



## Structural and functional analyses of *Mycobacterium tuberculosis* Rv3315c-encoded metal-dependent homotetrameric cytidine deaminase

Zilpa A. Sánchez-Quitian<sup>a,b</sup>, Cristopher Z. Schneider<sup>a</sup>, Rodrigo G. Ducati<sup>a,c</sup>, Walter F. de Azevedo Junior<sup>b</sup>, Carlos Bloch Junior<sup>d</sup>, Luiz A. Basso<sup>a,\*</sup>, Diógenes S. Santos<sup>a,\*</sup>

<sup>a</sup>Centro de Pesquisas em Biologia Molecular e Funcional (CPBMF), Instituto Nacional de Ciência e Tecnologia em Tuberculose (INCT-TB), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Av. Ipiranga, 6681, Porto Alegre, RS 90619-900, Brazil

<sup>b</sup>Programa de Pós-Graduação em Biologia Celular e Molecular, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil

<sup>c</sup>Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

<sup>d</sup>Laboratório de Espectrometria de Massas, Empresa Brasileira de Pesquisa Agropecuária-Recursos Genéticos e Biotecnologia, Estação Parque Biológico, Brasília, DF, Brazil

### ARTICLE INFO

#### Article history:

Received 14 September 2009

Received in revised form 19 November 2009

Accepted 17 December 2009

Available online 24 December 2009

#### Keywords:

Crystal structure

*Mycobacterium tuberculosis*

Pyrimidine salvage pathway

Cytidine deaminase

Enzyme kinetics

Metalloenzyme

### ABSTRACT

The emergence of drug-resistant strains of *Mycobacterium tuberculosis*, the causative agent of tuberculosis, has exacerbated the treatment and control of this disease. Cytidine deaminase (CDA) is a pyrimidine salvage pathway enzyme that recycles cytidine and 2'-deoxycytidine for uridine and 2'-deoxyuridine synthesis, respectively. A probable *M. tuberculosis* CDA-coding sequence (*cdd*, Rv3315c) was cloned, sequenced, expressed in *Escherichia coli* BL21(DE3), and purified to homogeneity. Mass spectrometry, N-terminal amino acid sequencing, gel filtration chromatography, and metal analysis of *M. tuberculosis* CDA (MtCDA) were carried out. These results and multiple sequence alignment demonstrate that MtCDA is a homotetrameric Zn<sup>2+</sup>-dependent metalloenzyme. Steady-state kinetic measurements yielded the following parameters:  $K_m = 1004 \mu\text{M}$  and  $k_{\text{cat}} = 4.8 \text{ s}^{-1}$  for cytidine, and  $K_m = 1059 \mu\text{M}$  and  $k_{\text{cat}} = 3.5 \text{ s}^{-1}$  for 2'-deoxycytidine. The pH dependence of  $k_{\text{cat}}$  and  $k_{\text{cat}}/K_M$  for cytidine indicate that protonation of a single ionizable group with apparent  $\text{p}K_a$  value of 4.3 abolishes activity, and protonation of a group with  $\text{p}K_a$  value of 4.7 reduces binding. MtCDA was crystallized and crystal diffracted at 2.0 Å resolution. Analysis of the crystallographic structure indicated the presence of a Zn<sup>2+</sup> coordinated by three conserved cysteines and the structure exhibits the canonical cytidine deaminase fold.

© 2009 Elsevier Inc. All rights reserved.

### 1. Introduction

Human tuberculosis (TB)<sup>1</sup>, caused primarily by *Mycobacterium tuberculosis*; is an infectious disease that remains a major challenge to public health systems worldwide. Over 9 million people develop TB and 2 million die annually (Dye et al., 1999). Based on TB skin tests, the World Health Organization estimates that one-third of the world's population is currently infected with *M. tuberculosis* in a latent form and, therefore, at risk of developing active disease (WHO, 2006). The emergence of multidrug-resistant strains (MDR,

resistant to at least isoniazid and rifampicin) and extensively drug-resistant strains (XDR, resistant to first- and second-line anti-TB drugs) of *M. tuberculosis* have exacerbated the treatment and control of TB. Approximately 0.5 million cases of MDR-TB emerged in 2006, and new cases of XDR-TB have been notified in 45 countries in all five continents (Centers for Disease Control and Prevention (CDC), 2006; Jain and Mondal, 2008). *M. tuberculosis* has been considered the world's most successful pathogen and this is largely due to the ability of the bacillus to persist in host tissues, where drugs that are rapidly bactericidal *in vitro* require prolonged administration to achieve comparable *in vivo* effects (Hingley-Wilson et al., 2003). More effective and less toxic anti-TB drugs acting on novel mycobacterial targets are needed to shorten the duration of current therapy, to improve the treatment of MDR- and XDR-TB cases, and to provide an efficient alternative to eliminate latent *M. tuberculosis* infection.

Pyrimidine nucleotides are essential for DNA replication and RNA transcription, as well as many other cellular functions, such as phospholipid biosynthesis and regulation of enzyme activity by uridylylation. There are two major pathways for pyrimidine nucleotide synthesis: the *de novo* and the salvage pathway. While the *de novo* pathway demands higher energy levels for pyrimidine

\* Corresponding authors. Address: Centro de Pesquisas em Biologia Molecular e Funcional (CPBMF), Instituto Nacional de Ciência e Tecnologia em Tuberculose (INCT-TB), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Av. Ipiranga 6681, TecnoPUC, Prédio 92A, Porto Alegre, RS 90619-900, Brazil. Fax: +55 51 33203629.

E-mail addresses: [luiz.basso@pucrs.br](mailto:luiz.basso@pucrs.br) (L.A. Basso), [diogenes@pucrs.br](mailto:diogenes@pucrs.br) (D.S. Santos).

<sup>1</sup> Abbreviations used: CDA, cytidine deaminase; ESI-MS, electrospray ionization mass spectrometry; IPTG, isopropyl-β-D-thiogalactopyranoside; MDR-TB, multidrug-resistant TB; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; TB, tuberculosis; XDR-TB, extensively drug-resistant TB.

biosynthesis from simple precursors, cells use the salvage pathway to reutilize pyrimidine bases and nucleosides derived from preformed nucleotides (Shambaugh, 1979). In addition, the pyrimidine salvage pathway enzymes enable organisms to acquire exogenously supplied pyrimidine bases and nucleosides, which may not be synthesized by certain species (Wheeler, 1990; Hyde, 2007). Although *M. tuberculosis* pyrimidine salvage pathway appears to be similar to that of *Escherichia coli* and even mammalian cells, it is expected that *M. tuberculosis* will exhibit unique regulatory properties in correlation with its pathogenicity and slow growth (Mizrahi et al., 2005).

A promising target for the rational drug design should be essential for survival of a pathogen and absent from its host. Alternatively, a promising target may play an important role in adaptation of the pathogen to a particular physiological state of the host. Since no transporters for bases, nucleosides or nucleotides of nucleic acids could be identified in *M. tuberculosis* (Braibant et al., 2000), it is thus likely that *M. tuberculosis* has to recycle bases, nucleosides, and/or nucleotides to survive in the hostile environment offered by the host macrophages. In general, pyrimidine bases and nucleosides, which are the transportable precursors of the nucleotides, are not available as exogenous nutrients to most bacteria (Wheeler, 1990). Accordingly, an understanding of the mode of action and role of *M. tuberculosis* pyrimidine salvage pathway enzymes could unveil new targets for the rational design of potent and selective antitubercular agents capable of, hopefully, preventing progression and reactivation of the disease.

Cytidine deaminase (CDA; EC 3.5.4.5), an evolutionarily conserved enzyme of the pyrimidine salvage pathway, catalyzes the hydrolytic deamination of cytidine and 2'-deoxycytidine to form, respectively, uridine and 2'-deoxyuridine (Carter, 1995). The hydrolytic deamination of cytidine nucleosides catalyzed by CDA has been characterized in different organisms, including *E. coli* (Ashley and Bartlett, 1984; Cohen and Wolfenden, 1971), *Bacillus subtilis* (Song and Neuhard, 1989), *Sacharomyces cerevisiae* (Itapa et al., 1970), and human liver (Wentworth and Wolfenden, 1975). The CDA protein family tends to be multi-subunit complexes. *E. coli* CDA occurs as a homodimer with the two active sites per dimer that are formed across the subunit interface (Betts et al., 1994). On the other hand, human CDA is a homotetramer with subunit molecular mass of 16.2 kDa (Vita et al., 1989), whereas *B. subtilis* CDA is a homotetramer with subunit molecular mass of 14.9 kDa (Song and Neuhard, 1989). Crystal structure of human CDA in complex with a tight-binding diazepam inhibitor shows that each active-site is made up of residues from three of four human CDA subunits, and all three contribute directly to recognition of the inhibitor (Chung et al., 2005). In *M. tuberculosis*, a putative *cdd* (*Rv3315c*) gene has been identified in the genome of the H37Rv strain by sequence similarity (Cole et al., 1998). However, the first step to try to understand the role of this gene product should be demonstration that the *Rv3315c* locus encodes a protein having CDA activity in *M. tuberculosis* as predicted by *in silico* analysis, and provide structural and functional analyses of the *cdd*-encoded protein.

In the present work, we report cloning, heterologous recombinant protein expression in *E. coli*, purification to homogeneity, N-terminal amino acid sequencing, electrospray ionization mass spectrometry analysis, and size exclusion chromatography of functional *cdd*-encoded *M. tuberculosis* CDA (MtCDA). We also report measurements of metal concentrations by inductively coupled plasma atomic emission spectrometry (ICP-AES), determination of steady-state kinetic parameters, pH-rate profiles, and X-ray crystallographic analysis of MtCDA using synchrotron radiation. These results provide a solid foundation on which to base gene replacement experiments aiming at understanding the role of MtCDA, if any, in survival and/or latency of the bacillus, and rational drug design.

