PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
FACULDADE DE BIOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

IDENTIFICAÇÃO DE LIGANTES DA CHIQUIMATO QUINASE DE *Mycobacterium tuberculosis* POR *DOCKING* MOLECULAR

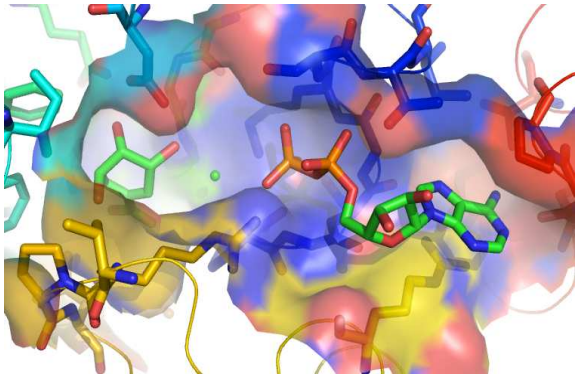
Carolina Pasa Vianna

Porto Alegre
2011

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
FACULDADE DE BIOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

IDENTIFICAÇÃO DE LIGANTES DA CHIQUIMATO QUINASE DE *Mycobacterium tuberculosis* POR *DOCKING* MOLECULAR

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Celular e Molecular como requisito para obtenção do grau de Mestre.



Pós-Graduando:
Carolina Pasa Vianna

Orientador:
Prof. Dr. Walter Filgueira de Azevedo Junior

A minha mãe Marlene Pasa Vianna e ao meu pai Julio de Oliveira Vianna (*in memoriam*) que tanto apoiaram e incentivaram o meu crescimento profissional.

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Walter Filgueira de Azevedo Junior, ao me receber em seu laboratório e por acreditar na minha capacidade.

Às minhas amigas, Cristiane Bereta e Andréia Valim, por me acompanharem nesta caminhada, sempre com palavras de coragem e incentivo.

Aos amigos que fiz ao longo do mestrado, em especial Mariel, Patrícia e Ivani.

Aos colegas do LabioQuest, em especial Rafael e Luis, pela ajuda e paciência durante a realização do projeto.

A minha prima Vivian Pasa e ao meu amigo Maiko Andrade, por me receber em sua casa, nos momentos que precisei de estadia para ficar em Porto Alegre.

A equipe diretiva da Escola Manoel Prestes, por me apoiar e me incentivar em todos os momentos.

SUMÁRIO

LISTA DE FIGURAS

LISTA DE SÍMBOLOS E ABREVIATURAS

AIDS	*Acquired immune deficiency syndrome* ou Síndrome da Imunodeficiência Adquirida
ADMET	Absorção, distribuição, metabolismo, excreção e toxicidade
ADP	Adenosina - difosfato
ATP	Adenosine - triphosphate
BCG	*Bacille Calmette-Guérin* ou vacina contra a tuberculose
HIV	*Human immunodeficiency virus* ou Vírus da imunodeficiência humana
MOLDOCK	*Molegro Virtual Docker*
MtSK	*Mycobacterium Tuberculosis Shikimate Kinase*
PDB	*Protein Data Bank* ou Banco de Dados de Proteínas
RMSD	Raiz do desvio médio quadrático
SK	*Shikimate kinase*
SKM	*Shikimate*
VS	Virtual *Screening* ou Triagem virtual
WHO	*World Health Organization* ou Organização Mundial da Saúde

RESUMO

Entre as doenças infecciosas, a tuberculose se destaca, como sendo a principal causa de morte humana, principalmente em países em desenvolvimento. Entre os alvos identificados no genoma de *Mycobacterium tuberculosis,* as enzimas da via do chiquimato merecem atenção especial, pois são essenciais para a sobrevivência dos microorganismos e ausente em mamíferos. O objeto do nosso estudo é a quinta enzima desta via, a Chiquimato Quinase (SK). Foram aplicados métodos de virtual *screening*, a fim de identificar novos potenciais inibidores para esta enzima. Neste trabalho nós empregamos o programa MOLDOCK em todas as simulações de *docking* molecular. A precisão do *docking* enzima-ligante foi validada em um conjunto de 12 complexos de SK- ligante, para aquelas estruturas cristalográficas que estavam disponíveis, gerando um RMSD abaixo de 2,0 Å. A aplicação deste protocolo em um banco de dados comercialmente disponível permitiu a identificação de novas moléculas com potencial para se tornar drogas contra tuberculose. Além disso, foram identificados os resíduos da cavidade de ligação que são essenciais para as interações intermoleculares desta enzima.

Palavras chave: Tuberculose - Via do Chiquimato - Chiquimato Quinase - *Docking* Molecular - *Moldock* - Virtual *Screening*

ABSTRACT

Tuberculosis (TB) is the major cause of human mortality from a curable infectious disease, attacking mainly in developing countries. Among targets identified in *Mycobacterium tuberculosis* genome, enzymes of the shikimate pathway deserve special attention, since they are essential to the survival of the microorganism and absent in mammals. The object of our study is shikimate kinase (SK), the fifth enzyme of this pathway. We applied virtual screening methods in order to identify new potential inhibitors for this enzyme. In this work we employed MOLDOCK program in all molecular docking simulations. Accuracy of enzyme-ligand docking was validated on a set of 12 SK-ligand complexes for which crystallographic structures were available, generating root-mean square deviations below 2.0 Å. Application of this protocol against a commercially available database allowed identification of new molecules with potential to become drugs against TB. Besides, we have identified the binding cavity residues that are essential to intermolecular interactions of this enzyme.

Keywords: Tuberculosis - Shikimate pathway - Shikimate Kinase - Docking Molecular - Moldock - Virtual Screening

Capítulo 1

INTRODUÇÃO

1.1) Tuberculose

As doenças infecciosas são a causa de sofrimento e morte de centenas de milhões de pessoas, especialmente em áreas tropicais e subtropicais do mundo, local onde ocorre cerca de 90% das mortes causadas por esse tipo de enfermidade. Entre as doenças infecciosas a tuberculose destaca-se, sendo a principal causa infecciosa de morte em humanos e está rapidamente tornando-se uma epidemia global. É estimado que em todo o mundo cem milhões de pessoas são infectadas anualmente. Aproximadamente dez milhões desenvolvem a doença, com cinco milhões destes progredindo para um estágio infeccioso, culminando com três milhões de mortes. De acordo com a Organização Mundial de Saúde (WHO Report 2009), a taxa global de incidência da tuberculose vem crescendo aproximadamente 0,3% ao ano.

A incidência de tuberculose teve um declínio rápido no início do século vinte nos países em desenvolvimento, devido à melhora nas condições sanitárias e de moradias. Essa tendência foi acelerada inicialmente pela introdução da vacinação BCG e da descoberta de antibióticos como a estreptomicina e, posteriormente, com a descoberta do ácido *p*-aminosalisílico (1946), isoniazida (1952) e rifampicina (1965). Mas está ocorrendo um ressurgimento dos casos de tuberculose nos países em desenvolvimento. O aumento no número de casos de tuberculose atualmente está relacionado a dois principais fatores. O primeiro é a ocorrência da tuberculose em pacientes co-infectados com o vírus HIV. Pacientes com AIDS, cujo sistema imunológico enfraquecido não pode controlar o crescimento do bacilo, apresentam um risco cem vezes maior de desenvolver a doença. O outro fator é a emergência de cepas resistentes aos antimicrobianos de primeira linha (isoniazida e rifampicina) utilizados no tratamento, devido às terapias inadequadas e o uso indiscriminado destes antibióticos (BASSO *et al.,* 2005).

Os mecanismos de resistência identificados até o momento são resultantes de mutações pontuais em genes codificadores das proteínas que são os alvos destes

agentes anti-tuberculose (BASSO *et al*,1998). Cepas de *M. tuberculosis* resistentes às drogas anti-tuberculose de primeira linha têm sido identificadas globalmente. Um aspecto preocupante da situação brasileira é que taxas superiores a 45% dos pacientes, previamente tratados, apresentam multi-resistência (definida como resistente à isoniazida e rifampicina) adquirida.

1.2) Via da Ácido chíquimico

Uma via sintética que tem levado a um grande interesse como alvo para desenho de novas drogas contra várias bactérias, outros organismos patogênicos e herbicidas, é a via do ácido chiquímico (ALIBHAI & STALLINGS, 2001; SCHÖNBRUNN *et al*., 2001; PARK *et al*., 2004). Esta via leva à síntese de aminoácidos aromáticos e outros compostos desta classe em bactérias, fungos, algas e plantas, porém ela é ausente em mamíferos. Já existem alguns herbicidas (Roundup® - MONSANTO - ALIBHAI & STALLINGS, 2001) que inibem fortemente, de maneira reversível, enzimas desta via sintética, o que a torna ainda mais interessante para o desenho de fármacos direcionados a inibí-la (BENTLEY, 1990). A via sintética do ácido chiquímico foi identificada em *M. tuberculosis* (RATLEDGE, 1982; ROBERTS *et al*., 1998; COLE, 1998), sendo composta por sete passos enzimáticos que catalisam sequencialmente a conversão da eritrose-4-fosfato e fosfoenolpiruvato em corismato (Figura 1). O corismato é o principal precursor para a biossíntese de PABA (ácido p-aminobenzóico, precursor do tetrahidrofolato), ácido p-hidroxibenzóico (precursor da coenzima Q ou ubiquinona), micobactinas e dos aminoácidos aromáticos essenciais (HERRMANN & WEAVER, 1999). Os intermediários da via do ácido chiquímico são ainda considerados como pontos de ramificação para outras vias metabólicas (HERRMANN & WEAVER, 1999). Algumas dessas enzimas de *M. tuberculosis*, como a 3-deoxi-D-arabino-heptulosonato 7-fosfato sintase (*Mt*DAHPS); corismato sintase (*Mt*CS); a 5-enolpiruvilchiquimato-3-fosfato sintase (*Mt*EPSPS) e chiquimato quinase (*Mt*SK), estão com seus cDNAs clonados, as proteínas foram expressas e purificadas com as respectivas atividades (OLIVEIRA *et al*., 2001; 2003; RIZZI *et al.*, 2005). Algumas destas enzimas de *M. tuberculosis* já foram cristalizadas e as estruturas em alta resolução determinadas; estudos de bioinformática para modelagem estrutural por homologia também foram

realizados (DE AZEVEDO *et al*., 2002a; PEREIRA *et al*., 2003; 2004; DIAS *et al*., 2004, 2006, 2007; OLIVEIRA *et al*., 2006). Pretendemos avançar nesses estudos usando a experiência já acumulada, para identificarmos novos inibidores, bem como estabelecer as bases estruturais para a ação dos inibidores já identificados, aliando-se estudos cristalográficos e *docking* molecular.

