi obtida com [(CQ)Au(PPh_3)][PF_6], que foi 9 vezes mais ativo que a Chloroquina difosfato contra a cepa de *P. falciparum* resistente à CQ, FcB1. Além disso, o composto foi também 5 vezes mais ativo contra a cepa mais agressiva FcB2, e 22 vezes mais ativo contra o parasita da Malária em roedores, *P. Berghei* (5). Outro fato interessante é que complexos metálicos de análgos a carboximidrazona apresentaram um atividade anti-TB aumentada *in vitro* contra a cepa H37Rv de *M. tuberculosis* (9). Entretanto, o(s) alvo(s) destes complexos metálicos são desconhecidos, e portanto, é dificil considerar esta abordagem como um desenho racional de drogas.

Como salientado por Gerard Jaouen (professor dequímica da “Ecole Nationale Superieure de Chimie de Paris”), “a área de química bio-inorgânica necessita provar que pode criar algo útil para a sociedade” (10). Como o campo da química medicinal inorgânica está em crescimento, não resta dúvida de que outras drogas contendo metais serão desenvolvidas em várias áreas terapêuticas, incluindo o tratamento de doençcas parasitárias. Eu espero que a abordagem descrita aqui gere frutos e auxilie no futuro desenvolvimento de agentes anti-TB.

No capítulo 5 foram apresentados os resultados de nossos esforços em compreender as bases estruturais dos mecanismos de resistência à isoniazida da enzima InhA. Com este objetivo, nós resolvemos as estruturas cristalinas da InhA espécie selvagem e dos mutantes resistente à isoniazida S94A, I21V e I47T InhA complexados com NADH.

A análise do bolsãode ligação a NADH mostrou que na estrutura do complexo S94A InhA-NADH, ao contrário do publicado previamente (11), não ocorre a rotação do grupo carbonila da Gli14, e somente a ponte de hidrogênio conservada entre molécula de água WAT1 e o grupo hidroxila da Ser94 foi perdida devido a mutação do resíduo Ser94 para Ala94. Assim, o rompimento desta ponte de hidrogênio é provavelmente a principal causa desta enzima mutante apresentar um redução de 6 vezes na sua afinidade por NADH, que aproximadamente corresponde a 2.5 kcal mol^{-1}(12). Este valor está em acordo com estimativas de energia de pontes de hidrogênio, que se encontram na faixa de 3 a 9 kcal mol^{-1}(13). Além disso, a conservada molécula de água WAT1 foi proposta como uma característica integral de domínios de ligação a dinucleotídeos do tipo Rossmann (“Rossmannn fold”), que contribui significativamente no reconhecimento do dinucleotídeo, proporcionando uma contribuição entálpica favorável à energia livre de ligação (14).

A enzima mutante I21V InhA possui uma mutação na alça de ligação a difosfato rica em glicina (Gli14 – Ala22 na InhA). Na enzima espécie selvagem, o átomo CD1 da cadeia lateral apolar do resíduo Ile21 realiza contatos de van der Waals com o anel nicotinamida, com a ribose da nicotinamida e com os oxigênios do grupo fofato. Já na enzima mutante e resistente à isoniazida I21V InhA, estes contatos de van der Waals foram perdidos porque a cadeia lateral do resíduo Val21 não possui o átomo CD1.

Por último, na enzima mutante I47T Inha, a cadeia lateral levemente polar do resíduo Thr47 possibilitou a entrada de uma molécula de água adicional (Water321, que é ausente nas estruturas dos complexos Inha espécie selvagem-NADH, I21V InhA-NADH e S94A InhA-NADH). Além disso, a mutação promoveu uma ponde de hidrogênio direta entre o átomo OG da Thr47 e o grupo carbonila da Ile15, resíduo este, que é parte da alça de ligação a difosfato rica em glicina (Gli14 – Ala22). Isto resultou em um enfraquecimento significativo das pontes de hidrogênio, mediadas pela água conservada WAT1, entre o grupo carbonila do resíduo Ile15 e os grupos hidroxila 2'-OH e 3'-OH da ribose da adenina.

Em comparação com a InhA espécie selvagem (sensível a isoniazida), os estudos cinéticos em estado pré-estacionário de ligação à NADH mostraram que os valores limitantes para a constante de velocidade de dissociação do NADH dos complexos binários foram 11 vezes, 5 vezes e 10 vezes mais altos para as enzimas I21V InhA, I47T InhA e S94A InhA, respectivamente. Em conclusão, os resultados descritos acima e no capítulo 5, são capazes de explicar a reduzida afinidade por NADH dos mutantes de InhA resistentes à isoniazida, e desse modo, explicam pelo menos parte dos mecanismos de resistência à isoniazida em *M. tuberculosis*. Em acordo com esta proposta, simulações de dinâmica molecular dos complexos binários formados entre NADH com a enzima Inha espécie selvagem, e com as enzimas InhA resistentes à isoniazida, que foram identificadas em isolados clínicos de *M. tuberulosis*, mostraram que as mutações na alça rica em glicina, I21V e I16T, resultaram em uma mudança no padrão de contatos por pontes de hidrogênio diretas com a parte pirofosfato do NADH (15). A parte pirofosfato apresentou mudanças de conformação consideráveis, reduzindo suas interações com o sítio ativo da InhA, o que provavelmente represente a fase inicial da expulsão do ligante, na qual o NADH se distância do sítio de ligação (50). Assim, o aumento nas constantes de dissociaçao em equilíbrio para NADH, que foram estimadas para os mutantes resistentes à isoniazida I21V InHA e I16T InhA (50), pode ser atribuído a uma dimuição do número de

interações por pontes de hidrogênio entre o NADH e os aminoácidos do sítio ativo, e também das interacões mediadas por moléculas de água. Em conclusão, a correlação entre as propriedades de ligação a NADH em solução, as estruturas cristalinas dos complexos binários com NADH, os estudos de dinâmica molecular e docking e a resistência à isoniazida de cepas de *M. tuberculosis* que possuem mutações no gene estrutural *inhA* produz uma explicação, a nivel molecular, para o mecanismo de resistência para esta droga clinicamente importante.

Eu espero que os resutlados apresentados na presente tese auxiliem no desenvolvimento de novoas agentes anti-TB, que possam superar a resistência à isoniazida.

REFERÊNCIAS

1. Rozwarski DA, Vilchèze C, Sugantino M, Bittman R, Sacchettini JC. Crystal structure of the Mycobacterium tuberculosis enoyl-ACP reductase, InhA, in complex with NAD^+ and a C16 fatty acyl substrate. J Biol Chem 1999; 274: 15582-89.

2. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, 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 MA, Rajandream M-A, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 1998; 393: 537-44.

3. Terwilliger TC, Park MS, Waldo GS, Berendzen J, Hung L-W, Kim C-Y, Smith CV, Sacchettini JC, Bellinzoni M, Bossi R, De Rossi E, Mattevi A, Milano A, Riccardi G, Rizzi M, Roberts MM, Coker AR, Fossati G, Mascagni P, Coates ARM, Wood SP, Goulding CW, Apostol MI, Anderson DH, Gill HS, Eisenberg DS, Taneja B, Mande S, Pohl E, Lamzin V, Tucker P, Wilmanns M, Colovos C, Meyer-Klaucke W, Munro AW, McLean KJ, Marshall KR, Leys D, Yang JK, Yoon H-J, Lee BI, Kwak JE, Han BW, Lee JY, Baek S-H, Suh SW, Komen

MM, Arcus VL, Baker EN, Lott JS, Jacobs WRJr, Alber T, Rupp B. The TB structural genomics consortium: a resource for *Mycobacterium tuberculosis* biology. Tuberculosis 2003; 83: 223-49.

4. Zhang CX, Lippard SJ. New metal complexes as potential therapeutics. Curr Opin Chem Biol 2003; 7:481-9.

5. Sanchez-Delgado RA, Anzellotti A. Metal complexes as chemotherapeutic agents against tropical diseases: trypanosomiasis, malaria and leishmaniasis. Mini Rev Med Chem. 2004; 4: 23-30.

6. Sanchez-Delgado RA, Navarro M, Perez H, Urbina JA. Toward a novel metal-based chemotherapy against tropical diseases. 2. Synthesis and antimalarial activity in vitro and in vivo of new ruthenium- and rhodium-chloroquine complexes. J Med Chem. 1996; 39:1095-99.

7. Navarro M, Perez H, Sanchez-Delgado RA. Toward a novel metal-based chemotherapy against tropical diseases. 3. Synthesis and antimalarial activity in vitro and in vivo of the new gold-chloroquine complex [Au(PPh3)(CQ)]PF6. J Med Chem. 1997; 40: 1937-9.

8. Navarro M, Vasquez F, Sanchez-Delgado RA, Perez H, Sinou V, Schrevel J. Toward a novel metal-based chemotherapy against tropical diseases. 7. Synthesis and in vitro antimalarial activity of new gold-chloroquine complexes. J Med Chem. 2004; 47:5204-9.

9. Sandbhor U, Padhye S, Billington D, Rathbone D, Franzblau S, Anson CE, Powell AK. Metal complexes of carboxamidrazone analogs as antitubercular agents. 1. Synthesis, X-ray crystal-structures, spectroscopic properties and antimycobacterial activity against *Mycobacterium tuberculosis* H37Rv.J Inorg Biochem. 2002; 90:127-36.

10. Dagani R. The bio side of organometallics. C&EN 2002; 80: 23-29.

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

12. Basso, L. A.; Zheng, R.; Musser, J. M.; Jacobs, W. R. Jr. & 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.

13. Fersht, A. (1999) Forces between molecules, and binding energies. In *Structure and mechanism in protein science* (Fersht, A.), pp. 324-348, W.H. Freeman and Company, New York.

14. Bottoms, C.A.; Smith, P.E. & Tanner, J.J. (2002). A structurally conserved water molecule in Rossmann dinucleotide-binding domains. Protein Sci., 11, 2125-2137.

15. Schroeder, E.K.; Basso, L.A.; Santos, D.S. and de Souza, O.N. (2005) Molecular dynamics simulation studies of the wild-type, I21V, and I16T mutants of isoniazid-resistant Mycobacterium tuberculosis enoyl reductase (InhA) in complex with NADH: toward the understanding of NADH-InhA different affinities. *Biophys. J.*, 89, 876-884.

7. ANEXOS - Outros artigos publicados durante o doutorado

- Molecular model of shikimate kinase from *Mycobacterium tuberculosis*. (2002) *Biochem. Biophys. Res. Commun.* **308**, 545-552.

- Cloning, overexpression, and purification of functional human purine nuceloside phosphorylase. (2003) *Protein Expr. Purif.* **27,** 158-164.

- Protective immune response against methicillin resistant *Staphylococcus aureus* in a murine model using a DNA vaccine approach. (2003) *Vaccine* **21,** 2661-2666.

- Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*. (2003) *Biochem. Biophys. Res. Commun.* **12,** 608-614.

- One-step purification of 5-enolpyruvylshikimate-3-phosphate synthase enzyme from *Mycobacterium tuberculosis*. (2003) *Protein Expr. Purif.* **28,** 287-292.

- Crystal structure of human purine nucleoside phosphorylase at 2.3 Å resolution. (2003) *Biochem. Biophys. Res. Commun.* **308,** 545-552.

- Molecular models for shikimate pathway enzymes of *Xylella fastidiosa*. (2004) *Biochem. Byophys. Res. Commun.* **320,** 979-991.

- Crystallographic structure of PNP from *Mycobacterium tuberculosis* at 1.9 Å resolution. *Biochem. Biophys. Res. Commun.* **324,** 789-794.

- Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2310-2319.
- Electron transfer kinetics and mechanistic study of the thionicotinamide coordinated to the pentacyanoferrate (III)/(II) complexes: a model system for the in vitro activation of thioamides anti-tuberculosis drugs. (2005) *J. Inorg. Biochem.* **99,** 368-375.

- DAHP synthase from *Mycobacterium tubercuosis* H37Rv: cloning, expression, and purification of functional enzyme. *Protein Expr. Purif.* **40,** 23-30.

ACADEMIC PRESS

Biochemical and Biophysical Research Communications 295 (2002) 142–148

BBRC

www.academicpress.com

Molecular model of shikimate kinase from *Mycobacterium tuberculosis*

Walter Filgueira de Azevedo Jr.,[a,b,*] Fernanda Canduri,[a,b] Jaim Simões de Oliveira,[c] Luiz Augusto Basso,[c] Mário Sérgio Palma,[b,d] José Henrique Pereira,[a] and Diógenes Santiago Santos[c]

[a] *Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil*
[b] *Center for Applied Toxicology, Instituto Butantan, São Paulo, SP 05503-900, Brazil*
[c] *Rede Brasileira de Pesquisa de Pesquisas em Tuberculose, Departamento de Biologia Molecular e Biotecnologia, UFRGS, Porto Alegre, RS 91501-970, Brazil*
[d] *Laboratory of Structural Biology and Zoochemistry/Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil*

Received 26 May 2002

Abstract

Tuberculosis (TB) resurged in the late 1980s and now kills approximately 3 million people a year. The reemergence of tuberculosis as a public health threat has created a need to develop new anti-mycobacterial agents. The shikimate pathway is an attractive target for herbicides and anti-microbial agents development because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals. Homologs to enzymes in the shikimate pathway have been identified in the genome sequence of *Mycobacterium tuberculosis*. Among them, the shikimate kinase I encoding gene (*aroK*) was proposed to be present by sequence homology. Accordingly, to pave the way for structural and functional efforts towards anti-mycobacterial agents development, here we describe the molecular modeling of *M. tuberculosis* shikimate kinase that should provide a structural framework on which the design of specific inhibitors may be based.

Keywords: Shikimate kinase; Bioinformatics; Structure; Drug design; *Mycobacterium tuberculosis*

The fifth annual report on global tuberculosis (TB) control of the World Health Organization found that there were an estimated 8.4 million new cases in 1999, up from 8.0 million in 1997 [1]. It is expected that there will be 10.2 million new cases in 2005 if the present trend continues. Approximately 3 million persons die from the disease each year [2]. Ninety percent of tuberculosis cases occur in developing countries, where few resources are available to ensure proper treatment and where human immunodeficiency virus (HIV) infection may be common. The concentration of deaths due to tuberculosis in demographically developing nations and mortality rate in the range from 25 to 54 years, the most economically fruitful years of life, causes substantial losses in productivity and contributes to the impoverishment of third-world countries [3]. The reemergence of TB as a public health threat, the high susceptibility of HIV-infected persons to the disease, and the proliferation of multi-drug (MDR) strains have created much scientific interest in developing new anti-mycobacterial agents to both treat *Mycobacterium tuberculosis* strains resistant to existing drugs and shorten the duration of short-course treatment to improve patient compliance [4].

The shikimate pathway is an attractive target for the development of herbicides and anti-microbial agents because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals [5]. In mycobacteria, the shikimate pathway leads to the biosynthesis of precursors for the synthesis of aromatic amino acids, naphthoquinones, menaquinones, and mycobactin [6]. Homologs to enzymes in the shikimate pathway have been identified in the complete genome sequence of *Mycobacterium tuberculosis* H37Rv strain [7]. Among them, the shikimate kinase I (mtSK, EC 2.7.1.71) encoding gene (*aroK*, Rv2539c) was proposed to be present

* Corresponding author. Fax: +55-17-221-2247.
E-mail address: walterfa@df.ibilce.unesp.br (W. Filgueira de Azevedo Jr.).

0006-291X/02/$ - see front matter
PII: S0006-291X(02)00632-0

by sequence homology. Shikimate kinase catalyzes a phosphate transfer from ATP to the carbon-3 hydroxyl group of shikimate resulting in the formation of shikimate-3-phosphate (S3P) and ADP.

The present paper describes the molecular model of *M. tuberculosis* shikimate kinase (mtSK) and analysis of SK and shikimate complex obtained by docking simulations. The homology modeling was performed using three crystallographic structures of SK from *Erwinia chrysanthemi*, solved to resolution better than 2.6 Å, as templates [8]. The mtSK has been cloned, sequenced, overexpressed in soluble and functional forms [9], thus allowing enzymological studies to be performed. The results presented here should provide a three-dimensional model of mtSK to both guide enzymological studies and aid the design of specific inhibitors.

Methods

Molecular modeling. For modeling of the mtSK we used restrained-based modeling implemented in the program MODELLER [10]. This program is an automated approach to comparative modeling by satisfaction of spatial restraints [11–13]. The modeling procedure begins with an alignment of the sequence to be modeled (target) with related known three-dimensional structures (templates). This alignment is usually the input to the program. The output is a three-dimensional model for the target sequence containing all main-chain and side-chain non-hydrogen atoms.

The degree of primary sequence identity between mtSK and *Erwinia chrysanthemi* shikiamte SK (ecSK) indicates that the crystallographic structures of ecSK are good models to be used as templates for mtSK. The atomic coordinates of three crystallographic ecSK structures (PDB access code: 1SHK, 2SHK, and 1E6C) [8,14], with two independent structures in each asymmetric unit, solved to resolution better than 2.6 Å were used to build up an ensemble of SK structures to be used as starting models for modeling of the mtSK. The atomic coordinates of all waters and ligands were removed from the ecSK structures. Next, the spatial restraints and CHARMM energy terms enforcing proper stereochemistry [15] were combined into an objective function. Finally, the model is obtained by optimizing the objective function in Cartesian space. The optimization is carried out by the use of the variable target function method [16] employing methods of conjugate gradients and molecular dynamics with simulated annealing. Several slightly different models can be calculated by varying the initial structure. A total of 500 models were generated for mtSK, and the final model was selected based on stereochemical quality.

Docking simulations. To obtain information about the docking of shikimate to ecSK and mtSK, several rigid docking simulations were performed using the geometric recognition algorithm, which was developed to identify molecular surface complementarity. The geometric recognition algorithm was implemented in the program GRAMM [17].

The atomic coordinates of shikimate, used in the docking simulations, were obtained from structure of 5-enolpyruvylshikimate-3-phosphate synthase liganded with shikimate-3-phosphate and glyphosate (PDB access code: 1G6S) [18]. To generate the ternary complex mtSK–shikimate–ADP/Mg^{2+} we superposed the atomic coordinates of the ADP/Mg^{2+} to the binary complex of mtSK–shikimate. The optimization of the complexes was carried out by the use of the variable target function method [10] employing methods of conjugate gradients and molecular dynamics with simulated annealing. All docking simulations and optimization process were performed on SGI Octane, R12000.

Analysis of the model. The overall stereochemical quality of the final model for mtSK complex was assessed by the program PROCHECK [19]. The cutoff for hydrogen bonds and salt bridges was 3.6 Å.

Results and discussion

Primary sequence comparison

The sequence alignment of ecSK (template) and mtSK (target) is shown in Fig. 1. The secondary structural elements are indicated in the figure. The sequence mtSK shows 34% of identity with the sequence of ecSK.

Quality of the model

Figs. 2A and B show the Ramachandran diagram ϕ–ψ plots for the mtSK structure and for three crystallographic SK structures solved to resolution better than 2.6 Å. The Ramachandran plot for the three ecSK structures was generated to compare the overall stereochemical quality of mtSK model against SK structures solved by biocrystallography. Analysis of the Ramachandran plot of the mtSK model shows that 91.1% of the residues lie in the most favorable regions and the

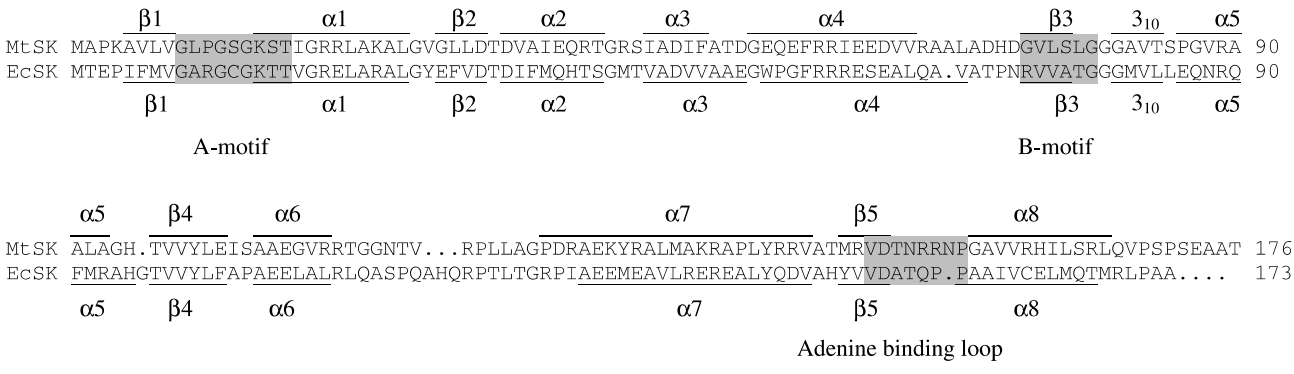


Fig. 1. The sequence alignment of ecSK and mtSK indicating the secondary structural elements. The sequence mtSK shows 34% of identity with the sequence of ecSK. The alignment was performed with the program CLUSTAL V [31].

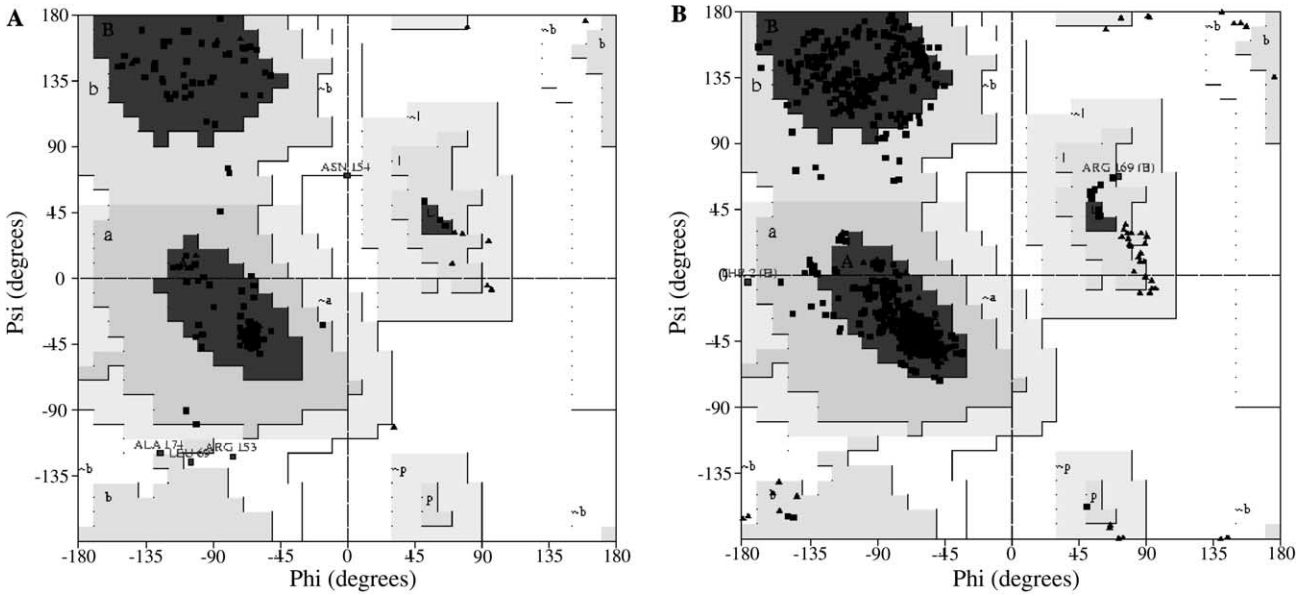


Fig. 2. (A) Ramachandran diagram ϕ–ψ plots for the mtSK structure and (B) for three crystallographic SK structures solved to resolution better than 2.6 Å.

remaining 8.9% in the additional allowed regions. The same analysis for three crystallographic ecSK structures (six chains) present 93.7% of residues in the most favorable, 6.1% additional allowed regions, and 0.7% generously allowed regions. The overall rating for the mtSK model is slightly poorer than the one obtained for the three structures of SK. However, it has over 90% of the residues in the most favorable regions.

Overall description

MtSK is an α/β protein consisting of a mixed β sheet surrounded by α helices. A central five stranded parallel β-sheet (β1–β5) presents the strand order 23145. The β-strands are flanked on either side by α helices (α1 and α8 on one side, α4, α5, and α7 on the other). Fig. 3 shows a schematic diagram of the mtSK structure, with shikimate and ADP/Mg^{2+} bound to the structure.

The ordering of the strands 23145 observed the mtSK structure classifies it as belonging to the same structural family as the nucleoside monophosphate (NMP) kinases. The mtSK structure exhibits the Walker A-motif located between β1 and α1 forming a canonical phosphate-binding loop (P-loop). The core of the mtSK structure forms a classical mononucleoside-binding fold [20].

It has been reported that NMP kinases undergo large confomation changes during catalysis. The regions responsible for this movement are NMP-binding site and the lid domain. The NMP-binding site is formed by a series of helices between strands 1 and 2 of the parallel β-sheet. The lid domain is a region of variable size and structure following the forth β-strand of the sheet [21,22]. The residues from 112 to 123 form the lid domain in the mtSK which has been reported to be highly dynamic and possibly flexible in solution in the ecSK structure [8].

ADP/Mg^{2+}-binding site

The molecular model for ternary complex mtSK–shikimate–ADP/Mg^{2+} indicates that ADP/Mg^{2+} is

Fig. 3. Ribbon diagram of the mtSK structure with shikimate and ADP/Mg^{2+} bound to the structure generated by MOLSCRIPT [32].

Table 1
Intermolecular hydrogen bonds between mtSK and ADP

Hydrogen bonds between active site and inhibitor			Distance (Å)
ADP	mtSK		
O1B	Leu10	O	3.49
O3B	Gly12	N	3.15
O1B	Ser13	N	3.08
O1B	Ser13	OG	2.66
O3A	Gly14	N	2.75
O3A	Lys15	N	3.35
O2B	Lys15	NZ	2.80
O3B	Lys15	NZ	2.54
O1B	Lys15	NZ	2.82
O2A	Ser16	OG	2.65
O2A	Ser16	N	3.40
O1A	Ser16	OG	2.98
O2A	Thr17	OG1	3.02
O2A	Thr17	N	3.08
N1	Arg110	NH1	3.05
O4	Arg110	NE	3.41
N3	Arg110	NH1	2.88
N7	Arg110	NH1	3.29
N6	Arg152	O	2.52
N1	Arg152	O	3.23
N1	Arg153	O	3.37
N6	Arg153	O	2.66
N3	Arg153	NE	2.80
N3	Arg153	NH2	2.68
N1	Arg153	N	3.20

tightly bound to the mtSK structure. The intermolecular hydrogen bonds are described in Table 1. Most of the intermolecular hydrogen bonds observed in the ecSK structure is conserved in the ternary complex mtSK–shikimate–ADP/Mg^{2+}. Fig. 4 shows the superposition of the ATP-binding site of mtSK and ecSK. As previously described a phosphate-binding loop (P-loop) accommodates the β-phosphate of ADP by donating hydrogen bonds from several backbone amides [8,14]. SKs contain a conserved stretch of sequence GXXXXGKT/S known as the Walker A-motif [23]. This motif forms the P-loop in the ecSK and mtSK structures. In addition to the Walker A-motif, the mtSK structure presents a modified Walker B-motif. The Walker B-motif is present in the majority of purine–nucleotide binding proteins. This motif, Z–Z–Asp–X–X–Gly (where Z is a hydrophobic residue and X is any residue) forms a loop around the γ-phosphate of the nucleotide. The B-motif present in the mtSK has a different conformation from that observed in proteins with full Walker B-motif [8], the Asp is replaced by Ser77. However, the conserved Gly80 of mtSK (Fig. 1) is in an almost identical position to the conserved Gly found in proteins with the full B-motif with its amide nitrogen hydrogen-bonded to the γ-phosphate of a bound ATP.

Shikimate-binding site

The crystallographic structure of ecSK indicated the presence of a strong electron density peak attributed to shikimate. However, the electron density was not clear enough to include shikimate in the molecular structure [8]. The docking simulations of shikimate to ecSK and mtSK identified that shikimate binds in a position analogous to nucleotide monophosphate in NMP kinases [8]. The shikimate binding domain, which follows strand β2, consists of helices α2 and α3 and the N-terminal region of helix α4. A total of four hydrogen bonds between ecSK and shikimate, in the model, involving the residues Lys15, Asp34, and Arg 136 was observed. For the mtSK model the same pattern was observed. Fig. 5A and B show the main residues involved in contact with shikimate in the complexes. Tables 2 and 3 show the intermolecular hydrogen bonds for both structures. The residues involved hydrogen

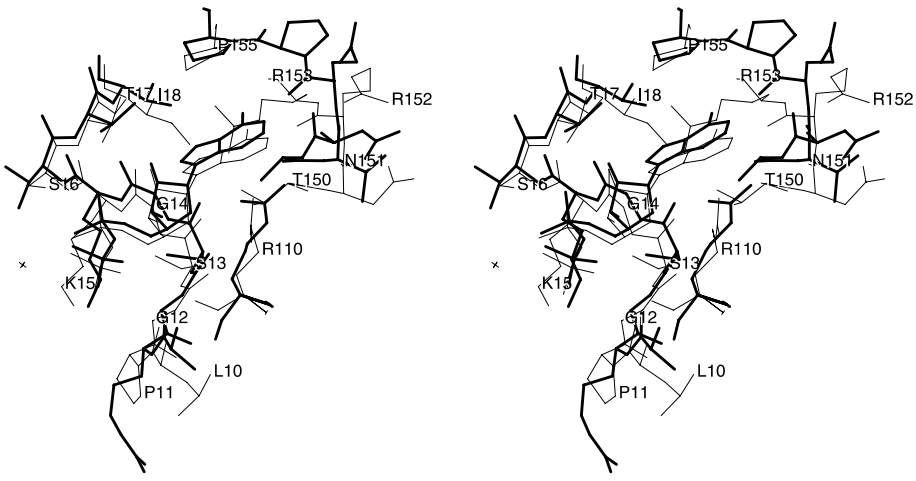


Fig. 4. Superimposed binding pockets of the ATP-binding site of mtSK (thick line) and ecSK (thin line).

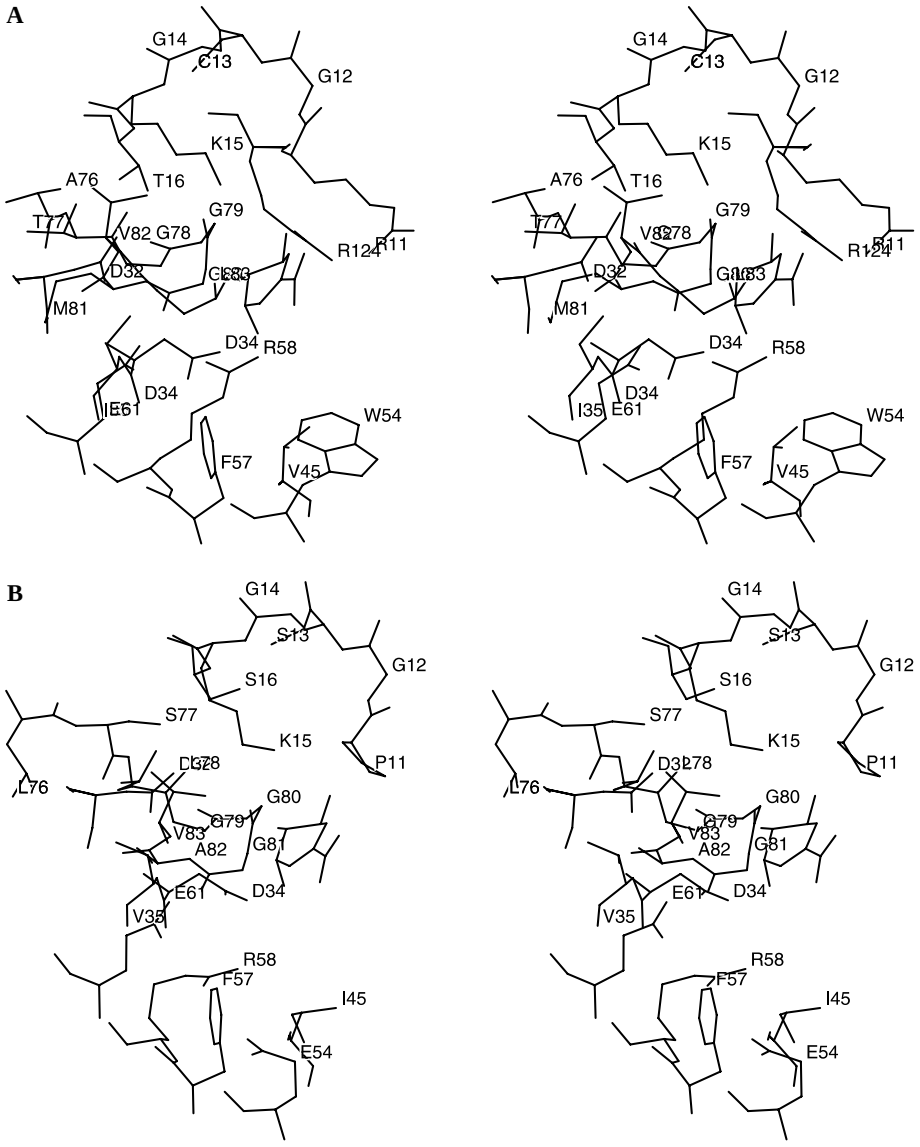


Fig. 5. Main residues involved in contact with shikimate in the complexes (A) mtSK:shikimate and (B) ecSK:shikimate.

bond identified from the docking simulation of shikimate to ecSK are in good agreement with that of crystallographic structure.

The electrostatic potential surface of the ecSK and mtSK complexed with shikimate calculated with GRASP [24] indicates the presence of some charge complementarity between shikimate and enzyme, nevertheless most of the residues in the binding pocket is hydrophobic in all structures. Figs. 6A and B show the molecular surfaces for mtSK and ecSK complexed with shikimate. The electrostatic potential surfaces of mtSK and ecSK show some striking differences. The main

Table 2
Intermolecular hydrogen bonds between mtSK and shikimate

Hydrogen bonds between active site and inhibitor			Distance (Å)
Shikimate	mtSK		
O1	Asp34	OD2	2.58
O2	Asp34	OD2	3.55
O3	Lys15	NZ	3.36
O5	Arg136	NH1	3.47

Table 3
Intermolecular hydrogen bonds between ecSK and shikimate

Hydrogen bonds between active site and inhibitor			Distance (Å)
Shikimate	ecSK		
O1	Asp34	OD2	2.64
O2	Asp34	OD2	3.17
O3	Lys15	NZ	3.46
O5	Arg11	NH1	2.77
O5	Arg136	NH1	3.51

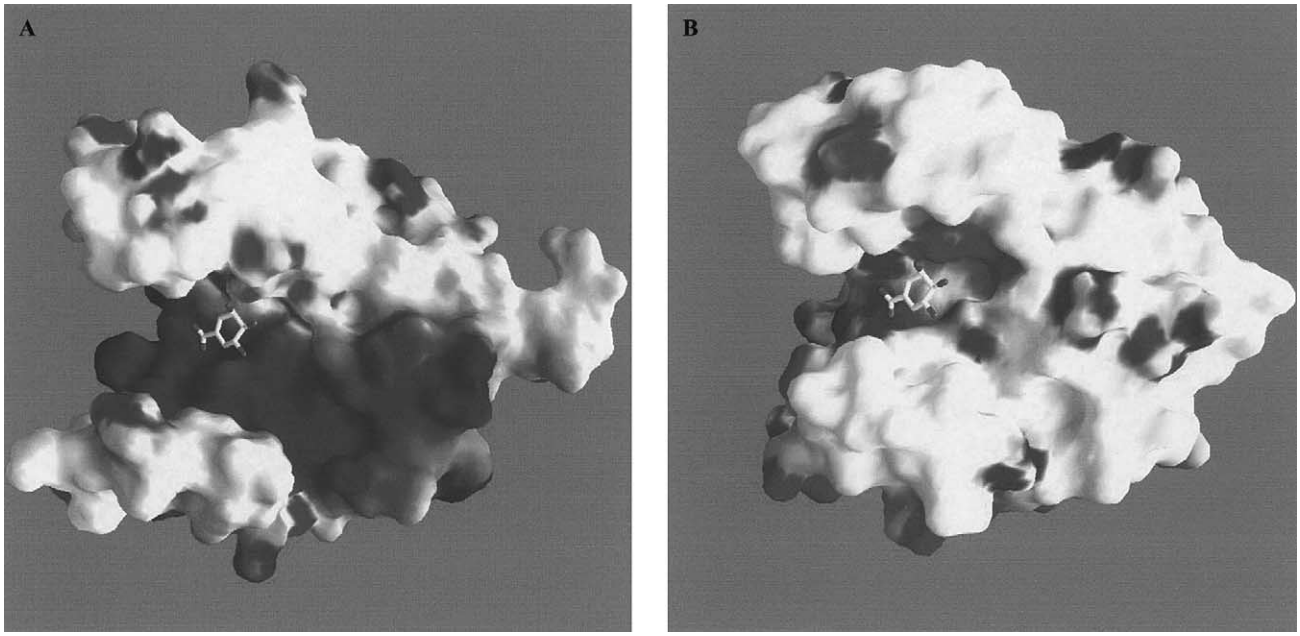


Fig. 6. (A) Molecular surfaces for mtSK and (B) ecSK complexed with shikimate generated with GRASP [24].

difference is the presence of a positive potential patch on the surface of mtSK, which is not observed in the ecSK surface. This positive patch indicates a concentration of positive charged residues in the mtSK structure, involving residues Arg21, Arg125, Lys128, Arg130, Lys135, Arg142, Arg147, Arg153, Arg160, His161, and Arg165. These residues are changed to polar or hydrophobic or negative charged residues in the ecSK sequence. Particularly interesting is the intermolecular hydrogen bond pattern between SK and shikimate. The residues Lys15, Asp34, and Arg136, involved in intermolecular hydrogen bonds, are conserved in both structures. The model strongly indicates that the shikimate binding domain is a well-conserved motif in SK structures. Furthermore, the alignment of 37 SK sequences, figure not shown, indicates that the main residues involved in intermolecular hydrogen bonds are conserved in all sequences. Such observation suggests that competitive inhibitors with shikimate will be able to inhibit most or even all SKs, since specificity and affinity between enzyme and its inhibitor depend on directional hydrogen bonds and ionic interactions, as well as on shape complementarity of the contact surfaces of both partners [25–30]. Further inhibition experiments may confirm this prediction.

Acknowledgments

This work was supported by grants from FAPESP (SMOLBNet), CNPq, CAPES, and Instituto do Milênio (CNPq-MCT). WFA (CNPq, 300851/98-7) and MSP (CNPq, 500079/90-0) are researchers for the Brazilian Council for Scientific and Technological Development.

References

[1] World Health Organization, Global Tuberculosis Control, WHO Report 2001, Geneva, Switzerland, WHO/CDS/TB/2001.287.
[2] N.E. Dunlap, J. Bass, P. Fujiwara, P. Hopewell, C.R. Horsburgh Jr., M. Salfinger, P.M. Simone, Diagnostic standards and classification of tuberculosis in adults and children, Am. J. Respir. Crit. Care Med. 161 (2000) 1376–1395.
[3] C.J.L. Murray, in: B.R. Bloom (Ed.), Tuberculosis: Pathogenesis, Protection, and Control, ASM, Washington, 1994, pp. 583–622.
[4] L.A. Basso, J.S. Blanchard, Resistance to antitubercular drugs, Adv. Exp. Med. Biol. 456 (1998) 115–144.
[5] R. Bentley, The shikimate pathway—a metabolic tree with many branches, Crit. Rev. Biochem. Mol. Biol. 25 (1990) 307–384.
[6] C. Ratledge, Nutrition, growth and metabolism, in: C. Ratledge, J.L. Stanford (Eds.), The Biology of the Mycobacteria, vol. 1, Academic Press, London, 1982, pp. 185–271.
[7] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B.G. Barrel, Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, Nature 393 (1998) 537–544.
[8] T. Krell, J.R. Coggins, A.J. Lapthorn, The three-dimensional structure of shikimate kinase, J. Mol. Biol. 278 (1998) 983–997.
[9] J.S. Oliveira, C.A. Pinto, L.A. Basso, D.S. Santos, Cloning and overexpression in soluble form of functional shikimate kinase and 5-enolpyruvylshikimate-3-phosphate synthase enzymes from *Mycobacterium tuberculosis*, Protein Express. Purif. 22 (2001) 430–435.
[10] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[11] A. Sali, J.P. Overington, Derivation of rules for comparative protein modeling from a database of protein structure alignments, Protein Sci. 3 (9) (1994) 1582–1596.
[12] A. Sali, L. Potterton, F. Yuan, H. van Vlijmen, M. Karplus, Evaluation of comparative protein modeling by MODELLER, Proteins 23 (3) (1995) 318–326.

[13] A. Sali, Modeling mutations and homologous proteins, Curr. Opin. Biotechnol. 6 (4) (1995) 437–451.
[14] T. Krell, J. Maclean, D.J. Boam, A. Cooper, M. Resmini, K. Brocklehurst, S.M. Kelly, N.C. Price, A.J. Lapthorn, J.R. Coggins, Biochemical and X-ray crystallographic studies on shikimate kinase: the important structural role of the P-loop lysine, Protein Sci. 10 (2001) 1137–1149.
[15] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy minimization and dynamics calculations, J. Comp. Chem. 4 (1983) 187–217.
[16] W. Braun, N. Go, Calculation of protein conformations by proton–proton distance constraints. A new efficient algorithm, J. Mol. Biol. 186 (3) (1985) 611–626.
[17] E. Katchalski-Katzir, I. Shariv, M. Eisenstein, A.A. Friesem, C. Aflalo, I.A. Vakser, Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques, Proc. Natl. Acad. Sci. USA 89 (1992) 2195–2199.
[18] E. Schönbrunn, S. Eschenburg, W.A. Shuttleworth, J.V. Schloss, N. Amrhein, J.N.S. Evans, W. Kabsch, Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate-3-phosphate synthase in atomic detail, Proc. Natl. Acad. Sci. USA 98 (2001) 1376–1380.
[19] R.A. Laskowski, M.W. MacArthur, D.K. Smith, D.T. Jones, E.G. Hutchinson, A.L. Morris, D. Naylor, D.S. Moss, J.M. Thorton, PROCHECK v.3.0—program to check the stereochemistry quality of protein structures—operating instructions, 1994.
[20] G.E. Schulz, Binding of nucleotides by proteins, Curr. Opin. Struct. Biol. 2 (1992) 61–67.
[21] C.W. Müller, G.J. Schlauderer, J. Reinstein, G.E. Schulz, Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding, Structure 4 (1996) 147–156.
[22] M. Gerstein, G. Schulz, C. Chothia, Domain closure in adenylate kinase Joints on either side of two helices close like neighboring fingers, J. Mol. Biol. 229 (2) (1993) 494–501.
[23] J.E. Walker, M. Saraste, M.J. Runswick, N.J. Gay, Distantly related sequences in the α and β subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold, EMBO J. 1 (1982) 945–951.
[24] A. Nicholls, K.A. Sharp, B. Honig, Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, Proteins 11 (4) (1991) 281–296.
[25] F. Canduri, L.G.V.L. Teodoro, C.C.B. Lorenzi, V. Hial, R.A.S. Gomes, J. Ruggiero Neto, W.F. de Azevedo Jr., Crystal structure of human uropepsin at 2.45 Å resolution, Acta Crystallogr. D 57 (2001) 1560–1570.
[26] W.F. De Azevedo Jr., H.J. Mueller-Dieckmann, U. Schulze-Gahmen, P.J. Worland, E. Sausville, S.-H. Kim, Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase, Proc. Natl. Acad. Sci. USA 93 (7) (1996) 2735–2740.
[27] W.F. De Azevedo Jr., S. Leclerc, L. Meijer, L. Havlicek, M. Strnad, S.-H. Kim, Inhibition of cyclin-dependent kinases by purine analogues: crystal structure of human CDK2 complexed with roscovitine, Eur. J. Biochem. 243 (1997) 518–526.
[28] W.F. De Azevedo Jr., F. Canduri, V. Fadel, L.G.V.L. Teodoro, V. Hial, R.A.S. Gomes, Molecular model for the binary complex of uropepsin and pepstatin, Biochem. Biophys. Res. Commun. 287 (1) (2001) 277–281.
[29] W.F. De Azevedo Jr., F. Canduri, N.J.F. da Silveira, Structural basis for inhibition of cyclin-dependent kinase 9 by flavopiridol, Biochem. Biophys. Res. Commun. 293 (1) (2002) 566–571.
[30] S.-H. Kim, U. Schulze-Gahmen, J. Brandsen, W.F. de Azevedo Jr., Structural basis for chemical inhibitor of CDK2, Prog. Cell Cycle Res. 2 (1996) 137–145.
[31] D.G. Higgins, A.J. Bleasby, R. Fuchs, CLUSTAL V: improved software for multiple sequence alignment, Comput. Appl. Biosci. 8 (2) (1992) 189–191.
[32] P. Kraulis, MOLSCRIPT: a program to produce both detailed and schematic plots of proteins, J. Appl. Cryst. 24 (1991) 946–950.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Protein Expression and Purification 27 (2003) 158–164
Protein Expression & Purification
www.elsevier.com/locate/yprep

Cloning, overexpression, and purification of functional human purine nucleoside phosphorylase

Rafael G. Silva,[a] Luiz Pedro S. Carvalho,[a] Jaim S. Oliveira,[a] Clotilde A. Pinto,[a] Maria A. Mendes,[b] Mário S. Palma,[b] Luiz A. Basso,[a,1] and Diógenes S. Santos[a,*]

[a] *Grupo de Microbiologia Molecular e Funcional, Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves, 9500, Porto Alegre-RS 91501-970, Brazil*
[b] *Laboratório de Biologia Estrutural e Zooquímica (CEIS), Departamento de Biologia, Instituto de Biociências, Universidade do Estado de São Paulo, Rio Claro, SP 13506-900, Brazil*

Received 26 June 2002, and in revised form 5 September 2002

Abstract

Purine nucleoside phosphorylase (PNP) catalyzes the phosphorolysis of the *N*-ribosidic bonds of purine nucleosides and deoxynucleosides. A genetic deficiency due to mutations in the gene encoding for human PNP causes T-cell deficiency as the major physiological defect. Inappropriate activation of T-cells has been implicated in several clinically relevant human conditions such as transplant tissue rejection, psoriasis, rheumatoid arthritis, lupus, and T-cell lymphomas. Human PNP is therefore a target for inhibitor development aiming at T-cell immune response modulation. In addition, bacterial PNP has been used as reactant in a fast and sensitive spectrophotometric method that allows both quantitation of inorganic phosphate (P_i) and continuous assay of reactions that generate P_i such as those catalyzed by ATPases and GTPases. Human PNP may therefore be an important biotechnological tool for P_i detection. However, low expression of human PNP in bacterial hosts, protein purification protocols involving many steps, and low protein yields represent technical obstacles to be overcome if human PNP is to be used in either high-throughput drug screening or as a reagent in an affordable P_i detection method. Here, we describe PCR amplification of human PNP from a liver cDNA library, cloning, expression in *Escherichia coli* host, purification, and activity measurement of homogeneous enzyme. Human PNP represented approximately 42% of total soluble cell proteins with no induction being necessary to express the target protein. Enzyme activity measurements demonstrated a 707-fold increase in specific activity of cloned human PNP as compared to control. Purification of cloned human PNP was achieved by a two-step purification protocol, yielding 48 mg homogeneous enzyme from 1 L cell culture, with a specific activity value of 80 U mg^{-1}.

Purine nucleoside phosphorylase (PNP)[2] catalyzes the reversible phosphorolysis of *N*-ribosidic bonds of both purine nucleosides and deoxynucleosides, except adenosine, generating purine base and ribose (or deoxyribose) 1-phosphate [1]. The major physiological substrates for mammalian PNP are inosine, guanosine, and 2′-deoxyguanosine [2]. PNP is specific for purine nucleosides in the β-configuration and exhibits a preference for ribosyl-containing nucleosides relative to the analogs containing the arabinose, xylose, and lyxose stereoisomers [3]. Moreover, PNP cleaves glycosidic bond with inversion of configuration to produce α-ribose 1-phosphate, as shown by its catalytic mechanism [4]. The human erythrocyte enzyme is active as a trimer, with each subunit presenting a molecular weight of 32,000 [5].

It has been reported that genetic deficiencies of PNP cause gradual decrease in T-cell immunity, though keeping B-cell immunity normal as well as other tissues [6]. The absence of PNP activity is thought to lead to accumulation of deoxyguanosine triphosphate, which inhibits the enzyme ribonucleotide reductase and ensuing DNA synthesis inhibition, thereby preventing cellular proliferation required for an immune response [7].

* Corresponding author. Fax: +55-51-3316-6234.
E-mail addresses: labasso@dna.cbiot.ufrgs.br (L.A. Basso), diogenes@dna.cbiot.ufrgs.br (D.S. Santos).
[1] Also corresponding author.
[2] *Abbreviations used:* ESI-MS, electrospray ionization mass spectrometry; IPTG, isopropyl β-D-thiogalactoside; LB, Luria–Bertani; MESG, 2-amino-6-mercapto-7-methylpurine ribonucleoside; P_i, inorganic phosphate; PNP, purine nucleoside phosphorylase.

1046-5928/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
PII: S1046-5928(02)00602-2

Thus, this enzyme is a potential target for drug development, which could induce immune suppression to treat, for instance, autoimmune diseases, T cell leukemia, and lymphoma and organ transplantation rejection [8]. Moreover, some PNP inhibitors have been tested in combination with nucleoside antiviral and anticancer drugs, showing the ability to potentiate the in vivo activity of these drugs [9].

In addition, a simple and rapid spectrophotometric method using 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG) and bacterial PNP has been developed to both quantitate P_i in solutions and to follow kinetics of P_i release from enzymatic reactions [10]. Although the reaction equilibrium favors the nucleoside formation with natural substrates, this synthetic substrate favors the cleavage of its own glycosidic bond [11]. Further studies have shown the possibility of using PNP as coupled enzyme in a continuous spectrophotometric assay for phosphorylase kinase [12] and protein phosphatase catalyzed reactions [13], as well as in combination with phosphodeoxyribomutase to remove P_i from solutions [14].

To pave the way for both using immobilized human PNP enzyme as a target for drug development in high-throughput drug screening and to test its viability as a tool for coupled enzymatic assay, the human PNP encoding cDNA was PCR amplified, cloned, and sequenced. The enzyme was overexpressed in soluble form in *Escherichia coli* BL21(DE3) host cells by a low-cost and simple protocol. The enzyme was purified by a two-step purification protocol, yielding cloned human PNP enzyme with the same specific activity found by other authors. This improved and simpler protocol for human PNP protein production in large quantities will contribute to current efforts towards the search for new drugs and the development of a low-cost P_i detection coupled assay.

Materials and methods

PCR amplification and cloning of human pnp cDNA

Synthetic oligonucleotide primers (First: 5′ a*atg*gagaacggatacacctatg 3′, second: 5′ a*tca*actggctttgtcagggag 3′, third: 5′ t**cat*atg***gagaacggatacacc 3′, and fourth: 5′ t**aagctt***tca* actggctttgtcag 3′) were designed based on the *pnp* cDNA sequence reported [15]. The first and third primers were complementary to the amino-terminal coding strand, and the second and fourth ones, to the carboxyl-terminal coding strand. The start and stop codons are shown in italics. The 5′ *Nde*I and 3′ *Hin*dIII restriction sites of the third and fourth primers are shown in bold. The first and second primers were designed to be complementary to 22 and 21 bases of, respectively, the N- and C-terminal ends of the PNP cDNA. A non-complementary A residue (deoxyadenyl-3′) at the 5′ end of each primer was added to protect the start and stop codons of primers 1 and 2, respectively, since, in our hands, occasionally a base is removed from the 5′-end of primers during PCR amplification experiments (data not shown). Removal of a nucleotide from the 3′-end of the amplification primers would have no effect on the gene product. Moreover, the extra four bases at the 3′-end of primers 1 and 2 were added to improve the likelihood of PCR amplification of the PNP cDNA. The primers (first and second) were used to amplify the *pnp* cDNA (870 bp) from a human liver cDNA expanded library in λ TriplEX phage (Clontec), using standard PCR conditions with a hot start at 99 °C for 10 min, and *Pfu* DNA polymerase (Stratagene). This PCR fragment was cloned into pCR-Blunt vector (Invitrogen) to be employed as DNA template for further amplification attempts. To introduce the 5′ *Nde*I and 3′ *Hin*dIII restriction sites, the third and fourth primers were used to amplify, using the same PCR conditions, the PNP cDNA that was cloned into the pCR-Blunt vector in the first round of PCR amplification. The PCR product was purified by electrophoresis on low-melting agarose, cloned into pCR-Blunt vector, digested with *Nde*I and *Hin*dIII, and ligated into a pET-23a(+) expression vector (Novagen), which had previously been digested with the same restiction enzymes. The DNA sequence of the amplified human *pnp* cDNA was determined using the MegaBace system (Amershan Pharmacia Biotech) to confirm the identity of the cloned DNA and to ensure that no mutations were introduced by the PCR amplification step.

Overexpression of recombinant human PNP

The pET23a(+)::*pnp* recombinant plasmid was transformed into electrocompetent *E. coli* BL21(DE3) cells and selected on LB agar plates containing 50 μg ml^{-1} carbenicillin. Control experiments were performed under the same experimental conditions, except that transformed *E. coli* cells harbored the expression vector lacking the target DNA insert. Single recombinant colonies were used to inoculate 5 ml LB medium containing carbenicillin (50 μg ml^{-1}). To evaluate human PNP protein expression as a function of cell growth phase, *E. coli* BL21(DE3) cells harboring pET-23a(+)::*pnp* recombinant plasmid were grown in the absence of IPTG and samples were removed at various times for OD_{600} measurements and for electrophoretic analysis. The cells were thus incubated at 180 rpm at 37 °C for 5 h10 min, 5 h40 min, 6 h20 min, 9 h20 min, 12 h20 min, and 28 h without addition of IPTG, harvested by centrifugation at 20,800*g* for 20 min, and stored at −20 °C. The stored cells were suspended in 50 mM Tris–HCl (pH 7.6), disrupted by sonication using two 10-s pulses to release the proteins, and cell debris

was separated by centrifugation at 23,500*g* for 30 min at 4 °C. The soluble protein content was analyzed by SDS–PAGE [16]. The proportion of recombinant human PNP to total soluble proteins in SDS–PAGE gels was estimated using a GS-700 imaging densitometer (Bio-Rad).

Purification of recombinant human PNP

Single colonies *E. coli* BL21(DE3) harboring the recombinant plasmid were used to inoculate 1 L LB medium containing 50 μg ml^{-1} carbenicillin and grown for 22 h at 37 °C and 180 rpm for large scale protein production. Cells (4 g) were harvested by centrifugation at 20,800*g* at 4 °C and stored at −20 °C. The cells were suspended in 15 ml of 50 mM Tris–HCl (pH 7.6) (buffer A), incubated for 30 min in the presence of lysozyme (0.2 mg ml^{-1}), disrupted by sonication as described above, and centrifuged at 48,000*g* for 60 min at 4 °C. The supernatant was incubated with 1% (w/v) of streptomycin sulfate and centrifuged at 48,000*g* for 30 min at 4 °C. The supernatant was dialyzed against buffer A and centrifuged at 48,000*g* for 30 min at 4 °C. The resulting supernatant was loaded on an FPLC Q-Sepharose Fast Flow (26 × 9.5) column (Amershan Pharmacia Biotech) pre-equilibrated with the same buffer. The column was washed with 10 volumes of the same buffer and the absorbed material was eluted with a linear gradient (0–100%) of 20 times the column volume of 50 mM Tris–HCl, 0.2 mM NaCl (pH 7.6) (buffer B). The recombinant protein eluted from the anion exchange column at 40% buffer B, the fractions were pooled (82 ml), concentrated down to 6.6 ml using an Amicon ultrafiltration cell (MWCO 10,000 Da), and applied to an FPLC Sephacryl S-200 (26 × 60) (Amersham Pharmacia Biotech) column pre-equilibrated with buffer A. The column was run with 1 volume of the same buffer. The recombinant human PNP protein eluted in a total volume of 10 ml and stored in 85% ammonium sulfate solution. The protein content was analyzed by SDS–PAGE [16].

