

**MODELAGEM MIA-QSAR DAS ATIVIDADES
BIOLÓGICAS DE INIBIDORES DE
FOSFODIESTERASE TIPO-5**

JOÃO EUSTÁQUIO ANTUNES

2009

JOÃO EUSTÁQUIO ANTUNES

**MODELAGEM MIA-QSAR DAS ATIVIDADES BIOLÓGICAS DE
INIBIDORES DE FOSFODIESTERASE TIPO-5**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Agroquímica, para a obtenção do título de “Mestre”.

Orientador
Prof. Dr. Matheus Puggina de Freitas

LAVRAS
MINAS GERAIS-BRASIL
2009

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Antunes, João Eustáquio.

Modelagem MIA-QSAR das atividades biológicas de
inibidores de fosfodiesterase tipo-5 / João Eustáquio Antunes. –
Lavras : UFLA, 2009.

104 p. : il.

Dissertação (Mestrado) – Universidade Federal de Lavras, 2009.

Orientador: Matheus Puggina de Freitas.

Bibliografia.

1. Disfunção erétil. 2. MIA-QSAR. 3. PLS. 4. Docking 5.
Inibidores da Fosfodiesterase tipo-5 I. Universidade Federal de
Lavras. II. Título.

CDD – 540.1133

JOÃO EUSTÁQUIO ANTUNES

**MODELAGEM MIA-QSAR DAS ATIVIDADES BIOLÓGICAS DE
INIBIDORES DE FOSFODIESTERASE TIPO-5**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Agroquímica, para a obtenção do título de “Mestre”.

APROVADA em 16 de fevereiro de 2009

Profª. Dra. Elaine Ferreira da Cunha UFLA

Prof. Dr. Antônio Chaufum Junior UFLA

Prof. Dr. Rene Luís de Oliveira Rigitano UFLA

Prof. Dr. Matheus Puggina de Freitas
UFLA
Orientador

LAVRAS
MINAS GERAIS-BRASIL

DEDICATÓRIA

A Deus

Aos meus pais: Júlio e Maria (*in memoria*)

A minha querida esposa Michelle

Ao professor Matheus Puggina de Freitas

Aos meus irmãos e familiares que me apoiaram

Aos amigos e colegas que me auxiliaram

AGRADECIMENTOS

Agradeço a Deus pela força maior, a vida e a capacidade de enfrentar os desafios. Nos momentos difíceis Ele esteve sempre do meu lado e nunca me abandonou.

Agradeço aos meus queridos pais Júlio (Vavá) pelo companheirismo nas partidas de sinuca e Maria (*in memoria*) por me mostrar na sua simplicidade e inteligência uma grande sabedoria, nos conselhos e exemplos dados por uma grande Mestre da educação. Hoje ela estaria muito feliz por esta grande conquista alcançada.

A minha querida esposa Michelle pelo apoio, amor e amizade. Seu incentivo foi muito importante para esta realização, sem sua ajuda e compreensão essa conquista jamais seria possível, eu a amo cada vez mais.

Ao Dr. Matheus Puggina de Freitas pela orientação e amizade. Por ter conduzido de forma brilhante este trabalho com muita inteligência e sabedoria. Pelo incentivo, apoio e por me mostrar como podemos realizar uma pesquisa de boa qualidade.

A Dra. Elaine Ferreira da Cunha pela colaboração, amizade e carinho.

Ao Dr. Roberto Rittner pela colaboração.

Aos professores do DQI/UFLA pelos ensinamentos e incentivos, em especial ao professor Dr. Teodorico pela amizade.

A todos aqueles que me ajudaram de alguma forma a desenvolver este projeto de pesquisa.

SUMÁRIO

	Página
LISTA DE FIGURAS.....	i
LISTA DE TABELAS.....	iii
LISTA DE ABREVIATURAS.....	iv
RESUMO GERAL.....	vi
GENERAL ABSTRACT.....	viii
CAPÍTULO 1: Introdução geral e revisão bibliográfica.....	1
1 Introdução geral.....	2
1.1 Aspectos gerais.....	2
1.2 Problemática.....	5
1.3 Tratamentos utilizados.....	6
1.4 Métodos computacionais – vantagens.....	8
2 Referencial teórico.....	10
2.1 Inibidores de fosfodiesterase tipo 5.....	10
2.1.1 Drogas mais utilizadas.....	11
2.1.2 Mecanismo de ação.....	16
2.1.3 Seletividade frente a outras fosfodiesterases.....	20
2.1.4 Efeitos colaterais.....	23
2.2 Métodos de modelagem.....	24
2.2.1 Química computacional.....	24
2.2.2 Aproximação do receptor (estudos de <i>docking</i>).....	26
2.2.3 Aproximação do ligante.....	30
2.3 Estudos QSAR de inibidores de PDE-5.....	50
3 Referências bibliográficas.....	53

CAPÍTULO 2: Avaliação preliminar do método MIA-QSAR para uma série de inibidores de PDE-5.....	60
1 Resumo	61
2 Abstract.....	62
3 Introdução	63
4 Métodos Computacionais	64
5 Objetivos.....	68
6 Resultados e Discussão	68
7 Conclusões.....	76
8 Referências Bibliográficas.....	77
CAPÍTULO 3: Proposição computacional de novos inibidores de PDE-5.....	78
1 Resumo	79
2 Abstract.....	80
3 Introdução	81
4 Métodos Computacionais	82
4.1 Construção do modelo MIA-QSAR	82
4.2 Cálculos de <i>Docking</i>	85
4.3 Cálculos ADME-Tox (absorção, distribuição, metabolismo, excreção e toxicidade)	86
5 Objetivos.....	88
6 Resultados e Discussão	89
6.1 MIA-QSAR	89
6.2 Estudos de <i>Docking</i>	94
6.3 Avaliação ADME-Tox.....	98
7 Conclusões.....	102
8 Referências Bibliográficas.....	103

LISTA DE FIGURAS

	Página
CAPÍTULO 1	
FIGURA 1	Importantes inibidores de PDE-5 disponíveis no mercado..... 5
FIGURA 2	Anatomia e mecanismo da ereção do pênis..... 17
FIGURA 3	Perspectiva celular da via de ereção..... 18
FIGURA 4	Representação gráfica de um modelo PARAFAC de dois componentes de arranjo dos dados X..... 46
FIGURA 5	A construção do arranjo 3D (adequado para tratamento usando métodos multimodos), que podem ser desdobrados para uma matriz 2D (matriz X), adequada para análises PLS..... 50
CAPÍTULO 2	
FIGURA 1	Estrutura molecular genérica dos derivados de guanina cíclica, inibidores de PDE-5 (1-59), desenhadas numa janela de dimensão 250 x 250 pixels. O ponto fixado na coordenada 50,100 está indicado na estrutura. O arranjo 3D foi desdobrado para a matriz X de dimensão 49 x 62500..... 65
FIGURA 2	Gráfico de RMSEC/RMSECV versus números de variáveis latentes..... 70
FIGURA 3	pIC ₅₀ calibrado/predito versus pIC ₅₀ experimental..... 73
FIGURA 4	Gráficos usados para os diagnósticos de <i>outliers</i> . A linha horizontal indica a região limitante..... 74
FIGURA 5	Moléculas propostas como novas drogas derivadas da

modelagem MIA-QSAR.....	75
-------------------------	----

CAPÍTULO 3

FIGURA 1	Estruturas dos inibidores de PDE-5 no conjunto de dados	85
FIGURA 2	Diagnóstico residual de Student para detecção de Outliers.....	92
FIGURA 3	Plotamento dos pIC_{50} experimental versus calibrado/predito.....	93
FIGURA 4	Inibidores PDE-5 propostos (pIC_{50} predito entre parênteses).....	94
FIGURA 5	Correlação entre energias intermoleculares (ΔE_s , kcal mol ⁻¹) e bioatividades experimentais (pIC_{50}).....	97
FIGURA 6	Estrutura do composto C docado dentro do sítio ativo da PDE-5. Os resíduos mostrados estão envolvidos em ligação de hidrogênio (Gln 817 e Tyr612) e interações π (Phe820) com o composto proposto.....	98

LISTA DE TABELAS

		Página
CAPÍTULO 1		
TABELA 1	Superfamílias das PDEs.....	22
TABELA 2	Efeitos adversos realizados em estudo comparando inibidores da PDE-5 com placebo em um grupo de pacientes (%)......	24
CAPÍTULO 2		
TABELA 1	Estrutura molecular dos compostos nos conjuntos treino e teste e valores experimentais e preditos de pIC_{50}	66
TABELA 2	Parâmetros estatísticos da modelagem MIA-QSAR.....	69
TABELA 3	Valores preditos pIC_{50} e residuais para o conjunto teste, usando os modelos MIA-QSAR e da literatura.....	72
CAPÍTULO 3		
TABELA 1	pIC_{50} experimental, calibrado e predito (LOOcv 20% out) para uma série de inibidores PDE-5.....	87
TABELA 2	Valores de energia ($kcal\ mol^{-1}$) entre PDE-5 e alguns inibidores da Figura 11.....	97
TABELA 3	pIC_{50} predito, energia de interação (E) e valores de energias de ligações de hidrogênio ($kcal\ mol^{-1}$) com a PDE-5, para Sildenafil, Tadalafil, Vardenafil e compostos propostos.....	98
TABELA 4	Resultados dos cálculos para a “regra dos cinco” de Lipinski e TPSA para os compostos de referência e os propostos.....	100
TABELA 5	Parâmetros ADME-TOX calculados para os compostos propostos e os de referência.....	102

LISTA DE ABREVIATURAS

QSAR	Relação quantitativa entre estrutura e atividade
MIA-QSAR	Relação quantitativa entre estrutura e atividade por análise de imagens. Multivariadas
QSSR	Relação quantitativa da seletividade das estruturas
MLR	Regressão linear múltipla
PLS	Mínimos quadrados parciais
CoMFA	Análise comparativa do campo molecular
CoMSIA	Análise comparativa do índice de similaridade molecular
D.E	Disfunção Erétil
PDE	Enzima fosfodiesterase
PDE-5	Enzima fosfodiesterase tipo 5
F.D.A	Entidade que regulamenta liberação de drogas e alimentos nos Estados Unidos
C_{max}	Concentração plasmática máxima
AUC	Área sobre a curva – proporção entre quantidade de droga que atinge a circulação e a depuração da droga
T_{max}	Tempo máximo para atingir o pico de concentração plasmática
CYP3A4	Citocromo P450 subtipo 3A4
DL₅₀	Dose letal 50% dos animais testados
DE₅₀	Dose efetiva 50%
VL	Variáveis latentes
LOOcv	Validação cruzada deixe uma de fora
RMSEC	Raiz quadrada da soma dos erros na calibração do modelo
RMSECV	Raiz quadrada da soma dos erros na validação cruzada
RMSEP	Raiz quadrada da soma dos erros na validação externa
2D	Bidimensional

3D Tridimensional
TPSA Área superficial topológica mede a polaridade de determinado composto, ou o tanto que o composto é disponível oralmente

RESUMO GERAL

ANTUNES, João Eustáquio. **Modelagem MIA-QSAR das atividades biológicas de inibidores de fosfodiesterase tipo-5**. 2009. 104 p. Dissertação (Mestrado em Agroquímica)-Universidade Federal de Lavras, Lavras, MG*

A aplicação de modelos teóricos para representar e manipular as estruturas de moléculas, estudar reações químicas e estabelecer relações entre a estrutura e propriedades da matéria constitui o domínio de atuação da modelagem molecular. Neste contexto, vários métodos de modelagem molecular são desenvolvidos e devem ser testados e validados quanto à sua aplicabilidade. Na primeira parte do trabalho, avaliou-se um método de modelagem molecular que relaciona quantitativamente a estrutura química com a atividade biológica, por meio de uma análise multivariada de imagem (MIA-QSAR). O modelo foi construído a partir de uma série de inibidores da fosfodiesterase tipo-5 (PDE-5), compostos úteis para o tratamento da disfunção erétil e comparado com outros métodos da literatura. A calibração e a validação do modelo foram realizadas por meio de regressão por mínimos quadrados parciais (PLS), apresentando alta correlação entre dados experimentais e valores estimados pelo modelo. Dessa forma, o modelo construído foi comprovado ser tão eficiente quanto os de outros métodos mais sofisticados, como CoMFA e CoMSIA, mostrando ser uma alternativa vantajosa, em vários aspectos, na área de QSAR. Na segunda parte do estudo, o método MIA-QSAR foi avaliado na prática para prever a potência inibitória de novos inibidores da fosfodiesterase tipo-5. Um conjunto de moléculas derivadas das estruturas do sildenafil (Viagra[®]), vardenafil (Levitra[®]) e tadalafil (Cialis[®]) foi usado na construção do modelo MIA-QSAR. Após calibração e validação do modelo construído, novos compostos foram propostos e suas atividades biológicas foram estimadas. Seis novos compostos apresentaram-se promissores, de acordo com a modelagem, exibindo elevadas capacidades inibitórias calculadas da fosfodiesterase, tanto quanto os compostos de referência (sildenafil, vardenafil e tadalafil). Um estudo de *docking* foi realizado para avaliar a afinidade entre os compostos mais potentes e a enzima PDE-5, tendo dois novos compostos sido acrescentados aos outros seis para serem submetidos ao processo de *docking*. A energia de interação entre os compostos propostos e os de referência foi bastante semelhante, mostrando boa afinidade pela enzima. Os oito compostos propostos e mais os compostos de referência foram submetidos aos estudos dos parâmetros farmacocinéticos ADME-TOX (absorção, distribuição, metabolismo, excreção e toxicidade) pelo

* Comitê Orientador: Matheus Puggina de Freitas – UFLA (Orientador)

uso das ferramentas ADME-Box e Tox-Box do software PharmaAlgorithms. Os resultados mostraram que os compostos propostos, especialmente dois deles, apresentaram parâmetros de absorção e toxicidade satisfatórios, comparáveis ou até melhores do que os fármacos de referência. O método MIA-QSAR mostrou ser tão eficiente quanto outros métodos QSAR baseados na estrutura do ligante e apresentou boa capacidade de estimar a atividade biológica de novos compostos inibidores da fosfodiesterase tipo 5. Pelo uso de outros recursos da química computacional, este estudo foi capaz de propor novos compostos, como drogas, exemplificando a eficiência da química computacional como ferramenta para a redução de tempo e de custo na elaboração de novos fármacos.

GENERAL ABSTRACT

ANTUNES, João Eustáquio. **MIA-QSAR MODELING biological activity of inhibitors of phosphodiesterase type-5**. 2009. 104 p. Dissertation (Master in Agrochemistry) -Universidade Federal de Lavras, Lavras, MG*

The use of theoretical models to represent and manipulate molecular structures, study chemical reactions and establish relationships between structure and matter properties constitutes the field of application of molecular modeling. In this context, several molecular modeling methods have been developed and should be tested and validated in regard to its applicability. In the first part of this work, a molecular modeling method, which correlates quantitatively a chemical structure with its biological activity through multivariate image analysis (MIA-QSAR), was evaluated. A model was built from a series of phosphodiesterase type-5 (PDE-5) inhibitors, useful compounds for the treatment of erectile dysfunction, and then compared with other methods of literature. Model calibration and validation was achieved through partial least squares (PLS) regression, demonstrating a high correlation between experimental and calculated data. Therefore, the model built was proved to be as efficient as the most sophisticated methods available today, such as CoMFA and CoMSIA, showing to be an advantageous alternative in many aspects in QSAR. In the second part, the MIA-QSAR method was evaluated in practice, to predict the inhibitory potency of novel phosphodiesterase type-5 inhibitors. A set of compounds derived from the structures of Sildenafil (Viagra[®]), Vardenafil (Levitra[®]) and Tadalafil (Cialis[®]) was used in the building of the MIA-QSAR model. After model calibration and validation, new compounds were proposed and their activities estimated. From this pool, six new chemical structures showed to be promising drug-like compounds, because they exhibited high calculated inhibition toward PDE-5, such as the reference compounds (Sildenafil, Vardenafil and Tadalafil). A *docking* study was carried out to evaluate the affinity of the proposed compounds with the PDE-5 enzyme, and two additional compounds were submitted to the *docking* process. The interaction energies of the proposed and reference compounds were similar, showing good affinity toward the enzyme. The eight proposed compounds, plus the reference ones, were submitted to pharmacokinetic ADME-Tox studies (Absorption, Distribution, Metabolism, Excretion and Toxicity) by using the ADME and Tox boxes of the PharmaAlgorithms program. The results showed that the predicted compounds, especially two of them, presented satisfactory

* Guidance Committee: Matheus Puggina de Freitas – UFLA (Adviser)

absorption and toxicity parameters, comparable or even better than the reference drugs. The MIA-QSAR method was shown to be as efficient as other ligand-based QSAR methods, demonstrating a great ability to estimate the biological activities of new PDE-5 inhibitors. By using different computational resources, this study was capable to propose new drugs, exemplifying the efficiency of computational chemistry as a tool to reduce time and cost during in the development of new drugs.

CAPÍTULO 1

INTRODUÇÃO GERAL E REVISÃO BIBLIOGRÁFICA

1 INTRODUÇÃO GERAL

1.1 Aspectos gerais

Químicos, biólogos, farmacêuticos e pesquisadores em geral estão sempre procurando por novas estruturas químicas que possuam grande especificidade em sistemas biológicos e que sejam capazes de ser compostos líderes. A identificação de tais ligantes é um passo crucial para se construir modelos, bem como para estudar e entender funções enzimáticas e para um eventual desenvolvimento de agentes terapêuticos. Como resposta a essa necessidade, novas metodologias baseadas em modelagem molecular e métodos capazes de manipular grande quantidade de informação utilizando ferramentas computacionais têm sido desenvolvidos, no intuito de predizer novas drogas, em detrimento de se testar experimentalmente e, às vezes, tentar predizer, intuitivamente, qual composto poderia apresentar uma atividade particular.

De acordo com a organização de produtores norte-americanos PhRMA, o custo médio do desenvolvimento de uma nova droga aumentou de US\$ 138 milhões, em 1975, para US\$ 802 milhões, em 2003. Este elevado custo, que foi revelado por um estudo realizado pela Universidade de Tufts, deve-se ao número de falhas durante o processo de desenvolvimento da droga: somente um, de cada 5.000 candidatos, chega ao mercado como um novo medicamento. Mesmo assim, não é uma garantia de sucesso, já que somente três de cada dez drogas produzidas produzem lucros que excedem os gastos do desenvolvimento.

O estudo da Universidade de Tufts também estimou que um período de 10 a 15 anos é necessário para se levar uma droga do laboratório a uma aprovação do FDA (US Food and Drug Administration) (Rees, 2003). A aplicação de métodos computacionais na predição de novas drogas ou de perfis farmacocinéticos mais adequados reduz o número de etapas durante o

desenvolvimento de fármacos, agilizando o processo e diminuindo os custos. Além disso, previsões computacionais podem permitir a introdução de um medicamento no mercado de mais maneira segura, pois elas podem revelar a toxicidade e efeitos colaterais de um candidato a fármaco.

Duas das mais reconhecidas e aplicadas metodologias em *Quantitative structure-activity relationship* (QSAR) são o *Comparative molecular field analysis* (CoMFA) (Cramer et al.; 1988; Jayatilleke et al., 2000; Singh et al., 2002; Nicolotti et al., 2002) e o *Comparative molecular similarity indices analysis* (CoMSIA) (Klebe et al., 1994). Um método igualmente capaz, baseado em regressão utilizando N-PLS e descritores gerados em um espaço tridimensional por meio do programa GRID (Goodford et al., 1995), de modo que representem interações estéricas, eletrostáticas e ligações de hidrogênio, também tem produzido notáveis resultados (Nilsson et al., 1998).

Em métodos como esses, descritores, que são parâmetros necessários para se correlacionar objetos (no caso desse projeto, as moléculas) com as variáveis dependentes (as atividades biológicas), são, na maioria, gerados em um espaço tridimensional – os formalismos 4D, 5D e 6D têm sido invocados em alguns modelos QSAR (Hopfinger et al., 1997; Vedani et al., 2002; Vedani et al., 2005) – para simular interações não covalentes entre sítios ativos e ligantes (aproximação dos ligantes). Mas, o exaustivo tratamento conformacional necessário para o sucesso de tais metodologias, juntamente com a complexidade de se alinhar os ligantes, bem como o custo para aquisição de tais softwares, encorajou pesquisadores a elaborarem um método simples e amplamente acessível para análise QSAR (Freitas et al., 2005). Este método é baseado na utilização de imagens bidimensionais para a geração de descritores (pixels) e posterior correlação dos mesmos com as variáveis dependentes, isto é, as atividades biológicas, por meio do algoritmo de regressão multivariada PLS (*Partial Least Squares*). A análise multivariada de imagens (*Multivariate Image*

Analysis ou MIA) é conhecida há algum tempo, mas sua aplicação tem se estendido a problemas químicos, principalmente na análise de dados espectrais, apesar de ela ter outras aplicações práticas em diversos campos (sensoriamento remoto, análises clínicas, etc.).

As pesquisas de Geladi e Esbensen (Geladi et al., 1988; Geladi et al., 1991; Geladi et al., 1996; Esbensen et al., 1996) ilustram bem esses casos. Tentativas promissoras para simplificar e oferecer vantagens sobre métodos tridimensionais têm surgido (Dixon et al., 2001), mas a obtenção de resultados comparáveis ainda é um problema não resolvido. O método aplicado neste trabalho, doravante denominado MIA-QSAR, no entanto, surge como nova possibilidade de oferecer alternativas para o desenvolvimento de novos fármacos, de maneira barata, rápida e fácil de manipular. Este método já foi preliminarmente testado e os resultados obtidos, por exemplo, demonstraram uma potencial capacidade para a predição de atividade biológica para uma série de moléculas com afinidade ao receptor dopamínico do tipo D2 (Freitas et al., 2005) e à enzima HIV protease (Freitas et al., 2006). Mas, uma análise mais refinada, em que a análise dos descritores possa ser mais bem explorada e parâmetros ADME-Tox (absorção, distribuição, metabolismo, excreção e toxicidade) possam ser estipulados na proposição de uma nova estrutura ativa, ainda tem de ser desenvolvida e aperfeiçoada.

Portanto, os compostos modelos a serem avaliados são uma série de inibidores do receptor fosfodiesterase tipo 5 (*phosphodiesterase type 5*, PDE-5) (Yang et al., 2006; Terret et al., 1996; Haning et al., 2002; Daugan et al., 2003), que estão relacionados à disfunção erétil (impotência) em humanos. Tais compostos são, também, de grande interesse social, visto que cerca de 20 a 30 milhões de norte-americanos sofrem de impotência (Fazio et al., 2004). Além disso, também é de interesse o desenvolvimento de tais compostos, pois poucos estudos QSAR relacionados a inibidores de PDE-5 têm sido desenvolvidos

(Yang et al., 2006) e pelo fato de os medicamentos mais importantes disponíveis no mercado (Figura 1) apresentarem sérios efeitos colaterais, como dor de cabeça, indigestão, dores nas costas e muscular e, principalmente, taquicardia.

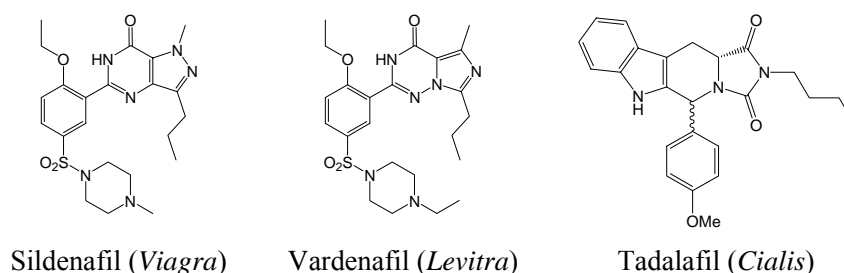


FIGURA 1. Importantes inibidores de PDE-5 disponíveis no mercado.

1.2 Problemática

A disfunção erétil, antes conhecida por impotência, é a incapacidade de se obter ou manter uma ereção adequada para a prática da relação sexual. A disfunção erétil não deve ser confundida com a falta ou diminuição no "apetite sexual", nem como dificuldade em ejacular ou em atingir o orgasmo. Esta não pode ser encontrada nas classificações internacionais de doenças com este nome genérico. Na realidade, o "Diagnostic and Statistical Manual of Mental Disorders" (DSM.IV) aborda o problema subdividindo o tema em vários tópicos. Fala-se em transtornos do desejo sexual, transtorno da excitação sexual, transtornos do orgasmo e transtornos de dor sexual. De qualquer forma, o ser humano teme muito qualquer tipo de disfunção erétil, qualquer rebaixamento em seu desempenho sexual e esse medo tem grande base cultural. Isso, provavelmente, porque a disfunção erétil sempre foi um assunto cercado de muitos tabus.

Mas, o que realmente significa a disfunção erétil? De modo geral, pode-se dizer que é uma disfunção sexual que incapacita o homem a obter ou manter a satisfação sexual. Alguns urologistas acreditam que as causas da disfunção erétil são, em 70% dos casos, originadas de problemas psicológicos assim distribuídos: 95% dos casos atingem pessoas com 20 anos, 70% aos 48 anos e 30% entre os 60 e 70 anos. Os restantes 30% dos casos seriam decorrentes de problemas orgânicos. Pesquisas realizadas nos EUA e Europa mostraram a prevalência de 30% a 56% entre os homens acima de 40 anos de idade com disfunção erétil. No Brasil, a taxa prevalece em torno de 46% dos homens na mesma faixa etária (Bestane et al., 2002).

