



UNIVERSIDADE FEDERAL DA BAHIA
FACULDADE DE FARMÁCIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS
FARMACÊUTICAS



HUMBERTO FONSECA DE FREITAS

DISSERTAÇÃO DE MESTRADO

ESTUDOS DE QSAR 2D PARA UMA SÉRIE DE DERIVADOS
AZÓLICOS ATIVOS CONTRA *Cryptococcus* E DE
MECANISMOS DE RESISTÊNCIA DE *Cryptococcus gattii*
ATRAVÉS DE MODELAGEM POR HOMOLOGIA

SALVADOR - BAHIA

2011

HUMBERTO FONSECA DE FREITAS

**ESTUDOS DE QSAR 2D PARA UMA SÉRIE DE DERIVADOS AZÓLICOS ATIVOS
CONTRA *Cryptococcus* E DE MECANISMOS DE RESISTÊNCIA DE *Cryptococcus
gattii* ATRAVÉS DE MODELAGEM POR HOMOLOGIA**

Dissertação apresentada junto ao programa de Pós-Graduação em Farmácia da Faculdade de Farmácia – UFBA, como requisito parcial para a obtenção do título de Mestre em Farmácia.

Orientador: Prof. Dr. Marcelo Santos Castilho
Co-orientadora: Profa. Dra. Tânia Fraga Barros

**Salvador
2011**

Sistema de Bibliotecas - UFBA

Freitas, Humberto Fonseca de.

Estudos de QSAR 2D para uma série de derivados azólicos ativos contra *Cryptococcus* e de mecanismos de resistência de *Cryptococcus gattii* através de modelagem por homologia / Humberto Fonseca de Freitas. - 2011.

146 f. : il.

Inclui apêndices.

Orientador: Prof. Dr. Marcelo Santos Castilho.

Co-orientadora: Prof^a. Dr^a. Tânia Fraga Barros.

Dissertação (mestrado) - Universidade Federal da Bahia, Faculdade de Farmácia, Salvador, 2011.

1. Fungos. 2. *Cryptococcus* spp. 3. Farmacologia. 4. Farmacologia - Planejamento.
I. Castilho, Marcelo Santos. II. Barros, Tânia Fraga. III. Universidade Federal da Bahia. Faculdade de Farmácia. IV. Título.

CDD - 615.1

CDU - 615.1



UNIVERSIDADE FEDERAL DA BAHIA
FACULDADE DE FARMÁCIA
PROGRAMA DE PÓS-GRADUAÇÃO EM FARMÁCIA



TERMO DE APROVAÇÃO

HUMBERTO FONSECA DE FREITAS

**ESTUDOS DE QSAR 2D PARA UMA SÉRIE DE DERIVADOS AZÓLICOS ATIVOS
CONTRA Cryptococcus E DE MECANISMOS DE RESISTÊNCIA DE Cryptococcus gattii
ATRAVÉS DE MODELAGEM POR HOMOLOGIA**

Dissertação submetida ao Programa de Pós-Graduação em Farmácia (nível Mestrado Acadêmico) da Faculdade de Farmácia da Universidade Federal da Bahia, como requisito parcial para a obtenção do título de Mestre em Farmácia.

Aprovada em 08 de julho de 2011.

BANCA EXAMINADORA

Marcelo Santos Castilho
UFBA
Orientador

Kátia Maria Honório
USP

Carlos Priminho Pirovani
UESC

Aos meus pais José Humberto e Ester, à minha irmã Gabriela e à Camilla pelo amor, compreensão, todo o incentivo e, especialmente, pela lição de vida que foram fundamentais para me tornar uma pessoa mais feliz e dedicada à minha vida e à minha carreira.

AGRADECIMENTOS

Agradeço inicialmente o meu orientador e amigo, Prof. Dr. Marcelo Santos Castilho, pela oportunidade de realizar este trabalho juntos, pela valiosa chance de aprender e crescer com suas idéias, pela dedicação, sabedoria e, principalmente, confiança no meu trabalho.

Agradeço também a Profa. Dra. Tânia Fraga Barros, minha co-orientadora e amiga, pela atenção, amizade, dedicação, pela contribuição fundamental para este trabalho e por todos os conselhos que sempre fiz questão de ouvir.

Aos meus orientadores, mais um agradecimento pela oportunidade da convivência com pessoas inteligentes, honradas e companheiras, contribuindo assim para a melhor formação de seus alunos como pessoas e profissionais.

Ao prof. Dr. Eudes da Silva Veloso, pelo apoio e pelas descontraídas conversas que sempre agregavam sabedoria e ensinamentos.

Agradeço à Pós-graduação em Ciências Farmacêuticas (PPGFAR), representada pelo coordenador Prof. Dr. Ajax Mercês Atta, pela oportunidade e contribuição deste curso na minha formação e crescimento profissional.

Aos meus familiares e amigos pela atenção, incentivo nos momentos difíceis e pelo companheirismo.

Aos meus colegas e amigos do LaBiMM e LPMC pelos momentos de descontração, aprendizado e ótimo ambiente de trabalho que pude vivenciar.

A todas as pessoas que direta ou indiretamente me auxiliaram para o desenvolvimento deste trabalho.

A FAPESB pelo suporte financeiro essencial para a realização deste mestrado.

RESUMO

Apesar dos avanços no desenvolvimento de fármacos antifúngicos, observa-se um crescimento de cepas resistentes aos fármacos de escolha (derivados azólicos). Diante disso, torna-se imprescindível racionalizar as propriedades físico-químicas e estruturais que governam a atividade antifúngica de derivados azólicos, bem como investigar os mecanismos moleculares responsáveis pela resistência. Visando elucidar algumas dessas questões os objetivos desse trabalho são: 1) Determinar a atividade biológica de 33 derivados azólicos, frente à *Cryptococcus gattii*, a fim de utilizá-la como variável dependente na construção de modelos classificatórios (KNN e SIMCA) e quantitativos (HQSAR e QSAR 2D); 2) Investigar, do ponto de vista estrutural, mutações pontuais no gene ERG11 que conferem resistência a derivados azólicos.

Os modelos classificatórios (KNN e SIMCA) mostram consistência interna razoável ao acertar a classificação de 60% dos compostos ativos e 75% dos compostos inativos do grupo teste. Adicionalmente, o modelo SIMCA sugere que a transferência de carga entre átomos distantes de 10 ligações (JGI10) é importante para diferenciar moléculas ativas das inativas contra *C. gattii*. Os modelos quantitativos baseados em fragmentos moleculares (hogramas QSAR) apresentam bom ajuste ($r^2=0,85$), porém baixo poder preditivo ($r^2_{pred}=0,38$), enquanto o melhor modelo de QSAR 2D baseado em descritores topológicos apresenta resultados estatísticos elevados ($r^2=0,95$ e $q^2=0,86$) com 3 PCs, além de bom poder preditivo ($r^2_{pred}=0,72$). Este modelo indica ainda que o aumento da potência é influenciado pela menor distribuição de carga de uma molécula na distância de dois átomos (GGI1) e pela menor diferença de eletronegatividade dos átomos na distância topológica 1 e 2 (GATS1e e MATS2e).

Os modelos comparativos de lanosterol 14 α -desmetilase de *C. gattii*, nativa e com mutações pontuais (G484S, H488Y e K156R), mostram que alterações no posicionamento do grupo heme e redução do número de interações entre a enzima e o grupo prostético, desestabilizam a interação entre o ferro do grupo heme e o nitrogênio do anel imidazólico ou triazólico, e modificações na flexibilidade conformacional do alvo terapêutico, reduzindo assim o acesso do antifúngico ao sítio ativo da enzima, são os principais mecanismos pelos quais as mutações reduzem a afinidade dos derivados azólicos pelo alvo terapêutico.

Palavras-chave: *Cryptococcus*. Derivados azólicos. Modelagem comparativa. QSAR 2D. Resistência.

ABSTRACT

Despite advances in the development of antifungal drugs, there is an upsurge of resistant strains against the drugs of choice (azole derivatives). Hence, it becomes essential to comprehend the physicochemical and structural properties that rule the antifungal activity of azole's derivatives, as well as to investigate the molecular mechanisms responsible for fungal resistance. Aiming at elucidating some of these questions the objectives of this study are: 1) To determine the biological activity of 33 azole derivatives against *Cryptococcus gattii*, in order to use it as a dependent variable in the development of classification (KNN and SIMCA) and quantitative (HQSAR and 2D QSAR) models; 2) To investigate, from a structural stand-point of view, mutations in the ERG11 gene which render *C. gattii* resistant to azole's derivatives.

Classification models (KNN and SIMCA) show reasonable internal consistency, correctly classifying 60% of the active compounds and 75% of inactive compounds from the test group. Additionally, the SIMCA model suggests that the charge transfer between atoms 10 bonds apart (JGI10) is important to differentiate azole derivatives that are either active or inactive against *C. gattii*. The quantitative models based on molecular fragments (hologram QSAR) shows good fit ($r^2 = 0.85$), but low predictive power ($r^2_{pred} = 0.38$), whereas the best 2D QSAR model, built from topological descriptors, shows outstanding statistical results ($r^2 = 0.95$ and $q^2 = 0.86$) with 3 PCs, along with good predictive power ($r^2_{pred} = 0.72$). This model also indicates that potency boost is influenced by reduced charge distribution between two atoms (GGI1) and by the lower electronegativity difference at the topological distance 1 and 2 (GATS1e and MATS2e).

Comparative models of native and mutated (G484S, H488Y and K156R) lanosterol 14 α -demethylases of *C. gattii* show that changes in the positioning of the heme group and reduced binding of the enzyme towards the prosthetic group, disrupt the interaction profile between the heme iron and the nitrogen of imidazole or triazole ring, as well as changes in the therapeutic target conformational flexibility, which reduces the access of the antifungal agent to the enzyme active site, are the main mechanisms by which affinity of azole derivatives is reduced.

Key-words: *Cryptococcus*. Azole derivatives. Comparative modeling. 2D QSAR. Resistance

LISTA DE ILUSTRAÇÕES

FIGURA 1.1 – Esquema comparativo das técnicas de SBDD e LBDD.....	Pag. 18
FIGURA 1.2 – Tipos de descritor molecular, onde “A” representa os átomos. (Adaptado de Willett e colaboradores, 1998).....	Pag. 19
FIGURA 1.3 – Principais antifúngicos disponíveis para uso clínico: Anfotericina (1); 5-flucitosina (2); Terbinafina (3); Itraconazol (4); Voriconazol (5); Posaconazol (6) e Fluconazol (7).....	Pag. 22
FIGURA 2.1 – Representação do método KNN para classificação de uma nova amostra considerando 1 vizinho (A) e 3 vizinhos (B).....	Pag. 27
FIGURA 2.2 – Projeção das componentes principais construídas em ordem decrescente da quantidade de variância (à esquerda) e das moléculas nas novas coordenadas definidas pelas PCs (à direita). Adaptado a partir de Waterbeemd e Rose (2003) e Sena e colaboradores (2000) (Em que, v_j ($j=1, 2, 3...m$) são as variáveis originais relacionadas à cada PC, x_i representa as amostras projetadas nas novas coordenadas $z_{i,1}$ e $z_{i,2}$ definidas pelas PCs).....	Pag. 28
FIGURA 2.3 – Esquema ilustrativo do processo de geração dos hologramas moleculares.....	Pag. 30
FIGURA 2.4 – Etapas envolvidas na geração de um modelo de HQSAR.....	Pag. 32
FIGURA 2.5 – Halos de inibição para o antifúngico Fluconazol em relação à <i>C. gattii</i> (A) e <i>C. neoformans</i> (B e C). Os símbolos I, II e III representam as concentrações de antifúngicos contidas nos discos: 16,4 mmol/L; 8,2 mmol/L e 4,1 mmol/L, respectivamente. O símbolo “C” representa o controle negativo do experimento contendo apenas DMSO, o solvente utilizado para a dissolução dos antifúngicos.....	Pag. 48
FIGURA 2.6 – Halos de inibição para o antifúngico Flutrimazol (A) e Hexaconazol (B) frente à <i>C. gattii</i> . Os símbolos I, II e III representam as concentrações de antifúngicos contidas nos discos: 16,4 mmol/L; 8,2 mmol/L e 4,1 mmol/L, respectivamente. O símbolo “C” representa o controle negativo do experimento contendo apenas DMSO, o solvente utilizado para a dissolução dos antifúngicos.....	Pag. 49
FIGURA 2.7 – Histograma de frequência dos antifúngicos em relação à variação no diâmetro de halo de inibição frente a <i>C. gattii</i> na menor concentração (4,1 mmol/L).....	Pag. 50
FIGURA 2.8 – Número de compostos classificados erroneamente para cada número de vizinhos utilizados para classificação.....	Pag. 52

FIGURA 2.9 – Coeficientes de poder discriminantes para os descritores presentes no modelo de SIMCA inicial. Destaque para os descritores que fizeram parte do modelo final (barras em azul).....	Pag. 54
FIGURA 2.10 – Coeficientes de poder de modelagem para os descritores presentes no modelo inicial de SIMCA. Destaque para os descritores que fizeram parte do modelo final (barras em azul).....	Pag. 55
FIGURA 2.11 – Distância interclasse dos compostos ativos e inativos dos grupos treino e teste segundo o modelo final de SIMCA.....	Pag. 56
FIGURA 2.12 – Perfil de atividade versus coeficientes de BEHe3 para os antifúngicos do grupo treino e teste.....	Pag. 57
FIGURA 2.13 – Alinhamento sequencial dos aminoácidos da enzima Lanosterol 14 α -desmetilase de fungos leveduriformes (<i>C. neoformans</i> , <i>C. gattii</i> , <i>C. albicans</i> e <i>C. tropicalis</i>) e fungos filamentosos (<i>M. perniciosus</i> , <i>A. fumigatus</i> e <i>S. reilianum</i>). Resíduos conservados em todas as seqüências estão destacados em cinza e resíduos pertencentes ao sítio ativo em azul. Alinhamento feito no CLUSTAL X 2.1 e análise da identidade e similaridade no BIOEDIT 7.05, através da matriz BLOSUM62.....	Pag. 59
FIGURA 2.14 – Concentração inibitória mínima do antifúngico Fluconazol frente a <i>C. neoformans</i> (4 μ g/mL) e <i>C. gattii</i> (32 μ g/mL).....	Pag. 60
FIGURA 2.15 – Histograma de frequência dos antifúngicos dos grupos treino e teste em relação à faixa de pCIM (ν).....	Pag. 64
FIGURA 2.16 – Valores experimentais versus valores preditos de pCIM para os 33 antifúngicos.....	Pag. 67
FIGURA 2.17 – Valores experimentais e preditos de pCIM para 33 antifúngicos de acordo com o melhor modelo de QSAR 2D baseado em descritores topológicos.....	Pag. 71
FIGURA 2.18 – Coeficiente do vetor de regressão para as variáveis do modelo final de QSAR 2D baseado em descritores topológicos.....	Pag. 72
FIGURA 2.19 – Coeficientes de GGII versus valores de CIM para 33 antifúngicos.....	Pag. 73
FIGURA 3.1 – Processo de construção de um modelo comparativo. Adaptado de Cavasotto e Phatak (2009).....	Pag. 77
FIGURA 3.2 – Representação da construção de modelos comparativos por restrições espaciais. Após o alinhamento das seqüências alvo e molde ocorre a extração das restrições espaciais da proteína molde que serão aplicadas sobre a seqüência da proteína alvo para gerar o modelo tridimensional. Adaptado de Sanchez e Sali. (2000).....	Pag. 79

FIGURA 3.3 – Representação gráfica dos ângulos phi e psi através do gráfico de Ramachandran mostrando as regiões favoráveis (vermelho) e permitidas (amarelo). Adaptado de Laskowski e colaboradores (1993).....	Pag. 80
FIGURA 3.4 – O VERIFY 3D deriva 6 ambientes químicos (E, P ₁ , P ₂ , B ₁ , B ₂ , B ₃) a partir do valores da área enterrada na proteína e da fração da cadeia lateral coberta por átomos polares (A). Um grupo de resíduos é, então, avaliado quanto a probabilidade de ser encontrado nesses ambientes químicos e representados de forma gráfica de pontuação 3D-1D (B). As categorias P ₁ e P ₂ sugerem regiões de ambientes polares da proteína; B ₁ , B ₂ , B ₃ sugerem regiões enterradas da proteína; e E, regiões exposta ao solvente. Adaptado de Bowie e colaboradores (1991).....	Pag. 81
FIGURA 3.5 – Eletroforese em gel de agarose 0,8%. (a) Padrão de peso molecular Lambda DNA/HindIII (4µL); (b) 2 µL de produto da amplificação do gene ERG11 de <i>C. gattii</i> ; (c) 2 µL de produto purificado da amplificação do gene ERG11 de <i>C. gattii</i>	Pag. 94
FIGURA 3.6 – Alinhamento sequencial da sequência genômica da ERG11 de <i>C. gattii</i> (Seq. <i>C. gattii</i>) contra as sequências obtidas após o sequenciamento com destaque às regiões conservadas (em cinza), o primer verso (em verde) e a sequência complementar do primer reversos (em azul).....	Pag. 95-96
FIGURA 3.7 – Alinhamento das sequências das proteínas moldes (2CIB e 3LD6) contra a sequência alvo (<i>C. gattii</i>) destacando as estruturas secundárias hélices alfa (vermelho) e fitas beta (azul) preditas pelo programa PSIPRED 3.0 (http://bioinf.cs.ucl.ac.uk/psipred/). As regiões sublinhadas tiveram suas sequências ajustadas manualmente.....	Pag. 99
FIGURA 3.8 – Alinhamento das sequências das proteínas moldes (2CIB e 3LD6) contra a sequência alvo (<i>C. gattii</i>) utilizado na construção dos modelos por homologia. Em destaque podem ser observados os resíduos conservados ao longo da evolução.....	Pag. 100
FIGURA 3.9 – Gráfico de Ramachandran para os modelos da enzima nativa antes da minimização (A) e após a minimização (B), e para as proteínas moldes 2CIB (C), 3LD6 (D). A porcentagem de resíduos nas regiões permitidas foi: 99,9% para as proteínas moldes (2CIB e 3LD6) e 98,2% para os modelos da enzima nativa antes e após a minimização. As regiões favoráveis estão destacadas em vermelho, as regiões permitidas em amarelo.....	Pag. 102
FIGURA 3.10 – Resíduos do modelo da enzima nativa encontrados na região não permitida do gráfico de Ramachandran (em amarelo) e o grupamento heme (em ciano).....	Pag. 103
FIGURA 3.11 – Gráfico de pontuação média 3D-1D, considerando uma janela de 21 aminoácidos, segundo o servidor VERIFY 3D, do modelo comparativo da enzima nativa.....	Pag. 104

FIGURA 3.12 – Modelo da Lanosterol 14 α -desmetilase de <i>C. gattii</i> nativa (em azul) sobreposta às estruturas moldes: 2CIB (em cinza) e 3LD6 (em verde). Em destaque (vermelho) o fragmento que apresenta baixa pontuação 3D-1D no servidor VERIFY 3D.....	Pag. 105
FIGURA 3.13 – Gráfico de pontuação média 3D-1D das proteína moldes (2CIB e 3LD6). Os resíduos 195 à 199 em 2CIB e 191 a 195 em 3LD6 correspondem a região delimitada entre os resíduos 221 e 224 no modelo da enzima nativa, vide FIGURA 3.11.....	Pag. 105
FIGURA 3.14 – Duas orientações do sítio ativo de lanosterol 14 α -desmetilase de <i>C. gattii</i> onde se destacam (amarelo) resíduos apolares (A) e polares (B), que interagem com o grupo heme (ciano).....	Pag. 106
FIGURA 3.15 – Gráfico de Ramachandran do modelo comparativo da enzima contendo a mutação G484S(A), resíduos destacados na região não permitida do gráfico de Ramachandran (B) e gráfico de pontuação 3D-1D média do VERIFY 3D (C).....	Pag. 107
FIGURA 3.16. – Choque estérico entre o ácido propanóico do grupo heme da enzima nativa (azul) e a serina (SER-484). O grupo heme da enzima mutante está na coloração cinza.....	Pag. 108
FIGURA 3.17 – Ângulo entre os átomos do grupo prostético da enzima nativa (azul) e da enzima contendo a mutação G484S (em cinza).....	Pag. 108
FIGURA 3.18 – Gráfico de Ramachandran (A) e modelo 3D da enzima com a mutação H488Y destacando os resíduos encontrados em regiões não permitidas (resíduos em amarelo) (B) e a região (em vermelho) de baixa pontuação 3D-1D do VERIFY 3D (C).....	Pag. 110
FIGURA 3.19 – Substituição da histidina (HIS-488) pela tirosina (TYR-488) causa perda da interação iônica com o ácido propanóico do grupo heme.....	Pag. 111
FIGURA 3.20 – Gráfico de Ramachandran (A) e estrutura 3D do modelo da enzima contendo a mutação K156R destacando a região de baixa pontuação 3D-1D do programa VERIFY 3D (alça em vermelho) (C) e os resíduos encontrados em regiões não permitidas (resíduos em amarelo) (B).....	Pag. 112
FIGURA 3.21 – Região de dobradiça (apontado pela seta) do modelo comparativo da enzima lanosterol 14 α -desmetilase nativa (A) que muda de posição no modelo da enzima com a mutação K156R (B).....	Pag. 113
FIGURA 3.22 – Modelo da proteína lanosterol 14 α -desmetilase de <i>C. gattii</i> destacando os resíduos que compõem a região de acesso do fármaco na proteína.....	Pag. 113

LISTA DE TABELAS

TABELA 2.1 – Parâmetros de distinção de fragmento no método HQSAR.....	Pag. 31
TABELA 2.2 – Códigos de derivados imidazólicos e triazólicos comerciais.....	Pag. 35
TABELA 2.3 – Reagentes utilizados nos ensaios de susceptibilidade por difusão em disco e microdiluição em caldo dos derivados azólicos frente <i>Cryptococcus</i>	Pag. 36
TABELA 2.4 – Diâmetro médio dos halos de inibição (milímetros) dos 33 derivados azólicos frente à <i>C. gattii</i> e <i>C. neoformans</i>	Pag. 47
TABELA 2.5 – Divisão dos antifúngicos em grupo treino e teste.....	Pag. 51
TABELA 2.6 – Correlação entre os descritores utilizados no modelo KNN.....	Pag. 53
TABELA 2.7 – Descritores presentes no modelo final de SIMCA.....	Pag. 55
TABELA 2.8 – Concentração inibitória mínima (CIM) para 33 antifúngicos derivados azólicos frente às cepas <i>C. gattii</i> e <i>C. neoformans</i>	Pag. 62
TABELA 2.9 – Valores de pCIM dos 33 antifúngicos frente à <i>C. gattii</i>	Pag. 64
TABELA 2.10 – Influência da distinção de fragmento sobre os parâmetros estatísticos do HQSAR, usando o tamanho de fragmento padrão (4 – 7).....	Pag. 66
TABELA 2.11 – Influência do tamanho do fragmento sobre os parâmetros estatísticos dos modelos de HQSAR.....	Pag. 66
TABELA 2.12 – Descritores presentes no modelo final de PLS.....	Pag. 70
TABELA 2.13 – Valores experimentais e preditos de pCIM dos antifúngicos do grupo teste para o modelo de QSAR 2D baseado em descritores topológicos.....	Pag. 71
TABELA 3.1 – Reagentes utilizados nos ensaios de biológicos.....	Pag. 85
TABELA 3.2 – Mutações no gene ERG11 de <i>C. albicans</i> e posição correspondente na sequência de <i>C. gattii</i>	Pag. 91
TABELA 3.3 – Estruturas cristalográficas de proteínas homologas à proteína alvo.....	Pag. 98

LISTA DE ABREVIATURAS E SIGLAS

QSAR	Relação quantitativa estrutura-atividade
HQSAR	Hologram Quantitative Structure-Activity Relationship
SBDD	Planejamento de fármacos baseado na estrutura do alvo macromolecular
LBDD	Planejamento de fármacos baseado não ligante
AIDS	Síndrome da imunodeficiência adquirida
ASD	Ágar Sabouraud Dextrose
YEPD	Extrato de levedura, peptona e dextrose
RMN	Ressonância Magnética Nuclear
3D	Tridimensional
CLSI	Instituto de Padrões Clínicos e Laboratoriais
KNN	K-nearest Neighbors
SIMCA	Soft Independent Modeling of Class Analogy
PC	Componente principal
CIM	Concentração Inibitória Mínima
PLS	Regressão multivariada por mínimos quadrados parcial
ATCC	American Type Culture Collection
DMSO	Dimetilsulfóxido
RPMI	Roswell Park Memorial Institute medium
pCIM	Logaritmo negativo da concentração inibitória mínima
2D	Bidimensional
RLM	Regressão Linear Múltipla
UFC	Unidade Formadora de Colônia
PCR	Reação em cadeia da polimerase
LPMC	Laboratório de Pesquisa em Microbiologia Clínica
LaBiMM	Laboratório de Bioinformática e Modelagem Molecular

SUMÁRIO

RESUMO

ABSTRACT

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

CAPÍTULO 1

1	PANORAMA GERAL SOBRE O DESENVOLVIMENTO DE FÁRMACOS ANTIFÚNGICOS.....	Pag. 16
1.1	Química medicinal e planejamento de fármacos.....	Pag. 16
1.2	Relação quantitativa entre a estrutura química e a atividade biológica (QSAR)	Pag. 18
1.3	Planejamento de fármacos antifúngicos.....	Pag. 20
1.4	Criptococose.....	Pag. 23

CAPÍTULO 2

2	ESTRATÉGIAS MODERNAS APLICADAS AO PLANEJAMENTO DE FÁRMACOS.....	Pag. 25
2.1	Métodos quimiométricos classificatórios.....	Pag. 26
2.2	Métodos quimiométricos quantitativos.....	Pag. 28
2.3	Objetivos.....	Pag. 33
2.4	Ensaio biológicos.....	Pag. 34
2.4.1	Microorganismos.....	Pag. 34
2.4.2	Reagentes.....	Pag. 34
2.4.3	Ensaio de susceptibilidade por difusão de disco.....	Pag. 36
2.4.3.1	Preparo de discos.....	Pag. 36
2.4.3.2	Preparo de inóculo.....	Pag. 37
2.4.3.3	Aplicação dos discos.....	Pag. 37
2.4.3.4	Controles de experimento.....	Pag. 38

2.4.4	Ensaio de susceptibilidade por microdiluição em caldo.....	Pag. 38
2.4.4.1	Preparo de soluções de antifúngicos.....	Pag. 38
2.4.4.2	Preparo de inóculo.....	Pag. 39
2.4.4.3	Determinação da concentração inibitória mínima.....	Pag. 39
2.4.4.4	Leitura dos resultados.....	Pag. 39
2.4.4.5	Controles de experimento.....	Pag. 40
2.5	Ensaio quimiométrico.....	Pag. 40
2.5.1	Conjunto de dados.....	Pag. 40
2.5.2	Cálculo de descritores.....	Pag. 41
2.5.3	Seleção de descritores.....	Pag. 41
2.5.4	Holograma QSAR.....	Pag. 42
2.5.5	Validação dos modelos.....	Pag. 43
2.5.5.1	Modelos classificatórios.....	Pag. 43
2.5.5.2	Modelos de QSAR clássico e HQSAR.....	Pag. 44
2.6	Resultado e discussão.....	Pag. 45
2.6.1	Ensaio de susceptibilidade por difusão de disco.....	Pag. 45
2.6.2	Modelos classificatórios.....	Pag. 50
2.6.2.1	Conjunto de dados.....	Pag. 50
2.6.2.2	Estudos de reconhecimento padrão supervisionados.....	Pag. 51
2.6.3	Ensaio de susceptibilidade por microdiluição em caldo.....	Pag. 60
2.6.4	Modelos quantitativos de relação estrutura-atividade.....	Pag. 63
2.6.4.1	Modelos de HQSAR.....	Pag. 65
2.6.4.2	Modelos de QSAR 2D clássico.....	Pag. 68

CAPÍTULO 3

3	ESTUDOS DA RESISTÊNCIA DE <i>CRYPTOCOCCUS</i> À DERIVADOS AZÓLICOS.....	Pag. 75
3.1	Objetivos.....	Pag. 83

3.2	Ensaio biológicos.....	Pag. 84
3.2.1	Microorganismos.....	Pag. 84
3.2.2	Reagentes.....	Pag. 84
3.2.3	Ensaio para determinação de Concentração Inibitória Mínima na presença de inibidores de bomba de efluxo.....	Pag. 86
3.2.4	Extração e sequenciamento do gene ERG11 de <i>C. gattii</i>	Pag. 86
3.2.4.1	Extração do DNA.....	Pag. 86
3.2.4.2	Amplificação do gene ERG11.....	Pag. 87
3.2.4.3	Sequenciamento do gene ERG11.....	Pag. 88
3.3	Construção de modelos 3D através de modelagem por homologia.....	Pag. 88
3.3.1	Identificação e seleção de proteínas homólogas à lanosterol 14 α -desmetilase de <i>Cryptococcus gattii</i>	Pag. 89
3.3.2	Alinhamento da sequência primária da proteína alvo e molde.....	Pag. 89
3.3.3	Construção do modelo por homologia.....	Pag. 90
3.3.4	Otimização dos modelos.....	Pag. 90
3.3.5	Validação dos modelos por homologia.....	Pag. 91
3.3.6	Mutações na enzima lanosterol 14 α -desmetilase que conferem resistência à derivados azólicos.....	Pag. 91
3.4	Análise dos modelos comparativos.....	Pag. 92
3.5	Resultados e discussão.....	Pag. 93

CAPÍTULO 4

4	CONSIDERAÇÕES FINAIS E PERSPECTIVAS.....	Pag. 114
	REFERÊNCIAS.....	Pag. 117
	APÊNDICES.....	Pag. 140

CAPÍTULO 1

1 PANORAMA GERAL SOBRE O DESENVOLVIMENTO DE FÁRMACOS ANTIFÚNGICOS

1.1 Química medicinal e planejamento de fármacos

A descoberta e o desenvolvimento de novos fármacos dependem de diversos fatores relacionados à inovação, tecnologia, gerenciamento e investimento em pesquisa e desenvolvimento (P&D).¹ Nesta perspectiva, o planejamento de fármacos visa desenvolver entidades químicas com alto índice de sucesso terapêutico e ação específica. Entretanto a obtenção de inovações farmacoterapêuticas exige altos investimentos e grande demanda de tempo.² De fato, o planejamento de fármacos é considerado altamente arriscado,^{3; 4} pois são necessários cerca de 15 anos de pesquisa, ao custo de aproximadamente 0,8 a 1,4 bilhão de dólares, para que compostos líderes percorram todas as etapas necessárias para se tornarem fármacos disponíveis comercialmente.⁵ Durante as últimas décadas, esforços foram realizados para tornar esse processo menos oneroso e mais ágil, o que tem se conseguido através da utilização de metodologias modernas que integram diversas áreas do conhecimento.^{6; 7; 8}

A química medicinal é uma área baseada na química, que envolve aspectos das ciências biológicas, médicas e farmacêuticas, com o intuito de auxiliar na descoberta, desenvolvimento, identificação e preparação de moléculas biologicamente ativas; no estudo do metabolismo; na interpretação do mecanismo de ação ao nível molecular e na elucidação de relações entre a estrutura química e a atividade biológica de moléculas bioativas.⁹ A química medicinal interage também com áreas relacionadas à matemática e, principalmente, informática, a partir de onde se delineiam novas áreas de conhecimento como a bioinformática e a quimioinformática.^{10; 11; 12; 13}

Do ponto de vista geral, a bio- e quimioinformática são abordagens computacionais que objetivam aplicar a tecnologia da informação e a capacidade de gerenciamento de dados para solucionar problemas biológicos e químicos, transformando assim dados em informação e informação em conhecimento.^{13; 14} No que se refere ao desenvolvimento de fármacos, essas

metodologias são usadas para a análise e visualização de moléculas bioativas, bem como para investigar suas propriedades físico-químicas, estruturais e biológicas. Os resultados obtidos através dessas técnicas permitem uma maior compreensão das exigências estéricas e eletrônicas necessárias para o efeito biológico desejado. Um exemplo de aplicação dessas metodologias é o planejamento de moléculas bioativas baseado na estrutura do alvo macromolecular (SBDD, do inglês *Structure-Based Drug Design*) (**FIGURA 1.1**).¹⁵

A partir da identificação e análise de propriedades estruturais dos sítios moduladores/ativos em proteínas pode-se empregar recursos computacionais para simular o acoplamento molecular de micromoléculas (ligantes) nesses sítios. Apesar das técnicas de SBDD serem úteis tanto para a descoberta de moléculas bioativas inéditas quanto para otimização de suas propriedades farmacodinâmicas, existem vários alvos macromoleculares para os quais ainda não há estrutura 3D disponível, como complexos poliproteicos¹⁶, proteínas acopladas à membrana¹⁷ ou transmembrana¹⁸. Nestes casos, pode-se recorrer a outra abordagem da quimioinformática, que agora se baseia em informações oriundas de uma série de ligantes conhecidos para os quais a atividade biológica e o mecanismo de ação, comuns a todas as moléculas da série, também são conhecidos. Essas técnicas, denominadas de planejamento de fármacos baseado no ligante (LBDD, do inglês *Ligand-based drug design*), podem ser aplicadas na identificação de novas moléculas bioativas ou na otimização de compostos líderes.¹⁹ Entre as técnicas de LBDD empregadas para esse último objetivo destacam-se os métodos que analisam quantitativamente as relações entre a estrutura química e a atividade biológica (QSAR, do inglês *Quantitative Structure-Activity Relationship*) (**FIGURA 1.1**).²⁰

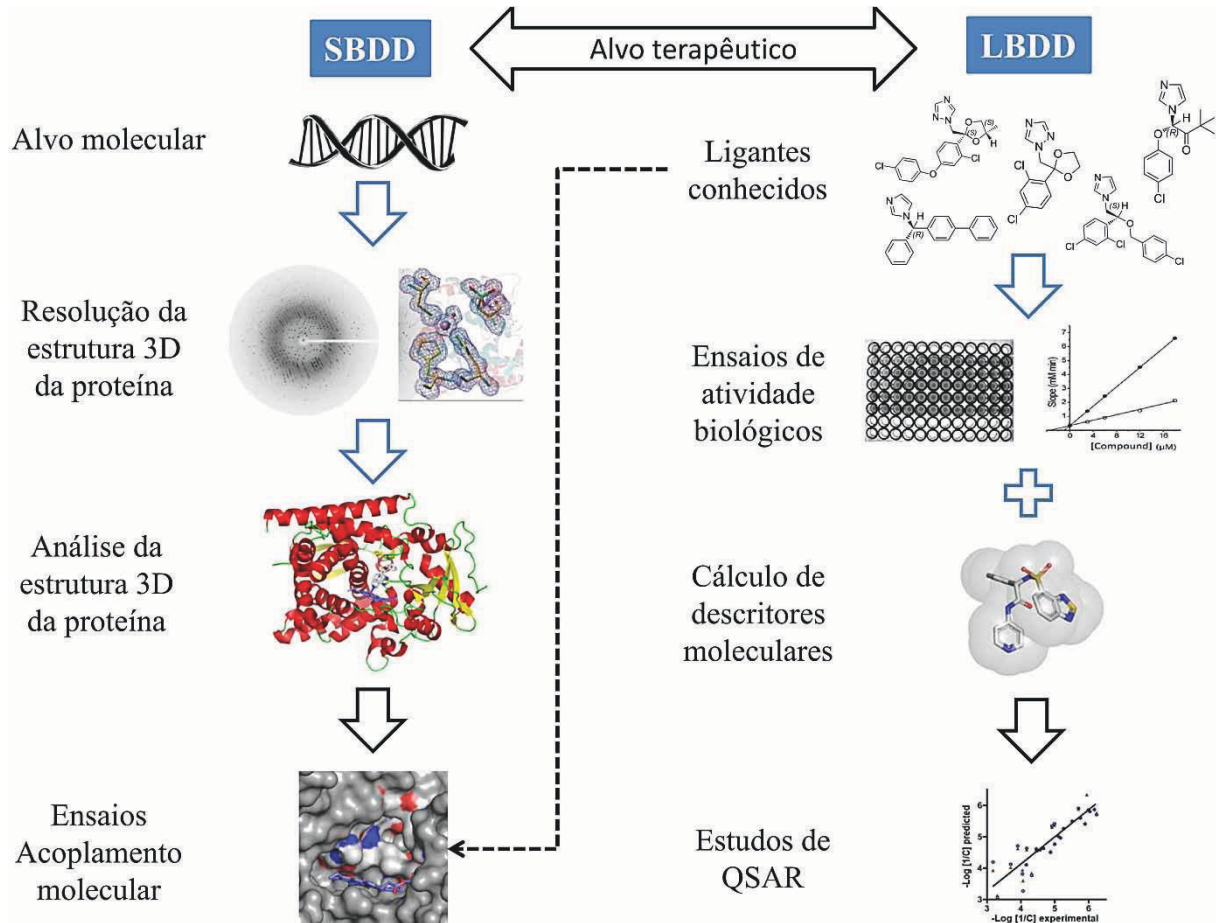


FIGURA 1.1 – Esquema comparativo das técnicas de SBDD e LBDD.

1.2 Relação quantitativa entre a estrutura química e a atividade biológica (QSAR)

Os métodos de QSAR têm como objetivo analisar e detectar propriedades estruturais e físico-químicas, que governam a atividade biológica de um conjunto de moléculas, com o intuito de elucidar o mecanismo de ação e estimar a atividade biológica de moléculas inéditas. Esse objetivo é alcançado através da criação de modelos matemáticos (equações de 1º grau) que quantificam quais propriedades, representadas na forma de descritores moleculares, são relevantes para a atividade biológica. Um descritor molecular codifica as informações químicas na forma de valores numéricos que representam uma determinada propriedade de uma molécula (Figura 1.2).²¹

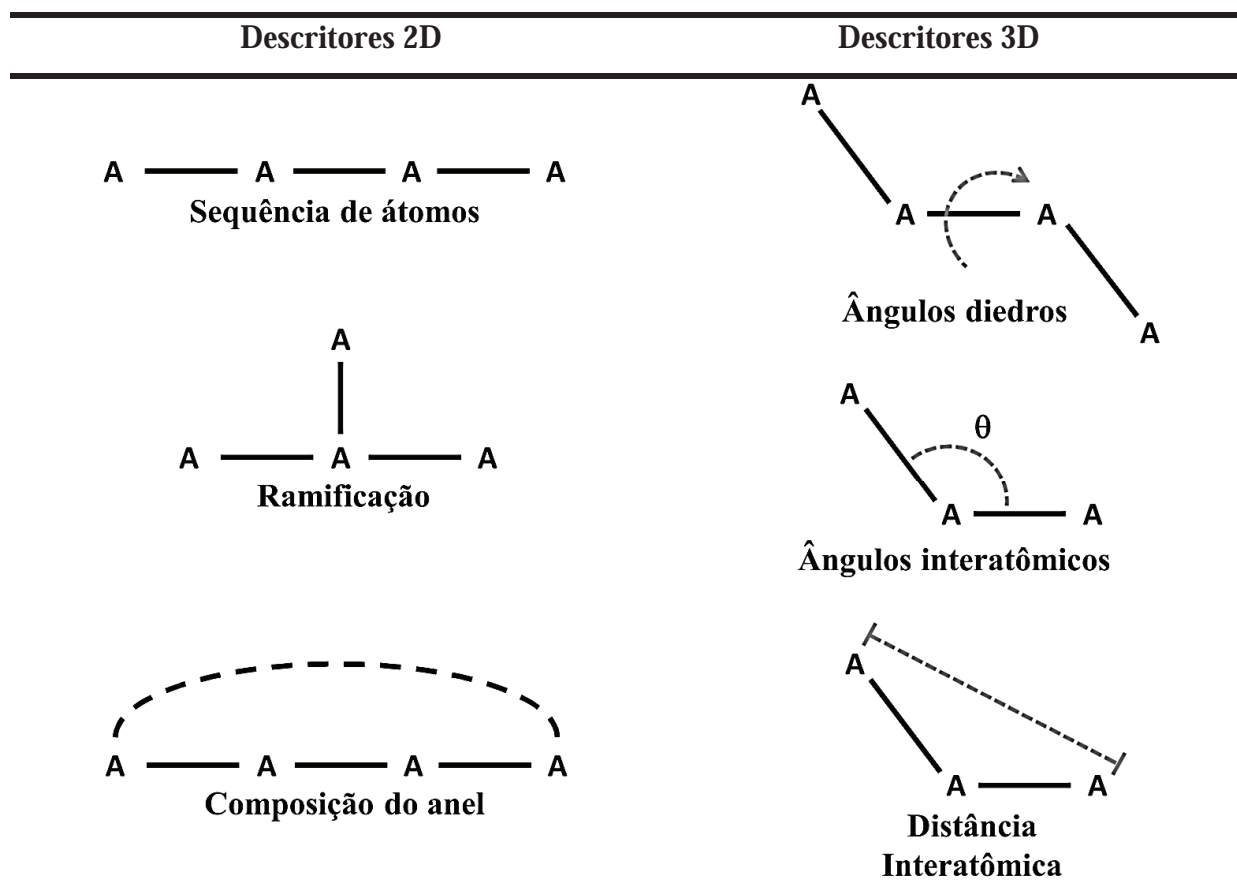


FIGURA 1.2 – Tipos de descritor molecular, onde “A” representa os átomos. Adaptado de Willett e colaboradores (1998)²²

Em princípio, descritores quânticos, calculados a partir de funções de onda das moléculas são considerados mais preciso e eficazes,^{23; 24} porém possuem aplicabilidade limitada devido ao tempo necessário para serem calculados, quando se investiga um grande número de moléculas.^{25; 26} Alternativamente, descritores que codificam tipo de átomos, conectividade, etc. são rapidamente calculados, porém são pouco eficazes na diferenciação molecular. Entre esses extremos, os índices topológicos parecem ser um compromisso aceitável entre o grau de informação contido e a agilidade com que podem ser calculados.^{27; 28; 29; 30} Atualmente, milhares de descritores bidimensionais estão disponíveis, porém apenas uma pequena parcela deles é útil para descrever a correlação existente entre a estrutura química e a atividade biológica de uma série de moléculas.^{31; 32; 33} Assim, etapas de seleção de descritores são consideradas cruciais para o desenvolvimento de modelos de QSAR robustos.^{34; 35; 36; 37; 38}

Entre as diversas estratégias de seleção de descritores, podem-se utilizar os métodos de seleção “para frente” (*forward selection*) e “para trás” (*backward selection*).^{39; 40}

O método *forward selection* é um algoritmo iterativo que adiciona, ao modelo de regressão, um descritor por vez até obtenção do modelo com melhor qualidade estatística. Em contrapartida, o método *backward selection* inicia o processo de construção dos modelos contendo todos os descritores disponíveis, os quais são sucessivamente eliminados até a identificação do modelo de melhor qualidade estatística. A cada inclusão ou eliminação de um descritor, um novo modelo é construído e a correlação dos descritores com a atividade biológica (ajuste), bem como, a significância estatística dos descritores para o modelo é investigada.³⁴ Apesar destas técnicas serem relativamente simples e práticas, apresentam limitações no processo de seleção de variáveis, uma vez que uma variável incluída (método *forward*) ou eliminada (método *backward*) não poderia ser em seguida descartada ou reincorporada ao modelo, impedindo a geração de modelos com combinações diferentes de descritores. Para contornar essa limitação, métodos baseados em algoritmos evolutivos, como o algoritmo genético têm sido largamente implementados na seleção de descritores.^{36; 41}

Além do tipo de descritor utilizado, a qualidade e aplicabilidade de modelos de QSAR dependem da: 1) Homogeneidade das moléculas, ou seja, elas devem ser química- e estruturalmente semelhante, exercendo seus efeitos biológicos através do mesmo mecanismo de ação;^{35; 42} 2) Variação da propriedade biológica, que deve ser determinada por meio de ensaios padronizados e apresentar variação de potência de 2 ou mais ordens de magnitude.^{43; 44}

Em função da gama de informações disponibilizadas por modelos de QSAR, esta metodologia tem sido empregada para o desenvolvimento de diferentes classes de fármacos.^{45; 46; 47} Entretanto, avanços semelhantes ainda não foram alcançados para a classe dos antifúngicos, como descrito a seguir, justificando, portanto o trabalho aqui descrito.