Protein determination

Protein concentration was determined by the method of Bradford et al. [17] using the Bio-Rad protein assay kit (Bio-Rad) and bovine serum albumin as standard.

Mass spectrometry analyses

The homogeneity of protein preparation was assessed by mass spectrometry (MS), employing some adaptations made to the system described by Chassaigne and Lobinski [18]. Samples were analyzed on a triple quadrupole mass spectrometer, model QUATTRO II, equipped with a standard electrospray (ESI) probe (Micromass, Altrinchan), adjusted to ca. 250 μl min^{-1}. The source temperature (80 °C) and needle voltage (3.6 kV) were maintained constant throughout the experimental data collection, applying a drying gas flow (nitrogen) of 200 L h^{-1} and a nebulizer gas flow of 20 L h^{-1}. The mass spectrometer was calibrated with intact horse heart myoglobin and its typical cone-voltage induced fragments. The subunit molecular mass of recombinant human PNP was determined by ESI-MS, adjusting the mass spectrometer to give a peak width at half-height of 1 mass unit, and the cone sample to skimmer lens voltage controlling the ion transfer to mass analyzer was set to 38 V. About 50 pmol (10 μl) of each sample was injected into electrospray transport solvent. The ESI spectrum was obtained in the multi-channel acquisition mode, scanning from 500 to 800 m/z at scan time of 5 s. The mass spectrometer is equipped with MassLynx and Transform softwares for data acquisition and spectra handling.

N-terminal amino acid sequencing

The N-terminal amino acid residues of purified recombinant human PNP were identified by automated Edman degradation sequencing using a PPSQ 21A gas-phase sequencer (Shimadzu).

Purine nucleoside phosphorylase assay

Recombinant human PNP was assayed in the forward direction in 50 mM Tris–HCl, pH 7.6. Enzyme activity was measured by the difference in absorbance between 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG) and the purine base product of its reaction with inorganic phosphate (P_i) catalyzed by PNP [10]. The bacterial PNP enzyme and MESG are commercially available as the Enzchek phosphate assay kit (Molecular Probes); the assay was performed by replacing the commercial PNP by the recombinant human PNP. This reaction gives an absorbance increase at 360 nm with an extinction coefficient value of 11,000 M^{-1} cm^{-1} at pH 7.6 [10]. Initial steady-state rates were calculated from the linear portion of the reaction curve for extracts of *E. coli* BL21(DE3) cells harboring pET-23a(+)::*pnp* plasmid and all fractions of the purification protocol.

Results and discussion

The nucleotide sequence analysis using the dideoxy-chain termination method of the PCR-amplified human *pnp* cDNA (870 bp) confirmed the identity of the coding DNA sequence of cloned fragment and demonstrated that no mutations were introduced by the PCR amplification steps. Although the strategy used here for cloning did not involve complementation [19] and

hybridization [15] previously used for cloning *pnp* cDNA, it proved to be suitable and efficient. Probably, owing to the lack of complete sequence complementarity between primers three and four and cDNA, no PCR product of the correct size (870 bp) could be obtained using these restriction site-containing primers under a number of experimental conditions tested for amplification. Hence, the *pnp* cDNA was first cloned without restriction sites and re-amplified from this recombinant plasmid using the restriction site-containing primers (third and fourth ones).

Human PNP was overexpressed in *E. coli* BL21(DE3) cells carrying pET-23a(+)::*pnp* recombinant plasmids. Analysis by SDS–PAGE with Coomassie blue staining indicated that the cell extracts contained a significant amount of protein with subunit molecular weight in agreement with the expected MW for human PNP [5] (Fig. 1). Densitometric quantification of the SDS–PAGE protein bands showed that recombinant human PNP constituted approximately 42% of total protein present in the soluble cell extract under the experimental conditions used. Enzyme activity measurements demonstrated that there was a 707-fold increase in specific activity for human PNP when *E. coli* harboring either pET-23a(+)::*pnp* or pET-23a(+) crude extracts were compared (Table 1). Contrary to what was previously described for this enzyme [20], where human PNP represented about 5% of total protein present in the soluble cell extracts with addition of IPTG (1 mM final concentration), overexpression was achieved here (after 22 h of cell growth) with no addition of IPTG.

Recombinant human PNP protein expression as a function of cell growth phase in the absence of IPTG could be detected at all time intervals tested (Fig. 2). However, recombinant human PNP protein expression

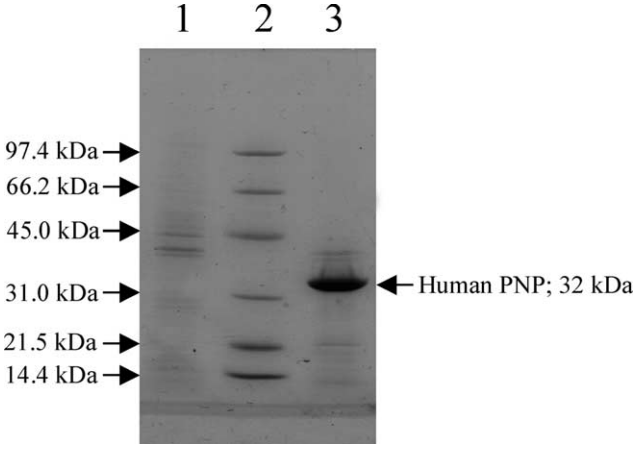


Fig. 1. SDS–PAGE analysis of protein-soluble crude extracts. Overexpression of human PNP after 22 h of cell growth in LB medium without addition of IPTG. Lane 1, *E. coli* BL21(DE3) [pET-23a(+)] (control); lane 2, MW markers; lane 3, *E. coli* BL21(DE3) [pET-23a(+)::*pnp*].

Table 1
Measurements of recombinant human purine nucleoside phosphorylase enzyme activity

Cell extract[a]	Sp act[b] (SA, U mg^{-1})	SA cloned/SA control
Control	0.02	1.00
PNP	17.04	707.10

[a] Cell crude extract in 50 mM Tris–HCl, pH 7.6.
[b] U ml^{-1}/mg ml^{-1}.

in *E. coli* BL21(DE3) host cells as compared to the control cells transformed with pET-23a(+) plasmid appeared to reach a maximum in the stationary phase (Fig. 2). The values for the proportion of recombinant human PNP to total soluble protein indicated in Fig. 2 are underestimates, since lower loads of protein on SDS–PAGE indicated a proportion of approximately 40% human PNP to total soluble protein after 22 h of growth (data not shown), in agreement with the results presented in Fig. 1.

Hosts for pET vectors (Novagen) harbor the highly processive T7 RNA polymerase under control of the IPTG-inducible *lac*UV5 promoter. Although it is often argued that the cost of IPTG limits the usefulness of *lac* promoter to high-added-value products, it was shown here that high levels of expression of human PNP could be obtained with pET vectors as cells entered the stationary phase without addition of inducer in LB medium. Similar experimental observations were reported using the pET system [21,22]. It has been demonstrated that when λDE3 hosts are grown to stationary phase in media lacking glucose, cyclic AMP mediated derepression of both the wild type and *lacUV5* promoters occurs [23]. These authors also proposed that cyclic AMP, acetate, and low pH are required to high-level expression in the absence of IPTG induction when cells approach stationary phase in complex medium, and that derepression of the *lac* operon in the absence of IPTG may be part of a general cellular response to nutrient limitation. As described in the pET System Manual (www.novagen.com), the *E. coli* BL21(DE3) host strain is a lysogen of bacteriophage DE3, a lambda derivative that carries a DNA fragment containing the *lacI* gene, the *lacUV5* promoter, and the gene for T7 RNA polymerase. This fragment is inserted into the *int* gene, preventing DE3 from integrating into or excising from the chromosome without a helper phage. Once a DE3 lysogen is formed, the only promoter known to direct transcription of the T7 RNA polymerase gene is the *lacUV5* promoter, which is inducible by IPTG. A characteristic of the pET-23a(+) expression vector is that it contains a "plain" T7 promoter. Background expression is minimal in the absence of T7 RNA polymerase because the host RNA polymerases do not initiate from T7 promoters and the cloning sites in pET plasmids are in regions weakly transcribed (if at all) by

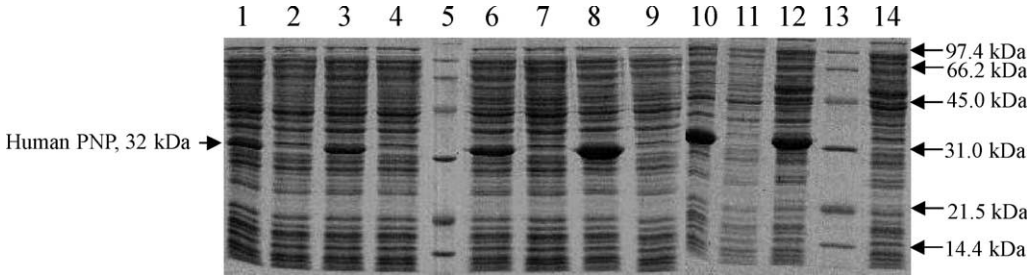


Fig. 2. SDS–PAGE analysis of total soluble protein as a function of growth time. A time course showing an increase of recombinant human PNP expression in the stationary phase. Lanes 5 and 13, MW markers; lanes 2, 4, 7, 9, 11, and 14, *E. coli* BL21(DE3) [pET-23a(+)] (control); lanes 1, 3, 6, 8, 10, and 12, *E. coli* BL21(DE3) [pET-23a(+)::*pnp*]. In brackets are given the values for, respectively, time intervals of sample removal, their OD_{600}, human PNP proportion to total soluble protein for the following lanes: 1 (5 h10 min, 0.315, 9.7%), 3 (5 h40 min, 0.380, 10.8%), 6 (6 h20 min, 0.511, 13.9%), 8 (9 h20 min, 0.846, 21.7%), 10 (12 h20 min, 1.440, 31.6%), and 12 (28 h, 4.358, 30.0%).

read-through activity of bacterial RNA polymerase. The fragment of DNA cloned into the pET-23a(+) vector harbors no other piece of DNA, except the coding DNA sequence of human PNP. It seems therefore unlikely that other promoter is present that might titrate out the *lac* repressor.

Human PNP was purified as described in Materials and methods and analyzed by SDS–PAGE with Coomassie blue staining. The relative mobility of the polypeptide chain in SDS–PAGE indicates a homogeneous protein with molecular weight value of approximately 32 kDa (Fig. 3). The enzymatic assay and protein concentration determination showed a specific activity of $80 \pm 3\,U\,mg^{-1}$ for the homogeneous target protein with MESG as substrate, indicating that the protein purification protocol used resulted in a 3.2-fold purification (Table 2). The specific activity value for the recombinant human PNP compares favorably with the value of $55\,U\,mg^{-1}$ for the homogeneous human erythrocyte PNP using guanosine as substrate [5]. Approximately 48 mg of homogeneous cloned human PNP protein could be obtained from 4 g of cells or, stated otherwise, approximately $48\,mg\,L^{-1}$ of LB medium (Table 2). Protocols for human PNP purification have been described either using erythrocytes [1,5,24] or prokaryotic host cells expressing recombinant human PNP [20,25].

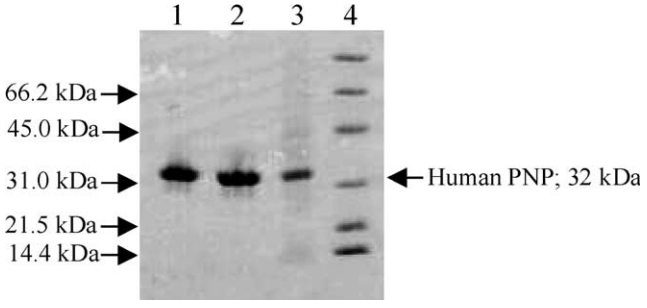


Fig. 3. SDS-PAGE analysis of pooled fractions from the various purification protocol steps. Lane 1, S-200 gel filtration; lane 2, Q-Sepharose Fast Flow ion exchange; lane 3, crude extract; lane 4, MW markers.

However, PNP is estimated to make up only 0.04% of the total protein in human erythrocytes and purification protocol involves a number of steps resulting in only 1.2 mg of homogeneous human PNP from 130 ml of freshly drawn blood [24]. Recombinant human PNP expressed in *E. coli* cells has been shown to represent approximately 5% of the soluble protein cell homogenates [20]. However, the purification protocol using a dye-matrix resin and isoelectrofocusing chromatography yielded about 2.6 mg of purified recombinant human PNP from 1.0 g of cells [20]. Therefore, to the best of our knowledge, the straightforward two-step purification protocol (ion exchange and gel filtration) associated with high expression of the target protein (42%), which results in a yield of approximately 12 mg of soluble and functional protein from 1.0 g of cells, represents a significant improvement on previous methods of expression and separation of recombinant human PNP.

The subunit molecular mass of active human PNP was determined to be 31,966 Da by electrospray ionization mass spectrometry (ESI-MS), consistent with the post-translational removal of the N-terminal methionine residue from the full length gene product (predicted mass: 32,097 Da). The ESI-MS result revealed no peak at the expected mass for *E. coli* PNP (25,950 Da), thus, providing evidence for both the identity and purity of the recombinant human protein. The first nine N-terminal amino acid residues of the recombinant protein were identified as ENGYTYEDY by the Edman degradation method. This result unambiguously identifies the recombinant protein as human PNP and confirms removal of the N-terminal methionine residue from it. A common type of co-/post-translational modification of proteins synthesized in prokaryotic cells is modification at their N termini. Methionine aminopeptidase catalyzed cleavage of initiator methionine is usually directed by the penultimate amino acid residues with the smallest side chain radii of gyration (glycine, alanine, serine, threonine, proline, valine, and cysteine) [27]. Interestingly, the *E. coli* expressed recombinant human PNP enzyme seems not to conform to this rule, since the

Table 2
Purification of human purine nucleoside phosphorylase from *E. coli* BL21(DE3) [pET23a(+)::*pnp*] cells

Purification step	Protein (mg)	Units (U)	Sp act[a] ($U mg^{-1}$)	Purification fold	Yield (%)
Crude extract	382.95	9569.92	25	1.0	100
Ion exchange	66.42	4994.12	75	3.0	52
Gel filtration	48.10	3851.84	80	3.2	41

[a] $U ml^{-1}/mg ml^{-1}$.

N-terminal methionine was removed, despite the penultimate amino acid residue being a charged residue (glutamate).

As previously pointed out, there are a number of potential applications for human PNP. From the viewpoint of biotechnological tools development, a bacterial enzyme has been used to follow P_i release kinetics [9] and human PNP has been studied for its application in purine nucleoside analog synthesis [25,26]. The commercially available "bacterial" protein used in continuous spectrophotometric assay for P_i detection has been shown to be about one-third PNP by weight [10] and to contain a large content of albumin [14], which may require further purification for its use in coupled assays. From the viewpoint of drug development, inappropriate activation of T-cells has been proposed or documented in several clinically relevant human conditions, such as transplant tissue rejection, psoriasis, rheumatoid arthritis, T-cell lymphomas, and lupus. Since genetic deficiency of human PNP causes T-cell deficiency as the major physiological defect, specific PNP inhibitors may provide useful agents for these disorders. Accordingly, a transition-state analog (Immucillin-H) that inhibits PNP enzyme activity has been shown to inhibit the growth of malignant T-cell leukemia cell lines with the induction of apoptosis [28].

The real-time BIA from Pharmacia Biosensor AB (BIACORE), which is a label-free technology for monitoring biomolecular interactions as they occur, was chosen for high-throughput drug screening studies. The detection principle of BIACORE equipment relies on surface plasmon resonance (SPR), an optical phenomenon that arises when light illuminates thin conducting films under specific concitions. Direct immobilization of protein ligands is possible through linkages between the *N*-hydroxy-succinimide (NHS) ester groups on a hydrophilic dextran matrix and amine groups on proteins. In protein molecules, NHS ester cross-linking reagents couple principally with the α-amines at the N-terminals and the ε-amines of lysine side chains [29]. Since human PNP protein has 12 lysine residues in its primary sequence, it is likely that immobilization of the recombinant protein will not present difficulties. The work presented here represents a crucial step in our efforts towards both using immobilized human PNP enzyme as a target for drug development in high-throughput drug screening and testing its suitability as a tool for coupled enzymatic assay. Moreover, since the deposited atomic coordinates of human PNP (PDB access codes: 1ULA and 1ULB) have been withdrawn due to low resolution, the work presented here also provides protein in quantities necessary for crystallographic studies. Indeed, we have recently deposited the atomic coordinates of the crystal structure of the recombinant human PNP protein solved at 2.3 Å resolution (PDB access code: 1M73).

Acknowledgments

Financial support for this work was provided by Millenium Initiative Program MCT-CNPq, Ministry of Health—Secretary of Health Policy (Brazil) to D.S.S. and L.A.B. D.S.S. and L.A.B. also acknowledge grants awarded by CNPq and FINEP. We thank Marcelo Brígido for his generous gift of human cDNA library, Denise Machado for assistance in human liver cDNA library expansion, and Deise Potrich for nucleotide sequence analysis.

References

[1] R.E. Parks Jr., R.P. Agarwal, in: P.D. Boyer (Ed.), The Enzymes, Academic Press, New York, 1972, pp. 483–514.
[2] V.L. Schramm, Enzymatic transition states and transition state analog design, Annu. Rev. Biochem. 67 (1998) 693–720.
[3] J.D. Stoeckler, C. Cambor, R.E. Parks Jr., Human erythrocytic purine nucleoside phosphorylase: reaction with sugar-modified nucleosides substrates, Biochemistry 19 (1980) 102–107.
[4] D.J.T. Porter, Purine nucleoside phosphorylase. Kinetic mechanism of the enzyme from calf spleen, J. Biol. Chem. 267 (1992) 7342–7351.
[5] A.S. Lewis, B.A. Lowy, Human erythrocytes purine nucleoside phosphorylase: molecular weight and physical properties, J. Biol. Chem. 254 (1979) 9927–9932.
[6] W. Stoop, B.J.M. Zegers, G.F.M. Hendrickx, L.H.S. van Heukelom, G.E.J. Staal, P.K. de Bree, S.K. Wadman, R.E. Ballieux, Purine nucleoside phosphorylase deficiency associated with selective cellular immunodeficiency, N. Engl. J. Med. 296 (1977) 651–655.
[7] B.S. Mitchell, E. Meijias, P.E. Daddona, W.N. Kelley, Purinogenic immunodeficiency diseases: selective toxicity of deoxyribonucleosides for T-cells, Proc. Natl. Acad. Sci. USA 75 (1978) 5011–5014.
[8] J.D. Stoeckler, in: R.I. Glazer (Ed.), Developments in Cancer Chemotherapy, CRC Press, Boca Raton, FL, 1984, pp. 35–60.
[9] L.L. Bennett Jr., P.W. Allan, P.E. Noker, L.M. Rose, S. Niwas, J.A. Montgomery, M.D. Erion, Purine nucleoside phosphorylase

inhibitors: biochemical and pharmacological studies with 9-benzyl-9-deazaguanine and related compounds, J. Pharm. Exp. Ther. 266 (1993) 707–714.
[10] M.R. Webb, A continuous spectrofotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems, Proc. Natl. Acad. Sci. USA 89 (1992) 4884–4887.
[11] M. Brune, J. Hunter, J.E.T. Corrie, M.R. Webb, Direct, real-time measurements of rapid inorganic phosphate release using a novel fluorescent probe and its application to actinomyosin subfragment 1 ATPase, Biochemistry 33 (1994) 8262–8271.
[12] Z.X. Wang, Q. Cheng, S.D. Killilea, A continuous spectrophotometric assay for phosphorylase kinase, Anal. Biochem. 230 (1995) 55–61.
[13] Q. Cheng, Z.X. Wang, S.D. Killilea, A continuous spectrophotometric assay for protein phosphatases, Anal. Biochem. 226 (1995) 68–73.
[14] A.E. Nixon, J.L. Hunter, G. Bonifacio, J.F. Eccleston, M.R. Webb, Purine nucleoside phosphorylase: its use in a spectroscopic assay for inorganic phosphate and for removing inorganic phosphate with aid of phosphodeoxyribomutase, Anal. Biochem. 265 (1998) 299–307.
[15] S.R. Williams, J.M. Goddard, D.W. Martin Jr., Human purine nucleoside phosphorylase cDNA sequence and genomic clone characterization, Nucleic Acids Res. 12 (1984) 5779–5787.
[16] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.
[17] M.M. Bradford, R.A. McRorie, W.L. Williams, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding, Anal. Biochem. 72 (1976) 248–254.
[18] H. Chassaigne, R. Lobinski, Characterization of horse kidney metallothionein isoforms by electrospray MS and reversed-phase HPLC-electrospray MS, Analyst. 123 (1998) 2125–2130.
[19] J.M. Goddard, D. Caput, S. Williams, D.W. Martin, Cloning of human purine nucleoside phosphorylase cDNA sequences by complementation in *Escherichia coli*, Proc. Natl. Acad. Sci. USA 80 (1983) 4281–4285.
[20] M.D. Erion, K. Takabayashi, H.B. Smith, J. Kessi, S. Wagner, S. Hönger, S. Shames, S. Ealick, Purine nucleoside phosphorylase. 1. Structure–function studies, Biochemistry 36 (1997) 11725–11734.
[21] K.C. Kelley, K.J. Huestis, D.A. Austen, C.T. Sanderson, M.A. Donoghue, S.K. Stickel, E.S. Kawasaki, M.S. Osburne, Regulation of *sCD4-183* gene expression from phage-T7-based vectors in *Escherichia coli*, Gene 156 (1995) 33–36.
[22] J.S. Oliveira, C.A. Pinto, L.A. Basso, D.S. Santos, Cloning and overexpression in soluble form of functional shikimate kinase and 5-enolpyruvylshikimate 3-phosphate synthase enzymes from *Mycobacterium tuberculosis*, Protein Expr. Purif. 22 (2001) 430–435.
[23] T.H. Grossman, E.S. Kawazaki, S.R. Punreddy, M.S. Osburne, Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability, Gene 209 (1998) 95–103.
[24] V. Zannis, D. Doyle, D.W Martin Jr., Purification and characterization of human erythrocyte purine nucleoside phosphorylase and its subunits, J. Biol. Chem. 253 (1978) 504–510.
[25] J.D. Stoeckler, F. Poirot, R.M. Smith, R.E. Parks Jr., S.E. Ealick, K. Takabayashi, M.D. Erion, Purine nucleoside phosphorylase. 3. Reversal of purine base specificity by site-directed mutagenesis, Biochemistry 36 (1997) 11749–11756.
[26] M.D. Erion, J.D. Stoeckler, W.C. Guida, R.L. Walter, S.E. Ealick, Purine nucleoside phosphorylase. 2. Catalytic mechanism, Biochemistry 36 (1997) 11735–11748.
[27] W.T. Lowther, B.W. Matthews, Structure and function of the methionine aminopeptidases, Biochim. Biophys. Acta 1477 (2000) 157–167.
[28] G. Kicska, L. Long, H. Hörig, G. Fairchild, P.C. Tyler, R.H. Furneaux, V.L. Schramm, H.L. Kaufman, Immucilli H, a powerful transition-state analog inhibitor of purine nucleoside phosphorylase, selectively inhibits human T lymphocytes, Proc. Natl. Acad. Sci. USA 98 (2001) 4593–4598.
[29] G.T. Hermanson, in: Bioconjugate Techniques, Academic Press, San Diego, 1996, pp. 137–168.

ELSEVIER

Vaccine 21 (2003) 2661–2666

Vaccine

www.elsevier.com/locate/vaccine

Protective immune response against methicillin resistant *Staphylococcus aureus* in a murine model using a DNA vaccine approach

José P.M. Senna [a], Daniela M. Roth [a], Jaim S. Oliveira [a], Denise C. Machado [b], Diógenes S. Santos [a,*]

[a] *Grupo de Microbiologia Molecular e Funcional, Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Cx.Postal 15005, Av. Bento Gonçalves, 9500, Porto Alegre RS, CEP 91501-970, Brazil*
[b] *Instituto de Pesquisas Biomédicas, Hospital São Lucas—Faculdade de Medicina, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil*

Received 17 April 2002; received in revised form 5 November 2002; accepted 18 November 2002

Abstract

Methicillin resistant *Staphylococcus aureus* (MRSA) are a major pathogen responsible for serious hospital infections worldwide. These bacteria are resistant to all beta-lactam antibiotics due to the production of an additional penicillin binding protein, the PBP2a, encoded by the *mecA* gene, which shows low affinity for this class of antibiotics. In this study, we cloned an internal region from the transpeptidase domain from the PBP2a into a mammalian expression vector, to be used as DNA vaccine in a Murine model. After three sets of DNA vaccination, the immune response represented by antibodies against a fragment of PBP2a was evaluated by enzyme linked immunosorbent assay (ELISA), showing a significant antibody response. The antibacterial effect of the DNA vaccine was evaluated by intraperitoneal immunization and challenge with a sublethal dose of MRSA for 7 days in mice. After the challenge, the number of bacteria from kidneys from immunized and non-immunized mice were determined. Kidneys from immunized mice had 1000 times less on bacteria than the positive controls (non-immunized mice). The response specificity indicates no effects against the normal PBPs from staphylococci and no effects against Gram positive rods from normal intestinal flora. Our results indicate that the immunization against the PBP2a from MRSA using a DNA vaccine approach could be used as a new strategy to efficiently fight these multiresistant bacteria.

Keywords: MRSA; DNA vaccine; PBP2a

1. Introduction

Methicillin resistant *Staphylococcus aureus* (MRSA) is an important pathogen responsible for serious nosocomial infections worldwide. This relatively virulent opportunist bacteria which is believed to be as virulent as methicillin sensitive *S. aureus* (MSSA) [1,2] is resistant to all beta-lactam antibiotics and the resistance is due to the production of an additional penicillin-binding protein (PBP) called PBP2′ or PBP2a encoded by the *mecA* gene, that shows a low affinity for beta-lactam antibiotics at concentrations that inhibit the beta-lactam-sensitive PBPs normally produced by *S. aureus* [3,4]. Despite the presence of the beta-lactam that inhibits the normal PBPs, MRSA can continue cell wall synthesis solely depending upon the uninhibited activity of PBP2a [5]. PBP2a is a multimodular class B PBP [6] with a molecular weight of 76 kDa encoded by the *mecA* gene located in the chromosome of MRSA [7]. Vancomycin, is a last resource antibiotic to treat MRSA infections. However, the report of vancomycin intermediate *S. aureus* (VISA) and very recent report of the isolation of a strain fully resistant to vancomycin (VRSA) [8,9], and the possible dissemination of untreatable staphylococcal infections indicates that it is necessary to search for alternative therapies to fight this bacteria. Several vaccine strategies against *S. aureus* have been described using bacterial structures—surface polysaccharide [10,11], whole cells or surface proteins—protein A [12] and fibronectin-binding protein [13] as target, but none has been fully successful until now.

DNA vaccines are considered a new approach to induce protective immunity. Since the publication of the first demonstration of in vivo protective efficacy in 1993 [14], this strategy has been applied against various pathogens [12–16]. DNA vaccines consist of a plasmid DNA encoding the proteins of interest with a strong eukaryotic promoter system. These vaccines could be administered by direct

* Corresponding author. Tel.: +55-51-3316-6072; fax: +55-51-3316-6234.
E-mail address: diogenes@dna.cbiot.ufrgs.br (D.S. Santos).

0264-410X/03/$ – see front matter © 2003 Published by Elsevier Science Ltd.
doi:10.1016/S0264-410X(02)00738-7

injection into the muscle leading to the production of protective antibodies and cellular immune responses [17,18].

In the present study, we evaluate the protective immune response in a murine model immunized with a DNA containing a fragment (249 bp) of the *mecA* gene and challenged with a MRSA strain. Penicillin binding proteins (PBPs) are located on the outer surface of Gram-positive bacteria to build the cell wall [19], and therefore the PBPs can be exposed to the host immune system.

2. Material and methods

2.1. Bacterial isolates

A MRSA clinical isolate (HPS-03) from the Hospital de Pronto Socorro, Porto Alegre, Brazil, was used in this study. The MRSA character was confirmed by PCR detection of *mecA* and *femA* genes [20]. A *S. aureus* American Type Culture Collection no. 25923, methicillin susceptible *S. aureus* was used as control.

2.2. Plasmid constructions

A 249 bp *mecA* fragment was amplified by PCR from HPS-03 MRSA. The foward primer, containing a restriction site for *Nhe* I and the reverse primer containing a restriction site for *Eco*RI were as follows: 5′-GCTAGCCAAGGAGGTCCAGCCATGAGTAACGAAGAA-3′ and 5′-TACGAATTCATATCTTGTAAC-3′. The DNA amplification was performed using *Taq* DNA polymerase (Cenbiot-Brazil) consisting of one cycle of 5 min at 94 °C, followed by 20 cycles of 1 min at 94 °C, 1 min at 52 °C, 1 min at 72 °C and one cycle of 5 min at 72 °C. The PCR product was digested with *Eco*RI and *Nhe* I and cloned into a pCI-Neo Mammalian Expression Vector (Promega) to use as DNA vaccine. The PCR product was also cloned into a *Sma* I site into pGEX-4T-2 Vector (Amersham Biosciences) to produce the recombinant peptide. The recombinant plasmids were electroporated into XL-1 blue *E. coli* cells and plated in Luria–Bertani agar containing ampicillin (50 μg/ml) and incubated overnight at 37 °C. Colonies was screened to detect the *mecA* fragment by PCR according to the conditions described above. The correct position of *mecA* fragment was determined by *Eco*RI digestion of pGEX-4T-2-*mecA*. The nucleotide sequence of the cloned *mecA* fragment were confirmed by Sanger's dideoxy mediated chain termination sequencing method [21] using the Thermo Sequenase Radiolabeled Terminator Cycle Sequencing kit (Amersham) of both recombinant plasmids, pCI-Neo-*mecA* and pGEX-4T-2-*mecA* fragment, showing to be identical to the reported sequence (Fig. 1).

2.3. Protein expression and purification

BL21 *E. coli* cells transfected with pGEX-4T-2-*mecA* fragment were grown in Luria–Bertani medium containing ampicillin (100 μg/ml) at 37 °C until exponential phase (optical density 0.6 at 600 nm), followed by induction with 0.1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG), and further incubation for 3 h. Cells were harvested and resuspended in phosphate buffered solution and submitted to sonication. After centrifugation, the recombinant *mecA* fragment expressed as GST fusion protein was purified from supernatant by using glutathione sepharose beads. The GST fusion protein was cleaved with thrombin to obtain the recombinant peptide (approximately 7 kDa) corresponding to the *mecA* fragment.

TASQGGPAMSNEEYNKLTDEKKEPLLNKFQITTSPGSTQKILTAM
IGLNN KTLDDKTSYKIDGKGWQKDKSWGGYNVTRYEFV

Fig. 1. Amino acid sequence corresponding to the amplified *mecA* region. The amino acids that was changed are marked in red. In blue, the conserved motif SXXK corresponding to the serine–protease motif.

2.4. Immunization procedures

The 8-week-old female BALB/c mice purchased from CEMIB (Unicamp, São Paulo, Brazil) were used in this experiment. Three days before the first immunization, bupivacain chloridrate 0.5% was injected at a dose of 2.5 μl/g of animal weight in the left quadriceps muscle. For DNA vaccination, three intramuscular injections with 10 μg of plasmid pCI-Neo-*mecA* or pCI-Neo only (negative control) diluted in 50 μl of sterile PBS were given at 2 weeks intervals into the left quadriceps muscle of mice. Blood was collected before (pre-immune serum) and after (immune serum) DNA vaccination (Fig. 2).

2.5. Analysis of the immune response

The levels of specific antibodies were evaluated by enzyme-linked immunosorbent assay. Plates were coated with 4.0 μg/ml of purified recombinant peptide diluted in 0.1 M carbonate/bicarbonate coating buffer (pH 9.6) and incubated overnight at 4 °C. Plates were washed three times with PBS containing 0.05% Tween 20, 200 μl of blocking buffer (PBS containing 5% non-fat dried milk) was added and the plates were incubated at 37 °C for 2 h. After removal of the blocking buffer, immune and control serum samples diluted 1:10 in blocking buffer were added to the plates and incubated at 37 °C for 2 h. After three washes

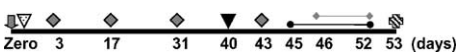


Fig. 2. Diagram showing the immunization schedule, sampling collection and MRSA challenge used to study the immune response and to evaluate the protective effect obtained by DNA vaccination: (⬇) bupivacain; (▽) collection of pre-immune blood; (◆) DNA vaccination (pCI-Neo or pCI-Neo-*mecA*); (▼) collection of immune sera; (●——●) oxacillin treatment period, (◆——◆) MRSA intraperitoneal infection with 2.0×10^6 cfu per day; (▨) mice sacrificed, sera and kidneys collection.

with PBS containing 0.05% Tween 20, 100 μl of anti-mouse polyvalent immunoglobulins peroxidase-conjugated (Dako) diluted 1:500, was added and the plates were incubated at 37 °C for 90 min. After three washes with PBS containing 0.05% Tween 20, 100 μl of substrate (0.4 mg of *o*-phenylenediamine hydrochloride diluted in 0.2 M citrate/phosphate buffer pH 5.0 containing 0.4 μl/ml of H_2O_2) was added and the plates were incubated at 37 °C for 15 min. Fifty microliters of a 2 M solution of sulfuric acid was added to quench the reaction. Absorbance was read at 490 nm wavelength with a microplate reader Benchmark (BioRad). A monoclonal α-GST antibody was used as positive control.

2.6. Analysis of DNA vaccine protection in mice

To assess the immune response against MRSA, mice immunized with pCI-Neo and pCI-Neo-*mecA* were challenged with MRSA (MRSA HPS-03). The lethal dose (5.0×10^8) and the median lethal dose (LD_{50}) (1.1×10^8) were calculated using the Reed and Muench method [18]. Mice received a sublethal dose of 2.0×10^6 cfu intraperitonially daily for 7 days. A non-immunized group receiving the same bacterial dose was included as positive control. Mice were immunized with pCI-Neo or pCI-Neo-*mec*A were subdivided in groups. Groups I and II received plasmid without *mecA* fragment and challenged with MRSA. Group I was treated with oxacillin (75 mg/kg) administered three times per day during the challenge period in the quadriceps muscle. Groups III–VI received plasmid containing *mecA* fragment. Groups IV and V were treated with the same dose of oxacillin as group I. Group V and VI were challenged with MRSA. Animals were sacrificed and the kidneys were excised, rinsed in sterile PBS and weighted.

2.7. Bacterial load

Each kidney was homogenized and diluted 1:10, 1:100 and 1:1000 in 1 μl of sterile LB broth. The 100 μl of each dilution was plated into LB agar plates containing 50 μg/ml of oxacillin and incubated at 37 °C. After 24 h the number of grown colonies was counted. DNA extraction was performed followed by PCR to detect the presence of *femA* and the *mecA* genes as earlier described [22].

2.8. Flow cytometry

Blood cells (150 μl) were stained with Fluorescein isothiocyanate (FITC)—conjugated hamster anti-mouse CD3e monoclonal antibody (PharMingen International, catalogue no. 01084A) and R-Phycoerythrin (R-PE)—conjugated rat anti-mouse CD8a (Ly-2) monoclonal antibody (PharMingen International, catalogue no. 01045A) diluted 1:5 in PBS buffer pH 7.2 to quantify the subpopulation of mature T lymphocytes MHC class I restricted T cells and with Fluorescein isothiocyanate—conjugated hamster anti-mouse CD3e monoclonal antibody (PharMingen International, catalogue no. 01084A) and R-Phycoerythrin—conjugated rat anti-mouse CD4 (L3T4) monoclonal antibody (PharMingen International, catalogue no. 09425A) diluted 1:5 in PBS buffer pH 7.2 to quantify the subpopulation of mature T lymphocytes MHC class II restricted T cells. After immunofluorescence staining for 20 min at room temperature, whole blood were lysed with FACS lysing solution (Becton and Dickinson, catalogue no. 349202) and the stained cells were analyzed by flow cytometry on a FACS Calibur apparatus (Becton and Dickinson). Data were analyzed with Cell Quest FACS analysis software (Becton and Dickinson).

2.9. Specificity of the immune response against PBP2a

(i) Gram stain of mice faeces before and after pCI-Neo-*mecA* immunization were proceeded to investigate the effect of other PBPs from Gram positive rods present in the intestinal normal flora.
(ii) To evaluate the effect of the normal PBPs present in *S. aureus*, one mouse immunized with pCI-Neo-*mecA* was challenged, according to earlier described protocol, with the methicillin susceptible *S. aureus* (ATCC 25923), and did not receive the oxacillin treatment. This mouse was included as a control of antibody specificity. The bacterial quantitation was proceeded in LB agar plates without oxacillin as described above.

2.10. Statistics

A Student's *t*-test was applied to determine the significance of the differences between vaccinated and control groups [23].

3. Results

3.1. Protection obtained in immunized mice and bacterial load

All the animals used in this study survived after the MRSA challenge. The number of bacteria in the kidneys from mice vaccinated or not were as follows: positive controls was 1.0 up to 9.3×10^4 cfu per kidney; from mice vaccinated with pCI-Neo and treated with oxacillin (Group I) was 1.3×10^3 up to 2.4×10^4 cfu per kidney; mice vaccinated with pCI-Neo and not treated with oxacillin (Group II) was 7.0×10 up to 2.0×10^4 cfu per kidney. The bacteria in kidneys from mice vaccinated with pCI-Neo-*mecA* and treated with oxacillin (Group V) varied from 0 (60% did not present bacteria in the kidneys) to 7.0×10^4 cfu per kidney. Mice vaccinated with pCI-Neo-*mecA* and not treated with oxacillin (Group VI) ranged from 0 (80% did not presented bacteria in the kidneys) to 1.3×10^2 cfu per kidney. The difference between

Table 1
The number of bacteria/kidney by group of mice challenged

Control group	Group I	Group II	Group V	Group VI
1.0×10^4	1.3×10^3	7.0×10	0	0
3.3×10^4	3.0×10^3	2.0×10^2	0	0
4.9×10^4	3.7×10^3	3.0×10^2	0	0
9.3×10^4	2.4×10^4	2.0×10^4	0	0
			0	0
			0	0
			1.0×10	0
			1.0×10	0
			8.0×10^2	3.0×10
			7.0×10^4	1.3×10^2

Group I: pCI-Neo + Oxa+ MRSA; Group II: pCI-Neo + MRSA; Group V: pCI-Neo-*mecA* + Oxa + MRSA; Group VI: pCI-Neo-*mecA* + MRSA.

the positive control and the vaccinated group (Group VI) was statistically significant ($P < 0.015$), as well as the difference between Groups II and VI ($P < 0.015$). The kidneys from groups that were not challenged (Groups III and IV) did not show any bacteria growth. The number of bacteria/kidney in each animal analyzed is shown in Table 1.

3.2. Kidney morphologic appearance from immunized and non-immunized mice

Morphologic differences could be observed among kidneys from immunized mice and the positive controls. Kidneys from vaccinated mice have the same appearence as the normal kidneys, and kidneys from the positive control group present hypertrophy and were colorless (data not shown).

3.3. Evaluation of the anti-PBP2a antibody production by ELISA

The specific antibody production, was measured by ELISA, with sera obtained before (pre-immune serum) and after challenge with MRSA from mice immunized with recombinant plasmid containing or not the *mecA* fragment in the presence or absence of oxacillin. The vaccinated group that received oxacillin and was challenged with MRSA (Group V) produced three times more anti-PBP2a specific antibody than the control group (Group I), and the vaccinated group that did not receive oxacillin and was challenged with MRSA (Group VI) produced an antibody response 1.9 times higher than Group I.

3.4. Evaluation of MHC classes I and II restricted T cells subpopulations by flow cytometry after vaccination

Humoral and cellular response evaluated by the presence of $CD3^+/CD4^+$ (MHC class II) and $CD3^+/CD8^+$ (MHC class I) T cells subpopulations, respectively, did not differ between vaccinated and non-vaccinated mice. Moreover, we were not able to detect activation of specific T cells subpopulations within the vaccinated groups that received different treatments.

3.5. Evaluation of the specificity of anti-PBP2a immune response

Differences on Gram stain of mice faeces from non-immunized and immunized mice were not observed. Gram-positive cocci and rods were present in both groups (data not shown). Vaccinated pCI-Neo-*mecA* mouse challenged with methicillin susceptible *S. aureus* (*mecA* negative strain) showed 1.2×10^5 cfu per kidney, meaning that the immune response obtained with pCI-Neo-*mecA* vaccination was specific against the PBP2a.

4. Discussion

Since the first use of antibiotics, we have observed a fight between the drugs and methicillin resistant *S. aureus*. Although its clinical virulence is approximately equal to MSSA [1,2], it has a surprising capability to respond to the selective pressure exerted by antimicrobial. This bacterium is suppose to become resistant to all drugs available, very soon. Recently, the US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, reported the isolation and characterization of a MRSA strain completely resistant to vancomycin. This particular strain was isolated from the foot ulcer and a catheter of a patient in Michigan. Accordingly, the Michigan VRSA acquired its resistance genes from *Enterococcus faecalis* fully vancomycin resistant [8]. The antibiotic therapy shows no signs to eradicate MRSA from nosocomial infections. In this work, we were able to develop and test in an animal model a new strategy to fight MRSA by activation of the immune system with a DNA vaccine. Different from other vaccine strategies against *S. aureus*, our target was the PBP2a, a protein responsible for the intrinsic resistance to all beta-lactam antibiotics in MRSA. This enzyme, is composed by a spanning region, responsible for the attachment to the cytoplasmic membrane, a non-penicillin binding domain of unknown function and a transpeptidase domain, being classified by Goffin and Ghuysen as a multimodular class B penicillin binding protein [6]. Due to the PBP2a location, which is outside to the cytoplasmic membrane [19], the PBP2a could be accessible to serum antibodies. The DNA vaccine approach was chosen due to its characteristics such as simple to work, stable and capable to elicit immune response, as reported by others [12,13]. We used an internal PBP2a fragment containing 249 bp and comprising the serino-protease domain (STQK) as antigen and our choice proved to be correct. We were able to induce anti-PBP2a specific antibody production and an effective protection, as shown by ELISA and challenging immunized mice with MRSA. The reduction of the bacterial number on kidneys from immunized mice was significant. Kidneys from 70% of the immunized mice did not shown the presence of bacteria after MRSA challenge. Since the antibodies raised against a fragment of PBP2a containing the transpeptidase active site would block the final step of

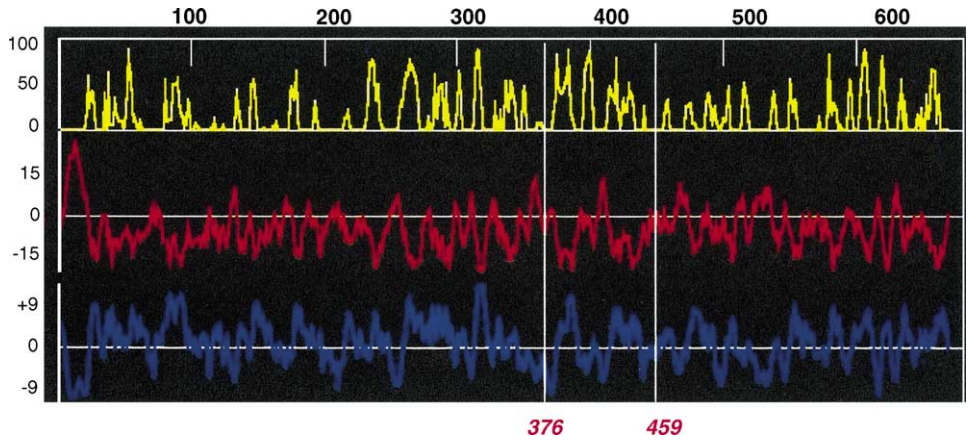


Fig. 3. Characteristics of hydropaticity and antigenicity from PBP2a predicted by the ANTHEPROT software (http://pbil.ibcp.fr/NPSA). Aminoacids positions are represented on the top. The positions 376 and 459 are indicated by horizontal white lines that corresponds to the peptide region used in this study. Yellow: antigencity characteristics, red: hydrophobicity characteristics and blue: hydrophilic characteristics.

cell wall assembly one should expect the results obtained in the present study. Kidneys from mice vaccinated with the pCI-Neo without the fragment of *mecA*, had a slight reduction of the bacterial number, probably due to the presence of the CpG motifs (approximately 10% of the pCI-Neo vector consists of the CpG motifs) [24]. Moreover, analysis using the ANTHEPROT software [25] predicts the antigenicity of the PBP2a fragment (Fig. 3) used in this study corroborate its use as immunogen. The specificity of immune response obtained against PBP2a was an important factor to be evaluated, since an immune response against PBPs from other then the Gram-positive rods that are present in the intestinal normal flora may cause undesirable side effects. In addition, it was not observed any difference between Gram stain of faeces from immunized and non-immunized mice. On other hand, mice immunized (pCI-Neo-*mecA* fragment) and challenged with a *mecA* negative *S. aureus* (MSSA) shows that the immune response obtained against MRSA is specific against PBP2a, not affecting the normal PBPs present in staphylococci that do not possess the *mecA* gene. The oxacillin administration (a beta-lactam antibiotic that are not affected by beta-lactamase) was used aiming to block the normal PBPs and to allow that only the PBP2a would assemble the peptidoglycan conferring a better PBP2a exposition to the antibody action. However, the protection obtained in mice that received the DNA vaccine was similar even in the absence of oxacillin. This observation led us to suggest that in MRSA, the PBP2a may have a more important role than the other PBPs to assemble the peptidoglycan cell wall. Ohwada et al. [26], using the whole *mecA* as antigen obtained protection in Balb/C mice infected with MRSA. The number of bacteria in the kidneys obtained from mice vaccinated with a plasmid expressing *mecA* was very low compared with mice vaccinated with a plasmid without the *mecA* sequence. Together these two observations suggest that the fragment of DNA containing the active site of the PBP2a or the entire *mecA* cloned sequence may be eligible as a candidate for DNA vaccination to fight MRSA.

In conclusion, we demonstrate that a fragment of PBP2a coding region, administered by DNA vaccination can elicits an immune response against MRSA as shown by the a strong antibacterial protection obtained in a DNA-vaccinated murine model infected with MRSA.

Acknowledgements

J.P.M. Senna and D.M. Roth, contributed equally to this work. This study was supported in part by Grant from the Brazilian Ministry of Health, Financiadora de Estudos e Projetos (FINEP) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). We acknowledge the expertise of Gaby Renard in the revision of this manuscript.

References

[1] Selvey LA, Whitby M, Johnson B. Nosocomial methicillin resistant *Staphylococcus aureus* bacteremia: is it any worse than nosocomial methicillin sensitive *Staphylococcus aureus* bacteremia. Infect Control Hosp Epidemiol 2000;10:645–8.

[2] French GL, Cheng AF, et al. Hong Kong strains of methicillin resistant and methicillin sensitive *Staphylococcus aureus* have similar virulence. J Hosp Infect 1990;15(2):117–25.

[3] Chambers HF. Methicillin resistance in staphylococci: molecular and biochemical basis and clinical implications. Clin Microbiol Rev 1997;10:781–91.

[4] Lim D, Strynadka CJ. Structural basis for the β-lactam resistance of PBP2a from methicillin-resistant *Staphylococcus aureus*. Nat Struct Biol 2002;9:870–6.

[5] Katayama Y, Ito T, Hiramatsu K. A new class of genetic element, *Staphylococcus* cassette chromosome mec, encodes methicillin resistance in *Staphylococcus aureus*. Antimicrob Agents Chemother 2000;44:1549–55.

[6] Goffin C, Ghuysen JM. Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs. Microbiol Molec Biol Rev 1998;62:1079–93.

[7] Berger-Bächi B. Expression of resistance to methicillin. Trends Microbiol 1994;2:389–93.
[8] Pearson, H. Bacteria defies last-resort antibiotic. Nature 2002:469.
[9] Tenover FC, Lancaster MV, Hill BC, Stedward CD, Stocker SA, Hancock GA, et al. Characterization of staphylococci with reduced susceptibilities to vancomycin and other glycopeptides. J Clin Microbiol 1998;36:1020–7.
[10] McKenney D, Pouliot KL, Wang Y, Murthy V, et al. Broadly protective vaccine for *Staphylococcus aureus* based on an in vivo-expressed antigen. Science 1999;284:1523–7.
[11] Lee JC, Perez NE, Hopkins CA, Pier GB. Purified capsular polyssaccharide-induced immunity to *Staphylococcus aureus* infection. J Infect Dis 1988;157:723–31.
[12] Geenberg D, Bayer A, Cheung A, Ward J. Protective efficacy of protein-A specific antibody against bacteriemic infection due to *Staphylococcus aureus* in an infant rat model. Infect Immun 1989;57:1113–20.
[13] Ulmer JB, Donnely JJ, Parker SE, et al. Heterologous protection against influenza by injection of DNA encoding a viral protein. Science 1993;259:1745–9.
[14] Mamo W, Jonsson P, Flock JI, Lindberg M, et al. Vaccination against *Staphylococcus aureus* mastitis: immunological response of mice vaccinated with fibronectin-binding protein (FnBP-A) to challenge with *S. aureus*. Vaccine 1994;12:988–92.
[15] Ulmer JB, Donnelly JJ, Liu MA. DNA vaccines promising: a new approach to inducing protective immunity. ASM News 1996;62:476–9.
[16] Lowrie DB, Tascon RE, Silva CL, et al. Therapy of tuberculosis in mice by DNA vaccination. Nature 1999;400:269–71.
[17] Shiver JW, Ulmer JB, Donnely JJ, Liu MA. Humoral and cellular immunities elicited by DNA vaccines: application to the human immunodeficiency virus and influenza. Adv Drug Delivery Rev 1996;21:19–31.
[18] Wild J, Gruner B, Metzger K, et al. Polyvalent vaccination against hepatitis B surface and core antigen using a dicistronic expression plasmid. Vaccine 1998;16:353–60.
[19] Navarre WW, Schneewind O. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. Microbiol Molec Biol Rev 1999;63:174–229.
[20] Vannuffel P, Gigi J, Ezzedine H, et al. Specific detection of methicillin-resistant *Staphylococcus aureus* species by multiplex-PCR. J Clin Microbio 1995;33:2864–7.
[21] Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbour (NY): Cold Spring Harbour Laboratory; 1989.
[22] Reed LJ, Muench HA. Simple method for estimating fifty per cent endpoints. Am J Hyg 1938;27:493–7.
[23] Leonard JM. The arithmetic of decision-making. London: The English University Press; 1971.
[24] Krieg AM. Immune effects and mechanisms of action of CpG motifs. Vaccine 2001;19:618–22.
[25] Deléage G, Combet C, Blanchet C, Geourjon C. ANTHEPROT: an integrated protein sequence analysis software with client/server capabilities. Comput Biol Med 2001;31:259–67.
[26] Ohwada A, Sekiya M, Hanaki H. DNA vaccination by *mecA* sequence evokes na antibacterial immune response against methicillin-resistant *Staphylococcus aureus*. J Antimicr Chemother 1999;44:767–74.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 312 (2003) 608–614
BBRC
www.elsevier.com/locate/ybbrc

Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*

José Henrique Pereira,[a] Fernanda Canduri,[a,b,*] Jaim Simões de Oliveira,[c] Nelson José Freitas da Silveira,[a] Luiz Augusto Basso,[c] Mário Sérgio Palma,[b,d] Walter Filgueira de Azevedo Jr.,[a,b,*] and Diógenes Santiago Santos[e,*]

[a] *Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil*
[b] *Center for Applied Toxinology, Instituto Butantan, São Paulo, SP 05503-900, Brazil*
[c] *Rede Brasileira de Pesquisa de Pesquisas em Tuberculose Grupo de Microbiologia Molecular e Funcional, Departamento de Biologia Molecular e Biotecnologia, UFRGS, Porto Alegre, RS 91501-970, Brazil*
[d] *Laboratory of Structural Biology and Zoochemistry, CEIS/Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil*
[e] *Faculdade de Farmácia/Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil*

Received 23 October 2003

Abstract

The shikimate pathway is an attractive target for herbicides and antimicrobial agent development because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals. Homologues to enzymes in the shikimate pathway have been identified in the genome sequence of *Mycobacterium tuberculosis*. Among them, the EPSP synthase was proposed to be present by sequence homology. Accordingly, in order to pave the way for structural and functional efforts towards anti-mycobacterial agent development, here we describe the molecular modeling of 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase isolated from *M. tuberculosis* that should provide a structural framework on which the design of specific inhibitors may be based on. Significant differences in the relative orientation of the domains in the two models result in "open" and "closed" conformations. The possible relevance of this structural transition in the ligand biding is discussed.

Keywords: EPSP synthase; Bioinformatics; Molecular modeling; *Mycobacterium tuberculosis*

Tuberculosis (TB) remains one of the most deadly infectious diseases in the world. It is estimated that approximately 1 billion individuals are infected with latent TB. Tuberculosis has made a steady global comeback in the late 1980s and now kills more than 2 million people each year worldwide, according to the World Health Organization (WHO) which in 1993 declared tuberculosis to be a global emergence [1].

Mycobacterium tuberculosis is by far the biggest killer among infectious agents, accounting for 7% of all deaths and 26% of all avoidable deaths. In impoverished third world countries where 95% of cases and 98% of deaths occur, 70–80% of TB cases are in the economically most productive 15–50 year age bracket [1,2].

The treatment of tuberculosis is a special problem in the field of chemotherapy. Many of the drugs employed to treat the disease are used only for treating infections caused by mycobacteria. Treatment of the active case of TB always includes simultaneous therapy with two or more of the frontline drugs: isoniazid, ethambutol, rifampicin, and streptomycin which are used to decrease the rate of emergence of resistant strains as well to increase the antibacterial effect [3–5].

Recent outbreaks of tuberculosis caused by multidrug-resistant (MDR) strains, mainly in individuals infected with HIV, have created a scaring element to the scenario and also created a worldwide interest in expanding current programs of development of new drugs

* Corresponding authors. Fax: +55172212247.
E-mail addresses: canduri@df.ibilce.unesp.br (F. Canduri), walterfa@df.unesp.br (W.F. Azevedo Jr.), diogenes@pucrs.br (D.S. Santos).

0006-291X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2003.10.175

different in kinds from existing ones and based on the principle of selective toxicity on enzymes or structure of the bacillus to both treat *M. tuberculosis* strains resistant to existing drugs and shorten the duration of short-course treatment to improve patient compliance [4,6].

