



**DANIELA RODRIGUES SILVA**

**MODELOS DE QSAR APLICADOS À ATIVIDADE  
ANTIDEPRESSIVA DE DERIVADOS DE HEXAHIDRO-  
PIRROLO-ISOQUINOLINA**

**LAVRAS – MG**

**2017**

**DANIELA RODRIGUES SILVA**

**MODELOS DE QSAR APLICADOS À ATIVIDADE ANTIDEPRESSIVA DE  
DERIVADOS DE HEXAHIDRO-PIRROLO-ISOQUINOLINA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agroquímica, área de concentração em Química/Bioquímica, para a obtenção do título de Mestre.

Prof.<sup>a</sup> Dra. Elaine Fontes Ferreira da Cunha  
Orientadora

**LAVRAS - MG  
2017**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca  
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Silva, Daniela Rodrigues.

Modelos de QSAR aplicados à atividade antidepressiva de  
derivados de hexahidro-pirrolo-isoquinolina / Daniela Rodrigues  
Silva. - 2017.

77 p. : il.

Orientador(a): Elaine Fontes Ferreira da Cunha.

.  
Dissertação (mestrado acadêmico) - Universidade Federal de  
Lavras, 2017.

Bibliografia.

1. Depressão. 2. Inibição dual H3/SERT. 3. QSAR-4D. I. da  
Cunha, Elaine Fontes Ferreira. . II. Título.

**DANIELA RODRIGUES SILVA**

**MODELOS DE QSAR APLICADOS À ATIVIDADE ANTIDEPRESSIVA DE  
DERIVADOS DE HEXAHIDRO-PIRROLO-ISOQUINOLINA**

**QSAR MODELS APPLIED TO THE ANTIDEPRESSANT ACTIVITY OF  
HEXAHYDRO-PYRROLO-ISOQUINOLINE COMPOUNDS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Agroquímica, área de concentração em Química/Bioquímica, para a obtenção do título de Mestre.

APROVADA em 10 de março de 2017.

Prof. <sup>a</sup> Dra. Ana Paula Guimarães	UFV
Prof. Dr. Teodorico de Castro Ramalho	UFLA

Prof.<sup>a</sup> Dra. Elaine Fontes Ferreira da Cunha  
Orientadora

**LAVRAS – MG  
2017**

## AGRADECIMENTOS

À Universidade Federal de Lavras, ao Departamento de Química, ao Programa de Pós-Graduação em Agroquímica e, especialmente, ao Laboratório de Modelagem Molecular, por possibilitarem o desenvolvimento desse trabalho.

À CAPES, pela concessão da bolsa de mestrado.

À professora Elaine, pela orientação, confiança, paciência, ensinamentos e apoio constantes. Foi uma honra trabalhar com você. Espero um dia poder retribuir tudo o que você fez por mim.

Aos professores Teodorico e Matheus, pela confiança, ensinamentos e apoio.

Aos companheiros do Laboratório de Modelagem Molecular, pelo suporte e momentos compartilhados. Sem dúvida vocês tornaram a caminhada muito mais prazerosa. E, de forma especial, às “meninas do QSAR”, Letícia Assis, Letícia Garcia, Giovanna, Tamiris e Mariene, pela amizade, esse trabalho não seria realizado sem o auxílio de vocês.

Aos amigos do coração, muito obrigada por fazerem parte da minha vida. Vocês são meus grandes exemplos. Se algum dia me tornar metade do que vocês são, já é o suficiente para mim.

Ao Lucas, pelo apoio, compreensão, amizade e carinho. Meu companheiro, meu braço direito e, acima de tudo, meu melhor amigo. Amo muito você!

À minha família como um todo, pelo apoio e por sempre acreditarem em mim.

Às minhas irmãs, Natalia e Ana Luisa, pela amizade e por estarem presentes em todos os momentos da minha vida.

Aos meus pais, Maria de Lourdes e Luiz, pela vida, confiança e pelos bons exemplos. Sem vocês eu não estaria onde estou.

À Deus, por tudo que Ele me proporcionou e pela força para concluir mais essa etapa.

## RESUMO

A depressão é uma desordem psicológica comum que aflige milhões de pessoas no mundo todo. Sua etiologia ainda não está completamente elucidada, porém sabe-se que os sistemas monoaminérgicos possuem papel fundamental e que, atualmente, constituem as bases do seu tratamento. Entretanto, apesar do número de antidepressivos disponíveis no mercado, esses medicamentos falham em atender uma parcela dos pacientes, são associados à efeitos adversos indesejados, além de causar uma melhora clínica apenas após semanas de tratamento. Portanto, a busca por novos compostos ativos se faz relevante, tanto para ampliar as opções terapêuticas como para propiciar uma melhor compreensão da fisiopatologia da doença. Nesse sentido, a inibição dual do transportador de serotonina (SERT) e do receptor de histamina H<sub>3</sub> surge como alternativa interessante, pois, além de aumentar os níveis de serotonina, auxilia no alívio de alguns sintomas cognitivos. Dessa forma, esse trabalho tem como objetivo desenvolver relações quantitativas em quatro dimensões entre a estrutura de derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória para os receptores H<sub>3</sub> e SERT, com o intuito de identificar características estruturais importantes para a atividade biológica desses compostos. Os modelos de QSAR para ambos receptores apresentaram parâmetros estatísticos satisfatórios que asseguram sua robustez e preditividade (H<sub>3</sub>R: R<sup>2</sup> = 0,808, Q<sup>2</sup> = 0,736, r<sup>2</sup><sub>m</sub> = 0,670 e R<sup>2</sup><sub>p</sub> = 0,771; SERT: R<sup>2</sup> = 0,818, Q<sup>2</sup> = 0,710, r<sup>2</sup><sub>m</sub> = 0,580 e R<sup>2</sup><sub>p</sub> = 0,728), além de grupos farmacofóricos importantes para a atividade biológica. A ligação dos compostos em estudo ao SERT também foi avaliada pelo ancoramento do composto mais ativo ao sítio de ligação do transportador, onde foi possível sugerir a existência de interações intermoleculares importantes. Portanto, em conjunto, essas informações podem ser utilizadas para guiar o planejamento de novas moléculas bioativas para o tratamento da depressão.

**Palavras-chave:** Depressão. Antidepressivos. Receptor de histamina H<sub>3</sub>. Transportador de serotonina. Inibição dual H<sub>3</sub>/SERT. (RI) QSAR-4D.

## ABSTRACT

Depression is a common psychological disorder that affects millions of people worldwide. Its etiology has not yet been fully elucidated, but it is known that monoaminergic systems play a critical role and are currently the basis of the treatment. However, despite the number of antidepressants available, these medicines fail to treat a portion of the patients, are associated with some undesirable side effects, and cause clinical improvement only after weeks of treatment. Therefore, the search for new active compounds becomes relevant, both to broaden the therapeutic options and to provide a better understanding of the pathophysiology of the disease. In this sense, the dual inhibition of serotonin transporter (SERT) and histamine H<sub>3</sub> receptor appears as an interesting alternative, because it assists in increasing the serotonin levels as well as in alleviating some cognitive symptoms. Thus, this work aims to develop quantitative relationships in four dimensions between the structure of hexahydro-pyrrolo-isoquinoline compounds and their respective inhibitory constants for H<sub>3</sub> and SERT, in order to identify important structural characteristics for the biological activity of these compounds. The QSAR models for both receptors have demonstrated satisfactory statistical parameters that assure their robustness and predictability (H<sub>3</sub>R: R<sup>2</sup> = 0.808, Q<sup>2</sup> = 0.736, r<sup>2</sup><sub>m</sub> = 0.670 and R<sup>2</sup><sub>p</sub> = 0.771; SERT: R<sup>2</sup> = 0.818, Q<sup>2</sup> = 0.710, r<sup>2</sup><sub>m</sub> = 0.580 and R<sup>2</sup><sub>p</sub> = 0.728), and pharmacophoric groups important for biological activity. The binding of the tested compounds to SERT was also evaluated by docking the most active compound to the transporter binding site, where it was possible to suggest the existence of important intermolecular interactions. Therefore, together, this information can be used to guide the development of new bioactive molecules for the treatment of depression.

**Keywords:** Depression. Antidepressants. Histamine H<sub>3</sub> receptor. Serotonin transporter. Dual inhibition H<sub>3</sub>/SERT. (RI) 4D-QSAR.

## LISTA DE FIGURAS

Figura 1 - A) Via de síntese da serotonina. B) Esquema de síntese, armazenamento, liberação, recaptação e inativação da serotonina.....	19
Figura 2 - Exemplos de compostos de cada classe de antidepressivos. ....	22
Figura 3 - A) Estrutura geral dos compostos derivados de hexahidro-pirrolo-isoquinolina desenvolvidos por Apodaca e colaboradores. B) Exemplo de inibidor dual SERT/H <sub>3</sub> com farmacóforo H <sub>3</sub> em vermelho.....	24
Figura 4 - Átomos que compõem os dez alinhamentos analisados utilizando a molécula 1A como referência.....	46
Figura 5 - (Esquerda) Descritores do alinhamento 4 representados no composto 1B, o mais ativo da série. (Direita) Sobreposição dos compostos do conjunto de dados. ....	52
Figura 6 - Os descritores do alinhamento 4 representados nos compostos com respectivo maior valor de frequência de ocupação.....	54
Figura 7 - Domínio de aplicabilidade para o modelo de QSAR do H <sub>3</sub> R representado pelo gráfico de influência <i>versus</i> resíduo de Student.....	57
Figura 8 - (Esquerda) Descritores do alinhamento 7 representados no composto 11A, o mais ativo da série. (Direita) Sobreposição dos compostos do conjunto de dados. ....	59
Figura 9 - Os descritores do alinhamento 7 representados nos compostos com respectivo maior valor de frequência de ocupação.....	61
Figura 10 - Domínio de aplicabilidade para o modelo de QSAR do SERT representado pelo gráfico de influência <i>versus</i> resíduo de Student.....	62
Figura 11 - (Esquerda) Estrutura tridimensional do hSERT com composto 11A ancorado no sítio de ligação central. (Direita) Principais resíduos de aminoácido do sítio ativo juntamente com os descritores do modelo de QSAR-4D. ....	63



## LISTA DE QUADROS

Quadro 1 - Compostos derivados de hexahidro-pirroló-isoquinolina e seus respectivos valores de constante inibitória, expressa como $pK_I$ , para receptores $H_3$ e SERT (negrito). .....	34
--	----

## LISTA DE TABELAS

Tabela 1 - Parâmetros estatísticos avaliados na metodologia QSAR.....	47
Tabela 2 - Parâmetros estatísticos médios da calibração, validação interna e externa dos modelos provenientes dos dez alinhamentos testados para o receptor H <sub>3</sub> .....	50
Tabela 3 - Matriz de correlação cruzada entre os descritores do alinhamento 4.....	52
Tabela 4 - Valores experimentais e preditos da constante inibitória juntamente com o resíduo dos compostos do conjunto de treinamento e teste (*) para o H <sub>3</sub> R.....	55
Tabela 5 - Parâmetros estatísticos médios da calibração, validação interna e externa dos modelos provenientes dos dez alinhamentos testados para o SERT.....	58
Tabela 6 - Matriz de correlação cruzada entre os descritores do alinhamento 7.....	59
Tabela 7 - Valores experimentais e preditos da constante inibitória juntamente com o resíduo dos compostos do conjunto de treinamento e teste (*) para o SERT.....	61
Tabela 8 - Energias obtidas do ancoramento molecular do composto 11A no sítio de ligação central do hSERT.....	62
Tabela 9 - Tabela completa com parâmetros estatísticos da calibração, validação interna e externa dos melhores modelos provenientes dos dez alinhamentos testados para H <sub>3</sub> R.....	72
Tabela 10 - Frequência de ocupação dos GCODs da eq. 9 para todos os compostos do conjunto de dados. ....	73
Tabela 11 - Tabela completa com parâmetros estatísticos da calibração, validação interna e externa dos melhores modelos provenientes dos dez alinhamentos testados para o SERT.....	76
Tabela 12 - Frequência de ocupação dos GCODs da eq. 10 para todos os compostos do conjunto de dados. ....	77

## LISTA DE SIGLAS

5-HIAA	5-hidroxindol acetaldeído
5-TH	5-hidroxitriptamina (serotonina)
CEP	Conformational Ensemble Profile
CoMFA	Comparative Molecular Field Analysis
DA	Domínio de aplicabilidade
DSM	Diagnostic and Statistical Manual of Mental Disorders
FDA	Food and Drug Administration
GA	Genetic Algorithm
GCOD	Grid Cell Occupancy Descriptor
GFA	Genetic Function Approximation
GPCR	G Protein-Coupled Receptor
H <sub>3</sub> R	Histamine H <sub>3</sub> Receptor
HPA	Hipotálamo-Pituitária-Adrenal
hSERT	Human Serotonin Transporter
ICD	International Statistical Classification of Diseases and Related Health Problems
IMAO	Inibidores da Monoamina Oxidase
IPE	Interaction Pharmacophore Element
LOF	Lack-Of-Fit
LOO	Leave-One-Out
LMO	Leave-Many-Out
LNO	Leave-N-Out
LSE	Least-Squares Error
MAO	Monoamina Oxidase
MAO-A	Monoamina Oxidase isoforma A
MAO-B	Monoamina Oxidase isoforma B
MDD	Major Depressive Disorder
MVD	Molegro Virtual Docker
NE	Noradrenalina
OMS	Organização Mundial da Saúde
PDB	Protein Data Bank

PLP	Piecewise Linear Potential
PLS	Partial Least Square
QSAR	Quantitative Structure-Activity Relationship
(RD) QSAR	Receptor Dependent QSAR
(RI) QSAR	Receptor Independent QSAR
SDM	Simulação de Dinâmica Molecular
SERT	Serotonin Transporter
SNC	Sistema Nervoso Central
SNP	Sistema Nervoso Periférico
SNRI	Selective Noradrenaline-Serotonin Reuptake Inhibitor
SSRI	Selective Serotonin Reuptake Inhibitor
TCA	Tricyclic Antidepressants
TM	Transmembranar
VL	Variáveis Latentes

## SUMÁRIO

1	INTRODUÇÃO .....	14
2	REFERENCIAL TEÓRICO .....	16
2.1	Depressão.....	16
2.1.1	Hipótese monoamínica .....	17
2.1.2	Serotonina .....	18
2.2	Antidepressivos.....	20
2.2.1	Antagonismo dual do transportador de serotonina e receptor de histamina H <sub>3</sub> .....	23
2.3	Química computacional no desenvolvimento de fármacos.....	24
2.4	Métodos quantitativos de relação estrutura-atividade.....	25
2.4.1	Desenvolvimento dos métodos QSAR.....	26
2.4.2	QSAR-4D.....	29
2.4.3	Regressão pelo método dos mínimos quadrados parciais.....	30
2.4.4	Aproximação da função genética .....	31
2.4.5	Validação dos modelos .....	31
3	METODOLOGIA.....	34
3.1	Dados Biológicos.....	34
3.2	Construção e minimização dos compostos.....	44
3.3	Conjuntos de treinamento e teste.....	44
3.4	Simulação de Dinâmica Molecular.....	45
3.5	Definição dos Elementos de Interação Farmacofórica.....	45
3.6	Definição do Alinhamento.....	45
3.7	Construção dos modelos QSAR-4D: Redução de dados e abordagem GFA.....	46
3.8	Validação dos Modelos.....	47
3.9	Domínio de aplicabilidade.....	48
3.10	Seleção da Conformação Representativa.....	49
3.11	Ancoramento molecular.....	49
4	RESULTADOS E DISCUSSÃO .....	50
4.1	Receptor de histamina H <sub>3</sub> .....	50
4.1.1	Definição do melhor alinhamento .....	50
4.1.2	Análise dos descritores .....	51
4.2	Transportador de serotonina.....	57
4.2.1	Definição do melhor alinhamento .....	57
4.2.2	Análise dos descritores .....	58
4.2.3	Ancoramento molecular .....	62
5	CONCLUSÃO.....	65
	REFERÊNCIAS .....	66
	APÊNDICE A .....	72
	APÊNDICE B.....	73
	APÊNDICE C .....	76
	APÊNDICE D .....	77

## 1 INTRODUÇÃO

A depressão é uma desordem psicológica comum que afeta milhões de pessoas pelo mundo todo (WHO, 2017). Estima-se que são gastos \$98,9 bilhões por ano nos EUA com a depressão, considerando apenas os gastos diretos com serviços médicos e farmacêuticos (GREENBERG et al., 2015; SOBOCKI et al., 2006). Sendo assim, a depressão representa um sério problema de saúde que aflige uma parcela considerável da população e que possui grande impacto socioeconômico (WHO, 2017).

A depressão é caracterizada por episódios recorrentes de humor depressivo, perda de interesse, perturbações no sono com consequente fadiga, dificuldade de concentração, entre outros sintomas. Embora seu mecanismo bioquímico ainda não esteja completamente elucidado, sabe-se que a neurotransmissão monoaminérgica, sobretudo envolvendo os sistemas serotoninérgico e noradrenérgico, possui participação fundamental (BONDY, 2002).

Os principais antidepressivos podem ser classificados em cinco categorias: i) inibidores da MAO (IMAO); ii) antidepressivos tricíclicos (TCA); iii) inibidores seletivos da recaptção de serotonina (SSRI); iv) inibidores da recaptção de serotonina-noradrenalina (SNRI) e v) antidepressivos atípicos. Atualmente, os SSRIs são considerados os agentes de primeira linha no tratamento da depressão, eles são resultado da busca por medicamentos tão eficientes quanto os TCAs, mas com perfil atenuado de efeitos colaterais. (MORENO; MORENO; SOARES, 1999; SHARMA; SANTRA; DUTTA, 2015). Houve um notável avanço no que se refere ao tratamento da depressão, entretanto, ainda existem algumas limitações. Uma delas é o atraso entre o início do tratamento e uma resposta positiva, pois a maioria dos fármacos causa uma melhora dos sintomas apenas após semanas de tratamento. Além disso, os medicamentos disponíveis têm taxa de resposta limitada, sendo que parte dos pacientes não apresentam melhora clínica (PYTKA et al., 2016; SHARMA; SANTRA; DUTTA, 2015). Portanto, faz-se necessário a realização de estudos adicionais objetivando o desenvolvimento de novos compostos para o tratamento da depressão, no intuito de aumentar as opções terapêuticas bem como propiciar uma melhor compreensão da doença.

Nesse contexto, compostos com propriedade inibitória dual para o receptor de histamina  $H_3$  e o transportador de serotonina (SERT) são uma alternativa interessante, visto que auxiliariam no alívio de alguns sintomas depressivos, como fadiga e distúrbios do sono, bem como na função serotoninérgica. A inibição do receptor  $H_3$  tem sido associada ao aumento na liberação de serotonina e, com o concomitante bloqueio do SERT, a reabsorção neuronal desse neurotransmissor será reduzida. Como resultado, maiores concentrações de serotonina

serão alcançadas, conduzindo assim a um efeito terapêutico melhorado. Com base nisso, Apodaca et al. (2012) desenvolveram uma série de compostos derivados de hexahidro-pirrolo-isoquinolina moduladores de receptores H<sub>3</sub> e SERT como possíveis candidatos à fármacos para o tratamento da depressão.

A descoberta e planejamento de novos fármacos é um processo complexo, laborioso e que demanda grandes investimentos. Dessa forma, os métodos computacionais são uma alternativa para simplificar o processo de planejamento de compostos bioativos de forma complementar às técnicas experimentais (KAPETANOVIC, 2008). Dentre as metodologias amplamente utilizados estão as relações quantitativas entre estrutura-atividade - QSAR. Métodos de QSAR são abordagens computacionais que buscam relacionar a estrutura química, diretamente ligada às propriedades físico-químicas, e a atividade biológica de certos compostos. Estas relações são expressas por modelos matemáticos que estabelecem uma correlação entre descritores moleculares de uma série de compostos análogos e a atividade biológica por eles desempenhada. Em química medicinal, esses métodos auxiliam, dentre outras coisas, na identificação de candidatos nocivos nas primeiras etapas de desenvolvimento, na otimização de compostos existentes e na predição de propriedades de compostos que ainda não foram testados (FERREIRA, 2002; TAVARES, 2004).

Utilizando esta metodologia, este trabalho visa desenvolver relações quantitativas entre a estrutura dos derivados de hexahidro-pirrolo-isoquinolina, propostas por Apodaca et al. (2012), e seus respectivos valores da constante inibitória para o receptor H<sub>3</sub> e SERT, com o objetivo de identificar características estruturais importantes para a atividade biológica desses compostos.

## 2 REFERENCIAL TEÓRICO

### 2.1 Depressão

Dados da Organização Mundial da Saúde (OMS) apontam que a depressão afeta mais de 300 milhões de pessoas no mundo todo (WHO, 2017). Qualquer um está sujeito a apresentar sintomas depressivos, porém, frequentemente eles começam em pessoas ainda jovens e atingem mais o sexo feminino. Além disso, a depressão se tornou a principal causa de anos perdidos por incapacidade (WHO, 2017).

Ademais, a depressão possui grande impacto econômico. Estima-se que são gastos por ano \$98,9 bilhões nos Estados Unidos e \$42 bilhões na Europa, considerando apenas os gastos diretos com serviços médicos e farmacêuticos. Sendo assim, a depressão representa um sério problema de saúde que aflige uma parcela considerável da população e que possui grande impacto social e econômico (GREENBERG et al., 2015; SOBOCKI et al., 2006).

A depressão é uma desordem psicológica comum caracterizada por sintomas como humor depressivo, perda de interesse ou prazer, alterações de peso e apetite, insônia, agitação ou retardação psicomotora, fadiga ou perda de energia, sentimentos de inutilidade, diminuição da capacidade de pensamento ou concentração e tendência suicida. De acordo com o perfil de sintomas apresentados, ela pode ser classificada em leve, moderada ou grave. O diagnóstico é geralmente baseado na avaliação clínica do paciente, de acordo com o DSM (do inglês, *Diagnostic and Statistical Manual of Mental Disorders*) ou pelo ICD (do inglês, *International Statistical Classification of Diseases and Related Health Problems*). A forma mais prevalente da depressão é o transtorno depressivo maior (MDD, do inglês *Major Depressive Disorder*), o qual apresenta um perfil mais severo com múltiplos episódios. Segundo o DSM-IV, o MDD é caracterizado por cinco dos nove sintomas citados acima, sendo um deles necessariamente o humor depressivo, e que persistem por mais de duas semanas (BONDY, 2002; EL-ALFY; ABOURASHED; MATSUMOTO, 2012; SHARMA; SANTRA; DUTTA, 2015).

A combinação de fatores ambientais e genéticos estão sendo associados à predisposição do indivíduo ao desenvolvimento da depressão. Dentre os cenários analisados, pode-se citar, por exemplo, situações de estresse e eventos traumáticos. Além disso, pessoas com casos de depressão na família têm a probabilidade maior de desenvolver a doença em algum momento da vida (BONDY, 2002; NESTLER et al., 2002; SAVEANU; NEMEROFF, 2012).



No pior dos casos, a depressão pode levar ao suicídio e também ao atentado à vida de outras pessoas, em que homicídio seguido de suicídio tem se tornado uma prática cada vez mais comum. Segundo a OMS, são perdidas quase 800,000 vidas por ano devido ao suicídio (WHO, 2017).

Em casos mais sérios, o tratamento da depressão é feito pela administração de antidepressivos, tópico que será discutido em mais detalhes posteriormente. Entretanto, nem todos os pacientes têm acesso ao tratamento, segundo dados da OMS, em alguns países a taxa é menor que 10% (WHO, 2017). Além disso, mesmo que a pessoa receba o tratamento, ainda existem algumas limitações. Por exemplo, o atraso entre o início do tratamento e uma resposta positiva, pois a maioria dos antidepressivos causa uma melhora dos sintomas apenas após semanas de tratamento. Ademais, alguns medicamentos possuem taxa de resposta limitada, ou seja, apenas uma porcentagem dos pacientes apresenta uma melhora clínica, além dos efeitos colaterais associados (APODACA et al., 2012; EL-ALFY; ABOURASHED; MATSUMOTO, 2012; SHARMA; SANTRA; DUTTA, 2015).

### **2.1.1 Hipótese monoamínica**

Embora a fisiopatologia da depressão ainda não esteja completamente elucidada, sabe-se que os sistemas monoaminérgicos exercem papel fundamental. Há cerca de cinquenta anos, foi proposta a teoria monoamínica da depressão, essa hipótese sugere que a doença estaria relacionada à deficiência de serotonina e noradrenalina em diferentes partes do cérebro. Portanto, o aumento da neurotransmissão dessas monoaminas aliviaria os principais sintomas depressivos (SCHILDKRAUT, 1965). Evidências experimentais dessa hipótese foram encontradas em modelos animais, bem como em seres humanos. Um exemplo é a administração de reserpina, um anti-hipertensivo que causa depleção pré-sináptica das monoaminas ao inibir seu transporte nas vesículas sinápticas, o que induz sintomas depressivos. Ao longo dos anos, essa teoria tem sido estendida, em que foram apontados os papéis de outros sistemas, como, por exemplo, dopaminérgicos, colinérgicos, glutamatérgicos, GABAérgicos e desregulação do eixo hipotálamo-pituitária-adrenal (HPA). No entanto, a hipótese monoamínica continua sendo a mais aceita e validada e constitui as bases do tratamento da depressão (BONDY, 2002; PYTKA et al., 2016).

A serotonina e a noradrenalina (NE) são neurotransmissores que desempenham papéis essenciais na regulação do humor, motivação, vigilância, percepção de dor e na função

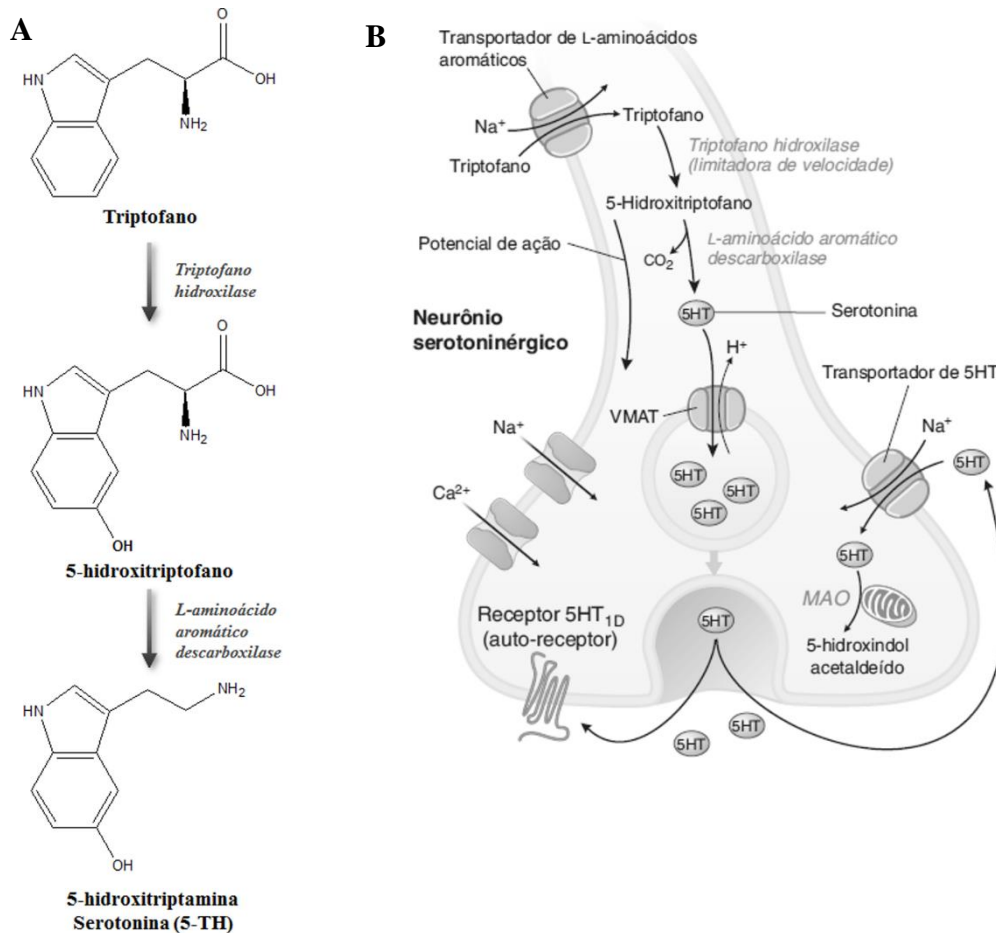
neuroendócrina. A síntese, secreção e metabolismo desses neurotransmissores ocorre de maneira semelhante (BONDY, 2002; SMITH; MARKS; LIEBERMAN, 2003).

