

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE FÍSICA DE SÃO CARLOS

RICARDO NASCIMENTO DOS SANTOS

Planejamento de moduladores de polimerização de  
microtúbulos com propriedades anticâncer, análise estrutural  
de macromoléculas e geração de uma base virtual de  
produtos naturais

São Carlos

2015



RICARDO NASCIMENTO DOS SANTOS

Planejamento de moduladores de polimerização de  
microtúbulos com propriedades anticâncer, análise estrutural  
de macromoléculas e geração de uma base virtual de  
produtos naturais

Tese apresentada ao Programa de Pós-  
Graduação em Física do Instituto de Física  
de São Carlos da Universidade de São  
Paulo, para obtenção do título de Doutor em  
Ciências.

Área de concentração: Física Aplicada  
Opção: Física Biomolecular  
Orientador: Prof. Dr. Adriano D. Andricopulo  
Coorientador: Prof. Dr. Rafael Guido

Versão Corrigida

(versão original na Unidade que aloja o Programa)

São Carlos

2015

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica revisada pelo Serviço de Biblioteca e Informação do IFSC,  
com os dados fornecidos pelo(a) autor(a)

Santos, Ricardo Nascimento dos  
Planejamento de moduladores de polimerização de microtúbulos com propriedades anticâncer, análise estrutural de macromoléculas e geração de uma base virtual de produtos naturais / Ricardo Nascimento dos Santos; orientador Adriano D. Andricopulo; co-orientador Rafael V. C. Guido - versão corrigida -- São Carlos, 2015.  
189 p.

Tese (Doutorado - Programa de Pós-Graduação em Física Biomolecular) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2015.

1. Planejamento de fármacos. 2. Câncer. 3. Predição de complexos proteicos. 4. Produtos naturais. 5. NuBBE database. I. D. Andricopulo, Adriano, orient. II. V. C. Guido, Rafael, co-orient. III. Título.

*Aos meus pais, Ronaldo e Dilma, pelo apoio incondicional.*



# Agradecimentos

Agradeço ao professor Adriano D. Andricopulo, pelas oportunidades, amizade, confiança e constante interesse e dedicação na realização dos projetos.

Ao professor José Nelson Onuchic, pela oportunidade em realizar parte do doutorado direto sobre sua orientação na Rice University, e a ele e sua esposa Mayra Onuchic, pela amizade e apoio durante os momentos difíceis longe de casa.

Ao professor Faruck Morcos e sua família, pela amizade, apoio e exemplo de profissionalismo.

À Raissa F. Gutierrez, pelo companheirismo e apoio ao longo de vários anos e pelas lições de vida, que levarei para sempre comigo.

Aos meus grandes amigos que sempre me ajudaram: Fernando Bachega, Carlos Sardinha, Leandro Farias, Henrique Hemerly, João Jorge Pereira, Heverton Nascimento, Marina Alves, Eduardo Gandini, Guilherme Lorenzen, Atílio Tomazini, Geraldo Sartori, Ivan Silva, Caio Reis, Cláudia Ferreira, Pedro Autreto e Marcelo Boareto.

Às companheiras de pesquisa, Daniara C. Fernandes e Marília Valli, pelas oportunidades de colaboração, companheirismo no laboratório e pela grande amizade cultivada ao longo destes anos.

Aos meus companheiros e técnicos do LQMC, pelas conversas e pelo aprendizado constante.

À minha família, pelo constante apoio e paciência, em especial meus pais Ronaldo e Dilma e aos meus avós Maria, Isabel e Airton.

Aos professores Fernando Coelho da Unicamp e Vanderlan Bolzani da UNESP, pela conversas e oportunidades de colaboração.

Ao professores Rafael V. C. Guido, pela coorientação e ao professor Glaucius Oliva, pelo aprendizado.

Aos colegas e ex-colegas do curso de graduação de Ciências Físicas e Biomoleculares da USP.

À Fundação de Amparo à Pesquisa de São Paulo (FAPESP), pelo auxílio financeiro necessário para a execução deste trabalho (processos 10/06741-4 e 13/11950-0).

Ao IFSC-USP pelo apoio e infraestrutura, em especial aos funcionários da CPG: Ricardo e Patricia, pela atenção e ajuda na solução de problemas acadêmicos do dia-a-dia e às funcionárias da biblioteca do IFSC, pela simpatia e ajuda na elaboração desta tese.





*“O maior bem do homem é uma mente inquieta.”*  
*(Isaac Asimov)*

*“A razão é o passo,  
o aumento da ciência o caminho,  
e o benefício da humanidade é o fim.”*  
*(Thomas Hobbes)*



# Resumo

SANTOS, R.N. **Planejamento de moduladores de polimerização de microtúbulos com propriedades anticâncer, análise estrutural de macromoléculas e geração de uma base virtual de produtos naturais.** 2015. 189 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2015.

Os trabalhos realizados e apresentados nesta tese de doutorado compreendem diversos estudos computacionais e experimentais aplicados ao planejamento de candidatos a novos fármacos para o tratamento do câncer, de uma metodologia inovadora para investigar a formação de complexos proteicos e de uma base de compostos naturais reunindo parte da biodiversidade brasileira com a finalidade de incentivar e auxiliar a descoberta e o desenvolvimento de novos fármacos no país. No primeiro capítulo, são descritos estudos que permitiram a identificação e o desenvolvimento de novas moléculas com atividade anti-câncer, através da integração de ensaios bioquímicos e métodos de modelagem molecular na área de química medicinal. Dessa forma, estudos de modelagem molecular e ensaios bioquímicos utilizando uma base de compostos disponibilizada pela colaboração com o Laboratório de Síntese de Produtos Naturais e Fármacos (LSPNF) da UNICAMP, permitiram identificar uma série de moléculas da classe ciclopenta- $\beta$ -indóis como inibidores da polimerização de microtúbulos com considerável atividade anti-câncer. Estes compostos apresentaram-se capazes de modular a polimerização de microtúbulos em ensaios *in vitro* frente ao alvo molecular e a células cancerígenas, com valores de  $IC_{50}$  na faixa de 20 a 30  $\mu M$ . Além disso, estudos experimentais permitiram identificar o sítio da colchicina na tubulina como a região de interação desta classe e ensaios de migração celular comprovaram sua atividade antitumoral. A partir dos resultados obtidos, estudos mais aprofundados de docking e dinâmica molecular permitiram elucidar as interações moleculares envolvidas no processo de ligação à proteína tubulina, e a utilização destes modelos moleculares no planejamento, síntese e avaliação de uma nova série de compostos. Com base nos dados obtidos por estudos computacionais, modificações foram propostas e novos inibidores da polimerização de tubulina foram planejados, sintetizados e avaliados, resultando na identificação de um inibidor de elevada atividade e perfil farmacodinâmico superior dentre as moléculas planejadas, com  $IC_{50}$  de 5  $\mu M$ . Concomitantemente, ensaios de citotoxicidade *in vitro* demonstraram uma interessante seletividade destes compostos por células cancerígenas em comparação a células saudáveis. Os estudos desenvolvidos com inibidores de tubulina aqui apresentados permitiram identificar moduladores da polimerização de microtúbulos com excelente perfil anti-câncer, que servirão como modelo para o desenvolvimento de novos tratamentos eficazes contra o câncer. No segundo capítulo é apresentado um novo método para prever modificações conformacionais e a formação de complexos multiméricos em sistemas proteicos. Este método foi elaborado durante os estudos desenvolvidos ao longo

de um programa de intercâmbio no laboratório The Center for Theoretical and Biological Physics (CTBP, Rice University, Estados Unidos), sob orientação do professor Dr. José Nelson Onuchic. Durante este projeto, estudos de modelagem computacional foram realizados utilizando métodos computacionais modernos desenvolvidos no próprio CTBP, tal como o método de Análise de Acoplamento Direto (DCA, do inglês *Direct-Coupling Analysis*) e um método de simulação conhecido como Modelagem Baseada em Estrutura (SBM, do inglês *Structure-Based Modeling*). Nos estudos aqui apresentados, os métodos DCA e SBM desenvolvidos no CTBP foram combinados, modificados e ampliados no desenvolvimento de uma nova metodologia que permite identificar mudanças conformacionais e elucidar mecanismos de enovelamento e oligomerização em proteínas. Os resultados obtidos através da predição de diversos complexos proteicos multiméricos com uma alta precisão mostram que este sistema é extremamente eficaz e confiável para identificar regiões de interface de contato entre proteínas a a estrutura quaternária de complexos macromoleculares. Esta nova metodologia permite a elucidação e caracterização de sistemas proteicos incapazes de serem determinados atualmente por métodos puramente experimentais. No terceiro capítulo desta tese de doutorado, é descrito a construção de uma base virtual de dados em uma iniciativa pioneira que tem como principal objetivo reunir e disponibilizar o máximo possível de toda a informação já obtida através do estudo da biodiversidade brasileira. Esta base, intitulada NuBBE DataBase, reúne diversas informações como estrutura molecular 2D e 3D e informações de atividades biológicas de diversas moléculas já isoladas pelo Núcleo de Bioensaios Biossíntese e Ecofisiologia de Produtos Naturais (NuBBE), localizado na Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP). A NuBBE<sub>DB</sub> será de grande utilidade para a comunidade científica, fornecendo a centros de pesquisa e indústrias farmacêuticas informações para estudos de modelagem molecular, metabolômica, derreplicação e principalmente para o planejamento e a identificação de novos compostos bioativos.

**Palavras-chave:** Planejamento de fármacos. Câncer. Predição de complexos proteicos. Produtos naturais. NuBBE database.

# Abstract

SANTOS, R.N. **Design of microtubule polymerization modulators with anticancer properties, structural analysis of macromolecules and development of a virtual database of natural products.** 2015. 189 p. Tese (Doutorado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2015.

The work developed during a doctorate program and shown here as a PhD thesis reports the accomplishment of a series of computational and experimental studies focused on the development of new anticancer agents, an innovative methodology for the investigation of protein complexes formation and of a new database for natural products based on the Brazilian biodiversity, in an effort to assist and encourage the discovery and development of new pharmaceutical drugs inside country. The first chapter describes studies that resulted in the identification and development of new molecules with anticancer activity through the integration of biochemical experiments and molecular modeling methods in the area of medicinal chemistry. Thus, molecular modeling studies and biochemical assays using a library of compounds provided by collaboration with the Laboratório de Síntese de Produtos Naturais e Fármacos (LSPNF) from the University of Campinas (Unicamp) have identified a number of molecules of the cyclopenta-b-indole class as inhibitors for microtubule polymerisation, with substantial anti-cancer activity. These compounds showed to be able to modulate microtubule polymerisation on *in vitro* assays against the molecular target and cancer cells with  $IC_{50}$  values in the range of 20 to 30  $\mu$ M. Moreover, experimental studies have identified the colchicine site of tubulin as in the region of interaction of this class and cell migration assays have proven their antitumour activity. Based on these results, further studies using molecular docking and molecular dynamics allowed to elucidate the molecular interactions involved in the binding process to tubulin protein, and these molecular models were used to guide the design, synthesis and evaluation of a novel series of compounds. From the data obtained by computational studies, modifications were proposed to design, synthesise and evaluate new tubulin polymerisation inhibitors, resulting in identification of a high-activity inhibitor and superior pharmacodynamic profile and  $IC_{50}$  of 5  $\mu$ M. Alongside, *in vitro* cytotoxicity assays demonstrated an interesting selectivity of these compounds for cancer cells when compared to healthy cells. The studies presented here with tubulin inhibitors allowed to identify modulators of microtubule polymerisation with excellent anti-cancer profile, that will provide a valuable scaffold for the development of new effective treatments against cancer. The second chapter presents a new method for predicting changes on the conformational and the formation of multimeric protein complexes. This method was developed during the studies carried out over an exchange program in the Center for Theoretical and Biological Physics (CTBT, Rice University, USA), under the supervision of professor Dr. José Nelson Onuchic. During this project, computer modeling studies were carried using modern methods devel-

oped in the CTBT itself, such as Direct-Coupling Analysis (DCA) and a simulation method known as Modeling Based Structure (SBM). In the studies presented here, the DCA and SBM methods developed in CTBP were combined, modified and expanded to develop a new methodology able to identify the conformational changes and to elucidate mechanisms folding and oligomerization of proteins. The results obtained through prediction of various multimeric protein complexes with high accuracy show that this system is extremely effective and reliable to identify interface contacts between proteins and to predict the quaternary structure of macromolecular complexes. This new method allows the characterization and elucidation of protein systems that are currently unable to be solely determined by experimental methods. The third chapter of this doctoral thesis describes the construction of a virtual database in a pioneering initiative that aims to gather and make available all the information already obtained through the study of Brazilian biodiversity. This database, entitled NuBBE DataBase, brings together various information such as 2D and 3D molecular structure and biological activity of several molecules already isolated by the Núcleo de Bioensaios Biossíntese e Ecofisiologia de Produtos Naturais(NuBBE), located at the Universidade Estadual Paulista Julio de Mesquita Filho (UNESP). The NuBBEDB will be useful to the scientific community, providing research and pharmaceutical centers information for molecular modeling studies, metabolomics, derreplication and principally for the planning and identification of new bioactive compounds.

**Keywords:** Drug design. Cancer. Prediction of protein complexes. Natural products. NuBBE database.

# Lista de Figuras

- Figura 1 – Número de novas entidade moleculares (NEM) regulamentadas nos Estados Unidos ao longo dos últimos anos desde o ano de 1993.<sup>18</sup> . . . . . 36
- Figura 2 – Etapas envolvidas no desenvolvimento racional de um novo fármaco nas indústrias farmacêuticas modernas. . . . . 37
- Figura 3 – (A) Processo de formação de um tumor a partir da divisão de células cancerígenas. Em alguns tipos de tumores, as células podem migrar para diferentes regiões do organismo formando tumores secundários, em um processo de metástase. (B) Fases da divisão celular. Células cancerígenas apresentam alterações em mecanismos essenciais na execução de processos biomoleculares, como a verificação e a manutenção do material genético ao longo do ciclo celular. . . . . 39
- Figura 4 – (A) Incidência de diferentes tipos de câncer em mulheres nos Estados Unidos entre 1930 e 2011. (B) Incidência de diferentes tipos de câncer em homens nos Estados Unidos entre 1930 e 2011.<sup>55</sup> . . . . . 41
- Figura 5 – Estruturas químicas de agentes anticâncer que atuam pela modulação da formação de microtúbulos. Estas moléculas inibem a polimerização dos microtúbulos ligando-se em sítios distintos da tubulina. (A) Paclitaxel é um produto natural extraído da árvore teixo do pacífico (*Taxus brevifolia*) que estabiliza os microtúbulos. (B) Vimblastina, obtida da planta vinca rosea (*Catharanthus roseus*), é o princípio ativo utilizado para o tratamento de tumores específicos, como linfoma de Hodgkin e câncer de testículos. (C) Colchicina, extraída das plantas colchicum (*Colchicum autumnale*), é um composto que vem sendo estudado pelas suas propriedades anticancerígenas, apesar de seus análogos geralmente apresentarem elevada toxicidade. . . . . 43
- Figura 6 – Principais sítios de ligação de moduladores da polimerização de tubulina e formação de microtúbulos. Figura gerada pela sobreposição de diversas estruturas cristalográficas contendo dímeros de  $\alpha\beta$ -tubulina em complexo com os compostos: vimblastina (código PDB: 1Z2B), colchicina (código PDB: 1SA0) e paclitaxel (código PDB: 1JFF). . . . . 43
- Figura 7 – Estratégia geral do planejamento de novos candidatos anticâncer desenvolvido neste trabalho, tendo como alvo a polimerização de microtúbulos. 46
- Figura 8 – Representação da interação entre um ligante (uma micromolécula) e um receptor molecular (uma macromolécula) na formação do complexo receptor-ligante. A interação do taxol com seu alvo  $\beta$  tubulina foi utilizada como exemplo. . . . . 48

Figura 9 – (A) Uso da triagem virtual na seleção dos compostos mais promissores para ensaios bioquímicos partindo de uma grande base de dados. Compostos ativos e inativos são representados por esferas vermelhas e azuis, respectivamente. (B) Utilização da docagem molecular na otimização da afinidade de ligação de um composto bioativo. . . . .	50
Figura 10 – Descrição geral da aplicação do método de algoritmo genético pelo programa GOLD na busca conformacional de um complexo receptor-ligante. (A) Cada conformação específica de uma molécula é descrita por uma combinação de coordenadas e diedros e representada em um vetor de forma análoga a um gene. (B) Diversas combinações possíveis são geradas para uma mesma molécula resultando em uma "população genética" de conformações. Esta população é submetida a operações que alteram alguns valores no vetor de forma análoga a operações de mutação e recombinação gênica no interior de células. Este processo é repetido até que uma convergência seja observada para uma única conformação de menor energia. . . . .	52
Figura 11 – Exemplo de um protomol gerado para o sítio de ligação da colchicina no dímero de tubulina na presença de GTP e do íon magnésio (código PDB: 1SA0) durante o processo de docagem do programa Surflex-Dock. Átomos de carbono, nitrogênio e oxigênio no protomol estão representados em preto (bege para o GTP), azul e vermelho respectivamente. Hidrogênios estão representados em branco. Todo o restante da proteína foi omitido para maior clareza. . . . .	53
Figura 12 – Ilustração do método utilizado pelo AutoGrid no cálculo dos mapas de energia durante a docagem molecular no programa AutoDock Vina de um complexo hipotético. O espaçamento da grade foi aumentado para melhor visualização. . . . .	54
Figura 13 – Conformações energéticas consideradas durante o processo de docagem molecular do programa AutoDock. (A) Estados não ligados da proteína e do ligante. (B) Estados ligados da proteína e do ligante. (C) Complexo formado. Possíveis interações inter- e intramoleculares são ilustradas por linhas tracejadas. . . . .	55
Figura 14 – (A) Exemplo de uma proteína solvatada em uma caixa cúbica com moléculas de água. (B) Etapas principais do processo de dinâmica molecular. . . . .	57



Figura 15 – Exemplo de uma equação de um campo de força utilizado para simulações de dinâmica molecular. Os termos ligados correspondem às interações entre átomos ligados através de ligações covalentes. As contribuições das variações de energia devido à alterações na distância ( $r$ ), ângulo ( $\theta$ ) e diedros ( $\phi$ ) em relação aos valores de equilíbrio são determinadas a partir das constantes  $K_r$ ,  $K_\theta$  e  $K_\phi$  ajustadas para reproduzir sistemas reais. O termo não ligado corresponde às interações hidrofóbicas e eletrostáticas entre átomos não ligados diretamente por ligação covalente. . . . . 58

Figura 16 – Estrutura cristalográfica de um dímero formado pela interação entre monômeros de  $\alpha$ - e  $\beta$ -tubulina (código de acesso PDB: 1JFF). As moléculas de GTP e GDP estão apresentadas em verde e magenta, respectivamente. A molécula de paclitaxel (em vermelho) pode ser observada interagindo com seu sítio de ligação na  $\beta$ -tubulina. . . . . 61

Figura 17 – Esquema de polimerização dos microtúbulos. A partir da associação de dímeros de  $\alpha\beta$ -tubulinas um protofilamento é formado. As interações laterais entre os dímeros do protofilamento levam à formação de uma folha de tubulinas. O crescimento dessa folha faz com que a mesma se feche formando um microtúbulos composto por 13 protofilamentos. A presença de moléculas de GTP (esferas verdes) expostas na superfície da subunidade  $\beta$ -tubulina faz com que a extremidade (+) tenha uma maior taxa de adição de novos dímeros. O equilíbrio entre as taxas de polimerização e a despolimerização permite controlar o tamanho dos microtúbulos no interior celular. . . . . 62

Figura 18 – (A) Estrutura molecular do composto fluorescente DAPI (4',6-Diamidino-2-fenilindol) utilizado como repórter no ensaio de detecção de polimerização da tubulina. (B) Descrição do ensaio de polimerização da tubulina baseado em fluorescência em três condições distintas: (i) na ausência de moduladores de microtúbulos (controle); (ii) na presença de paclitaxel, um estabilizador de microtúbulos; (iii) na presença de colchicina, um destabilizador de microtúbulos. As fases de nucleação (I), crescimento (II) e equilíbrio (III) estão destacadas no gráfico. . . . . 64

Figura 19 – (A) Fluorímetro Victor 3 1420, utilizado para monitorar a polimerização de tubulina. (B). Placa de leitura de 96 poços para ensaios de fluorescência. 65

- Figura 20 – (A) Ensaio de polimerização da tubulina baseado em espalhamento de luz, medido pela variação de absorbância em 340 nm. Como no ensaio de fluorescência, observam-se três etapas distintas da polimerização: (I) nucleação, (II) crescimento e (III) equilíbrio. (B) Ensaios de polimerização utilizados como controles: (i) na ausência de ligantes (controle com DMSO 1%); (ii) na presença de paclitaxel, um estabilizador de microtúbulos; (iii) na presença de nocodazol (ou colchicina), um desestabilizador de microtúbulos. . . . . 66
- Figura 21 – (A) Espectrofotômetro SpectraMax Plus 384 utilizado para monitorar a polimerização de tubulina por absorbância (espalhamento de luz). (B). Placa de leitura transparente de 96 poços utilizada para os ensaios de absorbância. . . . . 67
- Figura 22 – (A) Estruturas dos compostos flutax-1 e flutax-2 derivados pela ligação do paclitaxel ao marcador fluorescente fluoresceína e utilizados para monitorar a ligação ao sítio do paclitaxel, responsável por estabilizar os microtúbulos de tubulina. (B) Estrutura do composto BODIPY-vimblastina utilizados para monitorar a ligação ao sítio da vimblastina, responsável por inibir a polimerização de tubulina. . . . . 68
- Figura 23 – (A) Estrutura da colchicina fluorescente utilizada neste trabalho e obtida pela ligação da colchicina ao marcador fluorescente fluoresceína. (B) Espectros de absorbância e emissão da colchicina fluorescente.<sup>142</sup> . . . . . 69
- Figura 24 – (A) Descrição do ensaio competitivo para determinar a interação de um ligante com o sítio da colchicina. (A) Incubação da tubulina com colchicina fluorescente (controle negativo) e com colchicina fluorescente junto ao composto de estudo. (B) Coleta das frações da amostra em uma placa de fluorescência após aplicação em uma coluna cromatográfica de exclusão molecular. Nesta etapa a proteína tubulina ligada à colchicina fluorescente (e possivelmente ao composto) é separada dos outros componentes da solução, sendo eluída nas primeiras frações (C) Medida da emissão em 535 nm das frações coletadas em um fluorímetro. No ensaio padrão, observam-se os picos correspondentes à colchicina fluorescente ligada à tubulina (pico I) e livre em solução (pico II). . . . . 70
- Figura 25 – (A) Ensaio em que o composto de estudo inibidor da polimerização de microtúbulos não interage com o sítio da colchicina, mas se liga a algum outro sítio na tubulina. O sinal da fração de eluição da tubulina é intenso e similar ao do ensaio padrão. (B) Ensaio em que o composto avaliado se liga ao sítio da colchicina, apresentando uma competitividade pelo sítio que leva a uma diminuição do sinal de emissão da colchicina fluorescente na fração de eluição da tubulina, em comparação ao ensaio padrão. . . 71

Figura 26 – Propriedades físico-químicas previstas para os compostos da base de dados inicial sintetizada pelo LSPNF e para os fármacos paclitaxel, vimblastina e colchicina. $clogP$ : coeficiente de partição calculado. TPSA: área de superfície polar topológica (em Å <sup>2</sup> ). MM: massa molecular (em Da). . . . .	73
Figura 27 – Estruturas moleculares dos compostos selecionados a partir dos resultados de triagem virtual por docagem molecular (Tabela 3) para a realização de ensaios bioquímicos de polimerização de tubulina <i>in vitro</i> . . . . .	76
Figura 28 – Curva de concentração-resposta para a colchicina. O valor de IC <sub>50</sub> de 2 μM determinado para os ensaios de polimerização de microtúbulos baseados em fluorescência concorda com os valores reportados na literatura. <sup>144,145</sup> . . . . .	78
Figura 29 – Determinação do parâmetro quantitativo de modulação IC <sub>50</sub> para o derivado indólico <b>8</b> , através de ensaios de polimerização de tubulina na presença de diferentes concentrações do composto. Curvas de polimerização de controle (colchicina e taxol a 20 μM e DMSO) e na presença das seguintes concentrações do composto: 250, 100, 50, 25 e 10 μM. Todas as curvas correspondem a valores médios de triplicatas. A curva concentração-resposta permitiu determinar o valor de IC <sub>50</sub> igual a 25 ± 8 μM. . . . .	78
Figura 30 – Determinação do parâmetro quantitativo de modulação IC <sub>50</sub> para a espirocicloexadienona <b>9</b> , através de ensaios de polimerização de tubulina na presença de diferentes concentrações do composto. Curvas de polimerização de controle (colchicina e taxol a 20 μM e DMSO) e na presença das seguintes concentrações do composto: 500, 250, 100, 50, 25, 10 e 5 μM. Todas as curvas correspondem a valores médios de triplicatas. A curva concentração-resposta permitiu determinar um valor de IC <sub>50</sub> de 16 ± 3 μM. . . . .	79
Figura 31 – Novos compostos propostos para ensaios de modulação da polimerização de tubulina. (A) Derivado indólico <b>novo 1</b> (B) Derivado indólico <b>novo 2</b> . . . . .	80
Figura 32 – Ensaios de modulação da polimerização de tubulina na presença de diferentes concentrações do derivado indólico <b>novo 1</b> mostram que este composto não apresenta atividade considerável de modulação da polimerização de microtúbulos. . . . .	81

- Figura 33 – Determinação dos valores de  $IC_{50}$  para o derivado indólico **novο 2**. (A) Estrutura do derivado indólico **novο 2**. (B) Ensaios de modulação da polimerização de tubulina na presença de diferentes concentrações do derivado indólico **novο 2** mostram que este composto apresenta uma atividade de modulação da polimerização de microtúbulos que obedece a um comportamento do tipo concentração-resposta. (C) A curva concentração-resposta permitiu determinar um valor de  $IC_{50}$  de  $25 \pm 3 \mu M$ . . . . . 82
- Figura 34 – Ensaio de polimerização de tubulina na presença dos compostos derivado indólico **8** e derivado indólico **novο 2**, com concentrações de  $25 \mu M$  cada. As curvas correspondem aos valores médios de triplicatas. . . 82
- Figura 35 – Ensaios de migração de células tumorais (*wound healing*). Os ensaios foram executados na ausência de compostos moduladores de polimerização (controle negativo), na presença de  $50 \mu M$  dos compostos teste, identificados como capazes de inibir a polimerização de tubulina: derivado indólico **8**, derivado indólico **novο 1** e derivado indólico **novο 2**; e na presença de colchicina como controle positivo para inibição. Ensaios executados no LQMC pela aluna de doutorado Wanessa F. Altei. . . . . 84
- Figura 36 – Estudos celulares de imunofluorescência na presença dos derivados indólicos identificados como inibidores da polimerização de microtúbulos. Células cancerígenas MDA-MB-231 (células humanas de câncer de mama) foram tratadas com anticorpos para  $\alpha$ -tubulina (Sigma Aldrich, Saint Louis, MO, Estados Unidos) permitindo observar o citoesqueleto celular (em vermelho) em um microscópio confocal. Para distinguir o núcleo celular (em azul) o composto DAPI (4,6-diamidino-2-fenilindol) foi empregado. Ensaios executados no LQMC pela aluna de doutorado Wanessa F. Altei. . . . . 85
- Figura 37 – (A) Emissão em 535 nm do primeiro pico coletado no ensaio de competição e correspondente à colchicina fluorescente ligada à tubulina. No primeiro ensaio (padrão, em cinza) a incubação foi realizada na presença somente de colchicina fluorescente a  $0,5 \mu M$ . No segundo ensaio (em azul) o mesmo procedimento foi repetido adicionando  $10 \mu M$  de colchicina não fluorescente junto ao marcador fluorescente. No terceiro ensaio (em amarelo) o mesmo ensaio foi repetido na presença de  $25 \mu M$  do derivado indólico **8** junto à colchicina fluorescente. No quarto ensaio (em roxo) o mesmo procedimento foi repetido na presença de  $25 \mu M$  do derivado indólico **novο 2**. (B) Em ambos os ensaios na presença da colchicina e dos derivados indólicos observa-se uma acentuada redução do sinal máximo de emissão, indicando que eles são capazes de se ligar ao sítio da colchicina. . . . . 87

Figura 38 – Complexos preditos pelos estudos de docagem molecular utilizando o programa Autodock para os derivados indólicos <b>8</b> e <b>novos 2</b> no sítio do dímero de $\alpha\beta$ -tubulina (código de acesso PDB: 4O2B). (A) conformação do derivado indólico <b>8</b> ; (B) conformação do derivado indólico <b>novos 2</b> ; (C) sobreposição das conformações obtidas para ambos os derivados indólicos; (D) Comparação da conformações dos derivados indólicos com a conformação da colchicina cristalográfica no sítio de ligação. . . . .	89
Figura 39 – Curvas de RMSD apresentando a variação conformacional do dímero de $\alpha\beta$ -tubulina complexado com os ligantes de estudo durante o processo de simulação por dinâmica molecular. (A) Dímero associado ao ligante colchicina. (B) Dímero associado ao ligante derivado indólico <b>8</b> . (C) Dímero associado ao ligante derivado indólico <b>novos 2</b> . . . . .	92
Figura 40 – Desvios médios quadráticos (RMSDs) dos derivados indólicos e da colchicina em complexo com o dímero $\alpha\beta$ -tubulina. . . . .	93
Figura 41 – Conformações finais obtidas após 10 ns de simulação para os complexos formados pelos ligantes (A) colchicina, (B) derivado indólico <b>8</b> e (C) derivado indólico <b>novos 2</b> . As conformações iniciais de cada ligante são apresentadas para comparação (em cinza claro). . . . .	94
Figura 42 – Abertura do loop T7 da $\beta$ -tubulina na interface dimérica durante simulações de dinâmica molecular. Este movimento foi observado na presença do (A) derivado indólico <b>8</b> e do (B) derivado indólico <b>novos 2</b> . . . . .	95
Figura 43 – Curva de dose resposta e determinação do valor de $IC_{50}$ do derivado indólico <b>novos 6</b> em ensaios de polimerização de microtúbulos baseados em fluorescência. . . . .	99
Figura 44 – (A) Modelo de ligação proposto para o derivado indólico <b>novos 6</b> a partir de estudos de docagem molecular. (B) Comparação do modelo de ligação do derivado indólico <b>novos 6</b> com os derivados indólicos <b>8</b> e <b>novos 2</b> . . . . .	100
Figura 45 – Exemplos de complexos proteicos formados pela associação estável de diversas proteínas. . . . .	105

- Figura 46 – (A) Processo de acoplamento coevolutivo. No exemplo apresentado, um resíduo pertencente a um par estereoquimicamente complementar (esferas amarelas) é mutado para um resíduo de menor volume. Para que a proteína mantenha suas características estruturais e funcionais, o resíduo complementar deve ser mutado (no caso, para um resíduo mais volumoso) a fim de manter a complementariedade do par interagente. (B) Resíduos que encontram-se fisicamente em contato na estrutura tridimensional de uma proteína devem apresentar uma correlação de mutação. Isso pode ser observado analisando um conjunto de sequências de proteínas da mesma família. (C) As correlações entre resíduos podem ser quantificadas por um método de correlação (como por exemplo o DCA) e estes contatos preditos podem ser utilizados para guiar o enovelamento ou mudanças conformacionais de proteínas. . . . . 107
- Figura 47 – Processo de predição do complexo proteico tendo como exemplo a proteína tRNA metiltransferase (código PDB: 1UAL). (A) Os monômeros são separados a uma distância de 50 Å. Um dos monômeros é rotacionado para representar uma situação de orientação aleatória entre as macromoléculas. (B) As distâncias de equilíbrio das interações diméricas provenientes do método de DCA são reduzidas até que seja observada a formação de um complexo estável. . . . . 112
- Figura 48 – Parâmetros ajustáveis em potenciais de modelos baseados em estrutura (SBM). (A) Potencial do tipo Lennard-Jones. (B) Potencial do tipo gaussiano. Os parâmetros ajustáveis são definidos como:  $A$  = amplitude,  $d$  = distância de equilíbrio estável (mínima energia),  $w$  = largura do poço de potencial e  $\rho$  = volume de exclusão atômica. . . . . 115
- Figura 49 – Pares acoplados por co-evolução preditos por DCA e comparação com os mapas de contato dos complexos homodiméricos nativos e preditos por simulação. . . . . 118
- Figura 50 – RMSD das estruturas preditas (em relação aos homodímeros nativos) pelo método de simulação SBM/DCA, ao longo do tempo, em relação à estrutura do homodímero nativo. Diferentes gradientes de cinza representam as diferentes distâncias de equilíbrio e parâmetros das dos potenciais gaussianos utilizados em cada uma das 7 etapas de simulação (descritas na seção 2.2.4). Pode-se observar que todos os sistemas atingem uma conformação bastante estável e de baixo RMSD em todos os casos, indicando a formação de complexos proteicos. . . . . 120

Figura 51 – Estruturas dos complexos homodiméricos preditos pelo método de DCA acoplado ao SBM. Os sistemas escolhidos representam proteínas de diferentes níveis de complexidade: (A) proteínas globulares: proteína da tríade de histidina, aminomutase e GAPDH; (B) proteínas não globulares: tRNA metiltransferase, IspG, isocitrato desidrogenase e glicose 6-fosfato isomerase; (C) proteínas com multidomínios: álcool desidrogenase e cetoacil sintase; e (D) proteínas capazes de dimerizar de diferentes conformações: PhoB e ReX3. . . . .	121
Figura 52 – Comparação entre as estruturas dos complexos homodiméricos preditos para as proteínas reguladoras de resposta (A) PhoB e (B) ReX3. Estas proteínas são capazes de dimerizar de diferentes conformações. . . . .	124
Figura 53 – Estruturas moleculares de alguns fármacos de origem natural. (A) Paclitaxel, utilizado no tratamento de diversos tipos de tumores e isolado do teixo do pacífico. (B) Morfina, um potente anestésico e narcótico proveniente da papoila do ópio. (C) Quinina, um composto com atividades antitérmicas, antimaláricas e analgésicas extraído da árvore cinchona. . . . .	130
Figura 54 – Estruturas moleculares de alguns dos produtos naturais isolados pelo grupo NuBBE. . . . .	130
Figura 55 – Descrição da metodologia utilizada na organização e construção da base de dados NuBBE <sub>DB</sub> . . . . .	132
Figura 56 – Distribuição das fontes de origem dos compostos presentes na base NuBBE <sub>DB</sub> . . . . .	135
Figura 57 – Distribuições das características físico-químicas da base de dados NuBBE <sub>DB</sub> . (A) Massa molecular, em Daltons. (B) Volume molecular, em Å <sup>3</sup> . (C) Número de violações da “Regra dos 5” de Lipinski. (D) Número de ligações rotacionáveis. (E) Coeficiente de partição calculado, <i>clogP</i> . (F) Área de superfície polar topológica calculada, TPSA, em Å <sup>2</sup> . (G) Número de aceptores de ligação de hidrogênio (H) Número de doadores de ligação de hidrogênio. . . . .	136
Figura 58 – Apresentação da página de acesso da base de dados NuBBE <sub>DB</sub> . (A) Opção de busca de compostos por informações ou propriedades físico-químicas (B) Opção de busca pelo desenho de estruturas químicas (C) Exemplo de resultado de uma busca de compostos na NuBBE <sub>DB</sub> . (D) Exemplo das informações disponibilizadas de um composto da base. . . . .	138
Figura 59 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: aciloínas. . . . .	165
Figura 60 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: adutos de Morita-Baylis-Hilman - parte 1. . . . .	166
Figura 61 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: adutos de Morita-Baylis-Hilman - parte 2. . . . .	167

