**UNIVERSIDADE FEDERAL DE GOIÁS**

**PROGRAMA DE PÓS-GRADUAÇÃO EM MEDICINA TROPICAL E SAÚDE PÚBLICA**

**BRUNO JUNIOR NEVES**

# Reposicionamento de fármacos e planejamento de novos compostos ativos contra *Schistosoma mansoni*

**Goiânia**

**2016**

**TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS NA BIBLIOTECA DIGITAL DA UFG**

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

**1. Identificação do material bibliográfico**: **[ ] Dissertação** **[X] Tese**

**2. Identificação da Tese ou Dissertação**

Nome completo do autor: Bruno Junior Neves

Título do trabalho: Reposicionamento de fármacos e planejamento de novos compostos ativos contra *Schistosoma mansoni*

**3. Informações de acesso ao documento:**

Concorda com a liberação total do documento [ X ] SIM [ ] NÃO[1]

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.

Bruno Junior Neves

_____

Assinatura do autor[2]

Data: 17 / 12 / 2016

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

[2] A assinatura deve ser escaneada.

**BRUNO JUNIOR NEVES**

# Reposicionamento de fármacos e planejamento de novos compostos ativos contra *Schistosoma mansoni*

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Medicina Tropical e Saúde Pública da Universidade Federal de Goiás para obtenção do título de Doutor em Medicina Tropical e Saúde Pública.

Orientador: Dra. Carolina Horta Andrade

**Goiânia**

**2016**

Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Neves, Bruno Junior
Reposicionamento de fármacos e planejamento de novos compostos ativos contra Schistosoma mansoni [manuscrito] / Bruno Junior Neves. - 2016.
CLXII, 162 f.: il.

Orientador: Profa. Dra. Carolina Horta Andrade.
Tese (Doutorado) - Universidade Federal de Goiás, Instituto de Patologia Tropical e Saúde Pública (IPTSP), Programa de Pós Graduação em Medicina Tropical e Saúde Pública, Goiânia, 2016.
Bibliografia.
Inclui siglas, mapas, abreviaturas, símbolos, gráfico, lista de figuras.

1. esquistossomose. 2. QSAR. 3. homologia. 4. triagem virtual. 5. ensaios ex vivo. I. Andrade, Carolina Horta, orient. II. Título.

CDU 576.8

UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE PATOLOGIA TROPICAL E SAÚDE PÚBLICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MEDICINA TROPICAL E SAÚDE PÚBLICA
Rua 235, s/n – Setor Universitário - Goiânia/GO – CEP: 74.605-050
Fones: (62) 3209.6362 - 3209.6102 – Fax: (62) 3209.6363 - e-mail : ppgmtsp.ufg@gmail.com

**ATA DA REUNIÃO DA BANCA EXAMINADORA DA DEFESA DE BRUNO JÚNIOR NEVES** - Ao primeiro dia do mês de novembro do ano de 2016 (01/11/2016), às 13:30 horas, reuniram-se os componentes da Banca Examinadora: Profs. Drs. CAROLINA HORTA ANDRADE, FLORIANO PAES SILVA JÚNIOR, ADRIANO DEFINI ANDRICOPULO, RODOLPHO DE CAMPOS BRAGA e JOSÉ CLECILDO BARRETO BEZERRA, para, sob a presidência do primeiro, e em sessão pública realizada no INSTITUTO DE PATOLOGIA TROPICAL E SAÚDE PÚBLICA, procederem à avaliação da defesa de tese intitulada: **"REPOSICIONAMENTO DE FÁRMACOS E PLANEJAMENTO DE NOVOS COMPOSTOS ATIVOS CONTRA *Schistosoma mansoni*"** em nível de **DOUTORADO**, área de concentração em **PARASITOLOGIA**, de autoria de **BRUNO JÚNIOR NEVES**, discente do PROGRAMA DE PÓS-GRADUAÇÃO EM MEDICINA TROPICAL E SAÚDE PÚBLICA, da Universidade Federal de Goiás. A sessão foi aberta pela Orientadora Profa. Dra. CAROLINA HORTA ANDRADE, que fez a apresentação formal dos membros da Banca e orientou o Candidato sobre como utilizar o tempo durante a apresentação de seu trabalho. A palavra a seguir, foi concedida ao autor da tese que, em 30 minutos procedeu à apresentação de seu trabalho. Terminada a apresentação, cada membro da Banca argüiu o Candidato, tendo-se adotado o sistema de diálogo seqüencial. Terminada a fase de argüição, procedeu-se à avaliação da defesa. Tendo-se em vista o que consta na Resolução nº. 1034/2014 do Conselho de Ensino, Pesquisa, Extensão e Cultura (CEPEC), que regulamenta o Programa de Pós-Graduação em Medicina Tropical e Saúde Pública a Banca, em sessão secreta, expressou seu Julgamento, considerando o candidato **Aprovado** ou **Reprovado**:

| **Banca Examinadora** | **Aprovado / Reprovado** |
|---|---|
| Dra. Carolina Horta Andrade | Aprovado |
| Dr. Floriano Paes Silva Júnior | APROVADO |
| Dr. Adriano Defini Andricopulo | APROVADO |
| Dr. Rodolpho de Campos Braga | Aprovado |
| Dr. José Clecildo Barreto Bezerra | APROVADO |

Em face do resultado obtido, a Banca Examinadora considerou o candidato Habilitado, (**Habilitado ou não Habilitado**), cumprindo todos os requisitos para fins de obtenção do título de **DOUTOR EM MEDICINA TROPICAL E SAÚDE PÚBLICA**, na área de concentração em **PARASITOLOGIA**, pela Universidade Federal de Goiás. Cumpridas as formalidades de pauta, às 17 h 30 min, a presidência da mesa encerrou esta sessão de defesa de tese e para constar eu, JOSÉ CLEMENTINO DE OLIVEIRA NETO, secretário do Programa de Pós-Graduação em Medicina Tropical e Saúde Pública lavrei a presente Ata que depois de lida e aprovada, será assinada pelos membros da Banca Examinadora e por mim em duas vias de igual teor.
A Banca Examinadora aprovou a seguinte alteração no título da Tese:

_____

_____

_____

Dra. Carolina Horta Andrade (FF/UFG) _____

Dr. Floriano Paes Silva Júnior (IOC-FIOCRUZ/RJ) _____

Dr. Adriano Defini Andricopulo (IFSC/USP) _____

Dr. Rodolpho de Campos Braga (FF/UFG) _____

Dr. José Clecildo Barreto Bezerra (IPTSP/UFG) _____

Secretário da Pós-Graduação: _____

Profª Drª Regina Maria Bringel Martins
Subcoordenadora do Programa de Pós-Graduação
em Medicina Tropical e Saúde Pública
IPTSP/UFG

CONFERE COM O ORIGINAL
José Clementino de Oliveira Neto
Assistente em Administração
Pós-Graduação/IPTSP/UFG
SECRETÁRIO

i

***Dedico este trabalho, ao meu filho Raphael Junior Soares Neves, para que um dia ele possa reconhecer que a maior herança que um pai pode dar ao filho é a educação.***

# AGRADECIMENTOS

*Primeiramente agradeço a Deus, pela oportunidade de realizar mais este sonho.*

*À minha esposa Karla Karolina Soares Rocha Neves, pelo apoio, companheirismo e incentivo incondicional.*

*Aos meus pais, Sônia Maria da Silva Neves e Junior Moreira Neves, pelo carinho e por me incentivarem, desde a infância, a buscar o conhecimento e a lutar pelos meus sonhos.*

*À minha orientadora, Dra. Carolina Horta Andrade, pelo exemplo de profissionalismo, competência e pela orientação.*

*Aos Drs. Floriano Paes Silva Junior, Mario Roberto Senger, Rafael Ferreira Dantas, e demais colegas do LaBECFar, pelas enormes contribuições e aprendizado com os ensaios biológicos em S. mansoni.*

*Aos Drs. Ross Paveley e Nicholas Furnham pela colaboração com os ensaios biológicos em esquistossômulos de Schistosoma mansoni.*

*Aos Dr. Lee Kamentsky e Anne Elizabeth Carpenter, pela colaboração e automatização dos ensaios biológicos em vermes adultos de Schistosoma mansoni.*

*Ao Dr. Pedro Vitor Lemos Cravo, pela colaboração, ensinamento e grandes oportunidades oferecidas.*

*Ao Dr. Rodolpho de Campos Braga, pela disponibilização dos fluxos de trabalho (KSAR) utilizados para construção e validação de modelos de QSAR.*

*Ao Dr. Eugene Muratov, pelas importantes contribuições e colocações no desenvolvimento de modelos de QSAR e triagem virtual.*

*Ao Dr. José Clecildo Barreto Bezerra, pelos ensinamentos, acessibilidade e ideias motivacionais no campo da pesquisa científica e empreendedorismo.*

*Ao servidor técnico-administrativo José Clementido de Oliveira Neto, pela excelência de trabalho na secretaria do programa de pós-graduação.*

*Às farmacêuticas Kellen Farinelli e Luciana Resende Prudente, pela doação dos fármacos investigados nos ensaios biológicos.*

*Ao amigo Marcelo do Nascimento Gomes, pela companheirismo e orientações durante o curso de doutorado.*

*Aos amigos Flávia Cristina Silva, Cleber Camilo de Melo Filho, Marília Nunes do Nascimento e Arthur de Carvalho e Silva, pela amizade construída durante o curso de doutorado.*

*Aos demais colegas do LabMol, pelos momentos de descontração e trocas de conhecimento.*

*Aos amigos Hudson Henrique Silva, Leonardo Pereira Dias e Paulo Henrique Dias, pela amizade e incentivo.*

*Às empresas OpenEye e ChemAxon, pela disponibilização de licenças acadêmicas dos programas utilizados neste trabalho.*

*Às agências CAPES, CNPq e FAPEG, pelo apoio financeiro.*

*Ao Portal de Periódicos da CAPES.*

# SUMÁRIO

# TABELAS, FIGURAS E ANEXOS

# SÍMBOLOS, SIGLAS E ABREVIATURAS

| | |
|---|---|
| ·NO | Óxido nítrico |
| $\cdot O_2^-$ | Superóxido |
| ·OH | Hidroxil |
| µm | Micrômetro |
| µM | Micromolar |
| ºC | Grau Celsius |
| 1D | Unidimensional |
| 2D | Bidimensional |
| 3D | Tridimensional |
| CADD | Planejamento de fármacos auxiliado por computador |
| Cavβ | Subunidade β do canal de cálcio voltagem dependente |
| CCR | Taxa de classificação correta |
| cm | Centímetro |
| Cys | Cisteína |
| DA | Domínio de aplicabilidade |
| DNA | Ácido desoxirribonucleico |
| DNDi | *Drugs for Neglected Diseases initiative* |
| DTN | Doença tropical negligenciada |
| $EC_{50}$ | Concentração necessária para inibir 50% da motilidade máxima |
| ERNs | Espécies reativas de nitrogênio |
| EROs | Espécies reativas de oxigênio |
| FAD | Flavina adenina dinucleotídeo |
| FDA | *Food and Drug Administration* |
| $Fe^{2+}$ | Ferro no estado ferroso |
| Fer | Ferritina |
| FQ | Fitoquelatina |
| FQS | Fitoquelatina sintase |
| GPxs | Glutationa peroxidases |

| | |
|---|---|
| GR | Glutationa redutase |
| Grx | Glutarredoxina |
| GSH | Glutationa |
| GSSG | Glutationa dissulfeto |
| GST | Glutationa *S*-transferase |
| $H_2O_2$ | Peróxido de hidrogênio |
| hSERT | Proteína humana transportadora de serotonina |
| HTS | Triagem de alta vazão |
| $IC_{50}$ | Concentração necessária para inibir 50% da atividade |
| LBDD | Planejamento baseado na estrutura do ligante |
| LogP | Logaritmo decimal do coeficiente de partição |
| Met | Metionina |
| Met-O | Metionina sulfóxido |
| MSR | Metionina sulfóxido redutase |
| NADPH | Nicotinamida adenina dinucleótido fosfato |
| NCE | Nova entidade química |
| NOS II | Óxido nítrico redutase indutiva |
| OECD | Organização para a Cooperação e Desenvolvimento Econômico |
| OLT | Oltipraz |
| OMS | Organização Mundial de Saúde |
| $ONOO^-$ | Peroxinitrito |
| PAR | Paroxetina |
| PCA | Análise de componentes principais |
| PCS | Fitoquelatina sintase |
| PDB | *Protein Data Bank* |
| pKa | Logaritmo da constante de acidez no estado fundamental |
| Prxs | Peroxirredoxinas |
| PZQ | Praziquantel |
| QSAR | Relação quantitativa entre a estrutura e atividade |
| RF | *Random forest* |

| | |
|---|---|
| RNA | Ácido ribonucleico |
| RNAi | RNA de interferência |
| RNA-seq | Sequenciamento de RNA |
| Sec | Selenocisteína |
| *S. guineensis* | *Schistosoma guineensis* |
| *S. haematobium* | *Schistosoma haematobium* |
| *S. intercalatum* | *Schistosoma intercalatum* |
| *S. japonicum* | *Schistosoma japonicum* |
| *S. mansoni* | *Schistosoma mansoni* |
| *S. mekongi* | *Schistosoma mekongi* |
| SAR | Relação estrutura e atividade |
| SBDD | Planejamento baseado na estrutura |
| *Sj*TGR | Tiorredoxina glutationa redutase de *Shistosoma japonicum* |
| *Sm*TGR | Tiorredoxina glutationa redutase de *Shistosoma mansoni* |
| *Sm*SERT | Proteína transportadora de serotonina de *S. mansoni* |
| SNC | Sistema Nervoso Central |
| SOD | Superóxido dismutase |
| SVM | *Support vector Machine* |
| Trx | Tiorredoxina |
| TrxR | Tiorredoxina Redutase |
| Trx-$S_2$ | Tioredoxina oxidada |
| TTD | *Therapeutic Target Database* |
| VS | Triagem virtual |

# RESUMO

A esquistossomose é uma doença endêmica grave causada por trematódeos do gênero *Schistosoma*. Atualmente o controle desta doença é baseado exclusivamente no uso do fármaco praziquantel. Todavia, seu uso extensivo tem levantado a preocupação quanto ao surgimento de vermes resistentes e a necessidade de descobrir novos fármacos esquistossomicidas. Face ao exposto, o objetivo deste trabalho foi planejar novos inibidores da enzima tiorredoxina glutationa redutase de *S. mansoni* (*Sm*TGR) com atividade esquistossomicida utilizando ferramentas de quimioinformática e reposicionar novos fármacos para esquistossomose utilizando ferramentas de bioinformática. No estudo de planejamento de fármacos, modelos de QSAR binários foram construídos e validados para predição da atividade inibitória da *Sm*TGR, um alvo validado em esquistossomos. A partir dos modelos individuais, modelos de consenso e consenso rigoroso foram construídos (CCRs = 0,87 e 0,91, respectivamente) e utilizados na triagem virtual de 150 mil compostos. Ao final deste processo, 29 compostos foram priorizados e adquiridos para avaliação biológica. Como resultado, dois novos *hits* representando novos *scaffolds* moleculares apresentaram $EC_{50} \leq 3.5$ μM para esquistossômulos e $\leq 6.0$ μM para fêmeas de vermes adultos, baixa citotoxicidade em células WSS-1 de mamíferos ($IC_{50} > 16$ μM) e baixa reatividade com cisteíno proteases ($IC_{50} > 100$ μM). No estudo de reposicionamento de fármacos, 2.114 proteínas de *S. mansoni* foram utilizadas em uma busca por ortólogos em bases de dados de alvos terapêuticos de fármacos (DrugBank, TTD e STITCH). Como resultado, 215 fármacos foram preditos para interagir com 49 proteínas do esquistossomo, dos quais 47 já tinham atividade esquistossomicida reportada na literatura. Em seguida, a análise de componentes principais (PCA) e *k-means* demonstrou que 115 fármacos estavam dentro do espaço químico de agentes esquistossomicidas conhecidos, atribuindo maior confiabilidade as predições. Dentre eles, a paroxetina (PAR), um fármaco antidepressivo predito para inibir proteínas transportadoras de serotonina do *S. mansoni* (*Sm*SERTs), apresentou atividade esquistossomicida em esquistossômulos ($EC_{50} = 2,5$ μM) e vermes adultos ($EC_{50} = 5,1$ μM e 9,9 μM para machos e fêmeas) de *S. mansoni*. No final, estudos de docagem molecular com a *Sm*SERT-A e seu ortólogo em humanos (hSERT) exploraram as bases moleculares para a atividade esquistossomicida da PAR e forneceram informações para o planejamento de novos análogos mais potentes e seletivos.

# ABSTRACT

Schistosomiasis is a serious endemic disease caused by trematodes of the genus *Schistosoma*. Currently the control of this disease is based solely on the administration of praziquantel. However, its extensive use has raised concerns about the emergence of resistant worms and the need of discovering new anti-schistosomal drugs. Given the above, the main goal of this study was to design new *S. mansoni* thioredoxin glutathione reductase (*Sm*TGR) inhibitors showing anti-schistosomal activity through cheminformatics tools and to repurpose new drugs for schistosomiasis employing bioinformatics tools. At the stage of drug design, binary QSAR models were constructed and validated to predict inhibitory activity of *Sm*TGR, a validated target in schistosomes. From the individual models, consensus and consensus rigor models and were generated (CCRs = 0.87 and 0.91, respectively) and used for the virtual screening of 150,000 compounds. At the end of this process, 29 compounds were prioritized and purchased for biological evaluation. As a result, two new hits representing new molecular scaffolds showed $EC_{50} \leq 3.5$ μM to schistosomula and ≤ 6.0 μM for adult female worms, low cytotoxicity against WSS-1 mammalian cells ($IC_{50} > 16$ μM) and low reactivity with cysteine proteases ($IC_{50} > 100$ μM). In the second part of this work, that is drug repositioning, 2,114 proteins of *S. mansoni* were used in a search for orthologs in therapeutic drug target databases (DrugBank, TTD and STITCH). As a result, 215 drugs were predicted to interact with 49 schistosome proteins, of which 47 had already anti-schistosomal activity reported in the literature. Then, principal component analysis (PCA) and k-means showed that 115 drugs were in the chemical space of known anti-schistosomal agents, increasing the overall confidence of predictions. Among them, paroxetine (PAR), an antidepressant drug predicted to inhibit serotonin transporter proteins of *S. mansoni* (*Sm*SERTs) presented anti-schistosomal activity against schistosomula ($EC_{50} = 2.5$ μM) and adult worms ($EC_{50} = 5.1$ μM and 9.9 μM for males and females respectively) of *S. mansoni*. Lastly, molecular docking studies with *Sm*SERT-A and its ortholog in humans (hSERT) explored the molecular basis for anti-schistosomal activity of PAR and provided information for the design of more potent and selective analogs.

# 1. INTRODUÇÃO

## 1.1. Esquistossomose

A esquistossomose é uma doença tropical negligenciada (DTN) fortemente associada ao baixo índice de desenvolvimento socioeconômico, infraestrutura de saneamento básico precária e condições mínimas de higiene (WHO, 2016a). Estimativas da Organização Mundial de Saúde (OMS) indicam que a esquistossomose acomete cerca de 258 milhões de pessoas distribuídas em 78 países da África, Ásia e América Latina (Figura 1), dos quais 52 países apresentam transmissão da doença com intensidade o suficiente para justificar a quimioterapia profilática em grande escala (WHO, 2016b).

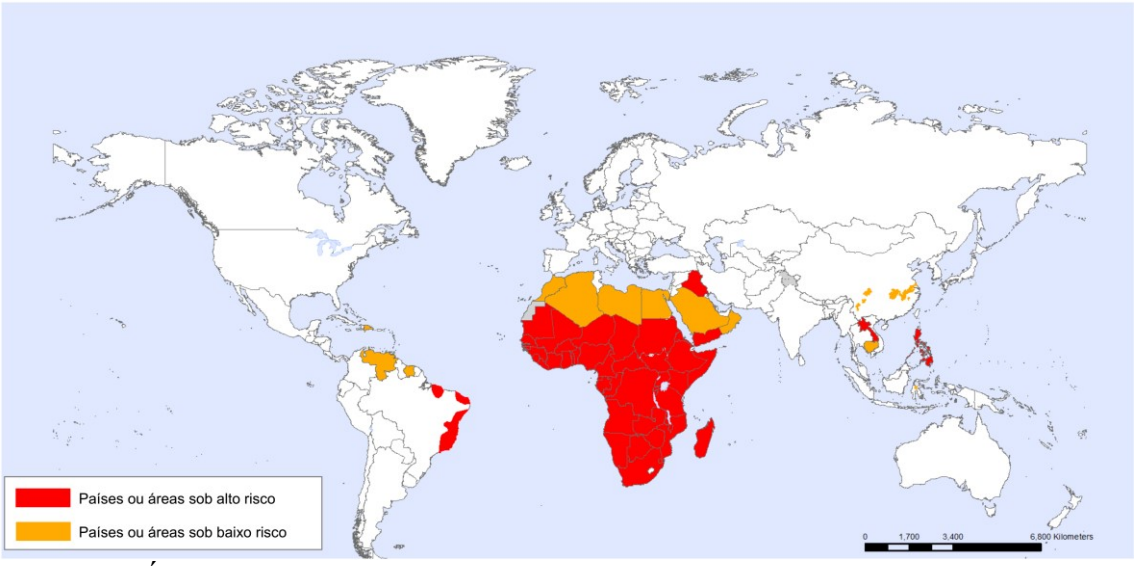



**Figura 1.** Áreas com risco de transmissão da esquistossomose (Adaptado de: WHO, 2016b).

No Brasil, cerca de 1,5 milhão de pessoas vivem em áreas sob o risco de contrair esquistossomose. A Figura 2 apresenta a distribuição da esquistossomose nos municípios brasileiros, de acordo com o percentual médio de positividade. Atualmente, a doença é detectada em 19 unidades federadas. As áreas endêmicas abrangem os estados de Alagoas, Bahia, Pernambuco, Rio Grande do Norte (faixa litorânea), Paraíba, Sergipe, Espírito Santo e Minas Gerais (predominantemente no Norte e Nordeste do estado). No Pará, Maranhão, Piauí, Ceará, Rio de Janeiro, São Paulo, Santa Catarina, Paraná, Rio Grande do Sul, Goiás e no Distrito Federal, a transmissão é focal e portanto, não atinge grandes áreas (MS, 2015).

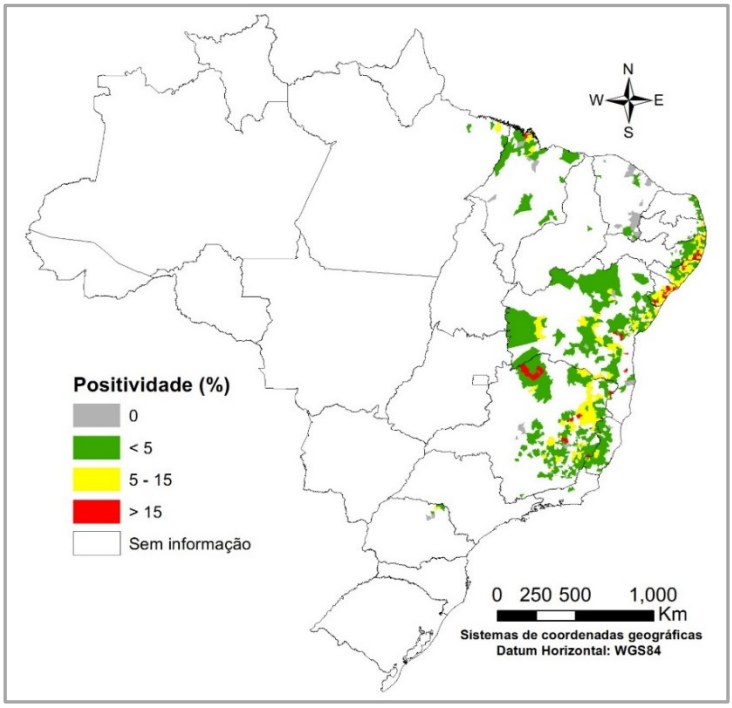



**Figura 2.** Percentual médio de positividade da esquistossomose mansônica nos municípios brasileiros (MS, 2015).

### 1.1.1. Etiologia e distribuição

A esquistossomose é causada por platelmintos trematódeos do gênero *Schistosoma.* Os vermes deste gênero apresentam um acentuado dimorfismo sexual, são heteroxênicos e parasitam vasos sanguíneos de mamíferos, os quais são seus hospedeiros definitivos. Atualmente, três espécies são responsáveis pela maioria das infecções humanas: *S. haematobium*, distribuído amplamente na África Subsaariana e partes do Oriente Médio; *S. japonicum*, com distribuição na Ásia, principalmente Filipinas e China; e *S. mansoni* com distribuição ampla na África, Oriente Médio e partes da América Latina. Outras três espécies distribuídas localmente também causam a doença no ser humano: *S. mekongi*, na bacia do rio Mekong (*i.e.* Camboja e Laos) e *S. intercalatum* e *S. guineensis* na África Ocidental e Central (ROSS et al., 2002; GRYSEELS et al., 2006; COLLEY et al., 2014).

A distribuição da esquistossomose deve-se em grande parte a fatores ambientais necessários para o desenvolvimento dos hospedeiros intermediários (Figura 3) em reservatórios de água estagnada (*i.e.* lagos, lagoas, represas e açudes), isto é, caramujos de água doce do gênero *Biomphalaria* (para *S. mansoni*), *Bulinus* (para *S. haematobium*) ou caramujos anfíbios do gênero *Oncomelania* (para *S. japonicum*), (PETERS; PASVOL, 2006; COLLEY et al., 2014).

**Figura 3.** Caramujos do gênero *Biomphalaria* (**A**), *Bulinus* (**B**) e *Oncomelania* (**C**) (Adaptado de: PETERS; PASVOL, 2006).

No Brasil, as principais zonas de distribuição da esquistossomose mansônica correspondem ao ambiente natural dos hospedeiros intermediários *B. glabrata* (nas regiões Nordeste, Sudeste e parte da região Sul), *B. straminea* (nas regiões Nordeste e Sudeste) e *B. tenagophila* (distribuição concentrada apenas no Sudeste e Sul do país) (SCHOLTE et al., 2012). Os focos endêmicos geralmente ocorrem na zona rural e reservatórios de água doce peridomiciliares. Consequentemente, a atividade humana nestas áreas consiste no principal fator de risco para transmissão e endemicidade da doença (VERCRUYSSE; SHAW; DE BONT, 2001; GRIMES et al., 2014, 2015).

### 1.1.2. Ciclo biológico

O esquistossomo possui um ciclo biológico complexo (Figura 4), alternando entre as formas assexuadas, presentes no hospedeiro intermediário (caramujo) e formas sexuadas, presentes no hospedeiro definitivo (mamíferos) (KING, 2009; GRYSEELS, 2012; FARRAR, 2014).

O ciclo biológico inicia após a deposição de fezes do hospedeiro definitivo contaminadas com ovos de esquistossomo nas imediações de reservatórios de água doce. (FARRAR, 2014; TOLEDO; FRIED, 2014). Sob estimulo da baixa osmolaridade do meio, e luminosidade e temperatura ambiente, o miracídio provoca a ruptura transversal da casca do ovo ao longo de uma linha conhecida como sutura (CARVALHO; COELHO; LENZI, 2008; TOLEDO; FRIED, 2014). O miracídio é uma larva com forma cilíndrocônica com a superfície composta por cílios, os quais permitem o movimento no ambiente aquático, uma papila apical onde se encontram as glândulas de adesão e reserva de glicogênio para manutenção do seu metabolismo energético (MAHMOUD, 2001; NEVES, 2005).

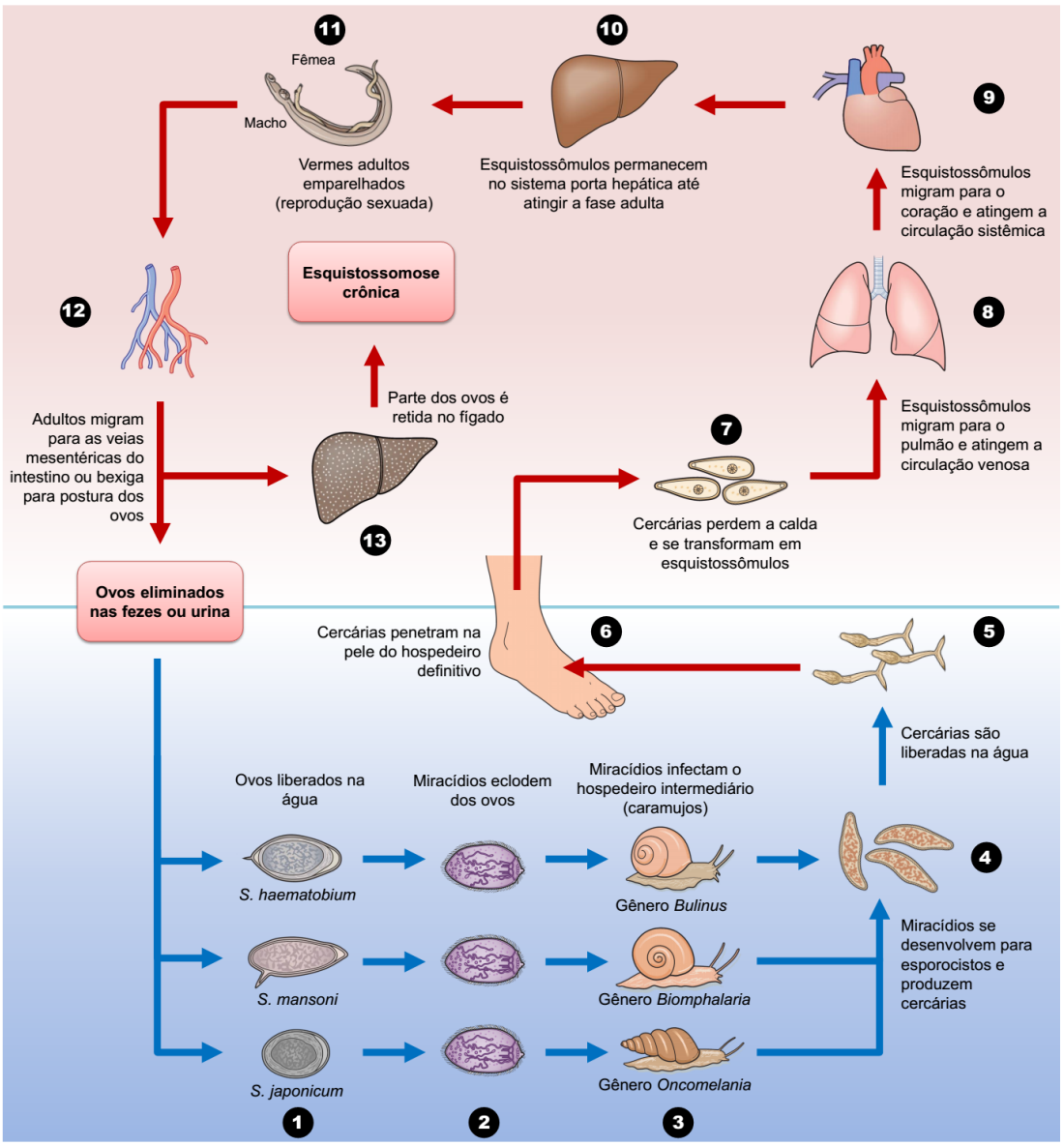



**Figura 4.** Ciclo biológico do esquistossomo (Adaptado de: ROSS et al., 2002). As setas vermelhas e azuis indicam os estádios de vida no hospedeiro vertebrado e invertebrado, respectivamente. Ovos eliminados nas fezes ou urina (**1**) liberam miracídios (**2**) que penetram no hospedeiro intermediário (**3**). Os miracídios se transformam em esporocistos I e II (**4**) que dão origem às cercárias (**5**). Essas por sua vez penetram na pele do hospedeiro definitivo (**6**) e se transformam em esquistossômulos (**7**). Os parasitos atingem a circulação, migram para o pulmão (**8**) e coração (**9**) e sistema porta-hepático (**10**) onde atingem maturidade sexual (**11**). Os vermes adultos acasalados (**12**) migram para veias mesentéricas ou plexo venoso da bexiga urinária (*S. haematobium*) para a postura dos ovos, os quais podem ser retidos nos tecidos (**13**).

Em seguida, o miracídio nada através de movimentos ciliares em direção ao hospedeiro intermediário, neste caso espécies de caramujos suscetíveis a infecção. O contato com as partes moles do caramujo faz com que sua papila apical assuma uma forma ventosa, ocorrendo, quase que simultaneamente, a descarga do conteúdo das glândulas de adesão. Após a adesão, o miracídio penetra nos tecidos moles do caramujo através da

ação combinada de movimentos contráteis e rotatórios e da ação de enzimas proteolíticas. (STIREWALT; HACKEY, 1956; MAHMOUD, 2001; BOGITSH; CARTER; OELTMANN, 2013; TOLEDO; FRIED, 2014).

Durante as 48 horas seguintes, o miracídio perde seus cílios, sistema nervoso e camadas musculares para se transformar em esporocisto primário, ou seja, um saco com paredes cuticulares contendo células germinativas ou reprodutivas. As células germinativas, em números de 50 a 100, iniciam então um intenso processo de multiplicação (poliembrionia). Na segunda semana da infecção, o esporocisto primário forma uma série de ramificações tubulares, que preenchem todos os espaços inter-tissulares do tecido conjuntivo do caramujo (MAHMOUD, 2001; BOGITSH; CARTER; OELTMANN, 2013; TOLEDO; FRIED, 2014).

Após um intervalo de 8 a 10 dias, ocorre a formação dos esporocistos secundários. O processo de formação inicia-se com um aglomerado de células germinativas nas paredes do esporocisto primário e vacuolização acentuada na parte central da larva. Esses aglomerados se reorganizam e dão origem a septos, ficando o esporocisto primário dividido em 150 a 200 camadas, cada septo ou camada já podendo ser considerado um esporocisto secundário. As paredes do esporocisto secundário são formadas por uma dupla camada muscular, constituída por fibras musculares longitudinais e transversais, logo abaixo de uma camada cuticular recoberta de espinhos, os quais possuem papel fundamental na motilidade e migração das larvas. Esta migração processa-se ativamente através dos tecidos do caramujo e termina nas glândulas digestivas (MAHMOUD, 2001; BOGITSH; CARTER; OELTMANN, 2013; TOLEDO; FRIED, 2014).

Encerrada a migração até as glândulas digestivas, ocorre a formação de cercárias. Este processo baseia-se na disposição das células germinativas em uma mórula, em cujo centro encontra-se uma grande célula basófila com um núcleo grande e vesicular. Com o crescimento da mórula, esta célula central se multiplica, constituindo o primórdio das glândulas de penetração das cercárias. As células externas da mórula são responsáveis pela formação das duas camadas celulares da cercária, constituída de fibras longitudinais e circulares. Ao mesmo tempo, também ocorre a formação de uma cutícula acelular e duas ventosas. Ao término destas transformações, as cercárias emergem do caramujo pelos espaços intercelulares e sistema venoso, estimuladas pela luminosidade e temperatura do ambiente aquático. Estima-se que um único miracídio possa dar origem a 300 mil cercárias e que para sua formação até sua saída para o meio aquático sejam

necessários de 20 a 35 dias para *S. mansoni* e *S. haematobium* e mais de 40 dias para *S. japonicum* (MAHMOUD, 2001; NEVES, 2005; TOLEDO; FRIED, 2014).

A cercária tem aspecto alongado e cilíndrico e uma cauda de extremidade curta e bifurcada fundamental para o seu deslocamento no ambiente aquático. O corpo cercariano apresenta uma ventosa oral, com terminações das chamadas glândulas de penetração, e ventosa ventral, responsável pela fixação da cercária na pele do hospedeiro definitivo (Figura 5A). Apesar da cercária sobreviver por até 72 horas no meio aquático, sua maior capacidade infectiva ocorre nas primeiras 12 horas de vida. Ao ficarem livres na água, nadam ativamente até serem atraídas por um hospedeiro definitivo. Em contato com a pele ou mucosa bucal do hospedeiro definitivo, a cercária toma a posição vertical, apoiando-se na pele pelas ventosas oral e ventral, e se fixa preferencialmente entre os folículos pilosos por ação lítica (glândulas de penetração) e ação mecânica (movimentos vibratórios intensos) promovem a penetração do corpo cercariano (MCKERROW, 2003; CARVALHO; COELHO; LENZI, 2008).

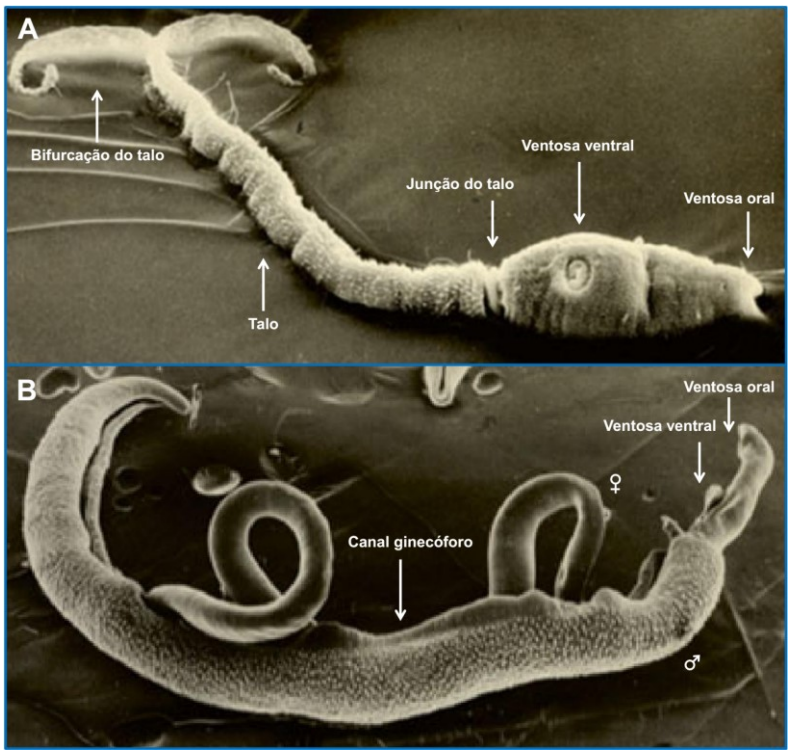



**Figura 5.** Microscopia eletrônica de varredura da forma infectante cercária (**A**) e de vermes adultos acasalados (**B**) (Adaptado de: TOLEDO; FRIED, 2014).

Após a penetração na pele do hospedeiro, diversas mudanças bioquímicas e morfológicas resultam na transformação de cercárias para esquistossômulos. As principais transformações são: perda da cauda, mudança do metabolismo aeróbio para anaeróbio, eliminação progressiva do glicocálice, esvaziamento das glândulas secretoras

e reorganização do tegumento. Concluídas estas mudanças, os esquistossômulos recém-transformados se deslocam da epiderme para a derme até atingir os pulmões. Uma vez na corrente sanguínea, pensa-se que a sua migração para os capilares sanguíneos do pulmão é inteiramente passivo. Entretanto, devido ao pequeno diâmetro dos vasos alveolares, os esquistossômulos são submetidos ao alongamento dos seus corpos e perda dos espinhos corporais. Neste momento, o organismo não se altera de modo perceptível em tamanho, mas apresenta-se morfologicamente alongado com aparência vermiforme. Entre o 7º a 21º dias de infecção, os esquistossômulos migram para o sistema porta-hepático e lá permanecem até atingir a maturidade sexual. Esse processo é concluído entre o 28º e 35º dia após a infecção (CARVALHO; COELHO; LENZI, 2008; WILSON, 2009; BOGITSH; CARTER; OELTMANN, 2013).

O macho adulto mede cerca de 1,5 cm de comprimento (2 cm para *S. japonicum*) e apresenta um sulco longitudinal, canal ginecóforo, e um tegumento claro recoberto por espinhos e tubérculos. A fêmea apresenta aspecto mais delgado, mede cerca de 2 cm de comprimento (3 cm para *S. japonicum*) e apresenta cor mais escura (acinzentada), em decorrência do sangue parcialmente digerido presente no ceco e do pigmento desta digestão, a hemozoína (Figura 5B). Tais características permitem ao macho carregar a fêmea em seu canal ginecóforo até as veias mesentéricas (plexo vesical da bexiga urinária para *S. haematobium*) e ancorar na parede do vaso sanguíneo para a oviposição (CHEEVER; DUVALL, 1982; FARRAR, 2014; TOLEDO; FRIED, 2014). Os primeiros ovos geralmente são produzidos em tomo do 35º dia após a infecção, entretanto eles só podem ser encontrados nas fezes ou urina após o 40º dia (FARRAR, 2014). O número de ovos produzidos diariamente varia de acordo com cada espécie, com a fêmea de *S. japonicum* produzindo até 3.500 ovos e as demais espécies (*S. mansoni* e S. *haematobium*) em torno de 300 ovos. Os ovos levam de seis a sete dias para tornarem-se maduros (ROSS et al., 2012; FARRAR, 2014).

O período desde a postura dos ovos até o momento de atingirem a luz intestinal ou urina (ovos de *S. haematobium*) é de aproximadamente 20 dias. Entretanto, parte dos ovos fica retida na parede intestinal, fígado ou trato urogenital provocando um processo inflamatório conhecido como reação granulomatosa. Esse processo é caracterizado por um infiltrado inflamatório celular que circunda o ovo. Consequentemente, a patologia da infecção advém principalmente do processo inflamatório nos tecidos desencadeado pela presença de ovos do esquistossomo (WARREN, 1978; ROSS et al., 2002; KING, 2009).

#### *1.1.2.1 Patologia da esquistossomose*

A esquistossomose apresenta-se no ser humano sob as formas aguda e crônica (KING, 2009). A forma aguda da esquistossomose tem maior incidência entre viajantes de curto prazo e migrantes que nunca foram expostos à infecção pelo esquistossomo. Os principais sintomas desta forma clínica podem ser caracterizados por uma reação de hipersensibilidade causada pela liberação de histamina e decorrente formação de manchas avermelhadas e prurido durante o processo de penetração do corpo cercariano na pele (CARVALHO; COELHO; LENZI, 2008; BARSOUM; ESMAT; EL-BAZ, 2013). Ainda na fase aguda, um quadro clínico denominado febre de Katayama resultante de reações de hipersensibilidade à migração de esquistossômulos e produção de ovos (14º até 84º dia pós-infecção) pode gerar sintomas como febre, cefaleia, dores abdominais, tosse não produtiva, fadiga, infiltrado pulmonar e eosinofilia (DOHERTY; MOODY; WRIGHT, 1996; ROSS et al., 2002, 2007; BOTTIEAU et al., 2006).

Na forma crônica aspectos comuns como anemia, deficiência no crescimento e desenvolvimento cognitivo são compartilhados (ROSS et al., 2002). Geralmente, a forma crônica intestinal da esquistossomose, resultante da retenção de ovos na parede do intestino, produz um quadro de inflamação, hiperplasia, ulceração, formação de microabscessos e pólipos intestinais. Os principais sintomas desta forma clínica incluem dor hipogástrica e na fossa ilíaca esquerda, estenose do cólon, presença de sangue nas fezes e intermitência entre diarreia e prisão de ventre (BOROS, 1970; CHEEVER; DUVALL, 1982; LAMBERTUCCI et al., 2000; ROSS et al., 2002; ELBAZ; ESMAT, 2013). A forma crônica hepatoesplênica causa um quadro clínico de hipertensão portal, hemorragia varicosa e esplenomegalia em decorrência da deposição de ovos no sistema porta-hepático seguida por reação inflamatória crônica granulomatosa e obstrução progressiva do fluxo de sangue (LAMBERTUCCI et al., 2000; ROSS et al., 2002; ELBAZ; ESMAT, 2013). A forma crônica no trato urinário, decorrente do acumulo de ovos de *S. haematobium* na camada muscular das paredes da bexiga urinária e ureteres, pode causar contração cicatricial nos ureteres e na junção ureterovesical. Os principais sinais clínicos desta patologia incluem refluxo urinário e infecção ascendente, disúria, hematúria, proteinúria, calcificações na bexiga, obstrução do ureter, cólica renal, hidronefrose e insuficiência renal (LAMBERTUCCI et al., 2000; ROSS et al., 2002; GRYSEELS et al., 2006; BARSOUM; ESMAT; EL-BAZ, 2013).

### 1.1.3. Controle da esquistossomose

O controle da esquistossomose representa um grande desafio nos países endêmicos, tendo em vista que a implementação de infraestrutura de saneamento básico e educação sanitária são demasiadamente caros e o controle do hospedeiro intermediário com moluscicidas é constantemente criticado pelas autoridades ambientais (TAKOUGANG et al., 2007; EVAN SECOR, 2014). Além disso, a implementação de políticas sociais que visam à implementação de hábitos seguros para uso da água muitas vezes é rejeitada por segmentos da sociedade que dependem das coleções hídricas para irrigação, atividades domésticas e lazer.

Outra importante frente de controle da esquistossomose poderia ser a vacinação, uma vez que a imunização resulta em proteção a longo prazo. O alto nível de proteção obtido pela vacinação com cercárias irradiadas (COULSON, 1997), a autocura relatada em macacos Rhesus (WILSON et al., 2008) e o uso de vacinas recombinantes veterinárias para helmintos como *Taenia ovis* (JOHNSON et al., 1989; DALTON; MULCAHY, 2001) sugerem que é possível o desenvolvimento de uma vacina contra esquistossomose.

Apesar dos grandes avanços obtidos no campo da vacinologia, ainda não existe uma vacina contra a esquistossomose aprovada para uso clínico. A ampla maioria dos antígenos vacinais estudados até o momento só reduziram parte da carga parasitária em modelos animais, mesmo com diferentes esquemas de imunização, formulação e modelos experimentais. Entretanto, três candidatos vacinais (*i.e.* glutationa *S*-transferase de 28 kD de *S. haematobium*, proteína de ligação a ácidos graxos Sm14 de *S. mansoni* e a tetraspanina de *S. mansoni*) encontram-se em estudos clínicos de fase I e II (TEBEJE et al., 2016).

Em termos gerais, a redução total da carga parasitária talvez seja inatingível com base no conhecimento científico atual e tecnologia em mãos, mas não é um pré-requisito obrigatório para uma vacina ser eficaz contra a esquistossomose. Esta afirmação um tanto antidogmática baseia-se no facto de que os esquistossomos não se replicam dentro dos seus hospedeiros definitivos e consequentemente, uma redução parcial na carga parasitária permite minimizar a patologia induzida pelos ovos e as taxas de transmissão da doença (TEBEJE et al., 2016).

### *1.1.3.1 Quimioterapia*

Atualmente, o fármaco praziquantel (PZQ) representa a única opção terapêutica para tratamento e controle da esquistossomose. Introduzido na prática clínica desde o ano de 1978, este fármaco apresenta-se como um derivado pirazino-isoquinolínico com baixa toxicidade, absorvido por via oral e eficaz contra as cinco espécies de esquistossomos que infectam humanos (GÖNNERT; ANDREWS, 1977; STELMA et al., 1995). Sua apresentação comercial é uma dos isômeros *R* (**1**) e *S* (**2**), na qual apenas o isômero *R* possui atividade farmacológica (XIAO; CATTO, 1989). As estruturas químicas de ambos os isômeros do PZQ estão destacadas na Figura 6.

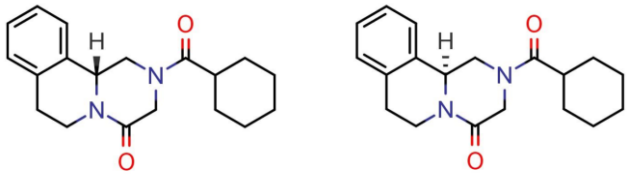



**Figura 6.** Estrutura química dos enantiômeros do PZQ.

O mecanismo de ação do PZQ ainda não é bem compreendido. Todavia, imediatamente após o contato com o fármaco, o verme apresenta atividade muscular aumentada seguida por forte contração muscular, que resulta na perda da capacidade de fixação do verme na parede do vaso sanguíneo. Danos ao tegumento também estão associados à ação do fármaco (CUPIT; CUNNINGHAM, 2015). A hipótese do mecanismo de ação mais aceita e estudada consiste na ativação da subunidade β do canal de cálcio voltagem dependente (Cavβ) e consequentemente aumento do influxo de cálcio para dentro das células (GREENBERG, 2005; NOGI et al., 2009). Outros mecanismos menos aceitos incluem ligação às proteínas miosina e actina, alteração na fluidez da membrana, inibição da recaptação de nucleotídeos, redução da concentração de glutationa (GSH) e inibição da renovação do tegumento (WIEST et al., 1992; LIMA et al., 1994; RIBEIRO et al., 1998; TALLIMA; EL RIDI, 2007; ANGELUCCI et al., 2007; GNANASEKAR et al., 2009).

Embora o PZQ seja altamente eficaz contra adultos e esquistossômulos nos sete primeiros dias de infecção, estudos *ex vivo* e *in vivo* demonstraram que os vermes jovens (*i.e.* entre 14º e 28º dia após a infecção) apresentam baixa resposta ao tratamento (SABAH et al., 1986). Portanto, indivíduos com alta carga parasitária (incluindo vermes jovens) e frequentemente expostos a locais de risco de infecção tendem a ter baixa

resposta terapêutica ao PZQ e frequentemente necessitam de repetição da quimioterapia. Consequentemente, esta característica vem levantando a preocupação da comunidade cientifica quanto ao surgimento de vermes resistentes ao PZQ. Nas últimas décadas, diversos estudos demonstraram que o esquistossomo poderia desenvolver resistência ao PZQ em condições laboratoriais controladas. Por exemplo, após administradas doses sub-terapêuticas de PZQ em camundongos ou caramujos infectados, verificou-se que as demais gerações de *S. mansoni* apresentavam menor sensibilidade ao fármaco (FALLON; DOENHOFF, 1994; ISMAIL et al., 1994; COUTO et al., 2011). Além disso, isolados de *S. mansoni* no Senegal (FALLON et al., 1995), Egito (ISMAIL et al., 1996) e Quênia (MELMAN et al., 2009) demonstraram ser menos suscetíveis ao PZQ, quando comparados com as linhagens controle.

Além disso, a quantidade de massa, tamanho, e sabor amargo (causado pelo isômero *S* na mistura racêmica) do comprimido de PZQ resulta em dificuldade de administração por via oral e em baixa adesão de crianças ao tratamento (MEYER et al., 2009). Portanto, a descoberta e planejamento de novos fármacos esquistossomicidas que apresentem novos mecanismos de ação e que possam ser utilizados em combinação com PZQ é de extrema importância (CAFFREY; SECOR, 2011; SECOR et al., 2015).

### 1.2. Planejamento e descoberta de novos fármacos

O planejamento e descoberta de novos fármacos constitui um processo complexo, competitivo e que depende da integração de várias áreas estratégicas, como inovação, conhecimento, tecnologia, gerenciamento e investimento (LOMBARDINO; LOWE, 2004). Estima-se que sejam necessários desde a concepção do projeto até a introdução de um único fármaco no mercado em média 10 anos e um investimento orçamentário que pode chegar a US$ 2,6 bilhões (MULLARD, 2014; AVORN, 2015; DIMASI; GRABOWSKI; HANSEN, 2015).

Nota-se ainda, que a competitividade e o potencial de crescimento da indústria farmacêutica está diretamente relacionado com a descoberta de *blockbusters* (*i.e.* termo em inglês utilizado para designar fármacos inovadores com vendas anuais superiores a US$ 1 bilhão por ano) (DEBNATH; AL-MAWSAWI; NEAMATI, 2010; JA, 2013). Portanto, os medicamentos líderes de vendas emergem de classes terapêuticas cujo maior mercado é constituído por países desenvolvidos, como fármacos hipoglicemiantes, antineoplásicos e moduladores do sistema nervoso central. Por outro lado, DTNs tradicionalmente não integram o quadro de investimento da indústria

farmacêutica uma vez estando associadas a condições de pobreza e sendo endêmicas apenas em países pobres.

Entretanto, os últimos anos tem testemunhado a expansão na descoberta de novas terapias para DTNs, com crescente integração e participação das indústrias farmacêuticas, universidades e organizações sem fins lucrativos como a DNDi (do inglês *Drugs for Neglected Diseases initiative*). Considerando os recursos financeiros limitados para execução de projetos nesta área, a chamada "ciência aberta", reconhecida pela implementação de um modelo de descoberta de fármacos enfatizado na partilha de informações como dados de ensaios biológicos e ensaios clínicos, tem emergido como uma alternativa para acelerar os esforços na descoberta de novos fármacos (CHATELAIN; IOSET, 2011; BOMBELLES; COAKER, 2015).

### 1.2.1. Gênese planejada de fármacos

A gênese planejada de fármacos é basicamente dividida em duas etapas, descoberta e desenvolvimento, de acordo com o esquema apresentado na Figura 7A.

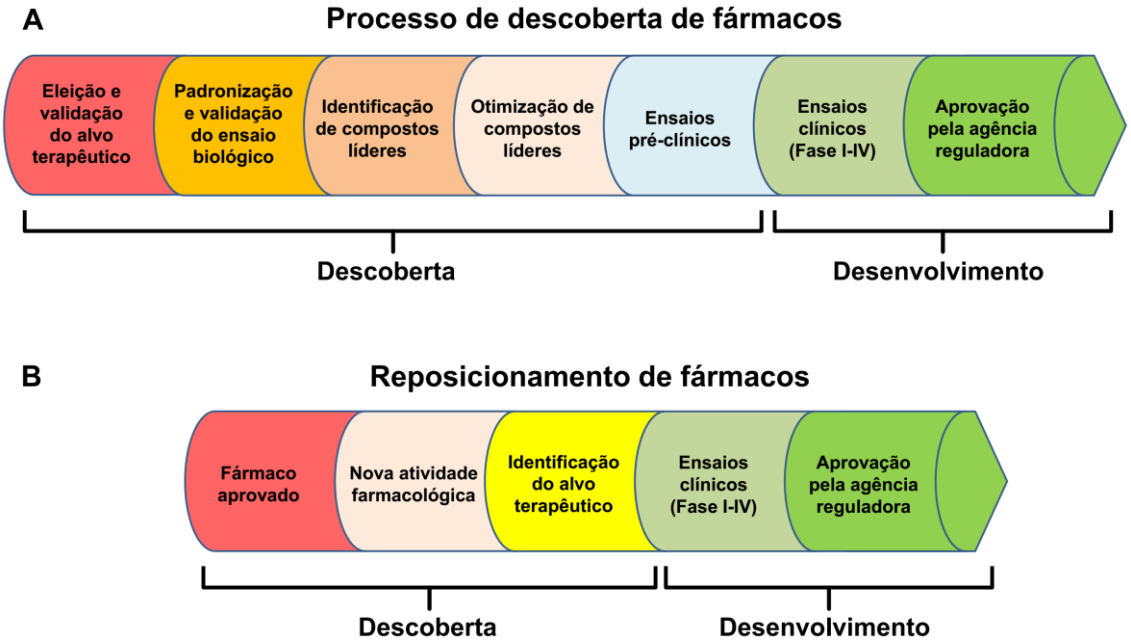



**Figura 7.** Etapas envolvidas na gênese planejada de novos fármacos (**A**) e reposicionamento de fármacos (**B**) (Adaptado de: HURLE et al., 2013).

Durante a etapa de descoberta, a pesquisa concentra-se na concepção de um novo projeto, escolha e validação do alvo terapêutico e na identificação de ligantes ativos (*hits*) para este alvo. O alvo terapêutico consiste em uma proteína (*e.g.* enzima, receptor ou transportador) ou ácido nucleico (*e.g.* DNA ou RNA) que demonstre um papel

importante no processo fisiopatológico estudado (PLOWRIGHT et al., 2012). Esta é uma etapa crítica tendo em vista que a biomacromolécula selecionada deve ser capaz de interagir com uma molécula pequena de forma que sua modulação possa produzir o efeito terapêutico desejado (HOPKINS; GROOM, 2002a). Uma vez validado o alvo (*e.g.* utilizando técnicas de eliminação do gene que codifica o alvo) a pesquisa concentra-se na padronização do ensaio biológico e na triagem de alta vazão (HTS, do inglês *High Throughput Screening*) ou triagem virtual (VS, do inglês *Virtual Screening*) de milhares de compostos com o objetivo de identificar ligantes (*hits*) que possuam o suficiente para modular ou inibir sua atividade *in vitro*. Após concluída a etapa de identificação de *hits*, os esforços são concentrados na geração de compostos líderes (do inglês, *hit-to-lead*) através do estudo da relação estrutura e atividade (SAR, do inglês *Structure-Activity Relationships*) e da relação quantitativa entre estrutura e atividade (QSAR, do inglês *Quantitative Structure-Activity Relationships*). Os compostos líderes apresentam características desejáveis para um fármaco (*i.e.* atividade biológica *in vivo*), principalmente potência e seletividade, mas ainda estão sujeitos a otimização por conta de propriedades farmacocinéticas indesejáveis *(e.g.* instabilidade metabólica e inibição de enzimas do complexo citocromo P450) e toxicidade (*e.g.* inibição da proteína hERG e hepatotoxicidade) (HUGHES et al., 2011; KATSUNO et al., 2015). Consequentemente, os líderes mais promissores podem ainda ser submetidos a otimização estrutural para correção destas propriedades.

Em termos gerais, o processo de descoberta e otimização de líderes não pode de modo algum ser considerado uma atividade rotineira. Durante as etapas de geração e otimização de compostos líderes apenas um ou dois compostos com perfil farmacodinâmico e farmacocinético adequados são selecionados como NCEs (do inglês, *New Chemical Entities*) para prosseguir para a fase de desenvolvimento clínico (HUGHES et al., 2011).

A etapa de desenvolvimento compreende estudos clínicos em humanos com o objetivo de determinar a toxicidade do candidato a fármaco em um grupo pequeno (20 a 100) de voluntários saudáveis (Fase I); avaliar a dosagem, eficácia e segurança em grupos (100 a 300) com a doença (Fase II); e comparar o candidato a fármaco com tratamento preconizado para determinar sua eficácia e propriedades farmacocinéticas em um grupo maior (até 3.000) de indivíduos (Fase III). O candidato a fármaco aprovado ao final da Fase III é então submetido à aprovação pelas agências reguladoras. Uma vez aprovado, têm-se início a fase IV, na qual o fármaco será comercializado e monitorado

quanto ao surgimento de eventuais efeitos adversos (RICK, 2015; FISHER; COTTINGHAM; KALBAUGH, 2015). Estima-se que taxa de insucesso nos estudos de Fase I chega a 36%, a 68% na Fase II e 40% na Fase III. Portanto, apenas 10% dos candidatos a fármacos em fases clínicas chegam ao mercado (HAY et al., 2014).

### 1.2.2. Alvos moleculares para o planejamento de novos fármacos

Graças aos avanços obtidos no campo da genômica e bioquímica do *S. mansoni* (BERRIMAN et al., 2009; PROTASIO et al., 2012), diversas vias metabólicas até então desconhecidas puderam ser estudadas como potenciais alvos biológicos para o planejamento de fármacos esquistossomicidas (KAWAMOTO et al., 1989; PATOCKA; RIBEIRO, 2007; BECKMANN et al., 2012; ROJO-ARREOLA et al., 2014; ZINIEL et al., 2015). Entre elas, os mecanismos bioquímicos envolvidos na relação esquistossomo-hospedeiro vem sendo considerados particularmente eficazes a medida que eles permitem ao esquistossomo neutralizar a resposta imune produzida pelo hospedeiro (PEARCE; SHER, 1987; MAIZELS et al., 1993; PEARCE; MACDONALD, 2002). A Figura 8 representa um modelo hipotético dos prováveis mecanismos de evasão do esquistossomo à resposta imune gerada pelo hospedeiro.

Os principais mecanismos de defesa estão envolvidos na: renovação da membrana tegumentar (WILSON; BARNES, 1977; BROUWERS et al., 1999; HAN et al., 2009); incorporação de proteínas séricas (*e.g.* albumina) do hospedeiro na sua superfície tegumentar (LIU et al., 2007); produção de antígenos com epítopos idênticos aos do hospedeiro e sua incorporação na superfície do tegumento (JONES et al., 2004; BRASCHI et al., 2006; BRASCHI; WILSON, 2006; LIU et al., 2006; PÉREZ-SÁNCHEZ et al., 2006); regulação da resposta imune gerada pelo hospedeiro, resultando na indução de apoptose de células de defesa (SALZET; CAPRON; STEFANO, 2000; CHEN et al., 2002; HAN et al., 2009); produção e expressão de proteases no tegumento e meio externo para proteólise de anticorpos (HAN et al., 2009); e a neutralização de espécies reativas de oxigênio (EROs) e nitrogênio (ERNs) por enzimas antioxidantes (MKOJI; SMITH; PRICHARD, 1988; CHIUMIENTO; BRUSCHI, 2009; PERBANDT; NDJONKA; LIEBAU, 2014).

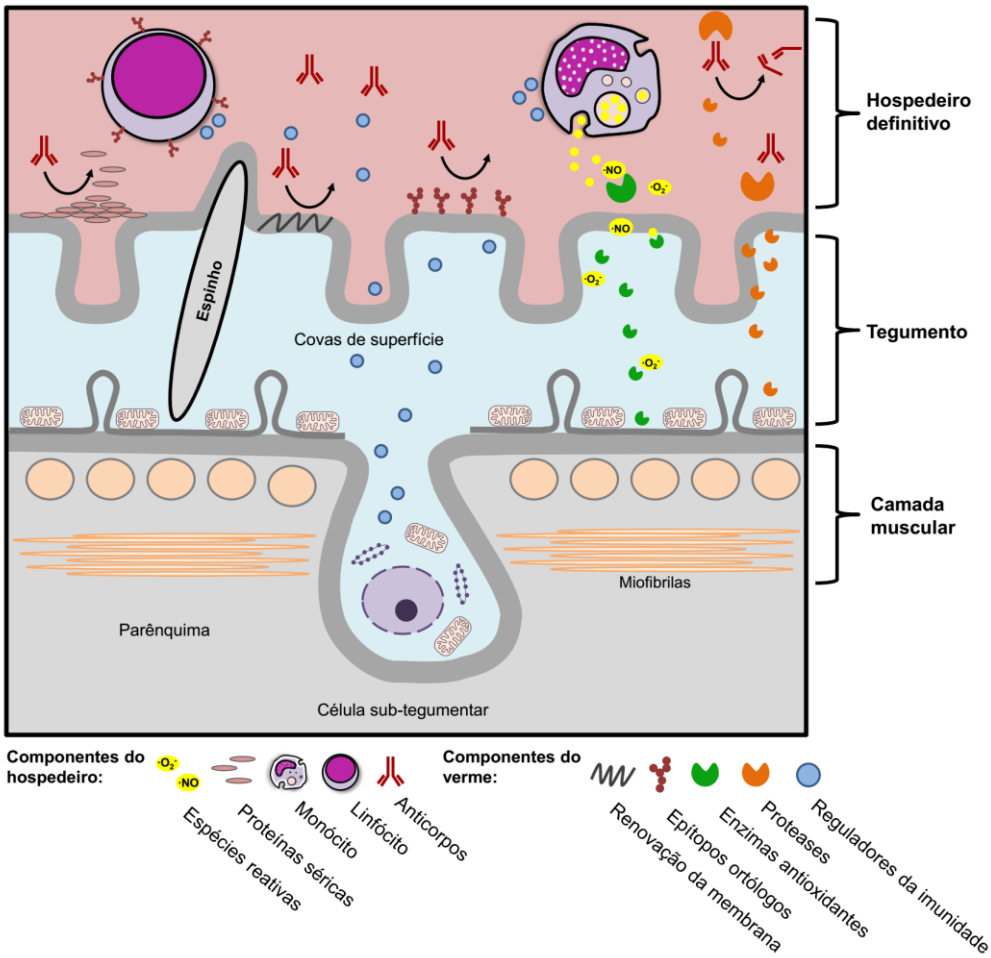



**Figura 8.** Modelo hipotético dos prováveis mecanismos de evasão do esquistossomo à resposta imune gerada pelo hospedeiro (Adaptado de: HAN et al., 2009).

### *1.2.2.1 O sistema antioxidante de esquistossomos*

As EROs e ERNs podem produzir respostas agressivas contra o esquistossomo e comprometer sua integridade metabólica e estrutural através de mecanismos inespecíficos como peroxidação de lipídeos, destruição da membrana celular, inativação de enzimas e ligação ao ácido desoxirribonucleico (DNA) (PAL; BANDYOPADHYAY, 2012; PERBANDT; NDJONKA; LIEBAU, 2014). Estas espécies reativas geralmente são produtos de enzimas específicas do hospedeiro, tais como a nicotinamida adenina dinucleótido fosfato (NADPH) oxidase de fagócitos e oxido nítrico redutase indutiva (NOS II) (CHIUMIENTO; BRUSCHI, 2009), de mecanismos combinatórios tais como a combinação de ·NO e $\cdot O_2^-$ para formar o radical peroxinitrito ($ONOO^-$) (VALKO et al., 2007) ou mecanismos reacionais como a produção do radical hidroxil (·OH) a partir da reação de peróxido de hidrogênio ($H_2O_2$) e ferro no estado ferroso ($Fe^{2+}$) (DRÖGE, 2002; PERBANDT; NDJONKA; LIEBAU, 2014), um

subproduto do consumo de hemoglobina pelo esquistossomo (ZUSSMAN; BAUMAN; PETRUSKA, 1970; TRUSCOTT et al., 2013).

Notadamente, os últimos anos tem testemunhado a caracterização, clonagem e expressão de várias enzimas do sistema antioxidante de *S. mansoni* (MEI et al., 1995, 1996; ALGER; WILLIAMS, 2002; SAYED; WILLIAMS, 2004; KUNTZ et al., 2007; OKE; MOSKOVITZ; WILLIAMS, 2009). Na Figura 9 estão representados os principais mecanismos bioquímicos desenvolvidos pelo esquistossomo para neutralizar EROs e ERNs.

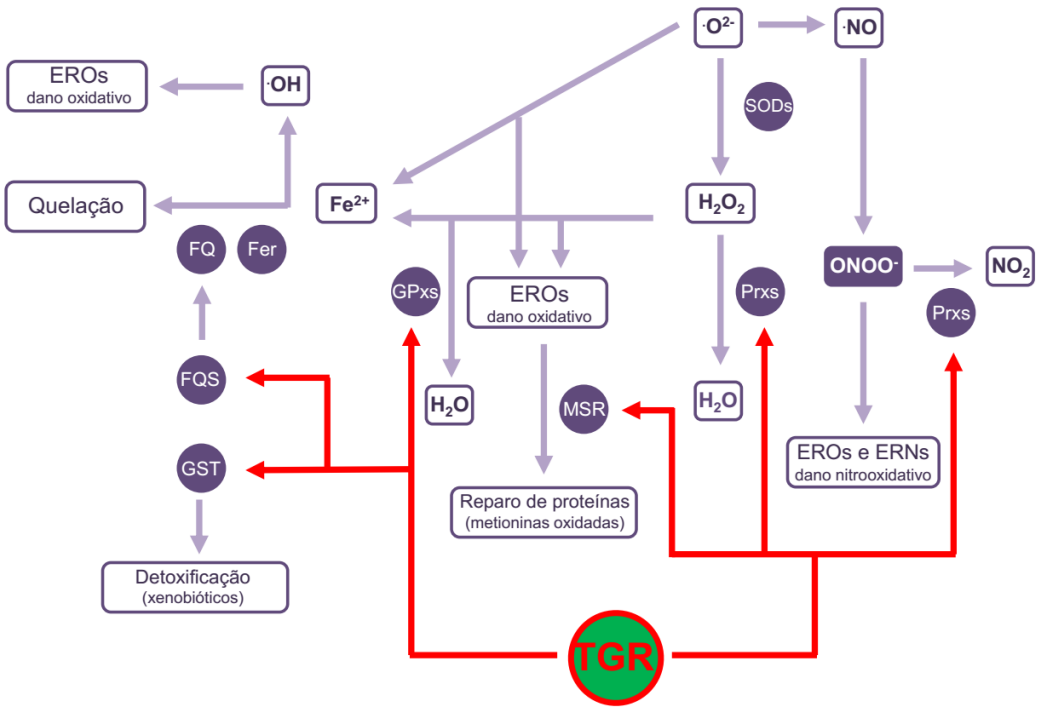



**Figura 9.** Representação dos mecanismos bioquímicos envolvidos no sistema antioxidante de esquistossomos. Superóxido ($\cdot O^{2-}$), óxido nítrico ($\cdot NO$), peroxinitrito (ONOO-), ferro no estado ferroso ($Fe^{2+}$), peróxido de hidrogênio ($H_2O_2$), espécies reativas de oxigênio (EROs), espécies reativas de nitrogênio (ERNs), superóxido dismutases (SODs), peroxirredoxinas (Prxs), glutationa peroxidases (GPxs), metionina sulfóxido redutase (MSR), glutationa *S*-transferase (GST), ferritina (Fer), fitoquelatina (FQ), fitoquelatina sintase (FQS) e tiorredoxina glutationa redutase (TGR).

A primeira etapa no processo de neutralização de EROs pelo esquistossomo consiste na conversão de $\cdot O^{2-}$ em $H_2O_2$ pelas enzimas superóxido dismutases (SODs) (HONG et al., 1992, 1993; MEI et al., 1995). Na etapa seguinte, $H_2O_2$ pode ser reduzido classes de enzimas: glutationa peroxidases (GPxs) e peroxirredoxinas (Prx) (WILLIAMS et al., 2013). As GPxs catalisam a redução de $H_2O_2$ ou hidroperóxidos orgânicos em água e álcoois utilizando o tripeptídeo glutationa (GSH) como equivalente redutor (MEI et al.,

1996; HUANG; RIGOUIN; WILLIAMS, 2012; WILLIAMS et al., 2013). A GPx é altamente expressa no tegumento e epitélio do intestino de esquistossomos, conferindo-lhes um provável papel na proteção das membranas celulares contra o dano oxidativo (MEI et al., 1996). O mecanismo de neutralização de $H_2O_2$ por Prxs é semelhante ao de GPxs. Entretanto, as Prxs podem utilizar GSH ou a proteína tiorredoxina (Txr) como equivalentes redutores (SAYED; WILLIAMS, 2004; RHEE; CHAE; KIM, 2005; SAYED; COOK; WILLIAMS, 2006).