## 2. Materials and methods

### 2.1. Bacterial strains, media, and growth conditions

*Escherichia coli* DH10B and BL21(DE3) (Novagen) strains were grown at 37 °C in Luria–Bertani (LB; Difco) medium at 180 rpm containing 50 µg/mL ampicillin.

### 2.2. PCR amplification, cloning, and DNA sequencing of the *cdd* coding sequence

Based on the published genome sequence from *M. tuberculosis* H37Rv (Cole et al., 1998), two oligonucleotide primers, CDD1 (5'-gc **cat atg** cct gat gtc gat tgg aat atg ctg-3') and CDD2 (5'-ga **aag ctt** tca ccg gcg ttc ccg ggg gag-3'), complementary to the 5' N-terminal and the 3' C-terminal ends of the probable *cdd* (*Rv3315c*) gene, were synthesized to contain, respectively, NdeI and HindIII restriction sites (bold). These primers were used to amplify the *cdd* coding sequence (402 bp) from *M. tuberculosis* H37Rv genomic DNA. Using standard PCR conditions (an initial denaturation step at 98 °C for 5 min, 35 cycles of denaturation at 98 °C for 30 s, annealing at 55 °C for 55 s, and extension at 72 °C for 2 min, and a final extension step for 10 min at 72 °C) and *Pfu* DNA polymerase for high-fidelity amplification (Stratagene), the amplified product was purified by electrophoresis on a low-melting agarose gel, digested with NdeI and HindIII (New England Biolabs), and ligated into the pET-23a(+) expression vector (Novagen). Nucleotide sequence of the *M. tuberculosis cdd* gene was determined by automated DNA sequencing to confirm the identity, integrity, and absence of PCR-introduced mutations in the cloned gene.

### 2.3. Expression of recombinant MtCDA in *E. coli*

The recombinant plasmid pET-23a(+):*cdd* was transformed into electrocompetent *E. coli* BL21(DE3) cells and selected on LB agar plates containing 50 µg/mL ampicillin (Sambrook and Russell, 2001). A single colony was used to inoculate 50 mL of LB medium in 250 mL flasks, and grown as mentioned above. As cell cultures reached an OD<sub>600</sub> value of 0.4–0.6, the cells were either induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) or not induced; and the cells collected at 3, 6, 9, 12, 24, and 48 h post-induction. The same procedure was employed for *E. coli* BL21(DE3) cells transformed with the pET-23a(+) vector only, which was used as experimental control. Cells were harvested at the indicated time points by centrifugation at 20,800g for 5 min at 4 °C, and the supernatant discarded. Cell pellets were suspended in 600 µL Tris–HCl 50 mM pH 7.5 (buffer A), disrupted by sonication at 4 °C (two pulses of 10 s each at an amplitude value of 21%) and cell debris was separated by centrifugation at 48,000g for 30 min at 4 °C. The soluble extract fractions were analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (Laemmli, 1970).

### 2.4. Purification of recombinant MtCDA

Single colonies were used to inoculate 700 mL of LB medium in 2 L flasks, and grown for 6 h past an OD<sub>600</sub> of 0.4–0.6, without addition of IPTG, which was the optimized protocol for MtCDA expression as discussed below. All subsequent steps were performed at 4 °C, unless stated otherwise. Cells were harvested by centrifugation at 10,879g for 30 min. The resulting cell paste (10 g) was suspended in buffer A (5 mL/g), cells were disrupted by sonication (15 pulses of 10 s each at an amplitude value of 60%) and centrifuged at 23,500g for 30 min. Nucleic acid precipitation present in the supernatant was carried out by addition of streptomycin sulfate 1% (final

concentration), stirring the mixture for 30 min, and centrifugation at 48,000g for 30 min. The supernatant containing soluble CDA was dialyzed against buffer A ( $3 \times 1$  L; 3 h each) using a dialysis tubing with a molecular weight cut off (MWCO) of 3.5 kDa and centrifuged at 48,000g for 30 min. The supernatant was loaded on a Q Sepharose Fast Flow column (GE Healthcare) pre-equilibrated with buffer A. The column was washed with 10 column volumes of buffer A and the adsorbed material was eluted with a linear gradient (0–100%) of 15 column volumes of 50 mM Tris–HCl pH 7.5 containing 500 mM NaCl (buffer B), and the target protein was eluted with approximately 0.9% of buffer B. The fractions containing the target protein were pooled and concentrated using an AMICON ultrafiltration membrane (MWCO = 30 kDa) and loaded on a Sephacryl S-200 column (GE Healthcare) pre-equilibrated with buffer A. The column was washed with 1.5 column volumes of buffer A at 0.25 mL/min flow-rate. The fractions containing the target protein were pooled, incubated with ammonium sulfate to a final concentration of 1 M, stirred for 30 min, and clarified by centrifugation at 23,500g for 30 min. The supernatant was loaded on a Butyl-Sepharose High Performance column (GE Healthcare) pre-equilibrated with 50 mM Tris–HCl pH 7.5 containing 1 M  $(\text{NH}_4)_2\text{SO}_4$  (buffer C). The column was washed with seven column volumes of buffer C and the adsorbed material was eluted with a linear gradient (0–100%) of 20 column volumes of buffer A, and the target protein eluted with approximately 44% of buffer C. All fractions were analyzed by SDS–PAGE and assayed for enzyme activity in the forward direction. The fractions containing homogeneous MtCDA were pooled and stored at  $-80^\circ\text{C}$ . Protein concentration was determined by the method of Bradford (1976) using the Bio-Rad protein assay kit (Bio-Rad) and bovine serum albumin as a standard.

### 2.5. Determination of MtCDA molecular mass

The molecular mass of native *M. tuberculosis* CDA recombinant protein was determined by gel filtration chromatography using a Superdex 200 (HR 10/30) (GE Healthcare) pre-equilibrated with 50 mM Tris–HCl pH 7.5 containing 200 mM NaCl at a flow-rate of 0.4 mL/min. The LMW and HMW Gel Filtration Calibration Kit (GE Healthcare) was used for protein molecular mass standards in the calibration curve, measuring the elution volumes ( $V_e$ ) of several standards (ferritin, catalase, aldolase, albumin, ovalbumin, chymotrypsinogen A, and Ribonuclease A), calculating their corresponding partition coefficients ( $K_{av}$ ), and plotting these values versus the logarithm of their molecular masses.  $K_{av}$  values were determined as  $K_{av} = (V_e - V_o)/(V_t - V_o)$ , where  $V_e$  and  $V_t$  are the elution volume of a sample and total bed volume of the column, respectively, and  $V_o$  is the void volume, which was determined by loading blue dextran 2000 (GE Healthcare) to the column. Protein elution was monitored at 215, 254, and 280 nm.

### 2.6. Mass spectrometry analysis

Recombinant MtCDA was analyzed by MALDI-TOF/TOF on an ABI 4700 Proteomics Analyzer (Applied Biosystems, Framingham, MA, USA), an Ultraflex II (Bruker Daltonics, Germany) and a Q-TOF Ultima API (Micromass, UK) as described elsewhere (Brand et al., 2006).

### 2.7. N-terminal amino acid sequencing

N-terminal amino acid residues of recombinant homogeneous MtCDA were determined by automated Edman degradation sequencing using a PPSQ 23 protein peptide sequencer (Shimadzu Co., Japan) (Brand et al., 2006).

### 2.8. Inductively coupled plasma atomic emission spectroscopy (ICP-AES) analysis of metals

Analysis of  $\text{Co}^{2+}$ ,  $\text{Cd}^{2+}$ ,  $\text{Cu}^{2+}$ ,  $\text{Fe}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Ni}^{2+}$ ,  $\text{Zn}^{2+}$ , and  $\text{Mg}^{2+}$  concentrations were carried out by ICP-AES (Spectro Ciros CCD). All measurements were in duplicate. Recombinant homogeneous MtCDA was extensively dialyzed against 50 mM Tris–HCl pH 7.5 and concentrated by ultrafiltration to a protein concentration of 6 mg/mL.

### 2.9. MtCDA activity assay and determination of steady-state kinetic parameters

MtCDA enzyme activity assays were performed for all purification steps following the direct spectrophotometric assay described by Cohen and Wolfenden (1971). In short, the time-dependent decrease in absorbance at 282 nm upon conversion of cytidine ( $\epsilon_{\text{cytidine}} = 3.6 \text{ M}^{-1} \text{ cm}^{-1}$ ) or 2'-deoxycytidine ( $\epsilon_{2'\text{-deoxycytidine}} = 6.4 \text{ M}^{-1} \text{ cm}^{-1}$ ) to, respectively, uridine or deoxyuridine, was continuously monitored by a UV-2550 UV/visible spectrophotometer (Shimadzu). Determination of steady-state kinetic parameters for cytidine and 2'-deoxycytidine were determined as a function of substrate concentrations ranging from 50 to 1100  $\mu\text{M}$ . The enzymatic assays were monitored using 0.5 cm pathlength quartz cuvettes, in 50 mM Tris–HCl buffer pH 7.5 at  $25^\circ\text{C}$ , and initiated with addition of 4.11  $\mu\text{M}$  of homogeneous recombinant MtCDA. One unit of enzyme activity (U) is defined as the amount of enzyme required to deaminate 1  $\mu\text{mol}$  of either cytidine or 2'-deoxycytidine per min.