Figura 62 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: adutos de Morita-Baylis-Hilman - parte 3. . . . .	168
Figura 63 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: adutos de Morita-Baylis-Hilman - parte 4. . . . .	169
Figura 64 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: amino-alcoóis. . . . .	170
Figura 65 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: $\beta$ -cetoésteres - parte 1. . . . .	171
Figura 66 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: $\beta$ -cetoésteres - parte 2. . . . .	172
Figura 67 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: derivados ciclopenta- $\beta$ -indólicos. . . . .	173
Figura 68 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: espirocicloexadienonas. . . . .	174
Figura 69 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: ftalídeos. . . . .	175
Figura 70 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: naftiridinas. . . . .	176
Figura 71 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: pirrozilidinas. . . . .	177
Figura 72 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: derivados de benzofurano, quinolina e aminoicosanóis. . . . .	178
Figura 73 – Triagem bioquímica inicial dos compostos selecionados por triagem virtual frente a microtúbulos: (A) Ensaio controle. (B-H) derivados indólicos.	181
Figura 74 – Triagem bioquímica inicial dos compostos selecionados por triagem virtual frente a microtúbulos: (A) Ensaio Controle. (B-C) Naftiridinas. (D-F) Espirocicloexadienonas. (G) $\beta$ -cetoéster. (H) Espisulosina. . . . .	182
Figura 75 – Curvas de RMSD para os estudos de dinâmica molecular para o sistema contendo colchicina em complexo com o dímero de tubulina. O perfil médio de 3 simulações é apresentado em vermelho e o desvio padrão em cinza. (A) Dímero de tubulina. (B) Colchicina. (C) Co-fator GTP. (D) Co-fator $Mg^{2+}$ . . . . .	185
Figura 76 – Curvas de RMSD para os estudos de dinâmica molecular para o sistema contendo o derivado indólico <b>8</b> em complexo com o dímero de tubulina. O perfil médio de 3 simulações é apresentado em vermelho e o desvio padrão em cinza. (A) Dímero de tubulina. (B) Derivado indólico <b>8</b> . (C) Co-fator GTP. (D) Co-fator $Mg^{2+}$ . . . . .	186



Figura 77 – Curvas de RMSD para os estudos de dinâmica molecular para o sistema contendo o derivado indólico **novoo 2** em complexo com o dímero de tubulina. O perfil médio de 3 simulações é apresentado em vermelho e o desvio padrão em cinza. (A) Dímero de tubulina. (B) Derivado indólico **novoo 2**. (C) Co-fator GTP. (D) Co-fator  $Mg^{2+}$ . . . . . 187



# Lista de tabelas

Tabela 1 – Exemplos de algumas classes estruturais sintetizadas e disponibilizadas pelo LSPNF. . . . .	47
Tabela 2 – Características dos ensaios de polimerização de tubulina baseados em fluorescência e espalhamento. . . . .	66
Tabela 3 – Classificação dos 10 compostos que apresentaram melhor pontuação consenso durante o processo de docagem molecular no sítio da colchicina utilizando os 3 programas de docagem molecular: GOLD, Auto-Dock Vina e Surflex-Dock. . . . .	75
Tabela 4 – Atividade dos compostos selecionados por triagem virtual em modular a polimerização de microtúbulos em ensaios bioquímicos. . . . .	77
Tabela 5 – Predição dos valores de coeficiente de partição ( $\log P$ ) para os derivados indólicos. . . . .	80
Tabela 6 – Avaliação da citotoxicidade dos derivados indólicos de estudo em células normais e cancerígenas. . . . .	86
Tabela 7 – Resultados dos estudos de docagem molecular das formas enantioméricas dos derivados indólicos <b>8</b> e <b>novo 2</b> . . . . .	90
Tabela 8 – Novos derivados indólicos planejados a partir dos estudos de modelagem molecular. . . . .	97
Tabela 9 – Ensaios baseados em fluorescência da polimerização de microtúbulos com os novos derivados indólicos planejados. . . . .	98
Tabela 10 – Sistemas selecionados para predição de homodímeros utilizando os métodos de SBM e DCA. . . . .	111
Tabela 11 – Parâmetros da função gaussiana utilizados ao longo do processo de simulação na formação de complexos homodiméricos. . . . .	116
Tabela 12 – Valores de RMSD e iRMSD obtidos para as estruturas moleculares preditas para os sistemas de estudo. . . . .	123
Tabela 13 – Ferramentas de Desenvolvimento. . . . .	157
Tabela 14 – Ferramentas de Elaboração de Texto. . . . .	157
Tabela 15 – Ambientes Operacionais. . . . .	158
Tabela 16 – Programas de Análise de Sistemas Moleculares. . . . .	158
Tabela 17 – Programas de Docagem e Dinâmica Molecular. . . . .	158
Tabela 18 – Programas de Desenho Bidimensional e Análise de Propriedades Físico Químicas de Pequenas Moléculas. . . . .	159
Tabela 19 – Programas Desenvolvidos durante o Projeto. . . . .	159
Tabela 20 – Programas Utilizados para Elaborar e Analisar os Gráficos Apresentados neste Trabalho. . . . .	159



# Lista de Abreviaturas e Siglas

BODIPY	Boro-dipirrometano
$\log P$	Logaritmo do Coeficiente de Partição Calculado
CTBP	Center for Theoretical Biological Physics
DAPI	4',6-Diamidino-2-fenilindol (Fluoróforo)
DCA	Direct Coupling Analysis (Análise de acoplamento direto)
DMSO	Solvente Dimetilsulfóxido
EGTA	Ácido etileno glicol-bis ( $\beta$ -amino-etil-éter) N,N,N',N'- tetra-acético
INCA	Instituto Nacional do Câncer
LBDD	Ligand-Based Drug Design (Planejamento baseado na estrutura do ligante)
LQMC	Laboratório de Química Medicinal e Computacional
LSPNF	Laboratório de Síntese de Produtos Naturais e Fármacos
NuBBE	Núcleo de Bioensaios, Biossíntese e Ecofisiologia de Produtos Naturais
NuBBE <sub>DB</sub>	NuBBE Database (Base de Dados NuBBE de Produtos Naturais)
PDB	Protein Data Bank (Banco de Dados de Proteínas)
PIPES	Piperazina-1,4-bis (Ácido 2-etanosulfídico)
QSAR	Quantitative Structure-Activity Relationships (Relações quantitativas entre estrutura e atividade)
RMSD	Root Mean Square Deviation (desvio quadrático médio)
SBDD	Structure-Based Drug Design (Planejamento baseado na estrutura do receptor)
TPSA	Topological Polar Surface Area (Área de Superfície Polar Topológica)
WHO	World Health Organization (Organização Mundial de Saúde )



# Sumário

<b>1</b>	<b>IDENTIFICAÇÃO DE NOVOS MODULADORES DE POLIMERIZAÇÃO DA TUBULINA</b>	<b>33</b>
<b>1.1</b>	<b>Introdução</b>	<b>35</b>
1.1.1	Planejamento Racional de Fármacos	35
1.1.2	Câncer	39
1.1.3	Microtúbulos como Alvo Molecular contra o Câncer	42
<b>1.2</b>	<b>Motivações e objetivos</b>	<b>44</b>
<b>1.3</b>	<b>Metodologia</b>	<b>44</b>
1.3.1	Estratégias de Planejamento	44
1.3.2	Métodos Computacionais	45
1.3.2.1	Geração da Base de Dados de Compostos Sintéticos	45
1.3.2.2	Docagem Molecular	48
1.3.2.2.1	O Programa GOLD	51
1.3.2.2.2	Docagem Molecular com o Surflex-Dock	52
1.3.2.2.3	Docagem Molecular com o AutoDock Vina	53
1.3.2.3	Triagem Virtual e Seleção de Compostos para Avaliação Bioquímica	54
1.3.2.4	Dinâmica Molecular	56
1.3.2.4.1	Princípios da Dinâmica Molecular	56
1.3.2.4.2	O Programa GROMACS	58
1.3.2.5	Estudos de Dinâmica Molecular com Inibidores da Polimerização de Microtúbulos	59
1.3.3	Métodos Experimentais	60
1.3.3.1	Polimerização de Microtúbulos	60
1.3.3.2	Quantificação de Polimerização da Tubulina por Fluorescência	63
1.3.3.3	Ensaio de Polimerização de Tubulina utilizando Fluorescência	63
1.3.3.4	Quantificação de Polimerização da Tubulina por Espalhamento	65
1.3.3.5	Ensaio de Polimerização de Tubulina utilizando Espalhamento	66
1.3.3.6	Identificação do Mecanismo Molecular de Compostos Moduladores de Tubulina	68
1.3.3.7	Identificação do Mecanismo Molecular de Modulação de Microtúbulos	70
<b>1.4</b>	<b>Resultados e Discussão</b>	<b>72</b>
1.4.1	Avaliação das Propriedades Físico-Químicas da Coleção de Compostos	72
1.4.2	Triagem Virtual de Compostos para Avaliação Bioquímica	74
1.4.3	Triagem Inicial de Moduladores de Polimerização de Microtúbulos através de Ensaio de Polimerização baseados em Fluorescência	74
1.4.4	Avaliação Quantitativa de Inibidores da Polimerização de Microtúbulos por meio de Ensaio Baseados em Fluorescência	77

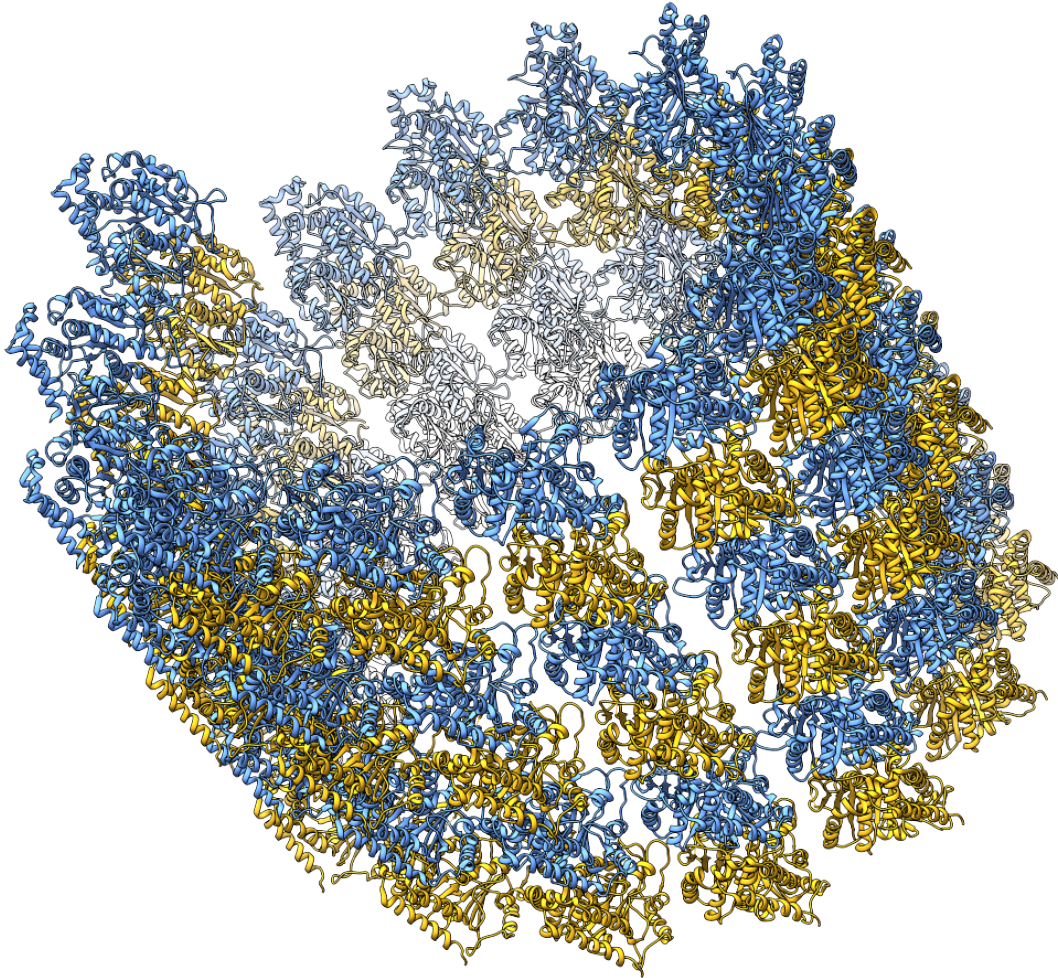
1.4.5	Planejamento de Novos Derivados Indólicos . . . . .	79
1.4.6	Ensaio de Polimerização baseados em Espalhamento . . . . .	81
1.4.7	Estudos de Migração Celular . . . . .	83
1.4.8	Estudos de Imunofluorescência com Microtúbulos . . . . .	83
1.4.9	Avaliação da Citotoxicidade dos Derivados Indólicos . . . . .	86
1.4.10	Estudos de Competição pelo Sítio da Colchicina . . . . .	87
1.4.11	Docagem Molecular dos Derivados Indólicos no Sítio da Colchicina . . . . .	88
1.4.12	Estudos de Dinâmica Molecular com os Derivados Indólicos . . . . .	91
1.4.13	Planejamento, Síntese e Avaliação de Novos Derivados Indólicos . . . . .	96
<b>1.5</b>	<b>Conclusão . . . . .</b>	<b>100</b>
<b>2</b>	<b>MODELAGEM BASEADA EM ESTRUTURA E ANÁLISE DE ACOPLAMENTO DIRETO NA PREDIÇÃO DE SISTEMAS DIMÉRICOS MACROMOLECULARES . . . . .</b>	<b>103</b>
<b>2.1</b>	<b>Introdução . . . . .</b>	<b>105</b>
2.1.1	Oligomerização em Sistemas Biológicos . . . . .	105
2.1.2	Predição de Interações entre Resíduos por Análise de Acoplamento Direto	106
2.1.3	Predição de Complexos Proteicos Utilizando Sinais de Coevolução em Simulações de Dinâmica Molecular com Modelos Baseados em Estrutura	108
<b>2.2</b>	<b>Motivações e objetivos . . . . .</b>	<b>109</b>
<b>2.3</b>	<b>Metodologia . . . . .</b>	<b>110</b>
2.3.1	Análise de Correlação Evolutiva de Sequências Proteicas . . . . .	110
2.3.2	Modelagem Estrutural de Sistemas Proteicos . . . . .	112
2.3.3	Modelagem Baseada em Estrutura Utilizando Potenciais Gaussianos . . . . .	113
2.3.4	Dinâmica Molecular . . . . .	115
<b>2.4</b>	<b>Resultados e Discussão . . . . .</b>	<b>116</b>
2.4.1	Estudos de Correlação Evolutiva em Sequências Proteicas . . . . .	116
2.4.2	Modelagem de Sistemas Proteicos Aplicando Modelos Baseados em Estrutura e Potenciais Gaussianos . . . . .	119
2.4.3	Identificação das Diferentes Interfaces de Proteínas Reguladoras de Resposta . . . . .	122
<b>2.5</b>	<b>Conclusão . . . . .</b>	<b>125</b>
<b>3</b>	<b>CONSTRUÇÃO DE UMA BASE DE DADOS VIRTUAL DE PRODUTOS NATURAIS . . . . .</b>	<b>127</b>
<b>3.1</b>	<b>Introdução . . . . .</b>	<b>129</b>
3.1.1	A importância dos Produtos Naturais no Desenvolvimento de Novos Fármacos . . . . .	129
3.1.2	A Biodiversidade no Brasil e os Compostos Identificados pelo NuBBE . . . . .	129
<b>3.2</b>	<b>Motivações e objetivos . . . . .</b>	<b>131</b>



<b>3.3</b>	<b>Metodologia</b> . . . . .	<b>131</b>
3.3.1	Construção de uma Base Virtual de Produtos Naturais de Livre Acesso	131
3.3.2	Organização dos Compostos e Construção das Estruturas Moleculares 3D . . . . .	133
3.3.3	Predição de Propriedades Físico-Químicas . . . . .	133
3.3.4	Montagem do Site de Busca NuBBE . . . . .	134
<b>3.4</b>	<b>Resultados e Discussão</b> . . . . .	<b>134</b>
3.4.1	Caracterização do Site de Busca NuBBE de Produtos Naturais . . . . .	134
<b>3.5</b>	<b>Conclusão</b> . . . . .	<b>138</b>
<b>4</b>	<b>ARTIGOS PUBLICADOS OU SUBMETIDOS EM PERIÓDICOS CIENTÍFICOS.</b> . . . .	<b>139</b>
<b>4.1</b>	<b>Artigos publicados</b> . . . . .	<b>139</b>
<b>4.2</b>	<b>Artigos submetidos para publicação.</b> . . . .	<b>140</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>141</b>
	<b>APÊNDICE – Recursos computacionais</b> . . . . .	<b>155</b>
	<b>ANEXOS</b>	<b>161</b>
	<b>ANEXO A – Estruturas Moleculares dos Compostos Avaliados na Identificação de Novos Moduladores de Polimerização de Microtúbulos.</b> . . . . .	<b>163</b>
	<b>ANEXO B – Triagem bioquímica inicial dos compostos selecionados por triagem virtual.</b> . . . . .	<b>165</b>
	<b>ANEXO C – Desvios quadráticos médios das simulações de dinâmica molecular com ligantes no sítio da colchicina do dímero de tubulina.</b> . . . . .	<b>181</b>



# 1 Identificação de Novos Moduladores de Polimerização da Tubulina





## 1.1 Introdução

### 1.1.1 Planejamento Racional de Fármacos

Os inúmeros avanços científicos e tecnológico alcançados nas últimas décadas têm contribuído significativamente para várias áreas que tem a inovação como elemento fundamental. O desenvolvimento de novos métodos experimentais e computacionais tem proporcionado uma revolução em diversos campos da pesquisa moderna, auxiliando o ser humano a modificar o ambiente ao seu redor.<sup>1,2</sup> Dentre os campos de pesquisa mais promissores, destaca-se o de descoberta de novos fármacos.<sup>1</sup>

Nos programas de pesquisa e desenvolvimento (P&D) de fármacos das indústrias farmacêuticas modernas, os modelos tradicionais de pesquisa básica, fundamentados historicamente em abordagens do tipo “tentativa e erro”, deram lugar a estratégias modernas avançadas. Estas estratégias buscam otimizar o processo de descoberta, diminuindo o tempo e os investimentos necessários para que um novo fármaco possa chegar ao mercado.<sup>1,3-8</sup>

Ao invés de reunir fisicamente uma grande biblioteca de compostos naturais ou sintéticos e testá-los aleatoriamente contra diversos alvos terapêuticos (em um processo árduo, demorado e de elevado custo), o desenvolvimento racional de fármacos propõe desenhar e planejar moléculas específicas para interagir com um alvo molecular de estudo. Esta abordagem utiliza uma combinação multidisciplinar de especialidades que permitem identificar e caracterizar um novo alvo molecular (por técnicas de biologia molecular e estrutural), desenhar e prever as interações de pequenas moléculas com estes alvos (por métodos computacionais de modelagem molecular) e avaliar a atividade destes compostos em ensaios bioquímicos.<sup>7-11</sup>

Ao longo da história, os medicamentos têm desempenhado um papel essencial para o aumento da expectativa e qualidade de vida humana.<sup>12-15</sup> Nos últimos 150 anos, a expectativa de vida média praticamente duplicou de 40 para 80 anos em países desenvolvidos. Sem dúvida, um dos principais motivos para esse alcance foi a descoberta dos antibióticos, vacinas e de novos tratamentos eficazes para diversas doenças de elevada mortalidade.<sup>14-17</sup> Juntamente com o aumento da expectativa de vida, o número de fármacos que chegaram ao mercado tem crescido significativamente (Figura 1).<sup>18-20</sup>

A descoberta e o desenvolvimento moderno de novos fármacos fundamenta-se no conhecimento prévio do processo fisiopatológico envolvido na doença em estudo e na seleção de um alvo molecular adequado para ter sua função modulada, proporcionando a resposta farmacológica e o efeito terapêutico desejados. Neste processo, a combinação de diversas especialidades aplicadas em conjunto a estudos de química medicinal é essencial. Esta multiplicidade se apresenta desde a descoberta e caracterização do alvo molecular por

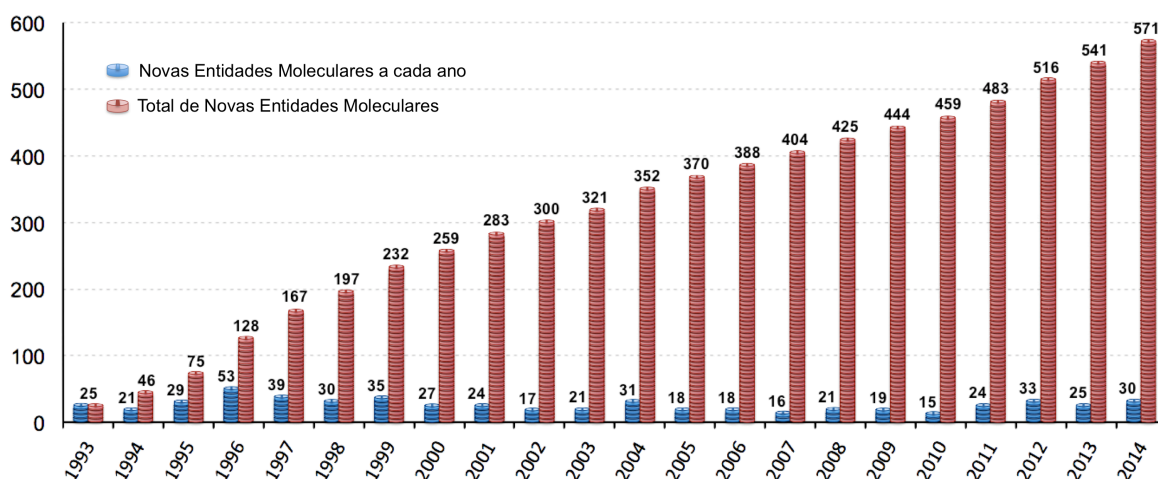


Figura 1 – Número de novas entidade moleculares (NEM) regulamentadas nos Estados Unidos ao longo dos últimos anos desde o ano de 1993.<sup>18</sup>

Fonte: Elaborada pelo autor.

estudos na área de biologia molecular e a determinação da estrutura molecular por biologia estrutural até o planejamento de inibidores por métodos de modelagem molecular em conjunto com estratégias de síntese orgânica e avaliação experimental.

O processo de desenvolvimento de novos fármacos leva geralmente mais de 10 anos e envolve investimentos da ordem de algumas centenas de milhões de dólares.<sup>21–24</sup> A Figura 2 ilustra algumas das principais etapas envolvidas neste complexo processo.

Após a validação e caracterização do alvo molecular (*e.g.* enzima, receptor celular), várias estratégias de química medicinal podem ser empregadas. Uma delas é a seleção e organização de uma bases de dados de compostos com propriedades físico-químicas adequadas, por meio da aplicação de filtros moleculares, como o cálculo da massa e volume molecular, coeficiente de partição (ou lipofilia,  $\log P$ ) e área de superfície polar (PSA, do inglês *polar surface area*). A aplicação destes filtros moleculares é uma etapa fundamental para excluir compostos com alta probabilidade de apresentar propriedades farmacocinéticas incompatíveis com o perfil de um fármaco. Por exemplo, moléculas muito apolares ou polares, ou que possuam grupos quimicamente reativos, mesmo que apresentem elevada afinidade e potência frente ao alvo molecular de interesse. Também devem ser considerados fatores como instabilidade metabólica, baixa biodisponibilidade, interação com proteínas plasmáticas, toxicidade *in vivo*, entre outros.<sup>25–27</sup>

A partir de uma coleção de compostos selecionados, diversas técnicas de planejamento baseadas na estrutura do ligante (LBDD, do inglês *ligand-based drug design*) e na estrutura do receptor (SBDD, do inglês, *structure-based drug design*) podem ser empregadas para identificar as moléculas mais promissoras para interagir com o alvo e apresentar a atividade desejada em ensaios biológicos.<sup>11,28</sup>

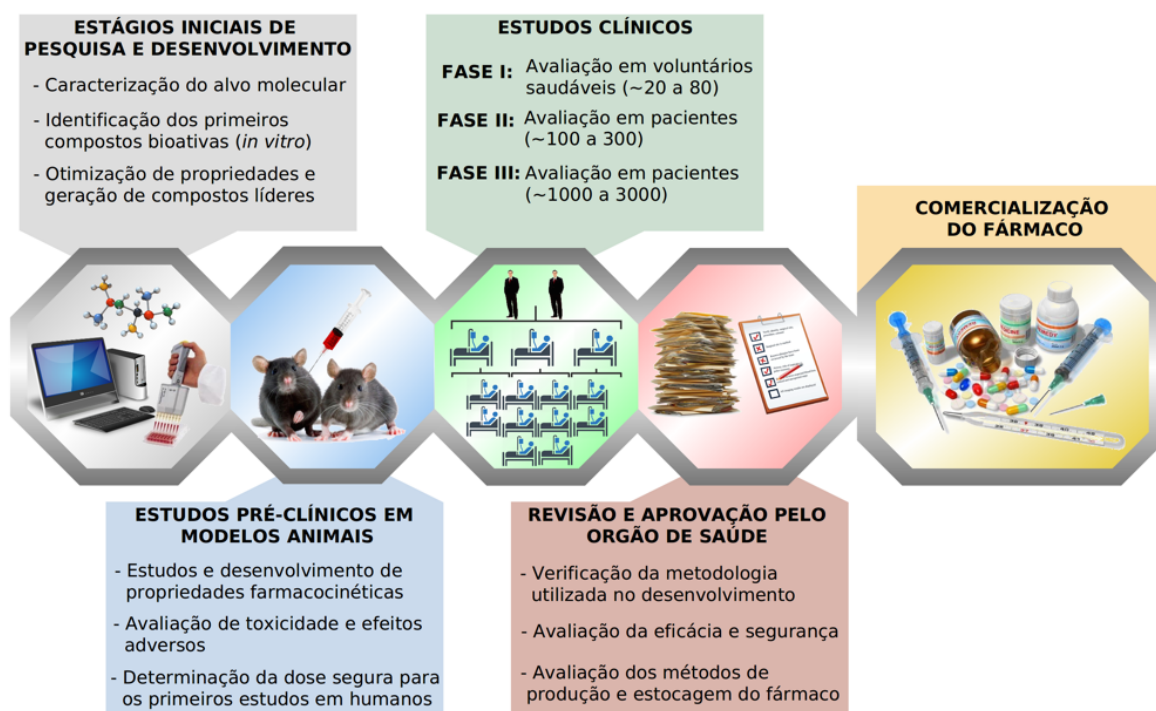


Figura 2 – Etapas envolvidas no desenvolvimento racional de um novo fármaco nas indústrias farmacêuticas modernas.

Fonte: Elaborada pelo autor.

Quando a estrutura tridimensional da macromolécula alvo é conhecida (determinada geralmente através de métodos experimentais, como cristalografia de raios-X e ressonância magnética nuclear), o SBDD encontra vasta aplicação com poderosas ferramentas computacionais de modelagem molecular. Neste caso, triagens virtuais podem ser desenvolvidas para investigar o modo de ligação e a complementariedade de moléculas frente ao sítio de ligação de um alvo molecular específico. Cada estrutura dentre um conjunto de moléculas pode ser encaixada na estrutura tridimensional do receptor em um processo iterativo de busca conformacional e predição de energias de interação, conhecido como docagem molecular (do inglês *molecular docking*). Aplicado repetitivamente para uma grande quantidade de moléculas em um mesmo alvo molecular, este método é bastante útil para selecionar dentre um grande conjunto de moléculas apenas aquelas com maior probabilidade de interagirem com o alvo, representando a fase farmacodinâmica de interação fármaco-receptor. Esta estratégia é conhecida como triagem virtual (também conhecida como VS, do inglês *virtual screening*).<sup>28,29</sup> A aplicação desta metodologia permite uma grande economia de tempo e uma redução substancial de custos. Por isso, é amplamente aplicada pelas grandes empresas farmacêuticas e centros de pesquisa em projetos de química medicinal e descoberta de fármacos.<sup>9,28,30-32</sup>

Após a seleção dos compostos com maior probabilidade de apresentar atividade biológica, estudos experimentais *in vitro* são realizados frente ao alvo molecular para confirmar

as predições computacionais. Outra importante estratégia de triagem é o desenvolvimento de ensaios miniaturizados e automatizados em larga escala, utilizando concentrações únicas dos compostos, em um processo conhecido como triagem em larga escala (HTS, do inglês *high-throughput screening*). O efeito biológico a ser medido deve ser uma propriedade facilmente detectável como, por exemplo, o crescimento celular ou a atividade catalítica de uma enzima. Os compostos com atividade biológica confirmada são submetidos a ensaios mais detalhados avaliando ainda o efeito *in vitro* sob diversas concentrações. Estes estudos buscam determinar quantitativamente a potência dos ligantes e muitas vezes permitem elucidar o mecanismo molecular envolvido no processo de modulação (*e.g.*, a inibição competitiva de uma enzima pela ligação do composto no mesmo sítio do substrato natural). Parâmetros quantitativos comumente utilizados são a  $IC_{50}$  (concentração de composto necessária para reduzir em 50% a atividade biológica monitorada) e o  $K_d$  (constante de dissociação do complexo receptor-ligante). Compostos bioativos e caracterizados quantitativamente quanto ao perfil farmacodinâmico podem ser utilizados como modelos iniciais para modificações estruturais e estudos de LBDD ou SBDD, gerando novas séries de compostos para serem avaliados.<sup>9, 28, 30, 33</sup>

Quando um número considerável de compostos de uma mesma classe já foram avaliados, a comparação entre suas características estruturais e atividades observadas permite identificar grupamentos químicos importantes para a interação do ligante como o receptor molecular. Este processo pode ser executado qualitativamente por estudos das relações entre estrutura e atividade (SAR, do inglês *structure-activity relationship*). Uma vez que o conjunto de compostos avaliados apresenta uma ampla faixa de atividade biológica e variações estruturais, estas relações podem ser determinadas numericamente por estudos quantitativos de estrutura-atividade (QSAR, do inglês *quantitative structure-activity relationships*). Os modelos de correlação gerados através destes estudos permitem racionalizar o planejamento de novos compostos mais potentes e seletivos, na busca de um composto líder com perfil farmacodinâmico necessário para avançar a etapas posteriores no desenvolvimento de fármacos.<sup>34-38</sup>

Uma vez que um composto ativo apresenta propriedades farmacodinâmicas satisfatórias, ele é submetido a estudos *in vivo* a fim de avaliar suas propriedades farmacocinéticas e a toxicidade em modelos animais (fase pré-clínica, Figura 2). Estes estudos envolvem a indução da condição clínica no animal e permitem avaliar vários efeitos dos compostos no organismo (*e.g.*, absorção, metabolismo, taxa de eliminação) e a capacidade do mesmo em eliminar os sintomas da doença. Caso os resultados sejam satisfatórios, novos testes são realizados em organismos cada vez mais complexos e semelhantes ao humano, até que a molécula entre em fase clínica (Figura 2). Compostos com baixa eficácia *in vivo* retornam novamente para estudos de modificações estruturais a fim de melhorar suas propriedades farmacocinéticas.<sup>39-41</sup> Estas etapas de testes *in vivo* são geralmente as que demandam maior tempo de pesquisa e que envolvem os maiores investimentos. Porém, uma vez que um fármaco consiga transpor as fases clínicas, apresentando eficácia e segurança, poderá ser sub-



metido à aprovação do órgão de saúde responsável e, caso seja aprovado, entrar finalmente em comercialização (Figura 2).

### 1.1.2 Câncer

Câncer é a denominação para um grupo de doenças devastadoras caracterizadas pelo crescimento e multiplicação descontrolada de células anormais, capazes de invadir estruturas próximas e se espalharem para diversas regiões do organismo. Estas células cancerígenas deixam de responder aos mecanismos de controle de multiplicação e crescimento celular, duplicando-se desordenadamente com elevada velocidade e danificando o organismo (Figura 3A).<sup>42-44</sup> Além de afetar tecidos adjacentes, alguns tipos de células cancerígenas podem desenvolver a habilidade de se desprender do tumor inicial e se deslocar para diferentes regiões do organismo, gerando tumores secundários (*e.g.*, através do deslocamento pelo sistema circulatório, Figura 3A). Este processo, conhecido como metástase, confere ao câncer uma agressividade muito maior, uma vez que resulta em um número elevado de tumores no organismo e na deslocalização da doença.<sup>42,45,46</sup>

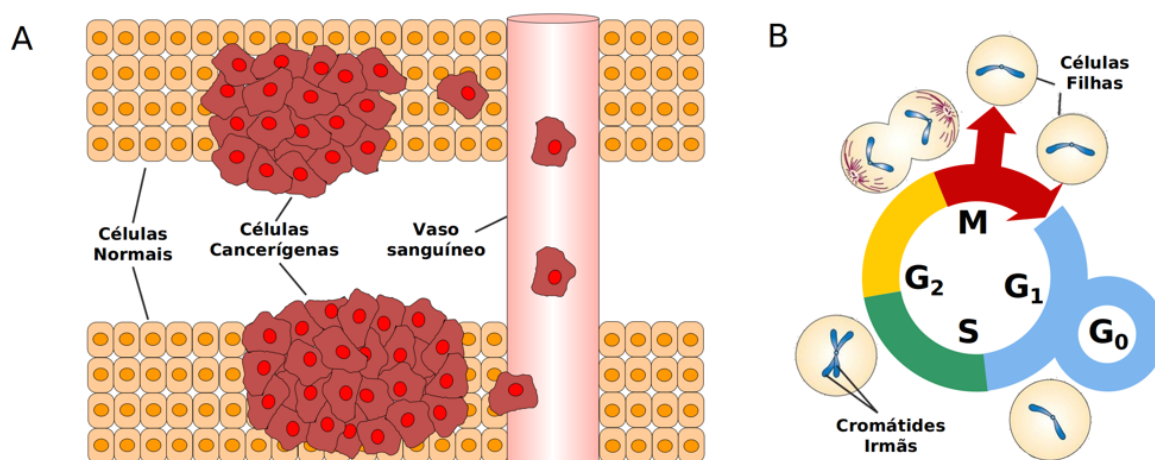


Figura 3 – (A) Processo de formação de um tumor a partir da divisão de células cancerígenas. Em alguns tipos de tumores, as células podem migrar para diferentes regiões do organismo formando tumores secundários, em um processo de metástase. (B) Fases da divisão celular. Células cancerígenas apresentam alterações em mecanismos essenciais na execução de processos biomoleculares, como a verificação e a manutenção do material genético ao longo do ciclo celular.