Em outra etapa da via bioquímica, $Fe^{2+}$ intracelular pode ser quelado pela enzima ferritina (Fer) (DIETZEL et al., 1992) e fitoquelatina (FQ), um oligopeptídeo produzido a partir da polimerização de GSHs pela enzima fitoquelatina sintase (FQS), ajudando desta forma a suprimir seu efeito pró-oxidante (RAY; WILLIAMS, 2011). Outros mecanismos como o reparo de proteínas danificadas por EROs pela enzima metionina sulfóxido redutase (MSR), que catalisa a redução de metionina sulfóxido (Met-O) em metionina (Met) utilizando Trx como equivalente redutor (RUAN et al., 2002; KANTOROW et al., 2007; OKE; MOSKOVITZ; WILLIAMS, 2009) e a redução de compostos eletrofílicos com caráter hidrofóbico (*e.g.* produtos da peroxidação lipídica) pela enzima GST (mediado pela doação de elétrons pela GSH ), também podem contornar e prevenir os dados causados por EROs (MEI; LOVERDE, 1997; WILLIAMS et al., 2013).

#### *1.2.2.1.1 Tiorredoxina glutationa redutase*

Dado o envolvimento da GSH e Trx no mecanismo reacional de neutralização de EROs e ERNs, seu papel pode ser considerado fundamental na manutenção da integridade e funcionamento sistema antioxidante de esquistossomos. Em mamíferos, o sistema Trx é regulado pela flavoenzima tiorredoxina redutase (TrxR), a qual é responsável pela transferência de elétrons a partir de NADPH para Txr oxidada (Trx-$S_2$), convertendo-a novamente para sua forma reduzida (Figura 10A). Este sistema fornece Trx para uma variedade de enzimas que contêm pares de Cisteína (Cys) ativos, tais como Prxs e MSR. Já o sistema GSH de mamíferos é regulado pela enzima glutationa redutase (GR). Tal como acontece com a TrxR, a GR transfere elétrons de NADPH exclusivamente para GSH dissulfeto (GSSG), convertendo-a novamente em GSH. Em seguida, a GSH pode transferir elétrons para glutarredoxina oxidada, convertendo-a para seu estado reduzido (Figura 10A). Além disso, a GSH funciona como um tampão e atua como

antioxidante, proporcionando elétrons para enzimas como GPxs, ou servindo como substrato para a síntese de FQ (WILLIAMS et al., 2013).

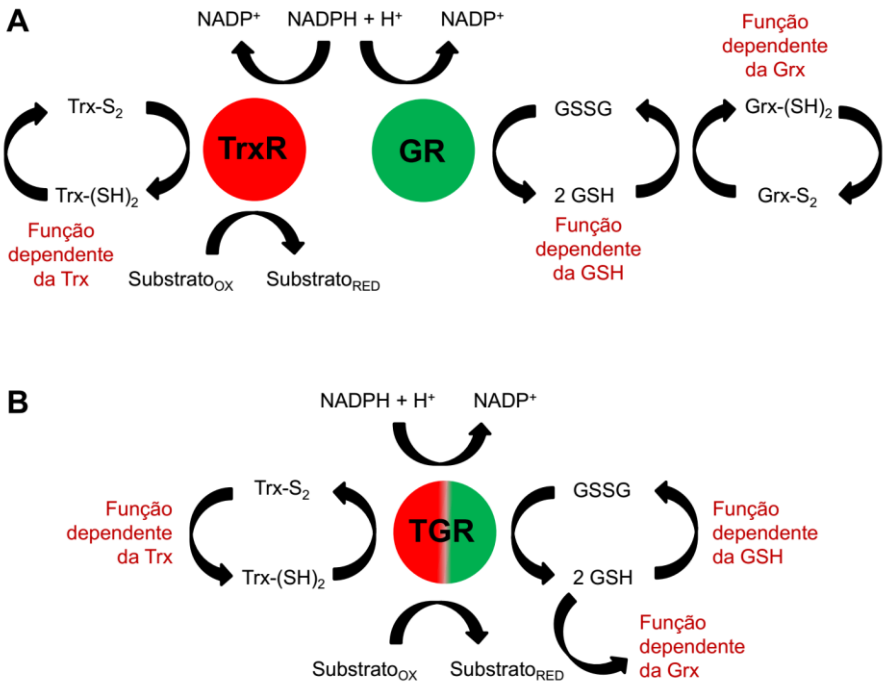



**Figura 10.** Regulação dos sistemas GSH e Trx em mamíferos (**A**) e no esquistossomo (**B**) (Adaptado de KUNTZ et al., 2007).

Em trematódeos, as enzimas GR e TrxR são substituídas por uma única enzima multifuncional (Figuras 9 e 10B), a tiorredoxina glutationa redutase (TGR), a qual é capaz de reduzir não apenas GSSG e Trx-$S_2$, mas compostos como $H_2O_2$ e proteínas glutationiladas (HUANG et al., 2011; PRAST-NIELSEN; HUANG; WILLIAMS, 2011). Estudos de silenciamento dos genes das TGRs de *S. mansoni* (*Sm*TGR) e *S. japonicum* (*Sj*TGR) utilizando a técnica de RNA interferente (RNAi) demonstraram que a inibição da sua expressão foi capaz de causar a morte de ambas as espécies. Além disso, a TGR humana possui baixa identidade sequencial com a *Sm*TGR e *Sj*TGR (≤ 54%), e sua provável função está relacionada à isomerização de ligações dissulfeto de proteínas que formam componentes estruturais dos espermatozoides (SUN et al., 2005; SU et al., 2005). Portanto, ambas *Sm*TGR e *Sj*TGR são alvos validados para o planejamento de novos agentes esquistossomicidas (KUNTZ et al., 2007; SONG et al., 2012; HAN et al., 2014).

A *Sm*TGR é uma flavoproteína homodimérica com monômeros organizadas de forma semelhante à letra "W", na qual cada monômero apresenta componentes do domínio TrxR e do domínio GR. A Figura 11 apresenta a estrutura tridimensional (3D) dos monômeros de *Sm*TGR obtidos por cristalografia de raios-X (Código PDB: 2X8C) e

seu arranjo bidimensional (2D). O domínio TrxR (destacado em vermelho) é composto de resíduos de ambos monômeros: flavina adenina dinucleotídeo (FAD); um par Cys154/Cys159 proveniente de um monômero; e um par Cys596'/Sec597' proveniente do outro monômero. Este último par se encontra na alça C-terminal de cada monômero (destacado em azul). Ainda no domínio TrxR, existe um par Cys520/Cys574, mas sua função ou envolvimento no mecanismo catalítico ainda não foi comprovada. Já o domínio GR (destacado em verde) contém apenas um par Cy-28/Cys31 (ANGELUCCI et al., 2010).

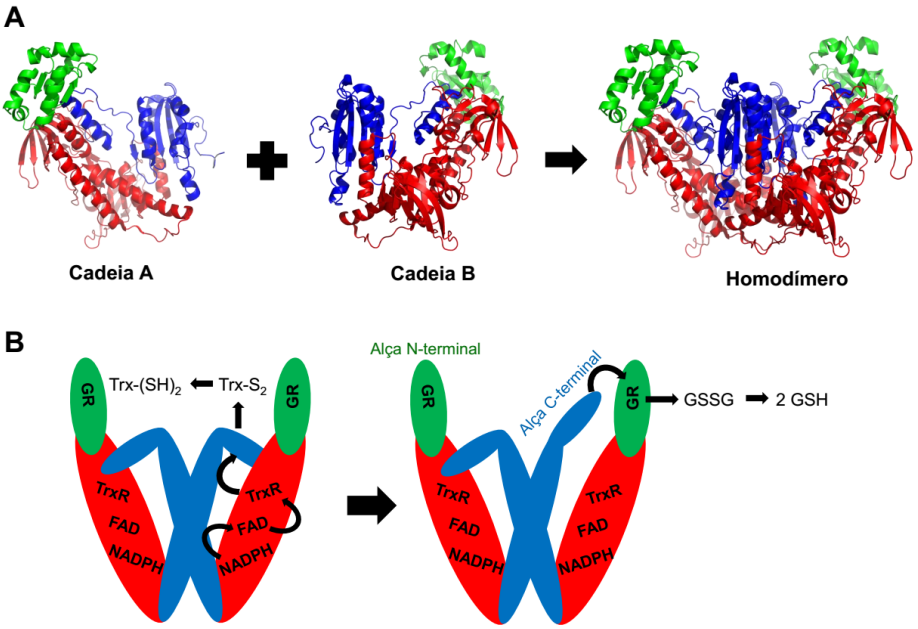



**Figura 11.** Representação do arranjo estrutural formado por monômeros da *Sm*TGR (Código PDB: 2X8C, **A**) e seus principais domínios (**B**) (Adaptado de: WILLIAMS et al., 2013). Domínios TrxR, GR e alça C-terminal estão destacados nas cores vermelha, verde e azul, respectivamente.

Embora o mecanismo catalítico da *Sm*TGR ainda não foi completamente elucidado, estudos bioquímicos e estruturais indicam que ele se baseia na transferência de elétrons entre os domínios estruturais. As figuras 11B e 12 apresentam um modelo hipotético do mecanismo catalítico da *Sm*TGR. Nele, os elétrons presentes no NADPH são transferidos para FAD e em seguida para um par de Cys154/Cys159 adjacente presente no domínio TrxR. Em seguida, os elétrons são transferidos para o par Cys596'/Sec597' da alça C-terminal, e se necessário, são transferidos para o par de Cys28/Cys31 presente no domínio GR. Portanto, acredita-se que a alça C-terminal apresente alta mobilidade para aceitar os elétrons do domínio TrxR e transferi-los para o domínio GR (ANGELUCCI et al., 2010; HUANG et al., 2011).

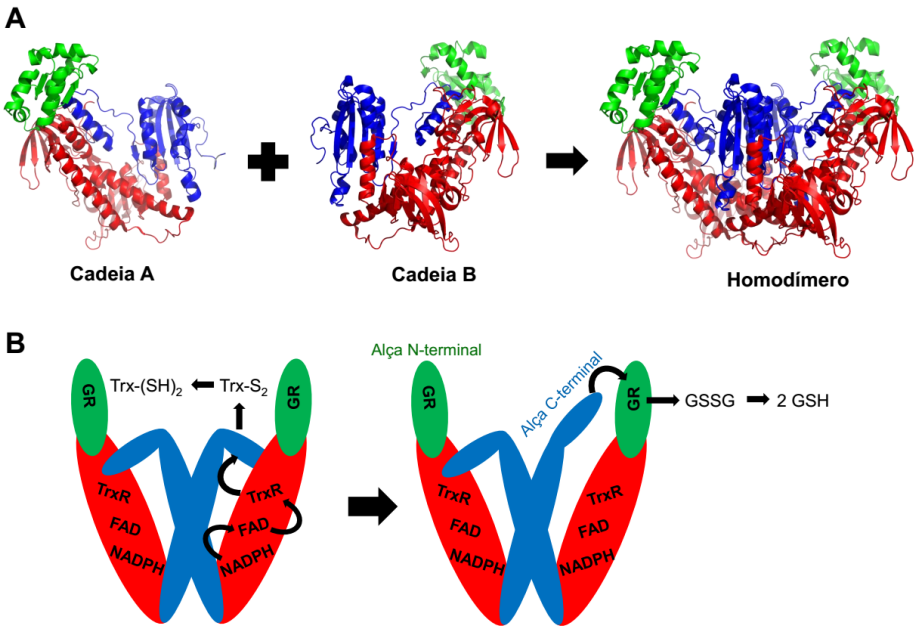



**Figura 12.** Representação do arranjo estrutural formado por monômeros da *Sm*TGR (Código PDB: 2X8C, **A**) e seus principais domínios (**B**) (Adaptado de: WILLIAMS et al., 2013). Domínios TrxR, GR e alça C-terminal estão destacados nas cores vermelha, verde e azul, respectivamente.

Dada estas características, observa-se que a *Sm*TGR pode aceitar dois pares de elétrons em cada subunidade, embora a forma cataliticamente ativa provavelmente apresenta apenas um par de elétrons no domínio TrxR, GR ou alça C-terminal (Figura 12) (ANGELUCCI et al., 2010; HUANG et al., 2011). O provável mecanismo catalítico envolvendo o domínio GR ocorre utilizando apenas uma Cys via um mecanismo de deglutationilação, que a princípio realiza um ataque nucleofílico do íon tiolato da Cys28 na ligação dissulfeto de GSSG ou proteínas glutationiladas, formando assim uma ligação dissulfeto entre GSH e Cys28. Em seguida, o íon tiolato de Cys31 realiza um ataque nucleofílico entre a ligação dissulfeto de GSH e Cys28, levando a formação de uma nova ligação entre Cys28/Cys31 e a liberação de GSH. Sugere-se ainda que o ambiente químico do domínio de TrxR ao redor do par Cys154/Cys159 pode facilitar a catálise de GSSG em GSH, apesar de sua eficiência catalítica ser relativamente menor. Curiosamente, nesta região também existe uma díade catalítica que poderia auxiliar na abstração do próton da Cys159 (PRAST-NIELSEN; HUANG; WILLIAMS, 2011). Ainda não existem dados detalhados sobre o mecanismo catalítico de Trx pelo domínio TrxR ou alça C-terminal da *Sm*TGR. Entretanto, sabe-se que mutação Sec597Cys reduz significativamente, mas não totalmente, a atividade catalítica da *Sm*TGR em relação ao

seu tipo selvagem. Portanto a Sec597 é importante, mas não essencial, para atividade catalítica da *Sm*TGR (PRAST-NIELSEN; HUANG; WILLIAMS, 2011).

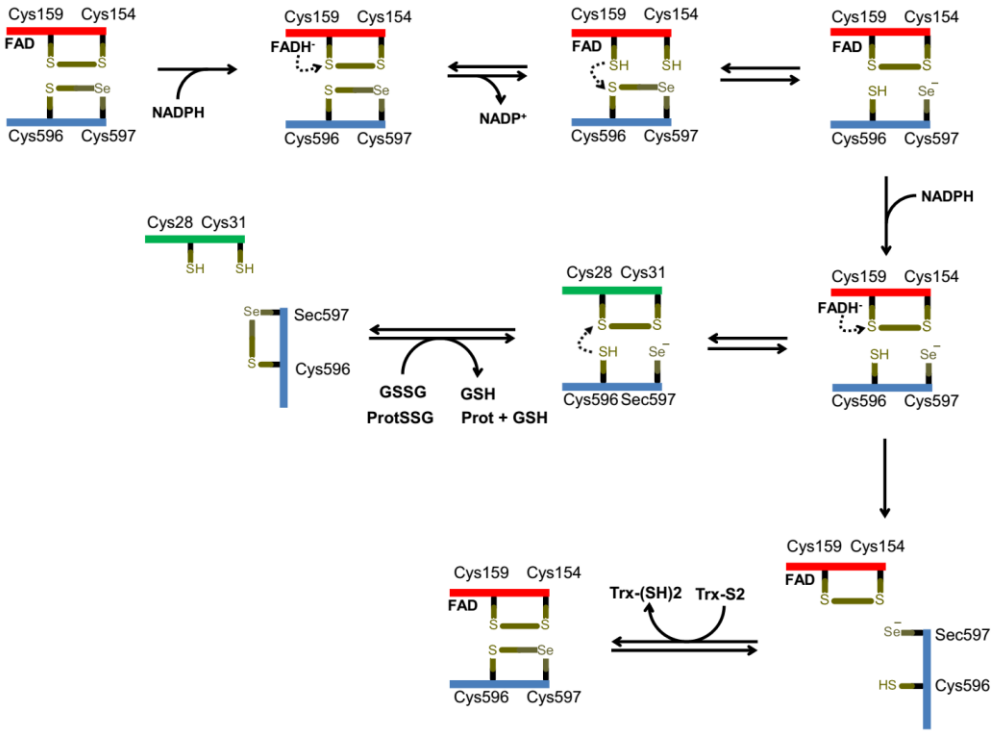



**Figura 13.** Modelo hipotético das principais etapas envolvidas no mecanismo catalítico da *Sm*TGR (Adaptado de: HUANG et al., 2011).

*1.2.2.1.1.1 Inibidores de Sm*TGR

Os primeiros inibidores da *Sm*TGR foram identificados no ano de 2007 a partir do desenvolvimento e padronização do ensaio enzimático. Neste estudo, a auranofina (**3**, Figura 13), um metalofármaco utilizado no tratamento da artrite reumatoide, foi identificada como um potente inibidor da *Sm*TGR ($IC_{50}$ = 7 nM). Posteriormente, um ensaio *ex vivo* com este inibidor a uma concentração de 10 µM causou o desemparelhamento de machos e fêmeas adultos após uma hora de exposição e 100% de mortalidade dos vermes após nove horas de exposição. Este estudo demonstrou pela primeira vez que a inibição da enzima *Sm*TGR poderia resultar na redução expressiva do número de vermes viáveis (KUNTZ et al., 2007).

Com a padronização do ensaio enzimático, diversas campanhas de HTS com a *Sm*TGR foram realizadas nos últimos anos (SIMEONOV et al., 2008; LEA et al., 2008; LI et al., 2015). Diversas classes de inibidores foram identificadas, incluindo o furoxano (**4**, Figura 13), um oxadiazol doador de óxido nítrico capaz de causar a morte de todos os estádios intra-hospedeiro de *S. mansoni* em estudos *ex vivo* e *in vivo* (RAI et al., 2009).

Atualmente, milhares de inibidores da *Sm*TGR estão sobre domínio público e, portanto, constituem as bases moleculares para o planejamento de novos fármacos esquistossomicidas.

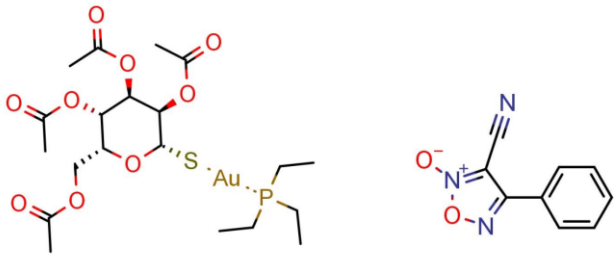



**Figura 14.** Estrutura química da auranofina (**3**) e do composto furoxano (**4**).

### 1.2.3. Estratégias *in silico* para o planejamento de novos fármacos

A concepção e desenvolvimento de projetos envolvendo ferramentas computacionais para a descoberta, planejamento e otimização novos candidatos a fármacos consiste em um campo de estudo em constante expansão denominado planejamento de fármacos auxiliado por computador (CADD, do inglês *Computer-Aided Drug Design*) (WERMUTH et al., 1998). Tal expansão se deve aos grandes avanços obtidos, nas últimas décadas, nas áreas de *software* e *hardware*. Somados a esses avanços, também se destaca o crescimento experimentado pelas áreas de genômica, proteômica, e biologia molecular e estrutural, que permitiram o acesso a um grande número de informações químicas e biológicas de ligantes e alvos biológicos. Atualmente, o uso de métodos computacionais permeia todos os aspectos da descoberta de fármacos e apresenta a vantagem de auxiliar no planejamento e descoberta de fármacos de forma mais rápida e com custos reduzidos (OOMS, 2000; TALELE; KHEDKAR; RIGBY, 2010; BRAGA et al., 2014a; JOY MACALINO et al., 2015).

Diferentes abordagens podem ser utilizadas para o planejamento de novos candidatos a fármacos, dependendo do nível de conhecimento disponível para execução do projeto. Quando a estrutura 3D do alvo macromolecular ou do complexo ligante-receptor estão disponíveis, o planejamento de fármacos é baseado na estrutura (SBDD, do inglês *Structure-Based Drug Design*). Entretanto, quando a estrutura 3D do alvo biológico não está disponível, o planejamento pode ser baseado na estrutura de ligantes conhecidos ou do ligante endógeno (LBDD, do inglês *Ligand-Based Drug Design*) (YOUNG, 2009; MERZ; RINGE; REYNOLDS, 2010; KLEBE, 2013). Em muitos casos,

o uso combinado das duas estratégias, pode ser útil no processo de descoberta de novos fármacos, gerando informações adicionais, fruto do sinergismo entre abordagens (BRAGA et al., 2014a).

#### *1.2.3.1 QSAR*

Nas últimas décadas, a capacidade de coletar, analisar e armazenar todos os tipos de informações químicas e biológicas se desenvolveu rapidamente. Técnicas modernas como a química combinatória e o HTS levaram ao acúmulo de informações químicas e biológicas em bases de dados. Atualmente, conjuntos de dados contendo milhares de compostos testados em múltiplos ensaios biológicos estão disponíveis em bases de dados como ChEMBL (GAULTON et al., 2012) e PubChem (WANG et al., 2012) para toda a comunidade científica. As informações contidas nestas bases de dados podem ser utilizadas para a construção e validação de modelos computacionais preditivos e úteis para a VS de novos ligantes (XIE, 2010; VAN NOORDEN, 2012). Nesse contexto, estudos de QSAR apresentam várias aplicações, tais como: predizer a atividade biológica de novos compostos (*e.g.* potência e seletividade), racionalizar o mecanismo de ação dos compostos estudados e discriminar compostos com propriedades biológicas indesejáveis (*e.g.* reativos, tóxicos e com efeitos farmacocinéticos inapropriados) (CHERKASOV et al., 2014; ALVES et al., 2016; RAIES; BAJIC, 2016).

##### *1.2.3.1.1 Princípios*

A abordagem de QSAR pode ser descrita como um método estatístico de análise de dados para desenvolver modelos matemáticos que possam predizer corretamente a atividade biológica de compostos baseando-se em relações estabelecidas entre estrutura química e atividade biológica. Portanto, dois tipos de informações são necessários para a construção de modelos de QSAR: (i) ter um conjunto de compostos com atividade biológica definida e (ii) descritores moleculares calculados para estes compostos (Figura 14). Obedecidas estas condições, uma relação pode ser estabelecida pela aplicação de métodos estatísticos que, genericamente a definem como $A_b = k'(D_1, D_2, ..., D_n)$, na qual $A_b$ é a atividade biológica do composto, D são descritores moleculares e $k'$ é um peso atribuído aos descritores por um método de aprendizado de máquina ou regressão. Em outras palavras, o método estatístico estabelece pesos aos descritores ajustando-os a equação que relaciona a estrutura química com a atividade biológica

(TROPSHA, 2010). Uma breve discussão dos descritores moleculares e métodos de aprendizado de máquina está disponível nas próximas seções.

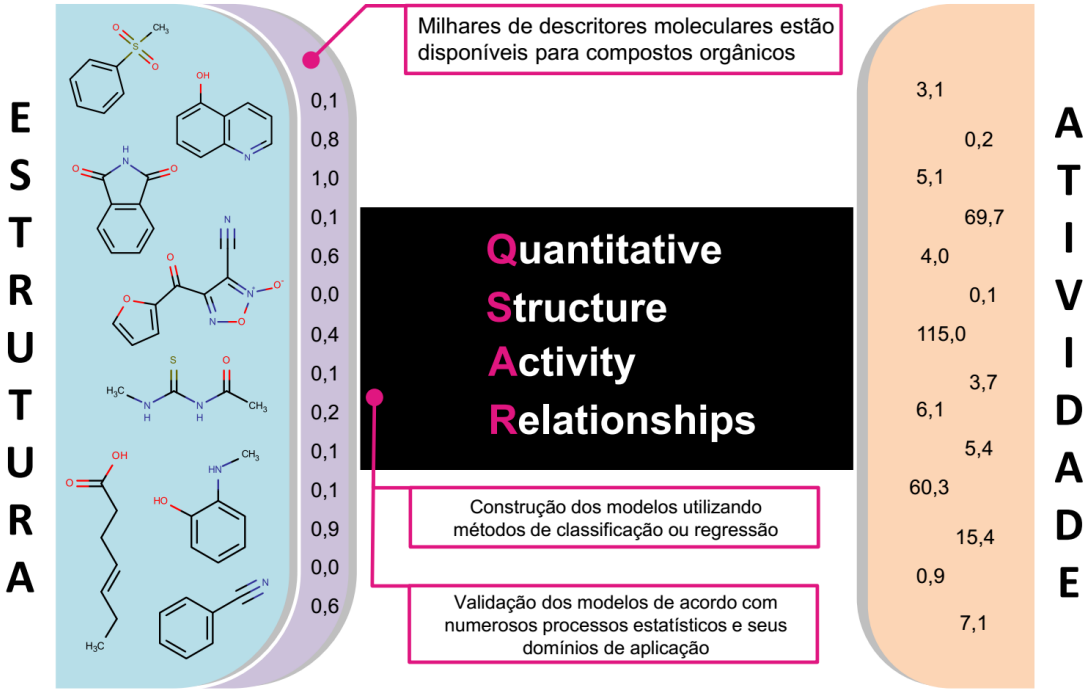



**Figura 15.** Representação do processo de construção de um modelo de QSAR (Adaptado de: TROPSHA, 2010).

#### *1.2.3.1.2 Descritores moleculares*

O descritor molecular é o resultado final de um procedimento matemático e lógico que transforma informação química codificada pela representação simbólica do composto em um número útil (TODESCHINI; CONSONNI, 2000, 2009; BUCKLE et al., 2013). O campo de estudo dos descritores moleculares é interdisciplinar e envolve diferentes tipos de teoria, tais como conhecimentos em álgebra, química orgânica, físico-química, teoria dos grafos, teoria da informação e química computacional. Atualmente, milhares de descritores moleculares estão disponíveis (TODESCHINI; CONSONNI, 2000, 2009; YAP, 2011; LANDRUM, 2014). Estes podem ser classificados em unidimensionais (1D), baseados em propriedades físico-químicas e da fórmula molecular (*e.g.* refratividade molar, logP e massa molecular); bidimensionais (2D), que descrevem propriedades que podem ser representadas como ligações e átomos (*e.g.* número de átomos de nitrogênio e oxigênio, número de ligações e índices de conectividade); e tridimensionais (3D), que dependem da conformação das moléculas (*e.g.* área de superfície acessível ao solvente e volume molecular) (XUE; BAJORATH, 2000). Outra classificação que diz respeito à natureza desses descritores, os classifica como: (i) quanto-

mecânicos (*e.g.* energia dos orbitais moleculares); (i) topológicos (*e.g.* índices de Randi'c); (iii) eletrostáticos (*e.g.* índices de polaridade e carga parciais); (iv) geométricos (*e.g.* volume molecular, área de superfície polar); (v) constitucionais (*e.g.* peso molecular, números de átomos e ligações); e (vi) impressões digitais moleculares (do inglês *fingerprints*), que calculam a presença ou frequência de determinados fragmentos na molécula (*e.g. fingerptints* Morgan e Avalon) (TODESCHINI; CONSONNI, 2009; ROGERS; HAHN, 2010; KOCKAR et al., 2010).

#### *1.2.3.1.3 Métodos de aprendizado de máquina*

O aprendizado de máquina é um ramo da inteligência artificial que estuda o desenvolvimento de sistemas capazes de adquirir conhecimento de forma automática. O sistema de aprendizado representa um algoritmo que toma decisões baseado em experiências acumuladas através da solução bem sucedida de problemas anteriores. Esse aprendizado é construído em três estágios: (i) representação dos dados; (ii) otimização da hipótese; e (iii) generalização. Por exemplo, durante a construção de modelos de QSAR, os compostos podem ser representados como descritores moleculares. Em seguida, uma hipótese (equação) é gerada na tentativa de estabelecer uma relação dos descritores com a atividade biológica e essa relação é otimizada. Supondo que a atividade biológica Y investigada é descrita em função de um único descritor molecular *x*, esta hipótese poderia ser estabelecida pela Equação 1 (WELLING, 2011; LAVECCHIA, 2015):

$$Y(x) = a + bx \qquad (\mathbf{1})$$

Com o objetivo de encontrar valores para os parâmetros *a* e *b*, algoritmos computacionais são utilizados para minimizar esses valores, de forma que se ajustem e sejam capazes de descrever os dados de entrada (conjunto treinamento). No final, é testada a capacidade de generalização da hipótese gerada. O termo generalizar diz respeito à habilidade da hipótese em predizer corretamente a atividade biológica de compostos não inseridos na construção do modelo (conjunto teste) (WELLING, 2011; MITCHELL, 2014).

#### *1.2.3.1.4 Boas práticas de desenvolvimento e validação em QSAR*

Nos últimos anos, várias diretrizes e recomendações de boas práticas de construção e validação de modelos de QSAR foram discutidas e publicadas (DEARDEN; CRONIN; KAISER, 2009; TROPSHA, 2010; CHERKASOV et al., 2014). A principal

delas, publicada no ano de 2004 pela OECD (do inglês *Organization for Economic Co-operation and Development*) sugere que o processo de construção e validação de modelos de QSAR para uso prático siga cinco princípios: (i) ter atividade biológica definida; (ii) utilizar algoritmos claros; (iii) ter domínio de aplicabilidade (DA) definido; (iv) apresentar robustez e preditividade apropriadas; e (v) realizar a interpretação mecanística dos modelos, ou seja, encontrar relações entre os descritores e a atividade biológica, visando a melhor compreensão do mecanismo de ação de uma estrutura química ou aprofundar o conhecimento biológico sobre a propriedade em estudo (OECD, 2004).

## 1.3. Reposicionamento de fármacos

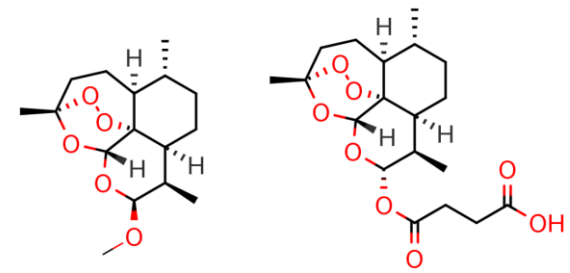
Tendo em vista o tempo, investimento e a elevada taxa de insucesso no processo de descoberta e desenvolvimento de novos fármacos, a estratégia de reposicionamento de fármacos, isto é, a descoberta de novas aplicações terapêuticas a partir de fármacos aprovados ou candidatos a fármacos em estudos clínicos, representa uma opção viável para reintrodução e aplicação de fármacos conhecidos para tratamento de doenças graves (ASHBURN; THOR, 2004). Esta abordagem tem valor comercial significativo, pois o tempo e custos necessários para a pesquisa são relativamente menores, uma vez que o desenvolvimento de formulações, a produção em escala comercial, o perfil farmacocinético e a disponibilidade de estudos clínicos de Fase I já são padronizados e, portanto, podem ser u utilizados e apresentados às agências reguladoras perante uma nova solicitação de registro (Figura 7B, página 12). Consequentemente, a abordagem de reposicionamento de fármacos pode reduzir significativamente o tempo necessário até sua aprovação pelas agências reguladoras (PAMMOLLI; MAGAZZINI; RICCABONI, 2011; HURLE et al., 2013). Exemplos de fármacos reposicionados incluem a sildenafila, originalmente desenvolvida para tratamento da hipertensão arterial e angina e reposicionada de forma revolucionaria para ser o primeiro tratamento da disfunção eréctil no homem; e miltefosina, originalmente utilizada no tratamento do câncer e reposicionada para tratamento da leishmaniose visceral (ASHBURN; THOR, 2004; BERMAN, 2015).

### 1.3.1. Reposicionamento de fármacos para esquistossomose

Nos últimos anos, grandes avanços foram obtidos no campo de reposicionamento de fármacos para a esquistossomose graças a avaliação biológica *ex vivo* de grandes bibliotecas de fármacos (ABDULLA et al., 2009; COWAN; KEISER,

2015; PANIC et al., 2015) em plataformas de triagem com alta vazão (MARCELLINO et al., 2012; PAVELEY et al., 2012; MANSOUR et al., 2016). Em paralelo, avanços no conhecimento acerca da biologia e bioquímica do verme também possibilitaram a identificação de vias metabólicas exploradas inicialmente em mamíferos que também são essenciais para sobrevivência do esquistossomo. Consequentemente, a atividade esquistossomicida de várias classes de fármacos foi identificada, incluindo estatinas (*i.e.* inibidores da 3-hidroxi-3-metil-glutaril-coenzima A redutase) (ROJO-ARREOLA et al., 2014), inibidores de enzimas do complexo citrocromo P450 (ZINIEL et al., 2015) e antineoplásicos (COWAN; KEISER, 2015; GELMEDIN; DISSOUS; GREVELDING, 2015; LEE; FAIRLIE, 2015)

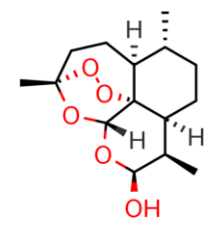
No entanto, as principais descobertas no campo foram obtidas ainda na década de 1980 com a identificação da atividade esquistossomicida dos antimaláricos artemeter (**5**), artesunato (**6**) e diidroartemisinina (**7**) frente a vermes jovens das três principais espécies de esquistossomo (Figura 15) (LIU et al., 2014; SAEED et al., 2016). Ambos os fármacos não apresentam atividade esquistossomicida satisfatória contra vermes adultos e o seu mecanismo exato de ação ainda não é bem definido. Non entanto, estudos preliminares sugerem que o mecanismo de ação das artemisininas esteja envolvido na clivagem do grupo endoperóxido do fármaco mediada pela interação com $Fe^{2+}$ de grupos heme. Esta interação resultaria na geração de espécies radicalares lesivas para os componentes celulares do verme (WU et al., 2003).



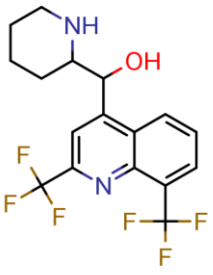
artemeter (**5**)



artesunato (**6**)



diidroartemisinina (**7**)



mefloquina (**8**)

**Figura 16**. Estrutura química dos fármacos antimaláricos com atividade esquistossomicida.

Mais recentemente, diversos estudos *ex vivo* e *in vivo* demonstraram que uma dose única do mefloquina (**8**, Figura 15) causa efeito semelhantes em vermes jovens e adultos de *S. mansoni*, *S. japonicum* e *S. haematobium* (KEISER et al., 2009; XIAO; MEI; JIAO, 2009; INGRAM; ELLIS; KEISER, 2012; XIAO, 2013) e danos morfológicos, histopatológicos e ultra estruturais no tegumento, musculatura, intestino e glândulas vitelinas (fêmeas) (XIAO, 2013). Embora seu mecanismo de ação não seja totalmente elucidado, estudos preliminares sugerem que a mefloquina possa impedir a formação do pigmento hemozoína, a principal via de detoxificação do verme (NCOKAZI; EGAN, 2005; CORRÊA SOARES et al., 2009).

*1.3.1.1 Quimiogenômica*

A quimiogenômica representa um novo campo da descoberta de fármacos que envolve a investigação de bibliotecas moleculares capazes de interagir com alvos biológicos de interesse utilizando bioensaios automatizados com alta vazão de dados (BREDEL; JACOBY, 2004). Com uma visão mais ampla da fase de descoberta, a quimiogenômica visa a identificação de todos os ligantes possíveis para todos os alvos biológicos possíveis (CARON et al., 2001; MESTRES, 2004). Em outras palavras, a análise de dados quimiogenômicos consiste num processo de aprendizagem infinito que visa completar todos os espaços em uma matriz (Figura 16), onde compostos e fármacos são reportados como linhas, alvos biológicos como colunas e os valores são tipicamente potências oriundas da avaliação biológica (ROGNAN, 2007).

Diversos aos avanços tecnológicos auxiliaram esta transformação, tais como: a formação de consórcios para sequenciamento de genomas, desenvolvimento de técnicas de biologia molecular e sistemas de cultura apropriados e disponibilidade de grandes bibliotecas comerciais de compostos (HURLE et al., 2013; LI et al., 2016). Esse aparato tecnológico levou à geração de grandes plataformas de bioensaios com alta vazão de dados, e consequentemente no armazenamento de seus dados em bases de dados de acesso público como *TDR Targets* (AGÜERO et al., 2008; MAGARIÑOS et al., 2012), DrugBank, *Therapeutic Target Database* (TTD) (CHEN; JI; CHEN, 2002; ZHU et al., 2010), ChEMBL (GAULTON et al., 2012), BindingDB (CHEN; LIU; GILSON, 2001) e PubChem (WANG et al., 2012), que atualmente armazenam dados biológicos, genéticos e farmacológicos de milhões de compostos. Na prática, a análise e preenchimento de

todos os espaços de uma matriz de dados quimiogenômicos ainda caracteriza um objetivo longe de ser concluído (KLABUNDE, 2007; ROGNAN, 2007). No entanto, a utilização do conhecimento químico, biológico e genético disponível na literatura representa um grande avanço nesta direção. Tais informações representam as bases moleculares e estruturais necessárias para a construção e validação de modelos computacionais úteis para inferir novas indicações terapêuticas para fármacos conhecidos (CHAVALI et al., 2008; EKINS et al., 2011, 2015).

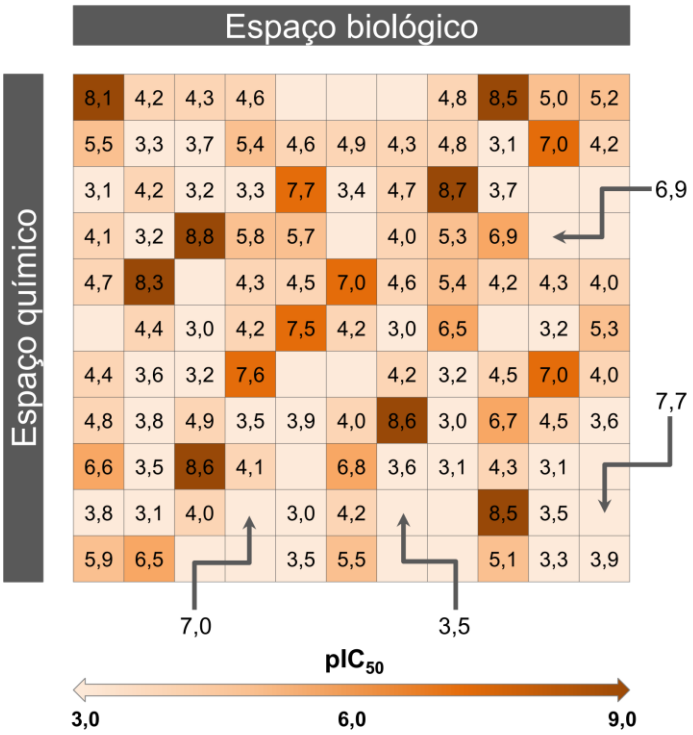



**Figura 17**. Espaços químico e biológico de uma matriz de dados quimiogenômicos.

### *1.3.2.2. Busca por homólogos em bases de dados*

Os recentes avanços no sequenciamento de genomas de esquistossomos (BERRIMAN et al., 2009; ZHOU et al., 2009; PROTASIO et al., 2012; YOUNG et al., 2012) proporcionaram um recurso inestimável para a para a priorização, identificação e estudo da funcionalidade de novos candidatos a alvos terapêuticos de fármacos (BUSCAGLIA; KISSINGER; AGÜERO, 2015) e ao mesmo tempo permitiram o desenvolvimento e aplicação de ferramentas computacionais de quimiogenômica aplicadas a estudos de reposicionamento de fármacos. Tendo em vista este conhecimento, o reposicionamento *in silico* de fármacos para esquistossomose pode ser conduzido através da busca por proteínas do verme que sejam ortólogas a alvos terapêuticos de

fármacos disponíveis nas bases de dados DrugBank (WISHART et al., 2006; KNOX et al., 2011), TTD (ZHU et al., 2010) e STITCH (KUHN et al., 2012).

Levando em consideração a hipótese de que proteínas que compartilham algum tipo de semelhança estrutural (*e.g.* identidade sequencial, sítios ativos e estruturas secundárias semelhantes) também podem compartilhar os mesmos ligantes (KLABUNDE, 2007; ROGNAN, 2007; POLLASTRI; CAMPBELL, 2011), diversos tipos de ferramentas computacionais podem ser aplicadas ao reposicionamento de fármacos para esquistossomose, tais como: alinhamento de proteínas (ALTSCHUL et al., 1997; LI, 2003; WANG et al., 2015), construção de árvores filogenéticas e análise de grafos (KRISTENSEN et al., 2011), análise de domínios e resíduos funcionais (GLASER et al., 2003; ASHKENAZY et al., 2010) e identificação de *chokepoints* em redes metabólicas (SINGH; MALIK; SHARMA, 2007; TAYLOR et al., 2013).

# 2. JUSTIFICATIVA

Tendo em vista o número alarmante de casos de no mundo, o surgimento de linhagens com baixa suscetibilidade ao PZQ bem como sua baixa eficácia contra vermes jovens, a descoberta e reposicionamento de novos fármacos para esquistossomose é eminente.

Graças aos avanços no conhecimento dos mecanismos envolvidos na relação esquistossomo-hospedeiro, diversas enzimas antioxidantes foram identificas. Entre elas, destaca-se a *Sm*TGR, um alvo estratégico envolvido na resposta antioxidante, e também expresso no tegumento dos principais estádios de vida do parasito. Centenas de milhares de compostos com valor de atividade inibitória sobre a *Sm*TGR estão disponíveis na literatura. A modelagem deste conhecimento utilizando ferramentas computacionais de LBDD tais como o QSAR, representa atualmente a principal estratégia para o planejamento e identificação de novos candidatos a fármacos esquistossomicidas. Portanto, para o planejamento de novos candidatos a fármacos esquistossomicidas, consideraram-se os seguintes aspectos:

1. O esquistossomo pode sobreviver por décadas na circulação sanguínea do hospedeiro devido a sua capacidade de neutralizar EROs e ERNs;
2. O sistema antioxidante de esquistossomos é regulado pela *Sm*TGR, um alvo validado que reduz equivalentes redutores necessários para manutenção da atividade catalítica de quase todas as enzimas do sistema antioxidante;
3. Os dados de inibição enzimática de milhares de inibidores de *Sm*TGR estão em bancos de dados públicos;
4. O planejamento de fármacos auxiliado por ferramentas computacionais tais como o QSAR pode ser realizado em curtos espaços de tempo e reduzir os custos operacionais envolvidos nos ensaios experimentais;
5. Embora a estrutura 3D da *Sm*TGR esteja disponível no PDB, a falta de conhecimento sobre o mecanismo reacional dos inibidores disponíveis bem como a presença de três prováveis sítios de interação em sua estrutura, inviabiliza o uso de ferramentas SBDD para VS de novos inibidores;
6. A obtenção de um fármaco inibidor da *Sm*TGR representa uma nova alternativa para o tratamento da esquistossomose com mecanismo de ação inovador.

Por outro lado, os recentes avanços no sequenciamento dos genomas de esquistossomos proporcionaram um recurso inestimável para a para a priorização, identificação e estudo da funcionalidade de novos candidatos a alvos terapêuticos de fármacos e ao mesmo tempo permitiram o desenvolvimento e aplicação de ferramentas computacionais de quimiogenômica aplicadas a estudos de reposicionamento de fármacos. Levando em consideração o fato de que proteínas que compartilham algum tipo de semelhança estrutural também podem compartilhar os mesmos ligantes, o estudo de reposicionamento *in silico* de novos fármacos para esquistossomose justifica-se pelos seguintes aspectos:

1. A versão completa do genoma de *S. mansoni* e a lista de todas as proteínas expressas nos principais estádios de vida do verme estão disponíveis na literatura;
2. Diversas bases de dados armazenam informações de alvos biológicos, além de informações químicas e farmacológicas de fármacos e candidatos a fármacos;
3. O reposicionamento *in silico* de fármacos para esquistossomose pode ser conduzido através da busca por proteínas do verme que tenham homologia com alvos terapêuticos de fármacos disponíveis nas bases de dados DrugBank, TTD e STITCH;
4. Estudos computacionais podem ser úteis para reduzir o custo operacional (*i.e.* número de vermes e animais) e tempo necessários para realização dos ensaios biológicos *ex vivo* em esquistossomos;
5. Os fármacos identificados para ter atividade esquistossomicida também podem servir como *scaffolds* para o planejamento e síntese de análogos estruturais com maior potência, seletividade e permeabilidade no tegumento do verme.

# 3. OBJETIVOS

Face ao exposto, o objetivo geral do presente trabalho foi planejar e identificar novos inibidores da enzima *Sm*TGR com atividade esquistossomicida utilizando ferramentas de quimioinformática e reposicionar novos fármacos para esquistossomose utilizando ferramentas de bioinformática. O trabalho foi dividido nas duas frentes de objetivos específicos, elencados a seguir:

## 3.1. Objetivos específicos

### 3.1.1. Planejar novos candidatos a fármacos esquistossomicidas

- Compilar, integrar e preparar o maior conjunto de dados publicamente disponível de inibidores da enzima *Sm*TGR;
- Desenvolver e validar modelos de QSAR binários para predição da inibição da *Sm*TGR através da combinação de vários descritores e métodos de aprendizado de máquina e seguindo as recomendações da OECD;
- Aplicar os modelos desenvolvidos na VS de novos inibidores da *Sm*TGR;
- Determinar a atividade esquistossomicida dos compostos identificados em esquistossômulos e vermes adultos de *S. mansoni*;
- Determinar a citotoxicidade e reatividade dos compostos em células de mamíferos e cisteíno proteases, respectivamente.

### 3.1.2. Reposicionar fármacos para esquistossomose

- Compilar, integrar e preparar um conjunto de proteínas expressas em esquistossômulos e adultos de *S. mansoni*;
- Realizar uma busca baseada em homologia em várias bases de dados, com o objetivo de identificar fármacos com potencial atividade esquistossomicida;
- Determinar a conservação estrutural das regiões funcionais para as proteínas de *S. mansoni* que apresentam considerável homologia para com alvos terapêuticos de fármacos;
- Priorizar fármacos para a avaliação biológica utilizando informações do espaço químico de compostos ativos e inativos conhecidos;
- Determinar a atividade esquistossomicida dos fármacos preditos em esquistossômulos e vermes adultos de *S. mansoni*.

# 4. ARTIGOS

## 4.1. Produtos da tese

**Artigo 1 –** *Discovery of New Anti-Schistosomal Hits by Integration of QSAR-Based Virtual Screening and High Content Screening*

Bruno J. Neves, Rafael F. Dantas, Mario R. Senger, Cleber C. Melo-Filho, Walter C. G. Valente, Ana C. M. de Almeida, João M. Rezende-Neto, Elid F. C. Lima, Ross Paveley, Nicholas Furnham, Eugene Muratov, Lee Kamentsky, Anne E. Carpenter, Rodolpho C. Braga, Floriano P. Silva-Junior e Carolina Horta Andrade

**Journal of Medicinal Chemistry**, v. 59, n. 15, p. 7075–7088, 2016. DOI: 10.1021/acs.jmedchem.5b02038 (Publicado).

**Artigo 2 –** *In Silico Repositioning-Chemogenomics Strategy Identifies New Drugs with Potential Activity against Multiple Life Stages of Schistosoma mansoni*

Bruno J. Neves, Rodolpho C. Braga, José C. B. Bezerra, Pedro V. L. Cravo e Carolina H. Andrade

**PLoS Neglected Tropical Diseases**, v. 9, n. 1, p. e3435, 2015. DOI: 10.1371/journal.pntd.0003435 (Publicado).

**Artigo 3 –** *The Antidepressant Drug Paroxetine as a New Lead Candidate in Schistosome Drug Discovery*

Bruno Junior Neves, Rafael F. Dantas, Mario R. Senger, Walter C. G. Valente, João M. Rezende-Neto, Willian T. Chaves, Lee Kamentsky, Anne Carpenter, Floriano P. Silva-Junior e Carolina H. Andrade

**Medicinal Chemistry Communcations**, v. 7, n. 6, p. 1176–1182, 2016. DOI: 10.1039/C5MD00596E (Publicado).

## 4.2. Artigos de revisão

**Artigo 4 –** *Modern Approaches to Accelerate Discovery of New Anti-Schistosomal Drugs*

Bruno J. Neves, Eugene Muratov, Renato B. Machado, Carolina H. Andrade e Pedro V. L. Cravo

**Expert Opinion on Drug Discovery**, v. 11, n. 6, p. 557–567, 2016. DOI: 10.1080/17460441.2016.1178230 (Publicado).

**Artigo 5 –** *Natural Products as Leads in Schistosome Drug Discovery*
Bruno J. Neves, Carolina H. Andrade e Pedro V. L. Cravo
**Molecules**, v. 20, n. 2, p. 1872–1903, 2015. DOI: 10.3390/molecules20021872 (Publicado).

**Artigo 6 –** *In Silico Chemogenomics Drug Repositioning Strategies for Neglected Tropical Diseases*
Carolina H. Andrade, Bruno J. Neves, Diego C. Silva, Cleber C. Melo-Filho, Juliana Rodrigues, Rodolpho C. Braga e Pedro V. L. Cravo
**Current Medicinal Chemistry**, 2017 (Submetido).

### 4.3. Artigos em colaborações

Durante o curso de doutorado, foram estabelecidas colaborações com outros grupos de pesquisa, vinculados a Universidade Federal de Goiás e instituições externas, que permitiram a produção dos seguintes artigos:

**Artigo 7 –** *QSAR-Driven Discovery of Novel Chemical Scaffolds Active against Schistosoma mansoni*
Cleber C. Melo-Filho, Rafael F. Dantas, Rodolpho C. Braga, Bruno J. Neves, Mario R. Senger, Walter C. G. Valente, João M. Rezende-Neto, Willian T. Chaves, Eugene N. Muratov, Ross A. Paveley, Nicholas Furnham, Lee Kamentsky, Anne E. Carpenter, Floriano P. Silva-Junior, e Carolina H. Andrade
**Journal of Chemical Information and Modeling**, v. 56, n. 7, p. 1357–1372, 2016. DOI: 10.1021/acs.jcim.6b00055 (Publicado).

**Artigo 8 –** *Analogues of marine guanidine alkaloids are in vitro effective against Trypanosoma cruzi and selectively eliminate Leishmania (L.) infantum intracellular amastigotes*
Ligia F. Martins, Juliana T. Mesquita, Erika G. Pinto, Thais A. Costa-Silva, Samanta E. T. Borborema, Andres J. Galisteo Junior, Bruno J. Neves, Carolina H. Andrade, Zainab Al Shuhaib, Elliot L. Bennett, Gregory P. Black, Philip M. Harper, Daniel M. Evans, Hisham S. Fituri, John P. Leyland, Claire Martin, Terence D. Roberts, Andrew J. Thornhill, Stephen A. Vale, Andrew

Howard-Jones, Dafydd A. Thomas, Harri L. Williams, Larry E. Overman, Roberto G. S. Berlinck, Patrick J. Murphy e Andre G. Tempone

**Journal of Natural Products**, v. 79, n. 9, p. 2202–2210, 2016. DOI: 10.1021/acs.jnatprod.6b00256 (Publicado).

**Artigo 9 –** *Blocking the L-type $Ca^{2+}$ channel (Cav1.2) is the key mechanism for the vascular relaxing effect of Pterodon spp. and its isolated diterpene methyl-6α-acetoxy-7β-hydroxyvouacapan-17β-oate*

Carolina F. Reis, Daniela M. L. Andrade, Bruno J. Neves, Leandra A. R. Oliveira, José F. Pinho, Leidiane P. Silva, Jader S. Cruz, Maria T. F. Bara, Carolina H. Andrade e Matheus L. Rocha

**Pharmacological Research**, v. 100, p. 242–249, 2015. DOI: 10.1016/j.phrs.2015.08.007 (Publicado).

**Artigo 10 –** *Illustrating and homology modeling the proteins of the Zika virus*

Sean Ekins, John Liebler, Bruno J. Neves, Warren G. Lewis, Megan Coffee, Rachelle Bienstock, Christopher Southan e Carolina H. Andrade

**F1000Research**, v. 5, p. 275, 2016. DOI: 10.12688/f1000research.8213.2 (Publicado).

**Artigo 11 –** *Antitrypanosomal activity and evaluation of the mechanism of action of dehydrodieugenol isolated from Nectandra leucantha (Lauraceae) and its methylated derivative against Trypanosoma cruzi*

Simone S. Grecco, Thais A. Costa-Silva, Gerold Jerz, Fernanda S. Sousa, Juliana T. Mesquita, Mariana K. Galuppo, Andre G. Tempone, Bruno J. Neves, Carolina H. Andrade, Rodrigo L. O. R. Cunha, Miriam Uemi, Patricia Sartorelli e João H. G. Lago

**Phytomedicine**, v. 24, p. 62–67, 2017. DOI: 10.1016/j.phymed.2016.11.015 (Publicado)

**Artigo 12 –** *QSAR-driven Design and Discovery of Potent and Selective Chalcone Derivatives with Antitubercular Activity*

Marcelo N. Gomes, Rodolpho C. Braga, Edyta M. Grzelak, Bruno J. Neves, Eugene Muratov, Rui Ma, Larry L. Klein, Sanghyun Cho, Guilherme R. Oliveira, Scott G. Franzblau e Carolina Horta Andrade

**Journal of Antimicrobial Chemotherapy**, 2017 (Submetido)

Journal of Medicinal Chemistry

Article

pubs.acs.org/jmc

# Discovery of New Anti-Schistosomal Hits by Integration of QSAR-Based Virtual Screening and High Content Screening

Bruno J. Neves,[†] Rafael F. Dantas,[‡] Mario R. Senger,[‡] Cleber C. Melo-Filho,[†] Walter C. G. Valente,[‡] Ana C. M. de Almeida,[‡] João M. Rezende-Neto,[‡] Elid F. C. Lima,[‡] Ross Paveley,[§] Nicholas Furnham,[§] Eugene Muratov,[∥] Lee Kamentsky,[⊥] Anne E. Carpenter,[⊥] Rodolpho C. Braga,[†] Floriano P. Silva-Junior,*[,‡] and Carolina Horta Andrade*[,†]

[†]LabMol—Laboratory for Molecular Modeling and Drug Design, Faculdade de Farmácia, Universidade Federal de Goiás, Rua 240, Qd.87, Setor Leste Universitário, Goiânia 74605-510, Brazil

[‡]LaBECFar—Laboratório de Bioquímica Experimental e Computacional de Fármacos, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Av. Brasil, 4365, Rio de Janeiro 21040-900, Rio de Janeiro, Brazil

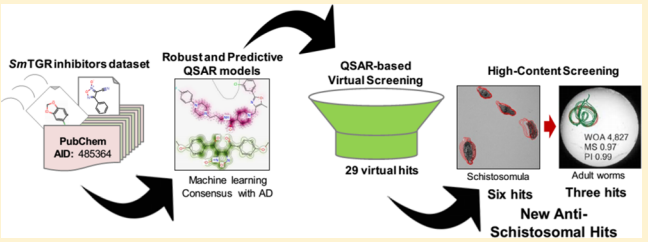
[§]Department of Infection and Immunity, London School of Hygiene and Tropical Medicine, London WC1E 7HT, United Kingdom

[∥]Laboratory for Molecular Modeling, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill North Carolina 27955-7568, United States

[⊥]Imaging Platform, Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts 02142, United States

Supporting Information

**ABSTRACT:** Schistosomiasis is a debilitating neglected tropical disease, caused by flatworms of *Schistosoma* genus. The treatment relies on a single drug, praziquantel (PZQ), making the discovery of new compounds extremely urgent. In this work, we integrated QSAR-based virtual screening (VS) of *Schistosoma mansoni* thioredoxin glutathione reductase (*Sm*TGR) inhibitors and high content screening (HCS) aiming to discover new antischistosomal agents. Initially, binary QSAR models for inhibition of *Sm*TGR were developed and validated using the Organization for Economic Co-operation and Development (OECD) guidance. Using these models, we prioritized 29 compounds for further testing in two HCS platforms based on image analysis of assay plates. Among them, 2-[2-(3-methyl-4-nitro-5-isoxazolyl)vinyl]pyridine and 2-(benzylsulfonyl)-1,3-benzothiazole, two compounds representing new chemical scaffolds have activity against schistosomula and adult worms at low micromolar concentrations and therefore represent promising antischistosomal hits for further hit-to-lead optimization.




## INTRODUCTION

Schistosomiasis is a neglected tropical disease caused by flatworms of the genus *Schistosoma*. These worms cause a chronic and often debilitating infection that impairs development and productivity, and exposure to these worms is strongly linked to extreme poverty.[1−4] Recent estimates of World Health Organization suggest that around 258 million people are infected resulting up to 200000 deaths annually. Currently, schistosomiasis is endemic in 78 countries worldwide, mainly in sub-Saharan Africa, the Middle East, the Caribbean, and South America, where infections are mediated through poor knowledge about the disease, poor sanitation, and lack of effective health policies.[5]

In the absence of a vaccine, the control of schistosomiasis relies on a single drug, praziquantel (PZQ), which has been used in clinical practice for almost four decades.[6] However, because of high incidence of reinfection, the widespread and repeated use of this drug in endemic areas raises concerns about the development of drug resistance by the parasite.[7−11] This problem is further emphasized by the known lack of efficacy of PZQ against juvenile worms,[12] which is a potential cause of treatment failure in endemic areas. Hence, there is an urgent need for new antischistosomal drugs with novel mechanisms of action.

The complete genome sequencing of *Schistosoma mansoni*,[13,14] *Schistosoma japonicum*,[15] and *Schistosoma hematobium*[16] has provided new information on their biological pathways, identifying potentially relevant targets for therapeutic intervention.[17] Thioredoxin glutathione reductase (TGR) is one of these targets; it plays a crucial role in the redox homeostasis of the parasite.[18] TGR is a multifunctional enzyme that acts in the detoxification of reactive oxygen species (ROS) generated by digestion of red blood cells[19,20] and by the host immune

Received: December 31, 2015
Published: July 11, 2016





DOI: 10.1021/acs.jmedchem.5b02038
J. Med. Chem. 2016, 59, 7075−7088

system.[21,22] In mammalian cells, there are two major systems to detoxify ROS, one is based on glutathione (GSH) and the other is based on thioredoxin (Trx). In both systems, NADPH provides reducing equivalents via two specialized oxidoreductase flavoenzymes. Glutathione reductase (GR) reduces glutathione disulfide (GSSG) and drives the GSH-dependent systems, whereas Trx reductases (TR) are pivotal in the Trx-dependent system. On the other hand, in schistosomes, thiol redox homeostasis is completely dependent on TGR, which controls the NADPH reduction of GSSG and Trx in both systems.[23−25] Given these characteristics, it is expected that the maintenance of the homeostatic levels of Trx and GSH in schistosomes play a key role in a variety of cellular processes such as defense against oxidative stress, DNA synthesis, detoxification, protein folding, and repair.[26] Moreover, RNA interference studies have showed that inactivation of TGR of *S. mansoni* (*Sm*TGR)[18] and TGR of *S. japonicum* (*Sj*TGR)[27,28] has profound effects on worm survival rates both in culture medium and infected mice.

Due to the importance of TGR in parasite's redox balance, we hypothesized that known *Sm*TGR inhibitors listed on publicly available databases may serve as the chemical basis to discover new antischistosomal compounds by virtual screening (VS). Docking-based and pharmacophore-based approaches are the most popular VS strategies to identify putative hits in chemical libraries. However, in recent years, quantitative structure−activity relationships (QSAR) models have been used widely in VS applications as well.[29−35]

The main goal of this study was the identification of new structurally dissimilar compounds with high antischistosomal activity. To achieve this goal, we designed a study with the following steps: (i) collection, rigorously curation, and integration of the largest possible data set of *Sm*TGR inhibitors, (ii) development of rigorously validated and mechanistically interpretable binary QSAR models, (iii) application of generated models for VS of three subsets from ChemBridge library (~150000 compounds), (iv) interpretation of developed models to derive structural rules useful for targeted design of new inhibitors, and (v) experimental validation of prioritized/designed hits on live schistosomula and adult worms in two distinct HCS platforms. As a result of this study, we found that the QSAR models were efficient for prediction of new *Sm*TGR inhibitors and identified six novel antischistosomal hit compounds active against schistosomula and three hits active against adult worms. Among them, two hits, 2-[2-(3-methyl-4-nitro-5-isoxazolyl)vinyl]pyridine (**3**) and 2-(benzylsulfonyl)-1,3-benzothiazole (**4**), representing new chemical scaffolds structurally dissimilar to known inhibitors of *S. mansoni*, could be considered as promising antischistosomal agents.

## RESULTS AND DISCUSSION

**Data Set Balancing.** Initially, thousands compounds with *Sm*TGR inhibition data were retrieved from the PubChem Bioassay Database (AID: 485364) and used to build binary QSAR models. Further, uncurated chemical structures were standardized, duplicates were removed, and 2854 compounds with reproducible potency ($IC_{50}$) ≤10 μM were considered as inhibitors, whereas the remaining 337327 compounds were considered as noninhibitors. Because the original data set was highly unbalanced, i.e., 2854 inhibitors and 337327 noninhibitors (1:118 ratio), it is not recommended for building binary QSAR models for the entire data set. During model building, most machine learning methods need equal weighting of the classes in terms of both the number of instances and the level of importance (i.e., active class has the same importance as inactive class). Consequently, when trying to predict a minority class in an unbalanced data set, machine learning methods are prone to assign most samples to the majority class, resulting in a large number of erroneous predictions for minority class.[36]

To reduce the number of the noninhibitors and ideally maintain the "chemical space" of the original data set, we evaluated the optimal number of representative compounds. To accomplish this task, we developed an undersampling workflow based on *k*-nearest neighbors (*k*NN) distances of the each noninhibitor to all inhibitors using the public available 166 substructures MACCS keys. We tested different sizes of the data set by removing noninhibitors and changing the inhibitors-to-noninhibitors ratios of 1:1 (balanced), 1:2, and 1:3.

To visualize the structural diversity of our data set before and after balancing, we performed a principal component analysis (PCA). PCA transforms the original measured variables into new orthogonal variables called principal components, which are a linear combination of the original variables. Detailed results of structural diversity investigation are shown in Supporting Information, Figure S1. The top two principal components retained 20% of the original information. Supporting Information, Figure S1A, represents the PCA plot of 2854 inhibitors (blue dots) vs all 337327 noninhibitors (gray dots). As we can see, the inhibitors are widely distributed across chemical space, reflecting significant chemical diversity. Supporting Information, Figure S1B−D, shows the noninhibitors selected with different ratios: 1:1 or 2854 noninhibitors, Figure S1B; 1:2 or 5705 noninhibitors, Figure S1C; and 1:3 or 8562 noninhibitors, Figure S1D. As we can see from the distribution of these dots, the most representative compounds were chosen that allowed minimal reduction of the original chemical space.

**Performance of Individual QSAR Models.** The balanced (ratio of 1:1) and unbalanced data sets (ratios of 1:2 and 1:3) were modeled by a combination of AtomPair,[37,38] molecular access system (MACCS),[39−41] and Morgan fingerprints,[38,42] chemistry development kit (CDK),[43] and Dragon descriptors[44,45] along with eight machine learning methods leading to 120 different binary QSAR models (Supporting Information, Tables S1, S2, and S3). According to the statistical results of a 5-fold external cross-validation procedure, we could draw three general conclusions: (i) random forest (RF), support vector machine (SVM), and gradient boosting machine (GBM) methods showed the best prediction ability among the eight tested machine learning methods; (ii) QSAR models built on balanced data sets are better than unbalanced (1:2 and 1:3 ratios) due to discrepant values between sensitivity (SE) and specificity (SP), the latter are prone to assign most samples as noninhibitors, resulting in a large number of erroneous predictions; and (iii) the QSAR models which were built from the balanced data set showed a high level of agreement between correct classification rate (CCR), SP, and SE values. Table 1 shows the detailed performances of the more predictive QSAR models derived from the balanced data set.

The combination of Morgan fingerprints with RF (Morgan−RF), MACCS key with RF (MACCS−RF), AtomPair fingerprints with SVM (AtomPair−SVM) and GBM (AtomPair−GBM), Dragon descriptors with SVM (Dragon−SVM) and GBM (Dragon−GBM), and CDK descriptors with SVM (CDK−SVM) led to more predictive QSAR models, with correct classification rate (CCR) ranging between 0.81 and 0.85


DOI: 10.1021/acs.jmedchem.5b02038
J. Med. Chem. 2016, 59, 7075−7088

**Table 1. Summarized Statistical Characteristics of QSAR Models Developed with Balanced Dataset**[a]

| model | CCR | *k* | SE | SP | coverage |
|---|---|---|---|---|---|
| Morgan−RF | 0.85 | 0.71 | 0.85 | 0.86 | 0.62 |
| MACCS−RF | 0.83 | 0.66 | 0.83 | 0.83 | 0.67 |
| AtomPair−SVM | 0.81 | 0.62 | 0.81 | 0.81 | 0.65 |
| AtomPair−GBM | 0.81 | 0.62 | 0.81 | 0.81 | 0.65 |
| Dragon−SVM | 0.85 | 0.70 | 0.85 | 0.84 | 0.69 |
| Dragon−GBM | 0.85 | 0.70 | 0.85 | 0.84 | 0.69 |
| CDK−SVM | 0.84 | 0.69 | 0.85 | 0.84 | 0.77 |
| consensus | 0.87 | 0.74 | 0.87 | 0.88 | 1.00 |
| consensus rigor | 0.91 | 0.81 | 0.96 | 0.87 | 0.38 |

[a]RF, random forest; SVM, support vector machine; GBM, gradient boosting machine; CCR, correct classification rate; *k*, Cohen's κ coefficient; SE, sensitivity; SP, specificity. Consensus and consensus rigor models were built by averaging the predicted values from the individual model for each machine learning technique (Morgan−RF, MACCS−RF, AtomPair−SVM, Dragon−SVM, and CDK−SVM).

and coverage of 0.62−0.77 (Table 1). The best individual model was built using the combination of Morgan−RF (CCR = 0.85, SE = 0.85, and SP = 0.86).

To ensure that the accuracy of the models was not due to chance correlation, 10 rounds of Y-randomization were performed for each data set (Supporting Information, Table S4). The results from this analysis (CCR values around 0.50) indicate that our models built using balanced data set are statistically robust.

**Performance of Consensus Models.** Several individual QSAR models were generated using multiple machine learning algorithms and descriptors/fingerprints. However, our previous experience suggests that consensus models that combine individual QSAR models are advantageous[46−49] and naturally minimize prediction errors during a VS campaign. Therefore, consensus models were built by averaging the predicted values obtained after combining the individual models built using the balanced data set. The detailed performances of 12 consensus models are given in Supporting Information, Table S5. Among them, the consensus model built by combining the Morgan−RF, MACCS−RF, AtomPair−SVM, Dragon−SVM, and CDK−SVM (Table 1 and Supporting Information, Table S5) showed the best performance among all constructed consensus models (CCR = 0.87, SE = 0.87, and SP = 0.88). This consensus model discriminates inhibitors and noninhibitors better than any of the individual QSAR models, with a 2% of increase in CCR, SE, and SP when compared with the best individual model (Morgan−RF).

In addition, the most rigorous consensus model (consensus rigor)[46] was built by combining five individual models with more restrictive conditions. A consensus rigor model only

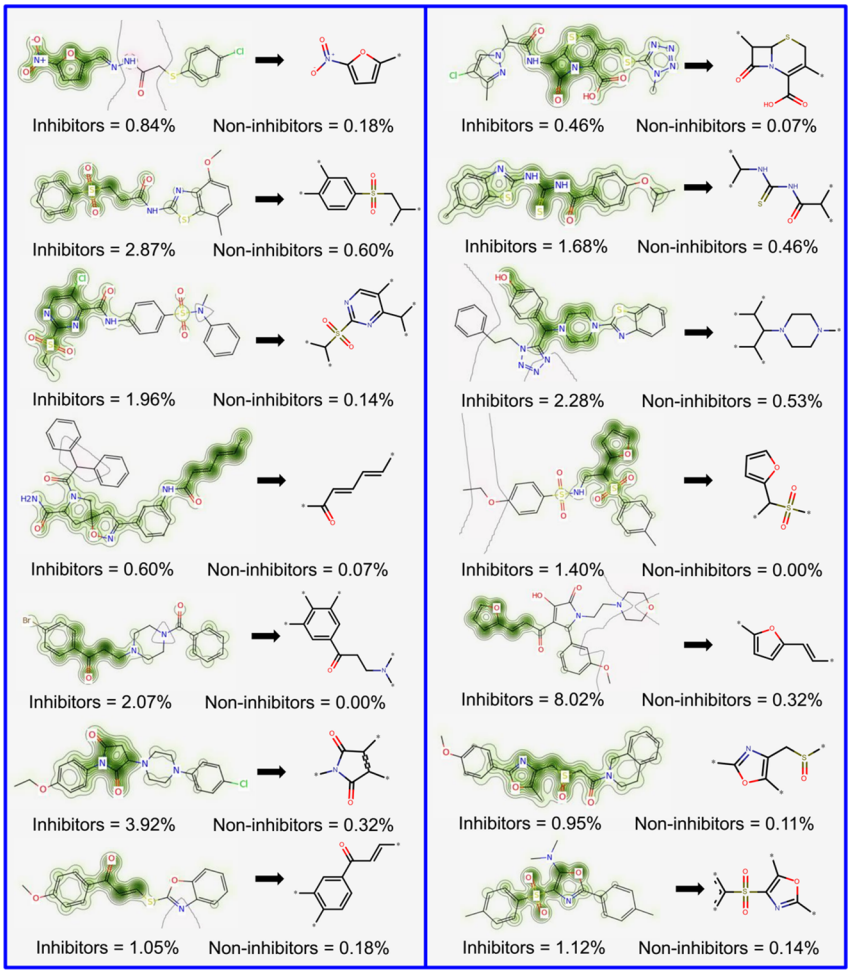



**Figure 1.** Favorable fragments (green) for *Sm*TGR inhibition predicted by the best individual QSAR model and their respective frequencies in inhibitors and noninhibitors sets.