1.3 Planejamento de fármacos antifúngicos

A história do planejamento de fármacos antifúngicos mostra o intenso esforço humano na tentativa de identificar alternativas terapêuticas seletivas, com boa biodisponibilidade sistêmica e baixa toxicidade. Durante os últimos 50 anos, os avanços no desenvolvimento de fármacos antifúngicos contrastaram com o crescimento alarmante de infecções fúngicas, em especial de *Candida*, *Aspergillus* e *Cryptococcus*.⁴⁸ Até a década 1970, a maioria das micoses, em geral de pele, não eram letais resultando numa baixa demanda de novos

antifúngicos.^{49; 50} No entanto, a expansão de outras áreas da medicina como a quimioterapia anticâncer, a introdução de imunossupressores para o tratamento de transplante de órgãos e, principalmente, o advento da AIDS alterou este panorama, bem como a importância das doenças fúngicas para a saúde pública.^{51; 52} De fato, as micoses invasivas (sistêmicas e oportunistas) tornaram-se alvo de preocupação dos órgãos mundiais de saúde e da indústria farmacêutica na busca por alternativas terapêuticas.⁵²

Diferentemente da história da quimioterapia bacteriana, os fármacos antifúngicos, em sua maioria, foram descobertos como inibidores do crescimento fúngico, sem grandes informações sobre o mecanismo de ação e/ou alvo terapêutico afetado.^{53; 54} Por este motivo, apenas 15 fármacos, relacionados a um pequeno número de alvos estão atualmente disponíveis para o tratamento das diversas micoses.⁵⁵

Adicionalmente, a maioria dos fármacos disponíveis (**FIGURA 1.3**) apresenta diversas limitações terapêuticas, tais como: alto índice de efeitos colaterais, especialmente em derivados polienicos (ex. Anfotericina B)⁴⁹; extensiva metabolização hepática, notória em derivados de alilaminas (ex. terbinafina), requerendo altas doses⁵⁶ ou rápido desenvolvimento de resistência (ex. 5-flucitosina).⁵⁷ Diante deste cenário, a introdução dos derivados azólicos tornou-se um marco da quimioterapia antifúngica, uma vez que, possibilitou que esses fármacos tivessem amplo espectro de ação, baixa toxicidade e atividade sistêmica.⁵⁴

Os derivados azólicos bloqueiam a via do ergosterol através da inibição da enzima Lanosterol 14-alfa desmetilase fúngica. Como resultado, ocorre a redução na produção de ergosterol e acumulação do substrato lanosterol e de derivados cujas características físico-químicas não são adequadas para a manutenção da estrutura da membrana celular, reduzindo a permeabilidade seletiva da célula e promovendo a morte do fungo.⁵⁰ Entre os principais fármacos, dessa classe, disponíveis para infecções sistêmicas e oportunistas destacam-se o itraconazol, fluconazol, voriconazol e posaconazol (**FIGURA 1.3**).⁵⁸

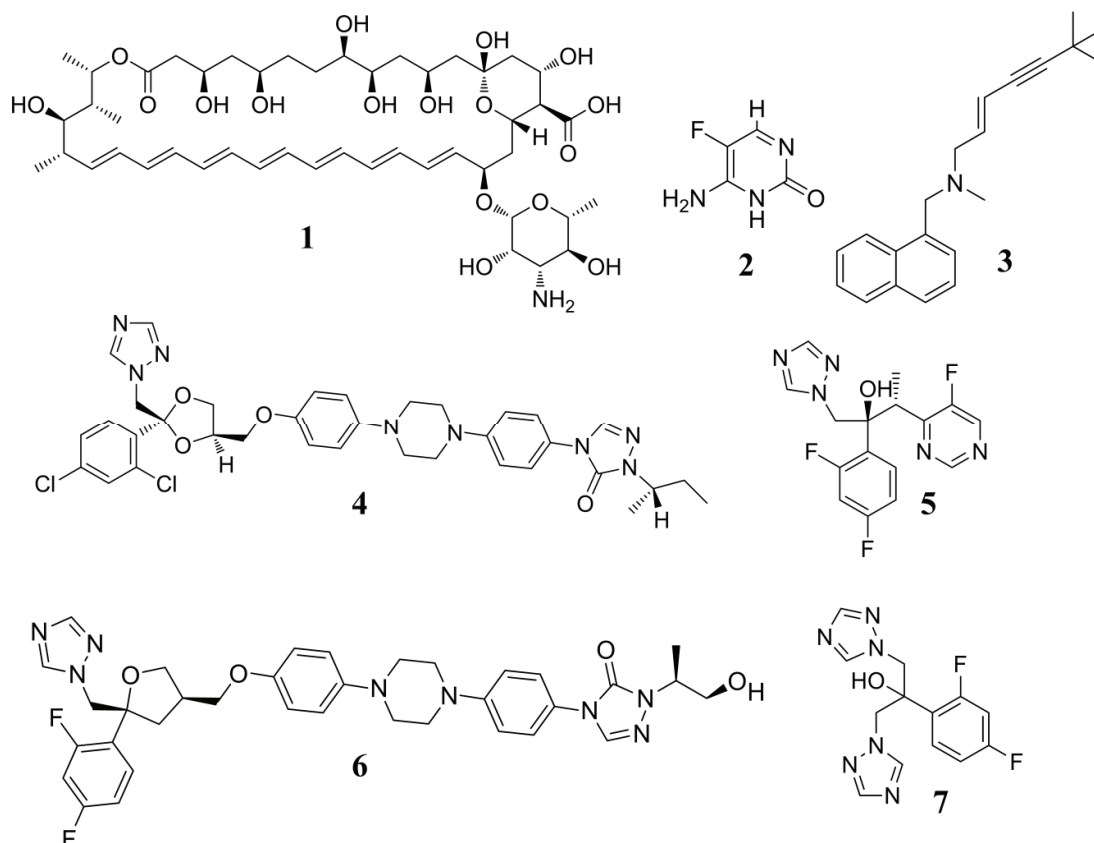


FIGURA 1.3 – Principais antifúngicos disponíveis para uso clínico: Anfotericina (1); 5-fluocitosina (2); Terbinafina (3); Itraconazol (4); Voriconazol (5); Posaconazol (6) e Fluconazol (7)

Apesar do desenvolvimento contínuo de novos antifúngicos, o aparecimento quase concomitante de cepas resistentes tem tornado o tratamento de infecções por *Aspergillus*, *Candida*, *Cryptococcus* cada vez mais difícil.^{59; 60; 61}

Por este motivo, estudos de QSAR 2D e HQSAR para cepas resistentes de *Candida albicans* tem sido frequentemente empregadas, possibilitando o desenvolvimento de modelos capazes de contribuir para a otimização de derivados azólicos.^{62; 63; 64} De forma semelhante, existem diversos estudos aplicando modelagem por homologia da enzima lanosterol 14 α -desmetilase de *C. albicans*.^{65; 66} Para destacar como mutações na enzima lanosterol 14 α -desmetilase poderiam influenciar na resistência à derivados azólicos e predizer as principais interações dos derivados triazólicos de segunda geração (ex. Posaconazol e Voriconazol) no alvo terapêutico.^{67; 68}

Embora os estudos com *Candida* mostrem que metodologias modernas de planejamento de fármaco podem ser utilizadas para o desenvolvimento de antifúngicos, estratégias semelhantes foram pouco exploradas no desenvolvimento de fármacos contra *Cryptococcus*.

1.4 Criptococose

A criptococose, também conhecida como tolurose, blastomicose européia e doença de Busse-buschke, é uma doença fúngica infecciosa, oportunista, de grande impacto sobre a saúde da população mundial e potencialmente fatal.⁶⁹

Antes da década de 1980, principalmente no período pré-epidêmico da AIDS (do inglês *Acquired Immune Deficiency Syndrome*), a criptococose era uma infecção sistêmica pouco comum, que afetava pacientes imunossuprimidos, a exemplo de portadores de diabetes mellitus, doença de Hodgkin's, lupus eritematoso sistêmico (LES) e terapias imunossupressivas (corticóides e anticancerígenos).⁷⁰

No entanto, nos últimos 20 anos, ocorreu um aumento expressivo no número de casos de criptococose paralelamente à elevação do número de paciente imunocomprometidos, particularmente portadores da AIDS, nos quais são encontrados com mais de 80% dos casos de criptococose.^{71; 72}

O agente etiológico da criptococose é o basídiomiceto do gênero *Cryptococcus* que é composto por mais de 30 espécies, das quais as leveduras encapsuladas *C. neoformans* e *C. gattii* são consideradas as mais relevantes para desenvolvimento da doença.⁷³ Estas leveduras ainda podem ser sub-classificadas em cinco sorotipos (A, B, C, D e AD) e oito tipos moleculares.^{74; 75; 76}

O *C. neoformans* é a principal causa de infecções oportunistas em pacientes imunocomprometidos, em especial portadores da síndrome da imunodeficiência adquirida (AIDS). *C. gattii*, por sua vez, é capaz de infectar pessoas imunocompetentes.^{77; 78; 79}

Segundo a Sociedade Americana de Doenças,⁸⁰ o tratamento da criptococose em pacientes HIV positivos é feita em três etapas: indução, consolidação e manutenção. Na primeira fase é comum se utilizar anfotericina B (AmB) juntamente com 5-flucitosina (5-FC). Porém no estágio de manutenção do tratamento, com duração de 6 a 12 meses, são utilizados principalmente derivados azólicos (fluconazol e itraconazol).

Diante do aparecimento de cepas resistentes aos derivados de azol, torna-se imprescindível investigar os mecanismos moleculares que são responsáveis por essa resistência bem como quais características físico-químicas e estruturais devem ser alteradas, nos antifúngicos, visando torná-los mais eficientes no combate da criptococose causada por cepas com baixa susceptibilidade a fluconazol. Os capítulos subsequentes relatam nossos esforços para alcançar esse objetivo através de duas abordagens distintas:

1) Emprego de técnicas de QSAR 2D, baseadas em hologramas moleculares e descritores topológicos, na investigação de propriedades químicas que distinguem antifúngicos ativos e inativos contra cepas de *C. gattii* resistentes a fluconazol e elucidação das características físico-químicas que determinam a potência desses antifúngicos frente a *C. gattii*. (Capítulo 2)

2) Utilização de modelos por homologia para explicar, do ponto de vista estrutural, como mutações pontuais na enzima lanosterol 14 α -desmetilase de *C. gattii* estão relacionadas com o aparecimento de resistência a derivados de azol. (Capítulo 3)

CAPÍTULO 2

2 ESTRATÉGIAS MODERNAS APLICADAS AO PLANEJAMENTO DE FÁRMACOS

Durante os últimos 20 anos, abordagens centradas no alvo terapêutico (ensaios enzimáticos) guiaram a identificação e otimização de compostos líderes. Entretanto, apesar da introdução de técnicas modernas de ensaio em larga escala (HTS, do inglês *High Throughput Screening*) e investimentos crescentes em P&D, não houve um aumento proporcional no número de novas entidades químicas (NCEs, do inglês *New Chemical Entities*) que chegaram ao mercado.⁸¹ Em parte, essa baixa taxa de sucesso pode ser atribuída ao fato de que ensaios enzimáticos negligenciam fatores farmacocinéticos importantes para atividade biológica, como metabolismo, distribuição e excreção. Por essa razão, ensaios celulares tem recebido maior atenção na identificação e seleção de compostos protótipos.⁸²

De modo geral, os ensaios celulares objetivam avaliar uma variedade de efeitos fenotípicos, tais como proliferação celular, toxicidade, mortalidade, alterações morfológicas e produção de um produto mensurável, sob a ação de um agente modulador (ex. antifúngico).⁸³ De fato, esses ensaios agregam mais um grau de complexidade ao processo avaliado, oferecendo maior conteúdo de informações e uma representação mais precisa e fiel da ação do fármaco na célula, uma vez que considera a passagem pela membrana e distribuição no meio intracelular como informações adicionais.^{84; 85}

No que se refere ao desenvolvimento de antifúngicos, ensaios de sensibilidade por difusão em disco desempenham um papel importante na etapa de triagem *in vitro*, em razão da simplicidade e baixo custo, além de fornecerem resultados quantitativos (diâmetro de halo de inibição), que permitem a categorização das substâncias avaliadas (Ex. ativo e inativo).⁸⁶ Todavia, esses ensaios são menos precisos que outros métodos de avaliação de atividade antifúngica, como por exemplo, os ensaios de concentração inibitória mínima (CIM) obtidos por microdiluição em caldo.⁸⁷ De fato, tendo-se em vista a reprodutibilidade, robustez e precisão dos resultados, o Instituto de Padrões Clínicos e Laboratoriais (do inglês *Clinical and Laboratory Standards Institute – CLSI*) recomenda que a avaliação da atividade antifúngica seja realizada por microdiluição em caldo.

Considerando que o grau de informação biológica disponível (contínua ou discreta) implica no emprego de metodologias quimiométricas distintas, a escolha e padronização dos ensaios biológicos constitui etapa fundamental para o estudo quantitativo das relações entre a estrutura química e a atividade biológica de antifúngicos.

Após a aquisição de dados biológicos para uma série congênere de moléculas, deve-se decidir que tipo de análise quimiométrica é mais adequado.

2.1 Métodos quimiométricos classificatórios

Quando a propriedade biológica é medida de forma discreta (ex. ativo e inativo), podem ser aplicados métodos de reconhecimento de padrões supervisionado que buscam regras classificatórias baseadas na divisão do espaço químico entre classes de compostos ou ao redor delas como, por exemplo, os métodos KNN (do inglês *K-Nearest Neighbors*) e SIMCA (do inglês *Soft Independent Modeling of Class Analogy*), respectivamente.^{88; 89}

O método KNN foi inicialmente introduzido em 1951 por Fix e Hodges⁹⁰ como método de classificação não paramétrico que utiliza a comparação entre duas amostras (moléculas), baseando-se na distância euclidiana entre elas.^{91; 92} A classificação de uma amostra externa (não utilizada na construção do modelo) pode ser determinada com base na distância desta nova amostra em relação às k-ésimas amostras vizinhas mais próximas no espaço químico definido pelos descritores utilizados, ou seja, a classificação da amostra desconhecida é determinada pela classe com maior número de vizinhos à ela relacionados.^{93;}

94

Uma representação ilustrativa desta técnica está apresentada na **FIGURA 2.1** que mostra a classificação de uma amostra de classe desconhecida (círculo) com base no número de vizinhos de cada classe (triângulos ou quadrados). Ao se considerar apenas um vizinho, a amostra desconhecida é atribuída à classe dos quadrados, enquanto que ao se considerar 3 vizinhos, a amostra é atribuída à classe dos triângulos.

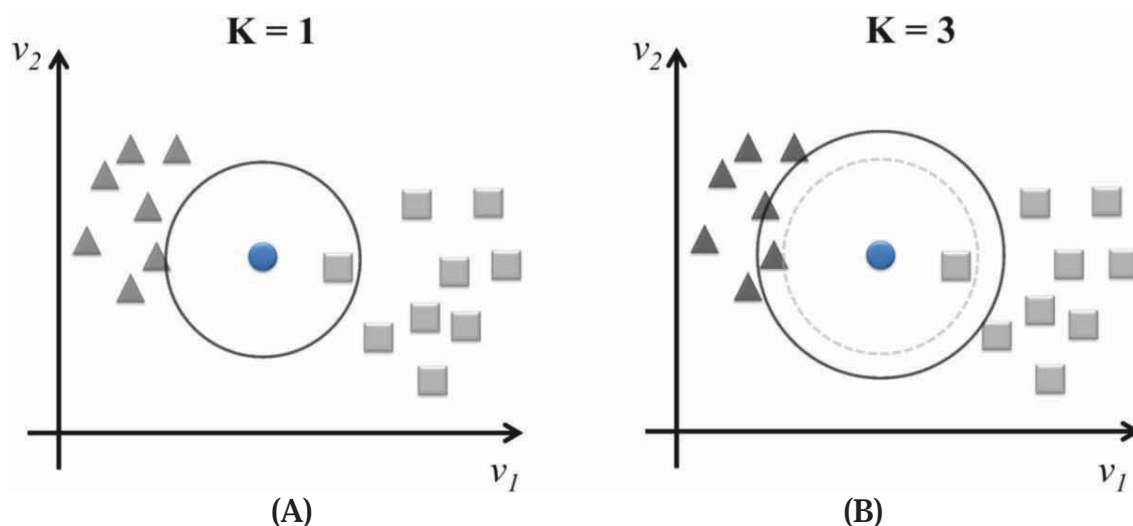


FIGURA 2.1 – Representação do método KNN para classificação de uma nova amostra considerando 1 vizinho (A) e 3 vizinhos (B)

Apesar da simplicidade e facilidade de implementação, o método KNN utiliza todos os descritores disponíveis para posicionar as amostras no espaço multidimensional, tornando o modelo sensível à presença de descritores pouco relevantes ou multicolineares, o que prejudica a classificação das amostras.⁹⁵ Além disso, esse método determina a classificação das amostras com base na regra classificatória pré-existente (ex. classe 1, classe 2), não permitindo a atribuição de uma amostra desconhecida a um nova classe.⁹⁶

Uma alternativa para contornar as limitações apresentadas pelo método KNN se baseia na utilização de componentes principais (PCA, do inglês *Principal Component Analysis*),⁹⁷ como implementado no método SIMCA, que foi introduzido por Syvante Wold em 1976.^{98; 99}

A PCA procura representar a informação original, geralmente contida em dezenas, ou até centenas de descritores, em um pequeno número de componentes principais (PC, do inglês *principal component*), não colineares, que nada mais são do que combinações lineares das variáveis originais (**FIGURA 2.2**). Na PCA as amostras (moléculas) podem ser projetadas (visualizadas) no novo subespaço químico definido pelas PCs, permitindo assim uma representação gráfica das relações inter-amostras em até 3 dimensões (3 PCs). Geralmente, isto permite a identificação de padrões e agrupamentos entre as amostras.^{97; 100}

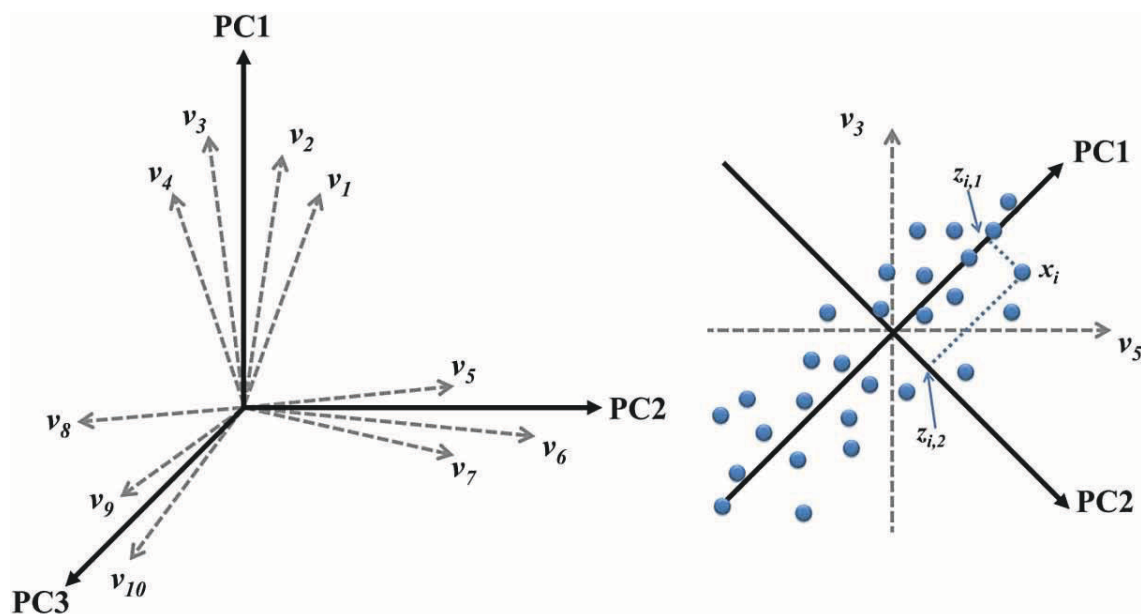


FIGURA 2.2 – Projeção das componentes principais construídas em ordem decrescente da quantidade de variância (à esquerda) e das moléculas nas novas coordenadas definidas pelas PCs (à direita). (Em que, v_j ($j=1, 2, 3...m$) são as variáveis originais relacionadas à cada PC, x_i representa as amostras projetadas nas novas coordenadas $z_{i,1}$ e $z_{i,2}$ definidas pelas PCs) Adaptado de Waterbeemd e Rose (2003), e Sena e colaboradores (2000)¹⁰¹

Diferentemente do método KNN, no método SIMCA pode identificar que uma amostra não faz parte de nenhuma das classes pré-determinadas, quando ela se encontra fora do espaço químico definido pelas PCs que descrevem as classes conhecidas.¹⁰²

2.2 Métodos quimiométricos quantitativos

Quando os dados de atividade biológica são mensurados de forma contínua (ex. K_i , IC_{50} , CIM) pode-se empregar modelos de regressão multivariada como regressão linear múltipla (RLM)¹⁰³ e regressão por mínimos quadrados parciais (PLS, do inglês *Partial Least Square*) com objetivo de identificar fatores físico-químicos e estruturais relevantes para a potência e prever a atividade de moléculas bioativas.^{104; 105}

No primeiro caso, os modelos são expressos na forma de um somatório das m contribuições individuais de parâmetros físico-químicos e estruturais (x) que explicam alguma ou toda a variação de uma variável dependente (y ou propriedade biológica) (**Equação 1**).²⁰

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_mx_m \quad (1),$$

onde $b_0, b_1, b_2 \dots b_m$ representam os coeficientes da regressão.

No entanto, os modelos de RLM têm aplicabilidade limitada no desenvolvimento de modelos com elevado número de descritores, em geral, altamente intercorrelacionados ou descrevendo informações redundantes.^{106; 107} Isto ocorre por que modelos de RLM são sensíveis a presença de colinearidade no conjunto de dados, produzindo um efeito denominado super-ajuste, que tem como característica a geração de modelos com altos valores de coeficiente da regressão, porém instáveis e, conseqüentemente, acompanhados de elevados erros padrões.¹⁰⁸

Alternativamente, métodos baseados em projeções como o PLS são considerados mais robustos que RLM, pois utilizam componentes principais para o desenvolvimento dos modelos.^{109; 110} No modelo de PLS, a variável dependente é expressa em função de um somatório de PCs, com expresso na **Equação 2**:

$$y = a_1PC_1 + a_2PC_2 + a_3PC_3 + \dots + a_mPC_m \quad (2),$$

em que, a_j ($j=1, 2, 3\dots m$) são os coeficientes da regressão e determinam o peso de cada PC no modelo. Cada PC_i (para $i=1, 2, 3,\dots,m$) é uma combinação linear das variáveis originais.^{104; 105; 111; 112}

Uma das técnicas de QSAR 2D que utiliza PLS é o método de holograma QSAR (HQSAR, do inglês *Hologram Quantitative Structure-Activity Relationship*).^{45; 113; 114}

A análise de HQSAR envolve 3 etapas fundamentais: a geração dos fragmentos subestruturais para cada molécula do conjunto de dados; a codificação destes fragmentos em hologramas; e a correlação dos hologramas com a atividade biológica das molécula. Inicialmente, a estrutura bidimensional de todas as moléculas do conjunto de dados é decomposto em todos os tipos de fragmentos possíveis (linear, ramificado, cíclico e sobreposto) (**FIGURA 2.3**).⁴⁴

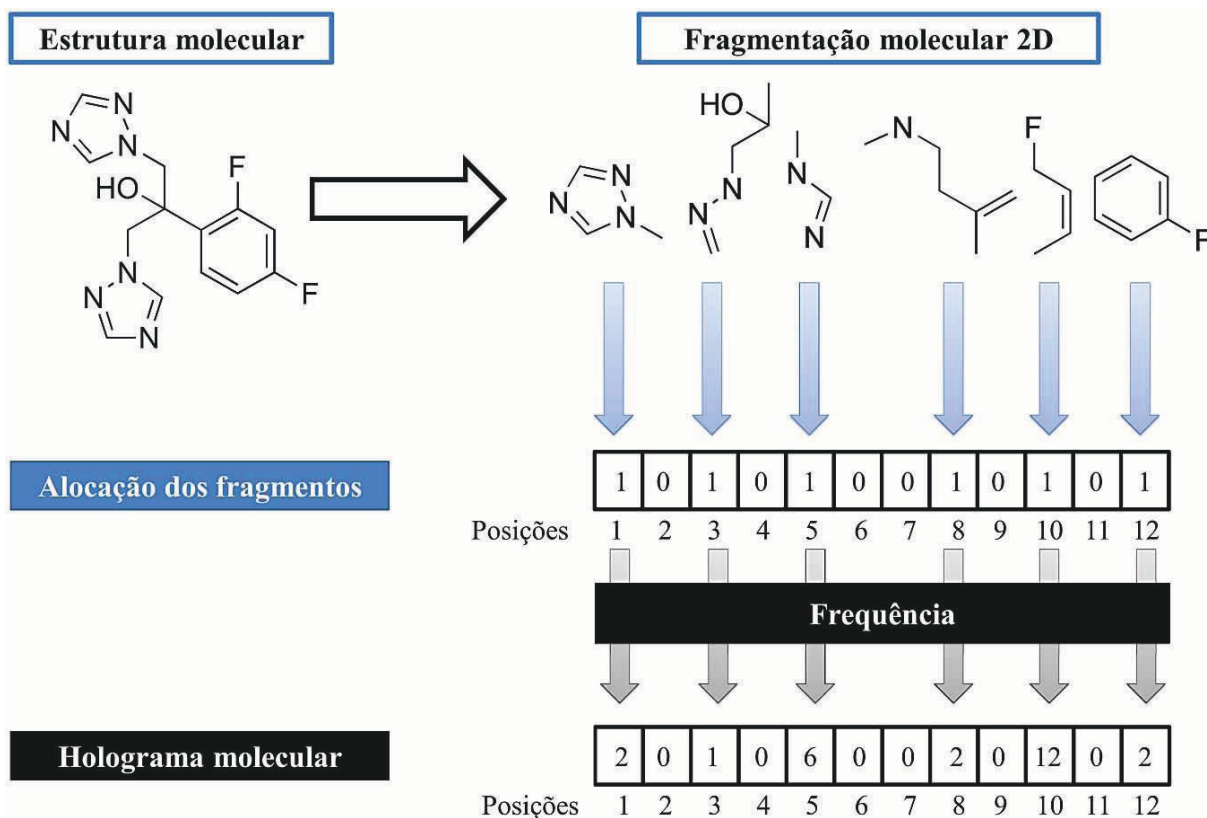


FIGURA 2.3 – Esquema ilustrativo do processo de geração dos hologramas moleculares

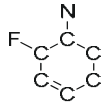
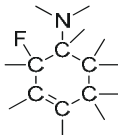
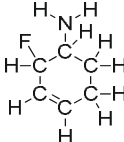
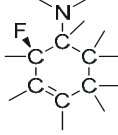
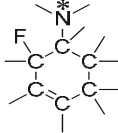
Para a codificação dos fragmentos em hologramas, vários parâmetros são ajustados, tais como comprimento do holograma (HL, do inglês *hologram's length*), tamanho dos fragmentos (FS, do inglês *fragment's size*), e distinção dos fragmentos (FD, do inglês *fragment's distinction*).

O comprimento do holograma determina o número de posições disponíveis para que diferentes fragmentos sejam incorporados, definindo, portanto, a dimensionalidade do espaço químico. Em decorrência do número de fragmentos ser tipicamente maior que o comprimento do holograma, alguns fragmentos com características subestruturais diferentes e, em muitos casos, importantes para a atividade biológica são alocados na mesma posição no holograma. Este fenômeno, denominado colisão de fragmentos, pode reduzir a capacidade de discriminação das características estruturais relevantes à atividade das moléculas, levando a obtenção de modelos com baixa capacidade preditiva e elevados valores de desvio padrão.^{115;}

116

O tamanho dos fragmentos é definido pelo número máximo (N) e mínimo (M) de átomos nos fragmentos gerados, sendo 4 a 7 átomos os valores padrão empregados. A distinção de fragmentos define as características moleculares que serão utilizadas para diferenciar um fragmento do outro (TABELA 2.1).^{117; 118}

TABELA 2.1 – Parâmetros de distinção de fragmento no método HQSAR

Distinção de fragmentos	Definição	Representação
Átomo (A)	Distinção baseada nos diferentes tipos de átomos;	CCCCCFN
Ligação (B, do inglês <i>Bond</i>)	Distinção baseada no tipo de ligação: simples, dupla, tripla ou aromática;	
Conectividade (C)	Distinção baseada na hibridização dos átomos contidos no fragmento;	
Hidrogênio (H)	Distinção baseada na presença de átomos de hidrogênio	
Quiralidade (Ch, do inglês <i>Chirality</i>)	Distinção baseada na presença de centros estereogênicos	
Doador eceptor (DA)	Distinção baseada na presença de átomos doadores ou aceptores de ligação de hidrogênio	

Considerando que os hologramas moleculares são considerados como variável independente do modelo, há necessidade de se utilizar PLS para relacioná-los com a propriedade biológica (**FIGURA 2.4**) e assim investigar quantitativamente as relações entre a estrutura química e a atividade biológica dos compostos em estudo. Por exemplo, a partir desse método, padrões de fragmentos subestruturais relevantes para a propriedade alvo podem ser identificados e empregados para prever a atividade biológica de moléculas congêneres que não foram utilizadas no desenvolvimento dos modelos.^{119; 120; 121}

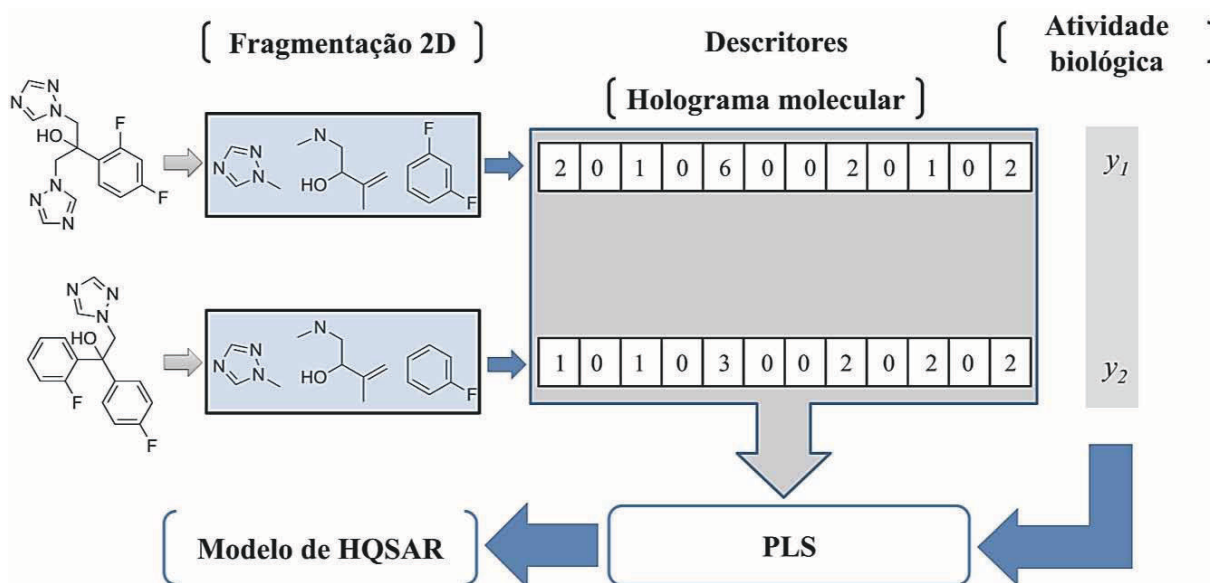


FIGURA 2.4 – Etapas envolvidas na geração de um modelo de HQSAR

Além de prever a propriedade alvo de moléculas ainda não testadas, o modelo de HQSAR pode fornecer informações importantes sobre quais fragmentos moleculares são importantes para a propriedade estudada, pois os mapas de contribuição indicam a contribuição de diferentes fragmentos moleculares para a propriedade em estudo.^{70; 76; 78}

A integração desse tipo de informação com aquela proveniente de modelos de QSAR 2D, baseados em descritores topológicos permite uma análise complementar que sugere quais são os determinantes físico-químicos responsáveis pelas contribuições positivas ou negativas observadas nos mapas de contribuição.^{47; 62}

2.3 Objetivos

Visando contribuir para o entendimento das relações entre a estrutura química e a atividade biológica de antifúngicos azólicos, potencialmente ativos contra *C. gattii*, os objetivos desse trabalho são:

- i) Avaliar a sensibilidade de cepas de *Cryptococcus* sensíveis e resistente a fluconazol frente a um conjunto de 33 derivados azólicos, disponíveis comercialmente;
- ii) Determinar a concentração inibitória mínima de 33 derivados azólicos para cepas de *Cryptococcus* sensíveis e resistentes a fluconazol;
- iii) Identificar as propriedades físico-químicas que diferenciam derivados azólicos ativos e inativos frente a *C. gattii* resistente à fluconazol;
- iv) Analisar os fatores estruturais e físico-químicos que contribuem para atividade biológica de derivados azólicos ativos contra *C. gattii* resistente à fluconazol.

2.4 Ensaios biológicos

Os ensaios de susceptibilidade aos antifúngicos por difusão em disco e por diluição em caldo foram desenvolvidos no Laboratório de Pesquisa em Microbiologia Clínica (LPMC) da Faculdade de Farmácia da Universidade Federal da Bahia (UFBA).

2.4.1 Microorganismos

Para o desenvolvimento dos estudos foi utilizado um isolado clínico *C. gattii* resistente à fluconazol disponível no banco de isolados do Laboratório de Pesquisa em Microbiologia Clínica e uma cepa controle *C. neoformans var. grubii* ATCC 90112, sensível à fluconazol.

2.4.2 Reagentes

Foram adquiridos 33 derivados azólicos comerciais disponíveis na empresa Sigma-Aldrich Fisher com grau de pureza igual ou superior a 95% (**TABELA 2.2**).

TABELA 2.2 – Códigos de derivados imidazólicos e triazólicos comerciais

Antifúngico	Código	Derivado	Antifúngico	Código	Derivado
Azaconazol	34045	Triazólico	Bifonazol	B8773	Imidazólico
Bromuconazol	46094	Triazólico	Cetoconazol	K1003	Imidazólico
Ciproconazol	46068	Triazólico	Climbazol	36127	Triazólico
Clotrimazol	33894	Imidazólico	Diclobutrazol	36764	Triazólico
Difenoconazol	36531	Triazólico	Diniconazol	46049	Triazólico
Econazol	E4632	Imidazólico	Epoxiconazol	36848	Triazólico
Etaconazol	45471	Triazólico	Fembuconazol	46096	Triazólico
Fluconazol	F8929	Triazólico	Fluotrimazol	45507	Triazólico
Fluquinconazol	46301	Triazólico	Flusilazol	45753	Triazólico
Flutriafol	34344	Triazólico	Hexaconazol	34348	Triazólico
Itraconazol	16657	Triazólico	Imazalil	36130	Imidazólico
Myclobutanil	34360	Triazólico	Metconazol	37909	Triazólico
Penconazol	36189	Triazólico	Miconazol	M3512	Imidazólico
Propiconazol	45642	Triazólico	Prochloraz	45631	Imidazólico
Sulconazol	S9632	Imidazólico	Prothioconazol	34232	Triazólico
Tebuconazol	36565	Triazólico	Triadimenol	46138	Triazólico
Triticonazol	34172	Triazólico			

Os demais reagentes utilizados nos ensaios biológicos estão listados na **TABELA 2.3**.

TABELA 2.3 – Reagentes utilizados nos ensaios de susceptibilidade por difusão em disco e microdiluição em caldo

Reagente	Marca
Ácido 4-Morfolinopropanosulfônico (MOPS)	Qhemis
Ácido clorídrico	Quimix
Ágar Batata dextrose	Acumedia
Ágar Mueller-Hinton	Difco
Azul de metileno	Nuclear
Cloreto de 2,3,5-Trifenil-Tetrazólio (TTC)	Nuclear
Cloreto de Sódio	Vetec
Dextrose	Synth
Dimetilsulfóxido (DMSO)	Sigma
Hidróxido de Sódio	Nuclear
RPMI-1640	Sigma
Triptona Bacteriológica	Sigma

2.4.3 Ensaios de susceptibilidade por difusão de disco

Os ensaios de susceptibilidade por difusão de disco foram realizados seguindo as recomendações estabelecidas pela norma M44-A2 descrita pelo Instituto de Padrões Clínicos e Laboratoriais (do inglês *Clinical and Laboratory Standards Institute – CLSI*).¹²²

2.4.3.1 Preparo de discos

Uma vez que não há informações sobre o perfil de atividade destes antifúngicos em relação à *Cryptococcus*, a concentração de cada antifúngico em dimetilsulfóxido (DMSO), exceto fluconazol, que foi solubilizado em água, foi padronizada de acordo com a concentração de fluconazol proposta pelo protocolo M-44A do CLSI para *Candida albicans*

(ATCC 90028) (25µg/disco correspondente a 8,2 mmol/L). A fim de avaliar a sensibilidade das cepas de *Cryptococcus*, também foram testadas concentrações equivalentes ao dobro da concentração padrão (50µg/disco correspondente a 16,4 mmol/L) e a metade da mesma (12,5µg/disco correspondente a 4,1 mmol/L). Discos de papel de filtro estéreis, de 6 mm diâmetro, foram impregnados com 10µL de cada solução preparada, sendo posteriormente colocados em estufa Fanem (Modelo 002CB) a 35°C, até todo solvente evaporar.

2.4.3.2 Preparo de inóculo

Os isolados foram semeados em Ágar Sabouraud Dextrose (ASD) e incubados a 28°C por 48 horas. Após o período de incubação, até 5 UFCs (Unidade formadoras de colônia) distintas, com diâmetro aproximado de 1 milímetro, foram suspensas em 5 mL de solução salina (0,85%) estéril. A suspensão foi homogeneizada por 15 segundos em vortex e sua turvação comparada e ajustada à escala padrão 0,5 McFarland (1 a 5 x 10⁶ células/mL). Em seguida, com auxílio de swab estéril, a suspensão foi semeada na superfície de Ágar Muller-Hinton acrescido de 0,5 µg/mL de azul de metileno e 2,0% glicose.

2.4.3.3 Aplicação dos discos

Com auxílio de uma pinça esterilizada, os discos contendo os antifúngicos, devidamente secos, foram colocados sobre a placa de Petri semeada, de forma a manter a distância mínima de 20 milímetros da borda e entre os discos. Em seguida, as placas foram incubadas a 35° C por 48 horas. Após o período de incubação, os diâmetros dos halos de inibição de crescimento fúngico ao redor dos discos foram medidos em milímetros (mm) utilizando uma régua, a partir das extremidades do halo com inibição total do crescimento.

2.4.3.4 Controles de experimento

Todos os experimentos foram realizados em triplicata, porém quando os halos diferiram mais de 2 mm para o mesmo antifúngico, o resultado foi desconsiderado e novo procedimento foi realizado. Para o controle de qualidade e reprodutibilidade foi averiguado o diâmetro do halo de inibição do disco comercial de fluconazol (25µg - CECON) frente às cepas padrão *Candida albicans* ATCC 90028 e *Candida parapsilosis* ATCC 22019 e, como controle negativo, foram utilizados discos de papel de filtro impregnados com 10µL do solvente DMSO.

2.4.4 Ensaio de susceptibilidade por microdiluição em caldo

Os ensaios de susceptibilidade por microdiluição em caldo foram feitos seguindo as recomendações estabelecidas pela Norma M27-A2 descrita pelo CLSI.¹²³

2.4.4.1 Preparo de soluções de antifúngicos

Com exceção do fluconazol, que foi dissolvido no meio RPMI 1640 estéril na concentração de 5120 µg/mL, os antifúngicos foram dissolvidos em DMSO na concentração de 6400 µg/mL. Posteriormente, a solução do fluconazol foi diluída a 1:40 em RPMI 1640 para obter a concentração de trabalho de 128 µg/mL, enquanto, que para os antifúngicos insolúveis em água, as soluções foram diluídas a 1:9 em uma solução de DMSO/RPMI 1640 (1:9) e, em seguida, a 1:5 em RPMI 1640, para obtenção da concentração de trabalho de 128 µg/mL.

2.4.4.2 Preparo de inóculo

A suspensão fúngica foi preparada de acordo com a **seção 2.4.3.2.** e, em seguida, duas diluições consecutivas de 1:50 e 1:20 foram realizadas em RPMI 1640 de forma a atingir um concentração celular de $5,0 \times 10^2$ a $2,5 \times 10^3$ células/mL.

2.4.4.3 Determinação da concentração inibitória mínima

O experimento foi realizado em placa de microdiluição de 96 poços com fundo em formato de “U”, onde foram dispensados 100 μ L de meio RPMI 1640, exceto nos poços da primeira fileira. Em seguida, 100 μ L da solução antifúngica (128 μ g/mL) foram dispensados nos poços da primeira e segunda fileira de todas as colunas. Após homogeneização dos poços da segunda fileira, foram procedidas diluições seriadas e sucessivas transferindo sempre 100 μ L da mistura de uma coluna para a seguinte, desprezando 100 μ L dos poços da última coluna. Em todos os poços foram inoculados 100 μ L da suspensão fúngica preparada na **seção 2.4.4.2.** A concentração final dos antifúngicos nos poços variou de 64 μ g/mL até 0,03 μ g/mL e a concentração fúngica final foi de $2,5 \times 10^2$ a $1,75 \times 10^3$ células/mL. Após esse procedimento, as placas foram homogeneizadas em vortex por 1 a 2 minutos e incubadas a 35° C por 72 horas.

2.4.4.4 Leitura dos resultados

Após o período de incubação, 15 μ L de cloreto de 2,3,5-Trifenil-Tetrazólio (TTC) a 0,5% foram adicionados a todos os poços da placa e a presença de crescimento fúngico foi observada através do aparecimento de coloração avermelhada nos poços. A CIM foi definida como a menor concentração do antifúngico que promove redução proeminente de crescimento, corresponde a 50% ou mais da inibição do crescimento, em comparação ao controle de crescimento fúngico sem antifúngico, como preconizado pelo protocolo do CLSI M27-A2.

2.4.4.5 Controles de experimento

Os experimentos foram feitos juntamente com controles de esterilidade do meio de cultura (100µL de RPMI 1640); controles do crescimento fúngico em caldo (100µL da suspensão fúngica com 100µL do RPMI 1640) e no diluente dos antifúngicos (20µL DMSO com 80µL de RPMI e 100µL da suspensão fúngica), onde foi observado crescimento total da levedura e controle de atividade antifúngica de fluconazol frente *Candida parapsilosis* ATCC 22019, nas concentrações estabelecidas pelo CLSI M27-A2.

2.5 Ensaios quimiométricos

Os estudos quimiométricos (modelos classificatórios e modelos quantitativos) foram desenvolvidos no Laboratório de Bioinformática e Modelagem Molecular (LaBiMM) da Faculdade de Farmácia da Universidade Federal da Bahia, utilizando estações de trabalho que empregam o sistema operacional Windows sete e onde estão instalados os programas utilizados no desenvolvimento do trabalho: SYBYL X 1.1 (Tripos Inc., St. Louis, USA), DRAGON 5.5 (Talete S.R.L, Milan, Italy), MOBYDIGS 1.0 (Talete S.R.L, Milan, Italy) e PIROUETTE 4.0 (Infometrix Inc., Washington, USA).

2.5.1 Conjunto de dados

O conjunto de dados utilizado para o desenvolvimento de modelos classificatórios consiste de 33 derivados azólicos congêneres e estruturalmente diversos (**TABELA 2.2 e APÊNDICE A**).

Os dados de atividade biológica frente à *C. gattii*, obtidos através do teste de susceptibilidade por difusão em disco descrito no **item 2.4.3** (4,1 mmol/L), foram utilizados para dividir as moléculas em grupo ativo e inativo. Será essa classificação que determinará as variáveis dependentes dos modelos classificatórios (**TABELA 2.4**). Assim, antifúngicos com

valores de diâmetro de halo de inibição acima de 25 milímetros foram considerados ativos (18 antifúngicos), caso contrário, inativos (15 antifúngicos).

Os valores de atividade biológica (Concentração Inibitória Mínima ou CIM) determinados por microdiluição em caldo como descrito na **seção 2.4.4** foram expressos na forma de unidades logarítmicas negativas, pCIM ($-\log[\text{CIM}]$), e utilizados como variável dependente no desenvolvimento de modelos de QSAR clássico e HQSAR (**TABELA 2.8 e TABELA 9**).

As estruturas tridimensionais dos derivados azólicos foram construídas com auxílio das ferramentas disponíveis no programa computacional SYBYL X 1.1 e submetidas aos procedimentos de minimização de energia empregando o campo de força Tripos¹²⁴ com critério de convergência de 0,01 kcal/mol e subsequentemente através do método semi-empírico PM3,¹²⁵ com convergência normal e otimização precisa e completa, como disponível no módulo MOPAC do programa SYBYL X 1.1 (Palavras-chave: 1SCF XYZ ESP NOINTER SCALE=1.4 NSURF=2 SCINCR=0.4 NOMM).

2.5.2 Cálculo de descritores

No desenvolvimento de modelos quimiométricos, descritores físico-químicos e estruturais são requeridos. Assim, as estruturas 3D dos derivados azólicos foram utilizadas para cálculo de 2489 descritores bidimensionais (2D), com auxílio programa Dragon 5.5. Descritores com variância interna menor que 0,1 e intercorrelação maior que 97% foram descartados. Esta estratégia manteve 412 descritores que foram usados como variáveis independentes no estudo quimiométrico.

2.5.3 Seleção de descritores

A seleção dos descritores utilizados na construção dos modelos classificatórios foi realizada através do peso de Fisher que calcula a distância entre duas categorias ou a distância normalizada entre duas classes, a partir da diferença entre as médias de um determinado descritor para cada classe, levando em consideração as variâncias das classes, **Equação 3**.

$$F = \frac{(\bar{x}_{\text{PSD},1} - \bar{x}_{\text{PSD},2})^2}{S_{\text{PSD},1} + S_{\text{PSD},2}} \quad (3),$$

onde $\bar{x}_{\text{PSD},1}$, $\bar{x}_{\text{PSD},2}$ descrevem os valores médios das classes 1 e 2, respectivamente, e $S_{\text{PSD},1}$, $S_{\text{PSD},2}$ descrevem os desvios padrões para as classes. ^{126; 127}

Esta metodologia permite selecionar descritores que individualmente possuem capacidade estatisticamente significativa para diferenciar as duas classes de antifúngicos (ativos e inativos). Foram selecionados os descritores com valores de peso de Fisher superiores a 0,25 (média mais duas vezes o desvio padrão de todos os 412 descritores), os quais foram reunidos, escalonados e utilizados em estudos de reconhecimento padrão supervisionado (KNN e SIMCA), disponíveis no programa PIROUETTE 4.0.

Para a construção dos modelos de QSAR clássico foi realizada uma estratégia diferente, na qual os 412 descritores 2D foram utilizados para construir modelos de Regressão Linear Múltipla (RLM) contendo até 4 descritores, como disponível no programa MOBYDIGS 1.0. ¹²⁸ Para tal, 10 populações contendo 2000 indivíduos (modelos) cada, evoluíram por cerca de 150 gerações através de mutações e permutações, de forma que os indivíduos mais aptos (melhores modelos) fossem identificados. Ao final desse processo, os melhores modelos que seguiram os parâmetros RQK foram selecionados de acordo com o valor de q^2 . ¹²⁹ Os parâmetros RQK são critérios que permitem a rejeição de modelos que apresentassem alta colinearidade entre os descritores e super-ajuste: Regra QUIK (0,005), regra do Q^2 Assintótico (-0,005), regra da redundância R^P (0,1) e regra do super-ajuste R^N (0,01). ^{129; 130}

Os descritores contidos nos 10 melhores modelos de cada população foram reunidos, autoescalonados e utilizados para análise supervisionada (PLS), disponível no programa PIROUETTE 4.0. O número ótimo de PCs a ser empregado na construção dos modelos foi determinado através do método de validação cruzada (q^2) LOO, do inglês *Leave-One-Out*.

2.5.4 Holograma QSAR

O módulo de Holograma QSAR da plataforma SYBYL X 1.1 foi empregado para a geração de modelos de QSAR baseado em fragmentos para a série de derivados azólicos apresentados na **seção 2.4.2**.

Os modelos de HQSAR são particularmente sensíveis a variações de parâmetros que especifiquem o tamanho, tipo de fragmento codificado e o comprimento do holograma molecular. Por este motivo, inicialmente foi avaliada a influência do tipo de fragmento (distinção de fragmento) sobre os parâmetros estatísticos, limitando o tamanho dos fragmentos entre 4 e 7 átomos para toda a série de tamanhos padrões de holograma (53-401). Em seguida, foi avaliada, para os melhores modelos, a interferência da variação no tamanho dos fragmentos (2-5, 3-6 e 5-8) sobre parâmetros estatísticos. A consistência interna (q^2), calculada por validação cruzada LOO, foi utilizada para avaliar a consistência interna dos modelos de HQSAR, enquanto sua capacidade preditiva foi avaliada através de validação externa (**vide seção 2.5.5.2**).

2.5.5 Validação dos modelos

Os modelos classificatórios e quantitativos foram avaliados utilizando validação cruzada (LOO) e validação externa, conforme mostrado abaixo.

2.5.5.1 Modelos classificatórios

O menor número (K) de vizinhos no modelo KNN e o número de PCs do modelo SIMCA que possam classificar de forma correta os compostos do grupo treino foram determinados por validação cruzada LOO (*Leave-One-Out*). Para tal, um composto do grupo treino é retirado do modelo inicial e um modelo novo é produzido com os compostos restantes. Logo em seguida o composto retirado é classificado pelo novo modelo. Este procedimento é repetido “n” vezes até que todos compostos do grupo treino tenham sido testados um por vez. No KNN esse procedimento é repetido para diferentes números de vizinhos até que seja identificado o número ótimo de vizinhos capaz de produzir menor erro na classificação dos compostos.¹³¹ Para SIMCA esse procedimento é feito para um número crescente de PCs até a identificação do modelo com maior consistência interna.¹⁰²

A partir do banco de dados contendo 33 derivados azólicos (**TABELA 2.4**) foram selecionados aleatoriamente 24 compostos para compor o grupo treino (**TABELA 2.5**),

empregados para a construção dos modelos quimiométricos. Os 9 derivados azólicos restantes (**TABELA 2.5**) fizeram parte do conjunto teste, sendo utilizados para avaliar a capacidade preditiva do modelo classificatório. Os compostos do grupo treino e teste foram selecionados aleatoriamente, porém com o cuidado para que o número de compostos em cada subgrupo (ativo e inativo) fosse proporcional. Dessa forma, o conjunto treino é formado por 13 compostos ativos e 11 inativos, enquanto o conjunto teste apresenta 5 compostos ativos e 4 inativos. A capacidade preditiva dos modelos foi expressa como a percentagem de acerto na classificação da atividade biológica.

2.5.5.2 Modelos de QSAR clássico e HQSAR

A capacidade preditiva dos modelos foi avaliada utilizando um conjunto de 9 compostos (equivalente à 27,3% do conjunto total de compostos), selecionado aleatoriamente, com base na sua distribuição de atividade biológica (**TABELA 2.9**).