The shikimate pathway is an attractive target for the development of herbicides and antimicrobial agents because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals [7]. In mycobacteria, the shikimate pathway leads to the biosynthesis of precursors for the synthesis of aromatic amino acids, naphthoquinones, menaquinones, and mycobactin [8]. Homologues to enzymes in the shikimate pathway have been identified in the complete genome sequence of *M. tuberculosis* H37Rv strain [9]. Among them, 5-enolpyruvylshikimate-3-phosphate (EPSP) synthase encoded by aroA gene, which catalyzes the transfer of the enolpyruvyl moiety from phosphoenolpyruvate (PEP) and inorganic phosphate. A valuable lead compound in the search of new inhibitors is glyphosate, which has proven to be potent and specific inhibitor of EPSP synthase [10]. Glyphosate is successfully used as a herbicide, being the active ingredient of the widely used weed control agent Roundup Ready, and was recently shown to inhibit the growth of the pathogenic parasities *Plasmodium falciparum*, *Toxoplasma gondii*, and *Cryptosporidium parvum* [11,12].

The present paper describes the two molecular models of *M. tuberculosis* EPSP synthase, one without any ligand and another in complex with 3-phosphoshikimate (S3P) and glyphosate. The homology modeling was performed using the crystallographic structures of EPSP synthase from *Escherichia coli* [12,13], as templates. The EPSP synthase has been cloned, sequenced, and overexpressed in soluble and functional form [14], thus allowing enzymological studies to be performed. The results presented here should provide a three-dimensional model of EPSP synthase to both guide enzymological studies and aid in the design of specific inhibitors.

Methods

Molecular modeling. For modeling of the EPSP synthase we used restrained-based modeling implemented in the program MODELLER [15]. This program is an automated approach to comparative modeling by satisfaction of spatial restraints [16–18]. The modeling procedure begins with an alignment of the sequence to be modeled (target) with related known three-dimensional structures (templates). This alignment is usually the input to the program. The output is a three-dimensional model for the target sequence containing all main-chain and side-chain non-hydrogen atoms.

The degree of primary sequence identity between MtEPSP synthase and EPSP synthase from *E. coli* indicates that the crystallographic structures of EcEPSP are good models to be used as templates for MtESPS synthase. The EPSP synthase isolated from *E. coli* showed two conformation states, open (unliganded structure) and closed (liganded structure) [12]. Models in the two conformations were generated, using the α-carbon atomic coordinates from 1EPSP [13] to generate the open structure and the atomic coordinates from 1G6S [12] to generate the closed structure. The atomic coordinates of all waters were removed from the EPSP structure. The 3-phosphoshikimate and glyphosate of the template were kept in the structure for the closed

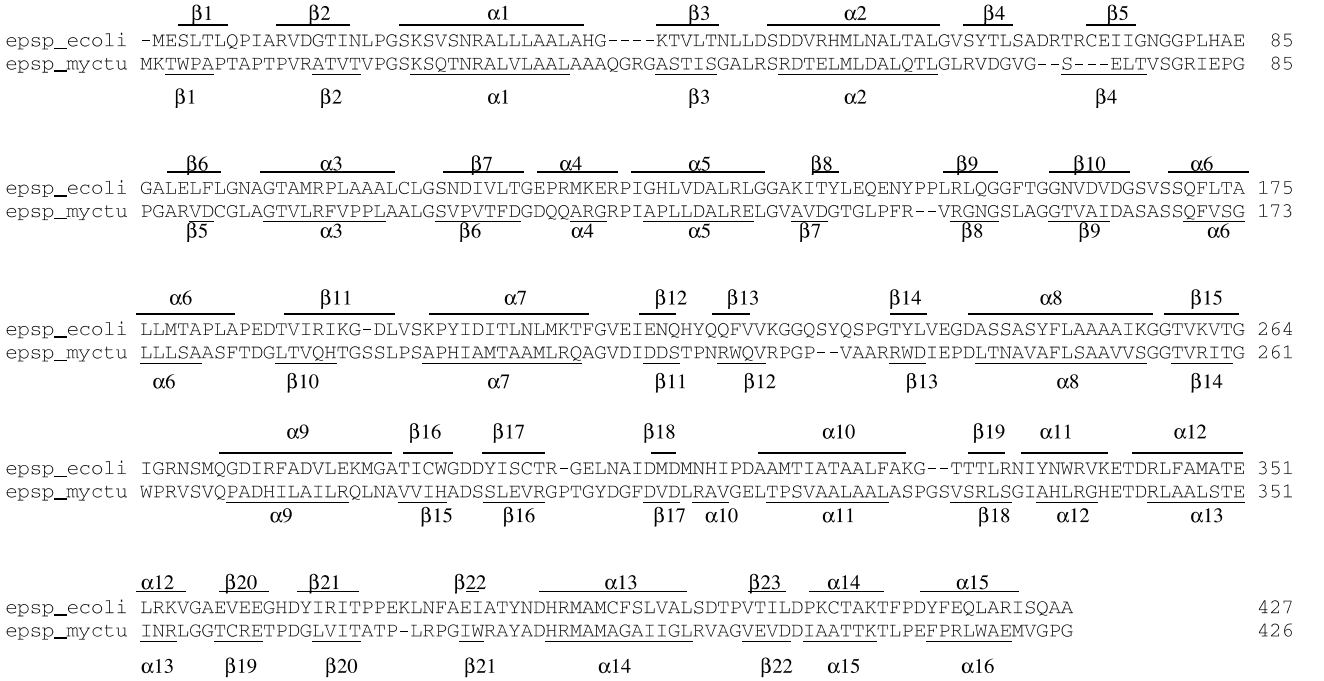


Fig. 1. The sequence alignment of EcEPSP and MtEPSP indicating the secondary structural elements. The sequence MtEPSP shows 31% of identity with the sequence of EcEPSP. The alignment was performed with the program CLUSTAL V [37].

conformation. Next, the spatial restraints and CHARMM energy terms enforcing proper stereochemistry [19] were combined into an objective function. Finally, the model is obtained by optimizing the objective function in Cartesian space. The optimization is carried out by the use of the variable target function method [20] employing methods of conjugate gradients. Several slightly different models can be calculated by varying the initial structure. A total of 1000 models were generated for each EPSP synthase conformation state, and the final models were selected based on stereochemical quality. The optimization of the complex was carried out by the use of the variable target function method [15] employing methods of conjugate gradients and molecular dynamics with simulated annealing. All modeling process was performed on a Beowulf cluster, with 16 nodes (B16/AMD Athlon 1800+; BioComp, São José do Rio Preto, SP, Brazil).

Analysis of the model. The overall stereochemical quality of the final model for EPSP complex was assessed by the program PROCHECK [21]. The cutoff for hydrogen bonds and salt bridges was 3.6 Å.

Results and discussion

Primary sequence comparison

The sequence alignment of EcEPSP (template) and MtEPSP (target) is shown in Fig. 1. The secondary structural elements are indicated in the figure. The sequence MtEPSP shows 31% identity with the sequence of EcEPSP.

Quality of the model

The Ramachandran plot for the 2 EcEPSP structures was generated in order to compare the overall stereochemical quality of MtEPSP model against EPSP

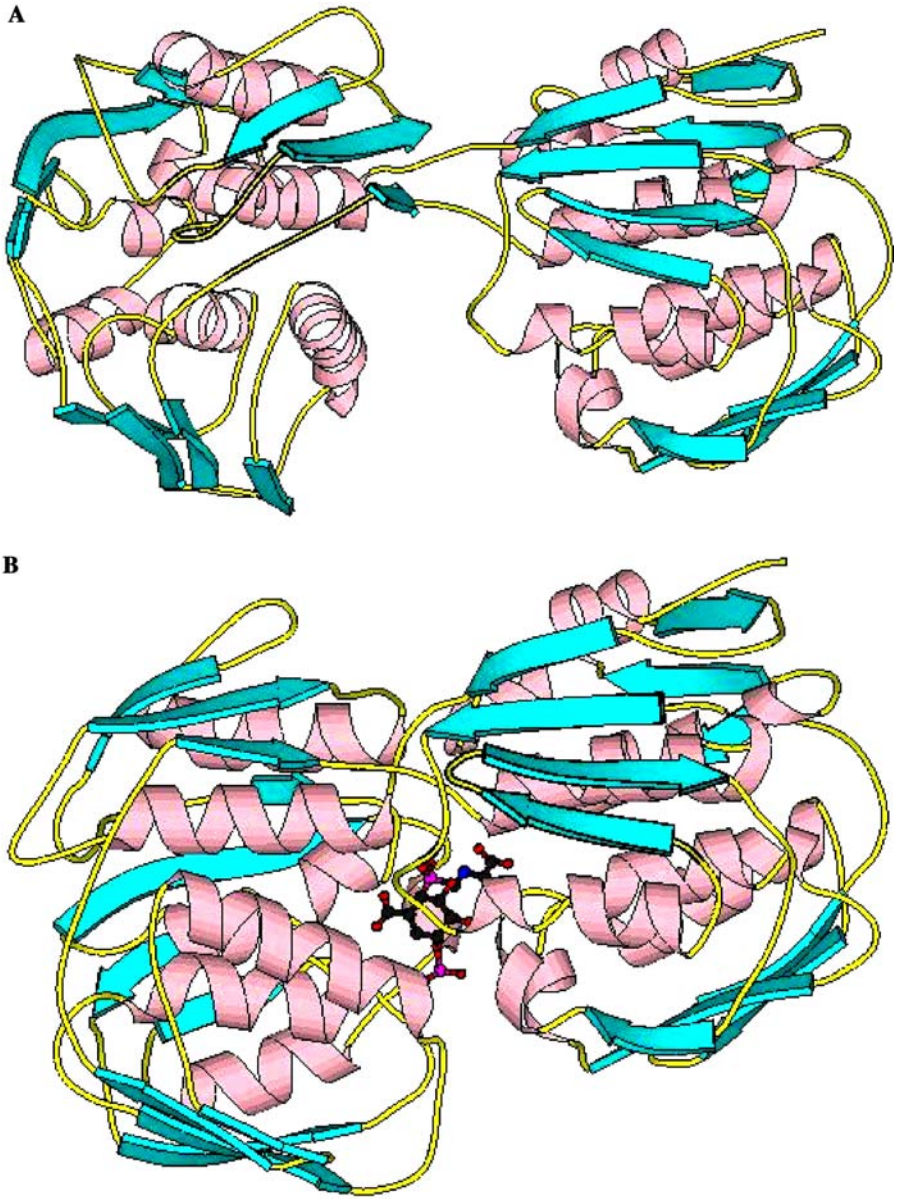


Fig. 2. Ribbon diagram of the MtEPSP structure in the open (A) and closed (B) conformations generated by Molscript [38].

structures solved by biocrystallography. Analysis of the Ramachandran plot of the MtEPSP models shows that 84.8% of the residues lie in the most favorable regions and the remaining 10.9% in the additional allowed regions for the closed structure, and 91.1% lie in the allowed regions for the open structure. The same analysis for two crystallographic EcEPSP structures present 91.2% of residues in the most favorable, 8.5% additional allowed regions, and 0.3% generously allowed regions. The overall rating for the MtEPSP model is slightly poorer than the one obtained for the two structures of EcEPSP. However, it has over 90% of the residues in the most favorable regions.

Overall description

Figs. 2A and B show the secondary structure of both models of MtEPSP synthase. EPSP synthase is an α/β protein consisting of a mixed β sheet surrounded by α helices. The structure folds into two similar domains, which are approximately hemispheric, each with a radius of about 21 Å. The domains are linked by two crossover chain segments with both the amino and carboxyl termini of the protein in the lower domain. The two flat surfaces of the hemispheres, which in projection form a "V," are almost normal and accommodate the amino termini of the six helices in each domain. The helical macrodipolar effects create a potential well that facilitates the binding of glyphosate and S3P [13]. The domains are composed of three copies of a $\beta\alpha\beta\alpha\beta\beta$-folding unit. The four-stranded β-sheet structures contain both parallel and antiparallel strands and the helices are parallel. In the closed structure the glyphosate and S3P molecules are bound in the interdomain cleft, promoting the closure of the structure.

Molecular hinge

Both crystal structures of EcEPSP synthase and the molecular models of MtEPSP synthase reveal the presence of two domains in which the domains, designated as domains A and B, undergo a translation and rotation from the open to the closed conformation. The center of mass of the two domains moves 4.6 Å from the open to the closed structure, for MtEPSP synthase. In the open and closed structures of MtEPSP synthase an angular difference at the interface of approximately 69° between the two domains results in different conformations, as illustrated in Fig. 3. Superposition of Cα atoms of the constituent domains of each structure results in rms differences of 0.20 and 0.39 Å for EcEPSP and MtEPSP synthases, respectively. The individual domains from both structures were superposed using the Cα atoms in the four major α-helices (residues 240–252, 270–280, 343–355, and 385–395). The low rmsd values, observed in the superpositions, indicate that the alternative conformation of the two-domain structures is due to relative motions of structurally conserved domains. However, the difference in conformation suggests a

Fig. 3. Molecular hinge. α-Carbon traces of the superimposed structures of MtEPSP synthase from the open conformation (dark line) and closed conformation (light line).

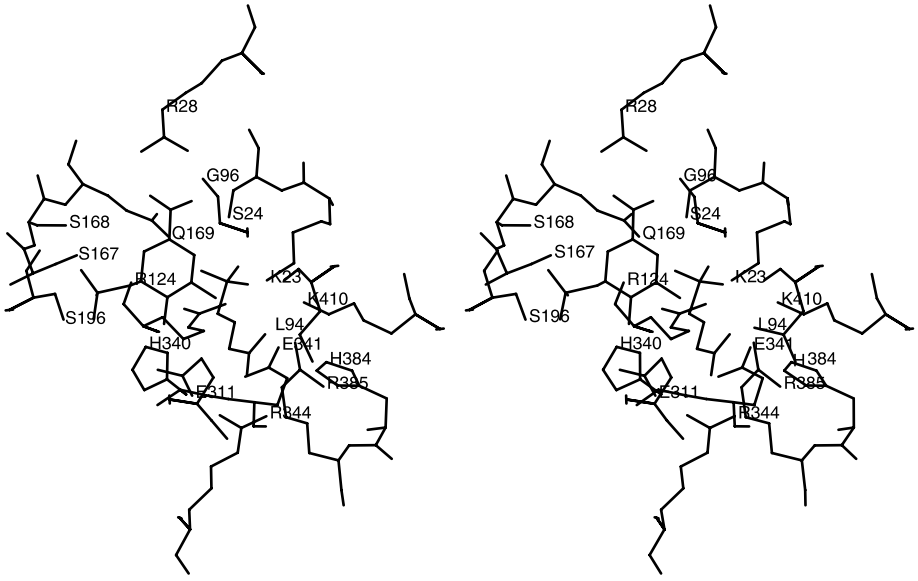


Fig. 4. Stereo view of the interdomain interface region illustrating the binding of glyphosate and S3P in the MtEPSP synthase structure.

degree of flexibility in the interdomain contact at the interface region, which acts as a hinge between the two monomers resulting in "open" and "closed" forms of the EPSP synthase structure. Similar molecular hinge is observed in PLA_2 structures [22–24].

3-Phosphoshikimate and glyphosate binding

Fig. 4 shows the active site of MtEPSP synthase. The glyphosate and S3P are close to each other and the 5-hydroxyl group of S3P is hydrogen bonded to the nitrogen atom of glyphosate with a distance of 2.82 Å. Analysis of the molecular surface of the MtEPSP synthase indicates that both ligands are buried in between the two domains, with a contact area of 29 and 134 $Å^2$ for glyphosate and 3-phosphoshikimate, respectively. The electrostatic potential surface of the MtEPSP synthase complexed with glyphosate and S3P calculated with GRASP [25] indicates the presence of some charge complementarity between ligands and enzyme (figure not shown), nevertheless most of the residues in the binding pocket are hydrophobic in all structures (see Fig. 4).

The analysis of the MtEPSP synthase complex indicates a total of 22 and 15 hydrogen bonds between the enzyme and glyphosate and S3P, respectively. Tables 1 and 2 show the intermolecular hydrogen bonds between

Table 1
Hydrogen bonds between EPSP:glyphosate

Hydrogen bonds between active site and inhibitor			Distance (Å)
Glyphosate	EPSP		
O4	Lys23	NZ	2.98
O1	Lys23	NZ	2.84
N1	Lys23	NZ	3.46
O3	Leu94	O	3.10
O2	Gly96	N	3.53
O3	Gly96	N	2.87
O3	Arg124	NH1	2.87
O2	Arg124	NH2	2.81
O3	Arg124	NH2	3.60
O2	Gln169	NE1	3.58
O2	Gln169	NE2	2.88
O5	Glu311	OE1	2.96
N1	Glu341	OE1	3.35
N1	Glu341	OE2	2.88
O4	Glu341	OE2	3.59
O5	Arg344	NH1	2.81
O5	Arg344	NH2	3.02
O4	His384	NE2	3.49
O4	Arg385	NE	3.59
O5	Arg385	NE	2.75
O4	Arg385	NH2	3.12
O3	Lys410	NZ	2.95

Table 2
Hydrogen bonds between EPSP:shikimate 3-phosphate

Hydrogen bonds between active site and ligand			Distance (Å)
Shikimate 3-phosphate	EPSP		
O3	Lys23	NZ	2.81
O5	Ser24	OG	2.70
O5	Arg28	NH1	2.86
O4	Arg28	NH2	2.78
O5	Arg28	NH2	3.60
O6	Ser167	OG	3.40
O8	Ser167	OG	2.70
O6	Ser168	OG	2.69
O6	Ser168	N	2.84
O1	Gln169	NE2	3.49
O6	Ser196	OG	3.59
O7	Ser196	OG	2.69
O2	Glu311	OE2	2.65
O2	His340	NE2	3.14
O1	His340	NE2	3.57

the enzyme and the ligands. Most of the hydrogen bonds between glyphosate and EPSP synthase involve residues Lys23, Arg124, Glu341, and Arg385, the same residues involved in hydrogen bonds between EcEPSP and glyphosate. Furthermore, the alignment of 69 EPSP synthase sequences indicates that the main residues involved in intermolecular hydrogen bonds, between the glyphosate and the enzyme, are conserved in all sequences (figure not shown). Such observation suggests that inhibitors derived from glyphosate will be able to inhibit most or even all EPSP synthases, since specificity and affinity between enzyme and its inhibitor depend on directional hydrogen bonds and ionic interactions, as well as on shape complementarity of the contact surfaces of both partners [26–36]. Further inhibition experiments may confirm this prediction.

Acknowledgments

This work was supported by Grants from FAPESP (SMOLBNet, Proc. 01/07532-0), CNPq, CAPES and Instituto do Milênio (CNPq-MCT) to DSS and LAB. WFA (CNPq, 300851/98-7) and MSP (CNPq, 500079/90-0) are researchers for the National Research Council.

References

[1] World Health Organization. Global Tuberculosis Control. WHO Report 2001. Geneva, Switzerland, WHO/CDS/TB/2001.287.
[2] N.E. Dunlap, J. Bass, P. Fujiwara, P. Hopewell, C.R. Horsburgh Jr., M. Salfinger, P.M. Simione, Diagnostic standards and classification of tuberculosis in adults and children, Am. J. Respir. Crit. Care Med. 161 (2000) 1376–1395.
[3] J.D. McKinney, W.R. Jacobs Jr., B.R. Bloom, Persisting problems in tuberculosis, in: R.M. Krause (Ed.), Emerging Infections, Academic Press, New York, 1998, pp. 51–146.
[4] C.J.L. Murray, Issues in operational, social, and economic research on tuberculosis, in: B.R. Bloom (Ed.), Tuberculosis: Pathogenesis, Protection, and Control, ASM Press, Washington, DC, 1994, pp. 583–622.
[5] B.R. Bloom, C.J.L. Murray, Tuberculosis: commentary on a reemergent killer, Science 257 (1992) 1055–1064.
[6] L.A. Basso, J.S. Blanchard, Resistance to antitubercular drugs, Adv. Exp. Med. Biol. 456 (1998) 115–144.
[7] R. Bentley, The shikimate pathway—A metabolic tree with many branches, Crit. Rev. Biochem. Mol. Biol. 25 (1990) 307–384.
[8] C. Ratledge, Nutrition, growth and metabolism, in: C. Ratledge, J.L. Stanford (Eds.), The Biology of the Mycobacteria, vol. 1, Academic Press, London, 1982, pp. 185–271.
[9] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B.G. Barrel, Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, Nature 393 (1998) 537–544.
[10] H.C. Steinrücken, N. Amrhein, The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid-3-phosphate synthase, Biochem. Biophys. Res. Commun. 94 (4) (1980) 1207–1212.
[11] F. Roberts, C.W. Roberts, J.J. Johnson, D.E. Kyle, T. Krell, J.R. Coggins, G.H. Coombs, W.K. Milhous, S. Tzipori, D.J. Ferguson, D. Chakrabarti, R. McLeod, Evidence for the shikimate pathway in apicomplexan parasites, Nature (London) 395 (6699) (1998) 801–805.
[12] E. Schönbrunn, S. Eschenburg, W.A. Shuttleworth, J.V. Schloss, N. Amrhein, J.N.S. Evans, W. Kabsch, Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail, Proc. Natl. Acad. Sci. USA 98 (2001) 1376–1380.
[13] W.C. Stallings, S.S. Abdel-Meguid, L.W. Lim, H.-S. Shieh, H.E. Dayringer, N.K. Leimgruber, R.A. Stegeman, K.S. Anderson, J.A. Sikorski, S.R. Padgette, G.M. Kishore, Structure and topological symmetry of the glyphosate target 5-enol-pyruvylshikimate-3-phosphate synthase: a distinctive protein fold, Proc. Natl. Acad. Sci. USA 88 (1991) 5046–5050.
[14] J.S. Oliveira, C.A. Pinto, L.A. Basso, D.S. Santos, Cloning and overexpression in soluble form of functional shikimate kinase and 5-enolpyruvylshikimate 3-phosphate synthase enzymes from *Mycobacterium tuberculosis*, Protein Express. Purif. 22 (2001) 430–435.
[15] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[16] A. Sali, J.P. Overington, Derivation of rules for comparative protein modeling from a database of protein structure alignments, Protein Sci. 3 (9) (1994) 1582–1596.
[17] A. Sali, L. Potterton, F. Yuan, H. van Vlijmen, M. Karplus, Evaluation of comparative protein modeling by MODELLER, Proteins 23 (3) (1995) 318–326.
[18] A. Sali, Modeling mutations and homologous proteins, Curr. Opin. Biotechnol. 6 (4) (1995) 437–451.
[19] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy minimization and dynamics calculations, J. Comp. Chem. 4 (1983) 187–217.
[20] W. Braun, N. Go, Calculation of protein conformations by proton–proton distance constraints. A new efficient algorithm, J. Mol. Biol. 186 (3) (1985) 611–626.
[21] R.A. Laskowski, M.W. MacArthur, D.K. Smith, D.T. Jones, E.G. Hutchinson, A.L. Morris, D. Naylor, D.S. Moss, J.M. Thorton, PROCHECK v.3.0—Program to Check the Stereochemistry Quality of Protein structures—Operating Instructions, 1994.
[22] W.F. De Azevedo Jr., R.J. Ward, F.R. Lombardi, J.R. Giglio, A.M. Soares, M.R.M. Fontes, R.K. Arni, Crystal structure of myotoxin-II: a myotoxic phospholipase A2-homologue from *Bothrops moojeni* venom, Protein Peptide Lett. 4 (5) (1997) 329–334.
[23] W.F. De Azevedo Jr., R.J. Ward, J.M. Gutierrez, R.K. Arni, Structure of a Lys49 phospholipase A2 homologue isolated from the venom of *Bothrops nummifer* (jumping viper), Toxicon 37 (2) (1999) 371–384.
[24] W.F. De Azevedo Jr., R.J. Ward, F. Canduri, A. Soares, J.R. Giglio, R.K. Arni, Crystal structure of piratoxin-I: a calcium-independent, myotoxic phospholipase A_2-homologue from *Bothrops pirajai* venom, Toxicon 36 (10) (1998) 1395–1406.
[25] A. Nicholls, K.A. Sharp, B. Honig, Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, Proteins 11 (4) (1991) 281–296.
[26] F. Canduri, L.G.V.L. Teodoro, C.C.B. Lorenzi, V. Hial, R.A.S. Gomes, J. Ruggiero Neto, W.F. de Azevedo Jr., Crystal structure of human uropepsin at 2.45 Å resolution, Acta Crystallogr. D 57 (2001) 1560–1570.
[27] W.F. De Azevedo Jr., H.J. Mueller-Dieckmann, U. Schulze-Gahmen, P.J. Worland, E. Sausville, S.-H. Kim, Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase, Proc. Natl. Acad. Sci. USA 93 (7) (1996) 2735–2740.
[28] W.F. De Azevedo Jr., S. Leclerc, L. Meijer, L. Havlicek, M. Strnad, S.-H. Kim, Inhibition of cyclin-dependent kinases by

purine analogues: crystal structure of human CDK2 complexed with roscovitine, Eur. J. Biochem. 243 (1997) 518–526.
[29] W.F. De Azevedo Jr., F. Canduri, V. Fadel, L.G.V.L. Teodoro, V. Hial, R.A.S. Gomes, Molecular model for the binary complex of uropepsin and pepstatin, Biochem. Biophys. Res. Commun. 287 (1) (2001) 277–281.
[30] W.F. De Azevedo Jr., F. Canduri, N.J.F. da Silveira, Structural basis for inhibition of cyclin-dependent kinase9 by Flavopiridol, Biochem. Biophys. Res. Commun. 293 (1) (2002) 566–571.
[31] S.-H. Kim, U. Schulze-Gahmen, J. Brandsen, W.F de Azevedo Jr., Structural basis for chemical inhibitor of CDK2, Prog. Cell Cycle Res. 2 (1996) 137–145.
[32] W.F. De Azevedo, R.T. Gaspar, F. Canduri, J.C. Camera, N.J.F. da Silveira, Molecular model of cyclin-dependent kinase 5 complexed with roscovitine, Biochem. Biophys. Res. Commun. 297 (2002) 1154–1158.
[33] W.F. De Azevedo Jr., J.S. de Oliveira, L.A. Basso, M.S. Palma, J.H. Pereira, F. Canduri, D.S. Santos, Molecular model of shikimate kinase from *Mycobacterium tuberculosis*, Biochem. Biophys. Res. Commun. 295 (1) (2002) 142–148.
[34] W.F. De Azevedo, F. Canduri, D.M. Santos, R.G. Silva, J.S. Oliveira, L.P.S. Carvalho, L.A. Basso, M.A. Mendes, M.S. Palma, D.S. Santos, Crystal structure of human purine nucleoside phosphorylase at 2.3 Å resolution, Biochem. Biophys. Res. Commun. 308 (3) (2003) 545–552.
[35] D.M. Santos, F. Canduri, J.H. Pereira, M.V.B. Dias, R.G. Silva, M.A. Mendes, M.S. Palma, L.A. Basso, W.F. de Azevedo, D.S. Santos, Crystal structure of human purine nucleoside phosphorylase complexed with acyclovir, Biochem. Biophys. Res. Commun. 308 (3) (2003) 553–559.
[36] W.F. De Azevedo, F. Canduri, D.M. Dos Santos, J.H. Pereira, M.V.B. Dias, R.G. Silva, M.A. Mendes, L.A. Basso, M.S. Palma, D.S. Santos, Structural basis for inhibition of human PNP by immucillin-H, Biochem. Biophys. Res. Commun. 309 (2003) 922–927.
[37] D.G. Higgins, A.J. Bleasby, R. Fuchs, CLUSTAL V: improved software for multiple sequence alignment, Comput. Appl. Biosci. 8 (2) (1992) 189–191.
[38] P. Kraulis, MOLSCRIPT: a program to produce both detailed and schematic plots of proteins, J. Appl. Cryst. 24 (1991) 946–950.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Protein Expression and Purification 28 (2003) 287–292
Protein Expression & Purification
www.elsevier.com/locate/yprep

One-step purification of 5-enolpyruvylshikimate-3-phosphate synthase enzyme from *Mycobacterium tuberculosis*

Jaim S. Oliveira,[a] Maria A. Mendes,[b] Mário S. Palma,[b] Luiz A. Basso,[a,*] and Diógenes S. Santos[a,1]

[a] *Grupo de Microbiologia Molecular e Funcional, Departamento de Biologia Molecular e Biotecnologia, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves 9500, Porto Alegre, RS 91501-970, Brazil*
[b] *Laboratório de Biologia Estrutural e Zooquímica (CEIS), Departamento de Biologia, Instituto de Biociências, Universidade do Estado de São Paulo, Rio Claro, SP 13506-900, Brazil*

Received 16 September 2002, and in revised form 4 December 2002

Abstract

Currently, there are 8 million new cases and 2 million deaths annually from tuberculosis, and it is expected that a total of 225 million new cases and 79 million deaths will occur between 1998 and 2030. The reemergence of tuberculosis as a public health threat, the high susceptibility of HIV-infected persons, and the proliferation of multi-drug-resistant strains have created a need to develop new antimycobacterial agents. The existence of homologues to the shikimate pathway enzymes has been predicted by the determination of the genome sequence of *Mycobacterium tuberculosis*. We have previously reported the cloning and overexpression of *M. tuberculosis aroA*-encoded EPSP synthase in both soluble and active forms, without IPTG induction. Here, we describe the purification of *M. tuberculosis* EPSP synthase (mtEPSPS) expressed in *Escherichia coli* BL21(DE3) host cells. Purification of mtEPSPS was achieved by a one-step purification protocol using an anion exchange column. The activity of the homogeneous enzyme was measured by a coupled assay using purified shikimate kinase and purine nucleoside phosphorylase proteins. A total of 53 mg of homogeneous enzyme could be obtained from 1 L of LB cell culture, with a specific activity value of approximately 18 U mg^{-1}. The results presented here provide protein in quantities necessary for structural and kinetic studies, which are currently underway in our laboratory.

Among deaths associated with infectious diseases, tuberculosis kills more adolescents and adults than any other single infection [1]. Currently, there are 8 million new cases and 2 million deaths annually from tuberculosis, and it is predicted that a total of 225 million new cases and 79 million deaths will occur between 1998 and 2030 [2]. The reemergence of TB as a public health threat, the high susceptibility of human immunodeficienty virus-infected persons to the disease, and the proliferation of multi-drug-resistant (MDR) strains have created much scientific interest in developing new antimycobacterial agents to both treat *Mycobacterium tuberculosis* strains resistant to existing drugs and shorten the duration of short-course treatment to improve patient compliance [3]. Moreover, treatment of patients infected with MDR *M. tuberculosis* must rely on second-line drugs, which are less effective and more expensive, and can cost up to $250,000 per person and take 2 years [4].

In the shikimate pathway, phosphoenolpyruvate (PEP)[2] and erythrose 4-phosphate (E4P) are converted to chorismate through seven enzymatic steps. This

* Corresponding author. Fax: +55-51-33166234.
E-mail addresses: labasso@dna.cbiot.ufrgs.br (L.A. Basso), diogenes@dna.cbiot.ufrgs.br (D.S. Santos).
[1] Also Corresponding author.

[2] *Abbreviations used*: EPSP, 5-enolpyruvylshikimate-3-phosphate; ESI-MS, electrospray ionization mass spectrometry; IPTG, isopropyl β-D-thiogalactoside; LB, Luria–Bertani; MESG, 2-amino-6-mercapto-7-methylpurine ribonucleoside; mtEPSPS, *Mycobacterium tuberculosis* EPSP synthase; mtSK, *Mycobacterium tuberculosis* shikimate kinase; PEP, phosphoenolpyruvate; P_i, inorganic phosphate; PNP, purine nucleoside phosphorylase.

1046-5928/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
doi:10.1016/S1046-5928(02)00708-8

pathway is a promising target for the development of potentially non-toxic herbicides and antimicrobial agents because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals [5]. Moreover, biochemical, genetic, and chemotherapeutic data have been presented for the existence of a functional shikimate pathway in apicomplexan parasites [6,7]. Thus, the shikimate pathway enzymes should also provide attractive targets for the development of new antiparasite agents. In mycobacteria, the branch-point compound chorismate is a precursor for the synthesis of aromatic amino acids, naphthoquinones, menaquinones, and mycobactins [8]. Although evidence that the shikimate pathway is essential in *M. tuberculosis* is lacking, the disruption of the *aroD* gene was used successfully to generate attenuated oral vaccine strains of *Salmonella typhi* and other bacteria [9]. Furthermore, the salicylate-derived mycobactin siderophores were recently shown to be essential for *M. tuberculosis* growth in macrophages [10].

Homologues to the shikimate pathway enzymes were identified in the complete genome sequence of *M. tuberculosis* H37Rv strain [11]. Among them, the 5-enolpyruvylshikimate-3-phosphate synthase (mtEPSPS, EC 2.5.1.19) encoding gene (*aroA*, Rv3227) has been cloned and the enzyme was overexpressed in both soluble and active forms in *E. coli* BL21(DE3) cells [12]. EPSPS catalyzes an unusual transfer reaction of the carboxyvinyl portion of phosphoenolpyruvate (PEP) regiospecifically to the 5-OH of shikimate 3-phosphate (S3P), forming the products 5-enolpyruvylshikimate-3-phosphate (EPSP) and inorganic phosphate (P_i) [13]. This reaction is the penultimate step of the shikimate pathway. EPSPS is in fact the site of action of glyphosate [*N*-(phosphonomethyl)glycine], which is a widely used broad-spectrum herbicide [14].

To pave the way for structural and functional efforts currently underway in our laboratory, the mtEPSPS enzyme, overexpressed in *E. coli* BL21(DE3) host cells, was purified by liquid chromatography and assayed by a coupled assay containing purified *M. tuberculosis* shikimate kinase (mtSK) and purine nucleoside phosphorylase (PNP) proteins. The purification was achieved by a one-step purification protocol using an anion exchange column, which yielded homogeneous mtEPSPS enzyme in large quantities. This work will allow steady-state and pre-steady-state kinetic studies to be performed as well as crystal structure determination of mtEPSPS enzyme. These kinetic studies, in combination with structural studies of mtEPSPS in complexes with its substrates, products, or inhibitors, will contribute to elucidation of controversial enzymatic mechanism of this enzyme and, consequently, to current efforts towards the development of new drugs for tuberculosis treatment.

Materials and methods

Overexpression of SK and EPSPS

The recombinant plasmid pET-23a(+)::*aro*A was transformed into electrocompetent *E. coli* BL21(DE3) cells and selected on LB agar plates containing $50\,\mu g\,mL^{-1}$ carbenicillin. Single colonies were used to inoculate 1.5 L of LB media containing $50\,\mu g\,mL^{-1}$ carbenicillin in 2-L shaking flasks. The mtEPSPS culture cells were grown without addition of isopropyl β-D-thiogalactopyranoside (IPTG) at 180 RPM and 37 °C for 23 h [12]. Cells were harvested by centrifugation at 6000*g* for 20 min at 4 °C and stored at −80 °C.

Purification of recombinant Mycobacterium tuberculosis EPSPS

The mtEPSPS cells (3.2 g) were suspended in 25 mL of 50 mM Tris–HCl (pH 7.8) and incubated for 30 min in the presence of lysozyme ($0.2\,mg\,mL^{-1}$). Cells were disrupted by sonication using four pulses of 20 s each at 60% amplitude with a 13 mm probe (High Intensity Ultrasonic Processor, 750 Watt model, Sonics Vibra Cell). Cell debris was removed by centrifugation at 48,000*g* for 60 min at 4 °C. The supernatant was incubated with 1% (w/v) of streptomycin sulfate and centrifuged at 48,000*g* for 30 min, at 4 °C. The supernatant was dialyzed twice against 2 L of the same buffer using a dialysis tubing with a molecular weight exclusion limit of 12,000–14,000. The sample was centrifuged at 48,000*g* for 30 min at 4 °C and the resulting supernatant (~30 mL) was loaded onto an FPLC Q Sepharose Fast Flow (26×9.5) column (Amersham Pharmacia Biotech). Isocratic elution of the recombinant mtEPSPS enzyme was achieved by washing the ion exchange resin with 10 column volumes (~500 mL) of 50 mM Tris–HCl (pH 7.8) at $1\,mL\,min^{-1}$. Protein purification was monitored by SDS–PAGE [15].

Protein determination

Protein concentration was determined by the method of Bradford [16] using the Bio-Rad Protein Assay kit (BIO-RAD) and bovine serum albumin as standard.

mtEPSPS activity assay

Recombinant mtEPSPS was assayed in the forward direction in 100 mM Tris–HCl, 50 mM KCl, and 5 mM $MgCl_2$, pH 7.6, at 25 °C. Enzyme activity was measured by estimating the rate of inorganic phosphate (P_i) release using a continuous spectrophotometric coupled assay with purine nucleoside phosphorylase (PNP; EC 2.4.2.1) and 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG) [17]. The PNP en-

zyme and MESG are commercially available as the Enzchek Phosphate Assay kit (Molecular Probes). This assay is based on the difference in absorbance between MESG and the purine base product of its reaction with P_i catalyzed by PNP. This reaction gives an absorbance increase at 360 nm with an extinction coefficient of $11,000\,M^{-1}\,cm^{-1}$ at pH 7.6 [17]. Since SK catalyzes a phosphate transfer from ATP to the carbon-3 hydroxyl group of shikimate forming shikimate-3-phosphate (S3P), the synthesis of the mtEPSPS substrate was achieved by adding homogeneous mtSK, shikimate, and ATP to the reaction mixture in the vial and the P_i released by mtEPSP enzyme activity was detected by PNP coupling enzyme. Typically, a 500 μL assay mixture contained $10\,U\,mL^{-1}$ mtSK, 0.2 mM MESG, $1\,U\,mL^{-1}$ PNP, 1 mM phosphoenolpyruvte (PEP), 2.4 mM shikimate, and 1 mM ATP. All components were mixed and incubated for 1 min at 25 °C, and the reactions were initiated by addition of mtEPSPS.

It should be pointed out that the amount of mtSK was chosen so that the reaction is 99% complete in a few seconds, assuming that the concentration of shikimate is smaller than its Michaelis–Menten constant as the reaction approaches completion. Accordingly, the required quantity of mtSK was calculated by using a value of 0.083 mM for mtSK ATP K_m [18] in the following equation [19]:

$$\ln\frac{s_0}{s_1} = \frac{V_m}{K_m}\cdot t,$$

where t is the total time (min^{-1}) for a decrease in s (ATP concentration) from 100% (s_0) to 1% (s_1) and V_m refers to the activity of mtSK in the buffer assay (μmol min^{-1} L^{-1}). A value of 1.15 s can be estimated by the above equation for conversion of approximately 99% of the initial concentration of substrate into product using $10\,U\,mL^{-1}$ mtSK and a final ATP concentration in the assay significantly smaller than its K_m value. Since the concentration of ATP used in the assay mixture (1 mM) is larger than its Michaelis–Menten constant, the time taken to reach 99% completion of the reaction is even shorter. One unit of enzyme activity (U) is defined for mtEPSPS, mtSK, and PNP enzymes as the amount of protein catalyzing the conversion of 1 μmol of substrate per minute.

PNP activity assay

Since the number of units for the coupling enzymes should be determined in the same assay conditions for the enzyme being studied, the PNP enzyme activity was determined in the same assay buffer utilized in mtEPSPS assays. In these conditions, the PNP enzyme activity was 3-fold lower than that measured in the assay buffer of Enzchek Phosphate Assay kit (20 mM Tris–HCl, 1 mM $MgCl_2$, pH 7.5, containing 0.1 mM sodium azide). Accordingly, the amount of protein corresponding to 1 U of PNP added to mtEPSPS assay mixture was three times larger.

Shikimate kinase activity assay

Purified mtSK enzyme activity, used in mtEPSPS assay, was assayed in the forward direction by coupling the ADP product formation to the pyruvate kinase (PK; EC 2.7.1.40) and lactate dehydrogenase (LDH; EC 1.1.1.27). Shikimate-dependent oxidation of NADH was continuously monitored at 340 nm. An extinction coefficient of NADH equal to $6.22 \times 10^3\,M^{-1}\,cm^{-1}$ was used for rate calculations. All reactions were carried out at 25 °C and initiated with addition of mtSK homogeneous protein. The 500 μL assay mixture contained 100 mM Tris–HCl buffer, pH 7.6, 50 mM KCl, 5 mM $MgCl_2$, 2.4 mM shikimic acid, 1 mM ATP, 1 mM phosphoenolpyruvate, 0.2 mM NADH, $3\,U\,mL^{-1}$ of pyruvate kinase, and $2.5\,U\,mL^{-1}$ of lactate dehydrogenase. PEP, (-)shikimate, LDH and PK enzymes, ATP, and NADH were all purchased from Sigma. The activities of LDH and PK enzymes ($U\,mL^{-1}$) were determined in the same assay buffer utilized for mtSK (100 mM Tris–HCl buffer, pH 7.6, 50 mM KCl, and 5 mM $MgCl_2$ at 25 °C).

Mass spectrometry analysis

The homogeneity of protein preparation was assessed by mass spectrometry (MS), employing some adaptations made to the system described by Chassaigne and Lobinski [20]. Samples were analyzed on a triple quadrupole mass spectrometer, model QUATTRO II, equipped with a standard electrospray (ESI) probe (Micromass, Altrinchan), adjusted to ca. 250 μL min^{-1}. The source temperature (80 °C) and needle voltage (3.6 kV) were maintained constant throughout the experimental data collection, applying a drying gas flow (nitrogen) of $200\,L\,h^{-1}$ and a nebulizer gas flow of $20\,L\,h^{-1}$. The mass spectrometer was calibrated with intact horse heart myoglobin and its typical cone-voltage induced fragments. The subunit molecular mass of recombinant mtEPSPS was determined by ESI-MS, adjusting the mass spectrometer to give a peak width at half-height of 1 mass unit, and the cone sample to skimmer lens voltage controlling the ion transfer to mass analyzer was set to 38 V. About 50 pmol (10 μL) of each sample was injected into electrospray transport solvent. The ESI spectrum was obtained in the multi-channel acquisition mode, scanning from 500 to 800 *m/z* at a scan time of 5 s. The mass spectrometer is equipped with MassLynx and Transform software for data acquisition and spectra handling.

N-Terminal amino acid sequencing

The N-terminal amino acid residues of purified recombinant mtEPSPS were identified by automated Edman degradation sequencing using a PPSQ 21A gas-phase sequencer (Shimadzu).

Results

Mycobacterium tuberculosis EPSP synthase was purified from *E. coli* BL21(DE3) cells carrying pET-23a(+)::*aro*A recombinant plasmid as described in Materials and methods and the samples of each step were analyzed by SDS–PAGE with Coomassie blue staining and assayed for enzyme activity in the forward reaction.

Purification of recombinant mtEPSPS

The mtEPSPS is a protein of 450 amino acids with 41 basic and 38 acidic residues (38 Arg plus 3 Lys and 25 Asp plus 13 Glu). In contrast to the protein purification protocol described for recombinant EPSP synthase enzyme from *E. coli* (ecEPSPS), which has 38 basic residues and 48 acidic residues [21], mtEPSPS does not adsorb to anionic exchange matrices equilibrated with buffer 50 mM Tris–HCl, pH 7.8. This is in accordance with the isoelectric point values of 9.6 and 5.18 calculated for, respectively, mtEPSPS and ecEPSPS enzymes. Hence, the mtEPSPS enzyme could be separated by a method of isocratic elution, while ecEPSP was probably adsorbed to anionic-exchange matrix Q-Sepharose Fast Flow. In purification procedures of non-adsorption in ion-exchange chromatography, one should apply a small volume of sample on column relative to volume of the column being used for better resolution [22]. Although a high sample to column volume ratio of 0.6 was used here (30 mL/50 mL), the significant retardation of mtEPSPS on anion-exchange column led to complete separation from the non-adsorbed fraction of proteins. The partition coefficient (α) defined as the fraction of protein adsorbed, which is the ratio between adsorbed molecules and total molecules (adsorbed + desorbed), gives values ranging from 0 to 1. For a particular protein that adsorbs partially ($0 < \alpha < 1$), the center of the protein peak will emerge after $1/(1-\alpha)$ times the column volume of buffer has passed through the column [22]. Since the volume of the anion exchange column (50 mL) and the center of the mtEPSPS elution peak (150 mL) values are known, a partition coefficient value of 0.67 can be estimated (assuming that the total adsorptive capacity of the columns is much greater than the total amount of protein in the sample). This value is consistent with a significant retardation of mtEPSPS protein on the anion exchange column, leading to separation from the non-adsorbed and strongly adsorbed fractions. The mtEPSPS from the Q-Sepharose was estimated by SDS–PAGE analysis to be more than 95% pure, as seen on an overloaded Coomassie stained 10–15% polyacrylamide gel (Fig. 1).

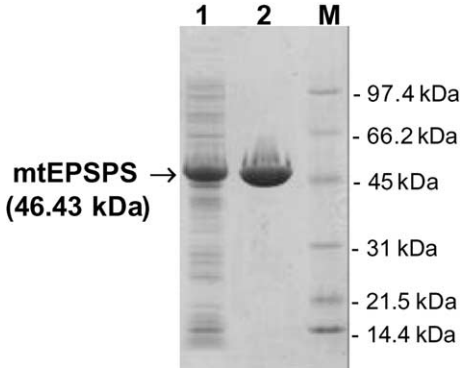


Fig. 1. SDS–PAGE analysis of pooled fractions from the purification protocol steps of *M. tuberculosis* EPSP synthase. Lane 1, crude extract; lane 2, Q-Sepharose fast flow anion exchange; lane 3, MW markers.

Recombinant mtEPSPS enzyme activity assay

The samples of each purification step were assayed for EPSPS enzyme activity in the forward reaction by using a continuous spectrophotometric coupled assay with mtSK and PNP proteins. It should be pointed out that the mtSK being used in the assays has been purified to homogeneity and estimated by gels of SDS–PAGE with Coomassie blue staining to be more than 95% pure (unpublished results). To show that the observed rate values in this coupled assay are actual rate values of mtEPSPS

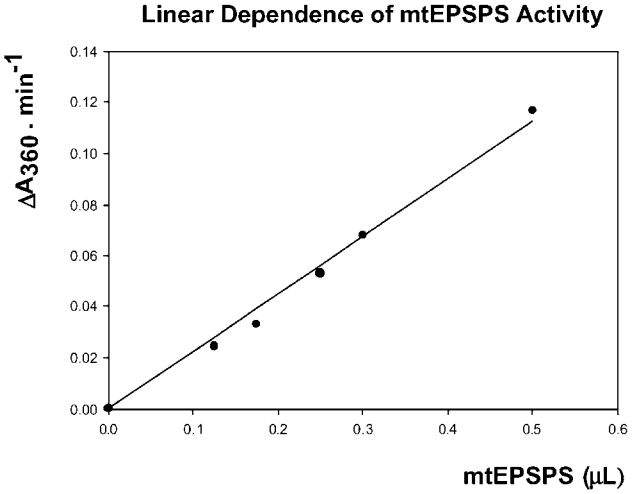


Fig. 2. Linear dependence of purified mtEPSPS activity. The rates of enzyme activity were carried out in the forward direction by continuously monitoring the increase of 2-amine-6-mercapto-7-methylpurine absorbance at 360 nm (slope = 0.238 ± 0.009).

Table 1
Purification of *M. tuberculosis* 5-enolpyruvylshikimate-3-phosphate synthase from *E. coli* BL21(DE3) [pET23a(+)::*aro*A] cells

Purification step	Protein (mg)	Units (U)	Sp act[a] ($U mg^{-1}$)	Purification fold	Yield (%)
Crude ext.[b]	244.93	1590.55	6.49	1.00	100
Anion exchange[c]	53.00	961.13	18.13	2.79	60

[a] $U mL^{-1}/mg mL^{-1}$ (buffer assays as described in Materials and methods).
[b] Cell crude extract in 50 mM Tris–HCl, pH 7.8.
[c] Homogeneous mtEPSPS after anion exchange in 50 mM Tris–HCl, pH 7.8.

in steady-state, some control experiments were made. First, the assay was carried out utilizing either half of mtSK ($5 U mL^{-1}$) or twice the amount of PNP ($2 U mL^{-1}$) in the vial. In both conditions, the specific activity observed was the same as that measured in typical assay conditions using $10 U mL^{-1}$ mtSK or $1 U mL^{-1}$ PNP. Second, when all assay components except either PEP or mtEPSPS were mixed and incubated at 25 °C, no changes in continuously measured values of absorbance at 360 nm were observed within 5 min. Moreover, the activity of mtEPSPS enzyme was linearly dependent on sample volume added to the reaction mixture (Fig. 2), thereby showing that the initial velocity is proportional to total enzyme concentration and that true initial velocities are being measured. The mtEPSPS enzyme was stable, since it could be stored as precipitated in ammonium sulfate saturated solution at 4 °C, with no observed decrease in specific activity after 4 months.

ESI-MS and N-terminal amino acid sequencing analyses

The subunit molecular mass of active mtEPSPS was determined to be 46,458.10 Da by electrospray ionization mass spectrometry (ESI-MS), consistent with the predicted mass for the full length gene product of 46,425.80 Da. The ESI-MS result revealed no peak at the expected mass for *E. coli* EPSP synthase (46,095.85 Da), thus providing evidence for both the identity and purity of the recombinant protein. The first 10 N-terminal amino acid residues of the recombinant protein were identified as MKTWPAPTAP by the Edman degradation method. This result unambiguously identifies the recombinant protein as *M. tuberculosis* EPSPS and confirms that of the N-terminal methionine residue was not removed from it. The non-modification at the N-terminal amino acid sequence of mtEPSPS is in accordance with the presence of the charged amino acid residue lysine in the second position, since the methionine aminopeptidase catalyzed cleavage of the N-terminal methionine is usually removed if the second residue is small and uncharged (glycine, alanine, serine, threonine, proline, valine, and cysteine) [23,24]. The results described above also confirm the fact that mtEPSPS enzyme was effectively separated from the ecEPSPS enzyme by the anion exchange chromatography step of the purification protocol.

Discussion

Table 1 shows the results of a typical protocol employed for mtEPSPS protein purification. The enzymatic assay and protein concentration determination showed a specific activity of 18.13 U/mg for the homogeneous target protein, indicating that the one-step protocol utilized resulted in a 2.8-fold purification. This result is in accordance with an mtEPSPS overexpression of approximately 40% of total cell protein, as previously reported by Oliveira et al. [12]. An amount of 53 mg of homogeneous cloned mtEPSPS protein could be obtained from 3 g of cells or, stated otherwise, approximately 53 mg/L of LB medium. Although it is often argued that the cost of IPTG limits the usefulness of the *lac* promoter to high-added-value products, we showed here that large amounts of mtEPSPS can be purified from cells grown in the absence of IPTG. In another report, T7 RNA polymerase expression system has been used for overexpression and purification of ecEPSPS in quantities equal to that reported here for mtEPSPS however; it was achieved by induction of *E. coli* BL21(DE3) cell culture with 0.4 mM IPTG [21].

In this report, we present the purification of EPSPS from *M. tuberculosis*. The purification was achieved by a simple and rapid purification protocol, which yielded mtEPSPS enzyme in large quantities. To raise antibodies, Garbe et al. [25] reported the cloning of the mycobacterial *aroA* gene and expression of an insoluble fusion protein between 345 amino acids of mtEPSPS and glutathione *S*-transferase. Therefore, mtEPSPS was characterized only by immunochemical analysis and not by functional studies as described here. Thus, to the best of our knowledge, the data presented here are the first report of purification in active form of mtEPSPS enzyme. Homogeneous mtEPSPS enzyme will provide protein in quantities necessary for both X-ray crystal structure determination and studies on the enzyme mechanism of action by steady-state and pre-steady-state kinetics aiming at antimycobacterial agent development.

Acknowledgments

This work was supported by FINEP and Millenium Institute, MCT/CNPq (Brazil) grants awarded to L.A.B.

and D.S.S. We thank Prof. John S. Blanchard for generously supporting this work.

References

[1] World Health Organization, World Health Organization Report on Infectious Diseases: Removing Obstacles to Healthy Development. Atar, Switzerland. WHO/CDS/99.1, 1999.
[2] C. Murray, J. Salomon, Modeling the impact of global tuberculosis control strategies, Proc. Natl. Acad. Sci. USA 95 (1998) 13881–13886.
[3] L.A. Basso, J.S. Blanchard, Resistance to antitubercular drugs, Adv. Exp. Med. Biol. 456 (1998) 115–144.
[4] E. Stokstad, Drug-resistant TB on the rise, Nature 287 (2000) 2391.
[5] R. Bentley, The shikimate pathway—a metabolic tree with many branches, Crit. Rev. Biochem. Mol. Biol. 25 (1990) 307–384.
[6] G.A. McConkey, Targeting the shikimate pathway in the malaria parasite *Plasmodium falciparum*, Antimicrob. Agents Chemother. 43 (1999) 175–177.
[7] F. Roberts, C.W. Roberts, J.J. Johnson, D.E. Kyle, T. Krell, J.R. Coggins, G.H. Coombs, W.K. Milhous, S. Tzipori, D.J.P. Fergunson, D. Chakrebarti, R. McLeod, Evidence for the shikimate pathway in apicomplexan parasites, Nature 393 (1998) 801–805.
[8] C. Ratledge, Nutrition, growth and metabolism, in: C. Ratledge, J.L. Stanford (Eds.), The Biology of the Mycobacteria, vol. 1, Academic Press, London, 1982, pp. 185–271.
[9] C.O. Tacket, M.B. Sztein, G.A. Losonsky, S.S. Wasserman, J.P. Nataro, R. Edelman, D. Pickard, G. Dougan, S.N. Chatfield, M.M. Levine, Safety of live oral *Salmonella typhi* vaccine strains with deletions in *htrA* and *aroC aroD* and immune response in humans, Infect. Immun. 65 (1997) 452–456.
[10] J.J. De Voos, K. Rutter, B.G. Schroder, H. Su, Y. Zhu, C.E. Barry III, The salicylate-derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for growth in macrophages, Proc. Natl. Acad. Sci. USA 97 (2000) 1252–1257.
[11] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B.G. Barrell, Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, Nature 393 (1998) 537–544.
[12] J.S. Oliveira, C.A. Pinto, L.A. Basso, D.S. Santos, Cloning and overexpression in soluble form of functional shikimate kinase and 5-enolpyruvylshikimate 3-phosphate synthase enzymes from *Mycobacterium tuberculosis*, Protein Expr. Purif. 22 (2001) 430–435.
[13] J.A. Sikorski, K.J. Gruys, Understanding glyphosate's molecular mode of action with EPSP synthase: evidence favoring an allosteric inhibitor model, Acc. Chem. Res. 30 (1997) 2–8.
[14] H.C. Steinrücken, N. Amrhein, The herbicide glyphosate is a potent inhibitor of 5-enol pyruvyl shikimic acid-3-phosphate synthase, Biochem. Biophys. Res. Commun. 94 (1980) 1207–1212.
[15] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.
[16] M.M. Bradford, R.A. McRorie, W.L. Williams, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding, Anal. Biochem. 72 (1976) 248–254.
[17] M.R. Webb, A continuous spectrophotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems, Proc. Natl. Acad. Sci. USA 89 (1992) 4884–4887.
[18] Y. Gu, L. Reshetnikova, Y. Li, Y. Wu, H. Yan, S. Singh, X. Ji, Crystal structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the dynamic role of the LID domain in catalysis, J. Mol. Biol. 319 (2002) 779–789.
[19] D.A. Harris, Spectrophotometric assays, in: D.A. Harris, C.L. Bashford (Eds.), Spectrophotometry & Spectrofluorimetry: a Pratical Approach, IRL Press, Oxford, 1988, pp. 49–90.
[20] H. Chassaigne, R. Lobinski, Characterization of horse kidney metallothionein isoforms by electrospray MS and reversed-phase HPLC-electrospray MS, Analyst 123 (1998) 2125–2130.
[21] W.A. Shuttleworth, C.D. Hough, K.P. Bertrand, J.N.S. Evans, Over-production of 5-enolpyruvylshikimate-3-phosphate synthase in *Escherichia coli*: use of T7 promoter, Protein Eng. 5 (1992) 461–466.
[22] R.K. Scopes, Separation by adsorption II: ion exchangers and nonspecific adsorbents, in: C.R. Cantor (Ed.), Protein Purification: Principles and Practice, Springer, New York, 1994, pp. 167–171.
[23] P.-H. Hirel, J.-M. Schmitter, P. Dessen, G. Fayat, S. Blanquet, Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the sis-chain length of the penultimate amino acid. Proc. Natl. Acad. Sci. USA 86 (1989) 8247–8251.
[24] W.T. Lowther, B.W. Matthews, Structure and function of the methionine aminopeptidases, Biochim. Biophys. Acta 1477 (2000) 157–167.
[25] T. Garbe, C. Joens, I. Charles, G. Dougan, D. Young, Cloning and characterization of the *aro*A gene from *Mycobacterium tuberculosis*, J. Bacteriol. 172 (1990) 6774–6782.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 308 (2003) 545–552
BBRC
www.elsevier.com/locate/ybbrc

Crystal structure of human purine nucleoside phosphorylase at 2.3 Å resolution

Walter Filgueira de Azevedo Jr.,[a,b,*] Fernanda Canduri,[a,b] Denis Marangoni dos Santos,[a,b] Rafael Guimarães Silva,[c] Jaim Simões de Oliveira,[c] Luiz Pedro Sório de Carvalho,[c] Luiz Augusto Basso,[c] Maria Anita Mendes,[b,d] Mário Sérgio Palma,[b,d] and Diógenes Santiago Santos[c,e,*]

[a] *Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil*
[b] *Center for Applied Toxinology, Instituto Butantan, São Paulo, SP 05503-900, Brazil*
[c] *Rede Brasileira de Pesquisas em Tuberculose, Departamento de Biologia Molecular e Biotecnologia, UFRGS, Porto Alegre, RS 91501-970, Brazil*
[d] *Laboratory of Structural Biology and Zoochemistry-CEIS/Department Biology, Institute of Biosciences,UNESP, Rio Claro, SP 13506-900, Brazil*
[e] *Faculdade de Farmácia/Instituto de Pesquisas Biomédicas, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil*

Received 24 June 2003

Abstract

Purine nucleoside phosphorylase (PNP) catalyzes the phosphorolysis of the *N*-ribosidic bonds of purine nucleosides and deoxynucleosides. In human, PNP is the only route for degradation of deoxyguanosine and genetic deficiency of this enzyme leads to profound T-cell mediated immunosuppression. PNP is therefore a target for inhibitor development aiming at T-cell immune response modulation and its low resolution structure has been used for drug design. Here we report the structure of human PNP solved to 2.3 Å resolution using synchrotron radiation and cryocrystallographic techniques. This structure allowed a more precise analysis of the active site, generating a more reliable model for substrate binding. The higher resolution data allowed the identification of water molecules in the active site, which suggests binding partners for potential ligands. Furthermore, the present structure may be used in the new structure-based design of PNP inhibitors.