DOI: 10.1021/acs.jmedchem.5b02038
J. Med. Chem. 2016, 59, 7075−7088

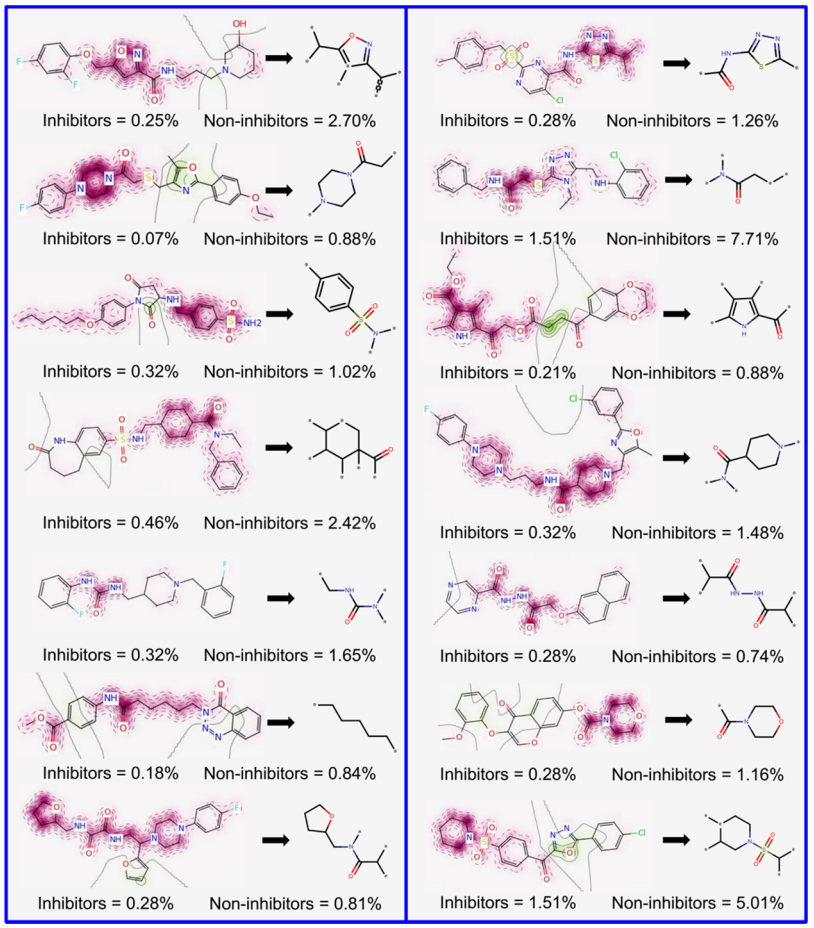



**Figure 2.** Unfavorable fragments (purple) for *Sm*TGR inhibition predicted by the best individual QSAR model and their respective frequencies in inhibitors and noninhibitors sets.

considers the outcome to be reliable when a compound was inside the applicability domain (AD) for the five models. If the compound was outside the AD for any model, then the outcome was specified as unreliable. Expectedly, the combination of Morgan−RF, MACCS−RF, AtomPair−SVM, Dragon−SVM, and CDK−SVM models (Tables 1 and Supporting Information, Table S5) also showed the best performance among all built consensus rigor models (CCR = 0.91, SE = 0.96, and SP = 0.87). In summary, the best consensus rigor model demonstrated better statistical results, with a 5% of increase in CCR, and 11% of increase in SE when compared with the best individual model (Morgan−RF). Although the AD of consensus rigor is limited only for certain chemical classes (coverage of 0.38), it has very high predictive power (CCR = 0.91).

**Model Interpretation.** The Morgan−RF model exhibited the best predictive performance, and, consequently, it possesses the features that are best correlated with *Sm*TGR inhibition activity. Therefore, we translated its features (fingerprints) into predicted probability maps (PPMs) and visualized the atomic and fragment contributions predicted by the QSAR model (Figures 1 and 2). Atoms and fragments promoting the inhibition are highlighted by green (Figure 1), atoms and fragments decreasing the inhibitory potential are highlighted by purple (Figure 2), and gray lines (Figures 1 and 2) delimit the region of split between the favorable and the unfavorable contributions.[50]

Analyzing the fragments with favorable contributions highlighted by PPMs, we noticed that 14 fragments were more frequent in the inhibitors set and absent in the noninhibitors set (Figure 1). Examples of favorable fragments for *Sm*TGR inhibition activity are nitrofuran, 2-ethenylfuran, (ethanesulfonyl) benzene, 2-(sulfonylmethyl) furan, carbonyl thiourea, and 4-methanesulfonyl-1,3-oxazole. By analyzing the fragments with unfavorable contribution into *Sm*TGR inhibition activity (Figure 2), several fragments, such as benzylsulfonamide, methylurea, morpholine-4-carbonyl, piperidine-4-carboxamide, 1-methanesulfonylpiperidine, and cyclohexanecarbonyl, were more frequent in the noninhibitors set. Compounds that contain these fragments may show a decreased *Sm*TGR inhibitory activity. This information could be useful for designing or optimizing new *Sm*TGR inhibitors by replacing unfavorable fragments by favorable fragments.

**Reaction Mechanism of *Sm*TGR Inhibition.** Although the inhibition mechanisms of most of the *Sm*TGR inhibitors are not well understood at the molecular level, the reaction mechanisms by which oxadiazole-2-oxides and cephalosporins operate could be identified according to a graphical interpretation of PPMs. However, for the best understanding of molecular inhibition mechanisms, it is important to highlight that the active site of *Sm*TGR is composed by a cysteine pair (Cys28/Cys31) in the glutharedoxin domain, a cysteine pair (Cys154/Cys159) in the thioredoxin domain, and a redox-active cysteine/selenocysteine pair (Cys596/Sec597) in the C-terminal tail. The latter should be highly mobile to accept


DOI: 10.1021/acs.jmedchem.5b02038
J. Med. Chem. 2016, 59, 7075−7088

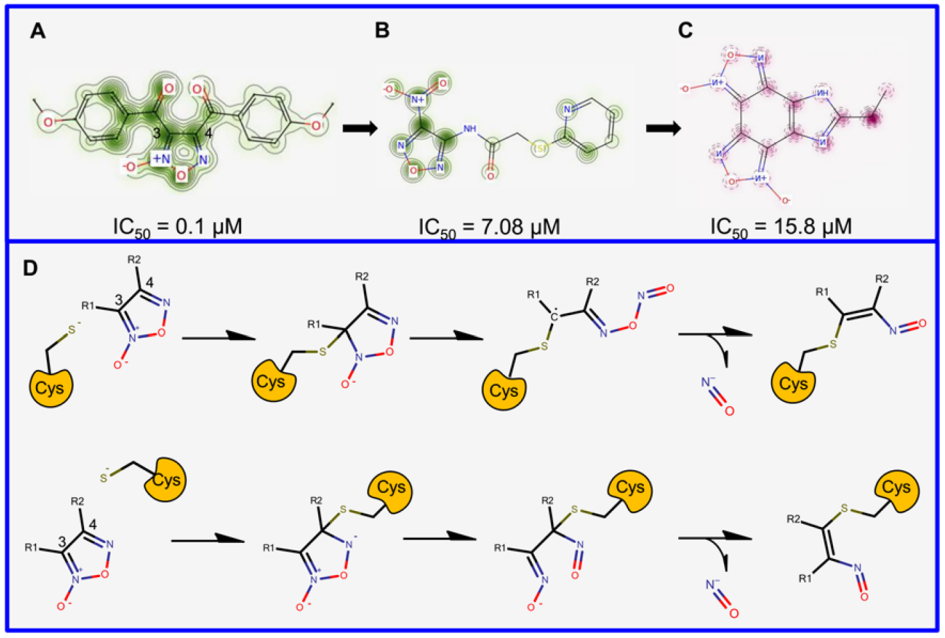



**Figure 3.** Predicted probability maps generated for oxadiazoles (A, B, and C) and their proposed reaction mechanism in the *Sm*TGR active site (D).

electrons from the Cys154/Cys159 pair and to donate electron pairs to Cys28/Cys31 pair.[24] These amino acids provide the perfect chemical environment for covalent inhibition. The higher nucleophilicity and low p$K_a$ of the selenol group of Sec are thought to confer Sec a catalytic advantage over Cys at the attacking position.[51−53] Nonetheless, the thioredoxin domain contains His571 and Glu576, a catalytic dyad that can facilitate proton abstraction of Cys159, thus impacting the catalytic efficiency of the thioredoxin domain of *Sm*TGR.[24]

We observed that the carbons 3 and 4 of the oxadiazole-2-oxide core presented the most important contributions for *Sm*TGR inhibition activity (Figure 3A−C). With PPMs information for this chemotype, a mechanistic rationale for inhibition was initiated through nucleophilic attack (presumably by a thiolate or selenoate of Cys or Sec, respectively) at either the position 3 or 4 of the oxadiazole ring and subsequent rearrangement of the heterocycle in a manner that allows release of the nitroxyl anion. An enzymatic oxidation is posited to transform this agent to nitric oxide (Figure 3D). These pieces of information corroborate with mechanism of inhibition proposed by Rai and colleagues[54] and mechanism of nitric oxide release in physiological solution under the action of thiols studied by Gasco and colleagues[55] In addition, PPMs indicated that the presence of amine-oxide group in core and electron-withdrawing substituents, such as carbonyl, at R1 and R2 positions are favorable for *Sm*TGR inhibition (Figure 3A), while removal of the amine-oxide group (Figure 3B) and presence of electron deficient substituents at R1 and R2 positions (Figure 3C) leads to modest potencies in terms of *Sm*TGR inhibition. These pieces of information corroborate with structure−activity relationships rules established by Rai and colleagues.[54]

The reaction mechanism by which cephalosporins exert their *Sm*TGR inhibition activity was also proposed using the PPMs information (Supporting Information, Figure S2A). For both compounds, the PPMs picked up the positive contributions of the basic core structure of cephalosporines, more specifically carbon 8 and nitrogen 5 of $\beta$-lactam ring, and partially positive contribution of 1-methyl-5-tetrazolethione core for inhibition of *Sm*TGR. On the basis of these results, we suggest that inhibition of *Sm*TGR by cephalosporins may occur via a mechanism similar to proposed by Triboulet and colleagues,[56] i.e., a nucleophilic attack of Cys or Sec on $\beta$-lactam carbonyl carbon, with formation of a tetrahedral intermediate, which collapses with $\beta$-lactam ring opening by N5−C8 bond fission. Then, the acyl-enzyme intermediate could hydrolyze or react further, with expulsion of the 1-methyl-5-tetrazolethione from carbon 3 generating a reactive methylene that could be trapped by other thiolate or selenoate (Supporting Information, Figure S2B).

**QSAR-Based Virtual Screening.** The QSAR-based VS was carried out following the workflow presented in Figure 4. Initially, 150000 compounds available on PremiumSet, DIVERSet-CL, and DIVERSet-EXP libraries of ChemBridge were downloaded and prepared for VS. As drug-like ligands are

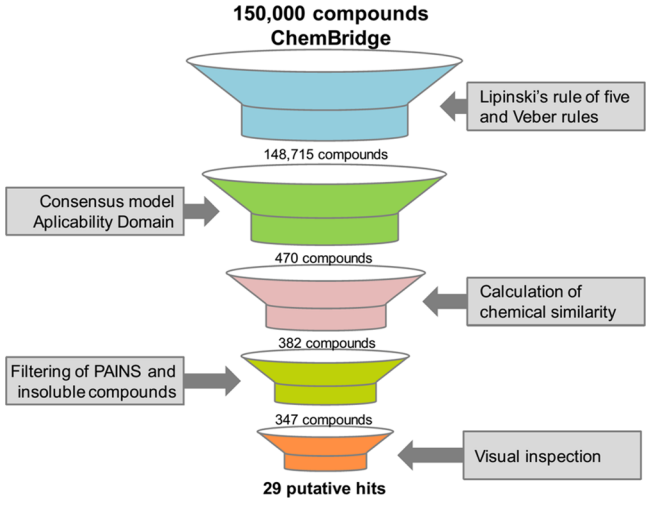



**Figure 4.** QSAR-based VS workflow used for identifying new compounds active against *S. mansoni*.



DOI: 10.1021/acs.jmedchem.5b02038
J. Med. Chem. 2016, 59, 7075−7088

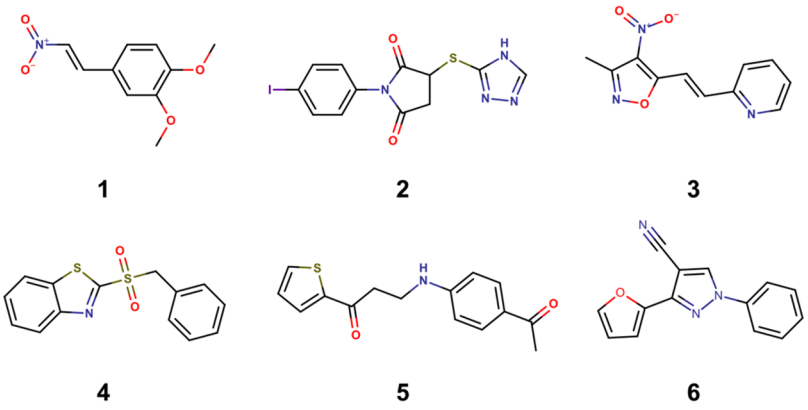



**Figure 5.** Chemical structures of six priority hits selected for further follow up.

highly desirable for the development of new leads with good oral bioavailability, we first filtered these libraries and excluded 1285 compounds that violated Veber[57] and Lipinski's rules.[58] The remaining compounds were predicted by the consensus and consensus rigor models. To narrow down the compounds list and to obtain the highest level of confidence for each prediction, we took both the consensus score (average class prediction) and consensus model coverage into consideration. Consensus model coverage was defined as a fraction of individual models for which a compound was found to fall within the respective ADs. In that sense, introducing probability cutoffs can lead to predictions with higher confidence. Therefore, only putative hits with an average class number prediction of 1.0 and consensus model coverage over 50% were selected (470 putative hits). In addition, we removed compounds with previous bioactivity data reported against *Sm*TGR or *S. mansoni* and pan-assay interference compounds (PAINs)[59,60] so that selected compounds would be novel *Sm*TGR inhibitors and contain no PAINs structures. Finally, the compounds were evaluated by predicting a panel of properties including high aqueous solubility (CIQPlogS),[61] acceptable binding to human serum albumin (QPlogKhsa),[61] acceptable brain/blood partition coefficient (QPlogBB),[61] nonblocking or weak blocking of hERG channel,[46,47] and absence of carcinogenicity and hepatotoxicity.[32] At the end of the VS workflow, 29 putative hits were visually inspected and acquired for biological evaluation (Supporting Information, Table S6).

**Ex Vivo Activity Against Schistosomula.** Compared to target-based VS approaches, the traditional whole-organism schistosome screening approach (phenotypic screening) is an old but indispensable method to discover new antischistosomal agents. This phenotypic approach may be used to validate if the predicted *Sm*TGR-inhibitor interaction has antischistosomal activity. Moreover, a validated compound from a phenotypic assay must have been able to reach its target within the assayed organism only after crossing several biological membranes and resisting degradation by detoxification enzymes. Hence, a hit coming from a phenotypic screen has much more biological value than one coming from a simple biochemical assay. Advances in automated microscopes, liquid handling systems, and computer-based image analysis programs have enabled the development of high-throughput phenotypic assays with cells or small whole organisms, a technique known as high-content screening (HCS).[62,63] HCS microscopes are able to capture high resolution images of live organisms in quick succession, a feature that has been explored to evaluate phenotypic and motility changes in schistosomula[64] or adult worms.[65,66]

Therefore, we employed a HCS assay to evaluate the biological activity of the selected compounds from virtual screening against the *S. mansoni* schistosomula. Assaying against this larval stage is commonly used as an initial screening step in antischistosomal drug discovery campaigns[67−72] because schistosomula are easier to obtain in larger numbers than adult worms. Of the 29 compounds tested against schistosomula, six were declared confirmed actives based on motility and phenotype scores at 20 $\mu$M after 48 h of exposure (Supporting Information, Table S6). The chemical structures of the six primary hits are shown in Figure 5.

Following the initial screening on schistosomula, the six primary hits were selected for determining half-maximal motility concentration ($EC_{50}$) at 0.31−20 $\mu$M range (Table 2 and Supporting Information, Figure S3). Among primary hits, 1,2-dimethoxy-4-(2-nitrovinyl)benzene (**1**), 1-(4-iodophenyl)-3-(4*H*-1,2,4-triazol-3-ylthio)-2,5-pyrrolidinedione (**2**), 3-[(4-acetylphenyl)amino]-1-(2-thienyl)-1-propanone (**5**), and 3-(2-furyl)-1-phenyl-1*H*-pyrazole-4-carbonitrile (**6**) only showed inhibition activity at the highest tested concentration (>20 $\mu$M). On the other hand, 2-[2-(3-methyl-4-nitro-5-isoxazolyl)-vinyl]pyridine (**3**) and 2-(benzylsulfonyl)-1,3-benzothiazole (**4**) showed efficacy in the same range of activity of the reference drug PZQ ($EC_{50}$ = 1.90 $\mu$M), with $EC_{50}$ values of 3.23 and 2.62 $\mu$M, respectively. This is an important feature for a new antischistosomal drug because modern lead discovery pipelines prioritize compounds that possess bioactivity across

**Table 2. Biological Activity Data for Hits of Interest**

| compd | schistosomula $EC_{50}$ ($\mu$M) | adult $EC_{50}$ ($\mu$M) male | adult $EC_{50}$ ($\mu$M) female | WSS-1 $CC_{50}$ ($\mu$M) | papain $IC_{50}$ ($\mu$M) |
|---|---|---|---|---|---|
| **1** | >20 | 29.8 | 5.77 | 17.48 | >100 |
| **2** | >20 | 10.2 | 17.9 | 133.40 | >100 |
| **3** | 3.23 | 6.43 | 5.68 | 16.38 | >100 |
| **4** | 2.62 | 21.1 | 4.91 | 28.49 | >100 |
| **5** | >20 | ND[a] | ND | ND | ND |
| **6** | >20 | ND | ND | ND | ND |
| PZQ | 1.90 | 0.22[b] | 0.64 | >400 | ND |

[a]ND: not determined. [b]$EC_{50}$ values produced for adult male after 72 h of exposure. WSS-1 human kidney epithelial cells were used to evaluate cytotoxicity.



DOI: 10.1021/acs.jmedchem.5b02038
J. Med. Chem. 2016, 59, 7075−7088

the entire developmental cycle of the parasite in the mammalian host.[73,74]

**Analysis of Phenotypic Profile.** Compounds **3** and **4** promoted the internal disruption of larvae as evidenced by the appearance of multiple vacuoles as well as the rounding and darkening of the schistosomula (Figure 6). To evaluate if schistosomula response profiles toward hits resemble those observed in the presence of known antischistosomal drugs (OLT, PZQ, dihydro-artemisinin, methyl-clonazepam, Ro15-5458, and oxamniquine), we applied a Bayesian treatment class model using phenotype scores.[64] This analysis indicated a shared target and/or mechanism of action between OLT and hits, and therefore, all six hits were classified as OLT-like compounds. At least in part, these results could be related to *Sm*TGR inhibition because OLT has already been identified as a noncompetitive inhibitor of this enzyme. It is also important to note that these phenotypic profile has been also observed after *Sm*TGR gene knockout.[18]

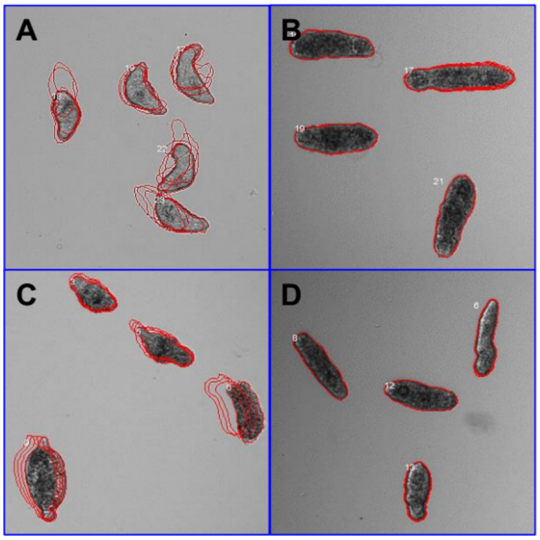



**Figure 6.** Phenotypes of schistosomula exposed for 48 h to 0.625% DMSO (control, A), 20 μM of **4** (B), and 10 μM of PZQ (C) and OLT (D). The outlines represent the position of each parasite over 5 time points (11 s interval).

**Ex Vivo Activity on Adult Worms.** Our next step was to investigate if the compounds identified as hits for schistosomula also had an effect on adult *S. mansoni* worms.[65,66] Therefore, we employed a new HCS platform recently developed by our group that allows for systematic evaluation of gender-, dose-, and time-dependent drug effects on individual male and female parasites by measuring over 100 image features related to worm motility and morphology. Previously, we have demonstrated the successful application of this platform in identification of potent antischistosomal hit compounds.[65,66] In this study, four compounds (**1**−**4**) were screened at 0.1−100 μM concentrations for incubation times varying from 0 h (immediately after compound addition to culture medium) to 72 h.

Inspection of the measured features suggested that at least three features were able to distinguish active from inactive compound concentrations or the DMSO control: the Overlap_RandIndex feature, which is related to motility, the intensity, and the area of the identified worm object. Figure 7 shows a 3D plot of these relevant features for individual female worms exposed to the investigated compounds at 20 μM concentration as well as for the PZQ and negative control (treated with 0.1% DMSO) after 48 h incubation. The sample images are shown to exemplify the phenotypes that can be captured by these features. In general, the feature most correlated to the antischistosomal activity of these compounds was the Overlap_RandIndex, which roughly measures the difference in worm position from one time-lapse frame to the next and is inversely proportional to worm motility in a scale varying from 0 to 1. For simplicity, we hereafter refer to this feature as the "motility score".

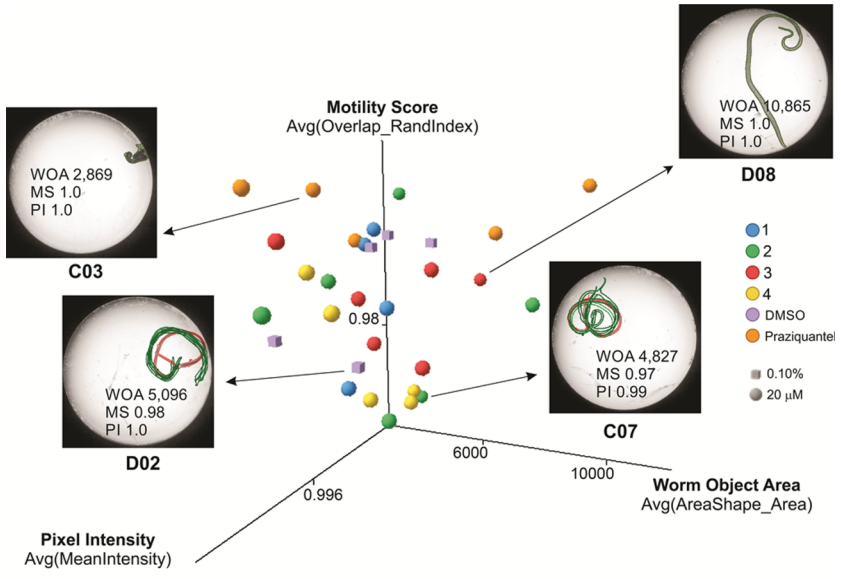



**Figure 7.** 3D scatter plot of the top three image features correlated to antischistosomal activity of the investigated compounds on female *S. mansoni* worms after 48 h drug exposure. Each point in the graph represents a well/condition in the assay. Sample images are shown for selected wells to illustrate the different phenotypes captured by these three parameters (OA, object area; MS, mobility score; PI, pixel intensity of the worm object). The green outlines represent the position of each parasite over five time points (3 s interval) overlaid on the initial position (red outline).



DOI: 10.1021/acs.jmedchem.5b02038
J. Med. Chem. 2016, 59, 7075−7088

To determine the potency of the hit compounds against adult worms with the reference drug PZQ, we have determined $EC_{50}$ values from dose response curves against male and female worms with varying incubation times (Supporting Information, Table S7 and Figure S7). Compounds showed motility inhibition potencies against adult worms ranging from 4.91 to 35 $\mu$M, depending on incubation time and gender (Table 2). Overall, inhibition was fully achieved after 48 h of incubation (Table 2). Compound **3** was the most active, with $EC_{50}$ around 6.00 $\mu$M for both genders. Compounds **1** and **4** showed satisfactory potencies (<10 $\mu$M) for female worms, with $EC_{50}$ = 5.77 and 4.91 $\mu$M, respectively, but not for male worms. Compound **2** was the less potent, with $EC_{50}$ values of 10.2 and 17.9 for male and female, respectively. Despite the satisfactory potencies displayed, all compounds had a less pronounced effect on adult worms than PZQ at all incubation times ($EC_{50}$ values ≤0.66 $\mu$M, see Supporting Information, Table S2). Results also indicated that female worms and schistosomula are slightly more sensitive to compounds action because they showed $EC_{50}$ values up to 5−8 times lower than those determined in males. In part, this could be due to a gender-specific expression pattern of *Sm*TGR and immature antioxidant system of the schistosomula. In fact, schistosomula express lower levels of *Sm*TGR than adults, which make them more susceptible to oxidative damage caused by inhibitors.[75,76]

**Cytotoxicity Against Human Cells.** Compounds **1**−**4** and PZQ were further evaluated for its cytotoxicity against human epithelial cells (WSS-1) from human kidney using a resazurin-based viability assay (Table 2). PZQ showed the lowest cytotoxicity, exhibiting half-maximal cytotoxic concentration ($CC_{50}$) above 400 $\mu$M. Compounds **2**, **3**, and **4** only were cytotoxic in concentrations higher than those necessary for antischistosomal activity. Compound **2** was the least cytotoxic compound ($CC_{50}$ = 133.40 $\mu$M), followed by **4** ($CC_{50}$ = 28.49 $\mu$M), **1** ($CC_{50}$ = 17.48 $\mu$M), and **3** ($CC_{50}$ = 16.38 $\mu$M).

**Controls for Nonspecific Inhibition and Off-Target Effects.** Colloidal aggregates have long plagued early drug discovery. When a colloid is formed, membrane and soluble proteins adsorb to its surface and are partially denatured, leading to nonspecific inhibition and occasionally activation.[77,78] Therefore, adult worms were coincubated with investigated compounds (at 20 and 100 $\mu$M) and detergent Triton X-100 (0.01%) and their antischistosomal effect was compared with activities obtained without detergent for excluding a possible promiscuous colloidal aggregate effect. No significant differences were observed after comparison of inhibition activities of both groups, showing that antischistosomal activity of the hit compounds is related to specific inhibition (Supporting Information, Figure S5). Further, we also investigated possible off-target effects of the hit compounds toward nucleophilic thiols in a papain inhibition assay. Again, none of the antischistosomal hits showed significant inhibition of papain at 100 $\mu$M while positive control E-64 fully inhibits this enzyme at 20 $\mu$M concentration (Table 2 and Supporting Information, Figure S6).

## CONCLUSIONS

To the best of our knowledge, this is the first study integrating QSAR-based VS and HCS methods to discover new antischistosomal agents. We have developed robust and predictive QSAR models for antischistosomal activity. Developed models were used in the most conservative way, i.e., in consensus fashion with the strictest AD criteria, for VS of three ChemBrigde data sets: DIVERSet-CL, DIVERSet-EXP, and PremiumSet. As a result, 470 putative *Sm*TGR inhibitors were identified. Then, 29 compounds were selected and tested against *S. mansoni* schistosomula using a HCS platform and six of them showed significant inhibition activities at 20 $\mu$M. Among them, compounds **3** and **4** showed inhibitory effect equivalent to PZQ, with $EC_{50}$ values around 2.50 $\mu$M. Both hits were also classified as OLT-like compounds, indicating a shared target with OLT, which has already been identified as an inhibitor of *Sm*TGR.[79] The results of gender-, dose-, and time-dependent inhibitory effect indicated that adult female worms of *S. mansoni* are slightly more sensitive than males to compounds action. Compounds **3** and **4** showed satisfactory potencies for female worms, with $EC_{50}$ values around 6.00 $\mu$M. Both compounds also demonstrated low cytotoxicity to WSS-1 mammalian cells ($CC_{50}$ > 16 $\mu$M) and inhibition of papain only in concentrations >100 $\mu$M. Finally, both compounds represent new chemical scaffolds which are structurally dissimilar to known inhibitors of *S. mansoni* and thus can be considered as new hit compounds for further chemical optimization.

## EXPERIMENTAL SECTION

**Computational.** *Data Set.* The QSAR models were developed according to best practices of predictive QSAR modeling,[80,81] which is fully compliant to Organization for Economic Co-operation and Development (OECD) guidance on development and validation of QSAR models such as (i) a defined end point, (ii) an unambiguous algorithm, (iii) a defined domain of applicability, (iv) appropriate measures of goodness-of-fit, robustness, and predictivity, and (v) mechanistic interpretation.[82] All in silico steps developed in this study were implemented in a publicly available KSAR workflow (http://labmol.farmacia.ufg.br/ksar). The KSAR workflow is tightly integrated with R and KNIME and includes many modules such as the module for preparing the data, PCA, building of QSAR models, and VS.[46,83] We first retrieved 359841 compounds containing half-maximal inhibitory concentration ($IC_{50}$) data for the *Sm*TGR enzyme from the PubChem BioAssay database (AID: 485364). Compounds with inconclusive $IC_{50}$ results were considered experimental errors and thus were not included in this study to avoid noise in model building. A total of 2854 out of these 359841 compounds had reproducible potency ($IC_{50} \leq 10$ $\mu$M) and were considered as inhibitors, whereas the remaining 356987 compounds were considered as noninhibitors.

*Data Set Curation.* Each compound of data set was carefully standardized according to the protocol proposed by Fourches and colleagues.[84,85] Briefly, explicit hydrogens were added and salts were removed, whereas specific chemotypes such as aromatic and nitro groups were normalized using ChemAxon Standardizer (v.6.1.2, ChemAxon, Budapest, Hungary, http://www.chemaxon.com). Polymers, inorganic salts, organometallic compounds, and mixtures were also removed. In addition, 4437 compounds with multiple *Sm*TGR measurements were identified during analyses of duplicates. Further analysis showed high concordance (99.9%) of duplicated records. In addition, 345 compounds with molecular weight greater than 700 Da were removed. In the end, the prepared data set contained 2854 inhibitors and 337327 noninhibitors.

*Molecular Fingerprints and Descriptors.* Three different types of fingerprints were used in this study: the Morgan fingerprint, a RDKit implementation[38] of the extended-connectivity fingerprints,[42] with radius of 2 and bit vector of 1024 bits; the MACCS structural key fingerprints;[39−41] and the AtomPair fingerprints (RDKit implementation[38] of the Carhart's atom pairs)[37] with bit vector of 1024 bits. All the fingerprints were calculated by the open-source cheminformatics toolkit RDKit v.2.4.0.[86] A brief description of Morgan, AtomPair, and MACCS fingerprints is available in Supporting Information.

The Chemistry Development Kit (CDK, v.1.4.19, GNU Lesser General Public License) descriptors and 0−2D descriptors were calculated using the PaDEL-Descriptor program[43] and DRAGON


DOI: 10.1021/acs.jmedchem.5b02038
J. Med. Chem. 2016, 59, 7075−7088

(v.5.5, Talete SRL, Milan, Italy), respectively. The complete list of CDK descriptors and a detailed discussion for DRAGON descriptors can be found elsewhere.[44,45] The descriptors matrix was then normalized and constant/near constant and highly correlated ($r \geq 0.9$) descriptors were removed.

*Data Set Analysis and Undersampling.* Because the original library was highly unbalanced (2854 inhibitors and 337327 noninhibitors), it is not recommended for building binary QSAR models for the entire data set. Thus, we decided to balance the data set. Unlike the traditional undersampling methods which randomly balance the data set, our linear undersampling strategy retains most of the representative structures of the noninhibitors set, thus ensuring as high as possible coverage of original chemical space. The basic principle here is to measure the whole inhibitors matrix represented by the MACCS key fingerprints evaluating the Euclidean distance to the MACCS key fingerprints of each noninhibitor using a *k*NN method,[87] implemented in R software v.3.0.3.[88] Then, the samples on noninhibitors set were linearly extracted over the whole set by using *k*-distances and were used to generate balanced and partially balanced data sets. Finally, we generated three undersampled data sets with inhibitor-to-noninhibitor ratios of 1:1 (2854 inhibitors and 2854 noninhibitors), 1:2 (2854 inhibitors and 5705 noninhibitors), and 1:3 (2854 inhibitors and 8562 noninhibitors).

*Machine Learning Implementation.* The building and optimization of statistically acceptable QSAR models requires a close combination between chemical information (i.e., fingerprints or descriptors) and several machine learning classifiers. For this reason, eight different machine learning classifiers, including the SVM with the radial basis Kernel function,[89] the RF,[90] GBM,[91] and partial least-squares discriminant analysis (PLS-DA)[92] approaches, classification and regression trees (CART),[93] *k*NN with Euclidean distance,[87] multilayer perceptron (MLP),[94] and multivariate adaptive regression splines (MARS)[95] were used. All machine learning classifiers were implemented using the R v.3.0.3.[88] A brief description about the theory of each machine learning method is described in Supporting Information.

*5-Fold External Cross-Validation.* The full data set of compounds with known inhibition activities is randomly divided into five subsets of equal size; then one of these subsets (20% of all compounds) is set aside as an external validation set and the remaining four sets together form the modeling set (80% of the full set). This procedure is repeated five times, allowing each of the five subsets to be used as external validation set. Models are built using the modeling set only, and it is important to emphasize that the compounds in momentary external set (fold) are not employed either to build or select the models.

*Applicability Domain.* The AD for each descriptor or fingerprint type was estimated based on the Euclidean distances among the training set of each QSAR model generated in the external 5-fold cross-validation procedure. The distance of a test compound to its nearest neighbor in the training set was compared to the predefined AD threshold level. If the distance was greater than this threshold level, the prediction was considered to be less trustworthy.[96] In this study, we defined AD as a distance threshold $D_T$ between a compound under prediction and its closest nearest neighbors of the training set. It was calculated as follows:

$$D_T = \bar{y} + Z\sigma \tag{1}$$

Here, $\bar{y}$ is the average Euclidean distance of the $k$ nearest neighbors of each compound within the training set, $\sigma$ is the standard deviation of these Euclidean distances, and $Z$ is an arbitrary parameter to control the significance level. We set the default value of this parameter $Z$ at 0.5. Thus, if the distance of the external compound from all of its nearest neighbors in the training set exceeds this threshold, the prediction is considered unreliable.

*Evaluation of Performance and Robustness.* To access the predictive performance of the binary QSAR models, SE, SP, and CCR were used. These statistic metrics are calculated by the following equations:

$$\mathrm{SE} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{2}$$

$$\mathrm{SP} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{3}$$

$$\mathrm{CCR} = \frac{\mathrm{SE} + \mathrm{SP}}{2} \tag{4}$$

Here, TP and TN represent the number of true positives (correct classifications of inhibitors), and true negatives (correct classifications of noninhibitors), respectively, while FP and FN represent the number of false positives (incorrect classifications of inhibitors) and false negatives (incorrect classifications of noninhibitors), respectively.

In addition to the above model evaluation metrics, Cohen's $\kappa$ ($k$) was used to measure the agreement between model predictions and experimental data.[97] This statistical parameter is calculated by the following equations:

$$\Pr(a) = \frac{\mathrm{TP} + \mathrm{TN}}{N} \tag{5}$$

$$\Pr(e) = [(\mathrm{TP} + \mathrm{FP}) \times (\mathrm{TP} + \mathrm{FN}) + (\mathrm{TN} + \mathrm{FN}) \times (\mathrm{TN} + \mathrm{FP})]/N \tag{6}$$

$$k = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)} \tag{7}$$

Here, $N$ denotes the total number of compounds, $\Pr(a)$ represents the relative observed agreement between the predicted classification of the model and the known classification, and $\Pr(e)$ is the hypothetical probability of chance agreement. In the end, $k$ analysis returns values between −1.0 (no agreement) and 1.0 (complete agreement), but values between 0.6 and 1.0 denote that the model is predictive. Finally, to further ensure that the robustness of the models was not due to chance correlation, 10 rounds of Y-randomization were performed for each constructed model.

*Consensus Modeling.* After the building of QSAR models using all pairwise combinations of different types of chemical descriptors/fingerprints and various machine learning methods, the best models were used for consensus modeling, which can be derived by calculating an average for individual models. In consensus modeling, the final predicted value for each compound is estimated by including an average of the predicted values from the set of QSAR models. Thus, the averaged predicted activity for each compound is in the [0, 1] range. Formally, compounds with the predicted activity higher than 0.5 are classified as inhibitors, and those with the predicted activity lower than 0.5 are classified as noninhibitors. Obviously, the closer the average predicted value is to 1 or 0, the higher the concordance among all models and the higher our confidence is in the classification of compounds as inhibitors or noninhibitors, respectively.

*Mechanistic Interpretation.* To explore favorable or unfavorable structural fragments for *Sm*TGR inhibition, the PPMs were generated to visualize the atomic and fragment contributions predicted by the best QSAR model.[50]

*Virtual Screening.* The purpose of VS is to identify in a library of chemicals a subset of compounds with the desired properties based on computational calculations. Here the DIVERSet-CL, DIVERSet-EXP, and PremiumSet diversity data sets taken from the ChemBridge database were screened to identify inhibitors of *Sm*TGR. Prior to screening, the data sets were curated in the same way as modeling set (see Data Set Curation section) and filtered using the Veber[57] and Lipinski's rules[58] to obtain drug-like compounds. Fingerprints and molecular descriptors were generated for all compounds and normalized (except fingerprints) based on the minimum and maximum values of each descriptor of the modeling set. Then, best consensus and consensus rigor models were used to predict the *Sm*TGR inhibition activity of compounds. The prediction results were accepted only when the compound was found within the applicability domains of more than 50% of all models used in consensus prediction. In addition, to estimate the structural novelty of putative hits, we



DOI: 10.1021/acs.jmedchem.5b02038
J. Med. Chem. 2016, 59, 7075−7088

calculated the pairwise Tanimoto coefficients (using MACCS key fingerprints) between each screened putative hit and compounds in the full data set of *Sm*TGR inhibitors. Then, putative hits with previous bioactivity data against *Sm*TGR or *S. mansoni* were identified and PAINS were removed using a workflow developed by Saubern and colleagues.[98] Finally, hits were imported into Maestro workspace v.9.3 and their aqueous solubility (CIQPlogS), binding to human serum albumin (QPlogKhsa), and brain/blood partition coefficient (QPlogBB) properties were predicted using QikProp v.3.4,[61] and hERG inhibition, carcinogenicity, and hepatotoxicity were predicted using the Pred-hERG server,[46,47,99] admetSAR server,[100,101] and PaDEL-DDPredictor program,[102,103] respectively.

**Experimental.** *Materials.* Investigated compounds were purchased from ChemBridge (San Diego-CA, USA), resuspended in 100% DMSO, and used immediately in the assays. It is important to mention that all chemical structures were confirmed using proton ($^1$H) NMR spectra at 300/400 MHz and liquid chromatography−mass spectrometry (LC-MS) analysis with evaporative light scattering and ultraviolet detectors confirmed a minimum purity of 95% for all compounds (spectra of compounds are listed in Supporting Information). DMEM and M169 media were purchased from Vitrocell Embriolife (Campinas-SP, Brazil). All other reagents were purchased from Sigma-Aldrich (St. Louis-MO, USA).

*Automated ex Vivo Larval S. mansoni HCS Assay.* Cercarie (*S. mansoni*, BH strain) were vortexed at maximum speed for 5 min for tail shedding and transformation into schistosomula by an adapted method from literature.[104,105] Briefly, schistosomula were resuspended in Medium 169, placed in 384-well plates (120 per well), and maintained in an incubator with 5% $CO_2$ overnight before compound addition. The worms were then incubated with investigated compounds and PZQ at 0.31−20 $\mu$M concentrations or DMSO (0.625%). The effect of the compounds on schistosomula motility and phenotypes was assessed at 48 h after compound addition using an automated bright-field ImageXpressMicro HCS microscope (IXM; Molecular Devices, Wokingham, UK). For motility analysis 5 × 11 s interval time-lapse images were collected using a 4× objective. For detailed morphology, a 10× objective was used to collect four adjacent images fields from within a well in order to increase the number of schistosomula for phenotype analysis. Analysis of both the larval phenotype and motility was then carried out in Pipeline Pilot 9 as described by Paveley and colleagues.[64] Phenotype analysis of individual parasites was carried out by a two class Laplacian-modified Bayesian categorization analysis of 80 image descriptors which constituted shape, size, image intensity, and texture statistics and compared to a training set of data comprising 20000 parasites. Motility analysis of individual parasites was also carried out by the average object displacement from the origin point in subsequent 4× image across the time frame series. Both the Bayesian phenotype and motility scores were subsequently adjusted to the control wells (DMSO treated) on each plate.[64]

*Automated ex Vivo Adult S. mansoni HCS Assay.* After 42−49 days post percutaneously infection of infant Swiss mice with 150 ± 10 *S. mansoni* cercariae (BH strain), animals were euthanized, and worms perfused from portal hepatic and mesenteric veins. Male and female parasites were rinsed and individually transferred into 96-well plates with complete DMEM media (i.e., DMEM plus 10% fetal calf serum, 2 mM L-glutamine, 100 $\mu$M/mL penicillin, 100 $\mu$g/mL streptomycin). The plates were maintained overnight at 37 °C in a humidified atmosphere of 5% $CO_2$. Further, worms were then incubated up to 72 h with 0.10−100 $\mu$M of selected compounds and PZQ or negative control DMSO at 0.1%. The effect of the compounds on adult worm motility or phenotype was assessed either immediately 24, 48, or 72 h after compound addition using a newly developed HCS assay. Our method uses 100 time-lapse images captured every 250−300 ms with an automated bright-field microscope using a 2× objective lens (ImageXpress Micro XLS, Molecular Devices, CA). Subsequent quantitative image analysis used a custom-developed pipeline for detecting changes in parasite motility and morphology using the open-source CellProfiler software v. 2.1.2.[106] The pipeline along with its validation will be thoroughly described in a subsequent publication, and the pipeline itself is freely available (www.cellprofiler.org/published_pipelines.shtml). Briefly, our strategy for motility measurement was based on sequential pairwise comparison of the 100 captured time-lapse images. The motility measurement called "AdjustedRandIndex" is calculated by comparing worm objects identified on images captured at times $t_n$ and $t_{n-1}$ with CellProfiler's CalculateImageOverlap module. This measure ranges from 0 to 1, with 1 meaning two objects are perfectly aligned (no movement). In addition to the "Overlap" mobility score, over 100 features related to size, shape, intensity, texture, and granularity are calculated for worm objects identified in the image analysis pipeline and saved in a database. These features are expected to describe different parasite phenotypes in response to drug exposure.

*Cytotoxicity Assay.* WSS-1 [WS-1](ATCCCRL-2029) epithelial cells derived from human kidney were grown in DMEM medium, supplemented with 4.5 g/L glucose, 50 $\mu$g/mL gentamicin, and 10% fetal bovine serum, and seeded into 96-well microplates at 5 × $10^4$ cells/mL. Twenty hours later, cells were exposed to 0.2−400 $\mu$M of PZQ, OLT, and LabMol compounds and kept under a humidified atmosphere (37 °C, 5% $CO_2$) for 48 h. To evaluate the cytotoxic effects of the compounds, the fluorescent viability dye resazurin was added to each well at a final concentration of 0.01 mg/mL 4 h before the end of the incubation. Resorufin fluorescence readings ($\lambda_{ex}$ = 560 nm, $\lambda_{em}$ = 590 nm) were performed immediately and 4 h after resazurin addition in a FlexStation 3 Benchtop multi-mode microplate reader (Molecular Devices, Sunnyvale, CA). The percentage of viable cell was calculated using cells treated only with DMSO (0.2−0.8%) as controls.

*Colloidal Aggregation Assay.* Adult worms were coincubated with compounds (at 20 and 100 $\mu$M) and detergent Triton X-100 (0.01%). The, motility measurements were performed after 48 h and 72 h, and their antischistosomal effect was compared with activities obtained without detergent.

*Papain Inhibition Assay.* Enzymatic assay was performed at 37 °C in 100 mM sodium acetate buffer, pH 3.5. Positive control E-64 and compounds were incubated at 20 and 100 $\mu$M concentrations for 5 min with papain (5 $\mu$g/mL), and the reaction was initiated with the addition of 50 $\mu$M Z-FR-AMC fluorogenic peptide substrate.

*Statistical Analysis.* One-way ANOVA followed by Tukey's multiple comparisons test was performed using GraphPad Prism v.5.00 (GraphPad Software, La Jolla California USA, www.graphpad.com). The $EC_{50}$ and $CC_{50}$ values were determined by four-parameter logist curve function using the same software. $EC_{50}$ values obtained for adult worms were calculated using TIBCO Spotfire software (Boston, MA).

*Ethics Statement.* Animal's maintenance and experiments were carried out in accordance with the Institutional Ethics Committee for Laboratory Animal Use at the Oswaldo Cruz Foundation (CEUA/FIOCRUZ, Brazil; license no. L-044/15).

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.5b02038.

More computational details regarding molecular fingerprints calculation and QSAR model development, as well as additional tables and figures of experimental results (PDF)

Molecular formula strings (CSV)

## ■ AUTHOR INFORMATION

### Corresponding Authors

*For C.H.A.: phone, + 55 62 3209-6451; fax, +55 62 3209-6037; E-mail, carolina@ufg.br.

*For F.P.S.-J.: phone, + 55 21 3865 8248; fax, +55 21 2590 3495; E-mail, floriano@ioc.fiocruz.br.



DOI: 10.1021/acs.jmedchem.5b02038
J. Med. Chem. 2016, 59, 7075−7088

### Author Contributions

B.J.N., R.F.D., and M.R.S. contributed equally.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Brazilian funding agencies, CNPq, CAPES, FAPEG, FAPERJ, and FIOCRUZ for financial support and fellowships. We are grateful to ChemAxon for providing academic license of their program. We also thank the Malacology Laboratory (Dr. Silvana C. Thiengo) from IOC/FIOCRUZ for providing *S. mansoni* cercariae, and the Bioassays and Drug Screening Platform (FIOCRUZ RPT11-I subunit), and NIH R01 GM095672 (Automated image analysis for high-throughput phenotypic screening in *Caenorhabditis elegans*) for technological support. B.J.N. was supported by a fellowship from the Coordination for the Improvement of Higher Education Personnel (CAPES). This work has been funded by the National Counsel of Technological and Scientific Development (CNPq), the State of Goiás Research Foundation (FAPEG), and State of Rio de Janeiro Research Foundation (FAPERJ). C.H.A. and F.P.S.Jr are CNPq productivity fellows. L.E.K. and A.E.C. were supported by a grant from the U.S. National Institutes of Health (GM095672). E.M. also acknowledges NIH (grants GM66940 and GM096967), CNPq (grant 400760/2014−2), and UNC for Junior Faculty Development Award for partial financial support. Authors also thank Molecular Devices for providing access to the HCS equipment. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## ■ ABBREVIATIONS USED

AD, applicability domain; CART, classification and regression trees; $CC_{50}$, half-maximal cytotoxic concentration; CDK, chemistry development kit; $EC_{50}$, half-maximal motility concentration; FN, false negatives; FP, false positives; GBM, gradient boosting machine; GR, glutathione reductase; GSH, glutathione; GSSG, glutathione disulfide; HCS, high content screening; $IC_{50}$, half-maximal inhibitory concentration; *k*NN, *k*-nearest neighbors; MACCS, Molecular ACCess System (MACCS) keys; MARS, multivariate adaptive regression splines; MLP, multilayer perceptron; NADPH, nicotinamide adenine dinucleotide phosphate; OECD, Organization for Economic Cooperation and Development; OLT, oltipraz; PCA, principal component analysis; PLS-DA, partial least-squares discriminant analysis; PPMs, predicted probability maps; PZQ, praziquantel; QSAR, quantitative structure−activity relationships; RF, random forest; *S. mansoni*, *Schistosoma mansoni*; SAR, structure−activity relationships; SE, sensitivity; SMARTS, SMILES arbitrary target specification; *Sm*TGR, *S. mansoni* TGR; SP, specificity; SVM, support vector machine; TGR, thioredoxin glutathione reductase; TN, true negatives; TP, true positives; TR, thioredoxin reductase; Trx, thioredoxin; and VS, virtual screening

## ■ REFERENCES

(1) Colley, D. G.; Bustinduy, A. L.; Secor, W. E.; King, C. H. Human Schistosomiasis. *Lancet* **2014**, *383* (9936), 2253−2264.

(2) Ross, A. G. P.; Bartley, P. B.; Sleigh, A. C.; Olds, G. R.; Li, Y.; Williams, G. M.; McManus, D. P. Schistosomiasis. *N. Engl. J. Med.* **2002**, *346* (16), 1212−1220.

(3) King, C. H. Toward the Elimination of Schistosomiasis. *N. Engl. J. Med.* **2009**, *360* (2), 106−109.

(4) Gryseels, B.; Polman, K.; Clerinx, J.; Kestens, L. Human Schistosomiasis. *Lancet* **2006**, *368* (9541), 1106−1118.

(5) *Schistosomiasis*; World Health Organization: Geneva, February 2016; http://www.who.int/mediacentre/factsheets/fs115/en (accessed June 13, 2016).

(6) Gönnert, R.; Andrews, P. Praziquantel, a New Board-Spectrum Antischistosomal Agent. *Z. Parasitenkd.* **1977**, *52* (2), 129−150.

(7) Ismail, M.; Metwally, A.; Farghaly, A.; Bruce, J.; Tao, L. F.; Bennett, J. L. Characterization of Isolates of *Schistosoma Mansoni* from Egyptian Villagers That Tolerate High Doses of Praziquantel. *Am. J. Trop. Med. Hyg.* **1996**, *55* (2), 214−218.

(8) Melman, S. D.; Steinauer, M. L.; Cunningham, C.; Kubatko, L. S.; Mwangi, I. N.; Wynn, N. B.; Mutuku, M. W.; Karanja, D. M. S.; Colley, D. G.; Black, C. L.; Secor, W. E.; Mkoji, G. M.; Loker, E. S. Reduced Susceptibility to Praziquantel among Naturally Occurring Kenyan Isolates of Schistosoma Mansoni. *PLoS Neglected Trop. Dis.* **2009**, *3* (8), e504.

(9) Fallon, P. G.; Sturrock, R. F.; Niang, A. C.; Doenhoff, M. J. Short Report: Diminished Susceptibility to Praziquantel in a Senegal Isolate of *Schistosoma Mansoni*. *Am. J. Trop. Med. Hyg.* **1995**, *53* (1), 61−62.

(10) Hagan, P.; Appleton, C. C.; Coles, G. C.; Kusel, J. R.; Tchuem-Tchuenté, L.-A. Schistosomiasis Control: Keep Taking the Tablets. *Trends Parasitol.* **2004**, *20* (2), 92−97.

(11) Loukas, A.; Bethony, J. M. New Drugs for an Ancient Parasite. *Nat. Med.* **2008**, *14* (4), 365−367.

(12) Wang, W.; Wang, L.; Liang, Y. Susceptibility or Resistance of Praziquantel in Human Schistosomiasis: A Review. *Parasitol. Res.* **2012**, *111* (5), 1871−1877.

(13) Berriman, M.; Haas, B. J.; LoVerde, P. T.; Wilson, R. A.; Dillon, G. P.; Cerqueira, G. C.; Mashiyama, S. T.; Al-Lazikani, B.; Andrade, L. F.; Ashton, P. D.; Aslett, M. a; Bartholomeu, D. C.; Blandin, G.; Caffrey, C. R.; Coghlan, A.; Coulson, R.; Day, T. a; Delcher, A.; DeMarco, R.; Djikeng, A.; Eyre, T.; Gamble, J. a; Ghedin, E.; Gu, Y.; Hertz-Fowler, C.; Hirai, H.; Hirai, Y.; Houston, R.; Ivens, A.; Johnston, D. a; Lacerda, D.; Macedo, C. D.; McVeigh, P.; Ning, Z.; Oliveira, G.; Overington, J. P.; Parkhill, J.; Pertea, M.; Pierce, R. J.; Protasio, A. V.; Quail, M. a; Rajandream, M.-A.; Rogers, J.; Sajid, M.; Salzberg, S. L.; Stanke, M.; Tivey, A. R.; White, O.; Williams, D. L.; Wortman, J.; Wu, W.; Zamanian, M.; Zerlotini, A.; Fraser-Liggett, C. M.; Barrell, B. G.; El-Sayed, N. M. The Genome of the Blood Fluke Schistosoma Mansoni. *Nature* **2009**, *460* (7253), 352−358.

(14) Protasio, A. V.; Tsai, I. J.; Babbage, A.; Nichol, S.; Hunt, M.; Aslett, M. a; De Silva, N.; Velarde, G. S.; Anderson, T. J. C.; Clark, R. C.; Davidson, C.; Dillon, G. P.; Holroyd, N. E.; LoVerde, P. T.; Lloyd, C.; McQuillan, J.; Oliveira, G.; Otto, T. D.; Parker-Manuel, S. J.; Quail, M. a; Wilson, R. A.; Zerlotini, A.; Dunne, D. W.; Berriman, M. A Systematically Improved High Quality Genome and Transcriptome of the Human Blood Fluke Schistosoma Mansoni. *PLoS Neglected Trop. Dis.* **2012**, *6* (1), e1455.

(15) Zhou, Y.; Zheng, H.; Chen, Y.; Zhang, L.; Wang, K.; Guo, J.; Huang, Z.; Zhang, B.; Huang, W.; Jin, K.; Dou, T.; Hasegawa, M.; Wang, L.; Zhang, Y.; Zhou, J.; Tao, L.; Cao, Z.; Li, Y.; Vinar, T.; Brejova, B.; Brown, D.; Li, M.; Miller, D. J.; Blair, D.; Zhong, Y.; Chen, Z.; Liu, F.; Hu, W.; Wang, Z.-Q.; Zhang, Q.-H.; Song, H.-D.; Chen, S.; Xu, X.; Xu, B.; Ju, C.; Huang, Y.; Brindley, P. J.; McManus, D. P.; Feng, Z.; Han, Z.-G.; Lu, G.; Ren, S.; Wang, Y.; Gu, W.; Kang, H.; Chen, J.; Chen, X.; Chen, S.; Wang, L.; Yan, J.; Wang, B.; Lv, X.; Jin, L.; Wang, B.; Pu, S.; Zhang, X.; Zhang, W.; Hu, Q.; Zhu, G.; Wang, J.; Yu, J.; Wang, J.; Yang, H.; Ning, Z.; Beriman, M.; Wei, C.-L.; Ruan, Y.; Zhao, G.; Wang, S.; Liu, F.; Zhou, Y.; Wang, Z.-Q.; Lu, G.; Zheng, H.; Brindley, P. J.; McManus, D. P.; Blair, D.; Zhang, Q.; Zhong, Y.; Wang, S.; Han, Z.-G.; Chen, Z.; Wang, S.; Han, Z.-G.; Chen, Z. The Schistosoma Japonicum Genome Reveals Features of Host−parasite Interplay. *Nature* **2009**, *460* (7253), 345−351.

(16) Young, N. D.; Jex, A. R.; Li, B.; Liu, S.; Yang, L.; Xiong, Z.; Li, Y.; Cantacessi, C.; Hall, R. S.; Xu, X.; Chen, F.; Wu, X.; Zerlotini, A.; Oliveira, G.; Hofmann, A.; Zhang, G.; Fang, X.; Kang, Y.; Campbell, B. E.; Loukas, A.; Ranganathan, S.; Rollinson, D.; Rinaldi, G.; Brindley, P. J.; Yang, H.; Wang, J.; Wang, J.; Gasser, R. B. Whole-Genome



DOI: 10.1021/acs.jmedchem.5b02038
*J. Med. Chem.* 2016, 59, 7075−7088

Sequence of Schistosoma Haematobium. *Nat. Genet.* **2012**, *44* (2), 221−225.
(17) Ferreira, L. G.; Oliva, G.; Andricopulo, A. D. Target-Based Molecular Modeling Strategies for Schistosomiasis Drug Discovery. *Future Med. Chem.* **2015**, *7* (6), 753−764.
(18) Kuntz, A. N.; Davioud-Charvet, E.; Sayed, A. A.; Califf, L. L.; Dessolin, J.; Arnér, E. S. J.; Williams, D. L. Thioredoxin Glutathione Reductase from Schistosoma Mansoni: An Essential Parasite Enzyme and a Key Drug Target. *PLoS Med.* **2007**, *4* (6), e206.
(19) TIMMS, A. R.; BUEDING, E. Studies of a Proteolytic Enzyme from Schistosoma Mansoni. *Br. J. Pharmacol. Chemother.* **1959**, *14* (1), 68−73.
(20) Hall, S. L.; Braschi, S.; Truscott, M.; Mathieson, W.; Cesari, I. M.; Wilson, R. A. Insights into Blood Feeding by Schistosomes from a Proteomic Analysis of Worm Vomitus. *Mol. Biochem. Parasitol.* **2011**, *179* (1), 18−29.
(21) LoVerde, P. Do Antioxidants Play a Role in Schistosome Host−Parasite Interactions? *Parasitol. Today* **1998**, *14* (7), 284−289.
(22) Callahan, H. L.; Crouch, R. K.; James, E. R. Helminth Anti-Oxidant Enzymes: A Protective Mechanism against Host Oxidants? *Parasitol. Today* **1988**, *4* (8), 218−225.
(23) Alger, H. M.; Williams, D. L. The Disulfide Redox System of Schistosoma Mansoni and the Importance of a Multifunctional Enzyme, Thioredoxin Glutathione Reductase. *Mol. Biochem. Parasitol.* **2002**, *121* (1), 129−139.
(24) Huang, H.; Day, L.; Cass, C. L.; Ballou, D. P.; Williams, C. H.; Williams, D. L. Investigations of the Catalytic Mechanism of Thioredoxin Glutathione Reductase from Schistosoma Mansoni. *Biochemistry* **2011**, *50* (26), 5870−5882.
(25) Bonilla, M.; Denicola, A.; Marino, S. M.; Gladyshev, V. N.; Salinas, G. Linked Thioredoxin-Glutathione Systems in Platyhelminth Parasites: Alternative Pathways for Glutathione Reduction and Deglutathionylation. *J. Biol. Chem.* **2011**, *286* (7), 4959−4967.
(26) Williams, D. L.; Bonilla, M.; Gladyshev, V. N.; Salinas, G. Thioredoxin Glutathione Reductase-Dependent Redox Networks in Platyhelminth Parasites. *Antioxid. Redox Signaling* **2013**, *19* (7), 735−745.
(27) Song, L.; Li, J.; Xie, S.; Qian, C.; Wang, J.; Zhang, W.; Yin, X.; Hua, Z.; Yu, C. Thioredoxin Glutathione Reductase as a Novel Drug Target: Evidence from Schistosoma Japonicum. *PLoS One* **2012**, *7* (2), e31456.
(28) Han, Y.; Fu, Z.; Hong, Y.; Zhang, M.; Han, H.; Lu, K.; Yang, J.; Li, X.; Lin, J. Inhibitory Effects and Analysis of RNA Interference on Thioredoxin Glutathione Reductase Expression in Schistosoma Japonicum. *J. Parasitol.* **2014**, *100* (4), 463−469.
(29) Zhang, L.; Fourches, D.; Sedykh, A.; Zhu, H.; Golbraikh, A.; Ekins, S.; Clark, J.; Connelly, M. C.; Sigal, M.; Hodges, D.; Guiguemde, A.; Guy, R. K.; Tropsha, A. Discovery of Novel Antimalarial Compounds Enabled by QSAR-Based Virtual Screening. *J. Chem. Inf. Model.* **2013**, *53* (2), 475−492.
(30) Neves, B. J.; Bueno, R. V.; Braga, R. C.; Andrade, C. H. Discovery of New Potential Hits of Plasmodium Falciparum Enoyl-ACP Reductase through Ligand- and Structure-Based Drug Design Approaches. *Bioorg. Med. Chem. Lett.* **2013**, *23* (8), 2436−2441.
(31) Bueno, R. V.; Toledo, N. R.; Neves, B. J.; Braga, R. C.; Andrade, C. H. Structural and Chemical Basis for Enhanced Affinity to a Series of Mycobacterial Thymidine Monophosphate Kinase Inhibitors: Fragment-Based QSAR and QM/MM Docking Studies. *J. Mol. Model.* **2013**, *19* (1), 179−192.
(32) Braga, R. C.; Alves, V. M.; Silva, A. C.; Nascimento, M. N.; Silva, F. C.; Liao, L. M.; Andrade, C. H. Virtual Screening Strategies in Medicinal Chemistry: The State of the Art and Current Challenges. *Curr. Top. Med. Chem.* **2014**, *14* (16), 1899−1912.
(33) Melo-Filho, C. C.; Braga, R. C.; Andrade, C. H. 3D-QSAR Approaches in Drug Design: Perspectives to Generate Reliable CoMFA Models. *Curr. Comput.-Aided Drug Des.* **2014**, *10* (2), 148−159.
(34) Artemenko, A. G.; Muratov, E. N.; Kuz'min, V. E.; Kovdienko, N. A.; Hromov, A. I.; Makarov, V. A.; Riabova, O. B.; Wutzler, P.; Schmidtke, M. Identification of Individual Structural Fragments of N,N′-(bis-5-Nitropyrimidyl)dispirotripiperazine Derivatives for Cytotoxicity and Antiherpetic Activity Allows the Prediction of New Highly Active Compounds. *J. Antimicrob. Chemother.* **2007**, *60* (1), 68−77.
(35) Kuz'min, V. E.; Artemenko, A. G.; Muratov, E. N.; Volineckaya, I. L.; Makarov, V. A.; Riabova, O. B.; Wutzler, P.; Schmidtke, M. Quantitative Structure-Activity Relationship Studies of [(Biphenyloxy)propyl]isoxazole Derivatives. Inhibitors of Human Rhinovirus 2 Replication. *J. Med. Chem.* **2007**, *50* (17), 4205−4213.
(36) Zakharov, A. V.; Peach, M. L.; Sitzmann, M.; Nicklaus, M. C. QSAR Modeling of Imbalanced High-Throughput Screening Data in PubChem. *J. Chem. Inf. Model.* **2014**, *54* (3), 705−712.
(37) Carhart, R. E.; Smith, D. H.; Venkataraghavan, R. Atom Pairs as Molecular Features in Structure-Activity Studies: Definition and Applications. *J. Chem. Inf. Model.* **1985**, *25* (4), 64−73.
(38) Riniker, S.; Landrum, G. a. Open-Source Platform to Benchmark Fingerprints for Ligand-Based Virtual Screening. *J. Cheminf.* **2013**, *5* (1), 26.
(39) Durant, J. L.; Leland, B. A.; Henry, D. R.; Nourse, J. G. Reoptimization of MDL Keys for Use in Drug Discovery. *J. Chem. Inf. Comput. Sci.* **2002**, *42* (6), 1273−1280.
(40) Dill, J. D. .; Hounshell, W. D. .; Marson, S. .; Peacock, S. .; Wipke, W. T. Search and Retrieval Using an Automated Molecular Access System. In *182nd National Meeting of the American Chemical Society, New York*; American Chemical Society: Washington, DC, 1981; pp 23−28.
(41) Anderson, S. Graphical Representation of Molecules and Substructure-Search Queries in MACCStm. *J. Mol. Graphics* **1984**, *2*, 83−90.
(42) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50* (5), 742−754.
(43) Yap, C. W. PaDEL-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints. *J. Comput. Chem.* **2011**, *32* (7), 1466−1474.
(44) *PaDEL-Descriptor*; http://padel.nus.edu.sg/software/padeldescriptor/ (accessed January 5, 2015).
(45) Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*, 1st ed.; Methods and Principles in Medicinal Chemistry; Todeschini, R., Consonni, V., Eds.; Wiley-VCH Verlag GmbH: Weinheim, Germany, 2000.
(46) Braga, R. C.; Alves, V. M.; Silva, M. F. B.; Muratov, E.; Fourches, D.; Tropsha, A.; Andrade, C. H. Tuning HERG out: Antitarget QSAR Models for Drug Development. *Curr. Top. Med. Chem.* **2014**, *14* (11), 1399−1415.
(47) Braga, R. C.; Alves, V. M.; Silva, M. F. B.; Muratov, E.; Fourches, D.; Lião, L. M.; Tropsha, A.; Andrade, C. H. Pred-hERG: A Novel Web-Accessible Computational Tool for Predicting Cardiac Toxicity. *Mol. Inf.* **2015**, *34* (10), 698−701.
(48) Alves, V. M.; Muratov, E.; Fourches, D.; Strickland, J.; Kleinstreuer, N.; Andrade, C. H.; Tropsha, A. Predicting Chemically-Induced Skin Reactions. Part I: QSAR Models of Skin Sensitization and Their Application to Identify Potentially Hazardous Compounds. *Toxicol. Appl. Pharmacol.* **2015**, *284* (2), 262−272.
(49) Alves, V. M.; Muratov, E.; Fourches, D.; Strickland, J.; Kleinstreuer, N.; Andrade, C. H.; Tropsha, A. Predicting Chemically-Induced Skin Reactions. Part II: QSAR Models of Skin Permeability and the Relationships between Skin Permeability and Skin Sensitization. *Toxicol. Appl. Pharmacol.* **2015**, *284* (2), 273−280.
(50) Riniker, S.; Landrum, G. a. Similarity Maps - A Visualization Strategy for Molecular Fingerprints and Machine-Learning Methods. *J. Cheminf.* **2013**, *5* (9), 42.
(51) Bonilla, M.; Denicola, A.; Novoselov, S. V.; Turanov, A. A.; Protasio, A.; Izmendi, D.; Gladyshev, V. N.; Salinas, G. Platyhelminth Mitochondrial and Cytosolic Redox Homeostasis Is Controlled by a Single Thioredoxin Glutathione Reductase and Dependent on Selenium and Glutathione. *J. Biol. Chem.* **2008**, *283* (26), 17898−17907.