O coeficiente de ajuste preditivo (r_{pred}^2) foi calculado de acordo com Schuurmann e colaboradores¹³² e está apresentado na **Equação 4**.

$$r_{pred}^2 = 1 - \frac{\sum_{i=1}^N (y_i^{pred} - y_i)^2}{\sum_{i=1}^N (y_i - y_{médio_teste})^2} \quad (4),$$

onde N é o número de moléculas “ i ”; y_i é o valor da atividade biológica (pCIM) para as “ i ” moléculas; y_i^{pred} é o valor predito de pCIM das “ i ” moléculas; $y_{médio_teste}$ é a média dos valores de pCIM das moléculas do grupo teste.¹³³

2.6 Resultado e discussão

Os ensaios celulares desempenham um papel fundamental para o planejamento de fármacos.⁸² Nesta perspectiva, compostos com atividade biológica conhecida e disponibilidade no mercado podem fornecer informações relevantes, de forma rápida e precisa, para a descoberta e otimização de novas moléculas bioativas.

Para o propósito deste trabalho, diversos derivados azólicos utilizados como agrotóxicos juntamente com fármacos aplicados clinicamente foram empregados em estudos *in silico* com o objetivo de identificar relações entre a estrutura química e sua atividade biológica. Para alcançar esse objetivo é imprescindível que a propriedade biológica medida para todos os antifúngicos seja comparável entre si. Embora alguns derivados azólicos disponíveis clinicamente (ex. itraconazol, fluconazol, voriconazol) tenham a atividade frente *Cryptococcus* determinada em diversos estudos,^{134; 135; 136; 137} não existem tais informações para os demais compostos estudados nesse trabalho. A fim de solucionar esse problema, a atividade antifúngica de todos os 33 derivados azólicos foi determinada através de ensaios padronizados descritos abaixo.

2.6.1 Ensaios de susceptibilidade por difusão de disco

Diante da falta de padronização para a maioria dos antifúngicos deste estudo, o protocolo M44-A2, descrito pelo CLSI, foi utilizado para avaliar o perfil de atividade antifúngica dos derivados azólicos frente ao isolado clínico *C. gattii*, resistente à fluconazol, e cepa controle *C. neoformans* var. *grubii* (ATCC 90112), sensível a fluconazol, através da técnica de difusão de disco. A concentração inicial testada (25µg/disco correspondente a 8,2 mmol/L) foi similar àquela preconizada para fluconazol frente à *Candida albicans* (ATCC 90028), mas também foram testadas concentrações equivalentes ao dobro da concentração padrão (50µg/disco correspondente a 16,4 mmol/L) e a metade da mesma (12,5µg/disco correspondente a 4,1 mmol/L).

Através dos resultados obtidos por difusão de disco é possível traçar o perfil de atividade dos 33 derivados azólicos em relação às cepas de *Cryptococcus* (TABELA 2.4). Por exemplo, os antifúngicos Fluconazol, Flutrimazol, Flutriafol não apresentaram inibição do crescimento fúngico frente à cepa *C. gattii* em nenhuma das concentrações, Azaconazol e

Procloraz em duas concentrações (4,1 e 8,2 mmol/L), ao passo que Myclobutanil e Triadimenol apenas na menor concentração testada (4,1 mmol/L). Em contrapartida, os antifúngicos Cetoconazol, Climbazol, Ciproconazol, Hexaconazol e Metconazol apresentaram maiores valores de halo de inibição (> 50 mm) em relação a *C. gattii*. Considerando a cepa sensível à fluconazol (*C. neoformans*), apenas o Flutrimazol não apresentou halo de inibição, ao passo que os antifúngicos Bromuconazol, Cetoconazol, Climbazol, Ciproconazol, Difenoconazol, Diniconazol, Epoxiconazol, Etaconazol, Fembuconazol, Fluquinconazol, Flusilazol, Hexaconazol, Metconazol, Penconazol e Prothioconazol apresentaram halo de inibição > 50 mm. Além disso, apenas o antifúngico Procloraz não apresentou halo de inibição na menor concentração frente a cepa *C. neoformans*.

TABELA 2.4 – Diâmetro médio dos halos de inibição (milímetros) dos 33 derivados azólicos frente à *C. gattii* e *C. neoformans*

Antifúngico	<i>C. gattii</i>			<i>C. neoformans</i>		
	Concentrações em mmol/L			Concentrações em mmol/L		
	16,4	8,2	4,1	16,4	8,2	4,1
Azaconazol	23,3	0,0	0,0	47,3	46,7	41,0
Bifonazol	14,3	12,3	11,3	25,3	25,3	23,3
Bromuconazol	47,3	43,3	31,3	66,7	64,3	56,7
Cetoconazol	59,0	44,0	37,3	60,7	55,0	48,7
Climbazol	57,7	46,7	41,0	68,0	63,3	58,7
Clotrimazol	35,0	33,3	32,3	47,0	45,3	43,3
Ciproconazol	60,7	57,7	57,0	71,0	68,0	66,3
Diclobutrazol	29,3	24,7	14,0	45,7	41,3	33,7
Difenoconazol	36,0	33,3	31,7	51,7	50,7	49,3
Diniconazol	38,7	36,7	34,3	53,3	53,7	52,0
Econazol	30,3	28,0	27,0	41,7	41,7	39,3
Epoxiconazol	34,7	35,0	34,3	59,3	58,0	58,0
Etaconazol	38,0	31,0	23,3	61,7	57,3	47,3
Fembuconazol	28,0	29,0	26,0	50,0	48,0	45,0
Fluconazol	0,0	0,0	0,0	41,0	33,0	26,7
Fluotrimazol	0,0	0,0	0,0	0,0	0,0	0,0
Fluquinconazol	38,3	37,0	35,0	52,7	52,0	48,7
Flusilazol	26,0	26,3	19,7	51,7	43,0	39,0
Flutriafol	0,0	0,0	0,0	40,3	30,3	22,0
Hexaconazol	63,7	54,0	55,3	76,3	73,0	67,0
Imazalil	29,7	21,7	12,0	44,7	38,7	29,3
Itraconazol	22,3	21,0	19,3	22,7	21,3	22,7
Metconazol	54,3	51,7	44,0	67,7	64,7	58,7
Miconazol	39,3	33,3	26,7	37,7	35,0	34,3
Myclobutanil	25,7	13,3	0,0	47,7	38,7	28,0
Penconazol	38,7	33,0	22,0	52,0	43,0	38,3
Prochloraz	17,0	0,0	0,0	23,3	18,3	0,0
Propriconazol	36,0	35,7	35,7	49,7	49,3	49,3
Prothioconazol	42,7	42,3	40,3	57,7	54,7	46,0
Sulconazol	39,0	38,0	37,7	43,7	40,7	40,3
Tebuconazol	39,3	33,3	26,7	65,7	61,0	55,0
Triadimenol	18,0	11,0	0,0	45,7	35,3	26,7
Triticonazol	26,7	26,0	19,3	49,0	45,0	37,3

Os resultados mostram que alguns antifúngicos possuem perfis de atividade bastante diferenciados em relação à *C. gattii* e *C. neoformans*, como pode ser evidenciado através da diferença nos diâmetros dos halos de inibição do antifúngico Fluconazol (FIGURA 2.5 A, B e C).

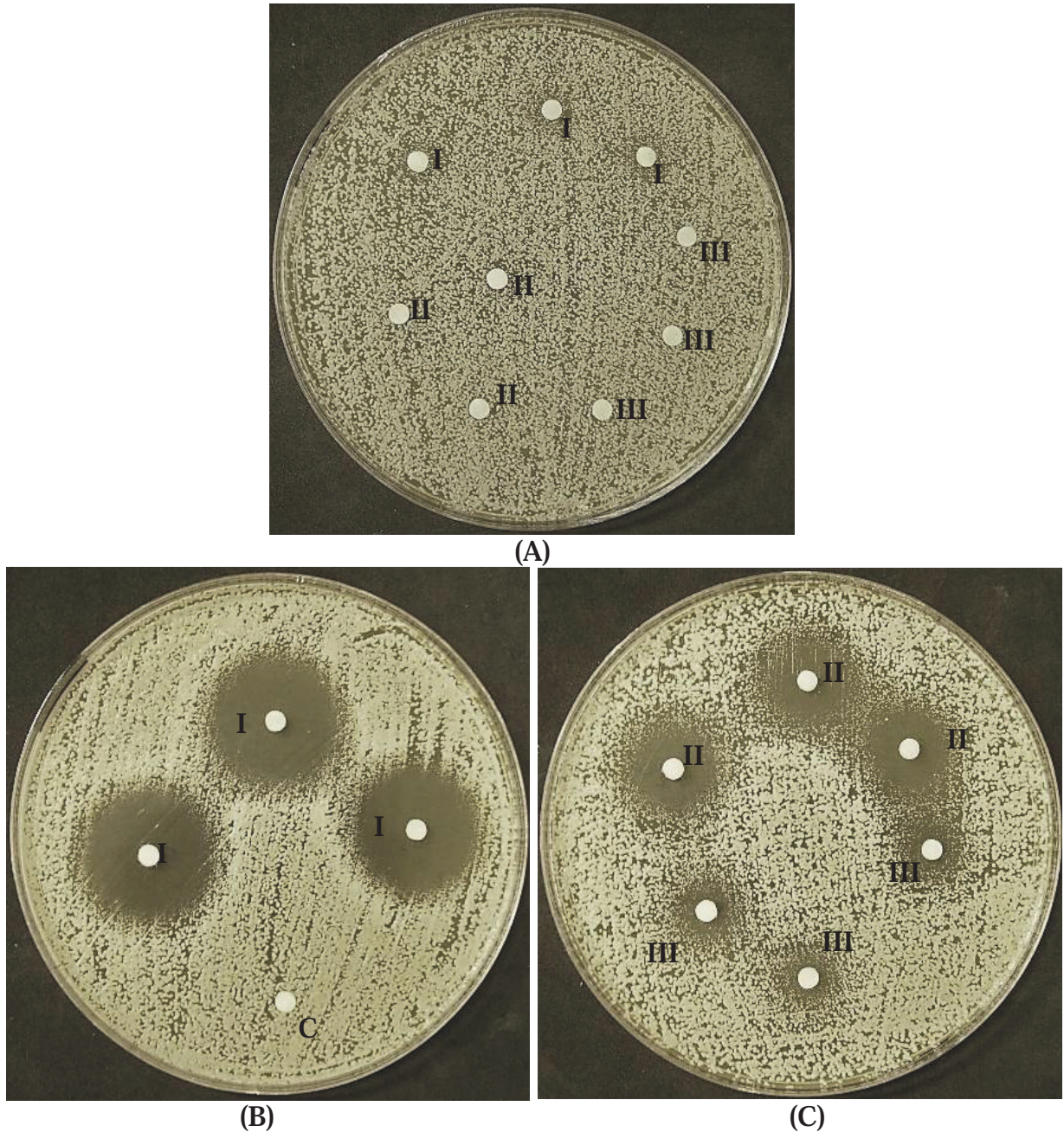


FIGURA 2.5 – Halos de inibição para o antifúngico Fluconazol em relação à *C. gattii* (A) e *C. neoformans* (B e C). Os símbolos I, II e III representam as concentrações de antifúngicos contidas nos discos: 16,4 mmol/L; 8,2 mmol/L e 4,1 mmol/L, respectivamente. O símbolo “C” representa o controle negativo do experimento contendo apenas DMSO, o solvente utilizado para a dissolução dos antifúngicos.

Fonte: O autor

De modo geral, *C. gattii* apresentou menor suscetibilidade aos antifúngicos testados, sugerindo que essa cepa apresenta bomba de efluxo ou mutações no alvo molecular dos derivados azólicos (Maiores detalhes sobre esses mecanismos de resistência serão discutidos no capítulo 3).

Considerando que o foco principal desse trabalho é a investigação das propriedades físico-químicas e estruturais que determinam a potência de antifúngicos azólicos frente a cepas resistentes de *Cryptococcus*, maior atenção foi dada aos resultados biológicos obtidos para *C. gattii*. De fato, pode-se observar que os antifúngicos testados podem ser claramente divididos quanto ao perfil de atividade apresentados, os quais variaram desde não inibição (FIGURA 2.6A) até alta atividade (FIGURA 2.6B).

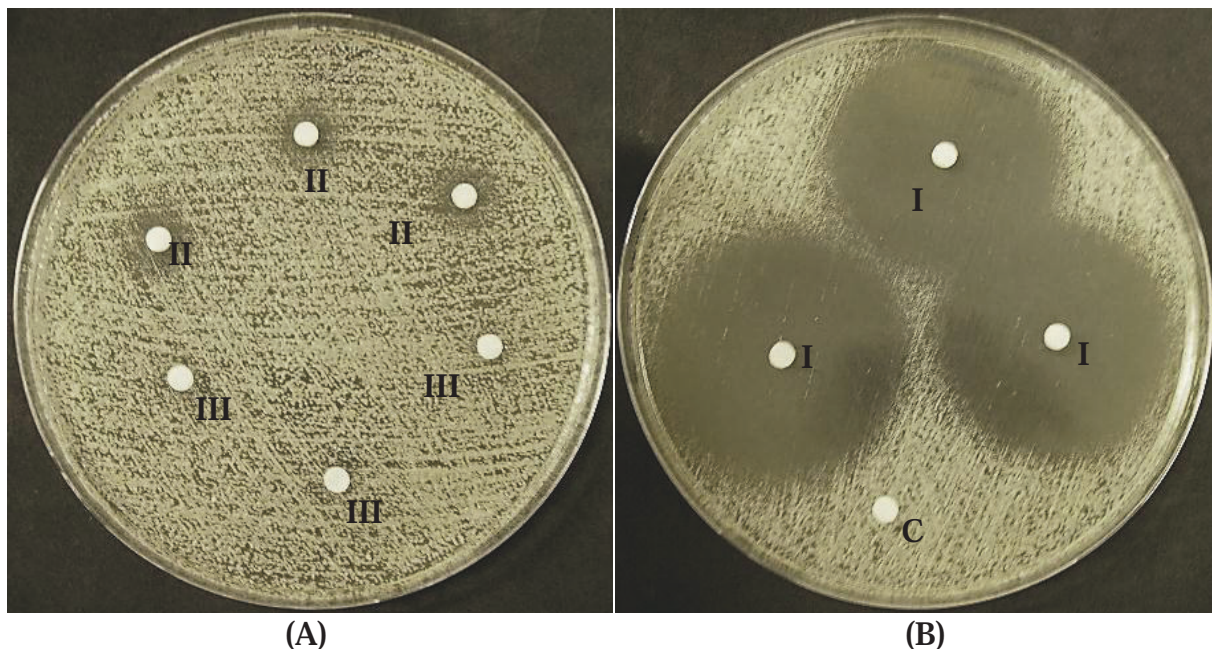


FIGURA 2.6 – Halos de inibição para o antifúngico Flutrimazol (A) e Hexaconazol (B) frente à *C. gattii*. Os símbolos I, II e III representam as concentrações de antifúngicos contidas nos discos: 16,4 mmol/L; 8,2 mmol/L e 4,1 mmol/L, respectivamente. O símbolo “C” representa o controle negativo do experimento contendo apenas DMSO, o solvente utilizado para a dissolução dos antifúngicos. Fonte: O autor

Estes resultados sugerem que diferenças estruturais e físico-químicas nestes antifúngicos podem estar relacionadas com a variação do perfil de atividade apresentados frente às cepas de *Cryptococcus*, muito embora a avaliação da atividade antifúngica por difusão em disco não permita mensurar a potência dos derivados azólicos estudados. No entanto, a investigação das propriedades químicas que separam compostos ativos dos inativos pode ser de grande utilidade para nortear estudos preliminares de identificação de novos

agentes antifúngicos, potencialmente úteis para o desenvolvimento de antifúngicos. Por essa razão, os dados oriundos dos testes de difusão em disco foram empregados para desenvolver modelos classificatórios descritos a seguir.

2.6.2 Modelos classificatórios

2.6.2.1 Conjunto de dados

Como um dos objetivos do planejamento de fármacos é identificar compostos com maior atividade, ou seja, cuja quantidade utilizada seja a menor possível em relação ao efeito desejado, optou-se por utilizar os dados obtidos com discos na concentração de 4,1 mmol/L para classificar os antifúngicos em ativos e inativos (**TABELA 2.4**). Essa divisão depende de um critério subjetivo, pois é o usuário/pesquisador que define o diâmetro de halo de inibição a ser empregado para divisão dos compostos. Considerando que modelos classificatórios devem ter um número aproximadamente igual de compostos em cada classe, utilizou o histograma de frequência mostrado na **FIGURA 2.7** para definir como ponto de corte halos de inibição de 25 mm.

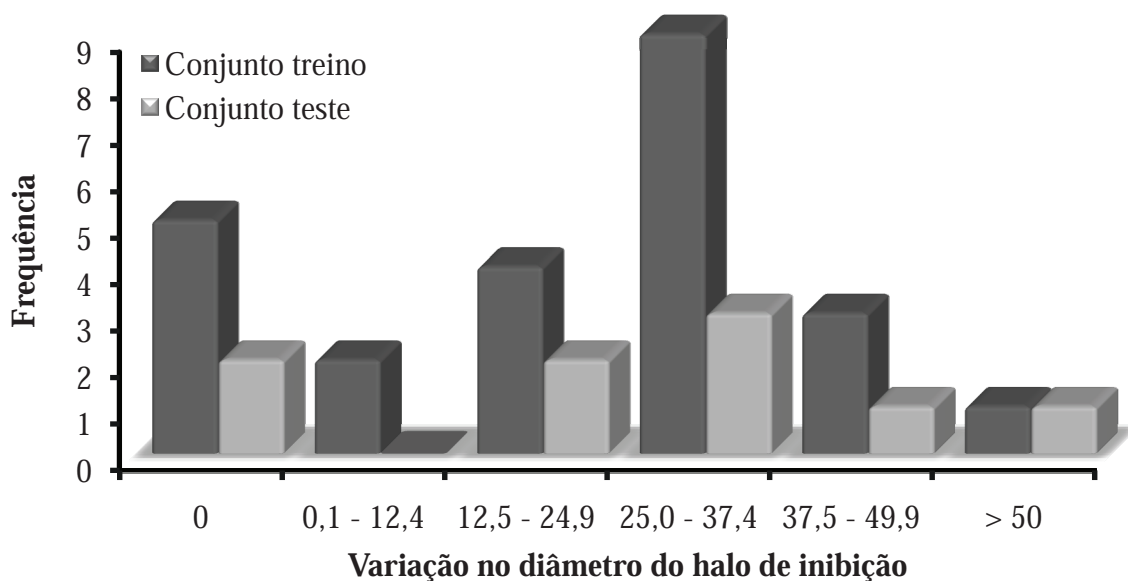


FIGURA 2.7 – Histograma de frequência dos antifúngicos em relação à variação no diâmetro de halo de inibição frente a *C. gattii* na menor concentração (4,1 mmol/L)

Portanto, compostos que apresentaram diâmetro médio de halo de inibição maiores que 25 mm foram classificados como ativos (18 antifúngicos ou 54,5%), caso contrário como inativos (15 antifúngicos ou 45,5%). Uma vez que geração de modelos estatísticos consistentes é dependente da composição dos grupos treino e teste,^{138; 139} os antifúngicos foram divididos em grupo treino, contendo 24 compostos (13 ativos e 11 inativos) e grupo teste, constituído de 9 compostos (5 ativos e 4 inativos) (TABELA 2.5).

TABELA 2.5 – Divisão dos antifúngicos em grupo treino e teste

Compostos ativos		Compostos inativos	
Treino	Teste	Treino	Teste
Bromuconazol	Difenoconazol	Azaconazol	Fluconazol
Cetoconazol	Fembuconazol	Bifonazol	Flusilazol
Climbazol	Hexaconazol	Diclobutrazol	Myclobutanil
Clotrimazol	Metconazol	Etaconazol	Triticonazol
Ciproconazol	Propriconazol	Fluotrimazol	
Diniconazol		Flutriafol	
Econazol		Imazalil	
Epoxiconazol		Itraconazol	
Fluquinconazol		Penconazol	
Miconazol		Prochloraz	
Prothioconazol		Triadimenol	
Sulconazol			
Tebuconazol			

2.6.2.2 Estudos de reconhecimento padrão supervisionados

O desenvolvimento de modelos classificatórios depende da descrição apropriada da variação estrutural no conjunto de dados, o que pode ser feito através de diversos descritores físico-químicos e estruturais que contém informações relevantes sobre a atividade biológica. Levando-se em consideração que descritores 2D não requerem a determinação da

conformação bioativas das moléculas e proporcionam uma descrição rápida e objetiva das características relevantes para atividade biológica, este tipo de descritor foi selecionado.

Assim, 412 descritores físico-químicos e estruturais, de um total de 2489 descritores disponíveis, foram calculados com auxílio do programa DRAGON 5.5. Logo em seguida, procedeu-se a seleção dos descritores, através do peso de Fisher.^{102; 140}

Esta estratégia resultou na seleção de 10 descritores, que foram reunidos e auto-escalados, visando assegurar que as influências relativas do diferentes descritores em todos os cálculos sejam independentes de suas unidades.¹⁴¹ Os descritores selecionados foram, em seguida, utilizados para o desenvolvimento de modelos de reconhecimento padrão supervisionado (KNN e SIMCA).

O método KNN classifica os compostos de acordo com a distância euclidiana entre eles. Um aspecto importante para o desenvolvimento de modelo de KNN é a determinação do menor número (K) de vizinhos capazes de classificar corretamente amostras desconhecidas. Nesse estudo, o número ótimo de vizinhos é 10 (**FIGURA 2.8**). Porém, é importante frisar que quanto mais próximo for o número ótimo de vizinhos do número total de compostos do grupo treino, maior a distância entre os compostos vizinhos no espaço multivariado;¹⁴²

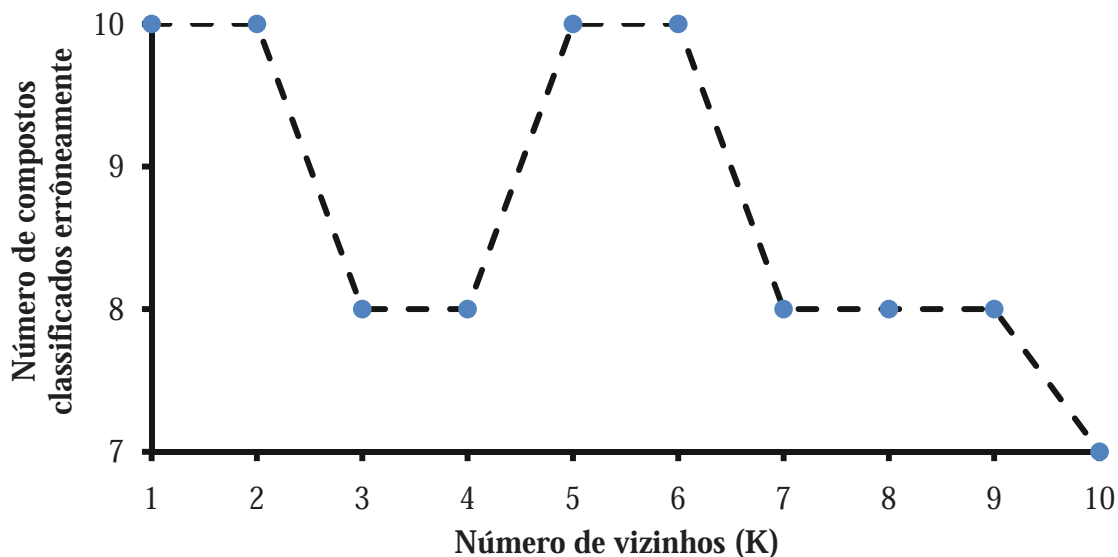


FIGURA 2.8 – Número de compostos classificados erroneamente para cada número de vizinhos utilizados para classificação. Fonte: Elaborado pelo próprio autor

O método KNN demonstrou consistência interna razoável ao acertar a classificação de 69% dos compostos ativos e 73% dos compostos inativos do grupo treino. Grau de acerto semelhante foi observado para o conjunto teste onde 60% dos compostos ativos e 75% dos compostos inativos foram classificados corretamente.

Um dos fatores que pode contribuir para a qualidade moderada do modelo KNN é a alta correlação entre os descritores utilizados para sua construção. De fato, é possível observar que alguns dos descritores destacados na **TABELA 2.6** são intercorrelacionados e podem acrescentar ruído à análise de KNN. Em virtude da falta de ferramentas diagnósticas¹⁴³ que permitam avaliar esta hipótese e identificar as razões pelas quais o método KNN não classifica corretamente alguns dos compostos, o conjunto de dados foi submetido à análise de reconhecimento padrão baseado em projeções, como o método SIMCA.

TABELA 2.6 – Correlação entre os descritores utilizados no modelo KNN

	nCL	PJI2	MATS8p	EEig05d	BEHm2	BEHm3	BEHe3	JGI10	nCs	C-008
nCL	1,00	0,25	-0,03	0,37	0,66	0,12	-0,13	0,22	0,29	0,40
PJI2		1,00	0,27	0,38	0,33	0,32	0,44	0,55	-0,15	0,19
MATS8p			1,00	0,44	0,32	0,74	0,51	0,04	-0,11	0,13
EEig05d				1,00	0,57	0,66	0,62	0,48	0,05	0,38
BEHm2					1,00	0,51	0,20	0,46	0,21	0,39
BEHm3						1,00	0,79	0,30	-0,04	0,18
BEHe3							1,00	0,40	-0,20	0,19
JGI10								1,00	0,07	0,06
nCs									1,00	0,35
C-008										1,00

Diferentemente do KNN, o método SIMCA baseia-se no desenvolvimento de modelos utilizando componentes principais (PCs) para determinar a categoria dos compostos. A utilização de PCs reduz a dimensionalidade dos modelos favorecendo o tratamento de dados multivariados.

No método SIMCA, os compostos de cada classe são projetados em subespaços diferentes formados pelas PCs de cada classe e o composto com categoria desconhecida é classificado de acordo com o desvio da distância do mesmo em relação ao subespaço formado pelas PCs e sua distância em relação às fronteiras de cada classe.^{144; 145} Isto torna a classificação dos compostos mais realista, em comparação ao KNN, pois o método SIMCA é capaz de identificar se os compostos fazem parte de uma das classes disponíveis, se não faz

parte de nenhuma das classes ou se não é possível determinar sua classe, uma vez que os resultados são expressos em termos probabilísticos.¹⁴⁶

O modelo SIMCA, construído com os mesmo descritores utilizados no modelo KNN, apresenta índice de acerto de 78% para os compostos ativos e 87% para os compostos inativos do grupo treino. Este resultado mostra que a utilização de PCs pode proporcionar resultados mais satisfatórios na classificação dos compostos. No entanto, a análise da distância entre as classes, que permite estimar a diferenciação entre as classes, é de apenas 0,71. Segundo Brereton (2009), duas classes podem ser consideradas separáveis quando a distância interclasse for maior do que 3.¹⁴⁷ Isto significa que apesar dos bons resultados de classificação dos compostos, o modelo ainda não possui capacidade para diferenciar as duas classes de forma adequada.

Contudo, no método SIMCA é possível identificar quais descritores mais contribuem para a separação das classes e quais tem grande influência para ambas as classes, utilizando para isso o poder discriminante (FIGURA 2.9) e poder de modelagem (FIGURA 2.10) dos descritores.¹⁴² Dessa forma, a análise desses fatores guiou a exclusão de descritores que pouco contribuem para a classificação correta dos compostos, proporcionando assim a otimização do modelo SIMCA que alcançou índice de acerto de 85% para os compostos ativos do treino e 100% para os compostos inativos.

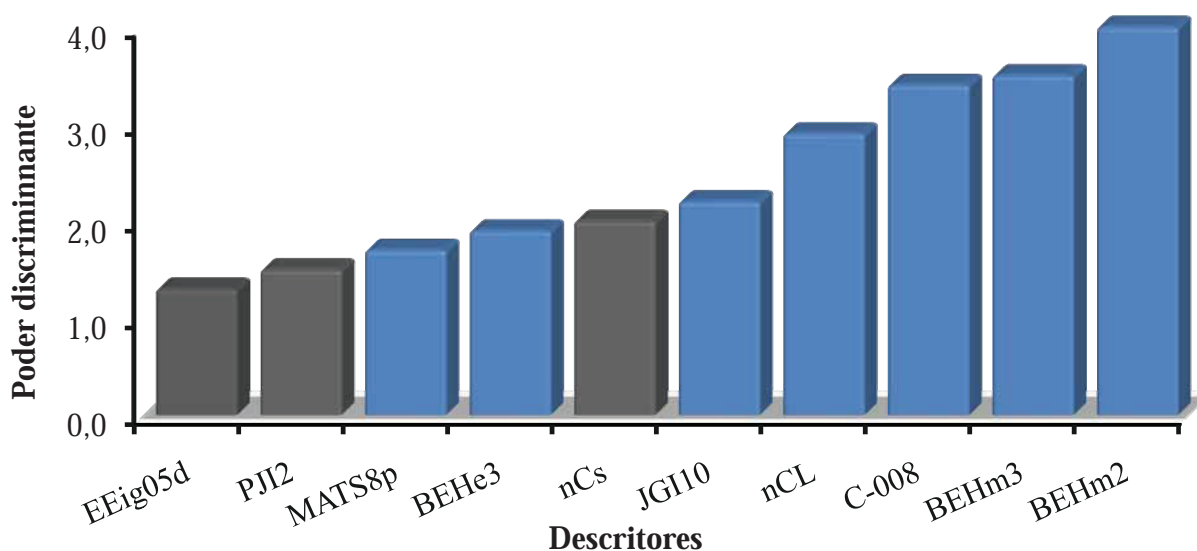


FIGURA 2.9 – Coeficientes de poder discriminantes para os descritores presentes no modelo de SIMCA inicial. Destaque para os descritores que fizeram parte do modelo final (barras em azul)

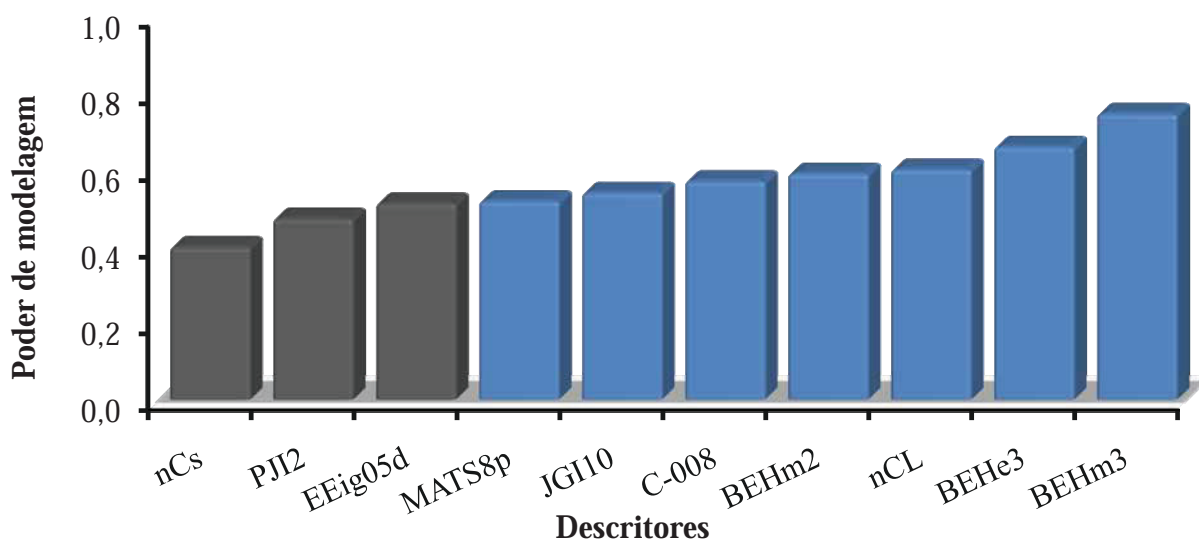


FIGURA 2.10 – Coeficientes de poder de modelagem para os descritores presentes no modelo inicial de SIMCA. Destaque para os descritores que fizeram parte do modelo final (barras em azul)

Uma breve descrição sobre os descritores encontrados no modelo final está disponível na TABELA 2.7.

TABELA 2.7 – Descritores presentes no modelo final de SIMCA

Descritores	Significado
C-008	Fragmento centrado no átomo (CHR ₂ X)*
nCL	Número de átomos de cloros
BEHm2	Maior autovalor de índice 2 da matriz de Burden, ajustado pela massa atômica
BEHm3	Maior autovalor de índice 3 da matriz de Burden, ajustado pela massa atômica
MATS8p	Autocorrelação de Moran calculado entre átomos separados por 8 ligações químicas ajustado pela polarizabilidade atômica
JGI10	Índice de carga topológica média calculada entre átomos separados por 10 ligações químicas
BEHe3	Maior autovalor de índice 3 da matriz de Burden, ajustado pela eletronegatividade de Sanderson

*R corresponde à carbono; X corresponde á heteroátomo (Ex. O, N, S)

O impacto desse protocolo também pode ser observado no aumento do valor de distância interclasse, que atingiu o valor de 1,86 (FIGURA 2.11).

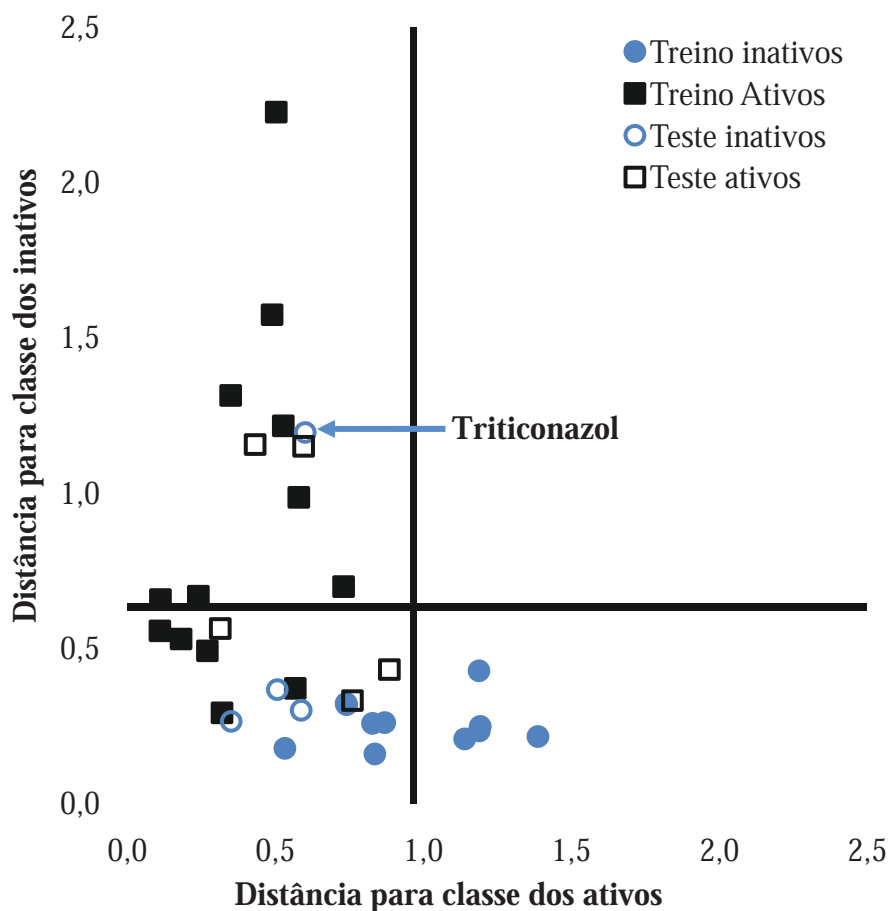


FIGURA 2.11 – Distância interclasse dos compostos ativos e inativos do grupo treino e teste segundo o modelo final de SIMCA

Os resultados obtidos sugerem que o modelo SIMCA apresenta grande consistência interna e pode ser útil na identificação de novos derivados azólicos ativos contra *C. gattii*. Visando testar essa hipótese, o modelo final foi empregado para a predição da classe dos compostos do grupo teste, apresentando resultados comparáveis ao do modelo KNN, ou seja, 60% de acerto para compostos ativos e 75% para compostos inativos.

O descritor mais importante para diferenciar as classes (JGI10) e aquele que melhor descreve a informação química contida em ambas as classes (BEHe3) foram utilizados para se fazer uma análise preliminar das propriedades físico-químicas e estruturais, desta classe de antifúngicos, que são cruciais para atividade biológica frente à *C. gattii*.

JGI10 é um índice de carga topológica média, proposto por Galvez e colaboradores, que avalia a transferência de cargas entre pares de átomos numa distância topológica k predefinida, e dessa forma, a transferência de carga global na molécula.^{148; 149}

Diversas “distâncias” entre os átomos podem ser calculadas, porém para JGI10, apenas as distâncias equivalentes à 10 ligações são consideradas, ou seja, moléculas cuja distância máxima entre os átomos é menor que 10 ligações, apresentam valor zero. De fato isso ocorre para maioria das moléculas inativas (73%), ao passo que 50% das moléculas ativas possuem valores acima de zero, sugerindo que o momento dipolo das moléculas pode contribuir para o aumento da atividade antifúngica.

O descritor BEHe3, por sua vez, é um autovalor obtido a partir de uma matriz de adjacência, cujo valor foi ajustado pela a eletronegatividade atômica.¹⁵⁰ A análise dos valores de BEHe3, para as moléculas do conjunto treino, sugere que valores extremos são encontrados apenas em compostos inativos (**FIGURA 2.12**).

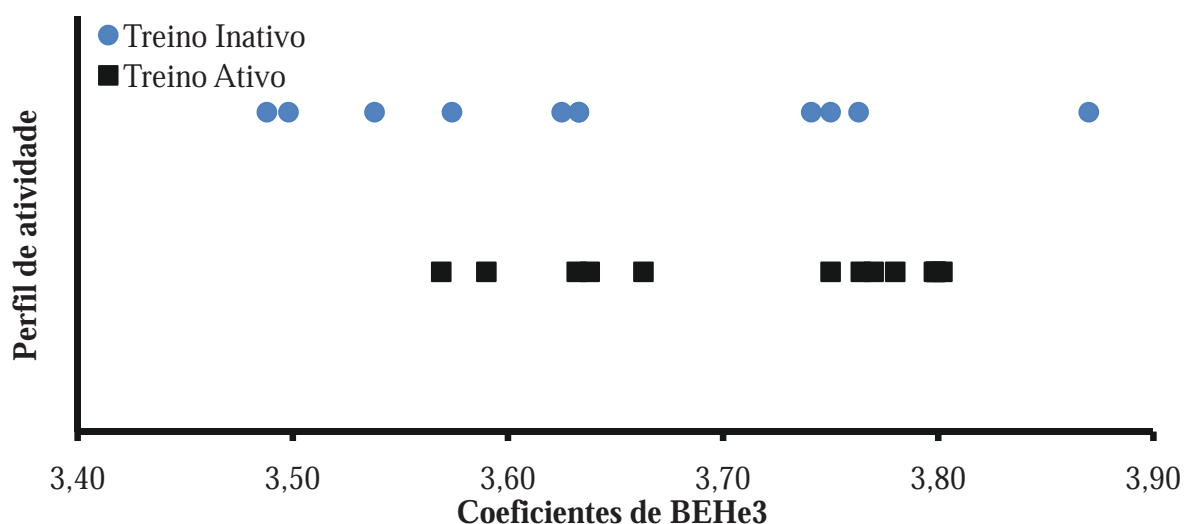


FIGURA 2.12 – Perfil de atividade versus coeficientes de BEHe3 para os antifúngicos do grupo treino e teste

Esses resultados corroboram com dados na literatura, nos quais a importância de interações eletrostáticas entre os átomos do nitrogênio heterocíclico dos derivados azólicos e o átomo de ferro da porção heme da proteína lanosterol 14 α -desmetilase é bem estabelecida.^{151; 152} Adicionalmente, Mota e colaboradores já demonstraram que esses tipos de descritor (JGI4 e BEHe3) são úteis na descrição do perfil de atividade de derivados azólicos frente ao fungo filamentososo *Moniliophthora perniciosa*.^{140; 153}

Estes achados mostram que as propriedades físico-químicas e estruturais destacadas neste trabalho podem ser importantes tanto na inibição do crescimento de fungos filamentosos quanto leveduriformes, o que pode estar relacionado com a alta identidade seqüencial encontrada no sítio ativo de lanosterol 14 α -desmetilase em diferentes espécies de fungo que ultrapassa 65% (**FIGURA 2.13**). Tomando *C. gattii* e *M. perniciosa* como exemplos, o alinhamento da seqüência completa de aminoácidos da enzima Lanosterol 14 α -desmetilase

destes dois fungos revela notável o grau conservação entre os resíduos que compõem o sítio ativo (identidade 81,3% e similaridade 88,4%) (**FIGURA 2.13 – em azul**). Por este motivo, é natural que as exigências estéricas e eletrostáticas necessárias para a inibição de lanosterol 14- α desmetilase em diferentes fungos sejam semelhantes, porém não idênticas.

<i>C. neoformans</i>	MS-----AII PQVQQL---GQVAQFFPPWFAALPTS LKVAIAVV-GIP--ALI IGLNVFQQLCLPRRKDLPPVVFHYI PWFGSAAYYGEDPYKFLFECRDKYGD LFTFILMGRRITVA	108
<i>C. gattii</i>	MS-----AII PQVQQL---GQLAQYI PPWFALPTS LKIVIAVI-GIP--AFI IGLNVFQQLCLPRTKDLPPVVFHYI PWFGSAAYYGEDPYKFLFECRDKYGD LFTFILMGRRITVA	108
<i>C. albicans</i>	M-----AIV-----ETVIDG--IN Y-----FLSLSVTQQIS ILLGVPFVY-----NLVWQYLYSLRKRDRAPLVFYI PWFGSAASYGQQPYEFFFESC RQKYGDVFSFMLLGKIMTVY	95
<i>C. tropicalis</i>	M-----AIV-----DTAIDG--IN Y-----FLSLSLTQQIT ILLVFPFIY-----NIAWQLLYSLRKRDRVPMVFYI PWFGSAASYGMQPYEFFFEC RRLKYGDVFSFMLLGKIMTVY	95
<i>M. perniciososa</i>	MSHVENLNGSTIPIVKDLWS-GYVAQAQE-----LSTRLLLLLA AINL-PVIAIVL--NVL RQVLI PRNKSEPPVVFHWLPI IGS AVSYGNDP INFFFKREKYGNVFTFILFGRRVTV A	111
<i>A. fumigatus</i>	M----GLIAIL-----D----GICKHCSTQSTWVLV-----GIGLLS I LAVSVI INVLQQLLFKNPH-EPPVVFHWF PFI GSTISYGI DPYKFFFD C RAKYGD IFTFILGKKT VY	98
<i>S. reilianum</i>	MVSTSSSSS ATASLL-DQLFALTP LDSSASVKA-LS-----LAVA--LPVLAIVL--NVA AQLLLPAPRNNPPVVFHFVPI IGS AVSYGIDPYKFFFE C RAKYGDVFTFVLLGRRITVA	108
<i>C. neoformans</i>	LGPKGNLNSLGGKISQVSAEEAYTHLTPVFGKGVVYDCPNEMLMQKKFKISGLTTESLQSYPPMITSECEDFFT----KEVGISPQKPSATL DLLKAMSELI ILTASRTLQ GKRE	223
<i>C. gattii</i>	LGPKGNLNSLGGKISQVSAEEAYTHLTPVFGKGVVYDCPNEMLMQKKFKISGLTTESLQSYPPMITSECEDFFT----KEVGISSQKPSATL DLLKAMSELI ILTASRTLQ GKRE	223
<i>C. albicans</i>	LGPKGHEFVFNALSDVSAEEAYKHLTPVFGKGVVYDCPN SRLMEQKKFAK FALTTDSFKRYVPKIREEILNYFVTDESFKLK----EKTHGVANVMKTQPEITIFTASRSLFGDEMRR	211
<i>C. tropicalis</i>	LGPKGHEFIYNAKLSDVSAEEAYTHLTPVFGKGVVYDCPN SRLMEQKKFAK FALTTDSFKTYVPKIREEVLNYFVNDVSFKTK----ERDHGVASVMKTQPEITIFTASRCLFGDEMRR	211
<i>M. perniciososa</i>	LGAQGNFIFLGGKSTVFSAAEAYTHLTPVFGKGVVYDVPNEVMEQKKFVKVGLSTENLRAVYGMIEDEVVEFIENDPTFKVYQDNDIN EWG SFDVLKILAEITILTASRTLQ GKQD VRS	231
<i>A. fumigatus</i>	LGTKGNDFILNGKLRDVC AEEVYSPLTPPVFGRHVYDCPN AKLMEQKKFVKYGLTSDALRSYVPLITDEVESFVKNSP-----AFQGHKGVFDVCKTIAEITIT YASRSLQ GKREVS	211
<i>S. reilianum</i>	LGPKGSNLVFNALHQVTAEDAYTHLTPVFGKEVVYDVPNAVMEQKKFVKVGLSDNFRVYVQIVDEVREYIKTDRPFAVQQG--NKSFSVDIFQAMSEVI ILTASRTLQ GKRE	226
<i>C. neoformans</i>	SLNGQFAKYVEDLDGGFTPLNFMFNLPLPSYKRRDEAQKAMSD FYLKIMENRRK----GESDHEHDMIENLQ S-CKYRNGVPLSDRDI AHIMIALLMAGQHTSSATSSWTLHLADR PD	338
<i>C. gattii</i>	SLNGQFAKYVEDLDGGFTPLNFMFNLPLPSYRRRDEAQKAMSD FYLKIMENRRK----GESDHEHDMIENLQ S-CKYRNGVPLSDRDI AHIMIALLMAGQHTSSATSSWTLHLADR PD	338
<i>C. albicans</i>	IFDRSFAQLYSDLDKGF TPI NFVFNLP LPHYWRRDAQKKISATYMK EIKSRRE---RGDIDPNRDLIDSLLIHSTYKDGVKMTDQEI ANLLIGILMGGQHTSASTSAWFLHLGKPH	328
<i>C. tropicalis</i>	SFDRSFAQLYADLDKGF TPI NFVFNLP LPHYWRRDAQQRKISAHYMK EIKRRRE---SGDIDPKRDLIDSLLVNSTYKDGVKMTDQEI ANLLIGVLMGGQHTSASTSAWFLHLAEQPQ	328
<i>M. perniciososa</i>	NLDKFAQLYNDLDGGFTPI NFVFNLP LPHYRKRDAQKKMSEFFQSFIEQRQ---GHSYDQDMIEALME-QRYRSGKGLKDHEI AHIMIALLMAGQHTSSATGSWALLHLADNPD	346
<i>A. fumigatus</i>	KFDSTFAELYHNLDMGFAPINFMPLWAPLPHNRKRDAQRKLTET YMEI I KARRQAGSKK DSE---DMVWNLMS-CVYKNGTPVPDEEI AHMMIALLMAGQHTSSSTASWIVLRLATRPD	327
<i>S. reilianum</i>	GLDKSFAQLYHDLDSGF TPI NFVFNLP LPSNFKRDAQKKMSQFYQDIVAKRRAAGHAGDGRGENDMIAALME-QRYKNGRALSGVEI AHMMIALLMAGQHTSSATSSWAFRLASRPE	345
<i>C. neoformans</i>	VVEALYQEQQKQLGNPDGTFRDYRYEDLK-ELPIMDSI IRETLRMHAPIHSIYRKVLSDI PVPPSL SAPSENGQYI I PKGHY IMAAPGVSQMDPRIWQDAKVWN PARWHDEKGF AAAAMV	457
<i>C. gattii</i>	IVEALYQEQQKQLGNPDGTFRDYRYEDLK-ELPIMDSI IRETLRMHAPIHSIYRKVLSDI PVPPSLAAPSENGQYI I PKGHY IMAAPGVSQMDPRIWQDAKVWN PARWHDEKGF AAAAMA	457
<i>C. albicans</i>	LQDVYIYQEVVELLKEKGGDLNDLTYEDLQ-KLPSVNNTI KE TLRMHMPLHSI FRKVTNPLRIPETN-----YIVPKGHYVLSVPGYAHTSERYFDNPEDFD PTRWDTAAAKANSVSF	439
<i>C. tropicalis</i>	LQDDLYEELTNLLKEKGGDLNDLTYEDLQ-KLPLVNNTI KE TLRMHMPLHSI FRKVMNPLRVNPTK-----YVI PKGHYVLSVAGYAHTSDRWFEHPEHFNRRWESDDTKASAVSF	439
<i>M. perniciososa</i>	IAEALYQEQQKHSNPDGSRFSMSYEELR-ELPVLD SVIRETLRVHPPHSI IMRYVRDDVPVPPTLSAPSEDRTYI VPKGHYVLSVPAVSQMDPLIWRNPEKWD PARWSDPEGIAAQAL K	465
<i>A. fumigatus</i>	IMEELYQEQIRVLG---SDLPPLTYDNLQ-KLDLHAKVI KE TLRRLHAPIHSI I RAVKNPMAVDGTS-----YVIPTSHNVLSSPGVTARSEEHPNPLEWNP HRWDEN--IAASA--	431
<i>S. reilianum</i>	IEELYAEQVKVYG---AELRELDYETQKTSVPLLDVAVK E TLRRLHPPHSI IMRYVKTDLAVPPTLASPAREEAYVI PKGHY IMAAPGVSQVDPALWKGADAFDEHRWLDAATAAAMQ--	460
<i>C. neoformans</i>	QYT--KAEQVYDYGFGSVSKGTESPYQPFAGRHRRCVGEQFAYTQLSTIFTYVVRNFTLK--LAVPKFPETNYRTMIVQPN-NPLVTFTLRN-AEVKQEV-	550
<i>C. gattii</i>	QYT--KAEQVYDYGFGSVSKGTESPYQPFAGRHRRCVGEQFAYTQLSTIFTYVVRNFTLK--LAVPKFPETNYRTMIVQPT-NPLVTFTLRN-AEVKQEV-	550
<i>C. albicans</i>	NSS---DEVYDYGFGKVS KGVSSPYLPFGGGRHRCIGEQFAYVQLGTILTTFFVYNLRWT--IDGYKVPDPDYSSMVVLPTEPAEIIWEKRE----TCMF-	528
<i>C. tropicalis</i>	NS---ED-TVYDYGFKIS KGVSSPYLPFGGGRHRCIGEQFAYVQLGTILTTYIYNFKWR--LNGDKVPD VDYQSMVTLPLEPAEIVWEKRD----TCMV-	528
<i>M. perniciososa</i>	TYVDENGEKI DYGFGAVSKGTESPYQPFAGKHRCI GEQFAYLQLGTLLSTFVRRMELR--IPTG-VPAQNYHTMITMPK-TPRRIHYRRR----KAN--	555
<i>A. fumigatus</i>	----EDDEKVYDYGGLVSKGTNSPYLPFGAGRHRCI GEQFAYLQLGTITAVLVRLFRFRNLPGVDGIPD TDYSSLFSKPLGRSFEVEFEKRE-SATKA---	523
<i>S. reilianum</i>	--DAGDD-KE DFGFGMISTGANSPYLPFGAGRHRCI GEQFAYLQIGVTVAVFVRMFKWQ--LDTK-FPDPDYQSMVVLPS-KQGSMIKLTVRGE-KLDLE	552

FIGURA 2.13 – Alinhamento sequencial dos aminoácidos da enzima Lanosterol 14 α -desmetilase de fungos leveduriformes (*C. neoformans*, *C. gattii*, *C. albicans* e *C. tropicalis*) e fungos filamentosos (*M. perniciososa*, *A. fumigatus* e *S. reilianum*). Resíduos conservados em todas as sequências estão destacados em cinza e resíduos pertencentes ao sítio ativo em azul. Alinhamento feito no CLUSTAL X 2.1 e análise da identidade e similaridade no BIOEDIT 7.05, através da matriz BLOSUM62

Embora promissores, os modelos classificatórios desenvolvidos não permitem prever a potência dos antifúngicos e, portanto, tem utilidade restrita no que se refere ao planejamento de novos derivados azólicos, ativos contra cepas de *C. gattii* resistentes à fluconazol. Visando contornar essa limitação, decidiu-se realizar ensaios de atividade por microdiluição em caldo para determinar a CIM dos antifúngicos avaliados até esse momento e, a partir da informação biológica medida (discreta), desenvolver modelos de QSAR que possam ser úteis para identificar fatores físico-químicos e estruturais relevantes para sua atividade antifúngica.

