



PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM MEDICINA E CIÊNCIAS DA SAÚDE

JULIANE DORS CORACINI

**SIMULAÇÃO COMPUTACIONAL DO SISTEMA PROTEÍNA-LIGANTE. ESTUDO
DA CHIQUIMATO QUINASE DE *MYCOBACTERIUM TUBERCULOSIS***

Porto Alegre

2012

JULIANE DORS CORACINI

**SIMULAÇÃO COMPUTACIONAL DO SISTEMA PROTEÍNA-LIGANTE. ESTUDO
DA CHIQUIMATO QUINASE DE *MYCOBACTERIUM TUBERCULOSIS***

Dissertação apresentada como requisito para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Medicina e Ciências da Saúde da Pontifícia Universidade Católica do Rio Grande do Sul.

Orientador: Prof. Dr. Walter Filgueira de Azevedo Jr.

Porto Alegre

2012

DADOS DE CATALOGAÇÃO

C787s Coracini, Juliane Dors

Simulação computacional do sistema proteína-ligante. Estudo da chiquimato quinase de *mycobacterium tuberculosis* / Juliane Dors Coracini. Porto Alegre: PUCRS, 2013.

91 f.: il. tab. Inclui artigo científico encaminhado para publicação no periódico *Current Drug Targets*.

Orientador: Prof. Dr. Walter Filgueira de Azevedo Júnior.

Dissertação (Mestrado) – Pontifícia Universidade Católica do Rio Grande do Sul. Faculdade de Medicina. Pós-Graduação em Medicina e Ciências da Saúde.

1. TUBERCULOSE. 2. CHIQUIMATO QUINASE. 3. SÍTIO DE LIGAÇÃO DO ATP. 4. *DOCKING* MOLECULAR. 5. *VIRTUAL SCREENING*. 6. SIMULAÇÃO POR COMPUTADOR. 7. MODELOS TEÓRICOS. I. Azevedo Júnior, Walter Filgueira de. II. Título.

CDD 574.88
CDU 616-002.5(043.3)
NLM WF 415

Isabel Merlo Crespo
Bibliotecária CRB 10/1201

JULIANE DORS CORACINI

**SIMULAÇÃO COMPUTACIONAL DO SISTEMA PROTEÍNA-LIGANTE. ESTUDO
DA CHIQUIMATO QUINASE DE *MYCOBACTERIUM TUBERCULOSIS***

Dissertação apresentada como requisito para obtenção do grau de Mestre pelo Programa de Pós-graduação em Medicina e Ciências da Saúde da Pontifícia Universidade Católica do Rio Grande do Sul.

Aprovada em: ____ de _____ de _____.

BANCA EXAMINADORA:

Dr. Luiz Augusto Basso (PUCRS)

Dr. Maurício Reis Bogo (PUCRS)

Dr. Hermes Luís Neubauer de Amorim (ULBRA)

Porto Alegre

2012

Dedico esta dissertação aos meus pais,
que tanto apoiaram e incentivaram
o meu crescimento profissional.

Agradecimentos

Agradeço ao meu orientador, o professor Dr. Walter Filgueira de Azevedo Jr pela compreensão, apoio e sabedoria. Principalmente por permitir que eu fizesse parte do seu grupo de pesquisa, o Laboratório de Bioquímica Estrutural (LaBioQuest), por aceitar me orientar no mestrado, sempre incentivando a busca do conhecimento científico.

A minha família, especialmente aos meus pais e irmãos, pelo apoio incondicional e coragem para eu seguir em frente.

Eu gostaria de agradecer a todos os meus colegas do laboratório, pela ajuda, amizade e companheirismo, principalmente aos colegas Bianca Villavicencio, Mariana Morrone Xavier e Dr. Rafael Caceres. Queria agradecer, em especial, à colega e amiga, Me. Fernanda Pretto Moraes, que sempre esteve por perto incentivando e descontraindo nos momentos difíceis.

À minha grande amiga, Vincenza Baiotto Soares, que mesmo fora do meio acadêmico, foi incansável em me ajudar, sempre pronta para tudo, sempre ao meu lado.

Também queria agradecer ao Ricardo Pereira Winckler pelo carinho em me ouvir e ajuda quando precisei.

Aos professores do curso pelos ensinamentos e por contribuírem com a minha formação.

A CAPES pela bolsa concedida.

Obrigada a todos!

*“Conheça todas as teorias, domine
todas as técnicas, mas ao tocar uma
alma humana, seja apenas outra alma humana.”*

Carl Jung

Resumo

A Tuberculose continua sendo a causa mais comum de morte em decorrência de um agente infeccioso. Entre os alvos identificados no genoma do *Mycobacterium tuberculosis*, enzimas da via do chiquimato merecem atenção especial. A chiquimato quinase é a quinta enzima da via do chiquimato, e tem sido identificada em fungos, organismos do filo apicomplexa, plantas e procariontes. Esta via metabólica é composta por sete passos, que catalisam sequencialmente a conversão de eritrose-4-fosfato e fosfoenolpiruvato em corismato; e é responsável pela biossíntese de aminoácidos aromáticos. Chiquimato quinase parece ser essencial para a sobrevivência do *Mycobacterium tuberculosis*, uma vez que é ausente no homem, esta enzima é considerada como um alvo para o desenvolvimento de quimioterápicos e medicamentos contra a tuberculose. O objetivo é identificar possíveis inibidores, focando as simulações de *docking* molecular no sítio de ligação do ATP da enzima. O programa usado nas simulações foi o *Molegro Virtual Docker* e a interação proteína-ligante foi testada em 12 estruturas cristalográficas e logo após, escolhido um protocolo de *docking* a partir de valores de RMSD abaixo de 2Å. O método foi validado usando o melhor protocolo de *re-docking* no *Virtual Screening* através do Fator de Enriquecimento que obteve resultado de 24,57%, que é considerado adequado para as simulações de *docking* molecular focados em quinases. O presente protocolo de *docking* foi aplicado em um banco de dados com mais de 80.000 moléculas. A análise dos resultados identificaram 5 potenciais inibidores da chiquimato quinase. Na análise das interações intermoleculares entre a enzima e os ligantes foram identificadas características estruturais responsáveis pela afinidade da ligação pelo ligante. Este é o primeiro estudo de *docking* molecular focado no bolsão de ligação do ATP da chiquimato quinase.

Palavras Chaves: Tuberculose, Chiquimato Quinase, sítio de ligação do ATP, *Docking* molecular, *Virtual Screening*.

Abstract

Tuberculosis remains the most common cause of death due to an infectious agent. Among targets identified in *Mycobacterium tuberculosis* genome, enzymes of the shikimate pathway deserve special attention. Shikimate kinase is the fifth enzyme of the shikimate pathway, which has been identified in fungi, apicomplexans, plants and prokaryotes. This metabolic route is composed of seven steps, which converts erythrose-4-phosphate and phosphoenol pyruvate to chorismic acid and is responsible for the biosynthesis of aromatic amino acids. Shikimate kinase has been shown to be essential to the survival of *Mycobacterium tuberculosis*, and since it is absent in human, this enzyme is considered to be a target for chemotherapeutic for development of antitubercular drugs. The aim here is to identify possible inhibitors, focusing on simulations of molecular docking in the ATP-binding site of the enzyme. The program used in the simulations was the Molegro Virtual Docker and protein-ligand interactions were tested in 12 crystallographic structures and then, it was chosen a protocol which generated docking RMSD values below 2 Å. Application of this docking protocol to a decoy dataset generated a enrichment factor of 24.57, which is considered adequate for molecular docking simulations focused on kinases. The present docking protocol was then applied to a small-molecule database with over 80,000 entries. Analysis of the results identified 5 potential shikimate kinase inhibitors. Examination of the intermolecular interaction between enzyme and the ligands identified the main structural features responsible for ligand-binding affinity. This is the first molecular docking study focused on the ATP-binding pocket of shikimate kinase.

Keywords: Tuberculosis, Shikimate Kinase, ATP binding site, *docking* molecular, *virtual screening*.

Lista de abreviaturas e Siglas

ADMET, Absorção, distribuição, metabolismo, excreção e toxicidade

ADP, Adenosina Difosfato

AE, Algoritmo Evolucionário

ATP, Adenosina Trifosfato

BCG, Bacille Calmette-Guérin ou vacina contra tuberculose

EF, Fator de Enriquecimento

HIV, Human immunodeficiency virus ou Vírus da Imunodeficiência Humana

MolDock, Molegro Virtual Docker

MtSK, Mycobacterium Tuberculosis Shikimate Kinase

MVD, Molegro Virtual Docker

PDB, Protein Data Bank ou banco de dados de proteínas

RMSD, Desvio Médio Quadrático

RO5, Lipinski's rule of five

SK, Shikimate Kinase ou Chiquimate Quinase

TB, Tuberculose

VS, Virtual Screening

WHO, World Health Organization ou Organização Mundial de Saúde

Lista de Ilustrações

Figura 1. Via do ácido chiquímico.

Figura 2. Reação enzimática catalisada pela enzima chiquimato quinase.

Figura 3. Chiquimato quinase complexada com ADP e chiquimato.

Figura 4. Sítio de ligação do ATP e do chiquimato da shikimate kinase.

Figura 5. Processo de Virtual Screening e as diferentes fases de seleção dos melhores resultados.

Figura 6. Esfera de docking (verde) usada nas simulações de docking molecular.

Figura 7. Estruturas moleculares dos 5 possíveis inibidores.

Figura 8. Ligplot do bolsão de ligação do ATP com a enzima chiquimato quinase.

Figura 9. Interações intermoleculares da molécula ZINC02135897 (LIG 02 C) com o bolsão de ligação do ATP da chiquimato quinase.

Figura 10. Interações intermoleculares da molécula ZINC02127309 (LIG 07 C) com o bolsão de ligação do ATP da chiquimato quinase.

Lista de Tabelas

Tabela 1. Tabela usada no processo de re-docking para as diferentes estruturas cristalográficas.

Tabela 2. Propriedades físico-químicas dos ligantes através da análise pelo filtro FAF-Drugs.

Tabela 3. Interações intermoleculares dos 5 melhores ligantes selecionados. A presença de um X indica que a interação ocorre. HB são as ligações de hidrogênio e VDW são os contatos de Van der Waals.

SUMÁRIO

1. INTRODUÇÃO	15
1.1 Tuberculose	15
1.2 Via do Ácido Chiquímico.....	16
1.3 Chiquimato Quinase	18
1.4 Sítio de ligação do ATP da Chiquimato Quinase	19
1.5 Desenvolvimento de Fármacos a partir de Computação Bio-Inspirada.....	21
2. JUSTIFICATIVA	24
3. OBJETIVOS	26
3.1 Objetivo geral	26
3.2 Objetivos específicos	26
4. MATERIAIS E MÉTODOS	28
4.1 Simulações de <i>Docking</i> molecular	28
4.2 <i>Re-docking</i> e <i>Cross-Docking</i>	29
4.3 <i>Virtual Screening</i>	32
4.3.1 Fator de Enriquecimento	34
5. RESULTADOS E DISCUSSÃO	36
5.1 <i>Re-docking</i> e <i>cross-docking</i>	36
5.2 <i>Virtual Screening</i>	37
5.2.1 Fator de Enriquecimento.....	38
5.2.2 Seleção e Análise dos melhores ligantes.....	38
5.3 Interações moleculares.....	40
6. CONCLUSÃO	46
REFERÊNCIAS	48

ANEXO	53
ANEXO- Shikimate kinase, a protein target for design of antitubercular drugs.....	54

Capítulo 1

Introdução

- 1.1 Tuberculose
 - 1.2 Via do Ácido Chiquímico
 - 1.3 Chiquimato Quinase
 - 1.4 Sítio de Ligação do ATP da Chiquimato Quinase
 - 1.5 Desenvolvimento de Fármacos a partir de Computação Bio-Inspirada
-

1. INTRODUÇÃO

1.1 Tuberculose

A tuberculose ainda é uma das doenças infecciosas curáveis mais temidas, graças a sua alta taxa de incidência, prevalência e mortalidade (**Zahrt, 2003**). Estima-se que um terço da população mundial esteja infectado pelo bacilo de Koch da tuberculose. Destacando a gravidade do problema dessa doença a nível mundial, a pesquisa para o desenvolvimento de novos fármacos para o seu tratamento é de extrema relevância, uma vez que desde a década de 80 não é desenvolvido nenhum medicamento novo contra essa doença. De acordo com a Organização Mundial de Saúde, há aproximadamente 8,8 milhões de casos incidentes, ocorrendo 1,4 milhões de mortes incluindo 0,35 milhões de mortes entre pessoas HIV positivas. Foram notificados recentemente 5,7 milhões de novos casos de TB e/ou recaída (**WHO Report 2010**). Apesar dos indicadores animadores em relação à tendência de queda da incidência e da mortalidade por tuberculose, seus números absolutos ainda causam indignação e trazem um desafio grandioso. No Brasil, são mais de 70 mil novos casos e o número de óbitos ultrapassa a cifra de 4,5 mil a cada ano (**Ministério da Saúde, 2011**).

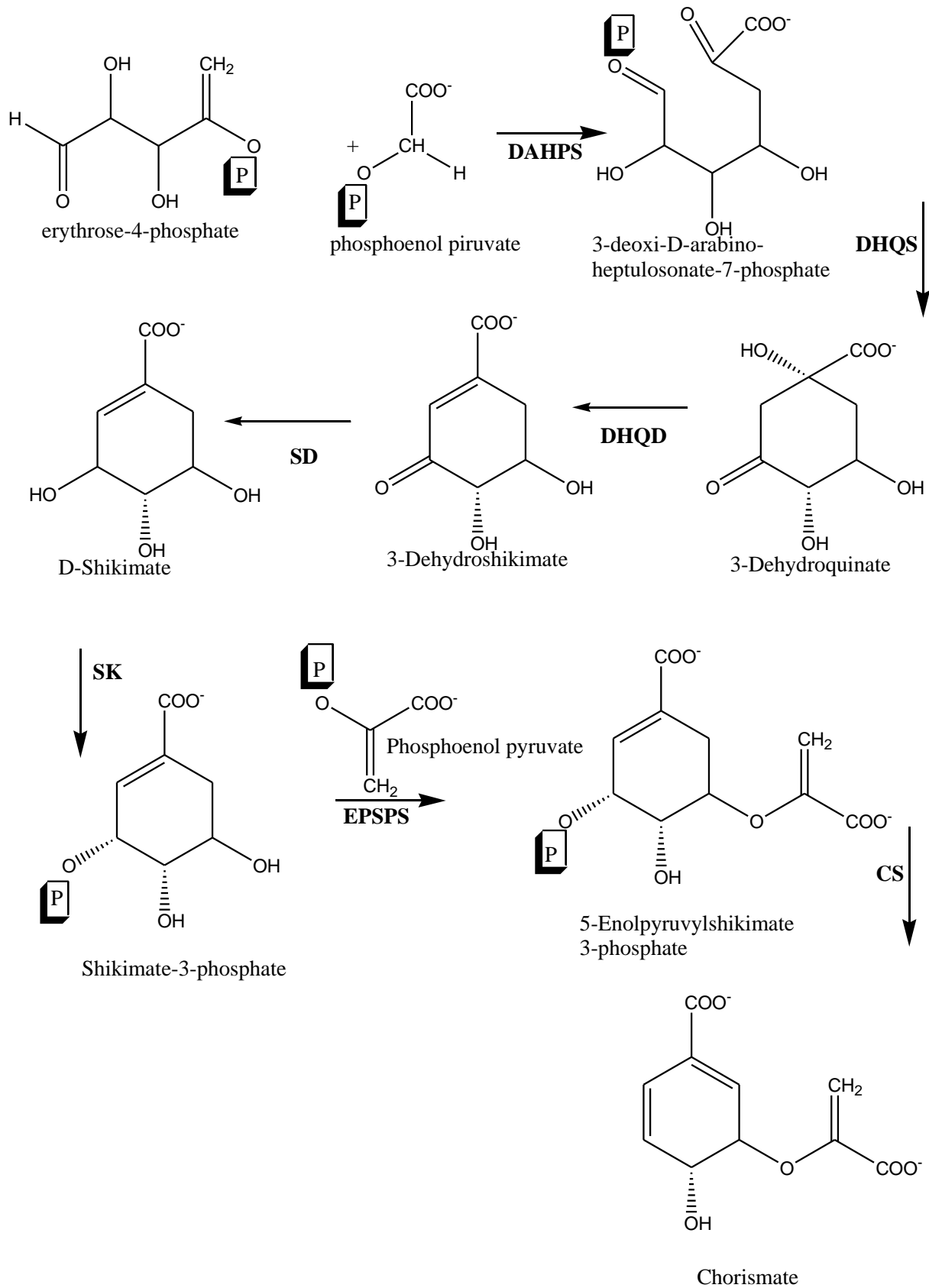
A infecção pelo *Mycobacterium tuberculosis* ocorre preferencialmente pelas vias aéreas, e dos indivíduos expostos 70% não desenvolvem a doença, mas 30% tornam-se infectados; destes, 40% desenvolvem a forma ativa da tuberculose e 60% estão em estágio latente da doença. Essa forma pode continuar latente, como também pode reativar a tuberculose tardiamente em 2-23% dos pacientes imunocompetentes. Nos pacientes com a síndrome da imunodeficiência adquirida, a reativação ocorre em aproximadamente 5-10% por ano (**Parish & Stoker, 2002**). Os indivíduos com a infecção latente são assintomáticos e mais difíceis de tratar, dificultando o controle da propagação da doença. O surgimento de cepas resistentes às drogas gerou formas diferentes de tuberculose que são chamadas MDR-TB (*MultiDrug Resistant*) que são resistentes a isoniazida e rifampicina, XDR-TB (*extensively-drug resistant*) que se mostram resistentes aos quimioterápicos tanto de primeira quanto de segunda linha e, mais recentemente TDR-TB (*totally-drug resistant*) (**Hargreaves, 2008**). Os mecanismos de resistências 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**). Estes levam ao ressurgimento da doença. A situação brasileira é preocupante, já que exhibe altas taxas dos pacientes, previamente tratados, com multi-resistência adquirida.

Pode-se dizer, então, que o aumento do número de casos de tuberculose atualmente está relacionado a dois fatores. O primeiro é a ocorrência da tuberculose em pacientes co-infectados com o vírus HIV. Esses pacientes, cujo sistema imunológico enfraquecido não pode controlar o crescimento do bacilo, apresentam risco bem maior de desenvolver a doença. Outro fator é a emergência de cepas resistentes aos antimicrobianos, utilizados no tratamento, devido às terapias inadequadas e o uso indiscriminado destes antibióticos (**Basso et al., 2005**). Em vista disso, há uma necessidade urgente em descobrir e desenvolver novos e melhores fármacos para o tratamento da tuberculose.

1.2 Via do Ácido Chiquímico

Um alvo de estudo para o descobrimento de novas drogas contra a tuberculose são as enzimas da via do ácido chiquímico (**Figura 1**), pois são essenciais para plantas, organismos do filo apicomplexa, parasitas e bactérias (**Roberts et al., 2002**), incluindo o *Mycobacterium tuberculosis* (**Parish & Stoker, 2002; Hartmann et al., 2006**), porém é ausente no homem, o que a torna um alvo bastante atrativo (**Hartmann et al., 2006; Bentley, 1990**).

Figura 1 - Via do ácido chiquímico



As estruturas foram desenhadas através do Programa ChemDraw Ultra 7.0.
 Fonte: Coracini (2012).

A via sintética do ácido chiquímico identificada no *Mycobacterium tuberculosis* (Cole, 1998) é constituída por sete passos enzimáticos, sendo cada um catalisado por uma enzima diversa e cada enzima codificada por um único gene (Ducati, 2007). Algumas dessas enzimas, já estão disponíveis como cristais e suas estruturas em alta resolução determinadas, assim como estudos de modelagem estrutural na área de bioinformática já foram realizados (De Azevedo *et al.*, 2002; Dias *et al.*, 2007; Oliveira *et al.*, 2006). O intuito é dar continuidade a esses estudos, a fim de identificarmos novos possíveis inibidores através de docking molecular e estudar as bases estruturais dos inibidores já identificados.

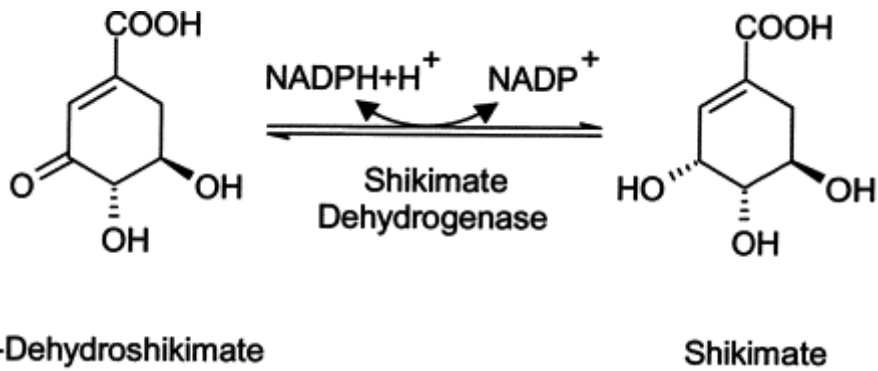
1.3 Chiquimato Quinase

O objeto do presente estudo é a enzima chiquimato quinase (SK) que catalisa a quinta reação da via do ácido chiquímico, convertendo adenosina trifosfato (ATP) e chiquimato em adenosina difosfato e chiquimato-3-fosfato, utilizando magnésio como cofator, realizando uma transferência de um fosfato doador (ATP) para o carbono do ácido chiquímico (Yan, 1997), (figura 2). O produto desta reação é subsequentemente fosforilado, resultando em precursores dos ácidos nucleicos.