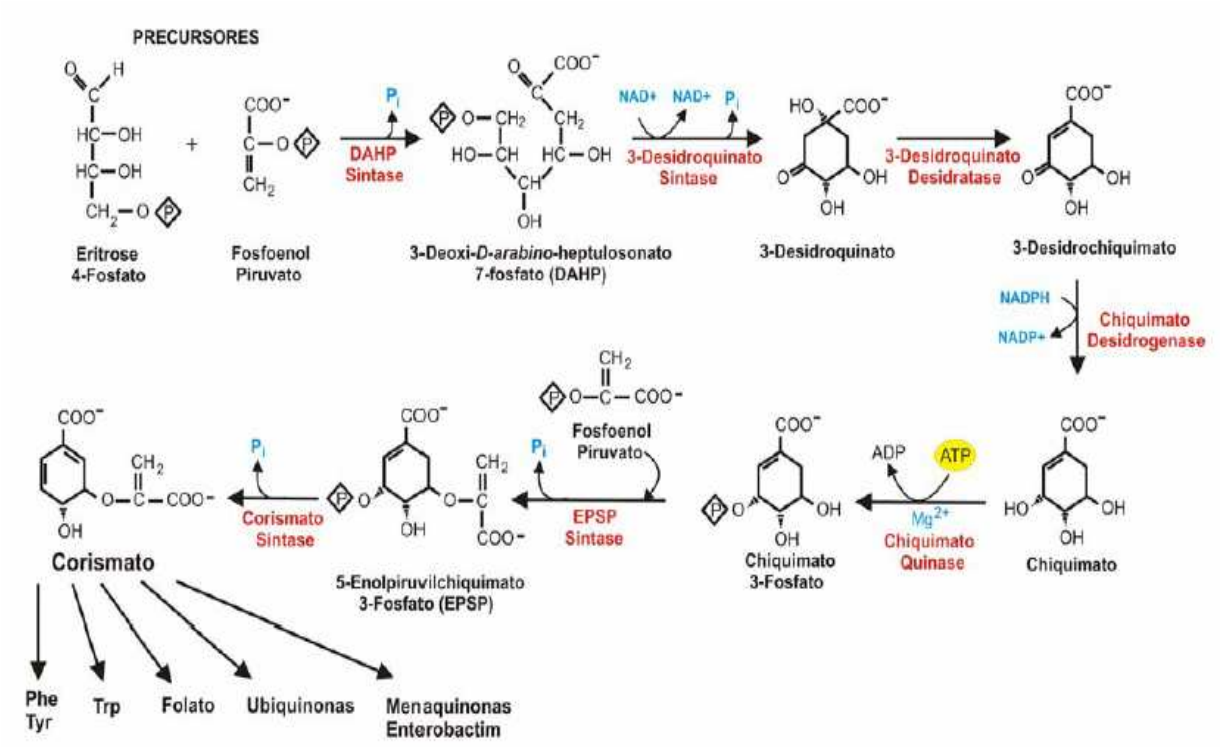


Figura 1: A via do ácido chiquimico. Fosfoenol-piruvato e eritrose 4-fosfato (precursores) são convertidos em ácido corismato. Corismato é essencial para a síntese de aminoácidos aromáticos (Phe, Tyr e Trp), folato, ubiquinonas, menaquinonas e entorobactium. Os sete passos são catalisados pelas enzimas: 3-deoxi-D-arabino-heptulosonato 7-fosfato sintase (DAHP Sintase); 3- desidroquinato sintase, 3- desidroquinato desidratase, chuiquimato desidrogenase, chiquimato quinase, 5-enolpiruvil chiquimato-3-fosfato sintase (EPSP Sintase) e corismato sintase. Figura modificada a partir de Mathews & van Holde. (MATHEWS & VAN HOLDE, 1990).

1.3) Chiquimato Quinase (SK)

A chiquimato quinase é a quinta enzima da via metabólica do ácido chiquímico (Figura 2), catalisa a fosforilação especifica do grupo 3-hidroxil do ácido chiquímico usando Adenosina Trifosfato (ATP) como um co-substrato e o chiquimato como substrato específico (KRELL et al., 1998; PEREIRA et al., 2004).

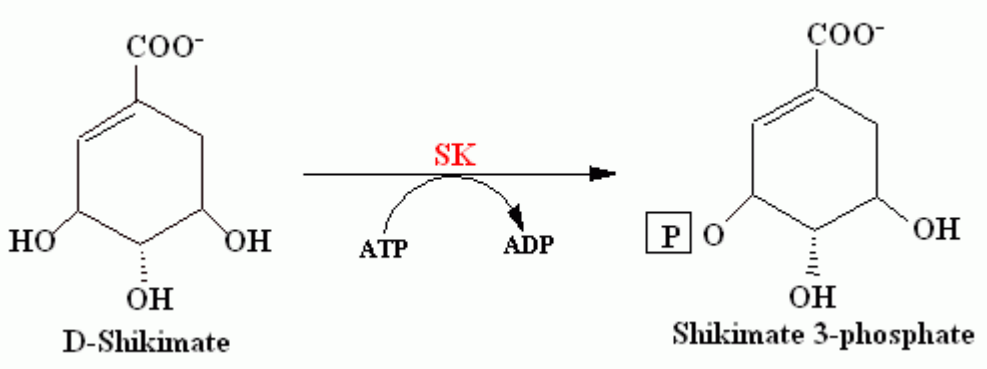


Figura 2: Reação catalisada pela chiquimato quinase.

A chiquimato quinase pertence à família estrutural das Nucleosídeo Monofosfato Quinases (NMP). Ela é uma proteína da classe α/β, consistindo de uma folha β central, formada por 5 fitas β paralelas circundadas por hélices-α, tendo uma topologia geral muito semelhante à adenilato quinase (PEREIRA et al., 2004 ; KRELL et al., 1998, 2001). Uma característica das NMP quinases é que elas sofrem grandes mudanças conformacionais durante a catálise (VONRHEIN et al., 1995). As NMP quinases são compostas por três domínios: 1) CORE, que contém um loop de ligação ao fosfato altamente conservado (P-loop); 2) o domínio LID, que sofre mudanças conformacionais substanciais devido a ligação ao substrato; e, 3) o domínio de ligação NMP, responsável pelo reconhecimento e ligação de um substrato específico (GU et al.,2002).

Com o objetivo de demonstrar a importância da via do ácido chiquímico em *Mycobacterium tuberculosis*, foi realizado a disrrupção do gene aroK, que codifica a chiquimato quinase. Os resultados obtidos através destes experimentos confirmaram que esta enzima é essencial para viabilidade do bacilo (PARISH & STOKER, 2002).

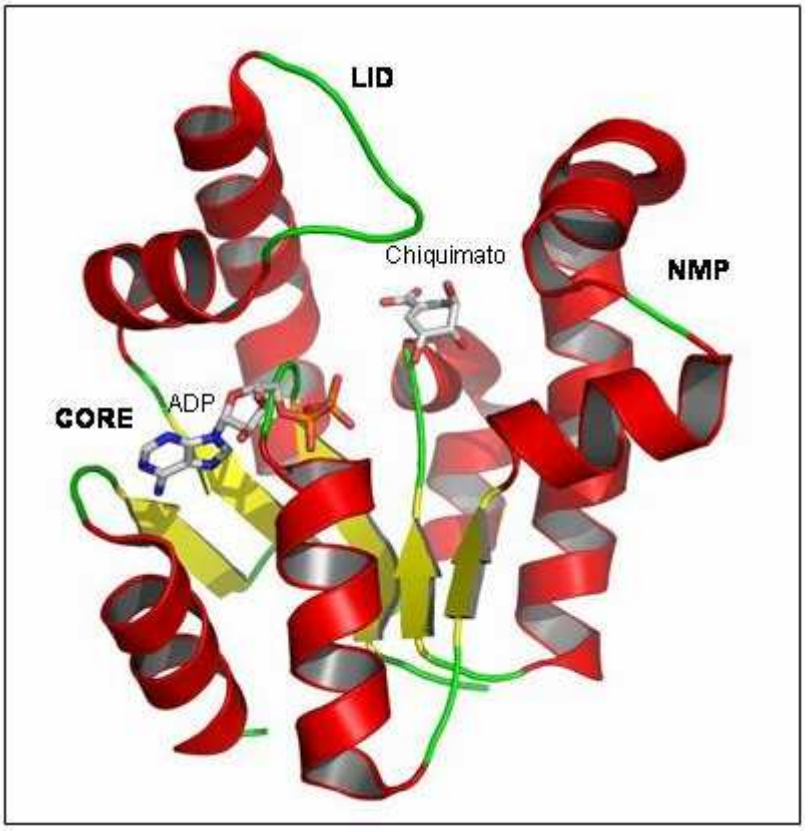


Figura 3: Chiquimato quinase complexada com ADP e chiquimato (código de acesso PDB: 2DFN) (DIAS et al., 2007).

1.4) Planejamento racional de fármacos *in silico*

O desenvolvimento racional de novas drogas, direcionado pela estrutura, baseia-se no conhecimento detalhado da estrutura tridimensional em alta resolução da proteína alvo, no entendimento da interação do alvo com o ligante natural e na racionalização de como o alvo poderá interagir com um fármaco potencial. Assim, o processo de interação proteína-ligante é fundamental para o desenho de novos fármacos e na otimização de um fármaco líder, aliando características de alta seletividade/especificidade e propriedades farmacocinéticas adequadas. Interações proteína-ligante participam da regulação da atividade protéica, por exemplo, a catálise enzimática é dependente da interação entre a enzima com seu substrato. O transporte de moléculas por proteínas é diretamente influenciado pelas propriedades da interação entre a proteína transportadora e a molécula transportada. A interação entre um ligante e a proteína receptora pode ocasionar mudanças conformacionais

nesta que transporta a sinalização celular tanto por membranas quanto no meio intracelular (SILVA, 2008). Portanto, é grande o interesse em compreender as forças que estabilizam uma determinada interação proteína-ligante para o desenvolvimento racional de novos fármacos. As principais forças moleculares que governam a interação proteína-ligante são: ligações de hidrogênio, interações hidrofóbicas e de Van der Walls.

O planejamento racional baseado em estrutura e no mecanismo de ação de enzimas específicas é a estratégia mais eficiente no desenvolvimento de novos fármacos, capaz de contribuir em todos os estágios do processo, desde a descoberta de protótipos, sua otimização, até a elaboração de compostos candidatos a testes clínicos. Esta estratégia é baseada no bloqueio ou estimulação da atividade de macromoléculas, como proteínas ou ácidos nucléicos, associados a diferentes processos patológicos. A informação estrutural do ligante permite a descoberta e síntese de compostos com complementariedade geométrica, hidrofóbica e eletrostática ao seu sítio de ligação, podendo vir a se tornar potenciais inibidores e futuros fármacos (SILVA, 2008). Essa abordagem caracteriza o planejamento racional de fármacos baseados em estrutura. O que a torna ainda mais atrativa, quando utilizada em proteínas, é o conhecimento de que 78% dos fármacos atuais têm como alvo receptor esse tipo de biomacromolécula (MARSHALL, 2004)

Desde a concepção do alvo biológico até a descoberta de um novo fármaco, um processo que pode levar em média 11 anos ou até mais, a bioinformática, juntamente com a química computacional, vem oferecendo um excelente direcionamento ao planejamento racional de fármacos, já que há inúmeros casos de sucesso envolvendo o emprego de simulações computacionais (MARSHALL, 2004), citando como exemplo os importantes fármacos: losartan, atorvastatina e celecoxib.

No planejamento racional de fármacos, destaca-se o *docking* molecular como um dos métodos mais empregados. No *docking*, são investigadas as possíveis orientações que determinada molécula assume no interior do sítio ativo de uma proteína. Os métodos de *docking*, em geral, envolvem uma função de energia contendo parâmetros eletrostáticos, van der Waals, ligações de hidrogênio, e algumas vezes, hidrofóbicos, os quais geram modelos matemáticos que predizem as

melhores orientações do ligante, segundo uma lista escore de energia (SILVA, 2008).

O *screening* virtual vem complementar o processo de descoberta de novos fármacos. Um dos mais recentes programas, Moldock (THOMSEN E CHRISTENSEN, 2006) considera a flexibilidade do ligante e é baseado em evolução diferencial dirigida e uma função escore baseada em campos de força. A partir dessa estratégia é possível selecionar por *screening* virtual compostos de bases de dados contendo tipicamente milhares de estruturas, eliminando assim ligantes não promissores antes que sejam sintetizados.

Há diversos casos de sucesso com o uso da abordagem de *screening* virtual, tal como a descoberta de isoflavonóides, como inibidores não-esteroidais da 5□ – redutase (BRENK *et. al*., 2003; CHEN *et al*., 2001). Porém o sucesso do *screening* dito *in silico*, e em geral das técnicas de *docking*, depende do conhecimento de detalhes estruturais do sítio de reconhecimento da biomacromolécula (CARLSON *et.al* 1999).

1.5) Propriedades físico-químicas

Nos últimos anos, percebeu-se um avanço considerável no desenvolvimento de técnicas de modelagem molecular que elevaram o poder de predição do comportamento dos ligantes em sistemas biológicos,que são aplicadas no estudo de diversas propriedades, como: absorção, distribuição, metabolismo, excreção e toxicidade (ERINS; ROSE, 2002).