2.6.3 Ensaios de susceptibilidade por microdiluição em caldo

Em geral, antifúngicos azólicos não apresentam os pontos extremos de CIM (100% de inibição) claramente distinguíveis, por esta razão foram utilizados pontos finais menos rigorosos (aceitando algum aparecimento de crescimento fúngico), como pontos finais equivalentes a 50% ou mais de inibição do crescimento fúngico como parâmetro para definir a CIM dos antifúngicos.^{134; 136} Exemplo disso, pode ser observado na **FIGURA 2.14**, que mostra a redução do crescimento de *C. neoformans* e *C. gattii* em consequência do aumento da concentração de fluconazol.

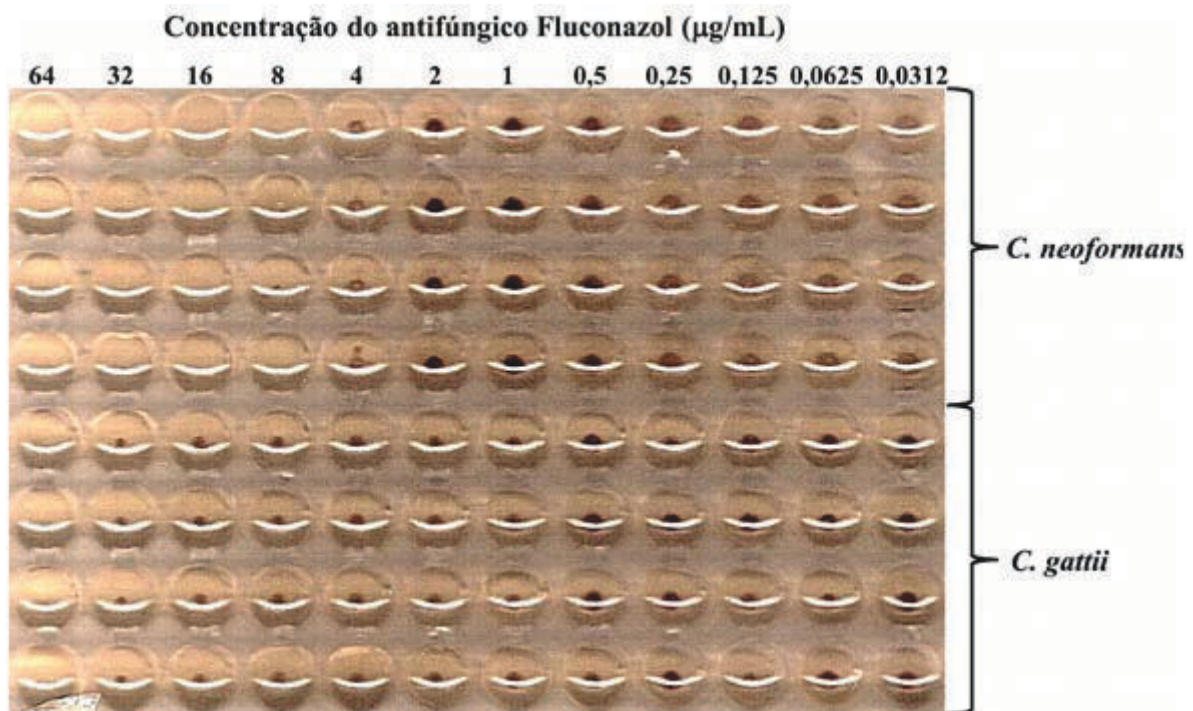


FIGURA 2.14 – Concentração inibitória mínima do antifúngico Fluconazol frente a *C. neoformans* (4 µg/mL) e *C. gattii* (32 µg/mL). Fonte: O autor

A concentração dos antifúngicos avaliada variou entre 64 até 0,031 $\mu\text{g/mL}$, entretanto, a fim de verificar a possibilidade de identificar compostos mais potentes, os antifúngicos que apresentaram CIM iguais a 0,031 $\mu\text{g/mL}$ tiveram seus experimentos refeitos e as concentrações iniciais (64 $\mu\text{g/mL}$) diluídas em 4 vezes para obter faixa de concentração do antifúngico entre 16 e 0,0078 $\mu\text{g/mL}$. Esta estratégia foi bem sucedida, uma vez que os antifúngicos Difenoconazol e Sulconazol apresentaram CIM equivalente à 0,0156 $\mu\text{g/mL}$ frente *C. neoformans* (**TABELA 2.8**).

Para possibilitar a comparação da potência dos antifúngicos, os valores de CIM em $\mu\text{g/mL}$ foram convertidos em μMolar ($\mu\text{mol/L}$) (**TABELA 2.8**).

TABELA 2.8 – Concentração inibitória mínima (CIM) para 33 antifúngicos derivados azólicos frente às cepas *C. gattii* e *C. neoformans*

Antifúngico	<i>C. gattii</i>		<i>C. neoformans</i>	
	µg/mL	µmol/L	µg/mL	µmol/L
Azaconazol	32	106,62	4	13,33
Bifonazol	>64	>206,19	64	206,19
Bromuconazol	8	21,22	2	5,30
Cetoconazol	0,25	0,47	0,0625	0,12
Climbazol	2	6,83	0,25	0,85
Clotrimazol	8	23,20	1	2,90
Ciproconazol	1	3,43	0,125	0,43
Diclobutrazol	8	24,37	1	3,05
Difenoconazol	1	2,46	<0,0156	<0,04
Diniconazol	2	6,13	0,25	0,77
Econazol	4	10,48	0,25	0,65
Epoconazol	1	3,03	0,125	0,38
Etaconazol	8	24,38	0,5	1,52
Fembuconazol	8	23,75	1	2,97
Fluconazol	32	104,48	4	13,06
Fluotrimazol	64	168,70	64	168,70
Fluquiconazol	16	42,53	1	2,66
Flusilazol	8	25,37	1	3,17
Flutriafol	>64	>212,42	8	26,55
Hexaconazol	2	6,37	0,25	0,80
Imazalil	16	53,84	4	13,46
Itraconazol	0,125	0,18	0,0625	0,09
Metconazol	1	3,13	0,0625	0,20
Miconazol	1	2,40	0,125	0,30
Myclobutanil	16	55,41	4	13,85
Penconazol	4	14,08	1	3,52
Prochloraz	32	84,96	16	42,48
Propiconazol	4	11,69	0,5	1,46
Prothioconazol	32	92,95	2	5,81
Sulconazol	0,25	0,63	<0,0156	<0,04
Tebuconazol	4	12,99	0,5	1,62
Triadimenol	4	10,11	1	2,53
Triticonazol	4	12,59	2	6,29

Os resultados obtidos mostram que a CIM dos antifúngicos frente à *C. neoformans* apresentou diferença superior a 5000 vezes, variando entre 0,04 e 206,19 $\mu\text{mol/L}$. Os compostos que apresentam menores valores de CIM (abaixo de 0,15 $\mu\text{mol/L}$) foram os antifúngicos Cetoconazol, Difenoconazol, Itraconazol e Sulconazol. Em contrapartida, os antifúngicos Bifonazol e Fluotrimazol apresentam valores de CIM acima de 100 $\mu\text{mol/L}$.

Em relação aos dados de CIM dos antifúngicos frente a *C. gattii*, observou-se uma variação entre 0,18 e 212,42 $\mu\text{mol/L}$, representando uma diferença acima de 1000 vezes. Cetoconazol, Itraconazol e Sulconazol apresentaram maior potência (CIM menor que 0,55 $\mu\text{mol/L}$), ao passo que os antifúngicos Bifonazol e Flutriafol apresentaram menor atividade, com valores de CIM maiores de 200 $\mu\text{mol/L}$.

Considerando que o foco principal desse trabalho é a investigação das propriedades físico-químicas e estruturais que determinam a atividade de antifúngicos azólicos frente a cepas resistentes de *Cryptococcus*, maior atenção foi dada aos resultados biológicos obtidos para *C. gattii*. Os dados de CIM obtidos para a série de 33 derivados azólicos ativos frente *C. gattii*, foram então utilizados para construção de modelos quantitativos de relação estrutura-atividade, utilizando para isso descritores topológicos (QSAR clássico) e descritores de fragmentos (holograma QSAR), como descrito abaixo.

2.6.4 Modelos quantitativos de relação estrutura-atividade

Para que os dados de atividade biológica fossem utilizados no desenvolvimento de modelos de QSAR, os valores de CIM (em $\mu\text{mol/L}$) foram convertidos nos valores correspondentes na escala logarítmica de pCIM ($-\log [\text{CIM}]$) e utilizados como variável dependente nos modelos de estatísticos desenvolvidos (**TABELA 2.9**).

Os valores de pCIM apresentam distribuição normal entre 3,67 e 6,75 (**FIGURA 2.15**), assegurando que os modelos obtidos serão aplicáveis para a predição da atividade biológica de moléculas cuja potência varia de mili até micromolar. A fim de avaliar a real capacidade preditiva dos modelos, os 33 antifúngicos foram divididos em dois subconjuntos: grupo treino (24 compostos) e grupo teste (9 compostos). O grupo treino é formado pelos compostos utilizados na construção dos modelos de QSAR, enquanto os antifúngicos que compõem o grupo teste, não utilizados na etapa de desenvolvimento dos modelos, participam

do processo de validação externa. A divisão foi feita de forma aleatória, mas com atenção especial para que todo espectro de potência estivesse amostrado no grupo teste (FIGURA 2.15).

TABELA 2.9 – Valores de pCIM dos 33 antifúngicos frente à *C. gattii*

Grupo Treino				Grupo Teste	
Antifúngico	pCIM	Antifúngico	pCIM	Antifúngico	pCIM
Azaconazol	3,97	Hexaconazol	5,20	Cetoconazol	6,33
Bifonazol	3,69	Imazalil	4,27	Climbazol	5,17
Bromuconazol	4,67	Itraconazol	6,75	Ciproconazol	5,47
Clotrimazol	4,64	Metconazol	5,50	Diclobutrazol	4,61
Difenoconazol	5,61	Miconazol	5,62	Econazol	4,98
Diniconazol	5,21	Penconazol	4,85	Fluconazol	3,98
Epoxiconazol	5,52	Prochloraz	4,07	Fluotrimazol	3,77
Etaconazol	4,61	Propriconazol	4,93	Myclobutanil	4,26
Fembuconazol	4,62	Sulconazol	6,20	Prothioconazol	4,03
Fluquiconazol	4,37	Tebuconazol	4,89		
Flusilazol	4,60	Triadimenol	5,00		
Flutriafol	3,67	Triticonazol	4,90		

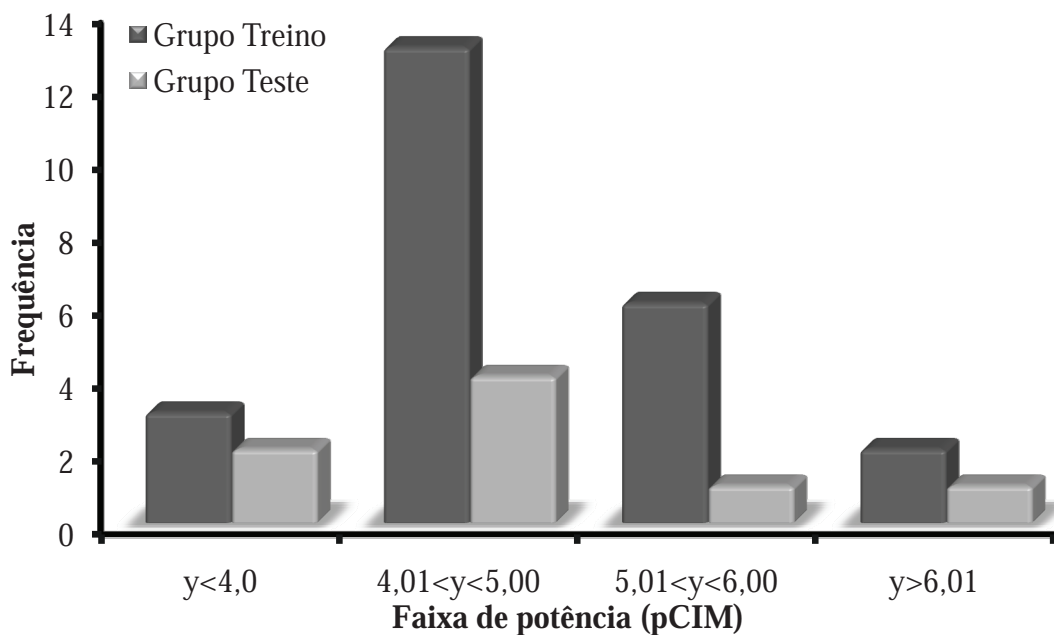


FIGURA 2.15 – Histograma de frequência dos antifúngicos do grupo treino e teste em relação à faixa de pCIM

2.6.4.1 Modelos de HQSAR

Inicialmente, optou-se por utilizar descritores derivados de fragmentos moleculares para a construção de modelos de QSAR, para tanto se empregou a técnica de holograma QSAR. Considerando que o padrão de fragmentos gerados para cada molécula tem grande impacto nos resultados estatísticos e este é influenciado por parâmetros como: tamanho do fragmento (FS), distinção de fragmento (FD) e comprimento do holograma molecular (HL), diversas combinações desses parâmetros foram avaliados durante o desenvolvimento dos modelos de HQSAR.

Em função do grande número de descritores obtidos e pela impossibilidade de se adotar estratégias para pré-seleção de descritores, os modelos de HQSAR foram desenvolvidos através de regressão por mínimos quadrados parciais (PLS). Todavia, visando evitar a obtenção de modelos com super-ajuste, o número máximo de PCs foi restringido a 4, quando se utilizou o método de validação cruzada LOO na determinação do número ótimo de PCs.

Inicialmente foi avaliada a influência da distinção de fragmento (FD) nos parâmetros estatísticos (**TABELA 2.10**). Com base nesses resultados, é possível observar que a distinção padrão (modelo 1- ABC) apresenta ajuste moderado e baixa consistência interna ($r^2=0,55$ e $q^2=0,11$, com 2 PCs). Na tentativa de melhorar este modelo, foi adicionado o parâmetro hidrogênio (H), seguido de quiralidade (Ch), e doador e acceptor (DA), que promoveram elevação discreta dos valores de r^2 e q^2 (comparar modelos **2**, **3** e **4** vs **1**). Diversas combinações foram avaliadas (modelos **5-11**), porém nenhuma delas produziu modelos com consistência interna significativa.

TABELA 2.10 – Influência da distinção de fragmento sobre os parâmetros estatísticos do HQSAR, usando o tamanho de fragmento padrão (4 – 7)

Modelo	FD	q^2	r^2	SEE	HL	N
1	ABC	0,11	0,55	0,52	61	2
2	ABCH	0,25	0,46	0,57	61	2
3	ABCHCh	0,25	0,88	0,28	307	4
4	ABCHChDA	0,28	0,81	0,36	59	4
5	ABH	0,20	0,41	0,59	71	2
6	ACH	0,23	0,45	0,58	59	2
7	ABCCh	0,21	0,60	0,49	61	2
8	ACHCh	0,23	0,48	0,56	61	2
9	AHCh	0,22	0,44	0,58	71	2
10	ABHCh	0,29	0,61	0,49	59	3
11	AHChDA	0,19	0,30	0,64	71	1

q^2 , coeficiente de correlação com validação cruzada (*LOO*); r^2 , coeficiente de correlação; SEE, desvio padrão da predição; *N*, número ótimo de componentes principais. Distinção de fragmento (FD): A, átomos; B, ligação; C, conectividade; H, átomos de hidrogênio; Ch, quiralidade; DA, doador e acceptor.

Na tentativa de obter modelos com melhores parâmetros estatísticos, a influência do tamanho dos fragmentos também foi investigada (**TABELA 2.11**) para o melhor modelo de HQSAR gerado (**modelo 10**). Embora o modelo gerado com fragmentos de 3 a 6 átomos (**modelo 13**) apresente bom ajuste, sua consistência interna é baixa ($r^2=0,85$, $q^2=0,35$, com 4 PCs).

TABELA 2.11 – Influência do tamanho do fragmento (FS) sobre os parâmetros estatísticos dos modelos de HQSAR

Modelo	FD	q^2	r^2	SEE	HL	N	FS
12	ABHCh	0,14	0,47	0,57	59	2	2-5
13	ABHCh	0,35	0,85	0,32	59	4	3-6
14	ABHCh	0,15	0,59	0,51	59	3	5-8

q^2 , coeficiente de correlação com validação cruzada (*LOO*); r^2 , coeficiente de correlação; SEE, desvio padrão da predição; *N*, número ótimo de componentes principais. Distinção de fragmento (FD): A, átomos; B, ligação; C, conectividade; H, átomos de hidrogênio; Ch, quiralidade; DA, doador e acceptor.

Portanto, esse modelo não apresenta capacidade preditiva para moléculas que não façam parte do conjunto treino. Isso pode ser comprovado pelo baixo poder preditivo desse modelo ($r^2_{pred} = 0,38$), quando utilizado para estimar a potência dos compostos do conjunto teste (FIGURA 2.16).

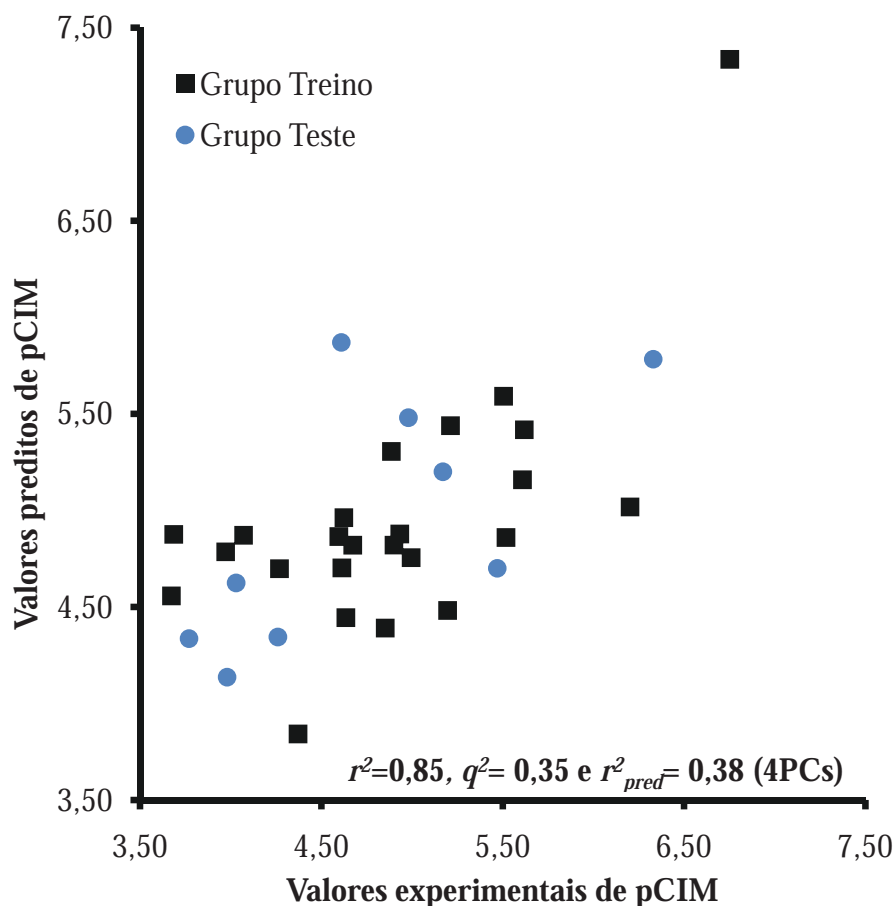


FIGURA 2.16 – Valores experimentais versus valores preditos de pCIM para os 33 antifúngicos

Esses resultados sugerem que modelos baseados exclusivamente em fragmentos moleculares não são capazes de produzir correlações úteis ou bem sucedidas com a atividade biológica dos derivados azólicos estudados. Diante disso, com o objetivo de desenvolver modelos com maior capacidade preditiva, descritores topológicos que descrevem informações estéricas e eletrostáticas foram utilizados para o desenvolvimento de modelos quantitativos de QSAR 2D, como descrito a seguir.

2.6.4.2 Modelos de QSAR baseado em descritores topológicos

Para o desenvolvimento de modelos de QSAR baseado em descritores topológicos um conjunto de 412 descritores físico-químicos e estruturais foram calculados com auxílio do programa DRAGON 5.5, como descrito na **seção 2.5.2**.

Em seguida, estes descritores foram utilizados como variáveis dependentes para a construção de modelos preliminares de regressão linear múltipla (RLM) com auxílio do programa MOBYDIGS 1.0.¹²⁸ Esta etapa teve por objetivo selecionar os descritores mais correlacionados com a atividade biológica (pCIM). Como descrito na **seção 2.5.3**, modelos de RLM com até 4 descritores foram gerados com auxílio do método de algoritmo genético.

O melhor modelo de RLM apresentou bons resultados estatísticos ($r^2=0,78$ e $q^2=0,74$), porém uma baixa capacidade preditiva ($r^2_{pred}=0,1$). Uma das explicações para isso pode estar relacionada à alta colinearidade dos descritores ou representação inadequada do espaço químico. Nesse segundo caso, seria necessário utilizar outro tipo de descritor molecular (ex. descritores 3D), pois a atividade biológica dependeria de características que não podem ser descritas de forma bidimensional (ex. conformação). Contudo, como pode ser comprovado abaixo, esse não é o problema encontrado neste trabalho.

A fim de eliminar problemas de colinearidade entre os descritores 2D, optou-se por desenvolver modelos baseados em regressão por mínimos quadrados parciais (PLS). Essa metodologia promove redução da dimensionalidade do modelo, uma vez que reescreve as coordenadas dos descritores originais em um novo eixo de coordenadas, os quais são ortogonais entre si.

Como o processo de geração e seleção de modelos ocorre de forma independente em cada uma das 10 populações, é comum encontrar modelos com combinações semelhantes de descritores em populações diferentes. Por este motivo, apenas 36 descritores encontrados nos 31 modelos com valor de $q^2 \geq 0,64$ (**APÊNDICE B**), foram reunidos, autoescalados e utilizados para o desenvolvimento de modelos de QSAR 2D por PLS.

O modelo inicial de PLS (considerando as PCs obtidas a partir dos 36 descritores originais) mostrou parâmetros estatísticos comparáveis aos modelos de RLM ($r^2=0,88$ e $q^2=0,69$) utilizando 3 componentes principais. Adicionalmente, o PLS disponibiliza ferramentas diagnósticas, como o vetor de regressão, que podem ser aplicadas para a exclusão de descritores que pouco contribuíram para o modelo, possibilitando a otimização do modelo e, conseqüente, melhora dos parâmetros estatísticos.

Esta estratégia guiou a exclusão iterativa de 12 descritores e resultou em um modelo final de PLS que apresenta excelente ajuste e ótima consistência interna ($r^2= 0,95$, $q^2= 0,86$ com 3 PCs) (**TABELA 2.12**), além de bom poder preditivo ($r^2_{pred}= 0,72$) (**FIGURA 2.17** e **TABELA 2.13**).

TABELA 2.12 – Descritores presentes no modelo final de PLS

Descritor	Significado
Me	Média atômica da eletronegatividade de Sanderson (Escalonada em átomos de carbono)
Ms	Média do estado eletrotológico
nDB	Número de duplas ligações
MAXDN	Varição eletrotológica negativa máxima
PW3	Índice de formato de Randic para distância topologica 3
T(F..F)	Soma das distâncias topológicas entre átomos de flúor
X2v	Índice (Chi-2) de conectividade de valências
X5v	Índice (Chi-5) de conectividade de valências
X2Av	Índice (Chi-2) de conectividade de valência média
IC3	Índice contendo informação sobre a simetria do vizinho de ordem 3
ATS6p	Autocorrelação de Broto-Moreau para distância topológica de índice 6 ponderado pela polarizabilidade atômica
MATS7m	Autocorrelação de Moran na distância topológica 7 ponderado pela massa atômica
MATS8m	Autocorrelação de Moran na distância topológica 8 ponderado pela massa atômica
MATS2e	Autocorrelação de Moran na distância topológica 2 ponderado pela eletronegatividade de Sanderson
GATS1m	Autocorrelação de Geary na distância topológica 1 ponderado pela massa atômica
GATS1e	Autocorrelação de Geary na distância topológica 1 ponderado pela eletronegatividade de Sanderson
GATS4p	Autocorrelação de Geary na distância topológica 4 ponderado pela polarizabilidade
ESpm03u	Momento espectral (03) do eixo da matriz adjacente
BELe8	Menor autovalor da matriz de Burden na distância tipológica 8 ponderado pela eletronegatividade de Sanderson
GGI ₁	Índice de carga topológica de ordem 1
SEigZ	Soma de autovalores do eixo Z ponderado pela matriz de distância de Barysz
nArX	Número de halogênios ligados à anel aromático
C-003	Fragmento centrado no átomo – CHR ₃ ^a
C-040	Fragmento centrado no átomo – R-C(=X)-X / R-C#X / X=C=X ^{a,b}

^a R representa qualquer agrupamento ligado através de carbonos. ^b “=” representa dupla ligação; “#” representa tripla ligação; X são heteroátomos (ie. O, N, S...)

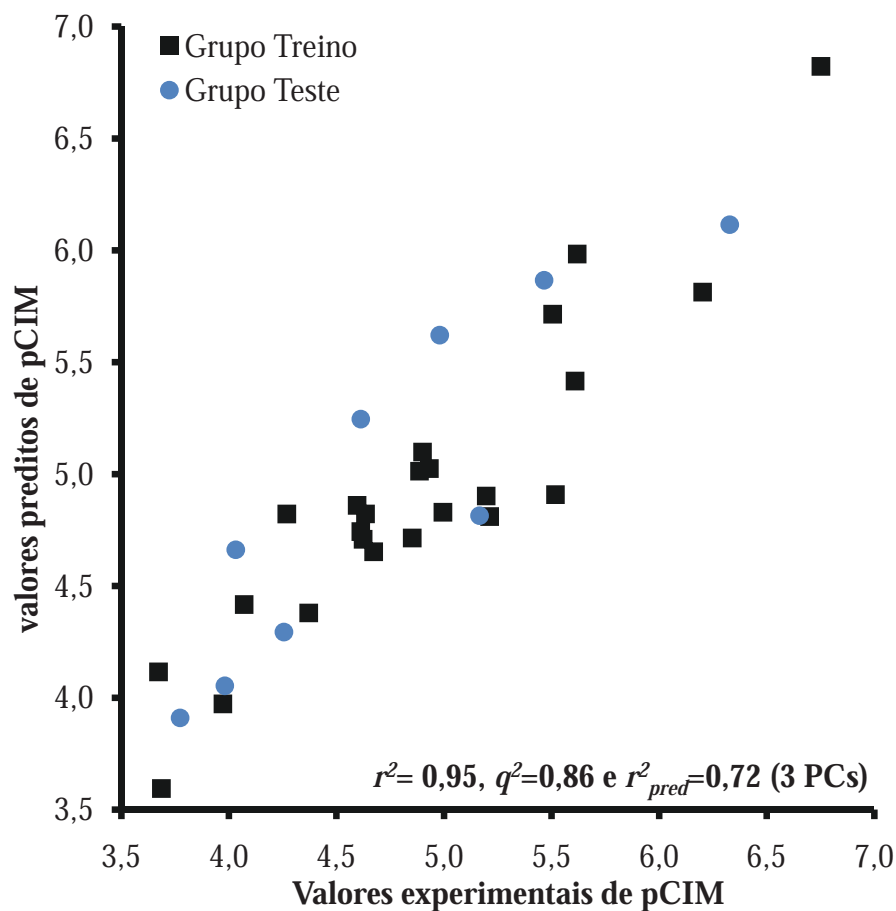


FIGURA 2.17 – Valores experimentais e preditos de pCIM para 33 antifúngicos de acordo com o melhor modelo de QSAR 2D baseado em descritores topológicos

TABELA 2.13 – Valores experimentais e preditos de pCIM dos antifúngicos do grupo teste para o modelo de QSAR 2D baseado em descritores topológicos

Antifúngico	pCIM		Resíduo
	Experimental	Predito	
Cetoconazol	6,33	6,11	-0,21
Climbazol	5,17	4,81	-0,35
Ciproconazol	5,47	5,87	0,40
Diclobutrazol	4,61	5,25	0,63
Econazol	4,98	5,62	0,64
Fluconazol	3,98	4,05	0,07
Fluotrimazol	3,77	3,91	0,14
Myclobutanil	4,26	4,29	0,04
Prothioconazol	4,03	4,66	0,63

Os resultados estatísticos alcançados sugerem que os descritores contidos no modelo final explicam as relações entre a estrutura química e a atividade biológica dos derivados azólicos estudados, podendo, portanto auxiliar na identificação de características físico-químicas e estruturais relevantes para a atividade de novos compostos potencialmente úteis como antifúngicos. Esse tipo de informação pode ser obtido através da análise do vetor de regressão do modelo, uma vez que descritores com coeficientes positivos possuem contribuição favorável à atividade, enquanto descritores com coeficientes negativos reduzem a atividade (**FIGURA 2.18**).

A análise preliminar do vetor de regressão mostra que 58,3% dos descritores estão relacionados a características eletrônicas (GATS1e, MATS2e, PW3, MS, Me, IC3, X5v, SEigZ, GATS4p, X2Av, ATS6p, BEL8e, X2v, GGI1) enquanto que características topológicas e estéricas compõem a porcentagem restante. Este dado sugere que parâmetros eletrônicos, como interações eletrostáticas, formação de dipolo e polarizabilidade de grupos funcionais são cruciais para a atividade antifúngica. De fato, sabe-se que interações eletrostáticas entre o nitrogênio 3 do anel imidazol (4 no triazol) com o ferro do grupo heme no sítio de ligação,^{49; 151} é fundamental para afinidade de inibidores de lanosterol 14- α desmetilase pelo seu alvo terapêutico. Adicionalmente, os descritores com maior contribuição positiva (GGI1) e negativa (GATS1e e MATS2e) para o vetor de regressão descrevem propriedades eletrônicas das moléculas (**FIGURA 2.18**). Assim, para exemplificar que tipo de informação química pode ser extraída de descritores topológicos, esses três descritores serão analisados em maior detalhe a seguir.

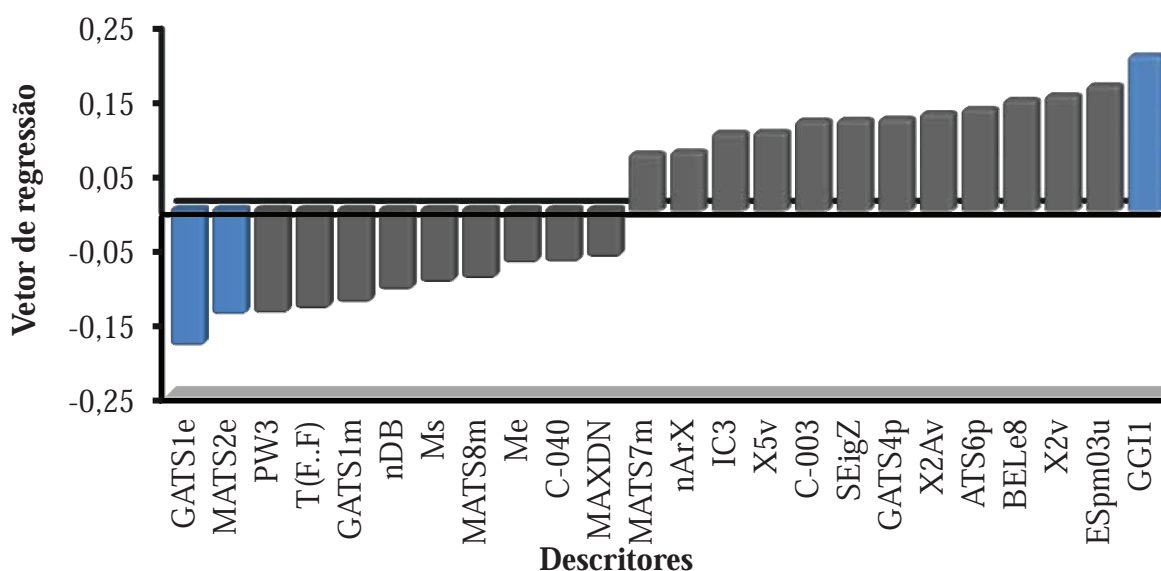


FIGURA 2.18 – Coeficiente do vetor de regressão para as variáveis do modelo final de QSAR 2D baseado em descritores topológicos

O descritor GGI1 é inversamente relacionado com a distribuição de carga global de uma molécula, a partir da transferência de carga entre átomos distante entre si por uma ligação química.^{148; 149} Dessa forma, quanto maior o valor de GGI1, menos distribuída encontra-se a densidade eletrônica entre os átomos, o que é detrimental para potência, uma vez que a interação com o ferro do grupamento heme de lanosterol 14 α -desmetilase depende da maior densidade eletrônica nos nitrogênios dos anéis imidazólicos e triazólicos.

Como pode ser observado na **FIGURA 2.19**, derivados azólicos potentes apresentam valores elevados de GGI1, como por exemplo, Cetoconazol (GGI1 = 8,5; CIM = 0,47 μ M) e Itraconazol (GGI1 = 10,5; CIM = 0,18 μ M). Por outro lado, compostos de baixa potência apresentam valores reduzidos de GGI1, por exemplo, Imazalil (GGI = 3,5; CIM = 53,84 μ M) e Bifonazol (GGI1 = 2,5; CIM = 206,19 μ M). Além disso, 75% dos compostos com valores de pCIM maiores que 4,85 possuem valores de GGI1 maiores que 5,0, ao passo que 73% dos compostos com valores de pCIM abaixo de 4,85 apresentam valores de GGI1 menores que 5,0.

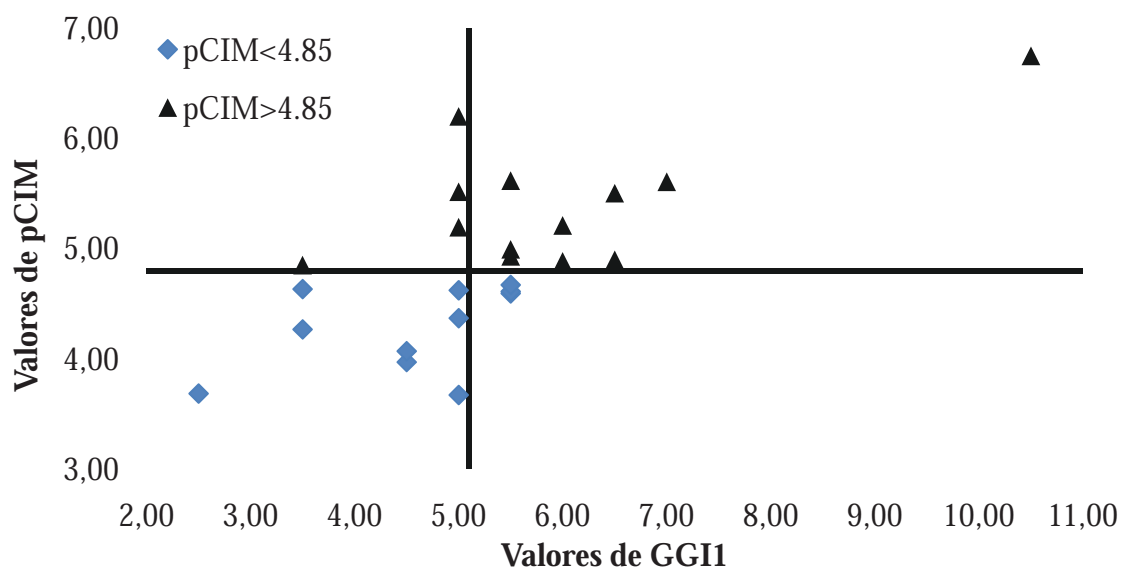


FIGURA 2.19 – Coeficientes de GGI1 versus valores de CIM para 33 antifúngicos

Abordagem similar à que foi utilizada para GGI1 pode ser aplicada para explicar como os descritores influenciam negativamente a atividade biológica desses compostos (GATS1e e MATS2e). Embora estes descritores destaquem a mesma característica em distâncias topológicas similares, eles não são matematicamente correlacionados ($r^2 = 0,05$). Descritores MATS são sensíveis à alterações globais na eletronegatividade, ao passo que, os descritores

GATS avaliam de forma mais eficiente alterações locais, por exemplo, decorrentes de ramificações.^{148; 149}

Quanto maiores são os valores de GATS1e e MATS2e maior será a diferença de eletronegatividade entre os átomos na distância topológica de 1 e 2 átomos, respectivamente. Como valores elevados destes descritores estão relacionados com valores baixos de pCIM, é possível concluir que pequenas diferenças na eletronegatividade entre os átomos dos derivados azólicos pode promover o aumento dos valores de pCIM.

CAPÍTULO 3

3 ESTUDOS DA RESISTÊNCIA DE *CRYPTOCOCCUS* À DERIVADOS AZÓLICOS

O combate às infecções fúngicas oportunistas tem se tornando um desafio cada vez mais difícil, em função do desenvolvimento acelerado de variedades resistentes de fungos aos antifúngicos disponíveis atualmente.^{59; 60; 61; 154}

Do ponto de vista clínico, a resistência pode ser dividida em primária (intrínseca), quando ocorre em todos os seres de um gênero ou espécie antes da exposição ao fármaco, e secundária (extrínseca), quando é desenvolvida após exposição ao fármaco.¹⁵⁵ Em *Cryptococcus*, a resistência intrínseca é muito pouco documentada, pois, sabe-se que esse microorganismo é imune a antifúngicos da classe das equinocandinas⁵⁹ e apresenta baixa susceptibilidade à alguns derivados azólicos.^{156; 157} Por outro lado, a resistência extrínseca à derivados azólicos, como o fluconazol, vem recebendo grande atenção.¹⁵⁵

Embora o aparecimento de resistência à derivados azólicos seja operacionalizado por múltiplos fatores, a redução do conteúdo intracelular, mediada por bombas de efluxo, é considerada um dos principais mecanismos de resistência.^{158; 159} Este mecanismo de resistência foi primeiramente documentado em *Candida albicans*,^{160; 161; 162; 163} e pode ser averiguado através da inclusão de moduladores de bomba, como verapamil, hormônios sexuais e ibuprofeno nos testes de determinação de CIM^{164; 165; 166}

Outro mecanismo de resistência, bem documentado em *Candida*, ocorre através de mutações na enzima lanosterol 14 α -desmetilase do fungo, os quais resultam em redução da afinidade de derivados azólicos pelo alvo macromolecular.^{154; 167; 168} Embora existam diversas mutações silenciosas distribuídas pelo gene ERG11 de *C. albicans*, mais de 30 mutações estão associadas à redução da afinidade da enzima lanosterol 14 α -desmetilase pelos derivados azólicos.^{158; 169}

Estudos acerca da resistência em *Cryptococcus*, por meio de bombas de efluxo ou de mutações na enzima lanosterol 14 α -desmetilase são escassos.^{59; 155} Venkateswarlu e colaboradores¹⁶¹ foram pioneiros nesse aspecto ao descrever a redução da afinidade de derivados azólicos pela enzima lanosterol 14 α -desmetilase e diminuição do conteúdo intracelular de derivados azólicos como possíveis causas para resistência de isolados clínicos

de *Cryptococcus*. Estudos mais recentes permitiram a identificação de uma mutação pontual no gene ERG11, que promove a substituição da glicina 484 por serina (G484S) na enzima lanosterol 14 α -desmetilase.⁶⁰ É interessante notar que estudos em *Candida albicans* sugerem que a mutação equivalente (G464S) promove alterações no posicionamento do grupamento heme, reduzindo assim a afinidade da enzima pelos derivados azólicos.^{168; 170}

Este tipo de estudo evidencia que informações estruturais podem ser úteis na elucidação do mecanismo de resistência em *Cryptococcus*. Entretanto, a caracterização estrutural de enzimas da via do ergosterol de fungos, em especial a lanosterol 14 α -desmetilase, tem apresentado resultados limitados. Até o presente momento, informações tridimensionais estão disponíveis para enzima de *Mycobacterium tuberculosis*, *Trypanosoma cruzi* e *Homo sapiens*. Contudo, deve-se avaliar com cautela as inferências feitas a partir dessas estruturas, uma vez que a identidade sequencial dessas proteínas com a enzima de *Cryptococcus* é da ordem de 25-35%.^{67; 68; 171; 172; 173; 174}

Enquanto não se dispõe de dados experimentais de alta resolução sobre a enzima de interesse, uma alternativa viável para agregar informações estruturais aos estudos sobre resistência é a utilização da modelagem comparativa, ou modelagem por homologia.

A técnica de modelagem por homologia consiste na predição da estrutura terciária de uma proteína desconhecida, a partir de sua sequência primária e de uma ou mais proteínas similares e relacionadas (homologas) cujas estruturas tenham sido determinadas experimentalmente.^{175; 176} Para tal, essa estratégia se baseia na conservação do enovelamento tridimensional entre membros de uma mesma família de proteínas.^{177; 178; 179} A conservação da estrutura terciária advém de uma consequência evolutiva, regida por restrições funcionais,¹⁸⁰ portanto, é esperado que a similaridade estrutural seja mais preservada que a similaridade sequencial.¹⁷⁶

O desenvolvimento de modelos por homologia é um processo interativo conduzido em 4 etapas principais (**FIGURA 3.1**): a) identificação e seleção de proteínas moldes; b) alinhamento das seqüências de aminoácidos das proteínas-moldes e da proteína cuja estrutura 3D desconhecida (alvo); c) construção das coordenadas do modelo; d) validação do modelo.

