

Crisiele Fontana

**Análise das Características Estruturais de  
Derivados de Indolocarbazolas Envolvidas na  
Complexação a Receptores Alvo**

Porto Alegre

2022



Crisiele Fontana

**Análise das Características Estruturais de  
Derivados de Indolocarbazolas Envolvidas na  
Complexação a Receptores Alvo**

Tese submetida ao Programa de Pós-Graduação em Biologia Celular e Molecular do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul como parte dos requisitos necessários para a obtenção do grau de Doutora em Biologia Celular e Molecular.

Orientador: Hugo Verli

Porto Alegre

2022

---

Crisciele Fontana

Análise das Características Estruturais de  
Derivados de Indolocarbazolas Envolvidas na  
Complexação a Receptores Alvo/ Crisciele Fontana. – Porto Alegre, 2022.  
179 f.

Orientador: Hugo Verli

Tese (Doutorado) – Universidade Federal do Rio Grande do Sul, Centro de  
Biotecnologia do Estado do Rio Grande do Sul, Programa de Pós-Graduação em  
Biologia Celular e Molecular, Porto Alegre, BR-RS, 2022.

1. Indolocarbazolas. 2. Parametrização. 3. Conformação. 4. Interação  
Intermolecular. 5. Dinâmica Molecular. I. Verli, Hugo, orient. II. Título

---

Crisiele Fontana

# **Análise das Características Estruturais de Derivados de Indolocarbazolas Envolvidas na Complexação a Receptores Alvo**

Tese submetida ao Programa de Pós-Graduação em Biologia Celular e Molecular do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul como parte dos requisitos necessários para a obtenção do grau de Doutora em Biologia Celular e Molecular.

Orientador: Hugo Verli

---

**Dr. Hugo Verli**  
Orientador

---

**Dr. Rodrigo Ligabue Braun**  
Depto. de Farmacociências  
UFCSPA

---

**Dr. Paulo Augusto Netz**  
Depto. de Físico-Química  
UFRGS

---

**Dr. Hubert Karl Stassen**  
Depto. de Físico-Química  
UFRGS

---

**Dr. José Fernando Bachega**  
(Suplente)  
UFCSPA, PPGBCM-UFRGS

Porto Alegre  
2022



*Este trabalho é dedicado às crianças adultas que,  
quando pequenas, sonharam em se tornar cientistas.*





# Agradecimentos

A Deus, pela dávida da vida, por mais essa oportunidade, pela presença constante, pelo inimaginável, por todas as graças providenciais.

Ao professor Hugo Verli, pela confiança e oportunidade do doutorado, pela inspiração científica e pelos mais diversos ensinamentos e conselhos.

À minha família, avós, avôs, tias, tios, pai e mãe, por todos os esforços e ensinamentos sobre os valores do trabalho, do desenvolvimento intelectual e da retidão de espírito que me fizeram chegar até aqui; aos meus irmãos e primos pelo suporte e compreensão durante essa trajetória e por tornarem os dias mais leves. Essa conquista é vossa também!

Aos meus amigos, pelas palavras de incentivo e confiança nos momentos de fragilidade e incerteza, pela disposição para o auxílio e a escuta e, principalmente, para a celebração de diferentes momentos.

Aos colegas do GBE Bianca, Jota, Felipe, Marcelo, Pablo e Ana Paula pelo conhecimento compartilhado, pelo auxílio prestado e pelas profundas conversas, foram essenciais para o início dessa caminhada; aos colegas Laís, Brisa, Shabir, Bruna, Vinícius e João, com quem concluo a jornada, sou muito grata pelo ambiente colaborativo e afetuoso e por compartilhar o 'retorno ao presencial' com vocês.

Aos professores Arthur Germano Fett-Neto e Hubert Karl Stassen, membros da Comissão de Acompanhamento, pela prestatividade, conselhos e auxílios.

Aos membros da banca de avaliação, pela disponibilidade e pelo conhecimento agregado aos trabalhos aqui apresentados.

À UFRGS, ao Centro de Biotecnologia e ao Programa de Pós-Graduação em Biologia Celular e Molecular pela excelência do ensino que muito agregou ao meu desenvolvimento acadêmico.

À Silvinha Centeno, pelo alto-astral, pelo sorriso solícito, pelas soluções afetuosas, por ser mais do que secretária.

À Alexandra Elbakyan e a todos que contribuem para uma ciência livre e colaborativa.

A todos os que não foram citados nominalmente, mas que tiveram influência na minha vida pessoal e profissional.



*“Cada pessoa deve trabalhar  
para seu próprio aperfeiçoamento  
e, ao mesmo tempo, participar  
da responsabilidade coletiva  
por toda a humanidade.  
(Marie Curie, 1867-1934.)*



# Resumo

Indolocarbazolas (ICZs) são produtos naturais glicosilados com ampla diversidade química e espectro de atividade biológica, notadamente, antitumoral, antimicrobiana e antiviral. Por diferentes abordagens sintéticas, esta classe de pequenos ligantes é complementada com uma variedade de modificações químicas que visam melhorar sua biodisponibilidade e especificidade. Embora sejam conhecidos complexos de ICZs com proteínas alvo, apoiando a otimização das propriedades bioativas dos ligantes baseada em estrutura, as propriedades conformacionais dos substituintes flexíveis em ICZs, como carboidratos, ainda são parcialmente inacessíveis aos métodos clássicos de biologia estrutural. Como complemento a esses métodos, cálculos de mecânica molecular podem oferecer dados exclusivos para acessar o conjunto conformacional de moléculas flexíveis e fornecer detalhes estruturais úteis para melhorar os recursos terapêuticos por meio de relações conformação-atividade e estrutura-atividade. A fim de avaliar as características conformacionais de ICZs e suas interações no solvente aquoso, simulações de dinâmica molecular foram realizadas através do campo de força GROMOS. Parâmetros específicos para a descrição de ICZs foram derivados de cálculos quânticos e ajustados a dados experimentais de espectroscopia de ressonância magnética nuclear para reproduzir na mecânica molecular a flexibilidade e o potencial eletrostático envolvido na interação intermolecular desses ligantes. Cálculos de metadinâmica apoiaram a amostragem ampliada dos sistemas e elucidaram o conjunto conformacional das estruturas. A partir dos novos parâmetros topológicos, foram realizadas simulações de 30 derivados de ICZs com diferentes padrões de substituição e porções sacarídicas. Algumas variações na rede de substituição mostraram estar envolvidas em modificações nas características de interação com a água e liberdade rotacional de diedros. A frequente interconversão entre conformeros de carboidratos ampliou a percepção do espaço conformacional de ICZs em solução, enquanto concorda bem com o princípio de seleção conformacional do ligante para a complexação ao alvo e revela o caráter limitante das técnicas experimentais clássicas para a descrição da conformação dessas moléculas. A presença de substituintes halogênicos e a combinação de grupos polares vizinhos aumentam a afinidade para ligações de hidrogênio com o solvente, o que pode afetar a energia livre de complexação a proteínas alvo, devido a variações nos custos energéticos de dessolvatação. O uso de simulações computacionais com parâmetros calibrados para dados experimentais pode fornecer informações precisas da dinâmica de ICZs em solução. A possibilidade única de acessar estados conformacionais e a disponibilidade para interações intermoleculares biologicamente relevantes pode orientar o desenho racional de ligantes com distribuição e seletividade melhoradas e assim, expandir o potencial terapêutico de ICZs.

**Palavras-chaves:** Indolocarbazolas. Parametrização. Conformação. Interação Intermolecular. Dinâmica Molecular.



# Abstract

Indolocarbazoles (ICZs) are glycosylated natural products presenting wide chemical diversity and a broad spectrum of biological activity, notably antitumor, antimicrobial and antiviral. Through different synthetic approaches, this class of small ligands has already been complemented with a variety of chemical modifications aiming to improve its bioavailability and specificity. While complexes of ICZs with target proteins are known, thus supporting structure-based optimization of ligands bioactive properties, the conformational properties of flexible substituents on ICZs, such as carbohydrates, are still partially inaccessible to classical structural biology methods. As a complement to these methods, molecular mechanics-based calculations can offer a unique set of data to assess the conformational ensemble of flexible molecules and provide useful structural details to improve therapeutic properties through conformational-activity and structure-activity relationships. In order to evaluate the conformational features of ICZs and their interactions in aqueous solvent, molecular dynamics simulations were performed using the GROMOS force field. Specific parameters for the description of ICZs were derived from quantum calculations and adjusted to experimental data from nuclear magnetic resonance spectroscopy to reproduce in molecular mechanics the flexibility and electrostatic potential involved in intermolecular interaction of these ligands. Metadynamics calculations supported the systems enhanced sampling and elucidated the compounds conformational ensemble. From the new topological parameters, it were carried out simulations of 30 ICZ derivatives presenting different substitution patterns and carbohydrate moieties. Some variations on the substitution network were shown to be involved in changes in water interaction features and dihedral rotational freedom. The frequent interconversion between carbohydrate conformers expanded the perception of the conformational space of ICZs in solution, while it matches with the conformational selection of the ligand for target complexation and exposes the limiting character of the classical approaches to describe these ligands conformations. The presence of halogen substituents and the combination of neighboring polar groups increase their affinity for hydrogen interactions with the solvent, which can affect the free energy to complexation with target proteins due to variations in desolvation energy costs. The use of computer simulations with parameters calibrated to experimental data can provide accurate information on the dynamics of ICZs in solution. The unique possibility to assess biologically relevant conformational states and availability for intermolecular interactions can guide the rational design of ligands with improved distribution and selectivity, and thus, to expand the therapeutic potential of ICZs.

**Key-words:** Indolocarbazoles. Parameterization. Conformation. Intermolecular Interaction. Molecular Dynamics.





# Lista de ilustrações

Figura 1 – Isômeros de indolocarbazolas e alguns organismos produtores . . . . .	24
Figura 2 – Estrutura molecular de rebecamicina, staurosporina e análogos. . . . .	25
Figura 3 – Complexos cristalográficos de staurosporina. . . . .	30
Figura 4 – Complexos cristalográficos de UCN-01. . . . .	31
Figura 5 – Complexo cristalográfico de midostaurin-DYRK1A. . . . .	33
Figura 6 – Complexo cristalográfico de SA315F-DNA-Top1. . . . .	34
Figura 7 – Complexo cristalográfico do substrato 4'- <i>O</i> -demetil-1,11-didescloro- rebecamicina ligado à enzima AtmM. . . . .	36
Figura 8 – Representação de condições periódicas de contorno e tempo de integração.	43
Figura 9 – Representação dos termos que compõem um campo de força. . . . .	45
Figura 10 – Estratégias do desenho de fármacos auxiliado por computador (CADD).	50
Figura 11 – Fluxograma para a construção de topologias. . . . .	56
Figura 12 – Fluxograma de geração de cargas atômicas parciais. . . . .	57
Figura 13 – Representação da energia livre durante cálculo de metadinâmica. . . . .	59



# Lista de abreviaturas e siglas

AGP2	Glicoproteína Ácida Alfa 2
ATP	Trifosfato de Adenosina
CADD	Desenho de Fármacos Auxiliado por Computador ( <i>Computer-Aided Drug Design</i> )
CDK2	Cinase 2 Dependente de Ciclina
Chk1	Proteína de Checkpoint Cinase 1
CV	Variáveis Coletivas ( <i>Collective Variables</i> )
DYRK1A	Cinase de Dupla Especificidade Regulada por Fosforilação de Tirosina 1A
EMA	<i>European Medicines Agency</i>
FDA	<i>Food and Drug Administration</i>
FES	Superfície de Energia Livre ( <i>Free-Energy Surface</i> )
FLT3	Fator Estimulador de Colônia de Macrófagos Tirocino Cinase ( <i>FMS-like tyrosine kinase 3</i> )
HTS	Triagem de Alto Rendimento ( <i>High-Throughput Screening</i> )
ICZ	Indolocarbazola
KIT	Proto-Oncogene Receptor de Tirocino Cinase
MD	Dinâmica Molecular ( <i>Molecular Dynamics</i> )
MM	Mecânica Molecular
NMR	Ressonância Magnética Nuclear ( <i>Nuclear Magnetic Resonance</i> )
PDGFR $\beta$	Receptor do Fator de Crescimento Derivado de Plaquetas $\beta$
PDK1	Cinase 1 Dependente de 3-fosfoinositídeo
PKA	Proteína Cinase Dependente de cAMP
PKC	Proteína Cinase C
QM	Mecânica Quântica ( <i>Quantum Mechanics</i> )

QSAR	Relação Estrutura-Atividade Quantitativa ( <i>Quantitative Structure Activity Relationship</i> )
RESP	Potencial Eletrostático Restrito, ( <i>Restrained Electrostatic Potential</i> )
Top1	Topoisomerase do tipo I
Top2	Topoisomerase do tipo II
vdW	Parâmetros de van der Waals
VEGFR2	Receptor do Fator de Crescimento Endotelial Vascular 2

# Lista de símbolos

$\alpha$	Alfa
$\beta$	Beta
$\gamma$	Gama
$\Delta$	Delta
$\theta$	Theta
$\nu$	Nu
$\pi$	Pi
$\sigma$	Sigma
$\phi$	Phi
$\omega$	Ômega
Å	Ângstron
Bar	Unidade de Pressão
f	Femto
J	Joule
$J$	Constante de Acoplamento
K	Kelvin
Hz	Hertz
$\mu$	Micro
n	Nano
p	Pico



# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>23</b>
<b>1.1</b>	<b>Indolocarbazolas</b>	<b>23</b>
1.1.1	Mecanismos de Ação Molecular	29
1.1.1.1	Inibidores de Cinases	29
1.1.1.2	Inibidores de Topoisomerases	32
<b>1.2</b>	<b>Resolução Estrutural de Ligantes Glicosilados</b>	<b>36</b>
1.2.1	Espectroscopia de Ressonância Magnética Nuclear	37
1.2.1.1	Constantes de Acoplamento Vicinal $^3J_{H,H}$	39
<b>1.3</b>	<b>Dinâmica Molecular</b>	<b>41</b>
1.3.1	Campos de Força	42
1.3.2	Construção de Topologias de Pequenos Ligantes	46
<b>1.4</b>	<b>Desenho de Fármacos Auxiliado por Computador</b>	<b>48</b>
<b>2</b>	<b>JUSTIFICATIVA</b>	<b>51</b>
<b>3</b>	<b>OBJETIVOS</b>	<b>53</b>
3.1	Objetivo Geral	53
3.2	Objetivos Específicos	53
<b>4</b>	<b>PROCEDIMENTOS METODOLÓGICOS</b>	<b>55</b>
<b>4.1</b>	<b>Construção das Topologias</b>	<b>55</b>
4.1.1	Derivação de Cargas Atômicas Parciais	55
4.1.2	Novos Potenciais Torcionais	56
<b>4.2</b>	<b>Dinâmica Molecular</b>	<b>57</b>
<b>4.3</b>	<b>Metadinâmica</b>	<b>58</b>
<b>5</b>	<b>RESULTADOS</b>	<b>61</b>
<b>5.1</b>	<b>Capítulo I: Aspectos Estruturais de Staurosporina e Análogos Implicados na Complexação a Receptores-Alvo</b>	<b>61</b>
5.1.1	Preâmbulo	61
<b>5.2</b>	<b>Capítulo II: Dinâmica Conformacional e da Afinidade de Interação com o Solvente de Rebecamicina e Derivados</b>	<b>82</b>
5.2.1	Preâmbulo	82
<b>6</b>	<b>DISCUSSÃO GERAL</b>	<b>117</b>
<b>6.1</b>	<b>Derivação dos Parâmetros Topológicos</b>	<b>117</b>
<b>6.2</b>	<b>Amostragem Conformacional</b>	<b>119</b>

6.3	Características de Interação com o Solvente . . . . .	120
7	CONCLUSÕES . . . . .	123
8	PERSPECTIVAS . . . . .	125
	REFERÊNCIAS . . . . .	127
	<b>APÊNDICES</b>	<b>149</b>
	APÊNDICE A – ATRACAMENTO MOLECULAR DE LIGANTES PIPERAZÍNICOS . . . . .	151
	APÊNDICE B – TRIAGEM VIRTUAL DE LIGANTES PEPTÍDI- COS ANTI-SARS-COV-2 . . . . .	165



# 1 Introdução

## 1.1 Indolocarbazolas

Indolocarbazolas (ICZs) constituem uma família de compostos heterocíclicos, isolados pela primeira vez em 1977 de actinobactérias *Streptomyces staurosporeus* [1]. Pouco tempo depois, compostos desta família puderam ser encontrados também em fungos [2], cianobactérias [3] e invertebrados marinhos [4]. Sua estrutura é composta por um sistema de anéis planos constituído de elementos indol e carbazola. Diferentes arranjos destes grupos químicos resultam em cinco principais sistemas de anéis isoméricos que compõem a família de ICZs. Os isômeros são definidos pela posição e orientação da fusão entre o indol (na ligação C-2-C-3) com um anel de benzeno da unidade de carbazola, nomeados de acordo com as regras de nomenclatura IUPAC como 11,12-di-hidroindolo[2,3-*a*]carbazola, 5,12-di-hidroindolo[3,2-*a*]carbazola, 5,7-di-hidroindolo[2,3-*b*]carbazola, 5,11-di-hidroindolo[3,2-*b*]carbazola e 5,8-di-hidroindolo[2,3-*c*]carbazola (Figura 1). Outros arranjos envolvendo grupos indol e carbazola também são possíveis, no entanto são de ocorrência rara e geralmente não associados à definição "clássica" de ICZ [5]. Por outro lado, outros compostos podem ser considerados parte da mesma família em função das estreitas relações biossintéticas, como é o caso de derivados de bis-indolilmaleimida [6].

A maioria dos ICZs isolados da natureza são derivados do isômero 11,12-di-hidroindolo[2,3-*a*]carbazola (por simplicidade, passa a ser referido no texto como indolo[2,3-*a*]carbazola), que por apresentar significativa atividade biológica, é a classe de compostos mais estudada e complementada por uma grande variedade de compostos sintéticos [7]. Os compostos estudados no contexto deste trabalho são derivados do isômero indolo[2,3-*a*]carbazola e apresentam uma hexopiranosose ligada ao sistema de anéis planos. A porção sacarídica é ligada à porção indol por meio de uma ou duas ligações glicosídicas com o átomo de nitrogênio deste grupamento. A seguir, os compostos são apresentados em grupos estruturalmente e quimicamente relacionados: rebecamicina e seus derivados, staurosporina e seus derivados e streptocarbazolas.

Desde a primeira descrição em 1977, mais de 130 ICZs foram isolados de diferentes organismos incluindo bactérias, fungos e invertebrados. Até o momento, pelo menos 23 derivados de rebecamicina foram isolados de invertebrados marinhos e bactérias, majoritariamente estreptomicetos e actinobactérias. Entre as staurosporinas, pelo menos 55 compostos de origem natural são conhecidos. Eles foram obtidos de tunicados, mariscos, ascídias, lesmas e bactérias, principalmente estreptomicetos, actinobactérias e cianobactérias [7].

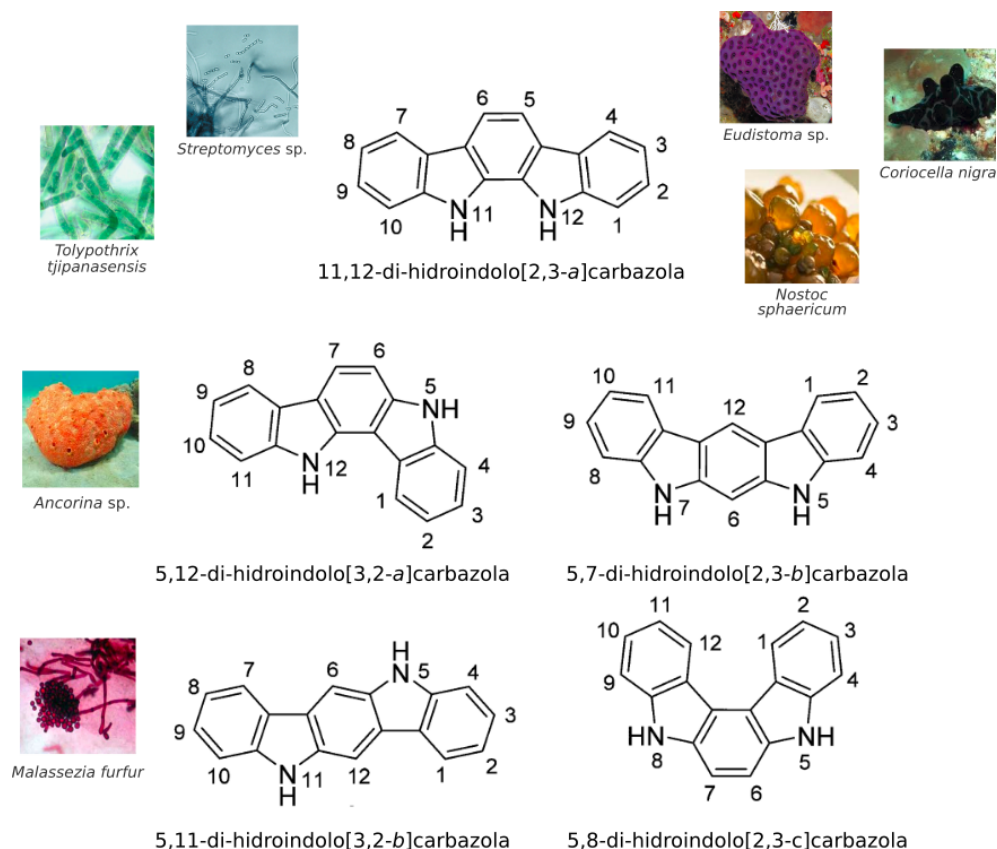


Figura 1 – Cinco principais isômeros indolocarbazólicos e alguns dos principais organismos produtores.

O composto rebecamicina foi descoberto em cultura de actinomicetos *Nocardia aerocolonigenes* (*Saccharothrix aerocolonigenes*) [8, 9, 10] e é o composto líder de uma série de análogos identificados posteriormente e derivados sintetizados. Os compostos deste grupo apresentam apenas uma ligação *N*-glicosídica ligando a porção sacarídica ao sistema plano de anéis (ou porção aglicona), que por sua vez é caracterizado pela presença de um anel de imida (Figura 2 A). Logo que isolada, a estrutura de rebecamicina foi determinada por cristalografia de raios-X e síntese total [8, 11].

No processo de síntese de rebecamicina por culturas de *S. aerocolonigenes*, ocorre naturalmente a produção de análogos desprovidos de alguns substituintes como cloro e metila. Tanto rebecamicina quanto esses análogos produzidos em menor quantidade, identificados como 11-descloro-rebecamicina, 1,11-didescloro-rebecamicina e 4-*O*-desmetil-1,11-didescloro-rebecamicina, apresentaram atividade antitumoral em diferentes linhagens celulares [10, 9]. Rebecamicina apresentou atividade contra leucemia e melanoma em murinos e inibiu o crescimento de células de adenocarcinoma de pulmão humano, produzindo quebras de fita simples no DNA dessas células [9]. O composto 11-descloro-rebecamicina apresenta ainda atividade antibacteriana [12].

Novos derivados sintéticos de indolo[2,3-*a*]carbazola puderam ser facilmente obtidos

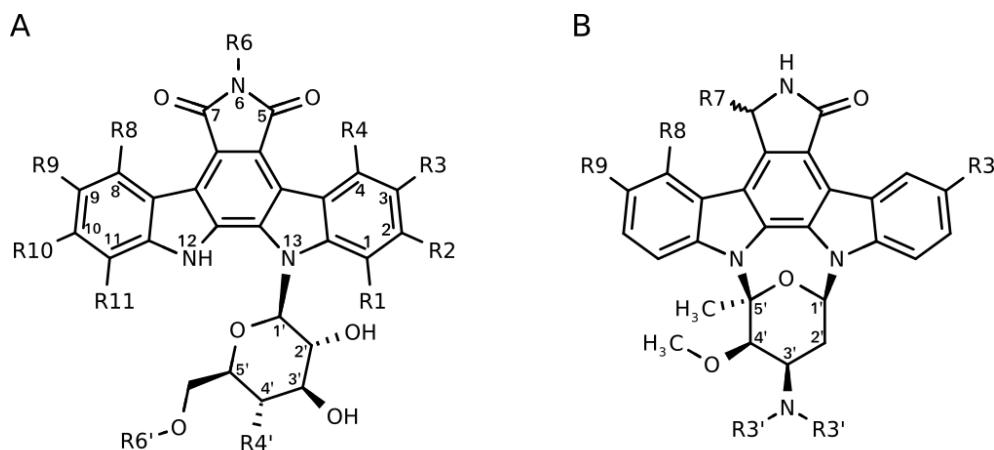


Figura 2 – Estruturas de rebecamicina e análogos (A) e de staurosporina e análogos (B).

através de mudanças nas condições de cultura, como substituição de substratos ou precursores em etapas específicas da síntese. A adição de brometo de potássio à cultura de *S. aerocolonigenes* permitiu a introdução de átomos de bromo na estrutura de rebecamicina, em substituição aos átomos de cloro de ocorrência natural [13]. No entanto, a simples adição de sais de flúor não levou à formação de derivados correspondentes. A produção de uma série de fluorindolocarbazolas foi possível através da adição de 7-fluoro-triptofano na cultura de *S. aerocolonigenes* [14]. Esforços posteriores resultaram na síntese de outros compostos fluorados com variações na porção sacarídica, além da presença do halogênio na porção aglicona dos fluorindolocarbazolas. São os compostos BMS-250749, um 4'-flúor derivado, e BMS-251873, um 6'-amino derivado, que mostraram surpreendente desempenho quanto à atividade antiproliferativa [15, 16], uma vez que os compostos com flúor apresentaram maior atividade antitumoral *in vivo* quando comparados com rebecamicina [10].

A biossíntese combinatória também é uma ferramenta empregada para a obtenção de novos derivados de indolo[2,3-*a*]carbazola. A clonagem de genes de diferentes organismos em um *cluster* funcional possibilitou a obtenção de múltiplos derivados de rebecamicina e staurosporina [17]. Além disso, estratégias envolvendo a combinação de genes em diferentes estruturas plasmidiais, *i.* uma designada para a síntese da porção aglicona e adição do resíduo de carboidrato a essa porção e *ii.* outra exclusiva para a síntese da porção sacarídica, possibilitaram expandir a variabilidade dos açúcares e das substituições ligadas a indolo[2,3-*a*]carbazola [18, 19].

O sucesso da engenharia genética na expansão da variabilidade de ICZs forneceu o composto ED-110, obtido através da adição de um resíduo de *D*-glicopirranose ao nitrogênio do indol do composto intermediário BE-13793C, formado pelo sistema indolo[2,3-*a*]carbazola e o grupo imida. A subsequente adição do substituinte NHCHO no átomo de

nitrogênio da imida originou o derivado NB-506 [20], que apresenta atividade citotóxica em células de linhagem tumoral e atividade antitumoral *in vivo* em xenoinxertos [21, 22]. A estrutura de NB-506, por sua vez, serviu de ponto de partida para etapas biossintéticas adicionais que originaram o composto edotecarin. A substituição do grupo NHCHO na imida de NB-506 pelo grupo NHCH(CH<sub>2</sub>OH)<sub>2</sub> e o deslocamento dos grupos hidroxila das posições 1 e 11 para as posições 2 e 10 [23], deu origem ao composto que apresenta intenso efeito inibitório sobre a topoisomerase I [24, 25]. O desempenho de edotecarin chegou a ser avaliado como agente antitumoral em ensaios clínicos de fase I e II, porém não foi eficaz contra tumores de diversas histogêneses e os testes foram encerrados [25, 26].

Becatecarin é um composto semi-sintético obtido pela introdução do substituinte (CH<sub>2</sub>)<sub>2</sub>NEt<sub>2</sub> no nitrogênio do grupo imida da molécula de rebecamicina [11, 27]. Becatecarin é um agente intercalante e inibidor de topoisomerasas I e II [28, 10]. Becatecarin apresentou atividade moderada em estudos clínicos de fase II. Embora benefícios clínicos tenham sido observados em pacientes com câncer biliar avançado, a toxicidade hematológica levou à redução de dose na maioria dos pacientes tratados [29] e, em geral, sua eficácia foi inferior àquela de medicamentos já existentes [30, 31]. Outro estudo, que empregou becatecarin no tratamento de câncer de pulmão de pequenas células, mostrou toxicidades gerenciáveis. Embora tenha sido descontinuado devido a dificuldades na produção do medicamento, os resultados não mostraram provável superioridade em comparação com os agentes já existentes para esta doença [32].

Mais recentemente, o derivado de rebecamicina 9-metoxi-rebecamicina foi isolado da cepa BCI2 de bactérias do gênero *Pseudonocardia*. Essas bactérias apresentaram um plasmídeo circular estável abrigando um *cluster* gênico necessário para a produção de rebecamicina, além de sequências que codificam outras enzimas, possivelmente responsáveis pela adição do grupo 9-metoxi específico para esta molécula [33].

Em relação aos análogos da staurosporina, estes são caracterizados pela presença de duas ligações *N*-glicosídicas conectando os átomos de carbono 1' e 5' da porção sacarídica com os nitrogênios indólicos da porção aglicona, que se caracteriza pela presença de um grupo amida [7] (Figura 2 B). Staurosporina foi o primeiro ICZ naturalmente isolado a partir de culturas de *Streptomyces staurosporeus* coletadas em uma amostra de solo [1]. A estrutura descrita por cristalografia de raios-X, embora inicialmente mostrasse a estereoquímica incorreta, revelou um cromóforo derivado de indolo[2,3-*a*]pirrolo[3,4-*c*]carbazona, ao qual se liga uma aminopiranosose através de duas ligações glicosídicas em ambos os nitrogênios dos grupos indol [34]. A estereoquímica correta foi estabelecida posteriormente [35] e confirmada pela síntese total [36].

Embora a staurosporina tenha apresentado atividade antimicótica no mesmo estudo em que foi isolada, a atividade antibacteriana não se confirmou [1]. Estudos posteriores mostraram ação citotóxica associada à capacidade de inibir proteína cinase C (PKC), no

entanto, com efeito inespecífico sobre diversas serino/treonino e tirosino cinases [37, 38, 39]. Mais recentemente, o composto também mostrou propriedades antiparasitárias ao induzir a morte de *Trypanosoma brucei* [40].

Um programa de triagem de inibidores seletivos de PKC, através da identificação de ligantes em diferentes isolados bacterianos do solo, mostrou que staurosporina, embora seja o produto principal, pode ser produzida juntamente com outros compostos relacionados. Assim, foram isolados os derivados 7-hidroxistaurosporina, UCN-01, um potente inibidor de PKC [41], e seu estereoisômero UCN-02 [42, 43]. UCN-01 mostrou atividade antitumoral *in vitro* e *in vivo*, bloqueando o crescimento de células de câncer colo-retal e potencializando a atividade do agente irinotecan [44]. O mesmo composto ainda mostrou atividade citoprotetora ao mediar uma parada específica e reversível do ciclo celular de células normais em modelo de câncer mamário em camundongos, diminuindo a toxicidade da quimioterapia citotóxica associada sem comprometer sua eficácia [45].

UCN-01 é um dos derivados de staurosporina mais investigados na prática clínica, seja como agente único ou em combinação com outros quimioterápicos. Ambas as estratégias mostraram respostas discretas e o manejo da toxicidade foi uma questão importante durante os estudos. O uso combinado de UCN-01 com carboplatina, em pacientes com tumores sólidos avançados, foi considerado pouco significativo [46], assim como a associação com topotecan, onde a avaliação contra câncer de ovário foi realizada em fase I e II [47, 48]. Outro estudo usando cisplatina e UCN-01 em pacientes com tumores sólidos malignos avançados precisou ser encerrado devido à alta toxicidade [49]. Por outro lado, avaliações posteriores mostraram toxicidade aceitável em combinação com irinotecan, em pacientes com tumores sólidos malignos [50] e com monofosfato de fludarabina, em pacientes com leucemia crônica [51].

Produzido em menor quantidade pela mesma via de síntese natural de staurosporina, o composto TAN-1030A também foi isolado de culturas de *Streptomyces* sp. e mostrou propriedades ativadoras de macrófagos associada a redução da progressão de células tumorais em camundongos [52]. Durante um estudo mais recente de extratos de *Streptomyces* sp. a staurosporina *N*-acetilada foi identificada. Na ocasião, a resolução da estrutura por cristalografia de raios-X, revelou que esse derivado apresenta a particularidade de ligar a porção sacarídica com a porção aglicona através de apenas uma ligação *N*-glicosídica. No mesmo estudo, a staurosporina *N*-acetilada mostrou citotoxicidade sobre linhagem de células tumorais prostáticas e potente atividade inibitória contra proteíno cinases [53].

Outras espécies, além de *Streptomyces* sp. também mostraram estar associadas à produção de derivados de staurosporina. O derivado ZHD-0501, naturalmente isolado de *Actinomadura* sp., apresenta um anel de oxazolona fundido à hexopiranosose, conferindo uma estrutura bicíclica na porção sacarídica. O mesmo estudo testou a citotoxicidade de

ZHD-0501, que inibiu a proliferação de células tumorais humanas (adenocarcinoma de pulmão, hepatocarcinoma e leucemia) e de camundongos (leucemia e câncer mamário) [54]. Alguns derivados de staurosporina, apresentando substituintes oxigenados na porção aglicona, foram isolados de tunicados marinhos *Cystodytes solitus*, no entanto, ainda não está claro se esses compostos são produzidos pelo próprio tunicado ou por microrganismos associados a essa espécie. Citotoxicidade significativa contra linhagens de carcinoma de pulmão humano, carcinoma colorretal e adenocarcinoma de mama foi observada no estudo da atividade desses produtos naturais [55].

Outro estudo, que visava a obtenção de inibidores seletivos de proteíno cinases, sintetizou o composto midostaurin (4'-*N*-benzoilstaurosporina), que foi obtido a partir da staurosporina por meio da introdução de um grupo benzoila na porção sacarídica, por meio da reação de acilação [56]. Atualmente, midostaurin é utilizado na prática clínica, empregado no tratamento de leucemia mielóide aguda e mastocitose sistêmica avançada [57, 58]. Midostaurin é um potente inibidor do fator estimulador de colônia de macrófagos tirosino cinase (FLT3, *FMS-like tyrosine kinase 3*) [59, 60] e tem atividade contra formas mutantes do receptor de tirosino cinase do proto-oncogene KIT, que leva à mastocitose sistêmica avançada [61, 62]. Por meio de uma série de colaborações entre a indústria e a academia, o composto midostaurin foi avaliado em combinação com quimioterapia padrão em um grande estudo de fase III com pacientes recém-diagnosticados com mutação em FLT3 [63, 64, 65]. Na mesma época, respostas persistentes também foram observadas em outros estudos que empregaram midostaurin no tratamento de pacientes com mastocitose sistêmica avançada [66, 67, 68]. Coletivamente, esses dados clínicos levaram à aprovação de midostaurin pela *Food and Drug Administration* (FDA) dos EUA e pela *European Medicines Agency* (EMA) para o tratamento de leucemia mielóide aguda com mutação em FLT3 e mastocitose sistêmica avançada [58].

Os compostos streptocarbazolas, também isolados de actinomicetos marinhos do gênero *Streptomyces* sp., trouxeram uma nova configuração em relação à glicosilação em ICZs. Streptocarbazolas apresentam duas ligações *N*-glicosídicas entre átomos de carbono 1' e 3' porção sacarídica e os átomos de nitrogênio dos grupos indol. As estruturas de streptocarbazola A e B, os primeiros a serem isolados, foram estabelecidas por métodos espectroscópicos associados a cálculos de mecânica quântica, que indicaram a presença de uma insaturação entre os carbonos 4' e 5' do carboidrato [69]. Streptocarbazola A mostrou citotoxicidade em cultivo de células tumorais humanas [69], enquanto streptocarbazola B tem ação antiparasitária contra amebas do gênero *Acanthamoeba* sp. [70]. Posteriormente, o composto streptocarbazola C foi isolado, não apresentando a insaturação na porção sacarídica como ocorre em streptocarbazola A e B, conforme revelou a cristalografia de raios-X. Streptocarbazola C não mostrou atividade inibitória contra PKC neste estudo preliminar [71].

## 1.1.1 Mecanismos de Ação Molecular

### 1.1.1.1 Inibidores de Cinases

As proteíno cinases constituem uma grande família de enzimas que catalisam a fosforilação de resíduos de aminoácidos em proteínas. Como resultado, geralmente ocorrem mudanças conformacionais na proteína alvo, afetando sua função. As proteíno cinases são responsáveis pelo controle de uma ampla variedade de processos de transdução de sinal dentro da célula [72]. A atividade desregulada de proteíno cinase é implicada na fisiopatologia de importantes doenças, incluindo câncer, distúrbios do sistema imunológico e distúrbios neurodegenerativos. Desde a descoberta de que staurosporina é um inibidor nanomolar de proteíno cinase [37], a busca de inibidores seletivos de cinases, com potencial terapêutico, é de grande interesse entre a comunidade envolvida no desenvolvimento de novos fármacos [7, 73].

Staurosporina é capaz de inibir proteíno cinase por competir com a molécula de trifosfato de adenosina (ATP) pelo seu sítio de ligação [37, 38]. A interação de staurosporina com o sítio de ATP, em cinases, foi confirmada pela determinação de estruturas cristalográficas do ligante complexado à cinase 2 dependente de ciclina (CDK2) [74] e à proteíno cinase dependente de cAMP (PKA) [75]. Mais tarde, foi demonstrado que o efeito inibitório de staurosporina é inespecífico e afeta a atividade de muitas serino/treonino e tirosino cinases [38, 39, 6]. Esses dados estão em boa concordância com a citotoxicidade deste composto mostrada em experimentos com diversas linhagens de células tumorais. Atualmente, staurosporina é considerado composto chave de todo o grupo indolo[2,3-*a*]carbazola [7].

Estruturas cristalográficas subsequentes confirmaram que staurosporina, bem como seus análogos, se ligam na fenda que compõe o sítio de ligação do ATP na proteíno cinase. A fração indolo[2,3-*a*]carbazola ocupa a bolsa hidrofóbica de ligação à adenina, com o grupo lactama imitando o padrão de ligação de hidrogênio da base de adenina, geralmente formando duas ligações de hidrogênio com resíduos presentes no fundo da cavidade. A porção sacarídica, na conformação de bote e perpendicular ao plano do heterociclo indolocarbazólico, forma contatos hidrofóbicos e ligações de hidrogênio no sítio de ligação à ribose do ATP [74, 75, 6]. A Figura 3 mostra staurosporina complexada à CDK2 e PKA e as ligações de hidrogênio formadas entre o ligante e os resíduos nas cinases.

Os alvos celulares de UCN-01 incluem isoformas de PKC, CDKs, proteíno de *checkpoint* cinase 1 (Chk1) e cinase 1 dependente de 3-fosfoinosítídeo (PDK1). Esses alvos provavelmente são responsáveis pelas variadas toxicidades observadas em ensaios clínicos [50]. Uma avaliação comparativa da especificidade de staurosporina e UCN-01 contra um painel de 29 cinases mostrou que, embora ambos ligantes sejam inibidores relativamente não específicos, a análise de *fingerprint* da inibição de UCN-01 apresenta um conjunto de proteíno cinases diferente quando comparado ao painel de inibição de staurosporina [76].

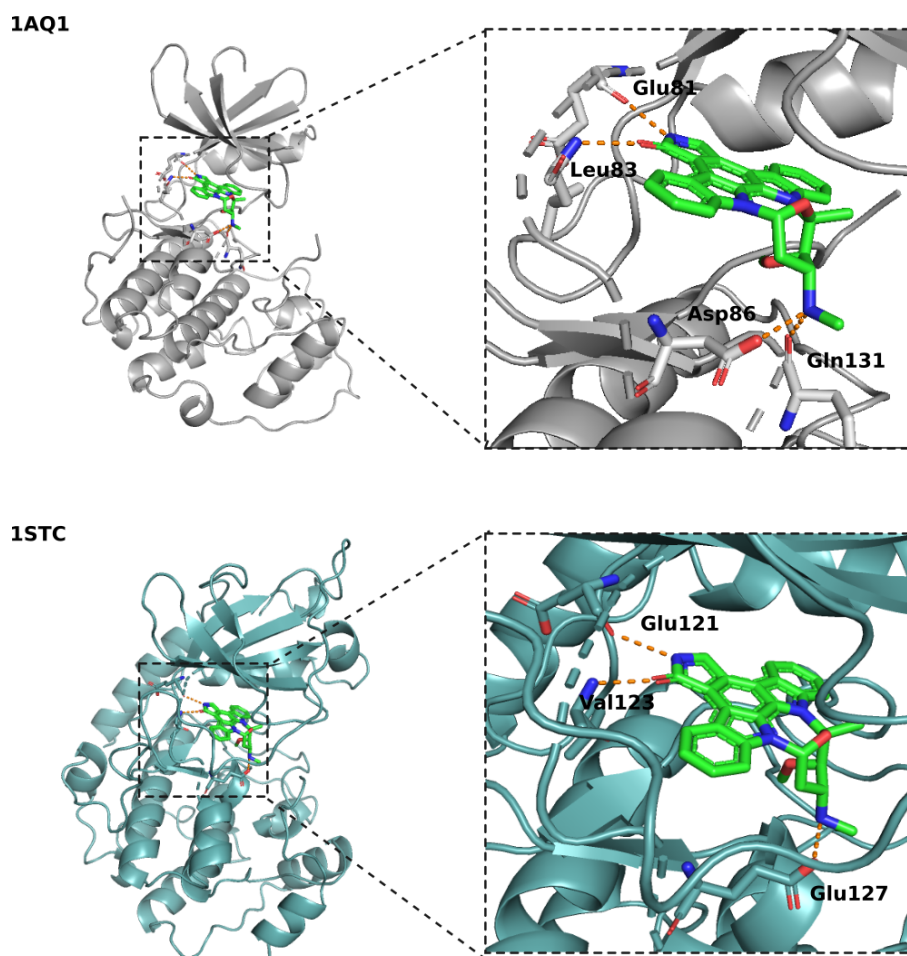


Figura 3 – Os complexos cristalográficos de staurosporina-CDK2 (*PDB ID* 1AQ1, resolução 2,00 Å) e staurosporina-PKA (*PDB ID* 1STC, resolução 2,30 Å) revelaram como o ligante combina muito bem com a forma do sítio de ligação ao ATP, o que o torna um inibidor altamente potente, mas relativamente inespecífico para proteíno cinases. Ligações de hidrogênio são representadas por linhas pontilhadas na cor magenta.

Uma análise dos complexos cristalográficos disponíveis mostrou que o análogo hidroxilado, UCN-01, ocupa a mesma posição ocupada por staurosporina, no sítio de ligação ao ATP de proteíno cinase, com geometria semelhante. Os complexos de ambos os ligantes com PDK1 mostram que as mesmas interações hidrofóbicas são estabelecidas pela porção heterocíclica e que as ligação de hidrogênio tanto do heterociclo quanto da porção sacarídica também são conservadas. As diferenças aparecem com relação às novas ligações de hidrogênio formadas pelo grupo 7-hidroxila de UCN-01, além de moléculas de água ordenadas em posições que entram em contato com esse grupo e o deslocamento da cadeia lateral de alguns resíduos da proteína em direção ao ligante, em comparação ao complexo com staurosporina. Essas mudanças resultam em mais espaço para acomodar o volumoso grupo 7-hidroxila em UCN-01 [76]. Complexada com a Chk1, UCN-01 estabelece ligações de hidrogênio semelhantes com resíduos equivalentes. Chk1 parece ter uma cavidade



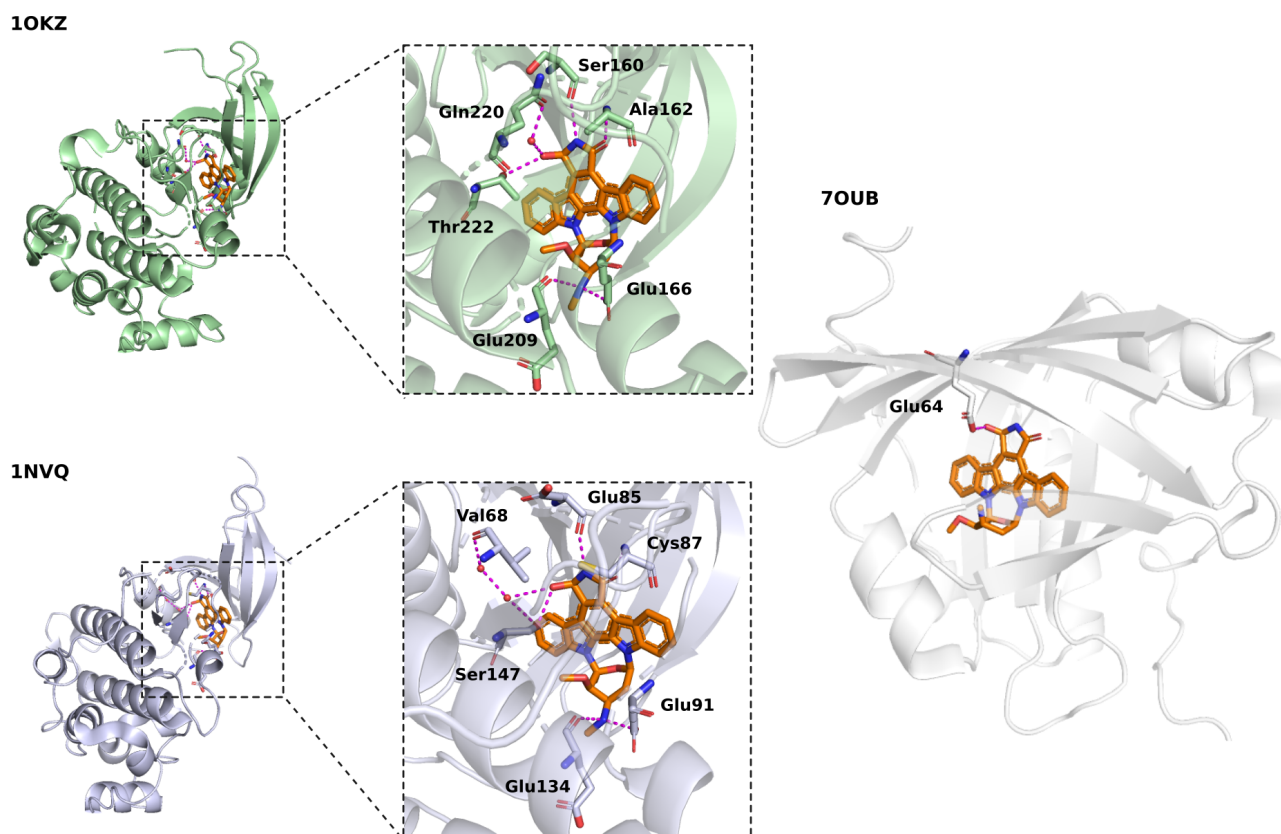


Figura 4 – Os complexos cristalográficos de UCN-01-PDK1 (*PDB ID* 10KZ, resolução 2,51 Å) e UCN-01-Chk1 (*PDB ID* 1NVQ, resolução 2,00 Å) mostram como o substituinte 7-hidroxila forma novas ligações de hidrogênio, inclusive com moléculas de água (representadas com esferas vermelhas) ordenadas a resíduos no sítio ativo das proteínas. O complexo UCN-01-AGP2 mostra a conformação em cadeira na porção sacarídica do ligante (*PDB ID* 7OUB, resolução 1,82 Å).

hidrofílica mais extensa, que permite a formação de uma rede de ligações de hidrogênio mediada por moléculas de água e envolvidas em interações com o grupo 7-hidroxila [77]. A Figura 4 mostra UCN-01 complexada à PDK1 e Chk1 e as ligações de hidrogênio formadas entre o ligante, os resíduos nas cinases e moléculas de água presentes na cavidade.

UCN-01 é reconhecido há muito tempo como potencializador do efeito de vários agentes citotóxicos [49, 47, 50, 51] e também se destaca como um agente protetor citostático [45]. No entanto, a forte afinidade de UCN-01 pela glicoproteína ácida alfa 2 (AGP2) tem um impacto significativo em sua eficácia, e essa interação é conhecida por causar um baixo volume de distribuição do fármaco [78]. Vários resíduos aromáticos revestem a bolsa de ligação de AGP2 a UCN-01, o que propicia diversas interações hidrofóbicas CH- $\pi$  e empilhamento  $\pi$  com o sistema de anéis aromáticos do ligante. O complexo apresenta apenas uma ligação de hidrogênio bem definida, envolvendo o grupo 7-hidroxila. Diferente do que ocorre nos complexos citados anteriormente, onde o carboidrato aparece na conformação de bote, quando UCN-01 complexa com AGP2, a conformação de cadeira está presente, como mostra o complexo 7OUB (resolução 1,82 Å) na Figura 4 [79].

O composto midostaurin, resultado da acilação do grupo metilamino [56] chegou a ser descrito como um inibidor seletivo de PKC, de KIT, de FLT3, do receptor do fator de crescimento endotelial vascular 2 (VEGFR2) e do receptor do fator de crescimento derivado de plaquetas  $\beta$  (PDGFR $\beta$ ) [80, 59, 60]. Esses estudos pré-clínicos apoiaram o prosseguimento de uma série de ensaios clínicos, embora a especificidade do alvo e a alta taxa de ligação a proteínas tenham sido reconhecidas como problemas [58].

A resolução da estrutura de midostaurin complexado à cinase de dupla especificidade regulada por fosforilação de tirosina 1A (DYRK1A) mostra a incomum formação de ligação dissulfeto e alterações conformacionais na proteína associadas à presença deste composto. Como outros derivados de staurosporina, midostaurin também se liga no sítio de ATP, no entanto através de apenas duas ligações de hidrogênio, mediadas pelos átomos de nitrogênio e oxigênio de grupo amida, em direção a aminoácidos presentes no fundo da cavidade. O complexo midostaurin-DYRK1A é caracterizado por diversas interações hidrofóbicas envolvendo o sistema heterocíclico e o grupo fenila (Figura 5) [81].

Em contraste com staurosporina, midostaurin não possui um grupo NH que poderia fazer uma ligação de hidrogênio estabilizadora ao carbonil de Glu291, por exemplo (como observado nos complexos da Figura 3). A substituição com o grupo benzoíla, induz um deslocamento conformacional no resíduo e coloca o grupo metila em direção ao anel fenila de midostaurin. Em média, midostaurin liga-se a proteíno cinases mais fracamente do que staurosporina [82] e a perda dessa ligação de hidrogênio provavelmente contribui para a menor afinidade na ligação [81].

Uma análise mais cuidadosa do complexo midostaurin-DYRK1A (Figura 5), permite observar que, como na maioria dos complexos anteriormente citados e disponíveis na literatura, a porção sacarídica desse derivado de staurosporina também aparece na conformação de bote (resolução 2,60 Å). Essa conformação favorece a interação hidrofóbica com o resíduo Glu291, à medida que posiciona o anel de fenila para fora da cavidade de ATP.

#### 1.1.1.2 Inibidores de Topoisomerases

As DNA topoisomerases são enzimas encontradas em todos os organismos vivos, desde arqueobactérias até humanos. Essas enzimas são agrupadas em duas classes, topoisomerases do tipo I (Top1) e topoisomerases do tipo II (Top2), e desempenham papel essencial em muitos processos celulares associados à separação de fitas de DNA, como replicação, transcrição, recombinação, remodelação da cromatina e reparo [83].

As Top1 fazem a clivagem transitória em uma única fita do duplex de DNA, com a função de aliviar sua tensão conformacional [83, 84]. Outra função de Top1 é promover a fosforilação dos fatores de *splicing* da família SR, graças a sua atividade de cinase [85, 86]. As Top2 são organizadas em dímeros e promovem a clivagem transitória das duas fitas

4NCT

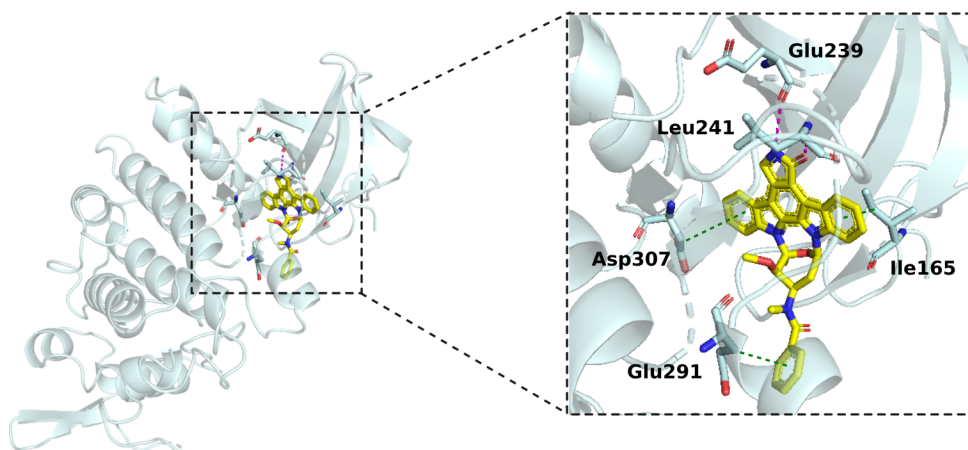


Figura 5 – O complexo cristalográfico midostaurin-DYRK1A (*PDB ID* 4NCT, resolução 2,60 Å) mostra que é marcante a influência das estruturas apolares de midostaurin na complexação à proteína alvo. Ligações de hidrogênio são representadas por linhas pontilhadas na cor magenta, enquanto as interações hidrofóbicas, na cor verde.

de DNA através de hidrólise de 2 moléculas de ATP em cada ciclo catalítico [87, 83, 88]. Os intermediários comuns da ação das topoisomerases são quebras de DNA ligadas às enzimas. Ao final dos processos que demandam o corte das fitas, as topoisomerases religam o DNA sem perda de base ou alteração na sequência. As DNA topoisomerases são alvo de diferentes classes de fármacos, com aplicação antitumoral e antibacteriana [89].

Quando o DNA é clivado por Top1, uma molécula análoga de rebecamicina estabiliza o complexo intermediário covalente DNA-Top1, impedindo a religação da fita. A ação do ligante resulta em extensa fragmentação no DNA celular, citotoxicidade celular e apoptose [9, 90]. Com essas descobertas, muitas ICZs antitumorais foram sintetizadas para atingir processos relacionados ao DNA, por meio de dois mecanismos primários: intercalação de DNA e bloqueio de Top1 [91, 92, 16]. Um terceiro mecanismo de ação atribuído a ICZs, seria a inibição direta da Top2 por staurosporina, que mostrou inibir a atividade catalítica da enzima bloqueando o sítio ativo de tirosino cinase [93].

Em contraste com a grande variedade estrutural encontrada entre os inibidores de cinase, as ICZs com significativa ligação ao DNA ou atividade de bloqueio de Top1 pertencem, quase exclusivamente, ao grupo de compostos glicosilados derivados de rebecamicina [6]. As relações estrutura-atividade foram exaustivamente estudadas e revisadas. Três domínios funcionais foram propostos para esses compostos: o domínio de interação com Top1 (a função imida no pirrol), o domínio de ligação intercalante de DNA (o sistema de anéis apolares) e o domínio de ligação ao sulco de DNA (a porção sacarídica). A inserção do cromóforo planar entre dois pares de bases consecutivos colocaria o resíduo de açúcar em um sulco da dupla hélice [94, 95, 91, 16].

1SEU

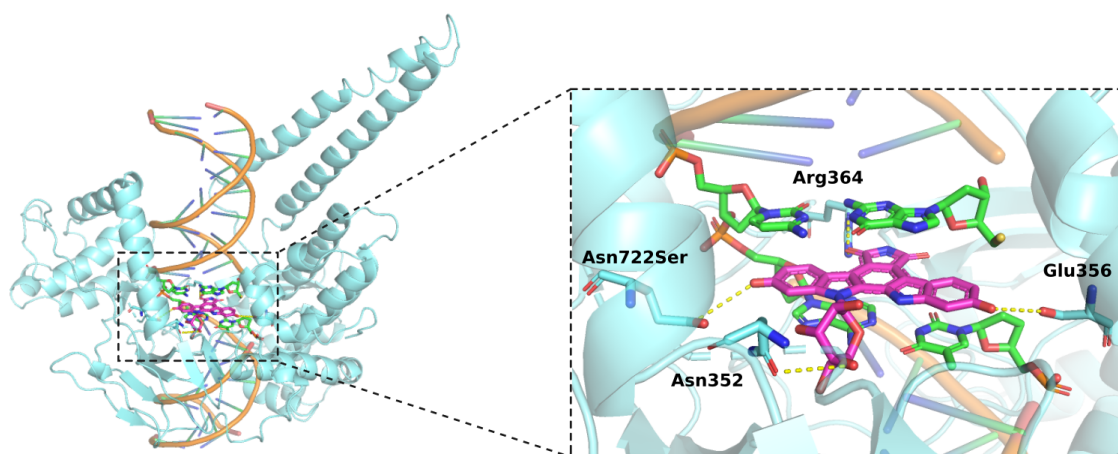


Figura 6 – O complexo cristalográfico SA315F-DNA-Top1 (*PDB ID* 1SEU, resolução 3,00 Å) revelou que o ligante (magenta) intercala entre dois pares de bases consecutivos do DNA (verde) enquanto estabelece ligações de hidrogênio com a enzima Top1 (ciano). A presença do carboidrato no sulco maior da cavidade de ligação intercalante ao DNA pode estabilizar o complexo SA315F-DNA-Top1.

A obtenção de estrutura cristalográfica do derivado de rebecamicina SA315F, ligado ao complexo covalente Top1-DNA, mostrou que o ligante SA315F intercala o DNA no local da clivagem mediada por Top1, conforme mostra a Figura 6. O grupo indol glicosilado estabelece empilhamento  $\pi$  com as bases no lado da fita íntegra do duplex de DNA, enquanto o grupo indol não glicosilado empilha-se com as bases no lado da fita clivada. As hidroxilas simétricas dos grupamentos indol fazem ligações de hidrogênio com resíduos da enzima Top1, assim como uma das carbonilas do anel maleimida, que se posiciona no sulco menor da cavidade de ligação intercalante ao DNA [96].

O resíduo de açúcar desempenha um papel essencial tanto para a interação com o DNA quanto para a inibição de Top1 [97, 94, 95]. Compostos indolocarbazólicos nos quais o resíduo de açúcar está ligado aos dois nitrogênios indólicos (como ocorre em staurosporina) geralmente mostraram interação muito fraca com o DNA e produziram efeitos mínimos em Top1 [90, 98, 99]. A remoção da porção glicosídica diminui a inibição de Top1 e a intercalação no DNA por indolocarbazóis [94]. O mapa da densidade globular mostrou a proeminência do substituinte sacarídico, que ocupa o sulco maior da cavidade de ligação ao DNA. É provável que essa porção do ligante estabilize o complexo SA315F-DNA-Top1 e contribua para a maior potência do ICZ quando comparado a outros agentes intercalantes de DNA não glicosilados, como camptotecina e indenoisoquinolina [96].

Os ICZs mais ativos na inibição de Top1 contêm substituintes na posição N6 do anel maleimida, que se posiciona no sulco menor da cavidade da ligação intercalante ao DNA adjacente. É o caso dos derivados NB-506 e edotecarin, que apresentam o grupo *N*-formil e o grupo di-hidroximetil, respectivamente, nessa posição [100, 101]. Um modelo,

do complexo ternário desses ligantes N6-substituídos com Top1-DNA, sugere que seria necessário um movimento de translação de menos de 1 Å do ligante ou um movimento conformacional de resíduos da enzima na vizinhança para melhor acomodar a complexação. O substituinte polar ainda forneceria grupos passíveis de estabelecer ligações de hidrogênio na região, contribuindo para a estabilidade do complexo [96]. A estrutura do modelo não está disponível.

Mais recentemente, a estrutura cristalográfica do intermediário de rebecamicina, 4'-*O*-demetil-1,11-didescloro-rebecamicina, complexado à uma enzima que cataliza a adição do grupo metila na posição 4' do carboidrato, foi depositada no PDB. A enzima RebM é caracterizada como metiltransferase responsável pela metilação no grupo 4'-hidroxila na porção sacarídica de derivados de rebecamicina, reação que constitui a última etapa da síntese desses compostos [102]. Com base em evidências de homologia com RebM, a enzima AtmM foi posteriormente caracterizada como uma 4'-*O*-metiltransferase capaz de metilar uma ampla variedade de ICZs *in vitro* [103]. Curiosamente, AtmM exibiu um escopo de substratos mais amplo do que RebM, o que beneficiou a produção de novos compostos a partir de sua atividade [104, 105].

Embora o trabalho que gerou o complexo 6UV6 ainda não tenha sido publicado, uma estimativa das interações intermoleculares, entre o ligante intermediário e a enzima AtmM, foram obtidas através de análises realizadas nos servidores *Protein-Ligand Interaction Profiler* (PLIP) [106] e *ProteinsPlus* [107] e mostradas na Figura 7. Nesta exploração, o ligante, ainda na forma de substrato, apresenta interações com a enzima em todas as suas porções. Interações  $\pi$  envolvem o sistema de anéis aromáticos, enquanto a porção sacarídica participa de uma rede ligações de hidrogênio que envolve os grupos 3'- e 4'-hidroxila, uma molécula de água e um resíduo de isoleucina, cuja cadeia lateral é rica em grupos CH<sub>n</sub>. As interações eletrostáticas na porção do açúcar são consistentes com a reação a ser catalizada pela enzima no grupo 4'-hidroxila, o que sugere uma boa adequação da geometria das estruturas envolvidas no complexo.

Diferentes ICZs análogos de rebecamicina, como 11-descloro-rebecamicina, NB-506 e di-Br-rebecamicina, exibiram simultaneamente uma gama de modos de ação sobre a função de DNA topoisomerases: prejudicando a atividade relaxante da tensão conformacional do DNA e potentemente inibindo a atividade cinase dessas enzimas, além de se mostrarem agentes intercalantes do DNA [86, 99, 108, 6]. O composto becatecarin, um derivado sintético de rebecamicina concebido para melhorar a sua solubilidade em água, seria um agente intercalante de DNA e um inibidor de Top1 e Top2 [27, 109]. BMS-250749 é um fluorindolocarbazola desenvolvido para atuar seletivamente sobre Top1. Em cultura celular, esse composto é cerca de 5 vezes mais potente que a camptotecina ao induzir quebras de fita simples no DNA, mediadas por Top1, com complexos de clivagem mais persistentes do que os observados no tratamento com camptotecina [10]. Ao nosso conhecimento, os

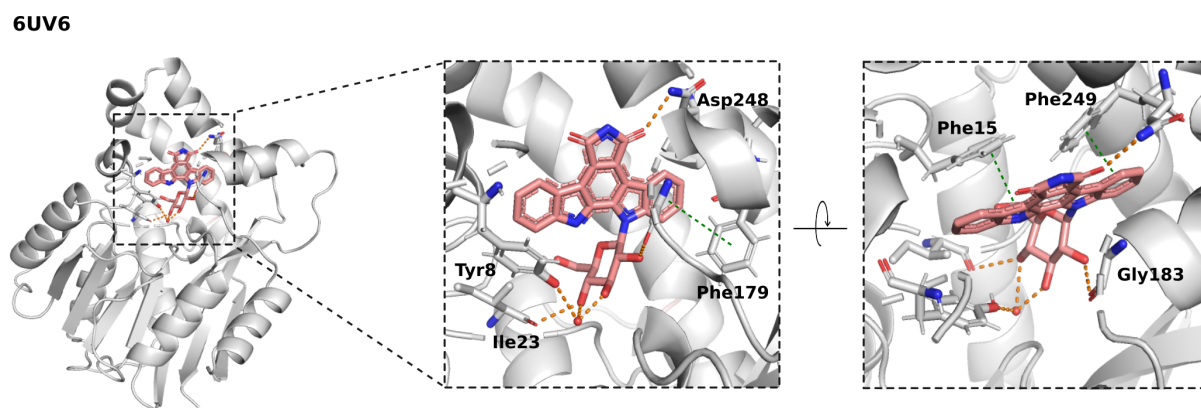


Figura 7 – O complexo cristalográfico do substrato 4'-*O*-demetil-1,11-didescloro-rebecamicina ligado à enzima AtmM (*PDB ID* 6UV6, resolução 2,72 Å) apresenta, a partir de um estudo exploratório, interações intermoleculares envolvendo todas as porções do ligante. Interações  $\pi$  (em verde) envolvem o cromóforo, enquanto a porção sacarídica participa de uma rede ligações de hidrogênio que envolve uma molécula de água (esfera em vermelho) e a cadeia lateral, rica em grupos  $\text{CH}_n$ , de um resíduo de isoleucina.

mecanismos de ação desses compostos ainda não foram elucidados em detalhes, assim como não se dispõe de estruturas que apresentem os modos de interação dos ligantes e seus alvos complexados.

## 1.2 Resolução Estrutural de Ligantes Glicosilados

Desde os primeiros passos do estudo da química, os químicos perceberam que a compreensão da estrutura das moléculas poderia levar à compreensão de muitos aspectos de suas propriedades químicas e reatividade. A ideia de estruturas tridimensionais para moléculas surgiu em meados do século XIX. A descoberta da quiralidade molecular por Louis Pasteur em 1848 foi baseada em observações experimentais de alguns cristais de moléculas orgânicas. Em 1865 foram introduzidos modelos de bolas e bastões para demonstrar conceitos estruturais, recurso que foi utilizado em 1874 para apresentar o princípio teórico do átomo de carbono tetraédrico assimétrico, quando surgiram as primeiras publicações sobre o tema [110].

A estrutura dos ligantes ICZs, analisados na presente tese, é composta por um sistema de anéis aromáticos heterocíclico, derivado de indolo[2,3-*a*]carbazola, e uma porção sacarídica, composta por anéis de hexopiranosose. O sistema de anéis insaturados constitui um cromóforo, pois contém muitos elétrons capazes de absorver energia ou luz visível na região do infravermelho ou ultravioleta e assim emitir cor quando os elétrons finalmente passam do estado excitado ao estado basal [111]. A porção sacarídica, cuja estrutura apresenta hexopiranososes com variados padrões de substituição nos compostos aqui estudados, desempenha um papel importante na estabilidade da formação dos complexos com seus

receptores alvo [74, 75, 94, 95, 97, 6]. Devido a suas características estruturais intrínsecas, como a alta flexibilidade, influenciada ainda pelo tipo de substituição presente nos anéis, os carboidratos continuam sendo os compostos menos estruturalmente caracterizados entre as principais classes de moléculas biológicas [112].

Carboidratos são muito difíceis de cristalizar e, na maioria dos casos, cristais únicos de qualidade suficiente para análise de raios-X não podem ser obtidos. Além disso, mesmo para uma minoria de cristalizações bem-sucedidas, foi relatado que a cristalografia de raios-X fornece estruturas mal resolvidas de porções glicanas [113, 114, 112]. Um levantamento realizado por nosso grupo caracterizou o estado atual de informações referentes a carboidratos no *Protein Data Bank* (PDB). A análise de estruturas de alta qualidade, mostrou que a maioria dos dados está concentrada em um pequeno conjunto de tipos de carboidratos, em suas formas monossacarídicas e conectados por uma pequena diversidade de ligações glicosídicas, sinalizando para a baixa variabilidade da disponibilidade de estruturas descritas com qualidade [115]. A espectroscopia de massas de carboidratos pode ser considerada uma técnica útil, porém não suficiente como ferramenta estrutural por si só, uma vez que a importante questão da estereoquímica de carboidratos não pode ser resolvida pelos métodos rotineiramente disponíveis [116].

Uma vantagem importante da espectroscopia de ressonância magnética nuclear (*nuclear magnetic resonance*, NMR) diz respeito à determinação da estrutura tridimensional de biomoléculas diretamente em solução aquosa, onde ocorrem os processos de relevância biológica e química. Ao contrário de muitos métodos analíticos de alto rendimento, a espectroscopia de NMR é capaz de tolerar a escassez de dados de referência e, portanto, desempenha um papel fundamental na elucidação estrutural primária de novas glicanas [117]. Por esses motivos, e devido a características próprias do princípio da espectroscopia de NMR, essa técnica se apresenta potencialmente poderosa para fornecer informações relativas às conformações dinâmicas de glicanas em solução, fornecendo informações valiosas sobre as ligações covalentes em nível atômico [116, 112].

### 1.2.1 Espectroscopia de Ressonância Magnética Nuclear

O estudo espectroscópico de moléculas fornece dados muito finos aplicáveis ao conhecimento detalhado de sua estrutura. Esta análise se propõe não apenas a determinar as posições relativas dos átomos que compõem uma molécula, mas também a intensidade das forças que os mantêm unidos e suas relações. A ressonância é o fenômeno pelo qual se produz uma vibração de frequência igual a outra vibração previamente estabelecida. No contexto da espectroscopia de NMR, a ressonância é consequência da transferência de energia entre um campo magnético radiofrequente e um núcleo sujeito a uma excitação eletromagnética constante e intensa o suficiente para desacoplá-lo de seus elétrons orbitais [118].

A espectroscopia de NMR foi desenvolvida no final da década de 1940 para estudar núcleos atômicos. Em 1951, os químicos descobriram que a espectroscopia de NMR poderia ser usada para determinar estruturas de compostos orgânicos. Esta técnica pode ser empregada apenas para estudar núcleos atômicos com número ímpar de prótons ou nêutrons (ou ambos), que ocorre nos átomos de  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{19}\text{F}$  e  $^{31}\text{P}$ , ou seja, núcleos magneticamente ativos, que possuem spin, assim como os elétrons. Como os núcleos possuem carga positiva e giram em torno de seu eixo, o efeito de spin faz com que eles se comportem como se fossem pequenos ímãs [119, 118].