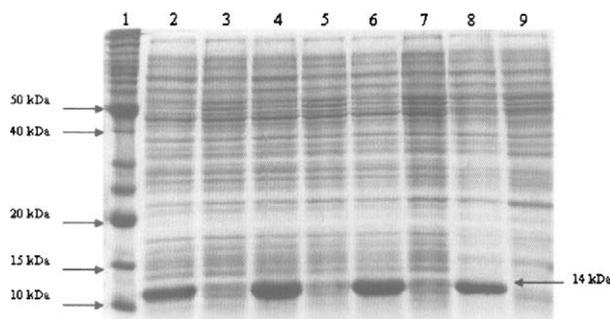
### 2.10. pH profiles

The dependence of the kinetic parameters on pH was studied using cytidine as the variable substrate in a buffer mixture of MES-HEPES-CHES over the following pH values: 4.0, 4.5, 5.0, 6.0, 7.5, 9.0, 10.0, and 11.0 (Cook and Cleland, 2007). Profiles were generated by plotting the either the  $\log k_{\text{cat}}$  or  $\log k_{\text{cat}}/K_M$  versus the pH values, and data fitted to Eq. (1), where  $y$  is the apparent kinetic parameter,  $C$  is the pH-independent plateau value of  $y$ ,  $H$  is the hydrogen ion concentration, and  $K_a$  is the apparent acid dissociation constant for ionizing groups. It should be pointed out that data fitting employed the pH range of 4.0–6.0 for  $\log k_{\text{cat}}$  versus pH plot and of 4–11 for  $\log k_{\text{cat}}/K_M$  versus the pH plot.

$$\log y = \log \left( \frac{C}{1 + \frac{H^+}{K_a}} \right) \quad (1)$$

### 2.11. Crystallization and X-ray data collection

MtCDA was concentrated to 12 mg/mL in sample buffer (20 mM Tris–HCl pH 7.5). Crystals of MtCDA were obtained using the vapor-diffusion hanging-drop method at 273 K using 24-well tissue-culture plates. Crystallization conditions were established using screening kits from Hampton Research (Crystal screens I and II) (Jancarik and Kim, 1991). Each hanging-drop was prepared by mixing 2  $\mu\text{L}$  protein solution and 2  $\mu\text{L}$  reservoir solution and placed over 400  $\mu\text{L}$  reservoir solution. The X-ray diffraction data for MtCDA crystal was collected at a wavelength of 1.437 Å using synchrotron radiation source (Station MX-1, LNLS, Campinas, Brazil) and a CCD detector (MAR CCD). The cryoprotectant contained 30% glycerol. The crystal was flash-frozen at 104 K in a cold nitrogen stream generated and maintained with Oxford Cryosystem. The oscillation range employed was  $1.0^\circ$ , the crystal-to-detector distance was 80 mm and the exposure time 1 min. A data set contained 180 frames was collected and processed to 1.99 Å resolution using the program MOSFLM and scaled with SCALA (CCP4, 1994).



**Fig. 1.** SDS–PAGE (15%) analysis of total soluble protein as a function of growth time. Lane 1, MW Bench marker (Invitrogen); Lane 2, 4, 6, and 8, *E. coli* BL21(DE3) [pET-23a(+):*cdd*]; lanes 3, 5, 7, and 9, *E. coli* BL21(DE3) [pET-23a(+)] (control). The times of cell growth were 3 h (lanes 2 and 3), 6 h (lanes 4 and 5), 9 h (lanes 6 and 7), and 12 h (lanes 8 and 9).

### 2.12. Molecular replacement and crystallographic refinement

The crystal structure of MtCDA was determined by standard molecular replacement methods using the program AMoRe (CCP4, 1994), using as search model the atomic coordinates of CDA from *Mus musculus* (PDB access code: 2FR5). All waters and ligands were removed from the search model and one monomer in the molecular replacement was employed, since analysis of the Matthews coefficient was consistent with one monomer per asymmetric unit in the MtCDA crystal. Further refinement continued with alternate cycles of positional restrained refinement and water molecules addition using the program REFMAC (CCP4, 1994).

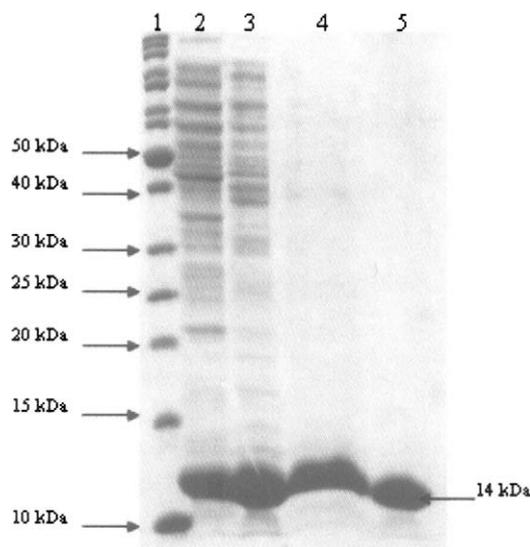
## 3. Results and discussion

### 3.1. *cdd* amplification, cloning, and sequencing

PCR amplification of a probable MtCDA coding sequence (*cdd*, Rv3315c) Cole et al., 1998 from *M. tuberculosis* H37Rv genomic DNA yielded a product with the expected size (402 bp), which was ligated into the pET-23a(+) expression vector. Nucleotide sequence analysis of *M. tuberculosis cdd* coding sequence confirmed both identity and integrity of the cloned product, showing that no mutations were introduced by the PCR amplification steps.

### 3.2. MtCDA expression and purification

SDS–PAGE analysis of *E. coli* BL21(DE3) cells harboring the recombinant pET-23a(+):*cdd* plasmid and grown in the absence of IPTG showed expression in the soluble fraction of a recombinant protein with an apparent molecular mass of 14 kDa, in agreement with the expected size of 14,072 Da for MtCDA (Fig. 1). The pET system makes use of the powerful T7 RNA polymerase, under control of IPTG-inducible *lacUV5* promoter, to transcribe the respective



**Fig. 2.** SDS–PAGE (15%) analysis of pooled fractions of the chromatographic steps of MtCDA protein purification protocol. Lane 1, MW Bench marker (Invitrogen); lane 2, crude extract (190.5 µg); lane 3, Q-Sepharose Fast Flow (207.4 µg); lane 4, Sephacryl S-200 (17.5 µg); lane 5, Butyl HP (11.6 µg). The values given between parentheses are for the amount of total protein loaded on each lane.

cloned target genes (Kelley et al., 1995). Expression of *cdd* gene product showed that high levels of protein production can be obtained in stationary phase for cells growing in the absence of inducer as has been previously reported (Oliveira et al., 2001; Silva et al., 2003; Rizzi et al., 2005). It has been suggested that this phenomenon is due to derepression of *lac* operon when cells approach stationary phase in complex medium and that require cyclic-AMP, acetate, and low pH to achieve high-level expression in the absence of IPTG induction, which may be part of a general cellular response to nutrition limitation (Grossman et al., 1998). However, more recently, it has been shown that unintended induction in the pET system is due to the presence of as little as 0.0001% of lactose in the medium (Studier, 2005).

The anion exchange chromatographic step resulted in 2.4-fold protein purification (Table 1), with removal of some contaminants with apparent subunit molecular mass values larger than 20 kDa (Fig. 2). The gel filtration chromatography resulted in 1.8-fold protein purification, with removal of a number of contaminants (Fig. 2). This decrease in the purification-fold might have been affected by the contaminants purified during this chromatographic step, for instance removal of MtCDA activators, or simply due to uncertainties in protein concentration determination that are inherent to every method. The third purification step (hydrophobic interaction) yielded approximately 10% (Table 1) of homogeneous MtCDA in solution (Fig. 2). This 2.3-fold purification protocol yielded approximately 35 mg (Table 1) of homogeneous *M. tuberculosis* CDA from 10 g of cells.

**Table 1**  
MtCDA protein purification protocol starting from 10 g of cells obtained from 2.8 L of cell culture.

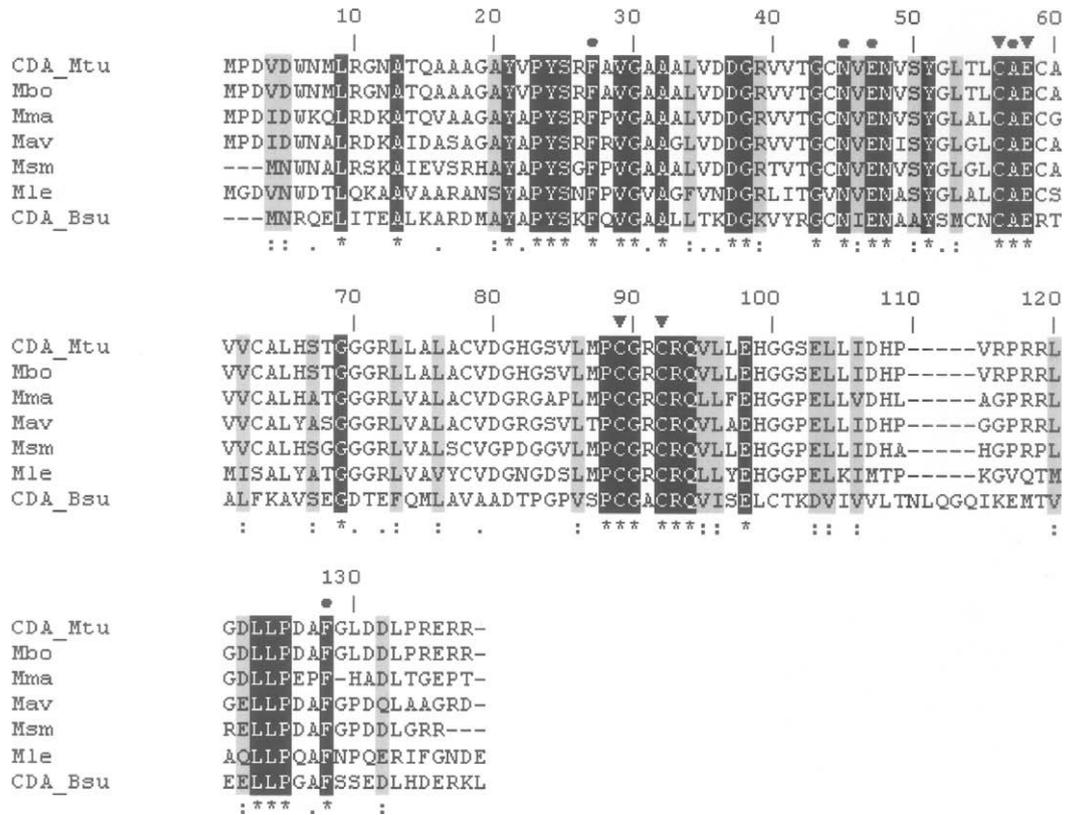
Purification step	Total protein (mg)	Total enzyme activity (U)	Specific activity (U mg <sup>-1</sup> )	Substrate	Purification fold	Yield (%)
Crude extract	819.15	1923.35	2.35	Cytidine	1	100
		1376.36	1.68	2'-deoxycytidine	1	100
Q Sepharose FF	207.4	1151.95	5.55	Cytidine	2.4	60
		799.16	3.85	2'-deoxycytidine	2.3	58
Sephacryl S-200	84.14	363.76	4.32	Cytidine	1.8	19
		259.49	3.08	2'-deoxycytidine	1.8	19
Butyl HP	34.98	190.17	5.44	Cytidine	2.3	10
		131.13	3.75	2'-deoxycytidine	2.2	10

