



Universidade Federal do Rio Grande

Programa de Pós-Graduação em Ciências da Saúde - Doutorado

Núcleo de Pesquisa em Microbiologia Médica - NUPEMM



**BIOINFORMÁTICA COMO FERRAMENTA PARA O ESTUDO DO
MECANISMO DE EFLUXO DA BOMBA MULTIDROGA AcrB**

LANDE VIEIRA DA SILVA JÚNIOR

Rio Grande – RS

Abril/ 2016



Universidade Federal do Rio Grande

Programa de Pós-Graduação em Ciências da Saúde - Doutorado

Núcleo de Pesquisa em Microbiologia Médica - NUPEMM



BIOINFORMÁTICA COMO FERRAMENTA PARA O ESTUDO DO MECANISMO DE EFLUXO DA BOMBA MULTIDROGA AcrB

Tese submetida ao Programa de Pós-Graduação em Ciências da Saúde da Universidade Federal do Rio Grande, como parte dos requisitos necessários para a obtenção do Título de Doutor em Ciências da Saúde.

Orientador: Dr. Pedro Eduardo Almeida da Silva

Rio Grande-RS

Abril/2016

Lande Vieira da Silva Júnior

Tese submetida ao Programa de Pós-Graduação em Ciências da Saúde da Universidade Federal do Rio Grande, como parte dos requisitos necessários para a obtenção do Título de Doutor em Ciências da Saúde.

BIOINFORMÁTICA COMO FERRAMENTA PARA O ESTUDO DO MECANISMO DE EFLUXO DA BOMBA MULTIDROGA AcrB

Banca examinadora:

Dr. Adriano Velasque Werhli (FURG)

Dra. Daniela Ramos (FURG)

Dr. Hermes Luís Neubauer de Amorim - Membro externo (ULBRA)

Dr. Andréa von Groll (FURG – suplente)

Dr. Pedro Eduardo Almeida da Silva (Orientador e Presidente)

**Dedico esta tese aos meus pais Olga e Lande (*in memoriam*), minhas irmãs,
Gisele e Janine, meus filhos Victor e Isadora e minha companheira
Fernanda.**

AGRADECIMENTOS

Agradeço ao Professor Doutor Pedro Eduardo Almeida da Silva por me acolher em seu grupo de pesquisa, pelas inúmeras horas de discussões sobre a proteína de efluxo AcrB, por permitir aplicar conhecimentos da física no impressionante campo da microbiologia, e principalmente pela sua visão de vanguarda de que a ciência deve ser praticada de maneira interdisciplinar.

A professora Doutora Karina Machado, pelas aulas, conselhos e pela co-orientação desse trabalho.

Ao professor PhD Terry P. Lybrand, meu supervisor durante o doutorado sanduíche em Nashville, Tennessee, EUA. Muito obrigado por ter me recebido em seu laboratório e pela oportunidade que me deu de conhecer um pouco mais sobre o vasto campo da biologia computacional.

Ao professor PhD Jonathan Sheeham, muito obrigado pelos ensinamentos, pela paciência e pela companhia nos almoços na Panera Bread junto com o amigo Chris Moth!

À Eric Dawson, professor de guitarra do Victor, pesquisador e músico de Nashville.

À Professora Doutora Andrea von Groll, sempre presente para mim e todos pós-graduandos.

A meus colegas do NUPEMM, João, Michel, Daniela, Ivy, Mariana, Fernanda, Clarice, Júlia, Tati, Priscila, Ana Bárbara, Lisiane e a todos os outros que participam desse grupo.

Aos professores Doutores Carlos Scaini, Melissa Xavier e Adriano Werhli, muito obrigado pelas ótimas aulas.

Aos funcionários da secretaria da pós-graduação, muito obrigado pela atenção prestada ao longo desses quatro anos.

Agradeço a meus amigos José Roberto e Cláudia, Marcelo e Márcia, Beto Vetromille, Hélio Mandeco, Sérgio e Elisa, Eduardo Marafiga, Marco Antônio e Gleci, Otávio e Cláudia, Gisele Simões, Nina e Paula, Brad Lee, Cristina Saboya e Ricardo, Paulo Valentin, Cristina Seidl, Renata Cruz, Paulo e Maria Inês.

Aos professores Francisco Fonseca e Marco Lessa por serem sempre lembrados como exemplos e modelos de pesquisadores e professores.

Em especial um agradecimento com muito carinho à professora e amiga por muitos anos “Dona” Iara Barragan!

A minha mãe Olga Maria Kaster Reis pela presença e pelo apoio dado ao longo de uma vida!

A meus filhos Victor e Isadora por estarem sempre presentes em todas as etapas destes estudos, por terem me acompanhado até os Estados Unidos quando no doutorado sanduíche, pelas dificuldades que, apesar de serem muito jovens, superaram junto comigo.

À Fernanda Pereira da Cruz, minha companheira, por ter me acompanhado em cada momento decisivo desta tese!

“Truth is like the sun. You can shut it out for a time, but it ain't goin' away.”
(Elvis Presley).

SUMÁRIO

Resumo	8
Abstract	9
Lista de Figuras	10
Abreviaturas e Siglas	11
Introdução	12
Revisão Bibliográfica	14
Objetivos	19
Materiais e Métodos	19
Artigo 1	25
Artigo 2	58
Resultados e Discussões	78
Conclusões	80
Referências	81
Anexo 1	87
Anexo 2	88

RESUMO

Este trabalho empregou ferramentas de bioinformática para estudar *in silico* o mecanismo de transporte da bomba de efluxo multidroga AcrB. Essa bomba faz parte do sistema de efluxo da bactéria gram-negativa *Escherichia coli* e é um modelo para estudos de bombas que pertencem à família RND.

Para que fosse estudado o mecanismo de efluxo, uma metodologia foi desenvolvida especialmente para os testes de docagem. Essa metodologia proporcionou a criação das curvas de energia e de diversas tabelas auxiliares que se caracterizaram como importantes ferramentas para a análise dos resultados.

As simulações computacionais de docagem e de dinâmica molecular indicaram que a inibição do mecanismo de efluxo pode ocorrer através de competição entre inibidor e substrato pelos mesmos sítios de ligação ou ainda por caminhos em comum no interior da bomba. Os estudos realizados também revelaram que certas regiões da AcrB parecem ter maior afinidade com os substratos no que se refere à intensidade da energia de ligação. Foram também identificados laços, hélices-alfa e folha- beta que parecem ser relevantes no transporte dos substratos.

Descritores: Bioinformática, Docagem, Dinâmica Molecular, Bomba de Efluxo Multidroga AcrB.

Vieira da Silva Jr., Lande – Bioinformática como Ferramenta para o Estudo do Mecanismo de Efluxo da Bomba Multidroga AcrB. 2016. 88 f. Tese (Doutorado) – Programa de Pós Graduação em Ciências da Saúde. Universidade Federal do Rio Grande, Rio Grande.

ABSTRACT

The present work used bioinformatics tools to study, *in silico*, the transport mechanism of the multidrug efflux pump AcrB. This pump is part of the efflux system of the gram-negative bacteria *Escherichia coli* and is a model for studying other pumps that belong to the RND family.

In order to study the efflux mechanism, a novel methodology was specially developed for the docking tests. This methodology led to the creation of the energy curves and several auxiliary tables that were of great importance as tools to analyze the results.

Docking and molecular dynamics simulations suggested that the efflux inhibition could be due to competition between inhibitor and substrate for the same binding sites or for common pathways in the internal part of the pump. The present studies also revealed that certain AcrB regions seem to have more affinity with the substrates regarding the intensity of the binding energy. Also, loops, α -helices and β -sheets that seemed important for a successful transport were identified.

Descriptors: Bioinformatics, Docking, Molecular Dynamics, AcrB Multidrug Efflux Pump.

Vieira da Silva Jr., Lande – Bioinformatics as a Tool to Study the Efflux Mechanism of the Multidrug Pump AcrB. 2016. 88 f. Thesis (Doctorate) – Graduated Program in Health Sciences. Universidade Federal do Rio Grande, Rio Grande.

LISTA DE FIGURAS

Figura 1 - Sistema de Efluxo AcrB-TolC	15
Figura 2 - Estrutura assimétrica 2GIF da proteína AcrB.....	16
Figura 3 - Estrutura simétrica 1IWG da proteína AcrB.....	17
Figura 4 - Caixas de docagem e definição do eixo z na estrutura 1IWG.....	24
Figura 5 - Curvas de Energia.....	79

LISTA DE ABREVIATURAS E SIGLAS

CLSI - Clinical and Laboratory Standards Institute

CPZ - Chlorpromazine

DOPE - Discrete Optimized Protein Energy

EI - Efflux Inhibitor

EtBr - Ethidium Bromide

FEB - Free Energy of Binding

MDR - Multidrug Resistant

MIC - Minimum Inhibitory Concentration

NMR - Nuclear Magnetic Resonance

PDB - Protein Data Bank

RDD - Rational Drug Design

RFF - Relative Final Fluorescence

RND - Resistance Nodulation Division

TMD - Targeted Molecular Dynamics

1 INTRODUÇÃO

Esta tese foi realizada no Núcleo de Pesquisa em Microbiologia Médica (NUPEMM) da Universidade Federal do Rio Grande (FURG), na cidade do Rio Grande no estado do Rio Grande do Sul e seu objetivo geral foi o de estudar, utilizando técnicas diversas de bioinformática, o mecanismo de efluxo da bomba multidroga AcrB, que faz parte do sistema de efluxo da bactéria *Escherichia coli*.

O efluxo de substâncias tóxicas em bactérias gram-negativas é considerado como um mecanismo de resistência que pode levar ao surgimento de microorganismos resistentes aos antimicrobianos atuais (Li e Nikaido, 2004; Nikaido e Takatsuka, 2009). Ressalta-se, por isso, a importância de estudos, realizados no menor tempo possível, que sejam capazes de desvendar a fisiologia das bombas de efluxo para melhor entender o seu funcionamento e as suas afinidades por determinadas moléculas. Nesse contexto, o emprego de técnicas computacionais mostra-se relevante para estudos da proteína transportadora AcrB sob o ponto de vista das interações energéticas substrato, inibidor e proteína de efluxo. As estruturas disponíveis da AcrB no banco de dados PDB, permitem identificar, através de simulações computacionais, as regiões onde os inibidores se ligam ao transportador e assim caracterizar estes locais em relação às energias de ligação envolvidas no processo. Esta etapa é importante para que se possa desenvolver moléculas capazes de inibir o mecanismo de efluxo da AcrB. A associação dos inúmeros recursos provenientes da bioinformática e de técnicas computacionais com resultados obtidos no laboratório, torna a busca por inibidores da bomba AcrB eficiente, rápida e racional. Justifica-se o emprego da proteína de efluxo AcrB como modelo para estudos de bombas da família RND, pois esta família tem homólogos em archaea, bactérias gram-positivas e células eucarióticas (Tseng *et al.*, 1999). Finalmente, é razoável supor que certas características da AcrB, como o mecanismo de transporte por exemplo, estarão presentes em outras bombas. Portanto, estudar o mecanismo de ação da proteína de efluxo AcrB, empregando técnicas de bioinformática, levará também a compreensão sobre o funcionamento de outras bombas de efluxo.

Para este trabalho foram utilizados bancos de dados da internet como, por exemplo, o Protein Data Bank (Protein Data Bank) para macromoléculas, de onde foram obtidas as estruturas 1IWG (Murakami *et al.*, 2002) e 4DX5 (Eicher *et al.*, 2012). Este banco de proteínas contém mais de 100.000 estruturas e é um repositório mundial para estrutura tridimensional de grandes moléculas e ácidos nucleicos. As estruturas proteicas estão

disponíveis na rede e são de livre acesso, permitindo ainda que pesquisadores as utilizem para estudos, bem como armazenamento e compartilhamento de informações sobre macromoléculas em abrangência mundial. Outro banco de dados importante foi o Zinc (Irwin *et al.*, 2012), de acesso gratuito que permite a obtenção de arquivos “.mol2” de compostos que podem ser utilizados em estudos *in silico*. Nesse ponto, cabe ressaltar a praticidade proporcionada pela bioinformática, onde, se pode, sem qualquer gasto e com grande economia de tempo, escolher estruturas proteicas e candidatos a substratos ou inibidores. Proteína e substrato podem também ter sua compatibilidade verificada através do uso de programas de docagem, como o AutoDock Vina (Trott e Olson, 2010), um programa que realiza os cálculos de maneira muito prática e eficiente para encontrar os locais nos quais proteína e substrato se ligam com altas magnitudes de energia de ligação. Com a etapa de docagem cumprida, o pesquisador pode inferir sobre o uso de certo composto como inibidor de efluxo. Tudo isso poderia ser feito sem qualquer representação gráfica, mas existem programas de distribuição livre que permitem a visualização tridimensional das estruturas proteicas e das moléculas individualmente, ou em conjunto, onde se ilustra a interação proteína-substrato. A inspeção visual de uma estrutura é importante, pois o pesquisador atento e experiente pode apenas com o olhar estimar o melhor lugar para se tentar realizar o procedimento de docagem. Para análise dos resultados de dinâmica molecular dirigida (Schlitter *et al.*, 1994), como a realizada nesse trabalho com o programa Amber 12 (D.A. Case *et al.*, 2012), onde se pode simular a movimentação dos resíduos de uma proteína de acordo com o tempo decorrido, temperatura, pressão e estados inicial e final, a visualização da estrutura é fundamental. Para esse propósito existem programas como o VMD (Humphrey *et al.*, 1996) e o Chimera (Pettersen *et al.*, 2004), ambos gratuitos que permitem, dentre diversas possibilidades, elaborar filmes mostrando como um substrato ou um inibidor se desloca no interior de uma proteína. É importante ressaltar que, devido às dimensões da proteína e também dos substratos e inibidores, microscopia, mesmo que de varredura ou por transmissão de elétrons, não apresenta poder de resolução suficiente para revelar sítios de ligação de um substrato sendo transportado por uma proteína. Entretanto, as ferramentas da bioinformática permitem, com grande grau de certeza, visualizar o que acontece nas dimensões de escala atômicas.

2 REVISÃO BIBLIOGRÁFICA

O aumento no número de casos de doenças infecciosas causadas por microrganismos resistentes aos antimicrobianos reforça a necessidade de desenvolver novas alternativas terapêuticas que possam ser utilizadas no tratamento destas enfermidades (Li e Nikaido, 2004; 2009).

A busca racional de novos fármacos procura reduzir o tempo e o custo do processo, aumentando a eficácia particularmente nas fases iniciais da identificação de um novo farmacóforo. As plataformas de desenvolvimento de fármacos tem sido amplamente reforçadas a partir do uso da biologia computacional e da bioinformática. Estes campos utilizam conhecimentos diversos, tais como: química, física, matemática, ciências da computação, dentre outros, para estudar sistemas biológicos, caracterizando-se como um notável campo interdisciplinar. Surge, então, um novo ambiente de estudos, além dos tradicionais *in vivo* e *in vitro*, chamado de *in silico*, ou seja, estudos virtuais realizados no computador. Nestes estudos através de simulações que envolvem modelos matemáticos pode-se simular o funcionamento de sistemas de organismos vivos. Neste sentido, estes estudos podem ser utilizados na compreensão de como interagem fármacos e proteínas (Lesk, 2005; Claverie e Notredame, 2007).

O sistema de efluxo constitui-se em um dos mais importantes mecanismos relacionados à resistência bacteriana a inúmeros compostos como corantes, detergentes, antibióticos e outras substâncias (Eicher *et al.*, 2009). Este sistema é composto por proteínas transportadoras que realizam a extrusão de substâncias do interior da célula (Martinez *et al.*, 2009). Por prevenir a acumulação de compostos tóxicos no interior da célula, o sistema de efluxo permite à bactéria sobreviver em meios inóspitos, aumentando seu tempo de vida e favorecendo o surgimento de mutantes resistentes aos antimicrobianos (Martinez e Baquero, 2000).

Estas proteínas, conhecidas como bombas de efluxo, pertencem a diversas famílias de transportadores sendo que as bombas da superfamília RND (Resistance Nodulation Division), são predominantemente encontradas em bactérias gram-negativas, e apresentam grande relevância para a resistência do tipo multidroga (Nikaido e Takatsuka, 2009; Fernandez e Hancock, 2012).

A bactéria *Escherichia coli* apresenta um sistema RND complexo composto por três proteínas distintas, AcrA, AcrB e TolC (Figura 1) (Seeger *et al.*, 2008), e que está relacionado

com o transporte de antibióticos de diferentes famílias como tetraciclina, cloranfenicol, beta-lactâmicos, rifamicinas e quinolonas (Okusu *et al.*, 1996; Webber e Piddock, 2003). Este complexo, responsável pela multiresistência, é referido na literatura como bomba de efluxo AcrAB-TolC e é capaz de capturar moléculas nocivas tanto no citoplasma quanto no espaço periplasmático, para posteriormente exportá-las para o meio extracelular (Nikaido, 1996). Quanto às funções de cada uma das partes acredita-se que a proteína AcrA tenha a função principal de conectar a AcrB com a TolC, a fim de estabilizar o conjunto. Já a proteína TolC posiciona-se imediatamente sobre a AcrA, de tal maneira que atravessa a membrana externa, ou seja, ela recebe as moléculas que são capturadas e transportadas pela AcrB e as conduz para o meio externo. Por fim, a parte fundamental deste sistema é a proteína AcrB (Figura 2), que é ativa e dependente de um fluxo de prótons para seu funcionamento (Seeger *et al.*, 2008).

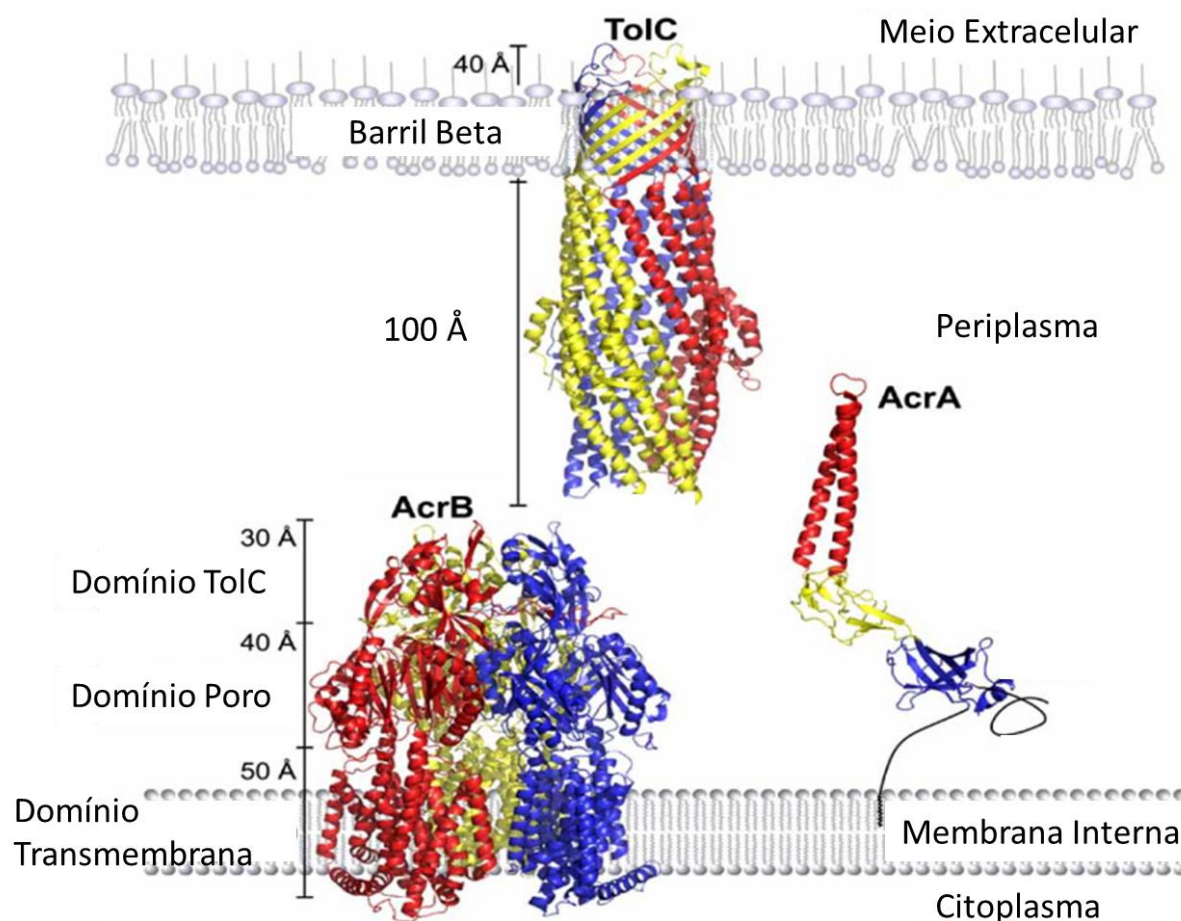


Figura 1: sistema de efluxo AcrAB-TolC da bactéria gram negativa *Escherichia coli*. Constituído por três partes distintas, uma proteína de conexão AcrA; uma outra que conduz as substâncias tóxicas para o exterior via membrana externa (TolC); e a AcrB onde ocorre o transporte ativo e acontece a captura de substratos no citoplasma ou no espaço periplasmático e subsequente transporte até a proteína TolC (figura adaptada de Seeger *et al.*, 2008).