Na ausência de um campo magnético, os spins nucleares se orientam aleatoriamente. No entanto, quando uma amostra é colocada em um campo magnético, os núcleos com spin positivo são orientados na mesma direção do campo, em um estado de energia mínima chamado estado de spin  $\alpha$ , enquanto os núcleos com spin negativo são orientados na direção oposta ao do campo magnético, em um estado de energia mais alto chamado estado de spin  $\beta$ . A diferença de energia entre os estados de spin  $\alpha$  e  $\beta$  depende da força do campo magnético aplicado. Quanto maior o campo magnético, maior a diferença de energia entre os dois estados de spin [118].

Quando uma amostra contendo um composto orgânico é brevemente irradiada por um forte pulso de radiação do espectro eletromagnético, cuja radiofrequência é definida, os núcleos no estado  $\alpha$  são promovidos ao estado  $\beta$ . Quando os núcleos retornam ao seu estado inicial, eles emitem sinais cuja frequência depende da diferença de energia ( $\Delta E$ ) entre os estados de spin  $\alpha$  e  $\beta$ . O termo ressonância magnética nuclear vem do fato de que os núcleos estão em ressonância com a radiofrequência. Ou seja, os núcleos passam de um estado de spin para outro em resposta à radiofrequência a qual estão sujeitos [118]. A relação entre a frequência do sinal ( $\nu$ ) e a força do campo magnético ( $H_o$ ) está definida pela equação

$$\Delta E = h\nu = h\frac{\gamma}{2\pi}H_o$$

onde  $h$  é a constante de Planck (valor aproximado de  $6,63 \times 10^{-34} \text{ J}\cdot\text{s}$ ) e  $\gamma$  representa o constante giromagnética, que depende do tipo de núcleo irradiado (no caso de  $^1\text{H}$  é igual a  $2,67 \times 10^8 \text{ rad}\cdot\text{T}^{-1}\cdot\text{s}^{-1}$  [119]).

Ao avaliar moléculas por espectroscopia de NMR, é realizada uma sequência de pulsos eletromagnéticos excitatórios capazes de manipular os spins de uma forma complexa e revelar influências externas sob um núcleo, como a proximidade e a ligação de outros átomos. A representação gráfica da distribuição da intensidade da radiação eletromagnética em função da sua frequência é registrada de forma gráfica em um espectro de NMR. As informações estruturais são interpretadas a partir da análise de características como a



largura, a intensidade e o deslocamento químico do sinal de cada núcleo registrado no espectro de NMR [118, 120].

A nuvem eletrônica dos átomos presentes nas vizinhanças de um núcleo é capaz de gerar campos magnéticos secundários que são os principais responsáveis pela alteração do deslocamento químico do núcleo, registrado ao longo do eixo de frequência de um espectro de NMR. A influência na magnetização de um núcleo causada pela ligação com outros átomos é conhecida por acoplamento escalar ou spin-spin, sendo representada pela constante de acoplamento  ${}^n J_{a,b}$ , onde  $n$  é o número de ligações covalentes separando os núcleos  $a$  e  $b$ . Normalmente, o acoplamento escalar se estabelece entre átomos separados por até três ligações químicas [120].

Um fator importante para a determinação da estrutura de moléculas por NMR é que as constantes de acoplamento  ${}^3 J_{a,b}$  dependem do ângulo de torção entre os átomos  $a$  e  $b$  acoplados. As informações relativas aos ângulos torcionais são bastante úteis na determinação da estrutura tridimensional de uma molécula. Assim, a espectroscopia por NMR em solução constitui uma das poucas técnicas experimentais capazes de fornecer informações da estrutura tridimensional de uma molécula com resolução atômica [120].

Os biólogos estruturais de proteínas obtêm grandes benefícios desta técnica que, aliada ao aprimoramento de equipamentos, *softwares* e sofisticadas metodologias de preparação de amostras, determinaram com sucesso as coordenadas atômicas desse tipo de biomolécula. A espectroscopia de NMR de carboidratos, no entanto, ainda apresenta alguns aspectos críticos. O número considerável de graus de liberdade, disponíveis para os movimentos internos das cadeias sacarídicas, dificulta as abordagens tradicionais para sua caracterização conformacional. Além disso, é frequente a sobreposição de sinais no espectro, devido à baixa variabilidade local das estruturas químicas. A resolução do espectro ainda sofre com o baixo índice de restrições conformacionais, em comparação com as estruturas proteicas [121, 122, 123].

Desta forma, devido à flexibilidade inerente dos carboidratos em solução, essa classe de biomoléculas é frequentemente melhor descrita como conjuntos de distintos confôrmeros. Por isso, os dados estruturais determinados por espectroscopia de NMR nem sempre são representativos do real comportamento de uma molécula, mesmo estando ela em solução.

#### 1.2.1.1 Constantes de Acoplamento Vicinal ${}^3 J_{H,H}$

Através de um espectro de NMR é possível observar sinais específicos de diferentes núcleos em diferentes ambientes químicos e ainda ligados a diferentes átomos. Assim, podem-se distinguir diversos grupos de átomos de hidrogênio através do seu perfil de deslocamento químico, que constitui um dos parâmetros mais importantes em estudos por NMR [120]. A equação de Karplus descreve a correlação entre as constantes de acoplamento  ${}^3 J_{H,H}$  e os ângulos de torção diedrais na espectroscopia de NMR, através da relação

$${}^3J_{H,H} = A\cos^2\phi + B\cos\phi + C$$

onde  ${}^3J_{H,H}$  define a constante de acoplamento;  $\phi$  é o ângulo diedral entre os átomos de hidrogênio, também conhecido como o ângulo de valência projetado, o ângulo de torção ou a torção; A, B e C são parâmetros derivados empiricamente cujos valores dependem da eletronegatividade, do comprimento e do ângulo de ligação entre os átomos envolvidos na substituição, embora as duas últimas propriedades moleculares tenham provado ter efeitos menores. Na relação  ${}^3J_{H,H}$ , o algarismo "3" sobrescrito indica que um átomo de  ${}^1\text{H}$  está acoplado a outro átomo de  ${}^1\text{H}$  a três ligações de distância. Em função das ligações H-C-C-H, esses hidrogênios ligados a átomos de carbono vizinhos são denominados vicinais [124, 125, 126].

A dependência de  ${}^3J_{H,H}$  na orientação da eletronegatividade dos substituintes foi abordada por Haasnoot, Altona e colegas em 1979, através da publicação de uma série de trabalhos em que os valores de  ${}^3J_{H,H}$  de etanos substituídos e outras moléculas foram investigadas detalhadamente [127, 128, 129]. Esses estudos, auxiliados pela sofisticação crescente da instrumentação de NMR, refletiram diferenças nas observações experimentais associadas à orientação de substituintes eletronegativos, indicando aumento da magnitude do acoplamento quando a orientação *gauche* estava presente, enquanto uma diminuição foi observada quando a orientação *trans* a substitui. A partir destas avaiiações, Haasnoot-Altona propuseram uma nova relação para a descrição do acoplamento  ${}^3J_{H,H}$ , dependente da orientação dos substituintes eletronegativos e cuidadosamente cobrindo todos os valores possíveis de  $\phi$  [130].

Após o entendimento de que  ${}^3J_{H,H}$  depende tanto da eletronegatividade e da orientação dos grupos substituintes, quanto da disponibilidade de um conjunto de constantes de aditividade (valores de  $\Delta J$ ), Haasnoot-Altona desenvolveram uma regra de aditividade simples para predição de valores de  ${}^3J_{H,H}$  em anéis de piranose. As constantes de aditividade foram derivadas de dados de  ${}^1\text{H}$  NMR de alta qualidade e podem ser definidas por  $\Delta J(x)$  para substituintes  $x$  em anéis de piranose. A partir de então a equação pôde ser reescrita da seguinte forma:

$$j_o^{-1} \cdot {}^3J_{H,H} = 13,86\cos^2\phi - 0,81\cos\phi + \sum_{i=1}^I \Delta\chi_i [0,56 - 2,32\cos^2(\xi_i\phi_o^{-1} + 17,9|\Delta\chi_i|)]$$

na qual  $j_o=1$  Hz,  $\phi_o=1^\circ$ ,  $\phi$  é o ângulo diedral entre os átomos de hidrogênio acoplados,  $I$  corresponde ao número de substituintes diferentes de hidrogênio do fragmento H-C-C-H,  $\Delta\chi_i = \chi_i - \chi_H$  é a diferença de eletronegatividade entre um átomo  $i$  e hidrogênio, assumindo efeitos aditivos dos substituintes ligados aos carbonos do fragmento H-C-C-H e dependentes da eletronegatividade relativa de Huggins [131]. Por fim, o parâmetro  $\xi_i$  assume o valor +1 ou -1, dependendo da orientação do substituinte eletronegativo, sendo negativo na conformação onde os átomos de hidrogênio acoplados são eclipsados [130].

Uma série de reparameterizações da equação de Karplus já foram propostas, no entanto a equação de Haasnoot-Altona é amplamente aceita e conhecida como equação de Karplus generalizada [132].

Essa relação entre a geometria local e a constante de acoplamento é de grande valor no âmbito da espectroscopia de NMR e é particularmente valiosa para determinar os ângulos de torção diedral entre átomos, e assim, descrever o perfil conformacional de moléculas em solução. A disponibilidade de dados experimentais das constantes de acoplamento  ${}^3J_{H,H}$  auxilia na validação dos modelos propostos para o estudo da dinâmica de ICZs em solução aquosa, além de inferir sobre a acurácia da descrição conformacional dessas moléculas. Estratégias semelhantes já foram empregadas em outros trabalhos do nosso grupo [133, 134].

## 1.3 Dinâmica Molecular

Simulações computadorizadas são amplamente utilizadas para obter informações sobre processos moleculares dinâmicos em uma escala de resolução próxima à atômica. Os métodos de dinâmica molecular (*molecular dynamics*, MD) podem ser usados para modelar a evolução temporal de sistemas biomoleculares compostos por proteínas, lipídios, ácidos nucleicos e carboidratos para estudar processos dinâmicos como dobramento de proteínas ou para estimar propriedades termodinâmicas, como energias livres de ligação de ligantes ao seu alvo. O principal desafio dessas simulações é justamente descrever as propriedades do sistema em termos das interações entre os átomos [135]. O nível de resolução atômico deve-se ao tamanho e complexidade dos sistemas biomoleculares e das escalas de tempo que devem ser alcançadas para modelar os processos de interesse. Para esse tipo de análises, o uso da mecânica clássica, também chamada de mecânica molecular (MM), em conjunto com um *campo de força* é normalmente o método de escolha, à medida que torna o custo computacional acessível e permite o acesso das propriedades analisadas, sem todavia, grande comprometimento da qualidade dos dados. Nessa abordagem, cada átomo em uma molécula é representado por um ponto no espaço com massa, carga parcial e parâmetros de van der Waals (vdW). Assim, os elétrons não são considerados explicitamente e assume-se que o sistema está no estado eletrônico fundamental. Nesse caso, a dinâmica do sistema pode ser descrita pelas equações de movimento de Newton [136].

Nas simulações de MD, as interações entre os átomos ou partículas do sistema são definidas por meio de uma função de energia potencial das coordenadas atômicas, também chamada de campo de força. Essa função de energia é empírica quando parametrizada, a partir de dados experimentais e/ou de cálculos de mecânica quântica de alto nível, para reproduzir uma gama de propriedades estruturais, energéticas ou termodinâmicas associadas a um conjunto de compostos modelo [135, 137]. Além da escolha do campo

de força mais apropriado ao tipo de análise que se deseja realizar, há outros aspectos importantes no emprego simulações por MD.

A água é o principal solvente de biomoléculas e uma das vantagens de estudos baseados em MD é a possibilidade de incluir a presença de moléculas de água nos sistemas modelados. Os modelos de água podem ser divididos em dois grupos: nos modelos explícitos, a estrutura da molécula de água, contendo os átomos, é incluída na simulação, enquanto nos modelos implícitos (ou contínuos) as moléculas são representadas através de propriedades dielétricas do solvente. Descrever eficientemente as interações da água com as moléculas analisadas em função do tempo, depende da escolha do modelo de água e das características da caixa de simulação (formato e volume de espaço onde ocorrerá a simulação) [120, 138, 139].

Em simulações por MD, a grande maioria dos átomos que compõem um sistema constitui o solvente. Por isso, a otimização no número de moléculas de água pode representar um ganho de economia no tempo de máquina para realizar uma simulação. O tamanho e a forma da caixa, usualmente centralizada no soluto, definirá a quantidade de solvente a ser inserida. Os formatos de caixas, mais frequentemente empregados estão relacionadas a uma estratégia denominada condições periódicas de contorno, capaz de fornecer continuidade no ambiente solúvel no qual se encontra o sistema. Réplicas periódicas da caixa, em todas as direções, mantêm a molécula em contato com o solvente, enquanto conservam constantes as condições da simulação e o número de moléculas no sistema [120, 138, 139]. A Figura 8A mostra uma representação das condições periódicas de contorno.

A evolução temporal da simulação é registrada em uma trajetória do comportamento molecular. Essa trajetória é constituída de pequenas etapas não contínuas, cujo tamanho é chamado de tempo de integração. Em uma escala de femtossegundos (fs), a definição do tempo de integração está diretamente associada ao tamanho da amostragem da simulação e, assim, ao custo computacional da mesma. Tempos de integração muito pequenos descrevem eventos lentamente e em grandes detalhes, mas demandam maior custo computacional. Por outro lado, tempos maiores são menos custosos, mas podem dar origem a instabilidades na simulação, de forma que o uso de valores intermediários são geralmente a alternativa de escolha [120, 138, 139]. A Figura 8B mostra uma representação do efeito do tempo de integração em três magnitudes.

### 1.3.1 Campos de Força

No âmbito da química e da modelagem molecular, um campo de força corresponde a um método computacional que engloba uma coleção de equações aliadas a um conjunto de parâmetros que permitem calcular a energia potencial de um sistema de átomos ou moléculas. A partir de uma abordagem Hamiltoniana, é possível reproduzir a geometria molecular e propriedades selecionadas de estruturas químicas em sistemas com intercon-

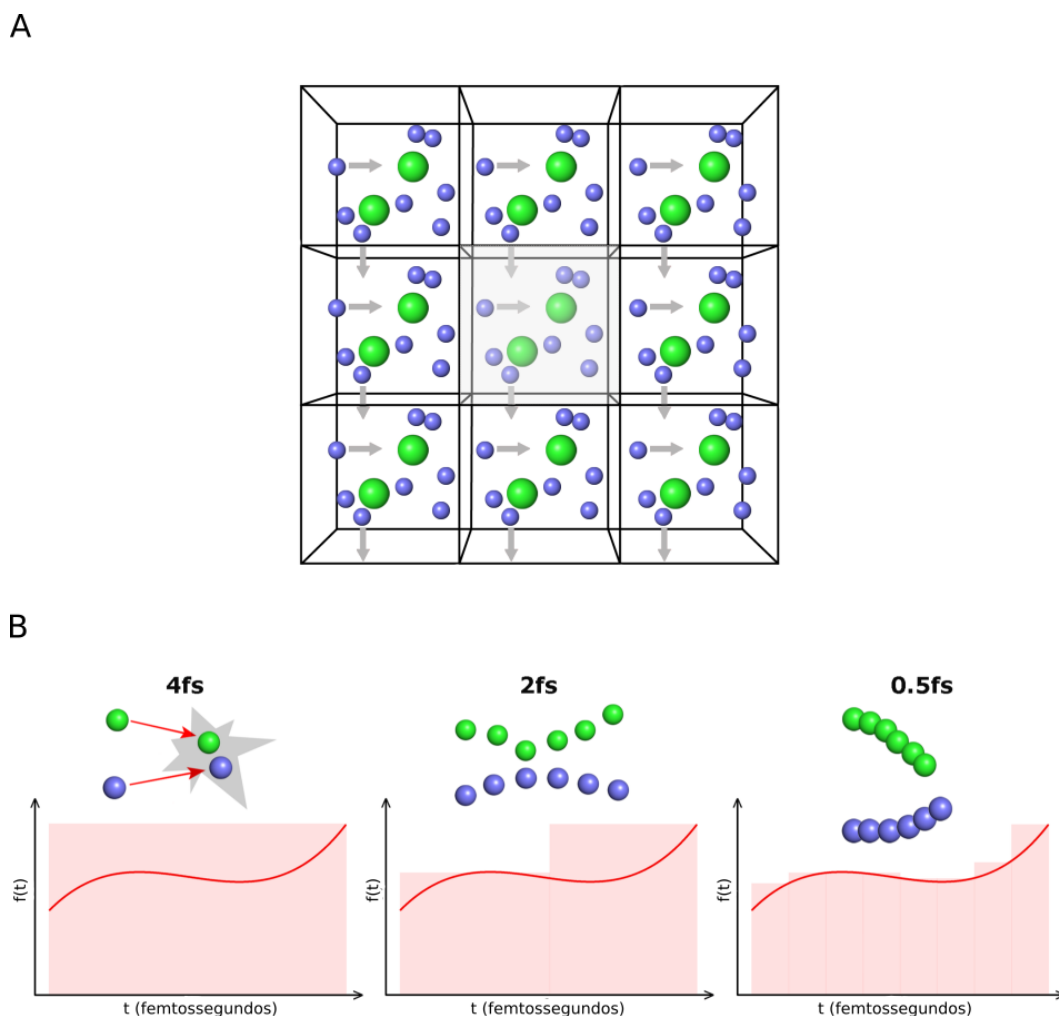


Figura 8 – (A) Representação das condições periódicas de contorno em uma simulação por dinâmica molecular, onde a caixa central é circundada por réplicas que garantem a continuidade do sistema. (B) Representação do efeito de diferentes tempos de integração. Valores mais elevados (4,0 fs) apresentam menores custos computacionais, embora possam gerar instabilidades. Valores muito pequenos (0,5 fs) descrevem eventos com mais detalhes, embora mais lentamente. Valores intermediários (2,0 fs) podem descrever adequadamente o evento em uma escala de tempo acessível. Adaptado de Verli, 2014.

versão entre as energias potencial e cinética durante a evolução temporal, como ocorre em simulações de dinâmica molecular [140].

Simulações de dinâmica molecular clássica fazem uso das leis do movimento de Newton para gerar uma trajetória, que descreve as posições e velocidades de um conjunto de moléculas ou átomos em função do tempo. Cada átomo em uma molécula é representado por um ponto no espaço constituído de massa, carga parcial e parâmetros de vdW. Os elétrons não são considerados explicitamente e as cargas parciais atômicas não variam, assume-se portanto que o sistema está no estado eletrônico fundamental. As forças que governam os movimentos atômicos são derivadas das interações átomo-átomo obtidas a partir das funções de energia potencial que compõem o campo de força [140, 141, 136].

Os campos de força clássicos são constituídos de centenas de termos, cuidadosamente escolhidos para melhor reproduzir as características do sistema. Os parâmetros adicionados aos termos que compõem uma função de energia podem ser derivados de experimentos físico-químicos, cálculos de mecânica quântica *ab initio* ou a partir da combinação de ambos. Assim, a função de energia potencial do campo de força tem caráter empírico, sendo parametrizada através do ajuste de um conjunto de propriedades selecionadas com base em dados experimentais para um grupo de sistemas moleculares modelo [140, 142]. O campo de força é dito efetivo, quando os parâmetros são ajustados empiricamente a dados experimentais e/ou dados de cálculos de mecânica quântica (QM) de nível superior [136].

Os campos de força clássicos utilizam essencialmente a mesma expressão empírica proposta por Levitt e Lifson em 1969 para aproximar a energia potencial total de um sistema [143]:

$$E_{pot}(\vec{r}) = E_{bond}(\vec{r}) + E_{angle}(\vec{r}) + E_{torsion}(\vec{r}) + E_{improper}(\vec{r}) + E_{elec}(\vec{r}) + E_{vdW}(\vec{r})$$

onde  $E_{pot}$  é a função de energia potencial do sistema e  $\vec{r}$  corresponde às coordenadas de todos os átomos do sistema. Os termos ligados ou covalentes descrevem o estiramento das ligação químicas ( $E_{bond}$ ), a deformação dos ângulos de ligação ( $E_{angle}$ ), a rotação dos ângulos diedrais próprios ( $E_{torsion}$ ) e a distorção dos ângulo diedrais impróprios ( $E_{improper}$ ) em moléculas. Os termos não ligados  $E_{elec}$  e  $E_{vdW}$  descrevem, respectivamente, as interações eletrostáticas e o componente de vdW entre átomos não ligados [136]. A Figura 9 mostra as funções associadas a cada um desses termos, com as respectivas representações gráficas dos mesmos.

As interações eletrostáticas têm sido geralmente representadas por interações de Coulomb entre cargas atômicas pontuais [144, 145], enquanto o componente de vdW, composto por um termo de repulsão e um termo de atração, é definido pela função de Lennard-Jones [146]. A mesma estratégia de aproximação da energia potencial tem sido usada por décadas, com maior incremento em termos de descrição dos potenciais

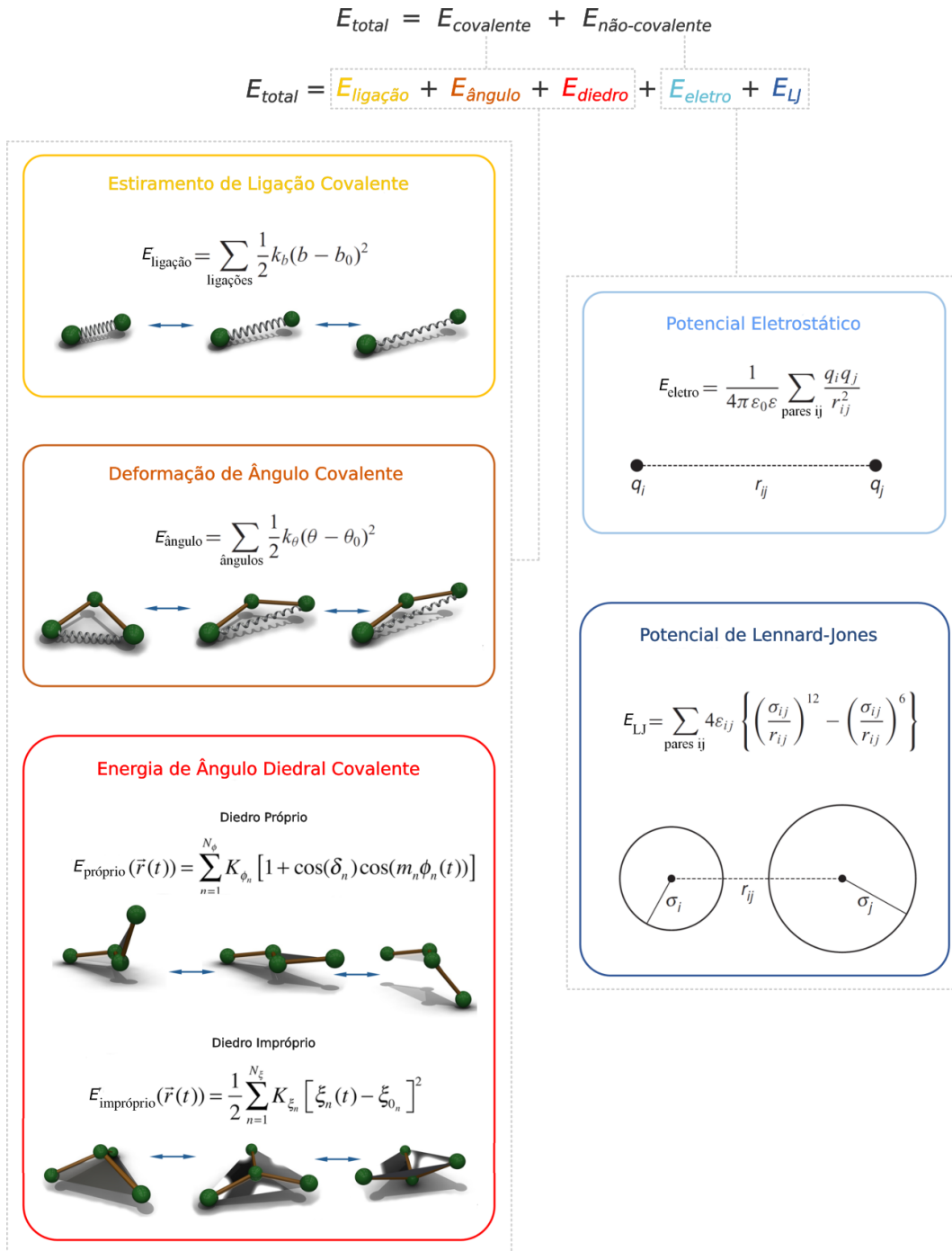


Figura 9 – Representaç\~{a}o dos termos que comp\~{o}em um campo de for\~{c}a e descrevem a energia potencial total de um sistema. Adaptado de Verli, 2014.

de Coulomb [147]. A partir do ano 2000, os campos de for\~{c}a polariz\~{a}veis, ou seja, que incorporaram a polariza\~{c}\~{a}o expl\~{i}cita dos \xc3\~{a}tomos tornaram-se mais frequentes [148], mas os campos de for\~{c}a cl\~{a}ssicos baseados em interaç\~{o}es Lennard-Jones e cargas pontuais ainda s\~{a}o dominantes [110].

Compartilhando a mesma forma funcional mostrada na equaç\~{a}o acima, os campos de for\~{c}a cl\~{a}ssicos AMBER [149, 150], OPLS [151, 152], CHARMM [153, 154] e GROMOS

[155, 156] foram amplamente aprimorados, testados e aplicados em uma infinidade de estudos de simulações moleculares em fase condensada, durante as últimas décadas [157, 110]. Quando o alvo desses estudos são biomoléculas como proteínas, ácidos nucleicos, lipídios e carboidratos, cuja característica estrutural é serem constituídas de um número relativamente pequeno de blocos de construção, o processo de parametrização e validação dos termos do campo de força é de certa forma simplificado. Em se tratando de outras moléculas orgânicas, no entanto, o cenário é diferente [136]. A citar, pequenos ligantes moleculares - com ação terapêutica ou candidatos a fármacos - apresentam um vasto espaço químico para tamanho e diversidade estrutural, o suficiente para dificultar a elaboração de blocos de construção e exigir parametrização individual de cada ligante, ou no mínimo, família de ligantes.

Uma abordagem comum para compor a função de energia em simulações moleculares de pequenas moléculas orgânicas, é o emprego de parâmetros pré-estabelecidos para termos ligados e não ligados em um campo de força, definidos a partir de uma estratégia de similaridade química e estrutural. Mesmo assim, cálculos de mecânica quântica (semiempírico ou *ab initio*) são necessários para gerar as cargas parciais das nova moléculas [157]. Exemplos de campos de força para moléculas orgânicas, que adotam essa filosofia, são o campo de força geral Amber (*General Amber Force Field*, GAFF) [158], o campo de força de todos os átomos OPLS (*OPLS All-Atom*, OPLS-AA) [159, 160], o campo de força geral CHARMM (*CHARMM General Force Field*, CGenFF) [161, 162] e o gerador automático de topologias do GROMOS (*Automated Topology Builder*, ATB) [135, 163, 164].

Campos de força ainda mais genéricos, que dispensam qualquer parametrização individual, são o campo de força molecular Merck (*Molecular Merck Force Field*, desenvolvido em 1994, MMFF94) [165] e o campo de força universal (*Universal Force Field*, UFF) [166]. As funções de energia centrais do MMFF94 são derivadas de cálculos de mecânica quântica e os parâmetros são atribuídos de acordo com resultados ajustados a dados experimentais ou de computação quântica de maior nível [167, 168, 169]. Embora os parâmetros possam ser classificados nos sete termos de energia que ocorrem nos demais campos de força anteriormente citados, esses campos de força empregam funções de energia potencial mais complexas e, portanto, não são compatíveis com a maioria dos campos de força biomoleculares [157].

### 1.3.2 Construção de Topologias de Pequenos Ligantes

Embora os protocolos listados acima sejam amplamente empregados na geração automática de topologias e simulações de pequenos ligantes, obter parâmetros confiáveis e compatíveis com esses campos de força para o estudo de classes heterogêneas de compostos como substratos, inibidores, cofatores e compostos bioativos, continua sendo um grande desafio [164]. Permanece em aberto também a questão em relação ao grau de confiabilidade



e precisão das interações produzidas e observadas através do uso desses parâmetros [135].

Os campos de força clássicos de carga fixa são campos de força efetivos, ou seja, se caracterizam pelo ajuste empírico dos parâmetros a dados provenientes de cálculos de QM de alto nível e/ou a dados experimentais, quando estes estiverem disponíveis. No entanto, a disponibilidade limitada de dados experimentais, somada ao número potencial de graus de liberdade de um modelo, seja em relação a sua flexibilidade, interação com o solvente ou mesmo aos efeitos eletrônicos omitidos, torna o processo de validação de parâmetros de campo de força altamente desafiador [164, 170].

Enquanto o campo de força MMFF94 envolve extensa parametrização para uma ampla gama de grupos funcionais orgânicos, ele foi, entretanto, desenvolvido com o objetivo de "definir um campo de força que descreva com precisão as propriedades moleculares em fase gasosa e que se comporte adequadamente quando o mesmo sistema é incorporado em fase condensada"[165]. Uma das potenciais fraquezas nessa abordagem, no entanto, é a omissão de dados de fase condensada, tanto nas etapas de parametrização quanto na testagem [110]. Mesmo com o extenso conjunto de bases de métodos de QM, a abordagem do MMFF94 se mostrou subestimada [171] e reafirmou a importância de incluir uma ampla gama de dados experimentais, amostrando propriedades intramoleculares e interações intermoleculares, incluindo fase condensada na derivação de parâmetros ou testagem do campos de força [110].

O processo de parametrização do campo de força GROMOS baseia-se na reprodução do comportamento molecular em fase condensada e de seus dados termodinâmicos. Na versão GROMOS53A6, os critérios usados para avaliar a precisão dos parâmetros derivados incluíram a densidade e o calor de vaporização [155, 172]. Dentro da filosofia do GROMOS, o servidor ATB é uma alternativa para a geração automática de parâmetros para pequenos ligantes compatível com o desenvolvimento de estudos no campo de força GROMOS. Embora o ATB forneça parâmetros passíveis de refinamento, com base em seu próprio banco de dados e dados de QM, a produção de parâmetros com base em similaridade nem sempre contempla o potencial de graus de liberdade de um grupo heterogêneo de moléculas, suas cargas atômicas e seus potenciais diedrais. Assim, o campo de força GROMOS não possui um conjunto de parâmetros oficial para o tratamento de estruturas heterogêneas. Avanços neste sentido têm sido propostos pelo nosso grupo e resultado em contribuições com parâmetros para o estudo de pequenos compostos. O desenvolvimento de parâmetros para melhor descrever a flexibilidade de carboidratos [173, 174] e a calibração de propriedades físico-químicas de anéis aromáticos comumente usados em química medicinal [175] já auxiliaram no estudo de outros grupos de compostos bioativos como chalconas, flavonóides [134] e ácidos siálicos (*em preparação*).

Na concepção do campo de força GROMOS, as moléculas são constituídas de uma coleção de grupos funcionais, o que permite alocar uma gama limitada de parâmetros,

aumentar a transferibilidade e mitigar a necessidade de sobreajustes [164]. Esse conceito de grupo empregado pelo GROMOS permite uma transferência versátil de frações químicas entre moléculas, ou seja, parâmetros conhecidos de grupos funcionais em comum podem ser aplicados a diferentes moléculas, mesmo em presença de outras diferenças estruturais entre elas. Assim, é possível construir a topologia de uma molécula, de forma consistente com a derivação do campo de força GROMOS, juntando os blocos de construção existentes [172, 136, 176]. A simetria na atribuição dos parâmetros, ou seja, garantir que átomos equivalentes recebam os mesmos parâmetros, independentemente da geometria, também deve ser consideradas quando as topologias forem empregadas em simulações onde as moléculas podem mudar de conformação [135].

Para os campos de força clássicos de carga fixa, apenas um conjunto de cargas atômicas parciais é usado. O ajuste para reprodução da superfície de Potencial Eletrostático Restrito (RESP), por ser um ajuste algorítmico de cargas pontuais centradas no átomo, possibilita a transferibilidade de cargas entre grupos funcionais comuns e é adequado para o uso em dinâmica molecular clássica [177]. O GROMOS usa grupos de carga, ou seja, grupos de aproximadamente cinco átomos cujo somatório de carga total é um número inteiro. As cargas parciais do GROMOS53A6 foram ajustadas para reproduzir a densidade, o calor de vaporização e as energias livres de solvatação em água e em alguns solventes orgânicos [155, 136].

No ATB, a determinação de parâmetros para descrição de ângulos diedrais é baseada na combinação dos tipos de átomos que os definem. Em casos ambíguos, várias opções são listadas a partir do banco de dados. Caso nenhuma correspondência adequada seja encontrada, um parâmetro diedral padrão com a menor constante de força e coeficientes correspondentes, derivados da arquitetura da região, é escolhido [135]. A parametrização de ângulos diedrais torcionáveis, no entanto, é menos trivial. O estudo da flexibilidade molecular de espécies heterogêneas, como é o caso dos pequenos ligantes, requer calibração acurada dos potenciais associados a ligações, considerando suas especificidades químicas e estruturais. Na década de 1990, os métodos de QM tornaram-se suficientemente rápidos para que os perfis de energia de torção pudessem ser calculados em maior escala. Embora esses cálculos nem sempre sejam realizados em fase condensada, mais recentemente, têm sido propostas estratégias de validação e ajuste a dados experimentais, como acoplamentos vicinais  $^3J$  provenientes de espectroscopia de NMR [136].

## 1.4 Desenho de Fármacos Auxiliado por Computador

Inovações em química combinatória contribuíram massivamente com o aumento de bancos de dados de compostos, permitiram a cobertura de grandes espaços químicos e, por fim, impulsionaram a expansão da descoberta de medicamentos e o desenvolvimento

de métodos de triagem de alto rendimento (*high-throughput screening*, HTS) [178, 179]. No entanto, apesar desses avanços, o número de novas entidades moleculares lançadas com sucesso no mercado não cresceu na mesma proporção, em grande parte devido ao desempenho inadequado em diferentes fases dos ensaios clínicos [180].

O processo de desenvolvimento de medicamentos, desde a descoberta até o lançamento no mercado leva, em média, de 10 a 15 anos e apresenta um custo aproximado de US\$ 800 milhões a US\$ 1,8 bilhão [181]. Essa tarefa árdua, demorada e cara, requer colaborações de muitas especialidades, incluindo química medicinal, química computacional, biologia e pesquisa clínica [182]. As metodologias de desenho de fármacos auxiliado por computador (*computer-aided drug design*, CADD) têm se apresentado como uma ferramenta essencial na descoberta de medicamentos, sendo base para o emprego de novas abordagens e tornando esse processo mais rápido e menos custoso financeiramente [183, 184]. Esses métodos são relevantes para diminuir o uso de modelos animais em pesquisas farmacológicas, para auxiliar o projeto racional de novos e seguros candidatos a medicamentos e para reposicionar medicamentos já comercializados, apoiando a trajetória da promoção de novas moléculas [185]. A Figura 10 apresenta algumas abordagens do CADD empregadas nas fases iniciais do processo de desenvolvimento de novos fármacos, antes que os ligantes candidatos sejam testados *in vitro* e *in vivo*.

Embora o HTS continue representando uma grande parte do processo de descoberta de medicamentos na indústria farmacêutica devido à sua alta taxa de sucesso, a falta de compreensão primária do mecanismo molecular envolvido na atividade dos compostos *hits* identificados pode dificultar a busca de candidatos promissores [186, 187]. Na fase inicial do processo de desenvolvimento de novos fármacos, é comum que exista pouca ou nenhuma informação sobre o alvo, os ligantes ou suas estruturas. As técnicas de CADD são capazes de fornecer essas informações, identificando quais proteínas podem ser alvo na patogênese e quais são os possíveis ligantes ativos que podem inibir essas proteínas [184]. Através do emprego de informações estruturais, seja do receptor alvo ou de um grupo de potenciais ligantes, em simulações computacionais, o CADD pode excluir compostos com propriedades indesejáveis e selecionar candidatos com mais chances de sucesso, fornecer informações valiosas sobre detalhes da interação responsável pela formação do complexo, bem como a afinidade de ligação [188, 189].

Dentre as estratégias de CADD, a utilização de simulações de dinâmica molecular pode contribuir com o entendimento da formação de complexos proteína-ligante [190]. O detalhamento da dinâmica conformacional de um ligante na complexação ao seu receptor-alvo é capaz de fornecer informações valiosas a respeito do modo de ligação. Oportunamente, o emprego de métodos de amostragem ampliada como a metadinâmica, através do emprego de vieses sobre determinadas variáveis coletivas, pode auxiliar na observação de eventos raros e/ou lentos, calcular energia livre e estimar parâmetros cinéticos associados à formação

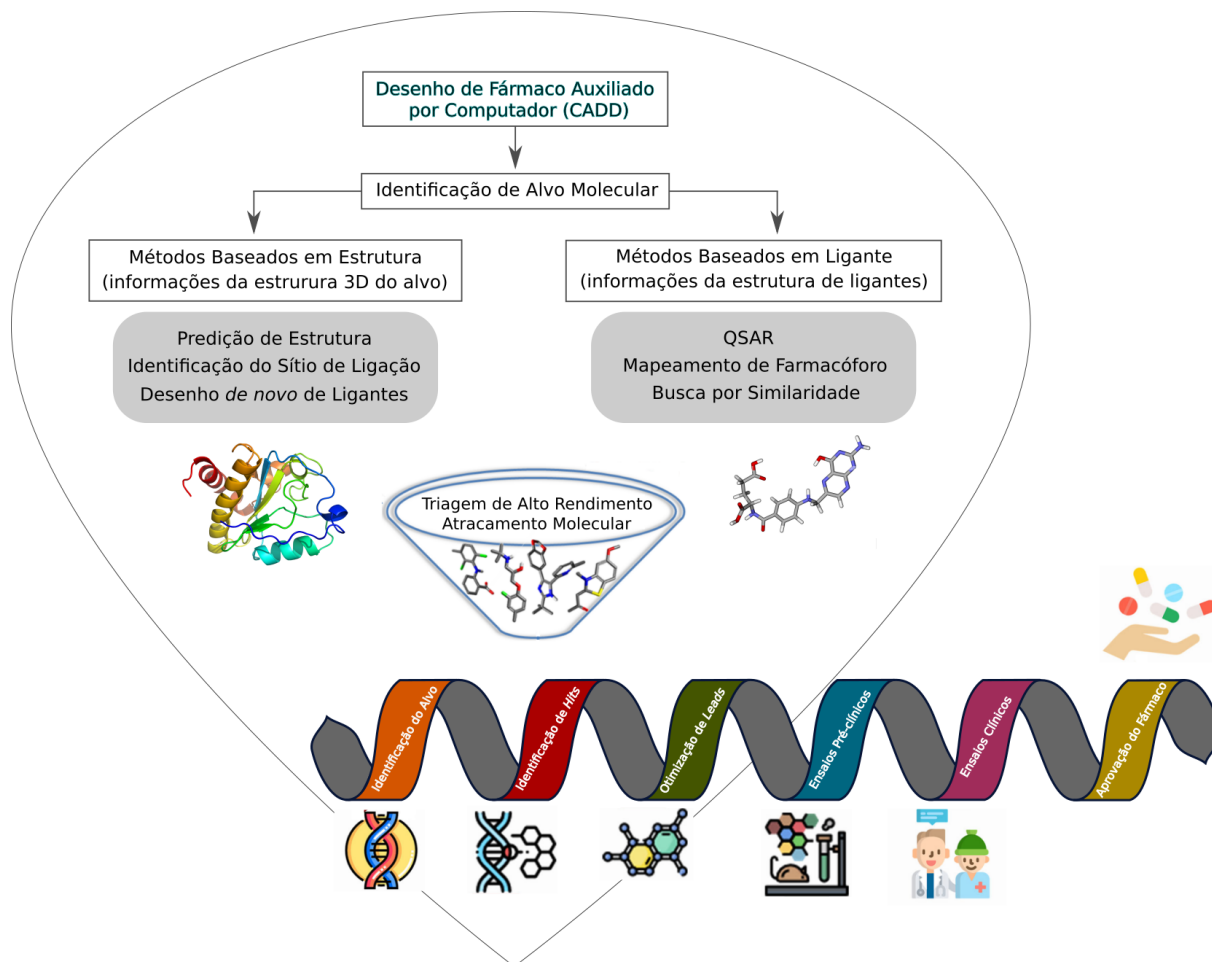


Figura 10 – Estratégias do desenho de fármacos auxiliado por computador (CADD) empregadas no desenvolvimento de fármacos. Os métodos de CADD são classificados em métodos baseados em estrutura e métodos baseados em ligantes. A abordagem depende da disponibilidade de informações estruturais do alvo e/ou de ligantes. Os acertos (*hits*) são identificados, filtrados e otimizados (*leads*) para obter potenciais candidatos a fármacos, antes que sejam testados em ensaios pré-clínicos. QSAR é relação estrutura-atividade quantitativa.

do complexo receptor-ligante [191, 192].

O crescente aumento da capacidade computacional e do desenvolvimento de campos de força com funções de energia mais acuradas é uma realidade [193]. Essa evolução vem ao encontro da necessidade da parametrização sistemática, usando propriedades em fase condensada no processo de calibração, para a melhor descrição de grupos específicos de moléculas, como é o caso de pequenos ligantes e produtos naturais. Métodos analíticos relacionados à dinâmica molecular, como o estudo da flexibilidade de moléculas bioativas em solução e a estimativa de medidas de interação com o solvente, podem se beneficiar de parâmetros acurados na reprodução do comportamento molecular e, conseqüentemente, contribuir no desenho racional de novos fármacos.

## 2 Justificativa

ICZs são produtos naturais produzidos por diversos organismos vivos, que apresentam propriedades biológicas associadas à inibição de proteíno cinases e topoisomerases. Embora essa família de ligantes compreenda um conjunto de moléculas de ampla variabilidade química e estrutural, pouco se conhece sobre diferenças no padrão de substituição química determinantes de diferenças no processo de complexação. Na verdade, esses ligantes estão associados à inibição não-seletiva de cinases e a diferentes graus de potência na parada do ciclo celular mediada pela inibição de topoisomerases. Os resultados de estudos pré-clínicos apoiaram o empenho de uma série de ensaios clínicos, no entanto, a falta de especificidade do alvo e a alta taxa de ligação a proteínas plasmáticas representaram problemas no seu seguimento. Como resultado, a persistência de efeitos adversos e toxicidade ainda têm impedido o avanço clínico de diversos análogos.

O uso de informações estruturais provenientes de complexos cristalográficos e de espectroscopia de NMR de ICZs forneceu informações úteis sobre a interação com receptores alvo e estimativas conformacionais dos ligantes. No entanto, dificuldades técnicas na cristalização da porção sacarídica desses ligantes podem levar a diferenças associadas à resolução da qualidade de estrutura, além da desvantagem de os dados se referirem ao estado sólido. Por outro lado, as informações estruturais provenientes de medidas realizadas em solução, por espectroscopia de NMR, fornecem médias conformacionais de uma população de conformeros sujeitos a variações conformacionais.

O processo de obtenção de parâmetros do campo de força GROMOS é baseado na reprodução do comportamento molecular em fase condensada e de seus dados termodinâmicos, sendo assim, uma abordagem promissora para reproduzir *in silico* dados obtidos empiricamente. Portanto, através de simulações de MD no campo de força GROMOS, é possível reproduzir, em fase condensada, as propriedades físico-químicas de sistemas moleculares. A descrição acurada das características termodinâmicas de ligantes ICZs é possível após o emprego de processos de parametrização, específicos para essa família de ligantes, baseados em cálculos quânticos e calibrados para dados experimentais de espectroscopia de RMN, tudo de acordo com a filosofia do GROMOS. Informações obtidas a partir dessas simulações, como características de interação com o solvente e da flexibilidade molecular, podem auxiliar no desenho racional e otimização de ligantes ICZs, uma vez que esses fatores são diretamente implicados no desempenho da complexação a alvos terapêuticos.



## 3 Objetivos

### 3.1 Objetivo Geral

Avaliar as características de flexibilidade e de interação com o solvente aquoso de ligantes indolocarbazólicos livres em solução através de simulações de dinâmica molecular. Tendo por base as características estruturais e funcionais dos compostos, o trabalho foi dividido em dois capítulos:

### 3.2 Objetivos Específicos

- Capítulo I - Indolocarbazolas inibidoras de cinases: staurosporina e derivados.
  - Parametrizar e validar topologias baseadas no campo de força GROMOS através de dados de espectroscopia de NMR.
  - Obter o conjunto conformacional de staurosporina e seus derivados em solução aquosa.
  - Observar, em nível atômico, características de interação com a água implicadas na complexação ao receptor alvo.
  
- Capítulo II - Indolocarbazolas inibidoras de topoisomerasas: rebecamicina e derivados.
  - Desenvolver parâmetros específicos para a descrição, no campo de força GROMOS, de aspectos estruturais envolvidos nos mecanismos de ação.
  - Observar a dinâmica das mudanças conformacionais envolvidas em processos bioativos.
  - Avaliar como as variações estruturais estão envolvidas em mudanças na solvatação e inferir sobre a farmacodinâmica das moléculas.





## 4 Procedimentos Metodológicos

### 4.1 Construção das Topologias

As geometrias das estruturas moleculares de cada composto ICZ foram retiradas do PDB, quando disponíveis, ou construídas usando o editor molecular Avogadro [194] com base em dados estereoquímicos consultados na literatura. Essas estruturas tridimensionais, serviram como referência para a construção da topologia de cada molécula, utilizando os parâmetros ligados e não-ligados dos campos de força GROMOS53A6 [155] e GROMOS53A6<sub>GLYC</sub> [156, 173]. Parâmetros específicos para compostos ICZs, compatíveis com o campo de força GROMOS, foram derivados de QM. As cargas atômicas parciais foram obtidas pelo método *least-square fit solution* [134] e novos potenciais diedrais foram calculados com o objetivo de garantir maior acurácia na descrição das preferências torcionais e conformacionais desses compostos. Um fluxograma da derivação dos novos parâmetros, com base em cálculos quânticos, é apresentado na Figura 11.

Os princípios de transferabilidade e simetria foram observados durante a construção das topologias. A construção em forma de blocos permitiu a transferência versátil de parâmetros para grupos funcionais semelhantes entre as moléculas [172]. A simetria da molécula foi observada para assegurar os mesmos parâmetros em átomos equivalentes, mesmo em geometrias diferentes [135].

#### 4.1.1 Derivação de Cargas Atômicas Parciais

As cargas atômicas parciais foram calculadas por métodos quânticos, através do Gaussian [195], utilizando a teoria de perturbação de segunda ordem de Møller-Plesset (MP2) [196] na base 6-31G\* [197], com solvatação implícita pelo Modelo do Contínuo Polarizável (PCM) [198]. O ajuste para reprodução da superfície de Potencial Eletrostático Restrito (RESP) possibilitou a transferência de cargas parciais entre grupos funcionais comuns entre as diferentes moléculas estudadas, por ser um ajuste algorítmico de cargas pontuais centradas no átomo, adequado para o uso em mecânica, dinâmica molecular e cálculos de energia livre em sistemas biológicos [177].

O vetor momento de dipolo do sistema de anéis aromáticos do cromóforo dos ICZs, calculado por métodos quânticos, foi utilizado como referência para derivação das cargas atômicas na mecânica clássica, assim como grupos de carga já parametrizados em trabalhos anteriores e calibrados para a reprodução de propriedades experimentais [175]. Para a derivação das cargas atômicas em regiões moleculares com diferentes padrões de substituição, foram construídos fragmentos moleculares de idêntica estrutura química

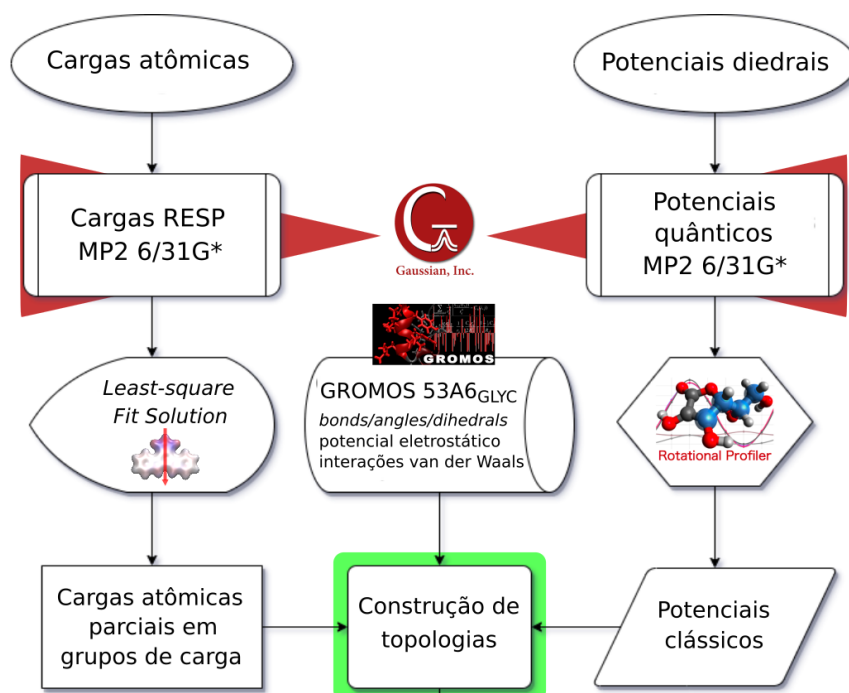


Figura 11 – A construção de topologias iniciou com o cálculo das cargas atômicas e dos potenciais diedrais através de método quântico. As cargas atômicas parciais foram ajustadas para a mecânica clássica através do protocolo *least-square fit solution*, enquanto os potenciais diedrais foram submetidos ao servidor *Rotational Profiler*. Outros parâmetros topológicos foram retirados do campo de força GROMOS53A6<sub>Glyc</sub>.

constituídos por pelo menos um anel, que tiveram as cargas atômicas calculadas pelo mesmo método quântico.

As cargas atômicas destes fragmentos foram ajustadas para a MM através do método *least-square fit solution*, utilizando limites de variação superiores e inferiores para as cargas de cada átomo a serem ajustadas e cargas fixas de átomos pertencentes ao cromóforo ou grupos de carga anteriormente parametrizados como citado acima. Um esquema representativo desta metodologia é apresentado na Figura 12. A solução do algoritmo retorna cargas atômicas parciais que, em conjunto, reproduzem a direção e o sentido do vetor momento de dipolo dos fragmentos obtido por QM [175, 134]. A definição dos grupos de carga, em anéis e seus substituintes, permitiu a transposição das cargas parciais desses grupos de átomos entre moléculas, com o mesmo padrão de substituição de grupos funcionais, de acordo com o princípio de transferabilidade do campo de força GROMOS [172].

#### 4.1.2 Novos Potenciais Torcionais

Novos parâmetros específicos para a descrição do perfil rotacional de ângulos diedrais foram derivados da QM para a construção das topologias de cada ICZ estudado.

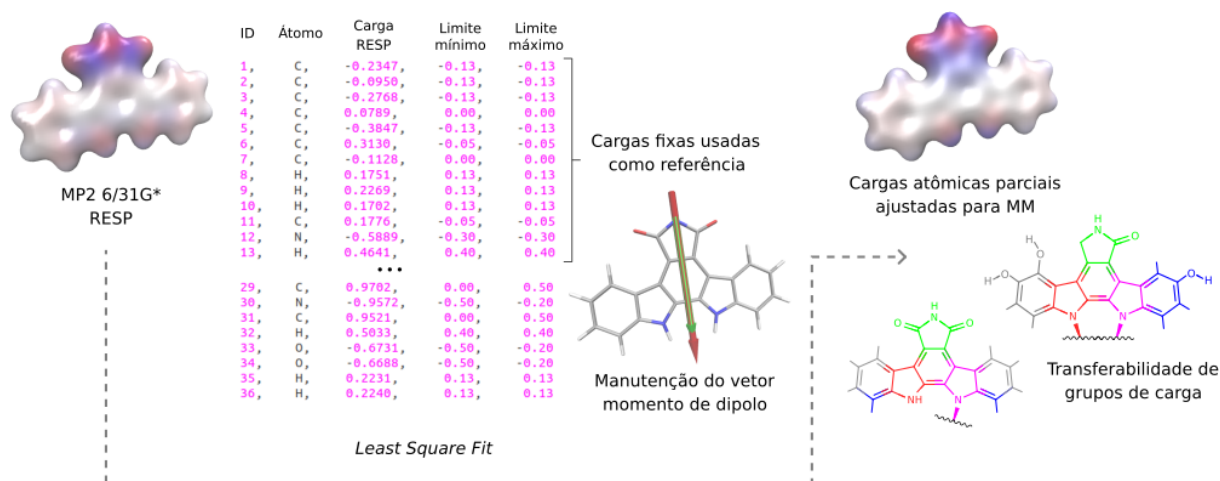


Figura 12 – Fluxo de geração de cargas atômicas parciais. As cargas foram calculadas por método quântico e ajustadas para a mecânica clássica, preservando o vetor momento de dipolo e utilizando limites inferiores e superiores. As novas cargas foram distribuídas em grupos de carga e o princípio de transferabilidade foi empregado na construção das topologias.

Os novos potenciais diedrais foram calculados na QM e ajustados para melhor serem reproduzidos na MM.

Os perfis rotacionais quânticos de cada ângulo diedral foram calculados usando Gaussian [195], utilizando o método MP2/6-31G\* [196, 197], após otimização da energia total de cada fragmento molecular, adotando o critério de convergência *tight*. A rotina *scan* foi utilizada para calcular a energia total em conformações diédricas a cada 30° e indicou a localização dos mínimos e a magnitude das barreiras energéticas a serem transpostas para o giro das ligações que definem cada ângulo diedral.

Para o cálculo do perfil rotacional na MM, as mesmas conformações diedrais empregadas no cálculo quântico foram mantidas fixas durante a minimização de energia utilizando o algoritmo *conjugate gradient* do GROMACS e o campo de força GROMOS53A6, adotando o critério de convergência de 1,0 kJ/mol [175]. O programa *gmx energy* calculou os mínimos e as barreiras energéticas associadas a cada conformação.

Para obter o *fit* dos potenciais da MM aos potenciais quânticos, ambos os perfis calculados foram submetidos ao servidor *on-line* Rotational Profile [199] que, através de regressão linear de mínimos quadrados, calculou os novos parâmetros para melhor reproduzir os potenciais quânticos em simulações clássicas de MD.

## 4.2 Dinâmica Molecular

A MD é um método de simulação no qual o movimento das partículas do sistema é calculado ao longo de um determinado período de tempo, permitindo que a evolução do

sistema seja investigada. As equações de movimento de Newton são normalmente usadas para descrever as trajetórias das partículas no sistema e a energia potencial é calculada com base nos parâmetros do campo de força.[200]

A etapa de simulação da dinâmica molecular de cada um dos compostos em solução biológica foi realizada através do pacote de simulação GROMACS 5.1, empregando os campos de força GROMOS53A6 [155] e GROMOS53A6<sub>GLYC</sub> [156, 173] para a descrição dos parâmetros físico-químicos das moléculas estudadas. Os sistemas de simulação foram montados em caixa dodecaédrica, onde uma única molécula de cada composto foi colocada e solvatada com o modelo explícito de água *extended simple point charge* (SPC/E) [198]. O tempo das simulações para o estudo das características conformacionais e de interação com o solvente foi definido em 1  $\mu$ s, com etapas de integração de 2 fs e condições periódicas de contorno. A minimização da energia do sistema foi alcançada por algoritmo de máximo declive. As ligações covalentes foram fixados pelo método de Lincs [201] e distâncias mínimas entre os átomos das moléculas e as bordas da caixa de simulação foram restritas a 1,0 nm. As interações eletrostáticas foram tratadas pelo algoritmo particle-mesh Ewald (PME) [202], enquanto os contatos de vdW com o *cutoff-scheme* de Verlet [203] para truncar todas as interações não ligadas em 1,2 nm [204].

### 4.3 Metadinâmica

Em simulações de dinâmica molecular convencionais, nem sempre é possível que o sistema visite todas as conformações energeticamente relevantes para seu estudo. As conformações relevantes podem estar separadas por altas barreiras de energia livre e para popular tais estados conformacionais, suplantando os custos energéticos da transição de um estado metaestável para outro, muitas vezes é necessária a ação de flutuações raras que impulsionam o sistema ao longo da energia livre. Outro fator que limita essa convergência do sistema é o longo tempo necessário para a difusão no espaço da configuração. Nesses casos, a presença de eventos lentos associados às características em estudo requerem uma quantidade muitas vezes indisponível de tempo computacional [205].

Para ampliar a amostragem em simulações de dinâmica molecular, e reproduzir a superfície de energia livre (*free-energy surface*, FES), a metadinâmica é uma ferramenta poderosa. Através da introdução de um potencial de *bias* (viés de força) em função de um selecionado número de graus de liberdade, frequentemente referido como variáveis coletivas (*collective variables*, CVs), é possível visitar tais eventos raros. Na metadinâmica, o potencial de *bias* é adicionado ao operador Hamiltoniano, que age sobre as CVs de uma forma dependente da evolução da simulação. Este potencial, correspondente à soma de Gaussianas depositadas ao longo da trajetória do sistema no espaço das CVs, é capaz de desencorajar o sistema a revisitar regiões que já foram amostradas à medida que estimula

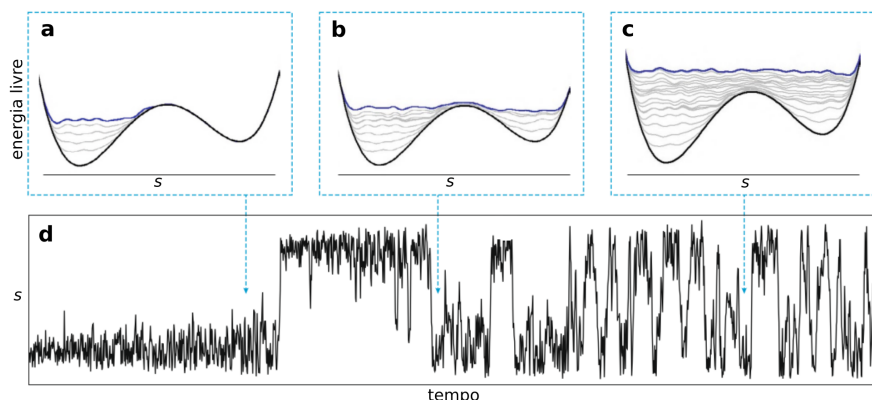


Figura 13 – Representação da energia livre de uma variável coletiva (CV)  $s$  durante uma simulação de metadinâmica. As linhas azuis representam a soma da energia livre depositada através das Gaussianas nos momentos **a**, **b** e **c** da simulação. Essas Gaussianas induzem flutuações cada vez maiores na CV (**d**), preenchendo o primeiro mínimo e em seguida o segundo mínimo, até que a CV começa a se difundir livremente entre os dois mínimos. Adaptado de Bussi e Laio, 2020.

a transposição de outras barreiras e permite a construção de uma efetiva FES em função das CVs escolhidas [192, 205, 206]. Uma representação da deposição das Gaussianas e da flutuação de energia livre a elas associada durante uma simulação é apresentada na Figura 13.

Para o estudo desenvolvido no contexto desta tese, a escolha das CVs teve como alvo a descrição de estados conformacionais dos compostos ICZs. Para observar os graus de liberdade da rotação da ligação glicosídica que une a porção sacarídica ao sistema de anéis do cromóforo, foi definido o ângulo diedral  $\phi$ , enquanto para observar a dinâmica do *puckering* de carboidratos ligados a esses compostos, foram escolhidas as coordenadas  $\theta$  e  $\phi$  definidas por Cremer–Pople [207].

A biblioteca PLUMED, versão 2.5 [208] foi interfaceada ao GROMACS para a execução de cálculos de metadinâmica *well-tempered* durante 300 ns usando um fator de *bias* igual a 10. Para calcular a torção das ligações glicosídicas, através da coordenada angular  $\phi$ , as dimensões das Gaussianas foram restritas a largura ( $\sigma$ ) de 0,35 rad, o intervalo de deposição foi de 500 fs com altura inicial ( $h$ ) de 1,2 kJ/mol, que corresponde a uma taxa de deposição ( $\omega$ ) de 2,4 J/mol.fs. Para o estudo do *puckering* de carboidratos, definiu-se  $\sigma = 0,1$  rad e  $h = 0,5$  kJ/mol, com  $\omega = 1,0$  J/mol.fs. A taxa de erro relacionada às CVs foi calculada por análise de média de bloco (*block-averaging*), abrangendo até 3% do valor da barreira. As imagens representativas do FES, que indica os estados conformacionais de hexopiranoses de maior estabilidade, foram construídas usando o utilitário *sum hills* e o pacote gráfico Gnuplot.



## 5 Resultados

### 5.1 Capítulo I: Aspectos Estruturais de Staurosporina e Análogos Implicados na Complexação a Receptores-Alvo

#### 5.1.1 Preâmbulo

Staurosporina e seus análogos são ICZs com propriedades de inibição não-seletiva de proteíno cinases que resultam em ação antimicrobiana e citostática. Em função do caráter indiscriminado da inibição de cinases, os efeitos adversos e a toxicidade ainda associados ao uso desses compostos têm impedido o avanço clínico de fármacos derivados desses análogos.

Através de simulações de dinâmica molecular no campo de força GROMOS, é possível reproduzir, em fase condensada, propriedades físico-químicas de sistemas moleculares. A descrição acurada das características termodinâmicas de staurosporina e seus análogos é possível após processos de parametrização baseados em cálculos quânticos e calibração para dados experimentais de espectroscopia de NMR para esses compostos, de acordo com a filosofia do GROMOS. Assim, avaliamos características moleculares desses ligantes livres em solução aquosa, como flexibilidade e padrões de interação com a água. O conhecimento dessas características moleculares, relacionadas ao processo de complexação com os receptores alvo, é substancial para o desenvolvimento racional de compostos bioativos com propriedades terapêuticas aprimoradas. A coexistência de múltiplos confôrmeros observada durante as simulações, e confirmada por cálculos de metadinâmica do *puckering* da fração sacarídica, expandiu a percepção do espaço conformacional de staurosporina e seus análogos. A fácil interconversão entre os confôrmeros, associada à baixa energia livre, sugere que o modelo de seleção conformacional do ligante é um mecanismo favorável no processo de complexação. Mudanças na disponibilidade para interações com o solvente aquoso, medidas através de ligação de hidrogênio e da distribuição do *bulk* de água, de acordo com diferenças nos substituintes dos ligantes mostrou efeitos aditivos na formação de ligações de hidrogênio com o solvente, que podem refletir em efeitos na energia livre de complexação devido à variação nos custos energéticos de dessolvatação. Além disso, mudanças na formação da rede de solvatação impactam na disponibilidade de moléculas de água de coordenação, que expandem o potencial de afinidade do ligante a seu receptor.

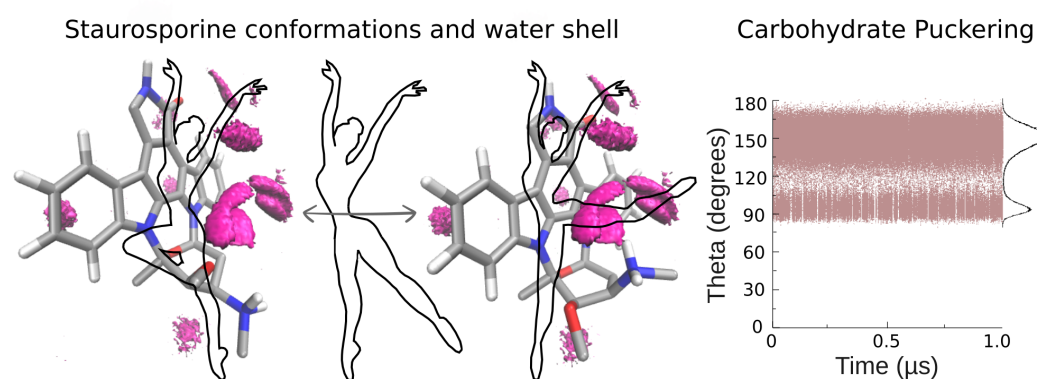
Os resultados referentes ao Capítulo I são apresentados a seguir, através do artigo científico *Theoretical models of staurosporine and analogs uncover detailed structural information in biological solution*, publicado no *Journal of Molecular Graphics and Modeling*.



## Theoretical models of staurosporine and analogs uncover detailed structural information in biological solution

Crisiele Fontana, João Luiz de Meirelles, Hugo Verli

### Graphical Abstract



### Highlights

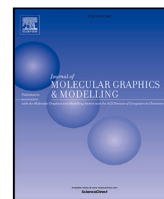
- Molecular dynamics showed coexistence of multiple staurosporine analogs conformers
- Spectroscopy data were reproduced using a tool based on Haasnoot-Altona equation
- Structural variations affects the intermolecular interaction with the water bulk
- Indolocarbazole parameterization provides framework to rational chemical modification





Contents lists available at ScienceDirect

## Journal of Molecular Graphics and Modelling

journal homepage: [www.elsevier.com/locate/jmglm](http://www.elsevier.com/locate/jmglm)

## Theoretical models of staurosporine and analogs uncover detailed structural information in biological solution

Crisciele Fontana, João Luiz de Meirelles, Hugo Verli\*

Universidade Federal do Rio Grande do Sul, Centro de Biotecnologia, Av. Bento Gonçalves, 9500 (Caixa Postal 15005), Porto Alegre, CEP 91501-970, RS, Brazil

## ARTICLE INFO

## Keywords:

Conformation  
 Indolocarbazole  
 Intermolecular interaction  
 Molecular dynamics  
 Parameterization  
 Water shell

## ABSTRACT

Staurosporine and its analogs (STA-analogs) are indolocarbazoles (ICZs) compounds able to inhibit kinase proteins in a non-specific way, while present antimicrobial and cytostatic properties. The knowledge of molecular features associated to the complexation, including the ligand shape in solution and thermodynamics of complexation, is substantial to the development of new bioactive ICZs with improved therapeutic properties. In this context, the empirical approach of GROMOS force field is able to accurately reproduce condensed phase physicochemical properties of molecular systems after parameterization. Hence, through parameterization under GROMOS force field and molecular simulations, we assessed STA-analogs dynamics in aqueous solution, as well as its interaction with water to probe conformational and structural features involved in complexation to therapeutic targets. The coexistence of multiple conformers observed in simulations, and confirmed by metadynamics calculations, expanding the conformational space knowledge of these ligands with potential implications in understanding the ligand conformational selection during complexation. Also, changes in availability to H-bonding concerning the different substituents and water can reflect on effects at complexation free energy due to variation at the desolvation energetic costs. Based on these results, we expect the obtained structural data provide systemic framework for rational chemical modification of STA-analogs.

## 1. Introduction

Indolocarbazoles (ICZs) are natural products isolated for the first time from actinomycetes, being also found in cyanobacteria, fungi, and marine invertebrates, among other organisms [1]. Since the first isolation of the staurosporine (1) in 1977 [2], these alkaloids received the attention of researchers and pharmaceutical companies because of its natural abundance, wide structural chemical variety and broad spectrum of biological activity, notably antitumor, antibacterial, antifungal, antiviral [3,4], antiparasitic [5] and neuroprotective properties [6]. The indolo[2,3-a]carbazole isomer, the most naturally abundant, has been reviewed extensively over the years, focusing particularly on its biosynthesis and development of analogs [1,7–12]. The results from these efforts already led some of these derivatives to clinical trials, which showed discrete responses against several oncological diseases, in a scenario where the toxicity management was an important issue during the studies [13–16].

The staurosporine analogs (STA-analogs) are known as kinase inhibitors due to competition with ATP molecules for the ATP-binding site. The drug midostaurin (2) is a STA-analog presenting a *N*-benzoyl hexopyranose moiety attached to the chromophore, and has been approved by both Food and Drug Administration (FDA), in the United

States, and European Medicines Agency (EMA), in Europe, for the treatment of acute myeloid leukemia and progressive systemic mastocytosis [13,14,17]. Still, the broad and indiscriminate ability to inhibit a wide range of kinases is an undesirable property of STA-analogs, and therefore it is necessary to search for new structures with higher specificity. Some success in this regard was achieved by compounds 2 and UCN-01 (3), whose selectivity was enhanced by the introduction of substituents at the amide ring and at the carbohydrate moiety, which act as donors/acceptors of hydrogen bonds (H-bonds) [12].

The development of new bioactive ICZs with improved therapeutic properties depends on the understanding of the specific molecular features associated to the complexation, which in turn may include the knowledge of ligands ensemble in solution [18]. Such information can both allow us to realize if the molecular recognition by its target receptor is based on conformational selection or induction [19,20], as well as to screen for the entropic and enthalpic costs of complexation [21–23].