1.3 Tratamentos utilizados

O principal neurotransmissor que facilita a função erétil é o óxido nítrico (NO). Atualmente, o tratamento da disfunção erétil pode incluir efetivos agentes orais como inibidores da fosfodiesterase (Figura 1). No passado, o tratamento consistia na injeção intracavernosa de vasoativos, como usos de vasodilatadores intrauretrais e dispositivos de constrição a vácuo (Raina et al., 2007). Este último tratamento foi descrito, primeiramente, em 1917 e consiste de um cilindro plástico limpo, uma bomba a vácuo e um anel de constrição. Depois da aplicação de um lubrificante, a parte final que abre o cilindro é colocada no pênis flácido e comprimida contra a parede do abdômen, para criar um selo hermético. A ereção é conseguida pela criação de um vácuo no interior do cilindro usando uma bomba diretamente conectada no cilindro ou conectada por tubulação. Depois que uma adequada ereção foi atingida, uma banda de constrição pode ser aplicada ao redor do pênis para ajudar manter a ereção. O dispositivo pode, então, ser removido e o paciente pode envolver-se em uma relação com a banda de constrição, mantendo a ereção. A ereção produzida por este dispositivo difere de uma ereção normal e está envolvida através da oclusão

venosa da banda de constrição, resultando num inchaço de todo o pênis, presumivelmente com preservação do influxo do sangue arterial (Raina et al., 2007).

A injeção intracavernosa de vasoativos como Alprostadil® ou tríplice mistura com Alprostadil® junto com papaverina e fentolamina promove ereção. Considerando a fentolamina um inibidor direto do adrenoreceptor, Alprostadil® e papaverina (um inibidor não específico da fosfodiesterase) (Fazio et al., 2004) modulam níveis de 3,5-ciclo-adenosina-monofosfatase nas células, eventualmente aumentando o fluxo de sangue no pênis por relaxamento da musculatura lisa nos corpos cavernosos. O uso de vasodilatadores intrauretrais, como o Minoxidil, promove uma ereção devido ao aumento do fluxo sanguíneo e, conseqüentemente, a ereção. Minoxidil é um agente anti-hipertensivo regulador do canal de potássio que produz uma significativa dilatação arterial.

Depois da introdução do primeiro agente efetivo oral para o tratamento da disfunção erétil, o citrato de sildenafil (Figura 1), em 1998, houve uma revolução no tratamento desta desordem e um aumento no número de homens presentes para avaliação e tratamento.

Sildenafil é efetivo na maioria dos homens com disfunção erétil na população em geral, incluindo aqueles com lesões na medula espinhal, *diabetes mellitus* e pacientes que tenham sido submetidos à prostatectomia radical. Novos agentes terapêuticos foram surgindo para o tratamento da disfunção erétil. Apomorfina sublingual em doses de 2 e 3 mg tem um mecanismo central de ação e, quando administrado sublingualmente, promove ereção. Há, atualmente, novos esforços para o desenho de novos inibidores da fosfodiesterase tipo-5, com aumento da potência e da seletividade, como o Tadalafil (Cialis®) e Vardenafil (Levitra®) (Raina et al., 2007).

Tratamentos alternativos aos inibidores da fosfodiesterase tipo-5 são indicados, caso a disfunção erétil seja determinada secundariamente a um estado

hipogonadal e, frequentemente, depois do uso sem sucesso de tais agentes terapêuticos, a terapia a base de testosterona exógena é indicado (Fazio et al., 2004). Atualmente, acredita-se que um metabólito da testosterona, a diidrotestosterona, aumenta a produção de óxido nítrico, produzindo o efeito clínico (Kalinchenko et al., 2003). Tratamento com segurança, sem uso de inibidores da PDE-5, pode ser conseguido pelo uso de ginseng coreano, um medicamento fitoterápico utilizado pela milenar medicina chinesa. Extratos fitoterápicos, como catuaba, marapuama e *Tribullus terrestris*, são terapias populares indicadas no tratamento da disfunção erétil, mas não há nenhuma pesquisa científica que comprovem sua efetividade (Raina et al., 2007).

Grandes avanços têm ocorrido no tratamento da disfunção erétil em relação à década passada. Muitos pacientes em tais condições são, agora, tratados por cuidados físicos primários e com outros especialistas, além dos urologistas. Uma correta avaliação clínica pode ajudar a diferenciar numerosas causas de disfunção erétil. Terapia oral com inibidores da fosfodiesterase tipo-5 (sildenafil, tadalafil e vardenafil) pode ser considerada de primeira linha para o tratamento dessa condição. Embora os inibidores da fosfodiesterase tipo-5 não possam ser administrados a pacientes em uso de nitratos, por causa do risco de hipotensão, no geral, eles são bem tolerados pela maioria dos pacientes (Fazio et al., 2004).

1.4 Métodos computacionais – vantagens

Métodos computacionais podem auxiliar, na área científica, tanto pelo acúmulo de dados (várias formas de memória) como em cálculos que processam informações em alta velocidade para propiciar respostas científicas apropriadas. Na Química não é diferente e uma das áreas que mais utilizam tais métodos é a Química Medicinal, em que é permitido criar modelos que auxiliarão na

produção de novos agentes terapêuticos e ou melhorar os já existentes, com menor tempo e custos financeiros.

Desde a introdução de descritores físico-químicos para a criação de modelos QSAR, alguns métodos de regressão têm sido aplicados para a correta linearidade experimental e a habilidade preditiva. Regressão linear múltipla (MRL) tem sido amplamente utilizada quando o número de amostras excede o montante de descritores, enquanto a regressão por PLS é mais comumente aplicada em métodos 3D-QSAR, como CoMFA (Cramer et al., 1988), em que um grande número de descritores é gerado (Freitas et al., 2008). A maioria dos métodos computacionais aplicados em QSAR, via aproximação do ligante, é baseada em aproximação 3D ou multidimensional, a qual fornece informações sobre a relação entre parâmetros físico-químicos, tais como campos estéricos e eletrostáticos, e bioatividade (Antunes et al., 2008a).

2 REFERENCIAL TEÓRICO

2.1 Inibidores de fosfodiesterase tipo 5

As fosfodiesterases (PDEs) compreendem onze famílias distintas de enzimas que catalisam a terminação de segundos-mensageiros ativos intracelulares por hidrólise da ligação fosfodiéster do monofosfato cíclico da adenosina (AMPc) ou monofosfato cíclico de guanosina (GMPc) em seus correspondentes monofosfato (AMP ou GMP). Inibidores das fosfodiesterases podem elevar os níveis intracelulares de AMPc e GMPc, ou ambos, dependendo do tipo de fosfodiesterase específica que é bloqueada.

A fosfodiesterase tipo-5 (PDE-5) é específica para GMPc e está presente em grandes concentrações em limitado número de tecidos, tais como: coração, músculo vascular sistêmico, corpo cavernoso do pênis, plaquetas e músculo liso visceral. Por estar presente nas células do músculo liso vascular é comum que, quando utilizados os inibidores da PDE-5, mesmo em doses terapêuticas para disfunção erétil, ocorra um aumento intracelular de GMPc e efeitos vasculares sistêmicos (Gupta et al., 2005). Entre as onze famílias diferentes de PDE (PDE-1 até PDE-11) identificadas e caracterizadas, PDE-5 é o alvo primário para o desenvolvimento de pequenas moléculas, como os inibidores da PDE-5, para o tratamento da disfunção erétil. O desenvolvimento de inibidores de PDE-5 para o tratamento terapêutico da disfunção erétil tem atraído grande atenção desde a introdução comercial do sildenafil (Viagra®), em 1998 e de outros dois inibidores da PDE-5 comercializados, que são o vardenafil (Levitra®) e o tadalafil (Cialis®), em 2003 (Cheng & Grande, 2007).

2.1.1 Drogas mais utilizadas

Citrato de sildenafil foi o primeiro inibidor de PDE-5 aprovado para o tratamento da disfunção erétil. Este composto é um potente inibidor da PDE-5 (IC_{50} de 3,9 nM), com alta seletividade ($PDE-n/PDE-5 > 1000$) da PDE-5 humana em relação a PDE-2, PDE-3 e PDE-4 e uma moderada seletividade ($PDE-1/PDE-5 > 80$) para PDE-1. Entretanto, sildenafil é apenas cerca de 10 vezes mais potente para PDE-5 em relação à PDE-6, a qual é encontrada nos fotorreceptores da retina humana. Esta baixa seletividade em relação à PDE-6 é presumida ser a causa de anormalidades na cor da visão, observada com altas doses ou níveis plasmáticos de sildenafil (Norwood et al., 2004).

Um estudo para avaliar a relação dose-resposta foi realizado com a utilização de doses orais de 25, 50 e 100 mg deste medicamento. De acordo com informações veiculadas nos Estados Unidos e na Europa, a dose inicial recomendada para a maioria dos pacientes é de 50 mg, devendo ser ingerida uma hora antes da relação sexual. Uma dose inicial mais baixa (25 mg) é recomendada em pacientes idosos, em sujeitos com comprometimento hepático ou falência renal (*clearance* de creatinina < 30 mL/min), ou durante administração concomitante com um potente inibidor da isoenzima CYP3A4 citocromo P-450 (CYP). Pesquisas realizadas na Europa sugerem cautela na coadministração com o inibidor da protease antirretroviral Ritonavir. Absorção e biodisponibilidade do sildenafil foram avaliadas em estudo científico. Sildenafil é relativamente lipofílico e contém uma base fraca na sua estrutura – piperazina ($pK_a = 6,5$), resultando em apenas uma ionização parcial para pH fisiológico.

De acordo com o Sistema de Classificação Farmacêutica do FDA, sildenafil é considerada droga de classe 1 (alta solubilidade/alta permeabilidade). Após a administração oral, ela é rapidamente absorvida, atingindo pico de concentração plasmática máxima dentro de uma hora (intervalo 0,5-2 horas). A biodisponibilidade oral absoluta média do sildenafil é de, aproximadamente,

38% a 41%. Estudos de absorção com carbono 14 marcado indicaram que 92% das doses administradas são realmente absorvidas; a biodisponibilidade oral relativamente baixa do sildenafil não é devido à incompleta absorção oral, mas sim ao metabolismo de primeira passagem no fígado. Apenas limitadas interações droga-alimento têm sido relatadas para este fármaco. Administração do medicamento após uma dieta com altos teores de gorduras causa redução na taxa de absorção e no grau de exposição sistêmica (Norwood et al., 2004).

A distribuição e a ligação proteica do sildenafil foram avaliadas. Ambos, sildenafil e seu principal metabólito ativo, são altamente ligados a proteínas plasmáticas ($\approx 96\%$) e a proteína ligada é independente da concentração da droga. Depois de uma administração intravenosa, o volume de distribuição médio do sildenafil é de 105 L, que excede o volume total de água do corpo (aproximadamente 42 L), indicando distribuição dentro dos tecidos e, possivelmente, ligação a proteínas extravasculares.

Em relação à eliminação, sildenafil é extensamente metabolizado; sem alteração, pode ser detectado tanto na urina quanto nas fezes. Depois de uma dose oral, metabólitos são predominantemente excretados nas fezes (73% a 88%) e uma menor parte na urina (6% a 15%). Uma depuração de 41L/h, determinada depois da administração intravenosa do medicamento em pessoas saudáveis, indica que o sildenafil pode ser classificado como uma droga de intermediária excreção hepática. A depuração, após uma administração oral em pacientes com disfunção erétil, é de 58,5L/h, com variabilidade de 29% entre os pacientes. A concentração plasmática de sildenafil e o metabólito *N*-dimetilsildenafil demonstraram diminuir exponencialmente, com uma média terminal de meia-vida de 3 para 5 horas para ambos, independente da via de administração (Norwood et al., 2004).

Cloridrato de vardenafil foi o primeiro da segunda geração de inibidores da PDE-5 que recebeu aprovação de marketing nos Estados Unidos. Vardenafil

tem alto potencial *in vitro* ($IC_{50} = 0,1$ a $0,7$ nM) e alta seletividade como inibidor da PDE-5, comparado com outras conhecidas fosfodiesterases (>15 vezes em relação à PDE-6, >130 para PDE-1, >300 para PDE-11 e >1000 para PDEs 2, 3, 4, 7, 8, 9 e 10). Ao contrário do sildenafil e tadalafil, vardenafil foi desenvolvido, desde o início, especificamente para uso como agente erectogênico (Norwood et al., 2004).

A relação entre dosagem e exposição-resposta foi avaliada para este medicamento. Doses orais de 5, 10 e 20 mg de vardenafil, administradas não mais que uma vez ao dia, demonstraram eficácia na triagem clínica. Para a maioria dos pacientes, a dose inicialmente recomendada é de 10 mg. Informações obtidas a respeito de pesquisas realizadas nos Estados Unidos mostram que a administração do fármaco deve ocorrer 60 minutos antes da relação sexual e, de acordo com a mesma informação na União Europeia, essa administração deve ocorrer de 25 a 60 minutos antes da relação sexual. A dose pode ser aumentada para, no máximo, 20mg ou diminuída para 5mg, com base na eficácia e na tolerabilidade. A dose inicial mais baixa é recomendada para homens idosos (>65 anos), em pacientes com ligeira a moderada insuficiência hepática e com severa insuficiência renal.

Na União Europeia e nos EUA, recomenda-se a utilização de doses reduzidas de vardenafil ou evitar completamente a coadministração com um forte inibidor CYP3A. Avaliações de dose-resposta limitam-se a dois pequenos estudos de fase II (fase de triagem clínica, na qual avalia-se a eficácia do medicamento num grupo de indivíduos com doença específica), comparando o efeito da dose oral de vardenafil de 10 mg *versus* 20 mg, e de 20 mg *versus* 40 mg na rigidez do pênis. Embora todas as doses tenham mostrado eficácia significativa, comparadas com o placebo, não houve nenhuma evidência de dose-dependência melhor do que a dosagem de 10mg (Norwood et al., 2004).

Vardenafil é considerada uma droga de classe 2, referente à baixa solubilidade e à alta permeabilidade de acordo com o FDA. Ela é rapidamente absorvida com concentração plasmática, sendo detectada em todos os indivíduos dentro de 8 a 15 minutos depois da administração oral. O pico da concentração plasmática foi observado no intervalo de 15 minutos até 3 horas depois da administração, com média de 42 minutos para a dosagem de 20mg e 40mg e ligeiramente depois, com 56 minutos, para a dosagem de 10mg. A biodisponibilidade absoluta para vardenafil foi descrita como, aproximadamente, 15% (95% de confiança). A determinação de apenas baixos níveis de metabólitos de vardenafil, após administração intravenosa, sugere que o efeito de primeira passagem do metabolismo hepático é o principal fator contribuinte para a baixa biodisponibilidade oral esta droga. A farmacocinética do vardenafil não é afetada pelo alimento rico em gorduras (Norwood et al., 2004).

Tadalafil é um seletivo e potente inibidor da PDE-5, com um IC_{50} de 0,94nM. Este fármaco exibe alta seletividade para PDE-5 comparado com outras PDEs: >600 vezes em relação à PDE-6, >10.000 em relação à PDE-1 até 4 e 7 até 10 e >5 vezes em relação à PDE-11.

Tadalafil é estruturalmente diferente de sildenafil e vardenafil (Figura 1) e as diferenças nas estruturas refletem no perfil farmacológico clínico desta droga. Como o sildenafil, o tadalafil foi desenvolvido inicialmente para uso em doenças cardiovasculares e foi, subsequentemente, empregado para o tratamento de disfunção erétil. Tadalafil foi o último dos três inibidores da PDE-5 introduzido no mercado farmacêutico americano, após ter sido aprovado pela FDA, em 2003 (Norwood et al., 2004).

A dose inicial recomendada para o Tadalafil varia em todos os mercados farmacêuticos do mundo. Nos Estados Unidos e na Europa, a dose recomendada para o tadalafil é de 10mg, anterior à relação sexual, com dose máxima recomendada de apenas uma única administração diária.

As doses podem ser ajustadas com base na eficácia e na tolerabilidade para mais alta (20 mg) e mais baixa (5 mg), ambas em uma única dose diária. Tadalafil é considerada droga de classe 2 (baixa solubilidade/alta permeabilidade), de acordo com o FDA. Este medicamento é rapidamente absorvido após administração oral, com tempo médio de concentração plasmática de 2 horas (intervalo de 0,6 até 6 horas). A biodisponibilidade absoluta do tadalafil após dosagem oral não tem sido relatada, mas, pelo menos 30% das doses são absorvidas a partir de uma solução oral. O tempo de absorção oral pode decorrer com êxito, se modelado por um rápido processo de primeira passagem. A absorção e as propriedades farmacodinâmicas do tadalafil não são afetadas por alimento ou álcool. Fumo e índice de massa corporal têm fraco efeito na sua farmacocinética (Norwood et al., 2004).

Tadalafil tem um volume de distribuição aparente de 60 para 70L, com variabilidade interindividual de 40% para 50%. Isso indica que ele é distribuído para os tecidos. A ligação a proteínas plasmáticas tem sido relatada como 94% com α_1 -ácido-glicoproteína e albumina como principal proteína ligante. Uma análise farmacocinética de uma população em pacientes tomando tadalafil sugere que o peso corporal influencia no volume de distribuição. Este estudo afirma que um aumento de 10kg no peso corporal aumenta 6,3% o volume de distribuição.

Tadalafil é excretado, primeiramente, como metabólito inativo, principalmente nas fezes (\approx 61% da dose oral administrada). O tempo médio de eliminação de meia-vida é de 17,5 horas e o aparente tempo médio de depuração oral é de 2,5L/h, em indivíduos saudáveis. A eliminação quase que exclusiva por metabolismo hepático e o relativo baixo valor para a depuração do medicamento por via oral indicam que este tem baixa depuração intrínseca, que se refere ao metabolismo hepático e pode ser classificada como uma droga de baixo raio de extração hepática (Norwood et al., 2004).

2.1.2 Mecanismo de ação

A ereção envolve a integração de uma função neural e vascular. Na essência, uma ereção ocorre quando o fluxo sanguíneo para o pênis ultrapassa a saída do sangue do pênis. As artérias cavernosas fornecem sangue para o corpo cavernoso do pênis (através da artéria pudental); as veias emissárias que correm através da túnica albugínea permitem a drenagem. Durante a ereção, o relaxamento da musculatura lisa trabecular resulta no aumento da entrada de sangue para os corpos cavernosos e expansão dos sinusoides presentes nesse local. Esta distensão causa compressão mecânica das veias emissárias, que dificulta sua capacidade de drenar sangue e, portanto, resulta na rigidez do pênis (Fazio et al., 2004) (Figura 2).

A via de ereção do pênis pode ser desencadeada por estimulação vaginal direta e por estimulação auditiva ou visual, que permite um aumento do fluxo sanguíneo. O pênis tem uma estrutura anatômica altamente especializada que permite um aumento maciço na entrada de sangue para ser aprisionada dentro de camadas em torno do pênis (a túnica albugínea), que causa rigidez e expansão da musculatura lisa cavernosa. O fluxo sanguíneo do pênis é controlado pelo centro autônomo de ereção que provém das contribuições simpáticas (T₁₂-L₂) e parassimpáticas (S₂-S₄) para o plexo pélvico, incluindo nervos cavernosos que inervam a artéria cavernosa e a musculatura lisa trabecular. Esses nervos são responsáveis pela disponibilização local de altas concentrações de óxido nítrico no músculo liso trabecular, que resulta no vasorrelaxamento (Fazio et al., 2004) (Figura 3).

O sinal (óxido nítrico) é liberado a partir de terminações nervosas ou de células endoteliais e ativa uma reação em cascata, que leva a um aumento na concentração de guanosina monofosfato cíclica (GMPc). Esta molécula, que é um segundo mensageiro, induz uma série de eventos que levam a um relaxamento da musculatura lisa por meio da redução na concentração do íon

cálcio intracelular. A enzima PDE-5 reverte este efeito pela metabolização do GMPc para GMP rapidamente.

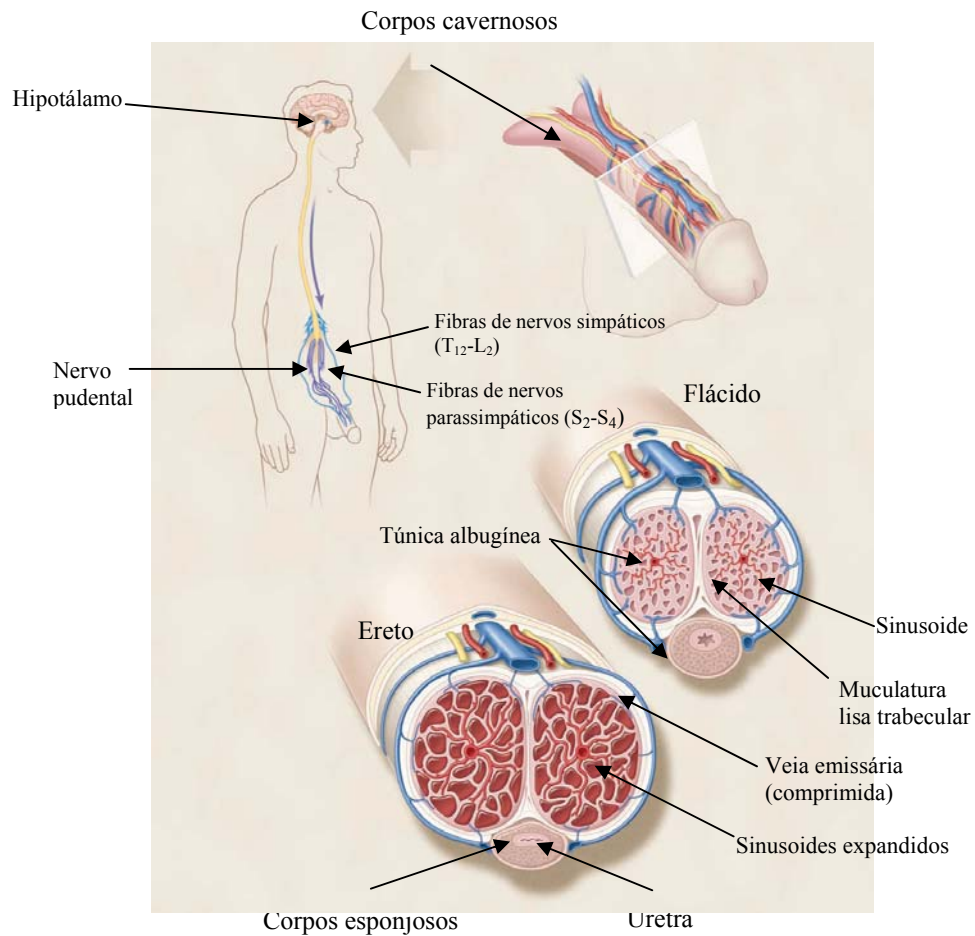


FIGURA 2. Anatomia e mecanismo da ereção do pênis. Adaptado de (Fazio et al., 2004).

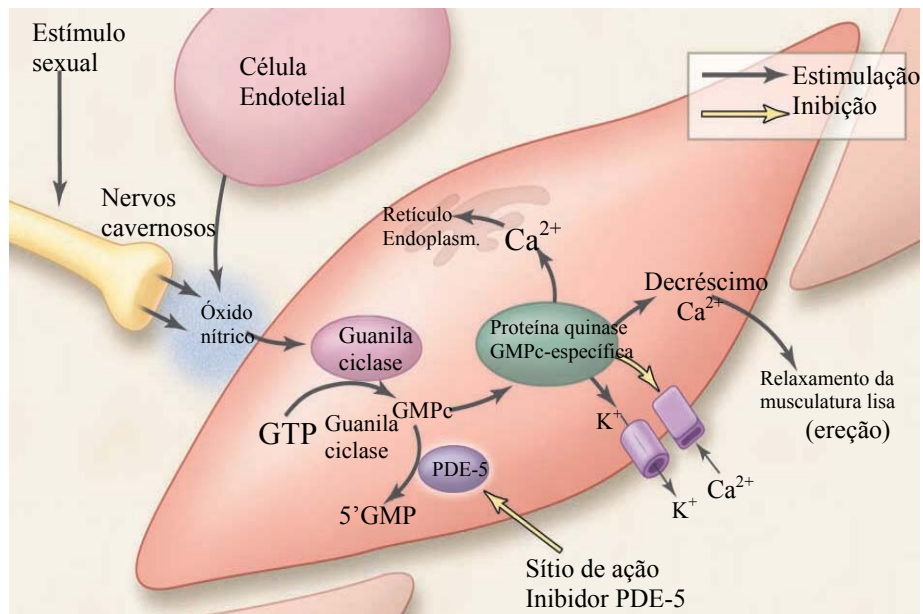


FIGURA 3. Perspectiva celular da via de ereção. Adaptado de (Fazio et al., 2004).

Os inibidores da PDE-5 (sildenafil, tadalafil e vardenafil) agem promovendo o relaxamento da musculatura lisa pela sua habilidade em permitir o acúmulo de GMPc quando óxido nítrico é liberado, como no caso de uma estimulação sexual (Fazio et al., 2004).

Óxido nítrico é difundido através da membrana do músculo liso e ativa a guanilato ciclase a produzir guanosina monofosfato cíclica (GMPc); a cascata bioquímica que resulta na alteração da permeabilidade dos canais de íons de cálcio e potássio, finalmente, leva à diminuição na concentração de cálcio citosólico, o que causa relaxamento da musculatura lisa e aumento regional do fluxo sanguíneo. Enzimas fosfodiesterases (PDEs) regulam esta via por

inativação do GMPc, convertendo este a GMP. PDE-5 é a mais importante isoenzima presente nos corpos cavernosos do pênis.