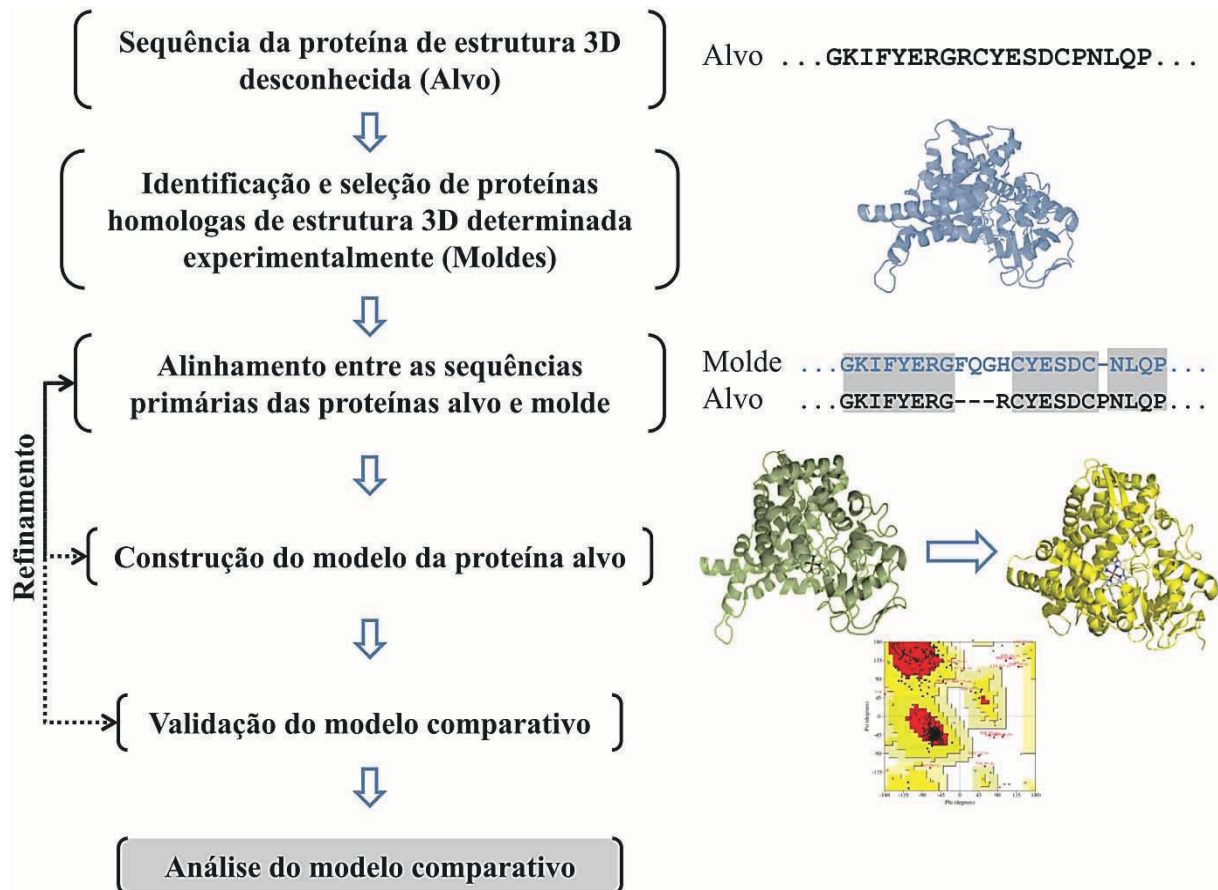


FIGURA 3.1 – Processo de construção de um modelo comparativo. Adaptado de Cavasotto e Phatak (2009) ¹⁸¹

Em primeiro lugar, é necessária a identificação de proteínas com estruturas tridimensionais conhecidas e correlacionadas com a seqüência-alvo e que podem servir como molde para a construção da estrutura tridimensional da proteína desconhecida (proteína-alvo). ¹⁸² Nesta etapa levam-se em consideração diversos aspectos como: informações acerca da estrutura secundária e terciária da proteína (ex. regiões altamente conservadas), funções semelhantes, identidade e similaridade seqüencial ou entre as seqüências. ¹⁸³ Essa etapa pode ser realizada com auxílio de ferramentas disponíveis como BLAST (do inglês *Basic Local Alignment Search Tool*) que realiza o alinhamento e comparação da seqüência de interesse com uma ou várias seqüências nucleotídicas e peptídicas. ¹⁸⁴

A partir da identificação das possíveis proteínas-moldes, é preciso selecionar as mais apropriadas para a construção do modelo por homologia, utilizando para isso os seguintes parâmetros: grau de identidade entre as seqüências primárias das proteínas alvo e molde; qualidade da estrutura cristalográfica (estrutura molde), a qual pode ser estimada pelos valores de R (do inglês *R-values*); resolução para estruturas cristalográficas ou número de restrições por resíduo em modelos de RMN. ¹⁸⁵

Posteriormente à seleção das proteínas moldes, é necessário fazer o ajuste do alinhamento das sequências primárias das proteínas moldes com a sequência alvo. Haja vista que os programas de modelagem dependem de equivalências estruturais entre os resíduos da proteína alvo e molde. Esta etapa é considerada uma das mais críticas no processo de modelagem.¹⁸⁶ O alinhamento seqüencial múltiplo permite identificar características estruturais conservadas, como elementos de estrutura secundária e resíduos catalíticos, que por sua vez, podem ser úteis na construção e interpretação do modelo por homologia.¹⁸⁷

Embora diversas ferramentas utilizem o alinhamento local para extrair equivalências entre as sequências alvo e molde, quando a identidade sequencial é menor que 40%, há a necessidade de intervenção manual para minimizar os pareamentos incorretos entre as sequências, o que leva em conta o conhecimento da estrutura da proteína em estudo. Por exemplo, deve-se evitar o aparecimento de lacunas (ou *gaps*) em regiões de estrutura secundária, como fitas beta-pregueadas ou hélices alfa, pois é esperada uma maior conservação da sequência nestas regiões do que em regiões de alça (do inglês *loops*).^{188; 189;}
¹⁹⁰ A utilização de alinhamento múltiplo, caso existam 2 ou mais sequências moldes disponíveis para o alinhamento, facilita essa tarefa ao mostrar regiões conservadas em uma mesma família de proteínas.¹⁹¹

Após o alinhamento entre as sequências alvo e moldes, inicia-se o processo de construção do modelo comparativo da proteína alvo, que podem ser realizado por método de corpos rígidos ou modelagem por restrições espaciais.¹⁹²

O método de corpos rígidos baseia-se no conceito que a proteína é formada por subestruturas conservadas (ex. fitas beta e hélices alfa) conectadas por regiões pouco conservadas (ex. alças e cadeias laterais). Dessa forma, após a sobreposição estrutural das proteínas moldes, as coordenadas espaciais médias dos átomos da cadeia principal são transferidas para os átomos dos resíduos correspondentes na sequência-alvo. Fragmentos que não possam ser modelados por essa metodologia (alças, inserções, etc.) tem sua conformação otimizada através de métodos de mecânica molecular. Por fim, a conformação das cadeias laterais e interações intermoleculares são otimizadas, também por mecânica molecular.^{193; 194;}
195

Na metodologia baseada na satisfação de restrições espaciais, disponível no programa MODELLER 9.8, restrições de distância e ângulos diedros são extraídas das estruturas moldes e empregadas para determinar a orientação espacial dos resíduos da sequência alvo (**FIGURA 3.2**).^{196; 197}

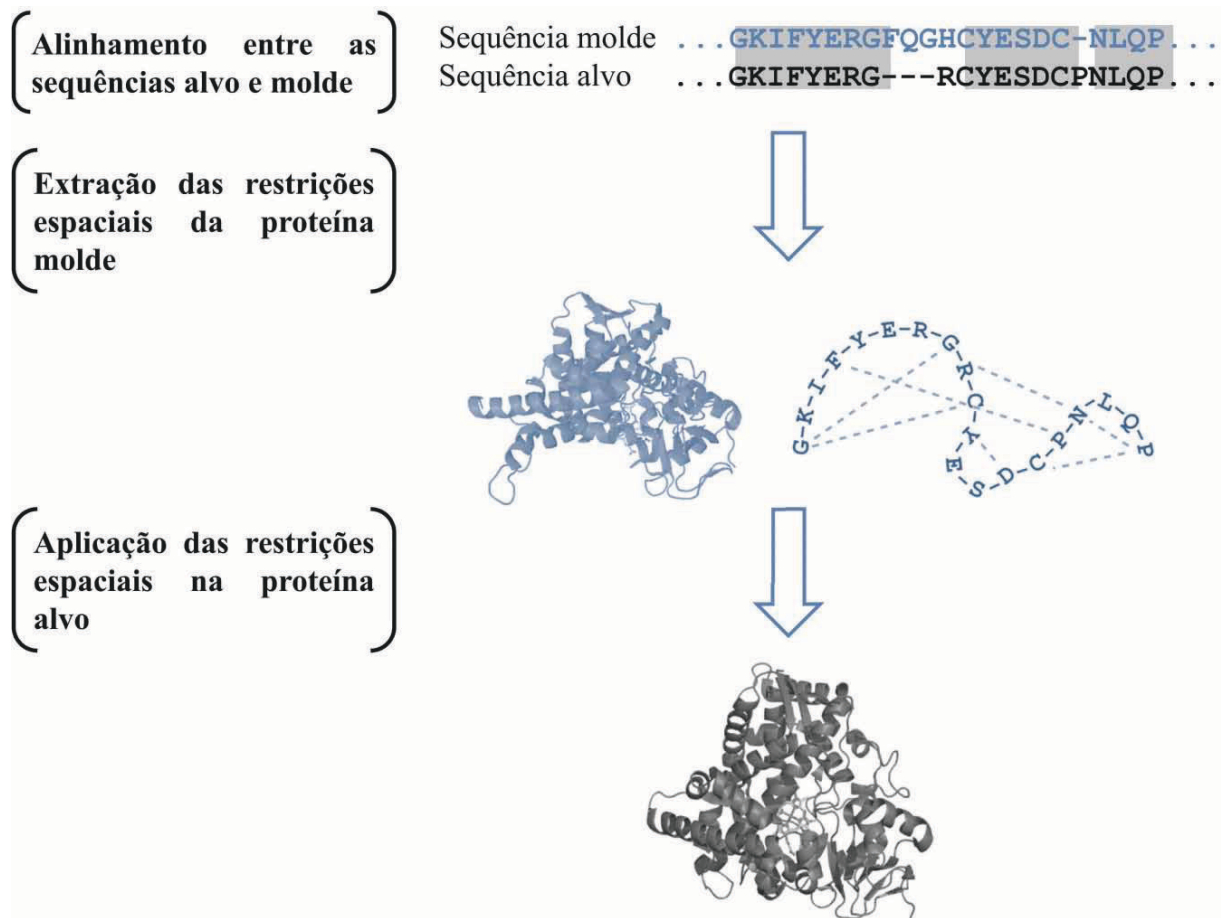


FIGURA 3.2 – Representação da construção de modelos comparativos por restrições espaciais. Após o alinhamento das sequências alvo e molde ocorre a extração das restrições espaciais da proteína molde que serão aplicadas sobre a sequência da proteína alvo para gerar o modelo tridimensional. Adaptado de Sanchez e Sali (2000) ¹⁹⁸

Em seguida, restrições espaciais e termos de campos de força CHARMM22 são combinados em uma função objetiva para a otimização da estereoquímica dos resíduos. ^{199;}
200

Cabe destacar que os modelos obtidos através dessas técnicas devem ser avaliados quanto a sua qualidade estereoquímica, empacotamento dos resíduos (enovelamento protéico) e ambiente químico onde os resíduos se encontram, a fim de garantir que a estrutura 3D produzida possa ter utilidade na interpretação de dados biológicos. ¹⁷⁶

Propriedades estereoquímicas como comprimentos de ligação, planaridade de anéis e de ligações peptídicas, quiralidade, impedimentos estericos etc. podem ser avaliados com auxílio do programa PROCHECK (http://www.jcsg.org/prod/scripts/validation/sv_final.cgi), ^{201;} ²⁰² o qual permite também analisar os ângulos torcionais phi (ϕ) e psi (ψ) dos resíduos da cadeia principal da proteína. A análise gráfica desses valores permite identificar resíduos que

apresentam choques estéricos visualizados em regiões não permitidas (energeticamente desfavoráveis) do gráfico de Ramachandran. (FIGURA 3.3)

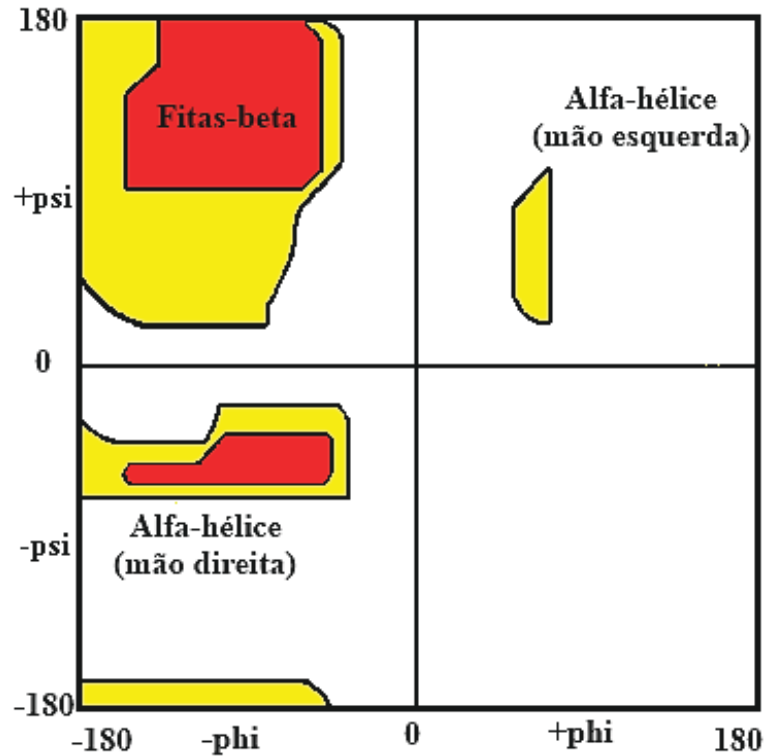


FIGURA 3.3 – Representação gráfica dos ângulos phi e psi através do gráfico de Ramachandran mostrando as regiões favoráveis (vermelho) e permitidas (amarelo). Adaptado de Laskowski e colaboradores (1993) ²⁰¹

O programa MODELLER 9.8 possui ferramentas para avaliação interna da qualidade do enovelamento do modelo, como por exemplo, a função de energia global denominada DOPE (do inglês *Discrete Optimized Protein Energy*), no qual valores pequenos de DOPE, indicam melhor qualidade do modelo gerado. ²⁰³

Também é possível avaliar a compatibilidade entre o modelo estrutural (3D) e a sequência primária de aminoácidos (1D) da proteína com o objetivo de obter informações acerca de erros no enovelamento geral ou local da proteína alvo através das ferramentas disponíveis no servidor VERIFY 3D (http://nihserver.mbi.ucla.edu/Verify_3D/).

Para tanto, o VERIFY 3D deriva 6 categorias de ambientes químicos a partir do valor da área enterrada na proteína e da fração da cadeia lateral coberta por átomos polares (FIGURA 3.4A). Como alguns resíduos adotam preferencialmente certas estruturas secundárias (hélices alfa, fitas beta), os ambientes são multiplicados por 3, levando aos 18 ambientes utilizados para pontuar os 20 tipos de aminoácidos da proteína. A seguir, cada aminoácido da proteína é analisado e recebe uma pontuação que descreve a probabilidade de

ser encontrado em um dos 18 ambientes químicos predefinidos. A somatória dessa pontuação fornece uma idéia geral do enovelamento da proteína, mas não permite avaliar erros localizados em regiões específicas da estrutura 3D.^{204; 205}

Visando contornar essa limitação, pode-se fazer uma análise gráfica da pontuação de cada resíduo versus sua posição na seqüência, considerando-se um ambiente químico composto por 21 aminoácidos (10 resíduos acima e abaixo do resíduo referência) (**FIGURA 3.4B**). Seqüências de valores abaixo de zero geralmente são encontradas em regiões cujos resíduos não estão em seus ambientes adequados, ou seja, apresentam enovelamento incorreto. Por outro lado, valores individuais acima de 0,2 indicam que o resíduo encontra-se num ambiente químico comumente observado em estruturas determinadas experimentalmente.

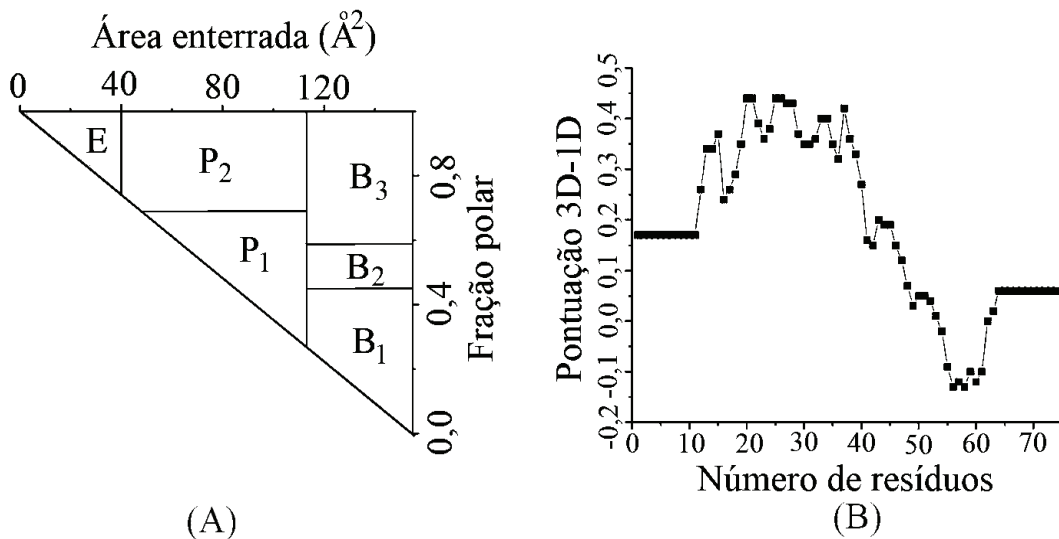


FIGURA 3.4 – O VERIFY 3D deriva 6 ambientes químicos (E, P₁, P₂, B₁, B₂, B₃) a partir do valores da área enterrada na proteína e da fração da cadeia lateral coberta por átomos polares (A). Um grupo de resíduos é, então, avaliado quanto a probabilidade de ser encontrado nesses ambientes químicos e representados de forma gráfica de pontuação 3D-1D (B). As categorias P₁ e P₂ sugerem regiões de ambientes polares da proteína; B₁, B₂, B₃ sugerem regiões enterradas da proteína; e E, regiões exposta ao solvente. Adaptado de Bowie e colaboradores (1991)²⁰⁵

Desde sua introdução, a modelagem por homologia tem se mostrado útil em diversas etapas do planejamento de fármacos e/ou para a compreensão de fenômenos biológicos e farmacológicos.¹⁸² Dentre as possíveis aplicações da modelagem comparativa, a obtenção de informações estruturais acerca de mecanismos de resistência merece destaque. De fato, essa estratégia tem sido amplamente utilizada para investigar o modo de interação de antifúngicos

no sítio ativo de lanosterol 14 α -desmetilase e/ou para explicar como mutações nessa enzima estão relacionadas com o aparecimento de resistência em *C. albicans*.^{66; 206; 207; 208; 209}. Contudo, não há estudos semelhantes sobre o papel de mutações, em lanosterol 14 α -desmetilase, no desenvolvimento de resistência em *Cryptococcus*.

3.1 Objetivos

Visando estudar, do ponto de vista estrutural, as mutações na enzima lanosterol 14 α -desmetilase de *C. gattii* que conferem resistência a antifúngicos imidazólicos e triazólicos, os objetivos específicos desse trabalho são:

- Identificar cepas de *C. gattii* que apresentam mutações na enzima lanosterol 14 α -desmetilase capazes de conferir resistência à fluconazol (antifúngico triazólico);
- Sequenciar o gene que codifica a lanosterol 14 α -desmetilase em cepas de *C. gattii* resistentes a fluconazol e que não apresentam bomba de efluxo;
- Construir e analisar modelos tridimensionais da enzima lanosterol 14 α -desmetilase de cepas de *C. gattii* com mutações que conferem resistência à fluconazol.

3.2 Ensaios biológicos

Os ensaios biológicos foram desenvolvidos no Laboratório de Pesquisa em Microbiologia Clínica (LPMC) da Faculdade de Farmácia da Universidade Federal da Bahia (UFBA).

3.2.1 Microorganismo

Para o desenvolvimento desse ensaio foi utilizado um isolado clínico *C. gattii* resistente à fluconazol disponível no banco de isolados do LPMC.

3.2.2 Reagentes

Os demais reagentes utilizados nos ensaios biológicos estão listados na **TABELA 3.1**.

TABELA 3.1 – Reagentes utilizados nos ensaios de biológicos

Reagente	Marca
Ácido 4-Morfolinopropanosulfônico (MOPS)	Qhemis
Ácido clorídrico	Quimex
Ágar Batata dextrose	Acumedia
Cloreto de 2,3,5-Trifenil-Tetrazólio (TTC)	Nuclear
Cloreto de Sódio	Vetec
Glicose	Synth
Dimetilsulfóxido (DMSO)	Sigma
Fluconazol	Sigma
Hidróxido de Sódio	Nuclear
Ibuprofeno	Sigma
RPMI-1640	Sigma
Verapamil	Sigma
Triptona Bacteriológica	Sigma
Extrato de levedura	BioRad
Agar bacteriológico	Acumedia
Uréia	Synth
Tris	Promega
Ácido etilenodiamino tetra-acético (EDTA)	Mallinckrodt Chemical
Dodecil sulfato de sódio (SDS)	USB
N-lauroilsarcosina	Sigma
Fenol/Cloroformio (1:1, pH=8,0)	Sigma
Isopropanol	Sigma
Agarose	Promega
Ácido bórico	Plus One
Brometo de etídio	Plus One
Peptona bacteriológica	Sigma

3.2.3 Ensaio para determinação de Concentração Inibitória Mínima na presença de inibidores de bomba de efluxo

Os ensaios de avaliação da presença de bombas de efluxo foram adaptados a partir do trabalho de Pinto e Silva e colaboradores (2009)¹⁶⁴ utilizando microdiluição em caldo, como preconizado pela Norma M27-A2 descrita pelo CLSI.¹²³ Resumidamente, a determinação da concentração inibitória mínima de fluconazol na presença dos inibidores de bomba de efluxo (solução de ibuprofeno 1,0 mmol/L ou verapamil 0,2 mmol/L em DMSO/RPMI (1:9)) foi realizada como descrito na **seção 2.4.4.3**, utilizando-se uma solução padrão de fluconazol em RPMI 1640 (1:40) com concentração final de 128 µg/mL (0,418 mmol/L). A suspensão fúngica foi preparada como disposta na **seção 2.4.3.2** e, em seguida, foram realizadas duas diluições consecutivas, primeiramente 1:50 em RPMI 1640 e a seguir 1:20 utilizando as soluções de ibuprofeno ou verapamil previamente produzidas, de forma a atingir uma concentração celular de $5,0 \times 10^2$ a $2,5 \times 10^3$ células/mL. Após a preparação das suspensões aguardou-se 30 minutos para que os inibidores de bomba de efluxo produzissem o efeito desejado. Todos os experimentos foram realizados em duplicata e a leitura dos resultados foi realizada como descritos na **seção 2.4.4.4**.

3.2.4 Extração e sequenciamento do gene ERG11 de *C. gattii*

3.2.4.1 Extração do DNA

O procedimento de extração do DNA de *C. gattii* foi adaptado a partir do trabalho descrito por Bolano e colaboradores (2001).²¹⁰

Inicialmente, uma unidade formadora de colônia de *C. gattii*, previamente repicada em ASD, foi incubada em meio YEPD (Extrato de levedura (1%), peptona (2%) e glicose (2%)) à temperatura ambiente sob agitação de 150 RPM por 18 horas, mantendo-se a proporção líquido-ar em 1:10. Após esse período, as células foram lavadas com água mili-Q gelada e incubadas em tampão de uréia (uréia 8 M, NaCl 0,5 M, Tris 20 mM, EDTA 20 mM, SDS 2%, pH 8,0) por 4 horas à temperatura ambiente e sob agitação de 60 RPM. O sedimento celular

foi centrifugado (centrifuga Sigma Laborzentrifugen, modelo 2-3) à 4000 xg e, em seguida, ressuspenso com 400 μ L de tampão de lise (0,5% de SDS, 0,5% de N-Lauroil Sarcosina em TE (Tris HCl 10 mM, EDTA 1 mM), pH 7,5), 400 μ L de mistura fenol/clorofórmio (1:1, pH 8,0) e 400 μ L de esferas de vidro (212-300 μ m) (Sigma). O sedimento celular foi submetido à lise mecânica em três ciclos de 3 minutos sob agitação de 2500 RPM com intervalos de 3 minutos à temperatura de 4°C. A massa celular resultante foi, então, centrifugada por 15 minutos a 17000 xg (Centrifuga Minispin plus Eppendorf) e a fase aquosa foi transferida para um novo tubo estéril onde foi adicionado isopropanol gelado (na mesma quantidade da fase aquosa transferida) para a precipitação do DNA. Após 24 horas, o DNA precipitado foi ressuspenso com 400 μ L de TE e conservado à -20°C.

O DNA extraído foi submetido à eletroforese (70 Volts por 2 horas) em gel de agarose 0,8% (p/v) contendo tampão TBE 1X (Tris-base 90 mM, ácido bórico 90 mM, EDTA 2 mM pH 8,0) e revelado em solução contendo brometo de etídio.

3.2.4.2 Amplificação do gene ERG11

Para a amplificação do gene ERG11, realizada através da técnica de PCR (do inglês *Polymerase Chain Reaction*), foi utilizado o protocolo descritor por Revankar e colaboradores (2004), bem como os iniciadores (do inglês *primer*): Sequência verso (5'-ATGTCGGCAATCATCCCCCAGGTC-3') e reverso (5'-TACCTCCTGCTTGACCTCGGC-3').²¹¹

A amplificação do DNA por PCR foi realizada em volumes finais de 50 μ L contendo: 50 ng do DNA genômico, 250 ng de cada primer, 0,2 nM de cada DNTp, 3 mM de cloreto de magnésio, 1,5 U de Taq Polimerase (Promega) e 10 μ L de 5X Green buffer (GoTaq® - Invitrogen).

A amplificação foi feita em termociclador MyClycle Thermal Clycle (BioRad) nas seguintes condições: 94°C por dois minutos para abertura inicial da fita seguido de um ciclo de 94°C por 30 segundos, 61°C por 30 segundos para a hibridização e 72°C por 2 minutos para extensão. Este ciclo foi repetido 25 vezes e processo de amplificação foi finalizado com uma extensão final de 10 minutos à 72°C.

Após análise em gel de agarose como descrito no **item 3.3.4.1**, foi feita a purificação do DNA de acordo com as especificações dispostas no protocolo do kit Wizard SV Gel and PCR clean-up system (Promega).

3.2.4.3 Sequenciamento do gene ERG11

O gene ERG11 foi sequenciado utilizando o estação ABI Prism 3100 da plataforma de sequenciamento do Centro de Pesquisas Gonçalo Muniz (CPqGM) da Fundação Osvaldo Cruz (FioCruz- Ba) em parceria com a profa. Dra. Marilda de Souza Gonçalves da Faculdade de Farmácia-UFBA. O seguimento de DNA (ERG11) foi sequenciado em triplicata, na concentração de 100 ng/ μ L.

3.3 Construção de modelos 3D através de modelagem por homologia

A construção dos modelos 3D por modelagem comparativa foi realizada no Laboratório de Bioinformática e Modelagem Molecular (LaBiMM) da Faculdade de Farmácia da Universidade Federal da Bahia, utilizando estações de trabalho que empregam o sistema operacional Windows sete e onde estão instalados os programas (MODELLER Versão 9.8 (Schrödinger, LLC., San Francisco); BioEdit versão 7.05 (Tom Hall, Carlsbad - EUA); CLUSTAL X versão 2.1 (<http://www.clustal.org/>); PYMOL versão 1.2.8 (<http://pymol.org/>); e SYBYL X versão 1.1 (Tripos Inc., St. Louis, USA)) e onde foram acessados via rede os servidores (PROCHECK (http://www.jcsg.org/prod/scripts/validation/sv_final.cgi); PSIPRED versão 3.0 (<http://bioinf.cs.ucl.ac.uk/psipred/>); VERIFY 3D (http://nihserver.mbi.ucla.edu/Verify_3D/) e HINGEPROT (<http://www.prc.boun.edu.tr/appserv/prc/hingeprot/hingeprot.html>)).

3.3.1 Identificação e seleção de proteínas homólogas à lanosterol 14 α -desmetilase de *Cryptococcus gattii*

A sequência primária da enzima lanosterol 14 α -desmetilase (EC:1.14.13.70) de *C. gattii* WM276 (Código de acesso: XP_003191854.1) foi alinhada contra sequências de proteínas cujas estruturas tenham sido resolvidas por difração de raio-X, com auxílio da ferramenta BLASTp, disponível no sítio <http://blast.ncbi.nlm.nih.gov>. Nessa etapa, utilizou-se o algoritmo PSI-BLAST (do inglês *Position-Specific Iterated BLAST*), em duas interações, para identificar proteínas homólogas a sequência alvo.²¹²

As estruturas de lanosterol 14 α -desmetilase identificadas foram, então, extraídas no sítio <http://www.rcsb.org> e o grau de identidade e similaridade entre as sequências moldes e alvo foram averiguadas com auxílio do programa BIOEDIT 7.05 utilizando a matriz BOLSUM62. As proteínas moldes utilizadas na modelagem foram selecionadas com base nos seguintes critérios: identidade (>24%) ou similaridade (>40%) sequencial; e resolução estrutural (< 3,0 Å). Este procedimento levou a seleção de duas proteínas que serviram como estruturas moldes para as etapas posteriores da modelagem.

3.3.2 Alinhamento da sequência primária da proteína alvo e molde

Nesta etapa foi feito o alinhamento múltiplo, com auxílio do programa CLUSTAL X 2.1, das sequências primárias das proteínas moldes selecionadas com a sequência alvo. O resultado do alinhamento foi inspecionado e sofreu ajuste manual a fim de evitar, por exemplo, lacunas em regiões conservadas da proteína.

3.3.3 Construção do modelo por homologia

Com base no alinhamento múltiplo entre as sequências das proteínas molde e a sequência alvo, previamente produzido, o modelo por homologia foi construído utilizando o método de restrições espaciais, disponível no programa MODELLER 9.8.

A partir das equivalências sequenciais obtidas pelo alinhamento, são extraídas restrições estereoquímicas e espaciais das proteínas moldes, que são utilizadas para a geração das coordenadas dos átomos da cadeia principal e lateral dos modelos tridimensionais da sequência alvo. As restrições espaciais são combinadas com termos de energia derivadas do campo de força CHARMM para a obtenção da função de energia empregada na otimização das coordenadas atômicas do modelo. Por último, os modelos são refinados utilizando dinâmica molecular onde ocorre 4 ciclos de aumento da energia do sistema seguido de redução lenta da energia, denominado recozimento simulado (do inglês *simulated annealing*). Foram gerados 10 modelos para cada alinhamento.

As etapas finais da modelagem foram feitas em quatro ciclos utilizando as seguintes opções: *a.library_schedule = autosched.slow* (otimização máxima) e *a.md_level = refine.slow_large* (refinamento máximo). O conjunto completo de instruções e comandos utilizados está disponível no **APÊNDICE C**.

A função objetiva DOPE foi empregada para a seleção dos melhores modelos (de menor valor), os quais foram otimizados, validados e analisados como descrito a seguir.

3.3.4 Otimização dos modelos

Cada modelo selecionado passou por nova etapa de relaxamento das cadeias laterais e do grupo prostético (Heme) com auxílio do módulo *biopolymer*, disponível no programa SYBYL X. 1.1. Após a inclusão dos hidrogênios dos aminoácidos, foi definida a protonação de histidinas, glutamatos, glutaminas e asparaginas, e calculadas as cargas parciais (AMBER7 FF02) para todos os átomos da proteína, exceto o ferro do grupo heme, que teve seu estado de oxidação predefinido como +2. Em seguida, as coordenadas dos átomos de hidrogênio foram submetidas à 100 ciclos de relaxamento e minimização utilizando o método de gradiente conjugados em campo de força AMBER 7 FF02 e constante dielétrica da água (78).

Este protocolo foi repetido, ampliando a flexibilidade conformacional para as coordenadas das cadeias laterais dos aminoácidos e, posteriormente para otimização do posicionamento do grupamento heme.

3.3.5 Validação dos modelos por homologia

Os modelos comparativos foram avaliados quanto à qualidade estereoquímica utilizando as informações disponibilizadas pelo servidor PROCHECK (http://www.jcsg.org/prod/scripts/validation/sv_final.cgi), especialmente, através do gráfico de Ramachandran. A compatibilidade entre a sequência primária e o enovelamento adotado, nas estruturas obtidas por modelagem por homologia foi inspecionada com auxílio do gráfico do pontuação 3D-1D do servidor VERIFY 3D (http://nihserver.mbi.ucla.edu/Verify_3D/).

3.3.6 Mutações na enzima lanosterol 14 α -desmetilase que conferem resistência à derivados azólicos

O protocolo previamente descrito foi repetido para a sequência alvo contendo 3 mutações pontuais identificadas em 3 sequências da lanosterol 14 α -desmetilase de isolados clínicos de *C. albicans*, resistentes a derivados azólicos.¹⁵⁸ Isto levou a construção de 3 modelos por homologia, cada qual contendo uma das mutações apresentadas na **TABELA 3.2**.

TABELA 3.2 – Mutações no gene ERG11 de *C. albicans* e posição correspondente na sequência de *C. gattii*

Mutações pontuais em <i>C. albicans</i>	Mutações equivalentes em <i>C. gattii</i>
G464S	G484S
H486Y	H488Y
K143R	K156R

3.4 Análise dos modelos comparativos

A influência das mutações nos modelos da lanosterol 14 α -desmetilase de *C. gattii* foram analisados visualmente com auxílio do programa PYMOL 1.2.8. Além disso, todos os modelos foram analisados utilizando o servidor HINGEPROT, disponível no sítio (<http://www.prc.boun.edu.tr/appserv/prc/hingeprot/hingeprot.html>), que busca regiões de flexibilidade da proteína, denominadas regiões de dobradiça,²¹³ através de algoritmos baseados no modelo de rede elástica, rede gaussiana ou de rede anisotrópica são empregados.^{214; 215; 216} O modelo de rede elástica identifica as regiões de flexão da proteína, o modelo de rede gaussiana prediz a magnitude relativa das flutuações e determina qual tipo de flexão predomina e o modelo de rede anisotrópica define a direção das flutuações em cada tipo de flexão.

3.5. Resultados e discussão

Embora diversos estudos sobre bombas de efluxo em *C. albicans* tenham sido desenvolvidos, pouca informação sobre esse tipo de mecanismo de resistência foi documentado até o momento para *Cryptococcus*.

Baseado nos dados de atividade dos derivados azólicos frente a cepa *C. gattii* (resistente à fluconazol) descritos no capítulo anterior, dois mecanismos de resistência são possíveis: redução do conteúdo intracelular de derivados azólicos através do desenvolvimento de bombas de efluxo e diminuição na afinidade dos fármacos pelo sítio de ligação (enzima lanosterol 14 α -desmetilase) por meio de mutações no gene ERG11. Considerando que existem poucos estudos sobre a influência de bombas de efluxo na resistência à derivados azólicos em *Cryptococcus*, decidiu-se investigar se esse mecanismo poderia ser responsável pela menor susceptibilidade à derivados azólicos, observada em nossos estudos (vide **TABELA 2.8**, pag. 62). Por essa razão, procedeu-se a determinação da concentração inibitória mínima de fluconazol frente à *C. gattii*, na presença de inibidores conhecidos de bomba de efluxo, como verapamil e ibuprofeno.

Verapamil é um conhecido inibidor da enzima glicoproteína-P humana, sendo utilizado para reversão de resistência à terapia com fármacos citostáticos. A enzima glicoproteína-P possui alto grau identidade sequencial com os aminoácidos de proteínas transportadoras “ABC” (do inglês *ATP Binding Cassette*) de fungos.¹⁶⁵ Embora seja conhecido o potencial do ibuprofeno como inibidor de bomba de efluxo, o mecanismo ainda não foi elucidado.²¹⁷ Estes ensaios foram realizados através da técnica de microdiluição em caldo, utilizando adaptações propostas por Pinto e Silva e colaboradores (2009) no protocolo M27-A2 descrito pelo CLSI.^{123; 164}

Os valores de CIM de fluconazol frente *C. gattii* realizados, na presença de inibidores de bomba utilizados para o estudo (ibuprofeno e verapamil) indicam que não houve alteração no perfil de susceptibilidade, sugerindo que não existe interferência deste mecanismo na resistência (**Dados não mostrados**). Alternativamente, esse resultado indica que alterações de base genética (ex. mutações no gene ERG11) podem ser a causa da resistência aos antifúngicos derivados azólicos.

Visando investigar essa hipótese, procedeu-se a extração do DNA seguido de amplificação e purificação do gene ERG11 do isolado *C. gattii* em estudo, como descritos nas **seções 3.2.4.1** e **3.2.4.2**, a fim de identificar a presença de mutações que pudessem estar

correlacionadas com o perfil de resistência observado. O resultado da amplificação e purificação do gene ERG11 pode ser observado na **FIGURA 3.5**.

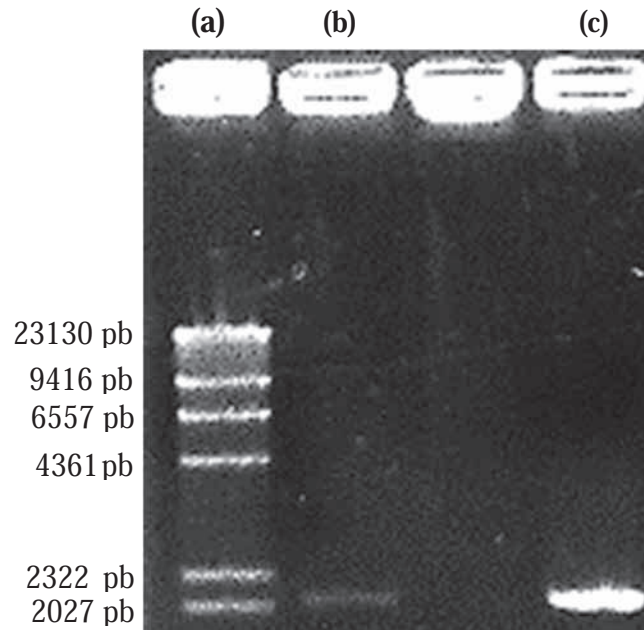


FIGURA 3.5 – Eletroforese em gel de agarose 0,8%. (a) Padrão de peso molecular Lambda DNA/HindIII (4 μ L); (b) 2 μ L de produto da amplificação do gene ERG11 de *C. gattii*; (c) 2 μ L de produto purificado da amplificação do gene ERG11 de *C. gattii*

De acordo com o padrão de peso molecular, a sequência amplificada possui tamanho aproximado de 2027 pares de base, o que está de acordo com o peso esperado segundo dados da literatura.²¹¹ A fim de confirmar que a sequência amplificada é referente ao gene de interesse, o mesmo foi sequenciado na plataforma de Sequenciamento do Centro de Pesquisas Gonçalo Muniz (CPqGM) e comparado com a sequência do gene ERG11 de *C. gattii* disponível no banco de genes (Código de acesso: XP_003191854.1) (**FIGURA 3.6**).

O alinhamento resultante indica que aproximadamente 750 pares de base no sentido do verso e do reverso têm alta identidade sequencial com o gene ERG11, confirmando assim a identidade do amplificado gênico. Contudo, não foi possível sequenciar a região central do gene, em decorrência de limitações na técnica e equipamentos disponíveis. Dessa forma, não foi possível identificar todas as oito regiões de introns presentes nessa sequência,²¹¹ tampouco averiguar a presença de mutações que pudessem estar relacionadas com a resistência, nesse primeiro momento.

Seq. C. gattii ATGTCGGCAATCATCCCC CAGGTC CAGCAACTGCTGG-GACAAGTGGCCCAATTTTCCACCGTGGTTTC 69
 Primer verso -----GGCAAGCAAAGTTGACTGAAGACAAGTGGCCCAATACTTCCACATGGTTTT 52
 Primer reverso ----- 1

Seq. C. gattii GCTGCCCTCCACCTCCTTGAAGTTCGGATCGCCGCTCGTTCGGTATCCCGCTCTCATCATTGGCTTGA 139
 Primer verso ACTGCCCTCCCTACCTCCTTGAAGATCGTGATCGGTGTCATTGGCATCCCTGCTTTTATCATTGGCTTGA 122
 Primer reverso ----- 1

Seq. C. gattii ACGTTTTTCAGCAACTTGTACGTATTGCATTGATGGGA-TTCAGCTCTTTCCCATGCTCATTGGATTCC 208
 Primer verso ACGTTTTTCACCAGCTTGTACGTGTATATTGATTGGCCTTTAGCTCTATCCTTATGCCCAATGGATTTC 192
 Primer reverso ----- 1

Seq. C. gattii CAAGCTGATCACAAGCTCCCTGACGCGTTAT-TCATCCGTGCGAGTGTCTTCCCGTAGAAAAGATCTTC 277
 Primer verso CAAGCTGACCACAAGCTCTCTGACGCGTTATATCGTCCGTACAGTGCCTTCCCGTAGAAGAGATCTTC 262
 Primer reverso ----- 1

Seq. C. gattii CTCTGTTGTCTTTCACTACATTCCCTGGTTTGGCTCAGCCGCTTACTATGGTGAAGATCCCTACAAGTT 347
 Primer verso CTCTGTTGTCTTTCACTATATTCCCTGGTTTGGCTCAGCTGCTTACTATGGTGAAGATCCCTACAAGTT 332
 Primer reverso ----- 1

Seq. C. gattii CTTGTTCGAATGCCGTGACAAATATGGAGACTTATTCACTTTCATCCTTATGGGTGCAAGGATTACCGTC 417
 Primer verso CTTGTTCGAGTGCCGTGACAAGTATGGAGACTTATTACTTTCATCCTTATGGGTGCAAGGATTACTGTC 402
 Primer reverso ----- 1

Seq. C. gattii GCGCTTGGTCTAAGGGTAACAACCTTTCTTTGGGTGAAAGATTTCTCAAGTCTCTGCCGAGGAAGCCT 487
 Primer verso GCGCTCGGACCCAAGGGTAACAACCTTTCTTTGGGTGAAAGATTTCTCAAGTCTCTGCCGAGGAAGCCT 472
 Primer reverso ----- 1

Seq. C. gattii ACACGTAAAGCTTATATGCTTTACTGGCTCAATATGGCTTACTTATTATCGCTTGGTAGCACTTGACTAC 557
 Primer verso ACACGTAAAGCTTATATGCTTCAATAGTTCAAGATGGCTTACTTACATTTGCTTGGTAGCACTTGACTAC 542
 Primer reverso ----- 1

Seq. C. gattii TCCCGTCTTTGGCAAAGGTGTTGTTTACGACTGCCCTAACGAGATGCTCATGCAGCAGAAGAAGTTTGTG 627
 Primer verso TCCCGTCTTTGGCAAAGGTGTTGTTTATGATTGCCCTAACGAGATGCTCATGCAGCAGAAGAAGTTTGTG 612
 Primer reverso ----- 1

Seq. C. gattii AGTGGATACCATTGCGAGCTTGATTCGCAAGCTCACTATTCTACAGATCAAGTCCGGTCTTACTACCGAG 697
 Primer verso AGTGGAAACCACACGCAG-TTGATTCACAAGCTCATTCTCTAAAGATCAAGTCCGGTCTTACAAGTGA 681
 Primer reverso ----- 1

Seq. C. gattii TCCCTTCAATCGTATCCTCCTATGATCACCAGCGAATGCGAAGATTTCTTCA CCAAAGAAGTCGGGATTT 767
 Primer verso TCCCTTCAATCTTATCCTCCATGATCACTAGCGAATGCGAAGATTTCTTCA----- 733
 Primer reverso ----- 1

Seq. C. gattii CTCCCCAGAAGCCTTCTGTACCCTCGACCTCCTCAAAGCCATGTCGGAGCTCATCATTCTTACTGCGTC 837
 Primer verso ----- 733
 Primer reverso ----- 1

Seq. C. gattii TCGTACTCTCCAGGAAAGGAAGTTCGTGAATCTCTTAATGGTCAGTTCGCCAAGTACTACGAGGATCTC 907
 Primer verso ----- 733
 Primer reverso ----- 1

Seq. C. gattii GACGGCGGTTTCACTCCTCTCAACTTTATGTTCCCAACTTGCCCTTCCAGTTACAAAAGGCGAGATG 977
 Primer verso ----- 733
 Primer reverso ----- 1

Seq. C. gattii AGGCTCAGAAGGCTATGAGCGACTTTTACTTGAAGATCATGGAGAACAGGAGGAAGGGTAAAAGTGACGT 1047
 Primer verso ----- 733
 Primer reverso ----- 1

Seq. C. gattii GAGTTGATTTCAAGTGTGAAGGAGAAAAGTCTGATTTGAGTAGCATGAGCACGACATGATTGAAAACC 1117
 Primer verso ----- 733
 Primer reverso ----- 1

Seq. C. gattii TCCAGAGCTGCAAGTACCAGAAACGGTGTCCCTCTCTGACCGTGACATTGCCACATCATGATTGCTCT 1187
 Primer verso ----- 733
 Primer reverso ----- 1

Seq. C. gattii TCTTATGGCTGGCCAGCACACTTCTCCGCTACTTCTTCTTGGACTCTTCTCCATCTTGTGACCGACCC 1257
 Primer verso ----- 733
 Primer reverso ----- 1

Seq. C. gattii GACGTTGTGTACGTTTTTTTTGGAGAAATTTGAATGGCGAAGCATAACTGACGAGGATGTAGAGAGGCTC 1327
 Primer verso ----- 733
 Primer reverso ----- 1

Continua...

Seq. <i>C. gattii</i>	TTTATCAGGAGCAGAAACAAAAGCTCGGTAAACCCTGACGGTACTTTCCGAGACTACAGATACGAGGACCT	1397
Primer verso	-----	733
Primer reverso	-----AGACTACAAGTACGAAGACCT	21
Seq. <i>C. gattii</i>	CAAGGAGTTGCCCATCATGGACTCTATCATCCGAGAAACTCTCCGAATGGTAAGTCAT-CCGACTTGTA	1466
Primer verso	-----	733
Primer reverso	CAAGGAGTTGCCCATCATGGACTCTATCATCCGAGAGACTCTTCGAATGGTAAGTCATTCTCGACTTTTA	91
Seq. <i>C. gattii</i>	TTTCGATACCTTACTAACC AAAATCCACTTCTAGCACGCTCCCATCCAGTAAGCCACGGACGCC---TCC	1533
Primer verso	-----	733
Primer reverso	TTTCGGATACCGTACTGACCTAACCCACTTTTAGCACGCTCCCATCCAGTAAGCTACGAATGCCACATTC	161
Seq. <i>C. gattii</i>	TTTATAAATGTACTAATAATAATAGCTCATCTATCGAAAAGTCTTTCCGACATCCCCGTCCCTCC	1603
Primer verso	-----	733
Primer reverso	ACCATATAAATCGTATTAACTTAAAGTAGCTCTATTTACCGAAAAGTCTTTCCGACATCCCCGTCCACCC	231
Seq. <i>C. gattii</i>	AGTCTTTCCGCCCTTCTGAGAACGGCCATACATCATCCCAAGGGCCACTACATATATGGCCGCCCTG	1673
Primer verso	-----	733
Primer reverso	AGTCTTGCCGCCCTTCCGAGAACGGCCAAATACATATCCCAAGGGTCACTACATCATGGCCGCCCTG	301
Seq. <i>C. gattii</i>	CGGTCTCCCAAATGGACCCTCGTATCTGGCAAGA GCCAAGGTCTGGAACCTGCCGATGGCAGCATGA	1743
Primer verso	-----	733
Primer reverso	CGGTCTCTCAAATGGACCCTCGGATCTGGCAAGATGCCAAGGTCTGGAATCCTGCCGATGGCAGCACGA	371
Seq. <i>C. gattii</i>	AAAGGGATTGCGCCGCCGCCATGGTCCAATACACTAAGGCGGAACAAGTCGACTATGGTTTCGGTTCT	1813
Primer verso	-----	733
Primer reverso	AAAGGGATTGCGCCGCCGCCATGGCCAGTACACCAAAGCCGAGCAAGTCGACTATGGTTTCGGTTCT	441
Seq. <i>C. gattii</i>	GTCAGCAAGGGTACCGAATCTCCTTACCAGCCTTTGGCGCCGGTAGGCACAGATGTGTGGTGAACAAT	1883
Primer verso	-----	733
Primer reverso	GTCAGCAAGGGTACCGAGTCTCCTTACCAGCCTTTGGTGCTGGCAGGCACAGGTGTGTGGTGAACAGT	511
Seq. <i>C. gattii</i>	TTGCGTACACTCAGCTTTCAACCATCTTACCTATGTTGTGAGGAATTTCACTTTGAAGCTTGCTGTTCC	1953
Primer verso	-----	733
Primer reverso	TTGCGTACACTCAGCTTTCCACCATCTTACCTATGTTGTGAGGAATTTCACTTTAAAGCTTGCCGTTCC	581
Seq. <i>C. gattii</i>	CAAGTTCCCGAAACCAATTACCGTGTAAGTCTTTGTGTTCCAATCTATCATGGGATGATTTGCTGATT	2023
Primer verso	-----	733
Primer reverso	CAAGTTCCCGAAACCAATTACCGTGTAAGTCTTTGTGTTCCAATCTACCC--GGATGATTTGTTGATT	649
Seq. <i>C. gattii</i>	CTTCATTGTGTAGACCATGATTGTCCAACCAACAACCCCTTGGTTACTTTACCCCTTCGAAACGCCGAG	2093
Primer verso	-----	733
Primer reverso	GCAT-TCGTGTAGACCATGATTGTCCAACCAACAACCCCTCATCTACAT-CATC-ATCGAAATC-----	711
Seq. <i>C. gattii</i>	GTCAAGCAGGAGGTATGA	2111
Primer verso	-----	733
Primer reverso	-----	711

FIGURA 3.6 – Alinhamento sequencial do da sequência genômica da ERG11 de *C. gattii* (Seq. *C. gattii*) contras as sequências obtidas após o sequenciamento com destaque às regiões conservadas (em cinza), o primer verso (em verde) e a sequência complementar do primer reversos (em azul)

Visando contornar essa limitação, projetos de pesquisa voltados para a extração do RNA de *C. gattii* e obtenção de uma biblioteca de cDNA, com o objetivo de amplificar e sequenciar apenas as regiões codificantes do gene ERG11 foram iniciados em colaboração com outros alunos de mestrado. Essa estratégia se mostra promissora, uma vez que seria necessário sequenciar apenas 800 pares de base e os *primers* necessários para esses experimentos são iguais aos utilizados inicialmente no protocolo de amplificação do DNA genômico. Apesar dos esforços realizados nos últimos 6 meses, ainda não foi possível obter a

sequência completa do gene ERG11, tampouco maiores informações acerca de mutações presentes nessa sequência.