### 3.3. MtCDA molecular mass determination

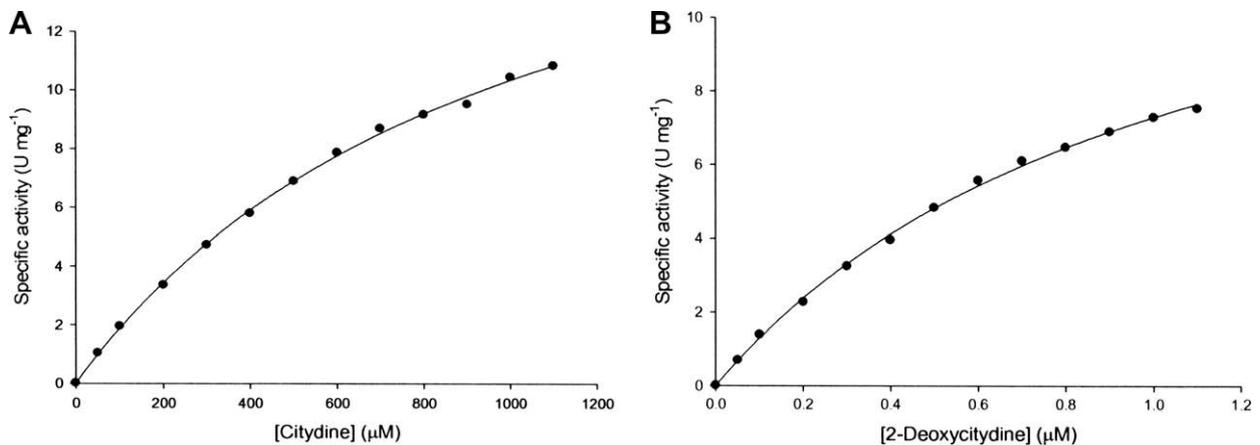
The subunit molecular mass of MtCDA was determined to be 13,938.55 Da by electrospray ionization mass spectrometry (ESI-MS), indicating removal of the N-terminal methionine residue (theoretical molecular mass 14,075.15 Da). These results provide evidence that confirms the identity and purity of the recombinant MtCDA. The protein native molecular mass determined by gel filtration chromatography showed a single peak with elution volume consistent with a molecular mass of 52.99 kDa, suggesting that the recombinant MtCDA protein is a tetramer in solution.

### 3.4. N-terminal amino acid sequencing and metal analysis by ICP-AES

The Edman degradation method identified the first 30 N-terminal amino acid residues of recombinant MtCDA as PDVDWNMLRGNATQAAAAGAYVVPYSRFAVGGAAALVDDGRVVTGCNVENVSYGLTLCAECA. This result unequivocally demonstrates that the homogeneous recombinant protein is MtCDA and confirms the removal of the N-terminal methionine. Protein N-terminal methionine excision is a common type of co-/post-translational modification process that occurs in the cytoplasm of many organisms and in two organelles (i.e. mitochondria and plastids) displaying protein synthesis. Methionine aminopeptidase



**Fig. 3.** Multiple sequence alignment of *M. tuberculosis* CDA (CDA\_Mtu) with homotetrameric *B. subtilis* CDA (CDA\_Bsu), whose crystal structure was determined, and other potential mycobacterial CDAs: *M. bovis* (Mbo), *M. marinum* (Mma), *M. avium* (Mav), *M. smegmatis* (Msm), and *M. leprae* (Mle) using the program CLUSTAL W (Carlow et al., 1999). Identical conserved residues are shown in white on a black background and are also indicated by asterisks below the alignment. Strongly similar and weakly similar residues are identified by colons and periods, respectively. Strongly similar residues are also shaded in gray. Circles and triangles above the alignment indicate, respectively, active-site residues involved in the presumptive ribose binding loop and zinc coordination.



**Fig. 4.** Determination of steady-state kinetic parameters. Substrate saturation curves for (A) cytidine and (B) 2'-deoxycytidine as the variable substrate.

catalyzed cleavage of initiator methionine is usually directed by the penultimate amino acid residues with the smallest side chain radii of gyration (Gly, Ala, Ser, Thr, Pro, Val, and Cys) (Hirel et al., 1989), which is in agreement with removal of the N-terminal methionine from MtCDA.

Metal concentration analysis by ICP-AES yielded the following results:  $\text{Co}^{2+}$ ,  $<0.01$  mg/L;  $\text{Cd}^{2+}$ ,  $<0.01$  mg/L;  $\text{Cu}^{2+}$ ,  $0.17 \pm 0.01$  mg/L;  $\text{Fe}^{2+}$ ,  $0.12 \pm 0.01$  mg/L;  $\text{Mn}^{2+}$ ,  $<0.01$  mg/L;  $\text{Ni}^{2+}$ ,  $<0.01$  mg/L;  $\text{Mg}^{2+}$ ,  $<0.01$  mg/L; and  $\text{Zn}^{2+}$ ,  $28.3 \pm 0.2$  mg/L. These results indicate the presence of one mol of  $\text{Zn}^{2+}$  ( $431 \mu\text{M}$ ) per mol of MtCDA enzyme subunit ( $430.5 \mu\text{M}$ ).

### 3.5. Sequence alignment

Sequence alignment of CDAs from *M. tuberculosis* H37Rv, *M. bovis* AF2122, *M. leprae* ML2174, *M. marinum* MMAR\_1204, *M. avium* 4296, *M. smegmatis*, and *B. subtilis* was carried out using CLUSTAL W (Thompson et al., 1994). The identity between the CDA sequences was 21.58% and almost all of the residues that are conserved among the CDAs from the seven organisms analyzed (Fig. 3) are involved in interactions that are important for the substrate, zinc ion binding and tetrameric interactions. The relationships between the conserved residues and their functional roles have been reported for CDAs from other organisms (Song and Neuhard, 1989; Carlow et al., 1999; Johansson et al., 2002, 2004). The residues Phe27, Asn45, Glu47, Ala57, and Phe128 (*M. tuberculosis* numbering) mediate substrate binding and are conserved in all sequences (Fig. 3). Residues Cys56, Glu58, Cys89, and Cys92 are

reported as probable zinc ligands for *B. subtilis* and *E. coli* CDAs, and were absolutely conserved in all sequences analyzed here (Betts et al., 1994; Johansson et al., 2002, 2004). The zinc ligand is coordinated to three cysteines in tetrameric *B. subtilis* CDA (Carlow et al., 1999), which is in agreement with the three conserved cysteine residues (Cys56, Cys89, and Cys92) and the tetrameric oligomeric state of MtCDA reported here. In *M. tuberculosis* Ser25, Arg93, Gln94, Glu98, and Leu125 are related to tetrameric interactions as was borne out by the structural analysis of MtCDA described below.