A proteína AcrB é composta por três monômeros iguais, e, por isso dita homotrimérica. As cadeias A, B e C – idênticas em relação à sequência de aminoácidos – da AcrB em três vistas diferentes podem ser visualizadas na Figura 2. Esta estrutura foi obtida por cristalografia de raios-X e seu código no Protein Data Bank (PDB) é 2GIF (Seeger *et al.*, 2006). As três cadeias estão em um estado assimétrico, o que levou a formulação de um modelo de transporte baseado na mudança cíclica de conformação entre os monômeros. Estes poderiam assumir configurações chamadas “Loose”, “Tight” e “Open”. O estado “Loose” refere-se a um estado relaxado, no qual o monômero está pronto para receber o substrato; o estado “Tight” refere-se a um estado no qual o substrato está fortemente ligado a um sítio de ligação e, finalmente o estado “Open” é o estado no qual a proteína libera o substrato para a proteína TolC, para que ele seja transportado para fora da membrana externa. Na Figura 3, obtida pela mesma técnica está a estrutura de código PDB 1IWG (Murakami *et al.*, 2002), cadeia A, também em três visualizações distintas. Com este monômero, compõe-se o trimer que então é formado por três partes iguais. A proteína assim formada é dita simétrica e é considerada no estado de repouso por não ter nenhum substrato ligado a ela quando resolvida.

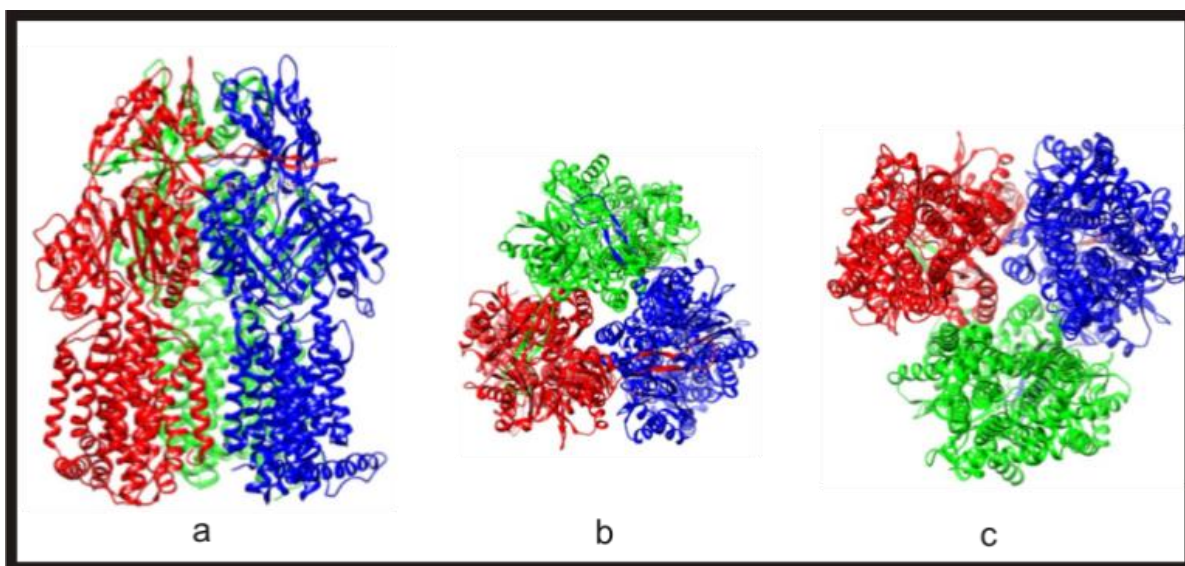


Figura 2: estrutura assimétrica 2GIF e as três cadeias A (vermelho), B (azul) e C (verde). a- vista lateral (referência), b- vista superior, c- vista inferior. Imagem gerada pelo programa Chimera.

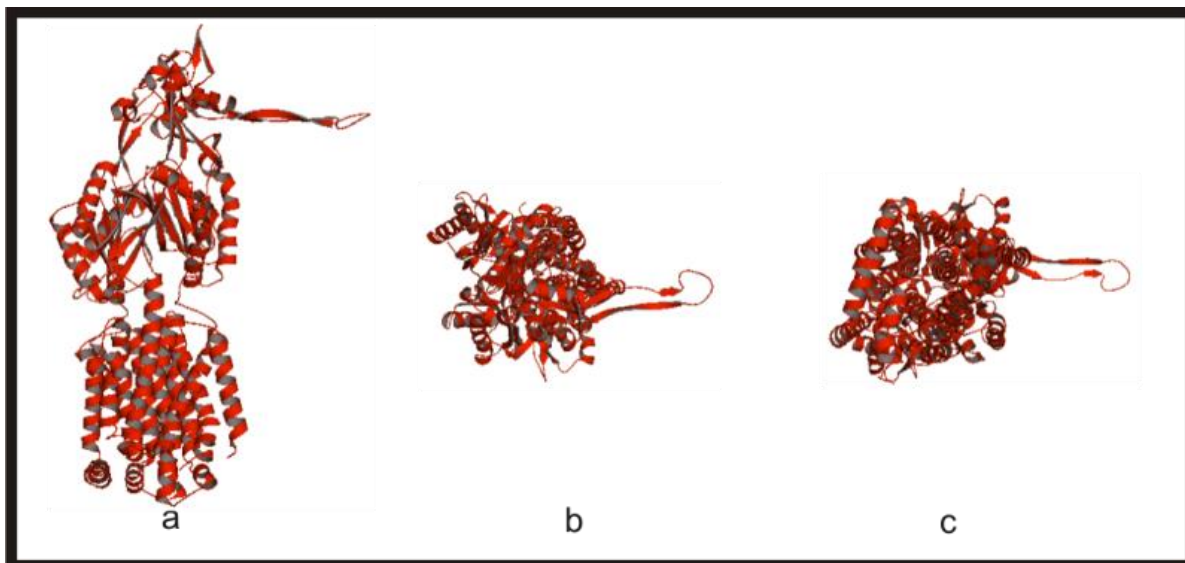


Figura 3: AcrB cadeia A em três vistas distintas, a- lateral (referência), b- superior, c- inferior. Imagem gerada pelo programa PyMol.

Embora todas as três proteínas (AcrA, AcrB e TolC) que compõe o complexo sejam importantes para o sucesso do efluxo, é na AcrB onde ocorre o acesso das moléculas para o interior da bomba e também onde ocorre a parte ativa do transporte (Seeger *et al.*, 2006; Pos, 2009). Para interferir de forma racional neste mecanismo de efluxo, torna-se importante entender o funcionamento da AcrB. Entretanto, esta é uma proteína grande, composta por 3147 aminoácidos (1049 por cada cadeia), o que traz certas dificuldades à compreensão mais clara do processo, como por exemplo, a variada gama de substratos da AcrB (Li *et al.*, 2015). Isoladamente, os estudos *in vitro* não são capazes de desvendar o que acontece durante o momento em que as moléculas são primeiramente atraídas para dentro da bomba ou ainda quando se ligam aos locais de alta afinidade e demandam energia para a retirada da molécula destas regiões de forte ligação. Um modelo, bastante aceito correntemente, que tenta explicar como a bomba de efluxo AcrB funciona é o modelo de rotação das 3 cadeias A, B e C, alternando entre os modos “entrada, ligação e saída”. Entretanto, o mecanismo de transporte da bomba de efluxo desde a ligação inicial do substrato até a saída, é ainda desconhecido (Seeger *et al.*, 2006; Eicher *et al.*, 2012).

Até o mês de março de 2016, 43 estruturas da AcrB foram depositadas no Protein Data Bank (Protein Data Bank), desde a publicação da estrutura 1IWG (Murakami *et al.*, 2002). Essas estruturas permitem o emprego de técnicas de biologia computacional para estudar o comportamento da AcrB frente a diferentes compostos.

A inibição de bombas de efluxo multidroga em bactérias gram-negativas, como a AcrB da *Escherichia coli*, justifica-se pelo fato destas bombas terem um papel relevante no mecanismo de resistência *per se* bem como precursoras no desenvolvimento de cepas bacterianas resistentes (Vargiu e Nikaido, 2012). Então, compreender como funciona o mecanismo de efluxo, e particularmente como atuam as proteínas transportadoras responsáveis, é fundamental para a inibição do mecanismo, permitindo não apenas o uso clínico de antibióticos que são substratos destas bombas, desenvolvendo estratégias para inibição e finalmente colaborando para a desaceleração no surgimento de microrganismos resistentes aos antibióticos.

Bombas de efluxo apresentam algumas características em comum, sendo uma delas universal: todas precisam capturar a molécula a ser exportada e conduzi-la para fora da célula. Desta forma, o conhecimento obtido e as técnicas desenvolvidas para se entender os processos de transporte e ligação de um inibidor com a AcrB poderão, posteriormente, ser estendidos para outras proteínas de efluxo, de outros microrganismos ou mesmo para tratamento de doenças humanas onde o efluxo é um mecanismo que dificulta o tratamento (Xia e Smith, 2012). Entretanto, como o sistema de efluxo tem a função natural de proteção das células, projetar inibidores que o inviabilizem, é algo complexo quando considerada a necessidade de toxicidade seletiva, mas desvendar o mecanismo de funcionamento de uma proteína de efluxo pode ajudar nesta questão. Pode-se, por exemplo, adaptar as estratégias para incapacitar a bomba de efluxo AcrB da *E. coli* para inibir a bomba de efluxo Tap do *Mycobacterium tuberculosis*, um microrganismo que apresenta um grande número de bombas de efluxo relacionadas à resistência tanto intrínseca quanto adquirida (Da Silva *et al.*, 2011). Ou ainda, utilizar os conhecimentos para a AcrB, em casos de *Pseudomonas aeruginosa* multirresistentes onde a característica “multiresistência” é atribuída, dentre outros fatores, a superexpressão de bombas de efluxo (Hirsch e Tam, 2010).

Embora uma patente até sugira formas gerais de moléculas inibidoras de bombas de efluxo (Chamberland, 2002) o fato de ainda desconhecermos, de forma detalhada, o mecanismo funcional das bombas torna difícil, senão impossível, o planejamento racional de inibidores que atuem inibindo apenas a proteína de efluxo de interesse.

Diante do exposto, ressalta-se a importância de estudos que sejam capazes de desvendar a fisiologia das bombas de efluxo para melhor entender o seu funcionamento e as suas afinidades por determinadas moléculas. Assim, justifica-se o uso da AcrB como modelo para estudos de bombas da família RND que tem homólogos em archea, bactérias gram-positivas e células eucarióticas (Tseng *et al.*, 1999). O emprego de técnicas computacionais é

fundamental para se estudar o transportador AcrB sob o ponto de vista das interações energéticas ligante (inibidor) e proteína de efluxo. As estruturas disponíveis da AcrB no banco de dados PDB, permitem a realização de varreduras virtuais de variados ligantes com essa proteína de transporte em um tempo pequeno. Com isso, torna-se possível identificar, utilizando modelos computacionais, as regiões onde os inibidores se ligam ao transportador para caracterizar estes locais em relação às energias envolvidas no processo. Esta etapa é fundamental para que se possa desenvolver moléculas capazes de inibir a AcrB. A associação dos inúmeros recursos provenientes da bioinformática e da biologia computacional com resultados obtidos no laboratório, torna a busca por inibidores da bomba AcrB eficiente, rápida e racional.

3 OBJETIVOS

3.1 Objetivo Geral

Estudar o mecanismo de transporte da AcrB utilizando técnicas de Bioinformática e Biologia Estrutural.

3.2 Objetivos Específicos

Propor uma metodologia para mostrar a distribuição das energias de ligação na proteína de substratos de interesse.

Identificar os sítios de alta energia de ligação tanto na superfície quanto no interior da proteína.

Revelar possíveis caminhos percorridos pelos substratos testados.

Identificar resíduos da proteína importantes para o sucesso do efluxo.

Sugerir maneiras de inibir o mecanismo de efluxo.

4 MATERIAIS E MÉTODOS

4.1 Bancos de Dados na Internet e Programas utilizados

A fim de cumprir os objetivos desse trabalho, fez-se necessário o emprego de diversas ferramentas de bioinformática, descritas a seguir, de acordo com sua finalidade.

Protein Data Bank – Banco de Dados de Proteínas.

Para que seja possível realizar simulações computacionais e revelar sítios de alta afinidade entre substrato e proteína, é necessário ter um arquivo que informe cada átomo que compõe a proteína bem como suas coordenadas cartesianas (x,y,z). Arquivos .pdb contém estas informações e podem ser obtidos na internet na página da internet do Protein Data Bank (Protein Data Bank), que é um banco de dados de estruturas proteicas. Esses arquivos ainda informam o resíduo da proteína ao qual o átomo pertence. O cabeçalho do arquivo fornece informações sobre como a estrutura foi obtida, autores, se possui ou não substratos ligados dentre outras. Nesse trabalho, foram utilizadas as estruturas 1IWG (Murakami *et al.*, 2002), simétrica no estado de repouso, e 4DX5 (Eicher *et al.*, 2012) assimétrica. Ambas foram obtidas com a técnica de cristalografia e difração de raios-X.

Zinc – Banco de Dados de Moléculas.

Os substratos para docagem e estudos de dinâmica molecular foram obtidos no banco de dados Zinc (Irwin *et al.*, 2012), que possui aproximadamente 90 milhões de compostos em forma biologicamente relevante, gratuitamente. Moléculas como o brometo de etídio, tetraciclina e clorpromazina foram obtidas nesse banco de dados. Entretanto, algumas moléculas, como a molécula NUNL02, derivada da tetra-hidropiridina, do artigo 1, tiveram que ser construídas com o programa HyperChem (HyperChem (TM) Professional 7.51, HyperCube Inc.) baseada em sua fórmula estrutural.

VMD, PyMol e Chimera – Programas de Visualização, Modelagem e Análise.

Existem diversos programas destinados à modelagem, visualização e análise de arquivos “.pdb” disponíveis na internet, gratuitamente, para “download”. Os três utilizados nesse trabalho foram os programas VMD (Humphrey *et al.*, 1996), PyMol (Schrodinger, 2015) e Chimera (Pettersen *et al.*, 2004). O programa VMD (Visual Molecular Dynamics), utilizado para confeccionar a maioria das imagens do artigo 2 deste trabalho, possibilita diferentes representações para a proteína, para seus resíduos e átomos ligados, o que facilitou a identificação de regiões de interesse. A ferramenta de análise permitiu identificar os resíduos envolvidos no transporte. Além disso, o programa possibilitou a visualização de

resultados de estudos de dinâmica molecular e a construção de animações exibindo o movimento dos substratos na proteína. O programa PyMol Molecular Graphics System é semelhante ao VMD, mas a distribuição gratuita não possui todas as ferramentas de análise e não permite a elaboração de imagens de alta qualidade para publicação, embora gere imagens razoáveis. Esse programa foi bastante utilizado para manipulação de arquivos .pdb, como por exemplo para a construção homotrimérica da estrutura IIWG, e análise dos estudos de docking e dinâmica. Finalmente, o programa Chimera da Universidade da Califórnia em São Francisco, também de distribuição gratuita, sem limitações, faz essencialmente tudo o que os outros dois fazem, mas possibilita uma melhor manipulação de estruturas com diversas moléculas ligadas. Esse programa foi fundamental na preparação de substratos para docagem no artigo 1 e para a criação dos complexos citados no artigo 2.

Embora parecidos e com muitas ferramentas em comum, cada um dos três programas mostra-se mais eficiente que os outros em certos aspectos, tanto no quesito facilidade de executar a tarefa desejada quanto no quesito utilização da unidade de processamento central (CPU) do computador: por exemplo, o programa Chimera também permite a visualização de arquivos de dinâmica molecular, mas o custo para a CPU é alto. Já o programa VMD executa a tarefa sem grande demanda para a CPU. Assim, a escolha sobre a melhor ferramenta depende do que se quer fazer, da habilidade do usuário com cada programa e da capacidade computacional de seu sistema.

Modeller – Modelagem de partes não resolvidas da estrutura IIWG.

Para esta finalidade, foi utilizado o programa Modeller 9.14 (Webb e Sali, 2014) que criou um loop para os resíduos 498 até 512 e 860 até 868, utilizando a configuração padrão. A escolha da melhor estrutura dentre as geradas foi baseada no valor DOPE, uma medida estatística usada para avaliar modelos gerados por homologia em predição de estruturas proteicas, e em inspeção visual da estrutura. A estrutura escolhida, um monômero derivado da IIWG original foi então transformada em uma estrutura homotrimérica, simétrica, com a utilização do programa PyMol. Essa estrutura foi utilizada no artigo 2 para os estudos de dinâmica molecular dirigida (Targeted Molecular Dynamics - TMD). Este programa é de distribuição livre.

AutoDock Tools e AutoDock Vina – Preparação da estrutura e ligantes para Docagem e Docagem respectivamente.

O programa utilizado para preparar as estruturas e ligantes para a docagem foi o Autodock Tools (Sanner, 1999; Morris *et al.*, 2009), que também permite a obtenção de imagens. O programa Autodock Vina (Trott e Olson, 2010) foi escolhido para a docagem. Os dois programas são gratuitos e disponíveis para “download”. AutoDock Tools é necessário para salvar os arquivos .pdbqt referentes à estrutura utilizada da proteína e dos substratos e também para dimensionar e posicionar a caixa no local da proteína onde, a partir dela serão feitas as docagens. O programa AutoDock Vina realiza, então, com os arquivos .pdbqt e com as dimensões e posicionamento da caixa a busca pelos sítios de melhor energia livre de ligação em cada caixa.

Amber versões 12 e 14 – Minimização e Dinâmica Molecular Dirigida.

Para o procedimento de dinâmica molecular dirigida (“Targeted Molecular Dynamics” ou TMD em inglês), minimização da estrutura 1IWG-47 foi utilizado o programa Amber versão 12 (D.A. Casee *et al.*, 2012). Para preparação dos complexos do artigo 2 foi utilizado uma ferramenta do Amber chamada de X-Leap, como interface gráfica ou ou T-Leap com acesso via terminal. T-Leap é mais eficiente, pois permite a escrever um pequeno programa que executa automaticamente diversos procedimentos, evitando assim a digitação repetitiva de comandos. Entretanto, algumas vezes o X-Leap deve ser utilizado. Para minimização da estrutura 4DX5 foi utilizado o programa Amber versão 14 (D.A. Case, 2015). Este programa não é de distribuição livre, necessita de uma licença para sua utilização. A dinâmica molecular é uma técnica muito utilizada para o estudo de macromoléculas biológicas. Sua metodologia está baseada na mecânica clássica newtoniana, mas essa ferramenta fornece informações sobre a movimentação dos resíduos da proteína que levam a diferentes configurações enquanto o tempo decorre. As forças de interação entre cada átomo dos resíduos que compõe a proteína são calculadas a partir de campos de força. Nesse estudo utilizamos para o artigo 1 o campo de força ff14SB e para o artigo 2 o ff12SB. Um típico campo de força apresenta-se conforme $V_{(r)} = \sum V 1 + \sum V 2 + \sum V 3 + \sum 4 + \sum 5$. Onde o termo $V 1$ refere-se a energia de afastamento entre dois átomos em relação à posição de equilíbrio, $V 2$ refere-se a energia devida ao ângulo da ligação entre o par de átomos, $V 3$ relaciona-se a energia de torção da ligação, $V 4$ refere-se a energia das interações de Van der

Waals e V_5 refere-se a energia eletrostática. A técnica utilizada no artigo 2 chama-se de dinâmica molecular dirigida (Targeted Molecular Dynamics) (Schlitter *et al.*, 1994) e simula a evolução da macromolécula entre dois estados, por exemplo com o substrato em uma posição de partida inicial em um complexo da macromolécula e em outro complexo o substrato encontra-se na posição final. Com essa técnica pode-se verificar como o substrato e a proteína interagem, uma vez que o substrato será conduzido ao estado final pelo caminho mais natural possível. Uma alternativa gratuita para a dinâmica molecular é o programa Gromacs (Van Der Spoel *et al.*, 2005).

LigPlot+ – Identificar os resíduos da proteína contatados pelos substratos.

Para esse fim, foi utilizado o programa LigPlot+ (Laskowski e Swindells, 2011), gratuito para uso acadêmico. Esse programa permitiu verificar quantos e quais resíduos da proteína fazem contatos com o brometo de etídio e com a molécula NUNL02. A partir dessas informações verificamos quais resíduos são comuns aos dois substratos por cada posição de melhor energia de ligação. O procedimento está descrito no artigo 1.

4.2 Metodologia

Os métodos utilizados para alcançar os resultados e conclusões dos artigos 1 e 2 estão detalhados nas respectivas seções de metodologia dos referidos artigos.