In this regard, computational techniques provide benefits such as cost and labor saving, reduce time to market and accelerate the development of new drugs [24,25]. Among these approaches, molecular dynamics (MD) simulations, represent an important tool to aid the

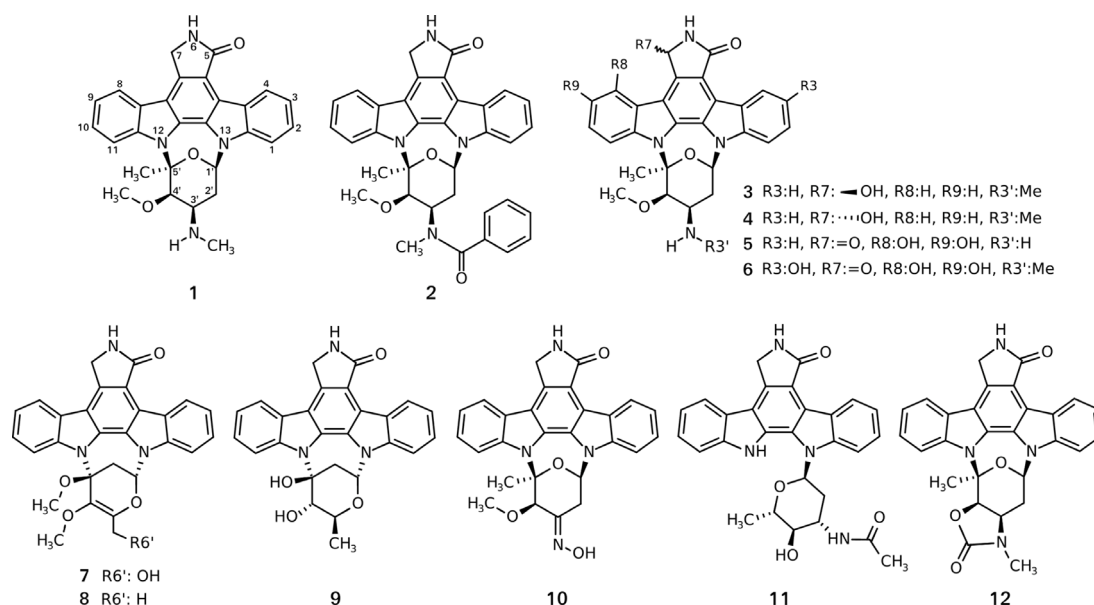
\* Corresponding author.

E-mail address: [hugoverli@gmail.com](mailto:hugoverli@gmail.com) (H. Verli).<https://doi.org/10.1016/j.jmglm.2023.108653>

Received 8 August 2023; Received in revised form 12 October 2023; Accepted 12 October 2023

Available online 26 October 2023

1093-3263/© 2023 Elsevier Inc. All rights reserved.



**Fig. 1.** Structures of compounds staurosporine (1), midostaurin (2), streptocarbazole A (7), B (8) and C (9), TAN-1030 A (10), *N*-acetylated staurosporine (11), and ZHD-0501 (12), whose dynamics in solution was assessed.

rational design and optimization of ligands by estimating the conformational ensemble and dynamics of chemical entities, generally inaccessible through most of the structural biology experiments [26]. In turn, the accuracy of results from MD simulations is directly associated with the calibration of ligands parameter set [27]. In our research group we have been produce parameters for multiple ligands under the GROMOS philosophy, including chalcones, flavonoids, hexopyranoses, aromatic and heterocyclic systems, focusing on calibration of torsions to quantum mechanics profiles and fit of atomic charges to condensed phase properties [28–30], as alternative to automated topology generators, aiming higher accuracy of experimental information [31, 32].

In the same way, the unusual shape of hexopyranose rings attached to STA-analogs, whose structures are presented in Fig. 1, reinforces the requirement for individual parameterization of topological terms to this molecule set. The empirical approach of GROMOS is promising to reproduce, in condensed phase, physicochemical properties of molecular systems after parameters calibration to experimental thermodynamic data, as can include NMR (nuclear magnetic resonance) spectroscopy data. Hence, by the use of molecular dynamics simulations, the current work aims to assess an STA-analogs ensemble in aqueous solution and its interaction with water to probe conformational and structural features involved in its complexation to biological targets. Complementary metadynamics calculations were also employed to reinforce the coexistence of different conformers in solution, supporting the development of a systemic framework able to be further applied on the rational chemical modification of STA-analogs.

## 2. Methods

### 2.1. Deriving the new torsional potentials

In order to obtain the best parameters to characterize the molecular flexibility of the compounds, the relative energy associated with the rotation of bonds at each dihedral angle was calculated by quantum mechanics (QM) and then fitted on molecular mechanics (MM) using relevant molecular fragments (shown in Fig. 3). Gaussian 09 [33] was employed to calculate the potential energies at progressive steps of 30°, describing the quantum system by Møller–Plesset perturbation theory at second-order level (MP2) and using the *scan* routine under 6-31G\* basis set, after geometry optimization adopting a *tight* convergence

criteria. The MP2/6-31G\* theory level has been shown to provide conformational data in agreement with experimental geometries for small molecules and complex systems [34,35], while being computationally accessible.

The potential energy associated with the torsion of the dihedral angle  $m$  in MM calculations based on GROMOS96 force field is described by the equation below:  $V_{\phi,m} = k_{\phi,m} [1 + \cos \delta_m \cos(n_m \phi_m)]$  whose terms were parameterized and represent the dihedral angle value  $\phi_m$ , the multiplicity of the term  $n_m$ , the associated phase shift  $\delta_m$ , and the corresponding force constant  $k_{\phi,m}$  [36]. The quantum new parameters were fitted to MM potentials using the Rotational Profiler server [37] under the GROMOS53A6 force field [36].

### 2.2. Topology building

The compounds structures were taken from the Protein Data Bank (PDB), when available, or built using the molecular editor Avogadro [38]. The GROMOS53A6 [36] and GROMOS53A6<sub>GLYC</sub> [30,39] force fields provided the topological parameters, such as bond stretching, bond-angle bending, improper dihedral deformation, and Lennard-Jones potentials that supported the MD simulations. The dihedral potentials of the hexopyranose rings were empirically adjusted by relaxation of the associated energy and analyzed through the  $xy$  dispersion in relation to the  $^3J_{H,H}$  vicinal coupling constants observed in NMR spectroscopy experiments, aiming to better reproduce its flexibility in solution. Additionally, the parameters to describe the bond angles formed between the atoms involved in bicyclic and tricyclic structures in the carbohydrate moiety were assigned after the geometry optimization in QM. Parameters for aromatic rings were obtained from a previous work from the group [29]. QM calculations were performed by Gaussian 09 [33] in order to obtain, for the relevant fragments, the partial atomic charges based on MP2 theory under 6-31G\* basis set using polarizable continuum model (PCM) solvent and fitted to Restrained Electrostatic Potential (RESP) charges. MM atomic charges were adjusted to better reproduce the QM dipole moment vector using an in house script searching for least-squares fit solutions previously developed [28]. In this process, the partial atomic charges from charge groups previously parameterized in GROMOS [28,29] were employed to restrict the fitting to the dipole moment vector, reinforcing the uniformity of the produced charges to the GROMOS 53A6 force field.

To validate the topological parameters employed to describe the hexopyranoses, the obtained conformational ensemble was compared to experimental data of  ${}^3J_{\text{H,H}}$ , taking into account the correlation between the vicinal coupling constants and the dihedral torsion angles [40]. Firstly, the hydrogen atoms were added to carbon atoms of the hexopyranose rings, directly on the trajectory simulation, by OpenBabel tool [41]. The geometry of the added hydrogens was inspected, and accurately reproduced the carbon atoms expected hybridization. The average  $\pm$  sd of the  ${}^3J_{\text{H,H}}$  constants from MD were calculated using an in house developed tool based on the Haasnoot–Altona parameterization of the Karplus equation [40,42,43]:  $j_o^{-1}$ .

$${}^3J_{\text{H,H}} = 13.86\cos^2\phi - 0.81\cos\phi + \sum_{i=1}^I \Delta\chi_i [0.56 - 2.32\cos^2(\xi_i\phi_o^{-1} + 17.9|\Delta\chi_i|)]$$

where  $j_o = 1$  Hz,  $\phi_o = 1^\circ$ ,  $\phi$  is the dihedral angle between the coupled hydrogen atoms,  $I$  corresponds to the number of substituents different from hydrogen on the fragment H–C–C–H,  $\Delta\chi_i = \chi_i - \chi_{\text{H}}$  is the difference in electronegativity between an atom  $i$  and hydrogen, assuming additive effects of the substituents attached to the carbons of the H–C–C–H fragment and dependent on Huggins' relative electronegativity [44]. Finally,  $\xi_i$  takes the value +1 or -1, depending on the orientation of the electronegative substituent, being negative in the conformation where the coupled hydrogen atoms are eclipsed [45].

### 2.3. Production simulation settings

In order to analyze the dynamics of STA-analogs in aqueous solution, MD simulations were performed using GROMACS suite in conjunction with GROMOS53A6<sub>GLYC</sub> force field during 1  $\mu\text{s}$ , aiming to reproduce the possible conformational states in solution, specially slow hexopyranose transitions [46,47]. The simulations were conceived with an integration step of 2 fs in a dodecahedron simulation box, where a single molecule of each compound was placed and solvated with the extended simple point charge (SPC/E) water model [48], under periodic boundary conditions, after an initial energy minimization by *steepest descents* algorithm. Electrostatic interactions were treated by the particle-mesh Ewald (PME) algorithm [49], while van der Waals contacts were evaluated with the Verlet *cutoff-scheme* [50] to truncate all non-bonded interactions at 1.2 nm [51]. The solvent and solute covalent bond lengths were constrained by the LINCS method [52,53], and the system temperature and pressure were kept constant at 298 K and 1 bar, respectively, using the V-rescale thermostat [54] and the Parrinello–Rahman barostat [55].

### 2.4. Obtaining the conformational ensemble of STA-analogs

The conformational ensemble of STA-analogs was assessed through the observation of the rotational profile of dihedral angles from substituents, glycosidic bonds, and the hexopyranose puckering. The distribution of dihedral angles was obtained from the *gmx angle* tool and distances between atoms establishing intramolecular electrostatic interactions from the *gmx pairdist* tool. The PLUMED library, version 2.5 [56], was interfaced to GROMACS to describe the hexopyranose puckering using the coordinates of  $\theta$  and  $\phi$  angles, whose distributions were further obtained from *gmx analyze* tool. The hexopyranose conformations populated during the simulations were identified using a protocol based on the Cremer–Pople  $\theta$  and  $\phi$  angular coordinates [57]. We refer to specific carbon and oxygen atoms in hexopyranose rings by its IUPAC numbering designation [58].

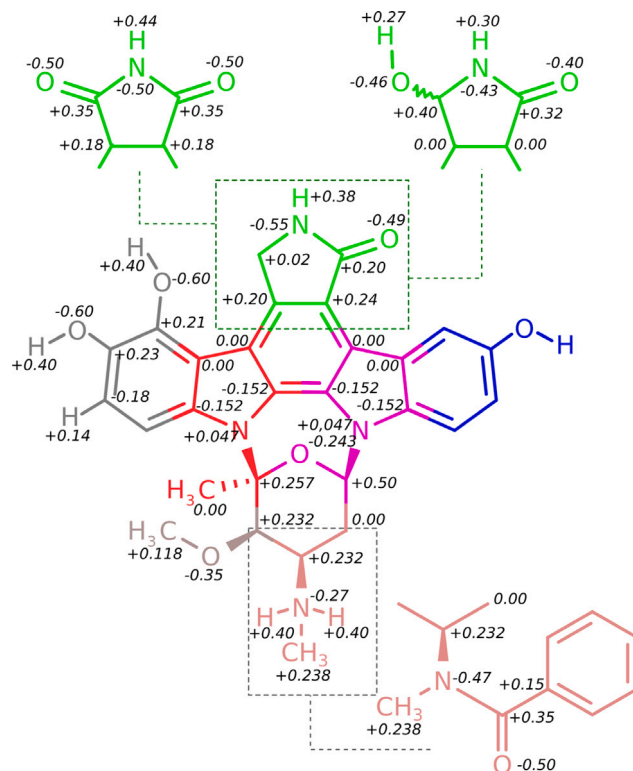


Fig. 2. New partial atomic charges derived to build the STA-analog topologies comprised into charge groups represented by different colors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### 2.5. Metadynamics

In order to reinforce the sampled conformational ensemble of STA-analogs hexopyranose in solution, particularly its rings conformations, well-tempered metadynamics was carried out by PLUMED. Metadynamics was a strategy chosen to reconstruct the free energy landscape (FEL) of the hexopyranoses as a function of selected collective variables (CVs). A bias potential as a function of CVs was added to the system Hamiltonian in a mode dependent on the evolution of the system [59]. The calculations were performed until 600 ns using the coordinates of  $\theta$  and  $\phi$  angles as CVs, under a bias factor of 10, after pilot evaluations to obtain the better convergence for all molecules through a standardized protocol. The Gaussians potential was deposited every 500 time steps with initial height  $h = 0.5$  kJ mol $^{-1}$  and width  $\sigma = 0.1$  rad. The errors ( $\epsilon$ ) related to the CVs were calculated by block-averaging analysis and comprised up to 2% of the barrier value. The FEL of the puckering was built using the *sum hills* utility and Gnuplot plotting package.

### 2.6. Assessing the interactions with water

The interaction between STA-analogs and hydration water molecules was analyzed to probe the availability of compounds functional groups to interact with its target receptors. The water accessibility to make H-bonds with polar atoms was observed by solvent Radial Distribution Function (RDF), using the *gmx rdf* tool. Polar atoms that showed first solvation shell within 0.5 nm were evaluated for H-bond averages ( $\text{Aver}_{\text{HB}}$ ) and for autocorrelation functions to estimate the H-bond lifetime ( $\text{Lifetime}_{\text{HB}}$ ) and H-bond free-energy of breakage ( $\Delta G_{\text{HB}}$ ) using the *gmx hbond* tool [60,61]. These H-bond descriptors were complemented by Spatial Distribution Functions (SDF) [62], calculated by the *gmx spatial* tool, to find out the water molecules location around STA-analogs. All the estimates were obtained from the entire simulation trajectory of each compound, recorded at each 1 ps.

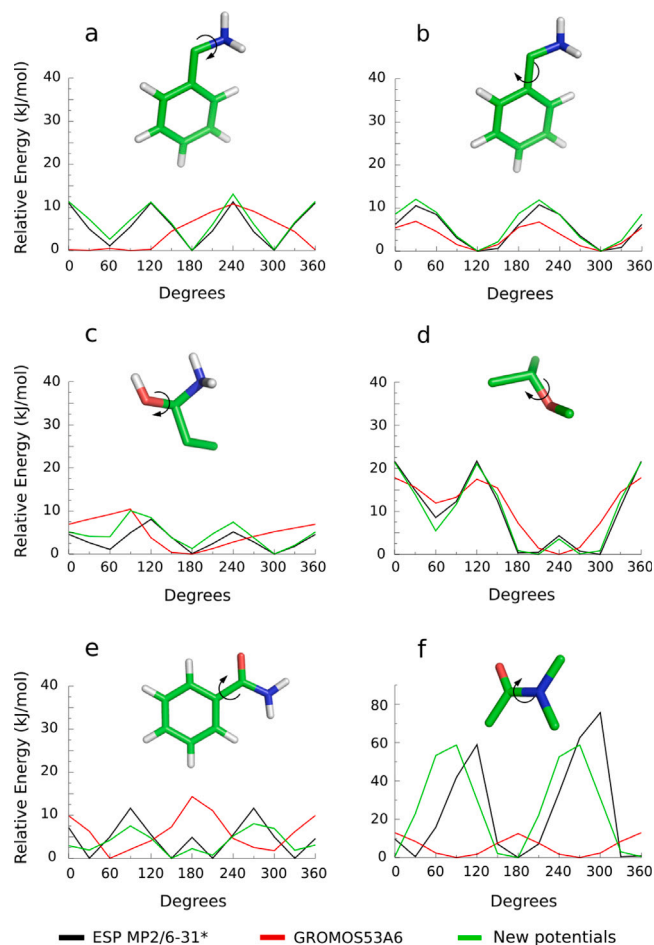


Fig. 3. Structures and energy potential curves applied to derive the new torsional parameters.

### 3. Results and discussion

#### 3.1. Topology building

The parameters set generated to describe the STA-analogs conformational ensemble in solution was empirically developed considering the GROMOS force field parameterization approach compatibility. The partial atomic charges distributed into charge groups, were presented by the representative structure in Fig. 2, which facilitates future efforts to parameterize similar compounds. Atomic charges not shown were taken from previous works [29,30]. The respective curves representing the energy minima and barriers of new dihedral potentials derived by fitting the MM profiles to the corresponding profiles calculated by QM and its fit are shown in Fig. 3, herewith the structures used to generate them. The fitted torsion parameters are indicated on Table 1.

From the novel torsional potentials and partial atomic charges, it was possible to perform MD simulations of a series of 12 compounds, presenting variable substituents and complexity in carbohydrate moieties (Fig. 1). The validation of the conceived topologies was performed comparing theoretical averages of  $^3J_{H,H}$  coupling, obtained from unbiased MD, to respective available experimental NMR data (see further in the text).

##### 3.1.1. Calibration of the hexopyranose torsions

When we initially analyzed the puckering of hexopyranoses attached to the ICZ chromophore by two *N*-glycosidic linkages, this moiety showed a noticeable rigidity, characterized by a single conformational state stability. Previous analysis suggested that the parameters

Table 1

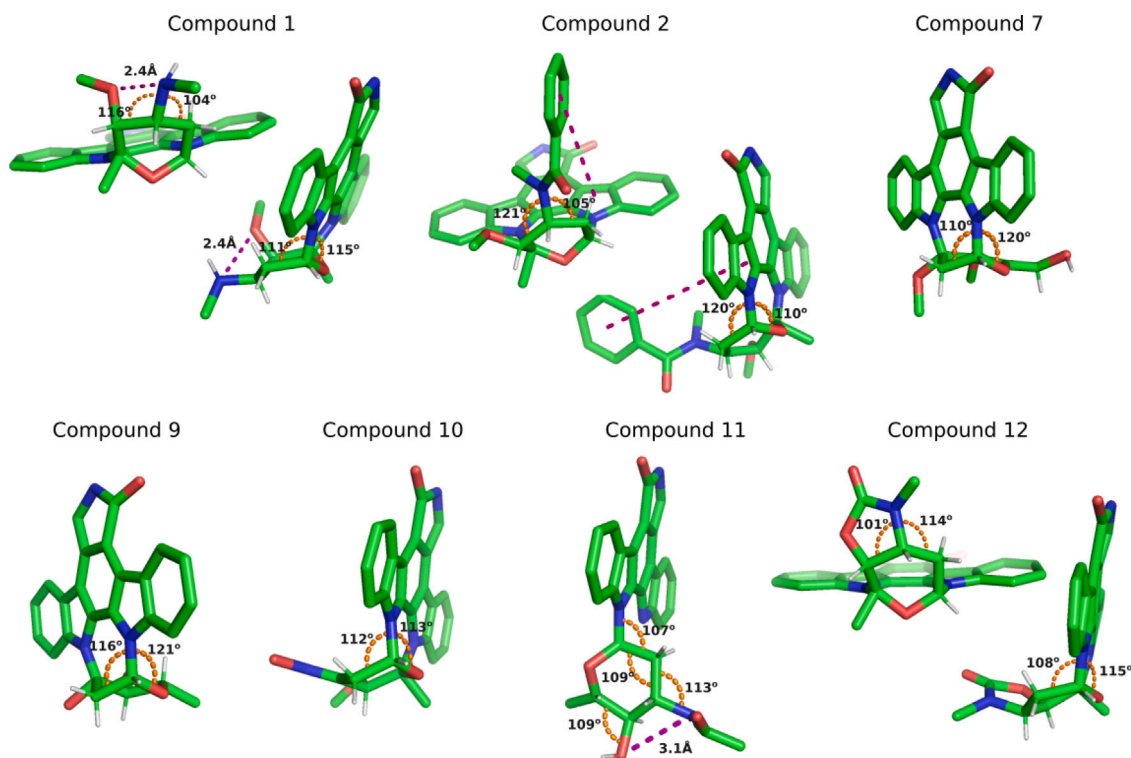
Dihedral potentials generated to describe the STA-analogs flexibility from MD simulations.

	Dihedral angle	$k_{\phi}$ (kJ mol <sup>-1</sup> )	$\delta$	$n$
a	C-CH <sub>2</sub> -N-H	-8.207	0	0
	C-CH <sub>2</sub> -N-H	5.910	0	1
	C-CH <sub>2</sub> -N-H	5.769	120	1
	C-CH <sub>2</sub> -N-H	5.951	0	3
b	C-C-CH <sub>2</sub> -NH <sub>2</sub>	-3.364	0	0
	C-C-CH <sub>2</sub> -NH <sub>2</sub>	2.138	0	2
	C-C-CH <sub>2</sub> -NH <sub>2</sub>	2.981	120	2
c	CH <sub>2</sub> -CH-O-H	-4.295	0	1
	CH <sub>2</sub> -CH-O-H	2.759	0	3
d	CH <sub>3</sub> -CH-O-CH <sub>3</sub>	-3.132	0	0
	CH <sub>3</sub> -CH-O-CH <sub>3</sub>	5.012	0	3
	CH <sub>3</sub> -CH-O-CH <sub>3</sub>	1.931	0	0
e	C-C-C-NH <sub>2</sub>	2.585	0	1
	C-C-C-NH <sub>2</sub>	-8.474	0	2
	C-C-C-NH <sub>2</sub>	0.925	0	6
	C-C-C-NH <sub>2</sub>	17.711	0	0
f	CH <sub>3</sub> -C-N-CH <sub>3</sub>	-28.630	0	2
	CH <sub>3</sub> -C-N-CH <sub>3</sub>	14.117	120	2
	CH <sub>3</sub> -C-N-CH <sub>3</sub>			

of the dihedral potentials employed by the GROMOS53A6<sub>GLYC</sub> force field to evaluate the hexopyranoses flexibility are associated with high barriers and large timescales to observe the transition of conformers [63,64]. In this sense, the presence of double glycosylation may have caused an additive effect on energy barriers and timescales to observe the complete conformational ensemble. The high barriers observed from metadynamics of the carbohydrate puckering, with original parameters of GROMOS53A6<sub>GLYC</sub> force field, motivated us to apply an empirical adjustment of the dihedral potentials exploring the relaxation of the energy involved in the flexibility of hexopyranose rings. Relaxing the empirical force constant to 10% allowed the conformers transition and brought the  $^3J_{H,H}$  theoretical values closer to the experimental ones, with lower standard deviation (Table S1), so this relaxed potentials were implemented in analyses of carbohydrates attached by double glycosylation.

As general feature for most compounds analyzed, the vicinal coupling average of H1' regarding H2'a and H2'e calculated from MD simulations are close to those observed by empirical measurements, as presented in Table 2. Concerning the compound 1, the main discrepancy occurs towards the vicinal coupling involved to the H3' signal. Attached to the same carbon atom, the 3'-amino group assumes the protonated form in aqueous solution. Differences in Haasnoot-Altona estimate, related to compounds presenting non-conjugated lone pairs of electrons in the  $\alpha$  substituent position, were already observed and would be associated to lack of calibration for predictions involving charged groups [70,71]. The Haasnoot-Altona equation takes into account the number of substituents bonded in  $\alpha$  and  $\beta$  positions belonging to the coupling pathway, besides its electronegativity and orientation [70]. Furthermore, interactions between substituents were found to affect  $^3J_{H,H}$  couplings by an important extent [72]. In this sense, the geometry of the aminopyranose allows the intramolecular electrostatic interaction between the 3'-amino group and 4'-oxygen, with average distance equal to  $0.24 \pm 0.02$ .

Other important factor affecting  $^3J_{H,H}$  couplings is the measure of the bond angles along the coupling pathway, that is, H-C-C and C-C-H angles [73]. An exploration of the bond angle bending in the carbohydrate nuclei is presented by snapshots in Fig. 4 and shows how the ring strain may affect the  $^3J_{H,H}$  vicinal coupling prediction by changes in bond angle measure over the canonical 109.5° of aliphatic carbons. It is possible to realize the magnitude of bond angle bending effect in compound 1 by comparing the vicinal coupling constants from its protonated form to the free base ones, as presented in Table 2. In the same way, it is likely that the unsaturated bond connecting the NOH group to C3' in compound 10 causes additional ring strain in this region, changing the magnitude of  $^3J_{H1',H2'}$  couplings when



**Fig. 4.** Bond angle bending, intramolecular interaction and ring strain in STA-analog carbohydrate nuclei may affect the  ${}^3J_{\text{H,H}}$  vicinal coupling prediction by the Haasnoot–Altona equation. Orange dashes indicates X–C–C bond angle, while purple dashes indicates intramolecular interaction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 2**

Comparison of empirical and theoretical vicinal coupling constants of STA-analogs sugar moieties employed to validate its topology.

Compound	1		1 (free base)		2		7	
	NMR [65]	MD	NMR [65]	MD	NMR	MD	NMR [66]	MD
${}^3J_{\text{H,H}}$								
H1'-H2' $\alpha$	9.4	7.9 $\pm$ 1.1	5.8	6.1 $\pm$ 1.2	–	7.7 $\pm$ 1.1	6.0	4.2 $\pm$ 1.0
H1'-H2' $e$	1.8	1.7 $\pm$ 1.3	1.3	1.7 $\pm$ 0.8	–	1.2 $\pm$ 0.8	3.5	2.6 $\pm$ 0.7
H2' $\alpha$ -H3'	7.4	7.9 $\pm$ 2.4	3.6	6.8 $\pm$ 1.5	–	9.6 $\pm$ 1.3	–	–
H2' $e$ -H3'	11.2	4.1 $\pm$ 4.4	3.6	1.2 $\pm$ 0.6	–	4.6 $\pm$ 2.6	–	–
H3'-H4'	1.6	6.4 $\pm$ 1.4	3.6	5.8 $\pm$ 1.0	–	7.9 $\pm$ 0.9	–	–
Compound	9		10		11		12	
${}^3J_{\text{H,H}}$	NMR [67]	MD	NMR	MD	NMR [68]	MD	NMR [69]	MD
H1'-H2' $\alpha$	3.9	2.7 $\pm$ 0.7	–	8.5 $\pm$ 0.5	11.7	9.9 $\pm$ 1.4	10.0	8.2 $\pm$ 1.1
H1'-H2' $e$	6.1	3.6 $\pm$ 0.8	–	0.79 $\pm$ 0.3	3.4	3.8 $\pm$ 1.8	6.0	1.6 $\pm$ 1.2
H2' $\alpha$ -H3'	–	–	–	–	–	8.3 $\pm$ 1.7	12.0	8.9 $\pm$ 1.3
H2' $e$ -H3'	–	–	–	–	–	4.8 $\pm$ 4.1	5.3	4.6 $\pm$ 3.6
H3'-H4'	–	–	–	–	–	6.6 $\pm$ 3.5	8.8	7.8 $\pm$ 0.8
H4'-H5'	9.9	9.2 $\pm$ 0.4	–	–	–	7.4 $\pm$ 2.5	–	–

NMR shows  ${}^3J_{\text{H,H}}$  in Hz from spectroscopy and MD shows average $\pm$ sd from simulations.

compared with compound 1, which has similar geometry in C1' and C2' (unfortunately, to our knowledge, there is not available data for singlet  ${}^3J_{\text{H,H}}$  coupling related to compound 10 [74–76]).

The sugar moiety of compound 2 differs from the aminopyranose of compound 1 by a *N*-methylbenzamide group attached to C3'. Although there is no available data for singlet  ${}^3J_{\text{H,H}}$  coupling related to compound 2 [77,78], it is possible to realize the predicted averages of  ${}^3J_{\text{H1',H2'}}$  are very close between compounds 1 and 2, as shown in Table 2, once the structure similarity in this molecular site allows for the comparison. On the other hand, there is a notable difference when comparing the values involving the 3' position. Once the *N*-methylbenzamide group is present in compound 2, it can change the stereo-electronic effect on the carbohydrate puckering. Furthermore, it is possible to predict differences between the calculated vicinal coupling and the experimental ones due to the intramolecular apolar

interaction involving the benzene ring at the *N*-methylbenzamide substituent and the pyrrole ring bonding C1', as presented in Fig. 6 A and discussed in the next topic.

Originating from marine-derived *Streptomyces sp.*, the streptocarbazoles A (7) and B (8) presented for the first time two *N*-glycosylations bonding C1' and C3' in hexopyranose to the indole groups nitrogens, in addition to one unsaturation between C4' and C5' [66]. Later reported, the streptocarbazole C (9) presents a similar glycosylation structure, but lacks the unsaturation and the methyl and hydroxyl exocyclic groups (Fig. 1) [79]. The theoretical averages for the  ${}^3J_{\text{H1',H2'\alpha}}$  and  ${}^3J_{\text{H1',H2'e}}$  vicinal coupling in compound 7 are close to the only two signals observable by NMR spectroscopy [66], due to the chemical configuration itself, as presented in Table 2.

Regarding the compound 9, the simplicity of the exocyclic groups, hinders any intramolecular interaction. On the other hand, it may favor the development of a partial double bond, which requires electron

donation by a suitably lone pair on the endocyclic oxygen, and is typically present in transition state conformations [80]. Geometry variations in this molecule site, possibly caused by the intermittent partial double bond, may explain the variation observed on the predicted averages of the coupling constants when compared to the empirical values [67,79] for the compound **9**, as presented in Table 2. The theoretical average of the  ${}^3J_{H4',H5'}$  coupling constant, however, agrees with the experimental data, indicating consistency in conformational dynamics in this molecular site.

The compound **12** is the first natural STA-analog presenting a heterocycle fused to the hexopyranose ring, which takes the  $B_{3,0}$  conformation in solution, as elucidated by NMR spectroscopy [69]. The theoretical  ${}^3J_{H,H}$  averages calculated from MD simulation, when compared to NMR spectroscopy data [69], showed considerable differences regarding  ${}^3J_{H1',H2'e}$  and  ${}^3J_{H2'a,H3'}$  values, as presented in Table 2. The stereoelectronic effect of the fused oxazolone ring, and lack of parameters for this kind of substituent in Haasnoot–Altona predictions may contribute to the differences observed. Meanwhile, the theoretical values of  ${}^3J_{H1',H2'a}$ ,  ${}^3J_{H2'e,H3'}$ , and  ${}^3J_{H3',H4'}$  are close to the experimental ones, which suggests an appropriate geometry in this molecular site.

The hexopyranose topology at compound **11** was built based on parameters developed for *N*-acetylated hexopyranoses [81,82], afterwards the GROMOS53A6<sub>GLYC</sub> force field. The lower energy associated to the dihedral angles avoided any empirical adjustment and displayed theoretical  ${}^3J_{H,H}$  averages very close those observed from NMR spectroscopy [68].

For all ligands analyzed, it is important to point out for possible solvent related differences. The solvation influence on the conformational properties of carbohydrates consists of a combination of specific interactions among both solute and solvent, including preferential solvation in more accessible groups and competition between intramolecular and solute–solvent interactions [83]. This complex combination of phenomena can change the conformers population equilibrium, adding up the intrinsic effects of the different solvents [84,85], which affects the calculation of the NMR spectroscopy properties when averaging over these molecular clusters [86].

### 3.2. STA-analogs conformational ensemble

Empirical observations agree that the carbohydrate moiety, attached at one or both nitrogens of indole group, plays a crucial role in ICZs biological activity [87–89]. With a few exceptions, the *N*-glycosylation is curiously restricted to the fused bis-indole-containing members [90], whose carbohydrate stereochemistry and complexity were able to affect the biological activity leading to changes in cytotoxicity [89,91]. It has been reported that glycosylated ICZs are more potent antitumor agents than its corresponding aglycone form, while structural modifications at these carbohydrates play a role on the selectivity towards different tumor cell lines [89,92].

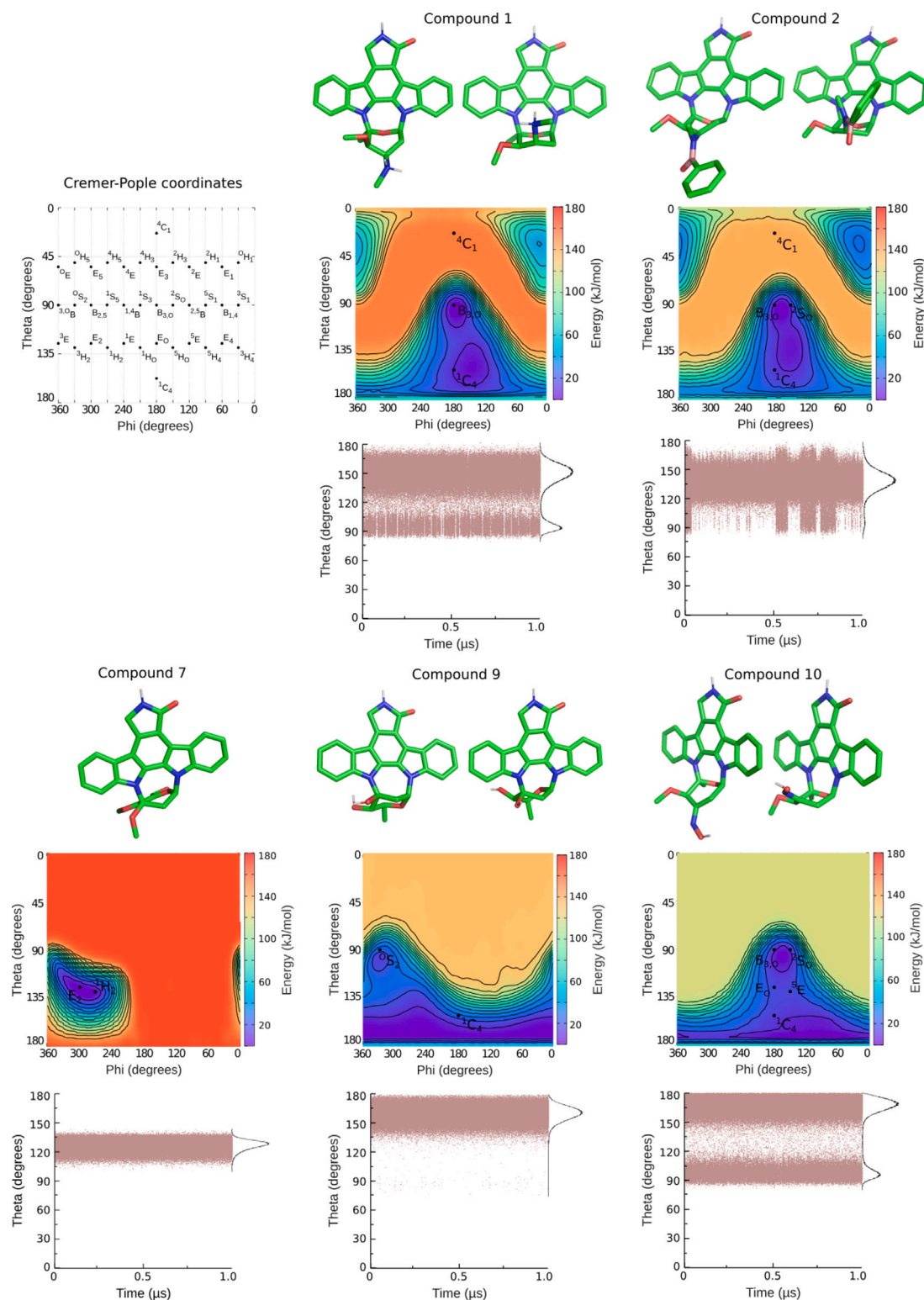
It is well reported that carbohydrate intrinsic structural features represent difficulties for structures description by experimental techniques, such as X-ray crystallography and NMR spectroscopy. In fact, the structural complexity of carbohydrate chains, including their heterogeneity and flexibility, combined with the non-availability of these compounds in sufficient quantities have hindered structural studies of carbohydrates [93]. Furthermore, its conformation depends on ring strain, stereoelectronic effects and changes in electronic density of the substituents, that can lower the barrier between conformers and allow dynamics conformational changes [94,95].

In addition to quantum mechanical methods, calculations based on well parameterized force fields, such as MD simulations, are able to provide detailed and accurate conformational description of carbohydrates in solution [28,96,97], while the comparison to NMR spectroscopy data can be numerically assessed resorting both vicinal coupling constants, estimated using the Haasnoot–Altona equation, and Cremer–Pople puckering coordinates.

The compound **1** and its hydroxylated and oxygenated derivatives, compounds **3–6**, share the same carbohydrate moiety, an aminopyranose with a 3'-amino group attached to the chromophore by two *N*-glycosylations. Previous analyses observed that this aminopyranose presents two main conformational states in solution, corresponding to boat and chair conformations, which are sensitive to the protonation state of the 3'-amine group [65]. Accordingly, our simulation showed the conformational transition between  $B_{3,0}$  and  ${}^1C_4$  states, showing the coexistence of both conformers in solution (Fig. 5). Two factors were previously reported as able to facilitate conformational transitions between  $B_{3,0}$  and  ${}^1C_4$  conformations. The axial 3'-amine group in  ${}^1C_4$  would act as a destabilizing factor of the chair form [98], and the cationic form favors the transition to the boat conformation, which allows for a better solvation of the ammonium group [99]. In water, the 3'-amine group is readily ionized, contributing to the variation of the conformational equilibrium due to the stereoelectronic effects, that involve dipolar component [100] and may be solvent sensitive [101]. The enhanced sampling simulations confirmed the  $B_{3,0}$  and  ${}^1C_4$  share the global minima, but also indicated a local minima corresponding to  ${}^4C_1$  ( $\epsilon_\theta < 0.9$ ,  $\epsilon_\phi < 0.6$ ), as showed the FEL in Fig. 5. The energy barrier between  $B_{3,0}$  and  ${}^1C_4$  states was observed to be  $< 20$  kJ mol<sup>-1</sup>, allowing a frequent interconversion during the unbiased simulation, while  ${}^4C_1$  state remains isolated by a barrier of at least 140 kJ mol<sup>-1</sup> and was not sampled during the 1  $\mu$ s simulated, since it is an unlikely state under normal solution conditions.

Compound **3** is a known enhancer of cytotoxic agents [16], and occupies the same position of compound **1** in kinase ATP-binding site, conserving carbohydrate boat geometry and interactions, as revealed by crystallographic complexes (*PDB ID* 1OKZ, 1NVQ, 1AQ1 and 1STC). However, the strong affinity of compound **3** for alpha-acid glycoprotein 2 (AGP2) leads a significant impact on its efficacy, due decreasing the drug distribution [102]. Unlike the boat conformation in complexes with kinases, it seems the chair is present when compound **3** complexes with AGP2, as revealed the crystallographic complex 7OUB [103]. Our data, estimates the high frequency and interconversion between both conformers in biological solutions and, by using the parameters proposed here, structural strategies to avoid the chair conformation could be implemented aiming to prevent the binding to nonspecific plasma proteins like AGP2.

In contrast to compound **1**, compound **2** lacks the 3'-amine group available to H-bonding. On average, compound **2** binds to protein kinases more weakly than compound **1** and the loss of this H-bond likely contributes to the lower binding affinity [104]. The additional *N*-methylbenzamide group in compound **2** also caused some modifications in FEL of the puckering metadynamics when compared to the compound **1**. The main difference was the formation of a single and bigger valley involving regions compatible to the  ${}^1C_4$ ,  $B_{3,0}$ , and  ${}^2S_0$  conformers ( $\epsilon_\theta < 0.8$ ,  $\epsilon_\phi < 0.7$ ) (Fig. 5). During the unbiased simulation the  ${}^1C_4$  conformer was prevalent, although  $B_{3,0}$ , and  ${}^2S_0$  also were populated. The prevalence of  ${}^1C_4$  state can be favored by the interaction between the benzene ring belonging to the *N*-methylbenzamide substituent and the five members ring from the indole (pyrrole) bonded to C1', potentially associated to a  $\pi - \pi$  interactions. The average distance between the rings was  $0.55 \pm 0.05$  nm, while the minimum distance was about 0.3 nm, consistent with molecular planes separated by interplanar distances of about 0.33 to 0.38 nm, to which the  $\pi - \pi$  interaction is attributed [105]. On the other hand, the maximum distance was about 0.7 nm and was reached when the carbohydrate  $\theta$  angle was compatible to the  $B_{3,0}$  state, whose geometry keeps away the benzene ring from the chromophore, as presented in Fig. 6 B. As occurs with other complexes involving STA-analogs, the compound **2** carbohydrate also appears in the boat state when complexed to dual-specificity tyrosine phosphorylation-regulated kinase 1 A (DYRK1 A, *PDB ID* 4NCT). This conformation places the phenyl ring outside the ATP cavity where it takes part of a hydrophobic interaction [104]. Supporting the apparent rigidity of this molecule moiety, the dihedrals



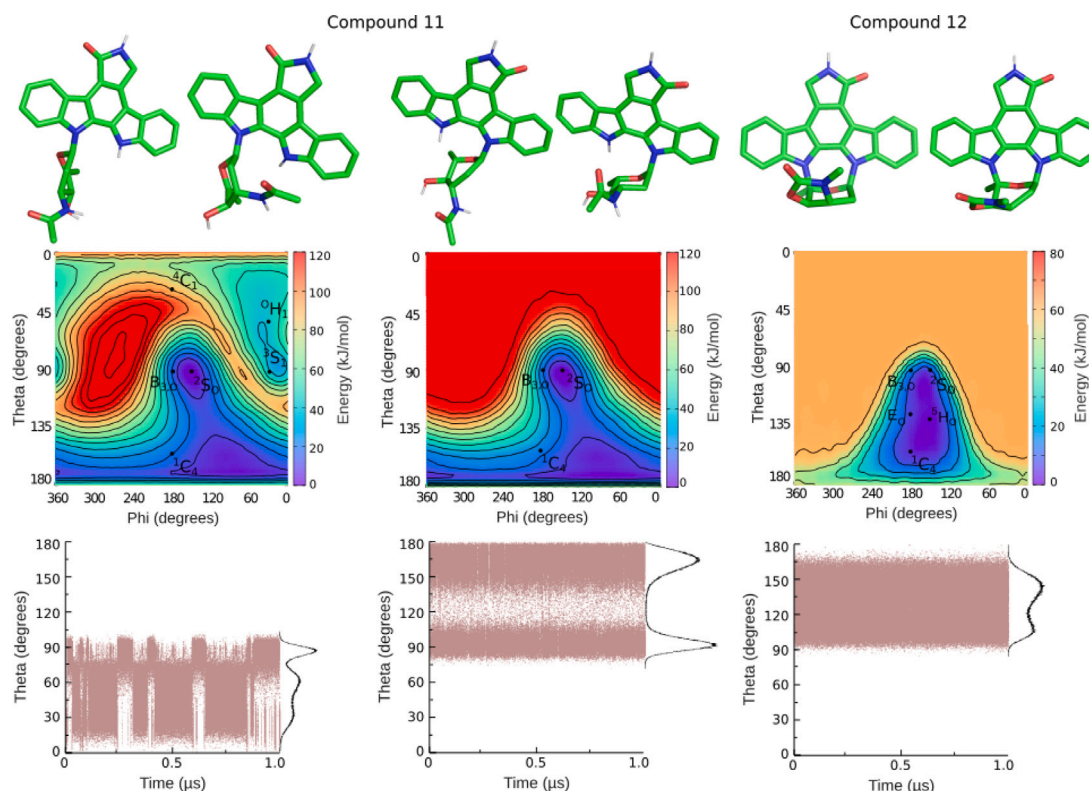


Fig. 5. (continued).

oxygen, contribute to the rigidity of the system observed during the MD simulation and evidenced in the metadynamics approach ( $\epsilon_\theta < 0.4$ ,  $\epsilon_\phi < 0.3$ ). It is likely that, this chemical configuration keeps the carbohydrate at  $E_2$  and  ${}^1H_2$  conformers, as showed the study of the compound 7 hexopyranose puckering presented in Fig. 5.

To illustrate the use of hexopyranose unusual conformations in drug design, let us to draw a parallel with the drug zanamivir (**13**), whose structure is shown in Fig. 7. This antiviral was designed to improve a well-known sialidase inhibitor, the Neu5Ac2en neuraminic acid, by the rational inclusion of a guanidine substituent in a hexopyranose ring to target a negatively charged pocket near the active site in sialidases [106]. The inhibitor, compound 13, which has a double bond between C2' and C3', had not present the expected success. It would not be able to assume the  ${}^4H_5$  transition state, predicted to bind to sialidases, while only the  $E_5$  conformer was observed in complexes [107]. The compound 13 was associated with reduced binding avidity to the enzyme, which still showed some catalytic competence [106,108]. The compound 7, on the other hand, presented both  $E_2$  and  ${}^1H_2$  conformers during the simulation. A careful comparison between the compounds 7 and 13 structures, point out the simplest exocyclic groups in the ICZ hexopyranose that are not large enough to establish intramolecular interactions and thus to influence conformational states [94,95]. Based on this comparison and on the presence of intramolecular interactions stabilizing conformers in hexopyranoses, observed from our simulations, a kind of intramolecular electrostatic effect would also be compatible in compound 13, promoting the  $E_5$  stabilization due to its large and polar substituents.

The absence of the endocyclic unsaturation in compound 9 relaxes the ring system rigidity of streptocarbazoles. The analysis of the hexopyranose  $\theta$  angle distribution showed the prevalence of the  ${}^1C_4$  conformer, while the  ${}^0S_2$  was more rarely populated. By enhanced sampling analysis it was observed that both conformers are separated by an energy barrier of at least  $20 \text{ kJ mol}^{-1}$ , relatively high to the easy transition between each other ( $\epsilon_\theta < 0.2$ ,  $\epsilon_\phi < 0.2$ ), justifying the poor frequency of the  ${}^0S_2$  state during the unbiased simulation (Fig. 5).

The analysis of  $\theta$  angle distribution in compound 10 hexopyranose showed the prevalence of  ${}^1C_4$ , while  $B_{3,0}$  and  ${}^2S_0$  conformers were also populated. The unsaturation bonding NOH group to C3' adds rigidity to this ring site, while maintains some symmetry during the conformational itinerary. Analyzing Fig. 5 one can realize the conformational states transit through chair, envelope and boat maintaining the regularity of  $\phi$  angle among  $160^\circ$  and  $180^\circ$ . By enhanced sampling analysis it was observed that the minima are separated by an energy barrier of less than  $20 \text{ kJ mol}^{-1}$ , allowing the conformers transition during the simulation ( $\epsilon_\theta < 0.3$ ,  $\epsilon_\phi < 0.3$ ).

NMR spectroscopy data suggest the most likely conformation of the compound 11 hexopyranose in solution would be  ${}^4C_1$  [68], thus the start geometry for simulations was built in this state. The MD showed  $\theta$  angles distribution ranging from  $\approx 0^\circ$  to  $\approx 100^\circ$ , thereby enhanced sampling calculations were performed to better identify these conformers based on Cremer–Pople coordinates. Surprisingly, the FEL did not showed the global minimum near the  ${}^4C_1$  region. Instead, two local minima were observed in  ${}^1C_4$  and  ${}^2S_0$  regions ( $\epsilon_\theta < 0.7$ ,  $\epsilon_\phi < 0.4$ ) as presented in Fig. 5. Based on the metadynamics FEL, it is possible to point out that the carbohydrate transited between the conformers  ${}^3S_1$ ,  ${}^0H_1$ , and  ${}^4C_1$  during the unbiased simulation. Its respective local minima valleys are poorly defined, allowing the spontaneous conformers interconversion in solution, within a region of about  $40 \text{ kJ mol}^{-1}$ . An energy barrier of at least  $90 \text{ kJ mol}^{-1}$  was observed to separate the sampled states from the global energy minima region, which is high enough to make difficult the observation of the  ${}^1C_4$  and  ${}^2S_0$  conformers during the unbiased simulation starting from  ${}^4C_1$ .

To better investigate the structural features implicated in the conformational dynamics of compound 11, simulations using  ${}^1C_4$  as hexopyranose starting geometry, analyses of *N*-glycosidic bond angle distribution and intramolecular interactions were carried out. The simulation starting from  ${}^1C_4$  geometry presented frequent transition between  ${}^1C_4$  and  ${}^2S_0$  conformers. The carbohydrate puckering metadynamics resulted in a FEL showing the same two energy minima, while it did not present the local minima observed when the simulation started from



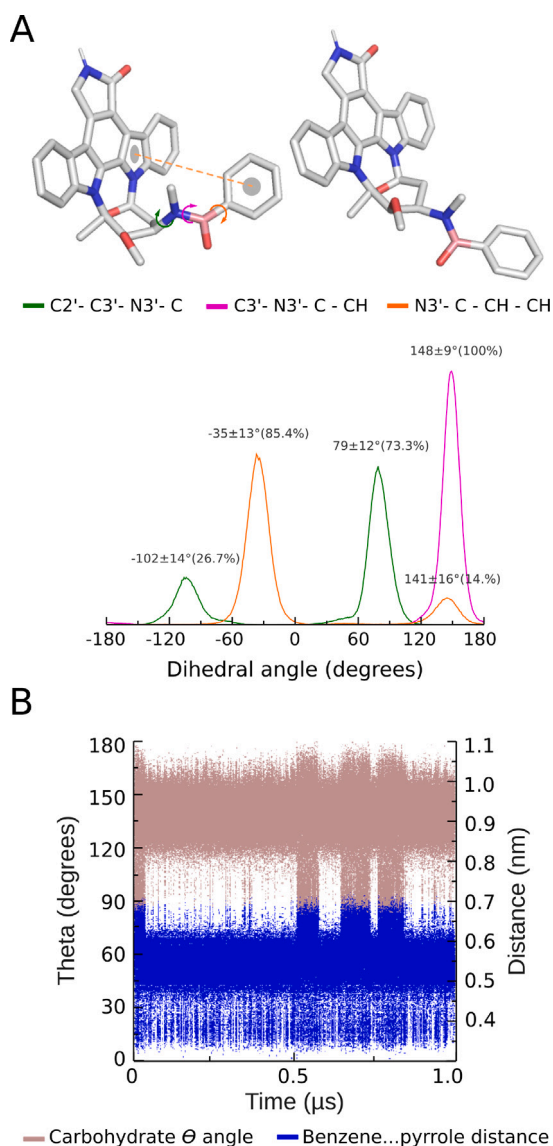


Fig. 6. The dihedral angle distribution of the compound 2 revealed small flexibility in the *N*-methylbenzamide substituent (A), while its orientation can affect the carbohydrate puckering due to possible  $\pi$ - $\pi$  interaction between the benzene and pyrrole rings (B), as suggested by the compatible molecular geometry.

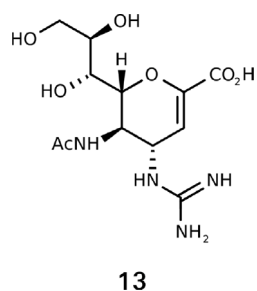


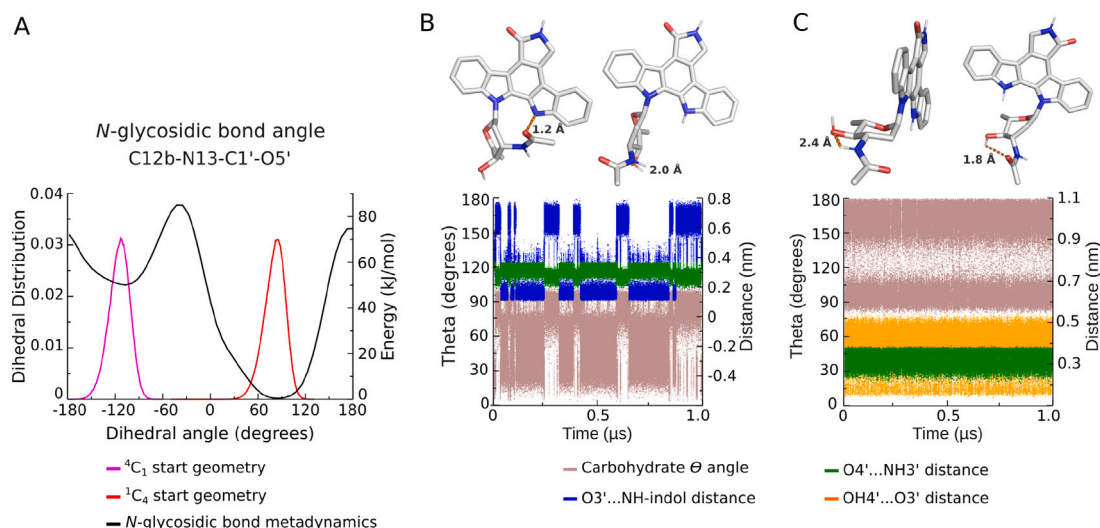
Fig. 7. Structure of zanamivir (13), the antiviral designed to inhibit sialidases resorting the rational inclusion of a guanidine substituent in a hexopyranose ring.

the  ${}^4C_1$  geometry ( $\epsilon_\theta < 0.4$ ,  $\epsilon_\phi < 0.3$ ), as presented in Fig. 5. The *N*-glycosidic dihedral angle average was  $90.5 \pm 52^\circ$  starting from  ${}^1C_4$ , while it was  $-119.5 \pm 34^\circ$  starting from  ${}^4C_1$ . The enhanced sampling of the *N*-glycosidic bond angle distribution showed two energy minima,

which coincide with the dihedral angles averages observed during both simulations, as showed in Fig. 8 A. The observed differences may be associated to factors that limit the rotational freedom degrees of the *N*-glycosidic bond and the free conformational interconversion of the carbohydrate. The energy cost for the *N*-glycosidic bond to populate the two main conformation is at least  $30 \text{ kJ mol}^{-1}$ , when starting from  ${}^4C_1$  (Fig. 8 A). However, starting from  ${}^1C_4$ , the cost is at least  $80 \text{ kJ mol}^{-1}$ , unlikely to be achieved under normal conditions in solution. In addition to the energetic component of *N*-glycosidic bond, intramolecular interactions can modulate the hexopyranose molecular flexibility, stabilizing some conformations. It was already observed *in silico* that the presence of exocyclic groups makes rich conformational landscapes, affecting the conformer transition of the hexopyranose rings. Intramolecular interactions in free  $\beta$ -*N*-acetylglucosamine stabilized the  ${}^5E$  conformation, favored by the proximity of *N*-acetyl arm (attached in C2') to the hydroxyl group at the neighboring carbon [95]. On the same direction, a detailed analysis of the distance between the polar groups showed that electrostatic interactions can modulate the *N*-acetylated hexopyranose puckering in compound 11. From the simulation started with  ${}^4C_1$  geometry, the minimum distance between indole amine group and acetyl oxygen was about 0.12 nm (average distance of  $0.35 \pm 0.02 \text{ nm}$ ) and occurred when  $\theta < 90^\circ$ , what is compatible with  ${}^0H_1$  and  ${}^4C_1$  conformers. On the other hand, indole amine group and acetyl oxygen presented increase of the distance between each other associated to  $\theta \approx 90^\circ$  and accompanied by approximation of the 4'-oxygen and acetyl amine groups until 0.20 nm (average distance of  $0.32 \pm 0.02 \text{ nm}$ ), what is compatible with the  ${}^3S_1$  conformer (Fig. 8 B). The simulation started from  ${}^1C_4$  showed that the minimum distance between 4'-oxygen and acetyl amine groups was 0.24 nm (average distance of  $0.32 \pm 0.02 \text{ nm}$ ), what is compatible with the  ${}^1C_4$  state. The minimum distance between 4'-hydroxyl to the acetyl oxygen was until 0.18 nm (average distance of  $0.38 \pm 0.06 \text{ nm}$ ) which is, on the other hand, compatible with  ${}^2S_0$  state (Fig. 8 C).

Regarding the compound 12, the ring strain generated from the complexity of the tricyclic structure, allied to the dynamic stereo-electronic and electrostatic changes [94,95], may contribute to favor the broad conformer interconversion observed during the simulation. The  $\theta$  angle was distributed between  $\approx 85^\circ$  and  $\approx 180^\circ$ , indicating a great flexibility in the conformational transition of the hexopyranose in compound 12. Moreover, the FEL from enhanced sampling calculations showed a large valley that holds conformations within the angular coordinates of  $\theta > 90^\circ$  combined with  $120^\circ < \phi < 210^\circ$ , comprising  $B_{3,0}$ ,  ${}^2S_0$ ,  $E_0$ ,  ${}^5H_0$ , and  ${}^1C_4$  and confirming the broad flexibility ( $\epsilon_\theta < 0.9$ ,  $\epsilon_\phi < 0.8$ ).

The anticancer and antimicrobial properties of these small ligands are involved in the inhibition of several protein kinases, including various isoforms and cyclin-dependent kinases in a competitive ATP-binding site fashion [6,109]. The boat state is expected to be the bioactive conformer, once crystallographic structures showed the inhibitor complexed to the ATP-binding site in this state, what is accompanied by conformational induced fit on residues in the target proteins to accommodate the bigger volume of the ligands when compared to the ATP substrate [104,110]. Our simulation showed the coexistence of different conformational states of STA-analogs hexopyranoses in aqueous solution, with variable frequencies, depending on the carbohydrate complexity. The accessibility of interconversion between conformational states observed for these ligands might correlate with a conformational selection as a model to describe kinetic data of complexation on these molecular systems. Furthermore, the interconversion equilibrium may be biased by the presence of a given protein target, thus ligands may bind its respective targets in an atroposelective-like way [111].



**Fig. 8.** Flexibility of the *N*-acetylated hexopyranose in compound **11**. The dihedral angle distribution of *N*-glycosidic bond is associated to the hexopyranose geometry (A). The puckering can be modulated by intramolecular interactions observed when simulations started from  ${}^4C_1$  geometry (B) and  ${}^1C_4$  geometry (C). The main conformations and interaction minima distances are indicated.

### 3.3. Interactions with water bulk

Direct ligand–receptor interaction is not the only factor determining the binding affinity between ligand and target molecule. In addition to electrostatic forces, one of the most important processes contributing to the binding free energy is desolvation. Water molecules play a crucial role in the mutual recognition of proteins and ligands during the complexation. Previously the complexation, both the protein binding site and the ligand are solvated, hence energy is needed to remove its water shells to allow the access of the ligand to the protein environment [112,113]. Have been suggested that aspects involving the ligand desolvation can result in differences in binding affinity. The ligand desolvation cost may be compensated by subsequent ligand–receptor interactions, while the exact balance between these contributions has an important role in determining the resulting affinity of the ligand for the protein [114]. In this regard, to assess aspects of the interaction between the free ligand with the water bulk can probe potential polar ligand–receptor interactions, as well can identify polar groups unavailable or less likely for interactions upon complexation. Such estimates can be accurately performed based on measures of the ligand energy cost for desolvation.

From MD simulations of STA-analogs in aqueous solvent, measures of its interaction with water were obtained from all polar atom and presented by  $Aver_{HB}$ ,  $Lifetime_{HB}$  and  $\Delta G_{HB}$  for interaction breakage, which can be consulted in Table S2 (supplementary material). Polar atoms that showed first solvation layer within 0.5 nm are discussed and enthalpies of desolvation are compared between the ligands to infer differences in binding affinities.

Regarding the chromophore in STA-analogs, the main polar interactions observed at compound **1** are centered at the lactam ring, through atoms 5-carbonyl and 6-amine (Table 3). Ranging from  $8.61 \pm 0.03$  to  $11.07 \pm 0.02$  kJ mol $^{-1}$ , these interactions undergo different interference levels by substitutions on the chromophore. The desolvation enthalpy can be reduced up to  $7.69 \pm 0.02$  kJ mol $^{-1}$  when a hydroxyl group is added at R7, directly on the lactam ring, as occurs in compound **3** and its stereoisomer UCN-02 (**4**), competing with 5-carbonyl and 6-amine for the solvent interaction. The SDF analysis showed the changes of the water molecules spacial distribution, that is, how the water bulk shifts from 5-carbonyl surrounding to nearby of the 7-hydroxyl groups. As shown in Fig. 9, the water bulk responds to the attractive electrostatic effect of the additional polar group, according to the 7-hydroxyl group stereochemistry of each compound.

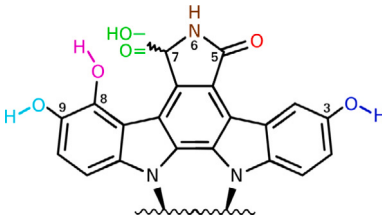
These changes in the hydration water structure are in agreement with data observed by X-ray diffraction. Compounds **1** and **3** occupy the same ATP-binding site in different kinase proteins, both compounds make very similar hydrophobic interactions involving the apolar ring system, and conserve the H-bonds from the lactam ring and from the carbohydrate. The differences occur regarding the 7-hydroxyl group at compound **3**, which makes additional H-bonds with surrounding residues in kinases, accompanied by coordination water molecules in distinct regions [110,115]. The SDF analysis showed the presence of water bulk in regions compatible with the location of these coordinate water molecules empirically described, as pointed out in Fig. 9 A and B. These observations reinforce the predictive power of using ligand–solvent interactions, obtained by MD simulations, to probe ligand–receptor interactions.

The kinases inhibited by compound **3** show residues presenting side chains able to make H-bond with the 7-hydroxyl group, in the hydroxyl-group-binding pocket site [110]. Regarding compound **4**, however, there is a little structural data available in the literature. A recent study points to the H-bond lost and probable clash with hydrophobic residues, besides shift of the ligand optimal binding pose when the reverse position of 7-hydroxyl group is present [103]. Alternative positions of hydroxyl group constitute valuable information for aiding the discovery of more specific analogs, since previous studies with compound **3** showed that the hydroxyl at lactam ring was responsible for differential interaction with kinases [110]. Indeed, as occurs with compound **3**, the structural water incorporation is thought to be a promising approach to enhance the binding affinity, once the binding site can be extended by water molecules mediation and additional hydrogen bonds can be formed [116].

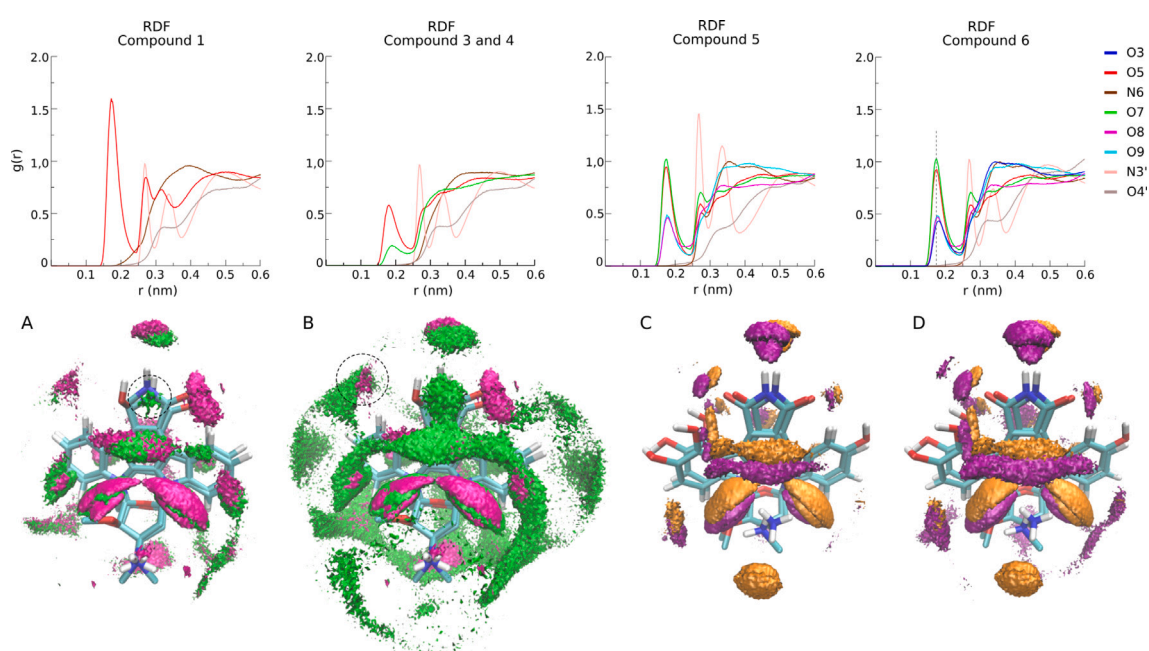
The substitution features at the lactam ring was observed to interfere in the solvation shell surrounding the amide atoms, suggesting a possible influence on both accessibility and strength of polar interactions of this ring and target molecules of STA-analogs. The inclusion of a second H-bond acceptor at the ring reduces the 5-oxygen interaction with the solvent by 3 kJ mol $^{-1}$ , while adds an interaction of at least  $7.69 \pm 0.04$  kJ mol $^{-1}$  at 7-hydroxyl (Table 3) and reflects in the RDF decline around 5-oxygen, in compounds **3** and **4**, as presented in Fig. 9. On the other hand, desolvation costs increasing by 5 kJ mol $^{-1}$  at 6-amine group is observed together the symmetry of maleimide ring in compounds **5** and **6**. The additional 7-carbonyl represents itself at least  $10.68 \pm 0.02$  kJ mol $^{-1}$  (Table 3) and contributes to define the first solvation layer of 6-nitrogen before 0.3 nm, as presented in Fig. 9. One second H-bonding group disrupts the solvent shell surrounding

**Table 3**

Desolvation enthalpies of the STA-analog chromophore suggest competitive or additive effects between nearby polar centers for the interaction with water.



$\Delta G_{\text{HB}} \pm \text{sd}$ (kJ mol <sup>-1</sup> )							
Analog	3-OH	5=O	6-NH	7-OH	7=O	8-OH	9-OH
1	–	11.07 ± 0.02	8.61 ± 0.03	–	–	–	–
3	–	7.92 ± 0.02	7.69 ± 0.02	7.99 ± 0.02	–	–	–
4	–	7.97 ± 0.02	7.72 ± 0.01	7.69 ± 0.04	–	–	–
5	–	9.99 ± 0.02	13.18 ± 0.04	–	10.68 ± 0.02	9.32 ± 0.06	7.94 ± 0.03
6	7.72 ± 0.03	9.90 ± 0.02	13.16 ± 0.04	–	10.69 ± 0.04	9.35 ± 0.06	7.97 ± 0.05



**Fig. 9.** Radial Distribution Functions (RDF) and Spatial Distribution Functions (SDF) of water. The SDF of compound 1 (magenta) and compound 3 (green) show solvation changes when 7-hydroxyl group is present, using isovalue 11 (A), while water bulk near the 5-carbonyl at compound 3 just appears with isovalue 9 (B). The solvation shell reduction around 5-carbonyl is evidenced in the RDF decline in compounds 3 and 4. The dotted circles show regions compatible with the location of coordinate water empirically described by X-ray diffraction. The SDF of compound 5 (orange) shows an additional water bulk around the carbohydrate amine group, when compared to that of the compound 6 (purple), using isovalue 14 (C), while RDF shows increase of solvation at 3'-nitrogen atom. The solvation shell around the 3-hydroxyl group at compound 6 just appears if expanded its SDF, using isovalue 12 (D), while the RDF shows the distance magnitude. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the lactam ring, thus, adding a hydroxyl or carbonyl group at R7 increases up to 14 kJ mol<sup>-1</sup> the global contribution for interactions with the solvent, which may help to understand the structure–activity relationship (SAR) for this group of compounds, particularly the effect of these modifications on STA-analogs activities.

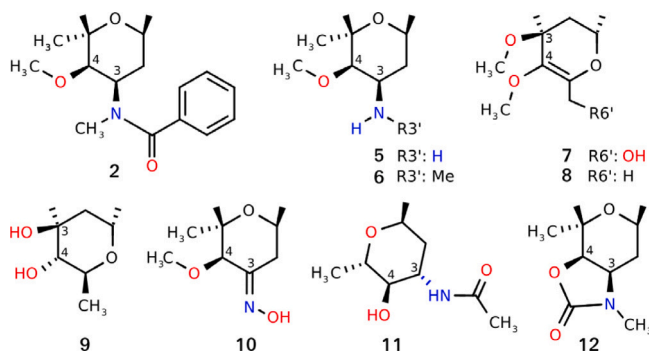
On the other hand, the inclusion of hydroxyl groups at the chromophore seems to be able to influence the ligand solvation reinforcing the interaction of the solvent with the maleimide polar centers, especially in the compounds 5 and 6. The proximity of the 8-hydroxyl to 7-carbonyl appears to trap the water shell between the groups, raising the desolvation cost in maleimide polar centers, as presented in Table 3. Farther from the maleimide ring, the first solvation layer of 3-hydroxyl at compound 6 is slightly displaced from the polar center, as shown in Fig. 9 by the dotted line in the RDF and after expansion of the SDF to make the solvent observable around this group (Fig. 9 C and D). This suggests that substitutions on the chromophore possess a small

potential to interfere negatively in the maleimide interactions with the target receptor, but could in turn potentiate such interactions.

Regarding the water bulk distribution around STA-analog hexopyranoses it was observed two different situations. From compounds with two glycosidic bond attaching the carbohydrate moiety, the indole nitrogens and 5'-oxygen have no good availability for H-bonding since they did not showed the first solvation layer well defined until 0.5 nm. In these cases, the availability for H-bonding belongs to the polar centers in the substitution groups, whose particularities are next presented. Differently, the compound 11, presenting just one glycosidic bond, has the 12-amine group available for H-bonding, which includes 13.17 ± 2.19 kJ mol<sup>-1</sup> in the chromophore enthalpy (Table S2). Moreover, this configuration in carbohydrate moiety, allows the 5'-oxygen to include 5.57 ± 0.07 kJ mol<sup>-1</sup> in the interaction with the solvent or target receptor, as presented in Table 4, which shows hexopyranoses desolvation costs.

**Table 4**

Desolvation enthalpies of the STA-analog hexopyranoses show differences in H-bonding affinity, that can reflect the strength to binding the target receptor.



Compound	$\Delta G_{\text{HB}} \pm \text{sd}$ (kJ mol <sup>-1</sup> )				
	3'-nitrogen	3'-oxygen	4'-oxygen	5'-oxygen	6'-oxygen
2	1.96 ± 0.59	12.36 ± 0.54	7.24 ± 0.21	–	–
5	14.24 ± 0.06	–	4.48 ± 0.10	–	–
6	16.72 ± 0.22	–	3.25 ± 0.18	–	–
7	–	3.29 ± 0.96	5.98 ± 0.24	–	10.54 ± 0.02
8	–	4.30 ± 1.10	6.82 ± 0.24	–	–
9	–	11.76 ± 0.03	11.97 ± 0.07	–	–
10	8.96 ± 0.04	12.35 ± 0.06	6.65 ± 0.09	–	–
11	11.54 ± 0.24	10.76 ± 0.31	10.89 ± 0.03	5.57 ± 0.07	–
12	1.76 ± 0.33	7.25 ± 0.04	6.42 ± 0.04	–	–

The measurements include hydrogen when colored in the figure.

