UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE CIENCIAS BÁSICAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS: BIOQUÍMICA

O estudo da enzima deidroquinato sintase de *Mycobacterium tuberculosis* H37Rv como alvo para o desenvolvimento de fármacos antituberculose.

Jordana Dutra de Mendonça

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Biológicas: Bioquímica, da Universidade Federal do Rio Grande do Sul, como pré-requisito para a obtenção do grau de Mestre em Bioquímica.

Orientador: Luiz Augusto Basso

Porto Alegre, 2010.

Este trabalho é dedicado a todos que junto comigo o fizeram tornar-se realidade.

Este trabalho foi realizado no Centro de Pesquisas em Biologia Molecular e Funcional, Instituto Nacional de Ciência e Tecnologia em Tuberculose da Pontifícia Universidade Católica do Rio Grande do Sul, sob orientação do Prof. Luiz Augusto Basso.

Agradecimentos

Aos Professores **Luiz Augusto Basso** e **Diógenes Santiago Santos**, pela confiança, orientação e oportunidade de participar de um grupo de pesquisa tão sólido e com um trabalho tão relevante para a ciência e para a sociedade.

Aos colegas do **CPBMF** e **4G**, pelo apoio, convívio e ensinamentos, científicos ou não. Em especial, à **Gisele Biazus**, pois, quando achei que tudo estava perdido, e se não fosse um dos seus diversos projetos de mestrado, eu não teria disponível a enzima DHase, fundamental para a realização dos ensaios cinéticos. Muito Obrigado! E também para minha amiga **Ardala Breda**, que sem seu apoio, amizade e incentivo, teria sido muito mais difícil (quase impossível) chegar ao fim dessa jornada.

Aos meus **amigos** sempre presentes, fisicamente ou em pensamento, sei que sempre estiveram torcendo por mim. **Fernanda Godinho, Luciana Eberle, Daiane Dal´Pra**, minhas fiéis escudeiras!

À **minha família, Jair, Ligia, Fernanda e Rafael,** e meu namorado, **Thiago**, que sempre compreenderam a minha ausência e me ajudaram no que foi preciso. Sem vocês não teria dado certo!

Ao **Programa de Pós-Graduação em Bioquímica**, pela oportunidade de desenvolver o mestrado.

Ao **CNPq/MCT**, pela bolsa de mestrado.

Sumário

PARTE I

1. **Resumo**

Apesar da incidência *per capita* da tuberculose (TB) ter se mantido estável em 2005, o número de novos casos que surgem a cada ano continua a aumentar no mundo todo. De acordo com a Organização Mundial de Saúde, foram estimados 9,4 milhões de novos casos de TB em 2008, dos quais 1,4 milhões eram HIV - positivos, e com 1,8 milhões de mortes - o equivalente a 4.500 mortes por dia. Fatores como migração, privação sócio-econômica, co-infecção TB-HIV e o aparecimento de cepas resistentes contribuíram para o aumento do número de casos de TB no mundo, principalmente nos países onde a TB já foi considerada erradicada, e criaram a necessidade do desenvolvimento de novas terapêuticas. Alvos moleculares específicos, que são essenciais para o patógeno, e ausentes no hospedeiro, como as enzimas da via do ácido chiquímico são alvos atraentes para o desenvolvimento de novas drogas antituberculose. Essa via leva à síntese de compostos aromáticos, como aminoácidos aromáticos, e é encontrada em plantas, fungos, bactérias e parasitas do *phylum Apicomplexa,* mas está ausente em humanos. No ano de 2000, foi comprovada a essencialidade dessa via para a viabilidade do bacilo, tornando todas essas enzimas alvos validados para estudo. A segunda enzima da via, deidroquinato sintase (DHQS), catalisa a conversão de 3-deoxi-*D*-arabino heptulosonato-7-fosfato em 3-deidroquinato, o primeiro composto cíclico. Neste trabalho, são descritos o requerimento de metais divalentes na reação e a determinação do mecanismo cinético da DHQS. Os parâmetros cinéticos verdadeiros foram determinados e, juntamente com os experimentos de ligação, o mecanismo rápido-equilíbrio aleatório foi proposto. O tratamento com EDTA aboliu completamente a atividade de DHQS, sendo que a adição de Co^{+2} e Zn^{+2} levam a recuperação total e parcial da atividade enzimática, respectivamente. O excesso de Zn^{+2} inibe a atividade DHQS, e os dados de ITC indicaram a presença de dois sítios seqüenciais de ligação, o que é consistente com a existência de um sítio secundário inibitório. O protocolo de cristalização foi estabelecido e experimentos em andamento proporcionarão a elucidação da estrutura tridimensional da DHQS, que irá beneficiar tanto o desenho de novos inibidores como uma análise detalhada dos rearranjos do domínio da proteína. Em conjunto, estes resultados representam um passo essencial para o desenho racional de inibidores específicos que podem fornecer uma alternativa promissora para um novo, eficaz, e mais curto de tratamento para TB.

Palavras-chave: deidroquinato sintase, caracterização cinética, titulação calorimétrica isotérmica, metaloenzima, *Mycobacterium tuberculosis,* via do ácido chiquímico, tuberculose.

2. **Abstract**

Although the estimated per capita tuberculosis (TB) incidence was stable in 2005, the number of new cases arising each year is still increasing globally. According with World Health Organization, there were estimated 9.4 million new TB cases in 2008, from which 1.4 million were HIV-positive, with 1.8 million deaths total – equal to 4500 deaths a day. Migration, socio-economic deprivation, HIV co-infection and the emergence of extensively-resistance strains, have all contributed to the increasing number of TB cases worldwide, mainly in countries where it was once considered eradicated, and have created an urgent need for the development of new therapeutics against TB. Specific molecular targets, that are essential to the pathogen, and absent in the host, like the enzymes of the shikimate pathway, are attractive targets to development of new antitubercular drugs. This pathway leads to the biosynthesis of aromatic compounds, including aromatic amino acids and it is found in plant, fungi, bacteria and *Apicomplexa* parasites, but is absent in humans. In 2000, this pathway was proved to be essential to the viability of the pathogen, which validates all its enzymes as potential targets. The second enzyme of this pathway, dehydroquinate synthase (DHQS), catalyzes the conversion of 3-deoxy-D-arabino-heptulosonate 7-phosphate in 3-dehydroquinate, the first cyclic compound. In this work, we described the metal requirement and kinetic mechanism determination of the dehydroquinate synthase. The determination of the true kinetic parameters was performed, and, in addition to ligand binding experiments, the rapid-equilibrium random mechanism was determined. The treatment with EDTA abolished completely the activity of DHQS, and the addition of Co^{+2} and Zn^{+2} leads to full and partial recovery of enzyme activity, respectively. Excess of Zn^{+2} inhibits the DHQS activity, and the ITC data revealed two sequential binding sites, which is consistent with the existence of a secondary inhibitory site. The crystallization protocol was established and ongoing experiments will provide the three-dimensional structure of *mt*DHQS, which will benefit both the design of novel inhibitors as well as detailed analysis of domain rearrangements of protein. Taken together, these results represent an essential step for the rational design of specific inhibitors that can provide a promising alternative to a new, effective, and shorter treatment for TB.

Keywords: dehydroquinate synthase, isothermal titration calorimetry, kinetic characterization, metalloenzyme, *Mycobacterium tuberculosis*, shikimate pathway, tuberculosis.

3. **Lista de Abreviaturas**

DAHP	3-deoxi-*D-arabino*-heptulosonato 7-fosfato
*mt*DHQS	deidroquinato sintase de *Mycobacterium tuberculosis*
DOTS	Terapia Diretamente Observada de Curto Prazo (do inglês Directly Observed Treatment Short Course)
DOIS	2-deoxi-scillo-inosose sintase
DQS	3-dehidroquinato
EDTA	ácido etilenodiamino tetra-acético (do inglês ethylenediamine tetraacetic acid)
EPTB	Tuberculose extrapulmonar
ITC	titulação calorimétrica isotérmica (do inglês Isothermal Titration Calorimetry)
K	constante de Michaelis
K_{cat}	constante catalítica
MDR-TB	linhagem de *M. tuberculosis* multirresistente a drogas
NAD^+	nicotinamida adenina dinucleotídeo
NADH	nicotinamida adenina dinucleotídeo forma reduzida
P_i	fosfato inorgânico
TB	tuberculose
V	velocidade máxima
XDR-TB	linhagem de *M. tuberculosis* extensivamente resistente a drogas
WHO	Organização Mundial da Saúde (do inglês World Health Organization)

4. **Lista de Tabelas**

5. **Lista de Figuras**

6. **Introdução**

A tuberculose (TB) é a causa de morte mais comum entre as doenças transmitidas por um único agente infeccioso, *Mycobacterium tuberculosis*, entre adultos, sendo responsável por altos índices de morbidade e mortalidade na população adulta em todo mundo, principalmente em países em desenvolvimento (Bloom e Murray, 1992; Corbett *et al.,* 2003). Essa calamidade é resultado de anos de negligência da epidemia pela comunidade de saúde internacional, particularmente nos países industrializados, onde a doença reapareceu como um problema público de saúde (Raviglione, 2003).

Apesar da incidência *per capita* de casos de TB estimada manter-se estável em 2005, o número de novos casos que surgem a cada ano continua aumentando em nível mundial, principalmente na África, Mediterrâneo Oriental e Sudeste Asiático (WHO, 2010). Entre os 15 países com as maiores taxas de incidência estimadas de TB, 13 são africanos, um fenômeno que pode ser relacionado com as altas taxas de co-infecção com HIV. De acordo com a Organização Mundial da Saúde, ocorreram 9,27 milhões de novos casos de TB no ano de 2007, dos quais 1,37 milhões (14%) eram HIV - positivos (WHO, 2009). A Figura 1 ilustra a incidência da TB no ano de 2007.

A pandemia do HIV alterou o cenário epidemiológico mundial, aumentando dramaticamente os casos de baciloscopia negativa e TB extrapulmonar, atualmente correspondendo à cerca de metade dos casos notificados de TB (WHO, 2008). A tuberculose é a principal causa de morte entre pessoas vivendo com HIV na África e a maior causa de morte em outros lugares, sendo responsável por 23% das 2 milhões de mortes de HIV em 2007 (cerca de 460 mil mortes) (WHO, 2009). Também é a doença mais

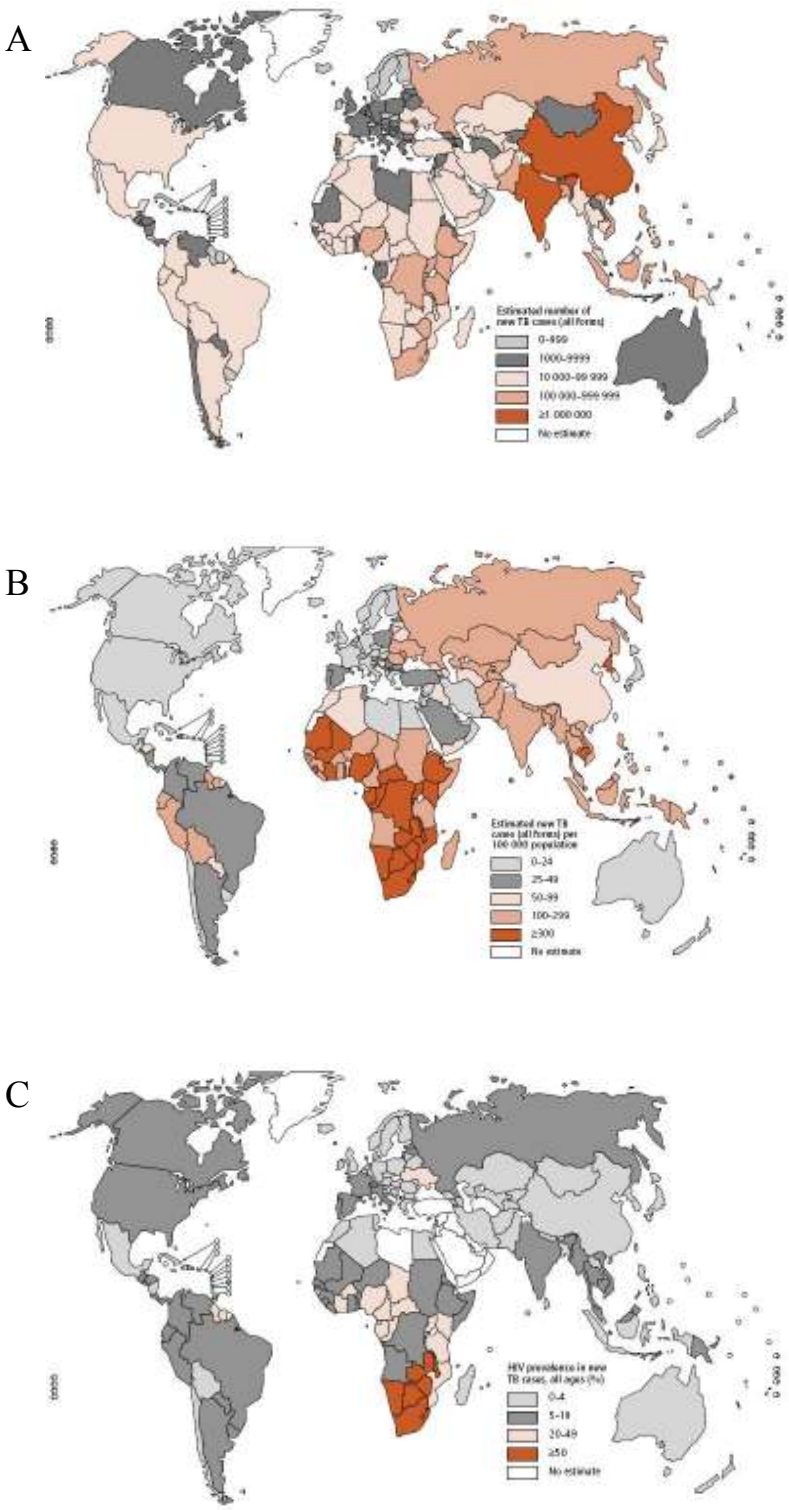


Figura 1: Os mapas acima mostram a incidência da tuberculose no mundo no ano de 2007. (A) Novos casos de TB em 2007, (B) Número total de casos de TB em 2007 e (C) Prevalência do HIV entre os novos casos de TB (WHO, 2009).

comum presente entre pessoas vivendo com HIV, que estão em tratamento com antiretrovirais. Estima-se a ocorrência de 1,37 milhões de casos de pacientes HIV positivos com TB no ano de 2007. Cerca de 80% desses pacientes vivem na África Subsaariana (WHO, 2009a).

A emergência de linhagens de *M. tuberculosis* multirresistentes a drogas (MDR-TB), resistentes a isoniazida e rifampicina, principais fármacos de primeira-linha utilizados no tratamento da TB, agravaram a situação da TB no mundo (Bloom e Murray, 1992). Em 2000, foi relatada a ocorrência de "superstrains", linhagens resistentes a pelo menos três dos quatro fármacos de primeira-linha, em um número significativo de casos (79% dos casos de MDR-TB) (WHO, 2004). Mais recentemente, foi relatada a ocorrência de linhagens MDR-TB resistentes a qualquer fluorquinolona e pelo menos uma das três drogas de segunda linha injetáveis usadas no tratamento da TB, as quais foram denominadas linhagens de *M. tuberculosis* extensivamente resistentes a drogas (XDR-TB) e já ditas virtualmente incuráveis (CDC, 2006). Até março de 2010, 58 países reportaram a ocorrência de pelo menos um caso de XDR-TB, caracterizando uma disseminação global e alarmando as autoridades públicas de saúde, conforme mostra a Figura 2 (WHO, 2010b).

Dessa forma, até 2020, estima-se que mais de um bilhão de pessoas serão infectadas com o bacilo, 200 milhões irão desenvolver a doença e 70 milhões morrerão caso as estratégias de controle não sejam otimizadas (Pasqualoto e Ferreira, 2001). Se a taxa de detecção e de tratamento de casos de MDR-TB aumentar sem que aumentem as taxas de cura de casos MDR-TB, os casos de XDR-TB podem aumentar exponencialmente e rapidamente se tornarem incontroláveis (Blower e Supervie, 2007). A rápida e quase completa mortalidade nessa população (pacientes XDR-TB) mostra a gravidade das

potenciais conseqüências da transmissão de linhagens de tuberculose resistentes a drogas em regiões de alta prevalência do HIV (Ganghi *et al.*, 2006).

Figura 2: Países que reportaram pelo menos um caso de XDR-TB até março de 2010 (WHO, 2010b).

Fatores como migração, privação socioeconômica, a ocorrência e desenvolvimento da resistência, co-infecção TB-HIV e o uso de terapias imunossupressoras têm contribuído para o aumento do número de casos de TB no mundo, principalmente em países onde essa doença era considerada erradicada (Goodman e Lipman, 2008). Dessa forma, é necessário o desenvolvimento de novas drogas antimicobacterianas, porém nenhum novo composto foi introduzido no tratamento da TB nos últimos 50 anos. Apesar da TB ser curada com a atual terapia preconizada, os seis meses necessários tornam o tratamento da doença muito longo e freqüentemente com a ocorrência de significativa toxicidade. Dessa forma, ocorre um abandono precoce do tratamento pelo paciente (Zhang, 2005).

Todos os fatos acima citados tornam evidente a importância do desenvolvimento de novas drogas que inibam alvos diferentes daqueles abordados pelas atuais drogas utilizadas na terapia e que reduzam o tempo de tratamento. Para evitar toxicidade significativa, os

alvos de inibição devem estar presentes no patógeno mas ausente no hospedeiro humano (Zhang, 2005).

6.1 O agente etiológico

O *M. tuberculosis* é considerado o principal agente etiológico da TB, sendo considerado o patógeno mais bem sucedido do mundo. O bacilo pode persistir em tecidos do hospedeiro, estabelecendo assim uma infecção latente, que pode ser reativada em momento onde o sistema imune do hospedeiro está comprometido (Gomez e McKinney, 2004; Hingley-Wilson *et al.,* 2006).

O *M. tuberculosis* é classificado como uma micobactéria com forma de bastonete delgado, aeróbico, não-formador de esporos e parasita obrigatório, descoberto por Robert Koch em 1882. Características como coloração álcool-ácido resistente, resistência a antimicrobianos e sua patogenicidade estão relacionadas com seu envelope celular diferenciado, o qual possui uma camada externa de ácidos micólicos, que formam uma camada cérea, resistente à água, e que impede a penetração de diversos antibióticos (Tortora *et al.,* 2000).

6.2 Transmissão e Patogênese

A TB é uma doença infecto-contagiosa, mais comumente adquirida inalando o bacilo da TB. A gotícula contendo *M. tuberculosis* pode ser lançada no ar por uma pessoa com a infecção ativa ao falar, tossir ou espirrar e, esta pode permanecer suspensa no ar por diversas horas (NSB Editorial Comment, 2000).

Os aerossóis lançados no ar alcançam os pulmões, onde geralmente são fagocitados pelos macrófagos nos alvéolos. Os macrófagos ativam-se pela presença dos bacilos e, em

geral, os destroem. Se os bacilos não forem destruídos, os fagócitos tendem a protegê-los dos anticorpos, permitindo assim que muitos bacilos sobrevivam e se multipliquem. Os macrófagos e outras células defensivas se acumulam no sítio de infecção, formando uma lesão fechada denominada *tubérculo*. Em sua maioria, os macrófagos presentes ao redor não são bem sucedidos em destruir as bactérias, mas liberam enzimas e citocinas que causam uma inflamação, lesivas ao pulmão. Após semanas, o interior do tubérculo torna-se caseoso, e os bacilos, crescem lentamente nessa localidade, e podem ficar dormentes durante anos, até uma possível reativação. Se as defesas do corpo são efetivas e interrompem a doença nesse ponto, as lesões caseosas lentamente cicatrizam e calcificam, sendo facilmente visualizadas nos raios X (Tortora, 2000).

Em 95% dos casos, o sistema imune do indivíduo suprime a replicação do bacilo antes dos sintomas aparecerem. De cada dez pessoas que entram em contato com o bacilo, apenas uma irá desenvolver a doença ativa durante a sua vida, ainda que algumas possam manter o bacilo em sua forma latente, caracterizando um grande e preocupante reservatório do bacilo pelo mundo (Gomez e McKinney, 2004).

6.3 Sintomas

Na TB ativa, o paciente pode ter sintomas gerais como mal-estar e tosse. Sempre que uma tosse com catarro persistir por três semanas ou mais, deve-se suspeitar de TB. Tosse, catarro amarelo, esverdeado ou até com sangue (hemoptise), dores no peito, falta de apetite, emagrecimento, febre, suores noturnos, indisposição e mal-estar são os sintomas relacionados à TB (Daniel, 2006).

6.4 TB extrapulmonar

Além da tuberculose pulmonar, apresentação clínica mais comum da doença, a tuberculose também pode infectar outros órgãos do corpo, sendo chamada de tuberculose extrapulmonar (EPTB) (Fuentes e Caminero, 2006). Cerca de 15% de pacientes com TB ativa apresentam EPTB, que normalmente resulta do crescimento excessivo da bactéria e sua disseminação pela corrente sanguínea (Ducati *et al.,* 2006).

O número desses casos relatados tem aumentado nesses anos, tanto em países industrializados como em países em desenvolvimento, principalmente em regiões em que há alta prevalência do HIV. Como os sintomas não são específicos, há uma grande dificuldade no diagnóstico da EPTB, sendo a histopatologia ainda o principal método diagnóstico (Fuentes e Caminero, 2006).

6.5 Tratamento

A TB é uma doença curável em praticamente 100% dos casos novos diagnosticados. O tratamento recomendado pela OMS consiste na administração combinada de isoniazida (INH), rifampicina (RIF), pirazinamida (PZA) e etambutol (ETB), durante os primeiros dois meses de tratamento, seguidos pela administração de INH e RIF por pelo menos quatro meses adicionais, sendo a falha no tratamento a principal causa dos altos índices de mortalidade da TB (Duncan, 2003). Porém, cinco mil mortes ocorrem no mundo por dia causadas pela TB (WHO, 2007).

O descobrimento da estreptomicina em 1944 inaugurou a época de ouro no desenvolvimento de fármacos anti-TB. Até a década de 70, foram descobertos/sintetizados os fármacos que são utilizados até hoje no tratamento da TB: ácido para-aminossalicílico (1946), INH (1952), cicloserina (1955), canamicina (1957), RIF (1965), etionamida (1966),

ETB (1968) e PZA (1970) (Duncan, 2003). A utilização da estreptomicina na terapêutica determinou uma melhora inicial nos pacientes, mas após poucos meses estes apresentavam recaídas. O ácido para-aminossalicílico, juntamente com a estreptomicina, produziu uma melhora no tratamento da TB. Com a utilização da isoniazida em 1952, foi possível um tratamento com os três medicamentos combinados (Lemos, 1994) revolucionando a terapia da TB e o prognóstico dos pacientes. O uso combinado desses antibióticos como tratamento padrão adotado no mundo ocidental, a partir de 1955, diminuiu consideravelmente as taxas de mortalidade da TB (Janin, 2007). A estrutura dos fármacos da época de ouro é mostrada na Figura 3.

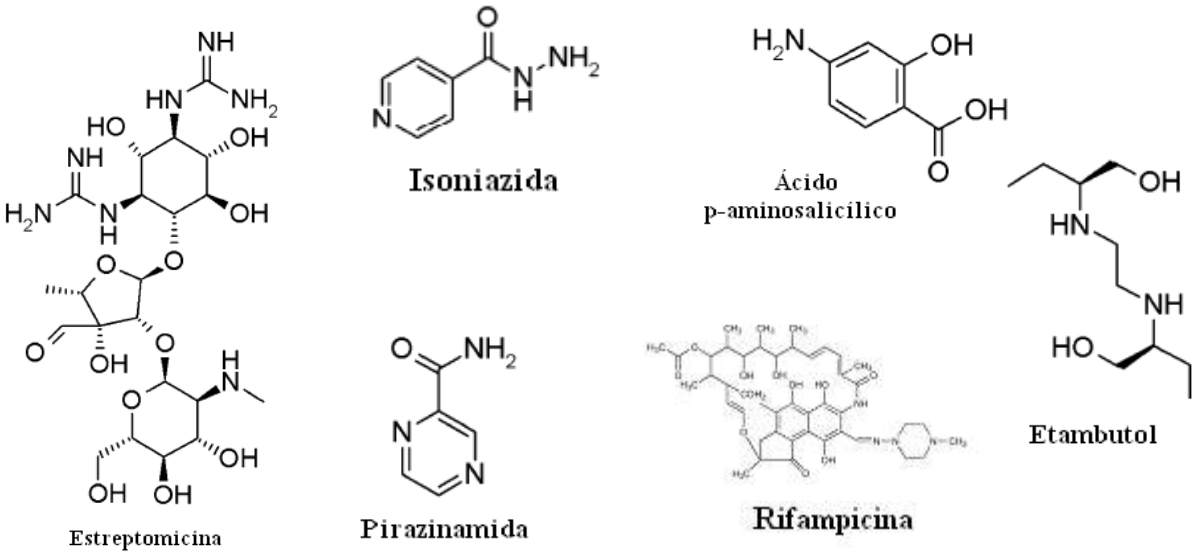


Figura 3: Fármacos da época de ouro de desenvolvimento de antituberculosos.

