

VICTOR DE SOUSA BATISTA

**Estudos computacionais de ligantes dos receptores nicotínicos de acetilcolina
do subtipo $\alpha 4\beta 2$**

Dissertação apresentada ao Instituto de
Química, Universidade Estadual Paulista,
como parte dos requisitos para obtenção do
título de Mestre em Química

Orientador: Prof. Dr. Nailton Monteiro do Nascimento Júnior

ARARAQUARA

2019

FICHA CATALOGRÁFICA

B333e Batista, Victor de Sousa
Estudos computacionais de ligantes dos receptores nicotínicos de acetilcolina do subtipo $\alpha 4\beta 2$ / Victor de Sousa Batista. – Araraquara: [s.n.], 2019
143 p.: il.

Dissertação (mestrado) – Universidade Estadual Paulista, Instituto de Química
Orientador: Nailton Monteiro do Nascimento-Júnior

1. Química medicinal. 2. Receptores nicotínicos.
3. Canais iônicos. 4. Estudos de validação. 5. Algoritmos genéticos. I. Título

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: "Estudos computacionais de ligantes dos receptores nicotínicos de acetilcolina do subtipo 4 β 2"

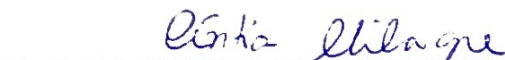
AUTOR: VICTOR DE SOUSA BATISTA

ORIENTADOR: NAILTON MONTEIRO DO NASCIMENTO JUNIOR

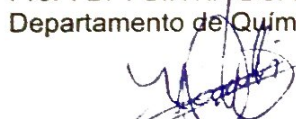
Aprovado como parte das exigências para obtenção do Título de Mestre em QUÍMICA, pela Comissão Examinadora:



Prof. Dr. NAILTON MONTEIRO DO NASCIMENTO JUNIOR
Departamento de Química Orgânica / Instituto de Química - UNESP - Araraquara



Profª. Drª. CINTIA DUARTE DE FREITAS MILAGRE
Departamento de Química Orgânica / Instituto de Química - UNESP - Araraquara



Dr. NATANAEL DANTE SEGRETTI
Inovação Radical / Aché Laboratórios Farmacêuticos - Guarulhos

Araraquara, 26 de fevereiro de 2019

DADOS CURRÍCULARES

Nome: Victor de Sousa Batista

Nome em citações bibliográficas: Batista, V. S.

Nascimento: 27/02/1994

Naturalidade: São Paulo/SP – Brasil

E-mail para contato: sbatista.victor@gmail.com

Endereço profissional: Universidade Estadual Paulista “Júlio de Mesquita Filho”. Instituto de Química de Araraquara. Departamento de Química Orgânica (LaQMedSOMM). Rua Professor Francisco Degni, 55. CEP: 14800-060. Araraquara/SP - Brasil.

FORMAÇÃO ACADÊMICA/TITULAÇÃO

Mestrado em Química: Universidade Estadual Paulista “Júlio de Mesquita Filho”, Araraquara/SP. Período: março 2017 – fevereiro 2019. Título da dissertação: Estudos computacionais de ligantes dos receptores nicotínicos de acetilcolina do subtipo $\alpha 4\beta 2$. Orientador: Prof. Dr. Nailton Monteiro do Nascimento Júnior, Bolsa: CAPES.

Graduação em Química: Universidade Estadual Paulista Júlio de Mesquita Filho, Araraquara/SP Período: fevereiro 2013 - dezembro 2016 Título do trabalho de conclusão de curso: Estudos de modelagem molecular de compostos bioativos frente ao receptor nicotínico de acetilcolina do subtipo $\alpha 4\beta 2$. Orientador: Prof. Dr. Nailton Monteiro do Nascimento Júnior, Bolsa: PROPe.

PRODUÇÃO BIBLIOGRÁFICA

Artigos completos publicados em periódicos

BATISTA, V. S.; NASCIMENTO-JÚNIOR, N. M. Molecular docking: considerations of a low cost and suitable methodology and some successful applications. *Medicinal & Analytical Chemistry International Journal*, 2, 000123, **2018**.

OKADA-JUNIOR, C. Y.; MONTEIRO, G. C.; AGUIAR, A. C. C.; BATISTA, V. S.; DE SOUZA, J. O.; SOUZA, G. E.; BUENO, R. V.; OLIVA, G.; NASCIMENTO-JÚNIOR, N. M.; GUIDO, R. V. C.; BOLZANI, V. S. Phthalimide Derivatives with Bioactivity against *P. falciparum*: Synthesis, Evaluation and Computational Studies Involving bc1 Cytochrome Inhibition. *ACS Omega*, 3, 9424-9430, **2018**.

BATISTA, V. S.; NASCIMENTO-JÚNIOR, N. M. Receptores Nicotínicos de Acetilcolina dos subtipos $\alpha 7$, $\alpha 4\beta 2$ e $\alpha 3\beta 4$: Características, Participação em Processos Patogênicos e Compostos Bioativos. *Revista Virtual de Química*, 8, 1721-1739, **2016**.

Trabalhos publicados em anais de eventos científicos

BATISTA, V. S.; NASCIMENTO-JÚNIOR, N. M. Molecular modeling studies of the nicotinic acetylcholine receptor (nAChRs) $\alpha 4\beta 2$: Validation through docking-based comparative intermolecular contacts analysis. 41^a Reunião Anual da Sociedade Brasileira de Química, **2018**, Foz do Iguaçu – Paraná.

PIROLLA, N. F. F.; BATISTA, V. S.; NASCIMENTO-JÚNIOR, N. M. Molecular modeling studies of the nicotinic acetylcholine receptor $\alpha 7$: Validation and analysis of bioactive compounds. 41^a Reunião Anual da Sociedade Brasileira de Química, **2018**, Foz do Iguaçu – Paraná.

RIBEIRO, T. R.; BATISTA, V. S.; NASCIMENTO-JÚNIOR, N. M. Molecular modeling studies of the *Penicillium Camemberti* Lipase (PCL) enzyme: Docking and validation

of bioactive compounds. 41ª Reunião Anual da Sociedade Brasileira de Química, **2018**, Foz do Iguaçu – Paraná.

BATISTA, V. S.; RODRIGUEZ, M. P.; PEZZA, H. R. Síntese e estudo da potencialidade do uso de nanopartículas de prata para a detecção de gentamicina. In: XXVII Congresso de Iniciação Científica da UNESP, **2015**, Araraquara – São Paulo.

Apresentação de trabalho e/ou palestra

BATISTA, V. S.; NASCIMENTO-JÚNIOR, N. M. Molecular modeling studies of the nicotinic acetylcholine receptor (nAChRs) $\alpha 4\beta 2$: Validation through docking-based comparative intermolecular contacts analysis. 41ª Reunião Anual da Sociedade Brasileira de Química, **2018**, Foz do Iguaçu – Paraná.

Ministrador do minicurso “Tutorial de química computacional no laboratório de informática” realizado no dia 11 de janeiro de **2019** durante a 2ª Escola de Modelagem Molecular da Unesp (2ª EMMUnesp), com carga horária de 4 horas.

BATISTA, V. S.; RODRIGUEZ, M. P.; PEZZA, H. R. Síntese e estudo da potencialidade do uso de nanopartículas de prata para a detecção de gentamicina. In: XXVII Congresso de Iniciação Científica da UNESP, **2015**, Araraquara – São Paulo.

PARTICIPAÇÃO EM EVENTOS CIENTÍFICOS

41ª Reunião Anual da Sociedade Brasileira de Química. Trabalho apresentado: Molecular modeling studies of the nicotinic acetylcholine receptor (nAChRs) $\alpha 4\beta 2$: Validation through docking-based comparative intermolecular contacts analysis, **2018**. Foz do Iguaçu – Paraná.

40ª Reunião Anual da Sociedade Brasileira de Química, **2017**. São Paulo – São Paulo/46th World Chemistry Congress (IUPAC), **2017**. São Paulo – São Paulo.

2º Simpósio de Química Medicinal da UNESP (2º SIMQUIMED), **2016**. São José do Rio Preto – São Paulo.

IV International Symposium on Drug Discovery, **2015**. Araraquara – São Paulo.

XXVII Congresso de Iniciação Científica da UNESP. Trabalho apresentado: Síntese e estudo da potencialidade do uso de nanopartículas de prata para a detecção de gentamicina, **2015**. Araraquara – São Paulo.

PRÊMIOS E TÍTULOS

Menção Honrosa de terceiro lugar para o trabalho: RIBEIRO, T. R.; BATISTA, V. S.; NASCIMENTO-JÚNIOR, N. M. Estudos de modelagem molecular da enzima *Penicillium Camemberti* Lipase (PCL): ancoragem e validação de compostos bioativos. III Simpósio de Química Medicinal – “Desafios na Pesquisa e Inovação de Antibióticos” (3º SIMQUIMED), **2018**. São José do Rio Preto – São Paulo.

DEMAIS ATIVIDADES ACADÊMICAS

Estágio de Docência UNESP- Curso de Bacharelado em Engenharia Química, auxiliando na disciplina de Química Orgânica Experimental, adquirindo experiência e desenvoltura na comunicação com os discentes e na prática de docência, supervisionado pelo Prof. Dr. Nailton Monteiro do Nascimento Júnior no ano de **2017** (Carga horária: 60h).

ORGANIZAÇÃO DE EVENTO CIENTÍFICO

Membro da comissão organizadora: Primeira Escola de Modelagem Molecular da Unesp (1ª EMMUnesp), Araraquara/SP, **2017**.

Membro da comissão organizadora: Segunda Escola de Modelagem Molecular da Unesp (2ª EMMUnesp), Araraquara/SP, **2019**.

AGRADECIMENTOS

Aos meus pais, Maria Beatriz de Sousa Batista e Oseas de Sousa Batista, por todo o amor, suporte e apoio incondicional durante minha formação pessoal e profissional. Ao meu irmão, Vinicius de Sousa Batista, por sempre estar presente desde a infância até hoje e por sempre me apoiar em todas as situações.

Ao meu orientador, Prof. Dr. Nailton Monteiro do Nascimento Júnior, pela excelente orientação e amizade que compartilhamos ao longo dos anos trabalhando juntos. Agradeço imensamente a confiança em mim depositada e aos ensinamentos que recebi de você, sem os quais eu não seria metade do pesquisador que sou hoje.

Às colegas do LaQMedSOMM, Thaiz R. Ribeiro, Natália F. Frediani e Marcelle V. Rocha, com as quais trabalhei ao longo desse projeto, agradeço pelas discussões científicas que permitiram o enriquecimento do trabalho realizado.

Ao Prof. Dr. Mutasem O. Taha, da unidade de descoberta de fármacos (*Drug Discovery Unit*) do departamento de ciências farmacêuticas da faculdade de farmácia da Universidade da Jordânia (Amman, Jordânia) pela pronta disponibilização dos scripts referentes à metodologia dbCICA e por todo o auxílio prestado durante a implementação da mesma.

A todos os meus “irmãos” da “República Atecubanos”, Caio Mathias (Fima), Matheus Mendes (Fetão), Kevin Lino (Marrone), Lucas Tavares (Valeta), Victor Carmo (Steak), Gustavo Sanches (Monge), Vitor Batistela (Zero Um), Gabriel Rodrigues (Xynelo), Gabriel Bernardeli (Mustache), Luiz Felipe Bertholdo (Perereka), Gabriel Marino (Waltinho) e Leandro Freitas (Set). Meus mais sinceros agradecimentos por todos os momentos inesquecíveis e também pelos momentos esquecíveis que passamos ao longo dessa jornada. Minha vida nessa universidade não seria a mesma sem vocês e certamente será mais triste longe de uma família como essa.

Em especial, agradeço ao meu grande amigo Guilherme G. Costa (Sumido) por todos os anos de camaradagem, desde os meus primeiros dias na universidade até hoje, sem a qual eu não teria forças pra alcançar meus objetivos. Obrigado, meu irmão, por todas as palavras de apoio ao longo dos anos e por sempre estar comigo nas horas boas, nas ruins e nas difíceis de classificar.

À banca examinadora, Dr. Natanael Dante Segretti, dá Aché Laboratórios Farmacêuticos e Prof^a. Dr^a Cíntia Duarte de Freitas Milagre, do Instituto de Química

da Unesp de Araraquara, pelas excelentes contribuições que possibilitaram um trabalho final mais completo.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa concedida e à Fundação de Amparo à pesquisa do Estado de São Paulo (FAPESP, Processo 2018/00187-7) pelos recursos fornecidos.

RESUMO

Os nAChRs (receptores nicotínicos de acetilcolina) são canais iônicos pentaméricos que respondem ao neurotransmissor endógeno acetilcolina, com os subtipos $\alpha 4\beta 2$ e $\alpha 7$ sendo expressos em grandes quantidades no cérebro humano. Esses receptores estão envolvidos em desordens neurológicas como a doença de Alzheimer, esquizofrenia, dor crônica e autismo, bem como no vício em nicotina. Neste contexto, o SBDD (*structure-based drug design*) é uma ferramenta poderosa para o desenvolvimento de novos candidatos a fármacos visando estes receptores. Neste trabalho, aplicou-se a metodologia dbCICA (*docking-based comparative intermolecular contacts analysis*) para a identificação das condições de ancoragem molecular ótimas para estes dois receptores. Em resumo, compostos bioativos atuando em cada um dos dois receptores foram coletados da literatura científica e ancorados nos receptores $\alpha 4\beta 2$ e $\alpha 7$, explorando diferentes condições de ancoragem. Em seguida, os resultados foram pós-processados através da metodologia dbCICA, uma abordagem baseada em algoritmo genético, visando a identificação de contatos críticos para o reconhecimento molecular entre os ligantes ancorados e os resíduos do sítio ativo. Através dessa análise, foi possível determinar as condições ótimas de ancoragem para ambos os receptores trabalhados. Os contatos críticos identificados através do dbCICA estão em acordo com relatos da literatura sobre interações críticas dessa classe de receptores. As condições ideais de ancoragem para o receptor $\alpha 4\beta 2$ foram utilizadas para selecionar 8 dentre 116 fragmentos moleculares inéditos com potencial atividade biológica frente ao mesmo para serem sintetizados em etapas futuras. Utilizando os dados obtidos para o subtipo $\alpha 4\beta 2$ também foi possível construir um mapa farmacofórico, validado através da curva ROC (*receiver operating characteristic curve*), que pode ser utilizado em triagens virtuais subsequentes e também em uma posterior etapa de incrementação estrutural. Além disso, foi possível triar 8 fragmentos moleculares inéditos, a partir dos mesmos 116, utilizando as condições de ancoragem ótimas identificadas para o receptor $\alpha 7$ com o intuito de avaliar a possibilidade de seletividade ou atividade dual.

Palavras – chave: nAChRs $\alpha 4\beta 2$. nAChRs $\alpha 7$. Ancoragem Molecular. dbCICA. Mapa farmacofórico. Química Medicinal.

ABSTRACT

The nAChRs (nicotinic acetylcholine receptors) are pentameric ligand-gated ionic channels that respond to the endogenous neurotransmitter acetylcholine, with the $\alpha 4\beta 2$ and $\alpha 7$ subtypes highly expressed in human brain. Those receptors are involved in many neurologic disorders such as Alzheimer's disease, Schizophrenia, chronic pain, autism, as well as in nicotine addiction. In this context, SBDD (structure-based drug design) is a powerful tool for designing novel drug candidates targeting those receptors. In this work, we applied dbCICA (docking-based comparative intermolecular contacts analysis) to identify optimal docking conditions for these two receptors. Briefly, bioactive compounds acting on each receptor were gathered from the scientific literature and docked into the binding sites of the $\alpha 4\beta 2$ and $\alpha 7$, exploring different docking conditions. Then, the results were post-processed through dbCICA, a genetic algorithm based approach, aiming to identify critical contacts for molecular recognition involving the docked ligands and binding site residues. Through this analysis it was possible to determine the optimal docking conditions for both receptors. The critical contacts identified by dbCICA are in agreement with the literature about critical interactions from this class of receptors. The optimal docking conditions for receptor $\alpha 4\beta 2$ were used to select 8 from 116 novel molecular fragments with potential biological activity towards this subtype, to be synthesized in future steps. Using the outputs from subtype $\alpha 4\beta 2$ it was possible to build a pharmacophoric map, validated by ROC curve (receiver-operating characteristic curve), which can be applied to further virtual screening and also in a fragment growth step latter on. Also, it was possible to select 8 from the same 116 novel molecular fragments using the optimal docking conditions for $\alpha 7$ nAChR aiming to evaluate if selectivity is possible, as well as dual activity.

Keywords: nAChRs $\alpha 4\beta 2$. nAChRs $\alpha 7$. Molecular docking. dbCICA. Pharmacophoric map. Medicinal Chemistry.

LISTA DE FIGURAS

Figura 1 - Estruturas químicas das moléculas de acetilcolina (1), muscarina (2), nicotina (3), ácido gama amino-butírico (4), glicina (5) e serotonina (6)	16
Figura 2 - Representação do nAChR (α_4) ₂ (β_2) ₃	17
Figura 3 – Representação simplificada dos nAChRs neuronais dos subtipos α_7 e (α_4) ₂ (β_2) ₃ e seus sítios ortostéricos, com a face principal em azul e a face complementar em verde.	19
Figura 4 - Representação tridimensional da sequência de aminoácidos da α -Btx, com a representação <i>cartoon</i> da mesma em magenta.	21
Figura 5 - Exemplos de compostos bioativos documentados na literatura que atuam nos nAChRs do subtipo α_7	22
Figura 6 - Curvas típicas de resposta monofásica (a) e bifásica (b) e estequiometrias correspondentes.....	25
Figura 7 - Exemplos de compostos bioativos documentados na literatura que atuam nos nAChRs $\alpha_4\beta_2$	26
Figura 8 – Fluxograma das etapas presentes em uma abordagem de SBDD.	28
Figura 9 – Métodos de introdução de flexibilidade para proteínas em ancoragem molecular.....	32
Figura 10 – Nicotina co-cristalizada na estrutura do nAChRs do subtipo $\alpha_4\beta_2$ (laranja, PDB ID: 5KXI) e pose de ancoragem bem-sucedida na avaliação por redocagem.....	37
Figura 11 – Exemplo de uma curva ROC hipotética.	38
Figura 12 - Compostos heterocíclicos com atividade frente aos nAChRs $\alpha_4\beta_2$ e α_7 , com destaque para os produtos naturais (-)-nicotina (3), (-)-citisina (17) e o fármaco vareniclina (13).	42
Figura 13 - Planejamento estrutural dos 116 fragmentos moleculares (21-24, a-zc), com padrões estruturais quinoxalínico e benzimidazólico, planejados como ligantes dos nAChRs $\alpha_4\beta_2$	45
Figura 14 - Corridas de docking realizadas para cada uma das 4 funções de pontuação do software GOLD. ...	48
Figura 15 – Exemplo de distâncias coletadas (esquerda) e matrizes binárias geradas (direita).	48
Figura 16 – Fluxograma do processo de seleção de variáveis através de um algoritmo genético.	50
Figura 17 – Esquema de validação cruzada 5-fold.	51
Figura 18 - Pose do ligante co-cristalizado (3) (em laranja) e resíduos da região do sítio ativo.....	55
Figura 19 - Pose de 3 co-cristalizado (laranja) sobreposta à 3 ancorado (ciano) e resíduos da região do sítio ativo. RMSD = 0,83 Å.	58
Figura 20 - Comparação entre a orientação espacial do ligante 13 a) ancorado no nAChRs $\alpha_4\beta_2$ e b) co-cristalizado na AChBP da espécie <i>Aplysia californica</i> (4AFT).	59
Figura 21 - Fórmula estrutural, massa molecular (MM), área de superfície polar (PSA) e coeficiente de partição (cLogP) dos 8 fragmentos selecionados para a etapa de síntese.	61
Figura 22 - Sobreposição das poses melhor ranqueadas para os 8 fragmentos moleculares (laranja) triados para a etapa de síntese.....	63
Figura 23 - a) Contatos positivos (verde) e negativos (azul) apontados pelo modelo nAChRs_ $\alpha_4\beta_2$ _1. b) Sobreposição dos ligantes bem-comportados 3, 13, 20, $\alpha_4\beta_2$ _6, $\alpha_4\beta_2$ _9, $\alpha_4\beta_2$ _13, $\alpha_4\beta_2$ _14, $\alpha_4\beta_2$ _34, $\alpha_4\beta_2$ _44, $\alpha_4\beta_2$ _45, $\alpha_4\beta_2$ _63, $\alpha_4\beta_2$ _71, $\alpha_4\beta_2$ _86, $\alpha_4\beta_2$ _88 e $\alpha_4\beta_2$ _92 dentro do sítio ativo mostrado em a)..	64
Figura 24 - Mapa farmafofocófico do modelo dbCICA nAChRs_ $\alpha_4\beta_2$ _1 combinado com as esferas de exclusão de nAChRs_ $\alpha_4\beta_2$ _2.....	65
Figura 25 - Curva ROC referente ao mapa farmafofocófico do modelo dbCICA nAChRs_ $\alpha_4\beta_2$ _1 combinado com as esferas de exclusão de nAChRs_ $\alpha_4\beta_2$ _2..	66
Figura 26 – Modo de ligação do ligante co-cristalizado 7 e resíduos da região do sítio ativo.....	68
Figura 27 - Pose de 7 co-cristalizado (laranja) sobreposta à 7 ancorado (ciano) e resíduos da região do sítio ativo..	71
Figura 28 - a) Sobreposição dos resíduos dos sítios ativos e ligantes co-cristalizados das estruturas 5AFH (cadeia B em azul, cadeia C em verde, 7 em laranja) e 5OUH (amarelo, 7 em ciano). b) Sobreposição da pose de 7 co-cristalizada em 5OUH (ciano) e pose de ancoragem obtida para 7 utilizando os parâmetros do modelo nAChRs_ α_7 _1 (verde), com RMSD = 0,65 Å.	73
Figura 29 - Fórmula estrutural, massa molecular (MM), área de superfície polar (PSA) e coeficiente de partição (cLogP) dos 8 fragmentos selecionados através da triagem virtual para o receptor α_7	74
Figura 30 - Sobreposição das poses melhor ranqueadas para os 8 fragmentos moleculares triados para o subtipo α_7	76

LISTA DE TABELAS

Tabela 1 - Possíveis receptores alvos no tratamento de algumas desordens neurológicas e região do cérebro afetada pela doença.	19
Tabela 2 - Exemplos de algoritmos de amostragem conformacional.	29
Tabela 3 - Exemplos de funções de pontuação.	33
Tabela 4 - Condições e estatística dos 10 melhores modelos dbCICA.	56
Tabela 5 - Contatos críticos dos 2 melhores modelos dbCICA para o subtipo $(\alpha 4)_2(\beta 2)_3$	56
Tabela 6 - Interações observadas para 3 em 5KXI, interações previstas para 3 e 13 e interações previstas para os 8 fragmentos selecionados para síntese.	61
Tabela 7 - Condições e estatística dos 10 melhores modelos dbCICA para o nAChRs do subtipo $\alpha 7$	69
Tabela 8 - Contatos críticos dos 2 melhores modelos dbCICA para o nAChRs do subtipo $\alpha 7$	69
Tabela 9 - Interações observadas para 7 em 5AFH, interações previstas para 7 e interações previstas para os 8 fragmentos selecionados pelo modelo nAChRs_ $\alpha 7_1$	74

LISTA DE ABREVIATURAS E SIGLAS

3D-QSAR: *tridimensional quantitative structure-activity relationship*

5-HT₃: receptor de serotonina 3

ACh: acetilcolina

AChBP: proteína ligante de acetilcolina

AChE: acetilcolinesterase humana

CaNMT: *Candida albicans* N-myristoyl transferase

Chk1: *checkpoint kinase 1* humana

CoMFA: *conformational molecular field analysis*

COPD: *chronic obstructive pulmonary disease*

CPU: *central processing unit*

Da: Dalton

DAT: transportadores de dopamina

dbCICA: *docking-based comparative intermolecular contacts analysis*

DNA: ácido desoxirribonucleico

EC₅₀: concentração mínima de um composto necessária para produzir metade do efeito máximo, *in vivo*.

ELIC: canal iônico ligante-dependente da bactéria *Erwinia chrysanthemi*

FBVS: *fragment-based virtual screening*

FGFR1: fator de crescimento fibroblástico humano

GABA_A e **GABA_C:** receptores de ácido gama-aminobutírico A e C

GB/SA: *Generalized Born Surface Area*

GK: glucoquinase humana

GLIC: canal iônico ligante-dependente da cianobactéria *Gloeobacter violaceus*

GP: glicogênio fosforilase humana

HKU-Cov 3CL^{pro}: *3C-like* protease do coronavírus da síndrome respiratória do oriente médio.

Hsp90: proteína de choque térmico 90 humana

HTS: *high throughput screening*

IC₅₀: concentração mínima de um composto necessária para reduzir o crescimento populacional de organismos em 50 por cento, *in vitro*.

K_i: constante de dissociação do complexo enzima-inibidor

mAChRs: receptores muscarínicos de acetilcolina

MM: Massa molecular
nAChRs: receptores nicotínicos de acetilcolina
PB/SA: *Poisson–Boltzmann Surface Area*
PDB: *Protein Data Bank*
PSA: área de superfície polar
 r^2 : coeficiente de correlação simples
 $r^2_{K\text{-fold}}$: coeficiente de correlação *K-fold*
 r^2_{Loo} : coeficiente de correlação *leave-one-out*
RAM: *random access memory*
RMSD: *root-mean-squared deviation*
RO3: regra dos 3 de Lipinsky
ROC: *receiver-operating characteristic*
SBDD: *structure-based drug design*
SERT: transportadores de serotonina
SNC: sistema nervoso central
SNP: sistema nervoso periférico
SUS: sistema único de saúde
 α -Btx: alfa-bungarotoxina
 α -Ctx: alfa-conotoxina

SUMÁRIO

1. INTRODUÇÃO.....	16
1.1. RECEPTORES NICOTÍNICOS DE ACETILCOLINA (nAChRs)	16
1.1.1. OS RECEPTORES NICOTÍNICOS DE ACETILCOLINA (nAChRs) DO SUBTIPO $\alpha 7$	20
1.1.2. OS RECEPTORES NICOTÍNICOS DE ACETILCOLINA (nAChRs) DO SUBTIPO $\alpha 4\beta 2$	23
1.2. INTRODUÇÃO À ANCORAGEM MOLECULAR NO CONTEXTO DA QUÍMICA MEDICINAL.....	27
1.2.1. AMOSTRAGEM CONFORMACIONAL DOS LIGANTES NA ANCORAGEM MOLECULAR	29
1.2.2. FLEXIBILIDADE DA PROTEÍNA NA ANCORAGEM MOLECULAR.....	30
1.2.3. FUNÇÕES DE PONTUAÇÃO.....	32
1.2.4. VALIDAÇÃO DE MODELOS DE ANCORAGEM MOLECULAR	36
1.2.5. ANÁLISE COMPARATIVA DE CONTATOS INTERMOLECULARES BASEADA EM ANCORAGEM MOLECULAR (dbCICA).....	39
2. JUSTIFICATIVA E OBJETIVOS.....	41
3. METODOLOGIA	44
3.1. HARDWARE UTILIZADO	44
3.2. PLANEJAMENTO ESTRUTURAL.....	44
3.3. OBTENÇÃO E PREPARAÇÃO DAS ESTRUTURAS TRIDIMENSIONAIS DOS nAChRs E DOS LIGANTES USADOS.....	45
3.4. ANCORAGEM MOLECULAR.....	46
3.5. VALIDAÇÃO POR dbCICA	47
3.6. TRIAGEM VIRTUAL DOS 116 FRAGMENTOS PROPOSTOS.....	52
3.7. MAPA FARMACOFÓRICO DO MODELO $(\alpha 4)_2(\beta 2)_3$	52
3.8. VALIDAÇÃO DO MAPA FARMACOFÓRICO $(\alpha 4)_2(\beta 2)_3$	53
4. RESULTADOS E DISCUSSÃO	54
4.1. SUBTIPO $(\alpha 4)_2(\beta 2)_3$	54
4.2. SUBTIPO $\alpha 7$	67
5. CONCLUSÕES.....	78
6. REFERÊNCIAS	80
7. ANEXOS.....	92
8. ARTIGOS PUBLICADOS NO PERÍODO DE DESENVOLVIMENTO DESTES TRABALHOS OU RELACIONADOS.....	105

1. Introdução

1.1. Receptores nicotínicos de acetilcolina (nAChRs)

Os receptores nicotínicos de acetilcolina (nAChRs) representam parte do grupo de receptores do neurotransmissor endógeno acetilcolina **(1)** (ACh) (**Figura 1**). Este grupo abrange, além dos já mencionados nAChRs, os receptores muscarínicos de acetilcolina (mAChRs). Os mAChRs são receptores metabotrópicos que são ativados pela toxina do cogumelo *Amanita muscaria*, a muscarina **(2)**, em um período de tempo relativamente longo, da ordem de milissegundos a segundos, enquanto os nAChRs são receptores ionotrópicos que são ativados em intervalos de tempo da ordem de sub-microsegundos a microssegundos pela molécula de nicotina **(3)**,¹⁻³ comumente proveniente da planta *Nicotiana tabacum*. Os nAChRs podem ser divididos em dois grupos, os receptores musculares e os neuronais,^{4,5} com os nAChRs neuronais majoritariamente expressos nas formas $\alpha 7$, $\alpha 4\beta 2$ e $\alpha 3\beta 4$, podendo exercer funções variadas, como regulação de excitabilidade neuronal e liberação de neurotransmissores.⁶⁻⁸

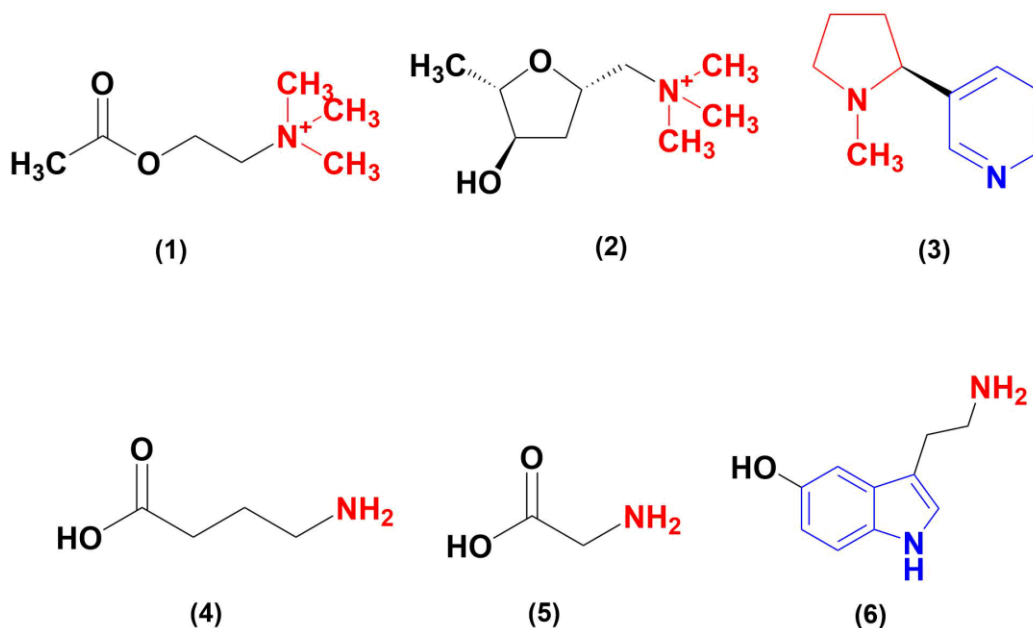


Figura 1 - Estruturas químicas das moléculas de acetilcolina **(1)**, muscarina **(2)**, nicotina **(3)**, ácido gama amino-butírico **(4)**, glicina **(5)** e serotonina **(6)**.

Obs: A seção da introdução sobre os nAChRs (seção 1.1) foi adaptada de BATISTA, V. S.; NASCIMENTO-JÚNIOR, N. M. *Revista Virtual de Química*, 8, 1721-1739 (2016), contido no final do documento. Para uma análise de relação estrutura-atividade de ligantes dos receptores discutidos neste trabalho, vide o anexo citado.

Estruturalmente, os nAChRs neuronais são glicoproteínas grandes (aproximadamente 290 kDa), que apresentam cinco subunidades transmembranares arranjadas em torno de um poro central permeável a cátions^{3,7,9} e fazem parte da família dos canais iônicos cys-loop, que inclui também os nAChRs musculares, receptores de ácido gama amino-butírico **(4)** (GABA_A e GABA_C), receptores de glicina **(5)** e receptores de serotonina **(6)** (5-HT₃).^{7,10} Cada receptor é dividido em três domínios: um grande domínio extracelular amino-terminal onde se encontram os sítios ortostéricos, um poro transmembranar e um pequeno domínio intracelular, tendo esse conjunto um comprimento total de aproximadamente 160 Å, conforme representado na **Figura 2**.

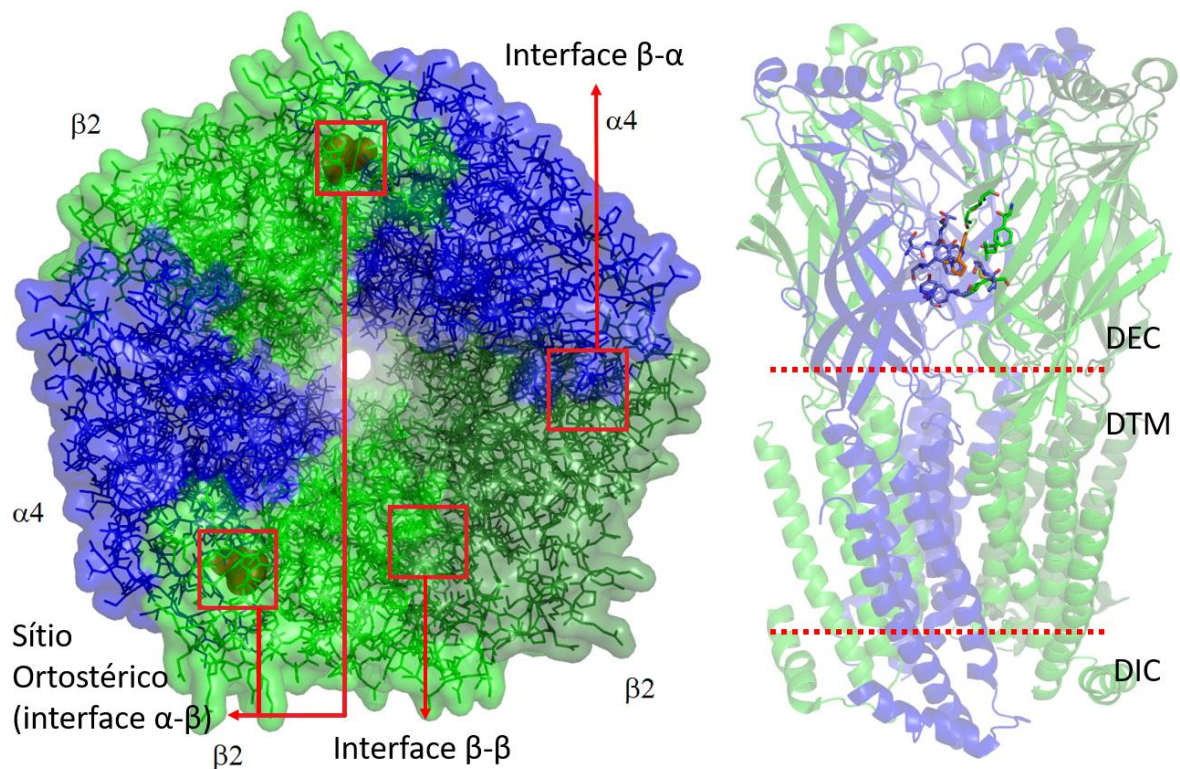


Figura 2 - Representação do nAChR $(\alpha 4)_2(\beta 2)_3$. À esquerda, quadrados em vermelho representam as interfaces β - α , β - β e α - β sendo o sítio ortostérico localizado na interface α - β . As seções mostradas no corte transversal à esquerda representam os domínios intracelular (DIC), domínio transmembranar (DTM) e domínio extracelular (DEC).

Os nAChRs neuronais apresentam uma grande variedade de funções biológicas que estão relacionadas com as diversas combinações possíveis entre as subunidades α e β , a saber, α ($\alpha 2$ - $\alpha 10$) e β ($\beta 2$ - $\beta 5$), podendo estas combinações serem homoméricas, como os receptores $\alpha 7$, $\alpha 8$ e $\alpha 9$ ou heteroméricas, como os receptores $\alpha 4\beta 2^*$, $\alpha 7\alpha 8^*$, $\alpha 9\alpha 10^*$, $\alpha 6\beta 2^*$, $\alpha 3\beta 4^*$, entre outros, com a presença obrigatória de uma subunidade α no receptor.^{3,5,7} O asterisco representa a

possibilidade de outras subunidades no complexo. Classifica-se as subunidades como α ou não- α baseando-se na presença ou ausência, respectivamente, de cisteínas vicinais conectadas por ligações dissulfeto no domínio extracelular do receptor.¹¹ Ao interagirem com seu neurotransmissor endógeno ou com um agonista, a estrutura dos nAChRs sofre mudança conformacional, sendo possível descrever o processo de ativação do receptor simplificada em três etapas conformacionais: **(a)** a conformação fechada, onde o receptor está apto a interagir com um agonista e não há a passagem de cátions pelo poro permeável, **(b)** a conformação aberta, onde o receptor já interagiu com um agonista e permite a passagem de cátions em um intervalo de tempo da ordem de milissegundos e **(c)** a conformação dessensibilizada, onde o receptor novamente deixa de permitir a passagem de cátions pelo poro central.¹² O poro central é permeável a cátions monovalentes e bivalentes, principalmente sódio e cálcio.^{13,14} Estudos sugerem que ao retornar da conformação dessensibilizada para a conformação fechada não ocorre a passagem pela conformação aberta, indicando um processo cíclico. Os mesmos estudos mostram também que existem múltiplos estados conformacionais fechado, aberto e dessensibilizado, conectados em uma sequência complexa que ainda precisa ser elucidada para a compreensão adequada do comportamento em nível atômico dos nAChRs.¹²

O sítio ortostérico dos nAChRs neuronais se encontra no domínio extracelular do receptor, nas interfaces entre as subunidades α e as demais subunidades que os constituem. As subunidades α são o componente primário do sítio ativo e as subunidades do tipo β são responsáveis pela parte complementar do mesmo, como exemplificado na **Figura 3**, sendo que a permeabilidade a cátions e afinidade por determinado agonista variam grandemente em função dos componentes complementares do receptor.¹⁵⁻¹⁷ Além dos componentes complementares do sítio ortostérico, os nAChRs podem apresentar uma subunidade assessora que não participa ativamente da formação do sítio ortostérico mas influencia nas propriedades farmacológicas e biofísicas do receptor, podendo esta ser dos subtipos $\alpha 3$ - $\alpha 5$ e $\beta 2$ - $\beta 5$.⁷ Dessa forma, fica clara a diversidade possível para os nAChRs, uma vez que as diversas subunidades existentes podem se combinar em diversas estequiometrias, formando receptores distintos. Também se nota porque é obrigatória a presença de uma subunidade do tipo α , visto que esta é a responsável pela parte principal do sítio ativo. Além do sítio ortostérico, os nAChRs apresentam também sítios alostéricos

onde moléculas podem interagir com o receptor e potencializar (PAM, *positive allosteric modulators*) ou inibir (NAM, *negative allosteric modulators*) a ativação causada por um agonista,¹⁸ sendo esta uma das possíveis abordagens utilizadas no planejamento de novas substâncias bioativas que tem como alvo os nAChRs.

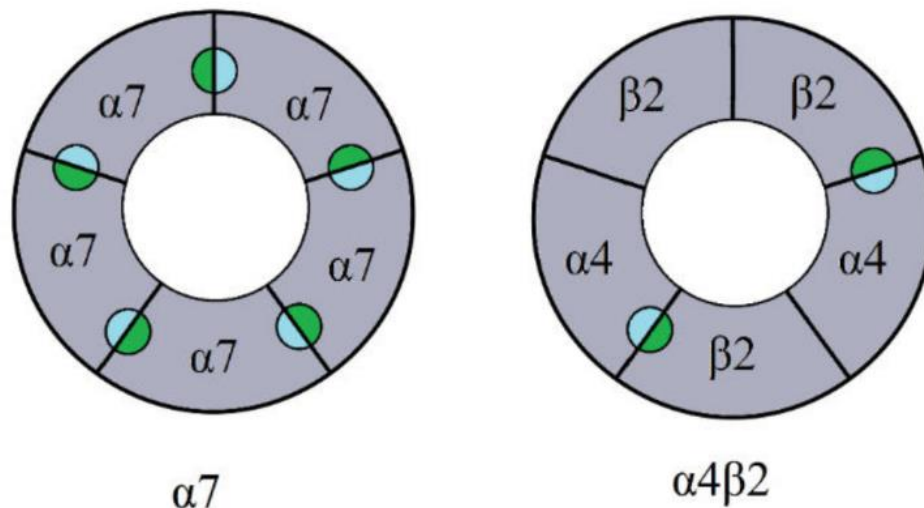


Figura 3 – Representação simplificada dos nAChRs neuronais dos subtipos $\alpha 7$ e $(\alpha 4)_2(\beta 2)_3$ e seus sítios ortostéricos, com a face principal em azul e a face complementar em verde. Outras estequiometrias entre mesmas subunidades apresentadas também são possíveis.³

Por conta de sua grande variedade de funções, os nAChRs neuronais são fundamentais no entendimento e tratamento de diversas desordens neuropsiquiátricas, como a síndrome de Tourette, autismo, déficit de atenção, desordem hiperativa, esquizofrenia, epilepsia, depressão, ansiedade, Alzheimer e Parkinson, conforme pode ser observado na **Tabela 1**.^{6,8,10,19–24}

Tabela 1 - Possíveis receptores alvos no tratamento de algumas desordens neurológicas e região do cérebro afetada pela doença. Dados obtidos da referência.²⁴

Desordem	Receptor	Região do cérebro
Epilepsia	$\alpha 4\beta 2^*$, $\alpha 7^*$, $\beta 4^*$, $\alpha 5^*$	Cortex Frontal, Hipocampo
Alzheimer	$\alpha 4\beta 2^*$, $\alpha 7^*$	Cortex, Hipocampo
Esquizofrenia	$\alpha 4\beta 2^*$, $\alpha 7^*$	Hipocampo
Ansiedade	$\alpha 4^*$, $\alpha 7^*$	Hipocampo
Parkinson	$\alpha 4\beta 2^*$, $\alpha 6\beta 2\beta 3^*$	Sistema Nigroestriatal
Dependência da nicotina	$\alpha 4\beta 2^*$, $\alpha 6\beta 2\beta 3^*$, $\beta 4$	Sistema Mesolímbico

A maior parte do conhecimento estrutural sobre esses receptores é proveniente do estudo de proteínas ligantes de acetilcolina (AChBP) homólogas obtidas de espécies como a *Torpedo marmorata* (PDB ID: 1OED), moluscos como as espécies

Lymnaea stagnalis (PDB ID: 1UW6) e *Aplysia californica* (PDB ID: 2BYN) e das proteínas procarióticas GLIC (canal iônico ligante-dependente da cianobactéria *Gloeobacter violaceus*, PDB ID: 4NPQ) e ELIC (canal iônico ligante-dependente da bactéria *Erwinia chrysanthemi*, PDB ID: 2VL0), entre outras.^{2,9,21,24–29} Apenas recentemente foi elucidada a estrutura cristalográfica dos nAChRs do subtipo $(\alpha 4)_2(\beta 2)_3$ através de difração de raios-X (PDB ID: 5KXI). Em trabalhos subsequentes foram obtidas as estruturas para as estequiometrias $(\alpha 4)_2(\beta 2)_3$ e $(\alpha 4)_3(\beta 2)_2$ (PDB ID: 6CNJ e 6CNK, respectivamente) através de microscopia eletrônica.^{29,30}

1.1.1. Os receptores nicotínicos de acetilcolina (nAChRs) do subtipo $\alpha 7$

Em geral, a localização dos nAChRs neuronais do tipo $\alpha 7$ no cérebro se conserva entre diferentes espécies de mamíferos apesar de diferenças significativas terem sido reportadas entre roedores e primatas em regiões associadas a cognição e processamento sensorial. Esses receptores são expressos grandemente nas regiões do hipocampo, córtex, substância negra e cerebelo.²¹

Estudos recentes com roedores indicam que agonistas parciais e totais dos nAChRs $\alpha 7$ são eficientes no tratamento de distúrbios neurológicos como a esquizofrenia e o mal de Parkinson e também no controle de desordens motoras provocadas por medicamentos, como as discinesias induzidas por levodopa que surgem no decorrer do tratamento do mal de Parkinson.^{6,21,23,31}

Os nAChRs neuronais do tipo $\alpha 7$ são estruturalmente e funcionalmente distintos dos demais nAChRs, apresentando menor afinidade por acetilcolina **(1)** e nicotina **(3)** que os demais, dessensibilização rápida e alta permeabilidade a Ca^{2+} , o que lhes permite regular diversos mecanismos celulares que dependem desse cátion.^{6,18} São constituídos exclusivamente por subunidades do tipo $\alpha 7$, o que confere a eles maior quantidade de sítios ortostéricos do que os demais nAChRs neuronais heteroméricos.^{15,32} Estudos indicam que a ativação dos nAChRs $\alpha 7$ desencadeia maior liberação de glutamato no hipocampo, demonstrando a importância desses receptores para os mecanismos sinápticos glutamatérgicos e modulação dos circuitos neurais do sistema nervoso central (SNC).^{33,34} O planejamento de novos agonistas para esse receptor é complexo porque o uso de grandes quantidades da substância desejada causará dessensibilização do mesmo, sendo necessário um estudo minucioso sobre as quantidades que permitam a ativação adequada do receptor com máxima resposta.

Uma das características dos nAChRs $\alpha 7$ é a sua sensibilidade única ao antagonista peptídico α -Bungarotoxina (**Figura 4**) (α -Btx) produzido pela cobra *Bungarus multicinctus*, que se liga permanentemente a esse receptor e impede a abertura do canal iônico. A utilização dessa substância em conjunto com técnicas de autorradiografia foi essencial na determinação inicial da localização desse receptor no cérebro.^{3,7,21}

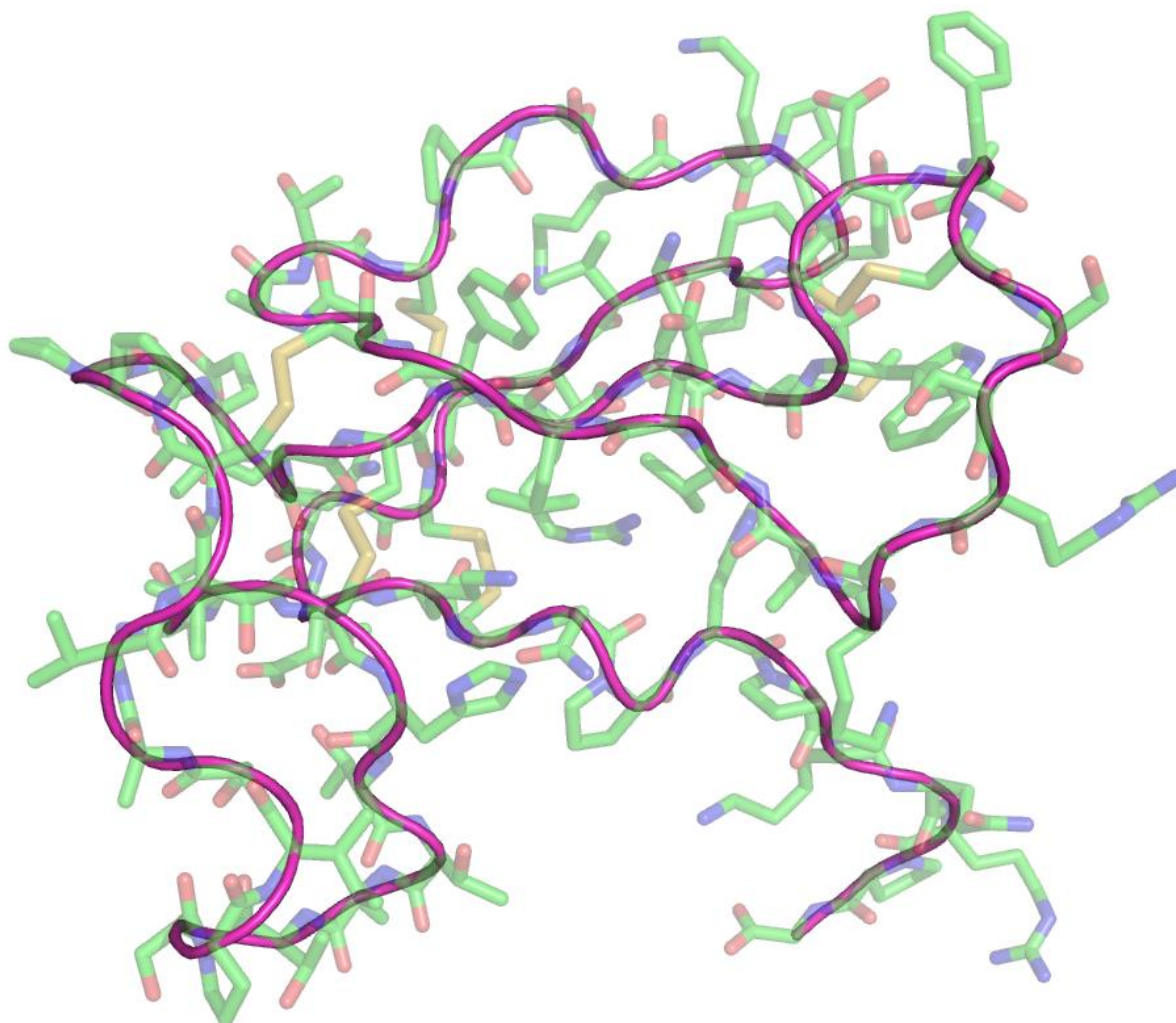


Figura 4 - Representação tridimensional da sequência de aminoácidos da α -Btx, com a representação *cartoon* da mesma em magenta. Sequência de aminoácidos: IV^CHTTATSPISAVT^GPPGENL^CYRKMW^CDAF^CSSRGKVVELG^GAAT^CPSKKPYEEVT^CSTDK^CNP HPKQRPG. Cisteínas destacadas em cores iguais na legenda estão ligadas por ligações dissulfeto. PDB ID: 1IDL.³⁵

Existem diversos compostos bioativos que atuam nos nAChRs $\alpha 7$ descritos na literatura, como exemplificado na **Figura 5**. A lobelina (**7**) é o principal alcaloide da planta *Lobelia inflata*. No passado, já foi um potencial candidato para o tratamento da dependência em nicotina (**3**), mas foi descontinuado devido à falta de evidências de eficiência para essa finalidade.³⁶ A molécula apresenta alta afinidade pelos subtipos

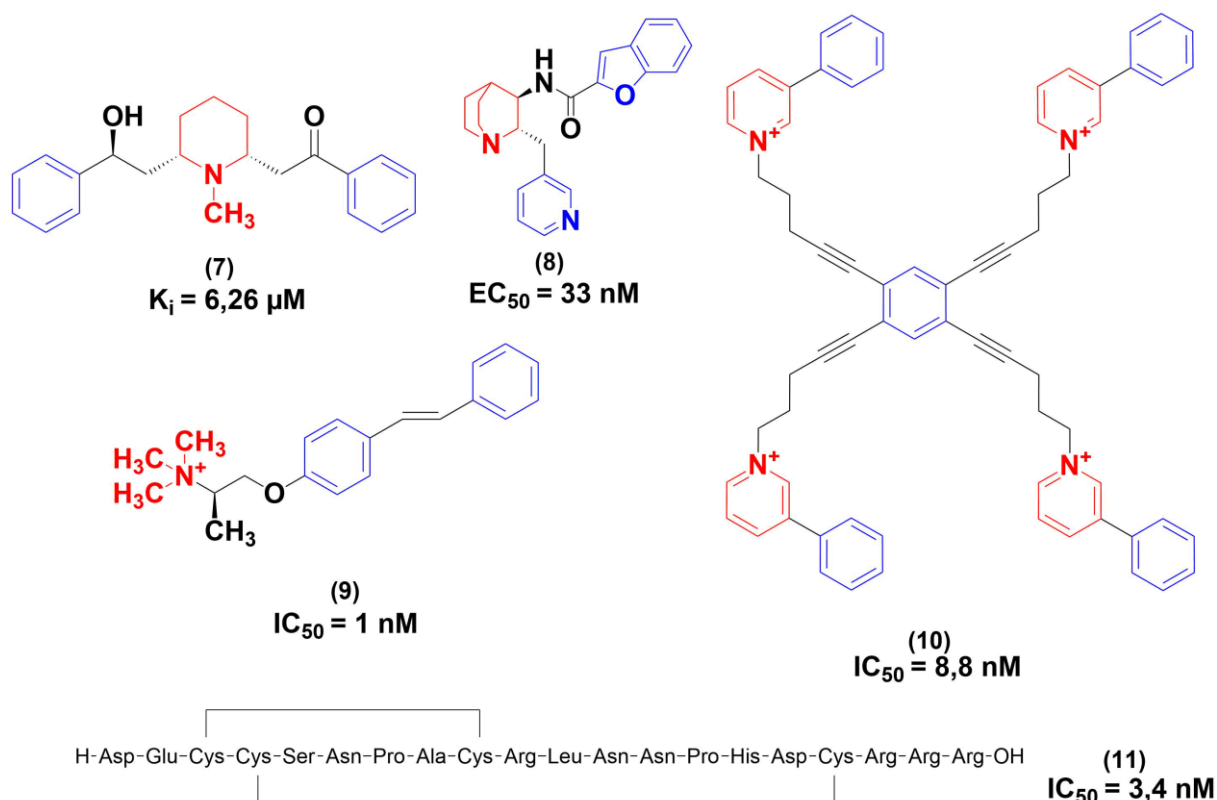


Figura 5 - Exemplos de compostos bioativos documentados na literatura que atuam nos nAChRs do subtipo $\alpha 7$.

$\alpha 7$, com ação antagonista para este subtipo, (**Figura 5**) e $\alpha 4\beta 2^*$ ($K_i = 0,004 \mu\text{M}$) e apresenta inibição do fluxo de dopamina causado pela nicotina em fatias estriatais de ratos. Além disso, a lobelina (**7**) interage com transportadores de monoamina, incluindo os transportadores de dopamina (DAT) e transportadores de serotonina (SERT) de maneira não-específica. Dessa forma, **7** pode ser uma alternativa para o tratamento da dependência em metanfetamina e a molécula atualmente está sendo avaliada para essa finalidade em ensaios clínicos.³⁷

A bradaniclina (**8**) é um agonista parcial desenvolvido, pela Targacept Inc. em colaboração com AstraZeneca, que apresenta boa afinidade pelos nAChRs $\alpha 7$ e baixa afinidade pelo receptor de serotonina 5-HT₃.³⁸ Testes preliminares utilizando roedores sugeriam que este composto apresentava potencial para o tratamento das três categorias de sintomas da esquizofrenia (positivos, negativos e cognitivos) e também efeitos sinérgicos com outros medicamentos antipsicóticos, como a clozapina (Clozaril®, Fazacol®). A bradaniclina também era um medicamento em potencial para o tratamento de Alzheimer e transtorno de déficit de atenção com hiperatividade, mas apesar dos resultados positivos em roedores o composto não obteve aprovação nos testes clínicos para nenhuma das três doenças, tendo seu desenvolvimento descontinuado.³⁹

O estilbeno substituído **9**, inicialmente foi proposto por GOTTI *et al.* (1998),⁴⁰ como um potente antagonista competitivo e seletivo dos nAChRs $\alpha 7$ presentes em cérebros de galinhas ($IC_{50} = 119$ nM). Em estudos posteriores por ANGELANTONIO *et al.* (2000)⁴¹ se demonstrou que o composto também atua nos nAChRs $\alpha 3\beta 4^*$ presentes em células cromafim de roedores, sendo o enantiômero R mais potente do que o enantiômero S nessa situação ($IC_{50} = 330$ nM e $1,5$ μ M, respectivamente). Em estudos subsequentes, GOTTI *et al.* (2000)⁴² demonstraram que o composto **9** apresenta maior afinidade por nAChRs $\alpha 7$ de mamíferos, expressos em ovócitos da espécie *Xenopus*, do que por nAChRs $\alpha 7$ de aves. Esses dados combinados evidenciam que existem diferenças de seletividade e especificidade significativas entre receptores provenientes de espécies distintas.

O composto **10** foi desenvolvido por ZHENG *et al.* (2011)⁴³ tendo como alvo os nAChRs do subtipo $\alpha 9\alpha 10$, que se acredita estarem associados à dores crônicas nas regiões do sistema nervoso periférico (SNP). A molécula apresentou boa inibição do receptor alvo ($IC_{50} = 0,56$ nM) e também do subtipo $\alpha 7$ (**Figura 5**), ambos expressos em ovócitos da espécie *Xenopus*, mas devido à natureza altamente polar da cabeça catiônica, espera-se que o composto apresente baixa biodisponibilidade no cérebro. Essa característica torna muito mais interessante o estudo da molécula **10** visando os nAChRs $\alpha 9\alpha 10$, que não estão presentes no cérebro, mas sim no SNP.

O composto **11** é um peptídeo sintético derivado da classe das α -conotoxinas (α -Ctx) documentado por WHITEAKER *et al.* (2007)⁴⁴ como um antagonista potente e seletivo dos nAChRs $\alpha 7$. As α -Ctxs naturais são peptídeos pequenos (13-19 resíduos de aminoácidos) produzidos por gastrópodes do gênero *Conus* e altamente específicas para alguns tipos de nAChRs. O peptídeo se liga reversivelmente ao receptor, com recuperação total da sensibilidade após 15 minutos, além de inibir tanto receptores de roedores quanto de humanos com potência comparável, o que faz com que esse antagonista seja potencialmente útil no estudo das funções dos nAChRs $\alpha 7$.⁴⁵

1.1.2. Os receptores nicotínicos de acetilcolina (nAChRs) do subtipo $\alpha 4\beta 2$

Os nAChRs $\alpha 4\beta 2$ são o subtipo predominante no cérebro dos mamíferos, sendo altamente sensíveis à acetilcolina (**1**) e nicotina (**3**), sendo considerado o maior responsável pela dependência relacionada a **3**.^{46,47} Quando expostos à nicotina (**3**) e

outros agonistas totais ou parciais por longos períodos de tempo, sofrem um processo chamado suprarregulação em resposta a dessensibilização progressiva provocada, sendo esse efeito muito mais pronunciado nesse receptor do que nos demais nAChRs neuronais.^{7,15,27} A ativação dos nAChRs $\alpha 4\beta 2$ aumenta a liberação de dopamina no sistema mesolímbico, proporcionando a sensação de recompensa característica de diversas drogas de abuso.⁴⁸

Estudos bioquímicos e eletrofisiológicos indicam que os nAChRs $\alpha 4\beta 2$ existem em razões estequiométricas distintas, podendo ser constituídos por duas subunidades α e três subunidades β , $(\alpha 4)_2(\beta 2)_3$, resultando em dois sítios ortostéricos nas interfaces $\alpha 4/\beta 2$, ou por três subunidades α e duas subunidades β , $(\alpha 4)_3(\beta 2)_2$, resultando em três sítios ortostéricos, sendo que esse terceiro sítio existe na interface $\alpha 4/\alpha 4$ e apresenta baixa afinidade por acetilcolina e outros agonistas. As duas estequiometrias são diferentes no que diz respeito a cinética de dessensibilização, permeabilidade a cálcio, condutância unitária, sensibilidade à exposição crônica de nicotina **(3)** e sensibilidade para diversos agonistas. As curvas de resposta à acetilcolina **(1)** são diferentes para cada estequiometria, sendo que o receptor $(\alpha 4)_3(\beta 2)_2$ apresenta uma curva bifásica **(Figura 6b)** característica devido a presença de sítios ortostéricos com afinidades distintas pelo neurotransmissor e o receptor $(\alpha 4)_2(\beta 2)_3$ apresenta uma curva de resposta monofásica **(Figura 6a)**. O receptor $(\alpha 4)_3(\beta 2)_2$ apresenta menor sensibilidade à agonistas em comparação com a outra estequiometria possível, uma vez que é necessária a presença de agonistas em todos os sítios ortostéricos para que o mesmo seja completamente ativado.^{17,32,46,49} Ambas as formas do receptor estão presentes no cérebro dos mamíferos, com estudos indicando que a estequiometria preferencialmente presente no cérebro humano é a de maior sensibilidade e que a proporção entre as duas varia quando ocorre exposição crônica a nicotina, aumentando a quantidade de receptores menos sensíveis com o passar do tempo.^{17,47} Além de mais sensível a nicotina, o receptor $(\alpha 4)_2(\beta 2)_3$ também é mais sensível a dessensibilização causada por exposição prolongada a baixas concentrações dessa substância, enquanto o receptor $(\alpha 4)_3(\beta 2)_2$ apresenta maior permeabilidade a Ca^{2+} e cinética de dessensibilização mais rápida.⁵⁰ Nas regiões do córtex e corpo estriado existe uma subpopulação desse receptor que contém a subunidade $\alpha 5$ na forma $(\alpha 4)_2(\beta 2)_2\alpha 5$.¹⁵

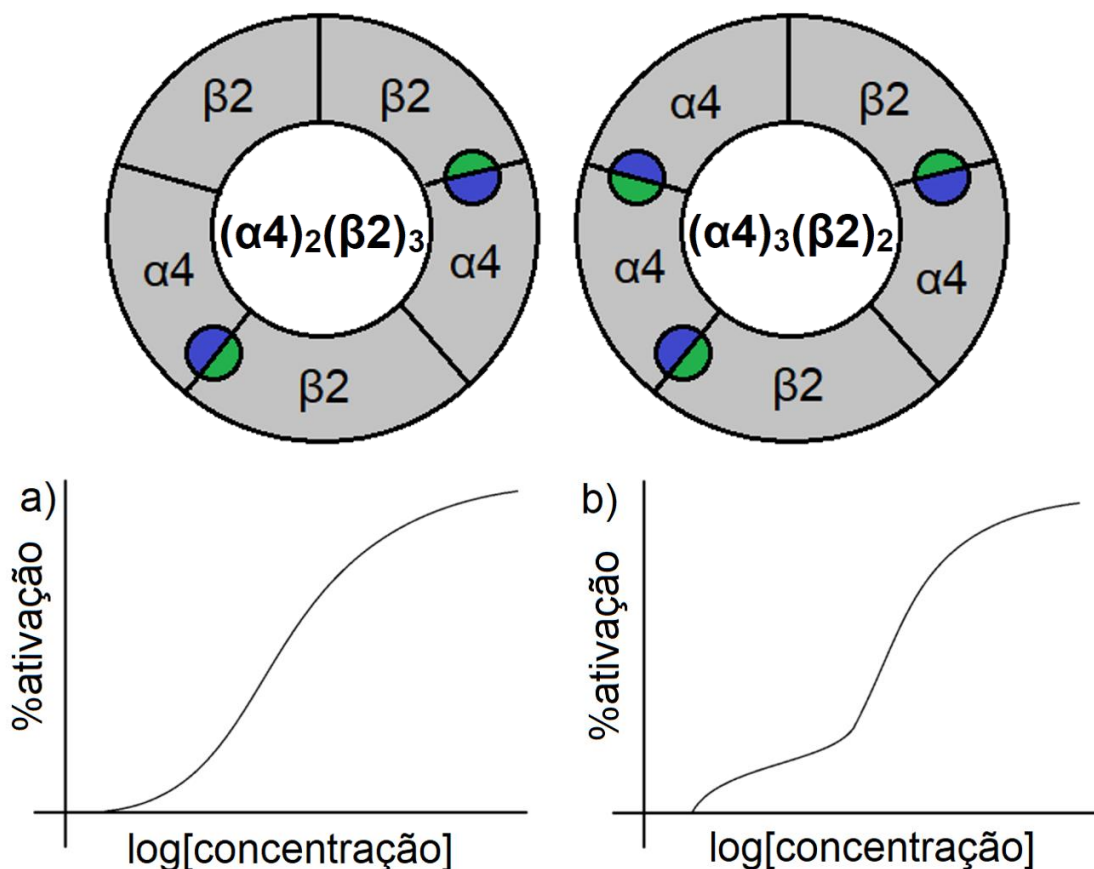


Figura 6 - Curvas típicas de resposta monofásica (a) e bifásica (b) e estequiometrias correspondentes.

