



Ministério da Educação
Universidade Federal de Goiás
Faculdade de Farmácia
Programa de Pós-Graduação em Ciências Farmacêuticas



KARLLA GONÇALVES DE MACEDO

**Estudos de SAR e QSAR para um conjunto de triazolopirimidinas inibidores da
enzima diidroorotato desidrogenase de *Plasmodium falciparum***

Goiânia

2014

KARLLA GONÇALVES DE MACEDO

**Estudos de SAR e QSAR para um conjunto de triazolopirimidinas inibidores da
enzima diidroorotato desidrogenase de *Plasmodium falciparum***

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas da Faculdade de Farmácia da Universidade Federal de Goiás, como requisito parcial para obtenção do título de Mestre em Ciências Farmacêuticas

Área de Concentração: Fármacos e Medicamentos

Orientadora: Profa. Dra. Carolina Horta Andrade

Goiânia

2014

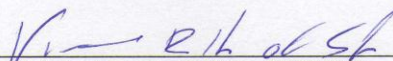
Folha de Aprovação

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas da Universidade Federal de Goiás, em 05 de agosto de 2014, pela mestranda Karlla Gonçalves de Macedo.

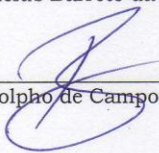
Banca Examinadora:



Profa. Dra. Carolina Horta Andrade (FF/UFG)
Presidente



Prof. Dr. Vinícius Barreto da Silva (PUC/GO)



Prof. Msc. Rodolpho de Campos Braga (IQ/UFG)

DEDICATÓRIA

Dedico este trabalho aos meus pais, Carlos (in memoriam) e Marcelina, pelo amor e dedicação. E aos meus anjos na Terra Tia Sueli e Tio Walter.

AGRADECIMENTOS

A Deus, por sempre nos permitir caminhar e alcançar nossos sonhos.

Ao meus pais, que nunca mediram esforços e incentivos para que pudesse trilhar minha carreira acadêmica, em especial minha mãe por todo amor e compreensão.

A minha família, que mesmo sem compreender o universo científico nunca deixou de apoiar.

Ao meu querido Marcelo Henrique, pelo amor, respeito, alegria e principalmente solidariedade nos momentos difíceis.

Aos amigos que fiz durante minha trajetória acadêmica, em especial a Ágatha, pelos tempos de Cepae; Ramon e Camila, pelo carinho e companheirismo durante a graduação; Renata, Larissa, Diego, Elias e Flávia, pela amizade, momentos de discussão científica, descontração e amizade construída, na pós graduação.

Aos professores que tive durante minha vida escolar, principalmente aqueles que marcaram minhas escolhas. Professora Marilene Barcelos, que foi minha professora de química no Ensino Médio e também me orientou no estágio, durante a graduação. Professoras Anna Benite e Olga Barros que me mostraram e me permitiram conhecer o universo da pesquisa, e claro a professora Carolina Horta, que me orientou nesse trabalho e me deu oportunidade de aprendizado. A você Carolina, o meu agradecimento, pela amizade, orientação e principalmente compreensão.

Aos colegas de Laboratório, pelo convívio, amizade, solidariedade, momentos de discussão científica.

Ao Rodolpho Braga, pela ajuda, por não medir esforços para a realização deste trabalho.

A todos os professores e funcionários da Faculdade de Farmácia/UFG.

As agências de fomento CAPES e CNPq. E FAPEG, pela concessão da bolsa.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

EPÍGRAFE

“e conhecereis a verdade, e a verdade vos libertará.”

João 8:32

SUMÁRIO

LISTA DE FIGURAS	IX
LISTA DE TABELAS	XI
LISTA DE ABREVIATURAS E SIGLAS	XIII
RESUMO	XV
ABSTRACT	XVI
1 INTRODUÇÃO	17
1.1 Malária	17
1.1.1 Epidemiologia	18
1.1.2 Ciclo biológico do parasito	19
1.1.3 Quimioterapia	22
1.1.3.1 4-aminoquinolinas	22
1.1.3.2 8-aminoquinolinas	23
1.1.3.3 Álcoois quinolínicos	24
1.1.3.4 Atovaquona	25
1.1.3.5 Antifolatos	25
1.1.3.6 Endoperóxidos	26
1.1.3.7 Novos fármacos antimaláricos em fase clínica de desenvolvimento	28
1.2 Planejamento Racional de Novos Antimaláricos	29
1.2.1 Bases Moleculares para o Planejamento Racional de Novos Antimaláricos	29
1.2.1.1 Diidroorotatodesidrogenase de <i>Plasmodium falciparum</i> (PfDHODH)	29
1.2.1.2 Triazolopirimidinas, inibidores PfDHODH	32
1.3 Pesquisa, Desenvolvimento & Inovação de novos fármacos	34
1.3.1 Química Medicinal e a Gênese Planejada de Fármacos	37
1.3.2 Relações Quantitativas entre Estrutura Química e Atividade biológica (QSAR)	38
1.3.2.1 Holograma QSAR (HQSAR)	41
2 OBJETIVOS	46
3 MATERIAIS E MÉTODOS	47
3.1 Material	47
3.1.1 Computadores e <i>software</i>	47
3.1.2 Conjunto de dados	48
3.2 Métodos	53
3.2.1 Padronização do conjunto de dados e preparo das estruturas	53

3.2.2	Estudo da relação entre estrutura e atividade (SAR)	53
3.2.2.1	Análise de agrupamentos hierárquicos	53
3.2.2.2	Análise de <i>cliffs</i> de atividade	54
3.2.3	Divisão dos conjunto treinamento e teste	54
3.2.4	HQSAR	54
3.2.5	QSAR – 2D	55
3.2.6	Validação externa dos modelos de HQSAR e QSAR-2D	56
3.2.6.1	<i>Coefficiente de correlação preditivo (r_{pred}^2)</i>	56
3.2.6.2	<i>Coefficiente de correlação modificado (r_m^2)</i>	57
3.2.6.3	<i>Coefficiente de correlação de concordância (CCC)</i>	58
3.2.6.4	<i>Método de Golbraikh e Tropsha (G&T)</i>	59
3.2.6.5	<i>Teste de Durbin Watson de correlação entre os resíduos (d)</i>	59
4	RESULTADOS E DISCUSSÃO	60
4.1	Caracterização do conjunto de dados	60
4.2	Estudo da Relação entre Estrutura e Atividade (SAR)	60
4.2.1	Análise de agrupamentos hierárquicos	60
4.2.2	Análise de <i>cliffs</i> de atividade	62
4.3	Divisão do conjunto de dados para construção dos modelos de QSAR	69
4.4	HQSAR	70
4.5	QSAR – 2D	79
5	CONCLUSÕES	82
6	REFERÊNCIAS BIBLIOGRÁFICAS	83

LISTA DE FIGURAS

Figura 1	Estrutura química do primeiro fármaco antimalárico utilizado, a quinina	17
Figura 2	Número de casos confirmados de malária no mundo em 2010 (WHO, 2012)	18
Figura 3	Ciclo de vida do <i>Plasmodium</i> no homem e no mosquito	20
Figura 4	Estruturas químicas da cloroquina, piperaquina e amodiaquina	23
Figura 5	Estrutura química da primaquina	24
Figura 6	Estruturas químicas da quinina e mefloquina	25
Figura 7	Estrutura química da atovaquona	25
Figura 8	Estruturas químicas da dapsona, sulfadoxina, proguanil e pirimetamina	26
Figura 9	Estruturas químicas da artemisinina, diidroartemisinina, artesunato e arteméter	27
Figura 10	Biossíntese de pirimidinas em Plasmodium pela via <i>de novo</i>	30
Figura 11	Representação em fita da estrutura 3D da PfDHODHco-cristalizada com DSM1. PDB ID 3I65	31
Figura 12	Estrutura do inibidor de HsDHODHleflunomida (A) e seu metabólito ativo (B)	32
Figura 13	Estruturas 3D obtidas a partir da cristalografia de raios-X da PfDHODH em complexo com os inibidores 22 (DSM 1), PDB ID 3I65, (azul); 26 (DMS 2), PDB ID 3I68, (rosa); 27 (DMS 74), PDB ID 3I6R, (alaranjado) (A) e 82 (DMS 267), PDB ID 3SFK (B)	33
Figura 14	Etapas do processo de descoberta e desenvolvimento de novos fármacos (LOMBARDINO; LOWE, 2004)	34
Figura 15	Processo de desenvolvimento do modelo de QSAR (adaptado(TROPSHA, 2010)	38
Figura 16	Esquema do processo de geração dos hologramas moleculares e dos modelos de HQSAR.	43
Figura 17	<i>Workflow</i> desenvolvido na plataforma Knime para geração dos modelos de QSAR – 2D	56
Figura 18	Gráficos de regressão com e sem interceptação em zero: dados preditos estão plotados do eixo X e experimentais no eixo Y (A); com inversão dos eixos (B)(ROY <i>et al.</i> , 2013).	58
Figura 19	Análise agrupamentos hierárquicos para o conjunto de dados: dendograma (A) e mapa de calor da matriz de distâncias organizadas com base na similaridade estrutural (B) (violeta/azul = similar; amarelo, vermelho = dissimilar).	61
Figura 20	Representação estrutural de alguns compostos do <i>cluster</i> a (A) e compostos do <i>cluster</i> b (B), ao meio o <i>scaffold</i> triazolopirimidínico, presente nos compostos.	62

Figura 21	Representação 3D da paisagem descontínua de atividade (<i>activity landscape</i>), calculada baseada na distância euclidiana	64
Figura 22	Interações mais importantes feitas com a <i>PfDHODH</i> (VYAS; GHATE, 2011)	67
Figura 23	(A) Sítio de interação da enzima <i>PfDHODH</i> co-cristalizada com o ligante (DMS 267) (PDB ID: 3SFK). (B) <i>Ligplot</i> das interações do inibidor com os aminoácidos da enzima.	68
Figura 24	Requisitos estruturais necessários para a atividade antimalárica de triazolopirimidinas, de acordo com o estudo de SAR.	69
Figura 25	Representação gráfica da distribuição dos valores de pIC_{50} nos conjuntos teste, treinamento e conjunto total.	70
Figura 26	Representação gráfica dos valores experimentais e preditos de pIC_{50} obtidos com os modelos (A) A/C/H, (B) A/H/DA, (C) A/B/C/H, (D) A/C/H/Ch (E) A/C/H/DA e (F) A/B/C/H/Ch de HQSAR para os compostos do conjunto de treinamento (preto) e teste (branco).	75
Figura 27	Representação gráfica dos mapas de contribuição de HQSAR para os compostos mais potentes do conjunto de dados (86, 87 e 98), utilizando o modelo gerado com os fragmentos A/C/H/DA. Em azul ciano está representado a máxima subestrutura comum dos compostos.	77
Figura 28	Representação gráfica dos mapas de contribuição de HQSAR para os compostos, menos potentes do conjunto, 98, 87, 110 e 128 utilizando o modelo gerado com os fragmentos A/C/H/DA. Em azul ciano está representado a máxima subestrutura comum dos compostos.	78
Figura 29	Representação gráfica dos valores experimentais e preditos de pIC_{50} obtidos com o modelo KRLS de QSAR- 2D para os compostos do conjunto de treinamento (rosa) e teste (azul).	80
Figura 30	Domínio de aplicabilidade dos compostos do conjunto teste	81

LISTA DE TABELAS

Tabela 1	Candidatos a fármacos antimaláricos que estão em estudos clínicos (Fonte: MMV,2014)	28
Tabela 2	Triazolopirimidinas e com os respectivos valores de IC ₅₀ em PfDHODH	33
Tabela 3	Parâmetros de distinção dos fragmentos no método de HQSAR	44
Tabela 4	Estruturas químicas e respectivos valores de IC ₅₀ e pIC ₅₀ para 132 triazolopirimidinas, inibidores da diidroorotato desidrogenase de <i>P.falciparum</i> (PfDHODH)	47
Tabela 5	Estruturas que representam <i>cliffs</i> de atividade	64
Tabela 6	Resultados das análises de HQSAR para várias combinações de distinção de fragmentos utilizando o tamanho padrão de fragmento (4-7), com conjunto de treinamento constituído por 102 compostos	71
Tabela 7	Modelos de HQSAR após a retirada dos <i>outliers</i>	72
Tabela 8	Influência dos diferentes tamanhos dos fragmentos nos parâmetros estatísticos nos melhores modelos de HQSAR	72
Tabela 9	Valores experimentais e preditos de pIC ₅₀ dos compostos do conjunto teste calculados para os seis melhores modelos.	73
Tabela 10	Desvio padrão dos valores de preditividade (σ), para os seis melhores modelos e capacidade preditiva.	74
Tabela 11	Resultados estatísticos da validação externa para os seis melhores modelos de HQSAR	76
Tabela 12	Características estatísticas dos modelos de QSAR-2D de para os inibidores de PfDHODH utilizando os diferentes métodos de regressão	79

LISTA DE ABREVIATURAS E SIGLAS

1D	Unidimensional
2D	Bidimensional ou segunda dimensão
3D	Tridimensional ou terceira dimensão
A	Átomos
a.C	Antes de Cristo
ACT	Carbamoiltransferase Aspartato
Asp	Ácido L-aspártico
B	Ligação
C	Conectividade
CA	Carbamoil Aspartato
CADD	Planejamento de Fármacos Auxiliado por Computador (do inglês, <i>Computer-Aided Drug Design</i>)
CCC	Coefficiente de Correlação de Concordância
Ch	Quiralidade
CP	Carbamoilfosfato
CPS	Carbamoil-fosfato Sintetase
CRC	Algoritmo Cíclico de Verificação de Redundância (do inglês, <i>cyclic redundancy check</i>)
d	<i>Durbin Watson</i>
DA	Doadores e Aceptores
DDT	Dicloro-difenil-tricloroetano
DHO	Diidroorotato
DHOtase	Diidroorotase
DHODH	Diidroorotato Desidrogenase
DNA	Ácido Desoxirribonucleico
FMN	Flavina Mononucleotídeo
GAT	Glutamina Amidotransferase Bifuncional
GBM	<i>Gradient Boosting Machine</i>
G&T	Método <i>Golbraikh e Tropsha</i>
H	Hidrogênio
<i>Hit</i>	Composto com potência inibitória inferior a 2 μ M
HL	Comprimento de holograma
HQSAR	Holograma QSAR
<i>Hs</i> DHODH	Diidroorotato Desidrogenase Humana

HTS	Triagem automatizada em larga escala (do inglês, <i>High-Throughput Screening</i>)
IC	Intervalo de Confiança
IC ₅₀	Concentração Inibitória de 50%
IUPAC	União Internacional de Química Pura e Aplicada (do inglês, <i>International Union of Pure and Applied Chemistry</i>)
KRLS	<i>Kernel-based Regularized Least Squares</i>
LabMol	Laboratório de Planejamento de Fármacos e Modelagem Molecular
LBDD	Planejamento de fármacos baseado no ligante (do inglês, <i>Ligand-Based Drug Design</i>)
LOO	Leave one out (Deixe um de fora)
MIF	Campos de interação molecular (do inglês, <i>Molecular Interaction Fields</i>)
<i>N</i>	Numero ótimo de componentes PLS
NCE	Nova Entidade Química (do inglês <i>New Chemical Entities</i>)
AO	Ácido Orótico
OMPDC	Orotidina-5'-monofosfato Descarboxilase
OMP	Orotidilato
OMS	Organização Mundial de Saúde
OPRT	Fosforibosiltransferase Orotato
ORO	Orotato
PD&I	Pesquisa, Desenvolvimento & Inovação
PDB	Banco de Dados de Proteínas (do inglês, <i>Protein Data Bank</i>)
<i>Pf</i> DHODH	Diidroorotato Desidrogenase de <i>Plasmodium falciparum</i>
PLS	Regressão por mínimos quadrados parciais (do inglês, <i>Partial Least Squares</i>)
q^2	Coefficiente de correlação da validação cruzada
QFM	Química Farmacêutica Medicinal
QSAR	Relações Quantitativas entre estrutura e atividade
QSAR – 2D	Relações Quantitativas entre estrutura e atividade bidimensional
r^2	Coefficiente de Correlação de Determinação Múltipla
r^2_{pred}	Coefficiente de Correlação Preditivo
RF	<i>Random Forest</i>
r_m^2	Coefficiente de Correlação Modificado
RMN	Ressonância Magnética Nuclear
RNA	Ácido Ribonucleico

SAR	Relações entre estrutura e atividade (do inglês, <i>Structure-Activity Relationships</i>)
SBDD	Planejamento de fármacos baseado na estrutura (do inglês, <i>Structure-Based Drug Design</i>)
SEE	Erro padrão estimado (do inglês, <i>Standard Error Estimate</i>)
SUS	Sistema Único de Saúde
SVM	<i>Support Vector Machines</i>
TZP	triazolopirimidina
UMP	Uridilato

RESUMO

O processo de planejamento e desenvolvimento de novos fármacos é um trabalho complexo, que demanda elevados investimentos de tempo e dinheiro. Estratégias de planejamento de fármacos auxiliadas por computador, CADD (*Computer-Aided Drug Design*) vêm se destacando, pois minimizam gastos e tempo, além de poder explorar um número maior de alvos biológicos e moléculas promissoras. A malária é uma doença endêmica grave na África e América do Sul, causada por protozoários do gênero *Plasmodium*. Em 2012 foram estimados 207 milhões de casos e 627.000 mortes, de acordo com a Organização Mundial da Saúde. A enzima diidroorotato desidrogenase (DHODH) atua na quarta etapa da biossíntese de pirimidinas, é um alvo validado para o planejamento de novos agentes antimaláricos. O objetivo geral deste trabalho foi desenvolver regras de relação entre estrutura e atividade (SAR) e modelos robustos e preditivos de relações quantitativas entre estrutura e atividade bidimensionais (QSAR-2D), utilizando um conjunto de triazolopirimidinas descritas na literatura como inibidores da DHODH de *P. falciparum* (PfDHODH). Foram desenvolvidas regras de SAR utilizando os métodos de análise de agrupamentos, *cliffs* de atividade e *landscapes* de atividade. Além disso, desenvolveu-se e validou-se vários modelos de QSAR-2D e de holograma QSAR (HQSAR). As análises de SAR, permitiram estabelecer requisitos estruturais essenciais para a atividade antimalárica das triazolopirimidinas, como substituintes haletos de alquila no anel triazolopimidínico, substituintes hidrofóbicos na posição *para* no anel benzênico, todos de acordo com o espaço químico da cavidade de interação da PfDHODH. Os modelos de HQSAR e QSAR-2D apresentaram bons parâmetros estatísticos e boa capacidade preditiva. Os mapas de contribuição de HQSAR também estão de acordo com o espaço químico da cavidade de interação da PfDHODH. Os dados obtidos servem como guia para o planejamento de novos antimaláricos com maior potência.

Palavras-chave: Malária, *Plasmodium falciparum*, planejamento de fármacos, diidroorotato desidrogenase, SAR, QSAR.

ABSTRACT

Drug discovery and development process requires high investments of both time and money. Strategies for drug design aided by computers, CADD (Computer-Aided Drug Design) have gained prominence over the last decades, in order to minimize the impact of those costs. CADD techniques also allow the exploration of a greater number of biological targets and promising molecules. Malaria is an endemic disease in Africa and in South American caused by the protozoa of the genus *Plasmodium*. In 2012, 207 million cases and 627,000 deaths were estimated, according to the World Health Organization. The enzyme dihydroorotate dehydrogenase (DHODH) catalyzes the fourth step of the pyrimidine biosynthesis, and consists in a validated target for the design of new antimalarial agents. The aim of this study was to develop structure-activity relationships (SAR) rules and to generate quantitative structure-activity relationships (QSAR) models using a set of triazolopyrimidines described in the literature as inhibitors of DHODH from *P. falciparum* (PfDHODH). SAR rules were established using methods of clustering, activity cliffs and activity landscapes. In addition, several models of 2D-QSAR and hologram QSAR (HQSAR) were developed and validated. The SAR analyses allowed the understanding of the basic structural requirements for the antimalarial activity of triazolopyrimidines, like alkyl halides substituents on the triazolopyrimidinic ring, hydrophobic substituents in the *para* position on the benzene ring, all in agreement with the chemical space inside the active site of the PfDHODH. The HQSAR and 2D-QSAR models showed good statistical parameters and good predictive ability. The HQSAR contour maps were also consistent with the chemical space of the active site of the enzyme. The results of this study could serve as guide for the design of new antimalarials with higher potency.

Keywords: Malaria, *Plasmodium falciparum*, drug design, dihydroorotate dehydrogenase, SAR, QSAR.

1. INTRODUÇÃO

1.1. Malária

A malária é uma doença endêmica grave que acomete o homem desde os primórdios da humanidade, representando uma ameaça de saúde pública principalmente para crianças e gestantes (BURROWS, 2012). Há relatos que a malária tenha sido a principal causa de morte entre os primatas precursores do *Homo sapiens*. Várias referências da literatura antiga já descreviam sintomas idênticos aos de malária. Em 2.700 a.C., o Cânon Chinês da medicina, Nei Ching, descreveu o quadro sintomático da malária e as relações entre febres e melancolias prolongadas. Da mesma forma, manuscritos do século VI a.C., mencionam febres mortais semelhantes à malária que afligiam a população da antiga Mesopotâmia (FRANÇA; SANTOS; FIGUEROA-VILLAR, 2008).

A primeira intervenção terapêutica utilizada surgiu apenas em 1639, quando padres jesuítas aproveitaram-se do conhecimento dos povos Maia de utilizar cascas da árvore *Cinchona officinalis* no tratamento de febres. Todavia, foi apenas em 1820 que os químicos franceses Pelletier e Caventou identificaram o alcalóide quinina (Figura 1) como o ingrediente ativo da casca de *Cinchona sp.* e a tornaram o principal quimioterápico disponível para o tratamento naquela época (BUTLER, KHAN, FERGUSON, 2010).

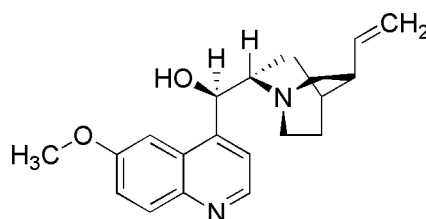


Figura 1 – Estrutura química do primeiro fármaco antimalárico utilizado, quinina.

Mesmo com vários relatos da história, a luta contra a malária só teve início após as descobertas de Laveran em 1880. Ele descobriu que a malária era causada por um parasito e Ross em 1898, o qual provou que a transmissão da doença ocorria através da picada de um mosquito infectado (FRANÇA; SANTOS; FIGUEROA-VILLAR, 2008). Estas descobertas abriram novos caminhos para medidas de controle e tratamento da doença, incluindo a instalação de telas em janelas, redução de criadouros do mosquito através de mudanças nos hábitos agrícolas, aplicação de inseticidas como o dicloro-difenil-tricloroetano (DDT) e descoberta de novos antimaláricos (HAY *et al.*, 2004). Em

1955, a Organização Mundial de Saúde (OMS) criou o programa mundial de erradicação da malária através da implementação de inseticidas e do fármaco cloroquina nas medidas iniciais de controle da doença. Quando o programa foi oficialmente encerrado em 1969, mais de 37 países foram declarados livres da doença, entretanto refletiu em um quadro de resistência generalizada a praticamente todos os inseticidas e à cloroquina (NÁJERA, SILVA, ALONSO, 2011).

1.1.1. Epidemiologia

De acordo com a OMS, estima-se que em 2012 ocorreram 207 milhões de casos de malária e 627.000 mortes em todo mundo. Os casos estão distribuídos nos países da África, Ásia, Oceania e América Latina. Os países Angola, República Democrática do Congo, Sudão, Etiópia, Moçambique, localizados na África, e a Índia, no continente Asiático, apresentaram os maiores números de casos confirmados de malária, superiores a 1 milhão em 2010 (WHO,2012) (figura 2).

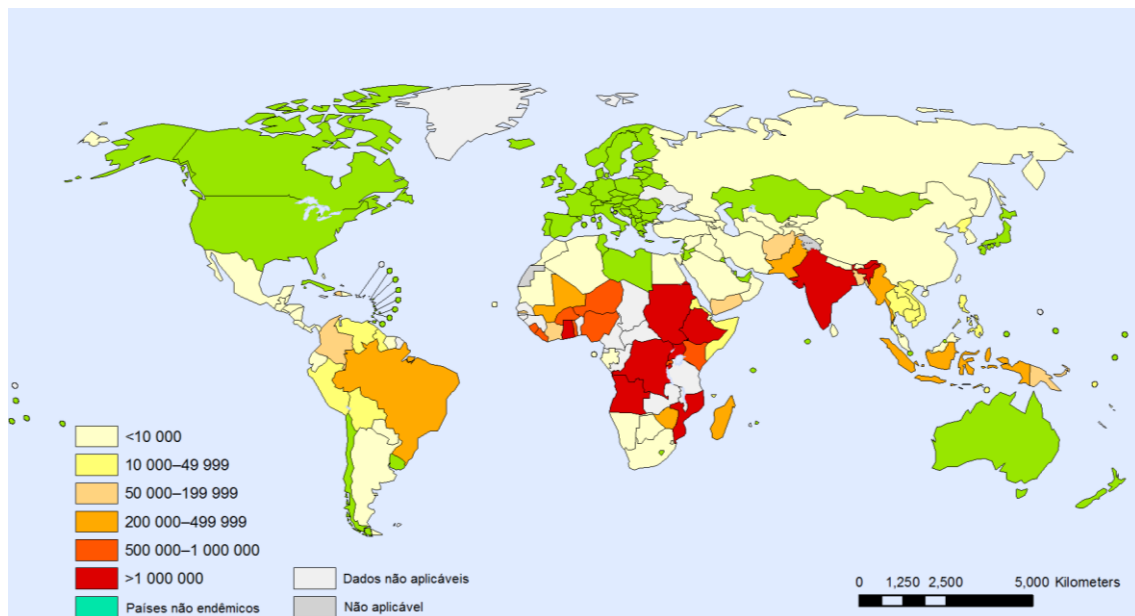


Figura 2 – Número de casos confirmados de malária no mundo em 2010 (WHO, 2012).

A malária é causada por protozoários pertencentes ao filo Apicomplexa, do gênero *Plasmodium spp.*, sendo transmitida através da picada de fêmeas do mosquito *Anopheles ssp.* (PHILLIPS; RATHOD, 2010). Há outras situações de transmissão menos frequentes, como transfusão de sangue, compartilhamento de agulhas contaminadas e transmissão vertical, antes ou durante o parto (WHO, 2014).

Há cerca de 100 espécies de *Plasmodium*, mas somente cinco são capazes de causar a doença em humanos: *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale* (África) e *P. knowlesi* (sudeste da Ásia) (BRAY, GARNHAM, 1982; PRICE *et al.*, 2007) . Entre eles, o *P. falciparum* é o parasito mais prevalente e responsável por mais de 90% das mortes, seguido pelo *P. vivax*, o segundo mais comum e o responsável pela forma recorrente da doença (SINGH *et al.*, 2008)

No Brasil, 99% dos casos de malária estão restritos à Amazônia Legal, totalizando cerca de 310 mil novos casos por ano, quase todos causados pelo *P. vivax* (75% dos casos). Apesar da baixa incidência, quando comparada aos 194 milhões de casos no continente africano, onde uma criança morre por minuto, o quadro endêmico nesta região representa gastos substanciais para o Sistema Único de Saúde (SUS), chegando a aproximadamente US\$ 120 milhões apenas em 2010. Apesar da redução de 26% das taxas de mortalidade entre os anos 2000 e 2010, a malária ainda continua sendo um problema grave de saúde pública, estando presente em 109 países de clima tropical e subtropical, nos quais as condições ambientais favorecem a manutenção e desenvolvimento dos vetores da doença (WHO, 2014).