### 2.1.2 Serotonina

A serotonina (5-hidroxitriptamina, 5-TH) é uma amina biogênica, mais especificamente uma indolamina, sintetizada a partir da hidroxilação e descarboxilação do aminoácido triptofano, o qual é obtido por meio da alimentação (Fig. 1A). Sua síntese ocorre no citosol do neurônio e, em seguida, a 5-TH é transportada em vesículas para armazenamento e liberação. A liberação das vesículas na fenda sináptica é desencadeada pelo impulso nervoso, que provoca a despolarização da membrana, fazendo com que os canais de cálcio se abram, permitindo assim um influxo de íons  $\text{Ca}^{2+}$  que causa a fusão da vesícula com a membrana neuronal e libera a 5-TH. Uma vez liberada, a 5-TH pode se ligar aos seus receptores pós-sinápticos ou aos autoreceptores pré-sinápticos, os quais regulam sua síntese e liberação. Grande parte da 5-TH liberada na fenda sináptica é transportada de volta para a célula pré-sináptica por meio do transportador de serotonina (SERT, do inglês *Serotonin Transporter*), a recaptação é acoplada ao co-transporte de íons  $\text{Na}^+$ . Após ser recaptada, a 5-TH pode seguir dois caminhos: ser novamente armazenada em vesículas para posterior uso na neurotransmissão ou sofrer inativação pela enzima monoaminaoxidase (MAO) (Fig. 1B) (NADAL-VICENS; CHYUNG; TURNER, 2009; SMITH; MARKS; LIEBERMAN, 2003).

A serotonina age por meio de uma variedade de receptores ligados à membrana, tanto no sistema nervoso central (SNC) e periférico (SNP), como em alguns tecidos não neuronais (por exemplo, o trato gastrointestinal, sistema cardiovascular e sangue). Até hoje foram identificados catorze subtipos diferentes de receptores de 5-TH, os quais são agrupados em sete famílias (5-TH<sub>1</sub> – 5-TH<sub>7</sub>). Com exceção da família 5-TH<sub>3</sub>, composta por canais iônicos regulados por ligante, todos os receptores são acoplados à proteína G (GPCRs, do inglês *G Protein-Coupled Receptors*). De forma geral, as classes de receptores 5-TH<sub>1</sub> (onde o 5-TH<sub>1D</sub> é um receptor pré-sináptico) e 5-TH<sub>5</sub> inibem a adenil ciclase, a classe 5-TH<sub>2</sub> aumenta a renovação do fosfatidilinositol, e as classes 5HT<sub>4</sub>, 5HT<sub>6</sub> e 5HT<sub>7</sub> estimulam a adenil ciclase. Os subtipos são expressos em diferentes regiões do cérebro (NADAL-VICENS; CHYUNG; TURNER, 2009; NICHOLS; NICHOLS, 2008; PYTKA et al., 2016).

Figura 1 - A) Via de síntese da serotonina. B) Esquema de síntese, armazenamento, liberação, recaptação e inativação da serotonina.



Fonte: Nadal-Vicens (2009)

Evidências experimentais apontam atividade reduzida dos sistemas serotoninérgicos em pacientes depressivos, em que ocorre uma diminuição tanto de transportadores como de alguns subtipos de receptores de serotonina. Além disso, também foram encontradas baixas concentrações de 5-hidroxiindol acetaldeído (5-HIAA), o principal metabólito da 5-TH, no fluido cerebrospinal de pacientes deprimidos. De fato, os inibidores seletivos da recaptação de serotonina (SSRI, do inglês *selective serotonin reuptake inhibitors*), que bloqueiam o SERT, representam um grupo bem-sucedido de antidepressivos, devido à alta eficiência e baixa incidência de efeitos colaterais. Enfatizando assim a importância dos sistemas serotoninérgicos no tratamento da depressão (DREVETS et al., 1999; MANN et al., 1996; PYTKA et al., 2016; SAVEANU; NEMEROFF, 2012).

## 2.2 Antidepressivos

Os antidepressivos têm em comum a capacidade de aumentar a disponibilidade sináptica de um ou mais neurotransmissores por meio da ação em receptores e enzimas específicas. Atualmente, os antidepressivos são agrupados preferencialmente segundo sua ação farmacológica, mas a classificação também pode ser feita a respeito da estrutura química. Os medicamentos antidepressivos aumentam a concentração de neurotransmissores principalmente por meio da inibição do metabolismo ou bloqueio da recaptação neuronal (MORENO; MORENO; SOARES, 1999; NADAL-VICENS; CHYUNG; TURNER, 2009).

Os principais medicamentos utilizados no tratamento da depressão podem ser classificados em cinco categorias: i) inibidores da MAO (IMAO); ii) antidepressivos tricíclicos (TCA); iii) inibidores seletivos da recaptação de serotonina (SSRI), iv) inibidores da recaptação de serotonina-noradrenalina (SNRI) e v) antidepressivos atípicos. Na Figura 2 estão representados alguns exemplos de cada classe. IMAO e TCA estão entre os primeiros antidepressivos desenvolvidos e ficaram conhecidos como antidepressivos tradicionais ou de primeira geração. O mecanismo de ação dos IMAO é baseado na inibição da degradação pela MAO, enzima mitocondrial responsável pela degradação de monoaminas. Com a redução da atividade da MAO, tem-se um aumento da concentração desses neurotransmissores, causando o efeito antidepressivo desejado. Os IMAO são classificados de acordo com sua especificidade às duas isoformas da MAO (MAO-A e MAO-B) e a reversibilidade de sua ligação. O efeito colateral mais acentuado dessa classe é a crise hipertensiva, resultado da inibição do catabolismo de tiraminas provenientes da dieta. Além disso, esses compostos também podem induzir inúmeras interações medicamentosas (NADAL-VICENS; CHYUNG; TURNER, 2009; SHARMA; SANTRA; DUTTA, 2015).

Os antidepressivos das classes TCA, SSRI e SNRI atuam pelo bloqueio da recaptação de monoaminas, o que aumenta a concentração desses neurotransmissores na fenda sináptica e, com isso, causa uma maior ativação dos receptores pós-sinápticos. Os TCAs possuem esse nome devido à sua estrutura química geral, que é composta por três anéis, dois anéis aromáticos conectados à um cicloheptano. Esses compostos atuam principalmente pela inibição da recaptação de serotonina e noradrenalina, porém o mecanismo de ação dessa classe de antidepressivos não está completamente elucidado. Os principais efeitos adversos dos TCAs estão associados à sua não seletividade, ou seja, esses compostos se ligam a outros receptores além dos seus alvos terapêuticos. Os TCAs podem afetar receptores muscarínicos

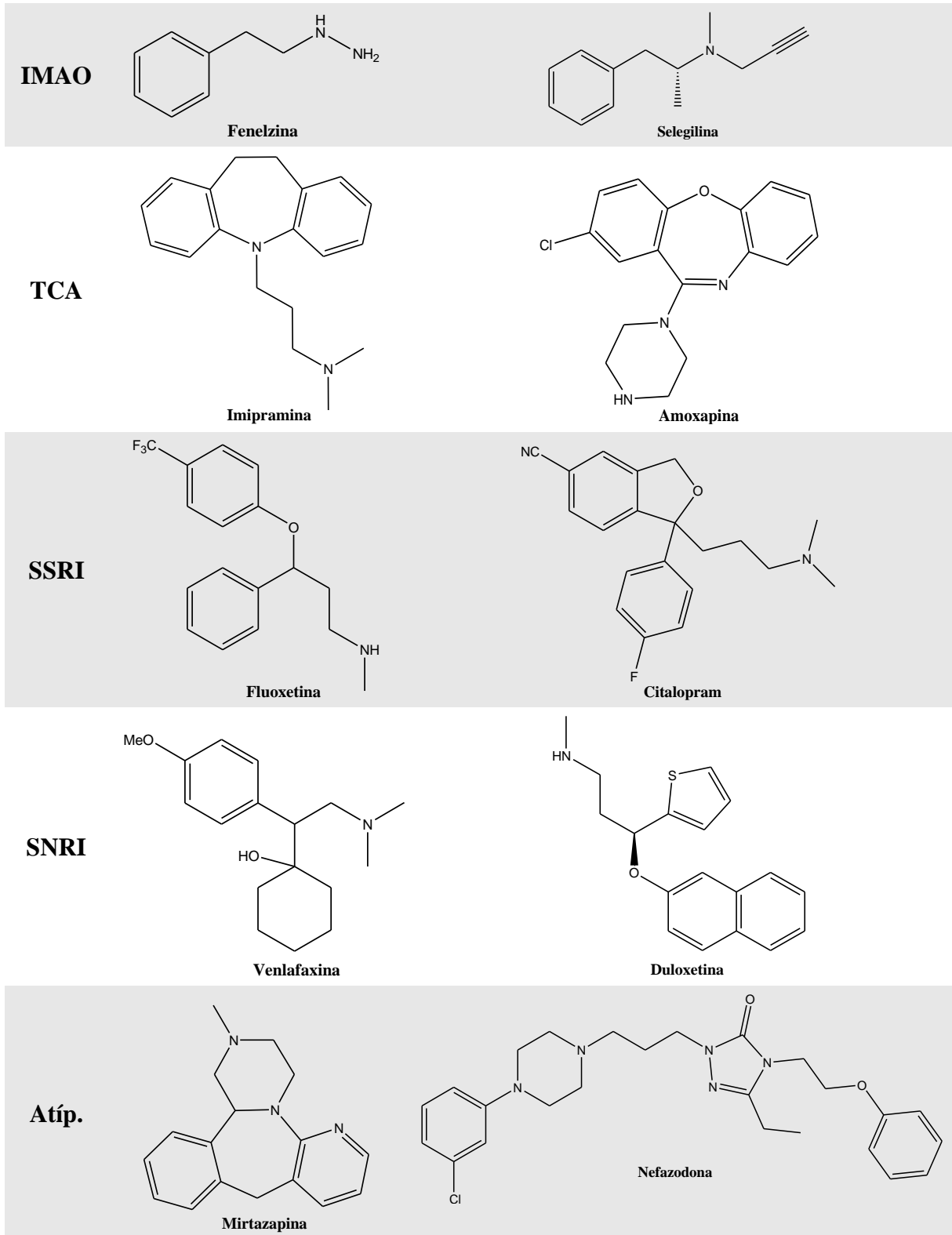
(colinérgicos), histamínicos, adrenérgicos, dopaminérgicos e canais de sódio (MORENO; MORENO; SOARES, 1999; NADAL-VICENS; CHYUNG; TURNER, 2009).

Os SSRIs são resultado da busca por medicamentos tão eficientes quanto os TCAs, mas com perfil atenuado de efeitos colaterais. Esses compostos ficaram conhecidos como antidepressivos de segunda geração e inauguraram uma nova era no tratamento da depressão. Essa classe de antidepressivos é caracterizada pela inibição potente e seletiva da recaptação de 5-TH, provocando uma melhora na neurotransmissão serotoninérgica e, conseqüentemente, dos sintomas depressivos. Embora a eficiência dos SSRIs seja comparável à dos TCAs, eles são mais seletivos e, portanto, apresentam um número menor de efeitos adversos. Por essa razão, são considerados agentes de primeira linha no tratamento da depressão (ARTIGAS, 2013; NADAL-VICENS; CHYUNG; TURNER, 2009).

Outra classe de antidepressivos são os SNRIs, os quais inibem a recaptação de 5-TH e NE por meio de um mecanismo dependente da concentração. Em baixas concentrações, esses compostos agem como SSRIs, enquanto que, em altas concentrações, eles também inibem a recaptação de NE. Por último, os antidepressivos atípicos representam os fármacos que não se enquadram nas outras classes. Esses antidepressivos interagem com múltiplos alvos e atuam por meio de mecanismos diferentes, em que muitos deles ainda não são completamente conhecidos. De forma geral, esses fármacos apresentam baixo perfil de efeitos colaterais e eficácia clínica semelhante (MORENO; MORENO; SOARES, 1999; NADAL-VICENS; CHYUNG; TURNER, 2009).

Apesar do avanço no que se refere ao tratamento da depressão e da quantidade de antidepressivos disponíveis no mercado, ainda existem algumas limitações. Um dos pontos em questão é o atraso entre o início do tratamento e uma resposta positiva, pois a maioria dos fármacos causa uma melhora dos sintomas apenas após semanas de tratamento. Além disso, os medicamentos disponíveis atualmente têm taxa de resposta limitada, sendo que parte dos pacientes não respondem ao tratamento. E, apesar dos avanços feitos em relação aos TCAs, os SSRIs não são totalmente desprovidos de efeitos colaterais. Destacando assim a necessidade de estudos adicionais objetivando o desenvolvimento de novos compostos para o tratamento da depressão (PYTKA et al., 2016; SHARMA; SANTRA; DUTTA, 2015).

Figura 2 - Exemplos de compostos de cada classe de antidepressivos.



Fonte: Do autor (2017)

### 2.2.1 Antagonismo dual do transportador de serotonina e receptor de histamina H<sub>3</sub>

Como mencionado acima, existe um atraso entre o início da administração de SSRIs e uma resposta positiva ao tratamento pelos pacientes. Portanto, como forma de amenizar alguns sintomas no período inicial, tem-se prescrito o uso conjunto de antidepressivos e estimulantes. Os estimulantes auxiliam no alívio dos sintomas até que a melhora clínica pela ação dos antidepressivos possa ser observada. Nesse sentido, o efeito combinado de bloqueio do receptor de histamina H<sub>3</sub> e do transportador de serotonina SERT pode ser uma estratégia interessante para o tratamento da depressão (STOCKING et al., 2010).

A histamina (2-[4-imidazolil]etilamina) é uma amina biogênica sintetizada a partir do aminoácido histidina, sob a ação da enzima L-histidina descarboxilase. A histamina atua como neurotransmissor no SNC e também está relacionada à resposta alérgica, processos inflamatórios e secreção do ácido gástrico. Os efeitos da histamina são mediados por sua ligação aos seus receptores. Atualmente, foram identificados quatro tipos de receptores (H<sub>1</sub>R, H<sub>2</sub>R, H<sub>3</sub>R e H<sub>4</sub>R), os quais pertencem à família GPCRs (NIKOLIC et al., 2014).

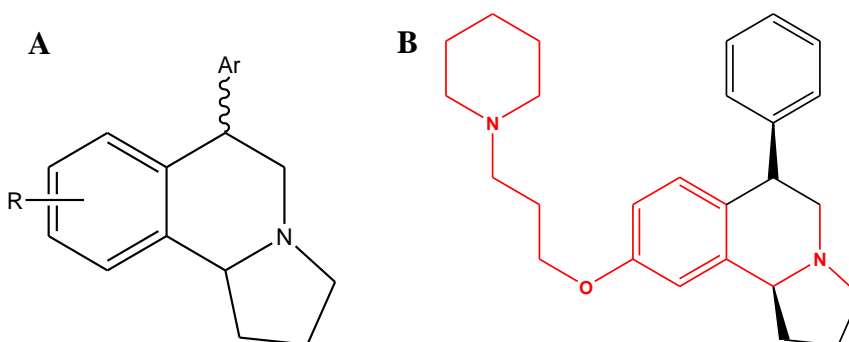
O receptor H<sub>3</sub> (H<sub>3</sub>R) possui papel fundamental na liberação de histamina e de outros neurotransmissores. Em neurônios histaminérgicos, o H<sub>3</sub>R atua como autoreceptor pré-sináptico que regula negativamente os níveis de histamina. Esse receptor também pode ser expresso em outros neurônios, onde atua como um heteroreceptor que modula a liberação de serotonina, noradrenalina, dopamina, acetilcolina, entre outros neurotransmissores (ESBENSHADE et al., 2008; TOKITA; TAKAHASHI; KOTANI, 2006).

Além disso, antagonistas do H<sub>3</sub>R têm sido avaliados para o tratamento de diferentes distúrbios cognitivos, uma vez que esse sistema está envolvido em processos de aprendizado, memória, atenção e vigilância (ESBENSHADE et al., 2008; NIKOLIC et al., 2014). Portanto, compostos com propriedade inibitória para o H<sub>3</sub>R e o SERT podem ser candidatos promissores para o tratamento da depressão, visto que auxiliariam no alívio de alguns sintomas, como fadiga e distúrbios do sono, bem como na função serotoninérgica. Como comentado anteriormente, a inibição do H<sub>3</sub>R tem sido associada ao aumento na liberação de serotonina e, com o concomitante bloqueio do SERT, a reabsorção neuronal desse neurotransmissor será reduzida. Como resultado, maiores concentrações de serotonina serão alcançadas, conduzindo a um efeito terapêutico melhorado (APODACA et al., 2012; STOCKING et al., 2010).

Nesse contexto, Apodaca et al. (2012) desenvolveram uma série de compostos derivados de hexahidro-pirrolo-isoquinolina moduladores de H<sub>3</sub>R e SERT (Fig. 3A). Em

primeiro lugar, os autores selecionaram uma estrutura padrão que possuísse afinidade para o SERT (hexahidro-pirrolo-isoquinolina) e, em seguida, introduziram farmacóforos conhecidos com atividade antagonista H<sub>3</sub>. O componente comum a esses farmacóforos é uma amina terciária ligada ao anel aromático contendo um segundo grupo amina (Fig. 3B) (KEITH et al., 2007a, 2007b; STOCKING et al., 2010).

Figura 3 - A) Estrutura geral dos compostos derivados de hexahidro-pirrolo-isoquinolina desenvolvidos por Apodaca et al. (2012). B) Exemplo de inibidor dual SERT/H<sub>3</sub> com farmacóforo H<sub>3</sub> em vermelho.



Fonte: Adaptado de Stocking (2010)

### 2.3 Química computacional no desenvolvimento de fármacos

A descoberta e desenvolvimento de novos fármacos é um processo complexo, laborioso e dispendioso. Estima-se que o custo médio para se trazer um novo fármaco de sucesso ao mercado ultrapassou US \$2,5 bilhões e requer um período superior a uma década (DIMASI; GRABOWSKI; HANSEN, 2014; DIMASI; HANSEN; GRABOWSKI, 2003; LOMBARDINO; LOWE III, 2004; MORGAN et al., 2011; MUNOS, 2009).

Nesse contexto, os métodos computacionais são uma alternativa para simplificar o processo de planejamento de compostos bioativos de forma complementar às técnicas experimentais. Esses métodos têm sua principal aplicação nas primeiras etapas, onde auxiliam na identificação e otimização dos compostos, selecionando os candidatos à fármacos mais promissores. Dessa forma, minimizando o tempo e os investimentos necessários (KAPETANOVIC, 2008).

Um exemplo de medicamento disponível no mercado que foi desenvolvido com o auxílio da química computacional é o E2020. Esse medicamento, comercialmente conhecido como Aricept<sup>®</sup>, foi aprovado pela Food and Drug Administration (FDA) nos Estado Unidos em 1996 como um potente inibidor da acetilcolinesterase utilizado no tratamento do mal de



Alzheimer. Esse medicamento foi desenvolvido, sintetizado e avaliado com base nos trabalhos de Cardozo et al. (1992), onde foram realizados estudos QSAR de uma série de indanone-benzilpiperidina (KRYGER; SILMAN; SUSSMAN, 1999).

## 2.4 Métodos quantitativos de relação estrutura-atividade

Métodos quantitativos de relação estrutura-atividade, ou QSAR (do inglês, *Quantitative Structure-Activity Relationship*), como o próprio nome sugere, são abordagens computacionais que buscam encontrar relações matemáticas entre uma propriedade de interesse (nesse caso a atividade biológica) e a estrutura química (representada por descritores moleculares) de uma série de compostos congêneres. Estas relações são obtidas por meio de métodos estatísticos e de análise de dados e, uma vez desenvolvido o modelo matemático, ele pode ser usado para prever propriedades de outros compostos análogos (TROPISHA, 2010; WALKER et al., 2003).

Originalmente, os métodos de QSAR foram baseados no princípio de que compostos com estruturas semelhantes induzem efeitos biológicos semelhantes. Esse conceito é conhecido há muito tempo em química orgânica, onde pequenas variações na estrutura química dentro de uma classe acarretam em pequenas mudanças nas propriedades físico-químicas (CRONIN, 2010; FERREIRA, 2002; TROPISHA, 2010).

Os modelos de QSAR são construídos utilizando-se a atividade biológica como variável resposta ( $y$ ) (propriedade a ser predita) e os descritores moleculares como as variáveis independentes ( $X$ ). Esses descritores são definidos por Todeschini e Consonni (2010, p. 30) como “o resultado final de um procedimento lógico e matemático que transforma a informação química codificado dentro de uma representação simbólica de uma molécula em um número útil”.

A representação matemática geral para o QSAR é dada pela equação 1:

$$\text{Atividade biológica} = f(x_1, x_2, \dots, x_N) \quad (1)$$

De forma que é preciso definir um conjunto de descritores ( $x_i$ ) adequados para descrever as características físicas e/ou químicas dos compostos e uma função matemática ( $f$ ) que consiga explicar os dados biológicos com precisão (SO, 2000).

O QSAR tem se aperfeiçoado continuamente desde sua sistematização na década de 60, por meio dos trabalhos de Hansch, Fujita, Free e Wilson, e se tornou uma ferramenta

amplamente utilizada no desenvolvimento de novos fármacos. Esses métodos auxiliam, dentre outras coisas, na identificação de candidatos nocivos nas primeiras etapas de desenvolvimento, na otimização de compostos existentes e na predição de propriedades de compostos que ainda não foram testados. Em virtude disso, essa metodologia desperta grande interesse, uma vez que contribui para redução do tempo e dos custos de desenvolvimento do produto e da necessidade de testes em animais (CRONIN, 2010; VERMA; KHEDKAR; COUTINHO, 2010).

Atualmente, os métodos QSAR podem ser classificados em QSAR-1D a QSAR-7D, segundo a representação estrutural do ligante e os descritores moleculares utilizados (LILL, 2007; VERMA; KHEDKAR; COUTINHO, 2010).

#### 2.4.1 Desenvolvimento dos métodos QSAR

Aparentemente, a origem das técnicas de QSAR se deu com a tese de Cros, da Universidade de Estrasburgo, em 1863. Cros observou que a toxicidade de álcoois em mamíferos aumentava à medida que sua solubilidade em água diminuía. Desde então, diversos trabalhos relacionando diferentes propriedades foram sendo desenvolvidos. Como, por exemplo, em 1868, Brown e Fraser, por meio de estudos com alcaloides, concluíram que a atividade fisiológica devia ser uma função da constituição química. Por volta de 1900, e de maneira independente, Meyer e Overton propuseram que existia uma relação entre a ação narcótica de alguns compostos orgânicos e sua solubilidade em água e em lipídeos (KUBINYI, 2002).

Em meados da década de 30, Hammett fez uma grande contribuição à metodologia QSAR. Em seu trabalho sobre a ionização de ácidos benzoicos substituídos, Hammett desenvolveu o primeiro procedimento de caráter geral, onde demonstrou uma relação linear entre o logaritmo da constante de ionização do ácido benzoico substituído e do ácido benzoico não substituído (equação 2):

$$\log K_X - \log K_0 = \rho\sigma \quad (2)$$

em que  $K_X$  é a constante de ionização do ácido benzoico *para* ou *meta* substituído,  $K_0$  a constante de ionização do ácido benzoico,  $\rho$  é a constante de reação e  $\sigma$  é a constante de grupo, mais conhecida como constante eletrônica de Hammett, a qual mede a influência eletrônica de um dado substituinte (TAVARES, 2004).

Hansch, a partir de estudos da lipofilicidade de ácido fenoxiacéticos substituídos, desenvolveu o parâmetro  $\pi$ , de forma similar à constante de Hammett (equação 3):

$$\log P_X - \log P_H = \pi_X \quad (3)$$

onde  $P_X$  é o coeficiente de partição em n-octanol/água do composto substituído,  $P_H$  o coeficiente de partição em n-octanol/água do composto não substituído e  $\pi_X$  é um parâmetro que reflete a contribuição hidrofóbica do grupo substituinte (KUBINYI, 2002).

Nesse ínterim, os descritores mais utilizados nas primeiras análises de QSAR foram o coeficiente de partição octanol/água ( $\log P$ ), a constante eletrônica de Hammett ( $\sigma$ ) e o parâmetro de Hansch ( $\pi$ ). Posteriormente, essas primeiras análises de QSAR foram classificadas como QSAR-1D, visto que consideram apenas uma propriedade físico-química para explicar o efeito biológico (LILL, 2007).

A sistematização da metodologia de QSAR teve início com os trabalhos de Free e Wilson (1964) e Hansch e Fujita (1964).

Em seu trabalho, Hansch e Fujita, ao invés de uma, combinaram várias propriedades em um modelo linear, conforme equação 4, metodologia que ficou conhecida com análise de Hansch:

$$\log 1/C = a.\pi + b.\sigma + \dots + \text{const.} \quad (4)$$

em que C corresponde à concentração molar que causa o efeito biológico desejado,  $\log 1/C$  a potência biológica, a e b os coeficiente de regressão (TAVARES, 2004).

Ao mesmo tempo e de maneira independente à Hansch e Fujita, Free e Wilson, em seu estudo de uma série de compostos análogos, postularam a equação 5, onde a contribuição de substituintes para a atividade biológica ocorre de maneira aditiva, dependendo apenas da posição e do tipo de grupo:

$$\log 1/C = \sum a_i + \mu \quad (5)$$

onde  $a_i$  são as contribuições individuais dos substituintes para a atividade biológica e  $\mu$  a atividade do composto de referência, que normalmente é o composto não substituído. A equação 5, apesar de mais simples, é mais restrita do que a proposta por Hansch e Fujita, uma

vez que pode ser aplicada apenas à moléculas cujos substituintes são combinações lineares daqueles do conjunto de treinamento (KUBINYI, 2002).