Entretanto, é importante ressaltar que esse trabalho teve como ponto de partida a construção das curvas de energia. Para este fim, foi desenvolvida uma metodologia que examina a proteína ao longo de seu eixo mais extenso o qual chamamos de eixo z , e que corresponde a direção vertical de acordo com a representação proposta. A bomba de efluxo é subdividida em várias caixas de mesmo tamanho (Figura 4). Calcula-se a melhor energia livre de ligação para cada caixa com o programa Autodock Vina. Então, para cada caixa temos um ponto que define a posição de melhor docagem. Finalmente, como resultado desse processo, construímos um gráfico no qual o eixo vertical representa a energia de ligação e o eixo horizontal, a respectiva caixa, construindo assim as curvas de energia (Figura 6). No artigo 2, o eixo horizontal não é a caixa, mas sim a coordenada média de ligação z da molécula em relação aos resíduos da proteína, para maiores detalhes, leia a metodologia do referido artigo. As curvas de energia possibilitaram as análises feitas no artigo 1 e foram o ponto de partida para a dinâmica molecular realizada no artigo 2. Posteriormente, o procedimento de docagem

desenvolvido para o primeiro artigo foi automatizado e transformado em uma ferramenta escrita em Python que realiza a docagem em posições predeterminadas pelo usuário e também fornece os resíduos contatados entre a proteína e o ligante (Seus e K. S.; Prates, 2016).

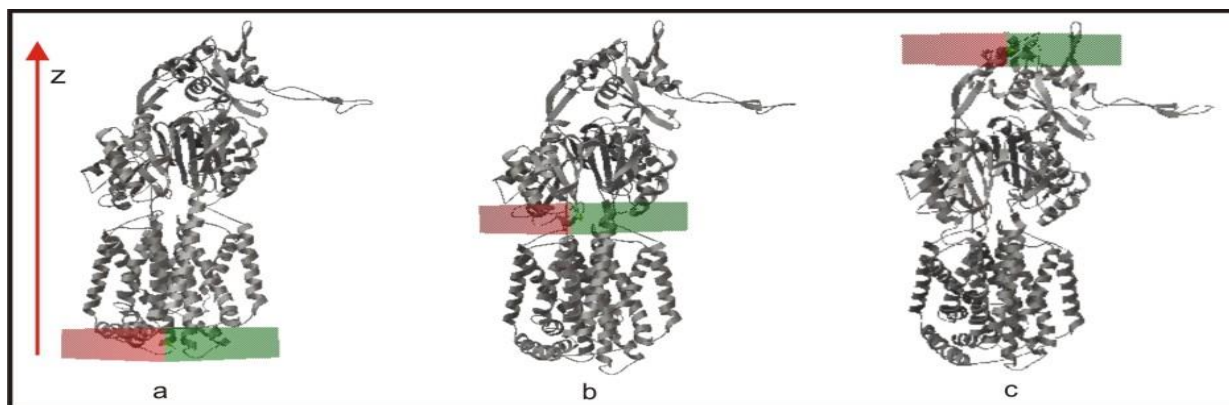


Figura 4: proteína AcrB e a direção do eixo z. As imagens a, b e c representam 3 das camadas onde são realizadas as docagens.

ARTIGO 1

Submetido para o International Journal of Antimicrobial Agents

Title: In vitro and in silico analysis of the efficiency of tetrahydropyridines as drug efflux inhibitors in *Escherichia coli*.

Lande Silva Jr.^a, Lillian Lucas Carrion^a, Andrea von Groll^a, Sofia Santos Costa^c, Elisabete Junqueira^c, Daniela Fernandes Ramos^a, Jéssica Cantos^a, Vinicius Rosa Seus^b, Isabel Leitão Couto^c, Liana da Silva Fernandes^d, Hélio Gauze Bonacorso^d, Marcos Antônio Pinto Martins^d, Nilo Zanatta^d, Miguel Viveiros^c, Karina S. Machado^b Pedro Eduardo Almeida da Silva^{#a}

^a*Núcleo de Pesquisa em Microbiologia Médica, Faculdade de Medicina, Universidade Federal do Rio Grande - FURG, Rio Grande, Brasil*

^b*Centro de Ciências Computacionais, Universidade Federal do Rio Grande, Rio Grande, Brasil*

^c*Grupo de Micobactérias, Unidade de Microbiologia Médica and Global Health and Tropical medicine (GHTM), Instituto de Higiene e Medicina Tropical, Universidade de Nova Lisboa, Lisboa, Portugal*

^d*Núcleo de Química de Heterociclos, Departamento de Química, Universidade Federal de Santa Maria, Santa Maria, Brasil*

Corresponding author: Pedro Eduardo Almeida da Silva pedrefurg@gmail.com

ABSTRACT

Objectives: To evaluate tetrahydropyridine derivatives for their ability to inhibit intrinsic and over-expressed AcrB efflux, and understand the mechanism action of the compounds by *in silico* studies. **Methods:** To address the objectives it was carried out the minimal inhibitory concentration determination, real-time fluorometric method and docking simulations.

Results: The results showed that compounds NUNL02, NUNL09 and 10 inhibit efflux mechanism. NUNL02 has a high affinity for the multidrug efflux pump AcrB and was effective even in the presence of glucose under similar physiological condition. Docking studies suggested that the action of NUNL02 in the accumulation assay is very likely because of competition with ethidium bromide for common binding sites and residues of the protein. This finding was supported by the fact that the best inhibitor had strong affinity for the regions where the substrate exhibited its strongest binding as well as for other positions.

Conclusion: Herein we showed for the first time the potential of tetrahydropyridines as efflux inhibitors and highlighted compound NUNL02 as an AcrB specific inhibitor. Docking studies suggest that competition is the putative mechanism action of these compounds

Key Words: Efflux, Inhibitors, AcrB, Competition.

INTRODUCTION

The phenomenon of multidrug resistance (MDR) is emerging in several clinically important microorganisms¹ and drug efflux is an important mechanism contributing to bacterial resistance to many antimicrobial and biocides².

RND (Resistance-Nodulation-cell Division) family transporters are tripartite systems responsible for the active efflux of many structurally dissimilar substrates, including clinically relevant antimicrobials³. The AcrAB-TolC system present in *Escherichia coli* consists of a transporter protein anchored in the plasma membrane that extends to the periplasm (AcrB), a protein that forms a channel through the outer membrane (TolC) and an adapter protein (AcrA) located in the periplasm⁴.

The role of efflux in subclinical resistance is poorly understood, although it was suggested that it facilitates the selection of high-level resistant strains^{5, 6}. Therefore, the use of substances capable to inhibit efflux represents a promising therapeutic to implement as adjuvants to antimicrobials, prevent the acquisition of another mechanism of resistance, or restore the effectiveness of an antimicrobial agent⁷.

Pyridines, an important class of organic compounds, act as proton transfer agents, and tetrahydropyridines have been shown to possess both therapeutic and pharmacological properties such as flow regulation of calcium ions⁸, being candidates to efflux inhibitors (EI) in bacteria.

Advances in computer modeling and bioinformatics have impacted the drug discovery process⁹⁻¹¹. Rational Drug Design (RDD)^{12, 13}, is an *in silico* step that reduces time and cost involved in drug discovery. RDD helps to elucidate the interaction between small compounds and a target receptor that can be simulated through molecular docking^{14, 15}. Hundreds of thousands of orientations and conformations of a ligand in the binding site of the receptor are tested and the best position and conformation of the drug candidate is estimated

in terms of free energy of binding (FEB) scores^{15, 16}. The more negative the FEB, the stronger the binding.

This study evaluated eighteen tetrahydropyridines derivatives for their ability to inhibit intrinsic (wild-type) and over-expressed AcrB. To understand the putative mechanism of action, we performed *in silico* docking simulations using an original approach.

MATERIALS AND METHODS

Derivatives of tetrahydropyridines

Eighteen synthetic compounds were obtained using a previously described method^{17, 18}. Table A.1 shows the structures of the tetrahydropyridines (NUNL) and the respective molecular formulas. All compounds were solubilized in 99.5% dimethyl sulfoxide at a concentration of 10,000 mg/l and stored at 4°C.

Chemical reagents

All reagents were acquired from Sigma-Aldrich. Stock solutions of glucose (20% p/v), ethidium bromide (EtBr) (10,000 mg/l), kanamycin (20,000 mg/l), tetracycline (10,000 mg/l), chlorpromazine (CPZ) (10,000 mg/l) and verapamil (30,000 mg/l) were prepared in distilled water, filtered with a 0.22 µm filter (Millipore Corporation USA), stored at 4°C and kept protected from light.

Microorganisms and growth conditions

Escherichia coli K-12 AG100, AG100_{TET} and AG100A strains were used in this study. K-12 AG100 is a wild-type strain that has several active intrinsic drug mechanisms, including the AcrAB-TolC system. The AG100_{TET} strain corresponds to strain AG100 adapted to tetracycline and over-expressing the AcrB, while AG100A corresponds to AG100 (Δ *acrAB*::Tn903 Kan^r) strain with an inactivated AcrAB¹⁹⁻²¹. The microorganisms were

grown in LB medium, supplemented with 8 mg/l of tetracycline for AG100_{TET} growth or with 100 mg/l of kanamycin for AG100A, and incubated for 18 hours at 37°C.

Determination of minimum inhibitory concentrations (MIC)

The MICs of the tetrahydropyridine compounds, CPZ and verapamil were determined using the two-fold broth microdilution method in 96-well microtiter plates according to the CLSI²². The range of concentrations tested were 0.75 - 400 mg/l for the tetrahydropyridine compounds, 0.45 - 240 mg/l for CPZ and 11.5 - 3000 mg/l for verapamil. After 18 hours of incubation at 37°C, the MICs were registered.

Real-time fluorometric method

A fluorometric method was used for detection of accumulation and efflux of the fluorochrome EtBr, a known substrate of the AcrB²², using a Rotor-Gene 3000TM thermocycler (Corbett Research, Sydney, Australia) in combination with real-time analysis software and filters for excitation and emission with wavelengths of 530 nm and 585 nm, respectively²³⁻²⁵.

The tetrahydropyridine inhibitory activity upon the efflux activity was assessed in two steps: (I) accumulation assays in the presence of the tetrahydropyridines and (II) efflux inhibition assays, performed as previously described²⁴⁻²⁶. For both assays a steady-state EtBr concentration, characteristic of each strain, was used to evaluate efflux inhibition as follows: 1 µg/mL (AG100 and AG100_{TET}) and 0.25 mg/l (AG100A). The reference efflux inhibitors (CPZ and verapamil) and the compounds were used at sub-inhibitory concentrations to guarantee no effect on cell viability (Table A.2).

The relative final fluorescence (RFF) index²⁵ was determined as $RFF = (F_{mol} - F_{control}) / F_{control}$, to compare the tetrahydropyridine efflux inhibitory activity. F corresponds to the final fluorescence values (at 60 min) of the EtBr accumulation curve in the presence of an inhibitor (F_{mol}), or the EtBr accumulation curve in the absence of the inhibitor ($F_{control}$).

Assays were performed in duplicates and the RFF mean and standard deviation were calculated with MedCal[®] v9.6.2.0. For each assay, the difference between the relative fluorescence, obtained under conditions of maximum efflux (glucose only), and minimum efflux (CPZ only) corresponds to 100% of efflux (i.e., 0% efflux inhibition). Thus, the percentage of EtBr efflux inhibition promoted by each compound corresponded to the difference between the maximum efflux and the ratio of relative fluorescence in the presence of the tetrahydropyridine compound relative to the maximum efflux condition.

Cytotoxicity assay

Cytotoxicity of NUNL 02, NUNL 09, NUNL 10 and NUNL 18 were assessed with the J774A.1 (ATCC TIB-67) and MCF-7 as recommended by Pavan et al. (2010)²⁷. The cells were cultivated at 37 °C with 5% CO₂ in modified DMEM medium (10,4 g DMEM, 900 ml water MilliQ, 2,0g NaHCO₃ - Sigma D5671) complemented with 10% Bovine Fetal Serum in order to obtain about 1×10^5 cell/mL and 5×10^4 cell/mL to J774A.1 and MCF-7, respectively. The compounds were dissolved in DMSO (5%) and subjected to two-fold serial dilution from 200 to 0.78 mg/l. Cells were exposed to the compounds at various concentrations for 24 h. Resazurin solution was then added to the cell cultures which were incubated for 6 h. Cell viability was detected by reduction of resazurin to resorufin²⁷.

Docking

Docking studies with Autodock Vina¹⁰ and AutoDock Tools^{28, 29} were performed with the AcrB structure, PDB ID 4DX5³⁰ (Protein Data Bank website <http://www.rcsb.org/pdb/home/home.do>), as the target receptor. Structure 4DX5 was chosen for the docking studies because of its high resolution (1.9 Å) and also because its three chains, A, B and C compose an asymmetric trimer, likely to correspond to one of the many possible

biological configurations of AcrB. The software Chimera³¹ was used to prepare each of the chains for docking. Finally, minimization was performed with software Amber 14, force field ff14SB³². This procedure generated three different receptors for the docking studies: 4DX5-A, 4DX5-B and 4DX5-C.

Since the number or location of the best binding sites for ligands in the AcrB protein is unknown, it is problematic to define a region to place a single grid in which the FEB for the best ligand positions and orientations will be found. In order to locate the site, or sites, of best binding a good strategy is to position many searching grids along the structure and calculate the FEB in each of them. The size of the grids was defined with Autodock Tools, and Autodock Vina was used to calculate the FEB values between EtBr, NUNL02, CPZ, tetracycline and structures 4DX5-A, B and C in each grid arranged along the z axis in (Figure 1). Each grid, hereafter called box, was constructed so that the length in the directions x and y (65 Å for each) determined the base of the box. Such dimensions allowed the resulting area to cover the protein in these directions, with the exception of the loop to the right. The direction z is the height of the box; a dimension of 15 Å was chosen for this parameter. The lengths chosen for x , y and z define the volume of the boxes and allow the ligands to assume all possible orientations inside them. The size of the box was kept constant for the docking studies. The first box was built at the lowest position of the protein, with the z coordinate defined as 260 (Figure 1a). Autodock Vina determined the best binding position for this box. Next, the box is kept with coordinates x and y unchanged, but 5 Å were added to the z coordinate, defining box 265. Autodock Vina calculates again the FEBs and determines the best binding position of the ligand for this box. This procedure is repeated until box 380 (Figure 2c), the last one defined. Adjacent boxes have a superposition of 10 Å and so the procedure allows that near boxes have the same, or similar, best binding positions. In total, 25

boxes were used to determine the best affinity sites between 4DX5-A, B, C and EtBr, NUNL02, CPZ and tetracycline.

Energy Distribution Curves

For each box in each of the three structures, the first model generated by Autodock Vina was considered to be the best binding position for the ligand. Through the coordinates of best binding, average coordinates (x , y , z) were calculated for each ligand in each box. Then, a single value for x , y and z was considered to define the average position between each ligand and the structure in each box. Plotting the z average coordinates against the FEB for each box, and joining adjacent points, allowed the construction of the “energy distribution curves”. The exact box, z average coordinates and FEBs used for making the curves can be found in Tables A.3, A.4 and A.5.

Substrate/Ligand competition for common binding sites

With the results obtained with LigPlot⁺³³, Tables 1, A.6 and A.7, which summarize the number of common residues between EtBr and NUNL02 for each EtBr box were made for all structures. In the left column of these tables is the number of the box where EtBr was docked, in the middle column, the number of the boxes where NUNL02 was docked that have common protein binding residues with EtBr, finally, in the right column, is the respective number of common residues (NCR) relative to the number of total EtBr binding contacts with protein residues at the respective box (EtR).

Graphics of the tridimensional structure of the AcrB were made with Autodock Tools²⁸ and PyMol³⁴.

RESULTS

MIC and Cytotoxicity

Most of the compounds presented, for all strains, MICs equal or higher than 200 mg/l, thus showing no relevant antibacterial activity. The MIC values were used to select the concentrations of tetrahydropyridines applied in the fluorometric assays. No cytotoxicity of NUNL02, NUNL09, NUNL10 and NUNL18 was observed in the used concentrations to both cell line tested.

Effect of the tetrahydropyridine derivatives on EtBr accumulation

Higher RFF values indicate a better compound capacity to promote EtBr accumulation. As shown in Figures A.1 and A.2a, CPZ, NUNL02 and NUNL10 had higher RFF values than verapamil and NUNL06, either in the absence or presence of glucose in AG100. Nevertheless, verapamil showed significant effect on AG100A (Figure A.2c). Interestingly, NUNL02 retained its capacity to promote EtBr accumulation in the presence of glucose.

For AG100_{TET}, again NUNL02, NUNL09 and NUNL18 were the most effective promoters of EtBr accumulation, with activity comparable to CPZ, either in the absence or presence of glucose (Figure A.2b). The other compounds and verapamil showed no significant effect. Regarding AG100A, the most effective compounds were NUNL09 and 10, either in the presence or absence of glucose. NUNL03, 14, 16 and 21 also displayed significant effect upon EtBr accumulation (Figure A.2c).

Efflux inhibition study

Table 2 shows the capacity of NUNL02, 9 and 10 to inhibit EtBr efflux in AG100 and AG100_{TET}. NUNL09 and NUNL10 presented a higher efflux inhibitory capacity.

However, this capacity was significantly reduced in the presence of glucose. NUNL02 had an efflux inhibitory capacity similar to CPZ for both strains and apparently was not so affected by glucose.

Docking

When there was an active AcrAB-To1C system, NUNL02 exhibited high RFFs, in contrast if this efflux system was not functional, it had very low RFFs, a fact that support the conclusion that NUNL02 is a substrate of the AcrB (Figure A.2). This is why NUNL02 was chosen for the docking studies

A visual inspection of the energy curves (Figure 2) permits to compare the FEBs and binding positions along the z axis between compounds. Tables 1, A.6 and A.7 point out the competition for same binding sites between EtBr and NUNL02. Between the tetrahydropyridine derivatives, the best EI in the laboratory tests was NUNL02: notice how its energy curves (Figure 2a, b, c) match with the ones of EtBr in terms of shape and FEB values. For CPZ, the FEB values are smaller than the ones for EtBr, however, the curves match fairly well. (Figure 2g, h, i)

The energy curve of EtBr against tetracycline was made only to compare the FEB values and z average coordinates of binding of these compounds (Figure 2d, e, f). Note that the FEBs have the same magnitude, with EtBr having better values in many positions. This justifies the use of EtBr, a cheap fluorescent compound, as a model for tetracycline efflux studies.

Energy curves made only for one substrate, show immediately the z coordinate of the best FEB values, which correspond to the deeper valleys in Figure 2 graphics. For example, for the

EtBr docked in 4DX5-A (Figure 2a), valleys are found in boxes 270, 320 and 360 with FEB values of -9.4, -8.2 and -7.6 kcal/mol.

When energy curves for different compounds are plotted in the same graphic, one can verify if their best FEBs are in the same or different z position and compare their magnitudes. For instance, the energy curves for EtBr against NUNL02 (Figure 2a, b, c) and CPZ (Figure 2g, h, i) suggest that substrate and ligand compete for the same region. To have a better insight on the hypothesis of competition as the mechanism of efflux inhibition, Tables 1, A.6 and A.7 must be used. In these tables, what calls attention is the fact that almost for all boxes where the positions of best FEB for EtBr were found, there is at least one box, but generally more, where NUNL02 made, in its best docking positions, bindings to protein residues that also are preferred by EtBr. This result is a strong indication that NUNL02 and EtBr are competing for the same or very near binding sites. For instance, in box 260 NUNL02 and EtBr found its best binding position in the same site. In the right column of Table 2, notice that the compounds share 7 residues in common (numerator) and that EtBr makes a total of 7 contacts with the protein residues (denominator). For EtBr in box 345, notice that five NUNL02 boxes had binding positions with common protein residues. Also, for structure 4DX5-A only two best binding positions of EtBr do not have any common residues with NUNL02, in boxes 340 and 370 (Table 1). Tables A.6 and A.7 also indicate that EtBr and NUNL02 are competing for similar protein sites.

The valleys, in the EtBr energy curves, correspond to places of best FEBs and could indicate important regions in the AcrB for a successful transport of substrates. Figure 2a, b and c show the valleys for EtBr and NUNL02 while Tables 3, A.8 and A.9 specify the z coordinates and the FEB of the valleys for EtBr, and the FEB for NUNL02 at these valleys. Notice that the FEB values for EtBr and NUNL02 are of the same magnitude for dockings at the same regions.

Although EtBr and NUNL02 are occupying the same spatial region most of the times, sometimes there is no superposition between them, as can be seen in Figure A.3a in boxes 295 and 320, in Figure A.3b boxes 325, 330, 335 and in Figure A.3c boxes 280, 285, 315 and 355. However, just visualizing this data would be misleading because in spite of not being superposed, the compounds can still be competing for same protein residues as is the case here. For instance, consulting Table 1, one can see that for the EtBr docked in box 295, there is one common protein residue with NUNL02 docked in box 305, a competition not for the same region, but for the same AcrB residue. The presence of NUNL02 in box 305 could collapse the region of box 295 where EtBr is docked, impairing EtBr efflux.

DISCUSSION

In this study, we found that some tetrahydropyridines have potential to be used as drug efflux inhibitors in *E. coli*. This finding was supported by real-time fluorometric assays that have been shown to be a sensitive tool for monitoring the accumulation and efflux^{35,36}.

