

## ESTUDOS DE QSAR 2D BASEADOS EM DESCRITORES TOPOLÓGICOS E FRAGMENTOS MOLECULARES PARA UMA SÉRIE DE DERIVADOS AZÓLICOS ATIVOS CONTRA *Candida albicans*

Jônathas G. Andrade, Humberto F. Freitas e Marcelo S. Castilho\*

Departamento de Medicamentos, Faculdade de Farmácia, Universidade Federal da Bahia, Campus Universitário de Ondina, 40170-115 Salvador - BA, Brasil

Recebido em 29/3/11; aceito em 5/9/11; publicado na web em 30/9/11

2D-QSAR STUDIES BASED ON TOPOLOGICAL DESCRIPTORS AND MOLECULAR FRAGMENTS FOR A SERIES OF AZOLE DERIVATIVES ACTIVE AGAINST *Candida albicans*. Azole derivatives are the main therapeutical resource against *Candida albicans* infection in immunocompromised patients. Nevertheless, the widespread use of azoles has led to reduced effectiveness and selection of resistant strains. In order to guide the development of novel antifungal drugs, 2D-QSAR models based on topological descriptors or molecular fragments were developed for a dataset of 74 molecules. The optimal fragment-based model ( $r^2 = 0.88$ ,  $q^2 = 0.73$  and  $r^2_{\text{pred}} = 0.62$  with 6PCs) and descriptor-based model ( $r^2 = 0.82$ ,  $q^2 = 0.79$  and  $r^2_{\text{pred}} = 0.70$  with 2 PCs), when analysed synergically, suggested that the triazolone ring and lipophilic properties are both important to antifungal activity.

Keywords: antifungal; azole derivatives; QSAR.

### INTRODUÇÃO

Durante as últimas duas décadas, o risco de infecções causadas pela levedura patogênica *C. albicans* tornou-se cada vez mais comum, especialmente em pacientes imunocomprometidos.<sup>1-3</sup> Por exemplo, cerca de 90% dos pacientes com HIV têm pelo menos um episódio de candidíase oral, cujo tratamento pode gerar gastos superiores a 55 mil dólares por pessoa.<sup>4,5</sup> Os fármacos de escolha para tratamento dessa infecção pertencem à classe dos derivados azólicos, que exercem sua função biológica através da inibição competitiva da enzima lanosterol 14 $\alpha$  desmetilase (CYP51).<sup>6</sup> Em função da ampla utilização desses medicamentos, observa-se uma redução na eficácia e aparecimento de cepas resistentes.<sup>7,8</sup> Dentre os principais mecanismos de resistência conhecidos em espécies de *Candida* estão: superexpressão de bombas de efluxo, diminuindo o acúmulo intracelular do fármaco; alterações na afinidade entre os derivados azólicos e alvo macromolecular, em decorrência de mutações envolvendo o gene ERG11, o qual codifica a enzima lanosterol 14 $\alpha$  desmetilase.<sup>9-12</sup> Como alternativa para superar esses problemas estão sendo desenvolvidos novos derivados azólicos ativos contra cepas resistentes.

Estratégias de planejamento baseado na estrutura do alvo terapêutico são amplamente utilizadas pelas grandes indústrias farmacêuticas para identificar moléculas protótipo cuja potência e seletividade devem ser otimizadas através da satisfação de exigências estereais e eletrônicas presentes apenas no alvo macromolecular.<sup>1,13</sup> Entretanto, essa abordagem é limitada pela ausência de informações estruturais da enzima lanosterol 14 $\alpha$ -desmetilase de *C. albicans*.<sup>14</sup> A fim de contornar essa restrição, técnicas de modelagem por homologia têm sido empregadas.<sup>14,15</sup> Contudo, a baixa identidade sequencial entre as proteínas molde e a proteína alvo, aproximadamente 30%, impede que informações detalhadas sobre a disposição espacial dos resíduos-chave para o reconhecimento molecular sejam identificadas de forma inequívoca.<sup>16,17</sup> Apesar disso, diversos estudos de acoplamento molecular e de QSAR 3D já foram realizados no intuito de identificar o provável modo de interação de derivados azólicos<sup>18,19</sup> e/ou as propriedades estereais e eletrônicas que determinam a atividade antifúngica dessa classe de moléculas.<sup>1,20</sup> Alternativamente, técni-

cas baseadas em informações provenientes de ligantes conhecidos podem fornecer informações cruciais sobre as propriedades físico-químicas e estruturais que estão presentes em inibidores da lanosterol 14 $\alpha$ -desmetilase de *C. albicans*. Por exemplo, Singla e Varadaraj<sup>21</sup> desenvolveram modelos de QSAR 2D, por regressão linear múltipla, que destacam a capacidade de descritores topológicos em prever a atividade antifúngica dos derivados triazólicos estudados. Similarmente, a integração de modelos de QSAR baseados em fragmentos moleculares como, por exemplo, holograma QSAR e descritores físico-químicos bidimensionais pode ser considerada uma estratégia interessante, uma vez que dispensa a necessidade de alinhamentos moleculares subjetivos; fornece modelos com qualidade estatística equivalente àquela encontrada em modelos de QSAR 3D, geralmente mais complexos que modelos 2D e, pode auxiliar na identificação de características estruturais e físico-químicas essenciais para a atividade biológica. Dessa forma, visando contribuir para o desenvolvimento de compostos com potente atividade antifúngica foram desenvolvidos modelos QSAR 2D robustos e preditivos, baseados em descritores físico-químicos ou fragmentos moleculares (holograma QSAR).