About the compound 5, the lack of the methyl group attached to carbohydrate amine increases the solvation concentration around this group. The raising in the RDF peaks is accompanied by a third water bulk around the nitrogen atom (Fig. 9 C and D), resulting from the additional H-bond donor availability which increases the  $\text{Aver}_{\text{HB}}$  in amine group to  $2.77 \pm 0.49$ , when compared to  $1.88 \pm 0.35$  in compound 6 methylamine group (Table S2). However, is arranged a consequent distribution of the desolvation cost between the H-bond donors and acceptor centers, as shown by the enthalpy change in nitrogen atom, from  $16.72 \pm 0.22$  kJ mol<sup>-1</sup> in compound 6 to  $14.24 \pm 0.06$  kJ mol<sup>-1</sup> in compound 5, as presented in Table 4.

Each hydroxyl group in hexopyranoses represent at least 10 kJ mol<sup>-1</sup> as observed in compounds 7, 9, 10, and 11 (Table 4), an important contribution to tightly H-bonding the water, and possibly the target receptor. The proximity between polar centers, belonging to the same substituent, in compounds 10 and 11 seems to cause an additive effect in its water affinity, once the average energy remains high, even in the presence of unsaturation. The same effect was not observed in compound 12, whose heterocycle fused to the hexopyranose ring can decrease the solvent accessibility, possibly due to geometry effects keeping the polar centers near the apolar chromophore and CH<sub>n</sub> groups.

The  $\Delta G_{\text{HB}}$  values obtained for the desolvation of polar atoms in STA-analogs have positive values, indicating the need of an energy gain to break the H-bond with water molecules. In this regard, to improve the affinity and selectivity of STA-analogs, the rational design of derivatives needs to be planned keeping in mind that the interaction energy of the ligand with the target should be large enough to compensate the desolvation enthalpy costs, or the ligand structure should be modified in order to reduce such costs. Still, the displacement of its hydration shell is expected to result in an affinity gain when the ligand binds to the target in a complementary way, that is, hydrophobic regions in the ligand interact with hydrophobic residues in the protein and polar atoms in the ligand interact with polar atoms in the protein [114,117, 118].

The differences of STA-analogs to H-bonding water can be thought as differences in ligand availability and strength to bind the target

receptor. The prediction of interaction with the solvent derived from MD simulations in water is useful to understand the solvent role in pre-complexation states of STA-analog ligands. The crucial role played by hydration had been showed through conformational changes observed when compared crystal and hydrated structures of the ligand [114], what supports the importance of observing the dynamics of STA-analogs in solution and not only consider crystallographic data of these ligands in approaches aiming to develop its application against therapeutic targets.

#### 4. Conclusions

In this article we present theoretical models empirically fitted to quantum mechanics to reproduce, through molecular mechanics, detailed structural information of STA-analogs analyzed by molecular simulations in aqueous solution. The obtained data, regarding the interaction of the molecules with the biological solvent and its molecular flexibility, helps to elucidate features involved in biological activity of the compounds.

The comprehensive puckering study of STA-analog carbohydrates presented, as well the multiple exocyclic groups and its degrees of freedom represent a potential factor to be exploited to expand the therapeutic potential of these compounds. In addition, the transition between conformers, observed in the carbohydrate moiety, expanded the perception of the conformational space of these ligands, and revealed the limiting character of classical experimental techniques, which generally indicate values of conformational averages, as occurs in NMR spectroscopy. The coordinate usage of computational chemistry and experimental data from NMR can provide access to biologically relevant conformations that can guide rational design of compounds or drive preferences for synthesis or biological characterization of new ligands.

The estimates of interaction between ligand and water, presented here in atomic level, constitute valuable information for the rational design of analogs with improved affinity and selectivity. In this sense, the changes regarding the availability to H-bonding, observed for different polar substituents can reflect on effects at complexation free energy due to variation on desolvation energetic costs. Furthermore,

the stereochemistry and small structural variations could dramatically influence the biological activities of STA-analogs, which requests further SAR studies aiming the development of new drugs, as selectivity is an unsolved problem for kinase inhibitors.

Finally, to know in detail the kinase inhibitors dynamic in solution is important not only for the treatment of diseases, but also for constitute scaffold boards in research to further understand the physiological roles of kinases. These structural data provide a systemic framework for rational chemical modification of STA-analog ligands and may contribute to the design of more selective kinase inhibitors.

### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hugo Verli reports financial support was provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ). Hugo Verli reports financial support was provided by Sistema Nacional de Processamento de Alto Desempenho (SINAPAD). Hugo Verli reports was provided by Fundação de Amparo a Pesquisa do Estado do Rio Grande do Sul (FAPERGS). Crisciele Fontana reports was provided by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

### Data availability

Data will be made available on request.

### Acknowledgments

The authors thank the Brazilian research funding agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), Brazil and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), Brazil. This research was developed with the support of the Sistema Nacional de Processamento de alto Desempenho (SINAPAD) and Centro Nacional de Supercomputação (CESUP).

### Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jmgm.2023.108653>.

### References

- [1] C. Sánchez, C. Méndez, J.A. Salas, Indolocarbazole natural products: Occurrence, biosynthesis, and biological activity, *Nat. Prod. Rep.* 23 (6) (2006) 1007–1045, <http://dx.doi.org/10.1039/b601930g>, URL: <http://pubs-rsc-org.ez45.periodicos.capes.gov.br/en/content/articlepdf/2006/np/b601930g>.
- [2] S. Omura, Y. Iwai, A. Hirano, A. Nakagawa, J. Awaya, H. Tsuchya, Y. Takahashi, R. Masuma, A new alkaloid AM-2282 OF Streptomyces origin. Taxonomy, fermentation, isolation and preliminary characterization, *J. Antibiotics* 30 (4) (1977) 275–282, URL: <http://www.ncbi.nlm.nih.gov/pubmed/863788>.
- [3] D. Nettleton, T. Doyle, B. Krishnan, G. Matsumoto, I. Clardy, Isolation and Structure Rebeccamycin - A new antitumor antibiotic from *Nocardia Aerocoligines*, *Tetrahedron Lett.* 26 (34) (1985) 4011–4014, URL: [https://ac.els-cdn.com/S0040403900892801/1-s2.0-S0040403900892801-main.pdf?\\_tid=6776a5a0-f423-4379-85c8-dd0b2c6ac4c8&acdnat=1539975381\\_303a91af7acd5bb4a32ac2a9cc87f44b](https://ac.els-cdn.com/S0040403900892801/1-s2.0-S0040403900892801-main.pdf?_tid=6776a5a0-f423-4379-85c8-dd0b2c6ac4c8&acdnat=1539975381_303a91af7acd5bb4a32ac2a9cc87f44b).
- [4] H. Nakano, S. Omura, Chemical biology of natural indolocarbazole products: 30 years since the discovery of staurosporine, *J. Antibiotics* 62 (1) (2009) 17–26, <http://dx.doi.org/10.1038/ja.2008.4>, URL: <http://www.nature.com/articles/ja20084>.
- [5] S.M. Pimentel-Elardo, S. Kozytzka, T.S. Bugni, C.M. Ireland, H. Moll, U. Hentschel, Anti-parasitic compounds from streptomyces sp. Strains isolated from mediterranean sponges, *Mar. Drugs* 8 (2) (2010) 373–380, <http://dx.doi.org/10.3390/md8020373>, URL: <http://www.mdpi.com/1660-3397/8/2/373>.
- [6] S. Akinaga, K. Sugiyama, T. Akiyama, UCN-01 (7-hydroxystaurosporine) and other indolocarbazole compounds: A new generation of anti-cancer agents for the new century? *Anti-Cancer Drug Des.* 15 (1) (2000) 43–52, URL: <http://www.ncbi.nlm.nih.gov/pubmed/10888035>.
- [7] T. Kaneko, H. Wong, J. Utzig, J. Schurig, T.W. Doyle, Water soluble derivatives of rebeccamycin, *J. Antibiotics* 43 (1) (1990) 125–127, <http://dx.doi.org/10.7164/antibiotics.43.125>, URL: <http://joi.jlc.jst.go.jp/JST.Journalarchive/antibiotics1968/43.125?from=CrossRef>.
- [8] M. Facompré, B. Baldeyrou, C. Bailly, F. Anizon, C. Marminon, M. Prudhomme, P. Colson, C. Houssier, DNA targeting of two new antitumor rebeccamycin derivatives, *Eur. J. Med. Chem.* 37 (12) (2002) 925–932, [http://dx.doi.org/10.1016/S0223-5234\(02\)01423-X](http://dx.doi.org/10.1016/S0223-5234(02)01423-X), [www.elsevier.com/locate/ejmech](http://www.elsevier.com/locate/ejmech), <http://linkinghub.elsevier.com/retrieve/pii/S022352340201423X>.
- [9] A.P. Salas, L. Zhu, C. Sánchez, A.F. Braña, J. Rohr, C. Méndez, J.A. Salas, Deciphering the late steps in the biosynthesis of the anti-tumour indolocarbazole staurosporine: Sugar donor substrate flexibility of the StaG glycosyltransferase, *Mol. Microbiol.* 58 (1) (2005) 17–27, <http://dx.doi.org/10.1111/j.1365-2958.2005.04777.x>, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881644/pdf/nihms205008.pdf>, <http://doi.wiley.com/10.1111/j.1365-2958.2005.04777.x>.
- [10] T. Janosik, A. Rannug, U. Rannug, N. Wahlström, J. Slätt, J. Bergman, Chemistry and properties of indolocarbazoles, *Chem. Rev.* 118 (18) (2018) 9058–9128, <http://dx.doi.org/10.1021/acs.chemrev.8b00186>, URL: <http://pubs.acs.org/doi/10.1021/acs.chemrev.8b00186>.
- [11] K. Pommerehne, J. Walisko, A. Ebersbach, R. Krull, The antitumor antibiotic rebeccamycin—Challenges and advanced approaches in production processes, *Appl. Microbiol. Biotechnol.* 103 (9) (2019) 3627–3636, <http://dx.doi.org/10.1007/s00253-019-09741-y>, URL: <http://link.springer.com/10.1007/s00253-019-09741-y>.
- [12] R.G. Zenkov, L.V. Ektova, O. Vlasova, G. Belitskiy, M.G. Yakubovskaya, K.I. Kirsanov, Indolo[2,3-a]carbazoles: diversity, biological properties, application in antitumor therapy, *Chem. Heterocyclic Compounds* 56 (6) (2020) 644–658, <http://dx.doi.org/10.1007/s10593-020-02714-4>, URL: <https://link.springer.com/10.1007/s10593-020-02714-4>.
- [13] B.K. Tomlinson, M.M. Gallogly, D.M. Kane, L. Metheny, H.M. Lazarus, B.M. William, M.D. Craig, M.J. Levis, B.W. Cooper, A phase II study of midostaurin and 5-azacitidine for untreated elderly and unfit patients with FLT3 wild-type acute myelogenous leukemia, *Clin. Lymphoma Myeloma Leukemia* 20 (4) (2020) 226–233.e1, <http://dx.doi.org/10.1016/j.clml.2019.10.018>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S2152265019320968>.
- [14] J. Gotlib, H.C. Kluin-Nelemans, T.I. George, C. Akin, K. Sotlar, O. Hermine, F.T. Awan, E. Hexner, M.J. Mauro, D.W. Sternberg, M. Villeneuve, A. Huntsman Labe, E.J. Stanek, K. Hartmann, H.-P. Horny, P. Valent, A. Reiter, Efficacy and safety of midostaurin in advanced systemic mastocytosis, *N. Engl. J. Med.* 374 (26) (2016) 2530–2541, <http://dx.doi.org/10.1056/NEJMoa1513098>, URL: <http://www.nejm.org/doi/10.1056/NEJMoa1513098>.
- [15] P.M. Fracasso, K.J. Williams, R.C. Chen, J. Picus, C.X. Ma, M.J. Ellis, B.R. Tan, T.J. Pluard, D.R. Adkins, M.J. Naughton, J.S. Rader, M.A. Arquette, J.W. Fleshman, A.N. Creekmore, S.A. Goodner, L.P. Wright, Z. Guo, C.E. Ryan, Y. Tao, E.M. Soares, S.-r. Cai, L. Lin, J. Dancy, M.A. Rudek, H.L. McLeod, H. Piwnica-Worms, A Phase 1 study of UCN-01 in combination with irinotecan in patients with resistant solid tumor malignancies, *Cancer Chemother. Pharmacol.* 67 (6) (2011) 1225–1237, <http://dx.doi.org/10.1007/s00280-010-1410-1>, URL: <http://link.springer.com/10.1007/s00280-010-1410-1>.
- [16] G.E. Marti, M. Stetler-Stevenson, N.D. Grant, T. White, W.D. Figg, T. Tohnya, E.S. Jaffe, K. Dunleavy, J.E. Janik, S.M. Steinberg, W.H. Wilson, Phase I trial of 7-hydroxystaurosporine and fludarabine phosphate: In vivo evidence of 7-hydroxystaurosporine induced apoptosis in chronic lymphocytic leukemia, *Leukemia Lymphoma* 52 (12) (2011) 2284–2292, <http://dx.doi.org/10.3109/10428194.2011.589547>, URL: <http://www.tandfonline.com/doi/full/10.3109/10428194.2011.589547>.
- [17] E.S. Kim, Midostaurin: First global approval, *Drugs* 77 (11) (2017) 1251–1259, <http://dx.doi.org/10.1007/s40265-017-0779-0>, URL: <http://link.springer.com/10.1007/s40265-017-0779-0>.
- [18] C.D. Blundell, T. Nowak, M.J. Watson, Measurement, Interpretation and Use of Free Ligand Solution Conformations in Drug Discovery, 2016, pp. 45–147, <http://dx.doi.org/10.1016/bs.pmch.2015.10.003>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0079646815000181>.
- [19] R.W. Hoffmann, Flexible molecules with defined shape? conformational design, *Angew. Chem., Int. Ed. Engl.* 31 (9) (1992) 1124–1134, <http://dx.doi.org/10.1002/anie.199211241>, URL: <https://onlinelibrary.wiley.com/doi/10.1002/anie.199211241>.
- [20] P. Csermely, R. Palotai, R. Nussinov, Induced fit, conformational selection and independent dynamic segments: An extended view of binding events, *Nat. Prec.* (2010) <http://dx.doi.org/10.1038/npre.2010.4422.1>, URL: <http://www.nature.com/articles/npre.2010.4422.1>.
- [21] G. Klebe, H.-J. Böhm, Energetic and entropic factors determining binding affinity in protein-ligand complexes, *J. Receptors Signal Transduction* 17 (1–3) (1997) 459–473, <http://dx.doi.org/10.3109/10799899709036621>, URL: <http://www.tandfonline.com/doi/full/10.3109/10799899709036621>.
- [22] C.-e.A. Chang, W. Chen, M.K. Gilson, Ligand configurational entropy and protein binding, *Proc. Natl. Acad. Sci.* 104 (5) (2007) 1534–1539, <http://dx.doi.org/10.1073/pnas.0610494104>, URL: <https://pnas.org/doi/full/10.1073/pnas.0610494104>.

- [23] B. Baum, L. Muley, M. Smolinski, A. Heine, D. Hangauer, G. Klebe, Non-additivity of functional group contributions in protein–ligand binding: A comprehensive study by crystallography and isothermal titration calorimetry, *J. Mol. Biol.* 397 (4) (2010) 1042–1054, <http://dx.doi.org/10.1016/j.jmb.2010.02.007>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0022283610001622>.
- [24] N. Bisht, B.K. Singh, Role of computer aided drug design in drug development and drug discovery, *Int. J. Pharmaceutical Sci. Res.* 9 (4) (2018) 1405–1415.
- [25] T. Behl, I. Kaur, A. Sehgal, S. Singh, S. Bhatia, A. Al-Harrasi, G. Zengin, E.E. Babes, C. Brisc, M. Stoicescu, M.M. Toma, C. Sava, S.G. Bungau, Bioinformatics accelerates the major tetrad: A real boost for the pharmaceutical industry, *Int. J. Mol. Sci.* 22 (12) (2021) 6184, <http://dx.doi.org/10.3390/ijms22126184>, URL: <https://www.mdpi.com/1422-0067/22/12/6184>.
- [26] F.D. Prieto-Martínez, E. López-López, K. Eurídice Juárez-Mercado, J.L. Medina-Franco, Computational drug design methods—Current and future perspectives, *Silico Drug Design*, Elsevier, 2019, pp. 19–44, <http://dx.doi.org/10.1016/B978-0-12-816125-8.00002-X>, URL: <https://linkinghub.elsevier.com/retrieve/pii/B978012816125800002X>.
- [27] J.A. Lemkul, W.J. Allen, D.R. Bevan, Practical considerations for building GROMOS-compatible small-molecule topologies, *J. Chem. Inf. Model.* 50 (12) (2010) 2221–2235, <http://dx.doi.org/10.1021/ci100335w>, URL: <https://pubs.acs.org/doi/10.1021/ci100335w>.
- [28] P.R. Arantes, M.D. Polêto, E.B.O. John, C. Pedebos, B.I. Grisci, M. Dorn, H. Verli, Development of GROMOS-compatible parameter set for simulations of chalcones and flavonoids, *J. Phys. Chem. B* 123 (5) (2019) 994–1008, <http://dx.doi.org/10.1021/acs.jpcc.8b10139>, URL: <http://pubs.acs.org/doi/10.1021/acs.jpcc.8b10139>.
- [29] M.D. Polêto, V.H. Rusu, B.I. Grisci, M. Dorn, R.D. Lins, H. Verli, Aromatic rings commonly used in medicinal chemistry: Force fields comparison and interactions with water toward the design of new chemical entities, *Front. Pharmacol.* 9 (2018) <http://dx.doi.org/10.3389/fphar.2018.00395>, URL: <http://journal.frontiersin.org/article/10.3389/fphar.2018.00395/full>.
- [30] L. Pol-Fachin, V.H. Rusu, H. Verli, R.D. Lins, GROMOS 53A6 <sub>GLYC</sub>, an improved GROMOS force field for hexopyranose-based carbohydrates, *J. Chem. Theory Comput.* 8 (11) (2012) 4681–4690, <http://dx.doi.org/10.1021/ct300479h>, URL: <https://pubs.acs.org/doi/10.1021/ct300479h>.
- [31] B.A.C. Horta, P.T. Merz, P.F.J. Fuchs, J. Dolenc, S. Riniker, P.H. Hünenberger, A GROMOS-compatible force field for small organic molecules in the condensed phase: The 2016H66 parameter set, *J. Chem. Theory Comput.* 12 (8) (2016) 3825–3850, <http://dx.doi.org/10.1021/acs.jctc.6b00187>, URL: <https://pubs.acs.org/doi/10.1021/acs.jctc.6b00187>.
- [32] W.F. van Gunsteren, X. Daura, N. Hansen, A.E. Mark, C. Oostenbrink, S. Riniker, L.J. Smith, Validation of molecular simulation: An overview of issues, *Angew. Chem., Int. Ed. Engl.* 57 (4) (2018) 884–902, <http://dx.doi.org/10.1002/anie.201702945>, URL: <https://onlinelibrary.wiley.com/doi/10.1002/anie.201702945>.
- [33] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery, J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, J.B. Farkas, J.V. Foresman, J.C. Ortiz, D.J. Fox, Gaussian09 Revision D.01, 2013.
- [34] L.J. Weiser, E.E. Santoso, A CGenFF-based force field for simulations of peptoids with both CIS and trans peptide bonds, *J. Comput. Chem.* 40 (22) (2019) 1946–1956, <http://dx.doi.org/10.1002/jcc.25850>, URL: <https://onlinelibrary.wiley.com/doi/10.1002/jcc.25850>.
- [35] C.G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid, J.C. Gumbart, Rapid parameterization of small molecules using the force field toolkit, *J. Comput. Chem.* 34 (32) (2013) 2757–2770, <http://dx.doi.org/10.1002/jcc.23422>, URL: <https://onlinelibrary.wiley.com/doi/10.1002/jcc.23422>.
- [36] C. Oostenbrink, A. Villa, A.E. Mark, W.F. Van Gunsteren, A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6, *J. Comput. Chem.* 25 (13) (2004) 1656–1676, <http://dx.doi.org/10.1002/jcc.20090>, URL: <http://doi.wiley.com/10.1002/jcc.20090>.
- [37] V.H. Rusu, D.E.S. Santos, M.D. Poletto, M.M. Galheigo, A.T.A. Gomes, H. Verli, T.A. Soares, R.D. Lins, Rotational Profiler: A fast, automated, and interactive server to derive torsional dihedral potentials for classical molecular simulations, *J. Chem. Inf. Model.* (2020) <http://dx.doi.org/10.1021/acs.jcim.0c01168>, <https://pubs.acs.org/doi/10.1021/acs.jcim.0c01168>, URL: <https://pubs.acs.org/doi/10.1021/acs.jcim.0c01168>.
- [38] M.D. Hanwell, D.E. Curtis, D.C. Lonie, T. Vandermeersch, E. Zurek, G.R. Hutchison, Avogadro: An advanced semantic chemical editor, visualization, and analysis platform, *J. Cheminformatics* 4 (1) (2012) 17, <http://dx.doi.org/10.1186/1758-2946-4-17>, URL: <https://jcheminf.springeropen.com/articles/10.1186/1758-2946-4-17>.
- [39] L. Pol-Fachin, H. Verli, R.D. Lins, Extension and validation of the GROMOS 53A6 glyco parameter set for glycoproteins, *J. Comput. Chem.* 35 (29) (2014) 2087–2095, <http://dx.doi.org/10.1002/jcc.23721>, URL: <http://doi.wiley.com/10.1002/jcc.23721>.
- [40] M. Karplus, Vicinal proton coupling in nuclear magnetic resonance, *J. Am. Chem. Soc.* 85 (18) (1963) 2870–2871, <http://dx.doi.org/10.1021/ja00901a059>, URL: <https://pubs.acs.org/doi/abs/10.1021/ja00901a059>.
- [41] N.M. O’Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, Open babel: An open chemical toolbox, *J. Cheminformatics* 3 (1) (2011) 33, <http://dx.doi.org/10.1186/1758-2946-3-33>, URL: <https://jcheminf.biomedcentral.com/articles/10.1186/1758-2946-3-33>.
- [42] C. Haasnoot, F. de Leeuw, C. Altona, The relationship between proton-proton NMR coupling constants and substituent electronegativities—I, *Tetrahedron* 36 (19) (1980) 2783–2792, [http://dx.doi.org/10.1016/0040-4020\(80\)80155-4](http://dx.doi.org/10.1016/0040-4020(80)80155-4), URL: <https://linkinghub.elsevier.com/retrieve/pii/0040402080801554>.
- [43] C.A.G. Haasnoot, F.A.A.M. de Leeuw, H.P.M. de Leeuw, C. Altona, The relationship between proton-proton NMR coupling constants and substituent electronegativities. II—conformational analysis of the sugar ring in nucleosides and nucleotides in solution using a generalized Karplus equation, *Organ. Magn. Resonance* 15 (1) (1981) 43–52, <http://dx.doi.org/10.1002/mrc.1270150111>, URL: <https://onlinelibrary.wiley.com/doi/10.1002/mrc.1270150111>.
- [44] M.L. Huggins, Bond energies and polarities I, *J. Am. Chem. Soc.* 75 (17) (1953) 4123–4126, <http://dx.doi.org/10.1021/ja01113a001>, URL: <https://pubs.acs.org/doi/abs/10.1021/ja01113a001>.
- [45] C. Altona, C.A.G. Haasnoot, Prediction of anti and gauche vicinal proton-proton coupling constants in carbohydrates: A simple additivity rule for pyranose rings, *Organ. Magn. Resonance* 13 (6) (1980) 417–429, <http://dx.doi.org/10.1002/mrc.1270130606>, URL: <https://onlinelibrary.wiley.com/doi/10.1002/mrc.1270130606>.
- [46] J. Landström, G. Widmalm, Glycan flexibility: Insights into nanosecond dynamics from a microsecond molecular dynamics simulation explaining an unusual nuclear Overhauser effect, *Carbohydr. Res.* 345 (2) (2010) 330–333, <http://dx.doi.org/10.1016/j.carres.2009.11.003>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0008621509005424>.
- [47] C. Seitz, L. Casalino, R. Konecny, G. Huber, R.E. Amaro, J.A. McCammon, Multiscale simulations examining glycan shield effects on drug binding to influenza neuraminidase, *Biophys. J.* 119 (11) (2020) 2275–2289, <http://dx.doi.org/10.1016/j.bpj.2020.10.024>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0006349520308468>.
- [48] H.J.C. Berendsen, J.R. Grigera, T.P. Straatsma, The missing term in effective pair potentials, *J. Phys. Chem.* 91 (24) (1987) 6269–6271, <http://dx.doi.org/10.1021/j100308a038>, URL: <https://pubs.acs.org/doi/10.1021/j100308a038>.
- [49] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N log(N) method for Ewald sums in large systems, *J. Chem. Phys.* 98 (12) (1993) 10089–10092, <http://dx.doi.org/10.1063/1.464397>, URL: <http://aip.scitation.org/doi/10.1063/1.464397>.
- [50] S. Páll, B. Hess, A flexible algorithm for calculating pair interactions on SIMD architectures, *Comput. Phys. Comm.* 184 (12) (2013) 2641–2650, <http://dx.doi.org/10.1016/j.cpc.2013.06.003>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0010465513001975>.
- [51] Y.M.H. Gonçalves, C. Senac, P.F.J. Fuchs, P.H. Hünenberger, B.A.C. Horta, Influence of the treatment of nonbonded interactions on the thermodynamic and transport properties of pure liquids calculated using the 2016H66 force field, *J. Chem. Theory Comput.* 15 (3) (2019) 1806–1826, <http://dx.doi.org/10.1021/acs.jctc.8b00425>, URL: <https://pubs.acs.org/doi/10.1021/acs.jctc.8b00425>.
- [52] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: A linear constraint solver for molecular simulations, *J. Comput. Chem.* 18 (12) (1997) 1463–1472, [http://dx.doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](http://dx.doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H), URL: <http://doi.wiley.com/10.1002/%28SICI%291096-987X%28199709%2918%3A12%3C1463%3A%3AAID-JCC4%3E3.0.CO;2-B>.
- [53] B. Hess, P-LINCS: A parallel linear constraint solver for molecular simulation, *J. Chem. Theory Comput.* 4 (1) (2008) 116–122, <http://dx.doi.org/10.1021/ct700200b>, URL: <https://pubs.acs.org/doi/10.1021/ct700200b>.
- [54] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.* 126 (1) (2007) 014101, <http://dx.doi.org/10.1063/1.2408420>, URL: <http://aip.scitation.org/doi/10.1063/1.2408420>.
- [55] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.* 52 (12) (1981) 7182–7190, <http://dx.doi.org/10.1063/1.328693>, URL: <http://aip.scitation.org/doi/10.1063/1.328693>.
- [56] G.A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, PLUMED 2: New feathers for an old bird, *Comput. Phys. Comm.* 185 (2) (2014) 604–613, <http://dx.doi.org/10.1016/j.cpc.2013.09.018>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0010465513003196>.
- [57] D. Cremer, J.A. Pople, General definition of ring puckering coordinates, *J. Am. Chem. Soc.* 97 (6) (1975) 1354–1358, <http://dx.doi.org/10.1021/ja00839a011>, URL: <https://pubs.acs.org/doi/abs/10.1021/ja00839a011>.



- [92] P. Moreau, S. Holbeck, M. Prudhomme, E.A. Sausville, Cytotoxicities of three rebeccamycin derivatives in the National Cancer Institute screening of 60 human tumor cell lines, *Anti-Cancer Drugs* 16 (2) (2005) 145–150, <http://dx.doi.org/10.1097/00001813-200502000-00005>, URL: <http://journals.lww.com/00001813-200502000-00005>.
- [93] T. Lakshmanan, D. Sriram, K. Priya, D. Loganathan, On the structural significance of the linkage region constituents of N-glycoproteins: An X-ray crystallographic investigation using models and analogs, *Biochem. Biophys. Res. Commun.* 312 (2) (2003) 405–413, <http://dx.doi.org/10.1016/j.bbrc.2003.10.149>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0006291X0302206X>.
- [94] H. Satoh, S. Manabe, Design of chemical glycosyl donors: Does changing ring conformation influence selectivity/reactivity? *Chem. Soc. Rev.* 42 (10) (2013) 4297, <http://dx.doi.org/10.1039/c3cs35457a>, URL: <http://xlink.rsc.org/?DOI=c3cs35457a>.
- [95] H.B. Mayes, L.J. Broadbelt, G.T. Beckham, How sugars pucker: Electronic structure calculations map the kinetic landscape of five biologically paramount monosaccharides and their implications for enzymatic catalysis, *J. Am. Chem. Soc.* 136 (3) (2014) 1008–1022, <http://dx.doi.org/10.1021/ja410264d>, URL: <https://pubs.acs.org/doi/10.1021/ja410264d>.
- [96] C. Pedebos, L. Pol-Fachin, H. Verli, Unrestrained conformational characterization of Stenocereus Eruca Saponins in aqueous and nonaqueous solvents, *J. Nat. Prod.* 75 (6) (2012) 1196–1200, <http://dx.doi.org/10.1021/np3000393>, URL: <http://pubs.acs.org/doi/10.1021/np3000393>.
- [97] P. Arantes, L. Sachett, C. Graebin, H. Verli, Conformational characterization of ipomotaosides and their recognition by COX-1 and 2, *Molecules* 19 (4) (2014) 5421–5433, <http://dx.doi.org/10.3390/molecules19045421>, URL: <http://www.mdpi.com/1420-3049/19/4/5421>.
- [98] R.B. Kinnel, P.J. Scheuer, 11-Hydroxystaurosporine: A highly cytotoxic, powerful protein kinase C inhibitor from a tunicate, *J. Organ. Chem.* 57 (23) (1992) 6327–6329, <http://dx.doi.org/10.1021/jo00049a049>, URL: <http://pubs.acs.org/doi/abs/10.1021/jo00049a049>.
- [99] P. Furet, G. Caravatti, N. Lydon, J.P. Priestle, J.M. Sowadski, U. Trinks, P. Traxler, Modelling study of protein kinase inhibitors: Binding mode of staurosporine and origin of the selectivity of CGP 52411, *J. Comput. Aided Mol. Des.* 9 (6) (1995) 465–472, <http://dx.doi.org/10.1007/BF00124317>, URL: <http://link.springer.com/10.1007/BF00124317>.
- [100] P.A. Petillo, L.E. Lerner, Origin and Quantitative Modeling of Anomeric Effect, 1993, pp. 156–175, <http://dx.doi.org/10.1021/bk-1993-0539.ch009>, URL: <https://pubs.acs.org/doi/abs/10.1021/bk-1993-0539.ch009>.
- [101] M.D. Walkinshaw, Variation in the hydrophilicity of hexapyranose sugars explains features of the anomeric effect, *J. Chem. Soc. Perkin Trans. 2* (12) (1987) 1903, <http://dx.doi.org/10.1039/p29870001903>, URL: <http://xlink.rsc.org/?DOI=p29870001903>.
- [102] C. Gambacorti-Passerini, M. Zucchetti, D. Russo, R. Frapolli, M. Verga, S. Bungaro, L. Tornaghi, F. Rossi, P. Pioltelli, E. Pogliani, D. Alberti, G. Corneo, M. D'Incalci, Alpha1 acid glycoprotein binds to imatinib (STI571) and substantially alters its pharmacokinetics in chronic myeloid leukemia patients, *Clin. Cancer Res. : Off. J. Am. Assoc. Cancer Res.* 9 (2) (2003) 625–632, URL: <http://www.ncbi.nlm.nih.gov/pubmed/12576428>.
- [103] E.J. Landin, C. Williams, S.A. Ryan, A. Bochel, N. Akter, C. Redfield, R.B. Sessions, N. Dedi, R.J. Taylor, M.P. Crump, The structural basis for high affinity binding of  $\alpha$ 1-acid glycoprotein to the potent antitumor compound UCN-01, *J. Biol. Chem.* 297 (6) (2021) 101392, <http://dx.doi.org/10.1016/j.jbc.2021.101392>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0021925821011984>.
- [104] M. Alexeeva, E. Åberg, R.A. Engh, U. Rothweiler, The structure of a dual-specificity tyrosine phosphorylation-regulated kinase 1A–PKC412 complex reveals disulfide-bridge formation with the anomalous catalytic loop HRD(HCD) cysteine, *Acta Crystallogr. D* 71 (5) (2015) 1207–1215, <http://dx.doi.org/10.1107/S1399004715005106>, URL: <http://scripts.iucr.org/cgi-bin/paper?S1399004715005106>.
- [105] C. Janiak, A critical account on  $\pi$ - $\pi$  stacking in metal complexes with aromatic nitrogen-containing Ligands †, *J. Chem. Soc. Dalton Trans.* (21) (2000) 3885–3896, <http://dx.doi.org/10.1039/b003010o>, URL: <http://xlink.rsc.org/?DOI=b003010o>.
- [106] M. von Itzstein, W.-Y. Wu, G.B. Kok, M.S. Pegg, J.C. Dyason, B. Jin, T. Van Phan, M.L. Smythe, H.F. White, S.W. Oliver, P.M. Colman, J.N. Varghese, D.M. Ryan, J.M. Woods, R.C. Bethell, V.J. Hotham, J.M. Cameron, C.R. Penn, Rational design of potent sialidase-based inhibitors of influenza virus replication, *Nature* 363 (6428) (1993) 418–423, <http://dx.doi.org/10.1038/363418a0>, URL: <http://www.nature.com/articles/363418a0>.
- [107] J.N. Varghese, V. Chandana Epa, P.M. Colman, Three-dimensional structure of the complex of 4-guanidino-Neu5Ac2en and influenza virus neuraminidase, *Prot. Sci.* 4 (6) (1995) 1081–1087, <http://dx.doi.org/10.1002/pro.5560040606>, URL: <https://onlinelibrary.wiley.com/doi/10.1002/pro.5560040606>.
- [108] F.S. Shidmoosavee, J.N. Watson, A.J. Bennett, Chemical insight into the emergence of influenza virus strains that are resistant to relenza, *J. Am. Chem. Soc.* 135 (36) (2013) 13254–13257, <http://dx.doi.org/10.1021/ja405916q>, URL: <https://pubs.acs.org/doi/10.1021/ja405916q>.
- [109] A.M. Lawrie, M.E. Noble, P. Tunnah, N.R. Brown, L.N. Johnson, J.A. Endicott, Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2, *Nat. Struct. Biol.* 4 (10) (1997) 796–801, <http://dx.doi.org/10.1038/nsb1097-796>, URL: <http://www.nature.com/doi/10.1038/nsb1097-796>.
- [110] D. Komander, G.S. Kular, J. Bain, M. Elliott, D.R. Alessi, D.M.F. Van Aalten, Structural basis for UCN-01 (7-hydroxystaurosporine) specificity and PDK1 (3-phosphoinositide-dependent protein kinase-1) inhibition, *Biochem. J.* 375 (2003) 255–262.
- [111] S.T. Toenjes, J.L. Gustafson, Atropisomerism in medicinal chemistry: Challenges and opportunities, *Fut. Med. Chem.* 10 (4) (2018) 409–422, <http://dx.doi.org/10.4155/fmc-2017-0152>, URL: <https://www.future-science.com/doi/10.4155/fmc-2017-0152>.
- [112] C. Chothia, J. Janin, Principles of protein–protein recognition, *Nature* 256 (5520) (1975) 705–708, <http://dx.doi.org/10.1038/256705a0>, URL: <http://www.nature.com/articles/256705a0>.
- [113] N. Shimokhina, A. Bronowska, S.W. Homans, Contribution of Ligand desolvation to binding thermodynamics in a Ligand–Protein interaction, *Angew. Chem.* 118 (38) (2006) 6522–6524, <http://dx.doi.org/10.1002/ange.200602227>, URL: <https://onlinelibrary.wiley.com/doi/10.1002/ange.200602227>.
- [114] C. Browning, E. Martin, C. Loch, J.-M. Wurtz, D. Moras, R.H. Stote, A.P. Dejaegere, I.M. Billas, Critical role of desolvation in the binding of 20-hydroxyecdysone to the ecdysone receptor, *J. Biol. Chem.* 282 (45) (2007) 32924–32934, <http://dx.doi.org/10.1074/jbc.M705559200>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0021925820613023>.
- [115] B. Zhao, M.J. Bower, P.J. McDevitt, H. Zhao, S.T. Davis, K.O. Johanson, S.M. Green, N.O. Concha, B.-B.S. Zhou, Structural Basis for Chk1 Inhibition by UCN-01, *J. Biol. Chem.* 277 (48) (2002) 46609–46615, <http://dx.doi.org/10.1074/jbc.M201233200>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0021925819332764>.
- [116] J.E. Ladbury, Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design, *Chem. Biol.* 3 (12) (1996) 973–980, [http://dx.doi.org/10.1016/S1074-5521\(96\)90164-7](http://dx.doi.org/10.1016/S1074-5521(96)90164-7), URL: <https://linkinghub.elsevier.com/retrieve/pii/S1074552196901647>.
- [117] M. Kolář, J. Fanfrlík, P. Hobza, Ligand conformational and solvation/desolvation free energy in ProteinLigand complex formation, *J. Phys. Chem. B* 115 (16) (2011) 4718–4724, <http://dx.doi.org/10.1021/jp2010265>, URL: <https://pubs.acs.org/doi/10.1021/jp2010265>.
- [118] A. Biela, M. Khayat, H. Tan, J. Kong, A. Heine, D. Hangauer, G. Klebe, Impact of Ligand and protein desolvation on Ligand binding to the S1 pocket of thrombin, *J. Mol. Biol.* 418 (5) (2012) 350–366, <http://dx.doi.org/10.1016/j.jmb.2012.01.054>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0022283612001234>.



## Supplementary Material

Table S1: Values resultings of  $xy$  dispersion and of diedral potential relaxation for the hexopyranoses puckering study in STA-analogs.

${}^3J_{H,H}$	Compound 1					Compound 12				
	NMR	original	10%	1.0%	0.1%	NMR	original	10%	1.0%	0.1%
H1'-H2' <i>a</i>	9.4	5.7±1.0	7.9±1.1	8.4±0.9	8.4±0.9	10.0	5.4±1.0	8.2±1.1	8.5±0.9	8.5±0.9
H1'-H2' <i>e</i>	1.8	2.1±0.8	1.7±1.3	2.1±1.6	2.2±1.6	6.0	2.3±0.8	1.6±1.2	1.6±1.2	1.7±1.2
H2' <i>a</i> -H3'	7.4	6.0±1.2	7.9±2.4	7.0±2.8	7.0±2.8	12.0	4.9±1.2	8.9±1.3	9.1±1.2	9.1±1.2
H2' <i>e</i> -H3'	11.2	1.3±0.5	4.1±4.4	6.9±4.8	6.9±4.8	5.3	1.9±0.7	4.6±3.6	5.7±3.4	5.9±3.3
H3'-H4'	1.6	4.9±0.9	6.4±1.4	6.1±1.7	6.1±1.7	8.8	6.5±0.6	7.8±0.8	7.8±0.8	7.8±0.8

Table S2: Measures of interaction between STA-analogs and water. The average of H-bonds formed during 1  $\mu$ s simulations are shown with the respective lifetime and free-energy of breakage at each polar atom.

Compound	Group	Aver <sub>HB</sub> $\pm$ sd	Lifetime <sub>HB</sub> $\pm$ sd (ps)	$\Delta$ G <sub>HB</sub> $\pm$ sd (kJ/mol)	AC $\chi^2$
<b>1</b>	5-oxygen	2.577 $\pm$ 0.67	14.011 $\pm$ 0.11	11.072 $\pm$ 0.02	0.002
	6-NH	0.916 $\pm$ 0.60	5.195 $\pm$ 0.08	8.613 $\pm$ 0.03	0.007
	12-nitrogen	0.020 $\pm$ 0.01	0.323 $\pm$ 0.04	1.725 $\pm$ 0.28	0.043
	13-nitrogen	0.007 $\pm$ 0.00	0.443 $\pm$ 0.45	2.511 $\pm$ 1.80	0.058
	3'-NH <sub>2</sub> <sup>+</sup>	1.893 $\pm$ 0.34	189.20 $\pm$ 17.52	17.525 $\pm$ 0.22	<0.001
	4'-oxygen	0.023 $\pm$ 0.15	0.410 $\pm$ 0.03	2.320 $\pm$ 0.18	0.026
	5'-oxygen	0.003 $\pm$ 0.00	6.178 $\pm$ 4.17	9.042 $\pm$ 8.20	0.062
<b>2</b>	5-oxygen	2.587 $\pm$ 0.61	14.165 $\pm$ 0.13	11.099 $\pm$ 0.02	0.002
	6-NH	0.914 $\pm$ 0.66	5.212 $\pm$ 0.07	8.621 $\pm$ 0.04	0.008
	12-nitrogen	0.026 $\pm$ 0.01	0.312 $\pm$ 0.04	1.642 $\pm$ 0.32	0.033
	13-nitrogen	0.009 $\pm$ 0.00	0.360 $\pm$ 0.43	1.995 $\pm$ 0.98	0.055
	3'-nitrogen	0.009 $\pm$ 0.09	0.356 $\pm$ 0.11	1.969 $\pm$ 0.59	0.044
	3'-oxygen	1.540 $\pm$ 0.54	23.616 $\pm$ 4.71	12.366 $\pm$ 0.54	0.003
	5'-oxygen	0.006 $\pm$ 0.00	0.439 $\pm$ 3.02	2.489 $\pm$ 5.19	0.058
<b>3</b>	5-oxygen	1.150 $\pm$ 0.65	3.938 $\pm$ 0.03	7.926 $\pm$ 0.02	0.009
	6-NH	0.670 $\pm$ 0.51	3.591 $\pm$ 0.06	7.697 $\pm$ 0.02	0.008
	7-OH	1.045 $\pm$ 0.74	4.044 $\pm$ 0.04	7.992 $\pm$ 0.02	0.016
	12-nitrogen	0.044 $\pm$ 0.00	0.485 $\pm$ 0.05	2.732 $\pm$ 0.16	0.031
	13-nitrogen	0.018 $\pm$ 0.00	0.310 $\pm$ 1.62	1.621 $\pm$ 0.36	0.045
	3'-NH <sub>2</sub> <sup>+</sup>	1.895 $\pm$ 0.34	179.34 $\pm$ 16.23	17.392 $\pm$ 0.21	<0.001
	5'-oxygen	0.004 $\pm$ 0.00	1.007 $\pm$ 4.54	4.544 $\pm$ 5.50	0.061
<b>4</b>	5-oxygen	1.160 $\pm$ 0.64	4.010 $\pm$ 0.03	7.971 $\pm$ 0.02	0.009
	6-NH	0.674 $\pm$ 0.51	3.626 $\pm$ 0.06	7.721 $\pm$ 0.01	0.008
	7-OH	1.079 $\pm$ 0.74	3.584 $\pm$ 0.03	7.692 $\pm$ 0.04	0.013
	12-nitrogen	0.040 $\pm$ 0.01	0.455 $\pm$ 0.05	2.574 $\pm$ 0.10	0.032
	13-nitrogen	0.014 $\pm$ 0.00	0.298 $\pm$ 0.05	1.531 $\pm$ 0.33	0.045
	3'-NH <sub>2</sub> <sup>+</sup>	1.904 $\pm$ 0.33	408.83 $\pm$ 23.24	19.435 $\pm$ 0.12	<0.001
	5'-oxygen	0.004 $\pm$ 0.00	1.133 $\pm$ 5.34	4.837 $\pm$ 5.31	0.061
<b>5</b>	5-oxygen	1.518 $\pm$ 0.58	9.080 $\pm$ 0.46	9.997 $\pm$ 0.02	0.003
	6-NH	0.963 $\pm$ 0.15	32.804 $\pm$ 0.65	13.181 $\pm$ 0.04	<0.001
	7-oxygen	1.678 $\pm$ 0.63	11.972 $\pm$ 0.21	10.682 $\pm$ 0.02	0.004
	8-OH	1.081 $\pm$ 0.53	6.927 $\pm$ 0.16	9.326 $\pm$ 0.06	0.011
	9-OH	1.106 $\pm$ 0.59	3.964 $\pm$ 0.08	7.942 $\pm$ 0.03	0.010
	12-nitrogen	0.008 $\pm$ 0.00	0.299 $\pm$ 0.04	1.532 $\pm$ 0.58	0.047
	13-nitrogen	0.006 $\pm$ 0.00	0.439 $\pm$ 0.04	2.485 $\pm$ 3.92	0.056
	3'-NH <sub>3</sub> <sup>+</sup>	2.772 $\pm$ 0.49	50.592 $\pm$ 15.05	14.243 $\pm$ 0.06	<0.001
	5'-oxygen	0.004 $\pm$ 0.00	1.145 $\pm$ 5.28	4.864 $\pm$ 4.90	0.055
<b>6</b>	3-OH	1.111 $\pm$ 0.56	3.627 $\pm$ 0.04	7.722 $\pm$ 0.03	0.011
	5-oxygen	1.499 $\pm$ 0.59	8.743 $\pm$ 0.47	9.903 $\pm$ 0.02	0.004
	6-NH	0.966 $\pm$ 0.15	32.639 $\pm$ 0.63	13.168 $\pm$ 0.04	<0.001
	7-oxygen	1.679 $\pm$ 0.62	12.040 $\pm$ 0.19	10.696 $\pm$ 0.04	0.004
	8-OH	1.083 $\pm$ 0.54	7.002 $\pm$ 0.17	9.352 $\pm$ 0.06	0.011
	9-OH	1.109 $\pm$ 0.59	4.019 $\pm$ 0.09	7.976 $\pm$ 0.05	0.010
	12-nitrogen	0.007 $\pm$ 0.00	0.292 $\pm$ 0.04	1.478 $\pm$ 0.52	0.048
	13-nitrogen	0.006 $\pm$ 0.00	0.444 $\pm$ 0.05	2.518 $\pm$ 1.09	0.059
	4'-oxygen	0.034 $\pm$ 0.18	0.598 $\pm$ 0.04	3.254 $\pm$ 0.18	0.021

	5'-oxygen	0.003±0.00	1.038±4.89	4.621±6.56	0.056
<b>7</b>	5-oxygen	2.048±0.64	9.617±0.12	10.139±0.03	0.003
	6-NH	0.890±0.60	5.001±1.13	8.518±0.04	0.007
	12-nitrogen	0.027±0.01	0.439±0.04	2.487±0.80	0.035
	13-nitrogen	0.021±0.00	0.303±0.04	1.566±0.29	0.043
	3'-oxygen	0.055±0.22	0.609±0.25	3.299±0.96	0.022
	4'-oxygen	0.178±0.38	1.803±0.18	5.989±0.24	0.027
	5'-oxygen	0.018±0.00	0.392±0.45	2.205±0.42	0.051
	6'-OH	2.191±0.64	11.320±0.11	10.543±0.02	0.003
<b>8</b>	5-oxygen	2.052±0.54	9.633±0.13	10.143±0.03	0.003
	6-NH	0.889±0.60	4.913±1.08	8.474±0.02	0.007
	12-nitrogen	0.050±0.02	0.635±0.03	3.402±0.91	0.033
	13-nitrogen	0.027±0.01	0.330±0.04	1.779±0.30	0.040
	3'-oxygen	0.090±0.03	0.913±0.28	4.302±1.10	0.023
	4'-oxygen	0.247±0.41	2.522±0.20	6.821±0.24	0.025
	5'-oxygen	0.013±0.00	0.519±0.55	2.903±1.42	0.057
	<b>9</b>	5-oxygen	2.047±0.65	9.672±0.13	10.153±0.02
6-NH		0.885±0.60	4.927±1.12	8.481±0.04	0.007
12-nitrogen		0.028±0.00	0.456±0.04	2.581±0.16	0.031
13-nitrogen		0.026±0.00	0.572±0.04	1.817±0.23	0.042
3'-OH		1.289±0.73	18.521±0.24	11.764±0.03	0.004
4'-OH		1.767±0.63	20.153±0.26	11.973±0.07	0.003
5'-oxygen		0.012±0.00	0.451±0.46	2.554±1.03	0.056
<b>10</b>		5-oxygen	2.572±0.60	13.601±0.12	10.998±0.02
	6-NH	0.908±0.63	5.037±0.07	8.536±0.03	0.008
	12-nitrogen	0.029±0.00	0.339±0.04	1.844±0.20	0.034
	13-nitrogen	0.008±0.00	0.339±0.04	1.844±0.20	0.055
	3'-nitrogen	0.986±0.65	5.981±0.11	8.962±0.04	0.010
	3'-OH	2.136±0.75	23.530±0.11	12.357±0.06	0.007
	4'-oxygen	0.185±0.39	2.354±0.25	6.650±0.09	0.030
	5'-oxygen	0.004±0.00	1.090±3.25	4.743±7.62	0.061
<b>11</b>	5-oxygen	2.561±0.60	13.683±0.11	11.013±0.02	0.002
	6-NH	0.912±0.63	5.113±0.06	8.573±0.04	0.008
	12-nitrogen	0.475±0.59	32.692±14.39	13.172±2.19	0.010
	13-nitrogen	0.024±0.00	0.311±0.04	1.631±0.44	0.043
	3'-NH	0.924±0.29	16.920±0.11	11.540±0.24	0.001
	3'-oxygen	1.219±0.75	12.401±0.11	10.769±0.31	0.004
	4'-OH	1.988±0.62	13.048±0.12	10.895±0.03	0.002
	5'-oxygen	0.215±0.41	1.527±4.37	5.578±0.07	0.017
<b>12</b>	5-oxygen	2.576±0.60	13.429±0.12	10.967±0.02	0.002
	6-NH	0.910±0.63	4.898±0.05	8.466±0.03	0.008
	12-nitrogen	0.026±0.00	0.354±0.05	1.954±0.11	0.043
	13-nitrogen	0.011±0.00	0.397±0.04	2.238±0.60	0.058
	3'-nitrogen	0.008±0.09	0.329±0.04	1.769±0.33	0.047
	4'-oxygen	0.386±0.50	2.145±0.04	6.420±0.04	0.020
	5'-oxygen	0.004±0.00	1.144±4.51	4.861±9.02	0.055
	1''-oxygen	0.990±0.07	3.002±0.05	7.253±0.04	0.021

## 5.2 Capítulo II: Dinâmica Conformacional e da Afinidade de Interação com o Solvente de Rebecamicina e Derivados

### 5.2.1 Preâmbulo

A atividade de bloqueio do ciclo celular mediada pelos ICZs rebecamicina e análogos está possivelmente associada a três domínios funcionais da estrutura dos ligantes. No cromóforo, o anel maleimida constitui um domínio de interação com a topoisomerase, enquanto o sistema de caráter apolar composto por grupos indolo e carbazola, o domínio de ligação intercalante de DNA. O domínio de ligação ao sulco de DNA e estabilização do complexo seria atribuído à porção sacarídica. As propriedades terapêuticas antitumoral, antibacteriana e antifúngica são resultado dessa capacidade intercalante ao DNA e inibitória de topoisomerasas, originando complexos do ligante ICZ intercalando o DNA clivado ligado à enzima. Enquanto o sistema apolar de anéis desempenha um papel importante ao intercalar o DNA, por outro lado reflete em dificuldades na distribuição tecidual de rebecamicina e análogos e na sua capacidade de atingir o alvo terapêutico, dentro da célula, o que tem prejudicado o avanço clínico desses fármacos.

Para melhorar a farmacocinética e farmacodinâmica desses ligantes, e assim promover a obtenção de fármacos mais eficazes, o CADD pode oferecer informações valiosas sobre variações estruturais associadas à sua complexação ao alvo. Assim, através de simulações de dinâmica molecular, avaliamos a flexibilidade de rebecamicina e análogos em solução biológica e sua interação com a água para investigar características estruturais e conformacionais envolvidas em suas propriedades bioativas. A abordagem empírica do campo de força GROMOS permite reproduzir, em fase condensada, propriedades físico-químicas de sistemas moleculares após parametrização baseada em cálculos de mecânica quântica.

A presença de substituintes de halogênio em rebecamicina e análogos refletiu no aumento da afinidade por interações eletrostáticas com a água, a partir de grupos polares vizinhos. A presença de halogênio em posições específicas, mostrou favorecer a manutenção da conformação bioativa desses ligantes, associada à geometria da ligação glicosídica. Com relação à porção sacarídica, a preferência conformacional de diferentes apresentações da  $\beta$ -glicose por  ${}^4C_1$  sugere a prevalência desse confômero em solução e diverge de dados cristalográficos envolvendo um análogo de rebecamicina. Assim, a conformação  ${}^4C_1$  da  $\beta$ -glicose deve ser considerada no planejamento racional de novos bloqueadores do ciclo celular baseados em rebecamicina, uma vez que representa um elemento chave para a inibição da topoisomerase.

Os resultados referentes ao Capítulo II são apresentados a seguir, sob a forma de manuscrito científico, preparado seguindo as normas do *Journal of Chemical Information and Modeling*.

# Structural evaluation of rebeccamycin analogs and its implication on topoisomerase inhibition

Crisiele Fontana,<sup>†</sup> João Luiz de Meirelles,<sup>†</sup> and Hugo Verli<sup>\*,†,‡</sup>

<sup>†</sup>*Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre*

<sup>‡</sup>*Current address: Av. Bento Gonçalves, 9500, Porto Alegre, RS, Caixa Postal 15005, CEP 91501-970, Brazil*

E-mail: hverli@cbiot.ufrgs.br

Phone: +55 51 3308 7770

## Abstract

Therapeutic properties such as antitumor, antibacterial, and antifungal are results of rebeccamycin and analogs (REB-analogs) DNA intercalative capacity and topoisomerase inhibition, disrupting the cellular cycle. Meanwhile, the apolar structure of REB-analogs reflect in difficulties in tissue distribution and arrival at the therapeutic target, hindering its clinical advance. To improve its pharmacokinetics and pharmacodynamics and obtain more effective drugs, computer-aided drug design can offer valuable information about structural changes in these ligands associated to its target complexation. Hence, through molecular dynamics simulations, we assessed the REB-analogs flexibility in biological solution, and their interaction with water to probe conformational and structural features involved to their bioactive properties. The empirical approach of GROMOS force field allows to reproduce in condensed phase, physicochemical properties of molecular systems after parameterization based on quantum mechanics. The presence of halogen substituents in REB-analogs was shown to enhance the polar groups water affinity, while favoring the maintenance of the bioactive

conformation of these ligands. The observed conformational preference of  $\beta$ -glucose for  ${}^4C_1$  should be considered in rational design of new analogs, since it represents a key element for topoisomerase inhibition.

## Introduction

A planar ring system consisting of indole and carbazole elements compose the characteristic structure of an heterocyclic class of compounds called indolocarbazoles (ICZs). The first member of this compound class were found in streptomycetes<sup>1</sup> and subsequently isolated from numerous representatives of flora and fauna.<sup>2,3</sup> The indolo[2,3-a]carbazoles, are the most extensive and biologically significant subclasses of ICZs, and include rebeccamycin (**1**) as a key compound, which was later isolated from culture of actinomycete *Nocardia aerocolonigenes* (*Saccharothrix aerocolonigenes*).<sup>4</sup>

The ICZs constitute an unique class of indole alkaloids, of natural or synthetic origin, presenting a variable number of therapeutic properties such as antitumor, antibacterial, antifungal, antiviral,<sup>4-7</sup> antiparasitic,<sup>8</sup> and neuroprotective<sup>9</sup> activities. Right after isolated, the compound **1** showed its cytostatic properties. It inhibited the growth of human lung adenocarcinoma cells, being associated with the formation of single-stranded DNA breaks,<sup>10</sup> and stabilizes the covalent enzyme and DNA intermediate complex, when DNA is cut by topoisomerase I (Top1).<sup>11</sup> After that, rebeccamycin analogs (REB-analogs) have been identified and a number of derivatives have been synthesized.<sup>2,3,12-16</sup>

The characteristic features of the REB-analogs structure is the presence of the maleimide ring in the planar ring system and one *N*-glycosidic bond, mostly binding a  $\beta$ -glucose.<sup>3</sup> Compound **1** is the lead compound of the series, however, due to its insolubility in water, only its analogs have been further investigated. Its analogs intercalate into DNA and through topoisomerases inhibition, can disrupt the cellular cycle. Among these compounds, edotecarin (**2**) was studied as an antitumor drug in phase I and II clinical trials, however it was not effective against tumors of various histogenesis, and the tests were terminated.<sup>17,18</sup> The synthetic

derivative, becatecarin (**3**), showed moderate activity in phase II studies. Although clinical benefits were observed, the hematologic toxicity led to the dose reduction for the majority of patients<sup>19</sup> and in general its effectiveness was lower than that of existing drugs.<sup>20,21</sup> Another study involving compound **3** was not continued due to insufficient drug supply.<sup>22</sup>

More than 1200 ICZ derivatives have been generated via total synthesis or semisynthesis,<sup>23,24</sup> aiming to improve the compound pharmacokinetics and to obtain more water-soluble drugs, through adding hydrophilic substitutions on the maleimide nitrogen or changing saccharide extensions.<sup>25</sup> Thus, the analysis of the contribution from each structural variation in the compounds behavior in biological solution can be greatly favored by the use of computational approaches such as molecular dynamics (MD). The observation of the conformational ensemble and dynamics of chemical entities, generally inaccessible through most classical structural biology experiments, constitute an important tool to aid the rational design and optimization of ligands.<sup>26,27</sup> However, it is important to notice that the development of a parameters set, specially to describe a selected compounds group, through careful calibration can offer more consistent information about the dynamic of small ligands in a given chemical environment.<sup>28</sup> In this sense, proper energies of interaction and conformational profiles in condensed phase had already been observed in previous works from our group.<sup>29-31</sup>

The empirical approach, i.e. based on the fitting to a set of selected properties from experimental data, of GROMOS force field parameters development process, is the chosen strategy.<sup>32-34</sup> The GROMOS philosophy marries very well with strategies to calibrate torsional potential and partial atomic charges of molecules in order to reproduce, in condensed phase, physicochemical properties and the dynamics of small molecules in solution. Thus, through the development of specific parameters for the GROMOS force field, the current analysis aims to assess the conformational ensemble of REB-analog compounds, in aqueous solution, and to observe their interaction with explicit water. From these observations, it is possible to probe, at atomic level, how structural determinants of the molecular flexibility and availability to interaction can change aspects in the complexation to the therapeutic

target.

## Methods

### Compounds topology building

Specific parameters, as relative energy associated with each dihedral angle and partial atomic charges, were calculated by quantum mechanics (QM) and then fitted on molecular mechanics (MM) aiming to better characterize REB-analogs dynamics in biological solution through GROMOS force field. After energy optimization of the geometries, Gaussian 09<sup>35</sup> calculated the potential energy associated with the rotation of each dihedral angle through Møller–Plesset perturbation theory at second-order level (MP2) under 6-31G\* basis set, while GROMACS 5.1 computed the dihedral potentials on MM using GROMOS53A6 force field.<sup>34</sup> The QM torsional parameters were fitted to reproduce on MM the dihedral rotational profile through the Rotational Profiler server.<sup>36</sup> The partial atomic charges were obtained from QM calculations under the same theory and basis, in implicit solvent, i.e. polarizable continuum model (PCM), using Restrained Electrostatic Potential (RESP) charges. Atomic charges of the chromophore were adjusted to better reproduce the quantum dipole moment vector on MM through a tool based on least-squares fit solution, previously developed.<sup>29</sup> Charge groups were defined using aromatic rings as reference and attending the chemical transferability into functional groups. The input geometries of REB-analogs were built using the molecular editor Avogadro.<sup>37</sup> Further topological parameters to describe atom types, bond stretching, bond-angle bending, improper dihedral deformation, and Lennard-Jones potentials, to support the MM analyses, have been provided by GROMOS53A6<sup>34</sup> and GROMOS53A6<sub>GLYC</sub><sup>31</sup> force fields and by a dataset dedicated to aromatic rings.<sup>30</sup>



## Production simulation settings

MD simulations of REB-analogs were performed using GROMACS 5.1 suite in conjunction with the GROMOS53A6<sub>GLYC</sub><sup>31,38</sup> force field during 1  $\mu$ s. Each molecule was placed in a dodecahedron box and solvated with the extended simple point charge (SPC/E) water model.<sup>39</sup> The calculations were carried out using integration steps of 2 fs, under periodic boundary conditions, after the system energy minimization was reached by *steepest descents* algorithm. The electrostatic interactions were treated by the particle-mesh Ewald (PME) algorithm<sup>40</sup> and the van der Waals contacts with the Verlet *cutoff-scheme*.<sup>41</sup> All covalent bonds lengths in the system were constrained by the LINCS method,<sup>42,43</sup> while the temperature (298 K) and pressure (1 bar) were kept constant, applying V-rescale thermostat<sup>44</sup> and Parrinello-Rahman barostat.<sup>45</sup>

Well-tempered metadynamics were the approach employed to enhance the sampling of the REB-analogs conformational ensemble. PLUMED library, version 2.5<sup>46</sup> was interfaced to GROMACS and carried out calculations performed during 500 ns using until two collective variables (CVs) and bias factor of 10. To calculate the torsion of glycosidic bond, the angular coordinate  $\phi$  was used as CV setting the Gaussian width ( $\sigma$ ) to 0.35 rad, the deposition interval was 500 fs with a starting Gaussian height ( $h$ ) of 1.2 kJ/mol, which corresponds to a deposition rate ( $\omega$ ) of 2.4 J/mol fs. To calculate the puckering of the sugar, the coordinates  $\theta$  and  $\phi$  were used as CVs with  $\sigma = 0.1$  rad and  $h = 0.5$  kJ/mol, with  $\omega = 1.0$  J/mol fs. The errors ( $\epsilon$ ) related to CVs were calculated by block-averaging analysis and comprised up to 3% of the value of the barrier. Estimates of free energy as a function of the dihedral angle and the free energy landscape (FEL) of the sugar were built using the *sum hills* utility and Gnuplot plotting package.

## Assessing conformational ensemble and water interaction

The compounds conformational ensemble was assessed through the observation of the rotational profile of dihedral angles at substituents and glycosidic bond, beyond the sugar pucker-

ing. Measures of the dihedral rotational degrees of freedom were obtained from GROMACS *angle* tool, while the distances between atoms establishing intramolecular interactions, from the *pairdist* tool. Both values are presented by average  $\pm$  standard deviation (sd). The sugar conformation populated during the simulations were identified through a protocol based on the Cremer–Pople angular coordinates  $\theta$  and  $\phi$ .<sup>47</sup> We refer to specific carbohydrates carbon and oxygen atoms by the IUPAC numbering designation.<sup>48</sup>

The REB-analogs interaction with the hydration water was analyzed to probe structural features involved in the hydrophilic character of the molecules. The water accessibility to establish hydrogen bonds (H-bonds) with polar atoms was observed by solvent Radial Distribution Function (RDF), using the *rdf* tool. Heteroatoms that showed first solvation shell within 0.5 nm were evaluated about their measures of interaction with aqueous solvent obtained from *hbond* tool. The H-bond averages ( $Aver_{HB}$ ) are shown through average  $\pm$  sd, while correlation between average lifetime ( $Lifetime_{HB}$ ) and average energy costs to H-bond breakage ( $\Delta G_{HB}$ ) are shown through  $\chi^2$ .<sup>49,50</sup>

## Results and discussion

In order to describe the molecular flexibility and the energies of interaction with the aqueous solvent of REB-analogs, specific topological parameters were empirically developed in accordance with GROMOS philosophy. The partial atomic charges, derived from QM and fitted to classical mechanics by a least-squares fit solution,<sup>29</sup> are shown in Figure 1, which presents by different colors how the charge groups were distributed.

To describe the rotational bonds, novel dihedral potentials were calculated through MM and fitted to the corresponding profiles obtained by QM. The structures used to derive both potentials together with the respective energy curves are shown in Figure 2. The final parameters from the fitting, applied to the topologies building, are presented in Table 1.

Using the new parameters to dihedral angles and partial atomic charges assigned to shared

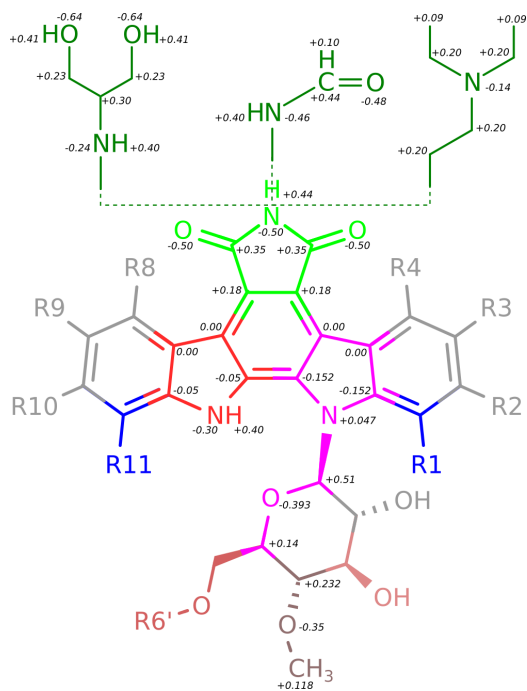


Figure 1: Representation of the new partial atomic charges for REB-analogs. The charges were derived from quantum mechanics and fitted to classical mechanics by a least-squares fit solution. The different colors represent the charge groups used to build the topologies.

Table 1: Novel parameters for dihedral potentials generated to describe the REB-analogs flexibility through molecular mechanics.

	Dihedral angle	$k_\phi$ (kJ mol <sup>-1</sup> )	$\delta$	$n$
a	CH <sub>3</sub> -CH-O-CH <sub>3</sub>	-3.132	0	0
	CH <sub>3</sub> -CH-O-CH <sub>3</sub>	5.012	0	3
b	C-NR-CH <sub>2</sub> -CH <sub>3</sub>	0.146	0	1
c	NH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -NH <sub>2</sub>	2.860	120	1
	NH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -NH <sub>2</sub>	8.647	0	3
d	C-NR-NH-CH <sub>3</sub>	7.119	0	2
e	N <sub>2</sub> -NH-CH-CH <sub>3</sub>	-4.994	0	0
	N <sub>2</sub> -NH-CH-CH <sub>3</sub>	6.510	0	1
	N <sub>2</sub> -NH-CH-CH <sub>3</sub>	9.641	0	3
f	NR-NH-C=O	35.760	0	0
	NR-NH-C=O	-44.414	0	2

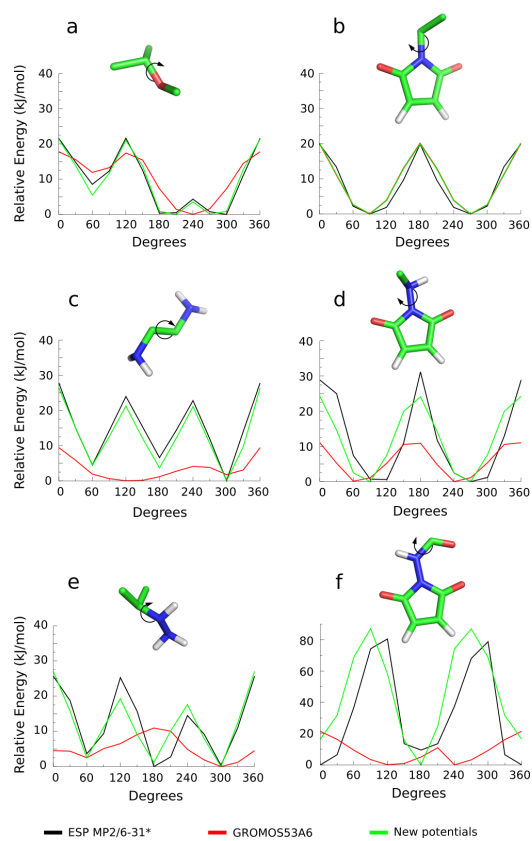
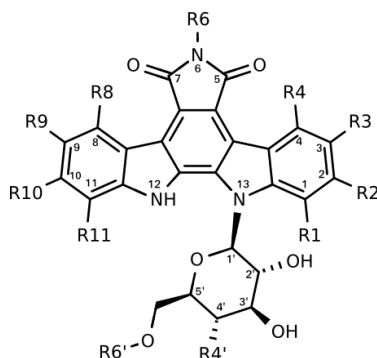


Figure 2: Structures and representative curves used to calculate the novel dihedral potentials for REB-analogs.

functional groups, it was possible to build topologies and carrying out MD simulations of a series of 18 molecules. From these simulations we assessed the conformational ensemble and the water interaction measures of the most researched REB-analog compounds, including those investigated in clinical trials. The chemical structure of REB-analogs and its variable substituents are shown in Table 2.

Table 2: Structures of the REB-analog compounds modeled and analyzed.