O alto índice de abandono no tratamento pelos indesejáveis efeitos adversos resultantes da quimioterapia fez a Organização Mundial da Saúde implementar a DOTS: Terapia Diretamente Observada de Curto Prazo, um programa no qual agentes de saúde domiciliares aconselham os pacientes, avaliando o progresso na terapia e certificando-se que a medicação está sendo corretamente tomada pelo paciente. De sua implementação até 2005, a DOTS pode ser observada em 182 países, cobrindo cerca de 77% da população mundial, levando a um aumento de 11% em 1995 para 45% em 2003 nas taxas de detecção de novos casos de TB (Dye *et al.,* 2005).

Nos casos de resistência, o tratamento é mais longo, geralmente 18 a 24 meses, e utiliza fármacos de segunda ou terceira linhas, os quais apresentam maior toxicidade em relação ao tratamento de escolha, além sempre dos fármacos de primeira linha que ainda se observa susceptibilidade (Mahmoudi e Iseman, 1993; Sharma e Mohan, 2004).

6.6 Busca de novas drogas

Desde a década de 1970, nenhum fármaco foi introduzido no tratamento, e no atual regime, apesar de eficaz, ainda morrem cinco mil pessoas por dia no mundo de TB (WHO, 2007). Muitos pacientes abandonam o tratamento principalmente devido aos efeitos adversos ou quando os sintomas da doença desaparecem pela ação do tratamento, julgando-se erroneamente curados (Duncan, 2003). Esse quadro leva a uma alta taxa de falência do tratamento e promove a emergência de linhagens do bacilo resistentes aos fármacos do tratamento.

Assim, a partir de todos os fatos citados acima, evidencia-se a necessidade de novos fármacos para o tratamento da TB que diminuam o tempo do tratamento, que sejam efetivos no tratamento da MDR/XDR-TB e promovam um melhor tratamento para TB

latente e extrapulmonar. Fica claro que, reduzindo o tempo de duração do tratamento, as taxas de adesão aumentarão, os efeitos adversos irão reduzir e o custo de implantação da DOTS será muito menor, prevenindo de forma eficiente o surgimento de novas linhagens resistentes ou expansão das linhagens existentes.

O seqüenciamento completo do genoma de *M. tuberculosis* H37Rv em 1998 proporcionou um melhor entendimento da biologia desse patógeno e também identificou potenciais alvos e vias metabólicas que podem ser de utilidade em intervenções profiláticas e terapêuticas (Cole *et al.,* 1998).

6.6.1 Via do Ácido Chiquímico

Alvos específicos que são indispensáveis para o patógeno e não estão presentes no hospedeiro, como as enzimas da via do ácido chiquímico, são atrativos para o desenvolvimento de novos antimicobacterianos. Em 2002, Parish e Stoker comprovaram que essa via é essencial para a viabilidade do bacilo, validando todas as sete enzimas da rota como potenciais alvos para o desenvolvimento de inibidores específicos com características promissoras de drogas anti-TB (Parish e Stoker, 2002).

A via do ácido chiquímico é composta por sete passos enzimáticos, que convertem fosfoenolpiruvato e eritrose-4-fosfato em corismato, o composto precursor de importantes estruturas aromáticas como os aminoácidos aromáticos triptofano, tirosina e fenilalanina, além de compostos como micobactinas, menaquinonas e naftoquinonas. Está presente em algas, plantas, bactérias e fungos, mas está ausente em mamíferos (Bentley, 1990), o que torna suas enzimas alvos validados para a caracterização e desenvolvimento de inibidores (Figura 4).

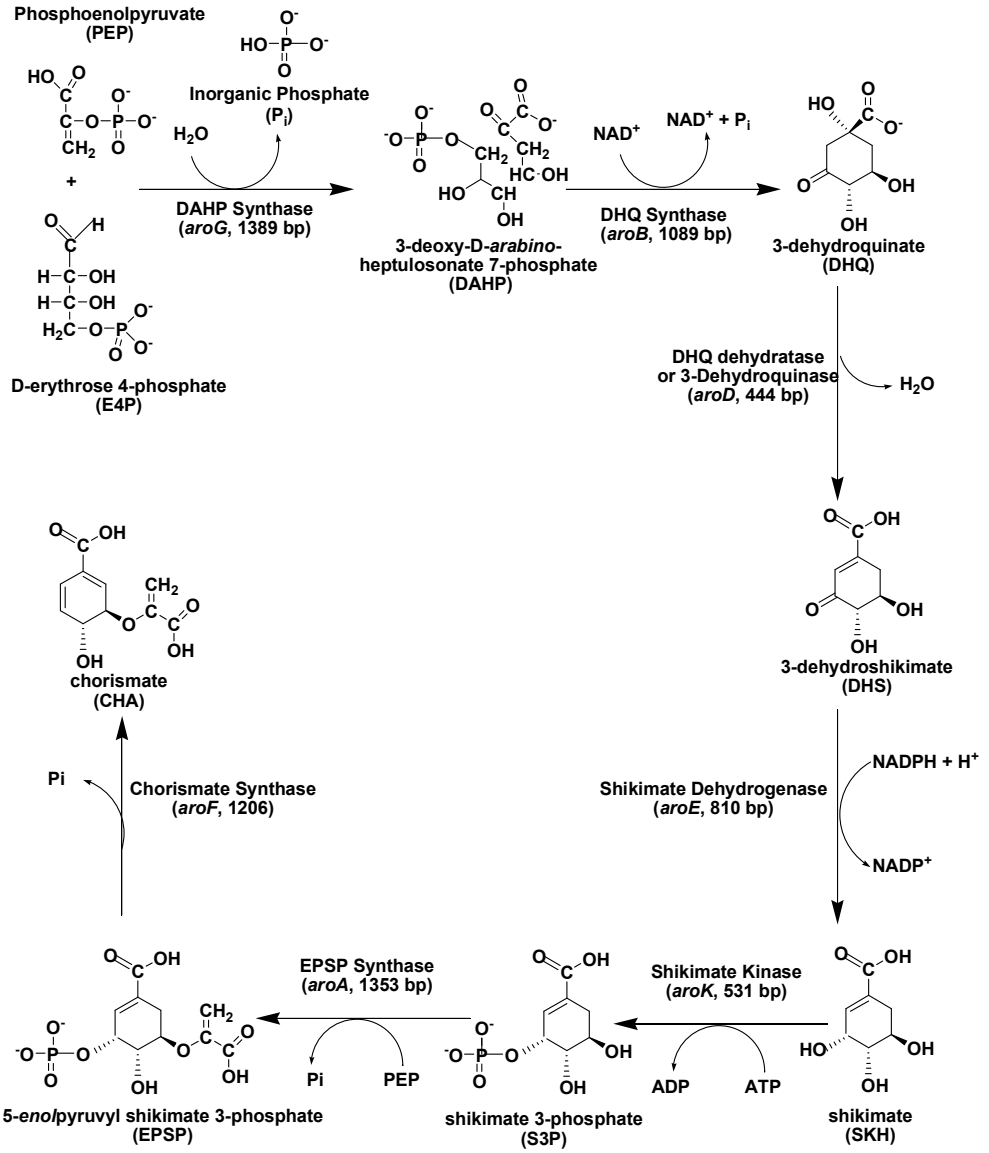


Figura 4: A via do ácido chiquímico.

6.6.2 A enzima deidroquinato sintase

A segunda enzima da via, deidroquinato sintase (DHQS), codificada pelo gene *aroB*, catalisa a conversão de 3-deoxy-D-arabino-heptulosonato-7-fosfato em 3-deidroquinato, primeiro composto cíclico da via. A reação catalisada pela DHQS envolve um mecanismo complexo incluindo reações de oxidação, β-eliminação, condensação aldólica intramolecular e redução, com participação de NAD^+ (Srinivasan *et al.,* 1963). Seu

mecanismo tem sido proposto como similar ao da enzima 2-deoxi-scillo-inosose sintase (DOIS), participante da síntese de antibióticos da classe dos aminoglicosídeos. Além das semelhanças do mecanismo, como a necessidade de metais divalentes para a atividade e a participação do NAD^+ como coenzima, o gene codificante da DOIS apresenta significante similaridade com a seqüência da DHQS, assim como do sitio catalítico e dos resíduos de ligação ao metal (Kudo et al, 1999; Kudo et al, 1999a).

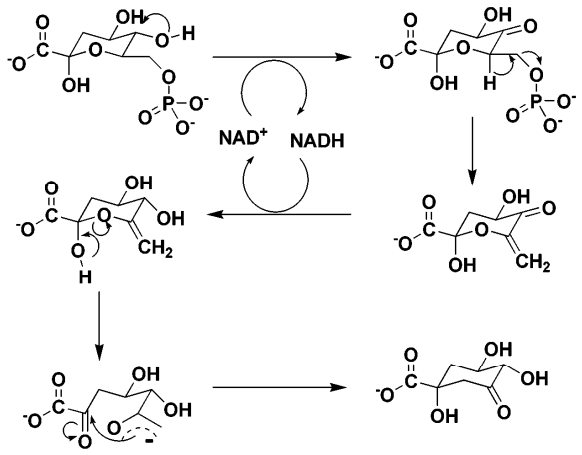


Figura 5: Reação catalisada pela DHQS.

A enzima de *M. tuberculosis* tem sido citada como um alvo importante e validado para o estudo e caracterização com objetivo de desenvolver futuros inibidores que possam futuramente ser introduzidos na terapia da TB por algumas Bases de Dados Virtuais, que visam facilitar a identificação e priorização de candidatos a alvos de drogas em microrganismos patogênicos como o *M. tuberculosis*. Como exemplo, pode-se citar o TDR Target Database, que classificou *mt*DHQS na lista top-50 (Agüero *et al.,* 2008), e o target-TB, que colocou a DHQS entre os 187 alvos identificados na classe de metabolismo intermediário e respiração (Raman *et al.,* 2008).

O conhecimento da estrutura molecular do sítio ativo e do mecanismo de ação da enzima é fundamental para o desenho de inibidores que possam ser empregados no tratamento anti-TB, e os resultados a seguir apresentados a cerca do mecanismo cinético da reação catalisada pela DHQS e do requerimento de cofatores para a atividade máxima dessa enzima representam um passo importante para o alcance desse objetivo final.

7. **Objetivos**

OBJETIVO GERAL:

Elucidação dos mecanismos cinético e químico, assim como a estrutura tridimensional, da enzima deidroquinato sintase para o desenho racional de inibidores baseados no estado de transição da enzima, que são estudos essenciais para o desenvolvimento de inibidores específicos que poderão ser usados no tratamento da TB.

OBJETIVOS ESPECÍFICOS

1- Ensaios espectrofotométricos e espectrofluorimétricos para a determinação das constantes cinéticas em estado pré-estacionário e estacionário.

2- Determinação da estrutura tridimensional da enzima deidroquinato sintase por cristalografia e/ou modelagem molecular.

PARTE II

Analysis of the metal requirement and kinetic mechanism determination of dehydroquinate synthase from *Mycobacterium tuberculosis* H37Rv: a crucial step in the rational design of anti-TB drugs.

Artigo a ser submetido para o periódico Molecular Biosystems, em 2010.

Analysis of the metal requirement and kinetic mechanism determination of dehydroquinate synthase from *Mycobacterium tuberculosis* H37Rv: a crucial step in the rational design of anti-TB drugs.

Jordana Dutra de Mendonça[a,b], Osao Adachi[c], Luiz A. Basso[a*], Diogenes S. Santos[a*]

[a]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), 6681/92-A Av. Ipiranga, 90619-900, Porto Alegre, RS, Brazil.

[b]Programa de Pós-Graduação em Ciências Biológicas: Bioquímica, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.

[c]Department of Biological Chemistry, Faculty of Agriculture, Yamaguchi University, Yamaguchi 753-8515, Japan

[*]Corresponding authors. Telephone/Fax: +55-51-33203629.

E-mail addresses: luiz.basso@pucrs.br (Luiz A. Basso); diogenes@pucrs.br (Diogenes S. Santos).

Running title: Kinetic mechanism of *mt*DHQS

Summary

Although the estimated per capita tuberculosis (TB) incidence was stable in 2005, the number of new cases arising each year is still increasing globally. Migration, socio-economic deprivation, HIV co-infection and the emergence of extensively-resistance strains, have all contributed to the increasing number of TB cases worldwide, mainly in countries where it was once considered eradicated. Specific molecular targets, that are essential to the pathogen, and absent in the host, like the enzymes of the shikimate pathway, are atractive targets to development of new antitubercular drugs. In this work, we described the metal requirement and kinetic mechanism determination of the dehydroquinate synthase. The determination of the true kinetic parameters was performed, and, in addition to ligand binding experiments, the rapid-equilibrium random mechanism was determined. The treatment with EDTA abolished completely the activity of DHQS, and the addition of Co^{+2} and Zn^{+2} leads to full and partial recovery of enzyme activity, respectively. Excess of Zn^{+2} inhibited the DHQS activity, and the ITC data revealed two sequential binding sites, which is consistent with the existence of a secondary inhibitory site. These results represent an important step for the rational design of potent DHQS inhibitors that can be further used as anti-TB drugs.

Introduction

Although the estimated per capita tuberculosis (TB) incidence was stable in 2005, the number of new cases arising each year is still increasing globally.[1] Among the 15 countries with the highest estimated TB incidence rates, 13 are located in Africa, a phenomenon linked to high rates of HIV co-infection, making TB the major cause of death in HIV-positive patients, which represented 23% of the estimated 2 million HIV deaths in 2007. According with World Health Organization, there were estimated 9.4 million new TB cases in 2008, from which 1.4 million were HIV-positive, with 1.8 million deaths total – equal to 4500 deaths a day.[2] Migration, socio-economic deprivation, HIV co-infection and the greater use of immunosuppressive agents in healthcare have all contributed to the increasing number of TB cases worldwide, mainly in countries where it was considered eradicated.[3] Furthermore, the recent report of the occurrence of at least one case of extensively-resistant drug TB – resistant to isoniazid and rifampicin plus resistance to any fluorquinolone and at least one of three injectable second line drugs used in TB treatment – in 57 countries worldwide, has characterized a global spread and alarmed the public health authorities.[4,5]

TB can be cured with the current therapy, but the long, six months required treatment, along with its highly significant toxicity, leads to patients' non adherence.[6] There is an urgent need for the development of both better vaccines and new and more efficient antimycobacterial agents with novel mechanisms of action. Future drugs should present selective toxicity, be active against drug-resistant and non-resistant strains, shorten short-

course treatment duration in order to improve patient´s compliance and, ideally, do not have pharmacological interactions with antiretroviral drugs commonly used to treat HIV.[7,8]

The complete sequencing *Mycobacterium tuberculosis* H37Rv genome has provided a better understanding of the biology of this pathogen and has also identified potential targets and biochemical pathways that may be used in prophylactic and therapeutic interventions.[9] Specific molecular targets that are indispensable to the pathogen and absent in the host, like the enzymes of the shikimate pathway, may be suitable to development of new antitubercular drugs. In 2002, Parish and Stoker proved that this pathway is essential to the viability of the bacilli, validating all such enzymes as potential targets to development of specific inhibitors with promising features of an anti-TB drug.[10]

The shikimate pathway second enzyme, dehydroquinate synthase (aro*B* encoded, DHQS, E.C 4.6.1.3), is responsible by the synthesis of the first cyclic compound (Fig. 1) that will be converted to chorismate, the precursor of aromatic compounds like the amino acids phenylalanine, tyrosine and tryptophan.[11].DHQS has been characterized in other microorganisms (*E. coli*, *B. circulans*, *N. crassa* and *A. nidulans*) and higher plants (*P. mungo and Sorghum*) as a metalloenzyme with preference for Co^{+2} and Zn^{+2}, that requires catalytic amounts of NAD^{+}.[12-17] Pathogenic bacteria witch presented mutations at this point of the pathway are attenuated for virulence.[18] The mechanism of the DHQS reaction appears to closely resemble to that of 2-deoxy-scyllo-inososе synthase (DOIS) in the 2-deoxystreptamine biosynthesis, as well as the protein sequence.[19,20]

*mt*DHQS has been cited as an important and validated target to development of specific inhibitors that could be used in the TB treatment by some virtual databases that aim to facilitate the identification and prioritization of candidate drug targets for pathogens like

M. tuberculosis, like the TDR Targets database, that ranked *mt*DHQS at the top-50 list, and Target-TB, that cited DHQS among 187 identified targets in intermediary metabolism and respiration class.[21,22]

Knowledge of the molecular structure of the active site and of the enzyme mechanism of action will aid the design of useful inhibitors that may be used as antimycobacterial agents. In this work, we describe the kinetic mechanism of reaction catalyzed by DHQS and define the cofactor requirements of this noteworthy enzyme from *M. tuberculosis*, an important step for the rational design of potent DHQS inhibitors that can further be used as anti-TB drugs.

Results

Enzyme assay and Initial velocity

The double-reciprocal plots showed a family of lines intersecting to the left of the *y*-axis (Fig. 2), which is consistent with ternary complex formation and a sequential mechanism.[23] Data were plotted in reciprocal form and fitted to the equation for a sequential initial velocity pattern: $v = VAB/(K_{ia}K_b + K_aB + K_bA + AB)$, where A and B are substrate concentrations, K_a and K_b are Michaelis constants for substrates A and B, respectively. The values of the true steady-state kinetic parameters were k_{cat} = 0.66 (± 0.02) s^{-1}, K_{DAHP} = 7.2 (± 0.5) μM, K_{NAD^+} = 79 (± 6) μM, k_{cat}/K_{DAHP} = 9.2 (± 0.6) x 10^3 $M^{-1}s^{-1}$, and k_{cat}/K_{NAD^+} = 8.4 (± 0.6) x 10^3 $M^{-1}s^{-1}$.

Metal requirement analysis

No improvement in the maximum activity of *mt*DHQS was observed in saturating concentration of DAHP and NAD^+ when non-chelated enzyme was assayed in presence of

different divalent metals, justifying the absence of these compounds in the mixture assay (data not shown).

Determination of the rate of the metal ion loss in the presence of EDTA. The addition of 0.1mM of EDTA to the enzyme in the absence of the substrate in the incubation mixture was capable to abolish the *mt*DHQS activity after 10 minutes on incubation. DAHP and NAD^+ were tested separately for their ability to protect the synthase against inactivation by EDTA. In the presence of saturating levels, DAHP gave a good protection, even in higher concentrations of EDTA (1mM), whereas NAD^+ was ineffective (data not show).

Metal reactivation. The formation of dehydroquinate from DAHP was abolished by chelating agents such as EDTA (0.1 mM) after 10 minutes of incubation and was restored completely by an excess of cobalt salt, and partially restored by zinc, excess of which showed an inhibitory action at 5 mM, Zn^{+2} inhibited 70% of the DHQS activity (data not show). Other divalent metals also supported the reaction to varying degrees (Table 1).

Atomic absorption analysis. Metal concentration analysis by ICP-AES yielded the following results: Ca^{+2}, 0.78 mg L^{-1}; Cd^{+2}, <0.004 mg L^{-1}; Co^{+2}, <0.004 mg L^{-1}; Cu^{+2}, 0.42 mg L^{-1}; Mg^{+2}, <0.001 mg L^{-1}; Mn^{+2}, <0.32 mg L^{-1}; Ni^{+2}, <0.01 mg L^{-1}; and Zn^{+2}, 3.74 mg L^{-1}. These results indicate the presence of 0.2 mol of Zn^{+2} per mol of DHQS.

Isotermal titration calorimetry

In order to determine the relative affinities of the metal binding site of the two main divalent metals responsible for the best activity of *mt*DHQS, Co^{+2} and Zn^{+2} was titrated to chelated enzyme (56 μM). Figure 3 shows the isotherm generated and the data after the

peak integration, subtraction of blank titration data (not shown), concentration normalization, and analysis by Origin 7.0 suite programs.

Data from Co^{+2} titration analysis indicated the presence of two sequential binding site (Fig.3A), where the primary binding site has a K_D of 16 μM and the secondary site has K_D of 781 μM. The overall binding isotherm for the interaction of DHQS with Zn^{+2} is biphasic (Fig. 3B) and it is best fitted to a model of two sequential binding sites. Binding to the first binding site is accompanied by a negative enthalpy change (K_D of 167 μM), whereas binding to the second site (K_D of 0.3 μM) is accompanied by a positive enthalpy change.

Equilibrium ligand binding to *mt*DHQS

The experiments with the intrinsic fluorescence of *mt*DHQS were carried out with objective to determine the order of substrate/product addition/dissociation in the catalytic site and distinguish the kinetic mechanism. The fluorescence of *mt*DHQS decrease upon addition of DAHP binding to the free enzyme, and the plot of concentration versus relative protein fluorescence variation (Fig. 4A) were hyperbolic yielding a dissociation constant of 73 (±7) μM.

No change in intrinsic protein fluorescence could be detected upon binding of NAD^+, which cannot be interpreted as an absence of interaction between NAD^+ and free *mt*DHQS. In order to elucidate this point, a competitive assay with NADH (a noticed inhibitor of DHQS) was performed.[14,24] The titration of NADH in *mt*DHQS solution causes a decrease in the intrinsic fluorescence of the protein, indicating the binding of NADH, with K_D= 215 (± 13) μM (Fig. 4B-inset). In the presence of NADH in concentration near of

K_D, NAD^+ was added and the nicotinamide fluorescence was measured. Plot of NAD^+ concentration versus relative nicotinamide fluorescence variation were hyperbolic, and the data were fitted for competitive equation: $f = FA/[Ka(1+I/Ki) + A]$, yielding a K_i for NAD^+ = 1400 (± 400) μM (Fig. 4B).

Both products, DHQ and Pi, showed changes in the intrinsic protein fluorescence, and the hyperbolic plots yields dissociation constants of 3100 (±135) μM (Fig. 4C) and 920 (± 80) μM (Fig. 4D), respectively.

Discussion

The development of DHQS inhibitors demands a detailed characterization of its catalytic mechanism. In the course of this conversion, which is mechanistically unusually diverse, the synthase apparently catalyzes an oxidation, a β-elimination, an intramolecular aldol condensation, and a reduction. It is impressive the number of steps catalyzed by a single enzyme during each catalyzed turnover of substrate DAHP into DHQ (Fig.1).[12] Because of the complex mechanism on that DHQS is involved, the enzyme has been speculated as a spectator of its own reaction, suggesting its catalysis as a simple oxireductase with several of the reactions occurring spontaneously.[25] However, the arrangement of the active site of DHQS indicates that the enzyme is not just a spectator in catalysis but stabilizes intermediates and prevents side reactions through its entire reaction pathway.[26]

The mechanism of the DHQS reaction is similar to the mechanism described for DOIS in the 2-deoxystreptamine biosynthesis, including the divalent metal ion requirement for the activity, being Co^{+2} the most effective to exert the enzyme activity, and the cyclic reduction and reoxidation of NAD^+ and NADH, respectively, during the enzyme-catalyzed

reaction prior to release of DHQ from the active site.[19] The *btrC* gene shows significant sequence similarity to various DHQS and the catalytic domain, and the metal binding residues are conserved between DOIS and DHQSs.[20]

The divalent metal cofactor has been proposed as an important factor during the catalysis of DHQS. In order to elucidate this role, analysis of the metal requirement of *mt*DHQS was performed. The absent necessity of addition of divalent metal in the standard assay with non-chelated enzyme can be explained by the general conditions of the expression system that provided the appropriate metal amount for the *mt*DHQS maximum activity. The treatment with EDTA results in a rapid formation of inactive apoenzyme, and the DAHP binding prevents the removal of the metal from the active site, which suggests that the metal is less accessible to EDTA when substrate is bound to the active site; conversely, NAD^+ has no protective effect. This protective behavior of DAHP was previously described in the *E. coli* and *N. crassa* enzymes.[12,14]

In the reactivation studies, the addition of Co^{+2} in the reaction mixture containing the chelating agent restored completely the *mt*DHQS activity, in accordance with *B. circulans* and *E. coli* enzymes, but differing from *N. crassa* that rapidly recovered the enzyme´s activity with Zn^{+2} rather than Co^{+2}.[12-14] In contrast, the presence of excess of Zn^{+2} in this case leads to decreased rates, which suggests that zinc may also bind to a second inhibitory site. This is similar to DHQS from *E. coli* and *A. nidulans* and was confirmed by the ITC assays, where the titration of Co^{+2} and Zn^{+2} in a DHQS solution indicated the existence of two sequential binding sites (Fig. 3).[12,27]

The atomic absorption analysis clearly showed the absence of any significant quantity of the cobalt content and the presence of a small, but significant, amount of zinc. The zinc content of the final preparation is 0.2 mol of Zn^{+2} per mol of *mt*DHQS, indicating

that the metal sites of the enzyme were only partially saturated in crude extract. This substoichiometric value obtained may represent the loss of the metal during the purification process, similar to described for other enzymes, like DAHP(Phe) synthase from *E. coli*.[28] It is interesting that the *mt*DHQS isolated enzyme did not contain predominantly cobalt, considering the preference of the enzyme showed in this work on the metal reactivation assays for Co^{+2}, but based on the higher bioavailability of Zn^{+2}, it seems likely that dehydroquinate synthase is naturally a zinc-mettaloenzyme dependent, and this result can be explained.[14] Taken together, these results demonstrate a central role for the metal ion in the catalytic mechanism, since each of chemical transformation mediated by dehydroquinate synthase has, in simpler enzymatic systems, been shown to involve metal cofactors.[12] The recent crystallization protocol obtained by our group from *mt*DHQS, and the structural data resulting from its analysis will allow the better understanding of the metal role in the catalysis (ongoing experiments).