Na **Figura 7** estão representados alguns exemplos de compostos bioativos que atuam nos nAChRs $\alpha 4\beta 2$. A epibatidina (**12**) é um alcaloide agonista dos nAChRs do subtipo $\alpha 4\beta 2$ que ocorre naturalmente na pele de sapos da espécie *Epipedobates tricolor* nativos do Equador.⁵¹ A substância é tóxica, com propriedades analgésicas atribuídas à interação com os nAChRs $\alpha 4\beta 2$ e causando paralisia através da interação com os mAChRs, podendo ser letal. A toxicidade acentuada da epibatidina (**12**) e sua baixa seletividade entre os nAChRs eliminou a possibilidade de que a utilizem para fins terapêuticos, mas a substância é largamente utilizada em diversos tipos de estudos acadêmicos, como análises de deslocamento de radioligantes, estudos da relação estrutura-atividade de análogos da molécula e avaliação das propriedades farmacológicas desses análogos.⁵²⁻⁵⁵

O fármaco vareniclina (Champix®, Chantix®) (**13**) é um agonista parcial dos nAChRs $\alpha 4\beta 2$ sintetizado por COE *et al.* (2005)⁵⁶, inspirado no produto natural (-)-citisina, que atenua a resposta dopaminérgica causada pela nicotina e, simultaneamente, reduz os sintomas de abstinência em fumantes. Obteve aprovação

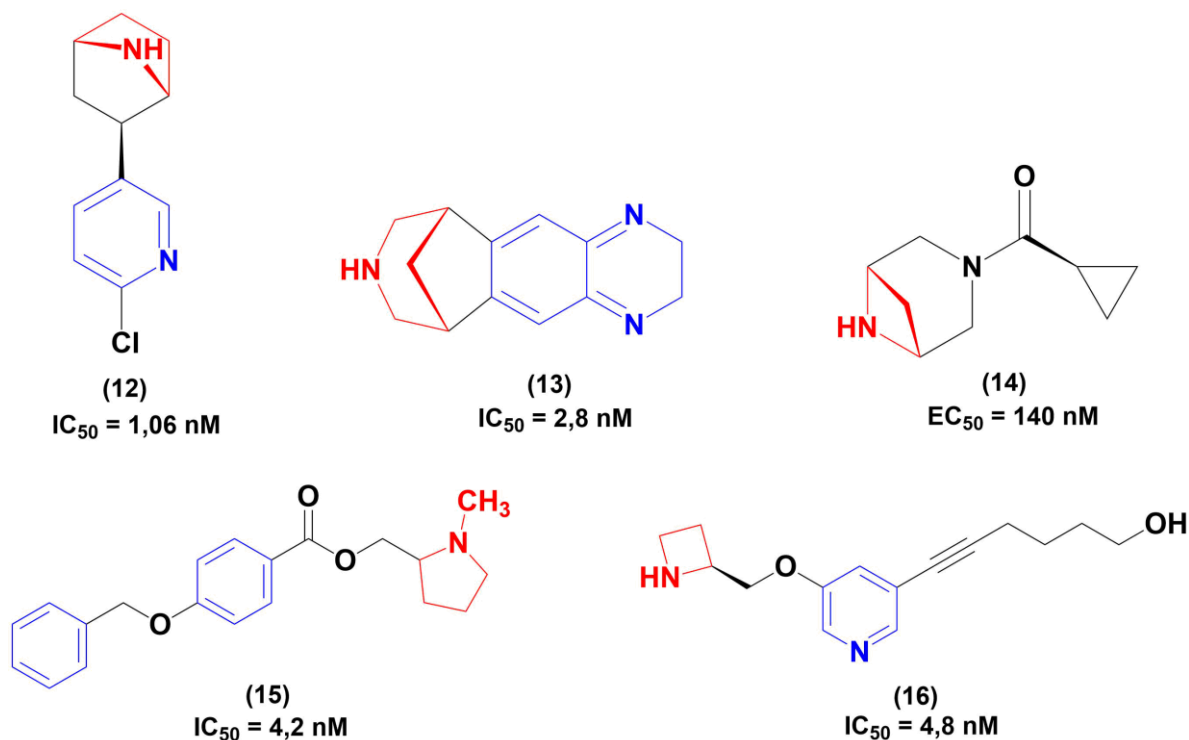


Figura 7 - Exemplos de compostos bioativos documentados na literatura que atuam nos nAChRs $\alpha 4\beta 2$.

para ser comercializado na Europa e EUA no ano de 2006 pela empresa farmacêutica Pfizer®.⁴⁸ É também um agonista total dos nAChRs dos subtipos $\alpha 7$ e $\alpha 4\beta 4$ e parcial do subtipo $\alpha 3\beta 4$.^{57,58} Atualmente, a vareniclina (**13**) e a bupropiona (Zyban®) são os únicos fármacos que não contêm nicotina (**3**) aprovados para o tratamento da dependência da nicotina (**3**).

O TC-8831 (**14**) é um agonista dos nAChRs $\alpha 4\beta 2$ e $\alpha 6/\alpha 3\beta 2\beta 3$ sintetizado por STRACHAN *et al.* (2014)⁵⁹ que apresenta boa afinidade por ambos os receptores ($K_i = 3$ e 21 nM , respectivamente). A molécula atua como um agonista total potente nos nAChRs $(\alpha 4)_2(\beta 2)_3$ (**Figura 7**), mas não é tão potente no receptor $(\alpha 4)_3(\beta 2)_2$ ($EC_{50} = 850 \text{ nM}$). Estudos realizados em primatas indicam que o TC-8831 tem efeitos positivos no tratamento das discinesias induzidas por levodopa, comumente presentes durante a utilização desse medicamento para o tratamento do mal de Parkinson.^{60,61}

O composto neonicotínico **15** foi sintetizado por FAUNDEZ-PARRAGUES *et al.* (2013)⁶², apresentando boa inibição dos nAChRs $\alpha 4\beta 2$, mas com baixa seletividade entre os subtipos desse receptor. Apresenta também baixa afinidade pelos nAChRs $\alpha 7$ ($IC_{50} = 2,6 \mu\text{M}$).

A sazetidina-A (**16**) é um agonista seletivo dos nAChRs $\alpha 4\beta 2$ que possui efeito analgésico e antidepressivo, apresentando atividade agonista total no receptor $(\alpha 4)_2(\beta 2)_3$ e pouca atividade agonista no receptor $(\alpha 4)_3(\beta 2)_2$.^{63,64} Especula-se que a

substância converte boa parte dos receptores $(\alpha 4)_2(\beta 2)_3$ altamente sensíveis para o estado dessensibilizado após exposição curta, sendo este o mecanismo responsável por seu longo efeito *in vivo*.⁶⁵ Há evidências de que parte da liberação dopaminérgica provocada por **16** ocorra devido a ativação de nAChRs do subtipo $\alpha 6\beta 2^*$.⁶⁶

1.2. Introdução à ancoragem molecular no contexto da Química Medicinal

A Química Medicinal é a área da química que lida com a descoberta, planejamento, desenvolvimento e caracterização analítica e farmacológica de novos candidatos à fármacos.⁶⁷ Os nAChRs, como mencionado anteriormente, estão relacionados à diversas desordens do sistema nervoso central e, conseqüentemente, são alvos muito interessantes para o desenvolvimento de novos agentes terapêuticos.

A descoberta de novos fármacos e formulações terapêuticas é um processo longo e dispendioso, estimando-se que um ciclo completo de descoberta de um novo fármaco, desde a identificação de um *hit* até o final dos testes clínicos, pode levar entre 10 e 15 anos⁶⁸ e custar mais de \$1,3 bilhões de dólares.⁶⁹ Na indústria farmacêutica, em diversos casos, esse processo é feito através da química combinatória, abordagem que permite a síntese rápida de centenas ou milhares de compostos, e *high-throughput screening* (HTS), onde uma grande biblioteca de compostos é testada frente à um alvo visando a obtenção de um *lead* que possa ser otimizado em etapas futuras do desenvolvimento do fármaco.^{70,71} Apesar dessas estratégias ainda serem amplamente utilizadas, pontos negativos como o alto investimento e risco de falha envolvidos no processo, em conjunto com os avanços em técnicas de espectroscopia biomolecular, cristalografia de raios-X, ressonância magnética nuclear e crio-microscopia eletrônica, tem popularizado métodos mais rápidos e racionais de desenvolvimento de fármacos.⁷⁰⁻⁷² Isso está diretamente relacionado à conclusão do projeto genoma humano e no conseqüente aumento no número de possíveis alvos terapêuticos conhecidos.⁷²⁻⁷⁴ Nesse contexto, uma das abordagens mais amplamente utilizadas no planejamento racional de fármacos, tanto na indústria quanto na academia, é o *structure-based drug design* (SBDD), que é a incorporação de dados sobre a estrutura do alvo desejado no processo de planejamento de fármacos.^{75,76} Um esquema representando as etapas de uma abordagem SBDD está representado na **Figura 8**.

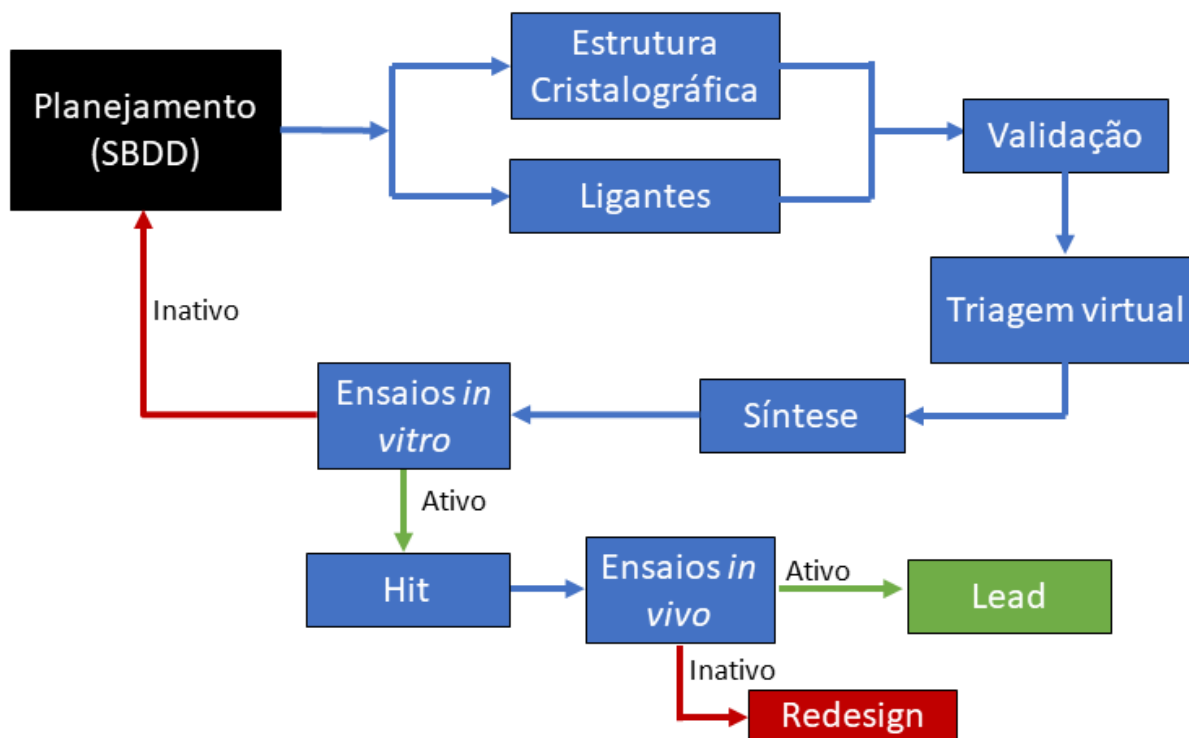


Figura 8 – Fluxograma das etapas presentes em uma abordagem de SBDD.

Dentro do campo do SBDD, a ancoragem molecular (*docking*) é atualmente a técnica mais utilizada.⁷⁷ Pode ser aplicada em fluxos de trabalho diversificados, como triagem virtual de grandes bancos de dados de moléculas, a otimização de compostos com afinidade pelo alvo de interesse já descrita e também auxiliar no entendimento das características do reconhecimento biomolecular a nível atômico. A ancoragem molecular é uma técnica que faz uso da estrutura tridimensional do alvo, geralmente uma proteína de interesse terapêutico, com o objetivo de prever as interações ligante-receptor através de métodos computacionais, em duas etapas correlacionadas: a amostragem das conformações do ligante dentro do sítio ativo e o ranqueamento das mesmas através de uma ou mais funções de pontuação, que estima a energia de ligação do complexo ligante-proteína.^{78,79} Idealmente, a etapa de amostragem deve ser capaz de buscar eficientemente por conformações dentro do espaço conformacional do ligante que se adequem ao perfil da superfície de energia livre do alvo, enquanto a função de pontuação deve ser capaz de associar corretamente as poses melhor ranqueadas com modos de ligação experimentalmente descritos para um par ligante-proteína, preservando interações com resíduos de interesse.^{78,80} O custo computacional decorrente da amostragem de todas as possíveis conformações de um ligante é imenso e maior ainda quando os graus de

liberdade do alvo também são considerados, de modo que é impraticável a avaliação de todas as combinações possíveis.⁷⁸⁻⁸⁰ Sendo assim, algumas abordagens de amostragem conformacional e de flexibilidade da proteína são comumente empregadas visando reduzir o tempo necessário para a conclusão do cálculo.

1.2.1. Amostragem conformacional dos ligantes na ancoragem molecular

Os métodos de amostragem conformacional se classificam em duas categorias: os métodos sistemáticos e os estocásticos. Na **Tabela 2** estão relacionados alguns programas de ancoragem, divididos pelo método de amostragem conformacional.

Tabela 2 - Exemplos de algoritmos de amostragem conformacional.

Sistemático		
Busca exaustiva	Glide ^{81,82}	
Fragmentação	FlexX ⁸³	Hammerhead ⁸⁴
<i>Ensemble</i> conformacional	FLOG ⁸⁵	EUDOC ⁸⁶
	MS-DOCK ⁸⁷	
Estocástico		
Monte Carlo	AutoDock ⁸⁸	ICM ⁸⁹
	QXP ⁹⁰	MCDOCK ⁹¹
<i>Tabu</i>	AutoDock Vina ⁹²	ROSETTALIGAND ⁹³
	PRO_LEADS ⁹⁴	PSI-DOCK ⁹⁵
Algoritmo evolucionário	GOLD ⁹⁶	AutoDock 3 & 4 ⁹⁷
	rDock ⁹⁸	
Otimização de <i>Swarm</i>	PLANTS ⁹⁹	SODOCK ¹⁰⁰
	PSOVina ¹⁰¹	

Métodos sistemáticos de amostragem conformacional associam uma série de valores discretos aos graus de liberdade do ligante e todos os valores de cada coordenada são explorados de maneira combinatória.¹⁰² O algoritmo busca dentro da superfície de energia do espaço conformacional e, após diversos ciclos, converge para uma solução de energia mínima. É possível que a solução encontrada seja um mínimo local e não o mínimo global, mas tal contratempo pode ser contornado através de buscas simultâneas que se iniciem a partir de conformações distintas do ligante.⁸⁰ Dentro dessa classe de algoritmos se englobam os de busca exaustiva, onde todas

as rotações livres do ligante são verificadas sistematicamente respeitando restrições e critérios de término predefinidas, os de fragmentação, onde o ligante é dividido em fragmentos que são ancorados em sequência resultando na construção incremental do ligante completo, e os algoritmos de *ensemble* conformacional, que ancoram uma série de conformações rígidas previamente geradas para um mesmo ligante, criando assim uma espécie de flexibilidade simplificada.

Os métodos estocásticos trabalham através de mudanças aleatórias nos graus de liberdade do ligante. Possuem o potencial de encontrar mínimos de energia rapidamente, porém também é possível que não encontrem a solução real devido à não realizarem uma busca completa do espaço conformacional, o que pode ser parcialmente solucionado através do aumento das iterações do algoritmo.⁷⁹ Dentro dessa classe se englobam os algoritmos do tipo Monte Carlo, onde a cada iteração do algoritmo uma mudança aleatória nos graus de liberdade é realizada e a solução resultante é aceita ou rejeitada com base em um critério estatístico de probabilidade e os métodos *Tabu*, que focam em evitar a amostragem de conformações já exploradas. Também nessa classe existem os algoritmos evolucionários, que aplicam operações matemáticas inspiradas em conceitos biológicos de reprodução para selecionar quais conformações estão mais aptas a se “reproduzirem” em cada ciclo e assim convergir à uma solução de menor energia, e os algoritmos de otimização de *Swarm* (enxame) que se baseiam na comunicação de enxames de insetos através de feromônios, onde cada feromônio corresponde à um grau de liberdade e insetos virtuais escolhem as melhores respostas com base nos valores desses feromônios.

1.2.2. Flexibilidade da proteína na ancoragem molecular

Quanto aos graus de liberdade do alvo, a flexibilidade da proteína costuma ser tratada de duas maneiras distintas: rígida ou flexível. Devido à capacidade computacional limitada a estratégia mais amplamente adotada é a utilização de um alvo rígido com ligante flexível, caracterizando ancoragem semi-rígida.^{78,79} Entretanto, existem evidências contundentes de que proteínas possuem alta flexibilidade e de que essa característica é de extrema importância para o reconhecimento molecular.^{103,104} Algumas abordagens para introduzir flexibilidade da proteína em estudos de ancoragem já foram exploradas e estão resumidas nos parágrafos abaixo.⁷⁹

Em primeiro lugar, se dividem as estratégias em dois grupos: (I) os que utilizam somente uma conformação do alvo e (II) os que utilizam múltiplas conformações. No

primeiro tipo de estratégia duas abordagens se destacam: (Ia) a utilização de potenciais leves (*soft docking*) e (Ib) a introdução de flexibilidade às cadeias laterais dos resíduos do sítio ativo. A primeira abordagem citada consiste na redução do termo de repulsão de van der Waals empregado pela função de pontuação, permitindo assim contatos mais próximos e conseqüentemente um complexo ligante-alvo mais compacto. A desvantagem dessa abordagem se encontra no fato do algoritmo aproximar somente movimentação mínima da proteína, o que pode resultar em poses irreais. Quanto à introdução de flexibilidade das cadeias laterais, a mesma ocorre, em geral, através da exploração de uma biblioteca de rotâmeros pré construída, mas a abordagem desconsidera completamente variações conformacionais do alvo.

Os métodos que utilizam múltiplas conformações do alvo podem ser empregados quando houver múltiplas estruturas resolvidas experimentalmente para o mesmo ou quando for possível a construção de um *ensemble* conformacional através de métodos computacionais. Dentro dessa classe de algoritmos se encaixam: os de *grid* médio, onde as estruturas do *ensemble* são usadas para a construção de uma “grade média” que pode ser do tipo simples ou com peso diferencial para cada proteína, e também os algoritmos de descrição unificada da proteína, onde as conformações são utilizadas na construção de proteínas quiméricas que obtenham a melhor performance possível. Também é possível considerar cada uma das proteínas do *ensemble* individualmente, avaliando o ligante de interesse em todas elas e buscando a melhor solução possível. Na **Figura 9** estão exemplificados dois dos métodos de flexibilidade da proteína mencionados acima.

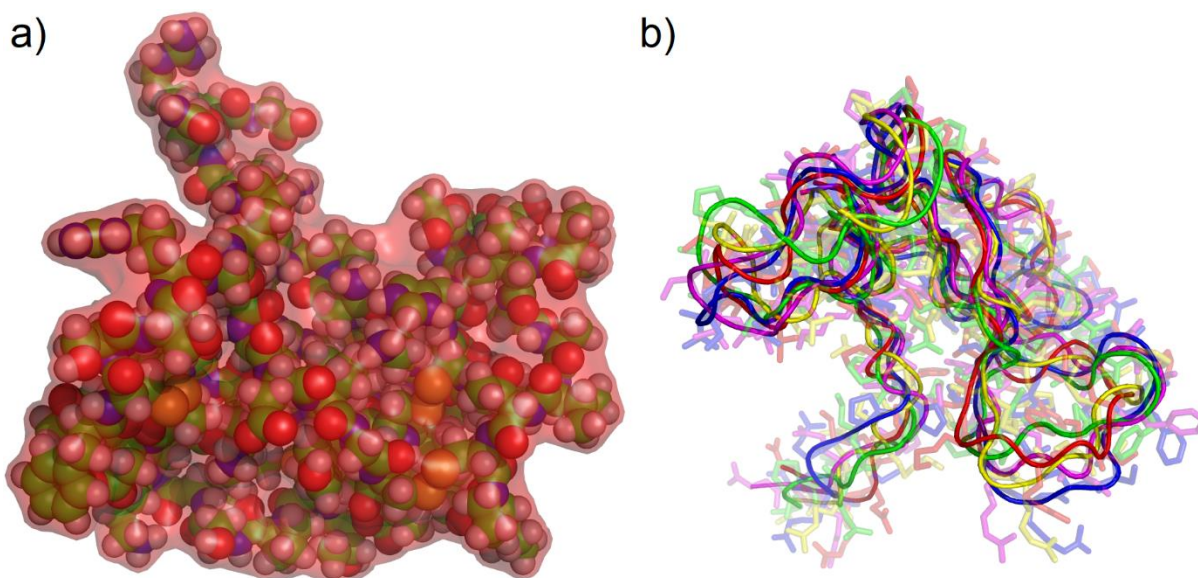


Figura 9 – Métodos de introdução de flexibilidade para proteínas em ancoragem molecular. a) potenciais leves (la), com a superfície em vermelho representando o tamanho original dos átomos e as esferas representando o tamanho corrigido após redução do termo de repulsão de van der Waals. Dessa forma, os átomos dos ligantes podem se aproximar mais dos resíduos de aminoácidos devido ao menor raio de van der Waals. b) exemplo de *ensemble* conformacional, onde cada uma das cores representa uma conformação experimentalmente descrita para uma proteína. PDB ID: 1IDL³⁵

1.2.3. Funções de pontuação

As funções de pontuação utilizadas atualmente para ancoragem molecular podem ser comumente classificadas em três categorias, de acordo com o método utilizado para construí-las: funções empíricas, funções de campo de força e funções baseadas em conhecimento (*knowledge-based*).¹⁰⁵ Algumas funções de pontuação e suas categorias estão relacionadas na **Tabela 3**.

Tabela 3 - Exemplos de funções de pontuação.

Campo de força	Empírica	Baseada em conhecimento
DOCK ¹⁰⁶	FlexX ⁸³	ITScore ^{107,108}
AutoDock ^{97,109}	GlideScore ⁸¹	PMF ¹¹⁰
GoldScore ⁹⁶	LUDI ¹¹¹	DrugScore ¹¹²
SYBYL/D-Score ¹¹³	PLP ¹¹⁴	DFIRE ¹¹⁵
SYBYL/G-Score ¹¹³	CHEMPLP ⁹⁹	SMoG ¹¹⁶
ICM ⁸⁹	ChemScore ¹¹⁷	BLEEP ^{118,119}
LigandFit ¹²⁰	SCORE ¹²¹	Mscore ¹²²
MedusaScore ¹²³	X-Score ¹²⁴	ASP ¹²⁵
	Surflex ¹²⁶	MotifScore ¹²⁷
	SYBYL/F-Score ⁸³	RFScore ¹²⁸
	LigScore ¹²⁹	PESD_SVM ¹³⁰
	AI-Score ¹³¹	PoseScore ¹³²
	SFCscore ¹³³	
	Fresno ¹³⁴	
	HYDE ¹³⁵	

As funções de pontuação baseadas em campo de força estimam a energia de ligação através de termos físicos para interações atômicas, incluindo termos eletrostáticos descritos por uma função coulombiana, interações de van der Waals tratadas por uma função potencial de Lennard-Jones e forças decorrentes de estiramento, torção e dobra das ligações químicas.^{72,78-80} Uma das vantagens desse tipo de função é que elas podem ser continuamente aprimoradas em conjunto com o estado da arte nos campos teóricos conforme novos campos de força, métodos quânticos e modelos de solvatação são desenvolvidos, mas devido à extrema complexidade do problema da predição de energia livre o sucesso tem sido modesto mesmo quando teorias de alto nível são aplicadas.¹⁰⁵ Parâmetros adicionais vem sendo incluídos para aprimorar as funções baseadas em campo de força, como termos de ligação de hidrogênio, solvatação e entrópicos, sendo que estes dois últimos se apresentam extremamente desafiadores no contexto atual.^{72,79} O termo entrópico de funções desse tipo geralmente é pouco exato, uma vez que ainda não se desenvolveu um modelo físico razoável para descrever essa propriedade.⁸⁰ O solvente pode ser tratado explicitamente, às custas da velocidade do cálculo, ou implicitamente através dos modelos de solvatação contínua como Poisson-Boltzmann (PB/SA) ou Born Generalizado (GB/SA). Métodos baseados em campo de força frequentemente necessitam de parâmetros empíricos para corrigir o peso dos termos da função devido

aos erros intrínsecos de cada termo de energia e também devido às diferentes dimensões de cada um desses termos.^{105,136} Tais métodos demandam mais tempo de cálculo devido à sua complexidade intrínseca.⁷⁸

Os métodos empíricos estimam a energia de ligação de um complexo através de funções matemáticas cujos termos são derivados de regressões lineares multivariadas. Consideram-se termos de van der Waals, eletrostáticos, ligação de hidrogênio, entropia e hidrofobicidade, entre outros, de modo que cada função desse tipo difere das demais quanto ao número e natureza de tais termos.^{72,80} Devido à simplicidade dos termos, funções de pontuação empíricas são consideravelmente mais rápidas do que as de campo de força.⁸⁰ Essas funções podem ser convenientemente ajustadas através da adição, remoção ou modificação dos termos de regressão de maneira mais funcional e intuitiva do que os termos de uma função baseada em campo de força, permitindo assim customizações para a solução de problemas mais específicos. Entretanto, frequentemente padrões de interação ligante-proteína específicos são ignorados por não serem significantes na análise de regressão, o que dificulta a descrição de todos os possíveis fatores envolvidos no reconhecimento molecular.¹⁰⁵ O crescimento exponencial de complexos ligante-proteína solucionados experimentalmente possibilita a utilização de grandes grupos de treinamento para o modelo de regressão, minimizando as preocupações iniciais quanto à validade de funções empíricas para sistemas diferentes do conjunto de treinamento.¹³⁶ Por serem empíricas, a qualidade desse tipo de função depende grandemente da qualidade e consistência dos dados de afinidade reportados na literatura científica, sendo este o maior entrave na construção de funções mais eficientes.¹⁰⁵

As funções de pontuação baseadas em conhecimento são construídas através da análise da frequência de ocorrência de diversos pares de átomos em estruturas cristalográficas experimentalmente resolvidas, sendo aqui definidos pares de átomos como quaisquer par de átomos dentro de uma distância limite um do outro que se considere representativo de uma interação ligante-proteína, considerando o ambiente químico onde se encontra tal par. Aplica-se a inversa da função de Boltzmann para correlacionar a frequência de ocorrência com o potencial de força média para os pares (*pairwise potential, potential of mean force*) e a pontuação obtida é a somatória de todas as contribuições dos pares de átomos observados entre o ligante e a proteína. Assim, obtém-se termos que recompensam e penalizam contatos com base nos

potenciais calculados, que são função das distâncias interatômicas entre eles. A premissa dessa abordagem se baseia na hipótese de que contatos observados com mais frequência do que em uma distribuição aleatória (distribuição normal) devem representar interações energeticamente favoráveis, enquanto contatos surgindo com menos frequência do que em uma distribuição normal representam interações desfavoráveis. Considera-se a situação de distribuição normal aquela na qual as interações interatômicas entre o par são nulas, caracterizando o estado de referência. Funções baseadas em conhecimento apresentam boa relação entre precisão e velocidade devido à simplicidade dos termos utilizados, uma vez que buscam capturar implicitamente os fatores energéticos do reconhecimento molecular através dos potenciais em pares. O fato de os potenciais serem extraídos diretamente das estruturas tridimensionais remove a necessidade de correlacionar dados experimentais de afinidade e devido ao banco de dados poder ser grande e diversificado essas funções são robustas e grandemente insensíveis ao grupo de treinamento utilizado.^{105,136} Dadas as bases teóricas do método, se entende que esse tipo de função é mais interessante para predição de poses do que para estimativas de afinidade ligante-proteína, mas essa correlação com a afinidade experimental vem sendo aplicadas com precisão comparável ou melhor do que a observada para outras abordagens e, em muitos casos, se introduzem termos empíricos que elevam os métodos baseados em conhecimento à métodos híbridos.¹⁰⁵ Um dos desafios relacionados à construção dessas funções é a definição do estado de referência, o estado onde a interação entre os pares é inexistente, uma vez que não existe sistema físico real onde essa condição seja observada e sendo então necessárias aproximações que podem comprometer a qualidade do resultado.¹³⁶

Recentemente, o advento de métodos de inteligência artificial resultou em uma nova classe de funções de pontuação baseadas em descritores, também chamadas de funções de pontuação baseada em aprendizado de máquina.^{105,136} Nessa classe, as funções são derivadas através da utilização de descritores de características químicas, como ligações de hidrogênio, interações eletrostáticas, empilhamentos aromáticos, descritores geométricos, entre outros, e os mesmos são trabalhados por um algoritmo de aprendizagem de máquina como a floresta aleatória, redes neurais, classificadores Bayesianos e máquinas de vetores de suporte, de modo que é possível construir modelos de regressão estatística através da seleção dos descritores inicialmente introduzidos. São similares às funções empíricas no sentido de

necessitarem de grupos de treinamento, consistindo de estruturas tridimensionais conhecidas e dados de afinidade, em sua elaboração. A grande diferença entre essa classe de algoritmos e as funções de pontuação clássicas é a abordagem não linear empregada, que é adequada para a modelagem de sistemas altamente complexos como o processo de reconhecimento molecular. Outra diferença marcante é que as funções clássicas têm formas funcionais inspiradas em teorias bem estabelecidas e definidas por humanos, enquanto os métodos baseados em descritores dependem exclusivamente da máquina para determinar a forma final do modelo.¹³⁷ Essa última característica pode ser um entrave na interpretação dos dados: os termos individuais utilizados nas funções clássicas normalmente possuem significados físicos interpretáveis enquanto o raciocínio por trás da seleção de variáveis realizada pela máquina é vago e difícil de correlacionar com termos físicos, uma vez que ela se retroalimenta a cada iteração.¹⁰⁵ Estudos mostram que as funções de pontuação baseadas em aprendizagem de máquina têm performance superior em relação às funções de pontuação discutidas anteriormente (clássicas) no quesito predição de afinidade ligante-proteína. As perspectivas para o desenvolvimento dessa abordagem são promissoras, dado o crescente volume de informação disponível atualmente e o interesse crescente no assunto, com cada vez mais pesquisadores aplicando e desenvolvendo metodologias baseadas em aprendizagem de máquina com extremo cuidado e competência. Entretanto, estudos adicionais se fazem necessários para confirmar ou refutar definitivamente se o aprendizado de máquina é de fato mais eficiente do que as funções de pontuações clássicas em quaisquer instâncias.^{137,138}

1.2.4. Validação de modelos de ancoragem molecular

Apesar da implementação de uma abordagem de ancoragem molecular ser muito prática, é necessário avaliar se as condições utilizadas no procedimento são adequadas para os fins desejados através de uma etapa de validação. Em geral, a validação de um protocolo de ancoragem se baseia em três abordagens distintas: (A) a redocagem (*redocking*), (B) curva característica de operação do receptor (*receiver-operating characteristic curve*, ROC) e (C) relação estrutura-atividade quantitativa tridimensional (*tridimensional quantitative structure-activity relationship*, 3D-QSAR).¹³⁹ Na primeira abordagem o ligante nativo da proteína a ser trabalhada é removido do sítio ativo e se pede ao algoritmo que calcule a orientação do ligante no sítio utilizando as condições à serem avaliadas para a ancoragem molecular. O sucesso da

abordagem se dá caso haja concordância entre o modo de ligação experimental frente ao calculado. Essa abordagem vem sendo amplamente utilizada, mas ignora o fato de que, para a maioria dos experimentos de ancoragem, o interesse está em avaliar ligantes distintos do ligante nativo da proteína, caracterizando ancoragem cruzada (*cross-docking*). Sabe-se que a precisão dos dados obtidos em uma ancoragem cruzada é pequena, o que reduz a confiabilidade dessa estratégia como método de validação.¹⁴⁰ Na **Figura 10** se encontra exemplificado um exemplo de redocagem bem sucedido.

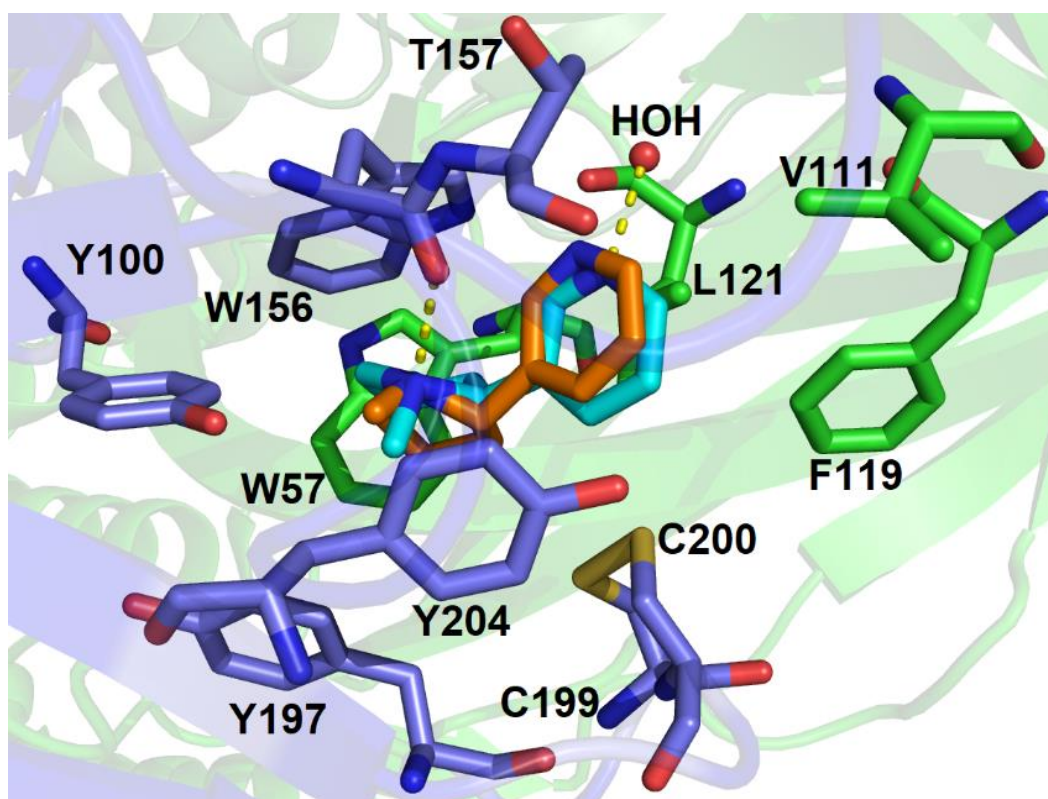


Figura 10 – Nicotina co-cristalizada na estrutura do nAChRs do subtipo $\alpha 4\beta 2$ (laranja, PDB ID: 5KXI) e pose de ancoragem bem-sucedida na avaliação por redocagem. Átomos de hidrogênio foram omitidos para maior clareza.

A validação por curva ROC avalia a capacidade do algoritmo de ancoragem de discernir entre ligantes ativos e inativos contidos dentro de um banco de dados, de modo que se obtém através da plotagem da curva dados referentes à, entre outras coisas, falsos positivos, falsos negativos e sobre ligantes corretamente classificados pelo modelo. Geralmente, tais bancos de dados são compostos por um grande número de ligantes inativos (*decoys*) e um número muito menor de ligantes ativos.¹⁴¹ Considera-se o protocolo de ancoragem válido caso os ligantes utilizados sejam

corretamente classificados como ativos e inativos, podendo este critério ser flexibilizado para considerar apenas regiões de interesse da curva. A desvantagem da abordagem está relacionada à prerrogativa de que os ligantes definidos como inativos são, de fato, inativos para o alvo em questão mesmo sem dados experimentais corroborando tal afirmação.¹⁴² Uma curva ROC hipotética está representada na

Figura 11.

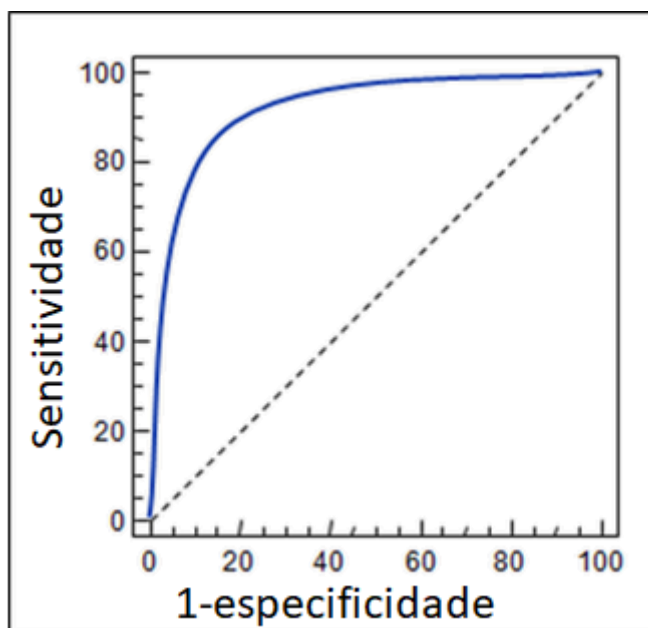


Figura 11 – Exemplo de uma curva ROC hipotética.

Quanto à validação por 3D/QSAR, se considera o protocolo de ancoragem válido caso seja bem sucedido quanto ao alinhamento de uma série de ligantes com bioatividade descrita para o alvo, de maneira que seja possível discriminar fatores que expliquem as variações de bioatividade observada para os mesmos através de técnicas computacionais estatísticas, como a análise conformacional do campo molecular (*conformational molecular field analysis*, CoMFA).¹⁴³ A desvantagem da aplicação de CoMFA para validação de estudos de ancoragem é a elevada complexidade e demanda computacional necessárias para a implementação da metodologia. Além disso, o alinhamento entre os ligantes ativos é uma etapa crítica da abordagem CoMFA e isso pode vir a ser um problema em casos onde as moléculas ancoradas não apresentem posicionamento conciso. Recentemente uma nova abordagem de 3D/QSAR para a validação de protocolos de ancoragem foi proposta, a análise comparativa de contatos intermoleculares baseada em ancoragem (*docking-*

based comparative intermolecular contacts analysis, dbCICA)¹⁴⁴ e a mesma será discutida na próxima seção.

1.2.5. Análise comparativa de contatos intermoleculares baseada em ancoragem molecular (dbCICA)

Como mencionado anteriormente, a análise comparativa de contatos intermoleculares baseada em ancoragem molecular é uma metodologia de pós processamento de resultados de ancoragem molecular que se baseia em 3D/QSAR e é implementada através de um algoritmo genético. A análise dbCICA se baseia na quantidade e qualidade dos contatos entre ligantes ancorados no receptor e os resíduos de aminoácidos dentro do sítio ativo, avaliando se uma configuração de ancoragem qualquer é capaz de alinhar os ligantes de tal forma que ligantes mais ativos participem de contatos com átomos distintos dos contatos observados para ligantes de menor atividade e vice-versa. Dessa forma, caso tal conjunto de contatos seja corretamente identificado para uma série de ligantes ativos ancorados é possível afirmar que a condição de ancoragem é válida, pois é capaz de posicionar as moléculas de forma que a variação observada na afinidade seja explicada. A análise se sustenta dada a constatação de que, em geral, funções de pontuação são bem-sucedidas na reprodução de modos de ligação experimentalmente elucidados durante experimentos de redocagem, o que sugere que seja possível correlacionar poses ancoradas com a bioatividade. A alta afinidade de um ligante frente à um alvo é relacionada a capacidade do mesmo de satisfazer certas interações críticas dentro do sítio ativo. Dessa forma, a identificação de tais contatos críticos não só auxilia na validação do protocolo de ancoragem como também pode ser utilizada para a elaboração de mapas farmacofóricos contendo as características químicas/estéricas necessárias para o reconhecimento molecular.¹⁴⁴

Nessa abordagem, se define um contato intermolecular com base na distância entre os átomos dos ligantes e do sítio ativo caso ela seja menor do que um limite predefinido. As distâncias intermoleculares são trabalhadas considerando todos os hidrogênios do ligante e do alvo explicitamente. Assume-se que as distâncias interatômicas entre átomos dos ligantes e da proteína sejam boas representações da interação atrativa que resulta em tal proximidade. O algoritmo dbCICA busca por uma somatória de tais contatos que seja diretamente proporcional a afinidade dos ligantes, permitindo peso diferencial para contatos positivamente correlacionados com a

afinidade, caracterizando assim contatos de maior relevância para o reconhecimento molecular. Espera-se que alguns contatos apresentem correlação inversamente proporcional à afinidade, caracterizando a presença de repulsões estéricas. Tais contatos são removidos da etapa inicial do pós processamento dos dados, onde se busca apenas a identificação da melhor somatória de contatos positivos, e reintroduzidos na etapa subsequente após um tratamento matemático simples. A avaliação dos contatos somados pelo algoritmo genético pode ser feita através de três métricas, o coeficiente de correlação (r^2), *leave-one-out* r^2 (r^2_{LOO}) ou *K-fold* r^2 ($r^2_{\text{K-fold}}$).

A metodologia dbCICA já foi aplicada com sucesso diversas vezes por TAHA *et. al.* para alvos distintos, resultando em moléculas com afinidade experimental confirmada.^{144–151} Através da modelagem por dbCICA os autores originais foram capazes de identificar novas entidades químicas com atividade frente à enzima N-myristoyl transferase da espécie *Candida albicans* (CaNMT),¹⁴⁴ glicogênio fosforilase da espécie humana (GP),¹⁴⁴ proteína de choque térmico 90 da espécie humana (Hsp90),¹⁴⁵ glucoquinase da espécie humana (GK),¹⁴⁶ *checkpoint kinase 1* da espécie humana (Chk1),¹⁴⁷ fator de crescimento fibroblástico 1 da espécie humana (FGFR1),¹⁴⁸ acetilcolinesterase humana (AChE),¹⁴⁹ *3C-like protease* do coronavírus da síndrome respiratória do oriente médio (HKU4-CoV 3CL^{pro})¹⁵⁰ e para a DNA-girase β .¹⁵¹ O número de estudos bem sucedidos ilustra a aplicabilidade do método à diversos sistemas não relacionados, reforçando sua utilidade como método de validação de resultados de ancoragem molecular.

2. Justificativa e Objetivos

Os efeitos de dependência ocasionados pela nicotina (**3**) são produzidos, em parte, através da sua ação como agonista no subtipo $\alpha 4\beta 2$ dos nAChRs.⁵⁶ Estudos combinando agonistas e antagonistas demonstraram que um efeito agonista parcial é o ideal para o tratamento da dependência quando um indivíduo para de fumar, visto que por meio desta abordagem terapêutica o indivíduo teria uma diminuição gradual da modulação do nAChR $\alpha 4\beta 2$ e da consequente ativação dopaminérgica. Por exemplo, o produto natural (-)-citisina (**17**), que é encontrado em numerosas espécies de plantas, tem atividade como agonista parcial do subtipo $\alpha 4\beta 2$ de nAChRs, mas apresenta baixa absorção, de modo a dificultar sua aplicação como medicamento.¹⁵² Alguns anos mais tarde, foram publicados resultados referentes ao fármaco vareniclina (Champix®)⁴⁸ (**13**), agonista parcial com elevada afinidade por nAChRs do subtipo $\alpha 4\beta 2$, que inibe a ativação dopaminérgica produzida pela nicotina enquanto alivia o desejo e os efeitos associados às crises de abstinência quando um indivíduo para de fumar.⁵⁶ A vareniclina (**13**) foi aprovada para ser comercializada como fármaco, sob a forma de tartarato, nos EUA e na Europa pela empresa farmacêutica Pfizer® em maio de 2006 para o tratamento da dependência da nicotina,⁴⁸ porém o tratamento ainda apresenta custo elevado¹⁵³

Na **Figura 12**, além da nicotina (**3**), da citisina (**17**) e do fármaco vareniclina (**13**), podemos observar os derivados **18-20**, que mostram o quanto é frequente nos ligantes de nAChRs do subtipo $\alpha 4\beta 2$ ao menos uma subunidade aromática (em azul) e átomos de nitrogênio básicos (em vermelho), geralmente em subunidades cíclicas. Também podemos constatar, observando os derivados **18-20**, que a subunidade (aza)bibílica presente em **13** não é essencial para observarmos alta afinidade frente aos nAChRs $\alpha 4\beta 2$, assim como para observarmos seletividade ao compararmos os nAChRs $\alpha 7$ e $\alpha 4\beta 2$.^{56,154}

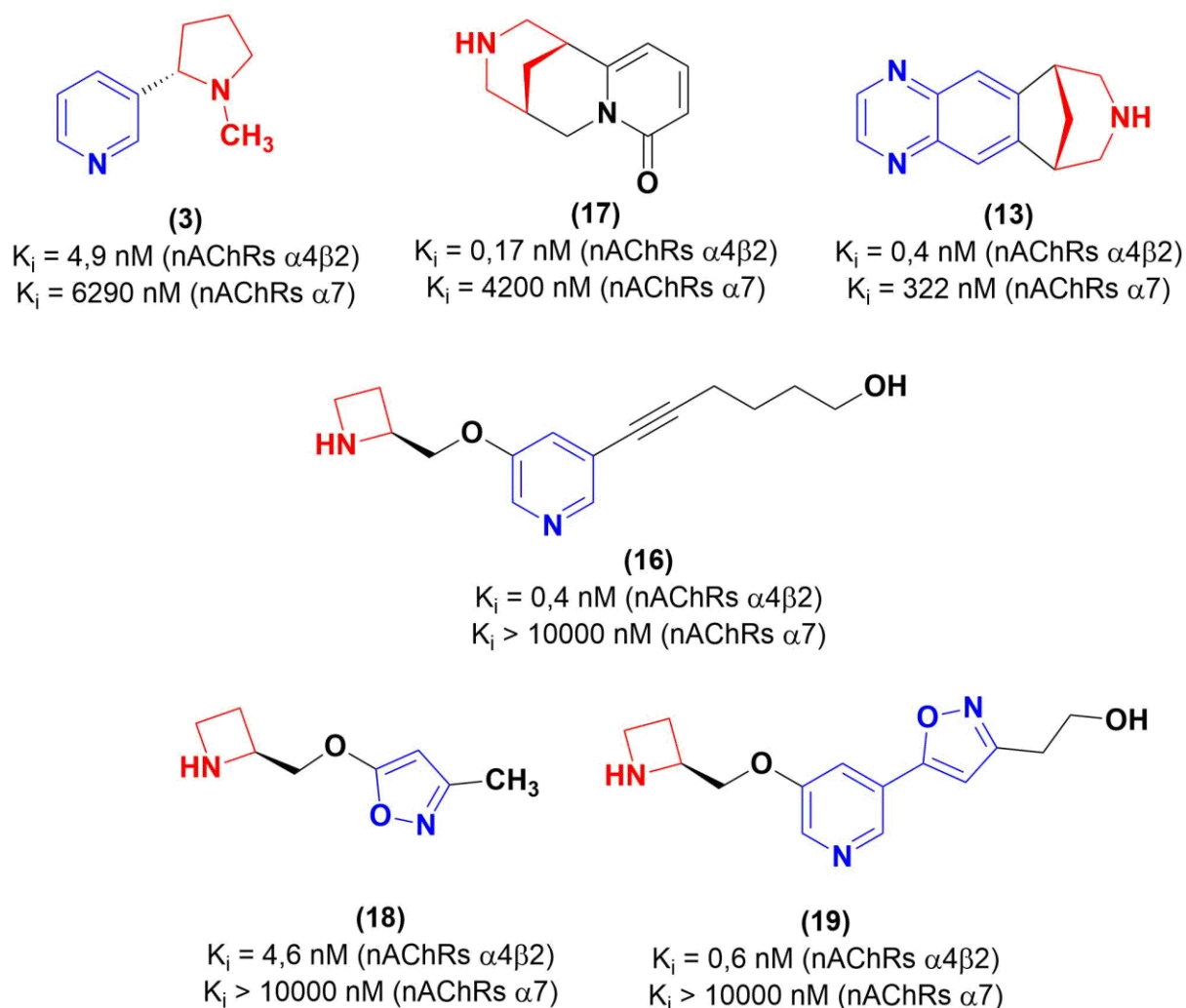


Figura 12 - Compostos heterocíclicos com atividade frente aos nAChRs $\alpha 4\beta 2$ e $\alpha 7$, com destaque para os produtos naturais (-)-nicotina (3), (-)-citisina (17) e o fármaco vareniclina (13).

Ao observarmos os aspectos socioeconômicos relacionados a esta doença, há alguns que merecem destaque. Estima-se que 1,3 bilhão de pessoas no mundo sejam fumantes, dentre as quais apenas 10% conseguem parar de fumar sem psicoterapia ou auxílio de medicamentos. Além disso, em 2015 o número anual de óbitos relacionados ao tabaco no mundo foi maior do que 6 milhões, sendo esses relacionados a doenças como câncer de pulmão, doença pulmonar obstrutiva crônica (COPD) e doenças cardiovasculares.¹⁵⁵ No Brasil, a dependência da nicotina afeta ca. 16% dos brasileiros e a mortalidade anual relacionada ao tabaco é de aproximadamente 200 mil pessoas. Quanto ao impacto nos cofres públicos, o Sistema Único de Saúde (SUS) tem prejuízo anual de R\$ 56,9 bilhões, sendo R\$ 39,4 bilhões gastos com despesas médicas e R\$ 17,5 bilhões com custos indiretos relacionados à

perda de produtividade causada por incapacitação de trabalhadores ou morte prematura.¹⁵⁶

Neste contexto, a busca de novos candidatos a fármacos pode vir a contribuir no desenvolvimento de novos tratamentos mais acessíveis à população. Dessa forma, vários grupos de pesquisa têm estudado melhor a fisiopatologia da dependência da nicotina e buscado novas substâncias agonistas parciais do subtipo $\alpha 4\beta 2$ dos nAChRs, visando ampliar o arsenal terapêutico para o tratamento dessa doença, assim como melhorias nas propriedades farmacocinéticas e diminuição de potenciais efeitos colaterais.^{154,157-159}

Dessa forma, o trabalho aqui descrito visa a obtenção de substâncias de padrão estrutural heterocíclico inéditas com atividade agonista parcial frente aos receptores nicotínicos de acetilcolina (nAChRs) do subtipo $\alpha 4\beta 2$, que possam ser úteis na descoberta de novas alternativas para o tratamento da dependência de nicotina, assim como outras doenças relacionadas a modulação deste biorreceptor, como o Alzheimer e o mal de Parkinson. Para tal, podemos descrever as metas específicas a seguir:

- Realização de estudos computacionais entre a estrutura cristalográfica do nAChRs $\alpha 4\beta 2$ (PDB ID: 5KXI) e ligantes da literatura para validação e compreensão de características estruturais relevantes dos receptores nicotínicos de acetilcolina;
- Utilização de modelos computacionais validados para a triagem virtual (*in silico*) dos compostos inéditos mais promissores para a etapa de síntese tendo em mente o alvo citado;
- Realização de estudos computacionais entre a estrutura cristalográfica do nAChRs $\alpha 7$ (PDB ID: 5AFH) e ligantes da literatura para validação e compreensão de características estruturais relevantes destes receptores nicotínicos de acetilcolina;
- Verificação da possibilidade de seletividade das moléculas planejadas para o subtipo $\alpha 4\beta 2$ através da comparação dos modelos computacionais construídos para ambos os nAChRs.

3. Metodologia

3.1. Hardware utilizado

Todos os cálculos realizados durante este trabalho foram realizados em uma estação de trabalho com sistema operacional Windows 10, processador Intel® Core™ i7-8700, CPU de 3,20 GHz, base 64 bits (x64) e 16 GB de memória RAM da marca DELL.

3.2. Planejamento estrutural

O planejamento estrutural desta classe de substâncias se originou na estrutura química do fármaco vareniclina (**13**), o qual apresenta elevada afinidade e alta seletividade pelos nAChRs $\alpha 4\beta 2$. Inicialmente, a subunidade heterobocíclica em **13** pode ser modificada através da estratégia de simplificação molecular, para originar as **Séries A e B (Figura 13)**, onde a subunidade benzimidazólica observada na **Série B** pode ser compreendida a partir da estratégia de modificação estrutural de bioisosterismo de anéis.¹⁶⁰ Também são propostos substituintes alifáticos de forma homóloga (**20-23, a-e**), seguido das variações em **20f-23f** e **20g-23g**, de modo a avaliarmos até que ponto o volume e superfície moleculares podem contribuir para as interações frente aos nAChRs $\alpha 4\beta 2$. Para os fragmentos **20-23 (h-zc)**, diversas subunidades aromáticas são exploradas, a não substituída e a substituída em diferentes posições (orto, meta e para), explorando os efeitos estérico e eletrônico em diferentes posições e com diferentes características (eletron-doadores e -retiradores), de modo a avaliarmos até que ponto estes substituintes podem contribuir para o reconhecimento molecular. A importância da restrição conformacional também é avaliada ao compararmos os fragmentos moleculares **20a-zc** e **22a-zc** com os fragmentos **21a-zc** e **23a-zc**, respectivamente. Estes fragmentos foram triados *in silico* por meio da estratégia de triagem virtual baseada em fragmentos moleculares (FBVS, *fragment based virtual screening*), onde os 8 melhores fragmentos foram eleitos para a etapa de síntese.^{161,162}

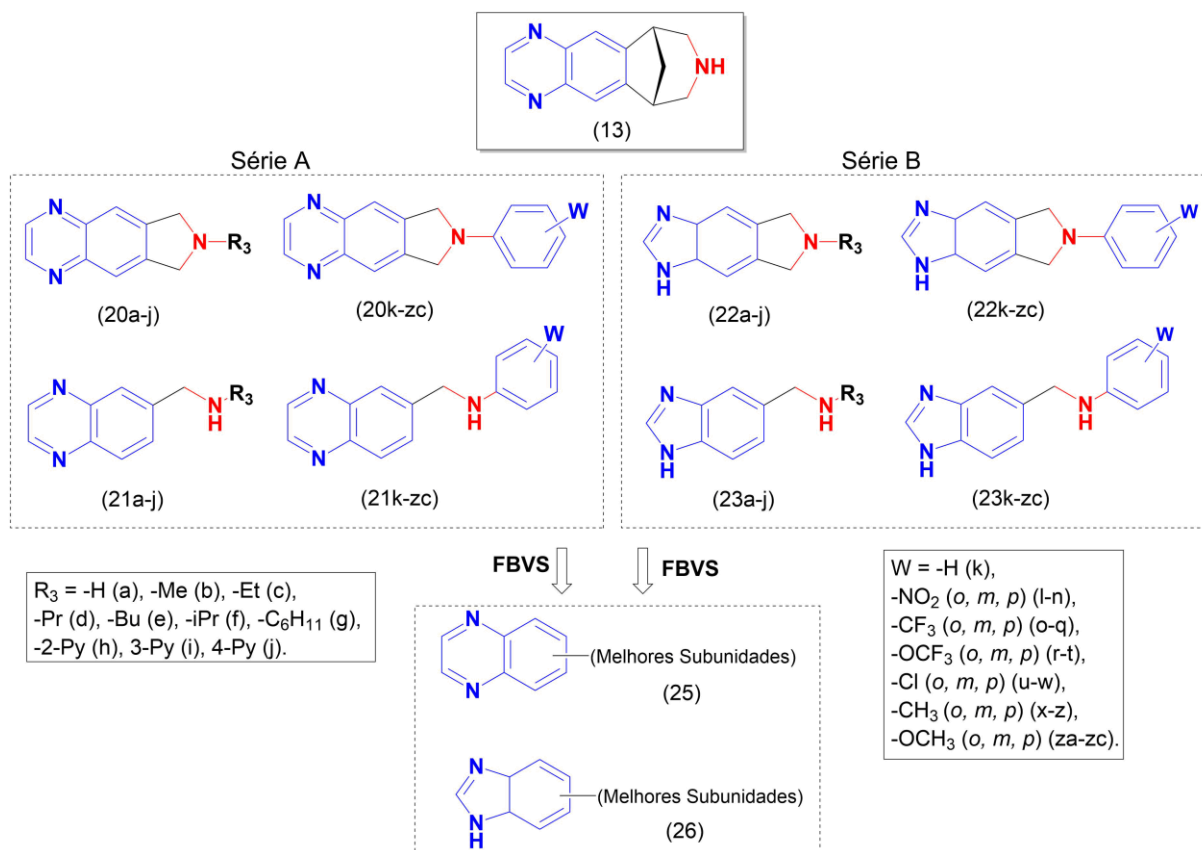


Figura 13 - Planejamento estrutural dos 116 fragmentos moleculares (21-24, a-zc), com padrões estruturais quinoxalínico e benzimidazólico, planejados como ligantes dos nAChRs $\alpha 4\beta 2$ e inspirados na vareniclina (**13**).

3.3. Obtenção e preparação das estruturas tridimensionais dos nAChRs e dos ligantes usados

Para a etapa de triagem virtual, utilizou-se uma metodologia baseada em ancoragem molecular com o intuito de selecionar os 8 fragmentos mais promissores para a etapa de síntese. A validação da metodologia citada se deu através da análise comparativa de contatos intermoleculares baseada em ancoragem molecular (dbCICA).¹⁴⁴ Inicialmente, se obteve a estrutura cristalográfica do nAChRs do subtipo $(\alpha 4)_2(\beta 2)_3$ do banco de dados Protein Data Bank (PDB ID: 5KXI, resolução 3,9 Å).^{29,163} O ligante co-cristalizado no sítio ativo é a nicotina (**3**). Obteve-se também a estrutura cristalográfica do nAChRs do subtipo $\alpha 7$ (PDB ID: 5AFH, resolução 2,4 Å)²⁷ para a construção do modelo de comparação. A estrutura 5AFH é uma proteína quimérica entre a AChBP da espécie *Lymnaea stagnalis* e o domínio extracelular do nAChRs $\alpha 7$ humano, com 71% de similaridade perante à proteína humana nativa. O ligante co-cristalizado em 5AFH é a lobelina (**7**).

Em seguida, realizou-se um levantamento bibliográfico de moléculas bioativas frente aos receptores alvo com o intuito de utilizá-las na etapa de validação dos modelos computacionais. Durante a seleção dessas moléculas atentou-se aos métodos utilizados para a realização dos ensaios farmacológicos de modo a selecionar somente moléculas cuja afinidade (K_i) foi medida através do mesmo protocolo farmacológico. Somente moléculas cuja estereoquímica absoluta foi bem definida pelos autores originais foram utilizadas para a construção dos modelos do subtipo $(\alpha 4)_2(\beta 2)_3$. Dessa forma, foi possível obter 98 moléculas com K_i entre 0,04 e 2176 nM para comporem o grupo de treinamento do modelo do subtipo $(\alpha 4)_2(\beta 2)_3$. Entre esses 98 ligantes estão incluídos a nicotina (**3**) e a vareniclina (**13**). Para a construção do modelo do subtipo $\alpha 7$ foi necessária a utilização de pares de enantiômeros para algumas moléculas a fim de aumentar o conjunto de treinamento, sendo que o enantiômero selecionado para a construção do modelo foi aquele que obteve maior pontuação de ancoragem. Assim, foi possível obter 93 moléculas com K_i entre 5,4 nM e 2180 μ M para comporem o grupo de treinamento do modelo desse subtipo. Os valores de K_i foram convertidos para pK_i ($-\log(K_i)$) para correlacionar linearmente a bioatividade dos ligantes com a variação da energia livre.¹⁴⁴ Estruturas químicas e K_i para as moléculas utilizadas na construção de ambos os modelos estão contidas nos **Anexos IA**^{164–171} e **IB**^{172–180}

As estruturas tridimensionais dos ligantes da literatura e dos ligantes planejados para este projeto foram construídas *in silico*, com todos os átomos explícitos, utilizando o software Discovery Studio Visualizer (v.17.0.2.1076).¹⁸¹ Os estados de protonação das moléculas foram determinados pelo software MarvinSketch (v.12.2.27.0)¹⁸² para pH 7,4. Os ligantes definidos como protonados no pH citado foram construídos nas duas possíveis formas enantioméricas em relação ao nitrogênio básico, quando aplicável. As estruturas foram minimizadas energeticamente pelo método semi-empírico PM7 através da interface Mercury CSD (v.3.9)¹⁸³ utilizando o MOPAC2016¹⁸⁴ e foram salvas no formato *.mol2 para serem utilizadas na ancoragem molecular.