### PARTE EXPERIMENTAL

#### Conjunto de dados

Um conjunto de 74 derivados triazólicos, com atividade biológica frente à *C. albicans* medida sob condições padronizadas, foi utilizado no desenvolvimento e validação de modelos de QSAR 2D (Tabela 1).<sup>1,22,23</sup> A fim de minimizar variações experimentais inerentes aos ensaios celulares, utilizou-se a relação  $\text{CIM80}_{\text{azol}}/\text{CIM80}_{\text{fluconazol}}$ , onde o valor de  $\text{CIM80}_{\text{azol}}$  representa a concentração mínima do derivado azólico necessária para inibir em 80% o crescimento fúngico e o valor de  $\text{CIM80}_{\text{fluconazol}}$  representa a concentração de fluconazol necessária para produzir o mesmo efeito segundo o protocolo M27-A2 do CLSI (controle do experimento).<sup>24</sup> Posteriormente, esses valores foram convertidos em  $\text{pCIM}_{80}$  ( $-\log [\text{CIM80}_{\text{azol}}]/[\text{CIM80}_{\text{fluconazol}}]$ ) e utilizados como variáveis dependentes nas análises de QSAR. Os valores de  $\text{pCIM}_{80}$  variam entre 0,19 e 2,69, sendo que compostos mais potentes que fluconazol têm valores elevados de  $\text{pCIM}_{80}$ , enquanto compostos com potência similar a fluconazol apresentam valores próximos de zero.

\*e-mail: castilho@ufba.br

**Tabela 1.** Estrutura química e valores de  $pCIM_{80}$  para o grupo treino e grupo teste (\*)

Comp.	R	$pCIM_{80}$
1	3-Cl	1,00
2	2-Br, 5-F	1,06
3	2-F, 4-Br	1,06
4	2-Cl	1,00
5	4-I	1,09
6	3,4-Cl	1,34
7	4-CF <sub>3</sub>	1,05
8	4-NH <sub>2</sub>	0,69
57*	2-Br, 4-F	1,06
58*	3-CF <sub>3</sub>	1,04

Comp.	R	$pCIM_{80}$
9	H	1,10
10	4-Cl	1,42
11	4-NO <sub>2</sub>	0,83
12	4-CH <sub>3</sub>	1,40
13	2,4-Cl	1,44
14	2-Cl	0,83
15	4-Br	1,75
59*	3-CH <sub>3</sub>	1,11

Comp.	R	$pCIM_{80}$
34	4-CH <sub>3</sub>	1,35
35	2-CH <sub>2</sub>	0,75
36	2-Cl	1,37
37	4-C(CH <sub>3</sub> ) <sub>3</sub>	0,79
38	3-CH <sub>3</sub>	1,35
39	4-OCH <sub>3</sub>	1,37
40	2-Br	1,41
41	4-OCH <sub>3</sub> Ph	1,44
42	4-OCF <sub>3</sub>	1,42
43	2-NO <sub>2</sub>	0,78
44	2-NH <sub>2</sub>	0,75
66*	4-F	1,36
67*	3-NO <sub>2</sub>	1,38
68*	4-NH <sub>2</sub>	0,75

Comp.	X	$pCIM_{80}$
45	(CH <sub>2</sub> ) <sub>4</sub>	0,19
46	(CH <sub>2</sub> ) <sub>5</sub>	0,80

Comp.	X	R	$pCIM_{80}$
16	CH <sub>2</sub>	4-F	2,05
17	CH <sub>2</sub>	4-C(CH <sub>3</sub> ) <sub>3</sub>	1,19
18	CH <sub>2</sub>	2-Cl	2,06
19	CH <sub>2</sub>	4-C <sub>2</sub> H <sub>5</sub>	1,76
20	CH <sub>2</sub>	2,4-Cl	2,39
21	CH <sub>2</sub>	4-Cl	2,36
22	CH <sub>2</sub>	H	1,15
23	CH <sub>2</sub>	3-I	2,43
24	CH <sub>2</sub>	4-Br	2,40
25	CH <sub>2</sub>	2-F, 4-Br	2,05
27	CH <sub>2</sub>	3-CF <sub>3</sub>	2,39
28	CH <sub>2</sub>	4-OCF <sub>3</sub>	2,40
29	CH <sub>2</sub> CO	4-Cl	2,09
30	CH <sub>2</sub> CO	4-Br	2,42
31	CH <sub>2</sub> CO	4-OCH <sub>3</sub>	1,48
32	CH <sub>2</sub> CO	2-Cl, 4-CH <sub>3</sub>	1,49
33	CH <sub>2</sub> CO	2,4-F	1,78
60*	CH <sub>2</sub>	3-Cl	1,46
61*	CH <sub>2</sub>	3-CH <sub>3</sub>	2,05
62*	CH <sub>2</sub>	2-Br, 4-F	2,11
63*	CH <sub>2</sub>	4-CF <sub>3</sub>	2,69
64*	CH <sub>2</sub> CO	4-C(CH <sub>3</sub> ) <sub>3</sub>	0,91
65*	CH <sub>2</sub>	4-F	2,07

Comp.	R	$pCIM_{80}$
47	3-Br	1,45
48	4-F	1,40
49	2-Cl	0,81
50	4-Br	1,45
51	4-NO <sub>2</sub>	0,82
52	4-C(CH <sub>3</sub> ) <sub>3</sub>	1,43
53	4-OCH <sub>2</sub> Ph	0,87
54	2-NO <sub>2</sub>	0,22
55	4-NH <sub>2</sub>	0,19
56	CH <sub>3</sub>	0,23
69*	2-CH <sub>3</sub>	0,79
70*	4-Cl	1,41
71*	2-Br	0,85
72*	3-NO <sub>2</sub>	0,82
73*	2-NH <sub>2</sub>	0,19
74*	Ph	0,28