Na busca de inovação e desenvolvimento de novos fármacos, é evidente a pressão do mercado sobre a otimização dos recursos financeiros. Pode-se destacar também a restrição, ou crescente dificuldade, com relação ao uso de animais para utilização dos testes de toxicidade. A maioria dos fármacos retirados do mercado ocorreram devido a estes fármacos causarem efeitos tóxicos indesejáveis como o antiinflamatório Vioxx® (REVISTA ÉPOCA,2005). Dessa forma, os métodos de predição de toxicidade *in silico* surgem como um importante e alternativa ferramenta na seleção ou priorização de moléculas promissoras a serem avaliadas com maior cautela em testes de toxicidade, reduzindo assim, os custos financeiros inerentes ao processo, o uso indiscriminado de animais e satisfazendo as precauções em relação

à toxicidade desde a fases iniciais do processo de desenvolvimento de fármacos (SILVA, 2008).

Para o desenvolvimento de novos fármacos, já é possível estimar propriedades farmacocinéticas de ligantes com potencial a se tornar fármacos. Como exemplo mais simples, a ‘ Regra dos Cinco’ (RO5), de Lipinski, preconiza que os fármacos que apresentam biodisponibilidade por via oral, seguem, a saber: peso molecular menor ou igual a 500, log P menor ou igual a 5, número de grupo de doadores de ligações de hidrogênio menor ou igual 5 e número de grupos de aceptores de ligações de hidrogênio menor ou igual a 10 (LIPINSKI et. al., 1997).

O Laboratório de Bioquímica Estrutural (LabBioqEst) da Faculdade de Biociências-PUCRS tem como objeto de estudo a interação proteína-ligante em nível molecular. A principal intenção é a determinação da estrutura tridimensional em alta resolução de proteínas nativas e em complexos com ligantes, para o posterior desenho racional de ligantes que possam interagir especificamente com as proteínas alvo. Estas informações serão importantes para ampliar a capacidade do grupo LabBioqEst-PUCRS em compreender o mecanismo de ação, a relação estrutura-função das proteínas alvo e auxiliar no desenho racional de novos fármacos.

1.6) OBJETIVO

1.6.1) Objetivo Geral

Este projeto de pesquisa pretende realizar estudos estruturais focados na interação proteína-ligante, focado na **Chiquimato Quinase** de *Mycobacterium tuberculosis*.

1.6.2) Objetivos específicos

1) Simular a interação proteína-ligante por meio de algoritmos de *docking* molecular;

2) Identificar aspectos estruturais determinantes para a especificidade do ligante pela enzima;

3) Propor novos ligantes que apresentem indicativos estruturais que mostrem aumento da especificidade deste pela enzima.

Capítulo 2

Artigo científico

Este capítulo apresenta uma cópia do artigo científico submetido ao *Journal of Molecular Modeling*, bem como cópia da confirmação da submissão.

>Dear Walter
>
>We have received the referee reports for your manuscript: "Identification of new potential Mycobacterium tuberculosis Shikimate Kinase inhibitors through molecular Docking simulations", which you submitted to Journal of Molecular Modeling
>
>TO VIEW REVIEWER ATTACHMENTS (should there be any), please login to the journal site as "Author" and access "Submission Needing Revision" from the Author Main Menu. On the next page display, navigate on the "Action Links" and select "View Reviewer Attachments" from the selection box. This will redirect you to the page that will allow you to "Download" and view the reviewer report attachments. Then, proceed with revising your manusript.
>
>As you can see, the referees have requested some revisions.
>Please revise the manuscript taking the referees remarks into account.
>To avoid misunderstandings it will be helpful if you provide a detailed response to the referee reports.
>
>I would appreciate a short receipt acknowledgement for this message.
>
>Thank you very much in advance. If you have further questions, please do not hesitate to contact me. The editorial office is looking forward to receiving the revised version of your manuscript.
>
>Kind regards,
>
>Andrzej Sokalski, Ph.D., D.Sc.
>Editor in Chief
>
>COMMENTS FOR THE AUTHOR:
>
>Reviewer #1: The manuscript describes the development of a virtual screening study for the Mycobacterium tuberculosis Shikimate Kinase.
>A lot of computational work has been done with technical diligence. The quality of the paper is quite good and it is acceptable for publication after minor changes:
>* In order to further measure the reliability of the VS protocol, you should use an enriched database, adding known Shikimate Kinase inhibitors to the Acros Organics collection and evaluating the number of known compounds recognized as active by the VS protocol.
>* The VS protocol reported in Figure 4, should be widely described in the text.
>* The authors decided to apply the ADMET filter after docking calculations. In

order to reduce the number of compounds to be docked (for larger database), could it be better to apply this filter before the docking calculations? Please discuss it.

>Reviewer #2: This manuscript reports a virtual screening study aimed to the identification of inhbitors of one of the enzymes of the shikimate pathway from Mycobacterium tuberculosis (shikimate kinase). Since the shikimate pathway is essential to M. tuberculosis but is absent in humans compounds inhibiting steps of this pathway are potential anti-tuberculosis drugs.
>Therefore, the rationale of the work is sound and the results derived interesting and of potential use for other researchers. Authors are experienced not only in the field of virtual screening but also have published a molecular model of the enzyme and have been involved in the obtention of the 3D-structure of M. tuberculosis shikimate kinase (MT_SK).
>Papers from other research groups have been previously published on the identification of MT_SK inhbitors as antitubercular drugs by similar procedures (Segura-Cabrera and Rodriguez-Perez (2008) Bioorg. Med. Chem. Lett. 18: 3152-3157 and Kumar et al. 2010 Chem. Biol. Drug. Des. 76: 277-284). Nevertheless, very limited information (experimental data) on the direct effect of these compounds on MT_SK is yet available.

>The novelty of the submitted paper relays on the method used in the virtual screening and the selection of compounds according to their pharmacological properties.
>Lead validation is a key issue in drug discovery and in my oppinion this is the missing point in this study. To increase the value of the submitted work and allow publication authors should give proof of the usefulness of the selected compounds as MT SK inhibitors.
>Ideally this could be done by testing the inhibitory effect of the best compounds, on the enzyme in their lab (Pereira et al. 2004 Acta Cryst. D60: 2310-2319) since compounds belong to a collection of commercially available molecules (Acros Organics). Alternatively,authors could analyse by their screening procedure compounds which have been already tested directly on MT_SK. For instance, three MT SK inhbitors have been recently identified and tested (Mulabagal and Calderon 2010 Anal. Chem. 82: 3616-3621). Analysing these compounds would allow validation of the method and the results obtained.

>Minor points:
>Authors should modifiy the Introduction and the Results and Discussion sections to include the data reported in the references mentioned above (other virtual screening studies and identification of MT_SK inhibitors)
>
>References format should be checked and adapted to JMM guidelines

Identification of new potential *Mycobacterium tuberculosis* shikimate kinase inhibitors through molecular docking simulations

Carolina Pasa Vianna[a,b] and Walter F. de Azevedo Jr.[a,b*]

[a] Faculdade de Biociências, Instituto Nacional de Ciência e Tecnologia em Tuberculose (INCT-TB), Laboratório de Bioquímica Estrutural (LaBioQuest), Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Av. Ipiranga 6681, Porto Alegre, RS 90619-900, Brazil

[b] Programa de Pós Graduação em Biologia Celular e Molecular, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil

[*]To whom correspondence may be addressed: Walter Filgueira de Azevedo Jr.
Av. Ipiranga 6681 – Faculdade de Biociências – Prédio 12C, Porto Alegre RS 90619-900, Brazil. Phone/Fax: +55 51 33204529; E-mail address: walter@azevedolab.net

ABSTRACT

Tuberculosis (TB) is the major cause of human mortality from a curable infectious disease, attacking mainly in developing countries. Among targets identified in *Mycobacterium tuberculosis* genome, enzymes of the shikimate pathway deserve special attention, since they are essential to the survival of the microorganism and absent in mammals. The object of our study is shikimate kinase (SK), the fifth enzyme of this pathway. We applied virtual screening methods in order to identify new potential inhibitors for this enzyme. In this work we employed MOLDOCK program in all molecular docking simulations. Accuracy of enzyme-ligand ***docking was validated*** on a set of 12 SK-ligand complexes for which crystallographic structures were available, generating root-mean square deviations below 2.0 Å. Application of this protocol against a commercially available database allowed identification of new molecules with potential to become drugs against TB. Besides, we have identified the binding cavity residues that are essential to intermolecular interactions of this enzyme.

Keywords: Tuberculosis; Shikimate pathway; Shikimate Kinase; Molecular Docking; Virtual Screening

Introduction

Tuberculosis (TB) is the most important cause of human death from a curable infectious disease. It is estimated that, worldwide, one hundred million people are infected annually and about ten million develop the disease, with five million of those progressing to an infectious stage, culminating with approximately three million deaths. According to the World Health Organization [1], the overall incidence of TB increases approximately 0.3% per year. The resurgence of this health problem occurred mainly due to the proliferation of multi (MDR-TB), extensively (XDR-TB), and recently, totally-drug (TDR-TB) resistant Mt strains. Besides, the high susceptibility of HIV/AIDS infected patients to TB is also a health problem. Therefore, there is an urgent need for the discovery and development of new and better drugs for the TB treatment [2].

Enzymes of the shikimate pathway (SP) are promising targets for the development of antimicrobial agents [3] and herbicides [4], because they are essential to the survival of algae, higher plants, bacteria, fungi, apicomplexan parasites and absent in mammals [5]. It is a seven-step biosynthetic route that converts erythrose 4-phosphate to chorismate, a precursor of aromatic amino acids and many other essential compounds [6].

The object of our study is the fifth enzyme of the SP, shikimate kinase (SK) (EC 2.7.1.71), which catalyzes the specific phosphorylation of the 3-hydroxy group of shikimate using ATP as a co-substrate resulting in shikimate-3-phosphate and ADP [7-8]. This enzyme is an established target against Mt, since Parish and Stoeker demonstrated that the SP is essential for the viability of Mt due to the disruption of the aroK gene, which codes for the SK enzyme [9].

SK is a member of the nucleoside monophosphate kinases (NMP kinases) family, which suffer large conformational changes during catalyses (Fig. 1) [10]. The enzymes of this family are composed of three domains: the CORE, which contains a highly conserved phosphate-binding loop (P-loop), the LID domain, which undergoes substantial structural changes upon substrate binding, and the NMP-binding domain which is responsible for the recognition and binding of a specific substrate [11].

Drugs are usually discovered by trial and error by means of high-throughput screening approaches that use *in vitro* experiments to evaluate the activity of a large

number of compounds against a known target. This procedure is very costly and time-consuming. If crystallographic information is available for the protein target, then molecular docking simulations can be a helpful computational approach in the drug-discovery process. [12]. Molecular docking is a simulation method that predicts the conformation of a receptor-ligand complex, in which the receptor can be either a protein or a nucleic acid, and the ligand is a small molecule. This computer simulation can generate many possible positions for the ligand in the receptor-binding pocket. Therefore, a criterion is necessary that will allow comparisons of all possible positions of the ligand, and then a selection can be made for the best position.

Our goal here is to find potential inhibitors against shikimate kinase from *Mycobacterium tuberculosis* MtSK using virtual screening (VS). VS can decrease costs and improve hits rates for lead discovery. For this, we used the MtSK structure [13] as a target for the molecular docking simulations with MOLDOCK [14]. Our docking protocol was validated against an ensemble of 12 crystallographic structures available for complexes of MtSK. The VS was validated by inclusion of a known SK inhibitor in the small-molecule database with over 4500 structures. We describe the results obtained in terms of the MOLDOCK scores, modes of interaction and discuss the importance of the active site residues in the ligand binding process.

Materials and Methods

Molecular docking simulations

One of the fundamental questions in structural biology is the study of protein-ligand interactions, particularly considering the pharmacological applications of such study in the design of drugs based on structure [15]. To simulate the interaction of MtSK with a library of ligands, we used the MOLDOCK program [14], an implementation of a variant of the evolutionary algorithm (EA). Recent evaluation of MOLDOCK strongly indicates that it is capable of finding the right position of a ligand. Furthermore, MOLDOCK exhibits better overall performance compared with SURFLEX, FLEXX, and GOLD [14]. In the present work, all simulations were performed in an iMac (Intel Processor Core 2 Duo, 2.66 GHz, 2 GB SDRAM DDR3 1066 MHz).