1.1.2. Ciclo biológico do parasito

O *Plasmodium* possui um ciclo de vida complexo, dividido em várias etapas: ciclo hepático, ciclo eritrocitário e ciclo esporocogônico. A Figura 3 apresenta as etapas envolvidas no ciclo de vida do *Plasmodium spp.*

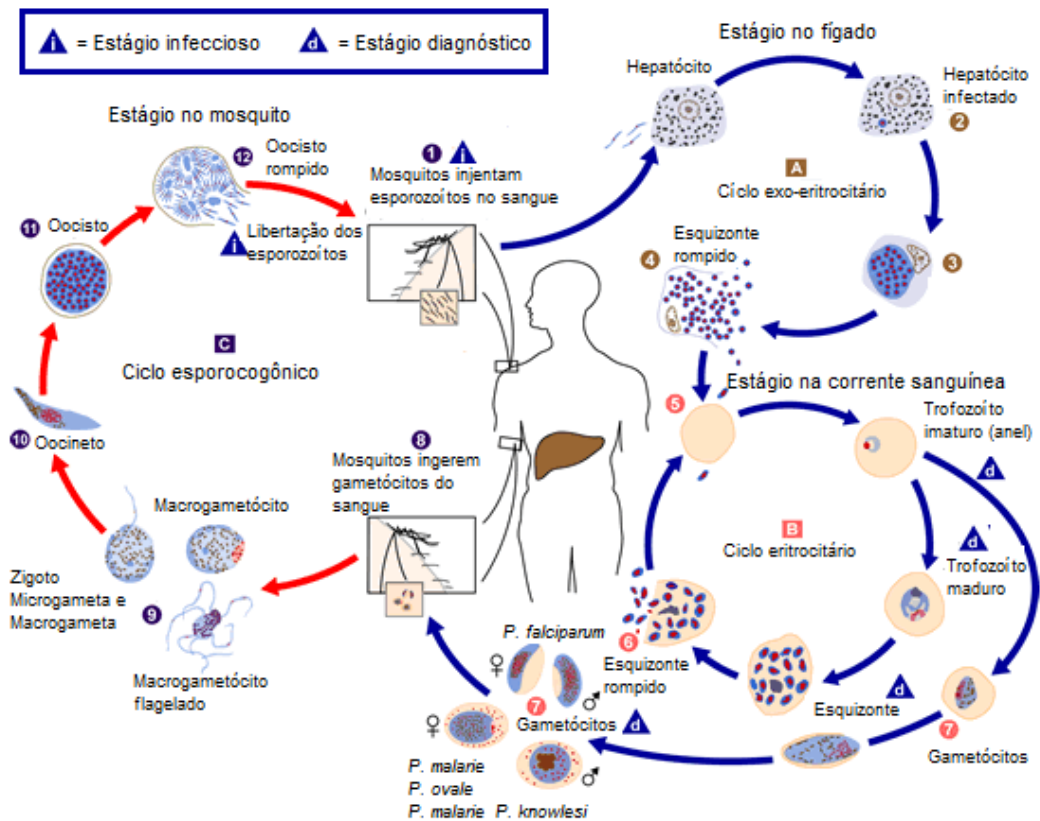


Figura 3 – Ciclo de vida do *Plasmodium* no homem e no mosquito. As setas azuis indicam as formas assexuadas hepáticas e eritrocíticas do parasito e as setas vermelhas indicam as formas sexuais. (1) o mosquito fêmea do gênero *Anopheles* liberam esporozoítos no sangue do hospedeiro, (2) que posteriormente invadem os hepatócitos, multiplicam-se assexuadamente e se transformam em esquizontes (3). Em seguida dezenas de esquizontes são liberados na circulação (4), invadem os eritrócitos, multiplicam-se assexuadamente (5) liberando trofozoítos que multiplicam-se sucessivamente em ciclos intermitentes (6). Trofozoítos se diferenciam em formas sexuais denominadas gametócitos (7) e são ingeridos por outro mosquito (8). No inseto, os gametócitos se diferenciam em macrogametócitos e microgametócitos, se fundem e dão origem a um zigoto (9). O zigoto se diferencia em oocineto (10), que por sua vez migra até o intestino e dá origem a um oocisto (11). Ao final oocistos migram até as glândulas salivares, tornando possível a infecção em outro hospedeiro humano (12) (Adaptado de: CDC, 2012).

O processo infeccioso inicia após a liberação de merozoítos das glândulas salivares do mosquito para a corrente sanguínea (1) do hospedeiro (YAMAUCHI *et al.*, 2007). Em seguida os parasitos percorrem a circulação sanguínea até chegar aos hepatócitos (2), local onde se multiplicam e diferenciam em esquizontes assexuadamente (ciclo hepático) por um período de 7-12 dias (3) (STURM *et al.*, 2006). Ao final deste período, dezenas de merozoítos são liberados na circulação sanguínea para infectar eritrócitos (4), dando início ao ciclo eritrocitário. (MOTA *et al.*, 2001; PRUDENCIO, RODRIGUEZ, MOTA, 2006). Cada eritrócito infectado é alvo de

repetidos ciclos de multiplicação assexuada (5), a medida que os trofozoítos formados rompem estas células liberando citocinas e uma enorme carga parasitária para invadir outros eritrócitos (6) (YAMAUCHI *et al.*, 2007; MILLER *et al.*, 2002; COWMAN, CRABB, 2002).

Os trofozoítos também se diferenciam em formas sexuadas chamadas de gametócitos (7). Estas formas constituem parte de um mecanismo essencial para disseminação da doença, ou seja, outro mosquito não contaminado pode picar o hospedeiro infectado (8) e ingerir gametócitos iniciando o ciclo esporogônico da doença. No mosquito, os gametócitos se diferenciam em macrogametócitos e microgametócitos e se fundem dando origem um zigoto (9). Em seguida o zigoto se diferencia na forma oocineto (10) e migra até a parede do intestino médio do mosquito dando origem a um oocisto (11). O tempo necessário para que o parasito complete o seu ciclo no mosquito varia de espécie para espécie. Os esporozoítos formados nos oocistos são liberados na hemolinfa do inseto e migram até as glândulas salivares, tornando possível a infecção em outro hospedeiro humano (12) (TUTEJA, 2007).

O período de incubação da doença varia de 7 a 14 dias, podendo, contudo, chegar a vários meses devido às formas latentes de *P. vivax* e *P. malariae*. A crise aguda da malária caracteriza-se por episódios de febre, calafrios, tosse, dificuldade respiratória, dor nas articulações, dor de cabeça, diarreia, vômitos e convulsões. Os intervalos entre as febres acompanham os ciclos repetidos de obstrução dos eritrócitos, repetindo-se a cada 48 horas para infecções por *P. vivax* e *P. falciparum* e de 72 horas para infecções causadas pelos demais parasitos (TUTEJA, 2007).

O quadro clínico da malária pode ser leve, moderado ou grave, dependendo da espécie do parasito, da quantidade de parasitos circulantes, do tempo de doença e do nível de imunidade adquirida pelo paciente. Os quadros graves estão sempre associados ao *P. falciparum*, uma vez que esse parasito acarreta complicações clínicas como icterícia, insuficiência renal, anemia grave e obstrução dos capilares que levam sangue ao cérebro (MILLER *et al.*, 2002; TUTEJA, 2007).

Na maioria dos casos a malária não representa um risco à vida, entretanto, esta doença é extremamente perigosa para mulheres grávidas e crianças. A invasão do parasito na placenta representa um problema durante a gravidez, estando sempre associada ao parto prematuro, ao baixo peso do bebê e a morte do recém-nascido (SNOW, MARSH, 1998; GEERTRUYDEN *et al.*, 2004).

1.1.3. Quimioterapia

Os fármacos antimaláricos são baseados em compostos naturais ou sintéticos identificados há mais de 70 anos. Os antimaláricos eritrocíticos atuam nas formas presentes nos eritrócitos do homem, enquanto os antimaláricos gametocíticos atuam nas formas sexuadas do parasita do indivíduo infectado, de forma que quando este é picado por outro mosquito se evita a disseminação da doença. Por último, existem os fármacos esporontocidas, que são aqueles que atuam nas formas hepáticas da doença (FRANÇA, SANTOS, VILLAR, 2008; GRIMBERG, MEHLOTRA, 2011).

A conduta terapêutica na utilização dos fármacos antimaláricos leva em consideração a tolerabilidade, segurança (especialmente em gestantes e crianças pequenas) e tratamento em regimes de curta duração. Apesar da grande quantidade de antimaláricos disponíveis, atualmente se preconiza o uso de associações entre antimaláricos com mecanismos de ação distintos, tornando o tratamento mais efetivo e evitando a expansão dos casos de resistência (PETERSEN, EASTMAN, LANZER, 2011). As principais opções terapêuticas utilizadas atualmente estão descritas abaixo.

1.1.3.1. 4-aminoquinolinas

A cloroquina (Figura 4) foi a primeira 4-aminoquinolina utilizada no tratamento da malária. Introduzida na terapêutica no final dos anos 40 para a prevenção de malária em larga escala, este fármaco já foi considerado o tratamento padrão devido a sua eficácia, acessibilidade e segurança, mesmo durante a gravidez. O mecanismo de ação da cloroquina foi alvo de intensas pesquisas durante décadas. Estudos sugerem que o mecanismo de ação da cloroquina envolve o sistema processo de polimerização do heme no vacúolo digestivo (COMBRINCK *et al.*, 2013). Acredita-se que cloroquina possa se ligar a hematina, um dímero de heme e impedir o processo de desintoxicação de heme livre, levando o parasito a morte pelo acúmulo de heme tóxico (EGAN, 2008; FITCH, 2004).

A cloroquina também possui a mais longa meia-vida entre os antimaláricos, proporcionando efeito quimioprolático durante quase 60 dias. Por outro lado isso também favorece o aumento da resistência, uma vez que o parasito pode ficar exposto a baixas concentrações deste fármaco por longos períodos de tempo. Atualmente, as cepas de *P. falciparum* resistentes a cloroquina já estão distribuídas em quase todas as regiões endêmicas do mundo (ALKADI, 2007; PETERSEN, EASTMAN, LANZER, 2011; WELLEMS, PLOWE, 2001). O principal mecanismo de resistência à cloroquina está

relacionado a mutações no gene *Plasmodium falciparum* cloroquine resistance transporter (*pfcr1*), que confere alterações nos transportadores de membrana; e mutações no gene *Plasmodium falciparum* multidrug resistance I (*pfmdr I*) podem aumentar os níveis de resistência, pois dificulta o acesso ao grupo heme (AVARBOGER; SHAPIRO 2005).

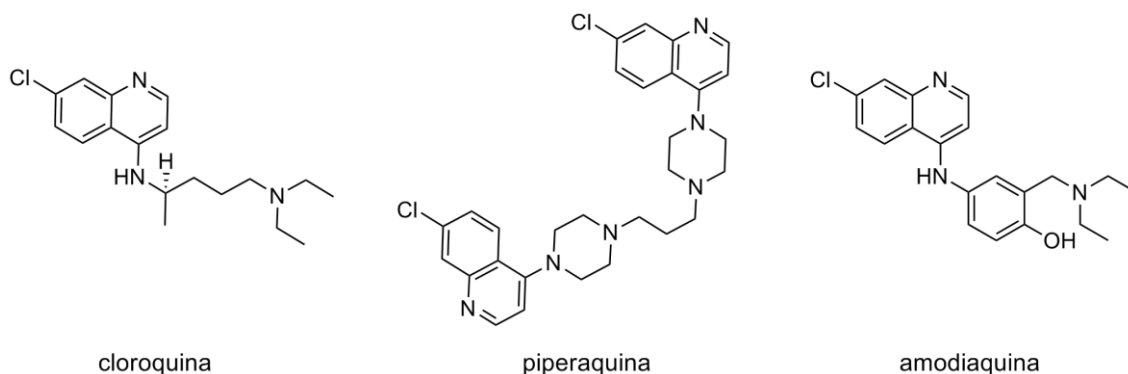


Figura 4 – Estruturas químicas da cloroquina, piperaquina e amodiaquina.

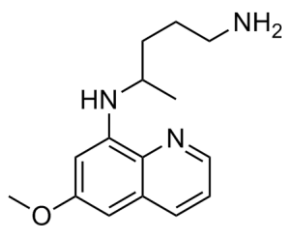
A piperaquina (Figura 4) é outro antimalárico da classe das 4-aminoquinolinas e possui meia-vida de 5 semanas. Estudos demonstram que este fármaco possui mecanismo de ação semelhante à cloroquina (FIVELMAN, ADAGU, WARHURST, 2007). A piperaquina é utilizada na China desde a década de 70, fato que resultou no aparecimento de cepas resistentes de *P. falciparum*. Hoje este fármaco é aplicado na terapia combinada com endoperóxidos (DAVIS *et al.*, 2005; WARHURST *et al.*, 2007; YEKA *et al.*, 2013).

A amodiaquina (Figura 4) é um pró-fármaco estruturalmente semelhante a da cloroquina, diferindo-se apenas por um grupo 4-hidróxi-anilino ligado ao anel quinolínico. Possui uma meia-vida curta (3 horas) e mecanismo de ação exercido pelo metabólito desetilamodiaquina de forma semelhante a cloroquina. Atualmente a amodiaquina apresenta um perfil de resistência semelhante à cloroquina, uma vez que ambos apresentam resistência cruzada (STEPNIEWSKA, WHITE, 2008; SÁ *et al.*, 2009; HAYESHI *et al.*, 2008).

1.1.3.2. 8-aminoquinolinas

A primaquina (Figura 5) é uma 8-aminoquinolina com meia-vida de 6 horas (EDWARDS *et al.*, 1993). Atualmente, este fármaco é a única terapia aprovada para o tratamento das formas hepáticas recorrentes de *P. vivax* (WELLS, BURROWS, BAIRD, 2010). Vários estudos demonstraram os benefícios da associação da

primaquina e cloroquina, levando a um estado de reversão da resistência à cloroquina (EGAN, 2006; BRAY *et al.*, 2005; SANCHEZ *et al.*, 2004). Entretanto, este fármaco é contra-indicado em indivíduos com deficiência da enzima glicose-6-fosfato desidrogenase, uma enzima responsável por manter níveis adequados de NADPH (Nicotinamida Adenina Dinucleotídeo Fosfato na forma reduzida), que se associa a quadros graves de anemia hemolítica (BEUTLER *et al.*, 2007; MANGANELLI *et al.*, 2013 VON SEIDLEIN *et al.*, 2013;).



primaquina

Figura 5 – Estrutura química da primaquina.

1.1.3.3. Álcoois quinolínicos

A quinina (Figura 6) é um dos fármacos mais antigos no tratamento da malária e mesmo assim ainda é utilizada para o tratamento dos casos graves de cepas resistentes de *P. falciparum*. O seu mecanismo de ação ainda não é completamente compreendido, sabe-se apenas que é semelhante ao da cloroquina, podendo se acumular no vacúolo digestivo do parasito e inibir a desintoxicação de heme. (FITCH, 2004). Adicionalmente, este fármaco apresenta meia-vida bem curta, o parasito fica menos tempo em contato com o fármaco, propriedade que provavelmente contribuiu para a escassez de resistência (BUTLER, KHAN, FERGUSON, 2010; MENG *et al.*, 2010; BRIOLANT *et al.*, 2011).

Outro álcool quinolínico bastante utilizado é a mefloquina (Figura 6), um antimalárico cuja meia vida varia entre 14-18 dias, 10 vezes mais potente que a quinina e utilizado no combate ao *P. falciparum* resistente à cloroquina (FRANÇA, SANTOS, VILLAR, 2008; STEPNIIEWSKA, WHITE, 2008). Embora tenha sido introduzido na terapêutica da malária desde a década de 70 (COWMAN, GALATIS, THOMPSON, 1994), seu mecanismo de ação ainda não está completamente elucidado. A única confirmação parte de experimentos *in vitro*, os quais demonstraram que a mefloquina pode se ligar ao heme e impedir a desintoxicação (COMBRINCK *et al.*, 2013; FITCH, CHAN, CHEVLI, 1979; EASTMAN, FIDOCK, 2009).

Polimorfismos no gene *Plasmodium falciparum multidrug resistance I (pfmdrI)* estão relacionados à resistência a quinina e mefloquina. (AVAR-BOGER; SHAPIRO, 2005).

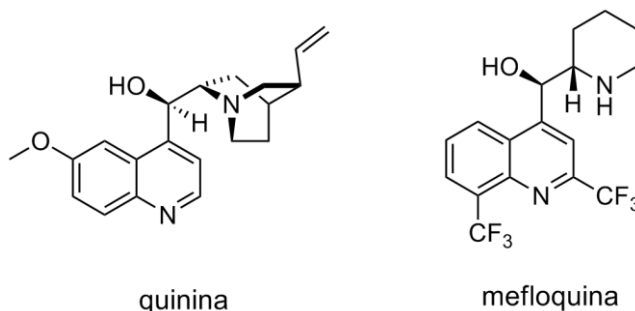


Figura 6 – Estruturas químicas da quinina e mefloquina.

1.1.3.4. Atovaquona

A atovaquona (Figura 7) é um fármaco amplamente utilizado no tratamento de doenças infecciosas, como a toxoplasmose e pneumonia causada por *Pneumocystis carinii*. Seu mecanismo de ação envolve o impedimento do transporte de elétrons na mitocôndria através da inibição do complexo III (citocromo bc1), inibindo a biossíntese de pirimidinas. (NIXON *et al.*, 2013). Este fármaco possui meia-vida entre 2-3 dias e seu uso é recomendado em associação com a proguanila. Trata-se de uma opção terapêutica de custo elevado e com alta susceptibilidade ao aparecimento de resistência (GIL *et al.*, 2003; PAINTER *et al.*, 2007; POROLLO *et al.*, 2012).

Mutações no gene *Citocromo b* conferem resistência a atovaquona, que ocasionam mutações no sítio ativo da enzima-alvo, resultando na diminuição da afinidade de ligação do fármaco. (AVAR-BOGER; SHAPIRO, 2005).

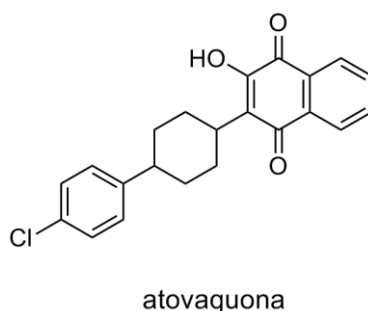


Figura 7 – Estrutura química da atovaquona.

1.1.3.5. Antifolatos

Os fármacos antifolatos são divididos em dois grupos de acordo com seus mecanismos de ação. O primeiro grupo inclui os inibidores competitivos da diidropteroato sintetase (DHPS). Exemplos de inibidores desta enzima são a dapsona e

sulfadoxina (Figura 8). O segundo grupo de fármacos antifolatos é representado pelos inibidores seletivos da enzima diidrofolato redutase (DHFR), a qual também é envolvida no processo de síntese de ácidos nucleicos e aminoácidos. Os principais inibidores desta enzima são proguanila e pirimetamina (Figura 8) (FRANÇA, SANTOS, VILLAR, 2008).

Mutações pontuais nos genes das enzimas-alvo *Pfdhfr* (DHFR) e *Pfdfps* (DHPS) estão associadas a resistência à pirimetamina e sulfadoxina, respectivamente, em *P. falciparum* (BASCO; RINGWALD, 1998).

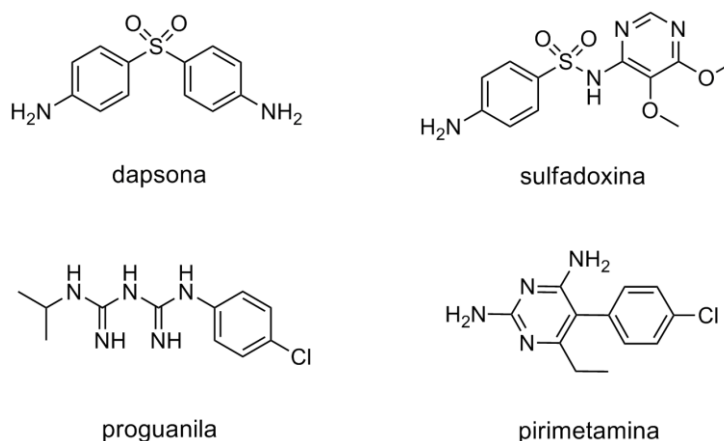


Figura 8 – Estruturas químicas da dapsona, sulfadoxina, proguanila e pirimetamina.

Devido ao efeito sinérgico obtido pela inibição das enzimas diidropteroato sintetase e diidrofolato redutase, os fármacos antifolatos são utilizados em associação. A combinação sulfadoxina/pirimetamina foi introduzida na década de 70 com nome comercial de Fancidar[®], após o surgimento de parasitos resistentes à cloroquina. Esta associação era considerada eficaz, barata, bem tolerada e com taxa adesão muito alta devido a sua dose única. Entretanto, devido ao rápido aparecimento de resistência, atualmente essa associação é usada exclusivamente para gestantes (UHLEMANN, KRISHNA, 2005). Já associação de dapsona e proguanila não é recomendada desde que se descobriu que ela poderia induzir hemólise em indivíduos com deficiência da enzima glicose-6-fosfato desidrogenase (LUZZATTO, 2010).

1.1.3.6. Endoperóxidos

A artemisinina e derivados (Figura 9) são fármacos esquizonticidas e gametocidas da classe dos endoperóxidos considerados juntamente com a quinina, o tratamento de escolha na quimioterapia de indivíduos infectados por cepas de *P.*

falciparum multirresistentes (BRAY, WARD, O' ILL, 2005). São lactonas sesquiterpênicas, com grupo endoperóxido essencial para atividade antimalárica. Desde a sua descoberta e isolamento da planta *Artemisia annua*, até sua caracterização como antimalárico eficaz, a artemisinina foi base para a concepção de novos derivados semissintéticos como diidroartemisinina, artesunato e arteméter (Figura 9), com solubilidade, absorção e distribuição tecidual otimizadas (EASTMAN, FIDOCK, 2009). O mecanismo de ação dos fármacos desta classe envolve a clivagem do grupo endoperóxido, levando a formação de espécies reativas de oxigênio capazes de alquilar proteínas específicas e fundamentais para a sobrevivência do parasito (O' NEILL, BARTON, WARD, 2010).

A resistência à artemisinina está relacionada a uma maior expressão do gene *pfmdr I* (AFONSO *et al.*, 2006) e a resistência ao artesunato foi comprovada que está ligada a uma mutação no gene da ubiquitina (HUNT *et al.*, 2007).

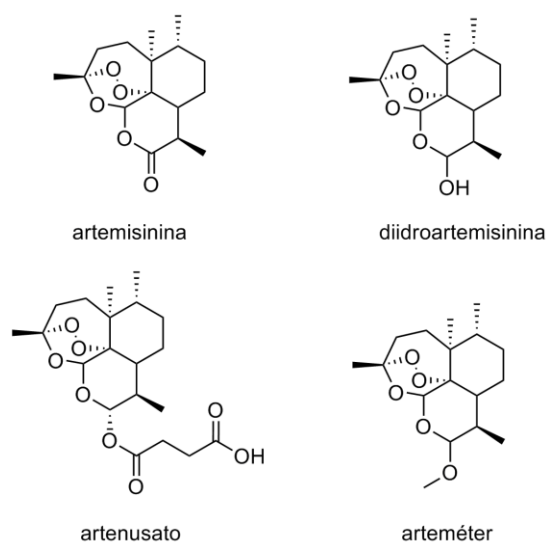


Figura 9 – Estruturas químicas da artemisinina, diidroartemisinina, artesunato e arteméter.

Recentemente, novos estudos apontam o aparecimento e expansão de cepas de *P. falciparum* resistentes aos fármacos endoperóxidos (CARRARA *et al.*, 2013; DONDORP *et al.*, 2009; FALL *et al.*, 2013; KYAW *et al.*, 2013; MALMBERG *et al.*, 2013) e caracterização do efeito neurotóxico apresentado através da administração intramuscular ou exposição por via oral a longo prazo (EFFERTH, KAINA, 2010; LI, HICKMAN, 2011).

1.1.3.7. *Novos fármacos antimaláricos em fase clínica de desenvolvimento*

Diante do aparecimento de cepas multirresistentes aos fármacos antimaláricos disponíveis, a pesquisa de novos fármacos tem sido impulsionada nos últimos anos. Entretanto, o panorama atual dos novos candidatos a fármacos antimaláricos envolvidos em estudos clínicos é prevalentemente representado por pesquisas envolvendo associações entre fármacos já existentes, adequação de doses para pacientes pediátricos e na pesquisa por novas formulações (formas dispersível, injetável e intrarretal). A Tabela 1 apresenta os principais candidatos a fármacos antimaláricos que estão em fase clínica de desenvolvimento, no ano de 2014 (MMV, 2014).

Tabela 1 – Candidatos a fármacos antimaláricos que estão em estudos clínicos (Fonte: MMV,2014).

Fase I	Fase II	Fase III	Registro	Fase IV
DSM 265 Takeda	OZ439/PQP (Monash University/UNMC/STI)	tafenoquina GSK	artemeter + lumefantrina Novartis	rectal artesunato MMV/WHO-TDR
CDRI 97-78 IPCA	OZ439/FQ sanofi	diidroartemisinina + piperaquina pediátrica Sigma-Tau	artemeter + lumefantrina dispersível Novartis	sulfadoxina pirimetanina + amodiaquina Guilin
MMV390048 (UCT)	ferroquina Sanofi Aventis	pironaridina + artesunato pediátrica University of Iowa	artesunato injetável Guilin	
n-terc-butil isoquina LSTM/GSK	fosmidomicina + piperaquina Jomaa Pharma	co-trimoxazol Institut of Tropical Medicine	diidroartemisinina + piperaquina Sigma-Tau	
	azul de metileno University of Heidelberg	arterolane + piperaquina Ranbaxy	pironaridina + artesunato Shin Poong/ University of Iowa	
	SAR97276 Sanofi Aventis	naftoquina + artemisinina KPC	ASAQ Sanofi Aventis /DNDi	
	artemisona UHKST			
	DF02 Dilafor			
	KAE609 Novartis			
	KAF 156 Novartis			

1.2 Planejamento racional de novos antimaláricos

O desenvolvimento das áreas “ômicas” como genômica (CROOKE, 1998), proteômica (WILKINS *et al.*, 1996), metabolômica (WISHART, 2008), juntamente com o avanço de técnicas da bioinformática e de métodos de elucidação estrutural, tais como cristalografia de raios-X e ressonância magnética nuclear (RMN), tem possibilitado a elucidação estrutural e mecanismos bioquímicos de um número cada vez maior de alvos moleculares (SEOANE *et al.* 2013). A descoberta de um alvo molecular promissor e a elucidação do seu papel no processo fisiopatológico estudado, por sua vez, é fundamental para o planejamento racional de fármacos (COHEN, 1996; PATRICK, 2013). A abordagem racional no planejamento de fármacos fundamenta-se na identificação de alvos moleculares para a ação de novos agentes terapêuticos, podendo buscar subsídios na aplicação de técnicas computacionais, para as quais a modelagem molecular é uma ferramenta importante (PATRICK, 2013).

O processo de planejamento de compostos com atividade antimalárica baseia-se, principalmente, na investigação e comparação dos mecanismos bioquímicos associados à sobrevivência do parasito. Graças aos avanços no conhecimento acerca da bioquímica do *P. falciparum*, alcançados pelo sequenciamento do seu genoma e proteoma, diversas vias metabólicas até então desconhecidas, puderam ser investigadas (GARDNER *et al.*, 2002).

1.2.1 Bases Moleculares para o Planejamento Racional de Novos Antimaláricos

1.2.1.1 Diidroorotato desidrogenase de *Plasmodium falciparum* (PfDHODH)

As pirimidinas podem ser sintetizadas por duas vias bioquímicas existentes na maioria dos organismos, a via de salvação e a via *de novo*. Porém, as espécies de *Plasmodium* obtêm pirimidinas unicamente pela via *de novo*, em contraste, às células humanas são capazes de sintetizar pirimidinas a partir das duas vias. Assim, a via *de novo* de biossíntese de pirimidinas é considerada uma via potencial para o desenvolvimento de agentes terapêuticos antimaláricos (CARVALHO, 2008; PHILLIPS, RATHOD, 2010; VYAS, GHATE, 2011). A enzima diidroorotato desidrogenase (DHODH) atua na quarta etapa de síntese de pirimidinas, compostos necessários para a síntese de DNA, RNA, glicoproteínas e fosfolipídios. É portanto, um alvo molecular, validado, para o planejamento de novos antimaláricos. É uma

flavoenzima que catalisa a oxidação de diidroorotato (DHO) em orotato (ORO) (Figura 10) (CARVALHO, 2008; PHILLIPS, RATHOD, 2010; VYAS, GHATE, 2011).

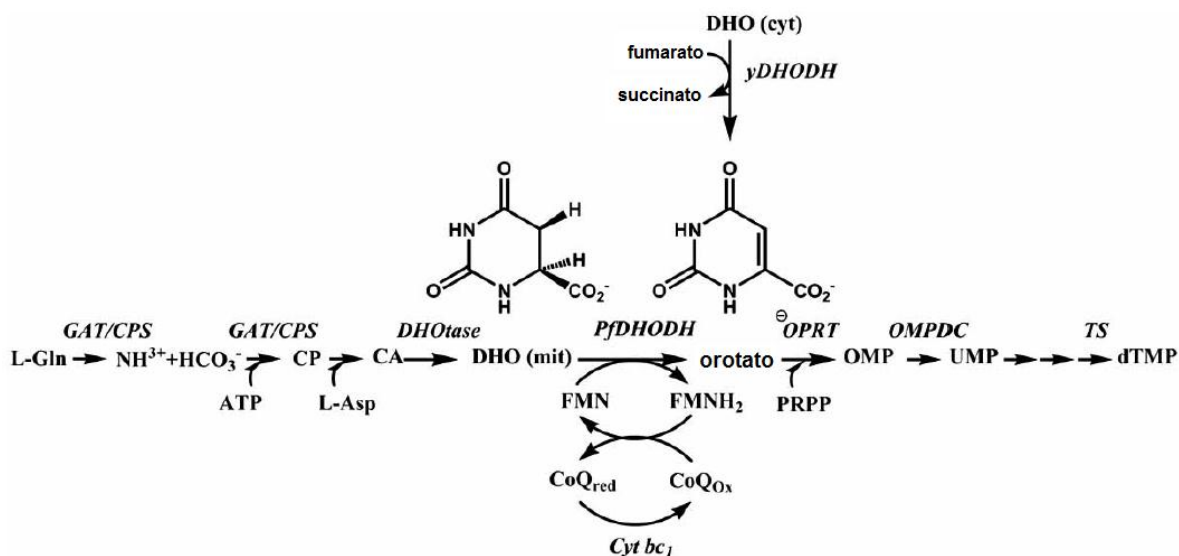


Figura 10 – Biossíntese de pirimidinas pela via *de novo* em *Plasmodium*. O processo envolve seis enzimas: a glutamina amidotransferase bifuncional (GAT), carbamoil-fosfato sintetase (CPS), carbamoiltransferase aspartato (ACT), diidroorotase (DHOTase), diidroorotato desidrogenase (DHODH), orotato fosforibosiltransferase (OPRT) e orotidina-5'-monofosfato descarboxilase (OMPDC). L-glutamina, L-glutamina; CP, carbamoilfosfato; Asp, ácido L-aspartático; CA, carbamoil aspartato; DHO, diidroorotato; OA, ácido orótico; OMP, orotidilato; UMP, uridilato. A (DHODH) atua na quarta etapa, utiliza a flavina mononucleotídeo (FMN) como cofator nas duas etapas da reação, redução e oxidação (Adaptado de: PHILLIPS, RATHOD, 2010).