DOI: 10.1021/acs.jmedchem.5b02038
*J. Med. Chem.* 2016, 59, 7075−7088

(52) Nauser, T.; Steinmann, D.; Koppenol, W. H. Why Do Proteins Use Selenocysteine instead of Cysteine? *Amino Acids* **2012**, *42* (1), 39−44.
(53) Johansson, L.; Gafvelin, G.; Arnér, E. S. J. Selenocysteine in Proteins-Properties and Biotechnological Use. *Biochim. Biophys. Acta, Gen. Subj.* **2005**, *1726* (1), 1−13.
(54) Rai, G.; Sayed, A. A.; Lea, W. A.; Luecke, H. F.; Chakrapani, H.; Prast-Nielsen, S.; Jadhav, A.; Leister, W.; Shen, M.; Inglese, J.; Austin, C. P.; Keefer, L.; Arnér, E. S. J.; Simeonov, A.; Maloney, D. J.; Williams, D. L.; Thomas, C. J. Structure Mechanism Insights and the Role of Nitric Oxide Donation Guide the Development of Oxadiazole-2-Oxides as Therapeutic Agents against Schistosomiasis. *J. Med. Chem.* **2009**, *52* (20), 6474−6483.
(55) Gasco, A.; Fruttero, R.; Sorba, G.; Di Stilo, A.; Calvino, R. NO Donors: Focus on Furoxan Derivatives. *Pure Appl. Chem.* **2004**, *76* (5), 973−981.
(56) Triboulet, S.; Dubée, V.; Lecoq, L.; Bougault, C.; Mainardi, J.-L.; Rice, L. B.; Ethève-Quelquejeu, M.; Gutmann, L.; Marie, A.; Dubost, L.; Hugonnet, J.-E.; Simorre, J.-P.; Arthur, M. Kinetic Features of L,D-Transpeptidase Inactivation Critical for $\beta$-Lactam Antibacterial Activity. *PLoS One* **2013**, *8* (7), e67831.
(57) Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **2002**, *45* (12), 2615−2623.
(58) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Delivery Rev.* **1997**, *23* (1−3), 3−25.
(59) Baell, J.; Walters, M. A. Chemistry: Chemical Con Artists Foil Drug Discovery. *Nature* **2014**, *513* (7519), 481−483.
(60) Baell, J. B.; Holloway, G. a. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53* (7), 2719−2740.
(61) *QikProp*, version 3.4; Schrödinger Inc.: New York, 2011; http://www.schrodinger.com/.
(62) Zanella, F.; Lorens, J. B.; Link, W. High Content Screening: Seeing Is Believing. *Trends Biotechnol.* **2010**, *28* (5), 237−245.
(63) Neves, B. J.; Muratov, E.; Machado, R. B.; Andrade, C. H.; Cravo, P. V. L Modern Approaches to Accelerate Discovery of New Antischistosomal Drugs. *Expert Opin. Drug Discovery* **2016**, *11* (6), 557−567.
(64) Paveley, R.; Mansour, N. R.; Hallyburton, I.; Bleicher, L. S.; Benn, A. E.; Mikic, I.; Guidi, A.; Gilbert, I. H.; Hopkins, A. L.; Bickle, Q. D. Whole Organism High-Content Screening by Label-Free, Image-Based Bayesian Classification for Parasitic Diseases. *PLoS Neglected Trop. Dis.* **2012**, *6* (7), e1762.
(65) Melo-Filho, C. C.; Dantas, R. F.; Braga, R. C.; Neves, B. J.; Senger, M. R.; Valente, W. C. G.; Rezende-Neto, J. M.; Chaves, W. T.; Muratov, E. N.; Paveley, R. A.; Furnham, N.; Kamentsky, L.; Carpenter, A. E.; Silva-Junior, F. P. S.; Andrade, C. H. QSAR-Driven Discovery of Novel Chemical Scaffolds Active Against Schistosoma Mansoni. *J. Chem. Inf. Model.* **2016** DOI: 10.1021/acs.jcim.6b00055.
(66) Neves, B. J.; Dantas, R. F.; Senger, M. R.; Valente, W. C. G.; Rezende-Neto, J. M.; Chaves, W. T.; Kamentsky, L.; Carpenter, A.; Silva-Junior, F. P.; Andrade, C. H. The Antidepressant Drug Paroxetine as a New Lead Candidate in Schistosome Drug Discovery. *MedChemComm* **2016**, *7*, 1176−1182.
(67) Panic, G.; Flores, D.; Ingram-Sieber, K.; Keiser, J. Fluorescence/luminescence-Based Markers for the Assessment of Schistosoma Mansoni Schistosomula Drug Assays. *Parasites Vectors* **2015**, *8* (8), 624.
(68) Cowan, N.; Keiser, J. Repurposing of Anticancer Drugs: In Vitro and in Vivo Activities against Schistosoma Mansoni. *Parasites Vectors* **2015**, *8*, 417.
(69) Panic, G.; Vargas, M.; Scandale, I.; Keiser, J. Activity Profile of an FDA-Approved Compound Library against Schistosoma Mansoni. *PLoS Neglected Trop. Dis.* **2015**, *9* (7), e0003962.
(70) Ingram-Sieber, K.; Cowan, N.; Panic, G.; Vargas, M.; Mansour, N. R.; Bickle, Q. D.; Wells, T. N. C.; Spangenberg, T.; Keiser, J. Orally Active Antischistosomal Early Leads Identified from the Open Access Malaria Box. *PLoS Neglected Trop. Dis.* **2014**, *8* (1), e2610.
(71) Paveley, R. A.; Bickle, Q. D. Automated Imaging and Other Developments in Whole-Organism Anthelmintic Screening. *Parasite Immunol.* **2013**, *35* (9−10), 302−313.
(72) Abdulla, M.-H.; Ruelas, D. S.; Wolff, B.; Snedecor, J.; Lim, K.-C.; Xu, F.; Renslo, A. R.; Williams, J.; McKerrow, J. H.; Caffrey, C. R. Drug Discovery for Schistosomiasis: Hit and Lead Compounds Identified in a Library of Known Drugs by Medium-Throughput Phenotypic Screening. *PLoS Neglected Trop. Dis.* **2009**, *3* (7), e478.
(73) Mansour, N. R.; Paveley, R.; Gardner, J. M. F.; Bell, A. S.; Parkinson, T.; Bickle, Q. High Throughput Screening Identifies Novel Lead Compounds with Activity against Larval, Juvenile and Adult Schistosoma Mansoni. *PLoS Neglected Trop. Dis.* **2016**, *10* (4), e0004659.
(74) Long, T.; Neitz, R. J.; Beasley, R.; Kalyanaraman, C.; Suzuki, B. M.; Jacobson, M. P.; Dissous, C.; McKerrow, J. H.; Drewry, D. H.; Zuercher, W. J.; Singh, R.; Caffrey, C. R. Structure-Bioactivity Relationship for Benzimidazole Thiophene Inhibitors of Polo-Like Kinase 1 (PLK1), a Potential Drug Target in Schistosoma Mansoni. *PLoS Neglected Trop. Dis.* **2016**, *10* (1), e0004356.
(75) Mkoji, G. M.; Smith, J. M.; Prichard, R. K. Antioxidant Systems in Schistosoma Mansoni: Evidence for Their Role in Protection of the Adult Worms against Oxidant Killing. *Int. J. Parasitol.* **1988**, *18* (5), 667−673.
(76) Nare, B.; Smith, J. M.; Prichard, R. K. Schistosoma Mansoni: Levels of Antioxidants and Resistance to Oxidants Increase during Development. *Exp. Parasitol.* **1990**, *70* (4), 389−397.
(77) Irwin, J. J.; Duan, D.; Torosyan, H.; Doak, A. K.; Ziebart, K. T.; Sterling, T.; Tumanian, G.; Shoichet, B. K. An Aggregation Advisor for Ligand Discovery. *J. Med. Chem.* **2015**, *58* (17), 7076−7087.
(78) Owen, S. C.; Doak, A. K.; Wassam, P.; Shoichet, M. S.; Shoichet, B. K. Colloidal Aggregation Affects the Efficacy of Anticancer Drugs in Cell Culture. *ACS Chem. Biol.* **2012**, *7* (8), 1429−1435.
(79) Kuntz, A. N.; Davioud-Charvet, E.; Sayed, A. A.; Califf, L. L.; Dessolin, J.; Arnér, E. S. J.; Williams, D. L. Thioredoxin Glutathione Reductase from Schistosoma Mansoni: An Essential Parasite Enzyme and a Key Drug Target. *PLoS Med.* **2007**, *4* (6), e206.
(80) Tropsha, A. Best Practices for QSAR Model Development, Validation, and Exploitation. *Mol. Inf.* **2010**, *29* (6−7), 476−488.
(81) Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, I. I.; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; Consonni, V.; Kuz'min, V. E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, *57* (12), 4977−5010.
(82) *OECD Principles for the Validation, For Regulatory Purposes, of (Quantitative) Structure−Activity Relationship Models*; Organisation for Economic Co-operation and Development: Paris, 2004; http://www.oecd.org/chemicalsafety/risk-assessment/37849783.pdf (accessed June 13, 2016).
(83) Neves, B. J.; Braga, R. C.; Bezerra, J. C. B.; Cravo, P. V. L; Andrade, C. H. In Silico Repositioning-Chemogenomics Strategy Identifies New Drugs with Potential Activity against Multiple Life Stages of Schistosoma Mansoni. *PLoS Neglected Trop. Dis.* **2015**, *9* (1), e3435.
(84) Fourches, D.; Muratov, E.; Tropsha, A. Trust, but Verify: On the Importance of Chemical Structure Curation in Cheminformatics and QSAR Modeling Research. *J. Chem. Inf. Model.* **2010**, *50* (7), 1189−1204.
(85) Fourches, D.; Muratov, E.; Tropsha, A. Curation of Chemogenomics Data. *Nat. Chem. Biol.* **2015**, *11* (8), 535.
(86) Landrum, G. *RDKit: Open-Source Cheminformatics Software*; SourceForge, 2014; http://www.rdkit.org/.
(87) Altman, N. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *Am. Stat.* **1992**, *46* (3), 175−185.


DOI: 10.1021/acs.jmedchem.5b02038
*J. Med. Chem.* 2016, 59, 7075−7088

(88) R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2008.
(89) Vapnik, V. *The Nature of Statistical Learning Theory*, 2nd ed.; Springer: New York, 2000.
(90) Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45* (1), 5−32.
(91) Friedman, J. H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29* (5), 1189−1232.
(92) Barker, M.; Rayens, W. Partial Least Squares for Discrimination. *J. Chemom.* **2003**, *17*, 166−173.
(93) Breiman, L.; Friedman, J.; Olshen, R. A.; Charles, J. S. *Classification and Regression Trees*, 1st ed.; Breiman, L., Ed.; Wadsworth & Brooks/Cole Advanced Books & Software: Monterey, CA, 1984.
(94) Rosenblatt, F. *Principles of Neurodynamics; Perceptrons and the Theory of Brain Mechanisms*, 1st ed.; Rosenblatt, F., Ed.; Spartan Books: Washington, DC, 1962.
(95) Friedman, J. H. Multivariate Adaptive Regression Splines. *Ann. Stat.* **1991**, *19* (1), 1−67.
(96) Zhang, S.; Golbraikh, A.; Oloff, S.; Kohn, H.; Tropsha, A. A Novel Automated Lazy Learning QSAR (ALL-QSAR) Approach: Method Development, Applications, and Virtual Screening of Chemical Databases Using Validated ALL-QSAR Models. *J. Chem. Inf. Model.* **2006**, *46* (5), 1984−1995.
(97) Cohen, J. A Coefficient of Agreement of Nominal Scales. *Educ. Psychol. Meas.* **1960**, *20*, 37−46.
(98) Saubern, S.; Guha, R.; Baell, J. B. KNIME Workflow to Assess PAINS Filters in SMARTS Format. Comparison of RDKit and Indigo Cheminformatics Libraries. *Mol. Inf.* **2011**, *30* (10), 847−850.
(99) Alves, V. M.; Braga, R. C.; Silva, M. B.; Muratov, E.; Fourches, D.; Tropsha, A.; Andrade, C. H. Pred-hERG: A Novel Web-Accessible Computational Tool for Predicting Cardiac Toxicity of Drug Candidates. In *Abstracts of Papers, 248th ACS National Meeting & Exposition, San Francisco, CA, United States, August 10−14, 2014*; American Chemical Society: Washington, DC, 2014; CINF-40.
(100) Cheng, F.; Li, W.; Zhou, Y.; Shen, J.; Wu, Z.; Liu, G.; Lee, P. W.; Tang, Y. admetSAR: A Comprehensive Source and Free Tool for Assessment of Chemical ADMET Properties. *J. Chem. Inf. Model.* **2012**, *52* (11), 3099−3105.
(101) Lagunin, A.; Filimonov, D.; Zakharov, A.; Xie, W.; Huang, Y.; Zhu, F.; Shen, T.; Yao, J.; Poroikov, V. Computer-Aided Prediction of Rodent Carcinogenicity by PASS and CISOC-PSCT. *QSAR Comb. Sci.* **2009**, *28* (8), 806−810.
(102) He, Y.; Liew, C. Y.; Sharma, N.; Woo, S. K.; Chau, Y. T.; Yap, C. W. PaDEL-DDPredictor: Open-Source Software for PD-PK-T Prediction. *J. Comput. Chem.* **2013**, *34* (7), 604−610.
(103) Liew, C. Y.; Lim, Y. C.; Yap, C. W. Mixed Learning Algorithms and Features Ensemble in Hepatotoxicity Prediction. *J. Comput.-Aided Mol. Des.* **2011**, *25* (9), 855−871.
(104) Marxer, M.; Ingram, K.; Keiser, J. Development of an in Vitro Drug Screening Assay Using Schistosoma Haematobium Schistosomula. *Parasites Vectors* **2012**, *5* (1), 165.
(105) Mansour, N. R.; Bickle, Q. D. Comparison of Microscopy and Alamar Blue Reduction in a Larval Based Assay for Schistosome Drug Screening. *PLoS Neglected Trop. Dis.* **2010**, *4* (8), e795.
(106) Kamentsky, L.; Jones, T. R.; Fraser, A.; Bray, M.-A.; Logan, D. J.; Madden, K. L.; Ljosa, V.; Rueden, C.; Eliceiri, K. W.; Carpenter, A. E. Improved Structure, Function and Compatibility for CellProfiler: Modular High-Throughput Image Analysis Software. *Bioinformatics* **2011**, *27* (8), 1179−1180.



DOI: 10.1021/acs.jmedchem.5b02038
*J. Med. Chem.* 2016, 59, 7075−7088

# Discovery of New Anti-Schistosomal Hits by Integration of QSAR-Based Virtual Screening and High Content Screening

*Bruno J. Neves,[†] Rafael F. Dantas,[‡] Mario R. Senger,[‡] Cleber C. Melo-Filho,[†] Walter C.G. Valente,[‡] Ana C. M. de Almeida,[‡] João M. Rezende-Neto,[‡] Elid F. C. Lima,[‡] Ross Paveley,[#] Nicholas Furnham,[#] Eugene Muratov,[§] Lee Kamentsky,[Ψ] Anne E. Carpenter,[Ψ] Rodolpho C. Braga,† Floriano P. Silva-Junior,[‡]* Carolina H. Andrade[†]**

[†]LabMol – Laboratory for Molecular Modeling and Drug Design, Faculdade de Farmácia, Universidade Federal de Goiás, Goiânia, Brazil

[‡]LaBECFar – Laboratório de Bioquímica Experimental e Computacional de Fármacos, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

[#]Department of Infection and Immunity, London School of Hygiene and Tropical Medicine, London, United Kingdom

[§]Laboratory for Molecular Modeling, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, USA

[Ψ]Imaging Platform, Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts, USA.

**Table Contents**

**Molecular Fingerprint Details.** In this study, three 2D fingerprint types were used: dictionary-based fingerprints, circular fingerprints, and path-based fingerprints. The public MACCS structural keys are a collection of 166 predefined substructures associated with the SMILES arbitrary target specification (SMARTS) pattern and belonging to the dictionary-based fingerprint class.[1] Morgan is a type of circular fingerprint that encodes circular atom environments up to determined radius from a central atom using Morgan algorithm and feature invariants. Atom features, such as chirality, atom, and bond types were used for generating Morgan fingerprints.[2–4] AtomPair fingerprints, implemented by Carhart and colleagues, are defined as a pair of atoms (AT) or descriptor centers separated by a fixed topological distance: $AT_i$-$AT_j$-$Dist_{ij}$, where $Dist_{ij}$ is the shortest path (the number of bonds) between $AT_i$ and $AT_j$. In addition, AtomPair fingerprints account for the information about element type, the number of bonded non-hydrogen neighbors and the number of $\pi$ electrons.[5,6]

**Machine Learning Details.** *Classification and Regression Trees (CART).* The CART algorithm, introduced by Breiman and colleagues,[7] is a non-parametric decision tree learning method that produces either classification or regression trees, depending on whether the dependent variable is binary or continuous, respectively. In this method, tree is built by dividing the root node, containing all samples, in two child nodes based on a split value for one descriptor/fingerprint of the matrix of dependent variables. Each split of descriptors within a tree is created based on the best partitioning that is possible. Then, child nodes become parents and each new parent node can give rise to two child nodes, etc. Nodes that are not split anymore (e.g. because they are homogeneous) are called terminal nodes or leafs. In general, the building of a CART model contains two steps. First, a big tree is built. Each node is split until pure terminal nodes are found. The tree obtained has a large number of terminal nodes and describes the

training set almost perfectly, but provides a poor predictive ability for new samples. To solve this problem, a last step, known as pruning, is performed. The pruning step consists of cutting away branches of the big tree to find smaller trees with improved predictive ability.

*Random Forest (RF).* The RF[8] algorithm is a tree bagging method that creates a large collection of decorrelated decision trees, and the final prediction is defined by majority voting from an ensemble of decision trees. In each tree, 1/3 of the training set is randomly extracted, while the remaining 2/3 of the training set is used for model building. Then, each tree in the forest is built by CART method,[7] and best split generated, among the randomly investigated descriptor in each node, is chosen. Each tree is grown to the largest possible extent without pruning. Last, the trained forest is then used to predict test set. The predicted classification values are defined by majority voting for one of the classes. The proportion of votes cast for a class may provide an indication of the probability of a label being correctly assigned, or of confidence in a prediction, but this should be considered an informal estimate only.

*Gradient Boosting Machine (GBM).* The GBM method[9] differs from bagging methods through the base learners, here classification or regression trees, are trained and combined sequentially. The principle idea behind this algorithm is generate models by computing a sequence of trees, in which each successive tree is built from the prediction residuals of the preceding tree. A simple (best) partitioning of the data is determined at each step in the boosting tree algorithm, and the deviations of the observed values from the respective residuals for each partition are computed. Given the preceding sequence of trees, the next tree will then be fitted to the residuals in order to find another partition that will further reduce the residual (error) variance for the data.

*Multivariate Adaptive Regression Splines (MARS).* MARS[10] is a multivariate nonparametric regression technique that can be extended to handle classification problems. This operating method is based on divide-and-conquer strategy partitioning the training data sets into separate regions, each of which gets its own classification. This makes MARS particularly suitable for problems with high input dimensions.

*Support Vector Machine (SVM).* SVM[11] is a kernel based approach first developed by Vapnik as a general data modeling methodology, aiming at minimizing structural risk and statistical learning theory. Briefly, SVM maps the data into a high-dimensional hyper plane (e.g. descriptors or fingerprints), using a kernel function that is typically linear, radial, or polynomial. The SVM seeks to find an optimal separation between two classes (e.g. inhibitors and non-inhibitors), such that each in their entirety lie on opposite sides of a separating hyper plane. Thus, SVM minimizes the empirical classification error and maximizes the geometric margin. This margin is defined as the distance from the separating hyper plane to its nearest sample. The hyper plane that defines such margin is called support hyper planes, and the data points that lie on these hyper planes are called support vectors. Thus, SVM is also known as a maximum margin classifier.

*Partial Least Squares – Discriminant Analysis (PLS-DA).* PLS-DA[12] is a linear and parametric method based on the PLS model in which the dependent variable is chosen to represent the class membership (e.g. inhibitors or non-inhibitors). First a classical PLS model is built with a training set. In PLS, the number of variables is reduced using PCA by creating new latent variables which maximize the covariance between the original variables and the response. Using the optimal number of latent variables, to build a linear regression model should provide the best predictive model. Opposite to classical PLS, where the response is quantitative and continuous, the

responses in PLS-DA are qualitative, discrete and coded in a vector with a class member. For an unknown sample, the predicted value obtained with the PLS-DA model is normally distributed around 0 or 1. To determine the limit from which a sample is considered to be in the inhibitors or non-inhibitors class, a threshold 0.5 is determined. When a value above the threshold is predicted, a sample is considered to belong to the inhibitors class, while a value below the threshold indicates that the sample belong to the non-inhibitors class.

*Multi-Layer Perceptron (MLP).* The MLP method[13] is a network of simple neurons called perceptrons. The basic concept of a single perceptron was introduced by Rosenblatt in 1958. The perceptron computes a single output from multiple real-valued inputs by forming a linear combination according to its input weights and then possibly putting the output through some nonlinear activation function. A typical MLP network consists of a set of source nodes forming the input layer (e.g. descriptors or fingerprints), one or more hidden layers of computation nodes, and an output layer related directly to the activity being predicted. The hidden layers and their number can generally vary depending on the problem at hand. To each of the hidden and output neurons, one virtual neuron can be assigned, called bias used as reference to activate or deactivate a neuron. In addition, MLPs are considered feed-forward neural networks since input signal propagates in only one direction, from the input to output layers. Finally, to train a network to predict values for given arguments, an iterative process that has information fed back from the output neurons to neurons in some layer before was performed, to enable further processing and adjustment of weights on the connections. In this study, training of MLP was performed with the back-propagation algorithm.

*k-Nearest Neighbors (kNN).* The *k*NN[14] is a non-parametric method that classifies samples based on a similarity measure. In this method, a sample could be classified by a majority vote of

its neighbors, with the sample investigated being assigned to the class most common amongst its $k$ nearest neighbors measured by a distance function. This distance is usually taken to be the Euclidean distance, though other metrics such as the Jaccard distance could be used. If $k = 3$, then the case is simply assigned to the class of its three nearest neighbors in a feature space. The samples, which in chemical applications are typically compounds, are described as position vectors in the feature space, which is usually of high dimensionality. It is helpful to scale the features so that distances measured in different directions in the space are comparable.

**Table S1.** Statistical characteristics of QSAR models developed with balanced dataset.

| Model | CCR | *k* | SE | SP | Coverage |
|---|---|---|---|---|---|
| Morgan-GBM | 0.84 | 0.68 | 0.85 | 0.84 | 0.62 |
| Morgan-SVM | 0.84 | 0.69 | 0.83 | 0.85 | 0.62 |
| Morgan-RF | 0.85 | 0.71 | 0.85 | 0.86 | 0.62 |
| Morgan-PLS-DA | 0.81 | 0.62 | 0.81 | 0.81 | 0.62 |
| Morgan-*k*NN | 0.78 | 0.55 | 0.90 | 0.66 | 0.62 |
| Morgan-CART | 0.73 | 0.47 | 0.73 | 0.73 | 0.62 |
| Morgan-MARS | 0.76 | 0.54 | 0.77 | 0.77 | 0.62 |
| Morgan-MLP | 0.82 | 0.65 | 0.84 | 0.81 | 0.62 |
| CDK-GBM | 0.84 | 0.67 | 0.84 | 0.83 | 0.77 |
| CDK-SVM | 0.84 | 0.69 | 0.85 | 0.84 | 0.77 |
| CDK-RF | 0.82 | 0.64 | 0.85 | 0.80 | 0.77 |
| CDK-PLS-DA | 0.77 | 0.54 | 0.76 | 0.78 | 0.77 |
| CDK-*k*NN | 0.78 | 0.57 | 0.78 | 0.78 | 0.77 |
| CDK-CART | 0.70 | 0.40 | 0.71 | 0.69 | 0.77 |
| CDK-MARS | 0.77 | 0.53 | 0.76 | 0.77 | 0.77 |
| CDK-MLP | 0.78 | 0.56 | 0.79 | 0.77 | 0.77 |
| Dragon-GBM | 0.85 | 0.70 | 0.85 | 0.84 | 0.69 |
| Dragon-SVM | 0.85 | 0.70 | 0.85 | 0.84 | 0.69 |
| Dragon-RF | 0.84 | 0.69 | 0.85 | 0.83 | 0.69 |
| Dragon-PLS-DA | 0.80 | 0.61 | 0.80 | 0.81 | 0.69 |
| Dragon-*k*NN | 0.80 | 0.59 | 0.83 | 0.76 | 0.69 |
| Dragon-CART | 0.76 | 0.53 | 0.79 | 0.74 | 0.69 |
| Dragon-MARS | 0.80 | 0.60 | 0.81 | 0.80 | 0.69 |
| Dragon-MLP | 0.80 | 0.61 | 0.81 | 0.79 | 0.69 |
| MACCS-GBM | 0.83 | 0.65 | 0.83 | 0.82 | 0.67 |
| MACCS-SVM | 0.83 | 0.65 | 0.82 | 0.83 | 0.67 |
| MACCS-RF | 0.83 | 0.66 | 0.83 | 0.83 | 0.67 |
| MACCS-PLS-DA | 0.76 | 0.52 | 0.76 | 0.76 | 0.67 |
| MACCS-*k*NN | 0.76 | 0.53 | 0.81 | 0.72 | 0.67 |
| MACCS-CART | 0.73 | 0.46 | 0.74 | 0.72 | 0.67 |
| MACCS-MARS | 0.74 | 0.48 | 0.73 | 0.75 | 0.67 |
| MACCS-MLP | 0.78 | 0.57 | 0.79 | 0.78 | 0.67 |
| AtomPair-GBM | 0.81 | 0.62 | 0.81 | 0.81 | 0.65 |
| AtomPair-SVM | 0.81 | 0.62 | 0.81 | 0.81 | 0.65 |
| AtomPair-RF | 0.80 | 0.61 | 0.79 | 0.82 | 0.65 |
| AtomPair-PLS-DA | 0.74 | 0.47 | 0.74 | 0.73 | 0.65 |
| AtomPair-*k*NN | 0.74 | 0.47 | 0.74 | 0.73 | 0.65 |
| AtomPair-CART | 0.69 | 0.37 | 0.72 | 0.65 | 0.65 |
| AtomPair-MARS | 0.70 | 0.41 | 0.70 | 0.70 | 0.65 |
| AtomPair-MLP | 0.76 | 0.52 | 0.77 | 0.75 | 0.65 |

CCR: correct classification rate; *k*: Cohen's kappa; SE: sensitivity; SP: specificity

**Table S2.** Statistical characteristics of QSAR models developed with unbalanced dataset (1:2).

| Model | CCR | *k* | SE | SP | Coverage |
|---|---|---|---|---|---|
| Morgan-GBM | 0.85 | 0.72 | 0.78 | 0.93 | 0.67 |
| Morgan-SVM | 0.86 | 0.73 | 0.77 | 0.94 | 0.67 |
| Morgan-RF | 0.86 | 0.74 | 0.76 | 0.95 | 0.67 |
| Morgan-PLS-DA | 0.81 | 0.64 | 0.71 | 0.91 | 0.67 |
| Morgan-*k*NN | 0.83 | 0.67 | 0.76 | 0.90 | 0.67 |
| Morgan-CART | 0.67 | 0.38 | 0.44 | 0.90 | 0.67 |
| Morgan-MARS | 0.75 | 0.53 | 0.62 | 0.88 | 0.67 |
| Morgan-MLP | 0.84 | 0.68 | 0.77 | 0.90 | 0.67 |
| CDK-GBM | 0.82 | 0.66 | 0.71 | 0.93 | 0.79 |
| CDK-SVM | 0.84 | 0.69 | 0.75 | 0.92 | 0.79 |
| CDK-RF | 0.82 | 0.66 | 0.70 | 0.94 | 0.79 |
| CDK-PLS-DA | 0.73 | 0.49 | 0.57 | 0.89 | 0.79 |
| CDK-*k*NN | 0.79 | 0.60 | 0.69 | 0.89 | 0.79 |
| CDK-CART | 0.70 | 0.41 | 0.52 | 0.87 | 0.79 |
| CDK-MARS | 0.76 | 0.53 | 0.63 | 0.88 | 0.79 |
| CDK-MLP | 0.78 | 0.56 | 0.70 | 0.86 | 0.79 |
| Dragon-GBM | 0.84 | 0.70 | 0.76 | 0.93 | 0.70 |
| Dragon-SVM | 0.85 | 0.72 | 0.78 | 0.93 | 0.70 |
| Dragon-RF | 0.84 | 0.74 | 0.74 | 0.94 | 0.70 |
| Dragon-PLS-DA | 0.78 | 0.57 | 0.65 | 0.90 | 0.70 |
| Dragon-*k*NN | 0.81 | 0.63 | 0.72 | 0.90 | 0.70 |
| Dragon-CART | 0.74 | 0.51 | 0.60 | 0.89 | 0.70 |
| Dragon-MARS | 0.78 | 0.58 | 0.68 | 0.89 | 0.70 |
| Dragon-MLP | 0.80 | 0.61 | 0.73 | 0.88 | 0.70 |
| MACCS-GBM | 0.82 | 0.66 | 0.74 | 0.91 | 0.64 |
| MACCS-SVM | 0.82 | 0.66 | 0.72 | 0.92 | 0.64 |
| MACCS-RF | 0.82 | 0.66 | 0.72 | 0.92 | 0.64 |
| MACCS-PLS-DA | 0.72 | 0.46 | 0.55 | 0.88 | 0.64 |
| MACCS-*k*NN | 0.77 | 0.54 | 0.67 | 0.86 | 0.64 |
| MACCS-CART | 0.70 | 0.42 | 0.53 | 0.87 | 0.64 |
| MACCS-MARS | 0.70 | 0.43 | 0.52 | 0.88 | 0.64 |
| MACCS-MLP | 0.79 | 0.58 | 0.72 | 0.86 | 0.64 |
| AtomPair-GBM | 0.81 | 0.64 | 0.70 | 0.92 | 0.66 |
| AtomPair-SVM | 0.81 | 0.65 | 0.71 | 0.92 | 0.66 |
| AtomPair-RF | 0.79 | 0.62 | 0.64 | 0.94 | 0.66 |
| AtomPair-PLS-DA | 0.72 | 0.47 | 0.58 | 0.87 | 0.66 |
| AtomPair-*k*NN | 0.78 | 0.55 | 0.71 | 0.85 | 0.66 |
| AtomPair-CART | 0.67 | 0.37 | 0.48 | 0.87 | 0.66 |
| AtomPair-MARS | 0.65 | 0.39 | 0.50 | 0.86 | 0.66 |
| AtomPair-MLP | 0.78 | 0.56 | 0.70 | 0.86 | 0.66 |

CCR: correct classification rate; *k*: Cohen's kappa; SE: sensitivity; SP: specificity

**Table S3.** Statistical characteristics of QSAR models developed with unbalanced dataset (1:3).

| Model | CCR | *k* | SE | SP | Coverage |
|---|---|---|---|---|---|
| Morgan-GBM | 0.85 | 0.73 | 0.74 | 0.95 | 0.65 |
| Morgan-SVM | 0.85 | 0.73 | 0.74 | 0.96 | 0.65 |
| Morgan-RF | 0.84 | 0.72 | 0.71 | 0.97 | 0.65 |
| Morgan-PLS-DA | 0.79 | 0.61 | 0.64 | 0.94 | 0.65 |
| Morgan-*k*NN | 0.82 | 0.64 | 0.72 | 0.91 | 0.65 |
| Morgan-CART | 0.72 | 0.49 | 0.51 | 0.93 | 0.65 |
| Morgan-MARS | 0.73 | 0.50 | 0.53 | 0.93 | 0.65 |
| Morgan-MLP | 0.83 | 0.66 | 0.73 | 0.92 | 0.65 |
| CDK-GBM | 0.80 | 0.65 | 0.65 | 0.95 | 0.80 |
| CDK-SVM | 0.83 | 0.69 | 0.71 | 0.95 | 0.80 |
| CDK-RF | 0.78 | 0.64 | 0.60 | 0.97 | 0.80 |
| CDK-PLS-DA | 0.69 | 0.44 | 0.45 | 0.94 | 0.80 |
| CDK-*k*NN | 0.78 | 0.59 | 0.62 | 0.93 | 0.80 |
| CDK-CART | 0.68 | 0.41 | 0.44 | 0.92 | 0.80 |
| CDK-MARS | 0.72 | 0.49 | 0.52 | 0.92 | 0.80 |
| CDK-MLP | 0.77 | 0.55 | 0.65 | 0.90 | 0.80 |
| Dragon-GBM | 0.83 | 0.71 | 0.71 | 0.96 | 0.71 |
| Dragon-SVM | 0.85 | 0.72 | 0.74 | 0.95 | 0.71 |
| Dragon-RF | 0.82 | 0.70 | 0.68 | 0.97 | 0.71 |
| Dragon-PLS-DA | 0.73 | 0.51 | 0.51 | 0.94 | 0.71 |
| Dragon-*k*NN | 0.79 | 0.60 | 0.65 | 0.93 | 0.71 |
| Dragon-CART | 0.69 | 0.43 | 0.43 | 0.94 | 0.71 |
| Dragon-MARS | 0.75 | 0.55 | 0.58 | 0.93 | 0.71 |
| Dragon-MLP | 0.80 | 0.61 | 0.69 | 0.91 | 0.71 |
| MACCS-GBM | 0.81 | 0.66 | 0.68 | 0.95 | 0.69 |
| MACCS-SVM | 0.81 | 0.66 | 0.68 | 0.95 | 0.69 |
| MACCS-RF | 0.81 | 0.67 | 0.67 | 0.96 | 0.69 |
| MACCS-PLS-DA | 0.68 | 0.40 | 0.42 | 0.93 | 0.69 |
| MACCS-*k*NN | 0.76 | 0.55 | 0.60 | 0.92 | 0.69 |
| MACCS-CART | 0.68 | 0.41 | 0.43 | 0.93 | 0.69 |
| MACCS-MARS | 0.66 | 0.37 | 0.40 | 0.93 | 0.69 |
| MACCS-MLP | 0.78 | 0.58 | 0.66 | 0.91 | 0.69 |
| AtomPair-GBM | 0.78 | 0.62 | 0.61 | 0.96 | 0.67 |
| AtomPair-SVM | 0.75 | 0.58 | 0.53 | 0.97 | 0.67 |
| AtomPair-RF | 0.75 | 0.58 | 0.53 | 0.97 | 0.67 |
| AtomPair-PLS-DA | 0.69 | 0.43 | 0.45 | 0.93 | 0.67 |
| AtomPair-*k*NN | 0.77 | 0.57 | 0.61 | 0.92 | 0.67 |
| AtomPair-CART | 0.64 | 0.34 | 0.36 | 0.93 | 0.67 |
| AtomPair-MARS | 0.64 | 0.34 | 0.36 | 0.93 | 0.67 |
| AtomPair-MLP | 0.75 | 0.53 | 0.60 | 0.90 | 0.67 |

CCR: correct classification rate; *k*: Cohen's kappa; SE: sensitivity; SP: specificity

**Table S4.** Statistical characteristics of Y-randomization models developed with balanced dataset.

| Model | CCR (SD) | k (SD) | SE (SD) | SP (SD) |
|---|---|---|---|---|
| Morgan–RF | 0.51 (0.01) | 0.01 (0.02) | 0.50 (0.01) | 0.51 (0.01) |
| MACCS–RF | 0.51 (0.01) | 0.02 (0.02) | 0.51 (0.04) | 0.51 (0.04) |
| AtomPair–SVM | 0.50 (0.01) | 0.01 (0.01) | 0.49 (0.01) | 0.50 (0.01) |
| AtomPair–GBM | 0.51 (0.01) | 0.01 (0.01) | 0.50 (0.02) | 0.51 (0.02) |
| Dragon–SVM | 0.50 (0.01) | 0.01 (0.01) | 0.51 (0.02) | 0.50 (0.02) |
| Dragon–GBM | 0.51 (0.01) | 0.01 (0.02) | 0.51 (0.02) | 0.51 (0.02) |
| CDK–SVM | 0.50 (0.01) | 0.00 (0.01) | 0.51 (0.01) | 0.50 (0.01) |

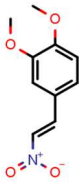
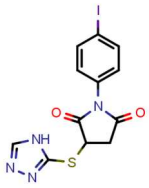
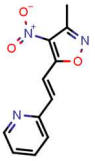
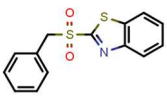
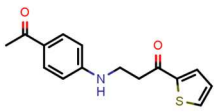
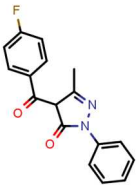
CCR: correct classification rate; *k*: Cohen's kappa; SE: sensitivity; SP: specificity

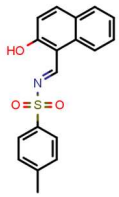
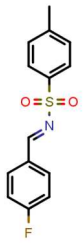
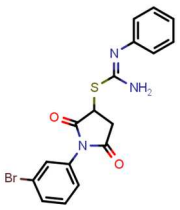
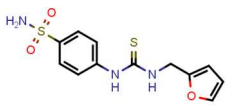
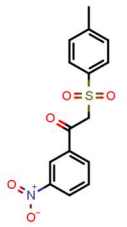
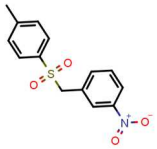
**Table S5.** Statistical characteristics of consensus and consensus rigor models developed.

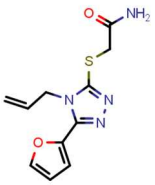
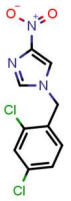
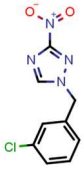
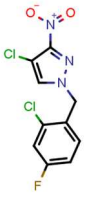
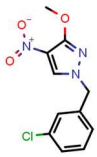
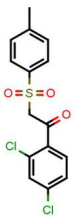
| Model | CCR | k | SE | SP | Coverage | Combination used |
|---|---|---|---|---|---|---|
| Consensus | 0.87 | 0.74 | 0.87 | 0.88 | 1.00 | Morgan-RF + MACCS-RF + AtomPair-SVM + Dragon-SVM + CDK-SVM |
| Consensus rigor | 0.91 | 0.81 | 0.96 | 0.87 | 0.38 | |
| Consensus | 0.87 | 0.73 | 0.86 | 0.87 | 1.00 | Morgan-RF + MACCS-RF + AtomPair-GBM + Dragon-GBM + CDK-GBM |
| Consensus rigor | 0.91 | 0.80 | 0.95 | 0.87 | 0.38 | |
| Consensus | 0.87 | 0.73 | 0.86 | 0.87 | 1.00 | Morgan-RF + MACCS-RF + AtomPair-RF + Dragon-SVM + CDK-SVM |
| Consensus rigor | 0.91 | 0.81 | 0.95 | 0.87 | 0.38 | |
| Consensus | 0.85 | 0.71 | 0.86 | 0.85 | 1.00 | Morgan-RF + MACCS-RF + AtomPair-RF + Dragon-RF + CDK-RF |
| Consensus rigor | 0.89 | 0.77 | 0.95 | 0.84 | 0.38 | |
| Consensus | 0.87 | 0.74 | 0.87 | 0.87 | 1.00 | Morgan-RF + MACCS-RF + AtomPair-SVM + Dragon-GBM + CDK-GBM |
| Consensus rigor | 0.91 | 0.80 | 0.95 | 0.87 | 0.38 | |
| Consensus | 0.87 | 0.74 | 0.87 | 0.88 | 1.00 | Morgan-RF + MACCS-RF + AtomPair-GBM + Dragon-SVM + CDK-SVM |
| Consensus rigor | 0.91 | 0.81 | 0.96 | 0.86 | 0.38 | |
| Consensus | 0.87 | 0.74 | 0.87 | 0.87 | 1.00 | Morgan-RF + MACCS-RF + AtomPair-GBM + Dragon-GBM + CDK-SVM |
| Consensus rigor | 0.91 | 0.81 | 0.95 | 0.87 | 0.38 | |
| Consensus | 0.87 | 0.74 | 0.87 | 0.87 | 1.00 | Morgan-GBM + MACCS-GBM + AtomPair-SVM+ Dragon-SVM + CDK-SVM |
| Consensus rigor | 0.91 | 0.81 | 0.96 | 0.87 | 0.38 | |
| Consensus | 0.87 | 0.74 | 0.87 | 0.87 | 1.00 | Morgan-GBM + MACCS-RF + AtomPair-SVM+ Dragon-SVM + CDK-SVM |
| Consensus rigor | 0.91 | 0.81 | 0.96 | 0.86 | 0.38 | |
| Consensus | 0.87 | 0.74 | 0.87 | 0.87 | 1.00 | Morgan-RF + MACCS-GBM + AtomPair-SVM+ Dragon-SVM + CDK-SVM |
| Consensus rigor | 0.91 | 0.81 | 0.96 | 0.87 | 0.38 | |
| Consensus | 0.87 | 0.73 | 0.86 | 0.87 | 1.00 | Morgan-GBM + MACCS-GBM + AtomPair-GBM+ Dragon-GBM + CDK-GBM |
| Consensus rigor | 0.90 | 0.78 | 0.95 | 0.86 | 0.38 | |
| Consensus | 0.87 | 0.74 | 0.86 | 0.88 | 1.00 | Morgan-SVM + MACCS-SVM + AtomPair-GBM+ Dragon-GBM+ CDK-GBM |
| Consensus rigor | 0.91 | 0.80 | 0.95 | 0.88 | 0.38 | |

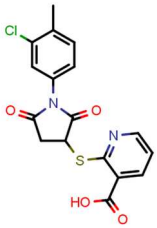
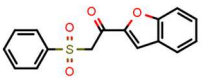
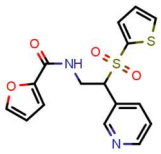
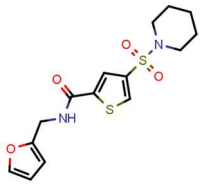
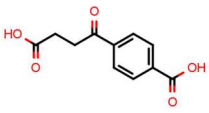
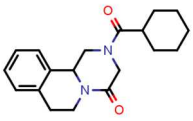
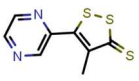
CCR: correct classification rate; *k*: Cohen's kappa; SE: sensitivity; SP: specificity

**Table S6.** The predictions for the PZQ, OLT, and 29 putative hits using more predictive consensus model and its motility and phenotype adjusted index values obtained for *S. mansoni* schistosomula exposed for 48h at a 20 μM concentration. The positive controls PZQ and OLT were tested at 10 μM concentration.

| Compound | Chemical structure | Prob. | AD | Motility adjusted index (mean ± SD) | Phenotype adjusted index (mean ± SD) |
|---|---|---|---|---|---|
| 1,2-dimethoxy-4-(2-nitrovinyl)benzene (LabMol-23, 1) |  | 1.0 | 0.6 | -0.74 ± 0.08 | -0.64 ± 0.01 |
| 1-(4-iodophenyl)-3-(4H-1,2,4-triazol-3-ylthio)-2,5-pyrrolidinedione (LabMol-28, 2) |  | 1.0 | 0.6 | -0.61 ± 0.09 | -0.48 ± 0.05 |
| 2-[2-(3-methyl-4-nitro-5-isoxazolyl)vinyl]pyridine (LabMol-37, 3) |  | 1.0 | 0.6 | -0.92 ± 0.04 | -0.58 ± 0.01 |
| 2-(benzylsulfonyl)-1,3-benzothiazole (LabMol-49, 4) |  | 1.0 | 0.8 | -0.93 ± 0.02 | -0.61 ± 0.09 |
| 3-[(4-acetylphenyl)amino]-1-(2-thienyl)-1-propanone (LabMol-50, 5) |  | 1.0 | 0.8 | -0.89 ± 0.02 | -0.40 ± 0.03 |
| 3-(2-furyl)-1-phenyl-1H-pyrazole-4-carbonitrile (LabMol-51, 6) |  | 1.0 | 1.0 | -0.33 ± 0.13 | -0.24 ± 0.05 |
| 4-(4-fluorobenzoyl)-5-methyl-2-phenyl-2,4-dihydro-3H-pyrazol-3-one (LabMol-24, 7) |  | 1.0 | 0.8 | 0.15 ± 0.05 | -0.17 ± 0.04 |

| | | | | | |
|---|---|---|---|---|---|
| N-[(2-hydroxy-1-naphthyl)methylene]-4-methylbenzenesulfonamide (LabMol-25, 8) |  | 1.0 | 0.6 | -0.12 ± 0.09 | -0.08 ± 0.00 |
| N-(4-fluorobenzylidene)-4-methylbenzenesulfonamide (LabMol-26, 9) |  | 1.0 | 0.6 | -0.04 ± 0.04 | -0.05 ± 0.02 |
| 1-(3-bromophenyl)-2,5-dioxo-3-pyrrolidinyl N'-phenylimidothiocarbamate (LabMol-27, 10) |  | 1.0 | 0.6 | -0.49 ± 0.06 | -0.08 ± 0.04 |
| 5-[(3-carboxyacryloyl)amino]-2-(4-morpholinyl)benzoic acid (LabMol-29, 11) |  | 1.0 | 0.6 | 0.34 ± 0.10 | -0.10 ± 0.14 |
| 4-({[(2-furylmethyl)amino]carbonothioyl}amino)benzenesulfonamide (LabMol-30, 12) |  | 1.0 | 0.6 | 0.13 ± 0.01 | -0.15 ± 0.04 |
| 2-[(4-methylphenyl)sulfonyl]-1-(3-nitrophenyl)ethanone (LabMol-31, 13) |  | 1.0 | 0.6 | 0.22 ± 0.04 | -0.11 ± 0.03 |
| 1-benzofuran-2-yl(4-methyl-3-nitrophenyl)methanone (LabMol-32, 14) |  | 1.0 | 0.8 | 0.06 ± 0.05 | -0.12 ± 0.01 |
| 1-{[(4-methylphenyl)sulfonyl]methyl}-3-nitrobenzene (LabMol-33, 15) |  | 1.0 | 0.8 | -0.07 ± 0.08 | -0.12 ± 0.06 |

| | | | | | |
|---|---|---|---|---|---|
| 2-{[4-allyl-5-(2-furyl)-4H-1,2,4-triazol-3-yl]thio}acetamide (LabMol-34, 16) |  | 1.0 | 0.6 | 0.17 ± 0.13 | 0.02 ± 0.15 |
| N-allyl-2-(2-furyl)-4-(phenylsulfonyl)-1,3-oxazol-5-amine (LabMol-35, 17) |  | 1.0 | 0.8 | -0.52 ± 0.02 | -0.08 ± 0.02 |
| 2-(2-phenylvinyl)-4-quinolinol (LabMol-36, 18) |  | 1.0 | 1.0 | -0.41 ± 0.10 | -0.15 ± 0.02 |
| 1-(2,4-dichlorobenzyl)-4-nitro-1H-imidazole (LabMol-38, 19) |  | 1.0 | 0.6 | 0.00 ± 0.15 | -0.12 ± 0.07 |
| 1-(3-chlorobenzyl)-3-nitro-1H-1,2,4-triazole (LabMol-39, 20) |  | 1.0 | 0.6 | 0.02 ± 0.06 | -0.04 ± 0.02 |
| 4-chloro-1-(2-chloro-4-fluorobenzyl)-3-nitro-1H-pyrazole (LabMol-40, 21) |  | 1.0 | 0.6 | -0.28 ± 0.01 | -0.13 ± 0.02 |
| 1-(3-chlorobenzyl)-3-methoxy-4-nitro-1H-pyrazole (LabMol-41, 22) |  | 1.0 | 0.8 | -0.26 ± 0.05 | -0.19 ± 0.04 |
| 1-(2,4-dichlorophenyl)-2-[(4-methylphenyl)sulfonyl]ethanone (LabMol-42, 23) |  | 1.0 | 0.8 | -0.05 ± 0.19 | -0.15 ± 0.08 |

| | | | | | |
|---|---|---|---|---|---|
| 2-{[1-(3-chloro-4-methylphenyl)-2,5-dioxo-3-pyrrolidinyl]thio}nicotinic acid (LabMol-43, 24) |  | 1.0 | 0.8 | 0.26 ± 0.10 | -0.13 ± 0.01 |
| 1-(4-chlorophenyl)-5-(2-furyl)-N-isopropyl-1H-1,2,3-triazole-4-carboxamide (LabMol-44, 25) |  | 1.0 | 0.8 | 0.20 ± 0.13 | -0.10 ± 0.02 |
| 1-(1-benzofuran-2-yl)-2-(phenylsulfonyl)ethanone (LabMol-45, 26) |  | 1.0 | 0.8 | 0.18 ± 0.12 | -0.14 ± 0.02 |
| N-[2-(3-pyridinyl)-2-(2-thienylsulfonyl)ethyl]-2-furamide (LabMol-46, 27) |  | 1.0 | 0.8 | 0.35 ± 0.04 | 0.00 ± 0.12 |
| N-(2-furylmethyl)-4-(1-piperidinylsulfonyl)-2-thiophenecarboxamide (LabMol-47, 28) |  | 1.0 | 0.8 | 0.10 ± 0.05 | -0.11 ± 0.03 |
| 4-(3-carboxypropanoyl)benzoic acid (LabMol-48, 29) |  | 1.0 | 0.6 | 0.12 ± 0.07 | -0.10 ± 0.01 |
| Praziquantel (PZQ) |  | 0.0 | 0.8 | -0.48 ± 0.04 | -0.17 ± 0.02 |
| Oltipraz (OLT) |  | 0.8 | 1.0 | -0.90 ± 0.04 | -0.34 ± 0.07 |

Prob.: probability; AD: applicability domain coverage; The identified hits are highlighted in bold fonts

**Table S7.** $EC_{50}$ values for the effect of investigated compounds and PZQ on the motility of male and female *S. mansoni* exposed for up to 72h.

| Time (h) | Compound | $EC_{50}$ (μM) | |
|---|---|---|---|
| | | Male | Female |
| 24h | PZQ | 0.26 | 0.28 |
| | 1 | 9.59 | 12.3 |
| | 2 | 17.0 | 19.8 |
| | 3 | 17.5 | 16.1 |
| | 4 | 35.1 | 21.9 |
| 48h | PZQ | No fit | 0.64 |
| | 1 | 29.8 | 5.77 |
| | 2 | 10.2 | 17.9 |
| | 3 | 6.43 | 5.68 |
| | 4 | 21.1 | 4.91 |
| 72h | PZQ | 0.22 | 0.59 |
| | 1 | 28.1 | No Fit |
| | 2 | 11.4 | No Fit |
| | 3 | No fit | 5.84 |
| | 4 | 20.4 | 5.83 |

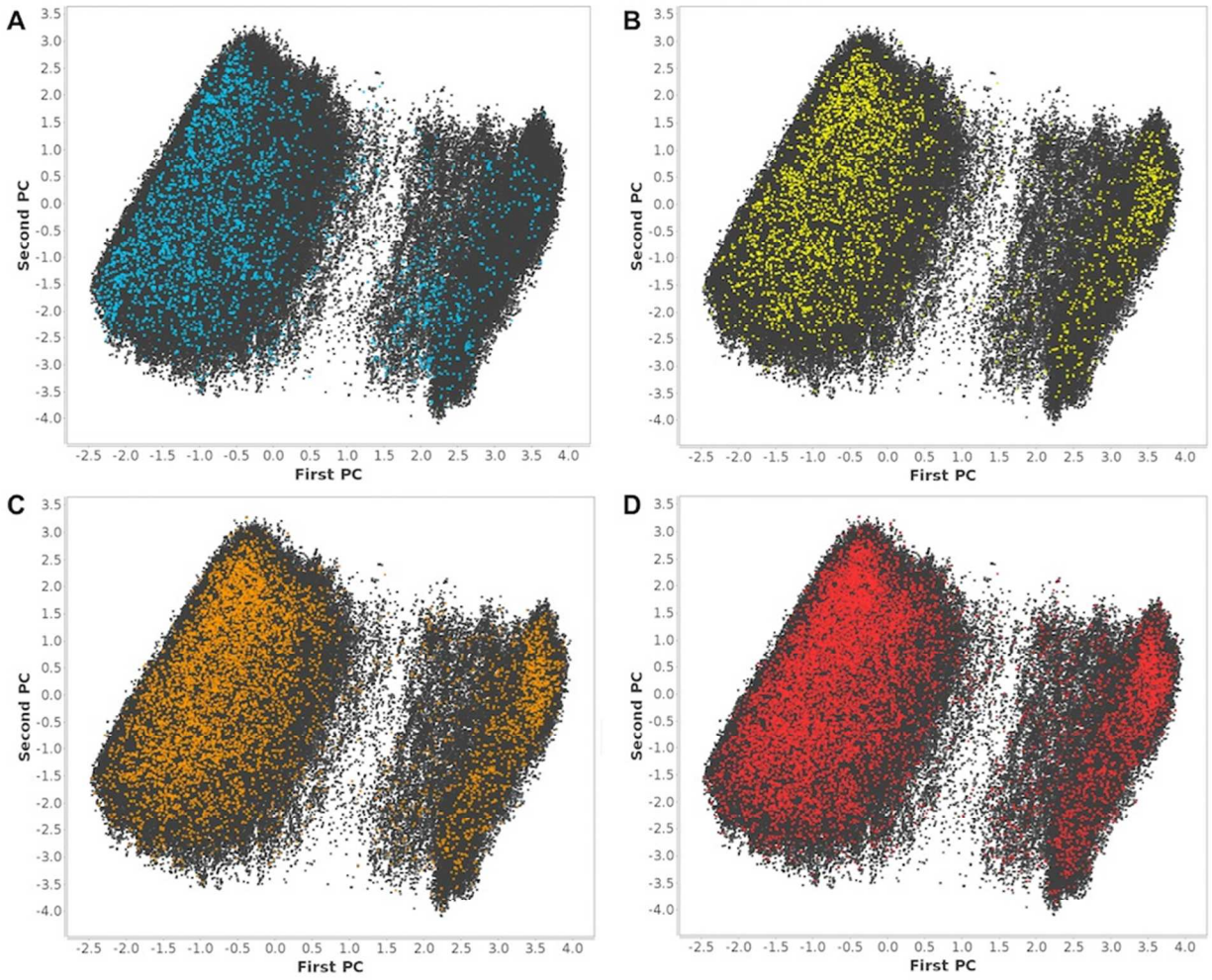



**Figure S1.** Chemical space of *Sm*TGR inhibitors and non-inhibitors into first two PCs computed using MACCS keys; blue dots representing the 2,854 inhibitors at 10 μM threshold (A); Yellow (B), orange (C), and red (D) dots represent non-inhibitors selected for dataset balancing with ratios of 1:1, 1:2, and 1:3 correspondingly; grey dots represent remaining non-inhibitors.

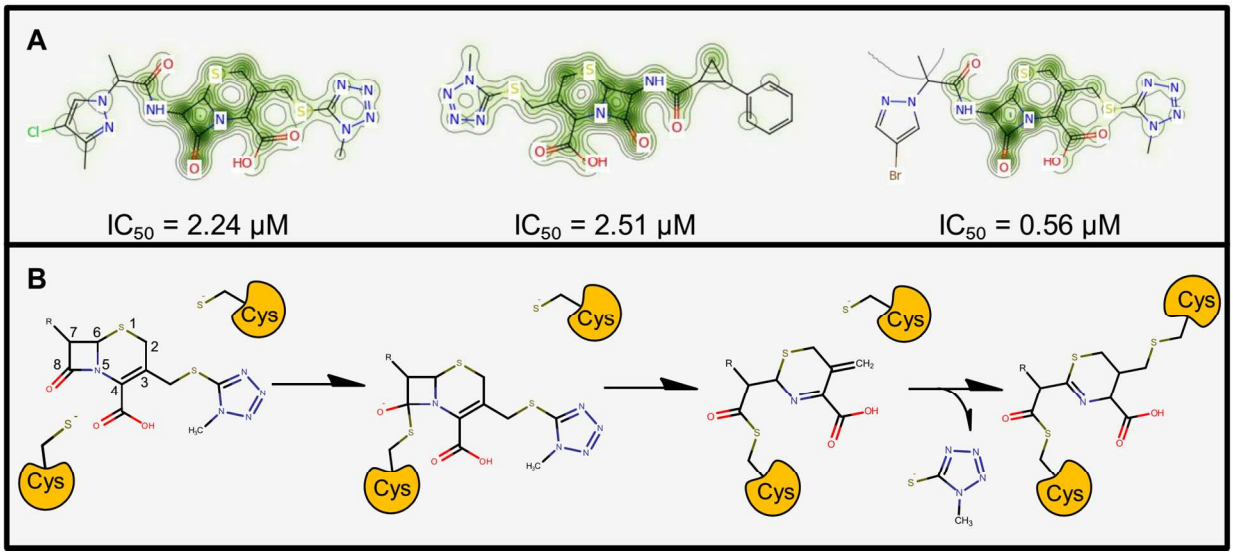



**Figure S2.** Predicted probability maps generated for cephalosporins (A) and their probable reaction mechanism in the *Sm*TGR active site (B).

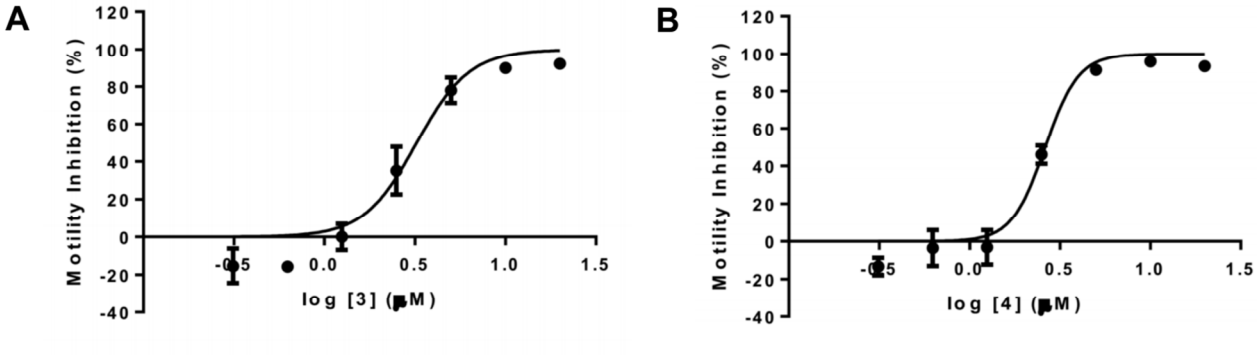



**Figure S3.** Motility dose-response curves for compounds **3** (A) and **4** (B) against *S. mansoni* larvae after 48h of incubation.

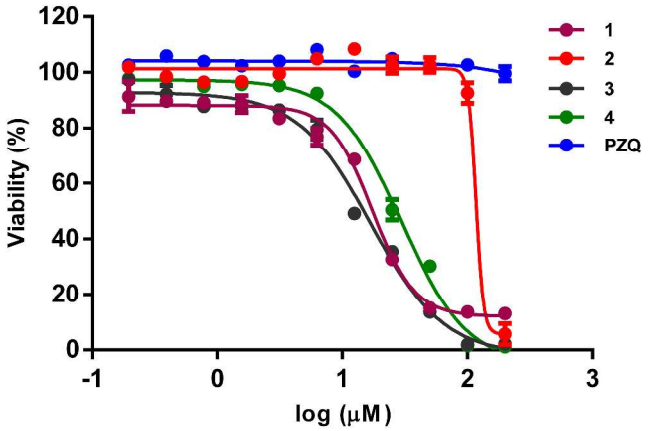



**Figure S4.** Viability dose-responses curves for compounds **1**−**4** and PZQ against WSS-1 human cells after 48h of incubation.

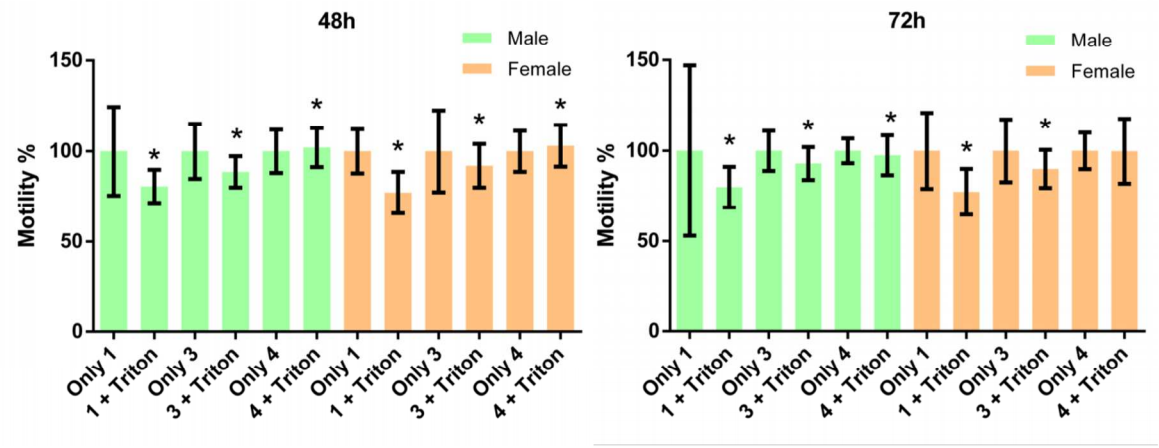



**Figure S5.** Effect of 0.01% Triton X-100 co-incubated with compounds **1**, **3** and **4** on adult male and female *S. mansoni* worms. Motility measurements were performed after 48h and 72h, and compared with group without detergent. Data expressed as mean ± standard deviation. * $P \leq 0.05$ using student-*t* test.

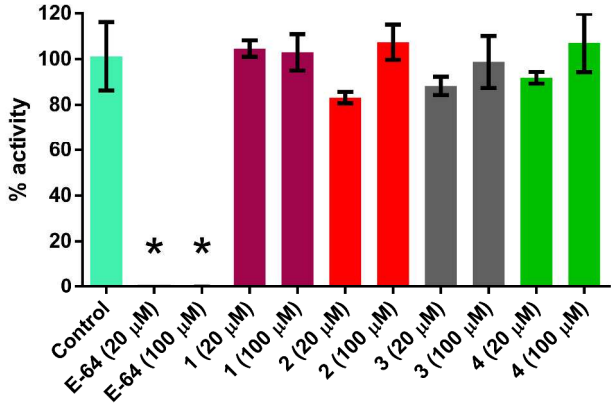



**Figure S6.** Papain activity in the presence of compounds **1**−**4** and E-64. Data expressed as mean ± standard deviation. * $P \leq 0.05$ relative to control using ANOVA followed by Dunnett´s test.

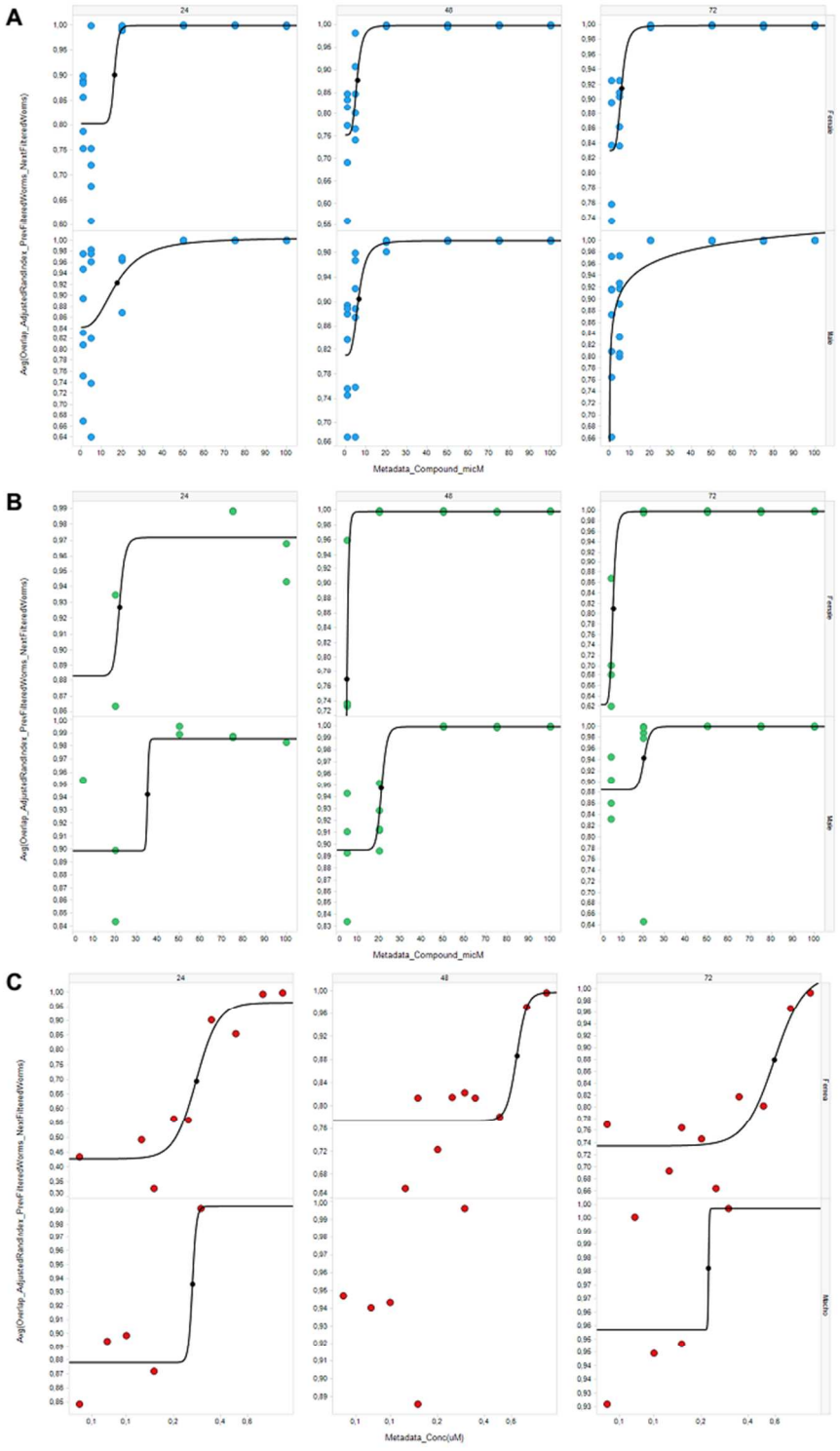
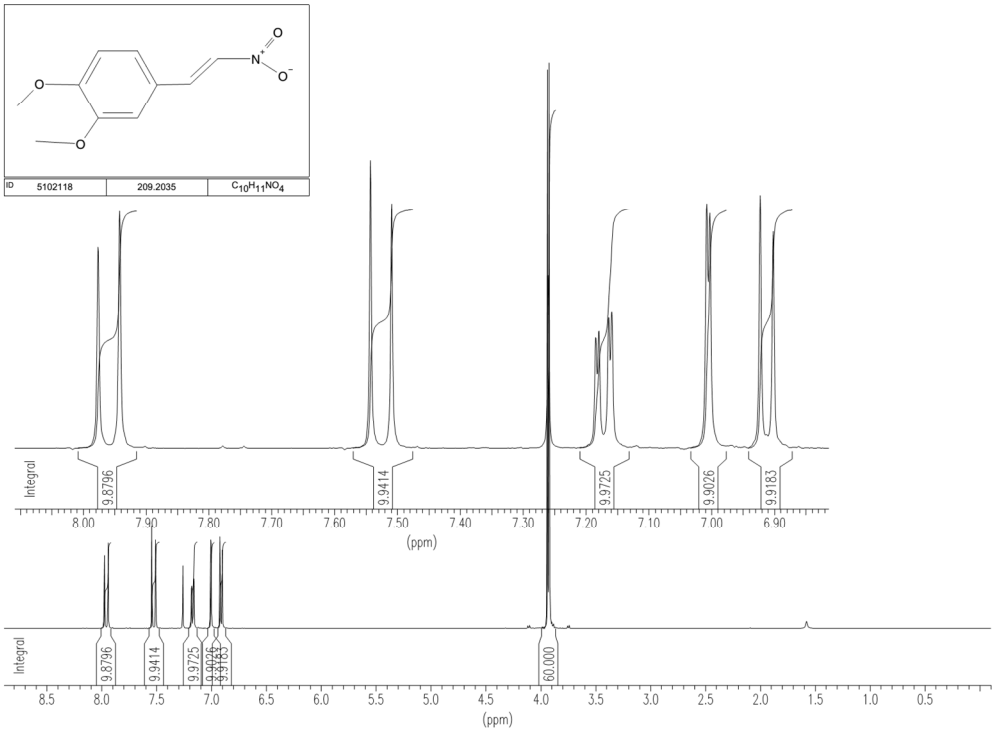


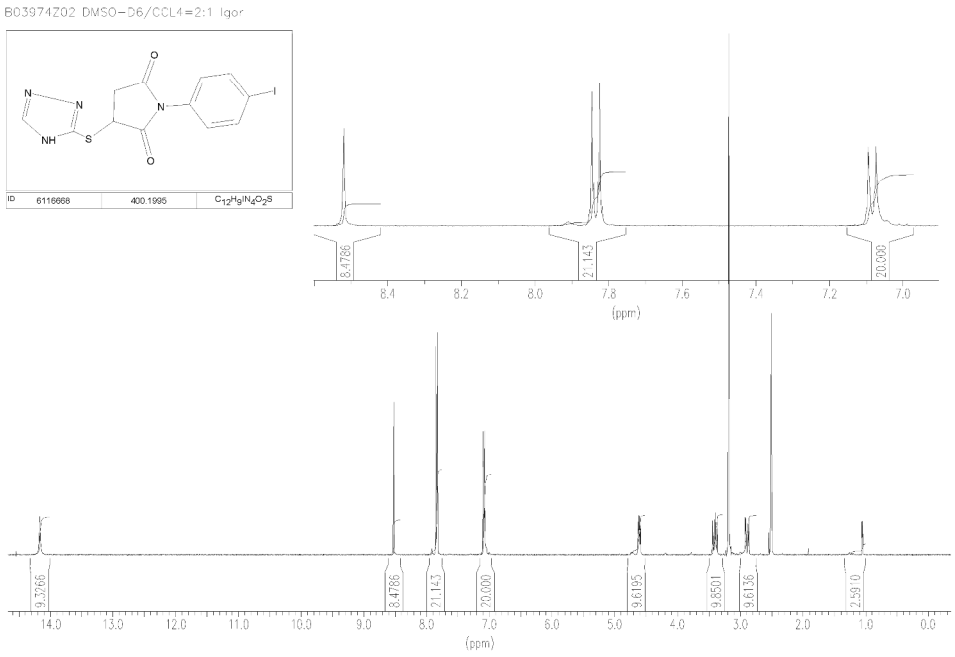

**Figure S7.** Dose-response curves for compounds **3** (A), **4** (B) and PZQ (C) on the motility of male and female *S. mansoni* exposed for up to 72h of incubation.

**Nuclear Magnetic Resonance (NMR) and Purity Data.** The chemical structure of all compounds purchased from ChemBridge was confirmed using proton ($^1$H) NMR spectra at 300/400 MHz. The $^1$H RMN spectrums of compounds are listed below. The Liquid Chromatography–Mass Spectrometry (LC-MS) analysis with evaporative light scattering and ultraviolet detectors confirmed a minimum purity of 95% for all samples.