Keywords: PNP; Synchrotron radiation; Structure; Drug design

Purine nucleoside phosphorylase (PNP) catalyzes the phosphorolysis of purine nucleosides to corresponding bases and sugar 1-phosphate. PNP plays a central role in purine metabolism, normally operating in the purine salvage pathway of cells. PNP is specific for purine nucleosides in the β-configuration and exhibits a preference for ribosyl-containing nucleosides relative to the analogs containing the arabinose, xylose, and lyxose stereoisomers [1]. Moreover, PNP cleaves glycosidic bond with inversion of configuration to produce α-ribose 1-phosphate [2].

PNP is a potential target for drug development, which could induce immune suppression to treat, for instance, autoimmune diseases, T-cell leukemia, lymphoma and organ transplantation rejection. Furthermore, PNP inhibitors can also be used to avoid cleavage of anticancer and antiviral drugs, since many of these drugs mimic natural purine nucleosides and can thereby be cleaved by PNP before accomplishing their therapeutic role [3,4].

The crystallographic structure of human PNP (HsPNP) was first determined in 1990 at 3.2 Å resolution [3]. Further crystallographic studies improved the resolution to 2.8 Å [4]. These atomic coordinates have been extensively used for structure-based design of PNP inhibitors [5–9].

We have now obtained higher-resolution X-ray diffraction data and refined the structure of the apoenzyme to 2.3 Å, using recombinant human PNP and synchro-

* Corresponding authors. Fax: +55-17-221-2247.
E-mail address: walterfa@df.ibilce.unesp.br (W.F. de Azevedo Jr.).

0006-291X/$ - see front matter © 2003 Published by Elsevier Inc.
doi:10.1016/S0006-291X(03)01431-1

tron radiation. Our analysis of the PNP structural data, hydration of residues in the binding pocket, and structural differences between the apoenzyme and PNP:guanine complex provides new insights into substrate binding, the purine-binding site, and can be used for future inhibitor design.

Materials and methods

Crystallization. Recombinant human PNP was expressed and purified as previously described [10]. HsPNP was crystallized using the experimental conditions described elsewhere [11]. In brief, a PNP solution was concentrated to 12 mg mL^{-1} against 10 mM potassium phosphate buffer (pH 7.1). Hanging drops were equilibrated by vapor diffusion at 25 °C against reservoir containing 19% saturated ammonium sulfate solution in 0.05 M citrate buffer (pH 5.3).

Data collection. Preliminary X-ray studies of HsPNP crystals showed that these crystals diffracted to 2.3 Å resolution, using synchrotron radiation, although they decayed when exposed to X-rays at room temperature [4]. In order to increase the resolution we collected data from a flash-cooled crystal at 104 K. Prior to flash cooling, glycerol was added, up to 50% by volume, to the crystallization drop for cryoprotection. X-ray diffraction data were collected at a wavelength of 1.4538 Å using the Synchrotron Radiation Source (Station PCr, Laboratório Nacional de Luz Síncrotron, LNLS, Campinas, Brazil) and a 34.5 cm MAR imaging plate detector (MAR Research) with an exposure time of 10 min per image at a crystal to detector distance of 200 mm. X-ray diffraction data were processed to 2.3 Å resolution using the program DENZO and scaled with the program SCALEPACK [12].

Upon cooling the cell parameters shrank from $a = b = 142.90$ Å, $c = 165.20$ Å to $a = b = 141.63$ Å, and $c = 163.16$ Å, a reduction of 3% in the volume of unit cell.

Crystal structure. The crystal structure of the HsPNP was determined by standard molecular replacement methods using the program AMoRe [13], using as search model the structure of HsPNP (PDB access code: 1ULA) [4]. Structure refinement was performed using X-PLOR [14]. The atomic positions obtained from molecular replacement were used to initiate the crystallographic refinement.

Root-mean-square deviation (R.m.s.d) differences from ideal geometries for bond lengths, angles, and dihedrals were calculated with X-PLOR 3.1 [14] and are presented in Table 2. The overall stereochemical quality of the final model for PNP was assessed by the program PROCHECK [15]. Atomic models were superposed using the program LSQKAB from CCP4 [16].

Results and discussion

Molecular replacement and crystallographic refinement

The standard procedure of molecular replacement using AMoRe [13] was used to solve the structure. After translation function computation the correlation was of 75.4% and the R_{factor} of 33.4%. The highest magnitude of the correlation coefficient function was obtained for the Euler angles $\alpha = 113.92°$, $\beta = 57.49°$, and $\gamma = 338.28°$. The fractional coordinates are $T_x = 0.4971$, $T_y = 0.2910$, and $T_z = 0.2003$. In the following rigid-body refinement, using X-PLOR, the R_{factor} decreased from 33.3% to 30.7%, using the same resolution range. At this stage $2F_{\text{obs}}$–F_{calc} maps were calculated. These maps showed clear electron density for the PNP structure. Further refinement in X-PLOR continued with simulated annealing using the slow-cooling protocol, followed by alternate cycles of positional refinement and manual rebuilding using XtalView [17]. Finally, the positions of water and sulfate molecules were checked and corrected in F_{obs}–F_{calc} maps. The final model has an R_{factor} of 20.5% and an R_{free} of 22.2%, with 68 water molecules and three sulfate ions. The HsPNP consists of 2251 non-hydrogen protein atoms.

Quality of the model

Analysis of the Ramachandran diagram ϕ–ψ plot for the present structure indicates that 87.7% of the residues are found to occur in the most favored regions, 11.9% in the additional allowed regions, and only one residue in the disallowed region of the plot. Analysis of the electron-density map ($2F_{\text{obs}} - F_{\text{calc}}$) agrees with the Thr221 positioning. The same analysis for two previously solved crystallographic PNP structures presents 73.9% of residues in the most favorable, 23.1% in the additional allowed regions, 1.6% in the generously allowed regions, and 1.4% in the disallowed region, which shows that the present structure has better overall stereochemistry.

Ignoring low-resolution data, a Luzzati plot [18] gives the best correlation between the observed and calculated data for a predicted mean coordinate error of 0.26 Å. The average B factor for main chain atoms is 16.21 Å^2, whereas that for side chain atoms is 17.46 Å^2. B factors for water molecules range from 6.88 to 46.75 Å^2, with an average of 29.35 Å^2 (Table 1).

Quaternary structure

Analysis of the crystallographic structures of HsPNP indicates a trimeric structure. However, in a number of instances the quaternary structure observed in the crystalline state is not conserved in solution [19]. Furthermore, in the case of HsPNP the low pH used in the crystallization condition may indicate that the crystallographic structure is in an environment distant to the physiological conditions. In addition, since the active site of the PNP is located near the interface of two subunits within the trimer, the precise information about the biological unit in solution is of capital importance to guide the structure-based design of inhibitors. Unpublished results of the quaternary structure of HsPNP in solution using SAXS indicated that the structure is also trimeric in solution in physiological conditions.

Overall description

Each PNP monomer displays an α-/β-fold consisting of a mixed β-sheet surrounded by α helices. The

Table 1
Data collection and refinement statistics

a (Å)	141.63
b (Å)	141.63
c (Å)	163.16
α (°)	90.00
β (°)	90.00
γ (°)	120.00
Space group	R32
Number of measurements with $I > 2\sigma(I)$	48,682
Number of independent reflections	26,429
Completeness in the range from 20.25 to 2.30 Å (%)	94.13
R_{sym}[a] (%)	8.7
Highest resolution shell (Å)	2.40–2.30
Completeness in the highest resolution shell (%)	87.6
R_{sym}[a] in the highest resolution shell (%)	19.7
Resolution range used in the refinement(Å)	8.0–2.3
R_{factor}[b] (%)	20.5
R_{free}[c] (%)	22.2
B values[d] (Å^2)	
Main chain	16.21
Side chains	17.46
Waters	29.35
Sulfate groups	23.09
Observed r.m.s.d. from ideal geometry	
Bond lengths (Å)	0.009
Bond angles (degrees)	1.48
Dihedrals (degrees)	25.18
No. of water molecules	68
No. of sulfate groups	3

[a] $R_{sym} = 100\ \Sigma|I(h) - < I(h) > /\Sigma I(h)$ with $I(h)$, observed intensity and $\langle I(h)\rangle$, mean intensity of reflection h overall measurement of $I(h)$.
[b] $R_{factor} = 100 \times \Sigma|F_{obs} - F_{calc}|/\Sigma(F_{obs})$, the sums being taken overall reflections with $F/\Sigma(F) > 2$ cutoff.
[c] $R_{free} = R_{factor}$ for 10% of the data, which were not included during crystallographic refinement.
[d] B values = average B values for all non-hydrogen atoms.

structure contains an eight-stranded mixed β-sheet and a five-stranded mixed β-sheet, which join to form a distorted β-barrel. The residues making up the eight-stranded sheet are 27–32, 43–50, 67–74, 76–83, 110–120, 129–138, 188–195, and 234–245. The five-stranded sheet consists of residues 116–120, 135–138, 192–196, 215–219, and 242–244. Seven α-helices surround the β-sheet structure. The α-helices are composed of residues 7–19, 36–42, 93–105, 168–182, 203–213, 222–231, and 262–282. Figs. 1A and B show schematic drawings of the PNP structure (trimer and monomer).

The contact area at interface between each subunit is 1124 Å^2, which indicates that the subunits are strongly bound to each other in the crystalline state. Analysis of the electrostatic potential surface at the subunit interface indicates good shape complementarity and some charge complementarity (figure not shown); however, most of the contacts are hydrophobic and involve residues Tyr88, Phe141, Phe159, Phe200, and Leu209.

Comparison with other PNPs

The amino acid sequence of HsPNP is compared to those of other PNPs in Fig. 2. Table 2 shows the R.m.s.d. and of the equivalent α-carbon atoms after superposition and identity of sequences. The structural similarity correlates with the similarity in the sequences for PNP isolated from non-human sources.

Ligand-binding conformational changes

The high-resolution structure of the PNP apoenzyme provides a more reliable structural model to assess conformational changes upon ligand binding. Analysis of the superposition between the present structure and human PNP complexed with guanine (PDB access code: 1ULB) indicates that there is a conformational change in the PNP structure upon binding of guanine to the active site. The overall change is relatively small, with an r.m.s.d. in the coordinates of all Cα of 1.4 Å after superposition. The largest movement was observed for His257 side chain, which partially occupies the purine subsite in the native enzyme. The residues 241–260 act as

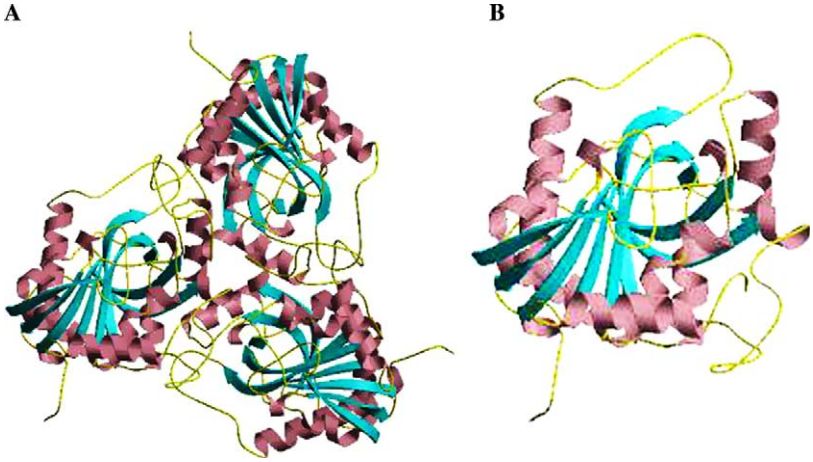


Fig. 1. Ribbon diagrams of HsPNP generated by Molscript [34] and Raster3d [35] for (A) trimer and (B) monomer.

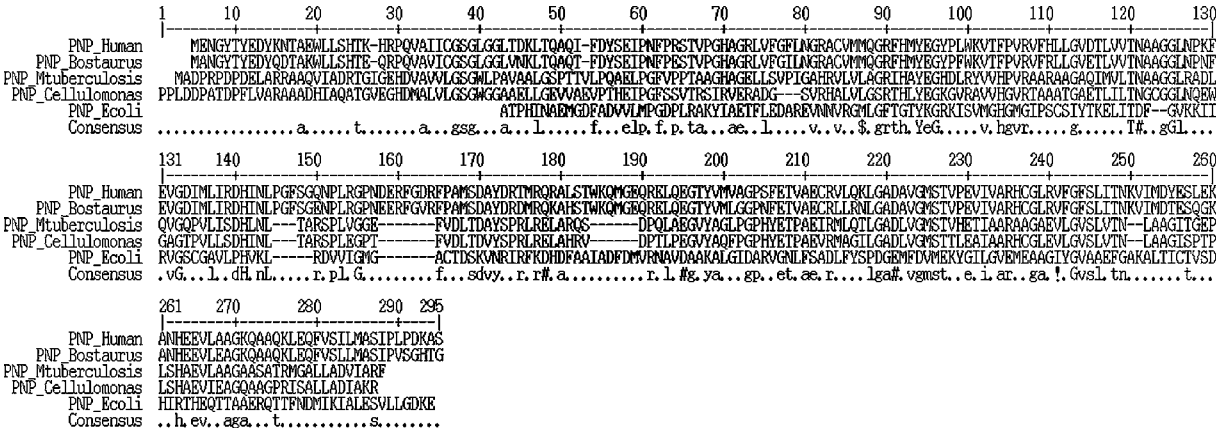


Fig. 2. Sequence alignment of HsPNP, BtPNP, PNP isolated from *Mycobacterium tuberculosis*, PNP from *Cellulomonas* sp., and PNP from *Escherichia coli.* The multiple alignment was performed using MULTALIGN [36].

Table 2
R.m.s. deviations after superposition of PNPs against atomic coordinates of 1M73

PDB acess code	Organism	Sequence identity with recombinant HsPNP (%)	R.m.s.d. (Å)
1ULA	*Homo sapiens*	99.7	0.74
1ULB	*Homo sapiens*	99.7	1.40
1B8O	*Bos taurus*	86.5	0.53
1G2O	*Mycobacterium tuberculosis*	37.8	1.21
1C3X	*Cellulomonas sp.*	35.0	3.14
1A69	*Escherichia coli*	14.7	3.38

a gate that opens during substrate binding. This gate is anchored near the central β-sheet at one end and near the C-terminal helix at the other end and it is responsible for controlling access to the active site. The gate movement involves a helical transformation of residues 257–265 in the transition apoenzyme complex. Comparison of the PNP apoenzyme with bovine PNP (BtPNP) complexed with immucillin-H also shows the gate movement and the helical transition in the same region.

Especially interesting is the position occupied by the Lys244 in the present structure. The hydrogen bond between the Lys244NZ and O6 of the purine ring was previously predicted from molecular modeling studies [20], since electron density was not detected for the Lys244 side-chain atoms in the low-resolution PNP structures [3,4]. The higher-resolution structure of HsPNP, here described, presents clear electron density for the Lys244 region. The ε-amino group of Lys244 forms hydrogen bonds with the carbonyl groups of Phe124 (3.3 Å) and Asn121 (2.7 Å), in the present structure, indicating that the side-chain of Lys244 is firmly locked in this region. Similar positioning for Lys244 is observed in the high-resolution structures of BtPNP [21,22]. Analysis of the same region in the low-resolution structure of HsPNP complexed with guanine (PDB access code: 1ULB) indicates a large movement in the Lys244 side chain upon binding of guanine, suggesting that the ε-amino group of Lys244 in the HsPNP moves 9.1 Å upon guanine binding.

Based on the high-resolution structures of HsPNP and BtPNP we propose that the binding of guanine and inhibitors to mammalian PNP does not generate large movement of Lys244 side chain, nor allows hydrogen bonding between ε-amino group of Lys244 and substrate, as speculated in previous reports [4,20]. Furthermore, the predicted salt-bridge between Lys244 and Glu201 is also unlikely to occur, since it was not observed, neither in the present structure nor in the high-resolution structures of BtPNP.

The present work strongly indicates that all previous modeling studies that showed the participation of the side chain of Lys244 in substrate binding should be revised [4,7,20,23]. Furthermore, analysis of Lys244Ala mutant exhibited kinetic parameters similar to the wild-type enzyme [24,25], which also indicates that Lys244 has no participation in substrate binding and catalysis. The authors of this work believe that such misjudgment of hydrogen-bonding pattern between substrate and HsPNP is due mainly to the low-resolution character of the previously determined structure and the poor stereochemical quality of the models, since Lys244 is located in the disallowed region of the Ramachandran plot for HsPNP complexed with guanine.

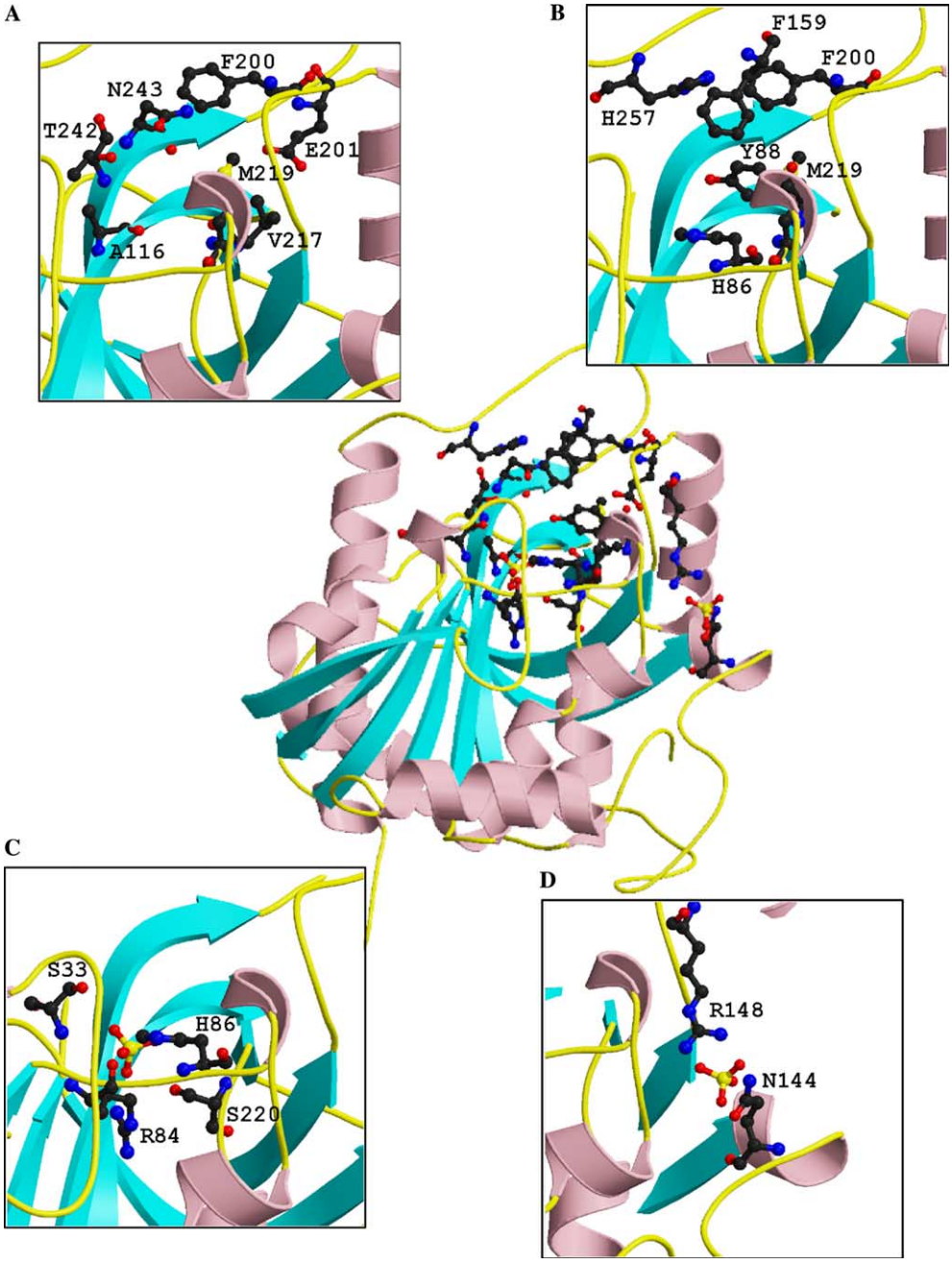


Fig. 3. Ligand-binding sites generated by Molscript [34] and Raster3d [35]. (A) Purine-binding site. (B) Ribose-binding site. (C) First phosphate-binding site. (D) Second regulatory phosphate-binding site.

Second phosphate regulatory binding site

The high-resolution structure of HsPNP shows clear electron-density peaks for three sulfate groups, which is present in high concentration in the crystallization experimental condition. Two of these sulfate groups have been previously identified in the low-resolution structure [3] and one new site was identified at subunit interface in the present structure. The first sulfate site, which is the catalytic phosphate-binding site, is positioned to form hydrogen bonds to Ser33, Arg84, His86, and S220. The second sulfate-binding site lies near Leu35 and Gly36 and is exposed to the solvent and whether it is mechanistically significant or an artifact resulting from the high sulfate concentration used to grow the crystals is not known. The third identified sulfate group makes three hydrogen bonds, involving residues Gln144 (3.21 Å) and Arg148 (2.68 and 3.22 Å) from adjacent subunit. A previous study of BtPNP activity as a function of phosphate concentration strongly indicates the presence of a second phosphate-binding site in the enzyme that may play a regulatory role [26]. Based on this

result we propose that the third phosphate-binding site identified in the present structure is the putative second regulatory phosphate-binding site.

It has also been proposed that the higher IC_{50} observed in experiments in higher phosphate concentrations may be explained by direct competition of phosphate groups and inhibitor for the phosphate-binding site (e.g., acyclovir diphosphate), while in other cases the reason is not so obvious [27]. It is tempting to speculate that the presence of a phosphate group in the second regulatory phosphate-binding site may offer further hindrance to the binding of substrates, which may also contribute to the larger IC_{50} observed for several inhibitors in the presence of higher phosphate concentrations.

An updated view of the active site

Analysis of the high-resolution structures HsPNP (present work) and BtPNP [21] strongly suggests that the composition of the purine-binding site needs to be updated. These structures indicate that the purine-binding site is composed of residues Ala116, Phe200, Glu201, Val217, Met219, Thr242, and Asn243. Ealick et al. [3] stated that Ala116 is located on one side of the purine ring and Phe200 on the other. The previously determined participation of Lys244 [4] in purine binding has not been identified in the present study. The ribose-binding site is mostly hydrophobic, which is composed of several aromatic amino acids, including Tyr88, Phe159 (of the adjacent subunit), Phe200, His86, His257, and Met219. Remarkably though, the 2.8 Å hydrogen bond between the 02′ oxygen and the amide of Met219 seems to be the only enzymatic contact relevant to ribosyl ring migration in the reaction coordinate motion [21]. Two phosphate-binding sites are present in the structure, one located near a glycine-rich loop (residues 32–37) is composed of residues Ser33, Arg84, His86, and Ser220 and a second at subunit interface near residues Gln144 and Arg148 (of the adjacent subunit). These phosphate-binding sites correspond, respectively, to the catalytic and regulatory sites in HsPNP. Fig. 3 shows the positions of all binding sites.

Implications for PNP inhibitor design

The specificity and affinity between enzyme and its inhibitor depend on directional hydrogen bonds and ionic interactions, as well as on shape complementarity of the contact surfaces of both partners [28–31]. The description of the PNP–guanine complex structure provides a list of potentially important interactions for ligand binding. However, other information such as binding affinities of ligand analogs, conformational flexibility of the ligand, and solvation of residues in the binding pocket of the apoenzyme is needed to identify functionally essential interactions. The purine interactions with PNP are mostly of hydrophobic nature with residues that are conserved throughout the PNP family, nevertheless several hydrogen bonds, between PNP and guanine, are observed. The fact that these hydrogen bonds are formed in a hydrophobic environment, which increases the interaction energy between two dipoles, most likely contribute to their importance in the overall binding affinity. The binding pocket for the purine in the apoenzyme shows two well-defined water molecules whose density of one water molecule overlaps with the purine ring in the superimposed guanine complex. This water molecule is substituting for guanine N1 atom forming a hydrogen bond with Glu201OE2, thereby confirming the importance of a hydrogen-bonding partner for Glu201OE2. Furthermore, the high-resolution crystallographic structures of complexes between PNP and immucillin-H show intermolecular hydrogen bonds involving the side chain of residue Glu201 in the complex structure [21,32]. In addition, kinetic data of purine-binding site mutants indicated that Glu201Ala mutant proved to be a less efficient enzyme relative to the wild type with a 2800-fold decrease in k_{cat}/K_m value in the phosphorolytic direction due to changes in both k_{cat} and K_m [20]. Protonation of this group prevents the substrate binding interaction at N1 and the proton-transfer bridge responsible for O6 protonation at the transition state [21,27]. The second water is close to Phe159 and Phe200 at subunit interface. Consistent with a role for Phe200, Phe200Ala mutation was shown to result in a large decrease in catalytic efficiency due predominantly to an increase in the K_m for inosine. Phe159 has been implicated in ribose subsite recognition, since the Phe159Ala mutant showed little change in k_{cat} value and an 8-fold increase in K_m value for inosine [20]. The large contribution of the guanine to binding affinity of several PNP inhibitors is probably due to the formation of essential buried hydrogen bonds and numerous van der Waals contacts in combination with smaller entropic cost of binding the rigid purine ring, as observed in the structure of CDK2 complexes [33].

Conclusions

Analysis of the high-resolution structure of PNP apoenzyme provides important information for the structure-based design of new drugs and the improvement of already identified lead compounds.

1. The use of cryocrystallography techniques and recombinant human PNP allowed an improvement in the resolution of 0.5 Å as compared to the HsPNP structures solved using crystals that were not frozen. Furthermore, the overall stereochemical quality for the present structure is considerably superior than the one obtained for the two previously determined

structures of HsPNP, which indicates that the cryogenic conditions used to collect the X-ray diffraction data and the recombinant human PNP not only increased the resolution but also increased substantially the overall stereochemical quality of the structure.
2. The water structure of the active site suggests binding partners for potential ligands.
3. Ligand-induced conformational changes in the protein are difficult to predict and need to be determined experimentally. In the case of PNP, there is a large movement of residues 240–260. These residues form a gate that opens during substrate binding. Furthermore, the present structure strongly indicates that previous molecular modeling studies of PNP need to be revised, since they predicted participation of the ε-amino group of Lys244 in substrate binding, such interaction is not observed in the high-resolution structure of BtPNP in complex with immucillin-H, which presents the side-chain of Lys244 in a conformation close to that observed in the present structure.
4. The identification of a second regulatory phosphate-binding site partially explains the phosphate dependency of IC_{50} observed for several PNP inhibitors.

Acknowledgments

We acknowledge the expertise of Denise Cantarelli Machado for the expansion of the cDNA library and Deise Potrich for the DNA sequencing. This work was supported by grants from FAPESP (SMOLBNet, Proc. 01/07532-0), CNPq, CAPES and Instituto do Milênio (CNPq-MCT). WFA (CNPq, 300851/98-7), MSP (CNPq, 500079/90-0), and LAB (CNPq, 520182/99-5) are researchers for the Brazilian Council for Scientific and Technological Development.

References

[1] J.D. Stoeckler, C. Cambor, R.E. Parks Jr., Human erythrocytic purine nucleoside phosphorylase: reaction with sugar-modified nucleosides substrates, Biochemistry 19 (1980) 102–107.
[2] D.J.T. Porter, Purine nucleoside phosphorylase. Kinetic mechanism of the enzyme from calf spleen, J. Biol. Chem. 267 (1992) 7342–7351.
[3] S.E. Ealick, S.A. Rule, D.C. Carter, T.J. Greenhough, Y.S. Babu, W.J. Cook, J. Habash, J.R. Helliwell, J.D. Stoeckler, R.E. Parks Jr., S.-F. Chen, C.E. Bugg, Three-dimensional structure of human erythrocytic purine nucleoside phosphorylase at 3.2 Å resolution, J. Biol. Chem. 265 (3) (1990) 1812–1820.
[4] S.E. Ealick, Y.S. Babu, C.E. Bugg, M.D. Erion, W.C. Guida, J.A. Montgomery, J.A. Secrist III, Application of crystallographic and modeling methods in the design of purine nucleoside phosphorylase inhibitors, Proc. Natl. Acad. Sci. USA 91 (1991) 11540–11544.
[5] P.W.K Woo, C.R. Kostlan, J.C. Sircar, M.K. Dong, R.B. Gilbertsen, Inhibitors of human purine nucleoside phosphorylase. Synthesis and biological activities of 8-amino-3-benzylhypoxanthine and related analogues, J. Med. Chem. 35 (1992) 1451–1457.
[6] J.-W. Chern, H.-Y. Lee, C.-S. Chen, Nucleosides. 5. Synthesis of guanine and formycin B derivatives as potential inhibitors of purine nucleoside phosphorylase, J. Med. Chem. 36 (1993) 1024–1031.
[7] W.C. Guida, R.D. Elliot, H.J. Thomas, J.A. Secrist III, Y.S. Babu, C.E. Bugg, M.D. Erion, S.E. Ealick, J.A. Montgomery, Structure-based design of inhibitors of purine nucleoside phosphorylase. 4. A study of phosphate mimics, J. Med. Chem. 37 (1994) 1109–1114.
[8] S. Niwas, P. Chand, V.P. Pathak, J.A. Montgomery, Structure-based design of inhibitors of purine nucleoside phosphorylase. 5. 9-Deazahypoxanthines, J. Med. Chem. 37 (1994) 2477–2480.
[9] P.E. Morris, A.J. Elliott, S.P. Walton, C.H. Williams, J.A. Montgomery, Synthesis and biological activity of a novel class of purine nucleoside phosphorylase inhibitors, Nucleosides, Nucleotides & Nucleic Acids 19 (2000) 379–404.
[10] R.G. Silva, L.P. Carvalho, J.S. Oliveira, C.A. Pinto, M.A. Mendes, M.S. Palma, L.A. Basso, D.S. Santos, Cloning, overexpression, and purification of functional human purine nucleoside phosphorylase, Protein Expr. Purif. 27 (1) (2003) 158–164.
[11] W.J. Cook, S.E. Ealick, C.E. Bugg, J.D. Stoeckler, R.E. Parks Jr., Crystallization and preliminary X-ray investigation of human erythrocytic purine nucleoside phosphorylase, J. Biol. Chem. 256 (1981) 4079–4080.
[12] D. Gewirth, The HKL Manual, Yale University, New Haven, CT 06511, USA, 1965.
[13] J. Navaza, AMoRe: an automated package for molecular replacement, Acta Cryst. A 50 (1994) 157–163.
[14] A.T. Brünger, X-PLOR Version 3.1: A System for Crystallography and NMR, Yale University Press, New Haven, 1992.
[15] R.A. Laskowski, M.W. MacArthur, D.K. Smith, D.T. Jones, E.G. Hutchinson, A.L. Morris, D. Naylor, D.S. Moss, J.M. Thorton, PROCHECK v.3.0—program to check the stereochemistry quality of protein structures—operating instructions, 1994.
[16] Collaborative computational project, No. 4, The CCP4 suite: programs for protein crystallography, Acta Crystallogr. D 50 (1994) 760–763.
[17] D.E. McRee, XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density, J. Struct. Biol. 125 (1999) 156–165.
[18] P.V. Luzzati, Traitement statistique des erreurs dans la determination des structures cristallines, Acta Crystallogr. 5 (1952) 802–810.
[19] D.I. Svergun, C. Barberato, M.H.J. Koch, L. Fetler, P. Vachette, Large differences are observed between the crystal and solution quaternary structures of allosteric aspartate transcarbamylase in the R state, Proteins Struct. Funct. Genet. 27 (1997) 110–117.
[20] M.D. Erion, K. Takabayashi, H.B. Smith, J. Kessi, S. Wagner, S. Hönger, S.L. Shames, S.E. Ealick, Purine nucleoside phosphorylase. 1. Structure–function studies, Biochemistry 36 (1997) 11725–11734.
[21] A. Fedorov, W. Shi, G.A. Kicska, E.V. Fedorov, P.C. Tyler, R.H. Furneaux, J.C. Hanson, G.J. Gainsford, J.Z. Larese, V.L. Schramm, S.C. Almo, Transition state structure of purine nucleoside phosphorylase and principles of atomic motion in enzymatic catalysis, Biochemistry 40 (2001) 853–860.
[22] C. Mao, W.J. Cook, M. Zhou, A.A. Federov, S.C. Almo, S.E. Ealick, Calf spleen purine nucleoside phosphorylase complexed with substrates and substrates analogues, Biochemistry 37 (1998) 7135–7146.
[23] J.A. Montgomery, Purine nucleoside phosphorylase: a target for drug design, Med. Res. Rev. 13 (3) (1993) 209–228.
[24] M.D. Erion, J.D. Stoeckler, W.C. Guida, R.L. Walter, S.E. Ealick, Purine nucleoside phosphorylase. 2. Catalytic mechanism, Biochemistry 36 (1997) 11735–11748.
[25] J.D. Stoeckler, A.F. Poirot, R.M. Smith, R.E. Parks Jr., S.E. Ealick, K. Takabayashi, M.D. Erion, Purine nucleoside phosphorylase. 3. Reversal of purine base specificity by site-directed mutagenesis, Biochemistry 36 (1997) 11749–11756.

[26] P.A. Ropp, T.W. Traut, Allosteric regulation of purine nucleoside phosphorylase, Arch. Biochem. Biophys. 288 (2) (1991) 614–620.
[27] L.A. Basso, D.S. Santos, W. Shi, R.H. Furneaux, P.C. Tyler, V.L. Schramm, J.S. Blanchard, Purine nucleoside phosphorylase from Mycobacterium tuberculosis. Analysis of inhibition by a transition-state analogue and dissection by parts, Biochemistry 40 (2001) 8196–8203.
[28] F. Canduri, L.G.V.L. Teodoro, C.C.B. Lorenzi, V. Hial, R.A.S. Gomes, J. Ruggiero Neto, W.F. de Azevedo Jr., Crystal structure of human uropepsin at 2.45 Å resolution, Acta Crystallogr. D 57 (2001) 1560–1570.
[29] W.F. de Azevedo Jr., F. Canduri, V. Fadel, L.G.V.L. Teodoro, V. Hial, R.A.S. Gomes, Molecular model for the binary complex of uropepsin and pepstatin, Biochem. Biophys. Res. Commun. 287 (1) (2001) 277–281.
[30] W.F. de Azevedo Jr., F. Canduri, N.J.F. da Silveira, Structural basis for inhibition of cyclin-dependent kinase 9 by flavopiridol, Biochem. Biophys. Res. Commun. 293 (2002) 566–571.
[31] W.F. de Azevedo Jr., J.S. de Oliveira, L.A. Basso, M.S. Palma, J.H. Pereira, F. Canduri, D.S. Santos, Molecular model of Shikimate kinase from *Mycobacterium tuberculosis*, Biochem. Biophys. Res. Commun. 295 (1) (2002) 142–148.
[32] W. Shi, L.A. Basso, D.S. Santos, P.C. Tyler, R.H. Furneaux, J.S. Blanchard, S.C. Almo, V.L. Schramm, Structures of purine nucleoside phosphorylase from *Mycobacterium tuberculosis* in complexes with immucillin-H and its pieces, Biochemistry 40 (28) (2001) 8204–8215.
[33] U. Schulze-Gahmen, H.L. De Bondt, S.H. Kim, High-resolution crystal structures of human cyclin-dependent kinase 2 with and without ATP: bound waters and natural ligand as guides for inhibitor design, J. Med. Chem. 39 (23) (1996) 4540–4546.
[34] P.J. Kraulis, MOLSCRIPT: a program to produce both detailed and schematic plots of proteins, J. Appl. Crystallogr. 24 (1991) 946–950.
[35] E.A. Merritt, D.J. Bacon, Raster3D: photorealistic molecular graphics, Methods Enzimol. 277 (1997) 505–524.
[36] F. Corpet, Multiple sequence alignment with hierarchical clustering, Nucleic Acids Res. 16 (1988) 10881–10890.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 320 (2004) 979–991
BBRC
www.elsevier.com/locate/ybbrc

Molecular models for shikimate pathway enzymes of *Xylella fastidiosa*

Helen Andrade Arcuri,[a,1] Fernanda Canduri,[a,d,1] José Henrique Pereira,[a] Nelson José Freitas da Silveira,[a] João Carlos Camera Jr.,[a] Jaim Simões de Oliveira,[b] Luiz Augusto Basso,[b] Mário Sérgio Palma,[c,d] Diógenes Santiago Santos,[e,*] and Walter Filgueira de Azevedo Jr.[a,d,*]

[a] *Department of Physics IBILCE/UNESP, São José do Rio Preto, SP 15054-000, Brazil*
[b] *Rede Brasileira de Pesquisas em Tuberculose, Department of Molecular Biology and Biotecnology, UFRGS, Porto Alegre, RS 91501-970, Brazil*
[c] *Laboratory of Structural Biology and Zoochemistry, Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil*
[d] *Center for Applied Toxicology, Institute Butantan, São Paulo, SP 05503-900, Brazil*
[e] *Center for Research and Development in Molecular, Structural and Functional Molecular Biology, PUCRS 90619-900, Porto Alegre, RS, Brazil*

Received 25 May 2004
Available online 25 June 2004

Abstract

The *Xylella fastidiosa* is a bacterium that is the cause of citrus variegated chlorosis (CVC). The shikimate pathway is of pivotal importance for production of a plethora of aromatic compounds in plants, bacteria, and fungi. Putative structural differences in the enzymes from the shikimate pathway, between the proteins of bacterial origin and those of plants, could be used for the development of a drug for the control of CVC. However, inhibitors for shikimate pathway enzymes should have high specificity for *X. fastidiosa* enzymes, since they are also present in plants. In order to pave the way for structural and functional efforts towards antimicrobial agent development, here we describe the molecular modeling of seven enzymes of the shikimate pathway of *X. fastidiosa.* The structural models of shikimate pathway enzymes, complexed with inhibitors, strongly indicate that the previously identified inhibitors may also inhibit the *X. fastidiosa* enzymes.

Keywords: Shikimate pathway; Structural bioinformatics; Drug design; *Xylella fastidiosa*

Xylella fastidiosa (*Xf*) is a fastidious and gram-negative, xylem-limited bacterium which causes the citrus variegated chlorosis (CVC), a serious disease of citrus fruit [1]. The disease spread rapidly by graft propagation with infected budwood and by sharpshooter vectors, and became widely distributed in all citrus growing regions of Brazil. CVC is one of the major concerns to the Brazilian citrus industry. Certain strains of *Xf* also affect other commercially important products such as coffee, nuts, and other fruits [2].

In plants and microorganisms, all the key aromatic compounds involved in primary metabolism, including the three aromatic amino acids found in proteins, are produced via shikimate pathway. The shikimate pathway is essential in algae, higher plants, bacteria, and fungi, but absent from mammals that depend on these compounds for your diet [3].

Xylella fastidiosa appears to be capable of synthesizing an extensive variety of enzyme cofactors and prosthetic groups, including biotin, folic acid, pantothenate and coenzyme A, ubiquinone, glutathione, thioredoxin, glutaredoxin, riboflavin, flavin mononucleotide (FMN), flavin adenine dinucleotide (FAD), pyrimidine nucleotides, porphyrin, thiamin, pyridoxal 5′-phosphate, and lipoate [1].

The shikimate pathway consists of seven enzymes that catalyze the sequential conversion of erythrose-4-phosphate and phosphoenolpyruvate to chorismate [4] (Fig. 1). All pathway intermediates can also be considered branch point compounds that may serve as substrates for other metabolic pathways [5].

* Corresponding author.
E-mail address: walterfa@df.ibilce.unesp.br (W.F. de Azevedo Jr.).
[1] H.A.C. and F.C. contributed equally to this work.

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.05.220

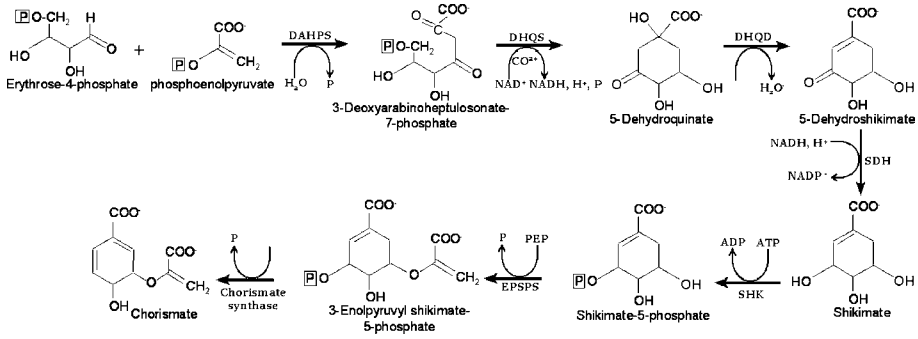


Fig. 1. The shikimate pathway in the sequence of seven metabolic steps, from phosphoenolpyruvate and erythrose 4-phosphate until the conversion to chorismate.

The molecular organization and structure of the shikimate pathway enzymes varies considerably between microrganism groups. Bacteria have seven individual polypeptides, each possessing a single enzyme activity, which are encoded by separated genes. Plants have a molecular arrangement similar to bacteria, i.e., separated enzymes encoded by separated genes, with the exception of dehydroquinase and shikimate dehydrogenase, which have been shown to be present as separated domains on a bifunctional polypeptide. Plant enzymes, although nuclear encoded, are largely active in the chloroplast and accordingly possess an N-terminal transit sequence. In contrast, all fungi examined to date have monofunctional 3-deoxy-D-arabino-heptulosonate-7-phosphate synthases and chorismate synthases and a pentafunctional polypeptide termed AROM. The AROM polypeptide has domains analogous to the bacterial enzymes: dehydroquinate synthase (DHQS), 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), shikimate kinase (SK), dehydroquinase (DHQD), and shikimate dehydrogenase (SDH) [4].

Chorismate is converted by five distinct enzymes to prephenate, anthranilate, aminodeoxychorismate, isochorismate, and *p*-hydroxybenzoate. These metabolites comprise the first committed intermediates in the biosynthesis of Phe, Tyr, Trp, folate, menaquinone, and the siderophore enterobactin, and ubiquinone, respectively. The synthesis of these precursors is in most cases highly regulated [6].

In monogastric animals, Phe and Trp are essential amino acids that have to come with the diet and Tyr is directly derived from Phe. Since bacteria use in excess 90% of their metabolic energy for protein biosynthesis, for most prokaryotes, the three aromatic amino acids represent nearly the entire output of aromatic biosynthesis, and regulatory mechanisms for shikimate pathway activity are triggered by the intracellular concentrations of Phe, Tyr, and Trp. This is not so in higher plants, in which the aromatic amino acids are the precursors for a large variety of secondary metabolites with aromatic ring structures that often make up a substantial amount of the total dry weight of a plant. Among the many aromatic secondary metabolites are flavonoids, many phytoalexins, indole acetate, alkaloids such as morphine, UV light protectants, and, most important, lignin [7].

In microrganisms, the shikimate pathway is regulated both by feedback inhibition and by repression of the first enzyme 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHPS). In higher plants, no physiological feedback inhibitor has been identified, suggesting that pathway regulation may occur exclusively at the genetic level. This difference between microrganisms and plants is reflected in the unusually large variation in the primary structures of the respective first enzymes [5].

Shikimate pathway enzymes are attractive targets for drug development, however any new inhibitor for shikimate pathway enzymes should have high specificity for *Xf* enzymes, since they are also present in the plants. Therefore, detailed structural information about the protein targets will help us to identify structural differences between enzymes of plant and *Xf*, which may serve as basis to design specific inhibitor against shikimate pathway enzymes of the *X. fastidiosa.* The seven enzymes modeled in the present work were 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (DAHPS), 3-dehydroquinate synthase (DHQS), 3-dehydroquinate dehydratase (DHQD), shikimate-5-dehydrogenase (SDH), shikimate kinase (SK), 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), and chorismate synthase (CS).

Knowledge of the three-dimensional structures of the enzymes will undoubtedly aid in the design of useful inhibitors that may be used as a bactericide against *X. fastidiosa.*

Methods

Molecular modeling. Homology modeling is usually the method of choice when there is a clear relationship of homology between the sequence of a target protein and at least one known structure. This computational technique is based on the assumption that tertiary structures of two proteins will be similar if their sequences are related, and it is the approach most likely to give accurate results [8]. There are two main approaches to homology modeling: (1) fragment-based comparative modeling [9,10] and (2) restrained-based modeling [11]. For modeling of the shikimate pathway we used the second approach.

Model building of shikimate pathway enzymes was carried out using the program MODELLER [11]. MODELLER is an implementation of an automated approach to comparative modeling by satisfaction of spatial restraints [12–14]. The modeling procedure begins with an alignment of the sequence to be modeled (target) with related known three-dimensional structures (templates). This alignment is usually the input to the program. The output is a three-dimensional model for the target sequence containing all main-chain and sidechain non-hydrogen atoms.

Since there are no three-dimensional structures available for shikimate pathway enzymes of the *X. fastidiosa* (targets) the models were built based on the atomic coordinates of shikimate pathway enzymes (templates), which were solved by crystallographic methods (Table 1). The percentage of primary sequence identity (Table 1) between enzymes of *Xf* and of templates indicates that the crystallographic structures of enzymes (templates) are good models to be used as templates for the enzymes of *X. fastidiosa* (targets). The alignment of enzymes of the *X. fastidiosa* (targets) and of templates is shown in Figs. 2A–G.

Next, the spatial restraints and CHARMM energy terms enforcing proper stereochemistry [15] were combined into an objective function. Finally, the model is obtained by optimizing the objective function in Cartesian space. The optimization is carried out by the use of the variable target function method [16] employing method of conjugate gradients and molecular dynamics with simulated annealing. Several slightly different models can be calculated by varying the initial structure. A total of 1000 models were generated for each enzyme and the final models were selected based on stereochemical quality. All optimization process was performed on a Beowulf cluster with 16 nodes (BioComp, AMD Athlon XP 1800+).

Analysis of the models. The overall stereochemical quality of the final models for each enzyme of the shikimate pathway of the *X. fastidiosa* was assessed by the program PROCHECK [17]. The root mean square deviation (rmsd) between Cα–Cα atom's distance was superposed using the program LSQKAB from CCP4 [18]. The cutoff for hydrogen bonds and salt bridges was 3.3 Å. The contact area for the complexes was calculated using AREAIMOL and RESAREA [18]. The root mean square deviation (rmsd) differences from ideal geometries for bond lengths and bond angles were calculated with X-PLOR [19,20]. The *G*-factor value is essentially just log-odds score based on the observed distributions of the stereochemical parameters. It is computed for the following properties: torsion angles (the analysis provided the observed distributions of ϕ–ψ, χ1–χ2, χ-1, χ-3, χ-4, and ω values for each of the 20 amino acid types) and covalent geometry (for the main-chain bond lengths and bond angles). These values' average was calculated using PROCHECK [17]. The Verify-3D measures the compatibility of a protein model with its sequence, using a 3D profile [21–23].

Results and discussion

Quality of the models

The atomic coordinates of crystallographic structures (Table 1) solved to resolution better than 2.0 Å (with two exceptions) were used to build up an ensemble of structures to be used as starting models for modeling of the *X. fastidiosa* structures. The atomic coordinates of all waters were removed from the templates.

The analysis of the Ramachandran diagram ϕ–ψ plots for the templates (Table 2) was used to compare the overall stereochemical quality of the *X. fastidiosa* structures against that of templates solved by biocrystallography. They present over 89.0% of the residues in the most favorable regions.

Overall description and interactions with inhibitors

Among the seven enzymes modeled in the present work four proteins were modeled with inhibitors (DAHPS:Mn^{2+}:PGL, DHQS:CRB:NAD, DHQD:FA1, and EPSPS:glyphosate:S3P), because only these templates have inhibitors in their crystallographic structure.

The structures of all molecular models were superposed with their templates (Figs. 6A–G). The rmsd for each superposition is given in Table 1. It was considered only Cα in the superposition. The rmsd value of bond lengths and bond angles is presented in Table 2.

The superpositions of the complexes are shown in Figs. 5A–E and the average *G*-factor and Verify 3D values are shown in Table 3.

3-Deoxy-D-arabino-heptulosonate-7-phosphate synthase

The first step of the shikimate pathway is the condensation of phosphoenolpyruvate (PEP) with the four-carbon sugar D-erythrose-4-phosphate (E4P) to produce 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP)

Table 1
The crystallographic structures of the shikimate pathway enzymes used as templates, the percentage of amino acid identity, and rmsd value after superposition of templates and targets

Enzymes	PDB access code	Resolution (Å)	References	Organism	Sequence identity (%)	rmsd (Å)	
						Site	Protein
DAHPS[a]	1GG1	2.00	[24]	*Escherichia coli* (*Ec*)	61.1	0.086	0.665
DHQS[a]	1DQS	1.80	[26]	*Emericella nidulans* (*En*)	31.8	0.055	0.039
DHQD[a]	1GU1	1.80	[30]	*Streptomyces coelicolor* (*Sc*)	42.5	0.059	0.055
SDH[a]	1NYT	1.50	[32]	*E. coli* (*Ec*)	42.3	—	0.166
SK[a]	1L4Y	2.00	[38]	*Mycobacterium turberculosis* (*Mt*)	32.6	—	0.452
	1WE2	2.30	[39]	*M. turberculosis* (*Mt*)	32.6	—	0.366
	2SHK	2.60	[40]	*Erwinia chrysanthemi* (*Ech*)	31.9	—	0.377
EPSPS[a]	1G6S	1.50	[46]	*E. coli* (*Ec*)	29.9	1.419	0.118
CS[a]	1QXO	2.00	[49]	*Streptococcus pneumoniae* (*Sp*)	30.6	0.582	0.650

[a] DAHPS, 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; DHQS, 3-dehydroquinate synthase; DHQD, 3-dehydroquinate dehydratase; SDH, shikimate-5-dehydrogenase; SK, shikimate kinase; EPSPS, 5-enolpyruvylshikimate-3-phosphate synthase; and CS, chorismate synthase.

Fig. 2. (A) The sequence alignment for DAHPS from *X. fastidiosa* and *E. coli* (1GG1); (B) DHQS from *X. fastidiosa* and *E. nidulans* (1DQS); (C) DHQD from *X. fastidiosa* and *S. coelicolor*; (D) SDH from *X. fastidiosa* and *E. coli* (1NYT); (E) SK from *X. fastidiosa* and *Mycobacterium turberculosis* (1L4Y–1WE2), *E. chrysanthemi* (2SHK); (F) EPSPS from *X. fastidiosa* and *E. coli* (1G6S); and (G) CS from *X. fastidiosa* and *Streptococcus pneumoniae* (1QXO). The alignments were performed with the program CLUSTAL W [54], the meaning of the special characters in the last line being: (*) indicates positions which have a single, fully conserved, (:) indicates that one of the following 'strong' groups is fully conserved, and (.) indicates that one of the following 'weaker' groups is fully conserved.

and inorganic phosphate [5]. This reaction is catalyzed by the enzyme DAHPS, a metal-activated enzyme, which in microrganisms is the target for negative-feedback regulation by pathway intermediates or by end products. The basic structure of DAHP was confirmed by several different chemical syntheses [7]. Typically, the comparison of the primary amino acid sequence of plant with DAHPS shows a surprisingly low 20% pairwise identity. In fact, PRETTYBOX (GCG program package), a program used to make multiple sequence alignments, identifies only 24 invariant residues for all known DAHPS [5].

The Fig. 3A shows the secondary structure of *Xf*DAHPS in the complex with manganese (II) ion (Mn^{2+}) and 2-phosphoglycolate (PGL). This model shows that the protein has 8 (β/α) barrel enhanced by several additional secondary structure elements, including a C-terminal tail that forms a three-stranded, anti-parallel β-sheet with the β6a/β6b strands [24]. The active site of *Xf*DAHPS consists of a channel between C-terminal and N-terminal region of the 8 (β/α) barrel with bound Mn^{2+} and PGL.