A chiquimato quinase pertence à família estrutural das nucleosídeo monofosfato quinases (NMP), é uma proteína da classe α/β , consistindo de folhas β circundadas por hélices- α , sendo bastante semelhante à adenilato quinase (Pereira *et al.*, 2004; Krell *et al.*, 1998, 2001). As NMP quinases sofrem grandes mudanças conformacionais durante a catálise (Vonrhein *et al.*, 1995).

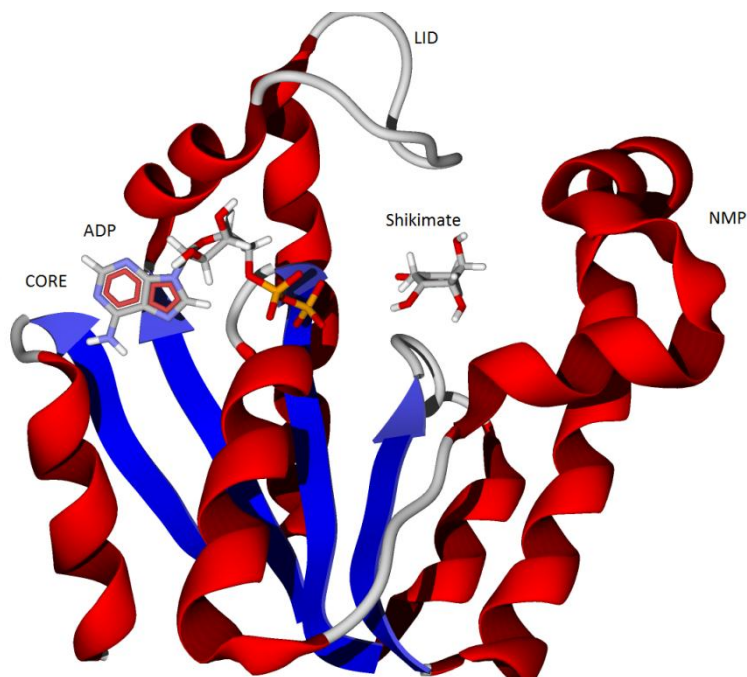
A família desta enzima possui três domínios conservados: CORE, contendo cinco folhas- β paralelas e o P-loop que formam o sítio de ligação dos nucleotídeos; o domínio LID, que se move sobre o sítio ativo e possui resíduos essenciais para a ligação do ATP; e o domínio de ligação NMP, que é responsável por reconhecer e se ligar ao nucleosídeo monofosfato (Yan H & Tsai MD, 1999; Gu Y *et al.*, 2002) (figura 3).

Figura 2 - Reação enzimática catalisada pela enzima chiquimato quinase.



Fonte: Adaptado de Padyana & Burley 2003.

Figura 3 - Chiquimato quinase complexada com ADP e chiquimato (código de acesso do PDB: 2DFN)



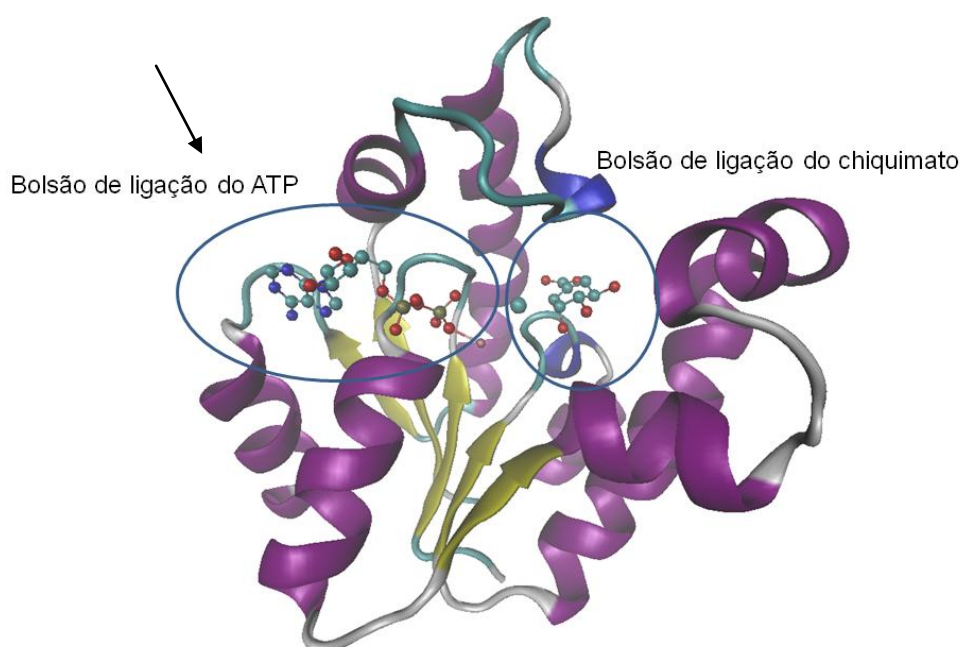
Fonte: Coracini (2012)

1.4 Sítio de Ligação do ATP da Chiquimato Quinase

A enzima chiquimato quinase foi validada como alvo molecular para desenho de drogas contra tuberculose (Parish & Stoker, 2002). Nosso laboratório recentemente realizou estudos focados nesta enzima no bolsão de ligação do chiquimato (Vianna & De Azevedo, 2011), assim como outros diversos estudos conduzidos para o mesmo bolsão de ligação (Segura &

Rodriguez, 2008; Saidemberg *et al.*, 2011; Hsu *et al.*, 2012; De Azevedo, 2011; Kumar *et al.*, 2010; Arora *et al.*, 2010; Mulabagal & Calderón, 2010; De Azevedo, 2010; Habig *et al.*, 2009; Pauli *et al.*, 2009; Han *et al.*, 2007; Silveira *et al.*, 2005; Filgueira & Canduri, 2002). No presente trabalho iremos ampliar tais estudos, focando as simulações de *docking* molecular no sítio de ligação de ATP da chiquimato quinase de *M. tuberculosis*. A chiquimato quinase apresenta dois bolsões de ligação, identificados nas estruturas cristalográficas (Pereira *et al.*, 2004). Um onde se liga o substrato, chamado bolsão de ligação do chiquimato. O outro é o bolsão de ligação do ATP. Estudos de simulação de *docking* molecular foram focados no bolsão de ligação do ácido chiquímico (Vianna & De Azevedo, 2011). A presente proposta visa o estudo do bolsão de ligação do ATP, com o potencial de beneficiar-se de estudos prévios de inibidores de quinase (Canduri & De Azevedo, 2005; De Azevedo *et al.*, 1997; 2002).

Figura 4 - Sítio de ligação do ATP e do chiquimato da chiquimato quinase.



Fonte: Coracini (2012)

Neste estudo, encontramos possíveis inibidores contra um alvo específico, focando as simulações de *docking* molecular no sítio de ligação do ATP da chiquimato quinase de *Mycobacterium tuberculosis* com a estrutura MtSK (*Mycobacterium tuberculosis* Shikimate Kinase) (Dias *et al.*, 2007), com código de acesso do PDB 2DFN, ampliando estudos recentes (Vianna & De Azevedo, 2011). Selecionamos os melhores inibidores com o processo de *Virtual screening*. Nosso protocolo de *docking* foi validado contra um conjunto de 12 estruturas cristalográficas disponíveis para os complexos MtSK. Foi utilizado um banco de dados de 89.425 estruturas e os resultados descritos foram obtidos em MOLDOCK scores, interações intermoleculares e discussão da importância dos resíduos de aminoácidos presentes no sítio ativo dos possíveis melhores ligantes.

1.5 Desenvolvimento de Fármacos a partir de Computação Bio-Inspirada

O desenvolvimento racional de fármacos baseia-se no conhecimento detalhado da estrutura tridimensional da proteína alvo, no conhecimento da sua interação com o ligante e na racionalização de como o alvo poderá interagir com um fármaco em potencial. Dessa forma, o processo de interação proteína-ligante tem de ser elementar para o desenho de fármacos e otimização de um fármaco líder, aliando características de alta seletividade, especificidade e propriedades farmacocinéticas adequadas (Thomsen & Christensen, 2006).

A descoberta de drogas é um processo de tentativa e erro em experimentos *in vitro*, utilizando um alvo conhecido e compostos candidatos na interação com esse alvo, avaliando sua atividade. Esta metodologia apresenta desvantagens como o alto custo financeiro, trabalho e tempo para testar os diversos compostos existentes, portanto há uma necessidade em desenvolver métodos que reduzam tais custos. A utilização de abordagens *in silico*, que são um meio de simular modelos biológicos em computador, vem sendo utilizada como uma forma de contribuir para o processo de descoberta de novas drogas de maneira mais rápida e barata (Lyne, 2002).

A estrutura e o mecanismo de ação das enzimas específicas no planejamento racional é a maneira mais eficiente de participar no desenvolvimento de fármacos, sendo capaz de contribuir em diversos estágios, como descoberta de protótipos, otimização, até a elaboração de compostos candidatos a testes clínicos. Baseando-se no bloqueio ou estimulação da atividade de macromoléculas, como proteínas ou ácidos nucleicos, associados a diferentes processos patológicos. A informação estrutural do ligante permite a descoberta e síntese de compostos com complementaridade geométrica, hidrofóbica e eletrostática ao seu sítio de ligação, sendo capazes de se tornar potenciais inibidores ou futuros fármacos (Silva, 2008).

Tem-se o conhecimento de que 78% dos fármacos atuais possuem como alvo receptor esse tipo de biomacromolécula biológica (Marshall, 2004).

A docagem (docking) molecular é um dos métodos mais empregados no desenvolvimento racional de fármacos. Através do *docking* molecular (Thomsen & Christensen, 2006; Heberlé & De Azevedo, 2011; De Azevedo, 2010) e de estruturas cristalográficas de uma proteína de interesse, pode-se procurar em uma base de dados possíveis ligantes que encaixem em seu sítio ativo e prever conformações do complexo proteína-ligante. Essa simulação computacional pode gerar diversas posições dos ligantes, e o objetivo é tentar achar as que tenham a menor distância relativa ao modelo experimental. Após, com o *virtual screening* (VS), procura-se possíveis ligantes através de certos critérios de seleção, que posteriormente podem ser testados *in vitro* ou *in vivo*, já com um maior direcionamento gerado a baixo custo e menor tempo (Peitsch, 2004).

A bioinformática, juntamente com a química computacional, tem mostrado um excelente direcionamento ao planejamento racional de fármacos através do sucesso envolvendo importantes fármacos como: antiretrovirais, um exemplo, inibidores do HIV protease (Kitchen *et al.*, 2004), e outros inúmeros, como inibidores não-esteroidais da 5 α -redutase (Brenk *et al.*, 2003; Chen *et al.*, 2001).

Capítulo 2

Justificativa

2. JUSTIFICATIVA

Descoberta há mais de um século, a tuberculose continua a preocupar. É uma das doenças infecciosas documentadas desde mais longa data e que continua a afligir a população, apesar de ser uma doença socialmente determinada. Em diversos países, houve a ideia de que, com a introdução da vacina BCG e a descoberta de antibióticos como a rifampicina (1965), nos dias de hoje, a doença estaria praticamente controlada e inexistente. No entanto, com o advento do vírus HIV e da AIDS, essa perspectiva mudou drasticamente. Conforme a Organização Mundial de Saúde (**WHO Report 2010**), atualmente com quase 9 milhões de novas contaminações, a tuberculose continua sendo um grave problema, que pode se agravar caso os países restrinjam verbas para o seu combate.

O surgimento de cepas resistentes aos antimicrobianos disponíveis no mercado, graças às terapias inadequadas e uso indiscriminado desses antibióticos e a ocorrência da tuberculose em pacientes infectados com o HIV, que tem seu sistema imune enfraquecido e um risco maior de desenvolver a doença, cria-se a necessidade da descoberta de novos agentes terapêuticos para o tratamento da tuberculose (**Basso et al., 2005**).

Entre os alvos identificados no genoma do *Mycobacterium tuberculosis*, a chiquimato quinase, quinta enzima da via do chiquimato, é um alvo bastante atrativo para o descobrimento de novos fármacos para o tratamento da tuberculose, pois ela está presente na micobactéria, porém é ausente no homem. O trabalho proposto tem o intuito de identificar novos potenciais inibidores da enzima chiquimato quinase utilizando ferramentas de bioinformática e modelagem molecular, visando à aplicação de abordagens computacionais para a previsão da interação proteína-ligante, bem como servir de ajuda para o desenvolvimento de novos fármacos. A proposta irá contribuir para a geração de conhecimento, com dados estruturais sobre proteínas alvos para o desenho de drogas e interação com diferentes ligantes. Estes dados serão disponibilizados na forma de banco de dados estruturais, conforme outros já desenvolvidos pelo coordenador do projeto e irá contribuir para o desenvolvimento de produtos estratégicos para o Ministério da Saúde (Linha de ação 9 do Ministério da Ciência e Tecnologia).

Capítulo 3

Objetivos

3.1 Objetivo Geral

3.2 Objetivos Específicos

3. OBJETIVOS

3.1 Objetivo Geral

Identificar possíveis inibidores para a enzima chiquimato quinase de *Mycobacterium tuberculosis*, com estudos estruturais focados na interação proteína-ligante.

3.2 Objetivos Específicos

- 1) Simular computacionalmente a interação proteína-ligante por meio de algoritmos de *docking* molecular;
- 2) Identificar aspectos estruturais determinantes para a especificidade do ligante pela enzima chiquimato quinase, focando as simulações de *docking* molecular no sítio de ligação do ATP;
- 3) Propor novos ligantes que apresentem indicativos estruturais que mostrem aumento da especificidade deste pela enzima.

Capítulo 4

Materiais e Métodos

- 4.1 Simulações de *Docking* Molecular
 - 4.2 *Re-docking* e *Cross Docking*
 - 4.3 Virtual Screening
 - 4.3.1 Fator de Enriquecimento
-

4. MATERIAIS E MÉTODOS

4.1 Simulações de *Docking* Molecular

No processo de simulação de *docking* molecular, são geradas várias possibilidades de encaixe proteína-ligante. A predição do ligante mais adequado pode ser feita a partir da aplicação da função escore empírica, responsável por analisar a interação deste complexo (Shoichet, 2004). Essas abordagens computacionais permitem a triagem *in silico* de bibliotecas de compostos, avaliando afinidade e a especificidade a partir de propriedades estruturais e químicas, como geometria, distribuição de cargas, polaridade, potencial de interações hidrofóbicas e ligações de hidrogênio. Dessa forma, a triagem de bancos de possíveis ligantes tem a função de identificar compostos que se ligam mais fortemente a uma proteína alvo em relação ao seu substrato natural, assim a reação bioquímica que a proteína alvo catalisa pode ser alterada ou impedida (inibição) (De Azevedo *et al.*, 1997).

Simulação de *docking* molecular é uma metodologia onde a estrutura tridimensional de um complexo proteína-ligante pode ser simulada computacionalmente. Há diversas metodologias para *docking* molecular (De Azevedo, 2010; De Azevedo & Dias, 2008). Neste estudo, foi usado o método de evolução diferencial guiada (*Guided differential evolution*, GDE) implementado no programa ‘Molegro Virtual Docker’ (MVD). Este é um dos programas disponíveis comercialmente para simulações de *docking* baseado em algoritmos evolucionários, uma técnica de otimização iterativa baseada na evolução darwiniana. No entanto, neste algoritmo a ideia de evolução é simplificada, tendo pouca semelhança com a evolução darwiniana. Recentemente, sabe-se que o MVD é capaz de encontrar a posição correta de um ligante. Também exibe um melhor desempenho comparado com AUTODOCK, SURFLEX, FLEXX e GOLD (Thomsen & Christensen, 2006; Heberlé & De Azevedo, 2011; De Azevedo, 2010; Araújo *et al.*, 2011). As simulações foram realizadas em computadores Dell (Intel Processador Core 2 Duo, 1.86 GHz, 2GB).

O MVD traz a implementação de quatro algoritmos de busca para encontrar a posição e a orientação do ligante. São eles: MOLDOCK optimizer, (implementação do algoritmo de evolução diferencial), MOLDOCK *Simplex Evolution (SE)*, *Iterated Simplex* e *Iterated Simplex* com otimização de colônias de formigas (Thomsen & Christensen, 2006; De Azevedo, 2010).

As simulações de *docking* foram realizadas através do MOLDOCK (Thomsen & Christensen, 2006), na qual o melhor complexo binário (proteína-ligante) é a posição mais próxima da estrutura cristalográfica. Por essa razão, devemos estabelecer uma metodologia

que avalia a distância da solução gerada por computador (pose) a estrutura cristalográfica. Esta distância pode ser calculada usando o Desvio Médio Quadrático (RMSD), que é uma medida das diferenças entre os valores previstos pelo modelo e os valores efetivamente observados a partir do objeto a ser modelado ou estimado (complexo de proteína-ligante). O RMSD é calculado entre dois conjuntos de coordenadas atômicas, neste caso, um para a estrutura cristalográfica (x_{ctal} , y_{ctal} , z_{ctal} ; o objeto que está sendo modelado) e outra para coordenadas atômicas obtidas das simulações de encaixe (x_{pose} , y_{pose} , z_{pose} ; previsto pelo modelo). A somatória é sobre todos os átomos de N que estão sendo comparados, utilizando a seguinte equação:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{ctal,i} - x_{pose,i})^2 + (y_{ctal,i} - y_{pose,i})^2 + (z_{ctal,i} - z_{pose,i})^2}.$$

Nas simulações de docking, o esperado é que os melhores resultados gerem um valor de RMSD inferior a 2.0 Å comparados com as estruturas cristalográficas (**Friesner *et al.*, 2004**). A estrutura do MtSK complexada com shikimate e ADP foram usadas nas simulações de docking. As moléculas de água e a shikimate foram removidos do modelo, então o substrato ADP foi usado sozinho no complexo MtSK-ADP.

4.2 Re-docking e Cross-Docking

Este procedimento de obter a posição cristalográfica do ligante é frequentemente chamado de “re-docking”, que é fundamentalmente um método de validação que determina se o algoritmo de docking molecular é capaz de recuperar a posição cristalográfica utilizando simulação computacional. Também pode ser usado um procedimento chamado “*cross-docking*” para validar ainda mais um protocolo de *docking*. Considerando que várias estruturas cristalográficas estão disponíveis para a mesma proteína, *cross-docking* pode ser aplicado. Pode-se dizer que o *cross-docking* são vários *re-dockings*. Este procedimento envolve uma série de *dockings* de ligantes encontrados em uma variedade de cristais de uma proteína idêntica para uma conformação cristalográfica de uma única proteína rígida (**Friesner *et al.*, 2004**). Quando um alvo da proteína apresenta grandes mudanças conformacionais sobre ligação do ligante, uma diferença significativa é esperada entre as estruturas cristalográficas e a estrutura “docada”. Nas simulações de *cross-docking* foram

utilizadas 12 estruturas cristalográficas (códigos de acesso PDB: 2DFN, 1U8A, 1ZYU, 2G1K, 2IYQ, 2IYR, 2IYS, 2IYX, 2IYY, 2IYZ, 3BAF, 1WE2). Este procedimento de validação *re-docking* e *cross-docking* é o estágio inicial do protocolo de VS (phase 1 virtual screening) descrito nas próximas etapas.

Tabela 1 - Tabela usada no processo de re-docking para as diferentes estruturas cristalográficas.