Re-docking and cross-docking

In molecular docking simulations, the best binary complex (protein-ligand) is the one closer to the crystallographic structure. For that reason we must establish a methodology that assesses the distance from the computer-generated solution (pose) to the crystallographic structure. This distance can be calculated using the root-mean-square deviation (RMSD), which is a measure of the differences between values predicted by a model and the values actually observed from the object being modeled or estimated (protein-ligand complex). The RMSD is calculated between two sets of atomic coordinates, in this case, one for the crystallographic structure (x_{ctal}, y_{ctal}, z_{ctal}; the object being modeled) and another for the atomic coordinates obtained from the docking simulations (x_{pose}, y_{pose}, z_{pose}; predicted model). A summation is then taken over all N atoms being compared, using the following equation:

$$\mathrm{RMSD} = \sqrt{\frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}(\mathrm{x}_{\mathrm{ctal,i}} - \mathrm{x}_{\mathrm{pose,i}})^2 + (\mathrm{y}_{\mathrm{ctal,i}} - \mathrm{y}_{\mathrm{pose,i}})^2 + (\mathrm{z}_{\mathrm{ctal,i}} - \mathrm{z}_{\mathrm{pose,i}})^2} \ .$$

In docking simulations, it is expected that the best results generate RMSD values less than 2.0 Å compared with crystallographic structures [16]. This procedure of obtaining the crystallographic position of the ligand is often called “re-docking,” which is fundamentally a validation method that determines whether the molecular docking algorithm is able to recover the crystallographic position using computer simulation. In this work, all RMSD calculations were calculated for non-hydrogen atoms.

In order to validate our docking protocol, we used the SK crystallographic coordinates available at the protein data bank (PDB), under the access code 2DFN [13]. We performed the docking simulation against the active site of MtSK and compared the docked poses with the crystallographic structure. We used the MOLDOCK default protocol with center at coordinates x=(-15.23), y=(-14.38), and z=(14.88) Å, and a docking sphere radius of 9 Å. Fig. 2 shows the docking sphere used in the simulations.

In the implementation of EA in MOLDOCK, computational approximations of an evolution course, called genetic operators, are applied to simulate the permanence of the most positive features. In a sample space, where there is a problem or a search routine and many different possible solutions (candidates), each option is ranked based on a set of parameters (scoring function or fitness function),

and only the best ranked solutions are kept for the next iteration. This cycle is repeated until an optimal solution can be found. In the molecular docking simulations, the optimal solution is the one with the best scoring function, which should be the closest to the crystallographic structure. MOLDOCK presents two biological inspired algorithms to perform positional searches in docking simulations. One is called the optimizer search algorithm (MOLDOCK Optimizer), which is based on an EA [14]. The second is a guided differential evolution algorithm (GDEA) called MOLDOCK SE. GDEA is based on an EA adjustment called differential evolution (DE), which provides a distinct method for selecting and modifying candidate solutions (individuals). We used MOLDOCK Optimizer as search algorithm.

In addition to re-docking, a procedure called “cross-docking” can also be used to further validate a docking protocol. Considering that several crystallographic structures are available for the same protein, cross-docking can be applied. This procedure involves docking a number of ligands found in a variety of crystal structures of a protein identical to a single rigid protein crystallographic conformation [17]. When a protein target presents major conformational changes upon ligand binding, a significant difference is expected between the crystallographic and docked structures. We identified 12 MtSK structures in PDB with ligands in the shikimate-binding site (PDB access codes: 2DFN, 1U8A, 1WE2, 1ZYU, 2G1K, 2IYQ, 2IYR, 2IYS, 2IYX, 2IYY, 2IYZ, and 3BAF). This search was performed on February, 24th 2011. This validation procedures, re-docking and cross-docking, is the initial stage of a virtual screening protocol (**phase 1**) described in the next sections.

Virtual screening

Our virtual screening (VS) protocol is divided in 4 phases as shown in Fig. 3. **Phase 1** is focused on selection and validation of a docking protocol, as described earlier in the section re-docking and cross-docking. **Phase 1** ends when an adequate protocol is found (selection criterion RMSD < 2.0 Å). It should be pointed out that the RMSD criterion is dependent on the number of torsion angles, and a less demanding criterion may be adopted for re-docking of a ligand with a number of torsion angles higher than 10 [14]. Once a docking protocol is chosen we select a small-molecule database to be used in the screening (**phase 2**). Here we used a ligand library commercially obtainable at Acros Organics. The ligands (mol2 format) were

downloaded from http://zinc.dock.org [18], with a total of 4579 small molecules. In addition to commercially oriented databases the ZINC database also provides an interface to build small-molecule databases based on molecular similarity, such as Tanimoto coefficient [19, 20].

In **phase 3**, we start docking simulations for each ligand present in the selected database. MOLDOCK program is the workhorse of the present protocol. It was used in all docking simulations described here. During a typical docking simulation several orientations can be obtained for each ligand. Here we selected the one with the lowest scoring function. The scoring function used by MOLDOCK improves accuracy of scoring functions with a new hydrogen bonding term and new charge schemes. Four scoring functions are implemented in the MOLDOCK, including MOLDOCK score and PLANTS score [14, 21]. These two functions offer grid-based versions, in which hydrogen bond directionality is not considered. In the present protocol we employed grid-based MOLDOCK score since it offers approximately four-fold greater speed by performing a precalculation of potential-energy values on an equally spaced cubic grid.

The MOLDOCK score is based on the piecewise linear potential (PLP) scoring functions developed by Yang et al. [22, 23]. The docking scoring function $E_{MOLDOCK\ SCORE}$ is defined as the following:

$$E_{\text{MOLDOCK SCORE}} = E_{\text{Intramol}} + E_{\text{intermol}}$$

where $E_{intermol}$ is the intermolecular interaction energy:

$$E_{\text{intermol}} = \sum_{i \in \text{ligand}} \sum_{j \in \text{protein}} \left[332 \frac{q_i q_j}{D r_{ij}} + E_{\text{PLP}}(r_{ij}) \right]$$

All non-hydrogen atoms in the ligand and protein are taken in the summation. The first term accounts for electrostatic interactions, in which the factor 332 is used to obtain energy in kJ/mol. *D* represents the dielectric constant, which is the following: $D = 4r_{ij}$. The second term (E_{PLP}) is a PLP, described elsewhere [22, 23]. To ensure that no energy term can be superior to the clash penalty, the electrostatic term is cut off at a level equivalent to the distance of 2.0 Å for distances less than 2.0 Å.

Intramolecular energy is given by the following equation:

$$E_{\text{intramol}} = E_{\text{penalty}} + \sum_{i \in \text{ligand}} \sum_{j \in \text{ligand}} E_{\text{PLP}}(r_{ij}) + \sum_{\text{single bonds}} A\left[1 - \cos\left(n\varphi - \varphi_0\right)\right]$$

The term $E_{penalty}$ is a penalty energy to be added to $E_{intramol}$ when two non-bonded atoms are closer than 2 Å (for non-hydrogen atoms). This term avoids unrealistic molecular topologies for the ligands. The second term is a PLP, already mentioned [22, 23]. The last term accounts for torsion energy, which is expressed as a periodic function. In this term, A, n, and ϕ_o are empirically determined [22, 23]. MOLDOCK defines a limiting sphere where the search is focused. If a ligand non-hydrogen atom is positioned outside this limiting sphere (the search space sphere), then a constant penalty of 10000 is added to the total energy (implemented for the grid-based version of the MOLDOCK score).

After identification of potential inhibitors by molecular docking simulations, the best scored ligands were submitted to the web server FAF-Drugs [24], in order to assess physical-chemical properties (**phase 4**). These are key properties that need to be considered in early stages of the drug discovery process, and FAF-Drugs allows users to filter molecules via simple rules such as molecular weight, polar surface area, logP and number of rotatable bonds. The ligands were filtered following the Lipinski's rule of five (RO5). RO5 advocates that drugs which present oral bioavailability, in general, follow: molecular weight less or equal to 500, Log*P* less or equal to 5, number of hydrogen bond donor groups less or equal to 5 and number of hydrogen bond acceptor groups less or equal to 10 [25].

In addition to the 4579 small molecules present in the Acros database we added staurosporine (PubChem Compound Identification: CID 44259) to the database to be used in the VS. This molecule has been already tested directly on MtSK [26]. This addition allows testing whether this VS protocol is able to identify SK inhibitors present in a database with over 4,500 ligands.

Results and Discussion

Docking and cross-docking

Re-docking simulations (phase 1 of the VS protocol) using the structure 2DFN generated an RMSD of 1.6 Å. In addition, cross-docking simulations generated RMSD ranging from 1 to 2 Å further validating the present docking protocol. These two tests indicated that the docking simulation was successful, and that the protocol is good enough to be used for the virtual screening process.

Virtual Screening

VS uses computational methodologies to identify biologically active molecules against a specific protein target. Two main methodologies are used in VS. Methods that search for similarity to validated ligands and molecular docking methods that require the use of crystallographic information of the target. Here we made use of the second approach. VS studies performed by other research groups have been previously published on the identification of MtSK inhibitors as antitubercular drugs by similar molecular docking procedures [27, 28]. Nevertheless, very limited information (experimental data) on the direct effect of these compounds on MtSK is yet available. The novelty of the present work relays on the method used in the VS and the selection of compounds according to their pharmacological properties. The VS simulations were carried out using the MOLDOCK program, having as target the MtSK (PDB access code 2DFN). The ligand library comprises 4580 molecules (Across database plus staurosporine). Addition of a known SK inhibitor allows testing the accuracy of the present protocol.

After docking simulations, we selected 20 top-scoring compounds from the initial set of 4580 compounds (selection based on MOLDOCK score). Staurosporine was present in the 20 top-scoring compounds obtained in the VS, with MOLDOCK score of -144.168. Identification of a known SK inhibitor among the best VS results gives further validation for this VS protocol. These 20 potential inhibitors were submitted to filter tests, available at the web server FAF-Drugs [24], to exclude those compounds that have known undesirable physical-chemical features to oral bioavailability. We could have applied this filter analysis previous to docking simulations, since it would reduce simulation time. Nevertheless, we kept filtering analysis after docking simulation, since the MOLDOCK protocol was fast enough to be run in less than a week of CPU time of an iMac (Intel Processor Core 2 Duo, 2.66 GHz, 2 GB SDRAM DDR3 1066 MHz). In addition, application of filtering analysis previous to docking simulations could eliminate candidates that fail to filtering analysis but present promising MOLDOCK score, which could have toxicity reduced by small modification in the structure.

Especially interesting is the fact that staurosporine is a well-known cyclin-dependent kinase (CDK) inhibitor that has a plethora of structural and functional studies [29-33]. Staurosporine is non-selective and too toxic for use in therapy, but

UCN-01, a hydroxylated form of staurosporine (7-hydroxystaurosporine), shows greater selectivity for CDK and is currently undergoing clinical trials in the United States and Japan [31]. This opens new possibilities to test new molecular moieties as potential SK inhibitors, the CDK inhibitors that have already shown low toxic effects make a promising dataset to be explored as potential SK inhibitors.

The FAF-Drugs parameters used were those of the Lipinski's role of five [25]. From the set with 20 selected molecules, 9 fit the Lipinski's role of five, which includes stauorosporine. Figures 4A-4I show the molecular structures for all 9 ligands. Since staurosporine is already a known SK inhibitor we excluded it from the rest of the analysis. Staurosporine was included only to test the VS protocol. The selected ligands are shown in Table 1. The MOLDOCK scores for these 8 molecules ranging from -144.208 to -151.943. All 8 ligands show MOLDOCK scores better than staurosporine.

Intermolecular interactions

In order to better understand the interactions of these 8 molecules with MtSK, we used the program LIGPLOT [34] to access the atoms of both, the small molecules and the protein ones that are responsible to make hydrogen bonds and van der Waals contacts. A comparison among the MOLDOCK score values obtained for these ligands, is not enough yet to predict activity, since in vitro assays are necessary to conclude this. Therefore it is not possible to say that the selected compounds, the ones with the best MOLDOCK scores, would be the most potent ones. We could observe, only, that among the selected compounds the best scores mean a greater potential to interact with the shikimate-binding cavity.