### 3.6. MtCDA kinetics

The activity of MtCDA was linearly dependent on recombinant protein concentration for cytidine and 2'-deoxycytidine at a constant concentration value of  $333.6 \mu\text{M}$  (data not shown), demonstrating that true initial velocities are being measured. Substrate saturation curves were hyperbolic (Fig. 4) and thus fitted to the Michaelis–Menten equation, yielding the following steady-state kinetic parameters:  $K_M = 1004 (\pm 53) \mu\text{M}$  and  $k_{\text{cat}} = 4.8 \pm (0.1) \text{s}^{-1}$

**Table 2**  
X-ray data collection and refinement statistics.

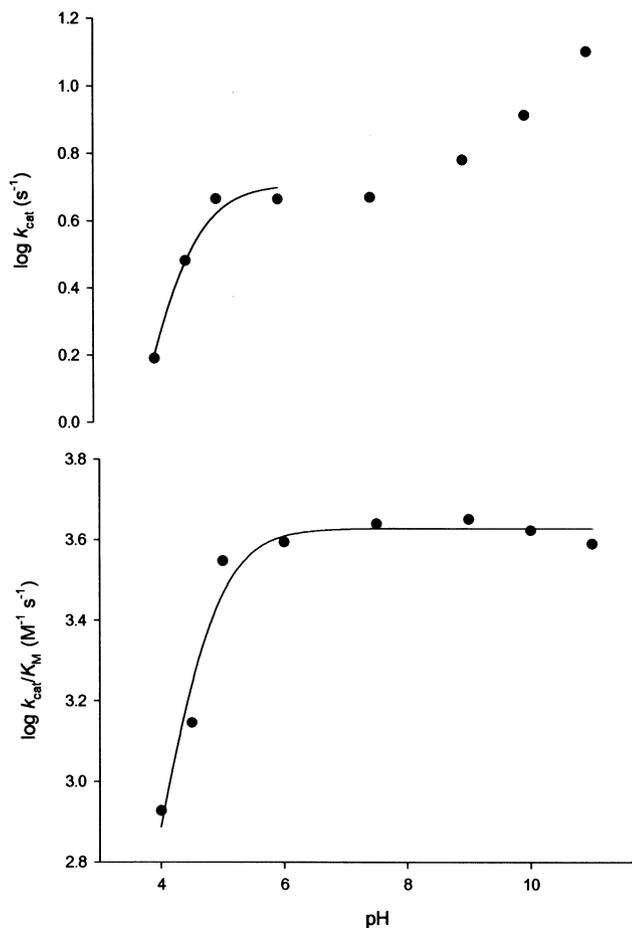
	PDB access code: 3jif
<i>Data collection</i>	
Space group	C222
<i>Cell dimensions</i>	
a, b, c (Å)	63.71, 75.34, 55.11
$\alpha\beta\gamma$ (°)	90.0, 90.0, 90.0
Resolution (Å)	55.13–1.99 (2.10–1.99)*
$R_{\text{sym}}$ (%)	8.6(33.2)
$I/\sigma I$	7.3(2.1)
Completeness (%)	97.0(92.2)
Multiplicity	2.3(2.1)
Number of observations read	49387
Number of unique reflections	9123
<i>Refinement</i>	
Resolution (Å)	18.24–1.99
Number of reflections	8621
$R_{\text{factor}}/R_{\text{free}}$ (%)	20.0/23.5
<i>No. atoms</i>	
Protein	893
Ligand/ion	1
Water	71
<i>B-factors</i>	
Protein (Å <sup>2</sup> )	16.48
$\text{Zn}^{2+}$ (Å <sup>2</sup> )	16.66
Water (Å <sup>2</sup> )	27.85
<i>R.m.s. deviations</i>	
Bond lengths (Å)	0.019
Bond angles (°)	1.985
<i>Ramachdran plot</i>	
Residues in most favored regions (%)	87.0
Residues in additional allowed regions (%)	11.0
Residues in generously allowed regions (%)	2.0
Residues in disallowed regions (%)	0.0
<i>Molprobrity output scores</i>	
Bad rotamers (%)	12.0
Ramachdran outliers (%)	1.7
Ramachdran favored (%)	95.9

$R_{\text{sym}} = \frac{\sum |I_i - \langle I_i \rangle|}{\sum I_i}$ , where  $I_i$  is the scaled intensity of the  $i$ th measurement and  $\langle I_i \rangle$  is the mean intensity for that reflection.

$R_{\text{factor}} = \frac{\sum |F_{\text{obs}}| - |F_{\text{calc}}|}{\sum F_{\text{obs}}}$ , where  $F_{\text{calc}}$  and  $F_{\text{obs}}$  are the calculated and observed structure factor amplitudes, respectively.

$R_{\text{free}}$  = as for  $R_{\text{factor}}$ , but for 5% of the total reflections chosen at random and omitted from refinement.

\* Values in parentheses are for highest-resolution shell.



**Fig. 5.** Dependence of kinetic parameters on pH for MtCDA-catalyzed deamination of cytidine. (A) pH dependence of  $\log k_{\text{cat}}$  on pH, (B) pH dependence of  $\log k_{\text{cat}}/K_M$ . Data fitting as described in Section 2.

for cytidine, and  $K_M = 1059 (\pm 64) \mu\text{M}$ , and  $k_{\text{cat}} = 3.5 (\pm 0.1) \text{s}^{-1}$  for 2'-deoxycytidine. The MtCDA  $K_M$  values are larger than *E. coli* CDA (220  $\mu\text{M}$  for cytidine and 58  $\mu\text{M}$  for 2'-deoxycytidine). In addition, the  $k_{\text{cat}}$  value for *M. tuberculosis* CDA was lower than the ones reported for *E. coli* and human liver CDAs (136.6 and 65  $\text{s}^{-1}$ , respectively) (Ashley and Bartlett, 1984). The specificity constant values ( $k_{\text{cat}}/K_M$ ) for both substrates are quite similar ( $4.8 \times 10^3 \text{M}^{-1} \text{s}^{-1}$  for cytidine and  $3.3 \times 10^3 \text{M}^{-1} \text{s}^{-1}$  for 2'-deoxycytidine), which indicates that MtCDA has no preference for any of these two substrates free in solution. On the other hand, the *E. coli* CDA has been shown to be more specific for 2'-deoxycytidine than cytidine (Ashley and Bartlett, 1984; Trimble and Maley, 1971). Interestingly, the average concentration of cytidine, which represents a compilation of published data on mammalian cells and fluids, has been found to be 4.7 ( $\pm 2.7$ )  $\mu\text{M}$  Traut, 1994. This value is approximately 200-fold lower than the MtCDA  $K_M$  value for cytidine, which would suggest that this enzyme activity is low in the physiological milieu of the host. However, it has been pointed out that the average concentration of bases and nucleosides in plasma and other extracellular fluids is generally in a concentration range that is usually lower than corresponding intracellular concentrations (Traut, 1994).

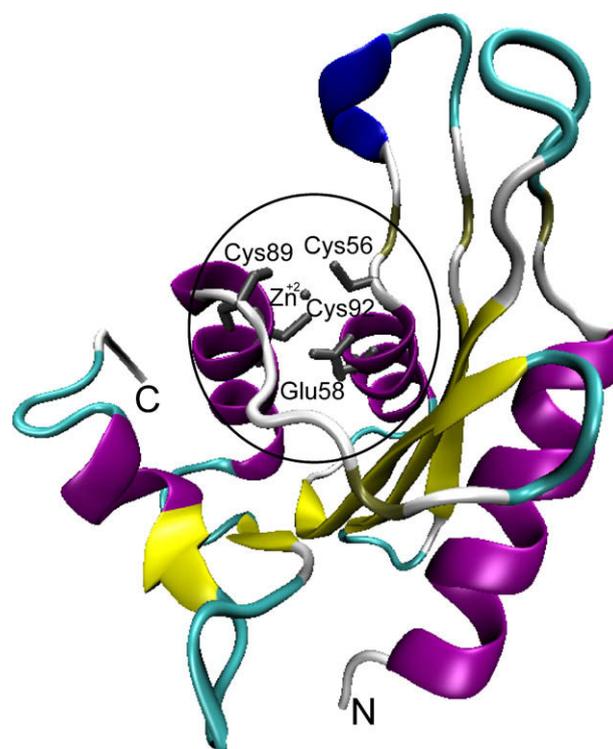
To assess the role of acid/base chemistry in the MtCDA enzymatic reaction, the pH dependence of  $k_{\text{cat}}$  and  $k_{\text{cat}}/K_M$  for cytidine was studied. The pH-rate profiles of the kinetic parameters using cytidine as substrate are shown in Fig. 5. It should be pointed out that data fitting employed the pH range of 4.0–6.0 for  $\log k_{\text{cat}}$  versus pH plot and of 4–11 for  $\log k_{\text{cat}}/K_M$  versus pH plot. The pH-rate profile for  $k_{\text{cat}}$  was quite complex. It showed a decrease at low pHs with a slope of 1 (Fig. 5A), demonstrating that protonation of a group with apparent  $\text{p}K_a$  value of 4.3 ( $\pm 1.0$ ) abolishes MtCDA activity for cytidine as substrate. Moreover, pH-rate profile for  $k_{\text{cat}}$  indicates that a group with apparent  $\text{p}K_a$  value of 8–9 increases even further the catalytic rate when deprotonated that could be ascribed to either of the three conserved cysteines (Cys56, Cys89, or Cys92) that are coordinated to a zinc atom. Even though pH dependence profiles exhibiting three  $\text{p}K_a$  values have been described (Cook and Cleland, 2007), measurements of enzyme activity at pH values larger than 11 would be required for MtCDA. However, most enzymes and proteins undergo denaturative changes in conformation when exposed to pH values above 10–12 (Englard and Seifter, 1990). At any rate, it is preliminary to invoke more complex models at this point and further experimental results should be provided. Notwithstanding, the  $k_{\text{cat}}/K_M$  pH-rate profile showed a decrease at low pH values with a slope of 1 (Fig. 5B), suggesting that protonation of a single ionizable group, exhibiting an apparent  $\text{p}K_a$  value of 4.7 ( $\pm 0.7$ ), reduces cytidine binding. The  $\text{p}K_a$  values of 4.3 for  $k_{\text{cat}}$  and of 4.7 for  $k_{\text{cat}}/K_M$  lie in the normal  $\text{p}K$  range for the  $\beta$ -carboxyl of aspartate and  $\gamma$ -carboxyl of glutamate amino acid residues. A glutamic acid residue has been shown to be conserved in dimeric and tetrameric CDAs from different organisms and was correlated with substrate binding (Carlow et al., 1999; Trimble and Maley, 1971). The pH-rate profiles of *E. coli* and *B. subtilis* indicated a decrease at low pHs, with  $\text{p}K_a$  values of 4.7 and 5.4, respectively, which also display a single protonated group (Carlow et al., 1999). It is thus tempting to suggest that either the conserved Glu47 or Glu58 residue of MtCDA (Fig. 3) plays an important role in substrate binding and/or catalysis. Site-directed mutagenesis, equilibrium binding, and measurements of steady-state kinetic parameters of mutant protein will be carried out to ascertain whether Glu47 or Glu58 plays any role in substrate binding and/or catalysis.