Molecular docking protocols (MOLDOCK)

Protocol	Scoring functions	Search algorithm	RMSD (Å)	RMSD (Å)	RMSD (Å)	RMSD (Å)
Sorting criteria			MolDock Score	Rerank Score	HBond	RMSD
1	MolDock Score	MolDock Optimizer				
2	MolDock Score	MolDock (Simplex Evolution) SE				
3	MolDock Score	Iterated Simplex				
4	MolDock Score	Iterated Simplex (Ant Colony Optimization)				
5	MolDock Score [GRID]	MolDock Optimizer				
6	MolDock Score [GRID]	MolDock (Simplex Evolution) SE				
7	MolDock Score [GRID]	Iterated Simplex				
8	MolDock Score [GRID]	Iterated Simplex (Ant Colony Optimization)				
Protocol	Scoring functions	Search algorithm	RMSD (Å)	RMSD (Å)	RMSD (Å)	RMSD (Å)
Sorting criteria			Plants Score	MolDock Score	Rerank Score	RMSD
9	PLANTS Score	MolDock Optimizer				
10	PLANTS Score	MolDock (Simplex Evolution) SE				
11	PLANTS Score	Iterated Simplex				
12	PLANTS Score	Iterated Simplex (Ant Colony Optimization)				
13	PLANTS Score [GRID]	MolDock Optimizer				
14	PLANTS Score [GRID]	MolDock (Simplex Evolution) SE				
15	PLANTS Score [GRID]	Iterated Simplex				
16	PLANTS Score [GRID]	Iterated Simplex (Ant Colony Optimization)				

Identification

PDB: _____

Chain: _____

Ligand: _____

Binding site

Center (Å)

X: _____

Y: _____

Z: _____

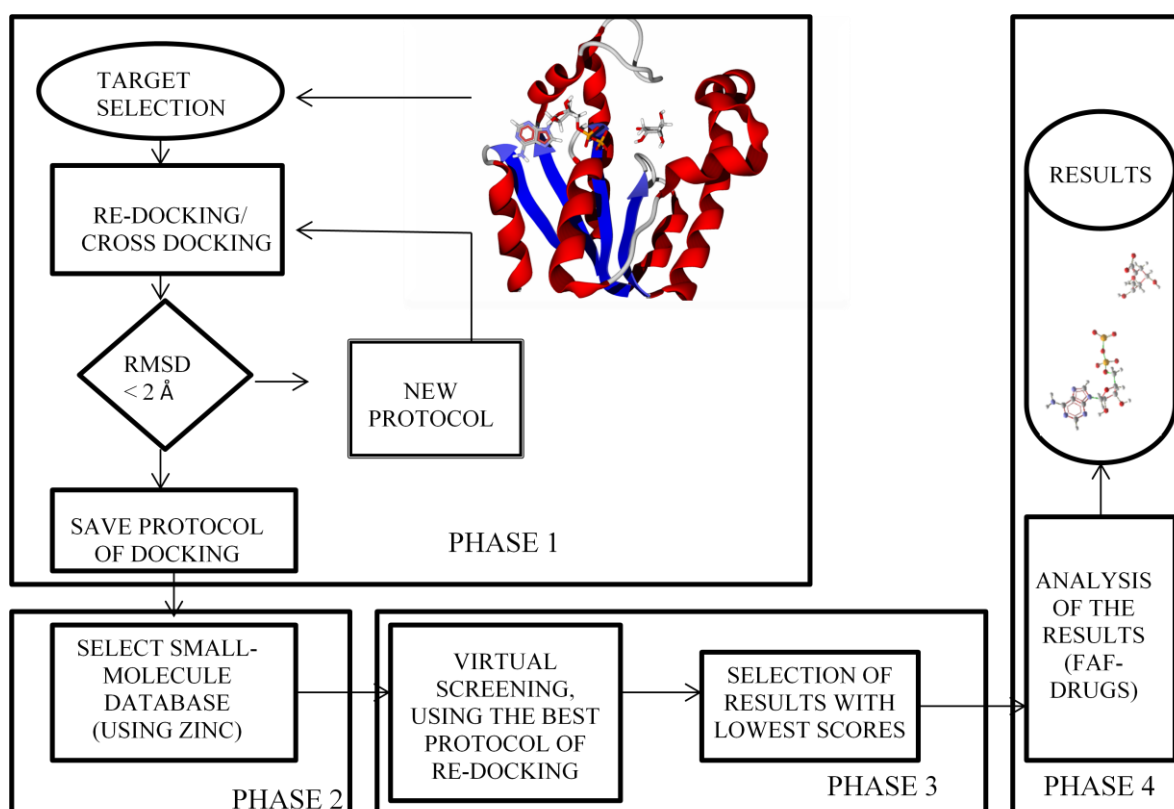
Radius (Å): _____

4.3 Virtual Screening

O processo de *Virtual Screening* (VS) utiliza a abordagem computacional executada através de softwares de simulação de *docking* molecular, nesse caso, como já mencionado, utilizamos o MVD (Thomsen & Christensen, 2006).

O Virtual Screening (VS) é dividido em 4 fases, como mostrado na figura 5.

Figura 5 - Processo de Virtual Screening e as diferentes fases de seleção dos melhores resultados.



Fonte: Coracini (2012)

A Seleção e a validação de um protocolo de docking (critério de seleção $RMSD < 2$, como descrito anteriormente, é a fase 1. Para validar o protocolo de docking, usamos as coordenadas cristalográficas da shikimate kinase disponíveis no protein data bank (PDB) através do código selecionado 2DFN. O protocolo escolhido para função score foi MolDock score [GRID] e algoritmo de busca MolDock Optimizer a partir do raio da esfera de docking de 10 Å e coordenadas ($x = -24.66$, $y = -3.96$ e $z = 15.93$) Å. Esse protocolo revelou melhores valores de RMSD. Salientando que o critério RMSD é dependente do número de ângulos de torção, e um critério menos exigente pode ser adotada para re-docking de um ligante com um número de ângulos de torção superior a 10 (Thomsen & Christensen, 2006). Depois que um

protocolo de docking é escolhido, seleciona-se um banco de dados de pequenas moléculas para ser usado no VS(fase 2). Foi usada uma biblioteca de ligantes provenientes de produtos naturais derivados de moléculas (ZINC Natural products) disponíveis comercialmente. Os ligantes (sdf file) são baixados através do site <http://zinc.dock.org> (Irwin & Shoichet, 2005), com total de 89.425 moléculas divididas em 23 conjuntos (*sets*).

Além de bases de dados comercialmente orientadas, o banco de dados ZINC também fornece uma interface para a construção de pequenas moléculas com base na similaridade molecular, tais como coeficiente de Tanimoto (De Azevedo, 2010; Timmers *et al.*, 2008).

Na fase 3, iniciam-se as simulações de docking para cada ligante presente no banco de dados selecionado. Durante as simulações, diversas orientações podem ser obtidas para cada ligante. Assim, selecionamos os menores scores obtidos. O critério de seleção escolhido foi Moldock Score. A função escore utilizada pelo MOLDOCK melhora a precisão de funções escores com as ligações de hidrogênio e novos sistemas de carga. Quatro funções *scores* são implementadas no MOLDOCK, incluindo *MOLDOCK Score* e *PLANTS Score* (De Azevedo, 2010). Estas duas funções oferecem versões baseadas em *grid*, em que a direcionalidade da ligação de hidrogênio não é considerada. No presente protocolo, usamos *grid-based MOLDOCK score*, uma vez que oferece uma velocidade aproximadamente quatro vezes maior através do pré-cálculo dos valores de potenciais de energia.

Os ligantes com melhores *scores* identificados pelas simulações de docking são submetidos ao programa FAF-Drugs (Miteva *et al.*, 2006), que funciona como um filtro através da análise das propriedades físico-químicas dos ligantes (fase 4). É baseado nos parâmetros para avaliar a biodisponibilidade oral de uma droga nos estágios iniciais do descobrimento de drogas. Os potenciais inibidores são avaliados através das Regras de Lipinsky (RO5), usadas para avaliar o potencial que uma molécula apresenta de ser absorvida oralmente, devendo satisfazer os critérios: peso molecular menor ou igual a 500, Log *P* menor ou igual a 5, número de grupos aceitadores de ligação hidrogênio menor ou igual a 10 e número de grupos doadores de hidrogênios menor ou igual a 5 (Eiben & Smith, 2003; Lipinski *et al.*, 2001). Também são aplicadas as Regras de Veber que afirmam que um composto para apresentar biodisponibilidade oral deve satisfazer os critérios: ligações flexíveis ≤ 10 e área de superfície polar ≤ 140 (Veber *et al.*, 2002).

Para melhor visualizar as interações intermoleculares entre resíduos da proteína-alvo e do ligante, com acesso aos resíduos da proteína que interagem com os átomos da pequena molécula por pontes de hidrogênio e interações hidrofóbicas, é usado o programa Ligplot (Wallace *et al.*, 1995). Assim, é possível avaliar os resíduos de aminoácidos da proteína que

interagem com os átomos da pequena molécula em teste, além de entender a natureza dessas interações.

4.3.1 Fator de Enriquecimento

Com o intuito de validar o protocolo de VS, é calculado o fator de enriquecimento (EF) que funciona como um controle de qualidade, onde em uma amostra aleatória se adiciona ligantes que se sabe que apresentariam um bom resultado e então é rodado o protocolo escolhido de VS. O EF é definido pela equação:

$$EF = \frac{H_a/H_t}{A/N}$$

Onde o H_a é o número de compostos ativos encontrados na posição H_t em um total de N compostos, nos quais A são ativos. O sucesso desse VS implica em $EF \gg 1$ (resultado considerado bom). Para essa simulação de validação é testado um subconjunto de *set* de quinases chamado decoys, ou seja, “armadilhas” que foram selecionados previamente com base na similaridade das moléculas. Esse *set* contém 1079 moléculas ($N=1079$) que foi enriquecido com 4 inibidores conhecidos MtSK ($A=4$) (Mulabagal & Calderón, 2010). Esse banco de dados para decoys MtSK e inibidores ativos MtSK está disponível para download no site <http://azevedolab.net/14.html>.

Capítulo 5

Resultados e Discussão

- 5.1 *Re-docking e Cross-docking*
 - 5.2 *Virtual Screening*
 - 5.2.1 Fator de Enriquecimento
 - 5.2.2 Seleção e Análise dos melhores ligantes
 - 5.3 Interações moleculares
-

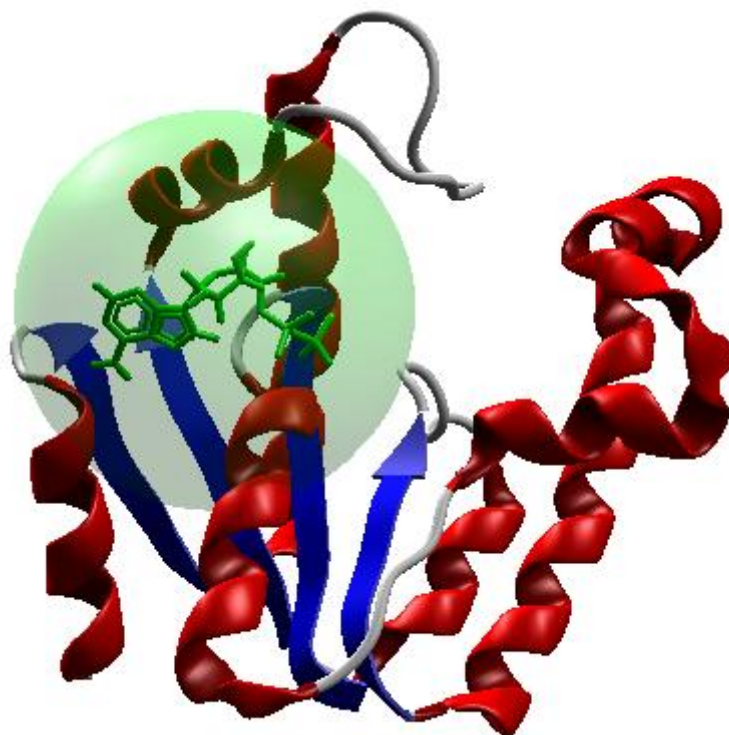
5. RESULTADOS E DISCUSSÃO

5.1 Re-docking e cross-docking

Inicialmente, as simulações de *cross-docking* foram realizadas com 12 estruturas cristalográficas (códigos de acesso PDB: 2DFN, 1U8A, 1ZYU, 2G1K, 2IYQ, 2IYR, 2IYS, 2IYX, 2IYY, 2IYZ, 3BAF, 1WE2). O PDB 2DFN apresentou os melhores valores de RMSD. As simulações de re-docking foram feitas com a estrutura 2DFN e a busca pelos melhores protocolos de dockings (fase 1 do protocolo de VS). O critério para melhor descrever a qualidade do programa de docking é o RMSD. A estrutura 2DFN, gerou um RMSD de 0.48 Å com função score MolDock Score [GRID], algoritmo de busca MolDock Optimizer. As análises dos resultados de re-docking com a combinação de quatro algoritmos de busca e quatro funções score (um total de 16 protocolos de docking) geraram RMSD de 0.46 a 10.48 Å. As simulações de cross docking obtiveram RMSD de 0.48 a 1.99 validando o protocolo de docking escolhido. Idealmente, resultados que, comparados à estrutura cristalográfica, resultam em um RMSD até 2 Å são considerados bons, podendo ter um valor maior se possuírem mais ângulos de torção.

Os parâmetros usados nos dockings, especialmente as características de busca são otimizadas por diversos runnings das simulações de docking de estruturas complexadas. Alguns dos parâmetros usados: raio da esfera de docking (10 Å), número de séries (10), tamanho da população máxima (100), coordenadas atômicas (X= -24.66, Y= -3.96, Z=15.93). A figura 6 mostra a esfera de docking usada nas simulações de *re-docking*.

Figura 6 - Esfera de *docking* (verde) usada nas simulações de *docking* molecular.



Fonte: Coracini (2012).

5.2 Virtual screening

Utiliza metodologias computacionais para identificar potenciais moléculas inibidoras contra uma proteína alvo específico, no caso a chiquimato quinase. Duas principais metodologias são usadas no VS, uma é o método que busca por semelhança com ligantes validados e outra é o método de docking molecular que requer informação cristalográfica do alvo, abordada neste estudo. Estudos de VS para identificação de inibidores MtSK como drogas contra a tuberculose foram realizados por outros grupos pelo procedimento similar de docking molecular (Segura & Rodriguez, 2008; Kumar, 2010). Porém, os dados experimentais ainda são muito limitados sobre o efeito destes compostos MtSK disponíveis. O presente estudo traz a novidade de focarmos nossas tentativas de docagem no bolsão de ligação do ATP do complexo MtSK usando diferentes métodos de VS e seleção de compostos de acordo com suas propriedades farmacológicas. As simulações VS foram realizadas utilizando o programa MVD. Primeiramente, foi realizado um VS com “decoys” para se testar a precisão e melhor validar o protocolo de docking. A partir do protocolo escolhido e validado, utilizou-se a biblioteca de produtos naturais do ZINC, um banco de dados gratuito com mais de 13 milhões de compostos comercialmente disponíveis, em formato pronto para o virtual screening, com 89.425 pequenos ligantes, à procura dos que melhor encaixem no bolsão de ligação do ATP da proteína de estrutura PDB 2DFN.

5.2.1 Fator de Enriquecimento

Em uma biblioteca com um total de 1.083 moléculas (“decoys”), incluindo 4 ligantes conhecidos, o programa selecionou 11 moléculas (Ht=11), dentro delas, 1 inibidor conhecido (Ha=1) e 10 falsos ligantes. Com isso,

$$EF = (1/11) / (4/1079) = 24.57\%$$

verificou-se a chance de 24.57 vezes maior de se escolher um composto ativo no banco de dados que um inativo. Resultados de estudos anteriores com kinases exibiram fatores de enriquecimento de 1.2-54, por exemplo (Brooijmans & Kuntz, 2003; Vadivelan *et al.*, 2007; Cavasotto & Abagyan, 2004; Gil-Redondo *et al.*, 2008), indicando que o presente protocolo de VS é adequado.

5.2.2 Seleção e Análise dos melhores ligantes

Após as simulações de docking, foram selecionados 17 possíveis ligantes, de um total de 89.425 moléculas com valores de energia ≤ -190.000 (critério de seleção: MOLDOCK score). Esses 17 inibidores (moléculas ranging -190.173 to -201.689) foram submetidos aos testes filtros, disponíveis na web Server FAF-Drugs (Miteva *et al.*, 2006), sendo que 11 satisfazem as Regras de Lipinsky e 9 satisfazem as Regras de Weber. Destes, 5 satisfazem ambas as regras, sendo considerados compostos de melhores características físico-químicas para uma boa disponibilidade oral. As simulações de 23 *sets* de compostos foram executados em 16 computadores iMac (Intel Processor Core 2 Duo, 236 2.66 GHz, 2 GB SDRAM DDR3 1066 MHz) durando aproximadamente 3 meses para rodar. Vale salientar, que alguns dos melhores MOLDOCK *scores* que não satisfazem as regras de biodisponibilidade oral, podem ter toxicidade diminuída com pequena modificação na sua estrutura. A figura 7 mostra a estrutura molecular dos 5 ligantes com melhores *scores* e baixa toxicidade. Os ligantes selecionados estão mostrados na tabela 2, na qual o MOLDOCK *score* dos 5 ligantes variam de -190.538 a -198.629. Os 5 compostos satisfazem as RO5 e as Regras de Veber, portanto, exibindo a performance das melhores estruturas com baixa toxicidade.

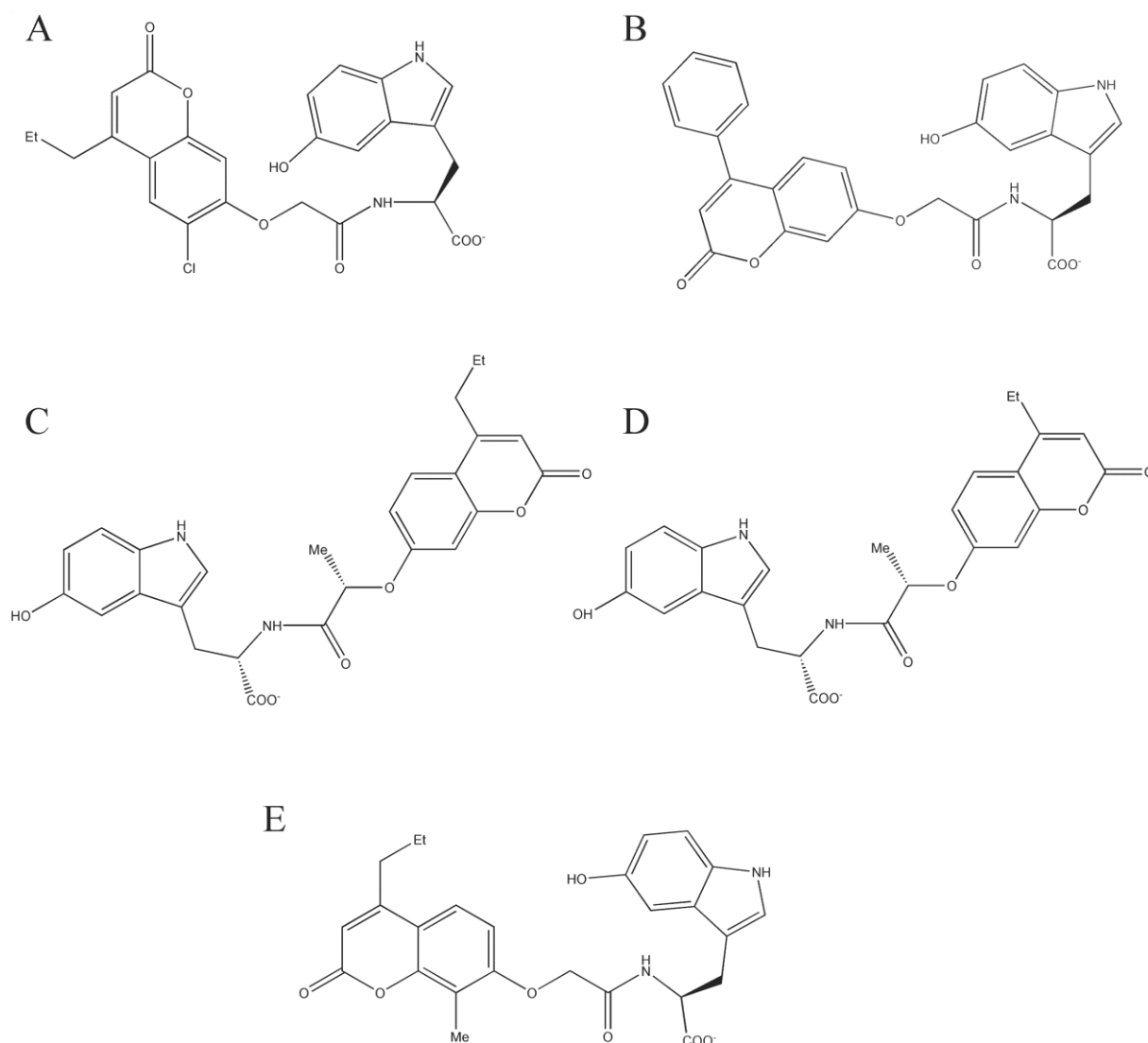
Tabela 2 - Propriedades físico-químicas dos ligantes através da análise pelo filtro FAF-Drugs.

Ligante	Código ZINC	Peso Molecular (Da)	Número de aceptadores de ligações de H	Número de doadores de H	X Log P	Ligações flexíveis	Área de superfície polar (PSA)
1	02135897	497.7	9	3	3.54	10	133.860
2	02127309	497.3	9	3	3.39	9	133.860
3	02133638	477.3	9	3	3.45	10	133.860
4	02130347	463.3	9	3	2.88	9	133.860
5	02160785	477.3	9	3	3.15	10	133.860

Fonte: Coracini (2012).

Os ligantes acima satisfazem as Regras de Lipinsky e as Regras de Weber.

Figura 7 - Estruturas moleculares dos 5 possíveis inibidores: **A)** ZINC02135897. **B)** ZINC02127309. **C)** ZINC02133638. **D)** ZINC02130347. **E)** ZINC 02160785.



As estruturas foram desenhadas através do Programa ChemDraw Ultra 7.0.

Fonte: Coracini (2012).