Fonte: Elaborada pelo autor.

O surgimento de células cancerígenas (ou tumores) é acompanhado por alterações genéticas no em regiões do DNA que codificam proteínas responsáveis pela execução de etapas essenciais na manutenção do ciclo celular (Figura 3B). O ciclo de crescimento e divisão celular é composto por 4 etapas caracterizadas pela predominância de processos metabólicos específicos: G<sub>1</sub> (intervalo 1), S (síntese), G<sub>2</sub> (intervalo 2) e M (mitose). As transições em

cada uma destas etapas são decididas de acordo com o balanço entre sinais químicos que promovem e inibem o crescimento celular. Na fase  $G_1$ , as células encontram-se em constante crescimento em tamanho e se preparando para copiar seu DNA (Figura 3B). Em seguida, ao entrar na fase S, elas promovem a síntese e duplicação do material genético. Após o DNA ser duplicado, as células entram em uma etapa conhecida como  $G_2$ , em que se preparam para o processo de divisão celular. Durante a fase  $G_2$ , as células realizam procedimentos de checagem para avaliar a qualidade do material genético copiado e executar reparos. Finalmente, após esta etapa de checagem, as células finalizam seu processo de divisão e cada uma das células filhas reinicia todo o processo descrito na Figura 3B. Por outro lado, células de tecidos que não precisam estar em constante multiplicação no organismo (*e.g.*, neurônios e células musculares cardíacas) progridem de  $G_1$  para um estado de dormência denominado  $G_0$ , ao invés de progredirem no ciclo de divisão celular (Figura 3B).<sup>42,47</sup>

A progressão a partir das etapas  $G_1$  e  $G_2$  é crucial na decisão do caminho que as células devem seguir. Caso um problema sério seja identificado em uma destas etapas, mecanismos celulares induzem as células a se autodestruírem (em um processo denominado apoptose), a fim de garantir a integridade do organismo. Para uma célula cancerígena conseguir se estabelecer ela precisa evitar estes pontos de checagem que atuam como barreiras de segurança, garantindo o perfeito funcionamento do metabolismo celular. Células cancerígenas apresentam mutações em genes responsáveis por gerar diversas proteínas envolvidas nos mecanismos de checagem (proto-oncogenes) e de diferenciação celular e, dessa forma, conseguem contornar a apoptose e se estabelecer como um tumor. Como consequência, células tumorais geralmente apresentam características metabólicas particulares, como por exemplo, níveis de sinalização anormais, insusceptibilidade a sinais de inibição de crescimento celular, habilidade de promover a formação de vasos sanguíneos (angiogênese) e capacidade ilimitada de divisão e migração celular entre tecidos.<sup>42-44</sup>

As mutações responsáveis por promover o surgimento de câncer em um organismo podem ter origem internas (*e.g.* pré-disposição devido à defeitos genéticos propagados entre gerações) e externas.<sup>48,49</sup> Diversos fatores externos podem gerar mutações no material genético das células e promover o surgimento de tumores. Exemplos deste fatores são a exposição a agentes químicos e alimentos cancerígenos, a radiação e a agentes infecciosos específicos.<sup>49-52</sup>

Apesar de todos os esforços em pesquisa ao longo das últimas décadas, o câncer é uma das doenças globais de maior mortalidade, sendo uma das mais temidas no mundo moderno. De acordo com a Organização Mundial de Saúde (WHO, do inglês *world health organization*), foram registrados aproximadamente 14 milhões de novos casos de câncer em todo o mundo em 2012, resultando na morte de 8,2 milhões de pessoas no mesmo ano.<sup>53</sup> Além disso, é estimado que a taxa de novos casos aumente em 70% ao longo das próximas duas décadas. Segundo os dados do Instituto Nacional do Câncer (INCA), aproximada-

mente 600 mil novos casos de câncer são diagnosticados no Brasil a cada ano, sendo a segunda maior causa de morte por doença no país.<sup>54</sup> De acordo com dados estimados para o ano de 2015 e que acompanham uma tendência mundial, o tipo de câncer com maior incidência no país será o de pele do tipo não melanoma (182 mil novos casos), seguido pelo câncer de próstata (69 mil) e de mama (57 mil); que atingem as populações masculina e feminina, respectivamente.<sup>54</sup> Em países desenvolvidos o câncer também é uma das principais causas de morte. Por exemplo, nos Estados Unidos ela representa a segunda causa mais comum, com uma estimativa de cerca de 1,66 milhões de novos casos e de 850 mil mortes em 2015.<sup>55</sup> Além disso, a incidência varia de acordo com o tipo de câncer, gênero e etnia da população (Figura 4).<sup>55</sup>

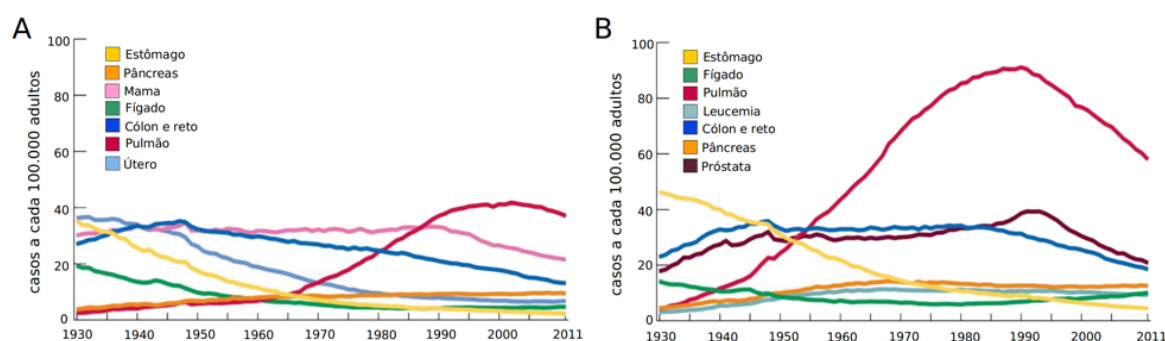


Figura 4 – (A) Incidência de diferentes tipos de câncer em mulheres nos Estados Unidos entre 1930 e 2011. (B) Incidência de diferentes tipos de câncer em homens nos Estados Unidos entre 1930 e 2011.<sup>55</sup>

Fonte: Elaborada pelo autor.

Considerando-se o tamanho, o estágio e a localização primária do câncer, o tratamento pode envolver a remoção cirúrgica do tumor, a aplicação de radioterapia, quimioterapia, terapia hormonal, terapia biológica orientada (anticorpos monoclonais), ou mesmo combinações específicas desses métodos.<sup>42,43</sup> Em pacientes com diagnóstico precoce de câncer e que possuem tumores localizados, é possível utilizar tratamentos localizados como a remoção cirúrgica ou a radioterapia. Entretanto, quando há a presença de tumores secundários em diferentes regiões do organismo (um processo conhecido como metástase), o tratamento quimioterápico é frequentemente o mais utilizado. Diversos fármacos estão disponíveis para o tratamento do câncer, atuando através de mecanismos moleculares distintos e variando quanto à eficácia de acordo com o tipo de câncer e o alvo molecular.<sup>56,57</sup> Alguns exemplos podem ser citados, como o fármaco imatinib, um inibidor da enzima tirosina quinase; o tamoxifeno, um antagonista capaz de se ligar ao receptor de estrogênio inibindo o crescimento celular; a flutamida, um antagonista de andrógeno e os fármacos de origem natural paclitaxel e vimblastina, que interferem no ciclo de divisão celular ao modular a dinâmica de polimerização de microtúbulos.<sup>58–61</sup>

### 1.1.3 Microtúbulos como Alvo Molecular contra o Câncer

Dentre as diversas proteínas estruturais encontradas em eucariotos, os microtúbulos desempenham papel essencial no desenvolvimento e manutenção da estrutura celular, no transporte de diversos componentes celulares (*e.g.* vesículas e mitocôndrias), nas vias de sinalização e no processo de mitose.<sup>62</sup> Os microtúbulos são formados pela polimerização e despolimerização das isoformas  $\alpha$  e  $\beta$  da proteína tubulina. Essenciais no processo de divisão celular, os microtúbulos são considerados um dos mais importantes alvos moleculares para o tratamento do câncer. Desde a descoberta do paclitaxel em 1971 (Figura 5A, comercializado como Taxol<sup>®</sup> pela empresa farmacêutica Bristol-Myers Squibb), um dos fármacos mais importantes para o tratamento do câncer até os dias de hoje, diversas classes de compostos que interferem com o processo de divisão celular pela alteração da dinâmica de formação de microtúbulos foram descobertos, variando quanto à forma de modulação e quanto aos sítios de ligação às unidades de tubulina.<sup>61,63,64</sup>

O princípio de funcionamento dos agentes anticâncer que atuam por meio da interação com microtúbulos baseia-se, de forma geral, em deslocar o equilíbrio dinâmico entre as taxas de polimerização e despolimerização de filamentos de microtúbulos no interior celular, de forma que a célula não consiga executar etapas cruciais de seu ciclo de divisão.<sup>42,43,63,64</sup> Quanto ao mecanismo de ação, esses fármacos podem ser classificados em dois grupos: (i) aqueles que estabilizam e (ii) os que desestabilizam os microtúbulos.<sup>62,64</sup> Os fármacos do primeiro grupo impedem que as células desfaçam as fibras do fuso após a separação dos cromossomos durante a mitose (Fase M, Figura 3B), tornando impossível para as células filhas procederem para a etapa final de divisão celular e promovendo um processo de apoptose. O paclitaxel é o agente mais famoso desse grupo (Figura 5A).<sup>61,63</sup> Por outro lado, os fármacos do segundo grupo, ao evitarem a formação de microtúbulos, impedem que a célula forme as fibras do fuso mitótico, necessárias para separar os cromossomos duplicados nas etapas finais da mitose. Como consequência, as células com todo o seu material duplicado são impossibilitadas de prosseguirem seu metabolismo de forma funcional. Exemplos de agentes que atuam dessa forma são a vimblastina e a colchicina (Figuras 5B e 5C, respectivamente).<sup>65,66</sup> Em ambos os mecanismos de ação, o bloqueio da divisão celular em um ponto em que as células encontram-se fragilizadas leva a um processo de apoptose celular.<sup>64-67</sup>

Atualmente são conhecidos cinco regiões distintas para ligação de moduladores da polimerização de tubulina com propriedades antimitóticas.<sup>64</sup> Duas delas promovem a polimerização e a estabilização dos microtúbulos por meio da interação com os sítios de ligação do paclitaxel e da laulimalida. As outras três são responsáveis por inibir a polimerização pela interação com os sítios da colchicina e da vimblastina e pela alquilação de resíduos da tubulina essenciais para polimerização (*e.g.*, o resíduo Cis-239 da  $\beta$ -tubulina). Dentre os mecanismos de modulação descritos, os baseados nos sítios do paclitaxel (de estabilização dos

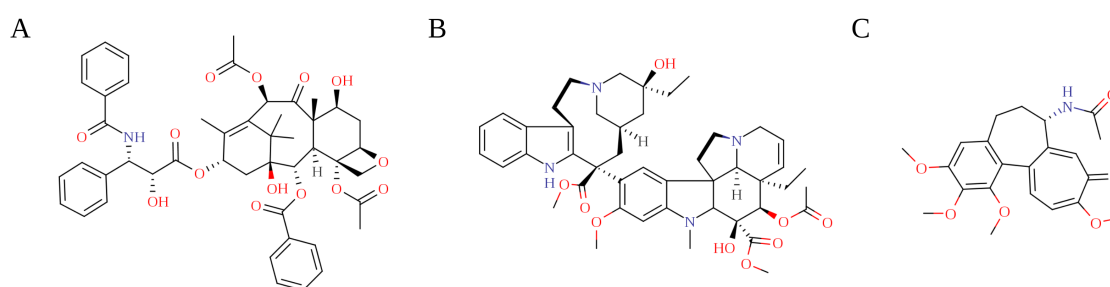


Figura 5 – Estruturas químicas de agentes anticâncer que atuam pela modulação da formação de microtúbulos. Estas moléculas inibem a polimerização dos microtúbulos ligando-se em sítios distintos da tubulina. (A) Paclitaxel é um produto natural extraído da árvore teixo do pacífico (*Taxus brevifolia*) que estabiliza os microtúbulos. (B) Vinorelbina, obtida da planta vinca rosea (*Catharanthus roseus*), é o princípio ativo utilizado para o tratamento de tumores específicos, como linfoma de Hodgkin e câncer de testículos. (C) Colchicina, extraída das plantas colchicum (*Colchicum autumnale*), é um composto que vem sendo estudado pelas suas propriedades anticancerígenas, apesar de seus análogos geralmente apresentarem elevada toxicidade.

Fonte: Elaborada pelo autor.

microtúbulos, Figura 6, em vermelho) e da vinorelbina e da colchicina (de desestabilização, Figura 6; em azul e verde, respectivamente) são os mais conhecidos e bem caracterizados.

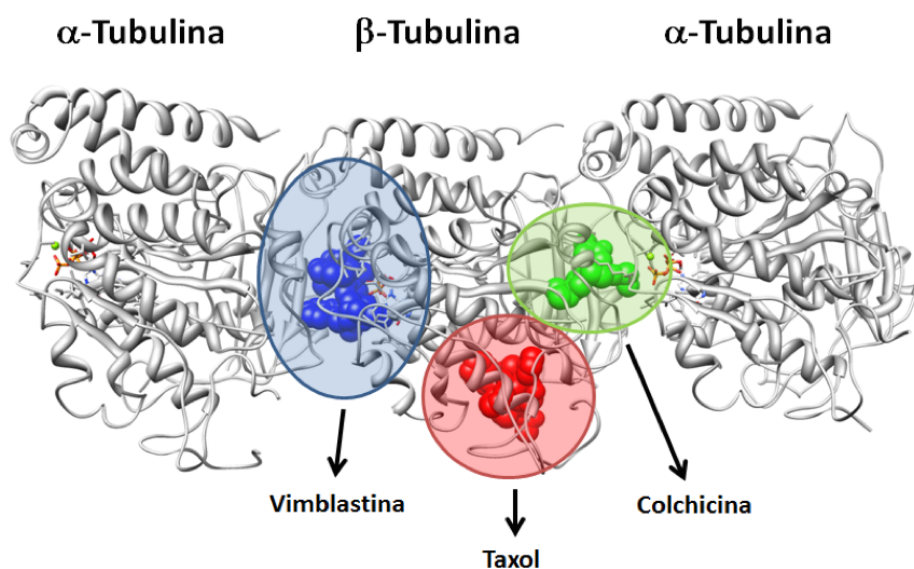


Figura 6 – Principais sítios de ligação de moduladores da polimerização de tubulina e formação de microtúbulos. Figura gerada pela sobreposição de diversas estruturas cristalográficas contendo dímeros de  $\alpha\beta$ -tubulina em complexo com os compostos: vinorelbina (código PDB: 1Z2B), colchicina (código PDB: 1SA0) e paclitaxel (código PDB: 1JFF).

Fonte: Elaborada pelo autor.

Mesmo com o considerável número de fármacos moduladores de microtúbulos disponíveis atualmente, diversas questões como a superexpressão de isoformas específicas, mu-

tações e outros mecanismos de resistência ainda não elucidados limitam a eficácia dos tratamentos terapêuticos atuais.<sup>62,64,68,69</sup> Diante disso, a busca de novos agentes moleculares que atuem contra células resistentes aos tratamentos atuais e que apresentem alta potência e seletividade é essencial para o desenvolvimento de uma nova geração de fármacos para o tratamento do câncer.<sup>62,64,68,70</sup> Atualmente, diversos estudos experimentais e computacionais têm sido desenvolvidos tendo os microtúbulos como alvo principal, resultando na descoberta de novos compostos com propriedades antitumorais e contribuindo consideravelmente para a elucidação dos processos moleculares relacionados ao câncer.<sup>71-76</sup>

## 1.2 Motivações e objetivos

Apesar do trágico panorama mundial de casos de câncer e dos constantes esforços na busca de novos métodos quimioterápicos, diversos fatores como resistência e diversidade fenotípica têm limitado a aplicação dos tratamentos disponíveis. Dentre as diversas estratégias para o desenvolvimento de novos agentes anticâncer, os microtúbulos têm se destacado como um dos alvos moleculares mais efetivos.

Os estudos desenvolvidos e apresentados nesta tese de doutorado têm como objetivo identificar novas moléculas moduladoras da polimerização de microtúbulos com atividade anticâncer. Os objetivos específicos incluem:

- Identificar novos moduladores de microtúbulos através de estudos computacionais e experimentais.
- Elucidar os mecanismos moleculares responsáveis pela atividade biológica observada.
- Empregar estratégias de planejamento racional de fármacos baseando-se em métodos de LBDD e SBDD.
- Avaliar a atividade de novos compostos propostos por meio de estudos experimentais.
- Obter um candidato a fármaco apto para o desenvolvimento de estudos *in vivo*.

## 1.3 Metodologia

### 1.3.1 Estratégias de Planejamento

Os estudos apresentados nesta tese de doutorado têm como foco principal o planejamento de novas classes de candidatos a fármacos, como etapa fundamental de um processo de desenvolvimento de novos agentes quimioterápicos para o tratamento do câncer. As

estratégias envolvidas nestes estudos englobam metodologias modernas para identificar moléculas bioativas, durante as etapas iniciais de um processo de descoberta e desenvolvimento de fármacos (Figura 2).<sup>28-30,33</sup> Esta abordagem envolve o uso de técnicas computacionais e experimentais atuando de forma complementar na identificação, avaliação, caracterização e planejamento molecular.<sup>28,33</sup>

Os microtúbulos foram selecionados como alvo molecular para estes estudos. As estruturas cristalográficas e os modelos moleculares disponíveis na literatura foram avaliados para um planejamento aplicando métodos de SBDD. Em seguida, uma coleção de estruturas tridimensionais de compostos foi gerada e utilizada juntamente com a estrutura do alvo em triagens virtuais para selecionar as moléculas com maior probabilidade de modular a polimerização de microtúbulos. Os compostos mais promissores, com base nas energias preditas de formação dos complexos ligante-receptor, foram adquiridos e avaliados em ensaios bioquímicos para confirmar a atividade biológica.

Em uma etapa seguinte, os compostos bioativos foram submetidos a estudos experimentais e computacionais para elucidar o mecanismo de ação e identificar características estruturais importantes para o processo de reconhecimento molecular. Estes estudos permitiram o planejamento e a avaliação de novos compostos em um processo iterativo. Por fim, os compostos com melhor perfil de atividade foram avaliados em ensaios frente a células tumorais para avaliar seus efeitos na estrutura intracelular, toxicidade e a atividade anticâncer.

Um breve resumo das estratégias de planejamento empregadas neste trabalho é apresentado na Figura 7.

## 1.3.2 Métodos Computacionais

### 1.3.2.1 Geração da Base de Dados de Compostos Sintéticos

Dentre os diversos grupos de pesquisa que realizam colaborações com o Laboratório de Química Medicinal e Computacional (LQMC), coordenado pelo professor Dr. Adriano D. Andricopulo, o Laboratório de Síntese de Produtos Naturais e Fármacos (LSPNF) da Universidade Estadual de Campinas (UNICAMP), dirigido pelo professor Dr. Fernando A. S. Coelho, tem se destacado pela experiência na síntese de compostos químicos de uma grande variedade de classes estruturais.<sup>77-79</sup> O LSPNF tem forte experiência nos estudos dos aspectos mecanísticos da reação de Morita-Baylis-Hillman e na aplicação de seus produtos na síntese de produtos naturais e fármacos.<sup>80,81</sup> Como parte desta colaboração, uma coleção única de compostos do LSPNF foi disponibilizada para estudos computacionais a fim de se identificar novos moduladores da polimerização de microtúbulos. O processo de identificação de compostos bioativos teve a participação da aluna de doutorado Daniara

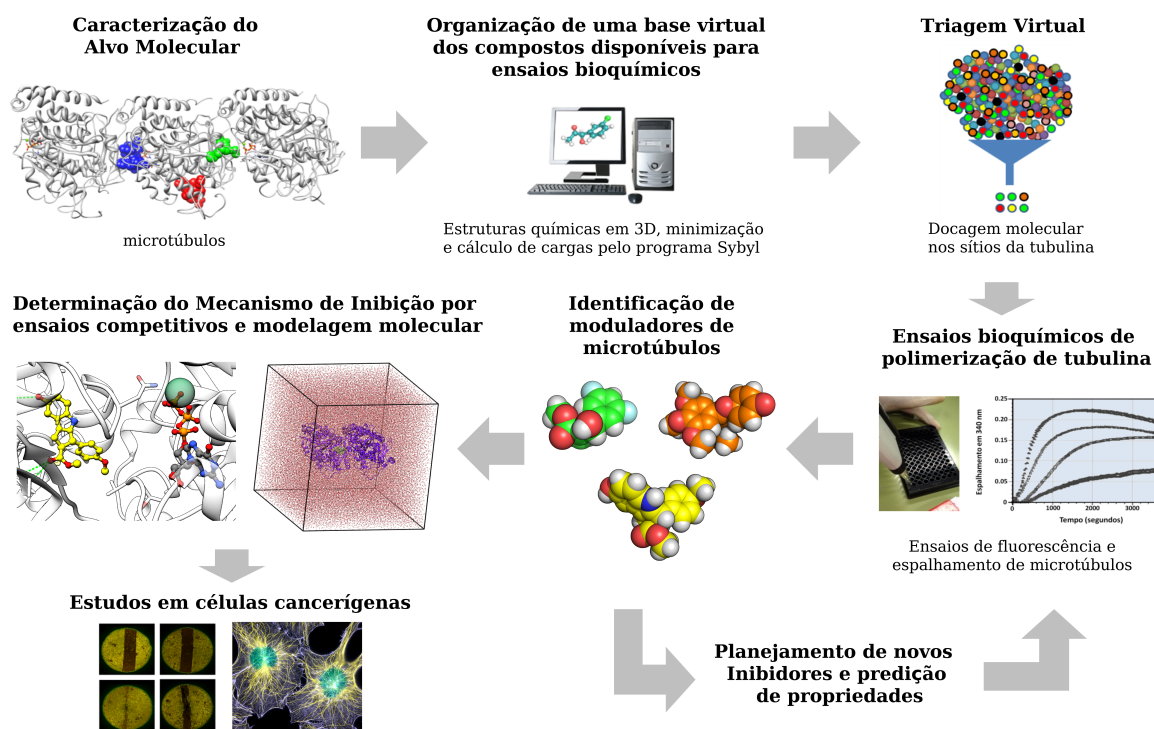


Figura 7 – Estratégia geral do planejamento de novos candidatos anticâncer desenvolvido neste trabalho, tendo como alvo a polimerização de microtúbulos.

Fonte: Elaborada pelo autor.

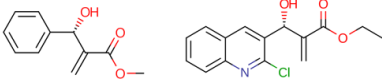
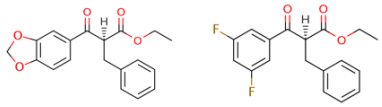
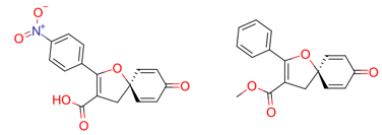
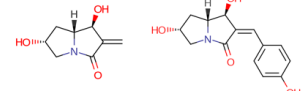
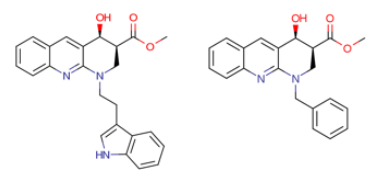
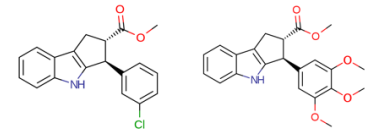
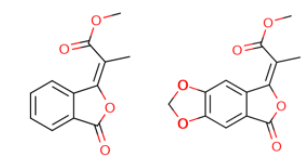
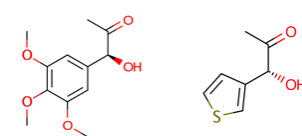
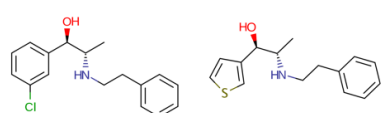
C. Fernandes (sob a orientação do professor Dr. Fernando A. S. Coelho e coorientação do professor Dr. Adriano D. Andricopulo, Processo: CNPq 140878/2011-4), responsável pela síntese de novos análogos direcionada pelos estudos desenvolvidos no LQMC.

Na primeira etapa do trabalho aqui apresentado foi realizada a geração de uma base de dados computacional (2D e 3D) a partir de substâncias sintetizadas no LSPNF, a fim de permitir a predição de propriedades físico-químicas, estudos de modelagem molecular e a seleção racional de compostos para a realização de ensaios bioquímicos. As substâncias já sintetizadas pelo LSPNF foram agrupadas de acordo com a classe estrutural, sendo divididas em 9 grupos, totalizando 258 estruturas. Uma amostragem geral das diferentes classes químicas representadas pelo conjunto de dados é apresentada na Tabela 1. As estruturas moleculares de todas as 258 moléculas utilizadas nestes estudos são apresentadas no Anexo A.

As estruturas tridimensionais dos compostos disponibilizados pelo LSPNF foram geradas utilizando o programa SYBYL 8.0 (Tripos Inc., Saint Louis, MO, Estados Unidos) levando à construção de uma base virtual de compostos.<sup>82</sup> Estas estruturas foram então preparadas para aplicação em estudos de modelagem molecular. A hibridização de cada átomo foi verificada utilizando a opção “Atom Types” do SYBYL e as cargas de todas as moléculas foram calculadas considerando-as em um pH neutro (7,0). A geometria de cada



Tabela 1 – Exemplos de algumas classes estruturais sintetizadas e disponibilizadas pelo LSPNF.

Classe química	Exemplos estruturais
Adutos de Morita-Baylis-Hillman	
$\beta$ -Cetoesteres	
Espirocicloexadienonas	
Pirrolizidinonas e pirrolizidinas substituídas	
1,8-Naftiridinas substituídas	
Derivados indólicos	
Ftalídeos	
Aciloínas	
1,2-aminoalcoóis	

Fonte: Elaborada pelo autor

uma das estruturas foi minimizada quanto à energia conformacional utilizando o campo de força Tripos e o método de minimização de Powell.<sup>83</sup> As distribuições parciais de cargas foram calculadas utilizando o método Gasteiger-Huckel, disponível no SYBYL 8.0.<sup>36,83-85</sup> Durante este processo, o ambiente de solvatação foi considerado de forma implícita através da utilização de uma constante dielétrica de 80,0 (correspondente à água na temperatura ambiente de 25 °C). A conformação 3D de cada molécula foi registrada no formato *Mol2*, compatível com a grande maioria dos programas de docagem molecular.

### 1.3.2.2 Docagem Molecular

A docagem molecular se destaca como um dos métodos computacionais mais amplamente utilizados em SBDD.<sup>28,86</sup> Este método consiste em prever a conformação bioativa de uma pequena molécula (ligante) no sítio de ligação de uma macromolécula (proteína-alvo), para a formação de um complexo receptor-ligante, como ilustrado na Figura 8.

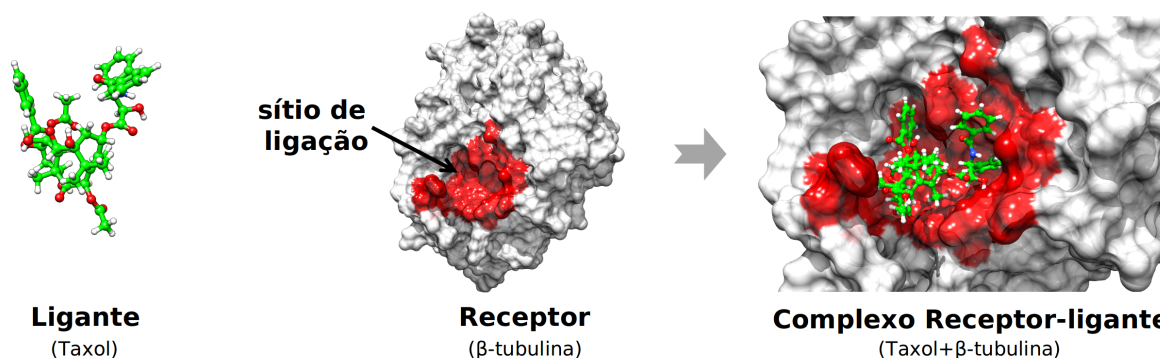


Figura 8 – Representação da interação entre um ligante (uma micromolécula) e um receptor molecular (uma macromolécula) na formação do complexo receptor-ligante. A interação do taxol com seu alvo  $\beta$  tubulina foi utilizada como exemplo.

Fonte: Elaborada pelo autor.

Além da predição do modo de ligação de um ligante em um sítio específico da macromolécula alvo (*e.g.*, uma enzima), o processo de docagem molecular envolve também a avaliação e a classificação do modo de ligação predito para o complexo molecular.<sup>28,86</sup> Por exemplo, para um dado receptor (R) em uma macromolécula alvo, os programas de modelagem molecular permitem estimar a estabilidade do complexo receptor-inibidor (RI). Este grau de estabilidade pode ser representado, por exemplo, pela constante de dissociação do complexo RI, denominada  $K_i$ , que é uma medida quantitativa da afinidade do inibidor (Equação 1.1).<sup>86</sup>

$$K_i = [R][I]/[RI] \quad (1.1)$$

Na equação 1.1,  $K_i$  é a constante de dissociação do complexo;  $[R]$  é a concentração de receptor livre (não ligado) no meio,  $[I]$  a concentração de inibidor livre e  $[RI]$  a concentração do complexo receptor-inibidor. O  $K_i$  pode ser diretamente relacionado com a energia livre de formação de um complexo específico através da Equação 1.2.

$$\Delta G_{lig} = RT \ln K_i \quad (1.2)$$

Na equação 1.2,  $\Delta G_{lig}$  é a variação da energia livre devido à formação (ligação) do complexo receptor-inibidor,  $R$  a constante universal dos gases,  $T$  a temperatura do sistema e  $K_i$  a constante de dissociação do complexo receptor-inibidor.

Os métodos de docagem molecular encontram variadas aplicações em diversas etapas básicas do processo de descoberta de fármacos, desde a identificação de novos ligantes com atividade biológica (também conhecidos como *hits*), até a otimização de compostos líderes.<sup>87</sup> No primeiro caso, coleções contendo um grande número de compostos (da ordem de centenas de milhares até alguns milhões) podem ser avaliadas computacionalmente contra um alvo molecular específico. Esta estratégia de triagem virtual permite a identificação e a seleção de compostos para uma etapa posterior de testes *in vitro* para caracterização da atividade biológica. São características deste método sua grande versatilidade, alto desempenho, baixo custo e economia de tempo (Figura 9A).<sup>87,88</sup>

Além da seleção de coleções moleculares, a docagem molecular também pode ser aplicada na otimização de compostos líderes. Neste caso, modificações moleculares são planejadas levando em conta aspectos químicos e estruturais das moléculas pequenas e do receptor alvo, com o objetivo de melhorar propriedades como afinidade e seletividade (Figura 9B).<sup>28,88,89</sup>

A docagem molecular consiste basicamente em duas etapas fundamentais: a busca conformacional e a pontuação.<sup>88,90</sup>

Na etapa de busca conformacional, um algoritmo de docagem altera a conformação da molécula, variando seus ângulos rotacionais e torcionais, de forma a acoplá-la no sítio de ligação. O procedimento de busca pode também se estender ao próprio sítio de ligação, levando em conta a flexibilidade dos resíduos. Este processo é bastante complicado, pois mesmo moléculas pequenas e de baixa complexidade podem apresentar um número significativo de conformações possíveis. Para ser considerado eficaz, o algoritmo de busca conformacional deve representar uma ampla faixa de combinações de valores para diversos parâmetros geométricos, sendo capaz de identificar as conformações que melhor se encaixam no sítio do receptor alvo. Além disso, esta busca conformacional deve ser efetuada de maneira rápida e eficiente, proporcionando um método aplicável e de custo computacional aceitável.<sup>88,90</sup>

A etapa de pontuação envolve a análise das conformações encontradas pelos méto-

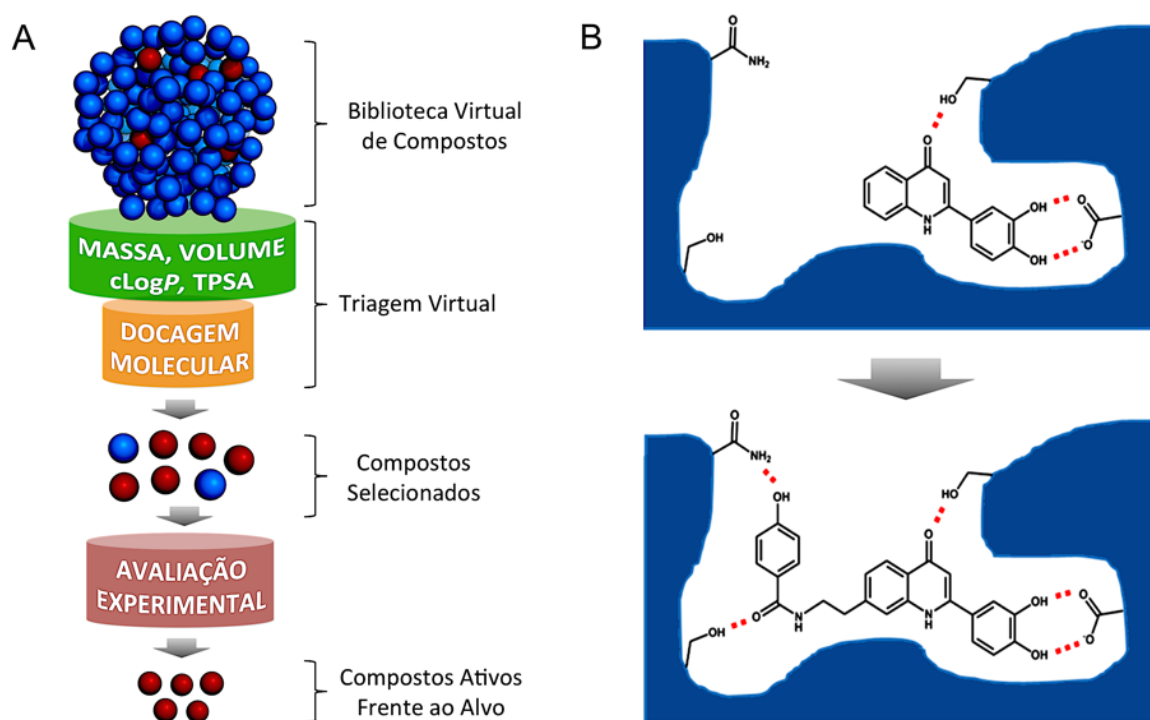


Figura 9 – (A) Uso da triagem virtual na seleção dos compostos mais promissores para ensaios bioquímicos partindo de uma grande base de dados. Compostos ativos e inativos são representados por esferas vermelhas e azuis, respectivamente. (B) Utilização da docagem molecular na otimização da afinidade de ligação de um composto bioativo.