The docking simulation results corroborate the importance of some shikimate-active site residues as responsible to establish intermolecular interactions with the substrate as well as with the tested ligands. The binding of shikimate to its cavity, presents pivotal residues that make protein-ligand interactions possible, as shown in Fig. 5. These residues are essential to the ligand binding and, finally, to the reaction catalyzed by the enzyme. The SK residues that perform intermolecular hydrogen bonds (HB) with the shikimate are: Gly80, Arg136 and Arg58. SK makes van der Waal contacts with residues: Ile45, Asp34, Pro11, Pro118, Gly79, Phe57, Leu119 and Gly81.

Information about intermolecular interactions for all 8 top-scoring compounds is summarized in Table 2. Analysis of shikimate-binding site indicated that all top-scoring compounds present interaction with residues Lys15, Ser16 and Arg117. Fig. 6 shows the intermolecular interactions for the top-scoring compound (ligand 1, ZINC15707201). This figure is representative of the positioning of all top-scoring compounds in the shikimate-binding pocket. Ligands 1, 2, 4, 5, 6 and 8 highlight the presence of residue Val116, suggesting it is also relevant to intermolecular interactions. Two previously published molecular docking studies focused on MtSK were able to identify intermolecular molecular interactions with the same residues [27, 28], further corroborating the pivotal importance of these residues for ligand-binding affinity. Especially interesting is the fact that these previous molecular docking studies analyzed completely different molecular moieties, such as dipeptides (arginine-aspartate/lysine-aspartate) [28] and triazole/tetrazole heteroaromatic systems [27]. These molecular structures were not present in the database used in the present study.

Conclusions

Advanced molecular docking algorithms available nowadays make possible to undertake larger virtual screening studies focused on small-molecules libraries up to millions of compounds. Here we described an efficient molecular docking protocol, which was able to recover crystallographic position of a ligand present in the active site of the SK. Re-docking and cross-docking simulations generated RMSD results below 2 Å. The virtual screening protocol was able to confirm a known SK inhibitor, staurosporine, as a top-scoring compound. Furthermore, the present work indicates new molecules with the potential to become drugs against TB. Besides, we identified the MtSK binding-cavity residues that are essential to make possible the interactions of this enzyme with a variety of molecules. Analysis of the top-scoring compounds also indicates that MtSK has the ability to bind a variety of molecular moieties not previously identified.

Acknowledgements

The authors would like to express their gratitude to the reviewers for their valuable comments and suggestions. This work was supported by National Institute

of Science and Technology on Tuberculosis (Decit/SCTIE/MS-MCT-CNPq-FNDCTCAPES). W.F.A. Jr. is research fellow of the National Council for Scientific and Technological Development of Brazil (CNPq). C.P.V. acknowledges a scholarship awarded by CAPES.

References

1. World Health Organization (2009) WHO Report. Geneva. Switzerland. WHO/CDS/TB/2009. http://www.who.int/topics/tuberculosis/en/. Accessed in October 2010

2. Basso LA, Pereira Da Silva LH, Fett-Neto AG, De Azevedo Jr WF, Moreira IS, Palma MS, Calixto JB, Astolfi Filho S, dos Santos RR, Soares MBP, Santos DS (2005) Mem Inst Oswaldo Cruz 100(6): 475-506

3. Davies GM, Barrett-Bee KJ, Jude DA, Lehan M, Nichols WW, Pinder PE, Thain JL, Watkins WJ, and Wilson RG (1994) Antimicrobial Agents and Chemotherapy 38: 403-406

4. Coggins JR (1989) The shikimate pathway as a target for herbicides. In : A.D.Dodge (ed) Herbicides and Plant metabolism. Cambridge University Press, Cambridge, UK, pp 97-112

5. Bentley R (1990) The shikimate pathway-a metabolic tree with many branches. Crit Rev Biochem Mol Biol 25**:**307-384

6. Ratledge C (1982) Nutrition, growth and metabolism. In: C Ratledge, JL Stanford (eds). The Biology of the Mycobacteria, vol 1. Academic Press, London, pp 185-271

7. Krell T, Coggins JR and Lapthorn AJ (1998) The three-dimensional structure of shikimate kinase. J Mol Biol 278: 983-997

8. Pereira JH, Oliveira JS, Canduri F, Dias MVB, Palma MS, Basso LA, Santos DS, & De Azevedo WF (2004) Acta Crystallogr Sect D-Biol Crystallogr 60: 2310-2319

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

10. Vonrhein C, Schlauderer GJ & Schulz GE (1995) Movie of the Structural Changes During a Catalytic Cycle of Nucleoside Monophosphate Kinases. Structure 3: 483-490

11. Gu Y, Reshetnikova L, Li Y, Wu Y, Yan H, Singh S, Ji X (2002) Crystal Structure of Shikimate Kinase from Mycobacterium tuberculosis Reveals the Dynamic Role of

the LID Domain in Catalysis. J Mol Biol 319: 779–789
12. Peitsch MC, (2004) Manuel Peitsch discusses knowledge management and informatics in drug discovery. Drug Discovery Today: Biocilico, v 02, p 94-96
13. Dias MV, Vasconcelos IB, Prado AM, Fadel V, Basso LA, De Azevedo WF Jr, Santos DS (2007) Acta Crystallogr Sect F Struct Biol Cryst Commun 63(Pt 1):1-6
14. Thomsen R, Christensen MH (2006) MolDock: A New Technique for High-Accuracy Molecular Docking. J Med Chem 11: 3315-3321
15. De Azevedo Jr WF, Canduri F, Da Silveira NJF (2002) Biochem Biophys Res Commun 293: 566-571
16. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shaw DE, Shelley M, Perry JK, Francis P, Shenkin PS, Glide (2004) A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. J Med Chem 47: 1739–1749
17. Thilagavathi R., Mancera RL (2010) Ligand-protein cross-docking with water molecules. J Chem Inf Model 50: 415-21
18. Irwin JJ, Shoichet BK (2005). ZINC - A Free Database of Commercially Available Compounds for Virtual Screening. J Chem Inf Model 45: 177-182
19. Timmers LFS, Pauli I, Caceres RA, de Azevedo Jr WF (2008) Drug-Binding Databases. Curr Drug Targets 9(12): 1092-1099
20. De Azevedo WF Jr (2010) Structure-Based Virtual Screening. Curr Drug Targets 11(3): 261-263
21. De Azevedo WF Jr (2010) MolDock Applied to Structure-Based Virtual Screening. Curr Drug Targets 11(3): 327-334
22. Yang JM, (2004) Development and evaluation of a generic evolutionary method for protein-ligand docking. J Comput Chem 25: 843-57
23. Yang JM, Chen CC (2004**)** GEMDOCK: a generic evolutionary method for molecular docking. Proteins 55: 288-304
24. Miteva MA, Violas S, Montes M, Gomez D, Tuffery P, Villoutreix BO (2006) FAF-drugs: free adme/tox filtering of compound collections. Nucleic Acids Res 34: W738–W744
25. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews 23: 3-25

26. Mulabagal V, Calderón AI (2010) Development of an ultrafiltration-liquid chromatography/mass spectrometry (UF-LC/MS) based ligand-binding assay and an LC/MS based functional assay for *Mycobacterium tuberculosis* shikimate kinase. Anal Chem 82(9):3616-21

27. Segura-Cabrera A and MA Rodriguez-Perez (2008) Structure-based prediction of *Mycobacterium tuberculosis* shikimate kinase inhibitors by high-throughput virtual screening. Bioorg Med Chem Lett 18:3152-3157

28. Kumar et al (2010) Chem Biol. Drug Des 76: 277-284

29. Noble ME, Endicott JA, Johnson LN (2004) Protein kinase inhibitors: insights into drug design from structure. Science 303:1800-1805

30. Canduri F, De Azevedo WF Jr (2005) Structural basis for interaction of inhibitors with Cyclin-Dependent Kinase 2. Curr Computer-Aided Drug Design 1(1): 53-64

31. Lawrie AM, Noble ME, Tunnah P, Brown NR, Johnson LN, Endicott JA (1997) Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2. Nat Struct Biol 4: 796-801

32. De Azevedo Jr WF, Mueller-Dieckmann H-J, Schulze-Gahmen U, Worland PJ, Sausville E, Kim SH (1996) Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase. Proc Natl Acad Sci. USA 93(7): 2735-2740

33. De Azevedo WF, Leclerc S, Meijer L, Havlicek L, Strnad M, Kim SH (1997) Inhibition of cyclin-dependent kinases by purine analogues: crystal structure of human cdk2 complexed with roscovitine. Eur J Biochem 243(1-2): 518-526

34. Wallace AC, Laskowski RA & Thornton JM (1995). LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. Prot. Eng. 8: 127-134

Figure Legends

Fig. 1 Structure of shikimate kinase in complex with ADP and shikimate (PDB access code: 2DFN).

Fig. 2 Search space sphere (green) defined for molecular docking simulations.

Fig. 3 Flowchart of virtual screening process.

Fig. 4 Molecular structures of the top-scoring compounds identified in the VS protocol. A) Staurosporine. B) ZINC15707201. C) ZINC20462780. D) ZINC15707234. E) ZINC15675581. F) ZINC15707188. G) ZINC22936889. H) ZINC20464408. I) ZINC22936937.

Fig. 5 Shikimate-binding pocket with main residues found in intermolecular interactions with shikimate.

Fig. 6 Shikimate-binding pocket with main residues found in intermolecular interactions with the top-scoring compound (ZINC15707201).

Table Legends

Table 1 Physical-chemical properties of ligands that fitted the Lipinki's role of five after analysis by FAF-Drugs.

Table 2 Detailed intermolecular interactions for the ligands selected in the VS procedure. The presence of an X indicates that the interaction of a certain ligand with another certain protein amino acid is observed. HB means hydrogen bonds and VDW means van der Waals contacts.