3.4. Ancoragem molecular

A ancoragem molecular foi realizada através do software GOLD Suite (v.5.4).⁹⁶ As estruturas cristalográficas dos receptores foram preparadas para o cálculo na interface do software citado pela remoção de ligantes de co-cristalização e adição dos

átomos de hidrogênio. O centro do raio do sítio ativo para o modelo $(\alpha 4)_2(\beta 2)_3$ foi definido nas coordenadas $x = 67,3680$, $y = -27,2034$ e $z = -39,0781$, sendo essa a posição do átomo de nitrogênio pirrolidínico de **1**. Para o modelo $\alpha 7$ o centro da região de busca foi definido nas coordenadas $x = -20,8987$, $y = -10,0015$ e $z = 8,5084$, sendo essa a posição do átomo de nitrogênio básico de **7**. O raio escolhido para ambos os modelos foi de 10 Å e a ancoragem molecular foi realizada no modo rígido, com 50 poses de ancoragem para cada ligante. Para o modelo $(\alpha 4)_2(\beta 2)_3$, uma molécula de água foi modelada na região do sítio ativo nas coordenadas $x = 69,3500$, $y = -22,8900$ e $z = -42,4450$ em acordo com trabalhos publicados na literatura.^{185,186} Quanto ao modelo do subtipo $\alpha 7$, a água foi modelada nas coordenadas $x = -17,5010$, $y = -12,1430$ e $z = 5,4700$ correspondendo à posição de uma molécula de água co-cristalizada na região do sítio ativo trabalhado. Permitiu-se que a molécula de água tivesse rotação livre e movimento translacional de até 2 Å nos eixos x , y e z em ambos os modelos. Quanto à flexibilidade dos ligantes, além da rotação livre para ligações simples permitiu-se que o software invertesse em 180° aminas e amidas planares, bem como ácidos carboxílicos, e foram permitidas variações nas conformações de anéis não aromáticos. Para a construção dos modelos dbCICA foram utilizadas as formas enantioméricas de maior pontuação nos casos onde havia duas possibilidades de ionização do nitrogênio básico e, no caso do modelo $\alpha 7$, enantiômeros de carbono também foram selecionados de maneira análoga, quando aplicável.

3.5. Validação por dbCICA

A validação por dbCICA foi realizada como segue: as 98 estruturas coletadas da literatura científica foram ancoradas no nAChRs $(\alpha 4)_2(\beta 2)_3$ em suas formas protonadas e desprotonada, na presença e na ausência de uma molécula de água no sítio ativo, e utilizando as quatro funções de pontuação do programa GOLD: ASP, ChemScore, GoldScore e CHEMPLP, totalizando 16 corridas de ancoragem. As mesmas condições citadas acima foram utilizadas para o ancoramento das 93 moléculas referentes ao modelo do subtipo $\alpha 7$. Um esquema ilustrando como foram realizadas as ancoragens moleculares para cada função de pontuação está contido na **Figura 14**.

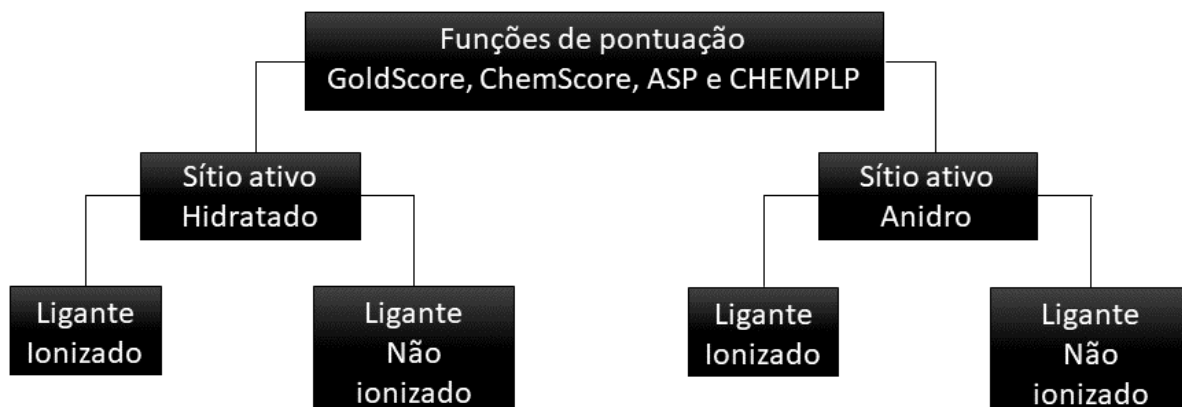


Figura 14 - Corridas de docking realizadas para cada uma das 4 funções de pontuação do software GOLD.

Em seguida se coletam as distâncias entre os átomos das poses melhor ranqueadas e do sítio ativo, considerando todos os hidrogênios explícitos. Caso tais distâncias fossem menores do que um limite predeterminado, sendo ele 2,5 ou 3,5 Å, caracteriza-se a presença de um contato intermolecular que recebe codificação 1 e caso elas fossem maiores não se considera que há contato e se atribui a codificação 0. Dessa forma, gera-se uma matriz binária cujas linhas representam poses dos ligantes e as colunas átomos do sítio ativo, sendo então que células com valor 1 correspondem à contatos entre as poses e os átomos do sítio ativo e células com valor 0 à ausência desses contatos. As distâncias coletadas e matrizes geradas estão exemplificadas na **Figura 15**.

E:ALA201:H - :9:H228	3.05909	Ligante	pKi	E:ALA201:CA	E:ALA201:CB	E:ALA201:H	E:ALA201:HA
E:ALA201:H - :9:H230	2.33556	:1	8,30980392	0	0	0	0
E:ALA201:HA - :13:O24	4.33778	:10	7,289036881	0	1	1	0
E:ALA201:HA - :13:C25	4.69905	:11	8,468521083	0	0	0	0
E:ALA201:HA - :13:H303	3.98847	:12	9	0	0	0	0
E:ALA201:HA - :13:H306	4.91796	:13	10	1	1	1	1
E:ALA201:HA - :14:H91	4.42245	:14	8,537602002	1	1	1	1
E:ALA201:HA - :14:H92	4.4734	:15	9,397940009	1	1	1	0
E:ALA201:HA - :14:N30	4.70623	:16	9,698970004	0	1	1	0
E:ALA201:HA - :14:C48	4.58351	:17	9,698970004	0	1	1	0
E:ALA201:HA - :14:H41	4.89771	:18	9,22184875	0	0	1	0
E:ALA201:HA - :14:H42	3.73303	:19	9,15490196	0	1	1	0
E:ALA201:HA - :14:H43	4.82344	:2	9,397940009	0	0	0	0
E:ALA201:HA - :15:O24	4.31309	:20	7,632644079	0	1	1	0
E:ALA201:HA - :16:O96	4.0917	:21	7,882728704	0	1	1	0
E:ALA201:HA - :16:C48	4.28785	:22	7,549750892	0	1	1	0
E:ALA201:HA - :16:H49	4.72093	:23	6,850780887	0	1	1	0
E:ALA201:HA - :16:H50	4.53201	:24	6,527243551	0	1	1	0
E:ALA201:HA - :16:H51	3.22088	:25	7,273272791	0	1	1	0
E:ALA201:HA - :17:O96	4.44986	:26	7,557520231	0	1	1	0
E:ALA201:HA - :17:H52	4.99064	:27	7,406713933	0	1	1	0
E:ALA201:HA - :17:H53	4.27102	:28	7,876148359	1	1	1	0
		:29	7,293282218	1	1	1	0

Figura 15 – Exemplo de distâncias coletadas (esquerda) e matrizes binárias geradas (direita). Nas matrizes, linhas representam ligantes enquanto colunas representam átomos do sítio ativo.

Foram geradas 32 matrizes binárias para cada um dos modelos, representando as combinações possíveis entre função de pontuação, protonação dos ligantes, hidratação do sítio ativo e distância limite de contatos. Para remover ruído estatístico do modelo $\alpha 7$, removeram-se colunas contendo contatos esparsos, *i. e.* colunas nas quais foram identificados 2 ou menos contatos com o total do conjunto de moléculas do grupo de treinamento. Cada coluna da matriz é regredida contra a bioatividade dos ligantes individualmente para determinar quais átomos do sítio ativo se correlacionam positivamente e negativamente com a atividade biológica. Em seguida, separam-se as colunas com correlação positiva das colunas com correlação negativa. Colunas com correlação negativa são invertidas, ou seja, células com valor 1 recebem valor 0 e vice-versa, e são excluídas da próxima etapa. Aplica-se então um algoritmo genético nas colunas positivas para determinar quais combinações entre elas melhor explicam a variação da bioatividade dos ligantes do grupo de treinamento. Isso é feito através da observação de qual soma de colunas apresenta melhor colinearidade com a bioatividade experimental. O algoritmo pode ser programado para retornar como resposta qualquer número de contatos positivos concomitantes. Para o modelo $(\alpha 4)_2(\beta 2)_3$ optou-se por utilizar entre 4 a 10 contatos positivos para o limite de distância 2,5 Å e entre 3 a 10 contatos para 3,5 Å. Para o modelo do subtipo $\alpha 7$ se utilizou entre 3 a 10 contatos para ambas as distâncias. Cada conjunto de colunas somadas é tratado individualmente para identificar o modelo ótimo para cada caso. O algoritmo também possibilita a atribuição de pesos para os contatos positivos, de modo que uma coluna qualquer pode aparecer até três vezes em um mesmo modelo dbCICA. Isso é feito através da atribuição de genes bivariados aos cromossomos. Após identificada a melhor combinação entre colunas positivas o algoritmo introduz as colunas negativas invertidas e busca qual combinação entre as colunas positivas ótimas e negativas melhor representa a variação da bioatividade dos ligantes. A escolha de quantos contatos negativos serão considerados cabe ao usuário, sendo que neste trabalho se optou por utilizar 5 ou 10 para ambos os modelos.

O algoritmo genético foi implementado utilizando o programa Matlab (Versão 2007b) e é constituído por quatro funções básicas: a função de criação, função de cruzamento, função de mutação e função de ajuste, conforme ilustrado na **Figura 16**.

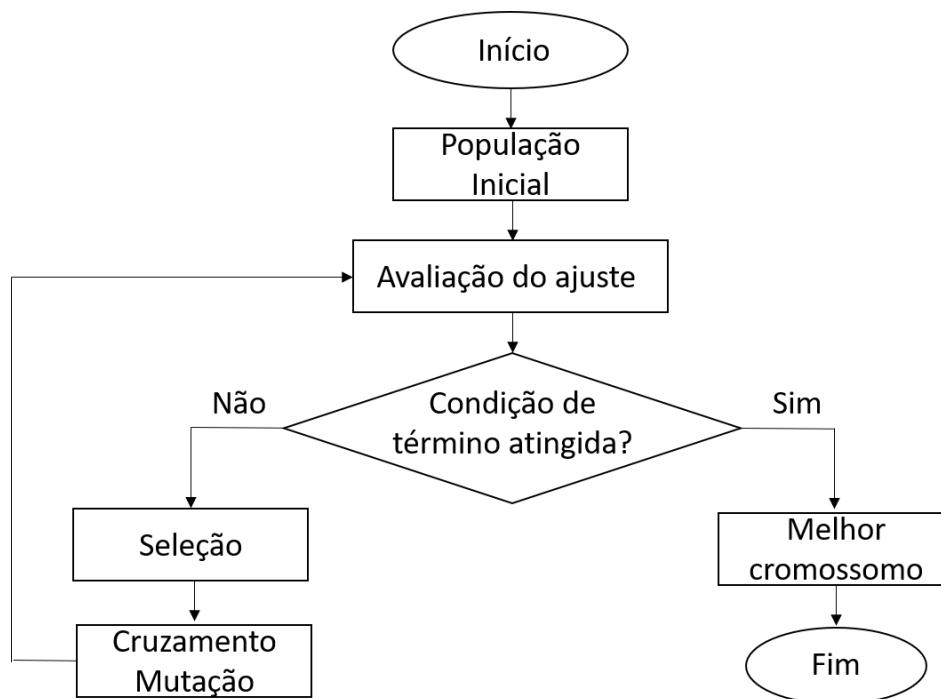


Figura 16 – Fluxograma do processo de seleção de variáveis através de um algoritmo genético.

A função de criação gera randomicamente uma população inicial de cromossomos com tamanho predefinido que corresponde à quantos contatos estão sendo somados, de modo que cada cromossomo representa uma possibilidade de combinação. Dessa forma, os cromossomos diferem entre si nos conjuntos de contatos que estão sendo somados e seus pesos. Os pesos de cada gene na primeira geração também são gerados aleatoriamente. As próximas gerações serão criadas através das funções de cruzamento e mutação da seguinte maneira: a função de cruzamento cria a prole a partir da população já existente combinando os genes de pares de cromossomos, gerando assim um par de novos cromossomos que são complementares entre si, enquanto a função de mutação aplica aleatoriamente mudanças nos genes de um cromossomo, produzindo somente um novo cromossomo. Cada cromossomo está associado à um valor de ajuste, que representa quão bem essa combinação de colunas explica a variação de bioatividade dos ligantes, sendo que esse valor é determinado pela função de ajuste e pode ser o coeficiente de correlação simples (r^2), coeficiente de correlação *leave-one out* (r^2_{LOO}) ou o coeficiente de correlação *K-fold* (r^2_{K-fold}). Neste trabalho, optou-se por implementar o coeficiente de correlação *5-fold* r^2 como método de ajuste de ambos os modelos. Para cada geração o cromossomo com melhor *5-fold* r^2 é definido como “criança elite” e é transmitido para a próxima geração sem alterações. Os parâmetros

escolhidos para o algoritmo genético foram: população inicial = 200, fração de cruzamento = 80%, criança elite = 1 e número máximo de gerações para finalização do algoritmo = 2000. Com base nas configurações descritas a quantidade e tipo da prole gerada a cada a geração é a seguinte: existe uma criança elite e 199 cromossomos além desse. Desses 199 cromossomos, 159 serão tratados pela função de cruzamento ($0,8 \times 199 = 159,2$, arredondado para 159) e os 40 remanescentes pela função de mutação. Dessa forma, o tamanho da população é mantido fixo no decorrer das gerações (1 elite + 159 cruzamentos + 40 mutações = 200 cromossomos).

A função de ajuste *5-fold* r^2 é utilizada para ranquear os cromossomos da seguinte forma: o grupo de treinamento é dividido aleatoriamente em dois subgrupos, o grupo de ajuste e o grupo de teste, de modo que o grupo de teste corresponda a 20% do total do grupo de treinamento. Essa operação se repete 5 vezes, gerando 5 pares de subgrupos, conforme mostrado na **Figura 17**.

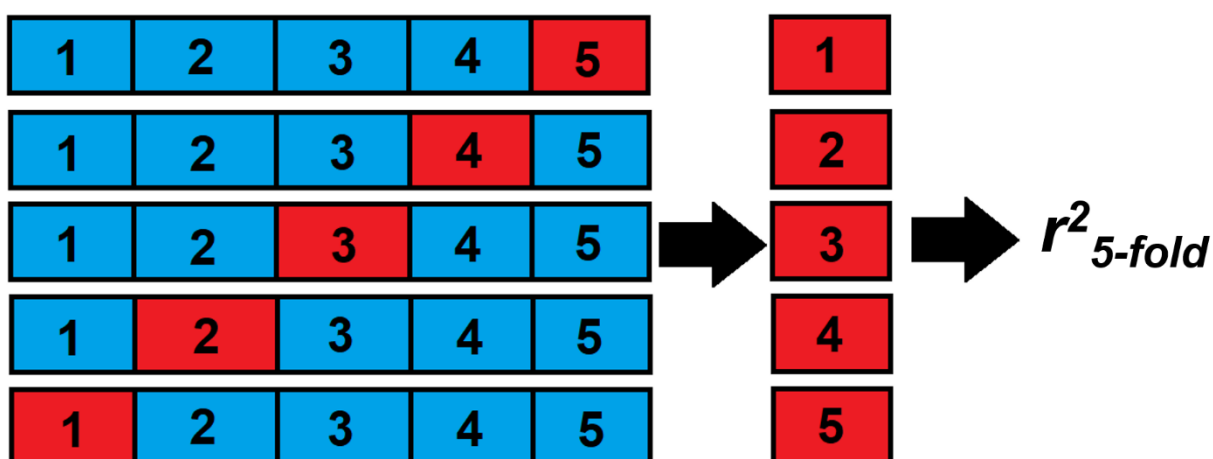


Figura 17 – Esquema de validação cruzada 5-fold.

O algoritmo é instruído a não repetir as moléculas presentes em cada grupo de teste para cada cromossomo de modo que cerca de 100% do grupo de treinamento esteja contido nos 5 subgrupos de teste. Os grupos de ajuste são então utilizados para construir 5 submodelos que serão utilizados na predição dos valores de pK_i dos respectivos subgrupos de teste. Os valores preditos para todos os cinco subgrupos são então correlacionados com os valores reais para a determinação do coeficiente de correlação *5-fold* r^2 para o cromossomo em questão. Ao final das iterações do algoritmo genético o cromossomo com maior *5-fold* r^2 é definido como o melhor modelo dbCICA.

3.6. Triagem virtual dos 116 fragmentos propostos

As condições de ancoragem molecular do melhor modelo dbCICA para o subtipo $(\alpha 4)_2(\beta 2)_3$ foram utilizadas para selecionar os fragmentos mais promissores para a síntese. Inicialmente se realizou a ancoragem molecular dos 116 fragmentos moleculares conforme descrito anteriormente. Tomou-se a pose melhor pontuada para cada um deles exceto quando houve grande discrepância na orientação das duas primeiras poses e a variação de pontuação entre elas foi menor do que 0,5. Nesses casos, a pose tomada como representativa foi aquela que apresentou o máximo de interações com resíduos de interesse, priorizando aquelas com mais contatos críticos apontados pelo modelo dbCICA. Definem-se resíduos de interesse como aqueles que interagem com **1** na estrutura cristalográfica 5KXI ou com **7** em 5AFH. Gerou-se então uma matriz binária para o conjunto de fragmentos, conforme explicado anteriormente e eles foram selecionados inicialmente com base nas interações envolvendo resíduos de interesse. Cabe ressaltar que as colunas negativas foram invertidas para que a soma dos contatos seja sempre positiva. Em seguida, utiliza-se a matriz binária para priorizar moléculas que participem de contatos críticos ao mesmo tempo que evitam contatos desfavoráveis apontados pelo modelo, dessa forma obtendo 8 candidatos promissores para a etapa de síntese. As condições ideais de ancoragem para o modelo $\alpha 7$ também foram utilizadas de maneira semelhante para determinar se algum dos 8 ligantes selecionados para o modelo $(\alpha 4)_2(\beta 2)_3$ poderia, em teoria, atuar em ambos os receptores.

3.7. Mapa farmacofórico do modelo $(\alpha 4)_2(\beta 2)_3$

Os resultados obtidos através da abordagem dbCICA para o subtipo $(\alpha 4)_2(\beta 2)_3$ foram utilizados na construção de um mapa farmacofórico que pode ser útil em posterior etapa de incremento estrutural dos fragmentos sintetizados (*fragment growth*) e na triagem virtual de novos ligantes desse subtipo. O modelo foi construído da seguinte forma: utilizando a ferramenta de regressão do Microsoft® EXCEL (365 ProPlus) se obteve a correlação entre a soma dos contatos apontados pelo melhor modelo dbCICA e a bioatividade experimental dos ligantes, de modo que foi possível prever os valores de pK_i para cada um deles. Em seguida, foram selecionados os ligantes mais ativos e bem-comportados dentro do modelo, ou seja, aqueles com

menor resíduo absoluto entre o valor real e predito. As poses de ancoragem desses ligantes foram sobrepostas no sítio ativo do receptor dentro do ambiente do software CSD CrossMiner (v. 1.4),¹⁸⁷ onde foram observadas sobreposições de alguns grupos químicos. A construção do mapa é feita então com base na complementariedade entre esses grupos sobrepostos e os átomos críticos mais próximos a eles através da criação de esferas que englobem as características essenciais para a ocorrência das interações de interesse, chamadas esferas de característica. Utilizam-se também os átomos dos contatos negativos para a criação de esferas de exclusão. O raio de tolerância permitido para as esferas de exclusão e de característica foi 1,6 Å, com exceção da esfera de característica que representa a projeção espacial associada a doadores de ligação de hidrogênio cujo raio foi de 2,2 Å.¹⁴⁴ Para a construção do mapa se utilizaram os contatos positivos e negativos do melhor modelo dbCICA e os contatos negativos do segundo melhor modelo a fim de refinar as zonas de exclusão.

3.8. Validação do mapa farmacofórico (α 4)₂(β 2)₃

A capacidade do mapa farmacofórico (α 4)₂(β 2)₃ de discernir entre ligantes ativos e inativos foi avaliada através da curva de característica de operação do receptor (ROC). Para tal, selecionou-se 34 moléculas dentre as 98 iniciais de modo que não houvesse moléculas com o mesmo valor de bioatividade e também sem distância Euclidiana igual a zero entre quaisquer outras moléculas do conjunto selecionado. A distância Euclidiana é uma medida de similaridade entre as moléculas e neste trabalho a métrica é aplicada para remover moléculas muito similares, diversificando o banco de dados de ativos. A equação que define a distância Euclidiana entre duas moléculas está descrita no **Anexo II**. Em seguida se utilizou o software DecoyFinder (v. 2.0)¹⁸⁸ para obter do banco de dados ZINC¹⁸⁹ 25 moléculas presumidamente inativas para cada um dos 34 ativos, totalizando 850 inativos. Os inativos foram selecionados com base em alguns critérios de similaridade frente aos seus respectivos ativos, em acordo com estudos da literatura que indicam que bancos de moléculas do tipo focados reduzem enriquecimento artificial dos modelos obtidos.¹⁴¹ A similaridade de Tanimoto entre ativos e inativos foi de no mínimo 0,85, a similaridade entre inativos e os outros 24 inativos do conjunto para um mesmo ativo também foi de no mínimo 0,85, se aceitando que os inativos tivessem ± 1 doador e ± 1 aceptor de ligação de hidrogênio em relação ao seus respectivos ativos. Permitiu-se também que o peso molecular dos inativos variasse em ± 48 Da em relação aos

seus ativos, ± 2 ligações rotacionáveis também foram aceitas, o coeficiente de partição água-octanol ($\log P$) pode variar em ± 1 e o desvio padrão máximo que os inativos poderiam ter em relação aos seus ativos foi de 1,5. As estruturas dos inativos foram trabalhadas para remover estados de protonação irreais para o pH de interesse (7,4) fazendo uso do software MarvinSketch para predição do microestado de ionização majoritário no pH indicado. Uma vez selecionados os inativos, o software Mercury foi utilizado para gerar até 500 confôrmeros para cada uma das 884 moléculas. Utilizando novamente o software CSD CrossMiner realizou-se então uma busca por *hits* dentro desse banco de confôrmeros. Os resultados dessa busca foram tratados no OriginPro¹⁹⁰ para a elaboração da curva ROC.

4. Resultados e discussão

4.1. Subtipo ($\alpha 4$)₂($\beta 2$)₃

Inicialmente se verificou os ângulos ϕ e Ψ dos aminoácidos da estrutura 5KXI através do servidor online RAMPAGE,¹⁹¹ concluindo-se através do gráfico de Ramachandran gerado que os resíduos com os quais **3** interagem não estão em regiões proibidas (**Anexo III**), sendo que o único que está em uma região permitida é A:W156 e os demais estão em regiões favoráveis. A nomenclatura adotada neste trabalho para a identificação de átomos do sítio ativo tem formato 1:2:3 onde 3 representa um átomo do sítio ativo, 2 o resíduo ao qual tal átomo pertence e 1 representa a subunidade na qual o resíduo indicado está localizado, sendo A a subunidade α e B a subunidade β . Os átomos foram nomeados automaticamente pelo software Discovery Studio Visualizer e a numeração dos resíduos respeita a numeração publicada originalmente para a estrutura utilizada. A estrutura foi inspecionada para verificar se as interações **3**-receptor eram mediadas por moléculas de água, concluindo-se que não são. Porém, devido à baixa resolução da estrutura cristalográfica e à estudos anteriores envolvendo proteínas homólogas e dinâmica molecular onde a água é essencial para o reconhecimento molecular, optou-se por modelar uma molécula de água no sítio ativo, conforme descrito na metodologia.^{185,186,192–194} Os resíduos que participam de interações com **3**, referidos daqui em diante como resíduos de interesse, e o tipo das mesmas são os seguintes: A:W156 (cátion- π , ligação de hidrogênio, hidrofóbica), A:Y204 (cátion- π , hidrofóbica), A:Y100 (ligação de hidrogênio não clássica), B:L121, A:T157, A:C200, B:W57 e

A:Y197 (interações do tipo hidrofóbicas para todos esses últimos). O modo de ligação no qual se encontra **3** co-cristalizado e os resíduos presentes na região estão representados na **Figura 18**, gerada através do software PyMol (v.099rc6),¹⁹⁵ que também foi utilizado para elaborar todas as visualizações tridimensionais demonstradas nesse trabalho, exceto quando diferentemente especificado.

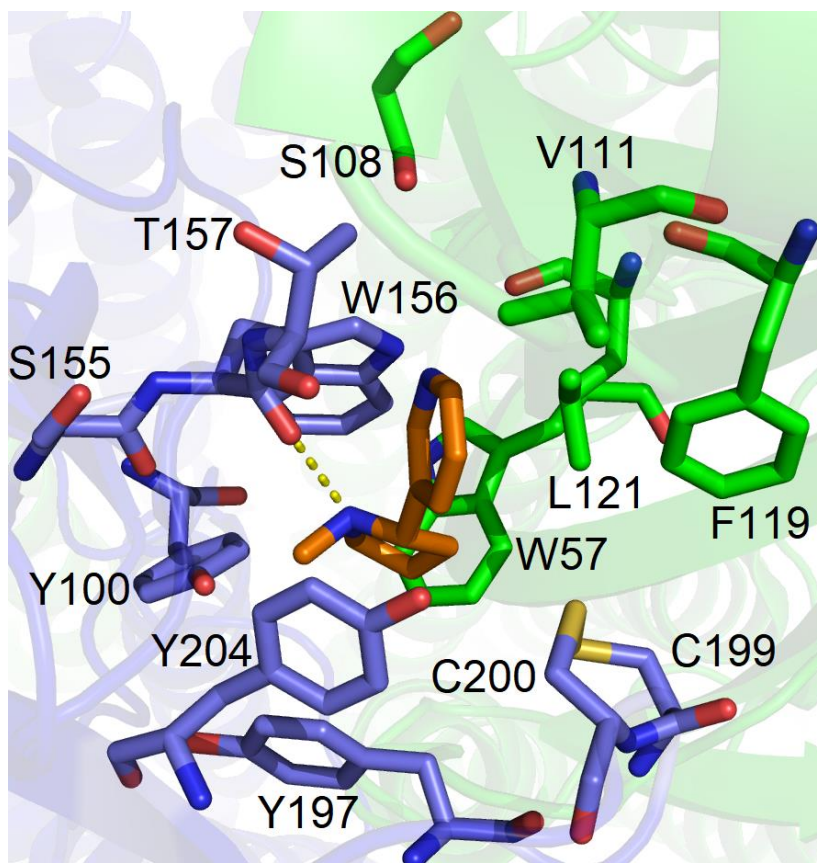


Figura 18 - Pose do ligante co-cristalizado (**3**) (em laranja) e resíduos da região do sítio ativo. Átomos de hidrogênio foram omitidos para maior clareza. Linhas tracejadas em amarelo representam átomos envolvidos em ligações de hidrogênio. Resíduos em azul fazem parte da subunidade α e os em verde da subunidade β . PDB ID: 5KXI

Em seguida, as 98 estruturas preparadas para a ancoragem molecular foram ancoradas no receptor de acordo com a metodologia descrita na seção anterior, em todas as condições de ancoragem citadas. Utilizando as poses e enantiômeros melhor pontuados para cada ligante foram gerados diversos modelos dbCICA conforme previamente descrito.

Os 10 modelos com maior *5-fold* r^2 estão descritos na **Tabela 4**. Os contatos críticos apontados pelos dois melhores modelos estão relacionados na **Tabela 5**.

Tabela 4 - Condições e estatística dos 10 melhores modelos dbCICA.

Modelo	Função de Pontuação	Água	Ionização	Distância dos contatos (Å)	Contatos (Positivos_Negativos)	5-fold r^2	Estatística F
nAChRs_α4β2_1	ChemScore	Sim	Sim	3,5	4_10	0,593	157,21
nAChRs_α4β2_2	ASP	Sim	Não	3,5	3_10	0,592	146,23
nAChRs_α4β2_3	CHEMPLP	Não	Não	3,5	4_10	0,582	139,92
nAChRs_α4β2_4	GOLDScore	Não	Sim	3,5	4_10	0,573	134,58
nAChRs_α4β2_5	ChemScore	Sim	Sim	3,5	3_10	0,568	141,80
nAChRs_α4β2_6	CHEMPLP	Não	Sim	3,5	4_5	0,565	127,31
nAChRs_α4β2_7	CHEMPLP	Não	Sim	3,5	3_10	0,563	136,79
nAChRs_α4β2_8	GOLDScore	Não	Sim	2,5	4_10	0,555	127,76
nAChRs_α4β2_9	CHEMPLP	Sim	Sim	3,5	3_10	0,555	125,06
nAChRs_α4β2_10	ASP	Sim	Não	3,5	4_10	0,549	125,65

Tabela 5 - Contatos críticos dos 2 melhores modelos dbCICA para o subtipo (α4)2(β2)3.

	Contatos positivos (peso)	Contatos negativos
nAChRs_α4β2_1	B:F119:CB (2), A:T157:HG23 (2), A:W156:C (3), A:Y204:CG (2)	A:A201:HB2, B:R81:HH22, B:K79:CD, B:W57:CE3, A:Y100:HD1, A:Y100:OH, A:Y197:HE1, A:Y204:CZ, A:Y204:HE1, B:V111:HG13
nAChRs_α4β2_2	HOH:1 (2), B:L121:HG (2), A:W156:C (2)	A:C200:HB32, B:L121:CD1, B:F119:HD2, B:F119:HE2, B:F119:HZ, A:W156:CB, A:W156:CG, A:W156:HH2, B:W57:CE2, A:Y204:HE2

Observando os dados apresentados na **Tabela 4** é possível constatar que a condição de ancoragem molecular que produziu melhor distinção entre as moléculas ativas e inativas se deu quando se utilizou a função ChemScore com os ligantes protonados em pH 7,4, uma molécula de água no sítio ativo e distância máxima de contatos igual a 3,5 Å, resultando em um *5-fold* r^2 igual a 0,593 e estatística F igual a 157,21. A estatística *5-fold* r^2 indica que aproximadamente 59% da variância observada para Y foi explicada através do modelo de regressão. Entende-se que esse valor é aceitável, uma vez que o modelo de regressão correlaciona a bioatividade, uma propriedade complexa, com a somatória dos contatos positivos e negativos invertidos, um único descritor e de alta simplicidade. Além disso, o objetivo da análise dbCICA é a identificação de quais contatos são determinantes para a bioatividade e o uso dessa informação para a triagem virtual, de modo que a predição dos valores de atividade é apenas uma ferramenta para a construção e ranqueamento dos modelo e não uma das propriedades que se deseja avaliar. Quanto à estatística F para o modelo nAChRs_α4β2_1, o valor de F crítico para os graus de liberdade do modelo (1 e 96) é igual a 3,94, com $p = 6,2 \times 10^{-22}$, indicando bom ajuste.

A presença de água não é tão crítica quanto a ionização dos ligantes, visto que para 5 dos 10 modelos a água não se encontra presente e apenas 3 dos 10 modelos apresentam ligantes não ionizados. Quanto aos contatos apontados pelos dois melhores modelos (**Tabela 5**), observa-se concordância quanto ao contato entre o átomo C do resíduo A:W156, reforçando sua importância. Tal contato representa uma ligação de hidrogênio entre os átomos de nitrogênio básicos dos ligantes e a carbonila da ligação peptídica do resíduo em questão, em acordo com diversos grupos de pesquisa que reportaram a importância dessa interação para o reconhecimento molecular de ligantes desse receptor.^{25,185,196,197} O contato do átomo CG do resíduo A:Y204 no modelo nAChRs_α4β2_1 representa interações hidrofóbicas que surgem em decorrência da ligação de hidrogênio citada anteriormente, pois ela tende a orientar os átomos de carbono α-N na direção desse resíduo. Além disso, quando regiões positivamente carregadas participam da ligação de hidrogênio citada ocorrem interações cátion-π envolvendo os resíduos aromáticos desta região. Estudos já publicados indicam que a interação cátion-π entre ligantes como acetilcolina, nicotina e carbamilcolina e a cadeia lateral do resíduo W156 é essencial para a atividade dessas moléculas.^{197,198} Entende-se que a proximidade dos ligantes ativos com o átomo HG23 do resíduo A:T157 ocorre devido à presença da molécula de água de hidratação no sítio ativo, uma vez que para maximizar as interações entre a mesma e os anéis heteroaromáticos os ligantes tendem a se aproximar do átomo citado. Quanto aos contatos negativos apontados pelo modelo é interessante notar que alguns deles também tem respaldo na literatura. Por exemplo, sabe-se através de estudos de mutagênese sítio dirigida que a mutação Y100F reduz a afinidade da acetilcolina e carbamilcolina devido à ausência do grupo hidroxila, que se encontra nas proximidades desses ligantes.^{199,200} Semelhantemente, mutações em Y197 reduzem a afinidade da acetilcolina pelo receptor.²⁰⁰⁻²⁰² Dessa forma se justificam tais contatos, uma vez que os mesmos representam choques estéricos com resíduos determinantes para a atividade biológica das moléculas. Constata-se então boa concordância entre os resultados obtidos para o modelo dbCICA nAChRs_α4β2_1 e dados da literatura, reforçando sua utilidade.

Como foi dito anteriormente, as moléculas **3** e **13** foram incluídas no conjunto de treinamento e a análise das suas melhores poses de ancoragem traz informações importantes a respeito da validade do modelo. A melhor pose de ancoragem para **3**

sobreposta à conformação obtida da estrutura cristalográfica original está representada na **Figura 19**.

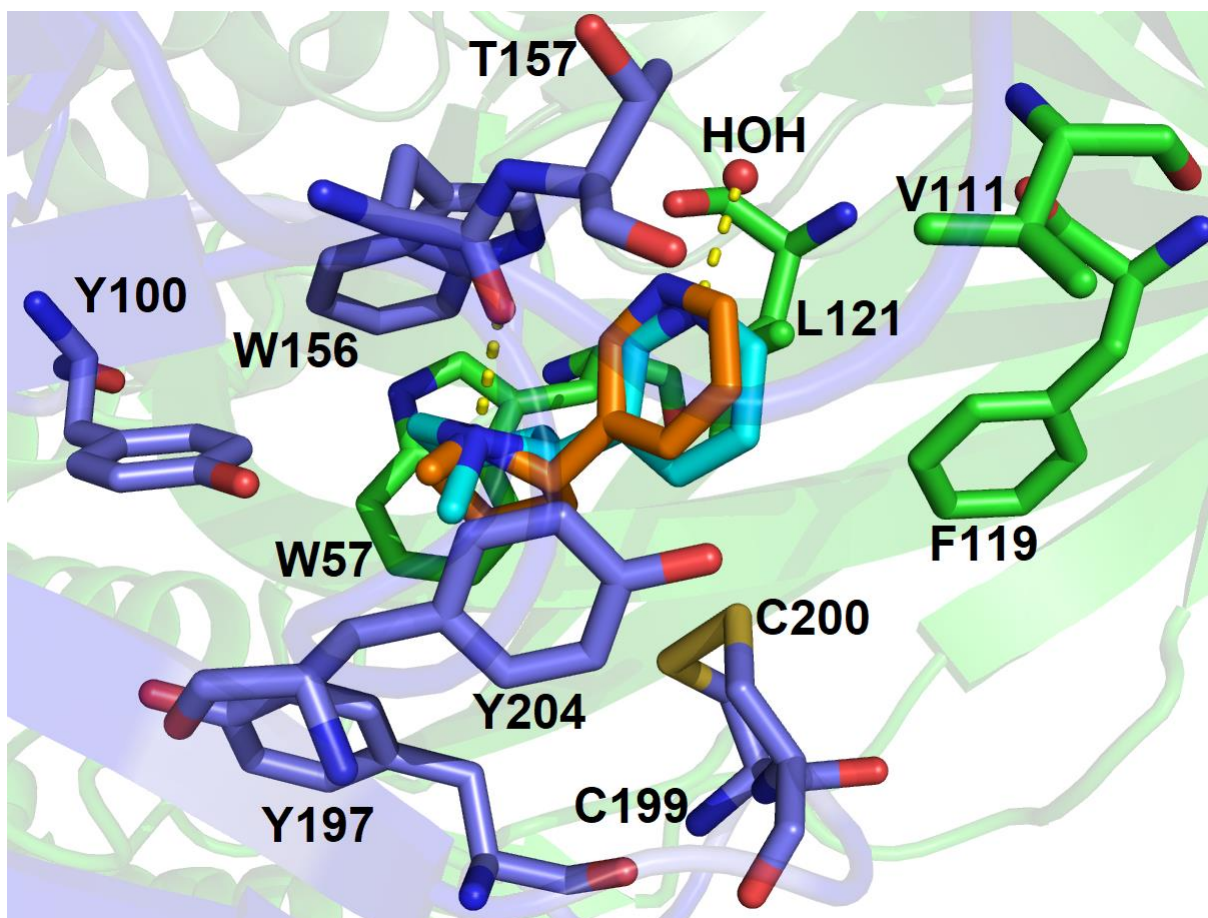


Figura 19 - Pose de **3** co-cristalizado (laranja) sobreposta à **3** ancorado (ciano) e resíduos da região do sítio ativo. RMSD = 0,83 Å. Átomos de hidrogênio foram omitidos para maior clareza. Linhas tracejadas em amarelo representam átomos envolvidos em ligações de hidrogênio. Resíduos em azul fazem parte da subunidade α e os em verde da subunidade β .

Através da **Figura 19** é possível observar que utilizando as condições de ancoragem molecular do melhor modelo dbCICA os grupos funcionais de **3** ancorado e co-cristalizado apresentam bom alinhamento entre si, com RMSD igual a 0,83 Å. A pontuação para essa pose foi de 29,0710 e as interações para ela previstas foram: Uma ligação de hidrogênio envolvendo o nitrogênio aromático e a água do sítio ativo, A:W156 (duas ligações de hidrogênio, uma comum e uma não clássica, cátion- π), A:Y204 (cátion- π e hidrofóbica), A:T157 (ligação de hidrogênio não clássica), B:L121, A:C200, sendo essas duas últimas somente hidrofóbicas.

A **Figura 20** mostra a comparação entre a pose experimental de **13** em uma AChBP homóloga proveniente da espécie *Aplysia californica* e o resultado da

ancoragem molecular realizada. O alinhamento tridimensional que possibilita tal comparação está contido no Anexo IV.

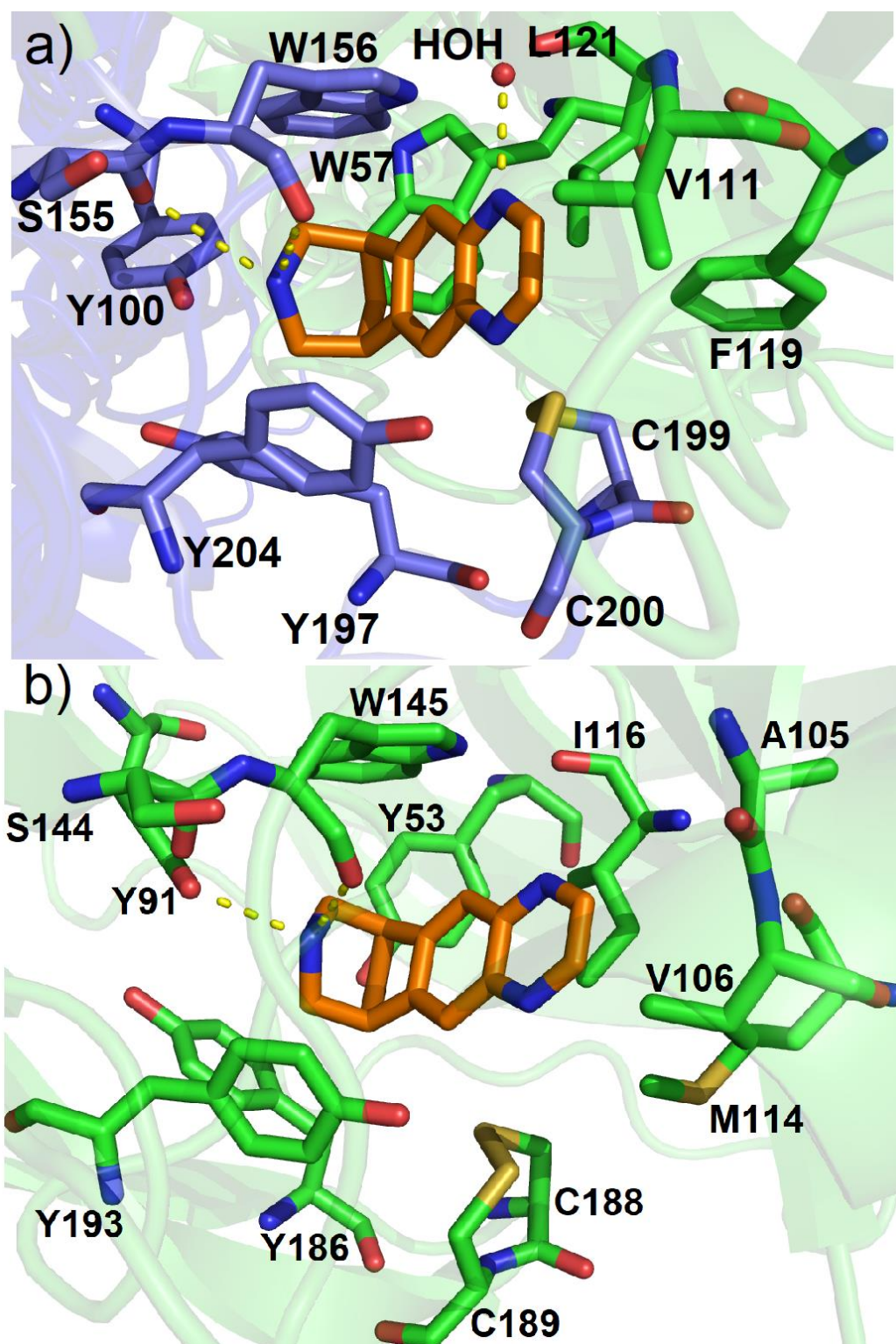


Figura 20 - Comparação entre a orientação espacial do ligante 13 a) ancorado no nAChRs $\alpha 4\beta 2$ e b) co-cristalizado na AChBP da espécie *Aplysia californica* (4AFT).¹⁹⁶ Átomos de hidrogênio foram omitidos para maior clareza. Linhas tracejadas em amarelo representam átomos envolvidos em ligações de hidrogênio.

De acordo com a sobreposição mostrada (**Anexo IV**), os resíduos de interesse A:W156, A:Y197 e A:Y204 do nAChRs estão conservados na mesma posição espacial na estrutura 4AFT, enquanto os resíduos A:C199 e A:C200 estão conservados mas em posição espacial diferente e os resíduos B:W57 e B:L121 estão substituídos por resíduos similares que ocupam a mesma posição espacial, correspondendo aos resíduos W145, Y186, Y193, C188, C189, Y53 e I116 da proteína homóloga. Dessa forma, entende-se que a estrutura homóloga é adequada para a ancoragem comparativa uma vez que 5 de 7 resíduos de interesse se encontram espacialmente conservados ou substituídos por resíduos de efeitos eletrônicos semelhantes. Para **13** ancorada, as interações são as seguintes: A:W156 (ligação de hidrogênio, cátion- π , empilhamento-T e hidrofóbica), A:S155 (ligação de hidrogênio), B:F119 (empilhamento-T), A:Y204 (cátion- π e empilhamento-T) B:L121, A:T157, A:C200, A:C199, B:W57, A:Y197, e B:V111 (do tipo hidrofóbico). O cálculo mostra também dois contatos repulsivos entre as cadeias laterais dos resíduos A:Y204 e A:Y100 e **3**, além de uma ligação de hidrogênio com a água de hidratação do sítio ativo. Apesar dos contatos estéricos, houve boa concordância entre o cálculo teórico e o resultado obtido da estrutura homóloga, reforçando a conclusão de que a condição de ancoragem molecular utilizada é adequada para a triagem virtual dos fragmentos moleculares.

Os 8 melhores fragmentos selecionados, conforme descrito na metodologia, tiveram suas propriedades físico-químicas calculadas e estão relacionados na **Figura 21**. As interações previstas pelo modelo computacional para cada um deles estão contidas na **Tabela 6**.

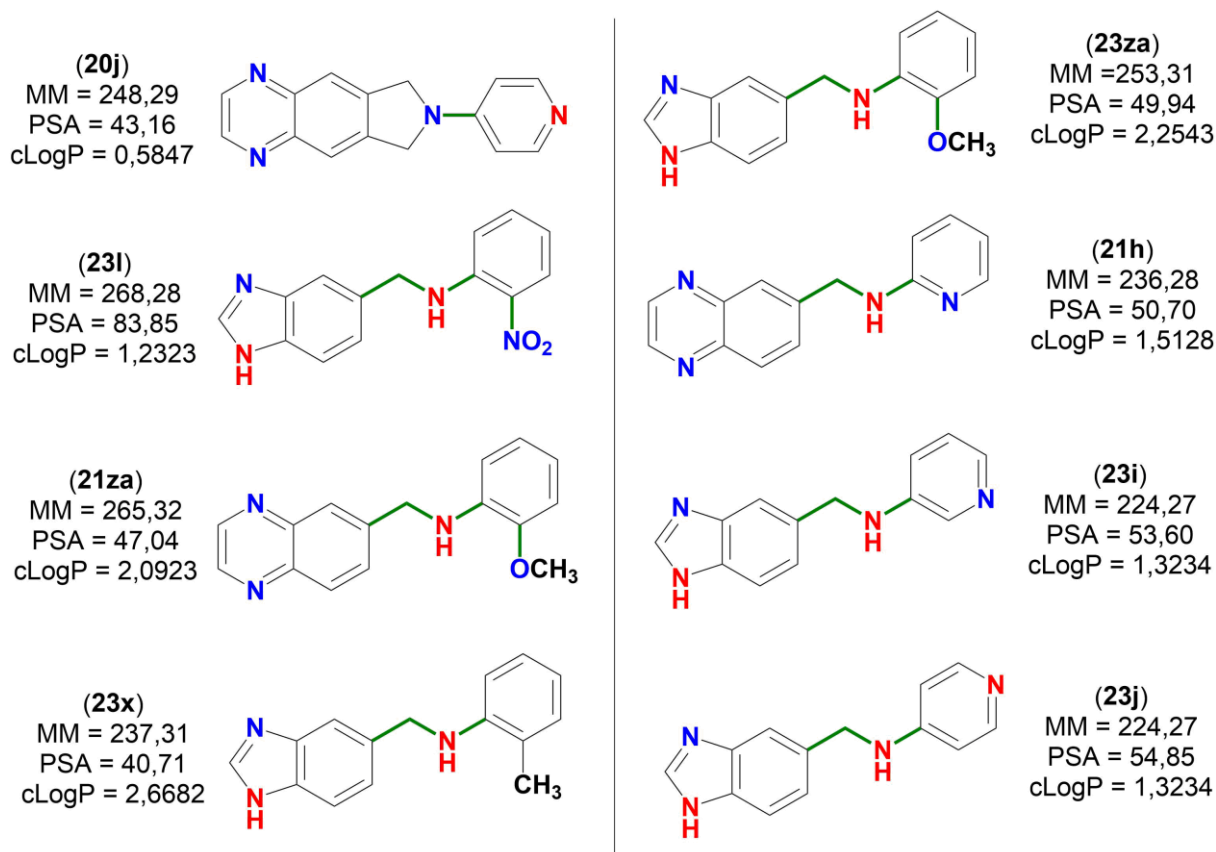


Figura 21 - Fórmula estrutural, massa molecular (MM), área de superfície polar (PSA) e coeficiente de partição (cLogP) dos 8 fragmentos seleccionados para a etapa de síntese. Átomos em vermelho são doadores e aceptores de ligação de hidrogênio, átomos em azul aceptores de ligação de hidrogênio e ligações químicas em rosa indicam rotação livre. Nitrogênios piridínicos em vermelho indicam que os mesmos foram previstos como protonados em pH 7,4 pelo software MarvinSketch. As propriedades físico-químicas foram calculadas nos softwares MarvinSketch (MM e PSA) e OSIRIS DataWarrior (cLogP, v. 4.7.3).²⁰³

Tabela 6 - Interações observadas para **3** em 5KXI, interações previstas para **3** e **13** e interações previstas para os 8 fragmentos seleccionados para síntese. À direita, soma dos contatos críticos apontados no modelo nAChRs_α4β2_1. Resíduos em vermelho são os de interesse.

Ligante	ChemScore	H ₂ O	W156	Y204	L121	T157	C200	C199	Y100	W57	Y197	V111	S155	F119	E202	A201	K79	D171	Soma
3 (5KXI)	-	-																	
3 (docking)	29,071																		
13	37,4297																		
20j	25,6033																		12
23l	25,5894																		11
21za	35,2376																		10
23x	34,4323																		10
23za	32,184																		10
21h	30,264																		10
23i	29,7049																		10
23j	28,5391																		10

Interações:

- Choque estérico
- Ligação de hidrogênio com átomos da ligação peptídica
- Ligação de hidrogênio com água
- π-Par isolado
- Atração eletrostática
- Cation-π
- Hidrofóbica
- Anion-π
- Ligação de hidrogênio não clássica
- Empilhamento-T
- Empilhamento-π
- π-Enxofre

Observando os dados relacionados na **Tabela 6** é possível constatar que as moléculas selecionadas, em sua maioria, interagem com os resíduos de interesse. Nota-se também através da **Figura 21** que as moléculas mais promissoras pertencem às séries com cadeia aberta, sendo que somente **20j** é conformacionalmente restrito. Optou-se por dar preferência para fragmentos que participassem da ligação de hidrogênio com a carbonila da ligação peptídica do resíduo A:W156, um contato crítico apontado pelo modelo dbCICA, visto que essa interação é bem estudada e relevante para o reconhecimento molecular de ligantes pelos nAChRs.^{25,185,196,197} Idealmente os ligantes deveriam participar de interações cátion- π com os resíduos aromáticos, visto que existem referências da literatura que endossam esses contatos,^{197,198} mas pelo fato da maioria dos fragmentos não existir em forma protonada em pH 7,4 isso não é observado, sendo que somente a molécula **23j** alcança esse critério. A **Figura 21** também apresenta algumas propriedades físico-químicas calculadas para os fragmentos moleculares selecionados a partir das quais é possível constatar se os mesmos respeitam a regra dos 3 (RO3). A RO3 é uma adaptação da regra dos 5 de Lipinsky que se aplica à fragmentos moleculares e postula que os mesmos, em geral, devem ter MM < 300, 3 ou menos doadores de ligação de hidrogênio, 3 ou menos aceptores de ligação de hidrogênio, 3 ou menos ligações com rotação livre, área de superfície polar ≤ 60 e cLogP ≤ 3 .²⁰⁴ Todas as moléculas selecionadas se enquadram na restrição de massa da RO3 e de doadores/aceptores de ligação de hidrogênio, sendo que somente **23i** falha no quesito de área de superfície polar. Quanto a rotações livres, as moléculas **23i**, **21za** e **23za** não se encaixam nos postulados da RO3. Os nitrogênios piridínicos de **20j** e **23j** são contados como doadores de ligação de hidrogênio devido às suas microespécies majoritárias calculadas para pH 7,4 estarem protonadas no átomo citado. Os dados a respeito dos contatos obtidos do modelo teórico e a boa concordância com a RO3 indicam que essas moléculas têm potencial biológico e podem ser utilizadas em uma etapa posterior de otimização, que fará uso do mapa farmacofórico elaborado utilizando os dados do dbCICA. A sobreposição dos modos de ligação preditos para os ligantes descritos na **Figura 21** e **Tabela 6** estão representados na **Figura 22**.

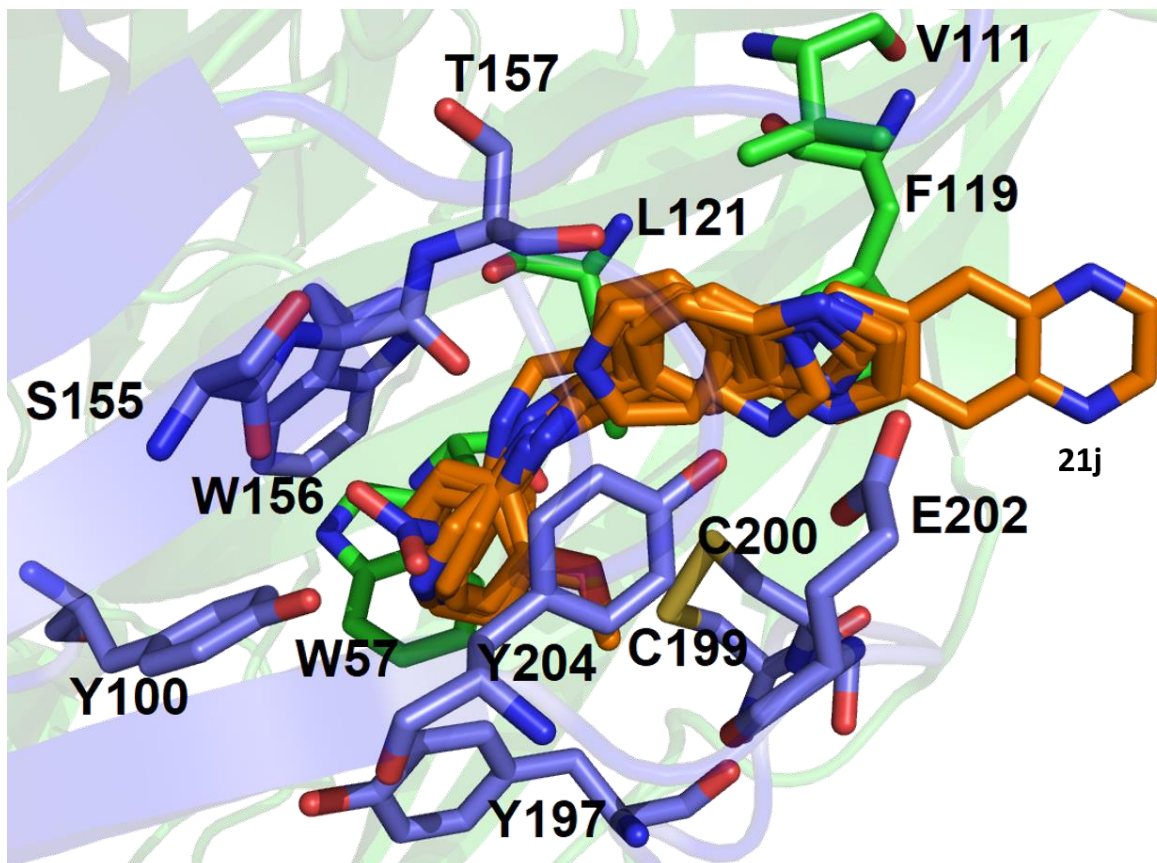


Figura 22 - Sobreposição das poses melhor ranqueadas para os 8 fragmentos moleculares (laranja) triados para a etapa de síntese. Átomos de hidrogênio foram omitidos para maior clareza. Resíduos em azul fazem parte da subunidade α e os em verde da cadeia β .

A sobreposição mostrada na **Figura 22** apresenta boa concordância entre as poses previstas para 7 dos ligantes, com exceção de **20j**. Esse fragmento se encontra na região mais externa do sítio ativo em decorrência da rigidez que impossibilita que o mesmo se posicione de maneira semelhante aos demais. Nota-se que os fragmentos *o*-substituídos se orientam na direção de A:Y204, em contato próximo com o átomo crítico CG do mesmo resíduo. Os compostos se orientam de modo a posicionar a região monocíclica aromática entre as cadeias laterais dos resíduos da cavidade aromática, A:W156, A:Y204, A:Y197, B:W57 e A:Y100, onde é possível maximizar as interações aromáticas. As subunidades quinoxalina e benzimidazola se encontram em uma região onde é possível que participem de empilhamentos- π com B:F119 e também de interações ânion- π com a cadeia lateral do resíduo A:E202. Nenhuma das moléculas selecionadas interage com a água modelada no sítio ativo, porém acredita-se que ela seja essencial para que as mesmas se orientem dessa maneira.

A **Figura 23** mostra os contatos críticos apontados e a sobreposição dos ligantes bem comportados no modelo nAChRs_α4β2_1. O mapa farmacofórico construído conforme descrito na metodologia se encontra representado na **Figura 24**.

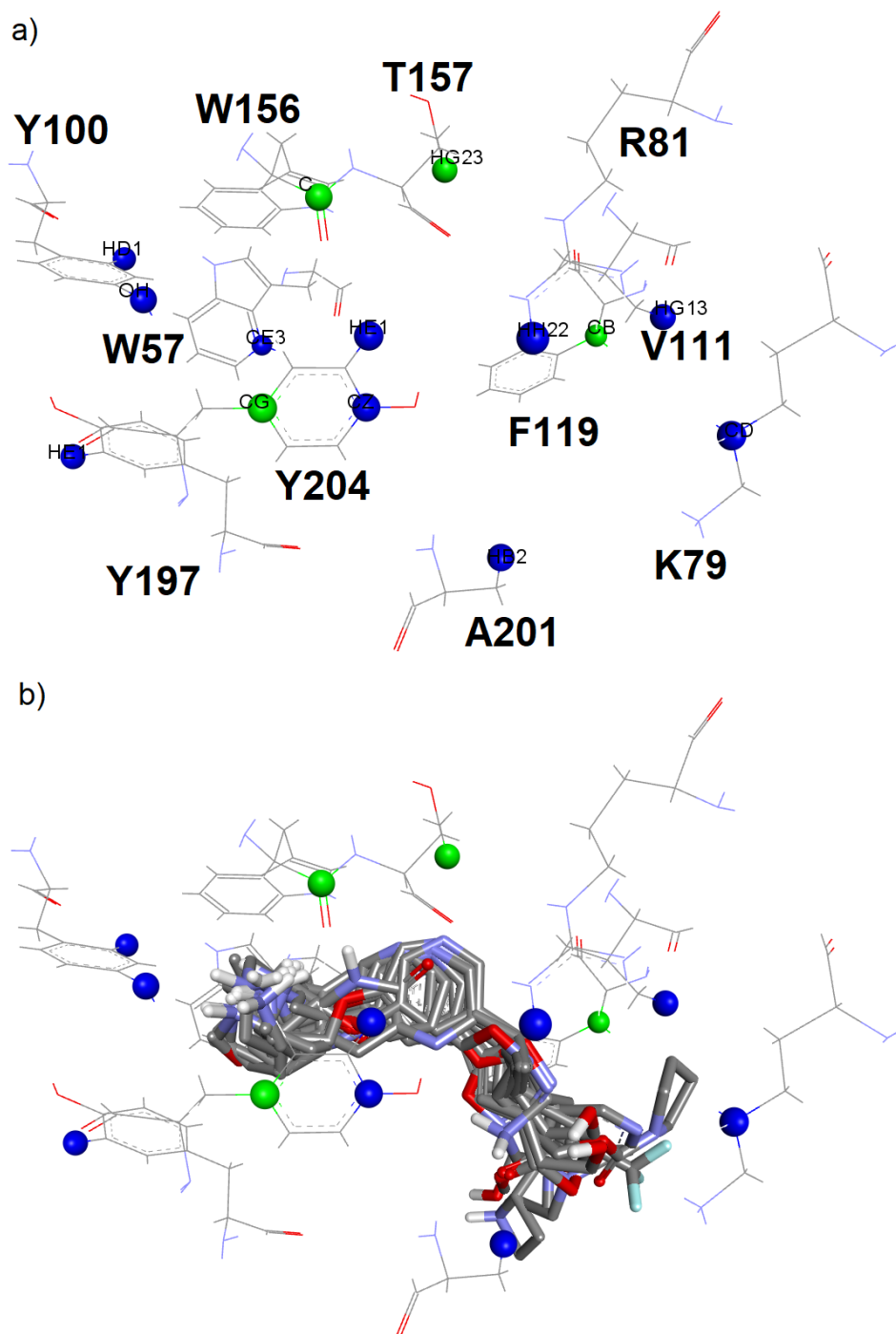


Figura 23 - a) Contatos positivos (verde) e negativos (azul) apontados pelo modelo nAChRs_α4β2_1. b) Sobreposição dos ligantes bem-comportados 3, 13, 20, α4β2_6, α4β2_9, α4β2_13, α4β2_14, α4β2_34, α4β2_44, α4β2_45, α4β2_63, α4β2_71, α4β2_86, α4β2_88 e α4β2_92 dentro do sítio ativo mostrado em a). Figura gerada no Discovery Studio Visualizer.