A ausência de pirimidinas pela via de salvação em *Plasmodium* foi primeiramente demonstrada pela análise da incorporação de precursores radiomarcados em ácidos nucleicos e por ensaio enzimático, e foi também confirmado pelo término do sequenciamento do genoma (PHILLIPS, RATHOD, 2010). A enzima DHODH catalisa as reações de oxidação/redução na biossíntese dos nucleotídeos de pirimidina. Durante essas reações, a DHODH utiliza a flavina mononucleotídeo (FMN) como cofator nas duas etapas da reação. Na primeira etapa, o diidroorotato (DHO) é reduzido para orotato (ORO), enquanto que a flavina mononucleotídeo (FMN) é reduzida (FMNH₂). Na segunda etapa, o FMN é reoxidado através do auxílio de um segundo substrato, um agente aceptor de elétrons (Figura 10) (CARVALHO, 2008; PHILLIPS, RATHOD, 2010; VYAS, GHATE, 2011).

As enzimas DHODH podem ser classificadas em duas classes, classe 1 e classe 2, considerando-se sua estrutura primária, localização celular, natureza do segundo substrato e origem biológica. As enzimas DHODH da classe 1 são encontradas principalmente em bactérias Gram-positivas, estão associadas ao citosol e são subdivididas nas classes 1A, 1B e 1S. As DHODH de classe 2 são encontradas principalmente em bactérias gram-negativas e eucariotos e estão relacionadas à membranas citosólicas e mitocondriais. Caracterizam-se pela extensão na região *N*-terminal que permite junção com membranas (CARVALHO, 2008; PHILLIPS, RATHOD, 2010).

De maneira geral, ambas classes de enzimas apresentam como elemento estrutural central o domínio β/α -barril em que está o sítio de ligação para o cofator FMN. Essa estrutura terciária consiste de uma região central formada por 8 fitas beta paralelas rodeadas por 8 alfa hélices (domínio catalítico). Na parte superior do barril, o sítio redutor do orotato é formado por três fitas β antiparalelas, que formam o bolso de ligação do substrato e do FMN. Nessa região também é encontrada a alça catalítica, que atua como “tampa” do barril, que controla a entrada do substrato e saída do produto do sítio catalítico (Figura 11). A base do barril é formada por um par de fitas β antiparalelas (CARVALHO, 2008; PHILLIPS, RATHOD, 2010).

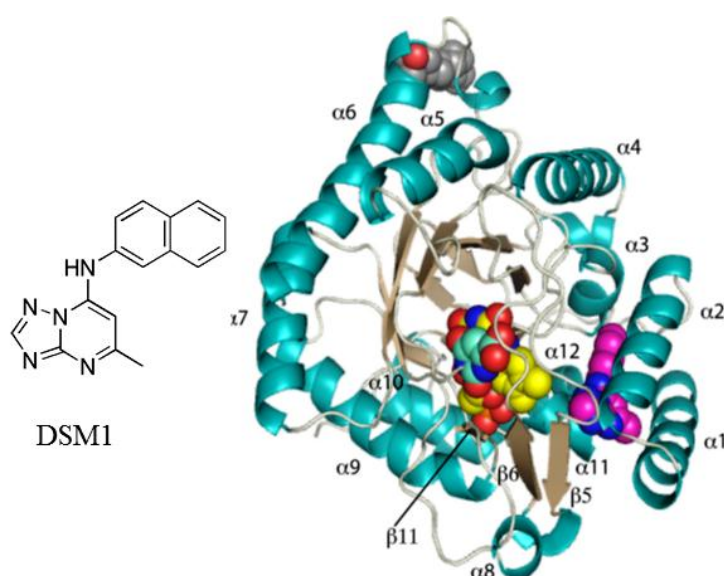


Figura 11 – Representação da estrutura 3D em fita da *PfDHODH* co-cristalizada com o inibidor DSM1 (PDB ID: 3I65). As α -hélices são apresentadas em azul ciano, as folhas- β em marrom, os ligantes são apresentados em modelo CPK: o FMN com átomos de carbono em azul, o DSM1 com átomos de carbono em rosa, o orotato em turquesa (DENG *et al.*, 2009).

A enzima DHODH é um alvo validado (GOLDENBERG *et al.*, 1999) para o tratamento de doenças imunossupressoras e seus inibidores possuem potencial terapêutico como agentes anti-inflamatórios e antiparasitários (PHILLIPS; RATHOD, 2010). O inibidor da DHODH humana (*HsDHODH*) leflunomida (ARAVA®) está em uso clínico para o tratamento da artrite reumatoide (seu metabólito ativo é o A77 1726) (Figura 12). O brequinar também é um inibidor da *HsDHODH* e possui ação antitumoral e imunossupressiva. Os inibidores da DHODH de *Helicobacter pylori* e da *Enterococcus faecalis* causam inibição do crescimento bacteriano (CHRISTOPHERSON, LYONS, WILSON, 2002; HEIKKILA *et al.*, 2006).

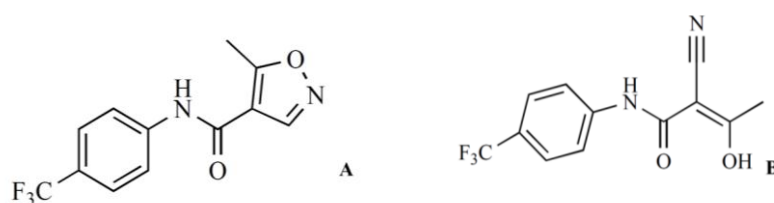


Figura 12 – Estrutura do inibidor de *HsDHODH* leflunomida (A) e seu metabólito ativo (B).

1.2.1.2 Triazolopirimidinas, inibidores *PfDHODH*

Estudos realizados por pesquisadores das Universidade do Texas South Western, Universidade de Washington e Monash Universidade, identificaram através de ensaios de triagem biológica automatizada em larga escala (HTS, do inglês, *high throughput screening*), o composto, (5-metil-[1,2,4] triazolo [1,5-*a*] pirimidin-7il)naftalen-2-amina (Composto 26), uma triazolopirimidina, TPZ, (IC₅₀ 0,047 μM), potente e seletiva para *PfDHODH* (PHILLIPS *et al.*, 2008).

A partir deste protótipo, uma série de TPZ foram sintetizadas e testadas em *PfDHODH* (Phillips *et al.*, 2008, GUJJAR *et al.*, 2009, DENG *et al.*, 2009, GUJJAR *et al.*, 2011, COTERON *et al.*, 2011, e MARWAHA *et al.*, 2012). Nestes estudos, todas as sínteses foram guiadas por estudos de relação estrutura-atividade (SAR). As análises de SAR mostram que substituintes hidrofóbicos nas posições R e R₁ favorecem a potência na enzima *PfDHODH* (Tabela 2).

Tabela 2 – Triazolopirimidinas com os respectivos valores de IC₅₀ em *Pf*DHODH.

Composto	R	R ₁	R ₂	R ₁	R ₂	IC ₅₀ (μM)
26	H	H	CH ₃	fenil	-	0,047±0,02
27	H	H	CH ₃	naftil	-	0,056±0,02
22	H	H	CH ₃	CF ₃	H	0,28±0,02
82	CF ₂ CH ₃	H	CH ₃	CF ₃	H	0,038
83	CF ₂ CH ₃	H	CH ₃	SF ₅	H	0,033

O modo de ligação dessas TPZs foi elucidado por cristalografia de raios-X da *Pf*DHODH com os seus inibidores (Compostos 26, 27, 22 e 82). Estes compostos se ligam através de duas cavidades hidrofóbicas do sítio de interação e por ligação de hidrogênio dos nitrogênios 1 e 3 (N) do anel triazolopirimidínico com His185 e o nitrogênio 5 com Arg 265 (DENG *et al.*, 2009b) (Figura 13 - A).

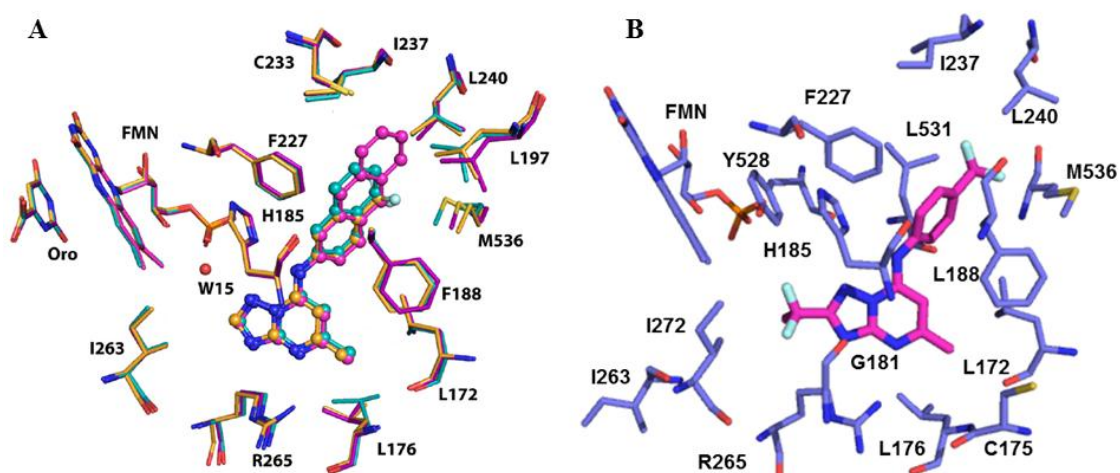


Figura 13 – Estruturas 3D obtidas a partir da cristalografia de raios-X da *Pf*DHODH em complexo com os inibidores 22 (DSM 1), PDB ID 3I65, (azul); 26 (DMS 2), PDB ID 3I68, (rosa); 27 (DMS 74), PDB ID 3I6R, (alaranjado) (A) e 82 (DMS 267), PDB ID 3SFK (B). (Adaptado: (COTERON *et al.*, 2011; DENG *et al.*, 2009b)

Estudos de QSAR realizados com TPZ relevaram que descritores espaciais, eletrônicos, termodinâmicos e estruturais são importantes para determinar a atividade inibitória das TPZs em *PfDHODH* (OJHA; ROY, 2010; SHAH *et al.*, 2012). Estudos de *docking* inferiram que hidrofobicidade é necessária para maior potência (OJHA; ROY, 2010). Os compostos planejados com esses estudos mostraram boa preditividade *in silico* de acordo com os modelos de QSAR (OJHA; ROY, 2010).

Atualmente o composto 83 (DMS 265) encontra-se em estudos clínicos de fase I. Este composto inibe a *PfDHODH* seletivamente em relação a seu ortólogo humano (IC_{50} 33nM e 2500nM, respectivamente). Ele é biodisponível por via oral (ratos: F=57-68%, $t_{1/2}$ = 12-28h), e eficaz *in vitro* (EC_{50} – 43nM (3D7) e *in vivo* em camundongos (ED_{50} = 2,8 mg/kg/dia) (COTERON *et al.*, 2011).

1.3. Pesquisa, Desenvolvimento & Inovação de novos fármacos

O processo de PD&I de novos fármacos pode ser dividido em duas grandes fases: (i) a fase pré-clínica, identificação e validação de alvos moleculares, descoberta de moléculas capazes de representar novas entidades químicas (NCE, do inglês *New Chemical Entities*) e otimização; e (ii) a fase clínica, visa garantir a segurança e eficácia de um novo candidato a fármaco em seres humanos (LOMBARDINO, LOWE, 2004; MENDOZA, 2011). As duas fases estão esquematizadas na Figura 14.

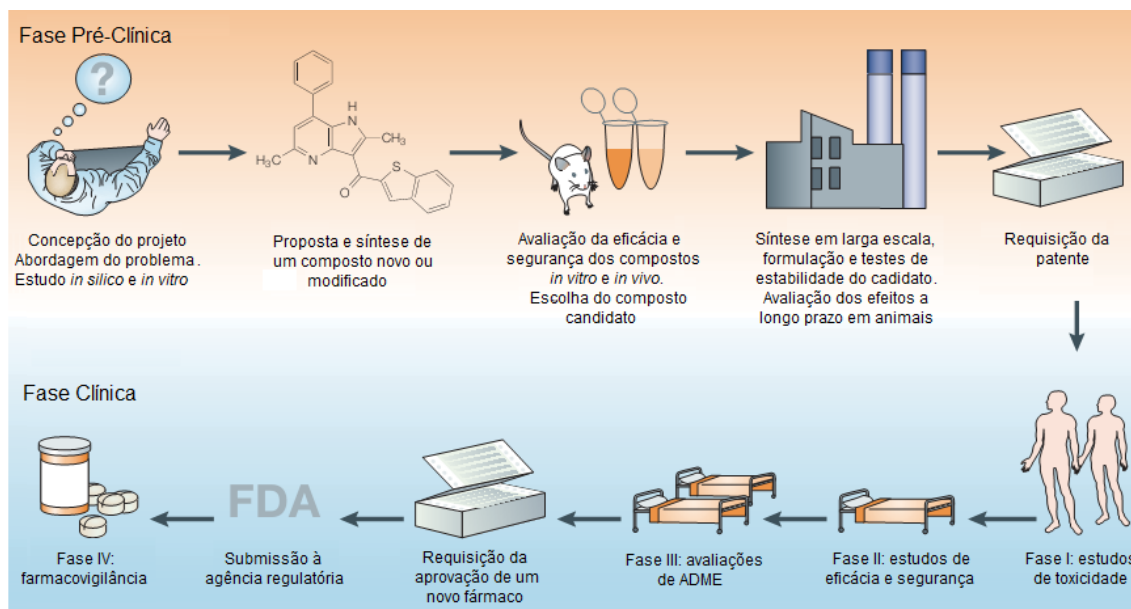


Figura 14 - Etapas do processo de descoberta e desenvolvimento de novos fármacos (LOMBARDINO; LOWE, 2004).

Na fase pré-clínica busca-se identificar e otimizar moléculas pequenas com atividade moduladora sobre o alvo biológico selecionado. Para tanto, a validação do alvo biológico é essencial na determinação do seu envolvimento na fisiopatologia da doença a ser tratada. O processo de validação pode ser realizado através da supressão da expressão do gene alvo, estabelecendo assim sua relevância no processo fisiopatológico e avaliando sua capacidade de gerar resposta farmacológica no controle da doença (ELEGRINB, GILL, PLOWRIGHT, 2012; TAMIMI, ELLIS, 2009; RENSLO, MCKERROW, 2006; SIGOILLOT, KING, 2011). Moléculas capazes de modular o alvo estabelecido, por sua vez, podem ser obtidas através de produtos naturais, síntese ou química combinatória. É possível também explorar o imenso espaço químico-biológico através de organização de bases de dados, aplicação de filtros moleculares e emprego de métodos como ensaios biológicos automatizados (HTS, do inglês, *High-Throughput Screening*) e triagem virtual (VS, do inglês, *Virtual Screening*) (BAJORATH, 2002; BADRINARAYAN, SASTRY, 2011; SCIOR *et al.*, 2012; SUKUMAR, DAS, 2011). Uma vez disponíveis e avaliadas experimentalmente quanto à capacidade moduladora do alvo proposto, as moléculas mais promissoras são selecionadas para posterior otimização, buscando-se melhorar suas propriedades farmacocinéticas e farmacodinâmicas em ensaios *in silico*, *in vitro* e *in vivo* (HUGHES *et al.*, 2011; LOMBARDINO, LOWE, 2004).

Após avaliação inicial do potencial terapêutico dos novos candidatos a protótipos de fármacos, estes são submetidos à fase clínica de desenvolvimento para confirmação da segurança e eficácia. A fase clínica é classicamente dividida em quatro fases (I-IV). Na fase I são realizados estudos de toxicidade em um grupo pequeno (20 a 100) de voluntários saudáveis. Na fase II os estudos são realizados em pacientes com a doença ou desordem (100 a 300 indivíduos) para se avaliar a dosagem, eficácia e segurança da composição. Já durante a fase III são realizadas avaliações farmacocinéticas em um grupo maior de indivíduos (300- 3.000 ou mais), provenientes de regiões geográficas distintas e escolhidos randomicamente. Nesta fase o candidato a fármaco é comparado com o tratamento preconizado e determina-se sua efetividade (FDA, 2006; MCGEE, 2006). Os candidatos a fármacos aprovados ao final da fase III são submetidos à análise juntamente com a documentação dos resultados obtidos e os protocolos utilizados nos ensaios, visando obter dos órgãos reguladores a autorização para comercialização. Uma vez aprovado e inserido no mercado, têm-se início a fase IV,

na qual o medicamento será comercializado e extensivamente monitorado quanto a eventuais efeitos indesejados (MCGEE, 2006).

A descoberta e o desenvolvimento de novos fármacos é um processo longo e complexo que demanda alto investimento (LOMBARDINO, LOWE, 2004). Estima-se para o lançamento de um único fármaco no mercado, em média, 12 a 15 anos gastos no processo de PD&I e custos de aproximadamente US\$800 milhões, podendo chegar a US\$ 1,8 bilhão (CBO, 2006; PAUL *et al.*, 2010). Os investimentos requeridos são justificados pelos altos custos dos estudos clínicos, bem como o grande número de compostos descartados do processo de desenvolvimento por falharem durante alguma das etapas do estudo. Estima-se que, de cada 250 compostos selecionados durante a fase pré-clínica, apenas 5 são aprovados para os estudos clínicos subsequentes, resultando, ao final, em um único fármaco aprovado para comercialização (PHRMA, 2011). Além disso, parte do montante contabilizado no processo de PD&I refere-se a custos direcionados ao *marketing* para a inserção do novo fármaco no mercado (ADEUSI, 2011).

A indústria farmacêutica investiu cerca de US\$51,1 bilhões em 2011 na busca de inovações terapêuticas (PHRMA, 2014). Apesar do investimento crescente em PD&I, não há um aumento proporcional no número de novos fármacos inseridos no mercado. Entre 1950 e 2008, estimativas apontam custos de desenvolvimento maiores e estagnação do número de fármacos aprovados pelos órgãos reguladores. Dentre os fatores responsáveis por este quadro, destacam-se o modelo de PD&I adotado, cujas etapas sequenciais das fases de estudo demandam muito tempo, bem como o aumento das exigências regulatórias (MUNOS, 2009).

Para manter a lucratividade, a indústria farmacêutica tem investido na pesquisa e desenvolvimento de *blockbusters*, termo inglês que designa fármacos com rendimentos anuais superiores a US\$1 bilhão (ADEUSI, 2011; EUROPEAN COMMISSION, 2008; PHRMA, 2011). Recentemente, observa-se também o aumento de investimentos em terapias personalizadas com a pesquisa de fármacos '*niche-busters*', os quais são desenvolvidos para mercados específicos, tais como pacientes transplantados, portadores de doença de *Parkinson* e portadores de síndromes genéticas (MITTRA, TAIT, 2012). Nota-se, ainda, que os medicamentos líderes de vendas emergem de classes terapêuticas cujo maior mercado é constituído por países ricos, como fármacos hipolipemiantes, antineoplásicos e moduladores do sistema nervoso central. Assim, o modelo de PD&I adotado pela indústria concentra-se prioritariamente nessas classes

terapêuticas, visando o lançamento de fármacos capazes de garantir rapidamente o retorno do investimento.

1.3.1. Química Medicinal e a Gênese Planejada de Fármacos

Segundo a definição da União Internacional de Química Pura e Aplicada (IUPAC, do inglês *International Union of Pure and Applied Chemistry*), a Química Medicinal possui caráter multidisciplinar e envolve a invenção, descoberta, planejamento e desenvolvimento de compostos biologicamente ativos, incluindo também o estudo do metabolismo e relações entre a estrutura química e a atividade biológica. Assim, a Química Medicinal constitui uma ciência essencial no processo de PD&I de novos fármacos, abrangendo múltiplas áreas de conhecimento (WERMUTH *et al.*, 1998).

O planejamento racional, baseado na elucidação da etiologia e dos processos bioquímicos envolvidos na doença, permite otimizar o processo de descoberta e desenvolvimento de novos fármacos. Através do uso de diversas estratégias é possível utilizar conhecimentos prévios do mecanismo de intervenção terapêutica pretendido e do alvo terapêutico escolhido para propor moléculas capazes de modular adequadamente o alvo eleito (COHEN, 1996; PATRICK, 2009).

O planejamento de novos candidatos a agentes terapêuticos e otimização destes pode ser obtido por diferentes estratégias da Química Medicinal. A escolha da estratégia a ser adotada depende da elucidação da estrutura do alvo terapêutico eleito (BARREIRO, 2009). O conhecimento das estruturas de alvos macromoleculares ou de complexos ligante-receptor permite a aplicação de estratégias de planejamento de fármacos baseado na estrutura (SBDD, do inglês *Structure-based Drug Design*). Quando a estrutura do alvo eleito não é conhecida, métodos de planejamento de fármacos baseado no ligante (LBDD, do inglês *Ligand-based Drug Design*) podem ser utilizados, explorando propriedades e características do substrato natural ou de séries de ligantes bioativos. As estratégias de SBDD e LBDD podem ser usadas conjuntamente no planejamento racional, uma vez que fornecem informações complementares (COHEN, 1996).

1.3.2. Relações Quantitativas entre estrutura química e atividade biológica (QSAR)

O planejamento, desenvolvimento e otimização de novos ligantes bioativos a partir de coleções de compostos pode ser realizado através de métodos que relacionem dados biológicos à estrutura química. Neste contexto, o emprego das relações entre a estrutura e atividade (SAR, do inglês, *Structure-Activity Relationships*) e os estudos de relações quantitativas entre estrutura e atividade (QSAR, do inglês *Quantitative Structure-Activity Relationships*) são ferramentas úteis à descoberta de novos candidatos a protótipos de fármacos. Métodos de QSAR, classificados de acordo com a dimensionalidade dos descritores que utilizam, estabelecem quantitativamente a correlação entre propriedades da estrutura química das moléculas e a atividade biológica que as mesmas possuem. Assim, através de um conjunto de dados são desenvolvidos modelos matemáticos capazes de prever a atividade biológica de moléculas externas ao conjunto que deu origem ao estudo (HANSCH, FUJITA, 1964). A figura 15 esquematiza o processo de desenvolvimento dos modelos de QSAR, em que são necessários: (i) compostos com atividade biológica/propriedade e (ii) descritores (TROPSHA, 2010).

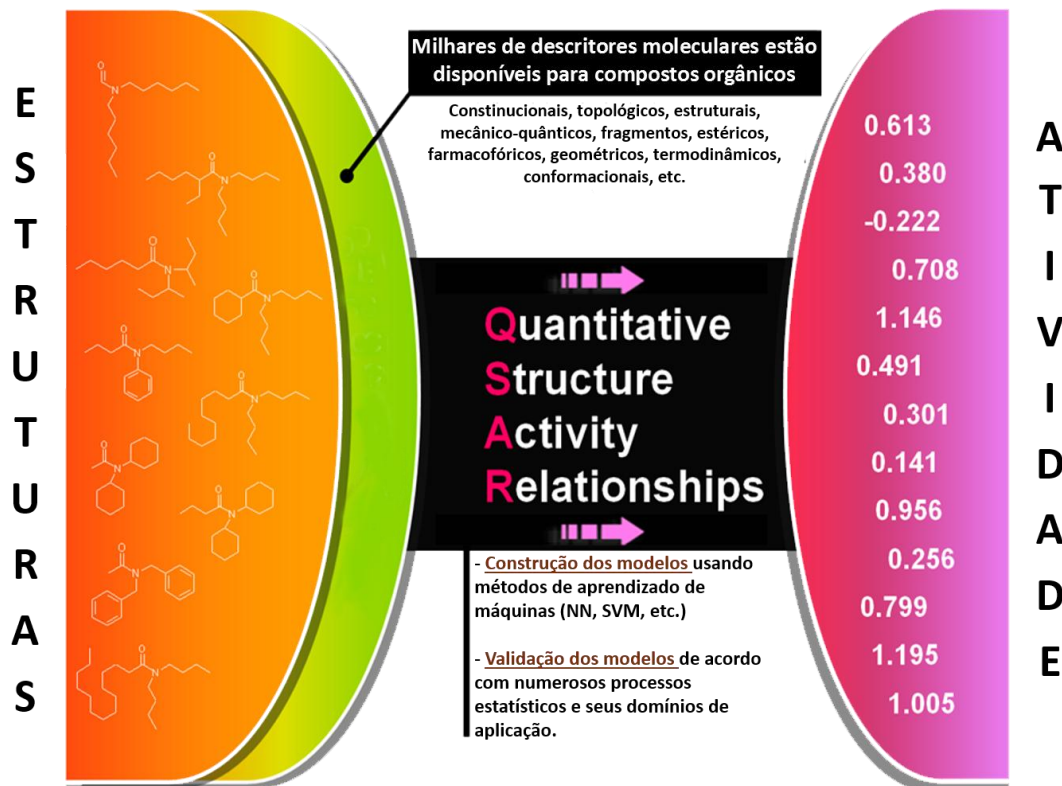


Figura 15 – Processo de desenvolvimento do modelo de QSAR (adaptado TROPSHA, 2010)

A metodologia de QSAR foi inicialmente estabelecida na década de 1960, pelos trabalhos de Hansch e Fujita. Esses pesquisadores, demonstraram que a atividade biológica pode ser linearmente correlacionada com diferentes parâmetros físico-químicos, sendo expressa pela Equação (1).

$$\text{Log } \frac{1}{C} = a\pi + b\delta + xEs + d \quad (1)$$

onde C representa a concentração molar do composto necessária para produzir uma resposta biológica definida, π a contribuição hidrofóbica, δ o efeito eletrônico, Es efeito estérico, d o termo independente e a , b e x os respectivos coeficientes determinados por análise de regressão (HANSCH, FUJITA, 1964; HANSCH, 1969).

Alguns anos depois, Hansch propôs a extensão do modelo linear para o parabólico, quando percebeu que a resposta biológica de compostos com caráter hidrofóbico estagnava ou reduzia após atingir um valor máximo. A partir deste problema, foi percebido que compostos com alta hidrofobicidade podem ficar retidos nas barreiras lipofílicas, portanto têm menor probabilidade de atingir os sítios biológicos de ação (HANSCH, 1969). O modelo parabólico proposto com base numa série de dados pressupõe uma relação de segunda ordem da hidrofobicidade, expressa como o logaritmo do coeficiente de partição entre octanol e água ($\text{Log } P$), com a atividade biológica (DEBNATH, 2001), conforme descrita pela Equação 2.

$$\text{Log } \frac{1}{C} = a(\text{Log } P)^2 + B\text{Log } P + x \quad (2)$$

Na década de 1970, Kubinyi desenvolveu um modelo bilinear que descreve a dependência não-linear da atividade biológica sobre o caráter hidrofóbico, no qual o termo β corresponde ao movimento do composto em sistemas multicompartimentados, como o sistema biológico, de acordo com o descrito pela Equação 3 (KUBINYI, 1997; TAVARES, 2004; DEBNATH, 2001).

$$\text{Log } \frac{1}{C} = a \text{Log } P - b\text{Log}(\beta P + 1) + x \quad (3)$$

No início do QSAR, apenas parâmetros físico-químicos e mecânico-quânticos estavam disponíveis para serem usados como descritores (SELASSIE; ONE,

2003). Então, percebeu-se que nem toda informação molecular era extraída por ensaios experimentais, o que impulsionou esforços para extrair informações codificadas na estrutura da molécula em números – chamados de descritores moleculares – através de um caminho teórico (CONSONNI; TODESCHINI, 2010).

Diferentes tipos de descritores químicos refletem diferentes níveis de representação estrutural. Esses descritores podem ser classificados quanto à sua “dimensionalidade” em unidimensionais (1D), baseados em propriedades físico-químicas e da fórmula molecular (ex., massa molecular, refratividade molar, logP, entre outros); bidimensionais (2D), que descrevem propriedades que podem ser calculadas de uma representação 2D (ex., número de átomos, número de ligações, índices de conectividade, entre outros); e tridimensionais (3D), que dependem da conformação das moléculas (ex., volume de Van der Waals, área de superfície acessível ao solvente, entre outros) (XUE; BAJORATH, 2000).

Outra classificação diz respeito à natureza desses descritores, que podem ser: (i) constitucionais, que são derivados da composição atômica do composto (ex., peso molecular, números de átomos e ligações); (ii) topológicos (ex., índices de Randić); (iii) geométricos, que são derivados de coordenadas 3D (ex., volume molecular, área de superfície polar, entre outros); (iv) eletrostáticos, que são derivados da substituição de cargas parciais (ex., índices de polaridade, carga parciais, entre outros); e (v) quântico-mecânicos, que são derivados das funções de onda dos elétrons (ex., energia dos orbitais moleculares) (CONSONNI; TODESCHINI, 2010).

Muitos programas estão disponíveis para o cálculo de descritores moleculares, como DRAGON[®] (Talete SRL, Milan, Italy), CDK[®] (<http://www.rguha.net/code/java/cdkdesc.html>), CODESSA[®] (<http://www.semichem.com/codessa/>), entre outros.