Com os trabalhos de Hansch, Fujita, Free e Wilson, a resposta biológica passou a ser relacionada a padrões estruturais, como índices de conectividade e grupos funcionais, todavia, não levando em conta a representação tridimensional do ligante. Posteriormente, essa metodologia foi classificada com QSAR-2D (VERMA; KHEDKAR; COUTINHO, 2010).

Na década de 1980, quando estruturas de difração de raios X de proteínas se tornaram disponíveis, foi possível a interpretação de parâmetros tridimensionais nas análises QSAR. O método de QSAR-3D mais conhecido é a Análise Comparativa de Campo Molecular (CoMFA, do inglês *Comparative Molecular Field Analysis*), desenvolvido por Cramer, Patterson e Bunce (1988). Nessa metodologia, as estruturas tridimensionais dos compostos são alinhadas dentro de uma caixa virtual, onde átomos de sonda são utilizados para computar a energia de interação de cada molécula em cada ponto da caixa. A energia de interação e as atividades biológicas são correlacionadas de forma quantitativa pela análise de regressão por mínimos quadrados parciais. Entretanto, essa metodologia é altamente dependente do alinhamento, da identificação da conformação “bioativa” e gera um grande número de descritores (CRAMER; PATTERSON; BUNCE, 1988; KUBINYI, 1997; LILL, 2007; VERMA; KHEDKAR; COUTINHO, 2010).

Em 1997, Hopfinger et al. (1997) propuseram o método de QSAR-4D. Nessa metodologia, a quarta dimensão corresponde à amostragem conformacional em função do tempo de simulação de dinâmica molecular (SDM), que permite analisar a flexibilidade conformacional de cada ligante. No QSAR-4D, um perfil de amostragem conformacional de cada ligante é usado para gerar os descritores, ao invés de apenas uma conformação inicial. Ademais, essa metodologia QSAR-4D pode ser executada de duas maneiras: pela análise QSAR dependente do receptor, (RD) QSAR-4D, quando a estrutura tridimensional do receptor está disponível, ou pela análise QSAR independente do receptor, (RI) QSAR-4D (ANDRADE et al., 2010; HOPFINGER et al., 1997).

A metodologia de QSAR-5D, além da múltipla representação do ligante (quarta dimensão do QSAR), considera a flexibilidade do receptor por meio de diferentes protocolos de encaixe induzido, a quinta dimensão (LILL, 2007; VEDANI; DOBLER, 2002). Ademais, a representação de diferentes modelos de solvatação, juntamente com os fundamentos do QSAR-5D, constitui a metodologia QSAR-6D (VEDANI; DOBLER; LILL, 2005).

Posteriormente, foi definida a metodologia de QSAR-7D, que considera modelos reais ou virtuais (como, por exemplo, modelagem por homologia) da interação ligante-receptor (POLANSKI, 2009).

#### 2.4.2 QSAR-4D

A metodologia do QSAR-4D proposta por Hopfinger et al. (1997) foi desenvolvida para lidar com problemas inerentes ao QSAR-3D, onde a quarta dimensão corresponde à amostragem conformacional. Nesse sentido, as maiores vantagens do QSAR-4D são a incorporação da flexibilidade conformacional do ligante, a avaliação de múltiplos alinhamentos e a interpretação de como as regiões individuais dos compostos contribuem para a atividade, por meio da definição das interações farmacofóricas (ANDRADE et al., 2010; HOPFINGER et al., 1997).

No QSAR-4D, um perfil de amostragem conformacional (CEP, do inglês *Conformational Ensemble Profile*) de cada ligante é usado para gerar os descritores moleculares, ao invés de apenas uma conformação inicial. Esse conjunto conformacional é obtido pela etapa de SDM, que permite obter um conjunto estatisticamente representativo de conformações distintas de baixa energia (HOPFINGER et al., 1997; PITA, 2006). Em síntese, a dinâmica molecular clássica consiste na resolução numérica, passo-a-passo, da equação de movimento de Newton para todos os átomos do sistema, em que o conjunto de coordenadas das partículas ao longo do tempo é denominado trajetória (JENSEN, 2007; PITA, 2006).

Os descritores moleculares na metodologia de QSAR-4D são definidos como o conjunto de descritores de ocupação da célula (GCOD, do inglês *Grid Cell Occupancy Descriptor*). De maneira semelhante ao CoMFA, o CEP para cada composto é alinhado dentro de uma caixa tridimensional virtual, que mimetiza o sítio de ligação da enzima ou receptor. A caixa é composta por células cúbicas que podem ter dimensões variadas definidas pelo usuário. Quanto menor o tamanho da célula, mais refinados serão os modelos, entretanto, o tempo de análise aumenta devido ao maior número de variáveis geradas. Os átomos de cada composto são classificados em elementos de interação farmacofórica (IPE, do inglês *Interaction Pharmacophore Element*), que correspondem aos tipos de interação que podem ocorrer no sítio ativo. Desse modo, os GCODs correspondem a frequência de ocupação de cada célula da caixa pelos diferentes tipos de elementos de interação farmacofórica (HOPFINGER et al., 1997; ROMEIRO et al., 2005).

Os GCODs são então processados em conjunto com os valores de atividade biológica utilizando uma metodologia que combina regressão por mínimos quadrados parciais e algoritmos genéticos.

### 2.4.3 Regressão pelo método dos mínimos quadrados parciais

De maneira geral, métodos lineares de regressão multivariada relacionam duas ou mais variáveis independentes ( $x$ ) à variável dependente ( $y$ ) como demonstrado na equação 6. A equação de regressão também pode ser escrita na notação matricial (eq.7), onde as linhas representam as amostras do conjunto de dados e as colunas correspondem às variáveis do modelo.

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n + e \quad (6)$$

$$Y = XB + E \quad (7)$$

em que  $b_i$  são os coeficiente de regressão,  $n$  o número de variáveis independentes e  $e$  o valor residual (FERREIRA, 2015; GELADI; KOWALSKI, 1986).

Nesse contexto, a regressão linear por mínimos quadrados parciais (PLS, do inglês *Partial Least Square*) não relaciona a matriz original  $X$  diretamente à variável resposta. O espaço das variáveis originais é aproximado por uma dimensão reduzida formada pelas variáveis latentes (VL). Em outras palavras, as VL são combinações lineares das variáveis originais construídas para explicar a variância em  $X$  ao mesmo tempo que a previsão da variável  $y$ , e são ortogonais entre si. Assim, o PLS reduz o número de variáveis a algumas poucas não correlacionadas. Essas VL são então utilizadas para modelar  $y$ , na forma descrita pelas equações acima, em que o número ótimo de componentes no modelo pode ser selecionado pela validação cruzada (FERREIRA, 2015; GELADI; KOWALSKI, 1986; PIRHADI; SHIRI; GHASEMI, 2015).

Com isso, o PLS permite relacionar um grande número de variáveis com uma ou mais variáveis resposta, sendo assim bastante empregado em estudos de QSAR. Além disso, a metodologia é especialmente interessante para modular dados com variáveis altamente correlacionadas (colineares), para os quais métodos mais simples de regressão, como a regressão linear múltipla, são mais sensíveis (FERREIRA; MONTANARI; GAUDIO, 2002; PIRHADI; SHIRI; GHASEMI, 2015).

#### 2.4.4 Aproximação da função genética

A aproximação da função genética (GFA, do inglês *Genetic Function Approximation*), algoritmo desenvolvido por Rogers e Hopfinger (1994), é um método baseado em algoritmo genético (GA, do inglês *Genetic Algorithm*), isto é, um algoritmo de otimização inspirado em princípios evolucionários.

Uma população inicial de “indivíduos” (modelos) é construída pela seleção aleatória de “genes” (descritores). Uma função de ajuste é aplicada para estimar a qualidade desses modelos, em que os melhores modelos têm maior probabilidade de propagar suas informações para as próximas “gerações” (SO, 2000). No caso da metodologia GFA, os modelos são classificados segundo os valores de LOF (do inglês, *lack-of-fit*), fator de desajuste de Friedman (ROGERS; HOPFINGER, 1994):

$$\text{LOF} = \frac{\text{LSE}}{\left(1 - \frac{c+dp}{M}\right)^2} \quad (8)$$

onde LSE é o erro de mínimos quadrados (do inglês, *least-squares error*),  $c$  é o número de funções de base,  $d$  é o *smoothing fator* (termo definido pelo usuário que controla o número de variáveis no modelo),  $p$  é o número de variáveis contidas nas funções de base e  $M$  o número de compostos no conjunto de treinamento. Portanto, o LOF é uma medida dos erros dos mínimos quadrados penalizados, isto é, quando dois modelos tem o mesmo valor de LSE, aquele com menor número de variáveis tem menor (melhor) valor de LOF (ROMEIRO et al., 2005).

Em seguida, os modelos sofrem processos de cruzamento e mutação, combinando os melhores termos e gerando modelos cada vez mais apurados. O processo se encerra quando o valor médio de LOF deixa de melhorar significativamente. Com isso, a grande vantagem da GFA é a construção e otimização de múltiplos ao invés de um único modelo (ROGERS; HOPFINGER, 1994).

#### 2.4.5 Validação dos modelos

Os parâmetros de ajuste do modelo, como o coeficiente de determinação ( $R^2$ ), fornecem uma medida da habilidade do modelo em reproduzir os dados do treinamento, porém sozinhos não são suficientes para avaliar sua qualidade. Para isso são necessários

alguns métodos de validação adicionais (FERREIRA; MONTANARI; GAUDIO, 2002; GRAMATICA, 2007).

Uma das validações mais comumente utilizadas é a validação cruzada. A validação cruzada avalia a robustez e preditividade interna do modelo, testando a significância estatística de cada componente. Em sua forma mais simples, um composto por vez do conjunto de calibração é omitido. O modelo é então recalculado e a atividade do composto excluído é predita. O processo é repetido até que todos os compostos tenham ficado de fora uma vez. Essa modalidade de validação cruzada é denominada *leave-one-out* (LOO). O procedimento também pode ser feito excluindo um conjunto de compostos de cada vez, ao invés de apenas um, nesse caso conhecido como *leave-many-out* (LMO) ou *leave-N-out* (LNO)(GRAMATICA, 2007; KIRALJ; FERREIRA, 2009).

Como um dos objetivos do QSAR é aplicar os modelos para estimar a atividade de outras moléculas, é necessário avaliar sua capacidade preditiva. Para isso, alguns autores argumentam que a melhor forma é utilizar um conjunto externo, que não foi incluído na construção do QSAR e para o qual os valores dos descritores e atividade biológica são conhecidos, e então calcular a atividade predita pelo modelo para esse grupo (CHERKASOV et al., 2014; TROPSHA, 2010). Geralmente, o conjunto de dados completo é dividido em dois subgrupos: o conjunto de treinamento (utilizado para construir os modelos de QSAR) e o conjunto de teste (utilizado para validar esses modelos). Recomenda-se que cerca de 20-30% dos dados componham o conjunto de teste. Além disso, esses compostos devem ser estruturalmente semelhantes aos do treinamento. Existem diferentes métodos descritos na literatura para a seleção desses conjuntos (GRAMATICA, 2007; KIRALJ; FERREIRA, 2009).

Outra validação bastante comum é a randomização, mais especificamente a Y-randomização. O processo consiste em desorganizar diversas vezes o bloco Y (atividades biológicas), mantendo o bloco de descritores X fixo. Com isso, espera-se que os modelos gerados nessas condições tenham baixa qualidade, ou seja, que o  $R^2$  do processo de randomização seja bem menor do que o  $R^2$  do modelo original. O intuito dessa validação é provar que não existe correlação por chance (*chance correlation*), que a relação entre os descritores e a atividade não é por acaso, em outras palavras, que eles não estão apenas estatisticamente correlacionados (KIRALJ; FERREIRA, 2009; RÜCKER; RÜCKER; MERINGER, 2007).

Por fim, é altamente recomendado na literatura a definição do domínio de aplicabilidade dos modelos. O domínio de aplicabilidade (DA) é um espaço teórico definido



pelos descritores e resposta biológica do treinamento. Apenas compostos que estejam dentro desse domínio serão bem preditos pelo QSAR. Além disso, uma vez que o DA é calculado com base nas informações do treinamento, ele é específico para cada modelo gerado (CHERKASOV et al., 2014; GRAMATICA, 2007).

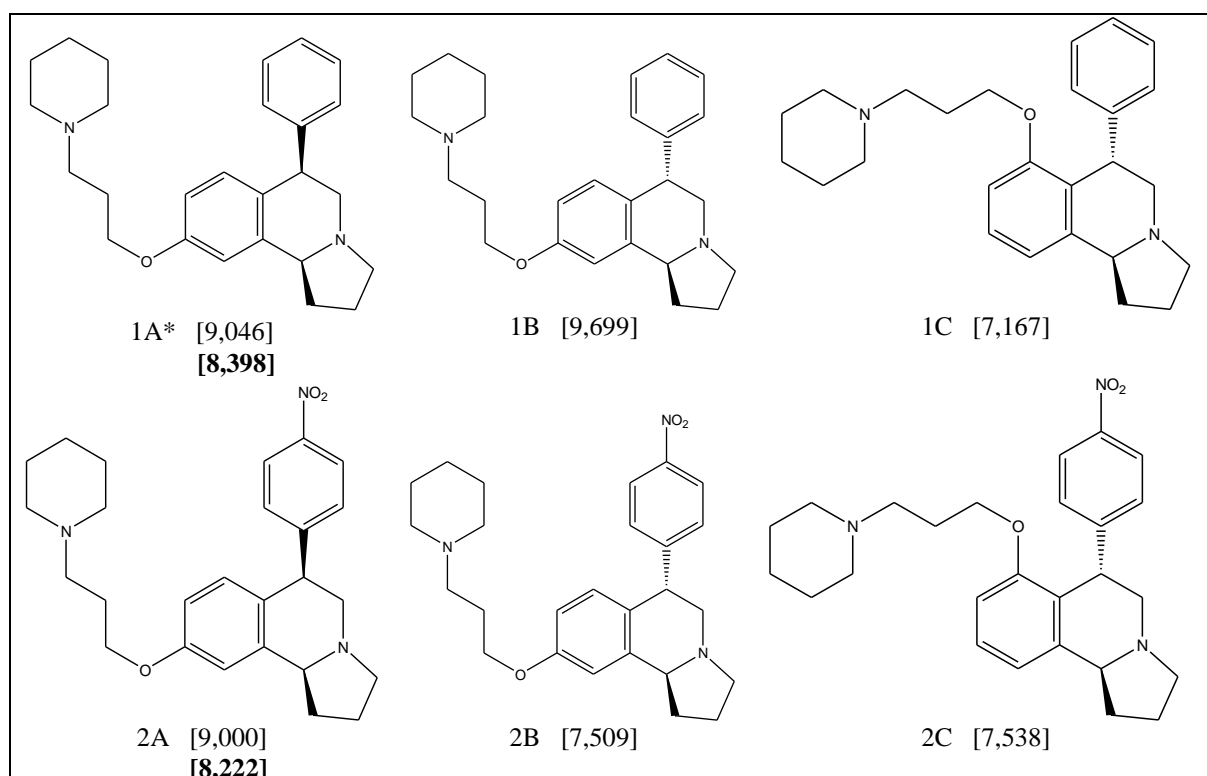
### 3 METODOLOGIA

O trabalho foi desenvolvido no Laboratório de Modelagem Molecular do Departamento de Química da Universidade Federal de Lavras.

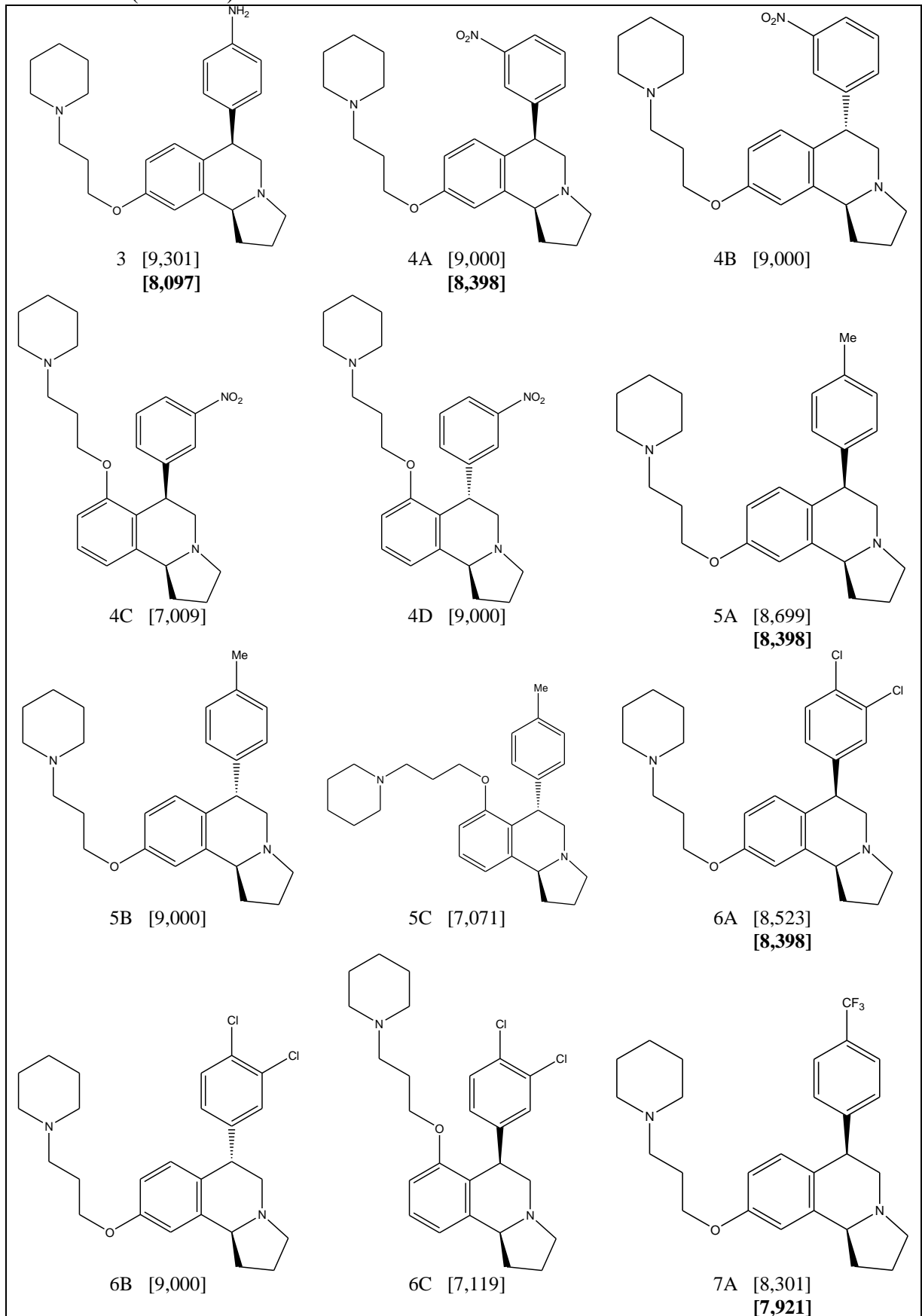
#### 3.1 Dados Biológicos

No Quadro 1 estão representados os compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de atividade para ambos os receptores em termos do logaritmo negativo da constante inibitória ( $-\log K_I = pK_I$ ). O conjunto de dados é composto por 108 estruturas com atividade biológica para o H<sub>3</sub>R e 43 para o SERT. Os dados foram extraídos da literatura (APODACA et al., 2012). É importante destacar que todos os dados foram obtidos pelo mesmo laboratório, eliminando assim potenciais erros caso fossem provenientes de fontes diferentes.

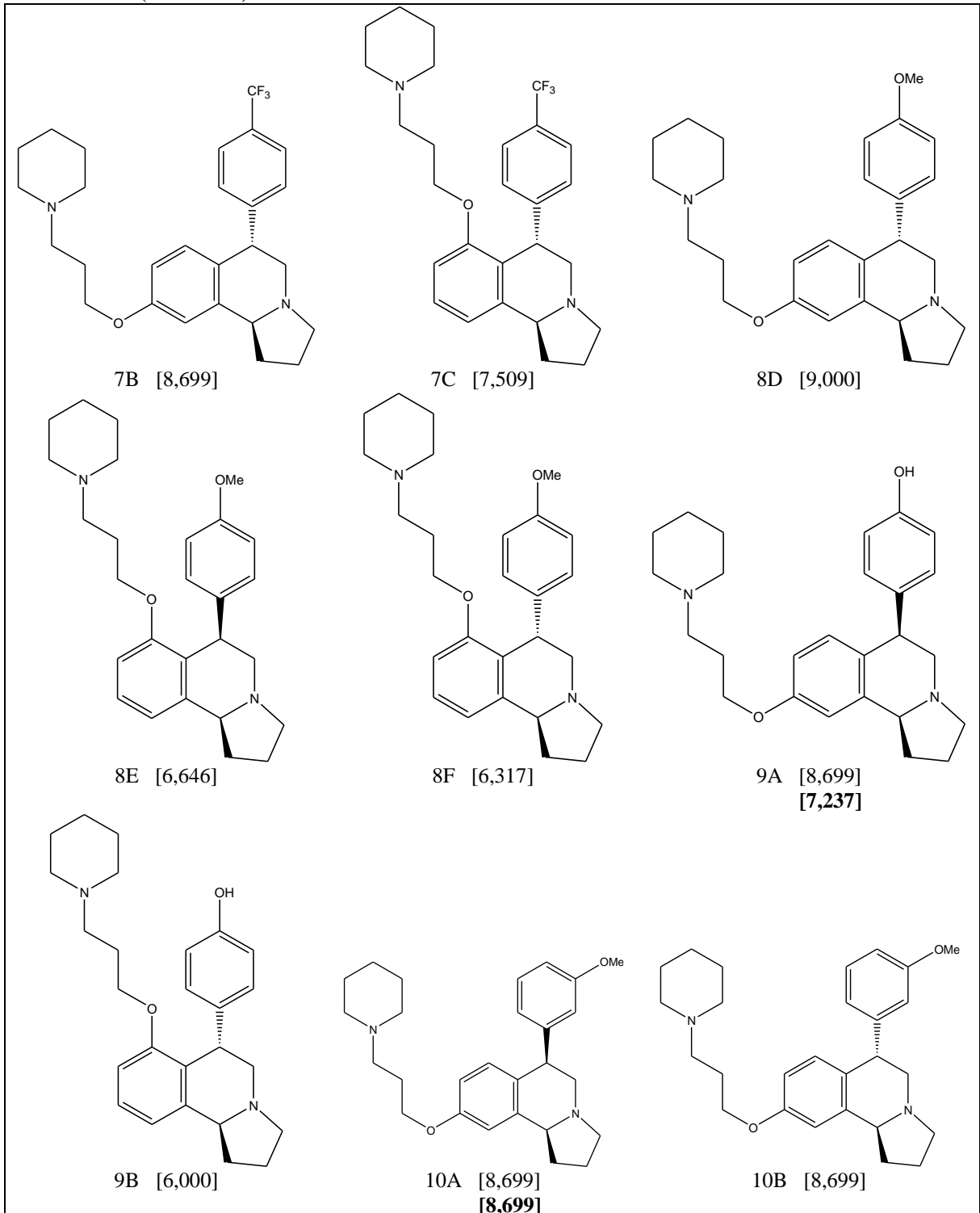
Quadro 1 - Compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória, expressa como  $pK_I$ , para o H<sub>3</sub>R e SERT (negrito).  
(Continua)



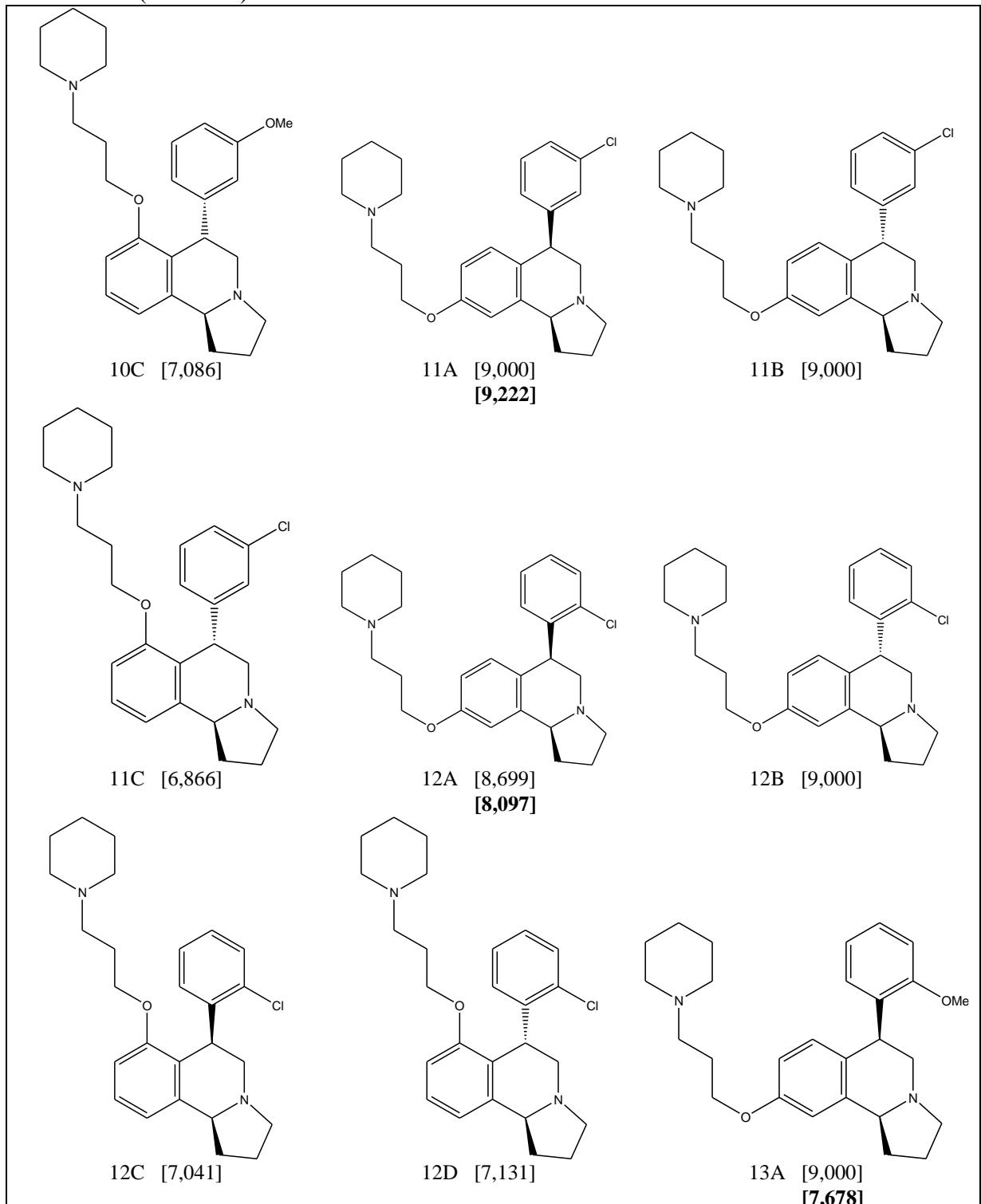
Quadro 1 - Compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória, expressa como  $pK_I$ , para o H<sub>3</sub>R e SERT (negrito). (Continua)