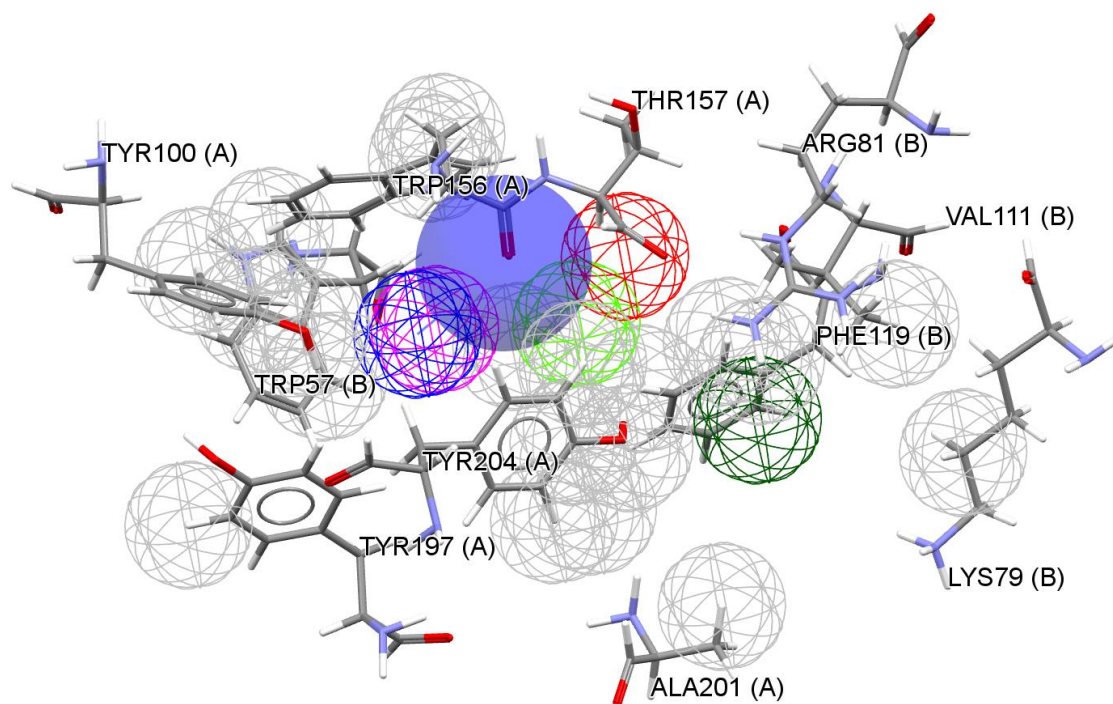


Figura 24 - Mapa farmacofórico do modelo dbCICA nAChRs_α4β2_1 combinado com as esferas de exclusão de nAChRs_α4β2_2. Esferas cinza são as de exclusão, verde claras representam características de anéis aromáticos, verde escuras hidrofóbicas, vermelhas aceptores de ligação de hidrogênio, rosa átomos positivamente carregados e em azul o vetor de doadores de ligação de hidrogênio, sendo que a esfera raiada representa a origem do vetor e a esfera opaca a cabeça para onde o mesmo aponta. Figura gerada no CSD CrossMiner.

A sobreposição mostrada na **Figura 23** revela algumas regiões em comum para os ligantes mais ativos que foram utilizados na construção do mapa farmacofórico mostrado na **Figura 24**. Optou-se por criar uma esfera de característica para átomos positivamente carregados onde houve sobreposição de grupos funcionais básicos, representando também o contato com A:Y204:CG pois é essa interação que resulta na proximidade dos átomos α-N dos ligantes com o resíduo citado. Nessa mesma região se modelou um vetor de doador de ligação de hidrogênio orientado na direção de A:W156:C. O contato com A:T157:HG23 está representado por uma esfera de aceitor de ligação de hidrogênio pois acredita-se que a proximidade com esse átomo se deve a molécula de água no sítio ativo. Próximo a esse ponto se construiu uma esfera de anel aromático onde houve sobreposição desse tipo de subunidade. O contato com B:F119:CB está representado por uma esfera hidrofóbica. Essa esfera se encontra um pouco mais distante do seu átomo complementar do que as demais pois se entende que esse contato existe devido à interação com o anel benzênico desse resíduo.

A capacidade do mapa construído de discernir entre moléculas ativas e inativas foi avaliada através da curva característica da operação do receptor (ROC) mostrada na **Figura 25**.

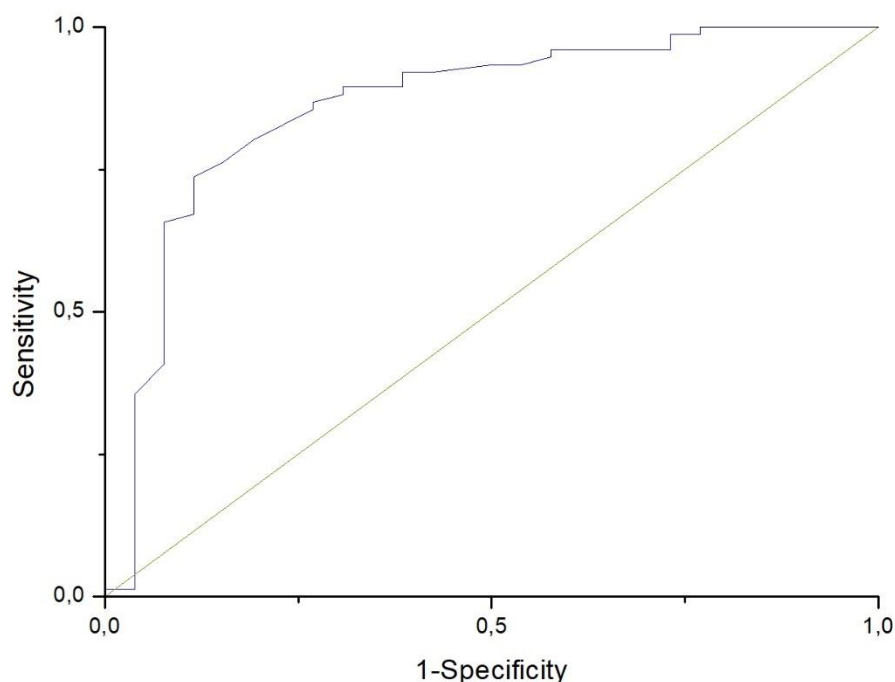


Figura 25 - Curva ROC referente ao mapa farmacofórico do modelo dbCICA nAChRs_α4β2_1 combinado com as esferas de exclusão de nAChRs_α4β2_2. A curva mostrada na **Figura 18** apresenta AUC = 0,86, Ya = 0,25 e ACC = 0,90.

As equações que definem esses parâmetros se encontram no **Anexo II**. A curva ROC mostra como a sensibilidade, que representa as moléculas ativas corretamente classificadas, varia em função da especificidade, representada como 1-Sp e correspondendo aos inativos corretamente descartados. As informações mostradas nesse tipo de curva são referentes ao percentual de moléculas corretamente classificadas pelo mapa farmacofórico quando se utiliza como limites de seleção os valores de RMSD obtidos para cada um dos ligantes ativos por ele capturados. Dessa forma, o limite ideal de seleção é aquele onde a relação entre ativos e inativos corretamente classificados se adequa aos propósitos do trabalho. Sendo que o mapa pode ser utilizado para buscar moléculas já existentes e não testadas frente ao nAChRs do subtipo α4β2, entende-se que o limite ideal seria aquele no qual se obtenham mais ativos do que inativos, respeitando a escala de quantas moléculas existiam no banco de dados inicial utilizado. De acordo com a curva

mostrada o limite ideal para decidir se um *hit* do banco de dados sendo triado tem potencial atividade frente ao receptor $(\alpha_4)_2(\beta_2)_3$ seria o ponto mais alto da inflexão da curva, correspondendo à aproximadamente 0,25 no eixo das abscissas e representado o limite de RMSD = 0,87.

4.2. Subtipo α_7

Inicialmente, foram verificados os ângulos ϕ e Ψ dos aminoácidos da estrutura 5AFH através do servidor online RAMPAGE, concluindo-se através do gráfico de Ramachandran gerado que os resíduos com os quais **7** interagem não estão em regiões proibidas (**Anexo III**). A nomenclatura adotada para identificação dos átomos segue o padrão definido na seção anterior, diferindo somente quanto à codificação das cadeias α , dado que para essa estrutura existem apenas subunidades desse tipo e ambas as letras B e C representam subunidades α . Desta forma, a letra B corresponde à subunidade que compõe o lado principal do sítio ativo enquanto a letra C representa a subunidade complementar do mesmo. Optou-se por utilizar o sítio ativo entre essas duas cadeias devido à ausência da ligação dissulfeto, característica de subunidades α no sítio localizado entre as cadeias A e B. Os átomos foram nomeados automaticamente pelo software Discovery Studio Visualizer e a numeração dos resíduos respeita a numeração publicada originalmente para a estrutura utilizada. A estrutura foi inspecionada para verificar se as interações **7**-receptor eram mediadas por moléculas de água, com a conclusão de que, para este sítio, não são. Entretanto, dentro desta mesma estrutura cristalográfica é possível observar em outros sítios ortostéricos interações **7**-receptor mediadas por moléculas de água. Além disso, é possível constatar a presença de uma molécula de água na região do sítio ortostérico escolhido para o trabalho, de modo que apesar da mesma não interagir com o ligante co-cristalizado é interessante considerar a mesma durante os cálculos de ancoragem molecular. A posição e graus de liberdade permitidos para essa molécula de água estão descritas na metodologia. Os resíduos que participam de interações com **7** serão referidos daqui em diante como resíduos de interesse. As interações observadas para **7** co-cristalizado e o tipo das mesmas são os seguintes: B:W145 (cátion- π , hidrofóbica, empilhamento-T), B:Y184 (cátion- π , hidrofóbica), C:W53 (ligação de hidrogênio não clássica, empilhamento-T), C:L116, B:Y191, B:C186 e B:C187 (interações do tipo hidrofóbicas para todos esses últimos). O modo de ligação

no qual se encontra **7** co-cristalizado e os resíduos presentes na região estão representados na **Figura 26**.

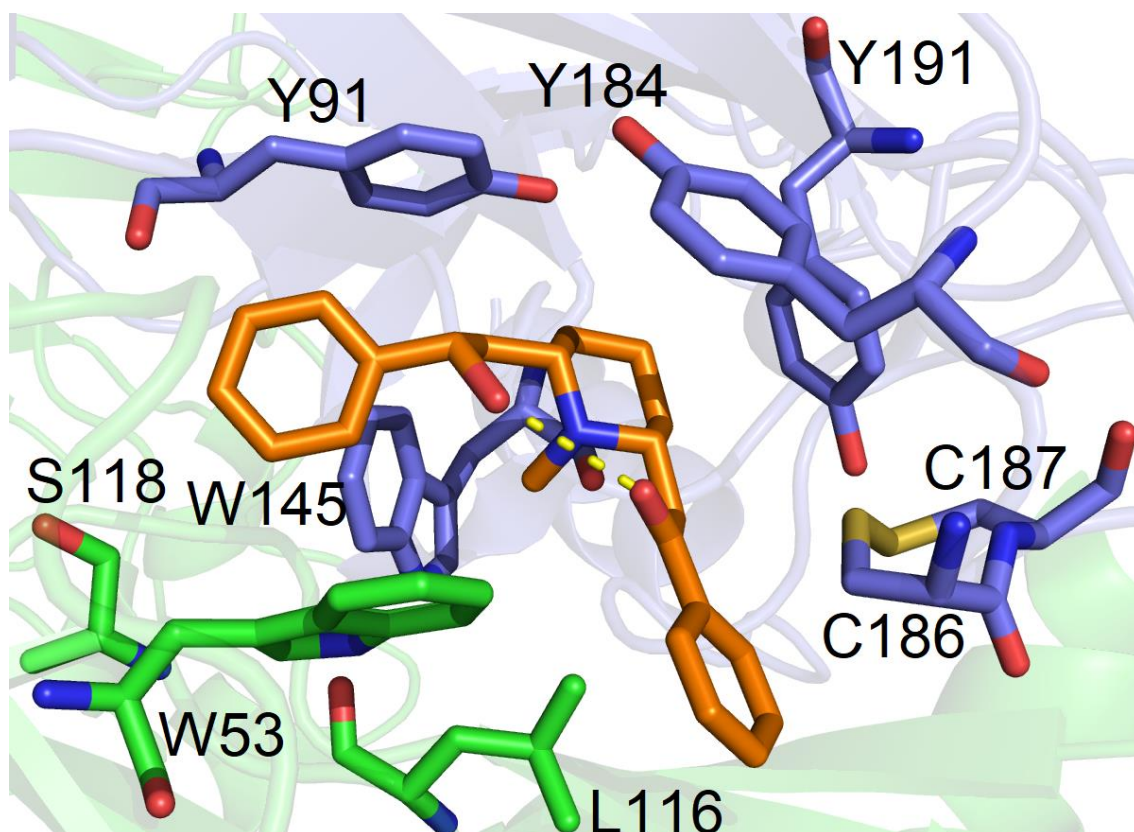


Figura 26 – Modo de ligação do ligante co-cristalizado **7** e resíduos da região do sítio ativo. Átomos de hidrogênio foram omitidos para maior clareza. Linhas tracejadas em amarelo representam átomos envolvidos em ligações de hidrogênio. Resíduos em azul fazem parte da cadeia B e os em verde da cadeia C.

Em seguida as 93 estruturas preparadas para a ancoragem molecular foram ancoradas no receptor de acordo com a metodologia descrita, em todas as condições de ancoragem citadas. Utilizando as poses e enantiômeros melhor pontuados para cada ligante, foram gerados diversos modelos dbCICA conforme previamente descrito. Os 10 modelos com maior *5-fold* r^2 estão descritos na **Tabela 7**. Os contatos críticos apontados pelos dois melhores modelos estão relacionados na **Tabela 8**.

Tabela 7 - Condições e estatística dos 10 melhores modelos dbCICA para o nAChRs do subtipo $\alpha 7$.

Modelo	Função de Pontuação	Água	Ionização	Distância dos contatos (Å)	Contatos (Positivos_Negativos)	5-fold r^2	Estatística F
nAChRs_ $\alpha 7$ _1	CHEMPLP	Não	Sim	2,5	3_10	0,582	134,37
nAChRs_ $\alpha 7$ _2	CHEMPLP	Não	Sim	3,5	3_10	0,534	118,08
nAChRs_ $\alpha 7$ _3	CHEMPLP	Não	Sim	2,5	10_10	0,530	114,28
nAChRs_ $\alpha 7$ _4	CHEMPLP	Sim	Sim	2,5	6_5	0,529	107,08
nAChRs_ $\alpha 7$ _5	CHEMPLP	Não	Sim	3,5	3_5	0,529	112,61
nAChRs_ $\alpha 7$ _6	CHEMPLP	Não	Sim	2,5	8_10	0,523	104,75
nAChRs_ $\alpha 7$ _7	CHEMPLP	Não	Sim	2,5	7_5	0,521	108,08
nAChRs_ $\alpha 7$ _8	CHEMPLP	Sim	Sim	2,5	8_10	0,516	102,42
nAChRs_ $\alpha 7$ _9	CHEMPLP	Sim	Sim	2,5	5_10	0,514	100,67
nAChRs_ $\alpha 7$ _10	CHEMPLP	Não	Sim	2,5	5_10	0,511	101,92

Tabela 8 - Contatos críticos dos 2 melhores modelos dbCICA para o nAChRs do subtipo $\alpha 7$.

	Contatos positivos (peso)	Contatos negativos
nAChRs_ $\alpha 7$ _1	B:Y184:CD1(2), B:Y191:CG(2), B:Y191:HD2(3)	C:A105:HA, C:L106:HG, B:W145:HA, C:W53:CZ2, B:Y184:CE1, B:Y184:CZ, B:Y184:HB2, B:Y191:CD1, B:Y191:HH, B:Y91:HB1
nAChRs_ $\alpha 7$ _2	B:K141:NZ(3), B:S144:O(2), B:Y184:HB2(2)	C:A105:CA, B:C186:HB1, B:C187:SG, C:L106:CA, B:W145:CA, C:W53:CZ3, C:Y115:N, B:Y91:C, B:Y91:CA, B:Y91:CE1

Observando os dados apresentados na **Tabela 7** se constata que a condição de ancoragem molecular que melhor explica a variação da bioatividade das moléculas do grupo de treinamento se deu quando utilizou-se a função CHEMPLP com os ligantes protonados em pH 7,4, sem a presença de água no sítio ativo e distância máxima de contatos igual a 2,5 Å, resultando em um 5-fold r^2 igual a 0,582 e estatística F igual a 134,37. A estatística 5-fold r^2 indica que aproximadamente 58% da variância observada para Y foi explicada através do modelo de regressão. Entende-se que o modelo é aceitável pelos mesmos motivos descritos anteriormente para a análise do modelo nAChRs_ $\alpha 4\beta 2$ _1. Quanto à estatística F para o modelo nAChRs_ $\alpha 7$ _1, o valor de F crítico para os graus de liberdade do modelo (1 e 93) é igual a 5,11, com $p = 1,29 \times 10^{-13}$, indicando bom ajuste.

No caso dos modelos construídos para este receptor, observa-se que a função CHEMPLP foi mais a mais eficiente, visto que todos os 10 melhores modelos calculados a utilizaram. O estado de ionização dos ligantes é extremamente importante, pois para todos os 10 modelos descritos na **Tabela 7** se observa essa característica. Quanto à presença de água se conclui que a mesma não é crítica para

o reconhecimento molecular, uma vez que 7 de 10 modelos não a utilizaram durante o cálculo. A análise dos contatos críticos apontados para o modelo nAChRs_α7_1 segue as mesmas premissas da análise para o modelo nAChRs_α4β2_1: Os contatos positivos mostrados representam interações envolvendo os resíduos do bolsão aromático do sítio ativo, onde átomos de nitrogênio positivamente carregados dos ligantes podem participar de interações cátion-π envolvendo tais resíduos. Além disso, tais resíduos podem participar de interações aromáticas de empilhamento-T e empilhamento-π quando o ligante possuir uma subunidade aromática, que são características de diversos ligantes dos nAChRs.

É interessante observar que no modelo nAChRs_α7_1 não aparecem contatos críticos envolvendo o grupo carbonila do resíduo B:W145, contato análogo ao contato A:W156:C do modelo nAChRs_α4β2_1, sendo este resíduo altamente conservado entre os nAChRs.²⁰⁵ Esta constatação corrobora com dados da literatura que indicam que o mecanismo pelo qual se dá a seletividade dos nAChRs do subtipo α4β2 frente ao subtipo α7 é decorrente da presença da ligação de hidrogênio supracitada para o primeiro receptor e ausência para o segundo. Devido à presença de um resíduo de lisina em uma região vizinha ao sítio ativo, se forma uma cadeia de ligações de hidrogênio que posicionam o triptofano do subtipo α4β2 de maneira que a interação mencionada seja favorecida. Para o subtipo α7 a lisina citada é substituída por uma glicina e estudos de dinâmica molecular e mutagênese sítio dirigida indicam que essa substituição faz com que a posição do triptofano em questão seja desfavorável para a formação dessa ligação de hidrogênio.^{10,206,207} Dessa forma, a validade do modelo é reforçada, uma vez que o mesmo foi capaz de indicar interações críticas bem documentadas na literatura científica, respeitando os mecanismos de seletividade já estabelecidos para esse conjunto de receptores.

O redocking da lobelina (**7**) foi realizado, utilizando as condições de ancoragem molecular obtidas para o modelo nAChRs_α7_1. A pose de ancoragem melhor pontuada para **7** está representada na **Figura 27**.

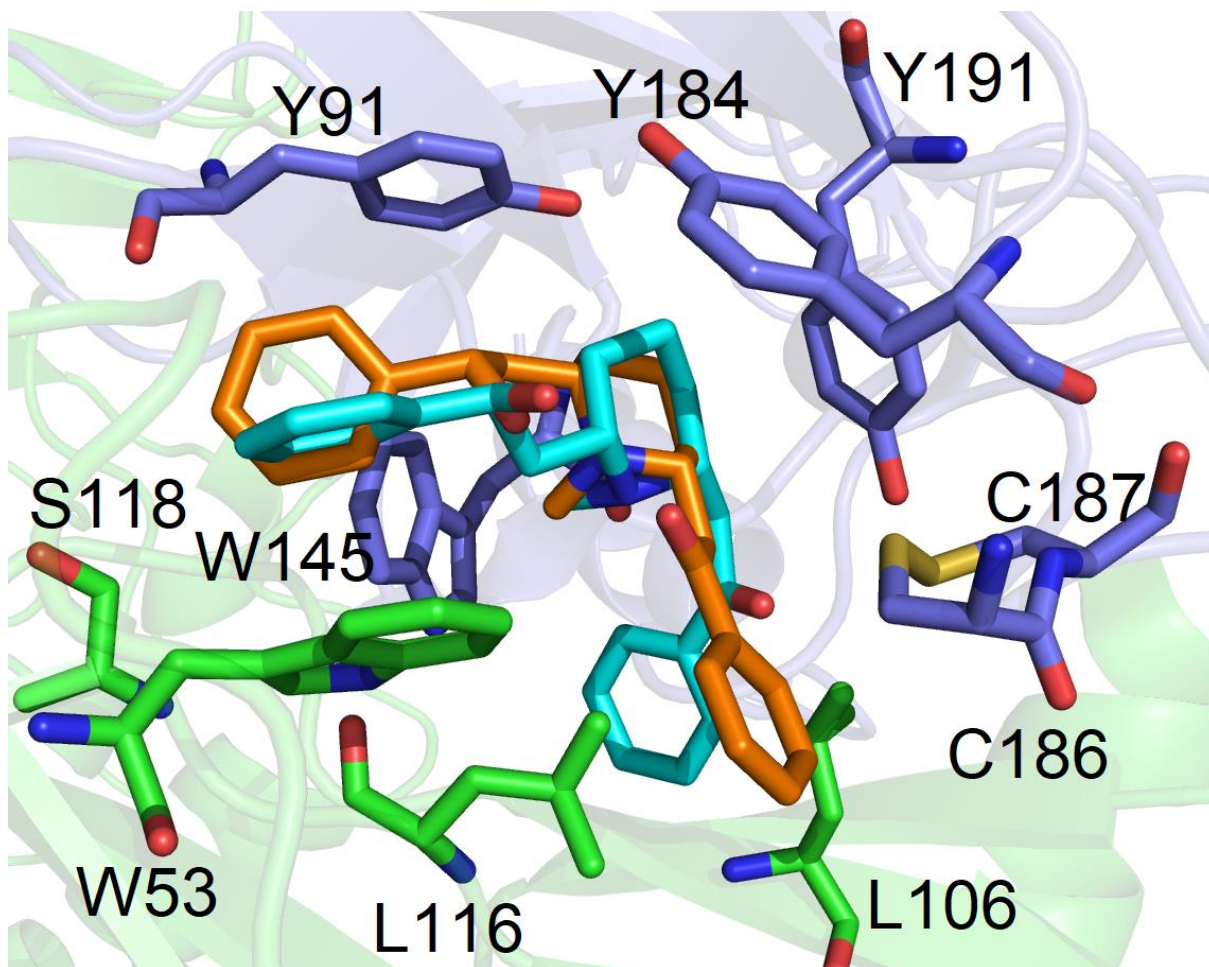


Figura 27 - Pose de **7** co-cristalizado (laranja) sobreposta à **7** ancorado (ciano) e resíduos da região do sítio ativo. RMSD = 3,45 Å. Átomos de hidrogênio foram omitidos para maior clareza. A ligação de hidrogênio intramolecular do ligante co-cristalizado foi omitida para maior clareza. Resíduos em azul fazem parte da cadeia B e os em verde da cadeia C.

A **Figura 27** mostra que o protocolo de ancoragem foi capaz de posicionar a região piperidínica de **7** corretamente, bem como a função hidroxila localizada à esquerda na figura. Para a pose ancorada a região aromática mostrada à esquerda se orienta de modo a maximizar interações aromáticas do tipo empilhamento- π envolvendo o resíduo C:W53, enquanto na pose co-cristalizada as interações envolvendo esse resíduo são do tipo empilhamento-T. Para **7** co-cristalizada, o anel aromático localizado à direita na **Figura 27** se orienta entre as cadeias laterais de C:L116 e B:C186, participando de contatos hidrofóbicos, enquanto para **7** ancorado esse mesmo anel se encontra entre os resíduos C:L116 e C:L106, também participando de interações hidrofóbicas. O oxigênio carbonílico diretamente ligado ao anel aromático, mostrado à direita na **Figura 27**, interage com a cadeia lateral do resíduo B:C187 em uma interação do tipo S-O. A pose de ancoragem de **7** apresenta

RMSD = 3,45 em relação ao ligante co-cristalizado, o que provavelmente se deve à elevada liberdade conformacional, com PLPscore = 81,6214. As interações observadas para essa pose são: B:W145 (ligação de hidrogênio não clássica), B:C187 (S-O), C:W53 (empilhamento- π), C:L106, C:L116, B:Y191 e B:Y184 (sendo estas do tipo hidrofóbico).

Quanto à validade das condições de ancoragem utilizadas para esse sistema, o redocking da lobelina traz algumas informações relevantes. Houve má concordância entre as regiões aromáticas da molécula ancorada em relação à molécula co-cristalizada. Entretanto, a molécula **7** apresenta elevada liberdade conformacional para esses anéis, sendo este o motivo para o alto valor de RMSD. É interessante ressaltar que a pose observada para **7** é consistente com a pose presente em uma estrutura cristalográfica publicada recentemente (PDB ID: 5OUH, resolução 2,5 Å)²⁰⁸ para o nAChRs do subtipo $\alpha 7$, conforme mostrado na **Figura 28b**, cuja sobreposição apresenta RMSD = 0,65. A proteína citada também é uma quimera do subtipo $\alpha 7$ humana com a AChBP da espécie *Lymnaea stagnalis*, com a introdução de mutações humanizantes adicionais em um sítio alostérico (F521I, V85K, L88I, A89L, A90L, S95E e K96R). Essas mutações resultam em pouca diferença na localização dos resíduos do sítio ativo, conforme mostrado na **Figura 28a**. No artigo publicado referente à essa estrutura, os autores chamam atenção para o fato de que o mapa de densidade eletrônica obtido dos dados de difração de raios-X é bem definido para a região piperidínica da molécula e menos proeminente para as regiões aromáticas devido à liberdade conformacional de **7**. A existência desses modos de ligação distintos para a lobelina em diferentes estruturas cristalográficas deixa clara a plasticidade do sítio ativo, que pode acomodar um mesmo ligante em diferentes poses apesar do arranjo semelhante dos resíduos aromáticos altamente conservados.²⁷ Constatados os fatos acima, acredita-se que a pose obtida para **7** utilizando os parâmetros de ancoragem do modelo nAChRs_ $\alpha 7$ _1 representa um modo de ligação plausível, confirmando a utilidade do modelo. Além disso, para a estrutura 5OUH não se observa a presença de moléculas de água no sítio ativo, corroborando as condições de ancoragem definidas como ótimas.

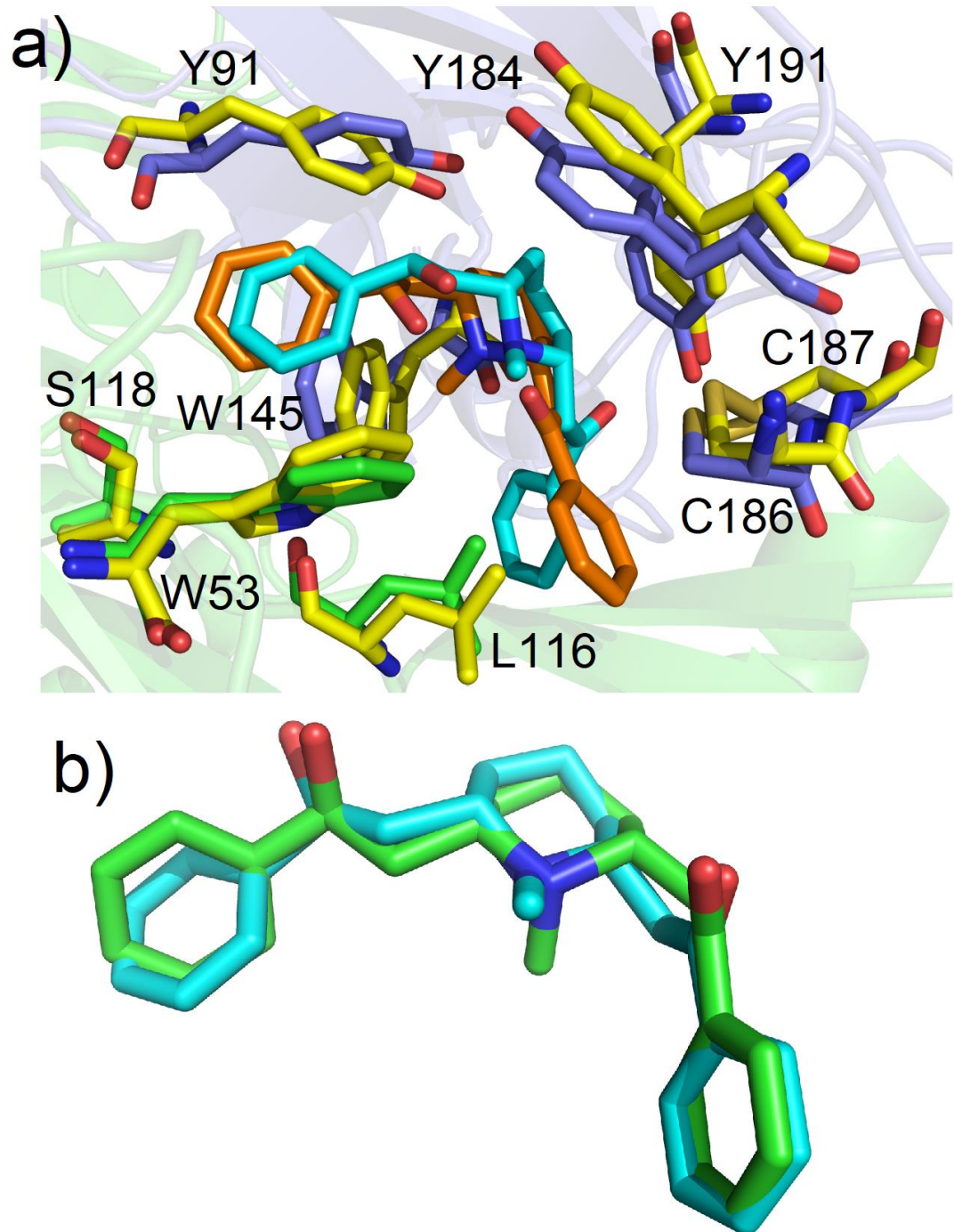


Figura 28 - a) Sobreposição dos resíduos dos sítios ativos e ligantes co-cristalizados das estruturas 5AFH (cadeia B em azul, cadeia C em verde, **7** em laranja) e 5OUH (amarelo, **7** em ciano). b) Sobreposição da pose de **7** co-cristalizada em 5OUH (ciano) e pose de ancoragem obtida para **7** utilizando os parâmetros do modelo nAChRs_α7_1 (verde), com RMSD = 0,65 Å.

Os 8 melhores fragmentos selecionados conforme descrito na metodologia e propriedades físico-químicas calculadas estão relacionados na **Figura 29**. As interações previstas pelo cálculo para cada um deles estão contidas na **Tabela 9**.

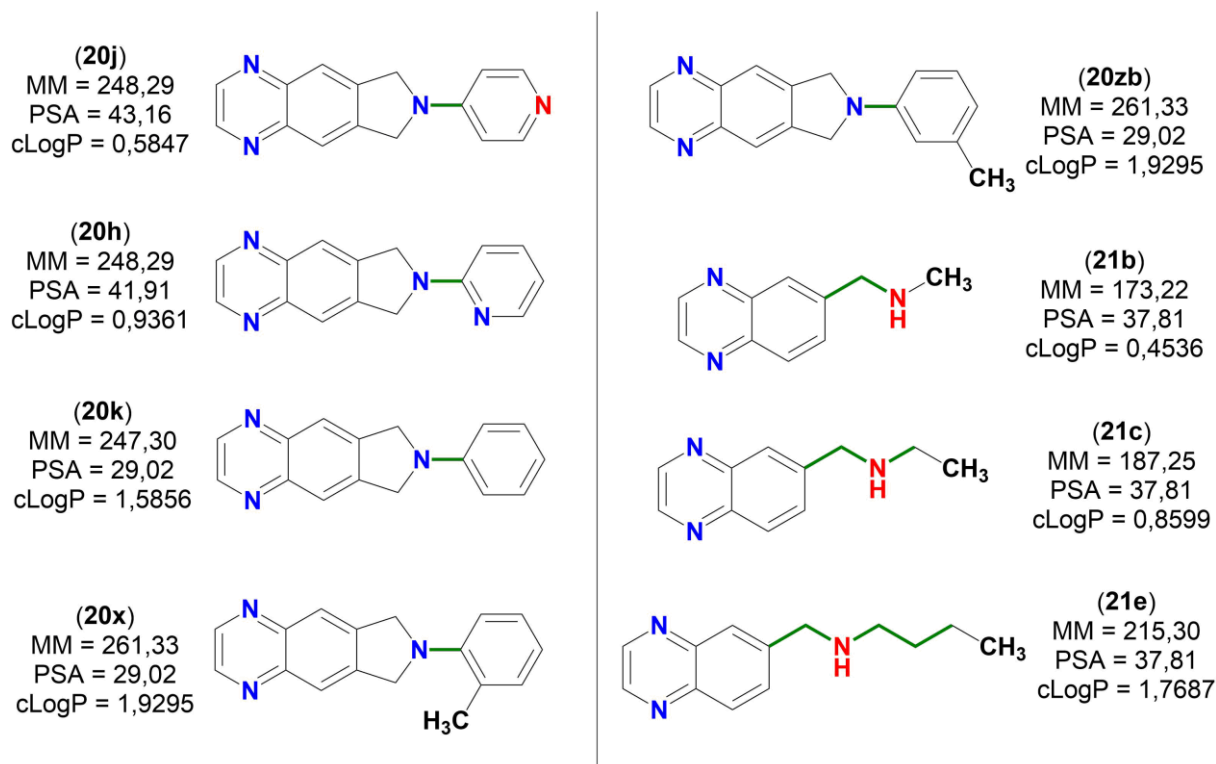


Figura 29 - Fórmula estrutural, massa molecular (MM), área de superfície polar (PSA) e coeficiente de partição (cLogP) dos 8 fragmentos selecionados através da triagem virtual para o receptor $\alpha 7$. Átomos em vermelho são doadores de ligação de hidrogênio, átomos em azul aceptores de ligação de hidrogênio e ligações químicas em rosa indicam rotação livre. Nitrogênios piridínicos em vermelho indicam que os mesmos foram previstos como protonados em pH 7,4 pelo software MarvinSketch. As propriedades físico-químicas foram calculadas nos softwares MarvinSketch (MM e PSA) e OSIRIS DataWarrior (cLogP).

Tabela 9 - Interações observadas para **7** em 5AFH, interações previstas para **7** e interações previstas para os 8 fragmentos selecionados pelo modelo nAChRs_ $\alpha 7$ _1. À direita, soma dos contatos críticos apontados no modelo nAChRs_ $\alpha 7$ _1. Resíduos em vermelho são os de interesse.

Ligante	PLP score	W53	W145	L116	Y184	Y191	C186	C187	Y91	L36	D160	S144	L106	Q55	Soma
7 (5AFH)	-														
7 (docking)	81,6214														
20j	57,3895														11
20h	55,6808														10
20k	56,5373														10
20x	60,6296														10
20zb	57,8204														10
21b	46,1190														10
21c	48,9445														10
21e	56,5060														10

Interações: ■ Ligação de hidrogênio com átomos da ligação peptídica ■ Ligação de hidrogênio com átomos da cadeia lateral do resíduo ■ Ligação de hidrogênio não clássica ■ π -Enxofre ■ Empilhamento- π ■ Cátion- π ■ Empilhamento-T ■ Hidrofóbica

Observando os dados mostrados na **Tabela 9** é possível constatar que as moléculas selecionadas interagem com os resíduos de interesse. Para a triagem virtual dessa série de moléculas, se observa que os fragmentos da série **20**, quinoxalinas conformacionalmente restrita, foram favorecidas pelo modelo, conforme mostrado na **Figura 29**. Somente as três últimas moléculas selecionadas, **21b**, **21c** e **21e** pertencem à classe de fragmentos conformacionalmente livres. As quinoxalinas obtiveram melhor resultado do que os benzimidazóis para esse receptor. Para essa série de fragmentos, se optou por priorizar compostos participando de empilhamentos aromáticos com os resíduos da região aromática do sítio ativo. Os fragmentos selecionados não participam de interações cátion- π com os resíduos da região aromática, com exceção de **20j**, **21c** e **21e**, uma interação considerada crítica para o reconhecimento molecular dessa classe de receptores. Isso ocorre devido às microespécies majoritárias calculadas para pH = 7,4 não serem protonadas para a maioria dos fragmentos planejados. Para as moléculas selecionadas, somente **20j**, **21b**, **21c** e **21e** estariam positivamente carregadas em pH intracelular de acordo com o cálculo realizado através do software MarvinSketch. Quanto às propriedades físico-químicas calculadas, todos os fragmentos se enquadram na restrição de massa imposta pela RO3 de Lipinsky, no critério de doadores de ligação de hidrogênio e também na restrição de área de superfície polar. O fragmento **20h** é o único que contém mais de 3 aceptores de ligação de hidrogênio, enquanto o fragmento **21e** é o único que possui mais de 3 ligações com rotação livre. É interessante observar que o fragmento **20j** foi selecionado para ambos os subtipos dos nAChRs trabalhados, o que pode significar que ele é um potencial candidato para a atividade dual. Nenhuma das demais moléculas selecionadas pelo modelo nAChRs_ α 7_1 foi selecionada pelo modelo nAChRs_ α 4 β 2_1 e vice versa, o que pode significar que as mesmas são seletivas para seus respectivos alvos. Entretanto, constata-se através da comparação da **Tabela 9** com a **Tabela 6** que as pontuações relacionadas para os fragmentos selecionados na primeira divergem muito em relação à pontuação de **7**, enquanto as pontuações mostradas para os fragmentos na segunda tabela estão muito mais próximos da pontuação calculada para **3**, o que pode significar que as moléculas selecionadas para o subtipo α 7 possuem interações menos favoráveis e, conseqüentemente, podem vir a ser menos ativas para esse alvo.

A sobreposição dos modos de ligação preditos para os fragmentos descritos na **Figura 29** e **Tabela 9** estão representados na **Figura 30**.

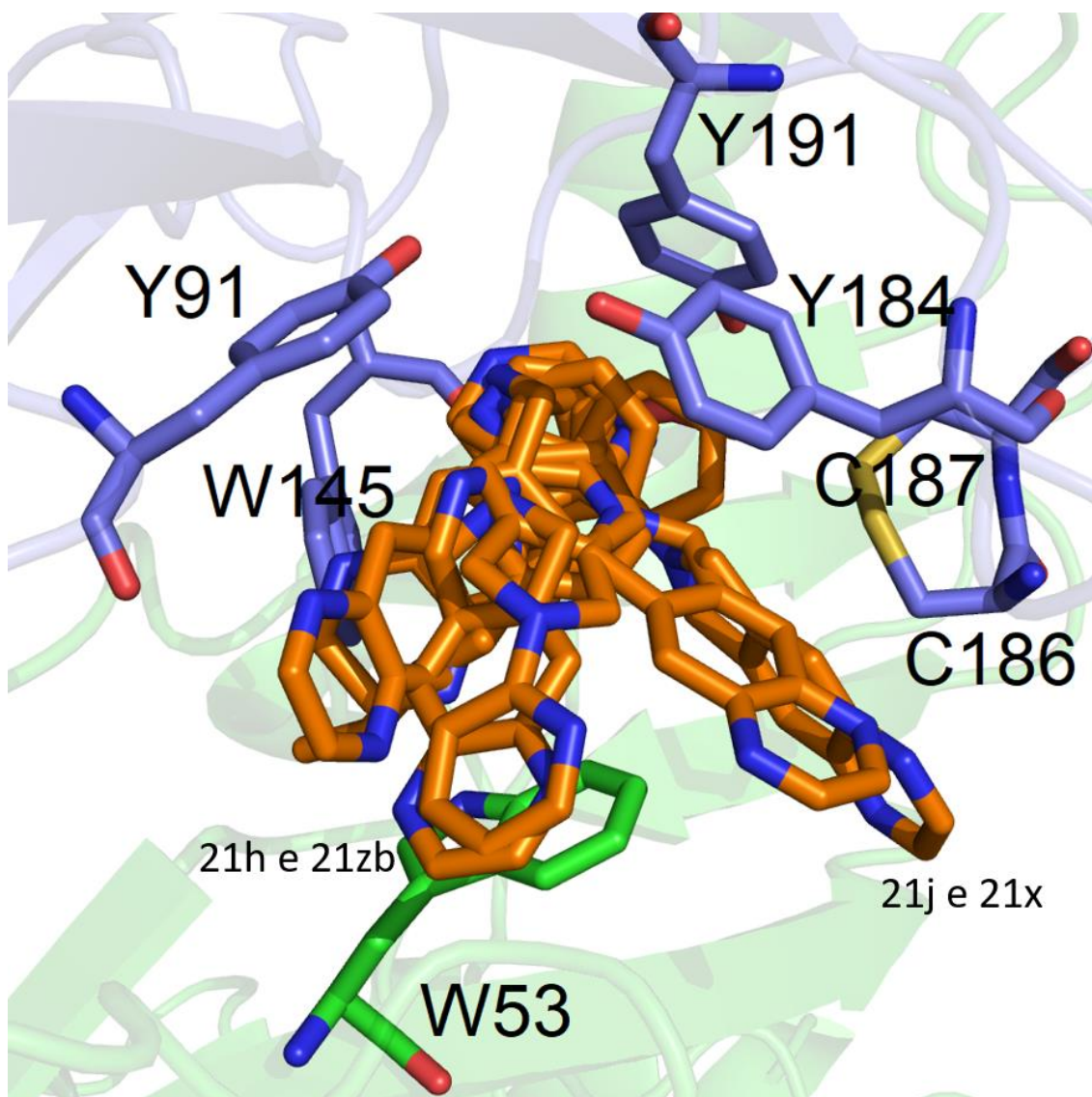


Figura 30 - Sobreposição das poses melhor ranqueadas para os 8 fragmentos moleculares triados para o subtipo $\alpha 7$. Átomos de hidrogênio foram omitidos para maior clareza. Resíduos em azul fazem parte da cadeia B e os em verde da cadeia C.

A sobreposição mostrada na **Figura 30** mostra que os fragmentos **20j** e **20x** se orientam de modo que a subunidade quinoxalina esteja fora do bolsão aromático do sítio ativo, enquanto as subunidades aromáticas ligadas ao nitrogênio terciário se posicionam no interior do bolsão e participam de interações aromáticas. Os fragmentos **20h** e **20zb** estão posicionados em direções opostas entre si: para o primeiro, o anel quinoxalínico se encontra no centro do bolsão aromático enquanto para o segundo ele se orienta sobre a cadeia lateral do resíduo C:W145. Os demais

fragmentos se encontram no interior do bolsão aromático, onde é possível que participem de interações de empilhamento com as cadeias laterais dos resíduos da região. A má sobreposição das poses de ancoragem dos 8 fragmentos selecionados pode estar relacionada com a já mencionada plasticidade do sítio ativo ou pode ser uma implicação das afirmações feitas acima quanto à pontuação calculada pela função CHEMPLP.

É possível construir um mapa farmacofórico para o modelo nAChRs_α7_1 conforme descrito na metodologia deste trabalho. Entretanto, o mapa não foi construído, visto que o foco do trabalho são os nAChRs do subtipo α4β2, sendo esta uma das etapas a serem exploradas para entender a seletividade.

5. Conclusões

Através da metodologia dbCICA foi possível selecionar dentre 116 fragmentos moleculares os 8 mais promissores para a etapa de síntese visando o nAChR do subtipo $\alpha 4\beta 2$. As condições de ancoragem molecular ideais obtidas para o melhor modelo nAChRs_ $\alpha 4\beta 2$ _1 foram: água explícita no sítio ativo, ligantes protonados para pH 7,4, 4 contatos positivos e 10 contatos negativos, assim como 3,5 Å de distância máxima para a existência de um contato. Utilizando essas condições foi possível obter RMSD = 0,83 Å para a redocagem da nicotina (**3**) e também uma pose com boa concordância frente à um modo de ligação determinado experimentalmente para a vareniclina (**13**) em uma AChBP homóloga proveniente da espécie *Aplysia californica*. O modelo computacional nAChRs_ $\alpha 4\beta 2$ _1 apresentou bom desempenho quanto às suas predições de contatos críticos, sendo que em sua maioria tais contatos puderam ser corroborados através de comparação com estudos existentes na literatura científica. Quanto às interações previstas para os 8 fragmentos selecionados, em sua maioria houve concordância quanto às interações com os resíduos de interesse, especialmente frente ao resíduo A:W156, que é de extrema importância para o reconhecimento molecular de ligantes dos nAChRs $\alpha 4\beta 2$ através de ligações de hidrogênio. Os dados obtidos através do dbCICA foram utilizados na construção de um mapa farmacofórico para o modelo nAChRs_ $\alpha 4\beta 2$ _1, validado através de curva ROC, com AUC = 0,86, Ya = 0,25 e ACC = 0,90.

Para o melhor modelo do subtipo $\alpha 7$ as condições ideais identificadas para a ancoragem foram: água ausente no sítio ativo, ligantes protonados para pH 7,4, 3 contatos positivos, 10 contatos negativos e 2,5 Å de distância máxima para a existência de um contato. Utilizando essas condições, obteve-se RMSD = 3,45 Å para o redocking da lobelina (**7**), o que se deve à elevada liberdade conformacional da molécula citada. A pose de ancoragem obtida para **7** apresenta boa concordância frente ao modo de ligação experimentalmente obtido de uma estrutura cristalográfica relacionada, com RMSD = 0,65 Å. Utilizando as condições ótimas, 8 dos 116 fragmentos moleculares foram selecionados de maneira análoga à triagem feita para o subtipo $\alpha 4\beta 2$, visando a verificação de possível seletividade. É interessante observar que os contatos críticos apontados por ambos os modelos dos nAChRs são consistentes quanto aos mecanismos de seletividade já descritos na literatura científica para esses dois receptores. Em uma etapa futura os dados obtidos para o

modelo nAChRs_α7_1 podem ser utilizados para a construção de um mapa farmacofórico, conforme feito para o modelo nAChRs_α4β2_1.

Frente às informações supracitadas se conclui que os modelos computacionais construídos foram adequadamente validados e que as moléculas selecionadas para o subtipo α4β2 são ligantes em potencial desse receptor. Sendo assim, as mesmas serão sintetizadas nas próximas etapas do trabalho. Após a síntese, caracterização e purificação de quantidade adequada das substâncias, as mesmas serão submetidas à ensaios farmacológicos *in vitro* e também ensaios de atividade funcional. Em seguida, as moléculas mais promissoras serão avaliadas em ensaios de estabilidade química e metabólica, bem como em ensaios de toxicidade *in vivo*. Caso um *hit* com atividade submicromolar seja identificado, é possível utilizá-lo em estudos de cristalografia de raios-X visando entender o reconhecimento molecular e também otimizá-lo através da abordagem *fragment growth*, utilizando o mapa farmacofórico já construído. Um estudo de dinâmica molecular também é interessante visando entender as mudanças conformacionais observadas nos nAChRs frente à ação agonista, antagonista e agonista parcial. Além disso, sabe-se que os nAChRs possuem sítios alostéricos onde atuam PAMs e NAMs, sendo a modulação destes sítios é de extremo interesse terapêutico, e dessa forma a construção de modelos computacionais para os mesmos pode ser útil na busca por novos candidatos à fármacos.

6. Referências

1. DANI, J. A. & BERTRAND, D. Nicotinic Acetylcholine Receptors and Nicotinic Cholinergic Mechanisms of the Central Nervous System. *Annual Review of Pharmacology and Toxicology*, 47, 699-729 (2007).
2. ALBUQUERQUE, E. X. *et al.* Mammalian Nicotinic Acetylcholine Receptors: From Structure to Function. *Physiological Reviews*, 89, 73-120 (2009).
3. GOTTI, C. *et al.* Brain nicotinic acetylcholine receptors: native subtypes and their relevance. *Trends in Pharmacological Sciences*, 27, 482-491 (2006).
4. HÖGGER, P. *et al.* The muscarinic acetylcholine receptors. in *Rhodopsin and G-protein Linked Receptors*, 2, 301–320 (JAI, 1996).
5. HOGG, R. C. *et al.* Nicotinic acetylcholine receptors: from structure to brain function. *Reviews of Physiology, Biochemistry and Pharmacology*, 147, 1–46 (2003).
6. QUIK, M. *et al.* Alpha7 nicotinic receptors as therapeutic targets for Parkinson's disease. *Biochemical Pharmacology*, 97, 399-407 (2015).
7. GOTTI, C. *et al.* Structural and functional diversity of native brain neuronal nicotinic receptors. *Biochemical Pharmacology*, 78, 703-11 (2009).
8. DINELEY, K. T. *et al.* Nicotinic ACh receptors as therapeutic targets in CNS disorders. *Trends in Pharmacological Sciences*, 36, 96–108 (2015).
9. UNWIN, N. Refined structure of the nicotinic acetylcholine receptor at 4 Å resolution. *Journal of Molecular Biology*, 346, 967-989 (2005).
10. GOTTI, C. & CLEMENTI, F. Neuronal nicotinic receptors: from structure to pathology. *Progress in Neurobiology*, 74, 363-396 (2004).
11. STOKES, C. *et al.* Looking below the surface of nicotinic acetylcholine receptors. *Trends in Pharmacological Sciences*, 36, 514-523 (2015).
12. AUERBACH, A. Activation of endplate nicotinic acetylcholine receptors by agonists. *Biochemical Pharmacology*, 97, 601-608 (2015).
13. DAJAS-BAILADOR, F. & WONNACOTT, S. Nicotinic acetylcholine receptors and the regulation of neuronal signalling. *Trends in Pharmacological Sciences*, 25, 317-324 (2004).
14. LAX, P. *et al.* Ca²⁺ permeability of human heteromeric nAChRs expressed by transfection in human cells. *Cell Calcium*, 32, 53-8 (2002).
15. GOTTI, C. *et al.* Heterogeneity and complexity of native brain nicotinic receptors. *Biochemical Pharmacology*, 74, 1102-1111 (2007).
16. R., H. *et al.* Nicotinic acetylcholine receptors: From basic science to therapeutics. *Pharmacology and Therapeutics*, 137, 22-54 (2013).
17. SON, C. D. *et al.* Nicotine Normalizes Intracellular Subunit Stoichiometry of Nicotinic Receptors Carrying Mutations Linked to Autosomal Dominant Nocturnal Frontal Lobe Epilepsy. *Molecular Pharmacology*, 75, 1137-1148 (2009).
18. CHATZIDAKI, A. & MILLAR, N. S. Allosteric modulation of nicotinic acetylcholine receptors. *Biochemical Pharmacology*, 97, 408-417 (2015).
19. DEUTSCH, S. I. *et al.* The $\alpha 7$ nicotinic acetylcholine receptor: A mediator of pathogenesis and therapeutic target in autism spectrum disorders and Down syndrome. *Biochemical Pharmacology*, 97, 363-377 (2015).
20. LOMBARDO, S. & MASKOS, U. Role of the nicotinic acetylcholine receptor in Alzheimer's disease pathology and treatment. *Neuropharmacology*, 96B, 255-262 (2015).
21. WALLACE, T. L. & BERTRAND, D. Neuronal alpha7 Nicotinic Receptors as a

- Target for the Treatment of Schizophrenia. *International Review of Neurobiology*, 124, 79–111 (2015).
22. KUTLU, M. G. *et al.* Nicotine Addiction and Psychiatric Disorders. *International Review of Neurobiology*, 124, 171–208 (2015).
 23. QUIK, M. *et al.* Nicotine and Nicotinic Receptor Drugs: Potential for Parkinson's Disease and Drug-Induced Movement Disorders. in *International Review of Neurobiology*, 124, 247-271 (2015).
 24. JENSEN, A. A. *et al.* Neuronal nicotinic acetylcholine receptors: Structural revelations, target identifications, and therapeutic inspirations. *Journal of Medicinal Chemistry*, 48, 4705-4745 (2005).
 25. BILLEN, B. *et al.* Molecular actions of smoking cessation drugs at $\alpha 2$ nicotinic receptors defined in crystal structures of a homologous binding protein. *Proceedings of the National Academy of Sciences*, (2012).
 26. BOURNE, Y. *et al.* Marine Macrocyclic Imines, Pinnatoxins A and G: Structural Determinants and Functional Properties to Distinguish Neuronal $\alpha 7$ from Muscle $\alpha 2\beta\gamma\delta$ nAChRs. *Structure*, 23, 1106-1115 (2015).
 27. SPURNY, R. *et al.* Molecular blueprint of allosteric binding sites in a homologue of the agonist-binding domain of the $\alpha 7$ nicotinic acetylcholine receptor. *Proceedings of the National Academy of Sciences*, 112, E2543-E2552 (2015).
 28. PAPKE, R. L. Merging old and new perspectives on nicotinic acetylcholine receptors. *Biochemical Pharmacology*, 89, 1-11 (2014).
 29. MORALES-PEREZ, C. L. *et al.* X-ray structure of the human $\alpha 4\beta 2$ nicotinic receptor. *Nature*, 538, 411-415 (2016).
 30. WALSH, R. M. *et al.* Structural principles of distinct assemblies of the human $\alpha 4\beta 2$ nicotinic receptor. *Nature*, 557, 261-265 (2018).
 31. KOHLHAAS, K. L. *et al.* Nicotinic modulation of auditory evoked potential electroencephalography in a rodent neurodevelopmental model of schizophrenia. *Biochemical Pharmacology*, 97, 482-487 (2015).
 32. WANG, J. *et al.* An accessory agonist binding site promotes activation of $\alpha 4\beta 2^*$ nicotinic acetylcholine receptors. *Journal of Biological Chemistry*, 290, 13907-13918 (2015).
 33. CHENG, Q. & YAKEL, J. L. The effect of $\alpha 7$ nicotinic receptor activation on glutamatergic transmission in the hippocampus. *Biochemical Pharmacology*, 97, 439-444 (2015).
 34. KOUKOULI, F. & MASKOS, U. The multiple roles of the $\alpha 7$ nicotinic acetylcholine receptor in modulating glutamatergic systems in the normal and diseased nervous system. *Biochemical Pharmacology*, 97, 378-387 (2015).
 35. ZENG, H. *et al.* The Solution Structure of the Complex Formed between α -Bungarotoxin and an 18-mer Cognate Peptide Derived from the $\alpha 1$ Subunit of the Nicotinic Acetylcholine Receptor from *Torpedo californica*. *Journal of Biological Chemistry*, 276, 22930–22940 (2001).
 36. STEAD LINDSAY, F. & HUGHES JOHN, R. Lobeline for smoking cessation. *Cochrane Database of Systematic Reviews*, 2, CD000124 (2012).
 37. HOJAHMAT, M. *et al.* Lobeline esters as novel ligands for neuronal nicotinic acetylcholine receptors and neurotransmitter transporters. *Bioorganic & Medicinal Chemistry*, 18, 640–649 (2010).
 38. MAZUROV, A. A. *et al.* Discovery and Development of $\alpha 7$ Nicotinic Acetylcholine Receptor Modulators. *Journal of Medicinal Chemistry*, 54, 7943–7961 (2011).

39. <https://adisinsight.springer.com/drugs/800026693> (Acesso em Fevereiro de **2019**).
40. GOTTI, C. *et al.* 4-Oxystilbene compounds are selective ligands for neuronal nicotinic α 7 receptors. *British Journal of Pharmacology*, 124, 1197–1206 (**1998**).
41. DI ANGELANTONIO, S. *et al.* Antagonism of nicotinic receptors of rat chromaffin cells by N,N,N-trimethyl-1-(4-trans-stilbenoxy)-2-propylammonium iodide: a patch clamp and ligand binding study. *British Journal of Pharmacology*, 129, 1771–1779 (**2000**).
42. GOTTI, C. *et al.* Drugs selective for nicotinic receptor subtypes: a real possibility or a dream? *Behavioural Brain Research*, 113, 183–192 (**2000**).
43. ZHENG, G. *et al.* Discovery of non-peptide, small molecule antagonists of α 9 α 10 nicotinic acetylcholine receptors as novel analgesics for the treatment of neuropathic and tonic inflammatory pain. *Bioorganic & medicinal chemistry letters*, 21, 2476–2479 (**2011**).
44. WHITEAKER, P. *et al.* Discovery, Synthesis, and Structure Activity of a Highly Selective α 7 Nicotinic Acetylcholine Receptor Antagonist. *Biochemistry*, 46, 6628–6638 (**2007**).
45. INNOCENT, N. *et al.* Alpha-conotoxin Arenatus IB[V11L,V16D] [corrected] is a potent and selective antagonist at rat and human native α 7 nicotinic acetylcholine receptors. *The Journal of Pharmacology and Experimental Therapeutics*, 327, 529–537 (**2008**).
46. MAZZAFERRO, S. *et al.* Additional acetylcholine (ACh) binding site at α 4/ α 4 interface of (α 4 β 2) α 4 nicotinic receptor influences agonist sensitivity. *The Journal of Biological Chemistry*, 286, 31043–31054 (**2011**).
47. NELSON, M. E. Alternate Stoichiometries of α 4 β 2 Nicotinic Acetylcholine Receptors. *Molecular Pharmacology*, 63, 332–341 (**2003**).
48. NIAURA, R. *et al.* Varenicline. *Nature Reviews Drug Discovery*, 5, 537–538 (**2006**).
49. HARPSOE, K. *et al.* Unraveling the High- and Low-Sensitivity Agonist Responses of Nicotinic Acetylcholine Receptors. *Journal of Neuroscience*, 31, 10759–10766 (**2011**).
50. GRUPE, M. *et al.* Targeting α 4 β 2 nicotinic acetylcholine receptors in central nervous system disorders: Perspectives on positive allosteric modulation as a therapeutic approach. *Basic and Clinical Pharmacology and Toxicology*, 116–187–200 (**2015**).
51. GERZANICH, V. *et al.* Comparative pharmacology of epibatidine: a potent agonist for neuronal nicotinic acetylcholine receptors. *Molecular Pharmacology*, 48, 774–782 (**1995**).
52. ULENS, C. *et al.* Use of acetylcholine binding protein in the search for novel α 7 nicotinic receptor ligands. *in silico* docking, pharmacological screening, and x-ray analysis. *Journal of Medicinal Chemistry*, 52, 2372–2383 (**2009**).
53. DALLANOCE, C. *et al.* Synthesis and binding affinity at α 4 β 2 and α 7 nicotinic acetylcholine receptors of new analogs of epibatidine and epiboxidine containing the 7-azabicyclo[2.2.1]hept-2-ene ring system. *Bioorganic & Medicinal Chemistry Letters*, 22, 829–832 (**2012**).
54. MU, L. *et al.* Synthesis and binding studies of epibatidine analogues as ligands for the nicotinic acetylcholine receptors. *European Journal of Medicinal Chemistry*, 41, 640–650 (**2006**).