Apesar dos problemas encontrados, decidiu-se prosseguir com o trabalho originalmente proposto, utilizando para isso dados já publicados na literatura.

Muito embora mutações no gene ERG11 sejam relativamente comuns, apenas uma pequena parcela causa resistência à derivados azólicos.¹⁵⁸ Exemplo clássico disso são as mutações pontuais G464S, H468Y e K143R, identificadas em isolados clínicos de *C. albicans* resistente à derivados azólicos,^{158; 168} sendo que apenas a primeira delas, identificada como G484S na sequência de *Cryptococcus*, já foi encontrada em isolados de *C. neoformans* resistentes a derivados azólicos.²¹¹ Visando investigar o impacto que as mutações G484S, H488Y e K156R teriam sobre a estrutura tridimensional de Lanosterol 14 α -desmetilase, utilizou-se a metodologia de modelagem por homologia para construir modelos tridimensionais da enzima nativa e de seus mutantes.

O primeiro requisito para se realizar essa tarefa é encontrar proteínas homólogas à sequência alvo que servissem de molde para construção do modelo. Assim, a partir da sequência primária da enzima lanosterol 14 α -desmetilase (EC: 1.14.13.70) de *C. gattii*, disponível no banco de genes (Código de acesso: XP_003191854.1), identificou-se 25 sequências homólogas, cujas estruturas cristalográficas estão disponíveis no sítio www.rcsb.org (**TABELA 3.3**).

TABELA 3.3 – Estruturas cristalográficas de proteínas homologas à proteína alvo

Código PDB	Organismo	Resolução	Identidade (%)	Similaridade (%) [*]
2W09	<i>M. tuberculosis</i>	1,57	24,55	40,79
2W0B	<i>M. tuberculosis</i>	1,56	24,55	40,79
2W0A	<i>M. tuberculosis</i>	1,60	24,36	40,61
2VKU	<i>M. tuberculosis</i>	1,95	24,55	40,79
2CIB	<i>M. tuberculosis</i>	1,50	24,55	40,79
2CI0	<i>M. tuberculosis</i>	1,53	24,55	40,79
1X8V	<i>M. tuberculosis</i>	1,55	24,55	40,79
1U13	<i>M. tuberculosis</i>	2,01	24,73	40,79
1H5Z	<i>M. tuberculosis</i>	2,05	24,55	40,79
1EA1	<i>M. tuberculosis</i>	2,21	24,55	40,79
1E9X	<i>M. tuberculosis</i>	2,10	24,55	40,79
3LD4	<i>L. infantum</i>	2,75	24,60	40,39
3JUS	<i>H. sapiens</i>	2,90	32,49	49,37
3JUV	<i>H. sapiens</i>	3,12	32,49	49,37
3LD6	<i>H. sapiens</i>	2,80	32,49	49,37
3K1O	<i>T. cruzi</i>	2,89	23,80	40,85
2WUZ	<i>T. cruzi</i>	2,35	23,98	41,56
2WX2	<i>T. cruzi</i>	2,27	23,98	41,56
3KHM	<i>T. cruzi</i>	2,85	23,98	41,21
3KSW	<i>T. cruzi</i>	3,05	23,80	40,85
2WV2	<i>T. brucei</i>	2,70	23,36	41,06
2X2N	<i>T. brucei</i>	2,60	23,19	40,70
3G1Q	<i>T. brucei</i>	1,89	22,94	40,14
3GW9	<i>T. brucei</i>	1,87	22,94	40,32

*O grau de similaridade foi calculado com base na matriz de similaridade BLOSUM62

A escolha das proteínas molde é uma etapa crucial, pois a precisão do modelo é diretamente afetada pelo grau de identidade/similaridade sequencial entre as sequências alvo e molde. Outro fator que influencia a qualidade do modelo é a resolução das proteínas molde, quanto maior a resolução, mais confiável é a informação estrutural utilizada na construção do modelo. Levando em consideração esses dois fatores, selecionaram-se as estruturas de

lanosterol 14 α -desmetilase de *Mycobacterium tuberculosis* (2CIB) e humana (3LD6) como moldes, por apresentarem respectivamente melhor resolução e maior identidade sequencial.

O alinhamento múltiplo das sequências primárias das proteínas alvo (*C. gattii*) e moldes (2CIB e 3LD6) mostra que as inserções/deleções encontradas na sequência de *C. gattii* não se encontram em regiões de estrutura secundária conservada, o que evidencia a qualidade do alinhamento obtido (FIGURA 3.7). Todavia, em função da baixa identidade sequencial observada (<33%), não se pode garantir que o alinhamento sequencial represente o alinhamento estrutural dessas proteínas.

2CIB	-----MSAVALPRVSGG	60
3LD6	-----MAKKTLPAGVKS	60
<i>C. gattii</i>	<u>MSAIIIPQVQQLLGQVAQF</u> <u>FPPWF</u> <u>AA</u> <u>LPTSLKVAIAVVGIPALIIGLNVFQQLCLPRRKDL</u>	60
2CIB	HDEHGHLEEFR-----TDPIGLMQRVRDELGDVGTFLAGKQVVLLSGSHANEFFFR	120
3LD6	PPYIFSPIPFLGHAIAFGKSPIEFLENAYEKYGPVFSFTMVGKTFITYLLGSDAALLFNS	120
<i>C. gattii</i>	PPVVFH <u>YIPWFGSAAYYGE</u> D <u>PKKFLFE</u> CRDKYGD <u>LF</u> T <u>FIL</u> MGR <u>RITVAL</u> G <u>PKGN</u> NLSLGG	120
2CIB	GDDLLDQAKAYPFMT-PIFGEGVVFDA SPERRKE --MLHNAALRGEQMKGHAATIEDQVR	180
3LD6	KNEDLNAEDVYSRLTTPVFGKGVAYDVPN PVFLEQKK MLKSGLNIAHFQHVSIIEKETK	180
<i>C. gattii</i>	KISQVSAEEAYTHLTPVFGKGVVYDCPN EMLMQQKFI KSGLT TESLQSY PPMITSECE	180
2CIB	RMIADW G-----EAGEIDLLDFFAELTIYTSSACLIGKKFRDQLDGRFAKLYHELERGT	240
3LD6	EYF-ESW GESG EK ----- NVFEALSELIILTASHCL HGKEIRSQLNEKVAQLYADLDGGF	240
<i>C. gattii</i>	DFFTKEV <u>GISPQKPSATLD</u> L LKAMSELIILTASRTL QKQEV RESLNGQFAKYEDLDGGF	240
2CIB	DPLAYVDP-YLPIES FRRRDEARNGLVALVADIMNGRI ANPPTDKSDR DMLDVLIA VKAE	300
3LD6	SHAAWLLPGWLPLP SFRRRDRAHREIKDIFYKAIQKRRQ S---QEKIDDILOTTLLDATYK	300
<i>C. gattii</i>	TPLNFMFP-NLPLPS YKRRDEAQKAMSDFYLKIMENRR KG--ESDHEHD MIENLQ SCKYR	300
2CIB	TGTPRF SADEITGMFISMMF AGHHTSSGTASWT LIELMRHRDAYAAVIDELDELY GDG--	360
3LD6	DGRP-LT DDEVAGMLIGLLLAGQHTSSTTS AWMGFFLARD KTLQKKCYLEQKTV CG---E	360
<i>C. gattii</i>	NGVP-L SDRDIAHIMIALLMAGQHTS SATSSWTLLHLA DRPDVVEALYQE QKQKLGNDPG	360
2CIB	--RSVS FHALRQI PQLENVLKETL RLHPLIILMRVAKGE FEV QG----- HRI HE	420
3LD6	NLPLPT YDQLKDLNLLDRCIKETL RLRPPIMIMRMAR-----TP QTV AG----- YT IIP	420
<i>C. gattii</i>	TFRDYR YEDL KELP IMDSI IRETLRMHAPIHSI YRKVL <u>SDIPVPPSLSAPSENGQYII</u> PK	420
2CIB	GD LVAASPAIS NRIPEDFPDPHDFVPARYEQPR-----QEDLLN	480
3LD6	GH QVCVSP TVNQRLKDSWVERLDFNPDRYLQD-----NPASGE	480
<i>C. gattii</i>	GH YIMAAPGVS QMDPRIWQDAKVNPARWHDE KGF <u>AAAAMVQYTKAEQVDYGF</u> GSVSKGT	480
2CIB	RWT WI PFGAGRHR CVGAAFAIMQIKAI F SVLLREY EFEMA QPPESYRNDHSMVQLAQP	540
3LD6	KFAYVPPF GAGRHR IGENFAYVQIKTIWSTMLRL YEF D LIDG -YFPTVNYTTMIHTPENP	540
<i>C. gattii</i>	ESPYQPF GAGRHR CVGEQFAYTQLSTIFTYVVR NF TLK LAVP-KFPE TNYRTMIVQ PNNP	540
2CIB	AAVRYRR RTGVHHHH 555	
3LD6	VIRYKRR STHHHHHH 555	
<i>C. gattii</i>	LVTF TLRNAEVKQEV 555	

FIGURA 3.7 – Alinhamento das sequências das proteínas moldes (2CIB e 3LD6) contra a sequência alvo (*C. gattii*) destacando as estruturas secundárias hélices alfa (vermelho) e fitas beta (azul) preditas pelo programa PSIPRED 3.0 (<http://bioinf.cs.ucl.ac.uk/psipred/>).^{228;229} As regiões sublinhadas tiveram suas sequências ajustadas manualmente

Por essa razão, foram realizadas intervenções manuais no alinhamento, como por exemplo, a remoção de 52 resíduos no N-terminal da sequência alvo. Essa região está ausente na enzima de *M. tuberculosis* por não estar ancorada na membrana e foi removida da enzima humana a fim de facilitar a determinação de sua estrutura cristalográfica.^{173; 218} De forma semelhante os resíduos que codificam a cauda de histidina (resíduos 548 á 555), necessária no processo de purificação das proteínas molde foram desconsiderados. Além disso, ajustes secundários no alinhamento dos resíduos 67 à 69, 187 à 199, e 397 à 414 foram feitos de forma a coincidir as regiões de alfa hélice e fitas beta das proteínas alvo e moldes, e a inserção composta pelos resíduos 454 à 474 foi excluída uma vez que não existia moldes disponíveis para a região. O alinhamento final, para a enzima nativa, utilizado para construção do modelo está representado na **FIGURA 3.8**.

```

2CIB      ALPRVSGGHDEHCHLEEFRTDPIGLMQVRDELGDVGTFLQLAGKQVLLSGSHANEFFFR 60
3LD6      SPPYIFSPPIPFLLGHAIAFGKSPLEFLENAYEKYGPVFSFTMVGKTFITYLLGSDAAALLFN 60
C. gattii LPPVVFHYIPWFGSAAYYGEDPYKFLFECRDKYGDLFTFILMGRRITVALGPKGNLNSLG 60

2CIB      AGDDDLQAKAYPFM--TPIFG-----LRGEQMKGHAATIEDQ 120
3LD6      -SKNEDLNAEDVYSRLTTPVFGKGVAYDVPNPVFLEQKMLKSGLNIAHFKQHVSIIEKE 120
C. gattii -GKISQVSAEEAYTHLTPVFGKGVVYDCPNEMLMQKKFIKSLTITESLQSYPPMITSE 120

2CIB      VRRMIADW-GEAGEI-----DLLDFFAELTIYTSACLI GKKFRDQLDGRFAKLYHELER 180
3LD6      TKEYFESW-GESGEK-----NVFEALSELIILTASHCLHGKEIRSQLEKVAQLYADLDG 180
C. gattii CEDFFTKEVGI SPQKPSATL DLLKAMSELIILTASRTLOGKEVRESLNGQFAKYEDLDG 180

2CIB      GTDPLAYVDP-YLPTESEFRRRDEARNGLVALVAD-IMNGRIANPDR--DMLDVLI AVKAE 240
3LD6      GFSHAAWLLPGWLPPLPSFRRRDRAHREIKDIFYKAIQKRRQSQ-EKIDDILOQLLD--AT 240
C. gattii GFTPLNFMFP-NLPLPSYKRRDEAQKAMSDFYLK-IMENRRKG-ESDHEHDMIENLQSC 240

2CIB      T--CTPRFSADEITGMFISMMFAGHHTSSGTASWTLIELMRHRDAYAAVIDELDELYG-- 300
3LD6      YKDCRPL-TDDEVAGMLIGLLLAGOHTSSSTTSAMGFFLARDKTLQKKCYLEQKTVCG-- 300
C. gattii YRNCVPL-SDRDIAHIMIALLMAGOHTSSATSSWTLHLADRDPVVEALYQEQKQKLG 300

2CIB      DG--RSVSFHAI RQIPQLENVLKETLRLHPPLIILMRVAKGEFEVQGHRIHEGDLVAASP 360
3LD6      -ENLPPLTYDQIKDLNLLDRCIKETLRLRPPIMIMMRMARTPQTVAGYTIIPGHQVCVSP 360
C. gattii DGTFRDYREDLKELPIMDSIIRETLRMHAPIHSIYRKVLSDIPVPPYIIPKGYHMAAP 360

2CIB      AISNRIPEDFPDPHDFVPAARYEQPRQEDLLNRWTWIPFGAGRHRVCGAAFAIMQIKATIFS 420
3LD6      TVNQRLKDSWVERLDFNPDRYLQD-NPASGEKFAYVPFGAGRHRVCI GENFAYVQIKTIWS 420
C. gattii GVSQMDPRIWQDAKVWNPARWH-DESVSKGTESPYQPFAGAGRHRVCGEQFAYTQLSTIFT 420

2CIB      VLLREYEFEMA--QPPEYRNDHSMVVQLAQPAAVRYRRR 461
3LD6      TMLRLYEFDLI--DGYF-PTVNYTTMIHTPENPVIRYKRRS 461
C. gattii YVVRN--FTLKLAVPKF-PETNYRTMIVQPNNPLVTFTLRN 461

```

FIGURA 3.8 – Alinhamento das sequências das proteínas moldes (2CIB e 3LD6) contra a sequência alvo (*C. gattii*) utilizado na construção dos modelos por homologia. Em destaque podem ser observados os resíduos conservados ao longo da evolução

A partir do alinhamento gerado para a enzima nativa, foram feitas modificações pontuais na seqüência primária da enzima lanosterol 14 α -desmetilase de *C. gattii*, visando simular as mutações encontradas em isolados clínicos de *C. albicans* resistentes à derivados azólicos (**TABELA 3.2**).¹⁵⁸ Esse procedimento deu origem a 3 alinhamentos que foram também utilizados para construir modelos por homologia da enzima-alvo, como descrito na **seção 3.3.5**. É importante salientar que algumas alterações no gene ERG11 de *C. albicans*, associados a resistência à derivados azólicos (ex. D116E e D278E) são constitutivas em *C. gattii* (E129 e E289), podendo portanto contribuir para a menor susceptibilidade de *C. gattii* frente à fluconazol.

Para cada alinhamento, foram construídos 10 modelos por homologia, os quais foram avaliados com base no valor da função de energia DOPE. Modelos com menor valor estão associados com melhor qualidade do enovelamento e, portanto, devem conter menor taxa de erros estereoquímicos na estrutura.²¹⁹ Apenas o modelo de menor energia para cada alinhamento foi considerado nas etapas seguintes.

Embora seja possível construir modelos por homologia que apresentam grupos protéticos (ex. grupo heme), deve-se analisar com restrições esses modelos, uma vez que as coordenadas do grupamento prostético são apenas transferidas da proteína-molde para a proteína alvo, sem que qualquer etapa de minimização e refinamento da estrutura seja executada. Por essa razão, mutações que sabidamente causam choques estéricos com o grupo Heme (ex. G484S) não afetam o posicionamento do grupo prostético no modelo gerado por essa metodologia. Visando corrigir essa limitação da técnica e assim gerar modelos que descrevam de forma mais adequada as estruturas de lanosterol 14 α -desmetilase mutantes, adotou-se a seguinte estratégia: As estruturas tridimensionais passaram por um processo de relaxamento da estrutura no qual foi dada liberdade rotacional aos hidrogênios. Numa segunda etapa, essa liberdade foi estendida aos resíduos próximos ao grupo heme e por fim ao próprio grupo heme.

Uma maneira de avaliar o impacto dessa abordagem sobre a estrutura tridimensional dos modelos é através de parâmetros estereoquímicos. Esse tipo de análise pode ser feito a partir do gráfico de Ramachandran, no qual os resíduos das proteínas são analisados em função dos seus ângulos phi e psi. A análise do gráfico de Ramachandran para o modelo da enzima nativa antes e depois desse procedimento mostra que não houve distorções graves na estrutura uma vez que a porcentagem de resíduos nas regiões permitidas não se alterou (98,2%). A diferença mais marcante entre o modelo da enzima antes e depois da minimização ocorreu no número de resíduos nas regiões generosamente permitidas (em amarelo claro no

gráfico de Ramachandran) que passou de 7,2% para 9,7%. O número de resíduos em regiões não permitidas foi igual (sete resíduos), equivalente à 1,8% do total de resíduos da proteína. (FIGURA 3.9A e FIGURA 3.9B). Por essa razão, os resultados apresentados de agora em diante sempre se referem as estruturas minimizadas.

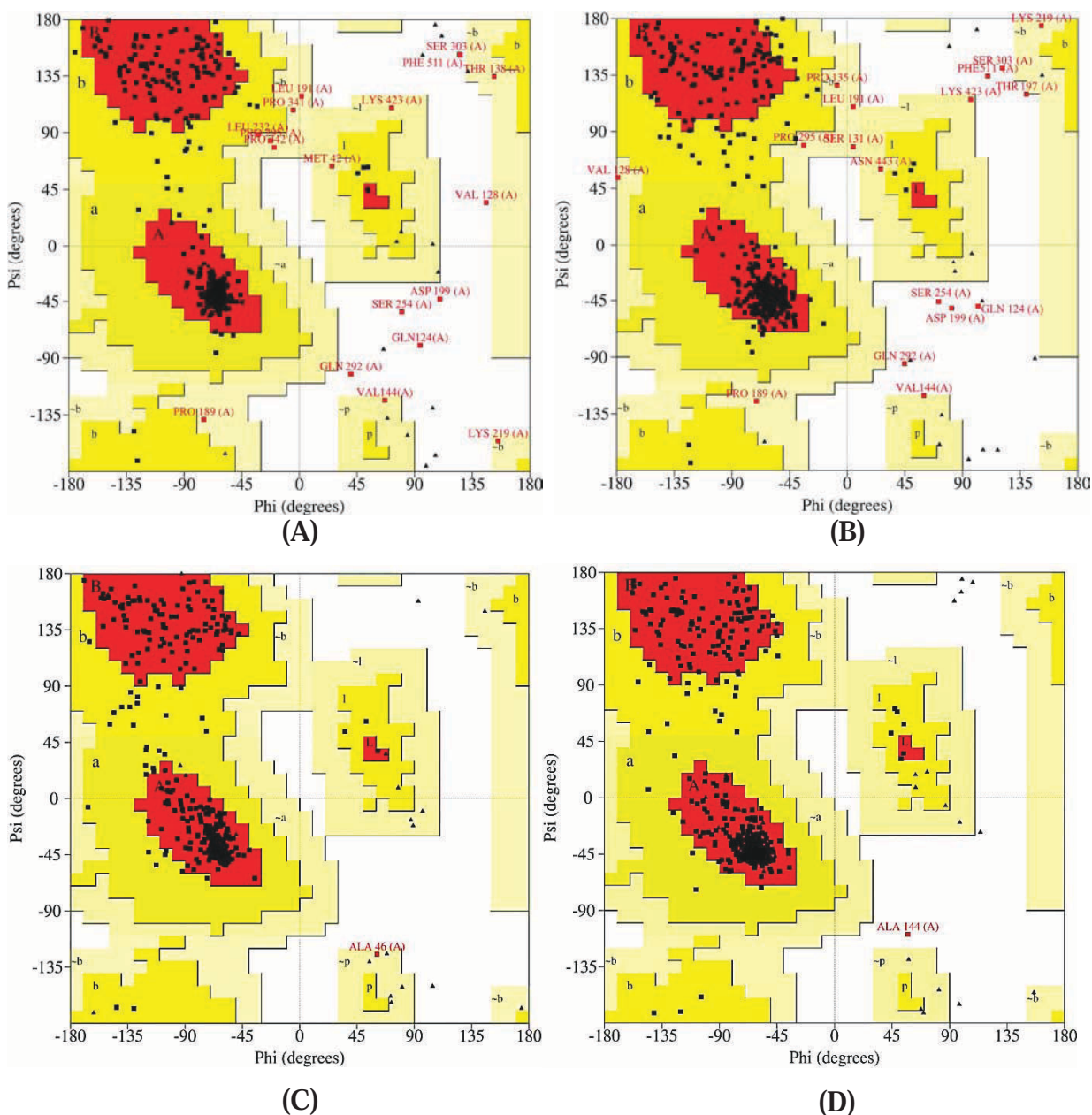


FIGURA 3.9 – Gráfico de Ramachandran para os modelos da enzima nativa antes da minimização (A) e após a minimização (B), e para as proteínas moldes 2CIB (C), 3LD6 (D). A porcentagem de resíduos nas regiões permitidas foi: 99,9% para as proteínas moldes (2CIB e 3LD6) e 98,2% para os modelos da enzima nativa antes e após a minimização. As regiões favoráveis estão destacadas em vermelho, as regiões permitidas em amarelo

Espera-se que a qualidade estereoquímica do modelo gerado por homologia seja comparável ao das proteínas moldes (**FIGURA 3.9C** e **FIGURA 3.9D**). Porém é notório que os modelos produzidos apresentam um número superior de resíduos nas regiões não permitidas quando comparados às proteínas molde. Portanto fica evidente, que os modelos gerados não representam, de forma correta, a estrutura completa da lanosterol 14 α -desmetilase de *C. gattii*. Entretanto, deve-se levar em conta que os resíduos encontrados em regiões não permitidas do gráfico de Ramachandran (GLN-124, THR-197, ASP-199, SER-254, GLN-292, SER-303, PHE-511), estão localizados em alças na superfície da enzima e, portanto, não devem afetar significativamente a conformação dos resíduos no sítio ativo da enzima (**FIGURA 3.10**).

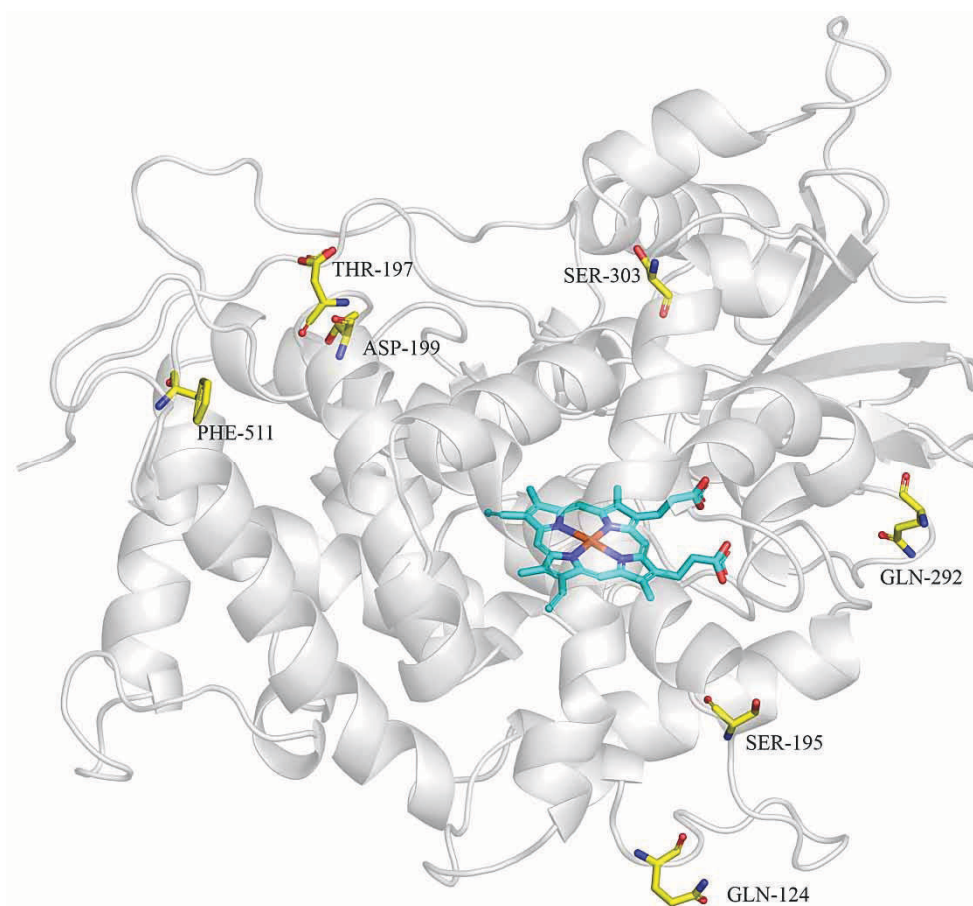


FIGURA 3.10 – Resíduos do modelo da enzima nativa encontrados na região não permitida do gráfico de Ramachandran (em amarelo) e o grupamento heme (em verde)

Resultados qualitativamente semelhantes são observados para os outros modelos, uma vez que a porcentagem de resíduos encontrados nas regiões permitidas e generosamente permitidas varia de 98,7% e 99,2%. Em todas as estruturas tridimensionais obtidas por

modelagem comparativa, não se observa resíduos que sofreram mutação em regiões não permitidas do gráfico de Ramachandran.

Embora o gráfico de Ramachandran seja uma boa medida da qualidade estereoquímica do modelo, ele não é suficiente para validar o modelo obtido. Por essa razão, os modelos também foram avaliados em relação à compatibilidade da sequência com o enovelamento proteico adotado, como disponível no servidor VERIFY 3D (http://nihserver.mbi.ucla.edu/Verify_3D/). Nessa análise, a pontuação 3D-1D avalia se a posição de cada resíduo no modelo 3D está de acordo com sua ocorrência esperada para um determinado ambiente químico. Dessa forma, resíduos que não se encontram em seu “ambiente preferencial” recebem pontuação mais baixa e podem ser facilmente identificados através de uma análise gráfica da pontuação 3D-1D, em função do número do resíduo na sequência primária.

De uma maneira geral, o modelo construído para a enzima nativa apresenta valores positivos de pontuação 3D-1D (**FIGURA 3.11**), o que sugere que o enovelamento proposto é compatível com a sequência primária utilizada. A região compreendida entre os resíduos 221 e 224, relativa a uma alça exposta ao solvente localizada entre duas hélices alfa (**FIGURA 3.12 – em vermelho**), representa uma exceção, uma vez que apresenta valores abaixo de zero.

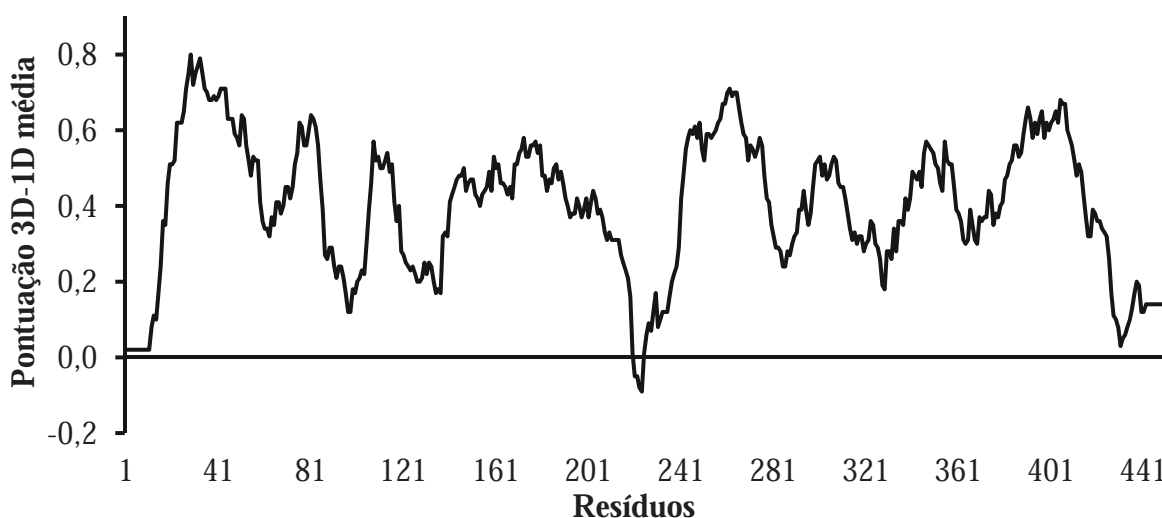


FIGURA 3.11 – Gráfico de pontuação média 3D-1D, considerando uma janela de 21 aminoácidos, segundo o servidor VERIFY 3D, do modelo comparativo da enzima nativa

Contudo, vale a pena destacar que a região correspondente nas proteínas molde também apresentam valores de pontuação 3D-1D comparativamente baixos (**Figura 3.13**). Portanto, é possível que a enzima lanosterol 14 α -desmetilase de *C. gattii*, de fato adote o enovelamento observado no modelo por homologia.

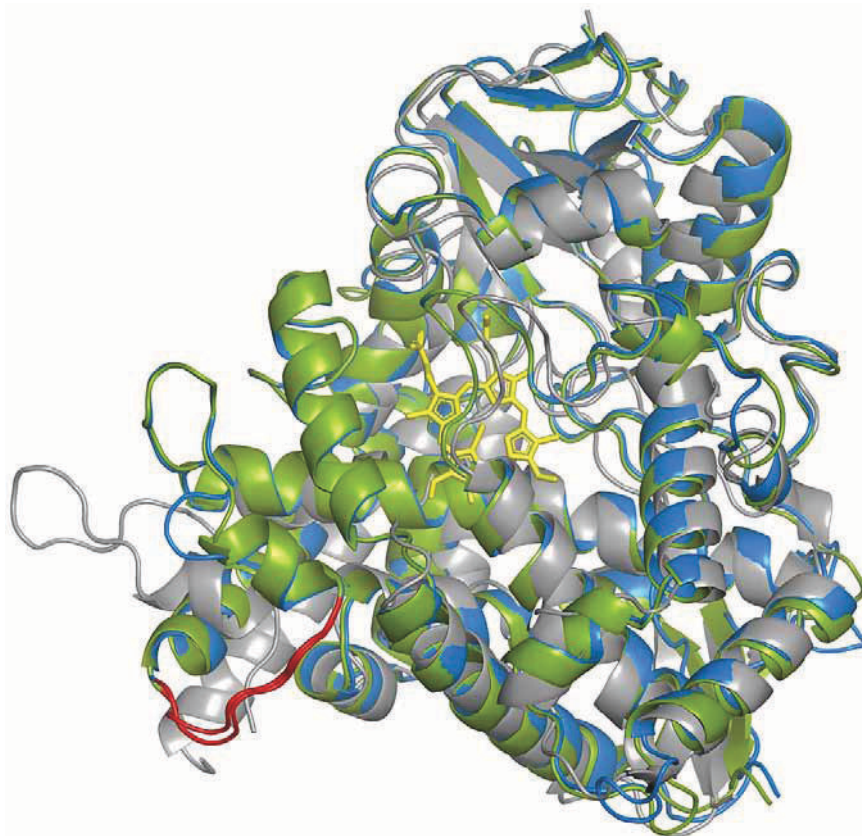


FIGURA 3.12 – Modelo da lanosterol 14 α -desmetilase de *C. gattii* nativa (em azul) sobreposta às estruturas moldes: 2CIB (em cinza) e 3LD6 (em verde). Em destaque (vermelho) o fragmento que apresenta baixa pontuação 3D-1D no programa VERIFY 3D

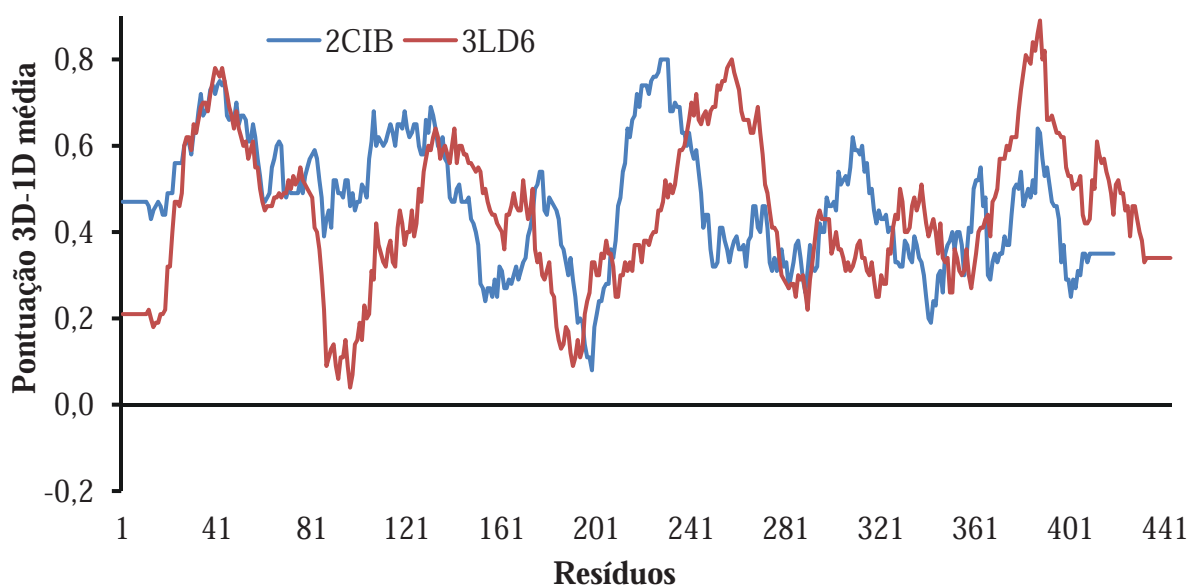


FIGURA 3.13 – Gráfico de pontuação média 3D-1D das proteína moldes (2CIB e 3LD6). Os resíduos 195 à 199 em 2CIB e 191 a 195 em 3LD6 correspondem a região delimitada entre os resíduos 221 e 224 no modelo da enzima nativa, vide FIGURA 3.11

A partir das análises realizadas, pode-se concluir que o modelo comparativo da lanosterol 14 α -desmetilase de *C. gattii* apresenta qualidade esteroquímica aceitável e enovelamento compatível, podendo, portanto fornecer informações relevantes sobre o papel de mutações pontuais no mecanismo de resistência à derivados azólicos.

Levando em consideração que a interação dos derivados azólicos com o ferro do grupo heme é fundamental para ação desses fármacos, realizou-se uma análise detalhada das interações que esse grupo prostético realiza na estrutura da enzima nativa. As cadeias alquílicas do grupo heme são circundadas por resíduos hidrofóbicos, com os quais ele realiza interações de Van Der Waals (ex. PRO-385; LEU-163, LEU-380; PHE-483; PHE-495; THR-321). As cadeias contendo ácidos propanóicos, por sua vez, fazem ligações de hidrogênio com os resíduos TYR-131, TYR-145, ARG-391 e HIS-488, enquanto o ferro se coordena com a CYS-490 (**FIGURA 3.14**).

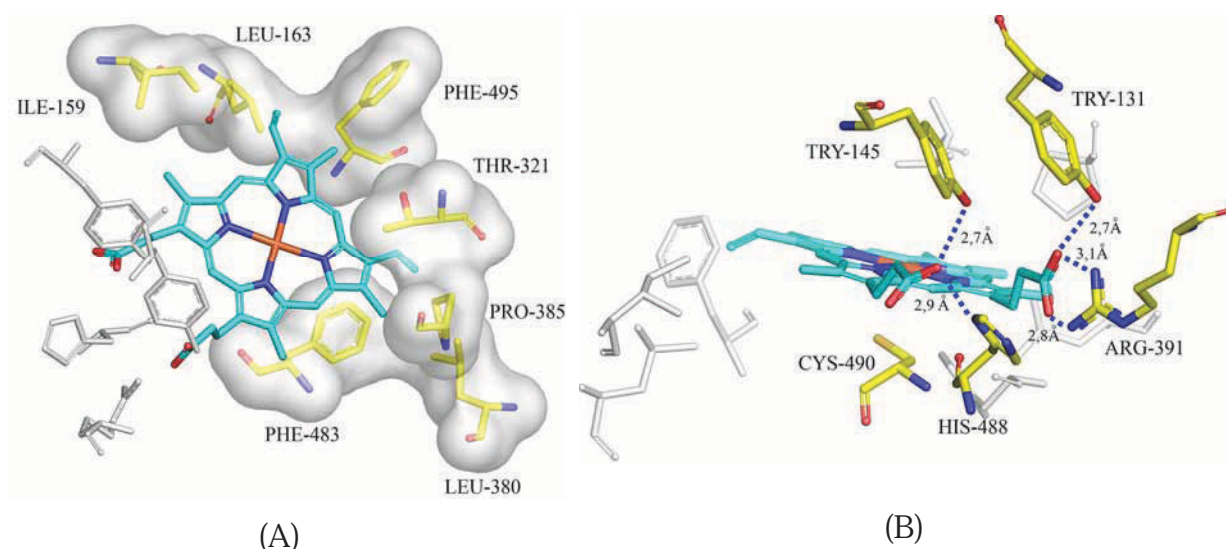


Figura 3.14 – Duas orientações do sítio ativo de lanosterol 14 α -desmetilase de *C. gattii* onde se destacam (amarelo) resíduos apolares (A) e polares (B), que interagem com o grupo heme (verde)

A seguir, iniciou-se a análise dos modelos das proteínas mutantes, mais especificamente pela avaliação do modelo contendo a mutação pontual G484S. A análise do gráfico de Ramachandran revelou que 84,1% dos resíduos estão em regiões favoráveis, 12,3% foram encontrados em regiões adicionalmente permitidas, 2,3% dos resíduos estão localizados na zona generosamente permitida, enquanto os resíduos ALA-196, PHE-511, ASP-199, GLN-124 e VAL-144 (1,3%), localizados em alças na proteína, são encontrados em regiões não permitidas (**FIGURA 3.15 A e B**).

A análise da compatibilidade entre a sequência de aminoácidos e o enovelamento observado mostra que nenhum resíduo possui valor abaixo de zero na função pontuação 3D-1D (FIGURA 3.15C), reforçando que o modelo comparativo desse mutante possui é adequado para análises estruturais.

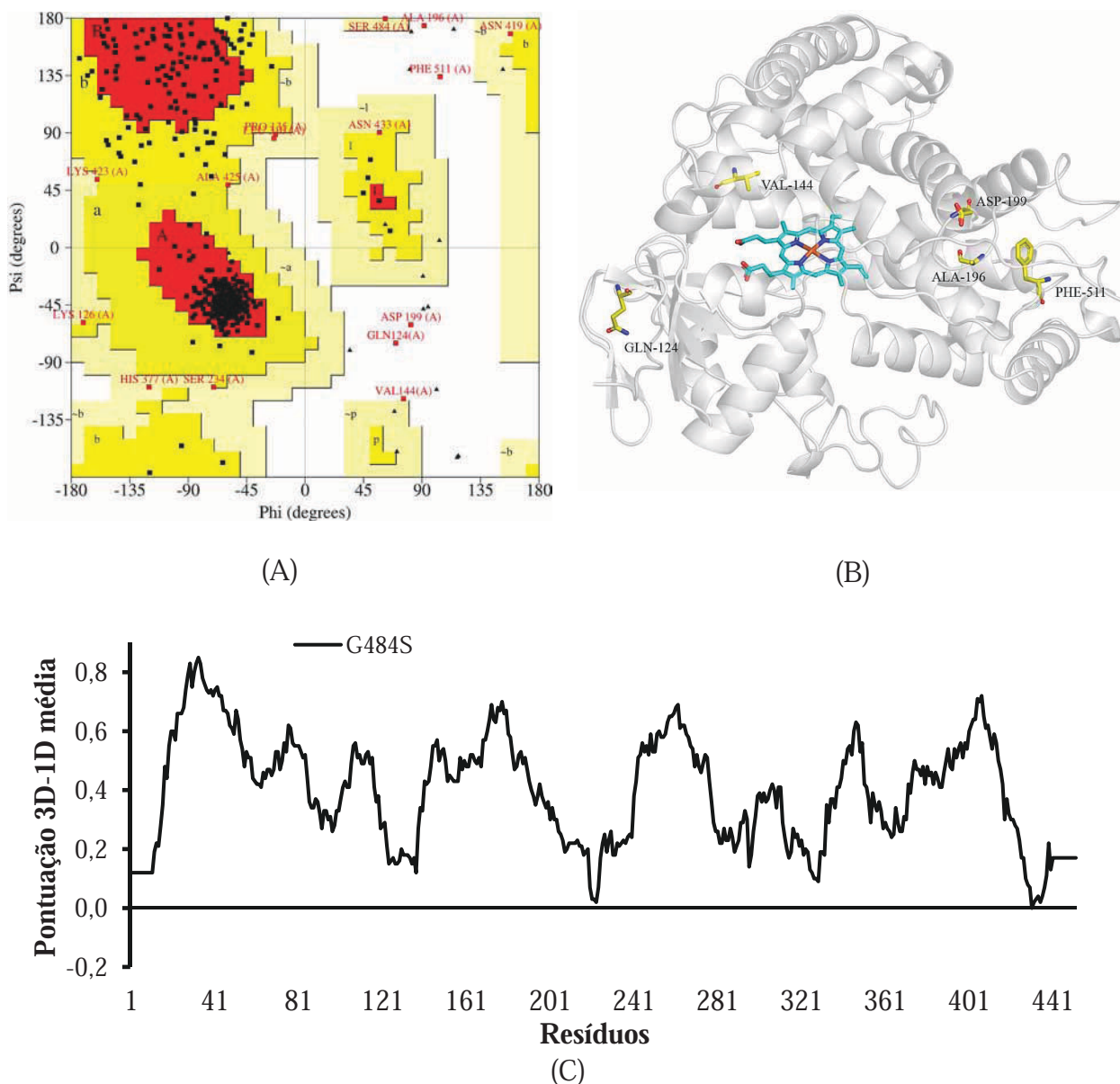


Figura 3.15 – Gráfico de Ramachandran do modelo comparativo da enzima contendo a mutação G484S(A), resíduos destacados na região não permitida do gráfico de Ramachandran (B) e gráfico de pontuação 3D-1D média do VERIFY 3D (C)

A mutação G484S, além de reduzir a flexibilidade da proteína, requerida para melhor acomodação de substratos e ligantes,²⁰⁸ deve modificar o posicionamento do grupo heme, uma vez que a serina, por apresentar cadeia lateral, se choca com a cadeia lateral do ácido propanóico presente no grupo heme (FIGURA 3.16).

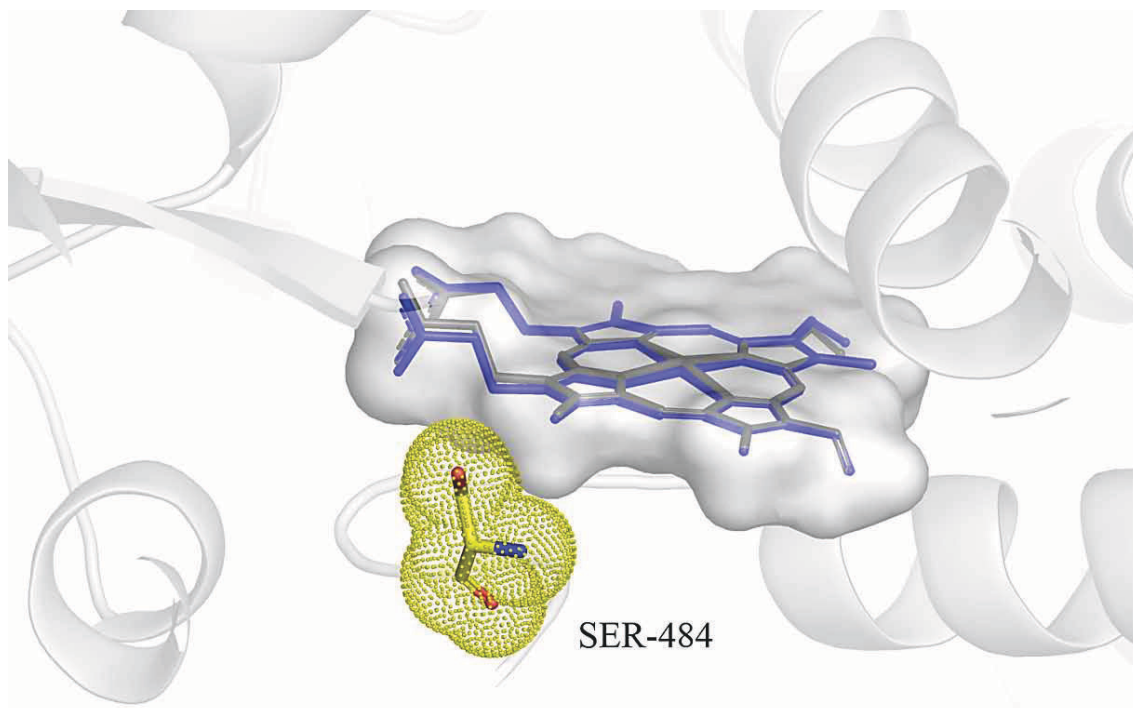


FIGURA 3.16. – Choque estérico entre o ácido propanóico do grupo heme da enzima nativa (azul) e a serina (SER-484). O grupo heme da enzima mutante está na coloração cinza

Haja vista que os derivados azólicos interagem de forma perpendicular ao ferro do grupo heme, desvios angulares no grupamento prostético, como sugeridas no modelo comparativo (FIGURA 3.17), podem proporcionar no fármaco choques estéricos secundários com resíduos do sítio ativo da proteína, dificultando a acomodação dos derivados azólicos no sítio de interação.