We observed a total of 12 hydrogen bonds between *Xf*DAHPS:Mn^{2+}:PGL, in binary model, involving the

Fig. 2. (*continued*)

residues Arg92, Lys97, Ala164, Arg165, Lys186, Arg234, and His268. For the *Ec*DAHPS:Mn^{2+}:PGL 12 hydrogen bonds involving the residues Arg92, Lys97, Ala164, Arg165, Lys186, Arg234, and His268 were observed. The contact areas for the complexes of *Ec*-DAHPS:Mn^{2+}:PGL and *Xf*DAHPS:Mn^{2+}:PGL are 96 and 81 $Å^2$, respectively.

3-Dehydroquinate synthase

The second enzyme of the shikimate pathway, 3-dehydroquinate synthase (DHQS), requires catalytic amounts of NAD^+ and a divalent cation for activity. For the bacterial enzyme, Co^{2+} and Zn^{2+} are the most active metal ions [25].

DHQS is a NAD^+-dependent metalloenzyme that converts DAHP to dehydroquinate [26]. DHQS is activated by inorganic phosphate, one of the reaction products, and the enzyme also requires catalytic amounts of NAD^+ for activity, even though the enzyme catalyzed reaction is redox neutral. Conversion of DAHP to DHQ proceeds by way of an intramolecular exchange of the DAHP ring oxygen with carbon 7, driven by the cleavage of the phosphoester. The reaction involves an oxidation, a β-elimination of inorganic phosphate, a reduction, a ring opening, and an intramolecular aldol condensation [5].

The Fig. 3B shows the secondary structure of *Xf*DHQS in complex with [1*R*-(1α, 3β, 4α, 5β]-5-(phosphonomethyl)-1,3,4-trihydroxycyclohexane-1-carboxylic

Fig. 2. (*continued*)

acid (CRB) and nicotinamide-adenine-dinucleotide (NAD^+). The structure is composed of an N-terminal α/βdomain and a C-terminal α-helical domain [26].

The N-terminal domain includes a Rossmann fold, but surprisingly binds NAD^+ in an inverted orientation to that observed in all other classic Rossmann-fold proteins. This domain consists of a seven-stranded β-sheet, with strand order 1296534 (Fig. 3B). Strands $\beta1$ and $\beta2$ form a β-hairpin which is connected by helix $\alpha1$ to a five-stranded Rossmann-fold structure [27,28]. This conformation orientates the NAD^+ so that the active site is on the opposite side of the sheet to that seen in all other known Rossmann type NAD^+-binding proteins [26], comprising a $\beta\alpha\beta$ unit and a $\beta\alpha\beta\alpha\beta$ unit. Strands $\beta7$

Table 2
Analysis of the Ramachandran plot and rmsd from ideal geometry for the templates and models of *X. fastidiosa*

Enzymes	Region of the Ramachandran plot				rmsd	
	Most favorable (%)	Additional allowed (%)	Generously allowed (%)	Disallowed (%)	Bond lengths (Å)	Bond angles (°)
DAHPS	94.4 (90.5)	4.9 (8.1)	0.7 (1.4)	0.0 (0.0)	0.019 (0.012)	2.92 (1.50)
DHQS	91.2 (92.5)	7.2 (7.5)	1.3 (0.0)	0.3 (0.0)	0.020 (0.13)	3.70 (2.30)
DHQD	96.8 (92.2)	3.2 (7.0)	0.0 (0.8)	0.0 (0.0)	0.020 (0.019)	4.12 (1.80)
SDH	93.2 (93.3)	5.6 (6.7)	0.9 (0.0)	0.4 (0.0)	0.032 (0.018)	4.51 (1.80)
SK	93.8 (90.5)	5.6 (8.8)	0.6 (0.7)	0.0 (0.0)	0.020 (0.005)	3.63 (1.20)
	(93.4)	(5.9)	(0.7)	(0.0)	(0.017)	(1.90)
	(95.6)	(4.4)	(0.0)	(0.0)	(0.013)	(1.70)
EPSPS	89.0 (90.8)	7.1 (8.9)	2.5 (0.3)	1.4 (0.0)	—	—
CS	90.7 (93.1)	7.3 (6.9)	1.7 (0.0)	0.3 (0.0)	0.022 (0.013)	3.92 (2.58)

In parentheses are values obtained for the templates used for modeling.

Table 3
Analysis of the 3D profile and *G*-factor for the templates and models *X. fastidiosa*

Enzymes	3D profile[a]			*G*-factor[b]	
	Total score	Ideal score	S_{ideal} score	Torsion angles	Covalent geometry
DAHPS	135.23 (153.33)	159.94 (154.87)	0.845*S* (0.990*S*)	0.12 (0.18)	−0.10 (0.45)
DHQS	150.27 (180.65)	169.15 (173.76	0.888*S* (1.039*S*)	−0.05 (0.07)	−0.36 (−0.02)
DHQD	58.79 (67.47)	67.17 (67.62)	0.875*S* (0.998*S*)	−0.03 (0.04)	−0.46 (−0.17)
SDH	125.98 (131.02)	125.42 (123.58)	1.004*S* (1.060*S*)	−0.12 (−0.05)	−0.83 (0.31)
SK	69.16 (82.02)	81.82 (74.95)	0.854*S* (1.094*S*)	0.02 (0.30)	−0.21 (0.57)
	(83.33)	(74.95)	(1.112*S*)	(0.08)	(0.23)
	(68.25)	(71.74)	(0.953*S*)	(−0.01)	(−0.51)
EPSPS	175.97 (230.44)	196.36 (195.44)	0.896*S* (1.179*S*)	−0.21 (0.08)	−0.87 (0.31)
CS	145.42 (163.31)	170.07 (171.92)	0.855*S* (0.949*S*)	−0.16 (0.00)	−0.73 (−0.16)

In parentheses are values obtained for the templates used for modeling.

[a] Total score: sum of the 3D–1D scores (statistical preferences) of each residue present in protein. Ideal score: $S_{ideal} = \exp(-0.83 + 1.008x\ln(L))$; where L is number of amino acids S_{ideal}. Score: compatibility of the sequence with their 3D structure. It is obtained total score/ideal score. S_{ideal} score above $0.45S_{ideal}$.

[b] Ideally, scores should be above −0.5. Values below −1.0 may need investigation.

and β8 form a separated β-hairpin. Both DHQS and the classic Rossmann-fold domains bind NAD(P)$^+$ at the C-terminal end of the sheet [27,28].

However, uniquely among known NAD(P)$^+$-binding protein structures, the phosphate moieties of NAD$^+$ in DHQS are associated with the glycine-rich first turn of the second βαβαβ unit, rather than the first (βα)βαβ unit [26–28].

We observed a total of 16 hydrogen bonds in the binary complex *Xf*DHQS:NAD$^+$:CRB, involving the residues Asp143, Lys149, Asn159, Glu191, Lys194, Lys233, Arg247, Asn251, His254, His276, and Lys335. For the *En*DHQS:NAD$^+$:CRB 16 hydrogen bonds involving the residues Asp146, Lys152, Asn162, Glu194, Lys197, Lys250, Arg264, Asn268, His271, His287, and Lys356 were observed. The contact areas for the complexes of *En*DHQS:NAD$^+$:CRB and *Xf*DHQS:NAD$^+$:CRB are 582 and 566 Å^2, respectively.

3-Dehydroquinate dehydratase

The third step of the shikimate pathway, dehydration of DHQ to give 3-dehydroshikimate (DHS), is catalyzed by DHQD that exists in two forms: types I and II [5]. In higher plants, the third and fourth steps of the shikimate pathway are catalyzed by a bifunctional enzyme [25]. This enzyme belongs to the superfamily of NAD(P)H-dependent oxidoreductases, which function in anabolic and catabolic enzyme pathways as well as in xenobiotic detoxification. This superfamily is usually subdivided into several families, including short chain dehydrogenases, medium chain dehydrogenases, aldo-keto reductases, iron-activated alcohol dehydrogenases, and long chain dehydrogenases [29].

The overall topology of the type I DHQD is an eight-stranded α/β barrel while the type II DHQD subunit consists of a five-stranded β-sheet core flanked by four α-helices. The type I enzymes catalyze a *cis* (*syn*)-dehydration involving loss of the 2-pro-*R* hydrogen via a covalent imine intermediate, while the type II enzymes catalyze a *trans* (*anti*)-dehydration which results in the loss of the 2-pro-*S* hydrogen, probably via an enolate intermediate [3].

Fig. 3C shows the secondary structure of *Xf*DHQD in complex with 2,3-anhydro-quinic acid (FA1). The enzyme is an α/β protein with a central five-stranded parallel β-sheet whose observed strand order is 21345, which is a flavodoxin-like fold, as defined in the SCOP

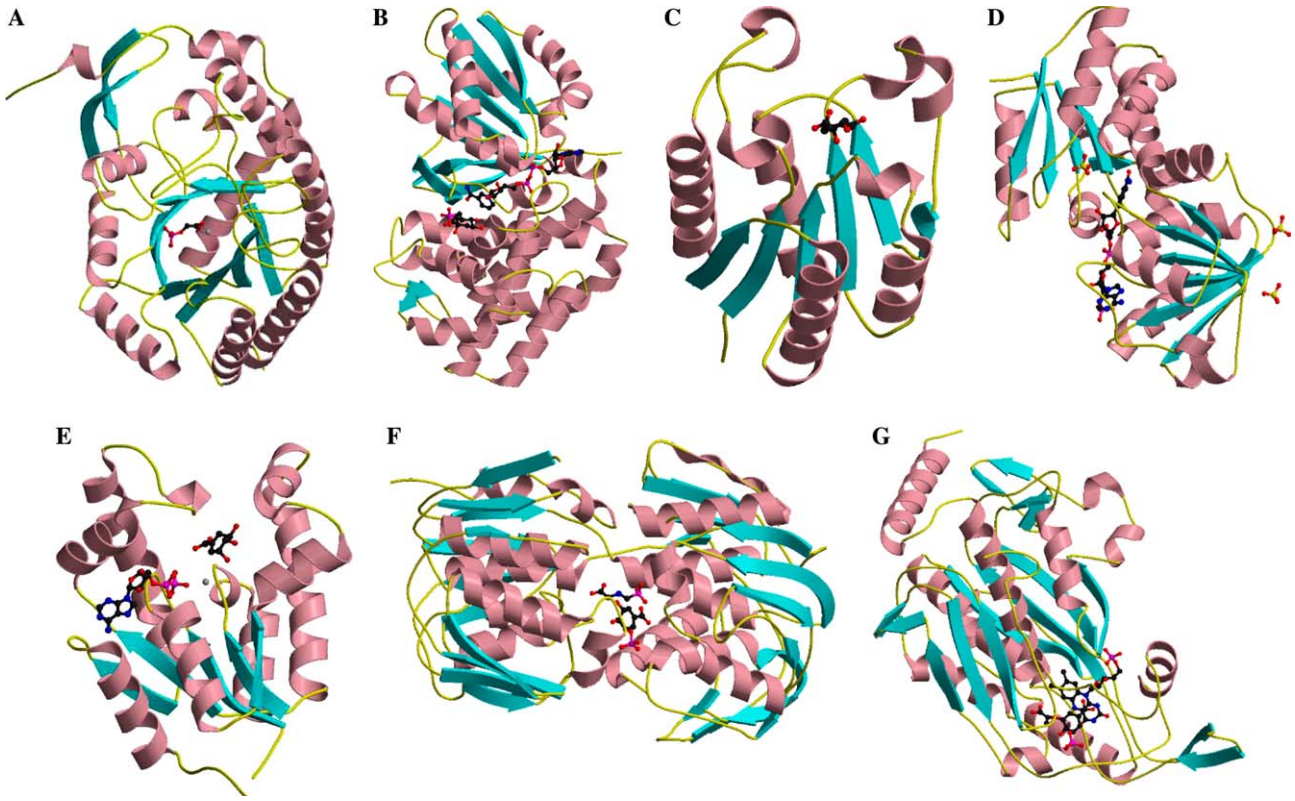


Fig. 3. Ribbon diagrams for: (A) *Xf*DAHPS:PGL:Mn^{2+}, (B) *Xf*DHQS:CRB:NAD^{+}, (C) *Xf*DHQD:FA1, (D) *Xf*SDH:NAD^{+}, (E) *Xf*SK:ATP:Mg^{2+}:SHK, (F) *Xf*EPSPS:glyphosate:S3P, and (G) *Xf*CS:FMN:EPSP. Figures were generated by Molscript [55] and Raster3d [56].

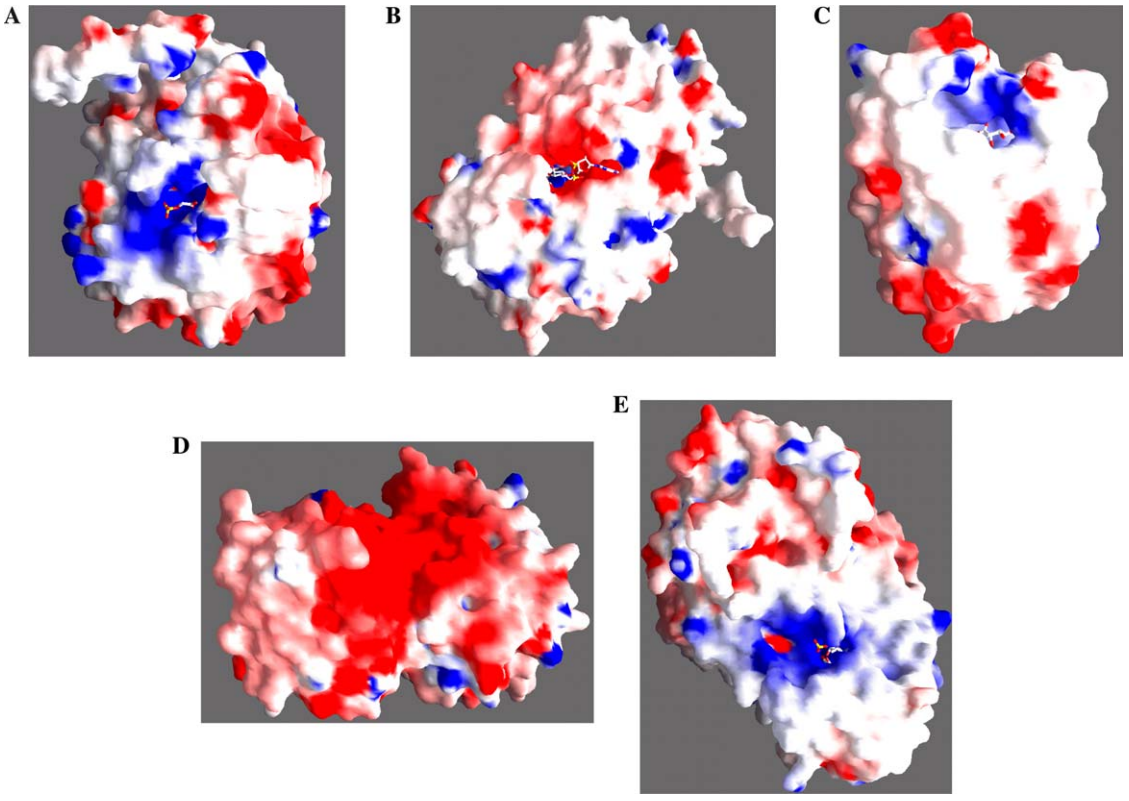


Fig. 4. Electrostatic potential surface for the complexes: (A) *Xf*DAHPS:PGL, (B) *Xf*DHQS:CRB, (C) *Xf*DHQD:FA1, (D) *Xf*EPSPS:glyphosate, and (E) *Xf*CS:FMN calculated with GRASP [53].

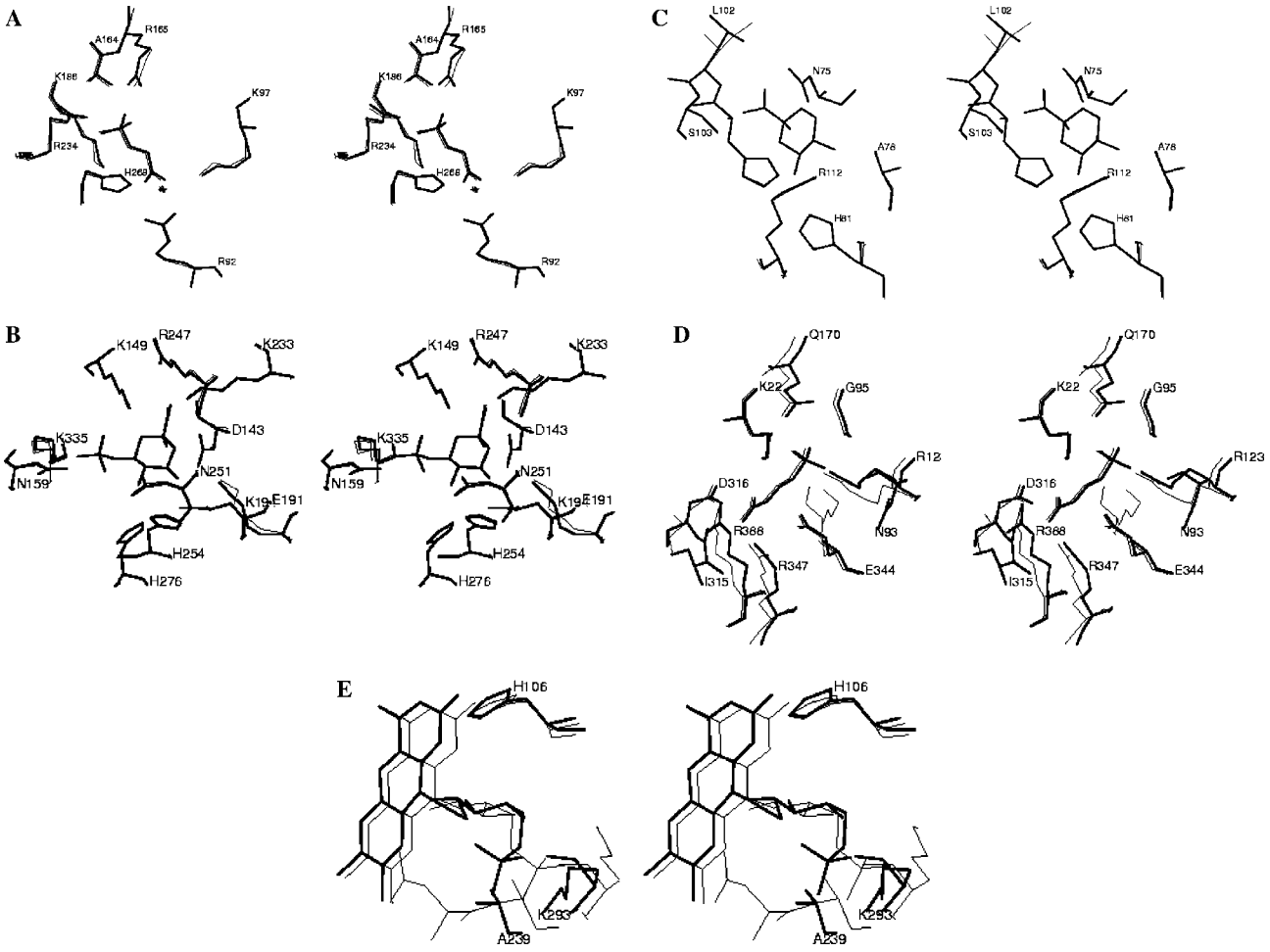


Fig. 5. Stereo views of the interdomain interface region illustrating the binding pockets of the molecular models (thick line) with their template structures (thin line). (A) *Xf*DAHPS:PGL, (B) *Xf*DHQS:CRB, (C) *Xf*DHQD:FA1, (D) *Xf*EPSPS:glyphosate, and (E) *Xf*CS:FMN.

[31] classification of protein folds. There are four main α-helices in the structure, with α1 and α4 on one face of the β-sheet and α2 and α3 on the other face. The active site in type II DHQD is located in the cleft formed at the carboxy edge of the β-sheet between strands β1 and β3, which is common not only to enzymes with a flavodoxin-like fold but also to α/β [30].

We observed a total of 11 hydrogen bonds between *Xf*DHQD and FA1 involving the residues Asn75, Ala78, His81, Leu102, Ser103, and Arg112. For the *Sc*DHQD:FA1 11 hydrogen bonds involving the residues Asn79, Ala82, His85, His106, Ile107, Ser108, and Arg117 were observed. The contact areas for the complexes of *Sc*DHQD:FA1 and *Xf*DHQD:FA1 are 105 and 144 Å^2, respectively.

Shikimate-5-dehydrogenase

The shikimate dehydrogenase catalyzes the fourth reaction in the shikimate pathway, the NADP-dependent reduction of 3-dehydroshikimate to shikimate [5].

This compound, which differs from shikimic acid only by the addition of a hydroxyl group at C1, is the precursor to the ubiquitous plant secondary product chlorogenate. To date, two independent families of quinate/shikimate dehydrogenases have been identified. The first consists of NAD-dependent dehydrogenases, and the second consists of membrane-associated dehydrogenases that utilize pyrrolo-quinoline-quinone as a cofactor [32].

Both types of dehydrogenases are involved in the catabolic quinate pathway, which allows growth of microrganisms with quinate as the sole carbon source by its conversion into protocatechuate and subsequent metabolism by the β-ketoadipate pathway [32].

Fig. 3D shows that the secondary structure of *Xf*SDH in complex with nicotinamide-adenine-dinucleotide (NAD^+) displays an architecture with two α/β domains separated by a wide cleft. The first three β-strands follow a regular α/β succession, with the helices α1 and α2 parallel to the β-strands, flanking opposite sides of the sheet. The domain is completed by a C-terminal α-helical hairpin (α9 and α10), which packs against the

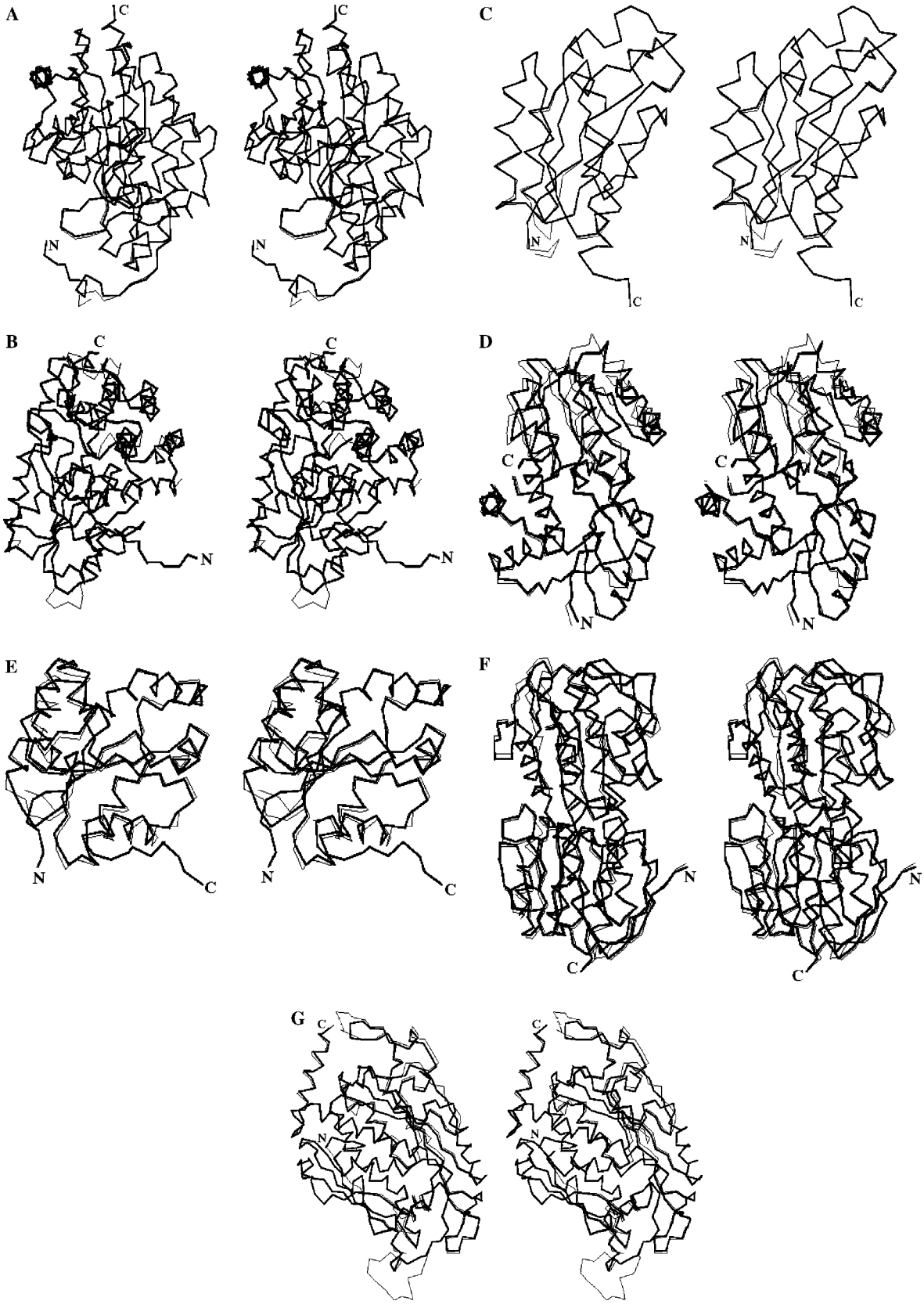


Fig. 6. Superposition of the molecular models (thick line) with their template structures (thin line). (A) *Xf*DAHPS, (B) *Xf*DHQS, (C) *Xf*DHQD, (D) *Xf*SDH, (E) *Xf*SHK, (F) *Xf*EPSPS, and (G) *Xf*CS.

β-sheet on the same side as α1 [32]. This structure reveals an enzyme with a deep cleft, which contains the active site formed at the junction of two domains. The C-terminal domain is easily recognizable as a Rossmann-fold dinucleotide-binding domain, responsible for binding the NADP cofactor. The N-terminal substrate binding and dimerization domain, an α–β–α sandwich, represents a unique topological fold [33].

Shikimate kinase

In the fifth step of the shikimate pathway, shikimate kinase catalyzes the phosphorylation of shikimate to yield shikimate 3-phosphate (S3P) [5].

Fig. 3E shows the secondary structure of *Xf*SK in the complex with shikimate and ADP/Mg^{2+}. It is an α/β protein consisting of a mixed β-sheet surrounded by α-helices. A central five-stranded parallel β-sheet (β1-β5) presents the strand order 23145. The β-strands are flanked on either side by α-helices (α1 and α8 on one side, α4, α5, and α7 on the other) [34]. The ordering of the strands 23145 classifies it as belonging to the same structural family as the nucleoside monophosphate (NMP) kinases. The *Xf*SK structure exhibits the walker-A motif located between β1 and α1 forming a canonical phosphate-binding loop (P-loop). The core of the SK structure forms a classical mononucleoside-binding fold [35]. It has been reported that NMP kinases undergo large conformational changes during catalysis. The regions responsible for this movement are NMP-binding site and the lid domain. The NMP-binding site is formed by a series of helices between strands 1 and 2 of the parallel β-sheet. The lid domain is a region of variable size and the structure following the fourth β-strand of the sheet [36,37]. The docking simulations of shikimate to *Erwinia chrysanthemi* and *Mycobacterium turberculosis* identified that shikimate binds in a position analogous to nucleotide monophosphate in NMP kinases [38–40]. The shikimate-binding domain, which follows strand β2, consists of helices α2 and α3, and the N-terminal region of helice α4.

5-Enolpyruvylshikimate-3-phosphate synthase

In the sixth step of the shikimate pathway, a second phosphoenolpyruvate (PEP) enters in the pathway. It is condensed with S3P to yield 5-enolpyruvylshikimate-3-phosphate (EPSP) and inorganic phosphate. EPSP synthase (EPSPS) is the only cellular target for the herbicide glyphosate. Glyphosate binding is competitive with PEP. For a long time, the ternary complex enzyme-S3P-glyphosate has been considered a transition state analog in which glyphosate takes the place of PEP [5]. Enzyme–ligand distances within the glyphosate-containing ternary complex were measured by NMR [41], and now reports from several laboratories indicate that the complex may not be a transition state analog, because PEP and glyphosate binding is apparently not identical [42,43].

The mode of action of glyphosate on EPSPS has been debated over the past two decades. While studies suggested that glyphosate is an intermediate-state analog of the EPSPS reaction it was later proposed that glyphosate might act as an allosteric inhibitor. This ambiguity was partly due to the lack of structural information about EPSPS. However, the nature of the binding site of either S3P or glyphosate remained obscure until recently, when the crystal structures of EPSPS complexed with S3P (binary complex) or complexed with S3P and glyphosate (ternary complex) were determined to atomic detail. Comparison of ligand-free EPSPS with the binary complex suggests that the two globular domains approach each other on S3P binding, such that the active site emerges in the interdomain cleft [44].

The two flat surfaces of the hemispheres, which in projection form a "V," are almost normal and accommodate the N-terminal of the six helices in each domain. The helical macrodipolar effects create a potential well that facilitates the binding of glyphosate and S3P. The domains are composed of three copies of a βαβαββ folding unit (Fig. 3F). The four-stranded β-sheet structures contain both parallel and antiparallel strands and the helices are parallel [45].

We observed a total of 16 hydrogen bonds in the ternary complex *Xf*EPSPS:S3P:glyphosate, involving the residues Lys22, Asn93, Arg123, Gln170, Ile315, Asp316, Arg347, Glu344, and Arg388. For the *Ec*EPSPS:S3P:glyphosate it was observed 15 hydrogen bonds involving the residues Lys22, Gly96, Arg124, Gln171, Asp313, Glu341, Arg344, Arg386, and Lys411. The contact areas for the complexes of *Ec*EPSPS:glyphosate and *Xf*EPSPS:glyphosate are 186 and 211 $Å^2$, respectively.

Chorismate synthase

The seventh and final step in the main trunk of the shikimate pathway is the *trans*-1,4 elimination of phosphate from EPSP. Under the influence of the enzyme CS, this enol ether undergoes loss of phosphoric acid to produce chorismate. In this reaction, the second of the three double bonds of the benzene ring is introduced. The reaction is catalyzed by CS and requires reduced flavin for activity even though the overall reaction is redox neutral [5].

In chorismate synthase catalysis, the reduced flavin is apparently directly involved in the mechanism of the reaction [5]. The CS catalyzed reaction requires reduced flavin mononucleotide (FMN) and substrate EPSP to produce chorismate, a phosphate molecule, and reduced FMN. It is an unusual reaction in that the redox state of the functional cofactor FMN remains unchanged. Another unusual feature of the CS catalyzed reaction is the

nonconcerted anti-1,4-elimination of the 3-phosphate and the C-(6-*pro*-*R*) hydrogen 4. It is expected that the three-dimensional structure of CS can significantly enhance our understanding of its enzymatic mechanism for this unusual chemistry [46,47].

Studies on CS from bacteria, fungi, and plants revealed that in these organisms the reduced FMN cofactor is made available in different ways to CS: chorismate synthases in fungi—in contrast to those in bacteria and plants—carry a second enzymatic activity which enables them to reduce FMN at the expense of NADPH [48].

The Fig. 3F shows the secondary structure of *Xf*CS in the complex with EPSP and FMN. *Xf*CS has been predicted to be an α–β barrel as a result, of secondary structure prediction efforts, is clearly very different from what was expected. It comprises a single large core domain, which is surrounded by loops and discrete stretches of secondary structure. The core consists of an internal two long α-helices, sandwiched between a pair of four-stranded antiparallel β-sheets [49].

We observed a total of four hydrogen bonds between the ternary complex *Xf*CS:EPSP:FMN, involving the residues His106, Ala239, and Lys293. For the *Sp*CS:EPSP:FMN 6 hydrogen bonds involving the residues His110, Asn251, Ala252, and Lys311 were observed. The contact areas for the complexes of *Sp*CS:EPSP:FMN and *Xf*CS:EPSP:FMN are 542 and 476 $Å^2$, respectively.

Conclusions

The specificity and affinity between enzyme and its inhibitor depend on directional hydrogen bounds and ionic interactions, as well as on shape complementarity of the contact surfaces of both partners [50–52]. The electrostatic potential surface of the all models was calculated with GRASP [53] and shown in Figs. 4A–E. Analysis of these structural models indicates that inhibitors such as glyphosate for EPSPS and FA1 for DHQD are strongly bound to the enzymes. All complexes show good shape and charge complementarity and exhibit a higher number of intermolecular hydrogen bonds, which strongly indicates that these inhibitors could also inhibit *X. fastidiosa* enzymes. Especially interesting is the *Xf*EPSPS:glyphosate:S3P complex that shows 16 intermolecular hydrogen bonds, more than that observed for *Escherichia coli* enzyme (14 bonds), and a higher contact area, which indicates that glyphosate may have higher affinity for *Xf*EPSPS than that observed for *Ec*EPSPS. Further inhibitor analysis may confirm this prediction.

Acknowledgments

This work was supported by grants from FAPESP (SMOLBNet 01/07532-0, 02/04383-7), CNPq, CAPES and Instituto do Milênio (CNPq-MCT). WFA (CNPq, 300851/98-7), and MSP (CNPq, 300337/2003-5) are researchers for the Brazilian Council for Scientific and Technological Development.

References

[1] A.J.G. Simpson et al., The genome sequence of the plant pathogen *Xylella fastidiosa*, Nature 406 (2000) 151–157.
[2] M. Bevan, The bugs from Brazil, Nature 406 (2000) 140–141.
[3] J.R. Coggins, C. Abell, L.B. Evans, M. Frederickson, D.A. Robinson, A.W. Roszak, A.P. Lapthorn, Experiences with the shikimate-pathway enzymes as targets for rational drug design, Biochem. Soc. Trans. 31 (2003) 548–552.
[4] S.A. Campbell, T.A. Richards, E.J. Mui, B.V. Samuel, J.R. Coggins, R. McLeod, C.W. Roberts, A complete shikimate pathway in *Toxoplasma gondii* an accident eukaryotic innovation, Int. J. Parasitol. 34 (2004) 5–13.
[5] K.M. Herman, L.M. Weaver, The shikimate pathway, Annu. Rev. Plant Mol. Biol. 50 (3) (1999) 473–503.
[6] F. Dosselaere, J. Vanderleyden, A metabolic node in action: chorismate-utilizing enzymes in microrganisms, Crit. Rev. Microbiol. 27 (2) (2001) 75–131.
[7] K.M. Herrmann, The shikimate pathway as an entry to aromatic secondary metabolism, Plant Physiol. 107 (1995) 7–12.
[8] R.T. Kroemer, S.W. Doughty, A.J. Robinson, W.G. Richards, Prediction of the three-dimensional structure of human interleukin-7 by homology modeling, Protein Eng. 9 (6) (1996) 493–498.
[9] T.L. Blundell, B.L. Sibanda, M.J. Sternberg, J.M. Thornton, Knowledge-based prediction of protein structures and the design of novel molecules, Nature 326 (6111) (1987) 347–352.
[10] T.L. Blundell, D. Carney, S. Gardner, F. Hayes, B. Howlin, T. Hubbard, J. Overington, D.A. Singh, B.L. Sibanda, M. SutCliffe, 18th Krebs, Hans lecture knowledge-based protein modeling and design, Eur. J. Biochem. 172 (3) (1988) 513–520.
[11] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[12] A. Sali, J.P. Overington, Derivation of rules for comparative protein modeling from a database of protein structure alignments, Protein Sci. 3 (9) (1994) 1582–1596.
[13] A. Sali, L. Potterton, F. Yuan, H. van Vlijmen, M. Karplus, Evaluation of comparative protein modeling by MODELLER, Proteins 23 (3) (1995) 318–326.
[14] A. Sali, Modeling mutations and homologous proteins, Curr. Opin. Biotechnol. 6 (4) (1995) 437–451.
[15] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy minimization and dynamics calculations, J. Comp. Chem. 4 (1983) 187–217.
[16] W. Braun, N. Go, Calculation of protein conformations by proton-proton distance constraints. A new efficient algorithm, J. Mol. Biol. 186 (3) (1985) 611–626.
[17] R.A. Laskowski, M.W. MacArthur, D.K. Smith, D.T. Jones, E.G. Hutchinson, A.L. Morris, D. Naylor, D.S. Moss, J.M. Thorton, PROCHECK v.3.0—Program to check the stereochemistry quality of Protein structures—Operating instructions (1994).
[18] Collaborative Computational Project No. 4, The CCP4 suite: programs for proteins crystallography, Acta Crystallogr. D 50 (1994) 760–763.
[19] C.D. Schwieters, J.J. Kuszewski, N. Tjandra, G.M. Clore, The Xplor-NIH NMR molecular structure determination package, J. Magn. Res. 160 (2003) 66–74.
[20] A.T. Brünger, X-PLOR Version 3.1: A System for Crystallography and NMR, Yale University Press, New Haven, 1992.

[21] J.U. Bowie, R. Luthy, D. Eisenberg, A method to identify protein sequences that fold into a known three-dimensional structure, Science 253 (1991) 164–170.
[22] R. Luthy, J. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, Nature 356 (1992) 83–85.
[23] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (12) (1983) 2577–2637.
[24] T. Wagner, I.A. Shumilin, R. Bauerle, R.H. Kretsinger, Structure of 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Escherichia coli*: comparison of the Mn^{2+}*2-phosphoglycolate and the Pb^{2+}*2-phosphoenolpyruvate complexes and implications for catalysis, J. Mol. Biol. 301 (2000) 389–399.
[25] K.A. Herrmann, The shikimate pathway: early steps in the biosynthesis of aromatic compounds, Plant Cell 7 (1995) 907–919.
[26] E.P. Carpenter, A.R. Hawkins, J.W. Frost, K.A. Brown, Structure of dehydroquinate synthase reveals an active site capable of multistep catalysis, Nature 394 (6690) (1998) 299–302.
[27] M.G. Rossmann, D. Moras, K.W. Olsen, Chemical and biological evolution of a nucleotide binding protein, Nature 250 (1974) 194–199.
[28] A.M. Lesk, NAD-binding domains of dehydrogenases, Curr. Opin. Struct. Biol. 5 (1995) 775–783.
[29] J. Benach, I. Lee, W. Edstrom, A.P. Kuzin, Y. Chiang, T.B. Actons, G.T. Monteliones, J.F. Hunt, The 2.3 Å crystal of the shikimate 5-deydrogenase orthologue YdiB from *E. coli* suggests a novel catalytic enviroment for an NAD-dependent dehydrogenase, J. Biol. Chem 278 (2003) 19176–19182.
[30] A.W. Roszak, D.A. Robinson, T. Krell, I.S. Hunter, M. Fredrickson, C. Abell, J.R. Coggins, A.J. Lapthorn, The structure and mechanism of the type II dehydroquinase from *Streptomyces coelicolor*, Structure (Camb.) 10 (4) (2002) 493–503.
[31] A.G. Murzin, S.E. Brenner, T. Hubbard, C. Chothia, SCOP: a structural classification of proteins database for the investigation of sequences and structures, J. Mol. Biol. 247 (1995) 536–540.
[32] G. Michel, A.W. Roszak, V. Sauvé, J. Maclean, A. Matte, J.R. Coggins, M. Cygler, A.J. Lapthorn, Structures of Shikimate dehydrogenase AroE and its paralog YdiB—a common structural framework for different activities, J. Biol. Chem. 278 (21) (2003) 19463–19472.
[33] E. Vogan, Shikimate dehydrogenase structure reveals novel fold, Structure 11 (2003) 902–903.
[34] G.E. Schulz, Binding of nucleotides by proteins, Curr. Opin. Struct. Biol. 2 (1992) 61–67.
[35] Y. Matsuo, K. Nishikawa, Protein structural similarities predicted by a sequence–structure compatibility method, Protein Sci. 3 (11) (1994) 2055–2063.
[36] C.W. Müller, G.J. Schlauderer, J. Reinstein, G.E. Schulz, Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding, Structure 4 (1996) 147–156.
[37] M. Gerstein, G. Schulz, C. Chothia, Domain closure in adenylate kinase. Joints on either side of two helices close like neighboring fingers, J. Mol. Biol. 229 (2) (1993) 494–501.
[38] Y. Gu, L. Reshetnikova, Y. Li, Y. Wu, H. Yan, S. Singh, X. Ji, Crystal structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the dynamic role of the LID domain in catalysis, J. Mol. Biol. 319 (3) (2002) 779–789.
[39] J.H. Pereira, J.S. de Oliveira, F. Canduri, M.V.B. Dias, M.S. Palma, L.A. Basso, W.F. de Azevedo Jr., D.S. Santos, X-ray crystallographic studies of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimate acid, Nat. Struct. Biol. (submitted).
[40] T. Krell, J.R. Coggins, A.J Lapthorn, The three-dimensional structure of shikimate kinase, J. Mol. Biol. 178 (5) (1998) 983–997.
[41] L.M. McDowell, A. Schmidt, E.R. Cohen, D.R. Studelska, J. Schaefer, Structural constraints on the ternary complex of 5-enolpyruvyl shikimate 3-phosphate synthase from rotational-echo double resonance NMR, J. Mol. Biol. 256 (1996) 160–171.
[42] R.D. Sammons, K.J. Gruys, K.S. Anderson, K.A. Johnson, J.A. Sikorski, Reevaluating glyphosate as a transition state inhibitor of EPSP synthase: identification of an EPSP synthase–EPSP–glyphosate ternary complex, Biochemistry 34 (1995) 6433–6440.
[43] L.M McDowell, C.A. Klug, D.D. Beusen, J. Schaefer, Ligand geometry of the ternary complex of 5-enolpyruvyl shikimate 3-phosphate synthase from rotational-echo double-resonance NMR, Biochemistry 35 (1995) 5395–5403.
[44] S. Eschenburg, M.L. Healy, M.A. Priestman, G.H. Lushington, E. Schönbrunn, How the mutation glycine96 to alanine confers glyphosate insensitivity to 5-enolpyruvyl shikimate-3-phosphate synthase from *Escherichia coli*, Planta 216 (2002) 129–135.
[45] J.H. Pereira, F. Canduri, J.S. de Oliveira, N.J. da Silveira, L.A. Basso, M.S. Palma, W.F. deAzevedo Jr., D.S. Santos, Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*, Biochem. Biophys. Res. Commun. 312 (3) (2003) 608–614.
[46] E. Schonbrunn, S. Eschenburg, W.A. Shuttleworth, J.V. Schloss, N. Amrhein, J.N. Evans, W. Kabsch, Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail, Proc. Natl. Acad. Sci. USA 98 (4) (2001) 1376–1380.
[47] C.M. Viola, V. Saridakis, D. Christendat, Crystal structure of chorismate synthase from *Auifex aeolicus* reveals a novel beta alpha beta sandwich topology, Proteins 54 (2004) 166–169.
[48] P. Macheroux, J. Schmid, N. Amrhein, A. Schaller, A unique reaction in a common pathway: mechanism and function of chorismate synthase in the shikimate pathway, Planta 207 (1999) 325–334.
[49] J. Maclean, S. Ali, The structure of chorismate synthase reveals a novel flavin binding site fundamental to a unique chemical reaction, Structure 11 (2003) 1499–1511.
[50] W.F. De Azevedo, F. Canduri, N.J.F. da Silveira, Structural basis for inhibition of cyclin-dependent kinase 9 by flavopiridol, Biochem. Biophys. Res. Commun. 293 (2002) 566–571.
[51] W.F. De Azevedo, R.T. Gaspar, F. Canduri, J.C. Camera, N.J.F. da Silveira, Molecular model of cyclin-dependent kinase 5 complexed with roscovitine, Biochem. Biophys. Res. Commun. 297 (2002) 1154–1158.
[52] W.F. De Azevedo, J.S. de Oliveira, L.A. Basso, M.S. Palma, J.H. Pereira, F. Canduri, D.S. Santos, Molecular model of shikimate kinase from *Mycobacterium tuberculosis*, Biochem. Biophys. Res. Commun. 295 (1) (2002) 142–148.
[53] A. Nicholls, K.A. Sharp, B. Honig, Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons, Proteins 11 (4) (1991) 281–296.
[54] J.D. Thompson, D.G. Higgins, T.J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice, Nucleic Acids Res. 22 (1994) 4673–4680.
[55] P.J. Kraulis, MOLSCRIPT: a program to produce both detailed and schematic plots of proteins, J. Appl. Crystallogr. 24 (1991) 946–950.
[56] E.A. Merritt, D.J. Bacon, Raster3D: photorealistic molecular graphics, Methods Enzymol. 277 (1997) 505–524.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 324 (2004) 789–794
BBRC
www.elsevier.com/locate/ybbrc

Crystallographic structure of PNP from *Mycobacterium tuberculosis* at 1.9 Å resolution

Diego O. Nolasco[a,b], Fernanda Canduri[a,b], José H. Pereira[a,b], Janaina R. Cortinóz[a], Mário S. Palma[c], Jaim S. Oliveira[d], Luiz A. Basso[d], Walter F. de Azevedo Jr.[a,*], Diógenes S. Santos[d,*]

[a] *Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil*
[b] *Programa de Pós-Graduação em Biofísica Molecular, UNESP, São José do Rio Preto, SP 15054-000, Brazil*
[c] *Laboratory of Structural Biology and Zoochemistry-CEIS, Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil*
[d] *Rede Brasileira de Pesquisa em Tuberculose, Centro de Pesquisa em Biologia Molecular e Functional, PUCRS, Porto Alegre, RS 90619-900, Brazil*

Received 1 September 2004
Available online 2 October 2004

Abstract

Even being a bacterial purine nucleoside phosphorylase (PNP), which normally shows hexameric folding, the *Mycobacterium tuberculosis* PNP (MtPNP) resembles the mammalian trimeric structure. The crystal structure of the MtPNP apoenzyme was solved at 1.9 Å resolution. The present work describes the first structure of MtPNP in complex with phosphate. In order to develop new insights into the rational drug design, conformational changes were profoundly analyzed and discussed. Comparisons over the binding sites were specially studied to improve the discussion about the selectivity of potential new drugs.

Keywords: Tuberculosis; PNP; Crystallography; Apoenzyme; Selectivity

Tuberculosis causes eight million new infections and kills two million people each year worldwide according to the World Health Organization (WHO). It is estimated that approximately two billion individuals, one-third of the world population, are infected with latent TB. Tuberculosis resurged in the late 1980s and was declared to be a global emergency by the WHO. The high susceptibility of human immunodeficiency virus-infected persons to the disease and the proliferation of multi-drug-resistant (MDR) strains have created a worldwide interest in expanding current programs in tuberculosis research. New antimycobacterial agents are needed to treat *Mycobacterium tuberculosis* strains resistant to existing drugs and to shorten the treatment course to improve patient compliance [1].

The TB is a chronic infectious disease caused by various alcohol–acid-resistant bacterium from the *Mycobacterium* gender. The more frequent cyclic form of the disease is the pulmonary tuberculosis, caused by the *M. tuberculosis* (Koch *bacillus*), but can also cause cerebral lesions, skinny (lupus), and ganglionaries, some times produced by the human origin *bacillus*, some times produced by the bovine origin *bacillus*. The Koch *bacillus* is an extremely small and resistant bacterium, in stick shape. It can live in dry conditions for months and can also resist disinfectants of moderate action.

Purine nucleoside phosphorylase (PNP) catalyzes the phosphorolysis of purine nucleosides to corresponding bases and ribose 1-phosphate. PNP plays a central role in purine metabolism, normally operating in the purine

* Corresponding authors.
E-mail addresses: walterfa@df.ibilce.unesp.br (W.F. de Azevedo Jr.), diogenes@pucrs.br (D.S. Santos).

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.09.137

salvage pathway of cells. PNP is specific for purine nucleosides in the β-configuration and exhibits a preference for ribosyl-containing nucleosides relative to the analogs containing arabinose, xylose, and lyxose stereoisomers. Moreover, PNP cleaves glycosidic bond with inversion of configuration to produce α-ribose 1-phosphate [2–4].

PNP is an enzyme of purine metabolism that functions in the salvage pathway, including even those of protozoan parasites, thus enabling the cells to utilize purine bases recovered from metabolized purine ribo- and deoxyribonucleosides to synthesize purine nucleotides [5–7]. The PNPs from various sources are members of a broader class of N-ribohydrolases and transferases, the transition states for which share ribosyl oxocarbenium like character, with cleavage of the C–N ribosyl bond [8,9].

A number of parasites lacking the ability to synthesize purine nucleotides de novo must utilize host purines, formed by PNP, for DNA synthesis. Thus, inhibition of PNP could prevent the spread of parasitic infection [10].

The human PNP is a potential target for drug development, which could induce immune suppression to treat, for instance, autoimmune diseases, T-cell leukemia, lymphoma, and organ transplantation rejection. Furthermore, PNP inhibitors can also be used to avoid cleavage of anticancer and antiviral drugs, since many of these drugs mimic natural purine nucleosides and can thereby be cleaved by PNP before accomplishing their therapeutic role [2,3].

Homologs to enzymes in the purine salvage pathway have been identified in the genome sequence of *M. tuberculosis*. The genome of *M. tuberculosis* comprised of 4,411,529 bp containing 3924 potential open reading frames [11]. On the basis of sequence homology, biochemical functions have been attributed to approximately 40% of the predicted proteins, while similarities to other described proteins were found for another 44% and the remaining 16% bore no resemblance to known proteins and may encode proteins with specific mycobacterial functions [11,12].

In the de novo synthesis of purine ribonucleotides, the formation of adenosine monophosphate (AMP) and guanine monophosphate (GMP) from inosine monophosphate (IMP) is irreversible, but purine bases, nucleosides, and nucleotides can be interconverted through the activities of purine nucleoside phosphorylase, adenosine deaminase, and hypoxanthine–guanine phosphoribosyl transferase. The specific inhibition of MtPNP could potentially lead to the accumulation of guanine nucleotides since a putative guanylate kinase and a nucleoside diphosphate kinase are encoded in the genome [13].

The crystallographic structure of the MtPNP was first determined in 2001 at 1.75 Å resolution with Immucillin-H [14,15] and inorganic phosphate, and also solved at 2.0 Å resolution with 9-deazahipoxanthine and iminoribitol [16].

We have now obtained X-ray diffraction data using synchrotron radiation and refined the structure of the apoenzyme at 1.9 Å resolution, using the recombinant MtPNP with inorganic phosphate (MtPNP · PO_4). Our analysis of the MtPNP · PO_4 structural data, and structural differences between the apoenzyme and the MtPNP · Immucillin-H · PO_4 complex provides new insights into substrate binding, the purine-binding site, and can be used for future inhibitors' design.

Methods

Crystallization. Recombinant MtPNP was expressed and purified as previously described [1]. MtPNP was crystallized using the experimental conditions described elsewhere [16]. In brief, the PNP (1 μL at a concentration of 25 mg ml^{-1}) containing 5 mM NaH_2PO_4 was mixed with an equal volume of the reservoir solution containing 100 mM Tris, pH 8.0, 25% PEG 3350, and 25 mM $MgCl_2$, and equilibrated against 1.0 mL of the reservoir solution. Diffraction from the crystals was consistent with the space group $P2_12_12$ ($a = 117.95$ Å, $b = 134.35$ Å, and $c = 44.20$ Å), with a trimer in the asymmetric unit ($V_m = 2.12$ $Å^3$ Da^{-1}; 41.92% solvent content).

Data collection, processing, and structure determination. Crystals were cryoprotected by transfer to crystallization solution with 20% glycerol and flash-cooled at 100 K. X-ray diffraction data were collected at 1.4270 Å wavelength on a CCD detector using synchrotron radiation at beam line CPR at the Synchrotron Radiation Source

Table 1
Data collection and refinement statistics

MtPNP · PO_4	
Data collection	
Resolution limits (Å)	40.0–1.9 (2.0–1.9)
Completeness (%)	90.7 (87.3)
Space group	$P2_12_12$
a	117.95
b	134.34
c	44.20
α	90°
β	90°
γ	90°
R_{sym}(%)[a]	6.2 (24.4)
No. of reflections	
Unique	52718
Total	524878
Structure refinement	
R_{factor}[b]	19.00
R_{free}[c]	26.30
No. of amino acids (trimmer)	788
No. of waters	457
No. of phosphate groups	3

Values in parentheses are for the highest resolution shell.

[a] $R_{sym} = 100\sum|I(h) - \langle I(h)\rangle|/\sum I(h)$, where $I(h)$ is the observed intensity and $\langle I(h)\rangle$ is the mean intensity of reflections h overall of $I(h)$.

[b] $R_{factor} = 100\sum|F_{obs} - F_{calc}|/\sum F_{obs}$. The sums are being taken over all reflections with $F/\sigma(F) > 2$ cutoff.

[c] $R_{free} = R_{factor}$ for 10% of the data, which were not included during the crystallographic refinement.

(Laboratório Nacional de Luz Síncrotron, LNLS, Campinas, Brazil). Data for the MtPNP were processed at 1.9 Å resolution using the program MOSFLM and scaled with the program SCALA [17] and were 90.7% complete with R_{sym} of 6.2%. In the highest resolution shell (1.9–2.0 Å) the reflections presented R_{sym} of 24.4%.

Structure determination and refinement. The structure of MtPNP · PO_4 was solved by molecular replacement with AMoRe software package [18] using the trimer of *M. tuberculosis* PNP · 9dH-X · IR · PO_4 (PDB ID code 1I80) complex as search model, the ligands and water molecules were removed from the model. The best solution after rigid-body refinement yielded an initial R_{factor} of 34.9% and a correlation coefficient of 66.4% using data in the resolution range of 8.0–4.0 Å. The atomic positions obtained from molecular replacement were used to initiate the crystallographic refinement. Model building was performed employing the program XtalView [19] using $2F_o - F_c$ and $F_o - F_c$ electron density maps. The densities for the phosphate groups were localized, one by monomer, and these groups were added. A total of 457 water molecules were added in the model, 50 molecules by each refinement step. The structure refinement of MtPNP · PO_4 was performed using Refmac5 [20]. The final model had R_{free} and R_{factor} of 26.30% and 19.00%, respectively (Table 1). The overall stereochemical quality of the final model was assessed by the programs PROCHECK [21] and Parmodel [22] (Table 2).

Table 2
Assessment of the final structure

MtPNP · PO_4	
Procheck	
Most favored regions	91.4%
Additional allowed regions	8.0%
Disallowed regions	0.6%
Parmodel	
3D Profile ideal score	121.29
3D Profile score	124.06
Main chain *B*-factor (Å^2)	18.39
Side chain *B*-factor (Å^2)	21.99
Average protein *B*-factor (Å^2)	20.16
Average water *B*-factor (Å^2)	26.10

Quality of the model. Analysis of the Ramachandran diagram $\Phi - \Psi$ plot for the present structure indicates that 91.4% of the residues are found to occur in the most favored regions, 8.0% in the additional allowed regions, and just 4 residues (0.6%) in the disallowed regions of the plot. Analysis of the electron-density map ($2F_{obs} - F_{calc}$) agrees with the Thr209, of the three monomers, and His68, of monomer A, positioning.

The final model of the MtPNP · PO_4 and the MtPNP complexes (MtPNP · ImmH · PO_4, MtPNP · IR · 9dHX · PO4) were superposed. Finally, the MtPNP · PO_4 and the *Homo sapiens* PNP (HsPNP · SO_4) were also superposed using the program PROFIT [23].