O nervo motor somático surge a partir da região sacro da medula espinhal, cujas fibras do nervo pudental aderem à inervação do músculo cavernoso, que são ativados durante a ejaculação e o clímax. Estimulação adrenérgica causa contração e intumescimento da musculatura lisa. Nervos colinérgicos podem contribuir para o processo de ereção pela inibição adrenérgica, assim como pela liberação do óxido nítrico do endotélio. Três mecanismos disparam este mecanismo vascular: psicogênico, reflexogênico e originado centralmente (ereção noturna).

Ereção psicogênica ocorre através de vias de estimulação (isto é, som, cheiro, visão e toque) que vão dos centros de ereção espinhal (T₁₁-L₂ e S₂-S₄) e induzem uma iniciação dopaminérgica de ereção vinda da área pré-óptica medial. Ereção reflexogênica é induzida diretamente pela estimulação genital, enviando mensagens ascendentes para centros de ereção central e mensagens diretas para os núcleos autônomos, que explicam atividade residual de ereção em pacientes com lesões superiores da medula espinhal. Suspeita-se que a ereção noturna, iniciada durante o sono REM (do inglês *rapid eye movement*, ou movimento rápido dos olhos) seja causada pela relativa diminuição na inibição simpática, com reforço da estimulação de centros pró-ereção. Apesar de recentes experimentações em modelos animais e voluntários humanos, informações na via central de ereção permanece sumária na melhor das hipóteses. Testosterona aumenta o interesse sexual e a frequência do ato sexual; este hormônio pode aumentar a frequência de ereção noturna, mas não altera a ereção reflexogênica ou psicogênica (Fazio et al., 2004).

O uso de terapia oral contra a disfunção erétil tem sido alvo de intensa investigação. Os inibidores da PDE-5 (sildenafil, tadalafil e vardenafil) são os principais agentes terapêuticos para o tratamento desta disfunção. PDE-5 é uma

enzima encontrada no músculo liso trabecular do pênis que catalisa a degradação de GMPc, que resulta na elevada concentração de cálcio citosólico e contração da musculatura lisa do pênis (Figura 3). Inibidores da PDE-5, portanto, bloqueiam esta via bioquímica para promover a ereção (Fazio et al., 2004).

2.1.3 Seletividade frente a outras fosfodiesterases

AMPc e GMPc são importantes mensageiros intracelulares envolvidos na transdução de diversas vias de fatores de crescimento e de estimulação fisiológica. Há ampla variedade de receptores acoplados funcionalmente à proteína G, os quais regulam a guanilato ciclase, a enzima responsável pela regeneração do AMPc. Ambos, receptor-associados e guanilato ciclase citosólica, direcionam a síntese de GMPc. Moduladores efetivos no nucleotídeo cíclico AMPc-dependente sinalizam, incluindo proteína quinase A (PKA), proteína quinase G (PKG envolvida na sinalização GMPc), AMPc regulado por fatores mutáveis (AMPc-GEF) e nucleotídeos cíclicos ativadores de canais. Há evidências sobre a funcionalidade celular que, uma vez disparada, pode mediar a sinalização comportamental do nucleotídeo cíclico. Proteína G acoplada a receptores associados à guanilato ciclase é confinada em uma discreta porção da membrana plasmática. Nucleotídeos cíclicos ativam canais, levando à sinalização do GMPc (Cheng, 2007).

O catabolismo do nucleotídeo cíclico é direcionado pelas fosfodiesterases (PDEs). Na maioria das células, a capacidade de hidrólise de nucleotídeo cíclico é amplamente superior à capacidade de síntese, indicando que os níveis intracelulares de nucleotídeo cíclico são largamente direcionados pela atividade da PDE. As superfamílias de PDE são grandes e complexas (PDE-1 até 11), contando, no total, onze famílias de genes estruturalmente relacionados e mais de sessenta isoformas distintas (Tabela 1). Cada uma das famílias de PDE contém de um a quatro genes e é responsável pela produção das

várias isoformas desta enzima. Algumas famílias de PDE hidrolisam exclusivamente AMPc (PDE-4, PDE-7 e PDE-8), exclusivamente GMPc (PDE-5, PDE-6 e PDE-9) ou ambos AMPc e GMPc (PDE-1, PDE-2, PDE-3, PDE-10 e PDE-11). Membros das famílias de PDE diferem na distribuição nos tecidos, na especificidade do inibidor e no modo de regulação. PDEs são reguladas por concentrações intracelulares de nucleotídeo cíclico, fosforilação, interação com proteínas regulatórias, compartimentalização subcelular e ligação Ca^{2+} /calmodulina, assim como também nas mudanças nas expressões gênicas (Cheng, 2007).

PDE ativa é encontrada não somente no citosol, mas também está localizada na membrana plasmática, retículo endoplasmático, membranas nucleares e citoesqueleto. Localizações intracelulares de PDEs contribuem para o processo de sinalização do nucleotídeo cíclico. Inibidores seletivos para famílias específicas de PDE têm sido empregados para demonstrar que PDEs desempenham papéis importantes no mecanismo de sinalização do nucleotídeo cíclico. Por exemplo, estudos indicam que hidrólises de AMPc em células glomerulares são, primeiramente, direcionadas pela PDE-3 e a PDE-4. Estudos mostram efeitos diferenciais na sinalização do AMPc de PDE-3 e PDE-4 (Cheng, 2007).

Inibidores de PDEs têm sido avaliados relativamente como agentes não tóxicos para vários tipos de doenças, como disfunção erétil e doenças renais, dentre outras (Tabela 1) (Cheng, 2007).

TABELA 1. Superfamílias das PDEs.

Família PDE	Genes	Substratos	Inibidores seletivos	Regulação	Aplicação experimental e clínica
PDE-1	1A, 1B, 1C	AMPc/GMPc	Vinpocetine Nicardipine	Ca ²⁺ /Calmodulina PKA/PKG	Demência, perda de memória.
PDE-2	2A	AMPc/GMPc	EHNA IC933	Estimulada pelo GMPc	Sépsis, Síndrome respiratória aguda, perda de memória.
PDE-3	3A, 3B	AMPc	Lixazinone Cilostamide Milirinone Cilostazol	Inibida pelo GMPc, fosforilada pela PKB	Glomerulonefrite, Insuficiência cardíaca congestiva, Trombose, Hipertensão pulmonar
PDE-4	4A, 4B, 4C, 4D	AMPc	Rolipram Denbufiline Cilomilast Roflumilast SCH 351591	Fosforilada pela PKA, fosforilada pela ERK (Kinase reguladas por sinais extracelulares).	Glomerulonefrite, Depressão bipolar, Encéfalomielites autoimune, Transplantes
PDE-5	5A	GMPc	Sildenafil Zaprinast Dipyridamole Vardenafil Tadalafil E4021	Ligadas a GMPc Fosforilada pela PKA Fosforilada pela PKG	Insuficiência renal crônica, Síndrome nefrítica, Disfunção erétil , Transplantes
PDE-6	6A, 6B, 6C	GMPc	Zaprinast Dipyridamole Vardenafil Tadalafil E4021	Ativada por transdução	—
PDE-7	7A, 7B	AMPc	Dipyridamole BRL 50481 IC242 BMS-586353 Thiadiazoles	Rolipram-insensível	—
PDE-8	8A, 8B	AMPc	Dipyridamole	Rolipram-insensível IBMX-insensível	—
PDE-9	9A	GMPc	Zaprinast SCH 51866	IBMX-insensível	—
PDE-10	10A	AMPc/GMPc	Dipyridamole Papaverine	Não conhecida	—
PDE-11	11A	AMPc/GMPc	Tadalafil Zaprinast Dipyridamole	Não conhecida	—

Adaptado de (Cheng, 2007).

2.1.4 Efeitos colaterais

A preocupação primária com a terapia para a disfunção erétil baseia-se nos efeitos adversos, principalmente cardiovasculares. Devido ao fato de inibidores da PDE-5 promoverem vaso-dilatação, eles, inerentemente, têm potencial para causar hipotensão. A preocupação maior tem sido em relação a pacientes idosos com doenças cardiovasculares pré-existentes. Estudos recentes demonstram que a incidência de infarto do miocárdio foi pequena numa triagem com grupos de pacientes com disfunção erétil usando sildenafil (6.884 pacientes por ano) e placebo (543 pacientes por ano), apresentando resultados de 0,8% a 1,1%, respectivamente. Portanto, praticamente não há variação entre os dois tratamentos. Resultado semelhante foi obtido quando se utilizou o tadalafil como droga no teste. Todos os inibidores da PDE-5 são contraindicados em homens que usam medicamentos à base de nitratos. Essas drogas potencializam os efeitos vasodilatadores dos nitratos, levando a um quadro de hipotensão arterial (Fazio et al., 2004).

Outros efeitos adversos associados à terapia com inibidores da PDE-5 não representam um número expressivo pela avaliação de um estudo nesse sentido (Tabela 2). Na triagem clínica, a taxa de indivíduos que não continuaram com a medicação por causa de efeito adverso foi pequena (<5%) para os três medicamentos, com pouco ou nenhum caso de priapismo (patologia associada à ereção dolorosa). Em relação à espermatogênese, não foi encontrado efeito significativo em pacientes que fizeram uso prolongado dos inibidores da PDE-5, em estudo dos hormônios e parâmetros dos espermatozoides (Fazio et al., 2004).

Fosfodiesterase tipo 6 (PDE-6) é encontrada na retina e é responsável pela fototransdução. Sildenafil é um potente inibidor da PDE-6 e, por isso, são relatados efeitos relacionados a perturbações visuais (isto é, brilho e embaçamento da visão). Tadalafil tem uma especificidade maior para PDE-5 do que PDE-6, que faz reduzir o potencial de toxicidade ocular (Fazio et al., 2004).

TABELA 2. Efeitos adversos detectados em estudo comparando inibidores da PDE-5 com placebo, em um grupo de pacientes (%).

Efeito adverso	Sildenafil (25-100mg)	Placebo	Vardenafil (5-20mg)	Placebo	Tadalafil (2.5-20mg)	Placebo
Dor de cabeça	14-30	4-6	6,8-15,3	3,9	7-21	6
Dispepsia	2-16	1-2	0,7-6,7	0	1-17	2
Rinite ou congestão nasal	1-11	2	2,8-7,3	3,3	4-6	4
Náuseas	1-6	1	NR	NR	NR	NR
Tonturas	NR	NR	NR	NR	NR	NR
Perturbações visuais	0-11	≤ 2	< 1-2	0	0.1	0
Dor nas costas	0-6	2	NR	NR	3-9	5
Mialgia	NR	NR	NR	NR	1-7	2
Bocejo	NR	NR	NR	NR	NR	NR
Sonolência	NR	NR	NR	NR	NR	NR
Faringite	NR	NR	NR	NR	NR	NR

OBS: NR (não relatado).

Adaptado de (Fazio et al., 2004).

2.2 Métodos de modelagem

2.2.1 Química computacional

Dentre outras aplicações da química computacional, o *design* de novas drogas tem se tornado uma ferramenta indispensável na aplicação prática de modelagem molecular para a pesquisa e o desenvolvimento de novos fármacos.

Novas técnicas para exploração de estruturas proteicas ou de dados referentes a conhecidos ligantes a alvos particulares podem ser usadas para novas predições.

A química computacional tem o importante papel de reduzir o trabalho analítico para pesquisa de novas moléculas; bastante sucesso de sua aplicação na indústria farmacêutica para escolha de compostos candidatos a novas drogas tem ocorrido, pois tais compostos são previamente desenhados e avaliados antes de serem sintetizados, reduzindo-se, assim, o custo e o tempo para a produção de um novo medicamento. A química computacional pode também complementar estudos experimentais sobre mecanismos de toxicidade. Bioquímicos estão sempre utilizando técnicas de modelagem molecular para ter acesso às interações em âmbito molecular e prever quais candidatos são mais ou menos tóxicos.

Os estudos QSAR (relação quantitativa entre estrutura e atividade) têm contribuído para descobertas no campo de toxicidade e oncogenicidade, assim como na seleção de compostos mais potentes. Em QSAR, em vez de serem utilizados apenas parâmetros estruturais, outras propriedades moleculares podem ser empregadas. Por exemplo, um grupo funcional particular pode ser um parâmetro estrutural que pode ser usado em predição de toxicidade, mas outros parâmetros, tais como potencial de ionização, orbitais HOMO-LUMO, pixel de uma imagem (Antunes et al., 2008b) e solubilidade aquosa são também importantes. A química computacional também pode auxiliar na predição de intermediários metabólicos via modelagem molecular, por meio do modelo de biodegradação ou reações de contaminantes biologicamente catalisados por redox. Uma vez testado contra os bem conhecidos sistemas de precisão, cálculos rápidos e baratos podem ser realizados para ajudar a prever qual metabólito intermediário pode ser formado em maior quantidade e como isso afetaria o meio ambiente.

Mais uma interessante ferramenta disponível que a química computacional tem fornecido é o importante papel da compreensão da sorção (são interpretações teóricas de como as moléculas empilham-se e fixam-se em superfície sólidas) (Dobrynski et al., 2007). Assim, tem-se compreendido melhor como poluentes interagem com materiais naturais, tais como minerais, sendo importante no uso em prol de estudos ambientais (Kubicki; 2005).

A química computacional também está presente na Biologia Molecular, por meio do emprego de técnicas de bioinformática. Métodos computacionais permitem interpretar e calcular equações que são aplicadas para simplificar modelos de expressão de genes e para compreender vias de sinalizações como vistos em modelos como *Escherichia coli* (Dobrzynski et al., 2007).

Portanto, a química computacional pode ser uma importante ferramenta para a ciência em geral. Dados termodinâmicos podem ser confirmados ou testados em muitas instâncias por métodos teóricos. Propriedades moleculares, tais como log Kw e orbitais HOMO e LUMO, podem ser computadas para ajudar a prever comportamento macroscópico dos compostos. Além disso, essa ciência pode ser empregada para ajudar a interpretar observações experimentais, assim como extrapolar resultados experimentais em sistemas nos quais experimentos são problemáticos por razões práticas (Kubicki; 2005).

2.2.2 Aproximação do receptor (estudos de *docking*)

Novas técnicas utilizadas na modelagem molecular têm surgido. Aproximação do receptor, conhecida por estudo de *docking*, é um recurso computacional importante para realizar comparações entre moléculas análogas e o sítio ativo da enzima de interesse. O modo de interação das moléculas com essas enzimas permite a opção de seletividade de compostos análogos frente a uma enzima alvo. A docagem e a dinâmica molecular dos sistemas em *docking* são realizadas por uma variedade de programas computacionais, como, por

exemplo, MVD (Molegro Virtual Docker) (Thomsen & Christensen, 2006) e SOFTDOCK (Jiang et al., 2002).

Funções biológicas de macromoléculas dependem de um preciso reconhecimento molecular. Processos de reconhecimento molecular são, muitas vezes, regulados e controlados de maneira precisa e, por sua vez, regula processos e vias biológicas nas células. A interação entre duas moléculas que se reconhecem pode formar um complexo molecular, vindo das energias ligantes da catálise enzimática, da inibição ou da ativação na enzima que causa uma mudança na molécula efetora e transmite um sinal através da membrana e assim por diante.

O mecanismo molecular tem sido extensamente estudado, experimental e teoricamente. Recentemente, tem havido muito incentivo para o progresso nos estudos teóricos, especialmente na área de reconhecimento molecular entre macromoléculas. Com o advento de mais cristais de estruturas de complexos moleculares, caracterização detalhada de reconhecimento de proteína tem sido calculada. Tem-se observado que a estabilização de proteína associada se deve a vários fatores, como o tamanho e o caráter químico das superfícies de proteínas, a complementaridade e a forma das superfícies presentes nas proteínas, e interações polares por meio de ligações de hidrogênio, iônicas e ou moléculas de água. O principal método de estudo destas grandezas é o chamado estudo de método de *docking* molecular, que se desenvolveu rapidamente na última década.

O método baseia-se no uso de coordenadas atômicas para representar uma molécula em *docking* com outras moléculas, utilizando campos de força atômicos simulados computacionalmente. Várias representações da superfície molecular e volume têm sido desenhadas. O método está, muitas vezes, relacionado à pesquisa de algoritmos de soluções espaciais. A solução espacial é definida pela rotação relativa e translação entre as moléculas em *docking* e

também outros graus de liberdade, incluindo a flexibilidade conformacional do ligante e do receptor.

Kunts et al. (1982) foram os primeiros a criar uma superfície molecular de uma molécula e, então, representá-la com um conjunto de esferas como imagens positivas ou negativas das superfícies. A imagem positiva do ligante foi então colocada em *docking* para a imagem negativa do receptor. A imagem negativa do receptor pode também ser usada para igualar com as coordenadas atômicas de pequenos ligantes de uma conhecida base de dados. Eficientes pesquisas e correspondentes algoritmos têm sido desenvolvidos por *docking* macromolecular com incorporação de sobreposição de volume. Formas e descritores químicos de superfície têm sido desenvolvidos com estas representações básicas (Meng et al., 1992).

Jiang & Kim (1991) utilizaram a superfície molecular de uma molécula para representar a superfície normal e os pontos que representam onde elas seriam atacadas. A superfície e o volume de uma molécula são, então, digitados em grades chamadas cubos de superfície e cubos de volume. Avaliando os dois conjuntos de cubos de superfícies moleculares e cubos de volumes são então achados através de exaustivas amostragens do espaço rotacional e translacional, com um rápido algoritmo de translação. A fim de acomodar certa flexibilidade conformacional, programas computacionais com método de *docking* são utilizados para avaliar e gerar soluções. Este recurso pode permitir (1) variação do tamanho do cubo e (2) cálculos dos ângulos no local da superfície que corresponde à complementaridade. A vantagem da representação desses cubos é captar as características mais importantes da superfície (utilizando pontos na superfície com áreas) e os detalhes da representação são variáveis de acordo com o tamanho dos cubos.

Outra forma de representação da superfície molecular é a utilização de descritores em forma quadrática (Goldman & Wipke 2000); (Free et al., 1964).

Os autores sugerem que uma forma explícita de algoritmo para *docking*, isto é, quando a forma de informação é utilizada explicitamente na geração de soluções, pode ser mais eficiente do que as outras formas de algoritmos implícitas já vistas.

Segundo Jiang et al. (2002), estudos indicam que muitos dos algoritmos de *docking* têm suas lacunas e não podem ser trabalhados universalmente em cada uma das moléculas complexadas. Entretanto, há necessidade de realizarem-se estudos sobre quando os algoritmos são efetivos e quando eles não são. Jiang et al. (2003) realizaram estudos com um grande número de complexos cristalinos, que são avaliados no banco de dados de proteínas (PDB) (Hobohm et al., 1991); (Hobohm & Sander, 1994).

Métodos de *docking* podem ser utilizados para recriar cristalograficamente a estrutura secundária de proteínas conhecidas, utilizando comparações geométricas. O resultado do *docking* pode revelar uma significativa percentagem da correta predição conformacional da proteína. Neste estudo, é utilizado o banco de dados de proteína (PDB) e o programa GRAMM¹ (Jiang et al., 2002) que iguala elementos da estrutura secundária com as bases de dados. A vantagem disso é permitir uma melhor compreensão dos princípios da estrutura de proteína que pode facilitar o desenvolvimento de melhores métodos para a predição das estruturas de proteínas (Jiang et al., 2002).

Estudos de *docking* são bem utilizados na predição de bioatividade de moléculas com potencial terapêutico para corroborar como ferramenta útil para este potencial. Compostos de referência e novas estruturas promissoras com fins terapêuticos podem ser ancorados no interior de sítio ativo de enzimas de interesse e cálculos poderão ser realizados por meio de programas computacionais para simular interação entre resíduos de aminoácido da enzima com átomos centrais dos compostos, como por meio de ligações de hidrogênio

¹ (<http://reco3.ams.sunysb.edu/gramm>)

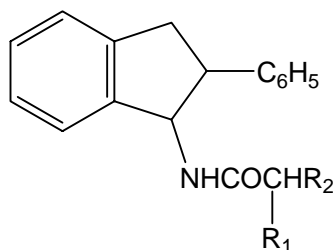
entre eles. São utilizados neste tipo de *docking* o banco de dados de proteína (PDB) e o programa específico para docking Molegro Virtual Docker (MVD), um programa para a predição da conformação mais provável como um ligante que se ligaria com uma macromolécula (Thomsen & Christensen, 2006; Antunes et al., 2008b).

2.2.3 Aproximação do ligante

a) Free-Wilson

Uma grande contribuição para o estudo da relação estrutura-atividade foi atribuída aos pesquisadores Spencer M. Free e James W. Wilson (Free & Wilson, 1964). Tais pesquisadores desenvolveram uma técnica matemática que sugere meios de descrever a relação entre estrutura e atividade de uma série de análogos químicos. Modelos matemáticos foram desenvolvidos para relacionar estrutura e atividade para um número de análogos por meio de duas mudanças diferentes para um simples carbono na molécula.

Por exemplo,



Fonte: Free & Wilson (1964).

em que R_1 é H ou CH_3 e R_2 é $\text{N}(\text{CH}_3)_2$ ou $\text{N}(\text{C}_2\text{H}_5)_2$.

Estes são compostos analgésicos cuja resposta biológica de interesse é a DL_{50} . Obtiveram-se os seguintes resultados de DL_{50} :

R_2	R_1		média
	H	CH_3	
$N(CH_3)_2$	2,13	1,64	1,885
$N(C_2H_5)_2$	1,28	0,85	1,065
	-----	-----	-----
média	1,705	1,245	1,475

O exemplo acima mostrou que o análogo com $R_1 = H$ e $R_2 = N(CH_3)_2$ tem uma DL_{50} de 2,13mg/10g e é, provavelmente, o melhor dos quatro. A contribuição de cada substituinte é fácil de ser determinada para esses dados. Comparando-se a média para o substituinte R_1 , poderia ser expressa a contribuição do H na posição R_1 como $1,705 - 1,475 = +0,23$ e a contribuição do CH_3 na posição R_1 como $1,245 - 1,475 = -0,23$. Neste caso, a simetria (-0,23 e +0,23) contribui para a solução. Igualmente para o substituinte R_2 , em que a contribuição média para o lado $N(CH_3)_2$ é +0,41 e, para o outro lado, é de -0,41. Este exemplo mostra a propriedade aditiva dos análogos. O modelo matemático proposto para esta situação é o seguinte:

$$\text{Resposta} = \text{média} + \text{efeito substituinte } R_1 + \text{efeito substituinte } R_2$$

$$DL_{50} = \mu + a[H] + a[CH_3] + b[N(CH_3)_2] + b[N(C_2H_5)_2] \quad (\text{Free \& Wilson, 1964}).$$

em que: $\mu =$ média total

$a[H] =$ contribuição do substituinte H para a posição R_1

$a[CH_3] =$ contribuição do substituinte CH_3 para a posição R_1

$b[N(CH_3)_2] =$ contribuição do substituinte $N(CH_3)_2$ para a posição R_2

$b[\text{N}(\text{C}_2\text{H}_5)_2]$ = contribuição do substituinte $\text{N}(\text{C}_2\text{H}_5)_2$ para a posição R_2

Substituindo-se os valores de DL_{50} na equação acima, foram produzidas quatro equações com cinco incógnitas. Entretanto, o exemplo mostra:

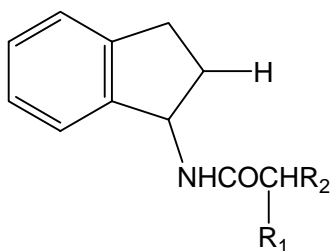
$$a[\text{H}] + a[\text{CH}_3] = 0 \text{ e } b[\text{N}(\text{CH}_3)_2] + b[\text{N}(\text{C}_2\text{H}_5)_2] = 0 \text{ (Free \& Wilson, 1964).}$$

ou $a[\text{H}] = -a[\text{CH}_3]$; $b[\text{N}(\text{CH}_3)_2] = -b[\text{N}(\text{C}_2\text{H}_5)_2]$ (Free & Wilson, 1964).

Assim, utilizando-se a igualdade acima, torna-se necessário determinar apenas os valores para μ , $a[\text{H}]$ e $b[\text{N}(\text{CH}_3)_2]$. Isso reduz o problema matemático para a solução de quatro equações e três incógnitas. Assim, fica fácil resolver por mínimos quadrados (Free & Wilson, 1964).

Por meio do que foi demonstrado, parece ser fácil solucionar o problema de quatro equações e três incógnitas, se for possível obter dados para três de quatro compostos.

Por exemplo,



Fonte. Free & Wilson (1964).

em que R_2 é $\text{N}(\text{CH}_3)_2$ ou $\text{N}(\text{C}_2\text{H}_5)_2$ e R_1 é H ou CH_3 . Os dados de DL_{50} são os seguintes:

R ₂	R ₁	
	H	CH ₃
N(CH ₃) ₂	2,75	?
N(C ₂ H ₅) ₂	1,90	1,55

Com base no modelo aditivo, era previsto estimar um valor acima de 2,00 para o exemplo acima. Porém, Takahashi et al. (1962) atribuíram um valor de 1,77, mostrando um problema existente no modelo. A dúvida estava sendo avaliada em relação à variação biológica, ou não aditividade, ou ambos. À medida que aumenta a variação em R₁ e R₂, mais problemas na prática seriam encontrados.