Fonte: Elaborada pelo autor.

dos de busca e a predição da afinidade de ligação do ligante para cada uma delas. Funções baseadas em campos de força são utilizadas para calcular a variação de energia livre na formação de cada complexo, contabilizando as contribuições entálpicas e entrópicas durante este processo. De maneira similar aos algoritmos de busca, as funções de pontuação devem ser eficientes, permitindo uma análise rápida e ao mesmo tempo considerando fenômenos físicos importantes envolvidos na formação dos complexos moleculares.<sup>28,88,90</sup>

Neste trabalho, um processo de triagem virtual da base de compostos do LSPNF foi desenvolvido utilizando os programas GOLD, Surflex-Dock e AutoDock Vina em conjunto. Estes estudos buscaram selecionar os compostos com maior probabilidade de interagir com cada um dos sítios de ligação da tubulina. Os resultados obtidos em cada programa foram analisados e utilizados para classificar os compostos por uma pontuação consenso, a fim de aumentar as chances de selecionar compostos que apresentem atividade em ensaios bioquímicos. A pontuação consenso (do inglês, *consensus score*) é um método bastante difundido e relevante para aumentar a confiabilidade do processo de triagem virtual, uma vez que reúne os resultados de programas de docagem molecular que utilizam diferentes abordagens para estimar a conformação e energia dos complexos proteína-ligante.<sup>91</sup> Os compostos são classificados pela média da classificação dos mesmos nos três programas utilizados. Dessa forma,

os compostos classificados nas primeiras posições na pontuação consenso são aqueles que foram bem pontuados em todos os programas de docagem.

#### 1.3.2.2.1 O Programa GOLD

O programa GOLD (do inglês *genetic optimisation for ligand docking*) é um método de docagem molecular que utiliza um algoritmo genético para realizar a etapa da busca conformacional na obtenção do complexo ligante-receptor. Este método tem apresentado excelente desempenho quando comparado a outros programas de docagem molecular.<sup>92,93</sup> De forma geral, a aplicação de um algoritmo genético na busca conformacional consiste em representar cada conformação do ligante de estudo como um gene (Figura 10A). Cada posição do gene pode conter, por exemplo, o valor para um diedro (ângulo formado por 4 átomos adjacentes) do ligante, permitindo a representação completa da conformação de uma molécula em um único gene. Na primeira etapa da busca conformacional, o programa de docagem molecular gera aleatoriamente uma população com vários genes correspondentes a conformações distintas de um único ligante. A complementariedade de cada uma destas conformações em relação ao sítio de ligação do receptor é avaliada por uma função de pontuação, durante uma etapa de classificação (Figura 10B). Em seguida, os genes (conformações) que apresentaram as piores classificações são descartados e substituídos por novos genes, gerados por pequenas alterações dos genes mais bem pontuados (mutação, inversão, crossover; Figura 10B). Dessa forma, uma nova população é então gerada pela associação dos novos genes obtidos a partir dos genes mais bem pontuados. Este processo é repetido iterativamente até que uma convergência conformacional da população seja observada ou se obtenha uma conformação que satisfaça critérios pré-estabelecidos pelo programa (*e.g.*, o cumprimento de um número máximo de iterações ou uma variação de energia entre iterações menor que um valor estipulado pelo usuário).

Para classificar os complexos obtidos de acordo com a energia resultante da ligação, o programa GOLD dispõe de funções baseadas em campos de força (GoldScore), semi-empíricas (ChemScore), empíricas (ChemPLP) e baseadas em conhecimento (ASP, do inglês *Astex Scoring Potential*).<sup>92,94</sup> A função de pontuação padrão do GOLD e utilizada neste estudo é a GoldScore. A pontuação de cada ligante calculada por esta função é representada por um termo denominado GOLD Fitness, que é proporcional à afinidade predita para o ligante no sítio de ligação (Equação 1.3).<sup>92</sup>

$$GOLDFitness = S_{lh\_ext} + S_{vdw\_ext} + S_{lh\_int} + S_{vdw\_int} \quad (1.3)$$

O GOLD Fitness é a pontuação fornecida pelo GOLD para avaliar a afinidade do ligante pelo sítio de estudo.  $S_{lh\_ext}$  e  $S_{vdw\_ext}$ : Contribuições devido às ligações de hidrogênio e interações de van der Waals (respectivamente) inter-moleculares, entre o ligante e o

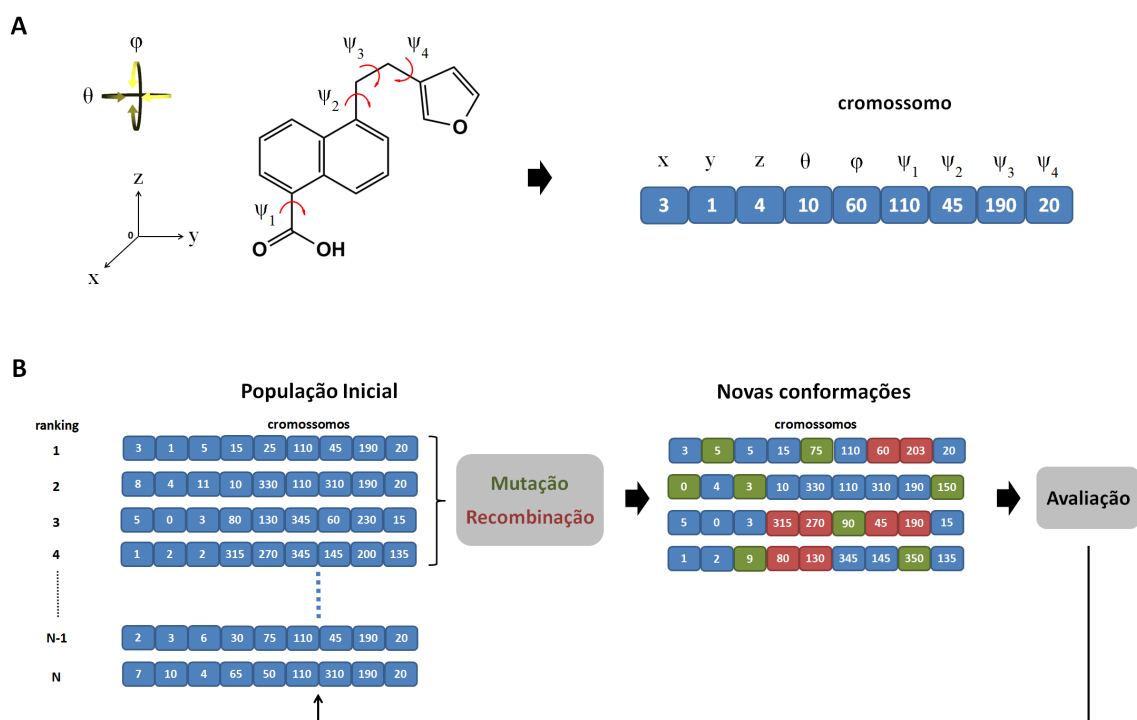


Figura 10 – Descrição geral da aplicação do método de algoritmo genético pelo programa GOLD na busca conformacional de um complexo receptor-ligante. (A) Cada conformação específica de uma molécula é descrita por uma combinação de coordenadas e diedros e representada em um vetor de forma análoga a um gene. (B) Diversas combinações possíveis são geradas para uma mesma molécula resultando em uma "população genética" de conformações. Esta população é submetida a operações que alteram alguns valores no vetor de forma análoga a operações de mutação e recombinação gênica no interior de células. Este processo é repetido até que uma convergência seja observada para uma única conformação de menor energia.

Fonte: Elaborada pelo autor.

receptor. *Slh\_int* e *Svdw\_int*: Contribuições devido às ligações de hidrogênio e interações de van der Waals (respectivamente) intramoleculares.

### 1.3.2.2.2 Docagem Molecular com o Surflex-Dock

O programa Surflex-Dock usa uma função de pontuação empírica e um método de busca conformacional baseado na similaridade molecular do ligante com um ligante fictício idealizado.<sup>95</sup> O ligante ideal utilizado para a busca é denominado *Protomol* e consiste em um conjunto de sondas moleculares representantes de grupos aceptores e doadores de ligação de hidrogênio (C=O e N-H, respectivamente) e grupos hidrofóbicos (CH<sub>4</sub>). Estes grupos preenchem o sítio de ligação e são posicionados de forma a maximizar as contribuições entálpicas do sistema (Figura 11). A partir da identificação das regiões de interação do sítio,

o programa tenta encaixar o ligante no sítio de forma a mimetizar as interações observadas no complexo Protomol-receptor.

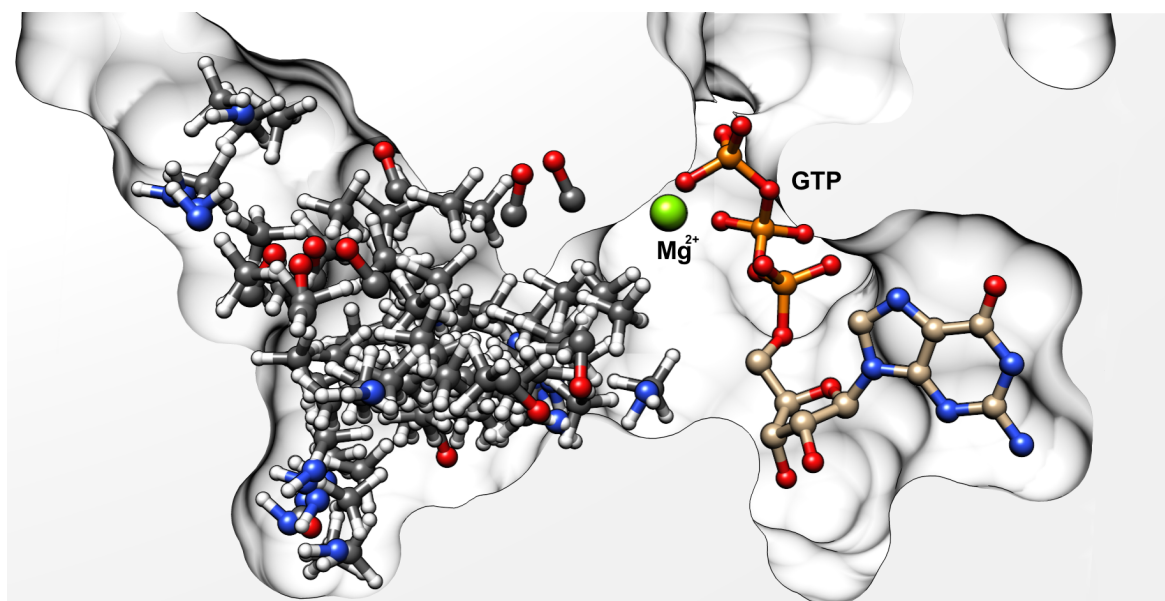


Figura 11 – Exemplo de um protomol gerado para o sítio de ligação da colchicina no dímero de tubulina na presença de GTP e do íon magnésio (código PDB: 1SA0) durante o processo de docagem do programa Surflex-Dock. Átomos de carbono, nitrogênio e oxigênio no protomol estão representados em preto (bege para o GTP), azul e vermelho respectivamente. Hidrogênios estão representados em branco. Todo o restante da proteína foi omitido para maior clareza.

Fonte: Elaborada pelo autor.

#### 1.3.2.2.3 Docagem Molecular com o AutoDock Vina

O programa AutoDock Vina é um software para docagem molecular e triagem virtual desenvolvido pelo laboratório Molecular Graphics situado no Scripps Research Institute, nos Estados Unidos. De livre distribuição, este programa destaca-se por sua fácil utilização e por permitir como padrão processamento em paralelo (utilizando todos os processadores disponíveis no computador). Este program tem apresentado precisão e eficiência comparáveis aos melhores programas de docagem molecular comerciais disponíveis.<sup>96</sup> A busca conformacional do AutoDock Vina, assim como a do programa GOLD, utiliza um algoritmo genético.

O método de pontuação se baseia em uma função semi-empírica que considera uma variedade de contribuições inter- e intramoleculares para o cálculo da variação de energia livre para a formação do complexo. As energias dos complexos são calculadas através do módulo AutoGrid, parte integrante do AutoDock, pela imersão do sistema em uma grade tridimensional e pelo cálculo dos potenciais de interação de cada átomo através do posicionamento de uma sonda ao longo dos pontos da rede 3D (Figura 12). Dessa forma, a

energia de uma configuração específica do ligante analisado pode ser avaliada diretamente utilizando o mapa gerado pela aplicação da rede tridimensional.

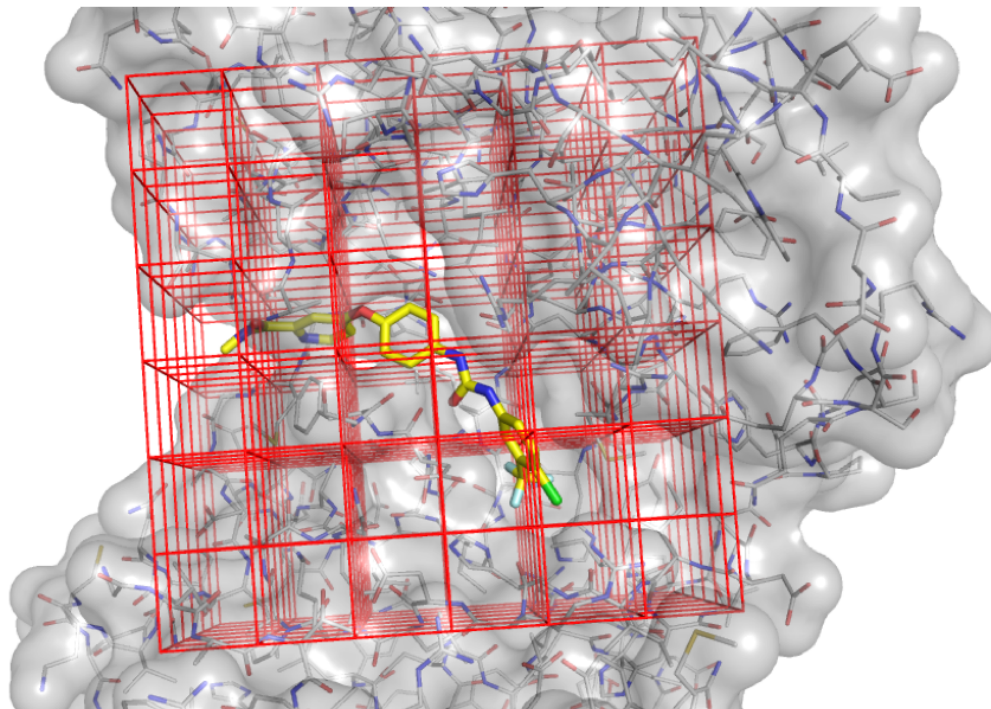


Figura 12 – Ilustração do método utilizado pelo AutoGrid no cálculo dos mapas de energia durante a docagem molecular no programa AutoDock Vina de um complexo hipotético. O espaçamento da grade foi aumentado para melhor visualização.

Fonte: Elaborada pelo autor.

O campo de força utilizado estima a energia de ligação em duas etapas (Figura 13). Na primeira etapa, as energias da proteína e do ligante são avaliadas e comparadas no estado não ligado (Figura 13A) e no estado ligado (Figura 13B). Na segunda etapa essas energias são comparadas com a energia do complexo receptor-ligante (Figura 13C). O resultado final fornecido pelo programa corresponde ao logaritmo da constante de dissociação,  $\log K_d$  ( $\log K_i$  no caso de um inibidor). Nesse caso, quanto menor o valor, maior a afinidade do ligante.<sup>96,97</sup>

### 1.3.2.3 Triagem Virtual e Seleção de Compostos para Avaliação Bioquímica

Em uma primeira etapa no processo de seleção de compostos para avaliação biológica, as estruturas moleculares construídas a partir da base de compostos do LSPNF foram utilizadas na predição de propriedades físico-químicas importantes para caracterizar o perfil farmacocinético do conjunto de compostos, tais como massa molecular, área de superfície polar, lipofilia ( $\log P$ ), número de doadores e aceptores de ligações de hidrogênio e a adequação às "regras dos cinco" de Lipinski.<sup>98,99</sup> Para esta análise, as estruturas dos compostos



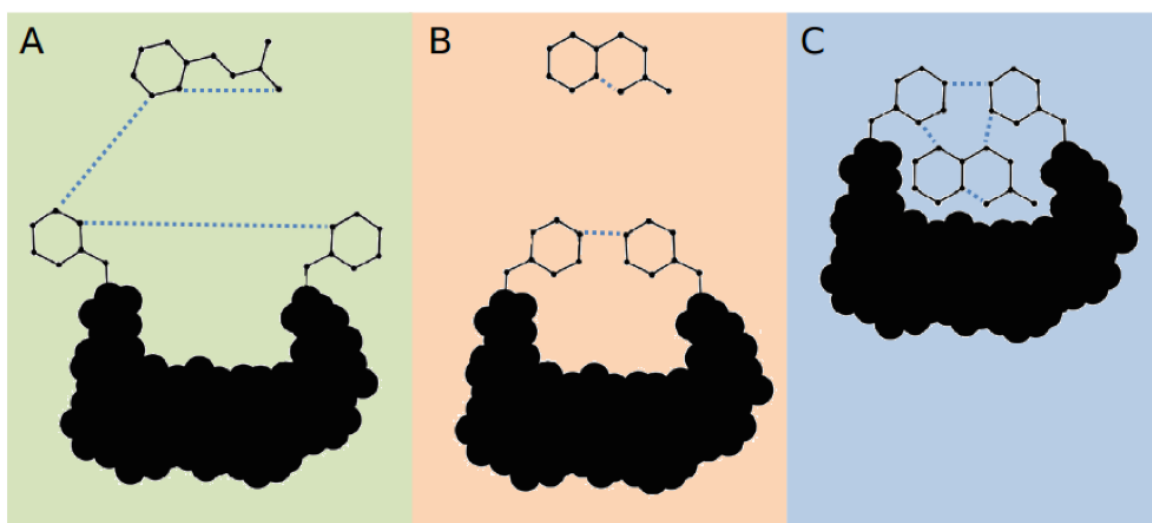


Figura 13 – Conformações energéticas consideradas durante o processo de docagem molecular do programa AutoDock. (A) Estados não ligados da proteína e do ligante. (B) Estados ligados da proteína e do ligante. (C) Complexo formado. Possíveis interações inter- e intra-moleculares são ilustradas por linhas tracejadas.

Fonte: Elaborada pelo autor.

foram submetidas ao servidor Molinspiration e comparadas com as propriedades de compostos já conhecidos capazes de interagir com o mesmo alvo.<sup>149</sup> Este processo foi utilizado para excluir perfis de interação desfavoráveis (Figura 9) e selecionar o perfil farmacodinâmico mais provável para interação dos compostos com o alvo.

A partir dos resultados dos estudos físico-químicos e da seleção do possível mecanismo de ligação, os programas GOLD, Surflex-Dock e AutoDock Vina foram utilizados para os estudos de docagem molecular durante a etapa de triagem virtual. As estruturas cristalográficas dos dímeros de tubulina em complexo com seus ligantes foram coletadas do servidor de estruturas de proteínas Protein Data Bank (PDB). Especificamente, correspondendo aos códigos PDB: 1JFF, 1Z2B e 1SA0 para os complexos com os ligantes paclitaxel, vimblastina e colchicina; respectivamente (Figura 6).<sup>100</sup> A estrutura em complexo com o ligante no sítio de ligação com maior probabilidade de interagir com os compostos da base, de acordo com as análises de propriedades físico-químicas, foi selecionada para estudos de docagem molecular.

O complexo cristalográfico selecionado foi preparado para docagem molecular por meio da retirada do ligante do sítio de ligação, adição de hidrogênios e verificação das conformações dos resíduos no sítio. A fim de maximizar a similaridade da simulação computacional com o ambiente molecular real, os estados de protonação dos resíduos foram avaliados. Os resíduos arginina e lisina foram considerados como protonados (carregados positivamente) enquanto os resíduos aspartato e glutamato foram considerados como desprotonados (negativamente carregados). Em todos os programas de docagem molecular as

coordenadas do ligante original do complexo foram utilizadas como referência para localização do sítio. Em todos os casos um número máximo de 30 conformações distintas foi estipulado para cada ligante da base de compostos.

No programa GOLD, o sítios de ligação corresponde a todos os resíduos presentes em um raio de 6 Å a partir das coordenadas centrais do ligante cristalográfico. A intensidade da busca conformacional durante o processo de docagem foi de 100% e a função de pontuação (do inglês, *Score*) utilizada foi a GoldScore. No programa Surflex-Dock, o sítio foi selecionado utilizando as coordenadas do ligante original para gerar o mapa de sondas protomol, necessário para a docagem. Este protomol foi construído utilizando os parâmetros do programa *threshold* (profundidade do sítio) e *bloat* (expansão do volume do sítio) de 0,5 e 0; respectivamente. A docagem no programa AutoDock Vina foi realizada utilizando um plugin com interface gráfica no programa de modelagem molecular Pymol 1.4 (DeLano Scientific, San Carlos, CA, USA).<sup>101,202</sup> Os sítios de ligação foram estipulados selecionando o volume de uma caixa cúbica com arestas de 15 Å; centralizada a partir do ligante original do complexo. Para calcular os mapas de energia dos sistemas durante a docagem, o módulo AutoGrid foi utilizado pelo AutoDock Vina com um espaçamento entre pontos da rede de 0,375 Å.<sup>96,97</sup>

A seleção dos compostos para testes iniciais a partir dos estudos de triagem virtual foi realizada considerando a pontuação de cada composto nos três programas de docagem molecular utilizados. As moléculas com melhor classificação geral nos três programas (classificação consenso) foram selecionadas e submetidas a ensaios bioquímicos de polimerização de tubulina.

#### 1.3.2.4 Dinâmica Molecular

##### 1.3.2.4.1 Princípios da Dinâmica Molecular

Todos os processos celulares necessários para promover o metabolismo celular e a vida de um organismo são desenvolvidos pela interação entre uma diversidade de moléculas, em um processo sincronizado e harmonioso.<sup>102</sup> Apesar de grande parte dos métodos experimentais como, por exemplo, a cristalografia de raios-X e a microscopia eletrônica permitirem determinar uma conformação específica de uma molécula, em um sistema real em solução as moléculas estão em constante movimento. As funções de uma grande parcela das proteínas de um organismo dependem de alterações conformacionais que exigem um elevado grau de flexibilidade, como por exemplo a abertura e o fechamento de um sítio catalítico de uma enzima.<sup>102,103</sup> Dessa forma, a elucidação detalhada do funcionamento destas proteínas requer a caracterização das suas diversas conformações adotadas em solução.

O entendimento do comportamento dinâmico de um sistema biológico tem grande

importância no desenvolvimento de fármacos.<sup>104</sup> Com os avanços nos estudos de sistemas biológicos, a antiga idéia do mecanismo do tipo “chave-fechadura”, em que o ligante se acomoda em um receptor rígido sem que ocorra nenhum rearranjo estrutural foi gradualmente substituída pela idéia de um modelo de ligação em que ambos o ligante e o receptor sofrem modificações estruturais na formação de um complexo molecular estável.<sup>104,105</sup> Métodos de simulação computacional, tais como a dinâmica molecular, permitem prever o comportamento dinâmico de moléculas em solução e sua trajetória com elevada precisão e nas mais variadas condições físico-químicas. Assim como um vídeo fornece muito mais informações de uma determinada situação do que uma fotografia, a trajetória de um sistema molecular permite obter muitas outras informações que seriam impossíveis de se determinar a partir de uma única estrutura estática.<sup>106</sup> Dessa forma, a partir da utilização destes métodos é possível propor modelos moleculares para explicar as alterações envolvidas nas interações entre moléculas.

De forma geral, qualquer fenômeno de um sistema molecular pode ser previsto aplicando os princípios da mecânica quântica. Apesar disso, a quantidade de cálculos necessários para descrever o comportamento de um sistema biológico, mesmo que pequeno, é excessiva e impraticável para o poder computacional disponível atualmente. A dinâmica molecular, desenvolvida no final da década de 70, busca contornar este problema por meio da simulação do comportamento de sistemas biológicos utilizando aproximações de mecânica clássica que se baseiam nas leis de movimento de Newton para descrever a trajetória de cada átomo em um sistema molecular.<sup>107</sup> Uma simplificação da metodologia da dinâmica molecular é descrita na Figura 14.

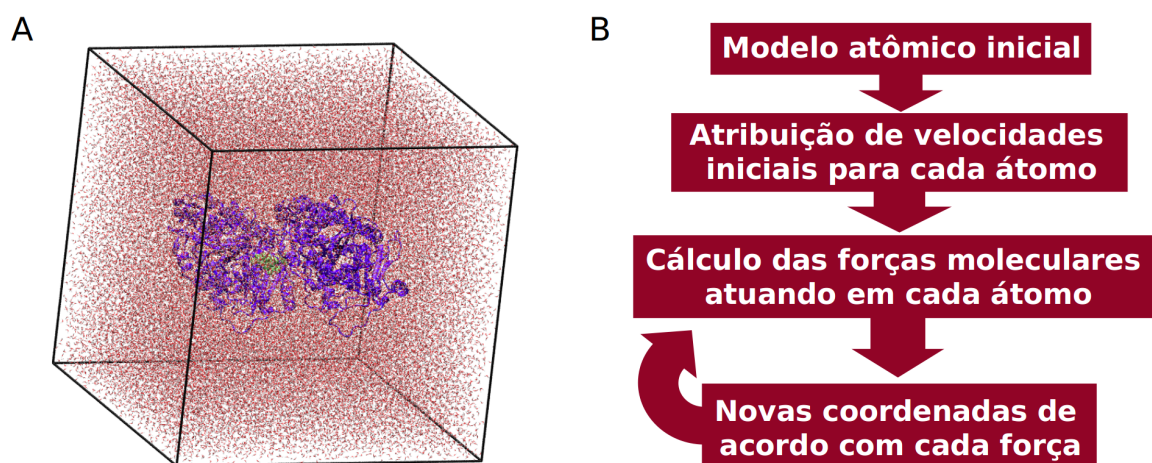


Figura 14 – (A) Exemplo de uma proteína solvatada em uma caixa cúbica com moléculas de água. (B) Etapas principais do processo de dinâmica molecular.

Fonte: Elaborada pelo autor.

Inicialmente a macromolécula de interesse é inserida em um solvente de acordo com as condições reais presentes no interior celular ou no ensaio experimental a ser simulado (Fi-

gura 14A). Este ambiente geralmente inclui explicitamente moléculas do solvente e possíveis íons e cofatores. Em seguida, velocidades iniciais são atribuídas a cada átomo do modelo inicial, de forma a representar uma distribuição de velocidades compatível com a temperatura e pressão do sistema real (Figura 14B).<sup>104,106</sup> A partir das coordenadas iniciais do sistema, as forças de interação entre cada átomo são calculadas por uma equação de campo de força (Figura 15).

$$E_{total} = \sum_{\text{ligação}} K_r (r - r_{eq})^2 + \sum_{\text{ângulos}} K_\theta (\theta - \theta_{eq})^2 + \sum_{\text{diedros}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] + \sum_{i < j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

Figura 15 – Exemplo de uma equação de um campo de força utilizado para simulações de dinâmica molecular. Os termos ligados correspondem às interações entre átomos ligados através de ligações covalentes. As contribuições das variações de energia devido à alterações na distância ( $r$ ), ângulo ( $\theta$ ) e diedros ( $\phi$ ) em relação aos valores de equilíbrio são determinadas a partir das constantes  $K_r$ ,  $K_\theta$  e  $K_\phi$  ajustadas para reproduzir sistemas reais. O termo não ligado corresponde às interações hidrofóbicas e eletrostáticas entre átomos não ligados diretamente por ligação covalente.

Fonte: Elaborada pelo autor.

Estes campos de força consideram as contribuições devidos às interações entre átomos ligados e não ligados (Figura 15). Para reproduzir fielmente as interações de um sistema real, os parâmetros do campo de força utilizado são geralmente ajustados utilizando dados experimentais ou preditos por métodos quânticos.<sup>104,106,107</sup>

A partir do cálculo das forças agindo sobre cada átomo do sistema, uma nova conformação para o sistema é gerada para um pequeno intervalo de tempo (da ordem de femtosegundos) de acordo com as leis de movimento de Newton (Figura 14B). Este processo é repetido iterativamente permitindo integrar toda a trajetória do sistema ao longo do tempo.

Diversos campos de força foram desenvolvidos para o desenvolvimento de simulação de dinâmica molecular com sistemas biológicos como, por exemplo, o GROMOS, o CHARMM e o AMBER.<sup>108–111</sup> Analogamente, diversos programas estão disponíveis para o desenvolvimento destas simulações, como GROMACS, NAMD, DESMOND e AMBER.<sup>112–115</sup>

#### 1.3.2.4.2 O Programa GROMACS

Dentre os programas disponíveis atualmente para o desenvolvimento de simulações de dinâmica molecular de sistemas biológicos, o GROMACS (do inglês *GRoningen Machine for*

*Chemical Simulations*, Groningen, Holanda) vem sendo amplamente utilizado. Este programa, de livre acesso, caracteriza-se por ser rápido e permitir a paralelização entre processadores e a utilização de unidades de processamento gráfico.<sup>112,116</sup>

O GROMACS é executado por linha de comando através de uma combinação de módulos que convertem estruturas e trajetórias moleculares para o formato compatível para o programa. Isto permite adicionar moléculas de um solvente, íons e outras moléculas específicas ao sistema, além de executar os cálculos de trajetória.

Este programa tem sido utilizado com sucesso na elucidação de diversos processos moleculares, como alterações conformacionais de proteínas envolvidas no transporte de moléculas através de membranas, em processos enzimáticos e na formação de complexos ligante-receptor no desenvolvimento de fármacos.<sup>117-121</sup>

#### 1.3.2.5 Estudos de Dinâmica Molecular com Inibidores da Polimerização de Microtúbulos

Os compostos identificados como moduladores da atividade de tubulina foram submetidos a estudos de dinâmica molecular a fim de avaliar a estabilidade dos complexos preditos por docagem molecular e identificar as interações responsáveis pela formação destes complexos. Para isso, uma nova docagem do ligante foi efetuada utilizando parâmetros mais refinados. A estrutura cristalográfica de maior resolução disponível de um complexo do ligante interagindo com seu sítio em um dímero de tubulina foi selecionada.<sup>122,123</sup> A estrutura do receptor foi preparada para a docagem pela retirada do ligante em complexo, adição de hidrogênios e pela verificação das conformações dos resíduos dos sítios, utilizando o programa Pymol 1.7.<sup>101,202</sup> A fim de maximizar a similaridade da simulação computacional com o ambiente molecular real, os estados de protonação dos resíduos foram avaliados, como descrito anteriormente no processo de triagem virtual. As coordenadas do ligante original do complexo foram utilizadas como referência para determinar o centro de um cubo com dimensões de 20 Å, que corresponde ao sítio de ligação a ser considerado pelo programa AutoDock Vina para executar o processo de busca conformacional. O mesmo procedimento foi também executado utilizando o ligante cristalográfico em um processo de validação.

Os complexos preditos pela docagem molecular dos compostos selecionados foram submetidos a um processo de dinâmica molecular com o programa GROMACS 4.5. A estrutura do receptor utilizada na docagem molecular foi convertida para o formato *gro*, compatível com o programa GROMACS. As topologias do receptor foram geradas utilizando o módulo *pdb2gmx* do programa, com o campo de força GROMOS96 43A1.<sup>124,125</sup> Analogamente, as coordenadas e as topologias dos ligantes (contendo os ângulos, diedros, massas e cargas dos átomos) foram geradas pelo servidor PRODRG2.<sup>126</sup> O mesmo processo foi exe-

cutado para o cofator GTP e o íon  $Mg^{2+}$ , presentes na estrutura original do complexo. Estes arquivos foram combinados aos obtidos para o receptor permitindo obter as coordenadas e topologia do complexo predito pelos estudos de docagem, necessários para realizar as simulações. O solvente foi considerado explicitamente adicionando-se moléculas de água do tipo TIP3 pelo módulo *genbox* do GROMACS e o sistema final foi neutralizado adicionando-se íons de sódio e cloro em coordenadas aleatórias em meio ao solvente pelo módulo *genion*, resultando em um sistema solvatado em uma caixa cúbica com margens espaçadas a 10 Å em relação à superfície do complexo.