NUNL02, NUNL09, NUNL10 and NUNL18 presented the best efflux inhibition activity, similar to CPZ, and were not toxic for eukaryotic cells. Out of these compounds, NUNL02 deserves special attention because it only showed inhibitory activity in the *E. coli* strains with a functional AcrB, suggesting that it is a strong EI candidate with AcrB specificity. This is highlighted because no NUNL02 efflux inhibitory activity was detected for AG100A, with a non-functional AcrB. In contrast, mainly for AG100, NUNL09 and NUNL10 presented efflux inhibition activity.

Molecular docking studies were performed to further elucidate the mechanics of the tetrahydropyridine derivative NUNL02 in efflux inhibition. The docking results suggest that the activity demonstrated by NUNL02 in the accumulation assay was most probably due to

efflux inhibition by competition for same binding sites and protein residues with EtBr in the AcrB transporter.

If a relevant binding site for efflux is occupied by a compound, the other cannot be at the same position, and so its efflux will be retarded until the first compound leaves this binding place. Thus, efflux could be being inhibited by a mechanism of competition. The energy curves of EtBr and NUNL02 have a good match in terms of shape, which refers to binding position, and FEB magnitudes, reinforcing the hypotheses of competition. On the other hand, the FEB values for EtBr are higher than the ones for CPZ, and this could be discouraging at first, to consider CPZ as an EI in this case. However, the energy curves of these compounds (Figure 2g, h, i) match fairly well and that could be an indication of preference for common binding sites. In our *in vitro* results the best EI was CPZ, a compound which is very likely to be a substrate of the AcrB³⁷. This could be an indication that the value of the FEB is neither the most important, nor the only thing to be considered when analyzing docking results for multidrug efflux pumps, since CPZ has worse binding energies than EtBr. Preliminary information on the matter of competition could be inferred from the energy curves by simply verifying if the curves match, or if the binding positions along the *z* axis are about the same.

Traditionally, when analyzing docking results there is a tendency to consider that the EIs with better FEBs would also be the best efflux inhibitors in laboratory experiments. However, on many occasions from 50% to 70% of the best inhibitors identified through simple docking do not correspond to those with good performances under *in vitro* testing³⁸.

It would be an erroneous interpretation to discard CPZ as an inhibitor of EtBr efflux or as an AcrB substrate based only on the FEB values. Nevertheless, the shapes of EtBr and CPZ energy curves have similar contours. This could be an indication of competition between EtBr and CPZ for the same residues and positions, as it was for EtBr and NUNL02 discussed

previously. The energy curves for CPZ and EtBr (Figure 5g, h and i) suggest that having common contact residues with the protein would be more important than the FEB when looking for efflux inhibitors.

Although this finding is preliminary and permits no definitive conclusions, it demands subsequent studies with different inhibitors, substrates and drug efflux proteins to verify whether these curves could be a new useful tool to aid in the interpretation of docking results and to verify whether a compound could be an EI.

In conclusion, herein we showed for the first time the potential of tetrahydropyridines as efflux inhibitors and highlighted compound NUNL02 as an AcrB specific inhibitor. Further, we proposed a novel method to interpret the docking results through the comparison of the shapes and FEBs of the energy distribution curves and suggest that the mechanism behind efflux inhibition could be by competition.

EtBr 4DX5-A BOX	NUNL02 BOX	NCR/EtR
260	260	7/7
265	260	7/7
270	265 - 270	10/14 – 10/14
275	275 - 280 - 285	8/10 – 8/10 – 2/10
280	275 - 280 - 285	8/10 – 7/10 – 1/10
285	275 - 280 - 285 - 295	1/4 – 2/4 – 4/4 – 1/4
290	285 - 290 - 295	1/7 – 6/7 – 6/7
295	290 - 295 - 305	5/9 – 6/9 – 1/9
300	290 - 295	6/7 – 5/7
305	310	5/11

310	310	5/10
315	315 - 320 - 325	10/14 – 9/14 – 2/14
320	315 - 320 - 325	9/13 – 8/13 – 2/13
325	325 - 330	5/6 – 2/6
330	325 - 330	5/6 – 2/6
335	325 - 330 - 335	1/10 – 7/10 – 2/10
340	-	-
345	335 - 340 - 345 - 350 - 355	1/11 – 2/11 – 8/11 – 7/11 – 1/11
350	335 - 340 - 345 - 350	2/11 – 3/11 – 7/11 – 7/11
355	355 - 360	7/8 – 7/8
360	355 - 360	7/9 – 7/9
365	365	6/8
370	-	-
375	375	2/7
380	370 - 380	2/8 – 2/8

Table 1: 4DX5-A. Left column - number of the box where EtBr was docked. Middle column: number of the boxes in which there exist common binding between EtBr and NUNL02. Right column: relative to NUNL02 boxes in the middle column, the respective number of common residues per number of EtBr bindings.

Compounds	AG100		AG100 _{TET}	
	Without glucose	With glucose	Without glucose	with glucose
NUNL02	73.99% (± 0.04)	68.24% (± 13.0)	130.5% (± 14.7)	74.6% (± 3.9)
NUNL09	120.4% (± 29.1)	18.5% (± 0.71)	145.4% (± 13.3)	54.0% (± 11.4)
NUNL10	106.1% (± 3.04)	23.9% (± 1.28)	144.8% (± 10.4)	71.2% (± 36.7)
CPZ	100% (± 0)	81.0% (± 6.6)	100% (± 0)	41.6% (± 0.2)

Table 2: Evaluation of the capacity of NUNL02, NUNL09 and NUNL10 to inhibit EtBr efflux in the *E. coli* strains AG100 and AG100_{TET} after the promotion of a maximum EtBr accumulation in the presence of CPZ.

4DX5-A EtBr			NUNL02	
Box	z coordinate	FEB	Box	FEB
270	267.034	-9.4	265 - 270	-9.3/-9.2
295	295.858	-7.8	290 - 295 - 305	-7.9/-7.7/-6.7
320	314.852	-8.2	315 - 320 - 325	-7.7/-7.8/-7.2
355 - 360	~356.90	-7.6	355 - 360	-7.9/-7.6

Table 3: The boxes with the z coordinate of the center of bindings for the minimum FEBs, the valleys in the energy curves, for EtBr in structure 4DX5-A. Also in the table, the corresponding boxes where NUNL02 and EtBr contacted common protein residues. The FEBs are listed for both compounds.

Funding

This work was supported by project 404081/2012-6 Apoio a Projetos de Pesquisa / Chamada MCTI/CNPq/MS-SCTIE - Decit N° 40/2012 - Pesquisa em Doenças Negligenciadas, Brazil, PTDC/BIA-MIC/121859/2010 from Fundação para a Ciência e a Tecnologia (FCT), Portugal. Sofia Santos Costa was supported by grants SFRH/BD/44214/2008 and SFRH/BPD/97508/2013 from FCT, Portugal. Lande Silva and

Daniela Ramos were supported by doctoral and post-doctoral grants, respectively, from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior/Ministério da Educação (CAPES/MEC), Brazil.

Reference List

1. Paphitou NI. Antimicrobial resistance: action to combat the rising microbial challenges. *Int J Antimicrob Agents* 2013; **42 Suppl**: S25-S28.
2. Higgins CF. Multiple molecular mechanisms for multidrug resistance transporters. *Nature* 2007; **446**: 749-57.
3. Blair JM, Piddock LJ. Structure, function and inhibition of RND efflux pumps in Gram-negative bacteria: an update. *Curr Opin Microbiol* 2009; **12**: 512-9.
4. Perez A, Poza M, Fernandez A, et al. Involvement of the AcrAB-TolC efflux pump in the resistance, fitness, and virulence of *Enterobacter cloacae*. *Antimicrob Agents Chemother* 2012; **56**: 2084-90.
5. Nakajima A, Sugimoto Y, Yoneyama H, et al. High-level fluoroquinolone resistance in *Pseudomonas aeruginosa* due to interplay of the MexAB-OprM efflux pump and the DNA gyrase mutation. *Microbiol Immunol* 2002; **46**: 391-5.
6. Ricci V, Tzakas P, Buckley A, et al. Ciprofloxacin-resistant *Salmonella enterica* serovar Typhimurium strains are difficult to select in the absence of AcrB and TolC. *Antimicrob Agents Chemother* 2006; **50**: 38-42.
7. Lomovskaya O, Watkins W. Inhibition of efflux pumps as a novel approach to combat drug resistance in bacteria. *J Mol Microbiol Biotechnol* 2001; **3**: 225-36.
8. Taylor MD, Badger EW, Steffen RP, et al. 2-(2-Aryl-2-oxoethylidene)-1,2,3,4-tetrahydropyridines. Novel isomers of 1,4-dihydropyridine calcium channel blockers. *J Med Chem* 1988; **31**: 1659-64.
9. Ewing TJ, Makino S, Skillman AG, et al. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *J Comput Aided Mol Des* 2001; **15**: 411-28.
10. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 2010; **31**: 455-61.

11. Verdonk ML, Cole JC, Hartshorn MJ, et al. Improved protein-ligand docking using GOLD. *Proteins* 2003; **52**: 609-23.
12. Kuntz ID. Structure-based strategies for drug design and discovery. *Science* 1992; **257**: 1078-82.
13. Mandal S, Moudgil M, Mandal SK. Rational drug design. *Eur J Pharmacol* 2009; **625**: 90-100.
14. Lengauer T, Rarey M. Computational methods for biomolecular docking. *Curr Opin Struct Biol* 1996; **6**: 402-6.
15. Meng XY, Zhang HX, Mezei M, et al. Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des* 2011; **7**: 146-57.
16. Yuriev E, Ramsland PA. Latest developments in molecular docking: 2010-2011 in review. *J Mol Recognit* 2013; **26**: 215-39.
17. Zanatta N, da Fernandes LS, Munchen S, et al. An Efficient Synthesis of Oxa- and Aza-Condensed Tetrahydropyridines from Cyclic Enones. *Synthesis* 2010; **2010**: 2348-54.
18. Zanatta N, Fernandes LS, Nachtigal FM, et al. Highly Chemoselective Synthesis of 6-Alkoxy-1-alkyl(aryl)-3-trifluoroacetyl-1,4,5,6-tetrahydropyridines and 1-Alkyl(aryl)-6-amino-3-trifluoroacetyl-1,4,5,6-tetrahydropyridines. *European Journal of Organic Chemistry* 2009; **2009**: 1435-44.
19. Okusu H, Ma D, Nikaido H. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of Escherichia coli multiple-antibiotic-resistance (Mar) mutants. *J Bacteriol* 1996; **178**: 306-8.
20. Viveiros M, Dupont M, Rodrigues L, et al. Antibiotic stress, genetic response and altered permeability of E. coli. *PLoS One* 2007; **2**: e365.
21. Viveiros M, Jesus A, Brito M, et al. Inducement and reversal of tetracycline resistance in Escherichia coli K-12 and expression of proton gradient-dependent multidrug efflux pump genes. *Antimicrob Agents Chemother* 2005; **49**: 3578-82.
22. Takatsuka Y, Chen C, Nikaido H. Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of Escherichia coli. *Proc Natl Acad Sci U S A* 2010; **107**: 6559-65.
23. Machado L, Spengler G, Evaristo M, et al. Biological activity of twenty-three hydantoin derivatives on intrinsic efflux pump system of Salmonella enterica serovar Enteritidis NCTC 13349. *In Vivo* 2011; **25**: 769-72.
24. Paixao L, Rodrigues L, Couto I, et al. Fluorometric determination of ethidium bromide efflux kinetics in Escherichia coli. *J Biol Eng* 2009; **3**: 18.

25. Viveiros M, Rodrigues L, Martins M, et al. Evaluation of efflux activity of bacteria by a semi-automated fluorometric system. *Methods Mol Biol* 2010; **642**: 159-72.
26. Viveiros M, Martins A, Paixao L, et al. Demonstration of intrinsic efflux activity of *Escherichia coli* K-12 AG100 by an automated ethidium bromide method. *Int J Antimicrob Agents* 2008; **31**: 458-62.
27. Pavan FR, da S Maia PI, Leite SR, et al. Thiosemicarbazones, semicarbazones, dithiocarbazates and hydrazide/hydrazones: anti-*Mycobacterium tuberculosis* activity and cytotoxicity. *Eur J Med Chem* 2010; **45**: 1898-905.
28. Morris GM, Huey R, Lindstrom W, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 2009; **30**: 2785-91.
29. Sanner MF. Phython: A Programming Language for Software Integration and Development. *J Mol Graphics Mod* 1999; **17**: 61.
30. Eicher T, Cha HJ, Seeger MA, et al. Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop. *Proc Natl Acad Sci U S A* 2012; **109**: 5687-92.
31. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 2004; **25**: 1605-12.
32. Case DA, Berryman JT, Betz RM, et al. AMBER 2015. In. University of California, San Francisco, 2015.
33. Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* 2011; **51**: 2778-86.
34. The PyMOL Molecular Graphics System, Version 1.8 Schrödinger LLC. In: 2016.
35. Costa SS, Falcao C, Viveiros M, et al. Exploring the contribution of efflux on the resistance to fluoroquinolones in clinical isolates of *Staphylococcus aureus*. *BMC Microbiol* 2011; **11**: 241.
36. Rodrigues L, Wagner D, Viveiros M, et al. Thioridazine and chlorpromazine inhibition of ethidium bromide efflux in *Mycobacterium avium* and *Mycobacterium smegmatis*. *J Antimicrob Chemother* 2008; **61**: 1076-82.
37. Bailey AM, Paulsen IT, Piddock LJ. RamA confers multidrug resistance in *Salmonella enterica* via increased expression of *acrB*, which is inhibited by chlorpromazine. *Antimicrob Agents Chemother* 2008; **52**: 3604-11.

38. Trotoev M, Abagyan R. Flexible ligand docking to multiple receptor conformations: a practical alternative. *Current opinion in structural biology* 2008; **18**: 178-84.

Figures

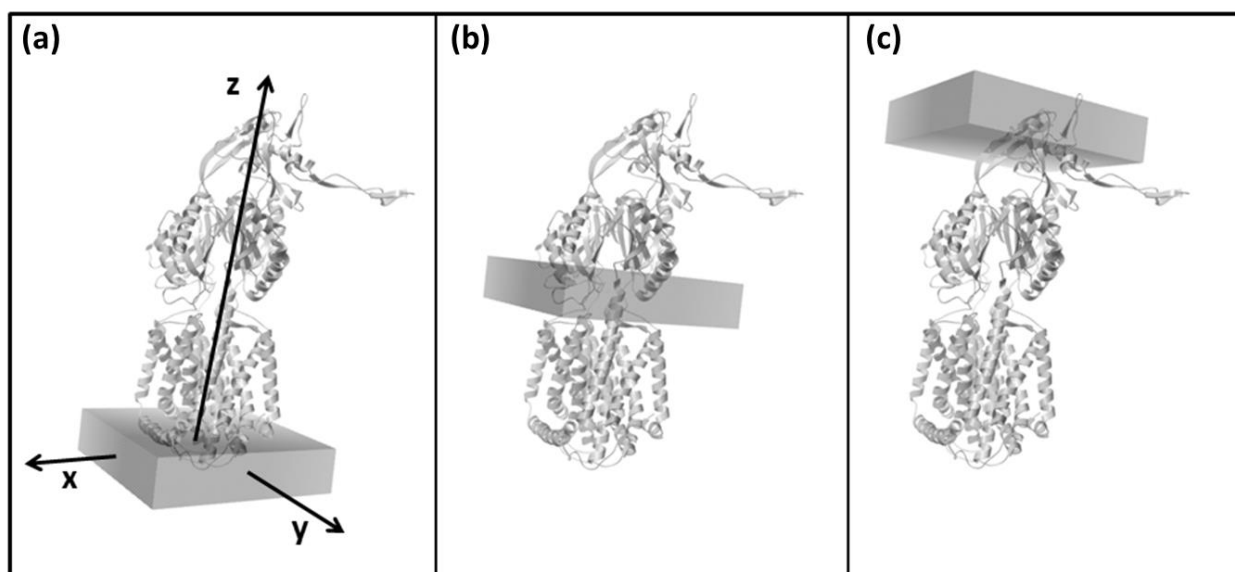


Figure 1: The search box is shown in three different positions. The height of the protein is defined in the direction z . a) The x , y and z directions defined at box number 260, the bottom of the protein and the starting position for the dockings. b) The box positioned about half of the height of the protein. c) Box 380, the last one, at the top of the protein. The dimensions chosen for x , y and z allow the ligands to assume all possible orientations and configurations inside the boxes.



Figure 2: Energy distribution curves. Vertical axis: FEB in kcal/mol. Horizontal axis: z average coordinates for the center of bindings. At structures 4DX5-A, B and C: (a) (b) (c) - Binding energies of EtBr against NUNL02; (d) (e) (f) - EtBr against tetracycline; (g) (h) (i) - EtBr against CPZ.

Appendix A

Table A.1. Tetrahydropyridines structural and molecular formula.

Compound	Molecular Formula	Compound	Molecular Formula
1	$C_{21}H_{21}F_3N_2O_3$	12	$C_{10}H_{14}F_3NO_2$
2	$C_{23}H_{24}F_3N_2O$	14	$C_{11}H_{14}F_3NO_2$
3	$C_{21}H_{21}F_3N_2O$	15	$C_{10}H_{12}F_3NO_2$
5	$C_{21}H_{21}F_3N_2O$	16	$C_{11}H_{14}F_3NO_2$
6	$C_{19}H_{17}F_3N_2O$	17	$C_9H_{11}F_3N_2O$
7	$C_{13}H_{10}F_3N_2O$	18	$C_{11}H_{15}F_3N_2O$
8	$C_{10}H_{12}F_3NO_2$	19	$C_{10}H_{13}F_3N_2O$
9	$C_{15}H_{16}F_3NO_2$	21	$C_{10}H_{13}F_3N_2O$
10	$C_{16}H_{18}F_3NO_2$	22	$C_9H_{10}F_3NO_2$

Table A.2. Concentrations in units of mg/l used for Real-Time Fluorometric Method.

	AG100	AG100 _{TET}	AG100 _A
Chlorpromazine	30	30	10
Verapamil	200	200	200
NUNL02	6.25	100	100
NUNL03	50	100	100
NUNL10	50	100	100
NUNL18	18	18	18
All other NUNL	100	100	100

Table A.3. Number of the box in Structure 4DX5-A where the docking was done, average z coordinate of the bindings between the ligand atoms and the atoms of the protein residues, and the free energy of binding values calculated in each box for the ethidium bromide and compound NUNL02.

4DX5-A Box	Average z		FEB (kcal/mol)	
	EtBr	NUNL02	EtBr	NUNL02
260	263.503	262.505	-6.6	-6.2
265	263.611	266.39	-7.	-9.3
270	267.034	266.735	-9.4	-9.2
275	278.003	277.997	-6.9	-6.5

280	277.909	278.799	-7.2	-7.
285	285.482	283.001	-6.2	-7.5
290	294.109	293.747	-7.6	-7.9
295	295.858	294.343	-7.8	-7.7
300	295.944	298.413	-7.5	-6.8
305	307.166	308.454	-7.4	-6.7
310	307.189	307.254	-7.4	-6.5
315	314.845	315.307	-8.1	-7.7
320	314.852	315.375	-8.2	-7.8
325	326.298	324.29	-7.1	-7.2
330	326.288	330.079	-7.1	-7.2
335	333.75	337.933	-7.4	-6.2
340	336.885	342.095	-7.	-6.8
345	349.406	348.495	-6.9	-7.4
350	348.994	348.173	-7.	-7.2
355	356.866	355.243	-7.6	-7.9
360	356.969	356.72	-7.6	-7.6
365	360.716	363.719	-6.	-5.6
370	366.629	370.174	-5.4	-4.8
375	374.692	370.26	-5.3	-4.8
380	379.509	379.44	-5.2	-4.1

Table A.4. Number of the box in Structure 4DX5-B where the docking was done, average z coordinate of the bindings between the ligand atoms and the atoms of the protein residues, and the free energy of binding values calculated in each box for the ethidium bromide and compound NUNL02.

4DX5-B Box	Average z		FEB (kcal/mol)	
	EtBr	NUNL02	EtBr	NUNL02
260	262.587	261.665	-6.8	-6.5
265	265.647	263.755	-8.3	-7.7
270	267.27	267.954	-7.6	-7.1
275	275.944	278.597	-6.2	-5.2
280	280.684	279.981	-6.2	-6.4
285	285.942	280.327	-6.1	-6.2
290	293.437	291.185	-7.6	-7.9
295	293.078	293.84	-7.5	-7.2
300	299.453	297.787	-7.4	-6.9
305	309.009	309.006	-8.2	-6.6
310	310.568	310.866	-8.6	-8.5
315	315.707	316.951	-8.4	-7.6
320	316.005	317.343	-8.4	-8.8
325	327.166	326.031	-9.5	-8.8
330	327.189	329.047	-9.5	-9.8
335	331.686	330.752	-8.8	-9.1
340	337.699	337.898	-8.7	-8.6
345	347.671	348.624	-7.	-7.9
350	354.833	354.395	-7.1	-6.7
355	355.407	355.966	-7.7	-6.7
360	354.617	355.522	-7.2	-7.6

365	362.588	361.829	-5.7	-5.3
370	375.08	370.314	-5.4	-5.5
375	375.743	369.797	-5.6	-5.3
380	379.228	379.794	-5.1	-4.2

Table A.5. Number of the box in Structure 4DX5-C where the docking was done, average z coordinate of the bindings between the ligand atoms and the atoms of the protein residues, and the free energy of binding values calculated in each box for the ethidium bromide and compound NUNL02.