55. CARROLL, F. I. Epibatidine structure-activity relationships. *Bioorganic & Medicinal Chemistry Letters*, 14, 1889–1896 (2004).
56. COE, J. W. *et al.* Varenicline: An $\alpha_4\beta_2$ nicotinic receptor partial agonist for smoking cessation. *Journal of Medicinal Chemistry*, 48, 3474–3477 (2005).
57. CIPPITELLI, A. *et al.* AT-1001: a high-affinity $\alpha_3\beta_4$ nAChR ligand with novel nicotine-suppressive pharmacology. *British Journal of Pharmacology*, 172, 1834–1845 (2015).
58. ARIAS, H. R. *et al.* Pharmacological and molecular studies on the interaction of varenicline with different nicotinic acetylcholine receptor subtypes. Potential mechanism underlying partial agonism at human $\alpha_4\beta_2$ and $\alpha_3\beta_4$ subtypes. *Biochimica et Biophysica Acta*, 1848, 731–741 (2015).
59. STRACHAN, J.-P. *et al.* Identification and pharmacological characterization of 3,6-diazabicyclo[3.1.1]heptane-3-carboxamides as novel ligands for the $\alpha_4\beta_2$ and $\alpha_6/\alpha_3\beta_2\beta_3$ nicotinic acetylcholine receptors (nAChRs). *European Journal of Medicinal Chemistry*, 86, 60–74 (2014).
60. JOHNSTON, T. H. *et al.* TC-8831, a nicotinic acetylcholine receptor agonist, reduces L-DOPA-induced dyskinesia in the MPTP macaque. *Neuropharmacology*, 73, 337–347 (2013).
61. ZHANG, D. *et al.* Nicotinic receptor agonists reduce L-DOPA-induced dyskinesias in a monkey model of Parkinson's disease. *The Journal of Pharmacology and Experimental Therapeutics*, 347, 225–234 (2013).
62. FAUNDEZ-PARRAGUEZ, M. *et al.* Neonicotinic analogues: selective antagonists for $\alpha_4\beta_2$ nicotinic acetylcholine receptors. *Bioorganic & medicinal chemistry*, 21, 2687–2694 (2013).
63. CUCCHIARO, G. *et al.* Analgesic effects of Sazetidine-A, a new nicotinic cholinergic drug. *Anesthesiology*, 109, 512–519 (2008).
64. KOZIKOWSKI, A. P. *et al.* Chemistry and pharmacology of nicotinic ligands based on 6-[5-(azetidino-2-ylmethoxy)pyridin-3-yl]hex-5-yn-1-ol (AMOP-H-OH) for possible use in depression. *ChemMedChem*, 4, 1279–1291 (2009).
65. XIAO, Y. *et al.* Pharmacological properties of sazetidine A, a selective ligand of $\alpha_4\beta_2$ nicotinic acetylcholine receptors. *Biochemical Pharmacology*, 82, 1029 (2011).
66. ZWART, R. *et al.* Sazetidine-A is a potent and selective agonist at native and recombinant $\alpha_4\beta_2$ nicotinic acetylcholine receptors. *Molecular Pharmacology*, 73, 1838–1843 (2008).
67. IMMING, P. Chapter 1 - Medicinal Chemistry: Definitions and Objectives, Drug Activity Phases, Drug Classification Systems. in (eds. Wermuth, C. G. *et al.*) 3–13 (Academic Press, 2015).
68. VAN NORMAN, G. A. Drugs, Devices, and the FDA: Part 1: An Overview of Approval Processes for Drugs. *JACC: Basic to Translational Science*, 1, 170–179 (2016).
69. DIMASI, J. A. *et al.* Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, 47, 20–33 (2016).
70. LAVECCHIA, A. & Giovanni, C. Virtual Screening Strategies in Drug Discovery: A Critical Review. *Current Medicinal Chemistry*, 20, 2839–2860 (2013).
71. LIONTA, E. *et al.* Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Current Topics in Medicinal Chemistry*, 14, 1923–1938 (2014).
72. MOITESSIER, N. *et al.* Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. *British Journal of*

- Pharmacology*, 153 Suppl, S7-26 (2008).
73. HOPKINS, A. L. & GROOM, C. R. The druggable genome. *Nature Reviews Drug Discovery*, 1, 727-730 (2002).
 74. CHENG, T. *et al.* Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. *The AAPS Journal*, 14, 133-141 (2012).
 75. MANDAL, S. *et al.* Rational drug design. *European Journal of Pharmacology*, 625, 90-100 (2009).
 76. ANDRICOPULO, A. *et al.* Structure-Based Drug Design Strategies in Medicinal Chemistry. *Current Topics in Medicinal Chemistry*, 9, 771-790 (2009).
 77. DE RUYCK, J. *et al.* Molecular docking as a popular tool in drug design, an in silico travel. *Advances and Applications in Bioinformatics and Chemistry*, 9, 1-11 (2016).
 78. MENG, X.-Y. *et al.* Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Current Computer Aided-Drug Design*, 7, 146-157 (2011).
 79. SALMASO, V. & MORO, S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. *Frontiers in Pharmacology*, 9:923 (2018).
 80. FERREIRA, L. G. *et al.* Molecular docking and structure-based drug design strategies. *Molecules*, 20, 13384-13421 (2015).
 81. FRIESNER, R. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*, 47, 1739-1749 (2004).
 82. HALGREN, T. A. *et al.* Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry*, 47, 1750-1759 (2004).
 83. RAREY, M. *et al.* A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology*, 261, 470-489 (1996).
 84. WELCH, W. *et al.* Hammerhead: Fast, fully automated docking of flexible ligands to protein binding sites. *Chemistry and Biology*, 3, 449-462 (1996).
 85. MILLER, M. D. *et al.* FLOG: A system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure. *Journal of Computer-Aided Molecular Design*, 8, 153-174 (1994).
 86. PANG, Y. P. *et al.* EUDOC: A computer program for identification of drug interaction sites in macromolecules and drug leads from chemical databases. *Journal of Computational Chemistry*, 22, 1750-1771 (2001).
 87. SAUTON, N. *et al.* MS-DOCK: Accurate multiple conformation generator and rigid docking protocol for multi-step virtual ligand screening. *BMC Bioinformatics*, 9:184 (2008).
 88. MORRIS, G. M. *et al.* Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. *Journal of Computer-Aided Molecular Design*, 10, 293-304 (1996).
 89. ABAGYAN, R. *et al.* ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *Journal of Computational Chemistry*, 15, 488-506 (1994).
 90. MCMARTIN, C. & BOHACEK, R. S. QXP: Powerful, rapid computer algorithms for structure-based drug design. *Journal of Computer-Aided Molecular Design*, 11, 333-344 (1997).
 91. LIU, M. & WANG, S. MCDOCK: A Monte Carlo simulation approach to the molecular docking problem. *Journal of Computer-Aided Molecular Design*, 435-

- 451 (1999).
92. TROTT, O. & OLSON, A. J. Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31, 455-461 (2010).
 93. MEILER, J. & BAKER, D. ROSETTALIGAND: Protein-small molecule docking with full side-chain flexibility. *Proteins: Structure, Function and Genetics*, 65, 538-548 (2006).
 94. BAXTER, C. A. *et al.* Flexible docking using Tabu search and an empirical estimate of binding affinity. *Proteins: Structure, Function and Genetics*, 33, 367-382 (1998).
 95. PEI, J. *et al.* PSI-DOCK: Towards highly efficient and accurate flexible ligand docking. *Proteins: Structure, Function and Genetics*, 62, 934-946 (2006).
 96. JONES, G. *et al.* Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology*, 267, 727-748 (1997).
 97. MORRIS, G. M. *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, 19, 1639-1662 (1998).
 98. RUIZ-CARMONA, S. *et al.* rDock: A Fast, Versatile and Open Source Program for Docking Ligands to Proteins and Nucleic Acids. *PLoS Computational Biology*, 10, e1003571 (2014).
 99. KORB, O. *et al.* Empirical scoring functions for advanced Protein-Ligand docking with PLANTS. *Journal of Chemical Information and Modeling*, 49, 84-96 (2009).
 100. CHEN, H. M. *et al.* SODOCK: Swarm optimization for highly flexible protein-ligand docking. *Journal of Computational Chemistry*, 28, 612-623 (2007).
 101. NG, M. C. K. *et al.* PSOVina: The hybrid particle swarm optimization algorithm for protein-ligand docking. *Journal of Bioinformatics and Computational Biology*, 13, 1541007 (2015).
 102. BROOIJMANS, N. & KUNTZ, I. D. Molecular Recognition and Docking Algorithms. *Annual Review of Biophysics and Biomolecular Structure*, 32, 335-373 (2003).
 103. TEAGUE, S. J. Implications of protein flexibility for drug discovery. *Nature Reviews Drug Discovery*, 2, 527-541 (2003).
 104. LIN, J.-H. Accommodating Protein Flexibility for Structure-Based Drug Design. *Current Topics in Medicinal Chemistry*, 11, 171-178 (2011).
 105. LIU, J. & WANG, R. Classification of current scoring functions. *Journal of Chemical Information and Modeling*, 55, 475-482 (2015).
 106. EWING, T. J. A. *et al.* DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases. *Journal of Computer-Aided Molecular Design*, 15, 411-428 (2001).
 107. HUANG, S.-Y & ZOU, X. An Iterative Knowledge-Based Scoring Function to Predict Protein-Ligand Interactions: I. Derivation of Interaction Potentials. *Journal of Computational Chemistry*, 27, 1866-1875 (2006).
 108. HUANG, S. Y. & ZOU, X. An iterative knowledge-based scoring function to predict protein-ligand interactions: II. Validation of the scoring function. *Journal of Computational Chemistry*, 27, 1875-1882 (2006).
 109. HUEY, R. *et al.* A Semiempirical Free Energy Force Field with Charge-Based Desolvation. *Journal of Computational Chemistry*, 28, 1145-1152 (2007).
 110. MUEGGE, I. *et al.* Evaluation of PMF scoring in docking weak ligands to the

- FK506 binding protein. *Journal of Medicinal Chemistry*, (1999).
111. BÖHM, H. J. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *Journal of Computer-Aided Molecular Design*, 8, 243-256 (1994).
 112. GOHLKE, H. *et al.* Knowledge-based scoring function to predict protein-ligand interactions. *Journal of Molecular Biology*, 295, 337-356 (2000).
 113. WANG, R. *et al.* Comparative evaluation of 11 scoring functions for molecular docking. *Journal of Medicinal Chemistry*, 46, 2287-2303 (2003).
 114. GEHLHAAR, D. K. *et al.* Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. *Chemistry and Biology*, 2, 317-324 (1995).
 115. ZHANG, C. *et al.* A knowledge-based energy function for protein-ligand, protein-protein, and protein-DNA complexes. *Journal of Medicinal Chemistry*, 48, 2325-2335 (2005).
 116. DEWITTE, R. S. & SHAKHNOVICH, E. I. SMOG: De novo design method based on simple, fast, and accurate free energy estimates. 1. Methodology and supporting evidence. *Journal of the American Chemical Society*, 118, 11733-11744 (1996).
 117. ELDRIDGE, M. D. *et al.* Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *Journal of Computer-Aided Molecular Design*, 11, 425-445 (1997).
 118. MITCHELL, J. B. O. *et al.* BLEEP - Potential of mean force describing protein-ligand interactions: I. Generating potential. *Journal of Computational Chemistry*, 20, 1165-1176 (1999).
 119. MITCHELL, J. B. O. *et al.* BLEEP - Potential of mean force describing protein-ligand interactions: II. Calculation of binding energies and comparison with experimental data. *Journal of Computational Chemistry*, 20, 1177-1185 (1999).
 120. VENKATACHALAM, C. M. *et al.* LigandFit: A novel method for the shape-directed rapid docking of ligands to protein active sites. *Journal of Molecular Graphics and Modelling*, 21, 289-307 (2003).
 121. WANG, R. *et al.* SCORE: A new empirical method for estimating the binding affinity of a protein-ligand complex. *Journal of Molecular Modeling*, 4, 379-394 (1998).
 122. YANG, C. Y. *et al.* M-score: A knowledge-based potential scoring function accounting for protein atom mobility. *Journal of Medicinal Chemistry*, 49, 5903-5911 (2006).
 123. YIN, S. *et al.* MedusaScore: An accurate force field-based scoring function for virtual drug screening. *Journal of Chemical Information and Modeling*, 48, 1656-1662(2008).
 124. WANG, R. *et al.* Further development and validation of empirical scoring functions for structure-based binding affinity prediction. *Journal of Computer-Aided Molecular Design*, 16, 11-26 (2002).
 125. MOOIJ, W. T. M. & VERDONK, M. L. General and targeted statistical potentials for protein-ligand interactions. *Proteins: Structure, Function and Genetics*, 61, 242-287 (2005).
 126. JAIN, A. N. Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. *Journal of Medicinal Chemistry*, 46, 499-511 (2003).

127. XIE, Z. R. & HWANG, M. J. An interaction-motif-based scoring function for protein-ligand docking. *BMC Bioinformatics*, 11:298 (2010).
128. BALLESTER, P. J. & MITCHELL, J. B. O. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. *Bioinformatics*, 26, 1169-1175 (2010).
129. KRAMMER, A. *et al.* LigScore: A novel scoring function for predicting binding affinities. *Journal of Molecular Graphics and Modelling*, 23, 395-407 (2005).
130. DAS, S. *et al.* Binding affinity prediction with property-encoded shape distribution signatures. *Journal of Chemical Information and Modeling*, 50, 298-308 (2010).
131. RAUB, S. *et al.* AIScore - Chemically diverse empirical scoring function employing quantum chemical binding energies of hydrogen-bonded complexes. *Journal of Chemical Information and Modeling*, 48, 1492-1510 (2008).
132. FAN, H. *et al.* Statistical potential for modeling and ranking of protein-ligand interactions. *Journal of Chemical Information and Modeling*, 51, 3078-3092 (2011).
133. SOTRIFTER, C. A. *et al.* SFCscore: Scoring functions for affinity prediction of protein-ligand complexes. *Proteins: Structure, Function and Genetics*, 73, 395-419 (2008).
134. ROGNAN, D. *et al.* Predicting binding affinities of protein ligands from three-dimensional models: Application to peptide binding to class I major histocompatibility proteins. *Journal of Medicinal Chemistry*, 42, 4650-4658 (1999).
135. REULECKE, I. *et al.* Towards an integrated description of hydrogen bonding and dehydration: Decreasing false positives in virtual screening with the HYDE scoring function. *ChemMedChem*, 3, 885-897 (2008).
136. HUANG, S. Y. *et al.* Scoring functions and their evaluation methods for protein-ligand docking: Recent advances and future directions. *Physical Chemistry Chemical Physics*, 12, 12899-12908 (2010).
137. WÓJCIKOWSKI, M. *et al.* Performance of machine-learning scoring functions in structure-based virtual screening. *Scientific Reports*, 7:46710 (2017).
138. AIN, Q. U. *et al.* Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 5, 405-424 (2015).
139. BATISTA, V. S. & NASCIMENTO-JÚNIOR, N. M. Molecular Docking : Considerations of a Low Cost and Suitable Methodology and Some Successful Applications. *Medicinal & Analytical Chemistry International Journal*, 2 (2018).
140. SUTHERLAND, J. J. *et al.* Lessons in molecular recognition. 2. Assessing and improving cross-docking accuracy. *Journal of Chemical Information and Modeling*, 47, 2293-302 (2007).
141. VERDONK, M. L. *et al.* Virtual screening using protein-ligand docking: Avoiding artificial enrichment. *Journal of Chemical Information and Computer Sciences*, 44, 793-806 (2004).
142. RÉAU, M. *et al.* Decoys selection in benchmarking datasets: Overview and perspectives. *Frontiers in Pharmacology*, 9:11 (2018).
143. ABU-HAMMAD, A. *et al.* Homology modeling of MCH1 receptor and validation by docking/scoring and protein-aligned CoMFA. *European Journal of Medicinal Chemistry*, 44, 2583–2596 (2009).
144. TAHA, M. O. *et al.* Docking-based comparative intermolecular contacts analysis as new 3-D QSAR concept for validating docking studies and in silico

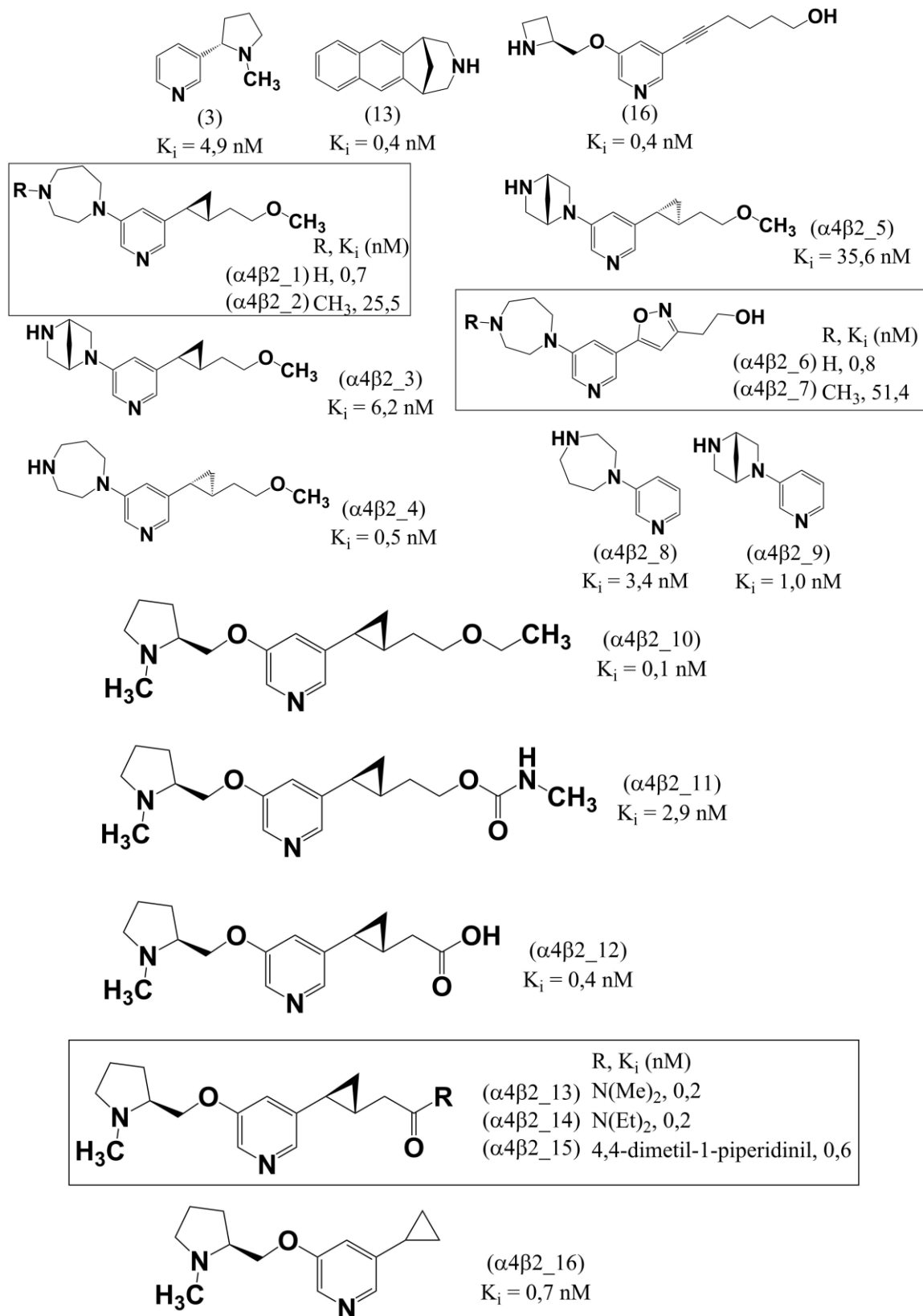
- screening: NMT and GP inhibitors as case studies. *Journal of Chemical Information and Modeling*, 51, 647-669 (2011).
145. AL-SHA'ER, M. A. & TAHA, M. O. Application of docking-based comparative intermolecular contacts analysis to validate Hsp90 α docking studies and subsequent *in silico* screening for inhibitors. *Journal of Molecular Modeling*, 18, 4843-4863 (2012).
146. TAHA, M. O. *et al.* The use of docking-based comparative intermolecular contacts analysis to identify optimal docking conditions within glucokinase and to discover of new GK activators. *Journal of Computer-Aided Molecular Design*, 28, 509-547 (2014).
147. JARADAT, N. J. *et al.* Combining docking-based comparative intermolecular contacts analysis and k-nearest neighbor correlation for the discovery of new check point kinase 1 inhibitors. *Journal of Computer-Aided Molecular Design*, 29, 561-581 (2015).
148. ALABED, S. J. *et al.* Computer-aided discovery of new FGFR-1 inhibitors followed by *in vitro* validation. *Future Medicinal Chemistry*, 8, 1841-1869 (2016).
149. HABASH, M. *et al.* Docking-based comparative intermolecular contacts analysis and *in silico* screening reveal new potent acetylcholinesterase inhibitors. *Medicinal Chemistry Research*, 26, 2768-2784 (2017).
150. ABUHAMMAD, A. *et al.* Computational modeling of the bat HKU4 coronavirus 3CLproinhibitors as a tool for the development of antivirals against the emerging Middle East respiratory syndrome (MERS) coronavirus. *Journal of Molecular Recognition*, 30, e2644 (2017).
151. AL-NADAF, A. H. *et al.* Discovery of new Gyrase beta inhibitors via structure based modeling. *Computational biology and chemistry*, 74, 263–272 (2018).
152. REAVILL, C. *et al.* Behavioural and pharmacokinetic studies on nicotine, cytisine and lobeline. *Neuropharmacology*, 29, 619-624 (1990).
153. <https://www.euroclinix.net/br/parar-de-fumar/champix> (Acesso em Fevereiro de 2019).
154. YU, L. F. *et al.* Recent developments in novel antidepressants targeting $\alpha 4\beta 2$ -nicotinic acetylcholine receptors. *Journal of Medicinal Chemistry*, 57, 8204-8223 (2014).
155. BRITTON, J. Death, disease, and tobacco. *The Lancet*, 389, 1861–1862 (2017).
156. <http://agenciabrasil.ebc.com.br/geral/noticia/2017-05/tabagismo-custa-r-569-bilhoes-por-ano-ao-brasil> (Acesso em Fevereiro de 2019).
157. HUANG, X. *et al.* Modeling subtype-selective agonists binding with $\alpha 4\beta 2$ and $\alpha 7$ nicotinic acetylcholine receptors: Effects of local binding and long-range electrostatic interactions. *Journal of Medicinal Chemistry*, 49, 7661-7674 (2006).
158. JI, J. *et al.* Synthesis and structure-activity relationship studies of 3,6-diazabicyclo[3.2.0]heptanes as novel $\alpha 4\beta 2$ nicotinic acetylcholine receptor selective agonists. *Journal of medicinal chemistry*, 50, 5493–5508 (2007).
159. CARROLL, F. I. *et al.* Synthesis and biological evaluation of bupropion analogues as potential pharmacotherapies for smoking cessation. *Journal of Medicinal Chemistry*, 53, 2204-2214 (2010).
160. LIMA, L. & BARREIRO, E. Bioisosterism: A Useful Strategy for Molecular Modification and Drug Design. *Current Medicinal Chemistry*, 12, 23-49 (2005).

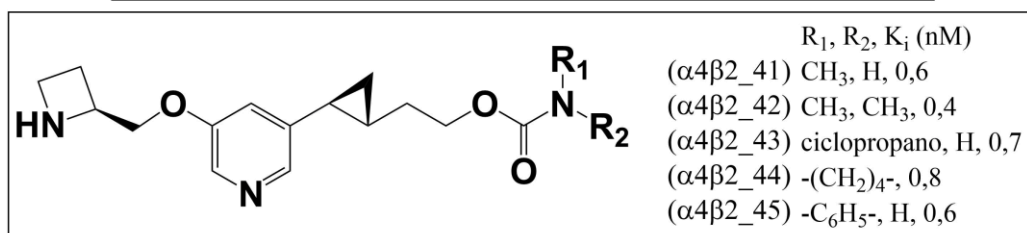
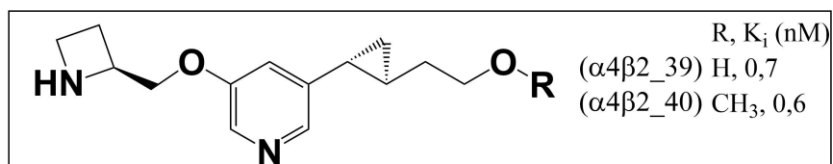
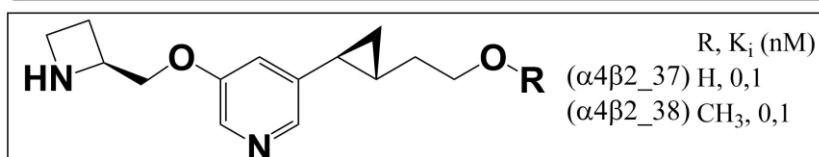
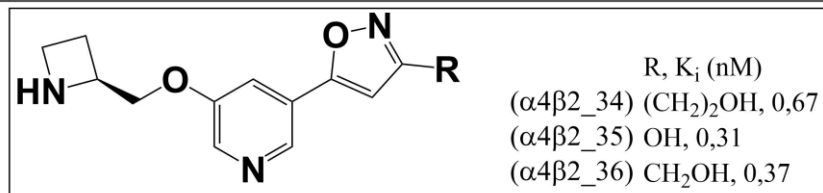
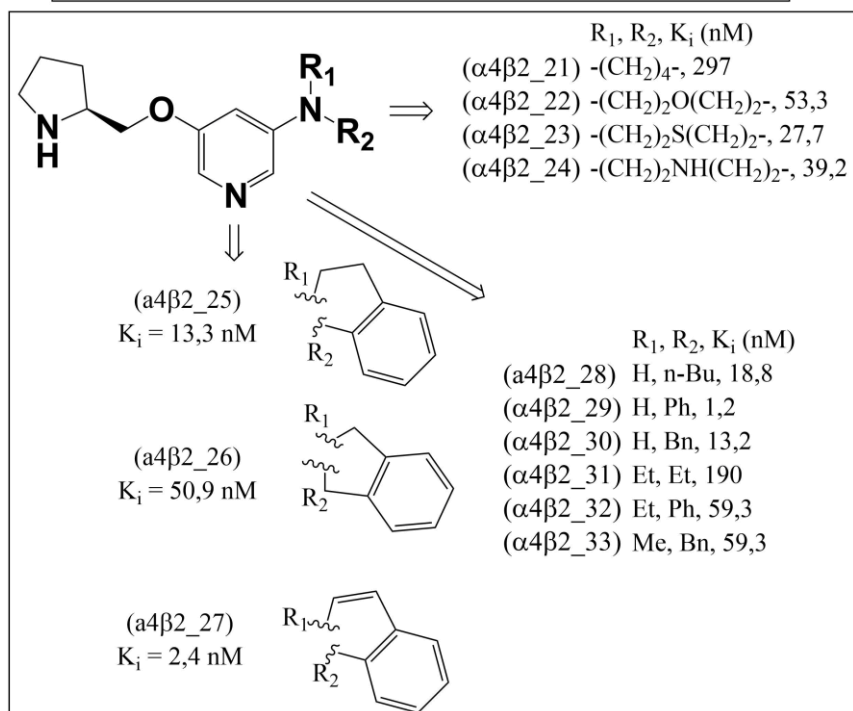
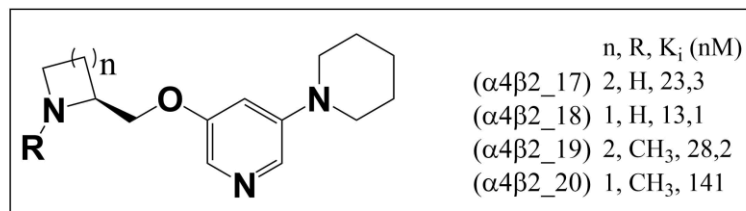
161. BARREIRO, E. J. L. & FRAGA, C. A. M. *Química Medicinal: As bases moleculares da ação dos fármacos*. (Artmed, **2015**).
162. SCOTT, D. E. *et al.* Fragment-based approaches in drug discovery and chemical biology. *Biochemistry*, 51, 4990-5003 (**2012**).
163. BERMAN, H. M. *et al.* The protein data bank. *Acta Crystallographica Section D: Biological Crystallography*, 28, 235-242 (**2002**).
164. ZHANG, H. K. *et al.* Synthesis and biological evaluation of novel hybrids of highly potent and selective $\alpha 4\beta 2$ -Nicotinic acetylcholine receptor (nAChR) partial agonists. *European Journal of Medicinal Chemistry*, 124, 689-697 (**2016**).
165. ONAJOLE, O. K. *et al.* Synthesis and Behavioral Studies of Chiral Cyclopropanes as Selective $\alpha 4\beta 2$ -Nicotinic Acetylcholine Receptor Partial Agonists Exhibiting an Antidepressant Profile. Part III. *ACS Chemical Neuroscience*, 7, 811-822 (**2016**).
166. LIU, J. *et al.* Chemistry and pharmacological characterization of novel nitrogen analogues of AMOP-H-OH (Sazetidine-A, 6-[5-(azetidin-2-ylmethoxy)pyridin-3-yl]hex-5-yn-1-ol) as $\alpha 4\beta 2$ -nicotinic acetylcholine receptor-selective partial agonists. *Journal of Medicinal Chemistry*, 53, 6973–6985 (**2010**).
167. LIU, J. *et al.* Discovery of isoxazole analogues of sazetidine-A as selective $\alpha 4\beta 2$ -nicotinic acetylcholine receptor partial agonists for the treatment of depression. *Journal of Medicinal Chemistry*, 54, 7280-7288 (**2011**).
168. ZHANG, H. *et al.* Chemistry and behavioral studies identify chiral cyclopropanes as selective $\alpha 4\beta 2$ -nicotinic acetylcholine receptor partial agonists exhibiting an antidepressant profile. *Journal of Medicinal Chemistry*, 55, 717-724 (**2012**).
169. YU, L. F. *et al.* Identification of novel $\alpha 4\beta 2$ -nicotinic acetylcholine receptor (nAChR) agonists based on an isoxazole ether scaffold that demonstrate antidepressant-like activity. *Journal of Medicinal Chemistry*, 55, 812-823 (**2012**).
170. YU, L. F. *et al.* Discovery of highly potent and selective $\alpha 4\beta 2$ -nicotinic acetylcholine receptor (nAChR) partial agonists containing an isoxazolylpyridine ether scaffold that demonstrate antidepressant-like activity. Part II. *Journal of Medicinal Chemistry*, 55, 9998-10009 (**2012**).
171. ZHANG, H. K. *et al.* Chemistry, pharmacology, and behavioral studies identify chiral cyclopropanes as selective $\alpha 4\beta 2$ -nicotinic acetylcholine receptor partial agonists exhibiting an antidepressant profile. part II. *Journal of Medicinal Chemistry*, 56, 5495-5504 (**2013**).
172. PALLAVICINI, M. *et al.* 5-(2-Pyrrolidinyl)oxazolidinones and 2-(2-pyrrolidinyl)benzodioxanes: Synthesis of all the stereoisomers and $\alpha 4\beta 2$ nicotinic affinity. *Bioorganic and Medicinal Chemistry Letters*, 19, 854–859 (**2009**).
173. CARBONNELLE, E. *et al.* Nitrogen substitution modifies the activity of cytosine on neuronal nicotinic receptor subtypes. *European Journal of Pharmacology*, 471, 85–96 (**2003**).
174. DALLANOCE, C. *et al.* Synthesis of epibatidine-related $\Delta 2$ -isoxazoline derivatives and evaluation of their binding affinity at neuronal nicotinic acetylcholine receptors. *European Journal of Organic Chemistry*, 6, 3746–3754 (**2006**).
175. DALLANOCE, C. *et al.* New analogues of epiboxidine incorporating the 4,5-dihydroisoxazole nucleus: Synthesis, binding affinity at neuronal nicotinic

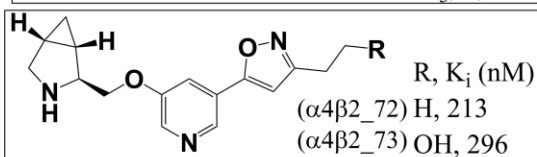
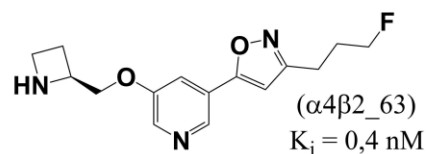
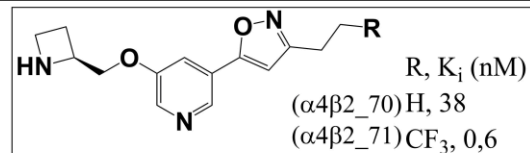
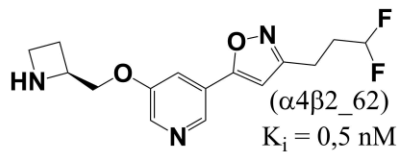
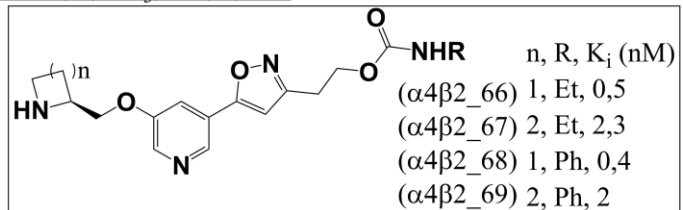
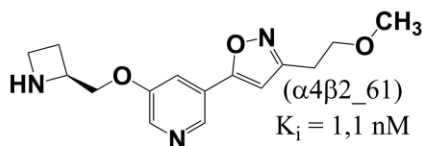
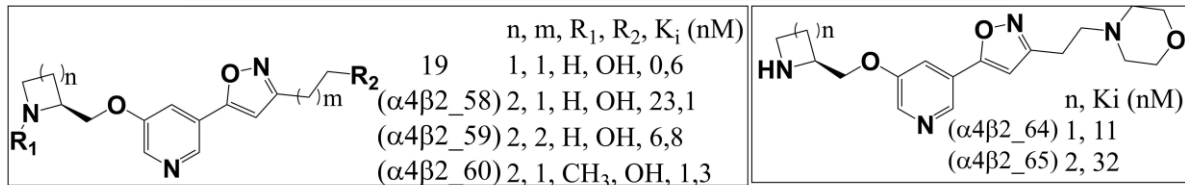
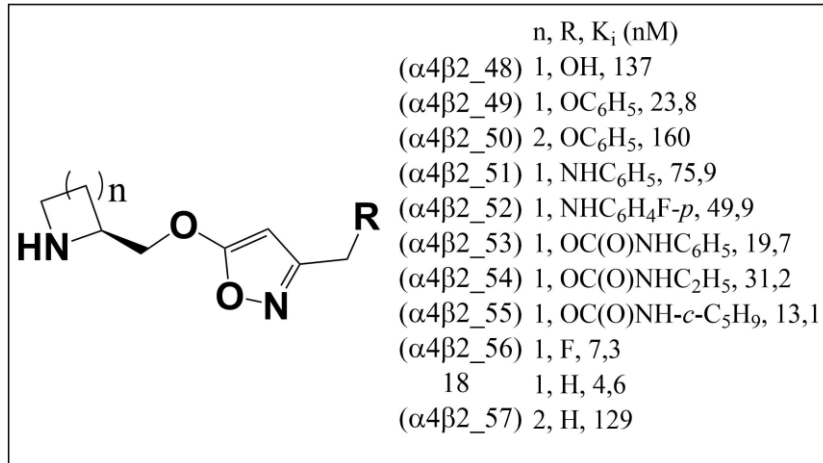
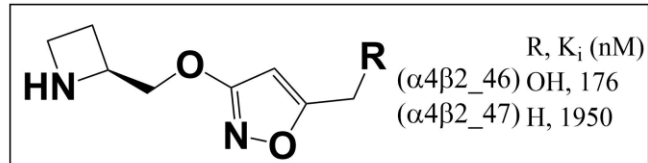
- acetylcholine receptors, and molecular modeling investigations. *Chemistry and Biodiversity*, 6, 244–259 (2009).
176. DALLANOCE, C. *et al.* Synthesis of novel chiral $\Delta 2$ -isoxazoline derivatives related to ABT-418 and estimation of their affinity at neuronal nicotinic acetylcholine receptor subtypes. *European Journal of Medicinal Chemistry*, 45, 5594–5601 (2010).
177. DALLANOCE, C. *et al.* Novel tricyclic $\Delta 2$ -isoxazoline and 3-oxo-2-methylisoxazolidine derivatives: Synthesis and binding affinity at neuronal nicotinic acetylcholine receptor subtypes. *Bioorganic and Medicinal Chemistry*, 18, 4498–4508 (2010).
178. DALLANOCE, C. *et al.* New spirocyclic $\Delta 2$ -isoxazoline derivatives related to selective agonists of $\alpha 7$ neuronal nicotinic acetylcholine receptors. *European Journal of Medicinal Chemistry*, 46, 5790–5799 (2011).
179. DALLANOCE, C. *et al.* Design, synthesis, and pharmacological characterization of novel spirocyclic quinuclidinyl- $\Delta 2$ -Isoxazoline derivatives as potent and selective agonists of $\alpha 7$ nicotinic acetylcholine receptors. *ChemMedChem*, 6, 889–903 (2011).
180. RIZZI, L. *et al.* A Small Library of 1,2,3-Triazole Analogs of CAP-55: Synthesis and Binding Affinity at Nicotinic Acetylcholine Receptors. *Chemistry and Biodiversity*, 15, e1800210 (2018).
181. Dassault Systèmes BIOVIA, Discovery Studio Visualizer, V.17.0.2.1076, San Diego: Dassault Systèmes, 2017.
182. <https://chemaxon.com/products/marvin> (Acesso em julho de 2018)..
183. MACRAE, C. F. *et al.* Mercury CSD 2.0 - New features for the visualization and investigation of crystal structures. *Journal of Applied Crystallography*, 41, 466-470 (2008).
184. STEWART, J. J. P. MOPAC2016. Stewart Computational Chemistry (2016).
185. CELIE, P. H. N. *et al.* Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. *Neuron*, 41, 90-914 (2004).
186. CRESTEY, F. *et al.* Dual Nicotinic Acetylcholine Receptor $\alpha 4\beta 2$ Antagonists/ $\alpha 7$ Agonists: Synthesis, Docking Studies, and Pharmacological Evaluation of Tetrahydroisoquinolines and Tetrahydroisoquinolinium Salts. *Journal of Medicinal Chemistry*, 61, 1719–1729 (2018).
187. KORB, O. *et al.* Interactive and Versatile Navigation of Structural Databases. *Journal of Medicinal Chemistry*, 59, 4257-4266 (2016).
188. CERETO-MASSAGUÉ, A. *et al.* DecoyFinder: An easy-to-use python GUI application for building target-specific decoy sets. *Bioinformatics*, 28, 1661-1662 (2012).
189. IRWIN, J. J. & SHOICHET, B. K. ZINC - A free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling*, 45, 177-182 (2005).
190. Origin (OriginLab, Northampton, MA)
191. <http://mordred.bioc.cam.ac.uk/~rapper/rampage.php> (Acesso em Julho de 2018).
192. SANDER, T. *et al.* Docking to flexible nicotinic acetylcholine receptors: A validation study using the acetylcholine binding protein. *Journal of Molecular Graphics and Modelling*, 29, 415-424 (2010).
193. AMIRI, S. *et al.* Molecular dynamics studies of AChBP with nicotine and

- carbamylcholine: The role of water in the binding pocket. *Protein Engineering, Design and Selection*, 20, 353-359 (2007).
194. RUCKTOOA, P. *et al.* Structural characterization of binding mode of smoking cessation drugs to nicotinic acetylcholine receptors through study of ligand complexes with acetylcholine-binding protein. *Journal of Biological Chemistry*, 287, 23283-23293 (2012).
 195. DELANO, W. L. The PyMOL Molecular Graphics System, Version 0.99 Schrödinger, LLC. Schrödinger LLC, (2002).
 196. HANSEN, S. B. *et al.* Structures of Aplysia AChBP complexes with nicotinic agonists and antagonists reveal distinctive binding interfaces and conformations. *EMBO Journal*, 24, 3635-3646 (2005).
 197. XIU, X. *et al.* Nicotine binding to brain receptors requires a strong cation-interaction. *Nature*, 458, 534-537 (2009).
 198. ZHONG, W. *et al.* From ab initio quantum mechanics to molecular neurobiology: A cation- binding site in the nicotinic receptor. *Proceedings of the National Academy of Sciences*, 95, 12088-12093 (1998).
 199. GALZI, J. L. *et al.* Functional significance of aromatic amino acids from three peptide loops of the $\alpha 7$ neuronal nicotinic receptor site investigated by site-directed mutagenesis. *FEBS Letters*, 294, 198-202 (1991).
 200. SINE, S. M. *et al.* Conserved tyrosines in the alpha subunit of the nicotinic acetylcholine receptor stabilize quaternary ammonium groups of agonists and curariform antagonists. *Journal of Biological Chemistry*, 269, 8808-8816 (1994).
 201. MIDDLETON, R. E. & COHEN, J. B. Mapping of the acetylcholine binding site of the nicotinic acetylcholine receptor: [3H]nicotine as an agonist photoaffinity label. *Biochemistry*, 30, 6987-6997 (1991).
 202. O'LEARY, M. E. & WHITE, M. M. Mutational analysis of ligand-induced activation of the Torpedo acetylcholine receptor. *Journal of Biological Chemistry*, 267, 8360-8365 (1992).
 203. SANDER, T. *et al.* DataWarrior: An open-source program for chemistry aware data visualization and analysis. *Journal of Chemical Information and Modeling*, 55, 460-473 (2015).
 204. CONGREVE, M. *et al.* A 'Rule of Three' for fragment-based lead discovery? *Drug Discovery Today*, 8, 876-877 (2003).
 205. WILLIAMS, D. K. *et al.* Differential Regulation of Receptor Activation and Agonist Selectivity by Highly Conserved Tryptophans in the Nicotinic Acetylcholine Receptor Binding Site. *Journal of Pharmacology and Experimental Therapeutics*, 330, 40-53 (2009).
 206. RUCKTOOA, P. *et al.* Insight in nAChR subtype selectivity from AChBP crystal structures. *Biochemical Pharmacology*, 78, 777-787 (2009).
 207. GRUTTER, T. *et al.* An H-bond between two residues from different loops of the acetylcholine binding site contributes to the activation mechanism of nicotinic receptors. *EMBO Journal*, 22, 1990-2003 (2003).
 208. DELBART, F. *et al.* An allosteric binding site of the $\alpha 7$ nicotinic acetylcholine receptor revealed in a humanized acetylcholine-binding protein. *Journal of Biological Chemistry*, 293, 2534-2545 (2018).

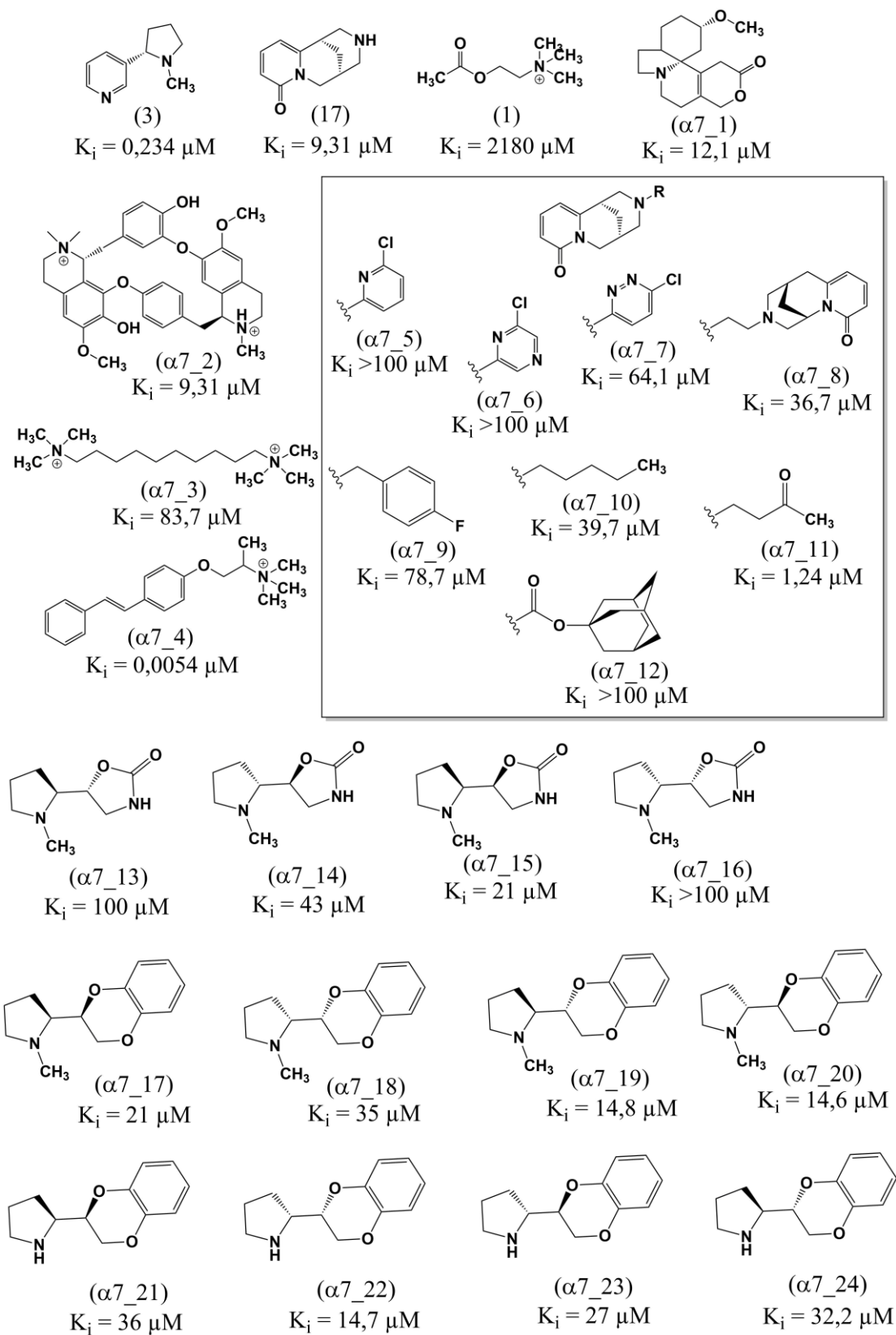
7. Anexos

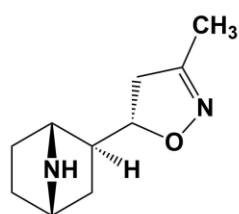
ANEXO IA – Estruturas químicas dos 98 ligantes utilizados para a construção dos modelos dbCICA para o nAChRs do subtipo $(\alpha 4)_2(\beta 2)_3$.



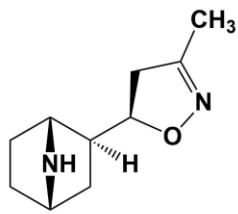


ANEXO IB – Estruturas químicas dos 93 ligantes utilizados para a construção dos modelos dbCICA para o nAChRs do subtipo $\alpha 7$.

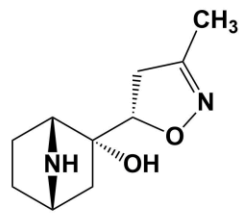




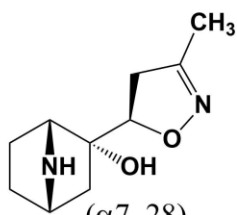
($\alpha 7_{25}$)
 $K_i = 32 \mu\text{M}$



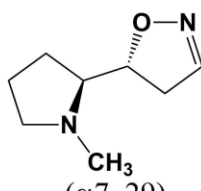
($\alpha 7_{26}$)
 $K_i = 40 \mu\text{M}$



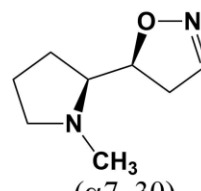
($\alpha 7_{27}$)
 $K_i = 77 \mu\text{M}$



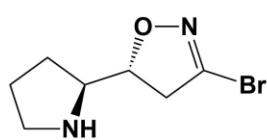
($\alpha 7_{28}$)
 $K_i = 32 \mu\text{M}$



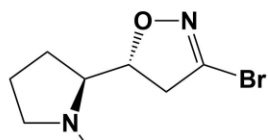
($\alpha 7_{29}$)
 $K_i = 40 \mu\text{M}$



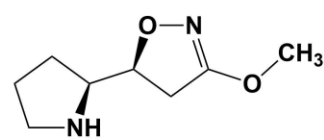
($\alpha 7_{30}$)
 $K_i = 0,44 \mu\text{M}$



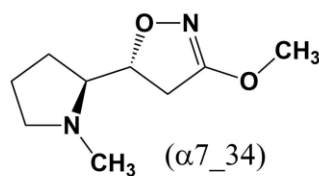
($\alpha 7_{31}$)
 $K_i = 35 \mu\text{M}$



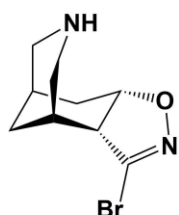
($\alpha 7_{32}$)
 $K_i = 43 \mu\text{M}$



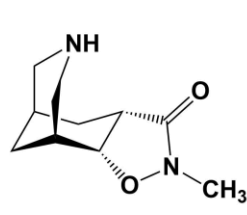
($\alpha 7_{33}$)
 $K_i = 11,3 \mu\text{M}$



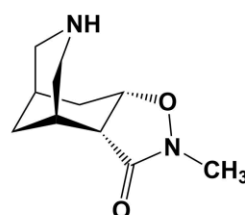
($\alpha 7_{34}$)
 $K_i = 2,9 \mu\text{M}$



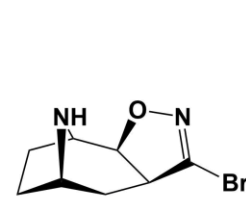
($\alpha 7_{35}$)
 $K_i = 28,5 \mu\text{M}$



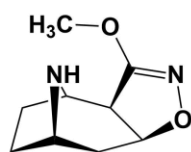
($\alpha 7_{36}$)
 $K_i = 9,5 \mu\text{M}$



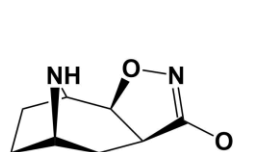
($\alpha 7_{37}$)
 $K_i > 100 \mu\text{M}$



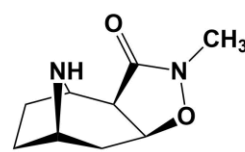
($\alpha 7_{38}$)
 $K_i = 20 \mu\text{M}$



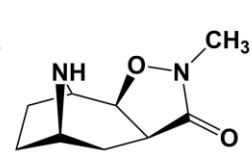
($\alpha 7_{39}$)
 $K_i = 68 \mu\text{M}$



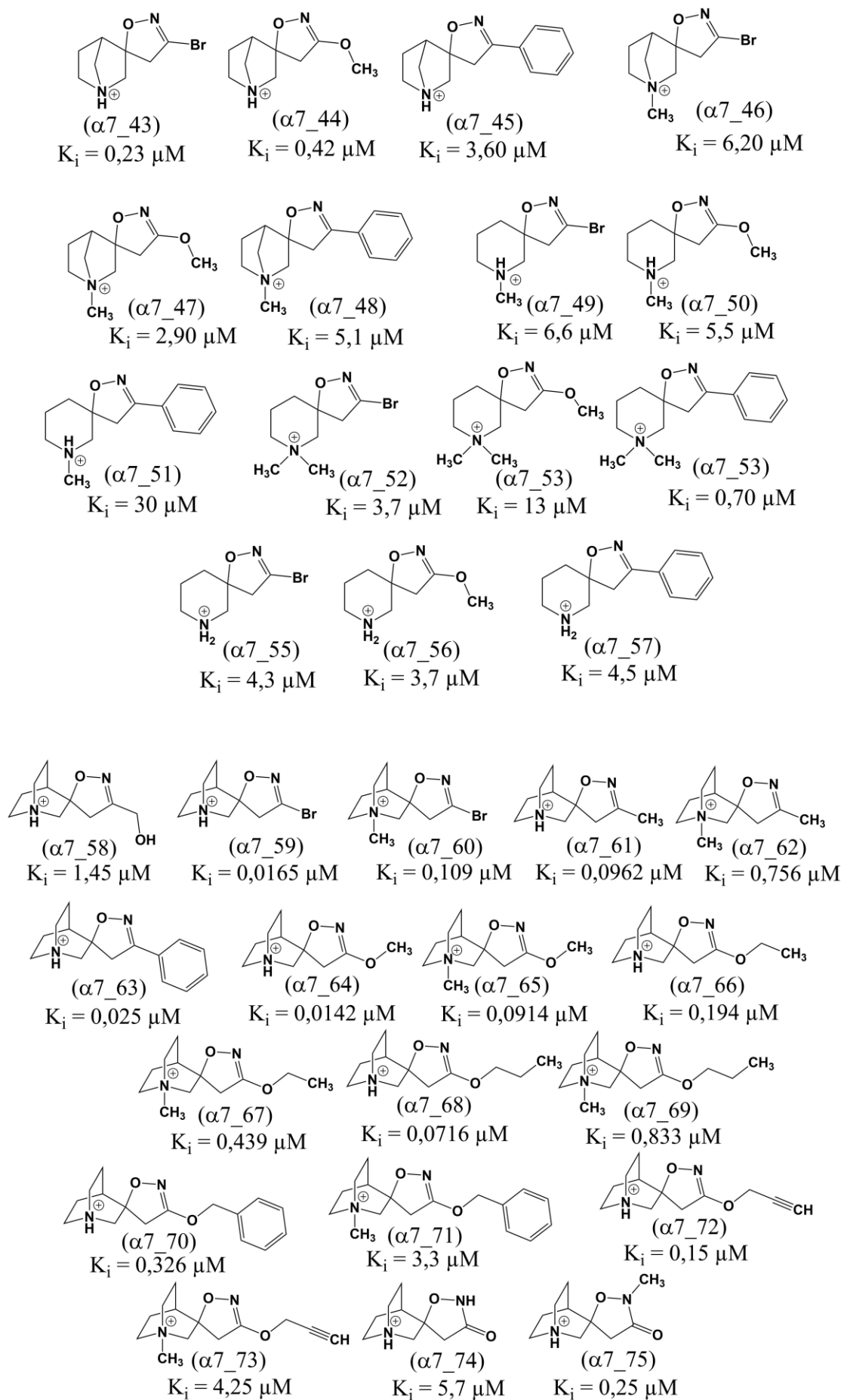
($\alpha 7_{40}$)
 $K_i = 28,5 \mu\text{M}$

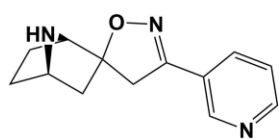


($\alpha 7_{41}$)
 $K_i = 28,2 \mu\text{M}$

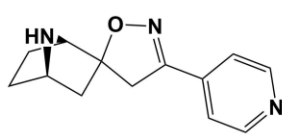


($\alpha 7_{42}$)
 $K_i > 100 \mu\text{M}$

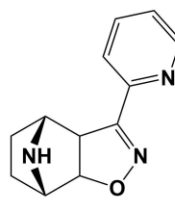




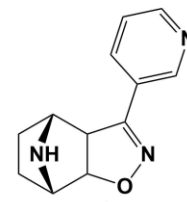
(α7_76)
K_i = 41 μM



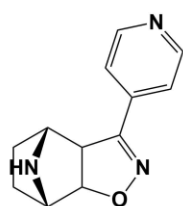
(α7_77)
K_i = 22 μM



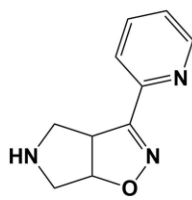
(α7_78)
K_i > 100 μM



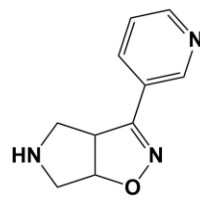
(α7_79)
K_i = 32 μM



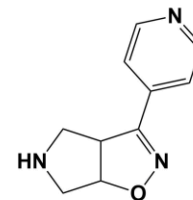
(α7_80)
K_i > 100 μM



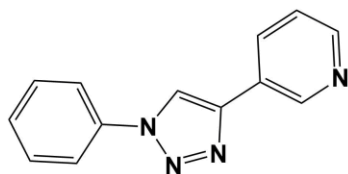
(α7_81)
K_i > 100 μM



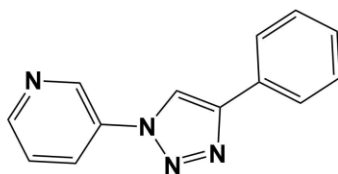
(α7_82)
K_i > 100 μM



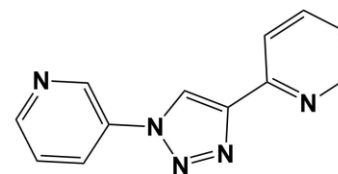
(α7_83)
K_i > 100 μM



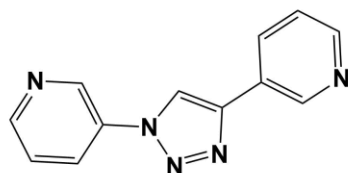
(α7_84)
K_i = 100 μM



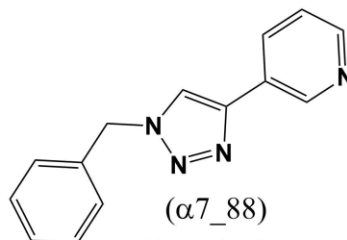
(α7_85)
K_i > 100 μM



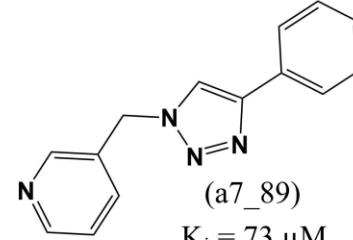
(α7_86)
K_i = 6,5 μM



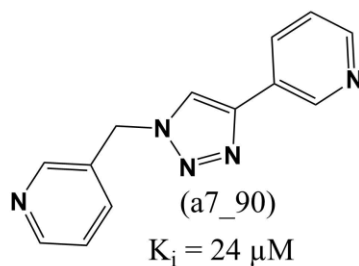
(α7_87)
K_i > 100 μM



(α7_88)
K_i = 91 μM



(α7_89)
K_i = 73 μM



(α7_90)
K_i = 24 μM

ANEXO II – Equações utilizadas no trabalho.

- Distância Euclidiana entre duas moléculas

$$D(i, j) = \sqrt{(HBD_i - HBD_j)^2 + (HBA_i - HBA_j)^2 + (NP_i + NP_j)^2}$$

HBD = Átomos doadores de ligação de hidrogênio

HBA = Átomos aceptores de ligação de hidrogênio

NP = Átomos pesados não polares

- Parâmetros calculados para a construção da curva ROC
 - Sensitividade

$$Se = \frac{TP}{TP + FN}$$

- Especificidade

$$Sp = \frac{TN}{TN + FP}$$

- Percentual de ativos

$$Ya = \frac{TP}{TO}$$

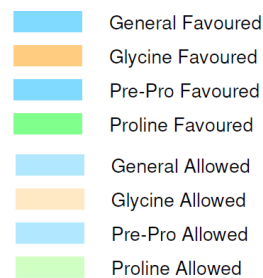
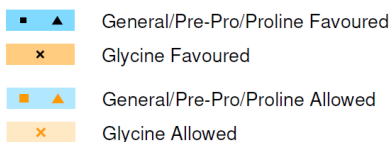
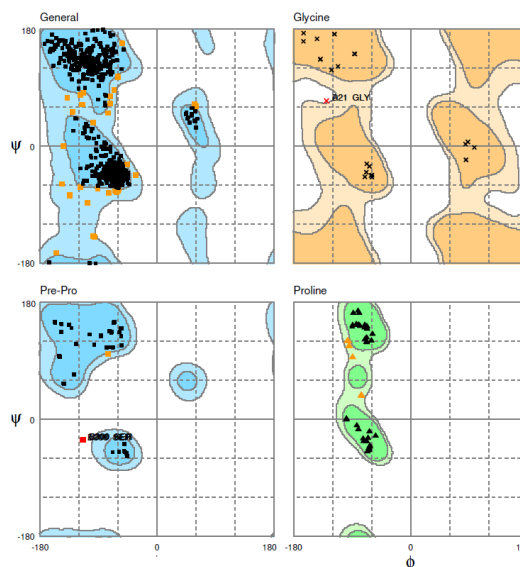
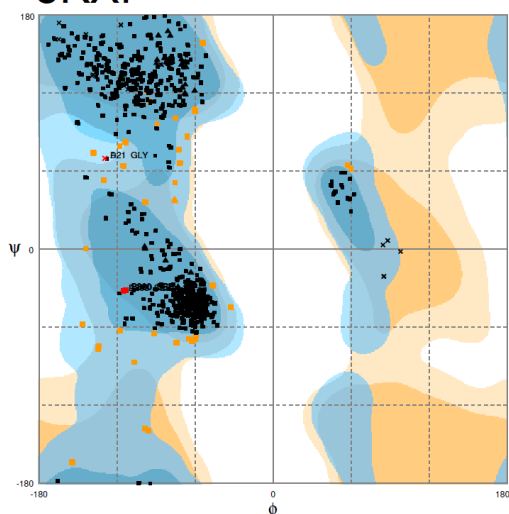
- Percentual de precisão

$$ACC = \frac{TP + TN}{TO}$$

Sendo TP o número de moléculas ativas capturadas pelo mapa farmacofórico, FN os ativos descartados, TN os inativos descartados, FP os falsos positivos, TC o total de moléculas ativas e inativas capturadas e TO o total de moléculas contidas no conjunto de dados.

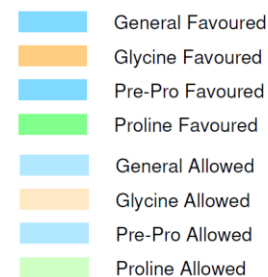
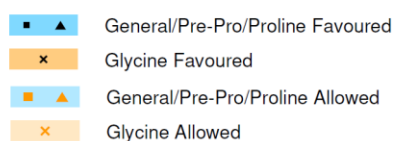
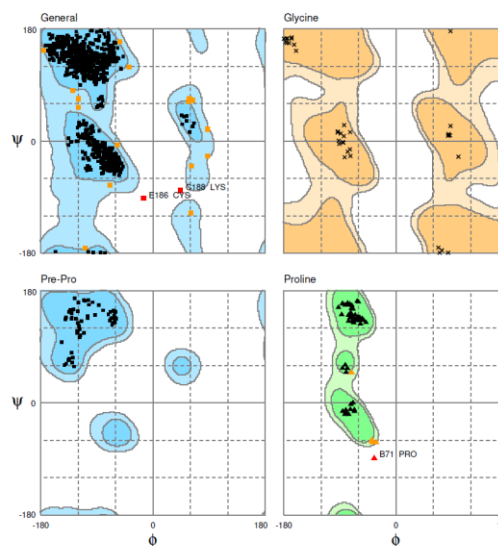
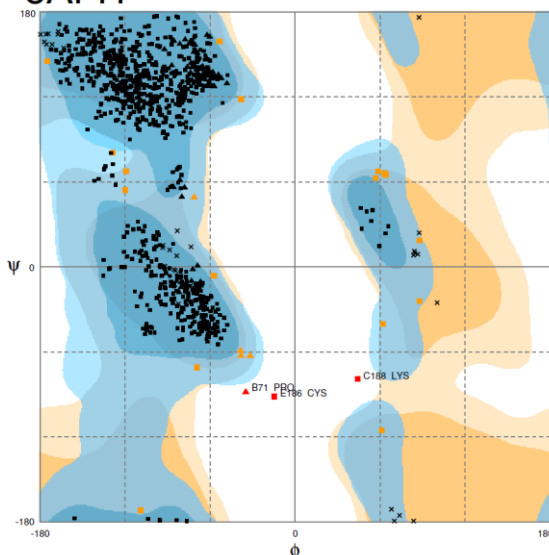
ANEXO III – Gráficos de Ramachandran para as estruturas cristalográficas 5KXI (($\alpha 4$)₂($\beta 2$)₃) e 5AFH ($\alpha 7$).

5KXI



Number of residues in favoured region (~98.0% expected) : 1685 (94.7%)
 Number of residues in allowed region (~2.0% expected) : 88 (4.9%)
 Number of residues in outlier region : 7 (0.4%)

5AFH



Number of residues in favoured region (~98.0% expected) : 987 (97.5%)
 Number of residues in allowed region (~2.0% expected) : 22 (2.2%)
 Number of residues in outlier region : 3 (0.3%)

ANEXO IV – Alinhamento das estruturas primárias das cadeias α (A) e β (B) da estrutura cristalográfica 5KXI ((α)₂(β)₃) e estrutura homóloga 4AFT (AChBP).