5.3 Interações Intermoleculares

Para uma melhor compreensão das interações dos compostos selecionados com MtSK, usamos o programa LIGPLOT (Wallace *et al.*, 1995). Ao analisarmos as interações intermoleculares entre proteína e pequenas moléculas, encontramos alguns resíduos de aminoácidos da proteína que frequentemente estão presentes na interação, seja em ligações de hidrogênio ou em contatos de Van der Waals, indicando resíduos chaves responsáveis pela especificidade da interação com o substrato e ligantes testados.

Os resíduos de aminoácidos presentes na interação intermolecular do ADP com a chiquimato quinase (figura 8), são responsáveis pela especificidade da ligação ligante. Ligações de hidrogênio envolvem o ligante e os resíduos Ser16, Arg117, Gly12, Lys15, Gly14, Thr17, Arg153 e contatos de Van der Waals estão presentes com os resíduos Arg110 e Pro155. A análise da ligação do ADP com a cavidade da proteína, indicou que os 5 possíveis ligantes apresentam interações com os resíduos Lys15, Gly14, Thr17, Gly12, Arg117, Ser16, Arg110 e Pro155, revelando a importância desses resíduos na ligação.

A tabela 3 mostra as interações intermoleculares dos 5 compostos selecionados. Os resíduos Thr17, Gly12 e Ser16 estão presentes nos 5 melhores ligantes, podendo ser com ligações de H ou contatos de VDW. No entanto, Lys15, Arg117 e Gly14 estão presentes em todos os ligantes, mas apenas em ligações de H; já os resíduos Pro155, Val158, Arg110 e Ile18 também estão presentes nos ligantes, porém somente em ligações de VDW. Isso nos faz pensar, a importância desses resíduos para a especificidade da ligação. Em todos os ligantes, observou-se que ocorrem 4 a 8 contatos de VDW e de 6 a 8 ligações de hidrogênios. Os resíduos Lys15 e Arg117 parecem ter importância, pois estão presentes em outro estudo (Gan *et al.*, 2006) com MtSK.

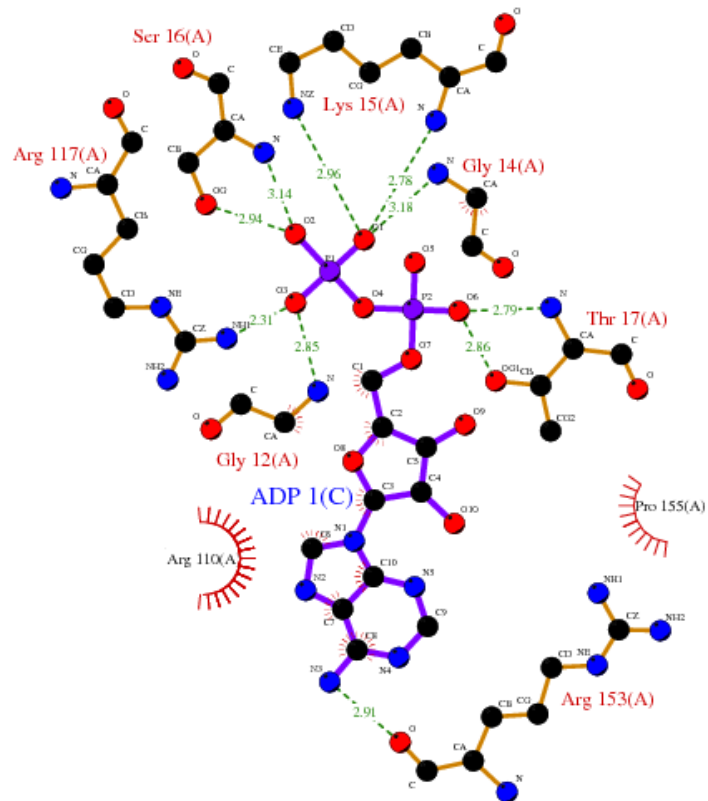
As figuras 9 e 10 mostram os ligplot dos ligantes ZINC02135897 e ZINC02127309, ou seja, as interações intermoleculares dos compostos com melhores *moldock score* entre os 5 compostos selecionados.

Tabela 3 - Interações intermoleculares dos 5 melhores ligantes selecionados. A presença de um X indica que a interação ocorre. HB são as ligações de hidrogênio e VDW são os contatos de Van der Waals.

Resíduos	Ligantes (códigos ZINC)				
	HB	ZINC	ZINC	ZINC	ZINC
	02135897	02127309	02133638	02130347	02160785
Thr 17	X	X	X	X	
Gly 12		X	X		X
Lys 15	X	X	X	X	X
Ser 16		X			
Arg 117	X	X	X	X	X
Thr 115		X			
Asp 34	X		X	X	X
Ser 13	X	X		X	
Gly 14	X	X	X	X	X
Arg 153					X
VDW					
Asn 154	X		X	X	X
Pro 155	X	X	X	X	X
Thr 17					X
Val 158	X	X	X	X	X
Arg 110	X	X	X	X	X
Ile 18	X	X	X	X	X
Ser 13					X
Gly 12	X			X	
Ser 16	X		X	X	X
Thr 150	X		X	X	

Fonte: Coracini (2012).

Figura 8 – Ligplot do bolsão de ligação do ATP com a enzima chiquimato quinase.

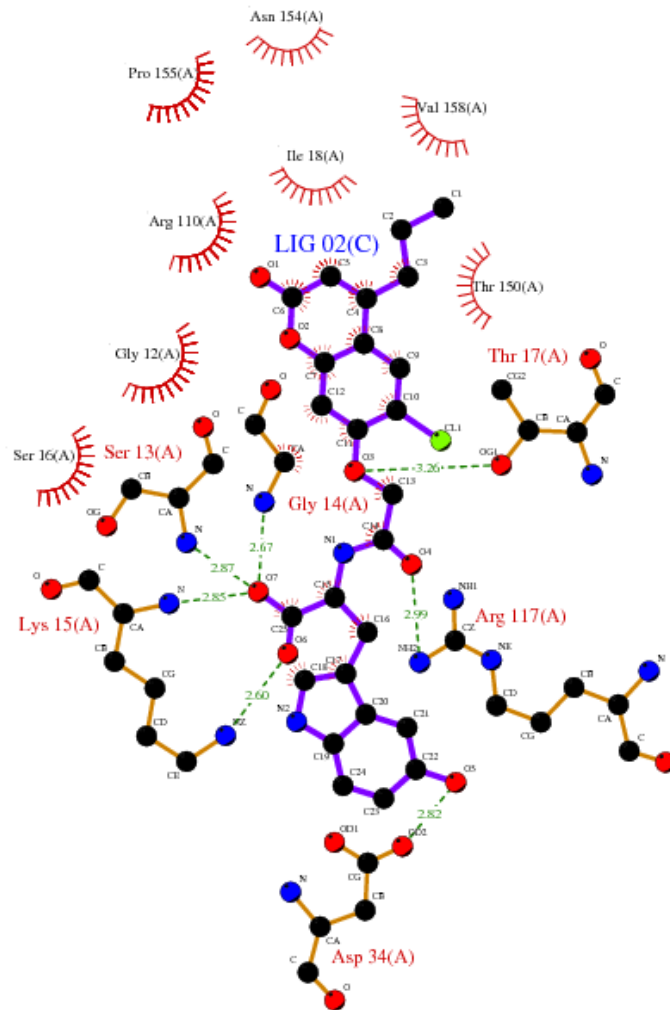


Key

-  Ligand bond
-  Non-ligand bond
-  Hydrogen bond and its length
-  Non-ligand residues involved in hydrophobic contact(s)
-  Corresponding atoms involved in hydrophobic contact(s)

Fonte: Coracini (2012).

Figura 9 - Interações intermoleculares da molécula ZINC02135897 (LIG 02 C) com o bolsão de ligação do ATP da chiquimato quinase.

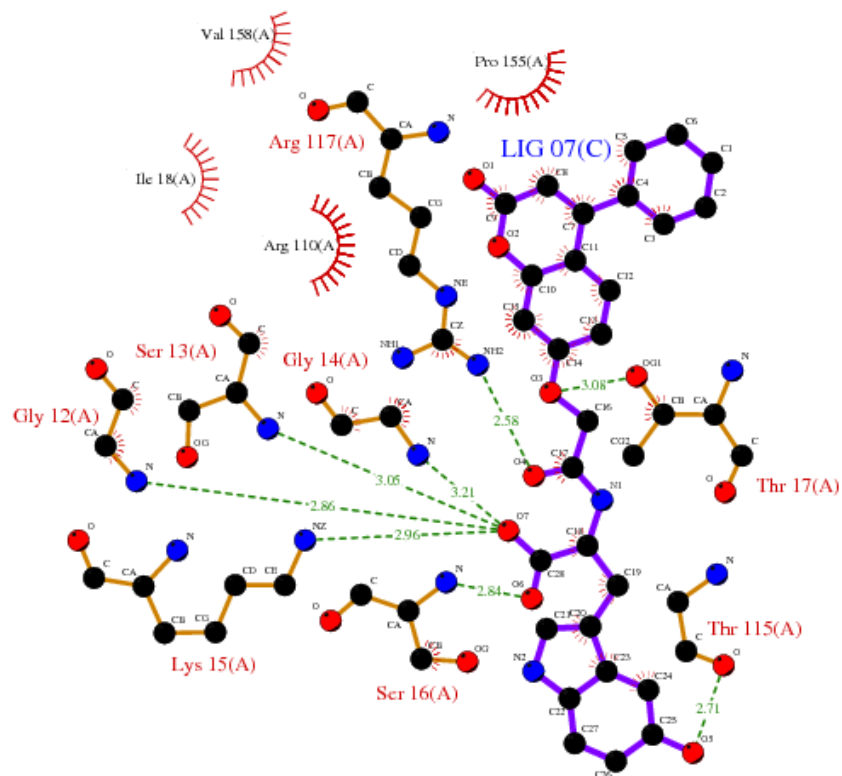


Key

- ● Ligand bond
- ● Non-ligand bond
- ● Hydrogen bond and its length
- Non-ligand residues involved in hydrophobic contact(s)
- Corresponding atoms involved in hydrophobic contact(s)

Fonte: Coracini (2012).

Figura 10 - Interações intermoleculares da molécula ZINC02127309 (LIG 07 C) com o bolsão de ligação do ATP da chiquimato quinase.



Key

- Ligand bond
- Non-ligand bond
- Hydrogen bond and its length
- His 53 Non-ligand residues involved in hydrophobic contact(s)
- Corresponding atoms involved in hydrophobic contact(s)

Fonte: Coracini (2012).

Capítulo 6

Conclusão

6.0 CONCLUSÃO

Diante das inúmeras moléculas disponíveis nas bibliotecas virtuais para testes como possíveis inibidores de diversas proteínas, utilizamos o processo virtual de encaixe-induzido com um excelente protocolo de docking que foi capaz de recuperar a posição cristalográfica do ligante presente no sítio ativo do ATP da chiquimato quinase. As simulações de cross-docking e re-docking geraram, na maioria das vezes, RMSD abaixo de 2 Å. Um aspecto inovador da presente abordagem, com relação a estudos anteriores sobre a MtSK (**Heberlé & De Azevedo, 2011; Vianna & De Azevedo, 2011**) foi o foco das simulações no sítio de ligação de ATP. Tal abordagem visa validar o uso de bibliotecas de decoys disponíveis para inibidores competitivos de quinases (**Huang *et al.*, 2006**). O protocolo de virtual screening é capaz de encontrar um inibidor da SK comprovado pelo VS realizado com decoys que foi capaz de gerar um fator de enriquecimento de 24.57, melhor que o identificado contra outra quinase, a CDK2 (**Huang *et al.*, 2006**). Identificamos uma molécula (ZINC02135897) com moldock *score* muito baixo, podendo ser de grande interesse para estudos posteriores de desenho de inibidores da SK. As 5 melhores moléculas testadas apresentam vários resíduos das interações em comum, revelando moléculas similares, o que nos leva a pensar que a seleção possa ter encontrado, de fato, as moléculas que melhor perfazem o encaixe proposto no estudo. Também encontramos vários resíduos nas interações intermoleculares dos compostos com melhor performance, que podem ser resíduos chaves nas ligações do MtSK com possíveis inibidores.

Com este processo, pode-se agilizar mais rapidamente o descobrimento de drogas, com um menor custo e alta eficiência, diminuindo o número de moléculas a serem testadas posteriormente *in vitro* e *in vivo*. É, sem dúvida, um processo inteligente que indica moléculas com potencial de se tornarem fármacos de baixa toxicidade e alta eficácia contra a tuberculose.

Referências

REFERÊNCIAS

- Araújo JQ, Lima JA, Pinto Ada C, de Alencastro RB, Albuquerque MG. Docking of the alkaloid geissospermine into acetylcholinesterase: a natural scaffold targeting the treatment of Alzheimer's disease. *J Mol Model*. 2011; 17(6): 1401-1412.
- Arora N, Banerjee AK, Murty US. In silico characterization of Shikimate Kinase of *Shigella flexneri*: a potential drug target. *Interdiscip Sci* 2010; 2: 280-290.
- 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 SD. The use of biodiversity as source of new chemical entities against defined molecular targets for treatment of malaria, tuberculosis, and T-cell mediated diseases-a review. *Mem Inst Oswaldo Cruz*. 2005; 100 (6):475-506.
- Bentley R. The shikimate pathway as a target for herbicidas. In: Dodge AD (ed) *Herbicides and Plant metabolism*. Cambridge, UK, 1990; pp 97-112.
- Brooijmans N, Kuntz ID. Molecular recognition and docking algorithms. *Annu Rev Biophys Biomol Struct*. 2003; 32: 335-373.
- Cavasotto CN, Abagyan RA. Protein flexibility in ligand docking and virtual screening to protein kinases. *J Mol Biol*. 2004; 337(1):209-225.
- De Azevedo WF Jr. MolDock applied to structure-based virtual screening. *Curr Drug Targets*. 2010; 11:327-334.
- De Azevedo W.F Jr, Dias, R. Evaluation of ligand-binding affinity using polynomial empirical scoring functions. *Bioorg Med Chem*. 2008; 16:9378-83.
- De Azevedo Jr WF. Structure-based virtual screening. *Curr Drug Targets*. 2010; 11:261-263.
- De Azevedo WF Jr. Molecular dynamics simulations of protein targets identified in *Mycobacterium tuberculosis*. *Curr Med Chem*. 2011; 18: 1353-1366.
- Dias MV, Vasconcelos IB, Prado AM, Fadel V, Basso LA, De Azevedo WF Jr, Santos DS. *Acta Crystallogr F Struct Biol Cryst Commun*. 2007; 63:1-6.
- Eiben AE, Smith JE. *Introduction to Evolutionary Computing*. Springer-Verlag: New York, 2003.
- Filgueira de Azevedo W Jr, Canduri F, Simões de Oliveira J, *et al*. Molecular model of shikimate kinase from *Mycobacterium tuberculosis*. *Biochem Biophys Res Commun*. 2002; 295: 142-148.
- 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: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem*. 2004; 47(7):1739-1749.

- Gan J, Gu Y, Li Y, Yan H, Ji X. Crystal structure of *Mycobacterium tuberculosis* shikimate kinase in complex with shikimic acid and an ATP analogue. *Biochemistry*. 2006; 5(28):8539-8545.
- Gil-Redondo R, Estrada J, Morreale A, Herranz F, Sancho J, Ortiz AR. VSDMIP: virtual screening data management on an integrated platform. *J Comput Aided Mol Des*. 2008; 23(3):171-184.
- Gu Y, Reshetnikova L, Li Y, Wu Y, Yan H, Singh S, Ji X. Crystal structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the dynamic role of the LID domain in catalysis. *J Mol Biol*. 2002; 319: 779-789.
- Habig M, Blechschmidt A, Dressler S, *et al*. Efficient elimination of nonstoichiometric enzyme inhibitors from HTS hit lists. *J Biomol Screen*. 2009; 14: 679-689.
- Han C, Zhang J, Chen L, *et al*. Discovery of *Helicobacter pylori* shikimate kinase inhibitors: bioassay and molecular modeling. *Bioorg Med Chem*. 2007; 15: 656-662.
- Hartmann MD, Bourenkov GP, Nicolai Strizhov AO, Bartunik HD. Mechanism of Phosphoryl Transfer catalyzed by Shikimate Kinase from *Mycobacterium tuberculosis*. *J Mol Biol*. 2006; 364: 411-423.
- Heberlé G, De Azevedo WF Jr. Bio-inspired algorithms applied to molecular docking simulations. *Curr Med Chem*. 2011; 18(9):1339-1352.
- Hsu KC, Cheng WC, Chen YF, *et al*. Core site-moiety maps reveal inhibitors and binding mechanisms of orthologous proteins by screening compound libraries. *PLoS One*. 2012; 7: e32142.
- Huang N, Shoichet BK, Irwin J.J. Benchmarking sets for molecular docking. *J Med Chem*. 2006, 49(23), 6789-801.
- Irwin JJ, Shoichet BK. ZINC – a free database of commercially available compounds for virtual screening. *J Chem Inf Model*. 2005; 45:177-182.
- Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov*. 2004; 3(11):935-949.
- Krell T, Coggins JR & Laphorn AJ. The three-dimensional structure of shikimate kinase. *J Mol Biol*. 1998; 278 (5): 983-997.
- Krell T, Maclean J, Boam DJ, Cooper A, Resmini M, Brocklehurst K, Kelly SM, Price NC, Laphorn AJ & Coggins J. Biochemical and X-ray crystallographic studies on shikimate kinase: the important structural role of the P-loop lysine. *Protein Sci*. 2001; 10(6):1137-49.
- Kumar M, Verma S, Sharma S, Srinivasan A, Singh TP, Kaur P. Structure-based in silico design of a high-affinity dipeptide inhibitor for novel protein drug target Shikimate kinase of *Mycobacterium tuberculosis*. *Chem Biol Drug Des*. 2010; 76:277-84.

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev.* 2001; 46(1-3):3-26.

Lyne P.D. Structure-based virtual screening: an overview. *Drug Discovery Today.* 2002; Vol. 7, No. 20.

Marshall GR. Introduction to chemoinformatics in drug discovery – A personal view. In: OPREA, T.I. *Chemoinformatics in drug discovery.* 2004; Weinheim: Wiley-VHC. 1-22.

Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância Epidemiológica. Tratamento diretamente observado (TOD) da tuberculose na atenção básica: protocolo de enfermagem. – Brasília: Ministério da Saúde, 2011. p. 11.

Miteva MA, Viola S, Montes M, Gomez D, Tuffery P, Villoutreix BO. FAF-drugs: free adme/tox filtering of compound collections. *Nucleic Acids Res.* 2006; 34:W738-W744.

Mulabagal V, Calderón AI. 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.* 2010; 82:3616-3621.

Padyana AK, Burley SK. Crystal Structure of Shikimate 5-Dehydrogenase (SDH) Bound to NADP: Insights into Function and Evolution. *Structure.* 2003; 11(8):1005-1013.

Parish T, Stoker NG. The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*. *Microbiology.* 2002; 148:3069–3077.

Pauli I, Caceres RA, de Azevedo WF Jr. Molecular modeling and dynamics studies of Shikimate Kinase from *Bacillus anthracis*. *Bioorg Med Chem.* 2008; 16: 8098-8108.

Peitsch MC. Manuel Peitsch discusses knowledge management and informatics in drug discovery. *Drug Discov Today Biosilico.* 2004; 02:94-96.

Pereira JH, Oliveira JS, Canduri F, Dias MVB, Palma MS, Basso LA, Santos DS & De Azevedo WF. Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid. *Acta Crystallogr D Biol Crystallogr.* 2004; 60: 2310-2319.

Roberts F, Roberts CW, Johnson JJ, Kyle DE, Krell T, Coggins JR, Coombs GH, Milhous WK, Tzipori S, Ferguson DJ, Chakrabarti D, McLeod R. Evidence for the shikimate in pathway in apicomplexan parasites. *Nature.* 2002; 393(6687):801-805.

Saidenberg DM, Passarelli AW, Rodrigues AV, *et al.* Shikimate kinase (EC 2.7.1.71) from *Mycobacterium tuberculosis*: kinetics and structural dynamics of a potential molecular target for drug development. *Curr Med Chem.* 2011; 18: 1299-310.

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

Silva VB (2008) Estudos de modelagem molecular e relação estrutura atividade da oncoproteína hnRNP K e ligantes. USP-Ribeirão Preto.

Silveira NJ, Uchôa HB, Pereira JH, *et al.* Molecular models of protein targets from *Mycobacterium tuberculosis*. J Mol Model. 2005; 11: 160-166.

Thomsen R, Christensen MH. MolDock: a new technique for high-accuracy molecular docking. J Med Chem. 2006; 49(11):3315-3321.

Timmers LFS, Pauli I, Caceres RA, De Azevedo Jr WF. Drug-binding databases. Curr Drug Targets. 2008; 9:1092-1099.

Vadivelan S, Sinha BN, Irudayam SJ, Jagarlapudi SA. Virtual screening studies to design potent CDK2-cyclin A inhibitors. J Chem Inf Model. 2007; 47(4):1526-1535.

Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem. 2002; 45(12):2615-2623.

Vianna CP, de Azevedo WF Jr. Identification of new potential *Mycobacterium tuberculosis* shikimate kinase inhibitors through molecular docking simulations. J Mol Model. 2011; 18(2):755-764.