Na década de 1980, os avanços em *hardware* e *software*, permitiram o estudo das propriedades tridimensionais (3D) (COHEN, 1996; OOMS, 2000). Em consequência, a metodologia descrita como QSAR tridimensional (3D) surgiu. Os descritores do QSAR-3D estão relacionados ao cálculo de propriedades correspondentes à estrutura 3D dos ligantes (método independente do receptor, IR) ou do complexo ligante-receptor (método dependente do receptor, DR). Estes descritores podem ser classificados em globais (moleculares) e locais (atômicos). Os descritores globais representam uma categoria mais simples, em que os descritores (volume molecular, área molecular superficial) são derivados da estrutura molecular como um todo, enquanto

que os descritores locais são derivados de fragmentos ou partes da estrutura molecular (átomos, grupos farmacofóricos) (ALBUQUERQUE *et al.*, 2007; VERLI *et al.*, 2002).

Nesta mesma década, Hopfinger introduziu a análise conformacional das estruturas e a obtenção de descritores globais 3D da forma molecular aos estudos de QSAR, criando o método denominado Análise da Forma Molecular (MSA, do inglês, *Molecular Shape Analysis*) (HOPFINGER, 1980). Em 1988, Cramer e colaboradores descreveram a metodologia de Análise Comparativa de Campos Moleculares (*Comparative Molecular Field Analysis*, CoMFA), a técnica de QSAR-3D mais utilizada no mundo. Essa técnica demonstra que a propriedade biológica dos compostos pode ser correlacionada com as energias estérica e eletrostática (descritores locais) provenientes das interações formadas com o ligante no sítio ativo do alvo biológico (CRAMER, PATTERSON, BUNCE, 1988).

Em 1997, a metodologia de QSAR-4D foi introduzida por Hopfinger e colaboradores (HOPFINGER *et al.*, 1997). Esta técnica baseia-se em na abordagem de QSAR-3D, que utiliza uma amostragem conformacional obtida através de simulação por Dinâmica Molecular (DM), reduzindo a dificuldade em identificar a conformação bioativa. Dimensões adicionais foram incorporadas aos métodos, QSAR, como por exemplo, o Quasar QSAR-5D (VEDANI, DOBLER, 2002), no qual a teoria do encaixe induzido é considerada, ou seja, a adaptação do sítio de ligação do receptor a cada ligante individual e, mais recentemente, o Quasar QSAR-6D (VEDANI, DOBLER, LILL, 2005), que considera os modelos de solvatação simultaneamente.

A utilização de técnicas de QSAR parte do preceito de que todos os compostos utilizados na análise interagem da mesma maneira no sítio ativo de determinado alvo biológico. Entretanto, cada método QSAR difere na forma de descrever as propriedades estruturais (descritores) dos compostos e como encontrar relações quantitativas entre as propriedades e a respectiva atividade (WERMUTH, 2008).

1.3.2.1 Holograma QSAR (HQSAR)

O HQSAR (HQSAR, Tripos®) é um método de QSAR-2D no qual a estrutura molecular bidimensional é convertida em notações lineares, ou seja, sequências numéricas que codificam a estrutura 2D das moléculas do conjunto de dados, e as informações codificadas pelos descritores bidimensionais são relacionadas

quantitativamente à atividade biológica (LOWIS, 1997; SEEL, TURNER, WILLETT, 1999).

O método se baseia no conceito de impressão digital molecular (do inglês, *fingerprint*) da estrutura 2D molecular, conhecida como holograma molecular (LOWIS, 1997; SEEL, TURNER, WILLETT, 1999), também é considerada uma técnica de QSAR *fragmente-based*.

Todos os fragmentos moleculares possíveis (lineares, ramificados, cíclicos e sobrepostos) são obtidos e posteriormente distribuídos pelo holograma. Cada posição do holograma contém a frequência com que um determinado fragmento aparece na molécula, sendo que fragmentos iguais que ocorrem mais de uma vez são sempre alocados na mesma posição. Informações tridimensionais, tais como hibridação e quiralidade, também são codificadas nos hologramas moleculares (LOWIS, 1997; Tripos, 2010a). Uma vez que as estruturas químicas são representadas em seus hologramas moleculares correspondentes, o HQSAR não requer o cálculo e a seleção de descritores físico-químicos comumente usados como descritores em outros modelos de QSAR-2D. Além disso, para a geração de modelos de HQSAR não são necessários o conhecimento das conformações bioativas das estruturas e o alinhamento molecular, o que torna o HQSAR um método rápido e aplicável a grandes conjuntos de dados (LOWIS, 1997; Tripos, 2010a).

Os modelos de HQSAR são desenvolvidos empregando-se o método de regressão multivariada por quadrados mínimos parciais (*PLS*, do inglês *Partial Least Squares*). Para análises de QSAR, o *PLS* é um método bastante utilizado por possibilitar a análise de conjuntos de dados com número de variáveis independentes muito superior ao número de variáveis dependentes, uma vez que o *PLS* reduz os descritores a variáveis latentes ortogonais (HASEGAWA, FUNATSU, 2012).

O processo de obtenção do holograma molecular e dos modelos de HQSAR está representado na Figura 16.

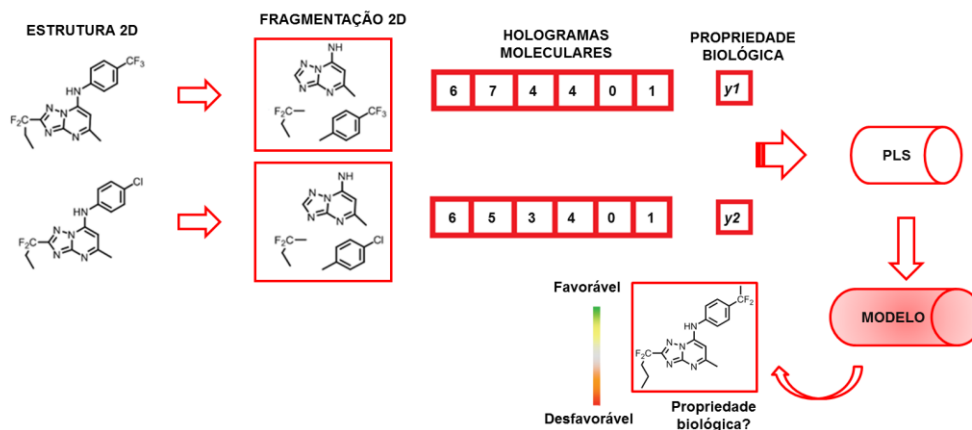


Figura 16 – Esquema do processo de geração dos hologramas moleculares e dos modelos de HQSAR.

Durante o processo de modelagem de HQSAR, vários parâmetros relevantes para a obtenção dos hologramas, que afetam a qualidade dos modelos, podem ser definidos, tais como:

- comprimento do holograma - define o número de posições disponíveis para que diferentes fragmentos sejam incorporados;
- tamanho dos fragmentos - define o número máximo (N) e mínimo (M) de átomos contidos em cada fragmento a ser gerado;
- distinção dos fragmentos – define o tipo de informação 2D a ser incluída na criação dos descritores.

Como, em geral, o comprimento do holograma é menor que o número total de fragmentos diferentes gerados, fragmentos distintos são alocados no mesmo holograma. A maneira com que os fragmentos são distribuídos nas posições do holograma influencia diretamente a qualidade dos modelos. O algoritmo cíclico de verificação de redundância (CRC, do inglês, *cyclic redundancy check*) avalia várias combinações de distribuição de fragmentos nas diferentes posições do holograma molecular para impedir que haja a colisão de fragmentos, ou seja, evita que fragmentos distintos e importantes sejam alocados em uma mesma posição no holograma. Além disso, diferentes números primos de comprimentos de holograma são utilizados, o que reduz a chance da colisão de fragmentos e consequente perda de informação (LOWIS, 1997; SEEL, TURNER, WILLETT, 1999).

Os parâmetros utilizados para distinguir os fragmentos são ajustáveis, e de fundamental importância, pois definem o tipo de informação 2D que será incluída na criação dos descritores. Na Tabela 3 são apresentados os diferentes parâmetros de

distinção dos fragmentos considerados no método HQSAR. Combinações desses parâmetros são utilizadas, para a escolha daquela que resulte nos melhores resultados estatísticos (LOWIS, 1997; SEEL, TURNER, WILLETT, 1999).

Tabela 3 – Parâmetros de distinção dos fragmentos no método de HQSAR

Distinção dos fragmentos	Definição dos parâmetros de distinção de fragmentos
Átomo (A)	Fragmentos diferenciados com base no tipo de átomo
Ligação (B)*	Fragmentos diferenciados com base no tipo de ligação química simples, dupla, tripla e aromática)
Conectividade (C)*	Fragmentos diferenciados com base no tipo de hibridização dos átomos contidos no fragmento
Hidrogênio (H)	Fragmentos são distinguidos com base na presença e número de átomos de hidrogênio
Quiralidade (Ch)	Fragmentos são distinguidos de acordo com a presença de centros estereogênicos
Doadores e aceptores (DA)	Fragmentos são diferenciados com base na presença átomos doadores e aceptores de ligação de hidrogênio

*Parâmetros semelhantes. Ao incluir o parâmetro Ligação (B), porém, a vizinhança dos átomos é levada em consideração, o que permite distinguir fragmentos que possuam o mesmo tipo de hibridização, mas estejam localizados em diferentes ambientes químicos. Por exemplo, H-C=C-H (etileno) e H-C:C-H (benzeno). Neste caso ambos os átomos de carbono possuem hibridização sp^2 , o que os diferencia é que no etileno a ligação é dupla e no benzeno a ligação é aromática.

Após definir os parâmetros a serem utilizados, o método de PLS é usado para a geração do modelo matemático de HQSAR, considerando os hologramas gerados, que são os descritores, como variáveis independentes e a atividade biológica como variável dependente, identificando-se os padrões de fragmentos subestruturais correlacionados com a resposta biológica. Como nos demais métodos de QSAR, a atividade biológica das moléculas externas ao conjunto de dados utilizado na construção do modelo, porém estruturalmente semelhantes, dentro do espaço químico ou do domínio de aplicabilidade a este conjunto, poderá ser prevista (LOWIS, 1997; SEEL, TURNER, WILLETT, 1999).

Um modelo de QSAR é capaz de prever a atividade de qualquer composto químico. No entanto, se sua estrutura química for muito diferente das estruturas utilizadas na construção do modelo, essa predição pode ser incorreta, pois o modelo não cobre o espaço químico de tal estrutura. Logo, a determinação do domínio de aplicabilidade é necessário para o modelo ser aceito (TROPISHA, 2010).

Além de permitir a previsão da atividade biológica, o HQSAR gera mapas de contribuição 2D, fornecendo dados topológicos a partir dos quais é possível visualizar os fragmentos moleculares importantes para a atividade biológica em estudo. Assim,

essas informações podem orientar a síntese de novas moléculas com propriedades otimizadas, o que representa uma vantagem adicional do HQSAR diante de outros métodos de QSAR-2D (HONORIO, GARRATT, ANDRICOPULO, 2005; LOWIS, 1997; SEEL, TURNER, WILLETT, 1999).

2. OBJETIVOS

O objetivo geral deste trabalho foi desenvolver regras de relação entre estrutura e atividade (SAR) e gerar modelos robustos e preditivos de QSAR-2D, utilizando um conjunto de triazolopirimidinas inibidores da *PfDHODH*, buscando-se elucidar os requisitos estruturais relevantes à atividade biológica desta classe de compostos, com vistas ao desenvolvimento de novos agentes antimaláricos.

Objetivos específicos:

1. Realizar estudos de relação entre estrutura e atividade (SAR) para um conjunto de 132 triazolopirimidinas, inibidores da *PfDHODH*;
2. Desenvolver modelos robustos e preditivos de QSAR-2D e HQSAR para um conjunto de 132 triazolopirimidinas inibidores da *PfDHODH*, a fim de se elucidar dos requisitos estruturais importantes à atividade biológica da classe investigada;
3. Analisar os mapas de contribuição 2D e utilizá-los como guias em futuros trabalhos no planejamento de novos antimaláricos.

3. MATERIAL E MÉTODOS

3.1. Material

3.1.1. Computadores e *softwares*

Todos os estudos foram realizados no Laboratório de Planejamento de Fármacos e Modelagem Molecular (LabMol), em estações computacionais operando em Linux. Os *softwares* utilizados estão disponíveis através de licenças adquiridas pelo LabMol, acadêmicas e/ou colaborativas.

3.1.2. Conjunto de dados

Para os estudos de modelagem, selecionou-se um conjunto de dados de triazolopirimidinas da literatura [(GUJJAR *et al.*, 2011), (GUJJAR *et al.*, 2009), (COTERON *et al.*, 2011) e (MARWAHA *et al.*, 2012b)] com valores de concentração inibitória de 50% (IC₅₀) que foram obtidos sobre as mesmas condições experimentais. Os valores de IC₅₀ dos compostos foram transformados para a escala logarítmica pIC₅₀ (-log IC₅₀). A faixa de atividade do conjunto de dados abrangeu três unidades logarítmicas, com valores variando de 4,24 a 7,72.

As estruturas químicas dos compostos do conjunto de dados e seus respectivos valores de atividade biológica, expressa em IC₅₀, estão representados na Tabela 4.

Tabela 4 – Estruturas químicas e respectivos valores de IC₅₀ e pIC₅₀ para 132 triazolopirimidinas inibidores da *Pf*DHODH.


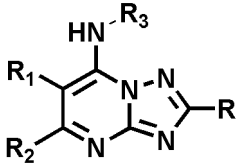
Composto						IC ₅₀ (μM) ^a	pIC ₅₀
	R	Ar	R ₁	R ₂	R ₃		
1	H	4-Br-Ph	H	CH ₃	-	0,78±0,25	6,11
2	H	4-CH ₃ -Ph	H	CH ₃	-	4,2±0,35	5,38
3	H	4-OCH ₃ -Ph	H	CH ₃	-	4,6±0,14	5,34
4	H	4-OCF ₃ -Ph	H	CH ₃	-	1,2±0,21	5,92
5	H	4-OCHF ₂ -Ph	H	CH ₃	-	2,8±0,28	5,55
6	H	4-NO ₂ -Ph	H	CH ₃	-	4,3±0,14	5,37
7	H	4- C ₆ H ₅ -Ph	H	CH ₃	-	5,3±0,8	5,28
8	H	4- C ₆ H ₅ CH ₂ - Ph	H	CH ₃	-	2,2±0,32	5,66
9	H	2,3-di-F-Ph	H	CH₃	-	39±7,5	4,41

Tabela 4 – Continuação

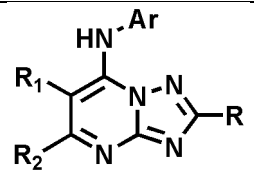
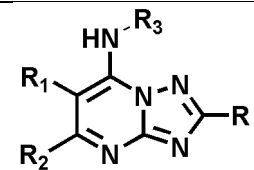
Composto							IC ₅₀ (μM) ^a	pIC ₅₀
	R	Ar	R ₁	R ₂	R ₃			
10	H	2,3,4-tri-F-Ph	H	CH ₃	-	29±4,4	4,54	
11	H	2,3,5-tri-F-Ph	H	CH₃	-	17±3,6	4,77	
12	H	2-F-4-CH ₃ -Ph	H	CH ₃	-	27±2,8	4,57	
13	H	3,5-di-F-Ph	H	CH₃	-	9,4±1,2	5,03	
14	H	3-CH ₃ -4-F-Ph	H	CH ₃	-	4,6 [*]	5,34	
15	H	3-CF₃-4-Cl-Ph	H	CH₃	-	0,8±0,28	6,10	
16	H	3-CF ₃ -4-Br-Ph	H	CH ₃	-	0,45±0,07	6,35	
17	H	3-CF ₃ -4-CH ₃ -Ph	H	CH ₃	-	1,55±0,49	5,81	
18	H	3,4-di-CH ₃ -Ph	H	CH ₃	-	0,35±0,21	6,46	
19	H	3-CF ₃ -4-CN-Ph	H	CH ₃	-	4,9±0,7	5,31	
20	H	3-F-4-CH₃-Ph	H	CH₃	-	0,86±0,09	6,07	
21	H	3-F-4-CF ₃ -Ph	H	CH ₃	-	0,077±0,03	7,11	
22	H	4-CF ₃ -Ph	H	CH ₃	-	0,28±0,02	6,55	
23	H	4-SF ₅ -Ph	H	CH ₃	-	0,13 [*]	6,89	
24	H	4-Cl-Ph	H	CH₃	-	1,6±0,24	5,8	
25	H	3-Cl-Ph	H	CH₃	-	1,4±0,13	5,85	
26	H	2-naftil	H	CH ₃	-	0,047±0,02	7,33	
27	H	2-antracênil	H	CH ₃	-	0,056±0,02	7,25	
28	H	3-F-Ph	H	CH₃	-	9,2±0,2	5,04	
29	H	3-CH ₃ -Ph	H	CH ₃	-	4,9±0,52	5,31	
30	H	3-CF ₃ -Ph	H	CH ₃	-	14 [*]	4,85	
31	H	4-F-Ph	H	CH ₃	-	19±2,0	4,72	
32	H	4-etil-Ph	H	CH ₃	-	0,61 [*]	6,21	
33	H	4- <i>n</i> -propil-Ph	H	CH ₃	-	0,99 [*]	6,0	
34	H	4- <i>n</i> -butil-Ph	H	CH ₃	-	0,28 [*]	6,55	
35	H	4- <i>n</i> -pentil-Ph	H	CH ₃	-	0,82 [*]	6,09	
36	H	4- <i>n</i> -hexil-Ph	H	CH ₃	-	4,2 [*]	5,38	
37	H	4- <i>n</i> -heptil-Ph	H	CH ₃	-	24 [*]	4,62	
38	H	4- <i>iso</i> -propil-Ph	H	CH ₃	-	0,22 [*]	6,66	

Tabela 4 – Continuação

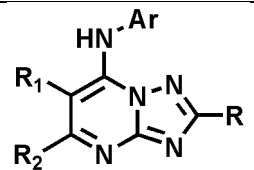
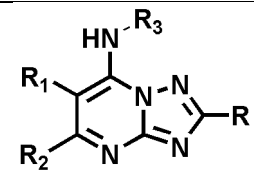
							
Composto	R	Ar	R ₁	R ₂	R ₃	IC ₅₀ (μM) ^a	pIC ₅₀
39	H	4- <i>terc</i> -butil-Ph	H	CH ₃	-	0,078*	7,11
40	H	4- <i>ciclo</i> -propil-Ph	H	CH ₃	-	0,42*	6,38
41	H	4-vinil-Ph	H	CH₃	-	0,80*	6,10
42	H	4-etinil-Ph	H	CH ₃	-	1,1*	5,96
43	H	4-CH₂CF₃-Ph	H	CH₃	-	2,5*	5,60
44	H	4-S-CF ₃ -Ph	H	CH ₃	-	0,61*	6,21
45	H	4-SF ₅ -Ph	H	CH ₃	-	0,13*	6,89
46	H	4-OCH ₂ C ₆ H ₅ -Ph	H	CH ₃	-	5,4*	5,27
47	H	4-N(CH ₃)-Ph	H	CH ₃	-	0,82*	6,09
48	H	4-pirrolidinil-Ph	H	CH ₃	-	3,5*	5,46
49	H	4-OC(CH ₃) ₃ -Ph	H	CH ₃	-	44*	4,36
50	H	4-C(CH ₃) ₂ CN-Ph	H	CH ₃	-	1,0*	6,0
51	H	3-F-4-C(CH ₃) ₃ -Ph	H	CH ₃	-	0,035*	7,46
52	H	3-F-4-Cl-Ph	H	CH ₃	-	0,55*	6,26
53	H	3,4-di-CF ₃ -Ph	H	CH ₃	-	0,54*	6,27
54	H	3,5-di-F-4-CF ₃ -Ph	H	CH ₃	-	0,19*	6,72
55	H	5,6,7,8-tetrahidro-2-naftil	H	CH ₃	-	0,064*	7,19
56	H	5-indanyl	H	CH ₃	-	0,22*	6,66
57	H	5-benzodioxoli	H	CH ₃	-	1,6*	5,80
58	H	5-benzotiofenil	H	CH ₃	-	0,09*	7,05
59	H	2-CF ₃ -5-benzoimidazolil	H	CH ₃	-	21*	4,68
60	H	4-CF ₃ -Ph	H	C ₂ H ₅	-	0,78*	6,11

Tabela 4 – Continuação

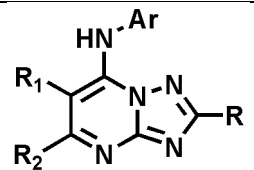
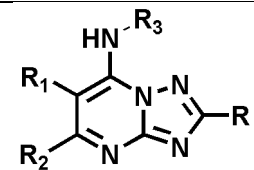
Composto							IC ₅₀ (μM) ^a	pIC ₅₀
	R	Ar	R ₁	R ₂	R ₃			
61	H	4-CF ₃ -Ph	H	CF ₃	-	1,7 [*]	5,77	
62	H	4-CF ₃ -Ph	H	Cl	-	0,17 [*]	6,77	
63	H	4-C(CH ₃) ₂ -Ph	C ₂ H ₅	CH ₃	-	1,7 [*]	5,77	
64	H	4-OCF ₃ -Ph	C ₂ H ₅	CH ₃	-	3,2 [*]	5,49	
65	H	3-F-4-CH₃-Ph	-C₃H₆-	-	-	1,4[*]	5,85	
66	iso-Pr	3-Cl-Ph	H	CH₃	-	0,12[*]	6,92	
67	<i>iso</i> -Pr	4-CF ₃ -Ph	H	CH ₃	-	0,11 [*]	6,96	
68	<i>iso</i> -Pr	4-SF ₅ -Ph	H	CH ₃	-	0,10 [*]	7	
69	<i>iso</i> -Bu	4-CF ₃ -Ph	H	CH ₃	-	2,0 [*]	5,7	
70	<i>tert</i> -Bu	4-CF ₃ -Ph	H	CH ₃	-	0,46 [*]	6,34	
71	<i>tert</i> -Bu	4-SF ₅ -Ph	H	CH ₃	-	0,55 [*]	6,26	
72	<i>sec</i> -Bu	4-CF ₃ -Ph	H	CH ₃	-	0,38 [*]	6,42	
73	<i>sec</i> -Bu	4-SF ₅ -Ph	H	CH ₃	-	0,33 [*]	6,48	
74	cicloPr	4-Cl-Ph	H	CH₃	-	0,88[*]	6,06	
75	cicloPr	3-Cl-Ph	H	CH ₃	-	0,82 [*]	6,09	
76	cicloPr	4-CF ₃ -Ph	H	CH ₃	-	0,14 [*]	6,85	
77	CH ₂ cicloPr	4-Cl-Ph	H	CH ₃	-	12 [*]	4,92	
78	CH ₂ cicloPr	4-CF ₃ -Ph	H	CH ₃	-	1,7 [*]	5,77	
79	CH ₂ cicloPr	4-SF ₅ -Ph	H	CH ₃	-	5,7 [*]	5,24	
80	CF₂CH₃	4-Cl-Ph	H	CH₃	-	0,038[*]	7,42	
81	CF ₂ CH ₃	3-Cl-Ph	H	CH ₃	-	0,054 [*]	7,27	
82	CF ₂ CH ₃	4-CF ₃ -Ph	H	CH ₃	-	0,038 [*]	7,42	
83	CF ₂ CH ₃	4-SF ₅ -Ph	H	CH ₃	-	0,033 [*]	7,48	
84	CF₂CH₂CH₃	4-Cl-Ph	H	CH₃	-	0,034[*]	7,47	
85	CF ₂ CH ₂ CH ₃	3-Cl-Ph	H	CH ₃	-	0,030 [*]	7,52	
86	CF ₂ CH ₂ CH ₃	4-CF ₃ -Ph	H	CH ₃	-	0,019 [*]	7,72	
87	CF ₂ CH ₂ CH ₃	4-SF ₅ -Ph	H	CH ₃	-	0,075 [*]	7,12	
88	CF₃	4-CF₃-Ph	H	CH₃	-	0,035[*]	7,46	
89	CF ₃	4-SF ₅ -Ph	H	CH ₃	-	0,030 [*]	7,52	
90	CH₂OCH₃	4-Cl-Ph	H	CH₃	-	2,2[*]	5,66	
91	CH₂OCH₃	3-Cl-Ph	H	CH₃	-	1,7[*]	5,77	
92	CH ₂ OCH ₃	4-CF ₃ -Ph	H	CH ₃	-	0,41 [*]	6,39	
93	CH ₂ OCH ₃	4-SF ₅ -Ph	H	CH ₃	-	0,39 [*]	6,41	
94	CH ₂ CH ₂ OCH ₃	4-CF ₃ -Ph	H	CH ₃	-	2,5 [*]	5,6	
95	CH ₂ CH ₂ OCH ₃	4-SF ₅ -Ph	H	CH ₃	-	2,5 [*]	5,6	
96	SCH ₃	4-Cl-Ph	H	CH ₃	-	0,030 [*]	7,52	
97	SCH₃	3-Cl-Ph	H	CH₃	-	0,026[*]	7,59	
98	SCH ₃	4-CF ₃ -Ph	H	CH ₃	-	0,036 [*]	7,44	
99	SCH ₃	4-SF ₅ -Ph	H	CH ₃	-	0,030 [*]	7,52	
100	SO ₂ CH ₃	4-Cl-Ph	H	CH ₃	-	0,045 [*]	7,35	

Tabela 4 – Continuação

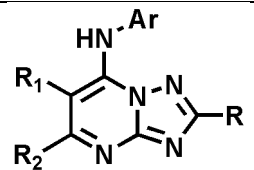
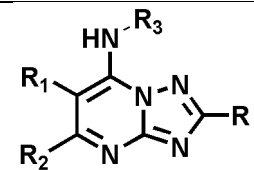
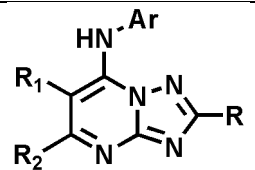
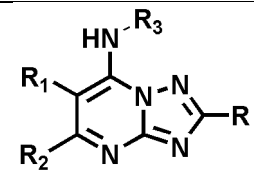
							
Composto	R	Ar	R ₁	R ₂	R ₃	IC ₅₀ (μM) ^a	pIC ₅₀
101	SO ₂ CH ₃	3-Cl-Ph	H	CH ₃	-	0,16 [*]	6,8
102	SO ₂ CH ₃	4-CF ₃ -Ph	H	CH ₃	-	0,090 [*]	7,05
103	SO ₂ CH ₃	4-SF ₅ -Ph	H	CH ₃	-	0,11 [*]	6,96
104	OCH₃	3-Cl-Ph	H	CH₃	-	0,50[*]	6,30
105	OCH ₃	4-Cl-Ph	H	CH ₃	-	0,46 [*]	6,34
106	OCH₂CH₃	3-Cl-Ph	H	CH₃	-	0,40[*]	6,40
107	OCH ₂ CH ₃	4-Cl-Ph	H	CH ₃	-	0,056 [*]	7,25
108	CH ₂ CH ₃	4-Cl-Ph	H	CH ₃	-	0,28 [*]	6,55
109	OCH₂CH₃	4-SF₅-Ph	H	CH₃	-	0,057[*]	7,24
110	OCH ₂ CH ₂ OC H ₃	3-Cl-Ph	H	CH ₃	-	4,9 [*]	5,31
111	CH₂CH₃	3-Cl-Ph	H	CH₃	-	0,24[*]	6,62
112	CH ₂ CH ₃	4-CF ₃ -Ph	H	CH ₃	-	0,087 [*]	7,06
113	CH₂OH	4-Cl-Ph	H	CH₃	-	5,3[*]	5,28
114	CH ₂ OH	3-Cl-Ph	H	CH ₃	-	6,1 [*]	5,21
115	CH₂CH₂OH	4-CF₃-Ph	H	CH₃	-	0,93[*]	6,03
116	H	-	H	CH ₃	<i>ciclo-</i> <i>dodecil</i>	58 [*]	4,24
117	H	-	H	CH ₃	2- <i>indanil</i>	0,5 [*]	6,3
118	H	-	H	CH ₃	5-F-2- <i>indanil</i>	0,61 [*]	6,21
119	H	-	H	CH ₃	5-Br-2- <i>indanil</i>	1,2 [*]	5,92
120	H	-	H	CH ₃	5,6-di- OCH ₃ - 2- <i>indanil</i>	4,6 [*]	5,34
121	H	-	H	CH ₃	1,2,3,4- tetrahid ro-2- nafti	0,54 [*]	6,27
122	H	-	H	CH ₃	6-F- 1,2,3,4- tetra- hidro2- nafti	0,63 [*]	6,20
123	H	-	H	CH ₃	6-Cl- 1,2,3,4- tetra- hidro2- nafti	0,67 [*]	6,17

Tabela 4 – Continuação

							
Composto	R	Ar	R ₁	R ₂	R ₃	IC ₅₀ (μM) ^a	pIC ₅₀
124	H	-	H	CH ₃	2-fenil-etil	15 [*]	4,82
125	H	-	H	CH ₃	3-fenil-propil	26 [*]	4,59
126	H	-	H	CH ₃	<i>n</i> -etil	8,6 [*]	5,07
127 ^{**}						4,8±0,38	5,32
128 ^{**}						27±6,3	4,57
129 ^{**}						0,077±0,08	7,11
130 ^{**}						4,8±0,7	5,32
131 ^{**}						0,4±0,09	6,4
132 ^{**}						0,44±0,17	6,36

^aerro dos resultados experimentais, ^{*} sem dados referentes ao erro experimental, ^{**} modificação bioisostérica do *scaffold* triazolopirimidínico. Em negrito, as estruturas do conjunto teste.

3.2. Métodos

3.2.1. Padronização do conjunto de dados e preparo das estruturas

O fluxo de trabalho referente a modelagem de QSAR desenvolvido neste trabalho foi realizado em três etapas principais: (i) padronização do conjunto de dados, preparo e análise (seleção de compostos e descritores) e (ii) construção de modelos, e (iii) validação e seleção dos modelos.