5KXI_A: NEURONAL ACETYLCHOLINE RECEPTOR SUBUNIT ALPHA-4

4AFT_A: SOLUBLE ACETYLCHOLINE RECEPTOR

```

      10      20      30      40      50      60
      .....|.....|.....|.....|.....|.....|
5KXI_A 11 EERLLKKL fsgynkwsrpvanisdvVLVRFglsiaqlIdvdeknqmmTTnvwvkqewhdY 70
4AFT_A  4 LMRLKSDLfnrpsmyp--gptkddpLVTlGftlqdiVkvdsstnevDLvyyeqqrwkln 61

      70      80      90      100     110     120
      .....|.....|.....|.....|.....|.....|
5KXI_A 71 KLRWDPaDYENVTSIRIPseliwRPDIVLYNNAgdg favthltKAHLFHDgRVQWTPPAI 130
4AFT_A 62 SLMWDPnEYGNITDFRTSAadiwTPDITAYSSTrpvqv-lspqIAVVTHDgSVMFIPAQR 120

      130     140     150     160     170     180
      .....|.....|.....|.....|.....|.....|
5KXI_A 131 YksscsidvtffpfdqqnctmKFGswydkakIDLVMhsvrdqldfwesgewvivDavg 190
4AFT_A 121 Lsfmcdptgvd-seegvtcavKFGswysgfEIDLKTdtdqvdLssyyasskyeilsatq 179

      190     200
      .....|.....|
5KXI_A 191 tyntnlyccbaeyPDITYAFVI 213
4AFT_A 180 trqvqlyccbaeyPDVNLVWKF 202

```

5KXI_B: NEURONAL ACETYLCHOLINE RECEPTOR SUBUNIT BETA-2

4AFT_B: SOLUBLE ACETYLCHOLINE RECEPTOR

```

      10      20      30      40      50      60
      .....|.....|.....|.....|.....|.....|
5KXI_B  4 EERLVEHLLdpsrynklirpatngselvTVQLMVslaqlisvhereqimttntWLTQEW 63
4AFT_B  4 LMRLKSDLfnrpsmypg----ptkddpLVTlGfTlqdivkvsstnevdlvYBQQRWK 59

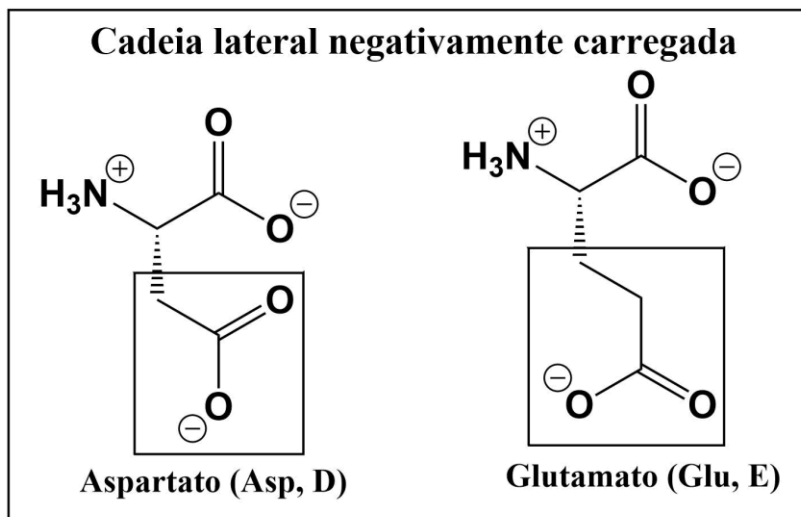
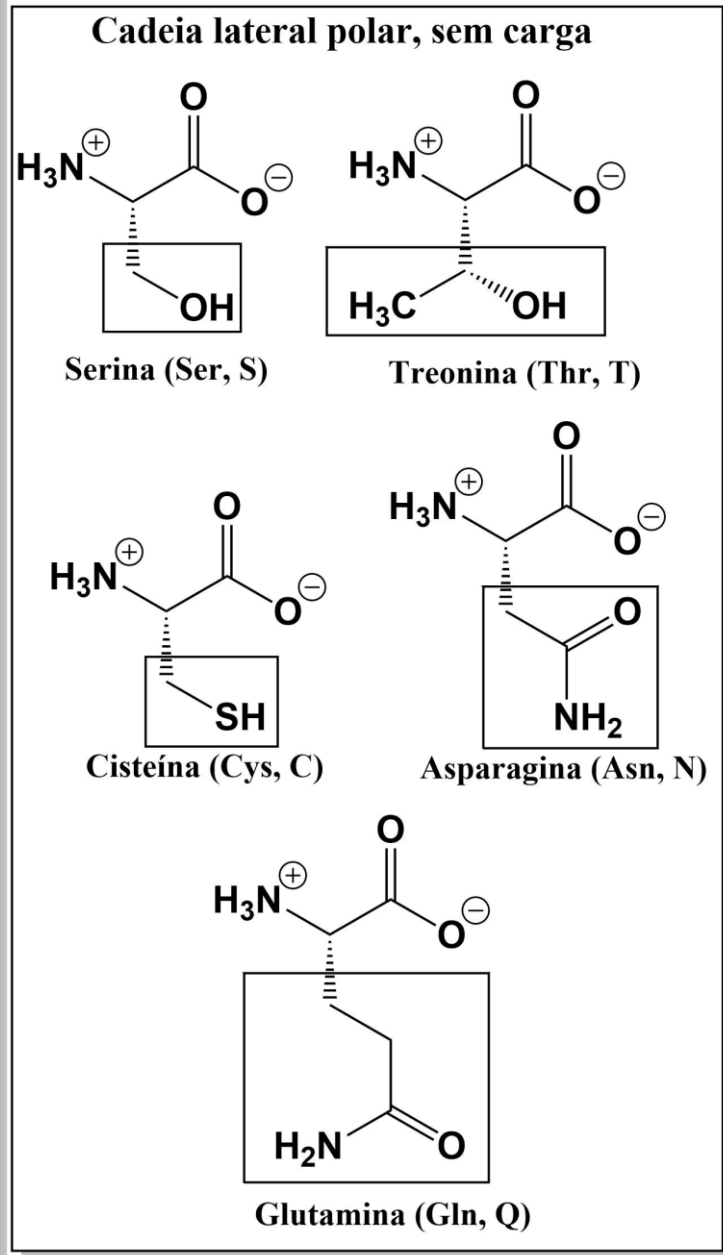
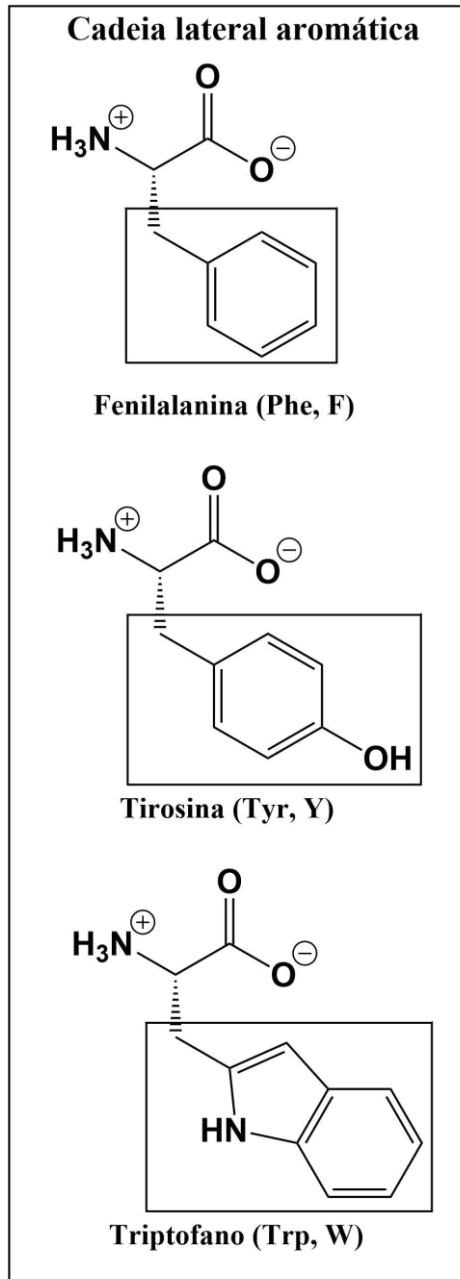
      70      80      90      100     110     120
      .....|.....|.....|.....|.....|.....|
5KXI_B 64 DYRLTWKpeeFDNMKKVRlpskhiwlpDVVLynnadgmyevsfysnAVVSYdgsiFLLP 123
4AFT_B 60 LNSLMNDpneyGNITDFRtsaadiwtpDITAYSstrpv-qvlsppiAVVTHdgsvMFI 118

      130     140     150     160     170     180
      .....|.....|.....|.....|.....|.....|
5KXI_B 124 AIyksackievkhfpfdqqnctmKFRswtydrteidlvkSEVAslddftpsgewdival 183
4AFT_B 119 QRlsfmcDptgv-dseegvtcavKFGswysgfEIDLKtdTDQVdlssyyasskyeilsa 177

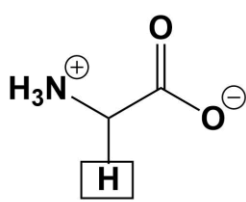
      190     200
      .....|.....|
5KXI_B 184 ---pgrnnpddstyvdiT 200
4AFT_B 178 tqtrqvqhysccpepyidvN 197

```

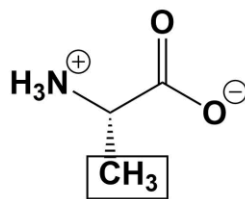
Resíduos em vermelho estão conservados e na mesma posição espacial, resíduos em preto estão conservados mas em posição espacial diferente e resíduos em azul indicam resíduos substituídos por resíduos homólogos.

ANEXO V – Aminoácidos: estruturas químicas e abreviações


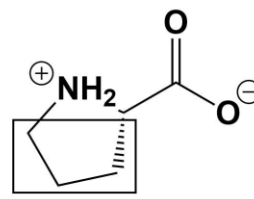
Cadeia lateral alifática e apolar



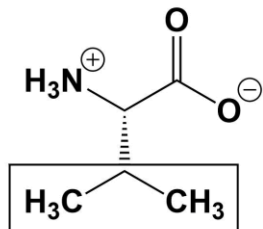
Glicina (Gly, G)



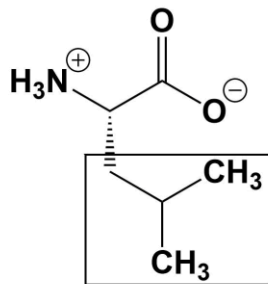
Alanina (Ala, A)



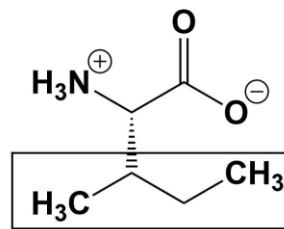
Prolina (Pro, P)



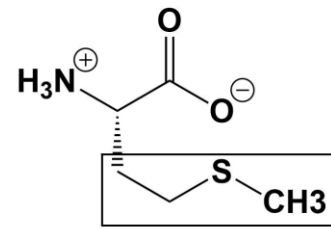
Valina (Val, V)



Leucina (Leu, L)

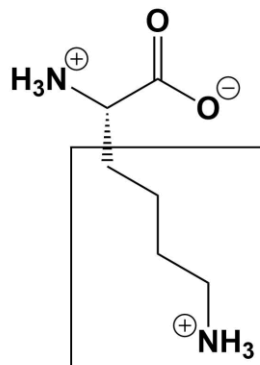


Isoleucina (Ile, I)

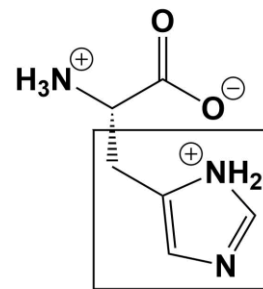


Metionina (Met, M)

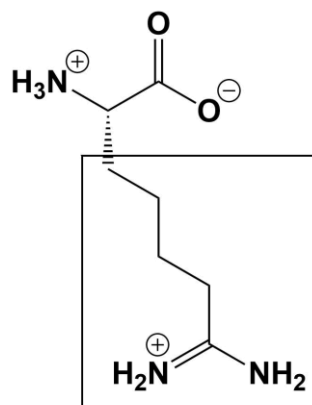
Cadeia lateral positivamente carregada



Lisina (Lys, K)



Histidina (His, H)



Arginina (Arg, R)

Artigo

Receptores Nicotínicos de Acetilcolina dos Subtipos $\alpha 7$, $\alpha 4\beta 2$ e $\alpha 3\beta 4$: Características, Participação em Processos Patogênicos e Compostos Bioativos

Batista, V. S.; Nascimento-Júnior, N. M.*

Rev. Virtual Quim., 2016, 8 (5), 1721-1739. Data de publicação na Web: 11 de julho de 2016<http://rvq.sbq.org.br>**The $\alpha 7$, $\alpha 4\beta 2$ and $\alpha 3\beta 4$ Nicotinic Acetylcholine Receptors: Characteristics, Role in Pathogenic Processes and Bioactive Compounds**

Abstract: Neuronal nicotinic acetylcholine receptors (nAChRs) are potential targets for the treatment of several central nervous system disorders, including Alzheimer's disease, Parkinson's disease and schizophrenia. On this review will be discussed the main characteristics of the nAChRs, more specifically of the $\alpha 7$, $\alpha 4\beta 2$ and $\alpha 3\beta 4$ subtypes, examples of bioactive compounds that act on these receptors and structure-activity relationships of some compounds on their respective target receptors.

Keywords: Medicinal chemistry; nicotinic acetylcholine receptors; bioactive compounds; structure-activity relationship.

Resumo

Os receptores nicotínicos de acetilcolina neuronais (nAChRs) são alvos em potencial para o tratamento de diversas doenças do sistema nervoso central (SNC), incluindo doença de Alzheimer, doença de Parkinson e esquizofrenia. Nesta revisão serão abordadas as características principais dos nAChRs, mais especificamente dos subtipos $\alpha 7$, $\alpha 4\beta 2$ e $\alpha 3\beta 4$, exemplos de compostos bioativos que atuam nesses receptores e a relação estrutura-atividade de alguns compostos em seus respectivos receptores alvo.

Palavras-chave: Química medicinal; receptores nicotínicos de acetilcolina; compostos bioativos; relação estrutura-atividade.

* Universidade Estadual Paulista "Júlio de Mesquita Filho", Laboratório de Química Medicinal, Síntese Orgânica e Modelagem Molecular (LaQMedSOMM), Instituto de Química, Departamento de Química Orgânica, Campus Araraquara, CEP 14800-060, Araraquara-SP, Brasil.

✉ nailtonjr@gmail.com

DOI: [10.21577/1984-6835.20160114](https://doi.org/10.21577/1984-6835.20160114)

Receptores Nicotínicos de Acetilcolina dos Subtipos $\alpha 7$, $\alpha 4\beta 2$ e $\alpha 3\beta 4$: Características, Participação em Processos Patogênicos e Compostos Bioativos

Victor de Sousa Batista, Nailton M. Nascimento-Júnior*

Universidade Estadual Paulista "Júlio de Mesquita Filho", Laboratório de Química Medicinal, Síntese Orgânica e Modelagem Molecular (LaQMedSOMM), Instituto de Química, Departamento de Química Orgânica, Campus Araraquara, CEP 14800-060, Araraquara-SP, Brasil.

* nailtonjr@gmail.com

Recebido em 23 de março de 2016. Aceito para publicação em 6 de julho de 2016

1. Introdução
2. A relação estrutura-atividade de ligantes de receptores nicotínicos de acetilcolina (nAChRs) do subtipo $\alpha 7$
3. A relação estrutura-atividade de ligantes de receptores nicotínicos de acetilcolina (nAChRs) do subtipo $\alpha 4\beta 2$
4. A relação estrutura-atividade de ligantes de receptores nicotínicos de acetilcolina (nAChRs) do subtipo $\alpha 3\beta 4$
5. Conclusão

1. Introdução

Os receptores nicotínicos de acetilcolina (nAChRs) são parte do grupo de receptores do neurotransmissor endógeno acetilcolina (ACh) **(1)** (**Figura 1**) que abrange, além dos já mencionados nAChRs, os receptores muscarínicos de acetilcolina (mAChRs), sendo os mAChRs receptores metabotrópicos que possuem alta afinidade pela muscarina **(2)** e baixa afinidade pela nicotina **(3)**, sendo o comportamento inverso observado nos nAChRs. Quanto a fisiologia da neurotransmissão colinérgica, a ativação dos mAChRs pela toxina do cogumelo *Amanita*

muscaria, a muscarina **(2)**, ocorre em um período de tempo relativamente longo, da ordem de milissegundos a segundos, enquanto os nAChRs são receptores ionotrópicos que são ativados em intervalos de tempo da ordem de sub-microsegundos a microssegundos pela molécula de nicotina **(3)**,¹⁻³ comumente proveniente da planta *Nicotiana tabacum*. Os nAChRs podem ser divididos em dois grupos, os receptores nicotínicos musculares e os neuronais,^{4,5} com os nAChRs neuronais majoritariamente expressos pelos subtipos $\alpha 7$, $\alpha 4\beta 2$ e $\alpha 3\beta 4$, podendo exercer funções variadas como regulação de excitabilidade neuronal e liberação de neurotransmissores.⁶⁻⁸

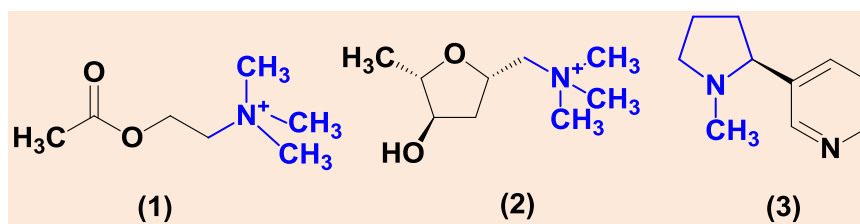


Figura 1. Estruturas das moléculas de acetilcolina (1), muscarina (2) e nicotina (3)

Estruturalmente, os nAChRs neuronais são glicoproteínas grandes (aproximadamente 290 kDa) que apresentam cinco subunidades transmembranares arranjadas em torno de um poro central permeável a cátions^{3,7,9} e fazem parte da família dos canais iônicos cys-

loop, que inclui também os nAChRs musculares, receptores de ácido gama aminobutírico (4) (GABA_A e GABA_C), receptores de glicina (5) e receptores de serotonina (5-HT₃) (6) (Figura 2).^{7,10}

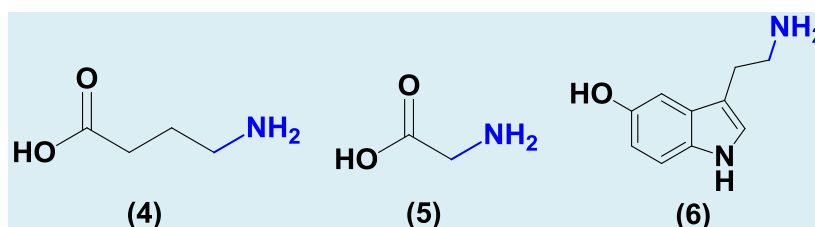


Figura 2. Estruturas das moléculas de ácido gama amino-butírico (4), glicina (5) e serotonina (6)

Cada receptor nicotínico de acetilcolina (nAChR) é dividido em três domínios: um grande domínio extracelular *N*-terminal onde encontram-se os sítios ortostéricos, um poro transmembranar e um pequeno domínio intracelular, tendo esse conjunto um comprimento total de aproximadamente 160 Å.⁹ Os nAChRs neuronais apresentam uma grande variedade de funções biológicas que estão relacionadas com as diversas combinações possíveis entre as subunidades α e β , a saber, α ($\alpha 2$ - $\alpha 10$) e β ($\beta 2$ - $\beta 4$), podendo estas combinações serem homoméricas, como os receptores $\alpha 7$, $\alpha 8$ e $\alpha 9$ ou heteroméricas, como os receptores $\alpha 4\beta 2^*$, $\alpha 7\alpha 8^*$, $\alpha 9\alpha 10^*$, $\alpha 6\beta 2^*$, $\alpha 3\beta 4^*$, entre outros, com a presença obrigatória de uma subunidade α no receptor.^{3,5,7} O asterisco representa a possibilidade de outras subunidades no complexo. Classifica-se as subunidades como α ou não- α baseando-se na presença ou ausência, respectivamente,

de cisteínas vicinais ligadas por pontes dissulfeto no domínio extracelular do receptor.¹¹

Ao interagirem com seu neurotransmissor endógeno **1** ou com um agonista, a estrutura dos nAChRs muda de conformação, de modo que é possível descrever, muito simplificada, o processo de ativação do receptor em três etapas: **(a)** a conformação fechada, onde o receptor está apto a interagir com um agonista e não há a passagem de cátions pelo poro permeável, **(b)** a conformação aberta, onde o receptor já interagiu com um agonista e permite a passagem de cátions em um intervalo de tempo da ordem de milissegundos e **(c)** a conformação dessensibilizada, onde o receptor novamente deixa de permitir a passagem de cátions pelo poro central.¹² O poro central é permeável a cátions monovalentes e bivalentes, principalmente sódio e cálcio.^{13,14,15} Estudos sugerem que ao

retornar da conformação dessensibilizada para a conformação fechada não ocorre a passagem pela conformação aberta, indicando um processo cíclico. Os mesmos estudos mostram também que existem múltiplos estados conformacionais fechado, aberto e dessensibilizado conectados em uma sequência complexa que ainda precisa ser elucidada para a compreensão completa do comportamento macroscópico de alguns nAChRs.¹²

O sítio ortostérico dos nAChRs neuronais encontra-se no domínio extracelular do receptor, nas interfaces entre as subunidades α e as demais subunidades que os constituem, de modo que as subunidades desse tipo são o componente primário do sítio ativo e as subunidades do tipo β são responsáveis pela parte complementar do mesmo, como exemplificado na **Figura 3**, sendo que a permeabilidade a cátions e

afinidade por determinado ligante bioativo variam grandemente em função dos componentes complementares do receptor.^{16,17,18} Além disso, os nAChRs costumam apresentar uma subunidade assessora que não participa ativamente da formação do sítio ortostérico mas influencia nas propriedades farmacológicas e biofísicas do receptor.⁷ Dessa forma, fica clara a abundância de configurações possíveis para os nAChRs, uma vez que as diversas subunidades existentes podem se combinar em inúmeras configurações. Além do sítio ortostérico, os nAChRs apresentam também diversos sítios alostéricos onde moléculas específicas podem interagir com o receptor e potencializar ou inibir a ativação causada por um agonista,¹⁹ sendo esta uma das possíveis abordagens utilizadas no planejamento de novas substâncias bioativas que tem como alvo os nAChRs.

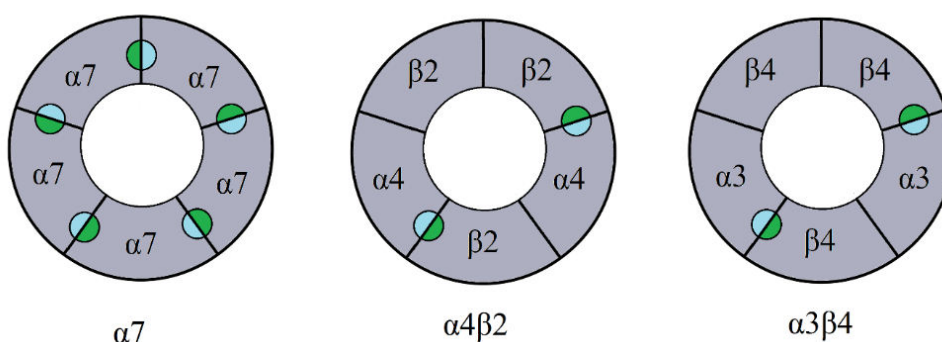


Figura 3. Representação dos nAChRs neuronais abordados nesse artigo e seus sítios ortostéricos, com a face principal em azul e a face complementar em verde. Outros arranjos entre mesmas subunidades apresentadas também são possíveis. Adaptado de Gotti e colaboradores³

Por conta da grande variedade de funções dos nAChRs neuronais, estes são fundamentais no entendimento e tratamento de diversas desordens neuropsiquiátricas, como a síndrome de Tourette, autismo, déficit de atenção, desordem hiperativa, esquizofrenia, epilepsia, depressão, ansiedade, doença de Alzheimer e doença de Parkinson, conforme pode ser observado na **Tabela 1**.^{6,8,10,20-25}

Até o presente momento, ainda não existem estruturas cristalográficas completas para os nAChRs neuronais dos mamíferos. A maior parte do conhecimento estrutural sobre esses receptores é proveniente do estudo de proteínas ligantes de acetilcolina (AChBP) homólogas obtidas de espécies de peixe como a *Torpedo marmorata* (PDB ID: 1OED), espécies moluscos como a *Lymnaea stagnalis* (PDB ID: 1UW6) e *Aplysia californica*

(PDB ID: 2BYN) e das proteínas procarióticas GLIC (canal iônico ligante-dependente da cianobactéria *Gloeobacter violaceus*, PDB ID: 4NPQ) e ELIC (canal iônico ligante-dependente da bactéria *Erwinia chrysanthemi*, PDB ID: 2VLO), entre outras.^{2,9,18,22,26-28}

Tabela 1. Possíveis receptores alvo no tratamento de algumas desordens neurológicas e região do cérebro afetada pela doença. Dados obtidos de Jensen e colaboradores²⁵

Desordem	Receptor	Região do cérebro
Epilepsia	$\alpha 4\beta 2^*$, $\alpha 7^*$, $\beta 4^*$, $\alpha 5^*$	Cortex frontal, Hipocampo
Doença de Alzheimer	$\alpha 4\beta 2^*$, $\alpha 7^*$	Cortex, Hipocampo
Esquizofrenia	$\alpha 4\beta 2^*$, $\alpha 7^*$	Hipocampo
Ansiedade	$\alpha 4^*$, $\alpha 7^*$	Hipocampo
Doença de Parkinson	$\alpha 4\beta 2^*$, $\alpha 6\beta 2\beta 3^*$	Sistema nigroestriatal
Dependência da nicotina	$\alpha 4\beta 2^*$, $\alpha 6\beta 2\beta 3^*$, $\beta 4$	Sistema mesolímbico

2. A relação estrutura-atividade de ligantes de receptores nicotínicos de acetilcolina (nAChRs) do subtipo $\alpha 7$

Em geral, a localização dos nAChRs neuronais do tipo $\alpha 7$ no cérebro se conserva entre diferentes espécies de mamíferos, apesar de diferenças significativas terem sido reportadas entre roedores e primatas em regiões associadas a cognição e processamento sensorial. Esses receptores são expressos grandemente nas regiões do hipocampo, córtex, substância negra e cerebelo.²²

Estudos recentes com roedores indicam que agonistas parciais e totais dos nAChRs $\alpha 7$ são eficientes no tratamento de distúrbios neurológicos como a esquizofrenia^{22,29} e o mal de Parkinson e também no controle de desordens motoras provocadas por medicamentos, como as discinesias induzidas por levodopa que surgem no decorrer do tratamento do mal de Parkinson.^{6,24}

Os nAChRs neuronais do tipo $\alpha 7$ são estruturalmente e funcionalmente distintos dos demais nAChRs, apresentando menor afinidade por nicotina e acetilcolina que os

demais, dessensibilização rápida e alta permeabilidade a Ca^{2+} , o que lhes permite regular diversos mecanismos celulares que dependem desse cátion.^{6,19} São constituídos exclusivamente por alças transmembranares do subtipo $\alpha 7$, o que confere a eles uma maior quantidade de sítios ortostéricos do que os demais nAChRs neuronais.^{16,30} A ativação dos nAChRs $\alpha 7$ desencadeia aumento na liberação de glutamato no hipocampo, demonstrando a importância desses receptores para os mecanismos sinápticos glutamatérgicos e modulação dos circuitos neurais do sistema nervoso central (SNC).^{31,32}

Uma das características dos nAChRs $\alpha 7$ é a sua sensibilidade única ao antagonista α -Bungarotoxina (α -Btx) produzido pela cobra *Bungarus multicinctus*, que se liga permanentemente a esse receptor e impede a abertura do canal iônico, sendo que a utilização dessa substância em conjunto com técnicas de autorradiografia foi essencial na determinação inicial da localização desse receptor no SNC.^{3,7,22}

Existem diversos compostos bioativos que atuam nos nAChRs $\alpha 7$ descritos na literatura, como exemplificado na **Figura 4**. O composto **7** faz parte de uma série de derivados contendo a subunidade estrutural dibenzosuberila inicialmente proposta por

Ulens e colaboradores³³ como um ligante do sítio ortostérico dos nAChRs $\alpha 7$ que apresentou alta afinidade ($K_i = 63,09$ nM) em estudos *in silico* e que não apresentava interação significativa com os nAChRs $\alpha 4\beta 2$ em estudos de deslocamento utilizando o radioligante [³H]epibatidina. Entretanto, os dados obtidos *in vitro* em conjunto com análises já mencionadas mostram que **7** atua como um inibidor não competitivo dos nAChRs $\alpha 7$ e $\alpha 4\beta 2$, bem como dos receptores 5-HT₃, indicando que **7** atua em sítios alostéricos dos receptores dessa família através de um mecanismo comum, provavelmente impedimento estérico do poro iônico central.

A bradaniclina (**8**) é um agonista parcial desenvolvido pela Targacept Inc. que apresenta alta afinidade pelos nAChRs $\alpha 7$ e baixa afinidade pelo receptor de serotonina 5-HT₃. Testes preliminares utilizando roedores sugeriam que **8** apresentava potencial para o tratamento das três categorias de sintomas da esquizofrenia (positivos, negativos e cognitivos) e também efeitos sinérgicos com outros medicamentos antipsicóticos, como a clozapina (Clozaril®, Fazacol®). A substância também foi considerada um medicamento em potencial para o tratamento de Alzheimer e transtorno de déficit de atenção com hiperatividade, mas apesar dos resultados positivos em roedores o composto não obteve aprovação nos testes clínicos para nenhuma das três desordens.³⁴

O estilbeno substituído (**9**) inicialmente foi proposto por Gotti e colaboradores³⁵ como um potente antagonista competitivo e seletivo dos nAChRs $\alpha 7$ presentes em cérebros de galinhas ($Cl_{50} = 119$ nM), mas em estudos posteriores por Angelantonio e colaboradores³⁶ demonstrou-se que **9** também atua nos nAChRs $\alpha 3\beta 4^*$ presentes em células do tipo cromafim de roedores, sendo o enantiômero de configuração *R* mais

potente do que seu atípodo ótico nessa situação ($Cl_{50} = 0,33$ μ M e $1,5$ μ M, respectivamente). Em estudos subsequentes, Gotti e colaboradores³⁷ demonstraram que o composto apresenta maior afinidade por nAChRs $\alpha 7$ de mamíferos, expressos em oócitos da espécie *Xenopus*, do que por nAChRs $\alpha 7$ de aves. Esses dados combinados evidenciam que existem diferenças de seletividade e especificidade significativas entre receptores provenientes de espécies distintas.

O composto **10** foi desenvolvido por Zheng e colaboradores³⁸ tendo como alvo os nAChRs do subtipo $\alpha 9\alpha 10$, que se acredita estarem associados dores crônicas nas regiões do sistema nervoso periférico (SNP). A molécula **10** apresentou inibição adequada do receptor alvo ($Cl_{50} = 0,56$ nM) e também do subtipo $\alpha 7$ (**Figura 4**), ambos expressos em oócitos da espécie *Xenopus*, mas devido à natureza altamente polar da subunidade estrutural catiônica espera-se que o composto apresente dificuldades em ultrapassar a barreira hematoencefálica (BHE). Essa característica torna muito mais interessante o estudo da molécula visando os nAChRs $\alpha 9\alpha 10$, que não estão presentes no SNC. O composto **11** é um peptídeo sintético derivado da classe das α -conotoxinas (α -Ctxs) documentado por Whiteaker e colaboradores³⁹ como um antagonista potente e seletivo dos nAChRs $\alpha 7$. As α -Ctxs naturais são peptídeos pequenos, contendo entre 13 e 19 resíduos de aminoácidos, produzidos por gastrópodes do gênero *Conus* e altamente seletivas para alguns tipos de nAChRs. O peptídeo **11** se liga reversivelmente ao receptor, com recuperação total da sensibilidade após 15 minutos, além de inibir tanto receptores de roedores quanto de humanos com potência comparável, o que faz com que esse antagonista seja potencialmente útil no estudo das funções dos nAChRs $\alpha 7$.⁴⁰

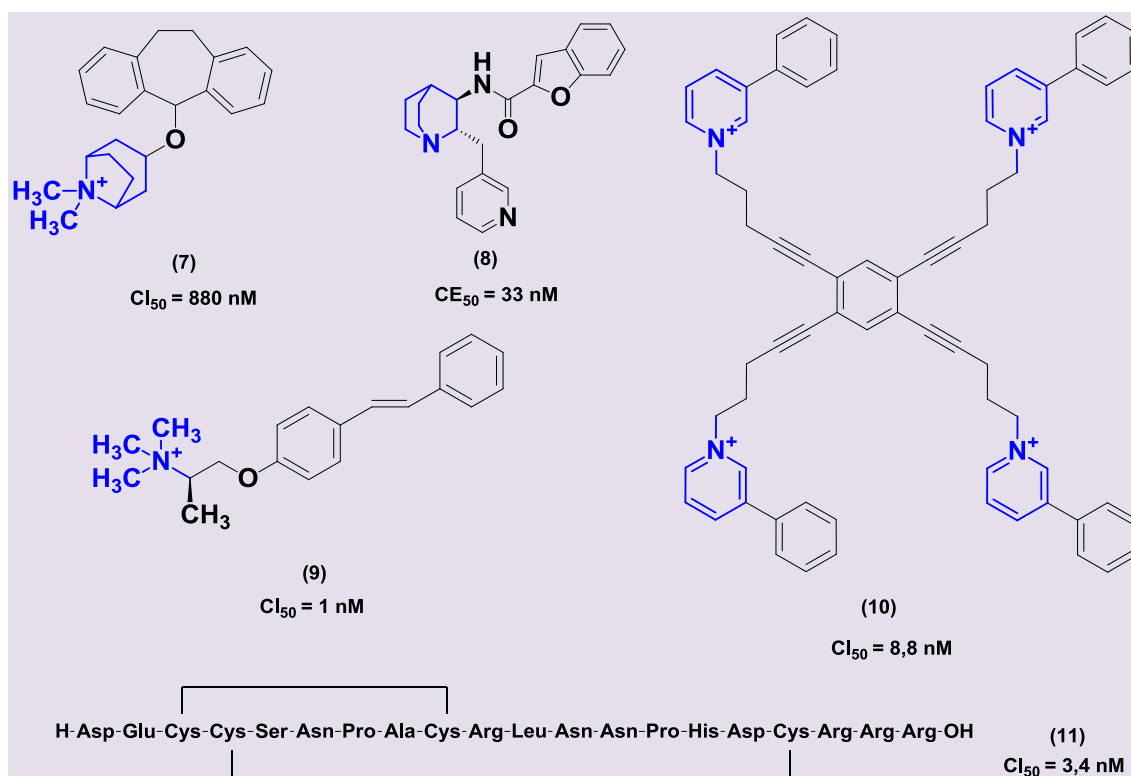


Figura 4. Estrutura química dos compostos **7**³³, **8**⁴¹, **9**³⁷, **10**³⁸, **11**⁴⁰, os quais atuam nos nAChRs do subtipo $\alpha 7$

Na **Figura 5** está representada uma quinuclidina substituída (**12**), análoga da bradaniclina (**8**), e alguns análogos estruturais obtidos por Mazurov e colaboradores⁴² que auxiliam no entendimento de como estes análogos estruturais se relacionam com a afinidade pelo nAChR $\alpha 7$. Substituições realizadas na região do anel piridínico (destacado em vermelho) alteram consideravelmente a afinidade do composto pelo receptor alvo devido a fatores estereoeletrônicos. A troca do substituinte 3-piridina **12** pelos substituintes 2-piridina **13**, 3-quinolina **16**, 2-furano **17** ou benzofuran-2-il **18** resulta em diminuição ou perda de afinidade. Esta perda de afinidade possivelmente se deve a existência de ligações de hidrogênio intramoleculares que conferem uma potencial restrição conformacional aos derivados **13**, **17** e **18** ou fatores estéricos, no caso dos derivados **16** e **18**, dificultando a interação com o sítio ativo.

No caso do composto **15**, a diminuição da afinidade pode estar relacionada à perda da

interação entre a subunidade piridínica e a cadeia lateral de um dos resíduos de aminoácidos do sítio ativo, podendo as interações cátion- π e hidrofóbicas serem as responsáveis pela interação entre a molécula e o biorreceptor alvo. Há pouca diferença de afinidade entre os compostos **12** e **14** devido a semelhança estrutural, visto que não há perda do ponto acceptor de ligação de hidrogênio e a aromaticidade é conservada. A substituição do grupo 4-bromobenzeno (destacado em azul) pelos grupos 4-fluorobenzeno **19** e benzeno **20** não resultou em perda de afinidade significativa, indicando que um halogênio na posição *para* em relação à cadeia principal é dispensável, provavelmente sendo as interações hidrofóbicas ou cátion- π as contribuições mais relevantes para a afinidade pelo receptor. Os compostos **21** e **22** tem menor afinidade por conta do tamanho dos substituintes. A substituição da função carbamato (destacada em verde) pelas funções carbamida **23** e amida **24** não resultou em diminuição significativa da

afinidade pelo receptor devido à semelhança entre esses três grupos, mas a substituição

pela sulfonamida teve como resultado pronunciada redução da afinidade.

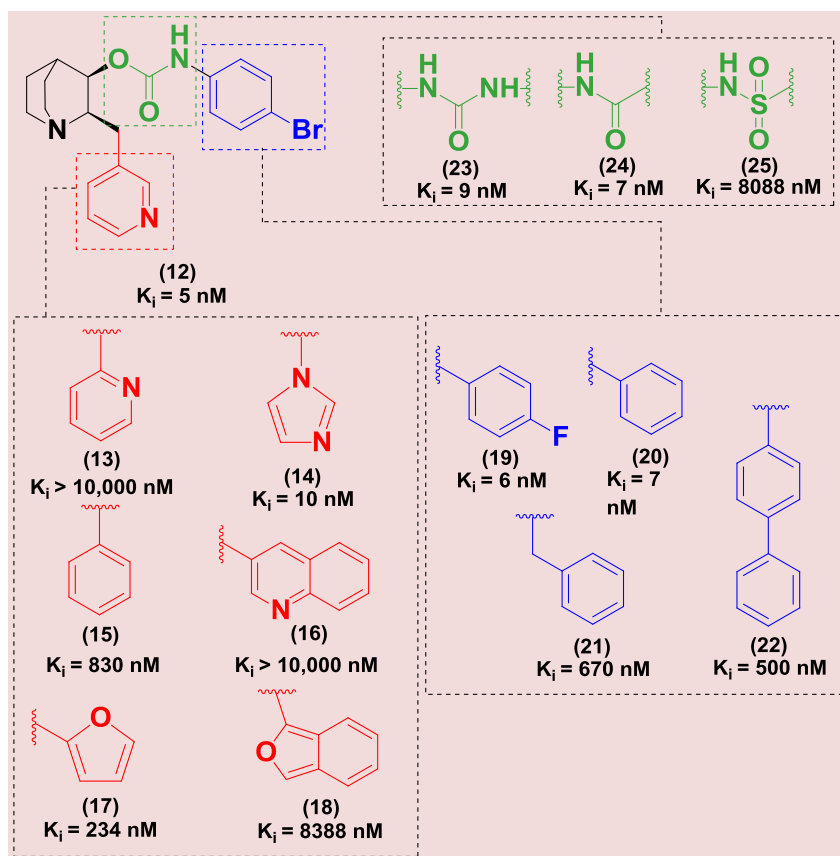


Figura 5. Modificações realizadas por MAZUROV *et al.*, (2005)⁴² na quinuclidina substituída (12) com seus respectivos valores de K_i frente ao nAChR $\alpha 7$

3. A relação estrutura-atividade de ligantes de receptores nicotínicos de acetilcolina (nAChRs) do subtipo $\alpha 4\beta 2$

Os nAChRs $\alpha 4\beta 2$ são o subtipo predominante no SNC dos mamíferos.^{43,44} Esse tipo de receptor é altamente sensível a acetilcolina (1) e nicotina ($CI_{50} = 1,2 \text{ nM}^{45}$) (3), estando estreitamente relacionado ao processo de desenvolvimento de dependência a esta substância em tabagistas.^{43,44} Quando expostos a nicotina e outros agonistas totais ou parciais por longos períodos de tempo sofrem um processo chamado suprarregulação em resposta a

dessensibilização progressiva provocada, sendo esse efeito muito mais pronunciado nesse receptor do que nos demais nAChRs neuronais.^{7,16,18} A ativação dos nAChRs $\alpha 4\beta 2$ aumenta a liberação de dopamina no sistema mesolímbico, proporcionando a sensação de recompensa característica de diversas drogas de abuso.⁴⁶

Estudos bioquímicos e eletrofisiológicos indicam que os nAChRs $\alpha 4\beta 2$ existem em razões estequiométricas distintas, podendo ser constituídos por duas subunidades α e três unidades β ($(\alpha 4\beta 2)_2\beta 2$), resultando em dois sítios ortostéricos nas interfaces $\alpha 4/\beta 2$, ou por três subunidades α e duas subunidades β ($(\alpha 4\beta 2)_2\alpha 4$), resultando em três sítios ortostéricos e sendo esse terceiro sítio localizado na interface $\alpha 4/\alpha 4$ e

apresentando baixa afinidade por acetilcolina e outros agonistas.⁴⁴ As duas relações estequiométricas são diferentes no que diz respeito a cinética de dessensibilização, permeabilidade a cálcio, condutância unitária, sensibilidade a exposição crônica de nicotina e sensibilidade a diversos agonistas. As curvas concentração-resposta da acetilcolina são diferentes para cada estequiometria, sendo que o receptor $(\alpha 4\beta 2)_2\alpha 4$ apresenta uma curva bifásica (**Figura 6b**) característica devido a presença de sítios ortostéricos com afinidades distintas pelo neurotransmissor e o receptor $(\alpha 4\beta 2)_2\beta 2$ apresenta uma curva concentração-resposta monofásica (**Figura 6a**). O receptor $(\alpha 4\beta 2)_2\alpha 4$ apresenta menor sensibilidade a agonistas em comparação com a outra estequiometria possível, uma vez que é necessária a presença de agonistas em todos os sítios ortostéricos para que o mesmo seja completamente ativado.^{30,43,47,48}

Ambas as formas do receptor estão presentes no SNC dos mamíferos, com estudos indicando que a estequiometria preferencialmente presente no cérebro humano é a de menor sensibilidade e que a proporção entre as duas varia quando ocorre exposição crônica a nicotina, aumentando a quantidade de receptores mais sensíveis com o passar do tempo.^{44,47}

Além de mais suscetível a nicotina, o receptor $(\alpha 4\beta 2)_2\beta 2$ também é mais suscetível a dessensibilização causada por exposição prolongada a baixas concentrações dessa substância, enquanto o receptor $(\alpha 4\beta 2)_2\alpha 4$ apresenta maior permeabilidade a íons Ca^{2+} e cinética de dessensibilização mais rápida.⁴⁹ Além destas duas estequiometrias presentes no SNC, nas regiões do córtex e corpo estriado existe uma subpopulação desse receptor que contém a subunidade $\alpha 5$ na forma $(\alpha 4\beta 2)_2\alpha 5$.¹⁶

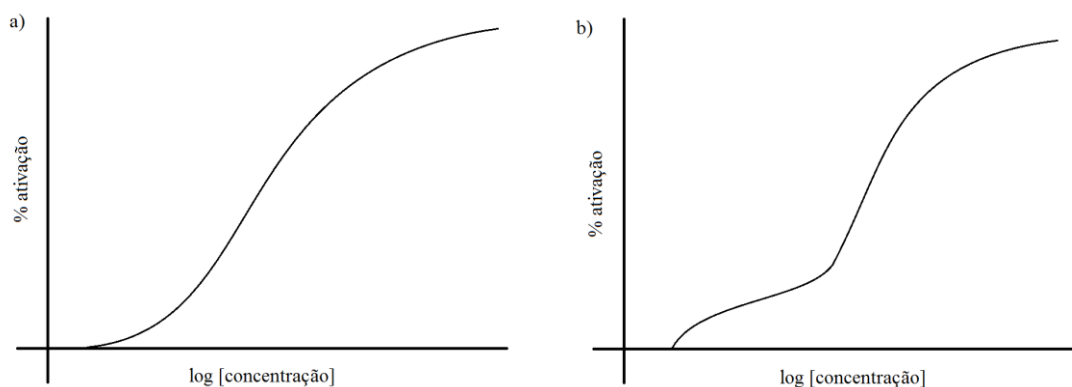


Figura 6. Curvas típicas de resposta monofásica (a) e bifásica (b)

Na **Figura 7** estão representados alguns compostos bioativos que atuam nos nAChRs $\alpha 4\beta 2$. A epibatidina (**26**) é um alcaloide que ocorre naturalmente na pele de sapos da espécie *Epipedobates tricolor* nativos do Equador.⁵⁰ A substância é tóxica, com propriedades analgésicas atribuídas à interação com os nAChRs $\alpha 4\beta 2$ e causando paralisia através da interação com os mAChRs, podendo ser letal. A toxicidade acentuada da epibatidina (**26**) e sua baixa seletividade entre os nAChRs impossibilitou a

sua utilização para fins terapêuticos, mas a substância é largamente utilizada em diversos estudos acadêmicos, como análises de deslocamento de radioligantes, estudos da relação estrutura-atividade de análogos de **26** e avaliação das propriedades farmacológicas desses análogos.^{33,51-53} O fármaco vareniclina (Champix®, Chantix®) (**27**) é um agonista parcial dos nAChRs $\alpha 4\beta 2$ sintetizado por Coe e colaboradores⁵⁴, inspirado no produto natural (-)-citisina, que diminui a liberação de dopamina induzida por

nicotina e induz a baixa liberação deste neurotransmissor no período de abstinência, obtendo aprovação para ser comercializado na Europa e EUA no ano de 2006 pela empresa farmacêutica Pfizer®.⁴⁶ O fármaco **27** é um agonista total dos nAChRs dos subtipos $\alpha 7$ e $\alpha 4\beta 4$ e parcial do subtipo $\alpha 3\beta 4$.^{55,56} Atualmente, a vareniclina (**27**) e a bupropiona (Zyban®) são os únicos fármacos que não contém nicotina (**3**) aprovados pelo FDA para o tratamento da dependência da nicotina (**3**).

O composto TC-8831 (**28**) é um agonista dos nAChRs $\alpha 4\beta 2$ e $\alpha 6/\alpha 3\beta 2\beta 3$ sintetizado por STRACHAN *et al.*, (2014)⁵⁷ que apresenta boa afinidade por ambos os receptores ($K_i = 3$ e 21 nM, respectivamente) (Figura 7). Estudos realizados em primatas indicam que **28** tem efeitos positivos no tratamento das discinesias induzidas por levodopa comumente presentes durante a utilização desse medicamento para o tratamento do

mal de Parkinson.^{58,59} O composto neonicotínico (**29**) foi sintetizado por FAUNDEZ-PARRAGUES *et al.*, (2013)⁴⁵, apresentando boa inibição dos nAChRs $\alpha 4\beta 2$, mas com baixa seletividade entre os subtipos desse receptor. Apresenta também baixa afinidade pelos nAChRs $\alpha 7$ ($Cl_{50} = 2,6 \mu M$). A sazetidina-A (**30**) é um agonista seletivo dos nAChRs $\alpha 4\beta 2$ que possui efeito analgésico e antidepressivo.^{60,61} O composto apresenta atividade agonista total no receptor $(\alpha 4\beta 2)_2\beta 2$ e pouca atividade agonista no receptor $(\alpha 4\beta 2)_2\alpha 4$. Especula-se que a substância converta boa parte dos receptores $(\alpha 4\beta 2)_2\beta 2$ altamente sensíveis para o estado dessensibilizado após uma exposição curta, sendo este o mecanismo responsável por seu longo efeito *in vivo*.⁶² Há evidências de que parte da liberação dopaminérgica provocada pela sazetidina-A (**29**) ocorra devido a ativação de nAChRs $\alpha 6\beta 2^*$.⁶³

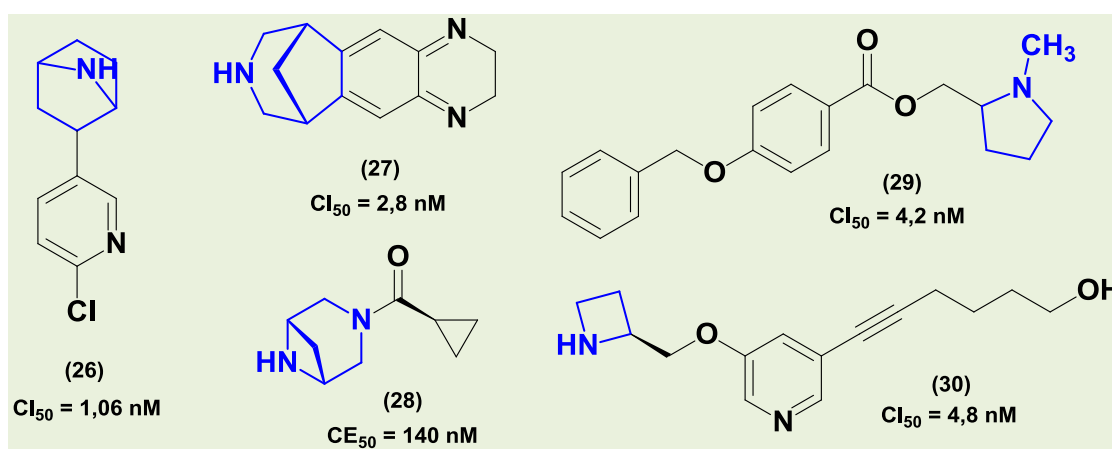


Figura 7. Estrutura química dos compostos bioativos **26**⁶⁴, **27**⁵⁵, **28**⁵⁷, **29**⁴⁵, **30**⁶⁵, os quais atuam nos nAChRs $\alpha 4\beta 2$

Tomando como ponto de partida o éster neonicotínico (**31**), análogo da molécula **29**, tem-se representada na Figura 8 algumas modificações realizadas por Faundez-Parragues e colaboradores⁴⁵ a respeito das quais podem ser feitas algumas considerações sobre a relação estrutural-atividade entre **31** e seus análogos estruturais e o receptor alvo (nAChR $\alpha 4\beta 2$). Primeiramente, os compostos sintetizados

apresentam afinidade e atividade muito maiores no receptor heteromérico $\alpha 4\beta 2$ do que no receptor homomérico $\alpha 7$ (K_i e $Cl_{50} > 1000 \text{ nM}$ para todos os compostos). Isso se deve provavelmente à subunidade N-metilpirrolidina (em azul), também presente na nicotina (**3**), que em pH fisiológico se encontra protonada e pode interagir com o receptor nAChR $\alpha 4\beta 2$ através de interações cátion- π . As substituições realizadas sugerem

que a interação entre o receptor e a região destacada da molécula **31** é de caráter hidrofóbico, uma vez que as alterações que introduziram substituintes de maior hidrofiliidade, como a quinolina **37**, piridinas substituídas **36** e **38** e nitrobenzeno **32** levaram a menores afinidades e potência. Substituições do grupo benzeno por subunidades mais hidrofóbicas produziram compostos com afinidade e potência igual ou maior, como pode ser observado para os compostos **29**, **34** e **35**.

O composto **33** em particular apresentou comportamento incompatível com o dos demais compostos dessa série, uma vez que a 3-fenilamina possui maior hidrofiliidade quando comparada ao benzeno. Logo, além de interações hidrofóbicas, a presença de um potencial doador e aceptor de ligações de hidrogênio favorece uma interação entre o composto e o receptor alvo, possibilidade esta que pode ser reforçada ao observarmos que o composto de maior afinidade dessa série, **29**, possui um oxigênio aceptor de ligação de hidrogênio em uma posição próxima à do nitrogênio da molécula **33**.

Interessantemente, alguns dos compostos sintetizados apresentaram potência inibitória diferenciada em relação às estequiometrias possíveis do receptor alvo. A molécula **31** apresenta maior potência inibitória para a estequiometria $(\alpha4\beta2)_2\beta2$ (2,5 μM) do que para as estequiometrias $(\alpha4\beta2)_2\alpha4$ (313 μM) e $(\alpha4\beta2)_2\alpha5$ (52,3 μM). Substituindo o grupo benzeno pelo grupo 3-fenilamina **33** obtém-se uma pequena diminuição da inibição para o receptor $(\alpha4\beta2)_2\beta2$ (4,3 μM), uma grande redução de atividade para a estequiometria $(\alpha4\beta2)_2\alpha5$ (480 μM) e um aumento significativo para a estequiometria $(\alpha4\beta2)_2\alpha5$ (1,02 μM). O composto **35** apresenta uma baixa afinidade para o receptor $(\alpha4\beta2)_2\beta2$ (3,6 μM) em comparação com **31** e aumento significativo para as estequiometrias $(\alpha4\beta2)_2\alpha5$ (6,0 μM) e $(\alpha4\beta2)_2\alpha4$ (95 μM). Por fim, a molécula **29** apresentou aumento de inibição para as estequiometrias $(\alpha4\beta2)_2\alpha4$ (3,6 μM) e $(\alpha4\beta2)_2\alpha5$ (17,8 μM) e diminuição para o receptor $(\alpha4\beta2)_2\beta2$ (25,8 μM) em comparação com **31**.

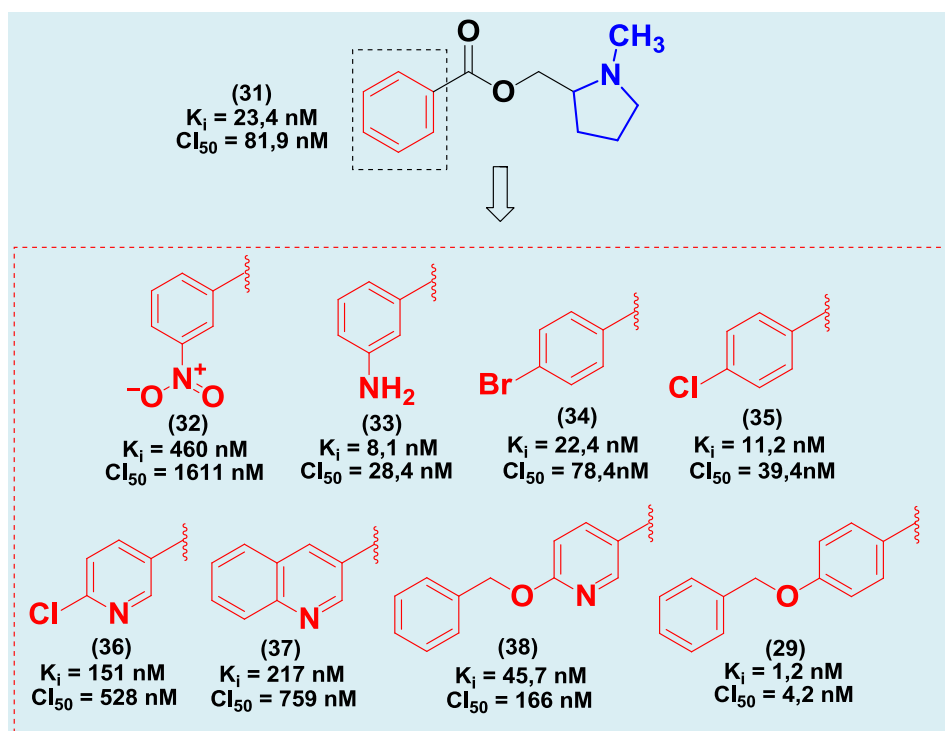


Figura 8. Modificações realizadas por FAUNDEZ-PARRAGUES *et al.*, (2013)⁴⁵ no composto neonicotínico (**31**) e seus respectivos valores de K_i e Cl_{50} para o $nAChR \alpha4\beta2$

4. A relação estrutura-atividade de ligantes de receptores nicotínicos de acetilcolina (nAChRs) do subtipo $\alpha 3\beta 4$

Os nAChRs subtipo $\alpha 3\beta 4$ estão localizados principalmente nas regiões do gânglio sensorial e autônomo, na habênula medial e núcleo interpeduncular.^{56,66,67} Estudos feitos com roedores, envolvendo agonistas e antagonistas parciais, indicam que esse receptor pode estar associado com dependência de diversas drogas de abuso, como nicotina, álcool, cocaína e metanfetamina.^{56,67,68} Também existem evidências de que esse receptor está envolvido no mecanismo que causa diminuição de apetite em tabagistas, sendo este um possível alvo para o planejamento de fármacos de controle de peso e antiobesidade.⁶⁹

As estruturas de alguns compostos bioativos que atuam nos nAChRs $\alpha 3\beta 4$ estão representadas na **Figura 9**. O composto **39**⁷⁰ é um antagonista dos nAChRs $\alpha 3\beta 4$ análogo da bupropiona, um fármaco utilizado no tratamento da dependência de nicotina e como antidepressivo. A substância **39** é um inibidor de recaptção de dopamina e norepinefrina mais potente do que a molécula base, sugerindo possível atividade antidepressiva. Apresenta também Cl_{50} no receptor $\alpha 3\beta 4$ menor (0,62 μM) do que o bupropiona (1,8 μM). O composto **40** é uma quinuclidina substituída disponível comercialmente cujas propriedades farmacológicas foram avaliadas por Kombo e colaboradores.⁷⁰ Ademais, **40** é um antagonista não competitivo dos nAChRs $\alpha 3\beta 4$ que também apresenta propriedades

anti-histamínicas semelhantes a antagonistas de histamina com estrutura similar, como a quifenadina (Fenkarol®).

O composto **41** foi sintetizado por Yu e colaboradores⁷¹ como um agonista dos nAChRs $\alpha 4\beta 2$ ($K_i = 4,6$ nM), apresentando maior afinidade e atividade sobre esse receptor do que nos nAChRs $\alpha 3\beta 4$. Estudos com roedores mostraram que o composto apresenta boa atividade antidepressiva e testes farmacocinéticos não fornecem dados relacionados a toxicidade. OAT-1001 (**42**) foi o primeiro composto a apresentar boa afinidade e seletividade pelos nAChRs $\alpha 3\beta 4$ em relação aos demais nAChRs.⁷² A substância **42** é um agonista parcial nos nAChRs $\alpha 3\beta 4$ que causa dessensibilização do receptor em concentrações nas quais provoca uma corrente interna, gerando um efeito global antagonista, o que classifica o composto como um antagonista funcional desse receptor.^{56,73} A substância **42** e seus análogos são candidatos em potencial para o tratamento da dependência em nicotina (**1**). O composto RegIIA (**43**) é uma α -Ctx presente no veneno do gastrópode *Conusregius* que antagoniza os nAChRs $\alpha 3\beta 4$. Essa toxina **43** também afeta o subtipo $\alpha 7$, mas não tem efeito em receptores do subtipo $\alpha 4\beta 2$. Adicionalmente, **43** apresenta uma sequência de aminoácidos praticamente idêntica a sequência da toxina OmIA da espécie *Conusomaria*, diferindo apenas na presença de um resíduo de glicina adicional na região C-terminal da OmIA. Essa pequena diferença confere as toxinas perfis de seletividade distintos. Essa toxina é uma ferramenta em potencial para o estudo das funções do receptor $\alpha 3\beta 4$ nos mecanismos envolvidos no desenvolvimento da dependência da nicotina, uma vez que não interage com os nAChRs do subtipo $\alpha 4\beta 2$.

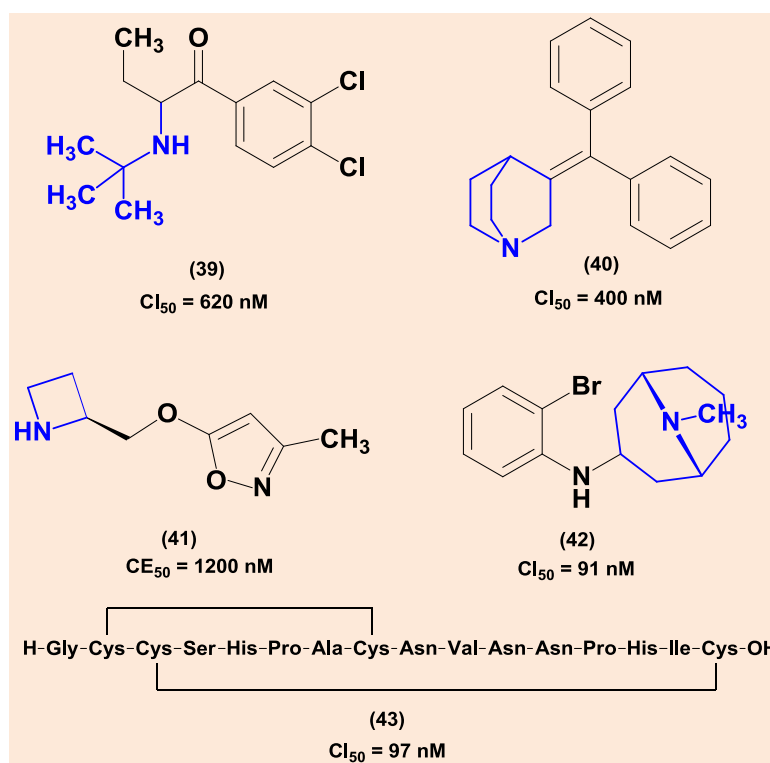


Figura 9. Estrutura química dos compostos bioativos **39**⁷⁴, **40**⁷⁰, **41**⁷¹, **42**⁵⁶, **43**⁷⁵, os quais atuam nos nAChRs $\alpha 3\beta 4$

A **Figura 10** apresenta uma série de compostos **44-54** sintetizados por Zaveri e colaboradores⁶⁶ que auxiliaram no entendimento da relação estrutura-atividade entre ligantes azabicíclicos e o receptor alvo (nAChRs $\alpha 3\beta 4$). A diidroindolinona substituída (**44**) (SR16584) apresenta alta afinidade pelo nAChRs $\alpha 3\beta 4$ e é seletiva para esse subtipo, sendo inativa frente ao subtipo $\alpha 4\beta 2$. A substituição da subunidade nonano aza-bicíclica de **44** por subunidades azepanobicíclicas **45** e **46** resultou em grande redução da afinidade pelo receptor alvo para **45** e em perda da afinidade para **46**, mantendo também a falta de afinidade pelo receptor $\alpha 4\beta 2$ em ambos os compostos ($K_i > 10.000$), provavelmente em decorrência de diferenças estéricas nesse biciclo, uma vez que o nitrogênio protonável dessa região estará localizado abaixo do plano do anel de seis membros, desfavorecendo as interações com o receptor. Além disso, é possível observar que a subunidade nonano aza-bicíclica apresenta interações mais favoráveis com o receptor do que a subunidade octano aza-bicíclica **49**, visto que também houve

uma diminuição considerável da afinidade pelo nAChRs $\alpha 3\beta 4$ em relação ao composto **44** e que a falta de afinidade pelo receptor $\alpha 4\beta 2$ se manteve inalterada. A substituição do grupo destacado em **44** por estruturas piperidínicas substituídas nos compostos **47** e **48** resultou em moléculas sem afinidade em ambos os receptores, indicando incompatibilidade desta subunidade estrutural frente ao receptor alvo.