Vonrhein C, Schlauderer GJ, Schulz GE. Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. Structure. 1995; 3(5): 483-90.

Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. 1995; 8(2):127-134.

World Health Organization. WHO Report. Geneva. Switzerland. WHO/CDS/TB/2010. <http://www.who.int/tb/en/> Accessed in July 2012.

Yan H, Tsai MD. Nucleoside monophosphate kinases: structure, mechanism, and substrate specificity. Adv Enzymol Relat Areas Mol Biol. 1999; 73:103-134.

Zahrt TC. Molecular mechanisms regulating persistent *Mycobacterium tuberculosis* infection. Microbes Infect. 2003; 5(2):159-67.

Anexo

Anexo

Shikimate kinase, a protein target for design of antitubercular drugs

Juliane Dors Coracini,
Walter Filgueira de Azevedo Jr.

Artigo submetido ao *Current Drugs
Targets*

Shikimate kinase, a protein target for design of antitubercular drugs

Juliane Dors Coracini^{1,2} and Walter Filgueira de Azevedo Jr.^{1,2*}

¹Faculdade de Biociências, Laboratório de Bioquímica Estrutural (LaBioQuest), Pontifícia Universidade Católica do Rio Grande do Sul, (PUCRS), Av. Ipiranga 6681, Porto Alegre, RS 90619-900, Brazil

²Programa de Pós-Graduação em Medicina e Ciências da Saúde, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, Brazil

*Corresponding author: walter@azevedolab.net

Faculdade de Biociências-PUCRS. 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

Shikimate kinase is the fifth enzyme of the shikimate pathway, which has been identified in fungi, apicomplexans, plants and prokaryotes. This metabolic route is composed of seven steps, which converts erythrose-4-phosphate and phosphoenol pyruvate to chorismic acid and is responsible for the biosynthesis of aromatic amino acids. Shikimate kinase has been shown to be essential to the survival of *Mycobacterium tuberculosis*, and since it is absent in human, this *enzyme is considered to be a* target for chemotherapeutic for development of antitubercular drugs. This review highlights the available crystallographic structures for shikimate kinases, which has been used to identify structural features for ligand-binding affinity. We also describe molecular docking studies focused on shikimate kinase. These computational studies were performed in order to identify the new generation of antitubercular drugs and several potential inhibitors have been described. In addition, a structural comparison of shikimate kinase ATP-binding pocket with human kinases is described. This structural analysis describes the potential beneficial aspects of abundant structural studies of human kinases and their inhibitors to bring further understanding of the ligand-binding specificity for shikimate kinase.

Keywords: shikimate kinase, kinase, *Mycobacterium tuberculosis*, drug-design, bioinformatics, molecular docking

Introduction

Shikimate pathway (SP) is an essential metabolic route for the survival of bacteria [1]. Based on the fact that it is absent in humans, SP enzymes have become valuable targets for development of a new generation of antibacterial drugs. This affirmation is anchored on the concept of selective toxicity, which is the fundamental principle behind Paul Ehrlich's proposal of the "magic bullet" [2]. This view that an illness could be cured by inhibiting a specific protein target is the pivotal concept in all computational drug-design methods discussed in this review. If we focus our analysis on antibacterial drugs only, we may say that this concept simply means that the majority of successful antibacterial drugs are those that target some dissimilarity in a cellular structure or metabolic process between the human cell and the target bacteria. On the other hand, the more alike the human cell and the bacteria are, more difficult it is to selectively strike solely the bacteria. Molecules that do not distinguish very well between types of cells are usually more toxic to the human than those that are more selective [3,4].

Generations of pharmacologists and medicinal chemists considered the Ehrlich's concept of the magic bullet as a molecule that targets a single essential protein in a restricted and highly specific mode [3]. SP enzymes fit quite well under this concept, since they are absent in animals and present in the bacteria to be targeted by the drugs. Therefore a "magic bullet" targeting SP enzymes is expected to miss the human protein targets, which may result in low toxicity for the drug. Nevertheless, this is only a source of inspiration for the initial prospective study of new potential inhibitors for these enzymes. There are exceptions to this concept in the SP, for instance, glyphosate (N-(phosphonomethyl)glycine) is an inhibitor of the enzyme 5-enolpyruvyl-shikimate-3-phosphate synthase (EPSPS) [5] the sixth step of SP, but it has been shown to be harmful to animals, with cancer producing effects [6]. The isopropylamine salt of glyphosate is the active ingredient of the herbicide Roundup®. This salt of glyphosate is a non-selective systematic, broad-spectrum herbicide, commercialized by the Monsanto Company. It is supposed to be only active when applied to the foliage and green bark of plants [5].

Nevertheless, a recent study described that cancer producing effects of glyphosate in rats fed with Roundup-tolerant genetically modified maize which was sprayed with the herbicide [6]. These results showed that both female and male rats were more prone to die than the control group. This study highlights that fifty percent of the male rats and about 70 percent of the females eating Monsanto genetic modified maize died earlier compared to 30 percent of males and 20 percent of females not eating genetically modified maize. In addition, this research

paper [6] has just been published and it has already been criticized for the statistical analysis of their results [7]. Yet, several other recent studies showed glyphosate toxicity for animals such as fishes, amphibians and also for human cells [8-16], which emphasizes the need of care in applying general principles as the *Paul Ehrlich's magic bullet concept* to the development of drugs.

The SP was revealed as a biosynthetic pathway through the pioneering work of Sprinson [17] and Davis and Mingioli [18]. The SP can be seen as a bridge uniting metabolism of carbohydrates to biosynthesis of aromatic compounds through seven metabolic steps. In the pathway, phosphoenolpyruvate and erythrose 4- phosphate are converted to chorismate [19-20]. SP is also present in plants, fungi and apicomplexan organisms [21]. Although it has been reported identification of all SP enzymes in apicomplexan organisms, such as *Toxoplasma gondii* (toxoplasmosis), and *Cryptosporidium parvum* (cryptosporidiosis) [22-24], SP enzymes have not been clearly identified the *Plasmodium falciparum* genome [25]. Only chorismate synthase, the last enzyme of the seven SP enzymes, could be conclusively found [26]. Nevertheless, careful bioinformatic sequence analysis was able to identify homologues to other enzymes in the shikimate pathway with low sequence homology in *Plasmodium falciparum* and other related apicomplexan parasites [25, 27], which somehow provides further support to the results about sensitivity to herbicide glyphosate observed for *Plasmodium falciparum* [24].

In this scenario, a lot of efforts have been concentrated on identification of inhibitors of the seven enzymes of SP. Furthermore, several research groups have undertaken structural and in silico studies, focuses on the fifth enzyme of SP, the shikimate kinase (EC 2.7.1.71).

The present modest review is focused on the SK. We describe recent developments on the structural biology and in silicon studies undertaken on the SKs from bacteria, with special attention to SK from *Mycobacterium tuberculosis*.

Shikimate pathway enzymes

Comparative analysis of the distribution of SP enzymes in different taxonomic groups indicates a variation in their presence [7, 28]. Bacteria present seven individual enzymes, which are encoded by separate genes (aroF, aroB, aroD, aroE, aroK (and aroL), aroA and aroC). All structural information about SP enzymes is available on an open database called SKPDB [29]. This database brings together molecular models for all the enzymes of SP obtained by homology modeling and the atomic coordinates obtained from crystallographic studies. In addition, the database describes each metabolic step of SP. We briefly illustrate

each reaction step of the SP here. The sequence of seven metabolic steps in the SP, from phosphoenolpyruvate and erythrose 4-phosphate to chorismate is shown in Fig. (1). Attached to each step in Fig. (1) is displayed a representative crystallographic structure of the enzyme responsible to the catalysis of that step. The product of these seven metabolic steps is chorismate, which is a general precursor for the synthesis of aromatic compounds, such as tyrosine, triptophane, phenylalanine, ubiquinone, menaquinones and folate. In the following paragraphs will briefly describe each enzymatic reaction. We also display the unique GenBank **ID for each gene of SP enzymes in** the genome of *Mycobacterium tuberculosis* [30].

The first enzyme of the SP is 3-deoxy-7-phosphoheptulonate synthase (DAHPS) (EC 2.5.1.54). For *Mycobacterium tuberculosis* the gene ID is 13318866 (updated on 23-Nov-2012). DAHPS is encoded by the *aroF* gene. This metabolic step is the condensation of erythrose-4-phosphate (E4P) and phosphoenolpyruvate (PEP) to 3-deoxy-D-arabinoheptulosonic acid-7-phosphate (DAHP). It was first identified in the genome of *Escherichia coli*, and the reaction has been revealed to be performed by three DAHP synthetase isoenzymes, the activity and formation of which are regulated by the aromatic amino acid end products [31-35]. This synthase requires a divalent cation for activity and catalyzes the aldol-like, stereospecific condensation of 2-phosphoenolpyruvate (PEP) and D-erythrose-4-phosphate (E4P) to 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) and inorganic phosphate [29, 31, 32]. In view of the fact that DAHPS is the first enzyme in the SP, it controls the quantity of carbon entering the metabolic route [36].

The second enzyme of SP is 3-dehydroquinate synthase (DHQS) (EC 4.2.3.4), which catalyzes the elimination of phosphate from DAHP to generate 3-dehydroquinate (DHQ) [36, 38], as shown on Fig. (1). This synthase is encoded by the *aroB* gene [39, 40] and its gene ID for *Mycobacterium tuberculosis* is 888392 (updated on 23-Nov-201). The enzyme DHQS needs divalent cations for optimal activity, such as Co^{2+} and Zn^{2+} and becomes active in the presence of inorganic phosphate, one of the reaction products and also requires NAD to become fully activated [29].

The following step of the metabolic route is catalyzed by the enzyme 3-dehydroquinate dehydratase (DHQD) (EC 4.2.1.10). This step is the dehydration of 3-dehydroquinate to 3-dehydroshikimate, which is metabolic stage shared by the catabolic quinate pathway. Furthermore, this reaction converts quinic acid to p-hydroxybenzoic acid that can be further metabolized via the β -keto-adipate pathway to acetyl-CoA and succinyl-CoA [40-43]. DHQD

is encoded by the *aroD* gene and its gene ID for *Mycobacterium tuberculosis* is 888397 (updated on 12-Oct-2012).

The enzyme shikimate-5-dehydrogenase (SDH) catalyzes the reduction of 3-dehydroshikimate to shikimate using NADH as a cofactor. It has been identified two classes of SDH in bacteria, known as *aroE* (EC 1.1.1.25) and *YdiB* (EC 1.1.1.282). SDH is characterized as a member of the superfamily of NAD(P)H-dependent oxidoreductases. This class of enzymes is typically subdivided into a number of families, including short chain dehydrogenases, medium chain dehydrogenases, aldo-keto reductases, and iron-activated alcohol dehydrogenases and long chain dehydrogenases [40, 44]. SDH from *Mycobacterium tuberculosis* has gene ID 887330 (updated on 12-Oct-2012).

SK catalyzes specific phosphorylation of the 3-hydroxyl group of D-shikimate to yield shikimate 3-phosphate using adenosine-5'-triphosphate (ATP) as a co-substrate [35, 44]. The *aroK* gene encodes shikimate kinase I (SK I) (E.C. 2.7.1.71), which has been shown to be essential for the survival of *Mycobacterium tuberculosis* [1]. In *Mycobacterium tuberculosis* its gene ID is 887434 (updated on 23-Nov-2012). An additional SK has been identified in *Escherichia coli*, the SK II, which is encoded by the *aroL* gene [46-49], its gene ID is 945031, (updated on 26-Nov-2012). The most important dissimilarity between the SKs is related to ligand-binding affinity. The K_m for shikimate is 20 mM for the SK I and 0.2 mM for the SK II enzyme [49]. It has been proposed that the SK II isoform is of pivotal importance in the SP, its expression is controlled by the *tyrR* regulator, and it is repressed by the presence of aromatic amino acids tyrosine and tryptophan [28, 48, 49]. On the other hand, it is not clear the physiological role of SK I in *E.coli*. In view of the fact that mutations in SK I are connected with sensitivity to the antibiotic Mecillinam [51], it has been proposed that SK I may work on a different pathway, unrelated to SP [28, 49]. As proposed by Parish and Stoker [1] and reviewed by Pereira and collaborators, if *M. tuberculosis* SK I has an analogous activity it is likely that disruption of this activity can be responsible for the observed inability of *M. tuberculosis* to grow in the absence of a functional copy of *aroK* gene [28].

The sixth enzyme of the SP is called EPSP synthase, and catalyzes the transfer of the enolpyruvyl moiety of phosphoenolpyruvate (PEP) to the 5-hydroxy position of shikimate-3-phosphate (S3P). The EPSP synthase (EC 2.5.1.19) is encoded by the *aroA* gene [28, 38, 52], its gene id for *Mycobacterium tuberculosis* is 888753 (updated on 12-Oct-2012).

The seventh and last step in the SP is the trans-1,4 elimination of phosphate from EPSP to yield chorismate. In this reaction, the second of the three double bonds of the benzene ring is introduced. The reaction is catalyzed by chorismate synthase (CS) (EC 4.2.3.5) and requires

reduced flavin for activity even though the overall reaction is redox neutral. In this respect, the enzyme is similar to DHQS, the second enzyme in the SP [36]. CS is encoded by the *aroC* gene [29], its gene ID in *Mycobacterium tuberculosis* is 925705 (updated on 11-Aug-2012).

As described above, bacteria have an isolated enzyme for each step of the SP and plants have a molecular assembly analogous to bacteria [53], with the exception of dehydroquinase (DHQase, third enzyme) and shikimate dehydrogenase (fourth enzyme) which have been shown to be present as separate domains on a bifunctional polypeptide chain [54]. On the other hand, fungi and apicomplexan organisms (*Toxoplasma gondii*) the SP have been shown to include monofunctional 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase and chorismate synthase (CS) enzymes and a pentafunctional polypeptide called *aroM*, which is responsible for the remaining five SP reactions [36, 55].

Crystallographic information about shikimate kinase

A search in the protein data bank (PDB) [56-58] for the keyword shikimate kinase resulted in 35 entries (December, 2012), as fully described in table 1. As we can see, most of the SK crystallographic structures were obtained from *Mycobacterium tuberculosis*, 20 out of 35 entries [45, 59-63]. Another interesting key aspect of the available SK structures; they were all solved by X-ray diffraction crystallography technique (molecular replacement method). Small proteins, with molecular weight lower than 25,000 such as SKs, are suitable to be submitted to automated nuclear magnetic resonance (NMR) structure determination [64]. Surprisingly, none SK structure has been obtained by NMR methods.

Besides SK from *Mycobacterium tuberculosis*, there are crystallographic structures of SK from *Erwinia chrysanthemi* (3 structures) [65, 66], *Escherichia coli* (1 structure) [67], *Campylobacter jejuni* (1 structure) (no primary citation for this crystallographic structure, PDB access code: 1VIA), *Helicobacter pylori* (6 structures) [68, 69], *Aquifex aeolicus* (1 structure) (to be published, PDB access code: 2PT5), *Arabidopsis thaliana* (1 structure) [70], *Coxiella burnetii* (1 structure) (to be published, PDB access code: 3TRF), and *Bacteroides thetaiotaomicron* (1 structure) (to be published, PDB access code: 3VAA). Analysis of the overall structural quality of the crystallographic information for SK, indicates resolution ranging from 1.35 to 2.9 Å.

Further analysis of the SK structural data, indicated that there are 28 structures of SK in complex with ligands. We considered a complexed SK structure, when at least one ligand was identified in the SK complexes by the program PoseView [71]. PoseView is fully integrated to the PDB page. It automatically creates comprehensive structural diagrams of molecular

binary complexes available at the PDB. A binary complex involves a protein structure and bound small molecules. A complex diagram shows the ligand, the amino acids of the protein interacting with the ligand and the intermolecular hydrogen bonds indicated as dashed lines between both molecules. The PoseView is based on a combinatorial optimization algorithm, which solves parts of the layout problem using non-heuristically approach. The molecules present in the binary complexes are shown as structural diagrams consistent with the chemical classification. Those molecules without intermolecular interactions identified by PoseView may be related to crystallization conditions, such as salts, glycerol, beta-mercaptoethanol, sulfate ions, phosphate ions. These chemicals are ordinarily used as crystallization additives in most of the protein crystallization experiments, as it has been described by Jancarik and Kim in 1991 [72].

*We only consider as ligands those with intermolecular interactions as shown by PoseView [71]. Those molecules without intermolecular interactions may be related to crystallization conditions, such as salts, glycerol, beta-mercaptoethanol, sulfate ions, phosphate ions.

Structure of shikimate kinase from *Mycobacterium tuberculosis*

The first X-ray crystallographic structure of SK complexed with ADP and shikimate was described by Pereira and collaborators and solved at 2.3 Å resolution using synchrotron radiation source [45]. The importance of this seminal work was recognized by Acta Crystallographica Section D editorial board, and the cover of December 2004 issue showed the first structure of SK from *Mycobacterium tuberculosis* in complex with shikimate and ADP. X-ray diffraction crystallographic studies SK from *Mycobacterium tuberculosis* [45, 59-63] revealed that the structure typically displays an alpha/beta-fold and shows five central parallel-strands flanked by alpha-helices, as shown in Fig. (2). This structural architecture is well conserved in all SKs solved so far.

The active site of SK can be divided in two binding pockets, shikimate-binding and ATP-binding pockets, as shown in Fig. (2). The continuous binding pockets present a volume of 197.625 Å³ as calculated by the program Molegro Virtual Docker (MVD) [73] and shown in Fig. (2). In order to compare, consider the ATP-binding pocket of human cyclin-dependent kinase 2 (CDK2) (PDB access code: 2A4L) [74-83]. This kinase presents a volume of 311.5 Å³, higher than the volume observed for SK, nevertheless the smaller SK protein is able to accommodate ATP and shikimate in its continuous binding pocket.

Shikimate-binding pocket

Careful analysis of the shikimate-binding pocket of complex SK-ADP-DHK in the 1WE2 structure has brought to light for the first time the unexpected details of the intermolecular interactions between shikimate and SK [45]. The shikimate-binding pocket follows strand beta2 and consists of helices alpha2 and 3 and the N-terminal region of helix alpha4 (residues ranging from 33 to 61). The nitrogen atoms from the side chains of the residues Arg 58 and Arg 136 and the NH backbone group of Gly 81 show intermolecular hydrogen bonds with the carboxyl group of shikimate, as shown in Fig. (3). The 3-hydroxyl group of shikimate forms intermolecular hydrogen bonds with the carboxyl group of Asp34 and the main-chain NH group of the residue Gly 80 (Fig. (3)). The 2-hydroxyl group of shikimate makes intermolecular hydrogen bonds to the side chain of Asp34.

We applied the program LIGPLOT [84] to analyze the relevant structural features present the SK from *Mycobacterium tuberculosis*. Most of the intermolecular hydrogen bonds at the shikimate-binding domain involve residues Asp 34, Arg 58, Gly 80, Gly 81, and Arg 136. Analysis of the van der Waals interactions between SK and ligands at this site indicated the participation of residues Arg 58, Gly 79, Gly 80, Gly 81, and Ala 82, as the major contact points with the ligands. Such structural features have been explored and confirmed in several molecular docking simulation studies. Molecular docking studies identified potential new SK inhibitors [85-98] such as [5-[[[(6S)-6-[3-[4-(trifluoromethoxy)phenyl]-1,2,4-oxadiazol-5-yl]-1,4,6,7-tetrahydroimidazo[5,4-d]py, 3-[(3S,5S)-5-[3-(1,3-benzodioxol-5-yl)-1,2,4-oxadiazol-5-yl]-1-methyl-pyrrolidin-3-yl]-1-isopropyl-u and staurosporine.

ATP-binding pocket

Analysis of the ATP-binding pocket of SK indicated a complex set of intermolecular interactions, as fully described by Pereira as collaborators [45]. The crystallographic structure of SK (PDB access code: 1WE2) revealed that the adenine moiety of ADP is sandwiched between residues Arg 110 and Pro 155, as shown in Fig. (4). The residue Arg 110 in SK is known to be the first residue of the LID domain observed for P-loop kinases [99]. In SK, the residue Arg 117 interacts with the alpha- and beta-phosphate groups. The residue Pro 155 is the last residue of the adenine-binding loop motif (residues ranging 148 to 155 in SK from *Mycobacterium tuberculosis*). This structural motif forms a loop that wraps around the adenine moiety of ATP, connecting the beta5-strand with the C-terminal alpha-helix. The second (aspartate) residue and the last (proline) residue are not conserved in the adenine-binding loops of *aroK*-encoded SKs from *Escherichia coli* [67] and *Campylobacter jejuni*

(PDB access code 1VIA). Further analysis of the structure, indicate that the carbonyl group of Arg 153 residue forms intermolecular hydrogen bonds to both the N6 atom of adenine and to a water molecule, which in turn hydrogen bonds to the N1 atom of adenine. Although both binding pockets can be equally accessed by molecular docking simulation, no previous in silico studies were focused on the ATP-binding pocket [85-98].