As estruturas dos compostos foram desenhadas no programa MarvinSketch versão 5.12 (ChemAxon), constituído por estruturas em 1D (SMILES) e 2D (bidimensional). Os compostos foram preparados usando o protocolo desenvolvido por Fourches e colaboradores (2010), no qual contra-íons foram removidos e quimiotipos específicos como grupamentos nitro e aromáticos foram normalizados usando o programa Standardizer (v.5.3, ChemAxon, Budapest, Hungry, www.chemaxon.com). A presença de duplicatas, isto é, compostos idênticos reportados mais de uma vez no conjunto de dados, foi identificada usando o ISIDA Duplicates (VARNEK *et al.*, 2008) e HiTQSAR (KUZ'MIN *et al.*, 2008). Quando as propriedades biológicas associadas a duas estruturas repetidas eram iguais, um dos compostos foi removido. Contudo, quando as propriedades biológicas eram significativamente diferentes, ambos foram removidos. Ao final deste processo de preparo de estruturas, obteve-se um conjunto com 132 triazolopirimidinas devidamente padronizadas para os estudos de modelagem molecular.

3.2.2. Estudo da relação entre estrutura e atividade (SAR)

3.2.2.1. Análise de agrupamentos hierárquicos

A HCA (do inglês, *hierarchical clustering analysis*) é um método de aprendizado não supervisionado. Nesse estudo empregou-se o método *Sequential Agglomerative Hierarchical Non-overlapping* implementado no software ISIDA Cluster (<http://infochim.u-strasbg.fr>) (VARNEK *et al.*, 2008). Nesse método, cada composto representa um grupo (do inglês, *cluster*) no início. Então, os n compostos são agrupados iterativamente em grupos usando suas distâncias euclidianas armazenadas em uma matriz quadrada de distâncias simétricas ($n * n$). Os dois objetos mais próximos (compostos ou grupos) são iterativamente identificados e agrupados para formar um novo grupo. A matriz de distâncias é atualizada com as novas distâncias separando o

novo grupo dos outros, de acordo com o tipo de acoplamento selecionado pelo usuário (nesse estudo foi usado o acoplamento completo). O processo se repete até que reste apenas um grupo. A relação sequencial de formação dos grupos resulta em um gráfico hierárquico chamado de dendograma. Esse software gera um dendograma e um mapa de calor da matriz de distância colorida de acordo com a similaridade/dissimilaridade entre os compostos.

3.2.2.2. *Análise de cliffs de atividade*

O fluxo de trabalho referente a modelagem de SAR desenvolvido neste trabalho foi realizado em três etapas principais: (i) preparo do conjunto de dados (estruturas 2D e CDK) (ii) geração de descritores *fingerprints*, (iii) cálculo da similaridade entre os compostos através do coeficiente de Tanimoto. Todo esse fluxo de trabalho foi desenvolvido na plataforma KNIME, através de um *workflow* desenvolvido pelo grupo de pesquisa do LabMol.

3.2.3. Divisão dos conjuntos treinamento e teste

O conjunto de dados foi dividido racionalmente em um conjunto de treinamento (80%) e um conjunto teste (20%) utilizando o algoritmo *Kennard-Stone* modificado, implementado no pacote *ksaR* v.0.7 para R, disponível na página web de nosso grupo de pesquisa (<http://labmol.farmacia.ufg.br/ksar>). Essa divisão foi utilizada para a construção dos modelos de HQSAR e QSAR-2D.

3.2.4. HQSAR

O módulo de HQSAR da plataforma SYBYL-X v.1.2 (Tripos, Inc., St. Louis, MO, USA) foi empregado na geração dos modelos de HQSAR para o conjunto de inibidores da enzima *PfDHODH*. As análises de HQSAR foram realizadas através de uma abordagem de triagem-erro, buscando-se a combinação que forneça a melhor consistência interna. Os modelos foram construídos com diferentes combinações dos parâmetros de distinção de fragmentos: átomos (A), ligações (B), conectividade (C), presença de átomos de hidrogênio (H), quiralidade (Ch) e doadores e aceptores de ligação de hidrogênio (DA). Tamanhos de comprimento de holograma disponíveis no programa (53, 59, 61, 71, 83, 97, 151, 199, 2557, 307, 353, 401), além de comprimentos adicionais (499, 673, 809 e 997), para evitar a colisão de fragmentos nos hologramas

moleculares. Variações do número máximo e mínimo de átomos nos fragmentos foram avaliados (2-5, 3-6, 4-7, 5-8, 6-9 e 7-10).

Em seguida, os padrões de frequência dos fragmentos nos hologramas gerados (variável independente) foram correlacionados às variáveis dependentes utilizando o método de regressão por mínimos quadrados parciais (*PLS*).

A validação interna dos modelos bem como a determinação do número ideal de componentes *PLS* foi realizada pela técnica de validação cruzada *leave-one-out* (*LOO*). Nesta técnica, cada composto é retirado uma vez e o modelo de HQSAR é reconstruído, ao final é feito uma média dos valores obtidos, gerando então o modelo final de HQSAR.

Para a determinação dos *outliers*, foi calculado o valor residual, que é a diferença entre o valor de atividade biológica experimental e a predita. Calculou-se então o desvio padrão médio dos valores de residuais, e compostos que excederam o valor de duas vezes o valor de desvio padrão, ou seja 95%, intervalo de confiança (IC), foram retirados. Após a retirada dos *outliers*, foram realizadas novas análises de *PLS* para a construção dos modelos finais de HQSAR.

O método de HQSAR também fornece informações visuais de quais fragmentos dos compostos são favoráveis ou desfavoráveis para atividade biológica, através de um mapa de cores. As cores na região do vermelho do espectro (vermelho, laranja avermelhado e laranja) refletem contribuições desfavoráveis ou negativas. Cores na região do verde (amarelo, azul esverdeado e verde) refletem contribuições favoráveis ou positivas. Átomos com contribuições intermediárias são representados em branco (HQSARTM MANUAL, 2010).

3.2.5. QSAR-2D

Os modelos de QSAR-2D foram construídos na plataforma KNIME, seguindo um *workflow* (Figura 17) desenvolvido em nosso grupo de pesquisa. Foram utilizados diferentes métodos de regressão para construção dos modelos de QSAR-2D, sendo eles *Partial Least Square* (PLS), *Random Forest* (RF), *Support Vector Machines* (SVM), *Gradient Boosting Machine* (GBM), GAMBoost e *Kernel-based Regularized Least Squares* (KRLS). Foram calculados descritores CDK-2D no software PaDEL (YAP, 2011) e utilizados como variáveis independentes na geração dos modelos de QSAR-2D.

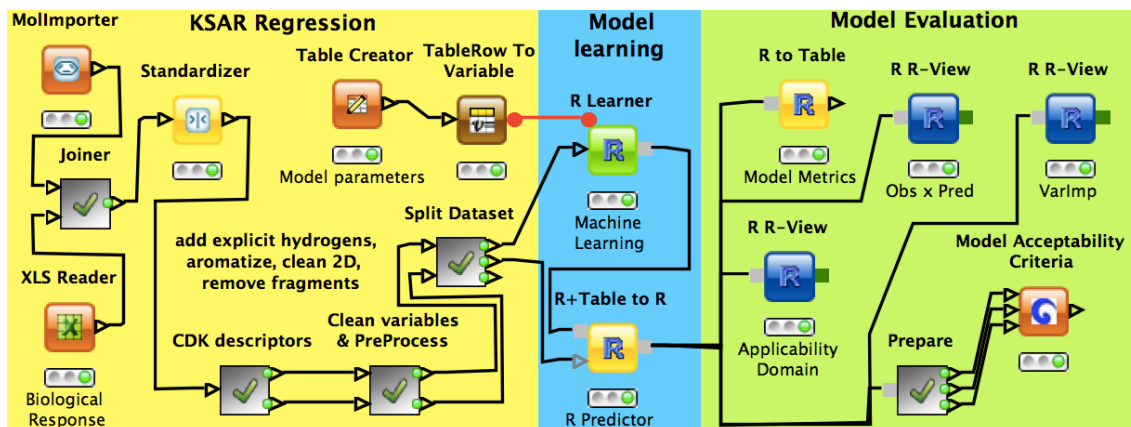


Figura 17 – *Workflow* desenvolvido na plataforma KNIME para geração dos modelos de QSAR – 2D

3.2.6. Validação externa dos modelos de HQSAR e QSAR-2D

A validação externa de todos os modelos de QSAR foi realizada empregando-se o conjunto teste contendo 26 compostos, a fim de se verificar a capacidade preditiva dos melhores modelos obtidos. A capacidade preditiva dos modelos gerados foi avaliada utilizando do coeficiente de correlação preditivo (r_{pred}^2); coeficiente de correlação modificado (r_m^2); diferença entre r_m^2 e $r'_m{}^2$ (Δr_m^2); coeficiente de correlação de concordância (CCC); teste de *Durbin Watson*, correlação entre os resíduos (d); método *Golbraikh e Tropsha* (G&T).

3.2.6.1. Coeficiente de correlação preditivo (r_{pred}^2)

O coeficiente de correlação preditivo (r_{pred}^2) foi calculado conforme a equação (4):

$$r_{pred}^2 = 1 - \frac{\sum_{i=1}^{N_{teste}} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{N_{teste}} (y_i - \bar{y}_{teste})^2} \quad (4)$$

onde N representa o número de compostos; \hat{y}_i é o valor de atividade biológica predita do conjunto teste; y_i é o valor de atividade biológica experimental do conjunto teste; e \bar{y}_{teste} é a media das atividades biológicas dos compostos do conjunto teste (CHIRICO, GRAMATICA, 2012; CONSONNI, BALLABIO, TODESCHINI, 2010; CONSONNI, BALLABIO, TODESCHINI, 2009).

3.2.6.2. Coeficiente de correlação modificado (r_m^2)

O r_m^2 é um parâmetro de validação externa adicional que considera a diferença real entre os valores preditos e experimentais (MITRA *et al.*, 2010; ROY, MITRA, 2012; ROY *et al.*, 2013; ROY *et al.*, 2012a; ROY *et al.*, 2012b). O cálculo de r_m^2 é descrito pela equação (5):

$$rm^2 = r^2 (1p\sqrt{r^2 - r_0^2}) \quad (5)$$

onde r^2 representa o coeficiente de correlação entre os valores observados (eixo-y) e preditos (eixo-x) com intersecção em zero no eixo das ordenadas e abscissas, o r_0^2 representa o coeficiente de correlação entre os valores observados (eixo-y) e preditos (eixo-x) sem intersecção em zero no eixo das ordenadas e abscissas.

O parâmetro r_m^2 possui algumas derivações, a equação (6) descreve a inversão dos eixos no plano cartesiano para a obtenção do valor inverso de r_m^2 (r'_m^2):

$$r' \frac{2}{m} = r^2 (1\sqrt{r^2 - r' \frac{2}{0}}) \quad (6)$$

onde r^2 representa o coeficiente de correlação entre os valores observados (eixo-y) e preditos (eixo-x) com intersecção em zero no eixo das ordenadas e abscissas, o r'_0^2 representa o coeficiente de correlação pela inversão entre os valores observados (eixo-x) e preditos (eixo-y) sem intersecção em zero no eixo das ordenadas e abscissas. A Figura 18 apresenta os coeficientes de correlação com e sem intersecção em zero, obtidos a partir da inversão dos eixos x e y .

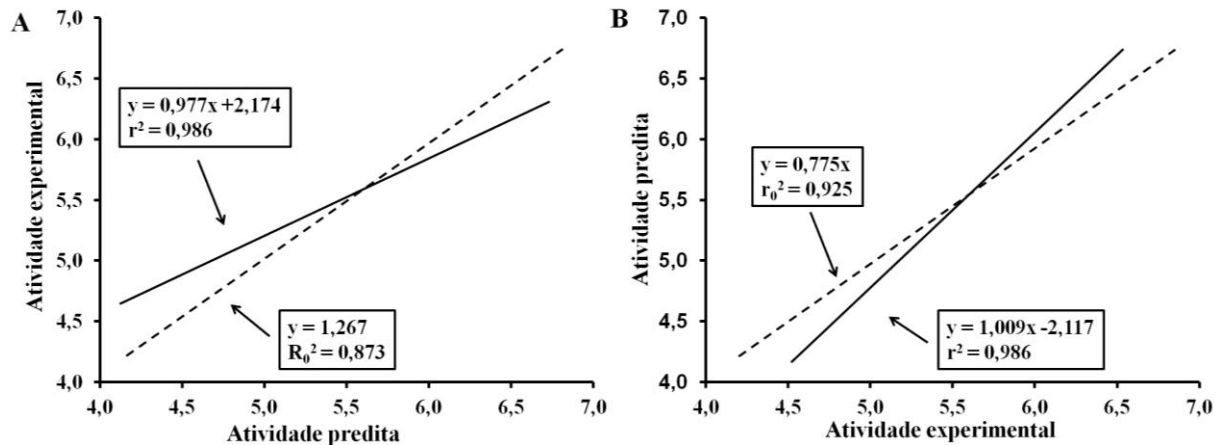


Figura 18 – Gráficos de regressão com e sem intercepção em zero: dados preditos estão plotados do eixo X e experimentais no eixo Y (A); com inversão dos eixos (B) (ROY *et al.*, 2013).

A correlação entre os valores experimentais (eixo- x) e preditos (eixo- y) é a mesma entre os valores experimentais (eixo- y) e preditos (eixo- x), entretanto isso não ocorre quando a intersecção é definida como zero. Portanto, o valor de r_m^2 será diferente do r'^2_m e a diferença de (r_m^2 delta) poderá ser utilizada como medida de previsão (ROY *et al.*, 2012), de acordo com o demonstrado na equação (7):

$$r_m^2 \text{ delta} = \text{abs}(r_m^2 - r'^2_m) \quad (7)$$

Além disso, como um dos valores de r_m^2 e r'^2_m pode penalizar fortemente a qualidade do modelo em termos de previsão, a média destas duas variáveis, representada pela equação (8) foi calculada.

$$r_m^2 \text{ médio} = \frac{(r_m^2 + r'^2_m)}{2} \quad (8)$$

3.2.6.3. Coeficiente de correlação de concordância (CCC)

O coeficiente de correlação de concordância (CCC) verifica o quão pequenas são as diferenças entre dados experimentais e dados externos previsões ajustados, independentemente da sua escala, conforme a equação 9. (CHIRICO; GRAMATICA, 2012)

$$CCC = \frac{2 \sum_{i=1}^{n_{EXT}} (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sum_{i=1}^{n_{EXT}} (y_i - \bar{y})^2 + \sum_{i=1}^{n_{EXT}} (\hat{y}_i - \bar{\hat{y}})^2 + n_{EXT} (\bar{y} - \bar{\hat{y}})^2} \quad (9)$$

N representa o número de compostos; \hat{y}_i é o valor de atividade biológica predita do conjunto teste; y_i é o valor de atividade biológica experimental do conjunto teste; \bar{y}_{teste} é a média das atividades biológicas dos compostos do conjunto teste; e $\bar{\hat{y}}$ média dos valores de predição.

3.2.6.4. Método de Golbraikh e Tropsha (G&T)

Golbraikh e Tropsha estabeleceram após uma cuidadosa análise que a validação externa, utilizando um conjunto que não participa da construção dos modelos é a única maneira de estabelecer um modelo de QSAR confiável. Portanto, para modelos de QSAR confiáveis o valor do coeficiente de correlação preditivo deve maior ou igual 0,7 (GOLBRAIKH; TROPSHA, 2002).

3.2.6.5. Teste de Durbin Watson de correlação entre os resíduos (d)

O teste de Durbin-Watson é utilizado para detectar a presença de auto correlação (dependência) nos resíduos de uma análise de regressão. Este teste é baseado na suposição de que os erros no modelo de regressão são gerados por um processo auto regressivo de primeira ordem (DURBIN; WATSON, 1950), conforme equações 10 e 11.

$$r_L = \frac{\sum_{j=1}^{n-k} \lambda_j w_j^2}{\sum_{j=1}^{n-k} w_j^2} \quad (10)$$

$$r_u = \frac{\sum_{j=k+1}^n \lambda_j w_j^2}{\sum_{j=1}^{n-k} w_j^2} \quad (11)$$

Então, $r_L \leq d \leq r_u$.

4. RESULTADOS E DISCUSSÃO

4.1. Caracterização do conjunto de dados

Os estudos de SAR e QSAR foram desenvolvidos para um conjunto de 132 triazolopirimidinas da literatura (GUJJAR *et al.*, 2011; GUJJAR *et al.*, 2009; COTERON *et al.*, 2011; MARWAHA *et al.*, 2012), com valores de concentração inibitória de 50% (IC₅₀) que foram obtidos sobre as mesmas condições experimentais, um requisito essencial para os estudos de QSAR e obtenção de modelos confiáveis (OECD, 2007). Os valores de IC₅₀ variam de 58µM a 0,019µM, uma faixa de variação considerada grande. Esses valores foram transformados para escala logarítmica, cuja a variação passou a ser de 3 unidades logarítmicas, variação normalizada, necessária para os estudos de QSAR.

4.2. Estudos de relação entre estrutura e atividade (SAR)

4.2.1. Análise de agrupamentos hierárquicos

A análise de agrupamentos hierárquicos dividiu os compostos em 2 grandes grupos (*clusters*), *cluster a* e *cluster b* (Figura 19). As estruturas de cada *cluster* são semelhantes, como observado no mapa de calor. No mapa de calor, as cores quentes, tons de vermelho e alaranjado, representam dissimilaridade entre os compostos, e cores frias, tons de azul e violeta, similaridade. Em cada *cluster*, a e b, as cores estão em tons de azul, logo compostos estruturalmente semelhantes. O mapa de calor, representando a matriz de distâncias das estruturas, e o dendograma do conjunto de dados são apresentados na Figura 19.

No *cluster a* estão os compostos com substituintes hidrofóbicos na posição *para* do anel benzênico (posição 1) menos volumosos, além substituintes no anel triazolpirimidínico (posição 5) (Figura 20). No *cluster b*, estão compostos com substituintes hidrofóbicos, volumosos (posição 1) e sem substituintes na posição 5 do anel triazolopirimidínico. Apesar dos compostos apresentar o mesmo *scaffold* (esqueleto) triazolopirimidínico, o conjunto é estruturalmente diverso. A distância Euclidiana entre os dois *clusters* varia de 193,48 a 1160,90, indicando a variedade dos compostos.

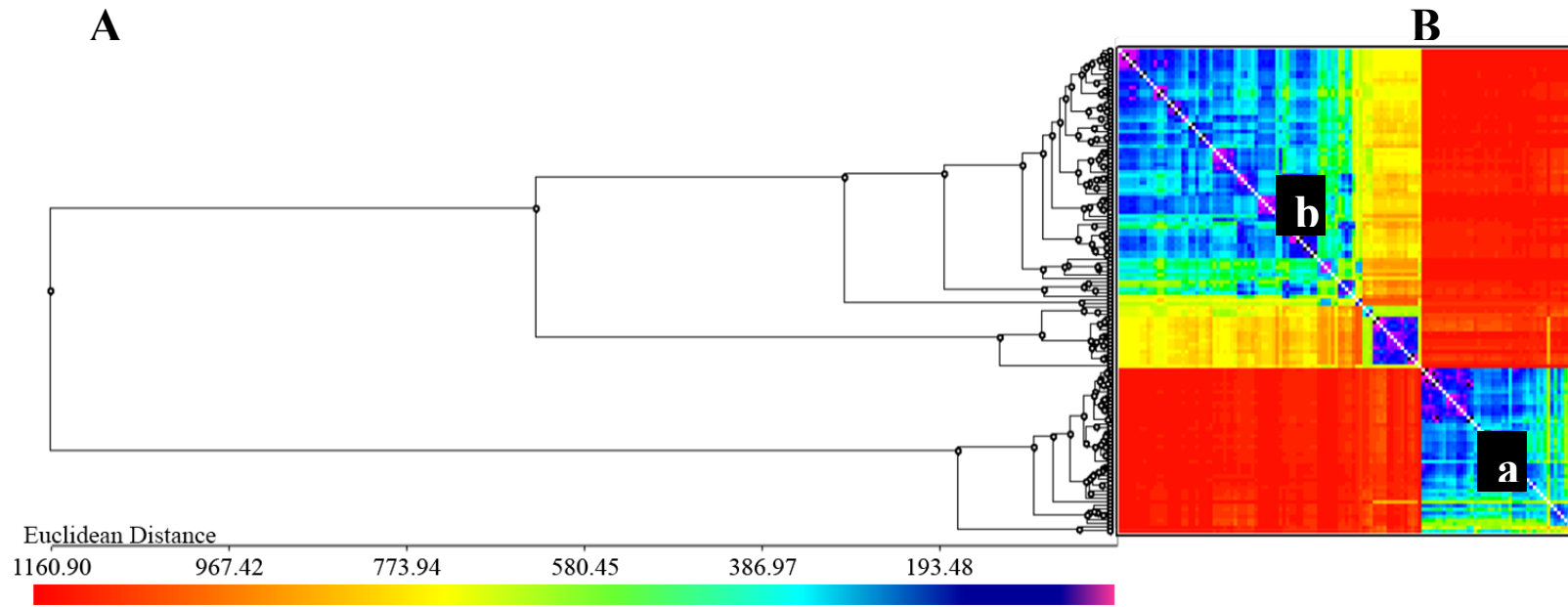


Figura 19 - Análise agrupamentos hierárquicos para o conjunto de dados: dendograma (A) e mapa de calor da matriz de distâncias organizadas com base na similaridade estrutural (B) (violeta/azul = similar; amarelo, vermelho = dissimilar).

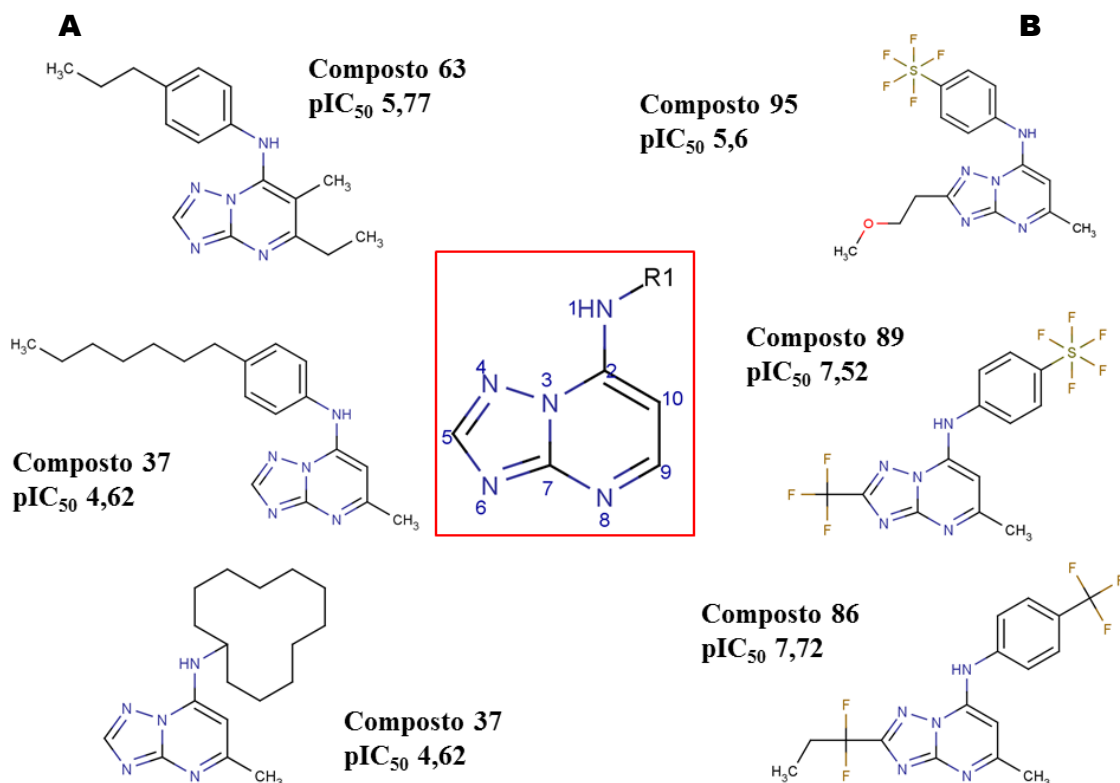


Figura 20 – Representação estrutural de alguns compostos do *cluster* a (A) e compostos do *cluster* b (B), ao meio o *scaffold* triazolopirimidínico, presente em todos os compostos.

4.2.2. Análise de *cliffs* de atividade

O estudo das características estruturais que influenciam na atividade biológica dos compostos permite estabelecer relações e assim propor novas estruturas que possam ser ainda mais potentes (WASSERMANN; WAWER; BAJORATH, 2010). Pequenas mudanças na estrutura dos compostos que geram grandes variações no valor da atividade biológica é denominada com *cliff* de atividade (WASSERMANN; WAWER; BAJORATH, 2010). Os *cliffs* de atividade explicam como alguns compostos estruturalmente semelhantes apresentam valores de atividade bastante diferentes. A Figura 20 apresenta a representação 3D da paisagem descontínua de atividade do conjunto de dados, calculada baseada na distância euclidiana. O plano *x/y* representa a projeção 2D das estruturas no espaço químico pelos descritores e o eixo *z* representa a potência dos compostos. Os picos e vales na mesma região (regiões de *cliffs*) correspondem a relação estrutura-atividade heterogênea dos compostos, ou seja, moléculas similares com potência alta e baixa. Região no mesmo espaço, com picos em

vermelho/amarelo representam os compostos com maior potência, enquanto que região com vales em verde representam os compostos com mais baixa potência.

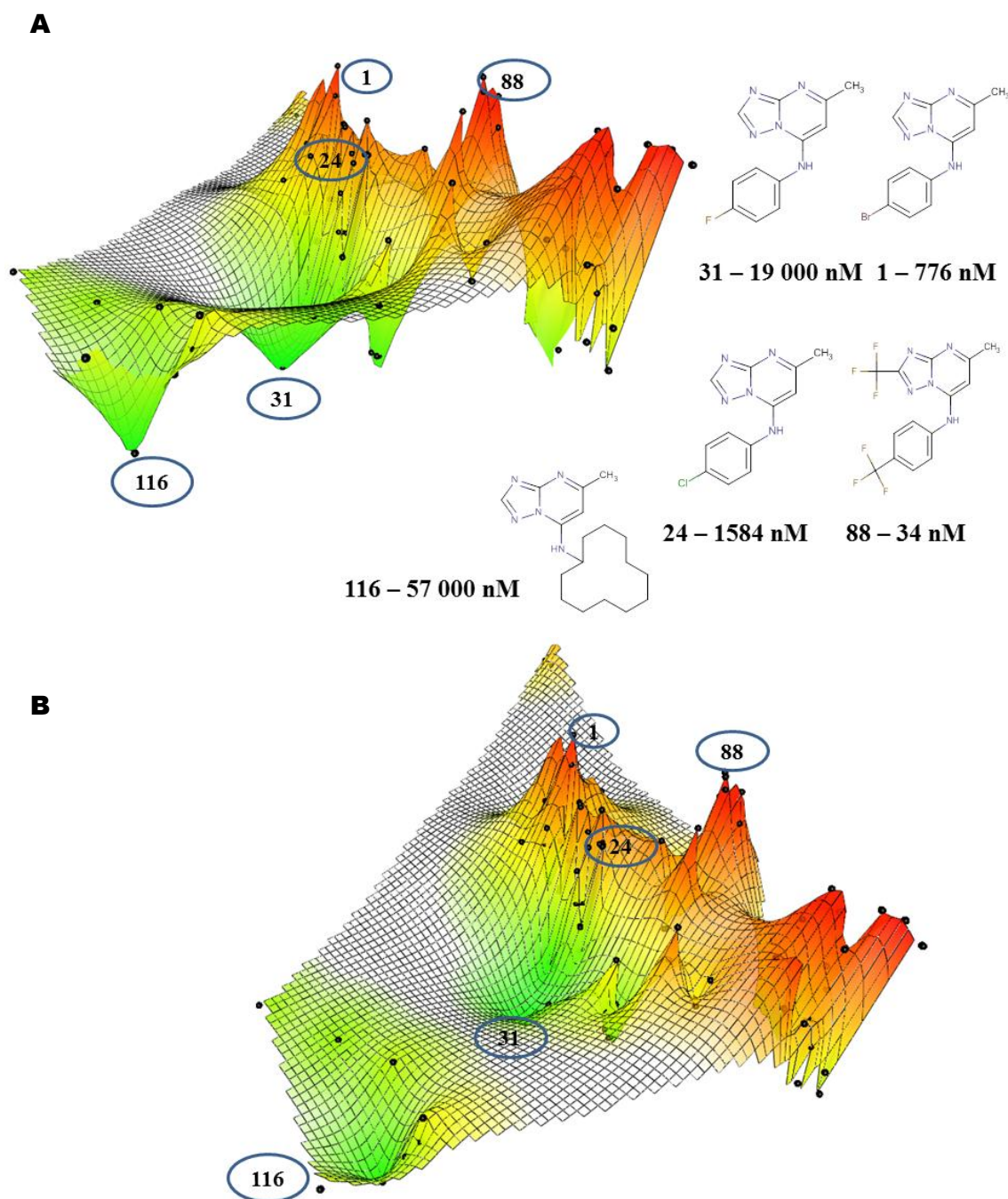


Figura 21 – As figuras A e B são representações 3D da paisagem descontínua de atividade (*activity landscape*), calculada baseada na distância euclidiana. O plano x/y representa a projeção 2D das estruturas no espaço químico pelos descritores e o eixo z representa a potência dos compostos. Os picos e vales na mesma região (regiões de *cliffs*) correspondem a relação estrutura-atividade heterogênea dos compostos, ou seja, moléculas similares com potência alta e baixa.