### 3.7. Structure analysis

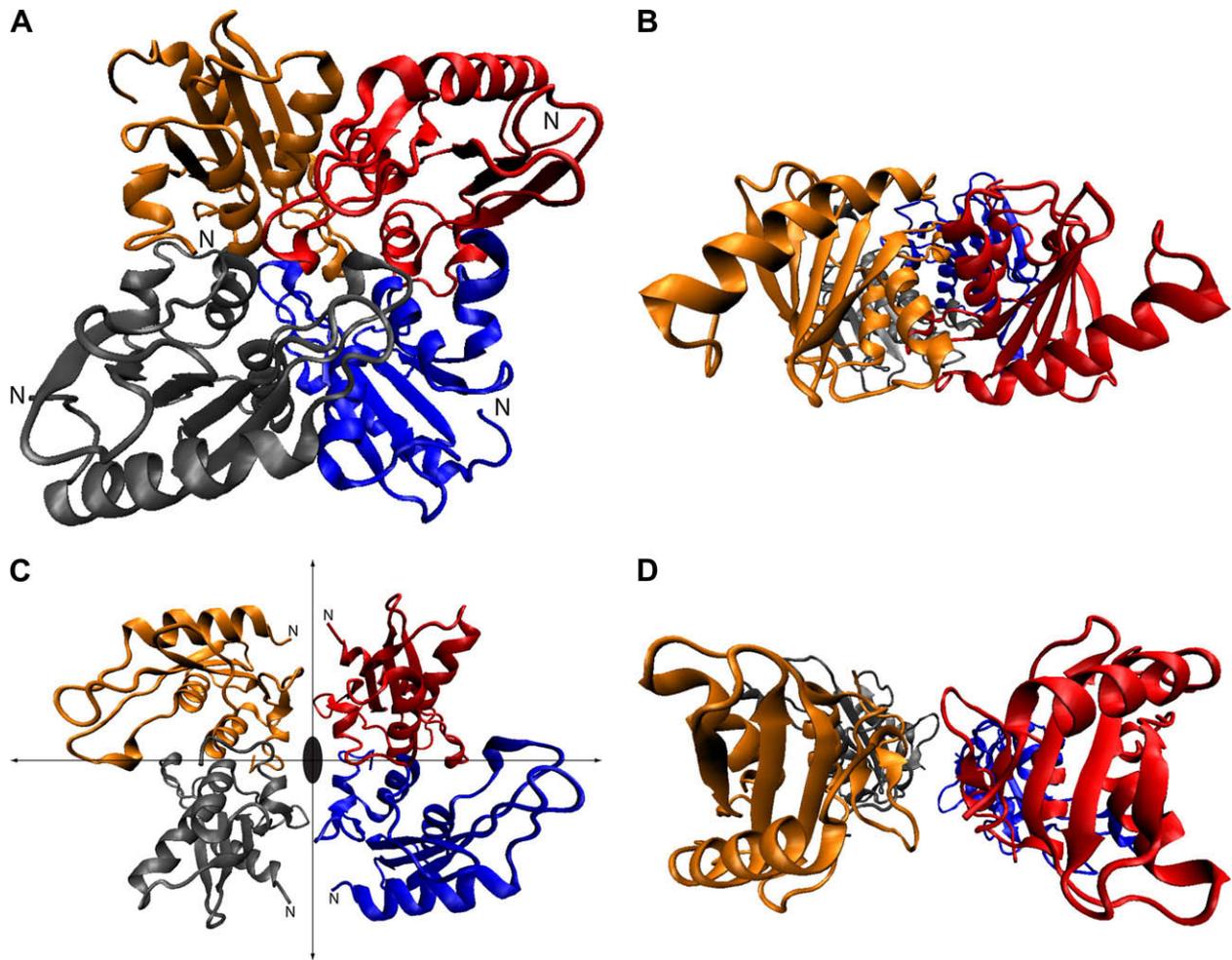
The MtCDA crystal structure was determined by molecular replacement methodology and refined to 2.0 Å resolution using

synchrotron radiation source. Data collection and refinement statistics for this structure are given in Table 2. The structure of MtCDA is compact and exhibits the canonical cytidine deaminase fold (Fig. 6). This structure is composed of five  $\beta$ -strands (Ala28–Val35, Arg39–Asn45, Arg72–Gly81, Leu104–Asp107, and Pro112–Leu115) and five helices (Trp6–Ala18, Ser50–Thr54, Cys56–Thr68, Cys89–Gly100, Leu115–Leu119). The MtCDA monomer comprises a single domain and its secondary structural elements are arranged in a three-layer core ( $\alpha/\beta/\alpha$ ) with a mixed  $\beta$ -sheet of five strands (order: 21345). This five-stranded  $\beta$ -sheet presents the last two strands forming a beta-wing in a highly flexible region.

There is one MtCDA monomer in the asymmetric unit; however, analysis of the crystal packing indicated the presence of intermolecular interactions forming a tetramer in the cell unit, consistent with the tetrameric oligomeric state determined by gel filtration chromatography of native MtCDA protein. There are 21 structures of CDAs deposited in the PDB (November, 2009), 17 are homotetramers and four homodimers. Among these 21 structures, we did not count the present MtCDA that has been deposited in the PDB (3IJF). Fig. 7A and B show two views of the canonical homotetrameric structure observed for CDAs from *S. cerevisiae* (1R5T), *B. subtilis* (1UX1, 1UWZ, 1UX0, 1JTK), *Mus musculus* (1ZAB, 2FR5, 2FR6), *Bacillus anthracis* (2D30), *Burkholderia pseudomallei* (3DMO), human (1MQ0), and *Aspergillus terreus* (1WN5, 1WN6, 2Z3G, 2Z3H, 2Z3I, 2Z3J). Fig. 7C and D show the MtCDA homotetrameric structure. Each monomer in the MtCDA tetrameric structure was generated by successive application of twofold rotation axes against the atomic coordinates of the content of the asymmetric unit. Structure comparison with other CDA structures indicates that the homotetrameric structure of MtCDA is different from the homotetrameric structures of previously deposited CDAs in at least two aspects. Previously solved CDAs present a more compact homotetramer, with a radius of gyration of 28 Å against 30 Å observed for



**Fig. 6.** Structure of cytidine deaminase from *Mycobacterium tuberculosis*. The Zn-binding pocket is indicated by an ellipse, in which the main residues involved in the coordination of  $\text{Zn}^{2+}$  can be seen. This figure and Figs. 7A–D and 8A and B were generated using VMD (Humphrey et al., 1996).



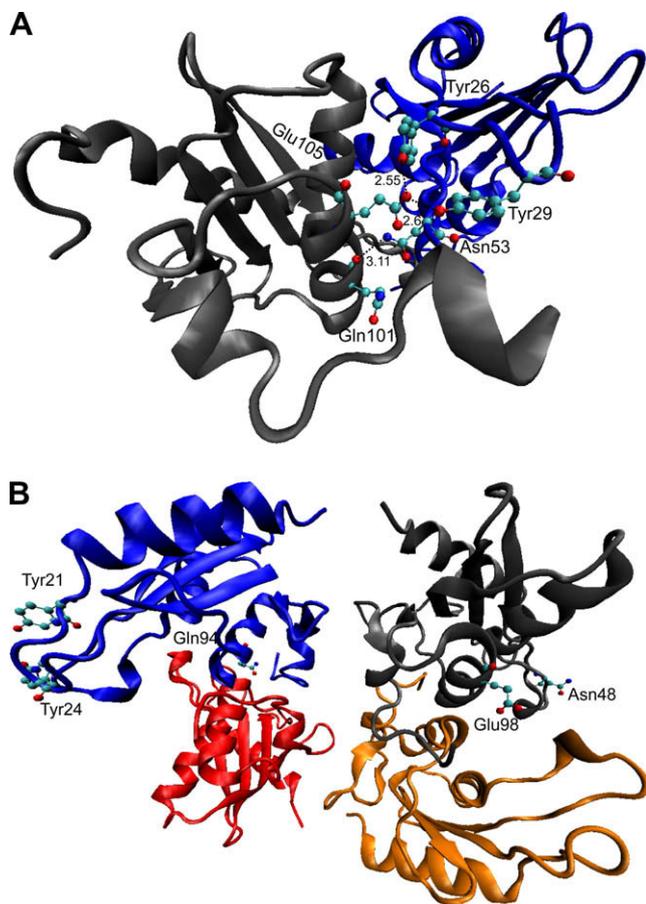
**Fig. 7.** Homotetramers of CDAs. (A) Canonical tetrameric structure. (B) Fig. 7A rotated 90°. (C) MtCDA tetrameric structure. (D) Fig. 7C rotated 90°. Chains A–D are colored in blue, red, gray and orange, respectively. Two-fold rotation axes employed to generate the crystallographic tetramer (MtCDA) are indicated in Fig. 7C.

MtCDA. In addition, the canonical tetramer presents N-terminal helices pointing towards the solvent. N-termini of chains A and B are 70 Å away from the equivalent termini of chains C and D, respectively (Fig. 7A and B). MtCDA presents N-terminal helices of chain A and B pointing towards N-terminal helices of chain C and D, respectively (Fig. 7C and D), and the N-termini are only 9 Å apart from each other (A–C and B–D).