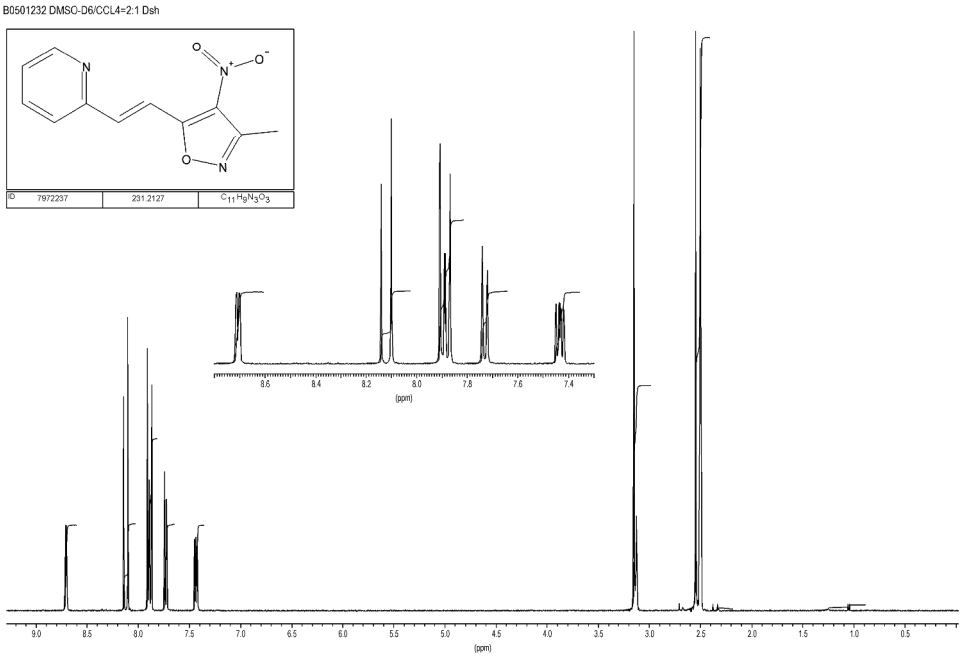
1,2-dimethoxy-4-(2-nitrovinyl)benzene (LabMol-23, **1**);



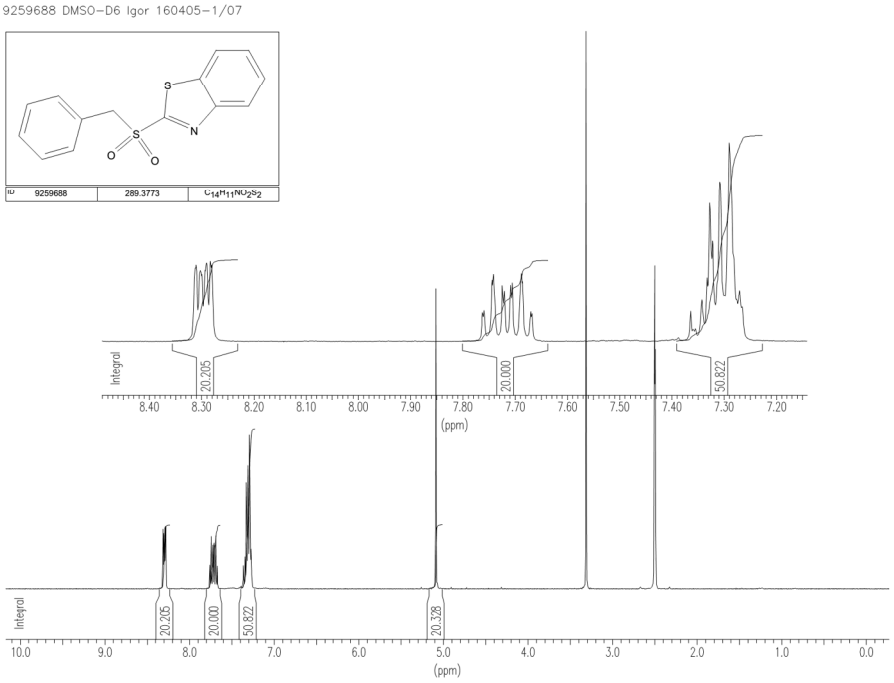
1-(4-iodophenyl)-3-(4H-1,2,4-triazol-3-ylthio)-2,5-pyrrolidinedione (LabMol-28, **2**);



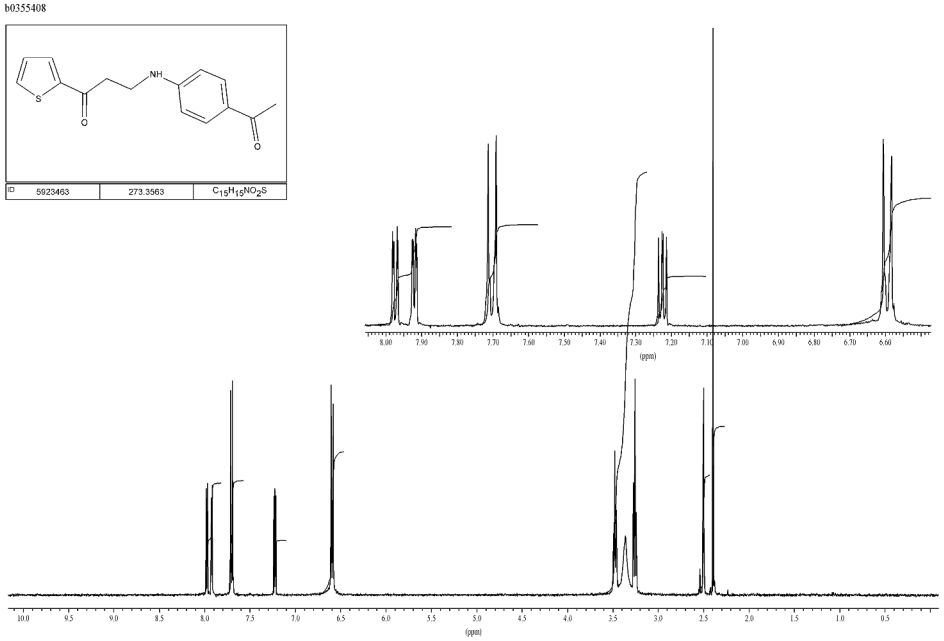

2-[2-(3-methyl-4-nitro-5-isoxazolyl)vinyl]pyridine (LabMol-37, **3**);



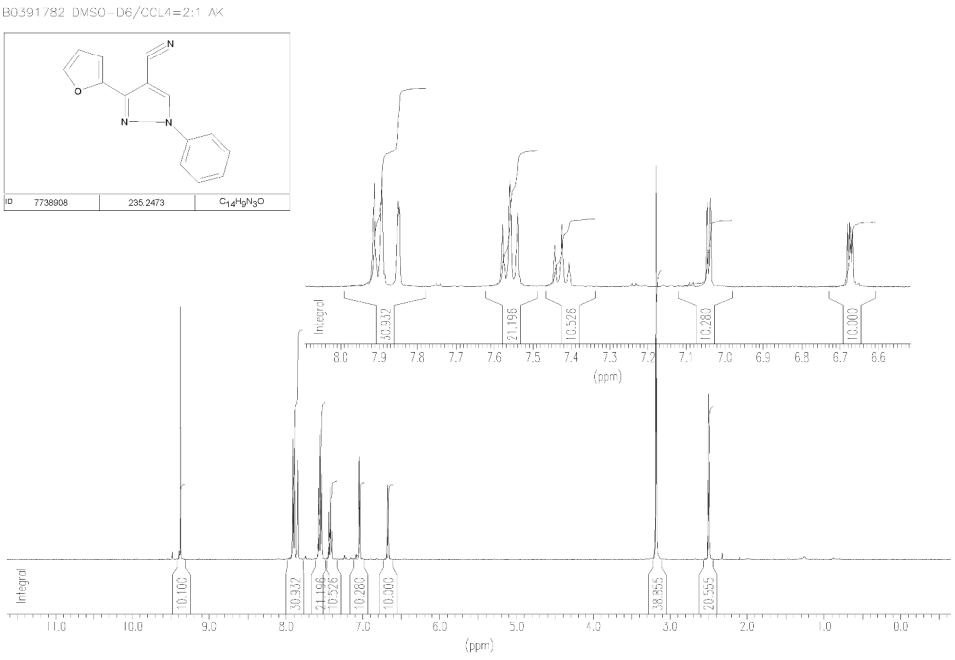
2-(benzylsulfonyl)-1,3-benzothiazole (LabMol-49, **4**);



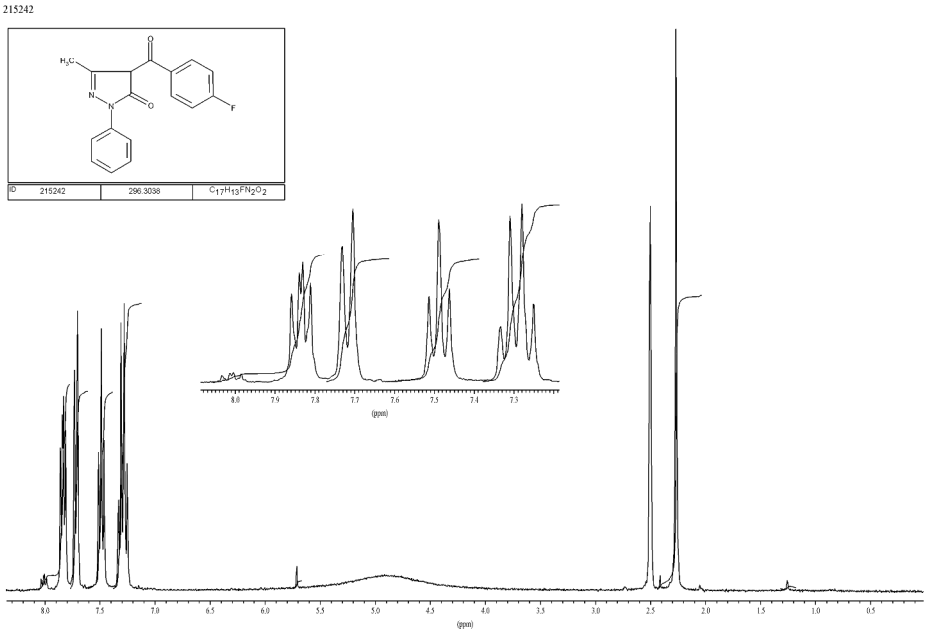

3-[(4-acetylphenyl)amino]-1-(2-thienyl)-1-propanone (LabMol-50, **5**);



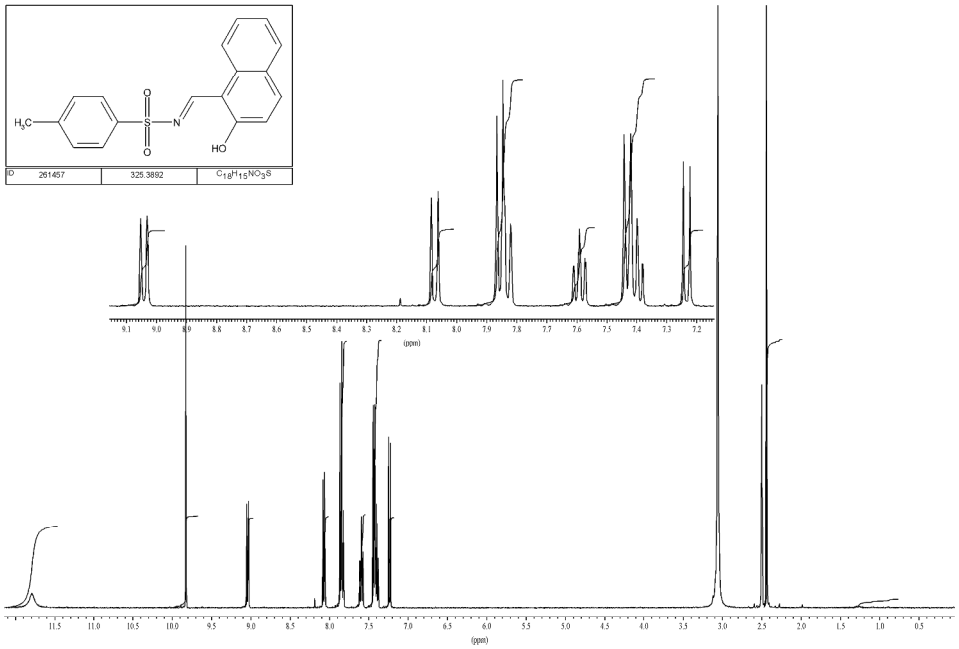
3-(2-furyl)-1-phenyl-1H-pyrazole-4-carbonitrile (LabMol-51, **6**);



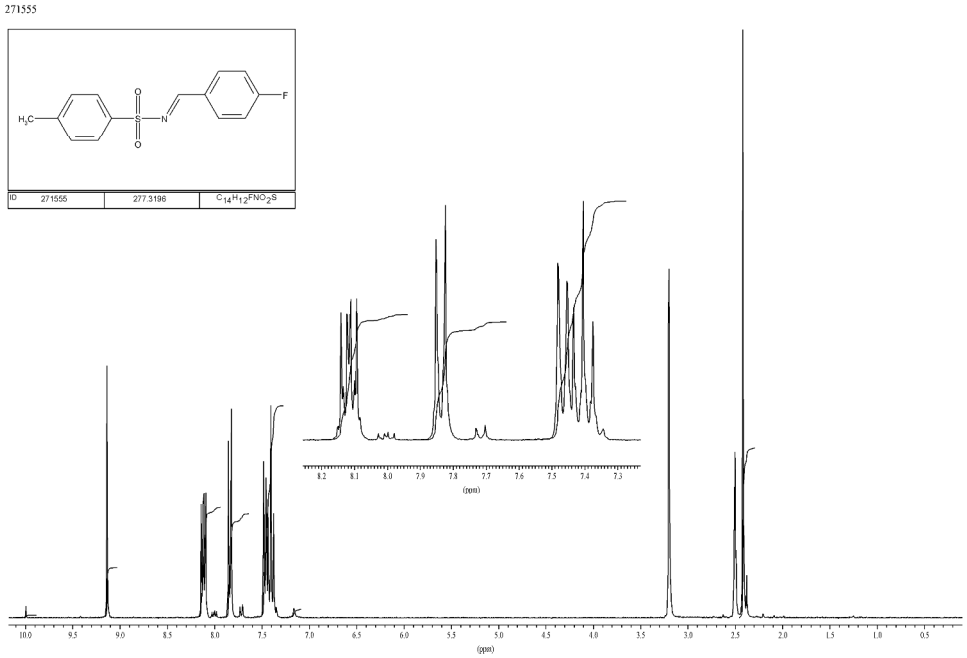

4-(4-fluorobenzoyl)-5-methyl-2-phenyl-2,4-dihydro-3H-pyrazol-3-one (LabMol-24, **7**);



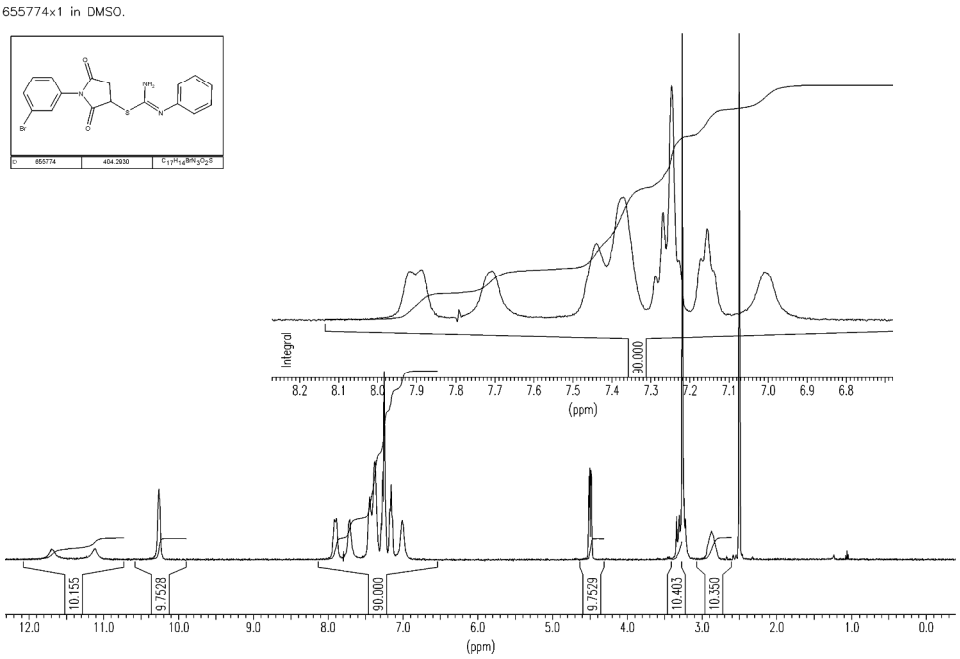
N-[(2-hydroxy-1-naphthyl)methylene]-4-methylbenzenesulfonamide (LabMol-25, **8**);



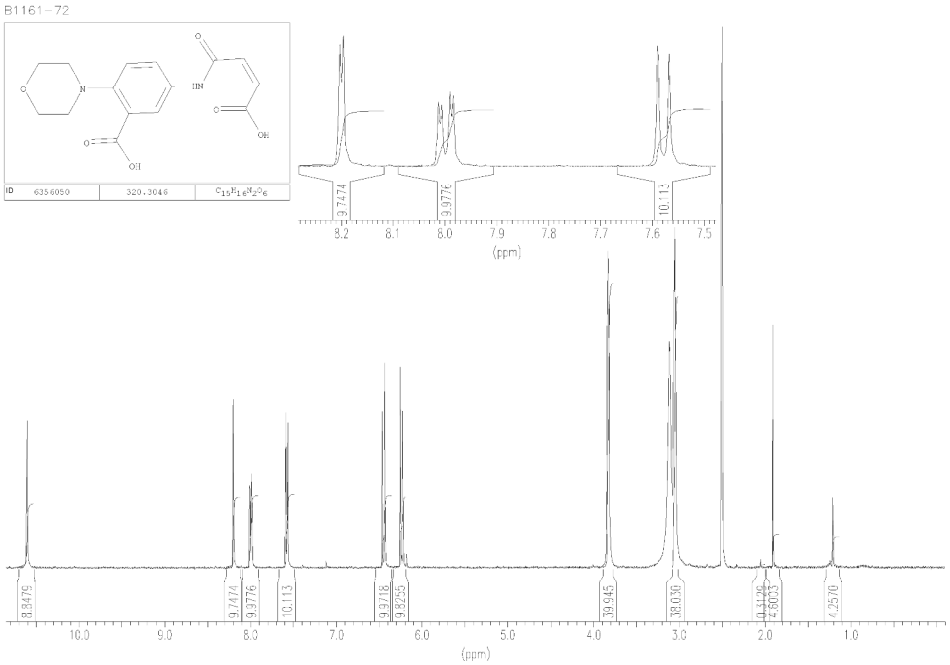

N-(4-fluorobenzylidene)-4-methylbenzenesulfonamide (LabMol-26, **9**);



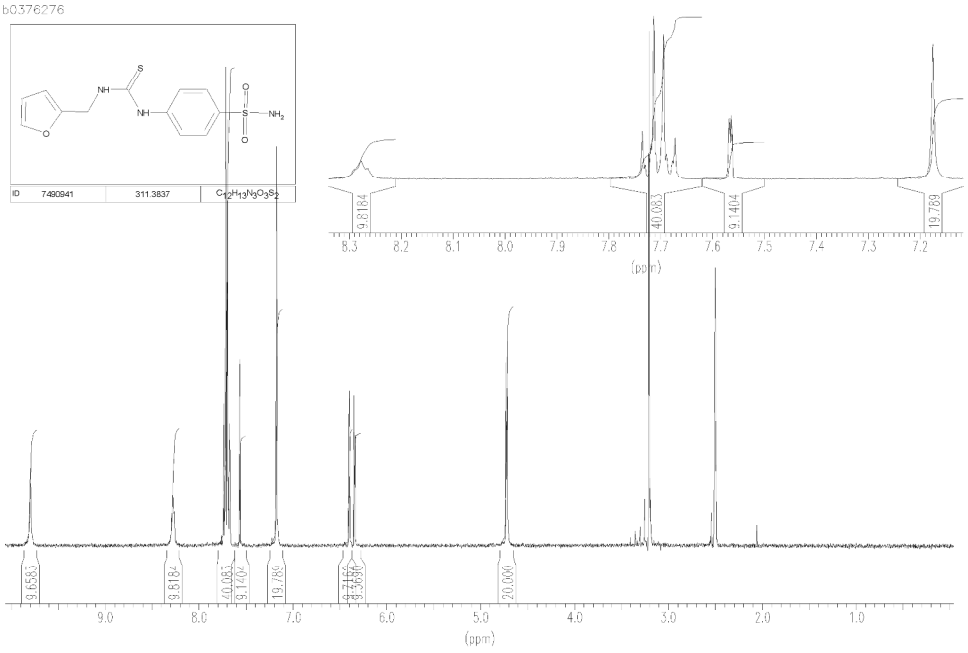
1-(3-bromophenyl)-2,5-dioxo-3-pyrrolidinyl N'-phenylimidothiocarbamate (LabMol-27, **10**);



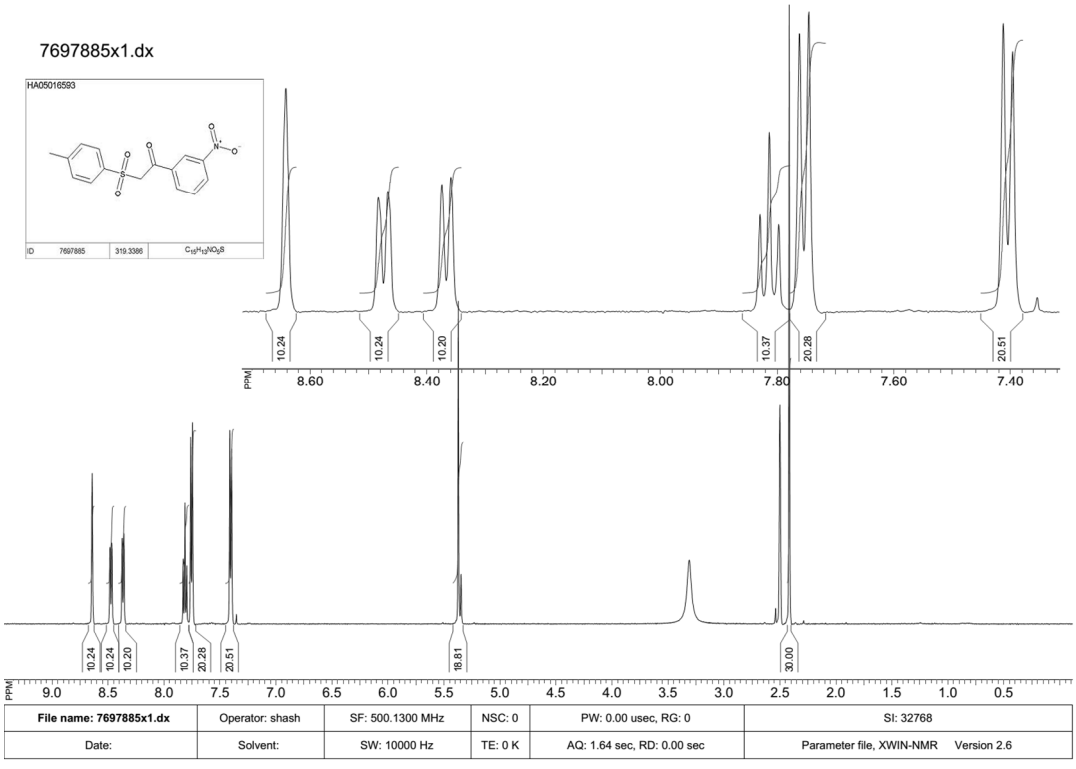

5-[(3-carboxyacryloyl)amino]-2-(4-morpholinyl)benzoic acid (LabMol-29, **11**);



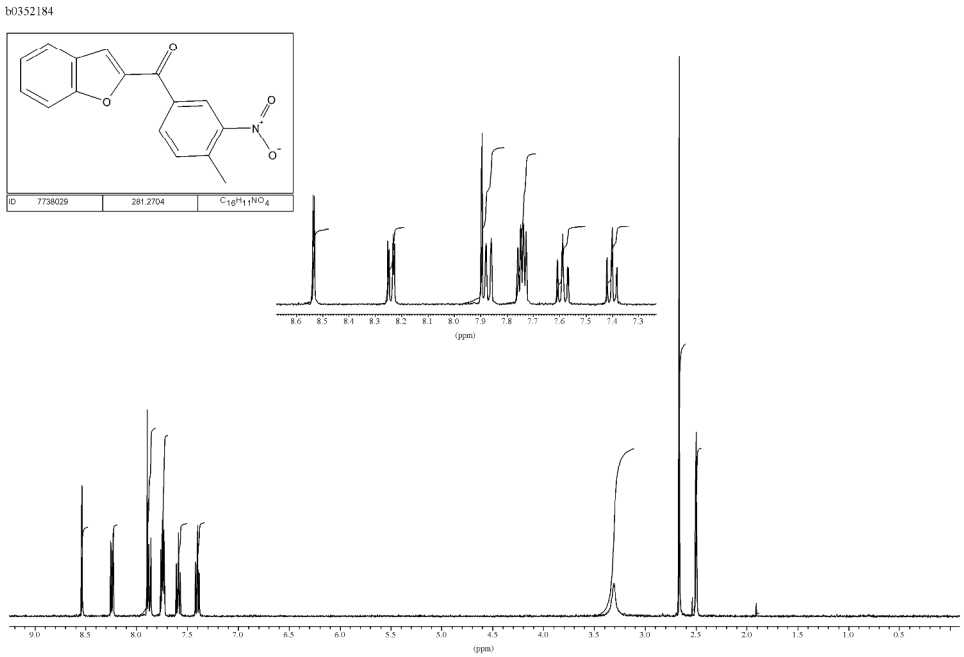
4-({[(2-furylmethyl)amino]carbonothioyl}amino)benzenesulfonamide (LabMol-30, **12**);



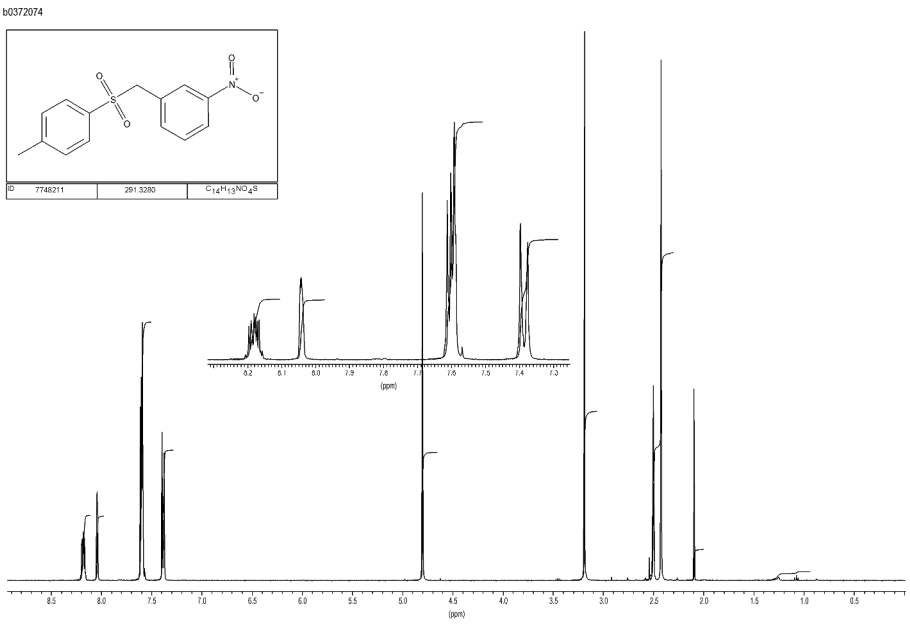

2-[(4-methylphenyl)sulfonyl]-1-(3-nitrophenyl)ethanone (LabMol-31, **13**);



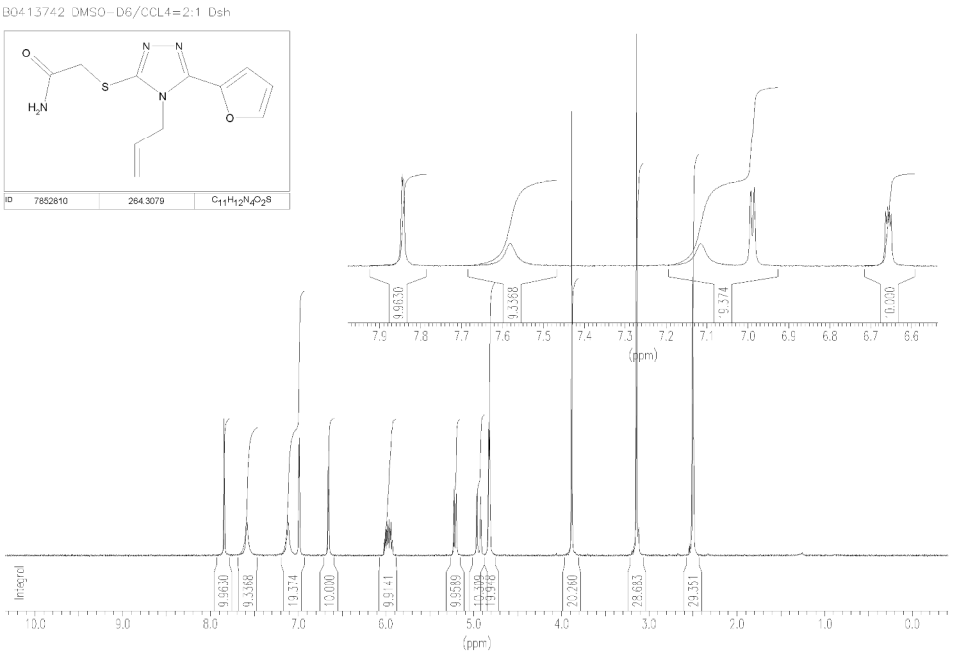
1-benzofuran-2-yl(4-methyl-3-nitrophenyl)methanone (LabMol-32, **14**);



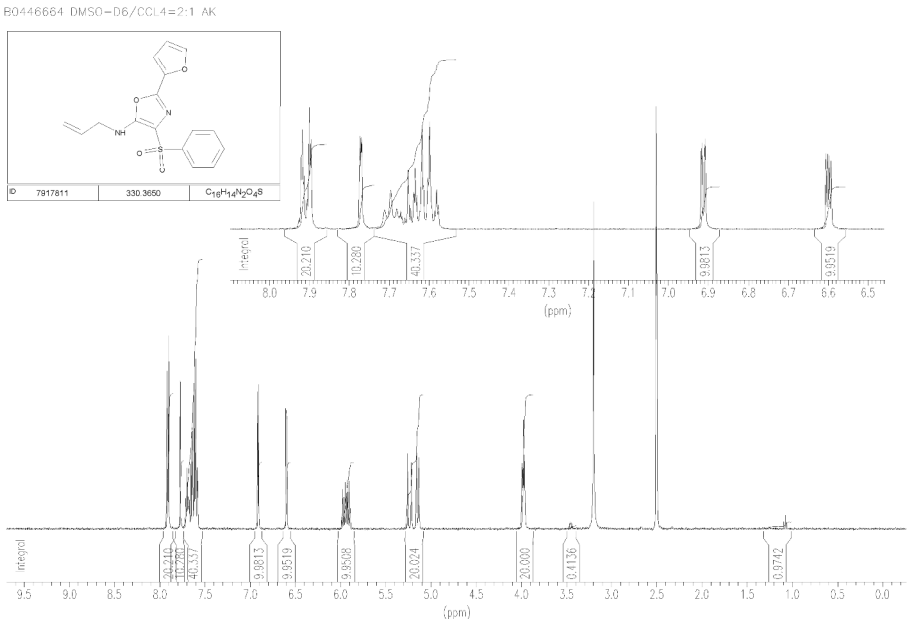

1-{[(4-methylphenyl)sulfonyl]methyl}-3-nitrobenzene (LabMol-33, **15**);



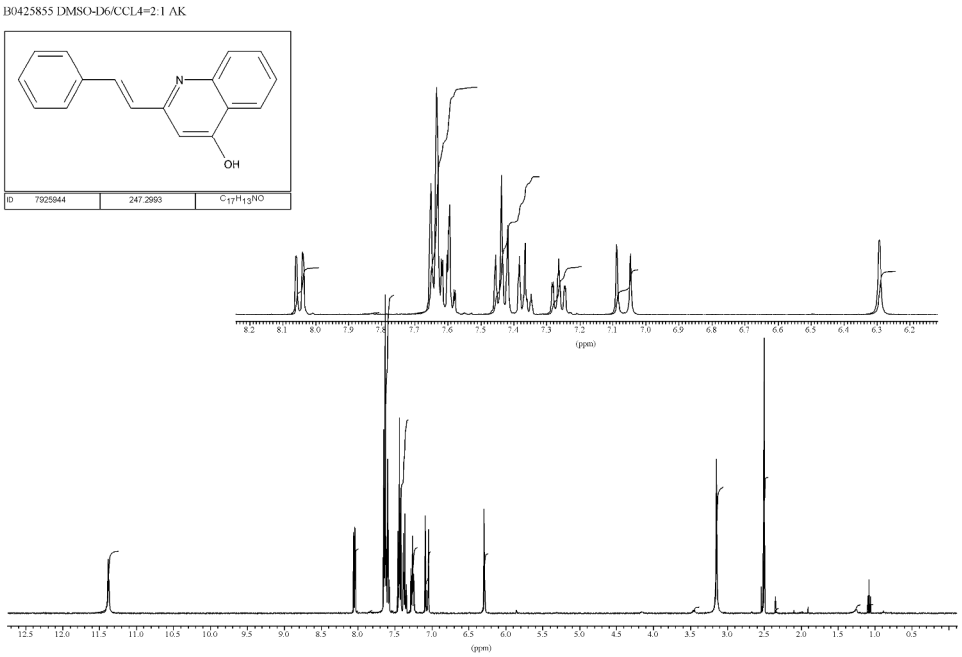
2-{[4-allyl-5-(2-furyl)-4H-1,2,4-triazol-3-yl]thio}acetamide (LabMol-34, **16**);



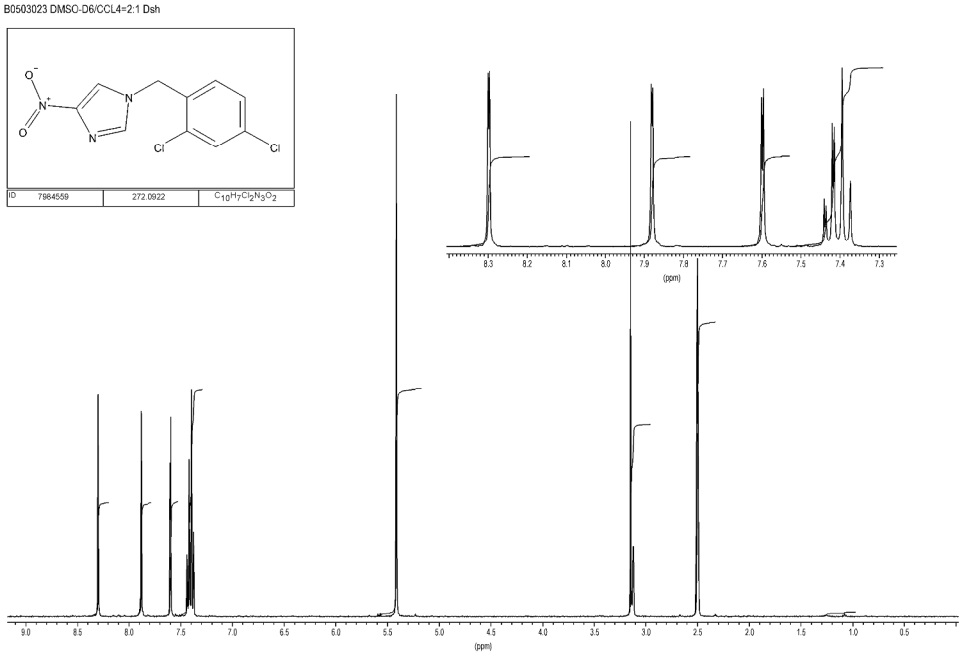

N-allyl-2-(2-furyl)-4-(phenylsulfonyl)-1,3-oxazol-5-amine (LabMol-35, **17**);



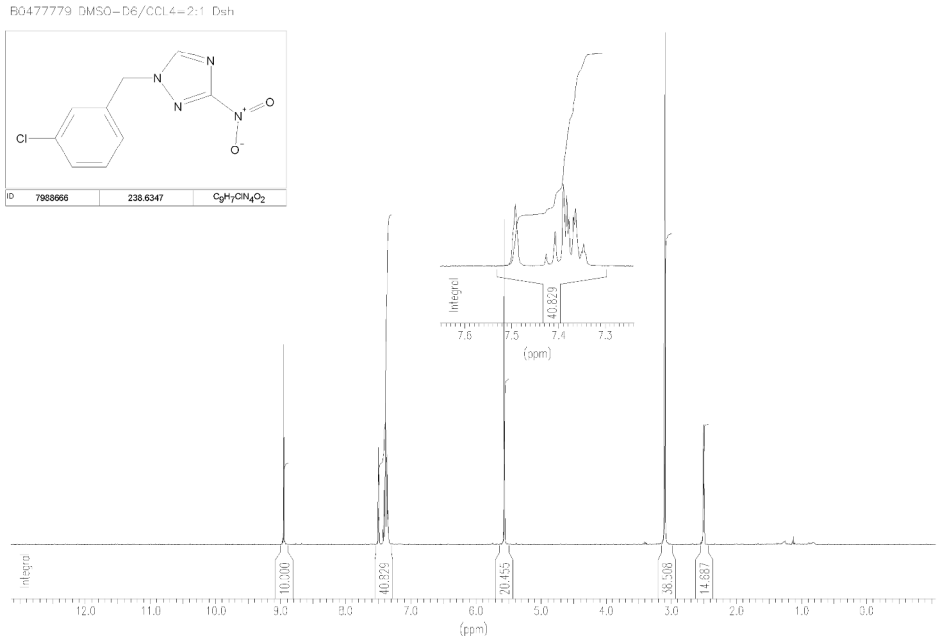
2-(2-phenylvinyl)-4-quinolinol (LabMol-36, **18**);



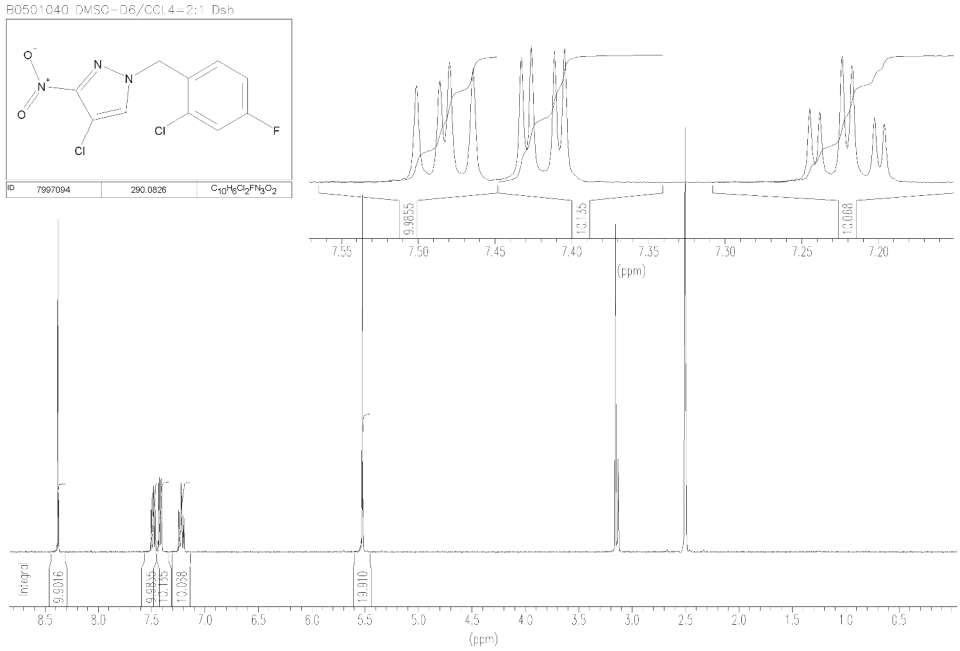

1-(2,4-dichlorobenzyl)-4-nitro-1H-imidazole (LabMol-38, **19**);



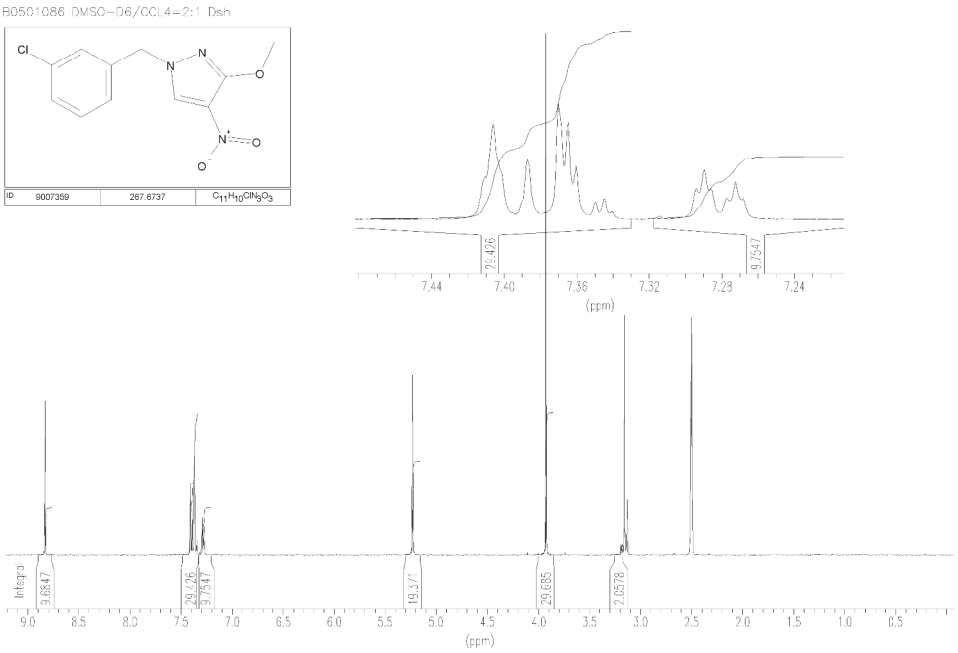
1-(3-chlorobenzyl)-3-nitro-1H-1,2,4-triazole (LabMol-39, **20**);



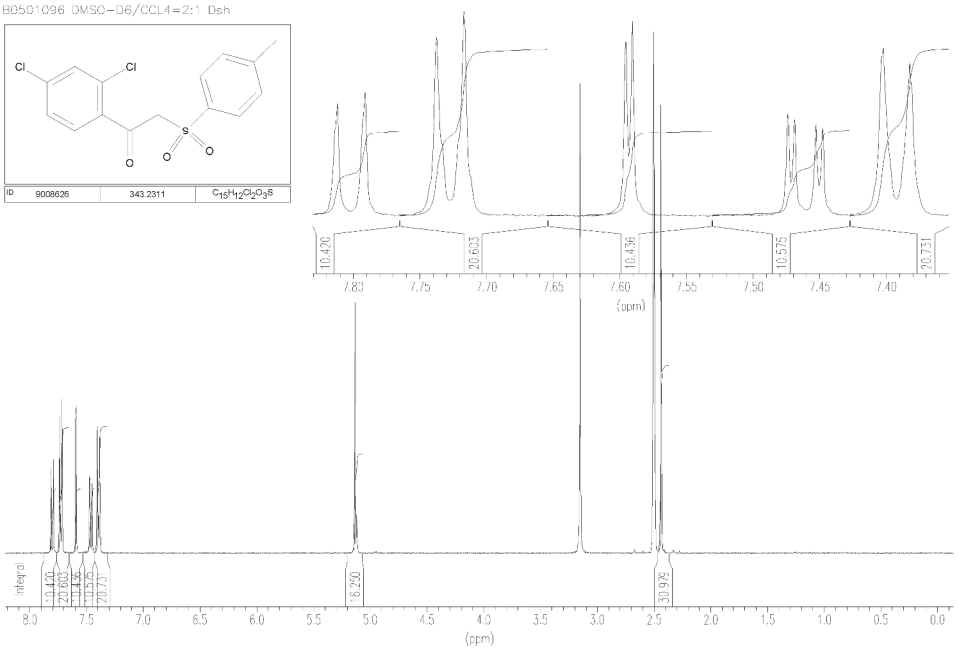

4-chloro-1-(2-chloro-4-fluorobenzyl)-3-nitro-1H-pyrazole (LabMol-40, **21**);



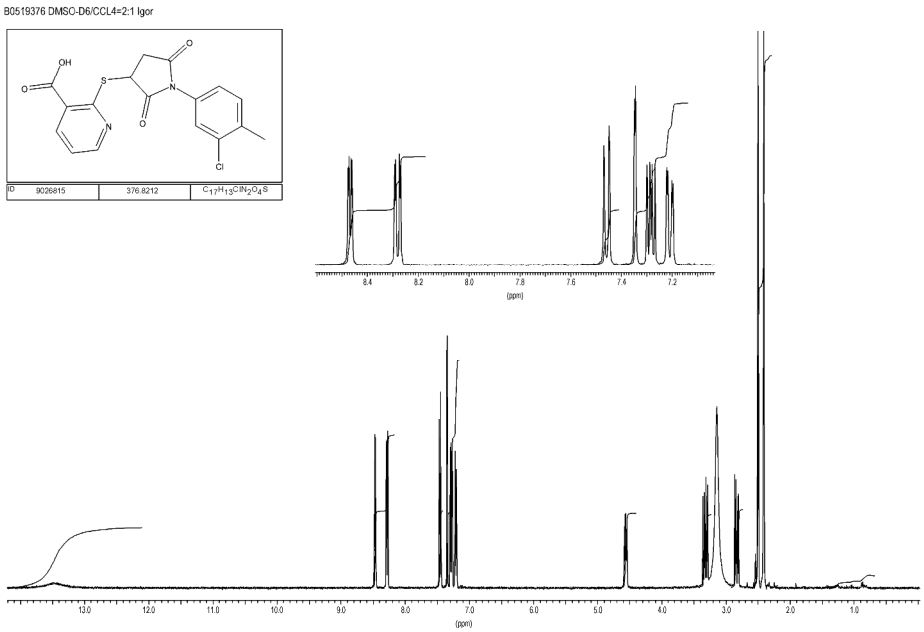
1-(3-chlorobenzyl)-3-methoxy-4-nitro-1H-pyrazole (LabMol-41, **22**);



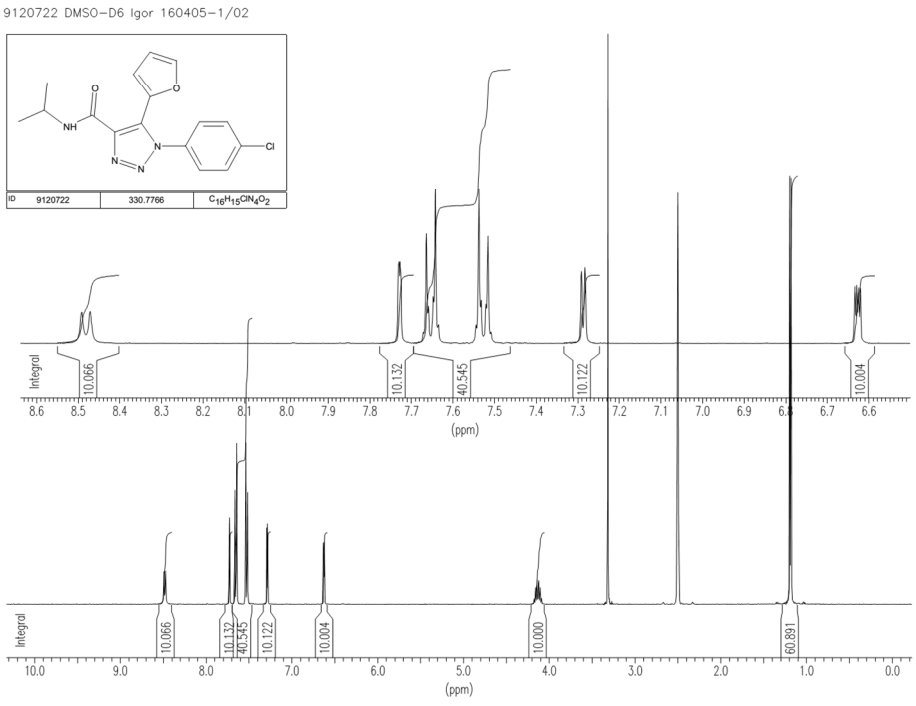

1-(2,4-dichlorophenyl)-2-[(4-methylphenyl)sulfonyl]ethanone (LabMol-42, **23**);



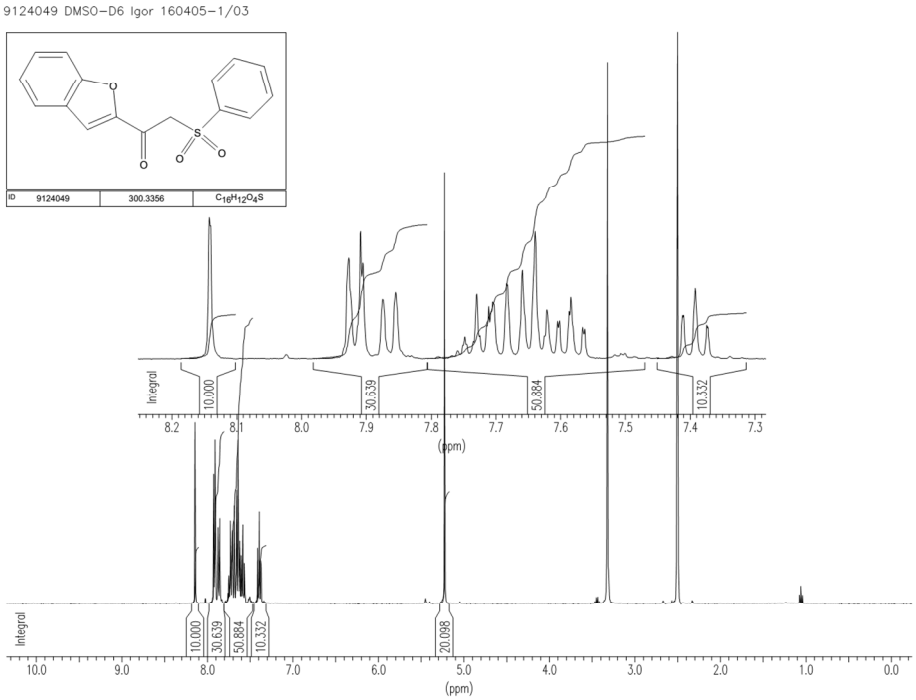
2-{[1-(3-chloro-4-methylphenyl)-2,5-dioxo-3-pyrrolidinyl]thio}nicotinic acid (LabMol-43, **24**);



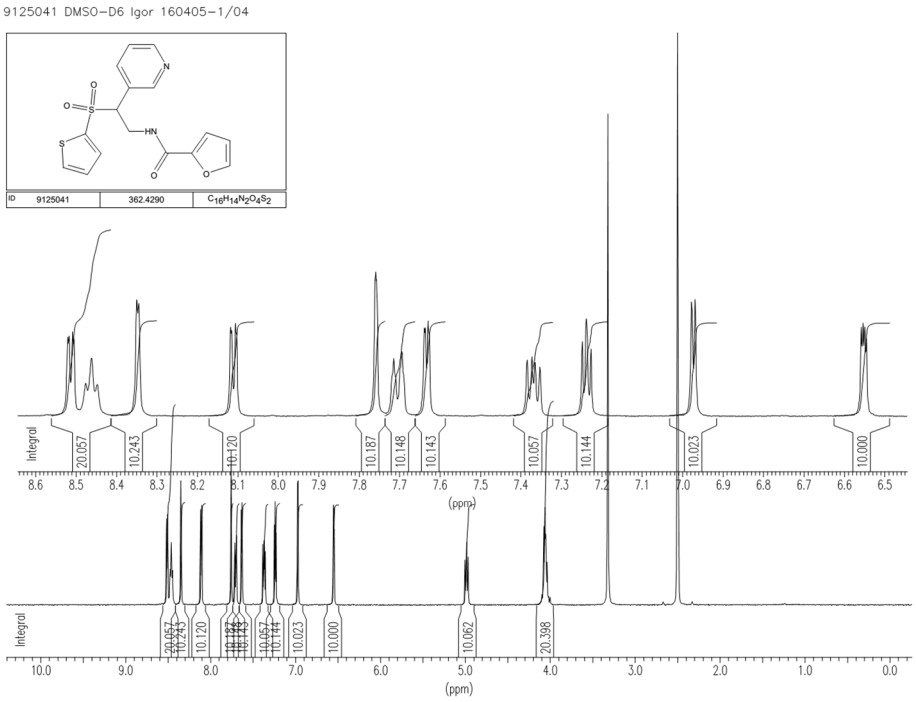

1-(4-chlorophenyl)-5-(2-furyl)-N-isopropyl-1H-1,2,3-triazole-4-carboxamide (LabMol-44, **25**);



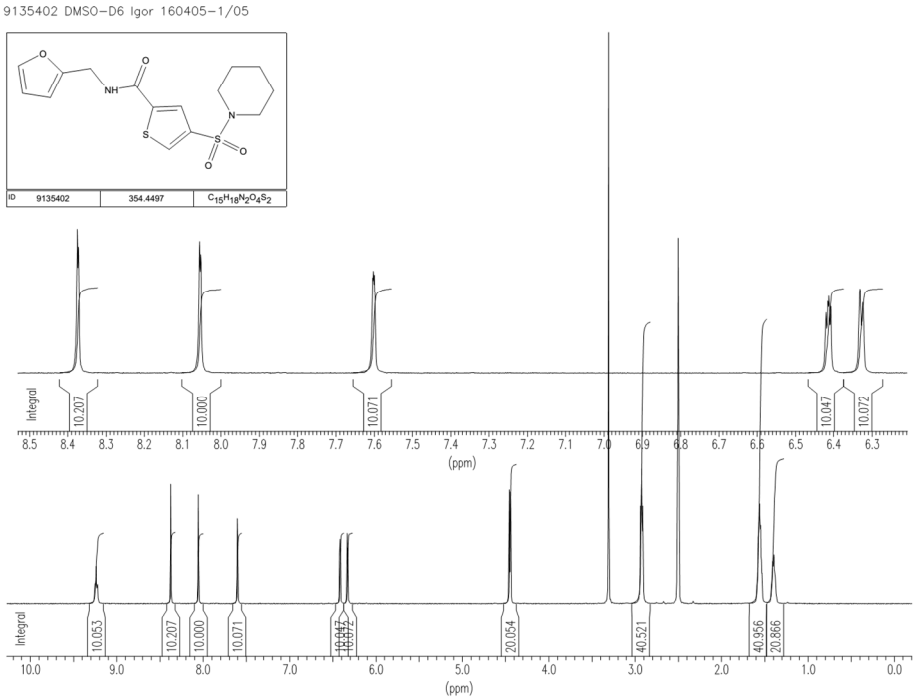
1-(1-benzofuran-2-yl)-2-(phenylsulfonyl)ethanone (LabMol-45, **26**);



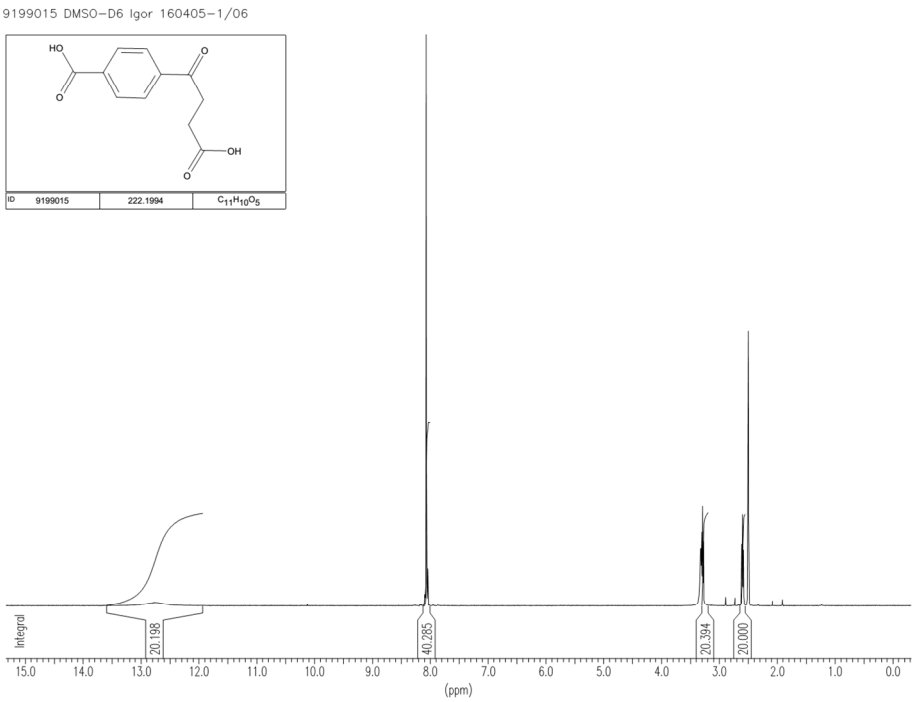

N-[2-(3-pyridinyl)-2-(2-thienylsulfonyl)ethyl]-2-furamide (LabMol-46, **27**);



N-(2-furylmethyl)-4-(1-piperidinylsulfonyl)-2-thiophenecarboxamide (LabMol-47, **28**);




4-(3-carboxypropanoyl)benzoic acid (LabMol-48, **29**);



**REFERENCES**

(1) MACCS Structural Keys. 2013, Accelrys, San Diego, CA.

(2) Morgan, H. L. The Generation of a Unique Machine Description for Chemical Structures-A Technique Developed at Chemical Abstracts Service. *J. Chem. Doc.* **1965**, *5* (2), 107–113.

(3) Riniker, S.; Landrum, G. a. Open-Source Platform to Benchmark Fingerprints for Ligand-Based Virtual Screening. *J. Cheminform.* **2013**, *5* (1), 26.

(4) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50* (5), 742–754.

(5) Baskin, I.; Varnek, A. Building a Chemical Space Based on Fragment Descriptors. *Comb. Chem. High Throughput Screen.* **2008**, *11* (8), 661–668.

(6) Carhart, R. E.; Smith, D. H.; Venkataraghavan, R. Atom Pairs as Molecular Features in Structure-Activity Studies: Definition and Applications. *J. Chem. Inf. Comput. Sci.* **1985**, *25* (4), 64–73.

(7) Breiman, L.; Friedman, J.; Olshen, R. A.; Charles, J. S. *Classification and Regression Trees*, 1nd ed.; Breiman, L., Ed.; Wadsworth & Brooks/Cole Advanced Books & Software: Monterey, 1984.

(8) Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45* (1), 5–32.

(9) Friedman, J. H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **2001**, *29* (5), 1189–1232.

(10) Friedman, J. H. Multivariate Adaptive Regression Splines. *Ann. Stat.* **1991**, *19* (1), 1–67.

(11) Vapnik, V. *The Nature of Statistical Learning Theory*, 2nd ed.; Springer: New York, 2000.

(12) Barker, M.; Rayens, W. Partial Least Squares for Discrimination. *J. Chemom.* **2003**, *17*, 166–173.

(13) Rosenblatt, F. *Principles of Neurodynamics; Perceptrons and the Theory of Brain Mechanisms*, 1nd ed.; Rosenblatt, F., Ed.; Spartan Books: Washington, 1962.

(14) Altman, N. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. *Am. Stat.* **1992**, *46* (3), 175–185.

OPEN ACCESS Freely available online

PLOS | NEGLECTED TROPICAL DISEASES

# *In Silico* Repositioning-Chemogenomics Strategy Identifies New Drugs with Potential Activity against Multiple Life Stages of *Schistosoma mansoni*

**Bruno J. Neves[1,2], Rodolpho C. Braga[1,3], José C. B. Bezerra[2], Pedro V. L. Cravo[2,4], Carolina H. Andrade[1,2]***

**1** LabMol – Laboratory for Drug Design and Modeling, Faculdade de Farmácia, Universidade Federal de Goiás, Goiânia, Brazil, **2** Instituto de Patologia Tropical e Saúde Pública, Universidade Federal de Goiás, Goiânia, Brazil, **3** Instituto de Química, Universidade Federal de Goiás, Goiaânia, Brazil, **4** Centro de Malária e Doenças Tropicais, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Lisboa, Portugal

## Abstract

Morbidity and mortality caused by schistosomiasis are serious public health problems in developing countries. Because praziquantel is the only drug in therapeutic use, the risk of drug resistance is a concern. In the search for new schistosomicidal drugs, we performed a target-based chemogenomics screen of a dataset of 2,114 proteins to identify drugs that are approved for clinical use in humans that may be active against multiple life stages of *Schistosoma mansoni*. Each of these proteins was treated as a potential drug target, and its amino acid sequence was used to interrogate three databases: Therapeutic Target Database (TTD), DrugBank and STITCH. Predicted drug-target interactions were refined using a combination of approaches, including pairwise alignment, conservation state of functional regions and chemical space analysis. To validate our strategy, several drugs previously shown to be active against *Schistosoma* species were correctly predicted, such as clonazepam, auranofin, nifedipine, and artesunate. We were also able to identify 115 drugs that have not yet been experimentally tested against schistosomes and that require further assessment. Some examples are aprindine, gentamicin, clotrimazole, tetrabenazine, griseofulvin, and cinnarizine. In conclusion, we have developed a systematic and focused computer-aided approach to propose approved drugs that may warrant testing and/or serve as lead compounds for the design of new drugs against schistosomes.

**Citation:** Neves BJ, Braga RC, Bezerra JCB, Cravo PVL, Andrade CH (2015) *In Silico* Repositioning-Chemogenomics Strategy Identifies New Drugs with Potential Activity against Multiple Life Stages of *Schistosoma mansoni*. PLoS Negl Trop Dis 9(1): e3435. doi:10.1371/journal.pntd.0003435

**Editor:** Banchob Sripa, Khon Kaen University, Thailand

**Received** June 12, 2014; **Accepted** November 23, 2014; **Published** January 8, 2015



**Data Availability:** The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

**Funding:** BJN was supported by a fellowship from the Coordination for the Improvement of Higher Education Personnel (CAPES). This work has been funded by the National Counsel of Technological and Scientific Development (CNPq) and the State of Goias Research Foundation (FAPEG). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

* carolina@ufg.br

## Introduction

Schistosomiasis is one of the main neglected tropical diseases affecting humans. It is caused by flatworms of the genus *Schistosoma* (*S. mansoni*, *S. japonicum*, *S. haematobium*, *S. intercalatum* and *S. mekongi*) [1–3]. This parasitic disease ranks second only after malaria in terms of its public health importance [4] because of its chronic and debilitating characteristics that result in a substantial burden on human health [5,6]. Recent estimates suggest that more than 249 million people were infected in 78 endemic countries located in sub-Saharan Africa, the Middle East, the Caribbean, and South America, resulting in 200,000 deaths annually [7].

Schistosomes have complex life cycles that involve vertebrate (often a mammal) and invertebrate (aquatic snail) hosts, in which sexual and asexual reproductive phases occur, respectively. Mammalian definitive hosts are infected via skin penetration by cercariae, which lose their bifurcated tail and become schistosomula [8,9]. After 5–7 days, schistosomula migrate from the lungs to the hepatic portal system via the blood stream and transform into adult worms. Male and female worms pair in the hepatic portal system and migrate to the mesenteric veins (except *S. haematobium*, which migrates to the urogenital system) to lay nearly 300 eggs per day. These eggs either pass into the gut lumen to be voided in the faeces and continue the life cycle or pass through the mesenteric veins and lodge in the liver, where they can cause granulomatous changes and fibrosis, both of which are key contributors to schistosomiasis [8,10].

In the absence of a vaccine, praziquantel (PZQ) has been the drug of choice recommended by the World Health Organization for the treatment and control of all the major *Schistosoma* species in mass drug administration programs for almost three decades [11]. More recently, the use of artemisinin derivatives alone or in combination with PZQ for the treatment and prevention of schistosomiasis has shown encouraging results [12], but it is unlikely to represent an ideal stand-alone drug-based control strategy. Moreover, the suboptimal efficacy of PZQ against immature worms that are present in newly acquired infections [13] and the prospect of drug resistance indicate a need to identify

## Author Summary

Schistosomiasis is a neglected tropical disease caused by schistosome parasites that affects millions of people worldwide. The current reliance on a single drug (Praziquantel) for treatment and control of the disease calls for the urgent discovery of novel schistosomicidal agents. One approach that can expedite drug discovery is to find new uses for existing approved drugs, a practice known as drug repositioning. Currently, modern drug repositioning strategies entail the search for compounds that act on a specific target, often a protein known or suspected to be required for survival of the parasite. Drug repositioning approaches for schistosomiasis are now greatly facilitated by the availability of comprehensive schistosome genome data in user-friendly databases. Here, we report a drug repositioning computational strategy that involves identification of novel schistosomicidal drug candidates using similarity between schistosome proteins and known drug targets. Researchers can now use the list of predicted drugs as a basis for deciding which potential schistosomicidal candidates can be tested experimentally.


new schistosomicidal drugs active against multiple stages of parasite life cycle [5,14–16].

One approach that can expedite drug discovery process is to find new uses for existing approved compounds, a practice commonly known as drug repositioning or repurposing [17]. Drug repositioning has proved to be an efficient way of identifying new therapies against neglected tropical diseases. A recent example of a repositioned drug is miltefosine, a drug that was originally developed to treat breast cancer and is now used against visceral leishmaniasis [18,19]. In addition to saving money and time, an advantage of drug repositioning is that the existing drugs have already been scrutinized in terms of pharmacokinetic and toxicity parameters [20–22].

Over the last few decades, advances in computer technologies have resulted in useful tools to assist early drug discovery and development. In this context, the use of *in silico* tools can reduce the cost and the time required to select the most promising candidates for *in vitro* and *in vivo* assays [20]. Our laboratory has been developing and applying many computer-assisted drug discovery (CADD) strategies in the hope of discovering new drug candidates for neglected tropical diseases [23–43].

Several *in silico* chemogenomic studies have demonstrated that genome-wide gene expression data might also represent a useful resource for identifying drugs and drug target genes that can potentially be used for drug repositioning [44–47]. The ultimate goal of chemogenomics is to establish the molecular relationship(s) between ligands and drug targets. Therefore, various publicly available databases, such as Therapeutic Target Database (TTD) [48], DrugBank [49], and STITCH [50], which integrate information about gene/protein–drug–disease interactions, are useful resources to develop these strategies. Based on the concept that "similar targets have similar ligands", homology-based searching using these databases helps to identify compounds that may act on a target for which there are no known active compounds but that are related by homology to one or more targets for which active compounds are known [51,52]. In such context, *S. mansoni* targets with structural homology similar to known targets of approved drugs are more likely to be susceptible to inhibitors listed in the drug target and drug databases.

Recently, Protasio et al. [53] used transcriptomic sequencing from four time points in the *S. mansoni* life cycle to refine gene predictions and establish expression profiles in the parasite. Consequently, a high-resolution map of the temporal changes in the transcription of genes was produced for all intra-mammalian life cycle stages of *S. mansoni*. These data have been compiled into a searchable format within the SchistoDB (www.schistodb.net) and GeneDB (www.genedb.org) databases [53]. Transcription profiling and genome sequencing data provide important fundamental information to support further advances in schistosome research. In the present study, we used an *in silico* target-based chemogenomics strategy, integrating *S. mansoni* genomics data with drug-target database resources to predict new drugs with potential activity against multiple life stages of *S. mansoni*.

## Materials and Methods

### Compilation of the list of *S. mansoni* genes

The target-based chemogenomics screening was performed on a dataset containing 2,076 genes that are differentially expressed among the 24 hour schistosomula *vs*. adult life stages obtained from Protasio et al. [53]. We also obtained 38 *S. mansoni* genes from the TDR Targets database [54] using the target search tool. We searched for targets with "any form of validation", which included "genetic", "pharmacological", and "observed phenotypes" (S1 Table). We focused on searching for genes that are expressed in "24 h schistosomula *vs*. adult" because they are intra-mammalian stages. However, some of these genes are also expressed in other temporal life cycle stages, such as "cercariae *vs*. 3 hour schistosomula" and "3 h *vs*. 24 h schistosomula". These genes are considered promising targets for prophylactic drugs because they are involved in the penetration through the mammalian host's skin, host adaptation, and differentiation and growth of the parasite. Therefore, genes were grouped according to the following division: group I was composed of genes differentially transcribed between "24 h schistosomula *vs*. adult" and also between "cercariae and 3 h schistosomula"; group II was composed of genes differentially transcribed between "24 h schistosomula *vs*. adults" as well as between "3 h and 24 h schistosomula"; group III was composed of genes transcribed between "24 h schistosomula *vs*. adult"; and group IV was composed of genes transcribed concurrently in all the life cycle stages (Fig. 1A). Information for individual genes or gene products (primary amino acid sequence in FASTA format, target name, and biological process/es) was then retrieved from the GeneDB *S. mansoni* genome database [55]. We then verified the annotation of each single putative protein and corrected it, if necessary, according to the recent updated annotations in the GeneDB database. For convenience, the *in silico* target-based chemogenomic pipeline is presented in Fig. 2.

### Identification of putative drug targets using publicly available drug databases

Each predicted protein from *S. mansoni* was used to interrogate three different publicly available databases that provide detailed information on drugs and their targets: TTD [48], DrugBank [49], and STITCH [50]. The search strategy for DrugBank and TTD was based on the principle of homology, whereby each query (*S. mansoni* target) was compared for matches to known drug targets contained in each of these databases. In cases where homologous drug targets were identified, all proteins with an output expectation value (E-value) of $\leq 10^{-20}$ for DrugBank and TTD were listed as "acceptable targets". This E-value represents the number of hits with an alignment score "Z" or equal or better than "Z" that would be expected by chance when searching a database. The E-value is the expected number of times a

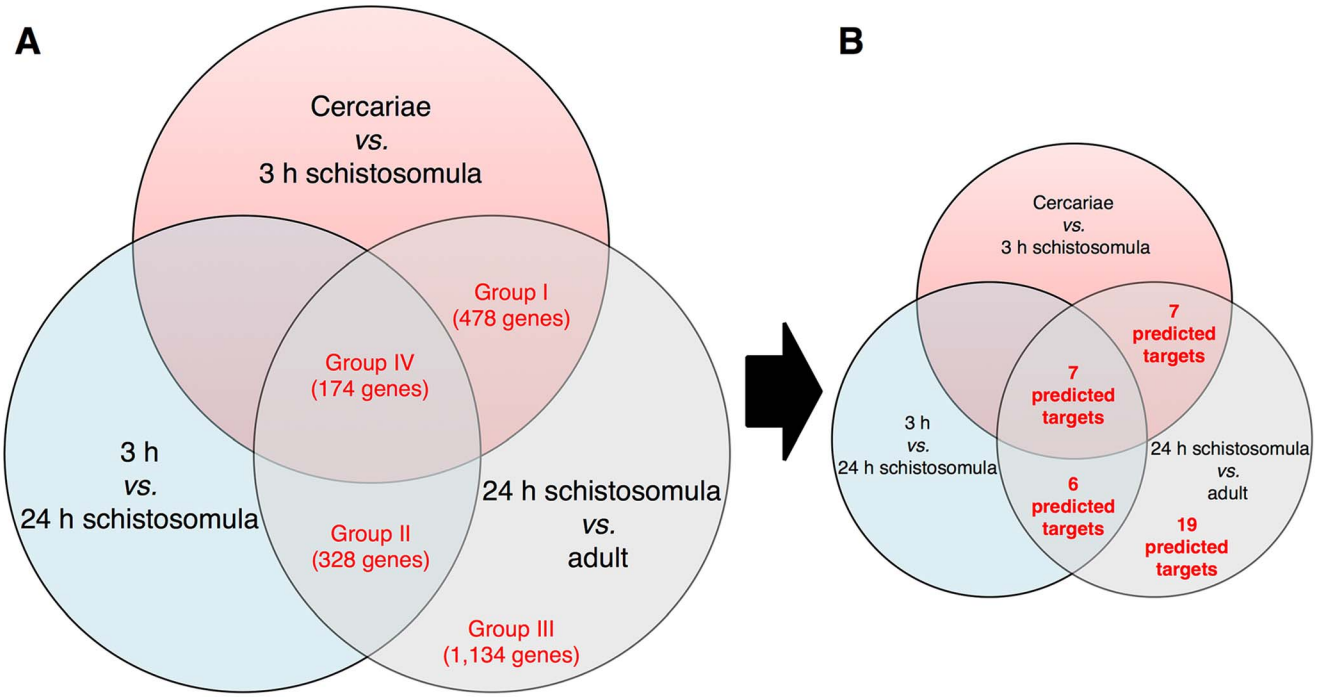



**Fig. 1. Distribution of genes from *S. mansoni* and predicted targets.** (**A**) Venn diagram represents the clustering of genes in four different groups according to temporal life stages of the parasite. (**B**) Rate of druggable *S. mansoni* targets identified for each group using the target-based chemogenomics approach.
doi:10.1371/journal.pntd.0003435.g001

homology match will occur at random in a given set of trials. However, the STITCH database integrates data from the literature and various databases of biological pathways, drug–target relationships, and binding affinities. The resultant network can be explored interactively or used as the basis for a confidence score ranging from 0–1. The confidence score is a set of high-confidence interactions between drugs and targets (i.e., proteins with shared Gene Ontology annotations) that is used as a reference for screening results. Therefore, in the case of the STITCH database, when significant matches were found, only targets with a score $\geq$0.8 were considered [50]. We filtered all predicted targets through inclusion in the list of only those proteins that were indicated to interact with approved drugs, excluding the nutraceutical class, as these compounds are unlikely to exhibit schistosomicidal activity [56].