4DX5-C Box	Average z		FEB (kcal/mol)	
	EtBr	NUNL02	EtBr	NUNL02
260	259.698	260.304	-5.7	-5.4
265	265.776	267.234	-6.2	-6.1
270	272.037	270.202	-6.	-5.3
275	277.853	279.524	-6.4	-7.2
280	283.505	280.642	-7.6	-7.1
285	283.48	285.637	-7.6	-7.
290	292.553	290.248	-6.4	-7.
295	295.913	291.152	-6.6	-6.3
300	299.53	298.436	-6.8	-7.1
305	307.473	308.044	-7.1	-6.7
310	313.716	310.122	-8.6	-6.9
315	314.762	313.586	-9.	-9.
320	317.813	315.913	-7.6	-9.6
325	325.447	326.513	-7.9	-7.7
330	325.896	326.513	-7.7	-7.

335	332.808	335.113	-6.9	-5.5
340	338.081	339.236	-6.2	-5.3
345	347.537	349.288	-6.8	-5.8
350	348.011	350.599	-6.8	-7.7
355	355.701	354.585	-8.1	-7.6
360	355.335	355.354	-7.4	-7.1
365	361.613	363.758	-6.	-4.8
370	373.883	371.489	-5.3	-4.7
375	378.397	378.278	-5.5	-5.9
380	378.369	378.841	-5.6	-4.2

Table A.6: 4DX5-B. Left column: number of the ethidium bromide box. Middle column: number of the boxes in which there exist common protein binding residues between the ethidium bromide and NUNL02. Right column: relative to the NUNL02 boxes in the middle column, the respective number of common residues per number of ethidium bindings.

Ethidium Bromide 4DX5-B BOX	NUNL02 4DX5-B BOX	C.R/Et.R
260	260 - 265	9/10 – 7/10
265	260 - 265	4/7 – 4/7
270	270	2/14
275	-	-
280	-	
285	275 - 280 - 285 - 290 - 295	2/7 – 2/7 – 3/7 – 2/7 – 1/7
290	290 - 295 - 300	7/7 – 5/7 – 2/7
295	290 - 295 - 300	8/8 – 7/8 – 4/8

300	290 - 295 - 300	3/10 – 5/10 – 8/10
305	300	2/13
310	300	1/10
315	315 - 320 - 325 - 330	7/13 – 10/13 – 2/13 – 1/13
320	315 - 320 - 325 - 330	9/15 – 11/15 – 1/15 – 1/15
325	315 - 320 - 325 - 330 - 335 - 340	2/10 – 3/10 – 9/10 – 8/10 – 8/10 – 4/10
330	315 - 320 - 325 - 330 - 335 - 340	2/9 – 3/19 – 8/9 – 8/9 – 7/9 – 3/9
335	325 - 330 - 335 - 340	5/7 – 5/7 – 5/7 – 5/7
340	325 - 330 - 335 - 340 - 345	1/8 – 4/8 – 3/8 – 8/8 – 1/8
345	-	-
350	355 - 360	1/7 – 4/7
355	350 - 355 - 360	1/8 – 4/8 – 7/8
360	345 - 350 - 355 - 360	1/8 – 1/8 – 3/8 – 6/8
365	365	5/6
370	380	1/5
375	370 - 380	1/5 - 1/5
380	370 - 380	1/7 – 5/7

Table A.7: 4DX5-C. Left column: number of the ethidium bromide box. Middle column: number of the boxes in which there exist common binding residues of the protein between the ethidium bromide and NUNL02. Right column: for each NUNL02 box listed in the middle column, the respective number of common residues relative to number of ethidium binding residues.

Ethidium Bromide 4DX5-C BOX	NUNL02 4DX5-C BOX	C.R/Et.R
260	260	7/7
265	260	5/9
270	-	-
275	280 - 285 - 295 - 300	3/4 – 4/4 – 1/4 – 1/4
280	280 - 285 - 290 - 295 - 300	3/4 – 4/4 – 1/4 – 1/4 – 1/4
285 = 280		
290	285 - 290 - 295	2/8 – 6/8 – 6/8
295	290 - 295	3/6 – 3/6
300	300	5/5
305	305 - 310	6/11 – 6/11
310	315 - 320 - 325	8/11 – 6/11 – 1/11
315	315 - 320 - 325	8/11 – 8/11 – 1/11
320	-	-
325	320 - 325 - 330	2/12 – 8/12 – 7/12
330	320 - 325 - 330	1/11 – 6/11 – 7/11
335	325 - 330	1/12 – 3/12
340	340 - 350	4/9 – 2/9
345	355	2/11
350	345 - 350	1/11 – 6/11

355	345 - 350 - 355 - 360	1/9 – 1/9 – 6/9 – 7/9
360	345 - 355 - 360	1/8 – 5/8 – 6/8
365	-	-
370	370 - 375	6/7 – 1/7
375	375 - 380	6/8 – 7/8
380 = 375	375 - 380	

Table A.8: The boxes with the z coordinate of the center of bindings for the minimum FEBs, the valleys in the energy curves, for EtBr in structure 4DX5-B. Also in the table, the corresponding boxes where NUNL02 and EtBr contacted common protein residues. The FEBs are listed for both compounds.

4DX5-B EtBr			NUNL02	
Box	z coordinate	FEB	Box	FEB
265	265.647	-8.3	260 - 265	-6.5/-7.7
325 - 330	~327.20	-9.5	315 up to 340	-7.6/-8.8/-8.8/-9.8/-9.1/-8.8
355	355.407	-7.7	350 - 355 - 360	-6.7/-6.7/-7.6

Table A.9: The boxes with the z coordinate of the center of bindings for the minimum FEBs, the valleys in the energy curves, for EtBr in structure 4DX5-C. Also in the table, the corresponding boxes where NUNL02 and EtBr contacted common protein residues. The FEBs are listed for both compounds.

4DX5-C EtBr			NUNL02	
Box	z coordinate	FEB	Box	FEB
280 - 285	~283.5	-7.6	280 up to 300	-7.1/-7.0/-7.0/-6.3/-7.1
315	314.762	-9.0	315 - 320 - 325	-9.0/-9.6/-7.7
355	355.701	-8.1	345 up to 360	-5.8/-7.7/-7.6/-7.1

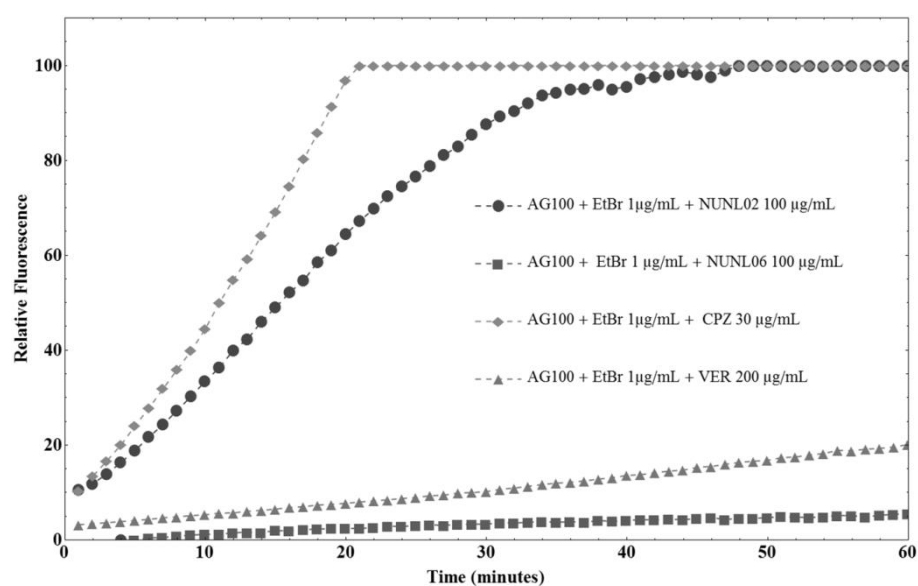


Figure A.1. Effect of the tetrahydropyridine compounds NUNL02 and NUNL06 on comparison with reference inhibitors CPZ and VER on the accumulation of EtBr by *E.coli* strain AG100.

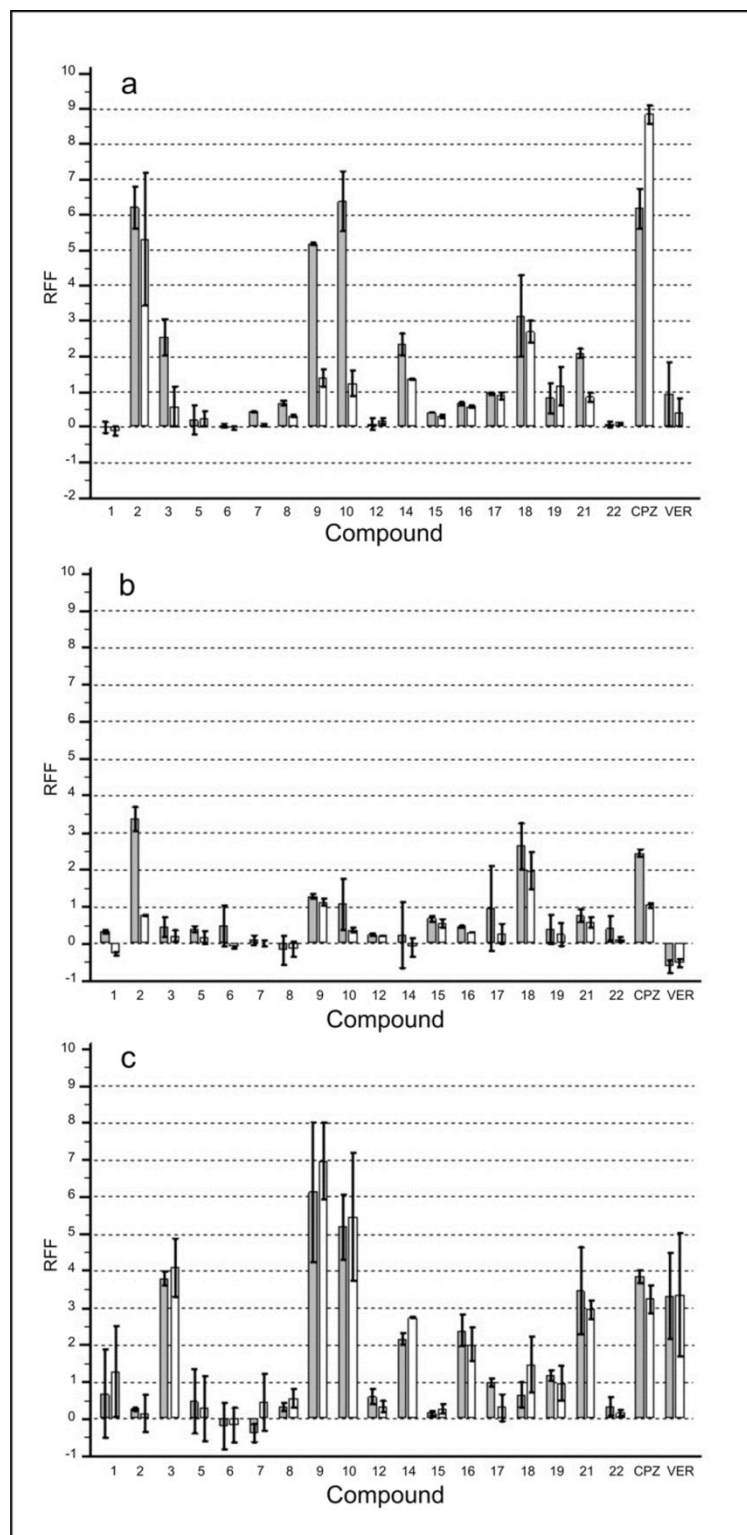


Figure A.2. RFF values based on EtBr accumulation assays for the tetrahydropyridines in the absence (gray bar) and presence (white bar) of glucose (0.4%) for *E. coli* strains AG100 (4a), AG100TET (4b) and AG100A (4c).

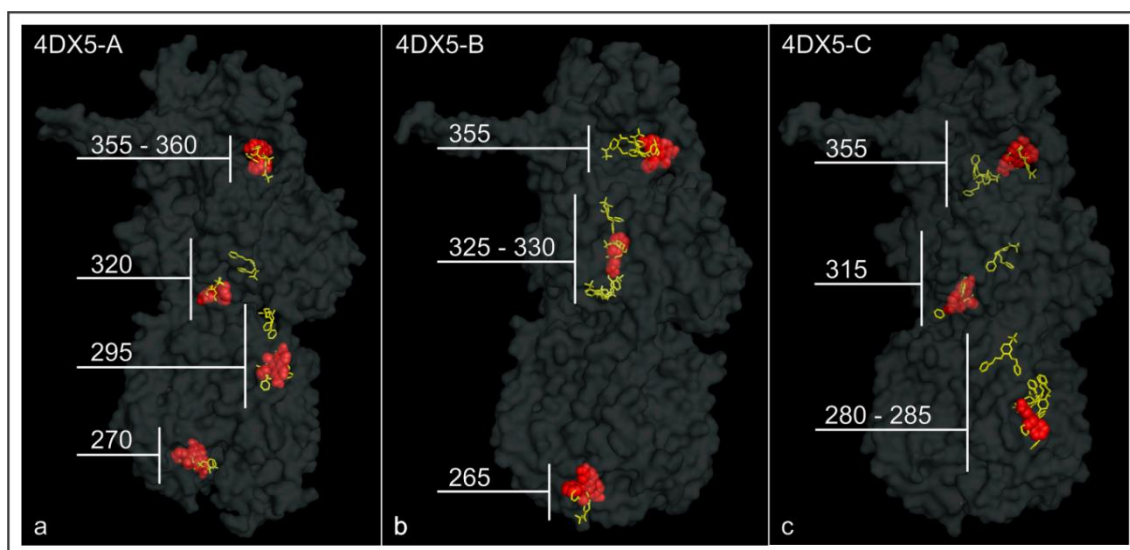


Figure A.3: The data from tables 8, 9 and 10. The EtBr is represented by red spheres in structures 4DX5-A, B and C in the positions corresponding to the valleys (sites of best binding) in the energy curves. NUNL02 is represented by yellow sticks. Notice the superposition between NUNL02 and EtBr indicating competition for the same binding region. Sometimes, as in box 270 and 295 (a), 325 – 330 and 355 (b) and 315 – 355 (c) there is no overlapping, but NUNL02 and EtBr still make some contact with common protein residues.

Fim do artigo 1

ARTIGO 2 – Pronto para submissão ao Jornal PLOS ONE.

Lande Vieira da Silva Jr ^a, Terry P. Lybrand ^b, Pedro Eduardo Almeida da Silva ^a, Karina Machado ^c.

^a*Núcleo de Pesquisa em Microbiologia Médica, Faculdade de Medicina, Universidade Federal do Rio Grande - FURG, Rio Grande, Brasil.*

^b*Center for Structural Biology, Vanderbilt University, Nashville, TN, USA.*

^c*Centro de Ciências Computacionais, Universidade Federal do Rio Grande, Rio Grande, Brasil.*

Targeted Molecular Dynamics Studies on the Mechanism of Efflux Inhibition of the Multidrug Efflux Pump AcrB.

Abstract

The technique of Targeted Molecular Dynamics was performed to the multidrug efflux pump AcrB, structure 1IWG, to study the mechanism of transport of two known substrates, ethidium bromide and a tetrahydropyridine derivative. The studies revealed a set of flexible loops, α -helices and β -sheets that seem to be of importance for the success of efflux.

Introduction

The multidrug efflux pump AcrB from *Escherichia coli*, belonging to the resistance-nodulation division family has been extensively studied as a model for RND efflux pumps that occur in gram-negative bacteria. It is responsible for the capture and extrusion of a wide variety of substrates from the cell (1). Since these efflux pumps contribute to bacterial resistance to many antimicrobial and biocides (2), thus constituting a health concern (3), it is important to use the best methods available for understanding the AcrB capture-extrusion mechanism, in order to inhibit efflux, improve the efficacy of antimicrobials and reduce the resistance level of gram-negative microorganisms.

Targeted Molecular Dynamics (TMD) is a method that induces conformational changes to a structure based solely on geometrical constraints (4). In this procedure, an initial spatial configuration is guided towards a predefined final configuration. TMD has been used before to find conformational changes (5) and pathways (6) due to the presence of ligands in macromolecules. These constraints correspond to a complex, or structure and ligand, in an initial (I) state and then in a final (F) state. The I and F states are known configurations of the

complex, and this knowledge can come from practical studies, i.e. the structure comes from X-Ray diffraction or NMR studies and has already the ligand bonded in different positions. If no structures of such kind exist, the I and F states could also come from docking simulations, as is the case for this work. Based on the docking results obtained for ethidium bromide (EtBr) and a tetrahydropyridine derivative NUNL02 (7, 8) in structure 1IWG, I and F states were defined and then TMD was performed to characterize possible pathways of these ligands while being transported by the protein.

The TMD simulations revealed interesting conformational changes in the protein backbone as the ligand finds its pathway across a series of loops, alpha helices and beta sheets and reinforce the idea of competition as a mechanism of efflux inhibition discussed in (9)(submitted).

Methodology

The structure with PDB: ID 1IWG (10), which has no ligands bound, and is considered to be in the resting state was the one chosen for the TMD studies. This X-ray crystal structure of the AcrB, is the symmetric conformation of the AcrB efflux protein. This ligand free structure has a resolution of 3.5 Angstroms, and 1049 residues for Chain A. The structure is divided in to three domains: a transmembrane domain, that seems to connect the inner cell with the periplasm, and a pore domain and the TolC domain that constitute the periplasmic headpiece (10, 11) (Figure 1). A small part of the amino acid sequence of the structure was not resolved by x-ray diffraction, and then, software Modeller 9.14 (12, 13), with default parameters, was used to construct loops in these missing parts.

Following the docking methodology described in (9) (submitted), Autodock Tools (11) and AutoDock Vina (12) were used to find positions of high affinity between the substrates EtBr, compound NUNL02 and the Modeller completed receptor structure 1IWG, chain A, based on the free energy of binding (FEB) scores. Positions 2 and 7, in the transmembrane domains, 13 and 20 in the pore domains and 28, in the TolC domains, were chosen (Figure 1) (Table 1). Position 2, has the best FEB for the region of the protein that is below the plasma membrane. Position 28 do not present a large magnitude FEB, but it was needed a position outside the structure to be the last F state. Theoretically, once in position 28, the molecule would be already inside the TolC protein. Both, EtBr and NUNL02 are substrates of the multidrug efflux pump AcrB (9, 11, 14). Note that EtBr (red spheres) and NUNL02 (yellow sphere) are superposed for these reference positions.

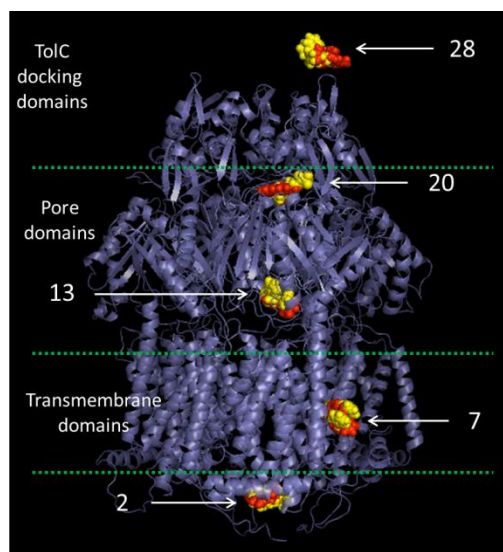


Figure 1: An overview of the positions chosen for the TMD procedure. EtBr is depicted as a red sphere and NUNL02 as a yellow one. The AcrB structure is represented as new cartoon in blue. Position 2 is located outside the structure, under the cell membrane, position 7 is at the side of the protein, sandwiched between the cell membrane and the AcrB in the transmembrane domains, position 13 is inside the structure in the pore domains, position 20 is between the pore and the TolC domains, finally position 28 is out the AcrB and already in the TolC protein.

	Position	2	7	13	20	28
Energy (kcal/mol)	EtBr	-5.8	-7.1	-8.2	-7.4	-2.3
	NUNL02	-7.2	-6.2	-7.6	-6.8	-1.1

Table 1: free energy of binding (FEB) for the positions that compose the pathways for EtBr and NUNL02.

The relatively low values found for the FEB of the positions (Table 1) could be due to the fact that, since structure 1IWG is ligand free, the protein is likely in a resting state, and the sites of high FEB magnitude in the structure are not in the optimal conformation. However, when, following the same methodology, docking was carried out for AcrB structure 4DX5 (15), which is asymmetrical and has ligands bound, the FEB values found with Autodock Vina were not much higher in magnitude (9).

The following pathways were constructed for the TMD simulations.

Pathway 1: for EtBr and NUNL02, with the departing point in a region which is inside the cell, below the plasma membrane, position 2, passing by intermediate positions 13 and 20, which were the positions of strongest binding in the docking procedures, with the final position located outside AcrB, position 28 (Figure 1a). In this simulation, we assume that the ligand could be captured by the AcrB from the cytoplasm, inside the cell (10).