The pattern of the double-reciprocal plots (Fig. 2) was consistent with ternary complex formation and a sequential mechanism. The mechanisms of ping-pong and rapid equilibrium ordered could be discarded, since these mechanisms display parallel lines and intersecting lines at the *y*-axis, respectively.[23] These data from initial velocity experiments in addition with the equilibrium binding analysis are consistent with a rapid-equilibrium random kinetic mechanism, in which both DAHP and NAD^+ bind to the free enzyme, and there is no preferential order for dissociation of the products, DHQ and Pi, from the active site (Fig.5). The true kinetic parameters showed that *mt*DHQS has a lower catalytic efficiency when compared to others DHQSs. The turnover numbers of 24 s^{-1} for *E. coli*, 19 s^{-1} for *N. crassa* and 6.8 s^{-1} for *A. nidulans*, are significantly higher when compared with 0.65 s^{-1} for *mt*DHQS.[14,27,29] The comparison of the K_{DAHP} values showed that they are

similar, with difference less than 5-fold when compared with *E. coli* (5.5 μM), *N. crassa* (1.4 μM) and *A. nidulans* monofunctional DHQS domain (21 μM), although the specificity constant for DAHP of 9.2 x 10^3 $M^{-1}s^{-1}$ showed that *mt*DHQS is less specific than the *E. coli* enzyme (K_{cat}/K_{DAHP} 2.5 X $10^7 M^{-1}$ s^{-1})[12,14,27]. In contrast, for the substrate NAD^+, the difference in K_M values is significantly higher, being almost 1000-fold larger than *E. coli* (K_{NAD}^+ 80nM), 400-fold larger than *N. crassa* (K_{NAD}^+ <0,2 μM) and 40-fold larger than *A. nidulans* monofunctional DHQS domain (K_{NAD}^+ 1.9 μM). [12,14,27]

Conclusions

Rational inhibitor design relies on mechanistic and structural information about the target enzyme, and the determination of kinetic parameters is essential to characterize enzymes of biological interest. ITC has been used as a very sensitive technique for the direct determination of thermodynamic and kinetic parameters of enzymatic reactions, mainly with enzymes that required coupled reaction, and specially when the purpose of the study involves the effect of activators, inhibitors, or stabilizers, and these compounds will affect only the enzyme of interest and not the other used in the coupled reaction.[30] These data in addition with availability of the tree-dimensional structure of *mt*DHQS will benefit both the design of novel inhibitors as well as detailed analysis of domain rearrangements of protein. Such structural information will be vital to obtain a better understanding of the mode of action of a drug at the molecular level, to appreciate the structural consequences of genetic variations, and to construct accurate pharmacophores for drug design.[31]

Experimental

Homogeneous *mt*DHQS

The homogeneous solution of recombinant *mt*DHQS was obtained as previously described in Mendonça et al 2007.[32] Protein concentration was determined with Bradford Protein Assay Kit (Bio-Rad Laboratories), using bovine serum albumin as standard.[33]

Enzyme assay and initial velocity

DHQS enzyme activity was monitored in an UV-2550 UV/Visible spectrophotometer (Shimadzu). DHQS activity was measured by coupling the conversion of 3-*deoxy*-D-arabino-heptulononate 7-phosphate (DAHP) into dehydroquinate (DHQ) to the dehydroquinate dehydratase-mediated conversion of DHQ into dehydroshikimate, and measuring the production of dehydroshikimate at 234 nm (ε_{234nm} = 12000 $M^{-1}cm^{-1}$).[29,34] The standard assay was performed in 5-mm pathlength quartz cuvettes at 25°C in a total volume of 1.5mL containing 50 mM Tris-HCl pH 7.5, 120 μM DAHP (Toronto Chemical Research), 600μM NAD^+ (Sigma), and 1 unit of dehydroquinate dehydratase (DHase) from *M. tuberculosis*. The true steady-state kinetic parameters were determined by varying concentrations of DAHP (3-50μM) against several fixed concentrations of NAD^+ (40-400μM). One unit of enzyme activity catalyzes the consumption of 1μmol of DAHP/min at 25°C.

Metal requirement analysis

Divalent metal ions were removed from all solutions by stirring with Chelex resin (BioRad), followed by filtration through a Milipore 0.22 μM sterilizing membrane.

Determination of the rate of the metal ion loss in the presence of EDTA. In an assay at 25°C, DHQS (30 μL of a solution of 35 μM) was added to 50 mM Tris-HCl buffer,

pH 7.5, containing NAD^+ (600μM) and EDTA (100 μM), in a final volume of 1.46mL. The reaction was initiated by the addition of DAHP (120μM) and DHase (1 unit), and the initial rate determined at different time reaction.

Metal reactivation. To study the requirement of divalent metal ion, each metal chloride including $ZnCl_2$, $MgCl_2.6H_2O$, $3CdSO_4.8H_2O$, $BaCl_2.2H_2O$, $CaCl_2.2\ H_2O$, $CoCl_2.6H_2O$, $MnCl_2.4H_2O$ and $NiCl_2.6H_2O$ (Sigma-Aldrich) was added in a final concentration of 1mM. Metal-free condition was made by adding EDTA to enzyme and NAD^+ (600μM) during 10 minutes at final concentration 0.1mM, and the reaction was started after 1 minute of incubation with the divalent metal by the addition of DAHP (120μM) and DHase (1 unit).

Atomic absorption analysis. Analysis of Ca^{+2}, Cd^{+2}, Co^{+2}, Cu^{+2}, Mg^{+2}, Mn^{+2}, Ni^{+2}, and Zn^{+2} concentrations of DHQS homogeneous solution in the final purification process were carried out by ICP-AES (Spectro Ciros CCD). All measures were in duplicate. DHQS was in Tris-HCl 50mM pH 7.6 in a final concentration of 13.25 mg/mL.

Isotermal titration calorimetry

The protein binding constants of protein/divalent metal were determined by ITC using a MicroCal ITC-200 microcalorimeter (Thermo). ITC measurements were carried out at 25^oC, and titrations were carried out using a 40μL-syringe and with stirring at 500rpm. Each titration consisted of a preliminary injection of 0.5 μL, followed by 19 injections of 2 μL into a cell containing 200μL of protein sample of 56 μM. To correct for dilution and mixture effects, a series of baselines were performed, where injections of metal were

carried out into buffer. Data were analyzed using MicroCal Origin 7.0 software and baseline subtracted from data to obtain accurate heat exchanges.

Equilibrium ligand binding to *mt*DHQS

Fluorescence measurements were carried out in a RF-5301 PC Spectrofluorophotometer (Shimadzu) at 25°C. Measurements of intrinsic DHQS protein fluorescence employed excitation wavelength at 300 nm in each binding experiment, and the emission wavelength ranged from 320 to 450 nm. For assays involving NADH and NAD^+, the nicotinamide fluorescence was monitored, with excitation wavelength at 300 nm and the emission wavelength ranged from 320 to 500 nm. The slits for excitation and emission were 1.5 and 15 nm, respectively. Fluorescence titrations of binary complex formation were carried out by making microliter additions of the following compounds to 2 mL containing 2 μM DHQS: 4 mM DAHP stock solution (1.999 – 181.34 μM final concentration); 200 mM NAD^+ stock solution (0.0989 – 4.877 mM final concentration); 500 mM DHQ solution (0.249 – 7.634 mM final concentration); 500 mM Pi stock solution (0.249 – 11.00 mM final concentration); 20 mM NADH stock solution (9.995 – 582.52 μM final concentration). Control experiments were employed to both determine the maximum ligand concentrations to be used with no significant inner filter effect and to account for any dilution effect on protein fluorescence.

Acknowledgements

This work was supported by the National Institute of Science and Technology on Tuberculosis (DECIT/SCTIE/MS-MCT-CNPq-FNDCT-CAPES) and the Millennium

Initiative Program (CNPq) to D.S.S. and L.A.B. D.S.S. (CNPq, 304051/1975-06), and L.A.B. (CNPq, 520182/99-5) are Research Career Awardees of the National Research Council of Brazil (CNPq). J.D.M. is the recipient of a MSc scholarship awarded by CNPq.

Notes and references

1 WHO, 2010a. http://www.who.int/mediacentre/factsheets/fs104/en/index.html. Acessed in 01-15-2010**.**

2 WHO, 2009. Global tuberculosis control: a short update to the 2009 report. WHO /HTM/TB/2009.426

3 A. Goodman and M. Lipman, *Tuberculosis Clinical Medicine*, 2008, **8(5)**, 531-534.

4 Centers for Disease Control and Prevention, *MMWR*., 2006, **55(11)**, 301-305.

5 WHO, 2010b. http://www.who.int/tb/challenges/xdr/xdr_map_sep09.pdf. Accessed in 01-21-2010.

6 Y . Zhang, *Annual Rev Pharmacol Toxicol.*, 2005, **45**, 529-564.

7 R. Ducati, L. A. Basso, D. S. Santos, *Current Drug Targets*, 2007, **8**, 423-435.

8 M. Jassal and W.R. Bishai, *The Lancet Infection Disease*, 2009, **9**, 19-30.

9 S. T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. 3rd 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 and B. G. Barrell, *Nature*. 1998, **393**, 537-544

10 T. Parish and N. G. Stoker, *Microbiology*, 2002, **148**, 3069-3077.

11 R. Bentley, *Crit. Rev. Biochem. Mol. Biol.*, 1990, **25**, 307-384.

12 S. L. Bender, S. Mehdi, and J. R. Knowles, *Biochemistry*, 1989, **28**, 7555-7560.

13 N. Hasan and E. W. Nester, *The Journal of Biological Chemistry*, 1978, **253(14)**, 4999-5004.

14 J. M. Lambert, M. R. Boocock and J. R. Coogins, *Biochem. J.*, 1985, **226**, 817-829

15 J. D. Moore, J. R. Coggins, R. Virden, and A. R. Hawkins, *Biochem J.*, 1994, **301**, 297-304.

16 E. Yamamoto, *Phytochemistry*, 1980, **19**, 779-781

17 R. Saijo, and T. Kosuge, *Phytochemistry*, 1978, **17**, 223-225

18 A. Gunel-Ozcan, K. A. Brown, A. G. Allen, and D. J. Maskel, *Microb Pathog*, 1997, 23, 311-316

19 F. Kudo, Y. Hosomi, H. Tamegai and K. Kakinuma, *The Journal of Antibiotics*, 1999a, **52(2)**, 81-88

20 F. Kudo, H. Tamegai, T. Fujiwara, U. Tagami, K. Hirayama and K. Kakinuma. *The Journal of Antibiotics*, 1999b, **52(6)**, 559-571

21 F. Agüero, B. Al-Lazikani, M. Aslett, M. Berriman, F. S. Buckner, R. K. Campbell, S. Carmona, I. M. Carruthers, A. W. Chan, F. Chen, G. J. Crowther, M. A. Doyle, C. Hertz-Fowler, A. L. Hopkins, G. McAllister, S. Nwaka, J. P. Overington, A. Pain, G. V. Paolini,

U. Pieper, S. A. Ralph, A. Riechers, D. S. Roos, A. Sali, D. Shanmugam, T. Suzuki, W. C. Van Voorhis, C. L. Verlinde, *Nat Rev Drug Discov.*, 2008, **7(11)**, 900-907.

22 K. Raman, K. Yeturu and N. Chandra, *BMC Syst Biol.*, 2008, **19**, 2-109.

23 I. H. Segel, in *Enzyme Kinetics, Behavior and analysis of rapid equilibrium and steady-state enzyme systems*, John Wiley & Sons, Inc., New York, 1975, 957p.

24 P. R. Srinivasan, J. Rothschild and B. D. Sprinson, *The Journal of Biological chemistry*, 1963, **238(10)**, 3176-3182.

25 P. A. Barlett and K. Satake, JACS, 1988, **110**, 1628-1630.

26 E. P. Carpenter, A. R. Hawkins, J. W Frost and K. A. Brown, *Nature*, 1998, **394**, 299-302.

27 A. Park, H. K. lamb, C. Nichols, J. D. Moore, K. A. Brown, A. Cooper, I. G. Charles, D. K. Stammers, and A. R. Hawkins, *Protein Science*, 2006, **13**, 2108-2119.

28 C. M. Sthephens and R. Bauerle, *The Journal of Biological Chemistry*, 1991, **266(31)**, 20810-20817.

29 U. S. Maitra and D. B. Sprinson, *J Biol Chem*., 1978, **253(15)**, 5426-30.

30 M. L. Bianconi, *Biophysical Chemistry*, 2007, **126**, 59-64

31 P. J. Whittle and T. L. Blundell, *Annu Rev Biophys Biomol Struct*, 1994, **23**, 349-375.

32 J. D. Mendonça, F. Ely, M. S. Palma, J. Frazzon, L. A. Basso and D. S. Santos, *Journal of Bacteriology*, 2007, **189(17)**, 6246-6252.

33 M. M. Bradford, *Anal. Biochem.,* 1976, **72,** 248-254.

34 S. Mistuhashi and B. D. Davis, *Biochim Biophys Acta.* 1954, **15(1)**, 54-61.

Table 1: Effect of divalent metal ions and EDTA on DHQS activity. Purified enzyme was incubated with 0.1 mM EDTA for 10 minutes, with subsequent addition of 1 mM of the metal ions. The reaction was initiated with DAHP and DHase.

Non-chelated DHQS	100%	Cd^{+2}	24%
None	<1%	Mg^{+2}	15%
Co^{+2}	94%	Mn^{+2}	11%
Zn^{+2}	43%	Ni^{+2}	4%
Ca^{+2}	38%	Ba^{+2}	4%

*Activities relative to the maximum activity of non-chelated DHQS.

Figures legends

Figure 1: Chemical reaction catalyzed by DHQS.

Figure 2: Double-reciprocal plots for DHQS with either DAHP (A) or NAD^+ (B) as the variable substrate. Each curve represents fixed-varying levels of the co-substrate, ranging from 40 to 400 μM for NAD^+ and 3 to 50 μM for DAHP.

Figure 3: ITC analysis of Co^{+2} (A) and Zn^{+2} (B) binding to DHQS. Raw data showing the heat pulses resulting from a titration of metal free DHQS (56 μM) in the calorimetric cell with a 0.5-μL injection of 1.5mM of each metal followed by 19 subsequent 2μL- injections, and the integrated heat pulses, normalized per mol of injectant, giving a differential binding curve that is adequately described by two-site sequential model.

Figure 4: Fluorescence spectroscopy on the equilibrium ligand binding to *mt*DHQS. Dependence of the enhancement in DHQS relative fluorescence upon its binding to the substrate (A) DAHP, and products (C) DHQ and (D) Pi. In (B), the dependence of the NADH-enzyme complex fluorescence in the competitive assay upon the NAD^+ concentration; inset shows the dependence of the enhancement in the nicotinamide relative fluorescence upon its binding to the enzyme.

Figure 5: Proposed kinetic mechanism for *mt*DHQS.

Figure 1

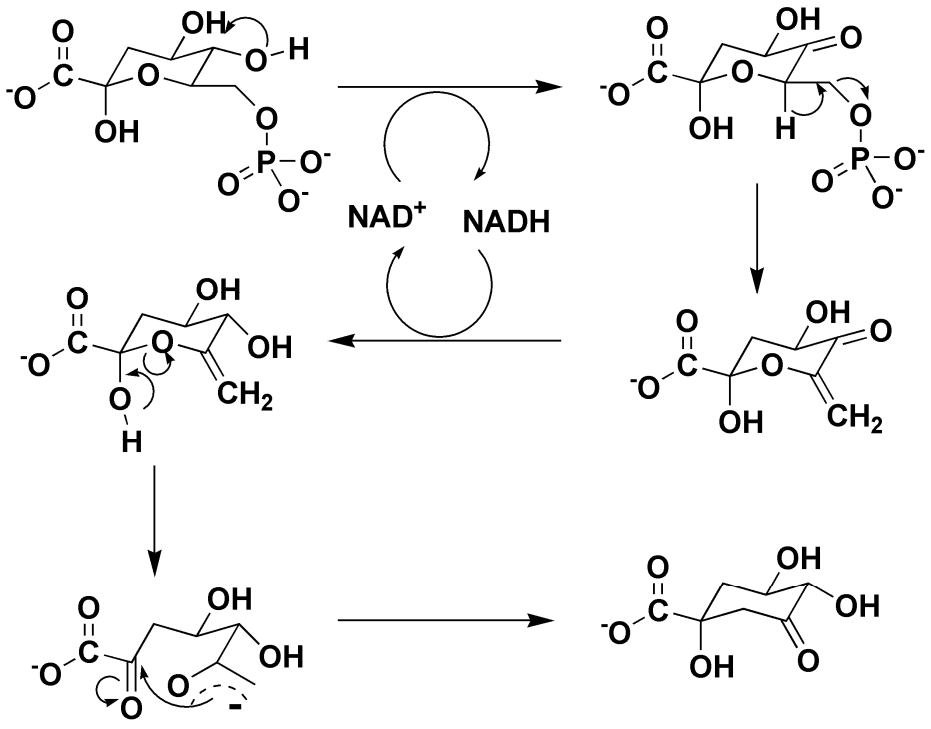


Figure 2

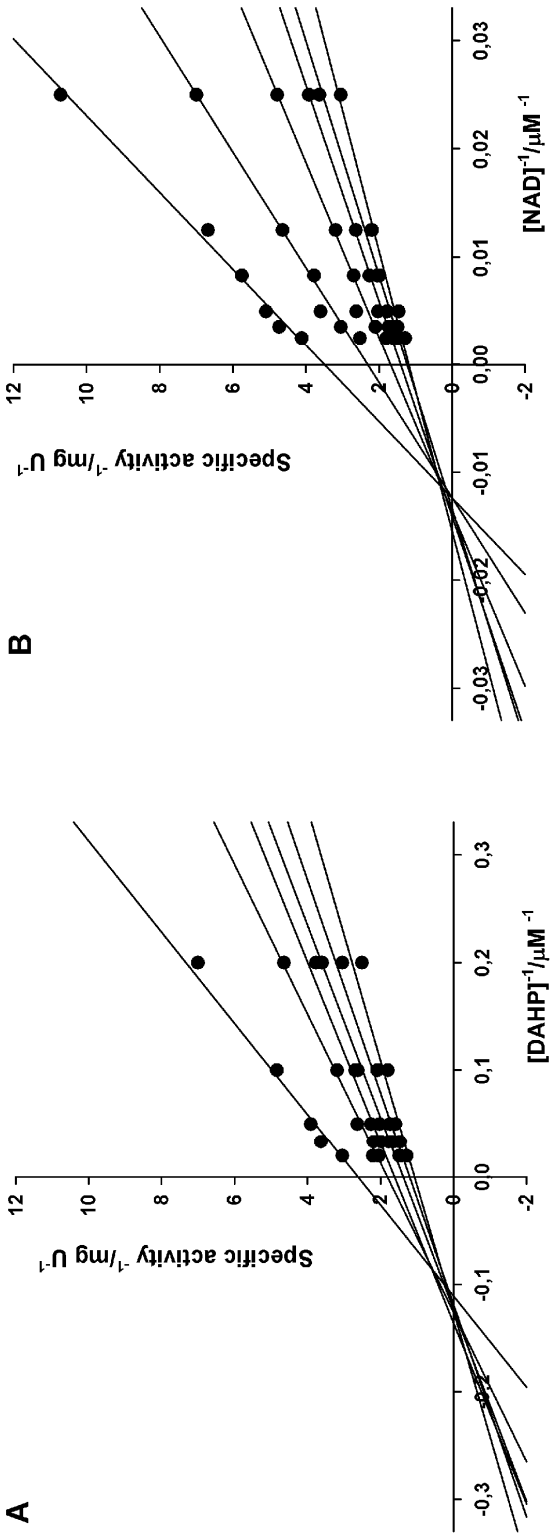

A
B
Specific activity -1/mg U-1
[DAHP]-1/µM -1
[NAD]-1/µM -1
12
10
8
6
4
2
0
-2
-0,3
-0,2
-0,1
0,0
0,1
0,2
0,3
-0,03
-0,02
-0,01
0,00
0,01
0,02
0,03

Figure 3

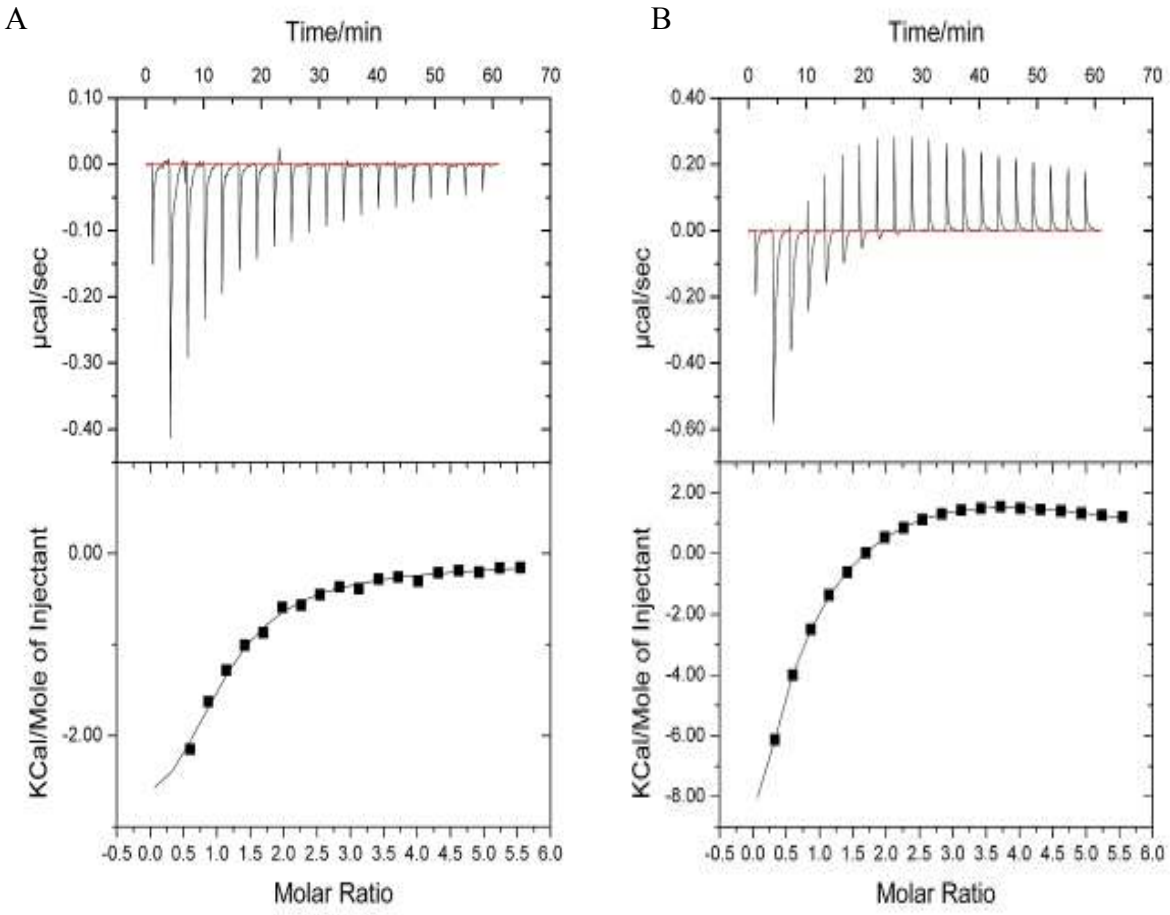

A
Time/min
µcal/sec
KCal/Mole of Injectant
Molar Ratio
B
Time/min
µcal/sec
KCal/Mole of Injectant
Molar Ratio