Os derivados triazólicos tiveram suas coordenadas 3D otimizadas sequencialmente por mecânica molecular, utilizando o campo de força Tripos, e através de cálculos semiempíricos AM1 (palavras chave: 1SCF XYZ ESP NOINTER SCALE=1.4 NSURF=2 SCINCR=0.4), como disponível na plataforma SYBYL-X 1.1. Para fins de validação externa dos modelos, o conjunto total de moléculas foi dividido em grupo treino (1-56) e teste (57-74) (Tabela 1) com base na diversidade estrutural e variação de potência das moléculas. Para tal, 337 descritores topológicos, calculados com auxílio do programa DRAGON 5.5 (vide desenvolvimento dos modelos de QSAR baseados em descritores topológicos), foram empregados para uma análise hierárquica de agrupamentos (HCA – *Hierarchical Cluster Analysis*), utilizando método de ligação completo e dados autoescalados. Representantes de todas as famílias foram selecionados aleatoriamente para compor o grupo teste, de forma que a diversidade estrutural contida no grupo teste representasse aquela encontrada no grupo treino. Durante essa etapa, foram tomados os cuidados necessários para que todas as faixas de potência estivessem representadas em ambos os grupos.

### Desenvolvimento de modelos de QSAR baseados em fragmentos

Na técnica de holograma de QSAR (HQSAR) as moléculas foram decompostas em fragmentos de acordo com parâmetros de distinção de fragmento (DF) utilizados ((átomos (A); ligações (B); conectividade (C); átomos de hidrogênio (H); quiralidade (Ch); doador/receptor (DA)). A seguir, considerando os tipos de fragmento, sua frequência e tamanho (TF), foram construídos hologramas moleculares, cujo comprimento pode variar entre 53 e 401 posições. Inicialmente, foram gerados hologramas moleculares com 3 ou mais distinções de fragmento, utilizando tamanho padrão de fragmento (4-7 átomos). Em seguida, os hologramas gerados com base nesses parâmetros foram correlacionados com a atividade biológica através de regressão por mínimos quadrados parciais (PLS – *Partial Least Square*) para a construção dos modelos de QSAR.<sup>25,26</sup> Uma vez identificada a melhor combinação de parâmetros de distinção de fragmentos e comprimento do holograma, esses parâmetros foram fixados para o estudo da influência da variação do tamanho de fragmento (TF) (2-5, 3-6 e 5-8 átomos) sobre os parâmetros estatísticos do melhor modelo de HQSAR. Durante o desenvolvimento dos modelos, utilizou-se a validação cruzada (LOO – *Leave-One-Out*) para determinar o número ótimo de componentes principais.

### Desenvolvimento de modelos de QSAR baseados em descritores topológicos

Para a construção dos modelos de QSAR baseados em descritores topológicos, a estrutura 3D das moléculas do conjunto de dados foi utilizada para o cálculo de 337 descritores físico-químicos e estruturais com auxílio do programa DRAGON 5.5. Em seguida, os descritores com baixa variância interna (< 0,01) ou com alta correlação entre si (>0,97) foram excluídos por não contribuírem com informações relevantes para o desenvolvimento dos modelos de QSAR 2D, uma vez que descritores com valores constantes ou quase constantes não são capazes de explicar a variação de potência observada e o emprego de descritores com alta colinearidade pode resultar em modelos instáveis e com baixo poder preditivo. Os descritores restantes foram utilizados na construção de modelos preliminares de QSAR, com até 4 descritores, através de regressão linear múltipla (RLM). A seleção dos descritores foi realizada com auxílio do método de algoritmo genético, como disponível no programa MOBYDIGS 1.0. Resumidamente, 10 populações de 100 indivíduos (equações com 4 descritores) evoluíram por pelo menos 1000 gerações, sendo

que a cada geração os modelos foram gerados aleatoriamente por mutação ou permutação (*crossover*). Os critérios utilizados para determinar quais indivíduos estão mais aptos (melhores modelos), a cada geração, foram: função QUIK (0,005), função  $Q^2$  assintótico (-0,005), função de redundância RP (0,1) e função de superajuste (*overfitting*) RN (0,01).<sup>27,28</sup>

Após esse processo, as melhores soluções (equações) foram selecionadas de acordo com seus valores de  $q^2$ .<sup>29</sup> Na etapa seguinte, os descritores presentes nos modelos com valor de  $q^2 > 0,4$  foram reunidos, autoescalados e utilizados para a construção de modelos de PLS, utilizando o programa PIROUETTE 4.0.

### Validação dos modelos de QSAR

A capacidade preditiva dos modelos de QSAR foi determinada por validação externa, utilizando um conjunto de 18 derivados triazólicos (compostos 57 a 74), não utilizados no desenvolvimento dos modelos. Os valores de  $r^2_{\text{pred}}$  foram calculados como descrito por Schüürmann e colaboradores.<sup>30,31</sup>