Molecular dynamics studies

One important feature for ligand recognition by SKs is that they undergo large conformational changes during catalysis. This structural rearrangement is dependent on an intricate pattern of intermolecular interactions between SK and shikimate and ATP [48]. These intermolecular interactions, such as hydrogen bonds, contact area and electrostatic interactions can be evaluated in complexes involving protein and ligand [100-109]. Nevertheless, precise analysis of protein-drug interactions from crystallographic structures is inadequate, since crystal structures are average structures obtained from the molecules packed in the crystal lattice. It is hard to identify directly from crystallographic structure the flexible parts of the molecule. Furthermore, the static view of the protein-ligand interaction is unrealistic since the proteins interact with ligands in a solvated environment and positioning of water molecules in crystallographic structures are limited to X-ray diffraction resolution parameters. One way to overcome this problem and obtain a more realistic view of protein-drug interactions comes from molecular dynamics (MD) simulations.

Three-dimensional structures obtained experimentally or by homology modeling can be submitted to MD simulations, where dynamical features of the complexes can be analyzed. Application of MD simulations to SK from *Mycobacterium tuberculosis* indicated large conformational changes due to ligand binding [28, 87, 89]. MD simulations of the structure of SK in complex with ADP and shikimate showed that the presence of the shikimate confers further stability to the structure of SK, allowing that the flexible loop that cover the shikimate binding site to be relatively stable during the simulation, when compared with the same region of the structure without shikimate.

Molecular docking simulations

Molecular docking simulations have been successfully applied to several different protein targets, for recent reviews see [93, 103, 110-118]. Most of the molecular docking simulations applied to identify new potential inhibitors for a protein target can be summarized in a four-

phase procedure, as has been previously described by Vianna & Azevedo [85]. This tandem process is called virtual screening (VS). Briefly, phase 1 of the VS is the validation phase where re-docking and cross-docking are carried out in order to identify the best docking protocol. In phase 2, a small-molecule dataset is chosen. In phase 3, the best molecular docking protocol, identified in phase 1, is used to screen the small-molecule database selected on phase 2. VS goal is to identify by molecular docking simulations new potential inhibitors for a protein target. Recently we applied this workflow to identify potential new inhibitors of SK [85]. Following identification of potential inhibitors, by the previously described VS protocol, the best scored ligands can be submitted to the web server FAF-Drugs [119], 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 may be filtered following the Lipinski' rules of five (RO5). RO5 advocates that drugs which present oral bioavailability, in general, follow: molecular weight less or equal to 500, LogP 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 [120].

The success of a VS protocol can be evaluated by its ability to enrich the quantity of identified inhibitors in the top ranks of a VS from among a larger number of decoy molecules in the screened database. Several researchers have assembled datasets of ligands and assumed decoys for several protein targets and used them to assess docking performance based on enrichment [121-125]. This database of decoys and known inhibitors is used to calculate the enrichment factor (EF), which is defined by the following equation,

$$EF = (Ha/Ht) / (A/N) ,$$

where Ha is the number of active compounds (known inhibitors) in the Ht top-ranked compounds of a total database of N compounds of which A are active [122]. It is expected in a successfully VS an $EF \gg 1$.

Molecular docking studies focused on the ATP-binding pocket

A recent study performed by our group identified potential new SK inhibitors [85] through molecular docking simulations focused on the shikimate-binding pocket of SK from *Mycobacterium tuberculosis*. In addition, several other studies have been conducted on the same binding pocket [86-98]. On the other hand, if we consider the overwhelming number of molecular docking simulation studies focused on ATP-binding pocket of human kinases, we

may benefit from this knowledge [80]. Especially if we take into account the specific small-molecule databases available for human kinases. In addition, several key features of protein-ligand interactions focused on human kinases have brought to light the common structural aspects necessary for binding to the ATP-binding pocket of human kinases. If we focus our attention on the most studied human kinase, the human CDK2, we could make a comparison of the structural determinants for ligand-binding specificity [74-83] between CDK2 and SK. Application of the same VS protocol previously described [85] to ATP-binding pocket of SK is able to identify potential SK inhibitors. In order to validate the docking protocol, we performed re-docking simulations (phase 1 of the VS protocol) using the crystallographic structure 2DFN [59], which resulted in an RMSD of 0.48 Å (scoring function: MOLDOCK Score Grid, search algorithm: MOLDOCK Optimizer). The re-docking simulations were restricted to a sphere with a radius of 10 Å, and center at X= -24.66, Y= -3.96, Z=15.93. Furthermore, cross-docking simulations generated RMSDs ranging from 0.48 to 2 Å, further validating the present docking protocol (PDB access codes used in the cross-docking: 2DFN, 1U8A, 1ZYU, 2G1K, 2IYQ, 2IYR, 2IYS, 2IYX, 2IYY, 2IYZ, 3BAF, 1WE2). These two tests indicated that the docking simulation was successful, and that the protocol is good enough to be used for the VS process.

To further validate the VS protocol, we used a database to calculate the EF, with 1083 decoys and four active molecules (N=1083). After application of the present VS protocol against this database, 11 molecules (Ht=11) were retrieved as hits (1.02 % of the database). Among these hits, one molecule was a known SK inhibitor (Ha=1). Thus, the enrichment factor was found to be 24.57, indicating that it is 24.57 times more likely to pick an active compound from the database than an inactive one. Previously published benchmarking sets for molecular docking of kinases exhibited EF ranging from 1.2 to 54 [122], indicating that the present VS protocol is adequate for VS purposes.

The VS simulations were carried out using the MVD program, having as target the SK (PDB access code 2DFN). The ligand library comprises 89,425 molecules obtained from ZINC database (Natural Products) [126]. After docking simulations, we selected 17 top-scoring compounds from the initial set of compounds (selection based on MOLDOCK score, ranging from -190.173 to -201.689). These 17 potential inhibitors were submitted to filter tests, available at the web server FAF-Drugs [119], to exclude those compounds that have known undesirable physical-chemical features to oral bioavailability, as established by Lipinski [120] and Veber rules [127]. After analysis, 5 compounds passed both tests and their structures are shown in Fig. (5).

Information about intermolecular interactions for all 5 top-scoring compounds is summarized in Table 2. Analysis of the ATP-binding site indicated that all top-scoring compounds present van der Waals contacts with residues Ile 18, Arg 110, Pro 155, and Val 158. Especially interesting is the intermolecular contacts with Arg 110 and Pro 155, as observed for the crystallographic structure 1WE2 at the ATP-binding pocket, all ligands are sandwiched between residues Arg110 and Pro155. This common structural feature indicates that this interaction is crucial for ligand-binding affinity. Furthermore, it has been identified that SK is inhibited by staurosporine [92], a well-known CDK2 inhibitor [78]. Based on this experimental evidence, we could expect to see common structural features between the ATP-binding pockets of CDK2 and SK. Considering the relative positioning of Arg 110 and Pro 155, as the major contact points for ligands at SK structure, we find a similar structural arrangement at ATP-binding pocket of CDK2. All inhibitors of CDK2, for which the crystallographic structure has been determined, indicate residues Ile 10 and Leu 134 as the major contact points, which form the majority of the van der Waals contacts in these complexes [78-80].

Analysis of the intermolecular hydrogen bonds in the docking results for SK at ATP-binding pocket shows that residues Gly 14, Lys 15, and Arg 117 participate in interactions in all top-score ligands. These residues have been shown to participate in hydrogen bonds with phosphate group of the ADP in the crystallographic structure 1WE2 [45]. Analysis of the intermolecular hydrogen bonds present in the crystallographic structures of CDK2 with different inhibitors, indicates the participation of main chain carbonyl group of Glu 81 with N3 of purine moiety as key point for ligand-binding affinity. Crystallographic structure of CDK2 in complex with the inhibitor deschloro-flavopiridol indicated the participation of main-chain carbonyl group of Glu 81 in an intermolecular hydrogen bond [74]. In the crystallographic structure 1WE2 the equivalent interaction is performed by the main chain carbonyl group of Arg 153. This hydrogen bond is observed only for the ligand ZINC 02160785, which suggests that further improvement of ligand-binding affinity could be explored by modification of the structures of others ligands, by adding a group to participate in a intermolecular hydrogen bond with Arg 154.

Final remarks

Computational approaches made possible to undertake larger VS studies focused on small-molecules libraries up to millions of compounds [128]. Here we reviewed the applications of these methodologies to SK. We also described the molecular docking study focused on the ATP-binding pocket of SK. Re-docking and cross-docking simulations generated RMSD

results below 2 Å, which indicate the ability of the protocol to recover the crystallographic position of a ligand. In addition, the VS protocol was able to identify potential new potential SK inhibitors. Furthermore, the present work indicates new molecules with the potential to become drugs against TB. Besides, we identified the SK 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 SK has the ability to bind a variety of molecular moieties not previously identified. Structural comparison of ATP-binding pockets of CDK2 and SK indicated some interesting similarities, such as a common structural feature that allows ligands to be sandwiched between two residues which are responsible for most of the van der Waals contacts observed in both structures. VS studies focused on ATP-binding pocket of SK will open the possibility to aggregate decades of structural studies focused on CDK2 to improve our understanding of the structural basis for inhibition of SKs.

Acknowledgments

This work was supported by National Institute of Science and Technology on Tuberculosis (Decit/SCTIE/MS-MCT-CNPq-FNDCTCAPES). W. F.A. Jr. is a research fellow of the National Council for Scientific and Technological Development of Brazil (CNPq). JCD acknowledges a scholarship awarded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (Brazil).

List of Abbreviations

ADP: Adenosine 5'-diphosphate
 ATP: Adenosine 5'-triphosphate
 CDK2: Cyclin-dependent kinase 2
 CS: chorismate synthase
 DAHP: 3-deoxy-D-arabinoheptulosonic acid-7- phosphate
 DAHPS: 3-deoxy-7-phosphoheptulonate synthase
 DHQS: 3-dehydroquininate synthase
 DHQ: 3-dehydroquininate
 DHQD: 3-dehydroquininate dehydratase
 E4P: erythrose-4- phosphate
 EF: Enrichment factor
 EPSPS: 5-enolpyruvyl-shikimate-3-phosphate synthase
 ID: Identification (Gene ID)

MD: Molecular dynamics
MVD: Molegro Virtual Docker
NMR: Nuclear magnetic resonance
PDB: Protein Data Bank
PEP: phosphoenolpyruvate
RMSD: Root mean square deviation
RO5: Lipinski's rule of five
S3P: shikimate-3-phosphate
SDH: shikimate-5-dehydrogenase
SK: Shikimate kinase
SP: Shikimate pathway
VS: Virtual screening

Disclosure: All the authors declared no competing interests.

References

- [1] Parish T, Stoker NG. The common aromatic amino acid biosynthesis pathway is essential in *Mycobacterium tuberculosis*. *Microbiology* 2002; 148(Pt 10): 3069-77.
- [2] **Ehrlich P.** Über den jetzigen Stand der Chemotherapie. *Ber Dtsch Chem Ges* 1909; **42: 17-47.**
- [3] Strebhardt K, Ullrich A. Paul Ehrlich's magic bullet concept: 100 years of progress. *Nat Rev Cancer* 2008; 8(6): 473-80.
- [4] Schwartz RS. Paul Ehrlich's magic bullets. *N Engl J Med* 2004; 350(11): 1079-80.
- [5] Steinrücken HC, Amrhein N. The herbicide glyphosate is a potent inhibitor of 5-enolpyruvyl-shikimic acid-3-phosphate synthase. *Biochem Biophys Res Commun* 1980; 94(4): 1207-12.
- [6] Séralini GE, Clair E, Mesnage R, *et al.* Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize. *Food Chem Toxicol* 2012; 50(11): 4221-31.
- [7] Ollivier, L. A Comment on "Séralini, G.-E, *et al.* Long term toxicity of a Roundup herbicide and a Roundup-tolerant genetically modified maize. *Food and Chemical Toxicology* 2012; pii: S0278-6915(12)00790-9.

- [8] Mesnage R, Bernay B, Séralini GE. Ethoxylated adjuvants of glyphosate-based herbicides are active principles of human cell toxicity. *Toxicology* 2012; pii: S0300-483X(12)00345-9.
- [9] Sribanditmongkol P, Jutavijittum P, Pongraveevongsa P, Wunnapuk K, Durongkadech P. Pathological and toxicological findings in glyphosate-surfactant herbicide fatality: a case report. *Am J Forensic Med Pathol* 2012; 33(3): 234-7.
- [10] Shiogiri NS, Paulino MG, Carraschi SP, *et al.* Acute exposure of a glyphosate-based herbicide affects the gills and liver of the Neotropical fish, *Piaractus mesopotamicus*. *Environ Toxicol Pharmacol* 2012; 34(2): 388-396.
- [11] Relyea RA. New effects of Roundup on amphibians: predators reduce herbicide mortality; herbicides induce antipredator morphology. *Ecol Appl* 2012; 22(2): 634-47.
- [12] Guilherme S, Santos MA, Barroso C, Gaivão I, Pacheco M. Differential genotoxicity of Roundup(®) formulation and its constituents in blood cells of fish (*Anguilla anguilla*): considerations on chemical interactions and DNA damaging mechanisms. *Ecotoxicology* 2012; 21(5): 1381-90.
- [13] Paetow LJ, Daniel McLaughlin J, Cue RI, Pauli BD, Marcogliese DJ. Effects of herbicides and the chytrid fungus *Batrachochytrium dendrobatidis* on the health of post-metamorphic northern leopard frogs (*Lithobates pipiens*). *Ecotoxicol Environ Saf* 2012; 80: 372-80.
- [14] Gui YX, Fan XN, Wang HM, Wang G, Chen SD. Glyphosate induced cell death through apoptotic and autophagic mechanisms. *Neurotoxicol Teratol* 2012; 34(3): 344-9.
- [15] Koller VJ, Fürhacker M, Nersesyan A, *et al.* Cytotoxic and DNA-damaging properties of glyphosate and Roundup in human-derived buccal epithelial cells. *Arch Toxicol* 2012; 86(5): 805-13.
- [16] Clair E, Mesnage R, Travert C, Séralini GÉ. A glyphosate-based herbicide induces necrosis and apoptosis in mature rat testicular cells in vitro, and testosterone decrease at lower levels. *Toxicol In Vitro* 2012; 26(2): 269-79.
- [17] Sprinson DB. The biosynthesis of aromatic compounds from D-glucose. *Adv Carbohydr Chem* 1960; 15: 235-70.
- [18] Davis BD, Mingioli ES. Aromatic biosynthesis. VII. Accumulation of two derivatives of shikimic acid by bacterial mutants. *J Bacteriol* 1953; 66(2): 129-36.
- [19] Pittard, AJ. In : F.C. Escherichia coli and *Salmonella typhimurium*: cellular and molecular biology. Neidhardt JL, Ingraham KB, Low B et al. Eds. Washington: American Society for Microbiology 1987; Vol. 1: pp 368-94.
- [20] Haslam E. Shikimic acid: metabolism and metabolites. Chichester: Wiley & Sons 1993.

- [21] Bentley R. The shikimate pathway—a metabolic tree with many branches. *Crit Rev Biochem Mol Biol* 1990; 25(5): 307-84.
- [22] McArthur JD, West NP, Cole JN, *et al.* An aromatic amino acid auxotrophic mutant of *Bordetella bronchiseptica* is attenuated and immunogenic in a mouse model of infection. *FEMS Microbiol Lett* 2003; 221(1): 7-16.
- [23] Roberts F, Roberts CW, Johnson JJ, *et al.* Evidence for the shikimate pathway in apicomplexan parasites. *Nature* 1998; 393: 801-5.
- [24] Kishore GM, Shah DM. Amino acid biosynthesis inhibitors as herbicides. *Annu Rev Biochem* 1988; 57: 627- 63.
- [25] McConkey GA, Pinney JW, Westhead DR, *et al.* Annotating the Plasmodium genome and the enigma of the shikimate pathway. *Trends Parasitol* 2004; 20(2): 60-5.
- [26] Tapas S, Kumar A, Dhindwal S, Preeti, Kumar P. Structural analysis of chorismate synthase from *Plasmodium falciparum*: a novel target for antimalaria drug discovery. *Int J Biol Macromol* 2011; 49(4): 767-77.
- [27] He QF, Li D, Xu QY, Zheng S. Predicted essential proteins of *Plasmodium falciparum* for potential drug targets. *Asian Pac J Trop Med* 2012; 5(5): 352-4.
- [28] Pereira JH, Vasconcelos IB, Oliveira JS, *et al.* Shikimate kinase: a potential target for development of novel antitubercular agents. *Curr Drug Targets* 2007; 8: 459-68.
- [29] Arcuri HA, Zafalon GFD, Marucci EA, *et al.* SKPDB: a structural database of shikimate pathway enzymes. *BMC Bioinformatics* 2010; 11: 12.
- [30] Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res* 2011; 39(Database issue):D32-7.
- [31] Smith LC, Ravel JM, Lax SR, Shive W. The control of 3-deoxy-D-arabino-heptulosonic acid 7-phosphate synthesis by phenylalanine and tyrosine. *J Biol Chem* 1962; 237: 3566-70.
- [32] Brown KD, Doy CH. End product regulation of the general aromatic pathway in *Escherichia coli* W. *Biochim Biophys Acta* 1963; 77: 170-2.
- [33] Doy CH, Brown KD. Control of aromatic biosynthesis: the multiplicity of 7-phospho-2-oxo-3 deoxy- D - arabino - heptonate - D - erythrose - 4 - phosphate - lyase (pyruvate phosphorylating) in *Escherichia coli* W. *Biochim Biophys Acta* 1965; 104: 377-89.
- [34] Wallace BJ, Pittard J. Genetic and biochemical analysis of the isoenzymes concerned in the first reaction of aromatic biosynthesis in *Escherichia coli*. *J Bacteriol* 1967; 93: 237-44.
- [35] Wallace BJ, Pittard J. Regulator gene controlling enzymes concerned in tyrosine biosynthesis in *Escherichia coli*. *J Bacteriol* 1969; 97(3): 1234-41.

- [36] Hermann KM, Weaver LM. The Shikimate Pathway. *Annu Rev Plant Mol Biol* 1999; 50(3): 473-503.
- [37] Wagner T, Shumilin IA, Bauerle R, Kretsinger RH. Structure of 3-Deoxy-D- arabino-heptulosonate-7-phosphate Synthase from *Escherichia coli*: Comparison of the Mn^{2+} -2-Phosphoglycolate and the Pb^{2+} -2-Phosphoenolpyruvate Complexes and Implications for Catalysis. *J Mol Biol* 2000; 301: 389-99.
- [38] Herrmann KM. The Shikimate Pathway: Early Steps in the Biosynthesis of Aromatic Compounds. *The Plant Cell* 1995; 7: 907-19.
- [39] Frost JW, Bender JL, Kadonaga JT, Knowles JR. Dehydroquinase synthase from *Escherichia coli*: purification, cloning, and construction of overproducers of the enzyme. *Biochemistry* 1984; 23(19): 4470-5.
- [40] Nasser D, Nester EW. Aromatic amino acid biosynthesis gene-enzyme relationships in *Bacillus subtilis*. *J Bacteriol* 1967; 94: 1706-14.
- [41] Basso LA, da Silva LHP, Fett-Neto AG, *et al.* The use of biodiversity as source of new chemical entities against defined molecular targets for treatment of malaria, tuberculosis, and T-cell mediated diseases - A Review. *Mem Inst Oswaldo Cruz* 2005; 100(6): 475-806.
- [42] Arcuri HA, Canduri F, Pereira JH, *et al.* Molecular Models for Shikimate Pathway Enzymes of *Xylella fastidiosa*. *Bioch Bioph Res Comm* 2004; 320: 979-91.
- [43] Roszak AW, Robinson DA, Krell T, *et al.* The structure and mechanism of the type II dehydroquinase from *Streptomyces coelicolor*. *Structure* 2002; 10(4): 493-503.
- [44] Benach J, Lee I, Edstrom W, *et al.* The 2.3 crystal of the shikimate 5-Dehydrogenase orthologue YdiB from *E.coli* suggests a novel catalytic environment for an NAD-dependent dehydrogenase. *J Biol Chem* 2003; 278: 19176-82.
- [45] Pereira JH, de Oliveira JS, Canduri F, *et al.* Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid. *Acta Crystallogr D Biol Crystallogr* 2004; 60: 2310-19.
- [46] Millar G, Lewendon A, Hunter MG, Coggins JR. The cloning and expression of the *aroL* gene from *Escherichia coli* K12. Purification and complete amino acid sequence of shikimate kinase II, the *aroL*-gene product. *Biochem J* 1986; 237(2): 427-37.
- [47] DeFeyter RC, Pittard J. Genetic and molecular analysis of *aroL*, the gene for shikimate kinase II in *Escherichia coli* K-12. *J Bacteriol* 1986; 165(1): 226-32.
- [48] DeFeyter RC, Davidson BE, Pittard J. Nucleotide sequence of the transcription unit containing the *aroL* and *aroM* genes from *Escherichia coli* K-12. *J Bacteriol* 1986; 165(1): 233-9.