Algumas interpretações de estrutura-atividade podem conter dados perdidos, assim como também variação nos dados avaliados. A escolha de um melhor composto pode ser difícil e a estimativa do efeito de cada substituinte era frequentemente ignorada. Por saber que alguma atividade biológica dos compostos se deve à aditividade, um modelo matemático pode ser escrito para ajudar a descrever esta atividade. Se há m variações em R₁ e n em R₂, há $m \times n$ diferentes compostos de interesse. O modelo requer dados para pelo menos $m + n - 1$ compostos. Substituintes adicionais para outras posições poderiam complicar o julgamento da estrutura-atividade.

Nem todas as posições devem ser restritas para atividade. Por isolar a posição aditiva, a solução poderá sugerir onde o efeito não aditivo pode ser não confiável. Novos exemplos foram demonstrados e, por meio do modelo proposto pela aditividade e resoluções de equações e incógnitas, podem-se estimar valores de atividades biológicas e ainda comparar com valores observados. Houve uma diferença entre valores observados e estimados para a atividade biológica, mas

isso deve ser explicado pela variação de procedimentos *in vitro* e ou um componente não aditivo na série de análogos (Hasegawa et al., 1999).

Este modelo matemático aproximado é sugerido como uma via para sumarizar dados gerados por uma série de análogos químicos. Esta aproximação quantitativa pode também servir como meio de transferir informações de uma série para outra. Este esquema mostrou não ser visto como um meio para identificar um único composto. O modelo também mostrou ser confiável para uma combinação de vários substituintes que produz resposta biológica para um excesso de estimativas aditivas. Tais resultados apareceriam em algumas séries de análogos e esta situação poderia ser justificada pelas diferenças das respostas estimadas menos resposta observada para todos os compostos. Finalmente, conclui-se que o modelo matemático não leva em consideração a tridimensionalidade de compostos, pH, pKa ou outras propriedades físicas similares (Hasegawa et al., 1999).

b) Descritores 2D clássicos estéricos, eletrostáticos (Hammett), lipofilicidade (Hansch)

Hammett (1940) foi o primeiro a reconhecer que uma reação química poderia ser tomada como um processo normal para propor uma equação (equação 1) para descrever a ionização do ácido benzoico em H₂O. Assim, efeitos estéricos não são colocados como papéis importantes nesta reação; a constante de substituinte σ foi considerada para representar alguns tipos de efeitos elétricos ou polares quando se substituiu -H por alguns substituintes nas posições *orto*, *meta* e *para* do anel benzênico (Ferreira, 2002).

$$\text{Log } K = \log K_0 + b_{1\sigma} \quad (1)$$

Nesta equação, $\log K_o$ e $\log K$ são constantes de dissociação dos ácidos benzoicos não substituídos e substituídos, respectivamente; σ é um parâmetro eletrônico de Hammett e b_1 é a constante de proporcionalidade. Hansch (Charton, 1999) estendeu a equação de Hammett em termos de propriedades eletrônicas dos substituintes como:

$$\text{Log } 1/C = b_0 + b_1\sigma + b_2 \log P \quad (2)$$

em que $\log 1/C$ na equação 2 representa a atividade biológica definida como logaritmo de uma recíproca da dose efetiva ($C = ED$) ou concentração inibitória ($C = IC$) que é necessária para produzir um certo efeito biológico, e P é o coeficiente de partição octanol/água. Os coeficientes b_0 , b_1 e b_2 podem ser determinados pelos recursos de métodos de regressão linear múltipla (MLR). Na forma de matriz, a equação de Hansch e as soluções por mínimos quadrados para estimar o vetor regressão \mathbf{b} pode ser escrita como:

$$\begin{bmatrix} \log \frac{1}{C} \end{bmatrix} = \begin{bmatrix} 1 & - & - \\ \vdots & \vdots & \vdots \\ 1 & - & - \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} \mathbf{e} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} b_0 + \begin{bmatrix} - \\ \vdots \\ - \end{bmatrix} b_1 + \begin{bmatrix} - \\ \vdots \\ - \end{bmatrix} b_2 + \begin{bmatrix} \mathbf{e} \end{bmatrix} \quad (3)$$

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e} = b_0 + b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \mathbf{e} \quad \Rightarrow \quad \hat{\mathbf{b}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}.$$

em que \mathbf{y} e \mathbf{e} são, respectivamente, o vetor da atividade biológica (propriedade físico-química, reatividade, etc.) e resíduos entre atividades medidas e estimadas. Os vetores \mathbf{x}_1 e \mathbf{x}_2 contêm, respectivamente, os valores de σ e $\log P$ para cada composto. \mathbf{X} é a matriz de variáveis independentes com uma coluna justaposta a sobrescritos \mathbf{T} e -1 indica ser a matriz inversa. A equação de Hansch estabeleceu um avanço na área de QSAR (Ferreira, 2002).

A vantagem desta equação é que ocorre para ambas as interações eletrônicas e hidrofóbicas. O pressuposto fundamental nesta abordagem são as contribuições adicionais para σ e $\log P$ de cada grupo de substituintes na mesma molécula, que significa que os substituintes são independentes (o efeito de um grupo substituinte para uma posição na molécula não afeta a substituição na outra posição). É conhecido que esse pressuposto é violado, por exemplo, quando ligações de hidrogênio ocorrem entre dois grupos adjacentes (Charton, 1999).

Entretanto, muito frequentemente, o conjunto de dados não se encaixa bem para este modelo linear e, por isso, o método foi estendido pela adição de termos não lineares e parâmetros experimentais/calculados, dentre eles os seguintes: i) várias propriedades da química quântica relatada para a distribuição de elétron na molécula (momentos dipolos, polarizabilidades, carga atômica, energia de orbital de fronteira, calor de formação, etc.); ii) parâmetros estéricos (volume molecular, área superficial, etc.); iii) peso molecular, refratividade molar; iv) índices de conectividades (topológicos) dados por um número simples que expressam como os átomos são arranjados na molécula e v) indicadores de variabilidades, ocorrendo devido à ausência ou à presença de funções químicas específicas (Charton, 1999).

Depois de incluir todos esses descritores, a próxima questão para ser respondida é: quais, entre todos esses parâmetros e funções, são realmente relatados para a atividade biológica e poderia ser incluída na equação de regressão? Para regressão linear múltipla (MLR), um ponto pode ser observado: as chances são altas para que altas correlações possam aparecer entre descritores e isso pode causar falsas soluções. Este problema pode ser resolvido de duas maneiras: fazendo uma seleção de variáveis e, então, empregando-se a regressão linear múltipla, como na equação de Hansch, ou pelo uso de técnicas

multivariadas baseadas em variáveis latentes, tais como PLS ou PCR (Freitas, 2008).

Outro ponto crucial a ser considerado em quaisquer dados é uma análise prévia. A matriz de dados originais, usualmente, não tem ótima distribuição de valores para as análises (i.e. tem diferentes unidades e variâncias nas variáveis), que requer algum tratamento prévio antes da análise de dados (Freitas, 2008).

c) Introdução aos descritores 3D e métodos de regressão

O algoritmo multilinear PLS (N-PLS) foi implementado como método de regressão em 3D-QSAR (Nilsson et al., 1998), utilizando estudo de um grupo de compostos com afinidade para receptores de subtipo D₂ para dopamina. Este estudo foi utilizado para validação do método multilinear PLS. Exaustivas análises conformacionais nos ligantes foram efetuadas para selecionar o análogo ativo que seria empregado para alinhá-los farmacologicamente em sua presumida conformação ativa. Foram utilizados descritores 3D gerados pelo programa GRID (Goodford, 1995). Neste trabalho, demonstrou-se que a introdução de um método multilinear PLS (N-PLS), como método de regressão em 3D-QSAR, apresentou várias vantagens, quando comparado ao método bilinear PLS, isto é, o tradicional PLS. Foi demonstrado que o multilinear PLS é mais estável, aumenta a habilidade preditiva e melhora a interpretação dos resultados (Nilsson et al., 1998).

d) Descritores 4D

Análises 4D-QSAR incorporam a liberdade do alinhamento desenvolvido em modelos 3D-QSAR, considerando uma média entre as conformações dos ligantes. Os descritores em análise 4D são as medidas das ocupâncias dos átomos de cada molécula numa grade espacial (*grid cell*),

considerando a conformação e o alinhamento dos ligantes (Hopfinger et al., 1997).

e) Descritores 5D e 6D

Com o auxílio de um novo software denominado QUASAR (Vedani et al., 2002), adiciona-se grau de liberdade – a quinta dimensão – pela representação múltipla da topologia do ligante. Foi demonstrado que o método permite um ajuste induzido da topologia do ligante ao sítio ativo do receptor. A diferença para o 6D-QSAR é que, neste último, podem-se obter considerações simultâneas de diferentes modelos de solvatação. Isto pode ser obtido mapeando-se partes da área superficial do ligante com propriedades do solvente, ou, então, fazê-lo implicitamente (Vedani et al., 2005).

f) CoMFA

Análises comparativas do campo ou domínio molecular CoMFA são uma importante ferramenta para correlacionar estrutura–atividade. Os principais recursos característicos para sua aplicação são: (1) representação das moléculas dos ligantes por seus campos estéricos e eletrostáticos, sendo amostrados pela interseção de um parâmetro tridimensional, (2) uma nova técnica de “ajuste de campo”, permitindo um alinhamento ótimo mútuo dentro de uma série de compostos, pela minimização das raízes quadradas médias das diferenças entre moléculas e os campos, (3) análise de dados por PLS (mínimos quadrados parciais), utilizando validação cruzada para maximizar a probabilidade de que os resultados tenham validade preditiva e (4) representação gráfica dos resultados (Cramer et al., 1988).

Duas observações devem ser ressaltadas, no que se relaciona ao método de campo molecular: 1) a interação que produz um observado efeito biológico é usualmente não covalente e 2) a maioria dos campos de força moleculares trata

interações não covalentes (não ligadas) apenas como efeitos estéricos e eletrostáticos e podem se correlacionar precisamente para uma grande variedade de propriedades moleculares observadas. Assim, é bem razoável que uma capacidade amostral de campos estéricos e eletrostáticos ao redor de um conjunto de ligantes (drogas) possa conter todas as informações necessárias para compreender suas propriedades biológicas experimentais. Entretanto, metodologias aplicadas nesse sentido, como CoMFA, tiveram auxílio de métodos de análise de dados, como o PLS, que podem derivar robustas equações lineares (Cramer et al., 1988).

O sucesso esperado pelo uso desta metodologia se deve à alta relevância dos descritores utilizados para resolver problemas na correlação entre estruturas e atividades. Nesse sentido, é importante considerar a natureza físico-química dos descritores. Efeitos estéricos e eletrostáticos são primariamente, mas, muitas vezes, as únicas interações não covalentes nos mecanismos de forças moleculares dos domínios. O importante papel na compreensão em relação à quantificação da ligação ligante-enzima pelo uso de tais campos de forças num tratamento perturbacional é consistente com as hipóteses subjacentes do CoMFA, em que uma amostragem adequada de interações de um ligante poderia ser suficiente para responder a mais questões sobre essas possíveis interações com o receptor.

Outro ponto relevante é considerar a natureza tridimensional de uma sequência de descritores do CoMFA. O parâmetro “análise da forma” utilizado em métodos 3D-QSAR anteriores ao CoMFA são, na verdade, alguns índices que descrevem “forma” apenas na mesma medida, por exemplo, que a “forma” da estrutura é descrita pela medida diferencial do peso ou volume. Porém, segundo proposto por Cramer et al. (1988), a interação droga/receptor deve ser universalmente considerada depender principalmente da forma de

complementaridade e os descritores devem realmente diferenciar detalhes da topografia da droga.

g) CoMSIA

O *Comparative molecular similarity indeces analysis*, ou CoMSIA, é uma metodologia 3D alternativa para computar propriedades de moléculas, baseando-se em índices de similaridade dos ligantes moléculas (drogas) por meio de um alinhamento comum. Os campos de diferentes propriedades físico-químicas utilizam uma dependência da distância tipo gaussiana e nenhuma singularidade ocorre nas posições dos átomos. Conseqüentemente, nenhuma das definições de limites de corte arbitrários e insuficiências devido a diferenças nas encostas dos campos são encontradas. Assim como no CoMFA, os campos são avaliados pelas análises de PLS.

No primeiro estudo utilizando esta metodologia (Klebe et al., 1994), dois conjuntos de dados de esteroides ligados à globulina ligante de corticosteroide e inibidores termolisinas foram analisados e os resultados foram comparados com os do método convencional CoMFA. Modelos de significância estatística comparável foram obtidos. Mapas de contribuição dos campos foram produzidos para os diferentes modelos. Devido aos cortes nas configurações do CoMFA e os difíceis passos para cálculos das potencialidades perto da superfície molecular, os mapas do CoMFA são frequentemente bastante fragmentados e não continuamente conectados. Isto torna suas interpretações mais difíceis.

Os mapas obtidos pelo método CoMSIA são superiores e mais fáceis de interpretar. Enquanto os mapas do CoMFA denotam regiões da molécula nas quais possíveis interações com o ambiente são esperadas, os mapas do CoMSIA realçam estas regiões dentro da área ocupada pelos esqueletos dos ligantes que requerem alguma propriedade físico-química particular importante para a

atividade. De acordo com Klebe et al. (1994), isso é mais significativo para traçar recursos que realmente importam, no que diz respeito ao *design* de novos compostos.

h) Detalhamento do método MIA-QSAR

Desde a introdução de descritores físico-químicos úteis em modelos QSAR, alguns métodos de regressão têm sido aplicados para correlacionar linearmente tais descritores com as respectivas bioatividades experimentais. Regressão linear múltipla (MLR) tem sido pouco utilizada quando o número de amostras (linhas) excede a quantidade de descritores (colunas), enquanto PLS é o método de regressão mais comumente aplicado em 3D-QSAR, como, por exemplo, CoMFA e outros métodos relatados, nos quais um maior número de descritores é gerado.

O método *Multivariate image analysis applied to QSAR*, ou MIA-QSAR (Freitas et al., 2005), é um caso especial (não único) em que o tratamento com os descritores (pixels) para cada composto ativo resulta num arranjo tridimensional, depois de agrupar amostras para dar um conjunto de dados. Tais arranjos podem ser tratados pelo uso de métodos multimodos, tais como PLS multilinear (N-PLS) e análise de fatores paralelos (PARAFAC). Esses métodos não têm sido muito explorados em estudos QSAR, apesar de suas supostas vantagens sobre outros métodos mais bem estabelecidos (Ferreira, 2002).

Modelos matemáticos têm sido utilizados para correlacionar estrutura química com propriedades/atividades biológicas. Usualmente, é desejável que descritores possuam significado físico-químico, a fim de melhor compreender a maneira pela a qual ligantes e receptores estão interagindo. Os descritores MIA têm mostrado conter informações químicas e podem ser atribuídos aos pixels de imagens em 2D, i.e. são números binários.

Imagens representam uma rica fonte de informações (informações espaciais) e podem ser usadas numa ampla variedade de aplicações, tais como na medicina, na química, no sensoriamento remoto, na agricultura e no controle qualitativo de alimentos e de produtos farmacêuticos. Quando uma série de imagens 2D é produzida e essas imagens são sobrepostas, uma tabela de dados em 3D é construída. Tal tabela tem enorme quantidade de dados altamente correlacionados. Na análise multivariada de imagens congruentes (MACI), cada imagem é representada por um perfil com informações quantitativas. Inicialmente, o desdobramento da tabela de dados em 3D para uma estrutura de dados em 2D é feito para facilitar a análise de dados (Freitas et al., 2008).

No caso do método MIA-QSAR, as imagens 2D são as próprias estruturas químicas bidimensionais construídas por meio de alguns softwares apropriados para desenhos químicos. Quando uma série de compostos é considerada, a porção variável das estruturas (os substituintes) pode ser tratada multivariadamente, a fim de estabelecer uma correlação entre estrutura química e bioatividades, isto é, os descritores MIA podem ser utilizados para realizar com sucesso previsões, o que é valioso para aqueles que estão projetando e sintetizando novos compostos ativos (Freitas et al., 2008).

A metodologia MIA-QSAR tem mostrado ser um método com algumas vantagens em vários aspectos, tendo o potencial de ser incorporada dentro de softwares que podem ser usados no dia-a-dia, tanto pelo usuário das ciências moleculares em trabalho computacional quanto experimental (Freitas et al., 2008).

No método MIA-QSAR, assim como em outros métodos nos quais uma grande quantidade de dados é gerada, tais como os já bem conhecidos CoMFA e CoMSIA, a regressão por PLS é o método mais usualmente aplicado para correlacionar amostras (representadas pelos descritores) com as correspondentes variáveis dependentes. Entretanto, o procedimento MIA-QSAR permite lidar

com arranjos tridimensionais, uma vez que a superposição de k imagens 2D (as estruturas químicas) gera um arranjo em 3D. Esse arranjo de dados pode ser tratado por meio de algoritmos apropriados para a decomposição de dados multimodais, tais como uma regressão multilinear por mínimos quadrados parciais (N-PLS) (Bro, 1996) e análise de fatores paralelos (PARAFAC) (Bro, 1997).

O método multilinear (N-PLS) tem demonstrado várias vantagens quando comparado ao bilinear (PLS), ou seja, N-PLS tem demonstrado ser mais estável, tem melhorado a interpretação dos resultados e aumentado a habilidade preditiva do modelo. Além do mais, a etapa de desdobramento do arranjo 3D não é necessária no N-PLS; em vez disso, a decomposição é realizada diretamente no conjunto de dados (Freitas et al., 2008).

Hasegawa et al. (1999) afirmaram que o principal problema, na maioria das metodologias 3D-QSAR, é que não são precisas para escolher a melhor conformação bioativa e a correspondente função de alinhamento das moléculas. Então, desenvolveram um método com o formalismo PLS 3-way para a resolução deste problema. Outros trabalhos aplicaram métodos PLS 3-way e 4-way em conjunto com 3D-QSAR para procurar pela melhor conformação bioativa e pelos parâmetros mais adequados sobre superfície molecular. A conclusão geral foi que 3-way e 4-way PLS são melhores do que o PLS tradicional para predição (Hasegawa et al., 2000, 2002, 2003).

Os resultados obtidos usando N-PLS e descritores 3D (gerados por meio do programa GRID) (Goodford, 1995), quando se realizou a modelagem de uma série de inibidores do receptor D₂ da dopamina, demonstraram que a modelagem 3D-QSAR com N-PLS é uma excelente alternativa para o tradicional método de regressão PLS. O modelo construído nesse estudo foi considerado altamente preditivo [ligeiramente superior para o MIA-QSAR/PLS modelado para as mesmas séries de compostos (Freitas et al., 2008)], estável (o q^2 da validação

cruzada não variou, independente do número de amostras deixadas de fora nesse procedimento) e interpretável (conformações dos ligantes foram encontradas para explicar as diferenças nos valores das atividades e os pesos das variáveis foram também racionalizados). A mais alta estabilidade é, provavelmente, um efeito originado do método N-PLS, enquanto o aumento na interpretação é diretamente relacionado com o desnecessário passo de desdobramento. Isto é importante no MIA-QSAR, em que os pixels (descritores binários) que compõem uma imagem (uma estrutura química) são desconectados depois do desdobramento do arranjo 3D para uma matriz 2D.

Apesar da grande quantidade de dados gerados da análise de imagem multivariada, que pode ser minimizada pela remoção das colunas com variância zero no modelo MIA-QSAR desdobrado, o arranjo 3D, frequentemente, não é grande o suficiente para tornar a N-PLS impraticável. Enquanto um conjunto de dados 3D é desdobrado para a forma de uma matriz 2D, em que cada fileira contém as variáveis (pixels = binários) descrevendo cada molécula, na tradicional modelagem por regressão PLS, o passo de desdobramento não é essencial em PLS multilinear; pelo contrário, a decomposição é realizada diretamente no arranjo 3D. Em PLS, a matriz \mathbf{X} é decomposta em vetores de “scores” t_1 ($K \times 1$) e vetores pesos (*weight*) w_1 ($IJ \times 1$), em que o vetor *score* é determinado para ter a propriedade de máxima covariância com a variável dependente \mathbf{y} , como descrito para o primeiro componente PLS. Subsequentemente, o vetor *score*, ou seja, as colunas em \mathbf{T} (t_1, \dots, t_A ; A é o número de componentes), é colocado no lugar das variáveis originais como regressores. Diferentemente, em N-PLS, a decomposição compreende um vetor *score* t_1 ($K \times 1$) e dois vetores pesos, w_I ($I \times 1$) e w_J ($J \times 1$), correspondendo à direção molecular, à direção x à direção y . Analogamente ao PLS bilinear, para dado \mathbf{W}^1 (w^I, \dots, w^J), os vetores scores \mathbf{T} (t_1, \dots, t_A) são determinados para ter a

propriedade de máxima covariância com \mathbf{y} . Assim, o modelo PLS multilinear é uma aproximação de $\underline{\mathbf{X}}$ por mínimos quadrados parciais (Freitas et al., 2008).

$$X_{kij} = t_k W_i^I W_j^J + e_{kij} \quad (4)$$

Quanto aos modelos de análises de fatores paralelos (PARAFAC) (Bro, 1997), os mesmos podem ser vistos como modelos de variáveis latentes, estendendo as análises de componentes principais (PCA) para dados multimodos e vêm tendo um uso generalizado, principalmente no campo de quimiometria, especialmente devido à sua propriedade de univocidade (uma única solução é obtida no ajuste do modelo), que pode ser essencial em certas aplicações de resoluções de curva e calibração de segunda ordem. Isto é, se os dados são realmente trilineares, as verdadeiras variáveis adjacentes poderiam ser encontradas se o número certo de componentes fosse usado.

Usualmente, mínimos quadrados alternados (ALS) é o algoritmo para resolver o modelo PARAFAC, mas ferramentas suplementares, tais como redes neurais artificiais (ANNs), podem ser empregadas para melhorar o modelo original PARAFAC. Este modelo está entre os mais promissores métodos multimodos a serem usados em QSAR. Entretanto, apenas um estudo usando PARAFAC em QSAR é conhecido (Bro, 1997). O modelo PARAFAC em 3D é dado por três matrizes de “loadings”, \mathbf{A} , \mathbf{B} e \mathbf{C} , com elementos a_{kj} , b_{ij} e c_{j} . O modelo trilinear deve minimizar a soma dos quadrados dos resíduos, e_{kij} no modelo (Freitas et al., 2008).

$$x_{kij} = \sum_{f=1}^F a_{kf} + b_{if} + c_{jf} + e_{kij} \quad (5)$$

Esta equação pode ser mostrada graficamente na Figura 4, para dois componentes ($F=2$).

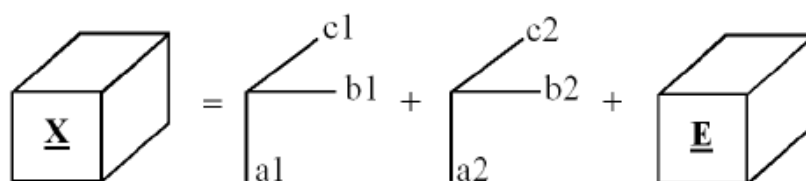


FIGURA 4. Representação gráfica de um modelo PARAFAC de dois componentes de arranjo dos dados **X**. Adaptado de (Freitas et al., 2008).

h.1) Construindo o modelo MIA-QSAR

Segundo Freitas et al. (2008), o primeiro passo na construção do modelo é o desenho das estruturas químicas. Esse é o passo chave para a construção de um modelo preditivo MIA-QSAR. Neste contexto, imagens (estruturas químicas) são matrizes especiais, consistindo de linhas e colunas das imagens (como em uma TV ou desenhos computacionais) em que cada par de coordenada (pixel) corresponde a um valor/intensidade. Essas imagens são matrizes de números; operações espaciais e transformações podem ser aplicadas a elas. Estruturas químicas podem ser desenhadas utilizando-se alguns softwares adequados para desenho de molécula, tais como ChemSketch, ChemDrawn, dentre outros.

As moléculas podem ser sistematicamente desenhadas, isto é, os mesmos fonte e tamanho das letras, assim como símbolos e números, devem ser usados para descrever os átomos na molécula e um critério único deve ser adotado para representar um dado substituinte, por exemplo, Me ou CH₃ (um ou outro), para se referir ao grupo metila. É importante utilizar o recurso de desenhar as estruturas padronizando o tipo e o detalhe de cada átomo ou grupo substituinte para que sejam sempre desenhados da mesma maneira. O fato é que

pequenas diferenças podem afetar o resultado e prejudicar a calibração do modelo.

Assim, alguns cuidados devem ser tomados nessa etapa da construção do modelo. Por exemplo, enantiômeros *R* e *S* de moléculas quirais podem ser esquematicamente distintas pelo tipo de ligação (projetada para frente ou para trás) e isto causaria, se for o caso, diferenças esperadas nas propriedades estudadas. Depois de uma série completa de moléculas ter sido desenhada da mesma maneira, elas podem ser convertidas em *bitmaps*; obviamente, quanto maior a resolução usada, mais preciso será o modelo construído.