## RESULTADOS E DISCUSSÃO

O desenvolvimento de resistência aos antifúngicos disponíveis na terapêutica é uma adaptação evolutiva esperada para micro-organismos expostos a ambientes inóspitos, mas que pode ter consequências dramáticas, especialmente para pacientes cujo sistema imune está debilitado. Do ponto de vista químico medicinal, esse cenário torna evidente a necessidade de esforços contínuos, voltados para o desenvolvimento de novas entidades químicas que sejam não só eficazes contra cepas resistentes, mas também apresentem baixa incidência de efeitos colaterais (ex. hepatotoxicidade). Esse objetivo pode ser alcançado de forma mais rápida e efetiva quando diversas estratégias de planejamento são utilizadas de forma sinérgica. Dentre as diversas técnicas disponíveis para guiar o planejamento de derivados azólicos mais potentes e seletivos, pode-se destacar aquelas que utilizam apenas informações provenientes de inibidores conhecidos de lanosterol 14 $\alpha$ -desmetilase, uma vez que a ausência de informação estrutural adequada acerca dessa macromolécula restringe a utilização de técnicas baseadas na estrutura do alvo terapêutico. Nesse contexto, a utilização de técnicas de QSAR bidimensionais oferece a vantagem adicional de dispensar a necessidade de se identificar a conformação bioativa, ou de como os derivados azólicos em questão se orientam no sítio ativo, necessitando apenas de informações acerca da estrutura química das moléculas em estudo e seus respectivos valores de propriedade biológica. De fato, a confiabilidade na determinação da propriedade biológica (variável dependente) é crucial para o desenvolvimento de modelos robustos e preditivos. Entretanto, o processo de determinação da concentração inibitória mínima, utilizando normas padronizadas e aceitas internacionalmente, como preconizado pelo Instituto de Padrões Clínicos e Laboratoriais (CLSI), está sujeito a pequenas variações. Por exemplo, em ensaios de microdiluição em caldo, o fluconazol apresenta valores de  $CIM_{80}$  entre 1 e 4  $\mu\text{g/mL}$  para *C. parapsilosis* ATCC 22019, de acordo com as recomendações do protocolo M27-A2 do CLSI.<sup>24</sup> Visando minimizar o erro associado a essas medidas, utilizou-se um padrão interno em cada determinação da propriedade biológica ( $CIM_{80}$  do fluconazol), o que nos levou a utilizar a relação  $-\log[CIM_{80, \text{azol}}/CIM_{80, \text{fluconazol}}]$ , ou simplesmente  $pCIM_{80}$ , como propriedade biológica utilizada no desenvolvimento dos modelos de QSAR. Fluconazol foi utilizado como padrão interno por ter sua faixa de atividade padronizada no protocolo M27-A2 do CLSI e por ter sua atividade biológica descrita nos três trabalhos que serviram de base para esse estudo. Adicionalmente, deve-se destacar que a relação  $-\log[CIM_{\text{azol}}/CIM_{\text{padrão}}]$  já foi empregada, com

sucesso, por Mota e colaboradores<sup>32</sup> no desenvolvimento de modelos quimiométricos.

A investigação das propriedades estruturais de derivados azólicos (Tabela 1) importantes para sua atividade biológica foi realizada através da técnica de holograma QSAR, na qual cada molécula foi decomposta em fragmentos moleculares. Esse processo é influenciado por diversos fatores, como distinção de fragmento (DF), comprimento do holograma (CH) e tamanho do holograma (TH), os quais foram avaliados durante o desenvolvimento do modelo, considerando o tamanho de fragmentos padrão de 4 a 7 átomos (Tabela 2). A incorporação de hidrogênios (H) ou quiralidade (Ch) à distinção padrão (ABC) teve impacto negativo sobre a consistência interna do modelo (compare **2** e **3** vs **1**). A redução da qualidade estatística pode ser explicada pelo aumento substancial de fragmentos obtidos ao se considerar os átomos de hidrogênio, o que pode introduzir ruído na análise de PLS (modelo 2), ou pela incorporação de uma distinção de fragmento (quiralidade) não relacionada com a atividade biológica e que, portanto, ao invés de melhorar a descrição da relação entre a estrutura química e a atividade biológica, atrapalha esse objetivo. Essa última hipótese merece estudos mais aprofundados, uma vez que a atividade biológica de moléculas enantiomericamente puras não foi descrita nos artigos que subsidiaram esse estudo e pelo fato de que fenômenos biológicos distintos (passagem pela membrana e interação com o alvo terapêutico) estão representados por uma única medida de propriedade biológica ( $CIM_{80}$ ). Por outro lado, a adição do parâmetro doador/aceptor (DA) restaurou a qualidade estatística do modelo (**5** vs **1**). Esse resultado se confirma ao eliminar os parâmetros Hidrogênio e quiralidade da distinção de fragmentos, o que resultou no melhor modelo obtido (**9** vs **5**). Essa melhora nos resultados estatísticos indica que ligações de hidrogênio ou interações eletrostáticas são importantes para a atividade biológica dessas moléculas, o que está de acordo com relatos da literatura onde a interação do nitrogênio 4 do anel triazólico, de derivados azólicos, com o átomo de ferro do grupo heme é considerada um das características fundamentais para a atividade antifúngica.<sup>1</sup>

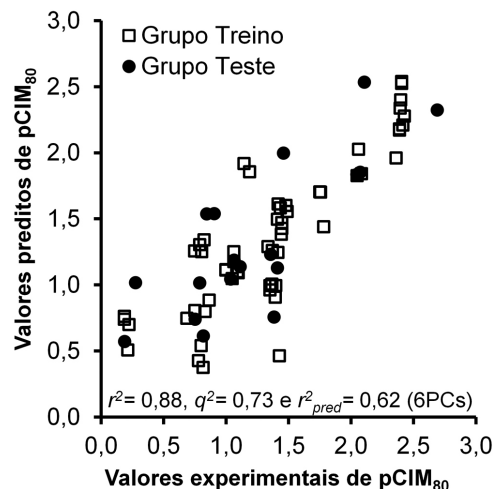
Numa segunda etapa, a influência da variação do tamanho dos fragmentos sobre os parâmetros estatísticos foi analisada (Modelos 10 a 12 - Tabela 2). Contudo, não se observou melhora estatisticamente significativa em nenhum dos modelos gerados.