- [49] DeFeyter RC, Pittard J. Purification and properties of shikimate kinase II from *Escherichia coli* K-12. *J Bacteriol* 1986; 165(1): 331-3.
- [50] Ely B, Pittard J. Aromatic amino acid biosynthesis: regulation of shikimate kinase in *Escherichia coli* K-12. *J Bacteriol* 1979; 138(3): 933-43.
- [51] Vinella D, Gagny B, Joseleau-Petit D, D'Ari R, Cashel M. Mecillinam resistance in *Escherichia coli* is conferred by loss of a second activity of the AroK protein. *J Bacteriol* 1996; 178(13): 3818-28.
- [52] Rogers SG, Brand LA, Holder SB, Sharps ES, Brackin MJ. Amplification of the *aroA* gene from *Escherichia coli* results in tolerance to the herbicide glyphosate. *Appl Environ Microbiol* 1983; 46(1): 37-43.
- [53] Butler JR, Alworth WL, Nugent MJ. Mechanism of dehydroquinase catalyzed dehydration. I. Formation of a Schiff base intermediate. *J Am Chem Soc* 1974; 96: 1617-8.
- [54] Mousdale DM, Campbell MS, Coggins JR. Purification and characterisation of a bi-functional dehydroquinase-shikimate: NADP oxidoreductase from pea seedlings. *Phytochemistry* 1987; 26: 2665-70.
- [55] Arcuri HA, Palma MS. Understanding the structure, activity and inhibition of chorismate synthase from *Mycobacterium tuberculosis*. *Curr Med Chem* 2011; 18(9): 1311-7.
- [56] Berman HM, Westbrook J, Feng Z, *et al.* The Protein Data Bank. *Nucleic Acids Res* 2000; 28: 235-42.
- [57] Berman HM, Battistuz T, Bhat TN, *et al.* The Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* 2002; 58(Pt 6 No 1): 899-907.
- [58] Westbrook J, Feng Z, Chen L, Yang H, Berman HM. The Protein Data Bank and structural genomics. *Nucleic Acids Res* 2003; 31(1): 489-91.
- [59] Dias MV, Faím LM, Vasconcelos IB, *et al.* Effects of the magnesium and chloride ions and shikimate on the structure of shikimate kinase from *Mycobacterium tuberculosis*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2007; 63: 1-6.
- [60] Hartmann MD, Bourenkov GP, Oberschall A, Strizhov N, Bartunik HD. Mechanism of phosphoryl transfer catalyzed by shikimate kinase from *Mycobacterium tuberculosis*. *J Mol Biol* 2006; 364: 411-23.
- [61] Gan J, Gu Y, Li Y, Yan H, Ji X. Crystal structure of *Mycobacterium tuberculosis* shikimate kinase in complex with shikimic acid and an ATP analogue. *Biochemistry* 2006; 45: 8539-45.

- [62] Dhaliwal B, Nichols CE, Ren J, *et al.* Crystallographic studies of shikimate binding and induced conformational changes in *Mycobacterium tuberculosis* shikimate kinase. *FEBS Lett* 2004; 574: 49-54.
- [63] Gu Y, Reshetnikova L, Li Y, *et al.* Crystal structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the dynamic role of the LID domain in catalysis. *J Mol Biol* 2002; 319: 779-89.
- [64] Fadel V, Bettendorff P, Herrmann T, *et al.* Automated NMR structure determination and disulfide bond identification of the myotoxin crostamine from *Crotalus durissus terrificus*. *Toxicon* 2005; 46(7): 759-67.
- [65] Krell T, Maclean J, Boam DJ, *et al.* Biochemical and X-ray crystallographic studies on shikimate kinase: the important structural role of the P-loop lysine. *Protein Sci* 2001; 10: 1137-49.
- [66] Krell T, Coyle JE, Horsburgh MJ, Coggins JR, Laphorn AJ. Crystallization and preliminary X-ray crystallographic analysis of shikimate kinase from *Erwinia chrysanthemi*. *Acta Crystallogr SectD* 1997; 53: 612-14.
- [67] Romanowski MJ, Burley SK. Crystal structure of the *Escherichia coli* shikimate kinase I (AroK) that confers sensitivity to mecillinam. *Proteins* 2002; 47: 558-62.
- [68] Cheng WC, Chen YF, Wang HJ, *et al.* Structures of *Helicobacter pylori* shikimate kinase reveal a selective inhibitor-induced-fit mechanism. *PLoS One* 2012; 7: e33481- e33481.
- [69] Cheng WC, Chang YN, Wang WC. Structural basis for shikimate-binding specificity of *Helicobacter pylori* shikimate kinase. *J Bacteriol* 2005; 187: 8156-63.
- [70] Fucile G, Garcia C, Carlsson J, Sunnerhagen M, Christendat D. Structural and biochemical investigation of two Arabidopsis shikimate kinases: the heat-inducible isoform is thermostable. *Protein Sci* 2011; 20: 1125-36.
- [71] Stierand K, Maass PC, Rarey M. Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. *Bioinformatics* 2006; 22(14): 1710-6.
- [72] Jancarik J, Kim S.-H. Sparse matrix sampling: a screening method for crystallization of proteins. *J Appl Cryst* 1991; 24: 409-411
- [73] Thomsen R, Christensen MH. MolDock: a new technique for high accuracy molecular docking. *J Med Chem* 2006; 49(11): 3315-21.
- [74] De Azevedo WF Jr, Mueller-Dieckmann HJ, Schulze-Gahmen U, *et al.* Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase. *Proc Natl Acad Sci USA* 1996; 93(7): 2735-40.

- [75] De Azevedo WF, Leclerc S, Meijer L, *et al.* Inhibition of cyclin-dependent kinases by purine analogues: crystal structure of human cdk2 complexed with roscovitine. *Eur J Biochem* 1997; 243(1-2): 518-26.
- [76] Krystof V, Uldrijan S. Cyclin-dependent kinase inhibitors as anticancer drugs. *Curr Drug Targets* 2010; 11(3): 291-302.
- [77] Hernandez MZ, Cavalcanti SM, Moreira DR, de Azevedo Jr. WF, Leite AC. Halogen atoms in the modern medicinal chemistry: hints for the drug design. *Curr Drug Targets* 2010; 11(3): 303-14.
- [78] Canduri F, Perez PC, Caceres RA, de Azevedo WF Jr. CDK9 a Potential Target for Drug Development. *Med Chem* 2008; 4(3): 210-18.
- [79] Krystof V, Cankar P, Hajdich M, *et al.* 4-Arylazo-3,5-diamino-1H-pyrazole CDK Inhibitors: SAR Study, Crystal Structure in Complex with CDK2, Selectivity, and Cellular Effects. *J Med Chem* 2006; 49(22): 6500-9.
- [80] Canduri F, de Azevedo WF Jr. Structural basis for interaction of inhibitors with Cyclin-Dependent Kinase 2. *Curr Computer-Aided Drug Design* 2005; 1(1): 53-64.
- [81] De Azevedo WF Jr, Canduri F, da Silveira NJ. Structural basis for inhibition of Cyclin-Dependent Kinase 9 by flavopiridol. *Biochem Biophys Res Commun* 2002; 293(1): 566-71.
- [82] Filgueira de Azevedo W Jr, Gaspar RT, Canduri F, Camera JC Jr, Freitas da Silveira NJ. Molecular model of cyclin-dependent kinase 5 complexed with roscovitine. *Biochem Biophys Res Commun* 2002; 297(5): 1154-8.
- [83] Canduri F, Uchoa HB, de Azevedo WF Jr. Molecular models of cyclin-dependent kinase 1 complexed with inhibitors. *Biochem Biophys Res Commun* 2004; 324(2): 661-6.
- [84] Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 1995; 8(2): 127-34.
- [85] Vianna CP, de Azevedo WF Jr. Identification of new potential *Mycobacterium tuberculosis* shikimate kinase inhibitors through molecular docking simulations. *J Mol Model* 2012; 18(2): 755-64.
- [86] Segura-Cabrera A, Rodríguez-Pérez MA. Structure-based prediction of *Mycobacterium tuberculosis* shikimate kinase inhibitors by high-throughput virtual screening. *Bioorg Med Chem Lett* 2008; 18: 3152-7.
- [87] Saidenberg DM, Passarelli AW, Rodrigues AV, *et al.* Shikimate kinase (EC 2.7.1.71) from *Mycobacterium tuberculosis*: kinetics and structural dynamics of a potential molecular target for drug development. *Curr Med Chem* 2011; 18: 1299-310.

- [88] Hsu KC, Cheng WC, Chen YF, *et al.* Core site-moiety maps reveal inhibitors and binding mechanisms of orthologous proteins by screening compound libraries. *PLoS One* 2012; 7: e32142.
- [89] De Azevedo WF Jr. Molecular dynamics simulations of protein targets identified in *Mycobacterium tuberculosis*. *Curr Med Chem* 2011; 18: 1353-66.
- [90] Kumar M, Verma S, Sharma S, *et al.* Structure-based in silico design of a high-affinity dipeptide inhibitor for novel protein drug target Shikimate kinase of *Mycobacterium tuberculosis*. *Chem Biol Drug Des* 2010; 76: 277-84.
- [91] Arora N, Banerjee AK, Murty US. In silico characterization of Shikimate Kinase of *Shigella flexneri*: a potential drug target. *Interdiscip Sci* 2010; 2(3): 280-90.
- [92] Mulabagal V, Calderón AI. 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* 2010; 82(9): 3616-21.
- [93] De Azevedo WF Jr. MolDock applied to structure-based virtual screening. *Curr Drug Targets* 2010; 11(3): 327-34.
- [94] Habig M, Blechschmidt A, Dressler S, *et al.* Efficient elimination of nonstoichiometric enzyme inhibitors from HTS hit lists. *J Biomol Screen* 2009; 14(6): 679-89.
- [95] Pauli I, Caceres RA, de Azevedo WF Jr. Molecular modeling and dynamics studies of Shikimate Kinase from *Bacillus anthracis*. *Bioorg Med Chem* 2008; 16(17): 8098-108.
- [96] Han C, Zhang J, Chen L, *et al.* Discovery of *Helicobacter pylori* shikimate kinase inhibitors: bioassay and molecular modeling. *Bioorg Med Chem* 2007; 15(2): 656-62.
- [97] Silveira NJ, Uchôa HB, Pereira JH, *et al.* Molecular models of protein targets from *Mycobacterium tuberculosis*. *J Mol Model* 2005; 11(2): 160-6.
- [98] Filgueira de Azevedo W Jr, Canduri F, Simões de Oliveira J, *et al.* Molecular model of shikimate kinase from *Mycobacterium tuberculosis*. *Biochem Biophys Res Commun* 2002; 295(1): 142-48.
- [99] Leipe DD, Koonin EV, Aravind L. Evolution and classification of P-loop kinases and related proteins. *J Mol Biol* 2003; 333(4): 781-815.
- [100] Timmers LF, Caceres RA, Vivan AL, *et al.* Structural studies of human purine nucleoside phosphorylase: Towards a new specific empirical scoring function. *Arch Biochem Biophys* 2008; 479: 28-38.
- [101] De Azevedo WF Jr. Protein-drug interactions. *Curr Drug Targets* 2008; 9(12): 1030.

- [102] De Azevedo WF Jr, Dias R. Computational methods for calculation ligand-binding affinity. *Curr Drug Targets* 2008; 9(12): 1031-9.
- [103] Dias R, de Azevedo WF Jr. Molecular docking algorithms. *Curr Drug Targets* 2008; 9(12): 1040-7.
- [104] Dias R, Timmers LFSM, Caceres RA, de Azevedo WF Jr. Evaluation of molecular docking using polynomial empirical scoring functions. *Curr Drug Targets* 2008; 9(12): 1062-70.
- [105] De Azevedo WF Jr, Dias R. Experimental approaches to evaluate the thermodynamics of protein-drug interactions. *Curr Drug Targets* 2008; 9(12): 1071-6.
- [106] Caceres RA, Pauli I, Timmers LFS, de Azevedo WF Jr. Molecular recognition models: A challenge to overcome. *Curr Drug Targets* 2008; 9(12): 1077-83.
- [107] De Azevedo WF Jr, Caceres RA, Pauli I, *et al.* Protein-drug interaction studies for development of drugs against *Plasmodium falciparum*. *Curr Drug Targets* 2009; 10(3): 271-8.
- [108] Ducati RG, Basso LA, Santos DS, de Azevedo WF Jr. Crystallographic and docking studies of purine nucleoside phosphorylase from *Mycobacterium tuberculosis*. *Bioorg Med Chem* 2010; 18(13): 4769-74.
- [109] De Azevedo WF Jr, Dias R. Evaluation of ligand-binding affinity using polynomial empirical scoring functions. *Bioorg Med Chem* 2008; 16(20): 9378-83.
- [110] Heberlé G, de Azevedo WF Jr. Bio-Inspired Algorithms Applied to Molecular Docking Simulations. *Curr Med Chem* 2011; 18(9): 1339-52.
- [111] Bellows ML, Floudas CA. Computational methods for de novo protein design and its applications to the human immunodeficiency virus 1, purine nucleoside phosphorylase, ubiquitin specific protease 7, and histone demethylases. *Curr Drug Targets* 2010; 11(3): 264-78.
- [112] Mitrasinovic PM. Advances in the structure-based design of the influenza A neuraminidase inhibitors. *Curr Drug Targets* 2010; 11(3): 315-26.
- [113] Kim D, Lee YH, Hwang HY, Kim KK, Park HJ. Z-DNA binding proteins as targets for structure-based virtual screening. *Curr Drug Targets* 2010; 11(3): 335-44.
- [114] De Azevedo WF Jr. Structure-based virtual screening. *Curr Drug Targets* 2010; 11(3): 261-3.
- [115] De Azevedo WF Jr, Dias R, Timmers LF, *et al.* Bioinformatics tools for screening of antiparasitic drugs. *Curr Drug Targets* 2009; 10(3): 232-9.
- [116] De Azevedo WF Jr, Caceres RA, Pauli I, *et al.* Protein-drug interaction studies for development of drugs against *Plasmodium falciparum*. *Curr Drug Targets* 2009; 10(3): 271-8.

- [117] Canduri F, de Azevedo WF. Protein crystallography in drug discovery. *Curr Drug Targets* 2008; 9(12): 1048-53.
- [118] Pauli I, Timmers LF, Caceres RA, Soares MB, de Azevedo WF Jr. *In silico* and *in vitro*: identifying new drugs. *Curr Drug Targets* 2008; 9(12): 1054-61.
- [119] Miteva MA, Viola S, Montes M, *et al.* FAF-drugs: free adme/tox filtering of compound collections. *Nucleic Acids Res* 2006; 34: W738-W744.
- [120] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 2001; 46(1-3): 3-26.
- [121] Huang N, Shoichet BK, Irwin JJ. Benchmarking sets for molecular docking. *J Med Chem* 2006; 49(23): 6789-801.
- [122] Brooijmans N, Sharp KA, Kuntz ID. Molecular recognition and docking algorithms. *Annu Rev Biophys Biomol Struct* 2003; 32: 335-73.
- [123] Bissantz C, Folkers G, Rognan D. Protein-based virtual screening of chemical databases: 1. Evaluation of different docking/scoring combinations. *J Med Chem* 2000; 43: 4759-67.
- [124] Kellenberger E, Rodrigo J, Muller P, Rognan D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* 2004; 57: 225-42.
- [125] Pham TA, Jain AN. Parameter estimation for scoring protein-ligand interactions using negative training data. *J Med Chem* 2006; 49(20): 5856-68.
- [126] Irwin JJ, Shoichet BK. ZINC--a free database of commercially available compounds for virtual screening. *J Chem Inf Model* 2005; 45(1): 177-82.
- [127] Veber DF, Johnson SR, Cheng HY, *et al.* Molecular properties that influence the oral bioavailability of drug candidates. *J Med Chem* 2002; 45(12): 2615-23.
- [128] Azevedo LS, Moraes FP, Xavier MM, *et al.* Recent Progress ff Molecular Docking Simulations Applied to Development of Drugs. *Curr Bioinformatics* 2012; 7: 352-65.

Figure Legends

Fig. 1 Sequence of seven metabolic steps in the SP, from phosphoenolpyruvate and erythrose 4-phosphate to chorismate. Attached to each step is displayed a representative crystallographic structure of the enzyme responsible to the catalysis of that step

Fig. 2 Crystallographic structure of SK from *Mycobacterium tuberculosis*.

Fig. 3 Shikimate-binding pocket of SK from *Mycobacterium tuberculosis*.

Fig. 4 ATP-binding pocket of SK from *Mycobacterium tuberculosis*.

Fig. 5 Molecular structures of 5 compounds identified as potential SK inhibitors.

Table captions

Table 1 List of available crystallographic structures for SK.

Table 2 Intermolecular interactions for top-scoring ligands. The X indicates the presence of the interaction. HB and VDW indicate hydrogen bonds and van der Waals contacts, respectively.