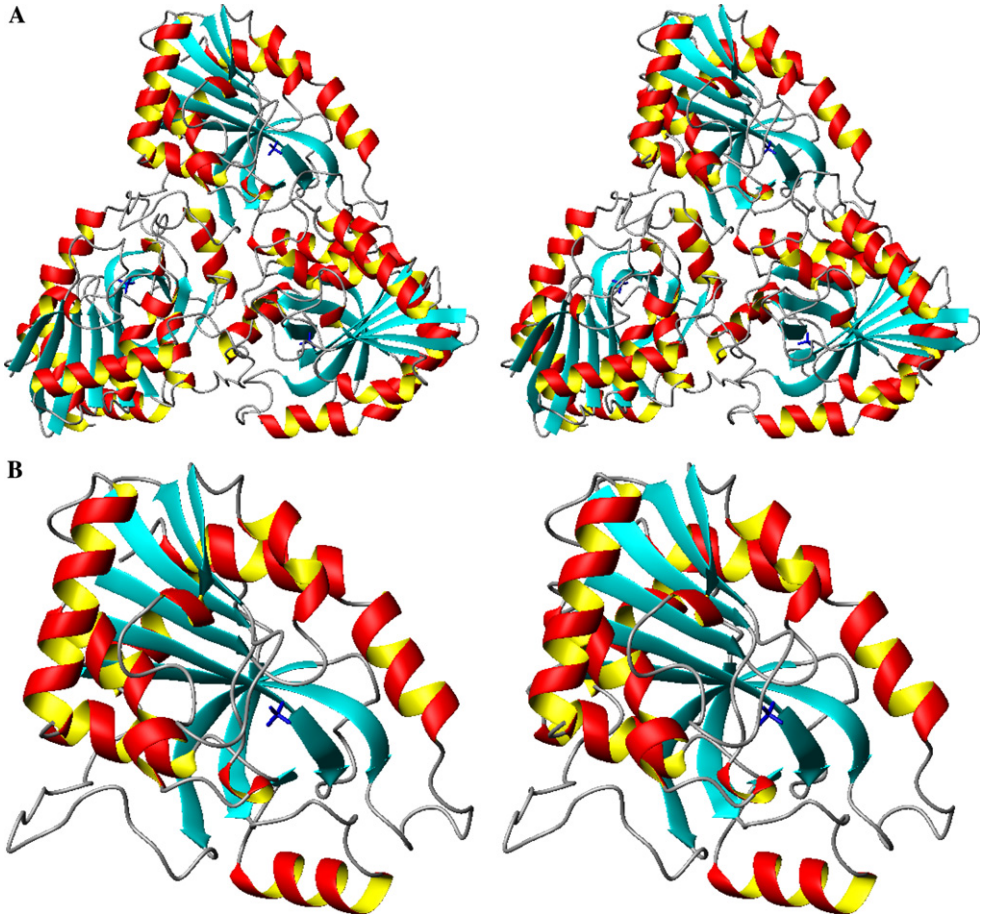


Fig. 1. Ribbon diagrams of the *M. tuberculosis* PNP trimer of the asymmetric unit (A) and monomer (B). The active sites are located near the trimer interfaces. These figures and the others were generated using MOLMOL [24].

Results and discussion

Overall structure of MtPNP · PO_4 apoenzyme

The protein is a symmetrical homotrimer with a triangular arrangement of subunits similar to the mammalian trimeric PNPs. Each monomer of the protein is folded into an α/β-fold consisting of 11 β sheet surrounded by eight α helices (Fig. 1). The structure of MtPNP · PO_4 shows clear electron-density peaks for three phosphate groups, which is present in high concentration in the crystallization experimental condition.

The three independent active sites lie near the subunit interfaces (Fig. 1A). Each active site is composed of five residues, Tyr92, Glu189, Met207, Asn231, and His243, which show a considerable different positioning in comparison with the MtPNP · ImmH · PO_4 complex. Each one of the three active sites contains one phosphate molecule.

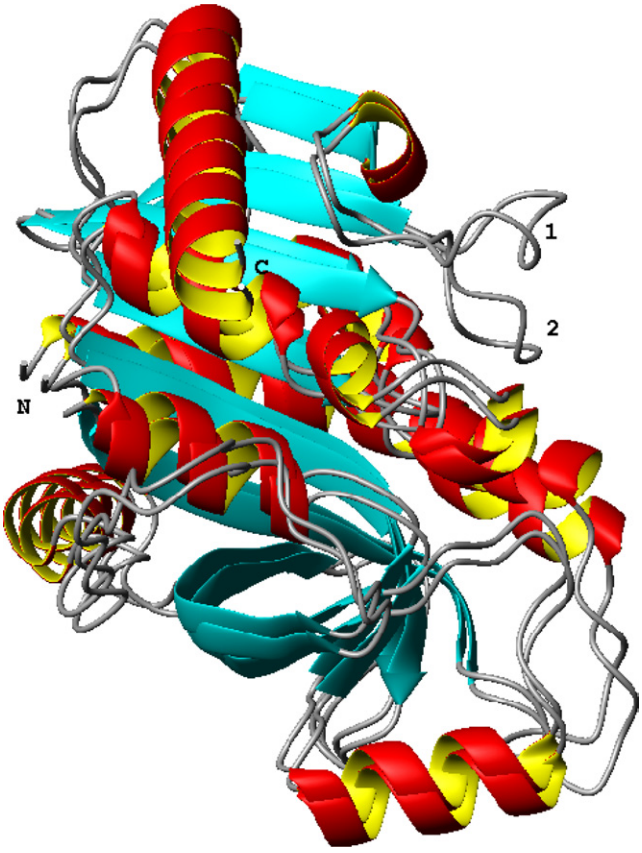


Fig. 2. Ribbon diagram of the superposition of the MtPNP · PO_4 (1) and MtPNP · ImmH · PO_4 (2) lid region.

Comparison of the MtPNP · PO_4 and the MtPNP · ImmH · PO_4 complex

A significant structural change is observed when both structures are superposed (Fig. 2), showing a kind of lid for the ligand. It is strongly believed that this change in the positions of residues from Pro62 to Gly70 is caused by the different crystallographic packing of the proteins, since the diffraction of the complex MtPNP · ImmH · PO_4 was consistent with the space group $P3_221$, different from the MtPNP · PO_4 space group, $P2_12_12$.

In a superposition of both structures, using the program PROFIT, a RMS deviation of 2.381 Å was observed. Fitting just the binding sites the observed RMS deviation was of 1.370 Å, showing that the catalytic site is more compact in the MtPNP · ImmH · PO_4 complex than in the MtPNP · PO_4 apoenzyme. The interaction forces of the residues Tyr92–Glu189–Met207–Asn231–His243 with the Immucillin-H molecule provide this compression.

The water molecule positions also show substantial differences and the phosphate groups are shifted because of the missing ligand, but the side chains of Arg88 and His90 continue participating in the hydrogen bond with the O1 of the phosphate, while Ser36 interacts with both O4 and O3 through side chain backbone atoms.

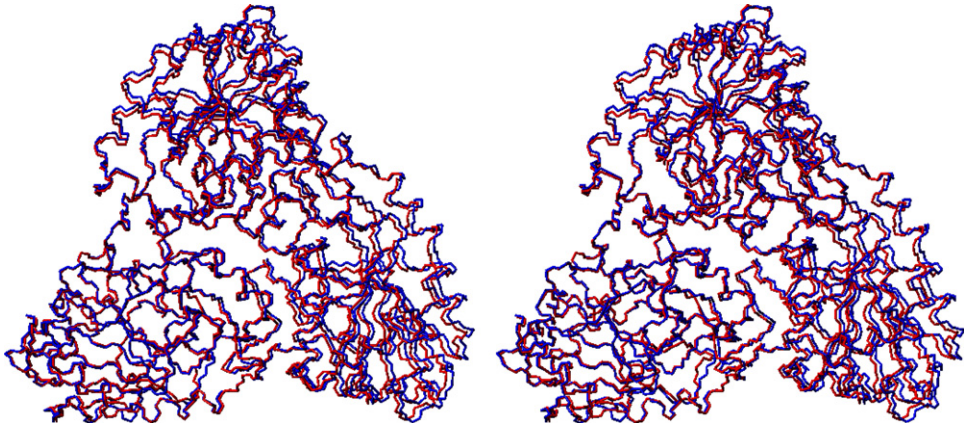


Fig. 3. Stereo diagram of the superposition of the MtPNP · PO_4 apoenzyme and the MtPNP · IR · 9dHX · PO_4 complex.

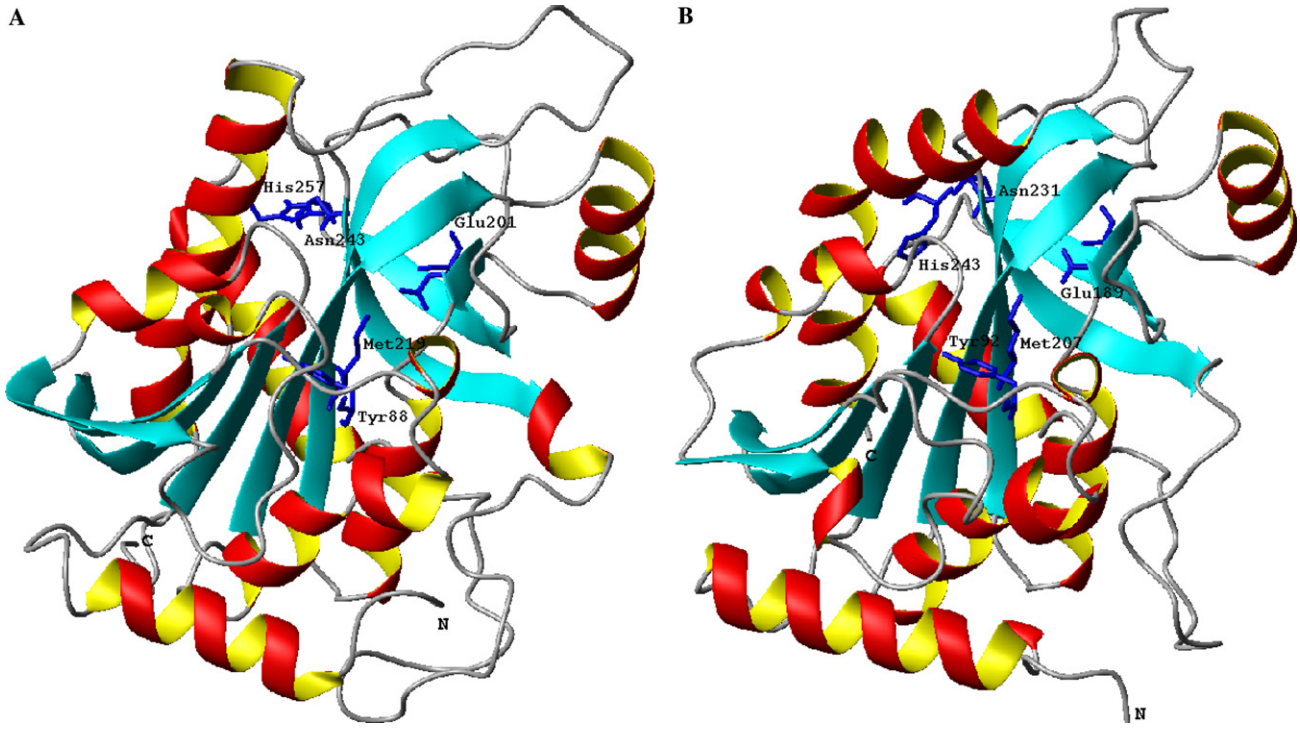


Fig. 4. Comparison of monomers from the HsPNP · SO_4 (A) and the MtPNP · PO_4 (B), and the residues of the active sites of both structures.

Comparison of the MtPNP · PO_4 apoenzyme and the MtPNP · IR · 9dHX · PO_4 complex

The region of the lid, residues from Pro62 to Gly70, fits perfectly, showing that the complete Immucillin-H provides much more influence over these residues than the 9-deazahypoxantine (Fig. 3). In a superposition of both structures the observed RMS deviation was of 0.654 Å, which corroborates with the idea of non-influence of the IR · 9dHX complex over the protein structure.

Comparison of the MtPNP · PO_4 and the HsPNP · SO_4

Even being from different origins, both PNPs show the mammalian trimeric structures. The active sites of the *M. tuberculosis* PNP and the human PNP are formed just by the same residues and show identical folding (Fig. 4). In a superposition of the structures, the observed RMS deviation was of 2.517 Å, and the lid regions agree perfectly, even with the changes of some residues, which are Val61 to Pro57–Pro62 to Arg58–Pro63 to Ser59–Ala65 to Val61 and Ala66 to Pro62.

Conclusions

This paper establishes that conformational differences between two stages of the same protein, complexed and apoenzyme, can be used to develop studies to improve the design of new more powerful inhibitors. The significance of this knowledge is that both structures can now be explored in attempts to ascend higher studies over rational drug design. The protein structure described here is one of the most important targets for drug development to treat tuberculosis.

Acknowledgments

This work was supported by grants from FAPESP (SMOLBNet, Proc. 01/07532-0, 02/04383-7, and 04/00217-0), CNPq, CAPES, and Instituto do Milênio (CNPq-MCT). W.F.A., D.S.S., M.S.P., and L.A.B. are researchers awardees from the National Research Council of Brazil–CNPq.

References

[1] L.A. Basso, D.S. Santos, W. Shi, R.H. Furneaux, P.C. Tyler, V.L. Schramm, J.S. Blanchard, Purine nucleoside phosphorylase from *Mycobacterium tuberculosis*. Inhibition by a transition-state analogue and dissection by parts, Biochemistry 40 (2001) 8196–8203.

[2] W.F. deAzevedo Jr., F. Canduri, D.M. dosSantos, R.G. Silva, J.S. deOliveira, L.P.S. deCarvalho, L.A. Basso, M.A. Mendes, M.S. Palma, D.S. Santos, Crystal structure of human purine nucleoside phosphorylase at 2.3 Å resolution, Biochem. Biophys. Res. Commun. 308 (2003) 545–552.

[3] A. Bzowska, E. Kulikowska, D. Shugar, Purine nucleoside phosphorylase: properties, functions and clinical aspects, Pharmacol. Ther. 88 (2000) 349–425.
[4] A. Bzowska, E. Kulikowska, D. Shugar, Properties of purine nucleoside phosphorylase (PNP) of mammalian and bacterial origin, Z. Naturforsch. Sect. C 45 (1990) 59–70.
[5] A. Bzowska, E. Kulikowska, D. Shugar, C. Bing-yi, B. Lindborg, N.G. Johansson, Acyclonucleoside analogue inhibitors of mammalian purine nucleoside phosphorylase, Biochem. Pharmacol. 41 (1991) 1791–1803.
[6] A. Bzowska, E. Kulikowska, D. Shugar, Formycins A and B and some analogues: selective inhibitors of bacterial *Escherichia coli* purine nucleoside phosphorylase, Biochim. Biophys. Acta 1120 (1992) 239–247.
[7] A. Bzowska, A.V. Ananiev, N. Ramzaeva, E. Alksins, J.A. Maurins, E. Kulikowska, D. Shugar, Purine nucleoside phosphorylase: inhibition by purine *N*(7)- and *N*(9)-acyclonucleosides; and substrate properties of 7-β-D-ribofuranosylguanine and 7-β-D -ribofuranosilhypoxanthine, Biochem. Pharmacol. 48 (1994) 937–947.
[8] S.E. Ealick, S.A. Rule, D.C. Carter, T.J. Greenhough, Y.S. Babu, W.J. Cook, J. Habash, J.R. Helliwell, J.D. Stoeckler, R.E. Parks Jr., S.-F. Chen, C.E. Bugg, Three-dimensional structure of human erythrocytic purine nucleoside phosphorylase at 3.2 Å resolution, J. Biol. Chem. 265 (3) (1990) 1812–1820.
[9] S.E. Ealick, Y.S. Babu, C.E. Bugg, M.D. Erion, W.C. Guida, J.A. Montgomery, J.A. Secrist III, Application of crystallographic and modeling methods in the design of purine nucleoside phosphorylase inhibitors, Proc. Natl. Acad. Sci. USA 91 (1991) 11540–11544.
[10] J.A. Montgomery, Purine nucleoside phosphorylase: a target for drug design, Med. Res. Rev. 13 (3) (1993) 209–228.
[11] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M.A. Quail, M.A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J.E. Sulston, K. Taylor, S. Whitehead, B.G. Barrell, Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, Nature 393 (1998) 537–544.
[12] Editorial, Nat. Struct. Biol. 7, 2000, 87–88.
[13] D. Avarbock, J. Salem, L. Li, Z. Wang, H. Rubin, Cloning and characterization of a bifunctional RelA/SpoT homologue from *Mycobacterium tuberculosis*, Gene 233 (1999) 261–269.
[14] V.L. Schramm, Enzymatic N-riboside scission in RNA precursors, Curr. Opin. Chem. Biol. (1997) 323–331.
[15] V.L. Schramm, Enzymatic transition-state analysis and transition-state analogs, Methods Enzymol. 308 (1999) 301–355.
[16] W. Shi, L.A. Basso, D.S. Santos, P.C. Tyler, R.H. Furneaux, J.S. Blanchard, S.C. Almo, V.L. Schramm, Structures of purine nucleoside phosphorylase from *Mycobacterium tuberculosis* in complexes with Immucillin-H and its pieces, Biochemistry 40 (2001) 8204–8215.
[17] Collaborative Computational Project, Number 4, The CCP4 suite: programs for protein, Acta Crystallogr. D 50 (1994) 760–763.
[18] J. Navaza, AmoRe: an automated package for molecular replacement, Acta Crystallogr. A 50 (1994) 157–163.
[19] D.E. McRee, XtalView/Xfit—aversatile program for manipulating atomic coordinates and electron density, J. Struct. Biol. 125 (1999) 156–165.
[20] A.T. Brünger, X-PLOR Version 3.1: A System for Crystallography and NMR, Yale University Press, New Haven, 1992.
[21] R.A. Laskowski, M.W. MacArthur, D.K. Smith, D.T. Jones, E.G. Hutchingson, A.L. Morris, D. Naylor, D.S. Moss, J.M. Thorton, PROCHECK v.3.0-program to check the stereochemistry quality of the protein structures-operating instructions, 1994.
[22] Parmodel, A web server for comparative modeling and quality assessment of proteins structures. Available from: <http://www.biocristalografia.df.ibilce.unesp.br/tools/parmodel/index.php>.
[23] A.C.R. Martin, Scitech Software ProFit, 1992–2001.
[24] R. Koradi, Institut fuer Molekularbiologie und Biophysik, ETH Zurich Spectrospin AG, Faellanden, Switzerland, MOLMOL 2.6, 2003.

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid

José Henrique Pereira,[a] Jaim Simões de Oliveira,[b] Fernanda Canduri,[a,c] Marcio Vinicius Bertacine Dias,[a] Mário Sérgio Palma,[c,d] Luiz Augusto Basso,[b,e] Diógenes Santiago Santos[e]* and Walter Filgueira de Azevedo Jr[a,d]*

[a]Programa de Pós-Graduação em Biofísica Molecular, Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil, [b]Rede Brasileira de Pesquisa em Tuberculose Grupo de Microbiologia Molecular e Funcional, Departamento de Biologia Molecular e Biotecnologia, UFRGS, Porto Alegre, RS 91501-970, Brazil, [c]Center for Applied Toxinology, Institute Butantan, São Paulo, SP 05503-900, Brazil, [d]Laboratory of Structural Biology and Zoochemistry, CEIS/Department of Biology, Institute of Biosciences, UNESP, Rio Claro, SP 13506-900, Brazil, and [e]Centro de Pesquisa e Desenvolvimento em Biologia Molecular e Funcional, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS 90619-900, Brazil

Correspondence e-mail: diogenes@pucrs.br, walterfa@df.ibilce.unesp.br

Received 24 August 2004
Accepted 5 October 2004

PDB Reference: shikimate kinase–MgADP–shikimate complex, 1we2, r1we2sf.

Tuberculosis made a resurgence in the mid-1980s and now kills approximately 3 million people a year. The re-emergence of tuberculosis as a public health threat, the high susceptibility of HIV-infected persons and the proliferation of multi-drug-resistant strains have created a need to develop new drugs. Shikimate kinase and other enzymes in the shikimate pathway are attractive targets for development of non-toxic antimicrobial agents, herbicides and anti-parasitic drugs, because the pathway is essential in these species whereas it is absent from mammals. The crystal structure of shikimate kinase from *Mycobacterium tuberculosis* (MtSK) complexed with MgADP and shikimic acid (shikimate) has been determined at 2.3 Å resolution, clearly revealing the amino-acid residues involved in shikimate binding. This is the first three-dimensional structure of shikimate kinase complexed with shikimate. In MtSK, the Glu61 residue that is strictly conserved in shikimate kinases forms a hydrogen bond and salt bridge with Arg58 and assists in positioning the guanidinium group of Arg58 for shikimate binding. The carboxyl group of shikimate interacts with Arg58, Gly81 and Arg136 and the hydroxyl groups interact with Asp34 and Gly80. The crystal structure of MtSK–MgADP–shikimate will provide crucial information for the elucidation of the mechanism of the shikimate kinase-catalyzed reaction and for the development of a new generation of drugs against tuberculosis.

1. Introduction

The shikimate pathway is a seven-step biosynthetic route that generates chorismic acid from phosphoenol pyruvate and erythrose 4-phosphate. The shikimate pathway is an attractive target for the development of antimicrobial agents (Davies *et al.*, 1994) and herbicides (Coggins, 1989) because it is essential in algae, higher plants, bacteria and fungi, whilst being absent from mammals (Bentley, 1990). Several enzymes of this pathway have been submitted to structural study (Arcuri *et al.*, 2004; Pereira *et al.*, 2003) in order to propose inhibitors for these enzymes. More recently, by disruption of the *aroK* gene, which codes for the shikimate kinase enzyme, the shikimate pathway has been shown to be essential for the viability of *Mycobacterium tuberculosis* (Parish & Stoker, 2002). Shikimate kinase (SK; EC 2.7.1.71), the fifth enzyme of the pathway, catalyses the specific phosphorylation of the 3-hydroxyl group of shikimic acid using ATP as a co-substrate.

Three previously solved structures of SK from *Erwinia chrysanthemi* (EcSK; Krell *et al.*, 1998, 2001) show that SK belongs to the same structural family as nucleoside monophosphate (NMP) kinases. The NMP kinases are composed of three domains: the CORE, LID and NMP-binding (NMPB) domains. A characteristic feature of the NMP kinases is that

they undergo large conformational changes during catalysis (Vonrhein *et al.*, 1995).

Three functional motifs of nucleotide-binding enzymes are recognizable in *M. tuberculosis* SK, including a Walker A motif, a Walker 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 conserved sequence G*XXXX*GKT/S (Walker *et al.*, 1982), where *X* represents any residue. This motif forms the phosphate-binding loop (P-loop; Smith & Rayment, 1996). In addition to the Walker A motif, a second conserved sequence *ZZDXX*G called the Walker B motif (Walker *et al.*, 1982) is observed, where *Z* represents a hydrophobic residue. An Asp→Ser replacement exists in MtSK. The Walker B motif consensus in SKs is *ZZZ*TGGG and the second glycine (Gly80 in MtSK) has been implicated in hydrogen bonding to the γ-phosphate of ATP. The adenine-binding loop motif may be described as an I/VD*XXX*(*X*)*X*P sequence stretch (Gu *et al.*, 2002).

The shikimate-binding domain has previously been assigned on the basis of the difference Fourier map of EcSK and structural comparison with NMP kinases. Shikimate was included in the crystallization conditions (co-crystallization); however, the electron density was not clear enough to include shikimate in the previously solved crystallographic structure (Krell *et al.*, 1998).

Two crystal structures of SK from *M. tuberculosis* (MtSK; Gu *et al.*, 2002; PDB codes 1l4y and 1l4u) have revealed the dynamic role of the LID domain in catalysis, but the precise position and interactions between shikimate and MtSK were not unequivocally demonstrated as the shikimate-binding site was not occupied by the substrate. A previously reported molecular-modelling study of the complex between MtSK and shikimate also failed to predict all the intermolecular hydrogen bonds (de Azevedo *et al.*, 2002). Here, we describe the crystal structure of the MtSK–MgADP–shikimate ternary complex at 2.3 Å resolution, unequivocally revealing in detail the interactions of amino-acid residues with bound shikimate and the conformational changes upon substrate binding. The crystal structure of MtSK–MgADP–shikimate will 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.

2. Materials and methods

2.1. Crystallization

Cloning, expression and purification have been reported elsewhere (Oliveira *et al.*, 2001). MtSK was concentrated and dialyzed against 50 m*M* Na HEPES buffer pH 7.5 containing 0.5 *M* NaCl and 5.0 m*M* $MgCl_2$. This protein solution was brought to 13.0 m*M* in shikimate and ADP by addition of the pure solids and centrifuged prior to crystallization. The protein concentration was about 17.0 mg ml^{-1}. Crystals were obtained by the hanging-drop vapour-diffusion method. The well solution contained 0.1 *M* Na HEPES buffer pH 7.5, 10% 2-propanol, 35% PEG 3350 and the drops consisted of a mixture of 1.0 µl well solution and 1.5 µl protein solution.

2.2. Data collection and processing

The data set for MtSK–MgADP–shikimate was collected at a wavelength of 1.431 Å using the Synchrotron Radiation Source (Station PCr, LNLS, Campinas, Brazil; Polikarpov *et al.*, 1998) and a CCD detector (MAR CCD). The cryoprotectant contained 15% glycerol, 12% PEG 3350 and 3.5% propanol. The crystal was flash-frozen at 104 K in a cold nitrogen stream generated and maintained with an Oxford Cryosystem. The oscillation range used was 1.0°, the crystal-to-detector distance was 90 mm and the exposure time was 50 s. A data set containing 160 frames was collected and processed to 2.3 Å resolution using the program *MOSFLM* (Leslie, 1992) and scaled with *SCALA* (Collaborative Computational Project, Number 4, 1994).

2.3. Molecular replacement and crystallographic refinement

The crystal structure of MtSK–MgADP–shikimate was determined by standard molecular-replacement methods using the program *AMoRe* (Navaza, 1994), using as a search model the structure of MtSK–MgADP (PDB code 1l4y; Gu *et al.*, 2002). After translation-function computation the correlation was 65% and the *R* factor was 38.3%. The highest magnitude of the correlation coefficient function was obtained for the Euler angles α = 54.85, β = 85.32, γ = 91.90°. The fractional coordinates are $T_x = 0.7336$, $T_y = 0.5326$, $T_z = 0.2768$. The atomic positions obtained from molecular replacement were used to initiate the crystallographic refinement. Structure refinement was performed using *X-PLOR* (Brünger, 1992). During rigid-body refinement, the *R* factor decreased from 38.3 to 35.8%. Further refinement continued with simulated annealing using the slow-cooling protocol, followed by alternate cycles of positional refinement and manual rebuilding using *XtalView* (McRee, 1999). Finally, the positions of waters, MgADP and shikimate were checked and corrected in $F_{obs} - F_{calc}$ maps. The final model has an *R* factor of 20.7% and an R_{free} of 28.7%.

Root-mean-square deviation differences from ideal geometries for bond lengths, angles and dihedrals were calculated with *X-PLOR* (Brünger, 1992). The overall stereochemical quality of the final model for MtSK–MgADP–shikimate was assessed by the program *PROCHECK* (Laskowski *et al.*, 1994). Atomic models were superposed using the program *LSQKAB* from *CCP*4 (Collaborative Computational Project, Number 4, 1994).

The molecular-surface areas were calculated using the program *Swiss PDB Viewer* v.3.7 (http://www.expasy.org/spdbv), a probe radius of 1.4 Å and a fixed radius for all atoms.

3. Results and discussion

The crystals of MtSK–MgADP–shikimate were grown in the presence of 5.0 m*M* $MgCl_2$, 13.0 m*M* ADP and 13.0 m*M* shikimate. The data set of MtSK–MgADP–shikimate was

Table 1
Summary of data-collection statistics for MtSK–MgADP–shikimate.

Values in parentheses are for the highest resolution shell.

Unit-cell parameters	
$a = b$ (Å)	62.91
c (Å)	90.92
Resolution (Å)	29.74–2.30 (2.41–2.30)
Space group	$P3_221$
No. of measurements with $I > 2\sigma(I)$	34274
No. of independent reflections	9563
Overall redundancy	2.4
Completeness (%)	98.7 (98.7)
R_{sym}† (%)	3.0 (7.2)

collected at 2.3 Å (Table 1) using the Synchrotron Radiation Source (Polikarpov *et al.*, 1998) and the structure was solved by molecular replacement. There is one molecule of MtSK–MgADP–shikimate in the asymmetric unit, containing residues 2–166, Mg^{2+}, ADP, shikimate, two Cl^- ions and 144 water molecules (Fig. 1). Therefore, the enzyme crystallized as a MtSK–MgADP–shikimate dead-end ternary complex. The N-terminal methionine residue is not observed and the ten C-terminal residues (NQIIHMLESN) are disordered. Details of the refinement and the final model statistics are presented in Table 2. Analysis of the Ramachandran diagram φ–ψ plot shows that 93.4% of non-glycine residues lie in most favoured regions and there are no residues in the disallowed region (Fig. 2). The average B factor for main-chain atoms is 34.64 Å^2, whereas that for side-chain atoms is 35.68 Å^2 (Table 2). In order to accurately determine the position of shikimate binding in the active site of MtSK, larger final

Table 2
Summary of refinement statistics for MtSK–MgADP–shikimate.

Values in parentheses are for the highest resolution shell.

Resolution range (Å)	6.00–2.30
Reflections used for refinement	8885
No. of residues	165
No. of water O atoms	144
No. of ADP molecules	1
No. of metal ions (Mg^{2+})	1
No. of Cl^-	2
No. of shikimate molecules	1
Final R factor† (%)	20.7 (31.8)
Final R_{free}‡ (%)	28.7 (36.9)
B factors§ (Å^2)	
Main chain	34.64
Side chains	35.68
Waters	39.63
ADP	21.81
Mg^{2+}	34.70
Cl^-	37.65
Shikimate	26.15
Observed r.m.s.d. from ideal geometry	
Bond lengths (Å)	0.017
Bond angles (°)	1.905
Dihedrals (°)	22.125
Ramachandran plot	
Most favoured φ/ψ angles (%)	93.4
Disallowed φ/ψ angles (%)	0

† R factor = $100 \times \sum |F_{obs} - F_{calc}| / \sum F_{obs}$, the sums being taken over all reflections with $F/\sigma(F) > 2\sigma(F)$. ‡ R_{free} = R factor for 10% of the data that were not included during crystallographic refinement. § B values = average B values for all non-H atoms.

concentrations of ADP and shikimate in the drop (8 mM) were used than those used by Gu *et al.* (2002) to obtain the crystals (4 mM). The average B-factor value of 26.15 Å^2 obtained for shikimate indicates a higher order and occupancy of this substrate in the present structure than in the previously solved SK structures, where either shikimate was not

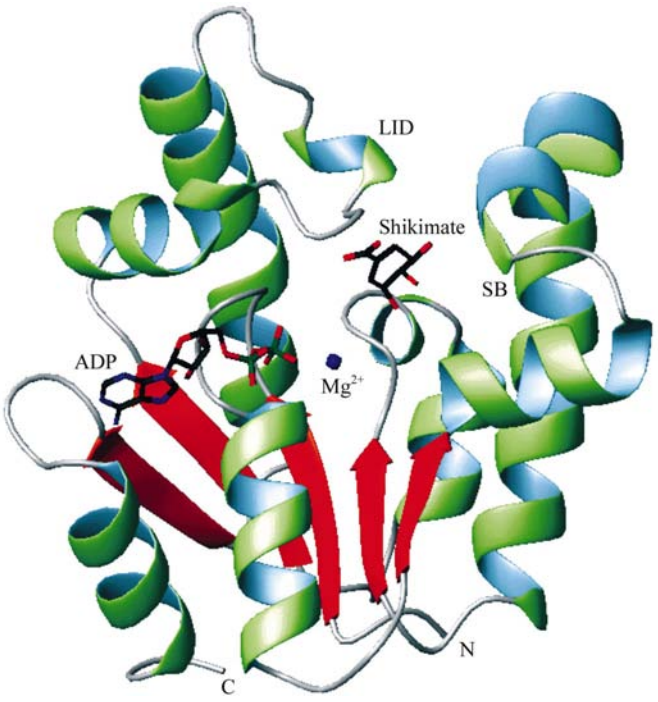


Figure 1
MtSK complexed with Mg^{2+}, ADP and shikimate. The LID (residues 112–124) and SB (residues 33–61) domains are responsible for large conformational changes during catalysis and are shown. The figure was prepared with *MOLMOL* (Koradi *et al.*, 1996).

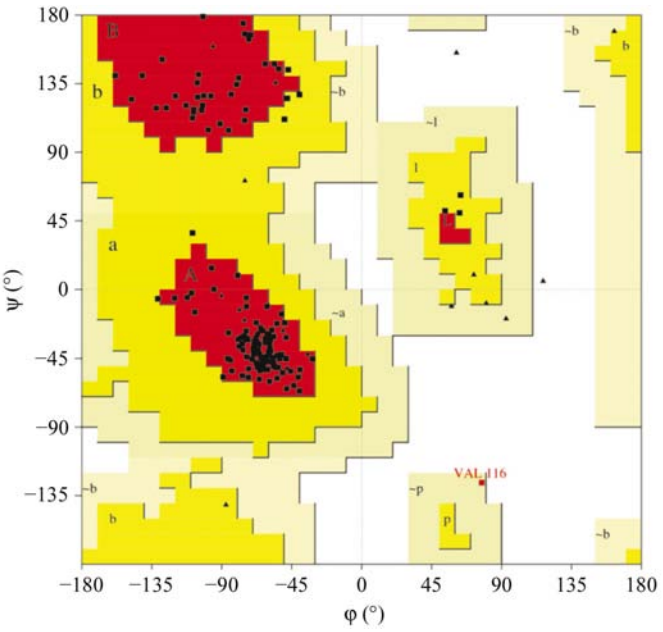


Figure 2
Ramachandran diagram for MtSK–MgADP–shikimate. 93.4% of non-glycine residues lie in the most favoured regions and no residues are in the disallowed region.

complexed or its electron density was too poor to accurately locate the substrate (Table 2) (Gu *et al.*, 2002; Krell *et al.*, 1998).

SK displays an α/β-fold and consists of five central parallel β-strands flanked by α-helices. As pointed out above, a characteristic feature of SKs is that they undergo large conformational changes during catalysis. There are two flexible regions of the structure that are responsible for movement: the SB and LID domains (Fig. 1), which in MtSK correspond to residues 33–61 and 112–124, respectively.

3.1. Interaction with ADP/Mg^{2+}

The structures of MtSK complexed with MgADP (Gu *et al.*, 2002) are highly similar, with a root-mean-square deviation of 0.19 Å for all pairs of C^{α} atoms. The adenine moiety of ADP is sandwiched between Arg110 and Pro155 as observed for the MtSK–MgADP structure (Gu *et al.*, 2002). The Arg110 in MtSK represents the first residue in a conserved motif, typically R*XX*(*X*)R, of the LID domain observed for P-loop kinases (Leipe *et al.*, 2003). The second conserved basic residue of this motif interacts with the γ-phosphate of ATP. In MtSK, this residue would be Arg117, which interacts with the α- and β-phosphate groups (Fig. 3), and thus the conserved motif of the LID domain for P-loop shikimate kinases would be $R(X)_{6-9}R$ (Fig. 4). The Arg117 residue may stabilize the transition state by neutralizing the developing negative charge on the β–γ bridge O atom (Hasemann *et al.*, 1996). The Pro155 is the last residue of the adenine-binding loop motif (residues 148–155 in MtSK), which was first recognized in AK and EcSK (Krell *et al.*, 1998) and has been described as an I/VD*XXX*(*X*)*X*P sequence stretch (Gu *et al.*, 2002). This motif forms a loop that wraps around the adenine moiety of ATP, connecting the β5-strand with the C-terminal α8-helix. The second (aspartate) residue and the last (proline) residue are not conserved in the adenine-binding loops of *aroK*-encoded SKs from *Escherichia coli* (Romanowski & Burley, 2002) and *Campylobacter jejuni* (PDB code 1via). However, they are

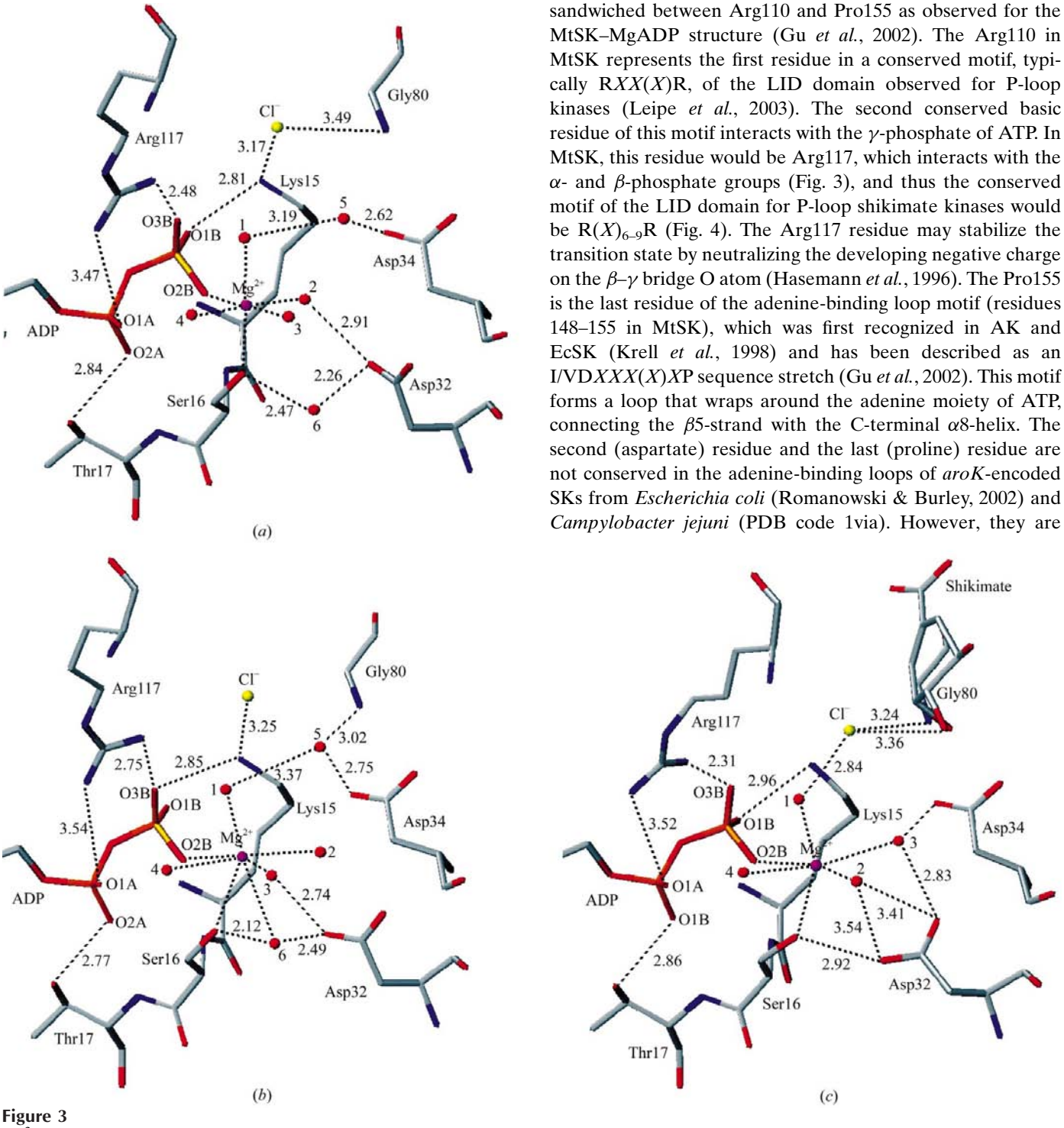


Figure 3
Mg^{2+} coordination. (*a*) 1l4y, (*b*) 1l4u (Gu *et al.*, 2002), (*c*) MtSK–MgADP–shikimate. For simplicity, only protein residues, atoms involved in binding of Mg^{2+} water and ADP are shown. Broken black lines represent hydrogen bonds or Mg^{2+} coordination. The distances are in Å. The following colour scheme was adopted: grey for carbon, red for oxygen, blue for nitrogen, yellow for chlorine, purple for magnesium and orange for phosphorus.

recognizable from structural alignment analysis using MtSK–MgADP as a reference structure. Accordingly, the I/VD-(*X*)*X*P primary sequence cannot be used to identify the adenine-binding loop sequences of all shikimate kinases. The carbonyl group of Arg153 residue forms hydrogen bonds to both the N6 atom of adenine and to a water molecule, which in turn hydrogen bonds to the N1 atom of adenine.

There are minor differences between the MtSK–MgADP and MtSK–shikimate–MgADP structures in the position of Mg^{2+}. In the MtSK–MgADP structure (PDB code 1l4y) a typical six-coordination is observed for Mg^{2+} (Fig. 3*a*), whereas magnesium six-coordination (or, more accurately, seven-coordination) is distorted in the structure of the Pt-derivative of MtSK–MgADP (PDB code 1l4u) (Fig. 3*b*). The structure of MtSK–MgADP–shikimate shows a distorted six-coordinated Mg^{2+} (Fig. 3*c*). Binding distances to Mg^{2+} are shown in Table 3.

The Mg^{2+} of MtSK–MgADP–shikimate interacts with a β-phosphate oxygen of ADP, Ser16 OG of the Walker A motif and four water molecules. In EcSK, the Mg^{2+} is four-coordinated, interacting with a β-phosphate O atom of ADP, Thr16 OG1, Asp32 OD2 and a water molecule held in position by Asp34 (Krell *et al.*, 1998). The Mg^{2+} binding in MtSK–MgADP (Gu *et al.*, 2002) and MtSK–MgADP–shikimate structures are somewhat similar: the six-coordinated magnesium ion structural water2 is held in place by a hydrogen bond to Asp32, which is conserved in all SKs (Fig. 4). Superimposition of MtSK–MgADP on MtSK–MgADP–shikimate showed that in the latter structure water1 and water4 are in equivalent positions but that shifts of 1.32 and 2.75 Å are observed for water2 and water3, respectively (Fig. 3). Water3 has moved to back the plane formed between the magnesium ion, the O2B atom of ADP and the side chain of Asp34. Water3 is held in place by Asp32 and Asp34 (Fig. 3). Asp32 also forms a hydrogen bond to Ser16 of the Walker A motif, whereas the interaction between Asp32 and Ser16 is *via* a bridging water molecule (water6) in the MtSK–MgADP structure (Gu *et al.*, 2002). A rotation of the Asp32 side-chain dihedral angles χ_1 and χ_2 leads to a direct interaction with Ser16 OG, which accounts for the exclusion of water6 from the magnesium-binding site in the ternary structure. The Mg^{2+}-coordinated water1 interacts with a chloride ion instead of interacting with Asp34 *via* a bridging water molecule (Fig. 3, water5) as observed in the MtSK–MgADP structure (Gu *et al.*, 2002). This chloride ion also interacts with the 3-hydroxyl group of shikimate and the backbone amide of Gly80. Moreover, Asp34 makes hydrogen bonds with the 2- and 3-hydroxyl groups of shikimate. Hence, the different mode of interaction observed for residue Asp34 arises from the presence of shikimate, which leads to the exclusion of water5 from the MtSK active site. In addition, the distance between ADP β-phosphate O2 and Mg^{2+} changes from 2.23 Å in our structure (Table 3) to 2.15 Å for the non-distorted six-coordinated magnesium ion in the MtSK–ADP structure (Gu *et al.*, 2002), thereby accounting for the slightly distorted six-coordination of Mg^{2+}. Interestingly, the two functions of the aspartate of the *ZZDXX*G Walker B motif (hydrogen bonding to Mg^{2+} water and to the hydroxy group of the serine or threonine of the Walker A motif) are taken over by the two conserved aspartate residues located at the last position of strand β2 (Asp32 in MtSK) and the first position of helix α2 (Asp34 in MtSK) of shikimate kinase and gluconate kinase (Kraft *et al.*, 2002). The Walker B motif consensus sequence in shikimate kinases has been proposed to be *ZZZ*TGGG (residues 75–81 in MtSK; Leipe *et al.*, 2003) and the second glycine (Gly80 in MtSK) has been implicated in hydrogen bonding the γ-phosphate of ATP. Nucleophilic attack on the γ-phosphate of ATP will be most facilitated by metal-ion binding

Table 3
Binding of Mg^{2+} in MtSKs.

n.o., not observed.

Atom	MtSK–MgADP–shikimate (Å)	1l4y (Å)	1l4u (Å)
Ser16 OG	2.36	2.15	2.57
ADP O2B	2.23	2.15	2.39
Water1 O	2.39	2.24	2.52
Water2 O	2.23	2.04	2.77
Water3 O	2.75	2.25	2.37
Water4 O	2.97	2.05	2.94
Water5 O	n.o.	n.o.	3.06

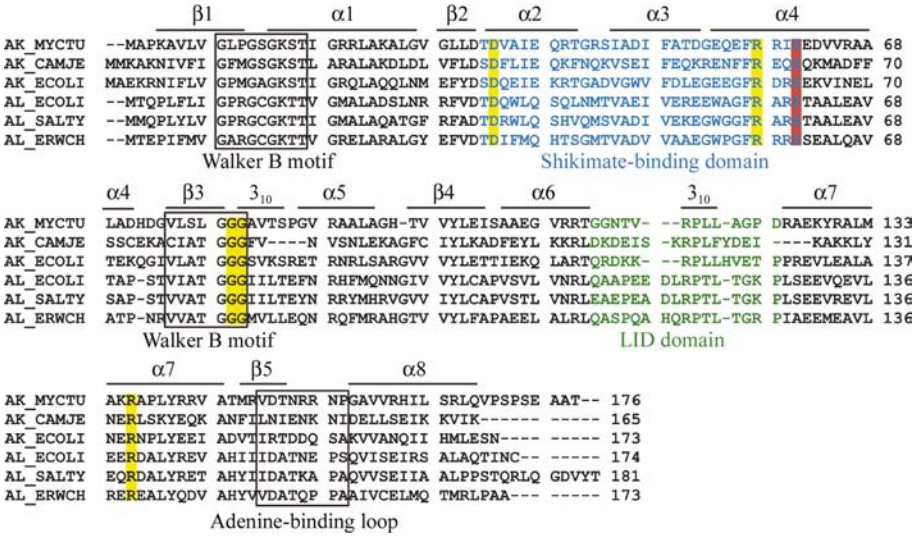


Figure 4
Alignment of shikimate kinases showing that the residues identified in binding of shikimate are conserved. The secondary-structure assignment for MtSK–shikimate is shown above the sequence and the locations of linear motifs are labelled. Shaded regions in yellow are the residues involved in binding of shikimate and the shaded region in red is the residue Glu61 (MtSK). Stretches in blue and green are the shikimate-binding domain and the LID domain, respectively. AK_MYCTU, *M. tuberculosis* SK I; AK_CAMJE, *C. jejuni* SK I; AK_ECOLI, *E. coli* SK I; AL_ECOLI, *E. coli* SK II; AL_SALTY, *S. typhimurium* SK II; AL_ERWCH, *E. chrysanthemi*.

Table 4
Direct and water-mediated hydrogen bonding of shikimate in MtSK.

All distances <3.6 Å are shown. n.o., not observed.

Shikimate	Atom	Distances (Å)	MtSK atom hydrogen bonded to water molecule	Distances (Å)
Hydroxyl groups				
O1	Gly80 N	3.10		
	Asp34 OD1	2.82		
	Asp34 OD2	2.57		
	Water320 O	2.46	Gly79 O	2.96
			Gly80 N	3.13
			Gly81 N	2.51
			Ala82 N	3.54
			Arg58 NH1	3.12
			Glu61 OE2	2.99
O2	Asp34 OD1	2.66		
	Asp34 OD2	2.85		
O3	n.o			
Carboxyl group				
O4	Gly81 N	3.22		
	Arg136 NH2	2.68		
O5	Arg58 NH2	2.67		
	Gly81 N	3.49		
	Arg136 NH2	2.34		

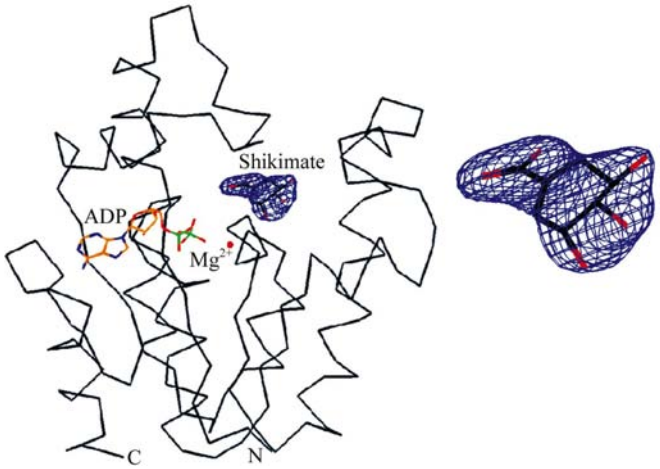


Figure 5
$F_{obs} - F_{calc}$ electron density contoured at 3.0σ showing the binding of shikimate. The figure was prepared with *XtalView* (McRee, 1999).

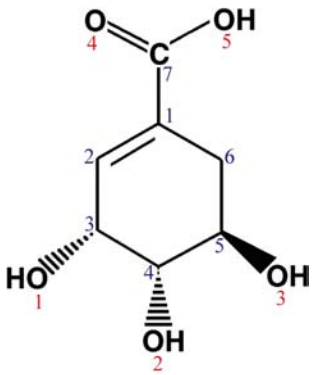


Figure 6
The molecular structure of shikimate shows the carboxyl group and three hydroxyl groups. C atoms are numbered in blue and O atoms in red.

(Mg^{2+}) to the β- and γ-phosphate groups, whereas departure of the leaving group will be most favoured in a structure with metal binding to the α- and β-phosphate groups (Jencks, 1975*a*). In enzyme reactions, the enzyme may facilitate the reaction and contribute to its specificity simply by favouring the binding of metal in the most favourable possible structure for the particular reaction that is being catalyzed. Although the mechanism of action of MtSK is still unknown, the interaction between Gly80 and the chloride ion may indicate that the latter occupies the γ-phosphate position as the chemical reaction proceeds, thereby suggesting that the interaction between Gly80 and γ-phosphate of ATP may play a role in the catalytic mechanism of MtSK, as discussed below.

3.2. Shikimate binding

The shikimate-binding domain, which follows strand β2, consists of helices α2 and α3 and the N-terminal region of helix α4 (residues 33–61). A peak of more than 3σ in the final $F_{obs} - F_{calc}$ difference Fourier map clearly indicates the position of bound shikimate in the electron density (Fig. 5).

The chemical structure of shikimic acid [3*R*-(3α,4α,5β)]-3,4,5-trihydroxy-1-cyclohexene-1-carboxylic acid] is shown in Fig. 6. 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, the main-chain NH group of Gly80 and a water molecule (Fig. 7). This water molecule in turn mediates interactions with the side chains of SB residues Arg58 and Glu61, Walker B residues Gly79, Gly80 and Gly81, and Ala82. The 2-hydroxyl group of shikimate hydrogen bonds to the side chain of Asp34. The distances of direct and water-mediated hydrogen bonds in the shikimate binding of MtSK are shown in Table 4.

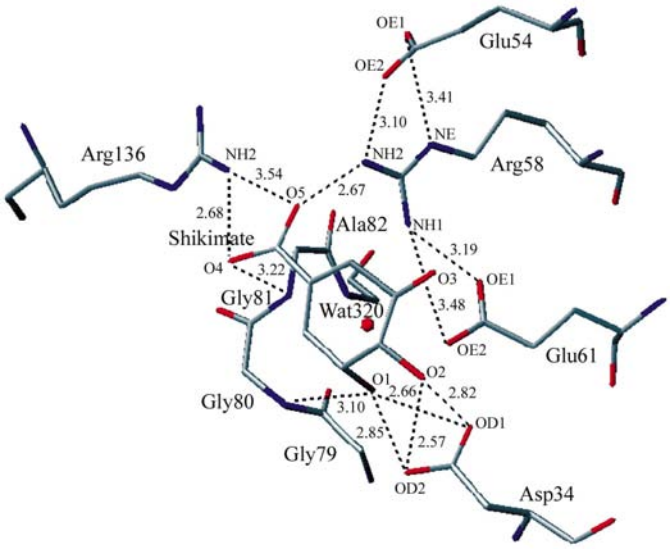


Figure 7
Polar interactions involved in the binding of shikimate to the MtSK active site. Hydrogen bonds between shikimate and protein groups and interactions between protein groups are represented as broken black lines. The interaction distances are in Å.

Glu61 is conserved in both *aroK*- and *aroL*-encoded shikimate kinase enzymes and has been implicated in shikimate binding (Gu *et al.*, 2002; Krell *et al.*, 1998; Romanowski & Burley, 2002). Krell and coworkers proposed that Glu61 is suitably positioned to bind the 5-hydroxyl group of shikimate in EcSK (Krell *et al.*, 1998). However, the amino-acid residues comprising the shikimate-binding domain were not clearly demonstrated because the electron density was not sufficient to position the shikimate molecule in the structure. In our structure, the Glu61 side chain forms a water-mediated interaction with the 3-hydroxyl group of shikimate. In addition, Glu61 forms a hydrogen bond and a salt bridge with the conserved Arg58 and assists in positioning the guanidinium group of Arg58 for substrate binding *via* interactions with the carboxylate group of shikimate. Glu61 OE2 makes a hydrogen bond with the amide N atom of Ala82. Moreover, Arg58 NH1 forms a hydrogen bond with Gly80 O. Therefore, Glu61 plays an important role in positioning a shikimate molecule, even though it is not directly involved in substrate binding. Furthermore, Glu61 is conserved in all SKs sequenced so far, which corroborates its role in anchoring Arg58 in the shikimate-binding site. The Glu54 carboxylate group also appears to anchor the guanidinium group of Arg58 for interaction with the shikimate carboxylate group, as suggested in the MtSK–MgADP structure (Gu *et al.*, 2002). However, Glu54 is not conserved (Fig. 4), except for *aroK*-encoded shikimate kinases.

3.3. Chloride ion binding

Chloride ions have been proposed to play an important role in determining the affinity of *E. chrysanthemi* SK for ATP and shikimate (Cerasoli *et al.*, 2003). However, no chloride ions in the active-site cavity are found in equivalent positions when the MtSK structures are superimposed on both *E. chrysanthemi* SK (Krell *et al.*, 1998) and the K15M mutant (Krell *et al.*, 2001) SKs. We have found only two chloride ions in MtSK–MgADP–shikimate that have equivalents in the MtSK–MgADP complex structures (1l4u and 1l4y; Gu *et al.*, 2002). One of them is bound to the active-site cavity and makes hydrogen bonds with 3-hydroxyl group of shikimate, Gly80 NH and water1 (Fig. 8), whereas the other is on the enzyme surface at a large distance from the active site. The conserved residue Arg117 in the LID domain is involved in ADP binding by forming two hydrogen bonds with the α- and β-phosphate O atoms (Fig. 7). Lys15 forms a hydrogen bond with a β-phosphate O atom (Fig. 3) and the chloride ion in the MtSK–MgADP structures (Fig. 8). The main-chain NH of Gly80 is hydrogen bonded to the chloride ion in the binary (1l4y) and ternary complex structures (Table 5). These residues are located in the vicinity of where the chemical reaction occurs and may thus play a critical role in transition-state stabilization. Consistent with this proposal, the K15M mutant of EcSK showed no detectable enzyme activity (Krell *et al.*, 2001), although it was in fact a K15M/P115L double mutant. Gly80 of the Walker B motif has been implicated in hydrogen bonding to the γ-phosphate of ATP (Krell *et al.*, 1998). The

Table 5
Binding of chloride ion in MtSK–MgADP and MtSK–MgADP–shikimate structures.