O *landscape*, desse conjunto, (Figura 21) mostra que compostos estruturalmente similares, estão no mesmo espaço químico, apresentam variações de potência, picos(maior potência) e vales (baixa potência), no mesmo espaço. Compostos com essa característica é denominado *cliffs* de atividade. Portanto, este é conjunto descontínuo.

Na Tabela 5 estão apresentados os compostos do conjunto que são considerados *cliffs*, dado obtido a partir do índice de relação estrutura-atividade SALI, do inglês, *Structure-Activity Landscape Index*. SALI, identifica os *cliffs*, pontos de inflexão importantes do *landscape* no mesmo espaço químico. O valor de SALI é calculado pela razão da diferença da potência de dois compostos pela diferença da similaridade entre eles. (BAJORATH *et al.*, 2009) (Equação 12).

$$SALI_{i,j} = \frac{|A_i - A_j|}{1 - sim(i,j)} \quad (12)$$

Onde, A_i e A_j , é o valor de potência, *sim* a similaridade entre as estruturas *i* e *j* entre elas. Valores alto de SALI, são compostos com alta semelhança estrutural, mas grandes diferenças de potência, portanto *cliff* (BAJORATH *et al.*, 2009).

Tabela 5 - Estruturas que representam *cliffs* de atividade.

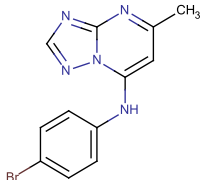
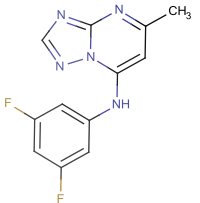
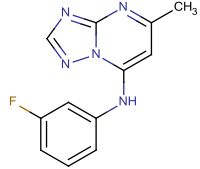
Cpd	Estrutura	IC ₅₀ (μM)	pIC ₅₀	Distância	SALI
1		0,78±0,25	6,11*	-	-
13		9,4± 1,2	5,03	0,051	21,060
28		9,2± 0,2	5,04	0,051	20,865

Tabela 5 – Continuação

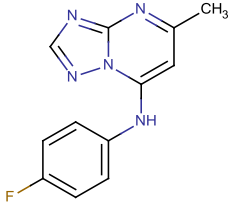
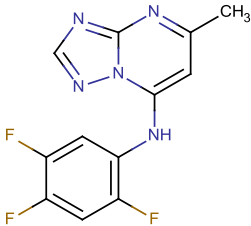
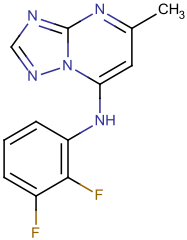
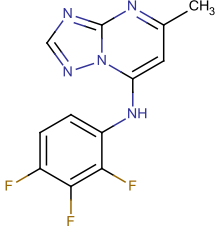
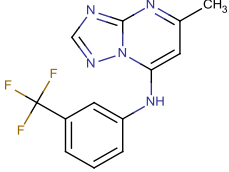
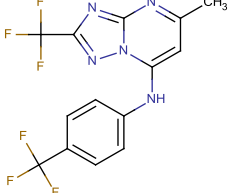
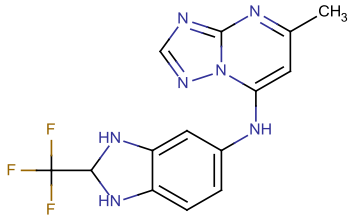
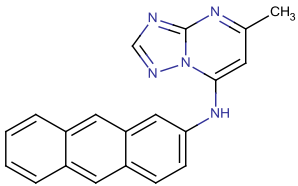
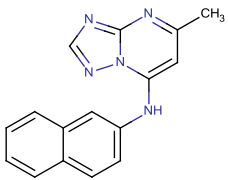
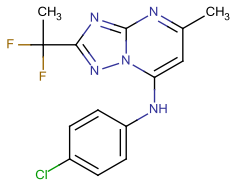
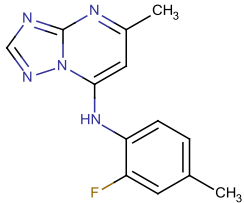
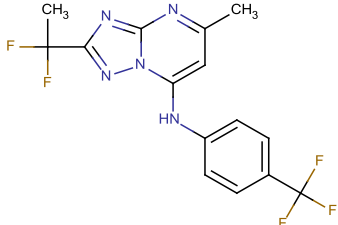
Cpd	Estrutura	IC ₅₀ (μM)	pIC ₅₀	Distância	SALI
31		19± 2,0	4,72	0,051	27,105
11		17± 3,6	4,77	0,098	13,735
9		39± 7,5	4,41	0,098	17,425
10		29± 4,4	4,54	0,098	16,093
30		14	4,85	0,100	12,600
88		0,035	7,46	0,100	13,500

Tabela 5 – Continuação

Cpd	Estrutura	IC ₅₀ (μM)	pIC ₅₀	Distância	SALI
59		21	4,68	0,100	14,300
27		0,056±0,02	7,25	0,105	10,830
26		0,047±0,02	7,33	0,105	11,590
80		0,038	7,42	0,119	11,004
12		27± 2,8	4,57	0,119	12,936
82		0,038	7,42	0,122	10,742

*Estrutura utilizada para comparação com as demais na análise de *cliffs*.

Os compostos que são considerados *cliffs* de atividade representam 12% do conjunto total. Analisando os compostos 1 e 31, na tabela 4, com substituintes bromo (Br) e flúor (F) na posição *para* do anel benzeno, o F e o Br fazem parte da família dos halogênios, e portanto, têm algumas características semelhantes do ponto de vista químico, logo se esperaria uma atividade biológica semelhante, para compostos com esses substituintes. No entanto, o composto 31, apresenta uma significativa diminuição da atividade, IC_{50} 19 μ M, com relação ao composto 1, 0,78 μ M. Vale ressaltar que o F é bem menos volumoso que o Br (raio atômico do F, 0,50Å, raio atômico do Br, 1,15) (MAHAN; MYERS, 1995), então o volume ocupado por esta região da molécula na cavidade de interação com a enzima é importante para a potência. De acordo com os estudos de raios-X da *PfDHODH*, (PDB ID: 3I6R, 3I65, 3I68 e 3SFK,) (COTERON *et al.*, 2011; DENG *et al.*, 2009) (Figura 22), compostos com substituintes hidrofóbicos no anel benzênico realizam interações com o bolso hidrofóbico da *PfDHODH*, o explica a maior potência de compostos com maior volume.

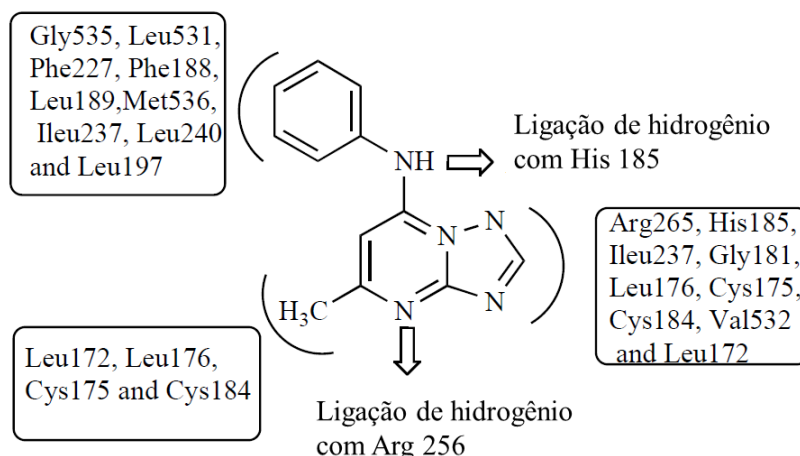


Figura 22 – Interações mais importantes feitas com a *PfDHODH* (VYAS; GHATE, 2011)

Os compostos 80, 82 e 88 também foram *cliffs*, são mais potentes que o composto 1, IC_{50} , 0,038 μ M, 0,038 μ M e 0,035 μ M, respectivamente. Estes compostos possuem um substituinte hidrofóbico trifluormetil, compostos 82 e 88, e cloro (Cl) na posição *para* do anel benzênico, requisito que favorece a potência, de acordo com o cristal, também possui substituinte hidrofóbico, haleto de alquila, composto 80 e 82, e trifluormetil, no anel triazolopirimidínico. Outras informações das estruturas de raio – X da *PfDHODH* (Figura 23) indica outra região de interação hidrofóbica da enzima com

estes substituintes dos compostos, favorecendo a potência. Logo, substituinte hidrofóbico nesse anel também favorece a potência.

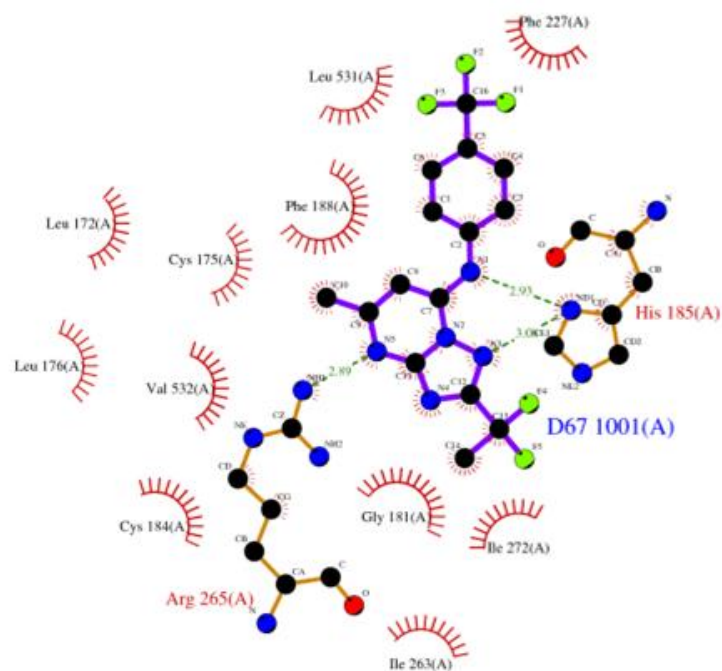
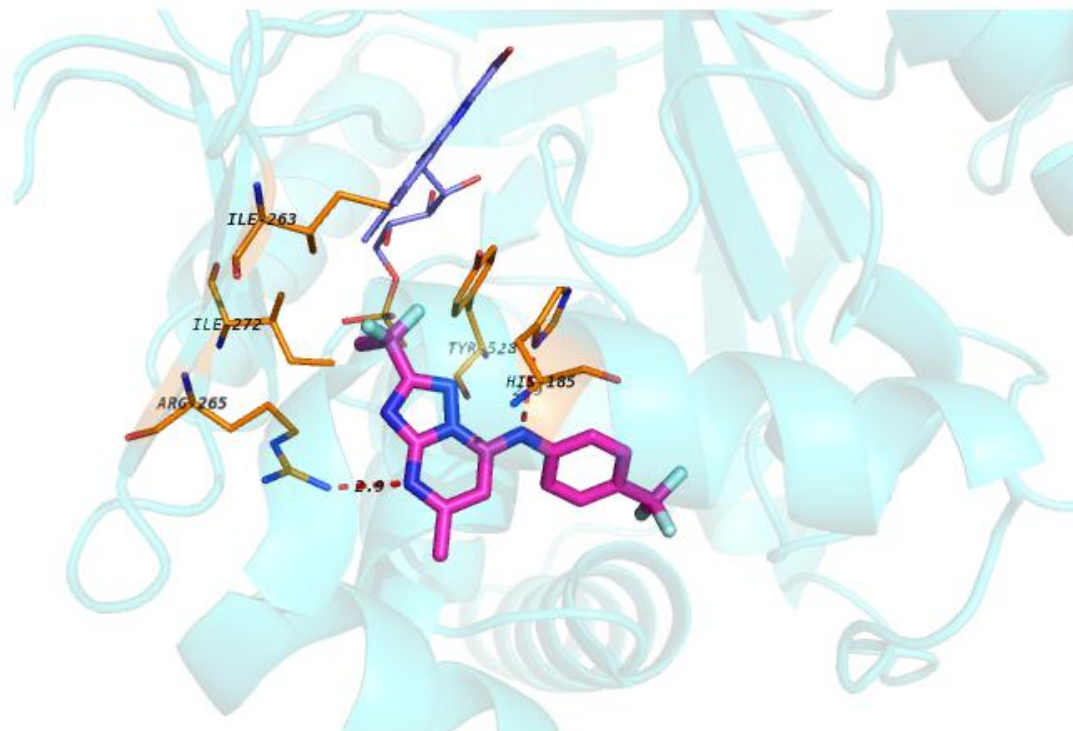


Figura 23 - (A) Sítio de interação da enzima *PfdHODH* co-cristalizada com o ligante (DMS 267) (PDB ID: 3SFK). (B) *Ligplot* das interações do inibidor com os aminoácidos da enzima.

Portanto, os compostos denominados *cliffs* apresentam essa característica devido ao espaço químico da cavidade de interação da *PfDHODH* com os ligantes.

Diante dessas características do conjunto de dados, alguns requisitos estruturais essenciais para a atividade biológica foram elucidados e estão representados na figura 24. Além disso, conclui-se que devido a diversidade do conjunto, há uma dificuldade na construção de modelos de QSAR, em que se preconiza o uso de um conjunto homogêneo.

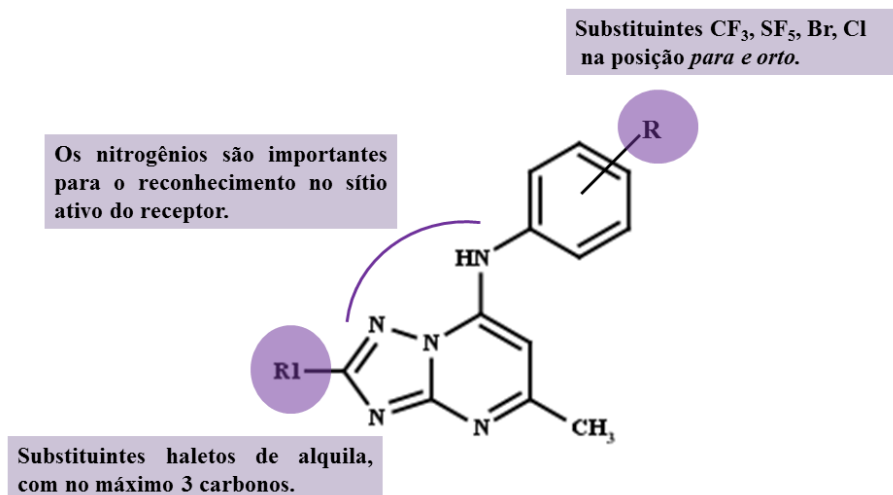


Figura 24 – Requisitos estruturais necessários para a atividade antimalárica de triazolopirimidinas, de acordo com o estudo de SAR.

4.3. Divisão do conjunto de dados para construção dos modelos de QSAR

O conjunto de dados foi dividido em conjunto treinamento e conjunto teste contendo 80% e 20% dos compostos, respectivamente. Para essa divisão utilizou-se o método racional com o algoritmo *Kernnard-Stone* modificado. Essa divisão racional pode assegurar uma maior diversidade estrutural dos compostos em ambos os conjuntos. O conjunto treinamento foi constituído por 102 compostos e o conjunto teste por 26 compostos. A faixa de distribuição da atividade biológica entre esses dois conjuntos está representada na Figura 25.

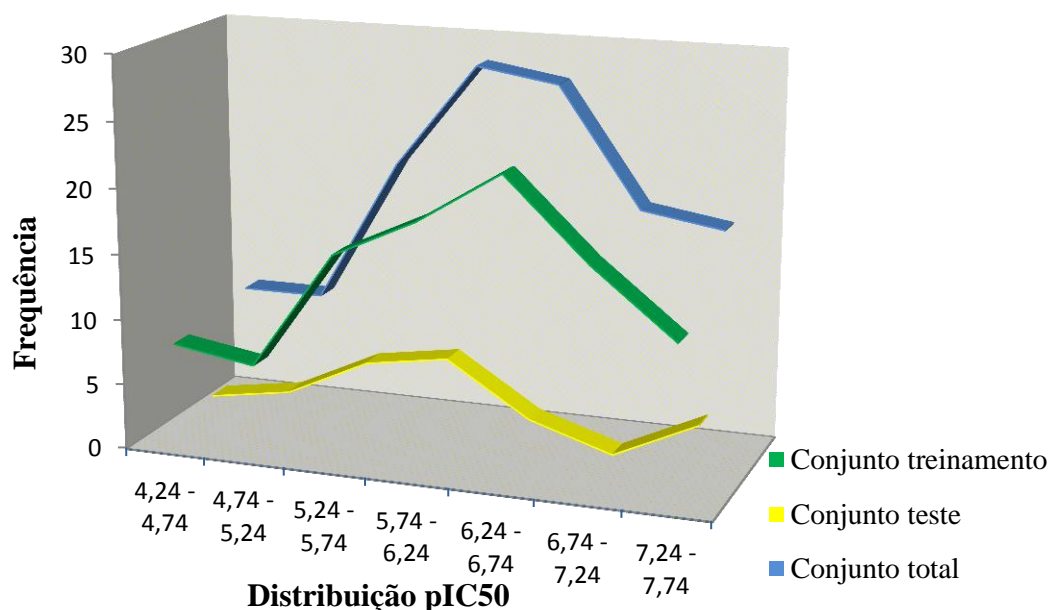


Figura 25 - Representação gráfica da distribuição dos valores de pIC₅₀ nos conjuntos teste, treinamento e conjunto total.

4.4. HQSAR

A construção de modelos HQSAR é sensível a variações nos parâmetros que especificam o comprimento do holograma molecular, tamanho e tipo de fragmento codificado, bem como as diferentes combinações destes parâmetros. Por isso, várias combinações de distinções de fragmento e comprimentos de holograma foram investigadas, buscando as combinações que apresentassem melhores parâmetros estatísticos.

A obtenção dos fragmentos moleculares foi realizada utilizando as seguintes distinções de fragmentos: átomos (A), ligações (B), conectividade (C), átomos de hidrogênio (H), quiralidade (Ch) e doadores e aceptores (DA). Inicialmente, essas distinções de fragmentos foram combinadas empregando o tamanho padrão de fragmento (4–7 átomos). As análises foram realizadas utilizando os 12 comprimentos de holograma disponíveis como padrão do SYBYL-X v.1.2, de 53 a 401 (53, 59, 61, 71, 83, 97, 151, 199, 257, 307, 353 e 401) (SEEL, TURNER, WILLETT, 1999). Adicionalmente, outros comprimentos de holograma maiores foram testados (601, 809 e

997), com a finalidade de se investigar a influência do aumento do comprimento do holograma na geração dos modelos de HQSAR. Os modelos que apresentaram resultados estatísticos com $q^2 \geq 0,5$ a partir das análises de PLS foram considerados para futuras análises e estão apresentados na (Tabela 6) (TROPSHA; GRAMATICA; GOMBAR, 2003).

Tabela 6 – Resultados das análises de HQSAR para várias combinações de distinção de fragmentos utilizando o tamanho padrão de fragmento (4-7), com conjunto de treinamento constituído por 102 compostos.

Modelo	Distinção	q^2	r^2	SEE	HL	N
1	A	0,50	0,81	0,63	353	7
2	A/H	0,50	0,82	0,63	401	8
3	A/DA	0,50	0,86	0,65	401	10
4	A/B/C	0,53	0,87	0,61	307	8
5	A/C/H	0,52	0,87	0,63	401	10
6	A/H/Ch	0,51	0,82	0,62	401	8
7	A/H/DA	0,54	0,88	0,61	353	10
8	A/B/C/H	0,52	0,85	0,62	151	9
9	A/B/C/Ch	0,53	0,87	0,61	307	8
10	A/C/H/Ch	0,52	0,86	0,62	401	9
11	A/C/H/DA	0,50	0,86	0,63	401	9
12	A/B/C/H/Ch	0,52	0,87	0,62	151	10
13	A/B/C/H/DA	0,53	0,92	0,62	353	12
14	A/B/C/H/Ch/DA	0,55	0,86	0,60	353	8

q^2 , coeficiente de correlação da validação cruzada (LOO); r^2 , coeficiente de correlação de determinação múltipla; SEE, erro padrão estimado; HL, comprimento do holograma; N, número de variáveis latentes ou componentes PLS. Distinção de fragmento: A, átomos; B, ligação; C, conectividade; H, hidrogênio; Ch, quiralidade; DA, doador e aceptor.

A análise dos compostos que não foram bem preditos pelos modelos foi feita através da análise dos valores residuais, sendo a diferença entre o valor de atividade experimental e predito. Os compostos que apresentaram valores residuais maiores do que duas vezes o desvio padrão dos resíduos foi considerado como *outlier*, ou seja, o composto não foi bem predito pelo modelo. Apenas seis modelos, representados na Tabela 7, apresentaram melhores resultados quando foram retirados compostos *outliers*.

Tabela 7 – Modelos de HQSAR gerados após a retirada dos *outliers*.

Modelo	Distinção	q^2	r^2	<i>SEE</i>	<i>HL</i>	<i>N</i>	Compostos removidos
5	A/C/H	0,67	0,97	0,53	401	17	114,132,30,34,37
7	A/H/DA	0,69	0,95	0,51	353	14	127,16,132,37,30
8	A/B/C/H	0,63	0,91	0,54	151	11	89,30,37
10	A/C/H/Ch	0,63	0,88	0,54	401	8	114,132,30,34
11	A/C/H/DA	0,67	0,93	0,52	401	12	132,34,37,58
12	A/B/C/H/Ch	0,65	0,93	0,53	151	12	89,131,30,37

q^2 , coeficiente de correlação da validação cruzada (LOO); r^2 , coeficiente de correlação de determinação múltipla; *SEE*, erro padrão estimado; *HL*, comprimento do holograma; *N*, número de variáveis latentes ou componentes PLS. Distinção de fragmento: A, átomos; B, ligação; C, conectividade; H, hidrogênio; Ch, quiralidade; DA, doador e aceptor.

Os compostos 127, 131, 132 apresentam o esqueleto triazolopiridiníco com mudança na posição dos nitrogênios, fato que pode explicar porque estes compostos não foram bem preditos pelos modelos, e portanto, e apresentaram comportamento de *outliers*. Já os compostos 30, 34, 37 apresentam substituintes do anel benzênico ligado a amida, de cadeia longa em relação aos demais substituinte do conjunto, o que pode explicar o fato de não serem bem preditos.

Em seguida, a variação do tamanho dos fragmentos foi investigada para os melhores modelos de HQSAR gerados (apresentados na Tabela 7). Estes parâmetros foram modificados de maneira a incluir fragmentos maiores ou menores dentro do holograma molecular. Os resultados desta análise (Tabela 8), demonstrou que apenas o modelo A/C/H teve melhora, dos parâmetros estatísticos, ao realizar a mudança do tamanho de fragmento para 3-6 átomos ($q^2_{LOO} = 0,70$, $r^2 = 0,96$ e $SEE = 0,51$) (Tabela 7).

Tabela 8 – Influência dos diferentes tamanhos dos fragmentos nos parâmetros estatísticos nos melhores modelos de HQSAR.

Modelo	Distinção	fragmento	q^2	r^2	<i>SEE</i>	<i>HL</i>	<i>N</i>
5	A/C/H	3-6	0,69	0,96	0,51	401	15
7	A/H/DA	4-7	0,69	0,95	0,51	353	14
8	A/B/C/H	4-7	0,63	0,91	0,54	151	11
10	A/C/H/Ch	4-7	0,63	0,88	0,54	401	8
11	A/C/H/DA	4-7	0,67	0,93	0,52	401	12
12	A/B/C/H/Ch	4-7	0,65	0,93	0,53	151	12

q^2 , coeficiente de correlação da validação cruzada (LOO); r^2 , coeficiente de correlação de determinação múltipla; *SEE*, erro padrão estimado; *HL*, comprimento do holograma; *N*, número de variáveis latentes ou componentes PLS. Distinção de fragmento: A, átomos; B, ligação; C, conectividade; H, hidrogênio; Ch, quiralidade; DA, doador e aceptor.

Por esta razão, os modelos 5 (A/C/H; $q_{LOO}^2 = 0,69$, $r^2 = 0,96$ e $SEE = 0,51$), modelo 7 (A/H/DA; $q_{LOO}^2 = 0,69$, $r^2 = 0,95$ e $SEE = 0,51$), modelo 8 (A/B/C/H; $q_{LOO}^2 = 0,63$, $r^2 = 0,91$ e $SEE = 0,54$), modelo 10 (A/C/H/Ch; $q_{LOO}^2 = 0,63$, $r^2 = 0,88$, $SEE = 0,54$) e modelo 11 (A/C/H/DA; $q_{LOO}^2 = 0,67$, $r^2 = 0,93$ e $SEE = 0,52$) e modelo 12 ($q_{LOO}^2 = 0,65$, $r^2 = 0,93$ e $SEE = 0,53$) foram considerados para análises posteriores.

Estes modelos foram avaliados quanto à sua capacidade preditiva dos compostos do conjunto teste, que não foram utilizados na construção dos modelos, e os resultados obtidos estão apresentados na Tabelas 9 e 10, e na Figura 26

Tabela 9 – Valores experimentais e preditos de pIC_{50} dos compostos do conjunto teste calculados para os seis melhores modelos.

Composto	pIC_{50}											
	Modelo 5 ^a		Modelo 7 ^b		Modelo 8 ^c		Modelo 10 ^d		Modelo 11 ^e		Modelo 12 ^f	
	Exp. [*]	Pred. [†]	Exp.	Pred.	Exp.	Pred.	Exp.	Pred.	Exp.	Pred.	Exp.	Pred.
9	4,41	4,97	4,41	5,44	4,41	4,99	4,41	5,55	4,41	5,10	4,41	4,94
11	4,77	4,52	4,77	4,65	4,77	4,93	4,77	5,30	4,77	4,89	4,77	4,73
13	5,03	5,20	5,03	5,06	5,03	5,20	5,03	5,53	5,03	5,26	5,03	5,08
15	6,10	6,50	6,10	5,75	6,10	6,42	6,10	6,21	6,10	5,69	6,10	6,34
20	6,07	5,46	6,07	5,99	6,07	5,35	6,07	5,52	6,07	5,48	6,07	5,31
24	5,80	6,39	5,80	6,18	5,80	6,16	5,80	6,26	5,80	5,93	5,80	6,21
25	5,85	6,35	5,85	5,92	5,85	6,29	5,85	6,25	5,85	5,74	5,85	6,33
28	5,04	5,47	5,04	5,68	5,04	5,44	5,04	5,78	5,04	5,40	5,04	5,43
41	6,10	6,16	6,10	6,02	6,10	5,64	6,10	6,33	6,10	6,39	6,10	5,79
43	5,60	5,90	5,60	5,74	5,60	7,17	5,60	5,70	5,60	5,91	5,60	7,26
65	5,85	5,23	5,85	5,24	5,85	5,06	5,85	5,48	5,85	5,28	5,85	4,80
66	6,92	6,76	6,92	6,71	6,92	6,68	6,92	6,92	6,92	6,69	6,92	6,53
74	6,06	6,33	6,06	6,48	6,06	6,19	6,06	6,43	6,06	6,33	6,06	6,14
80	7,42	7,18	7,42	7,40	7,42	6,91	7,42	7,23	7,42	7,45	7,42	6,99
84	7,47	7,29	7,47	7,47	7,47	7,12	7,47	6,94	7,47	7,50	7,47	7,17
88	7,46	7,54	7,46	7,59	7,46	6,34	7,46	7,11	7,46	7,38	7,46	6,28
90	5,66	6,07	5,66	6,00	5,66	5,76	5,66	6,02	5,66	5,85	5,66	5,82
91	5,77	6,03	5,77	5,74	5,77	5,89	5,77	6,01	5,77	5,67	5,77	5,94
97	7,59	7,28	7,59	7,02	7,59	7,38	7,59	7,00	7,59	7,06	7,59	7,47
104	6,30	6,39	6,30	6,16	6,30	6,24	6,30	6,31	6,30	6,19	6,30	6,28
106	6,40	6,53	6,40	6,20	6,40	6,31	6,40	6,44	6,40	6,32	6,40	6,41
109	7,24	6,83	7,24	6,76	7,24	6,67	7,24	6,84	7,24	7,05	7,24	6,69
111	6,62	6,60	6,62	6,25	6,62	6,52	6,62	6,55	6,62	6,27	6,62	6,52
113	5,28	6,29	5,28	5,45	5,28	5,65	5,28	6,25	5,28	5,70	5,28	5,64
115	6,03	6,88	6,03	6,28	6,03	6,23	6,03	6,66	6,03	6,83	6,03	6,15
131	6,82	6,67	6,82	6,74	6,82	7,52	6,82	6,70	6,82	7,00	6,82	7,77

distinção de fragmento: ^aA/C/H, ^bA/H/DA, ^cA/B/C/H, ^dA/C/H/Ch, ^eA/C/H/DA e ^fA/B/C/H/Ch. ^{*}Valor de pIC_{50} experimental, [†]Valor de pIC_{50} predito.

Foram calculados os valores de desvio padrão para cada modelo (Tabela 10). O desvio padrão é calculado através do valor de resíduo de cada composto do conjunto teste, resíduo é a diferença entre o valor experimental e o predito. Valores de resíduos que excederam uma vez o valor do desvio padrão, estão fora do intervalo de confiança de 99%, portanto são considerados mal preditos. Esse parâmetro permite avaliar a capacidade de predição de cada modelo. O modelo com maior capacidade de predição foi o A/B/C/H/Ch, 80,77% e A/B/C/H 73,08%. No entanto, apenas essa avaliação não permite caracterizar um modelo como bom, é necessário analisar outros parâmetros.