Figure 4

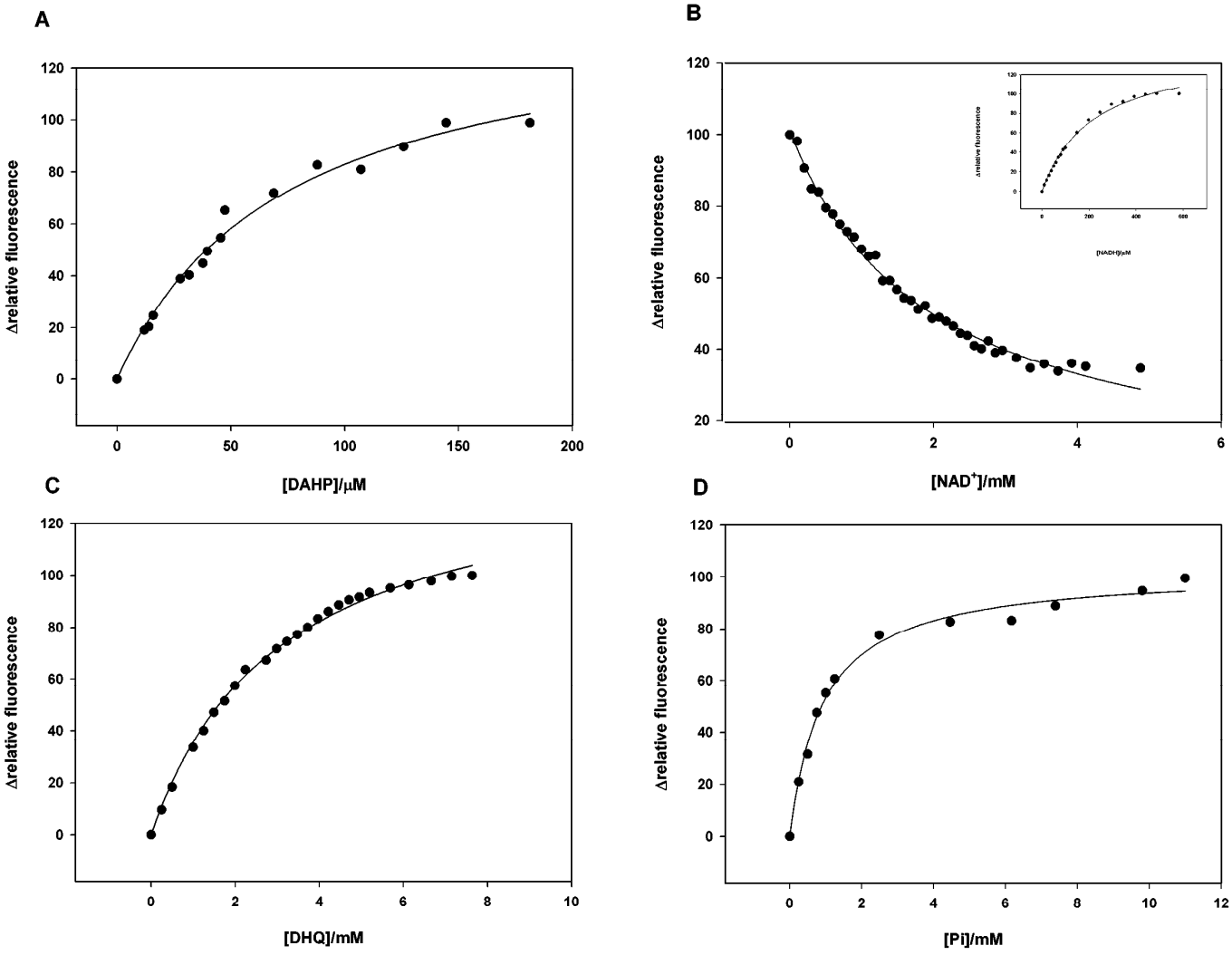
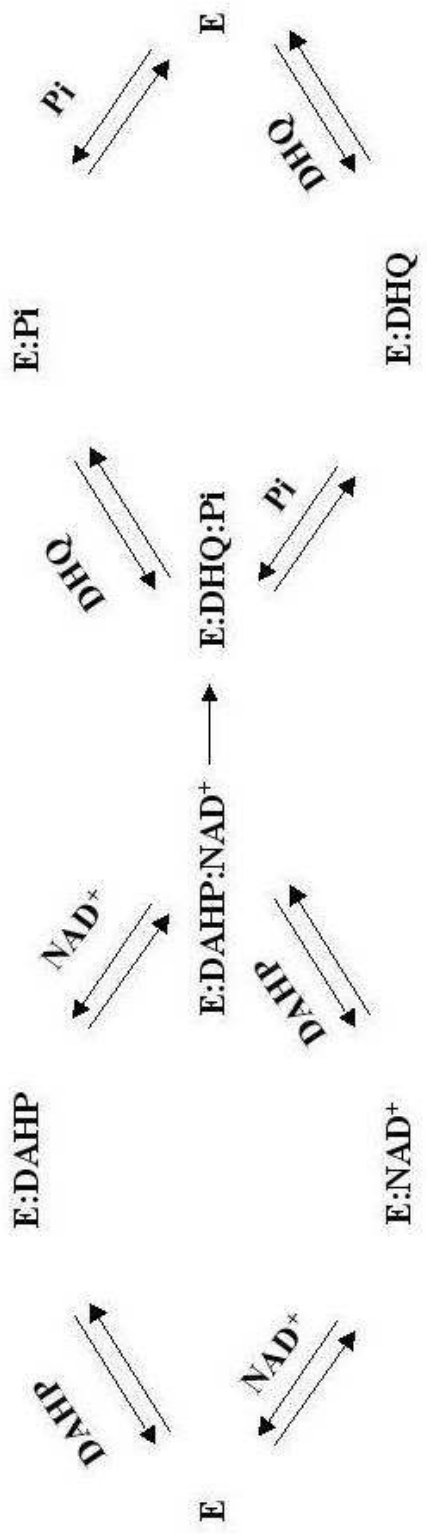

A
Δrelative fluorescence
[DAHP]/μM
B
Δrelative fluorescence
[NAD+]/mM
C
Δrelative fluorescence
[DHQ]/mM
D
Δrelative fluorescence
[Pi]/mM

Figure 5


E
DAHP
NAD+
E:DAHP
NAD+
E:NAD+
DAHP
NAD+
E:DAHP:NAD+
DAHP
Pi
DHQ
E:DHQ:Pi
Pi
DHQ
E:Pi
E:DHQ
Pi
DHQ
E

PARTE III

8. **Discussão**

A tuberculose é dita uma doença antiga, mas ainda presente nos dias atuais, preocupando as autoridades de saúde globais. Diante do quadro de resistência aos medicamentos utilizados no tratamento, sendo algumas linhagens ditas virtualmente incuráveis, são de extrema necessidade o estudo e a introdução de novas alternativas terapêuticas no tratamento da TB (Zhang, 2005; Goodman e Lipman, 2008).

O desenho racional de inibidores baseado no estado de transição de enzimas é uma moderna abordagem utilizada para o desenvolvimento de novos compostos com propriedades terapêuticas, e para tanto é necessária a caracterização detalhada do alvo escolhido, assim como do seu mecanismo catalítico. Esse alvo deve ser essencial ao patógeno e inexistente, ou com diferenças suficientes, no hospedeiro para que os efeitos adversos sejam nulos ou bastante diminuídos, obedecendo ao princípio da toxicidade seletiva. As enzimas da via do ácido chiquímico atendem a esses requisitos (Parish e Stoker, 2002).

O mecanismo proposto para a enzima DHQS é considerado incomum do ponto de vista mecanístico, onde em um único sítio catalítico, cinco diferentes reações ocorrem: oxidação, β-eliminação, redução, abertura do anel enol-piranose e condensação aldólica intramolecular (Bender *et al.*, 1989), sendo até proposta a passividade da enzima diante da reação, esta ocorrendo de forma espontânea. Porém, estudos estruturais evidenciaram não só a importância da participação da enzima como a sua ação em impedir que se formem metabólitos secundários (Barlett e Satake, 1988; Carpenter *et al.,* 1998).

Os metais divalentes têm um papel importante como cofator durante a reação catalisada pela DHQS e, para analisar esse papel na atividade da mtDHQS, estudos de

inativação e reativação com metais, titulação calorimétrica isotérmica (ITC) e absorção atômica foram realizados. Não foi necessária a adição de metal na reação quando nenhum tratamento quelante foi realizado, evidenciando que a enzima já se encontra na sua forma mais ativa e com a quantidade de metal necessária disponibilizada pelas condições de expressão.

A adição de EDTA ao meio de reação causou a diminuição da atividade da enzima, sendo observada a inativação de quase 100% da atividade após 10 minutos com 100μM de EDTA (Figura 6). Quando adicionado ao meio de incubação, em condições saturantes, o substrato DAHP evitou a inativação da enzima por até 60 minutos, o que sugere que o metal está menos acessível ao EDTA quando o DAHP liga-se ao sítio ativo. Em contrapartida, o NAD^+ não apresentou esse efeito protetor. Esses dados estão em acordo com o já reportado para as enzimas de *E. coli* e *N. crassa* (Bender *et al.*, 1989; Lambert *et al.*, 1985).

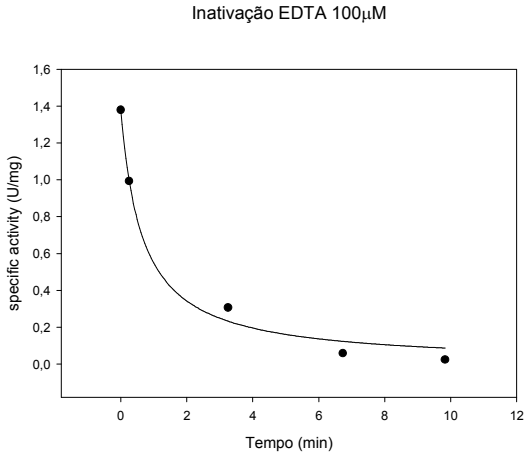


Figura 6: Gráfico representativo da inativação da enzima frente a 100μM EDTA.

Quando tratada com EDTA (100μM por 10 minutos), apenas a adição de 1 mM Co^{+2} foi capaz de restaurar completamente a atividade da enzima. O Zn^{+2} restaurou apenas parcialmente, mostrando ainda uma ação inibitória quando em excesso (5mM de Zn^{+2} reduziu 70% da atividade enzimática). O Co^{+2} também é responsável pela máxima atividade nas enzimas de *E. coli* (Bender *et* al, 1989) e *B. subtilis* (Hasan e Nester, 1978), além do Zn^{+2} ser inibitório para *E. coli* (Bender *et* al, 1989) e *A. nidulans* (Park *et al.,* 2006).

Os dados de ITC, ao titular Co^{+2} (Figura 3A do artigo) ou Zn^{+2} (Figura 3B do artigo) a uma solução de 56μM de DHQS, revelaram a existência de dois sítios seqüenciais para metal, o que é consistente com a proposição de um sítio secundário inibitório e explica essa ação do excesso de zinco (Park *et al.*, 2006). Ao contrário dos outros organismos, em *N. crassa*, o Zn^{+2} é metal de maior atividade (Lambert *et al.*, 1985).

Ao contrário do esperado pelos resultados de reativação acima descritos, o Co^{+2} não foi o metal mais abundante na solução pós-purificação, segundo os resultados da absorção atômica. Uma quantidade pequena, mas significativa, de Zn^{+2} foi encontrada, 0,2 mol de metal por mol de enzima, indicando que os sítios da enzima estão parcialmente saturados nas condições de análise. Essa quantidade subestequiométrica pode ser resultado da perda do metal durante o processo de purificação, como já descrito para outras enzimas como a DAHP (Phe) de *E. coli* (Sthephens e Bauerle, 1991). A presença de Zn^{+2} ao invés de Co^{+2} é atribuída, em outros trabalhos, a maior biodisponibilidade do primeiro, tornando a DHQS uma metaloenzima naturalmente dependente de zinco (Lambert *et al.*, 1985). Dessa forma, analisando os resultados referentes a metais, fica evidente a participação do metal no mecanismo catalítico da reação da DHQS, e estudos de cristalização que estão em

andamento irão auxiliar num melhor entendimento do papel do metal na reação e quão importante é esse metal.

Após a determinação do metal, e visto que em condições não-quelantes não há a necessidade de adicionar metal na reação, foram determinadas as constantes cinéticas verdadeiras. A partir do experimento de velocidade inicial, o dados foram aplicados à equação $v = VAB/(K_{ia}K_b + K_aB + K_bA + AB)$, onde A e B são concentrações de substratos, K_a e K_b são as constantes de Michaelis para os substratos A e B, respectivamente. Os valores obtidos foram: k_{cat} = 0,66 (± 0,02) s^{-1}, K_{DAHP} = 7,2 (± 0,5) µM, $K_{NAD}{}^{+}$ = 79 (± 6) µM, k_{cat}/K_{DAHP} = 9,2 (± 0,6) x 10^3 $M^{-1}s^{-1}$, and $k_{cat}/K_{NAD}{}^{+}$ = 8,4 (± 0.6) x 10^3 $M^{-1}s^{-1}$.

A comparação dos valores de k_{cat} mostra que a enzima de *M. tuberculosis* tem uma menor eficiência catalítica quando comparada com outras enzimas já caracterizadas (24 s^{-1} para *E. coli* (Maitra e Sprinson, 1978), 19 s^{-1} para *N. crassa* (Lambert *et al.,* 1985), 6.8 s^{-1} para *A. nidulans* (Park *et* al, 2006) e 0,66 s^{-1} para *mt*DHQS). Os valores de K_{DAHP} não evidenciam uma grande diferença de afinidade (cerca de 5 vezes), apesar de quando comparadas as constantes de especificidade - K_{cat}/K_{DAHP} 9,2 x 10^3 $M^{-1}s^{-1}$ de *mt*DHQS e 2,5 X $10^7 M^{-1}$ s^{-1} de *E. coli* (Bender *et al.,* 1989) – observa-se uma menor especificidade da *mt*DHQS. Já quando são comparados os valores para NAD^+, a diferença é maior (até 1000 vezes). A tabela 2 abaixo mostra um resumo dessa comparação.

Tabela 2: Comparação dos parâmetros cinéticos das deidroquinato sintase caracterizadas.

	B. subtillis	*E. coli*	*N. crassa*	*A. nidulans*	*mt*DHQS
K_{DAHP}	130µM	33 µM	1,4 µM	21 µM	7,2 µM
$K_{NAD}{}^{+}$	55µM	1µM	<0,2 µM	1,9 µM	79 µM
K_{cat}	-	24 s^{-1}	19 s^{-1}	6,8 s^{-1}	0,66 s^{-1}
Metal preferencial	Co^{+2} or Mn^{+2}	Co^{+2}	Zn^{+2}	Zn^{+2}	Co^{+2}
Referência	Hasan e Nester, 1978	Bender *et al*, 1989	Lambert *et al,* 1985	Park *et al,* 2006	

O padrão de retas obtido através do experimento de velocidade inicial indica a formação de um complexo ternário e a existência de um mecanismo seqüencial (Figura 2 do artigo). Com retas se interceptando a esquerda do eixo y, os mecanismos ping-pong (retas paralelas) e rápido-equilíbrio ordenado (retas interceptando no eixo y) puderam ser descartados (Segel, 1975).

Para a distinção entre os outros mecanismos, experimentos de ligação foram realizados no espectrofluorímetro mediante o monitoramento da fluorescência intrínseca da DHQS. A mudança na fluorescência da enzima mediante a titulação do substrato DAHP indicou que esse substrato é capaz de ligar-se à enzima livre, com K_D de 73 μM (Figura 4A do artigo). O mesmo não foi observado quando o segundo substrato, NAD^+, foi titulado, mas esse fato não pode ser interpretado como uma não-ligação à enzima livre. Dessa forma, um ensaio de competição foi idealizado frente ao NADH, um inibidor já descrito para outras DHQS (Srinivasan *et* al, 1963; Lambert *et al.,* 1985). A titulação de NADH na enzima livre, desta vez monitorando a fluorescência da nicotinamida, mostrou uma ligação do inibidor na enzima (Figura 4B inset do artigo). Após, NADH foi fixado em uma concentração próxima ao seu K_D (215μM) e NAD^+ foi titulado, e a mudança observada na fluorescência da nicotinamida evidenciou a ligação do NAD^+ à enzima livre, com K_D de 1400 μM (Figura 4B do artigo). A adição dos produtos DHQ e Pi à solução de enzima livre mostraram a capacidade de ambos de se ligarem à enzima, com K_D de 3100 μM (Figura 4C do artigo) e 920 μM (Figura 4D do artigo), respectivamente.

O padrão das retas do gráfico de duplos recíprocos e os dados dos experimentos de ligação são consistentes com o mecanismo cinético rápido-equilíbrio aleatório, onde ambos os substratos DAHP e NAD^+ podem ligar-se à enzima livre e não há uma ordem preferencial de dissociação dos produtos, DHQ e Pi, do sítio ativo (Figura 5 do artigo).

9. **Considerações finais e Perspectivas**

O desenho racional de inibidores depende fundamentalmente de informações sobre o mecanismo e a estrutura da enzima-alvo, sendo as etapas de determinação dos parâmetros cinéticos essenciais no processo de caracterização de alvos de interesse biológico.

A técnica de ITC é uma técnica sensível que tem sido amplamente empregada na determinação direta de parâmetros não só termodinâmicos como também cinéticos em reações enzimáticas, e tem sido de grande valia em casos onde não há a disponibilidade de ensaio direto para monitoramento da atividade das enzimas em estudo, especialmente quando o efeito de ativadores/inibidores/estabilizadores não se restringe à enzima em estudo ou não pode ser mensurado e/ou minimizado o seu efeito na enzima acoplada (Bianconi, 2007). Dessa forma, seria importante a padronização dessa técnica para o ensaio da DHQS, assim como a sua comparação e validação frente aos dados espectrofotométricos e acoplados obtidos nesse trabalho. Além disso, experimentos de estabilidade térmica e análise do envolvimento de prótons através do uso de tampões com diferentes energias de ionização também podem ser realizados para uma caracterização mais completa.

Além disso, os dados estruturais que serão obtidos a partir do protocolo de cristalização determinado (vide Anexo I) fornecerão a estrutura tridimensional da *mt*DHQS. Essa informação estrutural será de vital importância para obter uma melhor compreensão do modo de ação da enzima, a sua interação com seus substratos e com um possível inibidor em nível molecular, para apreciar as conseqüências estruturais das variações genéticas (mutações sítio-direcionadas).

10. **Referências**

[1] Agüero F, Al-Lazikani B, Aslett M, Berriman M, Buckner FS, Campbell RK, Carmona S, Carruthers IM, Chan AW, Chen F, Crowther GJ, Doyle MA, Hertz-Fowler C, Hopkins AL, McAllister G, Nwaka S, Overington JP, Pain A, Paolini GV, Pieper U, Ralph SA, Riechers A, Roos DS, Sali A, Shanmugam D, Suzuki T, Van Voorhis WC, Verlinde CL (2008) Genomic-scale prioritization of drug targets: the TDR Targets database. *Nat Rev Drug Discov*, 7(11), 900-907.

[2] Barlett PA, Satake K (1988) Does dehydroquinate synthase synthesize dehydroquinate? *JACS*, 110, 1628-1630.

[3] Bender SL, Mehdi S, Knowles JR (1989) Dehydroquinate synthase: role of divalent metal cations and of nicotinamide adenine dinucleotide in catalysis. *Biochemistry*, 28, 7555-7560.

[4] Bentley R (1990) The Shikimate pathway – A metabolic tree with many branches. *Crit Rer Biochem Mol Biol*, 25, 307 –384.

[5] Bianconi ML (2007) Calorimetry of enzyme-catalyzes reactions. *Biophysical Chemistry*, 126, 59-64

[6] Bloom BR, Murray CJL (1992) Tuberculosis: commentary on a reemergent killer. *Science*, 257, 1055-1064.

[7] Blower S, Supervie V (2007) Predicting the future of XDR tuberculosis. *The Lancet Infection Disease*, 7, 443.

[8] Carpenter EP, Hawkins AR, Frost JW, Brown KA (1998) Structure of dehydroquinate synthase reveals an active site capable of multistep catalysis. *Nature*, 394**,** 299-302.

[9] Centers for Disease Control and Prevention (2006) Emergence of Mycobacterium tuberculosis with extensive resistance to second-line drugs – Worldwide, 2000-2004. *MMWR*, 55(11), 301-305.

[10] Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE 3rd, TekaiaF , 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 MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*, 393, 537-544.

[11] Corbett EL, Watt CJ, Walker N, Maher D, Williams BG, Raviglione MC, Dye C (2003) The growing burden of tuberculosis. *Arch Intern Med,* 163, 1009-1021.

[12] Daniel TM (2006) The history of tuberculosis. *Resp Medicine*, 100, 1862-1870.

[13] Ducati, RG, Ruffino-Netto A, Basso LA, Santos DS (2006) The resumption of consumption – a review on tuberculosis. *Memórias Instituto Oswaldo Cruz*, 101, 697-714.

[14] Duncan K (2003) Progress in TB drug development and what is still needed. *Tuberculosis,* 83, 201-207.

[15] Dye C, Watt CJ, Bledd DM, Hosseini SM, Raviglione MC (2005) Evolution of tuberculosis control and prospects for reducing tuberculosis incidence, prevalence, and deaths globally. *JAMA***,** 293, 2767-2775.

[16] Fuentes ZM, Caminero JA (2006) Controversies in the treatment of extrapulmonary tuberculosis. *Arch Bronconeumol*, 42(4), 194-201.

[17] Gandhi NR, Moll A, Sturm AW, Pawinski R, Govender T, Lalloo U, Zeller K, Andrews J, Friedland G (2006) Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area od South Africa. *The Lancet*, 368, 1575-1580.

[18] Gomez JE, McKinney JD (2004) *M. tuberculosis* persistence, latency, and drug tolerance. *Tuberculosis*, 84, 29-44.

[19] Goodman A, Lipman M (2008) Tuberculosis. *Clinical Medicine*, 8(5), 531-534.

[20] Hasan N, Nester EW (1978) Dehydroquinate synthase in *Bacillus subtilis* - an enzyme associated with chorismate synthase and flavin reductase. *The Journal of Biological Chemistry*, 253(14), 4999-5004.

[21] Hingley-Wilson SM, Sambandamurthy VK, Jacobs Jr WR (2003) Survival perspectives from the world's most successful pathogen, *Mycobaterium tuberculosis*. *Nat Immunol*, 4, 949-955.

[21] Janin YL (2007) Antituberculosis durgs: ten years of research. *Bioorganic Medicine and Chemistry*. doi: 10.106/j.bmc.2007.01.030, 2007.

[22] Kudo F, Hosomi Y, Tamegai H, Kakinuma K (1999) Purification and characterization of 2-deoxy-scyllo-inososse synthase derived from *Bacillus circulans*. A crucial carbocyclization enzyme in the biosynthesis of 2-deoxystreptamine-containing aminoglycoside antibiotics. *The Journal of Antibiotics*, 52(2), 81-88.

[23] Kudo F, Tamegai H, Fujiwara T, Tagami U, Hirayama K, Kakinuma K (1999a) Molecular cloning of the gene for the key carbocycle-forming enzyme in the biosynthesis

of the 2-deoxystreptamine-containing aminocyclitol antibiotics and its comparision with dehydroquinate synthase. *The Journal of Antibiotics*, 52(6), 559-571.

[24] Lambert JM, Boocock MR, Coogins JR (1985) The 3-dehydroquinate synthase activity of the pentafunctional arom enzyme complex of *Neurospora crassa* is Zn^{2+}-dependent. *Biochem J*, 226, 817-829

[25] Lemos ACM (1994) Quimioterapia da TB. Em: Silva, P. Farmacologia. 4.ed. Rio de Janeiro:Guanabara Koogan, 1128-1136.

[26] Mahmoudi A, Iseman MD (1993) Pitfalls in the care of patients with tuberculosis. Common errors and their association with the acquisition of drug resistance. *JAMA*, 270, 65-68.

[27] Maitra US, Sprinson DB (1978) 5-Dehydro-3-deoxy-D-arabino-heptulosonic acid 7-phosphate. An intermediate in the 3-dehydroquinate synthase reaction. *J Biol Chem*, 253(15), 5426-30.

[28] NSB Editorial Comment (2000) Taming tuberculosis – again. *Nature Structural Biology,* 7(2), 87-88.

[29] Parish T, Stoker NG (2002) The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*. *Microbiology*, 148, 3069-3077.

[30] Park A, Lamb HK, Nichols C, Moore JD, Brown KA, Cooper A, Charles IG, Stammers DK, Hawkins AR (2006) Biophysical and kinetic analysis of wild-type and site-directed mutants of the isolated and native dehydroquinate synthase domain of the AROM protein. *Protein Science*, 13, 2108-2119.

[31] Pasqualoto KFM, Ferreira EI (2001) An approach for the rational design of new antituberculosis agents. *Current Drug Targets*, 2,427-437.

[32] Raman K, Yeturu K, Chandra N (2008) targetTB: a target identification pipeline for Mycobacterium tuberculosis through an interactome, reactome and genome-scale structural analysis. *BMC Syst Biol*, 19, 2-109

[33] Raviglione MC (2003) The TB epidemic from 1992 to 2002. *Tuberculosis*, 83, 4-14.

[34] Sharma SK, Mohan A (2004) Multidrug-resistant tuberculosis. *Indian Journal of Medical Research,* 120, 354-376.

[34] Sthephens CM, Bauerle R (1991) Analysis of the metal requirement of DAHPS from *Escherichia coli*. *The Journal of Biological Chemistry*, 266(31), 20810-20817.

[35] Segel IH (1975) *Enzyme Kinetics, Behavior and analysis of rapid equilibrium and steady-state enzyme systems*, John Wiley & Sons, Inc., New York, 957p.

[36] Srinivasan PR, Rothschild J, Sprinson DB (1963) The enzymic conversion of 3-deoxy-D-arabino-heptulosonic acid 7-phosphate to 5-dehydroquinate. *J Biol Chem*, 238, 3176-3182.

[37] Tortora GJ, Funke BR, Case CL. 2000. Microbiologia, 6.ed. Porto Alegre:Artes Médicas Sul.

[38] WHO (2004) Anti-tuberculosis drug resistance in the world: Third Global Report, Geneve Switzerland: World Health Organization.

[39] WHO (2007) 10 facts about tuberculosis. Disponível em www.who.org . Acessado em 22.04.2007.

[40] WHO (2008) Global tuberculosis control – surveillance, planning, financing. WHO/HTM/TB/2008.393. Geneva: World Health Organization.

[41] WHO (2009) Global tuberculosis control: a short update to the 2009 report. WHO /HTM/TB/2009.426. Geneva: World Health Organization.

[42] WHO (2009a) TB/HIV fact sheet 2009. Disponível em http://www.who.int/tb/challenges/hiv/en/index.html , acessado em 28/09/2009

[43] WHO (2009b) Countries that had reported at least oneXDR-TB case by end March 2009. World Health Organization.

[44] WHO (2010) Disponível em http://www.who.int/mediacentre/factsheets/fs104/en/index.html; acessado em 26/04/2010.