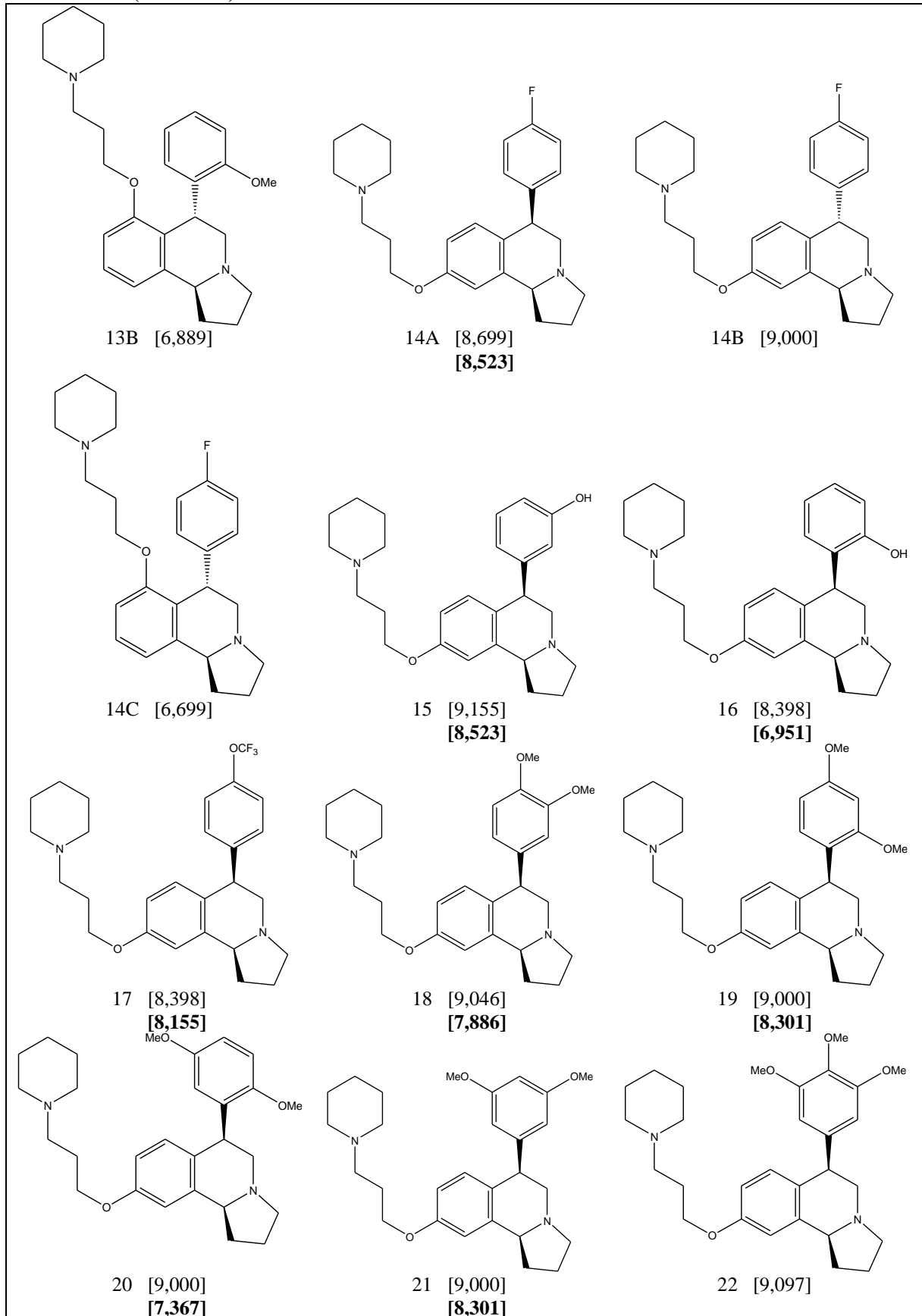
Quadro 1 - Compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória, expressa como  $pK_I$ , para o H<sub>3</sub>R e SERT (negrito). (Continua)



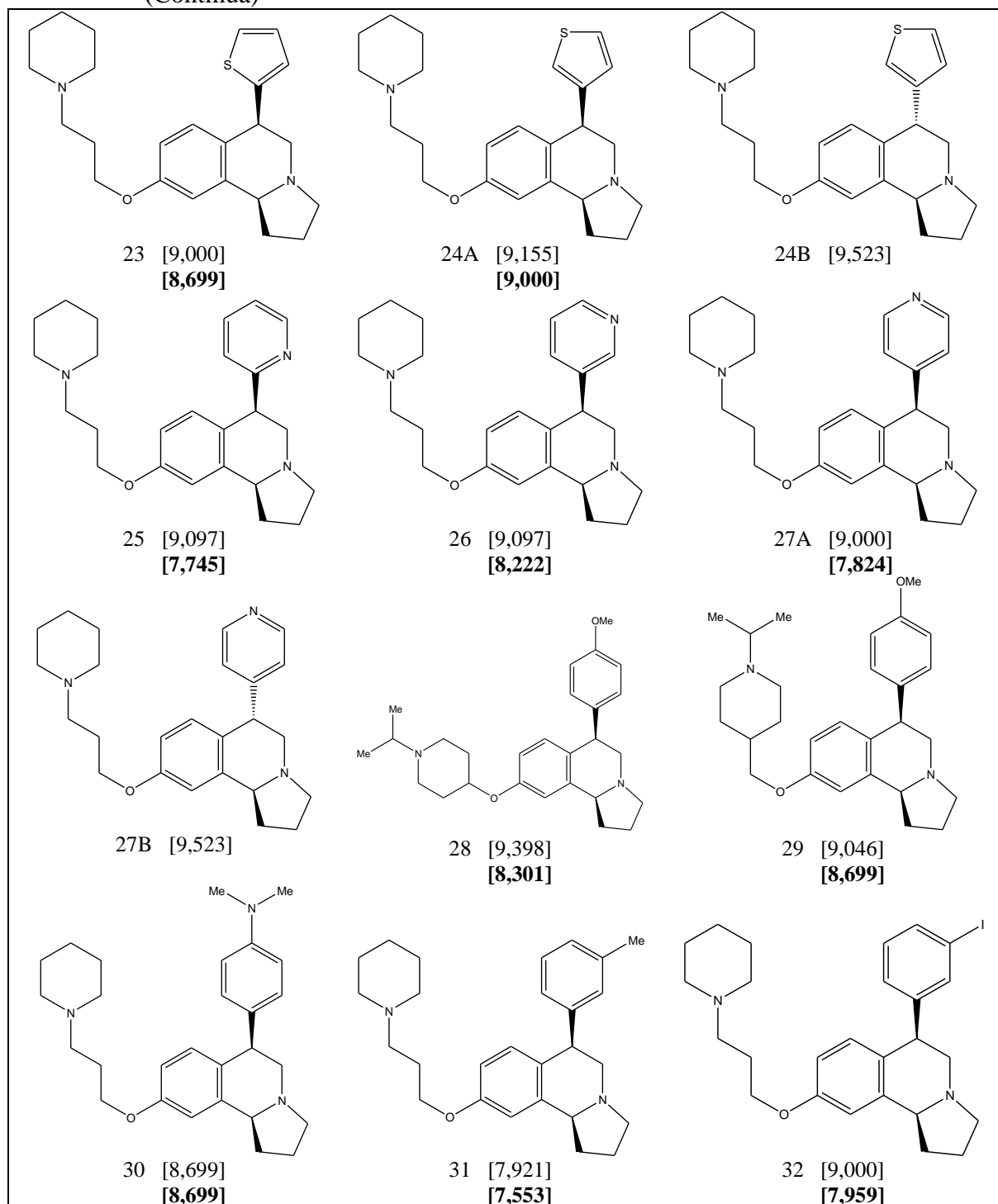
Quadro 1 - Compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória, expressa como  $pK_I$ , para o H<sub>3</sub>R e SERT (negrito). (Continua)



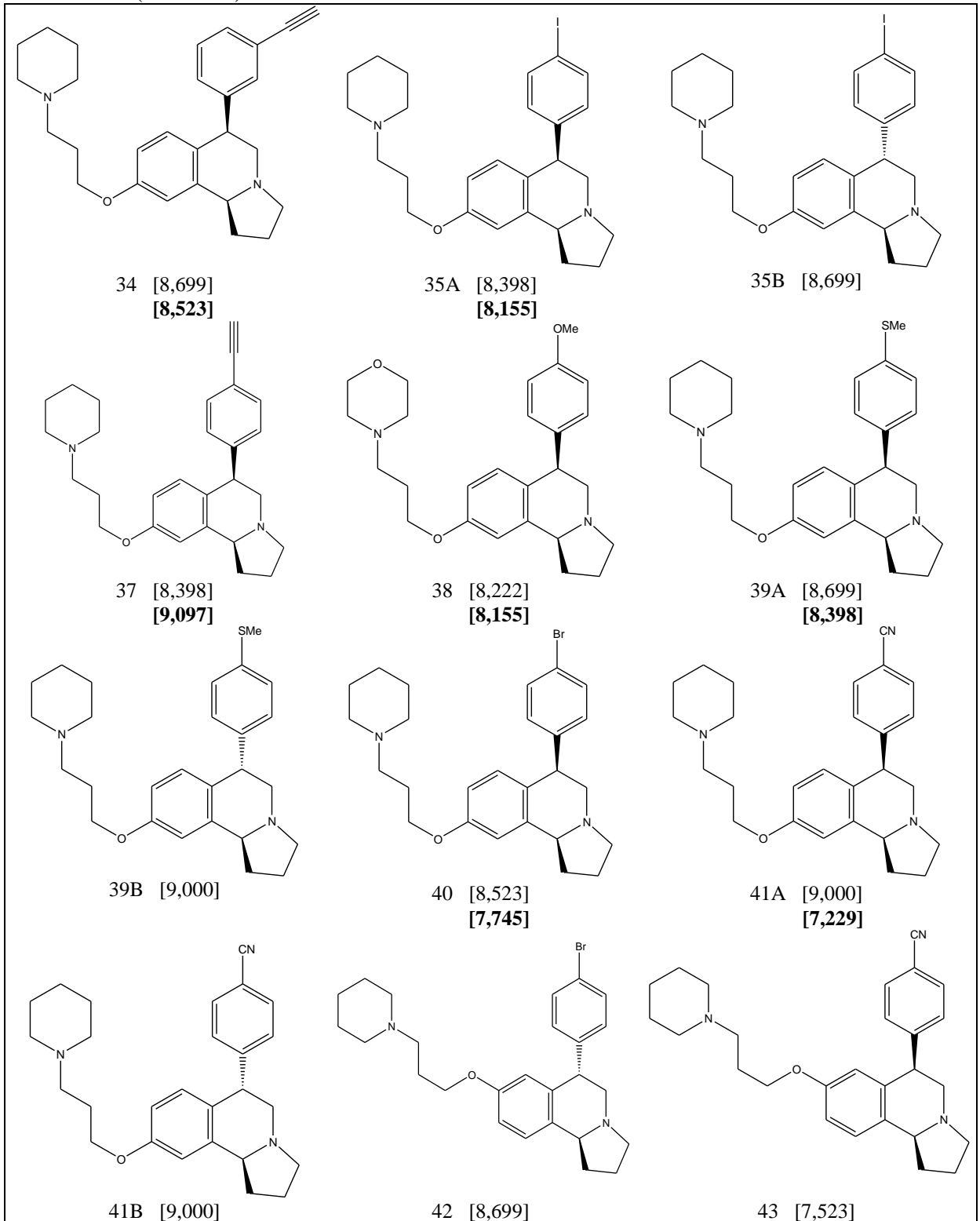
Quadro 1 - Compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória, expressa como  $pK_{i1}$ , para o H<sub>3</sub>R e SERT (negrito). (Continua)



Quadro 1 - Compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória, expressa como  $pK_I$ , para o H<sub>3</sub>R e SERT (negrito). (Continua)

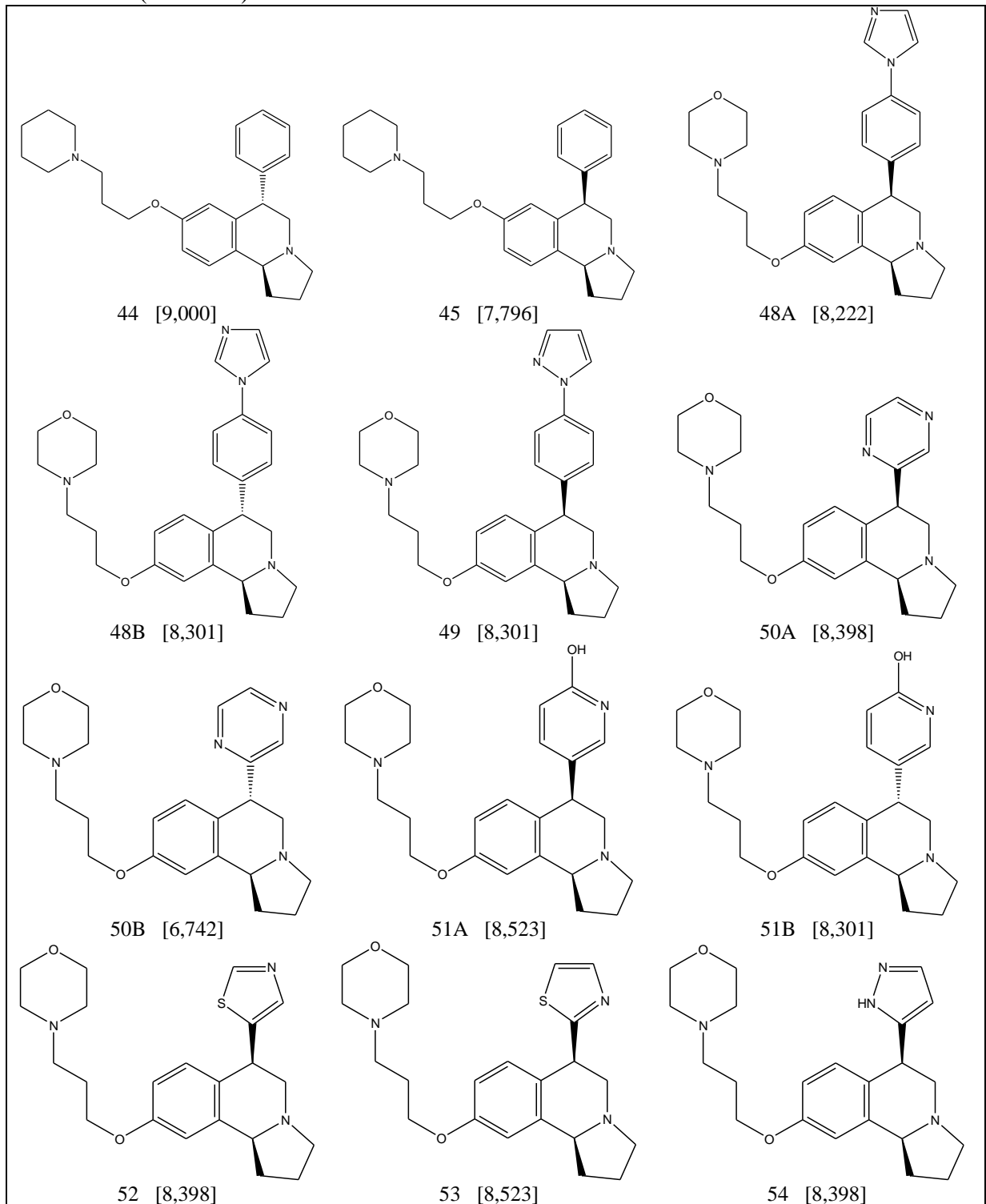


Quadro 1 - Compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória, expressa como  $pK_I$ , para o  $H_3R$  e SERT (negrito). (Continua)

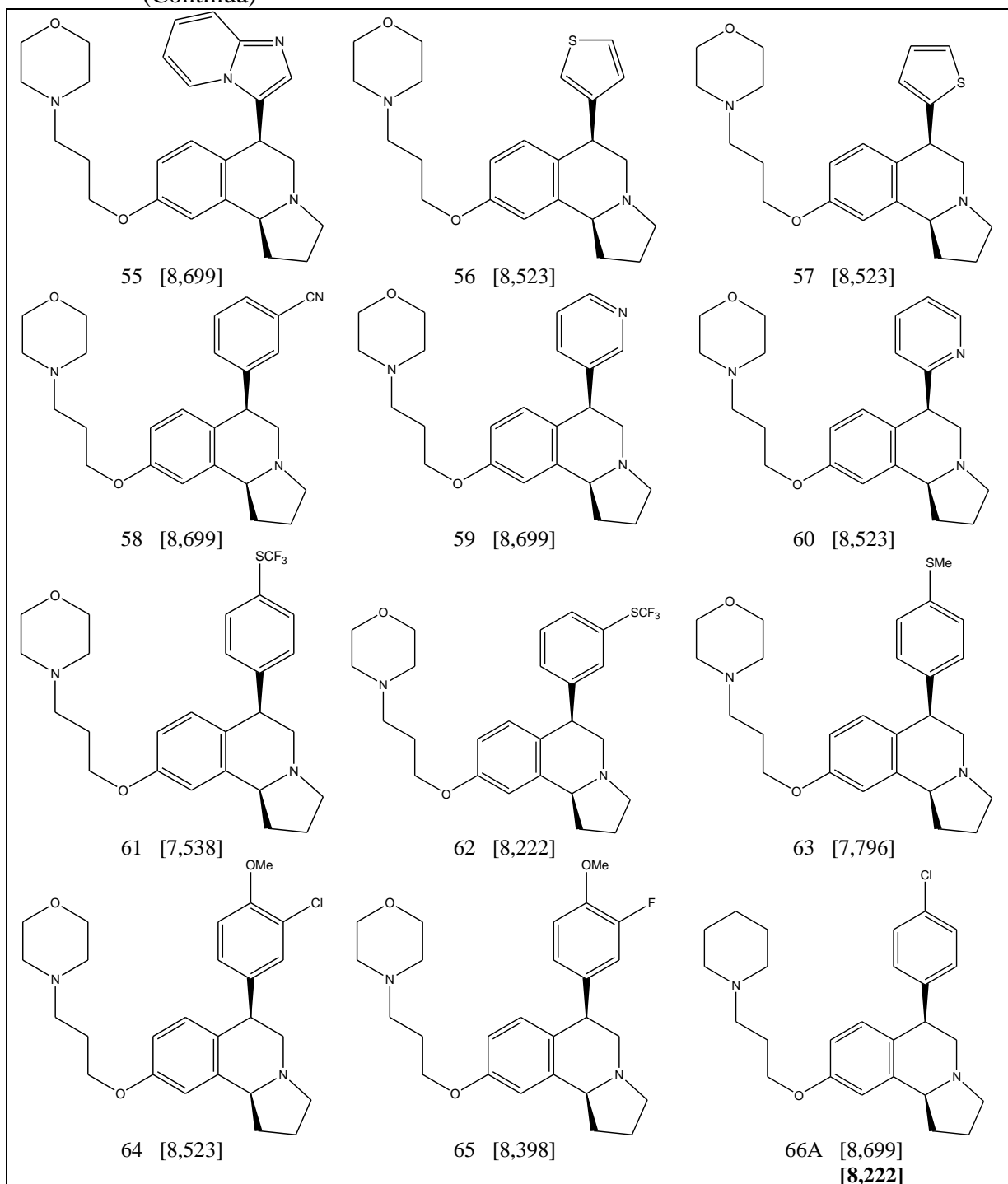




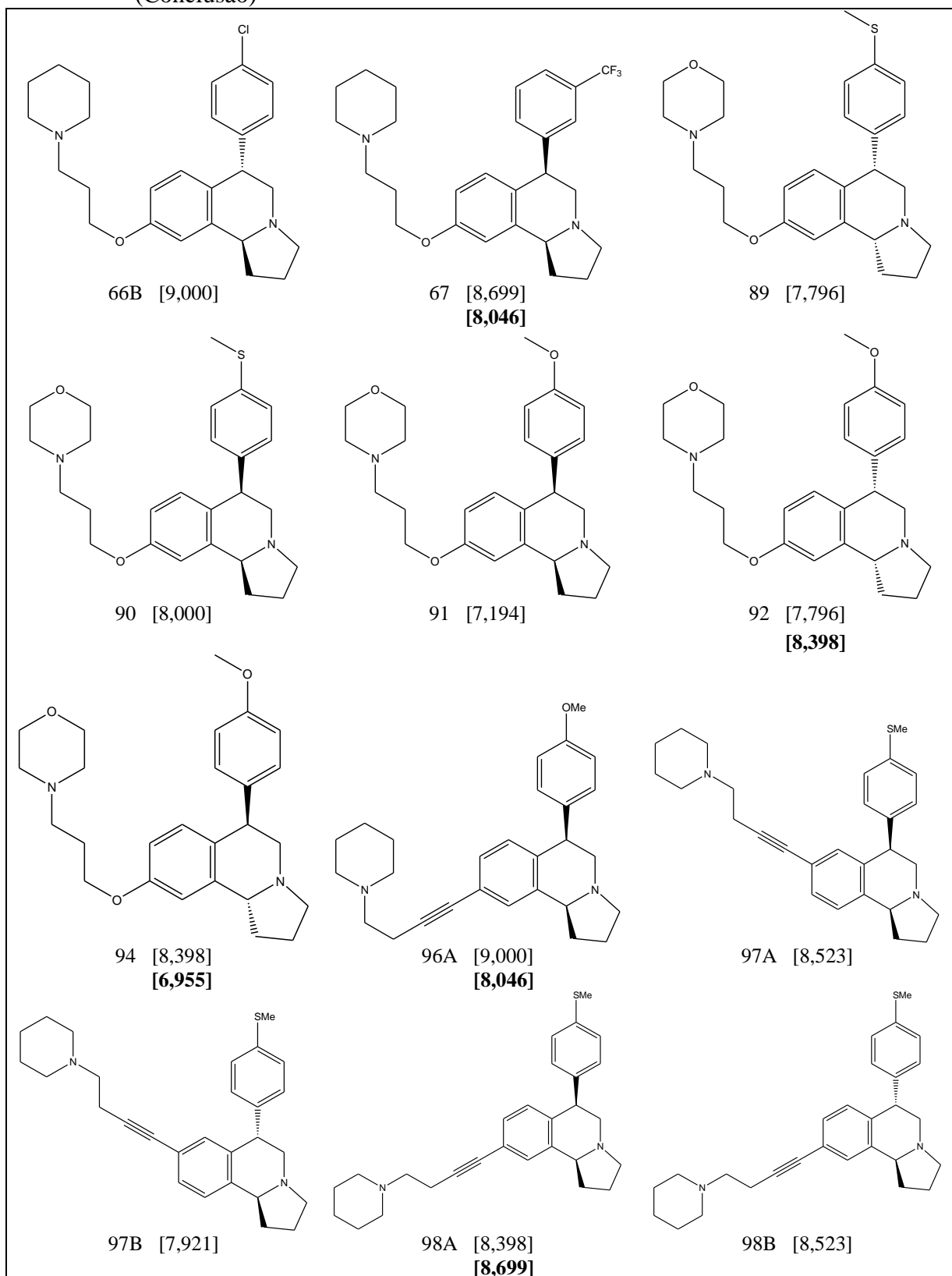
Quadro 1 - Compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória, expressa como  $pK_I$ , para o  $H_3R$  e SERT (negrito). (Continua)



Quadro 1 - Compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória, expressa como  $pK_I$ , para o H<sub>3</sub>R e SERT (negrito). (Continua)



Quadro 1 - Compostos derivados de hexahidro-pirrolo-isoquinolina e seus respectivos valores de constante inibitória, expressa como  $pK_I$ , para o H<sub>3</sub>R e SERT (negrito). (Conclusão)



\*A numeração foi mantida a mesma do trabalho original.

Fonte: Adaptado de Apodaca (2012)

Uma vez que os testes farmacológicos foram realizados em pH fisiológico, a porção piperidina dos compostos foi considerada em sua forma protonada. Além disso, a interação de grupos ionizáveis dos compostos com o receptor parece ser importante tanto para a ligação de antagonistas ao H<sub>3</sub>R (AXE; BEMBENEK; SZALMA, 2006; LEVOIN et al., 2013; SCHLEGEL et al., 2007) como para inibidores do SERT (ANDERSEN et al., 2009; COLEMAN; GREEN; GOUAUX, 2016).

### 3.2 Construção e minimização dos compostos

As estruturas tridimensionais dos compostos foram construídas e otimizadas utilizando o campo de força MM+ (ALLINGER, 1977). Em seguida, as cargas atômicas parciais foram calculadas usando o método semi-empírico RM1 (ROCHA et al., 2005).

### 3.3 Conjuntos de treinamento e teste

Os compostos foram divididos em dois grupos, o conjunto de treinamento ou calibração, os quais são utilizados na construção dos modelos de QSAR-4D; e que são validados externamente pelo conjunto de teste. Foram avaliados três conjuntos de teste para cada receptor, contendo aproximadamente 20% dos compostos. O primeiro conjunto foi selecionado manualmente com base na atividade biológica. Os compostos foram ordenados de acordo com valores crescentes de pK<sub>I</sub> e a seleção foi feita aleatoriamente de forma que abrangesse toda faixa de atividade (DA CUNHA et al., 2005). O segundo conjunto de teste foi definido pela seleção de Kennard-Stone, a qual é baseada na distância euclidiana das amostras (KENNARD; STONE, 1969). Por último, foi feita uma análise de cluster utilizando-se o quadrado da distância euclidiana e a máxima distância (*complete linkage*) entre as amostras.

Os conjuntos de teste selecionados para o receptor H<sub>3</sub> foram: teste 1 - 4C, 5A, 6C, 10B, 11A, 14B, 14C, 17, 22, 23, 24A, 34, 38, 41A, 48B, 50A, 51A, 55, 60, 61, 67 e 92; teste 2 - 2C, 3, 5C, 6A, 7A, 11B, 13A, 14B, 14C, 17, 32, 35A, 38, 39A, 48A, 56, 58, 60, 66A, 66B, 67 e 90; teste 3 - 2B, 10A, 12B, 12C, 13B, 14C, 15, 21, 26, 37, 39A, 42, 45, 49, 50A, 50B, 52, 55, 65, 67, 89 e 90. Por outro lado, os conjuntos de teste para o transportador SERT são: teste 1 - 2A, 4A, 10A, 12A, 13A, 16, 32, 34 e 98A; teste 2 - 6A, 9A, 16, 17, 32, 35A, 37, 38 e 67; teste 3 - 2A, 13A, 14A, 25, 26, 28, 30, 31 e 67.

### 3.4 Simulação de Dinâmica Molecular

O processo de SDM foi executado usando o programa Molsim 3.0 (MOLSIM, 1997) com o campo de força MM2 (WEINER et al., 1986), incorporado ao programa 4D-QSAR. A temperatura para cada SDM foi fixada em 300K, próxima da utilizada nos ensaios farmacológicos dos compostos em estudo. Foi definido um tempo de simulação de 50ps em intervalos de 0,001ps, totalizando assim 500 conformações para cada composto. Os cálculos de SDM foram realizados aplicando uma constante dielétrica dependente da distância de 3,5; com o intuito de tentar modular o efeito do solvente na ausência de um solvente explícito (DA CUNHA et al., 2004, 2005).