## Pairwise alignment and comparison of functional regions

Predicted *S. mansoni* targets were aligned with their homologue drug targets using pairwise BLAST [57]. We considered the *S. mansoni* targets for further evaluation only in cases where there was $\geq$80% overlap with the corresponding drug target. This filtering step enhances the probability of both proteins sharing the same active site. Subsequently, we performed the sequence alignment and compared the functional regions among the approved drug targets and *S. mansoni* targets using the ConSurf server [58]. This procedure was used to estimate the conservation of active sites between the proteins and the preservation of affinity for the predicted schistosomicidal drugs. The ConSurf server [58] is a bioinformatic tool for estimating the evolutionary conservation of amino acid positions in a protein based on the phylogenetic relationships between homologous sequences. Therefore, the degree of conservation of the amino acids within the active site of each approved drug target was estimated using 150 homologues from other organisms with similar sequences in the UniProt database [59]. The sequences were clustered and those presenting high sequence similarity (>95%) were excluded using the algorithm CD-HIT [60] to filter out redundant sequences. In the same way, the sequences that shared less than the given identity cutoff of <35% were also ignored [58]. A multiple sequence alignment (MSA) of the homologous sequences was constructed using the MAFFT-L-INS-I method [61], and the phylogenetic tree was built using the neighbor-joining algorithm [62]. Position-specific conservation scores were computed using the empirical Bayesian method [63]. Next, the functional regions of each approved drug target were visually compared with the corresponding *S. mansoni* target. The results were classified as functional residues with high ($\geq$80%) or moderate conservation (60–79%). In cases where the conservation between functional residues was less than 59%, the putative targets were excluded from further analyses.

## List of drugs yet to be tested against *Schistosoma* species

We carried out a literature search using the PubMed, PubChem Bioassay database, and SciFinder engines to identify approved drugs that have not been evaluated against *Schistosoma* species by querying all predicted drugs previously identified. The PubChem BioAssay database reports the available biological screening results for the chemical compounds described in the PubChem database, providing searchable descriptions of each bioassay, including descriptions of the conditions and readouts specific to that screening procedure. SciFinder is a chemistry research application that provides access to the world's most comprehensive and authoritative sources of references, chemical substances, and reactions in chemistry and is updated daily by Chemical Abstracts Service. Our definition of "evaluation" embraces biochemical or *in vitro* and *in vivo* assessments of one or more life stages of

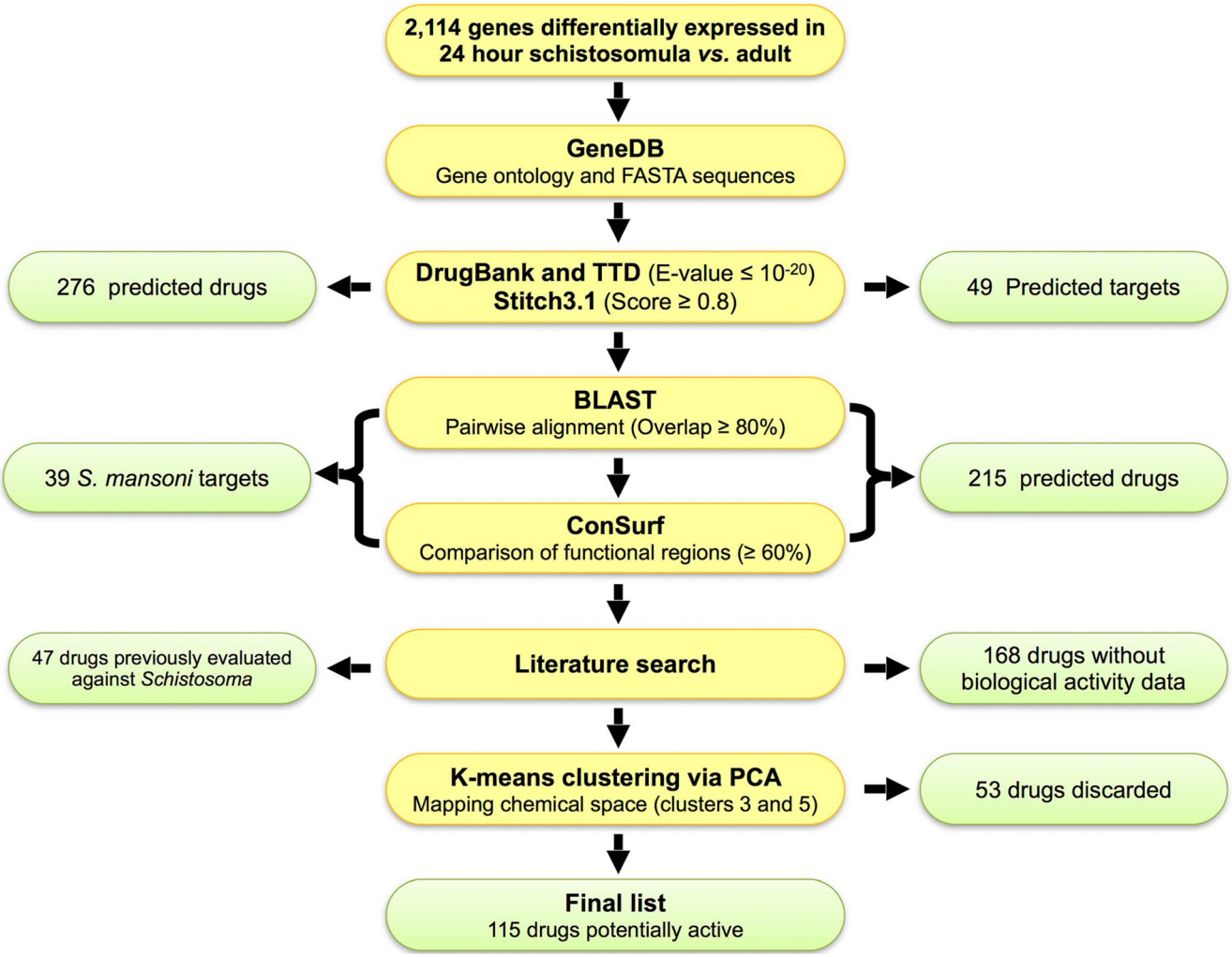


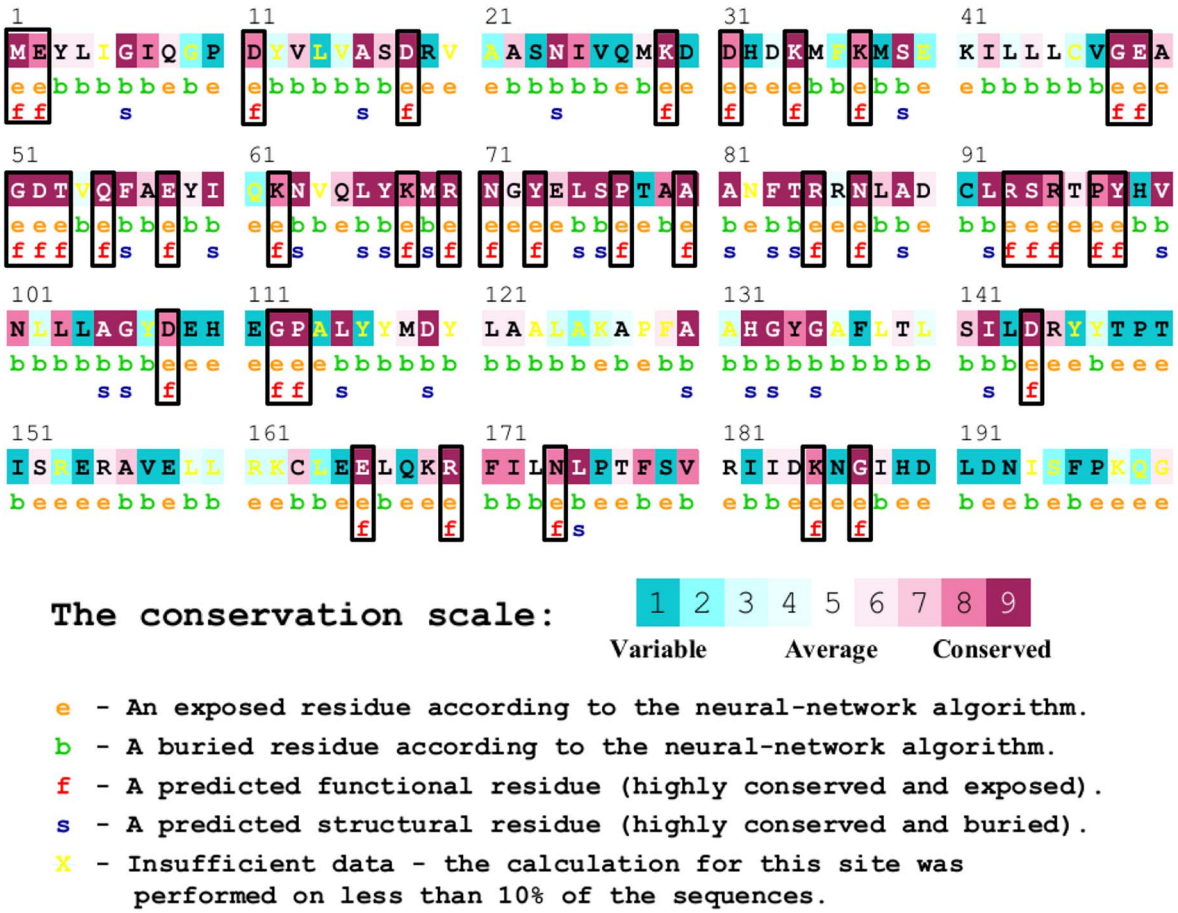

**Fig. 2. Flowchart summarizing the *in silico* repositioning chemogenomics strategy and corresponding results.** The green boxes represent the summarized results obtained at each stage of the study.
doi:10.1371/journal.pntd.0003435.g002

*Schistosoma* species. Therefore, if a given drug is noted as "not tested", it means that no publication record was found after either of the following search details: 1. ("drug name" [MeSH Terms] OR "drug name" [All Fields]) AND ("*Schistosoma*" [MeSH Terms] OR "*Schistosoma*" [All Fields]) and 2. ("drug name" [MeSH Terms] OR "drug name" [All Fields]) AND ("schistosomiasis" [MeSH Terms] OR "schistosomiasis" [All Fields]). It also might mean that the studies/assays retrieved were insufficiently informative to infer the potential usefulness of the drug as a schistosomicidal drug or lead compound.

## Chemical space analysis

We evaluated the "chemical space" of known active and inactive compounds against *Schistosoma*. The aim in using this strategy is to find whether predicted compounds share essential structural and physicochemical properties with schistosomicidal compounds. Initially, a dataset of active compounds with enzymatic, *in vitro*, and/or *in vivo* activity data for *Schistosoma* species was collected from the literature (S2 Table) using the PubMed and PubChem Bioassay databases [64]. Because of the differences in the experiments used for the biological activity evaluation, compounds of our database were considered to be active according to the specifications of each study or bioassay. In addition, a dataset of inactive compounds was compiled from a large dataset of non-inhibitors of the enzyme thioredoxin glutathione reductase of *S. mansoni* previously reported (PubChem Bioassay AID: 485364). These compounds were assumed to be inactive, because in the literature they were not reported to produce any schistosomicidal activity. All aforementioned chemical datasets were carefully curated and standardized according to the protocol proposed by Fourches et al. [65]. Structural normalization of specific chemotypes, such as aromatic and nitro groups, was performed using ChemAxon Standardizer (v.6.1.3, ChemAxon, Budapest, Hungary, http://www.chemaxon.com). Duplicates (i.e., identical compounds reported several times in the dataset) were identified using KSAR workflow (http://labmol.farmacia.ufg.br/ksar). The dataset is unbalanced, meaning that the size of the active and inactive classes does not match. Therefore, we used the algorithm *k*-Nearest Neighbor (*k*NN) developed in software R and the qsaR 1.5 package (http://labmol.farmacia.ufg.br/chemoinformatics) to equalize the number of compounds in different classes; this is referred to as "dataset balancing". The basic principle here is to evaluate the whole active dataset represented by the MACCS fingerprint matrix evaluating

**A**




**B**

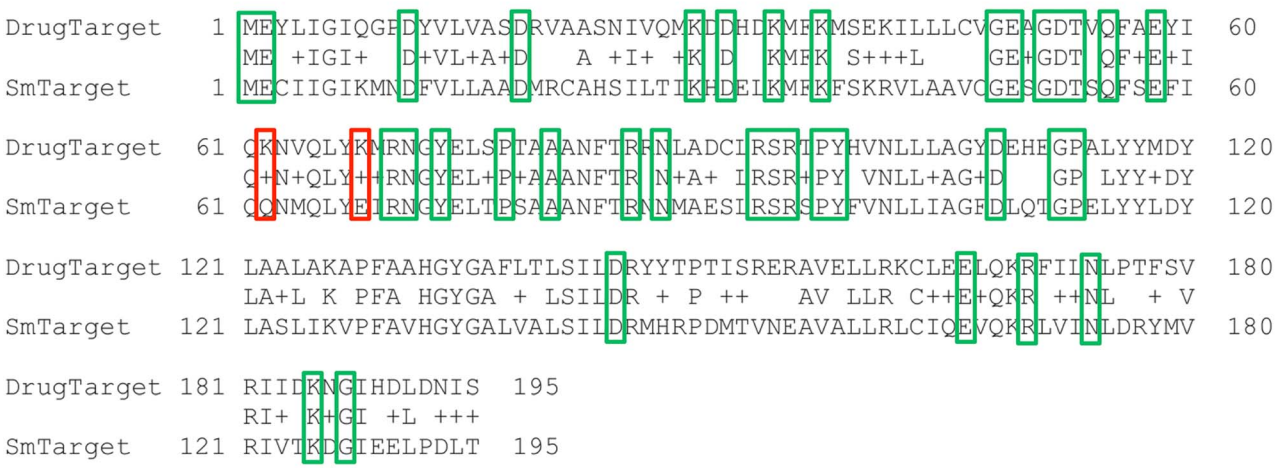



**Fig. 3. ConSurf analysis of the functional regions between an approved drug target and the corresponding *S. mansoni* target.** (**A**) ConSurf predictions demonstrated on human proteasome β type 2 (Gene ID: PSMB2), using 150 homologues obtained from the UNIPROT database. The sequence of the query protein is displayed with the residue conservation scores at each site color-coded onto it. The color-coding bar shows the coloring scheme; conserved amino acids are colored bordeaux, residues of average conservation are white, and variable amino acids are turquoise. The residues of the query sequence are numbered starting from 1 to 199. The first row below the sequence lists the predicted burial status of the site (*i.e.* ''b''– buried versus ''e'' – exposed). The second row indicates residues predicted to be structurally and functionally important: ''s'' and ''f'', respectively. (**B**) Analysis of the functional regions conserved with the corresponding *S. mansoni* proteasome β type 2. The green rectangles represent the conserved functional residues and red rectangles represent non-conserved functional residues.
doi:10.1371/journal.pntd.0003435.g003

the Euclidean distance to the MACCS fingerprint of each inactive compound. The compounds were reordered by nearest *k*-distance of the active compounds. Thereafter, a set representing 39 descriptors accounting for physicochemical properties were calculated using RDKit 2.4.0 [66]. The descriptor matrix was normalized, and low variance descriptors (variance upper bound set to 0.0) were removed before generating the model. The chemical space analysis of predicted drugs was generated using *k*-means clustering space using Principal Component Analysis (PCA) and employing the software R v.3.0.3 [67]. All steps of dataset balancing, processing, and chemical space analysis were implemented in R and KNIME, a graphical user interface that allows the assembly of nodes for modeling, data analysis, and visualization (S1 Fig.).

## Results

### Compilation of the genes list

The dataset of *S. mansoni* genes was compiled from the Protasio et al. study and the TDR Targets database, totalling 2,114 genes (S1 Table). We focused on searching drugs with potential activity in schistosomula and adult life cycle stages, which are all intra-mammalian stages. For this reason, genes that do not have a differential transcription between 24 h schistosomula and adults were not considered. Some of the genes differentially transcribed between these stages were also transcribed in other life cycle stages, including the cercariae; therefore, some of these genes are considered promising targets for drugs, as they are expected to be involved in penetration through mammalian host skin, adaptation, differentiation, and growth. The 2,114 genes were clustered in four main groups (I–IV) according to transcription in each life cycle stage (Fig. 1A). Totals of 478, 328, 1,134, and 174 transcribed genes were identified.

### Identification of putative drug targets using publicly available drug databases

The information about individual genes (primary amino acid sequence in FASTA format, target name, and biological process) was retrieved from the GeneDB *S. mansoni* genome database. Based on the FASTA sequence information, we predicted schistosomicidal drugs using the sequence similarity screening in three databases (DrugBank, STITCH 3.1, and TTD). In this step, numerical statistical probability parameters (E-value $\leq 10^{-20}$ or a score $\geq 0.8$) were adopted to provide high confidence for the data. We decided to use all three databases because each of them may contain different drug-target datasets and, consequently, the probability of targets and their drugs being missed is reduced. This analysis predicted 49 targets associated with 276 approved drugs (S3 Table).

### Pairwise alignment and comparison of functional regions

Pairwise sequence alignment was used to compare the *S. mansoni* targets previously identified with their approved homologous drug targets using BLAST alignment. Ten targets had less than an 80% overlap with their corresponding approved targets and were excluded from further analyses due to the improbability of both proteins sharing the same active site. Next, we performed sequence alignments and comparisons of functional regions for approved drug targets and predicted *S. mansoni* targets. This step allowed the identification of functionally relevant features and conserved residues necessary for catalysis and residues critical for binding. Fig. 3 shows an example of the ConSurf analysis of the functional regions between an approved drug target (human proteasome β type 2) and the corresponding *S. mansoni* target. Fig. 3A shows the predictions demonstrated on human proteasome β type 2 (Gene ID: PSMB2) using 150 homologues obtained from the UniProt database. This analysis revealed that 38 residues were predicted to be functionally important to the catalytic activity of the human enzyme. The functional regions of the human proteasome β type 2 were aligned to the respective *S. mansoni* orthologue. This comparison demonstrated that the active site predicted for *S. mansoni* proteasome β type 2 is conserved when compared to functional regions of its respective human target (Fig. 3B).

### Compilation of the "predicted targets list"

After running each of the *S. mansoni* protein sequences through the three databases, all proteins with negative results (negative hits) were excluded from further analyses, whilst predicted targets from each database were compiled into a single Excel file, hereafter called the "predicted targets list" (S3 Table). Each positive hit was examined further using BLAST pairwise alignment and the ConSurf server. This strategy resulted in a list of 39 potential druggable *S. mansoni* targets (~1.8% of the interrogated targets) that could interact with 215 approved drugs. The DrugBank, STICH 3.1 and TTD databases exclusively predicted 120 (56.0%), 6 (2.8%), and 18 (8.3%) of the approved drugs, respectively, whilst the remaining 71(32.9%) drugs were predicted by two or three of these databases. Detailed information about the predicted targets and their associated drugs are provided in S3 Table. The distribution of the 39 identified *S. mansoni* targets according to their expression group is shown in Fig. 1B. We found that 19 (48.7%) of the predicted *S. mansoni* targets are in group III, 7 (17.9%) are in group I, 6 (15.4%) are in group II, and 7 (17.9%) are in group IV (Fig. 1B and S3 Table).

### List of drugs yet to be tested against *Schistosoma* species

To investigate which of the predicted drugs have already been tested against *Schistosoma* species, we undertook a literature search of PubMed, PubChem Bioassay, and SciFinder. A total of 22 druggable targets associated with 47 drugs whose activity has been previously evaluated against *Schistosoma* were identified. Examples of some of these drugs and their corresponding targets are given in Table 1. Accordingly, we were confident that our chemogenomic strategy for identifying new schistosomicidal drugs is valid. Consequently, we predicted 168 drugs to be active against 33 druggable targets that have not yet been experimentally tested against schistosomes or that have not yet required further studies. The results are summarized in Fig. 2. The complete list of predicted drugs, their targets, alignment parameters, and conservation of the functional regions is given in S3 Table.

### Chemical space analysis

Finally, we used chemical space analysis to map the 168 predicted drugs in a multidimensional space using physicochemical descriptors for a dataset of active and inactive compounds against *Schistosoma* reported in the literature. The chemical space analysis is useful to refine the results and select the most promising drugs that share essential structural and physicochemical properties with schistosomicidal compounds. Because the original dataset was unbalanced, containing 355 active compounds *vs.* 331,228 inactive compounds (extracted from 101 bibliographic references, including 87 articles from PubMed and 14 from PubChem bioassays), it was not appropriate to build multivariate models [33]. For this reason, a balanced dataset containing 696 chemical structures (355 active compounds *vs.* 341 inactive compounds, S2 Table) was generated using the *k*NN method.

**Table 1.** Examples of approved drugs and their potential *S. mansoni* targets that were previously reported on the literature, correctly identified by our target-based chemogenomics strategy.

| Drug | Drug class | *S. mansoni* target (Target ID) | Biological process | Functional regions (%) | Activity data |
|---|---|---|---|---|---|
| methotrexate | antineoplasic | dihydrofolate reductase (Smp_175230.1) | synthesis of nucleic acid precursors | moderate conservation (78%) | $IC_{50}$ = 7 nM enzymatic assay [109] |
| flubendazole | antihelminthic | tubulin α chain, putative (Smp_016780.1) | cytoskeleton formation | high conservation (100%) | 79.5% of reduction of adult worms 25 days after infection (100 mg/kg) [110] |
| clonazepam | hypnotic and sedative | peripheral type benzodiazepine receptor (Smp_102510.1) | receptor activity | high conservation (83%) | $IC_{50}$ = 10 μM in adult worms [111] |
| auranofin | antirheumatic | thioredoxin glutathione reductase (Smp_048430.1) | cell redox homeostasis | high conservation (92%) | $IC_{50}$ = 0.7 nM enzymatic assay [112] |
| artesunate | antimalarial | $Ca^{2+}$ transporting ATPase (Smp_136710.1) | $Ca^{2+}$ homeostasis | high conservation (99%) | 97.1% reduction of adult worms 30 days after infection (30 mg/kg) [113] |
| nifedipine | antihypertensive | voltage-dependent calcium channel (Smp_159990.1) | $Ca^{2+}$ homeostasis | high conservation (88%) | Loss of motility and muscle contraction in adult worms (1 mg/kg) [82] |

doi:10.1371/journal.pntd.0003435.t001

The chemical space mapping was performed using *k*-means clustering via PCA using 39 physicochemical descriptors (Fig. 4). According to the PCA, the first and second principal components (PCs) explained 70.9% of the total variability of data and were categorized into five main clusters. Most of the compounds predicted to be active were located in regions marked in purple and green at the upper right corner of the score plot, totaling 215 active compounds (91.8%) and 19 inactive compounds (8.2%) (Fig. 4, clusters 3 and 5, respectively). Moreover, the inactive compounds were mostly delimited into the blue region containing 291 inactive compounds (74.4%) and 100 active compounds (25.6%) (Fig. 4, cluster 4). The red regions located in the center and at the upper right corner were flagged as inconclusive, as they contained similar proportions of both classes of compounds (Fig. 4, clusters 1 and 2, respectively). Remarkably, 115 drugs predicted by the proposed methodology are inside the overlapping area of the chemical space of the active compounds (Fig. 4, clusters 3 and 5) and are more likely to be active, whereas only 53 drugs were inside the overlapping area of the inactive compounds and the inconclusive clusters (Fig. 4, clusters 4, 1 and 2). Therefore, a "repurposing drug" located in the cluster regions 3 and 5 has a high probability (92%) of being active.

## Discussion

The main goal of this study was to identify drugs approved for clinical use in humans that may have potential schistosomicidal activity by performing a search of publicly available drug/target databases. However, since most target databases are only starting to be established, the predicted *S. mansoni* targets are not yet scored for druggability. The druggability concept adds a structural dimension and evaluates the likelihood that small drug-like molecules can bind a given target with sufficient potency to alter its activity [68–70]. Druggability is related to many factors, including the size of targets, the presence of pockets, and the overall charge and hydrophobicity of the interaction surface [68]. A number of factors were considered in this study in order to provide both confidence for the data generated and a solid basis from which to predict the druggability of individual *S. mansoni* targets. The predicted *S. mansoni* targets were considered druggable if they presented a sequence overlap ≥80%, a score ≥0.8, or an E-value $\leq 10^{-20}$ in relation to their predicted homologous targets and the high or moderate conservation of the functional regions. The overlapping sequences and analysis of functional regions among approved drug targets and *S. mansoni* targets revealed the importance of each position for the function of the predicted protein and the possible preservation of affinity for the predicted drug. Following this precondition, we were able to identify 168 drugs with the potential to inhibit their targets known to be transcribed in multiple life stages of *S. mansoni*. Moreover, in validation of the proposed chemogenomic strategy, several drugs previously demonstrated to be active against *Schistosoma* species in experimental assays were predicted by our methodology (Table 1 and S3 Table). Consequently, we were confident that our strategy for predicting schistosomicidal drugs is useful.

Additionally, we also evaluated the structural and physicochemical properties of known active and inactive schistosomicidal compounds to map the chemical space that is accessed by the 168 predicted drugs using the chemogenomics strategy. Chemical space is a term that is commonly used in place of "multi-dimensional descriptor space", which is a region defined by a particular choice of descriptors encompassing all the possible compounds that could be mapped to the coordinates of this multi-dimensional space [71,72]. This concept is closely related to the notion of chemical diversity. The diversity of a chemical library is a quantitative description of how different these compounds are from each other, with similar compounds and similar biological activities falling in the same chemical space region [73]. For this, we used *k*-means clustering via PCA, a method to compute the position of every compound in a two-dimensional coordinate system based on a set of computed properties. The PCA reduces high-dimensional data into a lower-dimensional space, thus making it more manageable and comprehensible by extracting essential information [74,75]. Indeed, PCA transforms the original measured variables, such as physicochemical descriptors, into new uncorrelated variables called PCs, which are a linear combination of the original measured variables. As a result, we found that 115 drugs predicted by this chemogenomics strategy are inside the chemical space of active schistosomicidal compounds, yielding a higher degree of confidence in the predictions.

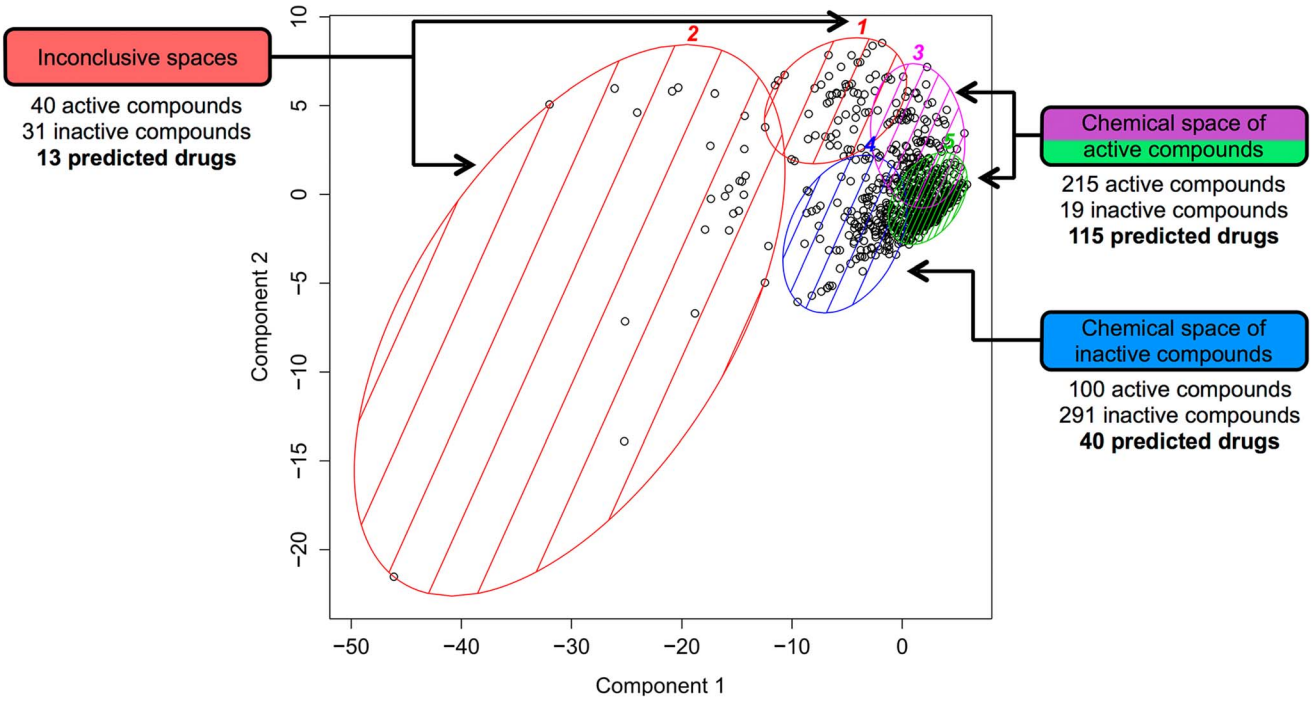



**Fig. 4. Chemical space analysis of the schistosomicidal compounds.** Purple and green regions (clusters 3 and 5, respectively) represent the chemical space of active compounds. 115 drugs predicted by our strategy are inside this area and are more likely to be active. The blue region (cluster 4) represents the chemical space of inactive compounds and red region (clusters 1 and 2) are inconclusive spaces. 53 predicted drugs are inside this area and are more likely to be inactive. The first and second components explain 70.9% of the total variability.
doi:10.1371/journal.pntd.0003435.g004

A previous chemogenomics screen in *S. mansoni* described by Caffrey et al. [47] identified 35 potential druggable targets for further investigation in drug discovery programs, showing the value of *in silico* approaches for drug discovery for schistosomiasis. Interestingly, only one drug target identified in that study (methionine aminopeptidase: Target ID = Smp_011120.1) is present amongst our predicted targets, which is likely because the studies differ significantly in their methodology. In addition, Berriman et al. [53] also reported an *in silico* approach to predict schistosomicidal drugs using the StARlite database and conservative parameters ($\leq$50% sequential identity and $\geq$80% overlap) for target exclusion. Only eight predicted drugs identified in our study (carbidopa, colchicine, dasatinib, deserpine, mycophenolate mofetil, mycophenolic acid, reserpine and ribavirin) are overlapping in both studies. This small number might relate to the different gene datasets, different databases, and different parameters for conservation filtering used [76].

One of our goals was to predict targets that control the muscle function and motility of the parasite. Schistosomes depend on their muscular systems for motility, penetration of host skin (cercariae), and migration (schistosomules). Additionally, schistosomes use their muscular system for pairing and mating and reproductive, digestive, and excretory processes. These responses are essential for the survival of the parasite. These behaviors require precise coordination not only of the musculature that enables the movement but also of the responses regulated by the neurotransmitters needed for successful motility [77]. Drugs that disrupt one or more of these motility functions would be expected to interfere with the normal life of the parasite and would consequently cause its death. It is noteworthy that PZQ, the drug of choice to treat schistosomiasis, disrupts the muscle function and causes paralysis of the worm [78]. Currently, the gold standard test for measuring drug activity for *S. mansoni* is the *in vitro* assessment of worm motility, measured visually through microscopy [79]. Furthermore, we suggest that 61 drugs have potential activity against the muscle function and motility of the parasite because they were predicted to interact with 4 neurotransmitter transporters (nicotinic acetylcholine receptor $\alpha$ subunit, Target ID = Smp_031680.1; $Na^+/Cl^-$ dependent transporter, Target ID = Smp_193800.1; vesicular amine transporter, Target ID = Smp_121920.1; and $Na^+/Cl^-$ dependent neurotransmitter transporter, and Target ID = Smp_160360.1); 2 ion channels ($Ca^{2+}$ transporting ATPase, Target ID = Smp_136710.1 and voltage-dependent $Ca^{2+}$ channel, Target ID = Smp_159990.1); and 2 indirectly related proteins (calmodulin, Target ID = Smp_026560.2; and acetylcholinesterase, Target ID = Smp_154600.1) (S3 Table). These drugs are attractive when related to the study reported by Smout et al. [79] that can simply and objectively assess anthelmintic activity by measuring parasite motility in real time in a fully automated high-throughput screening.

Another important aspect considered in this study has to do with intellectual property protection of the potential schistosomicidal drugs predicted by the proposed strategy, particularly for those drugs that are off-patents. Pre-existing patents could impede the commercialization of schistosomicidal repositioned drugs and make them uneconomical, given that schistosomiasis predominantly affects poor populations in low- and middle-income countries. Therefore, an extensive search was done to collect the patent information (expiration date) of the predicted drugs using the European Patent Office database, Google Patents, and SciFinder. We found that 80.9% of the predicted drugs are off-patents (S3 Table). Last, we refer specifically to six drugs that we suggest are candidates for pre-clinical (*in vitro* and *in vivo*) studies

**Table 2.** Examples of potential schistosomicidal drugs and their potential targets revealed in this study.

| Drug | Therapeutic class | *S. mansoni* target (Target ID) | Biological process | Functional regions (%) |
|---|---|---|---|---|
| cinnarizine | anti-allergic | voltage-dependent calcium channel (Smp_159990.1) | $Ca^{2+}$ homeostasis | high conservation (88%) |
| griseofulvin | antifungal | tubulin β chain, putative (Smp_192110.1) | cytoskeleton formation | high conservation (99%) |
| tetrabenazine | antipsychotic | vesicular amine transporter (Smp_121920.1) | neurotransmitter transport | moderate conservation (73%) |
| clotrimazole | antifungal | $Ca^{2+}$-activated $K^+$ channel (Smp_161450.1) | $K^+$ homeostasis | high conservation (80%) |
| gentamicin | antibacterial | heat shock protein 73 (Smp_106930.1) | protein folding | high conservation (98%) |
| aprindine | antiarrhythmic | calmodulin, putative (Smp_134500.1) | $Ca^{2+}$ binding messenger | high conservation (100%) |

doi:10.1371/journal.pntd.0003435.t002

(Table 2). The remaining drugs were not discussed in detail, because we found that pharmacokinetic and toxicity properties may render them less suitable as schistosomicidal drugs than the above chemicals. For example, ixabepilone and pralatrexate are antineoplastic drugs and thus might cause severe toxicity in humans. However, we consider that all predicted drugs identified in this study are attractive for further analysis.

Cinnarizine is an antagonist of the histamine H1 receptor used for the control of nausea due to motion sickness. This drug is also considered a potent dilator of peripheral vessels because of its ability to block $Ca^{2+}$ channels [80,81]. The present study suggests that cinnarizine may also be able to inhibit the *S. mansoni* voltage-dependent L-type calcium channel subunit alpha-1S (Target ID = Smp_159990.1; E-value = 0; functional regions = 88% conservation), which is homologous to the human enzyme. Curiously, PZQ is considered a $Ca^{2+}$ channel activator, which would allow more $Ca^{2+}$ channels to open and lead to the disruption of normal intracellular $Ca^{2+}$ levels. After exposure to PZQ, diverse effects become apparent in adult worms, such as muscular contraction and tegumental disruption, which subsequently leads to the exposure of parasite antigens on the worm's surface [78]. Despite the activator effect of PZQ on $Ca^{2+}$ channels, a recent study demonstrated that treatment with nifedipine, a $Ca^{2+}$ channel blocker, resulted in antischistosomal activity against schistosomula and significantly reduced their viability. Adult worms were also affected by nifedipine and exhibited impaired motility, several lesions on the tegument, intense contractility, and the reduction of egg deposition [82].

Griseofulvin is a fungistatic drug that is orally administered in the treatment of cutaneous mycoses. It was originally biosynthesized from *Penicillium griseofulvum* in 1939 [83], but its *in vivo* efficacy was first demonstrated only in 1958. Results from the present study suggest that griseofulvin might also be able to inhibit the *S. mansoni* tubulin-β chain (Target ID = Smp_192110.1; E-value = 0; score = 0.8; functional regions = 99% conservation), which is homologous to the *Candida albicans* protein and is expected to be involved in cytoskeleton formation. Griseofulvin is able to inhibit the growth of fungal, plant, and mammalian cells by blocking the cells at the G2/M phase of the cell cycle [84,85]. In fungi, griseofulvin deteriorates spindle and cytoplasmic microtubules, resulting in nuclear and mitotic abnormalities followed by distortions in hyphal morphology. Microtubules form a highly organized cellular cytoskeleton in eukaryotic cells, and their aggregation–disaggregation plays a key role in cell morphology and growth [86]. The concentration of griseofulvin required to deteriorate the spindle and cytoplasmic microtubules of fungal cells is much lower than that required to inhibit normal healthy mammalian cells due to its higher affinity for fungal tubulin as compared to mammalian tubulin [87–90]. Furthermore, griseofulvin selectively induces apoptosis in several cancer cell lines, sparing the normal healthy cells [85,91]. Therefore, we consider that griseofulvin has low toxicity against normal healthy cells, which makes it highly appropriate for clinical use.

Tetrabenazine is a reversible human vesicular monoamine transporter type 2 inhibitor. It acts within the basal ganglia and promotes the depletion of monoamine neurotransmitters, such as serotonin and dopamine, in nerve terminals [92]. In this study, we suggest that tetrabenazine might also be able to block the vesicular amine transporter of *S. mansoni* (Target ID = Smp_121920.1; E-value = $3.06^{-130}$; functional regions = 73% conservation). *S. mansoni* also has a sophisticated nervous system that includes both central and peripheral elements and employs a wide range of neurotransmitter transporters. Among them, there are vesicular amine transporters that are normally responsible for the uptake of cytosolic biogenic amines into synaptic vesicles. Serotonin and dopamine are largely responsible for neuromuscular signaling in the parasite, and therefore, the carriers are expected to be important components of the worm's motor control system [77,93]. It is worth noting that amine transport inhibitors have been shown to have strong effects on the parasite, as demonstrated in two medium-throughput drug screens of *S. mansoni* [94,95].

Clotrimazole is an antifungal drug commonly used to treat yeast infections of the vagina, mouth, and skin, such as athlete's foot, jock itch, and body ringworm. This drug is a potent inhibitor of 14-α demethylase, resulting in increased cellular permeability. It is also capable of inhibiting the movement of $Ca^{2+}$ and $K^+$ ions across the cell membrane by blocking the $Ca^{2+}$-activated $K^+$ channel [96]. The present study suggests that clotrimazole might also be able to block the *S. mansoni* $Ca^{2+}$-activated $K^+$ channel (Target ID = Smp_161450.1; E-value = $1.09^{-48}$; functional regions = 80% conservation). The $Ca^{2+}$-activated $K^+$ channel is essential for maintaining the membrane in a hyperpolarized state, thereby regulating neuronal excitability, smooth muscle contraction, and secretion [97,98]. Thus, the blocking of $Ca^{2+}$-activated $K^+$ channels in the muscle membranes of *S. mansoni* could be intimately involved in the dysfunction of rhythmic muscle activity. Due to its own chemical nature, clotrimazole is not well absorbed from the gastrointestinal tract. However, since clotrimazole is commercially available in powder form, it may be tested directly after dilution in an appropriate vehicle in experimental models of schistosomiasis with administration via a route other than oral, such as intra-peritoneal.

Gentamicin is an aminoglycoside drug composed of a mixture of related gentamicin components and is used to treat many types of bacterial infections, particularly those caused by gram-negative organisms. This drug binds the 30S subunit to the 16S ribosomal RNA (rRNA) of bacteria, but its affinity to the heat shock protein (HSP) 73 has also been well established [99]. We found that

gentamicin might be able to interfere with the heat shock protein 70 of *S. mansoni* (Target ID = Smp_106930.1; E-value = 0; functional regions = 98% conservation), a homologue of the human HSP73. HSPs are a family of proteins involved in basic life-protecting mechanisms against harmful extracellular effects, primarily heat shock response. Normally, the expression of these proteins is increased in response to cellular adaptation to high temperatures [100]. Among the HSP family, HSP70 is considered the most predominantly conserved with intracellular chaperone and extracellular immunoregulatory functions [101]. In *S. mansoni*, it is well established that HSP70 is involved in protein re-folding and the chaperone function as an adaptive response to the rapid temperature rise between fresh water ($\sim 28^{\circ}C$), in which the cercariae are found, and the warmer mammalian host ($\sim 37^{\circ}C$) [53].

Last, we refer to aprindine as a candidate. This is possibly one of the least obvious drugs to hold schistosomicidal activity because it is not an anti-infective agent but rather an anti-arrhythmic drug. An interesting fact is that aprindine has a binding affinity to calmodulin [102]. Thus, we suggest that aprindine may also be able to inhibit the *S. mansoni* calmodulin (Target ID = Smp_026560.2; E-value = $4.45^{-81}$; functional regions = 100% conservation). Calmodulin is the primary sensor of intracellular $Ca^{2+}$ levels that binds to and regulates a number of diverse target proteins involved in different functions, such as muscle contraction, apoptosis, and the immune response [103]. In *S. mansoni*, selective calmodulin inhibitors are known to disrupt egg hatching or cause miracidia to become vesiculated and die without undergoing transformation to the sporocyst stage [104,105]. $Ca^{2+}$ mobilization also plays a role in the cercarial penetration processes, possibly by calcium regulation of protease activities during infection [106,107]. It is important to mention that $Ca^{2+}$ ions are second messengers that are crucial for many biological functions, including muscle contraction, metabolism, and cell motility [108]. Importantly, visual inspection of the chemical structures allowed us to discover that aprindine is chemically similar (two aromatic centers and one aliphatic amine) to tricyclic drugs, a chemical class of the psychoactive drugs overactive against schistosomula stages [94].

Besides the drugs highlighted above, 109 other drugs are predicted to be active against *S. mansoni*. In all cases, we considered the numerical parameters (overlap, conserved functional regions, E-value, and score) for target homology sufficiently significant to infer drug predictions. Moreover, we verified that these drugs are "inside the chemical space" of active schistosomicidal compounds, making the predictions more reliable. Therefore, all 115 predicted drugs are candidates for drug repositioning and might be used as starting points for further *in vitro* and *in vivo* studies and schistosomicidal drug design because they are privileged structures and have established pharmacokinetic and toxicity profiles considered suitable for lead optimization.

## Conclusions

We used an *in silico* drug repositioning strategy based on the concept that "similar targets have similar ligands" to compile a list of drugs with potential activity against schistosomes. In doing so, we predicted 115 such compounds that we suggest justify evaluation as schistosomicidal drugs. We recognize that the activity of such compounds might be affected by appropriate chemical affinity with their predicted target. However, in addition to previous strategies, we used the criterion of conservation of functional residues among *S. mansoni* and its homologous targets and investigated the chemical space of known schistosomicidal compounds to further increase confidence in our predictions. Primary *in vitro* screens with these drugs might provide insights into their schistosomicidal activity. If promising activities are discovered, they could constitute important starting points for lead identification and optimization.

## Supporting Information

**S1 Fig Workflow of the chemical space analysis using the KNIME graphical user interface.** All steps of dataset balancing, processing, and chemical space analysis were implemented in R and KNIME, a graphical user interface that allows assembly of nodes for modeling, data analysis, and visualization.
(TIF)

**S1 Table List of *S. mansoni* genes with deferential expression in multiple life stages.**
(XLSX)

**S2 Table Dataset of known schistosomicidal compounds, with SMILES line notation for describing chemical structures, references, molecular descriptors and cluster analysis results.**
(XLSX)

**S3 Table List of predicted schistosomicidal drugs and target information.**
(XLSX)

## Author Contributions

Conceived and designed the experiments: BJN JCBB PVLC CHA. Performed the experiments: BJN RCB. Analyzed the data: BJN RCB PVLC CHA. Contributed reagents/materials/analysis tools: RCB JCBB PVLC CHA. Wrote the paper: BJN RCB PVLC CHA.

## References


1. Gryseels B, Polman K, Clerinx J, Kestens L (2006) Human schistosomiasis. Lancet 368: 1106–1118.
2. King CH, Dickman K, Tisch DJ (2005) Reassessment of the cost of chronic helmintic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis. Lancet 365: 1561–1569.
3. Colley DG, Bustinduy AL, Secor WE, King CH (2014) Human schistosomiasis. Lancet 383: 2253–2264. doi:10.1016/S0140-6736(13)61949-2.
4. Yang Z, Grinchuk V, Smith A, Qin B, Bohl JA, et al. (2013) Parasitic nematode-induced modulation of body weight and associated metabolic dysfunction in mouse models of obesity. Infect Immun 81: 1905–1914. doi:10.1128/IAI.00053-13.
5. Melman SD, Steinauer ML, Cunningham C, Kubatko LS, Mwangi IN, et al. (2009) Reduced susceptibility to praziquantel among naturally occurring Kenyan isolates of Schistosoma mansoni. PLoS Negl Trop Dis 3: e504. doi:10.1371/journal.pntd.0000504.
6. Rollinson D, Knopp S, Levitz S, Stothard JR, Tchuem Tchuenté L-A, et al. (2013) Time to set the agenda for schistosomiasis elimination. Acta Trop 128: 423–440.
7. WHO (2014) Schistosomiasis. Fact sheet: http://www.who.int/mediacentre/factsheets/fs115/en.
8. Gobert GN, Chai M, McManus DP (2007) Biology of the schistosome lung-stage schistosomulum. Parasitology 134: 453–460.
9. He Y-X, Salafsky B, Ramaswamy K (2005) Comparison of skin invasion among three major species of Schistosoma. Trends Parasitol 21: 201–203.
10. Burke ML, Jones MK, Gobert GN, Li YS, Ellis MK, et al. (2009) Immunopathogenesis of human schistosomiasis. Parasite Immunol 31: 163–176.

11. World Health Organization (2012) Progress report 2001–2011 and strategic plan 2012–2020. Geneva, Switzerland.
12. Pérez del Villar L, Burguillo FJ, López-Abán J, Muro A (2012) Systematic review and meta-analysis of artemisinin based therapies for the treatment and prevention of schistosomiasis. PLoS One 7: e45867.
13. Gönnert R, Andrews P (1977) Praziquantel, a new board-spectrum antischistosomal agent. Z Parasitenkd 52: 129–150.
14. Ismail M, Metwally A, Farghaly A, Bruce J, Tao LF, et al. (1996) Characterization of isolates of Schistosoma mansoni from Egyptian villagers that tolerate high doses of praziquantel. Am J Trop Med Hyg 55: 214–218.
15. Wang W, Wang L, Liang Y (2012) Susceptibility or resistance of praziquantel in human schistosomiasis: a review. Parasitol Res 111: 1871–1877. doi:10.1007/s00436-012-3151-z.
16. Gryseels B, Mbaye A, De Vlas SJ, Stelma FF, Guissé F, et al. (2001) Are poor responses to praziquantel for the treatment of Schistosoma mansoni infections in Senegal due to resistance? An overview of the evidence. Trop Med Int Heal 6: 864–873.
17. Chantree P, Phatsara M, Meemon K, Chaichanasak P, Changklungmoa N, et al. (2013) Vaccine potential of recombinant cathepsin B against Fasciola gigantica. Exp Parasitol 135: 102–109. doi:10.1016/j.exppara.2013.06.010.
18. Food and Drug Administration (2014) FDA approvals for the first 6 months of 2014. Nat Rev Drug Discov 13: 565–565. doi:10.1038/nrd4409.
19. Sundar S, Jha TK, Thakur CP, Engel J, Sindermann H, et al. (2002) Oral miltefosine for Indian visceral leishmaniasis. N Engl J Med 347: 1739–1746.
20. Ekins S, Williams AJ, Krasowski MD, Freundlich JS (2011) In silico repositioning of approved drugs for rare and neglected diseases. Drug Discov Today 16: 298–310.
21. Ma D-L, Chan DS-H, Leung C-H (2013) Drug repositioning by structure-based virtual screening. Chem Soc Rev 42: 2130–2141.
22. Li YY, Jones SJ (2012) Drug repositioning for personalized medicine. Genome Med 4: 27.
23. Andrade CH, Salum L, Pasqualoto KFM, Ferreira EI, Andricopulo AD (2008) Three-Dimensional Quantitative Structure-Activity Relationships for a Large Series of Potent Antitubercular Agents. Lett Drug Des Discov 5: 377–387. doi:10.2174/157018008785777289.
24. Andrade CH, Pasqualoto KFM, Zaim MH, Ferreira EI (2008) Abordagem racional no planejamento de novos tuberculostáticos: inibidores da InhA, enoil-ACP redutase do M. tuberculosis. Rev Bras ciencias Farm 44: 167–179. doi:10.1590/S1516-93322008000200002.
25. Andrade CH, Salum LDB, Castilho MS, Pasqualoto KFM, Ferreira EI, et al. (2008) Fragment-based and classical quantitative structure-activity relationships for a series of hydrazides as antituberculosis agents. Mol Divers 12: 47–59. doi:10.1007/s11030-008-9074-z.
26. Andrade CH, Pasqualoto KFM, Ferreira EI, Hopfinger AJ (2009) Rational design and 3D-pharmacophore mapping of 5′-thiourea-substituted alpha-thymidine analogues as mycobacterial TMPK inhibitors. J Chem Inf Model 49: 1070–1078. doi:10.1021/ci8004622.
27. Braga RC, Sabino JR, de Valeria O, Andrade CH (2010) Discovery of novel hit compounds for Trypanosoma cruzi sterol 14α-demethylase through structure-based virtual screening. Abstracts of Papers, 240th American Chemical Society National Meeting & Exposition, Boston, MA, United States, August 22–26. p. MEDI–379.
28. Andrade CH, Pasqualoto KFM, Ferreira EI, Hopfinger AJ (2010) 3D-Pharmacophore mapping of thymidine-based inhibitors of TMPK as potential antituberculosis agents. J Comput Aided Mol Des 24: 157–172. doi:10.1007/s10822-010-9323-y.
29. Andrade CH, Pasqualoto KFM, Ferreira EI, Hopfinger AJ (2010) 4D-QSAR: perspectives in drug design. Molecules 15: 3281–3294. doi:10.3390/molecules15053281.
30. Braga RC, Liao LM, Bezerra JC, Vinaud MC, Andrade CH (2012) Integrated Chemoinformatics approaches to virtual screening in the search of new lead compounds against Leishmania. Abstracts of Papers, 244th American Chemical Society National Meeting & Exposition, Philadelphia, PA, United States, August 19–23, 2012. p. CINF–46.
31. Neves BJ, Bueno R V, Braga RC, Andrade CH (2013) Discovery of new potential hits of Plasmodium falciparum enoyl-ACP reductase through ligand- and structure-based drug design approaches. Bioorg Med Chem Lett 23: 2436–2441. doi:10.1016/j.bmcl.2013.02.006.
32. Bueno R V, Toledo NR, Neves BJ, Braga RC, Andrade CH (2013) Structural and chemical basis for enhanced affinity to a series of mycobacterial thymidine monophosphate kinase inhibitors: fragment-based QSAR and QM/MM docking studies. J Mol Model 19: 179–192. doi:10.1007/s00894-012-1527-8.
33. Braga RC, Andrade CH (2013) Assessing the Performance of 3D Pharmacophore Models in Virtual Screening: How Good are They? Curr Top Med Chem 13: 1127–1138.
34. Melo-Filho CC, Braga RC, Andrade CH (2014) 3D-QSAR Approaches in Drug Design: Perspectives to Generate Reliable CoMFA Models. Curr Comput Aided Drug Des 10: 148–159.
35. Bueno R V, Braga RC, Segretti ND, Ferreira EI, Trossini GHG, et al. (2014) New tuberculostatic agents targeting nucleic acid biosynthesis: drug design using QSAR approaches. Curr Pharm Des 20: 4474–4485.
36. Braga RC, Alves VM, Silva AC, Nascimento MN, Silva FC, et al. (2014) Virtual Screening Strategies in Medicinal Chemistry: The state of the art and current challenges. Curr Top Med Chem 14.
37. Braga RC, Pinto EG, Martins LF, Tempone AG, Liao LM, et al. (2014) Multitarget virtual screening approach: Identification of new hit compounds against Leishmania infantum. Abstracts of Papers, 248th ACS National Meeting & Exposition, San Francisco, CA, United States, August 10–14, 2014. American Chemical Society. p. COMP–322.
38. Braga RC, Alves VM, Silva MFB, Muratov E, Fourches D, et al. (2014) Tuning hERG out: Antitarget QSAR Models for Drug Development. Curr Top Med Chem 14: 1399–1415.
39. Andrade CH, Silva DC, Braga RC (2014) In silico Prediction of Drug Metabolism by P450. Curr Drug Metab.
40. Andrade CH, Braga RC (2014) Drug Metabolism, Toxicology Experimental Determination And Theoretical Prediction: Challenges And Perspectives From A Medicinal Chemistry Point Of View. Curr Top Med Chem 14: 1323–1324.
41. Alves VM, Braga RC, Silva MB, Muratov E, Fourches D, et al. (2014) Pred-hERG: A novel web-accessible computational tool for predicting cardiac toxicity of drug candidates. Abstracts of Papers, 248th ACS National Meeting & Exposition, San Francisco, CA, United States, August 10–14, 2014. American Chemical Society. p. CINF–40.
42. Carneiro EO, Andrade CH, Braga RC, Tôrres ACB, Alves RO, et al. (2010) Structure-based prediction and biosynthesis of the major mammalian metabolite of the cardioactive prototype LASSBio-294. Bioorg Med Chem Lett 20: 3734–3736. doi:10.1016/j.bmcl.2010.04.073.
43. Braga RC, Andrade CH (2012) QSAR and QM/MM approaches applied to drug metabolism prediction. Mini Rev Med Chem 12: 573–582.
44. Iorio F, Isacchi A, di Bernardo D, Brunetti-Pierri N (2010) Identification of small molecules enhancing autophagic function from drug network analysis. Autophagy 6: 1204–1205.
45. Clark PM, Dawany N, Dampier W, Byers SW, Pestell RG, et al. (2012) Bioinformatics analysis reveals transcriptome and microRNA signatures and drug repositioning targets for IBD and other autoimmune diseases. Inflamm Bowel Dis 18: 2315–2333.
46. Bispo NA, Culleton R, Silva LA, Cravo P (2013) A Systematic In Silico Search for Target Similarity Identifies Several Approved Drugs with Potential Activity against the Plasmodium falciparum Apicoplast. PLoS One 8: e59288. doi:10.1371/journal.pone.0059288.
47. Caffrey CR, Rohwer A, Oellien F, Marhöfer RJ, Braschi S, et al. (2009) A comparative chemogenomics strategy to predict potential drug targets in the metazoan pathogen, Schistosoma mansoni. PLoS One 4: e4413. doi:10.1371/journal.pone.0004413.
48. Zhu F, Han B, Kumar P, Liu X, Ma X, et al. (2010) Update of TTD: Therapeutic Target Database. Nucleic Acids Res 38: D787–91.
49. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, et al. (2006) DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res 34: D668–72.
50. Kuhn M, Szklarczyk D, Franceschini A, von Mering C, Jensen LJ, et al. (2012) STITCH 3: zooming in on protein-chemical interactions. Nucleic Acids Res 40: D876–80.
51. Bredel M, Jacoby E (2004) Chemogenomics: an emerging strategy for rapid target and drug discovery. Nat Rev Genet 5: 262–275.
52. Rognan D (2007) Chemogenomic approaches to rational drug design. Br J Pharmacol 152: 38–52.
53. Protasio A V, Tsai IJ, Babbage A, Nichol S, Hunt M, et al. (2012) A systematically improved high quality genome and transcriptome of the human blood fluke Schistosoma mansoni. PLoS Negl Trop Dis 6: e1455. doi:10.1371/journal.pntd.0001455.
54. Agüero F, Al-Lazikani B, Aslett M, Berriman M, Buckner FS, et al. (2008) Genomic-scale prioritization of drug targets: the TDR Targets database. Nat Rev Drug Discov 7: 900–907.
55. Hertz-Fowler C, Peacock CS, Wood V, Aslett M, Kerhornou A, et al. (2004) GeneDB: a resource for prokaryotic and eukaryotic organisms. Nucleic Acids Res 32: D339–43.
56. Dong Y, Guha S, Sun X, Cao M, Wang X, et al. (2012) Nutraceutical interventions for promoting healthy aging in invertebrate models. Oxid Med Cell Longev 2012: 718491. doi:10.1155/2012/718491.
57. Agarwal P, States DJ (1998) Comparative accuracy of methods for protein sequence similarity search. Bioinformatics 14: 40–47.
58. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 38: W529–33.
59. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, et al. (2004) UniProt: the Universal Protein knowledgebase. Nucleic Acids Res 32: D115–9.
60. Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22: 1658–1659.
61. Katoh K, Kuma K, Miyata T, Toh H (2005) Improvement in the accuracy of multiple sequence alignment program MAFFT. Genome informatics 16: 22–33.
62. Pupko T, Bell RE, Mayrose I, Glaser F, Ben-Tal N (2002) Rate4Site: an algorithmic tool for the identification of functional regions in proteins by surface mapping of evolutionary determinants within their homologues. Bioinformatics 18 Suppl 1: S71–7.
63. Mayrose I, Graur D, Ben-Tal N, Pupko T (2004) Comparison of site-specific rate-inference methods for protein sequences: empirical Bayesian methods are superior. Mol Biol Evol 21: 1781–1791.

64. Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, et al. (2012) PubChem's BioAssay Database. Nucleic Acids Res 40: D400–12. doi:10.1093/nar/gkr1132.
65. Fourches D, Muratov E, Tropsha A (2010) Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research. J Chem Inf Model 50: 1189–1204. doi:10.1021/ci100176x.
66. Landrum G (2014) RDKit: Open-source cheminformatics: http://www.rdkit.org/.
67. R Development Core Team (2008) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
68. Hopkins AL, Groom CR (2002) The druggable genome. Nat Rev Drug Discov 1: 727–730.
69. Russ AP, Lampel S (2005) The druggable genome: an update. Drug Discov Today 10: 1607–1610.
70. Edfeldt FNB, Folmer RHA, Breeze AL (2011) Fragment screening to predict druggability (ligandability) and lead discovery success. Drug Discov Today 16: 284–287.
71. Dobson CM (2004) Chemical space and biology. Nature 432: 824–828.
72. Lipinski C, Hopkins A (2004) Navigating chemical space for biology and medicine. Nature 432.
73. Audisio D, Messaoudi S, Ijjaali I, Dubus E, Petitet F, et al. (2010) Assessing the chemical diversity of an hsp90 database. Eur J Med Chem 45: 2000–2009.
74. Lachance H, Wetzel S, Kumar K, Waldmann H (2012) Charting, navigating, and populating natural product chemical space for drug discovery. J Med Chem 55: 5989–6001.
75. Bon RS, Waldmann H (2010) Bioactivity-guided navigation of chemical space. Acc Chem Res 43: 1103–1114. doi:10.1021/ar100014h.
76. Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, et al. (2009) The genome of the blood fluke Schistosoma mansoni. Nature 460: 352–358. doi:10.1038/nature08160.
77. Ribeiro P, Patocka N (2013) Neurotransmitter transporters in schistosomes: structure, function and prospects for drug discovery. Parasitol Int 62: 629–638.
78. Doenhoff MJ, Cioli D, Utzinger J (2008) Praziquantel: mechanisms of action, resistance and new derivatives for schistosomiasis. Curr Opin Infect Dis 21: 659–667.
79. Smout MJ, Kotze AC, McCarthy JS, Loukas A (2010) A novel high throughput assay for anthelmintic drug screening and resistance diagnosis by real-time monitoring of parasite motility. PLoS Negl Trop Dis 4: e885.
80. Cohen CJ, Spires S, Van Skiver D (1992) Block of T-type Ca channels in guinea pig atrial cells by antiarrhythmic agents and Ca channel antagonists. J Gen Physiol 100: 703–728.
81. Singh BN (1986) The mechanism of action of calcium antagonists relative to their clinical applications. Br J Clin Pharmacol 21 Suppl 2: 109S–121S.
82. Silva-Moraes V, Couto FFB, Vasconcelos MM, Araújo N, Coelho PMZ, et al. (2013) Antischistosomal activity of a calcium channel antagonist on schistosomula and adult Schistosoma mansoni worms. Mem Inst Oswaldo Cruz 108: 600–604.
83. Oxford AE, Raistrick H, Simonart P (1939) Studies in the biochemistry of micro-organisms: Griseofulvin, C(17)H(17)O(6)Cl, a metabolic product of Penicillium griseo-fulvum Dierckx. Biochem J 33: 240–248.
84. Rathinasamy K, Jindal B, Asthana J, Singh P, Balaji P V, et al. (2010) Griseofulvin stabilizes microtubule dynamics, activates p53 and inhibits the proliferation of MCF-7 cells synergistically with vinblastine. BMC Cancer 10: 213.
85. Uen Y-H, Liu D-Z, Weng M-S, Ho Y-S, Lin S-Y (2007) NF-kappaB pathway is involved in griseofulvin-induced G2/M arrest and apoptosis in HL-60 cells. J Cell Biochem 101: 1165–1175.
86. Ruge E, Korting HC, Borelli C (2005) Current state of three-dimensional characterisation of antifungal targets and its use for molecular modelling in drug design. Int J Antimicrob Agents 26: 427–441.
87. Brian PW (1949) Studies on the Biological Activity of Griseofulvin. Ann Bot 13: 59–77.
88. Da Silva Barros ME, de Assis Santos D, Hamdan JS (2007) Evaluation of susceptibility of Trichophyton mentagrophytes and Trichophyton rubrum clinical isolates to antifungal drugs using a modified CLSI microdilution method (M38-A). J Med Microbiol 56: 514–518.
89. Chaudhuri a R, Ludueña RF (1996) Griseofulvin: a novel interaction with bovine brain tubulin. Biochem Pharmacol 51: 903–909.
90. Czymmek KJ, Bourett TM, Shao Y, DeZwaan TM, Sweigard JA, et al. (2005) Live-cell imaging of tubulin in the filamentous fungus Magnaporthe grisea treated with anti-microtubule and anti-microfilament agents. Protoplasma 225: 23–32.
91. Rebacz B, Larsen TO, Clausen MH, Rønnest MH, Löffler H, et al. (2007) Identification of griseofulvin as an inhibitor of centrosomal clustering in a phenotype-based screen. Cancer Res 67: 6342–6350. doi:10.1158/0008-5472.CAN-07-0663.
92. Jankovic J, Clarence-Smith K (2011) Tetrabenazine for the treatment of chorea and other hyperkinetic movement disorders. Expert Rev Neurother 11: 1509–1523.
93. Wimalasena K (2011) Vesicular monoamine transporters: structure-function, pharmacology, and medicinal chemistry. Med Res Rev 31: 483–519.
94. Abdulla M-H, Ruelas DS, Wolff B, Snedecor J, Lim K-C, et al. (2009) Drug discovery for schistosomiasis: hit and lead compounds identified in a library of known drugs by medium-throughput phenotypic screening. PLoS Negl Trop Dis 3: e478. doi:10.1371/journal.pntd.0000478.
95. Taft AS, Norante FA, Yoshino TP (2010) The identification of inhibitors of Schistosoma mansoni miracidial transformation by incorporating a medium-throughput small-molecule screen. Exp Parasitol 125: 84–94.
96. Brugnara C, de Franceschi L, Alper SL (1993) Inhibition of Ca(2+)-dependent K+ transport and cell dehydration in sickle erythrocytes by clotrimazole and other imidazole derivatives. J Clin Invest 92: 520–526.
97. Yazejian B, Sun XP, Grinnell AD (2000) Tracking presynaptic Ca2+ dynamics during neurotransmitter release with Ca2+-activated K+ channels. Nat Neurosci 3: 566–571.
98. Brayden JE, Nelson MT (1992) Regulation of arterial tone by activation of calcium-dependent potassium channels. Science 256: 532–535.
99. Miyazaki T, Sagawa R, Honma T, Noguchi S, Harada T, et al. (2004) 73-kDa molecular chaperone HSP73 is a direct target of antibiotic gentamicin. J Biol Chem 279: 17295–17300.
100. Oskay Halacli S, Halacli B, Altundag K (2013) The significance of heat shock proteins in breast cancer therapy. Med Oncol 30: 575.
101. Yang J, Yang L, Lv Z, Wang J, Zhang Q, et al. (2012) Molecular cloning and characterization of a HSP70 gene from Schistosoma japonicum. Parasitol Res 110: 1785–1793.
102. Levine SN, Hollier B (n.d.) Aprindine inhibits calmodulin-stimulated phosphodiesterase and Ca-ATPase activities. J Cardiovasc Pharmacol 5: 151–156.
103. Tidow H, Nissen P (2013) Structural diversity of calmodulin binding to its target sites. FEBS J 280: 5551–5565.
104. Katsumata T, Kohno S, Yamaguchi K, Hara K, Aoki Y (1989) Hatching of Schistosoma mansoni eggs is a Ca2+/calmodulin-dependent process. Parasitol Res 76: 90–91.
105. Katsumata T, Shimada M, Sato K, Aoki Y (1988) Possible involvement of calcium ions in the hatching of Schistosoma mansoni eggs in water. J Parasitol 74: 1040–1041.
106. Lewert RM, Hopkins DR, Mandlowitz S (1966) The role of calcium and magnesium ions in invasiveness of schistosome cercariae. Am J Trop Med Hyg 15: 314–323.
107. Fusco AC, Salafsky B, Vanderkooi G, Shibuya T (1991) Schistosoma mansoni: the role of calcium in the stimulation of cercarial proteinase release. J Parasitol 77: 649–657.
108. Salathe M (2007) Regulation of mammalian ciliary beating. Annu Rev Physiol 69: 401–422.
109. Jaffe JJ, McCormack JJ, Meymarian E (1972) Comparative properties of schistosomal and filarial dihydrofolate reductases. Biochem Pharmacol 21: 719–731.
110. Nessim NG, Hassan SI, William S, el-Baz H (2000) Effect of the broad spectrum anthelmintic drug flubendazole upon Schistosoma mansoni experimentally infected mice. Arzneimittelforschung 50: 1129–1133. doi:10.1055/s-0031-1300336.
111. Menezes CMS, Rivera G, Alves M a, do Amaral DN, Thibaut JPB, et al. (2012) Synthesis, biological evaluation, and structure-activity relationship of clonazepam, meclonazepam, and 1,4-benzodiazepine compounds with schistosomicidal activity. Chem Biol Drug Des 79: 943–949.
112. Angelucci F, Sayed AA, Williams DL, Boumis G, Brunori M, et al. (2009) Inhibition of Schistosoma mansoni thioredoxin-glutathione reductase by auranofin: structural and kinetic aspects. J Biol Chem 284: 28977–28985.
113. Li S, Wu L, Liu Z, Hu L, Xu P, et al. (1996) Studies on prophylactic effect of artesunate on schistosomiasis japonica. Chin Med J (Engl) 109: 848–853.

MedChemComm

RESEARCH ARTICLE

View Article Online
View Journal | View Issue



Cite this: *Med. Chem. Commun.*, 2016, 7, 1176

# The antidepressant drug paroxetine as a new lead candidate in schistosome drug discovery†

Bruno Junior Neves,‡[a] Rafael Ferreira Dantas,‡[b] Mario Roberto Senger,[b] Walter César Góes Valente,[b] João de Mello Rezende-Neto,[b] Willian Távora Chaves,[b] Lee Kamentsky,[c] Anne Carpenter,[c] Floriano Paes Silva-Junior*[b] and Carolina Horta Andrade*[a]

Received 23rd December 2015,
Accepted 21st April 2016

DOI: 10.1039/c5md00596e

www.rsc.org/medchemcomm

Recently, our *in silico* repositioning-chemogenomics approach predicted paroxetine (PAR), an antidepressant drug, as a inhibitor of *Schistosoma mansoni* serotonin transporters (*Sm*SERTs), and consequently, a new anti-schistosomal candidate. With the aim of determining the anti-schistosomal activity of this drug, we initially used a spectrophotometric assay to determine activity against schistosomula worms. During this investigation, we verified that PAR showed a pronounced effect on schistosomula viability ($IC_{50}$ = 2.5 μM) after 72 h of incubation. Then, we performed *ex vivo* studies with adult *S. mansoni* worms using a new automated image-based assay to accurately measure worm motility. As expected from the PAR's predicted mechanism of action, both male and female worms treated with low concentrations of PAR exhibited enhanced motility followed by reduction in motility as incubation time increased. PAR $EC_{50}$ values for motility reduction in male and female worms were 5.1 μM and 9.9 μM after 24 h of exposure, respectively, and this effect was maintained until the end of the experiment (72 h). Lastly, homology modeling and docking studies with *Sm*SERT-A and human SERT (*h*SERT) revealed insights into the chemical basis of PAR anti-schistosomal activity. These results provide crucial guidance for further studies to optimize PAR in terms of potency and selectivity.

## Introduction

Schistosomiasis is a neglected tropical disease caused by flatworms of the genus *Schistosoma*, with three species (*S. mansoni*, *S. haematobium*, and *S. japonicum*) accounting for the majority of human infections. These worms cause a chronic and often debilitating infection that impairs development and productivity, and this exposure is strongly linked to extreme poverty.[1–4] Recent estimates suggest that around 262 million people are infected, mainly in sub-Saharan Africa, the Middle East, the Caribbean, and South America, resulting in up to 200 000 deaths annually, while more than 42 million infected individuals experience high morbidity.[5]

In the absence of an effective vaccine, the control of schistosomiasis relies on a single drug, praziquantel (PZQ), which has been used in mass drug administration programs for almost four decades.[6] However, PZQ has low efficacy against juvenile worms, found during the early stages of infection (14–28 day post infection).[6–9] Moreover, reports of PZQ resistant or poorly sensitive isolates in African countries have been published[10–12] and a resistant strain was also induced in the laboratory.[13–15] Consequently, there are concerns that the widespread use of PZQ could potentially lead to development of resistant parasites in the near future. Hence, there is an urgent need for discovering new anti-schistosomal drugs with new mechanisms of action.