Pathway 2: for EtBr and NUNL02, the departing point is position 7, located at the side of the protein, in a region that is thought to be in the domain of the plasma membrane or immediately above it. The ending point is position 13, located inside the protein. In this simulation it is assumed that the substrate could be captured at the upper transmembrane domain (16).

From the completed structure obtained, a homotrimeric structure was built with Pymol (17). To this structure the substrates were added to the construction the complexes I and F for the TMD procedure. AMBER 12 with packages SANDER, xLeap and ANTECHAMBER was the molecular dynamics software used in this work (18). Then for each complex, with xLeap, missing hydrogens were placed and Na⁺ counter ions were added to neutralize the structure. After, the complex was placed in a truncated octahedron water box. Minimization of the complexes was performed with SANDER in two steps: first the protein is kept restrained and the water molecules were free to move (500 steps of steepest descent and 500 of conjugate gradient, cutoff set to 8) and secondly the whole complex is allowed to move (1000 steps of steepest descent and 1500 steps of conjugate gradient, cutoff set to 10). Force fields ff12SB and gaff were used, respectively, for protein structure and for the substrates. The atomic charges for the substrates were obtained with ANTECHAMBER. Heating from 0 to 300 K (20 ps and protein weakly restrained) and further equilibration at 300 K (100 ps) was performed for each complex before the TMD procedure that was then performed for pathways 1 and 2 with a very low driving force (0.5 kcal/mole per Å), small enough not to brute force the passage of the substrate through unlikely paths.

However, Amber requires the number of water molecules in the periodic water box to be the same in both I and F states. The way to accomplish this was by, initially, building a complex with the ligand bound in both the initial position and in the final position. This complex was named IF. The water box was added automatically and it contains a certain number of water molecules distributed equally around the structure. Then, from complex IF, with xLeap, the final position was removed, leaving the complex as protein plus ligand in the

initial position plus the water box (complex I). Note that the number of water molecules in complex I is the same of complex IF, it was not altered by this procedure. Then, complex I was prepared for the TMD as described previously. The same procedure, but removing the ligand from the initial position instead, allows the construction of a complex with the ligand bonded to the final position in a water box, or complex F with the same number of water molecules in the box as complex I. In this work one complex had to be built for each reference point, resulting in 10 complexes that can be used to create different pathways, as long as they pass at the reference points. For instance, with the 10 complexes generated, it would be possible to create a TMD run to analyze a pathway starting in position 2, then going to 20 and returning to 13. However, only the 2 pathways described previously were studied in this work. The results for region 20-28 in pathway 1 are for the EtBr only.

Results

Pathway 1

The results of TMD for this pathway show that the ligands are superposed most of the time, as they climb from position 2 to position 20, along the efflux protein. Figure 2 shows all frames from the TMD simulations for this pathway for ligands EtBr and NUNL02. Interestingly, from position 20 (Figure 2) up to 28, the paths are different.

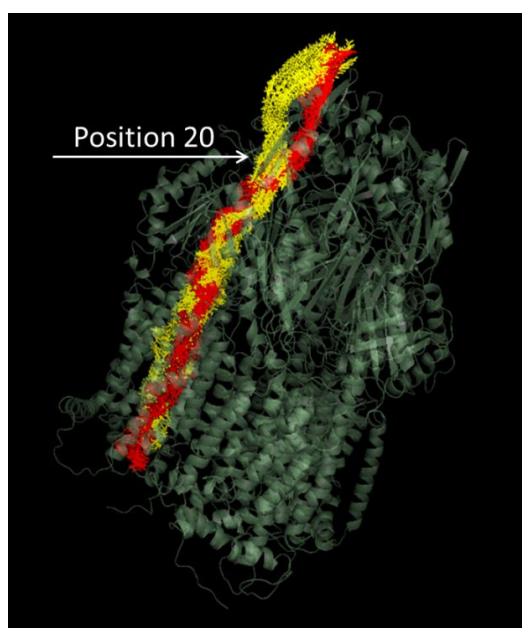


Figure 2: The EtBr and NUNL02 follow the same path inside the AcrB efflux pump, from position 2 to 13 to 20. Only between positions 20 (white arrow) and 28 the path of these substrates diverge.

The simulations with EtBr for pathway 1, revealed a structure that could be thought of as a tunnel, with many flexible loops, sheets and helices that exhibit switch like movements in some occasions. Dividing pathway 1 in three different areas made it easier to locate structures that, in the simulations, seemed to have an important role in transporting the substrates (Figure 3).

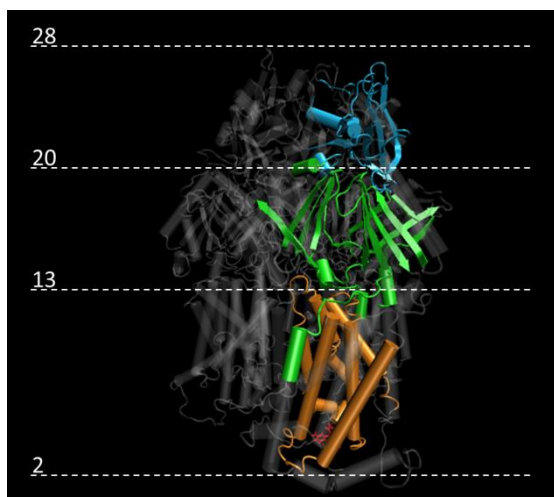


FIGURE 3: pathway 1 and the three regions. Orange: from position 2 to 13 (region 2-13), this region supposedly connects the inner cell with the periplasm, crossing the plasma membrane. It could be imagined as an access tunnel that capture and transport ligands from the cytoplasm. The TMD runs revealed a peristaltic like movement in this region. Green: from position 13 to 20 (region 13-20), located in the periplasm, a set of stiff β -sheets and flexible loops that open, close and adjust to let the substrates be transported. Blue: from position 20 to 28 (region 20-28), still in the periplasm, is the region that is connected to the TolC protein, show β -sheets, flexible loops and helices with switch like movements.

Region 2-13: depicted in Orange (Figure 4), this region is partly embedded in the plasma membrane. This region has a total of 12 transmembrane α -helices (Murakami et al. 2002), but only 9 were represented here for they seem to be the inner core of an entrance tunnel. Table 2 describes the residues in each helix.

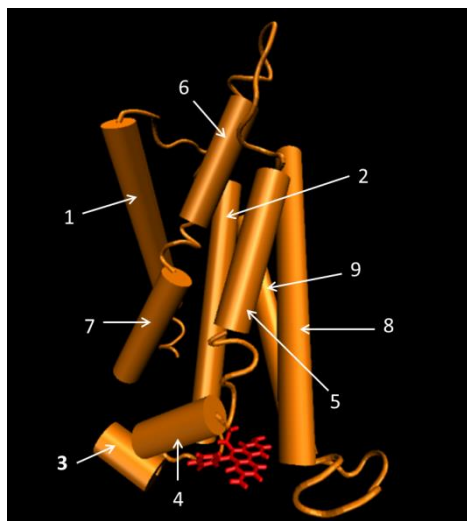


Figure 4: From pathway 1, region 2-13, the helices that form a tunnel structure with EtBr in position 2.

Helix 1	Helix 2	Helix 3	Helix 4	Helix 5	Helix 6	Helix 7	Helix 8	Helix 9
367-386	396-421	428-430	432-436	443-454	466-476	481-493	925-952	964-990

Table 2: residues for the helices in region 2-13.

Position 2 (Figure 5a), could be both a capture and an access point located inside the cell, thus allowing the AcrB efflux pump to extrude molecules that are in the cytoplasm. Also, a peristaltic movement seems to be well characterized in Region 2-13 as can be seen in Figure 5: in Figure 5a the helices are in a tighten position while EtBr is in position 2. The helices are in the transmembrane domain (10). When the molecule is inside the tunnel (Figure 5b) it is noticeable that the tunnel is wider, the helices made space for the passage of the substrate. In Fig 5c the tunnel shrinks again and finally in Figure 5d, when EtBr is almost in position 13, the helices are practically arranged as they were in Figure 5a. The helices in yellow (Figure 5) correspond to helices 6 and 7. These two helices are probably a single one that has a stretchable, flexible region that appears in the cartoon depiction as a loop comprehending residues 478 to 480.

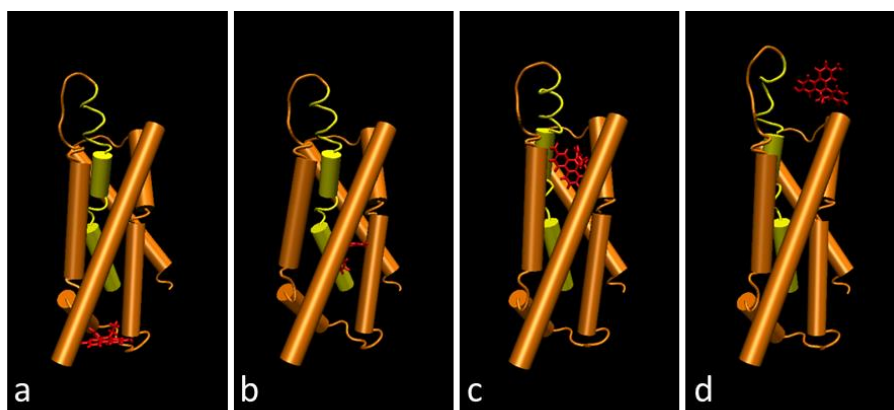


Figure 5: From pathway 1, side view of the tunnel with EtBr located in position 2. Notice the movement of the helices, widening from a to b and then shrinking from b to d.

Figure 6 illustrates three helices from the access tunnel and between positions 2 and 13, helix 6 in yellow and helices 1 and 2 in orange. Although there is a peristaltic like movement in region 2-13 involving all helices, helix 6, in yellow (Figure 6) has a pronounced displacement when compared to the neighbor helices. So, helix 6 could be of great importance for transport by allowing the passage of the substrate to the upper part of region 2-13.

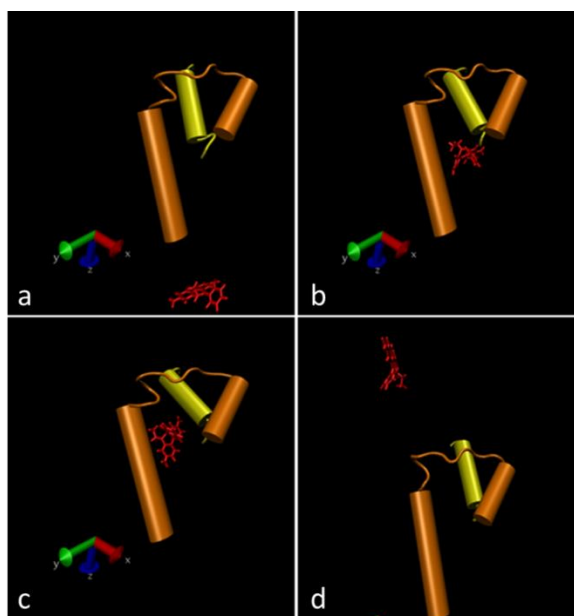


Figure 6: Helix 6 (in yellow), helix 1 and 2 (partially shown). From a to d, note the switch like movement of helix 6.

Region 13-20: this region (Figure 7 in green), in contrast to region 2-13, is mostly composed of beta sheets and flexible loops, it is within a region named pore domain (Murakami et al. 2002). The β -sheets are very stable and do not present great displacements

with the exception of the one indicated by an arrow in Figures 7 and 8a, with the residues in Table 3. In this region, there is still the impression of a tunnel shape, more or less delimited by β -sheets displaced in a way that resembles a thoracic box, giving a visual sensation of stiffness to the region. The flexible parts are indeed the loops. Loops 1 and 2 (Figure 7 and 8), whose residues are listed in Table 2 are on the borderline of region 2-13 and 13-20. They could be thought as the entrance to region 13-20. Loop 1 is the one with more prominent displacement, from Figure 8a to 8d, it is noticeable that it lifts up, giving space for the substrate and then goes down, once the substrate is farther, closing this passage. As loop 1 rises, so does the β -sheet in yellow (Figure 8b and 8c).

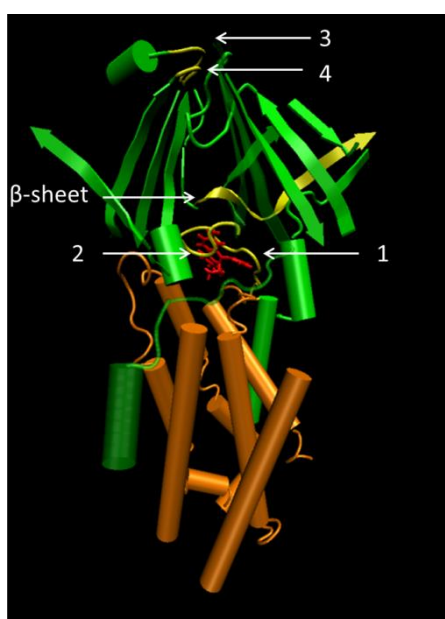


Figure 7: region 13-20 has many beta sheets arranged in a way that resembles a thoracic box, giving a visual feeling of stiffness to this region. Loop 1, 2, 3, 4 and the β -sheet marked presented interesting displacements. EtBr, in red is located right on position 13.

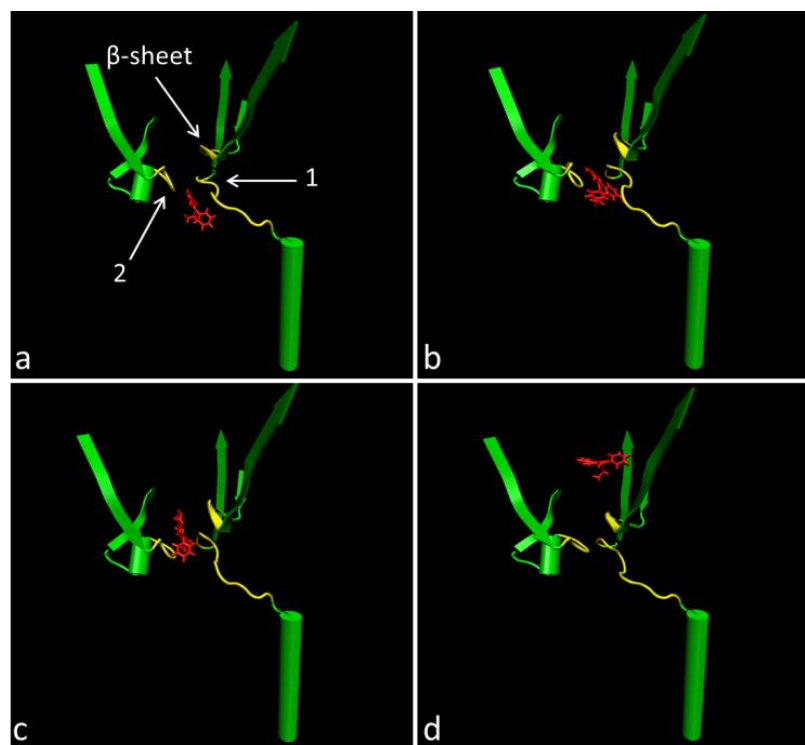


Figure 8: from region 13-20, a detail of loop 1 (arrow 1), loop 2 (arrow 2) and a beta sheet, that show a significant displacement as the substrate enters this region.

Loop 1	Loop 2	β -sheet	Loop 3	Loop 4
30-42	668-678	132-144	49-53	84-85

Table 3: residues of the loops and the β -sheet that seemed to be of importance for the transport in region 13-20.

On the borderline of region 13-20 and 20-28 are loops 3 and 4 (Figure 10, residues in Table 3). These loops are practically at position 20 and they could be thought of as an entrance to region 20-28. In Figure 9 from image a to d it is possible to note that the loop 3 lifts, allowing the substrate to reach region 20-28. Loop 4 lifts a little too (Figure 7 and 15).

Region 20-28: Depicted in blue in Figure 9, this region is in the TolC domain (10). It is composed of β -sheets, small α -helices and many loops. Helices 10 and 11 act like a switch, opening to let the substrate pass and loop 5 stretches allowing the substrate to reach position 28 (Figure 9, Table 5).

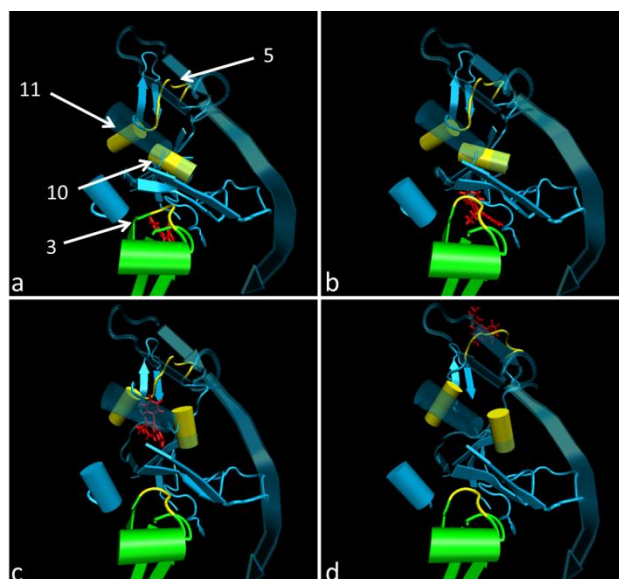


Figure 9: Region 20-28. Loop 3 is on the borderline of region 13-20 and 20-28. Helices 10 and 11, in yellow, are the helices with a more prominent displacement, already in region 20-28. Loop 5 exhibits flexibility and is located at the exit of the protein, near position 28.

Helix 10	Helix 11	Loop 5
752-754	203-208	189-202

Table 5: residues of Helix 10, 11 and loop 5.

The results for EtBr were chosen when looking for interacting residues in this region because NUNL02 moves to the outside of the protein. However, if the AcrB were attached to the TolC protein, as is believed (Seeger et al 11) NUNL02 would probably not behave so.

Pathway 2

Pathway 2 is comprehended in region 2-13 described for pathway 1. The difference is that the substrate capture point is in position 7, outside the structure, in the transmembrane region, sandwiched between the plasma membrane and the AcrB protein (figure 1). Position 13 is the same position described previously for pathway 1. TMD simulations for pathway 2 showed that the ligands do not enter the protein immediately from position 7. Instead, the ligands climb by the side of the protein until finding an entrance a little below the level of position 13 (Figure 11).

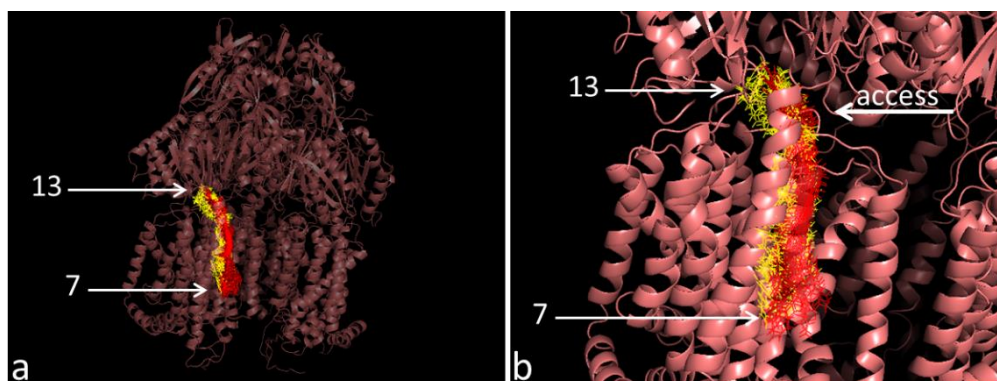


Figure 11: a) EtBr (in red) and NUNL02 (in yellow) at the AcrB structure. TMD revealed that both substrates shared the same path from position 7 to 13. b) Detail of the access position to the interior of the protein.

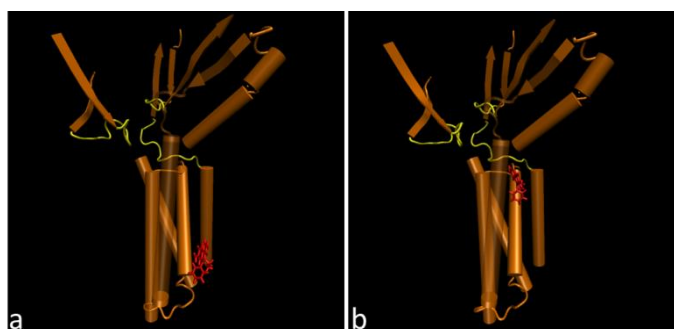
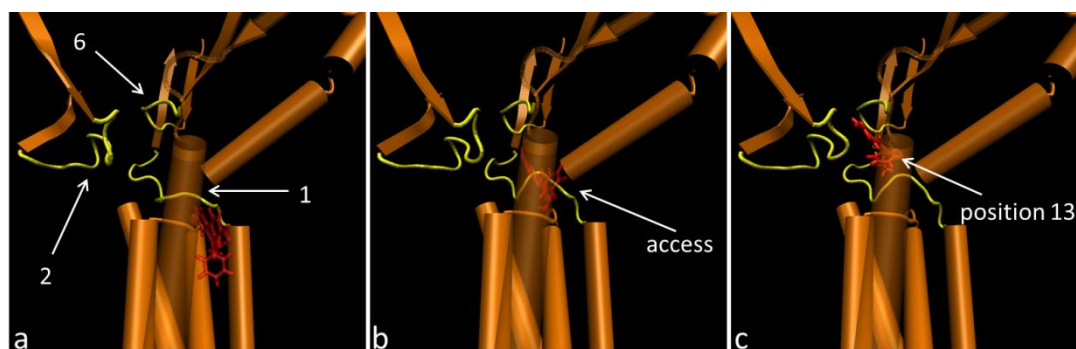


Figure 12: from position 7 (a) until the access point (b) there are not great configuration changes.