A substituição da subunidade diidroindolinona (em vermelho) pela 3-piridina, juntamente com a incorporação de uma subunidade éster resultam em perda de afinidade pelo subtipo $\alpha 3\beta 4$, como pode ser observado no composto **50**. O composto **51** apresenta boa afinidade pelos subtipos $\alpha 3\beta 4$ e $\alpha 4\beta 2$ ($K_i = 10,4 \text{ nM}$), mas a substituição do grupo 3-piridina pelo grupo 3-quinolina **52** resulta em diminuição de afinidade por ambos os receptores ($K_i = 423 \text{ nM}$ para o subtipo $\alpha 4\beta 2$) devido a fatores estéricos. Nos compostos **53** e **54**, a substituição da 3-piridina **53** pela 3-quinolina **54**, simultaneamente com a troca da posição da

ligação dupla no anel azabíclico resultam em diminuição da afinidade para ambos os receptores, novamente por fatores estéricos, apesar da diminuição ser muito menos pronunciada para o subtipo $\alpha 4\beta 2$, pelo qual as moléculas apresentam elevada afinidade

($K_i = 1,95$ nM para **53** e $K_i = 29,3$ nM para **54**). Esses resultados, em conjunto com o par **51** e **52** sugerem que o sítio ortostérico do subtipo $\alpha 4\beta 2$ é capaz de acomodar ligantes um pouco mais volumosos do que o sítio do receptor $\alpha 3\beta 4$.

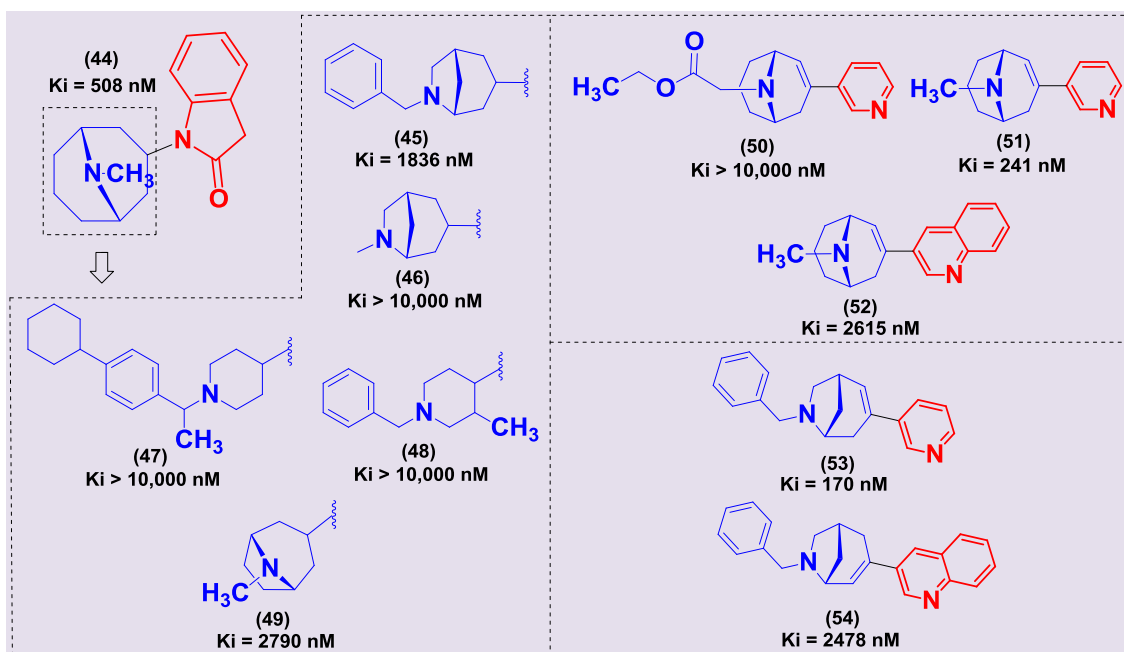


Figura 10. Modificações realizadas por ZAVERI et al., (2010)⁶⁶ no composto azabíclico (**41**) e seus respectivos valores de K_i

5. Conclusão

Os nAChRs neuronais dos subtipos $\alpha 7, \alpha 4\beta 2$ e $\alpha 3\beta 4$ são alvos promissores para o tratamento de diversas desordens do SNC, apresentando ampla quimiodiversidade de ligantes, variados resultados de bioatividade (K_i , Cl_{50} , EC_{50} , etc.), informações referentes a atividade intrínseca (agonista plena, agonista parcial e antagonista) e toxicidade, assim como sobre atuação nos sítios ortostéricos e alostéricos dos nAChRs. Estas informações podem ser utilizadas na obtenção de novos candidatos a fármacos mais promissores. Neste contexto, um dos desafios mais pertinentes envolvendo o estudo desses receptores é o desenvolvimento de compostos bioativos que sejam mais eficazes, seletivos e com problemas de toxicidades

reduzidos ou eliminados. No aspecto estrutural, a avaliação adequada da contribuição de cada subunidade do receptor ($\alpha 3$, $\alpha 4$, $\alpha 7$, $\beta 2$ e $\beta 4$) para a atividade de diferentes ligantes bioativos também merece destaque, visando a melhor compreensão da interação ligante-biorreceptor alvo de forma mais detalhada (nAChRs $\alpha 7, \alpha 4\beta 2$ e $\alpha 3\beta 4$).

Agradecimentos

Agradecemos os órgãos de fomento à pesquisa que apoiam o Laboratório de Química Medicinal, Síntese Orgânica e Modelagem Molecular (LaQMedSOMM). INCT – INOFAR. Processo: CNPq 573.564/2008-6 e FAPERJ E-

26/170.020/2008. **PROPe – UNESP**. Primeiros Projetos. Processo: # 840.

Referências Bibliográficas

¹ Dani, J. A.; Bertrand, D. Nicotinic acetylcholine receptors and nicotinic cholinergic mechanisms of the central nervous system. *Annual Reviews Pharmacology Toxicology* **2007**, *47*, 699. [[CrossRef](#)] [[PubMed](#)]

² Albuquerque, E. X.; Pereira, E. F. R.; Alkondon, M.; Rogers, S. W. Mammalian Nicotinic Acetylcholine Receptors: From Structure to Function. *Physiological Reviews* **2009**, *89*, 73. [[CrossRef](#)] [[PubMed](#)]

³ Gotti, C.; Zoli, M.; Clementi, F. Brain nicotinic acetylcholine receptors: native subtypes and their relevance. *Trends in Pharmacological Sciences* **2006**, *27*, 482. [[CrossRef](#)] [[PubMed](#)]

⁴ Högger, P.; Sadée, W.; Laméh, J. The muscarinic acetylcholine receptors. *Biomembranes* **1996**, *2B*, 301. [[CrossRef](#)]

⁵ Hogg, R. C.; Raggenbass, M.; Bertrand, D. Nicotinic acetylcholine receptors: from structure to brain function. *Reviews of Physiology, Biochemistry and Pharmacology* **2003**, *147*, 1. [[CrossRef](#)] [[PubMed](#)]

⁶ Quik, M.; Zhang, D.; McGregor, M.; Bordia, T. Alpha7 nicotinic receptors as therapeutic targets for Parkinson's disease. *Biochemical Pharmacology* **2015**, *97*, 399. [[CrossRef](#)] [[PubMed](#)]

⁷ Gotti, C.; Clementi, F.; Fornari, A.; Gaimarri, A.; Guiducci, S.; Manfredi, I.; Moretti, M.; Pedrazzi, P.; Pucci, L.; Zoli, M. Structural and functional diversity of native brain neuronal nicotinic receptors. *Biochemical Pharmacology* **2009**, *78*, 703. [[CrossRef](#)] [[PubMed](#)]

⁸ Dineley, K. T.; Pandya, A. A.; Yakel, J. L. Nicotinic ACh receptors as therapeutic targets in CNS disorders. *Trends in Pharmacological Sciences* **2015**, *36*, 96. [[CrossRef](#)] [[PubMed](#)]

⁹ Unwin, N. Refined structure of the nicotinic acetylcholine receptor at a 4Å resolution.

Journal of Molecular Biology **2005**, *346*, 967. [[CrossRef](#)] [[PubMed](#)]

¹⁰ Gotti, C.; Clementi, F. Neuronal nicotinic receptors: from structure to pathology. *Progress in Neurobiology* **2004**, *74*, 363. [[CrossRef](#)] [[PubMed](#)]

¹¹ Stokes, C.; Treinin, M.; Papke, R. L. Looking below the surface of nicotinic acetylcholine receptors. *Trends in Pharmacological Sciences* **2015**, *36*, 514. [[CrossRef](#)] [[PubMed](#)]

¹² Auerbach, A. Activation of endplate acetylcholine receptors by agonists. *Biochemical Pharmacology* **2015**, *97*, 601. [[CrossRef](#)] [[PubMed](#)]

¹³ Dajas-Bailador, F.; Wonnacott, S. Nicotinic acetylcholine receptors and the regulation of neuronal signaling. *Trends in Pharmacological Sciences* **2004**, *25*, 317. [[CrossRef](#)] [[PubMed](#)]

¹⁴ Lax, P.; Fucile, S.; Eusebi, F. Ca²⁺ Permeability of human heteromeric nAChRs expressed by transfection in human cells. *Cell Calcium* **2002**, *32*, 53. [[CrossRef](#)] [[PubMed](#)]

¹⁵ Nutter, T. J.; Adams, D. J. Monovalent and divalent cation permeability and block of neuronal nicotinic receptor channels in rat parasympathetic ganglia. *The Journal of General Physiology* **1995**, *105*, 701. [[PubMed](#)]

¹⁶ Gotti, C.; Moretti, M.; Gaimarri, A.; Zanardi, A.; Clementi, F.; Zoli, M. Heterogeneity and complexity of native brain nicotinic receptors. *Biochemical Pharmacology* **2007**, *74*, 1102. [[CrossRef](#)] [[PubMed](#)]

¹⁷ Hurst, R.; Rollema, H.; Bertrand, D. Nicotinic acetylcholine receptors: From basic science to therapeutics. *Pharmacology and Therapeutics* **2013**, *137*, 22. [[CrossRef](#)] [[PubMed](#)]

¹⁸ Papke, R. L. Merging old and new perspectives on nicotinic acetylcholine receptors. *Biochemical Pharmacology* **2014**, *89*, 1. [[CrossRef](#)] [[PubMed](#)]

¹⁹ Chatzidaki, A.; Millar, N. S. Allosteric modulation of nicotinic acetylcholine receptors. *Biochemical Pharmacology* **2015**, *97*, 408. [[CrossRef](#)] [[PubMed](#)]

²⁰ Deutsch, S. I.; Burket, J. A.; Urbano, R. M.; Benson, A. D. The α7 nicotinic acetylcholine receptor: A mediator of pathogenesis and therapeutic target in autism spectrum disorders and Down syndrome. *Biochemical*

- Pharmacology* **2015**, *97*, 363. [CrossRef] [PubMed]
- ²¹ Lombardo, S.; Maskos, U. Role of the nicotinic acetylcholine receptor in Alzheimer's disease pathology and treatment. *Neuropharmacology* **2015**, *96*, 255. [CrossRef] [PubMed]
- ²² Wallace, T. L.; Bertrand, D. Neuronal $\alpha 7$ Nicotinic Receptors as a Target for the Treatment of Schizophrenia. *International Review of Neurobiology* **2015**, *124*, 79. [CrossRef] [PubMed]
- ²³ Kutlu M. G.; Parikh V.; Gould, T. J. Nicotine Addiction and Psychiatric Disorders. *International Review of Neurobiology* **2015**, *124*, 171. [CrossRef] [PubMed]
- ²⁴ Quik, M.; Bordia, T.; Zhang D.; Perez, X. A. Nicotine and Nicotinic Receptor Drugs: Potential for Parkinson's Disease and Drug-Induced Movement Disorders. *International Review of Neurobiology* **2015**, *124*, 247. [CrossRef] [PubMed]
- ²⁵ Jensen, A. A.; Frølund, B.; Liljefors, T.; Krogsgaard-Larsen, P. Neuronal Nicotinic Acetylcholine Receptors: Structural Revelations, Target Identifications and Therapeutic Inspirations. *Journal of Medicinal Chemistry* **2005**, *48*, 4705. [CrossRef] [PubMed]
- ²⁶ Billen, B.; Spurny, R.; Brams, M.; van Elk, R.; Valera-Kummer, S.; Yakel, J. L.; Voets, T.; Bertrand, D.; Smit, A. B.; Ulens, C. Molecular actions of smoking cessation drugs at $\alpha 4\beta 2$ nicotinic receptors defined in crystal structures of a homologous binding protein. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 9173. [CrossRef] [PubMed]
- ²⁷ Bourne, Y.; Sulzenbacher, G.; Radic, Z.; Aráoz, R.; Reynaud, M.; Benoit, E.; Zakarian, A.; Servent, D.; Molgó, J.; Taylor, P.; Marchot, P. Marine macrocyclic imines, pinnatoxins A and G: structural determinants and functional properties to distinguish neuronal $\alpha 7$ from muscle $\alpha 1_2\beta\gamma\delta n$ AChRs. *Structure* **2015**, *23*, 1106. [CrossRef] [PubMed]
- ²⁸ Spurny, R.; Debaveye, S.; Farinha, A.; Veys, K.; Vos, A. M.; Gossas, T.; Atack, J.; Bertrand, S.; Bertrand, D.; Danielson, U. H.; Tresadern, G.; Ulens, C. Molecular blueprint of allosteric binding sites in a homologue of the agonist-binding domain of the $\alpha 7$ nicotinic acetylcholine receptor. *Proceedings of the National Academy of Sciences of the United States of America* **2015**, *112*, E2543. [CrossRef] [PubMed]
- ²⁹ Kohlhaas, K. L.; Robb, H. M.; Roderwald, V. A.; Rueter, L. E. Nicotinic modulation of auditory evoked potential electroencephalography in a rodent neurodevelopmental model of schizophrenia. *Biochemical Pharmacology* **2015**, *97*, 482. [CrossRef] [PubMed]
- ³⁰ Wang, J.; Kuryatov, A.; Sriram, A.; Jin, Z.; Kamenecka, T. M.; Kenny, P. J.; Lindstrom, J. An Accessory Agonist Binding Site Promotes Activation of $\alpha 4\beta 2$ *Nicotinic Acetylcholine Receptors. *The Journal of Biological Chemistry* **2015**, *290*, 13907. [CrossRef] [PubMed]
- ³¹ Cheng, Q.; Yakel, A. L. The effect of $\alpha 7$ nicotinic receptor activation on glutamatergic transmission in the hippocampus. *Biochemical Pharmacology* **2015**, *97*, 439. [CrossRef] [PubMed]
- ³² Koukouli, F.; Maskos, U. The multiple roles of the $\alpha 7$ nicotinic acetylcholine receptor in modulating glutamatergic systems in the normal and diseased nervous system. *Biochemical Pharmacology* **2015**, *97*, 378. [CrossRef] [PubMed]
- ³³ Ulens, C.; Akdemir, A.; Jongejan, A.; van Elk, R.; Bertrand, S.; Perrakis, A.; Leurs, R.; Smit, A. B.; Sixma, T. K.; Bertrand, D.; de Esch, J. P. Use Of Acetylcholine Binding Protein In The Search For Novel $\alpha 7$ Nicotinic Receptor Ligands. In Silico Docking, Pharmacological Screening, And X-Ray Analysis. *Journal of Medicinal Chemistry* **2009**, *52*, 2372. [CrossRef] [PubMed]
- ³⁴ Sítio Adis Insight. Disponível em<<http://adisinsight.springer.com/drugs/800026693>>. Acesso em: 20 janeiro 2016.
- ³⁵ Gotti, C.; Balestra, B.; Moretti, M.; Rovati, G. E.; Maggi, L.; Rossoni, G.; Berti, F.; Villa, L.; Pallavicini, M.; Clementi, F. 4-oxystilbene compounds are selective ligands for neuronal nicotinic acetylcholine α Bungarotoxin receptors. *British Journal of Pharmacology* **1998**, *124*, 1197. [CrossRef] [PubMed]

- ³⁶ Di Angelantonio, S.; Nistri, A.; Moretti, M.; Clementi, F.; Gotti, C. Antagonism of nicotinic receptors of rat chromaffin cells by *N,N,N*-trimethyl-1-(4-trans-stilbenoxy)-2-propylammonium iodide: a patch clamp and ligand binding study. *British Journal of Pharmacology* **2000**, *129*, 1771. [[CrossRef](#)] [[PubMed](#)]
- ³⁷ Gotti, C.; Carbonnelle, E.; Moretti, M.; Zwart, R.; Clementi, F. Drugs selective for nicotinic receptor subtypes: a real possibility or a dream? *Behavioural Brain Research* **2000**, *113*, 183. [[CrossRef](#)] [[PubMed](#)]
- ³⁸ Zheng, G.; Zhang, Z.; Dowell, C.; Wala, E.; Dvoskin, L. P.; Holtman, J. R.; McIntosh, J. M.; Crooks, P. A. Discovery of non-peptide, small molecule antagonists of $\alpha 9\alpha 10$ nicotinic acetylcholine receptors as novel analgesics for the treatment of neuropathic and tonic inflammatory pain. *Bioorganic & Medicinal Chemistry* **2011**, *21*, 2476. [[CrossRef](#)] [[PubMed](#)]
- ³⁹ Whiteaker, P.; Christensen, S.; Yoshikami, D.; Dowell, C.; Watkins, M.; Gulyas, J.; River, J.; Olivera, B. M.; McIntosh, J. M. Discovery, Synthesis, And Structure Activity of a Highly Selective $\alpha 7$ Nicotinic Acetylcholine Receptor Antagonist. *Biochemistry* **2007**, *46*, 6628. [[CrossRef](#)] [[PubMed](#)]
- ⁴⁰ Innocent, N.; Livingstone, P. D.; Hone, A.; Kimura, A.; Young, T.; Whiteaker, P.; McIntosh, J. M.; Wonnacott, S. α Conotoxin Ar1B[V11L, V16D] is a potent and selective antagonist at rat and human native $\alpha 7$ nicotinic acetylcholine receptors. *The Journal of Pharmacology and Experimental Therapeutics* **2008**, *327*, 529. [[CrossRef](#)] [[PubMed](#)]
- ⁴¹ Mazurov, A. A.; Speake, A. D.; Yohannes, D. Discovery and Development of $\alpha 7$ Nicotinic Acetylcholine Receptor Modulators. *Journal of Medicinal Chemistry* **2011**, *54*, 7943. [[CrossRef](#)] [[PubMed](#)]
- ⁴² Mazurov, A.; Klucik, J.; Miao, L.; Phillips, T. Y.; Seamans, A.; Schmitt, J. D.; Hauser, T. A.; Johnson, R. T.; Miller, C. 2-(Arylmethyl)-3-substituted quinuclidines as selective $\alpha 7$ nicotinic receptor ligands. *Bioorganic & Medicinal Chemistry Letters* **2005**, *15*, 2073. [[CrossRef](#)] [[PubMed](#)]
- ⁴³ Mazzaferro, S.; Benallegue, N.; Carbone, A.; Gasparri, F.; Vijayan, R.; Biggin, P. C.; Moroni, M.; Bermudez, I. Additional Acetylcholine (ACh) Binding Site at $\alpha 4/\alpha 4$ Interface of $(\alpha 4\beta 2)_2\alpha 4$ Nicotinic Receptor Influences Agonist Sensitivity. *The Journal of Biological Chemistry* **2011**, *286*, 31043. [[CrossRef](#)] [[PubMed](#)]
- ⁴⁴ Nelson, M. E.; Kuryatov, A.; Choi, C. H.; Zhou, Y.; Lindstrom, J. Alternate Stoichiometries of $\alpha 4\beta 2$ Nicotinic Acetylcholine Receptors. *Molecular Pharmacology* **2003**, *63*, 332. [[CrossRef](#)] [[PubMed](#)]
- ⁴⁵ Faundez-Parraguez, M.; Farias-Rabelo, N.; Gonzalez-Gutierrez, J. P.; Etcheverry-Berrios, A.; Alzate-Morales, J.; Adasme-Carreño, F.; Varas, R.; Bermudez, S.; Iturriaga-Vasquez, P. Neonicotinic analogues: Selective antagonists for $\alpha 4\beta 2$ nicotinic acetylcholine receptors. *Bioorganic Medicinal Chemistry* **2013**, *21*, 2687. [[CrossRef](#)] [[PubMed](#)]
- ⁴⁶ Niaura, R.; Jones, C. J.; Kirkpatrick, P. Varenicline. *Nature Reviews Drug Discovery* **2006**, *5*, 537. [[CrossRef](#)] [[PubMed](#)]
- ⁴⁷ Son, C. D.; Moss, F. J.; Cohen, B. N.; Lester, H. A. Nicotine Normalizes Intracellular Subunit Stoichiometry of Nicotinic Receptors Carrying Mutations Linked to Autosomal Dominant Nocturnal Frontal Lobe Epilepsy. *Molecular Pharmacology* **2009**, *75*, 1137. [[CrossRef](#)] [[PubMed](#)]
- ⁴⁸ Harpsøe, K.; Ahring, P. K.; Christensen, J. K.; Jensen, M. L.; Peters, D.; Balle, T. Unraveling the High-and Low-Sensitivity Agonist Responses of Nicotinic Acetylcholine Receptors. *The Journal of Neuroscience* **2011**, *31*, 10759. [[CrossRef](#)] [[PubMed](#)]
- ⁴⁹ Grupe, M.; Grunnet, M.; Bastlund, J. F.; Jensen, A. A. Targeting $\alpha 4\beta 2$ Nicotinic Acetylcholine Receptors in Central Nervous System Disorders: Perspectives on Positive Allosteric Modulation as a Therapeutic Approach. *Basic & Clinical Pharmacology & Toxicology* **2015**, *116*, 187. [[CrossRef](#)] [[PubMed](#)]
- ⁵⁰ Gerzanich, V.; Peng, X.; Wang, F.; Wells, G.; Anand, T.; Fletcher, S.; Lindstrom, J. Comparative Pharmacology of Epibatine: A Potent Agonist for Neuronal Nicotinic

- Acetylcholine Receptors. *Molecular Pharmacology* **1995**, *48*, 774. [PubMed]
- ⁵¹ Dallanoce, C.; Matera, C.; Pucci, L.; Gotti, C.; Clementi, F.; De Amici, M.; De Micheli, C. Synthesis and binding affinity at $\alpha 4\beta 2$ and $\alpha 7$ nicotinic acetylcholine receptors of new analogs of epibatidine and epiboxidine containing the 7-azabicyclo[2.2.1]hept-2-ene ring system. *Bioorganic & Medicinal Letters* **2012**, *22*, 829. [CrossRef] [PubMed]
- ⁵² Mu, L.; Drandarov, K.; Bisson, W. H.; Schibig, A.; Wirz, C.; Schubiger, P. A.; Westera, G. Synthesis and binding studies of epibatidine analogues as ligands for the nicotinic acetylcholine receptors. *European Journal of Medicinal Chemistry* **2006**, *41*, 640. [CrossRef] [PubMed]
- ⁵³ Carroll, F. I. Epibatidine structure-activity relationships. *Bioorganic & Medicinal Chemistry Letters* **2004**, *14*, 1889. [CrossRef] [PubMed]
- ⁵⁴ Coe, J. W.; Brooks, P. R.; Vetelino, M. G.; Wirtz, M. C.; Arnold, E. P.; Huang, J.; Sands, S. B.; Davis, T. I.; Lebel, L. A.; Fox, V. B.; Shrikhande, A.; Heym, J. H.; Shaeffer, E.; Rollema, H.; Lu, Y.; Mansbach, R. S.; Chambers, L. K.; Rovetti, C. C.; Shulz, D. W.; Tingley, F. D.; O'Neill, B. T. Varenicline: An $\alpha 4\beta 2$ Nicotinic Receptor Partial Agonist for Smoking Cessation. *Journal of Medicinal Chemistry* **2005**, *48*, 3474. [CrossRef] [PubMed]
- ⁵⁵ Arias, H. R.; Feuerbach, D.; Targowska-Duda, K.; Kackzor, A. A.; Poso, A.; Jozwiak, K. Pharmacological and molecular studies on the interaction of varenicline with different nicotinic acetylcholine receptor subtypes. Potential mechanism underlying partial agonism at human $\alpha 4\beta 2$ and $\alpha 3\beta 4$ subtypes. *Biochimica et Biophysica Acta - Biomembranes* **2015**, *1848*, 731. [CrossRef] [PubMed]
- ⁵⁶ Cippitelli, A.; Wu, J.; Gaiolini, K. A.; Mercatelli, D.; Schoch, J.; Gorman, M.; Ramirez, A.; Ciccocioppo, R.; Khroyan, T. V.; Yasuda, D.; Zaveri, N. T.; Pascual, C.; Xie, X.; Toll, L. AT-1001: A High-Affinity $\alpha 3\beta 4$ nAChR Ligand With Novel Nicotine-Suppressive Pharmacology. *British Journal of Pharmacology* **2015**, *172*, 1834. [CrossRef] [PubMed]
- ⁵⁷ Strachan, J. P.; Kombo, D. C.; Mazurov, A.; Heemstra, R.; Bhatti, B. S.; Akiredt, R.; Murthy, S.; Miao, L.; Jett, J. E.; Speake, J.; Bencherif, M. Identification and pharmacological characterization of 3,6-diazabicyclo[3.1.1]heptane-3-carboxamides as novel ligands for the $\alpha 4\beta 2$ and $\alpha 6/\alpha 3\beta 2\beta 3$ nicotinic acetylcholine receptors (nAChRs). *European Journal of Medicinal Chemistry* **2014**, *86*, 60. [CrossRef] [PubMed]
- ⁵⁸ Johnston, T. H.; Hout, P.; Fox, S. H.; Koprach, J. B.; Szeliga, K. T.; Jamed, J. W.; Graef, J. D.; Letchworth, S. R.; Jordan, K. G.; Hill, M. P.; Brotchiem J. M. TC-8831, a nicotinic acetylcholine receptor agonist, reduces L-DOPA-induced dyskinesia in the MPTP macaque. *Neuropharmacology* **2013**, *73*, 337. [CrossRef] [PubMed]
- ⁵⁹ Zhang, D.; Mallela, A.; Sohn, D.; Carrol, F. I.; Bencherif, M.; Letchworth, S.; Quik, M. Nicotinic Receptor Agonist Reduce L-DOPA-Induced Dyskinesias in a Monkey Model of Parkinson's Disease. *The Journal of Pharmacology and Experimental Therapeutics* **2013**, *347*, 225. [CrossRef] [PubMed]
- ⁶⁰ Cucchiari, G.; Xiao, Y.; Gonzales-Sulser, A.; Kellar, K. J. Analgesic Effects of Sazetidine-A, a New Nicotinic Cholinergic Drug. *Anesthesiology* **2008**, *109*, 512. [CrossRef] [PubMed]
- ⁶¹ Kozikowski, A. P.; Eaton, J. B.; Bajjuri, K. M.; Chellapan, A. K.; Chen, Y.; Karadi, S.; He, R.; Caldarone, B.; Manzano, M.; Yuen, P. W.; Lukas, R. J. Chemistry and Pharmacology of Nicotinic Ligands Based on 6-[5-(Azetidino-2-ylmethoxy)pyridine-3-yl]hex-5-yn-1-ol (AMOP-H-OH) for Possible Use in Depression. *ChemMedChem* **2009**, *4*, 1279. [CrossRef] [PubMed]
- ⁶² Xiao, Y.; Tuan, E.; Yasuda, R. P.; Sahibzada, N.; Wolfe, B. B.; Horton, L.; Tran, T.; Al-Muhtasib, N.; Iwueze, A. F.; Dipietro, J. R.; Xie, T.; Paige, M.; Brown, M. L.; Kellar, K. J. Pharmacological properties of sazetidine-A, a selective ligand of $\alpha 4\beta 2$ nicotinic acetylcholine receptors. *Biochemical Pharmacology* **2011**, *82*, 1029. [CrossRef]

- ⁶³ Zwart, R.; Carbone, A. L.; Moroni, M.; Bermudez, S.; Mogg, A. J.; Folly, E. A.; Broad, L. M.; Williams, A. C.; Zhang, D.; Ding, C.; Heinz, B. A.; Sher, E. Sazetidine-A Is a Potent and Selective Agonist at Native and Recombinant $\alpha 4\beta 2$ Nicotinic Acetylcholine Receptors. *Molecular Pharmacology* **2008**, *73*, 1838. [[CrossRef](#)] [[PubMed](#)]
- ⁶⁴ Sun, Q.; Yue, C. Q.; Ye, J.; Li, C. L.; Cheng, T. M.; Li, R. T. Unique spirocyclopiperazinium salt III: Further investigation of monospirocyclopiperazinium (MSPZ) salts as potential analgesics. *Bioorganic & Medicinal Chemistry Letters* **2007**, *17*, 6245. [[CrossRef](#)] [[PubMed](#)]
- ⁶⁵ Zhang, H.; Tückmantel, W.; Eaton, A. B.; Yuen, P. W.; Yu, L. F.; Bajjuri, K. M.; Fedolak, A.; Wang, D.; Ghavami, A.; Caldarone, B.; Paterson, N. E.; Lowe, D. A.; Brunner, D.; Lukas, R. J.; Kozikowski, A. P. Chemistry and Behavioral Studies Identify Chiral Cyclopropanes as Selective $\alpha 4\beta 2$ -Nicotinic Acetylcholine Receptor Partial Agonist Exhibiting an Antidepressant Profile. *Journal of Medicinal Chemistry* **2011**, *55*, 717. [[CrossRef](#)] [[PubMed](#)]
- ⁶⁶ Zaveri, N.; Jiang, F.; Olsen, C.; Polgar, W.; Toll, L. Novel $\alpha 3\beta 4$ Nicotinic Acetylcholine Receptor-Selective Ligands. Discovery, Structure-Activity Studies, and Pharmacological Evaluation. *Journal of Medicinal Chemistry* **2010**, *53*, 8187. [[CrossRef](#)] [[PubMed](#)]
- ⁶⁷ Cippitelli, A.; Brunori, G.; Gaiolini, K. A.; Zaveri, N. T.; Toll, L. Pharmacological Stress Is Required For The Anti-Alcohol Effect Of The $\alpha 3\beta 4^*$ Nachr Partial Agonist AT-1001. *Neuropharmacology* **2015**, *93*, 229. [[CrossRef](#)] [[PubMed](#)]
- ⁶⁸ Glick, S. D.; Maisonneuve, I. M.; Kitchen, B. A.; Fleck, M. W. Antagonism of $\alpha 3\beta 4$ nicotinic receptors as a strategy to reduce opioid and stimulant self-administration. *European Journal of Pharmacology* **2002**, *438*, 99-105. [[CrossRef](#)] [[PubMed](#)]
- ⁶⁹ Mineur, Y. S.; Abizaid, A.; Rao, Y.; Salas, R.; DiLeone, R. J.; Gündisch, D.; Diano, S.; De Biasi, M.; Horvath, T. L.; Gao, X. B.; Picciotto, M. R. Nicotine Decreases Food Intake Through Activation of POMC Neurons. *Science* **2011**, *332*, 1330 [[CrossRef](#)] [[PubMed](#)]
- ⁷⁰ Kombo, D. C.; Hauser, T. A.; Grinevich, V. P.; Melvin, M. S.; Strachan, J. P.; Sidach, S. S.; Chewing, J.; Fedorov, N.; Tallapragada, K.; Breining, S. R.; Mileer, C. H. Pharmacological properties and predicted binding mode of arylmethylenequinuclidine-like derivatives at the $\alpha 3\beta 4$ nicotinic acetylcholine receptor (nAChR). *Bioorganic & Medicinal Chemistry Letters* **2013**, *23*, 1450. [[CrossRef](#)] [[PubMed](#)]
- ⁷¹ Yu, L. F.; Tückmantel, W.; Eaton, J. B.; Caldarone, B.; Fedolak, A.; Hanania, T.; Brunner, D.; Lukas, R. J.; Kozikowski, A. P. Identification of Novel $\alpha 4\beta 2$ -Nicotinic Acetylcholine Receptor (nAChRs) Agonists Based on an Isoxazole Ether Scaffold that Demonstrate Antidepressant-like Activity. *Journal of Medicinal Chemistry* **2012**, *55*, 812. [[CrossRef](#)] [[PubMed](#)]
- ⁷² Toll, L.; Zaveri, N. T.; Polgar, W. E.; Jiang, F.; Khroyan, T. V.; Zhou, W.; Xie, X.; Stauber, G. B.; Costello, M. R.; Leslie, F. M. AT-1001: A High Affinity and Selective $\alpha 3\beta 4$ Nicotinic Acetylcholine Receptor Antagonist Blocks Nicotine Self-Administration in Rats. *Neuropsychopharmacology* **2012**, *37*, 1367. [[CrossRef](#)] [[PubMed](#)]
- ⁷³ Zaveri, N. T.; Bertrand, S.; Yasuda, D.; Bertrand, D. Functional Characterization of AT-1001, an $\alpha 3\beta 4$ Nicotinic Acetylcholine Receptor Ligand. At Human $\alpha 3\beta 4$ and $\alpha 4\beta 2$ nAChR. *Nicotine and Tobacco Research* **2015**, *17*, 361 [[CrossRef](#)] [[PubMed](#)]
- ⁷⁴ Carroll, F. I.; Blough, B. E.; Mascarella, S. W.; Navarro, H. A.; Eaton, J. B.; Lukas, R. J.; Damaj, M. I. Synthesis and Biological Evaluation of Bupropion Analogues as Potential Pharmacotherapies for Smoking Cessation. *Journal of Medicinal Chemistry* **2010**, *53*, 2204. [[CrossRef](#)] [[PubMed](#)]
- ⁷⁵ Franco, A.; Kompella, S. N.; Akondi, K. B.; Melaun, C.; Daly, N. L.; Luetje, C. W.; Alewood, P. F.; Craik, D. J.; Adams, D. J.; Marí, F. RegIIA: An $\alpha 4/7$ -conotoxin from the venom of *Conusregius* that potently blocks $\alpha 3\beta 4$ nAChRs. *Biochemical Pharmacology* **2012**, *83*, 419. [[CrossRef](#)] [[PubMed](#)]

Phthalimide Derivatives with Bioactivity against *Plasmodium falciparum*: Synthesis, Evaluation, and Computational Studies Involving bc_1 Cytochrome Inhibition

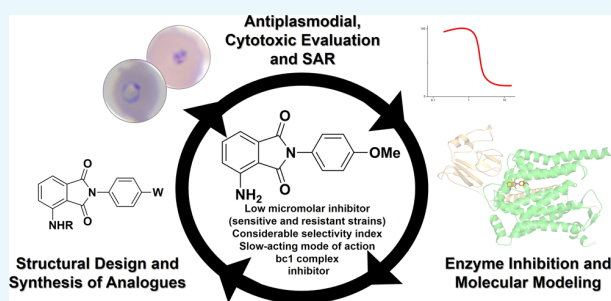
Celso Yassuo Okada-Junior,^{†,||} Gustavo Claro Monteiro,^{†,||} Anna Caroline Campos Aguiar,^{‡,||} Victor Sousa Batista,[§] Juliana Oliveira de Souza,[‡] Guilherme Eduardo Souza,[‡] Renata Vieira Bueno,[‡] Glaucius Oliva,[‡] Nailton M. Nascimento-Júnior,^{*,§} Rafael Victorio Carvalho Guido,^{*,‡,||} and Vanderlan Silva Bolzani^{*,†}

[†]Nuclei of Bioassays, Biosynthesis and Ecophysiology of Natural Products (NuBBE), Department of Organic Chemistry, Institute of Chemistry, and [§]Laboratory of Medicinal Chemistry, Organic Synthesis and Molecular Modeling (LaQMedSOMM), Department of Organic Chemistry, Institute of Chemistry, São Paulo State University—UNESP, Rua Professor Francisco Degni, 55, Jardim Quitandinha, 14800-060 Araraquara, São Paulo, Brazil

[‡]Sao Carlos Institute of Physics, University of Sao Paulo, Av. Joao Dagnone, 1100 Jardim Santa Angelina, Sao Carlos, São Paulo 13563-120, Brazil

Supporting Information

ABSTRACT: We describe herein the design and synthesis of *N*-phenyl phthalimide derivatives with inhibitory activities against *Plasmodium falciparum* (sensitive and resistant strains) in the low micromolar range and noticeable selectivity indices against human cells. The best inhibitor, 4-amino-2-(4-methoxyphenyl)isoindoline-1,3-dione (**10**), showed a slow-acting mechanism similar to that of atovaquone. Enzymatic assay indicated that **10** inhibited *P. falciparum* cytochrome bc_1 complex. Molecular docking studies suggested the binding mode of the best hit to Qo site of the cytochrome bc_1 complex. Our findings suggest that **10** is a promising candidate for hit-to-lead development.



INTRODUCTION

Malaria is a severe parasitic disease endemic in 91 countries. According to the WHO, 216 million cases and approximately 445,000 deaths globally were reported in 2016.¹ The disease can be caused by five species of *Plasmodium* spp (*Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae*, and *Plasmodium knowlesi*). *P. falciparum* is the major pathogen responsible for deaths, and it is associated with severe malaria.² The malaria parasite has developed a resistance to many of the current drugs, including emerging *P. falciparum* resistance to artemisinin derivatives (**1**), which is a component of artemisinin-based combination therapies that comprise the current first-line therapies.³ In this context, new safe and affordable antimalarial drugs are desperately needed. Structure-based approaches have been used to develop new antimalarial candidates with well-defined mechanisms of action.^{4–6} The cytochrome bc_1 complex contains a coenzyme Q, a cytochrome *c* reductase, and the ligand ubiquinone. The enzymatic complex catalyzes electron transfer in the mitochondria of the parasite, which maintains a stable membrane potential for this organelle.^{7,8} Atovaquone (**2**) (Chart 1), a ubiquinone analogue, was approved as a first line

drug for the treatment of malaria in 1992.⁹ Atovaquone has a well-defined mechanism of action; it binds to the oxidation site of the cytochrome bc_1 enzyme complex.^{2,8,10} As a competitive inhibitor of ubiquinone, the inhibitor blocks cellular respiration because of the reduction of electron transfer in the mitochondria of *Plasmodium* spp, which consequently promotes a collapse in the maintenance of the mitochondrial membrane potential.¹⁰ The ubiquinone binding mode indicates that the molecule interacts with a P center of the bc_1 enzymatic complex, inducing a conformational change in the Rieske protein at Fe–S bond, which allows the electron transfer. The atovaquone mode of action consists of the inhibition of the bc_1 complex, resulting in conformational changes that can affect the electronic transfer to the cytochrome *c*.^{9,11}

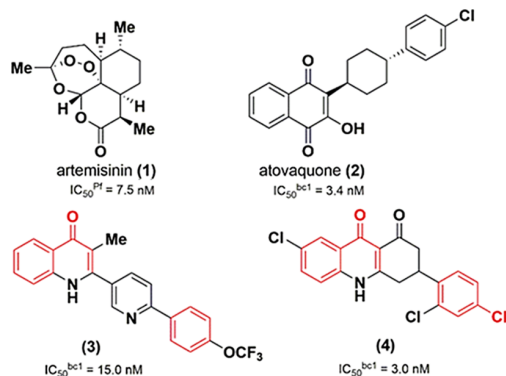
Recent studies have shown that the inhibition of the bc_1 complex also affects the electron acceptor enzyme known as dihydroorotate dehydrogenase (DHODH). This enzyme plays

Received: May 19, 2018

Accepted: August 2, 2018

Published: August 20, 2018

Chart 1. Approved Drugs for the Treatment of Malaria: Artemisinin (1) and Atovaquone (2); 3 (SL-2-25) and 4 (WR249685): Potent Compounds with in Vitro Inhibitory Activity against Whole Parasite (1) or the bc_1 Complex (2–4), and Their Similarities to 2 Are Highlighted in Red

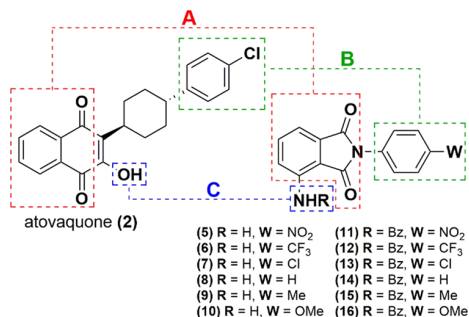


an important role in the biosynthesis of pyridines; therefore, inhibiting this enzymatic complex indirectly impacts the biosynthesis of pyridines in the parasite as well as the transport of electrons in the mitochondria.^{9,12}

Genetic mutations causing atovaquone resistance have been described since the early 1990s.^{13,14} The replacement of Tyr268 with serine (Y268S) or asparagine (Y268N) disrupts the chemical interactions between atovaquone and the active site residues of the bc_1 complex. Therefore, the investigation of new inhibitors is necessary to reduce the potential for drug resistance. Compounds 3 (SL-2-25) and 4 (WR249685) are examples of atovaquone analogues that are potent bc_1 complex inhibitors (Chart 1).^{15,16}

Here, we designed *N*-phenyl-substituted phthalimides (5–16) (Scheme 1) as surrogates for the atovaquone core. The

Scheme 1. Structural Design of the *N*-Phenyl Phthalimide Fragments from Atovaquone



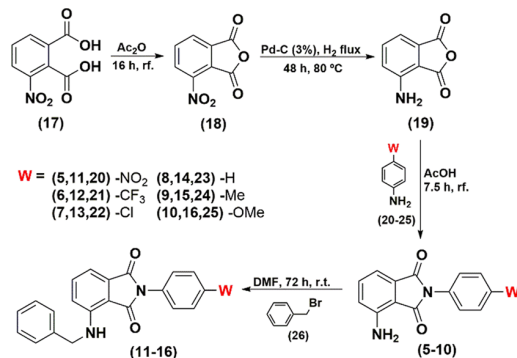
approach consisted of the maintenance of the two aromatic subunits in addition to the 1,4-naphthoquinone moiety (A) including a para-substituted phenyl ring (B) with different electronic and steric properties. The 3-hydroxy group on the 1,4-naphthoquinone moiety was replaced with a monosubstituted amine (C) to maintain the hydrogen-bonding abilities. Atovaquone has shown better oral bioavailability when the 3-hydroxyl substituent was replaced with a hydrophobic group (OCH₃),¹¹ and a benzyl group (C) was introduced to investigate its impact on the biological properties. It is worth mentioning that the designed compounds (5–16) follow the rule of three for fragment-based drug discovery, that is, a

molecular weight lower than 300 Da, up to three hydrogen bond acceptors and a calculated log *P* < 3.¹⁷

RESULTS AND DISCUSSION

In the synthesis of the *N*-phenyl-substituted phthalimide derivatives (5–16) (Scheme 2), 3-nitrophthalic acid (17) was

Scheme 2. Synthetic Route Used to Obtain the *N*-Phenylphthalimides (5–16)



refluxed in acetic anhydride for 16 h.¹⁸ Then, the nitro group of anhydride 18 underwent catalytic hydrogenation with Pd/C (3%) resulting in 19.¹⁹ In a divergent synthetic step, 3-aminophthalic anhydride (19) was refluxed for 7.5 h in glacial acetic acid with the appropriate para-substituted aniline (20–25) to afford fragments 5–10.²⁰ Finally, these aminophthalimides (5–10) were dissolved in *N,N*-dimethylformamide with benzyl bromide (26) to form the *N*-benzylphthalimides (11–16).²¹ See the Supporting Information for more details regarding the experimental procedures and the characterization data of the products and intermediates.

Compounds 5–16 were evaluated in vitro against *P. falciparum* (3D7 strain—chloroquine sensitive) by the SYBR green assay.²² The compounds were tested at single concentration (10 μM), and IC₅₀ values were determined for compounds that showed inhibitions >60%. Compounds 5–9 and 11–15 were inactive as they presented IC₅₀^{Pf} values higher than 10 μM (Table 1).

Conversely, compounds 10 and 16 inhibited in vitro *P. falciparum* growth at low micromolar concentrations (IC₅₀^{Pf} values of 4.2 and 6.8 μM, respectively). These data suggest that a para-methoxy substituent is beneficial for the inhibitory activity in this series. Moreover, the bulky benzyl substituent at site C was tolerated, suggesting that structural variation at this position can be further explored. Next, the cytotoxic effects of the active compounds were evaluated against a human hepatoma cell line (HepG2) by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide viability assay.²³ Compounds 10 and 16 showed low cytotoxicities (IC₅₀^{HepG2} > 250 μM) and considerable selectivity indices (SI values of >36 and >59, respectively) (Table 1). It is worth mentioning that both compounds were soluble at the highest concentration tested (250 μM).

To better characterize the antiparasitic activities and confirm the low cytotoxicities of the series, the most potent phthalimide derivative (10) was selected. The morphology of HepG2 cells was evaluated after 24 h of treatment with 250 μM of 10. At this concentration, we verified that compound 10

Table 1. Inhibitory and Cytotoxic Activities of the Phthalimide Derivatives against *P. falciparum* (3D7 Strain) and HepG2 Cells^a

Cpd	% of inhibition @ 10 μM	IC ₅₀ ^{Pf} (CI95%) (μM)	IC ₅₀ ^{HepG2} (μM)	SI*
5	15 \pm 2	>10	nd	nd
6	3 \pm 2	>10	nd	nd
7	5 \pm 3	>10	nd	nd
8	6 \pm 3	>10	nd	nd
9	4 \pm 2	>10	nd	nd
10	75 \pm 4	4.2 (3.2–5.2)	>250	>59
11	6 \pm 2	>10	nd	nd
12	5 \pm 2	>10	nd	nd
13	7 \pm 3	>10	nd	nd
14	23 \pm 3	>10	nd	nd
15	8 \pm 3	>10	nd	nd
16	65 \pm 5	6.8 (5.6–8.0)	>250	>36
artesunate		0.006 (0.005–0.008)	305 (281–329)	50.833

^aThe presented IC₅₀ values are the average of two independent experiments. *SI = IC₅₀^{HepG2}/IC₅₀^{Pf}; CI95% = 95% confidence interval; nd = not determined.

was not toxic to the cells as they showed a microscopic morphology similar to that of untreated cells (Figure 1A).

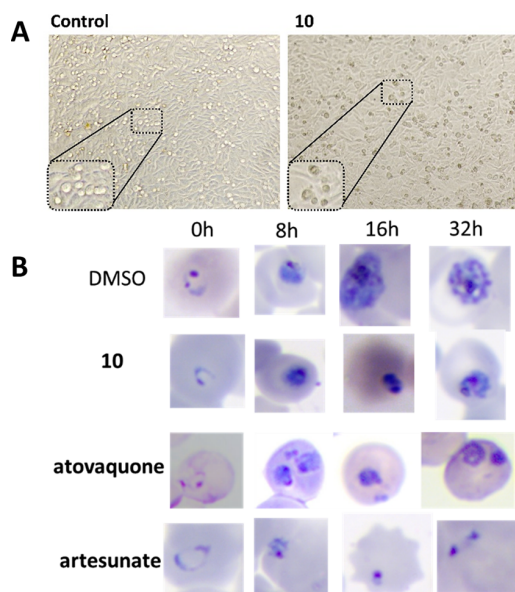
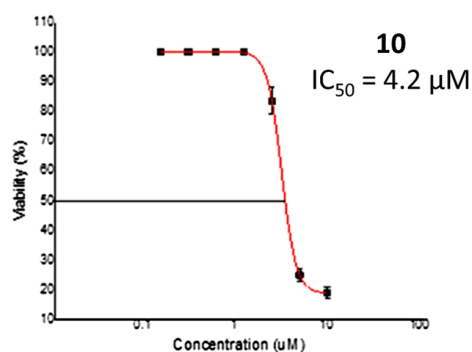


Figure 1. (A) Morphology of HepG2 cells before and after incubation with **10** at 250 μM . Insets: Expansion showing representative cell morphology; (B) microscopy of synchronized parasites continuously treated with **10** at 40 μM and the controls [dimethyl sulfoxide (DMSO), **10**, atovaquone, and artesunate]. Representative images of two independent experiments.

Next, the inhibitory activity of **10** was evaluated against the *P. falciparum* K1 strain, a strain that is multidrug-resistant to the antimalarial drugs chloroquine, pyrimethamine, and sulfadoxine. Compound **10** was active in the low micromolar range against the resistant parasite with an IC₅₀ value of 4.3 μM , which is approximately the same inhibitory activity observed against the sensitive 3D7 strain (Figure 2). To elucidate the stage-specific inhibitory activity of **10**, we incubated the compound with highly synchronized parasites at a concentration 10-fold greater than its IC₅₀ value and observed the morphological changes over time. As seen in Figure 1B, compound **10** showed activity in the later ring

3D7- sensitive parasite



K1- resistant parasite

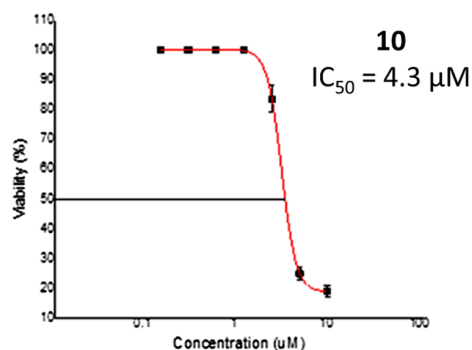


Figure 2. Representative concentration–response curves of **10** against sensitive (3D7 strain) and resistant parasites (K1 strain). The reported IC₅₀ values are the mean values of two independent experiments.

stages (late rings or early trophozoites) and it induced alterations in *P. falciparum* morphology between 16 and 32 h after incubation. These data suggest a stage-specific effect that may be related to a slow-acting mechanism in which the latter forms of *P. falciparum* in the intraerythrocytic cycle are susceptible to the effects of the compound. These results are in

good agreement with the stage-specific inhibitory activity of atovaquone, which is a slow-acting inhibitor of *P. falciparum*.²⁴

Aiming to elucidate the effects of compound **10** on *P. falciparum* mitochondrial respiration, an enzymatic assay was conducted to measure bc1 complex decylubiquinol-cytochrome *c* oxidoreductase activity.²⁵ Mitochondrial fractions were extracted from *P. falciparum* 3D7 parasites, and the assay was performed using a compound **10** concentration of 70 μM , in parallel with negative (DMSO) and positive (atovaquone at 6 μM) controls. Our results showed that **10** inhibited cytochrome bc1 enzymatic activity (74% at 70 μM) and demonstrated that cytochrome bc1 complex is a molecular target of **10** antiplasmodial activity (Figure 3).

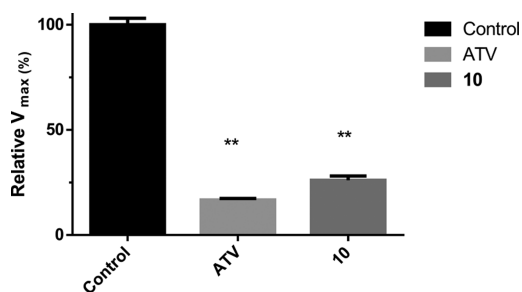


Figure 3. Inhibition of bc1 complex activity after incubation with atovaquone (ATV) at 6 μM (84 ± 1% of inhibition) or compound **10** at 70 μM (74 ± 2% of inhibition). DMSO was used as a negative control (1 ± 3% of inhibition). *** p < 0.005 by Mann–Whitney test.

The mode of atovaquone binding to the cytochrome bc1 complex indicated that the carbonyl group alpha to the hydroxyl group is important for ligand binding and recognition. In addition, the hydroxyl group is involved in a salt bridge with the nitrogen atom of the imidazole moiety of histidine in the Rieske protein.^{10,11} Finally, hydrophobic interactions and van der Waals forces are important for stabilizing atovaquone binding. For example, the cyclohexyl substituent and the 4-chlorophenyl substituent interact with the side chains of the V146, I269, P271, M139, L275, M295, and F296 residues from the *Saccharomyces cerevisiae* cytochrome bc1 complex.²⁶ Accordingly, on the basis of the structural similarities and in vitro cellular findings, which indicated that phthalimide derivative **10** shares a slow-acting inhibitory mechanism with atovaquone, we modeled the binding mode of **10** to the cytochrome bc1 complex.

The crystal structure of the *S. cerevisiae* cytochrome bc1 complex co-crystallized with stigmatellin was solved at 1.9 Å resolution (PDB ID: 3CX5)²⁷ and has been used as a molecular target for modeling studies. The observed binding mode of atovaquone co-crystallized with a cytochrome bc1 complex (PDB ID: 4PD4, resolution: 3.04 Å)²⁶ was used as a reference. Superposition of the two crystal structures showed no significant differences in the orientation of the amino acid residues within the protein binding sites, suggesting that the higher resolution structure is adequate for docking studies. The cytochrome bc1 complex from *S. cerevisiae* was prepared for rigid docking using GOLD software (v5.4).²⁸ The hydrogen atoms were added to the structure, and the water molecules and ligands were removed. The 3D structure of the hydroxyl group of atovaquone was built in the ionized state because previous studies showed that this microspecies is responsible for the activity.²⁶

To evaluate the accuracy of the method, including the ChemPLP scoring function,²⁹ we redocked atovaquone into the cytochrome bc1 complex binding site. The best-ranked binding mode was compared to the binding mode of atovaquone in the crystal structure, and both the orientation and interactions with key residues were reproduced (RMSD = 0.4 Å; PLPscore = 63.3), which confirms the accuracy of this strategy for the system under investigation (see the Supporting Information for details). On the basis of that, we used the validated protocol to predict the binding mode of **10** within the cytochrome bc1 complex binding site. According to the model, the phthalimide moiety of **10** establishes hydrogen-bonding interactions with the side chains of E272 and H181. Additionally, this moiety is in close contact with the side chains of I147, V146, and P271 and with the backbone atoms of W142. The methoxyphenyl substituent forms hydrophobic interactions with the side-chain atoms of M295 and L275. The ChemPLP score for this pose is 63.6, which is close to the calculated score for the atovaquone binding mode (Figure 4).

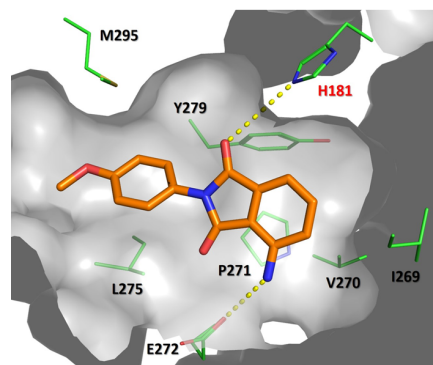


Figure 4. Docking pose of **10** (orange stick) in the binding site of the *S. cerevisiae* cytochrome bc1 complex (PDB ID 3CX5). Main interacting residues are indicated as green sticks. Hydrogen bonds are represented as yellow dotted lines. Residues I269, V270, P271, E272, L275, Y279, and M295 (black label) are located at the cytochrome *b* subunit, whereas H181 (red label) is located at the Rieske protein subunit. Residues W142, V146, I147, and all hydrogens atoms were omitted for clarity. Figure generated using PyMol (v 0.99).³⁰

As highlighted in the sequence alignment (Figure S52),^{31,32} all of the mentioned residues interacting with the phthalimide moiety are conserved in the *P. falciparum* cytochrome bc1 complex (Uniprot Q02768/Q8IL75), apart from M295 and L275. Residues M295 and L275 from bc1 of *S. cerevisiae* are replaced with V284 and F264, respectively. In *P. falciparum*, both substitutions can establish hydrophobic interactions with the methoxyphenyl substituent of **10**. Moreover, the alignment of *S. cerevisiae*, *P. falciparum*, and *Mycobacterium tuberculosis* bc1 sequences shows that the binding site of **10** is conserved within these species. The residues H181, I269, and P271 from bc1 complex of *S. cerevisiae* are conserved in *P. falciparum* and *M. tuberculosis*, whereas V146 and I147, conserved in *P. falciparum*, are conservative replaced by alanine and leucine, respectively, in *M. tuberculosis*. Most of nonconserved residues within the three species stabilizes compound **10** by hydrophobic and van der Waals interactions. Because a potent antimycobacterial activity is observed for phthalimide-containing compounds whose target is the bc1 complex of *M.*

tuberculosis,³³ the binding site identity is an indicative that the bc_1 complex might be the molecular target of compound **10**.

CONCLUSIONS

Phthalimide-based derivatives are attractive molecules for drug discovery and development. Because of the potential biological activity and pharmaceutical use, the privileged phthalimide scaffold was designed to be used in many therapeutic treatments, including antituberculosis as cytochrome bc_1 inhibitor,³³ anti-inflammatory compound developed as a phosphodiesterase-4 inhibitor³⁴ and for antimalarial as plasmepsin II & IV inhibitor.³⁵ In this work, the best phthalimide derivative, compound **10**, exhibited antimalarial activity against sensitive and resistant *P. falciparum* strains in the low micromolar range ($IC_{50} < 10 \mu M$), no cytotoxicity against HepG2 cells, a considerable selectivity index ($SI > 30$), and a slow-acting mechanism, which is consistent with the inhibitory stage of action of atovaquone. The modeled binding mode of **10** suggested the molecular determinants that might be related to the inhibitory activity of this series. Therefore, our findings indicate that 4-amino-2-(4-methoxyphenyl)-isoindoline-1,3-dione (**10**) is a new hit for the development of lead compounds with superior properties.

COMPUTATIONAL METHODS

Initially, all molecular structures of the *p*-phenyl-functionalized phthalimidic compounds were generated with Discovery Studio Visualizer free software (DSV/version: 17.2.0)³⁶ and were optimized using the semiempirical method PM6 in MOPAC free software (version 2016)³⁷ through the Mercury CSD software interface (v3.9), which is included in the CCDC GOLD Suite computational package. The commands used during optimization were: MMOK, XYZ, CHARGE = -1 [for docking of atovaquone (**2**)] and GNORM = 1. The format of the optimized compound file was *.mol2. Atovaquone, the reference ligand for this work, was modeled in the ionized state.²⁶ In parallel, the 3D file (*.pdb) of the *S. cerevisiae* cytochrome bc_1 crystal structure in complex with stigmatellin was obtained from Protein Data Bank (PDB ID: 3CX5) at 1.9 Å resolution.²⁷ The structure was chosen because atovaquone is a high affinity inhibitor of *S. cerevisiae* complex III ($K_i = 5$ nM), a nonpathogenic surrogate model for studying biological activities of compounds that target the parasite through this mechanism. Although the crystal structure contains a ligand other than atovaquone, we chose it because the cytochrome bc_1 crystal structure in complex with atovaquone (PDB ID: 4PD4)²⁶ lacks the Rieske protein and has lower resolution in comparison with the 3CX5 structure. Superposition of 3CX5 and 4PD4 indicates no significant differences in the position of the atoms in amino acid residues of the binding site, thereby suggesting that the structure is suitable for the docking calculations using atovaquone's binding mode in 4PD4 as a reference for evaluation of the model. The 3CX5.pdb file was then prepared for docking using GOLD software (v5.4),²⁸ where the hydrogen atoms were added to the structure. All water molecules and ligands were removed. The rigid docking was performed in GOLD using coordinates $x = 13.0858$, $y = 5.3280$, and $z = 17.4635$ as the center of a 10 Å search radius. GOLD was set to consider possible ligand internal hydrogen bonds in the calculations and score 50 binding modes using the ChemPLP scoring function. The selection of this scoring function was based on a previous article by Li et al. that

indicates best accuracy for ChemPLP in comparison with several other score functions.²⁹ The best GOLD docking results for each ligand were analyzed using Discovery Studio Visualizer free software to check the ligand–cytochrome bc_1 complex interaction types. The figures were generated in Pymol (v0.99).³⁰

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.8b01062.

Syntheses of the phthalimides derivatives **5–16**, analytical data, spectra of all synthesized compounds, description of the biological methods for evaluation of antimalarial activity, and computational modeling methodology of the compounds **10** and **16** in bc_1 cytochrome complex (PDF)

Accession Codes

Alphanumeric codes such as Y268S indicate mutations between amino acid residues, and codes such as H181 indicate the amino acid residue and the position in the protein.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: nailtonjr@iq.unesp.br (N.M.N.-J.).

*E-mail: rvcguido@usp.br (R.V.C.G.).

*E-mail: bolzaniv@iq.unesp.br (V.S.B.).

ORCID

Rafael Victorio Carvalho Guido: 0000-0002-7187-0818

Author Contributions

^{ll}C.Y.O.-J., G.C.M., and A.C.C.A. contributed equally to the work. C.Y.O.-J. and G.C.M. synthesized, isolated, purified, and characterized the compounds. A.C.C.A., G.E.S., and J.O.d.S., performed the in vitro studies. V.S.B., N.M.N.-J., and R.V.B. performed the molecular modeling studies. G.O., N.M.N.-J., R.V.C.G., and V.S.B. conceived the study, analyzed the data, contributed ideas, and wrote the paper.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was funded by (1) CNPq INCT-INOVAR Grant: 573.564/2008-6 and FAPERJ E-26/170.020/2008; (2) PROPe—UNESP. First Projects. Grant: 840; (3) INCT-BioNat Grant: 2014/50926-0 (FAPESP) and 465637/2014-0 (CNPq); (4) CEPID-CIBFar. Grant: 2013/07600-3 and 2014/50249-8; (5) CNPq Universal Grant: 405330/2016-2; and (6) Serrapilheira Institute Grant: Serra-1708-16250. We acknowledge the Institute of Chemistry of the São Paulo State University (UNESP) and the Physics Institute of São Carlos, University of São Paulo (USP) for the infrastructure that allowed the development of this manuscript.