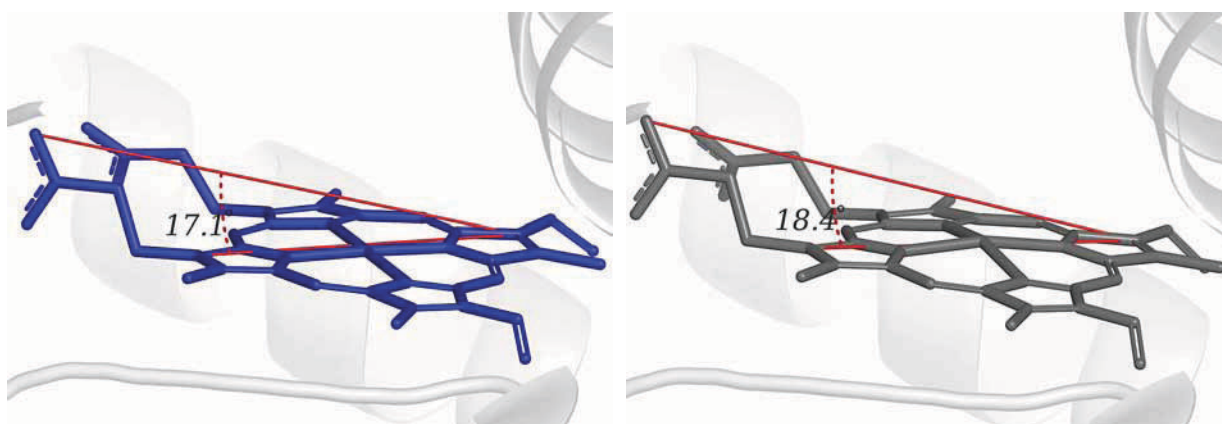


FIGURA 3.17 – Ângulo entre os átomos do grupo prostético da enzima nativa (azul) e da enzima contendo a mutação G484S (em cinza)

De fato, os efeitos desse tipo de mutação podem ser menores em fármacos com maior cadeia, tais como voriconazol e itraconazol, os quais são menos associados a eventos de resistência.^{49; 67} Estudos complementares sugerem que estes fármacos são capazes de fazer interações adicionais na enzima estabilizando a ligação entre o nitrogênio do triazol e o ferro do grupo heme reduzindo, assim, a influência de variações no posicionamento do grupo prostético em decorrência de mutações na enzima lanosterol 14 α -desmetilase.^{137; 220; 221}

Seguindo essa mesma estratégia, num segundo momento, o modelo comparativo que apresenta a mutação H488Y também foi investigado. A análise do gráfico de Ramachandran mostra que 84,1% dos resíduos estão em regiões mais favoráveis, 13,1% são encontrados em regiões adicionalmente permitidas e 1,8% em regiões generosamente permitidas (**FIGURA 3.18A**), ao passo que quatro resíduos (1%), todos localizadas em regiões de alça, foram encontrados em regiões não permitidas: PHE-511, LEU-536, GLN-124 e VAL-144 (**FIGURA 3.18B – resíduos em amarelo**). Apesar da alta qualidade estérica observada, o modelo mostrou uma região com baixa compatibilidade entre a sequência primária e a estrutura 3D adotada (**FIGURA 3.18B**). Entretanto, esta região pertence a uma alça localizada na superfície da proteína e distante do sítio de ligação, não interferindo, portanto, no estudo sobre o papel da mutação H488Y na resistência a derivados azólicos.

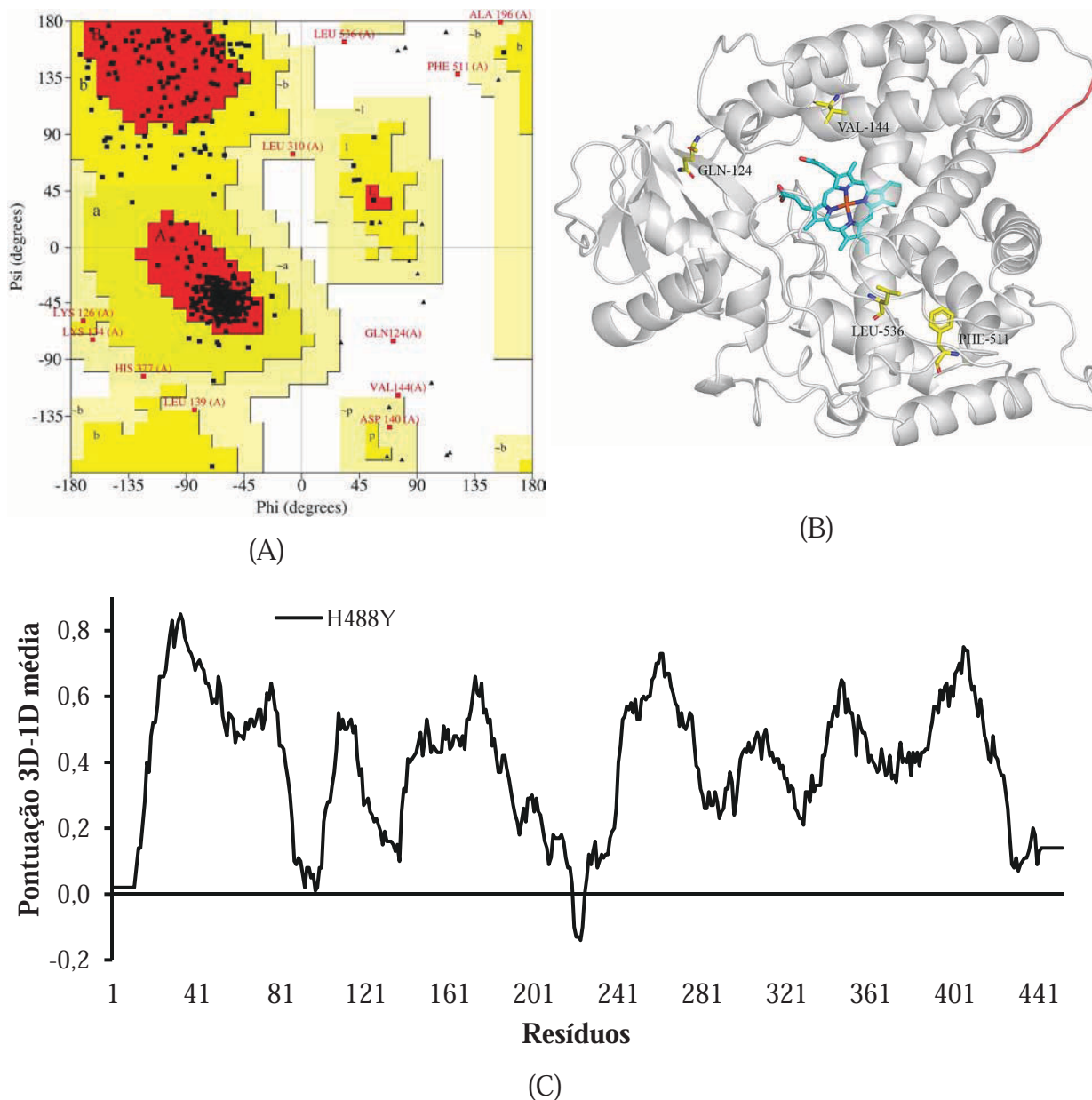


FIGURA 3.18 – Gráfico de Ramachandran (A) e modelo 3D da enzima com a mutação H488Y destacando os resíduos encontrados em regiões não permitidas (resíduos em amarelo) (B) e a região (em vermelho) de baixa pontuação 3D-1D do VERIFY 3D (C)

A mutação H488Y parece modificar o padrão de interações entre a proteína e o ácido propanóico do grupo heme, pois a interação iônica observada entre o resíduo histidina (protonado) e o carboxilato do ácido propanóico na estrutura nativa é perdida no modelo mutante, permitindo o deslocamento do grupo prostético (**FIGURA 3.19**) de forma semelhante ao observado no modelo da mutação G484S.

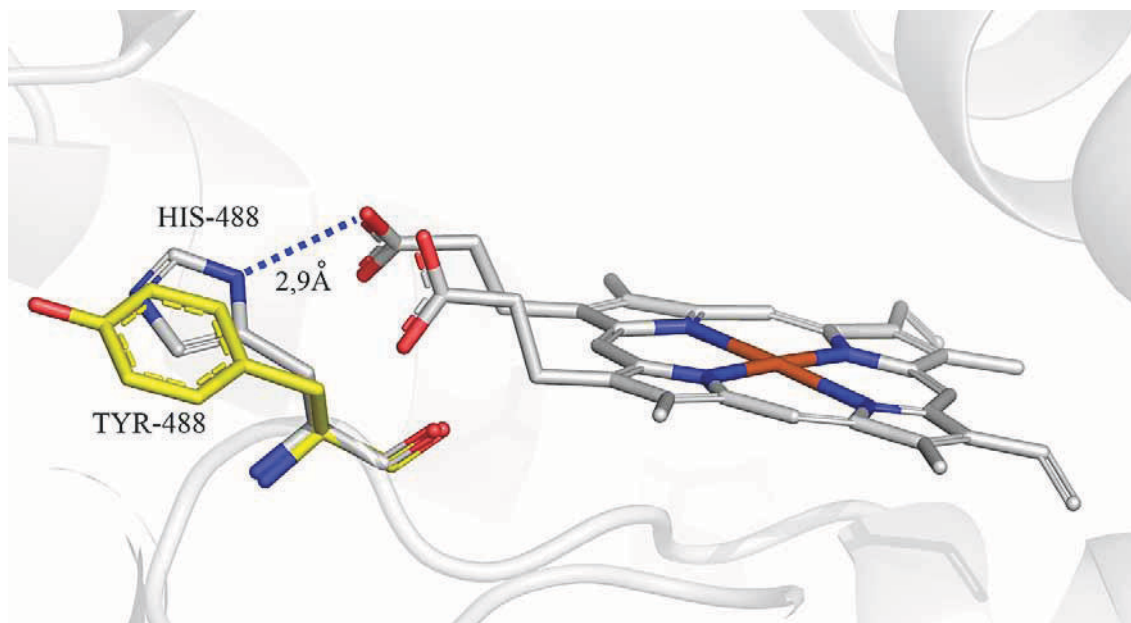


FIGURA 3.19 – Substituição da histidina (HIS-488) pela tirosina (TYR-488) causa perda da interação iônica com o ácido propanóico do grupo heme

As mutações investigadas até momento se localizam próximas ao grupo heme, alterando sua orientação. Por outro lado o modelo comparativo da enzima com a substituição de uma lisina por uma arginina na posição 156 (K156R) revela que essa mutação está distante do grupo prostético, não interagindo diretamente com ele. A qualidade estereoquímica desse modelo, bem como a adequação do enovelamento é semelhante ao observado nos demais modelos uma vez que 85,6% dos resíduos estão em regiões mais favoráveis, 12,6% foram encontrados em regiões adicionalmente permitidas e 1% em regiões generosamente permitidas. (FIGURA 3.20A), enquanto três resíduos (0,8%) foram encontrados em regiões não permitidas: ALA-196, PHE-511 e SER-303. A análise da compatibilidade entre a sequência e o seu enovelamento revelou que apenas três resíduos, já destacados em modelos anteriores, apresentam valor de pontuação 3D-1D abaixo de zero (FIGURA 3.20B).

Visando investigar os mecanismos indiretos que poderiam levar a resistência à derivados azólicos, simulou-se as mudanças conformacionais que a enzima lanosterol 14 α -desmetilase nativa e mutante (K156R) poderiam sofrer para permitir a entrada e acomodação do fármaco no sítio de ligação, uma vez que Lepesheva e colaboradores (2010) sugeriram que mutações que alteram o canal de entrada do fármaco, bem como sua superfície de interação podem resultar em redução da atividade do antifúngico.¹⁷⁴

A fim de realizar essa tarefa, utilizou-se o servidor HINGEPROT, que emprega um modelo elástico para calcular a flexibilidade da proteína em termos de corpos rígidos, unidos entre si por regiões flexíveis (dobradiça), a fim de determinar as principais mudanças

conformacionais da enzima lanosterol 14 α -desmetilase de *C. gattii* mutante em relação a enzima nativa.

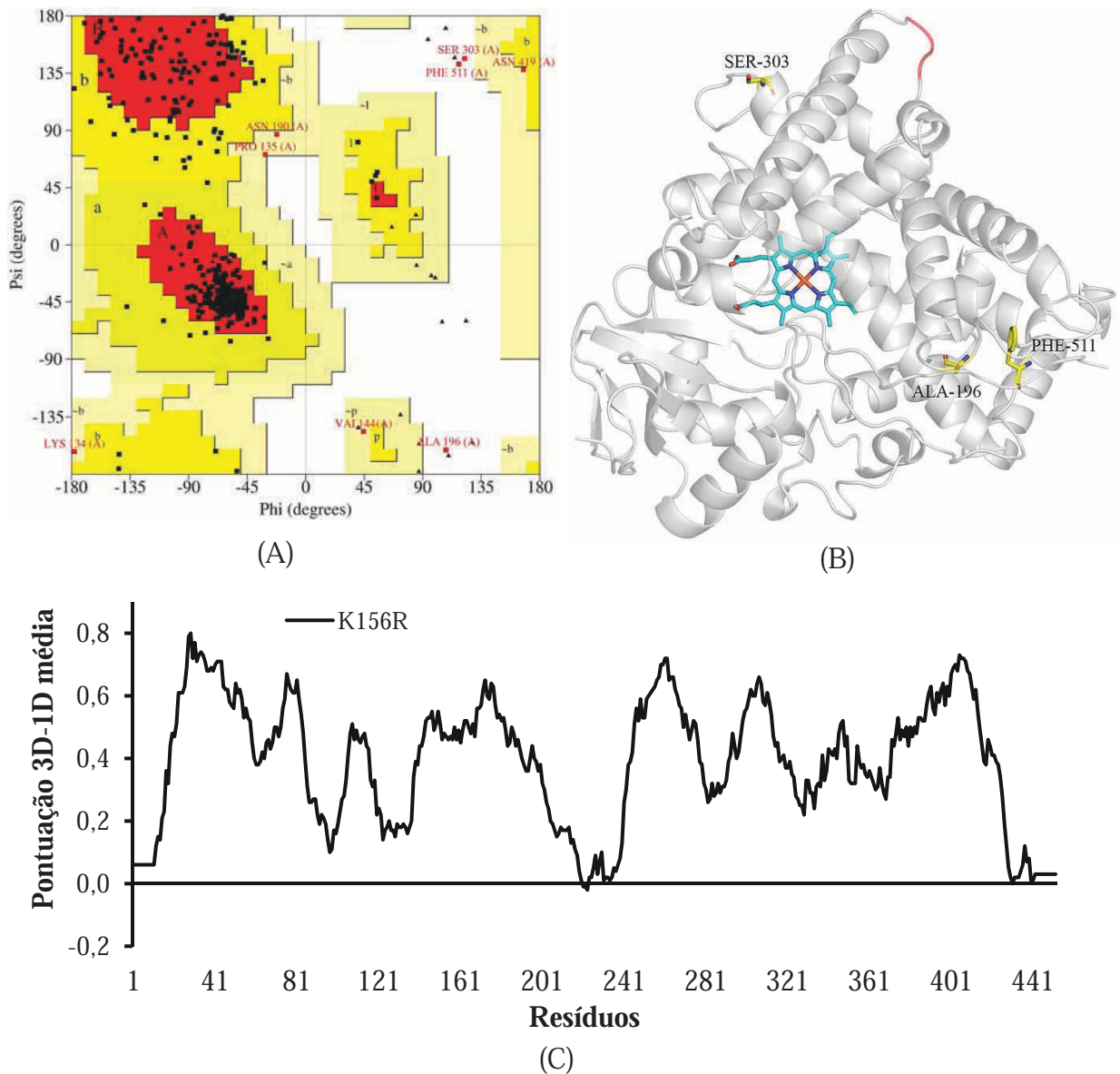


FIGURA 3.20 – Gráfico de Ramachandran (A) e estrutura 3D do modelo da enzima contendo a mutação K156R destacando a região de baixa pontuação 3D-1D do programa VERIFY 3D (alça em vermelho) (C) e os resíduos encontrados em regiões não permitidas (resíduos em amarelo) (B)

Segundo o servidor HINGEPROT, enzima nativa possui 4 domínios rígidos, unidos entre si por 3 regiões de dobradiça localizados entre os resíduos THR-135 e THR-136; THR-379 e LEU-380; e CIS-490 e VAL-491. Resultados aparentemente similares foram encontrados para a enzima mutante, com uma alteração importante: uma das dobradiças foi deslocada dos resíduos THR-135 e THR-136 para GLN-154 e GLN-155. Como consequência disso, a extensão de um dos domínios rígidos passou de 135 para 154 resíduos (**FIGURA 3.21**).

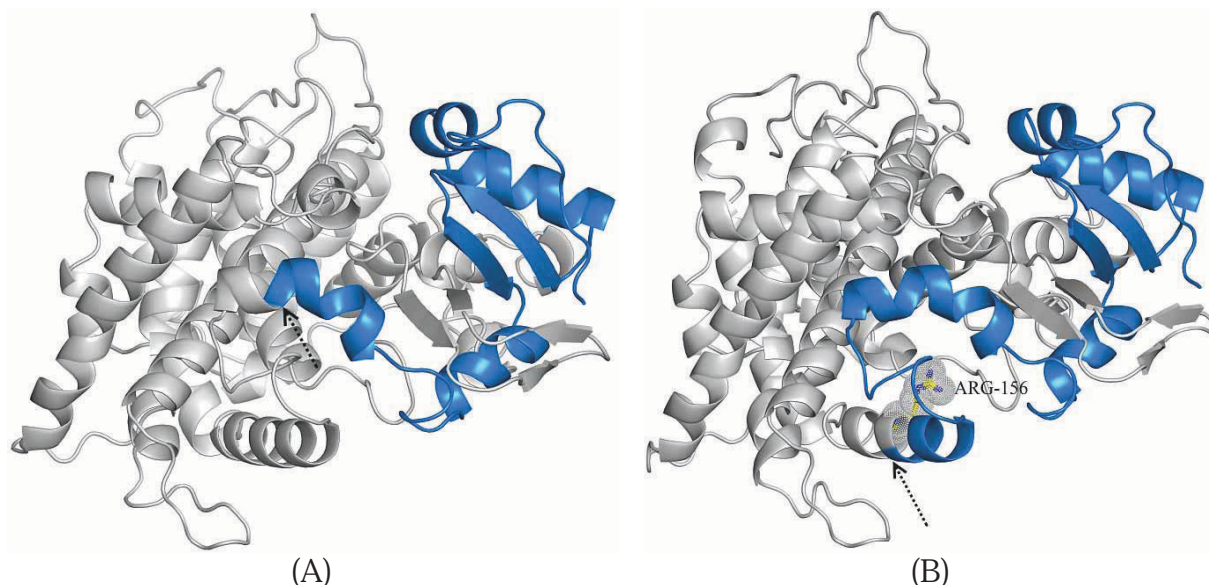


FIGURA 3.21 – Região de dobradiça (apontado pela seta) do modelo comparativo da enzima lanosterol 14 α -desmetilase nativa (A) que muda de posição no modelo da enzima com a mutação K156R (B)

Este tipo de mutação pode interferir no acesso ao sítio de interação dos derivados de azol, uma vez a entrada do fármaco é controlada, mesmo que parcialmente, por mudanças conformacionais em regiões específicas da proteína^{173; 208} (**FIGURA 3.22 – em vermelho**).

O deslocamento da região de dobradiça da proteína nativa (apontado pela seta na **FIGURA 3.21A**) para uma hélice alfa distante da região de entrada do fármaco (apontado pela seta na **FIGURA 3.21B**) sugere que a região inicialmente mais flexível, e possivelmente importante para a entrada do fármaco na enzima nativa, torne-se um domínio rígido, não apresentando a amplitude do deslocamento requerida para a entrada do fármaco no sítio de ligação.

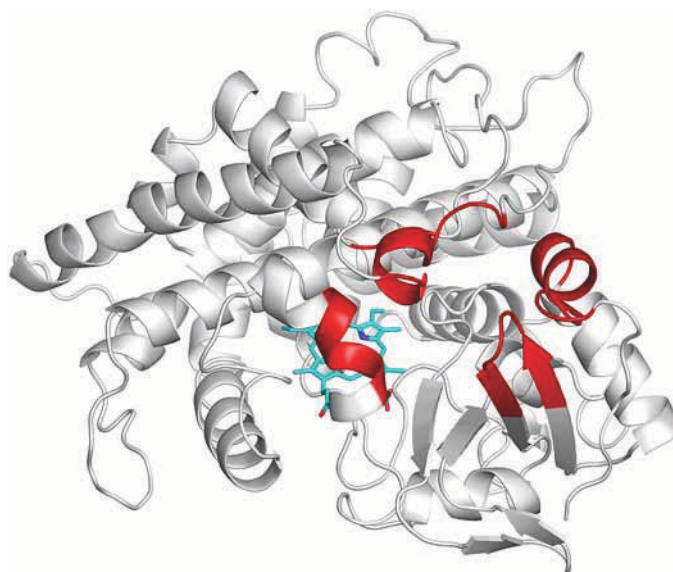


FIGURA 3.22 – Modelo da proteína lanosterol 14 α -desmetilase de *C. gattii* destacando os resíduos que compõem a região de acesso do fármaco na proteína

CAPÍTULO 4

4 CONSIDERAÇÕES FINAIS E PERSPECTIVAS

Ensaio celulares desempenham um papel fundamental na pesquisa e desenvolvimento de novos fármacos uma vez que avaliam uma variedade de efeitos fenotípicos, como a mortalidade de um microorganismo quando exposto a um agente modulador (antifúngico). Apesar disso, existem poucos estudos sobre o perfil de atividade dos derivados azólicos, estruturalmente diversos, frente *Cryptococcus*. Diante dessa lacuna, a susceptibilidade de cepas de *Cryptococcus* resistentes e sensíveis à fluconazol foi determinada para 33 derivados azólicos. Os resultados obtidos foram utilizados para classificar os antifúngicos em duas classes (ativos e inativos), com base nos valores de diâmetro de halo de inibição.

Considerando que o foco principal desse trabalho é a investigação das propriedades físico-químicas e estruturais que determinam a atividade de antifúngicos azólicos frente a cepas resistentes de *Cryptococcus*, maior atenção foi dada aos resultados biológicos obtidos para *C. gattii*. Métodos quimiométricos, como KNN e SIMCA, foram então empregados para investigar as propriedades químicas responsáveis pela atividade antifúngica observada e assim permitir a triagem *in silico* de derivados azólicos ainda não testados contra *C. gattii*.

O método KNN mostrou consistência interna razoável ao acertar a classificação de 60% dos compostos ativos e 75% dos compostos inativos do grupo teste (validação externa), considerando 10 vizinhos. O índice de acerto moderado do método KNN pode estar associado à alta colinearidade dos descritores utilizados. O melhor modelo SIMCA apresentou capacidade preditiva equivalente ao do método KNN, porém permitiu determinar quais características físico-químicas são importantes para separar os compostos ativos dos inativos: (JGI10) transferência de carga entre átomos distantes de 10 ligações. Estes resultados corroboram com dados previamente por Mota e colaboradores¹⁴⁰ para a atividade de antifúngicos frente ao fungo filamentoso *M. perniciosus*.

Embora promissores, os modelos classificatórios não permitem prever a potência dos antifúngicos, tendo, portanto utilidade restrita. A fim de contornar essa limitação, foi determinada a concentração inibitória mínima para os 33 derivados azólicos frente *Cryptococcus*. Os resultados mostram que há grande variação de potência entre os antifúngicos testados, permitindo que técnicas quantitativas e estudos das relações entre a

estrutura química e a atividade biológica sejam empregados. Modelos quantitativos baseados em fragmentos moleculares (hologramas QSAR) apresentaram bom ajuste ($r^2=0,85$), no entanto, baixa consistência interna ($q^2=0,35$) e poder preditivo ($r^2_{pred}=0,38$), com 4 PCs, sugerindo que fragmentos molecular não são capazes de expressar a atividade antifúngica senão para os antifúngicos do grupo treino. O melhor modelo de QSAR baseado em descritores topológicos apresentou ajuste excelente ($r^2=0,95$) e ótima consistência interna ($q^2=0,86$) com 3 componentes principais. Adicionalmente, o bom poder preditivo desse modelo ($r^2_{pred}=0,72$) reforça sua utilidade no planejamento de antifúngicos mais potentes contra *C. gattii*. Alcançar esse objetivo requer um entendimento profundo acerca das exigências estéreas e eletrostáticas relevantes para o aumento da potência destes antifúngicos. De acordo com o modelo de QSAR, o principal fator para o aumento da potência é a distribuição de carga de uma molécula na distância de dois átomos (GGI1). Esta informação sugere que a menor distribuição de elétrons na molécula possa ser importante para a atividade, pois compostos com maior potência tem maior densidade eletrônica no nitrogênio do imidazol (e do triazol) ligado ao ferro do grupo heme. Por outro lado, quanto maior a diferença de eletronegatividade dos átomos na distância topológica 1 e 2 (GATS1e e MATS2e), menor a potência dos antifúngicos.

Contudo, o sucesso terapêutico de um antifúngico não depende unicamente de sua potência, mas também de seu perfil de atividade frente a cepas resistentes a derivados azólicos. De fato, os resultados obtidos nos ensaios de atividade por microdiluição em caldo indicam que a cepa de *C. gattii* utilizada nesse trabalho é resistente a fluconazol. Estudos posteriores, utilizando inibidores conhecidos de bomba de efluxo (ibuprofeno e verapamil), sugerem que esse mecanismo de resistência não está presente no isolado clínico de *C. gattii* estudado. A partir dessa evidência, investigou-se a presença de mutações no gene ERG11 que pudessem estar associadas a menor suscetibilidade de *C. gattii* frente à fluconazol. No entanto, devido a problemas técnicos não foi possível ainda concluir essa etapa do trabalho.

Contudo, considerando a importância de se investigar, do ponto de vista estrutural, o efeito de mutações no gene ERG11 sobre a potência de derivados azólicos, decidiu-se utilizar mutações associadas a resistência, já descritas na literatura, para desenvolver modelos por homologia de lanosterol 14 α -desmetilase de *C. gattii* contendo mutações pontuais (G484S, H488Y e K156R).

Através da comparação dos modelos de enzima mutante com o modelo de enzima nativa pode-se sugerir que: i) A mutação G484S altera o posicionamento do grupo heme, promovendo um desvio angular no grupo prostético que desestabiliza a interação com o ferro

do grupo heme; ii) A mutação H488Y reduz o número de interações entre a enzima e o grupo prostético, favorecendo alterações no posicionamento do grupo heme, os quais podem afetar a acomodação dos ligantes no sítio ativo, tal como previsto para a mutação G484S; iii) Mutações que não interferem na orientação do grupo heme, como K156R, podem alterar a flexibilidade conformacional do alvo terapêutico, reduzindo assim o acesso do antifúngico ao sítio ativo da enzima.

Embora as ferramentas computacionais utilizadas tenham fornecido informações relevantes para a compreensão das propriedades físico-químicas e estruturais que modulam a atividade de derivados azólicos frente a *C. gattii*, bem como sobre os possíveis mecanismos de resistência, conferidos por mutações no gene ERG11, o desenvolvimento de antifúngicos mais potentes e ativos contra isolados clínicos de *C. gattii* resistentes a fluconazol depende de uma série etapas que não puderam ainda ser realizadas como: síntese de análogos imidazólicos e triazólicos, com base nas informações oriundas dos modelos de QSAR, avaliação de parâmetros farmacocinéticos dos derivados imidazólicos e triazólicos mais potentes que fluconazol, validação experimental da influência das mutações G484S, H488Y e K156R sobre a suscetibilidade de *C. gattii* frente a derivados azólicos, etc. Dessa forma, o trabalho apresentado constitui apenas a primeira das muitas etapas necessárias para ampliação do arsenal terapêutico disponível para o tratamento de criptococose.

Contudo, deve-se destacar que os resultados obtidos neste trabalho poderão contribuir para identificação de novos derivados azólicos ativos contra *C. gattii*, através de triagem virtual bem como para guiar o planejamento e síntese de inibidores de lanosterol 14 α -desmetilase que interajam de forma efetiva com o ferro do grupo heme, mesmo quando este estiver deslocado da posição encontrada na enzima nativa, ou seja, ativos contra cepas de *C. gattii* resistentes a fluconazol.

REFERÊNCIAS

- 1 GUIDO, R. V.; OLIVA, G.; ANDRICOPULO, A. D. Virtual screening and its integration with modern drug design technologies. **Current Medicinal Chemistry**, v. 15, n. 1, p. 37-46, 2008.
- 2 KAR, A. **Medicinal Chemistry**. 4. New Age International (P) Ltd., 2005. 960 ISBN 9788122415650.
- 3 LOMBARDINO, J. G.; LOWE, J. A. I. The role of the medicinal chemist in drug discovery - then and now. **Nature Reviews Drug Discovery**, v. 3, n. 10, p. 853-862, Oct 2004.
- 4 CUATRECASAS, P. Drug discovery in jeopardy. **The Journal of Clinical Investigation**, v. 116, n. 11, p. 2837-2842, Nov 2006.
- 5 CHONG, C. R.; SULLIVAN, D. J. J. New uses for old drugs. **Nature**, v. 448, n. 7154, p. 645-646, Aug 9 2007.
- 6 VAN DE WATERBEEMD, H. The history of drug research: From Hansch to the present. **Quantitative Structure-Activity Relationships**, v. 11, n. 2, p. 200-204, 1992.
- 7 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.
- 8 MASON, J. S. Introduction to the Volume and Overview of Computer-Assisted Drug Design in the Drug Discovery Process. In: TAYLOR, J. B. e TRIGGLE, D. J. (Ed.). **Comprehensive Medicinal Chemistry II**. Oxford: Elsevier Science, v.4, 2006. cap. 4.01, p.1-11. ISBN 9780080445175.
- 9 WERMUTH, G.; GANELLIN, C. R.; LINDBERG, P.; MITSCHER, L. A. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1998). **Pure and Applied Chemistry**, v. 70, n. 5, p. 1129-1143, May 1998.
- 10 JORGENSEN, W. L. The many roles of computation in drug discovery. **Science**, v. 303, n. 5665, p. 1813-1818, Mar 19 2004.
- 11 XU, J.; HAGLER, A. Chemoinformatics and Drug Discovery. **Molecules**, v. 7, n. 8, p. 566-600, 2002.

- 12 GAIKWAD, V. J. Application of chemoinformatics for innovative Drug discovery. **International Journal of Chemical Sciences and Applications**, v. 1, n. 1, p. 16-25, 2010.
- 13 PRAKASH, N. D., A. G. . Cheminformatics. **Journal of Proteomics & Bioinformatics**, v. 3, n. 8, p. 249-252, 2010.
- 14 GASTEIGER, J. The central role of chemoinformatics. **Chemometrics and Intelligent Laboratory Systems**, v. 82, n. 1-2, p. 200-209, May 26 2006.
- 15 ANDERSON, A. C. The Process of Structure-Based Drug Design. **Chemistry & Biology**, v. 10, n. 9, p. 787-797, 2003.
- 16 GAO, N.; ZAVIALOV, A. V.; EHRENBERG, M.; FRANK, J. Specific interaction between EF-G and RRF and its implication for GTP-dependent ribosome splitting into subunits. **Journal of Molecular Biology**, v. 374, n. 5, p. 1345-1358, Dec 14 2007.
- 17 LI, W. H.; TANG, Y.; LIU, H.; CHENG, J. G.; ZHU, W. L.; JIANG, H. L. Probing ligand binding modes of human cytochrome P450 2J2 by homology modeling, molecular dynamics simulation, and flexible molecular docking. **Proteins-Structure Function and Bioinformatics**, v. 71, n. 2, p. 938-949, May 1 2008.
- 18 GLOBISCH, C.; PAJEVA, I. K.; WIESE, M. Identification of putative binding sites of P-glycoprotein based on its homology model. **Chemmedchem**, v. 3, n. 2, p. 280-295, Feb 2008.
- 19 LENGAUER, T.; LEMMEN, C.; RAREY, M.; ZIMMERMANN, M. Novel technologies for virtual screening. **Drug Discovery Today**, v. 9, n. 1, p. 27-34, Jan 1 2004.
- 20 TROPSHA, A. Predictive QSAR (Quantitative Structure Activity Relationships) Modeling. In: TAYLOR, J. B. e TRIGGLE, D. J. (Ed.). **Comprehensive medicinal Chemistry II**. Oxford: Elsevier Science, v.4, 2006. cap. 4.07, p.149-165.
- 21 CONSONNI, V.; TODESCHINI, R.; PAVAN, M. Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 1. Theory of the novel 3D molecular descriptors. **Journal of Chemical Information and Computer Sciences**, v. 42, n. 3, p. 682-692, May-Jun 2002.
- 22 WILLETT, P.; BARNARD, J. M.; DOWNS, G. M. Chemical Similarity Searching. **Journal of Chemical Information and Computer Sciences**, v. 38, n. 6, p. 983-996, 1998.

- 23 LIVINGSTONE, D. J. The characterization of chemical structures using molecular properties. A survey. **Journal of Chemical Information and Computer Sciences**, v. 40, n. 2, p. 195-209, Mar-Apr 2000.
- 24 PETERS, M. B.; RAHA, K.; MERZ, K. M., JR. Quantum mechanics in structure-based drug design. **Current opinion in drug discovery and development**, v. 9, n. 3, p. 370-379, May 2006.
- 25 RAHA, K.; PETERS, M. B.; WANG, B.; YU, N.; WOLLACOTT, A. M.; WESTERHOFF, L. M.; MERZ, K. M., JR. The role of quantum mechanics in structure-based drug design. **Drug Discovery Today**, v. 12, n. 17-18, p. 725-731, Sep 2007.
- 26 DE VIVO, M. Bridging quantum mechanics and structure-based drug design. **Frontiers in Bioscience**, v. 16, p. 1619-1633, 2011.
- 27 KIER, L. B.; HALL, L. H.; MURRAY, W. J.; RANDIC, M. Molecular connectivity. I: Relationship to nonspecific local anesthesia. **Journal of pharmaceutical sciences**, v. 64, n. 12, p. 1971-1974, Dec 1975.
- 28 KIER, L. B.; MURRAY, W. J.; RANDIC, M.; HALL, L. H. Molecular connectivity V: connectivity series concept applied to density. **Journal of pharmaceutical sciences**, v. 65, n. 8, p. 1226-1230, Aug 1976.
- 29 JAIN, A. N.; KOILE, K.; CHAPMAN, D. Compass: predicting biological activities from molecular surface properties. Performance comparisons on a steroid benchmark. **Journal of medicinal chemistry**, v. 37, n. 15, p. 2315-2327, Jul 22 1994.
- 30 MARSHALL, G. R.; CRAMER, R. D., 3RD. Three-dimensional structure-activity relationships. **Trends in pharmacological sciences**, v. 9, n. 8, p. 285-289, Aug 1988.
- 31 RIDINGS, J. E.; BARRATT, M. D.; CARY, R.; EARNSHAW, C. G.; EGGINGTON, C. E.; ELLIS, M. K.; JUDSON, P. N.; LANGOWSKI, J. J.; MARCHANT, C. A.; PAYNE, M. P.; WATSON, W. P.; YIH, T. D. Computer prediction of possible toxic action from chemical structure: an update on the DEREK system. **Toxicology**, v. 106, n. 1-3, p. 267-279, Jan 8 1996.
- 32 SANDERSON, D. M.; EARNSHAW, C. G. Computer prediction of possible toxic action from chemical structure; the DEREK system. **Human & Experimental Toxicology**, v. 10, n. 4, p. 261-273, Jul 1991.

- 33 GAO, H.; LAJINESS, M. S.; VAN DRIE, J. Enhancement of binary QSAR analysis by a GA-based variable selection method. **Journal of molecular graphics & modelling**, v. 20, n. 4, p. 259-268, Jan 2002.
- 34 XU, L.; ZHANG, W. J. Comparison of different methods for variable selection. **Analytica Chimica Acta**, v. 446, n. 1-2, p. 477-483, Nov 19 2001.
- 35 KUBINYI, H. **QSAR : Hansch analysis and related approaches**. VCH, 1993. ISBN 9783527300358.
- 36 BARONI, M.; COSTANTINO, G.; CRUCIANI, G.; RIGANELLI, D.; VALIGI, R.; CLEMENTI, S. Generating Optimal Linear PLS Estimations (Golpe) - an Advanced Chemometric Tool for Handling 3D-Qsar Problems. **Quantitative Structure-Activity Relationships**, v. 12, n. 1, p. 9-20, Mar 1993.
- 37 KUBINYI, H. The key to the castle. II. Hansch analysis, 3D-QSAR and de novo design. **Pharmazie in unserer Zeit**, v. 23, n. 5, p. 281-290, Sep 1994.
- 38 GONZALEZ, M. P.; TERAN, C.; SAIZ-URRA, L.; TEIJEIRA, M. Variable selection methods in QSAR: an overview. **Current topics in medicinal chemistry**, v. 8, n. 18, p. 1606-1627, 2008.
- 39 STEYERBERG, E. W.; EIJKEMANS, M. J.; HABBEMA, J. D. Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. **Journal of clinical epidemiology**, v. 52, n. 10, p. 935-942, Oct 1999.
- 40 AGOSTINELLI, C. Robust stepwise regression. **Journal of Applied Statistics**, v. 29, n. 6, p. 825-840, Aug 2002.
- 41 WATERBEEMD, H. **Chemometric methods in molecular design**. VCH, 1995. ISBN 9783527300440.
- 42 HANSCH, C.; LEO, A.; HOEKMAN, D. **Exploring QSAR: Fundamentals and applications in chemistry and biology**. American Chemical Society, 1995. 580 ISBN 9780841229914.
- 43 LINUSSON, A.; GOTTFRIES, J.; LINDGREN, F.; WOLD, S. Statistical molecular design of building blocks for combinatorial chemistry. **Journal of Medicinal Chemistry**, v. 43, n. 7, p. 1320-1328, Apr 6 2000.

- 44 PERKINS, R.; FANG, H.; TONG, W.; WELSH, W. J. Quantitative structure-activity relationship methods: perspectives on drug discovery and toxicology. **Environmental Toxicology and Chemistry**, v. 22, n. 8, p. 1666-1679, Aug 2003.
- 45 ANDRADE, C. H.; SALUM LDE, B.; CASTILHO, M. S.; PASQUALOTO, K. F.; FERREIRA, E. I.; ANDRICOPULO, A. D. 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, Feb 2008.
- 46 VALADARES, N. F.; CASTILHO, M. S.; POLIKARPOV, I.; GARRATT, R. C. 2D QSAR studies on thyroid hormone receptor ligands. **Bioorganic & Medicinal Chemistry**, v. 15, n. 13, p. 4609-4617, Jul 1 2007.
- 47 CASTILHO, M. S.; C. GUIDO, R. V.; ANDRICOPULO, A. D. Classical and Hologram QSAR Studies on a Series of Tacrine Derivatives as Butyrylcholinesterase Inhibitors. **Letters in Drug Design & Discovery**, v. 4, p. 106-113, 2007.
- 48 PETRIKKOS, G.; SKIADA, A. Recent advances in antifungal chemotherapy. **International journal of antimicrobial agents**, v. 30, n. 2, p. 108-117, Aug 2007.
- 49 RUGE, E.; KORTING, H. C.; BORELLI, C. Current state of three-dimensional characterisation of antifungal targets and its use for molecular modelling in drug design. **International journal of antimicrobial agents**, v. 26, n. 6, p. 427-441, Dec 2005.
- 50 SUNDRIYAL, S.; SHARMA, R. K.; JAIN, R. Current advances in antifungal targets and drug development. **Current Medicinal Chemistry**, v. 13, n. 11, p. 1321-1335, 2006.
- 51 GRIMALDI, M.; DE ROSA, M.; DI MARINO, S.; SCRIMA, M.; POSTERARO, B.; SANGUINETTI, M.; FADDA, G.; SORIENTE, A.; D'URSI, A. M. Synthesis of new antifungal peptides selective against *Cryptococcus neoformans*. **Bioorganic & Medicinal Chemistry**, v. 18, n. 22, p. 7985-7990, 2010.
- 52 TANG, H.; ZHENG, C.; LV, J.; WU, J.; LI, Y.; YANG, H.; FU, B.; LI, C.; ZHOU, Y.; ZHU, J. Synthesis and antifungal activities in vitro of novel pyrazino [2,1-a] isoquinolin derivatives. **Bioorganic & Medicinal Chemistry Letters**, v. 20, n. 3, p. 979-982, Feb 1 2010.
- 53 BASTERT, J.; SCHALLER, M.; KORTING, H. C.; EVANS, E. G. V. Current and future approaches to antimycotic treatment in the era of resistant fungi and immunocompromised hosts. **International journal of antimicrobial agents**, v. 17, n. 2, p. 81-91, Feb 2001.

- 54 ODDS, F. C. Antifungal agents: their diversity and increasing sophistication. **Mycologist**, v. 17, n. 2, p. 51-55, 2003.
- 55 THOMPSON, G. R. I.; CADENA, J.; PATTERSON, T. F. Overview of antifungal agents. **Clinics in chest medicine**, v. 30, n. 2, p. 203-215, Jun 2009.
- 56 BELL, A. S. Major Antifungal Drugs. In: TAYLOR, J. B. e TRIGGLE, D. J. (Ed.). **Comprehensive Medicinal Chemistry II**. Oxford: Elsevier Science, v.7 2007. cap. 7.15, p.445-468. ISBN 978-0-08-045044-5.
- 57 SCORZONI, L.; BENADUCCI, T.; ALMEIDA, A. M. F.; SILVA, D. H. S.; BOLZANI, V. S.; MENDES-GIANNINI, M. J. S. Comparative study of disk diffusion and microdilution methods for evaluation of antifungal activity of natural compounds against medical yeasts *Candida* spp and *Cryptococcus* sp. **Journal of Basic and Applied Pharmaceutical Science**, v. 28, n. 1, p. 25-34, 2007.
- 58 OSTROSKY-ZEICHNER, L.; CASADEVALL, A.; GALGIANI, J. N.; ODDS, F. C.; REX, J. H. An insight into the antifungal pipeline: selected new molecules and beyond. **Nature Reviews Drug Discovery**, v. 9, n. 9, p. 719-727, Sep 2010.
- 59 PERFECT, J. R.; COX, G. M. Drug resistance in *Cryptococcus neoformans*. **Drug resistance updates**, v. 2, n. 4, p. 259-269, Aug 1999.
- 60 RODERO, L.; MELLADO, E.; RODRIGUEZ, A. C.; SALVE, A.; GUELFAND, L.; CAHN, P.; CUENCA-ESTRELLA, M.; DAVEL, G.; RODRIGUEZ-TUDELA, J. L. G484S amino acid substitution in lanosterol 14-alpha demethylase (ERG11) is related to fluconazole resistance in a recurrent *Cryptococcus neoformans* clinical isolate. **Antimicrobial Agents and Chemotherapy**, v. 47, n. 11, p. 3653-3656, Nov 2003.
- 61 IMBERT, F.; JARDIN, M.; FERNANDEZ, C.; GANTIER, J. C.; DROMER, F.; BARON, G.; MENTRE, F.; VAN BEIJSTERVELDT, L.; SINGLAS, E.; GIMENEZ, F. Effect of efflux inhibition on brain uptake of itraconazole in mice infected with *Cryptococcus neoformans*. **Drug Metabolism and Disposition**, v. 31, n. 3, p. 319-325, Mar 2003.
- 62 MOTA, S. G. R.; BARROS, T. F.; CASTILHO, M. S. 2D QSAR Studies on a Series of Bifonazole Derivatives with Antifungal Activity. **Journal of the Brazilian Chemical Society**, v. 20, n. 3, p. 451-459, 2009.
- 63 DUCHOWICZ, P. R.; VITALE, M. G.; CASTRO, E. A.; FERNANDEZ, M.; CABALLERO, J. QSAR analysis for heterocyclic antifungals. **Bioorganic & Medicinal Chemistry**, v. 15, n. 7, p. 2680-2689, Apr 1 2007.

- 64 KATRITZKY, A. R.; SLAVOV, S. H.; DOBCHEV, D. A.; KARELSON, M. QSAR modeling of the antifungal activity against *Candida albicans* for a diverse set of organic compounds. **Bioorganic & Medicinal Chemistry**, v. 16, n. 14, p. 7055-7069, Jul 15 2008.
- 65 JI, H.; ZHANG, W.; ZHOU, Y.; ZHANG, M.; ZHU, J.; SONG, Y.; LU, J. A three-dimensional model of lanosterol 14 α -demethylase of *Candida albicans* and its interaction with azole antifungals. **Journal of Medicinal Chemistry**, v. 43, n. 13, p. 2493-2505, Jun 29 2000.
- 66 TSUKUDA, T.; SHIRATORI, Y.; WATANABE, M.; ONTSUKA, H.; HATTORI, K.; SHIRAI, M.; SHIMMA, N. Modeling, synthesis and biological activity of novel antifungal agents (1). **Bioorganic & Medicinal Chemistry Letters**, v. 8, n. 14, p. 1819-1824, Jul 21 1998.
- 67 XIAO, L.; MADISON, V.; CHAU, A. S.; LOEBENBERG, D.; PALERMO, R. E.; MCNICHOLAS, P. M. Three-dimensional models of wild-type and mutated forms of cytochrome P450 14 α -sterol demethylases from *Aspergillus fumigatus* and *Candida albicans* provide insights into posaconazole binding. **Antimicrobial Agents and Chemotherapy**, v. 48, n. 2, p. 568-574, Feb 2004.
- 68 FUKUOKA, T.; JOHNSTON, D. A.; WINSLOW, C. A.; DE GROOT, M. J.; BURT, C.; HITCHCOCK, C. A.; FILLER, S. G. Genetic basis for differential activities of fluconazole and voriconazole against *Candida krusei*. **Antimicrobial Agents and Chemotherapy**, v. 47, n. 4, p. 1213-1219, Apr 2003.
- 69 REOLON, A.; PEREZ, L. R. R.; MEZZARI, A. Prevalência de *Cryptococcus neoformans* nos pombos urbanos da cidade de Porto Alegre, Rio Grande do Sul. **Jornal Brasileiro de Patologia e Medicina Laboratorial**, v. 40, p. 293-298, 2004.
- 70 MITCHELL, T. G.; PERFECT, J. R. Cryptococcosis in the Era of Aids - 100 Years after the Discovery of *Cryptococcus neoformans*. **Clinical Microbiology Reviews**, v. 8, n. 4, p. 515-548, Oct 1995.
- 71 MICOL, R.; LORTHOLARY, O.; SAR, B.; LAUREILLARD, D.; NGETH, C.; DOUSSET, J. P.; CHANROEUN, H.; TERRADINI, L.; GUERIN, P. J.; DROMER, F.; FONTANET, A. Prevalence, determinants of positivity, and clinical utility of cryptococcal antigenemia in Cambodian HIV-infected patients. **Journal of Acquired Immune Deficiency Syndromes**, v. 45, n. 5, p. 555-559, Aug 15 2007.