Chloride ion	Atom	Distance (Å)		
		1l4y	1l4u	MtSK–MgADP –shikimate
Cl180	Shikimate hydroxyl group O1	4.48†	4.66†	3.36
	Gly80 N	3.49	3.69	3.24
	Lys15 NZ	3.17	3.25	3.94
	Mg^{2+} Water1 O	3.87	3.75	2.84

† The distance was measured with the MtSK–MgADP structures (1l4y and 1l4u) superimposed on the MtSK–MgADP–shikimate structure.

distance between the chloride ion and the hydroxyl group bound to carbon 3 of shikimate (to which a phosphoryl group would be transferred) is 3.36 Å in the MtSK–MgADP–shikimate structure (Fig. 8). Lys15 and Gly80 are in slightly different positions in the ternary complex compared with the MtSK–MgADP binary complex. The distance between O1B of the β-phosphate of ATP (2.81 Å) and Lys15 is 2.96 Å in the ternary complex (Fig. 3). However, the distance between Lys15 and the chloride ion is greater in the ternary complex (3.94 Å, Table 5). The distance between Cl^- and the main-chain NH of Gly80 (3.49–3.69 Å) in the binary complex is shortened to 3.24 Å in the ternary complex. There appears to be a concerted movement of Lys15, Gly80 and the Cl^- ion (Fig. 8). This chloride ion appears to be shifted by 1.76 and

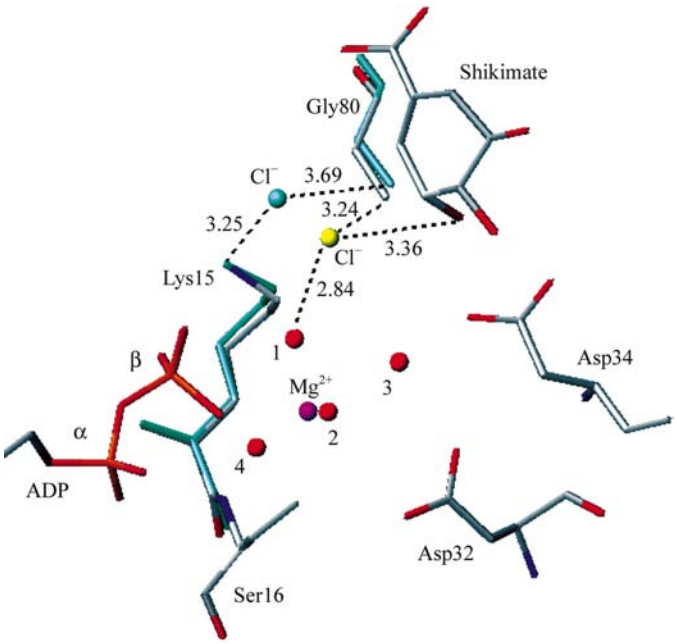


Figure 8
Chloride ions bound to the active sites of the MtSK–MgADP–shikimate and MtSK–MgADP (1l4u; Gu *et al.*, 2002) structures. For clarity, only residues Lys15, Ser16, Asp32, Asp34 and Gly80, shikimate, the α- and β-phosphates of ADP, Mg^{2+} and Cl^- ions and Mg^{2+}-coordinated waters are shown. The chloride ion hydrogen bonds are represented as broken black lines and distances are in Å. The Cl^- ion and the Lys15 and Gly80 residues of the superimposed MtSK–MgADP structure 1l4u are coloured cyan.

1.52 Å in MtSK–MgADP–shikimate compared with the MtSK–MgADP 1l4u and 1l4y complexes, respectively. It is thus tempting to suggest that the chloride ion in the ternary complex occupies the phosphoryl position in the reaction coordinate to shikimate 3-phosphate formation.

3.4. Conformational changes upon substrate binding

As pointed out above, MtSK belongs to the family of nucleoside monophosphate (NMP) kinases, which are composed of three domains: the CORE, LID and NMP-binding domains. Kinases should undergo large movements during catalysis to shield their active centre from water in order to avoid ATP hydrolysis (Jencks, 1975*b*). The NMP and LID domains of NMP kinases have been shown to undergo large motions that are independent, in agreement with the observed random bi-bi kinetics (Vonrhein *et al.*, 1995). Ligand-induced changes in the secondary structure of EcSK have been detected by comparing the circular-dichroism spectra of free enzyme, EcSK–shikimate binary complex and a ternary complex of EcSK, shikimate and adenylyl imidodiphosphate, an ATP analogue (Krell *et al.*, 1998). Alignment of C^{α} positions of the MtSK–MgADP–shikimate dead-end ternary complex and the MtSK–MgADP binary complex structures shows that the LID and SB domains undergo noticeable concerted movements towards each other (Fig. 9). The structural alignment included all residues and yielded r.m.s. deviation values of 0.56 and 0.54 Å for 1l4u and 1l4y, respectively. There is a shift of the LID domain, with an r.m.s. deviation of 1.33 Å for residues 112–124. The SB domain shift is somewhat smaller, with a calculated r.m.s. deviation of 0.74 Å for residues 33–61. The shikimate-binding cavity is delineated mainly by residues from the LID and SB domains, the Walker B motif and Arg136 from the α7 helix. A close inspection of the residues involved in these movements shows that the side chains of Val116, Pro118 and Leu119 from the LID domain and Ile45, Ala46, Glu54, Phe57 and Arg58 from the SB domain shifted upon shikimate binding to the MtSK–MgADP binary complex (Fig. 9). In MtSK with bound MgADP and shikimate, a cluster of hydrophobic contacts is formed between LID residues Val116, Pro118 and Leu119 and SB residues Ala46 and Phe49, which account for the stabilization of the partially closed shikimate-binding site cavity.

The conformational changes described above result in closure of the MtSK binding site, as shown by the reduction in molecular-surface area of MtSK–shikimate–MgADP compared with MtSK–MgADP (Fig. 10). The ADP, Mg^{2+}, shikimate, Cl^{-} ions and water molecules were removed prior to calculation. The calculated values are approximately 7246 $Å^2$ for MtSK complexed with MgADP (1l4u) and 6915 $Å^2$ for MtSK complexed with MgADP and shikimate. Thus, approximately 330 $Å^2$ of molecular surface is buried on shikimate binding. Since both MtSK–MgADP and MtSK–MgADP–shikimate have been crystallized in the same space group with similar values for the unit-cell parameters (Gu *et al.*, 2001) (Table 1) and residues 114–124 from the LID domain and residues 33–61 from the SB domain form no crystal contacts with symmetry-related MtSK molecules, the conformational changes observed cannot merely be a reflection of the different crystal-packing arrangements.

The MtSK–MgADP–shikimate structure should represent a partially, but not totally, closed structure, 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 MtSK enzyme prior to MgADP dissociation from its active site would result in an inactive abortive complex. Consistent with these structural results, measurements of EcSK intrinsic tryptophan fluorescence (Trp54) on shikimate binding to either EcSK or the EcSK–MgADP binary complex showed a modest synergism of binding between these substrates, since the dissociation constant value for shikimate (K_d = 0.72 m*M*) decreased to 0.3 m*M* in the presence of 1.5 m*M* ADP (Idziak *et al.*, 1997). Measurements of the quenching of protein fluorescence of the *aroL*-encoded EcSK upon nucleotide binding demonstrated the dissociation-constant values for ADP

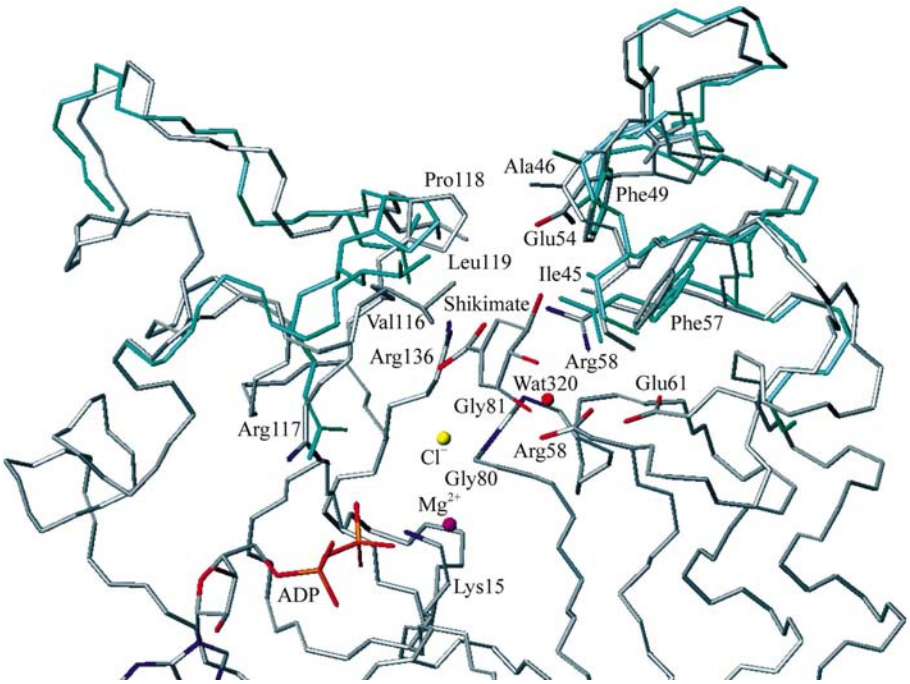


Figure 9
Shikimate binding in the MtSK–MgADP–shikimate structure. The C^{α} traces of MtSK–MgADP–shikimate (grey) and MtSK–MgADP (cyan) (1l4u; Gu *et al.*, 2002) were superimposed. The side chains with large shifts owing to shikimate binding are shown for both structures: Val116, Pro118 and Leu119 from LID domain and Ile45, Ala46, Glu54, Phe57 and Arg58 from the SB domain. For clarity, only residues from the LID and SB domains of 1l4u (cyan) are shown.

and ATP to be 1.7 and 2.6 m*M*, respectively (Krell *et al.*, 2001). The K_m for ATP (620 μ*M*) was found to be approximately four times lower than the dissociation constant in the absence of shikimate (Krell *et al.*, 2001). These results prompted the proposal that the conformational changes in the enzyme associated with the binding of the first substrate lead to an increase in the affinity for the second substrate. However, even if this holds for EcSK, it does not appear to hold for MtSK since the apparent dissociation-constant values for ATP (89 μ*M*) and shikimate (44 μ*M*) are similar to their K_m values, 83 μ*M* for ATP and 41 μ*M* for shikimate considering a random-order bi-bi enzyme mechanism (Gu *et al.*, 2002). Moreover, no evidence for synergism between shikimate and ATP could be observed in substrate binding to EcSK in a chloride-free buffer system (Cerasoli *et al.*, 2003). The K15M EcSK mutant has been crystallized in an open conformation that is proposed to presumably be equivalent to an apo-enzyme structure in which neither shikimate nor ADP (or ATP) would be bound (Krell *et al.*, 2001). This mutant was produced to evaluate the role of the conserved Lys15 of the Walker A motif (Krell *et al.*, 2001). However, an unwanted point mutation in the LID domain (Pro115Leu) was detected during refinement of the model and extensive contacts between LID domains of neighbouring EcSK enzymes were observed. It therefore appears unwarranted to consider the double K15M/P115L EcSK mutant a model for the apo enzyme. The incomplete LID-domain closure observed in the crystal structure presented here may suggest that the γ-phosphate of ATP plays a crucial role in the completion of the domain movement, as has also been proposed by others (Krell *et al.*, 1998). The equilibrium constant for the intramolecular hydrolysis of bound ATP to bound ADP and phosphate at enzyme active sites is considerably larger than the equilibrium constant for ATP hydrolysis in solution (Jencks, 1975*a*). Accordingly, the loss of two water molecules (water5 and water6) described in §3.1 is consistent with the exclusion of water molecules from the active site owing to the partial closure of MtSK upon shikimate binding in order to minimize ATP hydrolysis.

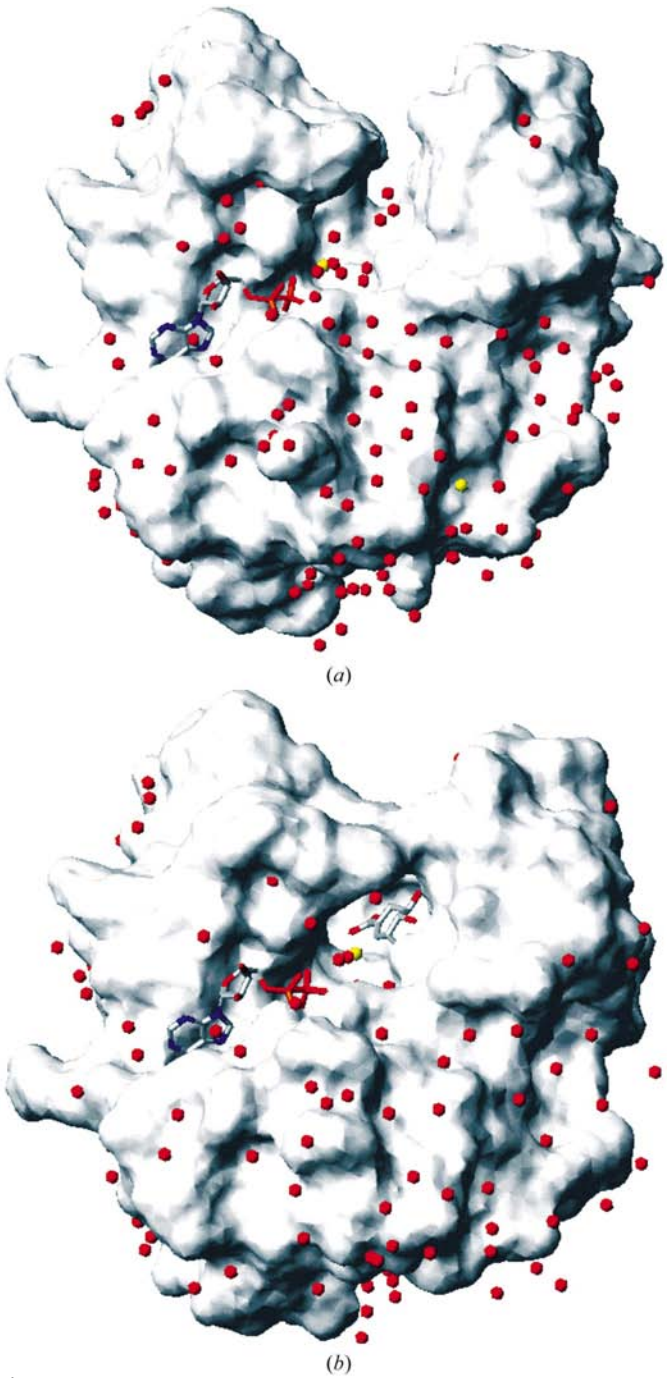


Figure 10
Molecular surface of (*a*) MtSK–MgADP (1l4u; Gu *et al.*, 2002) and (*b*) MtSK–MgADP–shikimate structures. Chloride ion is coloured yellow in MtSK–MgADP and MtSK–MgADP–shikimate. For water molecules only O atoms are shown.

4. Conclusions

The residues identified in the binding of shikimate whether directly or indirectly (Asp34, Arg58, Glu61, Gly79, Gly80, Gly81 and Arg136) are conserved in all SKs encoded by *aroK* and *aroL* genes (Fig. 4). The structures of SKs deposited in the PDB were superimposed and the positions of the residues involved in binding between shikimate and SK are highly conserved in *aroK*-encoded proteins (shikimate kinase from *M. tuberculosis* and *C. jejuni*) and *aroL*-encoded proteins (shikimate kinase from *E. chrysanthemi*). The conserved active site may account for the similar K_m values found for MtSK (*aroK*-encoded) and EcSK (*aroL*-encoded); however, it does not support the difference observed in the K_m value of SKI from *E. coli*. Romanowski & Burley (2002) proposed that the substitution of Leu83 (EcSK) for Lys86 in *E. coli* SKI disrupts the binding to shikimate. The loss of a stabilizing hydrophobic residue at this position may explain the significantly lower affinity of SKI from *E. coli* for shikimate compared with that of SKII. However, superimposition of SKI from *E. coli* with MtSK–shikimate indicates that Lys86 is distant from shikimate, which does not support Romanoswski and Burley's proposition. Moreover, even though a substitution of Leu83 (EcSK) for Thr84 occurs in MtSK, which would result in the loss of a stabilizing hydrophobic residue, the K_m values for MtSK and EcSK are similar.

Superimposition of SKI from *E. coli* on the MtSK–MgADP–shikimate structure shows the overlap of Leu123 onto the shikimate structure; however, we cannot affirm that the binding sites are different, since the structure of SKI from *E. coli* was crystallized without shikimate and shikimate binding in the SKI structure may promote further conformation changes that may include a displacement of the Leu123 side chain.

Here, we describe the residues involved in shikimate binding and conformational changes upon substrate binding to the MtSK–MgADP complex, resulting in a partially closed structure. A complete active-site closure could be achieved in a complex of MtSK with shikimate, Mg^{2+} and the non-hydrolysable ATP analogue adenosine 5′-(β,γ-methylene) triphosphate (AMP-PCP). 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 (Leipe *et al.*, 2003). The availability of the *M. tuberculosis* shikimate kinase structure complexed with shikimate should allow the molecular design of specific SK inhibitors that target either the shikimate- and ATP-binding sites or the shikimate-binding site only. Moreover, the knowledge of functional factors that lead to active-site closure could be used to design inhibitors that force MtSK into a closed conformation that would be unable to catalyze the phosphoryl transfer to shikimate.

This work was supported by grants from FAPESP (SMOLBNet, Proc. 01/07532-0, 02/05347-4, 04/00217-0), CNPq, CAPES and Instituto do Milênio (CNPq-MCT). WFA (CNPq, 300851/98-7), MSP (CNPq, 500079/90-0) and LAB (CNPq, 520182/99-5) are researchers of the Brazilian Council for Scientific and Technological Development.

References

Arcuri, H. A., Canduri, F., Pereira, J. H., Silveira, N. J. F., Camara, J. C. Jr, Oliveira, J. S., Basso, L. A., de Palma, M. S., Santos, D. S. & de Azevedo, W. F. (2004). *Biochem. Biophys. Res. Commun.* **320**, 979–991.

Azevedo, W. F. de, Canduri, F., de Oliveira, J. S., Basso, L. A., Palma, M. S., Pereira, J. H. & Santos, D. S. (2002). *Biochem. Biophys. Res. Commun.* **295**, 142–148.

Bentley, R. (1990). *Crit. Rev. Biochem. Mol. Biol.* **25**, 307–384.

Brünger, A. T. (1992). *X-PLOR Version 3.1: A System for Crystallography and NMR*. New Haven: Yale University Press.

Cerasoli, E., Kelly, S. M., Coggins, J. R., Lapthorn, A. J., Clarke, D. T. & Price, N. C. (2003). *Biochim. Biophys. Acta*, **1648**, 43–54.

Coggins, J. R. (1989). *Herbicides and Plant Metabolism*, edited by A. Dodge, pp. 97–112. Cambridge University Press.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Davies, G. M., Barrett-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**, 403–406.

Gu, Y., Reshetnikova, L., Li, Y., Wu, Y., Yan, H., Singh, S. & Ji, X. (2002). *J. Mol. Biol.* **319**, 779–789.

Gu, Y., Reshetnikova, L., Li, Y., Yan, H., Singh, S. V. & Ji, X. (2001). *Acta Cryst.* D**57**, 1870–1871.

Hasemann, C. A., Istvan, E. S., Uyeda, K. & Deisenhofer, J. (1996). *Structure*, **4**, 1017–1029.

Idziak, C., Price, N. C., Kelly, S. M., Krell, T., Boam, D. J., Lapthorn, A. J. & Coggins, J. R. (1997). *Biochem. Soc. Trans.* **25**, S627.

Jencks, W. P. (1975*a*). In *Catalysis in Chemistry and Enzymology*. New York: Dover.

Jencks, W. P. (1975*b*). *Adv. Enzymol.* **43**, 219–410.

Koradi, R., Billeter, M. & Wüthrich, K. (1996). *J. Mol. Graph.* **14**, 51–55.

Kraft, L., Sprenger, G. A. & Lindqvist, Y. (2002). *J. Mol. Biol.* **318**, 1057–1069.

Krell, T., Coggins, J. R. & Lapthorn, A. J. (1998). *J. Mol. Biol.* **278**, 983–997.

Krell, T., Maclean, J., Boam, D. J., Cooper, A., Resmini, M., Brocklehurst, K., Kelly, S. M., Price, N. C., Lapthorn, A. J. & Coggins, J. R. (2001). *Protein Sci.* **10**, 1137–1149.

Laskowski, R. A., MacArthur, M. W., Smith, D. K., Jones, D. T., Hutchinson, E. G., Morris, A. L., Naylor, D., Moss, D. S. & Thornton, J. M. (1994). *PROCHECK* v.3.0. *Program to Check the Stereochemistry Quality of Protein Structures. Operating Instructions.* http://www.ccp14.ac.uk/ccp/ccp14/ftp-mirror/llnlrupp/Procheck_NT/manual.txt.

Leipe, D. D., Koonin, E. V. & Aravind, L. (2003). *J. Mol. Biol.* **333**, 781–815.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.

McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Oliveira, J. S., Pinto, C. A., Basso, L. A. & Santos, D. S. (2001). *Protein Expr. Purif.* **22**, 430–435.

Parish, T. & Stoker, N. G. (2002). *Microbiology*, **148**, 3069–3077.

Pereira, J. H., Canduri, F., de Oliveira, J. S., da Silveira, N. J. F., Basso, L. A., Palma, M. S., de Azevedo, W. F. & Santos, D. S. (2003). *Biochem. Biophys. Res. Commun.* **312**, 608–614.

Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E. E., Garratt, R. C. & Craievich, A. (1998). *J. Synchrotron Rad.* **5**, 72–76.

Romanowski, M. J. & Burley, S. K. (2002). *Proteins Struct. Funct. Genet.* **47**, 558–562.

Smith, C. A. & Rayment, I. (1996). *Biophys. J.* **70**, 1590–1602.

Vonrhein, C., Schlauderer, G. J. & Schulz, G. E. (1995). *Structure*, **3**, 483–490.

Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). *EMBO J.* **1**, 945–951.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Inorganic Biochemistry 99 (2005) 368–375
JOURNAL OF Inorganic Biochemistry
www.elsevier.com/locate/jinorgbio

Electron transfer kinetics and mechanistic study of the thionicotinamide coordinated to the pentacyanoferrate(III)/(II) complexes: a model system for the in vitro activation of thioamides anti-tuberculosis drugs

Eduardo H.S. Sousa [a], Daniel L. Pontes [a], Izaura C.N. Diógenes [a], Luiz G.F. Lopes [a], Jaim S. Oliveira [b], Luiz A. Basso [b], Diógenes S. Santos [c], Ícaro S. Moreira [a,*]

[a] *Departamento de Química Orgânica e Inorgânica, Universidade Federal de Ceará, Cx. Postal 12200, Campus do Pici s/n, Fortaleza, CE 60455-760, Brazil*
[b] *Departamento de Biologia Molecular e Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS 91501-970, Brazil*
[c] *Centro de Pesquisa e Desenvolvimento em Biologia Molecular e Funcional, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS 90619-900, Brazil*

Received 15 June 2004; received in revised form 24 September 2004; accepted 5 October 2004
Available online 6 November 2004

Abstract

The mechanism of activation thioamide-pyridine anti-tuberculosis prodrugs is poorly described in the literature. It has recently been shown that ethionamide, an important component of second-line therapy for the treatment of multi-drug-resistant tuberculosis, is activated through an enzymatic electron transfer (ET) reaction. In an attempt to shed light on the activation of thioamide drugs, we have mimicked a redox process involving the thionicotinamide (thio) ligand, investigating its reactivity through coordination to the redox reversible $[Fe^{III/II}(CN)_5(H_2O)]^{2-/3-}$ metal center. The reaction of the Fe^{III} complex with thionicotinamide leads to the ligand conversion to the 3-cyanopyridine species coordinated to a Fe^{II} metal center. The rate constant, $k_{et} = 10\ s^{-1}$, was determined for this intra-molecular ET reaction. A kinetic study for the cross-reaction of thionicotinamide and $[Fe(CN)_6]^{3-}$ was also carried out. The oxidation of thionicotinamide by $[Fe(CN)_6]^{3-}$ leads to formation of mainly 3-cyanopyridine and $[Fe(CN)_6]^{4-}$ with a $k_{et} = (5.38 \pm 0.03)\ M^{-1}\ s^{-1}$ at 25 °C, pH 12.0. The rate of this reaction is strongly dependent on pH due to an acid–base equilibrium related to the deprotonation of the R-SH functional group of the imidothiol form of thionicotinamide. The kinetic results reinforced the assignment of an intra-molecular mechanism for the ET reaction of $[Fe^{III}(CN)_5(H_2O)]^{2-}$ and the thioamide ligand. These results can be valuable for the design of new thiocarbonyl-containing drugs against resistant strains of *Mycobacterium tuberculosis* by a self-activating mechanism.

Keywords: Thionicotinamide; Reactivity; Pentacyanoferrate(III)/(II); Electron transfer; Tuberculosis

1. Introduction

The *Mycobacterium tuberculosis inhA*-encoded NADH-dependent enoyl reductase enzyme is a bonafide target for isoniazid (INH), the primary drug used to treat tuberculosis, in *M. tuberculosis* and plays an important role in cell wall biosynthesis [1–4]. 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 [5–8]. The KatG-produced acylpyridine fragment of isoniazid is

* Corresponding author. Tel.: +55 85 2889412/9938; fax: +55 85 2889978.
E-mail address: icarosm@dqoi.ufc.br (Í.S. Moreira).

0162-0134/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jinorgbio.2004.10.004

covalently attached to the C4 position of NADH, and forms a binary complex with the wild-type enoyl reductase of *M. tuberculosis* [9] with a dissociation constant value lower than 0.4 nM [10]. Mutations in the *inhA* structural gene, in the *inhA* locus promoter region, and deletions of, or missense mutations in, the *katG* coding sequence have been associated with resistance to isoniazid in clinical isolates of *M. tuberculosis* [11]. In trying to find better alternatives to isoniazid, we have tested the ability of the pentacyano(isoniazid)ferrate(II) complex to inhibit both wild-type and isoniazid-resistant I21V mutant enzymes [12]. The in vitro kinetics of inactivation of both wild-type and I21V InhA enzymes by $[Fe^{II}(CN)_5(INH)]^{3-}$ demonstrate that this process requires no activation by *M. tuberculosis* KatG enzyme (a multifunctional hemeprotein catalase-peroxidase responsible for the activation of isoniazid prodrug) [9,13] no need for the presence of NADH, and its mechanism of action probably involves interaction with the NADH binding site of wild-type and mutant enoyl reductases [12]. This inorganic isoniazid analog may represent a new class of lead compounds to the development of anti-tubercular agents aiming at inhibition of a validated target.

Multi-drug-resistant tuberculosis (MDRTb) has been considered as an emerging epidemic that affects more that 50 million people worldwide [14–17]. 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, defined as resistant to isoniazid and rifampicin, and 79% of MDR-TB cases are now ''super strains'', resistant to at least three of the four main drugs used to treat TB [18]. Resistance to front-line drugs like INH has led to the use of thioamide drugs, which are less active and more toxic, as components of second-line therapies [19]. Recently, a flavoprotein monooxygenase, EtaA, was identified as the activating enzyme of thioamide anti-tuberculosis prodrugs [20,21]. As postulated for most of the anti-tuberculosis prodrugs, the proposed mechanism of activation involves an electron transfer (ET) reaction as a crucial pathway [9,13,22]. In spite of the important advances made in this research area, the mechanism of activation of thionamide-containing drugs is poorly understood. Since the activation of the prodrugs is postulated to occur through ET reactions, and based on the well known affinity of the nitrogen of the pyridine ring of the ligands towards the coordination to pentacyanoferrate moiety [23], we undertook a simple and suitable redox reversible $[Fe(CN)_5(H_2O)]^{2-/3-}$ probe metal center to mimic the mechanism of the prodrugs activation. The use of this system together with pyridinecarbothioamides-like prodrugs allows speculating whether the activation occurs through an inter- or intra-molecular ET process. Unfortunately, the reaction of ethionamide with pentacyanoferrate(II) moiety leads to the product complex coordinated through the thioamide group [24], which is the target for the prodrug ET investigation. To overcome this inconvenience, we have taken the thionicotinamide ligand which does not present steric hindrance around the nitrogen of the pyridine ring to carry out the experiments. Thionicotinamide shows a reducing property toward the superoxide ion, producing the 3-cyanopyridine (3-CNpy) corresponding nitrile with consequent formation of hydrogen peroxide and sulfur derivative species [25]. Despite the importance of isoniazid and ethionamide ligands as prodrugs for the treatment and prophylaxis of tuberculosis, little is known about their mechanisms of action on a molecular level [6,26,27]. The identities of these activated prodrugs are unknown, but details regarding the activation mechanisms are beginning to emerge [28]. A comprehension of the activation mechanisms could result on the design of improved versions of the prodrugs, providing an alternative method for the treatment of the multi-drug resistant (MDR) strains of *M. tuberculosis*. An understanding of the mechanism of activation may lead to the design of improved versions of the thioamide-containing prodrugs, providing an alternative method for the treatment of the MDR.

2. Experimental

2.1. Materials

All manipulations were carried out under inert atmosphere (N_2 or Ar) following conventional techniques. Milli-Q grade or doubly distilled water was used throughout. The thionicotinamide and 3-cyanopyridine ligands, from Aldrich, were used without further purification. The $Na_3[Fe(CN)_5(NH_3)] \cdot 3H_2O$ and $Na_2[Fe(CN)_5(NH_3)] \cdot H_2O$ complexes were synthesized as previously described [29,30] with minor modifications. Anal. Calc. for $Na_3[Fe(CN)_5(NH_3)] \cdot 3H_2O$: C, 18.42; H, 2.76; N, 25.77. Found: C, 18.38; H, 2.73; N, 25.62. Yield was better than 90%. Anal. Calc. for $Na_2[Fe(CN)_5(NH_3)] \cdot H_2O$: C, 22.47; H, 1.89; N, 31.47. Found: C, 22.56; H, 1.81; N, 31.22. Yield was better than 70%. All other chemicals were analytical reagent grade.

2.1.1. $Na_3[Fe(CN)_5(thio)] \cdot 3H_2O$

A 100 mg sample of $Na_3[Fe(CN)_5(NH_3)] \cdot 3H_2O$ was dissolved in 5 mL of a 30% dimethylformamide/water mixture, under argon flow. The thionicotinamide ligand (90 mg) dissolved in 2 mL of the same solvent mixture was added to the complex solution. Upon mixing, the resulting solution developed an orange color, and was allowed to react for 2 h, under argon, stirring, and in the absence of light. The reaction vessel was then cooled on ice bath, and a cold NaI saturated solution in ethanol

was added dropwise until complete precipitation. The red solid formed was collected by filtration, washed with ethanol and diethyl ether, dried and stored under vacuum in the absence of light. Anal. Calc. C, 29.50; H, 2.70; N, 21.90; S, 7.10%. Found: C, 29.41; H, 2.68; N, 22.00; S, 7.11. 1H NMR: $\delta(D_2O$, 300 MHz) 9.10, 9.17 (SH–NH), 9.08 (H-2), 9.02 (H-6), 7.95 (H-4) and 7.23 (H-5). Yield was better than 85%.

2.1.2. $Na_3[Fe(CN)_5(3\text{-}CNpy)]\cdot 3H_2O$

On the basis of the physical chemical properties, further shown on this work, the main product of the reaction of $Na_2[Fe^{III}(CN)_5(NH_3)]\cdot H_2O$ complex with thionicotinamide (reaction **1**) is the $Na_3[Fe(CN)_5(3\text{-}CNpy)]\cdot 3H_2O$ species. This complex was obtained by dissolving 100 mg of the $Na_2[Fe^{III}(CN)_5(NH_3)]\cdot H_2O$ complex in 5 mL of a 30% dimethylformamide/water mixture, under argon flow. The thionicotinamide ligand (100 mg) dissolved in 3 mL of the same solvent mixture was added to the complex solution. The reaction was allowed to stand for 2 h under argon flow, and stirring, in the absence of light. A dark red solid precipitates by adding a cold NaI/ethanol saturated solution. The precipitate was collected by filtration, washed with ethanol and diethyl ether, dried, and stored under vacuum in the absence of light. Anal. Calc. C, 32.01; H, 2.42; N, 23.73; S, 0.00. Found: C, 31.97 ; H, 2.45; N, 23.80; S, 0.00%. 1H NMR: $\delta(D_2O$, 300 MHz), 9.27 (H-2), 9.19 (H-6), 8.05 (H-4) and 7.33 (H-5). Yield was better than 80%.

For comparative purposes, the $Na_3[Fe(CN)_5(3\text{-}CNpy)]\cdot 3H_2O$ complex was also prepared following the literature method [31] by the direct reaction of $Na_3[Fe^{II}(CN)_5(NH_3)]\cdot 3H_2O$ with the 3-cyanopyridine ligand. Anal. Calc. C, 32.01; H, 2.42; N, 23.73. Found: C, 32.04; H, 2.40; N, 23.82. 1H NMR: $\delta(D_2O$, 300 MHz) 9.27 (H-2), 9.19 (H-6), 8.05 (H-4) and 7.33 (H-5).

2.2. Apparatus

The electronic absorption spectra were acquired with a Hewlett–Packard 8453 diode-array scanning spectrophotometer in the 1100–200 nm range. A quartz cell (1.0 cm) was used, keeping concentrations of samples between 1.0×10^{-3} and 1.0×10^{-2} M. Infrared spectra were taken on a Shimadzu FTIR-8300 spectrophotometer, using solid samples dispersed in KBr pellets. Electrochemical experiments were performed on a BAS 100BW electrochemical analyzer (Bioanalytical Systems-BAS, West Lafayette, IN). A conventional three-electrode configuration glass cell with a glassy carbon of 0.126 cm^2 geometrical area, and a Pt foil were used as working and auxiliary electrodes, respectively. A 0.1 M NaTFA solution (pH 4.0) was used as the electrolyte, at 25 ± 0.2 °C. Potential values were measured versus Ag/AgCl (3.0 M KCl) as reference electrode. The Mössbauer spectra were obtained with a conventional constant acceleration spectrometer operating in a triangular wave mode with a ^{57}Co source in a rhodium matrix. All spectra were taken with the source and absorber at room temperature, and sodium nitroprusside was taken as reference.

2.3. Kinetic measurements

The outer-sphere ET reaction rates were determined using a Hewlett–Packard 8453 spectrophotometer, at 25 ± 0.1 °C. Solutions of the reactants were freshly prepared under an argon atmosphere for each run. The buffer solutions consisting of H_2KPO_4 and NaOH were used. The ionic strength of the solutions (0.5 M) was adjusted by the addition of KCl. The pH measurements were performed on a Corning 440 pH meter equipped with a Cole Parmer glass electrode. Values of k_{obs} were calculated from a least-squares fitting to the equation $\ln(A_t - A_\infty) = \ln(A_0 - A_\infty) - k_{obs}t$, where A_t, A_0, and A_∞ are absorbance at time t, time zero (first point), and longer times ($\geqslant$4 half lives), respectively. The calculations were carried out by means of the kinetic software of the Hewlett–Packard 8453 spectrophotometer. The reaction progress was followed at 420 nm, the absorbance maximum of $[Fe(CN)_6]^{3-}$, under pseudo-first order conditions. Kinetic measurements for the intra-molecular reaction were performed on an Aminco-Morrow 126550-3 stopped-flow spectrophotometer.

3. Results and discussion

The reaction of thionicotinamide ligand with the $[Fe^{II}(CN)_5(H_2O)]^{3-}$ complex generates the $[Fe^{II}(CN)_5(thio)]^{3-}$ species, preserving the chemical composition of the thionicotinamide moiety [32]. However, the oxidized analogous species, $[Fe^{III}(CN)_5(thio)]^{2-}$, could not be isolated from the direct reaction of $[Fe^{III}(CN)_5(H_2O)]^{2-}$ ion with the same ligand. The main product of this reaction was characterized as the $[Fe^{II}(CN)_5(3\text{-}CNpy)]^{3-}$ complex, 3-CNpy = 3-cyanopyridine.

$$[Fe^{III}(CN)_5(H_2O)]^{2-} + C_6N_2SH_6 \rightarrow [Fe^{II}(CN)_5(C_6N_2H_4)]^{3-} + H_2S + H_2O \quad (1)$$

The electronic spectrum of the $[Fe^{II}(CN)_5(thio)]^{3-}$ complex displayed three bands. The shoulder near λ_{max} = 230 nm ($\varepsilon \cong 10^5$ $M^{-1}\,cm^{-1}$) is typically assigned to the $p\pi^*CN^- \leftarrow d\pi(Fe^{II})$ metal to ligand charge transfer (MLCT) transition. The broad band at λ_{max} = 290 nm ($\varepsilon = 5.5\times10^4$ $M^{-1}\,cm^{-1}$) is similar in energy and intensity to those observed for the free ligand and therefore was assigned to internal ligand transitions [33], whereas the solvent sensitive band at λ_{max} = 390 nm ($\varepsilon = 1.4\times10^3$ $M^{-1}\,cm^{-1}$) was attributed to the $p\pi^*(thio) \leftarrow d\pi(Fe^{II})$ MLCT transition.

The most prominent feature of the electronic spectrum for the product of reaction (**1**) of $Na_2[Fe^{III}(CN)_5(NH_3)] \cdot H_2O$ complex with thioamide ligand is a solvent sensitive band at 414 nm ($\varepsilon = 3.2 \times 10^3$ $M^{-1}\,cm^{-1}$), that is not usually observed in this spectral range for pentacyanoferrate(III) complexes with pyridine derivative ligands [34]. The energy of this band lies in the same position as that observed for the $[Fe^{II}(CN)_5(3\text{-}CNpy)]^{3-}$ complex [31], which is typically observed for the [pπ*(pyridine ring) ← dπ(Fe^{II})] MLCT transition, indicating the presence of Fe^{II} in the composition of this complex [35].

The infrared (IR) spectrum of the $Na_3[Fe^{II}(CN)_5(thio)] \cdot 3H_2O$ complex displayed several characteristic bands of the coordinated ligand in the imidothiol form [νSH, 2575 cm^{-1}; (δNH + νC=N + δCH), 1301 cm^{-1}; (νC=S + νCN), 1134 cm^{-1}; νCS, 738 cm^{-1}] [36]. The bands at 2052 and 572 cm^{-1} are assigned to the νC≡N and νFe–CN vibrational modes, respectively [37].

The IR spectrum for the product of reaction (**1**) showed bands due to the νCN (2051 cm^{-1}) and νFe–CN (573 cm^{-1}) modes at very similar frequencies as those observed for pentacyanoferrate(II) complexes [38]. An increase in the νCN stretching frequency of about 70 cm^{-1} is expected to occur for similar Fe(III) complexes [38]. Furthermore, the characteristic bands of the thionicotinamide ligand were not observed and an additional typical band of the nitrile stretching mode (νNC) appears at 2240 cm^{-1} [35].

The cyclic voltammetry for the $[Fe^{II}(CN)_5(thio)]^{3-}$ complex showed waves ($E_{1/2}$ = 240 mV, at 50 mV s^{-1}) characteristic of an irreversible oxidation process. The E_p values showed a small dependence on the scan rate and the plot of i_p versus $v^{1/2}$ was not linear over the experimental range between 10 and 150 mV s^{-1}. This is consistent with the anticipated ET reaction of thionicotinamide ligand coordinated to the Fe^{III} metal center.

Upon iron(II) metal center oxidation, the Fe^{III} electrochemically formed species induces the intra-molecular ET reaction to generate the $[Fe^{II}(CN)_5(3\text{-}CNpy)]^{3-}$ complex. The cyclic voltammetry for the product of the reaction (**1**) is identical to that for the $[Fe^{II}(CN)_5(3\text{-}CNpy)]^{3-}$ complex synthesized by the direct reaction of $Na_3[Fe^{II}(CN)_5(NH_3)] \cdot 3H_2O$ with 3-CNpy ligand, which exhibited one reversible $Fe^{III/II}$ process with $E_{1/2}$ = 283 mV.

The 1H NMR spectrum of the thioamide ligand exhibits four multiplets [δ 9.00 (H-2), 8.20 (H-4), 7.51 (H-5), 8.66 (H-6)] in an ABXY system, and two additional single signals [δ 9.87, δ 10.08] attributed to the hydrogen atoms of the imidothiol group (Chart 1) [39].

The 1H NMR spectrum of the $[Fe^{II}(CN)_5(thio)]^{3-}$ complex shows two single signals at δ 9.10 and δ 9.17 typical of the hydrogen atoms of the imidothiol group, indicating that the preferred imidothiol tautomer [39] of the ligand is preserved upon coordination. In the interpretation of the $[Fe^{II}(CN)_5(thio)]^{3-}$ spectrum, the effects of the cyanide ligands must also be taken into account. The anisotropy of the cyanides due to the perpendicular induced field at the C≡N bonds could deshield the H-2 and H-6 hydrogen atoms of the coordinated pyridine ring. This phenomenon is more effective for the H-2 and H-6 hydrogen atoms due to the unfavorable distance between the cyanide ligands and the H-4 and H-5 hydrogen atoms of the coordinated pyridine ring. On the basis of the preceding arguments the downfield shifted peaks centered at δ 9.08 and 9.02 were assigned to the H-2 and H-6 hydrogens. All other hydrogen signals have shown upfield shifts [δ 7.95 (H-4), 7.23 (H-5)]. Since the cyanide anisotropy is not expected to have a noticeable influence over the H-4 and H-5 hydrogens, the observed chemical shift variation ($\Delta\delta = 0.25$ and $\Delta\delta = 0.28$) compared to the free ligand spectrum could be substantially due to the [pπ*thio) ← dπ(Fe^{II})] back-bonding effect [40].

The 1H NMR spectrum for the product of reaction (**1**) presents one single signal at δ 9.27, two doublets centered at δ 9.19 and δ 8.05, and one multiplet centered at δ 7.35. The spectrum does not show the NMR signals typical of the thioamide ligand. Furthermore, the chemical shift values and the sharp shapes of the signals are not characteristic of NMR spectra of the Fe^{III} paramagnetic species. For example, the literature reports that the hydrogen spectrum of $[Fe^{III}(CN)_5imz]^{2-}$, imz = imidazole, exhibits broad signals at δ −19.1, 5.10 and 0.90, while for the Fe^{II} analogous complex the corresponding hydrogen signals were observed at δ 7.74, 7.20 and 7.00 [41,42].

Both 1H NMR spectra for the $[Fe(CN)_5(3\text{-}CNpy)]^{3-}$ complex, synthesized by the direct reaction of $[Fe(CN)_5(NH_3)]^{3-}$ with 3-CNpy ligand, and for the product of reaction (**1**) are identical and were attributed to the hydrogen atoms of the 3-CNpy coordinated ligand [δ 9.27 (H-2), 9.19 (H-6), 8.50 (H-4), 7.35 (H-5)].

Although these results strongly support that the main product of the reaction (**1**) is the $[Fe^{II}(CN)_5(3\text{-}CNpy)]^{3-}$ complex, Mössbauer spectroscopy was used to better investigate the electronic nature of the metal center.

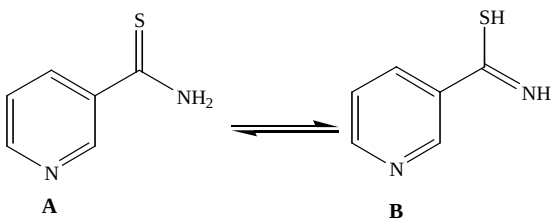


Chart 1. Structure representation of the thioamide (**A**) and imidothiol (**B**) tautomers.

The Mössbauer spectra of pentacyanoferrate complexes are doublets due to the presence of an electronic field gradient with non-cubic symmetry in the ^{57}Fe nucleus [43]. The non-cubic electronic configuration of the central ion comes from the asymmetric σ and π bonding involving only the axial ligand L, reflecting on the isomer shift (δ) Mössbauer parameter. The ligand σ donation mechanism increases the s electron population causing an increase of the electronic density at the metal [44]. The same effect is caused by the $p\pi^*(L) \rightarrow d\pi(M)$ back-donation that decreases the shielding effects. The σ and π bonding involving the metal and L axial ligand are responsible for the appearance of the electric field gradient (EFG) at the ^{57}Fe nucleus. The quadrupole splitting is caused by the interaction of the nuclear quadrupole moment of the iron atom with the Z component of the EFG [44]. The $[Fe^{III}(CN)_5L]$ complexes have a d^5 low-spin configuration that decreases the shielding effect of the s electrons compared to those of the d^6 low-spin Fe^{II} species. Also, the low-spin d^5 configuration induces a larger EFG due to the asymmetric electronic configuration. Thus, a smaller isomer shift and a larger quadrupole splitting are expected for Fe^{III} than those for the Fe^{II} analogous pentacyanoferrate complexes. The Mössbauer spectra of $Na_3[Fe(CN)_5(thio)] \cdot 3H_2O$ and for the product of reaction (**1**) are shown in Fig. 1. The quadrupole splitting (Δ) and the isomer shift (δ) values along with the data for similar complexes are shown in Table 1.

The isomer shift and quadrupole splitting values for Fe^{III} and Fe^{II} complexes are quite distinguishable. The δ and Δ Mössbauer parameters values for both complexes reported in this work strongly suggest the presence of an unique ^{57}Fe(II) Mössbauer nucleus, consistent with the proposed chemical compositions [45]. These results, along with those reported here for the full characterization of the complexes, indicate that an ET reaction involving the $[Fe^{III}(CN)_5(H_2O)]^{2-}$ complex and the thioamide ligand occurs with the consequent formation of the $[Fe^{II}(CN)_5(3\text{-CNpy})]^{3-}$ complex.

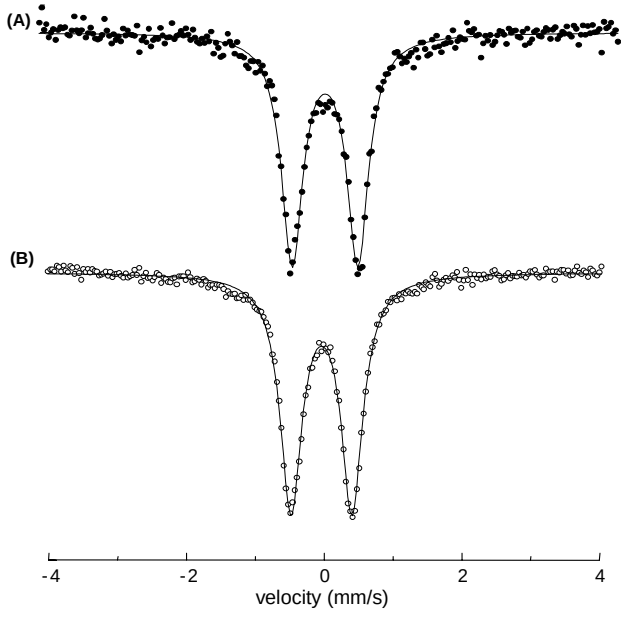


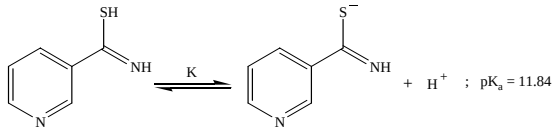
Fig. 1. Mössbauer spectra at 300 K of (**A**) $Na_3[Fe(CN)_5(thio)] \cdot 3H_2O$ and (**B**) $Na_3[Fe^{II}(CN)_5(3\text{-CNpy})] \cdot 3H_2O$ obtained in the reaction (**1**).

To provide experimental evidence on whether the ET reaction (**1**) occurs through an inter or intramolecular pathway, a kinetic study for the cross-reaction of thionicotinamide and $[Fe(CN)_6]^{3-}$ was first addressed. The oxidation of thionicotinamide by $[Fe(CN)_6]^{3-}$ leads to the formation of predominantly 3-cyanopyridine and $[Fe(CN)_6]^{4-}$ according to the following overall reaction (**2**):

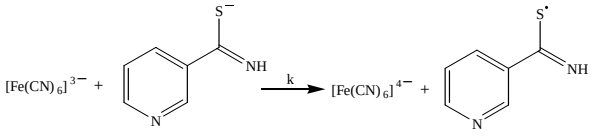
$$C_6N_2SH_6 + [Fe^{III}(CN)_6]^{3-} \rightarrow C_6N_2H_4 + [Fe^{II}(CN)_6]^{4-} + H_2S \quad (2)$$

The reaction was monitored by the absorbance decrease at 420 nm, the absorbance maximum of $[Fe(CN)_6]^{3-}$, under pseudo-first-order conditions. The rate of this reaction is strongly dependent on pH due to an acid–base equilibrium related to the deprotonation of -SH functional group ($pK_a = 11.84$) [33]. The pH value dependence observed for reaction (**2**), shown in Fig. 2, clearly demonstrates that the imidothiol tautomer is the main reactive species toward the ET processes involving the thionicotinamide ligand. A significant decrease in the rate constant is observed for pH < 11, for which the unreactive thioamide form becomes favored [25,33] (Chart 1). The protonation of $[Fe(CN)_6]^{3-}$ does not represent an important factor in the pH range selected for this investigation [46,47].

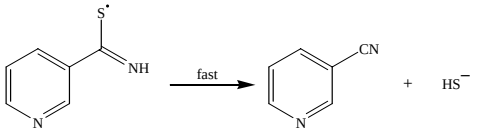
On the basis of the kinetic results we have proposed the following overall mechanism:



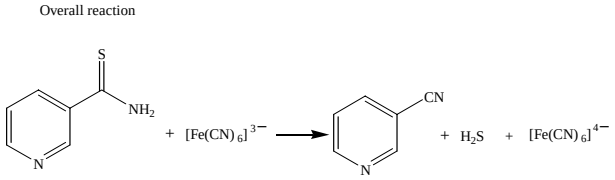
(3)



(4)



(5)



The rate law for this mechanism in (6) can, with the aid of Eq. (7), be simplified to (8) for the pH 12.0 experimental conditions.

$$-\frac{[\text{Fe(III)}]}{dt} = k[(\text{py})\text{-CS}^{-}\text{NH}][\text{Fe(III)}], \quad (6)$$

$$[(\text{py})\text{-CS}^{-}\text{NH}] = \frac{K_a[(\text{py})\text{-CSHNH}]}{[H^+]} \cong [(\text{py})\text{-CSHNH}], \quad (7)$$

$$k_{obs} = k[(\text{py})\text{-CSHNH}]. \quad (8)$$

Indeed, under this pH experimental condition, the first order rate constant k_{obs} depends linearly on [(py)-CSHNH] as shown in Fig. 3. The slope k_{obs}/[(py)-CSHNH] gives a value of $k = (5.38 \pm 0.03)$ M^{-1} s^{-1} at 25 °C, pH 12.0 (Phosphate buffer), μ = 0.50 M.

The ET reaction of $[Fe^{III}(CN)_5(H_2O)]^{2-}$ with the thionicotinamide ligand can be envisaged to occur by the inner- or outer-sphere pathways as illustrated in Scheme 1.

Table 1
Mössbauer parameter data, at 300 K, nitroprusside as reference

Complexes	$\delta \pm 10^{-4}$ (mm s^{-1})	$\Delta \pm 10^{-3}$ (mm s^{-1})
$Na_3[Fe^{II}(CN)_5(NH_3)] \cdot 3H_2O$	0.275	0.686
$Na_2[Fe^{III}(CN)_5(NH_3)] \cdot 3H_2O$	0.203	1.660
$Na_3[Fe^{II}(CN)_5(dtdp)] \cdot 4H_2O$[a]	0.291	0.886
$Na_2[Fe^{III}(CN)_5(dtdp)] \cdot 4H_2O$	0.220	1.709
$Na_3[Fe^{II}(CN)_5(thio)] \cdot 3H_2O$	0.212	0.912
$Na_3[Fe^{II}(CN)_5(3\text{-CNpy})] \cdot 3H_2O$[b]	0.219	0.887

[a] dtdp = 4,4′-dithiodipyridine [32].
[b] Generated from $[Fe^{III}(CN)_5(H_2O)]^{2-}$ reaction with thionicotinamide ligand.

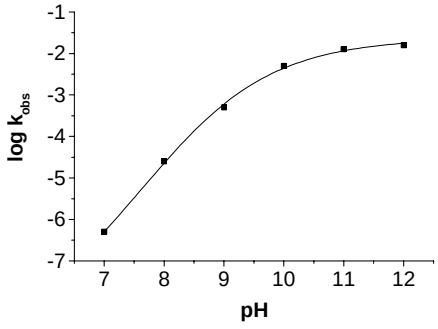


Fig. 2. pH dependence of k_{obs} for the oxidation of thionicotinamide by $[Fe(CN)_6]^{3-}$. $[Fe^{III}] = 1.0 \times 10^{-5}$ M; [thio] = 3.0×10^{-3} M; μ = 0.5 M; temperature = 25 °C.

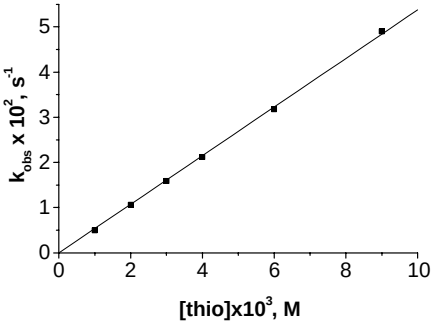
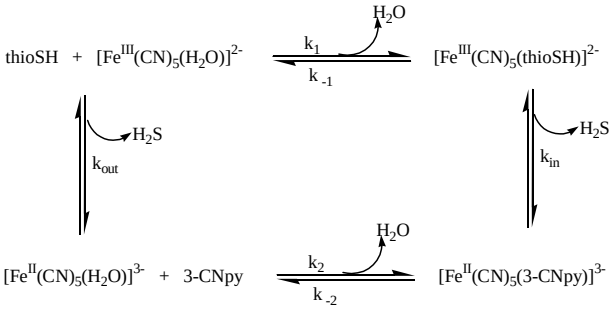


Fig. 3. Plot of k_{obs} versus [thio] for the oxidation of thionicotinamide by $[Fe(CN)_6]^{3-}$. [Fe(III) = 1.0×10^{-5} M; pH 12,0; μ = 0.5 M; temperature = 25 °C.