O passo de alinhamento 2D começa pela padronização do tamanho da janela em que as imagens (estruturas químicas) serão salvas como *bitmaps*. Já que o MIA-QSAR é um método de aproximação do ligante e, portanto, uma série congênica de compostos é considerada, um ponto em comum entre todas as estruturas deve ser tomado e fixado numa dada coordenada, de tal maneira que cada estrutura fique sobreposta uma sobre a outra por um ponto em comum entre elas. A única porção variante seria(m) o(s) substituinte(s). Por exemplo, um pixel existente no carbono carbonila presente em todos os compostos de uma série pode ser usado como referência. Se mais do que um ponto existe para um conjunto de dados, o pixel mais conveniente para ser usado na sobreposição é aquele que permite uma congruência máxima entre as estruturas, mas, quando se lida com séries congênicas (no caso do MIA-QSAR e métodos que requerem alinhamentos 3D), a porção não variável é sempre a mesma: o substituinte. A etapa do alinhamento 2D requer precisão manual, mas é mais simples do que o tratamento do alinhamento 3D, tais como no CoMFA, em que são previamente necessárias exaustivas varreduras conformacionais (Freitas et al., 2008).

O tratamento das imagens é importante no modelo MIA-QSAR (Freitas et al., 2008). Os *bitmaps* criados devem ser numericamente compreendidos, a fim de permitir a correlação entre a matriz **X** e o vetor coluna **y**, os blocos dos

descritores e das variáveis dependentes, respectivamente. Isso pode ser encontrado pela leitura das imagens previamente criadas e convertendo-as em duplo arranjo (*double array*), por exemplo, utilizando o programa Matlab. Tal conversão é necessária para desmembrar o que é codificado em linguagem de cor para notação de matriz. O resultado é um arranjo 2D ($I \times J$) para cada estrutura, composto pelos binários, por exemplo zero para um pixel e outro número inteiro para o espaço em branco da área de trabalho na qual as estruturas foram desenhadas e salvas. Além disso, para efeito de cálculos, é necessário conhecer o valor numérico correspondente para a cor primária e a respectiva localização na matriz, que é correlacionada com a estrutura química. A fim de evitar tratamentos com cores, deve-se trabalhar em escala cinza, que pode ser encontrada pela soma das três cores primárias, por meio do programa Matlab (Matlab 7.5, 2007); (Freitas et al., 2008).

Todos esses tratamentos consistem nas seguintes linhas de comando do Matlab:

```
[arquivo,MAP] = imread('arquivo.bmp','bmp');  
arquivo =double(arquivo);  
arquivo =(arquivo(:,:,1)+arquivo(:,:,2)+arquivo(:,:,3));
```

Entretanto, o tratamento acima não é um impedimento para outros estudos empregando cores como uma nova dimensão. Logo, diferentes cores podem ser aplicadas como novos descritores. Cada estrutura química pode ser descrita em termos de uma imagem RGB, do inglês “red-green-blue” (vermelho-verde-azul). Na prática, isso pode ser utilizado quando se classificam grupos de substituintes, por exemplo, grupos polares podem ser coloridos em azul, enquanto substituintes não polares ou volumosos são coloridos diferentemente. Isso mostra claramente as vantagens de tais procedimentos, mas requer uma manipulação de dados adicional (Freitas et al., 2008).

O próximo passo na construção do modelo é a construção da matriz \mathbf{X} e calibração. O conjunto de K compostos, descritos como arranjo 2D cada, pode ser agrupado para dar um arranjo 3D $K \times I \times J$. A calibração pode ser realizada neste estágio pela correlação deste arranjo 3D com um bloco \mathbf{y} (o vetor coluna de variáveis dependentes) utilizando um algoritmo de regressão “N-way”, tal como PLS multilinear (N-PLS). Embora N-PLS seja supostamente superior ao PLS desdobrado, devido à sua simplicidade (o número de variáveis pode ser efetivamente reduzido) e habilidade preditiva, o arranjo 3D construído é usualmente maior e, assim, a validação cruzada pode tornar-se impraticável, uma vez que a remoção de pontos não variantes das estruturas (informações irrelevantes) não é possível neste estágio. Consequentemente, cálculos utilizando métodos multimodais só seriam possíveis neste estágio para todo conjunto de dados. Então, é razoável desdobrar o arranjo 3D para a matriz $K \times (I \times J)$, como mostrado na Figura 5.

O desdobramento permite a aplicação de métodos de regressão como o PLS bilinear (o PLS tradicional) e a principal vantagem é a possibilidade de redução da dimensão da matriz \mathbf{X} , sem perda de informação, que pode ser obtida pela remoção de colunas com variância zero, que correspondem aos pontos em comum das estruturas químicas ou aqueles em branco da área de trabalho onde as estruturas foram desenhadas e salvas. Esse procedimento minimiza amplamente a memória usada, reduzindo custos computacionais. Por outro lado, a desvantagem é que torna o redobramento impossível; assim, informações sobre variáveis originais ficariam perdidas e as análises dos *loadings* se tornariam mais difíceis (Freitas et al., 2008).

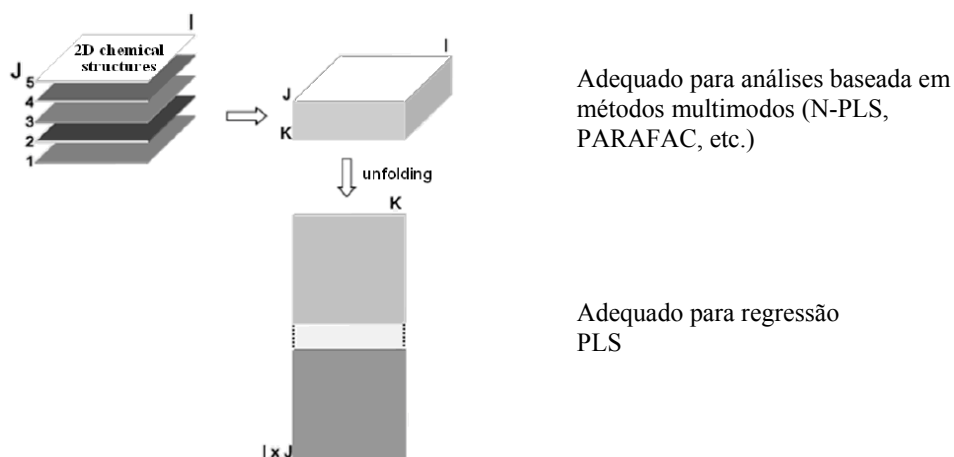


FIGURA 5. A construção do arranjo 3D, adequado para tratamento utilizando métodos multimodos, que podem ser desdobrados para uma matriz 2D (matriz **X**), adequada para análises PLS. Adaptado de (Freitas et al., 2008).

O modelo MIA-QSAR pode ser testado para, dessa forma, avaliar a capacidade de predição do modelo. Nesse sentido, os bem conhecidos testes de validação cruzada e validação externa podem ser realizados para mostrar que boas correlações eventualmente obtidas usando descritores MIA não resultem apenas do acaso. Isso também pode ser assegurado “bagunçando-se” o bloco **y** e sucedendo com a calibração (Freitas et al., 2008).

2.3 Estudos QSAR de inibidores de PDE-5

Além do presente trabalho, alguns estudos QSAR foram realizados com inibidores de PDE-5. Yang et al. (2006) utilizaram estudos de *docking* molecular e análises 3D-QSAR para entender como uma série de derivadas de guanina cíclicas interagem com PDE-5 e PDE-6. Utilizando a conformação de compostos revelados pelo *docking* molecular, análises CoMFA

e CoMSIA resultaram no primeiro modelo quantitativo da relação da estrutura e atividade e no primeiro modelo da relação estrutura seletividade QSSR (com altos valores de coeficientes de correlação de validação cruzada q^2 e calibração r^2) para a predição da atividade inibitória contra PDE-5 e seletividade contra PDE-6. Os altos valores de q^2 e r^2 indicam que os modelos 3D-QSAR e 3D-QSSR obtidos seriam avaliados para a predição de ambos os aspectos, atividade inibitória e seletividade de derivados de guanina cíclicas para estas proteínas alvos.

Um conjunto de contornos 3D baseados nos modelos 3D-QSAR e 3D-QSSR revelou algumas pistas úteis para melhorar a seletividade e a atividade, por meio de modificações realizadas nos compostos. Foi demonstrado que os fatores estéricos e eletrostáticos poderiam ser apropriadamente utilizados para um futuro *design* e o desenvolvimento de um composto mais ativo e seletivo inibidor da PDE-5 para o tratamento da disfunção erétil (Yang et al., 2006).

Srivani et al. (2007) realizaram um estudo 2D e 3D-QSAR, além de um estudo de *docking*, para uma série de inibidores da PDE-5. O modelo 2D-QSAR foi realizado utilizando-se um método de regressão MLR acoplado ao método CONDESSA (compreensão de descritores para análise de estrutura e estatística), obtendo-se bons resultados nos modelos lineares entre atividade inibitória e alguns descritores dos inibidores de PDE-5. Os resultados foram comparados aos obtidos pela modelagem realizada pelos modelos CoMFA e CoMSIA, para confirmar a potência de atividade e seletividade para PDE-5 em relação à PDE-6 de alguns compostos para o tratamento da disfunção erétil.

Eros et al. (2008) estudaram os inibidores da PDE-5, sendo o sildenafil (Viagra) o mais explorado, mas não excluindo os outros compostos desta classe. Foi realizada uma coleta de dados da literatura para os inibidores da PDE-5, formando um conjunto de 438 compostos recentemente publicados em artigos

científicos, sendo obtidos os seus respectivos IC_{50} . Um modelo QSAR foi construído para estes dados.

3 REFERÊNCIAS BIBLIOGRÁFICAS

ANTUNES, J.E.; FREITAS, M.P.; RITTNER, R. Bioactivities of a series of Phosphodiesterase type 5 (PDE-5) inhibitors as modelled by MIA-QSAR. **European Journal of Medicinal Chemistry**, Amsterdam, v. 43 p. 1632-1638, Aug. 2008a.

ANTUNES, J. E.; FREITAS, M. P.; da CUNHA, E. F. F.; RAMALHO, T. C.; RITTNER, R. In silico Prediction of Novel Phosphodiesterase Type-5 Inhibitors Derived from Sildenafil, Vardenafil and Tadalafil. **Bioorganic & Medicinal Chemistry**, Amsterdam, v. 16, p. 7599-7606, Aug. 2008b.

BESTANE, J. W.; BARTOLO, E. B.; FITTIPALDI, J. A.; Prevalence and determinants of erectile dysfunction in Santos, southeastern Brazil. **São Paulo Medicinal Journal**, São Paulo, v. 120, p. 49-54, Mar. 2002.

BRO, R. Multiway calibration. Multilinear PLS. **Journal of Chemometrics**, San Francisco, v. 10, p. 47-61, Jan. 1996.

BRO, R. PARAFAC. Tutorial and applications. **Chemometrics and Intelligent Laboratory Systems**, Texas, v. 38, p. 149-171, Mar. 1997.

CHARTON, M. Handbook of molecular descriptors In: CHARTON, M. (Ed.). **Advances in molecular structure research**. JAI , Greenwich, 1999, p. 25-88.

CHENG, J.; GRANDE, J.P. Cyclic nucleotide Phosphodiesterase (PDE) inhibitors: Novel therapeutic agents for progressive renal disease. **Society for Experimental Biology and Medicine**, New York, v. 232, p. 38-46, Jan. 2007.

CRAMER, D.R. III; PATTERSON, E.D.; BUNCE, D.J. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. **Journal of the American Chemistry Society**, Washington, v. 110, p. 5959-5967, Jan. 1988

DANA, S. A. **A disfunção erétil e a ereção do pênis**: as causas da disfunção erétil. Disponível em: <<http://www.virilplant.com/disfuncoes-sexuais.htm>> Acesso em: 27 set. 2008.

DAUGAN, A.; GRONDIN, P.; RUAULT, C.; DE GOUVILE, A. C. M.; COSTE, H.; KIRILOVSKY, J.; HYAFIL, F.; LABAUDINIERRE, R. The discovery of Tadalafil: A novel and highly selective PDE-5 inhibitor. 1: 5,6,11,11a-Tetrahydro-1*H*imidazo[1',5':1,6]pyrido[3,4-*b*]indole-1,3(2*H*)-dione analogues. **Journal of Medicinal Chemistry**, Washington, v. 46, p. 4525-4532, Dec. 2003.

DIXON, S. L.; MERZ, K. M. One-dimensional molecular representations and similarity calculations: methodology and validation. **Journal of Medicinal Chemistry**, Washington, v. 44, p. 3795-3809, Mar. 2001.

DOBZYNSKI M.; RODRIGUEZ J. V.; KAANDORP J. A.; BLOM J. G. Computational methods for diffusion-influenced biochemical reactions. **Systems Biology**, v. 12, p. 1969-1977, 2007.

EROS, D.; SZNTAI-KIS, C.; KISS, R.; KERI, G.; HEGYMEGI-BARAKONYI, B.; KOVESDI, I. Structure -activity relationships of PDE-5 inhibitors. **Current Medicinal Chemistry**, Sharjah, v. 15, p.1570-1585, 2008.

ESBENSEN, K.; HJELMEN, K. H.; KVAAL, K. The AMT approach in chemometrics - first forays. **Journal of Chemometrics**, San Francisco, v. 10, p. 569-590, Dec.1996.

FAZIO, L.; BROCK, G. Erectile dysfunction: management update. **Canadian Medical Association**, Vancouver, v. 170, p. 1429-1437, Apr. 2004.

FERREIRA, M. M. C. Multivariate QSAR. **Journal of the Brazil Chemical Society**, São Paulo, v. 13, p. 742, Dec. 2002.

FREE S. M.; WILSON J. W. A mathematical contribution to structure-activity studies. **Journal of Medicinal Chemistry**, Washington, v.7 , p. 395-399, July 1964.

FREITAS, M. P.; BROWN, S. D.; MARTINS, J. A. MIA-QSAR: a simple 2D image-based approach for quantitative structure–activity relationship analysis. **Journal of Molecular Structure**, Amsterdam, v. 738, p. 149-154, Jan. 2005.

FREITAS, M. P. MIA-QSAR modelling of anti-HIV-1 activities of some 2-amino-6-arylsulfonylbenzotrioles and their thio and sulfinyl congeners. **Organic & Biomolecular Chemistry**, London, v. 4, p. 1154-1159, Jan. 2006.

FREITAS, M. P.; da CUNHA, E. F. F.; RAMALHO, T. C.; GOODARZI, M. Multimode methods applied on MIA descriptors in QSAR. **Current Computer-Aided Drug Design**, Sharjah, v. 4, p. 273-282, Dec. 2008.

FREITAS, M.P. Multivariate image analysis applied to qsar: evaluation to a series of potential anxiolytic agents. **Chemometrics and Intelligent Laboratory Systems**, Texas, v. 91, p. 173-176, Jan. 2008.

GELADI, P.; ESBENSEN, K. Can image analysis provide information useful in chemistry? **Journal of Chemometrics**, San Francisco, v. 3 p. 419-429, Mar. 1988.

GELADI, P.; ESBENSEN, K. Regression on multivariate images: Principal component regression for modeling, prediction and visual diagnostic tools. **Journal of Chemometrics**, San Francisco, v. 5, p. 97-111, Mar. 1991.

GELADI, P.; GRAHN, H. **Multivariate image analysis**. Chichester: J.Wiley, 1996.

GOLDMAN, B. B.; WIPKE, W. T. QSD quadratic shape descriptors. 2. Molecular docking using quadratic shape descriptors (QSDock), **Proteins: structure, function and genetics**, San Francisco, v. 38, p. 79-94, Jan. 2000.

GOODFORD, J.P. **GRID**. Oxford: University of Oxford, 1995.

GUPTA, M.; KOVAR A. MEIBOHM B. The clinical pharmacokinetics of phosphodiesterase-5 inhibitors for erectile dysfunction. **The Journal of Clinical Pharmacology**, New York, v. 45, p. 987-1003, Sept. 2005.

HAMMETT, L. P. in **Hammett equation, Chap. VII "Physical Organic Chemistry"** 1^a ed., McGraw Hill, New York, 1940.

HANING, H.; NIEWOHRER, U.; SCHENKE, T.; ES-SAYED, M.; SCHMIDT, G.; LAMPE, T.; BISCHOFF, E. Imidazo[5,1-f][1,2,4]triazin-4(3H)-ones, a New Class of Potent PDE 5 Inhibitors, **Bioorganic Medicinal Chemistry Letter**, Amsterdam, v. 12, p. 865-868, Mar. 2002.

HARRISON, S. A.; REIFSNYDER, D. H.; GALLIS, B.; CADD, G. G.; BEAVO, J. A. Isolation and characterization of bovine cardiac muscle cGMP-inhibited phosphodiesterase: a receptor for new cardiotoxic drugs. **The American Society for Pharmacology**, Bethesda, v. 29, p. 506-514, Mar. 1986.

HASEGAWA, K.; ARAKAWA, M.; FUNATSO, K. 3D-QSAR study of insecticidal neonicotinoid compounds based on 3-way partial least squares model. **Chemometrics and Intelligent Laboratory Systems**, Texas, v. 47, p.33-40, Apr. 1999.

HASEGAWA, K.; ARAKAWA, M.; FUNATSO, K. Rational choice of bioactive conformations through use of conformation analysis and 3-way partial least squares modeling, **Chemometrics and Intelligent Laboratory Systems**, Texas, v. 50, p.253-261, Mar. 2000.

HASEGAWA, K.; MATSUOKA, S.; ARAKAWA, M.; FUNATSO, K. New molecular surface-based 3D-QSAR method using Kohonen neural network and 3-way PLS **Computational & Chemistry**, Amsterdam, v.26, p. 583-589, Feb. 2002.

HASEGAWA, K.; ARAKAWA, M.; FUNATSO, K. Simultaneous determination of bioactive conformations and alignment rules by multi-way PLS modeling **Computers Biology and Chemistry**, Amsterdam, v. 27, p.211-216, July 2003.

HOBOHM, U.; SCHARF, M.; SCHNEIDER, R.; SANDER, C. Selection of representative protein data set. **Protein Science**, Hoboken, v. 1, p. 409-417, Oct. 1991.

HOBOHM, U.; SANDER, C. Enlarged representative set of protein structures. **Protein Science**, Hoboken, v. 3, p. 522-524, Mar. 1994.

HOPFINGER, J.A.; WANG, S.; TOKARSKI, S.J.; JIN, B.; ALBUQUERQUE, M.; MADHAV, J.P.; DURAISWAMI, C. Construction of 3D-QSAR Models Using the 4D-QSAR Analysis Formalism. **Journal of the American Chemistry Society**, Washington, v. 119 p. 10509-10524, June 1997.

JAIN, A.N.; NICHOLLS A. Recommendations for evaluation of computational methods. **Journal of Computer-Aided Molecular Design**, New York, v. 22, p. 133-139, Mar. 2008.

JAYATILLEKE, P. R.; NAIR, A. C.; ZAUHAR, R.; WELSH, W. J. Computational Studies on HIV-1 Protease inhibitors: influence of calculated inhibitor-enzyme binding affinities on the statistical quality of 3D-QSAR CoMFA models. **Journal of Medicinal Chemistry**, Washington, v. 43, p. 4446-4451, Oct. 2000.

JIANG, F.; KIM, S. H. "Soft docking": Matching of molecular surface cubes. **Journal of Molecular Biology**, Amsterdam, v. 219, p. 79-102, May 1991.

JIANG, F.; LIN, W.; RAO, Z. SOFTDOCK: understanding of molecular recognition through a systematic docking study. **Protein Engineering Design & Selection**, Oxford, v. 15, p. 257-263, Apr. 2002.

JIANG, S.; TOVCHIGRECHKO, A.; VAKSER, I. A. The role of geometric complementarity in secondary structure packing: A systematic docking study. **Protein Science**, Hoboken, v. 12, p. 1646-1651, Aug. 2003.

KALINCHENKO, S. Y.; KOZLOV, G. I.; GONTCHAROV, N. P.; KATSIYA, G. V. Oral testosterone undecanoate reverses erectile dysfunction associated with diabetes mellitus in patients failing on sildenafil citrate therapy alone. **The Aging Male**, London, v. 6, p. 94-99, June 2003.

KLEBE, G.; ABRAHAM, U.; MIETZNER, T. Molecular similarity indices in a comparative analysis (comsia) of drug molecules to correlate and predict their biological activity. **Journal Medicinal Chemistry**, Washington, v.37, p. 4130, Mar. 1994.

KUBICKI, J. D. Computational chemistry applied to studies of organic contaminants in the environment: examples based on benzo[*a*]pyrene. **American Journal of Science**, New Haven, v. 305, p. 621-644, June 2005.

KUNTZ, I. D.; BLANEY, J. M.; OATLEY, S. J.; LANGRIDGE, R.; FERRINT, T. E., A geometric approach to macromolecule-ligand interactions, **Journal of Molecular Biology**, Amsterdam, v. 161, p. 269-288, Oct. 1982.

KUO, C. L.; ASSEFA, H.; KAMATH, S.; BRZOZOWSKI, Z.; SLAWINSKI, J.; SACZEWSKI, F.; BUOLAMWINI, J. K.; NEAMATI, N. Application of CoMFA and CoMSIA 3D-QSAR and docking studies in optimization of mercaptobenzenesulfonamides as HIV-1 integrase inhibitors. **Journal Medicinal of Chemistry**, Washington, v. 47, p. 385-399, Aug. 2003.

MATLAB VERSION 7.5. Natick, MA: MathWorks, 2007.

MENG, E. C.; SHOICHET, B. K.; KUNTZ, I. D. Automated docking with grid-based energy evaluation. **Journal of Computational Chemistry**, San Francisco, v. 13, p. 505-524, May 1992.

NICOLOTTI, O.; PELLEGRINI-CALACE, C.; CAROTTI, A.; CARRIERI, F.; SANZ, F. Ligands of Neuronal Nicotinic Acetylcholine Receptor (nAChR): Inferences from the Hansch and 3-D Quantitative Structure-Activity Relationship (QSAR) Models. **Current Medicinal Chemistry**, Sharjah, v. 9, p. 1-29, 2002.

NILSSON, J.; HOMAN, E. J.; SMILDE, A. K.; GROEL, C. J.; WIKSTROM, H. A multiway 3D QSAR analysis of a series of (*S*)-*N*-[(1-ethyl-2-pyrrolidinyl)methyl]-6-methoxybenzamides. **Journal of Computer-Aided Molecular Design**, New York, v. 12, p. 81-93, Jan. 1998.

NORWOOD, D. A.; ERAIKHUEMEN, N. E.; PIERRE, M. L.; BRANCH, E.; NORWOOD, R. M. Advances in the treatment of erectile dysfunction: a focus on tadalafil. **Journal of Pharmacy Practice**, New York, v. 17, p. 239-250, Aug. 2004.

RAINA, R.; PAHLAJANI, G.; ARGARWAL, A.; ZIPPE, C. D. Treatment of erectile dysfunction: update. **American Journal of Men's Health**, Thousand Oaks, v. 1, p. 126-138, Jun. 2007.

REES, P. Big pharma learns how to love it. **Scientific Computing World**, Cambridge, v.3, p. 16-18, Aug, 2003.

SINGH, J.; VAN VLIJMEN, H.; LEE, W.C.; LIAO, Y.; LIN, K.C.; ATEEQ, H.; CUERVO, J.; ZIMMERMAN, C.; HAMMOND, C.; KARPUSAS, M.; PALMER, R.; HATTOPADHYAY, T.; ADAMS, S. P. 3D QSAR (COMFA) of a series of potent and highly selective VLA-4 antagonists. **Journal of Computer Aided Molecular Design**, New York, v. 16, p. 201-211, Mar. 2002.

SRIVANI, P.; SRINIVAS, E.; RAGHU, R.; SASTRY, GN. Molecular modeling studies of pyridopurinone derivatives--potential phosphodiesterase 5 inhibitors. **Journal Molecular Graphic and Modelling**, Amsterdam, v. 26, p.378-401, Oct. 2007.

TAKAHASHI, T.; FUJIMURA, H.; OKAMURA, K., Syntheses of analgesics. XXX. Indanamine derivatives, **Journal of Pharmaceutical Society of Japan**, Tokyo, vol. 82, p. 1597-1603, Dec. 1962.

TERRETT, N. K.; BELL, A. S.; BROWN, D.; ELLIS, P. Sildenafil (Viagratm), a potent and selective inhibitor of type 5 cGMP phosphodiesterase with utility for the treatment of male erectile dysfunction. **Bioorganic & Medicinal Chemistry Letter**, Amsterdam, v. 6, p. 1819-1824, Aug. 1996.

THOMSEN, R.; CHRISTENSEN, M.H. MolDock: a new technique for high-accuracy molecular docking. **Journal of Medicinal Chemistry**, Washington, v. 49, 3315, Apr. 2006.

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

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

YANG, G. F.; LU, T.H.; XIONG, F.; ZHAN, C.G. Understanding the structure–activity and structure–selectivity correlation of cyclic guanine derivatives as phosphodiesterase-5 inhibitors by molecular docking, CoMFA and CoMSIA analyses. **Bioorganic & Medicinal Chemistry**, Amsterdam, v. 14 p. 1462-1473, Mar. 2006.