Tabela 10 – Desvio padrão dos valores de preditividade (σ), para os seis melhores modelos e capacidade preditiva.

Modelo	Distinção	Desvio padrão σ	Capacidade preditiva (%)
5	A/C/H	0,41	69,23
7	A/H/DA	0,37	69,23
8	A/B/C/H	0,55	73,08
10	A/C/H/Ch	0,46	65,38
11	A/C/H/DA	0,35	65,38
12	A/B/C/H/Ch	0,59	80,77

A Figura 26 apresenta a representação gráfica dos valores experimentais e preditos (calculados) de pIC_{50} para os compostos do conjunto de treinamento (quadrados pretos) e teste (quadrados brancos) para os seis melhores modelos de HQSAR.

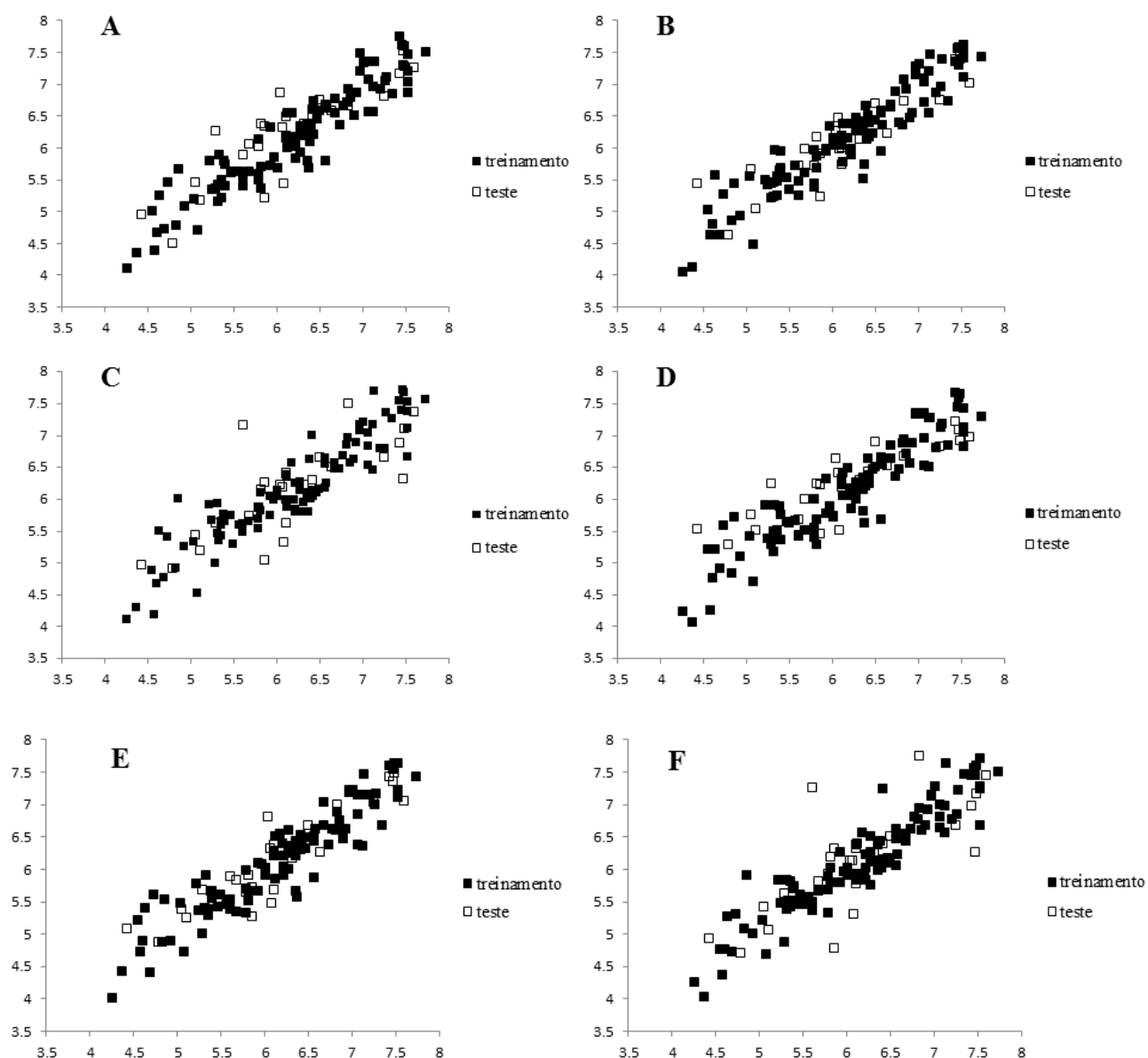


Figura 26 - Representação gráfica dos valores experimentais e preditos de pIC_{50} obtidos com os modelos (A) A/C/H, (B) A/H/DA, (C) A/B/C/H, (D) A/C/H/Ch (E) A/C/H/DA e (F) A/B/C/H/Ch de HQSAR para os compostos do conjunto de treinamento (preto) e teste (branco).

Com o objetivo de obter um valor mais representativo da capacidade preditiva, foi calculado o r_{pred}^2 (CHIRICO; GRAMATICA, 2012). O r_{pred}^2 é uma medida de ajuste que permite estimar a capacidade preditiva de um modelo considerando ambos conjuntos, de treinamento e teste, cujo valores preconizados para um bom modelo é maior que 0,6. Além de outros parâmetros como $r_m^2 \geq 0,5$, coeficiente de correlação modificado, maior $\Delta r_m^2 \leq 0,2$, diferença entre r_m^2 e $r'_m{}^2$, $CCC \geq 0,8$, coeficiente de correlação de concordância, $d = 2$, teste de *Durbin Watson*, correlação entre os resíduos. Na tabela 11 está representado o valor desses parâmetros para cada modelo, os modelos

que apresentaram os melhores resultados foram A/H/DA ($r_{pred}^2 = 0,822$; $CCC = 0,893$; $r_m^2 = 0,689$; $\Delta r_m^2 = 0,175$ e $d = 1,965$ e A/C/H/DA ($r_{pred}^2 = 0,834$; $CCC = 0,904$; $r_m^2 = 0,731$; $\Delta r_m^2 = 0,158$ e $d = 1,885$).

Tabela 11 – Resultados estatísticos da validação externa para os seis melhores modelos de HQSAR.

Modelo	Distinção	r_{pred}^2	CCC	r_m^2	Δr_m^2	d
5	A/C/H	0,751	0,857	0,660	0,203	1,787
7	A/H/DA	0,822	0,893	0,689	0,175	1,965
8	A/B/C/H	0,599	0,762	0,452	0,178	2,372
10	A/C/H/Ch	0,681	0,773	0,400	0,354	1,672
11	A/C/H/DA	0,834	0,904	0,731	0,158	1,885
12	A/B/C/H/Ch	0,549	0,753	0,430	0,068	2,289

r_{pred}^2 , coeficiente de correlação preditivo, r_m^2 , coeficiente de correlação modificado, Δr_m^2 , diferença entre r_m^2 e $r_{m'}^2$, CCC , coeficiente de correlação de concordância, d , teste de *Durbin Watson*, correlação entre os resíduos. Melhor modelo representado em negrito.

Além de prever a atividade biológica de compostos não incluídos no conjunto de treinamento, os modelos de HQSAR fornecem informações referentes às contribuições individuais sobre os fragmentos moleculares que estão diretamente relacionados com a atividade biológica. Essas informações são expressas através dos mapas de contribuição. Para verificar a contribuição dos diferentes fragmentos da molécula em relação à propriedade biológica avaliada, o módulo de HQSAR do SYBYL-X 1.2 utiliza um código de cores para atribuir aos átomos suas diferentes magnitudes de contribuição. Segundo o código utilizado, cores na região do vermelho do espectro (vermelho, laranja avermelhado e laranja) refletem contribuições desfavoráveis ou negativas, enquanto cores na região do verde (amarelo, azul esverdeado e verde) refletem contribuições favoráveis ou positivas. Átomos com contribuições intermediárias são representados em branco, e átomos constituintes da máxima estrutura comum, por sua vez, são mostrados em azul ciano.

A Figura 27 apresenta os mapas de contribuição obtidos do modelo de HQSAR gerado com as distinções de fragmentos A/C/H/DA, utilizando os compostos mais potente do conjunto de dados (composto 86, 98 e 87), e na figura 28 os compostos menos potentes (composto 116, 110 e 128). Este modelo foi selecionado pois apresentou os melhores parâmetros de preditividade externa de acordo com o preconizado, ($r_{pred}^2 \geq 0,6$, $r_{pred}^2 = 0,834$; $CCC = \geq 0,8$ $CCC = 0,904$; $r_m^2 = 0,6$, $r_m^2 =$

0,731; $\Delta r_m^2 \leq 0,2$, $\Delta r_m^2 = 0,158$, $d = 2$, $d = 1,885$) (CHIRICO; GRAMATICA, 2012) e mais informações na análise dos mapas de contribuição.

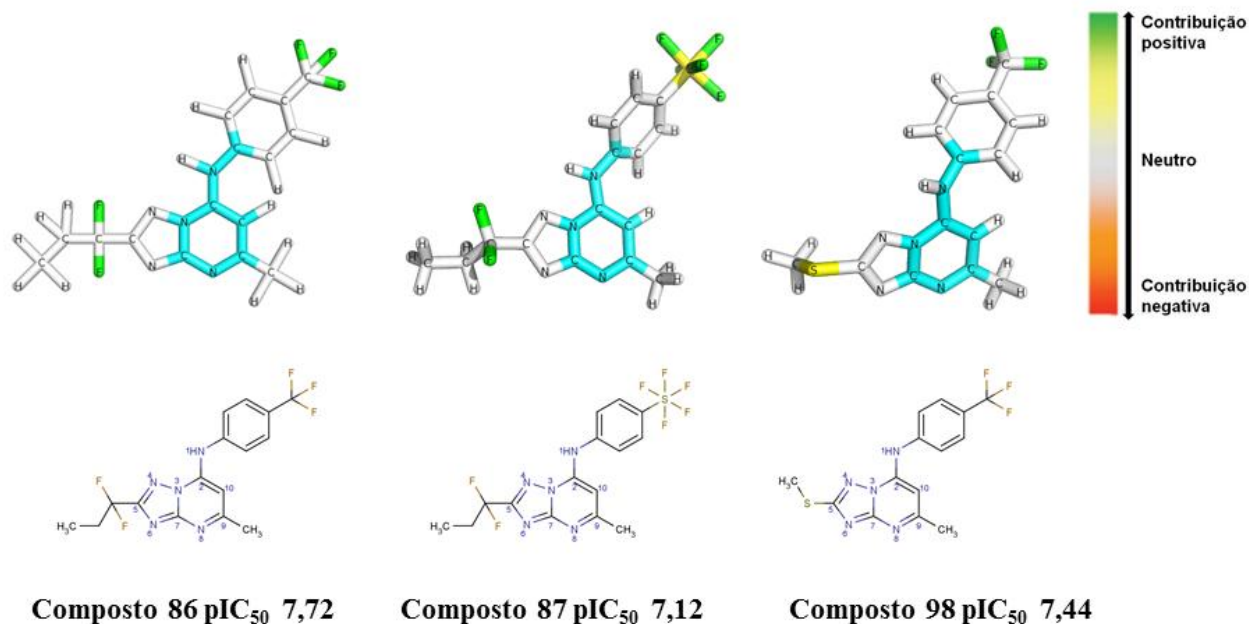


Figura 27 – Representação gráfica dos mapas de contribuição de QSAR para os compostos mais potentes do conjunto de dados (86, 87 e 98), utilizando o modelo gerado com os fragmentos A/C/H/DA. Em azul ciano está representado a máxima subestrutura comum dos compostos.

Os composto mais potente, do conjunto de dados apresentam substituintes hidrofóbicos com o grupo trifluormetila (CF_3) e penta-fluor-sulfono na posição *para* do anel benzênico em verde, favorável, sendo portanto, associado ao aumento de potência deste composto. O grupo haleto de alquila, nos composto 86 e 87, e o tiol, todos com característica hidrofóbica, ligado ao carbono na posição 5 do anel triazolopirimidínico também está associado ao aumento de potencia, pois os átomos de flúor, estão em verde, favorável e o enxofre em amarelo, também favorável. Todas essas informações de regiões favoráveis a potência estão em concordância com a cavidade de interação da *PfDHODH*, obtidos através dos das estruturas raio – X, além dos dados obtidos com os estudos de SAR deste trabalho.

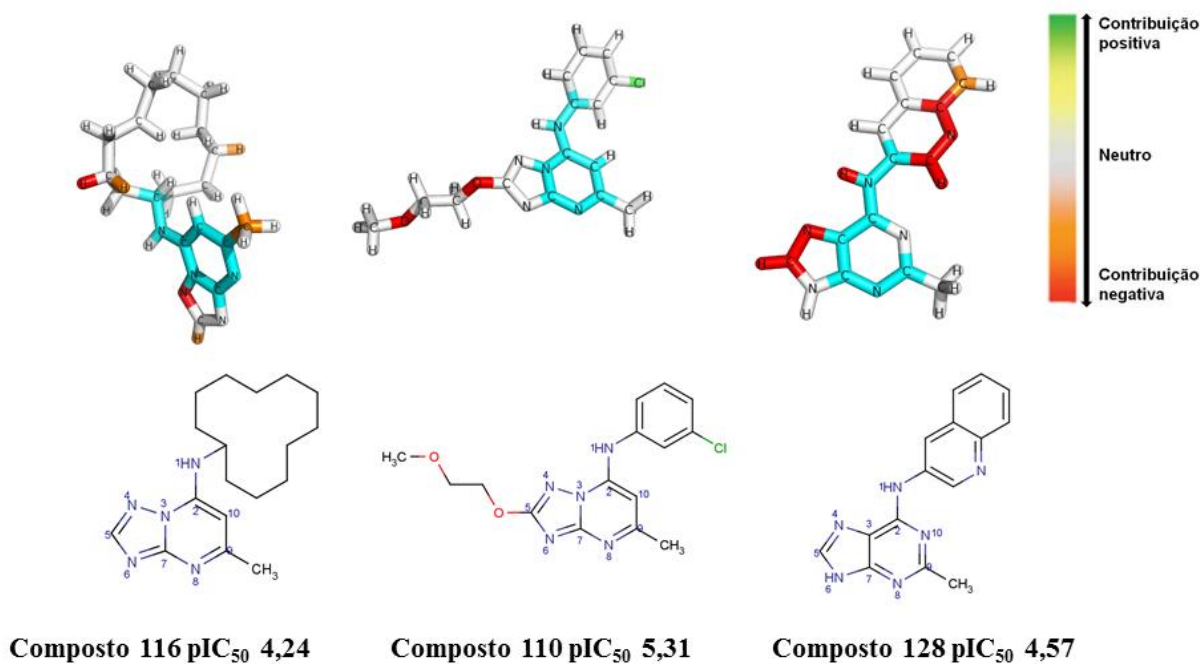


Figura 28 – Representação gráfica dos mapas de contribuição de HQSAR para os compostos, menos potentes do conjunto, 98, 87, 110 e 128 utilizando o modelo gerado com os fragmentos A/C/H/DA. Em azul ciano está representado a máxima subestrutura comum dos compostos.

A diminuição da potência está associada a mudança do tipo de anel ligado ao nitrogênio na posição 1, pois fragmentos do *ciclo*-dodecil estão em vermelho e alaranjado, sendo portanto desfavoráveis, como no composto 116, e fragmentos no heterocíclico ligado ao nitrogênio da posição 1, do composto 128, também apresenta região em vermelho desfavorável. O nitrogênio na posição 4 está em vermelho, desfavorável, para os compostos 116 e 128, o carbono na posição 5 e a metila ligada ao carbono na posição 9, estão em alaranjado, desfavorável no composto 116. Indicando que o anel triazolopirimidínico, que é o *scaffold* do conjunto de dados, se torna desfavorável ao aumento da potência dos compostos, quando não têm substituintes essenciais. Estes dados também estão em concordância com as estruturas cristalográficas descritas na literatura e os estudos de SAR realizados neste trabalho.

No entanto, quando o substituinte do carbono 5 possui átomos de oxigênio é desfavorável a potência, átomos de oxigênio em vermelho, composto 110, as características e tamanho do substituintes do carbono nessa é importante para o aumento da potência. Estes dados também estão em conformidade com o espaço químico da

cavidade de interação, em um bolso hidrofóbico faz interação em essa região do ligantes.

4.5 QSAR - 2D

A utilização de diferentes métodos de aprendizado de máquina para construção dos modelos de QSAR permite identificar qual o melhor método que pode gerar os melhores modelos. Os resultados estatísticos de modelos de QSAR-2D gerados para os conjuntos de treinamento e teste estão resumidos na Tabela 12.

Tabela 12 – Características estatísticas dos modelos de QSAR-2D de para os inibidores de PfdHODH utilizando os diferentes métodos de regressão.

	Conjunto treinamento ($n=102$)				Conjunto teste ($n=28$)					
	r^2	q^2	RMSE	CCC_{int}	r_{pred}^2	CCC	r_m^2	Dr_m^2	G&T	D
PLS ^a	0,94	0,65±0,16	0,56±0,14	0,97	0,40	0,69	0,33	0,04	R	2,00
GAMBOOST	0,88	0,63±0,17	0,55±0,12	0,92	0,73	0,80	0,46	0,30	R	1,77
RF ^b	0,95	0,61±0,19	0,57±0,15	0,95	0,83	0,87	0,59	0,20	R	1,87
GBM^c	0,98	0,69±0,14	0,51±0,17	0,99	0,85	0,91	0,74	0,14	A	1,93
SVM^d	0,99	0,73±0,14	0,47±0,12	0,99	0,84	0,90	0,73	0,15	A	1,85
KRLS^e	0,99	0,72±0,14	0,47±0,11	0,99	0,84	0,90	0,74	0,10	A	1,98

^aPartial Least Squares; ^bRandom Forest; ^cGradient Boosting Machine ; ^dSupport Vector Machines Radial Cost, ^eKernel-based Regularized Least Squares. q^2 , coeficiente de correlação da validação cruzada (LOO); r^2 , coeficiente de correlação de determinação múltipla; RMSE, erro padrão estimado; r_{pred}^2 , coeficiente de correlação preditivo, r_m^2 , coeficiente de correlação modificado, Dr_m^2 , diferença entre r_m^2 e r'_{m^2} , CCC, coeficiente de correlação de concordância, d , Durbin Watson, correlação entre os resíduos.

A utilização de diferentes métodos de aprendizado de máquina levou a modelos de QSAR-2D robustos e preditivos, com consistência interna e robustez adequados (Tabela 12). Os três melhores modelos foram gerados utilizando GBM, SVM e KLRS (destacados em negrito na Tabela 12).

Para a validação externa, utilizou-se os compostos do conjunto teste. Pode-se observar que três modelos se mostraram superiores, sendo eles GBM, SVM, KLRS. O modelo KRLS apresentou os resultados de predição externa mais satisfatórios ($r_{pred}^2 = 0,84$, $CCC = 0,90$, $r_m^2 = 0,74$, $Dr_m^2 = 0,10$, $d = 1,98$ e G&T = aceito), sendo portanto, selecionado como o melhor modelo de QSAR-2D. Na figura 29 está a representação gráfica dos valores experimentais e preditos de pIC_{50} obtidos com o modelo KLRS.

Este modelo poderá ser utilizado em estudos de triagem virtual, para auxiliar na seleção de compostos com atividade inibitória contra a enzima *PfDHODH*.

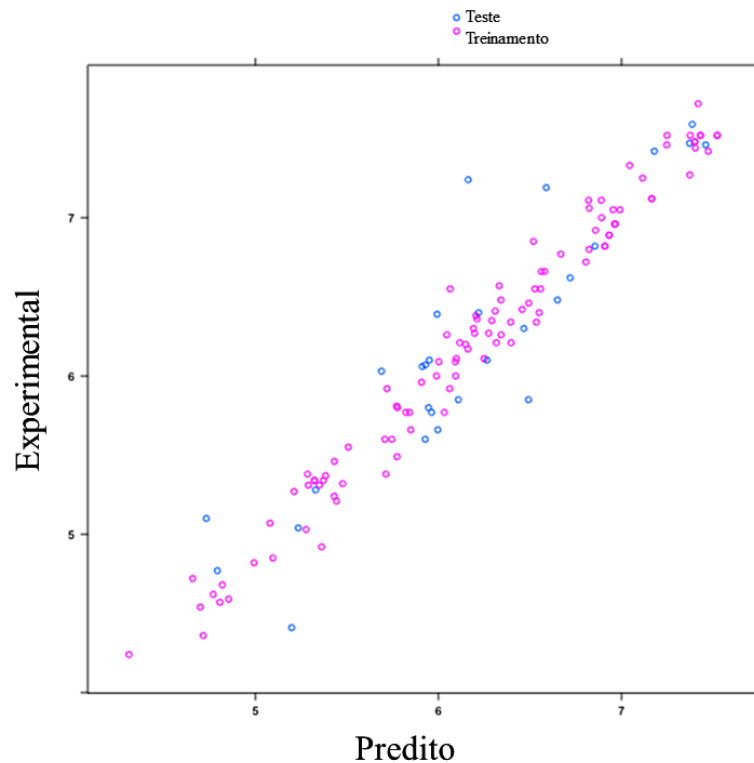


Figura 29 - Representação gráfica dos valores experimentais e preditos de pIC_{50} obtidos com o modelo KRLS de QSAR- 2D para os compostos do conjunto de treinamento (rosa) e teste (azul).

O domínio de aplicabilidade do conjunto teste calculado por distância Euclidiana kNN está representado na figura 30. Pode-se observar que somente o composto 41 está fora do domínio de aplicabilidade; portanto as estruturas do conjunto teste usado para validação externa dos modelos estão no mesmo espaço químico dos compostos utilizados na construção dos modelos, ou seja, do conjunto treinamento, logo são confiáveis as predições feitas.

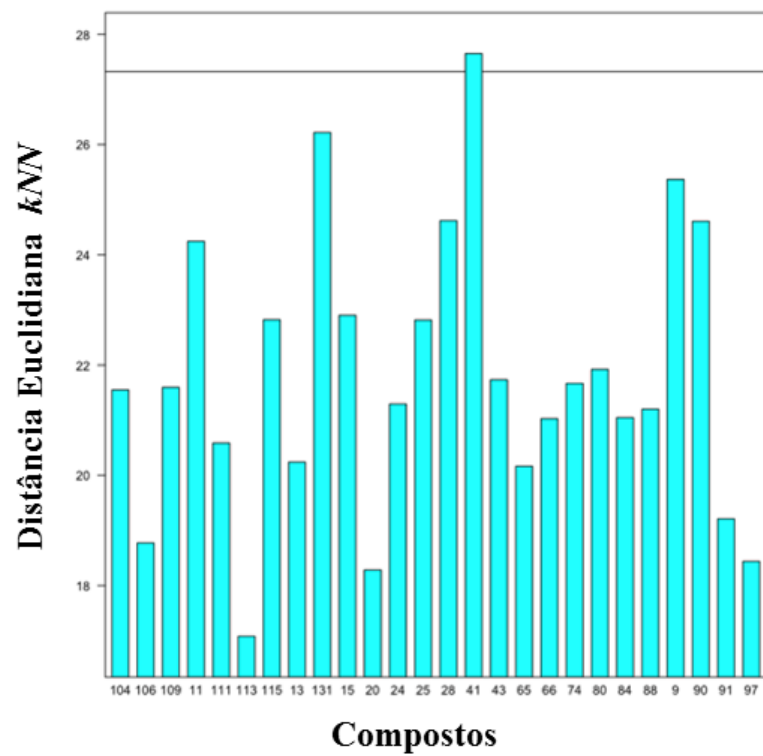


Figura 30 – Domínio de aplicabilidade calculado para os compostos do conjunto teste.

5. CONCLUSÕES

- As análises SAR para o conjunto de triazolopirimidinas permitiu identificar requisitos estruturais necessários para atividade biológica dos compostos.
- Análise de *cliffs* de atividade e de *landscape* de atividade caracterizou o conjunto de dados como descontínuo, em que pequenas mudanças nos compostos podem gerar grandes variações na potência.
- Os dados obtidos com o estudo de SAR e os mapas de contribuição de HQSAR estão em concordância com o espaço químico da cavidade de interação da *PfDHODH* e fornecem informações importantes para o planejamento de novos potenciais inibidores *PfDHODH*.
- Os modelos de HQSAR e QSAR-2D para os inibidores da *PfDHODH* apresentaram bons parâmetros estatísticos e boa capacidade preditiva, com elevada consistência interna e externa, podendo ser utilizados para a previsão da atividade biológica de compostos ainda não sintetizados.
- Os dados obtidos servem de guia para o planejamento de novos compostos com maior potência, evidenciando que as estratégias CADD são ferramentas importantes para o planejamento racional de novos candidatos a fármacos com propriedades otimizadas.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ADEUSI, S. **Pharmaceutical R & D : An Organizational Design Approach to Enhancing Productivity**. 2009. 112p. Dissertação (Mestrado em Ciências em Gerenciamento) – Massachusetts Institute of Technology, Massachusetts, USA, 2009.

AFONSO, A. et al. Resistance to Artemisinin but Lack Mutations in Candidate Genes *atp6* (Encoding the Sarcoplasmic and Endoplasmic Reticulum Ca^{2+} ATPase), *tctp* Malaria Parasites Can Develop Stable Resistance to Artemisinin but Lack Mutations in Candidate Genes *atp6* (Encoding the Sarcoplasmic and Endoplasmic Reticulum Ca^{2+} ATPase), *tctp* , *mdr1* , and *cg10*. **Antimicrobial Agentes and Chemotherapy** v. 50, p. 480-9, 2006.

ALBUQUERQUE, M.G. et al. Multidimensional- QSAR: Beyond the third-dimension in drug design. **Current Methods in Medicinal Chemistry and Biological Physics** v. 1, p. 1-10, 2007.

ALKADI, H. O. Antimalarial Drug Toxicity: a Review. **Chemotherapy** v. 53, n. 6, p. 385–391, 2007.

ANDERS, H. J.; VIELHAUER, V. Identifying and Validating Novel Targets with in Vivo Disease Models: Guidelines for Study Design. **Drug Discovery Today** v. 12, n. 11–12, p. 446–451, 2007.

ANDRADE, C. H. et al. Fragment-based and Classical Quantitative Structure-activity Relationships for a Series of Hydrazides as Antituberculosis Agents. **Molecular Diversity** v. 12, n. 1, p. 47–59, 2008.

ANDRICOPULO, A. D.; SALUM, L. L. B.; ABRAHAM, D. J. Structure-based Drug Design Strategies in Medicinal Chemistry. **Current Topics in Medicinal Chemistry** v. 9, n. 9, p. 771–790, 2009.

ARAV-BOGER, R.; SHAPIRO, T. A. Molecular mechanisms of resistance in antimalarial chemotherapy: the unmet challenge. **Annual review of pharmacology and toxicology**, v. 45, p. 565–85, jan. 2005.

BAJORATH, J. et al. Navigating structure-activity landscapes. **Drug discovery today**, v. 14, n. 13-14, p. 698–705, jul. 2009.

BARREIRO, E. J. A Química Medicinal e o Paradigma Do Composto Protótipo. **Revista Virtual De Química** v. 1, n. 1, p. 26–34, 2009.

BASCO, L. K.; RINGWALD, P. Molecular epidemiology of malaria in Yaounde, Cameroon I. Analysis of point mutations in the dihydrofolate reductase-thymidylate synthase gene of *Plasmodium falciparum*. **The American journal of tropical medicine and hygiene**, v. 58, n. 3, p. 369–73, mar. 1998.

BERMAN, H. M. et al. The Protein Data Bank and the Challenge of Structural Genomics. **Nature Structural Biology** v. 7 Suppl, n. november, p. 957–959, 2000.

BEUTLER, E. et al. Glucose-6-phosphate Dehydrogenase Deficiency and Antimalarial Drug Development. **American Journal of Tropical Medicine and Hygiene** v. 77, n. 4, p. 779–789, 2007.

BRAY P. G. et al. Primaquine Synergises the Activity of Chloroquine Against Chloroquine-resistant *P. Falciparum*. **Biochemical Pharmacology** v. 70, n. 8, p. 1158–1166, 2005.

BRAY, P. G.; WARD, S. A.; O'NEILL, P. M. Quinolines and Artemisinin: Chemistry, Biology and History. **Current Topics in Microbiology and Immunology** 2005.

BRAY, R. S.; GARNHAM, P. C. C. The Life-cycle of Primate Malaria Parasites. **British Medical Bulletin** v. 38, n. 2, p. 117–122, 1982.

BRIOLANT, S. et al. In Vitro Susceptibility to Quinine and Microsatellite Variations of the Plasmodium Falciparum Na⁺/H⁺ Exchanger (Pfnhe-1) Gene: The Absence of Association in Clinical Isolates from the Republic of Congo. **Malaria Journal** v. 10, n. 1, p. 37, 2011.

BUTLER, A.; KHAN, S.; FERGUSON, E. A Brief History of Malaria Chemotherapy. **The Journal of the Royal College of Physicians of Edinburgh** v. 40, n. 2, p. 172–177, 2010.

BURROWS, J. N. Antimalarial drug discovery: where next? **Future medicinal chemistry**, v. 4, n. 18, p. 2233–5, dez. 2012.