**Tabela 2.** Influência da distinção de fragmento (DF) para os parâmetros estatísticos dos modelos HQSAR

Modelo	DF	$r^2$	$q^2$	CH	PC	TF
1	ABC	0,85	0,65	97	6	4-7
2	ABCH	0,84	0,59	151	6	4-7
3	ABCCh	0,86	0,59	61	6	4-7
4	ABCHCh	0,86	0,60	61	6	4-7
5	ABCHChDA	0,84	0,67	151	5	4-7
6	ABCChDA	0,86	0,71	71	6	4-7
7	ABDA	0,74	0,67	353	3	4-7
8	ACDA	0,81	0,72	71	4	4-7
<b>9</b>	<b>ABCD</b>	<b>0,88</b>	<b>0,73</b>	<b>71</b>	<b>6</b>	<b>4-7</b>
10	ABCD	0,79	0,65	71	4	2-5
11	ABCD	0,85	0,67	71	6	3-6
12	ABCD	0,87	0,65	71	6	5-8

$r^2$ , coeficiente de correlação;  $q^2$ , coeficiente de correlação por validação cruzada; CH, comprimento do holograma; TF, tamanho do fragmento; PC, número ótimo de componentes principais. Distinções de fragmento (DF): A, átomos; B, ligações; C, conexões; H, átomos de hidrogênios; Ch, quiralidade; DA, doador e acceptor

Embora a consistência interna seja um pré-requisito para bons modelos de QSAR, esse atributo não é suficiente para garantir sua aplicabilidade no planejamento de moléculas mais potentes. De fato, essa característica é avaliada mais qualificadamente através de técnicas de validação externa. Através das mesmas, tenta-se mimetizar uma situação real na qual o modelo de QSAR é empregado para prever a atividade de moléculas não disponíveis para sua calibração.<sup>33</sup> Partindo deste pressuposto, o melhor modelo de HQSAR foi utilizado para prever a atividade de 18 moléculas do grupo teste, não utilizadas durante o desenvolvimento do modelo (Figura 1). Os resíduos encontrados foram inferiores a 0,64, exceto para duas moléculas (**71** -0,69 e **74** -0,74), corroborando com o poder preditivo encontrado para esse modelo ( $r^2_{pred} = 0,62$ ).

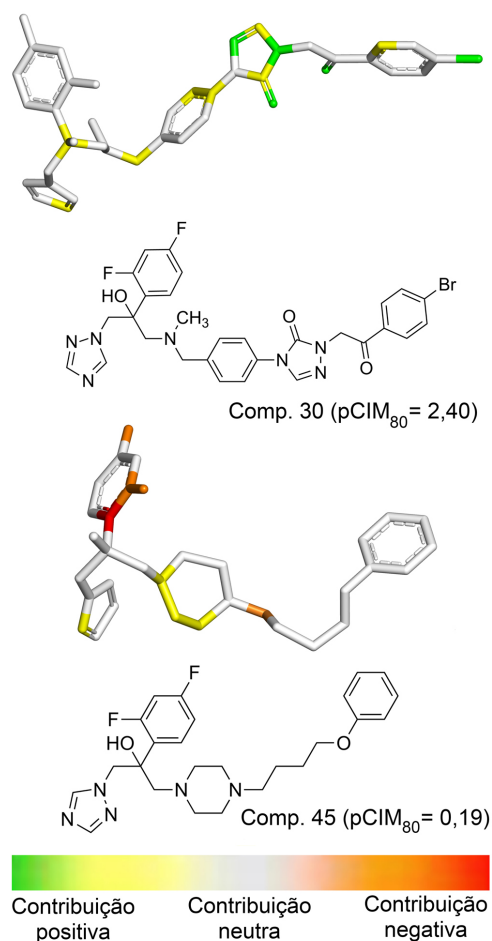


**Figura 1.** Valores de  $pCIM_{80}$  de acordo com modelo de QSAR baseado nos fragmentos

Dessa forma, fica patente que os fragmentos moleculares utilizados como descritores nos modelos de HQSAR estão relacionados com a atividade biológica e podem, portanto, fornecer informações pertinentes acerca das relações entre a estrutura química e a atividade biológica dos derivados azólicos em estudo. Essa informação pode ser extraída de mapas de contribuição, os quais representam graficamente as contribuições dos fragmentos para o modelo final de HQSAR (Figura 2). Nessa representação, um código de cores é usado para discriminar os fragmentos com contribuições negativas para atividade biológica, coloridos em vermelho ou laranja, daqueles com contribuições positivas, que estão representados em amarelo ou verde. Fragmentos comuns a todo o conjunto ou com contribuições intermediárias aparecem na cor cinza.<sup>26</sup>

Dessa forma, derivados azólicos potentes apresentam majoritariamente contribuições positivas para atividade, enquanto derivados com atividade antifúngica limitada têm predominância de fragmentos coloridos em vermelho e/ou laranja (Figura 2). A análise detalhada dos mapas de contribuição sugere que o anel triazolona tem efeito positivo sobre a potência, o que pode ser associado à capacidade de seus átomos atuarem como doadores ou aceptores de ligação de hidrogênio, corroborando a importância dessa distinção de fragmento sobre a qualidade estatística do melhor modelo de HQSAR. O anel fenila, substituído na posição *para* com um átomo de bromo, também apresenta contribuições positivas nas posições *para* e *orto* do anel. Sheng e colaboradores<sup>1</sup> também identificaram, através de modelos de QSAR 3D (CoMFA e CoMSIA) para uma série análoga de derivados azólicos, que esse padrão de substituição no anel resulta em moléculas com maior atividade antifúngica.

Embora, o modelo de HQSAR permita verificar visualmente os



**Figura 2.** Mapa de contribuição para derivados triazólicos ativos contra *C. albicans*

fragmentos que contribuem positiva ou negativamente para a atividade biológica, não é possível identificar os motivos pelos quais esses fragmentos foram destacados.