- 72 HAJJEH, R. A.; CONN, L. A.; STEPHENS, D. S.; BAUGHMAN, W.; HAMILL, R.; GRAVISS, E.; PAPPAS, P. G.; THOMAS, C.; REINGOLD, A.; ROTHROCK, G.; HUTWAGNER, L. C.; SCHUCHAT, A.; BRANDT, M. E.; PINNER, R. W.; GRP, C. A. S. Cryptococcosis: Population-based multistate active surveillance and risk factors in human immunodeficiency virus-infected persons. **Journal of Infectious Diseases**, v. 179, n. 2, p. 449-454, Feb 1999.
- 73 KWON-CHUNG, K. J.; BOEKHOUT, T.; FELL, J. W.; DIAZ, M. Proposal to Conserve the Name *Cryptococcus gattii* against *C. hondurianus* and *C. bacillisporus* (Basidiomycota, Hymenomycetes, Tremellomycetidae). **Taxon**, v. 51, n. 4, p. 804-806, 2002.
- 74 CHAYAKULKEEREE, M.; PERFECT, J. R. Cryptococcosis. **Infectious disease clinics of North America**, v. 20, n. 3, p. 507-544, Sep 2006.
- 75 LIN, X.; HEITMAN, J. The biology of the *Cryptococcus neoformans* species complex. **Annual review of microbiology**, v. 60, p. 69-105, 2006.
- 76 PAPPAS, P. G.; PERFECT, J. R.; CLOUD, G. A.; LARSEN, R. A.; PANKEY, G. A.; LANCASTER, D. J.; HENDERSON, H.; KAUFFMAN, C. A.; HAAS, D. W.; SACCENTE, M.; HAMILL, R. J.; HOLLOWAY, M. S.; WARREN, R. M.; DISMUKES, W. E. Cryptococcosis in human immunodeficiency virus-negative patients in the era of effective azole therapy. **Clinical infectious diseases**, v. 33, n. 5, p. 690-699, Sep 1 2001.
- 77 LAURENSEN, I. F.; TREVETT, A. J.; LALLOO, D. G.; NWOKOLO, N.; NARAQI, S.; BLACK, J.; TEFURANI, N.; SAWERI, A.; MAVO, B.; IGO, J.; WARRELL, D. A. Meningitis caused by *Cryptococcus neoformans* var. *gattii* and var. *neoformans* in Papua New Guinea. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 90, n. 1, p. 57-60, 1996.
- 78 CHEN, S.; SORRELL, T.; NIMMO, G.; SPEED, B.; CURRIE, B.; ELLIS, D.; MARRIOTT, D.; PFEIFFER, T.; PARR, D.; BYTH, K.; GROUP, A. C. S. Epidemiology and Host- and Variety-Dependent Characteristics of Infection Due to *Cryptococcus neoformans* in Australia and New Zealand. **Clinical Infectious Diseases**, v. 31, n. 2, p. 499-508, August 1, 2000 2000.
- 79 SPEED, B.; DUNT, D. Clinical and Host Differences Between Infections with the Two Varieties of *Cryptococcus neoformans*. **Clinical Infectious Diseases**, v. 21, n. 1, p. 28-34, July 1, 1995 1995.

- 80 PERFECT, J. R.; DISMUKES, W. E.; DROMER, F.; GOLDMAN, D. L.; GRAYBILL, J. R.; HAMILL, R. J.; HARRISON, T. S.; LARSEN, R. A.; LORTHOLARY, O.; NGUYEN, M. H.; PAPPAS, P. G.; POWDERLY, W. G.; SINGH, N.; SOBEL, J. D.; SORRELL, T. C. Clinical practice guidelines for the management of cryptococcal disease: 2010 update by the infectious diseases society of america. **Clinical infectious diseases**, v. 50, n. 3, p. 291-322, Feb 1 2010.
- 81 HIGGINS, M. J.; GRAHAM, S. J. Intellectual property. Balancing innovation and access: patent challenges tip the scales. **Science**, v. 326, n. 5951, p. 370-371, Oct 16 2009.
- 82 SMITH, A. M.; AMMAR, R.; NISLOW, C.; GIAEVER, G. A survey of yeast genomic assays for drug and target discovery. **Pharmacology & Therapeutics**, v. 127, n. 2, p. 156-164, Aug 2010.
- 83 MICHELINI, E.; CEVENINI, L.; MEZZANOTTE, L.; COPPA, A.; RODA, A. Cell-based assays: fuelling drug discovery. **Analytical and Bioanalytical Chemistry**, v. 398, n. 1, p. 227-238, Sep 2010.
- 84 EGLLEN, R. M.; GILCHRIST, A.; REISINE, T. An overview of drug screening using primary and embryonic stem cells. **Combinatorial Chemistry & High Throughput Screening**, v. 11, n. 7, p. 566-572, Aug 2008.
- 85 HARRIS, J. Letting the Target Determine your Compound Acquisition Strategy. 2009 2009. Disponível em: < http://www.ddw-online.com/enabling_technologies/302930/letting_the_target_determine_your_compound_acquisition_strategy.html >. Acesso em: 01 mar 2011.
- 86 ESPINEL-INGROFF, A. Standardized disk diffusion method for yeasts. **Clinical Microbiology Newsletter**, v. 29, n. 13, p. 97-100, 2007.
- 87 KIRAZ, N.; DAG, I.; OZ, Y.; YAMAC, M.; KIREMITCI, A.; KASIFOGLU, N. Correlation between broth microdilution and disk diffusion methods for antifungal susceptibility testing of caspofungin, voriconazole, amphotericin B, itraconazole and fluconazole against *Candida glabrata*. **Journal of Microbiological Methods**, v. 82, n. 2, p. 136-140, Aug 2010.
- 88 ALBANO, C.; DUNN, W.; EDLUND, U.; JOHANSSON, E.; NORDÉN, B.; SJÖSTRÖM, M.; WOLD, S. Four levels of pattern recognition. **Analytica Chimica Acta**, v. 103, n. 4, p. 429-443, 1978.
- 89 TOU, J. T.; GONZÁLEZ, R. C. **Pattern recognition principles**. Addison-Wesley, 1977. ISBN 9780201075861.

- 90 OTTO, M. **Chemometrics: statistics and computer application in analytical chemistry**. Wiley-VCH, 2007. ISBN 9783527314188.
- 91 ZHENG, W.; TROPSHA, A. Novel Variable Selection Quantitative Structure–Property Relationship Approach Based on the k-Nearest-Neighbor Principle. **Journal of Chemical Information and Computer Sciences**, v. 40, n. 1, p. 185-194, 1999.
- 92 KOWALSKI, B. R.; BENDER, C. F. Pattern recognition. Powerful approach to interpreting chemical data. **Journal of the American Chemical Society**, v. 94, n. 16, p. 5632-5639, 1972.
- 93 MOLFETTA, F. A.; BRUNI, A. T.; HONÓRIO, K. M.; DA SILVA, A. B. F. A structure-activity relationship study of quinone compounds with trypanocidal activity. **European Journal of Medicinal Chemistry**, v. 40, n. 4, p. 329-338, 2005.
- 94 DE MAESSCHALCK, R.; JOUAN-RIMBAUD, D.; MASSART, D. L. The Mahalanobis distance. **Chemometrics and Intelligent Laboratory Systems**, v. 50, n. 1, p. 1-18, 2000.
- 95 SCHEUBER, M. Potentials and limits of the k-nearest-neighbour method for regionalising sample-based data in forestry **European Journal of Forest Research**, v. 129, n. 5, p. 825-832, 2010.
- 96 GJERTSEN, A. K. Accuracy of forest mapping based on Landsat TM data and a kNN-based method. **Remote Sensing of Environment**, v. 110, n. 4, p. 420-430, 2007.
- 97 JOLLIFFE, I. **Principal component analysis**. Springer, 2002. ISBN 9780387954424.
- 98 TAN, C.; QIN, X.; LI, M. Comparison of chemometric methods for brand classification of cigarettes by near-infrared spectroscopy. **Vibrational Spectroscopy**, v. 51, n. 2, p. 276-282, 2009.
- 99 FLÅTEN, G. R.; GRUNG, B.; KVALHEIM, O. M. A method for validation of reference sets in SIMCA modelling. **Chemometrics and Intelligent Laboratory Systems**, v. 72, n. 1, p. 101-109, 2004.

- 100 SABIN, J. G.; FERRÃO, M. F.; FURTADO, J. C. Análise multivariada aplicada na identificação de fármacos antidepressivos. Parte II: Análise por componentes principais (PCA) e o método de classificação SIMCA. **Revista Brasileira de Ciências Farmacêuticas**, v. 40, p. 387-396, 2004.
- 101 SENA, M. M. D.; POPPI, R. J.; FRIGHETTO, R. T. S.; VALARINI, P. J. Avaliação do uso de métodos quimiométricos em análise de solos. **Química Nova**, v. 23, p. 547-556, 2000.
- 102 WEBER, K. C.; DA SILVA, A. B. F. A chemometric study of the 5-HT_{1A} receptor affinities presented by arylpiperazine compounds. **European Journal of Medicinal Chemistry**, v. 43, n. 2, p. 364-372, 2008.
- 103 MYERS, R. **Classical and modern regression with applications**. PWS-KENT, 1990. ISBN 9780534921781.
- 104 WOLD, S.; SJÖSTRÖM, M.; ERIKSSON, L. **Partial Least Squares Projections to Latent Structures (PLS) in Chemistry**. John Wiley & Sons, Ltd, 2002. ISBN 9780470845011.
- 105 WOLD, S.; TRYGG, J.; BERGLUND, A.; ANTTI, H. Some recent developments in PLS modeling. **Chemometrics and Intelligent Laboratory Systems**, v. 58, n. 2, p. 131-150, 2001.
- 106 TOPLISS, J.; COSTELLO, R. Chance correlations in structure-activity studies using multiple regression analysis. **Journal of Medicinal Chemistry**, v. 15, n. 10, p. 1066-1068, 1972.
- 107 TOPLISS, J.; EDWARDS, R. Chance factors in studies of quantitative structure-activity relationships. **Journal of Medicinal Chemistry**, v. 22, n. 10, p. 1238-1244, 1979.
- 108 LIVINGSTONE, D. **Data analysis for chemists: applications to QSAR and chemical product design**. Oxford University Press, 1995. ISBN 9780198557289.
- 109 CLARK, M.; CRAMER, R. D. The Probability of Chance Correlation Using Partial Least Squares (PLS). **Quantitative Structure-Activity Relationships**, v. 12, n. 2, p. 137-145, 1993.
- 110 WATERBEEMD, H. V.; ROSE, S. **Quantitative approaches to structure-activity relationships**. The practice of medicinal chemistry. WERMUTH, C. G.: Elsevier: 351-369 p. 2003.

- 111 LEACH, A.; GILLET, V. **An Introduction to Chemoinformatics**. Springer, 2007. ISBN 9781402062902.
- 112 GELADI, P.; KOWALSKI, B. R. Partial least-squares regression: a tutorial. **Analytica Chimica Acta**, v. 185, p. 1-17, 1986.
- 113 GUIDO, R.; CASTILHO, M.; MOTA, S.; OLIVA, G.; ANDRICOPULO, A. Classical and Hologram QSAR Studies on a Series of Inhibitors of Trypanosomatid Glyceraldehyde-3-Phosphate Dehydrogenase. **QSAR & Combinatorial Science**, v. 27, n. 6, p. 768-781, 2008.
- 114 CASTILHO, M. S.; GUIDO, R. V.; ANDRICOPULO, A. D. 2D Quantitative structure-activity relationship studies on a series of cholesteryl ester transfer protein inhibitors. **Bioorganic & Medicinal Chemistry**, v. 15, n. 18, p. 6242-52, Sep 15 2007.
- 115 WANG, X.; TANG, S.; LIU, S.; CUI, S.; WANG, L. Molecular hologram derived quantitative structure-property relationships to predict physico-chemical properties of polychlorinated biphenyls. **Chemosphere**, v. 51, n. 7, p. 617-632, May 2003.
- 116 MYINT, K. Z.; XIE, X. Q. Recent advances in fragment-based QSAR and multi-dimensional QSAR methods. **International journal of molecular sciences**, v. 11, n. 10, p. 3846-3866, 2010.
- 117 LOWIS, D. R. HQSAR: A New, Highly Predictive QSAR Technique. **Tripos Technical Notes**, v. 1, n. 5, p. 1-17, 1997.
- 118 SEEL, M.; TURNER, D. B.; WILLETT, P. Effect of parameter variations on the effectiveness of HQSAR analyses. **Quantitative Structure-Activity Relationships**, v. 18, n. 3, p. 245-252, Jul 1999.
- 119 TONG, W.; LOWIS, D. R.; PERKINS, R.; CHEN, Y.; WELSH, W. J.; GODDETTE, D. W.; HERITAGE, T. W.; SHEEHAN, D. M. Evaluation of quantitative structure-activity relationship methods for large-scale prediction of chemicals binding to the estrogen receptor. **Journal of Chemical Information and Computer Sciences**, v. 38, n. 4, p. 669-677, Jul-Aug 1998.
- 120 CASTILHO, M. S.; POSTIGO, M. P.; DE PAULA, C. B. V.; MONTANARI, C. A.; OLIVA, G.; ANDRICOPULO, A. D. Two- and three-dimensional quantitative structure-activity relationships for a series of purine nucleoside phosphorylase inhibitors. **Bioorganic & Medicinal Chemistry**, v. 14, n. 2, p. 516-527, 2006.

- 121 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, Jun 15 2005.
- 122 CLSI. **Method for Antifungal Disk Diffusion Susceptibility Testing of Yeasts. Document M44-A2**. INSTITUTE, C.-C. A. L. S. Pennsylvania: Clinical and Laboratory Standards Institute 29 2009.
- 123 _____. **Método de Referência para Testes de Diluição em Caldo para a Determinação da Sensibilidade a Terapia Antifúngica das Leveduras. Documento M27-A2**. INSTITUTE, C.-C. A. L. S. Pennsylvania: Clinical and Laboratory Standards Institute 22 2002.
- 124 CLARK, M.; CRAMER, R. D.; VAN OPDENBOSCH, N. Validation of the general purpose tripos 5.2 force field. **Journal of Computational Chemistry**, v. 10, n. 8, p. 982-1012, 1989.
- 125 STEWART, J. J. Optimization of parameters for semiempirical methods IV: extension of MNDO, AM1, and PM3 to more main group elements. **Journal of molecular modeling**, v. 10, n. 2, p. 155-164, Apr 2004.
- 126 CAVALCANTI, A. R. O.; SOARES LEITE, E.; NETO, B. B.; FERREIRA, R. On the Classes of Aminoacyl-tRNA Synthetases, Amino Acids and the Genetic Code. **Origins of Life and Evolution of Biospheres**, v. 34, n. 4, p. 407-420, 2004.
- 127 WANG, K. L.; WEN, Z. N.; NIE, F. S.; LI, M. L. A New Hybrid Model of Amino Acid Substitution for Protein Functional Classification. **Chinese Chemical Letters**, v. 16, n. 8, p. 1133-1136, 2005.
- 128 MOBYDIGS. **Mobydigs 1.0**. Milano: Talete_srl 2004.
- 129 TODESCHINI, R.; CONSONNI, V.; MAURI, A.; PAVAN, M. Detecting "bad" regression models: multicriteria fitness functions in regression analysis. **Analytica Chimica Acta**, v. 515, n. 1, p. 199-208, 2004.
- 130 TODESCHINI, R.; CONSONNI, V.; MAIOCCHI, A. The K correlation index: theory development and its application in chemometrics. **Chemometrics and Intelligent Laboratory Systems**, v. 46, n. 1, p. 13-29, Feb 15 1999.
- 131 SHAH, N. K.; GEMPERLINE, P. J. Combination of the Mahalanobis distance and residual variance pattern recognition techniques for classification of near-infrared reflectance spectra. **Analytical Chemistry**, v. 62, n. 5, p. 465-470, 1990.

- 132 SCHUURMANN, G.; EBERT, R. U.; CHEN, J.; WANG, B.; KUHNE, R. External validation and prediction employing the predictive squared correlation coefficient test set activity mean vs training set activity mean. **Journal of Chemical Information and Modeling**, v. 48, n. 11, p. 2140-2145, Nov 2008.
- 133 DODDAREDDY, M. R.; CHO, Y. S.; KOH, H. Y.; PAE, A. N. CoMFA and CoMSIA 3D QSAR analysis on N1-arylsulfonylindole compounds as 5-HT6 antagonists. **Bioorganic & Medicinal Chemistry**, v. 12, n. 15, p. 3977-3985, Aug 1 2004.
- 134 PFALLER, M. A.; BALE, M.; BUSCHELMAN, B.; LANCASTER, M.; ESPINEL-INGROFF, A.; REX, J. H.; RINALDI, M. G.; COOPER, C. R.; MCGINNIS, M. R. Quality control guidelines for National Committee for Clinical Laboratory Standards recommended broth macrodilution testing of amphotericin B, fluconazole, and flucytosine. **Journal of clinical microbiology**, v. 33, n. 5, p. 1104-1107, May 1995.
- 135 PFALLER, M. A.; ZHANG, J.; MESSER, S. A.; BRANDT, M. E.; HAJJEH, R. A.; JESSUP, C. J.; TUMBERLAND, M.; MBIDDE, E. K.; GHANNOUM, M. A. In vitro activities of voriconazole, fluconazole, and itraconazole against 566 clinical isolates of *Cryptococcus neoformans* from the United States and Africa. **Antimicrobial Agents and Chemotherapy**, v. 43, n. 1, p. 169-171, Jan 1999.
- 136 REX, J. H.; PFALLER, M. A.; LANCASTER, M.; ODDS, F. C.; BOLMSTROM, A.; RINALDI, M. G. Quality control guidelines for National Committee for Clinical Laboratory Standards - recommended broth macrodilution testing of ketoconazole and itraconazole. **Journal of clinical microbiology**, v. 34, n. 4, p. 816-817, Apr 1996.
- 137 SABATELLI, F.; PATEL, R.; MANN, P. A.; MENDRICK, C. A.; NORRIS, C. C.; HARE, R.; LOEBENBERG, D.; BLACK, T. A.; MCNICHOLAS, P. M. In vitro activities of posaconazole, fluconazole, itraconazole, voriconazole, and amphotericin B against a large collection of clinically important molds and yeasts. **Antimicrobial Agents and Chemotherapy**, v. 50, n. 6, p. 2009-2015, Jun 2006.
- 138 GUIDO, R.; TROSSINI, G.; CASTILHO, M.; OLIVA, G.; FERREIRA, E.; ANDRICOPULO, A. Structure-activity relationships for a class of selective inhibitors of the major cysteine protease from *Trypanosoma cruzi*. **Journal of Enzyme Inhibition and Medicinal Chemistry**, v. 23, n. 6, p. 964-973, 2008.
- 139 GUIDO, R. V.; OLIVA, G.; MONTANARI, C. A.; ANDRICOPULO, A. D. Structural basis for selective inhibition of trypanosomatid glyceraldehyde-3-phosphate dehydrogenase: molecular docking and 3D QSAR studies. **Journal of Chemical Information and Modeling**, v. 48, n. 4, p. 918-929, Apr 2008.

- 140 MOTA, S. G. R.; BARROS, T. F.; CASTILHO, M. S. In vitro Screening and Chemometrics Analysis on a Series of Azole Derivatives with Fungicide Activity Against *Moniliophthora perniciosa*. **Journal of the Brazilian Chemical Society**, v. 21, n. 3, p. 510-519, 2010.
- 141 SHARAF, M. A.; ILLMAN, D. L.; KOWALSKI, B. R. **Chemometrics**. Wiley, 1986. ISBN 9780471831068.
- 142 INFOMETRIX. **Pirouette - Multivariate Data Analysis**. Woodinville: Infometrix, Inc. 2008.
- 143 GONZALEZ-ARJONA, D.; GONZALEZ, A. G. Adaptation of linear discriminant analysis to second level-pattern recognition classification. **Analytica Chimica Acta**, v. 363, n. 1, p. 89-95, May 11 1998.
- 144 WOLD, S. Pattern recognition by means of disjoint principal components models. **Pattern Recognition**, v. 8, n. 3, p. 127-139, 1976.
- 145 DASZYKOWSKI, M.; KACZMAREK, K.; STANIMIROVA, I.; VANDER HEYDEN, Y.; WALCZAK, B. Robust SIMCA-bounding influence of outliers. **Chemometrics and Intelligent Laboratory Systems**, v. 87, n. 1, p. 95-103, 2007.
- 146 BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, M. B. **Chemometrics: a practical guide**. Wiley-Interscience, 1998. 360 ISBN 9780471124511.
- 147 BRERETON, R. G. **Chemometrics for pattern recognition**. Wiley 2009. ISBN 9780470987254.
- 148 GALVEZ, J.; GARCIA, R.; SALABERT, M. T.; SOLER, R. Charge Indexes - New Topological Descriptors. **Journal of Chemical Information and Computer Sciences**, v. 34, n. 3, p. 520-525, May-Jun 1994.
- 149 GALVEZ, J.; GARCIA-DOMENECH, R.; DE JULIAN-ORTIZ, J. V.; SOLER, R. Topological approach to drug design. **Journal of Chemical Information and Computer Sciences**, v. 35, n. 2, p. 272-284, Mar-Apr 1995.
- 150 BURDEN, F. R. A chemically intuitive molecular index based on the eigenvalues of a modified adjacency matrix. **Quantitative Structure-Activity Relationships**, v. 16, n. 4, p. 309-314, Aug 1997.

- 151 JI, H.; ZHANG, W.; ZHANG, M.; KUDO, M.; AOYAMA, Y.; YOSHIDA, Y.; SHENG, C.; SONG, Y.; YANG, S.; ZHOU, Y.; LÜ, J.; ZHU, J. Structure-Based de Novo Design, Synthesis, and Biological Evaluation of Non-Azole Inhibitors Specific for Lanosterol 14 α -Demethylase of Fungi. **Journal of Medicinal Chemistry**, v. 46, n. 4, p. 474-485, 2003.
- 152 SHENG, C.; ZHANG, W.; JI, H.; ZHANG, M.; SONG, Y.; XU, H.; ZHU, J.; MIAO, Z.; JIANG, Q.; YAO, J.; ZHOU, Y.; LU, J. Structure-based optimization of azole antifungal agents by CoMFA, CoMSIA, and molecular docking. **Journal of Medicinal Chemistry**, v. 49, n. 8, p. 2512-2525, Apr 20 2006.
- 153 MOTA, S. G. R. **Ensaio *in vitro* e análise quimimétrica de inibidores da enzima lanosterol 14 α -desmetilase de *Moniliophthora perniciosa*** 2009. 77 p. Dissertação (Mestrado). UEFS, Feira de Santana.
- 154 LUPETTI, A.; DANESI, R.; CAMPA, M.; DEL TACCA, M.; KELLY, S. Molecular basis of resistance to azole antifungals. **Trends in molecular medicine**, v. 8, n. 2, p. 76-81, Feb 2002.
- 155 KONTOYIANNIS, D. P.; LEWIS, R. E. Antifungal drug resistance of pathogenic fungi. **Lancet**, v. 359, n. 9312, p. 1135-1144, Mar 30 2002.
- 156 VARMA, A.; KWON-CHUNG, K. J. Heteroresistance of *Cryptococcus gattii* to fluconazole. **Antimicrobial Agents and Chemotherapy**, v. 54, n. 6, p. 2303-2311, Jun 2010.
- 157 PFALLER, M. A.; DIEKEMA, D. J. Epidemiology of Invasive Mycoses in North America. **Critical Reviews in Microbiology**, v. 36, n. 1, p. 1-53, Feb 2010.
- 158 MORIO, F.; LOGE, C.; BESSE, B.; HENNEQUIN, C.; LE PAPE, P. Screening for amino acid substitutions in the *Candida albicans* ERG11 protein of azole-susceptible and azole-resistant clinical isolates: new substitutions and a review of the literature. **Diagnostic microbiology and infectious disease**, v. 66, n. 4, p. 373-384, Apr 2010.
- 159 SANGLARD, D. Clinical relevance of mechanisms of antifungal drug resistance in yeasts. **Enfermedades infecciosas y microbiología clínica**, v. 20, n. 9, p. 462-469, Nov 2002.
- 160 MUKHERJEE, P. K.; CHANDRA, J.; KUHN, D. A.; GHANNOUM, M. A. Mechanism of fluconazole resistance in *Candida albicans* biofilms: Phase-specific role of efflux pumps and membrane sterols. **Infection and Immunity**, v. 71, n. 8, p. 4333-4340, Aug 2003.

- 161 VENKATESWARLU, K.; DENNING, D. W.; MANNING, N. J.; KELLY, S. L. Resistance to fluconazole in *Candida albicans* from AIDS patients correlated with reduced intracellular accumulation of drug. **FEMS microbiology letters**, v. 131, n. 3, p. 337-341, Sep 15 1995.
- 162 SANGLARD, D.; KUCHLER, K.; ISCHER, F.; PAGANI, J. L.; MONOD, M.; BILLE, J. Mechanisms of resistance to azole antifungal agents in *Candida albicans* isolates from AIDS patients involve specific multidrug transporters. **Antimicrobial agents and chemotherapy**, v. 39, n. 11, p. 2378-2386, Nov 1995.
- 163 ALBERTSON, G. D.; NIIMI, M.; CANNON, R. D.; JENKINSON, H. F. Multiple efflux mechanisms are involved in *Candida albicans* fluconazole resistance. **Antimicrobial Agents and Chemotherapy**, v. 40, n. 12, p. 2835-2841, Dec 1996.
- 164 PINTO E SILVA, A. T.; COSTA-DE-OLIVEIRA, S.; SILVA-DIAS, A.; PINA-VAZ, C.; RODRIGUES, A. G. Dynamics of in vitro acquisition of resistance by *Candida parapsilosis* to different azoles. **FEMS Yeast Research**, v. 9, n. 4, p. 626-633, 2009.
- 165 PINA-VAZ, C.; RODRIGUES, A. G.; COSTA-DE-OLIVEIRA, S.; RICARDO, E.; MÅRDH, P.-A. Potent synergic effect between ibuprofen and azoles on *Candida* resulting from blockade of efflux pumps as determined by FUN-1 staining and flow cytometry. **Journal of Antimicrobial Chemotherapy**, v. 56, n. 4, p. 678-685, October 2005 2005.
- 166 FORD, J. M.; HAIT, W. N. Pharmacology of Drugs That Alter Multidrug Resistance in Cancer. **Pharmacological Reviews**, v. 42, n. 3, p. 155-199, Sep 1990.
- 167 SANGLARD, D.; ODDS, F. C. Resistance of *Candida* species to antifungal agents: molecular mechanisms and clinical consequences. **Lancet Infectious Diseases**, v. 2, n. 2, p. 73-85, Feb 2002.
- 168 SANGLARD, D.; ISCHER, F.; KOYMANS, L.; BILLE, J. Amino acid substitutions in the cytochrome P-450 lanosterol 14 α -demethylase (CYP51A1) from azole-resistant *Candida albicans* clinical isolates contribute to resistance to azole antifungal agents. **Antimicrobial Agents and Chemotherapy**, v. 42, n. 2, p. 241-253, Feb 1998.
- 169 LEVIN, A. S.; COSTA, S. F.; MUSSI, N. S.; BASSO, M.; SINTO, S. I.; MACHADO, C.; GEIGER, D. C.; VILLARES, M. C. B.; SCHREIBER, A. Z.; BARONE, A. A.; BRANCHINI, M. L. M. *Candida parapsilosis* fungemia associated with implantable and semi-implantable central venous catheters and the hands of healthcare workers. **Diagnostic microbiology and infectious disease**, v. 30, n. 4, p. 243-249, Apr 1998.

- 170 KELLY, S. L.; LAMB, D. C.; LOEFFLER, J.; EINSELE, H.; KELLY, D. E. The G464S amino acid substitution in *Candida albicans* sterol 14 alpha-demethylase causes fluconazole resistance in the clinic through reduced affinity. **Biochemical and Biophysical Research Communications**, v. 262, n. 1, p. 174-179, Aug 19 1999.
- 171 GOLLAPUDY, R.; AJMANI, S.; KULKARNI, S. A. Modeling and interactions of *Aspergillus fumigatus* lanosterol 14-[alpha] demethylase `A' with azole antifungals. **Bioorganic & Medicinal Chemistry**, v. 12, n. 11, p. 2937-2950, 2004.
- 172 YOSHIDA, Y.; NOSHIRO, M.; AOYAMA, Y.; KAWAMOTO, T.; HORIUCHI, T.; GOTOH, O. Structural and evolutionary studies on sterol 14-demethylase P450 (CYP51), the most conserved P450 monooxygenase: II. Evolutionary analysis of protein and gene structures. **Journal of biochemistry**, v. 122, n. 6, p. 1122-1128, Dec 1997.
- 173 STRUSHKEVICH, N.; USANOV, S. A.; PARK, H.-W. Structural Basis of Human CYP51 Inhibition by Antifungal Azoles. **Journal of Molecular Biology**, v. 397, n. 4, p. 1067-1078, 2010.
- 174 LEPESHEVA, G. I.; HARGROVE, T. Y.; ANDERSON, S.; KLESHCHENKO, Y.; FURTAK, V.; WAWRZAK, Z.; VILLALTA, F.; WATERMAN, M. R. Structural Insights into Inhibition of Sterol 14 α -Demethylase in the Human Pathogen *Trypanosoma cruzi*. **Journal of Biological Chemistry**, v. 285, n. 33, p. 25582-25590, August 13, 2010 2010.
- 175 MARTINELLI, A.; TUCCINARDI, T. An overview of recent developments in GPCR modelling: methods and validation. **Expert Opinion on Drug Discovery**, v. 1, n. 5, p. 459-476, 2006.
- 176 HÖLTJE, H. D.; SIPPL, W.; ROGNAN, D.; FOLKERS, G. Introduction to comparative protein modeling. In: HÖLTJE, H. D. (Ed.). **Molecular modeling: basic principles and applications**. Weinheim: Wiley-VCH, 2008. p.111-125. ISBN 9783527315680.
- 177 NAYEEM, A.; SITKOFF, D.; KRYSSTEK, S., JR. A comparative study of available software for high-accuracy homology modeling: from sequence alignments to structural models. **Protein science**, v. 15, n. 4, p. 808-824, Apr 2006.
- 178 BLUNDELL, T. L.; SIBANDA, B. L.; STERNBERG, M. J. E.; THORNTON, J. M. Knowledge-Based Prediction of Protein Structures and the Design of Novel Molecules. **Nature**, v. 326, n. 6111, p. 347-352, Mar 26 1987.

- 179 GREER, J. Comparative Modeling Methods - Application to the Family of the Mammalian Serine Proteases. **Proteins-Structure Function and Genetics**, v. 7, n. 4, p. 317-334, 1990.
- 180 BENNER, S. A.; CANNAROZZI, G.; GERLOFF, D.; TURCOTTE, M.; CHELVANAYAGAM, G. Bona fide predictions of protein secondary structure using transparent analyses of multiple sequence alignments. **Chemical Reviews**, v. 97, n. 8, p. 2725-2843, Dec 1997.
- 181 CAVASOTTO, C. N.; PHATAK, S. S. Homology modeling in drug discovery: current trends and applications. **Drug Discovery Today**, v. 14, n. 13-14, p. 676-683, Jul 2009.
- 182 HILLISCH, A.; PINEDA, L. F.; HILGENFELD, R. Utility of homology models in the drug discovery process. **Drug Discovery Today**, v. 9, n. 15, p. 659-669, Aug 1 2004.
- 183 DEANE, C. M.; BLUNDELL, T. L. Protein comparative modelling and drug discovery. In: WERMUTH, C. G. (Ed.). **The Practice of Medicinal Chemistry**. 2. London: Elsevier/Academic Press, 2003. cap. 27, p.445-458. ISBN 9780123741943.
- 184 MOUNT, D. W. Using the Basic Local Alignment Search Tool (BLAST). **CSH protocols**, v. 2007, n. pdb.top17, p. -, 2007.
- 185 YOUNG, D. C. **Computational drug design: a guide for computational and medicinal chemists**. John Wiley & Sons, 2009. ISBN 9780470126851.
- 186 WEBSTER, D. M. **Protein structure prediction: methods and protocols**. Humana Press, 2000. ISBN 9780896036376.
- 187 MOSCA, R.; BRANNETTI, B.; SCHNEIDER, T. R. Alignment of protein structures in the presence of domain motions. **BMC Bioinformatics**, v. 9, p. -, Aug 27 2008.
- 188 JENNINGS, A. J.; EDGE, C. M.; STERNBERG, M. J. E. An approach to improving multiple alignments of protein sequences using predicted secondary structure. **Protein Engineering**, v. 14, n. 4, p. 227-231, Apr 2001.
- 189 BLAKE, J. D.; COHEN, F. E. Pairwise sequence alignment below the twilight zone. **Journal of Molecular Biology**, v. 307, n. 2, p. 721-735, Mar 23 2001.

- 190 SANCHEZ, R.; SALI, A. Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome. **Proceedings of the National Academy of Sciences of the United States of America**, v. 95, n. 23, p. 13597-13602, Nov 10 1998.
- 191 ESWAR, N.; JOHN, B.; MIRKOVIC, N.; FISER, A.; ILYIN, V. A.; PIEPER, U.; STUART, A. C.; MARTI-RENOM, M. A.; MADHUSUDHAN, M. S.; YERKOVICH, B.; SALI, A. Tools for comparative protein structure modeling and analysis. **Nucleic Acids Research**, v. 31, n. 13, p. 3375-3380, Jul 1 2003.
- 192 GUO, J. T.; ELLROTT, K.; XU, Y. A historical perspective of template-based protein structure prediction. **Methods in molecular biology**, v. 413, p. 3-42, 2008.
- 193 GUEX, N.; PEITSCH, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. **Electrophoresis**, v. 18, n. 15, p. 2714-2723, Dec 1997.
- 194 ARNOLD, K.; BORDOLI, L.; KOPP, J.; SCHWEDE, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. **Bioinformatics**, v. 22, n. 2, p. 195-201, January 15, 2006 2006.
- 195 SCHWEDE, T.; KOPP, J.; GUEX, N.; PEITSCH, M. C. SWISS-MODEL: An automated protein homology-modeling server. **Nucleic acids research**, v. 31, n. 13, p. 3381-3385, Jul 1 2003.
- 196 SALI, A. Comparative protein modeling by satisfaction of spatial restraints. **Molecular medicine today**, v. 1, n. 6, p. 270-277, Sep 1995.
- 197 SALI, A.; OVERINGTON, J. P. Derivation of rules for comparative protein modeling from a database of protein structure alignments. **Protein science**, v. 3, n. 9, p. 1582-1596, Sep 1994.
- 198 SANCHEZ, R.; SALI, A. Comparative protein structure modeling. Introduction and practical examples with modeller. **Methods in molecular biology**, v. 143, p. 97-129, 2000.

- 199 MACKERELL, A. D.; BASHFORD, D.; BELLOTT; DUNBRACK, R. L.; EVANSECK, J. D.; FIELD, M. J.; FISCHER, S.; GAO, J.; GUO, H.; HA, S.; JOSEPH-MCCARTHY, D.; KUCHNIR, L.; KUCZERA, K.; LAU, F. T. K.; MATTOS, C.; MICHNICK, S.; NGO, T.; NGUYEN, D. T.; PRODHOM, B.; REIHER, W. E.; ROUX, B.; SCHLENKRICH, M.; SMITH, J. C.; STOTE, R.; STRAUB, J.; WATANABE, M.; WIÓRKIEWICZ-KUCZERA, J.; YIN, D.; KARPLUS, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. **The Journal of Physical Chemistry B**, v. 102, n. 18, p. 3586-3616, 1998.
- 200 BROOKS, B. R.; BRUCCOLERI, R. E.; OLAFSON, B. D.; STATES, D. J.; SWAMINATHAN, S.; KARPLUS, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. **Journal of Computational Chemistry**, v. 4, n. 2, p. 187-217, 1983.
- 201 LASKOWSKI, R. A.; MACARTHUR, M. W.; MOSS, D. S.; THORNTON, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, n. 2, p. 283-291, 1993.
- 202 MORRIS, A. L.; MACARTHUR, M. W.; HUTCHINSON, E. G.; THORNTON, J. M. Stereochemical quality of protein structure coordinates. **Proteins-Structure Function and Bioinformatics**, v. 12, n. 4, p. 345-364, Apr 1992.
- 203 SHEN, M. Y.; SALI, A. Statistical potential for assessment and prediction of protein structures. **Protein science**, v. 15, n. 11, p. 2507-2524, Nov 2006.
- 204 LUTHY, R.; BOWIE, J. U.; EISENBERG, D. Assessment of protein models with three-dimensional profiles. **Nature**, v. 356, n. 6364, p. 83-85, 1992.
- 205 BOWIE, J. U.; LUTHY, R.; EISENBERG, D. A method to identify protein sequences that fold into a known three-dimensional structure. **Science**, v. 253, n. 5016, p. 164-170, July 12, 1991 1991.
- 206 BOSCOOTT, P. E.; GRANT, G. H. Modeling cytochrome P450 14 alpha demethylase (*Candida albicans*) from P450cam. **Journal of molecular graphics**, v. 12, n. 3, p. 185-192, Sep 1994.
- 207 HÖLTJE, H.-D.; FATTORUSSO, C. Construction of a model of the *Candida albicans* lanosterol 14-[alpha]-demethylase active site using the homology modelling technique. **Pharmaceutica Acta Helvetiae**, v. 72, n. 5, p. 271-277, 1998.

- 208 PODUST, L. M.; POULOS, T. L.; WATERMAN, M. R. Crystal structure of cytochrome P450 14alpha-sterol demethylase (CYP51) from *Mycobacterium tuberculosis* in complex with azole inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 6, p. 3068-3073, Mar 13 2001.
- 209 PODUST, L. M.; STOJAN, J.; POULOS, T. L.; WATERMAN, M. R. Substrate recognition sites in 14alpha-sterol demethylase from comparative analysis of amino acid sequences and X-ray structure of *Mycobacterium tuberculosis* CYP51. **Journal of inorganic biochemistry**, v. 87, n. 4, p. 227-235, Dec 15 2001.
- 210 BOLANO, A.; STINCHI, S.; PREZIOSI, R.; BISTONI, F.; ALLEGRUCCI, M.; BALDELLI, F.; MARTINI, A.; CARDINALI, G. Rapid methods to extract DNA and RNA from *Cryptococcus neoformans*. **FEMS Yeast Research**, v. 1, n. 3, p. 221-224, Dec 2001.
- 211 REVANKAR, S. G.; FU, J.; RINALDI, M. G.; KELLY, S. L.; KELLY, D. E.; LAMB, D. C.; KELLER, S. M.; WICKES, B. L. Cloning and characterization of the lanosterol 14alpha-demethylase (ERG11) gene in *Cryptococcus neoformans*. **Biochemical and biophysical research communications**, v. 324, n. 2, p. 719-728, Nov 12 2004.
- 212 ALTSCHUL, S. F.; MADDEN, T. L.; SCHAFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic acids research**, v. 25, n. 17, p. 3389-3402, Sep 1 1997.
- 213 GERSTEIN, M.; LESK, A. M.; CHOTHIA, C. Structural mechanisms for domain movements in proteins. **Biochemistry**, v. 33, n. 22, p. 6739-6749, Jun 7 1994.
- 214 ATILGAN, A. R.; DURELL, S. R.; JERNIGAN, R. L.; DEMIREL, M. C.; KESKIN, O.; BAHAR, I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. **Biophysical journal**, v. 80, n. 1, p. 505-515, Jan 2001.
- 215 BAHAR, I.; R., A. A.; ERMAN, B. Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. **Folding and Design**, v. 2, n. 3, p. 173-181, 1997.
- 216 HALILOGLU, T.; BAHAR, I.; ERMAN, B. Gaussian Dynamics of Folded Proteins. **Physical Review Letters**, v. 79, n. 16, p. 3090-3093, 1997.
- 217 SCOTT, E.; TARIQ, V.; MCCRORY, R. Demonstration of synergy with fluconazole and either ibuprofen, sodium salicylate, or propylparaben against *Candida albicans* in vitro. **Antimicrobial Agents and Chemotherapy**, v. 39, n. 12, p. 2610-2614, December 1, 1995 1995.

- 218 PODUST, L. M.; VON KRIES, J. P.; EDDINE, A. N.; KIM, Y.; YERMALITSKAYA, L. V.; KUEHNE, R.; OUELLET, H.; WARRIER, T.; ALTEKOSTER, M.; LEE, J. S.; RADEMANN, J.; OSCHKINAT, H.; KAUFMANN, S. H.; WATERMAN, M. R. Small-molecule scaffolds for CYP51 inhibitors identified by high-throughput screening and defined by X-ray crystallography. **Antimicrobial Agents and Chemotherapy**, v. 51, n. 11, p. 3915-3923, Nov 2007.
- 219 ESWAR, N.; WEBB, B.; MARTI-RENOM, M. A.; MADHUSUDHAN, M. S.; ERAMIAN, D.; SHEN, M.-Y.; PIEPER, U.; SALI, A. **Comparative Protein Structure Modeling Using MODELLER**. John Wiley & Sons, Inc., 2001. ISBN 9780471140863.
- 220 LI, X.; BROWN, N.; CHAU, A. S.; LOPEZ-RIBOT, J. L.; RUESGA, M. T.; QUINDOS, G.; MENDRICK, C. A.; HARE, R. S.; LOEBENBERG, D.; DIDOMENICO, B.; MCNICHOLAS, P. M. Changes in susceptibility to posaconazole in clinical isolates of *Candida albicans*. **The Journal of antimicrobial chemotherapy**, v. 53, n. 1, p. 74-80, Jan 2004.
- 221 AKINS, R. A. An update on antifungal targets and mechanisms of resistance in *Candida albicans*. **Medical Mycology**, v. 43, n. 4, p. 285-318, Jun 2005.

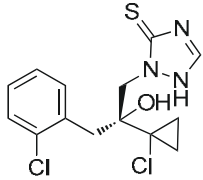
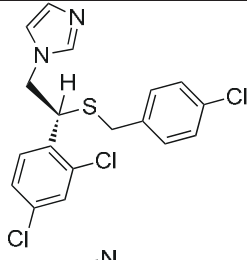
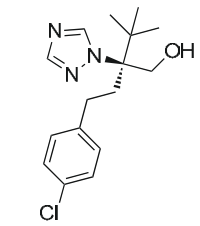
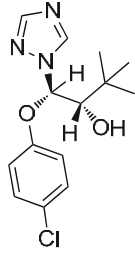
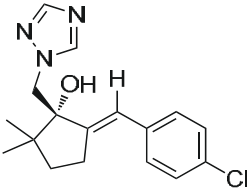
APÊNDICE A – Estrutura química dos antifúngicos utilizados nos estudos de avaliação da atividade antifúngicas frente *Cryptococcus* e estudos de QSAR

Nome	Estrutura	Nome	Estrutura
Azaconazol		Bifonazol	
Bromuconazol		Cetoconazol	
Climbazol		Clotrimazol	
Ciproconazol		Diclobutrazol	
Difenoconazol		Diniconazol	
Econazol		Epoxiconazol	
Etaconazol		Fenbuconazol	

Continua...

Nome	Estrutura	Nome	Estrutura
Fluconazol		Fluotrimazol	
Fluquinconazol		Flusilazol	
Flutriafol		Hexaconazol	
Imazalil		Itraconazol	
Metconazol		Miconazol	
Myclobutanil		Penconazol	
Prochloraz		Propiconazol	

Continua...

Nome	Estrutura	Nome	Estrutura
Prothioconazol		Sulconazol	
Tebuconazol		Triadimenol	
Triticonazol			

APÊNDICE B – Descritores selecionados por RLM

Descritores	Significado
Me	Média atômica da eletronegatividade de Sanderson (Escalonada em átomos de carbono)
Ms	Média do estado eletrotológico
nDB	Número de duplas ligações
J	Índice de distância e conectividade de Balaban
MAXDN	Variação eletrotológica negativa máxima
PW3	Índice de formato de Randic para distância topológica 3
T(F..F)	Soma das distâncias topológicas entre átomos de flúor
X2v	Índice (Chi-2) de conectividade de valências
X5v	Índice (Chi-5) de conectividade de valências
X2Av	Índice (Chi-2) de conectividade de valência média
SIC2	Índice contendo informações estruturais a simetria do vizinho de ordem 2
IC3	Índice contendo informação sobre a simetria do vizinho de ordem 3
ATS5e	Autocorrelação de Broto-Moreau para distância topológica de índice 5 ponderado pela eletronegatividade atômica de Sanderson
ATS6p	Autocorrelação de Broto-Moreau para distância topológica de índice 6 ponderado pela polarizabilidade atômica
MATS7m	Autocorrelação de Moran na distância topológica 7 ponderado pela massa atômica
MATS8m	Autocorrelação de Moran na distância topológica 8 ponderado pela massa atômica
MATS2e	Autocorrelação de Moran na distância topológica 2 ponderado pela eletronegatividade de Sanderson
GATS1m	Autocorrelação de Geary na distância topológica 1 ponderado pela massa atômica
GATS8m	Autocorrelação de Geary na distância topológica 8 ponderado pela massa atômica

Continua...

Descritores	Significado
GATS3v	Autocorrelação de Geary na distância topológica 3 ponderado pelo volume atômico de Van Der Walls
GATS1e	Autocorrelação de Geary na distância topológica 1 ponderado pela eletronegatividade de Sanderson
GATS5e	Autocorrelação de Geary na distância topológica 5 ponderado pela eletronegatividade de Sanderson
GATS4p	Autocorrelação de Geary na distância topológica 4 ponderado pela polarizabilidade
GATS6p	Autocorrelação de Geary na distância topológica 6 ponderado pela polarizabilidade
ESpm03u	Momento espectral (03) do eixo da matriz adjacente
BELe8	Menor autovalor da matriz de Burden na distância tipológica 8 ponderado pela eletronegatividade de Sanderson
GGI1	Índice de carga topológica de ordem 1
SEigZ	Soma de autovalores do eixo Z ponderado pela matriz de distância de Barysz
nCONN	Número de derivados de uréia
nN=C-N<	Número de derivados de amidina
nArX	Número de halogênios ligados à anel aromático
nCXr	Número de halogênios ligados à carbono (sp ³) de anel
C-003*	Fragmento centrado no átomo: CHR ₃ ^a
C-040*	Fragmento centrado no átomo: R-C(=X)-X ou R-C#X ou X=C=X ^{a,b}
H-052	Fragmento centrado no átomo: $\begin{array}{c} \text{R} \\ \\ \text{H}-\text{C}-\text{R}-\text{X}^{\text{a}} \\ \\ \text{R} \end{array}$
N-072	Fragmento centrado no átomo: RCO-N< ou >N-X=X

^a R representa qualquer agrupamento ligado através de carbonos. ^b “=” representa dupla ligação; “#” representa tripla ligação; X são heteroátomos (ie. O, N, S...)

APÊNDICE C – Instruções utilizadas para a construção do modelo comparativo no programa MODELER 9.8

```

# Homology modeling by the automodel class
from modeller import *          # Load standard Modeller classes
from modeller.automodel import * # Load the automodel class

log.verbose() # request verbose output
env = environ() # create a new MODELLER environment to build this model in

# directories for input atom files
env.io.atom_files_directory = ['.', './atom_files']

# Give less weight to all soft-sphere restraints:
env.schedule_scale = physical.values(default=1.0, soft_sphere=0.7)

# Read in HETATM records from template PDBs
env.io.hetatm = True

a = automodel(env,
               alnfile = 'Aln_seq_2CIB_3LD6.pir', # alignment filename
               knowns = ('2CIB_mod', '3LD6_mod'), # codes of the templates
               sequence = 'lan14ad', assess_methods=(assess.DOPE)) # code of the target
a.starting_model= 1 # index of the first model
a.ending_model = 10 # index of the last model
                  # (determines how many models to calculate)

#a.final_malign3d = True

# Very thorough VTFM optimization:
a.library_schedule = autosched.slow
a.max_var_iterations = 1000

# Thorough MD optimization:
a.md_level = refine.slow_large

# Repeat the whole cycle 4 times and do not stop unless obj.func. > 1E6
a.repeat_optimization = 4
a.max_molpdf = 1e6

a.make() # do the actual homology modelin

print "VALIDATION STARTS HERE"

# Get a list of all successfully built models from a.outputs
ok_models = filter(lambda x: x['failure'] is None, a.outputs)

# Rank the models by DOPE score
key = 'DOPE score'
ok_models.sort(lambda a,b: cmp(a[key], b[key]))

# Get top model
m = ok_models[0]
print "Top model: %s (DOPE score %.3f)" % (m['name'], m[key])

```