CARVALHO, S.G.C. **Estudos estruturais e funcionais de diidroorotato desidrogenases. 2008. 174f.** Tese (Doutorado em Física Aplicada) Instituto de Física de São Carlos da Universidade de São Paulo, São Carlos, 2008.

CENTER FOR DISEASE CONTROL AND PREVENTION. Malaria biology. 2012, Disponível em: <<http://www.cdc.gov/malaria/about/biology/index.html>>. Acesso em: 10 de julho de 2012.

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.

COHEN, N.C. **Guidebook on molecular modeling in drug design.** San Diego: Academic Press, p. 361, 1996.

CONSONNI, V.; BALLABIO, D.; TODESCHINI, R. Comments on the Definition of the Q₂ Parameter for QSAR Validation. **Journal of Chemical Information and Modeling** v. 49, n. 7, p. 1669–1678, 2009.

CONSONNI, V.; BALLABIO, D.; TODESCHINI, R. Evaluation of model predictive ability by external validation techniques. **Journal of Chemometrics**, v. 24, n. 3-4, p. 194–201, 17 fev. 2010.

CONSONNI, V.; TODESCHINI, R. Molecular descriptors. In: PUZYN, T.; LESZCZYNSKI, J.; CRONIN, M. T. (Eds.). **Recent Advances in QSAR Studies**. Dordrecht: Springer, 2010. p. 29–.

CLAUDIO, A. C.; ZANDONADE, E. Proposição, Validação e Análise Dos Modelos Que Correlacionam Estrutura Química e Atividade Biológica. **Química Nova** v. 24, n. 5, p. 658–671, 2001.

COHEN, F. J. Macro Trends in Pharmaceutical Innovation. **Nature Reviews. Drug Discovery** v. 4, n. 1, p. 78–84, 2005.

COMBRINCK, J. M. et al. Insights into the Role of Heme in the Mechanism of Action of Antimalarials. **ACS Chemical Biology** v. 8, n. 1, p. 133–137, 2013.

CONSONNI, V.; BALLABIO, D.; TODESCHINI, R. Comments on the Definition of the Q^2 Parameter for QSAR Validation. **Journal of Chemical Information and Modeling** v. 49, n. 7, p. 1669–1678, 2009.

CONSONNI, V.; BALLABIO, D.; TODESCHINI, R. Evaluation of Model Predictive Ability by External Validation Techniques. **Journal of Chemometrics** v. 24, n. 3, p. 194–201, 2010.

COTERON, J. M. et al. Structure-guided lead optimization of triazolopyrimidine-ring substituents identifies potent Plasmodium falciparum dihydroorotate dehydrogenase inhibitors with clinical candidate potential. **Journal of medicinal chemistry**, v. 54, n. 15, p. 5540–61, 11 ago. 2011.

COUZIN, J. Brains behind blockbuster. **Science**, v. 309, p. 728-730, 2005

COWMAN, A. F.; CRABB, B. S. The Plasmodium Falciparum Genome--a Blueprint for Erythrocyte Invasion. **Science** v. 298, n. 5591, p. 126–128, 2002.

COWMAN, A. F.; GALATIS, D.; THOMPSON, J. K. Selection for Mefloquine Resistance in Plasmodium Falciparum Is Linked to Amplification of the Pfmdr1 Gene and Cross-resistance to Halofantrine and Quinine. **Proceedings of the National Academy of Sciences of the United States of America** v. 91, n. 3, p. 1143–1147, 1994.

CRAMER, R. D.; PATTERSON, D. E.; BUNCE, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. **Journal of the American Chemical Society** v. 110, n. 18, p. 5959–5967, 1988.

CROOKE, S. T. Optimizing the impact of genomics on drug discovery and development. **Nature biotechnology**, v. 16 Suppl, p. 29–30, maio 1998.

DAVIS, T. M. E.; HUNG, T.-Y.; SIM, I.-K.; KARUNAJEEWA, H. A.; ILETT, K. F. Piperaquine: A Resurgent Antimalarial Drug. **Drugs** v. 65, n. 1, p. 75–87, 2005.

DEBNATH, A.K. Quantitative Structure-Activity Relationship (QSAR) paradigm - Hansch era to new millennium. **Mini-Reviews Medicinal Chemistry** v. 1, p. 187-195, 2001.

DENG, X. et al. Structural plasticity of malaria dihydroorotate dehydrogenase allows selective binding of diverse chemical scaffolds. **The Journal of biological chemistry**, v. 284, n. 39, p. 26999–7009, 25 set. 2009.

DONDORP, A. M. et al. Artemisinin Resistance in Plasmodium Falciparum Malaria. **New England Journal of Medicine** v. 361, n. 5, p. 455–467, 2009.

DUBIN, J.; WATSON, G. S. Testing for serial correlation in least square regression I. **Biometrika** v. 34, n 3, p. 409 - 428, 1950.

EASTMAN, R. T.; FIDOCK, D. A. Artemisinin-based Combination Therapies: A Vital Tool in Efforts to Eliminate Malaria. **Nature Reviews Microbiology** v. 7, n. 12, p. 864–874, 2009.

EDWARDS, G. et al. Interactions Among Primaquine, Malaria Infection and Other Antimalarials in Thai Subjects. **British Journal of Clinical Pharmacology** v. 35, n. 2, p. 193–198, 1993.

EFFERTH, T.; KAINA, B. Toxicity of the Antimalarial Artemisinin and Its Derivatives. **Critical Reviews in Toxicology** v. 40, n. 5, p. 405–421, 2010.

EGAN T. J. Haemozoin Formation. **Molecular and Biochemical Parasitology** v. 157, n. 2, p. 127–136, 2008.

EGAN, T. J. Chloroquine and Primaquine: Combining Old Drugs as a New Weapon Against *Falciparum* Malaria? **Trends in Parasitology** v. 22, n. 6, p. 235–237, 2006.

EUROPEAN COMMISSION. **Pharmaceutical Sector Inquiry Preliminary Report**. Brussels, Belgium, 2008. 426p.

FALL, B. et al. Plasmodium Falciparum Susceptibility to Anti-malarial Drugs in Dakar, Senegal, in 2010: An Ex Vivo and Drug Resistance Molecular Markers Study. **Malaria Journal** v. 12, p. 107, 2013.

FERREIRA, M. M. C. Multivariate QSAR. **Journal of the brazilian chemical society** v. 13, p. 742–753, 2002.

FITCH C. D. Ferriprotoporphyryn IX, Phospholipids, and the Antimalarial Actions of Quinoline Drugs. **Life Sciences** v. 74, n. 16, p. 1957–1972, 2004.

FITCH, C. D.; CHAN, R. L.; CHEVLI, R. Chloroquine Resistance in Malaria: Accessibility of Drug Receptors to Mefloquine. **Antimicrobial Agents and Chemotherapy** v. 15, n. 2, p. 258–262, 1979.

FIVELMAN, Q. L.; ADAGU, I. S.; WARHURST, D. C. Effects of Piperaquine, Chloroquine, and Amodiaquine on Drug Uptake and of These in Combination with Dihydroartemisinin Against Drug-sensitive and -resistant Plasmodium Falciparum Strains. **Antimicrobial Agents and Chemotherapy** v. 51, n. 6, p. 2265–2267, 2007.

FOLEY, M.; TILLEY, L. Quinoline Antimalarials: Mechanisms of Action and Resistance and Prospects for New Agents. **Pharmacology and Therapeutics** v. 79, n. 1, p. 55–87, 1998.

FRANÇA, T. C.; SANTOS, M. G.; VILLAR, J. F. Malária: Aspectos Históricos e Quimioterapia. **Química Nova** v. 31, n. 5, p. 1271–1278, 2008.

GARDNER, M. J. et al. Genome Sequence of the Human Malaria Parasite Plasmodium Falciparum. **Nature** v. 419, n. 6906, p. 498–511, 2002.

GEDECK, P.; LEWIS, R. A. Exploiting QSAR Models in Lead Optimization. **Current Opinion in Drug Discovery & Development** v. 11, n. 4, p. 569–575, 2008.

GEERTRUYDEN, V. J. P.; THOMAS, F.; ERHART, A.; D'ALESSANDRO, U. The Contribution of Malaria in Pregnancy to Perinatal Mortality. **The American Journal of Tropical Medicine and Hygiene** v. 71, n. 2 Suppl, p. 35–40, 2004.

GHOSE, A. K.; HERBERTZ, T.; SALVINO, J. M.; MALLAMO, J. P. Knowledge-based Chemoinformatic Approaches to Drug Discovery. **Drug Discovery Today** v. 11, n. 23–24, p. 1107–1114, 2006.

GIL, J. P. et al. Detection of Atovaquone and MalaroneTM Resistance Conferring Mutations in Plasmodium Falciparum Cytochrome b Gene (cytb). **Molecular and Cellular Probes** v. 17, n. 2-3, p. 85–89, 2003.

GOLBRAIKH, A.; TROPSHA, A. Beware of Q2! **Journal of Molecular Graphics & Modelling** v. 20, n. 4, p. 269–276, 2002.

GOLDENBERG, M. et al. New Drugs a Novel Immunomodulator of Active Rheumatoid Arthritis. v. 21, n. 11, p. 1837–1852, 1999.

GRIMBERG, B. T.; MEHLOTRA, R. K. Expanding the Antimalarial Drug Arsenal- Now, But How? **Pharmaceuticals** v. 4, n. 5, p. 681–712, 2011.

GUIDO, R. V. C.; OLIVA, G.; ANDRICOPULO, A. D. Modern Drug Discovery Technologies: Opportunities and Challenges in Lead Discovery. **Combinatorial Chemistry & High Throughput Screening** v. 14, n. 10, p. 830–839, 2011.

GUJJAR, R. et al. Identification of a metabolically stable triazolopyrimidine-based dihydroorotate dehydrogenase inhibitor with antimalarial activity in mice. **Journal of medicinal chemistry**, v. 52, n. 7, p. 1864–72, 9 abr. 2009.

GUJJAR, R. et al. Lead optimization of aryl and aralkyl amine-based triazolopyrimidine inhibitors of Plasmodium falciparum dihydroorotate dehydrogenase with antimalarial activity in mice. **Journal of medicinal chemistry**, v. 54, n. 11, p. 3935–49, 9 jun. 2011.

HANSCH, C.; FUJITA, T. RHO-SIGMA-PI Analysis. A method for the correlation of biological activity and chemical structure. **Journal American Chemical Society** v.86, p.1616-1626, 1964.

HANSCH, C. A quantitative approach to biochemical structure-activity relationships. **Chemical Research** v. 2, p. 232-239, 1969.

HASEGAWA, K.; FUNATSU, K. Evolution of PLS for Modeling SAR and Omics Data. **Molecular Informatics** v.31, n. 11-12, p. 766-775, 2012.

HASTINGS, I. M.; WATKINS, W. M. Tolerance Is the Key to Understanding Antimalarial Drug Resistance. **Trends in Parasitology** v. 22, n. 2, p. 71–77, 2006.

HAY, S. I.; GUERRA, C. A.; TATEM, A. J.; NOOR, A. M.; SNOW, R. W. The Global Distribution and Population at Risk of Malaria: Past, Present, and Future. **Lancet Infectious Diseases** v. 4, n. 6, p. 327–336, 2004.

HAYESHI, R.; MASIMIREMBWA, C.; MUKANGANYAMA, S.; UNGELL, A. L. B. Lysosomal Trapping of Amodiaquine: Impact on Transport Across Intestinal Epithelia Models. **Biopharmaceutics and Drug Disposition** v. 29, n. 6, p. 324–334, 2008.

HONORIO, K. M.; GARRATT, R. C.; ANDRICOPULO, A. D. Hologram Quantitative Structure-activity Relationships for a Series of Farnesoid X Receptor Activators. **Bioorganic & Medicinal Chemistry Letters** v. 15, n. 12, p. 3119–3125, 2005.

HOPFINGER, A.J. et al. Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. **Journal American Chemical Society** v. 119, n. 43, p. 10509-10524, 1997.

ITTARAT, I.; ASAWAMAHASAKDA, W.; BARTLETT, M. S.; SMITH, J. W.; MESHNICK, S. R. Effects of Atovaquone and Other Inhibitors on *Pneumocystis Carinii* Dihydroorotate Dehydrogenase. **Antimicrobial Agents and Chemotherapy** v. 39, n. 2, p. 325–328, 1995.

HUNT, P. et al. Gene encoding a deubiquitinating enzyme is mutated in artesunate- and chloroquine-resistant rodent malaria parasites. **Molecular microbiology**, v. 65, n. 1, p. 27–40, jul. 2007.

KIRALJ, R.; FERREIRA, M. M. C. Basic Validation Procedures for Regression Models in QSAR and QSPR Studies: Theory and Application. **Journal of the Brazilian Chemical Society** v. 20, n. 4, p. 770–787, 2009.

KOPEC, K. K.; BOZYCZKO-COYNE, D.; WILLIAMS, M. Target Identification and Validation in Drug Discovery: The Role of Proteomics. **Biochemical Pharmacology** v. 69, n. 8, p. 1133–1139, 2005.

KUBINYI, H. QSAR and 3D QSAR in drug design. **Drug Discovery Today** v. 2, p. 457-467, 1997.

KYAW, M. P. et al. Reduced Susceptibility of Plasmodium Falciparum to Artesunate in Southern Myanmar. **PLoS ONE** v. 8, n. 3, p. e57689, 2013.

LI, Q.; HICKMAN, M. Toxicokinetic and Toxicodynamic (TK/TD) Evaluation to Determine and Predict the Neurotoxicity of Artemisinin. **Toxicology** v. 279, n. 1–3, p. 1–9, 2011.

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.

LOWIS, D.R. HQSAR: a new, highly predictive QSAR technique. **Tripes Technical Notes** v.1, 1997.

LUZZATTO, L. The Rise and Fall of the Antimalarial Lapdap: A Lesson in Pharmacogenetics. **The Lancet** v. 376, n. 9742, p. 739–741, 2010.

MALMBERG, M. et al. Temporal Trends of Molecular Markers Associated with Artemether-lumefantrine Tolerance/resistance in Bagamoyo District, Tanzania. **Malaria Journal** v. 12, p. 103, 2013.

MAHAN, Bruce H; MYERS, Rollie J; TOMA, Henrique Eisi. **Química: um curso universitário**. 4^o ed. São Paulo: Edgard Bucher, 1995.

MANGANELLI, G.; MASULLO, U.; PASSARELLI, S.; FILOSA, S. Glucose-6-phosphate Dehydrogenase Deficiency: Disadvantages and Possible Benefits. **Cardiovascular & Hematological Disorders Drug Targets** v. 13, n. 1, p. 73–82, 2013.

MCGEE, P. Clinical Trials on the Move. **Drug Discovery & Development** v. 9, n. 6, p. 16–22, 2006.

MEDICINE FOR MALARIA VENTURE. Interactive R&D portfolio. 2012, Disponível em: <<http://www.mmv.org/research-development/rd-portfolio>>. Acesso em: 20 de maio de 2014.

MENG, H.; ZHANG, R.; YANG, H.; FAN, Q.; SU, X.; MIAO, J.; CUI, L.; YANG, Z. In Vitro Sensitivity of Plasmodium Falciparum Clinical Isolates from the China-Myanmar Border Area to Quinine and Association with Polymorphism in the Na⁺/H⁺ Exchanger. **Antimicrobial Agents and Chemotherapy** v. 54, n. 10, p. 4306–4313, 2010.

MILLER, L. H.; BARUCH, D. I.; MARSH, K.; DOUMBO, O. K. The Pathogenic Basis of Malaria. **Nature** v. 415, n. 6872, p. 673–679, 2002.

MITRA, I.; ROY, P. P.; KAR, S.; OJHA, P. K.; ROY, K. On Further Application of rM2 as a Metric for Validation of QSAR Models. **Journal of Chemometrics** v. 24, n. 1, p. 22–33, 2010.

MITTRA, J.; TAIT, J. Analysing Stratified Medicine Business Models and Value Systems: Innovation-regulation Interactions. **New Biotechnology** v. 29, n. 6, p. 709–719, 2012.

MOTA, M. M. et al. Migration of Plasmodium Sporozoites Through Cells Before Infection. **Science** v. 291, n. 5501, p. 141–144, 2001.

MUNOS, B. Lessons from 60 Years of Pharmaceutical Innovation. **Nature Reviews. Drug Discovery** v. 8, n. 12, p. 959–968, 2009.

NÁJERA, J. A.; SILVA, M. G.; ALONSO, P. L. Some Lessons for the Future from the Global Malaria Eradication Programme (1955-1969). **PLoS Medicine** v. 8, n. 1, p. 1–7, 2011.

NIXO, G. L. et al. Antimalarial Pharmacology and Therapeutics of Atovaquone. **The Journal of Antimicrobial Chemotherapy** v. 2, n. c, p. 1–9, 2013.

O'NEILL, P ; BARTON, V. E.; WARD, S. A. The Molecular Mechanism of Action of Artemisinin--the Debate Continues. **Molecules** v. 15, n. 3, p. 1705–1721, 2010.

OJHA, P. K.; ROY, K. Chemometric modeling, docking and in silico design of triazolopyrimidine-based dihydroorotate dehydrogenase inhibitors as antimalarials. **European journal of medicinal chemistry**, v. 45, n. 10, p. 4645–56, out. 2010.

CARRARA, V. I. et al. Malaria Burden and Artemisinin Resistance in the Mobile and Migrant Population on the Thai–Myanmar Border, 1999–2011: An Observational Study. **PLoS Med** v. 10, n. 3, p. 1–15, 2013.

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

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. Guidance document on the validation of (quantitative) structure-activity relationships [(Q)SAR] models relationships [(Q)SAR] models. 2007, Disponível em: <[http://search.oecd.org/officialdocuments/displaydocumentpdf/?doclanguage=en&ote=env/jm/mono\(2007\)2](http://search.oecd.org/officialdocuments/displaydocumentpdf/?doclanguage=en&ote=env/jm/mono(2007)2)>. Acesso em: 10 de fevereiro de 2013

OJHA, P. K.; MITRA, I.; DAS, R. N.; ROY, K. Further Exploring Rm2 Metrics for Validation of QSPR Models. **Chemometrics and Intelligent Laboratory Systems** v. 107, n. 1, p. 194–205, 2011.

PAINTER, H. J.; MORRISEY, J. M.; MATHER, M. W.; VAIDYA, A. B. Specific Role of Mitochondrial Electron Transport in Blood-stage Plasmodium Falciparum. **Nature** v. 446, n. 7131, p. 88–91, 2007.

PATRICK, G.L. An introduction to medicinal chemistry. Oxford: Oxford University Press, 3 Ed., p. 163-184, 2005.

PAUL, S. M. et al. How to Improve R&D Productivity: The Pharmaceutical Industry's Grand Challenge. **Nature Reviews. Drug Discovery** v. 9, n. 3, p. 203–214, 2010.

PETERSEN, I.; EASTMAN, R.; LANZER, M. Drug-resistant Malaria: Molecular Mechanisms and Implications for Public Health. **FEBS Letters** v. 585, n. 11, p. 1551–1562, 2011.

PHRMA, Pharmaceutical Research and Manufacturers of America. **Pharmaceutical Industry Profile 2011**. Washington DC, USA, 2011. 60p.

POROLLO, A. et al. Analysis of Current Antifungal Agents and Their Targets Within the *Pneumocystis Carinii* Genome. **Current Drug Targets** v. 13, n. 12, p. 1575–1585, 2012.

PRICE, R. N. et al. M. Vivax Malaria: Neglected and Not Benign. **The American Journal of Tropical Medicine and Hygiene** v. 77, n. 6, p. 79–87, 2007.

PRUDENCIO, M.; RODRIGUEZ, A.; MOTA, M. M. The Silent Path to Thousands of Merozoites: The Plasmodium Liver Stage. **Nature Reviews Microbiology** v. 4, n. 11, p. 849–856, 2006.

RAYNES, K. Bisquinoline Antimalarials: Their Role in Malaria Chemotherapy. **International Journal for Parasitology** v. 29, n. 3, p. 367–379, 1999.

ROY, K. et al. Comparative Studies on Some Metrics for External Validation of QSPR Models. **Journal of Chemical Information and Modeling** v. 52, n. 2, p. 396–408, 2011.

ROY, K; MITRA, I. On the Use of the Metric R_m^2 as an Effective Tool for Validation of QSAR Models in Computational Drug Design and Predictive Toxicology. **Mini Reviews in Medicinal Chemistry** v. 12, n. 6, p. 491–504, 2012.

ROY, K. et al. Some Case Studies on Application of “R(m) (2) ” Metrics for Judging Quality of Quantitative Structure-activity Relationship Predictions: Emphasis on Scaling of Response Data. **Journal of Computational Chemistry** p. 1–12, 2013.

ROY, K. et al. Introduction of $R_m^2(\text{rank})$ Metric Incorporating Rank-order Predictions as an Additional Tool for Validation of QSAR/QSPR Models. **Chemometrics and Intelligent Laboratory Systems** v. 118, p. 200–210, 2012a.

ROY, K. et al. Introduction of $R_m^2(\text{rank})$ Metric Incorporating Rank-order Predictions as an Additional Tool for Validation of QSAR/QSPR Models. **Chemometrics and Intelligent Laboratory Systems** v. 118, p. 200–210, 2012b.

SANCHEZ, C. P.; MCLEAN, J. E.; STEIN, W.; LANZER, M. Evidence for a Substrate Specific and Inhibitable Drug Efflux System in Chloroquine Resistant Plasmodium Falciparum Strains. **Biochemistry** v. 43, n. 51, p. 16365–16373, 2004.

SEEL, M.; TURNER, D. B.; WILLETT, P. Effect of Parameter Variations on the Effectiveness of HQSAR Analyses. **Molecular Informatics** v. 18, n. 3, p. 245–252, 1999.

SEOANE, J. A et al. New approaches in data integration for systems chemical biology. **Current topics in medicinal chemistry**, v. 13, n. 5, p. 591–601, jan. 2013.

SHAH, P. et al. 3D-QSAR studies of triazolopyrimidine derivatives of Plasmodium falciparum dihydroorotate dehydrogenase inhibitors using a combination of molecular dynamics, docking, and genetic algorithm-based methods. **Journal of chemical biology**, v. 5, n. 3, p. 91–103, jul. 2012.

SINGH, J. C. et al. Plasmodium Knowlesi Malaria in Humans Is Widely Distributed and Potentially Life Threatening. **Clinical Infectious Diseases** v. 46, n. 2, p. 165–171, 2008.

SNOW, R. W.; MARSH, K. New Insights into the Epidemiology of Malaria Relevant for Disease Control. **British Medical Bulletin** v. 54, n. 2, p. 293–309, 1998.

SRIVASTAVA, I. K.; VAIDYA, A. B. A Mechanism for the Synergistic Antimalarial Action of Atovaquone and Proguanil. **Antimicrobial Agents and Chemotherapy** v. 43, n. 6, p. 1334–1339, 1999.

STANTON, D. T. QSAR and QSPR Model Interpretation Using Partial Least Squares (PLS) Analysis. **Current Computer-aided Drug Design** v. 8, n. 2, p. 107–127, 2012.

STEPNIEWSKA, K.; WHITE, N. J. Pharmacokinetic Determinants of the Window of Selection for Antimalarial Drug Resistance. **Antimicrobial Agents and Chemotherapy** v. 52, n. 5, p. 1589–1596, 2008.

STURM, A. et al. Manipulation of Host Hepatocytes by the Malaria Parasite for Delivery into Liver Sinusoids. **Science** v. 313, n. 5791, p. 1287–1290, 2006.

SÁ, J. M. et al. E. Geographic Patterns of Plasmodium Falciparum Drug Resistance Distinguished by Differential Responses to Amodiaquine and Chloroquine. **Proceedings of the National Academy of Sciences of the United States of America** v. 106, n. 45, p. 18883–18889, 2009.

SYBYL-X, version 1.2: Tripos, Inc., St. Louis, MO, USA, 2010.

TAVARES, L.C. QSAR: A Abordagem de Hansch. **Química Nova** v. 27, n. 4, p. 631-639, 2004.

TRIPOS, Certara Inc. **HQSAR™ Manual: SYBYL®-X 1.1 2010**. Saint Louis, USA, 2010a. 92p.

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

TROPSHA, A.; GRAMATICA, P.; GOMBAR, V. The Importance of Being Earnest: Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. **QSAR & Combinatorial Science**, v. 22, n. 1, p. 69–77, abr. 2003.

TUTEJA, R. Malaria - an Overview. **The FEBS Journal** v. 274, n. 18, p. 4670–4679, 2007.

UHLEMANN, A. C.; KRISHNA, S. Antimalarial Multi-drug Resistance in Asia: Mechanisms and Assessment. **Current Topics in Microbiology and Immunology** 2005.

VARNEK, A. et al. ISIDA - Platform for Virtual Screening Based on Fragment and Pharmacophoric Descriptors. **Current Computer-Aided Drug Design**, v. 4, p. 1–8, 2008.

VEBER, D. F. et al. Molecular properties that influence the oral bioavailability of drug candidates. **Journal of Medicinal Chemistry** v. 45, p. 2615–2623, 2002.

VEDANI, A.; DOBLER, M. 5D-QSAR: The key for simulating induced fit? **Journal Medicinal Chemistry** v. 45, p. 2139-2149, 2002.

VEDANI, A.; DOBLER, M.; LILL, M.A. Combining protein modeling and 6D-QSAR. Simulating the binding of structurally diverse ligands to the estrogen receptor. **Journal Medicinal Chemistry** v. 48, p. 3700-3703, 2005.

VERLI, H.; ALBUQUERQUE, M. G.; ALENCASTRO, R. B.; BARREIRO, E. J. Local intersection volume: a new 3D descriptor applied to develop a 3D-QSAR pharmacophore model for benzodiazepine receptor ligands. **European Journal of Medicinal Chemistry** v. 37, p. 219-229, 2002.

VON SEIDLEIN, L. et al. Review of Key Knowledge Gaps in Glucose-6-phosphate Dehydrogenase Deficiency Detection with Regard to the Safe Clinical Deployment of 8-aminoquinoline Treatment Regimens: a Workshop Report. **Malaria Journal** v. 12, n. 1, p. 112, 2013.

VYAS, V. K.; GHATE, M. Recent developments in the medicinal chemistry and therapeutic potential of dihydroorotate dehydrogenase (DHODH) inhibitors. **Mini reviews in medicinal chemistry**, v. 11, n. 12, p. 1039–55, out. 2011.

WARHURST, D. C. et al. Activity of Piperaquine and Other 4-aminoquinoline Antiplasmodial Drugs Against Chloroquine-sensitive and Resistant Blood-stages of *Plasmodium Falciparum*. Role of B-haematin Inhibition and Drug Concentration in Vacuolar Water- and Lipid-phases. **Biochemical Pharmacology** v. 73, n. 12, p. 1910–1926, 2007.

WASSERMANN, A. M.; WAWER, M.; BAJORATH, J. Activity Landscape Representations for Structure-Activity Relationship Analysis. **Journal of medicinal chemistry**, p. 8209–8223, 16 set. 2010.

WELLEMS, T. E.; PLOWE, C. V. Chloroquine-resistant Malaria. **The Journal of Infectious Diseases** v. 184, n. 6, p. 770–776, 2001.

WELLS, T. N. C.; BURROWS, J. N.; BAIRD, J. K. Targeting the Hypnozoite Reservoir of *Plasmodium Vivax*: The Hidden Obstacle to Malaria Elimination. **Trends in Parasitology** v. 26, n. 3, p. 145–151, 2010.

WERMUTH, C.G., *The practice of medicinal chemistry*, 3 Ed. San Diego: Academic Press, p. 982, 2008.

WERMUTH, C.; GANELLIN, C.; LINDBERG, P. GLOSSARY OF TERMS USED IN MEDICINAL CHEMISTRY (IUPAC Recommendations 1998). **Pure and Applied Chemistry** v. 70, n. 5, p. 1129–1143, 1998.

WILKINS, M. R. et al. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. **Bio/technology (Nature Publishing Company)**, v. 14, n. 1, p. 61–5, jan. 1996.

WISHART, D. S. Applications of Metabolomics in Drug Discovery and Development. **Drug in R & D** v. 9 n. 5 p. 307-322, 2008.

WORD HEALTH ORGANIZATION. Malaria indicators. 2012, Disponível em: <<http://apps.who.int/ghodata/?vid=440>>. Acesso em: 15 de julho de 2012.

WORD HEALTH ORGANIZATION. Malaria indicators. 2012, Disponível em: <<http://apps.who.int/ghodata/?vid=440>>. Acesso em: maio de 2014.

WORD HEALTH ORGANIZATION. Country statistics. 2013, Disponível em: < http://www.who.int/malaria/publications/country-profiles/profile_bra_en.pdf.> Acesso em: maio de 2014

XUE, L.; BAJORATH, J. Molecular descriptors in chemoinformatics, computational combinatorial chemistry, and virtual screening. **Combinatorial chemistry & high throughput screening**, v. 3, n. 5, p. 363–72, out. 2000.

YAMAUCHI, L. M.; COPPI, A.; SNOUNOU, G.; SINNIS, P. Plasmodium Sporozoites Trickle Out of the Injection Site. **Cellular Microbiology** v. 9, n. 5, p. 1215–1222, 2007.

YAP, C. W. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. **Journal of computational chemistry**, v. 32, n. 7, p. 1466–74, maio 2011.

YEKA, A ; TIBEDERANA, J.; ACHAN J.; D’ALESSANDRO, U.; TALISNA, A. O. Efficacy of Quinine, Artemether-Lumefantrine and Dihydroartemisinin-Piperaquine as Rescue Treatment for Uncomplicated Malaria in Ugandan Children. **PLoS ONE** v. 8, n. 1, p. e53772, 2013.