CAPÍTULO 2

AVALIAÇÃO PRELIMINAR DO MÉTODO MIA-QSAR PARA UMA SÉRIE DE INIBIDORES DE PDE-5

1 RESUMO

As bioatividades de uma série de derivados de guanina cíclica, inibidores da fosfodiesterase tipo-5 (PDE-5) foram modeladas utilizando-se uma aproximação baseada em análise de imagens para relação quantitativa de estrutura e atividade (MIA-QSAR). A calibração do modelo mostrou ser robusta, com r^2 de 0,864 utilizando-se cinco componentes PLS. A habilidade preditiva do modelo foi testada por meio de validação cruzada *leave-one-out*, fornecendo um q^2 de 0,605. Este valor aumentou para 0,721 depois da remoção de dois *outliers*. Um conjunto de validação externa também foi utilizado para corroborar o potencial da modelagem e os resultados mostraram ser compatíveis com aqueles obtidos de uma metodologia 3D previamente aplicada para esta série de compostos. O método mostrou ser uma ferramenta potencial para a predição de novas drogas, como exemplificado pelo cálculo das bioatividades de dois novos congêneres propostos derivados do conjunto treino.

2 ABSTRACT

A series of cyclic guanine derivatives, phosphodiesterase type 5 (PDE-5) inhibitors, have been modelled using an image-based approach for quantitative structure activity relationships (MIA-QSAR). The calibration model showed to be robust with a r^2 of 0.864 using five PLS components. The predictive ability of the model was tested through leave-one-out cross-validation, giving a q^2 of 0.605 (q^2 improves to 0.721 after removing two outliers). An external validation set was also used to give an account for the modeling capability, and the results agreed with the ones obtained from a 3D methodology previously applied to this series of compounds. The method showed to be a potential tool for predicting new drug-like compounds, as exemplified by calculating the activities of two new proposed congeners derived from the training set.

3 INTRODUÇÃO

A metodologia MIA-QSAR, recentemente desenvolvida, baseada em análise de imagens 2D, evita varreduras conformacionais, alinhamentos tridimensionais e tem demonstrado ser confiável, preditiva e com algumas vantagens operacionais (Freitas et al., 2005). Assim, este método pode ser utilizado para a modelagem das atividades biológicas de uma série de compostos com potencial para tratamento da disfunção erétil, como um teste preliminar para predições de análogos de inibidores da PDE-5 comercialmente existentes. Isto se configura como uma possibilidade real no desenvolvimento de novos agentes terapêuticos. Para tal, as atividades biológicas de uma série de derivados de guaninas cíclicas, inibidores da PDE-5, foram modeladas empregando-se a metodologia MIA-QSAR. As estruturas químicas dos 49 derivados de guanina cíclica utilizados na modelagem MIA-QSAR, bem como suas respectivas atividades biológicas, foram obtidas na literatura (Yang et al., 2006).

4 MÉTODOS COMPUTACIONAIS

A série de estruturas **1-59** (Tabela 1) foi sistematicamente desenhada utilizando o software ChemSketch (ACD ChemSketch version 8.17, 2005) e, então, convertida para *bitmaps* numa janela de 250 x 250 pixels, com resolução de 102 x 102 pontos por polegadas. Dez estruturas moleculares foram descartadas devido à indisponibilidade de medidas experimentais. Todas as estruturas moleculares foram fixadas por um ponto comum entre elas, numa dada coordenada (Figura 1). Assim, as mesmas foram sobrepostas, como um alinhamento 2D para permitir máxima similaridade e congruência entre as estruturas. No conjunto de dados deste trabalho, o pixel localizado na coordenada 50,100 (sobre o carbono carbonila presente em todas a série) foi usado como referência no passo de alinhamento. Cada imagem 2D foi lida e convertida em binários (arranjo duplo no Matlab) e um arranjo 3D, o bloco de variáveis independentes, foi construído pelo agrupamento das 49 imagens tratadas, resultando num arranjo 3D de dimensão 49 x 250 x 250. O arranjo 3D foi então desdobrado para formar a matriz **X** (49 x 62500), a fim de ser correlacionada com o bloco **y** (vetor coluna das atividades) pelo uso do algoritmo NIPALS (Wold, 1966) para regressão por PLS. Entretanto, com a intenção de minimizar a memória usada, colunas com variância zero foram removidas, reduzindo a dimensão da matriz **X** para 49 x 3242. Esta matriz foi dividida em duas: conjuntos treino (37 x 3242) e teste (12 x 3242).

A validação do modelo de calibração foi feita por meio de validação cruzada deixe uma fora, do inglês *leave-one-out* (LOOcv), e validação externa (conjunto teste). A qualidade dos resultados foi avaliada pela análise de r^2 , q^2_{cv} e q^2 , os coeficientes de correlação ao quadrado das atividades experimentais *versus* os valores calibrados, validados (LOO_{CV}) e preditos, respectivamente.

Além disso, a soma dos quadrados médios dos erros da calibração, validação cruzada e validação externa (RMSEC, RMSECV e RMSEP, respectivamente) também foi empregada como parâmetro estatístico para avaliar a habilidade de predição do modelo. Tratamentos computacionais, incluindo diagnósticos de *outliers* usando análises baseadas nas distâncias entre as amostras (distância de Mahalanobis, resíduos de Student e Leverage), também foram aplicados, utilizando-se o programa Pirouette (Pirouette 3.11, 2003).

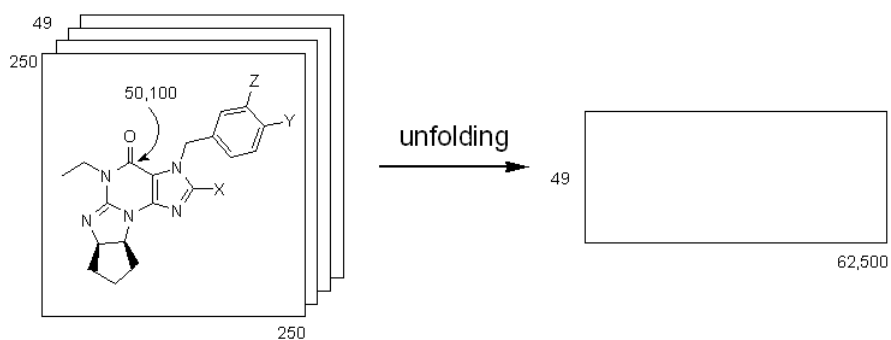
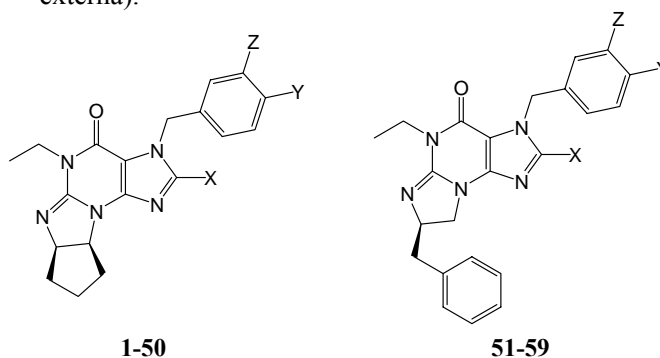


FIGURA 1. Estrutura molecular genérica dos derivados de guanina cíclica, inibidores de PDE-5 (**1-59**), desenhadas numa janela de dimensão 250 x 250 pixels. O ponto fixado na coordenada 50,100 está indicado na estrutura. O arranjo 3D foi desdobrado para a matriz **X** de dimensão 49 x 62500. Adaptado de (Antunes et al., 2008a)

TABELA 1. Estrutura molecular dos compostos nos conjuntos treino e teste e valores experimentais e preditos de pIC₅₀. ^a pIC₅₀ para inibidores de PDE-5 experimental, calibrado e predito (LOOcv e validação externa).



Comp.	X	Y	Z	Exp.	Calib.	LOO CV	Predito
1	CH ₂ -Ph	H	H	6,89	6,97	7,11	
2	CH ₂ -Ph	Me	H	7,40	7,44	7,17	
5	CH ₂ -Ph	Cl	H	6,80	6,73	7,18	
15	CH ₂ -Ph	OMe	Cl	7,24	7,28	7,32	
16	CH ₂ -Ph	3,4-OCH ₂ O		7,34	7,24	7,01	
18	CCPh	Cl	H	8,30	8,51	9,19	
19	CCPh	OMe	H	8,51	8,45	7,96	
20	CCPh	OH	H	9,52	9,27	8,36	
25	OEt	OMe	H	7,74	7,80	8,15	
26	SEt	OMe	H	7,89	7,82	8,11	
27	COOMe	OMe	H	8,17	8,19	8,23	
28	CN	OMe	H	8,20	8,26	8,23	
29	CONH ₂	OMe	H	8,40	8,41	8,50	
31	CF ₃	OH	H	8,15	8,07	8,59	
32	CONH ₂	OH	H	8,68	8,75	8,38	
33	SEt	OH	H	8,74	8,75	8,31	
36	H	OMe	Br	8,19	8,19	8,56	
37	H	OMe	Cl	8,08	8,02	8,56	
39	CONH ₂	OMe	Br	8,82	8,88	8,40	
40	CONH ₂	OH	Br	9,05	9,13	9,01	
41	CONH ₂	OH	Cl	8,96	8,94	9,08	
42	CONHMe	OMe	Cl	8,64	8,69	8,19	

...continua...

TABELA 1, Cont.

43	CONHMe	OH	Cl	8,89	8,96	8,88
44	OEt	OMe	Br	8,27	8,27	8,04
45	OEt	OMe	Cl	8,19	8,10	8,04
46	OEt	OH	Cl	8,68	8,66	8,90
47	OMe	OMe	Br	8,55	8,50	8,35
48	OMe	OMe	Cl	8,30	8,33	8,40
49	OMe	OH	Br	8,80	8,83	8,93
50	OMe	OH	Cl	8,70	8,64	9,00
51	CONH ₂	OMe	Cl	8,55	8,53	8,44
52	CONH ₂	OH	Cl	8,52	8,55	8,94
53	OBn	OH	Cl	8,68	8,69	7,85
54	OMe	OMe	CN	7,96	7,98	7,89
55	OMe	OMe	Br	8,66	8,64	8,37
58	OMe	OH	Br	8,89	8,88	8,71
59	OMe	OH	Cl	8,72	8,69	8,82
7*	CH ₂ -Ph	OMe	H	7,70		7,16
8*	CH ₂ -Ph	OH	H	8,41		7,16
9*	CH ₂ -Ph	NH ₂	H	6,26		7,20
17*	CCPh	H	H	8,10		8,94
23*	H	OMe	H	7,26		8,43
24*	CF ₃	OMe	H	7,32		8,14
30*	NH ₂	OH	H	7,66		8,72
34*	OEt	OH	H	9,05		8,49
35*	N ₃	OH	H	9,21		8,42
38*	H	OH	Br	8,20		9,12
56*	OMe	OMe	Cl	8,80		8,45
57*	OMe	OH	CN	7,37		8,18

* Compostos do grupo teste.

Legenda:

Comp. (Composto)

Adaptado de (Antunes et al., 2008a)

5 OBJETIVOS

Avaliar MIA-QSAR como método de modelagem das atividades de uma série de inibidores de PDE-5 (alguns derivados cíclicos de guanina).

6 RESULTADOS E DISCUSSÃO

No método MIA-QSAR, descritores (parâmetros requeridos para correlacionar as amostras – moléculas - com as correspondentes variáveis dependentes – as bioatividades) são pixels. Geladi & Esbensen (1989) demonstraram que análises de imagens podem conter muitas informações em química, embora os descritores não tenham um significado físico-químico direta, pois são números binários. Em QSAR, esses descritores têm mostrado conter informações químicas (Freitas et al., 2005; Freitas, 2006), permitindo a correlação entre estruturas químicas e bioatividades.

O ponto chave na modelagem MIA-QSAR para esta série de inibidores de PDE-5, que foram previamente avaliados por meio dos métodos CoMFA e CoMSIA (Yang et al., 2006), foi a construção da matriz **X**. A calibração foi realizada para o conjunto treino utilizando-se regressão por PLS. A soma dos quadrados médios dos erros da validação cruzada oscilou quando mais de 5 variáveis latentes (VL) foram usadas e um insignificante aumento foi observado na variância explicada dos dados usando um maior número de VLs (Tabela 2 e Figura 2). Então, a menor soma dos quadrados dos erros de predição acumulada (*cumpress*) foi alcançada para 5 VLs e um r^2 de 0,826 foi obtido.

TABELA 2. Parâmetros estatísticos da modelagem MIA-QSAR.

LV	% Cum. Var.	R^2	RMSEC	F	t	Q^2_{cv}	RMSECV	Q^2	RMSEP
1	83,7	0,024	0,66	-22,2	75,1	0,013	0,69	0,000	0,97
2	88,2	0,327	0,48	-13,7	102,7	0,159	0,56	0,192	0,85
3	91,1	0,621	0,37	40,4	134,5	0,387	0,49	0,125	0,84
4	93,1	0,830	0,24	155,8	204,7	0,572	0,39	0,140	0,87
5^a	94,6	0,864	0,22	219,9	224,5	0,605	0,38	0,137	0,87
6	95,9	0,914	0,17	352,5	288,6	0,609	0,39	0,145	0,85
7	96,2	0,945	0,13	600,6	361,8	0,604	0,40	0,269	0,76
8	96,6	0,962	0,11	896,4	437,6	0,616	0,40	0,265	0,77
9	96,9	0,975	0,09	1398,7	539,3	0,637	0,39	0,272	0,76
10	97,1	0,985	0,07	2310,4	687,9	0,642	0,38	0,294	0,73

^a Número ótimo de variáveis latentes.

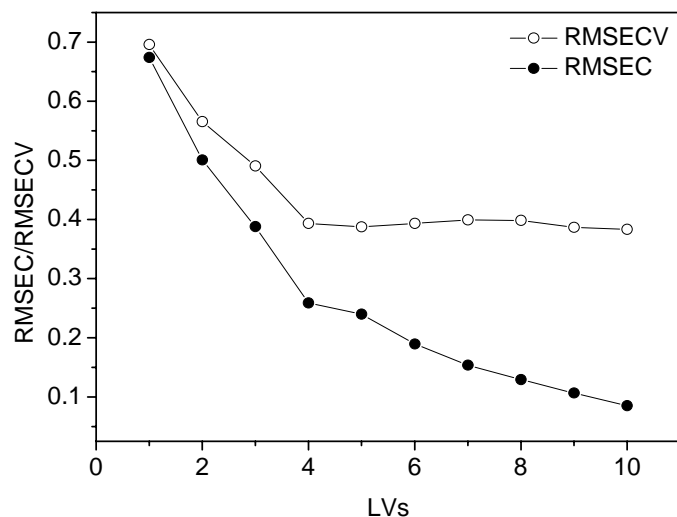


FIGURA 2. Gráfico de RMSEC/RMSECV versus números de variáveis latentes. Adaptado de (Antunes et al., 2008a)

Muitas variáveis dependentes e independentes podem ser correlacionadas utilizando-se tratamento por PLS. Dependendo do número de variáveis latentes usado, altos coeficientes de correlação podem ser obtidos utilizando regressões multivariadas, mesmo que, na verdade, pouca correlação exista entre os dois blocos. A fim de assegurar-se de que isso não ocorreu com o presente modelo, ou seja, para garantir que a modelagem MIA-QSAR para os inibidores da PDE-5 não tenha sido o resultado de falsas correlações entre estrutura e atividade, o bloco y (vetor coluna das atividades) foi “randomizado” de tal maneira que os compostos não estavam mais correspondendo aos seus respectivos valores de atividades biológicas. Este procedimento resultou numa diminuição significativa do r^2 (0,50) em relação ao valor real de r^2 para 5 VLs, confirmando que a calibração não foi uma correlação fortuita, permitindo, assim, assegurar a robustez do modelo.

A calibração do modelo foi validada por meio da validação cruzada (LOOcv), obtendo-se os valores preditos da Tabela 3 e Figura 3. A correlação

obtida com q^2_{cv} de 0,605 é condizente com o valor de um modelo “validável” (q^2_{cv} deveria ser acima de 0,500) (Golbraikh & Tropsha, 2002). Entretanto, duas amostras anômalas com altos resíduos em LOOCv foram identificadas por meio de análises de diagnósticos de *outliers*, nos quais os *scores* para ambas as amostras estão fora dos limites, de acordo com análises de distância de Mahalanobis, resíduos de Student e Leverage, como pode ser visualizado na Figura 4. Isto pode ser devido a uma provável imprecisão na medida experimental das bioatividades e ou devido à pobre predição dos efeitos de um ou mais substituintes em determinadas posições do anel benzênico dos compostos. O q^2_{cv} é significativamente elevado após a retirada dos *outliers*: 0,721.

A realização de uma validação externa é altamente recomendável, a fim de estabelecer um modelo QSAR (Golbraikh & Tropsha, 2002). Assim, doze compostos foram separados de acordo com o grupo teste previamente escolhido na literatura (Yang et al. 2006), que foram selecionados de forma aleatória a partir conjunto de moléculas; conseqüentemente, a predição das atividades desse grupo de compostos foi avaliada. Embora uma correlação estatística insignificante tenha sido obtida (q^2 de 0,137), essencialmente devido a algum substituinte não ou pobremente calibrado, os resultados concordam plenamente com aqueles obtidos por meio de sofisticadas metodologias 3D (Tabela 4). Entretanto, como em qualquer calibração, a confiabilidade do modelo deixa de existir quando ocorrem extrapolações, tal como para os compostos **9** e **35**, que possuem substituintes não calibrados. Assim, outros conjuntos testes foram simulados pela divisão do conjunto treino em 7 segmentos (chamado validação cruzada deixe 20% fora ou *leave-20%-out cross validation*). Cada segmento, composto de sete amostras selecionadas de forma aleatória, foi deixado de fora; em seguida, uma calibração foi realizada utilizando-se os seis segmentos remanescentes e, então, as bioatividades para amostras deixadas de fora foram

preditas empregando-se os parâmetros de regressão correspondentes aos seis segmentos calibrados. Este procedimento foi efetuado dez vezes, resultando num RMSEP médio de $0,60 \pm 0,03$, significativamente melhor do que o resultado obtido usando o conjunto teste da literatura (Yang et al., 2006).

TABELA 3. Valores preditos pIC_{50} e residuais para o conjunto teste, usando os modelos MIA-QSAR e da literatura^a. Adaptado de (Antunes et al., 2008a)

Comp.	Exp.	MIA-QSAR		CoMFA		CoMSIA	
		Pred.	Res.	Pred.	Res.	Pred.	Res.
7	7,70	7,16	0,54	6,89	0,81	7,17	0,53
8	8,41	7,16	1,25	7,35	1,06	7,66	0,75
9	6,26	7,20	-0,94	7,04	-0,78	6,87	-0,61
17	8,10	8,94	-0,84	8,58	-0,48	8,40	-0,3
23	7,26	8,43	-1,17	7,95	-0,69	7,83	-0,57
24	7,32	8,14	-0,82	7,89	-0,57	7,74	-0,42
30	7,66	8,72	-1,06	8,63	-0,97	8,21	-0,55
34	9,05	8,49	0,56	8,55	0,5	8,50	0,55
35	9,21	8,42	0,79	8,68	0,53	8,13	1,08
38	8,20	9,12	-0,92	7,95	0,25	8,54	-0,34
56	8,80	8,45	0,35	8,17	0,63	8,10	0,7
57	7,37	8,18	-0,81	8,38	-1,01	8,36	-0,99

Legenda:

-Comp. (composto),

-Exp. (atividade biológica experimental).

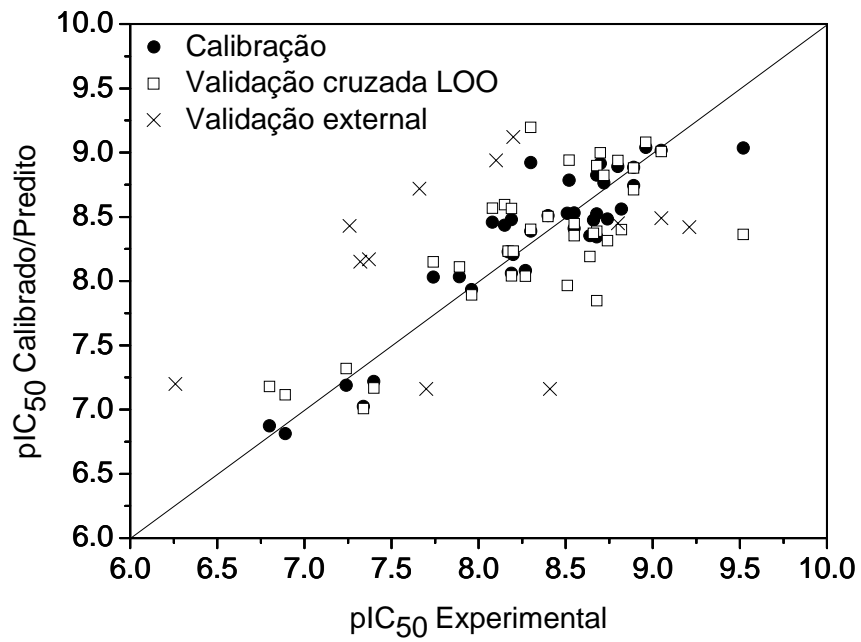


FIGURA 3. pIC_{50} calibrado/predito *versus* pIC_{50} experimental. Adaptado de (Antunes et al., 2008a)

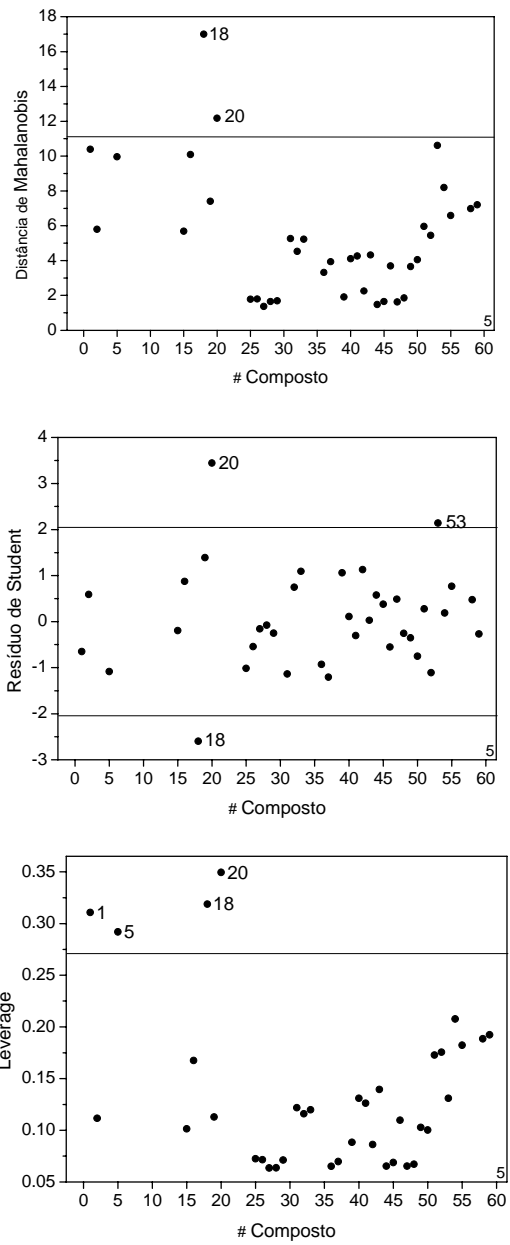


FIGURA 4. Gráficos utilizados para os diagnósticos de *outliers*. A linha horizontal indica a região limitante. Adaptado de (Antunes et al., 2008a)

O método MIA-QSAR tem o potencial para ser utilizado como uma ferramenta para a predição de novas drogas; estruturas de novos compostos podem ser propostas pela miscelânea de subestruturas dos congêneres experimentalmente mais ativos do conjunto treino das moléculas em estudo. A fim de ilustrar isso, subestruturas dos compostos **20**, **40** e **41** foram recombinadas para sugerir novos potentes ligantes. Neste exemplo, a estratégia oferece as estruturas **A** e **B** (Figura 5). Avaliação desses compostos propostos utilizando o modelo MIA-QSAR construído resultou em valores de pIC_{50} de 8,73 e 8,71, para **A** e **B**, respectivamente, sugerindo que eles seriam razoavelmente efetivos como inibidores da PDE-5.

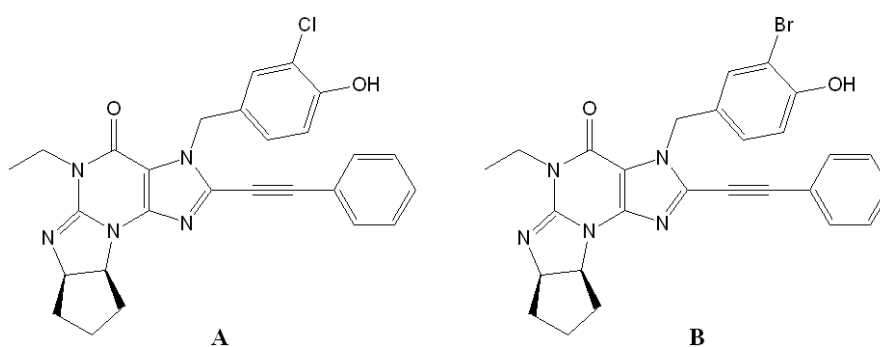


FIGURA 5. Moléculas propostas como novas drogas derivadas da modelagem MIA-QSAR. Adaptado de (Antunes et al., 2008a)

7 CONCLUSÕES

O modelo MIA-QSAR, construído a partir de uma série de inibidores da PDE-5, mostrou ser robusto, com habilidade preditiva comparável à de modelagens 3D. Dois novos compostos com a base estrutural de moléculas treino foram propostos, mostrando a potencialidade do método como uma ferramenta, por exemplo, para propor novos ligantes para síntese, que seriam uma miscelânea de subestruturas de duas ou mais diferentes séries congênicas contendo certa similaridade. O método apresentou resultados satisfatórios, sendo uma ferramenta valiosa em muitos aspectos, como, por exemplo, na aplicação de softwares de desenhos de estruturas químicas e os bem conhecidos métodos de regressão multivariada, tais como PLS, para dar informação química de uma maneira confiável e simples. Apesar da ausência de interpretação físico-química, os descritores MIA podem ser utilizados com sucesso para realizar predições de bioatividades de novos compostos.