An analysis of the of the intermolecular contacts using NCONT indicated participation of the residues Gly19, Tyr21, Tyr 24, Ser25, Val46, Glu47, Asn48, Val49, Ser50, Tyr51, Gly52, Leu53, Thr54, Leu55, Cys56, Cys59, Ala60, Val62, Cys63, His66, Ser67, Cys89, Arg93, Gln94, Val95, Glu98, His99, Phe123, and Leu125 in intermolecular contacts. In protein structures the interfaces frequently hide non-polar side chains, and polar groups originating from both main chain and the side chains. In the MtCDA tetramer most of the contacts are of hydrophobic nature. The residues Ser25, Arg93, Gln94, Glu98, and Leu124 have been previously identified as participating in intermolecular contacts in CDA structures (Carlow et al., 1999). The tetrameric CDAs contain two highly conserved tyrosine residues not found in any of the dimeric CDA sequences (Carlow et al., 1999), which correspond to the two highly conserved tyrosine residues (Tyr24 and Tyr51) identified in MtCDA sequence (Fig. 3), and is in agreement with the gel filtration and structural results here reported. Most of the residues involved in the intermolecular contacts (Fig. 3) are conserved in the CDA structures, for instance, the strong hydrogen bonds involving residues: Asp121:O-

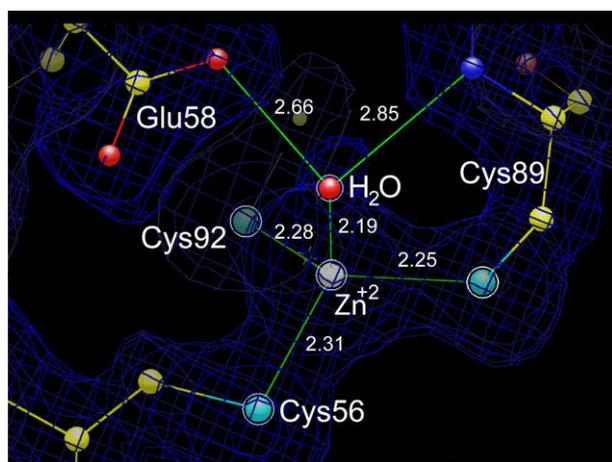
Arg93NH2 (2.9 Å), Pro120:O-Arg93:NH1(2.8 Å) and Gly90:N-Gln94:OE1(2.6 Å), which suggest the importance of these residues in the stabilization of the tetrameric structure. These common intermolecular contacts are observed in the interface of chains A–B and C–D. On the other hand, when we analyze the interfaces A–C and B–D the canonical CDA tetramers show intermolecular contacts involving Tyr26–Glu105, Tyr29–Glu105, and Gln101–Asn53 (we used here the numbering of CDA from *S. cerevisiae*). These three pairs of residues are conserved in all tetrameric CDAs, including MtCDA, and are representative of important intermolecular contacts. Fig. 8A shows these intermolecular hydrogen bonds for the structure of CDA from *S. cerevisiae* (1R5T) (chains A and C, which are equivalent to chains B and D). Analysis of equivalent region of MtCDA for pairs of residues Tyr21–Glu98, Tyr24–Glu98, and Gln94–Asn48, indicates that these intermolecular contacts are not observed in MtCDA (Fig. 8B). All chains are shown in Fig. 8B to make clear that none intermolecular involving these residues is observed in the MtCDA structure. Furthermore, MtCDA was crystallized in the space group C222, not observed for the previously solved CDA structures. Based on the evidence gathered here we show that the MtCDA tetramer presents a different structural arrangement, which may be due to unique crystallographic packing observed for this structure.

Analysis of the MtCDA structure indicated the zinc ion is coordinated to the side chain of Cys56, Cys89, Cys92, and a water molecule, with distances between zinc ion and sulfur atoms of cysteine



**Fig. 8.** Interface region for homotetrameric CDA structures. (A) Canonical tetrameric structures (only chains A and C). (B) For MtCDA structure all chains. Chains A–D are colored in blue, red, gray and orange, respectively. All distances are in Å.

residues ranging from 2.25 to 2.31 Å (Fig. 9). The side chain oxygen of Glu58 and main chain nitrogen of Cys89 are hydrogen bonded to the water molecule coordinating the zinc ion, such tetrahedral arrangement (Fig. 9) has been previously described in the CDA structures isolated from mouse (PDB access codes: 1ZAB and 2FR6) and from *B. anthracis* (2D30). In addition, this structural arrangement is consistent with both the pH-rate profiles described



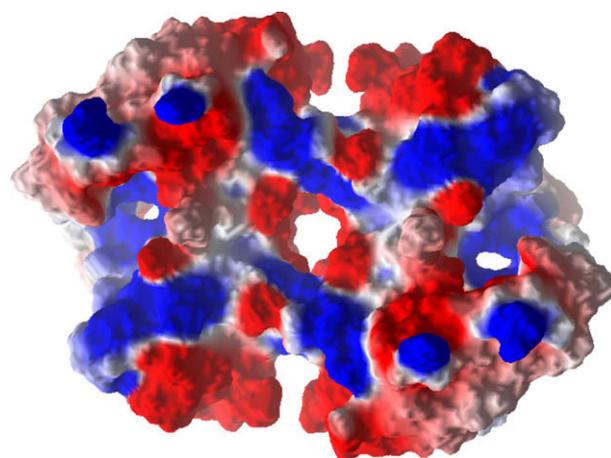
**Fig. 9.** 2fo-fc electron density map contoured at 2- $\sigma$  cutoff, where we can see Zinc coordination by Cys56, Cys89, Cys92, and a water molecule. This water is hydrogen bonded to Glu58 and Cys89. All distances are in Å.

above and the proposed mechanism of bacterial CDAs (Chung et al., 2005). In short, an active-site zinc atom is coordinated to a nucleophilic water/hydroxide and a conserved glutamate is envisaged to promote the initial attack at C4 of cytosine ring by protonating the adjacent N3-position and deprotonating the nucleophilic water, then again using general acid/base chemistry to facilitate breakdown of the tetrahedral intermediate.

Analysis of the presence of cavities using the moldock algorithm (Thomsen and Christensen, 2006) in the structure of the homotetrameric MtCDA indicated the presence of a major cavity in the center of the tetramer with a volume of 325.6 Å<sup>3</sup>. This major cavity is not observed in other CDA structures solved so far. Detailed analysis of this major cavity indicates that the side chain of Asp121 from two crystallographic related chains are pointing towards the center of the cavity. In addition, side chains of four symmetry related Asp117 are also pointing towards the center, these residues participate in a strong intramolecular hydrogen bond with Arg114. This prevalence of polar side chains creates a highly hydrophilic environment in this region, as can be seen in the electrostatic surface of the tetramer (Fig. 10), which is filled with 24-well-ordered water molecules.

#### 4. Concluding remarks

Enzyme kinetics and structural studies provide a framework on which to base the target-based rational design of new agents with antitubercular activity. However, the availability of sufficient amounts of proteins of *M. tuberculosis* still remains an essential and laborious step. Unfortunately, even when a genome can be sequenced, only up to 20% of the protein targets can produce soluble proteins under very basic experimental conditions (Lesley et al., 2002). Thus, expression of proteins in soluble form has been identified as an important bottleneck in efforts to determine biological activity and crystal structure of *M. tuberculosis* proteins (Vicentelli et al., 2003). The results presented here not only provide homogeneous MtCDA for further studies on the enzyme mechanism of action by, for instance, pre-steady-state kinetics and microcalorimetry, but also for X-ray crystal structure determination of MtCDA-ligand binary complexes. Even though CDA is present in the human host, it has been shown that the overall dissociation constant value of 3,4,5,6-tetrahydrouridine inhibitor is 0.24 μM for *E. coli* CDA (Ashley and Bartlett, 1984) as compared to 29 nM for human liver CDA (Wentworth and Wolfenden, 1975). The larger MtCDA  $K_M$  values (1004 μM for cytidine and 1059 μM for



**Fig. 10.** Electrostatic surface of MtCDA, blue indicates positively charged regions, red negatively charged regions and white is neutral. This figure was generated using moldock (Thomsen and Christensen, 2006).

2'-deoxycytidine) as compared to recombinant human liver CDA (39  $\mu\text{M}$  for both substrates) (Vincenzetti et al., 1996) indicate that there are differences in substrate recognition that can be exploited in inhibitor design. Moreover, it is feasible to design function-based species-specific inhibitors as has been demonstrated for bovine and human purine nucleoside phosphorylases despite these enzymes sharing 87% overall amino acid sequence identity and having totally conserved active-site residues (Taylor Ringia et al., 2006). The results reported here also provide experimental evidence that the *cdd* gene codes for a functional CDA enzyme in *M. tuberculosis*, which represents a solid foundation on which to base efforts towards understanding the role, if any, of pyrimidine salvage pathway in *M. tuberculosis* survival and/or persistence by gene replacement.

## Acknowledgments

This work was supported by National Institute of Science and Technology on Tuberculosis (Decit/SCTIE/MS-MCT-CNPq-FNDCT-CAPES) and Millennium Initiative Program (CNPq), Brazil, to D.S.S. and L.A.B., C.B. Jr. acknowledges financial support from "Embrapa Recursos Genéticos e Biotecnologia", Brazil. D.S.S. (304051/1975-06), L.A.B. (520182/99-5), C.B.J. (304034/2008-8), and W.F.A. Jr. (300851/98-7) are research career awardees of the National Council for Scientific and Technological Development of Brazil (CNPq). This research was partially supported by Laboratório Nacional de Luz Síncrotron (LNLS, Campinas, Brazil). Z.A.S.Q. acknowledges a scholarship awarded by CNPq.