A fim de investigar essa questão, modelos de QSAR 2D baseados em descritores moleculares foram desenvolvidos. Dentre os cerca de 2500 descritores 2D disponíveis no programa DRAGON 5.5, 337 apresentaram variância interna ( $>0,03$ ) e correlação mínima com a atividade biológica ( $r^2 > 0,10$ ) para serem úteis no desenvolvimento de modelos de QSAR. Como apenas uma pequena parcela desses descritores deve possuir correlação com a atividade biológica,<sup>34,35</sup> eles foram empregados para construir modelos preliminares de QSAR através de regressão linear múltipla com até 4 descritores, com auxílio do programa MOBYDIGS 1.0 (ver Parte Experimental). Os parâmetros estatísticos encontrados ( $r^2 < 0,55$  e  $q^2 < 0,47$ ) foram considerados insuficientes para permitir uma análise segura das relações entre a estrutura química e a atividade biológica dessa série de compostos, razão pela qual se decidiu empregar técnicas estatísticas mais robustas. Dessa forma, 25 descritores (Tabela 3) encontrados nos 24 melhores modelos de RLM ( $q^2 > 0,40$ ) foram agrupados, autoescalados e, depois, utilizados para a construção de modelos de QSAR através de regressão por mínimos quadrados parciais, como disponível no programa PIROUETTE 4.0.

Durante o desenvolvimento do modelo, foram eliminados progressivamente os descritores com contribuição desprezível para o vetor de regressão,<sup>32,36,37</sup> de forma que o modelo final contém apenas 15 das 25 variáveis originais (em destaque na Tabela 3). O melhor modelo apresenta ajuste e consistência interna comparáveis aos do

**Tabela 3.** Variáveis selecionadas pelo MOBYDIGS 1.0 (variáveis em destaque pertencem ao modelo final de QSAR baseado em descritores)

Descritor	Significado
<b>GATS4m</b>	Autocorrelação de Geary para uma distância topológica de 4, ajustada pelas massas atômicas
<b>GATS6m</b>	Autocorrelação de Geary para uma distância topológica de 6, ajustada pelas massas atômicas
GATS8e	Autocorrelação de Geary para uma distância topológica de 8, ajustada pelas eletronegatividades atômicas de Sanderson
GATS8v	Autocorrelação de Geary para uma distância topológica de 8, ajustada pelos volumes atômicos de van der Walls
<b>EEig05x</b>	Autovalor de corte 05 da matriz de corte adjacente
<b>EEig07d</b>	Autovalor de corte 07 da matriz de corte adjacente, ajustado pelos momentos dipolo
<b>BELv1</b>	Menor autovalor nº 1 de uma matriz de carga, ajustado pelos volumes de van der Walls
B08[O-I]	Existência ou não de ligação O-I na distância topológica 8
<b>BEHm6</b>	Maior autovalor de nº 6 de uma matriz de carga, ajustado pelas massas atômicas
<b>MLOGP</b>	Coefficiente de partição octanol/água de Moriguchi
<b>MATS8e</b>	Autocorrelação de Moran para uma distância topológica de 8, ajustada pelas eletronegatividades atômicas de Sanderson
BLFT96	Modelo de toxicidade de peixes (96 h) de Verharr baseado no MLOGP (mmol/L)
F01[C-O]	Frequência de C-O em distância topológica 1
<b>F08[C-N]</b>	Frequência de C-N em distância topológica 8
F08[N-F]	Frequência de N-F em distância topológica 8
<b>F10[C-C]</b>	Frequência de C-C em distância topológica 10
F10[C-F]	Frequência de C-F em distância topológica 10
<b>T(O..Br)</b>	Soma das distâncias topológicas entre O..Br
<b>T(N..F)</b>	Soma das distâncias topológicas entre N..F
<b>nX</b>	Número de átomos de halogênio
BELm8	Menor autovalor de nº 8 da matriz de carga ajustada pelas massas atômicas
X1Av	Índice médio de conectividade de valência ch-1
<b>MPC02</b>	Índice topológico
ARR	Número de anéis aromáticos
<b>nR05</b>	Número de anéis de cinco membros

melhor modelo de HQSAR ( $r^2 = 0,82$  e  $q^2(\text{LOO}) = 0,79$  com 2 PCs) (Figura 3), porém com capacidade preditiva superior ( $r^2_{\text{pred}} = 0,70$ ), sugerindo que descritores topológicos são úteis não só para compreensão das relações entre a estrutura química e a atividade biológica dos compostos do grupo treino, mas também de moléculas congêneres não utilizadas na calibração dos modelos de QSAR.

Embora diversos descritores topológicos não tenham significado químico aparente, a interpretação dos descritores com maior contribuição para o modelo segundo o vetor de regressão (Figura 4) pode, em alguns casos, auxiliar na melhor compreensão dos fatores que são relevantes para a atividade biológica. Por exemplo, a maioria dos descritores que tem contribuição negativa (GATS4m, GATS6m e BELv1) está relacionada a fatores estéreos (ex. impedimentos estéricos), enquanto descritores relacionados com lipofilicidade (nX e MLOGP) têm contribuição positiva (Figura 4).<sup>38,39</sup> É importante ressaltar que outros fatores além da interação ligante-macromolécula influenciam os valores de CIM<sub>80</sub>, portanto, é razoável supor que a contribuição positiva observada no gráfico de vetor de regressão reflita