8 REFERÊNCIAS BIBLIOGRÁFICAS

ACD/CHEMSKETCH VERSION 8.17. **Advanced chemistry development.** Toronto, Ont., Canada, 2005.

FREITAS, M.P. MIA-QSAR modelling of anti-HIV-1 activities of some 2-amino-6-arylsulfonylbenzotrioles and their thio and sulfinyl congeners. **Organic & Biomolecular Chemistry**, Amsterdam, v. 4, p. 1154–1159, Jan. 2006.

FREITAS, M. P.; BROWN, S. D.; MARTINS, J. A., MIA-QSAR: a simple 2D image-based approach for quantitative structure–activity relationship analysis. **Journal of Molecular Structure**, Amsterdam, vol. 738, p. 149-154, Jan. 2005.

GELADI, P.; ESBENSEN, K., Can image analysis provide information useful in chemistry? **Journal of Chemometrics**, San Francisco, v. 3, p. 419–429, Mar. 1989.

GOLBRAIKH, A.; TROPSHA, A. Beware of q^2 ! **Journal of Molecular Graphic and Modelling**, Amsterdam, v. 20, p. 269-273, Jan. 2002

PIROUETTE 3.11. X. Bothell, WA: Advanced Chemistry Development, 2003.

WOLD, H., Estimation of principal components and related models by iterative least squares. In: KRISHNAIAH, K.R. (Ed). **Multivariate analysis.** New York: Academic, 1966.p. 391-420.

YANG, G.F.; LU, H.T.; XIONG, Y.; ZHAN, C.G. Understanding the structure–activity and structure–selectivity correlation of cyclic guanine derivatives as phosphodiesterase-5 inhibitors by molecular docking, CoMFA and CoMSIA analyses, **Bioorganic & Medicinal Chemistry**, Amsterdam, v. 14, p. 1462-1473, Mar. 2006.

CAPÍTULO 3

PROPOSIÇÃO COMPUTACIONAL DE NOVOS INIBIDORES DE PDE-5

1 RESUMO

As atividades biológicas de uma série de compostos derivados de sildenafil, vardenafil e tadalafil foram modeladas por meio do método MIA-QSAR. Um modelo altamente preditivo foi construído e novos compostos, que são miscelâneas de subestruturas desses três representativos inibidores da PDE-5, foram preditos utilizando os parâmetros de calibração obtidos por meio de regressão por PLS. A alta bioatividade de oito compostos promissores foi corroborada por um estudo de *docking*. Perfis ADME-TOX (absorção, distribuição, metabolismo, excreção e toxicidade) foram calculados para tais compostos, sugerindo vantagens de alguns deles sobre as drogas mais comuns atualmente disponíveis para o tratamento de disfunção erétil.

2 ABSTRACT

The activities of a series of drug-like compounds derived from Sildenafil, Vardenafil and Tadalafil analogues were modeled through the MIA-QSAR (multivariate image analysis applied to quantitative structure-activity relationships) ligand-based approach. A highly predictive model was achieved and novel compounds, miscellany of substructures of these three representative phosphodiesterase type-5 (PDE-5) inhibitors were predicted using the calibration parameters obtained through partial least squares (PLS) regression. The high bioactivities of eight promising compounds were corroborated by docking evaluation. Calculated ADME-Tox (absorption, distribution, metabolism, excretion and toxicity) profiles for such compounds suggest advantages of some of them over the currently available, most common drugs used for the treatment of erectile dysfunction.

3 INTRODUÇÃO

Considerando que o teste preliminar apresentado no capítulo anterior sugere que o método MIA-QSAR é passível de aplicação para *design* de novos inibidores de PDE-5, um novo modelo QSAR foi construído, baseado em estruturas químicas utilizadas para o desenvolvimento dos fármacos sildenafil, vardenafil e tadalafil (Terrett et al., 1996; Haning et al., 2002; Daugan et al., 2003).

A seguir, são apresentadas a metodologia aplicada para a construção do modelo e a proposição de novos fármacos, e a discussão dos resultados obtidos.

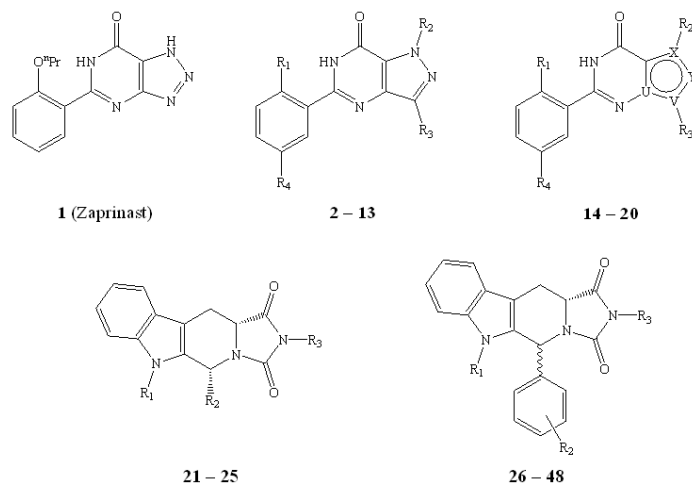
4 MÉTODOS COMPUTACIONAIS

4.1 Construção do modelo MIA-QSAR

O modelo MIA-QSAR foi construído a partir de estruturas análogas dos mais tradicionais inibidores de PDE-5 atualmente comercializados (Figura 1). Os descritores MIA, como apontado anteriormente, são números binários obtidos de pixels de estruturas químicas bidimensionais, que podem ser desenhadas pelo uso de programas especiais para esta função. Assim, as estruturas de 48 inibidores da PDE-5 (Figura 1), obtidas na literatura (Terrett et al., 1996; Haning et al., 2002; Daugan et al., 2003), foram desenhadas por meio do programa ChemSketch (ACD/ChemSketch 8.17, 2005). Cada estrutura química 2D foi sistematicamente construída, isto é, o esqueleto básico das estruturas (os anéis de 5 e 6 membros fundidos) foi mantido inalterado e os substituintes da Figura 1 foram, então, adicionados. Em seguida, as estruturas químicas foram transformadas em *bitmaps* e salvas numa janela de dimensão 400 x 400 pixels. Assim, o conjunto de dados utilizado é uma série de compostos congêneros, pois as estruturas químicas que possuem subestruturas em comum foram sobrepostas ao longo de toda a série. Isto foi feito pelo uso de um pixel em comum entre toda a série de compostos, utilizando a coordenada 200,100 das figuras.

Este alinhamento 2D é rápido e somente requer uma razoável precisão manual. As 48 imagens 2D foram lidas e convertidas em um arranjo duplo pelo uso do programa Matlab. Estas 48 amostras foram então agrupadas para formar um arranjo 3D (48 x 400 x 400), o qual foi desdobrado para um arranjo 2D (48 x 160,000). Dois *outliers* foram identificados por meio de diagnósticos de Students (compostos **1** e **36**) e, então, descartados do conjunto de dados, formando a matriz **X** (46 x 160,000).

A calibração do modelo foi realizada por meio da regressão por mínimos quadrados parciais (PLS) e a validação do modelo foi determinada pela validação cruzada “deixe uma de fora” (*leave-one-out* ou LOOCV). A validação externa é bastante recomendada (Golbraikh & Tropsha, 2002), assim, amostras selecionadas aleatoriamente, 20% do total da série de 46 compostos, foram usadas como um conjunto teste. Foram realizadas misturas aleatórias por 10 vezes e a predição média foi considerada. A habilidade preditiva foi estatisticamente avaliada pelos cálculos das somas dos quadrados dos erros de calibração (RMSEC), validação cruzada (RMSECV) e validação externa (RMSEP), assim como pelos coeficientes de correlação ao quadrado da linha de regressão dos valores experimentais *versus* calibrados/preditos (r^2 e q^2).



Cpd	R ₁	R ₂	R ₃	R ₄	U	V	Y	X	Cpd	R ₁	R ₂	R ₃
1	O ^t Pr	H		H					25	H	<i>cis</i> -phenyl	^t Bu
2	OEt	Me	Me	H					26	H	(<i>trans</i>) H	^t Bu
3	OEt	Me	ⁿ Pr	H					27	H	(<i>cis</i>) 4-O ^t Me	^t Bu
4	H	H	ⁿ Pr	H					28	H	(<i>trans</i>) 4-O ^t Me	^t Bu
5	H	Me	ⁿ Pr	H					29	H	(<i>cis</i>) 3-Cl	^t Bu
6	OH	Me	ⁿ Pr	H					30	H	(<i>trans</i>) 3-Cl	^t Bu
7	OCH ₂ ^t Pr	Me	ⁿ Pr	H					31	H	(<i>cis</i>) 4-Cl	^t Bu
8	NO ₂	Me	ⁿ Pr	H					32	H	(<i>trans</i>) 4-Cl	^t Bu
9	NHSO ₂ Me	Me	ⁿ Pr	H					33	H	(<i>cis</i>) 4-CN	^t Bu
10	OEt	Me	ⁿ Pr						34	H	(<i>trans</i>) 4-CN	^t Bu
11	OEt	Me	ⁿ Pr						35	H	(<i>trans</i>) 2-O ^t Me	^t Bu
12	OEt	Me	ⁿ Pr						36	Me	(<i>trans</i>) 4-O ^t Me	^t Bu
13	OEt	Me	ⁿ Pr						37	H	(<i>trans</i>) 4-OEt	H
14	OEt	Me	ⁿ Pr	H	N	C	N	C	38	H	(<i>cis</i>) 4-OEt	Me
15	OEt	Me	ⁿ Pr		N	C	N	C	39	H	(<i>cis</i>) 4-OEt	Et
16	OEt	Me	ⁿ Pr	H	C	C	N	N	40	H	(<i>trans</i>) 4-OEt	Et
17	OEt	Me	ⁿ Pr		C	C	N	N	41	H	(<i>cis</i>) 4-OEt	Bz
18	O ⁿ Pr		ⁿ Pr	H	C	N	C	N	42	H	(<i>trans</i>) 4-OEt	Bz
19	O ⁿ Pr		ⁿ Pr		C	N	C	N	43	H	(<i>cis</i>) 4-OEt	ⁿ hexyl
20	OEt	Me	ⁿ Pr		N	C	N	C	44	H	(<i>trans</i>) 4-OEt	ⁿ hexyl
21	H	<i>cis</i> -4-pyridinyl	^t Bu						45	H	(<i>trans</i>) 4-OEt	CH ₂ CH ₂ NMe ₂
22	H	<i>cis</i> -3-pyridinyl	^t Bu						46	H	(<i>trans</i>) 4-OEt	CH ₂ -3-pyridinyl
23	H	<i>cis</i> -3-thienyl	^t Bu						47	H	(<i>trans</i>) 4-OEt	CH ₂ CH ₂ -2-pyridinyl
24	H	<i>cis</i> -3-furanyl	^t Bu						48	H	(<i>trans</i>) 4-OEt	CH ₂ CH ₂ -pyrrolidinyl

FIGURA 1. Estruturas dos inibidores de PDE-5 no conjunto de dados. (Antunes et al., 2008b).

4.2 Cálculos de *Docking*

Coordenadas do cristal da enzima PDE-5 ligada com sildenafil e com tadalafil foram obtidas do banco de dados de proteínas (códigos PDB: 1 TBF e 1 XOZ, respectivamente). Os compostos sildenafil, vardenafil, tadalafil, **2-5, 10-13, 15, 21-25, 27, 29, 31, 33, 38, 41-43** (Tabela 1) e **A-H** foram submetidos aos estudos de *docking*, dentro do sítio ativo da PDE-5. As estruturas tridimensionais (3D) dos compostos foram obtidas a partir do programa PC Spartan módulo Pro/Builder (SpartanPro 1.0.1, 2001), utilizando-se os cristais das estruturas do sildenafil e vardenafil ligados à PDE-5 como modelos. Subsequentemente, as geometrias completamente otimizadas e os cálculos de distribuição das cargas atômicas dos ligantes foram realizados com o mesmo programa, utilizando o método semi-empírico AM1. Os compostos foram “docados” dentro do sítio ativo da enzima PDE-5, empregando-se o *Molegro Virtual Docker* (MVD) (Thomsen & Christensen, 2006), um programa que permite prever a conformação mais provável com que um ligante poderia se ligar a uma macromolécula. A função MolDock de “scoring” (*Mol Dock score*), utilizada pelo programa MVD, é derivada do *Potencial linear piecewise*, PLP e estendida no *Generic evolutionary method for molecular DOCK* (GEMDOCK) com um novo termo de ligações de hidrogênio e novos esquemas de cargas (Thomsen & Christensen, 2006). O algoritmo de busca do estudo de *docking* foi o MVD baseado em técnicas iterativas de otimizações, inspiradas pela teoria evolucionista de Darwin (algoritmos evolucionários – AE). O potencial sítio de ligação do receptor PDE-5 foi calculado utilizando-se o algoritmo do programa para detecção da cavidade do receptor. Moléculas ligantes e um conjunto de regiões composta de todos os resíduos de aminoácidos (cadeia lateral), tendo pelo menos um átomo distante pelo menos 12 Å do centro do sildenafil, foram considerados flexíveis durante a simulação de *docking*. Foi então selecionada a conformação de cada composto usando suas melhores similaridades espaciais

com o sildenafil (compostos **2-5, 10-13, 15**, Tabela 1) e o tadalafil (compostos **21-25, 27, 29, 31, 33, 38, 41-43**, Tabela 1), que foram representadas pelas estruturas com as energias de interações mais favoráveis entre o posicionamento do ligante e a proteína (da Cunha et al., 2008a; da Cunha et al., 2008b; da Cunha et al., 2008c).

4.3 Cálculos ADME-Tox (absorção, distribuição, metabolismo, excreção e toxicidade)

Os perfis farmacocinéticos e toxicológicos das drogas propostas foram calculados por meio dos pacotes ADME e TOX boxes do programa Pharma Algorithms (Pharma Algorithms, 2008). Os parâmetros da “regra dos 5” de Lipinski e a área total de superfície polar (TPSA), que tem demonstrado se correlacionar com absorção de drogas pelo organismo, foram obtidos pelo uso do programa Molinspiration (Molinspiration Cheminformatics).

TABELA 1. pIC50 experimental, calibrado e predito (LOOCV 20% out) para uma série de inibidores PDE-5. Adaptado de (Antunes et al., 2008b).

Composto	Exp	Calib.	LOO CV	Leave-20%-out CV
1 (I)	5,70	-	-	-
2 (II)	6,48	6,43	7,62	7,60
3 (III)	7,57	7,34	7,12	7,02
4 (IV)	7,09	7,28	7,57	7,51
5 (V)	5,35	5,59	6,23	6,32
6 (VI)	6,00	5,91	6,00	6,07
7 (VII)	6,02	5,94	6,94	6,99
8 (VIII)	5,36	5,25	5,57	5,64
9 (IX)	6,11	6,14	6,92	6,96
10 (X, Sildenafil)	8,44	8,33	8,24	8,32
11 (XI)	8,72	8,86	7,87	7,86
12 (XII)	8,68	8,59	8,42	8,38
13 (XIII)	8,24	8,31	8,57	8,37
14 (8)	8,30	8,54	7,87	7,78

...continua...

TABELA 1, Cont.

15 (9)	9,22	9,13	8,77	8,70
16 (10)	7,40	7,34	7,21	7,08
17 (11)	8,18	8,33	8,44	8,45
18 (12)	8,00	7,94	7,55	7,53
19 (13)	8,00	7,93	8,12	7,97
20 (Vardenafil)	9,15	9,18	9,39	9,31
21 (2)	6,52	6,49	7,07	7,11
22 (<i>cis</i> -6a)	7,04	6,97	6,89	6,95
23 (<i>cis</i> -6b)	7,52	7,23	6,98	6,97
24 (<i>cis</i> -6c)	7,00	7,17	7,51	7,29
25 (<i>cis</i> -6d)	7,22	7,41	7,50	7,52
26 (<i>trans</i> -6d)	7,70	7,47	7,31	7,35
27 (<i>cis</i> -6e, análogo do Tadalafil)	8,10	8,07	7,94	7,60
28 (<i>trans</i> -6e)	8,30	8,13	7,90	7,70
29 (<i>cis</i> -6f)	7,30	7,27	7,14	7,08
30 (<i>trans</i> -6f)	7,30	7,35	7,26	7,24
31 (<i>cis</i> -6g)	7,30	7,55	7,65	7,61
32 (<i>trans</i> -6g)	7,70	7,61	7,46	7,39
33 (<i>cis</i> -6h)	6,04	6,23	6,61	6,53
34 (<i>trans</i> -6h)	6,52	6,29	6,09	6,04
35 (<i>trans</i> -6i)	6,00	6,18	7,39	7,41
36 (<i>trans</i> -7)	5,70	-	-	-
37 (<i>trans</i> -6j)	7,70	7,56	7,15	7,11
38 (<i>cis</i> -6k)	8,00	7,98	7,98	7,97
39 (<i>cis</i> -6l)	8,00	8,06	8,10	8,04
40 (<i>trans</i> -6l)	8,15	8,13	8,01	7,91
41 (<i>cis</i> -6m)	8,40	7,98	7,55	7,56
42 (<i>trans</i> -6m)	7,74	8,02	8,21	8,20
43 (<i>cis</i> -6n)	8,15	8,19	8,38	8,33
44 (<i>trans</i> -6n)	8,52	8,32	8,11	8,15
45 (<i>trans</i> -9a)	7,52	7,78	8,02	7,93
46 (<i>trans</i> -9b)	7,70	7,89	8,09	7,98
47 (<i>trans</i> -9c)	8,22	8,17	7,89	7,94
48 (<i>trans</i> -9d)	7,00	7,05	8,04	8,01

5 OBJETIVOS

Os objetivos deste estudo foram:

- a) aplicar o método MIA-QSAR e estudos de *docking* para prever as atividades inibitórias de PDE-5 para novos compostos propostos, derivados das estruturas do sildenafil (Viagra®), vardenafil (Levitra®) e tadalafil (Cialis®);
- b) determinar os perfis farmacocinéticos e toxicológicos dos fármacos propostos.

6 RESULTADOS E DISCUSSÃO

6.1 MIA-QSAR

De acordo com o protocolo MIA-QSAR, as estruturas químicas de 48 inibidores da PDE-5 foram agrupadas de tal maneira que as subestruturas coincidentes (o “esqueleto” da molécula) foram todas congruentes, isto é, apenas substituintes representavam a variância explicada dos dados (99,7%) – as bases de um método QSAR baseado na estrutura do ligante. Os descritores MIA (pixels = binários) foram, então, correlacionados com o bloco **y** (vetor coluna de bioatividades) pelo uso da regressão por PLS. O número ótimo de variáveis latentes (VLs) foi então escolhido pela análise conjunta de RMSEC e RMSECV. O número de componentes PLS utilizado foi escolhido de tal maneira que os erros de calibração e validação cruzada não diminuían significativamente depois de se adicionarem mais componentes PLS (8 VLs).

Dois *outliers* (compostos **1** e **36**) foram identificados pela análise dos maiores resíduos de validação cruzada ($pIC_{50(\text{exp})} - pIC_{50(\text{LOOcv})} = -1,67$ e $-2,25$ respectivamente) e, por meio do diagnóstico de Student (Figura 2). O valor da atividade do composto **1** (Zaprinast) é bastante controverso na literatura (Terrett et al., 1996; Haning et al., 2002; Daugan et al., 2003) e **36** é o único composto do conjunto de dados que contém um grupo metila ligado ao nitrogênio do grupo indol, podendo ser considerado uma extrapolação do modelo durante o processo de validação cruzada. Esses dois *outliers* foram removidos do conjunto de dados para proceder a calibração.

A calibração obteve um alto coeficiente de correlação entre o pIC_{50} experimental e o calibrado, isto é, r^2 de 0,970 (RMSEC = 0,16). Muitas variáveis explicativas podem ser correlacionadas com variáveis dependentes utilizando PLS, mas, no intuito de assegurar que a alta correlação obtida não foi obra do

acaso, o bloco y foi misturado aleatoriamente, de tal maneira que os compostos não mais correspondiam aos seus respectivos valores de atividades biológicas. Este procedimento resultou num r^2 significativamente menor ($0,74 \pm 0,11$ foi a média obtida de 10 medidas) do que r^2 real para 8 VLs, confirmando que a calibração não foi uma correlação casual, demonstrando-se, assim, a robustez do modelo.

O modelo foi validado por meio da validação cruzada *leave-one-out* (LOOCV), obtendo-se um q^2 de 0,694 (RMSECV = 0,53). A fim de se obter outra garantia da validade do modelo, o conjunto treino foi misturado aleatoriamente e separado dentro de cinco segmentos, e predições foram realizadas para cada segmento, utilizando-se os parâmetros de regressão das amostras remanescentes. Este procedimento foi repetido dez vezes e a média do r^2_{test} obtida foi de 0,679 (RMSEP = 0,56).

Os resultados estatísticos, juntamente com os valores calibrados e validados da Tabela 2, também ilustrados na Figura 3, sugerem que o modelo MIA-QSAR construído é capaz de fazer predições com sucesso. Estruturas candidatas a novas drogas (inibidores da PDE-5) podem, então, ser determinadas pelo uso de parâmetros de regressão. A aplicação do modelo QSAR, criado para prever as atividades de novos compostos, que possuem uma base estrutural comum à dos análogos do sildenafil, vardenafil e tadalafil (previamente calibrados), permite que as contribuições de subestruturas de pelo menos dois dos reconhecidos fármacos possam originar novos compostos, pois todos os pixels (descritores) estariam calibrados.

De acordo com esse procedimento, compostos foram propostos e seis deles (Figura 4) exibiram atividades biológicas razoavelmente elevadas (pIC_{50} predito >8), comparáveis às dos compostos mais ativos do conjunto de dados. Assim, tais compostos foram submetidos à avaliação de *docking*, a fim de confirmar as tendências calculadas pelo método MIA-QSAR por meio da

comparação de suas afinidades pelo receptor com os compostos de referência sildenafil, vardenafil e tadalafil.

É bem conhecido que as drogas mais adequadas não são apenas aquelas que possuem alta atividade biológica; elas devem obedecer a alguns requisitos farmacológicos, tais como capacidade de absorção e baixa toxicidade. Assim, parâmetros ADME-Tox foram calculados para se proporem compostos que levem em consideração sua adequação como medicamentos úteis para o tratamento da disfunção erétil.

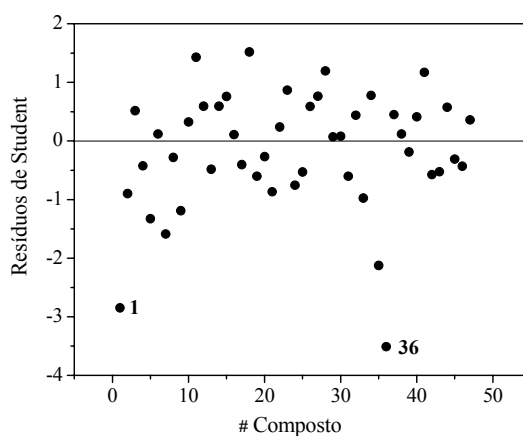


FIGURA 2. Diagnóstico residual de Student para a detecção de Outliers (Antunes et al., 2008b).

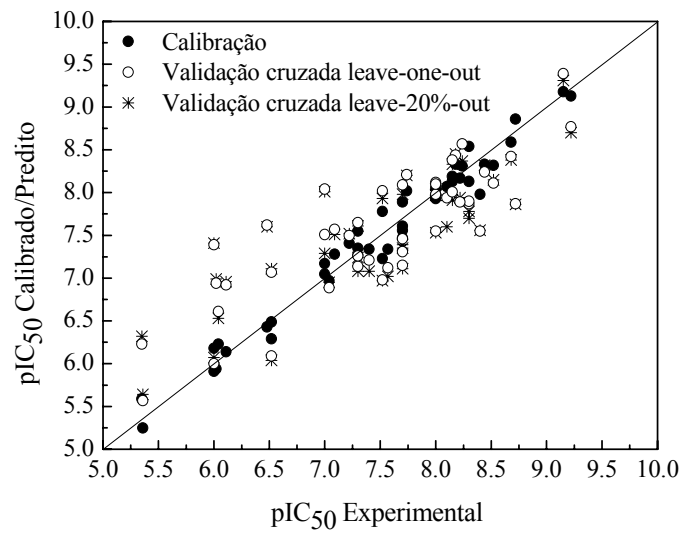


FIGURA 3. Plotamento dos pIC_{50} experimental versus calibrado/predito (Antunes et al., 2008b).