From the capture point at position 7, until the entrance of the protein (Figure 12), TMD did not reveal any major conformational change in the protein structure. However, once the substrate access the protein, loop 1 lifts to let the substrate go inside (Figure 13).



Figure

13: from a (immediately before the access) to c (inside the protein). a) Loops 1, 2 and 6. b) At the access point, loop 1 rises to allow the entrance of the substrate, loop 2 changes its

conformation a little. c) Loop 1 remains opened after the substrate reaches position 13.

Once inside the protein, when the substrate leaves position 13 to reach position 20, loop 6 lifts up (Figure 14) allowing the passage of the substrate.

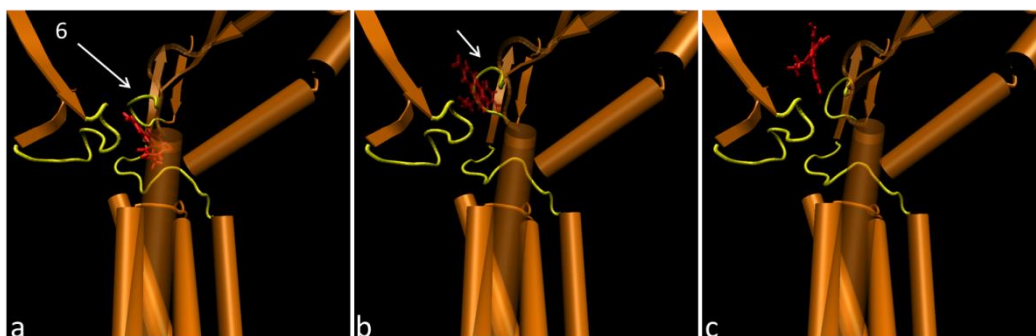


Figure 14: a) Loop 6, when the substrate is at position 13 remains in its original state. b) Once the substrate moves up, loop 6 rises. c) Loop 6 still interacts with the substrate that is on the way to position 20.

Loop 1	Loop 2	Loop 6
30-42	668-678	131-137

Table 3: Loops 1 and 2 (same as described in Table 2) and loop 6.

Discussion

The asymmetric configuration of the AcrB structure suggested a magnificent model for drug transport based on conformational cycling of the monomers between loose, tight and open configurations (11, 19). However, in this study, we used the symmetric structure 1IWG and the TMD results suggest that each monomer would be able to perform alone, independently, capturing and conducting the substrate towards the TolC protein to be extruded. The cycling model for the asymmetric structure would require a lot of coordination compared to a simpler model where the monomers could act individually. Besides, acting independently would probably make the efflux pump more efficient, for coordination is likely to demand extra time and energy. Reality may correspond to something in between models.

The first description of the AcrB structure 1IWG (10) divided it into three domains: transmembrane, pore and TolC. In this work, the division of the structure 1IWG in regions

was not based on what was originally described. Instead, the choice for the regions was based, primarily, on the initial and final positions for the TMD procedures as described, that was based on the best binding sites found by docking procedures. Further, after TMD was carried out, pathway 1 was subdivided in three regions, named 2-13, 13-20 and 20-28. The limiting points of these regions are positions 2, 13, 20 and 28 (Figure 1). There is a morphological difference between regions and there are access points between regions. The approach used in this work to define the regions, defined practically the same regions as the original division. That could be a coincidence or an indication that whenever there are morphological differences between regions of an efflux protein, points of strong binding and access/exit points between regions are likely to be found.

Although AcrB is thought to capture its substrates from the periplasm (20) and from within the outer leaflet of the cytoplasmic membrane (16), it is considered that the spatial configuration of the protein and its positioning in the cell, allows the capture of substrates from the cytoplasm (10). According to TMD performed here, position 2, could be a capture point located inside the cell membrane as well as an access point to region 2-13. This feature would increase the efficiency of the AcrB as a multidrug exporter for it would not depend on other membrane proteins, for instance EmrE and MdfA to transport the substrates from within the cell to the periplasm, where AcrB would then capture them (21). Position 13 of this study is located at the region named porter domain, it is inside the protein and was the position of strongest binding in the dockings. The access to this position, is delimited by loops 1, 2 if the substrate comes from position 2 (Figure 8), and by loops 1, 2 and 6 (Figure13) if the substrate comes from position 7. The set of loops are likely to be an access point to region 13-20, controlling the entrance of substrates region 13-20. Could there be a deep binding pocket at or around position 13, as the one described previously (15)? It is reasonable to think there could be. In this case loops 1, 2 and 6 would be the switch-loops mediating the entrance to this pocket. Still in region 13-20, immediately before entering region 20-28 there are loops 4 and 5, at what could be considering the exit of region 13-20, mediating the passing of the substrate. A switch-loop (Figure 15) that comprehends residues 614 to 623 mediates the entrance of the substrate to a deep binding pocket already described (15). However, loops 3 and 4 (Table 2), located a bit higher than the switch loop (Figure 15), mediate the access to region 20-28. If the switch-loop is considered to be at the entrance of the deep binding pocket, it is reasonable to conceive that loops 3 and 4 are likely to be at the exit of both, the binding pocket and region 13-20.

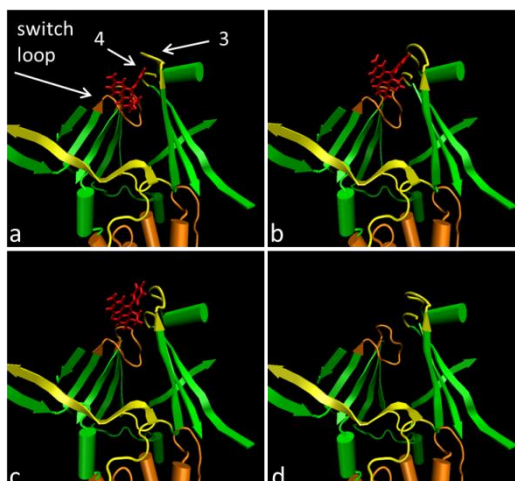


Figure 15: the switch-loop in orange (15) and loops 3 and 4. Notice how loop 3 lifts to allow the passage of the substrate, EtBr (licorice, in red). Loop 4 do not show a major displacement as loop 3, but it does move, as can be noticed comparing images a and b.

As for position 20-28, there are helices 10, 11 and loop 5 (Figure 9, Table 3), that interact with the substrate guiding it toward the TolC protein.

The transmembrane region corresponding to region 2-13 is mostly composed of α -helices, region 13-20, on the other hand, is predominantly composed of β -sheets. Region 20-18 is composed of small helices, sheets and loops.

The tunnel like appearance of pathway 1, between positions 2 and 13 is notable, and suggests that the substrates are likely to follow through this tunnel. The TMD results reinforce this point of view, showing that with a very low driving force, the substrates traverses pathway 1 very naturally. Note, in Figure 5d, that the tunnel seem to reset to its initial position in 5a, suggesting that the efflux pump monomer would be ready to capture another another from the cytoplasm. The tunnel exhibits a peristaltic kind of movement (Figure 5). However, it is hard to tell if the peristaltic movement is responsible for the conduction of the substrate, or if it is, actually, the interaction between substrate and protein residues that causes the peristaltic movement. If the second consideration is true, then the peristaltic movement is the consequence and not the cause of the substrate transport.

Pathway 2 is an interesting one because position 7, likely embedded in the upper layer of the cell membrane, in the transmembrane domain, could be a capture point on the outer surface of the protein. TMD results suggest that the substrate climbs the protein side until it finds an entrance in the periplasm domain and reach position 13. The access point would be at

the same height of position 13. Loop 1 mediates the access to position 13 and loop 6 mediates the exit. Loops 1, 2 and 6 surround position 13, the one with the higher FEB magnitude in the dockings (Table 1). It is tempting to state that position 13 is a position of importance for transport in the AcrB. The possibility of lateral capture of substrates that then would be laterally transported to an access point would increase the efficiency and power of the AcrB efflux protein. The substrate would have a larger area to dock, and not a specific target. This is an advantage when compared to entering the protein from specific access points, that are generally considered to be three for structure 1IWG (22). Thus, maybe any possible binding area on the outer surface the AcrB could be a potential via for its inside.

Also, according to the TMD results, the substrates EtBr and NUNL02 compete for the same trajectories in both pathways studied (Figures 2 and 11). It is known that NUNL02 has great affinity for with the AcrB protein and the results of the simulations reinforce the idea of inhibition by competition between substrates as proposed before (9). In pathway 1, EtBr and NUNL02 compete for the same path until position 28, when then, their ways diverge (Figure 2). Such divergence is a very good result for it indicates that TMD does not force the substrates to follow the same track once the initial and final positions are defined. The same superposition of EtBr and NUNL02 along the path, appear again in the simulations for pathway 2, when the substrates climb laterally until position 13. Considering competition, it would be possible to conceive a good inhibitor as one that has the same or, at least the closest path to the compound supposed to have its efflux inhibited and not only the same docking positions.

TMD results showed no evidence, in agreement with a previous study (20), that the substrates would go through the central pore.

Perhaps, to impair the efflux mechanism, instead of targeting single residues, it would be needed to target the loops, sheets and helices, the greater active parts described here and others that can be found in other studies. A simple substitution of a residue, for example, from an active loop, could not cause the loop to lose its functionality. A possible analogy could be that a hand does not lose its capacity of holding objects if one or two fingers are lost, however, of course some fingers are more important than others. Losing the thumb makes it harder to hold objects. Maybe the same reasoning is valid for residues and the greater parts they are part of. Besides, targeting residues could be more substrate specific than targeting entire flexible loops, sheets and helices that would have to be functional for the successful transport of any substrate. Looking only at the residues, may cause the missing of the greater

important parts. Should we look at the forest (loops, sheets and helices) or the trees (residues)?

Conclusion

The technique of targeted molecular dynamics was used to study how EtBr and NUNL02 could be captured and transported by the AcrB efflux pump. The simulations were done for two different pathways named 1 and 2 and revealed that loops 1, 2 and 6 (region 13-20), 3 and 4 (region 13-20), and 5 (region 20-28), α -helices 1 to 9 (region 2-13), 10 and 11 (region 20-28) and a β -sheet (region 13-20) had an active role in transport, interacting with the substrates as they were being extruded. Further, the simulations revealed that EtBr and NUNL02 compete for the same path both in pathway 1 and 2. This finding suggests that the mechanism of efflux inhibition by competition between molecules is likely. Subsequent studies could reveal which residues from these loops, sheets and helices described here are paramount for transport and so contribute for new strategies to inhibit the AcrB and other RND multidrug efflux pumps.

References

1. Li XZ, Plesiat P, Nikaido H. The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clin Microbiol Rev.* 2015;28(2):337-418.
2. Blair JM, Richmond GE, Piddock LJ. Multidrug efflux pumps in Gram-negative bacteria and their role in antibiotic resistance. *Future Microbiol.* 2014;9(10):1165-77.
3. Paphitou NI. Antimicrobial resistance: action to combat the rising microbial challenges. *Int J Antimicrob Agents.* 2013;42 Suppl:S25-8.
4. Schlitter J, Engels M, Kruger P. Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. *J Mol Graph.* 1994;12(2):84-9.
5. Lee HS, Robinson RC, Joo CH, Lee H, Kim YK, Choe H. Targeted molecular dynamics simulation studies of calcium binding and conformational change in the C-terminal half of gelsolin. *Biochem Biophys Res Commun.* 2006;342(3):702-9.
6. Aci-Seche S, Genest M, Garnier N. Ligand entry pathways in the ligand binding domain of PPARgamma receptor. *FEBS Lett.* 2011;585(16):2599-603.
7. Nilo Zanatta LdSF, Fabiane M. Nachtigall, Helena S. Coelho, Simone S. Amaral, Alex F. C. Flores, Helio G. Bonacorso and Marcos A. P. Martins. Highly Chemoselective Synthesis

- of 6-Alkoxy-1-alkyl(aryl)-3-trifluoroacetyl-1,4,5,6-tetrahydropyridines and 1-Alkyl(aryl)-6-amino-3-trifluoroacetyl-1,4,5,6-tetrahydropyridines. *European Journal of Organic Chemistry*. 2009(9):1435-44.
8. Nilo Zanatta LSdF, Sinara München, Helena S. Coelho, Simone S. Amaral, Leonardo Fantinel, Hélio G. Bonacorso, Marcos A. P. Martins. Highly Chemoselective Synthesis of 6-Alkoxy-1-alkyl(aryl)-3-trifluoroacetyl-1,4,5,6-tetrahydropyridines and 1-Alkyl(aryl)-6-amino-3-trifluoroacetyl-1,4,5,6-tetrahydropyridines. *Synthesis*. 2010(14):2348-54.
 9. Vieira da Silva Jr L, Carrion LL, Groll Av, Costa SS, Junqueira E, Ramos DF, et al. In vitro and in silico analysis of the efficiency of tetrahydropyridines as drug efflux inhibitors in *Escherichia coli*. 2016.
 10. Murakami S, Nakashima R, Yamashita E, Yamaguchi A. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature*. 2002;419(6907):587-93.
 11. Seeger MA, Diederichs K, Eicher T, Brandstatter L, Schiefner A, Verrey F, et al. The AcrB efflux pump: conformational cycling and peristalsis lead to multidrug resistance. *Curr Drug Targets*. 2008;9(9):729-49.
 12. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 1993;234(3):779-815.
 13. Webb B, Sali A. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics*. 2014;47:5 6 1-32.
 14. Takatsuka Y, Chen C, Nikaido H. Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of *Escherichia coli*. *Proc Natl Acad Sci U S A*. 2010;107(15):6559-65.
 15. Eicher T, Cha HJ, Seeger MA, Brandstatter L, El-Delik J, Bohnert JA, et al. Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop. *Proc Natl Acad Sci U S A*. 2012;109(15):5687-92.
 16. Nikaido H, Basina M, Nguyen V, Rosenberg EY. Multidrug efflux pump AcrAB of *Salmonella typhimurium* excretes only those beta-lactam antibiotics containing lipophilic side chains. *J Bacteriol*. 1998;180(17):4686-92.
 17. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. 2015.
 18. D.A. Casee TAD, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R., Luo RCW, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra., J. Swails AWG, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu., X. Wu SRB, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G., Cui DRR, D.H. Mathews,

M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T., Luchko SG, A. Kovalenko, and P.A. Kollman. Amber 12. University of California, San Francisco 2012.

19. Seeger MA, Schiefner A, Eicher T, Verrey F, Diederichs K, Pos KM. Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science*. 2006;313(5791):1295-8.

20. Sennhauser G, Amstutz P, Briand C, Storchenegger O, Grutter MG. Drug export pathway of multidrug exporter AcrB revealed by DARPIn inhibitors. *PLoS Biol*. 2007;5(1):e7.

21. Tal N, Schuldiner S. A coordinated network of transporters with overlapping specificities provides a robust survival strategy. *Proc Natl Acad Sci U S A*. 2009;106(22):9051-6.

22. Eicher T, Brandstatter L, Pos KM. Structural and functional aspects of the multidrug efflux pump AcrB. *Biol Chem*. 2009;390(8):693-9.

Fim do artigo 2.

5 Resultados e Discussões Gerais

5.1 Artigo 1

Nesse artigo, os resultados obtidos através das docagens tornaram possível a construção das curvas de energia (Figura 5) e as tabelas 1, A3 e A4, que indicam a preferência dos substratos pelos sítios de maior magnitude de energia de ligação na proteína. As curvas de energia e as tabelas passam a ser as principais ferramentas de análise das docagens. As curvas de energia facilitam a visualização dos valores das energias nos locais onde os substratos foram docados, indicando haver ou não competição pelo mesmo sítio. As tabelas se mostraram bastante úteis, pois identificam o número de resíduos da proteína contatados em comum pelos substratos mesmo quando aparentemente não há superposição entre os substratos na mesma caixa definida na proteína (Figura A3 do artigo 1).

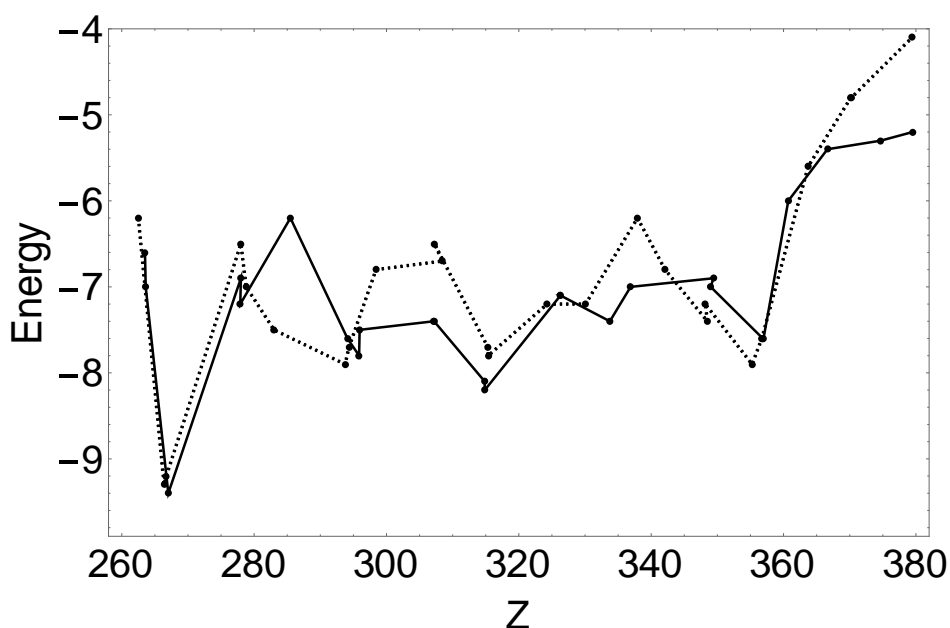


Figura 5: Eixo horizontal – Coordenada z onde se encontrou a melhor energia de ligação. Eixo Vertical – Energia de Ligação em unidades kcal/mol. Brometo de Etídio (linha cheia) e moléculas NUNL02.

Por exemplo, utilizando um programa de visualização, como o PyMol, percebe-se que diversas vezes a superposição entre substratos na AcrB é facilmente verificada, mesmo quando parcialmente. Entretanto, algumas vezes não há superposição, mas mesmos resíduos podem estar sendo compartilhados. Nesse trabalho não foi encontrado um número de resíduos

compartilhados maiores do que um, quando não há superposição. Entretanto, um resíduo ocupado por um ligante, se importante para o sucesso do efluxo, torna-se indisponível momentaneamente para se ligar ao outro substrato. Por isso, as curvas de energia e as tabelas devem ser usadas em conjunto para a análise dos resultados de docagem. Conhecer a quantidade de superposições ou de resíduos em comum entre substrato e composto pode ser um indicativo de que a inibição do efluxo se dá por competição. Sabe-se pelos testes de laboratório descritos no artigo 1 que, tanto o brometo de etídio quanto a molécula NUNL02 são substratos da bomba de efluxo AcrB e que a molécula NUNL02 inibiu o efluxo do brometo de etídio. Então, de acordo com os resultados obtidos, competição passa a ser um provável mecanismo de inibição, conforme proposto. O método de docagem permitiu encontrar, dentre outros, três locais de alta energia de ligação entre substrato e proteína, as posições 7, 13 e 20, que foram definidas e utilizadas no artigo 2.

Por fim, o artigo 1 apresenta também a hipótese de que, a semelhança entre as curvas de energia entre inibidor e substrato seria mais importante do que uma simples análise dos valores da energia livre de ligação, encontrados por docagem. Essa hipótese vem do fato da molécula clorpromazina ser um substrato reconhecido da AcrB (Bailey *et al.*, 2008) e que no entanto, seus valores de energia de ligação não são da mesma magnitude do EtBr. Não é possível explicar esse fato somente baseado em valores. Torna-se necessário buscar outra explicação e a comparação das curvas de energia do EtBr e da clorpromazina sugere que a competição por mesmos sítios, que se reflete na semelhança no desenho das curvas, pode ser uma possível explicação .