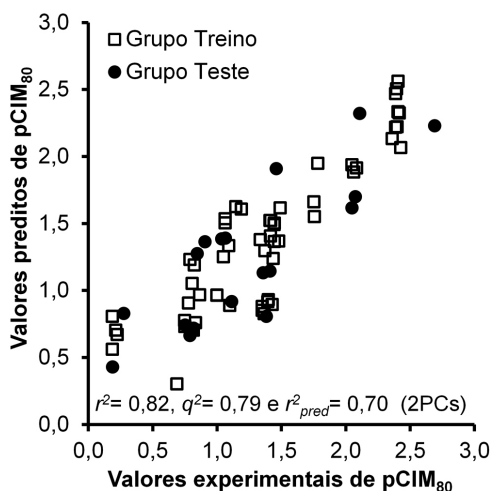


Figura 3. Valores preditos e experimentais de pCIM<sub>80</sub>, de acordo com o melhor modelo de QSAR baseado em descritores

também a capacidade dessas moléculas se partitionarem entre a fase aquosa e o ambiente lipofílico encontrado no sítio ativo da enzima lanosterol 14- $\alpha$  desmetilase de *C. albicans*.

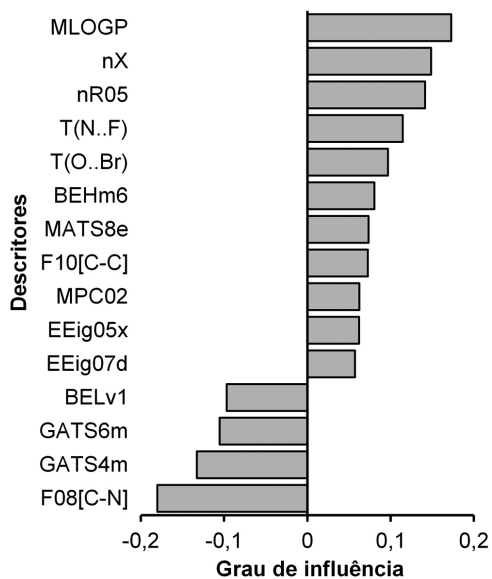


Figura 4. Análise do vetor de regressão

Outra característica importante para atividade antifúngica é a presença de anéis de 5 membros (nR05), representados nesse conjunto de dados pelos anéis triazolona ou triazol, este último comum a todas as moléculas do conjunto de dados. Por outro lado, moléculas potentes (pCIM<sub>80</sub> > 2,0) apresentam anel triazolona, enquanto o mesmo não é observado para moléculas com baixa atividade antifúngica (pCIM<sub>80</sub> < 0,99). Essa informação corrobora a contribuição positiva desse fragmento, observada nos mapas de contribuição de HQSAR (Figura 2) e está de acordo com resultados de acoplamento molecular, segundo os quais o anel triazolona interage, através de uma ligação de hidrogênio, com resíduos do canal de entrada do fármaco, permitindo a ampliação do espectro de ação destes compostos.<sup>40</sup>

## CONCLUSÃO

Os modelos de QSAR 2D baseados em fragmentos e descritores apresentam bom ajuste e consistência interna ( $r^2 = 0,88$  e  $q^2 = 0,73$ ),

contudo sua capacidade preditiva ( $r^2_{pred} = 0,62$ ) parece ser insuficiente para guiar o planejamento de moléculas mais potentes. Por outro lado, a inclusão de descritores topológicos, relacionados às propriedades físico-químicas, produz modelos com qualidade estatística equivalente ( $r^2 = 0,82$  e  $q^2 = 0,79$ ) e habilidade para prever a atividade biológica de moléculas congêneras de forma satisfatória ( $r^2_{pred} = 0,70$ ). Apesar das dificuldades encontradas na interpretação química de descritores topológicos, a análise dos descritores mais importantes para o modelo de QSAR baseado em descritores, segundo o gráfico do vetor de regressão (Figura 4), indica que a lipofilicidade das moléculas avaliadas é uma característica fundamental para sua atividade biológica. Paralelamente a isso, ambos os modelos sugerem que o anel triazolona contribui positivamente para potência, provavelmente por realizar ligações de hidrogênio com resíduos da enzima lanosterol 14 $\alpha$ -desmetilase. Portanto, os modelos de QSAR apresentados neste trabalho parecem ser úteis tanto para racionalizar as exigências moleculares necessárias para atividade antifúngica, como para guiar o planejamento de moléculas mais potentes contra *C. albicans*.