Figure 1

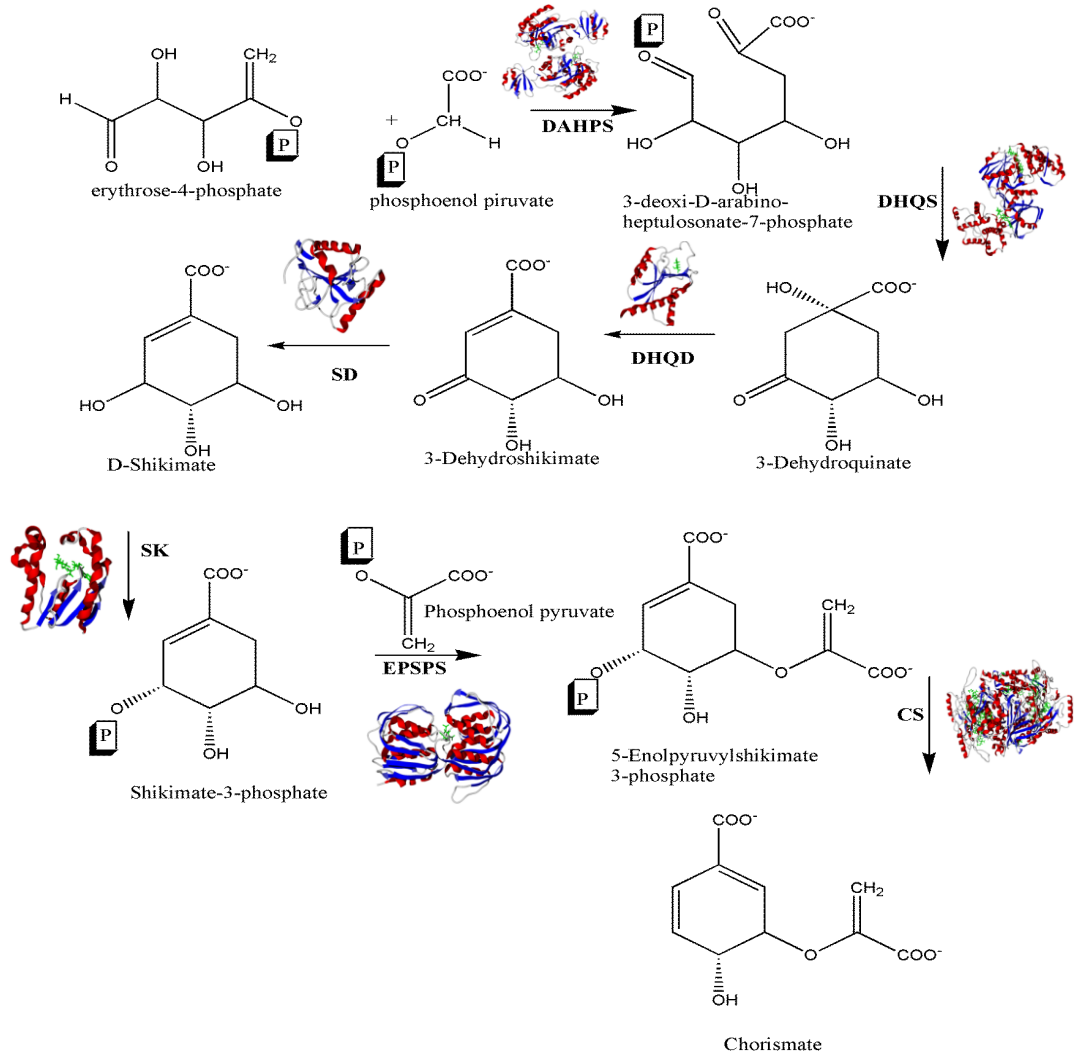


Figure 2

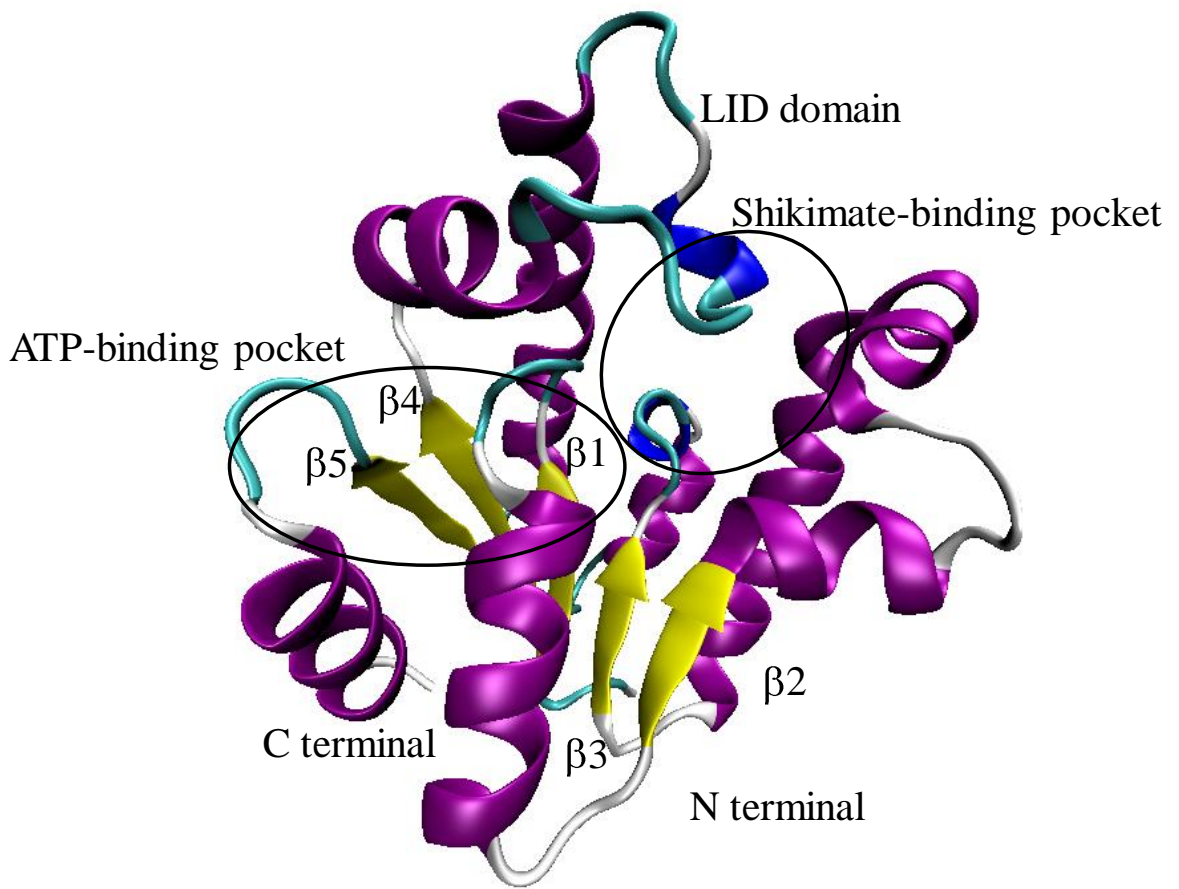


Figure 3

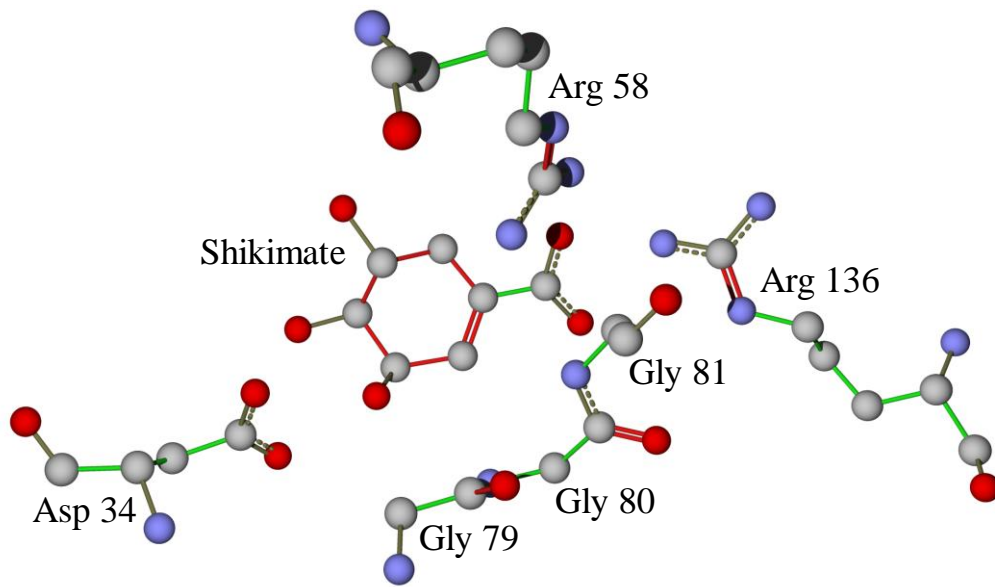


Figure 4

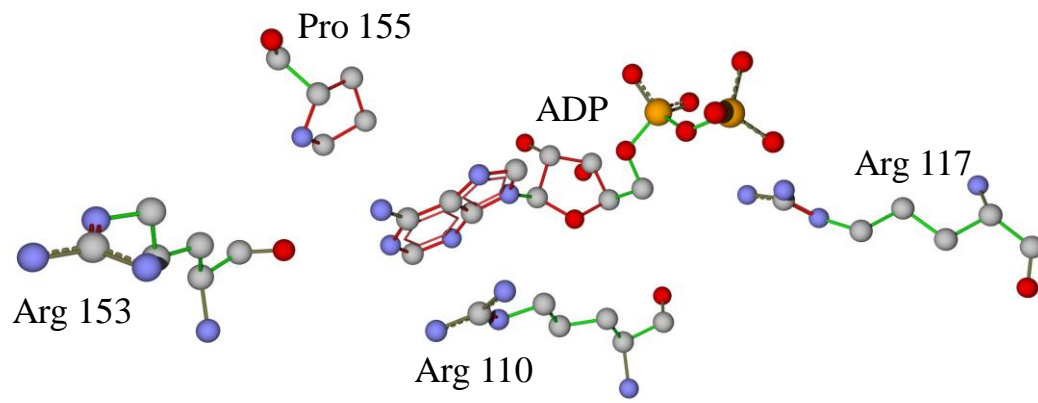
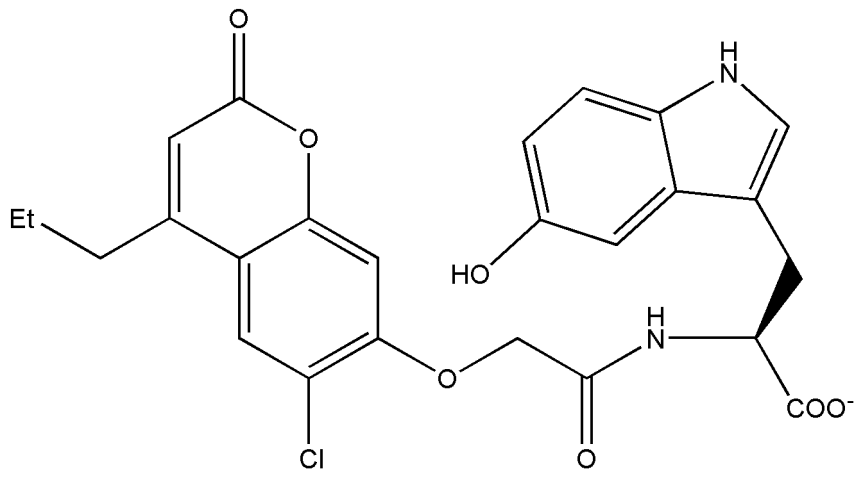
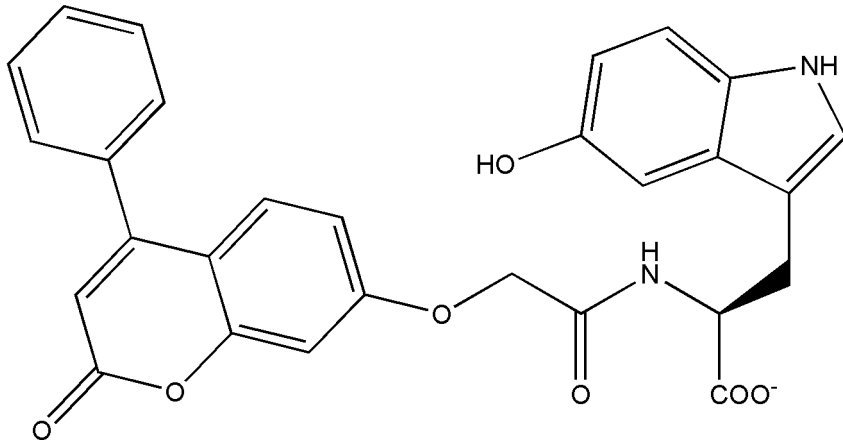


Figure 5A



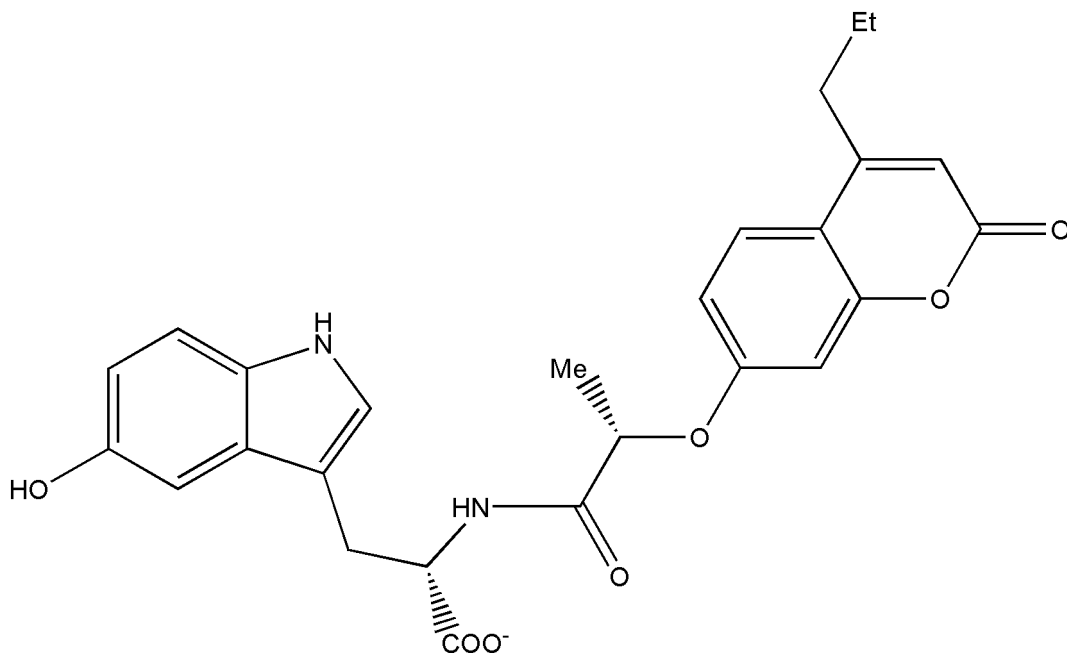
(ZINC 02135897)

Figure 5B



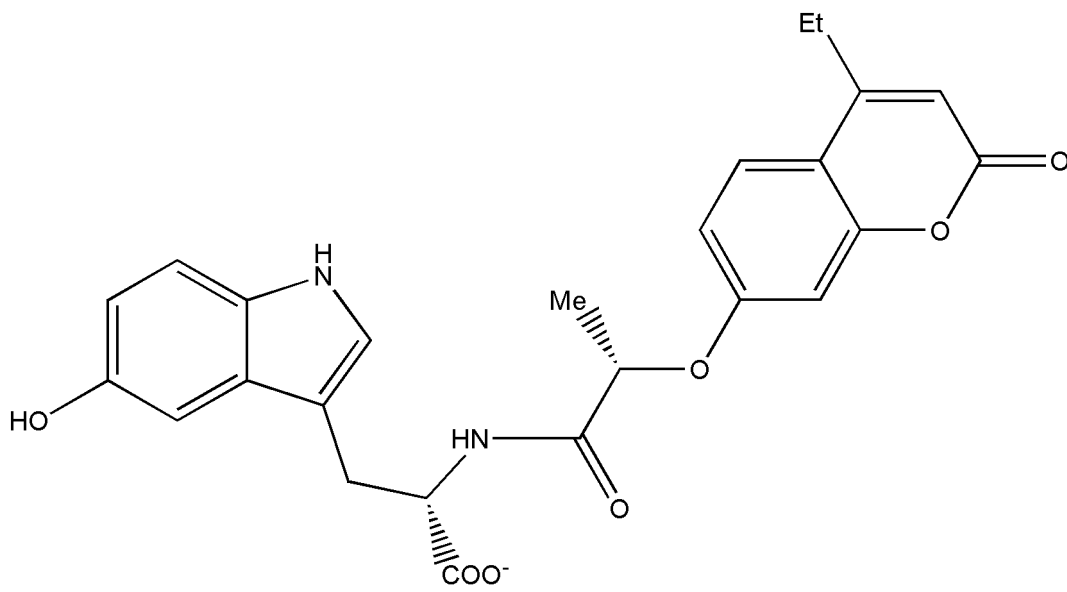
(ZINC 02127309)

Figure 5C



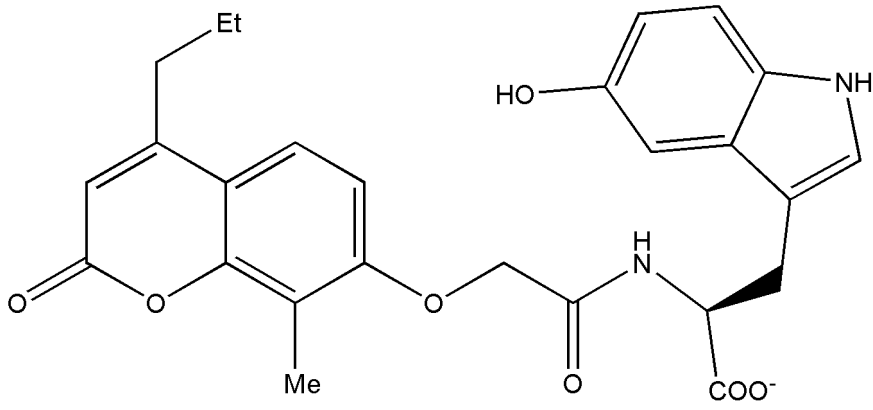
(ZINC 02127309)

Figure 5D



(ZINC02130347)

Figure 5E



(ZINC 02160785)

Table 1

PDB access codes	Organisms	Resolution (Å)	Deposition date (Year-Month-Day)	Ligands in the active site*	References
3VAA	<i>Bacteroides thetaiotaomicron</i>	1.7	2011-12-29	di(hydroxyethyl)ether	To be published
3TRF	<i>Coxiella burnetii</i>	2.6	2011-09-09	None	To be published
3NWJ	<i>Arabidopsis thaliana</i>	2.35	2010-07-09	None	[70]
3N2E	<i>Helicobacter pylori</i> 2669	2.53	2010-05-18	7-amino-4-hydroxy-3-[(E)-(5-hydroxy-7-sulfonaphthalen-2-yl)diazenyl]naphthalene-2-sulfonic acid and L(+)-Tartaric acid	[68]
3MRS	<i>Helicobacter pylori</i> 2669	2.40	2010-04-29	None	[68]
3MUF	<i>Helicobacter pylori</i>	2.30	2010-05-03	Adenosine-5'-diphosphate and shikimate-3-phosphate	[68]
3HR7	<i>Helicobacter pylori</i> 2669	1.80	2009-06-09	None	[68]
3BAF	<i>Mycobacterium tuberculosis</i>	2.25	2007-11-07	Phosphoaminophosphonic acid-adenylate ester and (3R, 4S, 5R)-3,4,5-Trihydroxycyclohex-1-ene-1-carboxylic acid	To be published
2PT5	<i>Aquifex aeolicus</i> VF5	2.10	2007-05-08	None	To be published
2DFT	<i>Mycobacterium tuberculosis</i>	2.8	2006-03-03	Adenosine-5'-diphosphate	[59]
2IYW	<i>Mycobacterium tuberculosis</i>	1.85	2006-07-22	Adenosine-5'-triphosphate	[60]
2IYQ	<i>Mycobacterium tuberculosis</i>	1.80	2006-07-21	Adenosine-5'-diphosphate And (3R, 4S, 5R)-3,4,5-Trihydroxycyclohex-1-ene-1-carboxylic acid	[60]
2IYR	<i>Mycobacterium tuberculosis</i>	1.98	2006-07-21	(3R, 4S, 5R)-3,4,5-Trihydroxycyclohex-1-ene-1-carboxylic acid	[60]
2IYS	<i>Mycobacterium tuberculosis</i>	1.40	2006-07-21	(3R, 4S, 5R)-3,4,5-Trihydroxycyclohex-1-ene-1-carboxylic acid	[60]
2IYT	<i>Mycobacterium tuberculosis</i>	1.47	2006-07-22	None	[60]
2IYU	<i>Mycobacterium tuberculosis</i>	1.85	2006-07-22	Adenosine-5'-diphosphate	[60]
2IYV	<i>Mycobacterium tuberculosis</i>	1.35	2006-07-22	Adenosine-5'-diphosphate	[60]
2IYX	<i>Mycobacterium</i>	1.49	2006-07-	(3R, 4S, 5R)-3,4,5-	[60]

	<i>tuberculosis</i>		22	Trihydroxycyclohex-1-ene-1-carboxylic acid	
2IYY	<i>Mycobacterium tuberculosis</i>	1.62	2006-07-22	Shikimate-3-phosphate	[60]
2IYZ	<i>Mycobacterium tuberculosis</i>	2.30	2006-07-22	Adenosine-5'-diphosphate and Shikimate-3-phosphate	[60]
2DFN	<i>Mycobacterium tuberculosis</i>	1.93	2006-03-03	Adenosine-5'-diphosphate And (3R, 4S, 5R)-3,4,5-Trihydroxycyclohex-1-ene-1-carboxylic acid	[59]
2G1J	<i>Mycobacterium tuberculosis</i>	2.00	2006-02-14	None	[61]
2G1K	<i>Mycobacterium tuberculosis</i>	1.75	2006-02-14	(3R, 4S, 5R)-3,4,5-Trihydroxycyclohex-1-ene-1-carboxylic acid	[61]
1ZYU	<i>Mycobacterium tuberculosis</i>	2.90	2005-06-11	Phosphomethylphosphonic acid adenylate ester And (3R, 4S, 5R)-3,4,5-Trihydroxycyclohex-1-ene-1-carboxylic acid	[61]
1ZUH	<i>Helicobacter pylori 2669</i>	1.80	2005-05-31	None	[69]
1ZUI	<i>Helicobacter pylori 2669</i>	2.30	2005-05-31	(3R, 4S, 5R)-3,4,5-Trihydroxycyclohex-1-ene-1-carboxylic acid	[69]
1WE2	<i>Mycobacterium tuberculosis</i>	2.30	2004-05-22	Adenosine-5'-diphosphate And 3-dehydroshikimate	[45]
1U8A	<i>Mycobacterium tuberculosis</i>	2.15	2004-08-05	Adenosine-5'-diphosphate and (3R, 4S, 5R)-3,4,5-Trihydroxycyclohex-1-ene-1-carboxylic acid	[62]
1VIA	<i>Campylobacter jejun</i>	1.57	2003-12-01	None	No Primary Citation for this Structure
1L4U	<i>Mycobacterium tuberculosis</i>	1.80	2002-03-05	Adenosine-5'-diphosphate And HEPES	[63]
1L4Y	<i>Mycobacterium tuberculosis</i>	2.00	2002-03-06	Adenosine-5'-diphosphate	[63]
1KAG	<i>Escherichia coli</i>	2.05	2001-11-01	None	[67]
1E6C	<i>Erwinia chrysanthemi</i>	1.80	2000-08-10	(4R)-2-Methylpentane-2,4-diol	[65]
1SHK	<i>Erwinia chrysanthemi</i>	1.90	1997-10-27	None	[66]
2SHK	<i>Erwinia chrysanthemi</i>	2.60	1997-10-27	Adenosine-5'-diphosphate	[66]

Table 2

SK					
Ligands (ZINC codes)					
HB	ZINC	ZINC	ZINC	ZINC	ZINC
	02135897	02127309	02133638	02130347	02160785
Thr 17	X	X	X	X	
Gly 12		X	X		X
Lys 15	X	X	X	X	X
Ser 16		X			
Arg 117	X	X	X	X	X
Thr 115		X			
Asp 34	X		X	X	X
Ser 13	X	X		X	
Gly 14	X	X	X	X	X
Arg 153					X
VDW					
Asn 154	X		X	X	X
Pro 155	X	X	X	X	X
Thr 17					X
Val 158	X	X	X	X	X
Arg 110	X	X	X	X	X
Ile 18	X	X	X	X	X
Ser 13					X
Gly 12	X			X	
Ser 16	X		X	X	X
Thr 150	X		X	X	

De: Current Drug Targets [cdt@benthamscience.org]
Enviado: quarta-feira, 12 de dezembro de 2012 16:24
Para: Walter Filgueira de A Junior
Assunto: [CDT] Submission Acknowledgement

Dear Dr. Dr. Walter F De AZEVEDO Jr.:

Thank you for submitting your manuscript to Current Drug Targets which we have received and will forward to the editorial board and external reviewers for peer reviewing.

As soon as we receive any feedback from the reviewers, we will get back to you with their comments, which may facilitate you with the revision of the article, if required. Before proceeding for peer-review, the editorial office will shortly contact you for further details and requirements.

Best regards,

Editorial Office
Bentham Science Publishers
Current Drug Targets
<http://bsp-cms.eurekaselect.com/index.php/CDT>