### 3.5 Definição dos Elementos de Interação Farmacofórica

Os átomos de cada composto foram classificados em sete tipos de IPEs, os quais correspondem aos tipos de interação que podem ocorrer entre o inibidor e o sítio ativo e estão relacionados com os grupos farmacofóricos: qualquer tipo (any), não polar (np), polar com densidade de carga positiva (p+), polar com densidade de carga negativa (p-), acceptor de ligação hidrogênio (hba), doador de ligação hidrogênio (hbd) e aromático (ar) (HOPFINGER et al., 1997).

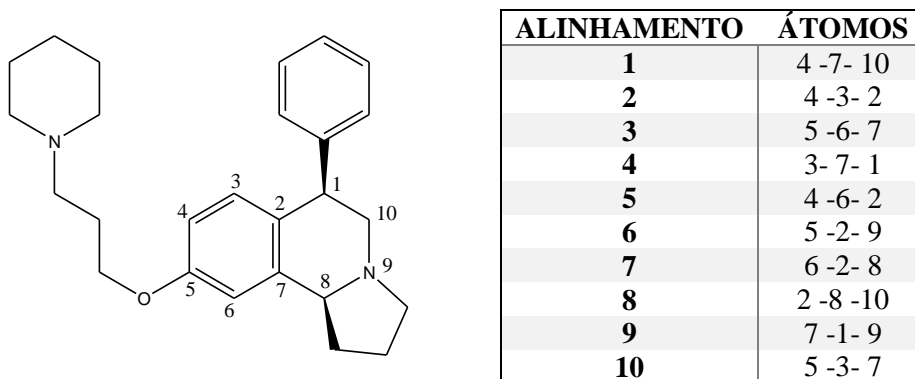
### 3.6 Definição do Alinhamento

A definição do alinhamento é uma das etapas mais importantes na metodologia de QSAR-4D, pois ele é o caminho pelo qual os compostos do conjunto de treinamento são comparados. Em geral, os alinhamentos são escolhidos a partir de átomos ou grupos funcionais de forma a contemplar a maior parte da estrutura comum a todos os compostos (PASQUALOTO et al., 2004; ROMEIRO et al., 2005). Neste trabalho, considerou-se que os compostos se ligam de maneira semelhante ao receptor, visto que são análogos estruturais. Serão analisados os dez melhores alinhamentos, selecionados com base no composto 1A (Fig. 4).

O CEP para cada composto, obtido após a etapa de SDM, foi então inserido dentro de uma caixa tridimensional virtual, conforme os alinhamentos selecionados. A caixa é composta por células cúbicas que registram a distribuição de ocupação espacial de cada átomo de cada ligante. Neste trabalho, a dimensão das células cúbicas foram definidas em 1,0Å de aresta

(ROMEIRO et al., 2005). A ocupação de cada célula de grade por cada tipo de IPE forma o conjunto de descritores de ocupação da célula (GCODs), que correspondem as variáveis independentes na metodologia QSAR-4D (DA CUNHA et al., 2005).

Figura 4 - Átomos que compõem os dez alinhamentos analisados utilizando a molécula 1A como referência.



Fonte: Do autor (2017)

### 3.7 Construção dos modelos QSAR-4D: Redução de dados e abordagem GFA

Para reduzir o número de descritores gerados na metodologia de QSAR-4D e melhorar sua qualidade, eliminando variáveis espúrias, dois critérios para a seleção de variáveis foram empregados antes da etapa de desenvolvimento dos modelos. No primeiro, os GCODs com variância igual a zero foram eliminados. No segundo, utilizou-se o método de PLS para selecionar os descritores com maior contribuição para explicar a variação da atividade biológica, em que foram selecionados os 2000 com maior peso no PLS.

Em seguida, os GCODs resultantes da redução de dados foram otimizados, juntamente com a atividade biológica, usando o método de GFA (ROGERS; HOPFINGER, 1994). Os cálculos foram iniciados com 100 a 500 modelos e com uma probabilidade de mutação de 50%. O *smoothing factor*, que controla o número de variáveis do modelo, foi variado de 0,1 a 1,0; e o número de cruzamentos de 30,000 a 90,000. Os modelos gerados foram ordenados segundo os valores de LOF e o melhor modelo proveniente de cada alinhamento foi submetido aos métodos de validação.

### 3.8 Validação dos Modelos

Para avaliar a qualidade dos modelos gerados foram empregadas as validações cruzada, externa e Y-randomização. Os parâmetros analisados estão representados na Tabela 1.

Tabela 1 - Parâmetros estatísticos avaliados na metodologia QSAR

Parâmetro	Definição
Fator de desajuste	$\text{LOF} = \frac{\text{LSE}}{\left(1 - \frac{c+d \cdot p}{M}\right)^2}$
Coefficiente de determinação	$R^2 = 1 - \frac{\sum_i (y_{ei} - y_{ci})^2}{\sum_i (y_{ei} - \langle y_e \rangle)^2}$
Coefficiente de determinação ajustado	$R^2_{\text{ajust}} = r^2 - \left(\frac{k-1}{M-k}\right) \cdot (1-r^2)$
Coefficiente de correlação da validação cruzada	$Q^2 = 1 - \frac{\sum_i (y_{ei} - y_{vi})^2}{\sum_i (y_{ei} - \langle y_e \rangle)^2}$
Coefficiente de correlação da validação cruzada ajustado	$Q^2_{\text{ajustado}} = \frac{1 - (1 - q^2)}{\frac{M-1}{M-k}}$
Coefficiente de correlação da validação externa	$R^2_{\text{pred}} = 1 - \frac{\sum_i (y_{eti} - y_{pi})^2}{\sum_i (y_{eti} - \langle y_e \rangle)^2}$
$r^2$ modificado	$r_m^2 = r^2 \cdot \left(1 - \sqrt{(r^2 - r_0^2)}\right)$
$R^2_p$	$R_p^2 = R^2 \cdot \sqrt{R^2 - R_r^2}$

onde: LSE é o erro de mínimos quadrados, c é o número de funções de base, d é o *smoothing factor*, p é o número de variáveis contidas nas funções de base, M o número de compostos no conjunto de treinamento,  $y_e$  valores experimentais de y para o conjunto de treinamento,  $y_c$  valores calculados de y para o conjunto de treinamento,  $y_v$  valores calculados de y da validação cruzada,  $y_{et}$  valores experimentais de y do conjunto de teste,  $y_p$  valores preditos de y do conjunto de teste,  $\langle y_e \rangle$  valor médio de  $y_e$  do conjunto de treinamento, k o número de variáveis do modelo,  $r^2$  é o coeficiente de correlação quadrático entre os valores observados e preditos do conjunto teste,  $r_0^2$  é o coeficiente de correlação quadrático do teste com intercepto em zero, e  $R_r^2$  coeficiente de correlação da randomização.

Fonte: Adaptado de Kiralj (2009)

A validação cruzada implementada no programa 4D-QSAR é a validação *leave-one-out*. Ao final do processo, é gerado o índice estatístico coeficiente de correlação quadrático da validação cruzada,  $Q^2$  (Tabela 1). Além disso, com o intuito de comparar o desempenho de modelos com números diferentes de termos, o coeficiente  $Q^2$  é convertido em  $Q^2_{ajustado}$  (Tabela 1). Para ambos coeficientes, o valor recomendado na literatura é  $> 0,5$  (KIRALJ; FERREIRA, 2009).

A validação externa foi feita com três conjuntos de teste (selecionados como descrito na sessão 3.3). O parâmetro utilizado para avaliar esse processo de validação é o coeficiente de correlação quadrático da validação externa,  $R^2_{pred}$  (Tabela 1). Entretanto, a avaliação desse parâmetro pode não ser suficiente para assegurar a capacidade preditiva, uma vez que ele depende da média das atividades do conjunto de treinamento,  $\langle y_e \rangle$ . Por essa razão, o  $r^2$  modificado ( $r^2_m$ ) foi introduzido (Tabela 1) e, da mesma forma que o  $R^2_{pred}$ , os modelos devem possuir esse coeficiente maior do que 0,5 (ROY et al., 2009).

O processo de Y-randomização, aleatorização do bloco Y, foi repetido por 10 vezes e ao final é gerado o coeficiente de correlação quadrático da randomização. Nesse sentido, o coeficiente  $R^2_p$  (Tabela 1) é utilizado para penalizar as diferenças entre o  $R^2$  do modelo e o  $R^2$  da randomização, onde é recomendado um valor de  $R^2_p > 0,5$  (ROY et al., 2009).

### 3.9 Domínio de aplicabilidade

Para avaliar o domínio de aplicabilidade dos modelos, foi utilizado o gráfico da influência (*leverage*, relacionada com os resíduos do bloco X) *versus* os resíduos de Student (resíduos padronizados da atividade) construído no programa Chemoface (NUNES et al., 2012).

A influência é uma medida da distância do composto ao centro do conjunto de treinamento, baseado no espaço definido pelos descritores do modelo. Compostos identificados como influentes possuem alto poder de alavancagem no modelo e são definidos como *outliers* estruturais. Por outro lado, *outliers* em Y são compostos que não foram bem preditos pelo modelo e, nesse trabalho, são definidos como aqueles cujo módulo do resíduo padronizado é superior a 2,5 (FERREIRA, 2015; GRAMATICA, 2007; TROPSHA, 2010).



### 3.10 Seleção da Conformação Representativa

Nessa etapa é definida uma conformação representativa para cada composto. Para isso, identifica-se todos os confôrmeros de cada composto que estejam dentro de um intervalo de energia  $\Delta E$ . Neste trabalho foi fixado um  $\Delta E = 5,0$  kcal/mol. Em seguida, dentre essas conformações, aquela com maior atividade conforme previsto pelo modelo foi determinada como a conformação representativa (PASQUALOTO et al., 2004).

### 3.11 Ancoramento molecular

Recentemente, a estrutura cristalográfica do transportador de serotonina humano (hSERT) complexada com dois antidepressivos foi depositada no banco de dados de proteína (PDB, do inglês *Protein Data Bank*) (COLEMAN; GREEN; GOUAUX, 2016). Dessa maneira, como forma de compreender melhor a ligação dos derivados de hexahidro-pirrolo-isoquinolina, o composto mais ativo da série foi ancorado dentro do sítio de ligação do transportador. O hSERT pertence à família de transportadores de neurotransmissores acoplado ao co-transporte de sódio. Sua estrutura é composta por doze alfa hélices transmembranares (TM), onde o sítio de ligação central é localizado na cavidade formada pelos resíduos de aminoácido das hélices TM1, TM3, TM6, TM8 e TM10 (COLEMAN; GREEN; GOUAUX, 2016).

De forma geral, a metodologia de ancoramento molecular tem como objetivo prever o modo de ligação de compostos ao sítio ativo do seu alvo biológico partindo do princípio que menores energias representam melhores interações ligante-receptor (GUEDES; DE MAGALHÃES; DARDENNE, 2014). Nesse trabalho, o ancoramento molecular foi feito no programa *Molegro Virtual Docker* (MVD) utilizando a estrutura de raios X do hSERT complexada com paroxetina com resolução de 3,14Å (PDB ID: 5I6X). A função de pontuação de energia empregada foi o MolDock Score [GRID] com resolução de 0,3Å. MolDock Score é uma extensão do potencial PLP (do inglês, *Piecewise Linear Potential*), que leva em consideração interações estéricas e de ligação de hidrogênio (THOMSEN; CHRISTENSEN, 2006). O algoritmo de busca utilizado foi o MolDock SE. Além disso, o número de cálculos foi definido em 100, em que cada cálculo retorna uma pose (conformação).

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Receptor de histamina H<sub>3</sub>

#### 4.1.1 Definição do melhor alinhamento

O conjunto de dados consistiu de derivados de hexahidro-pirrolo-isoquinolina, cujos valores da constante inibitória estão disponíveis na literatura (APODACA et al., 2012). Os compostos foram divididos em dois grupos, o conjunto de treinamento, o qual foi utilizado na construção dos modelos de QSAR; e que foi validado estatisticamente pelo conjunto de teste (contendo aproximadamente 20% dos compostos). Além disso, como forma de assegurar a qualidade dos modelos, foram analisados três conjuntos de teste.

A análise GFA gera vários modelos, no entanto, será analisado apenas o melhor modelo proveniente de cada alinhamento. Na Tabela 2 estão representados os parâmetros estatísticos médios dos modelos de QSAR-4D (a tabela completa com os valores de cada conjunto de teste é dada no Apêndice A).

Tabela 2 - Parâmetros estatísticos médios da calibração, validação interna e externa dos modelos provenientes dos dez alinhamentos testados para o receptor H<sub>3</sub>.

Alig	Des	R <sup>2</sup>	R <sup>2</sup> <sub>ajus</sub>	Q <sup>2</sup>	Q <sup>2</sup> <sub>ajus</sub>	R <sup>2</sup> <sub>pred</sub>	r <sup>2</sup> <sub>m</sub>	R <sup>2</sup> <sub>rand</sub>	R <sub>p</sub>
1	7	0,789	0,774	0,750	0,731	0,785	0,732	0,088	0,744
2	7	0,800	0,786	0,755	0,737	0,755	0,546	0,085	0,757
3	7	0,787	0,771	0,722	0,701	0,768	0,676	0,080	0,746
4	7	0,808	0,794	0,736	0,716	0,785	0,670	0,073	0,771
5	6	0,753	0,738	0,696	0,677	0,726	0,578	0,076	0,714
6	7	0,784	0,768	0,743	0,724	0,743	0,588	0,077	0,744
7	6	0,776	0,762	0,729	0,713	0,779	0,700	0,073	0,738
8	6	0,764	0,749	0,684	0,664	0,730	0,542	0,063	0,731
9	7	0,765	0,747	0,720	0,700	0,772	0,675	0,077	0,725
10	6	0,742	0,726	0,656	0,634	0,814	0,754	0,068	0,707

Analisando a tabela 2, observa-se que todos os modelos apresentam valores satisfatórios de R<sup>2</sup> e R<sup>2</sup><sub>ajus</sub>, variando de 0,742 a 0,808 e 0,726 a 0,794, respectivamente. Esses parâmetros refletem o grau de ajuste dos dados ao modelo, onde o R<sup>2</sup><sub>ajus</sub> é utilizado para comparar modelos com número de variáveis diferentes. A significância estatística de cada componente do modelo é avaliada pela validação cruzada. Os modelos também apresentam valores satisfatórios da validação cruzada, Q<sup>2</sup> e Q<sup>2</sup><sub>ajus</sub>, variando de 0,656 – 0,755 e 0,634 – 0,737, respectivamente. Para ambos os parâmetros, recomenda-se valores superiores à 0,5

(KIRALJ; FERREIRA, 2009). Quando se analisa os valores da randomização, todos os modelos possuem o  $R^2_{\text{rand}}$  inferior ao  $R^2$  do modelo e  $R^2_{\text{p}}$  superior ao valor recomendado na literatura ( $R^2_{\text{p}} > 0,5$ )(ROY et al., 2009). Por fim, no que se refere à validação externa, os coeficientes analisados são o  $R^2_{\text{pred}}$  e o  $r^2_{\text{m}}$ , cujos valores recomendados são superiores à 0,5 (ROY et al., 2009). Novamente, todos os modelos preenchem os requisitos necessários.

Uma vez que todos os modelos possuem valores estatísticos satisfatórios, a escolha do melhor alinhamento será feita com base no melhor balanço em todas as validações. Podendo assim destacar os alinhamentos 1 e 4, que estão entre os modelos com os maiores valores da calibração, validação cruzada e externa (Alinhamento 1:  $R^2=0,789$ ;  $Q^2=0,750$  e  $R^2_{\text{pred}}=0,785$ . Alinhamento 4:  $R^2=0,808$ ;  $Q^2=0,736$  e  $R^2_{\text{pred}}=0,785$ ). Entretanto, quando se analisa o domínio de aplicabilidade, o modelo proveniente do alinhamento 4 possui menor número de compostos *outliers*. Portanto, o alinhamento 4 foi escolhido para representar o conjunto de dados e prosseguir com a análise dos descritores.

#### 4.1.2 Análise dos descritores

A análise dos descritores contidos no modelo selecionado na seção anterior, proveniente do alinhamento 4, será feita no intuito de compreender as características estruturais importantes para a ligação dos compostos ao receptor de histamina  $H_3$ . Os modelos foram construídos utilizando  $pK_I$  como a variável resposta e os descritores da metodologia de QSAR-4D como as variáveis independentes. Na metodologia de QSAR-4D, os descritores são representados pelas coordenadas cartesianas e pelos elementos de interação farmacofórica, “(x,y,z,IPE)”, que correspondem as interações que podem ocorrer entre o ligante e o receptor. Nesse sentido, descritores com coeficiente positivo representam interações favoráveis entre o ligante-receptor, ao passo que descritores com coeficiente negativo indicam interações desfavoráveis.

O modelo de QSAR-4D do alinhamento 4 está representado na equação 9:

$$pK_I = 9,953 - 55,358 (2,-7,-2,p-) + 28,588 (-6,-8,-2,any) + 23,059 (4,-9,-1,any) + 4,962 (1,4,1,np) + 14,761 (-6,-8,2,np) - 2,458 (0,1,-1,any) - 50,054 (3,-3,-5,np) \quad (9)$$

$$M = 86 \quad LOF = 0.151 \quad LSE = 0.115$$

O modelo descrito acima é composto por sete descritores, quatro deles com coeficiente positivo e três com coeficiente negativo. Como forma de entender melhor a distribuição dos

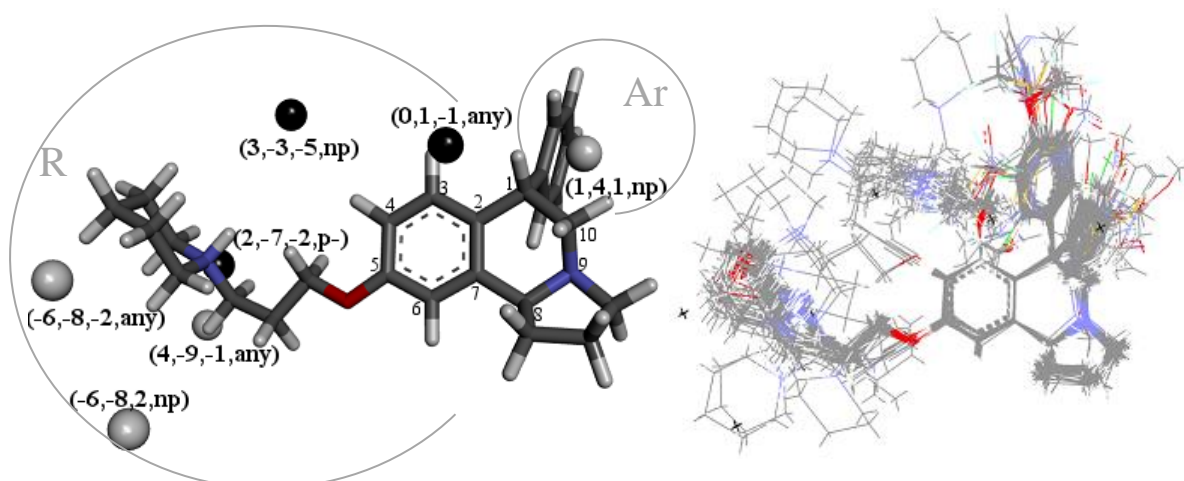
dados, a matriz de correlação entre os diferentes descritores é dada na Tabela 3, em que são considerados correlacionados aqueles descritores com  $r > 0,7$  (DA CUNHA et al., 2005).

Tabela 3 - Matriz de correlação cruzada entre os descritores do alinhamento 4.

	(2,-7,-2, p-)	(-6,-8,-2, any)	(4,-9,-1, any)	(1,4,1, np)	(-6,-8,2, np)	(0,1,-1, any)	(3,-3,-5, np)
(2,-7,-2,p-)	1,000						
(-6,-8,-2,any)	0,394	1,000					
(4,-9,-1,any)	-0,055	-0,107	1,000				
(1,4,1,np)	-0,182	-0,151	-0,080	1,000			
(-6,-8,2,np)	-0,158	-0,077	-0,067	0,043	1,000		
(0,1,-1,any)	-0,256	-0,328	-0,140	0,571	-0,078	1,000	
(3,-3,-5,np)	-0,229	-0,299	-0,083	0,575	-0,109	0,808	1,000

A Figura 5 mostra a posição dos descritores no composto mais ativo da série (à esquerda) juntamente com a sobreposição de todos os compostos do conjunto de dados (à direita). Os GCODs positivos e negativos são representados por esferas claras e escuras, respectivamente. Os valores das frequências de ocupação de todos os descritores são dados no Apêndice B. Além disso, na Figura 6 os descritores estão representados nos seus respectivos compostos com maior frequência de ocupação.

Figura 5 - (Esquerda) Descritores do alinhamento 4 representados no composto 1B, o mais ativo da série. (Direita) Sobreposição dos compostos do conjunto de dados.



Seis dos descritores estão localizados próximos ao substituinte R dos compostos, o qual varia quanto a sua posição no anel aromático e composição. O GCOD (0,1,-1,any) possui coeficiente negativo, um tipo de interação não específica e maior frequência de ocupação para

compostos com o substituinte R nas posições 3 e 4 do anel aromático (Fig.5). Portanto, esse descritor sugere que grupos nessas posições são prejudiciais para a atividade dos compostos. De fato, examinando os valores de atividade, a maioria dos compostos que apresentam o substituinte R na posição 3 possuem baixos valores de  $pK_I$ . Por exemplo, os compostos 1B e 1C e 4A e 4C possuem diferença de atividade de aproximadamente duas unidades logarítmicas. A mesma tendência pode ser observada quando o grupo R está na posição 4. Comparando os compostos 1A/45 e 1B/44, que variam apenas na posição de R, aqueles que possuem esse substituinte conectado ao carbono 4 apresentam menor valor de atividade. Única exceção quando um grupo alquino é adicionado à cadeia carbônica nos compostos 97A e 98A. Nesse caso, os dois compostos possuem valores semelhantes de  $pK_I$ .

Além disso, o GCOD (0,1,-1,any) possui um coeficiente de correlação de 0,808 com o GCOD (3,-3,-5,np). Esse descritor representa uma interação não polar, possui um coeficiente negativo e frequência de ocupação apenas para compostos com o grupo R na posição 3 (Fig.5). Indicando, de forma análoga ao descritor anterior, que substituintes nessa posição são prejudiciais a atividade biológica dos compostos em estudo.

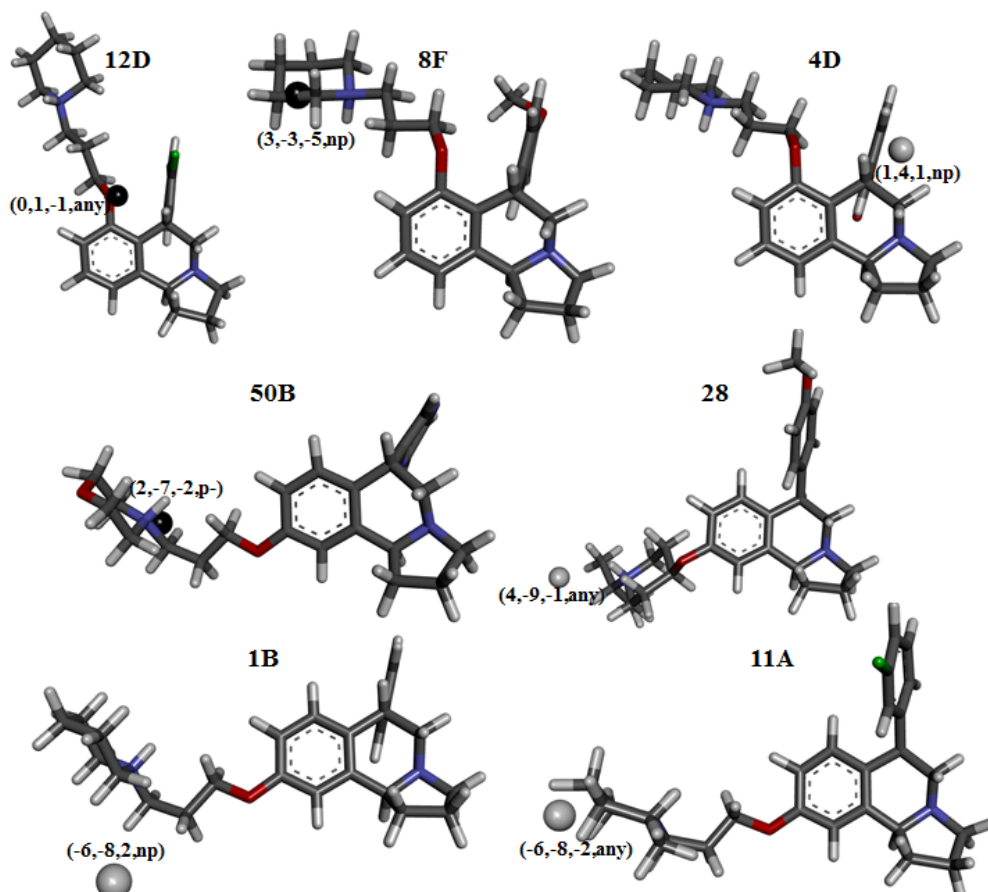
O GCOD (-6,-8,-2,any) apresenta coeficiente positivo, um tipo de interação não específica e maior frequência de ocupação para o composto 11A, o qual possui o grupo R na posição 5 do anel aromático (Fig.5). Esse GCOD está localizado próximo à porção final do grupo piperidina desse composto (Fig.6), indicando que substituintes nessa posição são favoráveis a atividade. O que corrobora com as informações fornecidas pelos dois descritores anteriores.

O GCOD (2,-7,-2,p-) apresenta um coeficiente negativo, elemento de interação farmacofórica polar com carga negativa e possui maior frequência de ocupação para as moléculas 11A e 50B. Nessas moléculas, esse descritor se encontra próximo ao nitrogênio do grupo piperidina e morfolina, respectivamente (Fig.6). Todavia, como todos os compostos possuem um átomo de nitrogênio nessa posição, acredita-se que a informação desse descritor esteja relacionada a orientação desse grupo. Uma das características estruturais de antagonistas  $H_3$  é a presença de pelo menos um grupo amino básico (LEBOIS; JONES; LINDSLEY, 2011), o qual aparentemente realiza interação iônica importante com resíduo do sítio ativo carregado negativamente (KISS; KESERÚ, 2016). Com isso, o GCOD (2,-7,-2,p-) parece estar relacionado ao impacto que a conformação adotada pelo substituinte R tem na atividade dos compostos.

Os descritores (4,-9,-1,any) e (-6,-8,2,np) também estão próximos ao substituinte R na posição 5. Ambos possuem coeficiente positivo, sendo que o primeiro representa uma

interação não específica e o último uma interação não polar. Entretanto, esses descritores estão um pouco afastados da estrutura dos compostos, podendo assim estarem relacionados à flexibilidade do substituinte R, que experimenta uma grande liberdade conformacional na etapa de simulação de dinâmica molecular, como pode ser observado na figura 5.