Figure 1

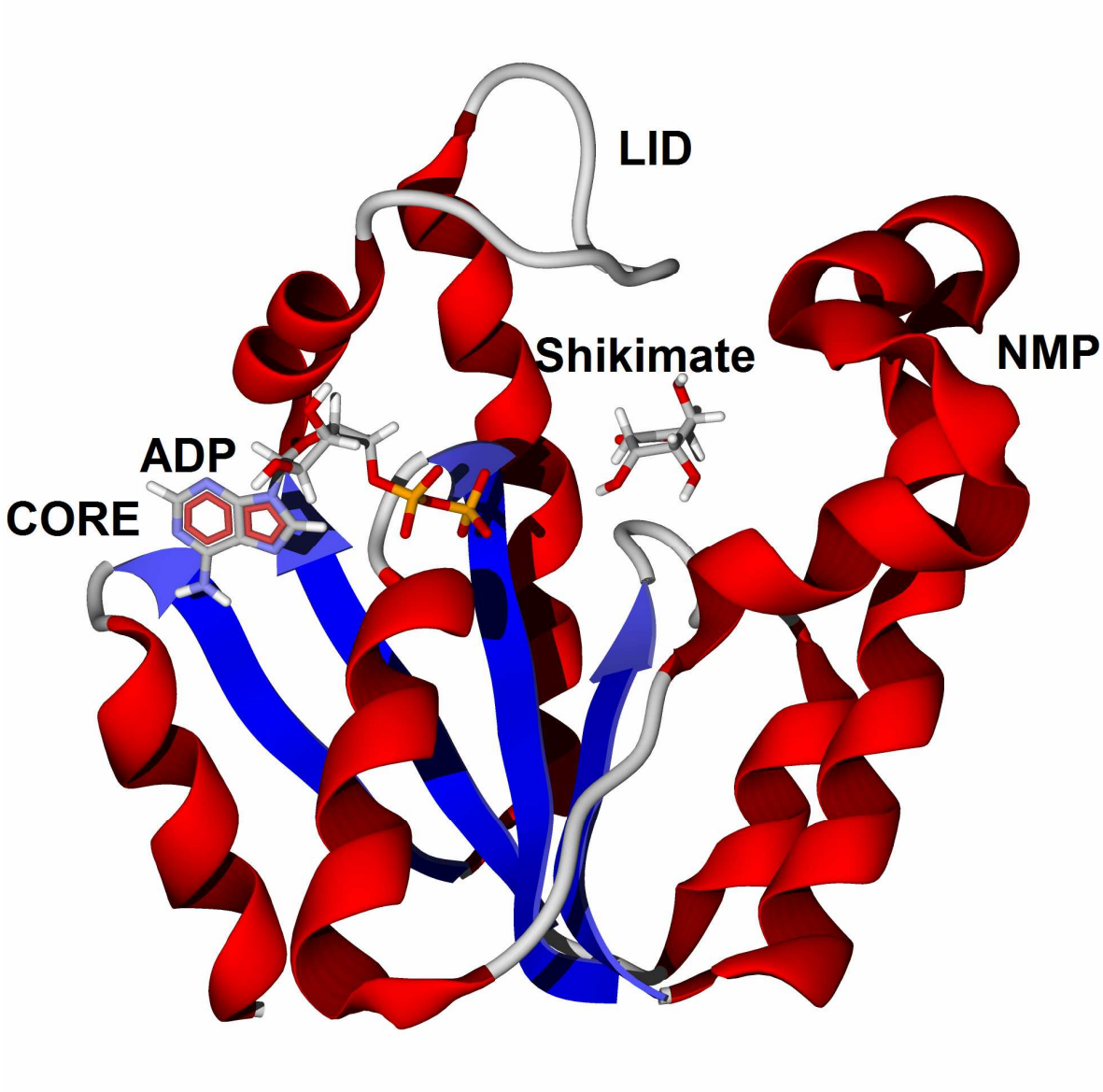


Figure 2

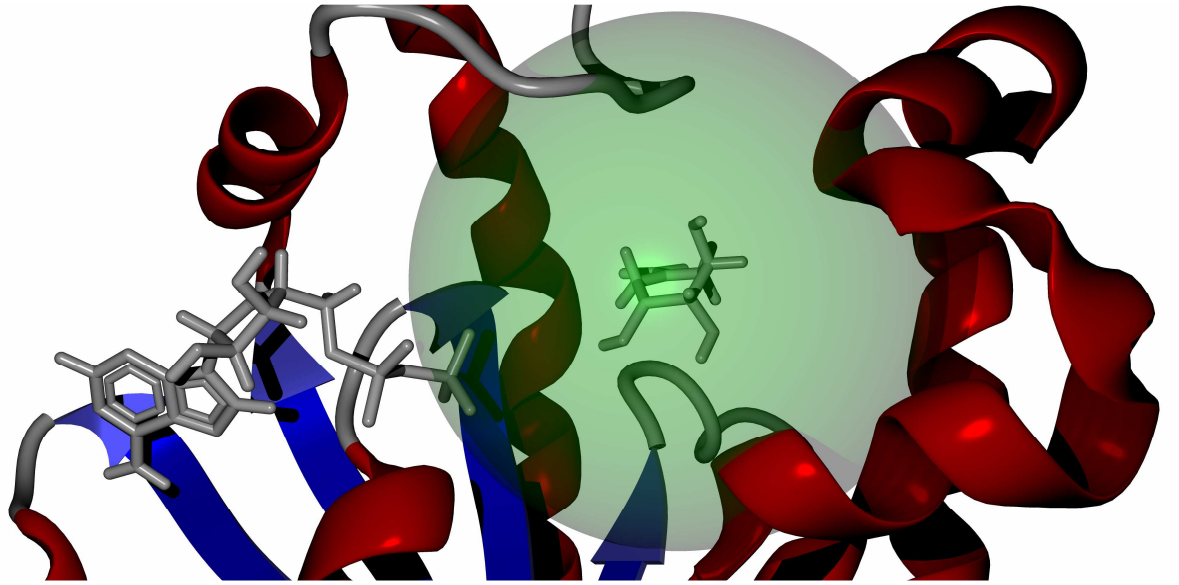


Figure 3

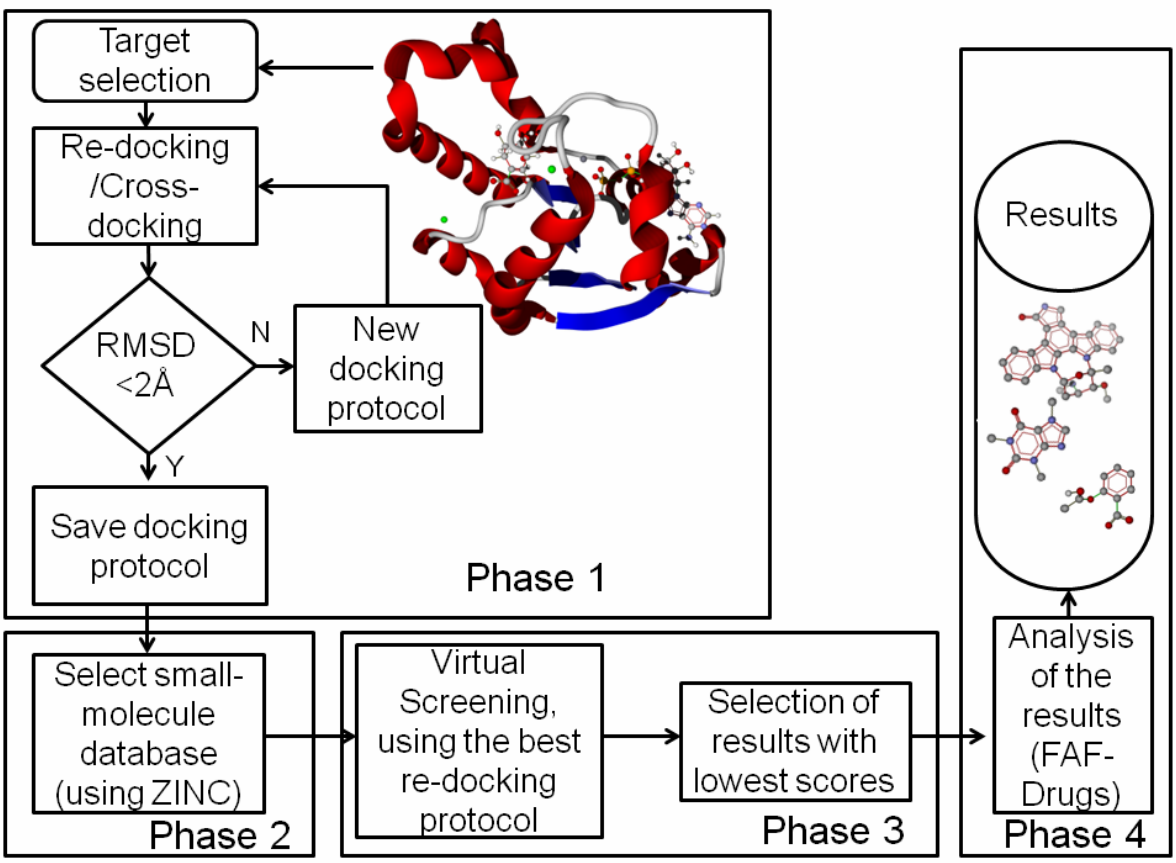

Target selection
Re-docking /Cross-docking
RMSD <2Å
N
New docking protocol
Y
Save docking protocol
Phase 1
Select small-molecule database (using ZINC)
Phase 2
Virtual Screening, using the best re-docking protocol
Selection of results with lowest scores
Phase 3
Results
Analysis of the results (FAF-Drugs)
Phase 4