Recently, using an *in silico* repositioning-chemogenomics approach,[16] our group predicted paroxetine (PAR), an antidepressant drug, as a *S. mansoni* serotonin transporter (*Sm*SERT) inhibitor. Studies on the physiological functions of plasma membrane SERT were first focused on the mammalian central nervous system, where it mediates re-uptake of the amine across the presynaptic membrane.[17] This is part of a mechanism that inactivates serotonin signalling by quickly removing it from the synaptic cleft.[18]

Studies dating back to the 1970s and 1980s have shown that among parasitic flatworms, in particularly *S. mansoni*, serotonin is an important modulator of neuromuscular function and increases metabolic activity by stimulating glucose

[a] *LabMol – Laboratory for Molecular Modeling and Design, Faculdade de Farmácia, Universidade Federal de Goiás, Goiânia - GO, Brazil. E-mail: carolina@ufg.br*
[b] *LaBECFar – Laboratório de Bioquímica Experimental e Computacional de Fármacos, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro - RJ, Brazil. E-mail: floriano@ioc.fiocruz.br*
[c] *Imaging Platform, Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts, USA*
† The authors declare no competing interests.
‡ These authors contributed equally.

uptake, glycogen breakdown, and lactate excretion.[19–22] The effects on motility are seen in intact, unpermeabilized worms, showing that serotonin either acts through surface receptors, or is taken internally via a specific transporter and acts directly on the worm musculature.[23] More recently, the existence of two *Sm*SERT isoforms was confirmed by molecular methods. These isoforms are structurally similar to other SERTs except for the presence of an additional 78 amino acids at the N-terminal end, which is presumably due to differential splicing of the transcript. In addition, both isoforms were cloned from *S. mansoni*, expressed heterologously in mammalian cells and shown to have selective, high-affinity serotonin transporter activity with a Km value comparable to that of mammalian SERTs.[23,24] This suggests that *S. mansoni* is dependent on *Sm*SERTs to regulate serotonin concentrations in the synaptic cleft and consequently its neuromuscular function.[23–25]

In this study, we validated the anti-schistosomal activity of PAR on schistosomula and adult *S. mansoni* worms using a spectrophotometric assay and a new automated image-based assay to accurately measure parasite viability and motility, respectively. Further, to guide the design and optimization of new anti-schistosomal drugs, we explored the interactions of PAR with the binding site of *Sm*SERTs and its respective host counterpart (*h*SERT) using molecular modeling strategies.

## Results and discussion

### Biological activity

The lack of activity against immature worms is a therapeutic limitation of PZQ which hampers the treatment in endemic areas since it does not prevent re-infection.[26] Therefore, we used a spectrophotometric, *i.e.*, resazurin-based assay, to determine the viability of schistosomula after the exposure to PAR for 72 h incubation period. This approach is commonly used as an initial screening step in antischistosomal lead discovery campaigns,[27–32] since schistosomula are relatively easier to obtain in larger amount when compared with adult worms. This analysis showed that PAR decreased schistosomula viability after 72 h of incubation ($IC_{50}$ = 2.5 μM) while PZQ had only negligible activity against this parasite stage at concentrations as high as 32 μM.[33] Remarkably, PAR effect on schistosomula viability corroborates with results of a screening campaign with 2160 compounds that identified several *h*SERT inhibitors as potent hits for *S. mansoni* schistosomula.[32]

After the initial screening, separated individual adult male and female *S. mansoni* worms were incubated with PAR for various time periods up to 72 h. Consistent and objective quantification of anti-schistosomal activity relied on an automated image-based motility assay; especially important given the limited numbers of adult worms that can be feasibly analysed and the tedious procedure for culturing and isolating them. As expected from the hypothesized mechanism of action for PAR in schistosomes, we observed for low drug concentrations an increase in worm motility of up to 9-fold in females and 2-fold in males immediately after addition of the drug in the culture medium, i.e., time 0 h (see Fig. 1).

This transient excitation in worm motility was only detected for the lower part of the tested drug concentration range (≤10 μM in females and ≤20 μM in males). This characteristic corroborates the effect produced by two classical serotonin transport inhibitors, fluoxetine and clomipramine, on schistosomula which were previously shown to induce a strongly hyperactive phenotype, corresponding to a 3-fold increase in larval motility, roughly the same effect as treatment with an excess of exogenous serotonin.[34] On the other hand, higher concentrations may be extremely toxic to the parasite causing immediate viability loss and a consequent reduction in worm motility. After 24 h incubation this hypermotility effect is reversed, with schistosomes showing decreased motility, which is expected to also be associated with diminished viability. This delayed effect would be consistent with the timing of serotonin receptor internalisation, a protective mechanism to avoid excitotoxicity.[35]

Using worm motility measurements obtained from our automated image-based assay, dose–response curves were also generated for each incubation time as exemplified in Fig. 2 for the 24 h time point.

The $EC_{50}$ values determined for PAR (Table 1) indicate that after 24 h incubation the effect is fully developed with values varying from 2.7 to 5.1 μM for male worms and 9.9 to 11.9 μM for female worms. The effect of PAR on adult males after 72 h of exposure is in the same range of the value determined for schistosomula. Our results also indicate that male worms are slightly more sensitive to PAR action since they showed on average $EC_{50}$ values 2–3 times lower than those determined in females. In part, this could be due to a gender-specific expression pattern of *Sm*SERT.[34]

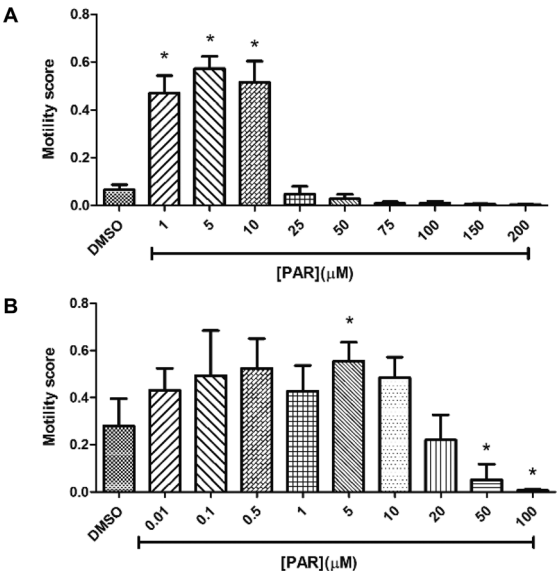



Fig. 1 PAR effect on motility of adult female (A) and male (B) *S. mansoni* worms immediately after addition to culture medium (0 h). Data expressed as mean ± standard deviation. * Statistically different from DMSO control ($p$-value < 0.05).

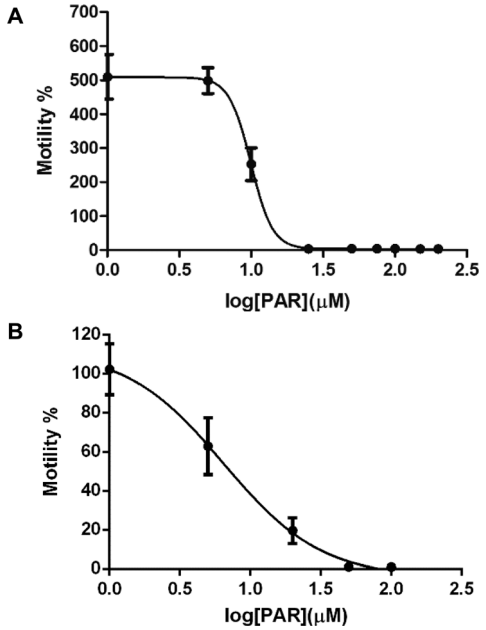



Fig. 2 PAR dose–response curves in female (A) and male (B) *S. mansoni* adult worms after 24 h of incubation. Data expressed as mean ± standard error for the mean. Percent values were calculated using the motility mean of the control group (DMSO 0.2%). PAR concentrations that increased worms motility are not shown in male dose–response curve.

Table 1 Time- and gender dependent $EC_{50}$ values (μM) for PAR and PZQ

| | PAR | | PZQ | |
|---|---|---|---|---|
| Time/gender | Male | Female | Male | Female |
| 0 h | 23.5 | 17.4 | 0.11 | 0.13 |
| 24 h | 5.1 | 9.9 | 0.12 | 0.30 |
| 48 h | 3.8 | 11.5 | 0.13 | No fit |
| 72 h | 2.7 | 11.9 | No fit | 0.66 |

On the other hand, $EC_{50}$ values indicate that PAR had a less pronounced effect on *S. mansoni* motility than PZQ at all incubation times ($EC_{50}$ values ≤ 0.66 μM, see Table 1). These results are in general accordance with the $EC_{50}$ value described for PZQ (0.2 μM) after 72 h of exposure using a visual analysis.[30] However, the alternate mechanism of action of PAR (and status as clinically-approved for human use) makes it an interesting molecular scaffold to develop more efficient compounds. Additionally, in contrast to PZQ, for which the molecular target remains unknown, new analogs of PAR may be rationally designed using 3D structural information of SERTs during the chemical optimization process (see docking results below).

### Homology modelling of *Sm*SERTs and *h*SERT

We developed homology models of *Sm*SERTs and *h*SERT to shed some light into PAR's binding mode and to explore structural differences between the two isoforms of *Sm*SERT and the human enzyme. The two *Sm*SERT isoforms differed in only three amino acids located in the protein's predicted intracellular N-terminal. *Sm*SERT-A has Leu99, Glu100 and Val118 while the corresponding amino acid residues in *Sm*SERT-B are Ser99, His100 and Iso118.[24] Initially, we selected suitable template protein structures in Protein Data Bank (PDB),[36] observing the following criteria: the template should have a high coverage (>70% of target aligned to template), sequence identity >30%, and X-ray crystallography resolution ≤3.0 Å. The *Sm*SERT-A and *Sm*SERT-B were built using *Drosophila melanogaster* dopamine transporter (*Dm*DAT)[37,38] available under PDB code 4M48 (sequence identity = 54.5%, and coverage = 73.3%). The *h*SERT structure was also modelled using *Dm*DAT available under PDB code 4M48, (sequence identity = 53.3% and coverage = 85.0%).

In order to identify structural distances between the modeled structures and their templates, the *Sm*SERT-A, *Sm*SERT-B, and *h*SERT models were superimposed onto their corresponding templates and root-mean-square deviations (RMSD) between backbone Cα-atoms were calculated. The *Sm*SERT models showed a RMSD value of 0.19 Å (Fig. 3A), while the *h*SERT model showed an RMSD of 0.06 Å (Fig. 3B). Most of the differences between the models and their templates were observed in the extracellular flexible loop. Remarkably, this extracellular loop of *Sm*SERTs is exceptionally long compared to the *h*SERT and therefore biologically relevant, but this region is far from the binding site of SERTs[23] and it is unlikely that these differences will affect the docking studies performed in the binding site. In addition, we did not observe structural and conformational differences between binding sites of *Sm*SERTs. Therefore, we decided to keep only the *Sm*SERT-A structure for the next modelling experiments. In order to evaluate the stereochemical quality of dihedral angles *phi* against *psi* of amino acid residues in the modeled protein structures and to identify sterically allowed regions for these angles, we used the PROCHECK analysis. Analysis of *Sm*SERT-A protein revealed that 93.8% residues are within the most favoured regions (red), 5.4% in additional allowed regions (yellow), 0.6% residues in generously allowed regions (beige), and only 0.2% residues in the disallowed regions (white) of the Ramachandran plot (Fig. 3C), showing the good quality of the generated homology model. Similarly, for *h*SERT structure the observed values were 93.1%, 5.9%, 0.8%, and 0.2%, respectively (Fig. 3D). Residues located in the disallowed regions are far from the binding site of both models, indicating that these residues may not affect the ligand–protein binding simulations. The quality of the homology models was further evaluated by the ERRAT scores of 80.8 (*Sm*SERT-A) and 87.2 (*h*SERT), which indicated an acceptable protein environment for the non-bonded interactions between different atom types.

### Molecular docking of PAR

To explore the possible intermolecular interactions of PAR in the binding sites of *Sm*SERT-A and *h*SERT, we performed molecular docking. Docking results showed that the predicted binding modes of PAR in both targets are similar with the experimental binding mode of the co-crystalized ligand of X-ray structure used as template in homology modeling.[32,33] In particular, the

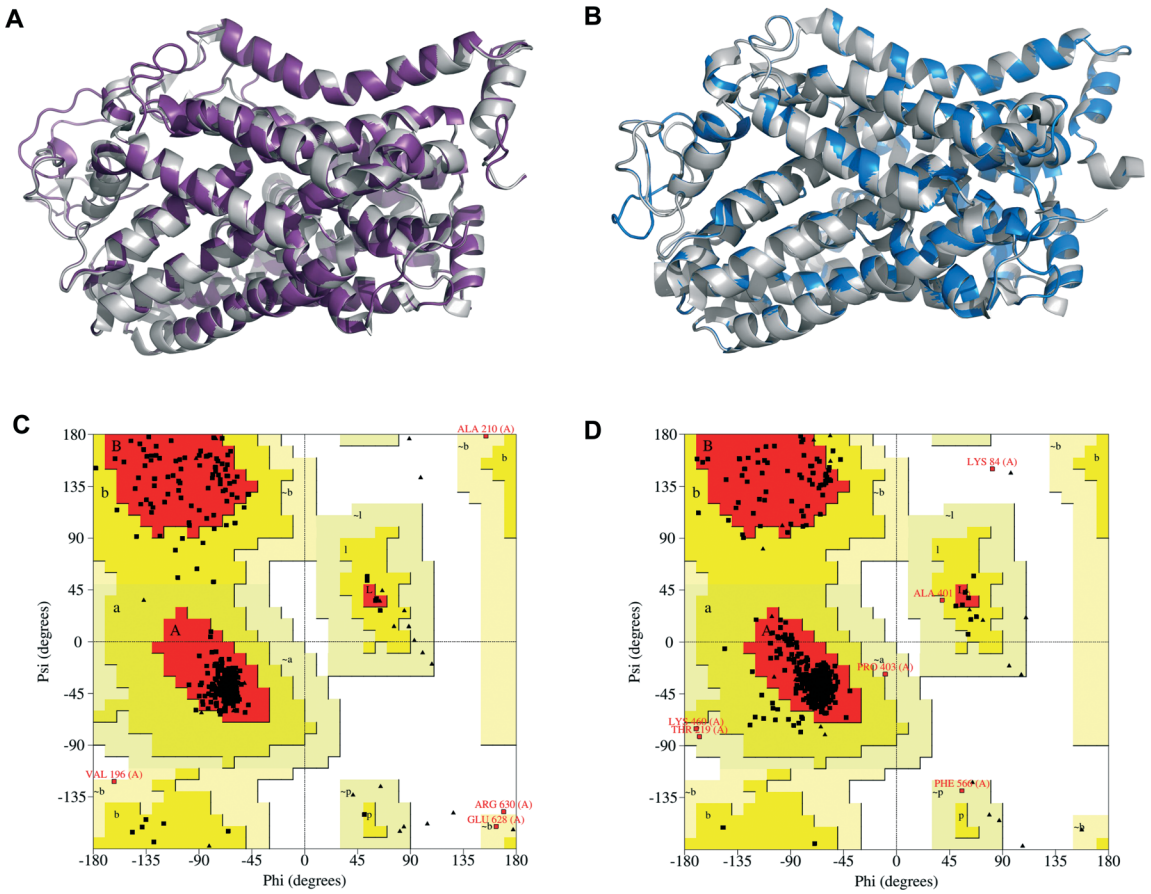



**Fig. 3** Superimposition of modeled *Sm*SERT-A (A, purple) and *h*SERT (B, blue) with their respective templates (gray). Ramachandran plots for the *Sm*SERT-A model (C) and *h*SERT model (D) obtained by PROCHECK, showing the dihedral angles *Psi* and *Phi* of amino acid residues. Red represents most favoured regions; yellow represents additional allowed regions; beige represents generously allowed regions; and white areas are disallowed regions.

aromatic rings of PAR interact with the hydrophobic pocket of *Sm*SERT-A formed by Tyr162, Phe346, Phe352, Ala506, and Thr158, whereas the protonated nitrogen of piperidine is able to form two hydrogen bonds (represented as green dashed lines) with the carbonyl groups of Phe81 and Asp84 (Fig. 4A). The docking of PAR in the binding site of *h*SERT showed similar interactions (Fig. 4B). Consequently, the predicted binding affinities of PAR by *Sm*SERT-A and *h*SERT were very similar, −11.3 and −12.6, respectively, showing that this ligand could have similar affinity with these two SERT proteins. Those observations corroborate the inhibition potencies of PAR measured by Fontana and colleagues ($IC_{50}$ for *Sm*SERTs around

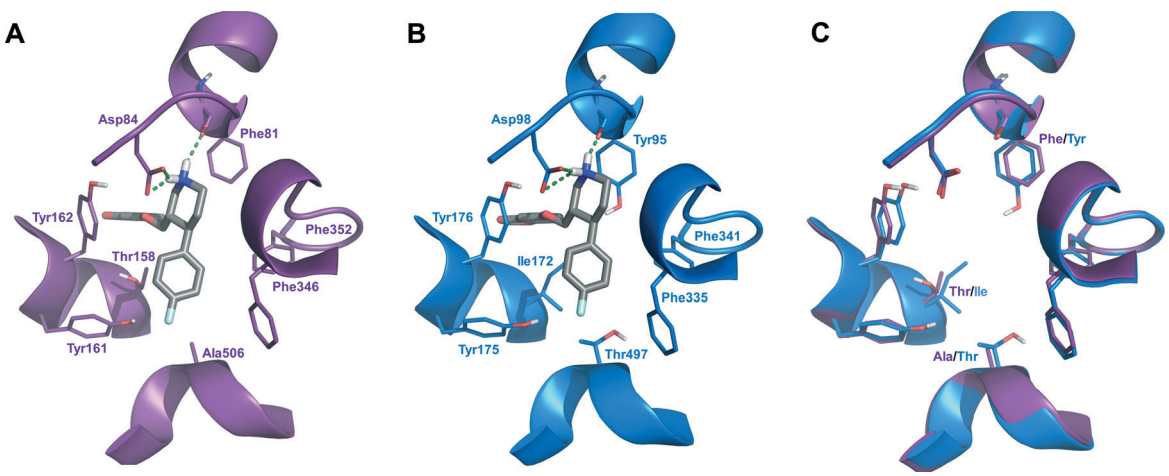



**Fig. 4** Intermolecular interactions of *Sm*SERT-A (A) and *h*SERT (B) with PAR, and structural differences between binding sites (C).

0.09 μM and $IC_{50}$ for *h*SERT = 0.02 μM) in a heterologous expression system.[20] It is worth noting that the protein flexibility should be taken into account to further explore the differences in the binding process of PAR with *Sm*SERT-A and *h*SERT in order to design a more selective *Sm*SERT inhibitor.

In addition, our analysis of the modeled proteins revealed differences in the hydrophobicity of the two binding sites. Several amino acid residues of the binding site (Thr158, Phe81 and Ala506) of *Sm*SERTs were substituted in *h*SERT by Ile172, Tyr95 and Thr497, respectively (Fig. 4C). It therefore appears that the binding site of *Sm*SERTs can accommodate bulkier ligands. These key differences between *Sm*SERTs and *h*SERT may be useful to design more potent and selective anti-schistosomal drug candidates.

# Experimental

### Reagents

PAR was purchased from Gemini (Anápolis-GO, Brazil) while PZQ was acquired from Sigma-Aldrich (St. Louis-MO, USA). Both drugs were diluted in 100% DMSO prior the assays. DMEM medium were purchased from Vitrocell (Campinas-SP, Brazil). All other reagents were acquired from Sigma-Aldrich (St. Louis-MO, USA).

### Ethics statement

Animal maintenance and experiments were carried out in accordance with the Institutional Ethics Committee for Laboratory Animal Use at the Oswaldo Cruz Foundation (CEUA/FIOCRUZ, Brazil; license number LW-78/12).

### *In vitro* schistosomula resazurin-based viability assay

Cercariae (*S. mansoni* BH strain) were shed in clean tap water from infected *Biomphalaria glabrata* snails exposed to direct illumination for 1 h. Cercariae were then vortexed at maximum speed for 5 min and transformed into schistosomula by an mechanical method adapted from literature.[33,39] Then, we performed the schistosomula resazurin-based viability assay as described previously, with minimal modifications.[33] Schistosomula were maintained in a complete M169 medium (120 per well) and incubated in black 384 well plates at 37 °C and 5% $CO_2$. The effect of PAR on schistosomula viability was assessed 72 h after drug exposure at concentrations of 0.05–200 μM. Lastly, 8 μL of resazurin (final concentration: 15 μg $mL^{-1}$) was added to each well 24 h before fluorescence measurement. The fluorescence intensity was measured using a FlexStation 3 Multi-Mode Microplate Reader (Molecular Devices) using an excitation wavelength of 530 nm and an emission wavelength of 580 nm.

### Mice infection and *ex vivo* assays on adult *S. mansoni*

Swiss mice (three- to six-days old) were infected percutaneously with 150 ± 10 cercariae (*S. mansoni* BH strain). The animals were placed, for a period of 30 min, into cylindrical vials under incandescent light with a thin water layer containing the cercariae. After *S. mansoni* maturity (42–49 days after infection), mice were euthanized, and worms were perfused with 0.85% sodium chloride and 0.75% sodium citrate solution. Male and female worms were maintained separately into 96 well plates with complete DMEM medium (*i.e.* DMEM plus 10% fetal calf serum, 2 mM L-glutamine, 100 μM $ml^{-1}$ penicillin, 100 μg $ml^{-1}$ streptomycin) throughout the entire experiment in an incubator at 37 °C with 5% $CO_2$. The effect of PAR (concentrations ranging between 0.01–200 μM) and positive control (*i.e.*, PZQ incubated with 0.05–1 μM) on adult worms was assessed either immediately or 24, 48 or 72 h after compound addition. Negative control wells contain DMEM with 0.2% of DMSO. The *ex vivo* drug sensitivity assay was based on motility measurements obtained with a newly developed anti-helminthic high-content screening platform briefly described below.

### Image acquisition and quantitative image analysis

Our method is based on sequential pairwise comparison of 100 time-lapse images captured every 250–300 ms using an automated bright-field microscope with a 2× objective lens (ImageXpress Micro XLS, Molecular Devices, CA). Subsequent quantitative image analysis used a custom-developed pipeline for detecting changes in parasite motility using the open-source CellProfiler software version 2.1.2.[40] The pipeline along with its validation will be thoroughly described in a subsequent publication and the pipeline itself is freely available at www.cellprofiler.org/published_pipelines.shtml.

Briefly, at each cycle of the pipeline, an image captured at a given instant ($t_n$) is compared with the image captured at the preceding instant ($t_{n-1}$) and so on until all images are processed. Two different motility measurements are calculated. First, a precursor metric, "AdjustedRandIndex" is calculated by comparing worm objects identified from images captured at times $t_n$ and $t_{n-1}$ with CellProfiler's Overlap module. This measure ranges from 0 to 1, with 1 meaning two objects are perfectly aligned (no movement). Thereafter, we calculate an "Overlap" mobility score, which is directly proportional to the amount of movement, by subtracting 1 − "AdjustedRandIndex". Another motility measure, "DiffWorms", is the mean pixel intensity of the image calculated from the absolute difference of the parasite images in $t_{n-1}$ and $t_n$. A higher DiffWorms score indicates higher parasite mobility. Both measures are iteratively taken for the 99 image pairs, and scores per well are calculated by averaging over all measurements.

### Statistical analysis

All statistical analysis and graphs were performed using GraphPad Prism version 5.00 for Windows (GraphPad Software, La Jolla California USA, www.graphpad.com).

### Homology modeling

The amino acid sequences of *Sm*SERT-A (ID: DQ220811), SmSERT-B (ID: DQ159205) and *h*SERT (ID: P31645) were retrieved from GenBank database[41] and used as targets for

homology modelling using the SWISS-MODEL server.[42] The latter performed the target-template sequence alignment after searching the putative X-ray template proteins in PDB for generating the 3D models for both target sequences. The built homology models are prone to contain internal constrains like unfavorable bond lengths, bond angles, torsion angles and contacts. Therefore, built models were exported to SAVES server (http://services.mbi.ucla.edu/SAVES/) and their overall stereochemical and structural quality was checked according to PROCHECK[43] and ERRAT scores.[44] PROCEHCK was applied to check the stereochemical quality of the model, including backbone torsional angles through the Ramachandran plot, while ERRAT, by means of its so-called overall quality factor, was used to check non-bonded atomic interactions, with higher scores indicating higher quality.

### Docking studies

The generated homology models were imported into Maestro v. 10.0 (Schrödinger, LLC, New York, NY, 2014) and prepared using the Protein Preparation Wizard workflow as follows: hydrogen atoms were added according to Epik v. 2.7[45] (pH 7.4 ± 1.0) and minimized using the OPLS-2005 force field.[46] Then, structure of PAR was also imported to Maestro and its most favorable ionization state was generated at pH 7.4 ± 1.0 using Epik. Subsequently, 300 conformations were generated using OMEGA v. 2.5.1,[47] while AM1-BCC charges[48] were added using QUACPAC v. 1.6.3.[49]

Previous to docking studies, a grid was defined to include the full ligand-binding site of *Sm*SERT, with dimensions of 18.6 Å × 16.0 Å × 17.6 Å (*x*, *y*, and *z*) and volume of 5276 Å$^3$ around Phe81, Asp84, Phe346, and Phe352 amino acid residues of binding site. The grid of *h*SERT was generated with dimensions of 17.0 Å × 16.3 Å × 17.0 Å and volume of 4720 Å$^3$ around Asp98, Tyr95, Phe335, and Phe341. Molecular docking of PAR on the built homology models was investigated using FRED software, available in OEDocking suite v. 3.0.1[50,51] using ChemGauss4 scoring function. FRED is a docking program that performs an exhaustive search, by systematically searching rotations and translations of each conformer of the ligand within the active site, filtering the possible poses for shape complementarity[52] and pharmacophoric features before selecting and optimizing poses using the Chemgauss4 scoring function.[50,53]

## Conclusions

We identified potent *in vitro* and *ex vivo* anti-schistosomal activity of PAR against schistosomula and adult life stages of *S. mansoni*, respectively. This drug offers a new biochemical pathway to kill schistosomes by disrupting serotonin signalling and its downstream events.[25,34,54] In addition, its mechanism of action may complement those currently postulated for PZQ (*i.e.*, increased $Ca^{2+}$ influx, inhibition of nucleoside uptake)[55] and thus could help to overcome resistance. Although *Sm*SERT has not been validated as the definitive target of PAR, it is likely the main mechanism of action on schistosomes. In absence of a structure obtained from X-ray crystallography or NMR experiments, we developed 3D homology models to shed some light on the structural differences between the binding sites of *Sm*SERTs and *h*SERT, and performed molecular docking studies with *Sm*SERT-A and *h*SERT to explore the possible interactions of PAR in the binding site of both SERTs. We observed similar predicted binding modes of PAR in both enzymes. In conclusion, we would recommend further medicinal chemistry efforts to optimize PAR scaffold in terms of potency, selectivity, and permeability to the tegument of the parasite.

## Acknowledgements

The authors would like to thank Brazilian funding agencies, CNPq, CAPES, FAPEG, FAPERJ and FIOCRUZ for financial support and fellowships, as well as the US National Institutes of Health (grant NIH GM095672 to AEC). We are grateful to OpenEye Scientific Software Inc. for providing academic license of their software. The authors also thank the Malacology Laboratory (Dr. Silvana C. Thiengo) from IOC/FIOCRUZ for providing *S. mansoni* cercarie and the Bioassays and Drug Screening Platform (FIOCRUZ RPT11-I subunit) for technological support. Special thanks to Molecular Devices for providing acess to HCS equipment.

## Notes and references

1 D. G. Colley, A. L. Bustinduy, W. E. Secor and C. H. King, *Lancet*, 2014, **383**, 2253–2264.
2 A. G. P. Ross, P. B. Bartley, A. C. Sleigh, G. R. Olds, Y. Li, G. M. Williams and D. P. McManus, *N. Engl. J. Med.*, 2002, **346**, 1212–1220.
3 C. H. King, *N. Engl. J. Med.*, 2009, **360**, 106–109.
4 B. Gryseels, K. Polman, J. Clerinx and L. Kestens, *Lancet*, 2006, **368**, 1106–1118.
5 World Health Organization. Schistosomiasis. Available from: http://www.who.int/mediacentre/factsheets/fs115/en, 2015.
6 R. Gönnert and P. Andrews, *Z. Parasitenkd.*, 1977, **52**, 129–150.
7 A. A. Sabah, C. Fletcher, G. Webbe and M. J. Doenhoff, *Exp. Parasitol.*, 1986, **61**, 294–303.
8 L. Pica-Mattoccia and D. Cioli, *Int. J. Parasitol.*, 2004, **34**, 527–533.
9 A. D. Aragon, R. A. Imani, V. R. Blackburn, P. M. Cupit, S. D. Melman, T. Goronga, T. Webb, E. S. Loker and C. Cunningham, *Mol. Biochem. Parasitol.*, 2009, **164**, 57–65.
10 M. Ismail, A. Metwally, A. Farghaly, J. Bruce, L. F. Tao and J. L. Bennett, *Am. J. Trop. Med. Hyg.*, 1996, **55**, 214–218.
11 S. D. Melman, M. L. Steinauer, C. Cunningham, L. S. Kubatko, I. N. Mwangi, N. B. Wynn, M. W. Mutuku, D. M. S. Karanja, D. G. Colley, C. L. Black, W. E. Secor, G. M. Mkoji and E. S. Loker, *PLoS Neglected Trop. Dis.*, 2009, **3**, e504.
12 P. G. Fallon, R. F. Sturrock, A. C. Niang and M. J. Doenhoff, *Am. J. Trop. Med. Hyg.*, 1995, **53**, 61–62.
13 F. F. B. Couto, P. M. Z. Coelho, N. Araújo, J. R. Kusel, N. Katz, L. K. Jannotti-Passos and A. C. A. Mattos, *Mem. Inst. Oswaldo Cruz*, 2011, **106**, 153–157.

14 P. G. Fallon and M. J. Doenhoff, *Am. J. Trop. Med. Hyg.*, 1994, **51**, 83–88.
15 M. M. Ismail, S. A. Taha, A. M. Farghaly and A. S. El-Azony, *J. Egypt. Soc. Parasitol.*, 1994, **24**, 685–695.
16 B. J. Neves, R. C. Braga, J. C. B. Bezerra, P. V. L. Cravo and C. H. Andrade, *PLoS Neglected Trop. Dis.*, 2015, **9**, e3435.
17 G. E. Torres, R. R. Gainetdinov and M. G. Caron, *Nat. Rev. Neurosci.*, 2003, **4**, 13–25.
18 B. J. Hoffman, S. R. Hansson, E. Mezey and M. Palkovits, *Front. Neuroendocrinol.*, 1998, **19**, 187–231.
19 T. E. Mansour, *Adv. Parasitol.*, 1985, **23**, 1–36.
20 M. S. Rahman, D. F. Mettrick and R. B. Podesta, *Exp. Parasitol.*, 1985, **60**, 10–17.
21 J. P. Boyle, J. V. Zaide and T. P. Yoshino, *Exp. Parasitol.*, 2000, **94**, 217–226.
22 J. P. Boyle and T. P. Yoshino, *J. Parasitol.*, 2005, **91**, 542–550.
23 N. Patocka and P. Ribeiro, *Mol. Biochem. Parasitol.*, 2007, **154**, 125–133.
24 a. C. K. Fontana, M. S. Sonders, O. S. Pereira-Junior, M. Knight, J. a. Javitch, V. Rodrigues, S. G. Amara and O. V. Mortensen, *Eur. J. Pharmacol.*, 2009, **616**, 48–57.
25 P. Ribeiro and N. Patocka, *Parasitol. Int.*, 2013, **62**, 629–638.
26 D. Cioli, L. Pica-Mattoccia, A. Basso and A. Guidi, *Mol. Biochem. Parasitol.*, 2014, **195**, 23–29.
27 G. Panic, D. Flores, K. Ingram-Sieber and J. Keiser, *Parasites Vectors*, 2015, **8**, 624.
28 N. Cowan and J. Keiser, *Parasites Vectors*, 2015, **13**, 417.
29 G. Panic, M. Vargas, I. Scandale and J. Keiser, *PLoS Neglected Trop. Dis.*, 2015, **9**, e0003962.
30 K. Ingram-Sieber, N. Cowan, G. Panic, M. Vargas, N. R. Mansour, Q. D. Bickle, T. N. C. Wells, T. Spangenberg and J. Keiser, *PLoS Neglected Trop. Dis.*, 2014, **8**, e2610.
31 R. A. Paveley and Q. D. Bickle, *Parasite Immunol.*, 2013, **35**, 302–313.
32 M.-H. Abdulla, D. S. Ruelas, B. Wolff, J. Snedecor, K.-C. Lim, F. Xu, A. R. Renslo, J. Williams, J. H. McKerrow and C. R. Caffrey, *PLoS Neglected Trop. Dis.*, 2009, **3**, e478.
33 N. R. Mansour and Q. D. Bickle, *PLoS Neglected Trop. Dis.*, 2010, **4**, e795.
34 N. Patocka and P. Ribeiro, *Mol. Biochem. Parasitol.*, 2013, **187**, 32–42.
35 M. Darmon, S. Al Awabdh, M.-B. Emerit and J. Masson, *Prog. Mol. Biol. Transl. Sci.*, 2015, **132**, 97–126.
36 T. Schmidt, A. Bergner and T. Schwede, *Drug Discovery Today*, 2013, **19**, 890–897.
37 A. Penmatsa, K. H. Wang and E. Gouaux, *Nat. Struct. Mol. Biol.*, 2015, **22**, 506–508.
38 A. Penmatsa, K. H. Wang and E. Gouaux, *Nature*, 2013, **503**, 85–90.
39 M. Marxer, K. Ingram and J. Keiser, *Parasites Vectors*, 2012, **5**, 165.
40 L. Kamentsky, T. R. Jones, A. Fraser, M.-A. Bray, D. J. Logan, K. L. Madden, V. Ljosa, C. Rueden, K. W. Eliceiri and A. E. Carpenter, *Bioinformatics*, 2011, **27**, 1179–1180.
41 D. A. Benson, M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell and E. W. Sayers, *Nucleic Acids Res.*, 2013, **41**, D36–D42.
42 M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. G. Cassarino, M. Bertoni, L. Bordoli and T. Schwede, *Nucleic Acids Res.*, 2014, **42**, W252–W258.
43 R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Crystallogr.*, 1993, **26**, 283–291.
44 C. Colovos and T. O. Yeates, *Protein Sci.*, 1993, **2**, 1511–1519.
45 J. C. Shelley, A. Cholleti, L. L. Frye, J. R. Greenwood, M. R. Timlin and M. Uchimaya, *J. Comput.-Aided Mol. Des.*, 2007, **21**, 681–691.
46 J. L. Banks, H. S. Beard, Y. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz, J. R. Maple, R. Murphy, D. M. Philipp, M. P. Repasky, L. Y. Zhang, B. J. Berne, R. A. Friesner, E. Gallicchio and R. M. Levy, *J. Comput. Chem.*, 2005, **26**, 1752–1780.
47 OpenEye Scientific Software Inc. OMEGA v. 2.5.1. Available from: http://www.eyesopen.com, 2013.
48 A. Jakalian, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2002, **23**, 1623–1641.
49 OpenEye Scientific Software Inc. QUACPAC v. 1.6.3. Available from: http://www.eyesopen.com, 2013.
50 M. McGann, *J. Chem. Inf. Model.*, 2011, **51**, 578–596.
51 OpenEye Scientific Software Inc. OEDocking v. 3.0.1. Available from: http://www.eyesopen.com, 2012.
52 M. R. McGann, H. R. Almond, A. Nicholls, J. A. Grant and F. K. Brown, *Biopolymers*, 2003, **68**, 76–90.
53 M. McGann, *J. Comput.-Aided Mol. Des.*, 2012, **26**, 897–906.
54 N. Patocka, N. Sharma, M. Rashid and P. Ribeiro, *PLoS Pathog.*, 2014, **10**, e1003878.
55 F. Angelucci, A. E. Miele, G. Boumis, M. Brunori, D. Dimastrogiovanni and A. Bellelli, *Curr. Top. Med. Chem.*, 2011, **11**, 2012–2028.

# 5. DISCUSSÃO

Na presente tese, duas abordagens distintas foram exploradas para identificação de novos agentes esquistossomicidas. Na primeira abordagem (Artigo 1), modelos de QSAR binários foram construídos e validados para predição da atividade inibitória da *Sm*TGR, um alvo validado em esquistossomos. A partir dos modelos individuais, modelos de consenso e consenso rigoroso foram construídos e utilizados na VS de 150 mil compostos. Ao final deste processo, 29 compostos foram priorizados e adquiridos para avaliação biológica. Como resultado, dois novos *hits* representando novos *scaffolds* moleculares apresentaram atividade contra esquistossômulos e vermes adultos de *S. mansoni*. Na segunda abordagem (Artigo 2), dados genômicos (proteínas expressas em esquistossômulos e vermes adultos) de *S. mansoni* foram utilizados em uma busca por ortólogos em bases de dados de alvos terapêuticos de fármacos (DrugBank, TTD e STITCH). Como resultado, 215 fármacos foram preditos para interagir com proteínas do esquistossomo, dos quais 47 já tinham atividade esquistossomicida reportada na literatura. Em seguida, a análise de componentes principais (PCA, do inglês *Principal Component Analysis*) e *k-means* demonstrou que 115 fármacos estavam dentro do espaço químico de agentes esquistossomicidas conhecidos, atribuindo maior confiabilidade as predições. Alguns destes fármacos foram então selecionados para avaliação biológica e a paroxetina (PAR), um antidepressivo predito para inibir proteínas transportadoras de serotonina do *S. mansoni* (*Sm*SERTs), apresentou os resultados mais promissores (Artigo 3). Em paralelo a estas abordagens, também foram produzidos três artigos de revisão abordando os seguintes temas:

- Revisão dos principais avanços na identificação de compostos líderes para esquistossomose utilizando plataformas de HTS e ensaios *ex vivo* de alta vazão, assim como na exploração de estratégias computacionais para a VS de novos ligantes em bases de dados. Também foram destacadas as principais vantagens e desvantagens de cada abordagem computacional e sugeridas soluções com vista na integração de várias estratégias, as quais contribuem para otimizar os resultados da investigação e levar a fluxos de trabalho eficientes e economicamente viáveis (Artigo 4);
- Revisão dos produtos naturais com atividade esquistossomicida bem como sua integração com ferramentas computacionais (*e.g.* LBDD e SBDD) para acelerar a descoberta de novos compostos líderes (Artigo 5);

- Revisão dos principais avanços no reposicionamento *in silico* de fármacos para DTNs utilizando ferramentas e conceitos do âmbito da quimiogenômica bem como os progressos obtidos com o preparo e modelagem de dados genômicos, biológicos e químicos. Ao final, também foram apresentados alguns exemplos de sucesso e sugeridas soluções para as limitações técnicas existentes (Artigo 6).

Dada a heterogeneidade das abordagens utilizadas, os resultados serão discutidos de forma separada nas próximas seções.

### 5.1. Planejamento de novos candidatos a fármacos esquistossomicidas

No presente trabalho (Artigo 1), modelos de QSAR binários foram desenvolvidos utilizando o maior conjunto de dados de compostos com atividade inibitória sobre a enzima *Sm*TGR disponível na literatura (SIMEONOV et al., 2008; LEA et al., 2008), se mostraram robustos e preditivos e subsequentemente foram aplicados a VS para identificar novos compostos com atividade esquistossomicida.

Na primeira etapa do estudo, o conjunto de dados original foi submetido a uma rigorosa preparação (Figura 17) para evitar que erros experimentais e de compilação sejam perpetuados e/ou interfiram na preditividade dos modelos (FOURCHES; MURATOV; TROPSHA, 2010, 2015, 2016). Inicialmente, compostos que não podem ser adequadamente tratados por técnicas de quimioinformática (*i.e.* organometálicos e compostos inorgânicos) foram removidos. Manter esses compostos no conjunto de dados poderia interferir negativamente na qualidade dos modelos, visto suas atividades biológicas para não poderia ser completamente correlacionada. Em seguida, quimiotipos específicos como anéis aromáticos e grupos nitro e formas tautoméricas foram padronizados e sais/fragmentos removidos. Ao final desta tarefa, as estruturas foram inspecionadas para garantir que todas estivessem corretas e uma análise de duplicatas foi realizada para remover estruturas presentes duas ou mais vezes no conjunto de dados (FOURCHES; MURATOV; TROPSHA, 2010, 2015, 2016).

Após identificadas as duplicatas, a análise do potencial binário mostrou boa concordância entre 99,9% dos dados, ou seja, apenas 0,1% das duplicatas apresentavam anotações diferentes. Geralmente, procedimentos como análise da variação intra- e interlaboratorial das atividades biológicas de duplicatas podem ser utilizados para verificar a qualidade do conjunto de dados. Espera-se através desta análise que as duplicatas apresentem potencias (*e.g.* $IC_{50}$, $EC_{50}$ e Ki) semelhantes ou que elas estejam dentro do mesmo limiar de atividade. Duplicatas com atividades discordantes geralmente

são oriundas de erros experimentais ou de compilação e, portanto, devem ser removidas do conjunto de dados ou corrigidas manualmente (FOURCHES; MURATOV; TROPSHA, 2010, 2015, 2016).

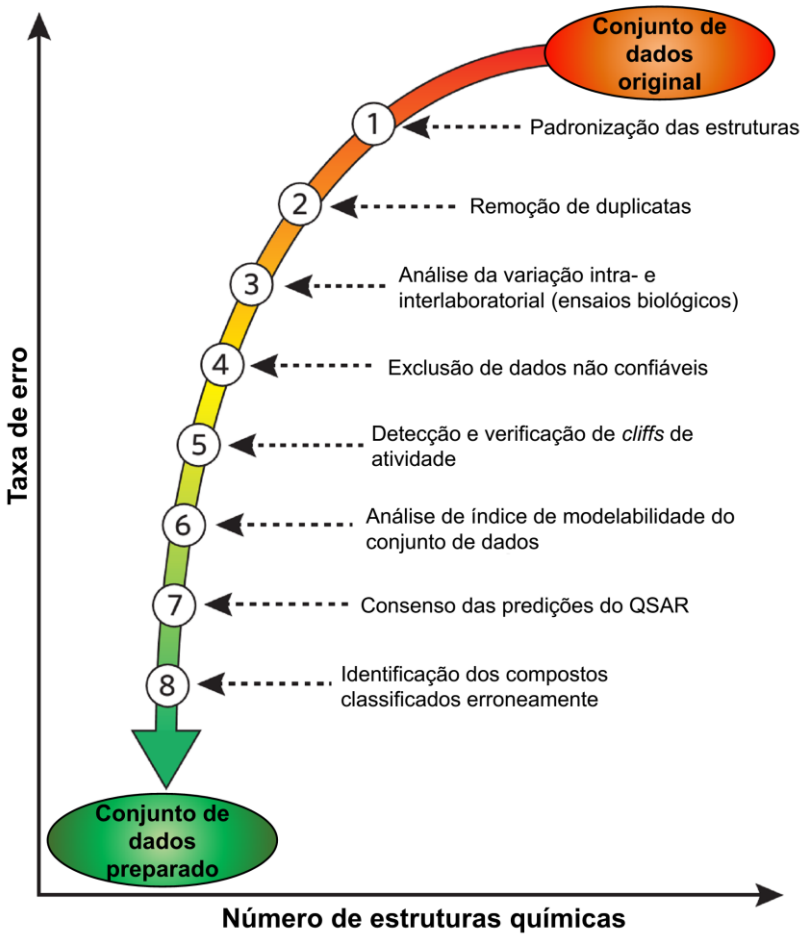



**Figura 18.** Representação do fluxo de trabalho para o preparo do conjunto de dados (Adaptado de: FOURCHES; MURATOV; TROPSHA, 2015).

Além disso, 14.580 compostos com dados de inibição inconclusivos (*i.e.* com apenas um ponto na curva de inibição ou que apresentam curva parcial de inibição) foram removidos do conjunto de dados. Geralmente, compostos não devem ser removidos do conjunto de dados de modelagem de QSAR sem explicação para isso. A remoção de *outliers* para melhorar os parâmetros estatísticos do modelo de QSAR é considerada uma estratégia de manipulação dos dados e essa melhoria pode não representar a real melhoria na preditividade externa do modelo (CHERKASOV et al., 2014). Todavia, a manutenção desses compostos no conjunto de dados poderia interferir negativamente na qualidade dos modelos de QSAR tendo em vista que suas atividades biológicas podem ser provenientes de erros experimentais.

O fluxo de trabalho utilizado para geração e validação dos modelos de QSAR foi totalmente complacente com as recomendações estabelecidas pela OECD: (i) o conjunto de dados utilizado apresenta atividade biológica definida; (ii) os modelos foram construídos utilizando algoritmos não ambíguos; (iii) o domínio de aplicabilidade (DA) foi definido para todos os modelos construído; (iv) os modelos apresentaram robustez e preditividade apropriadas; e (v) a interpretação mecanística foi realizada (OECD, 2004). Em consideração a esses princípios, alguns elementos chave aplicados durante o desenvolvimento e validação de modelos de QSAR serão discutidos a seguir.

Quanto a preditividade, é importante lembrar que o modelo de QSAR é resultado de uma representação teórica. Portanto seu desempenho e utilidade dependem diretamente de sua capacidade em predizer (generalizar), com elevada taxa de acerto, a atividade biológica de compostos que não foram utilizados para gerar os modelos (GRAMATICA, 2007; CHIRICO; GRAMATICA, 2011, 2012). Nesse trabalho, os modelos foram validados pela técnica de validação cruzada externa de 5-*fold*. Esta técnica divide o conjunto de dados total em cinco subgrupos de tamanhos idênticos. Um dos subgrupos (20% de todos os compostos) é definido como conjunto teste e os quatro conjuntos restantes são utilizados como conjunto de treinamento (80% de todo o conjunto de dados). Esse procedimento é repetido cinco vezes permitindo que cada um dos cinco subconjuntos seja usado como conjunto teste. Os modelos são gerados usando apenas o conjunto treinamento. É importante enfatizar que o conjunto teste nunca é empregado para construção dos modelos (ALVES et al., 2015a, 2015b).

Já a robustez dos modelos foi investigada utilizando o método de permutação da variável Y. Nesse procedimento, a variável Y é aleatorizada e modelos são gerados novamente. Caso modelos com variável Y aleatória também apresentem boa preditividade, ambos devem ser descartados, pois os descritores moleculares não descrevem bem Y, ou seja, não existe uma relação entre atividade biológica e descritores moleculares (RÜCKER; RÜCKER; MERINGER, 2007).

Do ponto de vista matemático, modelos de QSAR podem predizer a atividade biológica para qualquer composto, desde que cada descritor químico utilizado pelo modelo possa ser calculado para o mesmo. Entretanto, modelos de QSAR predizem a atividade de novos compostos a partir da interpolação dos compostos utilizados para construir o modelo. Assim, o DA, ou limite de similaridade, deve ser aplicado para impedir extrapolações extremas por parte do modelo quando esse prediz a atividade de compostos muito diferentes. Contudo, a análise e definição do DA representa um

requisito obrigatório para aceitabilidade de qualquer modelo (TROPSHA, 2010; CHERKASOV et al., 2014).

Embora as recomendações da OECD já sejam bastante criteriosas, outras estratégias como balanceamento do conjunto de dados, modelagem combinatória de descritores moleculares com diferentes métodos de aprendizado de máquina e construção de modelos por consenso também foram exploradas neste estudo visando obter modelos de QSAR otimizados.

Durante a construção dos modelos, a maioria dos métodos de aprendizado de máquina necessita de igual ponderação das classes em termos de número de compostos ou nível de importância. Quando o conjunto de dados utilizado para a modelagem é desbalanceado, estes métodos tendem a predizer maioria dos compostos como pertencentes a classe majoritária, resultando em um grande número de predições erradas para a classe minoritária (ZAKHAROV et al., 2014). Consequentemente, três proporções de balanceamento do conjunto de dados (1:1, 1:2 e 1:3) foram exploradas com o objetivo de reduzir o número de compostos na classe majoritária (não inibidores). Ao contrário das abordagens tradicionais, que convencionalmente selecionam estruturas de forma aleatória, abordagem utilizada aqui seleciona apenas as estruturas mais representativas da classe majoritária, preservando assim a diversidade química apresentada no conjunto de dados original.

Como resultado, os modelos de QSAR binários construídos com conjuntos de dados parcialmente desbalanceados (proporções de 1:2 e 1:3) apresentaram acentuada discrepância entre os valores de sensibilidade e especificidade, portanto não foram considerados em outras etapas do estudo. Por outro lado, modelos robustos e preditivos (conjunto de dados balanceado) foram obtidos pela combinação dos descritores Dragon, CDK, Morgan, AtomPair e MACCS com os métodos máquina de vetores de suporte (SVM, do inglês *support vector machine*) e RF (do inglês, *random forest*), apresentando valores de taxa de classificação correta (CCR) entre 0,81−0,85. Geralmente, modelos com valores de CCR maior que 0,70 são considerados preditivos (TROPSHA, 2010). O método de SVM utilizado mapeia os dados em um espaço altamente dimensional, utilizando uma função de Kernel radial. Em seguida, o algoritmo procura encontrar um plano de separação ótima entre duas classes através da maximização da margem entre os pontos mais próximos, conhecidos como suporte de vetores (Figura 18A). Em outras palavras, um modelo obtido por SVM representa os compostos (descritores) como pontos num espaço e os mapeia de maneira que cada classe (*i.e.* inibidores e não inibidores) seja

dividida por um espaço claro que seja tão amplo quanto possível. (VAPNIK, 2000). Já o método de RF é uma técnica de classificação baseada na construção de diversas arvores de decisão. O resultado de todas predições das árvores é agregado para se obter uma predição final (Figura 18B) (BREIMAN, 2001).

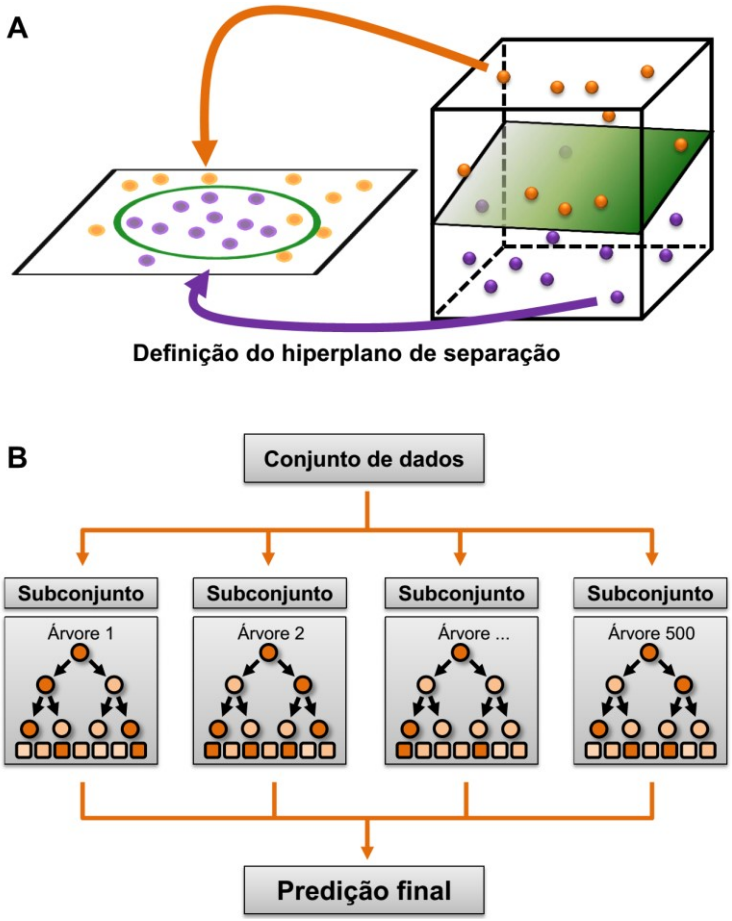



**Figura 19.** Representação do plano de separação gerado por SVM (**A**) e do consenso das predições obtidas pelas árvores de decisão no método de RF (**B**).

Após a construção e validação dos individuais, modelos de consenso e consenso rigoroso foram construídos (CCR iguais a 0,87 e 0,91, respectivamente). Modelos por consenso ou consenso rigoroso são bastante recomendados no âmbito da quimioinformática, dada a crescente evidência que a qualidade nas predições obtidas durante a VS são geralmente mais precisas e confiáveis (ZHU et al., 2008; TROPSHA, 2010; BRAGA et al., 2014b). O modelo de consenso pode ser gerado pela computação da média das predições individuais de cada modelo, ou seja, pela combinação de vários modelos de QSAR individuais. Já o consenso rigoroso requer que o mesmo composto esteja dentro do DA de todos os modelos utilizados para sua construção. Formalmente, os compostos com probabilidade predita superior a 0,5 são classificados como inibidores

e aqueles com a predição inferior a 0,5 são considerados como não-inibidores. Obviamente quando a predição for mais próxima de 1 ou 0, maior é a concordância entre todos os modelos e consequentemente sua confiabilidade para classificar compostos como inibidores ou não-inibidores, respectivamente.

Após construídos os modelos de consenso, ambos foram aplicados à VS de novos inibidores da *Sm*TGR utilizando três bases de dados com 50.000 compostos cada, (DIVERSet$^{TM}$-CL, DIVERSet$^{TM}$-EXP e PremiumSet) disponíveis na base de dados ChemBridge. Ao final deste processo, 29 compostos foram priorizados e adquiridos para avaliação biológica em esquistossômulos e vermes adultos de *S. mansoni*. Os ensaios foram realizados em colaboração com o Dr. Floriano Paes Silva Junior do Laboratório de Bioquímica Experimental e Computacional de Fármacos (LaBECFar) da Instituto Oswaldo Cruz. O método empregado baseia-se na quantificação da motilidade e efeitos fenotípicos em esquistossômulos e vermes adultos de *S. mansoni* utilizando um sistema de alta vazão baseado em bioimageamento por microscopia (PAVELEY et al., 2012; MELO-FILHO et al., 2016).

Como resultado, dois novos *hits* representando novos *scaffolds* moleculares apresentaram concentração necessária para inibir 50% da motilidade máxima ($EC_{50}$) ≤ 3,5 μM para esquistossômulos e ≤ 6,0 μM para fêmeas de vermes adultos, baixa citotoxicidade em células WSS-1 de mamíferos ($IC_{50}$ > 16 μM) e baixa reatividade com cisteíno proteases como a papaína ($IC_{50}$ > 100 μM). Além disso, ambos os *hits* promoveram fenótipos semelhantes em esquistossômulos, caracterizados pela a formação de vacúolos, arredondamento, escurecimento, e ruptura interna das larvas. Tais fenótipos também foram observados em vermes expostos ao composto OLT (*i.e.* um inibidor conhecido da *Sm*TGR) e durante o *knockout* do gene da *Sm*TGR (KUNTZ et al., 2007), sugerindo que o alvo destes compostos possa realmente ser a *Sm*TGR. Portanto é bem provável que estes *hits* tenham sido capazes de atingir seu alvo biológico dentro do esquistossomo depois de atravessar várias membranas biológicas e resistir à degradação por enzimas de detoxificação. Assim, um *hit* vindo de um ensaio fenotípico tem maior valor biológico quando comparado com um ensaio bioquímico simples (ZANELLA; LORENS; LINK, 2010).

### 5.2. Reposicionamento de fármacos para esquistossomose

Neste trabalho (Artigo 2), fármacos aprovados para uso clínico foram reposicionados *in silico* para o tratamento da esquistossomose baseando-se no conceito

de que "alvos biológicos semelhantes interagem com ligantes semelhantes" (KLABUNDE, 2007; ROGNAN, 2007). A exploração de alvos biológicos de esquistossomos com ortólogos em mamíferos tem sido evitada durante décadas visando eliminar possíveis problemas de seletividade e efeitos adversos durante o processo de descoberta de fármacos. No entanto, a situação é radicalmente diferente quando esta proteína do esquistossomo possui homologia com um alvo terapêutico de fármaco. Neste caso, dificuldade inicial pode ser transformada em oportunidade uma vez que seus ortólogos podem fornecer evidências de drogabilidade (BEGHYN et al., 2011; NJOROGE et al., 2014).

O termo drogabilidade (do inglês, *druggability*) é utilizado para avaliar a capacidade de um determinado alvo biológico interagir com um ligante de forma que sua modulação ou inibição seja capaz de produzir um efeito terapêutico (HOPKINS; GROOM, 2002b). De acordo com Crowther e colaboradores, proteínas de helmintos são consideradas drogáveis quando apresentam: ≥ 80% de cobertura para com o alvo terapêutico de fármaco, identidade sequencial ≥ 80% e valor esperado < $10^{-10}$ (CROWTHER et al., 2010). Todavia, estes valores são considerados bastante restritivos. No presente estudo, o conceito de drogabilidade foi reestruturado e a homologia entre as proteínas de *S. mansoni* e alvos terapêuticos de fármacos foi considerada satisfatória quando se obteve: ≥ 80% de cobertura para com o alvo terapêutico de fármaco, pontuação ≥ 0,8 (STITCH) ou valor esperado < $10^{-20}$ (TTD e DrugBank) e ≥ 60% de regiões funcionais conservadas.

Após aplicados estes critérios de inclusão e exclusão de ortólogos, 49 proteínas do verme que apresentaram baixa ou moderada identidade sequencial com ortólogos nas bases de dados, mas que apresentavam resíduos funcionais conservados, puderam ser incluídas na lista de potencias alvos de fármacos no esquistossomo. Além disso, 215 fármacos foram preditos para ter atividade esquistossomicida, destes 168 fármacos ainda não tinham atividade reportada na literatura e 47 fármacos já tinha atividade esquistossomicida reportada, validando assim metodologia.

Após a predição dos fármacos, propriedades físico-químicas e estruturais foram utilizadas para mapear o espaço químico de compostos com atividade esquistossomicida descrita na literatura, *decoys* e dos 168 fármacos preditos nesta abordagem utilizando análise de PCA integrada ao método de agrupamento *k-means*. A partir desta análise, foram reconhecidos padrões ou regiões específicas no espaço cartesiano 2D constituído predominantemente por agentes esquistossomicidas

conhecidos (espaço químico de compostos ativos) ou *decoys* (espaço químico de compostos inativos). Também verificou-se que 115 fármacos preditos (68,5%) estavam dentro do espaço químico de compostos esquistossomicidas conhecidos, atribuindo assim maior confiabilidade às predições. Em outras palavras, foi possível reduzir um espaço químico com alta dimensionalidade (*e.g.* descritores moleculares) para um espaço com dimensão menor, tornando-o mais interpretável para extrair informações biológicas importantes (BON; WALDMANN, 2010; LACHANCE et al., 2012).

Para verificar a confiabilidade das predições, os *hits* virtuais também foram comparados com *hits* experimentais identificados durante duas triagens *ex vivo* de alto rendimento (Figura 19) (ABDULLA et al., 2009; PANIC et al., 2015). Panic e colaboradores concluíram recentemente uma triagem *in vitro* com 1.600 fármacos aprovados pela FDA (do inglês, *Food and Drug Administration*). Embora os ensaios biológicos tenham sido conduzidos utilizando uma biblioteca maior de compostos, 21 *hits* virtuais estavam presentes entre os 121 *hits* experimentais. Em 2009, Abdulla e colaboradores desenvolveram um estudo semelhante, levando a identificação de 105 *hits*, dos quais 13 *hits* também foram preditos pela *in silico* desenvolvida na presente tese.

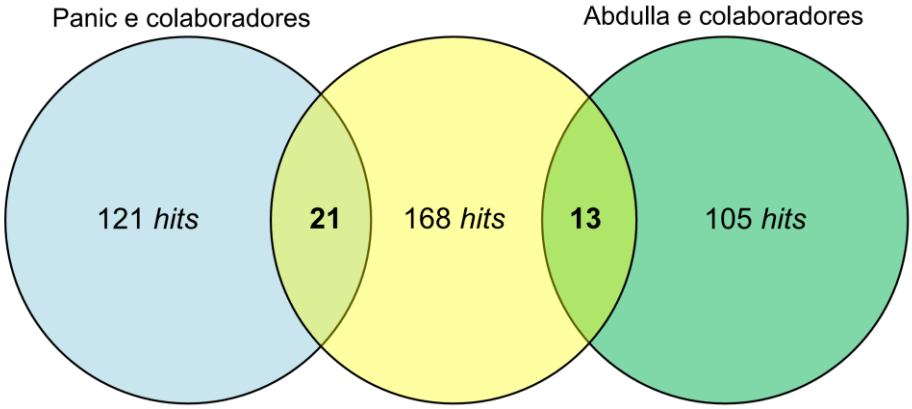



**Figura 20.** Diagrama de Venn representando o número de *hits* identificados nesta abordagem e sobrepostos em outros estudos reportados na literatura.

Durante a análise dos resultados, também foi possível constatar que 61 fármacos preditos pela abordagem *in silico* também poderiam afetar o controle da função muscular e motilidade do esquistossomo uma vez estando associados a alvos bioquímicos envolvidos no transporte de neurotransmissores e íons, ou enzimas envolvidas indiretamente no metabolismo e transporte $Ca^{2+}$. Diversos processos comportamentais estão associados com o sistema muscular de esquistossomos, tais como: motilidade, migração, emparelhamento e reprodução, digestão e excreção. Fármacos que

interrompem uma ou mais destas funções são considerados estrategicamente interessantes pelo fato de poderem interferir com a vida normal do verme e, consequentemente, provocar a sua morte. Além disso, o parâmetro padrão para avaliar a atividade esquistossomicida de fármacos ou compostos consiste na avaliação *ex vivo* da motilidade do verme por meio de microscopia (MARCELLINO et al., 2012; PAVELEY et al., 2012).

Após concluídas as análises *in silico*, alguns fármacos (*i.e.* anlodipino, alendronato, bupropiona, tramadol, paroxetina, venlafaxina e PAR) foram selecionados para avaliação biológica experimental *ex vivo* em colaboração com o Dr. Floriano Paes Silva Junior e equipe. Entre os fármacos testados, a PAR apresentou os resultados mais promissores, gerando um efeito pronunciado sobre a viabilidade de esquistossômulos de *S. mansoni* ($EC_{50}$ = 2,5 μM após 72 horas de incubação) em um ensaio com resazurina (MANSOUR; BICKLE, 2010). O efeito sobre vermes adultos de *S. mansoni* foi caracterizado por aumento de motilidade de até 9 vezes em fêmeas e 2 vezes em machos imediatamente após a exposição a baixas concentrações de PAR (≤10 μM para fêmeas e ≤ 20 μM para machos). Após 24 horas de incubação, uma redução acentuada da motilidade ($EC_{50}$ = 5,1 μM e 9,9 μM para machos e fêmeas, respectivamente após 24 horas de incubação) foi observada, condizendo plenamente com o mecanismo hipotético de inibição das *Sm*SERTs. Estudos da década de 1980 sugerem que as SERTs de platelmintos são responsáveis por inúmeros processos fisiológicos, tais como a modulação da função neuromuscular e do estímulo à absorção de glicose, quebra de glicogênio e excreção de lactato (MANSOUR, 1985; RAHMAN; METTRICK; PODESTA, 1985; BOYLE; ZAIDE; YOSHINO, 2000; BOYLE; YOSHINO, 2005). Esta foi a primeira evidência experimental de um inibidor de SERTs com atividade *ex vivo* em *S. mansoni*.

Em seguida, estudos de modelagem comparativa e docagem molecular foram realizados para determinar a estrutura 3D e predizer o modo de ligação e afinidade de interação da PAR nos sítios ativos das *Sm*SERTs e seu ortólogo em humanos (hSERT). Os estudos de docagem molecular podem ser realizados em duas etapas. Na primeira etapa, um algoritmo de busca explora todos os modos de ligação possíveis do ligante no sítio ativo da proteína. Em seguida, uma função de pontuação, baseada em energia de afinidade de interação ou não, atribui uma pontuação a cada uma das poses geradas para o ligante e as ordena conforme suas afinidades pelo sítio ativo (LIONTA et al., 2014; FERREIRA et al., 2015; CERQUEIRA et al., 2015). Após concluídas estas etapas, constatou-se que a PAR provavelmente apresenta modo de ligação e afinidade de

interação semelhante em ambas as proteínas, embora a *Sm*SERT-A tenha o sítio ativo maior e mais hidrofóbico quando comparado ao sítio ativo da hSERT. Estas informações poderão ser utilizadas como bases estruturais para o planejamento e síntese de novos análogos da PAR mais potentes e seletivos (BEGHYN et al., 2011; NJOROGE et al., 2014).

# 6. CONCLUSÕES

Na presente tese de doutorado foram apresentados estudos envolvendo a integração de métodos *in silico* e *ex vivo* para o reposicionamento de fármacos e planejamento e descoberta de novos compostos ativos contra *S. mansoni*. Diversas estratégias foram utilizadas, com destaque para modelos de QSAR binários construídos utilizando métodos de aprendizado de máquina, quimiogenômica, identificação de proteínas homólogas a alvos terapêuticos de fármacos em bases de dados, estudos de docagem molecular e ensaios biológicos *ex vivo* de alta vazão em esquistossômulos e vermes adultos de *S. mansoni*.

No planejamento e descoberta de novos compostos esquistossomicidas, foram gerados e validados modelos de QSAR robustos e preditivos para predição da inibição da enzima *Sm*TGR. A interpretação mecanística do melhor modelo gerado (combinação de descritores Morgan com método de aprendizado de máquina *Random Forest*) através dos mapas de probabilidade predita foi útil para compreender a relação entre grupos e/ou fragmentos estruturais com contribuição positiva e negativa para inibição da *Sm*TGR e estudar o mecanismo de inibição de diversas classes de inibidores. Os modelos gerados por consenso e consenso rigoroso foram úteis para a triagem virtual e seleção de 29 compostos, entre os quais dois *hits* apresentaram $EC_{50} \leq 3{,}5$ μM para esquistossômulos e ≤ 6,0 μM para fêmeas de vermes adultos nos ensaios *ex vivo*. A exposição a estes compostos gerou fenótipos semelhantes aos observados com o oltipraz, fármaco conhecido como inibidor de *Sm*TGR, e também ao fenótipo observado durante o *knockout* do gene da *Sm*TGR, sugerindo que a inibição da *Sm*TGR seja seu provável mecanismo de ação. Além disso, ambos compostos apresentaram baixa citotoxicidade em células de mamíferos WSS-1 ($IC_{50} > 16$ μM) e baixa reatividade com cisteíno-proteases ($IC_{50} > 100$ μM) e portanto, representam *scaffolds* moleculares atraentes para futuros estudos de otimização estrutural.

No reposicionamento de fármacos, diferentes critérios de inclusão e exclusão (*e.g.* valor esperado < $10^{-20}$, cobertura do alinhamento par a par entre sequências ≥ 80% e ≥ 60% de resíduos de aminoácidos funcionais conservados) foram considerados para a identificação de alvos terapêuticos de fármacos homólogos a proteínas expressas em esquistossômulos e adultos de *S. mansoni*. Através desta análise, 15 fármacos aprovados para uso clínico em humanos foram preditos com potencial atividade esquistossomicida. Em seguida, uma análise do espaço químico com agentes esquistossomicidas conhecidos

e *decoys* utilizando PCA e *k-means* permitiu transformar um espaço com alta dimensionalidade (*e.g.* descritores moleculares) em um espaço com dimensão menor, tornando-o mais interpretável para extrair informações biológicas importantes. Utilizando esta abordagem, foi possível demonstrar que 115 fármacos preditos estavam dentro do espaço químico de compostos esquistossomicidas conhecidos, atribuindo assim maior confiabilidade às predições. Destes, 61 apresentam potencial atividade no sistema neuromuscular do esquistossomo, uma vez sendo preditos para interagir com proteínas transportadoras de neurotransmissores, canais iônicos ou enzimas envolvidas indiretamente neste processo. Ao final das análises *in silico*, alguns fármacos foram selecionados e adquiridos para avaliação experimental, a partir da qual constatou-se que a paroxetina (PAR), predita anteriormente como inibidora de *Sm*SERT, apresentava os efeitos esquistossomicidas mais promissores ($EC_{50}$ = 2,5 μM, 5,1 μM e 9,9 μM para esquistossômulos, machos e fêmeas, respectivamente). Vale ressaltar que o efeito sobre vermes adultos foi caracterizado por aumento de motilidade de até 9 vezes em fêmeas e 2 vezes em machos imediatamente após a exposição a baixas concentrações de PAR, o condizendo com o mecanismo de ação aqui proposto. Ao final, estudos de modelagem comparativa e docagem molecular demonstraram que a PAR possui modo de ligação e afinidade semelhantes para ambas as SERTs (*i.e.* humana e de esquistossomo). Todavia, uma diferença de volume significativa entre os sítios ativos sugere que as *Sm*SERTs possam acomodar ligantes mais volumosos e hidrofóbicos.

As informações aqui destacadas representam bases moleculares e estruturais importantes para o planejamento de novos análogos mais potentes e seletivos. Tanto os estudos de reposicionamento (otimização estrutural da PAR), quanto o estudo de planejamento de novos compostos inibidores da *Sm*TGR serão conduzidos em outros trabalhos. Novas séries de compostos serão planejadas com auxílio de estudos de docagem e dinâmica molecular. Os compostos planejados serão sintetizados e avaliados em ensaios bioquímicos e fenotípicos com o intuito de se elucidar as relações entre estrutura e atividade. Após a otimização das propriedades biológicas em estudo, os compostos mais promissores serão avaliados em experimentos *in vivo*.