5.2 Artigo 2

Para o artigo 2, a importância das curvas de energia foi a de estabelecer as posições de partida e chegada para os substratos EtBr e NUNL02 na AcrB estrutura IHWG otimizada. Foram construídos, então, dois caminhos baseados nos valores de maior magnitude para a energia livre de ligação entre os substratos e a proteína. Nesses caminhos foi realizado o procedimento de TMD (Schlitter *et al.*, 1994), que desloca o substrato na proteína, entre duas posições predeterminadas, seguindo o caminho de menor energia, ou seja, o provável caminho que o substrato percorreria na bomba em uma situação real, com a proteína funcional na célula bacteriana. O primeiro caminho estabelecido, desloca os substratos do interior da célula, do citoplasma até fora da AcrB, no espaço que já corresponderia a proteína TolC. O segundo caminho, tem como ponto de partida uma posição que fica na região transmembrana

plasmática, onde o substrato fica entre a lateral da proteína e a membrana, daí o substrato é deslocado até o interior da célula, na altura da caixa da figura 5b. Com os resultados da TMD, foi possível a identificação de “loops”, folhas beta e hélices alfa que participaram ativamente do transporte dos substratos nas simulações e que não se encontram adequadamente identificados na literatura. É importante salientar que, muito provavelmente, já que não foram encontradas referências na literatura até o momento, é a primeira vez que se utiliza TMD para estudos da bomba de efluxo AcrB.

Os resultados das simulações também reforçaram a hipótese sobre a existência de competição entre inibidor (NUNL02) e substrato (EtBr) por mesmas trajetórias nos caminhos definidos na AcrB. As Figuras 2 e 11 do artigo 2 mostram as trajetórias compartilhadas pelo EtBr e pela NUNL02. No caminho 1, revelou-se movimentos peristálticos, e conjuntos de loops que atuam como mediadores entre regiões da proteína, definidas no artigo 2 que correspondem, praticamente aos domínios definidos na Figura 1. Outro resultado muito interessante foi a trajetória lateral seguida pelos ligantes no caminho 2. A menor trajetória entre as posições determinadas para esse caminho seria uma linha reta, mas para isso o substrato deveria entrar imediatamente na proteína. Entretanto, o substrato escala a superfície lateral da proteína até encontrar uma região de acesso, situada na altura da caixa da Figura 5b. Se esse resultado ocorre quando a proteína está em estado biológico, pontos de ligação externos à proteína podem ser pontos de captura de substratos. A partir desses pontos eles seriam conduzidos até algum ponto de acesso. Um substrato capturado está impossibilitado de adentrar na célula. Essa característica aumentaria a eficiência da proteína AcrB, pois ela poderia estar ativamente transportando um substrato em seu interior para fora da célula, mas já possuir outros capturados em sua superfície que está em contato com o periplasma e também em outros pontos de captura. Além disso, como aparentemente existem as estruturas que delimitam diferentes regiões, torna-se razoável supor que uma vez que o substrato deixe uma região e passe para outra, esta região estaria acessível para outro substrato. Assim, estas portas de comunicação entre regiões são também possíveis alvos para inibidores. Outro resultado interessante das simulações do artigo 2 foi a trajetória diferenciada dos substratos na última parte do caminho 1. Isso indica que o procedimento TMD, embora defina as posições inicial e final, não predetermina os substratos a seguirem a mesma trajetória.

Assim, resta baseado nos resultados obtidos, propor um mecanismo de inibição do efluxo.

Em primeiro lugar, propõe-se que seja construída a curva de energia para o composto a ter seu efluxo inibido, o substrato, e as respectivas tabelas que informam sobre a competição. Depois, faz-se o mesmo para os candidatos a inibidores. De acordo com as formas das curvas, os valores de energia obtidos ao longo da AcrB e a quantidade de resíduos comuns obtidos das tabelas, já se pode verificar se existe compatibilidade entre inibidor e bomba e competição entre inibidor e substrato. Essa primeira etapa pode ser feita somente com estudos de docagem seguindo a metodologia descrita no artigo 1. O emprego de TMD poderá vir logo a seguir, pois mostra a dinâmica do processo além de identificar resíduos e regiões que não são localizados por simples docagem.

8 CONCLUSÕES

As técnicas de bioinformática constituíram uma ferramenta útil para estudos da bomba de efluxo multidroga AcrB.

A técnica de docagem permitiu a construção das curvas de energia, que se mostrou uma ferramenta importante para se descobrir locais da proteína nos quais os substratos se ligam a ela com valores de grande magnitude de energia livre de ligação. As curvas de energia também sugerem, pela simples comparação de sua forma, se um substrato tem preferência pelos mesmos sítios de ligação que o outro. As curvas de energia possibilitaram ainda a construção dos caminhos a serem percorridos pelos substratos no segundo artigo.

A construção das tabelas do primeiro artigo possibilitou verificar a quantidade de competições por mesmos sítios entre substratos.

A dinâmica molecular dirigida do segundo artigo revelou uma série de “loops”, hélices α e folhas β que participam do transporte dos substratos e que podem ser, portanto, alvo de inibidores de efluxo. Além disso, essas simulações também mostraram que, dentre os caminhos escolhidos, os substratos percorreram as mesmas trajetórias.

Concluindo, a análise conjunta das curvas de energia, das tabelas, apresentadas no primeiro artigo e dos resultados da dinâmica molecular dirigida no segundo artigo, indicam que competição por mesmos sítios de ligação, ou caminho, podem ser um mecanismo de inibição de efluxo.

9 REFERÊNCIAS

BAILEY, A. M.; PAULSEN, I. T.; PIDDOCK, L. J. RamA confers multidrug resistance in *Salmonella enterica* via increased expression of *acrB*, which is inhibited by chlorpromazine. *Antimicrob Agents Chemother*, v. 52, n. 10, p. 3604-11, Oct 2008. ISSN 1098-6596 (Electronic)

0066-4804 (Linking). Disponível em:
<http://www.ncbi.nlm.nih.gov/pubmed/18694955>.

CHAMBERLAND, S. I. Y. L., V. J.; LEGER, R.; NAKAYAMA, K.; OHTA, T.; OHTSUKA, M.; RENAULT, T. E.; WATKINS, W. J. Efflux Pump Inhibitors. United States Patent US 6,399,629 B 2002.

CLAVERIE, J.-M.; NOTREDAME, C. *Bioinformatics for dummies*. 2nd. Hoboken, N.J.: Wiley Pub., 2007. xviii, 436 p. ISBN 9780470089859 (pbk.) 0470089857 (pbk.). Disponível em: Contributor biographical information
<http://www.loc.gov/catdir/enhancements/fy0827/2006934844-b.html> Publisher description
<http://www.loc.gov/catdir/enhancements/fy0827/2006934844-d.html>

Table of contents only <http://www.loc.gov/catdir/enhancements/fy0827/2006934844-t.html>.

D.A. CASE, J. T. B., R.M. BETZ, D.S. CERUTTI, T.E. CHEATHAM, III, T.A. DARDEN, R.E. DUKE, T.J. GIESE, H. GOHLKE, A.W. GOETZ, N. HOMEYER, S. IZADI, P. JANOWSKI, J. KAUS, A. KOVALENKO, T.S. LEE, S. LEGRAND, P. LI, T. LUCHKO, R. LUO, B. MADEJ, K.M. MERZ, G. MONARD, P. NEEDHAM, H. NGUYEN, H.T. NGUYEN, I. OMELIAN, A. ONUFRIEV, D.R. ROE, A. ROITBERG, R. SALOMON-FERRER, C.L. SIMMERLING, W. SMITH, J. SWAILS, R.C. WALKER, J. WANG, R.M. WOLF, X. WU, D.M. YORK AND P.A. KOLLMAN. *Amber 14*. 2015. Disponível em:
<http://ambermd.org/>.

D.A. CASE, T. A. D., T.E. CHEATHAM, III, C.L. SIMMERLING, J. WANG, R.E. DUKE, R. et al. *Amber 12*. University of California, San Francisco 2012.

DA SILVA, P. E. et al. Efflux as a mechanism for drug resistance in *Mycobacterium tuberculosis*. *FEMS Immunol Med Microbiol*, v. 63, n. 1, p. 1-9, Oct 2011. ISSN 1574-695X (Electronic) 0928-8244 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21668514>.

EICHER, T.; BRANDSTATTER, L.; POS, K. M. Structural and functional aspects of the multidrug efflux pump AcrB. *Biol Chem*, v. 390, n. 8, p. 693-9, Aug 2009. ISSN 1437-4315 (Electronic) 1431-6730 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19453279>.

EICHER, T. et al. Transport of drugs by the multidrug transporter AcrB involves an access and a deep binding pocket that are separated by a switch-loop. *Proc Natl Acad Sci U S A*, v. 109, n. 15, p. 5687-92, Apr 10 2012. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/22451937>.

FERNANDEZ, L.; HANCOCK, R. E. Adaptive and mutational resistance: role of porins and efflux pumps in drug resistance. *Clin Microbiol Rev*, v. 25, n. 4, p. 661-81, Oct 2012. ISSN 1098-6618 (Electronic) 0893-8512 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23034325> >.

HIRSCH, E. B.; TAM, V. H. Impact of multidrug-resistant *Pseudomonas aeruginosa* infection on patient outcomes. *Expert Rev Pharmacoecon Outcomes Res*, v. 10, n. 4, p. 441-51, Aug 2010. ISSN 1744-8379 (Electronic) 1473-7167 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/20715920>.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: visual molecular dynamics. *J Mol Graph*, v. 14, n. 1, p. 33-8, 27-8, Feb 1996. ISSN 0263-7855 (Print) 0263-7855 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/8744570>.

HyperChem (TM) Professional 7.51, HyperCube Inc.

IRWIN, J. J. et al. ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model*, v. 52, n. 7, p. 1757-68, Jul 23 2012. ISSN 1549-960X (Electronic) 1549-9596 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/22587354>.

LASKOWSKI, R. A.; SWINDELLS, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model*, v. 51, n. 10, p. 2778-86, Oct 24 2011. ISSN 1549-960X (Electronic) 1549-9596 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21919503>.

LESK, A. M. Introduction to bioinformatics. 2nd. Oxford ; New York: Oxford University Press, 2005. xviii, 360 p., 10 p. of plates ISBN 0199277877 (pbk.). Disponível em: Table of contents only <http://www.loc.gov/catdir/toc/fy055/2005279076.html> Publisher description <http://www.loc.gov/catdir/enhancements/fy0638/2005279076-d.html>.

LI, X. Z.; NIKAIDO, H. Efflux-mediated drug resistance in bacteria. *Drugs*, v. 64, n. 2, p. 159-204, 2004. ISSN 0012-6667 (Print) 0012-6667 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/14717618>.

Efflux-mediated drug resistance in bacteria: an update. *Drugs*, v. 69, n. 12, p. 1555-623, Aug 20 2009. ISSN 0012-6667 (Print) 0012-6667 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19678712>.

LI, X. Z.; PLESIAT, P.; NIKAIDO, H. The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clin Microbiol Rev*, v. 28, n. 2, p. 337-418, Apr 2015. ISSN 1098-6618 (Electronic) 0893-8512 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/25788514>.

MARTINEZ, J. L.; BAQUERO, F. Mutation frequencies and antibiotic resistance. *Antimicrob Agents Chemother*, v. 44, n. 7, p. 1771-7, Jul 2000. ISSN 0066-4804 (Print) 0066-4804 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/10858329>.

MARTINEZ, J. L. et al. Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. *FEMS Microbiol Rev*, v. 33, n. 2, p. 430-49, Mar 2009. ISSN 1574-6976 (Electronic) 0168-6445 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19207745>.

MORRIS, G. M. et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*, v. 30, n. 16, p. 2785-91, Dec 2009. ISSN 1096-987X (Electronic) 0192-8651 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19399780>.

MURAKAMI, S. et al. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature*, v. 419, n. 6907, p. 587-93, Oct 10 2002. ISSN 0028-0836 (Print) 0028-0836 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/12374972>.

NIKAIDO, H. Multidrug efflux pumps of gram-negative bacteria. *J Bacteriol*, v. 178, n. 20, p. 5853-9, Oct 1996. ISSN 0021-9193 (Print) 0021-9193 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/8830678>.

NIKAIDO, H.; TAKATSUKA, Y. Mechanisms of RND multidrug efflux pumps. *Biochim Biophys Acta*, v. 1794, n. 5, p. 769-81, May 2009. ISSN 0006-3002 (Print) 0006-3002 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19026770>.

OKUSU, H.; MA, D.; NIKAIIDO, H. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherichia coli* multiple-antibiotic-resistance (Mar) mutants. *J Bacteriol*, v. 178, n. 1, p. 306-8, Jan 1996. ISSN 0021-9193 (Print) 0021-9193 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/8550435>.

PETTERSEN, E. F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, v. 25, n. 13, p. 1605-12, Oct 2004. ISSN 0192-8651 (Print) 0192-8651 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/15264254>.

POS, K. M. Drug transport mechanism of the AcrB efflux pump. *Biochim Biophys Acta*, v. 1794, n. 5, p. 782-93, May 2009. ISSN 0006-3002 (Print) 0006-3002 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19166984>.

Protein Data Bank. Disponível em: <http://www.rcsb.org/pdb/home/home.do>. Acesso em: 26-03-2016.

SANNER, M. F. Python: a programming language for software integration and development. *J Mol Graph Model*, v. 17, n. 1, p. 57-61, Feb 1999. ISSN 1093-3263 (Print) 1093-3263 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/10660911>.

SCHLITTER, J.; ENGELS, M.; KRUGER, P. Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. *J Mol Graph*, v. 12, n. 2, p. 84-9, Jun 1994. ISSN 0263-7855 (Print) 0263-7855 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/7918256>.

SCHRODINGER, LLC. The PyMOL Molecular Graphics System, Version 1.8 2015.

SEEGER, M. A. et al. The AcrB efflux pump: conformational cycling and peristalsis lead to multidrug resistance. *Curr Drug Targets*, v. 9, n. 9, p. 729-49, Sep 2008. ISSN 1873-5592 (Electronic) 1389-4501 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/18781920>.

SEEGER, M. A. et al. Structural asymmetry of AcrB trimer suggests a peristaltic pump mechanism. *Science*, v. 313, n. 5791, p. 1295-8, Sep 1 2006. ISSN 1095-9203 (Electronic) 0036-8075 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/16946072>.

SEUS, V. R. S., L.; GOMES, J.; SILVA, P. E. A. DA; WERHLI, A. V.; MACHADO, K. S.; PRATES, N. S. Z., N. A Framework for Virtual Screening. SYMPOSIUM ON APPLIED COMPUTING - New York, NY, USA 2016.

TROTT, O.; OLSON, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*, v. 31, n. 2, p. 455-61, Jan 30 2010. ISSN 1096-987X (Electronic) 0192-8651 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/19499576>.

TSENG, T. T. et al. The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. *J Mol Microbiol Biotechnol*, v. 1, n. 1, p. 107-25, Aug 1999. ISSN 1464-1801 (Print) 1464-1801 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/10941792>.

VAN DER SPOEL, D. et al. GROMACS: fast, flexible, and free. *J Comput Chem*, v. 26, n. 16, p. 1701-18, Dec 2005. ISSN 0192-8651 (Print) 0192-8651 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/16211538>.

VARGIU, A. V.; NIKAIDO, H. Multidrug binding properties of the AcrB efflux pump characterized by molecular dynamics simulations. *Proc Natl Acad Sci U S A*, v. 109, n. 50, p. 20637-42, Dec 11 2012. ISSN 1091-6490 (Electronic) 0027-8424 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/23175790>.

WEBB, B.; SALI, A. Protein structure modeling with MODELLER. *Methods Mol Biol*, v. 1137, p. 1-15, 2014. ISSN 1940-6029 (Electronic) 1064-3745 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/24573470>.

WEBBER, M. A.; PIDDOCK, L. J. The importance of efflux pumps in bacterial antibiotic resistance. *J Antimicrob Chemother*, v. 51, n. 1, p. 9-11, Jan 2003. ISSN 0305-7453 (Print) 0305-7453 (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/12493781>.

XIA, C. Q.; SMITH, P. G. Drug efflux transporters and multidrug resistance in acute leukemia: therapeutic impact and novel approaches to mediation. *Mol Pharmacol*, v. 82, n. 6, p. 1008-21, Dec 2012. ISSN 1521-0111 (Electronic) 0026-895X (Linking). Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/22826468>.

9 Anexo 1

Os professores doutores Adriano Werhli e Karina Machado, do Programa de Pós-Graduação em Ciências Computacionais da FURG, gentilmente elaboraram uma descrição mais genérica para o procedimento de deslocamento da caixa de docagem. O procedimento original, descrito no artigo 1 desta tese, foi idealizado pelo autor deste trabalho durante a escrita do referido artigo. Segue abaixo a sugestão dos professores.

Posicionamento Variável da Caixa de docagem

Geralmente em simulações de docagem é necessária a definição de uma caixa de busca. Essa caixa define o espaço tridimensional no qual o algoritmo do programa de docagem utilizado procura pela melhor posição do ligante em relação à proteína receptora, de acordo com um escore fornecido pelo programa que geralmente está relacionado à energia livre de ligação. Uma característica muito útil do “framework” proposto é a possibilidade de se planejar com antecedência caixas específicas que serão analisadas pelo algoritmo de docagem. A simulação de docagem é realizada em todas as caixas que forem definidas. Dessa forma, ao final das simulações, tem-se o melhor resultado da docagem para cada caixa.

No programa Autodock Tools, utilizado neste trabalho, a caixa de docagem é definida pelo comprimento de cada uma de suas dimensões, chamadas x_b , y_b , z_b e pela posição de seu centro (x_c, y_c, z_c) .

No “framework” proposto, é possível variar a posição do centro da caixa em qualquer combinação de dimensões, ou seja, o usuário pode decidir se a variação será em uma, duas ou três dimensões. Para cada dimensão a qual o usuário desejar uma variação, ele deverá informar a coordenada final e um passo o qual indica o deslocamento da caixa na dimensão desejada. Generalizando, para a variação da posição da caixa, o usuário tem que fornecer a posição final do centro da caixa, $(x_{end}, y_{end}, z_{end})$, e o passo $(x_{step}, y_{step}, z_{step})$ que será usado em cada coordenada para alcançar a posição final. Definidos esses valores, um conjunto de pontos para cada coordenada é criado,

$$X = \{x_c, (x_c + x_{step}), (x_c + 2 \times x_{step}), \dots, (x_c + n_x \times x_{step})\}$$

$$Y = \{y_c, (y_c + y_{step}), (y_c + 2 \times y_{step}), \dots, (y_c + n_y \times y_{step})\}$$

$$Z = \{z_c, (z_c + z_{step}), (z_c + 2 \times z_{step}), \dots, (z_c + n_z \times z_{step})\}$$

where $x_{end} = x_c + n_x \times x_{step}$, $y_{end} = y_c + n_y \times y_{step}$ e $z_{end} = z_c + n_z \times z_{step}$.

Note que o número total de simulações de docagem que será realizado é dado por

$$N = n_x \times n_y \times n_z.$$

Se o usuário desejar variar o posicionamento da caixa em apenas uma dimensão, como foi feito no artigo 1 desta tese, ele deverá deixar os campos das coordenadas finais e passos relativos às dimensões restantes em branco.

10 Anexo 2

Carta eletrônica do editor da revista International Journal of Antimicrobial Agents (IJAA) afirmando que reconsiderará o artigo 1 desta tese se as correções solicitadas forem feitas. As correções estão sendo feitas.

From: **IJAA (ELS)** <IJAA@elsevier.com>
Date: 2016-03-30 8:04 GMT-03:00
Subject: RE: Your Submission IJAA-D-16-00239
To: Pedro Silva <pedrefurg@gmail.com>

Dear Pedro Silva,

With reference to your enquiry.

The Editor has advised that in this instance he would be happy to reconsider your paper if you are able to respond to the reviewers requests.

If you wish to go ahead, please let us know and we will initiate a rebuttal, which will allow you to revise your paper and resubmit.

Kind regards

Jacqui Merrison

From: Pedro Silva [mailto:pedrefurg@gmail.com]
Sent: 24 March 2016 16:06
To: IJAA (ELS)
Subject: Re: Your Submission IJAA-D-16-00239

Dear Editor

Thanks for your attention for our manuscript. Regard to the reviewer's comments we would like highlight that:

1. In fact, the structures are in Table A1, as indicate in line 89 and showing all compounds.
2. Both references are in indexed journals. 17 (European Journal of Organic Chemistry indexed in pubmed with Impact Factor: 3.065) and 18 (Synthesis indexed Current Contents/Life Science, Current Contents/Physical, Chemical and Earth Science, Chemical Abstracts, Science Citation Index, Research Alert, Scisearch, Index Chemicus, Chemistry Citation Index, Current Chemical Reactions, Reaction Citation Index, Reference Update, Biological Abstracts, EMBASE, CABS/BIOBASE, GEO

Abstracts/GEODATABASE). Both have DOI DOI: 10.1055/s-0029-1218779 and 10.1002/ejoc.200801119, **17 and 18** respectively.

3. In fact, cytotoxicity was done in J774A.1 (ATCC TIB-67), an eukaryotic cell and showed no cytotoxic.

Objectively, regarding the request of Reviewer:

(i) the necessary information for a correct assessment of the potential interest of these compounds as drugs (including structures, water solubility and other pertinent biophysical and biochemical properties).

We could show the structure of all compounds and describe the main physical and biochemical properties of these compounds, however the chemical formula is there (Appendix A)

(ii) the results of cytotoxicity assays with at least two lines of eukaryotic cells.

We could do cytotoxic assay with another line of eukaryotic cells, besides J774A.1

Best regards,

Pedro Eduardo Almeida Silva