[45] WHO (2010b) M/XDR-TB 2010 report fact sheet. Disponível em http://www.who.int/tb/features_archive/world_tb_day_2010/en/index.html, acessado em 26/04/2010.

[46] Zhang, Y (2005) The magic bullets and Tuberculosis drug targets. *Annual Rev Pharmacol Toxicol*. 45, 529-564.

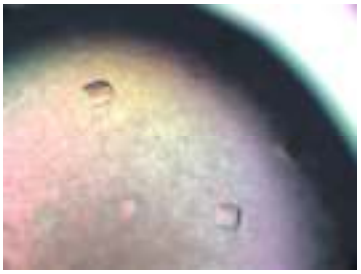
ANEXOS

Resultados obtidos durante o Mestrado que não fazem parte do corpo da dissertação.

11. ANEXO I – Fotos dos cristais de *mt*DHQS.

Foram obtidos cristais da DHQS nas suas formas baixo listadas, nas condições 27 PegIon, 30 do CS2 e 38 do CS2. Os cristais foram encaminhados ao Laboratório Nacional de Luz Síncontron em Campinas, SP, para coleta de dados.

- √ APO
- √ complexada DAHP
- √ complexada NAD^+
- √ complexada DHQ
- √ complexada Pi
- √ complexada NADH + DAHP.



DHQS + DAHP, PegIon 27

DHQS + NAD^+, CS2 30

DHQS + Pi, PegIon 27

12. ANEXO II - Participação em publicação: Werlang ICR, Schneider CZ, Mendonça JD, Palma MS, Basso LA, Santos DS (2009) Identification of Rv3852 as a nucleoid-associated protein in *Mycobacterium tuberculosis*, *Microbiology*, 155, 2652-2663.

O estudo da Rv3852, identificada no genoma como a provável histone-like protein (H-NS) de *Mycobacterium tuberculosis* (Cole *et al.*, 1998) tem o objetivo de caracterizar essa importante proteína envolvida em diferentes processos celulares, muitos relacionados a mudanças do ambiente e adaptação. Essa modulação ocorre através da regulação pleiotrópica, regulação da transcrição, recombinação e transposição (Atlung e Ingmer, 1997; Hommais *et al.*, 2001; Rimsky *et al.*, 2001; Dorman, 2004; Ward *et al.*, 2007). A ligação da H-NS ao DNA é independente da seqüência de nucleotídeos, sendo a estrutura dessa seqüência o fator determinante para o reconhecimento molecular, especialmente a curvatura intrínseca comumente encontrada nos promotores (Yamada *et al.*, 1991; Pérez-Martin *et al.*, 1994). Após a ligação, a H-NS pode influenciar a expressão gênica ao alterar a topologia do DNA (Owen-Hughes *et al.*, 1992; Tupper *et al.*, 1994) e também ao induzir a curvatura em um DNA não-curvado (Spurio *et al.*, 1997).

As proteínas nucleóide associadas (NAPs), ou proteínas tipo histonas (H-NS), são um grupo de pequenas proteínas (cerca de 14 KDa), normalmente básicas, ligantes a DNA que contribuem para a organização de nucleóides bacterianos e para o controle de diversos processos celulares chaves, como recombinação, replicação e transcrição (Dorman e Deighan, 2003; Dame, 2005; Luijterburg *et al.*, 2006). Recentemente, algumas NAPs para micobactérias têm sido descritas: Hlp ou MPD1 para *M. tuberculosis*, *M. bovis* e *M.*

smegmatis, e Lsr2 em *M. tuberculosis* e *M. smegmatis*. Lsr2 tem múltiplos papéis na resistência a antibióticos, na síntese da parede celular e na proteção do DNA contra espécies reativas de oxigênio. Estas características foram analisadas em detalhe no *M. tuberculosis*, fornecendo evidências de que esta é NAP de micobactéria funcionalmente equivalente a H-NS de bactérias Gram-negativas, apesar da falta de homologia significativa de seqüência (Gordon *et al.*, 2008; Colangeli *et al.*, 2009).

A H-NS tem sido extensivamente estudada em outros microrganismos patogênicos, como *E. coli* (Spurio *et al.*, 1997; Berlutti *et al.*, 1998), *Salmonella typhimurium* (Harrison *et al.*, 1994), *Shigella flexneri* (Falconi *et al.*, 2001; Beloin e Dorman, 2003), *Vibrio cholerae* (Nye *et al.*, 2000), *Serratia marcescens* (Franzon e Santos, 2004) e *Xylella fastidiosa* (Paula *et al.*, 2003), assim como a caracterização do seu papel na regulação dos fatores de virulência desses microrganismos. Em *E. coli*, por exemplo, o gene *hns* foi demonstrado contribuir para a resistência a drogas ao regular a expressão de genes de multirresistência (Nishino e Yamaguchi, 2004). Além disso, pesquisas envolvendo transferência gênica horizontal em bactérias confirmaram o importante papel da H-NS no silenciamento dos genes adquiridos (Luchinni *et al.*, 2006; Navarre *et al.*, 2006; Oshima *et al.*, 2006; Dorman, 2007).

Dessa forma, devido ao importante papel da H-NS e outras NAPs demonstrado nos mecanismos de virulência de outros patógenos, incluindo bactérias facultativas, torna-se importante caracterizar a função biológica da Rv3852 de *M. tuberculosis*, para desenvolvimento de inibidores específicos e também o desenvolvimento de vacinas.

Durante o ano de 2009, dentro desse trabalho, participei dos experimentos de complementação gênica. A complementação *in vivo* da susceptibilidade a serina em mutantes *hns* em *E. coli* tem sido utilizada para identificar as proteínas H-NS em diferentes

bactérias (Tendeng e Bertin, 2003; Rodriguez *et al.*, 2005). Como esperado, as células BSN26 cresceram a uma taxa maior que o mutante BSN27 (que possui um forte defeito de crescimento). Porém, nenhuma diferença foi observada no crescimento nas células transformadas BSN27 (pKK223-3) e BSN27 (pKH2) nas condições testadas, demonstrando assim que a Rv3852 de *M. tuberculosis* não corresponde a "clássica" H-NS, e reforça a proposição de se tratar de uma nova classe de NAP com necessidade de uma melhor caracterização da sua função biológica.

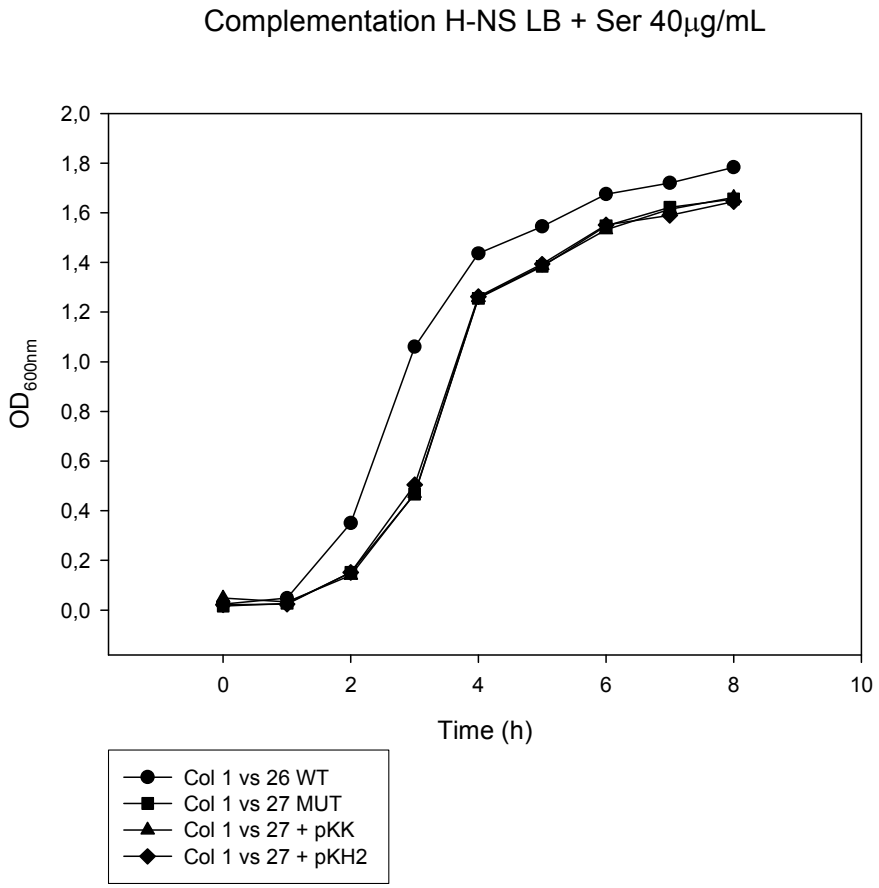


Figura: Curvas de crescimento. A transformação do mutante (knock-out *hns*) com o plasmídeo pKH2 (♦) não foi capaz de restaurar o fenótipo de crescimento comparado ao selvagem BSN26 (●), mantendo sua taxa de crescimento similar aos controles realizados (BSN27 (■) e BSN27 transformado com o plasmídeo controle pKK223-3 (▲).

Referências (demais, vide artigo)

Paula DP, Azzoni AR, Siqueira SF, Catani CF, Rosseli LK, de Souza AP (2003) Expression and purification of a putative H-NS nucleoid-associated protein from the phytopathogen Xyllela fastidiosa. *Protein Expr Purif*, 32, 61-67.

Nishino K, Yamaguchi A (2004) Role of histone-like protein HNS in multidrug resistance of *Escherichia coli*. *J Bacteriol*, 186, 1423-1429.

Luchinni S, Rowley G, Goldberg MD, Hurd D, Harrison M, Hinton JCD (2006) H-NS mediates the silencing of laterally acquired genees in bacteria. *PloS Pathog*, 2,746-752.

Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, Fang FC (2006) Selective silencing of foreing DNA with low GC content by the H-NS protein in Salmonella. *Science*, 313,236-238.

Oshima T, Ishikawa S, Kurokawa K, Aiba H, Ogasawara N (2006) *Escherichia coli* histone-like protein preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA Res*, 13, 141-153.

Microbiology (2009), **155**, 2652–2663 **DOI** 10.1099/mic.0.030148-0

Identification of Rv3852 as a nucleoid-associated protein in *Mycobacterium tuberculosis*

Isabel C. R. Werlang,[1,2] Cristopher Z. Schneider,[1] Jordana D. Mendonça,[1] Mario S. Palma,[3] Luiz A. Basso[1] and Diógenes S. Santos[1]

Correspondence
Luiz A. Basso
luiz.basso@pucrs.br
Diógenes S. Santos
diogenes@pucrs.br

[1]Centro de Pesquisas em Biologia Molecular e Funcional, Instituto Nacional de Ciência e Tecnologia em Tuberculose, Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga 6681, Porto Alegre, RS 90619-900, Brazil

[2]Programa de Pós-Graduação em Biologia Celular e Molecular, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves 9500, Porto Alegre, RS 91501-970, Brazil

[3]Laboratório de Biologia Estrutural e Zooquímica, Centro de Estudos de Insetos Sociais, Departamento de Biologia, Instituto de Biociências, Universidade Estadual Paulista, Rio Claro, SP 13506-900, Brazil

Tuberculosis remains the major cause of mortality due to a bacterial pathogen, *Mycobacterium tuberculosis*. The molecular mechanisms of infection and persistence have not been completely elucidated for this pathogen. Studies involving nucleoid-associated proteins (NAPs), which have been related to the control and influence of virulence genes in pathogenic bacteria, can help unveil the virulence process of *M. tuberculosis*. Here, we describe the initial characterization of an ORF for an *M. tuberculosis* putative NAP. The *Rv3852* gene was cloned and expressed, and its product purified to homogeneity. A qualitative protein–DNA binding assay was carried out by gel-retardation and the protein affinity for specific DNA sequences was assessed quantitatively by surface plasmon resonance (SPR). A stoichiometry of 10 molecules of monomeric protein per molecule of DNA was determined. The monophasic apparent dissociation rate constant values increased to a saturable level as a function of protein concentration, yielding two limiting values for the molecular recognition of proU2 DNA. A protein–DNA binding mechanism is proposed. In addition, functional complementation studies with an *Escherichia coli hns* mutant reinforce the likelihood that the Rv3852 protein represents a novel NAP in *M. tuberculosis*.

Received 20 April 2009
Revised 21 May 2009
Accepted 26 May 2009

INTRODUCTION

Tuberculosis (TB) is one of the major causes of death worldwide caused by a single infectious agent, *Mycobacterium tuberculosis*. TB resurgence in the late 1980s was caused by a combination of several factors, such as HIV co-infection, increased poverty in urban areas and emergence of *M. tuberculosis* multidrug-resistant strains (MDR-TB) (Raviglione, 2003). According to the 2008 Global TB Control Report of the World Health Organization (WHO, 2008), there were approximately 9.2 million new TB cases in 2006, of which 0.5 million were MDR-TB. Moreover, the emergence of extensively drug-resistant (XDR) TB cases (CDC, 2007), defined as cases in persons with TB whose isolates are MDR-TB as well as resistant to any one of the fluoroquinolone drugs and to at least one of the three injectable second-line drugs, amikacin, kanamycin or capreomycin, and their global distribution (Dorman & Chaisson, 2007), raise the prospect of virtually incurable TB worldwide. To compound the problem, it has been estimated that of 9.27 million incident TB cases in 2007, 1.37 million (15 %) were HIV-positive (WHO, 2009).

M. tuberculosis has been considered the world's most successful pathogen. It is able to resist macrophage killing and persist in body tissues, thereby establishing a latent infection which can be reactivated when the host immune system wanes (Gomez & McKinney, 2004; Hingley-Wilson *et al.*, 2003). The mechanism by which *M. tuberculosis* establishes latency and persistence is largely unknown and efforts have been made to address this and other issues, such as virulence (Andersen, 2007; Gandotra *et al.*, 2007; Saunders & Britton, 2007; Schnappinger *et al.*, 2006).

Nucleoid-associated proteins (NAPs), also known as histone-like proteins, are a diverse group of small, usually basic, DNA-binding proteins that contribute to the organization of bacterial nucleoids and the control of many key cellular processes, such as recombination,

Abbreviations: EMSA, electrophoretic mobility shift assay; NAP, nucleoid-associated protein; SPR, surface plasmon resonance; TB, tuberculosis.

replication and transcription (Dorman & Deighan, 2003; Dame, 2005; Luijsterburg *et al.*, 2006). The best studied NAPs are Fis (factor for inversion stimulation), H-NS (histone-like nucleoid structuring), HU (heat unstable) and IHF (integration host factor), although the existence of other similar proteins is becoming increasingly apparent (Luijsterburg *et al.*, 2008). Their structural and functional properties have been related to chromatin compaction (by bending, bridging and wrapping DNA) and global transcriptional regulation of several genes, including virulence genes of bacterial pathogens (Beloin & Dorman, 2003; Berlutti *et al.*, 1998; Falconi *et al.*, 2001; Franzon & Santos, 2004; Harrison *et al.*, 1994; Mangan *et al.*, 2006; Nye *et al.*, 2000; Schechter *et al.*, 2003; Stonehouse *et al.*, 2008; Wilson *et al.*, 2001). Most of these proteins have been reported to bind to non-specific DNA sequences (Bailly *et al.*, 1995; Krylov *et al.*, 2001; Yamada *et al.*, 1991), to exhibit low sequence or structural conservation (Luijsterburg *et al.*, 2008), and to induce DNA bending and other alterations (Dhavan *et al.*, 2002; Koh *et al.*, 2008; Schneider *et al.*, 2001; Spurio *et al.*, 1997; Swinger & Rice, 2004). Curved DNA is an integral element of promoter architecture and DNA bending is one of the major components of the control of bacterial gene expression (Pérez-Martín *et al.*, 1994).

The probable involvement of NAPs in both mycobacterial survival and adaptation to adverse host conditions should provide new insights to elucidate the molecular and cellular mechanisms by which *M. tuberculosis* establishes infection and persistence. In recent years, two NAPs have been identified in mycobacteria: Hlp (histone-like protein) or MDP1 (mycobacterial DNA-binding protein 1), which has been characterized in *M. tuberculosis* (Prabhakar *et al.*, 1998), *Mycobacterium smegmatis* (Lee *et al.*, 1998; Mukherjee *et al.*, 2008; Shires & Steyn, 2001) and *Mycobacterium bovis* (Lewin *et al.*, 2008); and Lsr2, which has been characterized in *M. tuberculosis* and *M. smegmatis* (Chen *et al.*, 2006, 2008; Colangeli *et al.*, 2007). Lsr2 has multiple roles in antibiotic resistance, cell wall biosynthesis and DNA protection against reactive oxygen intermediates, and these have been examined in detail in *M. tuberculosis*, providing evidence that this mycobacterial NAP is functionally equivalent to H-NS from Gram-negative bacteria, despite the lack of significant sequence homology (Colangeli *et al.*, 2009; Gordon *et al.*, 2008).

Another possible histone-like protein (Rv3852) has been proposed to be encoded in the genome of the *M. tuberculosis* H37Rv strain (Cole *et al.*, 1998). The *Rv3852* gene sequence consists of 405 nt that encode a small 134 aa protein with a predicted molecular mass of 13.8 kDa and a positive net charge. It is conserved among other mycobacterial species, showing some similarity at the N-terminal region with eukaryotic histones, e.g. histone H1 from *Trypanosoma cruzi* (Cole *et al.*, 1998, 2001; Stinear *et al.*, 2008). Together, these data suggest that *M. tuberculosis* Rv3852 and its homologues are members of a family of DNA-binding proteins. H-NS is a small chromatin-associated protein which is involved in many different cellular processes, most of them related to environmental changes and adaptation. H-NS modulates these processes through pleiotropic regulation of transcription, recombination and transposition (Atlung & Ingmer, 1997; Dorman, 2004; Hommais *et al.*, 2001; Rimsky *et al.*, 2001; Ward *et al.*, 2007). DNA binding by H-NS is independent of nucleotide sequence, despite its preference for curved sequences usually rich in AT nucleotides (Yamada *et al.*, 1991). It should, however, be pointed out that differential binding of H-NS to defined DNA sequence sites has recently been reported (Bouffartigues *et al.*, 2007). Considering the critical roles played by H-NS and other NAPs in adaptation to environmental changes, and in the growth and virulence mechanisms of pathogens, including facultative intracellular bacteria, attempts to establish/assign the biological function of the putative *M. tuberculosis* histone-like protein Rv3852 are worthwhile.

As a first step towards this goal, biochemical characterization of the putative histone-like protein Rv3852 from *M. tuberculosis* is required. We have thus amplified and cloned the corresponding *M. tuberculosis* ORF, proposed to be a possible histone-like protein, expressed the recombinant protein in *Escherichia coli* cells and purified it to homogeneity. We have also characterized the protein–DNA-binding interactions by a qualitative assay using a gel-retardation technique and verified the protein affinity for specific DNA sequences by surface plasmon resonance (SPR) using BIA-Core equipment. These studies have identified the product of the *M. tuberculosis Rv3852* DNA sequence as a DNA-binding protein that slightly prefers binding to curved DNA. In addition, we performed complementation studies to clarify whether the *M. tuberculosis* Rv3852 protein acts in a similar or different pattern to the well-characterized *E. coli* H-NS. Our results demonstrate that Rv3852 is a DNA-binding protein with unique structural and functional properties that nevertheless resembles other NAPs. We propose that Rv3852 is a member of a novel class of histone-like proteins present in *M. tuberculosis* and related mycobacteria.

METHODS

Cloning and construction of an Rv3852c expression plasmid. Synthetic oligonucleotide primers HNS *Nde*I (5′-AACATATG-CCAGACCCGCAGGATCGACCC-3′) and HNS *Hin*dIII (5′-GTAAGCTTTCAGCGGCGGCGCAGTTGCC-3′) were designed based on the coding sequence described for the *Rv3852* (*hns*) locus from *M. tuberculosis* H37Rv (Cole *et al.*, 1998). *Nde*I and *Hin*dIII restriction sites are underlined in the respective primers. Primers were used to amplify the target gene (405 bp) from *M. tuberculosis* genomic DNA using 10 % DMSO and *Pfu* DNA polymerase (Stratagene), and the PCR product was cloned into the pCR-Blunt vector (Invitrogen) and subcloned into the pET-23a(+) expression vector (Novagen). The nucleotide sequence of the cloned fragment was determined by automated DNA sequencing.

Expression of recombinant Rv3852 in *E. coli*. The recombinant pET-23a(+)::*hns* plasmid was transformed into electrocompetent *E.*

coli Rosetta (DE3) cells, and selected on Luria–Bertani (LB; Difco) agar plates containing 50 μg carbenicillin ml^{-1}. Control experiments were performed under the same experimental conditions, except that transformed *E. coli* cells harboured the expression vector lacking the target gene. A single recombinant colony was used to inoculate 50 ml LB medium containing 50 μg carbenicillin ml^{-1}. Cells were grown with stirring at 180 r.p.m. at 37 °C and, after reaching OD_{600} 0.3–0.5, they were grown for additional 3, 6, 9, 12, 18, 21 and 24 h without addition of IPTG. Cells (1.5 ml) were harvested by centrifugation at 20 800 ***g*** for 5 min and stored at −20 °C. The stored cells were suspended in 10 mM KH_2PO_4 (pH 7.2) and disrupted by sonication using three 10 s pulses and cell debris was separated by centrifugation at 20 800 ***g*** for 30 min at 4 °C. The soluble protein content was analysed by 15 % SDS-PAGE with Coomassie brilliant blue staining.

Purification of recombinant protein. For protein purification under native conditions, 15 g cells was resuspended in 75 ml lysis buffer [10 mM KH_2PO_4, pH 7.2, 500 mM NaCl, Complete Protease Inhibitor Cocktail (Roche)] and sonicated (15 times for 20 s at an amplitude of 60 % in a Vibra-Cell ultrasonic processor). Cell debris was pelleted by centrifugation (51 900 ***g*** for 30 min). The supernatant containing soluble protein was incubated with 1 % (w/v) streptomycin sulfate for 30 min and centrifuged at 51 900 ***g*** for 30 min. The supernatant was dialysed six times against 2 l 10 mM KH_2PO_4, pH 7.2 (buffer A), using dialysis tubing with a molecular mass exclusion limit of 3500 Da. This sample was clarified by centrifugation (51 900 ***g*** for 30 min) and loaded on an SP-Sepharose Fast Flow column (GE Healthcare) pre-equilibrated with buffer A. The column was washed with six column volumes of buffer A and the adsorbed material was eluted with a linear gradient (0–100 %) of 20 column volumes 10 mM KH_2PO_4, pH 7.2, 0.6 M KCl (buffer B) at 1 ml min^{-1}. The fractions containing recombinant protein were pooled (75 ml) and concentrated to 8 ml using an Amicon ultrafiltration membrane (molecular weight cut-off 3000 Da). The sample was loaded on a Sephacryl S-100 column (GE Healthcare) pre-equilibrated with buffer A. The Rv3852 recombinant protein was eluted in a total volume of 155 ml at a flow rate of 0.25 ml min^{-1} and loaded on a cation-exchange Mono-S column (GE Healthcare) pre-equilibrated with buffer A. The column was washed with six column volumes of the same buffer and the absorbed material was eluted with 20 column volumes of a linear gradient (0–100 %) of buffer B at 1 ml min^{-1}. Elution profiles were followed at 280 and 215 nm. Homogeneous Rv3852 protein was eluted in a total volume of 7 ml and stored at −80 °C. The recombinant protein purity was assessed by 15 % SDS-PAGE with Coomassie brilliant blue staining, and protein concentration was measured spectrophotometrically at 215 nm as described by Scopes (1994).

N-terminal amino acid sequencing. The N-terminal amino acid residues of the homogeneous recombinant Rv3852 protein preparation were identified by automated Edman degradation sequencing using a PPSQ 21A gas-phase sequencer (Shimadzu).

Determination of native Rv3852 molecular mass. The molecular mass of native recombinant protein was estimated by gel-permeation chromatography on a Superdex 200 HR column (GE Healthcare). The column was eluted with 10 mM KH_2PO_4, pH 7.2, at a flow rate of 0.5 ml min^{-1}. The eluate was monitored at 215 and 280 nm, and the column was calibrated with the following protein standards (GE Healthcare): RNase A (13 700 Da), chymotrypsinogen (25 000 Da), ovalbumin (43 000 Da) and albumin (67 000 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 masses.

Electrophoretic mobility shift assay (EMSA). Four DNA sequences were constructed for DNA–protein binding assay. proU1 (−200/+10) and proU2 (−20/+189) are sequences of the *proU* promoter region from *M. tuberculosis* H37Rv (http://genolist.pasteur.fr/TubercuList/index.html). Primers were designed, synthesized and used for amplifying each fragment by PCR using *Pfu* DNA polymerase (Stratagene). BENT (curved) and NC (noncurved) sequences were synthesized by annealing oligodeoxynucleotides (Table 1) of ~100 bp (Shimizu *et al.*, 1995). All fragments were cloned into the *Eco*RI restriction site of the pCR-Blunt vector (Invitrogen). Plasmid DNAs were cleaved with *Eco*RI (Invitrogen) and the inserts were purified from agarose gels. An equal amount (300 ng) of the DNA fragments was mixed with purified recombinant protein (0–1 μg) in binding buffer (10 mM Tris/HCl, pH 7.5, 80 mM NaCl, 1 mM DTT, 1 mM EDTA, 5 %, v/v, glycerol) in a total volume of 40 μl. After 15 min incubation at room temperature, these mixtures were loaded on a 1.5 % agarose gel and run in Tris-borate-EDTA buffer for 5 h. The gel was stained with 0.05 μg ethidium bromide ml^{-1} and visualized under UV light.