# 7. REFERÊNCIAS BIBLIOGRÁFICAS


ABDULLA, M.-H.; RUELAS, D. S.; WOLFF, B.; SNEDECOR, J.; LIM, K.-C.; XU, F.; RENSLO, A. R.; WILLIAMS, J.; MCKERROW, J. H.; CAFFREY, C. R. Drug discovery for schistosomiasis: hit and lead compounds identified in a library of known drugs by medium-throughput phenotypic screening. **PLoS Neglected Tropical Diseases**, v. 3, n. 7, p. e478, 2009.

AGÜERO, F.; AL-LAZIKANI, B.; ASLETT, M.; BERRIMAN, M.; BUCKNER, F. S.; CAMPBELL, R. K.; CARMONA, S.; CARRUTHERS, I. M.; CHAN, A. W. E.; CHEN, F.; CROWTHER, G. J.; DOYLE, M. A.; HERTZ-FOWLER, C.; HOPKINS, A. L.; MCALLISTER, G.; NWAKA, S.; OVERINGTON, J. P.; PAIN, A.; PAOLINI, G. V; PIEPER, U.; RALPH, S. A.; RIECHERS, A.; ROOS, D. S.; SALI, A.; SHANMUGAM, D.; SUZUKI, T.; VAN VOORHIS, W. C.; VERLINDE, C. L. M. J. Genomic-scale prioritization of drug targets: the TDR Targets database. **Nature reviews. Drug discovery**, v. 7, n. 11, p. 900–907, 2008.

ALGER, H. M.; WILLIAMS, D. L. The disulfide redox system of Schistosoma mansoni and the importance of a multifunctional enzyme, thioredoxin glutathione reductase. **Molecular and Biochemical Parasitology**, v. 121, n. 1, p. 129–139, 2002.

ALTSCHUL, S. F.; MADDEN, T. L.; SCHÄFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic acids research**, v. 25, n. 17, p. 3389–3402, 1997.

ALVES, V. M.; MURATOV, E.; FOURCHES, D.; STRICKLAND, J.; KLEINSTREUER, N.; ANDRADE, C. H.; TROPSHA, A. Predicting chemically-induced skin reactions. Part I: QSAR models of skin sensitization and their application to identify potentially hazardous compounds. **Toxicology and Applied Pharmacology**, v. 284, n. 2, p. 262–272, 2015a.

ALVES, V. M.; MURATOV, E.; FOURCHES, D.; STRICKLAND, J.; KLEINSTREUER, N.; ANDRADE, C. H.; TROPSHA, A. Predicting chemically-induced skin reactions. Part II: QSAR models of skin permeability and the relationships between skin permeability and skin sensitization. **Toxicology and Applied Pharmacology**, v. 284, n. 2, p. 273–280, 2015b.

ALVES, V. M.; MURATOV, E. N.; CAPUZZI, S. J.; POLITI, R.; LOW, Y.; BRAGA, R. C.; ZAKHAROV, A. V.; SEDYKH, A.; MOKSHYNA, E.; FARAG, S.; ANDRADE, C. H.; KUZ'MIN, V. E.; FOURCHES, D.; TROPSHA, A. Alarms about structural alerts. **Green chemistry**, v. 18, n. 16, p. 4348–4360, 2016.

ANGELUCCI, F.; BASSO, A.; BELLELLI, A.; BRUNORI, M.; PICA MATTOCCIA, L.; VALLE, C. The anti-schistosomal drug praziquantel is an adenosine antagonist. **Parasitology**, v. 134, n. 9, p. 1215–1221, 2007.

ANGELUCCI, F.; DIMASTROGIOVANNI, D.; BOUMIS, G.; BRUNORI, M.; MIELE, A. E.; SACCOCCIA, F.; BELLELLI, A. Mapping the catalytic cycle of Schistosoma mansoni thioredoxin glutathione reductase by X-ray crystallography. **The Journal of Biological Chemistry**, v. 285, n. 42, p. 32557–32567, 2010.

ASHBURN, T. T.; THOR, K. B. Drug repositioning: identifying and developing new uses for existing drugs. **Nature Reviews Drug Discovery**, v. 3, n. 8, p. 673–683, 2004.

ASHKENAZY, H.; EREZ, E.; MARTZ, E.; PUPKO, T.; BEN-TAL, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. **Nucleic acids research**, v. 38, p. W529-33, 2010.

AVORN, J. The $2.6 Billion Pill — Methodologic and Policy Considerations. **New England Journal of Medicine**, v. 372, n. 20, p. 1877–1879, 2015.

BARSOUM, R. S.; ESMAT, G.; EL-BAZ, T. Human Schistosomiasis: Clinical Perspective: Review. **Journal of Advanced Research**, v. 4, n. 5, p. 433–444, 2013.

BECKMANN, S.; LEUTNER, S.; GOUIGNARD, N.; DISSOUS, C.; GREVELDING, C. G. Protein kinases as potential targets for novel anti-schistosomal strategies.

**Current pharmaceutical design**, v. 18, n. 24, p. 3579–3594, 2012.

BEGHYN, T. B.; CHARTON, J.; LEROUX, F.; LACONDE, G.; BOURIN, A.; COS, P.; MAES, L.; DEPREZ, B. Drug to Genome to Drug: Discovery of New Antiplasmodial Compounds. **Journal of Medicinal Chemistry**, v. 54, n. 9, p. 3222–3240, 2011.

BERGQUIST, R.; AL-SHERBINY, M.; BARAKAT, R.; OLDS, R. Blueprint for schistosomiasis vaccine development. **Acta tropica**, v. 82, n. 2, p. 183–192, 2002.

BERMAN, J. Miltefosine, an FDA-approved drug for the "orphan disease", leishmaniasis. **Expert Opinion on Orphan Drugs**, v. 3, n. 6, p. 727–735, 2015.

BERRIMAN, M.; HAAS, B. J.; LOVERDE, P. T.; WILSON, R. A.; DILLON, G. P.; CERQUEIRA, G. C.; MASHIYAMA, S. T.; AL-LAZIKANI, B.; ANDRADE, L. F.; ASHTON, P. D.; ASLETT, M. A; BARTHOLOMEU, D. C.; BLANDIN, G.; CAFFREY, C. R.; COGHLAN, A.; COULSON, R.; DAY, T. A; DELCHER, A.; DEMARCO, R.; DJIKENG, A.; EYRE, T.; GAMBLE, J. A; GHEDIN, E.; GU, Y.; HERTZ-FOWLER, C.; HIRAI, H.; HIRAI, Y.; HOUSTON, R.; IVENS, A.; JOHNSTON, D. A; LACERDA, D.; MACEDO, C. D.; MCVEIGH, P.; NING, Z.; OLIVEIRA, G.; OVERINGTON, J. P.; PARKHILL, J.; PERTEA, M.; PIERCE, R. J.; PROTASIO, A. V; QUAIL, M. A; RAJANDREAM, M.-A.; ROGERS, J.; SAJID, M.; SALZBERG, S. L.; STANKE, M.; TIVEY, A. R.; WHITE, O.; WILLIAMS, D. L.; WORTMAN, J.; WU, W.; ZAMANIAN, M.; ZERLOTINI, A.; FRASER-LIGGETT, C. M.; BARRELL, B. G.; EL-SAYED, N. M. The genome of the blood fluke Schistosoma mansoni. **Nature**, v. 460, n. 7253, p. 352–358, 2009.

BOGITSH, B. J.; CARTER, C. E.; OELTMANN, T. N. **Human Parasitology**. 4. ed. Waltham: Elsevier, 2013.

BOMBELLES, T.; COAKER, H. Neglected tropical disease research: rethinking the drug discovery model. **Future Medicinal Chemistry**, v. 7, n. 6, p. 693–700, 2015.

BON, R. S.; WALDMANN, H. Bioactivity-guided navigation of chemical space. **Accounts of Chemical Research**, v. 43, n. 8, p. 1103–1114, 2010.

BOROS, D. L. Delayed Hypersensitivity-Type Granuloma Formation and Dermal Reaction Induced and Elicited by a Soluble Factor Isolated From Schistosoma Mansoni Eggs. **Journal of Experimental Medicine**, v. 132, n. 3, p. 488–507, 1970.

BOTTIEAU, E.; CLERINX, J.; DE VEGA, M. R.; VAN DEN ENDEN, E.; COLEBUNDERS, R.; VAN ESBROECK, M.; VERVOORT, T.; VAN GOMPEL, A.; VAN DEN ENDE, J. Imported Katayama fever: Clinical and biological features at presentation and during treatment. **Journal of Infection**, v. 52, n. 5, p. 339–345, 2006.

BOYLE, J. P.; YOSHINO, T. P. Serotonin-induced muscular activity in Schistosoma mansoni larval stages: importance of 5-HT transport and role in daughter sporocyst production. **The Journal of parasitology**, v. 91, n. 3, p. 542–550, 2005.

BOYLE, J. P.; ZAIDE, J. V; YOSHINO, T. P. Schistosoma mansoni: effects of serotonin and serotonin receptor antagonists on motility and length of primary sporocysts in vitro. **Experimental parasitology**, v. 94, n. 4, p. 217–226, 2000.

BRAGA, R. C.; ALVES, V. M.; SILVA, A. C.; NASCIMENTO, M. N.; SILVA, F. C.; LIAO, L. M.; ANDRADE, C. H. Virtual screening strategies in medicinal chemistry: the state of the art and current challenges. **Current Topics in Medicinal Chemistry**, v. 14, n. 16, p. 1899–1912, 2014a.

BRAGA, R. C.; ALVES, V. M.; SILVA, M. F. B.; MURATOV, E.; FOURCHES, D.; TROPSHA, A.; ANDRADE, C. H. Tuning HERG out: antitarget QSAR models for drug development. **Current Topics in Medicinal Chemistry**, v. 14, n. 11, p. 1399–1415, 2014b.

BRASCHI, S.; CURWEN, R. S.; ASHTON, P. D.; VERJOVSKI-ALMEIDA, S.; WILSON, A. The tegument surface membranes of the human blood parasite Schistosoma mansoni: A proteomic analysis after differential extraction. **Proteomics**, v. 6, n. 5, p. 1471–1482, 2006.

BRASCHI, S.; WILSON, R. A. Proteins exposed at the adult schistosome surface revealed by biotinylation. **Molecular & Cellular Proteomics**, v. 5, n. 2, p. 347–356, 2006.

BREDEL, M.; JACOBY, E. Chemogenomics: an emerging strategy for rapid target and drug discovery. **Nature reviews. Genetics**, v. 5, n. 4, p. 262–275, 2004.

BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.

BROUWERS, J. F.; SKELLY, P. J.; VAN GOLDE, L. M.; TIELENS, A. G. Studies on phospholipid turnover argue against sloughing of tegumental membranes in adult Schistosoma mansoni. **Parasitology**, v. 119, n. 3, p. 287–294, 1999.

BUCKLE, D. R.; ERHARDT, P. W.; GANELLIN, C. R.; KOBAYASHI, T.; PERUN, T. J.; PROUDFOOT, J.; SENN-BILFINGER, J. Glossary of Terms Used in Medicinal Chemistry Part II. In: **Annual Reports in Medicinal Chemistry**. [s.l: s.n.]. v. 48p. 387–418.

BUSCAGLIA, C. A.; KISSINGER, J. C.; AGÜERO, F. Neglected Tropical Diseases in the Post-Genomic Era. **Trends in Genetics**, v. 31, n. 10, p. 539–555, 2015.

CAFFREY, C. R.; SECOR, W. E. Schistosomiasis: from drug deployment to drug development. **Current opinion in infectious diseases**, v. 24, n. 5, p. 410–417, 2011.

CARDOSO, F. C.; MACEDO, G. C.; GAVA, E.; KITTEN, G. T.; MATI, V. L.; DE MELO, A. L.; CALIARI, M. V.; ALMEIDA, G. T.; VENANCIO, T. M.; VERJOVSKI-ALMEIDA, S.; OLIVEIRA, S. C. Schistosoma mansoni Tegument Protein Sm29 Is Able to Induce a Th1-Type of Immune Response and Protection against Parasite Infection. **PLoS Neglected Tropical Diseases**, v. 2, n. 10, p. e308, 2008.

CARON, P. R.; MULLICAN, M. D.; MASHAL, R. D.; WILSON, K. P.; SU, M. S.; MURCKO, M. A. Chemogenomic approaches to drug discovery. **Current opinion in chemical biology**, v. 5, p. 464–470, 2001.

CARVALHO, O. DO S.; COELHO, P. M. Z.; LENZI, H. L. **Schistosoma mansoni & Esquistossomose: uma visão multidisciplinar**. 1. ed. Rio de Janeiro: FioCruz, 2008.

CERQUEIRA, N. M. F. S. A.; GESTO, D.; OLIVEIRA, E. F.; SANTOS-MARTINS, D.; BRÁS, N. F.; SOUSA, S. F.; FERNANDES, P. A.; RAMOS, M. J. Receptor-based virtual screening protocol for drug discovery. **Archives of Biochemistry and Biophysics**, v. 582, p. 56–67, 2015.

CHATELAIN, E.; IOSET, J.-R. Drug discovery and development for neglected diseases: the DNDi model. **Drug design, development and therapy**, v. 5, n. 5, p. 175–181, 2011.

CHAVALI, A. K.; WHITTEMORE, J. D.; EDDY, J. A.; WILLIAMS, K. T.; PAPIN, J. A. Systems analysis of metabolism in the pathogenic trypanosomatid Leishmania major. **Molecular systems biology**, v. 4, p. 177, 2008.

CHEEVER, A. W.; DUVALL, R. H. Schistosoma japonicum: migration of adult worm pairs within the mesenteric veins of mice. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 76, n. 5, p. 641–645, 1982.

CHEN, L.; RAO, K. V. N.; HE, Y. X.; RAMASWAMY, K. Skin-stage schistosomula of Schistosoma mansoni produce an apoptosis-inducing factor that can cause apoptosis of T cells. **The Journal of Biological Chemistry**, v. 277, n. 37, p. 34329–34335, 2002.

CHEN, X.; JI, Z. L.; CHEN, Y. Z. TTD: Therapeutic Target Database. **Nucleic acids research**, v. 30, n. 1, p. 412–415, 2002.

CHEN, X.; LIU, M.; GILSON, M. K. BindingDB: a web-accessible molecular recognition database. **Combinatorial Chemistry & High Throughput Screening**, v. 4, n. 8, p. 719–725, 2001.

CHERKASOV, A.; MURATOV, E. N.; FOURCHES, D.; VARNEK, A.; BASKIN, I. I.; CRONIN, M.; DEARDEN, J.; GRAMATICA, P.; MARTIN, Y. C.; TODESCHINI, R.; CONSONNI, V.; KUZ'MIN, V. E.; CRAMER, R.; BENIGNI, R.; YANG, C.; RATHMAN, J.; TERFLOTH, L.; GASTEIGER, J.; RICHARD, A.; TROPSHA, A. QSAR modeling: where have you been? Where are you going to? **Journal of Medicinal Chemistry**, v. 57, n. 12, p. 4977–5010, 2014.

CHIRICO, N.; GRAMATICA, P. Real external predictivity of QSAR models: how to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. **Journal of Chemical Information and Modeling**, v. 51, n. 9, p. 2320–2335, 2011.

CHIRICO, N.; GRAMATICA, P. Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need

for scatter plot inspection. **Journal of Chemical Information and Modeling**, v. 52, n. 8, p. 2044–2058, 2012.

CHIUMIENTO, L.; BRUSCHI, F. Enzymatic antioxidant systems in helminth parasites. **Parasitology Research**, v. 105, n. 3, p. 593–603, 2009.

COLLEY, D. G.; BUSTINDUY, A. L.; SECOR, W. E.; KING, C. H. Human schistosomiasis. **The Lancet**, v. 383, n. 9936, p. 2253–2264, 2014.

CORRÊA SOARES, J. B. R.; MENEZES, D.; VANNIER-SANTOS, M. A; FERREIRA-PEREIRA, A.; ALMEIDA, G. T.; VENANCIO, T. M.; VERJOVSKI-ALMEIDA, S.; ZISHIRI, V. K.; KUTER, D.; HUNTER, R.; EGAN, T. J.; OLIVEIRA, M. F. Interference with hemozoin formation represents an important mechanism of schistosomicidal action of antimalarial quinoline methanols. **PLoS neglected tropical diseases**, v. 3, n. 7, p. e477, 2009.

COULSON, P. S. The radiation-attenuated vaccine against schistosomes in animal models: paradigm for a human vaccine? **Advances in parasitology**, v. 39, p. 271–336, 1997.

COUTO, F. F. B.; COELHO, P. M. Z.; ARAÚJO, N.; KUSEL, J. R.; KATZ, N.; JANNOTTI-PASSOS, L. K.; MATTOS, A. C. A. Schistosoma mansoni: a method for inducing resistance to praziquantel using infected Biomphalaria glabrata snails. **Memórias do Instituto Oswaldo Cruz**, v. 106, n. 2, p. 153–157, 2011.

COWAN, N.; KEISER, J. Repurposing of anticancer drugs: in vitro and in vivo activities against Schistosoma mansoni. **Parasites & vectors**, v. 13, n. 8, p. 417, 2015.

CROWTHER, G. J.; SHANMUGAM, D.; CARMONA, S. J.; DOYLE, M. A; HERTZ-FOWLER, C.; BERRIMAN, M.; NWAKA, S.; RALPH, S. A; ROOS, D. S.; VAN VOORHIS, W. C.; AGÜERO, F. Identification of attractive drug targets in neglected-disease pathogens using an in silico approach. **PLoS neglected tropical diseases**, v. 4, n. 8, p. e804, 2010.

CUPIT, P. M.; CUNNINGHAM, C. What is the mechanism of action of praziquantel and how might resistance strike? **Future Medicinal Chemistry**, v. 7, n. 6, p.

701–705, 2015.

DALTON, J. .; MULCAHY, G. Parasite vaccines — a reality? **Veterinary Parasitology**, v. 98, n. 1–3, p. 149–167, 2001.

DEARDEN, J. C.; CRONIN, M. T. D.; KAISER, K. L. E. How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR). **SAR and QSAR in Environmental Research**, v. 20, n. 3–4, p. 241–266, 2009.

DEBNATH, B.; AL-MAWSAWI, L. Q.; NEAMATI, N. Are we living in the end of the blockbuster drug era? **Drug News & Perspectives**, v. 23, n. 10, p. 670–684, 2010.

DIETZEL, J.; HIRZMANN, J.; PREIS, D.; SYMMONS, P.; KUNZ, W. Ferritins of Schistosoma mansoni: sequence comparison and expression in female and male worms. **Molecular and Biochemical Parasitology**, v. 50, n. 2, p. 245–254, 1992.

DIMASI, J. A.; GRABOWSKI, H. G.; HANSEN, R. W. The Cost of Drug Development. **New England Journal of Medicine**, v. 372, n. 20, p. 1972–1972, 2015.

DOHERTY, J. F.; MOODY, A. H.; WRIGHT, S. G. Katayama fever: an acute manifestation of schistosomiasis. **British Medical Journal**, v. 313, n. 7064, p. 1071–1072, 1996.

DRÖGE, W. Free radicals in the physiological control of cell function. **Physiological Reviews**, v. 82, n. 1, p. 47–95, 2002.

EKINS, S.; LAGE DE SIQUEIRA-NETO, J.; MCCALL, L.-I.; SARKER, M.; YADAV, M.; PONDER, E. L.; KALLEL, E. A.; KELLAR, D.; CHEN, S.; ARKIN, M.; BUNIN, B. A.; MCKERROW, J. H.; TALCOTT, C. Machine Learning Models and Pathway Genome Data Base for Trypanosoma cruzi Drug Discovery. **PLOS Neglected Tropical Diseases**, v. 9, n. 6, p. e0003878, 2015.

EKINS, S.; WILLIAMS, A. J.; KRASOWSKI, M. D.; FREUNDLICH, J. S. In silico repositioning of approved drugs for rare and neglected diseases. **Drug discovery today**, v. 16, n. 7–8, p. 298–310, 2011.

ELBAZ, T.; ESMAT, G. Hepatic and Intestinal Schistosomiasis: Review. **Journal of Advanced Research**, v. 4, n. 5, p. 445–452, 2013.

EVAN SECOR, W. Water-based interventions for schistosomiasis control. **Pathogens and Global Health**, v. 108, n. 5, p. 246–254, 2014.

FALLON, P. G.; DOENHOFF, M. J. Drug-resistant schistosomiasis: resistance to praziquantel and oxamniquine induced in Schistosoma mansoni in mice is drug specific. **The American Journal of Tropical Medicine and Hygiene**, v. 51, n. 1, p. 83–88, 1994.

FALLON, P. G.; STURROCK, R. F.; NIANG, A. C.; DOENHOFF, M. J. Short report: diminished susceptibility to praziquantel in a Senegal isolate of Schistosoma mansoni. **The American Journal of Tropical Medicine and Hygiene**, v. 53, n. 1, p. 61–62, 1995.

FARRAR, J. **Manson's Tropical Infectious Diseases**. 23. ed. New York: Elsevier, 2014.

FERREIRA, L.; DOS SANTOS, R.; OLIVA, G.; ANDRICOPULO, A. Molecular Docking and Structure-Based Drug Design Strategies. **Molecules**, v. 20, n. 7, p. 13384–13421, 2015.

FISHER, J. A.; COTTINGHAM, M. D.; KALBAUGH, C. A. Peering into the pharmaceutical "pipeline": investigational drugs, clinical trials, and industry priorities. **Social Science & Medicine (1982)**, v. 131, p. 322–330, 2015.

FOURCHES, D.; MURATOV, E.; TROPSHA, A. Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research. **Journal of Chemical Information and Modeling**, v. 50, n. 7, p. 1189–1204, 2010.

FOURCHES, D.; MURATOV, E.; TROPSHA, A. Curation of chemogenomics data. **Nature Chemical Biology**, v. 11, n. 8, p. 535–535, 2015.

FOURCHES, D.; MURATOV, E.; TROPSHA, A. Trust, but Verify II: A Practical Guide to Chemogenomics Data Curation. **Journal of Chemical Information and Modeling**, v. 56, n. 7, p. 1243–1252, 2016.

GAULTON, A.; BELLIS, L. J.; BENTO, A. P.; CHAMBERS, J.; DAVIES, M.; HERSEY, A.; LIGHT, Y.; MCGLINCHEY, S.; MICHALOVICH, D.; AL-

LAZIKANI, B.; OVERINGTON, J. P. ChEMBL: a large-scale bioactivity database for drug discovery. **Nucleic Acids Research**, v. 40, n. Database issue, p. D1100–D1107, 2012.

GELMEDIN, V.; DISSOUS, C.; GREVELDING, C. G. Re-positioning protein-kinase inhibitors against schistosomiasis. **Future Medicinal Chemistry**, v. 7, n. 6, p. 737–752, 2015.

GLASER, F.; PUPKO, T.; PAZ, I.; BELL, R. E.; BECHOR-SHENTAL, D.; MARTZ, E.; BEN-TAL, N. ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. **Bioinformatics (Oxford, England)**, v. 19, n. 1, p. 163–164, 2003.

GNANASEKAR, M.; SALUNKHE, A. M.; MALLIA, A. K.; HE, Y. X.; KALYANASUNDARAM, R. Praziquantel Affects the Regulatory Myosin Light Chain of Schistosoma mansoni. **Antimicrobial Agents and Chemotherapy**, v. 53, n. 3, p. 1054–1060, 2009.

GÖNNERT, R.; ANDREWS, P. Praziquantel, a new board-spectrum antischistosomal agent. **Zeitschrift für Parasitenkunde (Berlin, Germany)**, v. 52, n. 2, p. 129–150, 1977.

GRAMATICA, P. Principles of QSAR models validation: internal and external. **QSAR & Combinatorial Science**, v. 26, n. 5, p. 694–701, 2007.

GREENBERG, R. M. Are Ca2+ channels targets of praziquantel action? **International Journal for Parasitology**, v. 35, n. 1, p. 1–9, 2005.

GRIMES, J. E. T.; CROLL, D.; HARRISON, W. E.; UTZINGER, J.; FREEMAN, M. C.; TEMPLETON, M. R. The relationship between water, sanitation and schistosomiasis: a systematic review and meta-analysis. **PLoS Neglected Tropical Diseases**, v. 8, n. 12, p. e3296, 2014.

GRIMES, J. E. T.; CROLL, D.; HARRISON, W. E.; UTZINGER, J.; FREEMAN, M. C.; TEMPLETON, M. R. The roles of water, sanitation and hygiene in reducing schistosomiasis: a review. **Parasites & Vectors**, v. 8, n. 1, p. 156, 2015.

GRYSEELS, B. Schistosomiasis. **Infectious Disease Clinics of North America**, v. 26, n. 2, p. 383–397, 2012.

GRYSEELS, B.; POLMAN, K.; CLERINX, J.; KESTENS, L. Human schistosomiasis. **The Lancet**, v. 368, n. 9541, p. 1106–1118, 2006.

HAN, Y.; FU, Z.; HONG, Y.; ZHANG, M.; HAN, H.; LU, K.; YANG, J.; LI, X.; LIN, J. Inhibitory Effects and Analysis of RNA Interference on Thioredoxin Glutathione Reductase Expression in Schistosoma japonicum. **Journal of Parasitology**, v. 100, n. 4, p. 463–469, 2014.

HAN, Z.-G.; BRINDLEY, P. J.; WANG, S.-Y.; CHEN, Z. Schistosoma genomics: new perspectives on schistosome biology and host-parasite interaction. **Annual Review of Genomics and Human Genetics**, v. 10, n. 2, p. 211–240, 2009.

HAY, M.; THOMAS, D. W.; CRAIGHEAD, J. L.; ECONOMIDES, C.; ROSENTHAL, J. Clinical development success rates for investigational drugs. **Nature Biotechnology**, v. 32, n. 1, p. 40–51, 2014.

HONG, Z.; KOSMAN, D. J.; THAKUR, A.; REKOSH, D.; LOVERDE, P. T. Identification and purification of a second form of Cu/Zn superoxide dismutase from Schistosoma mansoni. **Infection and Immunity**, v. 60, n. 9, p. 3641–3651, 1992.

HONG, Z.; LOVERDE, P. T.; THAKUR, A.; HAMMARSKJÖLD, M. L.; REKOSH, D. Schistosoma mansoni: a Cu/Zn superoxide dismutase is glycosylated when expressed in mammalian cells and localizes to a subtegumental region in adult schistosomes. **Experimental Parasitology**, v. 76, n. 2, p. 101–114, 1993.

HOPKINS, A. L.; GROOM, C. R. The druggable genome. **Nature Reviews Drug Discovery**, v. 1, n. 9, p. 727–730, 2002a.

HOPKINS, A. L.; GROOM, C. R. The druggable genome. **Nature Reviews. Drug Discovery**, v. 1, n. 9, p. 727–730, 2002b.

HUANG, H.; DAY, L.; CASS, C. L.; BALLOU, D. P.; WILLIAMS, C. H.; WILLIAMS, D. L. Investigations of the Catalytic Mechanism of Thioredoxin Glutathione Reductase from Schistosoma mansoni. **Biochemistry**, v. 50, n. 26, p. 5870–5882, 2011.

HUANG, H.-H.; RIGOUIN, C.; WILLIAMS, D. L. The redox biology of schistosome parasites and applications for drug development. **Current Pharmaceutical Design**, v. 18, n. 24, p. 3595–3611, 2012.

HUGHES, J.; REES, S.; KALINDJIAN, S.; PHILPOTT, K. Principles of early drug discovery. **British Journal of Pharmacology**, v. 162, n. 6, p. 1239–1249, 2011.

HURLE, M. R.; YANG, L.; XIE, Q.; RAJPAL, D. K.; SANSEAU, P.; AGARWAL, P. Computational drug repositioning: from data to therapeutics. **Clinical Pharmacology and Therapeutics**, v. 93, n. 4, p. 335–341, 2013.

INGRAM, K.; ELLIS, W.; KEISER, J. Antischistosomal Activities of Mefloquine-Related Arylmethanols. **Antimicrobial Agents and Chemotherapy**, v. 56, n. 6, p. 3207–3215, 2012.

ISMAIL, M. M.; TAHA, S. A.; FARGHALY, A. M.; EL-AZONY, A. S. Laboratory induced resistance to praziquantel in experimental schistosomiasis. **Journal of the Egyptian Society of Parasitology**, v. 24, n. 3, p. 685–695, 1994.

ISMAIL, M.; METWALLY, A.; FARGHALY, A.; BRUCE, J.; TAO, L. F.; BENNETT, J. L. Characterization of isolates of Schistosoma mansoni from Egyptian villagers that tolerate high doses of praziquantel. **The American Journal of Tropical Medicine and Hygiene**, v. 55, n. 2, p. 214–218, 1996.

JA, I. MEga-trials for blockbusters. **JAMA**, v. 309, n. 3, p. 239–240, 2013.

JOHNSON, K. S.; HARRISON, G. B. L.; LIGHTOWLERS, M. W.; O'HOY, K. L.; COUGLE, W. G.; DEMPSTER, R. P.; LAWRENCE, S. B.; VINTON, J. G.; HEATH, D. D.; RICKARD, M. D. Vaccination against ovine cysticercosis using a defined recombinant antigen. **Nature**, v. 338, n. 6216, p. 585–587, 1989.

JONES, M. K.; GOBERT, G. N.; ZHANG, L.; SUNDERLAND, P.; MCMANUS, D. P. The cytoskeleton and motor proteins of human schistosomes and their roles in surface maintenance and host-parasite interactions. **BioEssays**, v. 26, n. 7, p. 752–765, 2004.

JOY MACALINO, S. Y.; VIJAYAKUMAR GOSU, B.; SUNHYE HONG, B.; SUN CHOI, B. Role of computer-aided drug design in modern drug discovery. **Archives of Pharmacal Research**, v. 38, n. 9, p. 1686–1701, 2015.

KANTOROW, M.; HAWSE, J. R.; COWELL, T. L.; BENHAMED, S.; PIZARRO, G. O.; REDDY, V. N.; HEJTMANCIK, J. F. Methionine sulfoxide reductase A is

important for lens cell viability and resistance to oxidative stress. **Proceedings of the National Academy of Sciences of the United States of America**, v. 101, n. 26, p. 9654–9659, 2007.

KATSUNO, K.; BURROWS, J. N.; DUNCAN, K.; VAN HUIJSDUIJNEN, R. H.; KANEKO, T.; KITA, K.; MOWBRAY, C. E.; SCHMATZ, D.; WARNER, P.; SLINGSBY, B. T. Hit and lead criteria in drug discovery for infectious diseases of the developing world. **Nature Reviews Drug Discovery**, v. 14, n. 11, p. 751–758, 2015.

KAWAMOTO, F.; SHOZAWA, A.; KUMADA, N.; KOJIMA, K. Possible roles of cAMP and Ca2+ in the regulation of miracidial transformation in Schistosoma mansoni. **Parasitology research**, v. 75, n. 5, p. 368–374, 1989.

KEISER, J.; CHOLLET, J.; XIAO, S.-H.; MEI, J.-Y.; JIAO, P.-Y.; UTZINGER, J.; TANNER, M. Mefloquine—An Aminoalcohol with Promising Antischistosomal Properties in Mice. **PLoS Neglected Tropical Diseases**, v. 3, n. 1, p. e350, 2009.

KING, C. H. Toward the elimination of schistosomiasis. **The New England journal of medicine**, v. 360, n. 2, p. 106–109, 2009.

KLABUNDE, T. Chemogenomic approaches to drug discovery: similar receptors bind similar ligands. **British Journal of Pharmacology**, v. 152, n. 1, p. 5–7, 2007.

KLEBE, G. **Drug Design: Methodology, Concepts, and Mode-of-Action**. 1. ed. Berlin, Heidelberg: Springer, 2013.

KNOX, C.; LAW, V.; JEWISON, T.; LIU, P.; LY, S.; FROLKIS, A.; PON, A.; BANCO, K.; MAK, C.; NEVEU, V.; DJOUMBOU, Y.; EISNER, R.; GUO, A. C.; WISHART, D. S. DrugBank 3.0: a comprehensive resource for “omics” research on drugs. **Nucleic Acids Research**, v. 39, p. D1035-41, 2011.

KOCKAR, H.; ALPER, M.; SAHIN, T.; HACIISMAILOGLU, M. S. Co-Fe films: effect of Fe content on their properties. **Journal of nanoscience and nanotechnology**, v. 10, n. 11, p. 7639–7642, 2010.

KRISTENSEN, D. M.; WOLF, Y. I.; MUSHEGIAN, A. R.; KOONIN, E. V. Computational methods for Gene Orthology inference. **Briefings in Bioinformatics**, v. 12, n. 5, p. 379–391, 2011.

KUHN, M.; SZKLARCZYK, D.; FRANCESCHINI, A.; VON MERING, C.; JENSEN, L. J.; BORK, P. STITCH 3: zooming in on protein-chemical interactions. **Nucleic Acids Research**, v. 40, n. Database issue, p. D876-80, 2012.

KUNTZ, A. N.; DAVIOUD-CHARVET, E.; SAYED, A. A.; CALIFF, L. L.; DESSOLIN, J.; ARNÉR, E. S. J.; WILLIAMS, D. L. Thioredoxin glutathione reductase from Schistosoma mansoni: an essential parasite enzyme and a key drug target. **PLoS Medicine**, v. 4, n. 6, p. e206, 2007.

LACHANCE, H.; WETZEL, S.; KUMAR, K.; WALDMANN, H. Charting, navigating, and populating natural product chemical space for drug discovery. **Journal of Medicinal Chemistry**, v. 55, n. 13, p. 5989–6001, 2012.

LAMBERTUCCI, J. R.; SERUFO, J. C.; GERSPACHER-LARA, R.; RAYES, A. A. .; TEIXEIRA, R.; NOBRE, V.; ANTUNES, C. M. . Schistosoma mansoni: assessment of morbidity before and after control. **Acta Tropica**, v. 77, n. 1, p. 101–109, 2000.

LANDRUM, G. **RDKit: Open-source cheminformatics**, 2014.

LAVECCHIA, A. Machine-learning approaches in drug discovery: methods and applications. **Drug Discovery Today**, v. 20, n. 3, p. 318–331, 2015.

LEA, W. A.; JADHAV, A.; RAI, G.; SAYED, A. A.; CASS, C. L.; INGLESE, J.; WILLIAMS, D. L.; AUSTIN, C. P.; SIMEONOV, A. A 1,536-well-based kinetic HTS assay for inhibitors of Schistosoma mansoni thioredoxin glutathione reductase. **Assay and Drug Development Technologies**, v. 6, n. 4, p. 551–555, 2008.

LEE, E. F.; FAIRLIE, W. D. Repurposing apoptosis-inducing cancer drugs to treat schistosomiasis. **Future Medicinal Chemistry**, v. 7, n. 6, p. 707–711, 2015.

LI, J.; ZHENG, S.; CHEN, B.; BUTTE, A. J.; SWAMIDASS, S. J.; LU, Z. A survey of current trends in computational drug repositioning. **Briefings in Bioinformatics**, v. 17, n. 1, p. 2–12, 2016.

LI, L. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. **Genome Research**, v. 13, n. 9, p. 2178–2189, 2003.

LI, T.; ZINIEL, P. D.; HE, P.; KOMMER, V. P.; CROWTHER, G. J.; HE, M.; LIU, Q.; VAN VOORHIS, W. C.; WILLIAMS, D. L.; WANG, M.-W. High-throughput

screening against thioredoxin glutathione reductase identifies novel inhibitors with potential therapeutic value for schistosomiasis. **Infectious Diseases of Poverty**, v. 4, n. 1, p. 40, 2015.

LIMA, S. F.; VIEIRA, L. Q.; HARDER, A.; KUSEL, J. R. Effects of culture and praziquantel on membrane fluidity parameters of adult Schistosoma mansoni. **Parasitology**, v. 109, n. 1, p. 57–64, 1994.

LIONTA, E.; SPYROU, G.; VASSILATIS, D. K.; COURNIA, Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. **Current Topics in Medicinal Chemistry**, v. 14, n. 16, p. 1923–1938, 2014.

LIU, F.; HU, W.; CUI, S. J.; CHI, M.; FANG, C. Y.; WANG, Z. Q.; YANG, P. Y.; HAN, Z. G. Insight into the host-parasite interplay by proteomic study of host proteins copurified with the human parasite, Schistosoma japonicum. **Proteomics**, v. 7, n. 3, p. 450–462, 2007.

LIU, F.; LU, J.; HU, W.; WANG, S. Y.; CUI, S. J.; CHI, M.; YAN, Q.; WANG, X. R.; SONG, H. D.; XU, X. N.; WANG, J. J.; ZHANG, X. L.; ZHANG, X.; WANG, Z. Q.; XUE, C. L.; BRINDLEY, P. J.; MCMANUS, D. P.; YANG, P. Y.; FENG, Z.; CHEN, Z.; HAN, Z. G. New perspectives on host-parasite interplay by comparative transcriptomic and proteomic analyses of Schistosoma japonicum. **PLoS Pathogens**, v. 2, n. 4, p. 268–281, 2006.

LIU, Y.-X.; WU, W.; LIANG, Y.-J.; JIE, Z.-L.; WANG, H.; WANG, W.; HUANG, Y.-X. New Uses for Old Drugs: The Tale of Artemisinin Derivatives in the Elimination of Schistosomiasis Japonica in China. **Molecules**, v. 19, n. 9, p. 15058–15074, 2014.

LOMBARDINO, J. G.; LOWE, J. A. The role of the medicinal chemist in drug discovery--then and now. **Nature Reviews. Drug Discovery**, v. 3, n. 10, p. 853–862, 2004.

MAGARIÑOS, M. P.; CARMONA, S. J.; CROWTHER, G. J.; RALPH, S. A.; ROOS, D. S.; SHANMUGAM, D.; VAN VOORHIS, W. C.; AGÜERO, F. TDR targets: A chemogenomics resource for neglected diseases. **Nucleic Acids Research**, v. 40, p. D1118–D1127, 2012.

MAHMOUD, A. A. F. **Schistosomiasis**. 1. ed. London: Imperial College, 2001.

MAIZELS, R. M.; BUNDY, D. A.; SELKIRK, M. E.; SMITH, D. F.; ANDERSON, R. M. Immunological modulation and evasion by helminth parasites in human populations. **Nature**, v. 365, n. 6449, p. 797–805, 1993.

MANSOUR, N. R.; BICKLE, Q. D. Comparison of microscopy and Alamar blue reduction in a larval based assay for schistosome drug screening. **PLoS neglected tropical diseases**, v. 4, n. 8, p. e795, 2010.

MANSOUR, N. R.; PAVELEY, R.; GARDNER, J. M. F.; BELL, A. S.; PARKINSON, T.; BICKLE, Q. High Throughput Screening Identifies Novel Lead Compounds with Activity against Larval, Juvenile and Adult Schistosoma mansoni. **PLOS Neglected Tropical Diseases**, v. 10, n. 4, p. e0004659, 2016.

MANSOUR, T. E. Serotonin Receptors in Parasitic Worms. **Advances in Parasitology**, v. 23, p. 1–36, 1985.

MARCELLINO, C.; GUT, J.; LIM, K. C.; SINGH, R.; MCKERROW, J.; SAKANARI, J. WormAssay: a novel computer application for whole-plate motion-based screening of macroscopic parasites. **PLoS neglected tropical diseases**, v. 6, n. 1, p. e1494, 2012.

MCKERROW, J. J. Invasion of skin by schistosome cercariae: some neglected facts. **Trends in Parasitology**, v. 19, n. 2, p. 66–68, 2003.

MEI, H.; HIRAI, H.; TANAKA, M.; HONG, Z.; REKOSH, D.; LOVERDE, P. T. Schistosoma mansoni: cloning and characterization of a gene encoding cytosolic Cu/Zn superoxide dismutase. **Experimental Parasitology**, v. 80, n. 2, p. 250–259, 1995.

MEI, H.; LOVERDE, P. T. Schistosoma mansoni: the developmental regulation and immunolocalization of antioxidant enzymes. **Experimental Parasitology**, v. 86, n. 1, p. 69–78, 1997.

MEI, H.; THAKUR, A.; SCHWARTZ, J.; LO VERDE, P. T. Expression and characterization of glutathione peroxidase activity in the human blood fluke Schistosoma mansoni. **Infection and Immunity**, v. 64, n. 10, p. 4299–4306, 1996.

MELMAN, S. D.; STEINAUER, M. L.; CUNNINGHAM, C.; KUBATKO, L. S.;

MWANGI, I. N.; WYNN, N. B.; MUTUKU, M. W.; KARANJA, D. M. S.; COLLEY, D. G.; BLACK, C. L.; SECOR, W. E.; MKOJI, G. M.; LOKER, E. S. Reduced susceptibility to praziquantel among naturally occurring Kenyan isolates of Schistosoma mansoni. **PLoS Neglected Tropical Diseases**, v. 3, n. 8, p. e504, 2009.

MELO-FILHO, C. C.; DANTAS, R. F.; BRAGA, R. C.; NEVES, B. J.; SENGER, M. R.; VALENTE, W. C. G.; REZENDE-NETO, J. M.; CHAVES, W. T.; MURATOV, E. N.; PAVELEY, R. A.; FURNHAM, N.; KAMENTSKY, L.; CARPENTER, A. E.; SILVA-JUNIOR, F. P.; ANDRADE, C. H. QSAR-Driven Discovery of Novel Chemical Scaffolds Active against Schistosoma mansoni. **Journal of Chemical Information and Modeling**, v. 56, n. 7, p. 1357–1372, 2016.

MERZ, J.; RINGE, D.; REYNOLDS, C. H. **Drug Design: Structure- and Ligand-Based Approaches**. 1. ed. Cambridge: Cambridge University Press, 2010.

MESTRES, J. Computational chemogenomics approaches to systematic knowledge-based drug discovery. **Current opinion in drug discovery & development**, v. 7, n. 3, p. 304–313, 2004.

MEYER, T.; SEKLJIC, H.; FUCHS, S.; BOTHE, H.; SCHOLLMEYER, D.; MICULKA, C. Taste, A New Incentive to Switch to (R)-Praziquantel in Schistosomiasis Treatment. **PLoS Neglected Tropical Diseases**, v. 3, n. 1, p. e357, 2009.

MITCHELL, J. B. O. Machine learning methods in chemoinformatics. **Wiley Interdisciplinary Reviews. Computational Molecular Science**, v. 4, n. 5, p. 468–481, 2014.

MKOJI, G. M.; SMITH, J. M.; PRICHARD, R. K. Antioxidant systems in Schistosoma mansoni: evidence for their role in protection of the adult worms against oxidant killing. **International Journal for Parasitology**, v. 18, n. 5, p. 667–673, 1988.

MS. **Ministério da Saúde. Vigilância da esquistossomose mansoni: Diretrizes técnicas**. Disponível em: <http://portalsaude.saude.gov.br/index.php/o-ministerio/principal/secretarias/svs/esquistossomose>. Acesso em: 4 out. 2016.

MULLARD, A. New drugs cost US$2.6 billion to develop. **Nature Reviews Drug Discovery**, v. 13, n. 12, p. 877–877, 2014.

NCOKAZI, K. K.; EGAN, T. J. A colorimetric high-throughput β-hematin inhibition screening assay for use in the search for antimalarial compounds. **Analytical Biochemistry**, v. 338, n. 2, p. 306–319, 2005.

NEVES, D. P. **Parasitologia Humana**. 11. ed. São Paulo: Atheneu, 2005.

NJOROGE, M.; NJUGUNA, N. M.; MUTAI, P.; ONGARORA, D. S. B.; SMITH, P. W.; CHIBALE, K. Recent Approaches to Chemical Discovery and Development Against Malaria and the Neglected Tropical Diseases Human African Trypanosomiasis and Schistosomiasis. **Chemical Reviews**, v. 114, p. 11138–11163, 2014.

NOGI, T.; ZHANG, D.; CHAN, J. D.; MARCHANT, J. S. A Novel Biological Activity of Praziquantel Requiring Voltage-Operated Ca2+ Channel β Subunits: Subversion of Flatworm Regenerative Polarity. **PLoS Neglected Tropical Diseases**, v. 3, n. 6, p. e464, 2009.

OECD. **OECD principles for the validation, for regulatory purposes, of (Quantitative) Structure-Activity Relationship models**. Disponível em: <http://www.oecd.org/chemicalsafety/risk-assessment/37849783.pdf>. Acesso em: 1 out. 2015.

OKE, T. T.; MOSKOVITZ, J.; WILLIAMS, D. L. Characterization of the methionine sulfoxide reductases of Schistosoma mansoni. **The Journal of Parasitology**, v. 95, n. 6, p. 1421–1428, 2009.

OOMS, F. Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. **Current medicinal chemistry**, v. 7, n. 2, p. 141–158, 2000.

PAL, C.; BANDYOPADHYAY, U. Redox-active antiparasitic drugs. **Antioxidants & Redox Signaling**, v. 17, n. 4, p. 555–582, 2012.

PAMMOLLI, F.; MAGAZZINI, L.; RICCABONI, M. The productivity crisis in pharmaceutical R&D. **Nature Reviews. Drug Discovery**, v. 10, n. 6, p. 428–438, 2011.

PANIC, G.; VARGAS, M.; SCANDALE, I.; KEISER, J. Activity Profile of an FDA-

Approved Compound Library against Schistosoma mansoni. **PLOS Neglected Tropical Diseases**, v. 9, n. 7, p. e0003962, 2015.

PATOCKA, N.; RIBEIRO, P. Characterization of a serotonin transporter in the parasitic flatworm, Schistosoma mansoni: cloning, expression and functional analysis. **Molecular and biochemical parasitology**, v. 154, n. 2, p. 125–133, 2007.

PAVELEY, R.; MANSOUR, N. R.; HALLYBURTON, I.; BLEICHER, L. S.; BENN, A. E.; MIKIC, I.; GUIDI, A.; GILBERT, I. H.; HOPKINS, A. L.; BICKLE, Q. D. Whole Organism High-Content Screening by Label-Free, Image-Based Bayesian Classification for Parasitic Diseases. **PLoS Neglected Tropical Diseases**, v. 6, n. 7, p. e1762, 2012.

PEARCE, E. J.; MACDONALD, A. S. The immunobiology of schistosomiasis. **Nature Reviews. Immunology**, v. 2, n. 7, p. 499–511, 2002.

PEARCE, E. J.; SHER, A. Mechanisms of immune evasion in schistosomiasis. **Contributions to Microbiology and Immunology**, v. 8, p. 219–232, 1987.

PERBANDT, M.; NDJONKA, D.; LIEBAU, E. Protective Mechanisms of Helminths Against Reactive Oxygen Species are Highly Promising Drug Targets. **Current Medicinal Chemistry**, v. 21, n. 15, p. 1794–1808, 2014.

PÉREZ-SÁNCHEZ, R.; RAMAJO-HERNÁNDEZ, A.; RAMAJO-MARTÍN, V.; OLEAGA, A. Proteomic analysis of the tegument and excretory-secretory products of adult Schistosoma bovis worms. **Proteomics**, v. 6 Suppl 1, p. S226–S236, 2006.

PETERS, W.; PASVOL, G. **Atlas of Tropical Medicine and Parasitology**. 6nd. ed. New York: Elsevier Ltd, 2006.

PLOWRIGHT, A. T.; JOHNSTONE, C.; KIHLBERG, J.; PETTERSSON, J.; ROBB, G.; THOMPSON, R. A. Hypothesis driven drug design: improving quality and effectiveness of the design-make-test-analyse cycle. **Drug Discovery Today**, v. 17, n. 1–2, p. 56–62, 2012.

POLLASTRI, M. P. M.; CAMPBELL, R. R. K. Target repurposing for neglected diseases. **Future medicinal chemistry**, v. 3, n. 10, p. 1307–1315, 2011.

PRAST-NIELSEN, S.; HUANG, H. H.; WILLIAMS, D. L. Thioredoxin glutathione reductase: Its role in redox biology and potential as a target for drugs against

neglected diseases. **Biochimica et Biophysica Acta**, v. 1810, n. 12, p. 1262–1271, 2011.

PROTASIO, A. V; TSAI, I. J.; BABBAGE, A.; NICHOL, S.; HUNT, M.; ASLETT, M. A; DE SILVA, N.; VELARDE, G. S.; ANDERSON, T. J. C.; CLARK, R. C.; DAVIDSON, C.; DILLON, G. P.; HOLROYD, N. E.; LOVERDE, P. T.; LLOYD, C.; MCQUILLAN, J.; OLIVEIRA, G.; OTTO, T. D.; PARKER-MANUEL, S. J.; QUAIL, M. A; WILSON, R. A.; ZERLOTINI, A.; DUNNE, D. W.; BERRIMAN, M. A systematically improved high quality genome and transcriptome of the human blood fluke Schistosoma mansoni. **PLoS Neglected Tropical Diseases**, v. 6, n. 1, p. e1455, 2012.

RAHMAN, M. S.; METTRICK, D. F.; PODESTA, R. B. Schistosoma mansoni: effects of in vitro serotonin (5-HT) on aerobic and anaerobic carbohydrate metabolism. **Experimental parasitology**, v. 60, n. 1, p. 10–17, 1985.

RAI, G.; SAYED, A. A.; LEA, W. A.; LUECKE, H. F.; CHAKRAPANI, H.; PRAST-NIELSEN, S.; JADHAV, A.; LEISTER, W.; SHEN, M.; INGLESE, J.; AUSTIN, C. P.; KEEFER, L.; ARNÉR, E. S. J.; SIMEONOV, A.; MALONEY, D. J.; WILLIAMS, D. L.; THOMAS, C. J. Structure Mechanism Insights and the Role of Nitric Oxide Donation Guide the Development of Oxadiazole-2-Oxides as Therapeutic Agents against Schistosomiasis. **Journal of Medicinal Chemistry**, v. 52, n. 20, p. 6474–6483, 2009.

RAIES, A. B.; BAJIC, V. B. In silico toxicology: computational methods for the prediction of chemical toxicity. **Wiley Interdisciplinary Reviews: Computational Molecular Science**, v. 6, n. 2, p. 147–172, 2016.

RAY, D.; WILLIAMS, D. L. Characterization of the phytochelatin synthase of Schistosoma mansoni. **PLoS Neglected Tropical Diseases**, v. 5, n. 5, p. e1168, 2011.

RHEE, S. G.; CHAE, H. Z.; KIM, K. Peroxiredoxins: a historical overview and speculative preview of novel mechanisms and emerging concepts in cell signaling. **Free Radical Biology & Medicine**, v. 38, n. 12, p. 1543–1552, 2005.

RIBEIRO, F.; COELHO, P. M. Z.; VIEIRA, L. Q.; WATSON, D. G.; KUSEL, J. R. The effect of praziquantel treatment on glutathione concentration in

Schistosoma mansoni. **Parasitology**, v. 116, n. 3, p. 229–236, 1998.

RICK, N. G. **Drugs: From Discovery to Approval**. New Jersey: John Wiley & Sons, 2015.

ROGERS, D.; HAHN, M. Extended-connectivity fingerprints. **Journal of Chemical Information and Modeling**, v. 50, n. 5, p. 742–754, 2010.

ROGNAN, D. Chemogenomic approaches to rational drug design. **British Journal of Pharmacology**, v. 152, n. 1, p. 38–52, 2007.

ROJO-ARREOLA, L.; LONG, T.; ASARNOW, D.; SUZUKI, B. M.; SINGH, R.; CAFFREY, C. R. Chemical and Genetic Validation of the Statin Drug Target to Treat the Helminth Disease, Schistosomiasis. **PLoS ONE**, v. 9, n. 1, p. e87594, 2014.

ROSS, A. G.; MCMANUS, D. P.; FARRAR, J.; HUNSTMAN, R. J.; GRAY, D. J.; LI, Y. S. Neuroschistosomiasis. **Journal of Neurology**, v. 259, n. 1, p. 22–32, 2012.

ROSS, A. G. P.; BARTLEY, P. B.; SLEIGH, A. C.; OLDS, G. R.; LI, Y.; WILLIAMS, G. M.; MCMANUS, D. P. Schistosomiasis. **The New England journal of medicine**, v. 346, n. 16, p. 1212–1220, 2002.

ROSS, A. G.; VICKERS, D.; OLDS, G. R.; SHAH, S. M.; MCMANUS, D. P. Katayama syndrome. **Lancet Infectious Diseases**, v. 7, n. 3, p. 218–224, 2007.

RUAN, H.; TANG, X. D.; CHEN, M.-L.; JOINER, M.-L. A; SUN, G.; BROT, N.; WEISSBACH, H.; HEINEMANN, S. H.; IVERSON, L.; WU, C.-F.; HOSHI, T.; CHEN, M.-L.; JOINER, M. A; HEINEMANN, S. H. High-quality life extension by the enzyme peptide methionine sulfoxide reductase. **Proceedings of the National Academy of Sciences of the United States of America**, v. 99, n. 5, p. 2748–2753, 2002.

RÜCKER, C.; RÜCKER, G.; MERINGER, M. Y-randomization and its variants in QSPR/QSAR. **Journal of Chemical Information and Modeling**, v. 47, n. 6, p. 2345–2357, 2007.

SABAH, A. A.; FLETCHER, C.; WEBBE, G.; DOENHOFF, M. J. Schistosoma mansoni: Chemotherapy of infections of different ages. **Experimental Parasitology**, v. 61, n. 3, p. 294–303, 1986.

SAEED, M. E. M.; KRISHNA, S.; GRETEN, H. J.; KREMSNER, P. G.; EFFERTH, T. Antischistosomal activity of artemisinin derivatives in vivo and in patients. **Pharmacological Research**, v. 110, p. 216–226, 2016.

SALZET, M.; CAPRON, A.; STEFANO, G. B. Molecular crosstalk in host-parasite relationships: Schistosome- and leech-host interactions. **Parasitology Today**, v. 16, n. 12, p. 536–540, 2000.

SAYED, A. A.; COOK, S. K.; WILLIAMS, D. L. Redox balance mechanisms in Schistosoma mansoni rely on peroxiredoxins and albumin and implicate peroxiredoxins as novel drug targets. **The Journal of Biological Chemistry**, v. 281, n. 25, p. 17001–17010, 2006.

SAYED, A. A; WILLIAMS, D. L. Biochemical characterization of 2-cys peroxiredoxins from Schistosoma mansoni. **The Journal of Biological Chemistry**, v. 279, n. 25, p. 26159–26166, 2004.

SCHOLTE, R. G. C.; CARVALHO, O. S.; MALONE, J. B.; UTZINGER, J.; VOUNATSOU, P. Spatial distribution of biomphalaria spp., the intermediate host snails of schistosoma mansoni, in Brazil. **Geospatial Health**, v. 6, n. 3, p. 95–101, 2012.

SECOR, W. E.; MONTGOMERY, S. P.; CAFFREY, C. R.; WANG, W.; WANG, L.; LIANG, Y.-S. Schistosomiasis and its treatment. **Future Medicinal Chemistry**, v. 7, n. 6, p. 681–684, 2015.

SHALABY, K. A.; YIN, L.; THAKUR, A.; CHRISTEN, L.; NILES, E. G.; LOVERDE, P. T. Protection against Schistosoma mansoni utilizing DNA vaccination with genes encoding Cu/Zn cytosolic superoxide dismutase, signal peptide-containing superoxide dismutase and glutathione peroxidase enzymes. **Vaccine**, v. 22, n. 1, p. 130–136, 2003.

SIMEONOV, A.; JADHAV, A.; SAYED, A. A.; WANG, Y.; NELSON, M. E.; THOMAS, C. J.; INGLESE, J.; WILLIAMS, D. L.; AUSTIN, C. P. Quantitative high-throughput screen identifies inhibitors of the Schistosoma mansoni redox cascade. **PLoS Neglected Tropical Diseases**, v. 2, n. 1, p. e127, 2008.

SINGH, S.; MALIK, B. K.; SHARMA, D. K. Choke point analysis of metabolic

pathways in E.histolytica: a computational approach for drug target identification. **Bioinformation**, v. 2, n. 2, p. 68–72, 2007.

SONG, L.; LI, J.; XIE, S.; QIAN, C.; WANG, J.; ZHANG, W.; YIN, X.; HUA, Z.; YU, C. Thioredoxin Glutathione Reductase as a Novel Drug Target: Evidence from Schistosoma japonicum. **PLoS One**, v. 7, n. 2, p. e31456, 2012.

STELMA, F. F.; TALLA, I.; SOW, S.; KONGS, A.; NIANG, M.; POLMAN, K.; DEELDER, A. M.; GRYSEELS, B. Efficacy and side effects of praziquantel in an epidemic focus of Schistosoma mansoni. **The American Journal of Tropical Medicine and Hygiene**, v. 53, n. 2, p. 167–170, 1995.

STIREWALT, M. A.; HACKEY, J. R. Penetration of Host Skin by Cercariae of Schistosoma mansoni. I. Observed Entry into Skin of Mouse, Hamster, Rat, Monkey and Man. **The Journal of Parasitology**, v. 42, n. 6, p. 565–580, 1956.

SU, D.; NOVOSELOV, S. V; SUN, Q. A.; MOUSTAFA, M. E.; ZHOU, Y.; OKO, R.; HATFIELD, D. L.; GLADYSHEV, V. N. Mammalian selenoprotein thioredoxin-glutathione reductase: Roles in bisulfide bond formation and sperm maturation. **The Journal of Biological Chemistry**, v. 280, n. 28, p. 26491–26498, 2005.

SUN, Q.-A.; SU, D.; NOVOSELOV, S. V; CARLSON, B. A; HATFIELD, D. L.; GLADYSHEV, V. N. Reaction mechanism and regulation of mammalian thioredoxin/glutathione reductase. **Biochemistry**, v. 44, n. 44, p. 14528–14537, 2005.

TAKOUGANG, I.; MELI, J.; WABO PONÉ, J.; ANGWAFO, F. Community acceptability of the use of low-dose niclosamide (Bayluscide), as a molluscicide in the control of human schistosomiasis in Sahelian Cameroon. **Annals of Tropical Medicine and Parasitology**, v. 101, n. 6, p. 479–486, 2007.

TALELE, T. T.; KHEDKAR, S. A.; RIGBY, A. C. Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. **Current topics in medicinal chemistry**, v. 10, n. 1, p. 127–141, 2010.

TALLIMA, H.; EL RIDI, R. Praziquantel binds Schistosoma mansoni adult worm actin. **International Journal of Antimicrobial Agents**, v. 29, n. 5, p. 570–575,

2007.

TAYLOR, C. M.; WANG, Q.; ROSA, B. A.; HUANG, S. C.-C.; POWELL, K.; SCHEDL, T.; PEARCE, E. J.; ABUBUCKER, S.; MITREVA, M. Discovery of anthelmintic drug targets and drugs using chokepoints in nematode metabolic pathways. **PLoS pathogens**, v. 9, n. 8, p. e1003505, 2013.

TEBEJE, B. M.; HARVIE, M.; YOU, H.; LOUKAS, A.; MCMANUS, D. P. Schistosomiasis vaccines: where do we stand? **Parasites & Vectors**, v. 9, n. 1, p. 528, 2016.

TODESCHINI, R.; CONSONNI, V. **Handbook of Molecular Descriptors**. 1. ed. Weinheim, Germany: Wiley-VCH Verlag GmbH, 2000.

TODESCHINI, R.; CONSONNI, V. **Molecular Descriptors for Chemoinformatics**. 1. ed. Weinheim, Germany: Wiley-VCH Verlag GmbH, 2009.

TOLEDO, R.; FRIED, B. **Digenetic Trematodes**. 1. ed. New York: Springer, 2014.

TRAN, M. H.; PEARSON, M. S.; BETHONY, J. M.; SMYTH, D. J.; JONES, M. K.; DUKE, M.; DON, T. A.; MCMANUS, D. P.; CORREA-OLIVEIRA, R.; LOUKAS, A. Tetraspanins on the surface of Schistosoma mansoni are protective antigens against schistosomiasis. **Nature medicine**, v. 12, n. 7, p. 835–840, 2006.

TROPSHA, A. Best Practices for QSAR Model Development, Validation, and Exploitation. **Molecular Informatics**, v. 29, n. 6–7, p. 476–488, 2010.

TRUSCOTT, M.; EVANS, D. A.; GUNN, M.; HOFFMANN, K. F. Schistosoma mansoni hemozoin modulates alternative activation of macrophages via specific suppression of Retnla expression and secretion. **Infection and Immunity**, v. 81, n. 1, p. 133–142, 2013.

VALKO, M.; LEIBFRITZ, D.; MONCOL, J.; CRONIN, M. T. D.; MAZUR, M.; TELSER, J. Free radicals and antioxidants in normal physiological functions and human disease. **The International Journal of Biochemistry & Cell Biology**, v. 39, n. 1, p. 44–84, 2007.

VAN NOORDEN, R. Chemistry's web of data expands. **Nature**, v. 483, n. 7391, p. 524–524, 2012.

VAPNIK, V. **The Nature of Statistical Learning Theory**. 2nd. ed. New York: Springer, 2000.

VERCRUYSSE, J.; SHAW, D. J.; DE BONT, J. Index of potential contamination for schistosomiasis. **Trends in Parasitology**, v. 17, n. 6, p. 256–261, 2001.

WANG, Y.; COLEMAN-DERR, D.; CHEN, G.; GU, Y. Q. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. **Nucleic Acids Research**, v. 43, n. W1, p. W78–W84, 2015.

WANG, Y.; XIAO, J.; SUZEK, T. O.; ZHANG, J.; WANG, J.; ZHOU, Z.; HAN, L.; KARAPETYAN, K.; DRACHEVA, S.; SHOEMAKER, B. A.; BOLTON, E.; GINDULYTE, A.; BRYANT, S. H. PubChem's BioAssay Database. **Nucleic Acids Research**, v. 40, n. Database issue, p. D400–D412, 2012.

WARREN, K. S. The pathology, pathobiology and pathogenesis of schistosomiasis. **Nature**, v. 273, n. 5664, p. 609–612, 1978.

WELLING, M. **A First Encounter with Machine Learning**. 1. ed. Irvine, CA: University of California, 2011.

WERMUTH, C.; GANELLIN, C.; LINDBERG, P.; MITSCHER, L. Glossary of terms used in medicinal chemistry. **Pure and Applied Chemistry**, v. 70, n. 5, p. 1129–1143, 1998.

WHO. **Neglected tropical diseases**. Disponível em: <http://www.who.int/gho/neglected_diseases/en/>. Acesso em: 21 nov. 2016a.

WHO. **World Health Organization. Schistosomiasis**. Disponível em: <http://www.who.int/mediacentre/factsheets/fs115/en/>. Acesso em: 4 out. 2016b.

WIEST, P. M.; LI, Y.; OLDS, G. R.; BOWEN, W. D. Inhibition of Phosphoinositide Turnover by Praziquantel in Schistosoma mansoni. **The Journal of Parasitology**, v. 78, n. 4, p. 753–755, 1992.

WILLIAMS, D. L.; BONILLA, M.; GLADYSHEV, V. N.; SALINAS, G. Thioredoxin glutathione reductase-dependent redox networks in platyhelminth parasites. **Antioxidants & Redox Signaling**, v. 19, n. 7, p. 735–745, 2013.

WILSON, R. A. The saga of schistosome migration and attrition. **Parasitology**, v. 136,

n. 12, p. 1581–1592, 2009.

WILSON, R. A.; BARNES, P. E. The formation and turnover of the membranocalyx on the tegument of Schistosoma mansoni. **Parasitology**, v. 74, n. 1, p. 61–71, 1977.

WILSON, R. A.; LANGERMANS, J. A. M.; VAN DAM, G. J.; VERVENNE, R. A.; HALL, S. L.; BORGES, W. C.; DILLON, G. P.; THOMAS, A. W.; COULSON, P. S. Elimination of Schistosoma mansoni Adult Worms by Rhesus Macaques: Basis for a Therapeutic Vaccine? **PLoS Neglected Tropical Diseases**, v. 2, n. 9, p. e290, 2008.

WISHART, D. S.; KNOX, C.; GUO, A. C.; SHRIVASTAVA, S.; HASSANALI, M.; STOTHARD, P.; CHANG, Z.; WOOLSEY, J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. **Nucleic Acids Research**, v. 34, n. Database issue, p. D668-72, 2006.

WU, W.-M.; CHEN, Y.-L.; ZHAI, Z.; XIAO, S.-H.; WU, Y.-L. Study on the mechanism of action of artemether against schistosomes: the identification of cysteine adducts of both carbon-centred free radicals derived from artemether. **Bioorganic & Medicinal Chemistry Letters**, v. 13, n. 10, p. 1645–1647, 2003.

XIAO, S. Mefloquine, a new type of compound against schistosomes and other helminthes in experimental studies. **Parasitology Research**, v. 112, n. 11, p. 3723–3740, 2013.

XIAO, S. H.; CATTO, B. A. Comparative in vitro and in vivo activity of racemic praziquantel and its levorotated isomer on Schistosoma mansoni. **The Journal of Infectious Diseases**, v. 159, n. 3, p. 589–592, 1989.

XIAO, S.; MEI, J.; JIAO, P. The in vitro effect of mefloquine and praziquantel against juvenile and adult Schistosoma japonicum. **Parasitology Research**, v. 106, n. 1, p. 237–246, 2009.

XIE, X.-Q. S. Exploiting PubChem for virtual screening. **Expert Opinion on Drug Discovery**, v. 5, n. 12, p. 1205–1220, 2010.

XUE, L.; BAJORATH, J. Molecular descriptors in chemoinformatics, computational combinatorial chemistry, and virtual screening. **Combinatorial Chemistry &**

**High Throughput Screening**, v. 3, n. 5, p. 363–372, 2000.

YAP, C. W. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. **Journal of Computational Chemistry**, v. 32, n. 7, p. 1466–1474, 2011.

YOUNG, D. C. **Computational Drug Design: A Guide for Computational and Medicinal Chemists**. 1. ed. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2009.

YOUNG, N. D.; JEX, A. R.; LI, B.; LIU, S.; YANG, L.; XIONG, Z.; LI, Y.; CANTACESSI, C.; HALL, R. S.; XU, X.; CHEN, F.; WU, X.; ZERLOTINI, A.; OLIVEIRA, G.; HOFMANN, A.; ZHANG, G.; FANG, X.; KANG, Y.; CAMPBELL, B. E.; LOUKAS, A.; RANGANATHAN, S.; ROLLINSON, D.; RINALDI, G.; BRINDLEY, P. J.; YANG, H.; WANG, J.; WANG, J.; GASSER, R. B. Whole-genome sequence of Schistosoma haematobium. **Nature Genetics**, v. 44, n. 2, p. 221–225, 2012.

ZAKHAROV, A. V.; PEACH, M. L.; SITZMANN, M.; NICKLAUS, M. C. QSAR modeling of imbalanced high-throughput screening data in PubChem. **Journal of Chemical Information and Modeling**, v. 54, n. 3, p. 705–712, 2014.

ZANELLA, F.; LORENS, J. B.; LINK, W. High content screening: seeing is believing. **Trends in Biotechnology**, v. 28, n. 5, p. 237–245, 2010.

ZHOU, Y.; ZHENG, H.; CHEN, Y.; ZHANG, L.; WANG, K.; GUO, J.; HUANG, Z.; ZHANG, B.; HUANG, W.; JIN, K.; DOU, T.; HASEGAWA, M.; WANG, L.; ZHANG, Y.; ZHOU, J.; TAO, L.; CAO, Z.; LI, Y.; VINAR, T.; BREJOVA, B.; BROWN, D.; LI, M.; MILLER, D. J.; BLAIR, D.; ZHONG, Y.; CHEN, Z.; LIU, F.; HU, W.; WANG, Z.-Q.; ZHANG, Q.-H.; SONG, H.-D.; CHEN, S.; XU, X.; XU, B.; JU, C.; HUANG, Y.; BRINDLEY, P. J.; MCMANUS, D. P.; FENG, Z.; HAN, Z.-G.; LU, G.; REN, S.; WANG, Y.; GU, W.; KANG, H.; CHEN, J.; CHEN, X.; CHEN, S.; WANG, L.; YAN, J.; WANG, B.; LV, X.; JIN, L.; WANG, B.; PU, S.; ZHANG, X.; ZHANG, W.; HU, Q.; ZHU, G.; WANG, J.; YU, J.; WANG, J.; YANG, H.; NING, Z.; BERIMAN, M.; WEI, C.-L.; RUAN, Y.; ZHAO, G.; WANG, S.; LIU, F.; ZHOU, Y.; WANG, Z.-Q.; LU, G.; ZHENG, H.; BRINDLEY, P. J.; MCMANUS, D. P.; BLAIR, D.; ZHANG, Q.; ZHONG, Y.; WANG, S.; HAN, Z.-G.; CHEN, Z.; WANG, S.;

HAN, Z.-G.; CHEN, Z. The Schistosoma japonicum genome reveals features of host–parasite interplay. **Nature**, v. 460, n. 7253, p. 345–351, 2009.

ZHU, F.; HAN, B.; KUMAR, P.; LIU, X.; MA, X.; WEI, X.; HUANG, L.; GUO, Y.; HAN, L.; ZHENG, C.; CHEN, Y. Update of TTD: Therapeutic Target Database. **Nucleic Acids Research**, v. 38, n. Database issue, p. D787–D791, 2010.

ZHU, H.; TROPSHA, A.; FOURCHES, D.; VARNEK, A.; PAPA, E.; GRAMATICA, P.; ÖBERG, T.; DAO, P.; CHERKASOV, A.; TETKO, I. V. Combinatorial QSAR Modeling of Chemical Toxicants Tested against Tetrahymena pyriformis. **Journal of Chemical Information and Modeling**, v. 48, n. 4, p. 766–784, 2008.

ZINIEL, P. D.; KARUMUDI, B.; BARNARD, A. H.; FISHER, E. M. S.; THATCHER, G. R. J.; PODUST, L. M.; WILLIAMS, D. L. The Schistosoma mansoni Cytochrome P450 (CYP3050A1) Is Essential for Worm Survival and Egg Development. **PLOS Neglected Tropical Diseases**, v. 9, n. 12, p. e0004279, 2015.

ZUSSMAN, R. A.; BAUMAN, P. M.; PETRUSKA, J. C. The role of ingested hemoglobin in the nutrition of Schistosoma mansoni. **The Journal of Parasitology**, v. 56, n. 1, p. 75–79, 1970.