Figure 4A

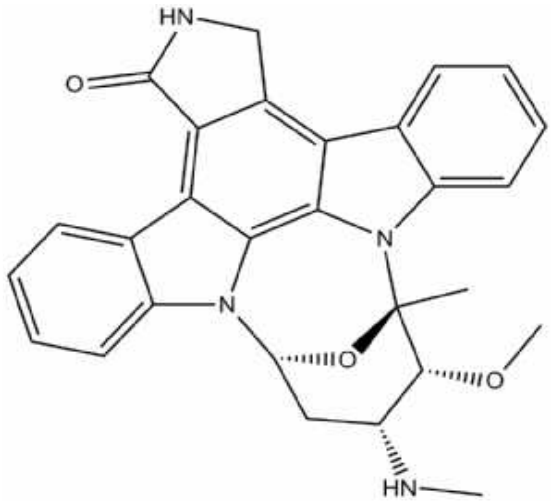


Figure 4B

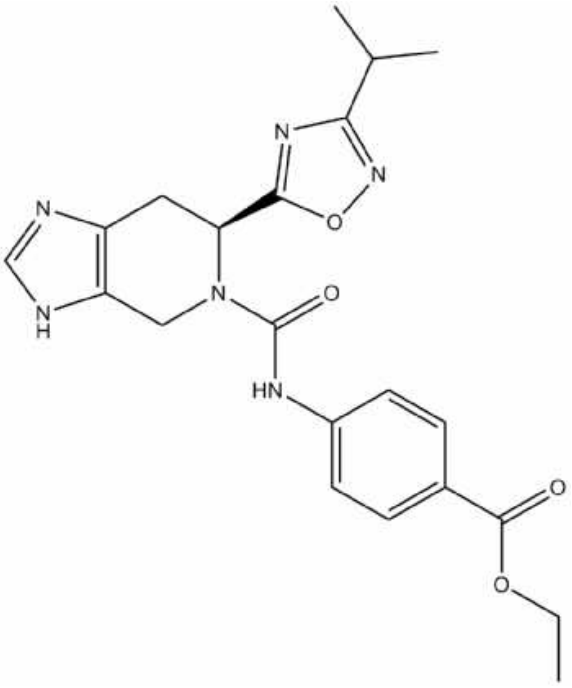


Figure 4C

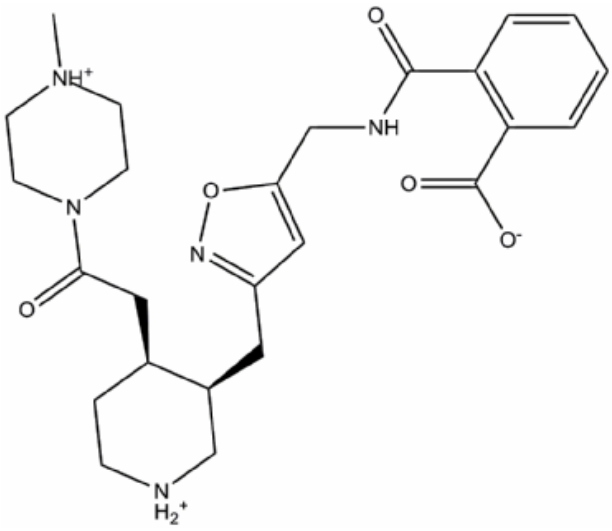


Figure 4D

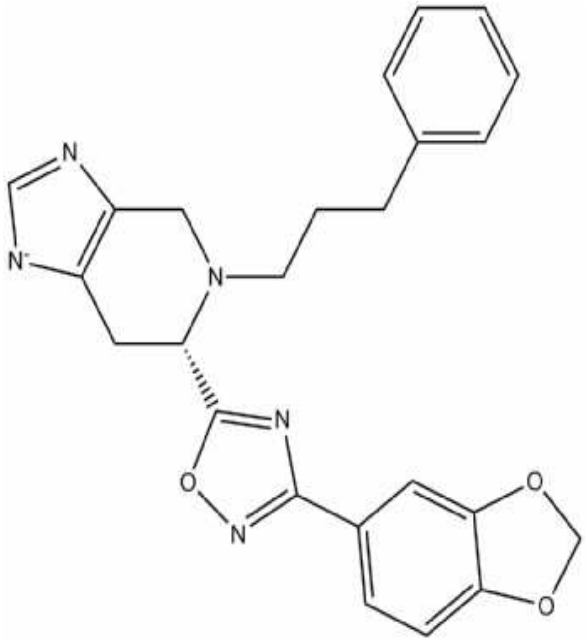


Figure 4E

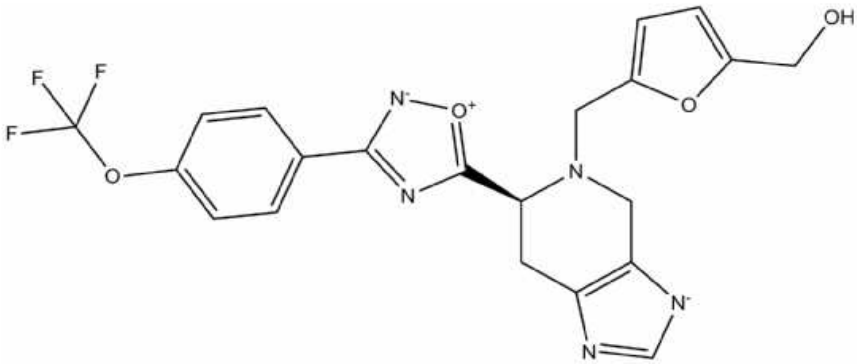


Figure 4F

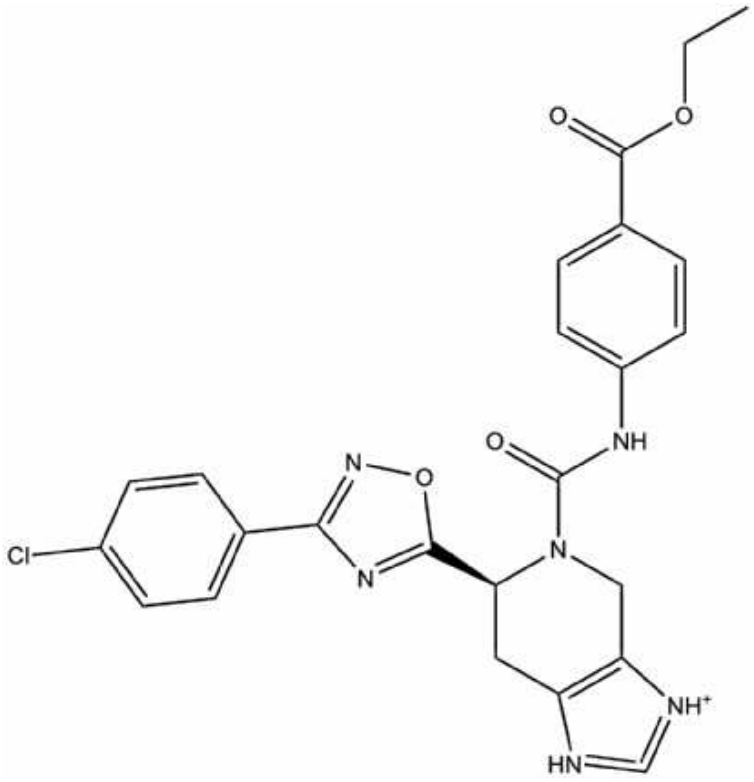


Figure 4G

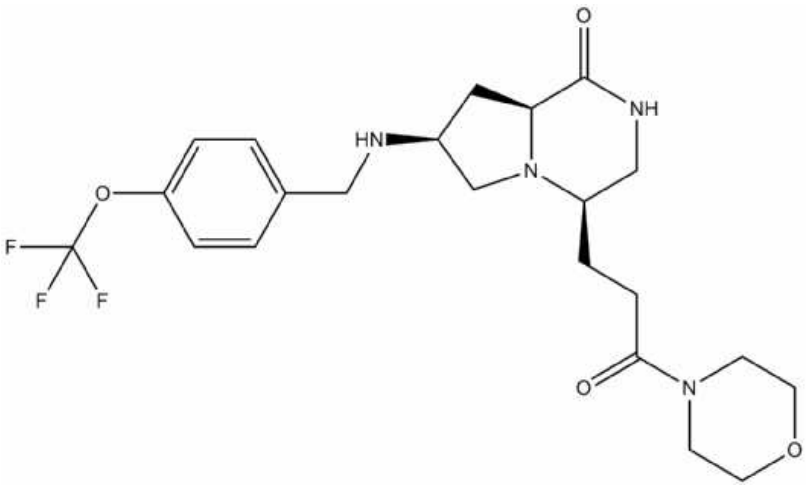


Figure 4H

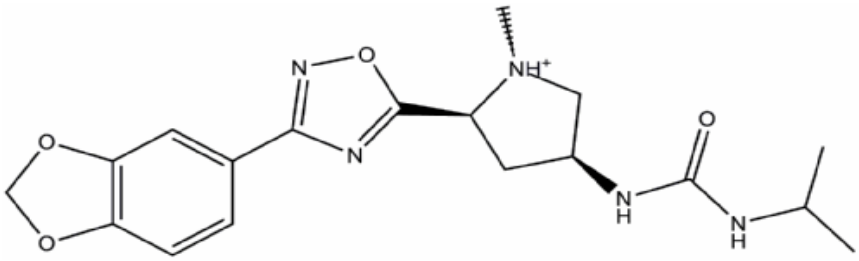


Figure 4I

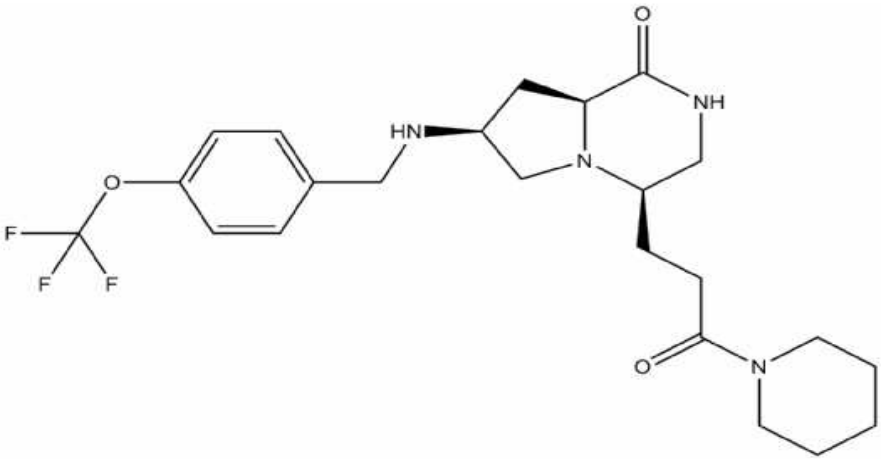


Figure 5

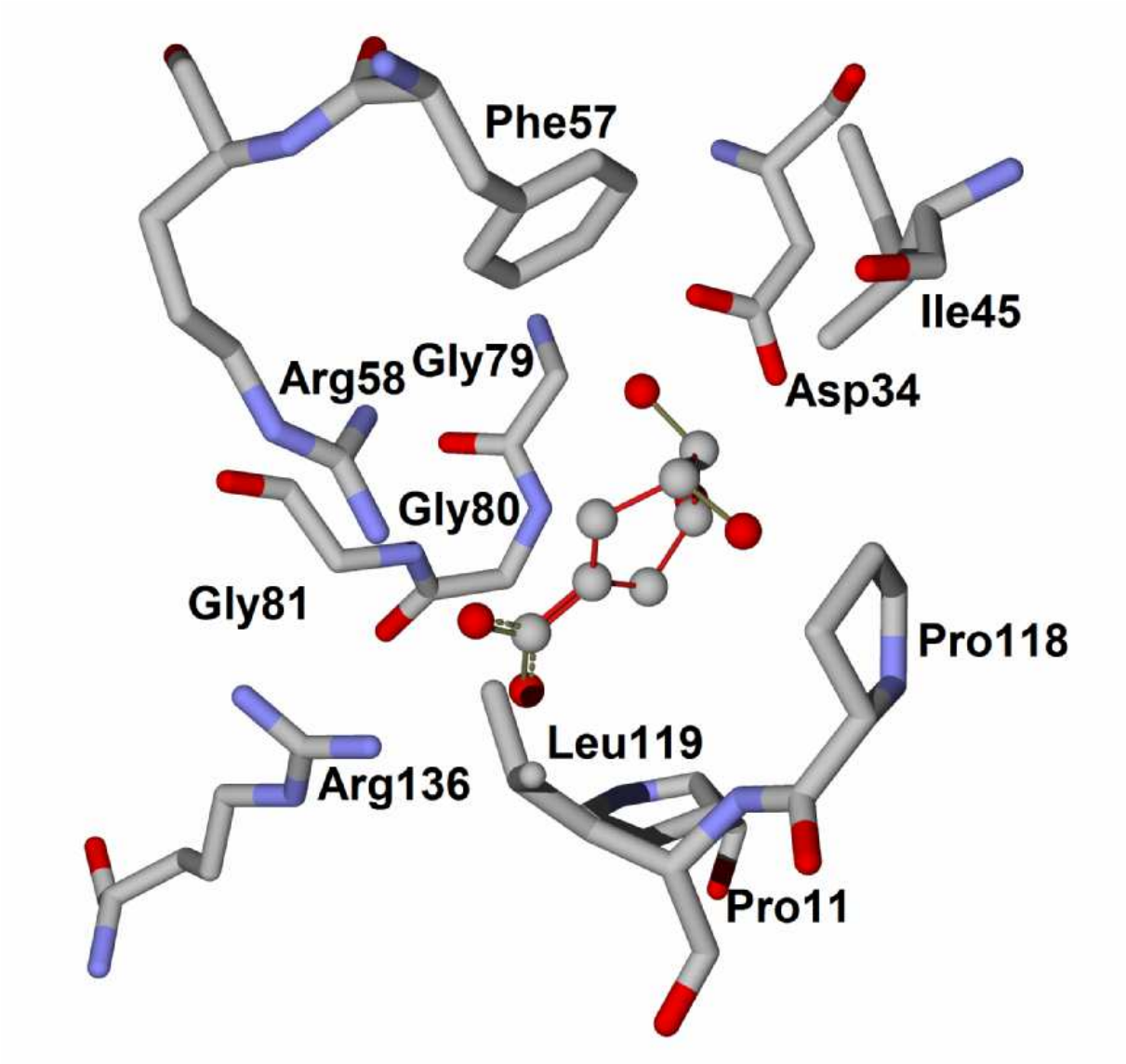

Phe57
Ile45
Asp34
Arg58
Gly79
Gly80
Gly81
Pro118
Leu119
Arg136
Pro11

Figure 6

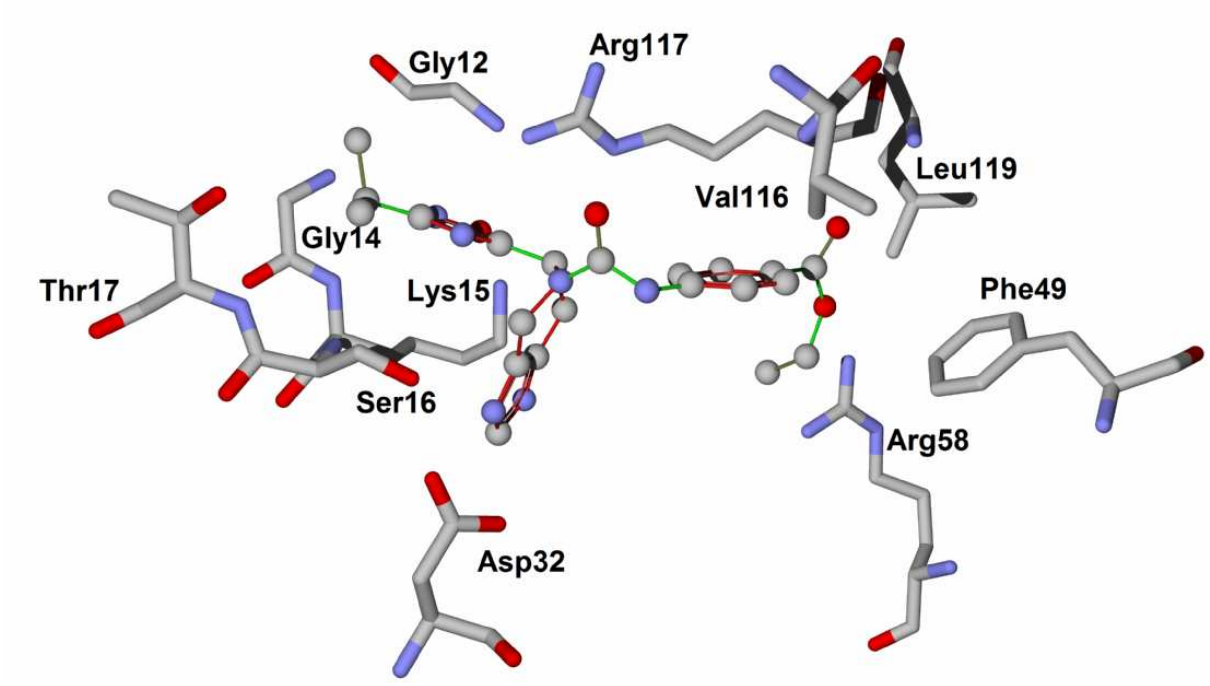

Gly12
Arg117
Leu119
Val116
Gly14
Thr17
Lys15
Phe49
Ser16
Arg58
Asp32

Table 1

Ligand	ZINC Code	Molecular weight (Da)	Number of HB acceptors	Number of HB donnors	LogP
1	15707201	423.3	10	1	0.82
2	20462780	485.3	10	4	-0.3
3	15707234	428.3	8	0	2.62
4	15675581	460.3	9	2	1.57
5	15707188	490.8	10	1	2.61
6	22936889	471.3	8	3	1.27
7	20464408	373.2	9	2	1.5
8	22936937	469.3	7	3	2.54