Kinetic analysis using a BIA-Core X system. Another 221 bp fragment of the proU2 sequence was amplified from the pCR-Blunt-proU2 recombinant plasmid using the forward primer labelled with biotin at the 5′ terminal end. The labelled fragment was immobilized (0.15 μg ml^{-1}) onto the second channel of a streptavidin sensor chip (SA Chip; GE Healthcare) and the first channel was used for reference. Unconjugated streptavidin was removed by three consecutive injections of 1 M NaCl, 50 mM NaOH (10 μl; 5 μl min^{-1}), and then 60 μl proU2-biotin was passed at 5 μl min^{-1} in HBS-EP buffer (BIA-Core; GE Healthcare) with 0.5 M NaCl. Several concentrations of protein (0, 0.05, 0.1, 0.5, 1, 2.5, 5 and 7.5 μM) were injected in HBS-EP buffer (80 μl; 50 μl min^{-1}) with a dissociation phase of 120 s. At the end of each interaction, the chip surface was regenerated by injection of 1 M NaCl (10 μl). The experiment was carried out in duplicate.

Competition assay. A new SA Chip was immobilized with the proU2-labelled fragment for the competition assay as described above. Protein concentration was fixed at 200 nM and this was

Table 1. DNA sequences for activity assays

DNA fragment	Sequence	Size (bp)	Molecular mass (Da)
proU1	−200 to +10 of *proU* promoter region*	221	136 604
proU2	−20 to +189 of *proU* promoter region*	221	136 615
BENT – $(A_6N_4)_{10}$	$(TTTTTTGCCG)_{10}$	112	69 231
NC – $(A_6N_9)_6$	$CCCCGGC(AAAAAACGGCCCGGC)_6$	109	67 400

*+1 is defined as the point of transcription initiation.

followed by preincubation with a 1-, 5-, 10-, 20-, 30- or 45-fold excess of non-biotinylated competitor DNAs (proU1, proU2, BENT, NC) in HBS-EP buffer for 15 min at room temperature prior to injection to the proU2-immobilized surface. Preincubated protein–DNA mixtures were injected in HBS-EP buffer (50 µl; 50 µl min^{-1}) with a dissociation phase of 120 s, and carried out in duplicate.

Complementation studies. A complementation plasmid based on the *M. tuberculosis Rv3852* gene was constructed by PCR amplification of the target sequence with primers HNS *Eco*RI (5′-CCGGAATTCATGCCAGACCCGCAGGATC-3′) and HNS *Hin*dIII (5′-GTAAGCTTTCAGCGGCGGCGCAGTTGCC-3′), which were designed to contain, respectively, *Eco*RI and *Hin*dIII restriction sites (underlined). The previously generated pET-23a(+)::*hns* expression vector was used as a template, and 10 % DMSO and native *Pfu* DNA polymerase (Stratagene) were employed under standard PCR conditions (cycling parameters were: an initial denaturation step at 96 °C for 5 min, followed by 30 cycles at 96 °C for 45 s, 55 °C for 45 s and 72 °C for 1 min, followed by a final extension step at 72 °C for 10 min). The 405 bp PCR product was gel-purified, cloned into the pCR-Blunt vector (Invitrogen) and subcloned into the pKK223-3 expression vector (Pharmacia) under the control of the strong *tac* promoter, resulting in pKH2 (pKK223-3::$Rv3852_{Mtb}$).

For the complementation assay, *E. coli* BSN26 (relevant genotype MC4100 *trp*::Tn*10*) and BSN27 (relevant genotype MC4100 *trp*::Tn*10* Δ*hns*) strains (Johansson *et al.*, 1998) were grown in LB medium or M63 minimal medium (MM) supplemented with 0.2 % glucose as a carbon source (Miller, 1992) and 40 µg L-serine ml^{-1}. Both strains were used as experimental controls. Plasmids pKK223-3 and pKH2 were separately electroporated into competent BSN27 cells, and transformants were isolated on LB agar plates containing 50 µg ampicillin ml^{-1}. Strains BSN26, BSN27, BSN27 (pKK223-3) and BSN27 (pKH2) were grown at 37 °C in 50 ml MM supplemented with serine (and 50 µg ampicillin ml^{-1}, where appropriate) in 250 ml flasks. Growth was monitored for serine sensitivity (from starting cultures that had reached OD_{600} 0.01) for up to 8 h. Three independent experiments were conducted for each strain under the above-mentioned conditions.

Data analysis. The dissociation constant for Rv3852 protein binding to immobilized proU2 DNA sequence was obtained by fitting the data to equation 1:

$$R_{eqi} = \frac{R_{max}[\text{Rv3852}]}{K_D + [\text{Rv3852}]} \quad (1)$$

where R_{eqi} is the equilibrium (maximum) resonance response unit for i concentration of Rv3852 protein (i=0, 0.05, 0.1, 0.5, 1, 2.5, 5 or 7.5 µM), R_{max} is the maximum resonance response unit at saturating protein concentrations, and K_D is the dissociation constant at equilibrium.

The stoichiometry for Rv3852 protein–proU2 complex formation was determined from fitting the data to equation 2 (Majka & Speck, 2007):

$$n = R_{max}\frac{MW_{DNA}}{R_{DNA}MW_{protein}} \quad (2)$$

where n is the number of protein molecules bound to DNA or the number of protein binding sites on DNA (assuming that a single protein molecule binds a single binding site), R_{max} is the resonance response for a saturating concentration of protein, R_{DNA} is the amount of immobilized DNA (RU), MW_{DNA} is the molecular mass of DNA (proU2=136 615 Da), and $MW_{protein}$ is the molecular mass of the Rv3852 protein from *M. tuberculosis* (13 822 Da).

Values for the monophasic apparent dissociation rate constants (single exponentials) were obtained by fitting the data to equation 3:

$$R(t) = R_0 e^{-k_{off}t} \quad (3)$$

where $R(t)$ is the resonance response at time t, R_0 is the resonance response at the analysis start point (which is not necessarily at the beginning of the dissociation phase), and k_{off} is the apparent dissociation rate constant.

The competition assay analysis was carried out as described by Teh *et al.* (2007). Briefly, the level of protein binding to immobilized proU2 was set at 100 %. The competition at a certain concentration of non-biotinylated competitor DNA was calculated in relation to the percentage binding value (% binding) as follows (equation 4):

$$\%\ \text{binding} = \left[\frac{R_i}{R_0}\right]100 \quad (4)$$

where R_i is the signal obtained at a certain concentration of competitor, and R_0 is the signal obtained in the absence of competitor DNA.

RESULTS AND DISCUSSION

Cloning, expression and purification of Rv3852

PCR amplification of the 405 bp corresponding to the Rv3852 coding sequence required the presence of 10 % DMSO, a co-solvent that improves denaturation of GC-rich DNA sequences (Pomp & Medrano, 1991), which is consistent with the high GC content of the *M. tuberculosis* H37Rv genome (Cole *et al.*, 1998). The amplified product was cloned into the pCR-Blunt cloning vector and subcloned into the pET-23a(+) expression vector between *Nde*I and *Hin*dIII restriction sites. Rv3852 recombinant protein was heterologously expressed in *E. coli* Rosetta (DE3) cells in the soluble fraction with no IPTG induction (data not shown). This unusual feature of the pET system has been verified in stationary phase for cells growing in the absence of appropriate inducer (Magalhães *et al.*, 2002; Mendonça *et al.*, 2007; Rizzi *et al.*, 2005; Silva *et al.*, 2003). It has been proposed that leaky protein expression in the pET system is part of a general cellular response to nutrition limitation when cells approach stationary phase, and that cyclic AMP, acetate and low pH are involved in derepression of the *lac* operon (Grossman *et al.*, 1998). However, it has more recently been shown that unintended induction of the pET system is due to the presence of as little as 0.0001 % lactose in the medium (Studier, 2005).

Rv3852 recombinant protein was purified to homogeneity by liquid chromatography. The purification protocol included 1 % streptomycin precipitation to reduce contaminant DNA. To obtain electrophoretically homogeneous protein, three chromatographic steps were required: SP-Sepharose Fast Flow, Sephacryl S-100 and Mono-S HR. The use of a cation-exchange column was based on the theoretical isoelectric point calculated for Rv3852, which is 11.35. The purification protocol for *M. tuberculosis* Rv3852 protein yielded ~3 mg homogeneous protein with an

apparent molecular mass of ~14 kDa from 15 g wet weight of cells (Fig. 1). The relatively low yield was probably due to the small amount of recombinant protein expressed in soluble form. Several *E. coli* strains were tested (data not shown) and the best result was obtained with *E. coli* Rosetta (DE3) cells. Even though we have no experimental evidence, it is tempting to propose that the pET vector suffers transcriptional repression by the expressed recombinant protein, since an inhibitory effect of H-NS-related proteins on the *lacUV5* promoter has been documented (Goransson *et al.*, 1990), as well as a likely deleterious effect on cell growth (Spurio *et al.*, 1997).

Protein characterization and amino acid sequence analysis of *M. tuberculosis* Rv3852

N-terminal sequencing of the homogeneous protein preparation unambiguously confirmed the identity of the first 23 aa and removal of the N-terminal methionine residue. The *M. tuberculosis* Rv3852 N-terminal region exhibits significant homology to eukaryotic H1 linker histones (Kasinsky *et al.*, 2001) due to the presence of proline, alanine and lysine residues clustered in tetrapeptide repeats (PAKK, KAAK). As seen in the multiple sequence alignment (Fig. 2), four PAKK sequences occur in both *M. tuberculosis* and *Mycobacterium bovis* (the causative agent of bovine TB), whereas the *Mycobacterium marinum* (a pathogen of fish and amphibia) and *Mycobacterium ulcerans* (the causative agent of Buruli ulcer, a severe skin disease) homologues possess seven PAKK sequences in their proteins. These repetitive amino acid sequences are known to bind DNA with high affinity, and it is interesting to note that a very similar protein organization has been described for the Hlp protein from *M. tuberculosis*, which has been characterized as a histone-like protein (Prabhakar *et al.*, 1998). However, the typical tetrapeptide repeats occur at the C-terminal domain of *M. tuberculosis* Hlp, with the N-terminal domain showing considerable sequence homology to HU proteins. *M. smegmatis* Hlp is organized according to the same pattern (Shires & Steyn, 2001). It has also been demonstrated that *M. smegmatis* Hlp lacking the entire C-terminal domain binds poorly to DNA and that the C-terminal repeats are responsible for DNA end-joining (Mukherjee *et al.*, 2008). Therefore, these important similarities point to a probable involvement of the *M. tuberculosis* Rv3852 protein in DNA binding and a structural and/or regulatory role in the mycobacterial chromosome.

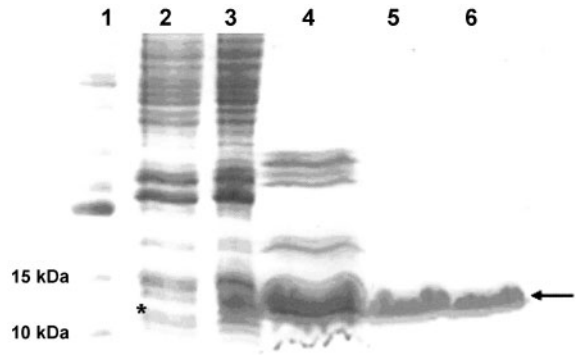


Fig. 1. SDS-PAGE analysis of pooled fractions from the recombinant protein purification protocol. Lanes: 1, BenchMark protein ladder (Invitrogen); 2, crude extract, soluble fraction; 3, crude extract after streptomycin sulfate precipitation and centrifugation; 4, after purification by SP Sepharose FF column; 5, after purification by Sephacryl S-100 column; 6, after purification by Mono S HR column. Homogeneous Rv3852 protein was visualized at 14 kDa (arrow). The asterisk marks the Rv3852 protein in soluble form in the crude extract.

Rv3852 from *M. tuberculosis* exhibits 30 alanine, 8 arginine, 16 lysine and 19 proline residues in its amino acid sequence (Cole *et al.*, 1998), with a calculated isoelectric point of 11.34. As recently described for another *M. tuberculosis* NAP, Lsr2, Rv3852 could make use of this high proportion of positively charged arginine and lysine residues to form salt bridges with the negative phosphate backbone of DNA (Chen *et al.*, 2008). Moreover, it is tempting to hypothesize that the lysine residues present in Rv3852 serve as acetylation sites, leading to potential mechanisms of differential gene regulation, since lysine acetylation has been considered an important, though underestimated, regulatory post-translational modification in bacteria (Zhang *et al.*, 2009).

To evaluate the native oligomeric state of Rv3852, we performed gel filtration on a calibrated Superdex 200 HR column. Elution of a single peak corresponding to a protein of ~25 200 ± 2520 Da indicated that the *M. tuberculosis* Rv3852 protein is a dimer in solution (subunit molecular mass is 13 822 Da), resembling other histone-like proteins (Dorman & Deighan, 2003).

DNA sequences selected for Rv3852 binding assays

Four DNA sequences were constructed to determine the DNA-binding activity of homogeneous Rv3852 protein (Table 1). We performed activity assays using two sequences of the *proU* operon from *M. tuberculosis* H37Rv. This operon, which encodes a transport system for the osmoprotectant glycine betaine and is induced by high osmolarity (Higgins *et al.*, 1987), is one of the most studied systems whose transcription is affected by H-NS in other pathogenic micro-organisms. Several reports have demonstrated that a sequence element (DRE), located downstream of the *proU* promoter (usually between +24 and +202), is preferred for binding by H-NS (Lucht *et al.*, 1994; Owen-Hughes *et al.*, 1992; Tupper *et al.*, 1994). Based upon these studies, we constructed two sequences, proU1 and proU2, that correspond to −200 to +10 and −20 to +189 of the *proU* promoter from *M. tuberculosis*, respectively, in order to confirm binding activity to these

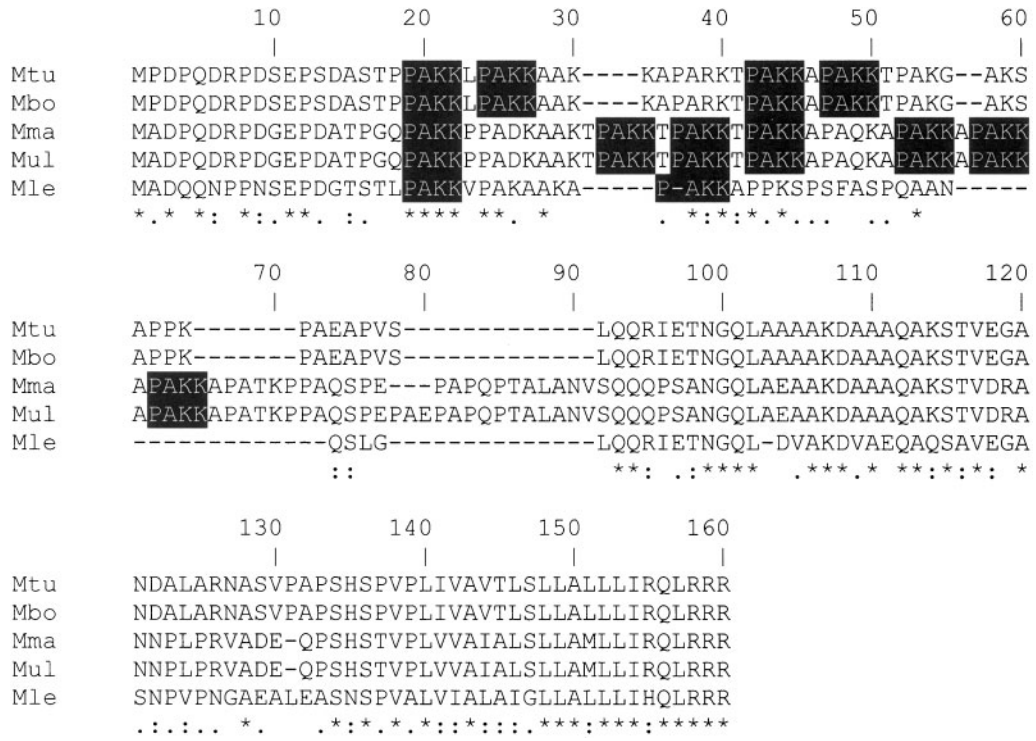


Fig. 2. Multiple sequence alignment of *M. tuberculosis* Rv3852 with its mycobacterial homologues, showing the PAKK repeats in the N-terminal region. *M. tuberculosis* (Mtu) histone-like protein Rv3852 was aligned with other potential mycobacterial homologues from *M. bovis* (Mbo), *M. marinum* (Mma), *M. ulcerans* (Mul) and *Mycobacterium leprae* (Mle) using the program CLUSTAL W (Thompson *et al.*, 1994). Tetrapeptide PAKK repeats (which are present in eukaryotic H1 linker histones and are known to bind DNA) are highlighted in white type on a black background. Identical conserved residues are indicated by asterisks below the alignment. Strongly similar and weakly similar residues are identified by colons and full stops, respectively.

regions by Rv3852 recombinant protein. The latter sequence is located at the beginning of the *proX* structural gene, the first gene of the *proU* operon (*proXVWZ*). Two other sequences were employed in DNA-binding assays: BENT, which is a curved sequence containing six contiguous A–T residues in phase (10 bp period), and NC, a noncurved sequence in which the adenine tracts are out of phase (15 bp period) (Koo *et al.*, 1986; Shimizu *et al.*, 1995). All sequences were cloned, excised from the agarose gel and purified prior to activity assays.

Activity of Rv3852 homogeneous protein by EMSAs

EMSA measurements at varying protein amounts (0–1 μg) showed that the Rv3852 protein had a similar relative affinity for all the sequences tested (Fig. 3). In spite of the preference of Rv3852 for curved DNA, some histone-like proteins can bind and bend noncurved DNA (Dame & Goosen, 2002; Dorman & Deighan, 2003; Schneider *et al.*, 2001; Spurio *et al.*, 1997; Swinger & Rice, 2004). We also verified that the protein is able to form multiple protein–DNA complexes, and that this process is concentration-dependent, suggesting the formation of higher forms of the dimeric protein when bound to DNA. To confirm that DNA binding was a particular property of the recombinant *M. tuberculosis* Rv3852 protein, two other *M. tuberculosis* homogeneous proteins were tested, InhA (Oliveira *et al.*, 2006) and PNP (Silva *et al.*, 2003). They showed no DNA-binding interaction under the same experimental conditions described for the Rv3852 protein assay (data not shown), thereby showing that this protein is indeed a DNA-binding protein.

Activity assay of Rv3852 homogeneous protein by SPR

Rv3852 protein affinity for proU1, proU2, BENT and NC DNA sequences was assessed by SPR using BIA-Core equipment. We immobilized 80 RU (resonance response units) of the proU2 DNA sequence onto an SA Sensor Chip (GE Healthcare). A low RU value (80) and high flow rate (50 μl min^{-1}) were employed to reduce mass transport effects that can limit the assay (Majka & Speck, 2007). Sensorgrams for varying protein concentrations (0, 0.05, 0.1, 0.5, 1, 2.5, 5 and 7.5 μM) in HBS-EP buffer were obtained in duplicate (Fig. 4, inset shows a typical sensorgram). Average values for equilibrium resonance

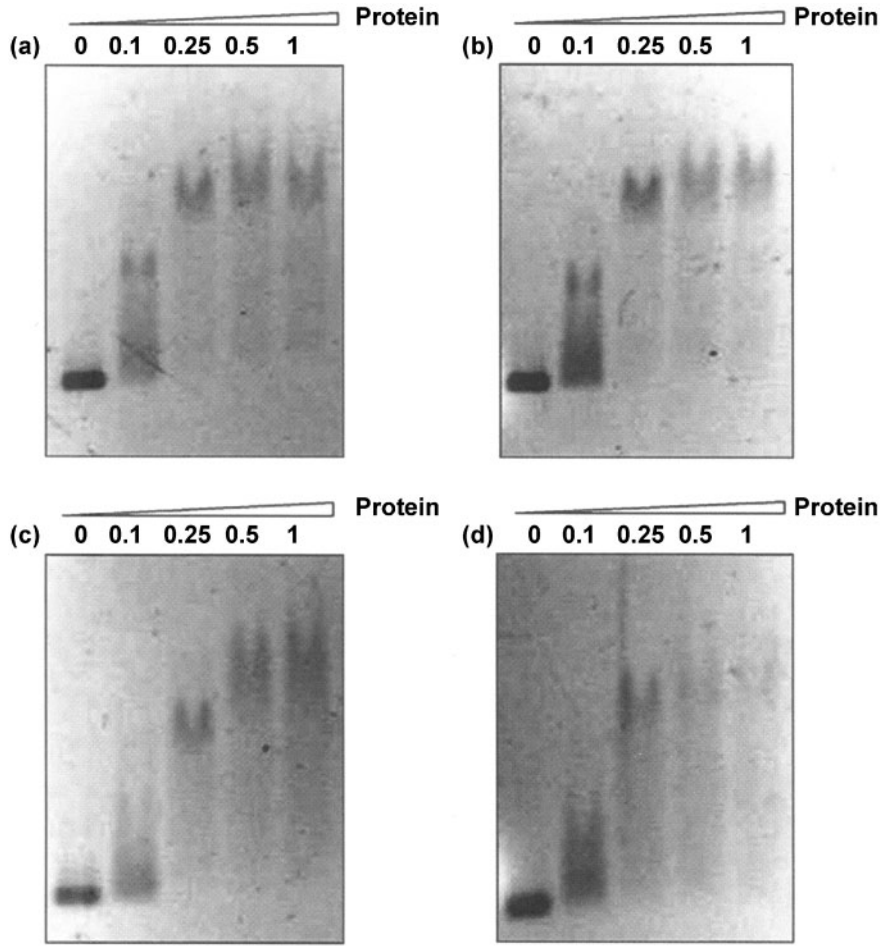


Fig. 3. Activity assay by EMSA. (a) Rv3852 protein–proU1 binding assay, (b) Rv3852 protein–proU2 binding assay, (c) Rv3852 protein–BENT binding assay, (d) Rv3852 protein–NC binding assay. All DNA sequences were tested with the following protein amounts: 0, 0.1, 0.25, 0.5 and 1 μg.

responses were fitted to Equation 1, yielding a dissociation constant (K_D) value of 289 (±84) nM and a saturation response (R_{max}) value of 80.43 (±4.67) RU for Rv3852 protein–proU2 complex formation (Fig. 4). The DNA–protein complex stoichiometry could also be estimated as ~10 molecules of protein monomers per molecule of DNA by fitting the data to Equation 2 (Majka & Speck, 2007).

The binding process was characterized by a fast association phase that did not allow reliable estimates to be obtained for the association rate constants. However, the dissociation phase was assessed using a delay of 120 s after the end of sample injection. The monophasic apparent dissociation rate constant (Fig. 5, inset) values were obtained by fitting data to Equation 3. These values were plotted against their respective protein concentrations and, fitting the data to a rectangular hyperbola that intersects at a finite value on the y axis, two limiting values could be determined: 0.0189 (±0.0010) s^{-1} at very low protein concentrations (approaching zero), and 0.0415 (±0.0020) s^{-1} at saturating protein concentrations (Fig. 5). It is thus tempting to propose that the value of 0.0189 s^{-1} is for the dissociation rate constant of the dimeric protein and the value of 0.0415 s^{-1} is for the dissociation rate constant of the oligomeric protein. We propose that the dimers of the recombinant Rv3852 protein are able to self-associate in higher oligomers along the DNA molecules and that this property is concentration-dependent, as we could verify by EMSA. We thus propose that higher oligomerization of the Rv3852 protein results in faster dissociation rate constant values.