## REFERÊNCIAS

- Sheng, C.; Zhang, W.; Ji, H.; Zhang, M.; Song, Y.; Xu, H.; Zhu, J.; Miao, Z.; Jiang, Q.; Yao, J.; Zhou, Y.; Lu, J.; *J. Med. Chem.* **2006**, *49*, 2512.
- Mavor, A. L.; Thewes, S.; Hube, B.; *Curr. Drug Targets* **2005**, *6*, 863.
- Lopez, M.; *Clin. Dermatol.* **2010**, *28*, 178.
- Trabulsi, L. R.; Alterthum, F.; *Microbiologia*, Atheneu Editora: São Paulo, 2008.
- Moran, C.; Grussemeyer, C. A.; Spalding, J. R.; Benjamin, D. K., Jr.; Reed, S. D.; *Am. J. infec. Control.* **2010**, *38*, 78.
- Aoyama, Y.; Yoshida, Y.; Sato, R.; *J. Biolog. Chem.* **1984**, *259*, 1661.
- Casalinuovo, I. A.; Francesco, P. D.; Garaci, E.; *Eur. Rev. Med. Pharmacol. Sci.* **2004**, *8*, 69.
- Morio, F.; Loge, C.; Besse, B.; Hennequin, C.; Le Pape, P.; *Diagn. Microbiol. Infect. Dis.* **2010**, *66*, 373.
- Marichal, P.; Koymans, L.; Willemsens, S.; Bellens, D.; Verhasselt, P.; Luyten, W.; Borgers, M.; Ramaekers, F. C. S.; Odds, F. C.; Vanden Bossche, H.; *Microbiology* **1999**, *145*, 2701.
- Sanglard, D.; *Enferm. Infec. Microbiol. Clin.* **2002**, *20*, 462; quiz 470, 479.
- Cannon, R. D.; Lamping, E.; Holmes, A. R.; Niimi, K.; Baret, P. V.; Keniya, M. V.; Tanabe, K.; Niimi, M.; Goffeau, A.; Monk, B. C.; *Clin. Microbiol. Rev.* **2009**, *22*, 291.
- Pina-Vaz, C.; Rodrigues, A. G.; Costa-de-Oliveira, S.; Ricardo, E.; Mårdh, P. A.; *J. Antimicrob. Chemother.* **2005**, *56*, 678.
- Davis, A. M.; Teague, S. J.; Kleywegt, G. J.; *Angew. Chem., Int. Ed.* **2003**, *42*, 2718.
- Ji, H.; Zhang, W.; Zhou, Y.; Zhang, M.; Zhu, J.; Song, Y.; Lü, J.; Zhu, J.; *J. Med. Chem.* **2000**, *43*, 2493.
- Santos Filho, O. A.; Alencastro, R. B.; *Quim. Nova* **2003**, *26*, 253.
- Jennings, A. J.; Edge, C. M.; Sternberg, M. J. E.; *Protein Eng.* **2001**, *14*, 227.
- Blake, J. D.; Cohen, F. E.; *J. Mol. Biol.* **2001**, *307*, 721.
- Chai, X.; Zhang, J.; Cao, Y.; Zou, Y.; Wu, Q.; Zhang, D.; Jiang, Y.; Sun, Q.; *Bioorg. Med. Chem. Lett.* **2011**, *21*, 686.
- Guan, Z.; Chai, X.; Yu, S.; Hu, H.; Jiang, Y.; Meng, Q.; Wu, Q.; *Chem. Biol. Drug. Des.* **2010**, *76*, 496.
- Chen, Q.; Zhu, X. L.; Jiang, L. L.; Liu, Z. M.; Yang, G. F.; *Eur. J. Med. Chem.* **2008**, *43*, 595.
- Singla, R. K.; Bhat, G. V.; *J. Enzyme Inhib. Med. Chem.* **2010**, *25*, 696.
- Wang, W.; Sheng, C.; Che, X.; Ji, H.; Cao, Y.; Miao, Z.; Yao, J.; Zhang, W.; *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5965.
- Che, X.; Sheng, C.; Wang, W.; Cao, Y.; Xu, Y.; Ji, H.; Dong, G.; Miao, Z.; Yao, J.; Zhang, W.; *Eur. J. Med. Chem.* **2009**, *44*, 4218.

24. CLSI - Clinical and Laboratory Standards Institute; *Reference Method for Broth Dilution Antifungal Susceptibility Testing of Yeasts*, Approved Standard, 2<sup>nd</sup> ed., Document M27-A2, CLSI: Pennsylvania, 2002.
25. Lowis, D. R.; *Tripos Technical Notes* **1997**, 1, 17.
26. Andrade, C. H.; Salum, L. D. E. B.; Castilho, M. S.; Pasqualoto, K. F.; Ferreira, E. I.; Andricopulo, A. D.; *Mol. Diversity* **2008**, 12, 47.
27. Mobydigs; *Mobydigs 1.0*; Talete srl, Milano, Italy, 2004.
28. Todeschini, R.; Consonni, V.; Maioiocchi, A.; *Chemom. Intell. Lab. Syst.* **1999**, 46, 13.
29. Todeschini, R.; Consonni, V.; Mauri, A.; Pavan, M.; *Anal. Chim. Acta* **2004**, 515, 199.
30. Schuurmann, G.; Ebert, R. U.; Chen, J.; Wang, B.; Kuhne, R.; *J. Chem. Inf. Model.* **2008**, 48, 2140
31. Doddareddy, M. R.; Cho, Y. S.; Koh, H. Y.; Pae, A. N.; *Bioorg. Med. Chem.* **2004**, 12, 3977.
32. Mota, S. G. R.; Barros, T. F.; Castilho, M. S.; *J. Braz. Chem. Soc.* **2009**, 20, 451.
33. Golbraikh, A.; Tropsha, A.; *J. Mol. Graph. Modell.* **2002**, 20, 269.
34. Gonzalez, M. P.; Teran, C.; Saiz-urra, L.; Teixeira, M.; *Curr. Top. Med. Chem.* **2008**, 8, 1606.
35. Gao, H.; Lajiness, M. S.; van Drie, J.; *J. Mol. Graph. Modell.* **2002**, 20, 259.
36. Guido, R.; Castilho, M.; Mota, S.; Oliva, G.; Andricopulo, A.; *QSAR Comb. Sci.* **2008**, 27, 768.
37. Freitas, H. F.; Paz, O. S.; Castilho, M. S.; *Quim. Nova* **2009**, 32, 2114.
38. Helguera, A. M.; Combes, R. D.; Gonzalez, M. P.; Cordeiro, M. N.; *Curr. Top. Med. Chem.* **2008**, 8, 1628.
39. Mazzatorta, P.; Estevez, M. D.; Coulet, M.; Schilter, B.; *J. Chem. Inf. Model.* **2008**, 48, 1949.
40. Yongwei, J.; Yongbin, C.; Jun, Z.; Yan, Z.; Xiaoyun, C.; Honggang, H.; Qingjie, Z.; Qiuye, W.; Dazhi, Z.; Yuanying, J.; Qingyan, S.; *Eur J Med Chem.* **2011**, 46, 3135.