In an effort to qualitatively elucidate the mechanism of the ET reaction of $[Fe^{III}(CN)_5(H_2O)]^{2-}$ with the thionicotinamide ligand, kinetic studies were performed at 25 °C, pH 7.0. The product solution was monitored spectrophotometrically by following the absorbance increase at 414 nm, i.e., the absorbance maximum of $[Fe^{II}(CN)_5(3\text{-CNpy})]^{3-}$ species. Contrary to that observed for rate constant of the cross-reaction of $[Fe^{III}(CN)_6]^{3-}$ and thionicotinamide ligand at the same conditions ($k_{obs} = 5.0 \times 10^{-7}$ s^{-1}), a rapid ET process is observed with a rate constant $k_{et} = 10$ s^{-1}. The rate was observed to be concentration independent. An attempt to determine experimentally the rate (k_{out}) of intermolecular ET reaction was not successful, mainly due to the absorption interference of the MLCT band of $[Fe^{II}(CN)_5(3\text{-CNpy})]^{3-}$ species that forms at higher rates. However, by a simple comparison with the rate value observed for the thionicotinamide ligand ET reaction with $[Fe^{III}(CN)_6]^{3-}$ that operates with a larger driving force ($\Delta G \cong 1.60$ V), a value smaller than 5.0×10^{-7} s^{-1} would be expected for k_{out} [48]. Additionally, the literature reports a rate constant $k_2 = 413$ M^{-1} s^{-1} for the formation of the $[Fe^{II}(CN)_5(3\text{-CNpy})]^{3-}$ species [31], also suggesting that the intermolecular pathway indicated in Scheme 1 is not preferred. These observations, together with the fact that the rate constant observed $k_{et} = 10$ s^{-1} is concentration independent, strongly suggest that coordination to pentacyanoferrate(III) com-



Scheme 1.

plex promotes the thioamide ligand activation through an intramolecular ET reaction.

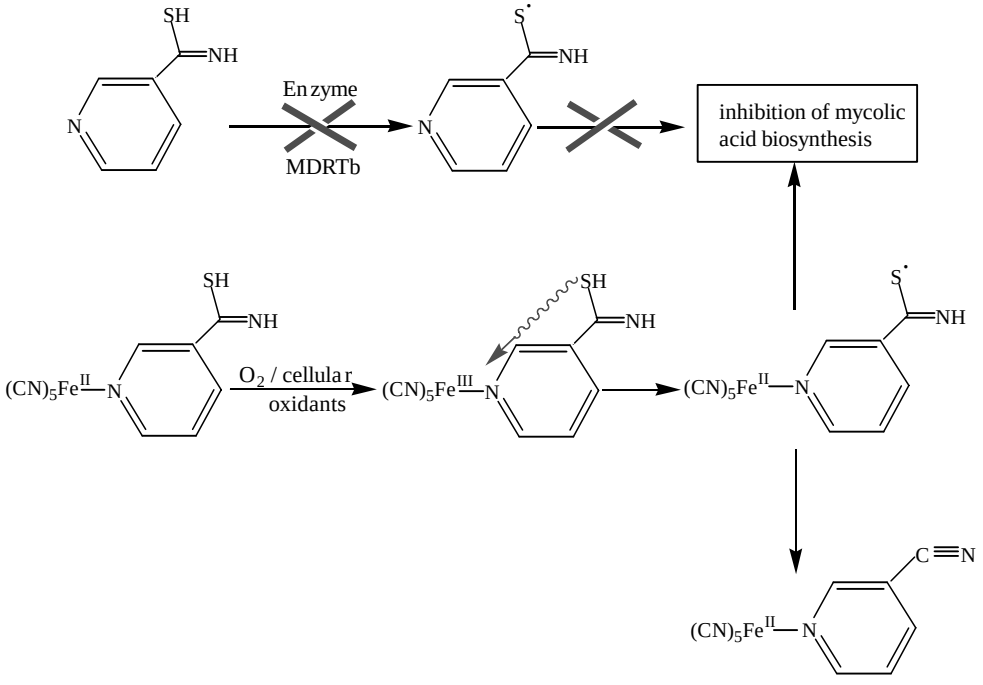
4. Conclusions

The data collected here, along with those reported in the literature [25] concerning the redox reaction of the thionicotinamide with superoxide ion lead us to propose the reaction patterns outlined in the Scheme 1 through an inner sphere pathway.

The in vitro activation of thionicotinamide and isoniazid through an intramolecular ET reaction in $[Fe^{III}(CN)_5(L)]$ complexes provide a suitable model system for the *M. tuberculosis* Etha (or EtaA) and KatG enzyme-dependent activation mechanisms [14]. Although the thionicotinamide in the tuberculosis therapy is not as effective as isoniazid and rifampicin, resistance to these front-line drugs has led to the use of less-active thioamides for the treatment of multi-drug-resistant tuberculosis [14,19,49]. Based on the mechanisms of activation propose for ethionamide and isoniazid, both via ET reaction [21], and on the results presented here, an alternative self-activation route can be proposed for designing new drugs for the treatment of multi-drug-resistant tuberculosis. These drugs would be activated by ET reactions before interacting with its cellular target. The oxidative activation of INH and ETH (or ETA) by cells of *M. tuberculosis* is mediated by the KatG and EthA (or EtaA) enzymes, respectively, that precedes inactivation of enzymes involved in mycolic acid biosynthesis that are constituents of the cell wall of the bacillus [21]. Most of the INH resistance is associated with *katG* structural gene alterations resulting in catalase-peroxidase mutant enzymes with impaired ability to form activated-INH intermediates. One approach to overcome this effect is the use of low oxidation state metal complexes coordinated to the prodrugs. For this purpose, we have found that the pentacyanoferrate(II) metal center and thionicotinamide ligand is a very first approximate model system on the perspective of a new approach for drug design for the treatment of multi-drug-resistant tuberculosis by proposing a self-activation mechanism illustrated in Scheme 2.

Upon oxidation of $[(CN)_5Fe^{II}(thio)]^{3-}$ complex ($E_{1/2}$ = 240 mV) by oxygen and/or cellular oxidants, it is postulated that the thermodynamically unstable iron(III) complex would undergo a rapid intramolecular ET reaction. The $[(py\text{-}CS^{\cdot}NH)Fe^{II}(CN)_5]^{3-}$ intermediate species decays to the $[(3\text{-}CNpy)Fe^{II}(CN)_5]^{3-}$ complex. This proposed self-activation mechanism is well supported by the literature that reports the electrochemical oxidation of thioamides yielding the corresponding nitrile and polysulfide species [25]. The same oxidation is reported to occur by the action of molecular oxygen, but the process is much slower. Other studies have detected nitriles as oxidation products of thioamides [50].

As reported recently, the incubation of both wild-type (WT) and isoniazid-resistant I21V mutant enoyl reductases from *M. tuberculosis* with $Na_3[Fe^{II}(CN)_5(IN-$



Scheme 2.

H)] · $4H_2O$ in the absence of NADH resulted in the time-dependent inactivation of the enzyme with the apparent first-order rate constant values of 327×10^{-3} and 315×10^{-3} min^{-1}, respectively [12]. The results suggest that the WT and I21V mutant inhA enzimes inactivation by $[Fe^{II}(CN)_5(INH)]^{3-}$ requires no activation by KatG, no need for the presence of NADH, and its mechanism of action probably involves interaction with the NADH binding site of the enzyme.

Pharmacological studies as well as similar experiments with isoniazid and ethionamide ligands including other metal complexes as ruthenium ammines are currently underway in our laboratories. The results will be soon reported in a forthcoming paper.

Acknowledgements

The authors acknowledge CNPq and CAPES, the Brazilian financial agencies. I.S.M. (CNPq, 477610/2003-0) L.A.B. (CNPq, 520182/99-5) and D.S.S. (CNPq, 304051/1975-06) are researcher awardees from the National Council for Scientific and Technological Development of Brazil.

References

[1] A. Quémard, J.C. Sacchettini, A. Dessen, C. Vilchèze, R. Bittman, W.R. Jacobs Jr., J.S. Blanchard, Biochemistry 34 (1995) 8235–8241.
[2] C. Vilchèze, H.R. Morbidoni, T.R. Weisbrod, H. Iwamoto, M. Kuo, J.C. Sacchettini, W.R. Jacobs Jr., J. Bacteriol. 182 (2000) 4059–4067.
[3] J.M. Musser, A. Amin, S. Ramaswamy, J. Infect. Dis. 180 (1999) 1751–1752.
[4] L.A. Basso, R. Zheng, J.M. Musser, W.R. Jacobs Jr., J.S. Blanchard, J. Infect. Dis. 178 (1998) 769–775.
[5] K. Johnsson, P.G. Schultz, J. Am. Chem. Soc. 116 (1994) 7425–7426.
[6] K. Johnsson, D.S. King, P.G. Schultz, J. Am. Chem. Soc. 117 (1995) 5009–5010.
[7] L.A. Basso, R. Zheng, J.S. Blanchard, J. Am. Chem. Soc. 118 (1996) 11301–11302.
[8] R.F. Zabinski, J.S. Blanchard, J. Am. Chem. Soc. 119 (1997) 2331–2332.
[9] D.A. Rozwarsky, G.A. Grant, D.H.R. Barton, W.R. Jacobs, J.C. Sacchettini, Science 279 (1998) 98–102.
[10] B. Lei, C.-J. Wei, S.-C. Tu, J. Biol. Chem. 275 (2000) 2520–2526.
[11] L.A. Basso, J.S. Blanchard, Adv. Exp. Med. Biol. 456 (1998) 115–144.
[12] J. Oliveira, E.H.S. Sousa, L.A. Basso, M. Palaci, R. Dietze, D.S. Santos, I.S. Moreira, Chem. Commun. (2004) 312–313.
[13] B. Heym, P.M. Alzari, N. Honore, S.T. Cole, Mol. Microbiol. 15 (1995) 235–245.
[14] P. Farmer, J. Bayona, M. Becerra, J. Furin, C. Henry, H. Hiatt, J.Y. Kim, C. Mitinick, E. Nardell, S. Shin, Int. J. Tuberc. Lung Dis. 2 (1998) 869–876.
[15] K. Viskum, A. Kok-Jensen, Int. J. Tuberc. Lung Dis. 1 (1997) 299–301.
[16] J.B. Bass Jr., L.S. Farer, P.C. Hopewell, R. O'Brien, R.F. Jacobs, F. Rubem, D.E. Snider Jr., G. Thornton, Am. J. Resp. Crit. Care Med. 149 (1994) 1359–1374.
[17] WHO. TB: Groups at Risk, WHO report on Tuberculosis Epidemic, World Health Organization, Geneva, 1996.
[18] World Health Organization, Anti-tuberculosis drug resistance in the world: Third Global Report, 2004.
[19] J. Crofton, O. Chaulet, D. Maher, J. Grosset, W. Harris, H. Norman, M. Iseman, B. Watt, Guidelines for the Management of Multi-drug-Resistant Tuberculosis, World Health Organization, Geneva, 1997.
[20] T.A. Vannelli, A. Dykman, O. Monttelano, J. Biol. Chem. 277 (2002) 12824–12829.
[21] A.E. DeBarber, K. Mdluli, M. Bosman, L.G. Bekker, C.E. Barry 3rd, Proc. Natl. Acad. Sci. USA 97 (2000) 9677–9682.
[22] R.S. Magliozzo, J.A. Marcinkeviciene, J. Am. Chem. Soc. 118 (1996) 11303–11304.
[23] H. Taube, Comment. Inorg. Chem. 1 (1981) 17–39, and references therein.
[24] E.H.S. de Sousa, Ph.D. Thesis, Universidade Federal do Ceará, 2003.
[25] O.A. Paez, C.M. Valdez, J. Tor, D.T. Sawyer, J. Org. Chem. 53 (1988) 2166–2170.
[26] A.R. Baulard, J.C. Betts, J. Engohang-Ndong, S. Quan, P.J. Brennan, C. Locht, G.S. Besra, J. Biol. Chem. 275 (2000) 28326–28331.
[27] E.S. Raper, Coord. Chem. Rev. 129 (1994) 91–156.
[28] N.L. Wengenack, H. Lopes, M.J. Kennedy, P. Tavares, A.S. Pereira, I. Moura, J.J.G. Moura, F. Rusnak, Biochemistry 39 (2000) 11508–11513.
[29] G. BrauerHandbook of Preparative Inorganic Chemistry, second ed., vol. 2, Academic Press, New York, 1965.
[30] J.H. Esperson, S.G. Wolenuk Jr., Inorg. Chem. 11 (1972) 2034–2041.
[31] A.P. Szecsy, S.S. Miller, A. Haim, Inorg. Chim. Acta 28 (1978) 189–200.
[32] I.S. Moreira, D.W. Franco, Inorg. Chem. 33 (1994) 1607–1613, and references therein.
[33] C. Tisser, M. Tisser, Bull. Soc. Chim. Fr. 10 (1970) 3752–3757.
[34] D.H. MacCartney, Rev. Inorg. Chem. 29 (1988) 101–134, and references therein.
[35] A.E. Almaraz, L.A. Gentil, L.M. Beraldo, J.A. Olabe, Inorg. Chem. 35 (1996) 7718–7727.
[36] V. Chauhan, S.K. Dikshit, Trans. Met. Chem. 13 (1988) 440–442.
[37] L. Tosi, Spectrochim. Acta A 29 (1973) 353–363.
[38] K. Nakamoto, Infrared and Raman Spectra of Inorganic Compounds, third ed., Wiley, New York, 1978.
[39] V.W. Brügel, Zeitschirift für Elektrochemie 66 (1962) 159–165.
[40] D.K. Lavalee, E.B. Fleischer, J. Am. Chem. Soc. 94 (1972) 2583–2599.
[41] F.J. Wu, D.M. Kurtz Jr., J. Am. Chem. Soc. 111 (1989) 6563–6572.
[42] R.E. Shepherd, J. Am. Chem. Soc. 98 (1976) 3329–3335.
[43] A.L. Coelho, I.S. Moreira, J.H. Araújo, M.A.B. Araújo, J. Radioanal. Nucl. Chem. 136 (1989) 299–309.
[44] D.P.E. Dickson, F.J. Berry, Mossbauer Spectroscopy, Cambridge University Press, London, 1986.
[45] I.S. Moreira, J.B. Lima, D.W. Franco, Coord. Chem. Rev. 196 (2000) 197–217.
[46] B. Bansch, P. Martinez, J. Zuluaga, D. Uribe, R. van Eldik, Z. Phys. Chem. 170 (1991) 59–71.
[47] J. Jordan, G.J. Ewing, Inorg. Chem. 1 (1962) 587–591.
[48] N. Sutin, Ann. Ver. Nucl. Sci. 12 (1962) 285–330.
[49] E.K. Schroeder, O.N. Souza, D.S. Santos, J.S. Blanchard, L.A. Basso, Curr. Pharm. Biotechnol. 3 (2002) 197–225.
[50] W. Walter, J. Voss, in: J. Zabicky (Ed.), The Chemistry of Thioamides, Interscience, New York, 1970, pp. 436–449.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Protein Expression and Purification 40 (2005) 23–30
Protein Expression & Purification
www.elsevier.com/locate/yprep

DAHP synthase from *Mycobacterium tuberculosis* H37Rv: cloning, expression, and purification of functional enzyme

Caroline Rizzi[a], Jeverson Frazzon[b], Fernanda Ely[a], Patrícia G. Weber[a], Isabel O. da Fonseca[a], Michelle Gallas[a], Jaim S. Oliveira[a], Maria A. Mendes[c], Bibiana M. de Souza[c], Mário S. Palma[c], Diógenes S. Santos[d,*], Luiz A. Basso[a,*]

[a] *Departamento de Biologia Molecular e Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS 91501-970, Brazil*
[b] *Departamento de Ciência dos Alimentos, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS 91501-970, Brazil*
[c] *Departamento de Biologia/CEIS, Universidade do Estado de São Paulo, Rio Claro, SP 13506-900, Brazil*
[d] *Centro de Pesquisa e Desenvolvimento em Biologia Molecular e Funcional, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS 90619-900, Brazil*

Received 13 April 2004, and in revised form 18 June 2004
Available online 8 December 2004

Abstract

Tuberculosis (TB), caused by *Mycobacterium tuberculosis*, remains the leading cause of mortality due to a bacterial pathogen. According to the 2004 Global TB Control Report of the World Health Organization, there are 300,000 new cases per year of multidrug resistant strains (MDR-TB), defined as resistant to isoniazid and rifampicin, and 79% of MDR-TB cases are now "super strains," resistant to at least three of the four main drugs used to treat TB. Thus there is a need for the development of effective new agents to treat TB. The shikimate pathway is an attractive target for the development of antimycobacterial agents because it has been shown to be essential for the viability of *M. tuberculosis*, but absent from mammals. The *M. tuberculosis aroG*-encoded 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (mtDAHPS) catalyzes the first committed step in this pathway. Here we describe the PCR amplification, cloning, and sequencing of *aroG* structural gene from *M. tuberculosis* H37Rv. The expression of recombinant mtDAHPS protein in the soluble form was obtained in *Escherichia coli* Rosetta-gami (DE3) host cells without IPTG induction. An approximately threefold purification protocol yielded homogeneous enzyme with a specific activity value of 0.47 U mg^{-1} under the experimental conditions used. Gel filtration chromatography results demonstrate that recombinant mtDAHPS is a pentamer in solution. The availability of homogeneous mtDAHPS will allow structural and kinetics studies to be performed aiming at antitubercular agents development.

Keywords: *Mycobacterium tuberculosis*; Shikimate pathway; DAHP synthase; Protein expression

Tuberculosis (TB)[1] 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

* Corresponding authors. Fax: +55 51 3166234.
E-mail addresses: diogenes@pucrs.br (D.S. Santos), labasso@dna.cbiot.ufrgs.br (L.A. Basso).

[1] *Abbreviations used:* TB, tuberculosis; MDR-TB, multidrug-resistant; PEP, phosphoenolpyruvate; E4P, D-erythrose-4-phosphate; DAHPS, 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase; DMSO, dimethyl sulfoxide; LB, Luria–Bertani; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; IPTG, isopropyl β-D-thiogalactoside; β-ME, β-mercaptoethanol; DAHPS(Phe), phenylalanine-regulated DAHPS; DAHPS(Try), tyrosine-regulated DAHPS; DAHPS(Trp), tryptophan-regulated DAHPS.

1046-5928/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2004.06.040

from TB high-burden countries [1]. 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 [2]. Approximately, 2 billion individuals are believed to harbor latent TB based on tuberculin skin test surveys [3], which represents a considerable reservoir of bacilli. Possible factors underlying the resurgence of TB worldwide include the HIV epidemic, increase in the homeless population, and decline in health care structures and national surveillance [4]. Another contributing factor is the evolution of multi-drug resistant strains (MDR-TB), defined as resistant to isoniazid and rifampicin, which are the most effective first-line drugs [5]. 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 [6]. 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 [7]. Hence, faster acting and effective new drugs to better combat TB, including MDR-TB, are needed.

The shikimate pathway is an attractive target for the development of herbicides and antimicrobial agents because it is essential in algae, higher plants, bacteria, and fungi, but absent from mammals [8]. In mycobacteria, the shikimate pathway leads to the biosynthesis of chorismic acid, which is a precursor for the synthesis of aromatic amino acids, naphthoquinones, menaquinones, and mycobactins [9]. The salicylate-derived mycobactin siderophores have been shown to be essential for *M. tuberculosis* growth in macrophages [10]. More recently, the shikimate pathway has been shown by disruption of *aroK* gene, which codes for the shikimate kinase enzyme, to be essential for the viability of *M. tuberculosis* [11]. The absence from the human host and essentiality of mycobacterial shikimate pathway indicate that any of its enzymes are promising targets for the development of potentially non-toxic antimycobacterial agents.

Homologues to enzymes in the shikimate pathway have been identified in the genome sequence of *M. tuberculosis* [12]. The first committed step in the shikimate pathway is catalyzed by 3-deoxy-D-*arabino*-heptulosonate 7-phosphate (DAHP) synthase (DAHPS; EC 4.1.2.15). DAHPS catalyzes the stereospecific condensation of phosphoenolpyruvate (PEP) and D-erythrose 4-phosphate (E4P), forming DAHP and inorganic phosphate [13]. Based on phylogenetic analysis, DAHPS has been divided into two classes, class I and class II [14]. *Escherichia coli* expresses three DAHPS isoenzymes that are representative of class II and require divalent metal for activity [15], which play a role in catalysis and/or structural integrity [16]. Each isoenzyme is specifically inhibited by one of the three aromatic amino acids [8]. DAHPS(Phe), a homotetramer encoded by the *aroG* gene, is feedback inhibited by phenylalanine; the *aroH*-encoded DAHPS(Trp) and *aroF*-encoded DAHPS(Tyr) are homodimers feedback inhibited by, respectively, tryptophan and tyrosine. In *M. tuberculosis* genome, however, only the *aroG* (Rv2178c) encoded DAHPS isoenzyme (mtDAHPS) has been proposed to be present by sequence homology.

To determine the mechanism of action of mtDAHPS by steady-state and pre-steady-state kinetics as well as for X-ray crystal structure determination aiming at the rational design of antimycobacterial agents, expression of *aroG* encoded mtDAHPS in functional form and in large quantity are needed. Accordingly, we here describe the PCR amplification, cloning, sequencing, expression in *E. coli* Rosetta-gami (DE3) cells, purification to homogeneity, oligomeric state determination, and assay of mtDAHP enzyme activity. Measurements of enzyme activity confirm the correct assignment to the structural gene encoding mtDAHPS in *M. tuberculosis*. The availability of mtDAHPS will allow enzyme kinetics and structural studies to be undertaken to provide a framework on which to base the design of new agents with antitubercular activity with, hopefully, low toxicity.

Materials and methods

PCR amplification and cloning of M. tuberculosis aroG gene

The design of synthetic oligonucleotide primers used for PCR amplification of *aroG* gene (5′-gga**catatg**aactgg accgtcgacatac-3′ and 5′-c**ggatcc**tcagtcccgcagcatctccgc-3′) was based on the complete genome sequence of *M. tuberculosis* H37Rv [12]. These primers were complementary to, respectively, the amino-terminal coding and carboxy-terminal noncoding strands of *aroG* gene containing 5′*Nde*I and 3′*Bam*HI restrictions sites, which are in bold. This pair of primers was used to amplify the *M. tuberculosis aroG* gene (1389 bp) from genomic DNA using standard PCR conditions and the enzyme *Pfu* DNA polymerase (Stratagene), which is a thermostable polymerase that exhibits low error rate, thus lowering the likelihood of introducing unwanted mutations. PCR amplification required the presence of 10% of dimethyl sulfoxide (DMSO) in the reaction mixture. The PCR product was purified by electrophoresis on low melting agarose, digested with *Nde*I and *Bam*HI (Boehringer–Mannheim), and cloned into pET-23a(+) (Novagen)

expression vector, which had previously been digested with the same restriction enzymes. To both confirm the identity of the cloned gene and ensure that no mutations were introduced by the PCR amplification step, the DNA sequence of the amplified *M. tuberculosis aroG* structural gene was determined by dideoxy-chain termination method [17], using the Thermo Sequenase radiolabeled terminator cycle sequencing kit (Amersham Biosciences).

Expression of mtDAHPS

The recombinant plasmid pET-23a(+)::*aroG* was transformed into electrocompetent *E. coli* Rosetta-gami (DE3) cells (Novagen), and selected on LB agar plates containing 50 μg mL^{-1} carbenicillin, 15 μg mL^{-1} kanamycin, 34 μg mL^{-1} chloramphenicol, and 12.5 μg mL^{-1} tetracycline. Single colonies were used to inoculate 500 mL LB medium, containing the same antibiotics and concentrations of LB solid medium, and grown at 37 °C and 180 rpm for 24 h, without addition of isopropyl β-D-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation at 48,000*g* for 20 min at 4 °C, and stored at −20 °C. For protein expression analysis, 10 mg of stored cells was resuspended in 500 μL Buffer A (50 mM Tris–HCl, pH 7.8), disrupted by sonification, and cell debris was removed by centrifugation. Both soluble and insoluble fractions were analyzed by SDS–PAGE 12% [18]. Control experiments were performed under the same experimental conditions except that *E. coli* host cells were transformed with the expression vector lacking the target gene.

Purification of recombinant mtDAHPS

Approximately, 36 g of cells was collected by centrifugation (48,000*g* for 20 min) from 6 L of LB medium. All subsequent steps were performed on ice or at 4 °C. Frozen cells (36 g) were thawed and resuspended in Buffer A (4 mL of buffer per gram of cell paste) containing 1 mM β-mercaptoethanol (β-ME) (Sigma) and 0.2 mg mL^{-1} of lysozyme, and the mixture was stirred for 30 min. Cells were disrupted by sonication, and cell debris was removed by centrifugation (48,000*g* for 30 min). The supernatant was incubated with 1% w/v of streptomycin sulfate for 15 min, and centrifuged (48,000*g* for 30 min). Solid ammonium sulfate was added to the supernatant fraction to a concentration of 25% saturation, incubated for 30 min, and centrifuged as above. The resultant pellet was resuspended in 70 mL Buffer A containing 1 mM β-ME and dialyzed against three changes of 2 L of the same buffer using a dialysis tubing with molecular weight cut-off of 12,000–4000 Da. The sample was clarified by centrifugation and loaded on a Q-Sepharose Fast Flow (2.6 cm × 8.2 cm) anion exchange column (Amersham Biosciences) previously equilibrated with Buffer A and fractionated using a 600 mL 0.0–0.6 M NaCl linear gradient. The fractions containing mtDAHPS (0.35–0.38 M NaCl) were pooled, concentrated to 8.0 mL using an Amicon ultrafiltration cell (MW 30,000 Da), and loaded on a Sephacryl S-200 HR (2.6 cm × 60 cm) gel filtration column (Amersham Biosciences) at 0.5 mL min^{-1}. The protein was eluted with Buffer B (Buffer A containing 200 mM NaCl) at the same flowrate. The active fractions were loaded on a Mono Q HR 16/10 anion exchange column (Amersham Biosciences) equilibrated with Buffer A and eluted with 400 mL linear 0.0–0.6 M NaCl gradient. The active fractions, which exhibited a single band on SDS–PAGE, were pooled, quickly frozen in liquid nitrogen, and stored at −80 °C.

Determination of protein concentration

Protein concentrations were determined using the Bio-Rad Laboratories protein assay kit (Bradford method) [19] and bovine serum albumin as standard.

Determination of mtDAHPS molecular mass

The molecular mass of native mtDAHPS homogeneous protein was determined by gel filtration chromatography using a Sephacryl S-200 (HR 10/30) (Amersham Biosciences) equilibrated with Buffer B at a flowrate of 0.4 mL min^{-1}. Protein molecular mass standards were from Gel Filtration LMW and HMW Calibration Kit from Amersham Biosciences. Protein elution was monitored at 280 nm.

mtDAHPS enzyme assay

Enzyme activity of recombinant mtDAPHS protein was assayed in the forward direction by a continuous spectrophotometric method described by Shoner and Hermann [20], monitoring the decrease in phosphoenolpyruvate (PEP) concentration at 232 nm ($\varepsilon = 2.8 \times 10^3$ M^{-1} cm^{-1}) on a Multi-Spec 1501 photodiode array spectrophotometer (Shimadzu). All reactions were carried out at 25 °C and initiated with addition of enzyme to a reaction mixture containing: 50 mM Tris–HCl, pH 7.0, 400 μM E4P (Sigma), 1 mM β-ME, and 200 μM PEP (Acrós Organics) in a total volume of 500 μL. One unit of enzyme activity (U) is defined as the amount of enzyme catalyzing the conversion of 1 μmol PEP/min at 25 °C.

N-terminal amino acid sequencing

The N-terminal amino acid residues of homogeneous recombinant mtDAHPS were identified by automated Edman degradation sequencing using a PPSQ 21A gas-phase sequencer (Shimadzu).

Mass spectrometry analysis

The homogeneity of recombinant protein preparation was assessed by mass spectrometry (MS), employing some adaptations made to the system described by Chassaigne and Lobinski [21]. Samples were analyzed on a triple quadrupole mass spectrometer, model QUATTRO II, equipped with standard electrospray (ESI) probe (Micromass, Altrinchan), adjusted to ca. 250 μL min^{-1}. The source temperature (80 °C) and needle voltage (3.6 kV) were maintained constant throughout the experimental data collection, applying a drying gas flow (nitrogen) of 200 Lh^{-1} and a nebulizer gas flow of 20 Lh^{-1}. The mass spectrometer was calibrated with intact horse heart myoglobin and its typical cone-voltage induced fragments. The subunit molecular mass of recombinant protein mtDAHPS was determined by ESI-MS, adjusting the mass spectrometer to give a peak with at half-height of 1 mass unit, and the cone sample to skimmer lens voltage controlling the ion transfer to mass analyzer was set to 38 V. About 50 pmol (10 μL) of each sample was injected into the electrospray transport solvent. The ESI spectrum was obtained in the multi-channel acquisition mode, scanning from 500 to 1800 *m/z* at a scan time of 7 s. The mass spectrometer is equipped with MassLynx and Transform software for data acquisition and spectra handling.

Results and discussion

The PCR amplification of *aroG* structural gene from *M. tuberculosis* H37Rv genomic DNA required the presence of 10% DMSO in the reaction mixture (data not shown). DMSO is a cosolvent that improves GC-rich DNA denaturation and helps to overcome the difficulties of polymerase extension through secondary structures, altering the structural conformation of DNA templates [22]. This result is consistent with the 65.6% G + C content of *M. tuberculosis* H37Rv genome [12]. PCR fragment was inserted into pET23a(+) expression vector [23] between *Nde*I and *Bam*HI restriction sites. DNA sequencing of the entire *aroG* structural gene by the dideoxy chain termination method both confirmed the identify of the cloned PCR product and showed that no mutations were introduced by the DNA amplification step.

Recombinant plasmids were introduced into *E. coli* BL21(DE3) host cells by electroporation. Unfortunately, recombinant mtDAHPS remained in the insoluble fraction (Fig. 1). Since one of the goals of the present work was to confirm the correct assignment to the structural gene encoding mtDAHPS, efforts were made to express recombinant *M. tuberculosis* DAHPS in its soluble, active form avoiding unfolding and refolding protocols because they cannot guarantee that they will yield large

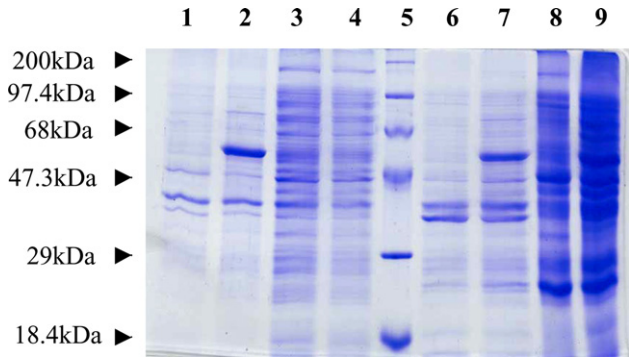


Fig. 1. SDS–PAGE (12%) of the soluble and insoluble fractions of the cell extracts of either BL21(DE3) or Rosetta-gami (DE3) host cells transformed with either pET-23a(+) (control) or pET-23a(+)::*aroG*. Expression conditions were 24 h at 37 °C without IPTG addition. Lane 1: insoluble fraction of BL21 (DE3) transformed with pET-23(+); lane 2: insoluble fraction of BL21(DE3) transformed with pET-23(+)::*aroG*; lane 3: soluble fraction of BL21(DE3) transformed with pET-23(+); lane 4: soluble fraction of BL21(DE3) transformed with pET-23(+)::*aroG*; lane 5: MW marker "High range" (Gibco-BRL); lane 6: insoluble fraction of Rosetta-gami (DE3) transformed with pET-23(+); lane 7: insoluble fraction of Rosetta-gami (DE3) transformed with pET-23(+)::*aroG*; lane 8: soluble fraction of Rosetta-gami (DE3) transformed with pET-23(+); and lane 9: soluble fraction of Rosetta-gami (DE3) transformed with pET-23(+)::*aroG*. Molecular mass of mtDAHPS is approximately 50.6 kDa.

amounts of biologically active product [24]. In addition, a number of protocols were tested to obtain mtDAHPS in the soluble fraction to no avail, including buffer additives (urea, deoxycholic acid, Triton X-100, and high NaCl concentrations) and reduced cultivation temperature (20, 25, and 30 °C). In practice, it is usually worthwhile to test several different vector/host combinations to obtain the best possible yield of protein in its desired form. Accordingly, a number of commercially available strains of *E. coli* host cells were tested in an attempt to produce mtDAHPS in the soluble fraction. Analysis of the relationship between codon preference and expression level led to the classification of *E. coli* genes into three main classes [25]. Class II genes, which correspond to genes highly and continuously expressed during exponential growth that is likely to resemble the tRNA population available for recombinant protein expression, have a number of avoided codons with frequencies of less than 6%. Insufficient tRNA pools can lead to premature translational termination, translation frameshifting or amino acid misincorporation that might result in expression of nonproperly folded recombinant protein [26]. Rare codons near the N-terminus of a coding sequence can have a severe effect on heterologous expression in *E. coli* [27]. Four rare codons for heterologous gene expression in *E. coli* are present near the N-terminus of *M. tuberculosis aroG* structural gene (1 × AUA for isoleucine, 3 × CCC for proline). To test whether these rare codons may have any effect on *aroG* expression, *E. coli* Rosetta (DE3) strain harboring tRNA genes for AGG, AGA, AUA, CUA, CCC, and

GGA rare codons on a chloramphenicol-resistant plasmid [28] was transformed with pET-23a(+)::*aroG* recombinant plasmid. Disappointingly, recombinant mtDAHPS remained in the insoluble fraction (data not shown), thereby discarding any effect of the mycobacterial *aroG* rare codons on recombinant protein expression. Although *E. coli* DAHPS has no disulfide bridges despite possessing three cysteine residues, there is no experimental evidence for the absence of disulfide bridges in mtDAHPS, which possesses five cysteine residues (Cys 87, Cys 231, Cys 365, Cys 420, and Cys 440), and a less reducing cytoplasmatic environment could improve mtDAHPS solubility. The Origami *E. coli* host strains (Novagen) have mutations in both the thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes, which greatly enhances disulfide bond formation in the cytoplasm [29,30]. Unfortunately, none of the protocols tested yielded soluble mtDAHPS. The Rosetta-gami (DE3) *E. coli* host strain (Novagen) combines the features of Rosetta and Origami strains. SDS–PAGE analysis showed that expression of recombinant mtDAHPS protein in its soluble form with the expected molecular mass (~51 kDa) could be achieved using the Rosetta-gami (DE3) cells grown at 37 °C for 24 h with no IPTG induction (Fig. 1). The underlying reason for this result is unclear; however, it underscores the need for optimization of vector/host combinations to achieve soluble recombinant protein expression before attempting any unfolding/refolding protocols.

It should be pointed out that a screening of experimental conditions was carried out to obtain high yield of recombinant protein expression, including temperature of growth, culture aeration, medium type, hours of growth after IPTG induction, and hours of growth in the absence of IPTG. The best results were obtained from Rosetta-gami (DE3) *E. coli* cells grown for 24 h at 37 °C in LB medium without IPTG induction as described above. In the pET vector system (Novagen), target genes are positioned downstream of the bacteriophage T7 late promoter. Typically, production hosts contain a prophage (λDE3) encoding the highly processive T7 RNA polymerase under control of the IPTG-inducible *lac*UV5 promoter that would ensure tight control of recombinant gene basal expression [31,32]. In agreement with the results presented here, leaky expression has been shown to occur in the pET system [33–37]. It has been proposed that leaky protein expression is a property of the lac-controlled system as cells approach stationary phase in complex medium and that cyclic AMP, acetate, and low pH are required to achieve high-level expression in the absence of IPTG induction, which may be part of a general cellular response to nutrient limitation [38].

Enzyme activity measurements demonstrated that there was a 92-fold increase in specific activity for mtDAHPS when Rosetta-gami (DE3) *E. coli* harboring either pET-23a(+)::*aroG* or pET-23a(+) crude extracts were compared (Table 1), indicating that mtDAHPS was expressed in its soluble and functional form. The purification protocol of recombinant mtDAHPS, protein determination, enzyme assay, and SDS–PAGE analysis were as described in Materials and methods. Recombinant mtDAHPS enzyme was purified approximately 3-fold (Table 2) to electrophoretic homogeneity (Fig. 2). The purification protocol yielded approximately 5 mg of homogeneous protein. A significant loss in protein yield occurred in the 25% ammonium sulfate precipitation step (Table 2) because some mtDAHPS remained in the supernatant. Protein precipitations by higher ammonium sulfate concentrations were also carried out, yielding larger amounts of recombinant mtDAHPS in the pellet (data not shown). However, a number of contaminants co-precipitated with mtDAHPS. In particular, a contaminant that co-eluted in subsequent chromatographic steps when larger than 25% ammonium sulfate concentrations were used. Accordingly, the 25% ammonium sulfate precipitation step was deemed more appropriate for the purification protocol because a significant amount of contaminants remained in the supernatant, while mtDAHPS with a lower protein-contaminating background remained in the pellet thus making the

Table 1
Measurements of recombinant DAHPS enzyme activity

Cell extract[a]	Specific activity[b] (SA, U mg^{-1})	AS cloned/SA control
Control	0.0018	1
DAHPS	0.1651	92

[a] Crude cell extract in 50 mM Tris–HCl, pH 7.0.
[b] U mL^{-1}/mg mL^{-1}.

Table 2
Purification of *M. tuberculosis* 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase expressed in *E. coli* Rosetta-gami(DE3) transformed with pET-23a(+)::*aroG*[a]

Purification step	Total protein (mg)	Total enzyme activity (U)	Specific activity[b] (U mg^{-1})	Purification fold	Yield (%)
Crude extract	2442.38	403.23	0.17	1.00	100
Ammonium sulfate	136.78	44.11	0.32	1.95	11
Q-Sepharose Fast Flow	21.42	7.00	0.33	1.98	2
Sephacryl S-200 HR	17.83	4.48	0.25	1.52	1
Mono Q HR 16/10	4.77	2.24	0.47	2.84	0.6

[a] Typical purification protocol starting from 36 g wet weight cells obtained from 6 L of culture.
[b] U mL^{-1}/mg mL^{-1}.

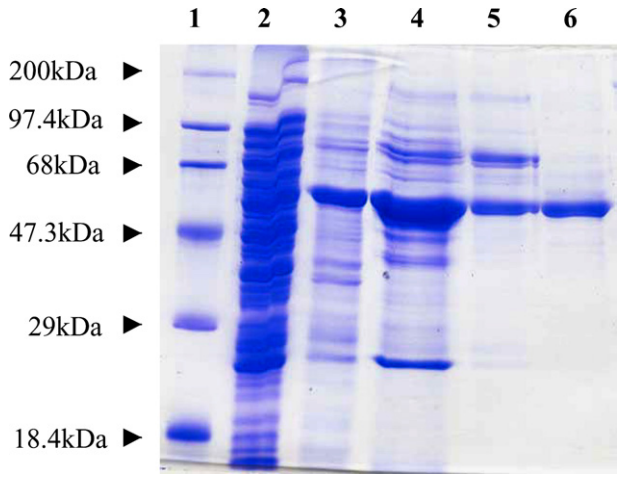


Fig. 2. SDS–PAGE analysis of pooled fractions from the purification steps of mtDAHPS. Lane 1, MW marker "High range" (Gibco-BRL); lane 2, crude extract; lane 3, ammonium sulfate precipitation; lane 4, Q-Sepharose Fast Flow ion exchange; lane 5, S-200 gel filtration; and lane 6, Mono Q ion exchange.

subsequent purification steps less demanding. The samples of each purification step were assayed for DAHPS enzyme activity and compared to control experiments to demonstrate that the observed values are actual velocities of mtDAHPS activity. Enzyme activity of homogeneous mtDAHPS was linearly dependent on sample volume added to the reaction mixture (Fig. 3), thereby showing that the initial velocity is proportional to total enzyme concentration and that true initial velocities are being measured. *E. coli* DAHPS(Phe) has been shown to be sensitive to oxidation, leading to inactivation [39]. Moreover, higher mtDAHPS enzyme activity values could be observed in the presence of β-ME (data not shown). Accordingly, β-ME was used in all steps of the mtDAHPS purification protocol. The mtDAHPS

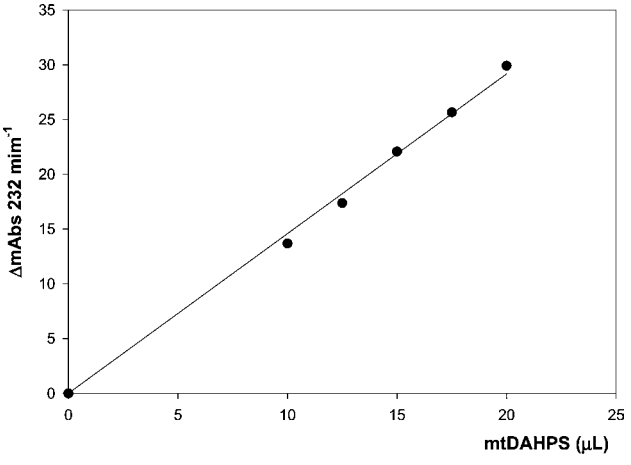


Fig. 3. Linear dependence of mtDAHPS activity on homogeneous protein volume. The rates of enzyme activity were performed in the forward direction by continuously monitoring the decrease of phosphoenolpyruvic acid at 232 nm.

specific activity has been found to be stable for at least two months when stored at −80 °C.

The subunit molecular mass of active mtDAHPS was determined to be 50510.38 Da by electrospray ionization mass spectrometry (ESI-MS), consistent with the post-translational removal of the N-terminal methionine residue from the full length gene product (predicted mass: 50641.51 Da). The ESI-MS result also revealed no peak at the expected mass for the three isoforms of *E. coli* DAHPS enzymes (38804.03, 38735.18, and 38009.53 Da), thus providing evidence for both the identity and purity of the recombinant protein. The first 11 N-terminal amino acid residues of the recombinant protein were identified as NWTVDIPIDQL by the Edman degradation chemistry protocol. This result unambiguously identifies the homogeneous recombinant protein as mtDAHPS and confirms removal of the N-terminal methionine residue from it. A common type of co-/post-translational modification of proteins synthesized in prokaryotic cells is modification at their N-termini. Methionine aminopeptidase catalyzed cleavage of initiator methionine is usually directed by the penultimate amino acid residues with the smallest side chain radii of gyration (glycine, alanine, serine, threonine, proline, valine, and cysteine) [40]. The N-terminal methionine was removed from the *E. coli* expressed recombinant mtDAHPS enzyme, consistent with the finding that some middle-sized penultimate amino acid residues (Asn, Asp, Leu, and Ile) undergo N-terminal processing [41].

A value of 253 ± 25 kDa was determined for the molecular mass of native mtDAHPS homogeneous protein by analytical gel filtration chromatography (data not shown), suggesting that mtDAHPS is a pentamer in solution. Whereas *E. coli* DAHPS(Phe) is a tetramer [42] and *E. coli* DAHPS(Trp) is a dimer [43]. Interestingly, more recently, recombinant DAHPS from *Pyrococcus furiosus* has been shown to be a dimer in solution and not to be inhibited by phenylalanine, tyrosine, or tryptophan [44].

DAHPS in most, not all, microorganisms is the target for pathway regulation by negative feedback inhibition, which controls carbon flow into the shikimate pathway. The most intensively investigated microorganism DAHPS has been the *E. coli* enzyme, which possesses three isoenzymes, each specifically regulated by one of three aromatic amino acid end products, either Phe, Tyr, or Trp [45]. The three isoforms have a common requirement for a metal cofactor, which can be similarly satisfied by a range of divalent metal ions [46]. DAHPS enzymes from a number of microorganisms have been studied, such as *Corynebacterium glutamicum* [47], *Thermotoga maritima* [48], *Bacillus subtilis* [49], *Saccharomyces cerevisiae* [50], and *P. furiosus* [44]. However, to the best of our knowledge, this is the first report on cloning, expression, and purification of functional DAHPS from *M. tuberculosis*.

Homogeneous mtDAHPS protein will provide protein in quantities necessary for studies on the enzyme mechanism of action by steady-state and pre-steady-state kinetics, its metal requirement, if any, and feedback inhibition by Phe, Tyr, and Trp. The expression and purification of mtDAHPS reported here will also provide protein for crystallization trials aiming at three-dimensional structure determination by X-ray diffraction. The three-dimensional structures of four forms of the Phe-regulated isoenzyme of *E. coli* DAHPS have been solved by X-ray crystallography [42,51–53], which should facilitate screening of experimental conditions to obtain crystals of mtDAHPS in complex with its substrates, possible metal cofactor and feedback inhibitors, if any. Expression of functional proteins in soluble form has been identified as an important bottleneck in efforts to determine biological activity and crystal structure of *M. tuberculosis* proteins [54]. We hope that the results reported here will contribute to efforts towards the structure determination of potential targets in *M. tuberculosis*. The enzymological and structural studies on mtDAHPS should help in the design of enzyme inhibitors to be tested as antimycobacterial agents.

Acknowledgments

Financial support for this work was provided by Millennium Initiative Program MCT-CNPq, Ministry of Health-Department of Science and Technology-UNESCO (Brazil) to D.S.S. and L.A.B. D.S.S. and L.A.B. also acknowledge grants awarded by PADCT, CNPq, and FINEP. L.A.B. (CNPq, 520182/99-5), D.S.S. (CNPq, 304051/1975-06), and M.S.P. (CNPq, 300337/2003-50) are researchers awardees from the National Council for Scientific and Technological Development of Brazil. M.A.M. is a post-doctoral fellow from FAPESP (01/05060-4).

References

[1] M.C. Raviglione, The TB epidemic from 1992 to 2002, Tuberculosis 83 (2003) 4–14.
[2] E.L Corbett., C.J. Watt, N. Walker, D. Maher, B.G. Williams, M.C. Raviglione, C. Dye, The growing burden of tuberculosis: global trends and interactions with the HIV epidemic, Arch. Intern. Med. 1639 (2003) 1009–1021.
[3] C. Dye, S. Scheele, P. Dolin, V. Pathania, M.C. Raviglione, Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country, JAMA 282 (1999) 677–686.
[4] B.R. Bloom, C.J.L. Murray, Tuberculosis: commentary on a reemergent killer, Science 257 (1998) 1055–1064.
[5] L.A. Basso, J.S. Blanchard, Resistance to antitubercular drugs, Adv. Exp. Med. Biol. 456 (1998) 115–144.
[6] World Health Organization, Anti-tuberculosis drug resistance in the world: Third Global Report, 2004.
[7] K. Duncan, Progress in TB drug development and what is still needed, Tuberculosis 83 (2003) 201–207.
[8] R. Bentley, The shikimate pathway—a metabolic tree with many branches, Crit. Rev. Biochem. Mol. Biol. 25 (1990) 307–384.
[9] C. Ratledge, Nutrition, growth and metabolism, in: C. Ratledge, J.L. Stanford (Eds.), The Biology of the Mycobacteria, vol. 1, Academic Press, London, 1982, pp. 185–271.
[10] J.J. de Voos, K. Rutter, B.G. Schroder, H. Su, Y. Zhu, C.E. Barry III, The salicylate-derived mycobactin siderophores of *Mycobacterium tuberculosis* are essential for growth in macrophages, Proc. Natl. Acad. Sci. USA 97 (2000) 1252–1257.
[11] T. Parish, N.G. Stoker, The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*, Microbiology 148 (2002) 3069–3077.
[12] S.T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S.V. Gordon, K. Eiglmeier, S. Gas, C.E. Barry III, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B.G. Barrell, Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence, Nature 393 (1998) 537–544.
[13] P.R. Srinivasan, D.B. Sprinson, 2-Keto-3-deoxy-D-*arabo*-heptonic acid 7-phosphate synthase, J. Biol. Chem. 234 (1959) 716–722.
[14] M.R. Birck, R.W. Woodard, *Aquifex aeolicus* 3-deoxy-D-*manno*-2-octulosonic acid 8-phosphate synthase: a new class of KDO 8-P synthase?, J. Mol. Evol. 52 (2001) 205–214.
[15] C.M. Stephens, R. Bauerle, Analysis of the metal requirement of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*, J. Biol. Chem. 266 (1991) 20810–20817.
[16] P.A. Jordan, S. Bohle, C. Ramilo, J.N.S. Evans, New insights into the metal center of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase, Biochemistry 40 (2001) 8387–8396.
[17] F. Sanger, S. Nicklen, R. Coulson, DNA sequencing with chain-terminating inhibitors, Pro. Natl. Acad. Sci. USA 74 (1977) 5436–5467.
[18] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680–685.
[19] M.M. Bradford, R.A. McRorie, W.L. Williams, A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem. 72 (1976) 248–254.
[20] R. Shoner, K.M. Hermann, 3-Deoxy-D-*arabino*-heptulosonate 7-phosphate synthase, J. Biol. Chem. 252 (1976) 5440–5447.
[21] H. Chassaigne, R. Lobinski, Characterization of horse kidney metallothionein isoforms by electrospray MS and reversed-phase HPLC-electrospray MS, Analyst 123 (1998) 2125–2130.
[22] D. Pomp, J.F. Medrano, Organic solvents as facilitators of polymerase chain reaction, Biotechniques 10 (1991) 58–59.
[23] F.W. Studier, A.H. Rosenberg, J.J. Dunn, J.W. Dubendorff, Use of T7 RNA polymerase to direct expression of cloned genes, Methods Enzymol. 185 (1990) 60–89.
[24] C.H. Schein, Production of soluble recombinant proteins in bacteria, Bio/technology 7 (1989) 1141–1149.
[25] A. Hénaut, A. Danchin, Analysis and predictions from *Escherichia coli* sequences, or *E. coli* in silico, in: F.C. Neidhardt (Ed.), *Eschechichia coli* and *Salmonella*: Cellular and Molecular Biology, ASM Press, Washington, DC, 1996, pp. 2047–2066.
[26] C. Kurland, J. Gallant, Error of heterologous protein expression, Curr. Opin. Biotechnol. 7 (1996) 489–493.
[27] Y. Nakamura, T. Gojobori, T. Ikemura, Codon usage tabulated from international DNA sequence databases: status for the year 2000, Nucleic Acids Res. 28 (2000) 292.
[28] R. Novy, D. Drott, K. Yaeger, R. Mierendorf, Overcoming the codons bias of *Escherichia coli* for enhanced protein expression, Innovations 12 (2001) 1–3.
[29] W.A. Prinz, F. Aslund, A. Holmgren, J. Beckwith, The role of the thioredoxin and glutaredoxin pathways in redox-facilitate attempts at crystallizing and elucidating the protein disulfide bonds in the *Escherichia coli* cytoplasm, J. Biol. Chem. 272 (1997) 15661–15667.

[30] F. Baneyx, Recombinant protein expression in *Escherichia coli*, Curr. Opin. Biotechnol. 10 (1999) 411–421.
[31] F.W. Studier, B.A. Moffatt, Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, J. Mol. Biol. 189 (1986) 113–130.
[32] J.W. Dubendorff, F.W. Studier, Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor, J. Mol. Biol. 219 (1991) 45–59.
[33] L.A. Basso, D.S. Santos, W. Shi, R.H. Furneaux, P.C. Tyler, V.L. Schramm, J.S. Blanchard, Purine nucleoside phosphorylase from *Mycobacterium tuberculosis*. Analysis of inhibition by a transition-state analogue and dissection by parts, Biochemistry 40 (2001) 8196–8203.
[34] J.S. Oliveira, C.A. Pinto, L.A. Basso, D.S. Santos, Cloning and overexpression in soluble form of functional shikimate kinase and 5-enoylpyruvylshikimate 3-phosphate synthase enzymes from *Mycobacterium tuberculosis*, Protein Expr. Purif. 3 (2001) 430–435.
[35] M.L.B. Magalhães, C.P. Pereira, L.A. Basso, D.S. Santos, Cloning and expression of functional shikimate dehydrogenase (EC 1.1.1.25) from *Mycobacterium tuberculosis* H37Rv, Protein Expr. Purif. 26 (2002) 59–64.
[36] R.G. Silva, L.P.S. Carvalho, J.S. Oliveira, C.A. Pinto, M.A. Mendes, M.S. Palma, L.A. Basso, D.S. Santos, Cloning, overexpression and purification of functional human purine nucleoside phosphorylase, Protein Expr. Purif. 27 (2003) 158–164.
[37] K.C. Kelley, K.J. Huestis, D.A. Austen, C.T. Sanderson, M.A. Donoghue, S.K. Stikel, E.S. Kawasaki, M.S. Osburne, Regulation of *sCD4-183* gene expression from phage T7 based vectors in *Escherichia coli*, Gene 156 (1995) 33–36.
[38] T.H. Grossman, E.S. Kawaski, S.R. Punreddy, M.S. Osburne, Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability, Gene 209 (1998) 95–103.
[39] R. Bauerle, O.K. Park, Metal-catalyzed oxidation of phenylalanine-sentitive 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*: inactivation and destabilization by oxidation of active-site cysteines, J. Bacteriol. 181 (1999) 1636–1642.
[40] W.T. Lowther, B.W. Matthews, Structure and function of the methionine aminopeptidases, Biochim. Biophys. Acta 1477 (2000) 157–167.
[41] P.-H. Hirel, J.-M. Schmitter, P. Dessen, G. Fayat, S. Blanquet, Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid, Proc. Natl. Acad. Sci. USA 86 (1989) 8247–8251.
[42] I.A. Shumilin, R.H. Kretsinger, R. Bauerle, Crystal structure of phenylalanine-regulated 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*, Struct. Folding Des. 7 (1999) 865–875.
[43] J.M. Ray, R. Bauerle, Purification and properties of tryptophan-sentitive 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*, J. Bacteriol. 173 (1991) 1894–1901.
[44] L.R. Schofield, M.L. Patchett, E.J. Parker, Expression, purification, and characterization of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase from *Pyrococcus furiosus*, Protein Expr. Purif. 34 (2004) 17–27.
[45] K.M. Herrmann, The shikimate pathway as an entry to aromatic secondary metabolism, Plant Physiol. 107 (1995) 7–12.
[46] C.M. Stephens, R. Bauerle, Analysis of the metal requirement of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*, J. Biol. Chem. 266 (1991) 20810–20817.
[47] H.F. Liao, L.L. Lin, H.R. Chin, W.H. Hsu, Serine 187 is a crucial residue for allosteric regulation of *Corynebacterium glutamicum* 3-deoxy-D-*arabino*-heptulosonate-7-phosphate, FEMS Microbiol. Lett. 184 (2001) 59–64.
[48] J. Wu, D.L. Howe, R.W. Woodard, *Thermotoga maritima* 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase: the ancestral eubacterial DAHP synthase?, J. Biol. Chem. 278 (2003) 27525–27531.
[49] R.A. Jensen, E.W. Nester, Regulatory enzymes of aromatic amino acid biosynthesis in *Bacillus subtilis*. I. Purification and properties of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthetase, J. Biol. Chem. 241 (1966) 3365–3372.
[50] G. Paravicini, T. Schmidheini, G. Braus, Purification and properties of 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase (phenylalanine-inhibitable) of *Saccharomyces cerevisiae*, Eur. J. Biochem. 186 (1989) 361–366.
[51] T. Wagner, I.A. Shumilin, R. Bauerle, R.H. Kretsinger, Structure of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase from *Escherichia coli*: comparison of the Mn^{2+} 2-phosphoglycolate and the Pb^{2+} 2-phosphoenolpyruvate complexes and implications for catalysis, J. Mol. Biol. 301 (2000) 389–399.
[52] I.A. Shumilin, C. Zhao, R. Bauerle, R.H. Kretsinger, Allosteric inhibition of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase alters the coordination of both substrates, J. Mol. Biol. 320 (2002) 1147–1156.
[53] I.A. Shumilin, R. Bauerle, R.H. Kretsinger, The high-resolution structure of 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase reveals a twist in plane of bound phosphoenolpyruvate, Biochemistry 42 (2003) 3766–3776.
[54] R. Vicentelli, C. Bignon, A. Gruez, S. Canaan, G. Sulzenbacher, M. Tegoni, V. Campanacci, C. Cambillau, Medium-scale structural genomics: strategies for protein expression and crystallization, Acc. Chem. Res. 36 (2003) 165–172.