ABBREVIATIONS

HepG2, liver hepatocellular carcinoma; WHO, World Health Organization; log *P*, log (partition coefficient); Pd/C, palladium on carbon; DMF, *N,N*-dimethylformamide; SYBR, *N,N'*-dimethyl-*N*-[4-[(*E*)-(3-methyl-1,3-benzothiazol-2-yl)dene)methyl]-1-phenylquinolin-1-ium-2-yl]-*N*-propylpropane-1,3-diamine; MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-

tetrazolium bromide; IC₅₀, half maximal inhibitory concentration; LD₅₀, median lethal dose; SI, selectivity index; PM6, parametric method 6; RMSD, root-mean-square deviation

REFERENCES

- (1) WHO. World Malaria Report. <http://www.who.int/malaria/publications/world-malaria-report2017/report/en/> (accessed 2017).
- (2) Biamonte, M. A.; Wanner, J.; Le Roch, K. G. Recent Advances in Malaria Drug Discovery. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 2829–2843.
- (3) Arie, F.; Witkowski, B.; Amaratunga, C.; Beghain, J.; Langlois, A.-C.; Khim, N.; Kim, S.; Duru, V.; Bouchier, C.; Ma, L.; Lim, P.; Leang, R.; Duong, S.; Sreng, S.; Suon, S.; Chuor, C. M.; Bout, D. M.; Ménard, S.; Rogers, W. O.; Genton, B.; Fandeur, T.; Miotto, O.; Ringwald, P.; Le Bras, J.; Berry, A.; Barale, J.-C.; Fairhurst, R. M.; Benoit-Vical, F.; Mercereau-Puijalon, O.; Ménard, D. A Molecular Marker of Artemisinin-Resistant *Plasmodium falciparum* Malaria. *Nature* **2014**, *505*, 50–55.
- (4) Aguiar, A. C.; de Sousa, L. R. F.; Garcia, C. R. S.; Oliva, G.; Guido, R. V. C. New Molecular Targets and Strategies for Antimalarial Discovery. *Curr. Med. Chem.* **2017**, *24*, DOI: 10.2174/0929867324666170830103003
- (5) Mishra, M.; Mishra, V. K.; Kashaw, V.; Iyer, A. K.; Kashaw, S. K. Comprehensive Review on Various Strategies for Antimalarial Drug Discovery. *Eur. J. Med. Chem.* **2017**, *125*, 1300–1320.
- (6) Drinkwater, N.; McGowan, S. From Crystal to Compound: Structure-Based Antimalarial Drug Discovery. *Biochem. J.* **2014**, *461*, 349–369.
- (7) El Hage, S.; Ane, M.; Stigliani, J.-L.; Marjorie, M.; Vial, H.; Baziard-Mouysset, G.; Payard, M. Synthesis and Antimalarial Activity of New Atovaquone Derivatives. *Eur. J. Med. Chem.* **2009**, *44*, 4778–4782.
- (8) Hughes, L. M.; Covian, R.; Gribble, G. W.; Trumpower, B. L. Probing binding determinants in center P of the cytochrome bc1 complex using novel hydroxy-naphthoquinones. *Biochim. Biophys. Acta, Bioenerg.* **2010**, *1797*, 38–43.
- (9) Fry, M.; Pudney, M. Site of action of the antimalarial hydroxynaphthoquinone, 2-[trans-4-(4'-chlorophenyl) cyclohexyl]-3-hydroxy-1,4-naphthoquinone (S66C80). *Biochem. Pharmacol.* **1992**, *43*, 1545–1553.
- (10) Barton, V.; Fisher, N.; Biagini, G. A.; Ward, S. A.; O'Neill, P. M. Inhibiting *Plasmodium* cytochrome bc1: a complex issue. *Curr. Opin. Chem. Biol.* **2010**, *14*, 440–446.
- (11) Zhou, J.; Duan, L.; Chen, H.; Ren, X.; Zhang, Z.; Zhou, F.; Liu, J.; Pei, D.; Ding, K. Atovaquone Derivatives as Potent Cytotoxic and Apoptosis Inducing Agents. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5091–5094.
- (12) Painter, H. J.; Morrissey, J. M.; Mather, M. W.; Vaidya, A. B. Specific Role of Mitochondrial Electron Transport in Blood-Stage *Plasmodium falciparum*. *Nature* **2007**, *446*, 88–91.
- (13) Srivastava, I. K.; Morrissey, J. M.; Darrouzet, E.; Daldal, F.; Vaidya, A. B. Resistance Mutations Reveal the Atovaquone-Binding Domain of Cytochrome *b* in Malaria Parasites. *Mol. Microbiol.* **1999**, *33*, 704–711.
- (14) Musset, L.; Bouchaud, O.; Matheron, S.; Massias, L.; Le Bras, J. Clinical Atovaquone-Proguanil Resistance of *Plasmodium falciparum* Associated with Cytochrome *b* Codon 268 Mutations. *Microbes Infect.* **2006**, *8*, 2599–2604.
- (15) Pidathala, C.; Amewu, R.; Pacorel, B.; Nixon, G. L.; Gibbons, P.; Hong, W. D.; Leung, S. C.; Berry, N. G.; Sharma, R.; Stocks, P. A.; Srivastava, A.; Shone, A. E.; Charoensutthivarakul, S.; Taylor, L.; Berger, O.; Mbekeani, A.; Hill, A.; Fisher, N. E.; Warman, A. J.; Biagini, G. A.; Ward, S. A.; O'Neill, P. M. Identification, Design and Biological Evaluation of Bisaryl Quinolones Targeting *Plasmodium falciparum* Type II NADH:Quinone Oxidoreductase (PfNDH2). *J. Med. Chem.* **2012**, *55*, 1831–1843.
- (16) Biagini, G. A.; Fisher, N.; Berry, N.; Stocks, P. A.; Meunier, B.; Williams, D. P.; Bonar-Law, R.; Bray, P. G.; Owen, A.; O'Neill, P. M.; Ward, S. A. Acridinediones: Selective and Potent Inhibitors of the Malaria Parasite Mitochondrial bc1 Complex. *Mol. Pharm.* **2008**, *73*, 1347–1355.
- (17) Scott, D. E.; Coyne, A. G.; Hudson, S. A.; Abell, C. Fragment-Based Approaches in Drug Discovery and Chemical Biology. *Biochemistry* **2012**, *51*, 4990–5003.
- (18) Nandhikonda, P.; Heagy, M. D. Dual Fluorescent N-Aryl-2,3-naphthalimides: Applications in Ratiometric DNA Detection and White Organic Light-Emitting Devices. *Org. Lett.* **2010**, *12*, 4796–4799.
- (19) Hudlicky, M. *Reductions in Organic Chemistry*; Ellis Horwood Limited: New York, 1984; pp 322.
- (20) Nascimento-Júnior, N. M.; Mendes, T. C. F.; Leal, D. M.; Corrêa, C. M. N.; Sudo, R. T.; Zapata-Sudo, G.; Barreiro, E. J.; Fraga, C. A. M. Microwave-Assisted Synthesis and Structure-Activity Relationships of Neuroactive Pyrazolo[3,4-b]pyrrolo[3,4-d]pyridine Derivatives. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 74–77.
- (21) Martínez-Montero, L.; Díaz-Rodríguez, A.; Gotor, V.; Gotor-Fernández, V.; Lavandera, I. Broadening the Chemical Scope of Laccases: Selective Deprotection of N-benzyl groups. *Green Chem.* **2015**, *17*, 2794–2798.
- (22) Smilkstein, M.; Sriwilaijaroen, N.; Kelly, J. X.; Wilairat, P.; Riscoe, M. Simple and Inexpensive Fluorescence-Based Technique for High-Throughput Antimalarial Drug Screening. *Antimicrob. Agents Chemother.* **2004**, *48*, 1803–1806.
- (23) Denizot, F.; Lang, R. Rapid colorimetric assay for cell growth and survival. *J. Immunol. Methods* **1986**, *89*, 271–277.
- (24) White, N. J. Assessment of the Pharmacodynamic Properties of Antimalarial Drugs In Vivo. *Antimicrob. Agents Chemother.* **1997**, *41*, 1413–1422.
- (25) Fisher, N.; Castleden, C. K.; Bourges, I.; Brasseur, G.; Dujardin, G.; Meunier, B. Human Disease-related Mutations in Cytochrome Studied in Yeast. *J. Biol. Chem.* **2004**, *279*, 12951–12958.
- (26) Birth, D.; Kao, W.-C.; Hunte, C. Structural Analysis of Atovaquone-Inhibited Cytochrome bc1 Complex Reveals the Molecular Basis of Antimalarial Drug Action. *Nat. Commun.* **2014**, *5*, 4029.
- (27) Solmaz, S. R. N.; Hunte, C. Structure of Complex III with Bound Cytochrome *c*1 Reduced State and Definition of a Minimal Core Interface for Electron Transfer. *J. Biol. Chem.* **2008**, *283*, 17542–17549.
- (28) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking 1 | Edited by F. E. Cohen. *J. Mol. Biol.* **1997**, *267*, 727–748.
- (29) Li, Y.; Han, L.; Liu, Z.; Wang, R. Comparative Assessment of Scoring Functions on an Updated Benchmark: 2. Evaluation Methods and General Results. *J. Chem. Inf. Model.* **2014**, *54*, 1717–1736.
- (30) *The PyMOL Molecular Graphics System*, Version 0.99; Schrödinger, LLC, 2002.
- (31) Notredame, C.; Higgins, D. G.; Heringa, J. T-coffee: a novel method for fast and accurate multiple sequence alignment 1 | Edited by J. Thornton. *J. Mol. Biol.* **2000**, *302*, 205–217.
- (32) Waterhouse, A. M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191.
- (33) van der Westhuyzen, R.; Winks, S.; Wilson, C. R.; Boyle, G. A.; Gessner, R. K.; Soares de Melo, C.; Taylor, D.; de Kock, C.; Njoroge, M.; Brunschwig, C.; Lawrence, N.; Rao, S. P. S.; Sirgel, F.; van Helden, P.; Seldon, R.; Moosa, A.; Warner, D. F.; Arista, L.; Manjunatha, U. H.; Smith, P. W.; Street, L. J.; Chibale, K. Pyrrolo[3,4-c]pyridine-1,3(2H)-diones: a novel antimycobacterial class targeting mycobacterial respiration. *J. Med. Chem.* **2015**, *58*, 9371–9381.
- (34) de Castro Barbosa, M. L.; Ramos, T. J. F.; de Arantes, A. C. S.; Martins, M. A.; Rodrigues e Silva, P. M.; Barreiro, E. J.; Lima, L. M. Synthesis and Pharmacological Evaluation of Novel Phenyl Sulfonamide Derivatives Designed as Modulators of Pulmonary Inflammatory Response. *Molecules* **2012**, *17*, 14651–14672.

(35) Singh, A. K.; Rathore, S.; Tang, Y.; Goldfarb, N. E.; Dunn, B. M.; Rajendran, V.; Ghosh, P. C.; Singh, N.; Latha, N.; Singh, B. K.; Rawat, M.; Rathi, B. Hydroxyethylamine Based Phthalimides as New Class of Plasmeprin Hits: Design, Synthesis and Antimalarial Evaluation. *PLoS One* **2015**, *10*, e0139347.

(36) Dassault Systèmes BIOVIA. *Discovery Studio Modeling Environment*, San Diego, Release 2017.

(37) Stewart, J. J. P. *MOPAC*, Stewart Computational Chemistry, Colorado Springs, CO, USA, [HTTP://OpenMOPAC.net](http://OpenMOPAC.net), 2016.

Molecular Docking: Considerations of a Low Cost and Suitable Methodology and Some Successful Applications

Batista VS and Nascimento-Júnior NM*

Department of Organic Chemistry, São Paulo State University, Brazil

***Corresponding author:** Nascimento-Júnior NM, Laboratory of Medicinal Chemistry, Organic Synthesis and Molecular Modeling (LaQMedSOMM), Department of Organic Chemistry, Institute of Chemistry, São Paulo State University – UNESP, Rua Professor Francisco Degni, 55, Jardim Quitandinha, 14800-060, Araraquara, São Paulo, Brazil, Tel: +55 21 3301-9897; Email: nailton.monteiro@unesp.br

Opinion

Volume 2 Issue 3

Received Date: September 25, 2018

Published Date: October 04, 2018

Abstract

Molecular docking can be a powerful tool, given the growing amount of available solved protein crystallographic structures and always improving computational tools but it can also develop in a more complex issue than one would initially imagine. The reasoning behind this statement comes from the simple fact that even if a good amount of data is readily available, care should be taken to ensure that the starting point of any docking job suits the purpose of the study at hand. There are several docking methodologies which can be useful in one or another work flow and to many researchers outside the field this can be elusive as there are several parameters that must be addressed beforehand. In the present study, the authors give an overview about our molecular docking protocol and analysis of the results, highlighting key aspects that must be addressed in each step of the process and setting the reader's mind to focus on what kind of information have to be considered before starting a docking job. Our focus is towards freeware software and web servers, resulting in a virtually free methodology that can be easily applied by newcomers to molecular docking. Also, examples of successful docking applications are given.

Keywords: Molecular Docking; Virtual Screening; Protein Obtention; Homology Modeling; Protein Energy Minimization

Abbreviations: SBDD: Structure-based drug design, PDB: Protein Data Bank; DSV: Discovery Studio Visualizer; GA: genetic algorithm; nAChRs: nicotinic acetylcholine receptors; NMR: Nuclear Magnetic Resonance; NPD: Natural Product Database; FDA: Food and Drug

Administration HTS: High Throughput Screening; SPR: Surface Plasmon Resonance; LE: Ligand Efficiency; LLE: Lipophilic Ligand Efficiency; SAR: Structure Activity Relationship.

Introduction

Structure-based drug design (SBDD) is receiving increasing attention, both in industry and academia, as structural genomics, computational tools and spectroscopic techniques become more sophisticated. This is a result of the increasing number of protein tridimensional structures elucidated and publicly available [1]. In this context, molecular docking is a wide spread method useful in applications such as virtual screening, lead optimization and to provide understanding of the intricate aspects of intermolecular recognition. This strategy focus on the prediction of the structure of protein-ligand and/or protein-protein complexes through two key steps: conformational sampling of the system and scoring the resulting complexes by a scoring function [2]. There are many different strategies to address both sampling and scoring in molecular docking, as well as several commercial and academic software packages to perform the calculations [3]. A side from the docking packages themselves, there are also other essential steps involved in the development of a docking methodology that relate to the protein structures used, ligand databases and docking parameters [4,5]. Such complexity has to be properly addressed in a successful docking job by careful consideration of the chosen methodology and its limitations.

The first and most relevant concern when envisaging a molecular docking study is, ultimately, why you are using it. There are a number of computational tools just as good (or more suitable) as docking for a given problem and none of them should be disregarded in the planning of a computational methodology. The molecular modeler must be aware of the strengths and weaknesses of a given

methodology, as well as the applicability of such approach into the setting of its own problem, and the ideal scenario would be the one where the researcher has all this in mind beforehand.

In the present manuscript we propose and discuss a docking methodology, its basis and key steps, focusing on software packages and web servers freely available to the academic community or subsidized by some funding agencies (e.g. CAPES, BR). The proposed methodology can be successfully applied to solve docking problems, as it will be show further (applications section). As well as some other examples of successful docking applications.

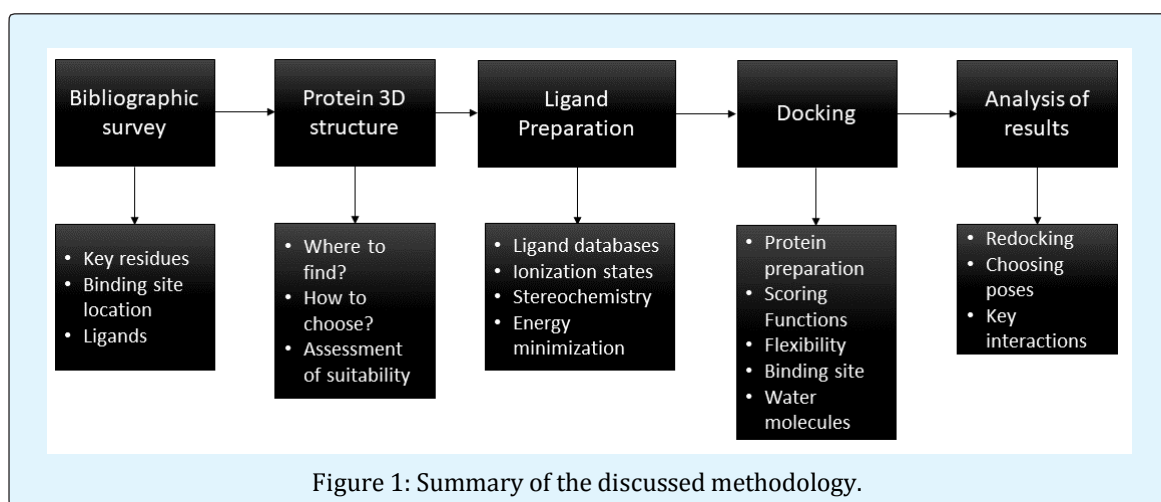
Software Packages and Web Servers

The following software packages and web servers recommended are shown below:

- Database for protein obtention: Protein Data Bank (PDB) [6];
- Database for ligand obtention: ChEMBL [7];
- Ramachandran plot: RAMPAGE [8];
- Homology modeling: Swiss Model [9];
- 3D structure visualizer and/or builder: Discovery Studio Visualizer (DSV) [10] or PyMol [11];
- Ligands ionization state prediction: Marvin Sketch [12];
- Ligands/protein energy minimization: MOPAC2016 [13];
- Docking package: CCDC GOLD Suite [14-17].

Discussed Methodology

The methodology described in this paper is summarized in Figure 1.



The starting point of any docking study should be an extensive research of the existing scientific literature regarding the desired molecular target. Any information about the properties of the system in study is of great importance. The researcher should be concerned with retrieving information like key residues for binding of ligands, where is the binding site located, what class of molecules binds to this specific binding site, presence or absence of water molecules into this site, what is the effect of the binding site environment on the ligand's ionizable moieties, existence of induced-fit effects and so on. A good amount of the mentioned information can be obtained just by analyzing the crystal structure of the desired target but care should be taken to avoid overlooking information that is not straightforward obtainable from this analysis. The presence of water molecules, for instance, can only be evaluated if the crystallographic resolution of the desired structure is good enough for this purpose.

The docking package used in this work is GOLD, available in the CCDC GOLD Suite, which have been extensively tested and has shown excellent performance for pose prediction and virtual screening. Its four native scoring functions, GoldScore [14], CHEMPLP [15], ChemScore [16], and ASP [17] were ranked 1th, 2th, 5th and 6th, respectively, in a study by Li, et al. [18] comparing several scoring functions from popular docking software like MOE [19] and GLIDE [20]. In addition, the CCDC GOLD Suite is subsidized by CAPES Brazilian funding agency to academic research institutions, which alongside with the use of freeware software packages and web servers, making the overall software cost for using this methodology virtually zero. GOLD uses a stochastic approach for conformational sampling, implemented through genetic algorithm (GA), alongside with the formerly mentioned scoring functions to score the ligand's predicted binding modes and has several available options for tuning the calculation as needed.

Obtention of the Targeted 3D Structure

Initially, a suitable tridimensional structure must be obtained. This can be done by either retrieving the data from a database, usually the PDB, or by computational modelling approaches, like homology modelling, and this choice depends on what information is already available for the chosen target. Homology modelling demands more effort and has more room for inaccuracies but this approach has been successfully applied in many cases

[21]. On the other hand, by using an already elucidated crystal structure, the advantage of experimental evidence supports the observed conformation of the system. Nevertheless, the choice of the starting structure is of major importance because changes in the position of the residues comprising the binding site can fool docking engines [22]. The evaluation of the suitability of a crystal structure is done by analyzing its Ramachandran plot, which can be done free of charge using RAMPAGE web server. It is important that the binding site residues are in favored or allowed regions of the graph, provided that the reason why a given residue is inside the allowed region instead of the favorable region is explainable, like to accommodate the presence of a ligand. The structure's resolution should be as good as possible. Also, crystallographic parameters contained in the PDB file header, like R-factor, should not be overlooked. A review by Wlodawer, et al. [23] that covers the basics of protein crystallography for those outside the field is recommended to better understand the implications of such parameters on the quality of the crystallographic data.

It is possible that the researcher has to choose from a variety of available crystal structures for the desired target either to start the docking or to build a homology model. In such cases, the most important thing to keep in mind is what your problem requires. For instance, when faced with several crystal structures for building a homology model the recommendation is that the template and target are of the same family of proteins and with at least 30% sequence identity between them but if the relevant residues in the homology model are properly positioned it is possible to find a situation where identities smaller than 30% are acceptable [24]. Additionally, it is possible to build homology models using pieces of different crystallized proteins to complete the gaps from each other. A homology model can be built online using the SwissModel web server. As for selecting among relevant crystal structures to a straightforward docking approach, the ideal scenario would be the one where you have a co-crystallized ligand to use as a reference. For example, when working with neuronal nicotinic acetylcholine receptors (nAChRs) it is advisable to start from a structure containing nicotine, a classic agonist, because the observed interactions can guide what are the relevant motifs for molecular recognition. Selecting a structure with a known co-crystallized ligand is also advantageous because it is easier to identify the region which should be treated as the binding site.

Setting up Ligands for Docking

With the appropriated crystal structure at hand, the next step is setting up ligands to dock into the desired binding site. Their tridimensional structures should be drawn using a software such as DSV or PyMol, the first being available free of charge. It is possible to obtain published ligands online on the ChEMBL database but it is important to always check the original paper for biological activity and absolute stereochemistry data. In docking approaches the absolute stereochemistry is mandatory for a good level of confidence in the obtained results. Thus, ligands acting in racemic form or with absolute stereochemistry poorly defined by the original authors should be avoided. If unavoidable, it is possible to set the docking software to score all possible stereoisomeric forms of any molecule by building and docking them into the binding site and then sorting the best option based on the calculated score for each one.

It is important that, when building a database of ligands, one takes into account what the problem to be solved requires. First of all the researcher has to be sure that the selected ligands all bind into the same binding site in order to compare the differences in their binding modes. The number of ligands needed depends on the desired application. Docking-based virtual screening, for instance, requires a good amount of molecules while the understanding of the binding motifs for a given molecule might only need one or two. When the interest resides in the protocol's ability to differentiate between active and inactive compounds, it is best to use a focused database, meaning that the actives and inactives ought to have some degree of similarity. If the researcher is interested in using docking in a 3D-QSAR study; the bioassays for the selected ligands must all have been performed through the same *in vitro* protocol in order for the bioactivities to be comparable. These examples are only a few of the possible criteria that one has to take into account for the selection of ligands suitable to the study.

Another matter to be addressed when building ligands for docking is their ionization state. There are many ligand-receptor interactions that only occur due to the existence of an ionized moiety in the ligand. It is possible to predict the ionization state of a given molecule using ChemAxon's MarvinSketch software which is freely available for academic research. It is important to notice that the binding site's micro environment can play a significant role in the ionization state of ligands so the calculation should only be used when there is no

information regarding the active micro species for a given molecule.

The ligand's starting conformation should be reasonably low in energy so an energetic minimization step is advised. This can be done by quantum mechanics, semi-empirical, empirical methods or molecular mechanics. The choice of method is not so important, provided that the conformation is realistic, since the software is going to perform a stochastic search of the ligand's conformational space. We recommend the semi-empirical approach because of it is easy of use, presents good accuracy, has low computational cost and because the state-of-the-art algorithm MOPAC's PM7 is freely available to the academic community and also implementable in many third party software, like Mercury [25], which is a part of the CCDC GOLD Suite.

Protein Preparation

Once the ligands have been built, the preparation of the protein structure for docking is started. The docking calculation can be performed using different software packages, like Auto DockVina [26], GLIDE and GOLD, each one having its own particularities. Although the choice of which software to use is up to the researcher, the focus of this work will be directed towards GOLD because of the reasons formerly explained. Nevertheless, the guidelines presented here apply for most docking engines.

In GOLD, the initial step for protein preparation is the addition of hydrogen atoms, which are not explicit in structures obtained through X-Ray crystallography. If the initial structure already has hydrogens, like the ones obtained through Nuclear Magnetic Resonance (NMR), this step can be safely ignored. The addition of such atoms is automatic upon input of the required command by the user. It is important to notice that some amino acid residues, like histidine (His, H), have more than one possible tautomeric state and this should be reflected in the docking calculation by using the most likely tautomer for the system under evaluation. If no information about tautomerization of binding site residues is available, it is reasonable to test for all of them and select the one yielding better results.

The next step is the removal of water molecules and co-crystallized ligands. All waters inside the binding site must be removed even if they are going to be used in the calculation. The removed waters can be saved to a separate coordinate file which can be used further down the workflow to define its position. The co-crystallized

ligands located inside the binding site must also be removed before docking and the ones outside have no influence in the calculation since the software only considers a small predefined section of the whole protein. Ligands left in the binding site are going to be treated as an additional protein chain so cofactors can be used in calculations.

Setting up the Docking

After the protein and ligand's are prepared, the docking job itself can begin. GOLD has configuration templates for some classes of proteins, like kinases and cytochromes belonging the P450 super family, although for most docking jobs the user will have to set the desired configurations manually.

In GOLD the molecular modeler has the option to perform rigid or flexible docking. Rigid docking has the advantage of experimental evidence to support the obtained results but lacks the freedom of movement that would better represent the real biomolecular system. The opposite applies to flexible docking: the flexibility of the true biological system is mimicked but as the amino acids side chains move, the uncertainties revolving their positions increase and the output conformation might not be a good representative of the reality, thus increasing the number of false positives (i.e. ligands that appear to dock well but in reality do not bind). Also, the computational time required for a flexible docking greatly increases in comparison with the rigid approach. GOLD can make up to ten side chains flexible for a given protein and has also the option for introducing localized backbone movements through rotation of the improper torsion defined by the atom sequence $C\alpha-N-C-C\alpha$. The software has two other options to introduce protein flexibility, namely soft potentials and ensemble docking. The first approach uses alternative Lennard-Jones potentials in the external Vander Waals contributions to the fitness function in order to allow shorter contacts, while the second uses two or more superimposed forms of the same protein in the same docking job to consider different protein conformations. Protein flexibility should be introduced preferably when there is experimental evidence supporting it.

The center and size of the search radius – the binding site –, must be defined by the user. GOLD allows setting an atom, point, ligand or list of atoms as the center of the active site. The binding site radius should cover the size of the largest ligand to be docked. As mentioned before, if there is a co-crystallized ligand complexed with the

biomacro molecule, it is possible to obtain information of where to define this binding site. The user has to be assured that the ligand is indeed in the orthosteric site and not inside an allosteric one, or the opposite depending on the site of interest.

After the definition of the binding site, the user has to load the ligands into the software and define how many GA runs will be performed for each of them. Every GA run results in a different binding mode with a score associated to the predicted ligand-protein interactions. If one docks the same ligand originally bound to the active site, it is possible to set its original conformation as reference, i.e. redocking, which is going to be addressed further down this manuscript. Ligands can be treated as rigid or flexible during the calculation, with flexibility defined as free rotation of rotatable bonds. There are other options for ligand flexibility such as flipping ring corners, amide bonds, pyramidal nitrogen, planar R-NR1R2 and also protonated carboxylic acids. There is also an option to detect possible ligand's internal hydrogen bonds. A ligand should always be treated as flexible, the only exception being cases where the evaluation of the interactions of a specific conformation is desired.

Water molecules in a docking job using GOLD can be treated as fixed, free to rotate, free to translate within a 2 Å radius or both the latter. It is also possible to let the software decide whether or not the water should be bound or displaced by the ligand during the job. One should be cautious when deciding if a water molecule should be included since many ligand-protein interactions are mediated by water. As mentioned earlier, only structures having adequate resolution (≥ 2.7 Å) [23] give useful information about the presence of water molecules inside the binding site. Information about whether or not to include a water molecule in the calculation can be obtained by checking the interactions observed in homologous crystallized proteins or through molecular dynamics studies.

The docking can be run using four different scoring functions in GOLD, namely ChemPLP, GoldScore, ChemScore and ASP. Each one of them has its strengths and weaknesses and their applicability to a given problem has to be evaluated. Guidelines for choosing among them will be described in the next section of this work. The software allows rescoring the solutions of a job using different scoring functions and this can be used to compensate the deficits from the original run. It is possible to enforce diversity by setting the software to generate diverse solutions for the same ligand. Also, the

procedure can be programmed to terminate early if a set of solutions are all within a predefined RMSD threshold from each other and this is particularly useful when working with easy flexible ligands, i.e. ligands with little conformational freedom. All these options have impact on the speed and efficiency of the search. The GA parameters can be optimized to balance the ration between speed and efficiency as needed.

GOLD has several additional parameters that can be tuned to address specific situations, like docking with metalloproteins and complexed metal ions, soft potentials for works involving discrete backbone and side chain movements, covalent docking and others that will not be addressed here. As stated before, extensive research is useful to decide whether or not to introduce them into the calculation. Detailed information about those parameters and how to use them is available on the software's user guide.

Analysis of the Docking Results

A big concern regarding molecular docking is the validity of the results. As stated before, the applicability of a scoring function and the chosen docking configuration has to be assessed beforehand. This is usually done by evaluating their ability to differentiate between active and inactive molecules from a database, e.g. ROC curve analysis [27], by its capacity to align a set of ligands in a way that the variations in their bioactivities is explained, e.g. CoMFA [28], or by redocking the co-crystallized ligand into the binding site to compare the calculated coordinates and interactions with the experimental data. This work will focus on redocking, as it is the simplest of the validation techniques, and information about how the validation protocol was performed for the two first examples is given elsewhere [29,30].

Redocking analysis is very straightforward: one only need to compare the atomic positions of the redocked ligand with the original conformation extracted from the PDB file via RMSD. This can be done easily in GOLD itself by selecting the ligand removed from the binding site in the previous steps as reference. Usually, when the RMSD is low there will be good agreement between the interactions observed for the reference and the ones calculated by the software. More flexible ligands are likely to have higher RMSD values but as long as the pharmacophoric groups are correctly aligned it is safe to assume good accuracy. Most docking jobs are interested in evaluating interactions for ligands other than the reference, i.e. cross-docking. This limits the applicability

of validation through redocking because ligands completely different from the reference might not be well predicted by the computational method. In parallel, the binding site can undergo conformational changes in response to the presence of a ligand, thus reducing the validity of cross-docking situations even if the redocking was successful. In any case, the choice of scoring function based on redocking is performed by assessing which one yields lowest RMSD for the top ranked pose of the redocked ligand.

The results of a docking run are a set of predicted binding modes (poses) for each ligand ranked by the score function according to the interactions between the ligands and the protein. Being so, the expectation is that the top ranked pose represents the conformation and interactions of the true biological system. This is not always true as there will be occasions where the score for the two top ranked binding modes is similar but the poses themselves are distinct. A useful criteria to select which one of them is representative is to evaluate the predicted interactions for both and select the pose that has the most interactions involving key residues, i.e. residues known to be of importance to molecular recognition. We recommend this type of analysis when the difference in score between the top and second ranked binding modes is below 0.5 units. The selected docking poses can be extracted to a PDB file to be analyzed in a structure visualizer such as DSV or PyMol. For analysis of the results DSV is very suitable, first because it is freely available and second because of its automated options for ligand-protein interactions.

Usually docking algorithms fail to correlate the calculated score with the experimental binding affinity and thus it is important to notice that this score can only be used to estimate the affinity of a compound in relation to others [31]. This is the reason why when analyzing the results of a docking job one should consider interactions with key residues as well as the overall orientation of the docked ligands when compared to the reference.

Successful Docking Applications

The first example herein presented of successful molecular docking application is a study by Okada-Junior, et al. [32] targeting malaria disease through the cytochrome *bc*₁ complex of *P. falciparum*, a protozoa from the Plasmodium spp associated with severe malaria [33]. Atovaquone (1) is a first line treatment for malaria disease but resistance has been reported since the early 1990s, thus making the search for new treatments of

paramount importance [34]. The drug is an analogue of ubiquinone and acts by competitively inhibiting the endogenous substrate at the cytochrome bc₁ enzyme complex, thus hampering cellular respiration due to reduction of electron transfer in the mitochondria and ultimately promoting the collapse of the mitochondrial membrane potential [35]. In the mentioned work, a new series of phthalimidic derivatives were synthesized and tested for *in vitro* inhibition of *P. falciparum* from the 3D7 strain (chloroquine sensitive), resulting in two promising compounds with low micromolar inhibitory concentration (IC₅₀) of 4.2 and 6.8 μM, respectively named 2 and 3 in this work (10 and 16 in the original work). Compound 2 was also tested for inhibition of *P. falciparum* K1 strain, which is resistant to the antimalarial drugs chloroquine, pyrimethamine and sulfadoxine, displaying approximately the same inhibition as for the sensitive strain (IC₅₀ = 4.3 μM). Aiming to understand the molecular basis of such results, a docking study using 2 was performed. The methodology follow the guidelines presented in this paper: both Atovaquone and 2's 3D structures were built in DSV and optimized through the semi-empirical algorithm PM6 [36] with MOPAC and the docking was performed in GOLD using ChemPLP scoring function. No 3D structure was available for *P. falciparum* cytochrome bc₁. The more suitable protein structure was found to be the bc₁ cytochrome from *S. cerevisiae* complexed with stigmatellin at 1.9 Å resolution (PDB ID 3CX5) [37]. Atovaquone binds to the complex III of *S. cerevisiae* with high affinity (K_i = 5nM), which led to its development as a nonpathogenic surrogate model for studying bioactive compounds targeting the parasite through this mechanism. A crystallographic structure of the same protein complex crystallized with atovaquone exists (PDB ID: 4PD4) [38] but it has lower resolution and lacks the Rieske protein, which is essential for molecular recognition and thus was discarded for docking. The binding mode of atovaquone in 4PD4 was used as reference after confirmation that the difference in atomic positions of binding site residues was minimal. Also, it is important to mention that atovaquone was modelled in its ionized form due to studies claiming that this is the active microspecies. Docking of atovaquone into 3CX5's binding site resulted in interactions very similar to the ones observed for the reference structure, yielding RMSD= 0.4 and PLPscore= 63.3, indicating good accuracy of the computational approach. Compound 2 was then docked using the validated conditions and its calculated score was 63.6. The binding modes of atovaquone and 2 are presented in Figure 2. Molecular structures for 1, 2, and 3 are presented in Figure 3.

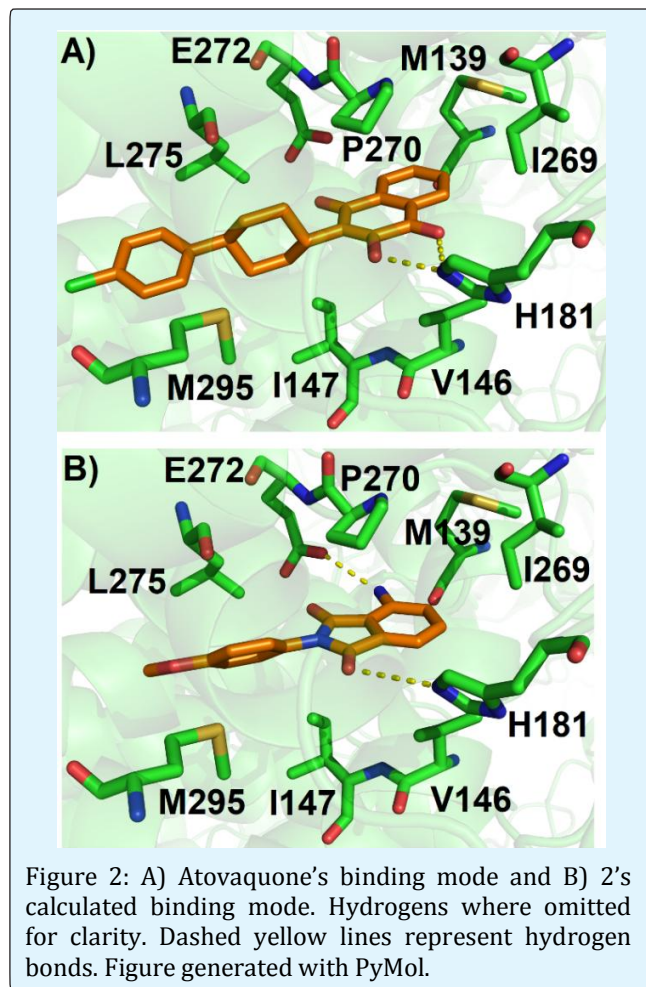


Figure 2: A) Atovaquone's binding mode and B) 2's calculated binding mode. Hydrogens were omitted for clarity. Dashed yellow lines represent hydrogen bonds. Figure generated with PyMol.

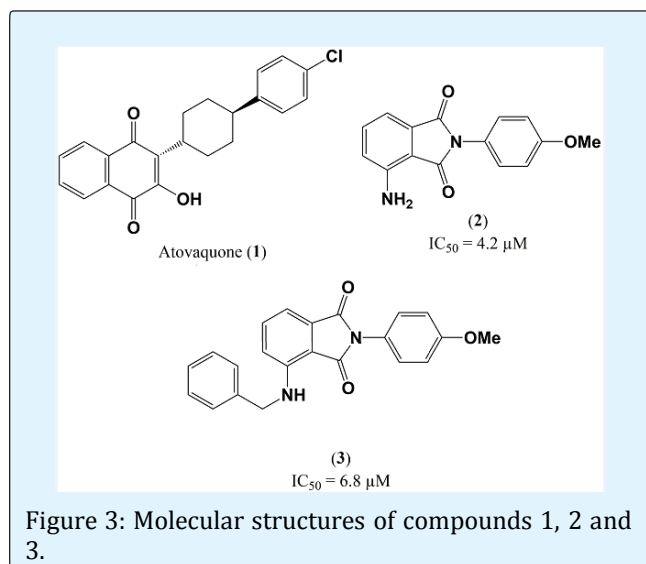


Figure 3: Molecular structures of compounds 1, 2 and 3.

The computational model predicted the occurrence of hydrogen bonds involving the phthalimidic moiety and the side chains of E272 and H181, as well as close contacts with the side chains of residues I147, V146, P271 and with backbone atoms of W142. The methoxyphenyl moiety takes part in hydrophobic interactions with the side chains of M295 and L275. The hydrogen bond involving carbonylic oxygen from the ligand and the side chain of residue H181 is observed for the reference. The interaction of 2 with E272 is not present in the crystal structure of atovaquone but exists in other inhibitor's complexes, like stigmatellin and 5-n-heptyl-6-hydroxy-4,7-Dioxobenzothiazole (HDBT), although in some cases being mediated by a water molecule. The ligand's phthalimidic ring mimics atovaquone's hydrophobic contacts with the side chains of V146 and P271. The methoxyphenyl moiety is positioned close to where atovaquone's cyclohexyl moiety is, simulating the reference's hydrophobic contacts with residues L275 and M295. These results alongside with an enzymatic assay measuring how compound 2 affects the activity of the bc_1 complex decylubiquinol-cytochrome c oxidoreductase, which resulted in 74% inhibition at 70 μM , indicate that the proposed phthalimidic derivative indeed acts through this mechanism of action and can be a promising lead to further development in malaria drug discovery research. This real life example gives insight on how docking can be useful in the understanding of relevant interaction motifs and mechanism of action of newly synthesized compounds.

The second example was developed by Kayode, et al. [39] which used molecular docking as a tool for the identification of inhibitors of mesotrypsin through virtual screening. Mesotrypsin (PRSS3) is one of the digestive enzymes produced and secreted by the human pancreas and its over expression has been observed in several different types of cancer. It also is related to metastasis in some types of cancer, like prostate and pancreatic. This enzyme is very different from other human trypsins because it has almost total resistance to biological trypsin inhibitors [40]. In the mentioned work the researchers performed an ensemble docking using three crystal structures of mesotrypsin, with PDB ID's 3P92 [41], 3P95 [41] and 1HW4 [42] using Glide XP algorithm to screen the Natural Product Database (NPD) and the Food and Drug Administration (FDA) approved Drug Database. Ligands where prepared automatically through Schrodinger's LigPrep module. Molecules having docking scores lower than -10 kcal/mol versus one or more receptors were selected and visually inspected for

hydrogen bonds with residues D189, S190, G192, R193 and G216 from the binding pocket, thus resulting in 28 promising candidates. Among those, 12 were readily commercially available and therefore were obtained and evaluated for inhibitory activity towards mesotrypsin with two of them indeed displaying inhibition in the low micromolar range, namely diminazene (4) ($K_i = 3.66 \mu\text{M}$) and hydroxystilbamidine (5) ($K_i = 10.57 \mu\text{M}$). Diminazene was successfully crystallized with mesotrypsin (PDB ID: 5TP0) and its experimental binding mode matched the calculated one with RMSD = 0.614 over the well-defined electron density region, thus validating the proposed methodology. The second benzamidine moiety from 4, which is located within the solvent channel, presented poor definition of the electron density in that region of the crystal structure, indicating that the drug is able to adopt multiple conformations within this channel and in agreement with other published crystal structures for this molecule complexed with bovine trypsin [43]. The experimental binding mode of the well-defined electron density region is shown in Figure 4. Molecular structures for 4 and 5 are presented in Figure 5.

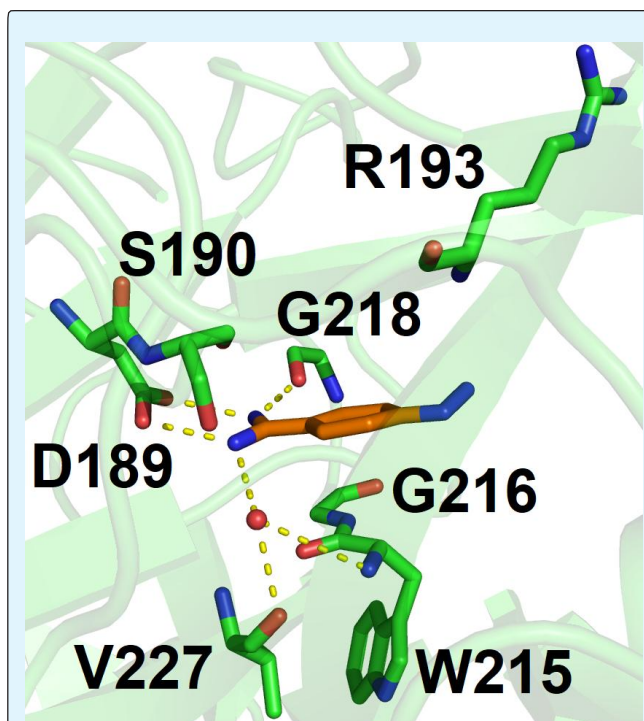
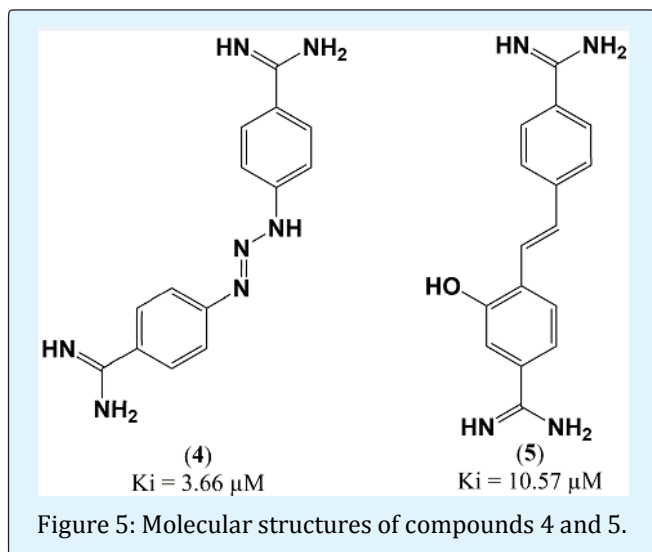


Figure 4: Representation of the binding mode observed for diminazene (4) complexed with PRSS3 (PDB ID: 5TP0). Hydrogens were omitted for clarity. Dashed yellow lines represent hydrogen bonds. Figure generated with PyMol.



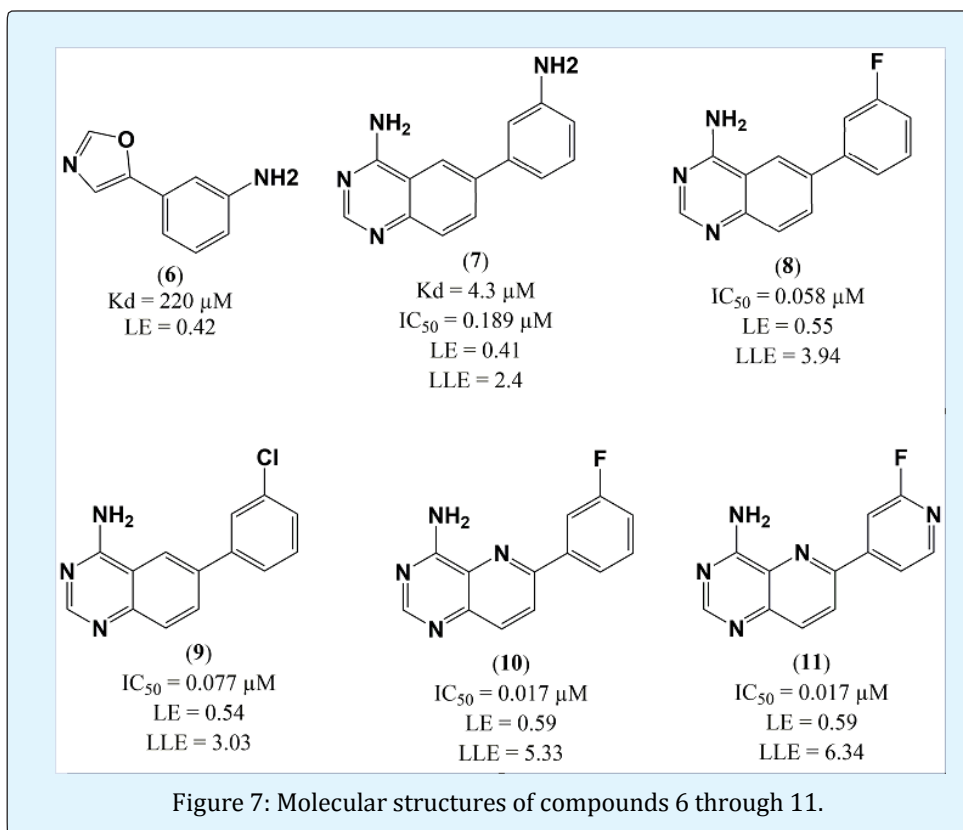
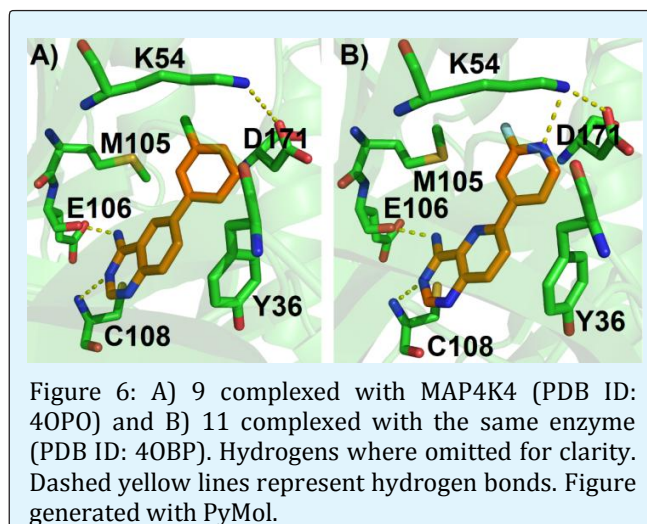
The interaction with R193, not shown in Figure 4 due to poor definition of the electron density in this region, is one of the most critical interactions for this enzyme as this residue is located on a region of the binding site where almost all other trypsin and chymotrypsin family serine proteases have a highly conserved glycine. Therefore, further optimization of this interaction may be a strategy to address selectivity for mesotrypsin over other trypsins. Thus, this molecule presents itself as an excellent starting point for selective mesotrypsin inhibitors, as there is currently almost none [44]. Looking at this example the power of docking as a tool for virtual screening becomes clear. Then, that docking can be useful to drive the design of selective inhibitors towards a target over other related proteins.

The final application discussed here is a work by Crawford, et al. [45] which used molecular docking to enhance the affinity of a lead compound targeting mitogen-activated protein kinase, kinase, kinase and kinase 4 (MAP4K4, a.k.a. HGK). MAP4K4 expression and function are linked to focal adhesion dynamics regulation, embryonic development, insulin sensitivity, systemic inflammation, lung inflammation, atherosclerosis and type-2 diabetes [46]. It has been recently reported that this enzyme is involved in lung adenocarcinoma maintenance through regulation of the MAPK/ERK pathway, acting via inhibition of protein phosphatase 2 [47]. Initially the researchers used high throughput screening (HTS) to search a fragment library, using surface plasmon resonance (SPR), aiming to identify a hit for further development and they reported the

progression for an oxazole fragment, herein named 6 (1 in the original work), with $K_d = 220 \mu\text{M}$ and ligand efficiency (LE) of 0.42. Based on the comparison between the fragment and other kinase inhibitors they postulated that the oxazole acted competitively at the enzyme's ATP site and thus performed a molecular docking targeting this site. The docking was performed with Glide SP algorithm through Schrodinger's Maestro interface, with LigPrep module being the method for ligand preparation and using an unpublished in-house MAP4K4 structure complexed with one of the HTS hits as 3D coordinates (resolution of 2.35 Å). Binding poses were evaluated based on Glide's docking score and by assessing the formation of hydrogen bonds to hinge residues E106 and C108 and other favorable intermolecular interactions with binding site residues. The docking software predicted the oxazole ring to be taking part in a conventional hydrogen bond with the backbone nitrogen of residue C108 and also a non-classic hydrogen bond with the carbonyl oxygen from E106's backbone. Aiming to maximize these hydrogen bonds they synthesized biaryl compounds which had complementary hydrogen bond donor/acceptor characteristics. Restrictions were applied when designing these fragments as to keep them inside an ideal drug-like space, thus permitting only fragment growths leading to molecules with molecular weight below 350 Da and cLog P below 3.5. This step led to the identification of a quinazoline 7 (8 in the original work) having 55-fold increase in potency ($K_d = 4.3 \mu\text{M}$ and $\text{IC}_{50} = 0.189 \mu\text{M}$) that still had comparable LE to the lead fragment 6 (LE = 0.41) and also having good lipophilic ligand efficiency (LLE) of 2.4. Then a structure activity relationship (SAR) study was conducted using the identified scaffold as template for exploring the effects of different substituted aryl groups. *Ortho* and *para* substitutions resulted in decreased potency while introduction of halogens in *meta* position yielded more potent molecules. From this series, molecules 8 and 9 (19 and 22 in the original publication) displayed the best enhancement, with IC_{50} of 0.058 and 0.077, LE of 0.55 and 0.54 and LLE of 3.94 and 3.03, respectively. The X-ray co-crystal structure of compound 9 was solved (PDB ID: 4OBO) and revealed the molecular basis for the observed results from this series. First, the amino quinazoline core indeed addressed the desired hydrogen bonds initially observed for 6 and an additional non-classic hydrogen bond exists between this moiety and the backbone carbonyl oxygen of C108. Substitutions in *para* position were detrimental due to a salt bridge between K54 and D171 that blocked further expansions in this position and *ortho* position was not well tolerated due to unfavorable torsion strains arising from binding into a nearly planar

binding pocket. The side chain of residue Y36, which is part of the P-loop, moved down to interact via T-stacking with the quinazoline core. The halogen atom in *meta* position takes part in a favorable interaction with the gatekeeper residue M105. From the data obtained in this step it became clear that although potency was enhanced, the molecules were deviating from the ideal drug-like space initially proposed for cLogP and thus to increase polarity a nitrogen walk around the quinazoline core was undertaken, replacing carbon for nitrogen atoms in positions 5, 7 and 8. The results of these modifications showed that Y36 had major impacts on the ligand's potencies due to the formerly mentioned T-stacking being unfavorable for the substitutions in positions 7 and 8, mainly because of repulsion involving the pi system and the aromatic lone pairs. Substitution in position 5, producing compound 10 (26 in the original work), indeed enhanced both potency and LLE ($IC_{50} = 0.017 \mu\text{M}$ and $LLE = 5.33$), driving the series back into the desired drug-like space. Then, a new series of close analogues of 10 were synthesized, culminated in compound 11 (29 in the original work) having comparable potency of $0.017 \mu\text{M}$ and superior LLE of 6.34, as well as favorable *in vivo* pharmacokinetic properties. The crystal structure for

MAP4K4 complexed with 11 was also solved (PDB ID: 40BP) and displayed virtually the same interactions, having only an additional hydrogen bond with the side chain of K79. The binding modes of both crystallized molecules are displayed in Figure 6. Molecular structures for compounds 6 through 11 are displayed in Figure 7.



This example illustrates how interactions predicted by molecular docking can be useful to drive synthetic efforts towards enhancing potency of a lead compound.

Conclusion

A low cost docking methodology was discussed with the goal of introducing key aspects of molecular docking to newcomers to the field, like how to choose a starting crystal structure, build ligands for docking correctly, how to interpret the results and other aspects. The discussed methodology was successfully applied and published in combination with other results. In addition, other examples illustrating the applicability of molecular docking were also discussed. Those examples show the importance of extensive bibliographic survey about the system to be modelled. This methodology can serve as a guide and definitely has room for modifications, which we strongly advise in order to properly address the docking problem to be solved.

Acknowledgments

The authors would like to thank São Paulo Research Foundation (FAPESP, process number: 2018/00187-7), for the financial support. Also, we thank Guilherme Gonçalves Costa (IQ/Unicamp) for English language revision.

References

- Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and Scoring in Virtual Screening for Drug Discovery Methods and Applications. *Nat Rev Drug Discov* 3(11): 935-949.
- Meng XY, Zhang HX, Mezei M, Cui M (2011) Molecular Docking: A Powerful Approach for Structure-based Drug Discovery. *Curr Comput Aided Drug Des* 7(2): 146-157.
- Pagadala NS, Syed K, Tuszynski J (2017) Software for molecular docking: a review. *Biophys Rev* 9(2): 91-102.
- Elokely KM, Doerksen RJ (2013) Docking Challenge: Protein sampling and Molecular Docking Performance. *J Chem Inf Model* 53(8): 1934-1945.
- Huang N, Shoichet BK, Irwin JJ (2006) Benchmarking Sets for Molecular Docking. *J Med Chem* 49(23): 6789-6801.
- <https://www.rcsb.org/>
- <https://www.ebi.ac.uk/chembl/>
- <http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>
- <https://swissmodel.expasy.org/>
- Dassault Systèmes BIOVIA , Discovery Studio Modeling Environment, San Diego, Release 2017.
- The PyMOL Molecular Graphics System , Version 0.99, Schrödinger, LLC, 2002.
- MarvinSketch (version 6.2.2, calculation module developed by ChemAxon, <http://www.chemaxon.com/products/marvin/marvinsketch/>, 2014.
- Stewart JP MOPAC (2016) Stewart Computational Chemistry, Colorado Springs, CO, USA.
- Jones G (1997) Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 267(3): 727-748.
- Korb O, Stutzle T, Exner TE (2009) Empirical Scoring Functions for Advanced Protein-Ligand Docking with PLANTS. *J Chem Inf Model* 49(1): 84-96.
- Baxter CA, Murray CW, Clark DE, Westhead DR, Eldridge MD (1998) Flexible Docking Using Tabu Search and an Empirical Estimate of Binding Affinity. *Proteins* 33(3): 367-382.
- Mooij WT, Verdonk ML (2005) General Targeted Statistical Potentials for Protein-Ligand Interactions. *Proteins* 61(2): 272-287.
- Li Y (2014) Comparative Assessment of Scoring Functions on an Updated Benchmark: 2. Evaluation Methods and General Results. *J Chem Inf Model* 54(6): 1717-1736.
- Molecular Operating Environment (MOE) (2013) Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2018.
- Glide Schrödinger (2018) LLC, New York.
- Vyas VK, Ukawala RD, Ghatge M, Chintcha C (2012) Homology Modeling a Fast Tool for Drug Discovery:

- Current Perspectives. *Drug Discov Today* 14(13-14): 673-678.
22. McGovern SL, Shoichet BK (2003) Information Decay in Molecular Docking Screens Against Holo, Apo, and Modeled Conformations of Enzymes. *J Med Chem* 46(14): 2895-2907.
 23. Wlodawer A, Minor W, Dauter Z, Jaskolski M (2008) Protein Crystallography for non-crystallographers, or How to Get the Best (but not more) from Published Macromolecular Structures. *FEBS J* 275(1): 1-21.
 24. Saxena et al. (2013) Fundamentals of Homology Modeling Steps and Comparison among Important Bioinformatic Tools: An Overview. *Sci Int* 1(7): 237-252.
 25. Macrae CF (2008) Mercury CSD 2.0 - new features for the visualization and investigation of crystal structures. *J Appl Crystallogr* 41(2): 466-470.
 26. Trott O, Olson AJ (2010) AutoDockVina: Improving the Speed and Accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31(2): 455-461.
 27. Empereur-mot C, Guillemain H, Latouche A, Zaguray JF, Viallon V, et al. (2015) Predictiveness Curves in Virtual Screening. *J Cheminform* 7: 52.
 28. Cramer RD, Patterson DE, Bunce JD (1988) Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J Am Chem Soc* 110(18): 5959-5967.
 29. Hevener KE, Zhao W, Ball DM, Babaoglu K, Qi J, et al. (2009) Validation of molecular Docking Programs for Virtual Screening Against Dihydropteroate Synthase. *J Chem Inf Model* 49(2): 444-460.
 30. Abu-Hammad A, Zalloum WA, Zalloum H, Abu-Sheikha G, Taha MO (2009) Homology Modeling of MCH1 Receptor and Validation by docking/scoring and protein-aligned CoMFA. *Eur J Med Chem* 44(6): 2583-2596.
 31. Vilar S, Constanzi S (2012) Predicting Biological Activities through QSAR Analysis and Docking-based Scoring Methods *Mol Biol* 914: 271-284.
 32. Okada-Junior CY (2018) Phthalimide Derivatives with Bioactivity against Plasmodium falciparum: Synthesis, Evaluation, and Computational Studies Involving bc1 Cytochrome Inhibition. *ACS Omega* 3(8): 9424-9430.
 33. Biamonte MA (2013) Recent Advances in Malaria Drug Discovery. *Bioorg Med Chem Lett* 23(10): 2829-2843.
 34. Vaidya AB, Mather MW (2000) Atovaquone Resistance in Malaria Parasites. *Drug Resist Updat* 3(5): 283-287.
 35. Barton V, Fisher N, Biagini GA, Ward SA, O'Neill PM (2010) Inhibiting Plasmodium Cytochrome bc₁: a Complex Issue. *Curr Opin Chem Biol* 14(4): 440-446.
 36. Stewart JP (2007) Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J Mol Model* 13(12): 1173-1213.
 37. Solmar SR, Hunte C (2008) Structure of Complex III with Bound Cytochrome c in Reduced State and Definition of a Minimal Core Interface for Electron Transfer. *J Biol Chem* 283(25): 17542-17549.
 38. Birth D, Kao WC, Hunte C (2014) Structural Analysis of Atovaquone-Inhibited Cytochrome bc₁ Complex Reveals the Molecular Basis of Antimalarial Drug Action. *Nat Commun* 5: 4029.
 39. Kayode O, Huang Z, Soares AS, Caulfield TR, Dong Z, et al. (2017) Small Molecule Inhibitors of Mesotrypsin from a Structure-based Docking Screen. *PLoS One* 12(5): e0176694.
 40. Salameh MA, Radisky ES (2013) Biochemical and Structural Insights into Mesotrypsin: an Unusual Human Trypsin. *Int J Biochem Mol Biol* 4(3): 129-139.
 41. Salmeh MA, Soares AS, Hockla A, Radisky DC, Radisky ES (2011) The P2' Residue is a Key Determinant of Mesotrypsin Specificity: Engineering High-affinity Inhibitor with Anticancer Activity. *Biochem J* 440(1): 95-105.
 42. Katona G, Berglund GI, Hajdu J, Garf L, Szilagyi L (2002) Crystal structure reveals basis for the inhibitor resistance of human brain trypsin. *J Mol Biol* 315(5): 1209-1218.
 43. Perilo CS, Pereira MT, Santoro MM, Nagem RAP (2010) Structural Binding Evidence of the Trypanocidal Drugs Berenil® and Pentacarinat®

- Active Principles to a Serine Protease Model. *Int J Biol Macromol* 46(5): 502-511.
44. deVeer SJ, Li CY, Swedberg JE, Schroeder EI, Craik DJ (2018) Engineering potent mesotrypsin inhibitors based on the plant-derived cyclic peptide, sunflower trypsin inhibitor-1. *Eur J Med Chem* 155: 695-704.
45. Crawford TD, Ndubaku CO, Chen H, Boggs JW, Bravo BJ, et al. (2014) Discovery of Selective 4-Aminopyridopyrimidine Inhibitors of MAP4K4 Using Fragment-Based Lead Identification and Optimization. *J Med Chem* 57(8): 3484-3493.
46. Gao X, Gao C, Liu G, Hu J (2016) MAP4K4: An Emerging Therapeutic Target in Cancer. *Cell Bio sci* 6: 56.
47. Gao X, Chen G, Gao C, Zhang DH, Kuan SF, et al. (2017) MAP4K4 is a novel MAPK/ERK pathway regulator required for lung adenocarcinoma maintenance. *Mol Oncol* 11(6): 628-639.