Figura 6 - Os descritores do alinhamento 4 representados nos compostos com respectivo maior valor de frequência de ocupação.



Por fim, o GCOD (1,4,1,np) está localizado próximo ao substituinte aromático dos compostos em estudo (Fig. 5). Esse descritor possui coeficiente positivo, elemento de interação farmacofórica não polar e maior frequência de ocupação para o composto 4D (Fig.6). Interessante notar que esse descritor possui frequência de ocupação apenas para compostos em que o anel aromático está na posição *trans*, independente da natureza química e da posição do grupo substituinte no anel. De fato, em geral, compostos com o anel aromático na posição *trans* são igualmente ou mais potentes que seus correspondentes na posição *cis*. Como é o caso, por exemplo, dos compostos 1A/1B, 5A/5B e 7A/7B. Além disso, a constante inibitória não é muito sensível a posição do substituinte no anel aromático. Por exemplo,

quando a hidroxila está na posição *para* (9A,  $pK_I = 8,699$ ), *meta* (15,  $pK_I = 9,155$ ) ou *orto* (16,  $pK_I = 8,398$ ), a atividade tem variação menor que uma unidade logarítmica.

Lebois, Jones e Lindsley (2011) desenvolveram um modelo farmacofórico de antagonistas H<sub>3</sub>R em que os autores identificaram características conservadas entre os compostos. De forma geral, as estruturas constituem de uma amina básica, uma cadeia alquila, um núcleo aromático conectado à uma porção final. Esse grupamento final possui grande diversidade estrutural, podendo acomodar um grupo polar, uma segunda amina básica ou resíduo lipofílico (LEBOIS; JONES; LINDSLEY, 2011). Nesse sentido, as informações fornecidas pelo descritor (1,4,1,np) concordam com o modelo farmacofórico proposto por esses autores, onde a natureza dos substituintes no anel aromático tem pouco efeito na atividade biológica dos compostos.

A Eq. 6 foi utilizada para computar os valores de  $pK_I$  preditos. O  $pK_I$  experimental, predito e o valor residual para o treinamento e teste estão contidos nas Tabela 4.

Tabela 4 - Valores experimentais e preditos da constante inibitória juntamente com o resíduo dos compostos do conjunto de treinamento e teste (\*) para o H<sub>3</sub>R. (Continua)

Comp	$pK_I$	$pK_I$ pred	Res	Comp	$pK_I$	$pK_I$ pred	Res	Comp	$pK_I$	$pK_I$ pred	Res
<b>1A</b>	9,046	8,713	0,333	<b>13B</b>	6,889	6,909	-0,020	<b>44</b>	9,000	8,338	0,662
<b>1B</b>	9,699	9,595	0,104	<b>14A</b>	8,699	8,654	0,045	<b>45</b>	7,796	7,834	-0,038
<b>1C</b>	7,167	6,741	0,426	<b>14B*</b>	9,000	9,003	-0,003	<b>48A</b>	8,222	8,308	-0,086
<b>2A</b>	9,000	8,365	0,635	<b>14C*</b>	6,699	6,794	-0,095	<b>48B*</b>	8,301	9,041	-0,740
<b>2B</b>	7,509	8,174	-0,665	<b>15</b>	9,155	9,020	0,135	<b>49</b>	8,301	8,362	-0,061
<b>2C</b>	7,538	7,883	-0,345	<b>16</b>	8,398	8,532	-0,134	<b>50A*</b>	8,398	8,458	-0,060
<b>3</b>	9,301	8,822	0,479	<b>17*</b>	8,398	7,667	0,731	<b>50B</b>	6,742	7,508	-0,766
<b>4A</b>	9,000	9,259	-0,259	<b>18</b>	9,046	8,627	0,419	<b>51A*</b>	8,523	8,517	0,006
<b>4B</b>	9,000	8,424	0,576	<b>19</b>	9,000	8,238	0,762	<b>51B</b>	8,301	7,955	0,346
<b>4C*</b>	7,009	6,843	0,166	<b>20</b>	9,000	8,795	0,205	<b>52</b>	8,398	8,542	-0,144
<b>4D</b>	9,000	8,738	0,262	<b>21</b>	9,000	9,200	-0,200	<b>53</b>	8,523	8,886	-0,363
<b>5A*</b>	8,699	8,888	-0,189	<b>22*</b>	9,097	8,893	0,204	<b>54</b>	8,398	8,783	-0,385
<b>5B</b>	9,000	8,747	0,253	<b>23*</b>	9,000	8,317	0,683	<b>55*</b>	8,699	8,721	-0,022
<b>5C</b>	7,071	6,803	0,268	<b>24A*</b>	9,155	8,645	0,510	<b>56</b>	8,523	8,237	0,286
<b>6A</b>	8,523	8,899	-0,376	<b>24B</b>	9,523	8,808	0,715	<b>57</b>	8,523	8,261	0,262
<b>6B</b>	9,000	9,141	-0,141	<b>25</b>	9,097	8,970	0,127	<b>58</b>	8,699	8,471	0,228
<b>6C*</b>	7,119	6,628	0,491	<b>26</b>	9,097	8,821	0,276	<b>59</b>	8,699	8,257	0,442
<b>7A</b>	8,301	8,593	-0,292	<b>27A</b>	9,000	8,810	0,190	<b>60*</b>	8,523	8,562	-0,039

Tabela 4 - Valores experimentais e preditos da constante inibitória juntamente com o resíduo dos compostos do conjunto de treinamento e teste (\*) para o H<sub>3</sub>R. (Conclusão)

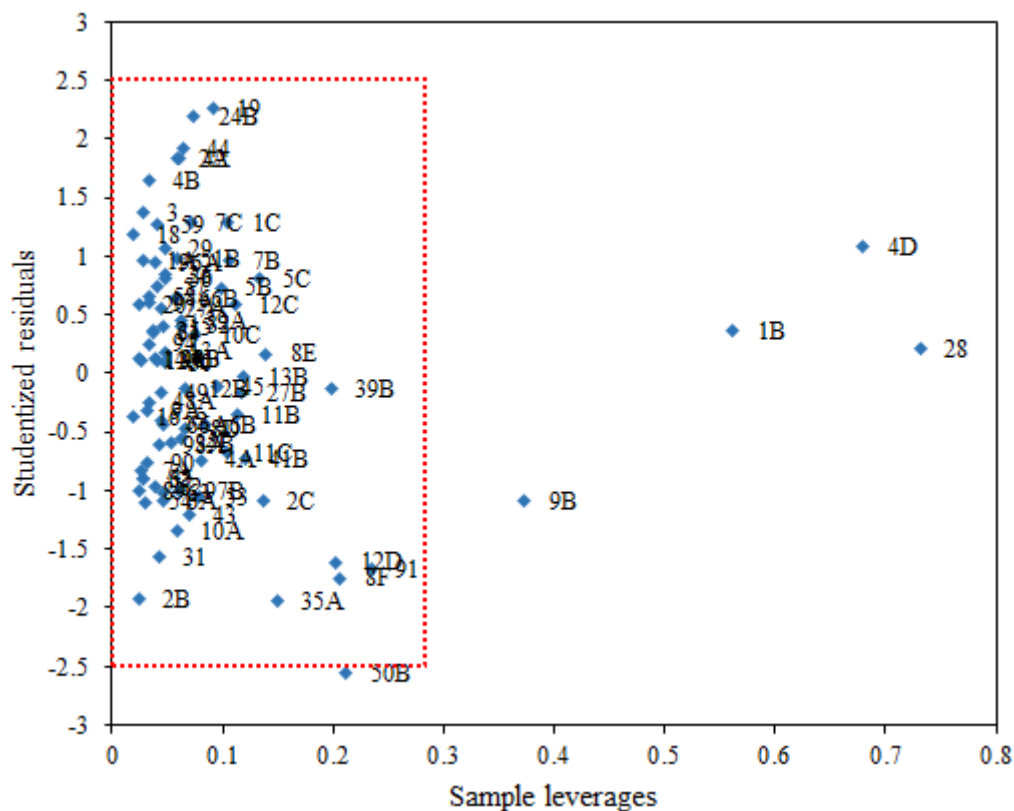
Comp	pK <sub>I</sub>	pK <sub>I</sub> pred	Res	Comp	pK <sub>I</sub>	pK <sub>I</sub> pred	Res	Comp	pK <sub>I</sub>	pK <sub>I</sub> pred	Res
<b>7B</b>	8,699	8,362	0,337	<b>27B</b>	9,523	9,567	-0,044	<b>61*</b>	7,538	7,620	-0,082
<b>7C</b>	7,509	7,070	0,439	<b>28</b>	9,398	9,430	-0,032	<b>62</b>	8,222	8,562	-0,340
<b>8D</b>	9,000	9,154	-0,154	<b>29</b>	9,046	8,675	0,371	<b>63</b>	7,796	8,153	-0,357
<b>8E</b>	6,646	6,603	0,043	<b>30</b>	8,699	8,662	0,037	<b>64</b>	8,523	8,404	0,119
<b>8F</b>	6,317	6,866	-0,549	<b>31</b>	7,921	8,468	-0,547	<b>65</b>	8,398	8,720	-0,322
<b>9A</b>	8,699	8,814	-0,115	<b>32</b>	9,000	8,857	0,143	<b>66A</b>	8,699	8,850	-0,151
<b>9B</b>	6,000	6,321	-0,321	<b>34*</b>	8,699	8,938	-0,239	<b>66B</b>	9,000	8,774	0,226
<b>10A</b>	8,699	9,166	-0,467	<b>35A</b>	8,398	9,061	-0,663	<b>67*</b>	8,699	8,662	0,037
<b>10B*</b>	8,699	8,930	-0,231	<b>35B</b>	8,699	8,902	-0,203	<b>89</b>	7,796	8,150	-0,354
<b>10C</b>	7,086	6,976	0,110	<b>37</b>	8,398	8,720	-0,322	<b>90</b>	8,000	8,269	-0,269
<b>11A*</b>	9,000	8,762	0,238	<b>38*</b>	8,222	8,424	-0,202	<b>91</b>	7,194	7,697	-0,503
<b>11B</b>	9,000	9,103	-0,103	<b>39A</b>	8,699	8,534	0,165	<b>92*</b>	7,796	8,286	-0,490
<b>11C</b>	6,866	7,098	-0,232	<b>39B</b>	9,000	9,031	-0,031	<b>94</b>	8,398	8,311	0,087
<b>12A</b>	8,699	8,665	0,034	<b>40</b>	8,523	8,491	0,032	<b>96A</b>	9,000	8,671	0,329
<b>12B</b>	9,000	9,041	-0,041	<b>41A*</b>	9,000	8,729	0,271	<b>97A</b>	8,523	8,311	0,212
<b>12C</b>	7,041	6,853	0,188	<b>41B</b>	9,000	9,245	-0,245	<b>97B</b>	7,921	8,264	-0,343
<b>12D</b>	7,131	7,634	-0,503	<b>42</b>	8,699	8,071	0,628	<b>98A</b>	8,398	8,626	-0,228
<b>13A</b>	9,000	8,933	0,067	<b>43</b>	7,523	7,937	-0,414	<b>98B</b>	8,523	8,478	0,045

Além disso, foi definido o domínio de aplicabilidade do QSAR, com intuito de identificar compostos com comportamento distinto do restante do treinamento. Desse modo, o gráfico de influência *versus* resíduo de Student está representado na figura 7.

Analisando o valor residual, o composto 50B se encontra um pouco abaixo do limite inferior. A atividade desse composto é superestimada pelo modelo (resíduo igual à 0,766, tabela 4). Compostos com o grupo piridina em geral possuem altos valores de atividade (pK<sub>I</sub> > 8). A adição do segundo nitrogênio, formando o grupo pirazina do composto 50B, reduz seu pK<sub>I</sub> para 6,742. Portanto, parece que o modelo teve dificuldade em modular essa diferença de comportamento. No que se refere à influência, quatro amostras possuem valor superior ao limite estabelecido, são elas os compostos 9B, 1B, 4D e 28. Compostos com valores altos possuem grande influência na construção do modelo e são definidos como *outliers* estruturais. No composto 28, por exemplo, que possui a influência mais extrema, o substituinte R é diferente de todos os demais compostos do conjunto de treinamento, como pode ser observado no quadro 1.



Figura 7 - Domínio de aplicabilidade para o modelo de QSAR do H<sub>3</sub>R representado pelo gráfico de influência *versus* resíduo de Student.



## 4.2 Transportador de serotonina

### 4.2.1 Definição do melhor alinhamento

O mesmo procedimento descrito para o H<sub>3</sub>R foi aplicado ao transportador de serotonina. Os derivados de hexahidro-pirroló-isoquinolina foram divididos em conjunto de treinamento e teste (contendo aproximadamente 20% dos compostos). Na Tabela 5 estão os parâmetros estatísticos médios para os dez alinhamentos testados (a tabela completa se encontra no Apêndice C).

Os modelos descritos na tabela 5 apresentam valores de  $R^2$  e  $R^2_{ajus}$  variando de 0,613 a 0,818 e 0,544 a 0,786, respectivamente. Dentre os alinhamentos avaliados, os alinhamentos 1, 5 e 6 possuem  $R^2$  abaixo de 0,7. Além disso, os alinhamentos 1, 2, 5, 6, 8 e 9 possuem  $R^2_{ajus} < 0,7$ . Por essa razão, esses alinhamentos não serão considerados em análises futuras. Todos os quatro alinhamentos restantes apresentam  $Q^2$  e  $Q^2_{ajus}$  superiores à 0,5. No que se refere à randomização, os alinhamentos também possuem valores estatísticos satisfatórios, com  $R^2_{rand}$  inferior ao  $R^2$  do modelo e  $R^2_p$  maior que 0,5. Por fim, os parâmetros da validação externa, apenas os alinhamentos 4 e 7 possuem  $R^2_{pred} > 0,5$ , de 0,769 e 0,756, respectivamente. Porém,

quando se avalia o valor de  $r^2_m$ , apenas o alinhamento 7 preenche os requisitos necessários, com  $r^2_m = 0,580$ . Com isso, o alinhamento 7 foi selecionado para representar o conjunto de dados e prosseguir com a análise dos descritores.

Tabela 5 - Parâmetros estatísticos médios da calibração, validação interna e externa dos modelos provenientes dos dez alinhamentos testados para o SERT.

Alig	Des	R <sup>2</sup>	R <sup>2</sup> <sub>ajus</sub>	Q <sup>2</sup>	Q <sup>2</sup> <sub>ajus</sub>	R <sup>2</sup> <sub>pred</sub>	r <sup>2</sup> <sub>m</sub>	R <sup>2</sup> <sub>rand</sub>	R <sub>p</sub>
1	4	0,680	0,648	0,566	0,523	0,340	0,108	0,119	0,616
2	5	0,732	0,695	0,620	0,568	0,333	0,118	0,159	0,648
3	6	0,754	0,711	0,607	0,537	0,351	0,232	0,166	0,666
4	6	0,761	0,718	0,607	0,537	0,769	0,478	0,196	0,655
5	6	0,613	0,544	0,360	0,245	0,214	0,031	0,173	0,519
6	4	0,642	0,606	0,469	0,416	0,353	0,144	0,115	0,581
7	6	0,818	0,786	0,710	0,659	0,756	0,580	0,169	0,728
8	6	0,708	0,656	0,569	0,492	0,399	0,095	0,156	0,625
9	6	0,729	0,681	0,503	0,414	0,118	0,000	0,184	0,630
10	6	0,757	0,714	0,630	0,564	0,475	0,105	0,215	0,641

#### 4.2.2 Análise dos descritores

O modelo de QSAR-4D do alinhamento 7 está representado na equação 10:

$$\begin{aligned}
 pK_I = & 8,933 + 28,270 (6,-7,-2,any) - 17,795 (5,-7,-3,np) - 11,757 (1,0,7,any) \\
 & + 2,514 (2,-1,5,np) + 9,725 (2,-3,10,np) + 2,091 (4,-2,4,any) \quad (10) \\
 M = & 34 \quad \quad \quad LOF = 0.107 \quad \quad \quad LSE = 0.051
 \end{aligned}$$

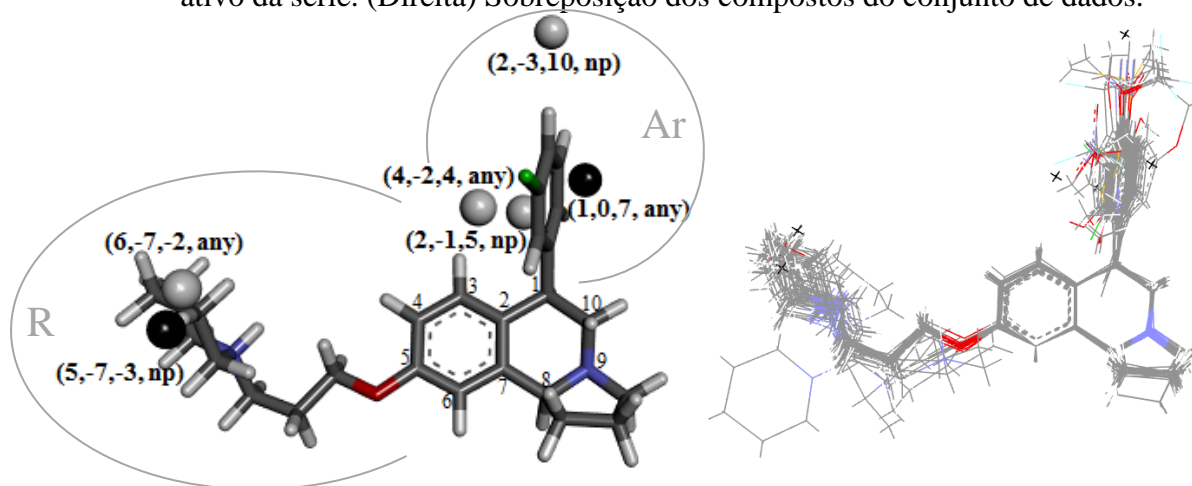
O modelo descrito acima é composto por seis descritores, dois deles com coeficiente negativo, que correspondem a interações desfavoráveis entre ligante-receptor, e quatro descritores com coeficiente positivo, indicando interações favoráveis. Também foi feita a matriz de correlação entre os descritores do modelo (Tabela 6), com o intuito de entender melhor a distribuição dos dados.

A posição dos descritores no composto mais ativo da série juntamente com a sobreposição de todos os compostos do conjunto de dados é ilustrada na Figura 8. Os valores das frequências de ocupação dos descritores da Eq. 10 são fornecidos no Apêndice D. Ademais, na Figura 9 estão representados os descritores nos seus respectivos compostos com maior frequência de ocupação.

Tabela 6 - Matriz de correlação cruzada entre os descritores do alinhamento 7.

	(6,-7,-2, any)	(5,-7,-3, np)	(1,0,7, any)	(2,-1,5, np)	(2,-3,10, np)	(4,-2,4, any)
(6,-7,-2,any)	1,000					
(5,-7,-3,np)	0,865	1,000				
(1,0,7,any)	-0,127	-0,132	1,000			
(2,-1,5,np)	0,001	-0,027	-0,163	1,000		
(2,-3,10,np)	-0,284	-0,378	0,291	-0,013	1,000	
(4,-2,4,any)	0,038	0,165	0,310	-0,178	-0,075	1,000

Figura 8 - (Esquerda) Descritores do alinhamento 7 representados no composto 11A, o mais ativo da série. (Direita) Sobreposição dos compostos do conjunto de dados.



Como pode ser observado na figura 8, quatro dos descritores moleculares estão localizados ao redor do substituinte aromático dos compostos analisados. O GCOD (1,0,7,any) apresenta coeficiente negativo, tipo de interação não específica e maior frequência de ocupação para a molécula 18 (Fig.9). Esse descritor possui frequência de ocupação para todos os compostos, exceto para o 92. Esse composto é o único com estereoquímica *S* do carbono ligado ao anel aromático. Comparando-se as estruturas 38 ( $pK_I = 8,155$ ), 92 ( $pK_I = 8,398$ ) e 94 ( $pK_I = 6,955$ ), que variam apenas no arranjo espacial dos grupos, o composto 92 é o mais ativo. Enquanto que o composto 94, com estereoquímica *R* nos dois carbonos quirais, é um dos menos ativos da série. Com isso, o descritor (1,0,7,any) parece destacar o efeito da estereoquímica na potência dos compostos. Entretanto, existe pouca informação no conjunto de dados referente a isto, o que não permite afirmar claramente seu papel para a atividade desses compostos.

O GCOD (2,-1,5,np) representa uma interação não polar, possui coeficiente positivo e maior frequência de ocupação para a molécula 23 (Fig.9). Nesse composto, o descritor se

encontra sobre o átomo de enxofre do tiofeno, indicando que grupos nessa posição são favoráveis para a atividade dos compostos. Interessante notar que o descritor (2,-1,5,np) não possui ocupação para compostos com o substituinte na posição *orto* do anel aromático. De fato, quando se compara a diferença de atividade conforme a posição do substituinte no anel é variada, aqueles compostos com grupo *orto* possuem menores valores de  $pK_I$ . Como é o caso, por exemplo, dos compostos 9A ( $pK_I = 7,237$ ), 15 ( $pK_I = 8,523$ ) e 16 ( $pK_I = 6,951$ ), cuja diferença consiste apenas na posição da hidroxila no anel, em que o composto 16 (OH *orto*) possui a menor atividade do conjunto de dados. Comportamento semelhante também pode ser observado para os compostos 66A ( $pK_I = 8,222$ ), 11A ( $pK_I = 9,222$ ) e 12A ( $pK_I = 8,097$ ).

O descritor (4,-2,4,any) apresenta coeficiente positivo, interação não específica e maior frequência de ocupação para o composto 67 (Fig.9). O GCOD está posicionado próximo ao substituinte *meta* desse composto, indicando que grupos nessa posição são favoráveis para a atividade biológica. Em geral, grupos polares na posição *meta* possuem maiores valores de  $pK_I$ . Alguns exemplos são as estruturas: 67 ( $pK_I = 8,046$ ) e 7A ( $pK_I = 7,921$ ), 15 ( $pK_I = 8,523$ ) e 9A ( $pK_I = 7,237$ ), bem como 11A ( $pK_I = 9,222$ ) e 66A ( $pK_I = 8,222$ ). Em todos os casos, aquelas estruturas com substituinte *meta* são mais ativas que seus correspondentes na posição *para*.

O GCOD (2,-3,10,np) representa interação não polar, possui coeficiente positivo e maior frequência de ocupação para o composto 98A (Fig.9). Nesse composto, o descritor está próximo ao metil do tioéter *para* substituído, sugerindo a influência desse grupo nessa posição para a atividade dos compostos. Quando se analisa os valores de  $pK_I$ , aqueles compostos com éter ou tioéter na posição *para* possuem altos valores de atividade ( $pK_I > 8$ ).

Os dois últimos descritores, (6,-7,-2,any) e (5,-7,-3,np), encontram-se próximos ao grupo R dos compostos e possuem correlação de 0,865. O primeiro apresenta coeficiente positivo, interação não específica e maior frequência de ocupação para os compostos 11A e 10A (Fig.9), onde está localizado perto da porção piperidina. Já o GCOD (5,-7,-3,np) representa uma interação não polar, possui coeficiente negativo e maior frequência de ocupação para o composto 4A (Fig.9), onde também se encontra próximo à piperidina. Com isso, esses descritores parecem estar relacionados a orientação desse substituinte no espaço.

A Eq. 10 foi utilizada para computar os valores de  $pK_I$  preditos. O  $pK_I$  experimental, predito e o valor residual para o treinamento e teste estão contidos nas Tabela 7.

Figura 9 - Os descritores do alinhamento 7 representados nos compostos com respectivo maior valor de frequência de ocupação.

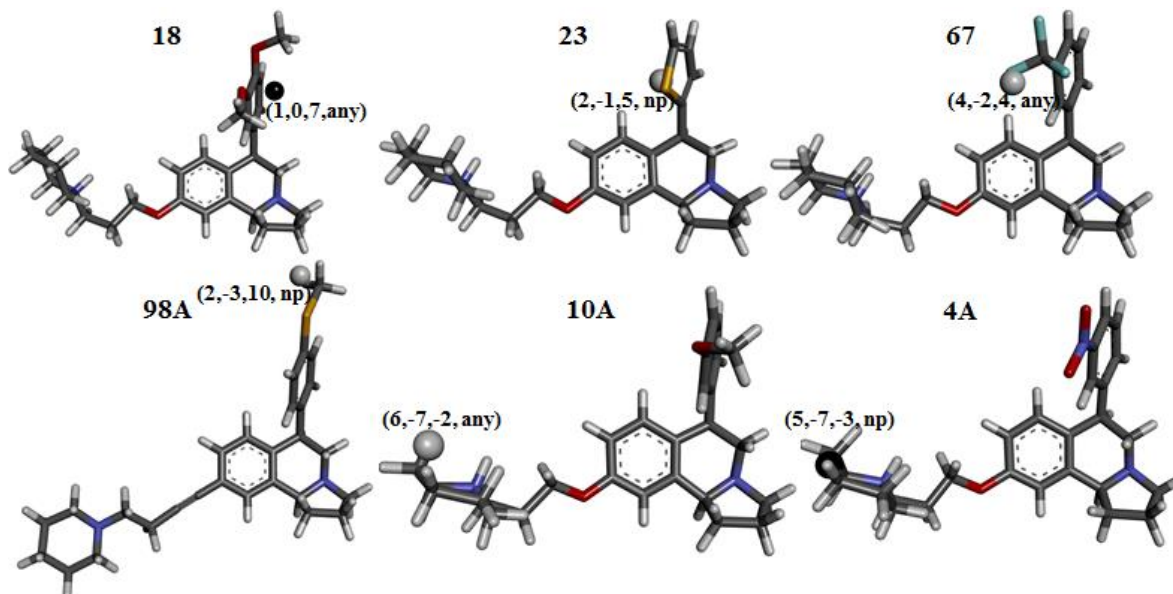
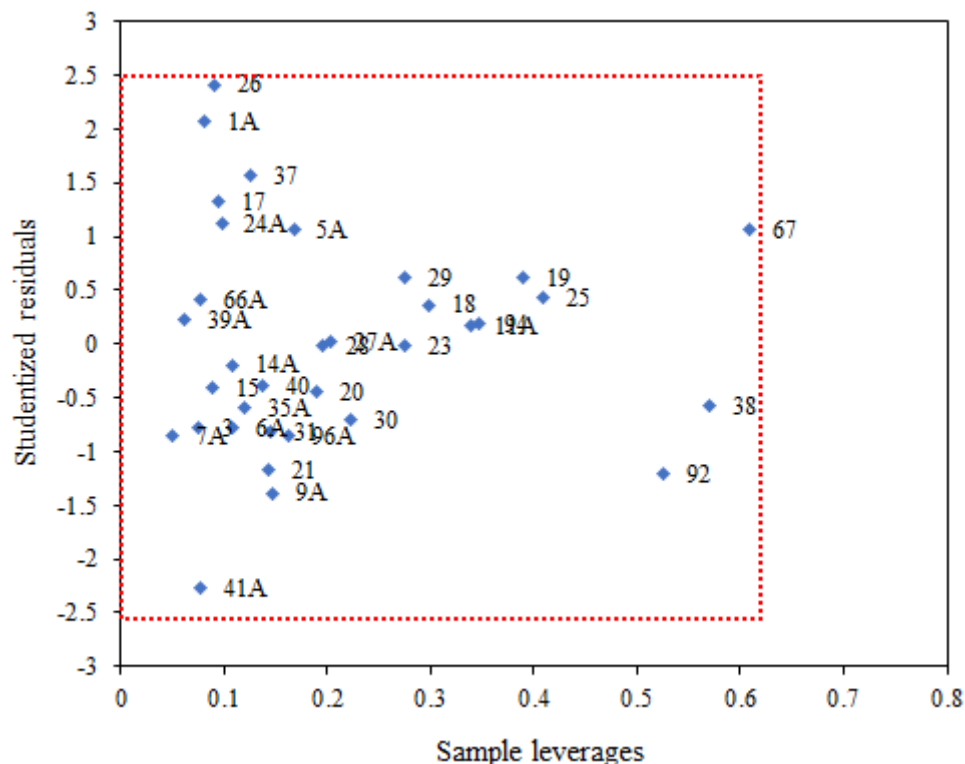


Tabela 7 - Valores experimentais e preditos da constante inibitória juntamente com o resíduo dos compostos do conjunto de treinamento e teste (\*) para o SERT.