	R1	R2	R3	R4	R6	R8	R9	R10	R11	R4'	R6'
1	Cl	H	H	H	H	H	H	H	Cl	OMe	OH
2	H	OH	H	H	NHCH(CH <sub>2</sub> OH) <sub>2</sub>	H	H	OH	H	OH	OH
3	Cl	H	H	H	(CH <sub>2</sub> ) <sub>2</sub> N(CH <sub>2</sub> CH <sub>3</sub> ) <sub>2</sub>	H	H	H	Cl	OMe	OH
4	OH	H	H	H	NHCH(=O)	H	H	H	OH	OH	OH
5	Cl	H	OMe	H	H	H	H	H	Cl	OMe	OH
6	OH	H	H	H	H	H	H	H	OH	OH	OH
7	H	OH	H	H	H	H	H	OH	H	OH	OH
8	Cl	H	H	H	H	H	H	H	H	OMe	OH
9	H	H	H	H	H	H	H	H	H	OMe	OH
10	H	H	H	H	H	H	H	H	H	OH	OH
11	Br	H	H	H	H	H	H	H	Br	OMe	OH
12	H	H	Br	H	H	H	Br	H	H	OMe	OH
13	H	F	H	H	H	H	H	F	H	OMe	OH
14	H	F	H	H	H	H	H	F	H	OH	OH
15	H	H	F	H	H	H	F	H	H	OH	OH
16	H	H	H	F	H	F	H	H	H	OH	OH
17	H	H	F	H	H	H	F	H	H	F	OH
18	H	H	F	H	H	H	F	H	H	OH	NH <sub>2</sub>

The compounds are <sup>1</sup>rebeccamycin, <sup>2</sup>edotecarin, <sup>3</sup>becatecarin, <sup>4</sup>NB-506, <sup>5</sup>9-methoxy-rebeccamycin, <sup>6</sup>ED-110, <sup>7</sup>SA315F, <sup>8</sup>11-deschloro-rebeccamycin, <sup>9</sup>1,11-dideschloro-rebeccamycin, <sup>10</sup>4'-*O*-demethyl-1,11-dideschloro-rebeccamycin, <sup>11</sup>bromorebeccamycin, <sup>12</sup>1,11-dideschloro-3,9-dibromo-rebeccamycin, <sup>13-16</sup>fluoroindolocarbazoles, <sup>17</sup>BMS-250749, and <sup>18</sup>BMS-251873.

## Assessing water interactions

From these analogs, the lead compound **1** has natural poor water solubility, which is why it is not perfectly suitable for use as a drug in the human body, then right after its discovery,

the design of analogs with improved water solubility was required for its clinical use.<sup>51</sup> Some of these compounds, resulting from a effort union with the purpose of discovering analogs of compound **1** more water soluble and with *in vivo* activity, are the compounds **2**, **3**, and compound NB-506 **4**. The 6-N-formylamino derivative, the compound **4**, was more soluble and showed remarkable antineoplastic activities.<sup>52</sup> In addition, it was the basis to design the compound **2**, through the formylamino substituent replacement by a more polar group on the maleimide nitrogen, increasing its aqueous solubility, plasma stability, and reinforcing its cytotoxicity without penalizing the action on Top1.<sup>53</sup> The most studied, however, is the compound **3**. In this semi-synthetic analog, the shift by a 6-N-diethylaminoethyl group improved the pharmacokinetic and the anticancer properties and led it to be pursued as a clinical candidate.<sup>16,21,51,54</sup> These approaches evidenced that changing the substituents on the maleimide ring do not cause a loss of activity, but can improve its water solubility.<sup>55</sup> This observation motivated us to investigate how these structural differences can affect the interaction with the water bulk and its magnitude in terms of Lifetime<sub>HB</sub> and desolvation cost,  $\Delta G_{HB}$ . The complete data of  $A_{verHB}$ , Lifetime<sub>HB</sub> and  $\Delta G_{HB}$  for breakage of all polar atom that showed first solvation shell within 0.5 nm can be consulted in Table S1 (supplementary material).

The analysis of the solvent organization, through RDF, allowed to compare the effect of the maleimide ring, in compound **1**, with its substituted derivatives in 6-nitrogen, and to observe changes in the organization of the solvation water layers, as presented in Figure 3. The compound **2** presented a well organized hydration shell regarding the first, second and third solvation layers of both the oxygen substituents, which is accompanied by a global high desolvation cost, about 42 kJ/mol, due the substituent at 6-nitrogen, as presented in Table 3. Around the compound **4**, the solvent presented a similar distribution in the first layer and more disorganization in the distal layers (Figure 3). In this compound, it is possible that the chemical configuration of the 6-N-formylamino group potentiates the resonant effect in the region, dislocating the nitrogens lone pairs and making them more unavailable for H-bond.

This character can be reflected in smaller Lifetime<sub>HB</sub> and desolvation costs when compared to compound **2** (Table 3).

The inclusion of a protonable 6-N-diethylaminoethyl group, as occurs in compound **3**, provides a very strong interaction with water from this substituent, with Lifetime<sub>HB</sub> of 531.15±171.17 ps and breakage energy mean of 20.08±1.09 kJ/mol, in a similar magnitude that occurs in the ionized nitrogen of the compound **2**, which presents two H-bond donors group (Table 3). Although H-bond life time is very expressive in the 6-N-diethylaminoethyl substituent presence, the solvent RDF showed a displacement of the probability to finding solvent more concentrated to further distance from polar centers, as shown in Figure 3. This displacement of the solvation shell can contribute to tone down the global effect of the significant increase in desolvation costs, since this location penetrates into less organized regions of the water bulk and the H-bond formation dynamics may be affected by the compensatory entropy contribution.<sup>56</sup> The improvement of water solubility in compound **3** altered the physical and pharmacokinetic properties, when compared to the performance of the compound **1**, and was responsible for uncovering the compound **3** as a strong DNA intercalating agent with distal site activity, required for *in vivo* therapeutic usage.<sup>51</sup>

Table 3: Differences in affinity to H-bonding and desolvation costs at REB-analogs substituents can reflect in changes regarding the strength to binding the target receptor.

Compound	Atoms	Lifetime <sub>HB</sub> ±sd (ps)	ΔG <sub>HB</sub> ±sd (kJ/mol)
1	5,7-oxygen	8.21±0.10	9.78±0.03
	6-NH	24.85±0.35	12.49±0.04
2	5,7-oxygen	7.61±0.10	9.53±0.03
	6-nitrogen	0.00±0.00	0.00±0.00
	6α-NH <sup>2+</sup>	579.34±39.94	20.29±0.17
	6δ <sub>a</sub> ,6δ <sub>b</sub> -OH	12.78±0.30	10.67±0.06
3	5,7-oxygen	8.32±0.11	9.78±0.03
	6-nitrogen	0.53±0.02	2.93±0.06
	6γ-NH <sup>+</sup>	531.15±171.17	20.08±1.09
4	5,7-oxygen	6.72±0.10	9.24±0.03
	6-nitrogen	0.27±0.01	1.28±0.03
	6α-NH	12.07±0.15	10.70±0.03
	6γ-oxygen	6.11±0.12	9.01±0.03

As clinically useful compounds may have enhanced water solubility, ICZs with more potent biological activities have been obtained through strategies of structural modifications

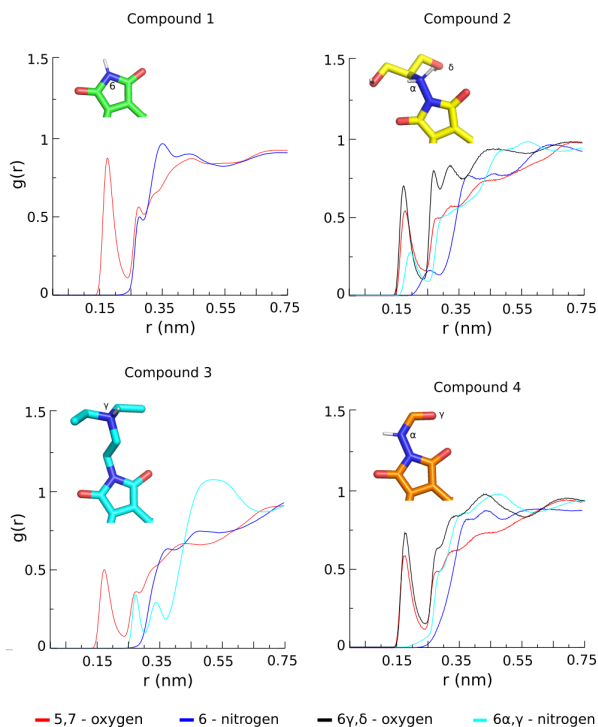


Figure 3: Radial Distribution Function of water around polar atoms at maleimide ring of the compounds rebeccamycin (**1**), edotecarin (**2**), becatecarin (**3**), and NB-506 **4**.

by adding substituents on the planar ring system or introducing a hydrophilic group on the maleimide nitrogen. Approaches to replace the uncharged sugar residue of REB-analogs with a positively charged amino-carbohydrate and to elongate the carbohydrate side chain were also applied.<sup>57</sup> In this sense, our data suggest that also halogen groups attached to the REB-analogs chromophore seem to play an influence on the dynamics of H-bond formation in neighboring polar groups.

The compounds 11-deschloro-rebeccamycin (**8**) and 1,11-dideschloro-rebeccamycin (**9**), showed shorter  $\text{Lifetime}_{\text{HB}}$  and lower  $\Delta G_{\text{HB}}$  breakage on hexopyranose 6'-hydroxyl ( $17.24 \pm 0.75$  ps and  $11.58 \pm 0.06$  kJ/mol,  $17.01 \pm 0.58$  ps and  $11.55 \pm 0.06$  kJ/mol, respectively) when compared to the compound **1**, which has Cl in both positions ( $25.84 \pm 0.81$  ps and  $12.59 \pm 0.08$  kJ/mol). The absence of Cl in position 1, as occur in **9**, showed to affect also H-bonds on 2'-hydroxyl in a minor scale, from  $12.95 \pm 0.13$  ps and  $10.88 \pm 0.02$  kJ/mol to  $11.06 \pm 0.14$  ps and  $10.49 \pm 0.01$  kJ/mol. This results are in the same direction of previous data, that



suggested that the chlorine group can enhance the availability of neighboring heteroatoms to H-bonding.<sup>58</sup> An equal influence due to halogen position was observed when Cl is replaced by Br. The 6'-hydroxyl group on hexopyranose attached to bromorebeccamycin (**11**) showed more affinity with water ( $25.31\pm 0.69$  ps and  $12.30\pm 0.07$  kJ/mol) when compared with 1,11-dideschloro-3,9-dibromo-rebeccamycin (**12**) ( $16.57\pm 0.65$  ps and  $10.89\pm 0.06$  kJ/mol), where Br positions were changed from 1 to 3 and from 11 to 9, followed by a discrete reduction of the water affinity on hexopyranose 2'-hydroxyl of these compounds ( $12.86\pm 0.14$  ps and  $10.86\pm 0.02$  kJ/mol,  $12.25\pm 0.13$  ps and  $9.76\pm 0.03$  kJ/mol, respectively). As covalently-bonded halogen atoms can interact attractively with both nucleophile and electrophile centers,<sup>59</sup> as is the water, our data suggest that the presence of this functional groups are able to improve the molecular affinity to water H-bonding. The sugar hydroxyl groups closer of the halogen (2'-hydroxyl and 6'-hydroxyl) showed higher water affinity averages, which could be favored by the formation of a solvent networking on bulk water around these atoms.<sup>60</sup> Our observation regarding the influence of halogen groups attached to REB-analogs on their water affinity is useful to the drug rational design, once halogenated molecules have been recognized as able to directly bind biomolecular targets via halogen bonds. Then, to know specific and directional non-covalent interactions that these groups can establish, even as to estimate desolvation costs of halogenated ligands is essential.<sup>61,62</sup>

Regarding compounds with fluorine, the halogen on positions 4 and 8 of the fluoroindolo-carbazole **16** was able to increase the maleimide oxygens Lifetime<sub>HB</sub> and  $\Delta G_{HB}$  breakage to  $11.01\pm 0.12$  ps and  $10.47\pm 0.03$  kJ/mol (Table S1), when compared to the compounds with this positions occupied by hydrogens, that presented averages about 8.3 ps and 9.7 kJ/mol (Table 3). The presence of the 4'-fluor group in the compound the BMS-250749 (**17**) changes, on the carbohydrate hydroxyl groups, the water interaction measures in a similar way to those observed in compounds presenting hydroxyl group in the same position. The slight affinity reduction observed on 2',3' and 6' hydroxyl groups attached to the sugar, is a probable consequence of the redistribution of availability to H-bonding in the carbohydrate,

due to the electronegative character of fluorine which can modify the water shell hydration. Although fluorine is generally not considered to form halogen bonds, due its tendency to engage significant *sp* hybridization,<sup>63</sup> the combination of this halogen and oxygen substituents was shown to tune halogen bonds.<sup>64</sup> Such substitutions can dramatically increase the halogen bond strength to a comparable degree of the H-bond strength one by making substitutions of polar groups near the halogen. This tunability of halogen bonds has been shown mainly as result from the modulation of the electrostatic contribution arising from the overall interaction, a property that classifies halogen bonds a very promising mode of interaction for the design of novel pharmaceutical ligands.<sup>64,65</sup>

## Conformational features affecting target receptors interactions

Regarding the structure of REB-analogs and their action mechanism on the Top1-DNA complex blocking, three functional domains have been proposed. While the maleimide ring constitute the Top1-interacting domain, the ICZ ring system is the DNA-intercalative domain.<sup>66,67</sup> The insertion of the planar chromophore between two consecutive base pairs would place the carbohydrate residue into a groove of the double helix. Thus, the sugar moiety constitute the third domain, the DNA-groove binding domain.<sup>23,67,68</sup> It was already observed that the sugar moiety plays an essential role for both DNA interaction and Top1 inhibition, once analogs for which the sugar is attached to the two indole nitrogens (as occurs in staurosporine) generally showed very weak interaction with DNA and produced minimal effects on Top1.<sup>11,69</sup>

Compounds with a single *N*-glycosidic bond, like REB-analogs, were related to be balance between two conformations. The closed pose shows an H-bond between the indole NH group and the sugar endocyclic oxygen (NH–O5' interaction), while at the open one the indole NH group is free to make H-bond with the solvent. Previous empirical studies observed that there is a conformational preference for the closed geometry, against the open one, between REB-analogs, in several solvents.<sup>58,70</sup> Based in that observations, it is possible that these

compounds have few degrees of freedom in the carbohydrate moiety. Thus, to investigate how structural aspects are involved in modulation of the REB-analogs flexibility is useful to infer if the forces that influence its conformational status in solution could favor a preferential conformation when the ligand binds to its biological target. Therefore, we analyzed the REB-analogs flexibility through the observation of the electrostatic intramolecular interactions, the *N*-glycosidic bond dihedral angle rotation, and the carbohydrate puckering.

Regarding the H-bond resulting from the NH–O5' interaction, involved in the closed conformation, we investigate if the chlorine groups presence have some hindrance effect on this electrostatic interaction. From the simulation, the H-bond average distances observed showed equal measures in compounds **1** (with both chlorine group) and **9** (lacks both chlorine group),  $0.21 \pm 0.01$  nm. Analysing the compound **9**, it was also possible to observe the bifurcated H-bond formation between 6'-OH group and the indole NH, at  $0.26 \pm 0.06$  nm, that occurs in absence of the chlorine hindrance.<sup>58</sup> The compound **18**, which presents the 6'-amino group, an exclusively donor of H, preserved the closed conformation, even with a slight distance of the NH–O5' interaction,  $0.30 \pm 0.01$  nm, confirming the main importance of this interaction for the maintenance of the closed conformation. The additional H-bond which results in the bifurcated H-bond, however, contribute to stabilize the closed conformation, that seems to be required to inhibition of Top1,<sup>67,71,72</sup> furthermore, the addition of a methyl group on indole nitrogen led to a loss of the closed conformation, reduced the interaction with DNA, abolished Top1 inhibition, and reflected on decreased cytotoxicity.<sup>70</sup>

To probe the effect of the methyl group to destabilize the closed conformation, its addition at the indole nitrogen in compound **9** was implemented. The presence of the 12-N-methyl group harm the close conformation, probably due to a steric hindrance which shifted the *N*-glycosidic bond dihedral angle to  $87.3 \pm 8.5^\circ$ , when compared to compounds able to establish the NH–O5' interaction, whose averages were about  $60^\circ$ , as presented in Figure 4. The presence of halogen group at the positions 1 and 11 of the chromophore, on the other hand, showed minimal affects on the freedom degrees of the *N*-glycosidic bond, as shown

in Figure 4 in terms of dihedral distribution regarding compounds **1** and **9**. The angular averages were very similar between the compounds with and without chlorine groups, **1** ( $60.5 \pm 7.6^\circ$ ), **8** ( $58.6 \pm 7.4^\circ$ ), **9** ( $54.5 \pm 8.2^\circ$ ), and **10** ( $54.7 \pm 8.2^\circ$ ).

As previous observations suggested that the chlorine group, peri to the *N*-glycosidic bond, would favor the closed conformation<sup>58</sup> we investigate why the *N*-glycosidic bond dihedral angles measures were so similar in presence and absence of halogen groups. Thus, an enhanced sampling of the *N*-glycosidic bond dihedral angle flexibility was performed to ensure the complete conformational space observation. Through metadynamics calculations, it was possible to observe that the halogen presence in positions 1 and 11 (compound **1**) can form steric hindrance to the *N*-glycosidic bond rotation, since only a global minimum near  $60^\circ$  was populated. On the other hand, the absence of halogen groups (compound **9**) resulted in two possible energy minima for the *N*-glycosidic bond rotation, whereas the global minimum when the sampling was enhanced was near  $-120^\circ$ . The high energy barrier that separates the two conformational minima (at least 50 kJ/mol) can explain why the conformers transition was not observed during the unbiased simulation, even in absence of halogen groups (Figure 4). In solution, it is possible that the compounds with no halogen groups in positions 1 and 11 coexist at two conformations, which do not interconvert with each other under usual conditions, since the energy barrier is too high and the transition is not observable, even in very long simulations. This feature can be associated to an atropisomerism-like effect, due to the steric hindrance and/or electronic factors account for the slow rate of interconversion.<sup>73</sup> Furthermore, from input geometries with either *N*-glycosidic bond dihedral angle of  $60^\circ$  or  $-120^\circ$ , the dihedral angle remained restricted at each start angle during the simulation, corroborating the hindrance of dihedral flexibility due the high energy costs associated to the rotation, as represented in Figure 4.

The closed conformation in REB-analogs is possible when *N*-glycosidic bond dihedral angle is nearly to  $60^\circ$ . A compatible geometry of the compound **10** is observed in the ligand bounded to the methyltransferase AtmM, resulting in the complex 6UV6 (to be published),

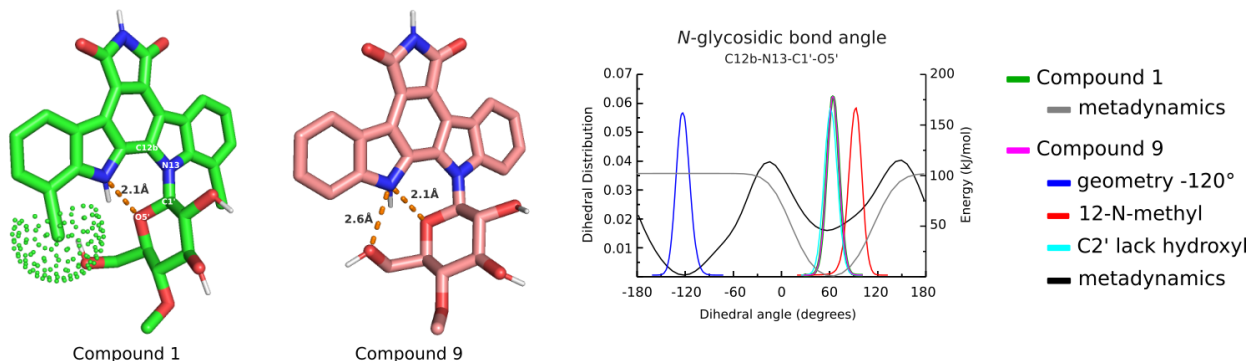


Figure 4: The distribution of the *N*-glycosidic bond angle in rebeccamycin (**1**) and 1,11-dideschloro-rebeccamycin (**9**) show how structural changes affect the conformational dynamics of the bioactive pose, the closed conformation involving the NH–O5' interaction.

available in Protein Data Bank (PDB). The methyltransferase AtmM methylates a variety of ICZs *in vitro*<sup>74</sup> and the compound **10**, in the substrate form, presents the closed conformation in the crystal, with H-bond distance between the NH–O5' interaction of 0.27 nm, *N*-glycosidic dihedral angle of 60.6° and sugar ring conformation is a <sup>4</sup>C<sub>1</sub> (resolution of 2.72 Å). To probe the closed conformation observed during our simulation, we merged its geometry into the ligand electron density map available from the 6UV6 complex. The conformation of the compound **10**, with average angle of 54.7 ± 8.2°, is very close to the crystal (60.6°) and settles well within its electron density map, as shown in Figure 5 A. The addition of the 4'-*O*-methyl is the last step of the REB-analogs biosynthesis.<sup>75</sup> It is likely that the conformational selection of the closed conformation is present in this biosynthesis stage, what could contribute to release the bioactive conformer in solution.

Other crystal structure available in PDB presents the REB-analog SA315F (**7**) complexed to the target Top1 and clivated DNA in entry 1SEU.<sup>76</sup> In this complex, the carbohydrate moiety of the ligand is shown in a different way, the H-bond distance between the NH–O5' interaction is 0.27 nm, the *N*-glycosidic dihedral angle is -65°, while the β-glucose ring is presented through the energetically unfavorable conformation <sup>1</sup>C<sub>4</sub> (resolution of 3.00 Å), although the work lacks structural discussion about the carbohydrate moiety of the ligand. Given that it is not available the electron density map of this complex, an exploratory

overlay was carried out to observe how the ligand complexed to the target Top1–DNA would fit in the electron density map from the complex with the enzyme AtmM, aiming to analyze the carbohydrate moiety. The hexopyranose almost fits within the electron density map, presenting a little region out, depending on the observation angle, as shown in Figure 5 B.

Sugars intrinsic structural features represent difficulties for classical techniques of structures description, such as X-ray crystallography and NMR spectroscopy. The structural complexity of sugar chains, including their heterogeneity and flexibility, combined with the non-availability of these compounds in sufficient quantities and scarcity of reference data have made difficult the structural studies of carbohydrates.<sup>77</sup>

The sugar moiety of REB-analogs is mainly characterized by the presence of a  $\beta$ -glucose attached to the planar ring system through one *N*-glycosidic bond.<sup>3</sup> The sugar puckering analyses of the compounds included in this study was performed, for analogs presenting  $\beta$ -glucose and  $\beta$ -glucose with exocyclic groups variation, to probe if the stereoelectronic effect and changes in electronic density of the substituents could affect the rings conformations and allow conformational changes.<sup>78,79</sup>

The calculation of the  $\beta$ -glucose puckering, obtained from MD simulations, showed a conformational preference for  ${}^4C_1$ , as defined by the Cremer–Pople puckering coordinates  $0^\circ < \theta < 45^\circ$  and  $0^\circ < \phi < 360^\circ$ .<sup>47</sup> The angular averages of the  $\beta$ -glucose were  $\theta = 9.8^\circ \pm 4.7$  and  $\phi = 158.0^\circ \pm 83.8$  and it were very similar to that calculated from the substituted sugars attached to the compounds **17**, with a 4'-fluorine group, and **18**, with a 6'-amino group (data not shown). Through the enhanced sampling, the conformational preference for  ${}^4C_1$ , was confirmed in both compounds, which constitutes the global and single minimum of the FEL ( $\epsilon_\theta < 0.56$ ,  $\epsilon_\phi < 0.25$ ) (Figure 5 C).

Previous empirical studies showed the importance of the sugar moiety, its presence and configuration, for the Top1 interference and formation of intercalating complexes mediated by REB-analogs. Glycosylated ICZs are more potent cytostatic agents than their corresponding aglycone. In addition to that, selectivity toward the cell lines was observed for compounds

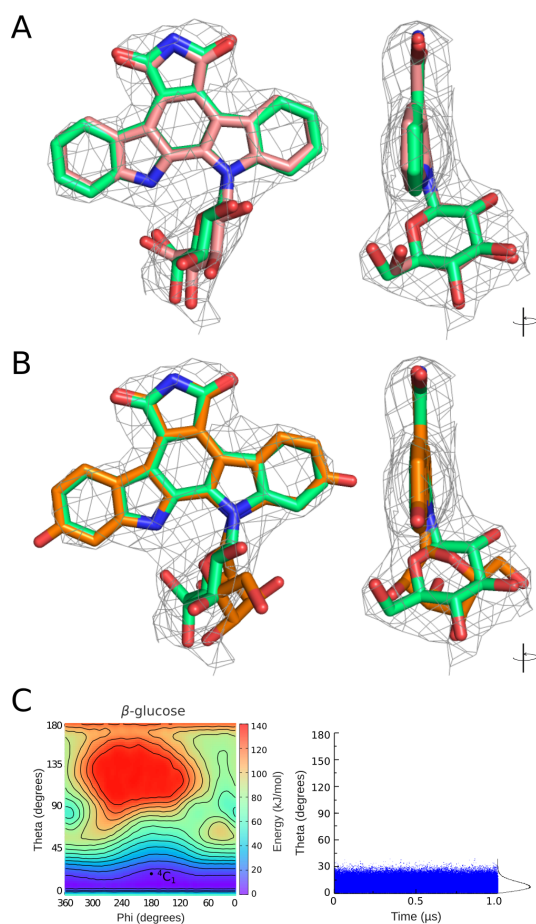


Figure 5: Carbohydrate conformation of REB-analogs free in solution and complexed to proteins. (A) The average conformation of the free compound **10** (pink) settles well in the electron density map of the ligand from the complex 6UV6 (green), while (B) the ligand from the complex 1SEU (orange) presents an outside region. (C) The carbohydrate conformation at <sup>4</sup>C<sub>1</sub> was stable during 1 μs of molecular dynamics and it preferential conformation was confirmed by metadynamics, performed during 500 ns.

presenting the sugar, but not for compounds lacking this substituent.<sup>23,66,80</sup> Compounds possessing an  $\alpha$ -*N*-glycosidic bond do not behave as intercalating agents and have less effect on Top1, while the  $\beta$ -*N*-glycosidic bond represents a key element for its inhibition.<sup>23,81</sup> Due to the evidences associated to carbohydrate geometry and the bioactivity of REB-analogs, to understand the conformational dynamics of this molecular moiety can contribute in the improvement of the efficiency of these ligands.

## Conclusions

Through this study, the models conceived to assess structural information required to improve the therapeutic properties of REB-analogs elucidate chemical and conformational features involved in bioactive processes and might help to expand the therapeutic potential of these compounds. In the same way, changes regarding H-bonds lifetimes with the solvent could potentially affect the complexation free energy due to the desolvation energetic costs. The broad array of REB-analogs conformation represents a potential factor to be exploited aiming the efficacy of these compounds. Together, the dynamic behavior observed for the REB-analogs elucidate structural features involved in its complexation process and can contribute to the clinical advancement of drugs from these analogs.

## Acknowledgement

The authors thank the Brazilian research funding agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo a Pesquisa do Estado do Rio Grande do Sul (FAPERGS). This research was developed with the support of the Sistema Nacional de Processamento de alto Desempenho (SINAPAD) and Centro Nacional de Supercomputação (CESUP).



## References

- (1) Omura, S.; Iwai, Y.; Hirano, A.; Nakagawa, A.; Awaya, J.; Tsuchya, H.; Takahashi, Y.; Masuma, R. A new alkaloid AM-2282 OF Streptomyces origin. Taxonomy, fermentation, isolation and preliminary characterization. *The Journal of antibiotics* **1977**, *30*, 275–82.
- (2) Sánchez, C.; Méndez, C.; Salas, J. A. Indolocarbazole natural products: occurrence, biosynthesis, and biological activity. *Natural Product Reports* **2006**, *23*, 1007–1045.
- (3) Zenkov, R. G.; Ektova, L. V.; Vlasova, O. ; Belitskiy, G. ; Yakubovskaya, M. G.; Kirsanov, K. I. Indolo[2,3-a]carbazoles: diversity, biological properties, application in antitumor therapy. *Chemistry of Heterocyclic Compounds* **2020**, *56*, 644–658.
- (4) Nettleton, D.; Doyle, T.; Krishnan, B.; Matsumoto, G.; Clardy, I. *Isolation and Structure Rebeccamycin - a New Antitumor Antibiotic from Nocardia Aerocoligenes*; 1985; Vol. 26; pp 4011–4014.
- (5) Gribble, G. W.; Berthel, S. J. *A Survey of Indolo [2,3-a] carbazole Alkaloids and Related Natural Products*; 1993; pp 365–409.
- (6) Nakano, H.; Ōmura, S. Chemical biology of natural indolocarbazole products: 30 years since the discovery of staurosporine. *The Journal of Antibiotics* **2009**, *62*, 17–26.
- (7) Knübel, G.; Larsen, L. K.; Moore, R. E.; Levine, I. A.; Patterson, G. M. Cytotoxic, antiviral indolocarbazoles from a blue-green alga belonging to the Nostocaceae. *The Journal of antibiotics* **1990**, *43*, 1236–9.
- (8) Pimentel-Elardo, S. M.; Kozytska, S.; Bugni, T. S.; Ireland, C. M.; Moll, H.; Hentschel, U. Anti-Parasitic Compounds from Streptomyces sp. Strains Isolated from Mediterranean Sponges. *Marine Drugs* **2010**, *8*, 373–380.
- (9) Akinaga, S.; Sugiyama, K.; Akiyama, T. UCN-01 (7-hydroxystaurosporine) and other

- indolocarbazole compounds: a new generation of anti-cancer agents for the new century? *Anti-cancer drug design* **2000**, *15*, 43–52.
- (10) Bush, J. A.; Long, B. H.; Catino, J. J.; Bradner, W. T.; Tomita, K. Production and biological activity of rebeccamycin, a novel antitumor agent. *The Journal of antibiotics* **1987**, *40*, 668–78.
- (11) Yamashita, Y.; Fujii, N.; Murakata, C.; Ashizawa, T.; Okabe, M.; Nakano, H. Induction of mammalian DNA topoisomerase I mediated DNA cleavage by antitumor indolocarbazole derivatives. *Biochemistry* **1992**, *31*, 12069–12075.
- (12) KANEKO, T.; WONG, H.; UTZIG, J.; SCHURIG, J.; DOYLE, T. W. Water soluble derivatives of rebeccamycin. *The Journal of Antibiotics* **1990**, *43*, 125–127.
- (13) Facompré, M.; Baldeyrou, B.; Bailly, C.; Anizon, F.; Marminon, C.; Prudhomme, M.; Colson, P.; Houssier, C. DNA targeting of two new antitumour rebeccamycin derivatives. *European Journal of Medicinal Chemistry* **2002**, *37*, 925–932.
- (14) Salas, A. P.; Zhu, L.; Sánchez, C.; Braña, A. F.; Rohr, J.; Méndez, C.; Salas, J. A. Deciphering the late steps in the biosynthesis of the anti-tumour indolocarbazole staurosporine: sugar donor substrate flexibility of the StaG glycosyltransferase. *Molecular Microbiology* **2005**, *58*, 17–27.
- (15) Janosik, T.; Rannug, A.; Rannug, U.; Wahlström, N.; Slätt, J.; Bergman, J. Chemistry and Properties of Indolocarbazoles. *Chemical Reviews* **2018**, *118*, 9058–9128.
- (16) Pommerehne, K.; Walisko, J.; Ebersbach, A.; Krull, R. The antitumor antibiotic rebeccamycin—challenges and advanced approaches in production processes. *Applied Microbiology and Biotechnology* **2019**, *103*, 3627–3636.
- (17) Saif, M. W.; Diasio, R. B. Edotecarin: A Novel Topoisomerase I Inhibitor. *Clinical Colorectal Cancer* **2005**, *5*, 27–36.

- (18) Saif, M. W.; Sellers, S.; Diasio, R. B.; Douillard, J.-Y. A phase I dose-escalation study of edotecarin (J-107088) combined with infusional 5-fluorouracil and leucovorin in patients with advanced/metastatic solid tumors. *Anti-Cancer Drugs* **2010**, *21*, 716–723.
- (19) Dowlati, A.; Posey, J.; Ramanathan, R. K.; Rath, L.; Fu, P.; Chak, A.; Krishnamurthi, S.; Brell, J.; Ingalls, S.; Hoppel, C. L.; Ivy, P.; Remick, S. C. Phase II and pharmacokinetic trial of rebeccamycin analog in advanced biliary cancers. *Cancer Chemotherapy and Pharmacology* **2009**, *65*, 73–78.
- (20) Burstein, H. J.; Overmoyer, B.; Gelman, R.; Silverman, P.; Savoie, J.; Clarke, K.; Dumadag, L.; Younger, J.; Ivy, P.; Winer, E. P. Rebeccamycin analog for refractory breast cancer: A randomized phase II trial of dosing schedules. *Investigational New Drugs* **2006**, *25*, 161–164.
- (21) Langevin, A.-M.; Bernstein, M.; Kuhn, J. G.; Blaney, S. M.; Ivy, P.; Sun, J.; Chen, Z.; Adamson, P. C. A phase II trial of rebeccamycin analogue (NSC #655649) in children with solid tumors: A Children’s Oncology Group study. *Pediatric Blood & Cancer* **2008**, *50*, 577–580.
- (22) Schwandt, A.; Mekhail, T.; Halmos, B.; O’Brien, T.; Ma, P. C.; Fu, P.; Ivy, P.; Dowlati, A. Phase-II Trial of Rebeccamycin Analog, a Dual Topoisomerase-I and -II Inhibitor, in Relapsed “Sensitive” Small Cell Lung Cancer. *Journal of Thoracic Oncology* **2012**, *7*, 751–754.
- (23) Prudhomme, M. Rebeccamycin analogues as anti-cancer agents. *European Journal of Medicinal Chemistry* **2003**, *38*, 123–140.
- (24) Salas, J. A.; Mé Ndez, C. Indolocarbazole antitumour compounds by combinatorial biosynthesis. *Current Opinion in Chemical Biology* **2009**, *13*, 152–160.
- (25) Thibodeaux, C.; Liu, H.-W.; Thorson, J. *Comprehensive Glycoscience*; Elsevier, 2007; pp 373–396.

- (26) Yu, W.; MacKerell, A. D. *Computer-Aided Drug Design Methods*; 2017; pp 85–106.
- (27) Prieto-Martínez, F. D.; López-López, E.; Eurídice Juárez-Mercado, K.; Medina-Franco, J. L. *In Silico Drug Design*; Elsevier, 2019; pp 19–44.
- (28) Lemkul, J. A.; Allen, W. J.; Bevan, D. R. Practical Considerations for Building GROMOS-Compatible Small-Molecule Topologies. *Journal of Chemical Information and Modeling* **2010**, *50*, 2221–2235.
- (29) Arantes, P. R.; Polêto, M. D.; John, E. B. O.; Pedebos, C.; Grisci, B. I.; Dorn, M.; Verli, H. Development of GROMOS-Compatible Parameter Set for Simulations of Chalcones and Flavonoids. *The Journal of Physical Chemistry B* **2019**, *123*, 994–1008.
- (30) Polêto, M. D.; Rusu, V. H.; Grisci, B. I.; Dorn, M.; Lins, R. D.; Verli, H. Aromatic Rings Commonly Used in Medicinal Chemistry: Force Fields Comparison and Interactions With Water Toward the Design of New Chemical Entities. *Frontiers in Pharmacology* **2018**, *9*.
- (31) Pol-Fachin, L.; Rusu, V. H.; Verli, H.; Lins, R. D. GROMOS 53A6 <sub>GLYC</sub> , an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates. *Journal of Chemical Theory and Computation* **2012**, *8*, 4681–4690.
- (32) Daura, X.; Mark, A. E.; Van Gunsteren, W. F. Parametrization of aliphatic CH<sub>n</sub> united atoms of GROMOS96 force field. *Journal of Computational Chemistry* **1998**, *19*, 535–547.
- (33) VanGunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hunenberger, P. H.; Kruger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: the GROMOS96 Manual and User Guide*; vdf Hochschulverlag AG an der ETH Zurich and BIOMOS b.v.: Zurich, Groningen, 1996; pp 1–1042.

- (34) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry* **2004**, *25*, 1656–1676.
- (35) Frisch, M. J. et al. Gaussian09 Revision D.01. 2013.
- (36) Rusu, V. H.; Santos, D. E. S.; Poletto, M. D.; Galheigo, M. M.; Gomes, A. T. A.; Verli, H.; Soares, T. A.; Lins, R. D. Rotational Profiler: A Fast, Automated, and Interactive Server to Derive Torsional Dihedral Potentials for Classical Molecular Simulations. *Journal of Chemical Information and Modeling* **2020**, acs.jcim.0c01168.
- (37) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics* **2012**, *4*, 17.
- (38) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *European Biophysics Journal* **2011**, *40*, 843–856.
- (39) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *The Journal of Physical Chemistry* **1987**, *91*, 6269–6271.
- (40) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An  $N \log(N)$  method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98*, 10089–10092.
- (41) Páll, S.; Hess, B. A flexible algorithm for calculating pair interactions on SIMD architectures. *Computer Physics Communications* **2013**, *184*, 2641–2650.
- (42) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry* **1997**, *18*, 1463–1472.

- (43) Hess, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation* **2008**, *4*, 116–122.
- (44) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **2007**, *126*, 014101.
- (45) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics* **1981**, *52*, 7182–7190.
- (46) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Computer Physics Communications* **2014**, *185*, 604–613.
- (47) Cremer, D.; Pople, J. A. General definition of ring puckering coordinates. *Journal of the American Chemical Society* **1975**, *97*, 1354–1358.
- (48) McNaught, A. D. Nomenclature of carbohydrates (IUPAC Recommendations 1996). *Pure and Applied Chemistry* **1996**, *68*, 1919–2008.
- (49) Luzar, A.; Chandler, D. Hydrogen-bond kinetics in liquid water. *Nature* **1996**, *379*, 55–57.
- (50) van der Spoel, D.; van Maaren, P. J.; Larsson, P.; Timneanu, N. Thermodynamics of Hydrogen Bonding in Hydrophilic and Hydrophobic Media. *The Journal of Physical Chemistry B* **2006**, *110*, 4393–4398.
- (51) Long, B.; Rose, W.; Vyas, D.; Matson, J.; Forenza, S. Discovery of Antitumor Indolocarbazoles: Rebeccamycin, NSC 655649, and Fluoroindolocarbazoles. *Current Medicinal Chemistry-Anti-Cancer Agents* **2002**, *2*, 255–266.
- (52) Yoshinari, T.; Matsumoto, M.; Arakawa, H.; Okada, H.; Noguchi, K.; Suda, H.; Okura, A.; Nishimura, S. Novel antitumor indolocarbazole compound 6-N-formylamino-12,13-dihydro-1,11-dihydroxy-13-(beta-D-glucopyranosyl)-5H-indolo[2,3-a]pyrrolo[3,4-

- c]carbazole-5,7(6H)-dione (NB-506): induction of topoisomerase I-mediated DNA cleavage and mechanisms of cell li. *Cancer Res* **1995**, *55*, 1310–5.
- (53) Ohkubo, M.; Nishimura, T.; Honma, T.; Nishimura, I.; Ito, S.; Yoshinari, T.; Arakawa, H.; Suda, H.; Morishima, H.; Nishimura, S. Synthesis and biological activities of NB-506 analogues: Effects of the positions of two hydroxyl groups at the indole rings. *Bioorganic Medicinal Chemistry Letters* **1999**, *9*, 3307–3312.
- (54) Nock, C. J.; Brell, J. M.; Bokar, J. A.; Cooney, M. M.; Cooper, B.; Gibbons, J.; Krishnamurthi, S.; Manda, S.; Savvides, P.; Remick, S. C.; Ivy, P.; Dowlati, A. A phase I study of rebeccamycin analog in combination with oxaliplatin in patients with refractory solid tumors. *Investigational New Drugs* **2011**, *29*, 126–130.
- (55) Bailly, C.; Colson, P.; Houssier, C.; Rodrigues-Pereira, E.; Prudhomme, M.; Waring, M. J. Recognition of specific sequences in DNA by a topoisomerase I inhibitor derived from the antitumor drug rebeccamycin. *Molecular pharmacology* **1998**, *53*, 77–87.
- (56) Lumry, R.; Rajender, S. Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: A ubiquitous property of water. *Biopolymers* **1970**, *9*, 1125–1227.
- (57) Rao, B. P. C.; Oliveira, O. N.; Varala, R. *Scope of Selective Heterocycles from Organic and Pharmaceutical Perspective*; InTech, 2016.
- (58) Gilbert, E. J.; Chisholm, J. D.; Van Vranken, D. L. Conformational Control in the Rebeccamycin Class of Indolocarbazole Glycosides. *The Journal of Organic Chemistry* **1999**, *64*, 5670–5676.
- (59) Politzer, P.; Murray, J. S.; Clark, T. Halogen bonding and other  $\sigma$ -hole interactions: A perspective. *Physical Chemistry Chemical Physics* **2013**, *15*, 11178–11189.

- (60) Jong, K.; Hassanali, A. A. A Data Science Approach to Understanding Water Networks Around Biomolecules: The Case of Tri-Alanine in Liquid Water. *The Journal of Physical Chemistry B* **2018**, *122*, 7895–7906.
- (61) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. Halogen bonds in biological molecules. *Proceedings of the National Academy of Sciences* **2004**, *101*, 16789–16794.
- (62) Xu, Z.; Yang, Z.; Liu, Y.; Lu, Y.; Chen, K.; Zhu, W. Halogen Bond: Its Role beyond Drug–Target Binding Affinity for Drug Discovery and Development. *Journal of Chemical Information and Modeling* **2014**, *54*, 69–78.
- (63) Politzer, P.; Lane, P.; Concha, M. C.; Ma, Y.; Murray, J. S. An overview of halogen bonding. *Journal of Molecular Modeling* **2007**, *13*, 305–311.
- (64) Riley, K. E.; Murray, J. S.; Fanfrlík, J.; Řezáč, J.; Solá, R. J.; Concha, M. C.; Ramos, F. M.; Politzer, P. Halogen bond tunability I: the effects of aromatic fluorine substitution on the strengths of halogen-bonding interactions involving chlorine, bromine, and iodine. *Journal of Molecular Modeling* **2011**, *17*, 3309–3318.
- (65) Riley, K. E.; Ford, C. L.; Demouchet, K. Comparison of hydrogen bonds, halogen bonds, C H $\pi$  interactions, and C X $\pi$  interactions using high-level ab initio methods. *Chemical Physics Letters* **2015**, *621*, 165–170.
- (66) Bailly, C.; Riou, J.-F.; Colson, P.; Houssier, C.; Rodrigues-Pereira, E.; Prudhomme, M. DNA Cleavage by Topoisomerase I in the Presence of Indolocarbazole Derivatives of Rebeccamycin †. *Biochemistry* **1997**, *36*, 3917–3929.
- (67) Bailly, C.; Qu, X.; Graves, D. E.; Prudhomme, M.; Chaires, J. B. Calories from carbohydrates: Energetic contribution of the carbohydrate moiety of rebeccamycin to DNA binding and the effect of its orientation on topoisomerase I inhibition. *Chemistry and Biology* **1999**, *6*, 277–286.



- (68) Saulnier, M. G. et al. Discovery of a Fluoroindolo[2,3- a ]carbazole Clinical Candidate with Broad Spectrum Antitumor Activity in Preclinical Tumor Models Superior to the Marketed Oncology Drug, CPT-11. *Journal of Medicinal Chemistry* **2005**, *48*, 2258–2261.
- (69) Anizon, F.; Belin, L.; Moreau, P.; Sancelme, M.; Voldoire, A.; Prudhomme, M.; Ollier, M.; Severe, D.; Riou, J.-F.; Bailly, C.; Fabbro, D.; Meyer, T. Syntheses and Biological Activities (Topoisomerase Inhibition and Antitumor and Antimicrobial Properties) of Rebeccamycin Analogues Bearing Modified Sugar Moieties and Substituted on the Imide Nitrogen with a Methyl Group. *Journal of Medicinal Chemistry* **1997**, *40*, 3456–3465.
- (70) Facompré, M.; Carrasco, C.; Vezin, H.; Chisholm, J. D.; Yoburn, J. C.; Van Vranken, D. L.; Bailly, C. Indolocarbazole Glycosides in Inactive Conformations. *ChemBioChem* **2003**, *4*, 386–395.
- (71) Moreau, P.; Anizon, F.; Sancelme, M.; Prudhomme, M.; Bailly, C.; Carrasco, C.; Ollier, M.; Severe, D.; Riou, J. F.; Fabbro, D.; Meyer, T.; Aubertin, A. M. Syntheses and biological evaluation of indolocarbazoles, analogues of rebeccamycin, modified at the imide heterocycle. *Journal of medicinal chemistry* **1998**, *41*, 1631–40.
- (72) Zhang, G.; Shen, J.; Cheng, H.; Zhu, L.; Fang, L.; Luo, S.; Muller, M. T.; Lee, G. E.; Wei, L.; Du, Y.; Sun, D.; Wang, P. G. Syntheses and Biological Activities of Rebeccamycin Analogues with Uncommon Sugars. *Journal of Medicinal Chemistry* **2005**, *48*, 2600–2611.
- (73) LaPlante, S. R.; Fader, L. D.; Fandrick, K. R.; Fandrick, D. R.; Hucke, O.; Kemper, R.; Miller, S. P. F.; Edwards, P. J. Assessing Atropisomer Axial Chirality in Drug Discovery and Development. *Journal of Medicinal Chemistry* **2011**, *54*, 7005–7022.
- (74) Gao, Q.; Zhang, C.; Blanchard, S.; Thorson, J. S. Deciphering Indolocarbazole and

- Enediylne Aminodideoxypentose Biosynthesis through Comparative Genomics: Insights from the AT2433 Biosynthetic Locus. *Chemistry Biology* **2006**, *13*, 733–743.
- (75) ONAKA, H.; TANIGUCHI, S.-i.; IGARASHI, Y.; FURUMAI, T. Characterization of the Biosynthetic Gene Cluster of Rebeccamycin from *Lechevalieria aerocolonigenes* ATCC 39243. *Bioscience, Biotechnology, and Biochemistry* **2003**, *67*, 127–138.
- (76) Staker, B. L.; Feese, M. D.; Cushman, M.; Pommier, Y.; Zembower, D.; Stewart, L.; Burgin, A. B. Structures of Three Classes of Anticancer Agents Bound to the Human Topoisomerase II DNA Covalent Complex. *Journal of Medicinal Chemistry* **2005**, *48*, 2336–2345.
- (77) Lakshmanan, T.; Sriram, D.; Priya, K.; Loganathan, D. On the structural significance of the linkage region constituents of N-glycoproteins: an X-ray crystallographic investigation using models and analogs. *Biochemical and Biophysical Research Communications* **2003**, *312*, 405–413.
- (78) Satoh, H.; Manabe, S. Design of chemical glycosyl donors: does changing ring conformation influence selectivity/reactivity? *Chemical Society Reviews* **2013**, *42*, 4297.
- (79) Mayes, H. B.; Broadbelt, L. J.; Beckham, G. T. How Sugars Pucker: Electronic Structure Calculations Map the Kinetic Landscape of Five Biologically Paramount Monosaccharides and Their Implications for Enzymatic Catalysis. *Journal of the American Chemical Society* **2014**, *136*, 1008–1022.
- (80) Moreau, P.; Holbeck, S.; Prudhomme, M.; Sausville, E. A. Cytotoxicities of three rebeccamycin derivatives in the National Cancer Institute screening of 60 human tumor cell lines. *Anti-Cancer Drugs* **2005**, *16*, 145–150.
- (81) Ohkubo, M.; Nishimura, T.; Kawamoto, H.; Nakano, M.; Honma, T.; Yoshinari, T.; Arakawa, H.; Suda, H.; Morishima, H.; Nishimura, S. Synthesis and biological activities

of NB-506 analogues modified at the glucose group. *Bioorganic and Medicinal Chemistry Letters* **2000**, *10*, 419–422.

## Supplementary Material

Table S1: Measures of interaction between REB-analogs and water. The average of H-bonds formed during 1  $\mu$ s simulations are shown with the respective lifetime and free-energy of breakage at each polar atom.

Compound	Atom	Aver <sub>HB</sub> $\pm$ sd	Lifetime <sub>HB</sub> $\pm$ sd (ps)	$\Delta$ G <sub>HB</sub> $\pm$ sd (kJ/mol)	AC $\chi^2$
<b>1</b>	O5	1.424 $\pm$ 0.58	8.127 $\pm$ 0.10	9.772 $\pm$ 0.03	0.004
	NH6	0.924 $\pm$ 0.16	24.846 $\pm$ 0.35	12.492 $\pm$ 0.04	<0.001
	O7	1.461 $\pm$ 0.58	8.334 $\pm$ 0.11	9.784 $\pm$ 0.03	0.003
	OH2'	1.584 $\pm$ 0.50	12.950 $\pm$ 0.13	10.877 $\pm$ 0.02	0.003
	OH3'	1.592 $\pm$ 0.48	13.457 $\pm$ 0.23	10.966 $\pm$ 0.04	0.003
	O4'	0.205 $\pm$ 0.42	1.101 $\pm$ 0.02	4.767 $\pm$ 0.05	0.014
	O5'	0.025 $\pm$ 0.01	0.422 $\pm$ 0.02	2.474 $\pm$ 0.11	0.028
	OH6'	1.814 $\pm$ 0.50	25.845 $\pm$ 0.81	12.590 $\pm$ 0.08	0.002
<b>2</b>	OH2	1.075 $\pm$ 0.56	3.691 $\pm$ 0.32	7.765 $\pm$ 0.04	0.011
	O5	0.639 $\pm$ 0.53	7.644 $\pm$ 0.09	9.570 $\pm$ 0.03	0.008
	N6	0.000 $\pm$ 0.00	0.000 $\pm$ 0.00	0.000 $\pm$ 0.00	0.062
	NH <sup>2+</sup> <sub>6<math>\alpha</math></sub>	1.919 $\pm$ 0.31	579.34 $\pm$ 39.94	20.299 $\pm$ 0.17	<0.001
	OH <sub>6<math>\delta</math>a}</sub>	1.509 $\pm$ 0.53	12.013 $\pm$ 0.31	10.691 $\pm$ 0.06	0.003
	OH <sub>6<math>\delta</math>b}</sub>	1.504 $\pm$ 0.53	11.853 $\pm$ 0.30	10.657 $\pm$ 0.06	0.003
	O7	0.640 $\pm$ 0.53	7.549 $\pm$ 0.10	9.539 $\pm$ 0.03	0.008
	O10	1.968 $\pm$ 0.60	13.094 $\pm$ 0.15	10.904 $\pm$ 0.03	0.004
	OH2'	1.611 $\pm$ 0.54	12.502 $\pm$ 0.18	10.734 $\pm$ 0.05	0.007
	OH3'	1.720 $\pm$ 0.50	12.992 $\pm$ 0.15	10.973 $\pm$ 0.03	0.003
	OH4'	1.763 $\pm$ 0.48	10.986 $\pm$ 0.17	10.469 $\pm$ 0.03	0.003
	O5'	0.025 $\pm$ 0.01	0.420 $\pm$ 0.02	2.452 $\pm$ 0.10	0.027
	OH6'	1.687 $\pm$ 0.57	17.941 $\pm$ 0.78	11.685 $\pm$ 0.07	0.002
<b>3</b>	O5	0.829 $\pm$ 0.52	8.358 $\pm$ 0.10	9.791 $\pm$ 0.03	0.004
	N6	0.001 $\pm$ 0.34	0.525 $\pm$ 0.02	2.930 $\pm$ 0.06	0.058
	NH <sup>+</sup> <sub>6<math>\gamma</math></sub>	0.936 $\pm$ 0.19	531.15 $\pm$ 171.17	20.084 $\pm$ 1.09	<0.001
	O7	0.874 $\pm$ 0.52	8.298 $\pm$ 0.11	9.773 $\pm$ 0.03	0.004
	OH2'	1.579 $\pm$ 0.49	12.653 $\pm$ 0.15	10.774 $\pm$ 0.03	0.003
	OH3'	1.581 $\pm$ 0.48	13.271 $\pm$ 0.19	10.948 $\pm$ 0.04	0.003
	O4'	0.201 $\pm$ 0.41	1.036 $\pm$ 0.01	4.616 $\pm$ 0.03	0.014
	O5'	0.026 $\pm$ 0.01	0.431 $\pm$ 0.01	2.532 $\pm$ 0.11	0.028
	OH6'	1.967 $\pm$ 0.50	23.410 $\pm$ 0.75	12.963 $\pm$ 0.09	0.003
<b>4</b>	OH1	1.510 $\pm$ 0.57	12.324 $\pm$ 0.28	10.754 $\pm$ 0.04	0.005
	O5	1.017 $\pm$ 0.60	6.656 $\pm$ 0.10	9.227 $\pm$ 0.03	0.006
	N6	0.022 $\pm$ 0.14	0.270 $\pm$ 0.01	1.283 $\pm$ 0.03	0.036
	NH <sub>6<math>\alpha</math></sub>	0.971 $\pm$ 0.27	12.070 $\pm$ 0.15	10.702 $\pm$ 0.03	0.002
	O <sub>6<math>\gamma</math></sub>	1.226 $\pm$ 0.61	6.108 $\pm$ 0.12	9.014 $\pm$ 0.03	0.008
	O7	1.056 $\pm$ 0.60	6.768 $\pm$ 0.11	9.268 $\pm$ 0.03	0.006
	OH11	1.626 $\pm$ 0.60	12.579 $\pm$ 0.15	10.805 $\pm$ 0.04	0.005
	OH2'	1.600 $\pm$ 0.52	12.063 $\pm$ 0.20	9.702 $\pm$ 0.06	0.006
	OH3'	1.737 $\pm$ 0.52	12.092 $\pm$ 0.17	9.995 $\pm$ 0.03	0.005
	OH4'	1.763 $\pm$ 0.46	11.006 $\pm$ 0.12	10.474 $\pm$ 0.03	0.003
	O5'	0.026 $\pm$ 0.01	0.432 $\pm$ 0.01	2.534 $\pm$ 0.11	0.003
	OH6'	1.779 $\pm$ 0.62	16.369 $\pm$ 0.75	11.458 $\pm$ 0.05	0.002
<b>5</b>	O3	0.292 $\pm$ 0.46	0.771 $\pm$ 0.23	3.884 $\pm$ 0.02	0.013
	O5	1.557 $\pm$ 0.58	9.660 $\pm$ 0.11	10.150 $\pm$ 0.03	0.003
	NH6	0.925 $\pm$ 0.15	25.397 $\pm$ 0.37	12.235 $\pm$ 0.03	<0.001
	O7	1.543 $\pm$ 0.58	8.576 $\pm$ 0.10	9.855 $\pm$ 0.03	0.003

	OH2'	1.601±0.49	12.396±0.14	10.588±0.02	0.007
	OH3'	1.511±0.59	12.635±0.22	10.789±0.04	0.004
	O4'	0.225±0.42	1.125±0.02	4.820±0.05	0.014
	O5'	0.024±0.01	0.398±0.01	2.532±0.10	0.030
	OH6'	1.033±0.50	13.945±0.65	11.060±0.06	0.002
<b>6</b>	OH1	0.655±0.50	2.234±0.30	6.521±0.03	0.004
	O5	1.531±0.58	8.510±0.10	9.836±0.04	0.003
	NH6	0.971±0.16	25.992±0.29	12.604±0.05	<0.001
	O7	1.536±0.58	8.796±0.11	9.918±0.03	0.002
	OH11	0.956±0.56	4.539±0.15	8.278±0.05	0.011
	OH2'	1.704±0.52	12.566±0.18	9.852±0.05	0.005
	OH3'	1.697±0.51	12.273±0.19	9.080±0.04	0.004
	OH4'	1.763±0.46	10.986±0.20	10.469±0.03	0.003
	O5'	0.025±0.01	0.412±0.01	2.487±0.10	0.002
	OH6'	2.023±0.59	16.533±0.69	11.482±0.04	0.003
<b>7</b>	OH2	1.085±0.46	3.959±0.28	7.939±0.03	0.006
	O5	1.543±0.58	8.652±0.32	9.877±0.02	0.003
	NH6	0.926±0.15	24.406±0.32	12.293±0.04	<0.001
	O7	1.547±0.58	8.797±0.12	9.918±0.04	0.003
	O10	1.976±0.59	13.216±0.14	10.622±0.03	0.006
	OH2'	1.797±0.51	13.289±0.16	11.086±0.04	0.004
	OH3'	1.797±0.51	13.289±0.18	11.086±0.03	0.004
	OH4'	1.745±0.46	12.918±0.15	10.952±0.03	0.002
	O5'	0.024±0.01	0.435±0.01	2.437±0.11	0.029
	OH6'	1.734±0.58	17.486±0.75	11.518±0.08	0.004
<b>8</b>	O5	1.418±0.58	8.089±0.11	9.710±0.04	0.004
	NH6	0.925±0.16	25.458±0.34	11.952±0.03	<0.001
	O7	1.456±0.58	8.355±0.11	9.791±0.04	0.004
	OH2'	1.687±0.50	13.100±0.13	10.905±0.03	0.003
	OH3'	1.716±0.49	13.309±0.17	10.945±0.02	0.003
	O4'	0.211±0.42	1.108±0.02	4.782±0.04	0.014
	O5'	0.033±0.01	0.411±0.02	2.323±0.10	0.025
	OH6'	1.774±0.55	17.246±0.75	11.587±0.06	0.002
<b>9</b>	O5	1.430±0.58	8.319±0.12	9.780±0.02	0.003
	NH6	0.024±0.15	24.635±0.18	12.467±0.02	<0.001
	O7	1.461±0.58	8.363±0.10	9.793±0.03	0.003
	OH2'	1.600±0.54	11.058±0.14	10.485±0.01	0.005
	OH3'	1.717±0.49	12.964±0.15	10.880±0.02	0.003
	O4'	0.213±0.42	1.127±0.01	4.825±0.02	0.014
	O5'	0.037±0.01	0.457±0.02	2.585±0.11	0.024
	OH6'	1.750±0.55	17.009±0.58	11.553±0.06	0.002
<b>10</b>	O5	1.435±0.58	8.338±0.10	9.785±0.03	0.003
	NH6	0.925±0.16	25.491±0.27	12.765±0.03	<0.001
	O7	1.458±0.58	8.347±0.11	9.788±0.02	0.003
	OH2'	1.618±0.53	11.184±0.15	10.513±0.03	0.004
	OH3'	1.624±0.51	9.189±0.12	10.026±0.01	0.004
	OH4'	1.761±0.46	10.950±0.11	10.461±0.02	0.003
	O5'	0.042±0.01	0.476±0.02	2.688±0.12	0.023
	OH6'	1.788±0.57	15.795±0.67	11.369±0.05	0.005
<b>11</b>	O5	1.415±0.58	8.177±0.09	9.737±0.04	0.004
	NH6	0.925±0.16	23.555±0.41	12.068±0.03	<0.001
	O7	1.461±0.58	8.401±0.11	9.804±0.03	0.003
	OH2'	1.574±0.49	13.143±0.12	10.171±0.02	0.004
	OH3'	1.563±0.48	13.594±0.17	10.159±0.03	0.003
	O4'	0.204±0.42	1.119±0.02	4.808±0.03	0.014
	O5'	0.032±0.02	0.432±0.01	2.675±0.10	0.029
	OH6'	1.905±0.49	25.310±0.69	12.292±0.07	0.002
<b>12</b>	O5	1.418±0.58	9.245±0.11	10.041±0.02	0.004

	NH6	0.926±0.15	24.417±0.37	12.361±0.03	<0.001
	O7	1.456±0.58	9.216±0.11	10.033±0.03	0.003
	OH2'	1.663±0.52	12.858±0.14	10.859±0.02	0.004
	OH3'	1.725±0.49	13.052±0.19	10.896±0.03	0.003
	O4'	0.213±0.42	1.108±0.01	4.782±0.04	0.014
	O5'	0.041±0.02	0.473±0.02	2.673±0.10	0.024
	OH6'	1.630±0.57	16.568±0.65	10.896±0.06	0.003
<b>13</b>	O5	1.424±0.58	8.283±0.13	9.769±0.02	0.003
	NH6	0.924±0.16	21.586±0.28	13.201±0.02	<0.001
	O7	1.457±0.58	8.396±0.11	9.803±0.14	0.003
	OH2'	1.682±0.52	12.249±0.13	9.759±0.03	0.006
	OH3'	1.817±0.49	13.309±0.12	9.922±0.03	0.003
	O4'	0.210±0.42	1.113±0.01	4.793±0.02	0.014
	O5'	0.035±0.01	0.485±0.02	2.745±0.11	0.035
	OH6'	1.649±0.57	15.620±0.55	10.963±0.04	0.004
<b>14</b>	O5	1.419±0.58	8.285±0.14	9.769±0.02	0.003
	NH6	0.924±0.16	21.504±0.28	13.231±0.02	<0.001
	O7	1.459±0.58	8.452±0.10	9.819±0.14	0.003
	OH2'	1.776±0.52	12.292±0.13	9.772±0.02	0.006
	OH3'	1.739±0.51	12.113±0.13	9.616±0.02	0.005
	OH4'	1.767±0.46	11.568±0.11	9.853±0.03	0.003
	O5'	0.037±0.02	0.547±0.01	2.862±0.10	0.039
	OH6'	1.688±0.58	14.891±0.51	10.313±0.03	0.004
<b>15</b>	O5	1.415±0.58	8.425±0.11	9.811±0.02	0.003
	NH6	0.925±0.15	21.470±0.27	12.954±0.02	<0.001
	O7	1.442±0.58	8.472±0.10	9.825±0.15	0.003
	OH2'	1.782±0.52	12.326±0.12	9.812±0.01	0.006
	OH3'	1.739±0.52	12.127±0.11	9.021±0.02	0.006
	OH4'	1.768±0.45	11.584±0.11	9.858±0.02	0.002
	O5'	0.042±0.01	0.487±0.02	2.978±0.11	0.023
	OH6'	1.740±0.58	14.961±0.56	10.338±0.04	0.004
<b>16</b>	O5	1.639±0.56	11.011±0.12	10.475±0.03	0.003
	NH6	0.927±0.16	21.458±0.26	12.591±0.02	<0.001
	O7	1.679±0.56	11.003±0.13	10.473±0.03	0.003
	OH2'	1.783±0.51	12.651±0.10	9.877±0.02	0.006
	OH3'	1.733±0.51	12.098±0.12	9.910±0.02	0.005
	OH4'	1.861±0.46	11.509±0.11	9.836±0.03	0.003
	O5'	0.037±0.01	0.503±0.01	2.934±0.10	0.029
	OH6'	1.723±0.57	14.722±0.53	10.497±0.05	0.004
<b>17</b>	O5	1.411±0.58	8.384±0.10	9.799±0.02	0.003
	NH6	0.924±0.15	21.448±0.25	12.537±0.02	<0.001
	O7	1.444±0.58	8.431±0.09	9.813±0.02	0.003
	OH2'	1.670±0.52	12.697±0.03	10.828±0.16	0.004
	OH3'	1.760±0.51	10.759±0.02	10.417±0.15	0.003
	O5'	0.041±0.02	0.480±0.02	2.706±0.12	0.023
	OH6'	1.670±0.57	15.435±0.48	11.312±0.05	0.002
<b>18</b>	O5	1.419±0.58	8.484±0.12	9.828±0.02	0.003
	NH6	0.925±0.16	21.402±0.24	12.271±0.03	<0.001
	O7	1.450±0.58	8.570±0.11	9.853±0.03	0.003
	OH2'	1.693±0.51	13.175±0.11	10.920±0.04	0.003
	OH3'	1.623±0.51	9.373±0.09	10.076±0.02	0.004
	OH4'	1.601±0.51	10.355±0.10	10.323±0.03	0.003
	O5'	0.005±0.001	0.262±0.01	1.209±0.01	0.030
	NH <sub>3</sub> +6'	2.738±0.85	25.007±0.56	12.508±0.12	0.001

## 6 Discussão Geral

ICZs compõem uma família de compostos heterocíclicos glicosilados produzidos por diferentes microorganismos e uma série de invertebrados marinhos [1, 2, 3, 4]. Desde o primeiro isolado de staurosporina em 1977, a partir de actinobactérias coletadas no solo [1], a variabilidade de compostos dessa família de ligantes expandiu-se através do emprego de diversas abordagens biossintéticas, de síntese química ou engenharia genética em organismos produtores [7]. Através dos estudos que compuseram os capítulos I e II, esta tese buscou compreender a dinâmica conformacional e as características de interação com o solvente de compostos ICZs livres em solução aquosa.

### 6.1 Derivação dos Parâmetros Topológicos

A fiel reprodutibilidade de dados experimentais a partir de simulações de MD é dependente da parametrização do campo de força utilizado, assim como da derivação dos parâmetros topológicos apropriados para a descrição dos sistemas e suas particularidades. Nesse contexto, o processo de elaboração de topologias para descrever novas moléculas é uma etapa crucial para garantir simulações confiáveis e que forneçam corretas previsões do comportamento molecular em estudo. Um problema no desenvolvimento de campos de força é a grande variedade de compostos químicos, para os quais parâmetros adequados precisam ser derivados. Essa questão pode ser, até certo ponto, suavizada se os parâmetros puderem ser transferíveis entre átomos ou grupos de átomos em moléculas diferentes, mantendo-os simples, o que facilita inclusive a eficiência da simulação e a amostragem das variações dinâmicas [170]. Assim, os parâmetros são especificados dentro de fragmentos moleculares representativos dos grupos funcionais presentes. As topologias moleculares são construídas pela montagem de blocos, a partir do princípio da transferibilidade [176].

Para a geração de novos parâmetros para o estudo de pequenos ligantes, visando o desenvolvimento de fármacos, o ATB produz topologias compatíveis com o campo de força GROMOS de forma sistemática e automatizada. Ao utilizar essa ferramenta, os parâmetros para descrever cargas atômicas parciais, ligações, ângulos de ligação e diedros impróprios são derivados de cálculos quânticos, enquanto que os parâmetros para ângulos torcionais são atribuídos a partir de dados já existentes no banco de dados do campo de força GROMOS, basta que o usuário forneça a estrutura do ligante [135]. Em se tratando de moléculas flexíveis, entretanto, algumas limitações resultantes da abordagem do ATB merecem ser pontuadas, como a sensibilidade das cargas atômicas às mudanças conformacionais da molécula e da flexibilidade molecular à presença de água explícita no sistema em análise. Esses aspectos podem dificultar a derivação sistemática e até mesmo

o emprego da transferibilidade de um conjunto de cargas parciais para moléculas com alto grau de flexibilidade [175]. A omissão de estratégia de derivação de parâmetros para ângulos torcionais para novas estruturas moleculares representa perda na acurácia da reprodução do comportamento molecular do sistema, uma vez que as variáveis que definem os potenciais diedrais são sensíveis à configuração funcional e estrutural do fragmento, além dos efeitos aditivos da vizinhança química.

Assim, a derivação de parâmetros específicos para a descrição de ligantes ICZs no campo de força GROMOS, a partir de cálculos quânticos, e o ajuste empírico através da calibração a dados de espectroscopia de NMR constituem uma estratégia para refinar a reprodutibilidade do comportamento molecular desses ligantes em simulações de MD.

A derivação de cargas parciais para ICZs, utilizando como referência o vetor momento de dipolo obtido por cálculos quânticos, seguiu o protocolo desenvolvido por Polêto *et al.* que forneceu parâmetros topológicos robustos para sistemas aromáticos ao derivar novas cargas através de um algoritmo baseado em *least-square fit solution* em conjunto com módulos de cargas já parametrizados [175]. A estratégia vai ao encontro do princípio de transferibilidade, dinamizando com segurança a construção de topologias, uma vez que permite a extrapolação dos parâmetros dos grupos funcionais calibrados para novas moléculas. Além disso, uma melhor descrição das propriedades termodinâmicas de ICZs pode ser atingida através do uso de modelos bem calibrados para a descrição de propriedades físico-químicas em fase condensada. A isso, soma-se a vantagem do uso do campo de força GROMOS e sua filosofia de parametrização que faz uso de propriedades em fase condensada, como entalpia de vaporização e energia livre de solvatação, e a sensibilidade às cargas atômicas parciais e aos parâmetros de Lennard-Jones, das energias de interação, tanto intra quanto intermoleculares, e de seus componentes entálpicos [164, 172, 155].

A geração de novos parâmetros para ligações torcionáveis representa um impacto direto sobre a descrição conformacional de ICZs. A obtenção dos potenciais diedrais a partir de cálculos quânticos e o ajuste para sua reprodução na mecânica clássica, através de regressão linear de mínimos quadrados empregada pelo protocolo de Rusu *et al.*, permitiram a manutenção dos potenciais máximos e mínimos da rotação diedral e simplicidade dos parâmetros [199], o que favoreceu a eficiência da simulação e a ampla amostragem das variações conformacionais dos compostos.

O ajuste empírico dos potenciais diedrais para carboidratos permitiu uma melhor amostragem do conjunto conformacional de hexopiranoses e aproximou os dados teóricos de  $^3J_{H,H}$  aos sinais observados por espectroscopia de NMR. Esses resultados concordam com a previsão de Hansen e Hünenberger de que a parametrização de campos de força para descrição de hexopiranoses dependeria principalmente de observáveis experimentais espectroscópicos, termodinâmicos e cinéticos em solução para a adequada descrição da energia livre da conformação desses anéis, o que representa um importante desafio em



vista da relativa escassez de dados experimentais sobre carboidratos solvatados [209]. As discrepâncias pontuais na reprodução dos sinais de acoplamento  ${}^3J_{H,H}$  podem estar refletindo a falta de calibração da equação de Haasnoot-Altona para o cálculo de acoplamentos envolvendo grupos carregados, além da interação entre substituintes e das flexões irregulares de ângulos de ligação entre os átomos que compõem a via de acoplamento [210, 126, 132], eventos observados durante as simulações de diferentes compostos ICZs. Soma-se a isso, a conhecida incerteza de 1 a 2 Hz na parametrização da relação de Karplus para converter ângulos diedrais em constantes de acoplamento [211] e a ambiguidade intrínseca a valores de acoplamento em torno de 7 Hz, que podem ser o resultado de uma média conformacional, dificultando a distinção entre confômeros [136].