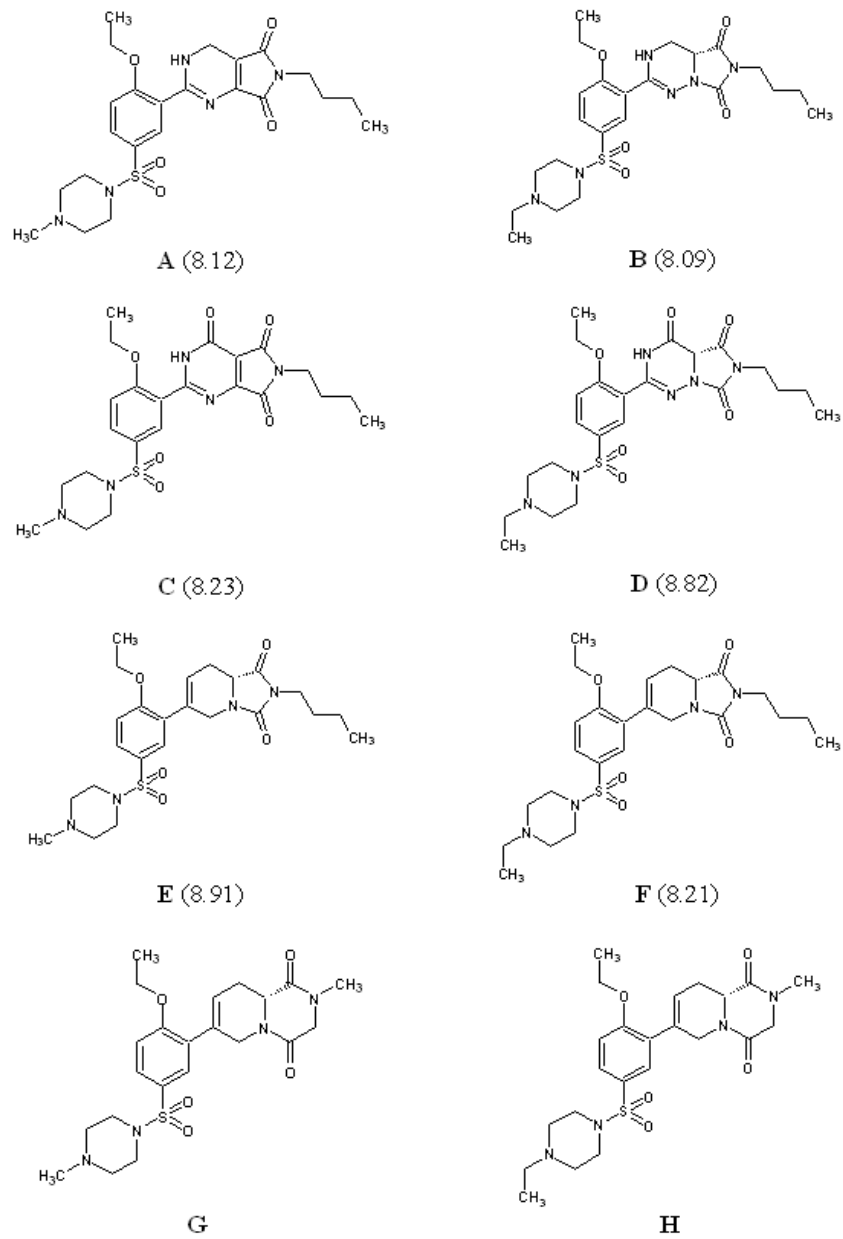


FIGURA 4. Inibidores PDE-5 propostos (pIC_{50} predito entre parênteses) (Antunes et al., 2008b).

6.2 Estudos de *Docking*

A função “scoring MolDock” foi aplicada para avaliar a afinidade entre o receptor PDE-5 e 22 inibidores selecionados, mais o sildenafil, tadalafil e vardenafil. Na Tabela 2 estão listadas as energias de interação calculadas. A relação entre as energias de interações calculadas (ΔE) e a atividade inibitória dos PDE-5 (valores de pIC_{50}), incluindo sildenafil e tadalafil, é mostrada na Figura 15. Uma correlação foi encontrada entre ΔE s e os pIC_{50} via análise de regressão linear ($r^2 = 0,56$). Esta razoável correlação demonstra que pIC_{50} pode ser obtido a partir dos valores de energias de interações modelados para os inibidores de PDE-5. Assim, a correlação entre dados teóricos e experimentais é satisfatória.

O modelo MIA-QSAR construído para a série de inibidores da PDE-5 da Tabela 1 foi validado e foi utilizado para prever a atividade biológica de seis novos compostos (A-F, Figura 4). A exatidão das predições e a alta potência dos compostos propostos foram reforçadas pela comparação de suas afinidades pelo receptor com os correspondentes valores para sildenafil, tadalafil e vardenafil (Tabela 2). As tendências nos valores de afinidade não correspondem exatamente aos valores de pIC_{50} obtidos pelo método MIA-QSAR, apesar de os desvios estarem próximos dos erros calculados (RMSECV). Entretanto, as afinidades calculadas são, no mínimo, parecidas (significativamente maiores em alguns casos) do que para o sildenafil, assim como previsto pela aproximação do ligante. Assim, os resultados do estudo de *docking* corroboraram os encontrados pelo estudo MIA-QSAR.

Dois compostos adicionais, G e H, foram posteriormente submetidos ao estudo de *docking* e as afinidades correspondentes foram medidas. As atividades desses compostos não foram calculadas pelo uso do modelo MIA-QSAR porque o anel fundido de cinco membros da estrutura base comum do conjunto treino foi substituído por um anel piperazinodiona, o qual não foi calibrado. Esta

modificação estrutural torna **G** e **H** mais similar ao tadalafil, que apresenta vantagens sobre o sildenafil (Daugan et al., 2003).

De acordo com a avaliação de *docking*, ambos, **G** e **H**, apresentaram valores de energias de interações maiores do que o vardenafil. Devido aos resultados satisfatórios para os oito compostos propostos, todos eles foram submetidos à avaliação ADME-Tox.

Os compostos **A-H** são miscelâneas dos compostos mais representativos do conjunto treino: sildenafil, vardenafil e análogo tadalafil. É conhecido que grupos polares ligados ao nitrogênio sulfonilpiperazínico, tais como **11** (grupo etoxila) e **12** (grupo amida), potencializam a bioatividade, devido a uma interação hidrofílica com o solvente. A principal interação que governa os valores de energia da Tabela 2 é a ligação de hidrogênio entre o sítio ativo da PDE-5 e os ligantes. O grupo etoxila de **A**, **B**, **D**, **E** e **F** interage com o sítio ativo da PDE-5, formando ligações de hidrogênio com Gln817. Além disso, o grupo carbonila da imida dos compostos **C** e **D** também interage de maneira similar com esse resíduo de aminoácido. Ligações de hidrogênio adicionais são observadas entre os grupo carbonila correspondentes dos compostos **A**, **B**, **D**, **E** e **F** e o resíduo Tyr612. O anel pirimidinona do composto **C** também interage com Tyr612 por meio de ligação de hidrogênio. Finalmente, os compostos **G** e **H** interagem com Gln817 por meio do grupo carbonila.

É importante ressaltar, entretanto, que outras interações intermoleculares podem ser relevantes. Nesse sentido, foram observadas interações π entre o anel central de todos os compostos (**A-H**) e Phe820. A partir dos dados da Tabela 3, torna-se evidente que os escores de energia dos compostos decrescem de acordo com a sequência: **C**>**D**>**A**>**B**>**F**>**E**>tadalafil>**H**>sildenafil>**G**>vardeafil, isto é, o composto **C** interage mais fortemente com PDE-5 do que os compostos remanescentes.

O composto **C** no interior do sítio ativo da enzima PDE-5 e os resíduos de aminoácidos que interagem por meio de ligação de hidrogênio são mostrados na Figura 5.

TABELA 2. Valores de energia (kcal mol^{-1}) entre PDE-5 e alguns inibidores da Figura 11 (Antunes et al., 2008b).

Compound	E (docking)	Compound	E (docking)
2	-131,21	23	-144,49
3	-143,24	24	-151,28
4	-108,01	25	-152,94
5	-112,16	27	-163,09
10	-176,90	29	-171,19
11	-177,90	31	-159,50
12	-182,24	33	-153,02
13	-186,50	38	-159,95
15	-176,03	41	-181,43
21	-148,93	42	-185,27
22	-159,44	43	-173,63

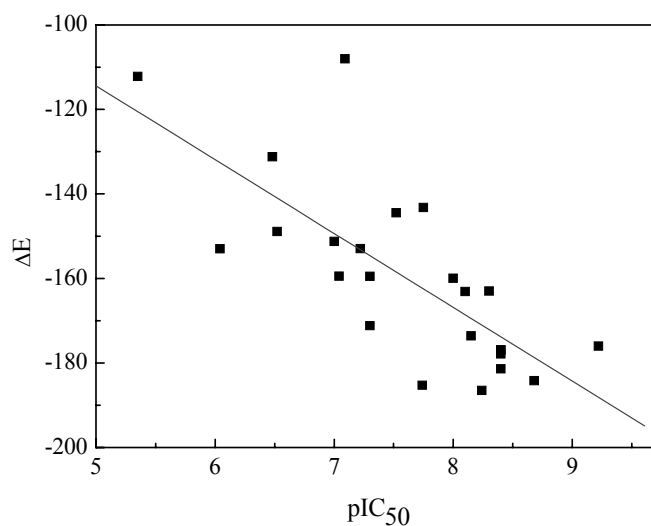


FIGURA 5. Correlação entre energias intermoleculares (ΔE s, kcal mol^{-1}) e bioatividades experimentais (pIC_{50}) (Antunes et al., 2008b).

TABELA 3. pIC₅₀ predito, energia de interação (E) e valores de energias de ligações de hidrogênio (kcal mol⁻¹) com a PDE-5, para sildenafil, tadalafil, vardenafil e compostos propostos (Antunes et al., 2008b).

Composto	pIC ₅₀ (MIA-QSAR)	E (docking)	Ligação de H
C	8,23	-175,70	-5,30
D	8,82	-172,25	-4,71
A	8,12	-169,70	-4,38
B	8,09	-167,56	-1,85
F	8,21	-165,60	-2,50
E	8,91	-163,90	-2,10
Tadalafil	8,30 ^a	-163,01	-2,26
H		-162,23	-2,74
Sildenafil	8,44 ^a	-161,63	-4,60
G		-161,09	-2,50
Vardenafil	9,15 ^a	-160,69	-5,01

^a Valores experimentais

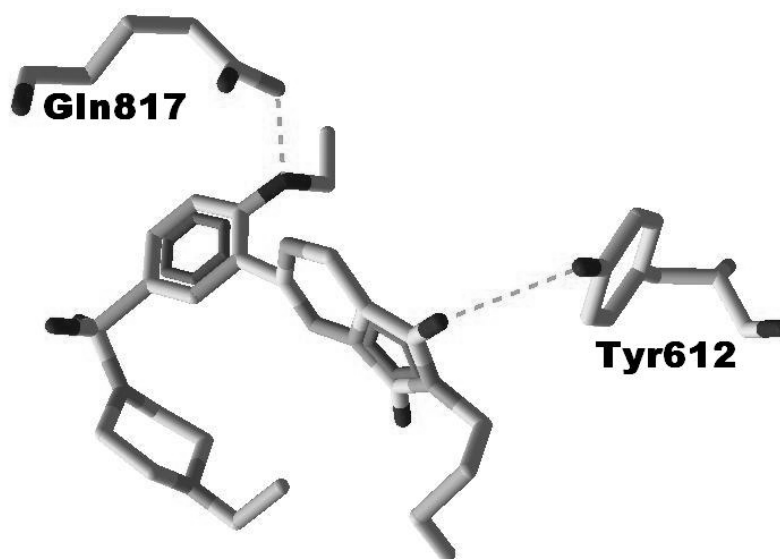


FIGURA 6. Estrutura do composto C docado dentro do sítio ativo da PDE-5. Os resíduos mostrados estão envolvidos em ligação de hidrogênio (Gln 817 e Tyr612) e interações π (Phe820) com o composto proposto (Antunes et al., 2008b).

6.3 Avaliação ADME-Tox

Lipinski et al. (2007) propuseram uma série de regras impondo limitações a alguns parâmetros, tais como $\log P$ (logaritmo do coeficiente de partição octanol/água), peso molecular e o número de grupos aceptores e doadores de ligações de hidrogênio, conhecidas como “regra dos cinco”. A regra diz que, para que um composto apresente-se como um fármaco adequado, deve ter: $\log P \leq 5$, peso molecular ≤ 500 , número de aceptores de ligações de hidrogênio ≤ 10 e número de doadores de ligações de hidrogênio ≤ 5 . Moléculas que violarem mais do que uma dessas regras podem ter problemas com biodisponibilidade.

Pelos dados da Tabela 4 observa-se que os compostos de referência não violam nenhuma dessas regras, enquanto os compostos propostos **B**, **C** e **D** apresentam peso molecular maior e maior número de aceptores de ligações de hidrogênio do que o permitido para considerá-los com perfil de bons fármacos. Também, tais compostos possuem altos valores de TPSA (área superficial polar total da molécula), sugerindo que suas biodisponibilidades orais podem ser piores do que as dos compostos remanescentes da Tabela 4 (a biodisponibilidade oral é inversamente proporcional à TPSA). Entretanto, os compostos preditos **A**, **E** e **F** não violaram nenhuma regra e exibiram baixos valores de TPSA (especialmente **E** e **F**) e, portanto, são supostamente mais biodisponíveis do que o sildenafil e vardenafil.

Isto é, claramente, resultado da remoção de um grupo carbonila e de um átomo de nitrogênio do anel de seis membros fundidos. Modificações do sistema indol-1,3(2*H*)-diona em **E** e **F**, pela troca do anel de cinco membros por uma piperazinodiona (tal como no tadalafil, estrutura da Figura 1), resulta nos compostos **G** e **H**, que possuem parâmetros da “regra de cinco” e TPSA similares aos dos compostos **E** e **F**. Essas modificações mostraram ser benéficas durante o desenvolvimento do tadalafil (Daugan et al., 2003).

Fica evidente, a partir da análise dos dados da Tabela 4, que a droga de última geração tadalafil exibe peso molecular, números de aceptores de ligações de hidrogênio e TPSA significativamente menores do que todos os compostos da Tabela 4, os quais, supostamente, contribuem para as conhecidas vantagens sobre o sildenafil e o vardenafil. Além disso, o coeficiente de partição octanol/água do tadalafil é comparável ao do sildenafil, vardenafil, **E** e **F**, mas maior do que o de **A-D**, **G** e **H**. Estes últimos compostos estão mais distantes de violarem a regra para $\log P$, mas eles deveriam ter razoável lipofilicidade para permear nas membranas.

Uma análise mais específica dos parâmetros farmacológicos pode ser mais bem detalhada por meio de uma completa avaliação ADME-Tox.

TABELA 4. Resultados dos cálculos para a “regra dos cinco” de Lipinski e TPSA, para os compostos de referência e os propostos. Adaptado de (Antunes et al., 2008b).

Parâmetro ^a	Sildenafil I	Vardenafil I	Tadalafil I	A	B	C	D	E	F	G
$\log P$	2,511	2,557	2,360	0,98	1,25	1,85	0,57	2,39	2,77	0,41
P.M.	474,6	488,6	389,4	489, 6	506, 6	503, 6	520, 6	490, 6	504, 6	462, 6
# A.L.H.	10	10	7	10	11	11	12	9	9	9
# D.L.H.	1	1	1	1	1	1	1	0	0	0
TPSA	113,4	112,9	74,9	113, 3	114, 9	134, 7	131, 9	90,5	90,5	90,5

^a $\log P$: logaritmo do coeficiente de partição octanol/água; P.M. peso molecular; # A.L.H e # D.L.H: número de aceptores e doadores de ligação de hidrogênio, respectivamente; TPSA: área superficial polar total da molécula.

Os dados da Tabela 5 retratam alguns parâmetros específicos relacionados com absorção, distribuição, metabolismo, excreção e toxicidade para os compostos de referência e os preditos. No geral, todos os compostos apresentaram vantagens e desvantagens, quando comparados uns com os outros.

Nenhuma diferença marcante nos efeitos sobre a saúde e no perfil de toxicidade em roedores foi observada entre os compostos. Entretanto, os parâmetros relacionados à absorção chamaram a atenção, já que os parâmetros calculados para os promissores compostos **E** e **F** indicam que eles devam ser tão solúveis em água e em tampão quanto os compostos de referência e são, presumivelmente, mais biodisponíveis oralmente e devem ter uma absorção significativamente maior do que o sildenafil e vardenafil. Além disso, esses valores são comparáveis com aqueles obtidos para o inibidor de PDE-5 de última geração, o tadalafil. Comportamento similar foi observado para os compostos **G** e **H**.

TABELA 5. Parâmetros ADME-TOX^a calculados para os compostos propostos e os de referência. Adaptado de (Antunes et al., 2008b).

ADME-Tox	Sildenafil	Vardenafil	Tadalafil	A	B	C	D	E	F	G	H
Sol. H ₂ O (mg mL ⁻¹)	0,43	0,86	0,088	0,069	0,023	0,17	0,086	0,18	0,10	0,51	0,32
Log sol. pH 1.7 (estômago)	-0,22	0,44	-3,65	-0,09	-0,45	-2,36	-0,82	-0,54	-0,74	-0,16	-0,32
Log sol. pH 4.6 (duodeno)	-1,15	-0,62	-3,65	-0,09	-0,44	-2,47	-1,59	-1,55	-1,49	-1,05	-1,00
Log sol. pH 6.5 (intestino delgado)	-2,74	-2,28	-3,65	-1,24	-2,94	-2,21	-3,31	-3,15	-3,21	-2,66	-2,70
Log sol. pH 7.4 (sangue)	-2,93	-2,66	-3,65	-2,32	-3,92	-1,59	-3,69	-3,40	-3,59	-2,90	-3,08
Log sol. pH 8.0 (colon)	-2,81	-2,73	-3,65	-2,91	-4,20	-1,03	-3,75	-3,43	-3,66	-2,94	-3,15
% Biodisponibilidade oral	30-70	<30	30-70	<30	<30	<30	<30	30-70	30-70	30-70	30-70
Absorção (cm s ⁻¹)	1,58×10 ⁻⁴	1,23×10 ⁻⁴	1,85×10 ⁻⁴	1,9×10 ⁻³	1,44×10 ⁻⁴	0,62×10 ⁻⁴	1,00×10 ⁻⁴	3,61×10 ⁻⁴	3,91×10 ⁻⁴	1,09×10 ⁻⁴	1,26×10 ⁻⁴
Distribuição (L kg ⁻¹) ^b	1,1(93%)	1,9(92%)	2,1(93%)	2,4(76%)	2,4(73%)	0,3(85%)	1,6(71%)	1,9(82%)	1,1(94%)	1,4(78%)	1,5(81%)
LD50 rato/camundongo (mg kg ⁻¹ , oral)	540/470	1100/950	570/1000	180/360	1300/990	400/490	1100/1200	900/850	950/880	570/720	760/740
LD50 rato/camundongo (mg kg ⁻¹ , intraperitoneal)	60/84	170/240	750/460	25/35	200/280	39/51	240/310	98/210	95/200	750/200	92/190
LD50 camundongo (mg kg ⁻¹ , intravenoso)	45	72	120	20	84	27	100	66	61	74	68
LD50 camundongo (mg kg ⁻¹ , subcutâneo)	180	430	420	57	330	76	330	350	360	290	390
Teste Ames (genotoxicidade, %)	0,014	0,006	0,001	0,002	0,002	0,003	0,002	0	0	0	0
Prob. de efeito no sangue	0,55	0,53	0,49	0,91	0,44	0,92	0,45	0,48	0,46	0,78	0,77
Prob. no sistema cardiovascular	0,93	0,92	0,94	0,95	0,93	0,94	0,91	0,96	0,96	0,96	0,96
Prob. no sistema gastrointestinal	0,86	0,86	0,49	0,99	0,97	0,93	0,34	0,73	0,76	0,77	0,76
Prob. de efeito nos rins	0,13	0,13	0,98	0,32	0,09	0,30	0,08	0,31	0,31	0,34	0,33
Prob. de efeito no fígado	0,40	0,39	0,99	0,76	0,49	0,74	0,47	0,38	0,38	0,37	0,37
Prob. de efeito no intestino	0,65	0,74	0,14	0,79	0,52	0,85	0,22	0,63	0,61	0,43	0,42

^a De acordo com os cálculos, nenhum composto é transportado pelo transportador intestinal para peptídeos (PepT1) e pela proteína apical transportadora de sódio dependente de ácido biliar (ASBT).

^b A ligação à proteína plasmática é dada entre parêntesis.

7 CONCLUSÕES

O modelo MIA-QSAR, construído a partir de uma série de análogos de sildenafil, vardenafil e tadalafil, foi utilizado para prever as atividades biológicas de novos inibidores de PDE-5, sendo validado e corroborado por estudos de *docking*. Estruturas químicas foram propostas baseadas numa estratégia combinatória, em que oito compostos resultantes deste estudo apresentaram-se altamente potentes. Os perfis ADME-Tox apresentaram-se bastante satisfatórios para quatro desses compostos, representados pelas letras **E**, **F**, **G** e **H**, assim, sugerindo que eventuais sínteses de inibidores de PDE-5 sejam amparadas por este estudo.

8 REFERÊNCIAS BIBLIOGRÁFICAS

ACD/CHEMSKETCH VERSION 8.17. Toronto: Advanced Chemistry Development, 2005. Software.

ANTUNES, J. E.; FREITAS, M. P.; da CUNHA, E. F. F.; RAMALHO, T. C.; RITTNER, R. In silico prediction of novel Phosphodiesterase type-5 inhibitors derived from Sildenafil, Vardenafil and Tadalafil. **Bioorganic & Medicinal Chemistry**, Amsterdam, v. 16, p. 7599-7606, Aug. 2008b.

DA CUNHA, E. F. F.; RAMALHO, T. C.; REYNOLDS, R. C. Binding Mode Analysis of 2,4-diamino-5-methyl-5-deaza-6-substituted Pteridines with *Mycobacterium tuberculosis* and Human Dihydrofolate Reductases. *Journal of Biomolecular Structure & Dynamics*, New Jersey, v. 25, p. 377-386, Feb. 2008a.

DA CUNHA, E. F. F.; RAMALHO, T. C.; SOUZA, T. C. S.; JOSA, D.; CAETANO, M. S. Molecular modelling of *Mycobacterium tuberculosis* acetolactate synthase catalytic subunit and its molecular docking study with inhibitors, **Molecular Simulation**, New York, v. 34, p. 707-713, June 2008b.

DA CUNHA, E. F. F.; RAMALHO, T. C.; SOUZA, T. C. S.; JOSA, D.; CAETANO, M. S. Hypothesis Paper: Homology Modeling of Wild-type, D516V, and H526L *Mycobacterium Tuberculosis* RNA Polymerase and Their Molecular Docking Study with Inhibitors, **Journal of Biomolecular Structure & Dynamics**, New Jersey, v. 25, p.373-376, Feb. 2008c.

DAUGAN, A.; GRONDIN, P.; RUAULT, C.; DE GOUVILE, A. C. M.; COSTE, H.; KIRILOVSKY, J.; HYAFIL, F.; LABAUDINIERRE, R. The discovery of tadalafil: a novel and highly selective PDE-5 Inhibitor. 1: 5,6,11,11a-Tetrahydro-1*H*imidazo[1',5':1,6]pyrido[3,4-*b*]indole-1,3(2*H*)-dione Analogues. **Journal of Medicinal Chemistry**, Washington, v. 46, p. 4525-4532, Dec. 2003.

HANING, H.; NIEWOHNER, U.; SCHENKE, T.; ES-SAYED, M.; SCHMIDT, G.; LAMPE, T.; BISCHOFF, E. Imidazo[5,1-*f*], and]triazin-4(3*H*)-ones, a new class of potent PDE 5 inhibitors, **Bioorganic & Medicinal Chemistry Letter**, Amsterdam, v. 12, p. 865-868, Mar. 2002.

LIPINSKI, C. A.; LOMBARDO, F.; DOMINY, B. W.; FEENEY, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Advanced Drug Delivery Reviews**, Amsterdam, v. 23, p. 3-25, Jan. 1997.

MILBURN, M. V.; KIM, S.; SCHLESSINGER, J.; BOLLAG, G. A glutamine switch mechanism for nucleotide selectivity by phosphodiesterases. **Molecular Cell**, Massachusetts, v. 15, p. 279-286, July 2004.

MOLINSPIRATION CHEMINFORMATICS. Bratislava, Slovak Republic. Disponível em. <<http://www.molinspiration.com>> Acesso em: 10 jan 2009. Software.

PHARMA ALGORITHMS. Toronto, 2008. Software.

SPARTANPRO 1.0.1. Wavefunction, Irvine, CA, 2001. Software.

TERRETT, N. K.; BELL, A. S.; BROWN, D.; ELLIS, P. Sildenafil (Viagratm), a potent and selective inhibitor of type 5 Cgmp Phosphodiesterase with utility for the treatment of male erectile dysfunction. **Bioorganic & Medicinal Chemistry Letter**, Amsterdam, v. 6, p. 1819-1824, Aug. 1996.

THOMSEN, R.; CHRISTENSEN, M. H. MolDock: a new technique for high-accuracy molecular docking. **Journal of Medicinal Chemistry**, Washington, v. 49, p. 3315-3321, June 2006.

ZHANG, K. Y. J.; CARD, G. L.; SUZUKI, Y.; ARTIS, D. R.; FONG, D.; GILLETTE, S.; HSIEH, D.; NEIMAN, J.; WEST, B. L.; ZHANG, C.; MILBURN, M. V.; KIM, S.; SCHLESSINGER, J.; BOLLAG, G. A glutamine switch mechanism for nucleotide selectivity by phosphodiesterases. **Molecular Cell**, Massachusetts, v. 15, p. 279-286, July 2004.