Table 2

Resídues	Ligands							
HB	**Zinc15675581**	**Zinc15707188**	**Zinc15707201**	**Zinc15707234**	**Zinc20462780**	**Zinc20464408**	**Zinc22936889**	**Zinc22936937**
Gly12						X		
Ser13						X		
Gly14						X		
Lys15	X	X	X	X	X	X		
Ser16		X	X					X
Thr17	X					X	X	X
Asp32		X	X					
Asp34	X						X	X
Leu78		X						
Arg117	X	X		X	X	X		
Leu119		X	X	X		X	X	
VDW								
Leu10					X			
Gly12	X	X	X	X		X		
Gly14	X	X	X	X		X	X	X
Lys15							X	X
Ser16	X			X	X	X	X	
Thr17		X	X	X	X			
Val35	X							
Gly38	X			X				
Phe49	X	X	X	X		X		X
Arg58			X					X
Gly80					X			
Val116	X	X	X		X		X	X
Arg117			X				X	X

Capítulo 3
CONSIDERAÇÕES FINAIS

Há uma grande busca por terapias mais eficientes para as doenças que acometem a humanidade. Sempre haverá a necessidade da introdução de novos fármacos no arsenal terapêutico, seja pela falta de eficiência dos fármacos atuais, seja pelo alto nível de toxicidade, pelo surgimento de novas doenças e de cepas resistentes aos atuais fármacos utilizados. Aliás, talvez, não seja só a descoberta de novas moléculas promissoras, mas sim a descobertas de propriedades e interações de moléculas já existentes.

A tuberculose está em destaque no que tange ao desenvolvimento racional de drogas, devido ao aumento de casos, em conjunto com o surgimento de cepas resistentes aos medicamentos existentes e também ao surgimento da HIV/AIDS, aumentando assim os casos de tuberculose devido a baixa imunidade destes pacientes.

Com a descoberta da seqüência dos genes da *Mycobacterium tuberculosis*, o bacilo causador da tuberculose, foram feitos nockout de alguns genes. Um deles foi o aroK, gene codificador da enzima chiquimato quinase, e conforme estudos feitos por Parisch et al., esta enzima mostrou-se de suma importância, pois a sua ausência sozinha, é capaz de interromper a via metabólica do ácido chiquimico, inviabilizando o bacilo, pois esta via é responsável por sintetizar aminoácidos aromáticos e outros compostos essenciais para a sobrevivência do bacilo.

A aplicação da química computacional, aliada a bioinformática, tem oferecido um excelente suporte para o desenvolvimento de novos fármacos. Com o poder computacional e a tecnologia disponível atualmente pode ser realizado um direcionamento nos estudos, facilitado pela capacidade de predição virtual de interações e propriedades farmacológicas.

A realização deste trabalho foi de suma importância para a identificação de novas moléculas com potencial de se tornarem fármacos contra a Tuberculose. Além disso, identificamos resíduos do sítio ativo da Chiquimato Quinase que são essenciais para a interação da proteína com estes tipos de ligantes. Esperamos em trabalhos futuros realizar testes *in vitro* com estes ligantes afim de confirmar nossas predições *in silico*.

REFERÊNCIAS BIBLIOGRÁFICAS

ALIBHAI, M.F. & STALLINGS, W.C. (2001) *Proc. Natl. Acad. Sci. U.S.A.,* 98, 2944-2946.

BASSO, L.A., ZHENG, R., MUSSER, J.M., JACOBS, W.R. Jr. & BLANCHARD, J.S. (1998) *J. Infect. Dis*., 178, 769-775.

BASSO, L. A., PEREIRA DA SILVA, L. H., FETT-NETO, A. G., DE AZEVEDO JR., W. F., MOREIRA, I. S., PALMA, M. S., CALIXTO, J. B., ASTOLFI FILHO, S., dos SANTOS, R. R., SOARES, M. B. P., SANTOS, D. S. (2005) *Mem. Inst. Oswaldo Cruz,* 100(6): 475-506.

BENTLEY, R. (1990) *Crit. Rev. Biochem. Mol. Biol.,* 25, 307-384.

BRENK, R.; NAERUM,L.; GRAEDLER, U.; GERBER,H.; GARCIA,G.A. (2003).*Virtual screening for submicromolar leads of tRNA-guanine transglycosylase based on a new unexpected binding mode detected by crystal structure analysis*. Journal of Medicinal Chemistry, v. 46, p 1133 – 1143,

COLE, S.T., BROSCH, R., PARKHILl, J. *et al.*, (1998) *Nature,* 393, 537-544.

CARLSON,H.; MASUKAWA,K.M.; McCAMMON,J. A. (1999). *Method for including the dynamic fluctuations of a protein in a computer-aided drug desing*. Journal of Physical Chemistry A, v. 103, p. 10213-10219,

CHEN, G.S.; CHANG,C.S.; KAN,W.M.; CHANG, C.L.; WANG, K.C.; CHERN, J.N. (2001). *Novel lead generation through hypothetical pharmacophore three-dimensional database searching: discovery of isoflavonoids as nonsteroidal inhibitors of rat 5□-reductase*. Journal of Physical Chemistry, v.44, p. 3759 – 3763,

DE AZEVEDO, W.F. Jr., CANDURi, F., SIMÕES DE OLIVEIRA, J., BASSO, L.A., Palma, M.S., Pereira, J.H. & Santos, D.S. (2002a) *Biochem. Biophys. Res. Commun.*, 295, 142-148.

DIAS, M.V.B., Ely, F., CANDURI, F., PEREIRA, J.H., FRAZZON, J., BASSO, L.A., Palma, M.S., DE AZEVEDO, W.F., & Santos, D.S. (2004). *Acta Crystallogr. Sect. D.-Biol. Crystallogr.* 60, 2003-2005.

DIAS, M.V., BORGES, J.C., ELY, F., PEREIRA, J.H., CANDURI F., RAMOS CH, FRAZZON,J., PALMA M S, BASSO LA, SANTOS DS, DE AZEVEDO WF Jr. (2006) *J. Struc. Biol.* 154(2), 130-143.

DIAS MV, VASCONCELOS IB, PRADO AM, FADEL V, BASSO LA, DE AZEVEDO WF Jr, SANTOS DS. (2007) *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 63(Pt 1):1-6.

ERINS, S.;ROSE, J. (2002). *In silico ADME/Tox: the atate of the art*. Journal of Molecular Graphics and Modelling, v. 20, p. 305 – 309.

GU, Y., RESHETNIKOVA,L., Li, Y., WU, Y,. YAN, H., SINGH, S. & JI, S. (2002). *J Mol. Biol*. 319, 779-789.

HERRMANN, K.M. & WEAVER, L.M. (1999). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50, 473-503.

KRELL, T., COGGINS, J. R. & LAPTHORN, A. J. (1998) *J. Mol. Biol*. 278, 983-997.

KRELL, T., MACLEAN,J. BOAM,D.J., COOPER, A., RESMINI,M., BROCKLEHURST,K., KELLY, S.M., PRICE, N.C., LAPTHORN, A. J. & COGGINS, J. (2001). *Protein Sci*.10,1137-1149.

LIPINSKI, C. A.; LOMBARDO, F.; DOMINY, B. W.; FEENEY, P.J. (1997) *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings*. Advanced Drug Delivery Reviews, v. 23, p. 3-25.

MATHEWS, C.K. & VAN HOLDE, K.E. (1990) *Biochemistry*, Ed. Benjamin Cummings, 1130p., California, USA.

MARSHALL,G.R. (2004) *Introduction to chemoinformatics in drug discovery – A personal view*. In: OPREA, T.I. Chemoinformatics in drug discovery. Weinheim: Wiley – VHC. p. 1-22.

OLIVEIRA, J.S., PINTO, C.A., BASSO, L.A. & SANTOS, D.S. (2001) *Protein Expr. Purif.,* 22, 430-435.

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) *J. Mol. Biol.* 359(3), 646-666.

PARK, H., HILSENBECK, J.L., KIM, H.J., SHUTTLEWORTH, W.A., *et al.* (2004) *Mol. Microbiol.,* 51, 963-971.

PARISH, T.; STOKER, N. G. (2002) *Microbiology*, 148, 3069.

PEREIRA, J.H., CANDURI, F., DE OLIVEIRA, J.S., DA SILVEIRA, N.J., BASSO, L.A., PALMA, M.S., DE AZEVEDO, W.F. Jr, & SANTOS, D.S. (2003) *Biochem. Biophys. Res. Commun.,* 312, 608-614.

PEREIRA, J.H., OLIVEIRA, J.S., CANDURI, F., DIAS, M.V.B., PALMA, M.S., BASSO, L.A., SANTOS, D.S., & DE AZEVEDO, W.F. (2004) *Acta Crystallogr. Sect. D.-Biol. Crystallogr.* 60, 2310-2319.

RATLEDGE, C. (1982) Nutrition, growtl and metabolism, in: C. Ratlefge, J.L. Stanford (eds), *The Biology of the Mycobacteria*, vol. 1. Academic Press, London, pp. 185-271.

REVISTA ÉPOCA, v.380, Ago. 2005.

RIZZI, C., FRAZZON, F., Ely, F., WEBER, P.G., FONSECA, I.O., GALLAS, M., Oliveira, J.S., MENDES, M.A., SOUZA, B.M., PALMA, M.S., SANTOS, D.S. & BASSO, L.A. (2005). *Protein Expr. Purif.,* 40(1), 23-30.

ROBERTS, F., ROBERTS, C.W., JOHNSON, J.J., KYLE, D.E., *et al.* (1998) *Nature,* 393, 801-805.

SCHÖNBRUNN, E., ESCHENBURG, S., SHUTTLEWORTH, W.A., SCHLOSS, J.V., *et al.* (2001) *Proc. Natl. Acad. Sci. U.S.A.,* 98, 1376-1380.

SILVA, V. B. (2008) *Estudos de modelagem molecular e relação estrutura atividade da oncoproteína hnRNP K e ligantes*. USP- Ribeirão Preto.

THOMSEN, R.; CHRISTENSEN, M.H. *J. Med. Chem*. (2006) 11: 315-3321.

UCHOA, H.B., JORGE, G.E., DA SILVEIRA, N.J., CAMERA, J.C., CANDURI, F. & DE AZEVEDO, W.F. (2004) *Biochem. Biophys. Res. Commun.* 325, 1481-1486.

VONRHEIN, C., SCHLAUDERER, G. J. & SCHULTZ, G. F. (1995). *Struture*, 3, 483-490.

WORLD HEALTH ORGANIZATION (2009) WHO Report. Geneva. Switzerland. WHO/CDS/TB/2009. http://www.who.int/topics/tuberculosis/en/. Acessado em outubro de 2010.