O sistema final foi submetido a um processo de minimização utilizando uma função do tipo *Steepest-Descent*, a fim de corrigir qualquer colisão estereoquímica de alta energia.<sup>127</sup> Em seguida, o sistema foi equilibrado em duas etapas: (i) por uma simulação a volume e temperatura constante (*NVT*), em que a temperatura do sistema foi elevada até a temperatura do sistema real, e em seguida (ii) por uma simulação a pressão e temperatura constante (*NPT*).<sup>125</sup> Em ambas as etapas a temperatura considerada foi de 310 K (37 °C), a mesma temperatura dos ensaios de avaliação da atividade dos ligantes frente à tubulina. Durante todo o processo de equilíbrio, os ligantes presentes no sistema (molécula bioativa selecionado, GTP e íon  $Mg^{2+}$ ) foram mantidos fixos para conservar o estado conformacional inicial do complexo. O sistema final equilibrado a uma temperatura de 310 K e pressão de 1 atm foi então submetido a simulações (*NPT*) de 10 ns. Para cada um dos ligantes o processo de simulação descrito foi repetido em triplicata.<sup>124,125</sup> Como validação da metodologia de dinâmica molecular, o mesmo procedimento também foi repetido para o ligante cristalográfico.

### 1.3.3 Métodos Experimentais

#### 1.3.3.1 Polimerização de Microtúbulos

Os microtúbulos são estruturas longas e filamentosas, compostas basicamente por dímeros de  $\alpha$ - e  $\beta$ -tubulina (Figura 16). Ambas as subunidades são encontradas em todos os organismos eucariotos, sendo extremamente conservadas e possuindo massa molecular de aproximadamente 55 kDa.<sup>62,64,128</sup>

O equilíbrio entre as diversas formas oligoméricas de tubulina ( $\alpha$ -tubulina,  $\beta$ -tubulina, dímero de  $\alpha\beta$ -tubulina e filamentos de microtúbulos) no interior celular é controlado principalmente pela concentração de GTP (trifosfato de guanosina) em solução. Na presença de baixas concentrações de GTP,  $\alpha$ - e  $\beta$ -tubulinas são encontradas como heterodímeros em sua forma funcional, preparados para atuar como blocos na construção de microtúbulos (Figura 16). Cada tubulina possui um sítio de ligação de GTP. Ao se ligar no sítio da  $\alpha$ -tubulina, o GTP estabiliza a proteína em uma interface que induz a associação da  $\beta$ -

tubulina e a formação do dímero funcional. Desta forma o GTP fica estabilizado em uma interface intra-dimérica (Figura 16).

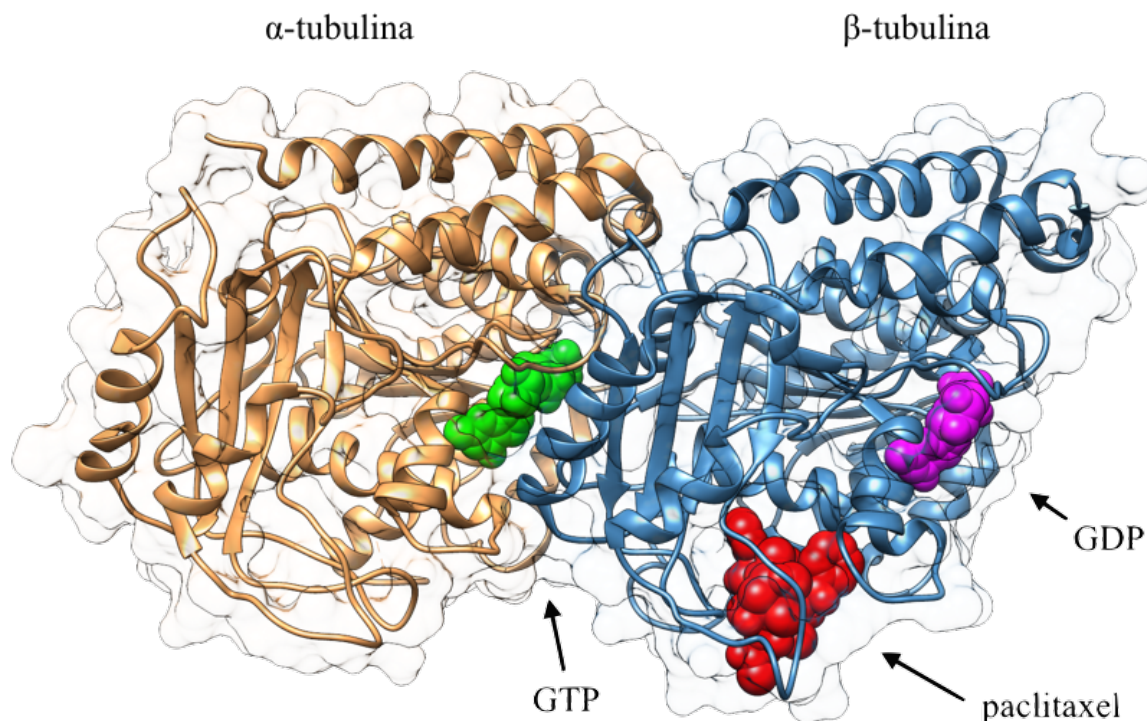


Figura 16 – Estrutura cristalográfica de um dímero formado pela interação entre monômeros de  $\alpha$ - e  $\beta$ -tubulina (código de acesso PDB: 1JFF). As moléculas de GTP e GDP estão apresentadas em verde e magenta, respectivamente. A molécula de paclitaxel (em vermelho) pode ser observada interagindo com seu sítio de ligação na  $\beta$ -tubulina.

Fonte: Elaborada pelo autor.

Na presença de altas concentrações de GTP, o sítio da  $\beta$ -tubulina em uma das extremidades do heterodímero também é ocupado (Figura 16). Assim como no sítio da  $\alpha$ -tubulina, a presença do GTP neste sítio também induz a formação de uma interface entre  $\alpha$ - e  $\beta$ -tubulinas e resulta na associação entre dois dímeros. Este processo de associação iterativa de dímeros permite a formação dos microtúbulos. Por outro lado, moléculas de GTP ligadas ao sítio da  $\beta$ -tubulina não apresentam a mesma estabilidade do que na interface intra-dimérica e podem ser hidrolisadas para GDP (difosfato de guanosina).

Em sua forma não hidrolisada, as moléculas de GTP presentes nas interfaces entre dímeros induzem a associação longitudinal de novos dímeros livres de  $\alpha\beta$ -tubulina, resultando na formação de pequenos protofilamentos estabilizados em uma forma reta (Figura 17A). Estes protofilamentos se associam lateralmente formando folhas estáveis de tubulina (Figura 17B), que crescem até se fecharem em um microtúbulo de 13 protofilamentos (Figura 17C). O arranjo dos dímeros de tubulina ao longo de um protofilamento confere uma polaridade aos microtúbulos que permite designar as duas extremidades como (+) e (-) (Figura 17C). A

presença do GTP ligado à extremidade (+) da tubulina faz com que os dímeros de tubulina sejam adicionados preferencialmente na extremidade (+) e se dissociem preferencialmente na extremidade (-) de cada microtúbulo (Figura 17C).<sup>47, 128, 129</sup>

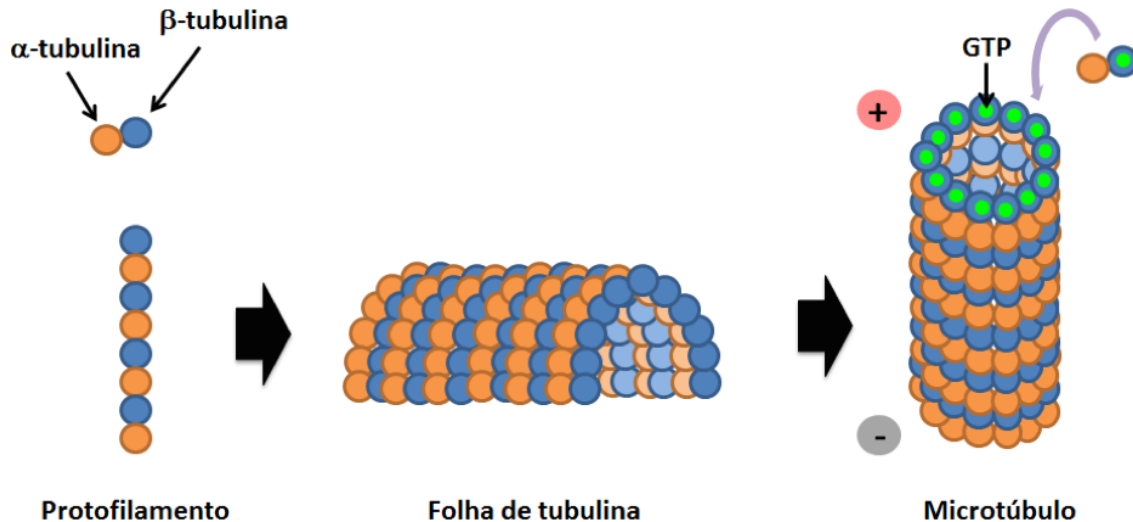


Figura 17 – Esquema de polimerização dos microtúbulos. A partir da associação de dímeros de  $\alpha\beta$ -tubulinas um protofilamento é formado. As interações laterais entre os dímeros do protofilamento levam à formação de uma folha de tubulinas. O crescimento dessa folha faz com que a mesma se feche formando um microtúbulos composto por 13 protofilamentos. A presença de moléculas de GTP (esferas verdes) expostas na superfície da subunidade  $\beta$ -tubulina faz com que a extremidade (+) tenha uma maior taxa de adição de novos dímeros. O equilíbrio entre as taxas de polimerização e a despolimerização permite controlar o tamanho dos microtúbulos no interior celular.

Fonte: Elaborada pelo autor.

Ao longo do processo de crescimento dos microtúbulos, as moléculas de GTP nas interfaces dos dímeros já associados podem ser hidrolisadas em GDP, desestabilizando a forma reta dos filamentos e gerando uma tensão ao longo dos microtúbulos. Apesar disso, as interações laterais entre os filamentos e a alta fração de dímeros com GTP não hidrolisado nas extremidades do microtúbulo impedem a despolimerização. Neste processo, o tempo de hidrólise de GTP para GDP é um fator chave para a dinâmica de polimerização. Quando a concentração de GTP em solução começa a diminuir, a taxa de crescimento e a quantidade de dímeros associados por GTP na extremidade de crescimento (+) decaem e o microtúbulo começa a perder estabilidade. Uma vez que os GTPs na extremidade do microtúbulo são hidrolisados, a tensão acumulada nos protofilamentos é liberada e o microtúbulo se desfaz rapidamente.<sup>47, 128, 129</sup> Este fenômeno de crescimento lento e despolimerização rápida dos microtúbulos é conhecido como instabilidade dinâmica.<sup>130</sup>



### 1.3.3.2 Quantificação de Polimerização da Tubulina por Fluorescência

As taxas de polimerização e despolimerização de microtúbulos na presença de ligantes podem ser avaliadas por diversos métodos experimentais. Um método bastante sensível e que permite determinar quantitativamente a influência de um determinado ligante na polimerização de microtúbulos é o baseado em fluorescência.<sup>131-133</sup> Nesse método, o composto repórter e fluorescente DAPI (4',6-Diamidino-2-fenilindol; Figura 18A) é adicionado à reação de polimerização de tubulina. Esse composto possui a capacidade de se incorporar a microtúbulos com uma afinidade muito maior do que a dímeros livres de tubulina e apresenta a característica de absorver e emitir luz com máximos de comprimento de onda em 350 nm e 460 nm, respectivamente. Quando associados à uma macromolécula, as taxas de absorção e emissão do DAPI são amplificadas. Desta forma, durante a polimerização da tubulina na presença de DAPI é possível observar um aumento de emissão de luz no comprimento de 460 nm devido à incorporação do repórter fluorescente nos microtúbulos, à medida que os mesmos são polimerizados. O processo de polimerização é caracterizado por três fases distintas (Figura 18B). Na fase I, conhecida como nucleação, centros de associação de dímeros surgem, formando protofilamentos e folhas de tubulina. Na fase II, denominada de crescimento, os microtúbulos que surgem a partir do fechamento das folhas de tubulina começam a crescer continuamente, pela adição de dímeros de  $\alpha\beta$ -tubulina. Esta etapa é caracterizada pelo aumento de sinal devido ao número crescente de moléculas repórter associadas aos microtúbulos. A fase III é conhecida como fase de equilíbrio, em que uma equiparação entre as taxas de polimerização e despolimerização é estabelecida, a qual depende principalmente da concentração de tubulina e GTP no ensaio e da temperatura do sistema.<sup>131,132</sup>

### 1.3.3.3 Ensaios de Polimerização de Tubulina utilizando Fluorescência

Para estudar a capacidade dos compostos selecionados por triagem virtual em modular a dinâmica de formação de microtúbulos, ensaios de polimerização de tubulina foram conduzidos. Os ensaios foram realizados com kits comerciais de polimerização de tubulina (Cytoskeleton Inc., Denver, CO, EUA). O kit de polimerização (código BK011P, baseado em fluorescência) consiste em 10 mg de tubulina liofilizada (em suas isoformas  $\alpha$  e  $\beta$ ) purificada de cérebro de porco (pureza > 99%), alíquotas de GTP, paclitaxel, DMSO e soluções tamponantes específicos para reconstituição e polimerização de tubulina.<sup>131</sup> Os ensaios foram realizados de acordo com protocolos já estabelecidos para identificação de moduladores de polimerização de tubulina.<sup>131-133</sup> A reação disponibilizada no kit comercial para ensaio consiste em: 80 mM PIPES (piperazina-1,4-bis(ácido 2-etanosulfônico)), 0,5 mM EGTA (ácido etileno glicol-bis ( $\beta$ -amino-etil-éter)  $N,N,N',N'$ -tetra-acético), 1 mM GTP, 2 mM  $MgCl_2$ , 15% de glicerol e 10  $\mu$ M de DAPI; em pH 6,9.

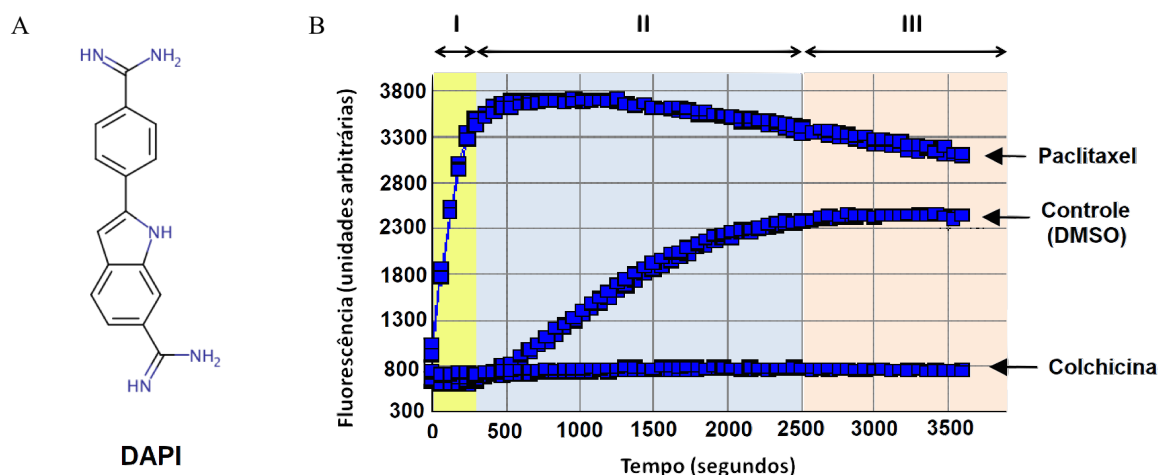


Figura 18 – (A) Estrutura molecular do composto fluorescente DAPI (4',6-Diamidino-2-fenilindol) utilizado como repórter no ensaio de detecção de polimerização da tubulina. (B) Descrição do ensaio de polimerização da tubulina baseado em fluorescência em três condições distintas: (i) na ausência de moduladores de microtúbulos (controle); (ii) na presença de paclitaxel, um estabilizador de microtúbulos; (iii) na presença de colchicina, um desestabilizador de microtúbulos. As fases de nucleação (I), crescimento (II) e equilíbrio (III) estão destacadas no gráfico.

Fonte: Elaborada pelo autor.

As medidas de fluorescência foram executadas utilizando o leitor de placas Victor 3 1420 (Perkin Elmer Inc, Waltham, MA, Estados Unidos) disponível em nosso laboratório (Figura 19A), juntamente com placas OptiPlate-1536 F Black (Perkin Elmer, Waltham, Estados Unidos) opacas e de 96 poços (Figura 19B). Com o intuito de padronizar todas as medidas de polimerização da tubulina na presença de diversos ligantes, todos os ensaios foram realizados utilizando reações de 50  $\mu\text{L}$ /poço da placa e diluindo os compostos (originalmente dissolvidos em 100% de DMSO) em concentrações de 1% de DMSO. Esta concentração de DMSO foi determinada a fim de permitir a solubilização dos compostos no meio reacional sem afetar a cinética de reação dos ensaios.<sup>134</sup>

Na primeira etapa de realização dos ensaios, a temperatura do leitor de placas foi pré-ajustada para 37 °C. Aliquotas de 5  $\mu\text{L}$  de soluções preparadas para cada ligante (10 vezes concentradas e em 10% DMSO) foram transferidas para cada poço. Em seguida, uma reação de polimerização foi preparada pela adição de GTP e tubulina a um tampão com glicerol a baixa temperatura (no gelo, a 0 °C), levando a uma mistura com 2 mg/mL de tubulina e 1 mM de GTP. Imediatamente após preparar a mistura, alíquotas de 45  $\mu\text{L}$  da reação foram adicionadas à placa utilizando-se uma pipeta multicanal. A placa de leitura foi inserida no fluorímetro e o processo de excitação a 355 nm e leitura de emissão a 410-480 nm foi iniciado. Foram executados 70 ciclos de medidas em todos os poços da placa, com intervalos de aproximadamente 1 min entre duas leituras do mesmo poço. Os ensaios foram realizados em triplicata, utilizando 3 poços adjacentes.

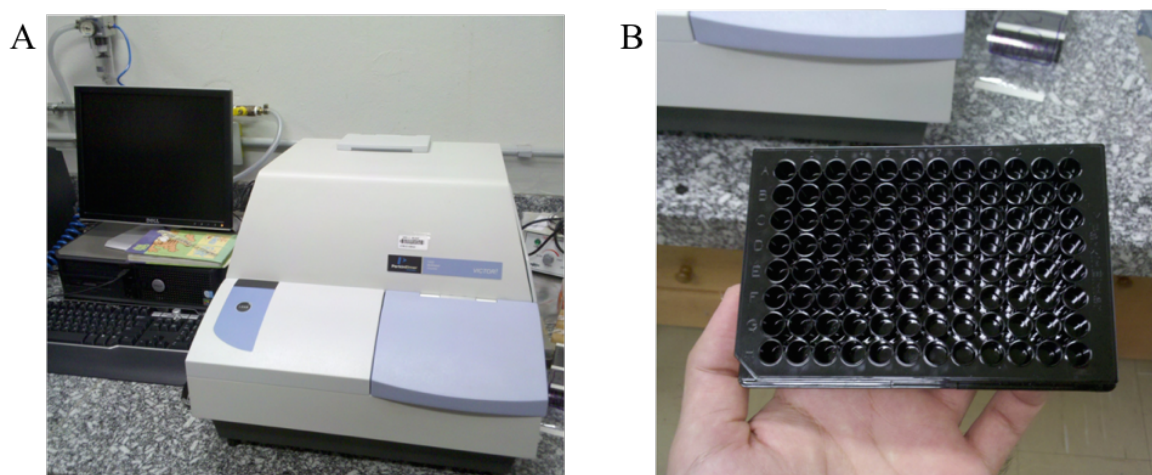


Figura 19 – (A) Fluorímetro Victor 3 1420, utilizado para monitorar a polimerização de tubulina. (B). Placa de leitura de 96 poços para ensaios de fluorescência.

Fonte: Elaborada pelo autor.

A inibição dos microtúbulos foi avaliada a partir do valor do sinal de fluorescência medido 20 minutos após o início do ensaio (região em que se inicia a fase de estabilização da taxa de polimerização de tubulina, Figura 18).<sup>135</sup> A porcentagem de inibição da formação de microtúbulos foi calculada pela razão entre a diferença de sinais de fluorescência entre o composto e o controle positivo para inibição (colchicina) e entre o controle negativo (DMSO) e o controle positivo (colchicina), como descrito na equação 1.4.

$$\%Inibição = 1 - (F_{composto} - F_{colchicina}) / (F_{DMSO} - F_{colchicina}) \quad (1.4)$$

#### 1.3.3.4 Quantificação de Polimerização da Tubulina por Espalhamento

O ensaio clássico para quantificar a polimerização de tubulina consiste em monitorar a absorbância do sistema com comprimento de onda na faixa de 340 nm. Esse ensaio baseia-se no fato de que o espalhamento de luz neste comprimento de onda é proporcional ao aumento da turbidez e da absorbância do meio devido à formação de microtúbulos (Figura 20A).<sup>136</sup> O comportamento observado para a polimerização neste ensaio é o mesmo do ensaio de fluorescência, apresentando fases de nucleação, crescimento e equilíbrio (fases I, II e III respectivamente, Figura 20A). Analogamente ao ensaio de fluorescência, compostos já conhecidos e capazes de modular a polimerização de tubulina podem ser utilizados como padrão positivo (paclitaxel, Figura 20B) e negativo (nocodazol, Figura 20B), junto ao ensaio controle de polimerização padrão (DMSO, Figura 20B) para avaliar a capacidade de compostos teste em modular a formação de microtúbulos.

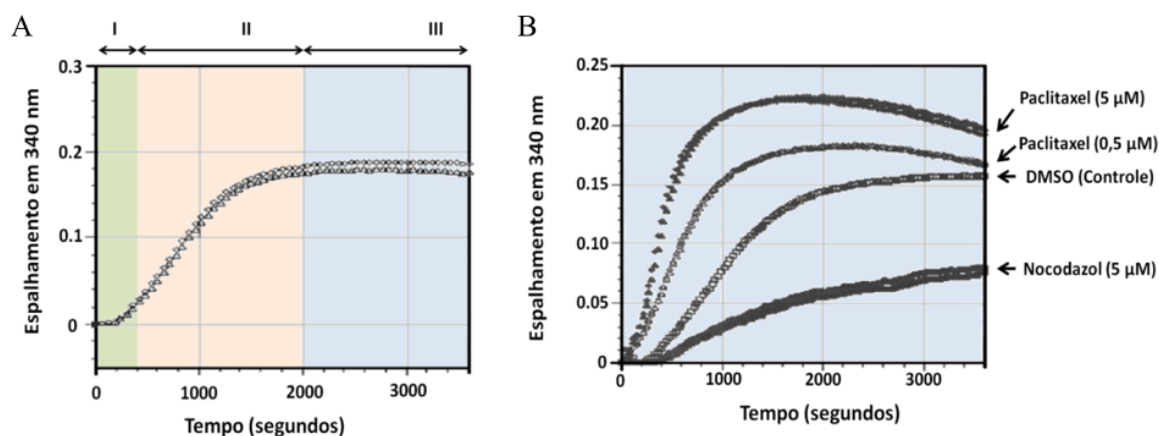


Figura 20 – (A) Ensaio de polimerização da tubulina baseado em espalhamento de luz, medido pela variação de absorvância em 340 nm. Como no ensaio de fluorescência, observam-se três etapas distintas da polimerização: (I) nucleação, (II) crescimento e (III) equilíbrio. (B) Ensaio de polimerização utilizados como controles: (i) na ausência de ligantes (controle com DMSO 1%); (ii) na presença de paclitaxel, um estabilizador de microtúbulos; (iii) na presença de nocodazol (ou colchicina), um desestabilizador de microtúbulos.

Fonte: Elaborada pelo autor.

A desvantagem do ensaio clássico de espalhamento consiste na necessidade de utilizar tubulina em volume e quantidade maiores do que no ensaio de fluorescência (Tabela 2). Como a variação do sinal de espalhamento de luz (proporcional à absorvância) durante a polimerização é bastante sutil, a concentração de tubulina necessária para observar uma variação de sinal considerável deve ser alta. Esta variação de sinal também é proporcional ao caminho óptico percorrido, exigindo um volume maior de solução em cada ensaio (Tabela 2). Apesar destas desvantagens, o ensaio clássico pode ser utilizado junto com o ensaio baseado em fluorescência para confirmar a capacidade de modular a polimerização observada e representa uma alternativa ao ensaio de fluorescência quando o composto de estudo interage com o DAPI ou emite no mesmo comprimento de onda.

Tabela 2 – Características dos ensaios de polimerização de tubulina baseados em fluorescência e espalhamento.

Características	Fluorescência	Espalhamento
Concentração de tubulina por ensaio	2 mg/mL	3 mg/mL
Volume da reação	50 $\mu$ L	100-200 $\mu$ L
Razão entre sinal e ruído (S/R)	4	2

Fonte: Elaborada pelo autor

### 1.3.3.5 Ensaio de Polimerização de Tubulina utilizando Espalhamento

O ensaio de polimerização clássico, baseado em espalhamento, é muito similar ao ensaio baseado em fluorescência descrito anteriormente, caracterizando-se pela ausência

do marcador fluorescente DAPI, bem como pelo maior volume e concentração de tubulina em cada ensaio (Tabela 2) e pela concentração de glicerol (5%). A tubulina utilizada neste ensaio (código HTS02-B, Cytoskeleton Inc., Denver, CO, Estados Unidos) consiste em 44 mg de proteína liofilizada, purificada de cérebro bovino (com 97% pureza). O tampão utilizado (preparado em laboratório) contém os seguintes componentes: 80 mM PIPES, 1 mM EGTA, 1 mM GTP, 1 mM MgCl<sub>2</sub> e 5% glicerol; em pH 7,0. As medidas de absorbância foram executadas utilizando o espectrofotômetro de placa SpectraMax Plus 384 (Molecular Devices, Sunnyvale, CA, Estados Unidos) disponível no laboratório LQMC (Figura 21A), juntamente com placas TPP92696 Tissue Culture Plates (TPP, Saint Louis, MO, Estados Unidos) de 96 poços e transparentes para medição de absorbância (Figura 21B).

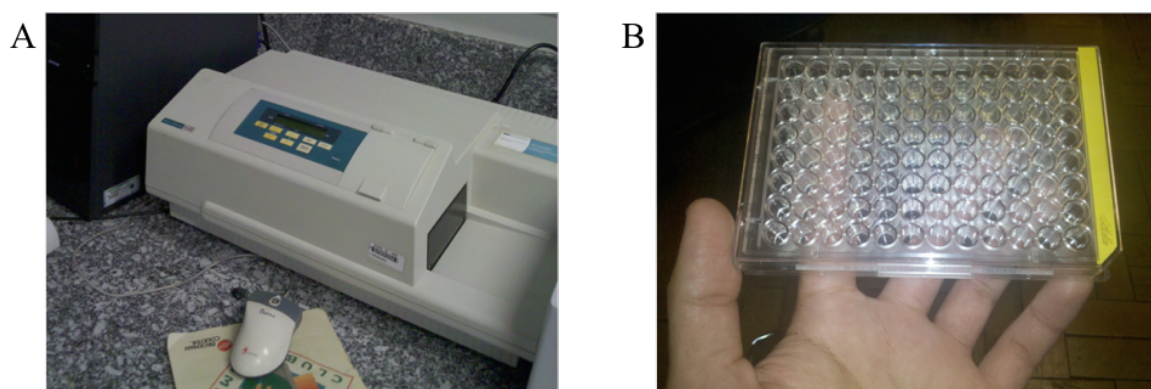


Figura 21 – (A) Espectrofotômetro SpectraMax Plus 384 utilizado para monitorar a polimerização de tubulina por absorbância (espalhamento de luz). (B). Placa de leitura transparente de 96 poços utilizada para os ensaios de absorbância.

Fonte: Elaborada pelo autor.

Os ensaios foram realizados seguindo protocolos já estabelecidos na literatura.<sup>136, 138</sup> A temperatura do leitor de placas foi pré-ajustada para 37 °C e alíquotas de 20 µL de soluções preparadas para cada ligante (10 vezes concentradas e em 10% DMSO) foram transferidas para cada poço. A reação de polimerização foi preparada pela adição de GTP e tubulina ao tampão de polimerização a baixa temperatura, levando a uma mistura com 3 mg/mL de tubulina e 1 mM de GTP. Imediatamente após preparar a mistura, alíquotas de 180 µL da reação foram adicionadas à placa utilizando-se uma pipeta multicanal. A placa de leitura foi inserida no espectrofotômetro e o processo de leitura do espalhamento pela absorbância em 340 nm foi iniciado. Foram executados 121 ciclos de medidas dos ensaios da placa (começando em  $t = 0$  s), com intervalos de aproximadamente 30 s entre duas leituras do mesmo poço. Os ensaios foram realizados em triplicata, utilizando 3 poços adjacentes.

## 1.3.3.6 Identificação do Mecanismo Molecular de Compostos Moduladores de Tubulina

A presença de vários sítios possíveis de modulação na tubulina, tanto na desestabilização quanto na estabilização de microtúbulos, torna necessário o desenvolvimento de outros estudos para identificar o mecanismo molecular envolvido na atividade observada para um composto. Um método de identificação do sítio de ligação bastante simples e escolhido para ser desenvolvido neste projeto consiste em analisar a capacidade de um composto teste em reduzir o sinal de uma molécula fluorescente que apresente atividade de modulação conhecida frente à tubulina.<sup>64, 136–139</sup> Os compostos fluorescentes descritos na literatura utilizados para monitorar o sinal de ligação aos microtúbulos são geralmente os próprios moduladores já conhecidos (como o paclitaxel, a colchicina e a vimblastina) fusionados à uma molécula com alta capacidade de emitir luz fluorescente em um intervalo específico de comprimento de onda.<sup>140</sup> Os compostos fluorescentes comumente utilizados para monitorar a ligação de estabilizadores de microtúbulos no sítio do paclitaxel (Taxol<sup>®</sup>) são o flutax-1 e flutax-2, e correspondem ao paclitaxel fusionado a derivados do marcador fluorescente fluoresceína (Figura 22A).<sup>137, 141</sup> A interação com o sítio da vimblastina também pode ser monitorada utilizando um derivado deste composto ligado ao grupo fluorescente BODIPY (boro-dipirrometano) conhecido como BODIPY-Vimblastina (Figura 22B).<sup>64, 137–139</sup>

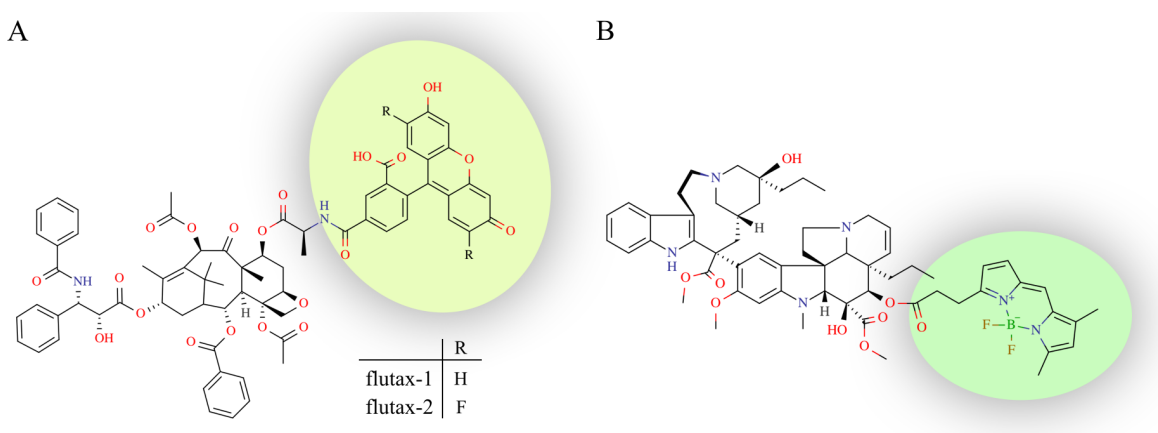


Figura 22 – (A) Estruturas dos compostos flutax-1 e flutax-2 derivados pela ligação do paclitaxel ao marcador fluorescente fluoresceína e utilizados para monitorar a ligação ao sítio do paclitaxel, responsável por estabilizar os microtúbulos de tubulina. (B) Estrutura do composto BODIPY-vimblastina utilizados para monitorar a ligação ao sítio da vimblastina, responsável por inibir a polimerização de tubulina.

Fonte: Elaborada pelo autor.

Analogamente, a interação de compostos no sítio de ligação da colchicina (selecionado para os estudos deste trabalho) pode ser verificada por ensaios utilizando a colchicina fluorescente (colchicina ligada a um derivado do marcador fluoresceína, Figura 23A).<sup>64, 139, 140, 142</sup> A colchicina fluorescente apresenta um máximo de absorção na faixa de 485 nm e emissão na faixa de 500-540 nm (Figura 23B).<sup>138, 142</sup>

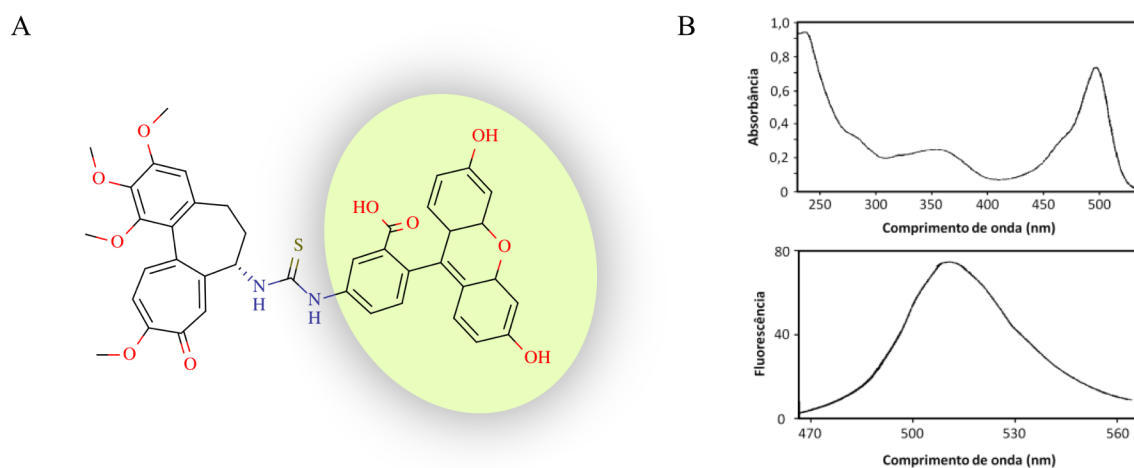


Figura 23 – (A) Estrutura da colchicina fluorescente utilizada neste trabalho e obtida pela ligação da colchicina ao marcador fluorescente fluoresceína. (B) Espectros de absorvância e emissão da colchicina fluorescente.<sup>142</sup>

Fonte: Elaborada pelo autor.