Competition assay

To investigate the relative affinity for the DNA sequences evaluated in our studies, we immobilized a new SA Sensor Chip and carried out a competition assay. A value of 110 RU was obtained for the 221 bp proU2 DNA. Considering that 0.78 ng of a DNA molecule bound at

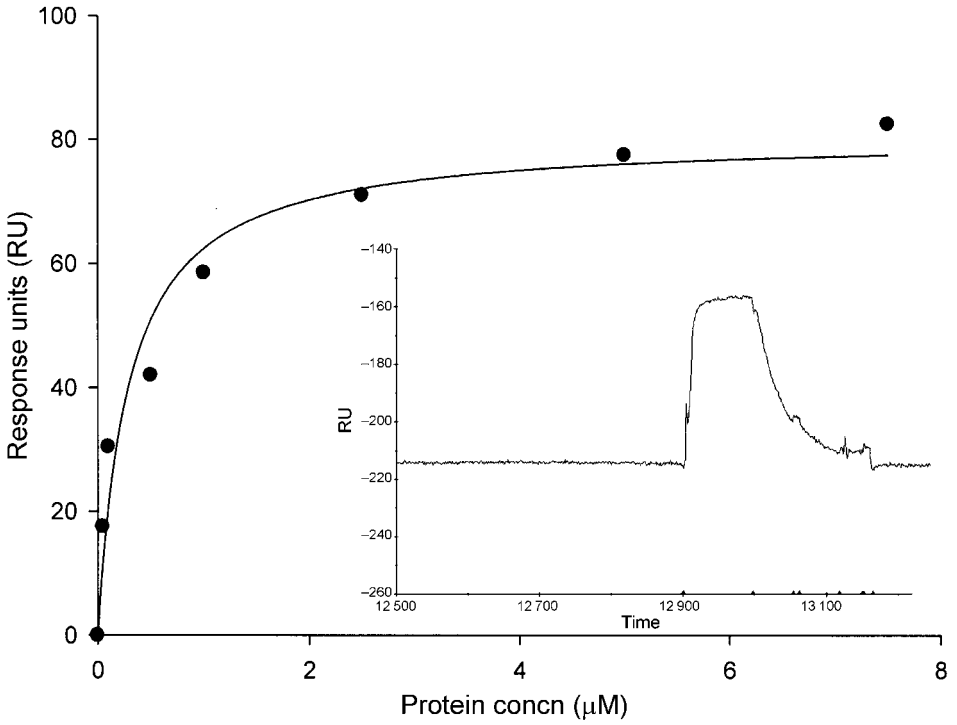


Fig. 4. Saturation curve of the Rv3852 protein–proU2 complex. Mean RU values were plotted against the corresponding protein concentrations and the analysis was performed using the SigmaPlot program. Fitting the data to a rectangular hyperbola (Equation 1) yielded values of 289 (±84) nM for the equilibrium dissociation constant (K_D) and of 80.43 (±4.67) RU for the maximum resonance at saturating protein concentrations (R_{max}). The inset shows an illustrative sensorgram obtained with 1 μM protein.

the surface gives a response of 1000 RU (Bouffartigues *et al.*, 2007), we can expect ~0.086 ng or 6.3×10^{-16} mol of proU2 DNA (136 615 g mol^{-1}) distributed within the dextran layer. Protein concentration was fixed at 200 nM and was incubated with a 1-, 5-, 10-, 20-, 30- or 45-fold excess of non-biotinylated competitor DNAs (proU1, proU2, BENT and NC). This competition assay approach allows assessment of the relative affinity of a large number of DNA sequences with one surface preparation (Teh *et al.*, 2007).

Among the tested DNA sequences, fitting the data to Equation 4, proU1 showed the most pronounced reduction in Rv3852 protein binding to immobilized proU2, indicating that this sequence has the strongest affinity for the recombinant protein as compared with the other sequences tested (Fig. 6). We found the following affinity order: proU1>proU2>BENT>NC. The competition assay showed that proU1 has a slightly higher affinity than proU2 for recombinant *M. tuberculosis* Rv3852 protein. It is likely that this observed preference for proU1 is because it corresponds to the promoter region of the *proU* operon, in contrast to proU2, which corresponds to a sequence of the first gene of the operon. The preference of the protein for these two sequences as compared with BENT and NC can be explained by additional structural characteristics, besides DNA curvature, that may favour DNA–protein interactions. This experiment also demonstrated that the recombinant protein shows a slight preference for the curved sequence rather than the noncurved. This can be explained by the fact that the DNA-binding protein expends less energy to bind a curved sequence, since to exert its activity it first needs to bend a noncurved sequence. *E. coli* H-NS protein binds with relatively high affinity to DNA of any sequence (Dorman, 2007), although, as mentioned above, a well-defined preferential DNA-binding site has been reported (Bouffartigues *et al.*, 2007). The competition results presented here are in agreement with this.

Lack of functional complementation of an *E. coli hns* mutant by *M. tuberculosis* Rv3852

Since we initially did not know whether Rv3852 would represent an authentic H-NS-like homologue or another type of NAP, we performed a complementation assay to test the ability of the *M. tuberculosis* protein to restore an *hns* mutant phenotype. *In vivo* complementation of serine susceptibility of *E. coli hns* mutants has been commonly employed to identify H-NS-like proteins in a number of bacteria (Rodríguez *et al.*, 2005; Tendeng & Bertin, 2003).

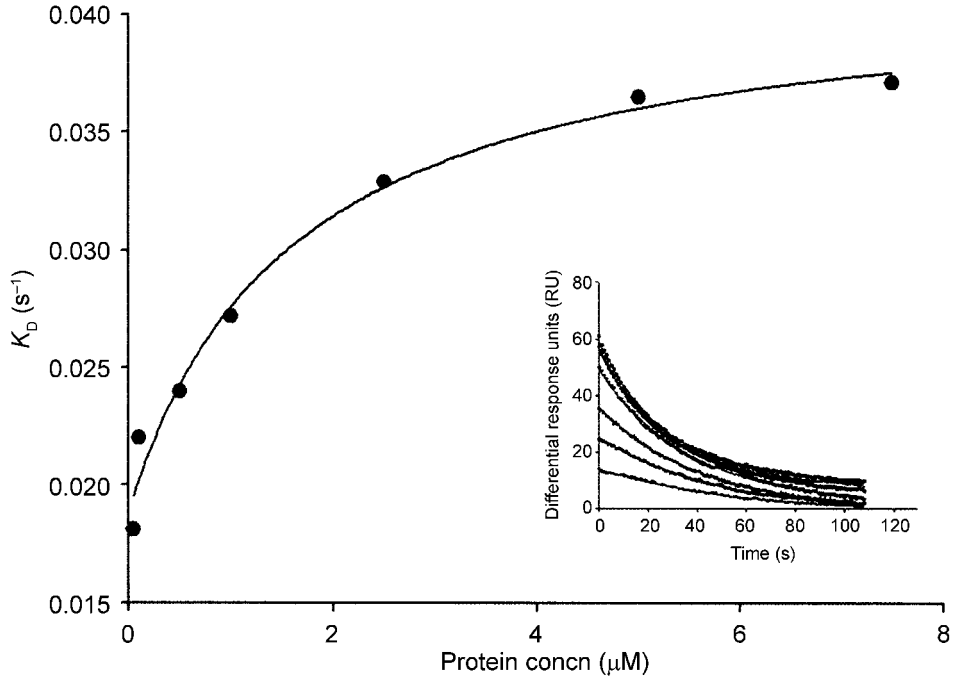


Fig. 5. Rv3852 protein dissociation rate constants for the proU2-immobilized DNA. The data were obtained by SPR from a delay of 120 s after the end of sample injection. Each apparent dissociation rate constant was plotted against its corresponding protein concentration. The data were fitted to a rectangular hyperbola that intersects at a finite value on the *y* axis by SigmaPlot, from which we can determine two limiting values for the dissociation rate constants: 0.0189 s^{-1} for dimeric protein at concentrations approaching zero, and 0.0415 s^{-1} for oligomeric protein at saturating concentrations. Inset: traces showing the monophasic apparent dissociation rate constants for varying protein concentrations (0.05–7.5 μM).

Accordingly, a pKK223-3-based complementation plasmid containing the Rv3852-coding sequence (pKH2) cloned under control of the strong *tac* promoter was constructed. Plasmids pKK223-3 (control) and pKH2 were separately electroporated into BSN27 (Δ*hns*) cells. *E. coli* BSN26 (wild-type) and BSN27 not transformed with any plasmid were included as controls. As expected, BSN26 cells grew at a higher rate than the BSN27 mutant (which exhibits a

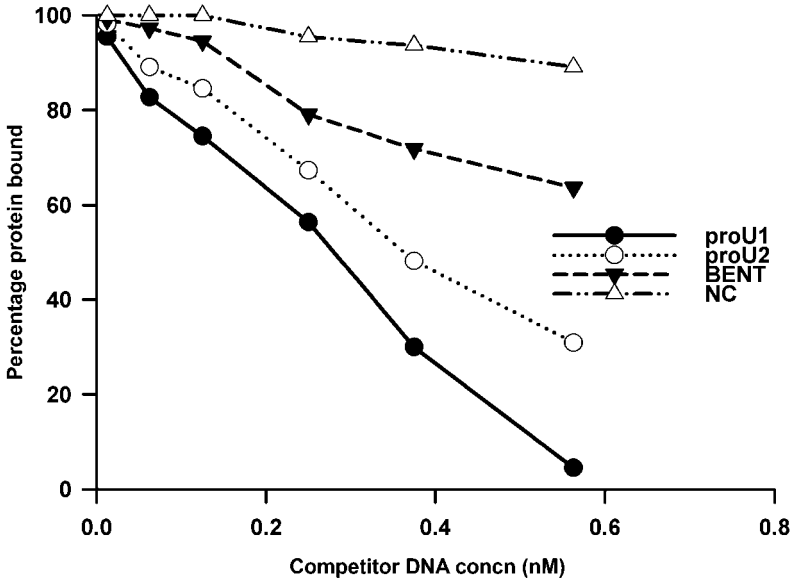


Fig. 6. Competition profile of the Rv3852 protein binding to the proU2-immobilized DNA in the presence of competitor DNAs. The protein binding signal in the absence of competitor DNAs was set as 100 %. Protein binding was tested at a 1-, 5-, 10-, 20-, 30- or 45-fold excess of competitor DNA.

severe growth defect), but no difference between the latter and BSN27 (pKK223-3) or BSN27 (pKH2) transformants could be observed under the conditions tested (data not shown), demonstrating that *M. tuberculosis* Rv3852 does not correspond to a 'classical' H-NS-like protein, and thus reinforcing the probability that it represents a novel NAP of unknown function.

Expression of *Rv3852* has been shown to be altered under some stress conditions. In a nutrient starvation model in *M. tuberculosis*, *Rv3852* mRNA levels have been shown to be upregulated after 4 h of starvation (Betts *et al.*, 2002). In another study, in which the heat-shock response of *M. tuberculosis* was analysed using an *hspR* (which encodes the transcriptional repressor HspR) deletion mutant, *Rv3852* expression was also upregulated compared with the wild-type strain (Stewart *et al.*, 2002). These data suggest that, as exemplified by other NAPs, transcription of *Rv3852* is most probably regulated by environmental stresses, such as nutrient availability, heat- and cold-shock responses, and oxidative damage.

Conclusion

In this work, we present the initial characterization of the putative ORF for an NAP (Rv3852) from *M. tuberculosis*. Its amino acid sequence in the N-terminal region shows the presence of tetrapeptide repeats that are found in H1 histones, and this feature, which is similar to that of another histone-like protein from *M. tuberculosis*, MDP1, suggests that the protein encoded by the ORF Rv3852 is involved in the structuring of the mycobacterial DNA and similar functions, contributing to the molecular machinery of the TB bacillus. We have shown that the Rv3852 recombinant protein binds to the *proU* promoter region, an important environmentally controlled system that has been shown to be regulated by H-NS in other pathogenic bacteria. The dimeric structure of the protein is suggested to self-associate to form higher oligomers, as indicated by the multiple protein–DNA complexes observed by EMSA. A stoichiometry of 10 molecules monomeric protein per molecule of DNA was determined by SPR. Analysis by SPR of dissociation rate constant values as a function of Rv3852 concentration suggests that the dimeric protein dissociates from the proU2 DNA sequence more slowly than the oligomeric protein. Complementation studies confirmed that the ORF annotated as a possible H-NS protein from *M. tuberculosis* does not correspond to this protein. Recent reports have demonstrated that the Lrs2 protein has H-NS activity in *M. tuberculosis*, based on its ability to complement several phenotypes from an *E. coli hns* mutant (Gordon *et al.*, 2008). We therefore propose that the recombinant protein characterized here has regulatory activity(ies) related to adaptation or even virulence genes, as have several other described NAPs. Notwithstanding these findings, these issues will need to be clarified in future research.

ACKNOWLEDGEMENTS

Financial support for this work was provided by the National Institute of Science and Technology on Tuberculosis (Decit/SCTIE/MS-MCT-CNPq-FNDCT-CAPES) and the Millennium Initiative Program MCT-CNPq Ministry of Health – Secretary of Science, Technology and Strategic Materials (SCTIE/DECIT, Brazil) to D. S. S. and L. A. B. D. S. S. also acknowledges MCT/CNPq 02/2006-Universal for financial support. D. S. S. (CNPq, 304051/1975-06), L. A. B. (CNPq, 520182/99-5) and M. S. P. (500079/90-0) are research career awardees from CNPq. I. C. R. W. was supported by a PhD scholarship from CNPq. C. Z. S. was funded by BNDES. J. D. M. was supported by CNPq. We are grateful to Dr Antonio Juárez and Dr Cristina Madrid, both of the Universidad de Barcelona, for kindly providing the *E. coli* BSN26 and BSN27 strains. We thank Rafael G. Silva for helpful discussions on the BIA-Core assays.

REFERENCES

Andersen, P. (2007). Tuberculosis vaccines – an update. *Nat Rev Microbiol* **5**, 484–487.

Atlung, T. & Ingmer, H. (1997). H-NS: a modulator of environmentally regulated gene expression. *Mol Microbiol* **24**, 7–17.

Bailly, C., Waring, M. J. & Travers, A. A. (1995). Effects of base substitutions on the binding of a DNA-bending protein. *J Mol Biol* **253**, 1–7.

Beloin, C. & Dorman, C. J. (2003). An extended role for the nucleoid structuring protein H-NS in the virulence gene regulatory cascade of *Shigella flexneri*. *Mol Microbiol* **47**, 825–838.

Berlutti, F., Casalino, M., Zagaglia, C., Fradiani, P. A., Visca, P. & Nicoletti, M. (1998). Expression of the virulence plasmid-carried apyrase gene (*apy*) of enteroinvasive *Escherichia coli* and *Shigella flexneri* is under the control of H-NS and the VirF and VirB regulatory cascade. *Infect Immun* **66**, 4957–4964.

Betts, J. C., Lukey, P. T., Robb, L. C., McAdam, R. A. & Duncan, K. (2002). Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol Microbiol* **43**, 717–731.

Bouffartigues, E., Leh, H., Anger-Leroy, M., Rimsky, S. & Buckle, M. (2007). Rapid coupling of surface plasmon resonance (SPR and SPRi) and ProteinChip based mass spectrometry for the identification of proteins in nucleoprotein interactions. *Nucleic Acids Res* **35**, e39.

CDC (2007). Extensively drug-resistant tuberculosis – United States, 1993–2006. *MMWR Morb Mortal Wkly Rep* **56**, 250–253.

Chen, J. M., German, G. J., Alexander, D. C., Ren, H., Tran, T. & Liu, J. (2006). Roles of Lsr2 in colony morphology and biofilm formation of *Mycobacterium smegmatis*. *J Bacteriol* **188**, 633–641.

Chen, J. M., Ren, H., Shaw, J. E., Wang, Y. J., Li, M., Leung, A. S., Tran, V., Berbenetz, N. M., Kocíncová, D. & other authors (2008). Lsr2 of *Mycobacterium tuberculosis* is a DNA-bridging protein. *Nucleic Acids Res* **36**, 2123–2135.

Colangeli, R., Helb, D., Vilchèze, C., Hazbón, M. H., Lee, C.-G., Safi, H., Sayers, B., Sardone, I., Jones, M. B. & other authors (2007). Transcriptional regulation of multi-drug tolerance and antibiotic-induced responses by the histone-like protein Lsr2 in *M. tuberculosis*. *PLoS Pathog* **3**, e87.

Colangeli, R., Haq, A., Arcus, V. L., Summers, E., Magliozzo, R. S., McBride, A., Mitra, A. K., Radjainia, M., Khajo, A. & other authors (2009). The multifunctional histone-like protein Lsr2 protects mycobacteria against reactive oxygen intermediates. *Proc Natl Acad Sci U S A* **106**, 4414–4418.

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S. & other authors (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537–544.

Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honore, N., Garnier, T., Churcher, C. & other authors (2001). Massive gene decay in the leprosy bacillus. *Nature* **409**, 1007–1011.

Dame, R. T. (2005). The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol Microbiol* **56**, 858–870.

Dame, R. T. & Goosen, N. (2002). HU: promoting or counteracting DNA compaction? *FEBS Lett* **529**, 151–156.

Dhavan, G. M., Crothers, D. M., Chance, M. R. & Brenowitz, M. (2002). Concerted binding and bending of DNA by *Escherichia coli* integration host factor. *J Mol Biol* **315**, 1027–1037.

Dorman, C. J. (2004). H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol* **2**, 391–400.

Dorman, C. J. (2007). H-NS, the genome sentinel. *Nat Rev Microbiol* **5**, 157–161.

Dorman, S. E. & Chaisson, R. E. (2007). From magic bullets back to the magic mountain: the rise of extensively drug-resistant tuberculosis. *Nat Med* **13**, 295–298.

Dorman, C. J. & Deighan, P. (2003). Regulation of gene expression by histone-like proteins in bacteria. *Curr Opin Genet Dev* **13**, 179–184.

Falconi, M., Prosseda, G., Giangrossi, M., Beghetto, E. & Colonna, B. (2001). Involvement of Fis in the H-NS-mediated regulation of *virF* gene of *Shigella* and enteroinvasive *Escherichia coli*. *Mol Microbiol* **42**, 439–452.

Franzon, J. H. & Santos, D. S. (2004). A role for histone-like protein H1 (H-NS) in the regulation of hemolysin expression by *Serratia marcescens*. *Braz J Med Biol Res* **37**, 1763–1769.

Gandotra, S., Schnappinger, D., Monteleone, M., Hillen, W. & Ehrt, S. (2007). *In vivo* gene silencing identifies the *Mycobacterium tuberculosis* proteasome as essential for the bacteria to persist in mice. *Nat Med* **13**, 1515–1520.

Gomez, J. E. & McKinney, J. D. (2004). *M. tuberculosis* persistence, latency, and drug tolerance. *Tuberculosis* (*Edinb*) **84**, 29–44.

Goransson, M., Sonden, B., Nilsson, P., Dagberg, B., Forsman, K., Emanuelsson, K. & Uhlin, B. E. (1990). Transcriptional silencing and thermoregulation of gene expression in *Escherichia coli*. *Nature* **344**, 682–685.

Gordon, B. R. G., Imperial, R., Wang, L., Navarre, W. W. & Liu, J. (2008). Lsr2 of *Mycobacterium* represents a novel class of H-NS-like proteins. *J Bacteriol* **190**, 7052–7059.

Grossman, T. H., Kawasaki, E. S., Punreddy, S. R. & Osburne, M. S. (1998). Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene* **209**, 95–103.

Harrison, J. A., Pickard, D., Higgins, C. F., Khan, A., Chatfield, S. N., Ali, T., Dorman, C. J., Hormaeche, C. E. & Dougan, G. (1994). Role of *hns* in the virulence phenotype of pathogenic salmonellae. *Mol Microbiol* **13**, 133–140.

Higgins, C. F., Cairney, J., Stirling, D. A., Sutherland, L. & Booth, I. R. (1987). Osmotic regulation of gene expression: ionic strength as an intracellular signal? *Trends Biochem Sci* **12**, 339–344.

Hingley-Wilson, S. M., Sambandamurthy, V. K. & Jacobs, W. R., Jr (2003). Survival perspectives from the world's most successful pathogen, *Mycobacterium tuberculosis*. *Nat Immunol* **4**, 949–955.

Hommais, F., Krin, E., Laurent-Winter, C., Soutourina, O., Malpertuy, A., Le Caer, J. P., Danchin, A. & Bertin, P. (2001). Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. *Mol Microbiol* **40**, 20–36.

Johansson, J., Dagberg, B., Richet, E. & Uhlin, B. E. (1998). H-NS and StpA proteins stimulate expression of the maltose regulon in *Escherichia coli*. *J Bacteriol* **180**, 6117–6125.

Kasinsky, H. E., Lewis, J. D., Dacks, J. B. & Ausió, J. (2001). Origin of H1 linker histones. *FASEB J* **15**, 34–42.

Koh, J., Saecker, R. M. & Record, M. T., Jr (2008). DNA binding mode transitions of *Escherichia coli* $HU_{\alpha\beta}$: evidence for formation of a bent DNA–protein complex on intact, linear duplex DNA. *J Mol Biol* **383**, 324–346.

Koo, H. S., Wu, H. M. & Crothers, D. M. (1986). DNA bending at adenine.thymine tracts. *Nature* **320**, 501–506.

Krylov, A. S., Zasedateleva, O. A., Prokopenko, D. V., Rouviere-Yaniv, J. & Mirzabekov, A. D. (2001). Massive parallel analysis of the binding specificity of histone-like protein HU to single- and double-stranded DNA with generic oligodeoxyribonucleotide microchips. *Nucleic Acids Res* **29**, 2654–2660.

Lee, B. H., Murugasu-Oei, B. & Dick, T. (1998). Upregulation of a histone-like protein in dormant *Mycobacterium smegmatis*. *Mol Gen Genet* **260**, 475–479.

Lewin, A., Baus, D., Kamal, E., Bon, F., Kunisch, R., Maurischat, S., Adonopoulou, M. & Eich, K. (2008). The mycobacterial DNA-binding protein 1 (MDP1) from *Mycobacterium bovis* BCG influences various growth characteristics. *BMC Microbiol* **8**, 91.

Lucht, J. M., Dersch, P., Kempf, B. & Bremer, E. (1994). Interactions of the nucleoid-associated DNA-binding protein H-NS with the regulatory region of the osmotically controlled *proU* operon of *Escherichia coli*. *J Biol Chem* **269**, 6578–6586.

Luijsterburg, M. S., Noom, M. C., Wuite, G. J. L. & Dame, R. T. (2006). The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J Struct Biol* **156**, 262–272.

Luijsterburg, M. S., White, M. F., van Driel, R. & Dame, R. T. (2008). The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol* **43**, 393–418.

Magalhães, M. L. B., Pereira, C. P., Basso, L. A. & Santos, D. S. (2002). Cloning and expression of functional shikimate dehydrogenase (EC 1.1.1.25) from *Mycobacterium tuberculosis* H37Rv. *Protein Expr Purif* **26**, 59–64.

Majka, J. & Speck, C. (2007). Analysis of protein–DNA interactions using surface plasmon resonance. *Adv Biochem Eng Biotechnol* **104**, 13–36.

Mangan, M. W., Lucchini, S., Danino, V., Cróinín, T. O., Hinton, J. C. D. & Dorman, C. J. (2006). The integration host factor (IHF) integrates stationary-phase and virulence gene expression in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol* **59**, 1831–1847.

Mendonça, J. D., Ely, F., Palma, M. S., Frazzon, J., Basso, L. A. & Santos, D. S. (2007). Functional characterization by genetic complementation of *aroB*-encoded dehydroquinate synthase from *Mycobacterium tuberculosis* H37Rv and its heterologous expression and purification. *J Bacteriol* **189**, 6246–6252.

Miller, J. H. (1992). *A Short Course in Bacterial Genetics: a Laboratory Manual and Handbook for Escherichia coli and Related Bacteria*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Mukherjee, A., Bhattacharyya, G. & Grove, A. (2008). The C-terminal domain of HU-related histone-like protein Hlp from *Mycobacterium smegmatis* mediates DNA end-joining. *Biochemistry* **47**, 8744–8753.

Nye, M. B., Pfau, J. D., Skorupski, K. & Taylor, R. K. (2000). *Vibrio cholerae* H-NS silences virulence gene expression at multiple steps in the ToxR regulatory cascade. *J Bacteriol* **182**, 4295–4303.

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

Owen-Hughes, T. A., Pavitt, G. D., Santos, D. S., Sidebotham, J. M., Hulton, C. S. J., Hinton, J. C. D. & Higgins, C. F. (1992). The chromatin-associated protein H-NS interacts with curved DNA and influences DNA topology and gene expression. *Cell* **71**, 255–265.

Pérez-Martín, J., Rojo, F. & de Lorenzo, V. (1994). Promoters responsive to DNA bending: a common theme in prokaryotic gene expression. *Microbiol Rev* **58**, 268–290.

Pomp, D. & Medrano, J. F. (1991). Organic solvents as facilitators of polymerase chain reaction. *Biotechniques* **10**, 58–59.

Prabhakar, S., Annapurna, P. S., Jain, N. K., Dey, A. B., Tyagi, J. S. & Prasad, H. K. (1998). Identification of an immunogenic histone-like protein (HLP_{Mt}) of *Mycobacterium tuberculosis*. *Tuber Lung Dis* **79**, 43–53.

Raviglione, M. C. (2003). The TB epidemic from 1992 to 2002. *Tuberculosis* (*Edinb*) **83**, 4–14.

Rimsky, S., Zuber, F., Buckle, M. & Buc, H. (2001). A molecular mechanism for the repression of transcription by the H-NS protein. *Mol Microbiol* **42**, 1311–1323.

Rizzi, C., Frazzon, J., Ely, F., Weber, P. G., Fonseca, I. O., Gallas, M., Oliveira, J. S., Mendes, M., Souza, B. M. & other authors (2005). DAHP synthase from *Mycobacterium tuberculosis* H37Rv: cloning, expression, and purification of functional enzyme. *Protein Expr Purif* **40**, 23–30.

Rodríguez, S., Nieto, J. M., Madrid, C. & Juárez, A. (2005). Functional replacement of the oligomerization domain of H-NS by the Hha protein of *Escherichia coli*. *J Bacteriol* **187**, 5452–5459.

Saunders, B. M. & Britton, W. J. (2007). Life and death in the granuloma: immunopathology of tuberculosis. *Immunol Cell Biol* **85**, 103–111.

Schechter, L. M., Jain, S., Akbar, S. & Lee, C. A. (2003). The small nucleoid-binding proteins H-NS, HU, and Fis affect *hilA* expression in *Salmonella enterica* serovar Typhimurium. *Infect Immun* **71**, 5432–5435.

Schnappinger, D., Schoolnik, G. K. & Ehrt, S. (2006). Expression profiling of host pathogen interactions: how *Mycobacterium tuberculosis* and the macrophage adapt to one another. *Microbes Infect* **8**, 1132–1140.