Dessa forma, a proposta desta tese fornece um conjunto de parâmetros calibrados para a simulação de ICZs livres em solução biológica e visa uma adequada reprodução da dinâmica conformacional e das energias de interação entre o ligante e o solvente aquoso, utilizando a filosofia da mecânica molecular clássica. Espera-se que uma adequada predição de energias livres de interação e dos estados conformacionais pré-complexação proporcionem uma melhor compreensão dos modos de ligação aos receptor-alvo de ICZs e auxiliem na expansão do potencial clínico desses ligantes.

## 6.2 Amostragem Conformacional

A flexibilidade molecular está diretamente associada ao número de ligações químicas torcionáveis entre os átomos de sua estrutura, impactando na variabilidade do conjunto conformacional que um determinado composto pode apresentar. A transição entre os confômeros que compõem essa população, no entanto, ocorre em escalas de tempo muito pequenas para que possam ser facilmente observadas experimentalmente. Por outro lado, o emprego de simulações computacionais permitem a observação de fenômenos estruturais em uma escala de tempo e espaço que dificilmente podem ser atingidas experimentalmente.

A população conformacional dos ligantes, ou seja, a quantidade de confômeros possível no meio biológico, impacta diretamente o custo entrópico para a formação do complexo receptor-ligante. Nesse sentido, quanto maior o teor de confômeros similares à conformação bioativa (complexada), menor será o custo entrópico para a complexação. Assim, o estudo da dinâmica conformacional de ICZs em meio biológico é de grande importância para melhorar a compreensão do processo de formação de complexos e torná-lo mais eficiente [212]. Para mapear a dinâmica conformacional desses ligantes, de forma a reproduzir ao máximo as condições biológicas no estado pré-complexação, utilizou-se água como solvente durante as simulações.

Apesar de serem constituídos por um sistema de anéis aromáticos, com conseqüente rigidez nessa região, os ICZs possuem uma região de alta flexibilidade, a porção sacarídica.

O importante papel desse domínio molecular no mecanismo bioativo desses compostos, como mimetização de ATP e estabilização do complexo receptor-ligante, já é bem descrito [74, 75, 94, 95, 91]. As simulações em meio biológico mostraram diferentes confôrmeros nessa região, múltiplas conformações de hexopiranoses e variações na distribuição do ângulo diedral da ligação glicosídica. Para aumentar a amostragem e preencher todo o espaço conformacional dos ICZs estudados, assim como identificar a energia livre associada a cada conformação e a barreira que separa um confôrmero de outro, cálculos de metadinâmica foram empregados. A amostragem ampliada pela adição de viés, sobre as coordenadas que definem a conformação dos anéis e o ângulo da ligação glicosídica, permitiu acesso ao conjunto de conformações que populam cada perfil, nos quais as conformações mais prevalentes tendem a ser aquelas com menor energia livre [205, 192].

Por outro lado, a amostragem conformacional avaliada é diretamente dependente da acurácia dos parâmetros topológicos utilizados, em especial dos potenciais torcionais e das cargas atômicas parciais. Assim, a acurácia da calibração das torções impacta diretamente na adequada descrição de suas populações conformacionais e, por consequência, na entropia do ligante livre em solução [172]. Além disso, outros fatores podem atuar na variação conformacional da porção sacarídica de ICZs, dentre eles, a configuração dos grupos exocíclicos nos carboidratos, o impacto das substituições no cromóforo sobre as conformações preferenciais, a presença de ligações de hidrogênio intramoleculares e as interações mediadas por moléculas do solvente aquoso. Esses fatores podem desempenhar um papel crucial na mudança da abundância relativa de certas conformações e constituem um desafio analítico devido à difícil previsibilidade e dependência da dinâmica do sistema.

A identificação dos possíveis confôrmeros e suas frequências relativas em solução permitem inferir sobre as suas respectivas disponibilidades para complexação com o receptor-alvo em meio biológico. A presença de múltiplos confôrmeros de hexopiranoses observada durante as simulações, com pelo menos dois confôrmeros bem definidos e relativa baixa energia livre para interconversão combina bem com mecanismos de seleção conformacional de ligantes, devido ao baixo custo entálpico associado. Estudos envolvendo simulações da dinâmica molecular da complexação receptor-ligante podem beneficiar-se dos parâmetros propostos nesta tese, uma vez que o mapeamento do espaço conformacional do ligante permite estudar os processos que envolvem a seleção conformacional do mesmo, um tópico ainda pouco abundante na literatura.

### 6.3 Características de Interação com o Solvente

O processo de formação do complexo receptor-ligante está intimamente relacionado com a dinâmica do solvente ao redor de ambos, receptor e ligante. As variações entálpicas que ocorrem durante esse processo resultam das dinâmicas de formação e rompimento

de um grande número de interações, dentre elas ligações de hidrogênio e interações de vdW entre receptor e ligante e de ambos com o solvente circundante. Da mesma forma, as variações entrópicas envolvem o rearranjo do solvente e as mudanças conformacionais do receptor e do ligante durante o processo de complexação [213].

Para inferir sobre a participação entálpica das ligações de hidrogênio que os grupos polares em ICZs estabelecem com o solvente aquoso, foram calculadas propriedades derivadas dessas interações, como o número médio de ligações de hidrogênio, seus tempo de residência e meia-vida, além da energia livre de rompimento da ligação. Variações nessas propriedades também puderam ser observadas analisando mudanças no padrão de substituição nas vizinhanças desses grupos, com resultados que variaram desde o aumento da afinidade com a água até a competição pela interação e consequente redistribuição dessa afinidade e reorganização do solvente na camada de solvatação.

Esses resultados permitem compreender melhor o papel da dinâmica de substituintes e do solvente na descrição das interações de ICZs com o solvente biológico, além de fornecer informações que podem ser utilizadas em estudos da termodinâmica do processo de complexação. Assim, os dados obtidos nessa análise podem ser empregados em estratégias que visam a otimização dos compostos ICZs em termos de especificidade de alvo, relacionada à disponibilidade para interações com diferentes resíduos em proteíno cinases, e biodisponibilidade, devido à solubilidade em água desses ligantes ser um aspecto alvo de aprimoramento.



## 7 Conclusões

Este trabalho se propôs a estudar as características de flexibilidade e de interação com o solvente aquoso de ligantes indolocarbazólicos livres em solução através de simulações de dinâmica molecular, como meio de prospecção das propriedades de interações ligante-receptor. Os resultados apresentados permitem traçar as seguintes conclusões:

- Capítulo I - Indolocarbazolas inibidoras de cinases: staurosporina e derivados.
  - A validação dos parâmetros desenvolvidos para o campo de força GROMOS, utilizando-se de dados de espectroscopia de NMR, indicou a adequação das topologias para o estudo da MD desses ligantes;
  - A ampla percepção do conjunto conformacional de staurosporina e seus derivados, em solução aquosa, representa um fator a ser explorado visando a seletividade e eficiência desses compostos;
  - As características de interação com a água, observadas em nível atômico, podem afetar a energia livre de complexação ao receptor alvo, devido a variações nos custos de dessolvatação associados.
  
- Capítulo II - Indolocarbazolas inibidoras de topoisomerasas: rebecamicina e derivados.
  - A geração de parâmetros específicos e baseados em dados quânticos para a descrição, no campo de força GROMOS, de aspectos estruturais envolvidos nos mecanismos de ação permitiu a descrição conformacional acurada das moléculas estudadas;
  - A análise da dinâmica conformacional, envolvida no mecanismo de ação dos ligantes, elucidou o efeito de substituições vizinhas na barreira rotacional de ligações e permitiu compreender melhor a dinâmica da conformação bioativa das moléculas estudadas;
  - A avaliação das variações estruturais e seus impactos sobre a rede de solvatação estão implicadas na farmacodinâmica e, conseqüentemente, em mudanças nas propriedades bioativas das moléculas.



## 8 Perspectivas

O desenvolvimento deste trabalho e as informações elucidadas abrem perspectivas, tanto para aprofundar o estudo de ligantes ICZs, quanto para expandir investigações da dinâmica de outras família de pequenos ligantes, no âmbito da química medicinal e do desenvolvimento de novos fármacos. Neste sentido, podem-se elencar os seguintes cenários:

- Implementação de MD de complexos ICZ-proteína, após a triagem de derivados contra potenciais alvos, através de atracamento molecular, ou modelos preditivos baseados em QSAR, para investigar variações estruturais implicadas na melhoria da afinidade e seletividade;
- Expandir a análise conformacional e de interação com o solvente para outras famílias de pequenos ligantes, através das abordagens de parametrização específica, dentro da filosofia GROMOS, aqui empregadas;
- O desenvolvimento de rotinas automatizadas, para a construção de topologias de pequenos ligantes e validação de parâmetros com base em dados de espectroscopia de NMR, são encorajados a partir dos resultados aqui expostos;
- O avanço de estudos comparativos, para avaliar a confiabilidade da descrição de sistemas biológicos teóricos, podem beneficiar-se do uso de metodologias computacionais que fazem uso de dados experimentais para a pesquisa de moléculas de interesse terapêutico.





## Referências

- 1 OMURA, S. et al. A new alkaloid AM-2282 OF Streptomyces origin. Taxonomy, fermentation, isolation and preliminary characterization. *The Journal of antibiotics*, v. 30, n. 4, p. 275–82, apr 1977. ISSN 0021-8820. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/863788>>. Citado 3 vezes nas páginas 23, 26 e 117.
- 2 STEGLICH, W. et al. Indole Pigments from the Fruiting Bodies of the Slime Mold *Arcyria denudata*. *Angewandte Chemie International Edition in English*, v. 19, n. 6, p. 459–460, jun 1980. ISSN 0570-0833. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/anie.198004591>>. Citado 2 vezes nas páginas 23 e 117.
- 3 KNÜBEL, G. et al. Cytotoxic, antiviral indolocarbazoles from a blue-green alga belonging to the Nostocaceae. *The Journal of antibiotics*, v. 43, n. 10, p. 1236–9, oct 1990. ISSN 0021-8820. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/2175302>>. Citado 2 vezes nas páginas 23 e 117.
- 4 KINNEL, R. B.; SCHEUER, P. J. 11-Hydroxystaurosporine: a highly cytotoxic, powerful protein kinase C inhibitor from a tunicate. *The Journal of Organic Chemistry*, v. 57, n. 23, p. 6327–6329, nov 1992. ISSN 0022-3263. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/jo00049a049>>. Citado 2 vezes nas páginas 23 e 117.
- 5 JANOSIK, T. et al. Chemistry and Properties of Indolocarbazoles. *Chemical Reviews*, v. 118, n. 18, p. 9058–9128, sep 2018. ISSN 0009-2665. Disponível em: <<http://pubs.acs.org/doi/10.1021/acs.chemrev.8b00186>>. Citado na página 23.
- 6 SÁNCHEZ, C.; MÉNDEZ, C.; SALAS, J. A. Indolocarbazole natural products: occurrence, biosynthesis, and biological activity. *Natural Product Reports*, v. 23, n. 6, p. 1007–1045, 2006. Disponível em: <<http://pubs-rsc-org.ez45.periodicos.capes.gov.br/en/content/articlepdf/2006/np/b601930g>>. Citado 5 vezes nas páginas 23, 29, 33, 35 e 37.
- 7 ZENKOV, R. G. et al. Indolo[2,3-a]carbazoles: diversity, biological properties, application in antitumor therapy. *Chemistry of Heterocyclic Compounds*, v. 56, n. 6, p. 644–658, jun 2020. ISSN 0009-3122. Disponível em: <<https://link.springer.com/10.1007/s10593-020-02714-4>>. Citado 4 vezes nas páginas 23, 26, 29 e 117.
- 8 NETTLETON, D. et al. *Isolation and Structure Rebecamycin - a New Antitumor Antibiotic from Nocardia Aerocoligenes*. [S.l.], 1985. v. 26, n. 34, 4011–4014 p. Disponível em: <[https://ac.els-cdn.com/S0040403900892801/1-s2.0-S0040403900892801-main.pdf?{\\\_}tid=6776a5a0-f423-4379-85c8-dd0b2c6ac4c8{\&}acdnat=1539975381{\\\_}303a91af7acd5bb4a32ac2a](https://ac.els-cdn.com/S0040403900892801/1-s2.0-S0040403900892801-main.pdf?{\_}tid=6776a5a0-f423-4379-85c8-dd0b2c6ac4c8{\&}acdnat=1539975381{\_}303a91af7acd5bb4a32ac2a)>. Citado na página 24.
- 9 BUSH, J. A. et al. Production and biological activity of rebecamycin, a novel antitumor agent. *The Journal of antibiotics*, v. 40, n. 5, p. 668–78, may 1987. ISSN 0021-8820. Disponível em: <[http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed{\&}Cmd=Retrieve{\&}list{\\\_}uids=3112080{\&}dopt=abstractplushttp://www.ncbi.nlm.nih.gov/put](http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed{\&}Cmd=Retrieve{\&}list{\_}uids=3112080{\&}dopt=abstractplushttp://www.ncbi.nlm.nih.gov/put)>. Citado 2 vezes nas páginas 24 e 33.

- 10 LONG, B. et al. Discovery of Antitumor Indolocarbazoles: Rebeccamycin, NSC 655649, and Fluoroindolocarbazoles. *Current Medicinal Chemistry-Anti-Cancer Agents*, v. 2, n. 2, p. 255–266, mar 2002. ISSN 15680118. Disponível em: <<http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1568-0118&volume=2&issue=2&spage=255>>. Citado 4 vezes nas páginas 24, 25, 26 e 35.
- 11 KANEKO, T.; WONG, H. Two Synthetic Approaches to Rebeccamycin. *Tetrahedron Letters*, v. 26, n. 34, p. 4015–4018, 1985. Citado 2 vezes nas páginas 24 e 26.
- 12 MATSON, J. A. et al. AT2433-A1, AT2433-A2, AT2433-B1, and AT2433-B2 novel antitumor antibiotic compounds produced by actinomadura melliura. Taxonomy, fermentation, isolation and biological properties. *The Journal of Antibiotics*, v. 42, n. 11, p. 1547–1555, 1989. ISSN 18811469. Disponível em: <<http://joi.jlc.jst.go.jp/JST.Journalarchive/antibiotics1968/42.1547?from=CrossRef>>. Citado na página 24.
- 13 LAM, K. S. et al. Isolation of a bromo analog of rebeccamycin from *Saccharothrix aerocolonigenes*. *The Journal of Antibiotics*, v. 44, n. 9, p. 934–939, 1991. ISSN 0021-8820. Disponível em: <<http://joi.jlc.jst.go.jp/JST.Journalarchive/antibiotics1968/44.934?from=CrossRef>>. Citado na página 25.
- 14 LAM, K. S. et al. *Production, Isolation and Structure Determination of Novel Fluoroindolocarbazoles from Saccharothrix aerocolonigenes ATCC39243*. [S.l.], 2001. v. 54, n. 1, 1–9 p. Disponível em: <[https://www.jstage.jst.go.jp/article/antibiotics1968/54/1/54{\\\_}1{\\\_}1/{\\\_}pdf](https://www.jstage.jst.go.jp/article/antibiotics1968/54/1/54{\_}1{\_}1/{\_}pdf)>. Citado na página 25.
- 15 BALASUBRAMANIAN, B. N. et al. Design and Synthesis of a Fluoroindolocarbazole Series as Selective Topoisomerase I Active Agents. Discovery of Water-Soluble 3,9-Difluoro-12,13-dihydro-13-[6-amino-β- <math>\text{D}</math>-glucopyranosyl]-5 <math>\text{H}</math>,13 <math>\text{H}</math>-benzo[ <math>\text{b}</math>- thienyl[2,3-. *Journal of Medicinal Chemistry*, v. 47, n. 7, p. 1609–1612, mar 2004. ISSN 0022-2623. Disponível em: <<https://pubs.acs.org/doi/10.1021/jm034197s>>. Citado na página 25.
- 16 SAULNIER, M. G. et al. Discovery of a Fluoroindolo[2,3- a ]carbazole Clinical Candidate with Broad Spectrum Antitumor Activity in Preclinical Tumor Models Superior to the Marketed Oncology Drug, CPT-11. *Journal of Medicinal Chemistry*, v. 48, n. 7, p. 2258–2261, apr 2005. ISSN 0022-2623. Disponível em: <<https://pubs.acs.org/doi/10.1021/jm049090z>>. Citado 2 vezes nas páginas 25 e 33.
- 17 SANCHEZ, C. et al. From The Cover: Combinatorial biosynthesis of antitumor indolocarbazole compounds. *Proceedings of the National Academy of Sciences*, v. 102, n. 2, p. 461–466, jan 2005. ISSN 0027-8424. Disponível em: <<http://www.pnas.org/content/pnas/102/2/461.full.pdfhttp://www.pnas.org/cgi/doi/10.1073/pnas.0407809102>>. Citado na página 25.
- 18 RODRÍGUEZ, L. et al. Engineering Deoxysugar Biosynthetic Pathways from Antibiotic-Producing Microorganisms. *Chemistry Biology*, v. 9, n. 6, p. 721–729, jun 2002. ISSN 10745521. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1074552102001540>>. Citado na página 25.
- 19 SALAS, A. P. et al. Deciphering the late steps in the biosynthesis of the anti-tumour indolocarbazole staurosporine: sugar donor substrate flexibility of the

- StaG glycosyltransferase. *Molecular Microbiology*, v. 58, n. 1, p. 17–27, jul 2005. ISSN 0950382X. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881644/pdf/nihms205008.pdf>><<http://doi.wiley.com/10.1111/j.1365-2958.2005.04777.x>>. Citado na página 25.
- 20 TANAKA, S. et al. A new indolopyrrolocarbazole antitumor substance, ED-110, a derivative of BE-13793C. *The Journal of Antibiotics*, v. 45, n. 11, p. 1797–1798, 1992. ISSN 0021-8820. Disponível em: <<http://joi.jlc.jst.go.jp/JST.Journalarchive/antibiotics1968/45.1797?from=CrossRef>>. Citado na página 26.
- 21 YOSHINARI, T. et al. Novel antitumor indolocarbazole compound 6-N-formylamino-12,13-dihydro-1,11-dihydroxy-13-(beta-D-glucopyranosyl)-5H-indolo[2,3-a]pyrrolo[3,4-c]carbazole-5,7(6H)-dione (NB-506): induction of topoisomerase I-mediated DNA cleavage and mechanisms of cell li. *Cancer Res*, v. 55, n. 6, p. 1310–5, 1995. Citado na página 26.
- 22 KANZAWA, F. et al. Antitumor Activities of a New Indolocarbazole Substance, NB-506, and Establishment of NB-506-resistant Cell Lines, SBC-3/NB. *CANCER RESEARCH*, v. 55, p. 2806–2813, 1995. Citado na página 26.
- 23 OHKUBO, M. et al. Synthesis and biological activities of NB-506 analogues: Effects of the positions of two hydroxyl groups at the indole rings. *Bioorganic Medicinal Chemistry Letters*, v. 9, n. 23, p. 3307–3312, dec 1999. ISSN 0960894X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0960894X99005958>>. Citado na página 26.
- 24 OHKUBO, M. et al. Synthesis and biological activities of NB-506 analogues modified at the glucose group. *Bioorganic and Medicinal Chemistry Letters*, v. 10, n. 5, p. 419–422, 2000. ISSN 0960894X. Citado na página 26.
- 25 SAIF, M. W.; DIASIO, R. B. Edotecarin: A Novel Topoisomerase I Inhibitor. *Clinical Colorectal Cancer*, v. 5, n. 1, p. 27–36, may 2005. ISSN 15330028. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1533002811701633>>. Citado na página 26.
- 26 SAIF, M. W. et al. A phase I dose-escalation study of edotecarin (J-107088) combined with infusional 5-fluorouracil and leucovorin in patients with advanced/metastatic solid tumors. *Anti-Cancer Drugs*, v. 21, n. 7, p. 716–723, aug 2010. ISSN 0959-4973. Disponível em: <<https://journals.lww.com/00001813-201008000-00008>>. Citado na página 26.
- 27 KANEKO, T. et al. Water soluble derivatives of rebeccamycin. *The Journal of Antibiotics*, v. 43, n. 1, p. 125–127, 1990. ISSN 0021-8820. Disponível em: <<http://joi.jlc.jst.go.jp/JST.Journalarchive/antibiotics1968/43.125?from=CrossRef>>. Citado 2 vezes nas páginas 26 e 35.
- 28 Weitman, Steven M.D., Ph.D; Moore, R. B. et al. In Vitro Antitumor Activity of Rebeccamycin Analog (NSC 655649) Against Pediatric Solid Tumors. *Journal of Pediatric Hematology/Oncology*, v. 20, n. 2, p. 136–139, 1998. Citado na página 26.
- 29 DOWLATI, A. et al. Phase II and pharmacokinetic trial of rebeccamycin analog in advanced biliary cancers. *Cancer Chemotherapy and Pharmacology*, v. 65, n. 1, p. 73–78, dec 2009. ISSN 0344-5704. Disponível em: <<http://link.springer.com/10.1007/s00280-009-1005-x>>. Citado na página 26.

- 30 BURSTEIN, H. J. et al. Rebeccamycin analog for refractory breast cancer: A randomized phase II trial of dosing schedules. *Investigational New Drugs*, v. 25, n. 2, p. 161–164, dec 2006. ISSN 0167-6997. Disponível em: <<http://link.springer.com/10.1007/s10637-006-9007-6>>. Citado na página 26.
- 31 LANGEVIN, A. M. et al. A phase II trial of rebeccamycin analogue (NSC 655649) in children with solid tumors: A Children's Oncology Group study. *Pediatric Blood and Cancer*, v. 50, n. 3, p. 577–580, mar 2008. ISSN 15455009. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/pbc.21274>>. Citado na página 26.
- 32 SCHWANDT, A. et al. Phase-II Trial of Rebeccamycin Analog, a Dual Topoisomerase-I and -II Inhibitor, in Relapsed "Sensitive" Small Cell Lung Cancer. *Journal of Thoracic Oncology*, v. 7, n. 4, p. 751–754, apr 2012. ISSN 15560864. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1556086415329191>>. Citado na página 26.
- 33 Van Arnem, E. B. et al. A Rebeccamycin Analog Provides Plasmid-Encoded Niche Defense. *Journal of the American Chemical Society*, v. 137, n. 45, p. 14272–14274, nov 2015. ISSN 0002-7863. Disponível em: <<http://pubs.acs.org/http://pubs.acs.org/doi/10.1021/jacs.5b09794>>. Citado na página 26.
- 34 FURUSAKI, A. et al. X-Ray crystal structure of staurosporine: a new alkaloid from a *Streptomyces* strain. *J. Chem. Soc., Chem. Commun.*, n. 18, p. 800–801, 1978. ISSN 0022-4936. Disponível em: <<http://xlink.rsc.org/?DOI=C39780000800>>. Citado na página 26.
- 35 FUNATO, N. et al. Absolute Configuration of Staurosporine By X-Ray Analysis. *Tetrahedron Letters*, v. 35, n. 8, p. 1251–1254, 1994. Citado na página 26.
- 36 LINK, J. T. et al. Staurosporine and ent -Staurosporine: The First Total Syntheses, Prospects for a Regioselective Approach, and Activity Profiles 1. *Journal of the American Chemical Society*, v. 118, n. 12, p. 2825–2842, jan 1996. ISSN 0002-7863. Disponível em: <<https://pubs.acs.org/doi/10.1021/ja952907g>>. Citado na página 26.
- 37 TAMAOKI, T. et al. Staurosporine, a potent inhibitor of phospholipidCa<sup>++</sup>dependent protein kinase. *Biochemical and Biophysical Research Communications*, v. 135, n. 2, p. 397–402, mar 1986. ISSN 0006291X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/0006291X86900082>>. Citado 2 vezes nas páginas 27 e 29.
- 38 OMURA, S. et al. Staurosporine, a Potentially Important Gift from a Microorganism. *The Journal of Antibiotics*, v. 48, n. 7, p. 535–548, 1995. ISSN 0021-8820. Disponível em: <<http://joi.jlc.jst.go.jp/JST.Journalarchive/antibiotics1968/48.535?from=CrossRef>>. Citado 2 vezes nas páginas 27 e 29.
- 39 RÜEGG, U. T.; GILLIAN, B. Staurosporine, K-252 and UCN-01: potent but nonspecific inhibitors of protein kinases. *Trends in Pharmacological Sciences*, v. 10, n. 6, p. 218–220, jun 1989. ISSN 01656147. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/0165614789902630>>. Citado 2 vezes nas páginas 27 e 29.
- 40 BARTH, T. et al. Staurosporine-Induced Cell Death in *Trypanosoma brucei*; and the Role of Endonuclease G during Apoptosis. *Open Journal*

- of *Apoptosis*, v. 03, n. 02, p. 16–31, 2014. ISSN 2168-3832. Disponível em: <<http://www.scirp.org/journal/doi.aspx?DOI=10.4236/ojapo.2014.32003>>. Citado na página 27.
- 41 CAI, Y. et al. Further Minor Metabolites of Staurosporine Produced by a *Streptomyces longisporoflavus* Strain. *The Journal of Antibiotics*, v. 49, n. 6, p. 519–526, 1996. ISSN 0021-8820. Disponível em: <<http://joi.jlc.jst.go.jp/JST.Journalarchive/antibiotics1968/49.519?from=CrossRef>>. Citado na página 27.
- 42 TAKAHASHI, I. et al. UCN-01 and UCN-02, new selective inhibitors of protein kinase C. I. Screening, producing organism and fermentation. *The Journal of Antibiotics*, v. 42, n. 4, p. 564–570, 1989. ISSN 0021-8820. Disponível em: <<http://joi.jlc.jst.go.jp/JST.Journalarchive/antibiotics1968/42.564?from=CrossRef>>. Citado na página 27.
- 43 TAKAHASHI, I. et al. UCN-01 and UCN-02, new selective inhibitors of protein kinase C. II. Purification, physico-chemical properties, structural determination and biological activities. *The Journal of Antibiotics*, v. 42, n. 4, p. 571–576, 1989. ISSN 0021-8820. Disponível em: <<http://joi.jlc.jst.go.jp/JST.Journalarchive/antibiotics1968/42.571?from=CrossRef>>. Citado na página 27.
- 44 SIGNORE, M. et al. UCN-01 enhances cytotoxicity of irinotecan in colorectal cancer stem-like cells by impairing DNA damage response. *Oncotarget*, v. 7, n. 28, p. 44113–44128, jul 2016. ISSN 1949-2553. Disponível em: <<https://www.oncotarget.com/lookup/doi/10.18632/oncotarget.9859>>. Citado na página 27.
- 45 MULL, B. B. et al. Specific, reversible G1 arrest by UCN-01 in vivo provides cytostatic protection of normal cells against cytotoxic chemotherapy in breast cancer. *British Journal of Cancer*, v. 122, n. 6, p. 812–822, mar 2020. ISSN 0007-0920. Disponível em: <<http://www.nature.com/articles/s41416-019-0707-z>>. Citado 2 vezes nas páginas 27 e 31.
- 46 EDELMAN, M. J. et al. Phase I and Pharmacokinetic Study of 7-Hydroxystaurosporine and Carboplatin in Advanced Solid Tumors. *Clinical Cancer Research*, v. 13, n. 9, p. 2667–2674, may 2007. ISSN 1078-0432. Disponível em: <<http://clincancerres.aacrjournals.org/lookup/doi/10.1158/1078-0432.CCR-06-1832>>. Citado na página 27.
- 47 HOTTE, S. et al. Phase I trial of UCN-01 in combination with topotecan in patients with advanced solid cancers: a Princess Margaret Hospital Phase II Consortium study. *Annals of Oncology*, v. 17, n. 2, p. 334–340, feb 2006. ISSN 09237534. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0923753419403876>>. Citado 2 vezes nas páginas 27 e 31.
- 48 WELCH, S. et al. UCN-01 in combination with topotecan in patients with advanced recurrent ovarian cancer: A study of the Princess Margaret Hospital Phase II consortium. *Gynecologic Oncology*, v. 106, n. 2, p. 305–310, aug 2007. ISSN 00908258. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0090825807001461>>. Citado na página 27.
- 49 LARA, P. N. et al. The Cyclin-Dependent Kinase Inhibitor UCN-01 Plus Cisplatin in Advanced Solid Tumors: A California Cancer Consortium Phase I Pharmacokinetic and Molecular Correlative Trial. *Clinical Cancer Research*,

- v. 11, n. 12, p. 4444–4450, jun 2005. ISSN 1078-0432. Disponível em: <<http://clincancerres.aacrjournals.org/lookup/doi/10.1158/1078-0432.CCR-04-2602>>. Citado 2 vezes nas páginas 27 e 31.
- 50 FRACASSO, P. M. et al. A Phase 1 study of UCN-01 in combination with irinotecan in patients with resistant solid tumor malignancies. *Cancer Chemotherapy and Pharmacology*, v. 67, n. 6, p. 1225–1237, jun 2011. ISSN 0344-5704. Disponível em: <<http://link.springer.com/10.1007/s00280-010-1410-1>>. Citado 3 vezes nas páginas 27, 29 e 31.
- 51 MARTI, G. E. et al. Phase I trial of 7-hydroxystaurosporine and fludarabine phosphate: in vivo evidence of 7-hydroxystaurosporine induced apoptosis in chronic lymphocytic leukemia. *Leukemia Lymphoma*, v. 52, n. 12, p. 2284–2292, dec 2011. ISSN 1042-8194. Disponível em: <<http://www.tandfonline.com/doi/full/10.3109/10428194.2011.589547>>. Citado 2 vezes nas páginas 27 e 31.
- 52 TANIDA, S. et al. TAN-999 and TAN-1030A, new indolocarbazole alkaloids with macrophage-activating properties. *The Journal of Antibiotics*, v. 42, n. 11, p. 1619–1630, 1989. ISSN 0021-8820. Disponível em: <<http://joi.jlc.jst.go.jp/JST.Journalarchive/antibiotics1968/42.1619?from=CrossRef>>. Citado na página 27.
- 53 QIN, L.-L. et al. Bioactive metabolites from marine-derived *Streptomyces* sp. A68 and its Rifampicin resistant mutant strain R-M1. *Phytochemistry Letters*, v. 23, p. 46–51, feb 2018. ISSN 18743900. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1874390017305888>>. Citado na página 27.
- 54 HAN, X.-X. et al. ZHD-0501, a novel naturally occurring staurosporine analog from *Actinomadura* sp. 007. *Tetrahedron Letters*, v. 46, n. 36, p. 6137–6140, sep 2005. ISSN 00404039. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0040403905014528>>. Citado na página 28.
- 55 REYES, F. et al. Cytotoxic Staurosporines from the Marine Ascidian *Cystodytes solitus*. *Journal of Natural Products*, v. 71, n. 6, p. 1046–1048, jun 2008. ISSN 0163-3864. Disponível em: <<https://pubs.acs.org/doi/10.1021/np700748h>>. Citado na página 28.
- 56 CARAVATTI, G. et al. Inhibitory activity and selectivity of staurosporine derivatives towards protein kinase C. *Bioorganic Medicinal Chemistry Letters*, v. 4, n. 3, p. 399–404, feb 1994. ISSN 0960894X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/0960894X94800049>>. Citado 2 vezes nas páginas 28 e 32.
- 57 KIM, E. S. Midostaurin: First Global Approval. *Drugs*, v. 77, n. 11, p. 1251–1259, jul 2017. ISSN 0012-6667. Disponível em: <<http://link.springer.com/10.1007/s40265-017-0779-0>>. Citado na página 28.
- 58 STONE, R. M. et al. Midostaurin: its odyssey from discovery to approval for treating acute myeloid leukemia and advanced systemic mastocytosis. *Blood Advances*, v. 2, n. 4, p. 444–453, feb 2018. ISSN 2473-9529. Disponível em: <<https://ashpublications.org/bloodadvances/article/2/4/444/15859/Midostaurin-its-odyssey-from-discovery-to-approval>>. Citado 2 vezes nas páginas 28 e 32.

- 59 WEISBERG, E. et al. Inhibition of mutant FLT3 receptors in leukemia cells by the small molecule tyrosine kinase inhibitor PKC412. *Cancer Cell*, v. 1, n. 5, p. 433–443, jun 2002. ISSN 15356108. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1535610802000697>>. Citado 2 vezes nas páginas 28 e 32.
- 60 KAWASE, T. et al. Effect of Fms-like tyrosine kinase 3 (FLT3) ligand (FL) on antitumor activity of gilteritinib, a FLT3 inhibitor, in mice xenografted with FL-overexpressing cells. *Oncotarget*, v. 10, n. 58, p. 6111–6123, oct 2019. ISSN 1949-2553. Disponível em: <<https://www.oncotarget.com/lookup/doi/10.18632/oncotarget.27222>>. Citado 2 vezes nas páginas 28 e 32.
- 61 GROWNEY, J. D. et al. Activation mutations of human c-KIT resistant to imatinib mesylate are sensitive to the tyrosine kinase inhibitor PKC412. *Blood*, v. 106, n. 2, p. 721–724, jul 2005. ISSN 0006-4971. Disponível em: <<https://ashpublications.org/blood/article/106/2/721/21319/Activation-mutations-of-human-cKIT-resistant-to>>. Citado na página 28.
- 62 SOTLAR, K. et al. Variable presence of KIT D816V in clonal haematological non-mast cell lineage diseases associated with systemic mastocytosis (SM-AHNMD). *The Journal of Pathology*, v. 220, n. 5, p. 586–595, apr 2010. ISSN 00223417. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/path.2677>>. Citado na página 28.
- 63 STONE, R. M. et al. Patients with acute myeloid leukemia and an activating mutation in FLT3 respond to a small-molecule FLT3 tyrosine kinase inhibitor, PKC412. *Blood*, v. 105, n. 1, p. 54–60, jan 2005. ISSN 0006-4971. Disponível em: <<https://ashpublications.org/blood/article/105/1/54/19995/Patients-with-acute-myeloid-leukemia-and-an>>. Citado na página 28.
- 64 STONE, R. M. et al. Phase IB study of the FLT3 kinase inhibitor midostaurin with chemotherapy in younger newly diagnosed adult patients with acute myeloid leukemia. *Leukemia*, v. 26, n. 9, p. 2061–2068, sep 2012. ISSN 0887-6924. Disponível em: <<https://www.nature.com/articles/leu2012115>>. Citado na página 28.
- 65 STONE, R. M. et al. Midostaurin plus Chemotherapy for Acute Myeloid Leukemia with a FLT3 Mutation. *New England Journal of Medicine*, v. 377, n. 5, p. 454–464, aug 2017. ISSN 0028-4793. Disponível em: <<http://www.nejm.org/doi/10.1056/NEJMoa1614359>>. Citado na página 28.
- 66 GOTLIB, J. Activity of the tyrosine kinase inhibitor PKC412 in a patient with mast cell leukemia with the D816V KIT mutation. *Blood*, v. 106, n. 8, p. 2865–2870, oct 2005. ISSN 0006-4971. Disponível em: <<http://www.bloodjournal.org/cgi/doi/10.1182/blood-2005-04-1568>>. Citado na página 28.
- 67 GOTLIB, J. et al. Efficacy and Safety of Midostaurin in Advanced Systemic Mastocytosis. *New England Journal of Medicine*, v. 374, n. 26, p. 2530–2541, jun 2016. ISSN 0028-4793. Disponível em: <<http://www.nejm.org/doi/10.1056/NEJMoa1513098>>. Citado na página 28.
- 68 DEANGELO, D. J. et al. Efficacy and safety of midostaurin in patients with advanced systemic mastocytosis: 10-year median follow-up of a phase II trial. *Leukemia*, v. 32, n. 2, p. 470–478, feb 2018. ISSN 0887-6924. Disponível em: <<http://www.nature.com/articles/leu2017234>>. Citado na página 28.

- 69 FU, P. et al. Streptocarbazoles A and B, Two Novel Indolocarbazoles from the Marine-Derived Actinomycete Strain *Streptomyces* sp. FMA. *Organic Letters*, v. 14, n. 9, p. 2422–2425, may 2012. ISSN 1523-7060. Disponível em: <<https://pubs.acs.org/sharingguidelines><http://pubs.acs.org/doi/10.1021/ol3008638><https://pubs-acs-org.ez45.periodicos.capes.gov.br/doi/pdf/10.1021/ol3008638>>. Citado na página 28.
- 70 CARTUCHE, L. et al. Antiamoebic Activities of Indolocarbazole Metabolites Isolated from *Streptomyces sanyensis* Cultures. *Marine Drugs*, v. 17, n. 10, p. 588, oct 2019. ISSN 1660-3397. Disponível em: <<https://www.mdpi.com/1660-3397/17/10/588>>. Citado na página 28.
- 71 ZHOU, B. et al. Cytotoxic indolocarbazoles alkaloids from the streptomyces sp. A65. *Tetrahedron*, Elsevier Ltd, v. 74, n. 7, p. 726–730, 2017. ISSN 14645416. Disponível em: <<https://doi.org/10.1016/j.tet.2017.12.048>>. Citado na página 28.
- 72 PAWSON, T.; SCOTT, J. D. Protein phosphorylation in signaling – 50 years and counting. *Trends in Biochemical Sciences*, v. 30, n. 6, p. 286–290, jun 2005. ISSN 09680004. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0968000405001210>>. Citado na página 29.
- 73 KOLPAKSIDI, I. P. et al. ANTITUMOR DRUGS BASED ON INDOLOCARBAZOL DERIVATIVES. *Pharmacy Pharmacology*, v. 9, n. 4, p. 252–265, sep 2021. ISSN 2413-2241. Disponível em: <<https://www.pharmpharm.ru/jour/article/view/881>>. Citado na página 29.
- 74 LAWRIE, A. M. et al. Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2. *Nature Structural Biology*, v. 4, n. 10, p. 796–801, oct 1997. ISSN 1072-8368. Disponível em: <<http://www.nature.com/doi/10.1038/nsb1097-796>>. Citado 3 vezes nas páginas 29, 37 e 120.
- 75 PRADE, L. et al. Staurosporine-induced conformational changes of cAMP-dependent protein kinase catalytic subunit explain inhibitory potential. *Structure*, v. 5, n. 12, p. 1627–1637, dec 1997. ISSN 09692126. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0969212697003109>>. Citado 3 vezes nas páginas 29, 37 e 120.
- 76 KOMANDER, D. et al. *Structural basis for UCN-01 (7-hydroxystaurosporine) specificity and PDK1 (3-phosphoinositide-dependent protein kinase-1) inhibition*. [S.l.], 2003. v. 375, 255–262 p. Citado 2 vezes nas páginas 29 e 30.
- 77 ZHAO, B. et al. Structural Basis for Chk1 Inhibition by UCN-01. *Journal of Biological Chemistry*, v. 277, n. 48, p. 46609–46615, nov 2002. ISSN 00219258. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0021925819332764>>. Citado na página 31.
- 78 GAMBACORTI-PASSERINI, C. et al. Alpha1 acid glycoprotein binds to imatinib (STI571) and substantially alters its pharmacokinetics in chronic myeloid leukemia patients. *Clinical cancer research : an official journal of the American Association for Cancer Research*, v. 9, n. 2, p. 625–32, feb 2003. ISSN 1078-0432. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/12576428>>. Citado na página 31.



- 79 LANDIN, E. J. et al. The structural basis for high affinity binding of  $\alpha$ 1-acid glycoprotein to the potent antitumor compound UCN-01. *Journal of Biological Chemistry*, v. 297, n. 6, p. 101392, dec 2021. ISSN 00219258. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0021925821011984>>. Citado na página 31.
- 80 FABBRO, D. et al. PKC412—a protein kinase inhibitor with a broad therapeutic potential. *Anti-cancer drug design*, v. 15, n. 1, p. 17–28, feb 2000. ISSN 0266-9536. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/10888033>>. Citado na página 32.
- 81 ALEXEEVA, M. et al. The structure of a dual-specificity tyrosine phosphorylation-regulated kinase 1A–PKC412 complex reveals disulfide-bridge formation with the anomalous catalytic loop HRD(HCD) cysteine. *Acta Crystallographica Section D Biological Crystallography*, v. 71, n. 5, p. 1207–1215, may 2015. ISSN 1399-0047. Disponível em: <<http://scripts.iucr.org/cgi-bin/paper?S1399004715005106>>. Citado na página 32.
- 82 GANI, O. A. B. S. M.; ENGH, R. A. Protein kinase inhibition of clinically important staurosporine analogues. *Natural Product Reports*, v. 27, n. 4, p. 489, 2010. ISSN 0265-0568. Disponível em: <<http://xlink.rsc.org/?DOI=b923848b>>. Citado na página 32.
- 83 CHAMPOUX, J. J. DNA Topoisomerases: Structure, Function, and Mechanism. *Annual Review of Biochemistry*, v. 70, n. 1, p. 369–413, jun 2001. ISSN 0066-4154. Disponível em: <<http://www.annualreviews.org/doi/10.1146/annurev.biochem.70.1.369>>. Citado 2 vezes nas páginas 32 e 33.
- 84 LEPPARD, J. B.; CHAMPOUX, J. J. Human DNA topoisomerase I: relaxation, roles, and damage control. *Chromosoma*, v. 114, n. 2, p. 75–85, jul 2005. ISSN 0009-5915. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15830206>>. Citado na página 32.
- 85 ROSSI, F. et al. Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. *Nature*, v. 381, n. 6577, p. 80–2, may 1996. ISSN 0028-0836. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/8609994>>. Citado na página 32.
- 86 LABOURIER, E. et al. Interaction between the N-terminal domain of human DNA topoisomerase I and the arginine-serine domain of its substrate determines phosphorylation of SF2/ASF splicing factor. *Nucleic acids research*, v. 26, n. 12, p. 2955–62, jun 1998. ISSN 0305-1048. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/9611241http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC147637>>. Citado 2 vezes nas páginas 32 e 35.
- 87 ALMOUZNI, G.; MÉCHALI, M. Assembly of spaced chromatin involvement of ATP and DNA topoisomerase activity. *The EMBO journal*, v. 7, n. 13, p. 4355–65, dec 1988. ISSN 0261-4189. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/2854062http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC455159>>. Citado na página 33.
- 88 STRICK, R. et al. Cation-chromatin binding as shown by ion microscopy is essential for the structural integrity of chromosomes. *The Journal of cell biology*, v. 155, n. 6, p. 899–910, dec 2001. ISSN 0021-9525. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/>

11739403<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2150894>>. Citado na página 33.

89 MITSCHER, L. A. Bacterial Topoisomerase Inhibitors: Quinolone and Pyridone Antibacterial Agents. *Chemical Reviews*, v. 105, n. 2, p. 559–592, feb 2005. ISSN 0009-2665. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/cr030101q>>. Citado na página 33.

90 YAMASHITA, Y. et al. Induction of mammalian DNA topoisomerase I mediated DNA cleavage by antitumor indolocarbazole derivatives. *Biochemistry*, v. 31, n. 48, p. 12069–12075, dec 1992. ISSN 0006-2960. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/bi00163a015>>. Citado 2 vezes nas páginas 33 e 34.

91 PRUDHOMME, M. Rebeccamycin analogues as anti-cancer agents. *European Journal of Medicinal Chemistry*, v. 38, n. 2, p. 123–140, feb 2003. ISSN 02235234. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0223523403000114>>. Citado 2 vezes nas páginas 33 e 120.

92 LING-HUA, M.; ZHI-YONG, L.; POMMIER, Y. Non-Camptothecin DNA Topoisomerase I Inhibitors in Cancer Therapy. *Current Topics in Medicinal Chemistry*, v. 3, n. 3, p. 305–320, jan 2003. ISSN 15680266. Disponível em: <<http://www.eurekaselect.com/openurl/content.php?genre=article&iissn=1568-0266&volume=3&issue=3&spage=305>>. Citado na página 33.

93 LASSOTA, P.; SINGH, G.; KRAMER, R. Mechanism of Topoisomerase II Inhibition by Staurosporine and Other Protein Kinase Inhibitors. *Journal of Biological Chemistry*, v. 271, n. 42, p. 26418–26423, oct 1996. ISSN 00219258. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0021925818399241>>. Citado na página 33.

94 BAILLY, C. et al. DNA Cleavage by Topoisomerase I in the Presence of Indolocarbazole Derivatives of Rebeccamycin †. *Biochemistry*, v. 36, n. 13, p. 3917–3929, apr 1997. ISSN 0006-2960. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/bi9624898>>. Citado 4 vezes nas páginas 33, 34, 37 e 120.

95 BAILLY, C. et al. Calories from carbohydrates: Energetic contribution of the carbohydrate moiety of rebeccamycin to DNA binding and the effect of its orientation on topoisomerase I inhibition. *Chemistry and Biology*, v. 6, n. 5, p. 277–286, 1999. ISSN 10745521. Citado 4 vezes nas páginas 33, 34, 37 e 120.

96 STAKER, B. L. et al. Structures of Three Classes of Anticancer Agents Bound to the Human Topoisomerase IDNA Covalent Complex. *Journal of Medicinal Chemistry*, v. 48, n. 7, p. 2336–2345, apr 2005. ISSN 0022-2623. Disponível em: <<https://pubs.acs.org/doi/10.1021/jm049146p>>. Citado 2 vezes nas páginas 34 e 35.

97 ZHANG, G. et al. Syntheses and Biological Activities of Rebeccamycin Analogues with Uncommon Sugars. *Journal of Medicinal Chemistry*, v. 48, n. 7, p. 2600–2611, apr 2005. ISSN 0022-2623. Disponível em: <<https://pubs-acscs.org.ez45.periodicos.capes.gov.br/doi/pdf/10.1021/jm0493764http://pubs.acs.org/doi/abs/10.1021/jm0493764>>. Citado 2 vezes nas páginas 34 e 37.

98 ANIZON, F. et al. Rebeccamycin analogues bearing amine substituents or other groups on the sugar moiety. *Bioorganic & Medicinal Chemistry*, v. 11, n. 17, p. 3709–3722,

aug 2003. ISSN 09680896. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0968089603003432>>. Citado na página 34.

99 MOREAU, P. et al. Semi-synthesis, topoisomerase I and kinases inhibitory properties, and antiproliferative activities of new rebeccamycin derivatives. *Bioorganic & Medicinal Chemistry*, v. 11, n. 23, p. 4871–4879, nov 2003. ISSN 09680896. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S096808960300628X>>. Citado 2 vezes nas páginas 34 e 35.

100 FUKASAWA, K. et al. Sequence-selective DNA cleavage by a topoisomerase I poison, NB-506. *International Journal of Cancer*, v. 75, n. 1, p. 145–150, jan 1998. ISSN 00207136. Disponível em: <[https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1097-0215\(19980105\)75:1%3C145::AID-IJC22%3E3.0.CO;2-E](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0215(19980105)75:1%3C145::AID-IJC22%3E3.0.CO;2-E)>. Citado na página 34.

101 YOSHINARI, T. et al. Mode of action of a new indolocarbazole anticancer agent, J-107088, targeting topoisomerase I. *Cancer research*, v. 59, n. 17, p. 4271–5, sep 1999. ISSN 0008-5472. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/10485471>>. Citado na página 34.

102 ONAKA, H. et al. Characterization of the Biosynthetic Gene Cluster of Rebeccamycin from *Lechevalieria aerocolonigenes* ATCC 39243. *Bioscience, Biotechnology, and Biochemistry*, v. 67, n. 1, p. 127–138, jan 2003. ISSN 0916-8451. Disponível em: <<https://academic.oup.com/bbb/article/67/1/127-138/5944355>>. Citado na página 35.

103 GAO, Q. et al. Deciphering Indolocarbazole and Eneidyne Aminodideoxypentose Biosynthesis through Comparative Genomics: Insights from the AT2433 Biosynthetic Locus. *Chemistry Biology*, v. 13, n. 7, p. 733–743, jul 2006. ISSN 10745521. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1074552106001712>>. Citado na página 35.

104 ZHANG, C. et al. RebG- and RebM-Catalyzed Indolocarbazole Diversification. *ChemBioChem*, v. 7, n. 5, p. 795–804, may 2006. ISSN 14394227. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/cbic.200500504>>. Citado na página 35.

105 ZHANG, C. et al. Natural Product Diversification Using a Non-natural Cofactor Analogue of S -Adenosyl- l -methionine. *Journal of the American Chemical Society*, v. 128, n. 9, p. 2760–2761, mar 2006. ISSN 0002-7863. Disponível em: <<https://pubs.acs.org/doi/10.1021/ja056231t>>. Citado na página 35.

106 ADASME, M. F. et al. PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic Acids Research*, v. 49, n. W1, p. W530–W534, jul 2021. ISSN 0305-1048. Disponível em: <<https://academic.oup.com/nar/article/49/W1/W530/6266421>>. Citado na página 35.

107 SCHÖNING-STIERAND, K. et al. ProteinsPlus: interactive analysis of protein–ligand binding interfaces. *Nucleic Acids Research*, v. 48, n. W1, p. W48–W53, jul 2020. ISSN 0305-1048. Disponível em: <<https://academic.oup.com/nar/article/48/W1/W48/5820880>>. Citado na página 35.

108 PILCH, B. et al. Specific Inhibition of Serine- and Arginine-rich Splicing Factors Phosphorylation, Spliceosome Assembly, and Splicing by the Antitumor Drug NB-5061. *CANCER RESEARCH*, v. 61, n. 18, p. 6876–6884, 2001. Citado na página 35.

- 109 NOCK, C. J. et al. A phase I study of rebeccamycin analog in combination with oxaliplatin in patients with refractory solid tumors. *Investigational New Drugs*, v. 29, n. 1, p. 126–130, feb 2011. ISSN 0167-6997. Disponível em: <<http://link.springer.com/10.1007/s10637-009-9322-9>>. Citado na página 35.
- 110 DAUBER-OSGUTHORPE, P.; HAGLER, A. T. Biomolecular force fields: where have we been, where are we now, where do we need to go and how do we get there? *Journal of Computer-Aided Molecular Design*, v. 33, n. 2, p. 133–203, feb 2019. ISSN 0920-654X. Disponível em: <<http://link.springer.com/10.1007/s10822-018-0111-4>>. Citado 4 vezes nas páginas 36, 45, 46 e 47.
- 111 MCNAUGHT, A. D.; WILKINSON, A. *Compendium of Chemical Terminology: IUPAC Recommendations*. 2nd ed.. ed. Oxford: Blackwell Science, 1997. 464 p. ISBN 0-86542-684-8. Citado na página 36.
- 112 TOUKACH, F. V.; ANANIKOV, V. P. Recent advances in computational predictions of NMR parameters for the structure elucidation of carbohydrates: methods and limitations. *Chemical Society Reviews*, v. 42, n. 21, p. 8376, 2013. ISSN 0306-0012. Disponível em: <<http://xlink.rsc.org/?DOI=c3cs60073d>>. Citado na página 37.
- 113 IMBERTY, A.; PÉREZ, S. Structure, Conformation, and Dynamics of Bioactive Oligosaccharides: Theoretical Approaches and Experimental Validations. *Chemical Reviews*, v. 100, n. 12, p. 4567–4588, dec 2000. ISSN 0009-2665. Disponível em: <<https://pubs.acs.org/doi/10.1021/cr990343j>>. Citado na página 37.
- 114 HRICOVINI, M. Structural Aspects of Carbohydrates and the Relation with their Biological Properties. *Current Medicinal Chemistry*, v. 11, n. 19, p. 2565–2583, oct 2004. ISSN 09298673. Disponível em: <<https://www.eurekaselect.com/62188/article>>. Citado na página 37.
- 115 MEIRELLES, J. L. de et al. Current Status of Carbohydrates Information in the Protein Data Bank. *Journal of Chemical Information and Modeling*, v. 60, n. 2, p. 684–699, feb 2020. ISSN 1549-9596. Disponível em: <<https://pubs.acs.org/doi/10.1021/acs.jcim.9b00874>>. Citado na página 37.
- 116 DUUS, J. Ø.; GOTFREDSEN, C. H.; BOCK, K. Carbohydrate Structural Determination by NMR Spectroscopy: Modern Methods and Limitations. *Chemical Reviews*, v. 100, n. 12, p. 4589–4614, dec 2000. ISSN 0009-2665. Disponível em: <<https://pubs.acs.org/doi/10.1021/cr990302n>>. Citado na página 37.
- 117 FRANK, M.; SCHLOISSNIG, S. Bioinformatics and molecular modeling in glycobiology. *Cellular and Molecular Life Sciences*, v. 67, n. 16, p. 2749–2772, aug 2010. ISSN 1420-682X. Disponível em: <<http://link.springer.com/10.1007/s00018-010-0352-4>>. Citado na página 37.
- 118 VALCÁRCEL, M. J. P.; GONZÁLEZ, S. H.; VALCÁRCEL, M. D. H. Espectroscopia. fundamento de la resonancia magnética nuclear. *Entornos*, v. 1, n. 21, p. 89–94, 2008. Citado 3 vezes nas páginas 37, 38 e 39.
- 119 KEELER, J. *Understanding NMR Spectroscopy*. 1st ed.. ed. Cambridge: Wiley, 2005. 476 p. ISBN 978-0470017876. Citado na página 38.

- 120 VERLI, H. *Bioinformática da Biologia à flexibilidade molecular*. 1a. ed. ed. São Paulo: SBBq, 2014. 282 p. ISBN 978-85-69288-00-8. Citado 2 vezes nas páginas 39 e 42.
- 121 WOODS, R. J. Three-dimensional structures of oligosaccharides. *Current Opinion in Structural Biology*, v. 5, n. 5, p. 591–598, oct 1995. ISSN 0959440X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/0959440X95800492>>. Citado na página 39.
- 122 KAMIYA, Y.; SATOH, T.; KATO, K. Recent advances in glycoprotein production for structural biology: toward tailored design of glycoforms. *Current Opinion in Structural Biology*, v. 26, p. 44–53, jun 2014. ISSN 0959440X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0959440X14000372>>. Citado na página 39.
- 123 KATO, K.; YAMAGUCHI, T. Paramagnetic NMR probes for characterization of the dynamic conformations and interactions of oligosaccharides. *Glycoconjugate Journal*, v. 32, n. 7, p. 505–513, oct 2015. ISSN 0282-0080. Disponível em: <<http://link.springer.com/10.1007/s10719-015-9599-1>>. Citado na página 39.
- 124 KARPLUS, M. Contact Electron-Spin Coupling of Nuclear Magnetic Moments. *The Journal of Chemical Physics*, v. 30, n. 1, p. 11–15, jan 1959. ISSN 0021-9606. Disponível em: <<http://aip.scitation.org/doi/10.1063/1.1729860>>. Citado na página 40.
- 125 KARPLUS, M. Vicinal Proton Coupling in Nuclear Magnetic Resonance. *Journal of the American Chemical Society*, v. 85, n. 18, p. 2870–2871, sep 1963. ISSN 0002-7863. Disponível em: <<https://pubs.acs.org/doi/abs/10.1021/ja00901a059>>. Citado na página 40.
- 126 COXON, B. Chapter 3 Developments in the Karplus Equation as they Relate to the NMR Coupling Constants of Carbohydrates. In: . [s.n.], 2009. p. 17–82. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0065231809000031>>. Citado 2 vezes nas páginas 40 e 119.
- 127 HAASNOOT, C. A. G. et al. Interpretation of vicinal proton-proton coupling constants by a generalized Karplus relation. Conformational analysis of the exocyclic C4-C5 bond in nucleosides and nucleotides: Preliminary Communication. *Recueil des Travaux Chimiques des Pays-Bas*, v. 98, n. 12, p. 576–577, 1979. ISSN 01650513. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/recl.19790981206>>. Citado na página 40.
- 128 HAASNOOT, C.; LEEUW, F. de; ALTONA, C. The relationship between proton-proton NMR coupling constants and substituent electronegativities—I. *Tetrahedron*, v. 36, n. 19, p. 2783–2792, jan 1980. ISSN 00404020. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/0040402080801554>>. Citado na página 40.
- 129 HAASNOOT, C. A. G. et al. The relationship between proton-proton NMR coupling constants and substituent electronegativities. II—conformational analysis of the sugar ring in nucleosides and nucleotides in solution using a generalized Karplus equation. *Organic Magnetic Resonance*, v. 15, n. 1, p. 43–52, jan 1981. ISSN 0030-4921. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/mrc.1270150111>>. Citado na página 40.
- 130 ALTONA, C.; HAASNOOT, C. A. G. Prediction of anti and gauche vicinal proton-proton coupling constants in carbohydrates: A simple additivity rule for pyranose

- rings. *Organic Magnetic Resonance*, v. 13, n. 6, p. 417–429, jun 1980. ISSN 0030-4921. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/mrc.1270130606>>. Citado na página 40.
- 131 HUGGINS, M. L. Bond Energies and Polarities 1. *Journal of the American Chemical Society*, v. 75, n. 17, p. 4123–4126, sep 1953. ISSN 0002-7863. Disponível em: <<https://pubs.acs.org/doi/abs/10.1021/ja01113a001>>. Citado na página 40.
- 132 CONTRERAS, R. Angular dependence of spin–spin coupling constants. *Progress in Nuclear Magnetic Resonance Spectroscopy*, v. 37, n. 4, p. 321–425, nov 2000. ISSN 00796565. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0079656500000273>>. Citado 2 vezes nas páginas 41 e 119.
- 133 PEDEBOS, C.; POL-FACHIN, L.; VERLI, H. Unrestrained Conformational Characterization of Stenocereus eruca Saponins in Aqueous and Nonaqueous Solvents. *Journal of Natural Products*, v. 75, n. 6, p. 1196–1200, jun 2012. ISSN 0163-3864. Disponível em: <<https://pubs.acs.org/doi/10.1021/np3000393>>. Citado na página 41.
- 134 ARANTES, P. R. et al. Development of GROMOS-Compatible Parameter Set for Simulations of Chalcones and Flavonoids. *The Journal of Physical Chemistry B*, v. 123, n. 5, p. 994–1008, feb 2019. ISSN 1520-6106. Disponível em: <<http://pubs.acs.org/doi/10.1021/acs.jpcc.8b10139>>. Citado 4 vezes nas páginas 41, 47, 55 e 56.
- 135 MALDE, A. K. et al. An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0. *Journal of Chemical Theory and Computation*, v. 7, n. 12, p. 4026–4037, dec 2011. ISSN 1549-9618. Disponível em: <<https://pubs.acs.org/doi/10.1021/ct200196m>>. Citado 6 vezes nas páginas 41, 46, 47, 48, 55 e 117.
- 136 RINIKER, S. Fixed-Charge Atomistic Force Fields for Molecular Dynamics Simulations in the Condensed Phase: An Overview. *Journal of Chemical Information and Modeling*, v. 58, n. 3, p. 565–578, mar 2018. ISSN 1549-9596. Disponível em: <<https://pubs.acs.org/doi/10.1021/acs.jcim.8b00042>>. Citado 5 vezes nas páginas 41, 44, 46, 48 e 119.
- 137 HORTA, B. A. C. et al. A GROMOS-Compatible Force Field for Small Organic Molecules in the Condensed Phase: The 2016H66 Parameter Set. *Journal of Chemical Theory and Computation*, v. 12, n. 8, p. 3825–3850, aug 2016. ISSN 1549-9618. Disponível em: <<https://pubs.acs.org/doi/10.1021/acs.jctc.6b00187>>. Citado na página 41.
- 138 SCHLICK, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide*. New York, NY: Springer New York, 2010. v. 21. (Interdisciplinary Applied Mathematics, v. 21). ISBN 978-1-4419-6350-5. Disponível em: <<http://link.springer.com/10.1007/978-1-4419-6351-2>>. Citado na página 42.
- 139 LEACH, A. *Molecular Modelling: Principles and Applications*. Prentice Hall, 2001. ISBN 9780582382107. Disponível em: <<https://books.google.com.br/books?id=kB7jsbV-uhkC>>. Citado na página 42.
- 140 FRENKEL, D.; SMIT, B. *Understanding Molecular Simulation From Algorithms to Applications*. 2nd. ed. San Diego: Academic Press, 2001. 658 p. ISBN 9780122673511. Citado na página 44.

- 141 SOARES, T. A. et al. Validation of the GROMOS force-field parameter set 45A3 against nuclear magnetic resonance data of hen egg lysozyme. *Journal of Biomolecular NMR*, v. 30, n. 4, p. 407–422, dec 2004. ISSN 0925-2738. Disponível em: <<http://link.springer.com/10.1007/s10858-004-5430-1>>. Citado na página 44.
- 142 DAURA, X.; MARK, A. E.; Van Gunsteren, W. F. Parametrization of aliphatic CH<sub>n</sub> united atoms of GROMOS96 force field. *Journal of Computational Chemistry*, v. 19, n. 5, p. 535–547, apr 1998. ISSN 0192-8651. Disponível em: <[https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1096-987X\(19980415\)19:5%3C535::AID-JCC6%3E3.0.CO;2-N](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1096-987X(19980415)19:5%3C535::AID-JCC6%3E3.0.CO;2-N)>. Citado na página 44.
- 143 LEVITT, M.; LIFSON, S. Refinement of protein conformations using a macromolecular energy minimization procedure. *Journal of Molecular Biology*, v. 46, n. 2, p. 269–279, dec 1969. ISSN 00222836. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/002228369904215>>. Citado na página 44.
- 144 SCOTT, R. A.; SCHERAGA, H. A. Conformational Analysis of Macromolecules. II. The Rotational Isomeric States of the Normal Hydrocarbons. *The Journal of Chemical Physics*, v. 44, n. 8, p. 3054–3069, apr 1966. ISSN 0021-9606. Disponível em: <<http://aip.scitation.org/doi/10.1063/1.1727180>>. Citado na página 44.
- 145 SCOTT, R. A.; SCHERAGA, H. A. Conformational analysis of macromolecules. I. Ethane, propane, n-butane, and n-pentane. *Biopolymers*, v. 4, n. 2, p. 237–238, feb 1966. ISSN 0006-3525. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/bip.1966.360040210>>. Citado na página 44.
- 146 JONES, J. On the determination of molecular fields. —II. From the equation of state of a gas. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, v. 106, n. 738, p. 463–477, oct 1924. ISSN 0950-1207. Disponível em: <<https://royalsocietypublishing.org/doi/10.1098/rspa.1924.0082>>. Citado na página 44.
- 147 SPOEL, D. van der. Systematic design of biomolecular force fields. *Current Opinion in Structural Biology*, v. 67, p. 18–24, apr 2021. ISSN 0959440X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0959440X2030141X>>. Citado na página 45.
- 148 JING, Z. et al. Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. *Annual Review of Biophysics*, v. 48, n. 1, p. 371–394, may 2019. ISSN 1936-122X. Disponível em: <<https://www.annualreviews.org/doi/10.1146/annurev-biophys-070317-033349>>. Citado na página 45.
- 149 CORNELL, W. D. et al. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society*, v. 117, n. 19, p. 5179–5197, may 1995. ISSN 0002-7863. Disponível em: <<https://pubs.acs.org/doi/abs/10.1021/ja00124a002>>. Citado na página 45.
- 150 LINDORFF-LARSEN, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics*, v. 78, n. 8, p. 1950–1958, jun 2010. ISSN 08873585. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/prot.22711>>. Citado na página 45.

- 151 JORGENSEN, W. L.; TIRADO-RIVES, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, v. 110, n. 6, p. 1657–1666, mar 1988. ISSN 0002-7863. Disponível em: <<https://pubs.acs.org/doi/abs/10.1021/ja00214a001>>. Citado na página 45.
- 152 HARDER, E. et al. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *Journal of Chemical Theory and Computation*, v. 12, n. 1, p. 281–296, jan 2016. ISSN 1549-9618. Disponível em: <<https://pubs.acs.org/doi/10.1021/acs.jctc.5b00864>>. Citado na página 45.
- 153 MACKERELL, A. D. et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *The Journal of Physical Chemistry B*, v. 102, n. 18, p. 3586–3616, apr 1998. ISSN 1520-6106. Disponível em: <<https://pubs.acs.org/doi/10.1021/jp973084f>>. Citado na página 45.
- 154 VANOMMESLAEGHE, K. et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of Computational Chemistry*, p. NA–NA, 2009. ISSN 01928651. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/jcc.21367>>. Citado na página 45.
- 155 OOSTENBRINK, C. et al. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, v. 25, n. 13, p. 1656–1676, oct 2004. ISSN 0192-8651. Disponível em: <<http://doi.wiley.com/10.1002/jcc.20090>>. Citado 6 vezes nas páginas 46, 47, 48, 55, 58 e 118.
- 156 SCHMID, N. et al. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *European Biophysics Journal*, v. 40, n. 7, p. 843–856, jul 2011. ISSN 0175-7571. Disponível em: <<http://link.springer.com/10.1007/s00249-011-0700-9>>. Citado 3 vezes nas páginas 46, 55 e 58.
- 157 BLEIZIFFER, P.; SCHALLER, K.; RINIKER, S. Machine Learning of Partial Charges Derived from High-Quality Quantum-Mechanical Calculations. *Journal of Chemical Information and Modeling*, v. 58, n. 3, p. 579–590, mar 2018. ISSN 1549-9596. Disponível em: <<https://pubs.acs.org/doi/10.1021/acs.jcim.7b00663>>. Citado na página 46.
- 158 WANG, J. et al. Development and testing of a general amber force field. *Journal of Computational Chemistry*, v. 25, n. 9, p. 1157–1174, jul 2004. ISSN 0192-8651. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/jcc.20035>>. Citado na página 46.
- 159 JORGENSEN, W. L.; MAXWELL, D. S.; TIRADO-RIVES, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society*, v. 118, n. 45, p. 11225–11236, nov 1996. ISSN 0002-7863. Disponível em: <<https://pubs.acs.org/doi/10.1021/ja9621760>>. Citado na página 46.
- 160 JORGENSEN, W. L.; TIRADO-RIVES, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proceedings of the National Academy of Sciences*, v. 102, n. 19, p. 6665–6670, may 2005. ISSN 0027-8424. Disponível em: <<http://www.pnas.org/cgi/doi/10.1073/pnas.0408037102>>. Citado na página 46.



- 161 VANOMMESLAEGHE, K.; MACKERELL, A. D. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *Journal of Chemical Information and Modeling*, v. 52, n. 12, p. 3144–3154, dec 2012. ISSN 1549-9596. Disponível em: <<https://pubs.acs.org/doi/10.1021/ci300363c>>. Citado na página 46.
- 162 VANOMMESLAEGHE, K.; RAMAN, E. P.; MACKERELL, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded Parameters and Partial Atomic Charges. *Journal of Chemical Information and Modeling*, v. 52, n. 12, p. 3155–3168, dec 2012. ISSN 1549-9596. Disponível em: <<https://pubs.acs.org/doi/10.1021/ci3003649>>. Citado na página 46.
- 163 KOZIARA, K. B. et al. Testing and validation of the Automated Topology Builder (ATB) version 2.0: prediction of hydration free enthalpies. *Journal of Computer-Aided Molecular Design*, v. 28, n. 3, p. 221–233, mar 2014. ISSN 0920-654X. Disponível em: <<http://link.springer.com/10.1007/s10822-014-9713-7>>. Citado na página 46.
- 164 STROET, M. et al. Automated Topology Builder Version 3.0: Prediction of Solvation Free Enthalpies in Water and Hexane. *Journal of Chemical Theory and Computation*, v. 14, n. 11, p. 5834–5845, nov 2018. ISSN 1549-9618. Disponível em: <<https://pubs.acs.org/doi/10.1021/acs.jctc.8b00768>>. Citado 4 vezes nas páginas 46, 47, 48 e 118.
- 165 HALGREN, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, v. 17, n. 5-6, p. 490–519, apr 1996. ISSN 01928651. Disponível em: <[https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1096-987X\(199604\)17:5/6%3C490::AID-JCC1%3E3.0.CO;2-P](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1096-987X(199604)17:5/6%3C490::AID-JCC1%3E3.0.CO;2-P)>. Citado 2 vezes nas páginas 46 e 47.
- 166 CASEWIT, C. J.; COLWELL, K. S.; RAPPE, A. K. Application of a universal force field to organic molecules. *Journal of the American Chemical Society*, v. 114, n. 25, p. 10035–10046, dec 1992. ISSN 0002-7863. Disponível em: <<https://pubs.acs.org/doi/abs/10.1021/ja00051a041>>. Citado na página 46.
- 167 HALGREN, T. A. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *Journal of Computational Chemistry*, v. 17, n. 5-6, p. 520–552, apr 1996. ISSN 01928651. Disponível em: <[https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1096-987X\(199604\)17:5/6%3C520::AID-JCC2%3E3.0.CO;2-W](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1096-987X(199604)17:5/6%3C520::AID-JCC2%3E3.0.CO;2-W)>. Citado na página 46.
- 168 HALGREN, T. A. Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. *Journal of Computational Chemistry*, v. 17, n. 5-6, p. 553–586, apr 1996. ISSN 01928651. Disponível em: <[https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1096-987X\(199604\)17:5/6%3C553::AID-JCC3%3E3.0.CO;2-T](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1096-987X(199604)17:5/6%3C553::AID-JCC3%3E3.0.CO;2-T)>. Citado na página 46.
- 169 HALGREN, T. A.; NACHBAR, R. B. Merck molecular force field. IV. conformational energies and geometries for MMFF94. *Journal of Computational Chemistry*, v. 17, n. 5-6, p. 587–615, apr 1996. ISSN 01928651. Disponível em: <[https://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1096-987X\(199604\)17:5/6%3C587::AID-JCC4%3E3.0.CO;2-Q](https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1096-987X(199604)17:5/6%3C587::AID-JCC4%3E3.0.CO;2-Q)>. Citado na página 46.