## References

- Ashley, G.W., Bartlett, P.A., 1984. Purification and properties of cytidine deaminase from *Escherichia coli*. *J. Biol. Chem.* 259, 13615–13620.
- Ashley, G.W., Bartlett, P.A., 1984. Inhibition of *Escherichia coli* cytidine deaminase by a phosphapyrimidine nucleoside. *J. Biol. Chem.* 259, 13621–13627.
- Betts, L., Xiang, S., Short, S.A., Wolfenden, R., Carter, C.W., 1994. Cytidine deaminase. The 2.3 Å crystal structure of an enzyme: transition-state analog complex. *J. Mol. Biol.* 235, 635–656.
- Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254.
- Braibant, M., Gilot, P., Content, J., 2000. The ATP binding cassette (ABC) transport systems of *Mycobacterium tuberculosis*. *FEMS Microbiol. Rev.* 24, 449–467.
- Brand, G.D., Krause, F.C., Silva, L.P., Leite, J.R., Melo, J.A., Prates, M.V., Pesquero, J.B., Santos, E.L., Nakaie, C.R., Costa-Neto, C.M., Bloch Jr., C., 2006. Bradykinin related peptides from *Phyllomedusa hypochondrialis*. *Peptides* 27, 2137–2146.
- Carlow, D.C., Carter, C.W., Mejhede, N., Wolfenden, R., 1999. Cytidine deaminases from *B. subtilis* and *E. coli*: compensating effects of changing zinc coordination and quaternary structure. *Biochemistry* 38, 12258–12265.
- Carter, C.W., 1995. The nucleoside deaminases for cytidine and adenosine: structure, transition state stabilization, mechanism, and evolution. *Biochimie* 77, 92–98.
- CCP4, Collaborative computational project number 4, *Acta Cryst. D* 50 (1994) 760–763.
- Centers for Disease Control and Prevention (CDC), 2006. Emergence of *Mycobacterium tuberculosis* with extensive resistance to second-line drugs worldwide, 2000–2004. *Morb. Mortal. Wkly. Rep.* 55, 301–305.
- Chung, S.J., Fromme, J.C., Verdine, G.L., 2005. Structure of human cytidine deaminase bound to a potent inhibitor. *J. Med. Chem.* 48, 658–660.
- Cohen, R.M., Wolfenden, R., 1971. Cytidine deaminase *Escherichia coli*. *J. Biol. Chem.* 246, 7561–7565.
- Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E., Tekaiia, 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., Barrell, B.G., 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393, 537–544.
- Cook, P.F., Cleland, W.W., 2007. PH Dependence of kinetic parameters and isotope effects. In: Cook, P.F., Cleland, W.W. (Eds.), *Enzyme Kinetics and Mechanisms*. Garland Science, London and New York, pp. 325–366.
- Dye, C., Scheele, S., Dolin, P., Pathania, V., Ravigione, M.C., 1999. Global burden of tuberculosis: Estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. *J. Am. Med. Ass.* 282, 677–686.
- England, S., Seifter, S., 1990. Precipitation techniques. *Methods Enzymol.* 182, 285–300.
- Grossman, T.H., Kawasaki, E.S., Punreddy, S.R., Osburne, M.S., 1998. Spontaneous cAMP-dependent depression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene* 209, 95–103.
- Hingley-Wilson, S.M., Sambandamurthy, V.K., Jacobs Jr., W.R., 2003. Survival perspectives from the world's most successful pathogen, *Mycobacterium tuberculosis*. *Nat. Immunol.* 4, 949–955.
- Hirel, P.H., Schmitter, M.J., Dessen, P., Fayat, G., Blanquet, S., 1989. 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, 8247–8251.
- Humphrey, W., Dalke, A., Schulten, K., 1996. VMD – visual molecular dynamics. *J. Mol. Graph.* 14, 33–38.
- Hyde, J.E., 2007. Targeting purine and pyrimidine metabolism in human apicomplexan parasites. *Curr. Drug Targets* 8, 31–47.
- Itapa, P.L., Cercignani, G., Balestrieri, E., 1970. Partial purification and properties of Cytidine deaminase from Baker's yeast. *Biochemistry* 9, 3390–3395.
- Jain, A., Mondal, R., 2008. Extensively drug-resistant tuberculosis: current challenges and threats. *FEMS Immunol. Med. Microbiol.* 53, 145–150.
- Jancarik, J., Kim, S.-H., 1991. Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.* 24, 409–411.
- Johansson, E., Mejlhede, N., Larsen, S., 2002. Crystal structure of the cytidine deaminase from *Bacillus subtilis* at 2.0 Å resolution. *Biochemistry* 41, 2563–2570.
- Johansson, E., Neuhaud, J., Willemoe, M., Larsen, S., 2004. Structural, kinetic, and mutational studies of the zinc ion environment in tetrameric cytidine deaminase. *Biochemistry* 43, 6020–6029.
- Kelley, K.C., Huestis, K.J., Austen, D.A., Sanderson, C.T., Donohue, M.A., Stickel, S.K., Kawasaki, E.S., Osburne, M.S., 1995. Regulation of CD4-183 gene expression from phage-T7-based vectors in *Escherichia coli*. *Gene* 156, 33–36.
- Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
- Lesley, S.A., Kuhn, P., Godzik, A., Deacon, A.M., Mathews, I., Kreusch, A., Spraggon, G., Klock, H.E., McMullan, D., Shin, T., Vincent, J., Robb, A., Brinen, L.S., Miller, M.D., McPhillips, T.M., Miller, M.A., Scheibe, D., Canaves, J.M., Guda, C., Jaroszewski, L., Selby, T.L., Elsliger, M.A., Wooley, J., Taylor, S.S., Hodgson, K.O., Wilson, I.A., Schultz, P.G., Stevens, R.C., 2002. Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline. *Proc. Natl. Acad. Sci. USA* 99, 11664–11669.
- Mizrahi, V., Buckstein, M., Rubin, H., 2005. Nucleic acid metabolism. In: Cole, S.T., Eisenach, K.D., McMurray, D.N., Jacobs, W.R., Jr. (Eds.), *Tuberculosis and the Tubercle Bacillus*. ASM Press, Washington, DC, pp. 369–378.
- Oliveira, J.S., Pinto, C.A., Basso, L.A., Santos, D.S., 2001. Cloning and overexpression in soluble form of functional shikimate kinase and 5-enolpyruvylshikimate 3-phosphate synthase enzymes from *Mycobacterium tuberculosis*. *Protein Expr. Purif.* 22, 430–435.
- Rizzi, C., Frazzon, J., Ely, F., Weber, P.G., Fonseca, I.O., Gallas, M., Oliveira, J.S., Mendes, M.A., Souza, B.M., Palma, M.S., Santos, D.S., Basso, L.A., 2005. DHP synthase from *Mycobacterium tuberculosis* H37Rv: cloning, expression, and purification of functional enzyme. *Protein Expr. Purif.* 40, 23–30.
- Sambrook, J., Russell, D.W., 2001. *Molecular Cloning a Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
- Shambaugh, G.E., 1979. Pyrimidine biosynthesis. *Am. J. Clin. Nutr.* 32, 1290–1297.
- Silva, R.G., Carvalho, L.P., Oliveira, J.S., Pinto, C.A., Mendes, M.A., Palma, M.S., Basso, L.A., Santos, D.S., 2003. Cloning, overexpression, and purification of functional human purine nucleoside phosphorylase. *Protein Expr. Purif.* 27, 158–164.
- Song, B.H., Neuhaud, J., 1989. Chromosomal location, cloning and nucleotide sequence of the *Bacillus subtilis* *cdd* gene encoding cytidine/deoxycytidine deaminase. *Mol. Gen. Genet.* 216, 462–468.
- Studier, F.W., 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41, 207–234.
- Taylor Ringia, E.A., Tyler, P.C., Evans, G.B., Furneaux, R.H., Murkin, A.S., Schramm, V.L., 2006. Transition state analogue discrimination by stabilized purine nucleoside phosphorylases. *J. Am. Chem. Soc.* 128, 7126–7127.
- WHO/Stop TB Partnership. The Global Stop TB 2006–2015. WHO/HTM/STB/2006.35, World Health Organization. Available from: <<http://www.who.int>>.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. 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, 4673–4680.
- Thomsen, R., Christensen, M.H., 2006. MolDock: a new technique for high-accuracy molecular docking. *J. Med. Chem.* 49, 3315–3321.
- Traut, T.W., 1994. Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.* 140, 1–22.
- Trimble, R.B., Maley, F., 1971. Metabolism of 4-N-hydroxy-cytidine in *Escherichia coli*. *J. Bacteriol.* 108, 145–153.
- Vicentelli, R., Bignon, C., Gruez, A., Cnaan, S., Sulzenbacher, G., Tegoni, M., Campanacci, V., Cambillau, C., 2003. Medium-scale structural genomics: strategies for protein expression and crystallization. *Acc. Chem. Res.* 36, 197–225.

- Vincenzetti, S., Cambi, A., Neuhard, J., Garattini, E., Vita, A., 1996. Recombinant human cytidine deaminase: expression, purification, and characterization. *Prot. Protein Exp. Purif.* 8, 247–253.
- Vita, A., Cacciamani, T., Natalini, P., Puggieri, S., Magn, G., 1989. Cytidine deaminase from human spleen. *Adv. Exp. Med. Biol.* 253, 71–77.
- Wentworth, D.F., Wolfenden, R., 1975. On the interaction of 3,4,5,6-tetrahydrouridine with human liver cytidine deaminase. *Biochemistry* 14, 5099–5105.
- Wheeler, P.R., 1990. Biosynthesis and scavenging of pyrimidines by pathogenic mycobacteria. *J. Gen. Microbiol.* 136, 189–201.