Comp	$pK_I$	$pK_I$ pred	Res	Comp	$pK_I$	$pK_I$ pred	Res	Comp	$pK_I$	$pK_I$ pred	Res
1A	8,398	8,068	0,330	17	8,155	7,826	0,329	32*	7,959	7,467	0,492
2A*	8,222	7,896	0,326	18	7,886	7,652	0,234	34*	8,523	8,748	-0,225
3	8,097	8,324	-0,227	19	8,301	8,010	0,291	35A	8,155	8,327	-0,172
4A*	8,398	7,662	0,736	20	7,367	7,513	-0,146	37	9,097	8,796	0,301
5A	8,398	8,185	0,213	21	8,301	8,518	-0,217	38	8,155	8,377	-0,222
6A	8,398	8,470	-0,072	23	8,699	8,597	0,102	39A	8,398	8,355	0,043
7A	7,921	8,134	-0,213	24A	9,000	8,783	0,217	40	7,745	7,759	-0,014
9A	7,237	7,575	-0,338	25	7,745	7,681	0,064	41A	7,229	7,749	-0,520
10A*	8,699	8,847	-0,148	26	8,222	7,756	0,466	66A	8,222	8,160	0,062
11A	9,222	9,027	0,195	27A	7,824	7,983	-0,159	67	8,046	7,927	0,119
12A*	8,097	8,390	-0,293	28	8,301	8,235	0,066	92	8,398	8,826	-0,428
13A*	7,678	7,172	0,506	29	8,699	8,502	0,197	94	6,955	6,886	0,069
14A	8,523	8,657	-0,134	30	8,699	8,712	-0,013	96A	8,046	8,162	-0,116
15	8,523	8,604	-0,081	31	7,553	7,711	-0,158	98A*	8,699	9,306	-0,607
16*	6,951	6,402	0,549								

Da mesma maneira que para o H<sub>3</sub>R, foi definido o domínio de aplicabilidade do modelo de QSAR descrito na eq. 10, o qual está ilustrado na figura 10. Como pode ser observado, todos os compostos se encontram dentro da região delimitada, não havendo assim nenhum *outlier*.

Figura 10 - Domínio de aplicabilidade para o modelo de QSAR do SERT representado pelo gráfico de influência *versus* resíduo de Student.



#### 4.2.3 Ancoramento molecular

Recentemente, a estrutura cristalográfica do hSERT ligado aos antidepressivos escitalopram e paroxetina foi reportada (COLEMAN; GREEN; GOUAUX, 2016). Com isso, para entender melhor o comportamento dos derivados de hexahidro-pirroló-isoquinolina analisados nesse trabalho, o composto 11A (mais ativo da série) foi ancorado no sítio ativo do transportador utilizando o programa MVD (THOMSEN; CHRISTENSEN, 2006). As energias obtidas estão representadas na Tabela 8.

Tabela 8: Energias obtidas no ancoramento molecular do composto 11A no sítio de ligação central do hSERT.

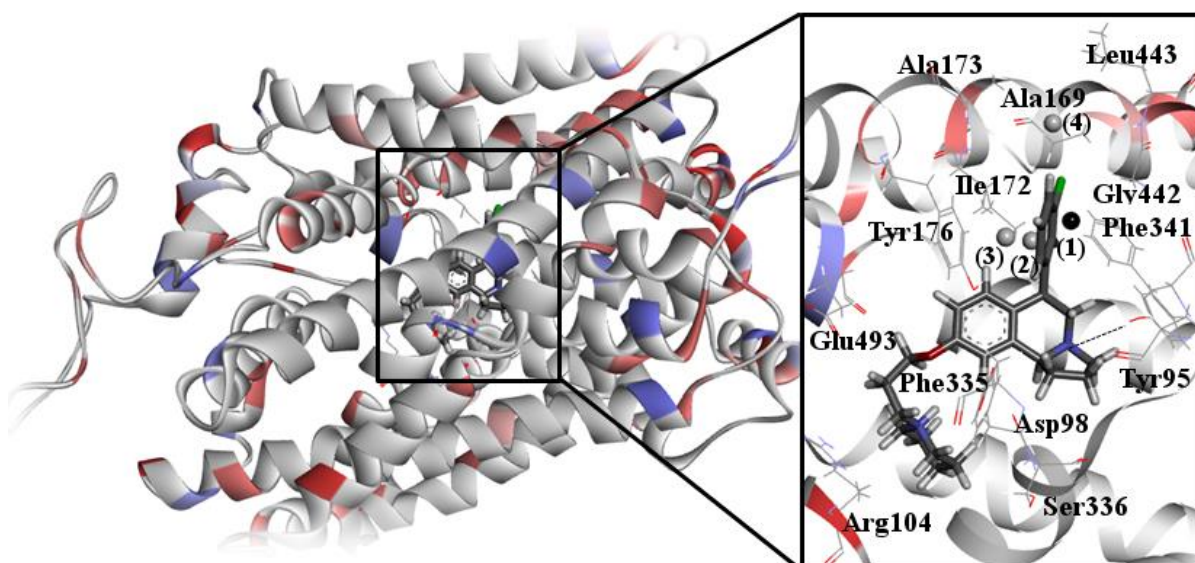
Composto	MolDock Score	Rerank Score	Energia de interação
<b>11A</b>	-143,596	-116,786	-157,615

A Figura 11 ilustra o composto 11A no sítio ativo do hSERT, os principais resíduos de aminoácido estão em destaque. Como pode ser observado, o grupo amino da porção hexahidro-pirroló-isoquinolina ocupa a região delimitada pelos resíduos de aminoácido

Tyr95, Asp98 e Ser336, similar ao grupo amino do antidepressivo paroxetina (COLEMAN; GREEN; GOUAUX, 2016). A hidroxila do resíduo de Tyr95 se encontra a uma distância de 3,52Å do nitrogênio do composto, formando uma ligação de hidrogênio. O composto 11A adota uma conformação em que o grupamento R interage com os resíduos de Phe335, Arg104, Ser336, e Glu493 na parte mais externa do sítio de ligação. Por outro lado, o substituinte aromático ocupa uma região mais hidrofóbica, formada pelas cadeias laterais dos resíduos de Tyr176, Ile172, Gly442, Phe341 e Leu443.

Na Figura 11 também estão representados os descritores do modelo da eq.10, como forma de relacionar com os resultados de QSAR. Os GCODs (6,-7,-2,any) e (5,-7,-3,np) não são ilustrados porque a conformação adotada pelo substituinte R é diferente da analisada no modelo de QSAR-4D. Na figura 11, o descritor (1,0,7,any) se encontra na parte posterior à porção hexahidro-pirrólo-isoquinolina, perto do resíduos de Ile172, Phe341 e Gly442. Os GCODs (2,-1,5,np) e (4,-2,4,any) estão posicionados entre o substituinte aromático do composto 11A e a cadeia lateral do resíduo de Tyr176, em que os anéis aromáticos estão quase paralelos entre si, possivelmente formando uma interação do tipo  $\pi$ -stacking. Além disso, esse resíduo é o que mais contribui para a estabilidade do composto no sítio ativo. Por último, o descritor (2,-3,10,np) está na parte superior do substituinte aromático do 11A, próximo aos aminoácidos Ala169, Ala173 e Leu443. Portanto, substituintes não polares na posição *para* do anel poderiam aumentar a interação do composto com esses resíduos.

Figura 11 – (Esquerda) Estrutura tridimensional do hSERT com composto 11A ancorado no sítio de ligação central. (Direita) Principais resíduos de aminoácido do sítio ativo juntamente com os descritores do modelo de QSAR-4D (1: GCOD (1,0,7,any), 2: GCOD (2,-1,5,np), 3: GCOD (4,-2,4,any) e 4: GCOD (2,-3,10,np)).



A análise de QSAR-4D e de ancoramento molecular foram feitas com o intuito de identificar características estruturais importantes para a ligação dos derivados de hexahidropirrolo-isoquinolina a ambos H<sub>3</sub>R e SERT. As informações contidas nos descritores moleculares do modelo de QSAR-4D para o H<sub>3</sub>R indicam que a posição do grupo R é fundamental para a atividade dos compostos, ao passo que substituições no anel aromático tem menor efeito. Por outro lado, o modelo de QSAR-4D para o SERT aponta que a natureza e posição das substituições no anel são determinantes. Além disso, o ancoramento do composto 11A ao sítio ativo do SERT permitiu identificar interações importantes para a ligação desses compostos e esclarecer algumas informações contidas nos descritores do QSAR, em que a posição dos GCODs correlaciona com a interação de alguns resíduos específicos. Portanto, em conjunto, essas informações podem ser utilizadas para guiar o planejamento de novas moléculas bioativas para o tratamento da depressão.



## 5 CONCLUSÃO

Apesar da quantidade de antidepressivos disponíveis no mercado, o tratamento da depressão ainda enfrenta algumas dificuldades. Salientando assim a necessidade de desenvolvimento de novos compostos bioativos. Nesse sentido, a inibição dual do transportador de serotonina e do receptor de histamina H<sub>3</sub> é uma alternativa interessante, uma vez que auxilia no alívio de alguns sintomas depressivos e na neurotransmissão serotoninérgica. Com isso, esse trabalho teve como objetivo avaliar características estruturais importantes para ligação dos derivados de hexahidro-pirroló-isoquinolina ao H<sub>3</sub>R e ao SERT por meio da análise de QSAR-4D e ancoramento molecular. Os modelos de QSAR-4D para ambos receptores apresentaram parâmetros estatísticos satisfatórios que asseguram sua robustez e preditividade (H<sub>3</sub>R: R<sup>2</sup> = 0,808, Q<sup>2</sup> = 0,736, r<sup>2</sup>m = 0,670 e R<sup>2</sup>p = 0,771; SERT: R<sup>2</sup> = 0,818, Q<sup>2</sup> = 0,710, r<sup>2</sup>m = 0,580 e R<sup>2</sup>p = 0,728), além de grupos farmacofóricos importantes para a atividade biológica. A ligação dos compostos em estudo ao SERT também foi avaliada pelo ancoramento do composto mais ativo ao sítio de ligação do transportador, onde foi possível sugerir a existência de interações intermoleculares importantes. Portanto, em conjunto, essas informações podem ser utilizadas para guiar o planejamento de novas moléculas bioativas para o tratamento da depressão.

## REFERÊNCIAS

- ALLINGER, N. L. Conformational Analysis. 130. MM2. A Hydrocarbon Force Field Utilizing V1 and V2 Torsional Terms. **Journal of American Chemical Society**, Washington, v. 99, n. 25, p. 8127–8134, 1977.
- ANDERSEN, J. et al. Location of the antidepressant binding site in the serotonin transporter: Importance of Ser-438 in recognition of citalopram and tricyclic antidepressants. **Journal of Biological Chemistry**, Bethesda, v. 284, n. 15, p. 10276–10284, 2009.
- ANDRADE, C. H. et al. 4D-QSAR: Perspectives in Drug Design. **Molecules**, Basel, v. 15, n. 5, p. 3281–3294, maio 2010.
- APODACA, R. et al. **Hexahydro-pyrrolo-isoquinoline compounds**. US 2012/0321559 A1, 20 dez. 2012.
- ARTIGAS, F. Future directions for serotonin and antidepressants. **ACS Chemical Neuroscience**, Washington, v. 4, n. 1, p. 5–8, 2013.
- AXE, F. U.; BEMBENEK, S. D.; SZALMA, S. Three-dimensional models of histamine H3 receptor antagonist complexes and their pharmacophore. **Journal of Molecular Graphics and Modelling**, New York, v. 24, n. 6, p. 456–464, 2006.
- BONDY, B. Pathophysiology of depression and mechanisms of treatment. **Dialogues in Clinical Neuroscience**, Paris, v. 4, n. 1, p. 7–20, 2002.
- CARDOZO, M. G. et al. Conformational Analyses and Molecular-Shape Comparisons of a Series of Indanone-Benzylpiperidine Inhibitors of Acetylcholinesterase. **Journal of Medicinal Chemistry**, Washington, v. 35, n. 3, p. 590–601, 1992a.
- CARDOZO, M. G. et al. QSAR Analyses of the Substituted Indanone and Benzylpiperidine Rings of a Series of Indanone-Benzylpiperidine Inhibitors of Acetylcholinesterase. **Journal of Medicinal Chemistry**, Washington, v. 35, n. 3, p. 584–589, 7 fev. 1992b.
- CHERKASOV, A. et al. QSAR modeling: Where have you been? Where are you going to? **Journal of Medicinal Chemistry**, Washington, v. 57, n. 12, p. 4977–5010, 2014.
- COLEMAN, J. A.; GREEN, E. M.; GOUAUX, E. X-ray structures and mechanism of the human serotonin transporter. **Nature**, London, v. 532, n. 7599, p. 334–339, 2016.
- CONSONNI, V.; TODESCHINI, R. Molecular Descriptors. In: PUZYN, T.; LESZCZYNSKI, J.; CRONIN, M. (Eds.). **Recent Advances in QSAR Studies**. 1 ed. [s.l.]: Springer, 2010. v. 8, p. 29–102.
- CRAMER, R. D.; PATTERSON, D. E.; BUNCE, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. **Journal of American Chemical Society**, Washington, v. 110, n. 18, p. 5959–5967, 1 ago. 1988.
- CRONIN, M. Quantitative Structure–Activity Relationships (QSARs) – Applications and Methodology. In: PUZYN, T.; LESZCZYNSKI, J.; CRONIN, M. (Eds.). **Recent Advances in QSAR Studies**. 1 ed. [s.l.]: Springer, 2010. v. 8, p. 3–11.

DA CUNHA, E. F. F. et al. LIV-3D-QSAR model for estrogen receptor ligands. **Journal of Molecular Modeling**, New York, v. 10, n. 4, p. 297–304, 31 jul. 2004.

DA CUNHA, E. F. F. et al. 4D-QSAR Models of HOE/BAY-793 Analogues as HIV-1 Protease Inhibitors. **QSAR & Combinatorial Science**, Weinheim, v. 24, n. 2, p. 240–253, mar. 2005.

DIMASI, J. A.; GRABOWSKI, H. G.; HANSEN, R. W. Cost to Develop and Win Marketing Approval for a New Drug Is \$2.6 Billion. **Tufts Center for the Study of Drug Development**, Boston, 18 nov. 2014. Disponível em: <[http://csdd.tufts.edu/news/complete\\_story/pr\\_tufts\\_csdd\\_2014\\_cost\\_study](http://csdd.tufts.edu/news/complete_story/pr_tufts_csdd_2014_cost_study)>. Acesso em: 20 nov. 2015.

DIMASI, J. A.; HANSEN, R. W.; GRABOWSKI, H. G. The price of innovation: new estimates of drug development costs. **Journal of Health Economics**, Amsterdam, v. 22, p. 151–185, 2003.

DREVETS, W. C. et al. Pet imaging of serotonin 1A receptor binding in depression. **Biological Psychiatry**, New York, v. 46, n. 10, p. 1375–1387, 15 nov. 1999.

EL-ALFY, A. T.; ABOURASHED, E. A.; MATSUMOTO, R. R. Nature Against Depression. **Current Medicinal Chemistry**, Sharjah, v. 19, n. 14, p. 2229–2241, 2012.

ESBENSHADE, T. A. et al. The histamine H3 receptor: an attractive target for the treatment of cognitive disorders. **British journal of pharmacology**, Malden, v. 154, n. 6, p. 1166–1181, 2008.

FERREIRA, M. M. C. Multivariate QSAR. **Journal of the Brazilian Chemical Society**, São Paulo, v. 13, n. 6, p. 742–753, 2002.

FERREIRA, M. M. C. Calibração - Métodos de Regressão. In: \_\_\_\_ **Quimiometria: Conceitos, Métodos e Aplicações**. Campinas: Editora da Unicamp, 2015. p. 251–404.

FERREIRA, M. M. C.; MONTANARI, C. A.; GAUDIO, A. C. Seleção de variáveis em QSAR. **Química Nova**, São Paulo, v. 25, n. 3, p. 439–448, 2002.

FREE, S. M. J.; WILSON, J. W. A Mathematical Contribution to Structure-Activity Studies. **Journal of Medicinal Chemistry**, Washington, v. 7, n. 4, p. 395–399, 1964.

GELADI, P.; KOWALSKI, B. R. Partial least-squares regression: a tutorial. **Analytica Chimica Acta**, Amsterdam, v. 185, p. 1–17, 1986.

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

GREENBERG, P. E. et al. The Economic Burden of Adults With Major Depressive Disorder in the United States (2005 and 2010). **The Journal of Clinical Psychiatry**, Memphis, v. 76, n. 2, p. 155–162, 2015.

GUEDES, I. A.; DE MAGALHÃES, C. S.; DARDENNE, L. E. Receptor–ligand molecular docking. **Biophysical Reviews**, [s.l.], v. 6, n. 1, p. 75–87, 2014.

HANSCH, C.; FUJITA, T.  $\rho$ - $\sigma$ - $\pi$  Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. **Journal of American Chemical Society**, Washington, v. 86, p. 1616–1626, 1964.

HOPFINGER, A. J. et al. Construction of 3D-QSAR Models Using the 4D-QSAR Analysis Formalism. **Journal of American Chemical Society**, Washington, v. 119, n. 43, p. 10509–10524, 1997.

JENSEN, F. Simulation Techniques. In: \_\_\_\_ **Introduction to Computational Chemistry**. 2 ed. [s.l.]: John Wiley & Sons Ltd, 2007. p. 445–486.

KAPETANOVIC, I. M. Computer-aided drug discovery and development (CADD): In silico-chemico-biological approach. **Chemico-Biological Interactions**, Clare, v. 171, n. 2, p. 165–176, 2008.

KEITH, J. M. et al. Pyrrolidino-tetrahydroisoquinolines bearing pendant heterocycles as potent dual H3 antagonist and serotonin transporter inhibitors. **Bioorganic and Medicinal Chemistry Letters**, Oxford, v. 17, n. 15, p. 4374–4377, 2007a.

KEITH, J. M. et al. Pyrrolidino-tetrahydroisoquinolines as potent dual H3 antagonist and serotonin transporter inhibitors. **Bioorganic and Medicinal Chemistry Letters**, Oxford, v. 17, n. 15, p. 2603–2607, 2007b.

KENNARD, R. W.; STONE, L. A. Computer Aided Design of Experiments. **Technometrics**, [s.l.] v. 11, n. 1, p. 137–148, 1969.

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

KISS, R.; KESERŰ, G. M. Structure-based discovery and binding site analysis of histamine receptor ligands. **Expert Opinion on Drug Discovery**, Oxon, v. 11, n. 12, p. 1165–1185, 2016.

KRYGER, G.; SILMAN, I.; SUSSMAN, J. L. Structure of acetylcholinesterase complexed with E2020 (Aricept): implications for the design of new anti-Alzheimer drugs. **Structure with Folding & Design**, London, v. 7, n. 3, p. 297–307, 15 mar. 1999.

KUBINYI, H. QSAR and 3D QSAR in drug design. Part 1: methodology. **Drug Discovery Today**, Oxford, v. 2, n. 11, p. 457–467, nov. 1997.

KUBINYI, H. From Narcosis to Hyperspace: The History of QSAR. **Quantitative Structure-Activity Relationships**, Weinheim, v. 21, p. 348–356, 2002.

LEBOIS, E. P.; JONES, C. K.; LINDSLEY, C. W. The Evolution of Histamine H3 Antagonists/Inverse Agonists. **Current Topics in Medicinal Chemistry**, Sharjah, v. 11, n. 6, p. 648–660, 2011.

LEVOIN, N. et al. Determination of the binding mode and interacting amino-acids for dibasic H3 receptor antagonists. **Bioorganic and Medicinal Chemistry**, Oxford, v. 21, n. 15, p. 4526–4529, 2013.

LILL, M. A. Multi-dimensional QSAR in drug discovery. **Drug Discovery Today**, Oxford, v. 12, n. 23/24, p. 1013–1017, dez. 2007.

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

MANN, J. J. et al. Attempted suicide characteristics and cerebrospinal fluid amine metabolites in depressed inpatients. **Neuropsychopharmacology**, New York, v. 15, n. 6, p. 576–586, 1996.

MOLSIM. **Molsim User's Guide** 1780 Wilson Dr., Lake Forest, IL 60045 Doherty D.C. & The Chem21 Group Inc., , 1997.

MORENO, R. A.; MORENO, D. H.; SOARES, M. B. DE M. Psicofarmacologia de antidepressivos. **Revista Brasileira de Psiquiatria**, São Paulo, v. 21, p. 24–40, 1999.

MORGAN, S. et al. The cost of drug development: A systematic review. **Health policy**, Clare, v. 100, n. 1, p. 4–17, 2011.

MUNOS, B. Lessons from 60 years of pharmaceutical innovation. **Nature Reviews Drug Discovery**, London, v. 8, n. 12, p. 959–968, dez. 2009.

NADAL-VICENS, M.; CHYUNG, J. H.; TURNER, T. J. Farmacologia da Neurotransmissão Serotoninérgica e Adrenérgica Central. In: GOLAN, D.E. et al. **Princípios de Farmacologia - A Base Fisiopatológica da Farmacoterapia**. 2 ed. [s.l.]: Editora Guanabara Koogan, 2009. p. 186–202.

NESTLER, E. J. et al. Neurobiology of Depression. **Neuron**, Cambridge, v. 34, n. 1, p. 13–25, 2002.

NICHOLS, D. E.; NICHOLS, C. D. Serotonin Receptors. **Chemical Reviews**, Washington, v. 108, n. 5, p. 1614–1641, 2008.

NIKOLIC, K. et al. Procognitive properties of drugs with single and multitargeting H3 receptor antagonist activities. **CNS Neuroscience & Therapeutics**, Hoboken, v. 20, n. 7, p. 613–623, 2014.

NUNES, C. A. et al. Chemoface: a Novel Free User-Friendly Interface for Chemometrics. **Journal of the Brazilian Chemical Society**, São Paulo, v. 23, n. 11, p. 2003–2010, 2012.

PASQUALOTO, K. F. M. et al. Rational Design of New Antituberculosis Agents: Receptor-Independent Four-Dimensional Quantitative Structure-Activity Relationship Analysis of a Set of Isoniazid Derivatives. **Journal of Medicinal Chemistry**, Washington, v. 47, n. 15, p. 3755–3764, 15 jul. 2004.

PIRHADI, S.; SHIRI, F.; GHASEMI, J. B. Multivariate statistical analysis methods in QSAR. **RSC Advances**, Cambridge, v. 5, n. 127, p. 104635–104665, 2015.

PITA, S. S. da R. **Modelos de QSAR-4D Dependente do Receptor de Inibidores Peptídicos da Tripanotona Redutase**. 2006. 110 p. Dissertação (Mestrado em Química Orgânica)-Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

POLANSKI, J. Receptor Dependent Multidimensional QSAR for Modeling Drug - Receptor Interactions. **Current Medicinal Chemistry**, Sharjah, v. 16, n. 25, p. 3243–3257, jan. 2009.

PYTKA, K. et al. The role of serotonergic, adrenergic and dopaminergic receptors in antidepressant-like effect. **Pharmacological Reports**, Kraków, v. 68, n. 2, p. 263–274, 2016.

ROCHA, G. B. et al. RM1: A Reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br and I. **Journal of Computational Chemistry**, Hoboken, v. 27, n. 10, p. 1101–1111, 2005.

ROGERS, D.; HOPFINGER, A. J. Application of Genetic Function Approximation to Quantitative Structure-Activity Relationships and Quantitative Structure-Property Relationships. **Journal of Chemical Information and Computer Sciences**, Washington, v. 34, n. 4, p. 854–866, 1 jul. 1994.

ROMEIRO, N. C. et al. Construction of 4D-QSAR models for use in the design of novel p38-MAPK inhibitors. **Journal of Computer-Aided Molecular Design**, Dordrecht, v. 19, n. 6, p. 385–400, jun. 2005.

ROY, P. P. et al. On two novel parameters for validation of predictive QSAR models. **Molecules**, Basel, v. 14, n. 5, p. 1660–1701, 2009.

RÜCKER, C.; RÜCKER, G.; MERINGER, M.  $\gamma$ -Randomization and Its Variants in QSPR / QSAR. **Journal of Chemical Information and Modeling**, Washington, v. 47, n. 6, p. 2345–2357, 2007.

SAVEANU, R. V.; NEMEROFF, C. B. Etiology of Depression: Genetic and Environmental Factors. **Psychiatric Clinics of North America**, Philadelphia, v. 35, n. 1, p. 51–71, 2012.

SCHILDKRAUT, J. J. The Catecholamine Hypothesis of Affective Disorders: A Review of Supporting Evidence. **American Journal of Psychiatry**, Washington, v. 122, n. 5, p. 509–522, 1 nov. 1965.

SCHLEGEL, B. et al. Generation of a homology model of the human histamine H3 receptor for ligand docking and pharmacophore-based screening. **Journal of Computer-Aided Molecular Design**, Dordrecht, v. 21, n. 8, p. 437–453, 2007.

SHARMA, H.; SANTRA, S.; DUTTA, A. Triple reuptake inhibitors as potential next-generation antidepressants: a new hope? **Future Medicinal Chemistry**, London, v. 7, n. 17, p. 2384–2405, 2015.

SMITH, C. M.; MARKS, A. D.; LIEBERMAN, M. A. Metabolism of the Nervous System. In: \_\_\_\_ **Marks' Basic Medical Biochemistry: A Clinical Approach**. 2 ed. [s.l.] Lippincott Williams & Wilkins, 2003. p. 881–904.

SO, S. S. Quantitative Structure-Activity Relationships. In: CLARK, D. E. (Ed.). **Evolutionary Algorithms in Molecular Design**. Weinheim: Wiley-VCH, 2000. p. 71–97.

SOBOCKI, P. et al. Cost of depression in Europe. **Journal of Mental Health Policy and Economics**, Milano, v. 9, n. 2, p. 87–98, jun. 2006.