- 170 GUNSTEREN, W. F. van et al. Validation of Molecular Simulation: An Overview of Issues. *Angewandte Chemie International Edition*, v. 57, n. 4, p. 884–902, jan 2018. ISSN 14337851. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/anie.201702945>>. Citado 2 vezes nas páginas 47 e 117.
- 171 BORDNER, A. J.; CAVASOTTO, C. N.; ABAGYAN, R. A. Direct Derivation of van der Waals Force Field Parameters from Quantum Mechanical Interaction Energies. *The Journal of Physical Chemistry B*, v. 107, n. 35, p. 9601–9609, sep 2003. ISSN 1520-6106. Disponível em: <<https://pubs.acs.org/doi/10.1021/jp034164o>>. Citado na página 47.
- 172 LEMKUL, J. A.; ALLEN, W. J.; BEVAN, D. R. Practical Considerations for Building GROMOS-Compatible Small-Molecule Topologies. *Journal of Chemical Information and Modeling*, v. 50, n. 12, p. 2221–2235, dec 2010. ISSN 1549-9596. Disponível em: <<https://pubs.acs.org/doi/10.1021/ci100335w>>. Citado 6 vezes nas páginas 47, 48, 55, 56, 118 e 120.
- 173 POL-FACHIN, L. et al. GROMOS 53A6 <sub>GLYC</sub>, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates. *Journal of Chemical Theory and Computation*, v. 8, n. 11, p. 4681–4690, nov 2012. ISSN 1549-9618. Disponível em: <<http://pubs.acs.org/doi/10.1021/ct300479h>>. Citado 3 vezes nas páginas 47, 55 e 58.
- 174 POL-FACHIN, L.; VERLI, H.; LINS, R. D. Extension and validation of the GROMOS 53A6 glyc parameter set for glycoproteins. *Journal of Computational Chemistry*, v. 35, n. 29, p. 2087–2095, nov 2014. ISSN 01928651. Disponível em: <<http://doi.wiley.com/10.1002/jcc.23721>>. Citado na página 47.
- 175 POLÊTO, M. D. et al. Aromatic Rings Commonly Used in Medicinal Chemistry: Force Fields Comparison and Interactions With Water Toward the Design of New Chemical Entities. *Frontiers in Pharmacology*, v. 9, apr 2018. ISSN 1663-9812. Disponível em: <<http://journal.frontiersin.org/article/10.3389/fphar.2018.00395/full>>. Citado 5 vezes nas páginas 47, 55, 56, 57 e 118.
- 176 P. Oliveira, M.; HÜNENBERGER, P. H. Systematic optimization of a fragment-based force field against experimental pure-liquid properties considering large compound families: application to oxygen and nitrogen compounds. *Physical Chemistry Chemical Physics*, v. 23, n. 33, p. 17774–17793, 2021. ISSN 1463-9076. Disponível em: <<http://xlink.rsc.org/?DOI=D1CP02001C>>. Citado 2 vezes nas páginas 48 e 117.
- 177 BAYLY, C. I. et al. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *The Journal of Physical Chemistry*, v. 97, n. 40, p. 10269–10280, oct 1993. ISSN 0022-3654. Disponível em: <<http://pubs.acs.org/doi/abs/10.1021/j100142a004>>. Citado 2 vezes nas páginas 48 e 55.
- 178 JHOTI, H.; REES, S.; SOLARI, R. High-throughput screening and structure-based approaches to hit discovery: is there a clear winner? *Expert Opinion on Drug Discovery*, v. 8, n. 12, p. 1449–1453, dec 2013. ISSN 1746-0441. Disponível em: <<http://www.tandfonline.com/doi/full/10.1517/17460441.2013.857654>>. Citado na página 49.

- 179 LAVECCHIA, A.; GIOVANNI, C. Virtual Screening Strategies in Drug Discovery: A Critical Review. *Current Medicinal Chemistry*, v. 20, n. 23, p. 2839–2860, jun 2013. ISSN 09298673. Disponível em: <<http://www.eurekaselect.com/openurl/content.php?genre=article&issn=0929-8673&volume=20&issue=23&spage=2839>>. Citado na página 49.
- 180 LAVECCHIA, A.; CERCHIA, C. In silico methods to address polypharmacology: current status, applications and future perspectives. *Drug Discovery Today*, v. 21, n. 2, p. 288–298, feb 2016. ISSN 13596446. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S1359644615004596>>. Citado na página 49.
- 181 MACALINO, S. J. Y. et al. Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research*, v. 38, n. 9, p. 1686–1701, sep 2015. ISSN 0253-6269. Disponível em: <<http://link.springer.com/10.1007/s12272-015-0640-5>>. Citado na página 49.
- 182 CUI, W. et al. Discovering Anti-Cancer Drugs via Computational Methods. *Frontiers in Pharmacology*, v. 11, may 2020. ISSN 1663-9812. Disponível em: <<https://www.frontiersin.org/article/10.3389/fphar.2020.00733/full>>. Citado na página 49.
- 183 SLIWOSKI, G. et al. Computational Methods in Drug Discovery. *Pharmacological Reviews*, v. 66, n. 1, p. 334–395, jan 2014. ISSN 0031-6997. Disponível em: <<http://pharmrev.aspetjournals.org/lookup/doi/10.1124/pr.112.007336>>. Citado na página 49.
- 184 MAIA, E. H. B. et al. Structure-Based Virtual Screening: From Classical to Artificial Intelligence. *Frontiers in Chemistry*, v. 8, apr 2020. ISSN 2296-2646. Disponível em: <<https://www.frontiersin.org/article/10.3389/fchem.2020.00343/full>>. Citado na página 49.
- 185 BROGI, S. et al. Editorial: In silico Methods for Drug Design and Discovery. *Frontiers in Chemistry*, v. 8, aug 2020. ISSN 2296-2646. Disponível em: <<https://www.frontiersin.org/article/10.3389/fchem.2020.00612/full>>. Citado na página 49.
- 186 MACARRON, R. et al. Impact of high-throughput screening in biomedical research. *Nature Reviews Drug Discovery*, v. 10, n. 3, p. 188–195, mar 2011. ISSN 1474-1776. Disponível em: <<http://www.nature.com/articles/nrd3368>>. Citado na página 49.
- 187 LIONTA, E. et al. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Current Topics in Medicinal Chemistry*, v. 14, n. 16, p. 1923–1938, oct 2014. ISSN 15680266. Disponível em: <<http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1568-0266&volume=14&issue=16&spage=1923>>. Citado na página 49.
- 188 LEELANANDA, S. P.; LINDERT, S. Computational methods in drug discovery. *Beilstein Journal of Organic Chemistry*, v. 12, p. 2694–2718, dec 2016. ISSN 1860-5397. Disponível em: <<https://www.beilstein-journals.org/bjoc/articles/12/267>>. Citado na página 49.
- 189 DAS, P. S. A REVIEW ON COMPUTER AIDED DRUG DESIGN IN DRUG DISCOVERY. *World Journal of Pharmacy and Pharmaceutical Sciences*, p. 279–291, jul 2017. ISSN 22784357. Disponível em: <[http://wjpps.com/wjpps\\_controller/abstract\\_id/7294](http://wjpps.com/wjpps_controller/abstract_id/7294)>. Citado na página 49.

- 190 NERO, T. L.; PARKER, M. W.; MORTON, C. J. Protein structure and computational drug discovery. *Biochemical Society Transactions*, v. 46, n. 5, p. 1367–1379, oct 2018. ISSN 0300-5127. Disponível em: <<https://portlandpress.com/biochemsoctrans/article/46/5/1367/67796/Protein-structure-and-computational-drug-discovery>>. Citado na página 49.
- 191 BLUNDELL, C. D.; NOWAK, T.; WATSON, M. J. Measurement, Interpretation and Use of Free Ligand Solution Conformations in Drug Discovery. In: . [s.n.], 2016. p. 45–147. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0079646815000181>>. Citado na página 50.
- 192 BUSSI, G.; LAIO, A. Using metadynamics to explore complex free-energy landscapes. *Nature Reviews Physics*, v. 2, n. 4, p. 200–212, apr 2020. ISSN 2522-5820. Disponível em: <<http://www.nature.com/articles/s42254-020-0153-0>>. Citado 3 vezes nas páginas 50, 59 e 120.
- 193 NERENBERG, P. S.; HEAD-GORDON, T. New developments in force fields for biomolecular simulations. *Current Opinion in Structural Biology*, v. 49, p. 129–138, apr 2018. ISSN 0959440X. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0959440X17301604>>. Citado na página 50.
- 194 HANWELL, M. D. et al. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, v. 4, n. 1, p. 17, dec 2012. ISSN 1758-2946. Disponível em: <<https://jcheminf.springeropen.com/articles/10.1186/1758-2946-4-17>>. Citado na página 55.
- 195 FRISCH, M. J. et al. *Gaussian09 Revision D.01*. 2013. Citado 2 vezes nas páginas 55 e 57.
- 196 MØLLER, C.; PLESSET, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Physical Review*, v. 46, n. 7, p. 618–622, oct 1934. ISSN 0031-899X. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRev.46.618>>. Citado 2 vezes nas páginas 55 e 57.
- 197 PETERSSON, G. A. et al. A complete basis set model chemistry. I. The total energies of closed-shell atoms and hydrides of the first-row elements. *The Journal of Chemical Physics*, v. 89, n. 4, p. 2193–2218, aug 1988. ISSN 0021-9606. Disponível em: <<http://aip.scitation.org/doi/10.1063/1.455064>>. Citado 2 vezes nas páginas 55 e 57.
- 198 BERENDSEN, H. J. C.; GRIGERA, J. R.; STRAATSMA, T. P. The missing term in effective pair potentials. *The Journal of Physical Chemistry*, v. 91, n. 24, p. 6269–6271, nov 1987. ISSN 0022-3654. Disponível em: <<https://pubs.acs.org/doi/10.1021/j100308a038>>. Citado 2 vezes nas páginas 55 e 58.
- 199 RUSU, V. H. et al. Rotational Profiler: A Fast, Automated, and Interactive Server to Derive Torsional Dihedral Potentials for Classical Molecular Simulations. *Journal of Chemical Information and Modeling*, p. acs.jcim.0c01168, nov 2020. ISSN 1549-9596. Disponível em: <<https://pubs.acs.org/doi/10.1021/acs.jcim.0c01168>>. Citado 2 vezes nas páginas 57 e 118.
- 200 NING, Y. et al. Pore-Scale Modeling and Simulation in Shale Gas Formations. In: *Petrophysical Characterization and Fluids Transport in Unconventional Reservoirs*.

Elsevier, 2019. p. 217–246. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/B9780128166987000115>>. Citado na página 58.

201 HESS, B. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation*, v. 4, n. 1, p. 116–122, jan 2008. ISSN 1549-9618. Disponível em: <<https://pubs.acs.org/doi/10.1021/ct700200b>>. Citado na página 58.

202 DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald: An  $\log(N)$  method for Ewald sums in large systems. *The Journal of Chemical Physics*, v. 98, n. 12, p. 10089–10092, jun 1993. ISSN 0021-9606. Disponível em: <<http://aip.scitation.org/doi/10.1063/1.464397>>. Citado na página 58.

203 PÁLL, S.; HESS, B. A flexible algorithm for calculating pair interactions on SIMD architectures. *Computer Physics Communications*, v. 184, n. 12, p. 2641–2650, dec 2013. ISSN 00104655. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0010465513001975>>. Citado na página 58.

204 GONÇALVES, Y. M. H. et al. Influence of the Treatment of Nonbonded Interactions on the Thermodynamic and Transport Properties of Pure Liquids Calculated Using the 2016H66 Force Field. *Journal of Chemical Theory and Computation*, v. 15, n. 3, p. 1806–1826, mar 2019. ISSN 1549-9618. Disponível em: <<https://pubs.acs.org/doi/10.1021/acs.jctc.8b00425>>. Citado na página 58.

205 BARDUCCI, A.; BONOMI, M.; PARRINELLO, M. Metadynamics. *WIREs Computational Molecular Science*, v. 1, n. 5, p. 826–843, sep 2011. ISSN 1759-0876. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/wcms.31>>. Citado 3 vezes nas páginas 58, 59 e 120.

206 SUTTO, L.; MARSILI, S.; GERVASIO, F. L. New advances in metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, v. 2, n. 5, p. 771–779, sep 2012. ISSN 17590876. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/wcms.1103>>. Citado na página 59.

207 CREMER, D.; POPLE, J. A. General definition of ring puckering coordinates. *Journal of the American Chemical Society*, v. 97, n. 6, p. 1354–1358, mar 1975. ISSN 0002-7863. Disponível em: <<https://pubs.acs.org/doi/abs/10.1021/ja00839a011>>. Citado na página 59.

208 TRIBELLO, G. A. et al. PLUMED 2: New feathers for an old bird. *Computer Physics Communications*, v. 185, n. 2, p. 604–613, feb 2014. ISSN 00104655. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0010465513003196>>. Citado na página 59.

209 HANSEN, H. S.; HÜNENBERGER, P. H. A reoptimized GROMOS force field for hexopyranose-based carbohydrates accounting for the relative free energies of ring conformers, anomers, epimers, hydroxymethyl rotamers, and glycosidic linkage conformers. *Journal of Computational Chemistry*, v. 32, n. 6, p. 998–1032, apr 2011. ISSN 01928651. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/jcc.21675>>. Citado na página 119.

- 210 ALTONA, C. et al. Empirical group electronegativities for vicinal NMR proton-proton couplings along a C-C bond: Solvent effects and reparameterization of the Haasnoot equation. *Magnetic Resonance in Chemistry*, v. 32, n. 11, p. 670–678, nov 1994. ISSN 0749-1581. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/mrc.1260321107>>. Citado na página 119.
- 211 STEINER, D. et al. On the calculation of 3 J  $\alpha\beta$ -coupling constants for side chains in proteins. *Journal of Biomolecular NMR*, v. 53, n. 3, p. 223–246, jul 2012. ISSN 0925-2738. Disponível em: <<http://link.springer.com/10.1007/s10858-012-9634-5>>. Citado na página 119.
- 212 ATKINS, P.; PAULA, J. D.; KEELER, J. *Atkins' physical chemistry*. 11. ed. London, England: Oxford University Press, 2017. ISBN 9780198769866. Citado na página 119.
- 213 DU, X. et al. Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. *International Journal of Molecular Sciences*, v. 17, n. 2, p. 144, jan 2016. ISSN 1422-0067. Disponível em: <<http://www.mdpi.com/1422-0067/17/2/144>>. Citado na página 121.

# Apêndices





# APÊNDICE A – Atracamento Molecular de Ligantes Piperazínicos

Em colaboração com o grupo do Prof. Dr. Ricardo Menegatti, do Laboratório de Química Farmacêutica Medicinal na Universidade Federal de Goiás, foi realizado um estudo de atracamento molecular para avaliar as interações de compostos bioativos da classe das piperazinas com os receptores alvo GABA<sub>A</sub> e 5-HT<sub>1A</sub>.

Esta análise elucidou os mecanismos envolvidos na atividade ansiolítica de LQFM032 e LQFM212, resultante do efeito agonista da interação de ambos os ligantes com o sítio benzodiazepínico presente no complexo do receptor GABA<sub>A</sub>. O ensaio de atracamento molecular contribuiu ainda para compreender a diferença observada com relação aos efeitos de LQFM032 e LQFM212 sobre o receptor 5-HT<sub>1A</sub>. A interação diferencial desses ligantes, com resíduos envolvidos no agonismo e no antagonismo do receptor serotoninérgico, esclareceu como o efeito ansiolítico de LQFM212 pode ser mediado pela ativação do receptor 5-HT<sub>1A</sub> em caráter agonista. Diferentemente do que ocorre com o ligante LQFM032, que mostrou um padrão de interação antagonista com os resíduos do mesmo receptor.

Este trabalho resultou na publicação do artigo intitulado "*Neuropharmacological assessment in mice and molecular docking of piperazine derivative LQFM212*" na revista *Behavioural Brain Research*.



## Neuropharmacological assessment in mice and molecular docking of piperazine derivative LQFM212



Lorrane Kelle Moreira<sup>a</sup>, Adriane Ferreira de Brito<sup>a</sup>, Crisciele Fontana<sup>e</sup>, Flávio Souza de Carvalho<sup>b</sup>, Germán Sanz<sup>c</sup>, Boniek Gontijo Vaz<sup>c</sup>, Luciano Morais Lião<sup>c</sup>, Fábio Fagundes da Rocha<sup>d</sup>, Hugo Verli<sup>e</sup>, Ricardo Menegatti<sup>b</sup>, Elson Alves Costa<sup>a,\*</sup>

<sup>a</sup> Laboratory of Pharmacology of Natural and Synthetic Products, Institute of Biological Sciences, Federal University of Goiás, Campus Samambaia, Goiânia, GO, Brazil

<sup>b</sup> Laboratory of Medicinal Pharmaceutical Chemistry, Faculty of Pharmacy, Federal University of Goiás, Goiânia, GO, Brazil

<sup>c</sup> Chemistry Institute, Federal University of Goiás, Campus Samambaia, Goiânia, GO, Brazil

<sup>d</sup> Department of Physiological Sciences, Institute of Biology, Federal Rural University of Rio de Janeiro, Seropédica, RJ, Brazil

<sup>e</sup> Center of Biotechnology, Federal University of Rio Grande do Sul, RS, Brazil

### ARTICLE INFO

#### Keywords:

Anxiolytic-like activity  
Benzodiazepine site  
Nicotinic receptor  
Serotonergic pathway

### ABSTRACT

Piperazine derivatives are an attractive class of chemical compounds for the treatment of various mental illness. Herein, we demonstrated the synthesis of LQFM212, a piperazine derivative, behavioral evaluation in mice and computational studies. In neuropharmacological assessment, LQFM212 treatment at doses of 18, 54 or 162  $\mu\text{mol/kg}$  increased the sleep duration in sodium pentobarbital-induced sleep test. LQFM212 at dose of 162  $\mu\text{mol/kg}$  increased climbing time in the chimney test and decreased the number of squares crossed in the open field test, suggesting that LQFM212 in high doses reduces spontaneous movement. However, LQFM212 treatment at the doses of 18 or 54  $\mu\text{mol/kg}$  increased the preference for the center of field which could be indicative of anxiolytic-like effects. In elevated plus maze and light-dark box tests, LQFM212 treatment altered all parameters observed that demonstrate anxiolytic-like activity. These effects were reversed by flumazenil, mecamlamine, WAY-100635 and PCPA, but not with ketanserin, showing that anxiolytic-like activity involve benzodiazepine site of GABA<sub>A</sub> receptor, nicotinic and serotonergic pathways. Molecular docking of LQFM212 showed that the ligand has more interactions with GABA<sub>A</sub> receptor than with 5-HT<sub>1A</sub> receptor. Despite the involvement of benzodiazepine site on anxiolytic-like effect of LQFM212, treatment with this compound did not alter cognitive function in the step-down avoidance test. In this sense, this piperazine derivative is a good prototype for treating anxiety disorders with putative mechanism of action.

### 1. Introduction

Mental illness, including anxiety disorders, are among the most prevalent and burdensome diseases worldwide with serious impairment to affected patients, resulting in significant morbidity and impairment of daily activities [1]. In anxiety disorders, occur neurochemical changes that contribute to extensive excitation of the central nervous system (CNS) and several modulators can affect brain activity in specific areas and throughout the neurocircuit system, allowing for many potential pharmacological targets [2,3].

The  $\gamma$ -aminobutyric acid (GABA) is the main inhibitory neurotransmitter in CNS, acts in control emotionality and changes in GABAergic function system appear to be a common denominator of

anxiety disorders [4–6]. The relationship between GABAergic system and behavior arose from the recognition of allosteric modulation of GABA<sub>A</sub> receptors by drugs such as benzodiazepines (BZDs), which due many adverse effects, e.g. memory impairment, motor incoordination, tolerance and dependence, are restricted to short-term treatment of anxiety disorders [7].

The first long-term treatment line for anxiety disorders consists mainly of selective serotonin reuptake inhibitors (SSRIs) because they have less adverse effects than BZDs but have therapeutic delay [7]. SSRIs restore the abnormalities of serotonergic neurotransmission by acting on the serotonin reuptake transporter (SERT). Besides that, other damage on serotonergic pathway can occur and include serotonin (5-HT) release dysregulation, related to presynaptic autoreceptors 5-HT<sub>1A</sub>/

\* Corresponding author : Laboratório de Farmacologia de Produtos Naturais, — ICB-2, Sala 209, Universidade Federal de Goiás, Campos Samambaia, CP 131, CEP 74001-970, Goiânia, GO, Brasil.

E-mail address: [xico@ufg.br](mailto:xico@ufg.br) (E.A. Costa).

<https://doi.org/10.1016/j.bbr.2020.112827>

Received 16 December 2019; Received in revised form 25 June 2020; Accepted 20 July 2020

Available online 28 July 2020

0166-4328/ © 2020 Elsevier B.V. All rights reserved.

<sup>1</sup>D, deficits in postsynaptic serotonergic receptors responsiveness and alteration of serotonin neurotransmission in non-serotonergic systems, since there are 5-HT heteroreceptors in neurons of other neurochemical systems, for example, activation of 5-HT<sub>2A/2C</sub> receptors located in GABAergic interneurons leads to an increase in release of GABA [2,8,9].

The large variety of studies both human and animals have shown that in addition to the classical neurotransmitters mentioned above, other pathways can be unregulated in mental disorders and make potential new targets for the treatment of conditions such as anxiety disorders. Alterations in signaling through nicotinic acetylcholine receptors (nAChRs) can affect many aspects of emotionality. nAChRs are expressed on many cell types throughout the CNS, e.g. neurons of hippocampus, locus coeruleus and amygdala, and modulate multiple brain circuits. Nicotine is known to stimulate the release of many different neurotransmitters including glutamate, GABA, ACh, dopamine, norepinephrine and 5-HT by activating nAChRs [10,11].

Evidence on literature shows that piperazine derivatives constitute an attractive class of chemical compounds for treatment of several mental disorders, including anxiety disorders (for review see [12]). These compounds have small size which confer the ability to cross the blood-brain barrier (BBB) and act in the CNS, producing anxiolytic [13], e.g. buspirone, antidepressant [14], e.g. trazodone, and anti-psychotic effects [15,16], e.g. clozapine.

The present paper describes the synthesis of the piperazine compound 2,6-di-*tert*-butyl-4-((4-(2-hydroxyethyl)piperazin-1-yl)methyl)phenol (3) (LQFM212), which was designed through molecular hybridization strategy between LQFM032 (1) and butylated hydroxytoluene (2) (BHT). The compound (1) has an anxiolytic-like activity mediated through GABA<sub>A</sub> receptor benzodiazepine site and nicotinic receptors but not by 5-HT<sub>1A</sub> receptors [17], and (2) is considered an antioxidant gold standard [18,19]. As illustrated in Fig. 1A, we changed 1-(phenyl)-1*H*-pyrazole (A and B) present in LQFM032 (1) by 2,6-di-*tert*-butylphenol from BHT (2), aiming produce a new compound with pharmacological activity and additional mechanisms of action.

The main objective of the present study was to assess the neuropharmacological activity of the new piperazine derivative LQFM212 (3) and the mechanisms involved in the observed activities. In the first phase, we carry out a screening CNS-related activity using well-validated animal models, followed to specific tests to assess anxiolytic-like effects and discriminate the activity found. The second phase was carried out with the experimental assays directed to evaluate the possible mechanisms of action *in vivo* using previous administration of antagonists and enzymatic inhibitors of the, GABAergic, nicotinic and serotonergic pathways. Besides that, were performed experimental session to check for possible cognitive damage. In the third phase, computational studies were performed to better clarify and establish the mechanism of action of LQFM212 (3).

## 2. Material and methods

### 2.1. Chemistry

#### 2.1.1. General methods

Reactions were monitored by TLC using commercially available precoated plates (Whatman 60 F254 silica) and developed plates were examined under UV light (254 and 365 nm). <sup>1</sup>H and <sup>13</sup>C-NMR spectra were recorded in the indicated solvent on Bruker Avance III 500 MHz spectrometer. Chemical shifts are quoted in parts per million downfield from TMS and the coupling constants are in Hertz. All assignments of the signals of <sup>1</sup>H and <sup>13</sup>C NMR spectra are consistent with the chemical structures of the products described. Infrared spectra were recorded on a Perkin-Elmer Spectrum Bx-II FT-IR System spectrophotometer instrument as films on KBr discs. Mass spectra (MS) were obtained with a microTOF III (Bruker Daltonics Bremen, Germany). The sample preparation for MS analysis consisted of diluting 1 µg of sample in 1 ml of methanol. To perform the analysis in positive mode, 1 µL of formic acid

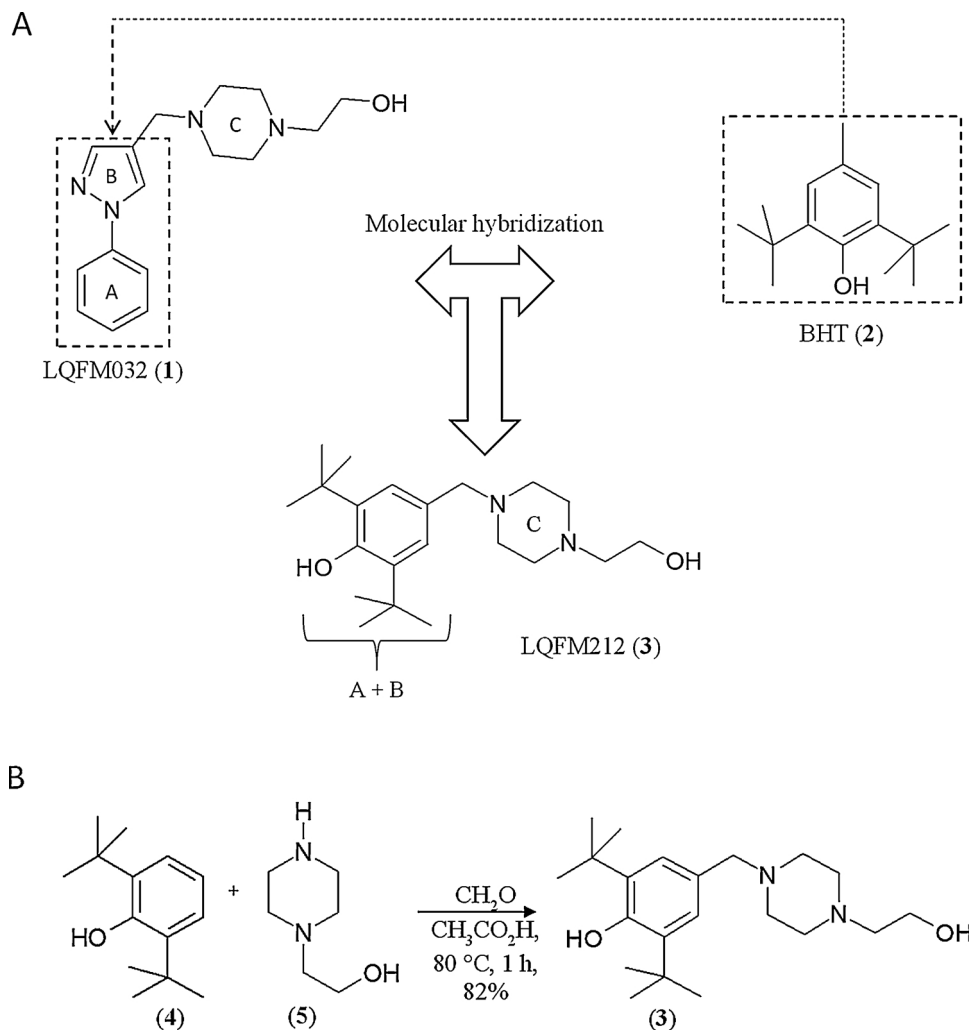
was added to the sample. The solution obtained was directly infused at a flow rate of 3 µL/min into the ESI source. ESI(+) source conditions were as follows: nebulizer nitrogen gas temperature and pressure of 2.0 bar and 200 °C, capillary voltage of -4,5 kV, transfer capillary temperature of 200 °C; drying gas of 4 L.min<sup>-1</sup>; end plate offset of -500 V; skimmer of 35 V and collision voltage of -1.5 V. Each spectrum was acquired using 2 microscans. The resolving power (m/Δm50 % 16.500,00, where Δm50 % is the peak full width at half-maximum peak height). Mass spectra were acquired and processed with data analysis software (Bruker Daltonics, Bremen, Germany). Melting points were performed using a Marté melting point apparatus, and the results were uncorrected. The organic solutions were dried over anhydrous sodium sulfate and organic solvents were removed under reduced pressure in a rotary evaporator.

#### 2.1.2. Synthesis of LQFM212 (3)

To a stirred heterogeneous mixture of 200 µl of acetic acid, (130.2 mg, 1 mmol) 1-(2-hydroxyethyl) piperazine (5), with cooling in an ice bath was added (90 µL, 1.0 mmol) 37 % aqueous of formaldehyde solution. The flask was removed from the ice bath and equipped with a reflux condenser, through which (206.3 mg, 1 mol) of 2,6-di-*tert*-butylphenol (4) was added all at once. In turn the mixture was stirred at temperature of 80 °C for 1 h (Fig. 1B) [20]. The mixture cooled to room temperature and the residue was partitioned between water and CH<sub>2</sub>Cl<sub>2</sub>. The phases were separated and the aqueous layer was extracted 3 × 15 ml with CH<sub>2</sub>Cl<sub>2</sub> and the combined organic layers were dried (Na<sub>2</sub>SO<sub>4</sub>), concentrated in vacuo, and the crude product was purified by column chromatography (SiO<sub>2</sub>, hexane/EtOAc = 73:27) to afford 2,6-di-*tert*-butyl-4-((4-(2-hydroxyethyl)piperazin-1-yl)methyl)phenol (3) (285 mg, 82 %) as a brown oil, R<sub>f</sub> = 0.65 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH, 90:10); IRmax (KBr) cm<sup>-1</sup>: 3600 (ν O-H), 3500–2500 (ν O-H), 2956 (ν C-H), 1575 (ν C = C) (Figure S1, supplementary material); <sup>1</sup>H NMR (500.13 MHz) CDCl<sub>3</sub> δ: 1.43 (s, 18H, H7' and 8'); 2.02 (s, 1H, H10); 2.65 (t, 2H, H8); 2.58 (m, 4H, H4 and H6); 2.69 (m, 4H, H3 and H7); 3.49 (s, 2H, H1); 3.67 (m, 2H, H9); 5.16 (s, 1H, H11); 7.07 (s, 2H, H2' and 6') (Figure S2 and Table 1, supplementary material); <sup>13</sup>C NMR (125.76 MHz) CDCl<sub>3</sub> δ: 30.5 (C-7' and 8'); 34.4 (C-9' and 10'); 50.4 (C-4 and 6); 52.3 (C-3 and 7); 57.4 (C-9); 59.7 (C-8); 62.4 (C-1); 124.7 (C-1'); 126.4 (C-2' and 6'); 135.7 (C-3' and 5') (Figure S3-4 and Table 1, supplementary material); [M<sup>+</sup>H]<sup>+</sup> m/z = 349.2795; purity > 99 % (Figure S5, supplementary material).

### 2.2. Drugs and treatments

The chemicals used were: LQFM212 (2,6-di-*tert*-butyl-4-((4-(2-hydroxyethyl)piperazin-1-yl)methyl)phenol) (3) – molecular weight: 348.28 g/mol; Tween 80® (Vetec, Brazil); NaCl (Belga); Diazepam (Teuto, Brazil); Sodium pentobarbital (Sigma-Aldrich, USA); 2,5-dimethoxy-4-iodoamphetamine (DOI - Sigma-Aldrich, USA); Flumazenil (União Química, Brazil); Mecamylamine hydrochloride (Sigma-Aldrich, USA); *p*-chlorophenylalanine (PCPA - Sigma-Aldrich, USA); WAY-100635 (Sigma-Aldrich, USA); Ketanserin (Sigma-Aldrich, USA). LQFM212 (3) was solubilized in 2% Tween 80® and administered orally (p.o.) at doses of 18, 54 or 162 µmol/kg. These doses are equimolar to the doses LQFM032 (1) [17]. Control group was treated orally with 2% Tween 80® (10 mL/kg) in distilled water, as vehicle. Diazepam were prepared in distilled water and administered orally at doses of 3.5 or 18 µmol/kg. Sodium pentobarbital, DOI, mecamylamine, PCPA, WAY-100635 and ketanserin were prepared in 0.9 % saline (NaCl) and administered intraperitoneally (i.p.). Doses of the drugs were chosen according to the literature data and we also confirmed the selected doses in previous studies [17,21]. In particular, experiments for 5-HT<sub>2A/2C</sub> receptors, a prior standardization was performed to find dose of ketanserin antagonist that did not interfere with results but could reverse effect of DOI agonist (Figure S6-7, supplementary material).



**Fig. 1.** Design of 2,6-di-*tert*-butyl-4-((4-(2-hydroxyethyl)piperazin-1-yl)methyl)phenol (3) (LQFM212). (A) Structural design of LQFM212 (3) from LQFM032 (1) and BHT (2). (B) Synthetic route to preparation LQFM212 (3).

### 2.3. Neuropharmacology assessment

#### 2.3.1. Animals

Experiments were conducted using adult male Swiss mice, weighing approximately 30–45 g (6–8 weeks old), provided by the Central Animal House of Universidade Federal de Goiás. The animals were maintained under controlled conditions of temperature ( $22 \pm 1^\circ\text{C}$ ), 12:12 h light/dark cycle and free access to water and food and acclimatized for 7 days before the beginning of the experiments. All protocols carried out between 8:00 a.m. and 4:00 p.m., with eight mice per group, and were approved by the Ethics Committee on Animal Use of Universidade Federal de Goiás-Brazil (no. 021/13).

#### 2.3.2. Pharmacological screening of central activity

Initially, a screening CNS-related activity was carried out, in which three tests were performed, namely: sodium pentobarbital-induced sleep, chimney and open field test (OFT). Sixty minutes after oral treatment with vehicle, LQFM212 (3) at doses of 18, 54 or 162  $\mu\text{mol}/\text{kg}$  or diazepam 18  $\mu\text{mol}/\text{kg}$ , in the sodium pentobarbital-induced sleep, the animals received an intraperitoneal administration of sodium pentobarbital at dose of 200  $\mu\text{mol}/\text{kg}$  to evaluate possible central depressant or stimulant effects of LQFM212 (3) due to synergistic or antagonism effect with barbiturate. The latency time, from pentobarbital injection to sleep onset (when occurs to loss of the righting reflex) and the duration of the sleeping time (the time required to recover the

righting reflex) were recorded for each animal [22]. The chimney test was performed in order to verify possible impairment in motor coordination. Sixty minutes after the same oral treatments, the animals were placed, individually, in a glass tube with a predetermined mark. The time that the animal took to climb backwards in the glass tube in vertical position to reach the predetermined mark has been recorded, with cut-off to climbing of 30 s. Animals that could not perform this activity were considered with impaired in motor coordination [23]. In addition, OFT was carry out to assess the mice ambulatory behavior. The apparatus is a circular arena with the floor divided into eight equal squares and sixty minutes after the same oral treatments each animal was placed at the center of arena and allowed to explore freely for 5 min. The number of squares crossed, rearing number, grooming frequency, immobility time (s), crossings frequency and time spent at the center (%CrCe and %Tce, respectively) of the arena were recorded [21,24].

#### 2.3.3. Evaluation of anxiolytic-like activity

Specific tests were performed to evaluate the anxiolytic-like activity of LQFM212 (3): elevated plus maze (EPM) and light-dark box (LDB) tests. Sixty minutes after oral treatment with vehicle, LQFM212 (3) (54  $\mu\text{mol}/\text{kg}$ ) or diazepam (3.5  $\mu\text{mol}/\text{kg}$ ), the animals were submitted to EPM and LDB tests. EPM apparatus consists of two open arms and two enclosed arms connected by a central platform and raised from the floor. The test is based on the natural aversion of mice for open and elevated

areas, as well as on their natural spontaneous exploratory behavior in novel environments [25]. In EPM protocol, the animals were placed individually at the center of the EPM, facing the closed arms and the experiment was recorded for 5 min. Parameters such as percentage of open arms entries and percentage of time spent in open arms, percentage of time spent on central platform and ethological evaluations, stretched-attend postures (SAP) and head dips, were used as a measure of anxiety-like activity [26,27]. LDB test assesses the rodent's natural preference for the dark and aversion to bright lights since the apparatus consists of two compartments, one dark and one illuminated with the floor divided into nine equal squares, separated by an opening in the central wall. In LDB protocol, the animals were placed individually at center of the light area facing the opening of the dark area. Time of second latency (first transition from dark to light area), number of transitions between the compartments, time spent and number of squares crossed in the light area were recorded over 5 min [28,29].

#### 2.3.4. Mechanism of action involved on anxiolytic-like activity

To evaluate the role of benzodiazepine site of GABA<sub>A</sub> receptor and nicotinic receptors, animals were pretreated by intraperitoneal administration with 0.9 % saline, flumazenil (6.6 µmol/kg), a benzodiazepine site antagonist, or mecamylamine (30 µmol/kg), a non-selective nicotinic receptors antagonist. After 30 min, the animals were treated orally with vehicle or LQFM212 (3) (54 µmol/kg). Sixty minutes after treatment, mice were submitted to EPM.

To evaluate the role of serotonergic pathway were used antagonists of 5-HT<sub>1A</sub> and 5-HT<sub>2</sub> receptors, and inhibitor of tryptophan hydroxylase. For participation of 5-HT<sub>1A</sub> receptors, animals were pretreated by intraperitoneal administration with 0.9 % saline or WAY-100635 (0.7 µmol/kg), a 5-HT<sub>1A</sub> receptor antagonist, and 30 min after were treated orally with vehicle or LQFM212 (3) (54 µmol/kg). For participation of 5-HT<sub>2A/2C</sub> receptors, animals were pretreated by intraperitoneal administration with 0.9 % saline or ketanserin (0.25 µmol/kg), a 5-HT<sub>2A/2C</sub> antagonist, and 30 min after, were treated with vehicle, LQFM212 (3) (54 µmol/kg, p.o.) or 2,5-dimethoxy-4-iodoamphetamine - DOI (3.35 µmol/kg, i.p.), a 5-HT<sub>2A/2C</sub> agonist. In another protocol to evaluate the influence of endogenous serotonin, animals were pretreated (i.p.) during 4 consecutive days (once a day) with 0.9 % saline or *p*-chlorophenylalanine - PCPA (500 µmol/kg), a serotonin synthesis inhibitor by inhibition of tryptophan hydroxylase. One hour after the last pretreatment, the animals were treated orally with vehicle or LQFM212 (3) (54 µmol/kg). In all protocols, one hour after treatment with vehicle or LQFM212 (3), and thirty minutes after treatment with DOI, mice were submitted to EPM.

#### 2.3.5. Evaluation of cognition activity on step-down avoidance (SDA) test

The protocol of SDA test is divided into sections, first the animals are subjected to a training session, where sixty minutes after oral treatment with vehicle, LQFM212 (3) at doses of 18, 54 or 162 µmol/kg or diazepam 18 µmol/kg mice were placed on the safe platform and the step-down from the platform with all four paws was immediately followed by a foot-shock (0.5 mA) for 1 s. The latency to step-down was recorded for basal information. Afterward, test session was carried out after an interval of 90 min and 24 h from the training session, which was conducted in the absence of shock, and the step-down latency from the platform was recorded (cut-off 180 s) and evaluated as indicative of memory retention [30].

#### 2.4. Computational studies

The molecular docking study was performed on the GABA<sub>A</sub> and 5-HT<sub>1A</sub> receptors. The GABA<sub>A</sub> structure came from the 6HUO PDB entry, that provides the α1β3γ2 GABA<sub>A</sub> receptor [31]. For the analyses involving the receptor 5-HT<sub>1A</sub> receptor we applied a model predicted by GPCR-I-TASSER designer of 3D structure prediction of G protein-coupled receptors, available in <https://zhanglab.ccmb.med.umich.edu>

[32]. The protein receptors were ionized at pH = 7.4 and the protonation state was assigned through PROPKA 3.1 web server [33,34]. The ligands structures were built in Avogadro molecular editor [35]. Beyond the ligands LQFM032 (1), BHT (2) and LQFM212 (3) tested at both receptors, diazepam and flumazenil were tested at GABA<sub>A</sub>, while 8-OH-DPAT and WAY-100635 were tested at 5-HT<sub>1A</sub>, as agonist and antagonist control, respectively.

Molecular docking was carried out through DockThor [36,37] using a grid to limit the search space to 5-HT<sub>1A</sub> inner channel, taking as references the PDB entries of 5-HT<sub>1B</sub> protein-ligand complexes 5V54 [38] and 6G79 [39] and taking care of reach as the top as the deep inner channel [40], using the grid size of 34 × 20 × 20 Å. The grid at GABA<sub>A</sub> was placed to encompass the BZD-binding site, as have showed early experimental studies involving BZDs [31,41,42], such dimensions were 20 × 20 × 28 Å, except for flumazenil whose docking required a slightly grid size reduction, measuring 20 Å at each axis. Non-covalent interaction results among the receptors and the ligands were evaluated with Protein-Ligand Interaction Profiler (PLIP) [43] and images were generated with PyMOL Molecular Graphics System 1.8 and PoseView [44,45].

#### 2.5. Statistical analysis

Results were expressed as mean ± standard error of the mean (SEM). Differences between two groups were detected by Student's T-test and among three or more groups by the one-way ANOVA followed by Newman-Keuls test. Statistical difference was considered when *p* < 0.05. GraphPad Prism software was used for all analyzes.

### 3. Results

#### 3.1. Synthesis of LQFM212 (3)

Compound LQFM212 (3) was obtained in 82 % of yields by Mannich reaction (Fig. 1B). And it was duly identified by the spectroscopic techniques described above (Figure S1-5, supplementary material).

#### 3.2. Effects of LQFM212 (3) on pharmacological screening of central activity

As described in Table 1, none of the tested doses of LQFM212 (3) was able to decrease the sleep latency, but all doses of LQFM212 (3) increased the sleep duration around 28 % in the sodium pentobarbital-induced test. However, positive control diazepam (18 µmol/kg) decreased the sleep latency and increased sleep duration. The dose of 162 µmol/kg of LQFM212 (3) increased the climbing time by 305 % in chimney test, similarly to positive control diazepam (18 µmol/kg). In OFT, LQFM212 (3) treatment at doses of 18 and 54 µmol/kg increased %CrCe by 27 % and 38 %, and %TcE by 46 % and 69 %, respectively. The dose of 162 µmol/kg decreased the number of squares crossed, grooming frequency and rearings number by 26 %, 22 % and 53 %, respectively, similarly to positive control diazepam (18 µmol/kg) which decreased all parameters observed.

#### 3.3. Anxiolytic-like activity of LQFM212 (3)

LQFM212 (3) treatment modified all behaviors observed in EPM and LDB, as described in Table 2. Concerning parameters in EPM, percentage of open arms entries and percentage of time spent in open arms increased by 29 % and 27 %, respectively, and number of head dips increased by 112 %. Furthermore, decreased percentage of time spent on central platform and number of SAP by 29 % and 80 %, respectively. In LDB, number of transitions, time spent and squares crossed in the light area increased by 32 %, 62 % and 26 %, respectively, and decreased the time of second latency by 33 %. All these results observed with LQFM212 (3) are similar to the positive control

**Table 1**  
Effects of LQFM212 (3) on pharmacological screening of central activity.

	Vehicle 10 mL/kg	LQFM212 18 µmol/kg	LQFM212 54 µmol/kg	LQFM212 162 µmol/kg	Diazepam 18 µmol/kg
<i>Sodium pentobarbital- induce sleep test</i>					
Latency (s)	173.9 ± 12.85	169.8 ± 5.11	188.7 ± 14.92	172.1 ± 8.01	127 ± 8.02**
Duration (min)	39.89 ± 3.38	50.75 ± 3.53*	51.71 ± 3.29*	50.9 ± 3.23*	104.7 ± 11.16***
<i>Chimney test</i>					
Climbing time (s)	4.62 ± 0.56	4.78 ± 0.66	4.25 ± 0.52	18.71 ± 4.11***	30.00 ± 0.0***
<i>Open field test (OFT)</i>					
Squares crossed	70.78 ± 3.09	62.40 ± 5.76	66.13 ± 5.46	52.67 ± 5.78*	52.29 ± 7.65*
Rearings	54.13 ± 2.98	57.33 ± 4.20	53.00 ± 3.20	42.40 ± 3.75*	33.25 ± 3.89***
Grooming	4.29 ± 0.68	3.57 ± 0.37	2.50 ± 0.42*	2.00 ± 0.38*	2.22 ± 0.55*
Immobility (s)	2.33 ± 0.75	5.46 ± 1.33	3.86 ± 1.64	17.16 ± 6.89	46.99 ± 6.86***
%CrCe	37.41 ± 2.19	47.40 ± 1.89*	51.27 ± 2.67***	39.85 ± 2.95	15.88 ± 2.19***
%TCe	21.35 ± 2.08	31.24 ± 3.34*	36.15 ± 3.99**	16.72 ± 2.15	12.37 ± 2.37*

Results are expressed as mean ± SEM. \*p < 0.05, \*\*p < 0.01 and \*\*\*p < 0.001 – compared with the control group (one-way ANOVA followed by Newman-Keuls test). %CrCe, percentage of center crosses in the OFT. %TCe, percentage of center time in the OFT.

diazepam (3.5 µmol/kg).

### 3.4. Mechanism of action involved on anxiolytic-like activity

Percentage of open arms entries and percentage of time spent in open arms in EPM increased by treatment with LQFM212 (3). Flumazenil or mecamlamine pretreatment reverted these parameters by 21 % and 31 % (Fig. 2A-B), and by 22 % and 35 %, (Fig. 4A-B), respectively. Furthermore, flumazenil or mecamlamine, *per se*, did not change the observed parameters. In evaluation of participation of serotonergic pathway,

WAY-100635 pretreatment reverted the anxiolytic-like effect of LQFM212 (3) in 18 % and 24 %, respectively (Fig. 5A-B), whereas, *per se*, WAY-100,635 did not change the observed behavior. Similar to LQFM212 (3) anxiolytic-like effect, DOI treatment in EPM increased the percentage of open arms entries and percentage of time spent in open arms by 68 % and 87 %, respectively. Ketanserin treatment reverted partial the anxiolytic-like effect of DOI. On the other hand, this reversal was not observed with LQFM212 (3) treatment (Fig. 5C-D). Furthermore, *per se*, ketanserin did not alter the observed parameters. PCPA pretreatment also reverted the anxiolytic-like effect of LQFM212 (3) as it decreased the percentage of time in open arms in EPM by 29 % (Fig. 5F), whereas, *per se*, PCPA did not alter the observed parameters.

### 3.5. Effects of LQFM212 (3) on SDA test

LQFM212 (3) did not cause memory impairment on short- and long-term memory (after 90 min and 24 h, respectively; Fig. 3) as demonstrated by an increase in latency to step-down during test session. However, diazepam (18 µmol/kg) decreased latency to step-down on SDA after 90 min and 24 h by 61 % and 63 %, respectively.

**Table 2**  
Anxiolytic-like effect of LQFM212 (3) in EPM and LDB.

	Vehicle 10 mL/kg	LQFM212 54 µmol/kg	Diazepam 3.5 µmol/kg
<i>Elevated plus maze (EPM)</i>			
Open arms entries (%)	47.25 ± 5.14	61.09 ± 2.92*	66.67 ± 3.75*
Time spent in open arms (%)	42.91 ± 4.22	54.56 ± 2.65*	59.37 ± 4.49*
Time spent on central platform (%)	23.39 ± 2.64	6.50 ± 1.35*	14.28 ± 1.90*
Number of SAP	19.00 ± 2.73	3.70 ± 0.76***	1.89 ± 0.71***
Number of head dips	22.86 ± 3.62	48.50 ± 2.85***	62.43 ± 3.75***
<i>Light-dark box (LDB)</i>			
Number of transitions	18.78 ± 1.11	24.89 ± 0.99*	29.57 ± 3.81**
Second latency (s)	14.17 ± 1.12	9.49 ± 1.42*	7.58 ± 1.20**
Squares crossed in light area	26.00 ± 3.12	42.13 ± 3.62*	45.75 ± 4.97**
Time spent in light area (%)	40.43 ± 4.84	50.99 ± 1.97*	56.08 ± 2.35**

Results are expressed as mean ± SEM. \*p < 0.05, \*\*p < 0.01 and \*\*\*p < 0.001 – compared with the vehicle group (one-way ANOVA followed by Newman-Keuls test). SAP, stretched-attend postures.

### 3.6. Computational studies

The tested ligands LQFM032 (1), BHT (2), LQFM212 (3) and the agonist and antagonist controls, diazepam and flumazenil, have showed interaction with residues from α1-D and γ2-C chains in GABA<sub>A</sub> receptor, as defined as BZD-binding site [41,42,46] (Figure S8, supplementary material). Flumazenil has formed hydrophobic interaction with Tyr58, Phe77, Phe100, and Thr207 in addition to H-bonds with Tyr160 at 3.59 Å and Ser205 at 3.82 Å, corroborating previous results [42] (Figure S8C, supplementary material). LQFM032 (1) and BHT (2), could interact with similar residues when compared to the agonist diazepam: Phe100, Phe77, Tyr58, and Ser205 (Figure S8B, D-E, supplementary material). LQFM032 (1) showed a T-shaped π-stacking at 60.57° between the imidazole ring and Tyr58 and a parallel displaced π-stacking at 23.53° between the benzene ring and Tyr210, besides five electrostatic interactions that were H-bond with Ser205 at 3.26 Å, Gln204 at 2.94 Å, and Glu189 at 2.64 Å plus a double salt bridge with this same residue (at 3.67 Å and 5.38). BHT (2) has presented a parallel displaced π-stacking at 17.54° with Phe77 and a H-bond at 3.16 Å with Ser205 and further hydrophobic interactions with Thr142 and Val203. The two compounds combination to obtain LQFM212 (3) has preserved from BHT (2) the hydrophobic interaction between one *tert*-butyl group and Tyr58 and Val203, inherited from LQFM032 (1) interaction of the same type with Phe77 plus the electrostatic interactions with Glu189 (single salt bridge at 3.78 Å and H-bond at 2.61 Å) and H-bond with Gln204 at 2.71 Å, and finally has yield a weaker H-bond with Asn60 at 3.47 Å (Fig. 6 and Figure S8F, supplementary material).

Regarding, docking studies on 5-HT<sub>1A</sub> receptor, both ligands LQFM032 (1), BHT (2), LQFM212 (3) and the controls 8-OH-DPAT and WAY-100635 tested were able to establish interaction in a binding site located in the top of receptor transmembrane helices (TMH) between TMH3, TMH6, TMH7 (Figure S9, supplementary material) on the same direction from previous docking results [47–49]. WAY-100635

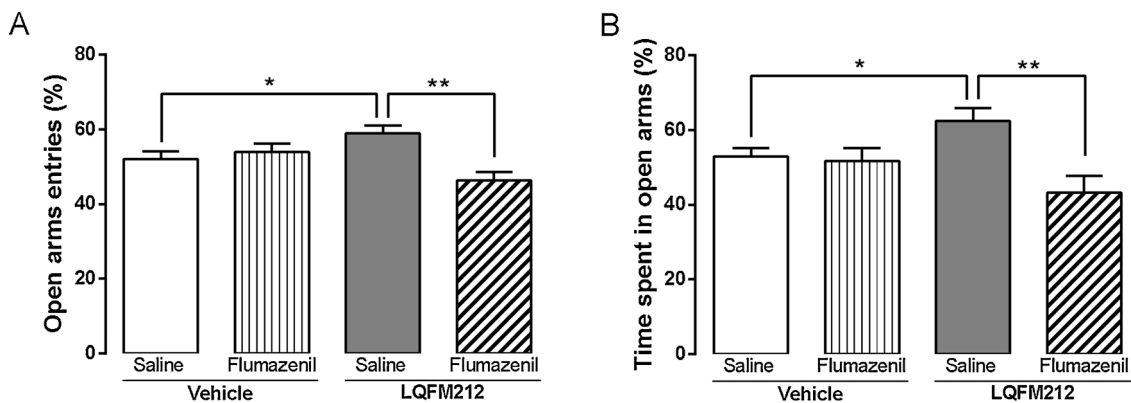


Fig. 2. Effects of pretreatment with flumazenil on the effects of treatments with vehicle or LQFM212 (3) in the EPM. Results are expressed as mean  $\pm$  SEM. \* $p < 0.05$  and \*\* $p < 0.01$  by *t*-test. (A) Percentage of open arms entries. (B) Percentage of time spent in open arms.

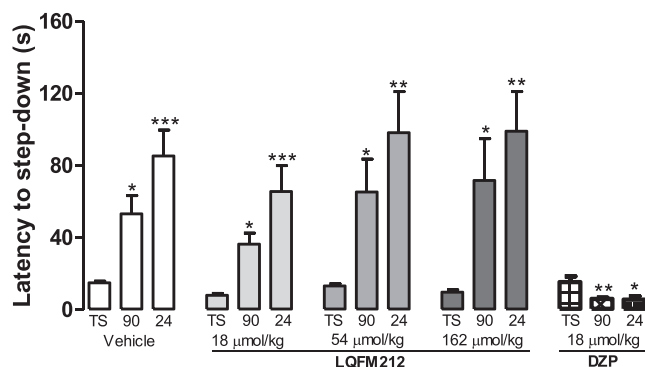


Fig. 3. Effects of LQFM212 (3) on SDA. Results are expressed as mean  $\pm$  SEM. \* $p < 0.05$ , \*\* $p < 0.01$  and \*\*\*  $p < 0.001$  by one-way ANOVA followed Newman-Keuls test. TS, training session.

presented hydrophobic interactions with residues Val117, Phe362, indeed a T-shaped  $\pi$ -stacking at  $84.62^\circ$  between the phenol ring and Phe361, according with previous observations [48]. The known antagonist also showed one H-bond among Asp116 and phenol hydroxyl group at  $2.55 \text{ \AA}$ . When analyzing LQFM032 (1) interactions, PLIP was able to show a salt bridge (at  $5.42 \text{ \AA}$ ) between Asp116 and the piperazine nitrogen, although did not show H-bond formation among the same groups. Even more, it was found a weaker H-bond between Gln97 and hydroxyl group at  $4.07 \text{ \AA}$  (Figure S9C-D, supplementary material). LQFM032 (1) and BHT (2) have shared hydrophobic interaction with the same residues Val117 and Ile113. Moreover, Phe361 also interacted with both ligands, through a T-shaped  $\pi$ -stacking at  $81.34^\circ$  with the LQFM032 (1) phenyl ring and a hydrophobic interaction with

BHT (2), that in turn has displayed also a hydrophobic interaction with Ile189 at extra cellular looping 2 (ECL2) (Figure S9E, supplementary material). The 5-HT<sub>1A</sub> control agonist 8-OH-DPAT have showed H-bond with Asp116 at  $2.58 \text{ \AA}$  and Asn386 at  $3.44 \text{ \AA}$  from hydroxyl group, besides hydrophobic interactions with Ala365, Phe112 and Phe361 (Figure S9B, supplementary material). LQFM212 (3) displayed hydrophobic interactions with Ala383 and Asn386, besides with Phe19 present at ECL1. Further that, also formed H-bond (at  $3.66 \text{ \AA}$ ) with the same residue Asp116, plus a salt bridge (at  $4.67 \text{ \AA}$ ) from the nitrogen atom, closer the hydroxyl group H-bond former, reinforcing the electrostatic interaction (Fig. 6 and Figure S9F, supplementary material).

#### 4. Discussion

Behavioral effects of LQFM212 (3) were examined herein using well-validated animal models to screening CNS-related activities. In sodium pentobarbital-induced sleep test all doses at LQFM212 (3) increased sleep time without altering the latency time, which is indicative of central depressant effect probably related to the synergistic effect with barbiturate [22]. Our findings in chimney test suggest that at highest dose tested of LQFM212 (3) leads to motor impairment. In OFT, LQFM212 (3) showed preference for the center of field which could be indicative of anxiolytic-like effects [23]. However, LQFM212 (3) treatment at highest dose decreased the number of squares crossed, suggesting that LQFM212 (3) in high doses reduces spontaneous movement and corroborate with results detected in chimney test. These data are similar to positive control, diazepam, which has anxiolytic effect at lower doses and produces sedation and myorelaxant effect at higher doses [50]. The intermediate dose of LQFM212 (3) was chosen

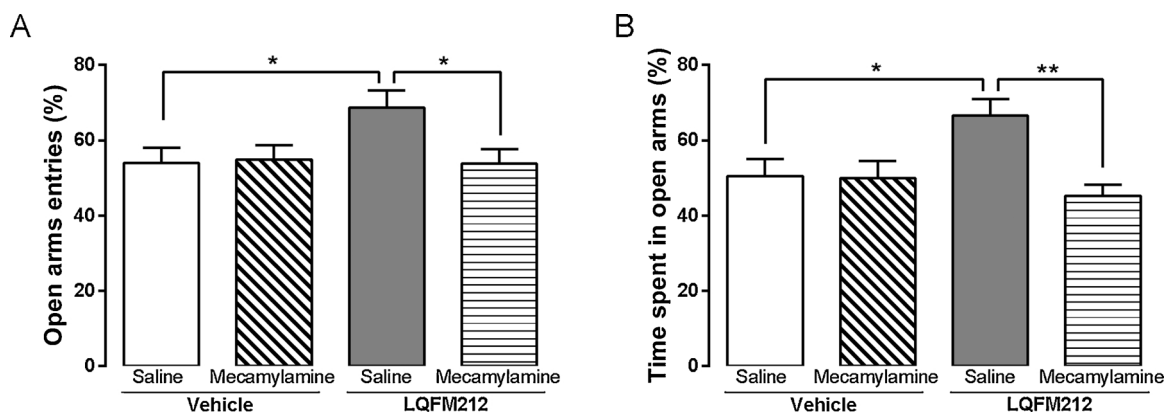


Fig. 4. Effects of pretreatment with mecamlamine on the effects of treatments with vehicle or LQFM212 (3) in the EPM. Results are expressed as mean  $\pm$  SEM. \* $p < 0.05$  and \*\* $p < 0.01$  by *t*-test. (A) Percentage of open arms entries. (B) Percentage of time spent in open arms.

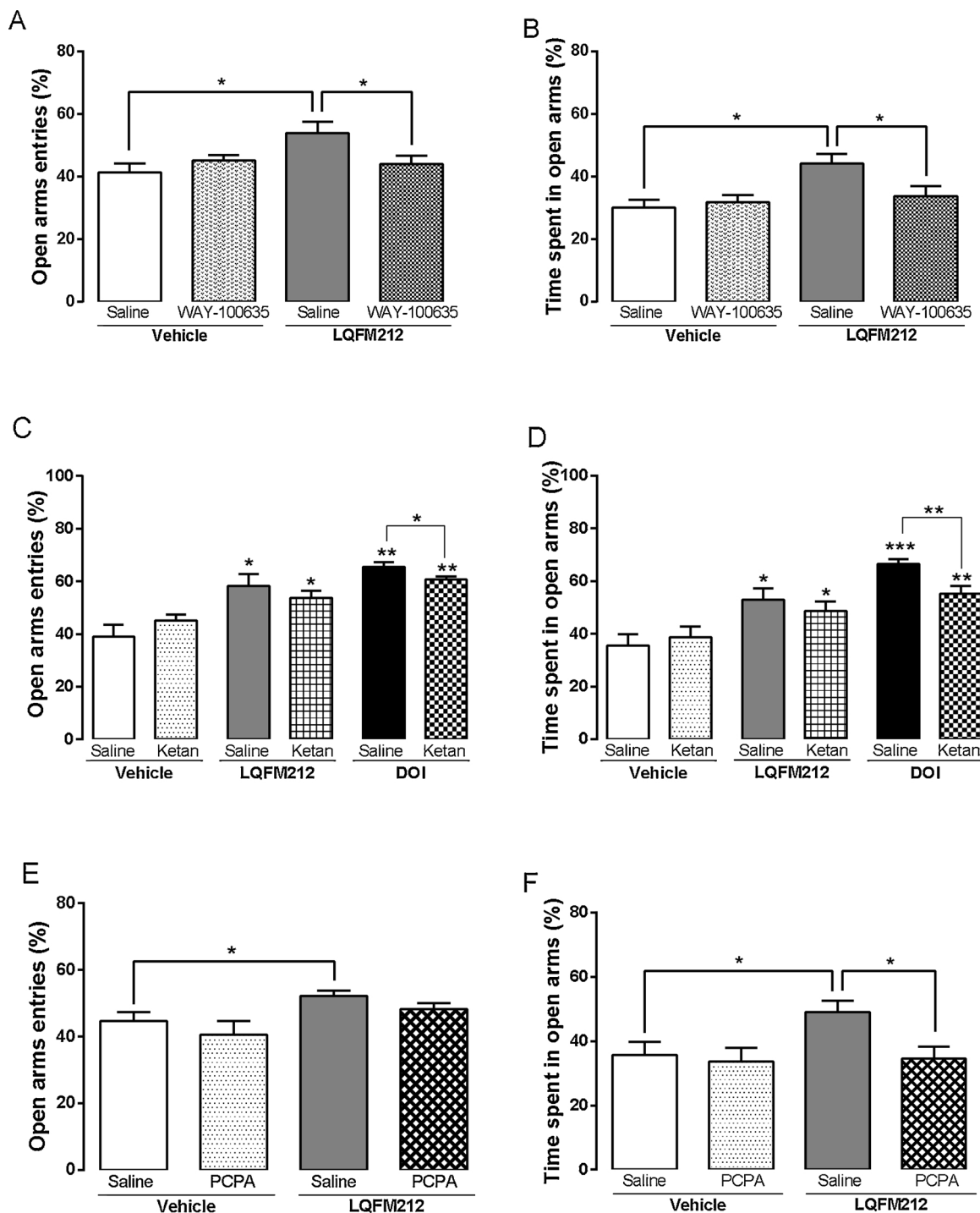


Fig. 5. Effects of pretreatment with WAY-100635, ketanserin or PCPA on the effects of treatments with vehicle or LQFM212 (3) in the EPM. Results are expressed as mean ± SEM. \*p < 0.05 and \*\*\*p < 0.001 by t-test. (A) Percentage of open arms entries. (B) Percentage of time spent in open arms.

for the subsequent tests, because results previous tests suggested anxiolytic-like effect without altering the locomotor activity, which excludes the possibility of false positives in specific tests to evaluation of the anxiolytic-like activity: EPM and LDB. These tests are intensively used in the study of the neurobiological bases of anxiety and in the screening of anxiolytic compounds, and both methodologies are based on the natural behavior of the animals [51]. LQFM212 (3) showed anxiolytic-like effect in both tests, similarly to other piperazine derivatives [13,17].

To elucidate the action mechanisms involved on anxiolytic-like activity of LQFM212 (3), firstly was evaluated the participation of

benzodiazepine site present in the GABA<sub>A</sub> receptor complex. Benzodiazepine agonists act as positive modulators by opening the chloride channel, resulting in increased chloride influx and hyperpolarization of neuronal membrane, with consequently anxiolytic, anticonvulsive, myorelaxant and/or hypno-sedative effects [52–54]. Our data showed that the effects of LQFM212 (3) were abolished with flumazenil pretreatment, suggesting benzodiazepine site of GABA<sub>A</sub> receptor can contribute to anxiolytic-like effect observed with LQFM212 (3), similarly the effects found with LQFM032 (1) [17], one lead compound used to design LQFM212 (3). When docked to GABA<sub>A</sub> receptor, LQFM212 (3) was able to preserve some relevant interactions



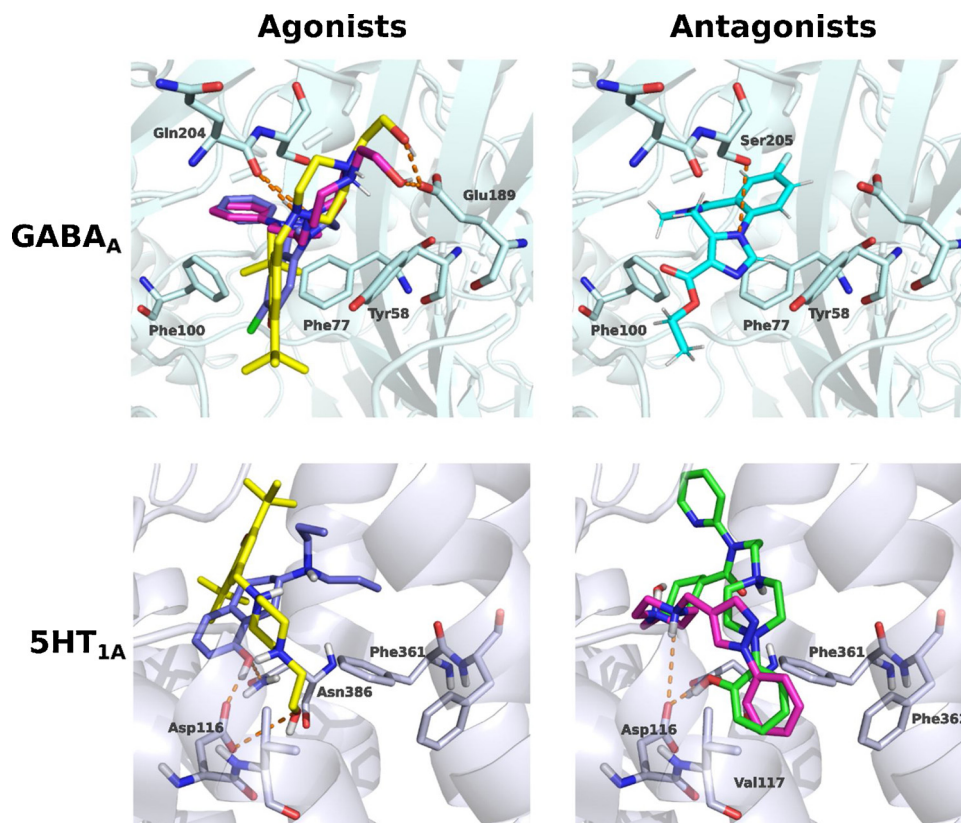


Fig. 6. Ligands binding poses docked in GABA<sub>A</sub> and 5-HT<sub>1A</sub> receptors. Orange dashed lines are H-bond. Ligands are LQFM032 (1) (magenta), LQFM212 (3) (yellow), WAY-100635 (green), flumazenil (cyan) and agonist controls (purple).

from the precursors LQFM032 (1) and BHT (2). Such interactions reflect an interesting similarity to those observed between GABA<sub>A</sub> receptor and the diazepam agonist and LQFM032 (1), that were Tyr58, Phe77, Glu189, and Gln204. Among the GABA<sub>A</sub> residues implicated on the LQFM212 (3) interaction rescued from docking, Tyr58, Phe77, and Asn60 have already shown to be involved on interaction with diazepam, that indicate also the residues Phe100, Ser205 and Tyr210 [31] matching with our agonist control analyses. Regarding flumazenil action to abolish the LQFM032 (1) and LQFM212 (3) effects, the early interaction with the same residues implicated on agonists interaction, highlighting Tyr58, Phe77, and Phe100, seems to have occupied the BZD-binding site and blocked the later ligands tested access. Our results corroborate previous data from electron microscopy structures that clarified how the competitive antagonist flumazenil can to block the agonist effect on GABA<sub>A</sub> receptor interacting with the same residues, outlined by aromatic residues to nest flumazenil, including Tyr58, Phe77, Phe100, and Tyr160 [42].

BZDs are often associated with undesirable anterograde amnesia stem primarily from an impair acquisition processes of information [55,56]. Regarding cognition impairment caused by BZDs was evaluated if LQFM212 (3) treatment damage mice performance on SDA. Inhibitory avoidance involves learning to inhibit a response in order to avoid an aversive stimulus, e.g. foot-shock during the training session on SDA. In this regard, is possible evaluate a good memory acquisition shown by a higher latency to descend from a platform to a previously electrified grid, and treatment that reduces this latency can be considered amnesic [57]. LQFM212 (3) treatment did not alter cognitive function and vehicle group showed the same effect. However, treatment with diazepam reduced latency to step-down on SDA at both assessments in the test session. This result is consistent with the adverse effect on memory of classic BZDs described above.

GABA<sub>A</sub> receptor is pentameric ligand-gated ion channel formed by various combinations of different subunits which bestow certain

physiological characteristics as well as pharmacological properties. For instance,  $\alpha 1$  subunit is related mainly to sleep and sedating effects,  $\alpha 2$  subunit with anxiolytic effects,  $\alpha 3$  with myorelaxation and  $\alpha 5$  with memory and learning [52]. Benzodiazepine site involvement in the anxiolytic-like effect of LQFM212 (3) was well characterized by flumazenil antagonism and docking assays. However, none of the tested doses of LQFM212 (3) affected cognition on SDA, as occurred with diazepam treatment. It is known that  $\alpha 5$ -containing GABA<sub>A</sub> receptors is one of the minimal requirements for diazepam action (for review see [52]) which might explain its amnesic effects. Experimental findings related  $\alpha 5$  knockout mice, compared to the wild-type animals, show enhanced platform learning in the Morris water maze [58] while an inverse agonist selective for  $\alpha 5$ -containing GABA<sub>A</sub> receptors facilitated the acquisition in a similar protocol of working memory [59]. In this regard, it is plausible to suggest that perhaps LQFM212 (3) has greater selectivity of action on GABA<sub>A</sub> receptor subunits than diazepam. Zolpidem as well as another nonbenzodiazepine hypnotics, zaleplon and eszopiclone, has action on  $\alpha 1$ -containing GABA<sub>A</sub> receptors and are known to be useful in the treatment of insomnia and may offer a better safety profile over BZD. Moreover, eszopiclone is also a piperazine derivative predominantly selective for the  $\alpha 2$  and  $\alpha 3$  subtypes with less sensitivity for  $\alpha 1$  subtype and, consequently, has anxiolytic effects in addition to hypnotics [60].