A verificação do sítio da colchicina como o sítio de ligação do composto analisado pode ser realizada pela comparação entre um ensaio na presença do ligante e um ensaio padrão na ausência do mesmo, monitorando a emissão em 535 nm após excitação a 485 nm. Para realizar o ensaio uma amostra de tubulina é incubada na presença de colchicina fluorescente. O composto de estudo é então adicionado e a mistura é incubada novamente (Figura 24A). A solução incubada é aplicada a uma coluna de exclusão molecular, responsável por separar a tubulina dos outros componentes da amostra, e frações são coletadas em uma placa contendo vários poços (Figura 24B). Ao fim da coleta, a placa é inserida em um fluorímetro e o sinal da colchicina fluorescente (emissão em 535 nm) é medido em cada uma das frações coletadas (Figura 24C). Como a colchicina fluorescente possui capacidade de se ligar ao sítio de ligação da colchicina, a tubulina coletada nas primeiras frações da placa em um ensaio padrão (sem um ligante teste) apresenta-se saturada com o marcador (colchicina fluorescente) e um pico de emissão é observado (pico I, Figura 24C). Nas frações seguintes, este sinal diminui juntamente com o término da eluição de tubulina, mas volta a subir quando a colchicina fluorescente excedente, não ligada e livre em solução, é eluída (pico II, Figura 24C). Na presença de um ligante teste incapaz de se ligar ao sítio da colchicina, o sinal do ensaio é o mesmo que o apresentado no ensaio padrão (Figura 25A), uma vez que sua presença não interfere a associação da colchicina fluorescente ao seu sítio de ligação. Porém, se o composto é capaz de se ligar ao sítio da colchicina, o deslocamento da colchicina fluorescente devido à competição por ligação com este composto de estudo resulta em uma diminuição do sinal de emissão do primeiro pico (Figura 25B).

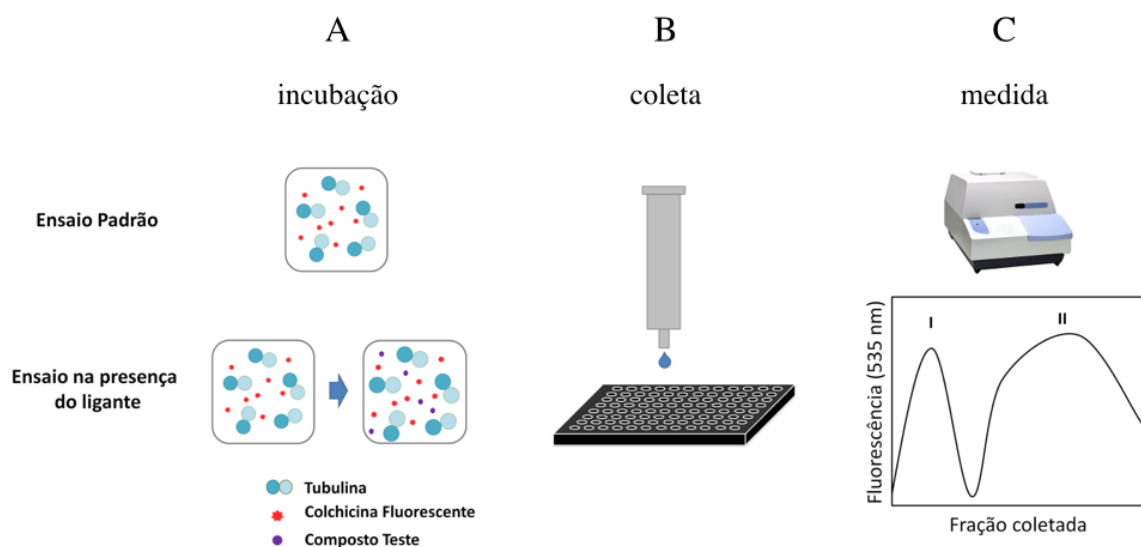


Figura 24 – (A) Descrição do ensaio competitivo para determinar a interação de um ligante com o sítio da colchicina. (A) Incubação da tubulina com colchicina fluorescente (controle negativo) e com colchicina fluorescente junto ao composto de estudo. (B) Coleta das frações da amostra em uma placa de fluorescência após aplicação em uma coluna cromatográfica de exclusão molecular. Nesta etapa a proteína tubulina ligada à colchicina fluorescente (e possivelmente ao composto) é separada dos outros componentes da solução, sendo eluída nas primeiras frações (C) Medida da emissão em 535 nm das frações coletadas em um fluorímetro. No ensaio padrão, observam-se os picos correspondentes à colchicina fluorescente ligada à tubulina (pico I) e livre em solução (pico II).

Fonte: Elaborada pelo autor.

### 1.3.3.7 Identificação do Mecanismo Molecular de Modulação de Microtúbulos

Os ensaios para identificar o sítio de ligação dos compostos teste foram realizados monitorando o sinal da colchicina fluorescente presente em frações de soluções contendo tubulina, colchicina fluorescente e o composto teste. Este ensaio foi estabelecido baseando-se em ensaios descritos na literatura.<sup>138</sup> A resina utilizada para separar as frações coletadas foi a de exclusão molecular cromatográfica Sephadex G-25 Medium (Sigma-Aldrich, Saint Louis, MO Estados Unidos). Esta resina foi empacotada em uma coluna cromatográfica resultando em um volume de resina de 2 mL com um comprimento (caminho de separação) de 5 cm.

O tampão utilizado no ensaio para incubar a tubulina com colchicina fluorescente e os compostos de estudo apresenta os seguintes componentes: 0,25 mM PIPES, 0,05 mM GTP, 0,25 mM MgCl<sub>2</sub>; em pH 6,9. Cada reação foi realizada com 2 mL de tubulina a 1 mg/ml no tampão descrito. A tubulina utilizada para o ensaio foi a mesma do ensaio de polimerização monitorada por espalhamento (Cytoskeleton, código HTS02-B, tubo de 44 mg de proteína liofilizada, 97% de pureza).

Na primeira etapa do ensaio, 200  $\mu$ L de tubulina a 10 mg/ml, solubilizada no tam-



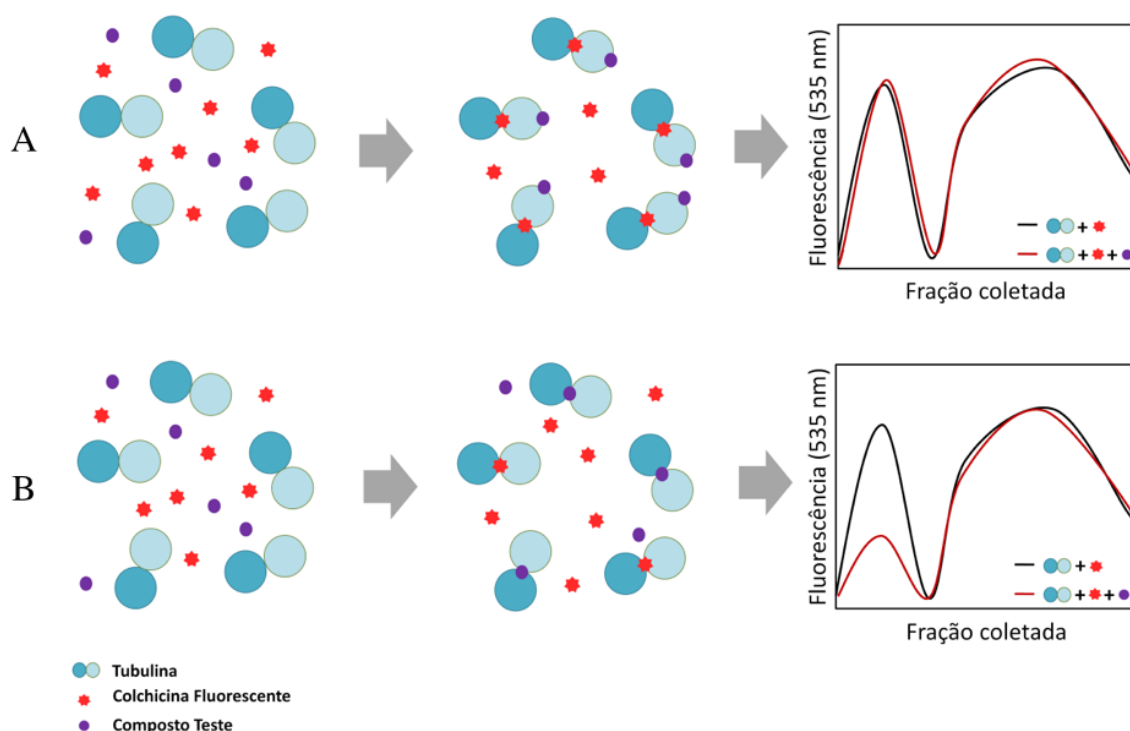


Figura 25 – (A) Ensaio em que o composto de estudo inibidor da polimerização de microtúbulos não interage com o sítio da colchicina, mas se liga a algum outro sítio na tubulina. O sinal da fração de eluição da tubulina é intenso e similar ao do ensaio padrão. (B) Ensaio em que o composto avaliado se liga ao sítio da colchicina, apresentando uma competitividade pelo sítio que leva a uma diminuição do sinal de emissão da colchicina fluorescente na fração de eluição da tubulina, em comparação ao ensaio padrão.

Fonte: Elaborada pelo autor.

pão do ensaio e em baixa temperatura (0 °C, no gelo) foram adicionados em um tubo. Em seguida, 1  $\mu\text{L}$  de GTP na concentração de 100 mM foi adicionado à solução seguido da adição e leve agitação de 1,8 mL do tampão do ensaio. Em seguida, 20  $\mu\text{L}$  de colchicina fluorescente recém-preparada (a 50  $\mu\text{M}$ ) foram adicionados. A mistura destes componentes resultou em 2 mL de uma solução contendo tubulina na concentração de 1 mg/mL e 0,05 mM de GTP, além do próprio tampão. Esta solução foi incubada a 37 °C durante 30 min. Após este intervalo de tempo, 20  $\mu\text{L}$  do composto teste foram adicionados, resultando em uma concentração de 25  $\mu\text{M}$  e 1% DMSO. A solução foi levemente agitada até a completa solubilização do composto e novamente incubada por 30 minutos.

Ao final da etapa de incubação, a mistura resultante foi aplicada na coluna cromatográfica contendo a resina G25 e frações de aproximadamente 150  $\mu\text{L}$  (correspondente a 6 gotas) foram coletadas em cada poço de uma placa de fluorescência de 96 poços (OptiPlate-1536 F). Ao término da eluição da amostra, volumes de tampão foram aplicados na resina e coletados até completar a total eluição de 2 volumes da amostra. As medidas da fluorescência das frações coletadas foram realizadas no mesmo equipamento utilizado nos ensaios.

de polimerização de microtúbulos baseado em fluorescência (Victor 3 1420) através da excitação em 485 nm e da monitoração de emissão em 535 nm (Figura 19).

## 1.4 Resultados e Discussão

### 1.4.1 Avaliação das Propriedades Físico-Químicas da Coleção de Compostos

Em uma primeira etapa do processo de triagem virtual, as propriedades físico-químicas de todas as moléculas presentes na coleção disponibilizada pelo LSPNF foram analisadas através da submissão das suas estruturas moleculares ao servidor Molinspiration.<sup>149</sup> Estas propriedades foram comparadas com as de compostos já conhecidos e capazes de interagir com microtúbulos (paclitaxel, vimblastina e colchicina), como critério na seleção de um dos três sítios para realização de estudos iniciais de docagem molecular e bioquímicos frente à tubulina. Uma vez que as propriedades físico-químicas de um composto são fatores importantes que determinam o seu perfil farmacocinético, moléculas que apresentam características físico-químicas muito diferentes das apresentadas por moléculas bioativas de fármacos com atividade biológica comprovada possuem baixa probabilidade de exibir um perfil farmacocinético parecido e adequado para aplicações terapêuticas frente ao mesmo alvo molecular.

No gráfico da Figura 26, são apresentados os valores calculados para as propriedades físico-químicas da base de compostos são apresentados: (i) coeficiente de partição (ou lipofilicidade,  $\log P$ ), (ii) área de superfície polar topológica (TPSA, em  $\text{Å}^2$ ), (iii) massa molecular (em daltons, Da), (iv) número de grupos aceptores de ligação de hidrogênio, (v) número de grupos doadores de ligação de hidrogênio e (vi) número de violações da regra dos cinco de Lipinsky.<sup>98, 143</sup> Estas propriedades são comparadas às de moléculas já conhecidas capazes de interagir com cada sítio do dímero de tubulina (paclitaxel, vimblastina e colchicina, Figura 26).

Uma análise das curvas do gráfico de propriedades apresentado na Figura 26 mostra que, dentre as moléculas conhecidas para cada sítio de ligação e avaliadas neste estudo, a colchicina é a que apresentou maior similaridade com os compostos presentes na base de dados. As moléculas da base de dados possuem massas moleculares distribuídas entre 100 e 400 Da (média de 265,8 Da), chegando a valores similares com a massa molecular da colchicina (de 399,4 Da), ao passo que os compostos paclitaxel e vimblastina apresentam massas moleculares de 853,92 e 825,0 Da; respectivamente. Outra característica bastante distinta apresentada pelos compostos paclitaxel e vimblastina foi a alta lipofilicidade. Enquanto os compostos da base apresentam um valor de  $\log P$  médio de 2,4, a colchicina possui um valor de apenas 1,1, e o paclitaxel e a vimblastina apresentaram altos valores de  $\log P$  de 5,0

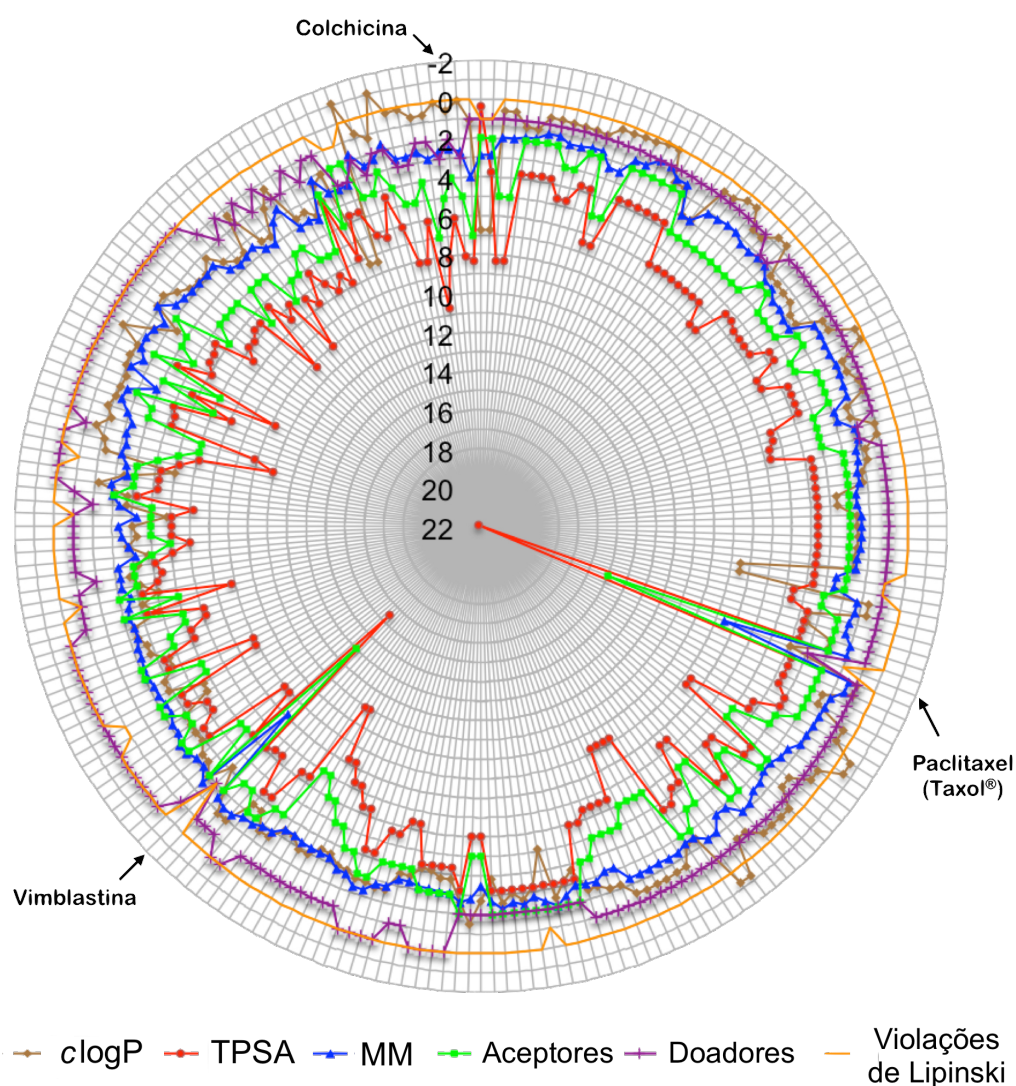


Figura 26 – Propriedades físico-químicas previstas para os compostos da base de dados inicial sintetizada pelo LSPNF e para os fármacos paclitaxel, vimblastina e colchicina.  $clogP$ : coeficiente de partição calculado. TPSA: área de superfície polar topológica (em  $\text{\AA}^2$ ). MM: massa molecular (em Da).

Fonte: Elaborada pelo autor.

e 6,1.

Dessa forma, a partir da comparação das propriedades físico-químicas da base de dados e de compostos já existentes capazes de interagir com microtúbulos, o sítio da colchicina foi escolhido como o mais adequado para a realização de estudos iniciais de modelagem molecular e ensaios bioquímicos.

### 1.4.2 Triagem Virtual de Compostos para Avaliação Bioquímica

A coleção de compostos sintéticos disponibilizados pelo LSPNF foi utilizada para construir uma base de dados e realizar estudos de docagem molecular frente às estruturas cristalográficas de dímeros de tubulina. O sítio da colchicina foi escolhido como o de maior probabilidade de interagir com os compostos da base, baseando-se nos resultados de estudos de propriedades físico-químicas (Figura 26).

Os 10 primeiros compostos com melhor classificação em uma pontuação consenso, considerando-se os 3 programas de docagem molecular utilizados (GOLD, AutoDock Vina e Surflex-Dock) para cada um dos sítios de ligação, são apresentadas na Tabela 3.

Dentre os compostos da base de dados identificados pelos estudos de triagem virtual, os derivados indólicos (ciclopenta- $\beta$ -indóis) foram os que apresentaram maior predominância, correspondendo a 60% dos compostos com melhor classificação (*i.e.*, maior probabilidade de interagir com o sítio da colchicina).

Considerando a disponibilidade inicial destes e de outros compostos da mesma classe, os 10 compostos selecionados por triagem virtual e mais 4 análogos de mesma classe com alta similaridade estrutural foram selecionados para ensaios bioquímicos de modulação da polimerização de tubulina em concentração única. As estruturas dos compostos selecionados para avaliação são apresentadas na Figura 27.

### 1.4.3 Triagem Inicial de Moduladores de Polimerização de Microtúbulos através de Ensaios de Polimerização baseados em Fluorescência

Os compostos selecionados a partir dos estudos de triagem virtual (14 moléculas, Figura 27) foram testados em ensaios de polimerização de tubulina baseados em fluorescência. Todos os compostos foram avaliados em triplicata com concentração única de 250  $\mu$ M. Os resultados são apresentados na Tabela 4. As curvas de polimerização de cada ensaio são apresentadas no Anexo B (Figuras 73 e 74).

Os resultados obtidos na presença de derivados indólicos indicam que alguns compostos desta classe apresentaram uma considerável capacidade de inibição da polimerização de microtúbulos na concentração testada (Tabela 4 e Anexo B). Os ensaios com os compostos derivado indólico 1 apresentaram uma considerável diminuição no sinal de fluorescência ao longo do tempo em comparação aos ensaios padrão (utilizando DMSO como controle negativo de atividade, Figura Anexo 1 A), indicando uma redução no nível de polimerização de tubulina ao longo do tempo. Esse comportamento é ainda mais pronunciado na presença dos derivados indólicos **6**, **7** e **14** (Figuras Anexo B). Surpreendentemente, a formação de tubulina parece ser totalmente inibida na presença do derivado indólico **8** (Figura 73F), resultando em uma curva que coincide com a obtida na presença de 20  $\mu$ M de colchicina. Por

Tabela 3 – Classificação dos 10 compostos que apresentaram melhor pontuação consenso durante o processo de docagem molecular no sítio da colchicina utilizando os 3 programas de docagem molecular: GOLD, AutoDock Vina e Surflex-Dock.

Sítio	Colchicina			Pontuação		
	Classificação	Composto	GOLD (Fitness)	AutoDock Vina $\log K_d$ (kcal/mol)	SurflexDock $\log K_d$ (kcal/mol)	
1	derivado indólico 7		55,85	-9,3	9,81	
2	1,8-Naftiridina 4		54,45	-9,4	7,36	
3	derivado indólico 5		52,38	-8,8	7,09	
4	derivado indólico 6		51,90	-9,0	6,39	
5	derivado indólico 8		68,17	-8,2	6,37	
6	$\beta$ -Cetoster 17		53,89	-8,5	6,18	
7	derivado indólico 14		50,65	-8,4	7,08	
8	episulosina		69,35	-5,6	9,96	
9	derivado indólico 1		45,42	-8,8	5,02	
10	espirocicloexadienona 9		42,91	-8,1	5,54	
	colchicina		58,55	-8,4	8,11	

Fonte: Elaborada pelo autor

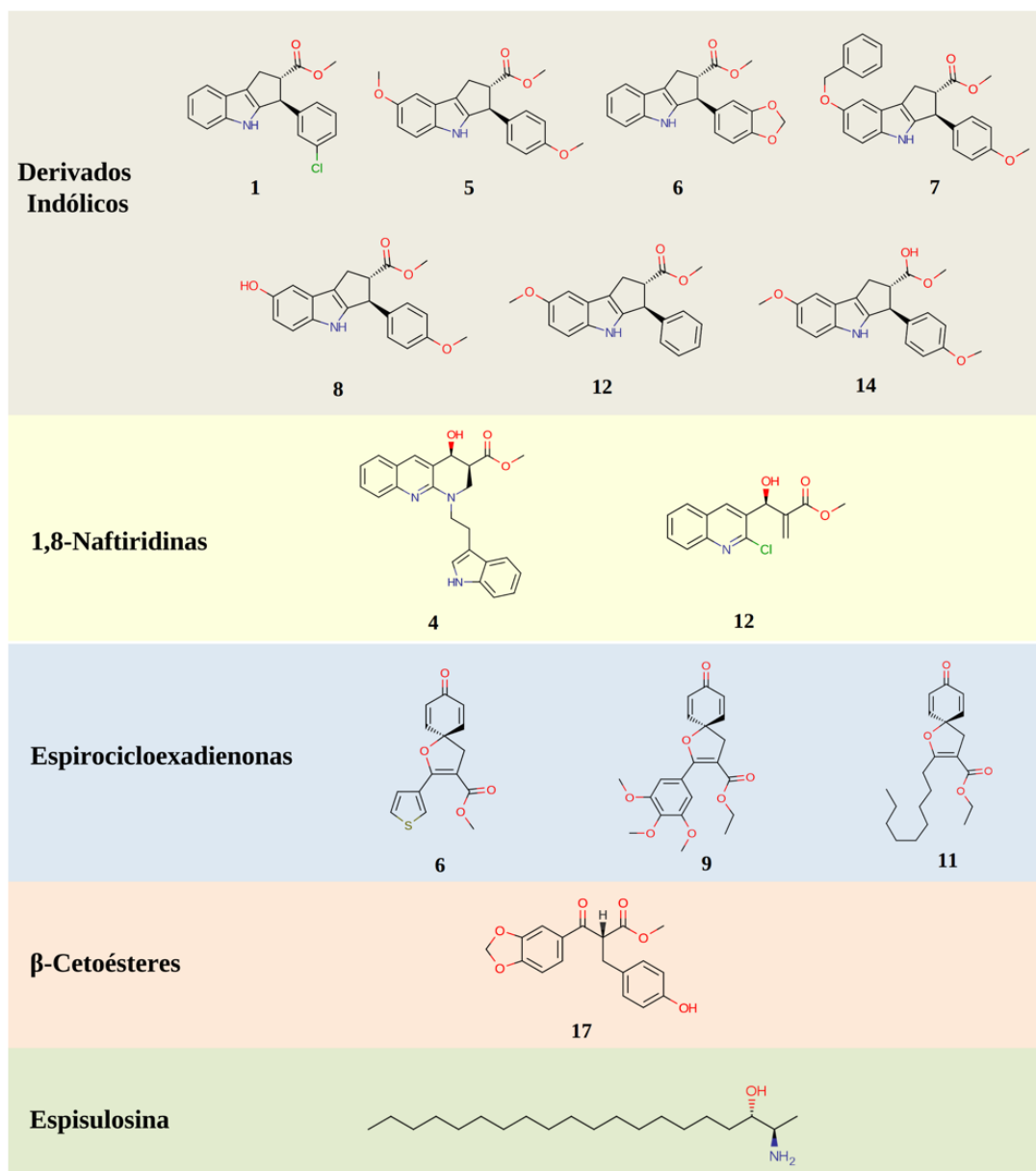


Figura 27 – Estruturas moleculares dos compostos selecionados a partir dos resultados de triagem virtual por docagem molecular (Tabela 3) para a realização de ensaios bioquímicos de polimerização de tubulina *in vitro*.

Fonte: Elaborada pelo autor.

outro lado, nenhuma atividade substancial de modulação da tubulina foi observada para os derivados indólicos **5** e **12** (Figuras Anexo B).

Os ensaios de polimerização na presença de derivados de naftiridinas (Anexo B) permitiram identificar uma capacidade de inibição de polimerização para a 1,8-naftiridina **4** (Anexo B), apesar da inibição não ter sido total, mesmo com a alta concentração de 250  $\mu\text{M}$ . O padrão apresentado pela 1,8-naftiridina **12** não foi conclusivo (Anexo B), mas considerado similar ao controle, na ausência de composto teste (Anexo B).

Tabela 4 – Atividade dos compostos selecionados por triagem virtual em modular a polimerização de microtúbulos em ensaios bioquímicos.

<b>Composto</b>	<b>Atividade</b>
derivado indólico 1	inibição moderada
derivado indólico 5	ausente
derivado indólico 6	inibição moderada
derivado indólico 7	inibição moderada
derivado indólico 8	inibição alta
derivado indólico 12	ausente
derivado indólico 14	inibição baixa
1-8-natdiridina 4	inibição baixa
1-8-naftiridina 12	inibição baixa
Espirocicloexadienona 6	ausente
Espirocicloexadienona 9	inibição alta
Espirocicloexadienona 10	inibição moderada
$\beta$ -cetoester 17	inibição baixa
Espisulosina	pequeno aumento

Fonte: Elaborada pelo autor

A série de compostos pertencentes à classe de espirocicloexadienonas apresentou um padrão promissor de inibição da polimerização de microtúbulos (Anexo B). Dentre estes compostos, a espirocicloexadienona **9** inibiu totalmente a polimerização de tubulina em todos os ensaios na concentração testada (74E), de forma similar ao derivado indólico 8 (Anexo B). Ensaios na presença dos compostos  $\beta$ -cetoester **17** e espisulosina apresentaram baixas atividades de inibição e de aumento da polimerização de microtúbulos, respectivamente (Anexo B).

#### 1.4.4 Avaliação Quantitativa de Inibidores da Polimerização de Microtúbulos por meio de Ensaios Baseados em Fluorescência

A partir dos resultados obtidos nos ensaios de polimerização na triagem bioquímica, os compostos: derivado indólico 8 (Tabela 4) e espirocicloexadienona 9 (Tabela 4) foram escolhidos para novos ensaios baseados em concentração-resposta, buscando a determinação de valores de  $IC_{50}$  (concentração necessária para se obter 50% de inibição da atividade biológica observada, neste caso a polimerização de microtúbulos). As determinações dos valores de  $IC_{50}$  foram realizadas repetindo-se o ensaio anterior, com a variação da concentração dos compostos em estudo.

Inicialmente, o valor de  $IC_{50}$  foi determinado, como um padrão inicial do perfil de inibição de microtúbulos pela interação com o sítio da colchicina e um parâmetro de comparação de potência. As curvas de polimerização de microtúbulos na presença de diferentes concentrações de colchicina (entre 20 e 0,1  $\mu M$ ) e a curva de determinação do valor de  $IC_{50}$

são apresentadas na Figura 28. O valor de  $IC_{50}$  determinado para a colchicina em nossos experimentos é bastante próximo do valor reportado na literatura, de  $5 \mu M$ .<sup>144, 145</sup>

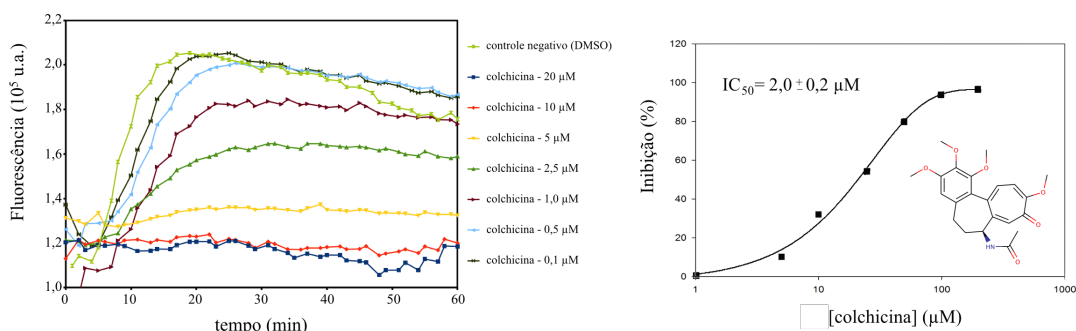


Figura 28 – Curva de concentração-resposta para a colchicina. O valor de  $IC_{50}$  de  $2 \mu M$  determinado para os ensaios de polimerização de microtúbulos baseados em fluorescência concorda com os valores reportados na literatura.<sup>144, 145</sup>

Fonte: Elaborada pelo autor.

Em seguida, os valores de  $IC_{50}$  dos compostos selecionados (derivados indólico **8** e espirocicloexadienona **9**) foram determinados utilizando o mesmo procedimento. A concentração do derivado indólico **8** foi variada entre 250 e  $10 \mu M$  (em  $500 \mu M$ , o composto apresentou-se insolúvel em 10% DMSO) e a concentração da espirocicloexadienona **9** foi variada de 500 a  $5 \mu M$ . Os valores de atividade dos compostos e suas respectivas concentrações foram utilizados na obtenção de gráficos com curvas de concentração-resposta e na determinação dos valores de  $IC_{50}$ . As curvas dos ensaios para os compostos derivado indólico **8** e espirocicloexadienona **9**, e os aos valores de  $IC_{50}$  determinados, são apresentadas nas Figuras 29 e 30, respectivamente.

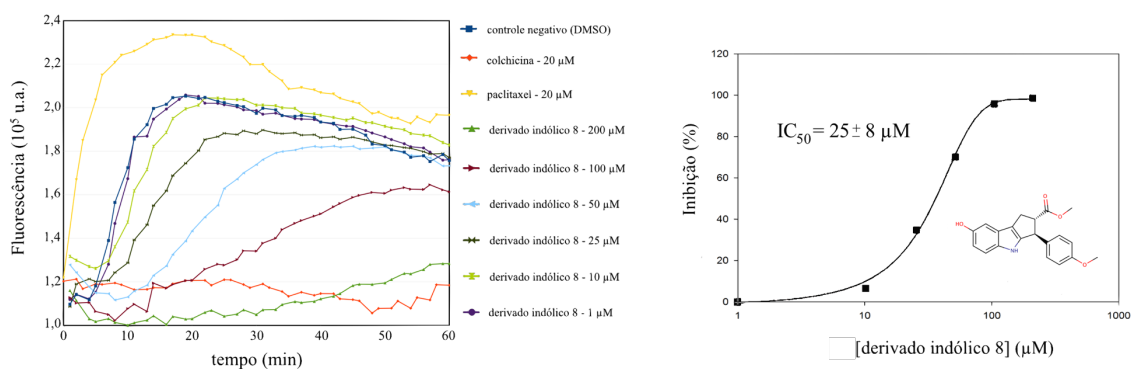


Figura 29 – Determinação do parâmetro quantitativo de modulação  $IC_{50}$  para o derivado indólico **8**, através de ensaios de polimerização de tubulina na presença de diferentes concentrações do composto. Curvas de polimerização de controle (colchicina e taxol a  $20 \mu M$  e DMSO) e na presença das seguintes concentrações do composto: 250, 100, 50, 25 e  $10 \mu M$ . Todas as curvas correspondem a valores médios de triplicatas. A curva concentração-resposta permitiu determinar o valor de  $IC_{50}$  igual a  $25 \pm 8 \mu M$ .

Fonte: Elaborada pelo autor.



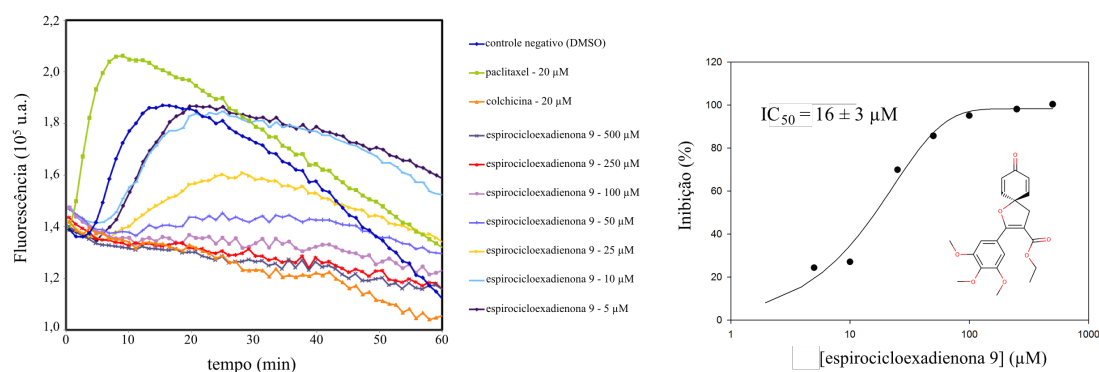


Figura 30 – Determinação do parâmetro quantitativo de modulação  $IC_{50}$  para a espirocicloexadienona **9**, através de ensaios de polimerização de tubulina na presença de diferentes concentrações do composto. Curvas de polimerização de controle (colchicina e taxol a  $20 \mu\text{M}$  e DMSO) e na presença das seguintes concentrações do composto: 500, 250, 100, 50, 25, 10 e  $5 \mu\text{M}$ . Todas as curvas correspondem a valores médios de triplicatas. A curva concentração-resposta permitiu determinar um valor de  $IC_{50}$  de  $16 \pm 3 \mu\text{M}$ .