Schneider, R., Lurz, R., Luder, G., Tolksdorf, C., Travers, A. & Muskhelishvili, G. (2001). An architectural role of the *Escherichia coli* chromatin protein FIS in organising DNA. *Nucleic Acids Res* **29**, 5107–5114.

Scopes, R. K. (1994). *Protein Purification: Principles and Practice.* New York: Springer.

Shimizu, M., Miyake, M., Kanke, F., Matsumoto, U. & Shindo, H. (1995). Characterization of the binding of HU and IHF, homologous histone-like proteins of *Escherichia coli*, to curved and uncurved DNA. *Biochim Biophys Acta* **1264**, 330–336.

Shires, K. & Steyn, L. (2001). The cold-shock stress response in *Mycobacterium smegmatis* induces the expression of a histone-like protein. *Mol Microbiol* **39**, 994–1009.

Silva, R. G., Carvalho, L. P. S., Oliveira, J. S., Pinto, C. A., Mendes, M. A., Palma, M. S., Basso, L. A. & Santos, D. S. (2003). Cloning, expression and purification of functional human purine nucleoside phosphorylase. *Protein Expr Purif* **27**, 158–164.

Spurio, R., Falconi, M., Brandi, A., Pon, C. L. & Gualerzi, C. O. (1997). The oligomeric structure of nucleoid protein H-NS is necessary for recognition of intrinsically curved DNA and for DNA bending. *EMBO J* **16**, 1795–1805.

Stewart, G. R., Wernisch, L., Stabler, R., Mangan, J. A., Hinds, J., Laing, K. G., Young, D. B. & Butcher, P. D. (2002). Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. *Microbiology* **148**, 3129–3138.

Stinear, T. P., Seemann, T., Harrison, P. F., Jenkin, G. A., Davies, J. K., Johnson, P. D. R., Abdellah, Z., Arrowsmith, C. & other authors (2008). Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. *Genome Res* **18**, 729–741.

Stonehouse, E., Kovacikova, G., Taylor, R. K. & Skorupski, K. (2008). Integration host factor positively regulates virulence gene expression in *Vibrio cholerae*. *J Bacteriol* **190**, 4736–4748.

Studier, F. W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif* **41**, 207–234.

Swinger, K. K. & Rice, P. A. (2004). IHF and HU: flexible architects of bent DNA. *Curr Opin Struct Biol* **14**, 28–35.

Teh, H. F., Peh, W. Y. X., Su, X. & Thomsen, J. S. (2007). Characterization of protein–DNA interactions using surface plasmon resonance spectroscopy with various assay schemes. *Biochemistry* **46**, 2127–2135.

Tendeng, C. & Bertin, P. N. (2003). H-NS in Gram-negative bacteria: a family of multifaceted proteins. *Trends Microbiol* **11**, 511–518.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Tupper, A. E., Owen-Hughes, T. A., Ussery, D. W., Santos, D. S., Ferguson, D. J. P., Sidebotham, J. M., Hinton, J. C. D. & Higgins, C. F. (1994). The chromatin-associated protein H-NS alters DNA topology *in vitro*. *EMBO J* **13**, 258–268.

Ward, C. M., Wardle, S. J., Singh, R. K. & Haniford, D. B. (2007). The global regulator H-NS binds to two distinct classes of sites within the Tn10 transpososome to promote transposition. *Mol Microbiol* **64**, 1000–1013.

WHO (2008). *Global Tuberculosis Control: Surveillance, Planning, Financing.* Geneva: World Health Organization.

WHO (2009). *Global Tuberculosis Control: Epidemiology, Strategy, Financing.* Geneva: World Health Organization.

Wilson, R. L., Libby, S. J., Freet, A. M., Boddicker, J. D., Fahlen, T. F. & Jones, B. D. (2001). Fis, a DNA nucleoid-associated protein, is involved in *Salmonella typhimurium* SPI-1 invasion gene expression. *Mol Microbiol* **39**, 79–88.

Yamada, H., Yoshida, T., Tanaka, K., Sasakawa, C. & Mizuno, T. (1991). Molecular analysis of the *Escherichia coli hns* gene encoding a DNA-binding protein, which preferentially recognizes curved DNA sequences. *Mol Gen Genet* **230**, 332–336.

Zhang, J., Sprung, R., Pei, J., Tan, X., Kim, S., Zhu, H., Liu, C. F., Grishin, N. V. & Zhao, Y. (2009). Lysine acetylation is a highly abundant and evolutionarily conserved modification in *Escherichia coli*. *Mol Cell Proteomics* **8**, 215–225.

Edited by: S. V. Gordon

13. ANEXO III – Instruções aos autores para submissão de artigos a Molecular Biosystems.

Article Layout

Guidelines for Layout of Articles for Submission†

Also see: www.rsc.org/authorguidelines

CONTENTS

1.0 Organization of material

Every latitude, consistent with brevity, in the form and style of papers is permitted, and no rigid pattern for either is prescribed. **The suggestions outlined here are for guidance only**.

1.1 Full articles

1.1.1 Title. A paper should have a short, straightforward title directed at the general reader. Lengthy systematic names and complicated and numerous chemical formulae should therefore be avoided where possible. The use of non-standard abbreviations and symbols in a title is not encouraged. Brevity in a title, though desirable, should be balanced against its accuracy and usefulness.

The use of Series titles and Part numbers in titles of papers is discouraged. Instead the Series title and Part number can be included as a footnote to the first page together with a reference (reference 1) to the preceding Part.

When the preceding part has been submitted to the Society but is not yet published, the paper reference number should be given.

† For more detailed information on this topic, as well as links to useful websites and software resources, see: http://www.rsc.org/authorguidelines.

1.1.2 Author names. Full names for all the authors of an article should be given; initials should not be used.

1.1.3 Graphical contents entry. Graphics are included in the contents list. The format incorporates, a small graphic (maximum size 8 cm wide × 4 cm high) alongside one sentence of text, which should be presented in such a way as to encourage further perusal of the article, by highlighting the novelty and main feature(s) of interest; excessive lists of results and, in particular, cumbersome formulae should therefore be avoided.

In view of the space available graphics should be as clear as possible. Simple schematic diagrams or reaction schemes are preferred to ORTEP-style crystal structure depictions and complicated graphs, for example. The graphic used in the Contents entry need not necessarily appear in the article itself. Authors should bear in mind the final size of any lettering on the graphic. For examples of graphical contents entries check the online version of the appropriate journal.

1.1.4 Summary. Every paper must be accompanied by a summary (50–250 words) setting out briefly and clearly the main objects and results of the work; it should give the reader a clear idea of what has been achieved. The summary should be essentially independent of the main text; however, names, partial names or linear formulae of compounds may be accompanied by the numbers referring to the corresponding displayed formulae in the body of the text.

1.1.5 Introduction. This should give clearly and briefly, with relevant references, both the nature of the problem under investigation and its background.

1.1.6 Results and discussion. It is usual for the results to be presented first, followed by a discussion of their significance. Only strictly relevant results should be presented and figures, tables, and equations should be used for purposes of clarity and brevity. The use of flow diagrams and reaction schemes is encouraged. Data must not be reproduced in more than one form, *e.g.* in both figures and tables, without good reason.

1.1.7 Experimental. Descriptions of experiments should be given in detail sufficient to enable experienced experimental workers to repeat them; the degree of purity of materials should be given, as should the relative quantities used. Descriptions of established procedures are unnecessary. Standard techniques and methods used throughout the work should be stated at the beginning of the section. Apparatus should be described only if it is non-standard; commercially available instruments are referred

to by their stock numbers (*e.g.* Perkin-Elmer 457 or Varian HA-100 spectrometers). The accuracy of primary measurements should be stated. Unexpected hazards encountered during the experimental work should be noted. In general there is no need to report unsuccessful experiments.

1.1.8 Conclusion. This is for interpretation and to highlight the novelty and significance of the work. The conclusions should *not* summarise information already present in the text or abstract.

1.1.9 Acknowledgements. Contributors other than co-authors may be acknowledged in a separate paragraph at the end of the paper; acknowledgements should be as brief as possible.

1.1.10 Dedications. Personal dedications of an appropriate nature may be included as a footnote to the title of the paper. Dedications for significant birthdays (from 60 years onwards) and *in memoriam* dedications would be considered appropriate. Other forms of dedication may require approval of the relevant journal's Editorial Board.

1.1.11 Bibliographic references and notes. These should be listed at the end of the manuscript in numerical order.

1.2 Communications

Individual articles should be as brief as possible; depending on the journal in question, a page limit my apply. Formatting should be as for Full Articles, except for the following topics.

1.2.1 Summary. This is restricted to one sentence of text.

1.2.2 Article. No section headings are used in Communications. Brief details of key experiments are permitted and should include the amounts of reagents used in chemical reactions. Extensive spectroscopic and other supporting data are not required, but authors are encouraged to supply such data as Electronic Supplementary Information to aid the referees in their assessment of the work. Description for routine procedures should *not* be included.

1.2.3 Notes and bibliographic references. These should not be extensive and inclusion of 5–10 references is recommended.

1.2.4 Figures. These should be kept to a minimum bearing in mind the restrictions to the length of most Communications.

2.0 Style and presentation

2.1 Brevity

For reasons of economy, brevity in the presentation of papers is essential. Authors should note that the following practices are likely grounds for rejection of a manuscript, or acceptance only after substantial revision.

- Unnecessary division of work into separate parts of a series of papers.
- Submission of fragmentary work which can be included in a larger article.
- Undue elaboration of hypotheses.
- Over-detailed and verbose exposition of ideas.
- Excessive use of diagrams; for example, a straight-line plot can be adequately expressed as an equation together with, if necessary, a table of deviations.
- Duplication of data in text, tables and figures, *etc.*
- Descriptions of slight variations of essentially the same technique.

2.2 Linguistic and typographical conventions

2.2.1 Grammar and spelling. Standard English or American spelling is used but consistency should be maintained within a paper.

2.2.2 Abbreviations. The use of common or standard abbreviations is encouraged.

2.2.3 Use of italics. Foreign words and phrases and Latin abbreviations are given in italics: *e.g.*, *in toto*, *in vivo*, *ca.*, *cf.*, *i.e.*

In the names of chemical compounds or radicals italics are used for prefixes (other than numerals or symbols) when they define the positions of named substituents, or when they define stereoisomers: other prefixes are printed in roman. (*Note:* Initial capital letters are not to be used with italic prefixes or single-letter prefixes: full stops are not to be associated with letter prefixes.) For example, *o*-, *m*- and *p*-nitrotoluenes, but *ortho*-, *meta*- and *para*- compounds (*o*-, *m*- and *p*- are used only with specific names; *ortho*-, *meta*- and *para*- are used with classes), *N*,*N*-dimethylaniline, *trans*- and *cis*-bis(glycinato)platinum(II), *gem*- and *vic*-diols, benzil *anti*-oxime.

The names of periodicals or their abbreviations are set in italics.

2.2.4 Headings.

(a) Main sections (Experimental, Results and discussion, *etc.*): side-heading, bold, first initial capital letter only, no final fullstop.
(b) Main side-heading: bold, first initial capital letter only, no final fullstop.
(c) Subsidiary side-heading: bold, first initial capital letter only, final fullstop.
(d) Further subdivision: italic, first initial capital letter only, final fullstop.

For example:

Experimental

Preparation of the thiolate complexes

Bis(benzenethiolato)bis(dimethyldithiocarbamato)[*N*,*N*-dimethylhydrazido(2–)]molybdenum(VI) 10. Method 1. Benzenethiol (1 g) was added to . . .

3.0 Presentation of experimental data

3.1 Physical characteristics of compounds

Data associated with particular compounds should be listed after the name of the compound concerned, following the description of its preparation. **The following is suggested as the order in which the most commonly encountered data for a new compound should be cited**: yield, melting point, optical rotation, refractive index, elemental analysis, UV absorptions, IR absorptions, NMR spectrum, mass spectrum. Appropriate formats for the citation of each are as follows.

3.1.1 Yield. In parentheses after the compound name (or its equivalent). Weight and percentage are separated by a comma, *e.g.* the lactone (7.1 g, 56%).

3.1.2 Melting point. In the form mp 75 °C (from EtOH), *i.e.* the crystallization solvent in parentheses. If an identical mixed melting point is to be recorded, the form mp and mixed mp 75 °C is appropriate.

3.1.3 Optical rotation. The *units* should be stated in the preamble to the Experimental section, *e.g.* $[\alpha]_D$ values are given in 10^{-1} deg cm^2 g^{-1}. Shown in the form $[\alpha]_D^{22}$ –22.5 (*c* 0.95 in EtOH), *i.e.* concentration and solvent in parentheses.

3.1.4 Refractive index. Given in the form n_D^{22} 1.653.

3.1.5 Elemental analysis. In the presentation of elemental analyses, both forms (Found: C, 63.1; H, 5.4. $C_{13}H_{13}NO_4$ requires C, 63.2; H, 5.3%) and (Found: C, 62.95; H, 5.4. Calc. for $C_{13}H_{13}NO_4$: C, 63.2; H, 5.3%) are acceptable. Analyses are normally quoted to the nearest 0.1%, but a 5 in the second place of decimals is retained. For identification purposes for new compounds, an accuracy to within ±0.3% is expected, and in exceptional cases, to within ±0.5% is required.

If a molecular weight is to be included, the appropriate form is: [Found: C, 63.1; H, 5.4%; M (mass spectrum), 352 (or simply M^+, 352). $C_{13}H_{13}NO_4$ requires C, 63.2; H, 5.3%; M, 352].

3.1.6 UV absorptions. These are given in the form λ_{max}(EtOH)/nm 228 (ε/dm^3 mol^{-1} cm^{-1} 40 900), 262 (19 200) and 302 (11 500). Inflections and shoulders are specified as 228infl or 262sh. Alternatively the following form may be used:

λ_{max} (EtOH)/nm 228, 262 and 302 (ε/dm^3 mol^{-1} cm^{-1} 40 900, 19 200 and 11 500). log ε may be quoted instead of ε.

3.1.7 IR absorptions. As follows: ν_{max}/cm^{-1} 3460 and 3330 (NH), 2200 (conj. CN), 1650 (CO) and 1620 (CN). The type of signal (s, w, vs, br) can be indicated by appended letters (*e.g.* 1760vs).

3.1.8 NMR data. For all spectra δ values should be used, with the nucleus indicated by subscript if necessary (*e.g.* δ_H, δ_C). A statement specifying the units of the coupling constants should be given in the preamble to the Experimental section, *e.g. J* values are given in Hz. Instrument frequency, solvent, and standard should be specified. For example: δ_H(100 MHz; $CDCl_3$; Me_4Si) 2.3 (3 H, s, Me), 2.5 (3 H, s, COMe), 3.16 (3 H, s, NMe) and 7.3–7.6 (5 H, m, Ph). A broad signal may be denoted by br, *e.g.* 2.43 (1 H, br s, NH). Order of citation in parentheses: (i) number of equivalent nuclei (by integration), (ii) multiplicity (s, d, t, q), (iii) coupling constant, *e.g.* $J_{1,2}$ 2, J_{AB} 4, (iv) assignment; italicisation can be used to specify the nuclei concerned (*e.g.* CH_3CH_2). The proton attached to C-6 may be designated C(6)H or 6-H; the methyl attached to C-6, 6-Me or C(6)Me. Mutually coupled protons in ^{1}H NMR spectra must be quoted with precisely matching *J* values, in order to assist thorough interpretation. In instances of any ambiguities when taking readings from computer print-outs, mean *J* values should be quoted, rounded to the nearest decimal point.

3.1.9 Mass spectrometry data. Given in the form: *m*/*z* 183 (M^+, 41%), 168 (38), 154 (9), 138 (31) *etc*. The molecular ion may be specified as shown if desired. Relative intensities in parentheses (% only included once). Other assignments may be included in the form *m*/*z* 152 (33, M – CH_3CONH_2). Metastable peaks may be listed as: M* 160 (189→174), 147 (176→161), *etc*. The type of spectrum (field desorption, electron impact, *etc*.) should be indicated. Exact masses quoted for identification purposes should be accurate to within 5 ppm (EI and CI) or 10 ppm (FAB or LSIMS).

3.1.10 Literature citations. If comparison is to be made with literature values, these should be quoted in parentheses, *e.g.* mp 157 °C (from chloroform) (lit.,[19] 156 °C), or ν_{max}/cm^{-1} 2020 and 1592 (lit.,[24] 2015 and 1600).

3.1.11 Experiments involving microorganisms. For work involving microorganisms, sufficient detail should be provided to identify the species being used.

3.1.12 A typical experimental section. The paragraph below exemplifies many of the points made in the preceding sections.

(*E*,*E*)-Undeca-3,8-diene-1,11-diol 16

A solution of the undeca-3,8-diyne-1,11-diol (2 g, 11 mmol) in tetrahydrofuran (2 cm^3) was added cautiously dropwise over 5 min to a stirred solution of lithium aluminium hydride (8.3 g, 22 mmol) in tetrahydrofuran (200 cm^3) at 0 °C. The mixture was heated under reflux in an atmosphere of nitrogen for 20 h, after which it was cooled and quenched by careful addition of saturated aqueous sodium sulfate. The mixture was acidified with dilute hydrochloric acid (100 cm^3) and then extracted with diethyl ether (4 × 100 cm^3). The combined extracts were dried and evaporated under reduced pressure to leave a yellow oil, which was purified by column chromatography on silica using diethyl ether as eluent to give the dienediol **16** (1.9 g, 95%) as a colourless oil (Found: C, 71.5; H, 11.4. $C_{11}H_{20}O_2$ requires C, 71.7; H, 10.9%); ν_{max}(film)/cm^{-1} 3342, 1777, 1711, 968 and 766; δ_H(250 MHz; $CDCl_3$; Me_4Si) 1.45 (2 H, quintet, *J* 7.4, $CH_2CH_2CH_2$), 1.65 (2 H, br s, OH), 2.03 (4 H, dt, *J* 6.6 and 7.4, 2 × CH_2CH=CH), 2.27 (4 H, dt, *J* 7.6 and 6.3, 2 × CH_2CH=CH), 3.63 (4 H, t, *J* 6.3, 2 × C H_2OH) and 5.34–5.61 (4 H, m, 2 × CHCH); δ_C(67.8 MHz; $CDCl_3$; Me_4Si) 29.3 (t), 32.3 (2 × t), 36.2 (2 × t), 62.3 (2 × t), 126.6 (2 × d) and 133.4 (2 × d); *m*/*z* (EI) 154.1334 (M^+ – CH_2O. $C_{10}H_{18}O$ requires 154.1358), 135 (7%), 125 (7), 107 (24), 98 (38) and 81 (100).

3.2 Characterisation of new compounds

It is the responsibility of authors to provide fully convincing evidence for the homogeneity and identity of all compounds they claim as new. Evidence of both purity and identity is required to establish that the properties and constants reported are those of the compound with the new structure claimed.

A compound is considered as new (a) if it has not been prepared before, (b) if it has been prepared before but not adequately purified, (c) if it has been purified but not adequately characterized, (d) if, earlier, it has been assigned an erroneous constitution, or (e) if it is a natural product isolated or synthesized for the first time. In preliminary communications compounds are often recorded with limited characterizing data; in spite of (c) above later preparations of such compounds are not considered as new if the properties previously reported are confirmed; the same applies to patents.

Referees will assess, as a whole, the evidence in support of the homogeneity and structure of all new compounds. No hard and fast rules can be laid down to cover all types of compound, but evidence for the unequivocal identification of new compounds should wherever possible include good elemental analytical data; an accurate mass measurement of a molecular ion does not provide evidence of purity of a compound and must be accompanied by independent evidence of homogeneity *e.g.* HPLC. Low-resolution mass spectrometry must be treated with even more reserve in the absence of firm evidence to distinguish between alternative molecular formulae. Where elemental analytical data cannot be obtained, appropriate evidence which is convincing to an expert in the field may be acceptable, but authors should include, for the referees, an explanation of the special nature of their problem.

Spectroscopic information necessary to the assignment of structure should be given. Just how complete this information should be must depend upon the circumstances; the structure of a compound obtained from an unusual reaction or isolated from a natural source needs much stronger supporting evidence than one derived by a standard reaction from a precursor of undisputed structure. Authors are reminded that full spectroscopic assignments may always be treated as Supplementary Data (see Section 3.5) where their importance does not justify their inclusion in the published paper.

Particular care should be taken in supporting the assignments of stereochemistry (both relative and absolute) of chiral compounds reported, for example by NMR spectroscopy, X-ray crystallography, polarimetry or correlation with known compounds of undisputed configuration. In cases where mixtures of isomers are generated (*e.g. E/Z* isomers, enantiomers, diastereoisomers), the constitution of the mixture should usually be established using appropriate analytical techniques (*e.g.* NMR spectroscopy, GLC, HPLC) and reported in an unambiguous fashion. In the case of asymmetric reactions in which enantiomeric mixtures are prepared, the direct measurement of the enantiomer ratio and its reporting expressed as an enantiomeric excess (ee) is recommended, and is preferred to (less reliable) polarimetry methods.

3.3 Characterisation within chemical biology

Where compounds are synthesised for testing in biological systems, sufficient evidence for purity and identity must be provided such that the results of the experiment may be trusted.

The homogeneity of oligomeric compounds (peptides, saccharides, nucleotides *etc*.) should be determined by HPLC analyses or by other appropriate analytical methods (*e.g.* capillary electrophoresis) with a purity of not less than 95%.

4.0 Bibliographic references, notes and footnotes

Footnotes or Notes may be used to present material which, if included in the body of the text, would disrupt the flow of the argument but which is, nevertheless, of importance in qualifying or amplifying the textual material. Footnotes are referred to with the following symbols: †,‡,§,¶,‖, *etc.* Alternatively the information may be included as Notes (end-notes) to appear in the Notes/references section of the manuscript. Notes should be numbered using the same numbering system as the bibliographic references.

Bibliographic reference to the source of statements in the text is made by use of *superior numerals* at the appropriate place. The reference numbers should be cited in the correct sequence through the text (including those in tables and figure captions, numbered according to where the table or figure is designated to appear). The

references themselves are given at the end of the final printed text along with any Notes.

Authors are encouraged to check the RSC Reviews web site to ensure that they have cited relevant recent reviews.†

4.1 Journals

The style of journal abbreviations to be used in the Society's publications is that defined in Chemical Abstracts Service Source Index (CASSI). The abbreviations listed in CASSI are based upon internationally recognised systems. A list of CASSI-style abbreviations covering the most commonly cited journals is available from our web site.† It is not, of course, a full list; CASSI plus its quarterly supplements run to more than 2000 pages.

If you cannot locate an authoritative abbreviation for a journal, and if it is not obvious how the title should be abbreviated, please cite the full title.

Bibliographic details should be cited in the order: **year**, **volume**, **page**. Where possible, page number ranges are preferred over single values, but either format is acceptable.

Where page numbers are not yet known, articles may be cited by DOI (Digital Object Identifier). *e.g.* A. R. Jones, *Dalton Trans.*, 2005, DOI: 10.1039/B503459J.

Please note that journal citations in articles submitted to the journal *Photochemical & Photobiological Sciences* should include the article titles.

4.2 Books

For example:

> J. Barker, in *Catalyst Deactivation*, ed. B. Delmon and C. Froment, Elsevier, Amsterdam, 2nd edn., 1987, vol. 1, ch. 4, pp. 253–255.

4.3 Patents

Patents should be indicated in the following form:

> *Br. Pat*., 357 450, 1986. *US Pat*., 1 171 230, 1990.

4.4 Reports and bulletins, *etc.*

For example:

> R. A. Allen, D. B. Smith and J. E. Hiscott, *Radioisotope Data*, UKAEA Research Group Report AERE-R 2938, H.M.S.O., London, 1961.

4.5 Material presented at meetings

For example:

> H. C. Freeman, Proceedings of the 21st International Conference on Coordination Chemistry, Toulouse, 1980.

4.6 Theses

For example:

> A. D. Mount, Ph.D. Thesis, University of London, 1977.

4.7 Reference to unpublished material

For material presented at a meeting, congress or before a Society, *etc*., but not published, the following form is used:

> A. R. Jones, presented in part at the 28th Congress of the International Union of Pure and Applied Chemistry, Vancouver, August, 1981.

For material accepted for publication, but not yet published, the following form

> A. R. Jones, *Dalton Trans.*, 2003, DOI: 10.1039/paperno.

is used for RSC journals, and

> A. R. Jones, *Angew. Chem.*, in press.

is used for non-RSC journals. If DOI numbers are known these should be cited in the form recommended by the publisher

For material submitted for publication but not yet accepted the following form is used:

> A. R. Jones, *Angew. Chem.*, submitted.

For personal communications the following is used:

> G. B. Ball, personal communication.

If material is to be published but has yet to be submitted the following form is used:

> G. B. Ball, unpublished work.

Reference to unpublished work should not be made without the permission of those by whom the work was performed.

4.8 Names

The names and initials of all authors are always given in the reference; they must not be replaced by the phrase *et al*. This does not prevent some, or all, of the names being mentioned at their first citation in the cursive text: initials are not necessary in the text.

4.9 Composite references

Whenever possible, composite references should be used rather than a series of individual references. The style for composite references is as follows:

> A. B. Jones, *J. Am. Chem. Soc.*, 1956, **78**, 1234–1246; A. B. Jones and C. D. Brown, *J. Am. Chem. Soc.*, 1957, **79**, 567–569; A. B. Jones and E. F. Green, *J. Am. Chem. Soc*., 1957, **79**, 999–1048.

Idem, *loc. cit*., and *op. cit* are not used in references.