STOCKING, E. M. et al. Exploration of structure-activity relationships for dual serotonin transporter reuptake inhibitors-histamine H3 receptor antagonists. **Current Topics in**

**Medicinal Chemistry**, Sharjah, v. 10, n. 5, p. 596–616, 2010.

TAVARES, L. C. QSAR: A Abordagem de Hansch. **Quimica Nova**, São Paulo, v. 27, n. 4, p. 631–639, 2004.

THOMSEN, R.; CHRISTENSEN, M. H. MolDock: a new technique for high-accuracy molecular docking. **Journal of Medicinal Chemistry**, Washington, v. 49, n. 11, p. 3315–21, 2006.

TOKITA, S.; TAKAHASHI, K.; KOTANI, H. Recent Advances in Molecular Pharmacology of the Histamine Systems: Physiology and Pharmacology of Histamine H3 Receptor: Roles in feeding regulation and therapeutic potential for metabolic disorder. **Journal of Pharmacological Sciences**, Kyoto, v. 101, p. 12–18, 2006.

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

VEDANI, A.; DOBLER, M. 5D-QSAR: The Key for Simulating Induced Fit? **Journal of Medicinal Chemistry**, Washington, v. 45, n. 11, p. 2139–2149, 23 maio 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, n. 11, p. 3700–3703, 2 jun. 2005.

VERMA, J.; KHEDKAR, V. M.; COUTINHO, E. C. 3D-QSAR in Drug Design - A Review. **Current Topics in Medicinal Chemistry**, Sharjah, v. 10, n. 1, p. 95–115, jan. 2010.

WALKER, J. D. et al. Guidelines for Developing and Using Quantitative Structure – Activity Relationships. **Environmental Toxicology and Chemistry**, Malden, v. 22, n. 8, p. 1653–1665, 2003.

WEINER, S. J. et al. An all atom force field for simulations of proteins and nucleic acids. **Journal of Computational Chemistry**, Hoboken, v. 7, n. 2, p. 230–252, abr. 1986.

WHO. **Depression**. Disponível em: <<http://www.who.int/topics/depression/en/>>. Acesso em: 17 mar. 2017.

**APÊNDICE A – Modelos de QSAR-4D para o receptor H<sub>3</sub>**

Tabela 9 – Tabela completa com parâmetros estatísticos da calibração, validação interna e externa dos melhores modelos provenientes dos dez alinhamentos testados para H<sub>3</sub>R.

<b>Alig</b>	<b>Des</b>	<b>R<sup>2</sup></b>	<b>R<sup>2</sup><sub>ajus</sub></b>	<b>Q<sup>2</sup></b>	<b>Q<sup>2</sup><sub>ajus</sub></b>	<b>R<sup>2</sup><sub>pred</sub></b>	<b>r<sup>2</sup><sub>m</sub></b>	<b>R<sup>2</sup><sub>rand</sub></b>	<b>R<sub>p</sub></b>
<b>1.1</b>	7	0,791	0,775	0,751	0,732	0,777	0,723	0,072	0,754
<b>1.2</b>	7	0,779	0,763	0,740	0,721	0,825	0,734	0,124	0,714
<b>1.3</b>	7	0,798	0,783	0,758	0,740	0,754	0,739	0,066	0,764
<b>1med</b>	7	0,789	0,774	0,750	0,731	0,785	0,732	0,088	0,744
<b>2.1</b>	7	0,812	0,798	0,771	0,754	0,696	0,575	0,091	0,765
<b>2.2</b>	7	0,790	0,774	0,737	0,718	0,812	0,636	0,083	0,747
<b>2.3</b>	7	0,800	0,785	0,758	0,740	0,756	0,427	0,081	0,758
<b>2med</b>	7	0,800	0,786	0,755	0,737	0,755	0,546	0,085	0,757
<b>3.1</b>	7	0,803	0,789	0,757	0,739	0,660	0,567	0,071	0,767
<b>3.2</b>	7	0,774	0,757	0,699	0,677	0,848	0,762	0,095	0,725
<b>3.3</b>	7	0,782	0,766	0,709	0,688	0,796	0,701	0,074	0,745
<b>3med</b>	7	0,787	0,771	0,722	0,701	0,768	0,676	0,080	0,746
<b>4.1</b>	7	0,810	0,796	0,759	0,741	0,776	0,724	0,077	0,771
<b>4.2</b>	7	0,808	0,794	0,715	0,693	0,777	0,694	0,078	0,768
<b>4.3</b>	7	0,805	0,790	0,733	0,713	0,801	0,592	0,063	0,773
<b>4med</b>	7	0,808	0,794	0,736	0,716	0,785	0,670	0,073	0,771
<b>5.1</b>	6	0,770	0,756	0,731	0,714	0,632	0,458	0,081	0,728
<b>5.2</b>	6	0,745	0,729	0,684	0,665	0,798	0,654	0,072	0,708
<b>5.3</b>	6	0,745	0,730	0,673	0,653	0,749	0,623	0,077	0,706
<b>5med</b>	6	0,753	0,738	0,696	0,677	0,726	0,578	0,076	0,714
<b>6.1</b>	7	0,789	0,774	0,746	0,727	0,726	0,626	0,071	0,753
<b>6.2</b>	7	0,782	0,766	0,746	0,727	0,746	0,482	0,067	0,748
<b>6.3</b>	7	0,779	0,763	0,738	0,719	0,758	0,656	0,093	0,731
<b>6med</b>	7	0,784	0,768	0,743	0,724	0,743	0,588	0,077	0,744
<b>7.1</b>	6	0,788	0,775	0,746	0,731	0,749	0,738	0,086	0,743
<b>7.2</b>	6	0,770	0,756	0,725	0,708	0,825	0,784	0,064	0,738
<b>7.3</b>	6	0,770	0,756	0,717	0,700	0,764	0,579	0,069	0,734
<b>7med</b>	6	0,776	0,762	0,729	0,713	0,779	0,700	0,073	0,738
<b>8.1</b>	6	0,769	0,755	0,661	0,640	0,716	0,576	0,086	0,725
<b>8.2</b>	6	0,759	0,744	0,673	0,653	0,765	0,615	0,048	0,734
<b>8.3</b>	6	0,764	0,749	0,717	0,700	0,710	0,435	0,056	0,735
<b>8med</b>	6	0,764	0,749	0,684	0,664	0,730	0,542	0,063	0,731
<b>9.1</b>	7	0,780	0,764	0,741	0,722	0,702	0,644	0,077	0,741
<b>9.2</b>	7	0,751	0,732	0,704	0,682	0,874	0,822	0,084	0,708
<b>9.3</b>	7	0,763	0,745	0,716	0,695	0,740	0,558	0,072	0,726
<b>9med</b>	7	0,765	0,747	0,720	0,700	0,772	0,675	0,077	0,725
<b>10.1</b>	6	0,758	0,743	0,683	0,664	0,746	0,667	0,060	0,727
<b>10.2</b>	6	0,736	0,720	0,656	0,635	0,872	0,803	0,061	0,705
<b>10.3</b>	6	0,731	0,715	0,627	0,605	0,826	0,791	0,083	0,689
<b>10med</b>	6	0,742	0,726	0,656	0,634	0,814	0,754	0,068	0,707



**APÊNDICE B – Frequência de ocupação dos GCODs do alinhamento 4**

Tabela 10 – Frequência de ocupação dos GCODs da eq. 9 para todos os compostos do conjunto de dados. (Continua)

<b>Comp.</b>	<b>(2,-7,-2, p-)</b>	<b>(-6,-8,-2, any)</b>	<b>(4,-9,-1, any)</b>	<b>(1,4,1, np)</b>	<b>(-6,-8,2, np)</b>	<b>(0,1,-1, any)</b>	<b>(3,-3,-5, np)</b>
<b>1A</b>	0,000	0,012	0,000	0,000	0,000	0,644	0,000
<b>1B</b>	0,000	0,002	0,000	0,042	0,064	0,638	0,000
<b>1C</b>	0,000	0,000	0,000	0,068	0,000	0,996	0,022
<b>2A</b>	0,010	0,018	0,000	0,000	0,000	0,630	0,000
<b>2B</b>	0,004	0,004	0,000	0,002	0,000	0,684	0,000
<b>2C</b>	0,000	0,000	0,000	0,152	0,000	0,986	0,008
<b>3</b>	0,002	0,014	0,000	0,000	0,000	0,578	0,000
<b>4A</b>	0,000	0,028	0,000	0,000	0,000	0,608	0,000
<b>4B</b>	0,006	0,018	0,000	0,002	0,000	0,700	0,000
<b>4C</b>	0,000	0,000	0,000	0,000	0,000	0,980	0,014
<b>4D</b>	0,000	0,000	0,000	0,380	0,000	0,976	0,014
<b>5A</b>	0,000	0,014	0,000	0,000	0,000	0,596	0,000
<b>5B</b>	0,000	0,002	0,000	0,002	0,026	0,674	0,000
<b>5C</b>	0,000	0,000	0,000	0,122	0,000	0,998	0,026
<b>6A</b>	0,006	0,026	0,000	0,000	0,000	0,596	0,000
<b>6B</b>	0,000	0,016	0,000	0,100	0,000	0,718	0,000
<b>6C</b>	0,000	0,000	0,000	0,000	0,000	0,986	0,018
<b>7A</b>	0,006	0,016	0,000	0,000	0,000	0,604	0,000
<b>7B</b>	0,012	0,022	0,002	0,026	0,000	0,704	0,000
<b>7C</b>	0,000	0,000	0,000	0,094	0,000	0,996	0,018
<b>8D</b>	0,004	0,026	0,000	0,036	0,012	0,682	0,000
<b>8E</b>	0,000	0,000	0,000	0,000	0,000	0,996	0,018
<b>8F</b>	0,000	0,000	0,000	0,172	0,000	0,992	0,030
<b>9A</b>	0,000	0,014	0,000	0,000	0,000	0,626	0,000
<b>9B</b>	0,000	0,000	0,000	0,136	0,000	0,978	0,038
<b>10A</b>	0,000	0,022	0,000	0,000	0,000	0,576	0,000
<b>10B</b>	0,000	0,004	0,000	0,072	0,014	0,692	0,000
<b>10C</b>	0,000	0,000	0,000	0,076	0,000	0,998	0,018
<b>11A</b>	0,018	0,046	0,000	0,000	0,000	0,614	0,000
<b>11B</b>	0,008	0,034	0,000	0,064	0,002	0,702	0,000
<b>11C</b>	0,000	0,000	0,000	0,018	0,000	0,994	0,010
<b>12A</b>	0,000	0,010	0,000	0,000	0,000	0,640	0,000
<b>12B</b>	0,004	0,030	0,000	0,012	0,002	0,666	0,000
<b>12C</b>	0,000	0,000	0,000	0,000	0,000	0,976	0,014
<b>12D</b>	0,000	0,000	0,000	0,028	0,000	1,000	0,000

Tabela 10 – Frequência de ocupação dos GCODs da eq. 9 para todos os compostos do conjunto de dados. (Continua)

<b>Comp.</b>	<b>(2,-7,-2, p-)</b>	<b>(-6,-8,-2, any)</b>	<b>(4,-9,-1, any)</b>	<b>(1,4,1, np)</b>	<b>(-6,-8,2, np)</b>	<b>(0,1,-1, any)</b>	<b>(3,-3,-5, np)</b>
<b>13A</b>	0,006	0,026	0,000	0,000	0,000	0,582	0,000
<b>13B</b>	0,000	0,000	0,000	0,000	0,000	0,994	0,012
<b>14A</b>	0,004	0,012	0,000	0,000	0,000	0,578	0,000
<b>14B</b>	0,000	0,002	0,000	0,006	0,042	0,674	0,000
<b>14C</b>	0,000	0,000	0,000	0,100	0,000	0,998	0,024
<b>15</b>	0,000	0,020	0,000	0,000	0,000	0,612	0,000
<b>16</b>	0,002	0,008	0,000	0,000	0,000	0,626	0,000
<b>17</b>	0,014	0,002	0,000	0,000	0,000	0,638	0,000
<b>18</b>	0,002	0,012	0,000	0,000	0,000	0,634	0,000
<b>19</b>	0,012	0,014	0,000	0,000	0,000	0,590	0,000
<b>20</b>	0,002	0,016	0,000	0,000	0,000	0,612	0,000
<b>21</b>	0,000	0,022	0,000	0,000	0,000	0,562	0,000
<b>22</b>	0,000	0,014	0,000	0,000	0,000	0,594	0,000
<b>23</b>	0,010	0,010	0,004	0,000	0,000	0,594	0,000
<b>24A</b>	0,004	0,016	0,000	0,000	0,000	0,628	0,000
<b>24B</b>	0,000	0,006	0,012	0,006	0,000	0,660	0,000
<b>25</b>	0,002	0,018	0,000	0,000	0,000	0,564	0,000
<b>26</b>	0,000	0,004	0,008	0,000	0,000	0,582	0,000
<b>27A</b>	0,000	0,008	0,000	0,000	0,000	0,558	0,000
<b>27B</b>	0,000	0,022	0,000	0,054	0,024	0,666	0,000
<b>28</b>	0,000	0,000	0,040	0,000	0,000	0,588	0,000
<b>29</b>	0,000	0,000	0,000	0,000	0,012	0,592	0,000
<b>30</b>	0,008	0,024	0,000	0,000	0,000	0,624	0,000
<b>31</b>	0,000	0,000	0,000	0,000	0,000	0,604	0,000
<b>32</b>	0,008	0,028	0,002	0,000	0,000	0,610	0,000
<b>34</b>	0,000	0,018	0,000	0,000	0,000	0,622	0,000
<b>35A</b>	0,000	0,004	0,018	0,000	0,000	0,578	0,000
<b>35B</b>	0,006	0,028	0,000	0,010	0,000	0,638	0,000
<b>37</b>	0,000	0,010	0,000	0,000	0,000	0,618	0,000
<b>38</b>	0,000	0,000	0,000	0,000	0,000	0,622	0,000
<b>39A</b>	0,010	0,022	0,000	0,000	0,000	0,608	0,000
<b>39B</b>	0,000	0,004	0,000	0,024	0,038	0,698	0,000
<b>40</b>	0,004	0,002	0,006	0,000	0,000	0,584	0,000
<b>41A</b>	0,000	0,010	0,000	0,000	0,000	0,614	0,000
<b>41B</b>	0,002	0,036	0,000	0,018	0,000	0,698	0,000
<b>42</b>	0,000	0,000	0,000	0,022	0,000	0,810	0,000
<b>43</b>	0,000	0,000	0,000	0,000	0,000	0,820	0,000
<b>44</b>	0,000	0,000	0,000	0,064	0,000	0,786	0,000

Tabela 10 – Frequência de ocupação dos GCODs da eq. 9 para todos os compostos do conjunto de dados. (Conclusão)

<b>Comp.</b>	<b>(2,-7,-2, p-)</b>	<b>(-6,-8,-2, any)</b>	<b>(4,-9,-1, any)</b>	<b>(1,4,1, np)</b>	<b>(-6,-8,2, np)</b>	<b>(0,1,-1, any)</b>	<b>(3,-3,-5, np)</b>
<b>45</b>	0,000	0,000	0,000	0,000	0,000	0,862	0,000
<b>48A</b>	0,002	0,000	0,000	0,000	0,000	0,624	0,000
<b>48B</b>	0,000	0,006	0,000	0,056	0,014	0,638	0,000
<b>49</b>	0,004	0,000	0,002	0,000	0,000	0,576	0,000
<b>50A</b>	0,000	0,000	0,000	0,000	0,000	0,608	0,000
<b>50B</b>	0,016	0,006	0,000	0,006	0,000	0,716	0,000
<b>51A</b>	0,000	0,000	0,000	0,000	0,000	0,584	0,000
<b>51B</b>	0,008	0,002	0,000	0,000	0,000	0,656	0,000
<b>52</b>	0,000	0,000	0,000	0,000	0,002	0,586	0,000
<b>53</b>	0,000	0,000	0,000	0,000	0,016	0,530	0,000
<b>54</b>	0,002	0,008	0,000	0,000	0,010	0,584	0,000
<b>55</b>	0,004	0,014	0,000	0,000	0,004	0,598	0,000
<b>56</b>	0,006	0,002	0,000	0,000	0,000	0,586	0,000
<b>57</b>	0,004	0,000	0,000	0,000	0,000	0,598	0,000
<b>58</b>	0,000	0,002	0,000	0,000	0,000	0,626	0,000
<b>59</b>	0,004	0,000	0,000	0,000	0,000	0,600	0,000
<b>60</b>	0,000	0,000	0,000	0,000	0,000	0,566	0,000
<b>61</b>	0,016	0,002	0,000	0,000	0,000	0,612	0,000
<b>62</b>	0,000	0,000	0,000	0,000	0,006	0,602	0,000
<b>63</b>	0,006	0,000	0,002	0,000	0,000	0,616	0,000
<b>64</b>	0,000	0,000	0,000	0,000	0,000	0,630	0,000
<b>65</b>	0,000	0,010	0,000	0,000	0,000	0,618	0,000
<b>66A</b>	0,006	0,026	0,000	0,000	0,000	0,616	0,000
<b>66B</b>	0,006	0,026	0,002	0,014	0,000	0,694	0,000
<b>67</b>	0,000	0,008	0,000	0,000	0,000	0,618	0,000
<b>89</b>	0,004	0,004	0,000	0,000	0,000	0,690	0,000
<b>90</b>	0,002	0,000	0,000	0,000	0,000	0,640	0,000
<b>91</b>	0,016	0,000	0,002	0,000	0,000	0,576	0,000
<b>92</b>	0,002	0,000	0,002	0,000	0,000	0,652	0,000
<b>94</b>	0,000	0,000	0,000	0,000	0,000	0,668	0,000
<b>96A</b>	0,000	0,004	0,000	0,000	0,002	0,580	0,000
<b>97A</b>	0,000	0,000	0,000	0,000	0,000	0,668	0,000
<b>97B</b>	0,000	0,000	0,000	0,054	0,000	0,796	0,000
<b>98A</b>	0,000	0,000	0,006	0,000	0,000	0,596	0,000
<b>98B</b>	0,000	0,000	0,000	0,012	0,000	0,624	0,000

### APÊNDICE C – Modelos de QSAR-4D para o SERT

Tabela 11 - Tabela completa com parâmetros estatísticos da calibração, validação interna e externa dos melhores modelos provenientes dos dez alinhamentos testados para o SERT.

<b>Alig</b>	<b>Des</b>	<b>R<sup>2</sup></b>	<b>R<sup>2</sup><sub>ajus</sub></b>	<b>Q<sup>2</sup></b>	<b>Q<sup>2</sup><sub>ajus</sub></b>	<b>R<sup>2</sup><sub>pred</sub></b>	<b>r<sup>2</sup><sub>m</sub></b>	<b>R<sup>2</sup><sub>rand</sub></b>	<b>R<sub>p</sub></b>
<b>1.1</b>	4	0,725	0,698	0,659	0,625	0,083	0,000	0,081	0,684
<b>1.2</b>	4	0,654	0,619	0,524	0,476	0,494	0,000	0,109	0,597
<b>1.3</b>	4	0,659	0,625	0,516	0,468	0,445	0,325	0,132	0,590
<b>1med</b>	4	0,680	0,648	0,566	0,523	0,340	0,108	0,119	0,616
<b>2.1</b>	5	0,801	0,774	0,717	0,678	0,153	0,000	0,137	0,729
<b>2.2</b>	5	0,760	0,727	0,671	0,626	0,261	0,000	0,175	0,667
<b>2.3</b>	5	0,635	0,585	0,472	0,399	0,586	0,354	0,166	0,546
<b>2med</b>	5	0,732	0,695	0,620	0,568	0,333	0,118	0,159	0,648
<b>3.1</b>	6	0,836	0,807	0,755	0,711	0,286	0,246	0,181	0,739
<b>3.2</b>	5	0,743	0,698	0,663	0,603	0,245	0,000	0,176	0,650
<b>3.3</b>	6	0,684	0,628	0,403	0,297	0,521	0,451	0,142	0,609
<b>3med</b>	6	0,754	0,711	0,607	0,537	0,351	0,232	0,166	0,666
<b>4.1</b>	6	0,766	0,725	0,635	0,570	0,624	0,548	0,214	0,650
<b>4.2</b>	6	0,764	0,722	0,660	0,600	0,782	0,101	0,196	0,659
<b>4.3</b>	6	0,751	0,707	0,525	0,440	0,899	0,785	0,178	0,656
<b>4med</b>	6	0,761	0,718	0,607	0,537	0,769	0,478	0,196	0,655
<b>5.1</b>	6	0,799	0,763	0,708	0,656	0,096	0,095	0,195	0,695
<b>5.2</b>	6	0,531	0,448	0,137	-0,02	0,390	0,000	0,160	0,444
<b>5.3</b>	6	0,509	0,421	0,234	0,097	0,156	0,000	0,163	0,419
<b>5med</b>	6	0,613	0,544	0,360	0,245	0,214	0,031	0,173	0,519
<b>6.1</b>	4	0,710	0,681	0,586	0,545	0,163	0,000	0,106	0,654
<b>6.2</b>	4	0,699	0,669	0,556	0,512	0,221	0,000	0,105	0,645
<b>6.3</b>	4	0,517	0,469	0,264	0,191	0,675	0,432	0,135	0,444
<b>6med</b>	4	0,642	0,606	0,469	0,416	0,353	0,144	0,115	0,581
<b>7.1</b>	6	0,833	0,803	0,759	0,716	0,852	0,584	0,153	0,752
<b>7.2</b>	6	0,802	0,767	0,674	0,616	0,759	0,592	0,188	0,702
<b>7.3</b>	6	0,819	0,786	0,699	0,645	0,656	0,565	0,165	0,731
<b>7med</b>	6	0,818	0,786	0,710	0,659	0,756	0,580	0,169	0,728
<b>8.1</b>	6	0,823	0,791	0,759	0,716	0,240	0,153	0,155	0,741
<b>8.2</b>	6	0,615	0,546	0,412	0,307	0,729	0,132	0,141	0,540
<b>8.3</b>	6	0,686	0,630	0,536	0,454	0,229	0,000	0,173	0,593
<b>8med</b>	6	0,708	0,656	0,569	0,492	0,399	0,095	0,156	0,625
<b>9.1</b>	6	0,895	0,877	0,764	0,722	0,064	0,000	0,186	0,797
<b>9.2</b>	6	0,649	0,587	0,379	0,268	0,247	0,000	0,174	0,556
<b>9.3</b>	6	0,642	0,579	0,365	0,252	0,043	0,000	0,193	0,537
<b>9med</b>	6	0,729	0,681	0,503	0,414	0,118	0,000	0,184	0,630
<b>10.1</b>	6	0,806	0,771	0,733	0,686	0,289	0,000	0,221	0,687
<b>10.2</b>	6	0,761	0,718	0,579	0,503	0,552	0,000	0,218	0,642
<b>10.3</b>	6	0,705	0,652	0,579	0,504	0,583	0,316	0,206	0,592
<b>10med</b>	6	0,757	0,714	0,630	0,564	0,475	0,105	0,215	0,641

**APÊNDICE D - Frequência de ocupação dos GCODs do alinhamento 7**

Tabela 12 - Frequência de ocupação dos GCODs da eq. 10 para todos os compostos do conjunto de dados. (Continua)

	<b>(6,-7,-2, any)</b>	<b>(5,-7,-3, np)</b>	<b>(1,0,7, any)</b>	<b>(2,-1,5, np)</b>	<b>(2,-3,10, any)</b>	<b>(4,-2,4, any)</b>
<b>1A</b>	0,012	0,044	0,070	0,160	0,000	0,000
<b>2A</b>	0,024	0,056	0,106	0,210	0,000	0,000
<b>3</b>	0,022	0,054	0,082	0,276	0,000	0,000
<b>4A</b>	0,066	0,126	0,110	0,112	0,000	0,056
<b>5A</b>	0,036	0,084	0,094	0,332	0,000	0,000
<b>6A</b>	0,046	0,066	0,100	0,230	0,000	0,004
<b>7A</b>	0,020	0,040	0,088	0,152	0,000	0,000
<b>9A</b>	0,014	0,056	0,114	0,232	0,000	0,000
<b>10A</b>	0,068	0,080	0,106	0,150	0,000	0,136
<b>11A</b>	0,080	0,104	0,074	0,220	0,000	0,000
<b>12A</b>	0,020	0,040	0,082	0,226	0,000	0,000
<b>13A</b>	0,052	0,084	0,158	0,000	0,000	0,058
<b>14A</b>	0,018	0,036	0,052	0,186	0,000	0,000
<b>15</b>	0,046	0,080	0,072	0,220	0,000	0,042
<b>16</b>	0,006	0,050	0,154	0,000	0,000	0,000
<b>17</b>	0,012	0,026	0,120	0,170	0,000	0,000
<b>18</b>	0,008	0,028	0,190	0,196	0,022	0,248
<b>19</b>	0,050	0,076	0,138	0,000	0,042	0,110
<b>20</b>	0,024	0,070	0,096	0,000	0,000	0,132
<b>21</b>	0,034	0,066	0,098	0,230	0,000	0,178
<b>23</b>	0,042	0,070	0,116	0,432	0,000	0,000
<b>24A</b>	0,032	0,052	0,058	0,220	0,000	0,000
<b>25</b>	0,062	0,124	0,070	0,010	0,000	0,000
<b>26</b>	0,004	0,024	0,108	0,162	0,000	0,000
<b>27A</b>	0,010	0,060	0,070	0,262	0,000	0,000
<b>28</b>	0,000	0,000	0,142	0,224	0,042	0,000
<b>29</b>	0,004	0,002	0,134	0,208	0,056	0,000
<b>30</b>	0,046	0,064	0,116	0,220	0,044	0,000
<b>31</b>	0,012	0,044	0,130	0,122	0,000	0,212
<b>32</b>	0,046	0,088	0,132	0,140	0,000	0,000
<b>34</b>	0,028	0,066	0,040	0,264	0,000	0,002
<b>35A</b>	0,002	0,008	0,090	0,214	0,000	0,000
<b>37</b>	0,024	0,034	0,058	0,180	0,002	0,000
<b>38</b>	0,000	0,042	0,096	0,276	0,064	0,002
<b>39A</b>	0,032	0,064	0,084	0,256	0,000	0,000
<b>40</b>	0,012	0,024	0,136	0,204	0,000	0,000

Tabela 12 - Frequência de ocupação dos GCODs da eq. 10 para todos os compostos do conjunto de dados. (Conclusão)

	(6,-7,-2, any)	(5,-7,-3, np)	(1,0,7, any)	(2,-1,5, np)	(2,-3,10, any)	(4,-2,4, any)
<b>41A</b>	0,020	0,056	0,096	0,148	0,000	0,002
<b>66A</b>	0,036	0,078	0,080	0,214	0,000	0,000
<b>67</b>	0,014	0,072	0,122	0,150	0,000	0,448
<b>92</b>	0,000	0,006	0,000	0,000	0,000	0,000
<b>94</b>	0,002	0,022	0,146	0,002	0,000	0,000
<b>96A</b>	0,006	0,000	0,122	0,158	0,010	0,000
<b>98A</b>	0,000	0,000	0,130	0,122	0,164	0,000