Fonte: Elaborada pelo autor.

Os ensaios de determinação de  $IC_{50}$  de ambos os compostos derivado indólico **8** e espirocicloexadienona **9** confirmaram um comportamento concentração-resposta e a considerável capacidade destes compostos em inibir a formação de microtúbulos, com valores de  $IC_{50}$  abaixo de  $30 \mu\text{M}$ . Apesar do interessante potencial apresentado pela molécula espirocicloexadienona **9** (Figura 30), uma análise mais detalhada da literatura mostrou que compostos da classe espirocarbocíclicos apresentam uma reatividade considerável.<sup>146–148</sup> Devido à esta propriedade desfavorável para o planejamento de fármacos, a classe dos derivados indólicos foi escolhida como prioridade para estudos posteriores de planejamento.

Dessa forma, o derivado indólico **8** foi selecionado para novos ensaios de polimerização baseados em espalhamento e de competição frente à ao sítio da colchicina, com o propósito de confirmar a capacidade de modulação, e identificar o possível mecanismo de inibição desta classe.

#### 1.4.5 Planejamento de Novos Derivados Indólicos

Devido aos problemas de solubilidade observados para o derivado indólico **8**, novas modificações foram propostas a fim de se obter uma nova série de compostos mais solúveis para a realização de ensaios bioquímicos. Em uma primeira análise, a comparação estrutural dos derivados indólicos avaliados e das respectivas atividades apresentadas (Figura 27 e Tabela 4) permite observar uma característica estrutural que diferencia o derivado indólico **8** dos outros compostos da mesma classe: a presença de um substituinte hidroxila na posição 5 do anel benzênico do sistema indolólico (Figura 27). Esta hidroxila, que apresenta as funcionalidades de acceptor e doador de ligação de hidrogênio, está presente somente neste

composto, entre todos os avaliados para esta classe. Portanto, para confirmar a importância desta hidroxila na atividade observada e aumentar a solubilidade desta classe, dois novos compostos foram sintetizados e avaliados em ensaios de polimerização. Estes compostos foram denominados derivado indólico **novo 1** e derivado indólico **novo 2** e suas estruturas moleculares são apresentadas na Figura 31.

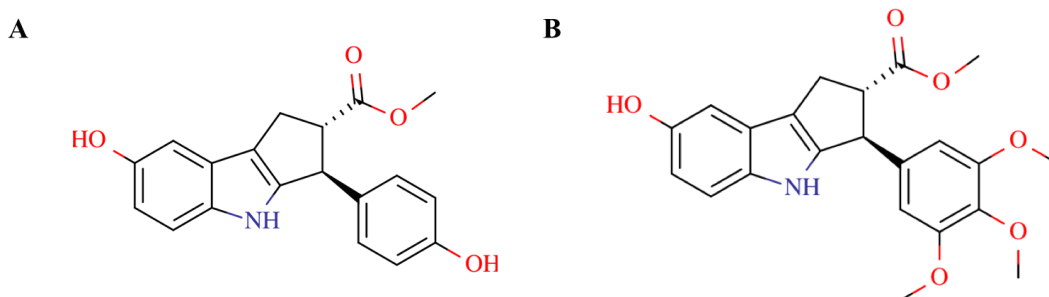


Figura 31 – Novos compostos propostos para ensaios de modulação da polimerização de tubulina. (A) Derivado indólico **novo 1** (B) Derivado indólico **novo 2**.

Fonte: Elaborada pelo autor.

Com a finalidade de estimar o melhor perfil de solubilidade dos derivados indólicos propostos, um estudo computacional de solubilidade foi desenvolvido utilizando os servidores Molinspiration e OSIRIS.<sup>149,150</sup> Os valores de  $clogP$  preditos ( $clogP$ ) por ambos os programas confirmaram um possível aumento considerável da solubilidade dos compostos propostos em relação aos avaliados anteriormente (Tabela 5), considerando a solubilidade de um composto como inversamente proporcional ao seu coeficiente de partição  $\log P$ .

Tabela 5 – Predição dos valores de coeficiente de partição ( $clogP$ ) para os derivados indólicos.

Composto	$clogP$ (Molinspiration)	$clogP$ (OSIRIS)
Derivado indólico 1	4,82	3,87
Derivado indólico 6	4,06	3,36
Derivado indólico 7	5,85	4,20
Derivado indólico 8	3,72	2,85
Derivado indólico novo 1	3,18	2,66
Derivado indólico novo 2	3,29	2,64

Fonte: Elaborada pelo autor

Os compostos propostos (derivados indólicos **novo 1** e **novo 2**, Figura 31) foram sintetizados e avaliados para a determinação de  $IC_{50}$ . As concentrações testadas variaram de 500 a 1  $\mu M$ , e os compostos apresentaram-se solúveis em todas as concentrações empregadas. Todos os ensaios foram realizados em triplicata.

O composto derivado indólico **novo 1** não apresentou atividade substancial em todas as concentrações testadas (Figura 32). A comparação entre este composto e o derivado indólico **8**, mostra que a única diferença do novo composto proposto (derivado indólico

**novo 1**) é a presença de um grupo hidroxila ao invés de um metoxila na posição *para* da fenila substituinte na posição 3 do anel pentano fusionado ao indol (ver Figuras 26 e 30 A). Esse resultado permitiu classificar como prejudicial à atividade a substituição de um grupo somente acceptor de ligação de hidrogênio por um grupo dual, tanto acceptor como doador de ligação de hidrogênio na posição descrita, uma vez que essa modificação resultou na total perda de atividade de inibição.

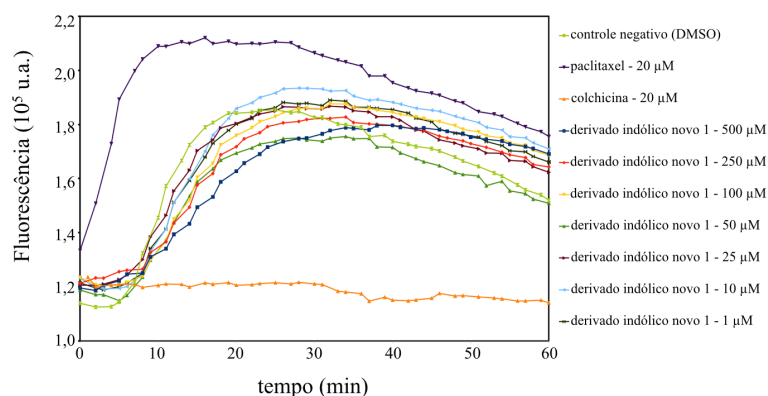


Figura 32 – Ensaios de modulação da polimerização de tubulina na presença de diferentes concentrações do derivado indólico **novo 1** mostram que este composto não apresenta atividade considerável de modulação da polimerização de microtúbulos.

Fonte: Elaborada pelo autor.

A determinação do valor de  $IC_{50}$  para o composto derivado indólico **novo 2** (Figura 33) mostrou que o mesmo apresenta uma atividade inibitória muito similar à apresentada pelo derivado indólico **8**, com um  $IC_{50}$  de  $25 \pm 3 \mu M$ . A comparação entre as estruturas do derivado indólico **8** e o derivado indólico **novo 2** proposto (Figuras 26 e 30 B) mostra que o último apresenta a presença adicional de 2 grupos metoxila nas posições *meta* da fenila pertencente ao anel pentano, fusionado ao grupo indol.

#### 1.4.6 Ensaios de Polimerização baseados em Espalhamento

Para comprovar novamente a capacidade de inibição de polimerização de tubulina apresentada pelos compostos derivado indólico **8** e derivado indólico **novo 2** nos ensaios de fluorescência na presença de DAPI, ensaios de polimerização baseados em espalhamento de luz na presença destes ligantes foram desenvolvidos. Nestes ensaios o composto paclitaxel (taxol) foi utilizado como padrão positivo para o máximo de polimerização, com concentração de  $20 \mu M$ . A colchicina, por sua vez, foi utilizada como padrão positivo para inibição de polimerização, com uma concentração de  $5 \mu M$ . O mesmo volume contendo DMSO 1%, presente em todos os compostos (solubilizados em DMSO 1%) foi utilizado como controle negativo para compostos que não modulam a formação de microtúbulos. Os compostos fo-

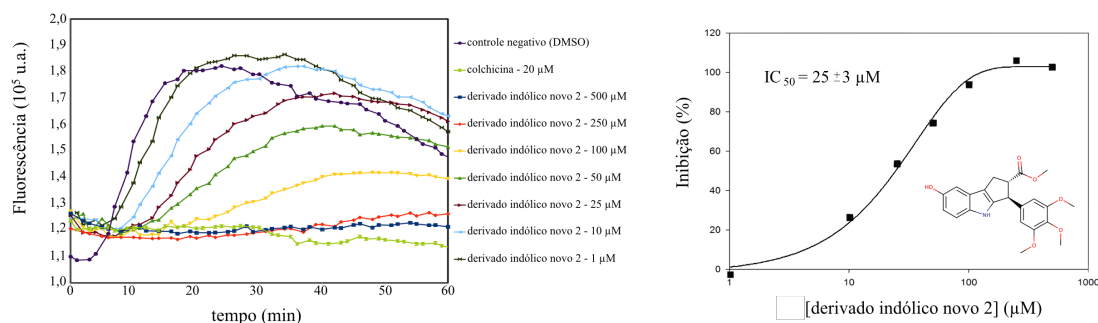


Figura 33 – Determinação dos valores de  $\text{IC}_{50}$  para o derivado indólico **novo 2**. (A) Estrutura do derivado indólico **novo 2**. (B) Ensaios de modulação da polimerização de tubulina na presença de diferentes concentrações do derivado indólico **novo 2** mostram que este composto apresenta uma atividade de modulação da polimerização de microtúbulos que obedece a um comportamento do tipo concentração-resposta. (C) A curva concentração-resposta permitiu determinar um valor de  $\text{IC}_{50}$  de  $25 \pm 3 \mu\text{M}$ .

Fonte: Elaborada pelo autor.

ram analisados em concentrações de  $25 \mu\text{M}$  e os resultados (médias de três ensaios idênticos), junto com as curvas de controle, são apresentados na Figura 34.

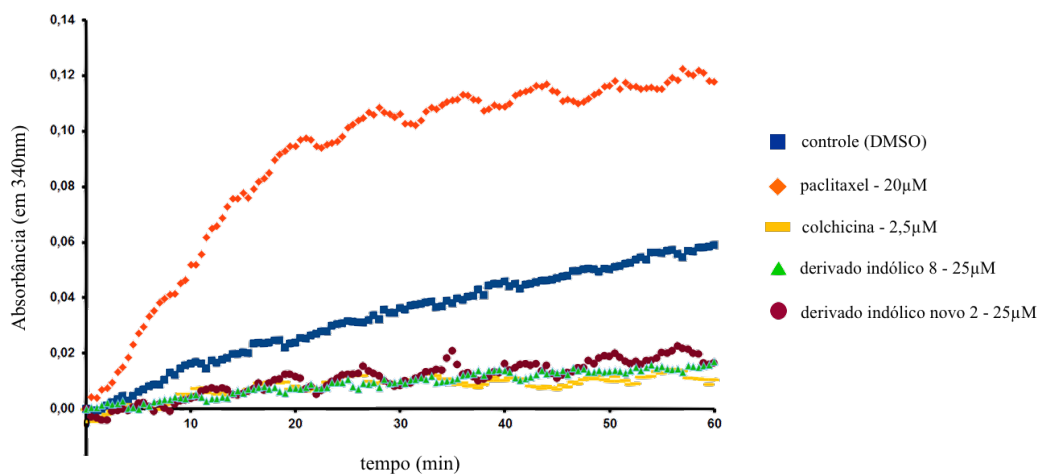


Figura 34 – Ensaio de polimerização de tubulina na presença dos compostos derivado indólico **8** e derivado indólico **novo 2**, com concentrações de  $25 \mu\text{M}$  cada. As curvas correspondem aos valores médios de triplicatas.

Fonte: Elaborada pelo autor.

Como pode ser observado na Figura 34, os ensaios baseados em espalhamento com os compostos derivado indólico **8** e derivado indólico **novo 2** apresentaram um perfil de inibição da polimerização de tubulina muito similar ao do ensaio na presença de colchicina. Esses resultados concordam com os resultados obtidos nos ensaios de fluorescência e reafirmam a capacidade destes compostos em inibir a polimerização de tubulina em concentrações na faixa de micromolar.

### 1.4.7 Estudos de Migração Celular

Para avaliar o efeito dos compostos identificados com atividade de inibição de polimerização de microtúbulos em células cancerígenas, os derivados indólicos **8**, **novo 1** e **novo 2** foram submetidos a ensaios de migração celular em células cancerígenas MDA-MB-231 (células humanas de câncer de mama). Estes ensaios foram realizados pela aluna de doutorado Wanessa F. Altei (processo FAPESP 2009/15320-5), sob orientação do professor Dr. Adriano D. Andricopulo. O ensaio, conhecido como *wound healing*, consiste em retirar uma área específica de uma cultura de células tumorais (causando uma fenda) e medir o preenchimento desta área pela migração das células ao longo do tempo. Esse procedimento é repetido na presença do ligante de estudo e o preenchimento da área é comparado ao apresentado no ensaio controle no mesmo intervalo de tempo. Nos ensaios desenvolvidos com os derivados indólicos, a área livre da cultura de células tumorais foi monitorada após 22 h, na presença de 50  $\mu\text{M}$  de cada um dos ligantes. Os resultados, apresentados na Figura 35, mostram que os derivados indólicos **8** e **novo 2** apresentam alta capacidade em inibir a migração celular na concentração avaliada (inibição de 75 e 70%, respectivamente), de forma similar à atividade da colchicina em 1  $\mu\text{M}$  (inibição de 74%). Por outro lado, o composto derivado indólico **novo 1** não apresentou atividade de inibição de migração. Estes resultados coincidem exatamente com os apresentados nos ensaios *in vitro* de polimerização da tubulina, em que somente o derivado **novo 1** não apresentou capacidade de modular a polimerização de tubulina (Figura 32).

### 1.4.8 Estudos de Imunofluorescência com Microtúbulos

A elevada atividade inibitória apresentada pelos derivados indólicos em ensaios de polimerização de microtúbulos e de migração celular estimularam o desenvolvimento de novos estudos para avaliar mais detalhadamente o comportamento estrutural dos microtúbulos no interior celular e na presença destes compostos. Dessa forma, ensaios de imunofluorescência foram realizados em colaboração com aluna de doutorado Wanessa F. Altei. Nestes ensaios, células cancerígenas MDA-MB-231 (as mesmas utilizadas nos ensaios de *wound-healing* já descritos) foram fixadas e incubadas por 1 h com um anticorpo primário capaz de se ligar à  $\alpha$ -tubulina. Em seguida, o sistema foi incubado com o anticorpo secundário Alexa Flúor 555, capaz de se ligar ao anticorpo de  $\alpha$ -tubulina e ser monitorado pela sua emissão em um comprimento de onda de 565 nm.<sup>151</sup>

A Figura 36 apresenta os resultados obtidos para os estudos de imunofluorescência. Enquanto células cultivadas na ausência dos compostos avaliados (controle negativo) apresentaram um citoesqueleto desenvolvido e com forma definida por uma rica rede de microtúbulos (Figura 36A), células cultivadas na presença de colchicina (controle positivo) em diferentes concentrações apresentaram um drástica redução no tamanho e complexi-

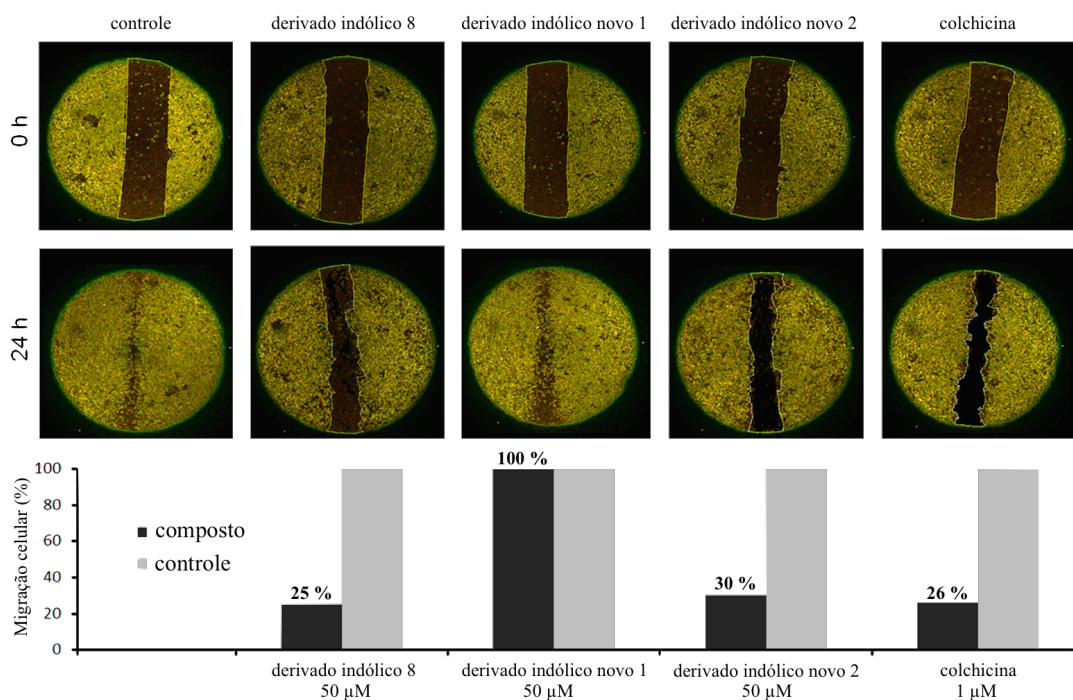


Figura 35 – Ensaio de migração de células tumorais (*wound healing*). Os ensaios foram executados na ausência de compostos moduladores de polimerização (controle negativo), na presença de 50  $\mu$ M dos compostos teste, identificados como capazes de inibir a polimerização de tubulina: derivado indólico **8**, derivado indólico **novo 1** e derivado indólico **novo 2**; e na presença de colchicina como controle positivo para inibição. Ensaio executado no LQMC pela aluna de doutorado Wanessa F. Alteí.

Fonte: Elaborada pelo autor.

dade do citoesqueleto (Figura 36B-C). Esta mesma capacidade de desorganizar e dissolver a rede de microtúbulos responsável pelo citoesqueleto intracelular foi observada para os derivados indólicos. O derivado indólico **8** (Figura 36 D-F) e o derivado indólico **novo 2** (Figura 36G-I) promoveram uma substancial redução da rede de microtúbulos em diversas concentrações na faixa de micromolar (10, 50 e 100 $\mu$ M).

Além da dissolução do citoesqueleto celular na presença destes inibidores de microtúbulos, é possível observar a presença de células com mais de um núcleo celular (marcados em azul, pela emissão do DAPI, Figura 36). Esta característica celular é comum às células sujeitas a compostos moderadores da polimerização de microtúbulos, uma vez que a incapacidade de estabelecer um equilíbrio funcional de polimerização e despolimerização de tubulina impede que a célula consiga se dividir completamente. Isto resulta em um aprisionamento da célula ao final do seu processo de divisão, como descrito detalhadamente na Figura 3B, levando à apoptose.

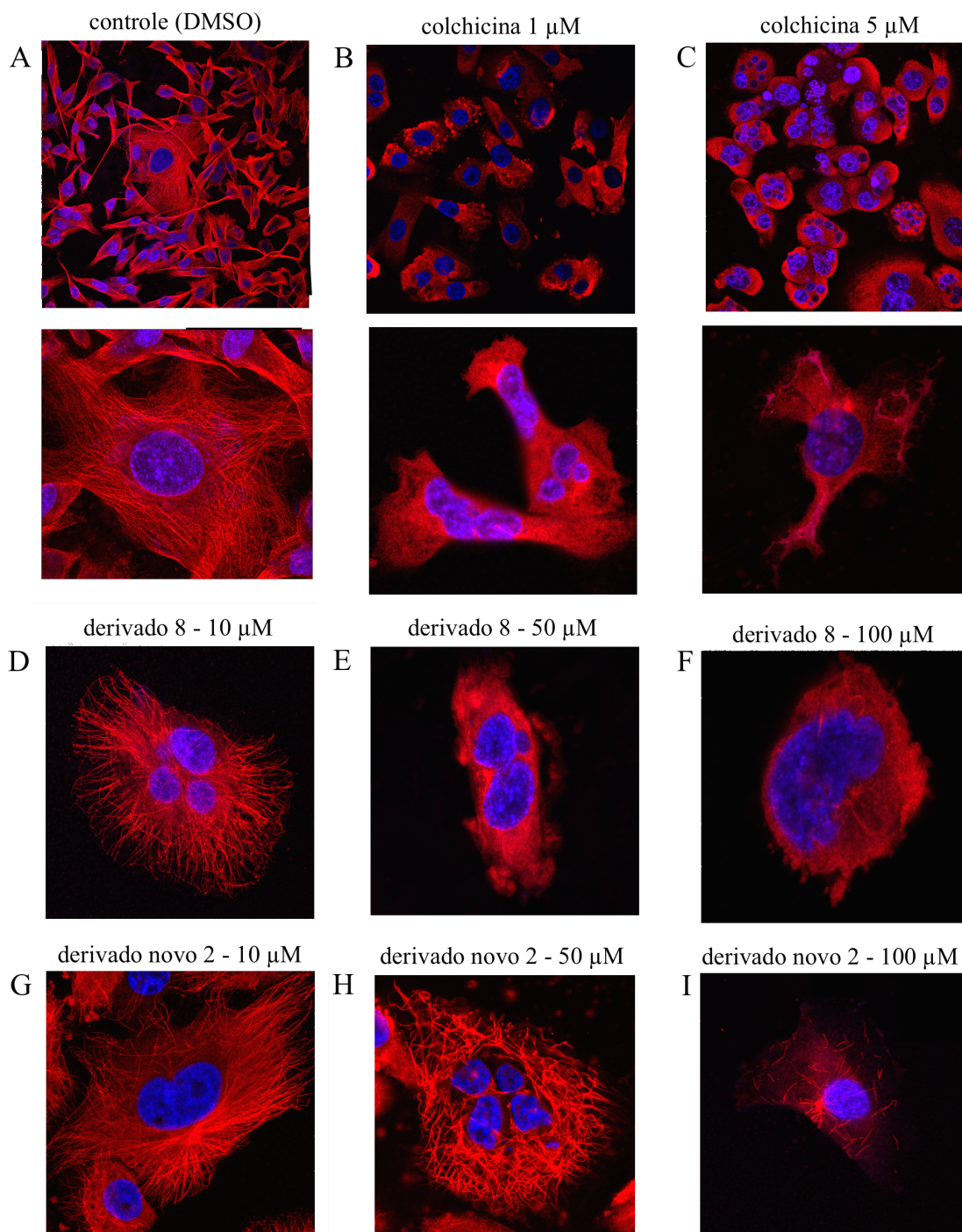


Figura 36 – Estudos celulares de imunofluorescência na presença dos derivados indólicos identificados como inibidores da polimerização de microtúbulos. Células cancerígenas MDA-MB-231 (células humanas de câncer de mama) foram tratadas com anticorpos para  $\alpha$ -tubulina (Sigma Aldrich, Saint Louis, MO, Estados Unidos) permitindo observar o citoesqueleto celular (em vermelho) em um microscópio confocal. Para distinguir o núcleo celular (em azul) o composto DAPI (4,6-diamidino-2-fenilindol) foi empregado. Ensaios executados no LQMC pela aluna de doutorado Wanessa F. Altei.

Fonte: Elaborada pelo autor.

### 1.4.9 Avaliação da Citotoxicidade dos Derivados Indólicos

A considerável capacidade dos derivados indólicos identificados em inibir a polimerização de microtúbulos, tanto em ensaios bioquímicos frente ao alvo molecular quanto em estudos celulares de imunofluorescência, juntamente com a capacidade observada de inibir a migração de células tumorais evidenciam o potencial destas moléculas para o desenvolvimento de novos agentes para o tratamento do câncer.

Para avaliar a toxicidade destes compostos em diferentes linhagens celulares, estudos de citotoxicidade foram desenvolvidos pela aluna de doutorado Luma G. Magalhães, sob orientação do professor Dr. Adriano D. Andricopulo. O método utilizado, conhecido como MTS, mede a viabilidade de uma cultura celular para análise espectrofotométrica da absorbância de um composto tetrazólio (3-(4,5-dimetiltiazol-2-il)-5-(3-carboximetoxifenil)-2-(4-sulfonil)-2H-tetrazólio), mais conhecido como MTS.<sup>152</sup> Células viáveis em cultura são capazes de reduzir o MTS e, portanto, a quantidade de MTS detectada é proporcional ao número de células em cultura.

O efeito dos derivados indólicos **8** e **novo 2** e da colchicina (controle), na viabilidade de fibroblastos de gengiva humanas (células saudáveis, linhagem FGH) e de células humanas de câncer de mama (linhagem MDA-MB-231) são apresentados na Tabela 6.

Tabela 6 – Avaliação da citotoxicidade dos derivados indólicos de estudo em células normais e cancerígenas.

<b>Composto</b>	<b>EC<sub>50</sub> MDA-MB-231<sup>a</sup></b>	<b>EC<sub>50</sub> FGH<sup>a</sup></b>
Derivado indólico 8	8 ± 2	>100
Derivado indólico novo 2	27 ± 12	>100
Colchicina	0,014 ± 0,001	<0,02

<sup>a</sup> Experimentos realizados em triplicata.

Fonte: Elaborada pelo autor

Enquanto a colchicina apresenta uma toxicidade extremamente alta (na faixa de 10<sup>1</sup> nM), tanto em células saudáveis quanto cancerígenas, os derivados indólicos apresentaram uma toxicidade muito maior em células cancerígenas. Apesar dos derivados indólicos terem apresentado valores de EC<sub>50</sub> (concentração necessária para promover a apoptose de 50% das células em cultura) na faixa de poucos micromolares (de 8 e 27 μM, respectivamente; Tabela 6), estes compostos se mostraram pouco tóxicos em concentrações menores que 100 μM (máxima concentração em que se apresentaram solúveis nas condições do ensaio).

Estes resultados demonstram a propriedade de seletividade destes compostos sintéticos, muito importante para o desenvolvimento de novas alternativas para o tratamento do câncer.



### 1.4.10 Estudos de Competição pelo Sítio da Colchicina

Os resultados obtidos por triagem virtual da base de dados e as atividades dos derivados indólicos **8** e **novo 2** em inibir a formação de microtúbulos, destacaram a presença dos substituintes metoxila e tri-metoxila (respectivamente) na fenila substituinte na posição 3 do anel pentano fundido ao indol. Nota-se ainda a presença deste mesmo grupamento na estrutura da colchicina (Figura 5 C), o que sugere que estes compostos possam exibir um mecanismo molecular similar de interação com o sítio de ligação da colchicina. Para confirmar esta hipótese, foram desenvolvidos ensaios de competição frente ao sítio da colchicina, empregando como marcador uma colchicina fluorescente sintética.

A capacidade dos derivados indólicos em se ligar ao sítio da colchicina foi determinada por meio de amostras de tubulina que foram incubadas com estes ligantes e com colchicina fluorescente (colchicina fusionada à molécula fluoresceína), sendo depois separadas por cromatografia de exclusão molecular. Desta forma, ao se comparar o sinal de fluorescência emitido pelas amostras contendo tubulina em complexo com colchicina fluorescente na ausência e na presença dos compostos de estudo, é possível determinar a capacidade de ligação ao sítio da colchicina pela variação do sinal emitido. Os resultados do ensaio de competição obtidos utilizando os derivados indólicos **8** e **novo 2** com uma concentração de  $25 \mu\text{M}$  e colchicina fluorescente a  $0,5 \mu\text{M}$ , junto com o ensaio padrão utilizando colchicina não fluorescente são apresentados na Figura 37.

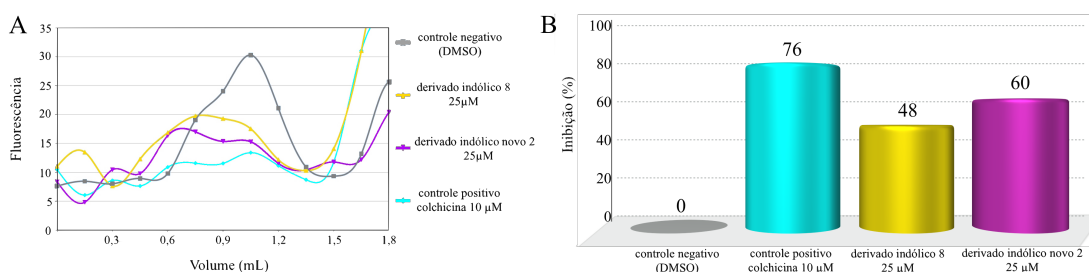


Figura 37 – (A) Emissão em 535 nm do primeiro pico coletado no ensaio de competição e correspondente à colchicina fluorescente ligada à tubulina. No primeiro ensaio (padrão, em cinza) a incubação foi realizada na presença somente de colchicina fluorescente a  $0,5 \mu\text{M}$ . No segundo ensaio (em azul) o mesmo procedimento foi repetido adicionando  $10 \mu\text{M}$  de colchicina não fluorescente junto ao marcador fluorescente. No terceiro ensaio (em amarelo) o mesmo ensaio foi repetido na presença de  $25 \mu\text{M}$  do derivado indólico **8** junto à colchicina fluorescente. No quarto ensaio (em roxo) o mesmo procedimento foi repetido na presença de  $25 \mu\text{M}$  do derivado indólico **novo 2**. (B) Em ambos os ensaios na presença da colchicina e dos derivados indólicos observa-se uma acentuada redução do sinal máximo de emissão, indicando que eles são capazes de se ligar ao sítio da colchicina.

Fonte: Elaborada pelo autor.

Como pode ser observado na Figura 37, o ensaio na presença do derivado indólico **8** resultou em uma considerável diminuição no sinal de emissão correspondente à colchi-

cina fluorescente ligada à tubulina. A redução do sinal de emissão da cochicina fluorescente presente na fração de eluição da tubulina pode ser melhor visualizada analisando os primeiros picos de emissão de cada ensaio (Figura 37). Em comparação ao sinal do ensaio controle, em que a tubulina é incubada somente com a colchicina fluorescente (Figura 37, barra azul), a presença de colchicina não fluorescente (10  $\mu\text{M}$ , Figura 37, barra cinza) levou a uma redução de aproximadamente 76% no sinal de emissão. Analogamente, a presença do derivado indólico **8** (25  $\mu\text{M}$ , Figura 37, barra amarela) reduziu o sinal em 48%. O derivado indólico **novo 2** (25  $\mu\text{M}$ , Figura 37, barra roxa) reduziu o sinal em 60%. A redução de sinal observada é bastante similar à apresentada pelo ensaio na presença de colchicina.

Esses resultados mostram que os derivados indólicos **8** e **novo 2**, que apresentaram capacidade de inibir a polimerização de tubulina, são capazes de se ligar ao sítio da colchicina.

#### 1.4.11 Docagem Molecular dos Derivados Indólicos no Sítio da Colchicina

Tendo em vista a comprovação experimental do sítio da colchicina como alvo de interação dos derivados indólicos que apresentaram atividade em inibir a polimerização de microtúbulos, estudos mais aprofundados de modelagem molecular foram desenvolvidos a fim de permitir a elaboração de um modelo computacional para explicar o mecanismo molecular envolvido na atividade de modulação destes inibidores.

Primeiramente, estudos mais detalhados de docagem molecular foram executados com o propósito de se obter um complexo plausível dos compostos ligados ao sítio da colchicina para ser utilizado como conformação inicial em estudos posteriores de modelagem. Além disso, uma vez que os derivados indólicos avaliados e identificados pertenciam a uma mistura racêmica, esta análise mais detalhada permitiu identificar a forma enantiomérica com maior complementariedade estrutural e eletrônica ao sítio da colchicina. Para isso, uma nova estrutura do dímero de tubulina em complexo com a colchicina foi selecionada (código de acesso PDB: 4O2B) e preparada de acordo com os procedimentos já descritos na sessão de metodologia computacional.<sup>36,153</sup> A estrutura do receptor foi preparada para a docagem pela retirada do ligante em complexo (colchicina), adição de hidrogênios e pela verificação das conformações dos resíduos dos sítios, utilizando o programa Pymol 1.7.<sup>202</sup> As moléculas de água w606, w623 e w728 foram consideradas como águas cristalográficas e utilizadas como parte do sítio de ligação, por apresentarem um baixo valor do fator de temperatura reportado e por estarem interagindo diretamente com a colchicina no complexo cristalográfico original. Ambas as formas enantioméricas (*RR* e *SS*) dos derivados indólicos **8** e **novo 2** foram submetidas ao processo de docagem molecular utilizando o programa AutoDock Vina com o nível máximo de busca conformacional.

De acordo com os resultados obtidos pelo estudo de docagem molecular, os derivados indólicos na forma enantiomérica *SS* apresentaram uma complementariedade superior à apresentada pela forma *RR*, como apresentado na Tabela 7. Enquanto a diferença de afinidade predita para os enantiômeros do derivado indólico **8** é bastante pequena e inconclusiva, os valores de afinidade preditos para o derivado indólico **novo 2** indicaram que a forma *SS* deve apresentar uma afinidade aproximadamente 10 vezes superior à da forma *RR*. Dessa forma, de acordo com os resultados de docagem molecular observados, a forma enantiomérica *SS* dos derivados indólicos foi selecionada como a forma ativa mais provável de ser responsável pela inibição da polimerização de microtúbulos e utilizada nos estudos posteriores de modelagem molecular.

As conformações de menor energia (maior afinidade de ligação) preditas pelo processo de docagem molecular são apresentados na Figura 38. Pode-se observar que ambos os ligantes apresentaram um padrão de interação similar, com a região do grupamento fenólico dos derivados indólicos inserido em um bolsão hidrofóbico formado principalmente pelos resíduos Ala248, Leu253 e Ile316 da  $\beta$ -tubulina (Figura 38 A, B e C). Além disso, ambos os derivados tiveram uma conformação predita com o grupamento metoxila na posição *para* do anel benzênico substituinte interagindo com o resíduo Asn256 da  $\beta$ -tubulina, por meio de uma ligação de hidrogênio (Figura 38C).