Alterations in cholinergic signaling might be involved in mood regulation and learning [for review see [10]]. Nicotine has been reported enhance passive avoidance learning [61] and has anxiolytic-like effects [62] in animal models. Role of intrinsic cholinergic tone in anxiety can be tested by administering the mecamylamine nicotinic receptor antagonist. Our results suggest that mecamylamine pretreatment reverses effects of LQFM212 (3) in EPM. Therefore, nicotinic receptors would be relevant to LQFM212 (3) effects. These results, as well as those found on SDA, are similar to the results shown by LQFM032 (1) [17]. Brito et al. suggest that the participation of the cholinergic

pathway through nicotinic receptors contributes to anxiolytic-like activity and non-impairment of cognition through a co-transmission involved nicotinic pathway and modulated benzodiazepine site [17]. ABT 418 is a potent agonist at  $\alpha 4/\beta 2$ -containing nAChRs which enhances retention of the inhibitory avoidance response in mice [63] and has anxiolytic-like effects in EPM [64]. LQFM212 (3) and LQFM032 (1) are structurally similar to the JWB1-84-1 choline analog [65] and can be explaining the agonist activity at nicotinic receptors of both piperazine derivatives.

Structural studies relating to the pharmacology of nAChRs is a challenge given the variety of subtypes and heterogeneity of subunits. However, 3D-QSAR pharmacophore models display two essential pharmacophoric elements to interaction with nAChRs: a cationic center and a hydrogen bond acceptor [66]. B-973, a piperazine derivative, acts as an agonist and positive allosteric modulator (ago-PAM) of  $\alpha 7$ -containing nAChRs. This subtype of nAChR which implicates in the beneficial effects of nicotine on attention and has become a target for improving cognitive deficits [67,68]. Allosteric activation of nAChRs represents an attractive alternative to orthosteric agonists and its limiting efficacy as tachyphylaxis, upregulation of receptor, and possible cellular toxicity associated with chronic activation [67]. Several piperazine derivatives have been investigated in this line [69–71], and further studies to determine the action of LQFM212 on nAChRs have great importance. In general, nAChRs are widely expressed both presynaptically and postsynaptically and it has been suggested that the most important role in CNS is the modulation of neurotransmitter release modulated by the nAChRs expressed in cholinergic, glutamatergic, and GABAergic terminals. Nicotinic receptor activity directly and indirectly initiates an intracellular  $Ca^{2+}$  signal that contributes to neurotransmitter release [72]. Nicotine has also been shown to inhibit the firing of serotonergic neurons in the dorsal raphe which project to many areas of the CNS [73].

Further investigation the involvement of 5-HT in the mechanism of action of anxiolytic-like effect of LQFM212 (3) we evaluated direct participation some serotonergic receptors, 5-HT<sub>1A</sub> and 5-HT<sub>2A/2C</sub>, and the contribution of endogenous serotonin. 5-HT<sub>1A</sub> receptor plays an important role in anxiety and partial agonists of this receptor such as buspirone are used to treat this condition [74]. This receptor is linked to the inhibition of adenylate cyclase and mainly function as inhibitory presynaptic mainly in serotonin neurons of the midbrain raphe nuclei and as postsynaptic receptors mainly in the hippocampus, septum, amygdala and corticolimbic areas [75]. Pretreatment with WAY-100635 reversed the effect of LQFM212 (3), suggest that anxiolytic-like effect of LQFM212 (3) can be mediated by activation of 5-HT<sub>1A</sub> receptor, differently than occurs with LQFM032 (1) [17].

To support this evidence, we performed molecular docking of 5-HT<sub>1A</sub> receptor and the ligands WAY-100635, LQFM032 (1) and LQFM212 (3). All G-protein-coupled monoaminergic receptors contain a highly conserved Asp residue at similar positions in TMH3, corresponding to Asp116 in the 5-HT<sub>1A</sub> receptor [47]. The presence of both H-bond and salt bridge between Asp116 and LQFM212 (3) piperazine nitrogen, similar with 8-OH-DPAT agonist, can contribute to stabilize the receptor-ligand complexation, although in accordance with results *in vivo* by WAY-100635 pretreatment. Our results suggest that WAY-100635 and LQFM032 (1) are both antagonists of 5-HT<sub>1A</sub> and have interacted with similar residues, specially the residues Phe361 and Phe362 corroborate with previous results that have indicated that phenylalanine residues in TMH6 are crucial for ligand binding [76]. Furthermore, both established hydrophobic interaction with Val117 indeed LQFM032 (1) has presented two electrostatic interactions, while WAY-100635 has presented only one. The docking assay helped us to understand the difference observed in results by LQFM212 (3) and LQFM032 (1) for 5HT<sub>1A</sub> receptor, as the anxiolytic-like effect of LQFM032 (1) was not reversed by NAN-190 pretreatment [17].

5-HT<sub>2</sub> receptors has been implicated as playing a role in anxiety [8,77,78], this receptor family is preferentially coupled to G<sub>q/11</sub> to

increase inositol phosphates and cytosolic  $[Ca^{2+}]$ . They are expressed mainly in cortical areas and can also be expressed in GABAergic and glutamatergic interneurons, besides involved in tonic modulation of dopaminergic activity [77]. Literature data suggest that modifications of neurotransmission mediated by 5-HT<sub>2A</sub> agonists may assist anxiolytic effects of SSRIs and benzodiazepines [79,80]. Here, we demonstrated that pretreatment with ketanserin was not able to revert the anxiolytic-like effect observed with LQFM212 (3). On the other hand, the effect of DOI was reverted partial by pretreatment with ketanserin, demonstrating that the anxiolytic-like effect of DOI depends, at least in part, of 5-HT<sub>2A/2C</sub> receptor.

Serotonin depletion by inhibiting its synthesis is a tool that can be used to investigate the functionality of the serotonergic system. Pretreatment with PCPA leads to a slow and cumulative depletion of 5-HT, which can range from 60 to 80 % reduction in endogenous serotonin levels [81,82]. Pretreatment with PCPA reversed the anxiolytic-like effect of LQFM212 (3), characterizing the involvement of serotonergic pathway and 5-HT bioavailability contributes to the action mechanism of LQFM212 (3).

In summary, the findings in present paper suggest that piperazine derivative, LQFM212, has anxiolytic-like activity with putative action mechanisms, which involves the benzodiazepine site of GABA<sub>A</sub> receptor, nicotinic receptors, and serotonergic pathways, without to alter cognitive function. Molecular docking of LQFM212 showed that the ligand has more interactions with GABA<sub>A</sub> receptor than with 5-HT<sub>1A</sub> receptor. In this sense, this piperazine derivative is a good prototype for treating anxiety disorders.

#### Authors statement

Lorrane Kelle da Silva Moreira conceived and designed research, conducted experiments, analyzed data and wrote the manuscript.

Adriane Ferreira de Brito conducted experiments and analyzed data.

Crisciele Fontana conducted experiments, analyzed data and wrote the manuscript.

Flávio Souza de Carvalho, Germán Sanz, Boniek Gontijo Vaz and Luciano Moraes Lião conducted experiments.

Fábio Fagundes da Rocha analyzed data and wrote the manuscript.

Hugo Verli, Ricardo Menegatti and Elson Alves Costa conceived and designed research, analyzed data and wrote the manuscript. All authors read and approved the manuscript.

#### Declarations of Competing Interest

None.

#### references

[12].

#### Acknowledgments

The authors are grateful to CAPES and CNPq for financial support.

#### References

- [1] D.L. Hoffman, E.M. Dukes, H.U. Wittchen, Human and economic burden of generalized anxiety disorder, *Depress. Anxiety* 25 (1) (2008) 72–90.
- [2] J.M. Kent, S.J. Mathew, J.M. Gorman, Molecular targets in the treatment of anxiety, *Biol. Psychiatry* 52 (2002) 1008–1030.
- [3] M. Bourin, Animal models for screening anxiolytic-like drugs: a perspective, *Dialogues Clin. Neurosci.* 17 (2015) 295–303.
- [4] A.M. Depino, T. Tsetsenis, C. Gross, GABA homeostasis contributes to the developmental programming of anxiety-related behavior, *Brain Res.* 1210 (2008) 189–199.
- [5] H. Möhler, The GABA system in anxiety and depression and its therapeutic potential, *Neuropharmacology* 62 (2012) 42–53.
- [6] H. Möhler, Physiology and pharmacology of the GABA system: focus on GABA

- receptors, in: J. Monti, S. Pandi-Perumal, H. Möhler (Eds.), *GABA and Sleep*, Springer, Basel, 2010, pp. 3–24.
- [7] J.R. Strawn, L. Geraciotti, N. Rajdev, K. Clemenza, A. Levine, Pharmacotherapy for generalized anxiety disorder in adult and pediatric patients: an evidence-based treatment review, *Expert Opin. Pharmacother.* 19 (10) (2018) 1057–1070.
- [8] B. Goitia, M.C. Rivero-Echeto, N.V. Weisstaub, J.A. Gingrich, E. Garcia-Rill, V. Bisagno, F.J. Urbano, Modulation of GABA release from the thalamic reticular nucleus by cocaine and caffeine: role of serotonin receptors, *J. Neurochem.* 136 (2016) 526–535.
- [9] J. Bockaert, C. Bécamel, Contrôle de l'anxiété par les récepteurs de la sérotonine 5-HT<sub>2c</sub> de la strie terminale, *Med. Sci. (Paris)* 33 (2017) 87–89.
- [10] M.R. Picciotto, D.H. Brunzell, B.J. Caldarone, Effect of nicotine and nicotinic receptors on anxiety and depression, *Neuroreport* 13 (9) (2002) 1097–1106.
- [11] M.R. Picciotto, A.S. Lewis, G.I. van Schalkwyk, Y.S. Mineur, Mood and anxiety regulation by nicotinic acetylcholine receptors: a potential pathway to modulate aggression and related behavioral states, *Neuropharmacology* 96 (2015) 235–243.
- [12] A.F. Brito, L.K.S. Moreira, R. Menegatti, E.A. Costa, Piperazine derivatives with central pharmacological activity used as therapeutic tools, *Fundam. Clin. Pharmacol.* 33 (2019) 13–24.
- [13] A.F. Brito, J.L.R. Martins, J.O. Fajemiroye, et al., Central pharmacological activity of a new piperazine derivative: 4-(1-phenyl-1H-pyrazol-4-ylmethyl)-piperazine-1-carboxylic acid ethyl ester, *Life Sci.* 90 (2012) 910–916.
- [14] P.M. Galdino, D.R. Oliveira, L.F. Florentino, et al., Involvement of the monoamine system in antidepressant-like properties of 4-(1-phenyl-1H-pyrazol-4-ylmethyl)-piperazine-1-carboxylic acid ethyl ester, *Life Sci.* 143 (2015) 187–193.
- [15] G. Neves, C.B. Antonio, A.H. Betti, M.A. et al., New insights into pharmacological profile of LASSBio-579, a multi-target N-phenylpiperazine derivative active on animal models of schizophrenia, *Behavioural Brain Res* 237 (2013) 86–95.
- [16] D. Siskind, L. McCartney, R. Goldschlager, S. Kisely, Clozapine v. First- and second-generation antipsychotics in treatment-refractory schizophrenia: systematic review and meta-analysis, *Br. J. Psychiatry* 209 (2016) 385–392.
- [17] A.F. Brito, J.O. Fajemiroye, H.F.S. Neri, et al., Anxiolytic-like effect of 2-((1-phenyl-1H-pyrazol-4-yl)methyl)piperazin-1-yl)ethan-1-ol is mediated through the benzodiazepine and nicotinic pathways, *Chem. Biol. Drug Des.* 90 (2017) 432–442.
- [18] A.M. Pisoschi, A. Pop, The role of antioxidants in the chemistry of oxidative stress: a review, *Eur. J. Med. Chem.* 97 (2015) 55–74.
- [19] W.A. Yehye, N.A. Rahman, A. Ariffin, S.B.A. Hamid, A.A. Alhadi, F.A. Kadir, M. Yaeghoobi, Understanding the chemistry behind the antioxidant activities of butylated hydroxytoluene (BHT): a review, *Eur. J. Med. Chem.* 101 (2015) 295–312.
- [20] E.L. Eliel, M.T. Fisk, 5-Methylfurfuryldimethylamine, *Org Synth Coll* 35 (1955) 78.
- [21] D.M. da Silva, G. Sández, B.G. Vaz, et al., Tert-butyl 4-((1-phenyl-1H-pyrazol-4-yl)methyl)piperazine-1-carboxylate (LQFM104) – new piperazine derivative with antianxiety and antidepressant-like effects: putative role of serotonin system, *Biomed. Pharmacother.* 103 (2018) 546–552.
- [22] E.A. Carlini, V. Burgos, Screening farmacológico de ansiolíticos: metodologia laboratorial e comparação entre diazepam e clorbenzepam, *Rev Assoc Bras Psiquiatr* 1 (1979) 25–31.
- [23] A.J.S.C. Lapa, M.T.R. Lima-Landman, M.S.A. Castro, Caracterização farmacológica da atividade farmacológica no sistema nervoso central, *Campinas, Sociedade Brasileira de Plantas Mediciniais* (2008).
- [24] J. Archer, Tests for emotionality in rats and mice: a review, *Anim. Behav.* 21 (1973) 205–235.
- [25] M. Komada, K. Takao, T. Miyakawa, Elevated plus maze for mice, *J. Vis. Exp.* 22 (2008) e1088.
- [26] R.G. Lister, The use of a plus maze to measure anxiety in the mouse, *Psychopharmacology* 92 (1987) 180–185.
- [27] S. Pellow, P. Chopin, S. File, M. Briley, Validation of open-closed arm entries in an elevated plus-maze as a measure of anxiety in the rat, *J. Neurosci. Methods* 14 (1985) 149–167.
- [28] B. Costall, B.J. Jones, M.E. Kelly, R.J. Naylor, D.M. Tomkins, Exploration of mice in a black and white box: validation as a model of anxiety, *Pharmacol. Biochem. Behav.* 32 (1989) 777–785.
- [29] S.M. Miller, C.C. Piasecki, J.S. Lonstein, Use of the light-dark box to compare the anxiety-related behavior of virgin and postpartum female rats, *Pharmacol. Biochem. Behav.* 100 (2011) 130–137.
- [30] R. Roesler, R. Walz, J. Quevedo, F. de-Paris, S. Zanata, E. Graner, et al., Normal inhibitory avoidance learning and anxiety, but increased locomotor activity in mice devoid of PrP(C), *Mol Brain Res* 71 (2) (1999) 349–353.
- [31] S. Masiulis, R. Desai, T. Uchański, I. Serna Martin, D. Laverty, et al., GABA<sub>A</sub> receptor signalling mechanisms revealed by structural pharmacology, *Nature* 565 (2019) 454–459.
- [32] J. Zhang, J. Yang, R. Jang, Y. Zhang, Correspondence, GPCR-I TASSER: a hybrid approach to G protein-coupled receptor structure modeling and the application to the human genome, *STFODE* 23 (2015) 1538–1549.
- [33] M.H. Olsson, C.R. Søndergaard, M. Rostkowski, J.H. Jensen, PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions, *J. Chem. Theory Comput.* 7 (2) (2011) 525–537.
- [34] C.R. Søndergaard, M.H. Olsson, M. Rostkowski, J.H. Jensen, Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values, *J. Chem. Theory Comput.* 7 (7) (2011) 2284–2295.
- [35] M.D. Hanwell, D.E. Curtis, D.C. Lonie, T. Vandermeersch, E. Zurek, G.R. Hutchison, Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, *J. Cheminformatics* 4 (2012) 17.
- [36] F.L. Custódio, H.J. Barbosa, L.E. Dardenne, A multiple minima genetic algorithm for protein structure prediction, *Appl. Soft Comput.* 15 (2014) 88–99.
- [37] C.S. Magalhães, D.M. Almeida, H.J.C. Barbosa, L.E. Dardenne, A dynamic niching genetic algorithm strategy for docking highly flexible ligands, *Inf. Sci. (Ny)* 289 (2014) 206–224.
- [38] W. Yin, X. Zhou, D. Yang, P. de Waal, M. Wang, et al., A common antagonistic mechanism for class A GPCRs revealed by the structure of the human 5-HT<sub>1B</sub> serotonin receptor bound to an antagonist, *Cell Discov.* (2018).
- [39] J. Garcia-Nafria, R. Nehme, P. Edwards, C. Tate, Cryo-EM structure of the serotonin 5-HT<sub>1B</sub> receptor coupled to heterotrimeric G<sub>β</sub>, *Nature* 558 (2018) 620–623.
- [40] Y.D. Paila, S. Tiwari, D. Sengupta, A. Chattopadhyay, Molecular modeling of the human serotonin<sub>1A</sub> receptor: role of membrane cholesterol in ligand binding of the receptor, *Mol. Biosyst.* 7 (2011) 224–234.
- [41] E. Sigel, A. Buhr, The benzodiazepine binding site of GABA<sub>A</sub> receptors, *Trends Pharmacol. Sci.* 18 (1997) 425–429.
- [42] S. Zhu, C.M. Novello, J. Teng, R.M. Walsh, J.J. Kim, R.E. Hibbs, Structure of a human synaptic GABA<sub>A</sub> receptor, *Nature* 559 (2018) 67–72.
- [43] S. Salentin, S. Schreiber, V.J. Haupt, M.F. Adasme, M. Schroeder, PLIP: fully automated protein-ligand interaction profiler, *Nucleic Acids Res.* 43 (2015) 443–447.
- [44] K. Stierand, P.C. Maass, M. Rarey, Molecular complexes at a glance: automated generation of two-dimensional complex diagrams, *Bioinformatics* 22 (2006) 1710–1716.
- [45] P.C. Fricker, M. Gastreich, M. Rarey, Automated drawing of structural molecular formulas under constraints, *J. Chem. Inf. Comput. Sci.* 44 (2004) 1065–1078.
- [46] P.J. Whiting, R.M. McKernan, K.A. Wafford, Structure and pharmacology of vertebrate GABA<sub>A</sub> receptor subtypes, *Int. Rev. Neurobiol.* 38 (1995) 95–138.
- [47] J. Dabrowska, M. Brylinski, Stereoselectivity of 8-OH-DPAT toward the serotonin 5-HT<sub>1A</sub> receptor: Biochemical and molecular modeling study, *Biochem. Pharmacol.* 72 (2006) 498–511.
- [48] S. Dilly, J.F. Liegeois, Structural insights into 5-HT<sub>1A</sub>/D<sub>4</sub> selectivity of WAY-100635 analogues: molecular modeling, synthesis, and in vitro binding, *J. Chem. Inf. Model.* 56 (2016) 1324–1331.
- [49] D. Warszycki, M. Rueda, S. Mordalski, K. Kristiansen, G. Sata la, et al., From homology models to a set of predictive binding pockets 5-HT<sub>1A</sub> receptor case study, *J. Chem. Inf. Model.* 57 (2017) 311–321.
- [50] M.L. Novas, C. Wolfman, J.H. Medina, E. de Robertis, Proconvulsant and anxiogenic effects of n-butyl-β-carboline-3-carboxylate on endogenous benzodiazepine binding inhibitor from brain, *Pharmacol. Biochem. Behav.* 30 (1988) 331–336.
- [51] A. Ennaceur, P.L. Chazot, Preclinical animal anxiety research - flaws and prejudices, *Pharmacol. Res. Perspect.* 4 (2016) e00223.
- [52] E.R. Korpi, G. Grunder, H. Luddens, Drug interactions at GABA(A) receptors, *Prog. Neurobiol.* 67 (2002) 113–159.
- [53] J.J. Seo, S.H. Lee, Y.S. Lee, et al., Anxiolytic-like effects of obovatol isolated from *Magnolia obovata*: involvement of GABA/benzodiazepine receptors complex, *Prog. Neuropsychopharmacol. Biol. Psychiatry* 31 (2007) 1363–1369.
- [54] S. Taliani, B. Cosimelli, F. da Settimo, et al., Identification of anxiolytic/nonsedative agents among Indol-3-ylglyoxylamides acting as functionally selective agonists at the γ-aminobutyric acid-α (GABA<sub>A</sub>) α2 benzodiazepine receptor, *J. Med. Chem.* 62 (2009) 3723–3734.
- [55] D. Beracochea, Anterograde and retrograde effects of benzodiazepines on memory, *Transfus. Apher. Sci.* 6 (2006) 1460–1465.
- [56] L.P. Longo, B. Johnson, Addiction: part I. Benzodiazepines - side effects, abuse risk and alternatives, *Am. Fam. Physician* 61 (7) (2000) 2121–2128.
- [57] J. Quillfeldt, Behavioral methods to study learning and memory in rats, in: M. Andersen, S. Tufik (Eds.), *Rodent Model as Tools in Ethical Biomedical Research*, 2015.
- [58] N. Collinson, F.M. Kuenzi, W. Jarolimek, K.A. Maubach, et al., Enhanced learning and memory and altered GABAergic synaptic transmission in mice lacking the alpha 5 subunit of the GABA<sub>A</sub> receptor, *J. Neurosci.* 22 (13) (2002) 5572–5580.
- [59] N. Collinson, J.R. Atack, P. Laughton, G.R. Dawson, D.N. Stephens, An inverse agonist selective for alpha5 subunit-containing GABA<sub>A</sub> receptors improves encoding and recall but not consolidation in the Morris water maze, *Psychopharmacology (Berl.)* 188 (4) (2006) 619–628.
- [60] G.M. Anis, M. Thomas, M.A. Henderson, Pharmacotherapy treatment options for insomnia: a primer for clinicians, *Int. J. Mol. Sci.* 17 (1) (2016) 50.
- [61] G. Üzümlü, A.S. Dfler, N. Bahçekapili, M. Tasyürekli, Y.Z. Ziyilhan, Nicotine improves learning and memory in rats: morphological evidence for acetylcholine involvement, *Int. J. Neurosci.* 114 (9) (2004) 1163–1179.
- [62] A. Cohen, R.W. Young, M.A. Velazquez, et al., Anxiolytic effects of nicotine in a rodent test of approach-avoidance conflict, *Psychopharmacology* 204 (3) (2009) 541–549.
- [63] M.V. Decker, P. Curzon, J.D. Brioni, S.P. Arneric, Effects of ABT-418, a novel cholinergic channel ligand, on place learning in septal-lesioned rats, *Eur. J. Pharmacol.* 261 (1–2) (1994) 217–222.
- [64] S.P. Arneric, J.P. Sullivan, C.A. Briggs, D. Donnelly-Roberts, D.J. Anderson, et al., (S)-3-methyl-5-(1-methyl-2-pyrrolidinyl) isoxazole (ABT 418): a novel cholinergic ligand with cognition-enhancing and anxiolytic activities: I. In vitro characterization, *J. Pharmacol. Exp. Ther.* 270 (1) (1994) 310–318.
- [65] A. Sood, J. Warrenbeach, S. Webster, A. Terry, J. Buccafusco, The effects of JW184-1 on memory-related task performance by amyloid Aβ transgenic mice and by young and aged monkeys, *Neuropharmacology* 53 (5) (2007) 588–600.
- [66] I. Tomassoli, D. Gündisch, The twin drug approach for novel nicotinic acetylcholine receptor ligands, *Bioorg. Med. Chem.* 23 (15) (2015) 4375–4389.
- [67] D.J. Post-Munson, R.L. Pieschl, T.F. Molski, et al., B-973, a novel piperazine positive allosteric modulator of the α7 nicotinic acetylcholine receptor, *Eur. J. Pharmacol.* 799 (2017) 16–25.
- [68] J.W. Young, K. Finlayson, C. Spratt, et al., Nicotine improves sustained attention in mice: evidence for involvement of the alpha7 nicotinic acetylcholine receptor,

- Neuropsychopharmacology 29 (2004) 891–900.
- [69] M. Quadri, R.L. Papke, N.A. Horenstein, Dissection of N,N-diethyl-N'-phenylpiperazines as  $\alpha 7$  nicotinic receptor silent agonists, *Bioorg. Med. Chem.* 24 (2) (2016) 286–293.
- [70] D. Purohit, V. Saini, S. Kumar, et al., 3D-QSAR and docking study of 2-(pyridin-3-yl)oxy)methyl piperazines as  $\alpha 7$  nicotinic acetylcholine receptor modulators for the treatment of inflammatory disorders, *Mini Rev. Med. Chem.* (2019).
- [71] A. Gulsevin, R.L. Papke, C. Stokes, et al., Allosteric Agonism of  $\alpha 7$  Nicotinic Acetylcholine Receptors: Receptor Modulation Outside the Orthosteric Site, *Mol. Pharmacol.* 95 (6) (2019) 606–614.
- [72] M. Zoli, S. Pucci, A. Vilella, C. Gotti, Neuronal and extraneuronal nicotinic acetylcholine receptors, *Current Neuropharmacology* 16 (2018) 338–349.
- [73] G. Engberg, S. Erhardt, T. Sharp, et al., Nicotine inhibits firing activity of dorsal raphe 5-HT neurones *in vivo*, *Naunyn-Schmiedeberg Arch Pharmacol* 362 (2000) 41–45.
- [74] M.D. Soon-Tae Lee, B.S. Jong-Ha Park, M.D. Manho Kim, Efficacy of the 5-HT<sub>1A</sub> agonist, buspirone hydrochloride, in migraineurs with anxiety: a randomized, prospective, parallel group, double-blind, placebo-controlled study, *Headache* 45 (2005) 1004–1011.
- [75] F. Artigas, Serotonin receptors involved in antidepressant effects, *Pharmacol. Ther.* 137 (2013) 119–131.
- [76] W. Cho, L.P. Taylor, A. Mansour, H. Akil, Hydrophobic residues of the D2 dopamine receptor are important for binding and signal transduction, *J. Neurochem.* 65 (2002) 2105–2115.
- [77] E.L. Gibson, A.M.C. Barnfield, G. Curzon, Evidence that mCPP-induced anxiety in the plus-maze is mediated by postsynaptic 5-HT<sub>2C</sub> receptors but not by sympathomimetic effects, *Neuropharmacology* 33 (1994) 457–465.
- [78] B.A. Nic Dhonnchadha, M. Bourin, M. Hascoet, Anxiolytic-like effects of 5-HT<sub>2</sub> ligands on three mouse models of anxiety, *Behav. Brain Res.* 140 (2003) 203–214.
- [79] B.A. Nic Dhonnchadha, N. Ripoll, F. Clenet, M. Hascoet, M. Bourin, Implication of 5-HT<sub>2</sub> receptor subtypes in the mechanism of action of antidepressants in the four plates test, *Psychopharmacology* 179 (2005) 418–429.
- [80] F. Massé, B.A. Nic Dhonnchadha, M. Hascoët, M. Bourin, Anxiolytic-like effect of 5-HT(2) ligands and benzodiazepines co-administration: comparison of two animal models of anxiety (the four-plate test and the elevated plus maze), *Behav. Brain Res.* 177 (2007) 214–226.
- [81] J.P. Redrobe, M. Bourin, M.C. Colombel, G.B. Baker, Dose-dependent noradrenergic and serotonergic properties of venlafaxine in animal models indicative of antidepressant activity, *Psychopharmacology (Berl.)* 138 (1998) 1–8.
- [82] O.F. O'Leary, A.J. Bechtholt, J.J. Crowley, T.E. Hill, M.E. Page, I. Lucki, Depletion of serotonin and catecholamines block the acute behavioral response to different classes of antidepressant drugs in the mouse tail suspension test, *Psychopharmacology* 192 (2007).

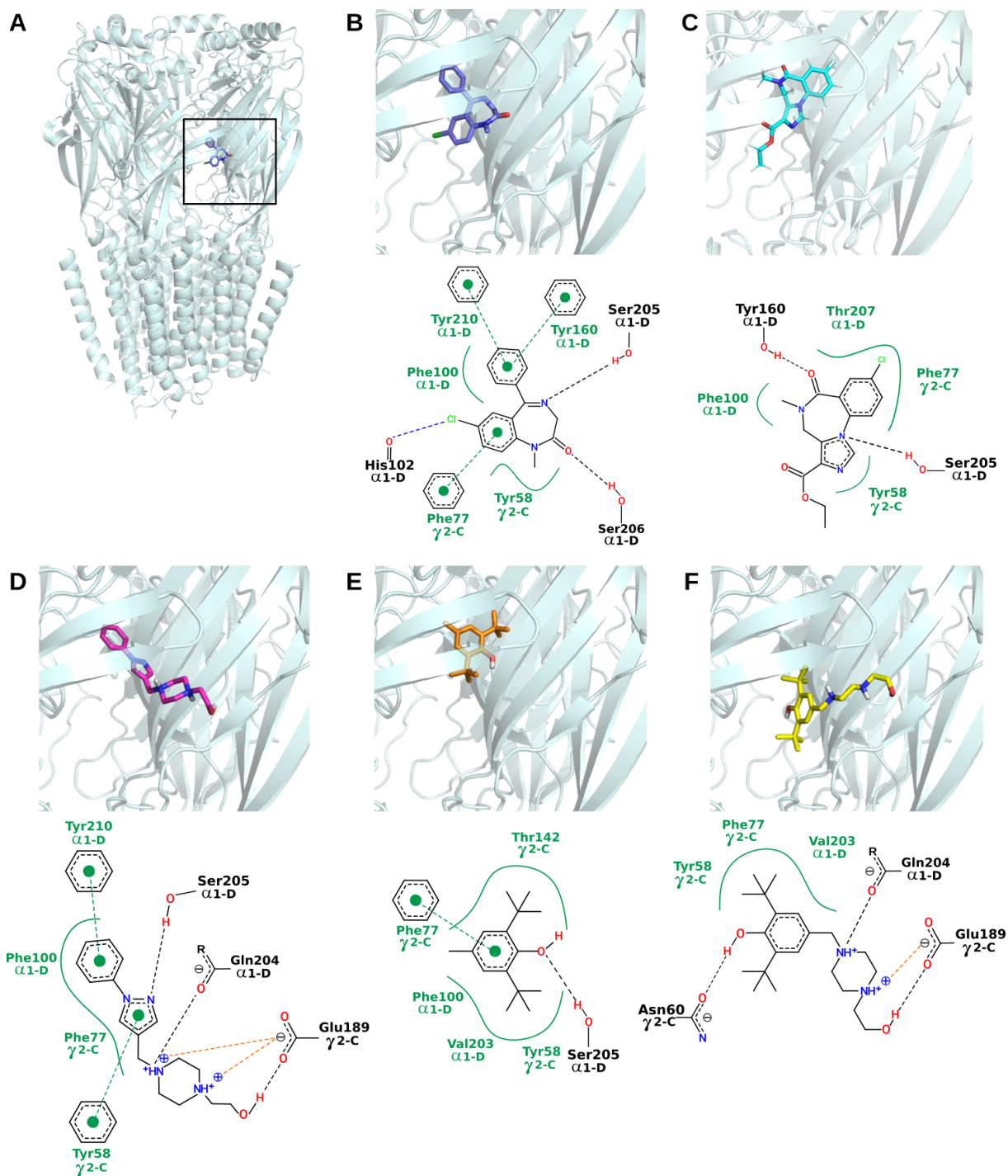


Figura S8 – Receptor GABA<sub>A</sub> com as poses de ligação dos ligantes e o detalhamento das interações observadas. A) O receptor (ciano) e a região de atracamento em detalhe. Os ligantes são B) controle do agonismo diazepam, C) controle do antagonismo flumazenil, D) LQFM032 (1), E) BHT (2) e LQFM212 (3). As linhas tracejadas representam interações hidrofóbicas (verde, as esferas são empilhamento  $\pi$ ), ligação de hidrogênio (preto), ponte salina (laranja) e ligação de halogênio (azul).

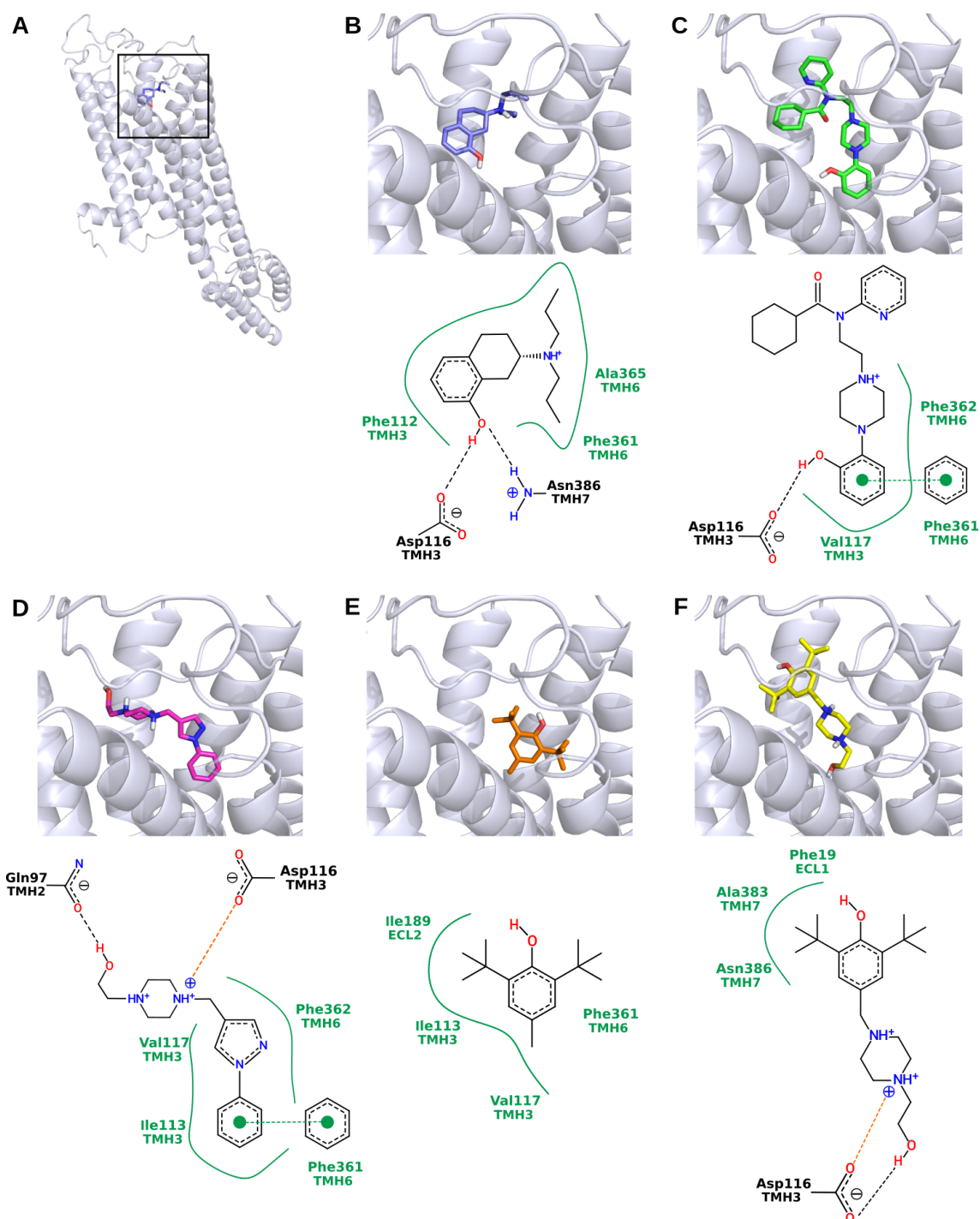


Figura S9 – Receptor 5-HT<sub>1A</sub> com as poses de ligação dos ligantes e o detalhamento das interações observadas. A) O receptor (prata) e a região de atracamento em detalhe. Os ligantes são B) controle do agonismo 8-OH-DPAT, C) controle do antagonismo WAY-100635, D) LQFM032 (1), E) BHT (2) e LQFM212 (3). As linhas tracejadas representam interações hidrofóbicas (verde, esferas são empilhamento  $\pi$ ), ligação de hidrogênio (preto) e ponte salina (laranja).

## APÊNDICE B – Triagem Virtual de Ligantes Peptídicos anti-SARS-CoV-2

Em colaboração com os grupos do professor Dr. Alexandre Macedo, do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul, e dos professores Dr. Walter Beys e Dra. Lucélia Santi, do Centro de Pesquisa Experimental do Hospital de Clínicas de Porto Alegre, a estratégia de HTS foi realizada para identificar, a partir de um banco composto por pelo menos 5000 peptídeos multiespécie, ligantes peptídicos capazes de bloquear a interação de SARS-CoV-2 com seu principal receptor celular, a enzima conversora de angiotensina 2 (ECA2).

Os peptídeos foram testados quanto à afinidade no atracamento molecular em duas regiões de ECA2: no domínio de interação com a proteína *spike* do SARS-CoV-2, a hélice *N*-terminal, e na fenda catalítica da enzima. A triagem virtual em dois domínios de ECA2 teve por objetivo selecionar peptídeos capazes de impedir a interação com o vírus e evitar ligantes que comprometessem a atividade da enzima. Os 10 peptídeos com melhor afinidade teórica, foram selecionados, sintetizados e testados em cultura celular. A partir de análises *in vitro*, dois peptídeos mostraram significativa propriedade antiviral. Análises das características de interação dos complexos proteína-ligante, envolvendo os peptídeos bioativos, foram realizadas e revelaram diferenças na estabilidade da interação que podem estar diretamente relacionadas a diferenças na  $IC_{50}$  observada para a contenção da proliferação viral.

Neste trabalho estão sendo apresentados dados parciais associados a resultados preliminares das análises *in vitro*, que ainda estão sendo conduzidas em laboratório de apoio.

# *Virtual Screening* of Peptide Ligands Against SARS-CoV-2 Infection

Crisciele Fontana, João Luiz de Meirelles, Hugo Verli

---

---

## 1. Introduction

Since the end of 2019 the world faces a global pandemic of acute respiratory tract infections caused by the novel coronavirus SARS-CoV-2, first identified in Wuhan, Hubei Province, China [1, 2]. This virus, responsible for the outbreak denominated coronavirus disease 2019 (COVID-19), is much more contagious than previous viruses involved in severe acute respiratory syndrome (SARS) coronavirus (SARS-CoV) in 2002 and Middle East Respiratory Syndrome (MERS) coronavirus (MERS-CoV) in 2012 [3].

The SARS-CoV-2 is a RNA enveloped virus containing 29.9 kb presenting until 79.6% of identity with SARS-CoV [4], and both viruses use the angiotensin-converting enzyme 2 (ACE2) as a cellular entry receptor [5, 6]. The coronavirus are comprised by the structural proteins that compose the envelope (E), the membrane (M), the nucleocapsid (N), and spike (S) proteins. These latter are characterized by large protrusions from the virus surface, giving the appearance of a crown, what gave rise to the name coronavirus [7]. The S protein consists of subunits S1 and S2, responsible for the attachment and membrane fusion, respectively. The SARS-CoV-2 S protein binds to human ACE2 in the cell membrane through the S1 subunit of the receptor-binding domain (RBD) with a 10- to 20-fold higher affinity when compared to SARS-CoV RBD binding capacity [8, 6]. This enhanced affinity for human ACE2 may contribute to SARS-CoV-2 higher infectivity, as COVID-19 is globally widespread and new cases are constantly reported [7].

The SARS-CoV-2 binding to the host ACE2 receptor is the start step to the virus entry, replication, and spread. Accordingly, to prevent this interaction is an approach highly promising therapy to treat the viral infection through the blocking of the virus accessibility to the membrane-bound ACE2 [9]. In order to disrupting extended protein binding interfaces, peptides offer



28 a synthetically accessible solution to disrupt protein-protein interaction by  
29 binding at interface regions containing multiple contact hot spots [10].

30 In this context, we dispose a library of multi-species peptides to be ex-  
31 plored in order to identify inhibitors of the SARS-CoV-2 binding to its cell  
32 receptor ACE2. The library is constituted with natural peptides or generated  
33 from biochemical-enzymatic processes, using proteins from different sources:  
34 animal, vegetable, fungi, yeast and bacteria [11, 12, 13, 14, 15, 16, 17]. The  
35 search of protein nature inhibitors, such as peptides and miniproteins, to  
36 prevent the SARS-CoV-2-ACE2 complexation is a strategy that has been  
37 explored since the beginning of the pandemic [18, 19, 20, 21, 22, 23, 24].

## 38 2. Methods

### 39 2.1. Preparing the peptides and ACE2 structures

40 The peptide sequences from the database were initially converted to  
41 FASTA format, before the PeptideBuilder library [25] for Python generates  
42 the extended peptide structure in PDB format. The dihedral angles used  
43 were  $\phi = -120^\circ$ ,  $\psi = 140^\circ$  and  $\omega = 180^\circ$ , based on the rotamer library by  
44 Shapovalov Dunbrack [26]. The peptides were ionized at pH = 7.4 through  
45 the LigPrep ionizer from Maestro Schrödinger [27].

46 The ACE2 structures available in protein data bank entries 6M0J [28] and  
47 1R4L [29] were accessed. The complex 1R4L was used as inhibitory control of  
48 ACE2 through a crossdocking of its ligand at the catalytic site of the protein  
49 structure present in 6M0J, that was applied to the screening of peptides.  
50 In this way, the inhibitory effect in the pre-complexation conformation of  
51 ACE2 were controlled. ACE2 structure from 6M0J was ionized at pH = 7.4  
52 and the protonation state was assigned through PROPKA [30]. PrepWizard  
53 from Maestro Schrödinger [27] optimized the structure and the hydrogens  
54 network before curated by visual inspection at the target domains, both the  
55 catalytic site and S protein interaction domain. The  $\text{Zn}^{2+}$  coordination in  
56 the catalytic site was kept.

### 57 2.2. Running the virtual screening

58 The extended structure of the peptide was used as input to the ligand  
59 flexible virtual screening through the DockThor program [31, 32, 33] using  
60 ACE2 as target protein.

61 Peptide ligands, containing from 6 to 14 residues, were grouped by the  
62 sequence size and jobs were performed separately for each group in two dif-  
63 ferent ACE2 domain. The search space was limited by grids positioned on  
64 the domain of interaction with the S protein, in the *N*-terminal helix [34],  
65 and in the catalytic cavity [29] of ACE2, taking care to include the residues  
66 involved with interactions as showed in original complexes. This approach  
67 was chosen to select peptides able to inhibit the virus interaction with its  
68 receptor and to exclude peptides that interfere in ACE2 catalytic activity.

69 The top 1 to 5 ligands from each run were imported in Maestro Schrödinger  
70 [27] to a visual inspection of the complexes. The presence of interactions was  
71 used as a criterion for the run inclusion while avoiding the presence of ligands  
72 with positive intramolecular energy. The top 5 ligands from each run were  
73 pooled and reordered according to the affinity energy [33]. The 10 peptides  
74 with the best predicted binding affinity to the site of interaction of ACE2  
75 with the S protein, were synthesized and tested for the antiviral activity in  
76 culture of pulmonary fibroblasts (MRC5 strain) in a support laboratory.

### 77 2.3. Running the molecular dynamics

78 Based on the results from the *in vitro* analyses, molecular dynamics sim-  
79 ulations of peptide–ACE2 complexes of the most biologically active ligands  
80 were performed through GROMOS53A6 [35] force field.

81 In order to refine the interaction aspects between the peptides and the  
82 protein target ACE2, molecular dynamics of 100 ns were performed after  
83 the system energy minimization and equilibration. The temperature and  
84 pressure were kept constant, at 298 K e at 1 bar, respectively, using the  
85 V-rescale thermostat [36] and Parrinello-Rahman barostat [37]. The simula-  
86 tions were carried out in aqueous solvent under periodic boundary conditions  
87 and integration step of 2 fs. Electrostatic interactions were treated by the  
88 particle-mesh Ewald (PME) algorithm [38], while the van der Waals con-  
89 tacts with the Verlet *cutoff-scheme* [39]. The covalent bonds lengths were  
90 constrained by the LINCS method [40].

## 91 3. Results

92 Assessing the SARS-CoV-2–ACE2 co-crystal structure, shows that the  
93 interaction interface spans a large elongated surface area, as is common for  
94 protein-protein interactions [19]. Considering this interface of interaction and  
95 the presence of hot spot residues involved in this complexation, were selected

106 10 peptides with better affinity predictions to binding ACE2 *N*-terminal  
 107 helix to be synthesized and tested in cell culture. The affinity energy of the  
 108 peptides to binding the ACE2 *N*-terminal helix and its catalytic cavity are  
 109 shown in Table 1.

100 Table 1: Affinity measures of peptide ligands predicted to binding to the  
 101 *N*-terminal helix and in the catalytic cavity of ACE2.

Peptide ID	size (residues)	<i>N</i> -terminal affinity (kcal/mol)	Catalytic cavity affinity (kcal/mol)
1362 peptide	10	-10.123	-7.611
1459 peptide	10	-10.110	-8.677
1546 peptide	10	-9.825	-10.310
1761 peptide	10	-9.809	-9.055
1753 peptide	10	-9.685	-10.271
3181 peptide	12	-9.221	-9.976
3818 peptide	13	-8.819	-9.963
2286 peptide	11	-8.759	-9.464
2529 peptide	11	-8.709	-9.710
2374 peptide	11	-8.707	-9.162

102 From the analyses carried out in culture of pulmonary fibroblasts, the  
 103 peptides 1459 and 1761 showed bioactive properties against the SARS-CoV-  
 104 2 infection, presenting IC<sub>50</sub> of 470.47  $\mu$ M and 296.29  $\mu$ M, respectively. The  
 105 analysis of peptide–ACE2 complexes from molecular docking showed electro-  
 106 static interactions between these ligands and residues in the ACE2 domain  
 107 where the SARS-CoV-2 S protein binds, mainly Asp30, Glu35, Gln42, and  
 108 Lys353 [34, 28], as presented in Figure 1. Molecular dynamics of these com-  
 109 plexes revealed differences regarding the stability of the peptide ligands at  
 110 the interaction domain with the ACE2 *N*-terminal helix. The complex with  
 111 1761 ligand showed stability on interaction site, as evaluated through root-  
 112 mean-square deviation of atomic positions (RMSD), when compared to the  
 113 complex with 1459 ligand. Both structures fluctuation are showed in Figure  
 114 2, which can be directly related to the ligands IC<sub>50</sub> obtained in cell culture.

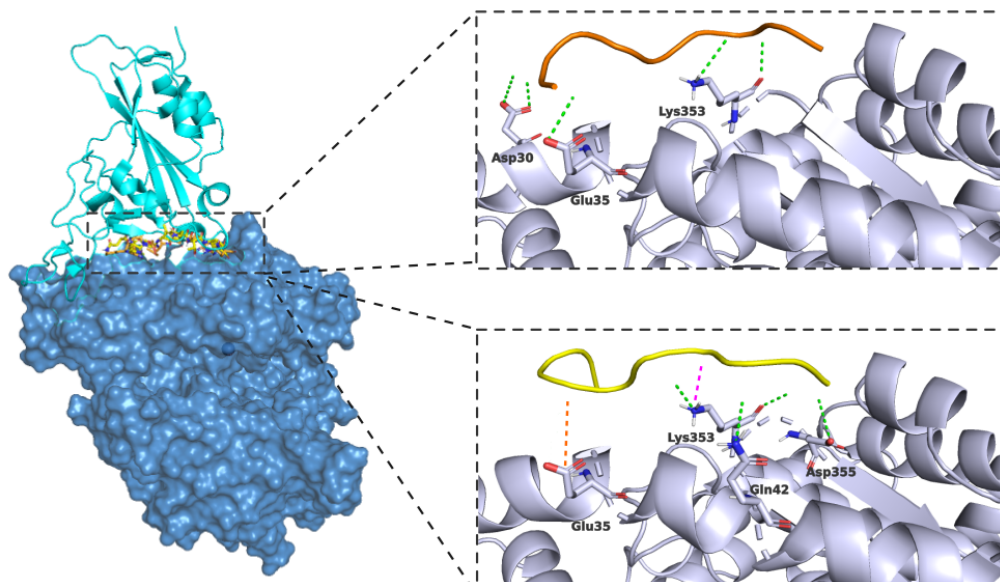


Figure 1: The peptide 1459 (orange) makes hydrogen bonds (green dashes) with residues in the ACE2 domain where the SARS-CoV-2 S protein binds. The peptide 1761 (yellow) makes even cation- $\pi$  interaction (magenta dashes) and salt-bridge (orange dashes).

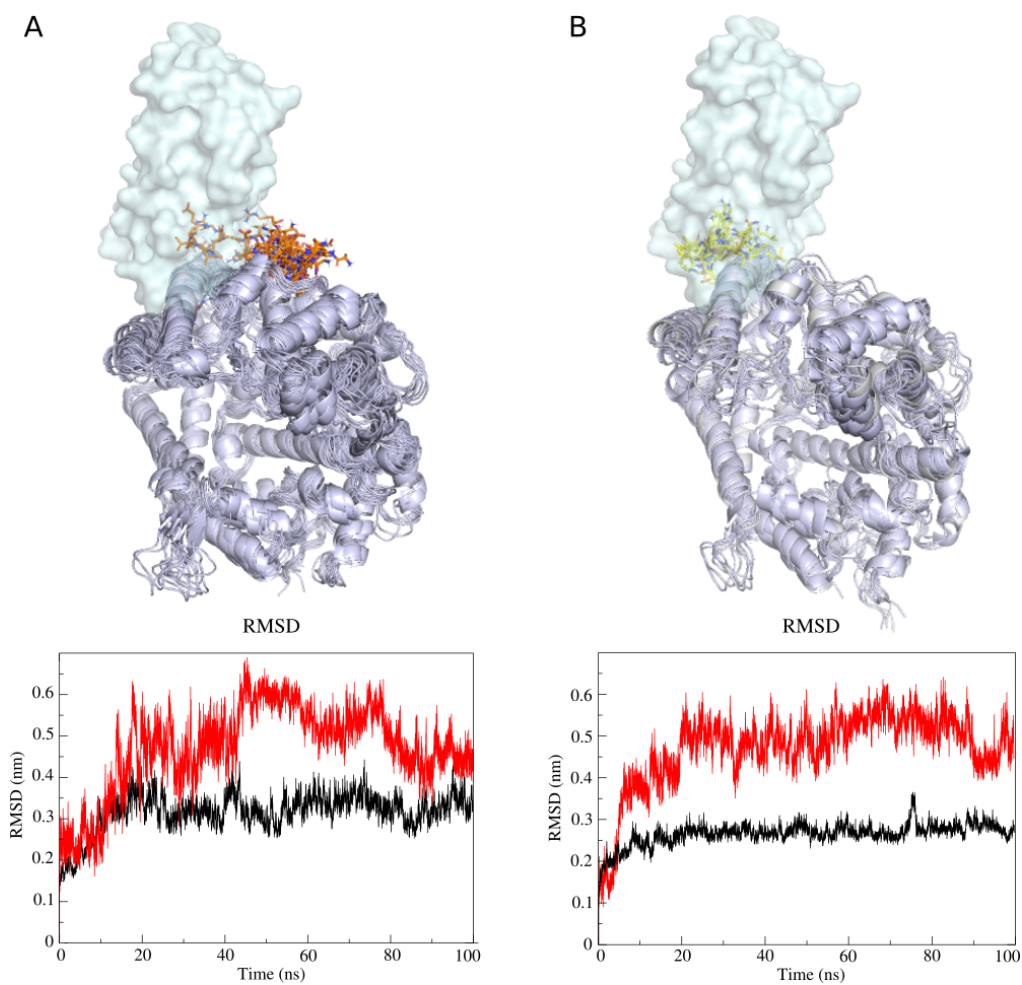


Figure 2: Molecular dynamics simulation of peptide–ACE2 complexes from docking showed a displacement of the ligand 1459 from the ACE2 interaction surface with the SARS-CoV-2 S protein (A), while the ligand 1761 was more stable in the binding domain (B) during the 100 ns observed. RMSD analyses of the ACE2 receptor is shown in black and RMSD of the peptide ligands in red. S protein was after positioned in the complex to indicate the interaction surface.

#### 115 4. Discussion

116 When one compares the ligands predicted affinities to binding ACE2, it  
117 is possible notice a very similar free energy to binding the *N*-terminal helix  
118 and the catalytic cavity. Although these predictions helped to chose the most  
119 likely peptides to prevent the S protein binding, this measure should not be  
120 rigidly considered, due to variations in the scoring function employed to de-  
121 scribe the binding event, to predict affinity data using experimental measures,  
122 and finally to ranking the ligands based on these descriptors [41, 33]. Then,  
123 we chose the first 10 ligands with the best predict affinity to binding the  
124 *N*-terminal helix, to be tested in cell culture, after careful visual inspection.  
125 These ligands does not displayed important interactions with residues in the  
126 catalytic cavity, so were included as candidate to prevent e SARS-CoV-2-  
127 ACE2 complexation. Previous reports suggested the enzymatic activity and  
128 viral receptor property of ACE2 are independent, since most known ACE2  
129 inhibitors reduce ACE2 enzymatic activity with no effects on the SARS-CoV  
130 S protein binding [42, 9]. Moreover, studies indicated that the binding of a  
131 virus to ACE2 does not necessarily interfere on its enzymatic activity [43, 44].

132 The *in vitro* analysis indicated the 1459 and 1761 ligands were able to  
133 contain the SARS-CoV-2 infection. The ligand 1459 presented electrostatic  
134 interactions with the residues Asp30 and Glu35, in the ACE2 receptor *N*-  
135 terminal helix, and with Lys353, in a nearby coil present in the protein S  
136 binding surface. The ligand 1761 presented electrostatic interactions with  
137 the residues Glu35 and Gln42, in the *N*-terminal helix, and with Lys353  
138 and Asp355, in the nearby coil. Both peptide ligands showed electrostatic  
139 interaction with the Glu35 side chain and two contacts with the Lys353  
140 residue, with the backbone carboxyl group and with the side chain amino  
141 group. The ligand 1761 even showed a cation- $\pi$  interaction with the amino  
142 group in Lys353, while a salt-bridge was observed with the Glu35 residue  
143 (Figure 1).

144 To better understand how complexation features are associated with the  
145 differences observed in  $IC_{50}$  of the ligands, molecular dynamics simulations of  
146 peptide-ACE2 complexes were carried out to refine the information regarding  
147 the interactions data observed from molecular docking. During the simula-  
148 tion, the complex 1459-ACE2 preserved interactions with residues Asp30  
149 and Lys353, even additional contact with His34, Glu37, Tyr41, and Asp355.  
150 While the complex 1761-ACE2 preserved interactions just with Lys353, but  
151 showed new ones with Asp30, His34, Glu37, and Asp38. As can be observed

152 in the 6M0J complex, the S protein establishes electrostatic contacts with  
153 the residues Asp30, Glu35, Asp38, Tyr41, Gln42, Lys353, and Arg393, be-  
154 sides others residues [28]. While other complex, the 6M17, indicated contacts  
155 with Gln24, Asp30, His34, Tyr41, Gln42, Lys353, and Arg357 [34]. Interac-  
156 tions between the peptides and the same residues which the S protein makes  
157 contact, as data from 6M0J and 6M17 complexes, corroborate the bioactive  
158 properties of both 1459 and 1761 ligands contain the infection by SARS-  
159 CoV-2 through a competitive way.

160 The RMSD analyses showed less conformational variation related to 1761-  
161 ACE2 complex, indicating a more stable complexation when compared to the  
162 complex with the 1459 peptide, which slipped of the surface interaction (Fig-  
163 ure 2. The permanence of 1761 ligand on the contact region with S protein  
164 may be a factor which contributed to the lower IC<sub>50</sub> to contain the viral  
165 proliferation when compared to 1459 ligand.

## 166 5. Conclusion

167 From an extensive library of multispecies peptides, it was possible to  
168 screen ligands with potential application on prevention of the SARS-CoV-2-  
169 ACE2 interaction and consequently to contain the viral infection causing the  
170 worldwide prevalent COVID-19. Data generated for prediction of binding  
171 affinity and molecular dynamics between peptide-ACE2 interaction agree  
172 with results obtained from *in vitro* analysis of the bioactive properties of the  
173 ligands. This alignment reinforces the appropriate structures preparation  
174 strategies as well as the virtual screening protocol developed for this search.

175 In order to better understand the structural determinants of the complex-  
176 ation and to improve the efficacy of antiviral peptides through a rational drug  
177 design aspect, additional analysis and new search strategies should be em-  
178 ployed. New virtual screening including new peptide candidates through an  
179 enriched search by similarity, using the results from this round as positional  
180 restrictions, is a promising approach.

## 181 References

- 182 [1] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang,  
183 W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F.  
184 Gao, W. Tan, A Novel Coronavirus from Patients with Pneumonia in  
185 China, 2019, New England Journal of Medicine 382 (2020) 727-733.

- 186 [2] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan,  
187 J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin,  
188 H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang,  
189 Z. Gao, Q. Jin, J. Wang, B. Cao, Clinical features of patients infected  
190 with 2019 novel coronavirus in Wuhan, China, *The Lancet* 395 (2020)  
191 497–506.
- 192 [3] W. Ni, X. Yang, D. Yang, J. Bao, R. Li, Y. Xiao, C. Hou, H. Wang,  
193 J. Liu, D. Yang, Y. Xu, Z. Cao, Z. Gao, Role of angiotensin-converting  
194 enzyme 2 (ACE2) in COVID-19, *Critical Care* 24 (2020) 422.
- 195 [4] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R.  
196 Si, Y. Zhu, B. Li, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo,  
197 R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng,  
198 K. Zhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhan, Y.-Y.  
199 Wang, G.-F. Xiao, Z.-L. Shi, A pneumonia outbreak associated with a  
200 new coronavirus of probable bat origin, *Nature* 579 (2020) 270–273.
- 201 [5] W. Li, M. J. Moore, N. Vasilieva, J. Sui, S. K. Wong, M. A. Berne, M. So-  
202 masundaran, J. L. Sullivan, K. Luzuriaga, T. C. Greenough, H. Choe,  
203 M. Farzan, Angiotensin-converting enzyme 2 is a functional receptor for  
204 the SARS coronavirus, *Nature* 426 (2003) 450–454.
- 205 [6] W. Tai, L. He, X. Zhang, J. Pu, D. Voronin, S. Jiang, Y. Zhou, L. Du,  
206 Characterization of the receptor-binding domain (RBD) of 2019 novel  
207 coronavirus: implication for development of RBD protein as a viral at-  
208 tachment inhibitor and vaccine, *Cellular Molecular Immunology* 17  
209 (2020) 613–620.
- 210 [7] S. Beyerstedt, E. B. Casaro, É. B. Rangel, COVID-19: angiotensin-  
211 converting enzyme 2 (ACE2) expression and tissue susceptibility to  
212 SARS-CoV-2 infection, *European Journal of Clinical Microbiology In-*  
213 *fectious Diseases* 40 (2021) 905–919.
- 214 [8] D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C.-L. Hsieh,  
215 O. Abiona, B. S. Graham, J. S. McLellan, Cryo-EM structure of the  
216 2019-nCoV spike in the prefusion conformation, *Science* 367 (2020)  
217 1260–1263.



- 218 [9] H. Jia, E. Neptune, H. Cui, Targeting ACE2 for COVID-19 Therapy:  
219 Opportunities and Challenges, *American Journal of Respiratory Cell  
220 and Molecular Biology* 64 (2021) 416–425.
- 221 [10] K. Josephson, A. Ricardo, J. W. Szostak, mRNA display: from ba-  
222 sic principles to macrocycle drug discovery, *Drug Discovery Today* 19  
223 (2014) 388–399.
- 224 [11] L. Santi, W. O. Beys-da Silva, M. Berger, D. Calzolari, J. A. Guimarães,  
225 J. J. Moresco, J. R. Yates, Proteomic Profile of *Cryptococcus neoformans*  
226 Biofilm Reveals Changes in Metabolic Processes, *Journal of Proteome  
227 Research* 13 (2014) 1545–1559.
- 228 [12] L. Santi, C. J. B. Coutinho-Rodrigues, M. Berger, L. A. S. Klein, E. M.  
229 De Souza, R. L. Rosa, J. A. Guimarães, J. R. Yates, W. M. S. Perinotto,  
230 V. R. E. P. Bittencourt, W. O. Beys-da Silva, Secretomic analysis of  
231 *Beauveria bassiana* related to cattle tick, *Rhipicephalus microplus*, in-  
232 fection, *Folia Microbiologica* 64 (2019) 361–372.
- 233 [13] W. O. Beys-da Silva, L. Santi, M. Berger, D. Calzolari, D. O. Passos,  
234 J. A. Guimarães, J. J. Moresco, J. R. Yates, Secretome of the Biocontrol  
235 Agent *Metarhizium anisopliae* Induced by the Cuticle of the Cotton Pest  
236 *Dysdercus peruvianus* Reveals New Insights into Infection, *Journal of  
237 Proteome Research* 13 (2014) 2282–2296.
- 238 [14] R. Fontoura, D. J. Daroit, A. P. F. Corrêa, K. S. Moresco, L. Santi,  
239 W. O. Beys-da Silva, J. R. Yates, J. C. F. Moreira, A. Brandelli,  
240 Characterization of a novel antioxidant peptide from feather keratin  
241 hydrolysates, *New Biotechnology* 49 (2019) 71–76.
- 242 [15] R. L. Rosa, M. Berger, L. Santi, D. Driemeier, P. Barros Terraciano,  
243 A. R. Campos, J. A. Guimarães, M. H. Vainstein, J. R. Yates, W. O.  
244 Beys-da Silva, Proteomics of Rat Lungs Infected by *Cryptococcus gattii*  
245 Reveals a Potential Warburg-like Effect, *Journal of Proteome Research*  
246 18 (2019) 3885–3895.
- 247 [16] A. Weber, L. E. Argenti, A. P. B. de Souza, L. Santi, W. O. Beys-da  
248 Silva, J. R. Yates, I. C. Bustamante-Filho, Ready for the journey: a  
249 comparative proteome profiling of porcine cauda epididymal fluid and  
250 spermatozoa, *Cell and Tissue Research* 379 (2020) 389–405.

- 251 [17] É. A. Blasi, G. Buffon, A. G. Rativa, M. C. Lopes, M. Berger, L. Santi,  
252 M. Lavallée-Adam, J. R. Yates, J. Schwambach, W. O. Beys-da Silva,  
253 R. A. Sperotto, High infestation levels of *Schizotetranychus oryzae*  
254 severely affects rice metabolism, *Journal of Plant Physiology* 219 (2017)  
255 100–111.
- 256 [18] Y. Han, P. Král, Computational Design of ACE2-Based Peptide In-  
257 hibitors of SARS-CoV-2, *ACS Nano* 14 (2020) 5143–5147.
- 258 [19] G. Zhang, S. Pomplun, A. R. Loftis, A. Loas, B. L. Pentelute, The  
259 first-in-class peptide binder to the SARS-CoV-2 spike protein, 2020.
- 260 [20] M. Wolfe, S. Webb, Y. Chushak, R. Krabacher, Y. Liu, N. Swami,  
261 S. Harbaugh, J. Chávez, A high-throughput pipeline for design and  
262 selection of peptides targeting the SARS-Cov-2 Spike protein, *Scientific*  
263 *Reports* 11 (2021) 21768.
- 264 [21] Y.-N. Zhang, Y. Zhang, S. Su, H.-Y. Zhu, W. Xu, L. Wang, M. Wu,  
265 K. Chen, F.-Q. Yu, T.-K. Xi, Q. Zhou, Y.-H. Xie, X. Qin, H. Ge, L. Lu,  
266 J. Qing, G.-M. Fang, Neutralizing SARS-CoV-2 by dimeric side chain-  
267 to-side chain cross-linked ACE2 peptide mimetics, *Chemical Communi-*  
268 *cations* 58 (2022) 1804–1807.
- 269 [22] B. Jawad, P. Adhikari, K. Cheng, R. Podgornik, W.-Y. Ching, Compu-  
270 tational Design of Miniproteins as SARS-CoV-2 Therapeutic Inhibitors,  
271 *International Journal of Molecular Sciences* 23 (2022) 838.
- 272 [23] T. E. Tallei, Fatimawali, A. A. Adam, M. M. Elseehy, A. M. El-  
273 Shehawi, E. A. Mahmoud, A. D. Tania, N. J. Niode, D. Kusumawaty,  
274 S. Rahimah, Y. Effendi, R. Idroes, I. Celik, M. J. Hossain, T. B. Emran,  
275 Fruit Bromelain-Derived Peptide Potentially Restrains the Attachment  
276 of SARS-CoV-2 Variants to hACE2: A Pharmacoinformatics Approach,  
277 *Molecules* 27 (2022) 260.
- 278 [24] S. Rajpoot, K. Solanki, A. Kumar, K. Y. J. Zhang, S. S. Pullamsetti,  
279 R. Savai, S. M. Faisal, Q. Pan, M. S. Baig, In-Silico Design of a  
280 Novel Tridecapeptide Targeting Spike Protein of SARS-CoV-2 Variants  
281 of Concern, *International Journal of Peptide Research and Therapeutics*  
282 28 (2022) 28.

- 283 [25] M. Z. Tien, D. K. Sydykova, A. G. Meyer, C. O. Wilke, PeptideBuilder:  
284 A simple Python library to generate model peptides, *PeerJ* 1 (2013)  
285 e80.
- 286 [26] M. V. Shapovalov, R. L. Dunbrack, A Smoothed Backbone-Dependent  
287 Rotamer Library for Proteins Derived from Adaptive Kernel Density  
288 Estimates and Regressions, *Structure* 19 (2011) 844–858.
- 289 [27] G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sher-  
290 man, Protein and ligand preparation: parameters, protocols, and in-  
291 fluence on virtual screening enrichments, *Journal of Computer-Aided*  
292 *Molecular Design* 27 (2013) 221–234.
- 293 [28] J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi,  
294 Q. Wang, L. Zhang, X. Wang, Structure of the SARS-CoV-2 spike  
295 receptor-binding domain bound to the ACE2 receptor, *Nature* 581  
296 (2020) 215–220.
- 297 [29] P. Towler, B. Staker, S. G. Prasad, S. Menon, J. Tang, T. Parsons,  
298 D. Ryan, M. Fisher, D. Williams, N. A. Dales, M. A. Patane, M. W.  
299 Pantoliano, ACE2 X-Ray Structures Reveal a Large Hinge-bending Mo-  
300 tion Important for Inhibitor Binding and Catalysis, *Journal of Biological*  
301 *Chemistry* 279 (2004) 17996–18007.
- 302 [30] M. H. M. Olsson, C. R. Søndergaard, M. Rostkowski, J. H. Jensen,  
303 PROPKA3: Consistent Treatment of Internal and Surface Residues in  
304 Empirical pK<sub>a</sub> Predictions, *Journal of Chemical Theory and Compu-*  
305 *tation* 7 (2011) 525–537.
- 306 [31] F. L. Custódio, H. J. Barbosa, L. E. Dardenne, A multiple minima  
307 genetic algorithm for protein structure prediction, *Applied Soft Com-*  
308 *puting* 15 (2014) 88–99.
- 309 [32] C. S. de Magalhães, D. M. Almeida, H. J. C. Barbosa, L. E. Dardenne,  
310 A dynamic niching genetic algorithm strategy for docking highly flexible  
311 ligands, *Information Sciences* 289 (2014) 206–224.
- 312 [33] K. B. Santos, I. A. Guedes, A. L. M. Karl, L. E. Dardenne, Highly  
313 Flexible Ligand Docking: Benchmarking of the DockThor Program on  
314 the LEADS-PEP Protein–Peptide Data Set, *Journal of Chemical Infor-*  
315 *mation and Modeling* 60 (2020) 667–683.

- 316 [34] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Structural basis  
317 for the recognition of SARS-CoV-2 by full-length human ACE2, *Science*  
318 367 (2020) 1444–1448.
- 319 [35] C. Oostenbrink, A. Villa, A. E. Mark, W. F. Van Gunsteren, A biomolec-  
320 ular force field based on the free enthalpy of hydration and solvation:  
321 The GROMOS force-field parameter sets 53A5 and 53A6, *Journal of*  
322 *Computational Chemistry* 25 (2004) 1656–1676.
- 323 [36] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through ve-  
324 locity rescaling, *The Journal of Chemical Physics* 126 (2007) 014101.
- 325 [37] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A  
326 new molecular dynamics method, *Journal of Applied Physics* 52 (1981)  
327 7182–7190.
- 328 [38] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An  $N \log(N)$   
329 method for Ewald sums in large systems, *The Journal of Chemical*  
330 *Physics* 98 (1993) 10089–10092.
- 331 [39] S. Páll, B. Hess, A flexible algorithm for calculating pair interactions  
332 on SIMD architectures, *Computer Physics Communications* 184 (2013)  
333 2641–2650.
- 334 [40] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A  
335 linear constraint solver for molecular simulations, *Journal of Computa-*  
336 *tional Chemistry* 18 (1997) 1463–1472.
- 337 [41] I. A. Guedes, F. S. S. Pereira, L. E. Dardenne, Empirical Scoring Func-  
338 tions for Structure-Based Virtual Screening: Applications, Critical As-  
339 pects, and Challenges, *Frontiers in Pharmacology* 9 (2018).
- 340 [42] M. J. Huentelman, J. Zubcevic, J. A. Hernandez Prada, X. Xiao, D. S.  
341 Dimitrov, M. K. Raizada, D. A. Ostrov, Structure-Based Discovery of  
342 a Novel Angiotensin-Converting Enzyme 2 Inhibitor, *Hypertension* 44  
343 (2004) 903–906.
- 344 [43] W. Li, C. Zhang, J. Sui, J. H. Kuhn, M. J. Moore, S. Luo, S.-K. Wong,  
345 I.-C. Huang, K. Xu, N. Vasilieva, A. Murakami, Y. He, W. A. Marasco,  
346 Y. Guan, H. Choe, M. Farzan, Receptor and viral determinants of

- 347 SARS-coronavirus adaptation to human ACE2, *The EMBO Journal* 24  
348 (2005) 1634–1643.
- 349 [44] A. M. Davidson, J. Wysocki, D. Battle, Interaction of SARS-CoV-2 and  
350 Other Coronavirus With ACE (Angiotensin-Converting Enzyme)-2 as  
351 Their Main Receptor, *Hypertension* 76 (2020) 1339–1349.