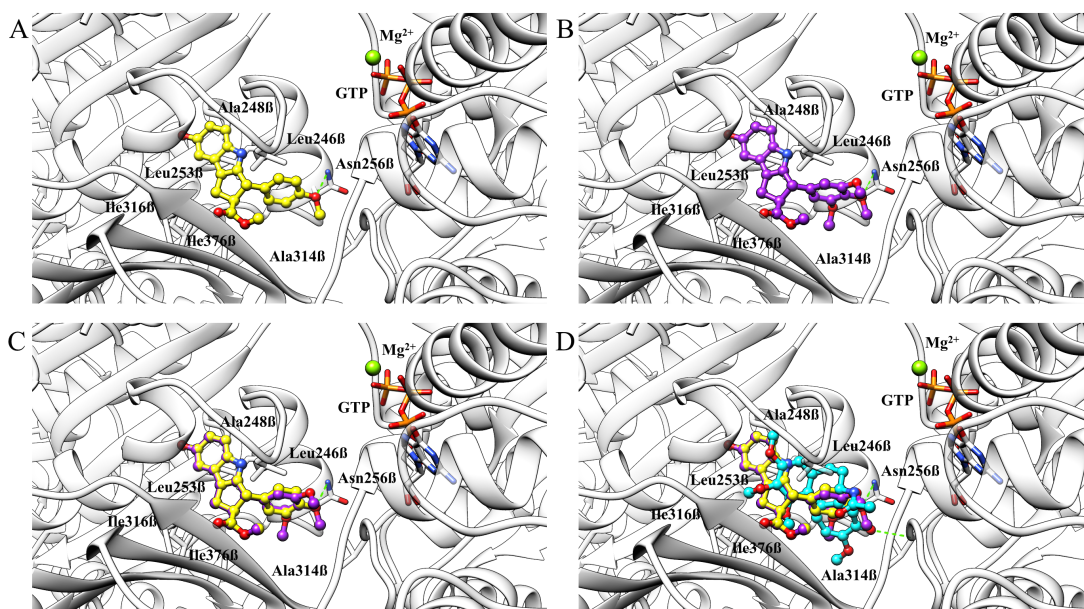
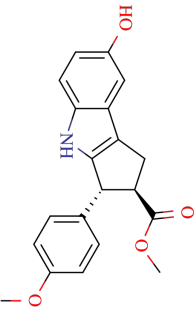
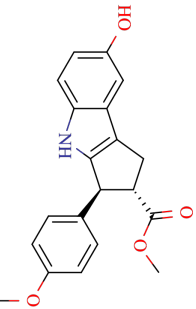
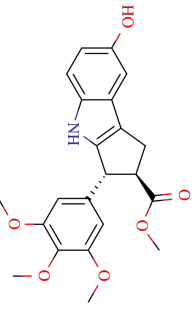
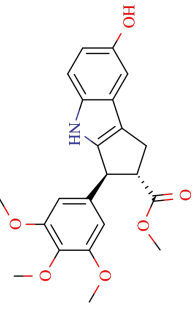


Figura 38 – Complexos preditos pelos estudos de docagem molecular utilizando o programa Auto-dock para os derivados indólicos **8** e **novo 2** no sítio do dímero de  $\alpha\beta$ -tubulina (código de acesso PDB: 4O2B). (A) conformação do derivado indólico **8**; (B) conformação do derivado indólico **novo 2**; (C) sobreposição das conformações obtidas para ambos os derivados indólicos; (D) Comparação da conformações dos derivados indólicos com a conformação da colchicina cristalográfica no sítio de ligação.

Fonte: Elaborada pelo autor.

Tabela 7 – Resultados dos estudos de docagem molecular das formas enantioméricas dos derivados indólicos 8 e novo 2.

Composto	Enantiômero	Estrutura Molecular	Pontuação do AutoDock Vina $\log K_d$ (kcal/mol)
Derivado indólico 8	<i>RR</i>		-8,1
	<i>SS</i>		-8,2
Derivado indólico novo 2	<i>RR</i>		-6,8
	<i>SS</i>		-7,7

Fonte: Elaborada pelo autor

A conformação de ambas as estruturas dos derivados indólicos **8** e **novo 2** no sítio mostram uma sobreposição total do grupamento ciclopenta- $\beta$ -indol, indicando que o mesmo é essencial para o padrão de ligação desta classe (Figura 38C). Além disso, as orientações dos grupamentos tri-metoxila no anel benzênico substituinte do derivado indólico **novo 2**, e do grupamento metoxila do derivado indólico **8**, interagindo em um bolsão formado pelos resíduos Asn256 e Ala314 da  $\beta$ -tubulina, indicam que os mesmos sejam importantes para às atividades de inibição observadas. A comparação das conformações previstas para os derivados com a colchicina mostra um padrão diferente de interação para estes ligantes (Figura 38D), em que os grupos tri-metoxila ligados ao anel benzênico assumem orientações distintas no sítio de ligação. Enquanto ambos os derivados apresentam-se mais inseridos na cavidade formada pela  $\beta$ -tubulina na interface dimérica, interagindo somente com esta cadeia protéica, a colchicina apresenta-se mais próxima da  $\alpha$ -tubulina e da molécula de GTP no sítio. Além disso, ao invés de formar uma ligação de hidrogênio com o resíduo Asn256 da  $\beta$ -tubulina, a colchicina interage com o resíduo Val181 da  $\alpha$ -tubulina.

#### 1.4.12 Estudos de Dinâmica Molecular com os Derivados Indólicos

Uma propriedade intrínseca e bastante interessante da família das tubulinas é a sua elevada plasticidade. Esta capacidade permite que estas proteínas assumam diversos estados conformacionais que possibilitam sua associação e dissociação na formação de microtúbulos em um processo dinâmico e altamente complexo. Assim como a interação entre cadeias protéicas resultam em modificações consideráveis na conformação das tubulinas, a sua interação com pequenas moléculas também resulta em consideráveis modificações no sítio de ligação.

Considerando esta elevada plasticidade, espera-se que na estrutura cristalográfica considerada para os estudos de docagem molecular com os derivados indólicos (código PDB: 4O2B), os resíduos presentes no sítio de ligação da colchicina na interface do dímero de  $\alpha\beta$ -tubulina estejam adaptados para promover uma máxima complementariedade com o ligante cristalográfico colchicina.<sup>154</sup> Dessa forma, apesar de interagirem no mesmo sítio da colchicina, é bastante provável que a presença das moléculas de estudo identificadas (derivados indólicos) promovam modificações consideráveis nos resíduos presentes na interface  $\alpha\beta$ -tubulina, que não são capazes de serem consideradas por métodos usuais de docagem molecular. Devido à esta limitação, as conformações obtidas foram submetidas a estudos de dinâmica molecular a fim de se obter um modelo de interação mais preciso, capaz de elucidar os mecanismos moleculares responsáveis pela atividade de inibição observada e direcionar o planejamento de novos compostos com maior afinidade pelo sítio de ligação.

Os complexos preditos pelos estudos de docagem molecular para os derivados indólicos no sítio da colchicina (Figura 38) foram submetidos a estudos de dinâmica mole-

cular utilizando o programa GROMACS. A estrutura de dímeros de tubulina selecionada para docagem (código PDB: 4O2B) não apresenta um loop entre os resíduos 273 e 284 da  $\beta$ -tubulina. Para a realização dos estudos de dinâmica molecular estes resíduos ausentes foram modelados através do servidor SWISS-MODEL.<sup>155</sup> Como validação do método, o complexo cristalográfico contendo a colchicina ligado ao seu sítio no dímero de tubulina também foi submetido a estudos de simulação utilizando exatamente a mesma metodologia. Todos os complexos foram submetidos a simulações de 10 ns mantendo-se a temperatura e pressão constante a fim de reproduzir as condições dos ensaios experimentais em que os compostos foram avaliados frente ao alvo (37 °C e 1 atm, ensaios de fluorescência e espalhamento).

Os valores de desvio médio quadrático (RMSD, do inglês *Root-Mean-Square Deviation*) correspondentes à média de 3 simulações para o dímero de  $\alpha\beta$ -tubulina associado a cada um dos ligantes são apresentados nas Figura 39. Como pode-se observar pelas curvas de RMSD (em vermelho), em todos os sistemas o dímero de tubulina apresentou um padrão de variação estrutural muito similar, atingindo um equilíbrio conformacional após aproximadamente 4 ns.

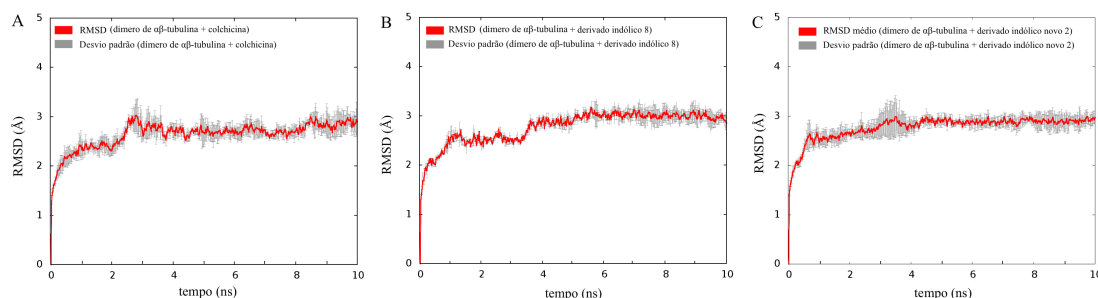


Figura 39 – Curvas de RMSD apresentando a variação conformacional do dímero de  $\alpha\beta$ -tubulina complexado com os ligantes de estudo durante o processo de simulação por dinâmica molecular. (A) Dímero associado ao ligante colchicina. (B) Dímero associado ao ligante derivado indólico **8**. (C) Dímero associado ao ligante derivado indólico **novo 2**.

Fonte: Elaborada pelo autor.

Os valores médios dos RMSDs resultantes de simulações em triplicata para cada uma das moléculas, em relação à conformação inicial, são apresentados na Figura 40.

Como esperado, uma vez que o sítio de ligação da estrutura cristalográfica do dímero de tubulina já se encontrava em uma conformação otimizada para a colchicina, esta molécula foi a que apresentou a menor variação conformacional ao longo de toda a simulação, com um RMSD médio de apenas 1,5 Å (Figura 40, curva em azul claro). Por outro lado, ambos os derivados indólicos apresentaram uma considerável modificação conformacional durante a simulação de dinâmica molecular, com RMSDs médios de 4,1 e 2,7 Å para os derivados indólicos **8** e **novo 2**, respectivamente. A estabilização dos RMSDs de ambos os derivados indólicos após 4 ns de simulação (Figura 40, curvas amarela e roxa) indicam que

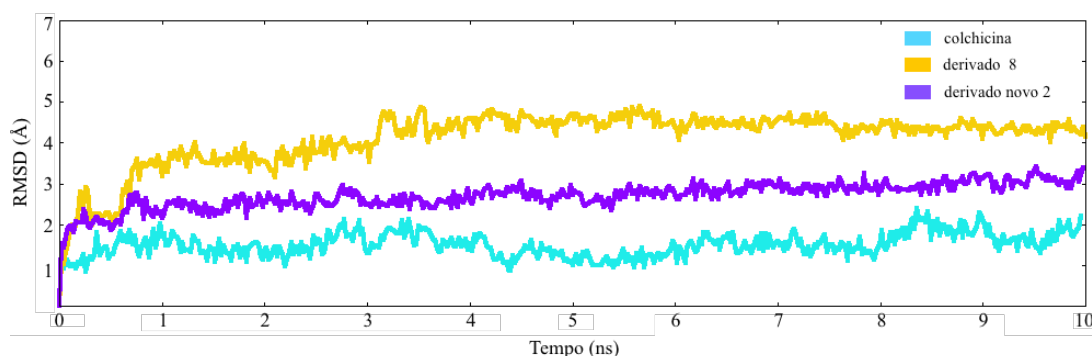


Figura 40 – Desvios médios quadráticos (RMSDs) dos derivados indólicos e da colchicina em complexo com o dímero  $\alpha\beta$ -tubulina.

Fonte: Elaborada pelo autor.

o sistema encontrou uma conformação estável em que ambos os ligantes são estabilizados por interações com o sítio da colchicina. Este ponto de equilíbrio está de acordo com a estabilização conformacional dos dímeros de tubulina em todas as simulações, indicando que a estabilização total dos sistemas se deve principalmente às alterações no sítio de ligação e à formação de novos padrões de interação entre a proteína e os ligantes.

Os perfis de RMSD para os ligantes, o co-fator GTP presente no sítio da colchicina e o íon de magnésio associado a ele, juntamente com os desvios correspondente à simulação em triplicata, são apresentados nas figuras do Anexo C. Os gráficos apresentando os RMSD do GTP e do íon magnésio (Anexo C) mostram que em todas as simulações tanto o co-fator quanto o íon associado permaneceram fortemente ligados às suas posições iniciais no sítio de ligação e apresentaram mínimas alterações conformacionais. A alta estabilidade do íon magnésio ligado ao GTP em todas as simulações é um indicativo de que as simulações reproduziram de forma plausível o comportamento do sistema real de estudo.

As conformações finais de cada complexo após 10 ns de simulação são apresentadas na Figura 41. As conformações iniciais dos ligantes (em cinza claro) são também apresentadas para comparação.

Como já observado na pela curva de RMSD na Figura 40, a conformação final da colchicina após 10 ns de simulação é bastante similar à conformação apresentada na estrutura cristalográfica (Figura 41A). A estrutura refinada difere apenas pelo rearranjo do grupamento acetoamida da colchicina, que permitiu a formação de uma ligação extra de hidrogênio da cadeia principal do resíduo Leu246 da  $\beta$ -tubulina, além da interação de hidrogênio já presente no complexo cristalográfico (e mantida durante a simulação) entre a carbonila no grupo heptaleno da colchicina e a amida da cadeia principal do resíduo Val181 pertencente à  $\alpha$ -tubulina (Figura 41A).

As simulações de dinâmica molecular para os derivados indólicos apresentaram variações conformacionais consideráveis na interface dimérica, principalmente no loop da

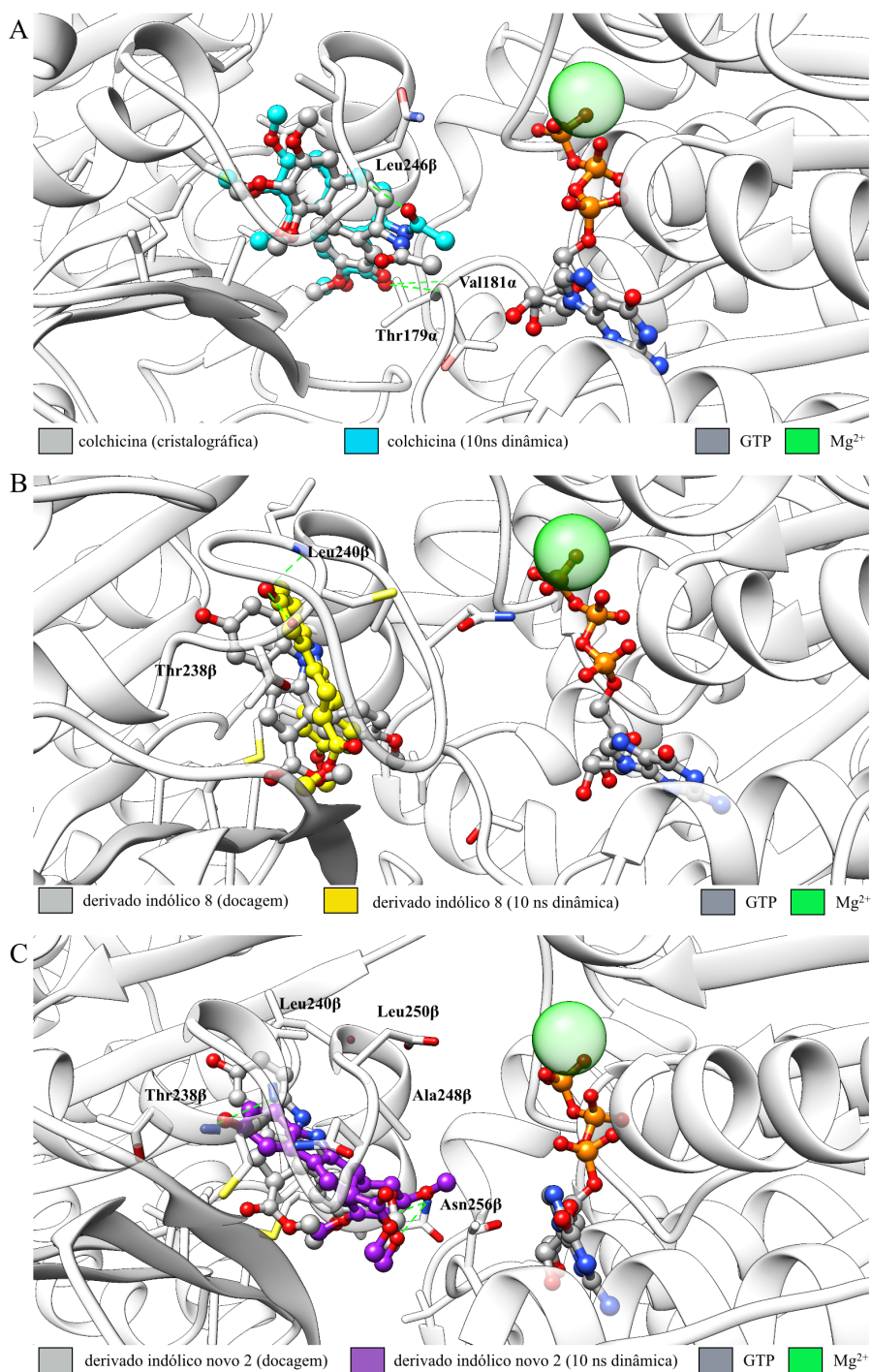


Figura 41 – Conformações finais obtidas após 10 ns de simulação para os complexos formados pelos ligantes (A) colchicina, (B) derivado indólico **8** e (C) derivado indólico **novo 2**. As conformações iniciais de cada ligante são apresentadas para comparação (em cinza claro).

Fonte: Elaborada pelo autor.

$\beta$ -tubulina formado pelos resíduos Leu240 a Ala248. As modificações conformacionais e o movimento deste loop são indicadas na Figura 42.

Esta região da  $\beta$ -tubulina, descrita na literatura como loop T7, é bastante flexível



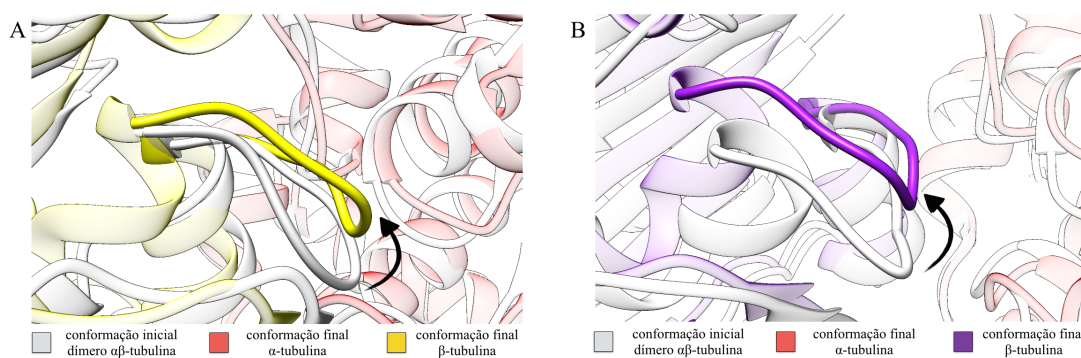


Figura 42 – Abertura do loop T7 da  $\beta$ -tubulina na interface dimérica durante simulações de dinâmica molecular. Este movimento foi observado na presença do (A) derivado indólico **8** e do (B) derivado indólico **novo 2**.

Fonte: Elaborada pelo autor.

e estudos têm demonstrado que seu movimento é induzido pela ligação da colchicina à interface dimérica.<sup>156</sup> Estes estudos sugerem que através da sua abertura em direção à  $\alpha$ -tubulina, induzida pela ligação da colchicina ou outro ligante no sítio da colchicina, o loop T7 da  $\beta$ -tubulina promove um impedimento estérico que evita a mudança do dímero de tubulina da forma torcida (presente no dímero em solução) para a conformação reta, adotada pelo dímero ao se associar ao microtúbulo.<sup>156</sup>

Interessantemente, durante a simulação de ambos os derivados indólicos observou-se uma leve abertura do loop T7 para uma posição mais próxima da cadeia da  $\alpha$ -tubulina (Figura 42). Este processo, além de concordar com o mecanismo proposto para a colchicina para inibir a formação de microtúbulos, permite que os derivados indólicos **8** e **novo 2** assumam uma conformação mais aprofundada na cavidade da  $\beta$ -tubulina na interface dimérica. Esta alteração conformacional permite que o grupo hidroxila presente no anel benzênico de ambos os ligantes formem novas ligações de hidrogênio com os grupamentos amida das cadeias principais dos resíduos Thr238 $\beta$  e Leu240 $\beta$  (Figura 41B-C). Além destas interações, de acordo com os resultados dos estudos de dinâmica molecular, o derivado indólico **novo 2** apresenta mais uma ligação de hidrogênio com o nitrogênio da cadeia lateral do resíduo Asn256 $\beta$  (Figura 41C).

Além das ligações de hidrogênio, a conformação mais estável predita por simulação apresenta interações hidrofóbicas entre o anel ciclopenta- $\beta$ -indol dos derivados indólicos e um bolsão hidrofóbico formado pelos resíduos Leu253, Met257, Cis239 e Ile316 da  $\beta$ -tubulina. As trajetórias das simulações para cada um dos ligantes podem ser observadas nos endereços eletrônicos disponibilizados no Anexo 3.

### 1.4.13 Planejamento, Síntese e Avaliação de Novos Derivados Indólicos

A partir da análise dos resultados experimentais e de modelagem molecular obtidos para a classe dos ciclopenta- $\beta$ -indóis, e da caracterização dos derivados indólicos **8** e **novo 2** como principais representantes como inibidores da polimerização de microtúbulos, novos compostos foram planejados e sintetizados para avaliação bioquímica. O planejamento destas moléculas teve como objetivos principais avaliar a influência de grupamentos descritos como fundamentais para a atividade nos estudos de modelagem e propor um composto com perfil farmacodinâmico superior aos já identificados.

A Tabela 8 descreve os novos compostos planejados, sintetizados e avaliados em ensaios de polimerização de microtúbulos nesta nova etapa.

De acordo com o modelo de ligação proposto para os derivados indólicos **8** e **novo 2** pelos estudos de docagem e dinâmica molecular, a hidroxila fenólica destas moléculas e parte integrante do grupo indol é fundamental para estabilizar os complexos através de sua interação com resíduos no sítio da colchicina, como a Thr238 e a Leu240 da  $\beta$ -tubulina (Figura 41). A avaliação do derivado indólico **novo 3** proposto (Tabela 8) busca corroborar a influência deste grupo e o modelo de ligação proposto, através da retirada da hidroxila do derivado indólico **novo 2** (Figura 31B).

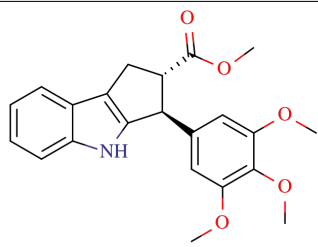
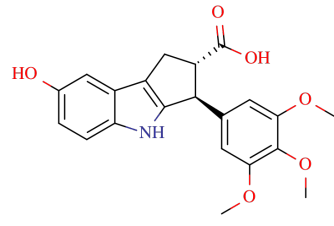
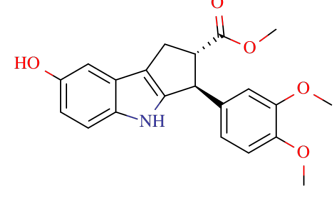
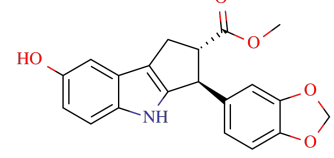
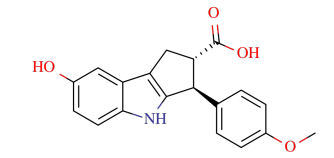
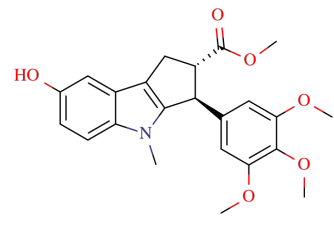
Outra característica a ser investigada nestes estudos foi a importância do grupamento éster substituinte no anel ciclopenteno fusionado ao grupo indol destes compostos. Dessa forma, o derivado indólico **novo 4** foi proposto pela substituição deste éster por ácido acético (Tabela 8).

As interessantes atividades dos derivados indólicos **8** e **novo 2** e a alta similaridade estrutural entre ambos, diferenciando apenas pela presença de uma mono-metoxila e uma tri-metoxila presentes no anel aromático ligado ao ciclopenteno, nos direcionou a derivação e síntese do derivado indólico **novo 5**, contendo uma di-metoxila (Tabela 8). Além disso, para investigar a conformação destes substituinte metoxila dos derivados indólicos e avaliar suas conformações propostas nos modelos computacionais, o derivado indólico **novo 6** foi proposto pela fusão dos grupos metoxi nas posições *para* e *meta* em um anel dioxalano (Tabela 8).

O derivado **novo 7**, por sua vez, foi proposto pela substituição do grupamento éter por um ácido acético a fim de averiguar a importância desse substituinte na atividade. Por fim, o derivado **novo 8** foi proposto adicionando um grupo metila ao nitrogênio indólico para confirmar a importância da amina secundária como um doador e receptor de ligações de hidrogênio.

Os resultados da avaliação destes novos compostos propostos em ensaios bioquímicos baseados em fluorescência na polimerização de microtúbulos são apresentados na Ta-

Tabela 8 – Novos derivados indólicos planejados a partir dos estudos de modelagem molecular.

Composto	Estrutura Molecular <sup>a</sup>	clogP
Derivado indólico novo 3		3,80
Derivado indólico novo 4		2,68
Derivado indólico novo 5		3,31
Derivado indólico novo 6		3,55
Derivado indólico novo 7		3,10
Derivado indólico novo 8		3,36

<sup>a</sup> Mistura racêmica.

Fonte: Elaborada pelo autor

bela 9.

Tabela 9 – Ensaios baseados em fluorescência da polimerização de microtúbulos com os novos derivados indólicos planejados.

Composto	% inibição a 50 $\mu$ M <sup>a</sup>
Derivado indólico novo 3	10
Derivado indólico novo 4	42
Derivado indólico novo 5	35
Derivado indólico novo 6	87
Derivado indólico novo 7	32
Derivado indólico novo 8	65

<sup>a</sup> Valor médio observado em ensaios em triplicata.

Fonte: Elaborada pelo autor

A comparação dos resultados apresentados na Tabela 9 com as estruturas dos derivados indólicos planejados permite confirmar diversas relações entre a estrutura e a atividade já observadas em estudos de modelagem molecular.

A baixa atividade do derivado indólico **novo 3** demonstra a importância da hidroxila substituinte na posição 5 do indol. Como os derivados indólicos **novo 2** e **novo 3** diferem estruturalmente apenas pela presença desta hidroxila no primeiro, pode-se afirmar que este grupo é essencial para atividade desta classe. Além disso, esta observação corrobora os modelos de ligação dos derivados no sítio da colchicina (Figura 41), em que a hidroxila realiza importantes interações com o sítio através de ligações de hidrogênio com os resíduos Thr238 e a Leu240 da  $\beta$ -tubulina.

A substituição do grupo éster pelo ácido acético como substituinte no anel ciclopenteno dos derivados **8** e **novo 2**, levando à síntese dos derivados indólicos **novo 4** e **novo 7**, resultou em uma perda significativa de atividade, com ambos os compostos planejados apresentando valores de IC<sub>50</sub> maiores que 50  $\mu$ M. Apesar da perda de atividade não apresentar uma correlação óbvia com os modelos computacionais de ligação, pode-se sugerir que a presença de um ácido carboxílico nesta posição resulte em interações menos favoráveis com o resíduo Asn256 $\beta$  do que as obtidas na presença do grupamento éster.

Um resultado bastante interessante é a perda substancial de atividade do derivado indólico **novo 5**, que apresenta duas metilas no anel benzênico substituinte no anel penteno (Tabela 8). Uma vez que os derivados indólicos **8** e **novo 2** apresentam elevada atividade e possuem uma e três metilas nesta região, respectivamente, era esperado que o derivado indólico **novo 5** apresentasse um perfil farmacodinâmico similar. Esta aparente discrepância poderá ser investigada em trabalhos posteriores.

Os resultados obtidos para o derivado indólico novo **8** sugerem que a presença de um grupo doador de ligação de hidrogênio no nitrogênio do grupamento indol não é essencial para a atividade desta série, uma vez que a presença de uma metila (levando a amina

secundária para a forma terciária) não reduziu a atividade desta molécula, em comparação com o derivado indólico **novo 2**. Este resultado também encontra-se de acordo com o modelo de ligação predito pelos estudos de dinâmica molecular, em que não se observa um papel fundamental desta amina secundária em realizar interações com resíduos do sítio da colchicina.

A elevada atividade apresentada pelo derivado indólico **novo 6** direcionou a avaliação deste composto em ensaios de concentração-resposta a fim de determinar o valor de  $IC_{50}$  e caracterizar o perfil farmacodinâmico deste composto. Os resultados desta avaliação são apresentados na Figura 43.

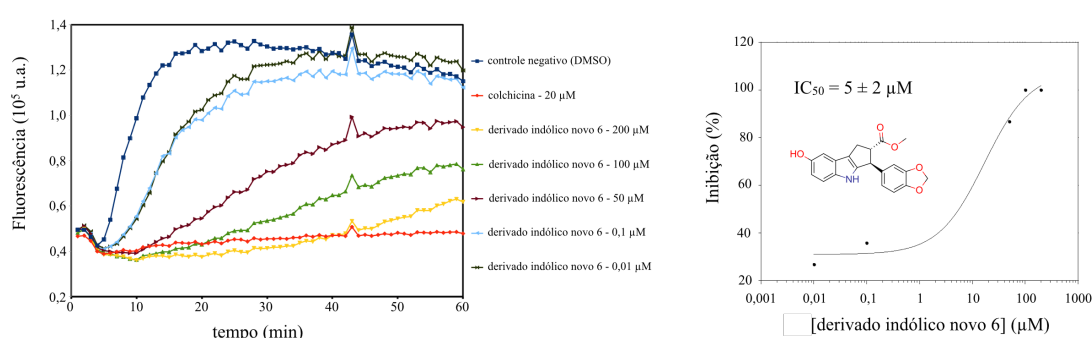


Figura 43 – Curva de dose resposta e determinação do valor de  $IC_{50}$  do derivado indólico **novo 6** em ensaios de polimerização de microtúbulos baseados em fluorescência.

Fonte: Elaborada pelo autor.

O valor de  $IC_{50}$  determinado para o derivado indólico **novo 6**, de  $5 \pm 2 \mu$ M, demonstra que a fusão dos dois grupamentos metila em um anel dioxalano resultou em uma melhora significativa de atividade e na obtenção de um composto ainda mais potente em inibir a polimerização de microtúbulos, em comparação com os compostos anteriormente caracterizados. A comparação destes resultados com o modelo de interação proposto pelos estudos de modelagem molecular indicam que a fusão dos grupamentos metila promove uma orientação ótima dos átomos de oxigênio para interação com resíduos do sítio, além de reduzir a entropia do complexo formado através do congelamento das ligações rotacionáveis dos substituintes metila.

Esta hipótese está de acordo com o modelo de ligação obtido pelo programa Auto-Dock Vina para o derivado indólico **novo 6**, apresentado na Figura 44.

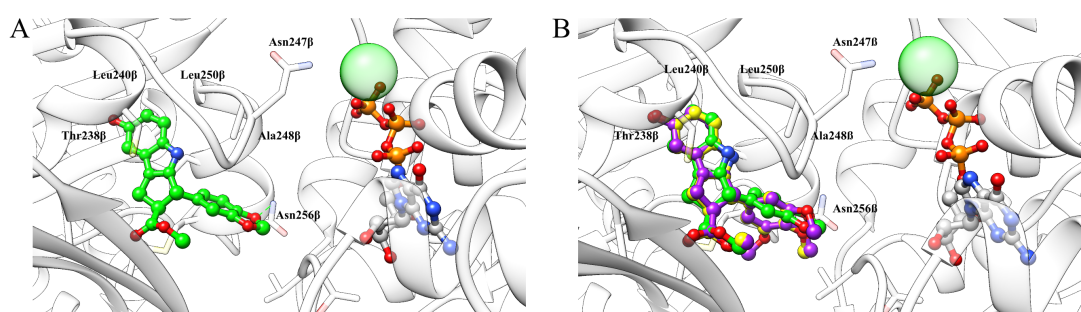


Figura 44 – (A) Modelo de ligação proposto para o derivado indólico **novo 6** a partir de estudos de docagem molecular. (B) Comparação do modelo de ligação do derivado indólico **novo 6** com os derivados indólicos **8** e **novo 2**.

Fonte: Elaborada pelo autor.

## 1.5 Conclusão

Considerando-se o panorama global de alta incidência de câncer e a relevância desta doença como uma das principais causas de morte no mundo moderno, o desenvolvimento de novos tratamentos eficazes é essencial. Apesar desse panorama e dos constantes esforços na busca de novos tratamentos, diversos fatores como resistência e diversidade fenotípica têm limitado a aplicação dos tratamentos disponíveis. O sucesso obtido por um número considerável de fármacos que modulam a ação de microtúbulos, como por exemplo, o paclitaxel e a vimblastina, corrobora a importância desse alvo molecular para o desenvolvimento de novos agentes quimioterápicos.

A identificação de novos moduladores da polimerização de tubulina com propriedades otimizadas e com atividade contra tumores resistentes aos fármacos disponíveis foi um dos principais objetivos deste trabalho. Nesta tese de doutorado é descrita a utilização de uma metodologia completa e bastante moderna na identificação e planejamento de novos fármacos. Os estudos realizados permitiram uma combinação de diversos métodos computacionais e experimentais que englobam várias áreas da ciência, em um programa altamente multidisciplinar. Os estudos obtidos permitiram a identificação e caracterização de uma nova série de derivados indólicos (ciclopenta- $\beta$ -indóis) com interessante perfil de inibição da polimerização de tubulina na formação de microtúbulos. Os compostos identificados apresentaram elevada atividade de inibição na faixa de micromolar tanto em ensaios de polimerização quanto de migração de células tumorais.

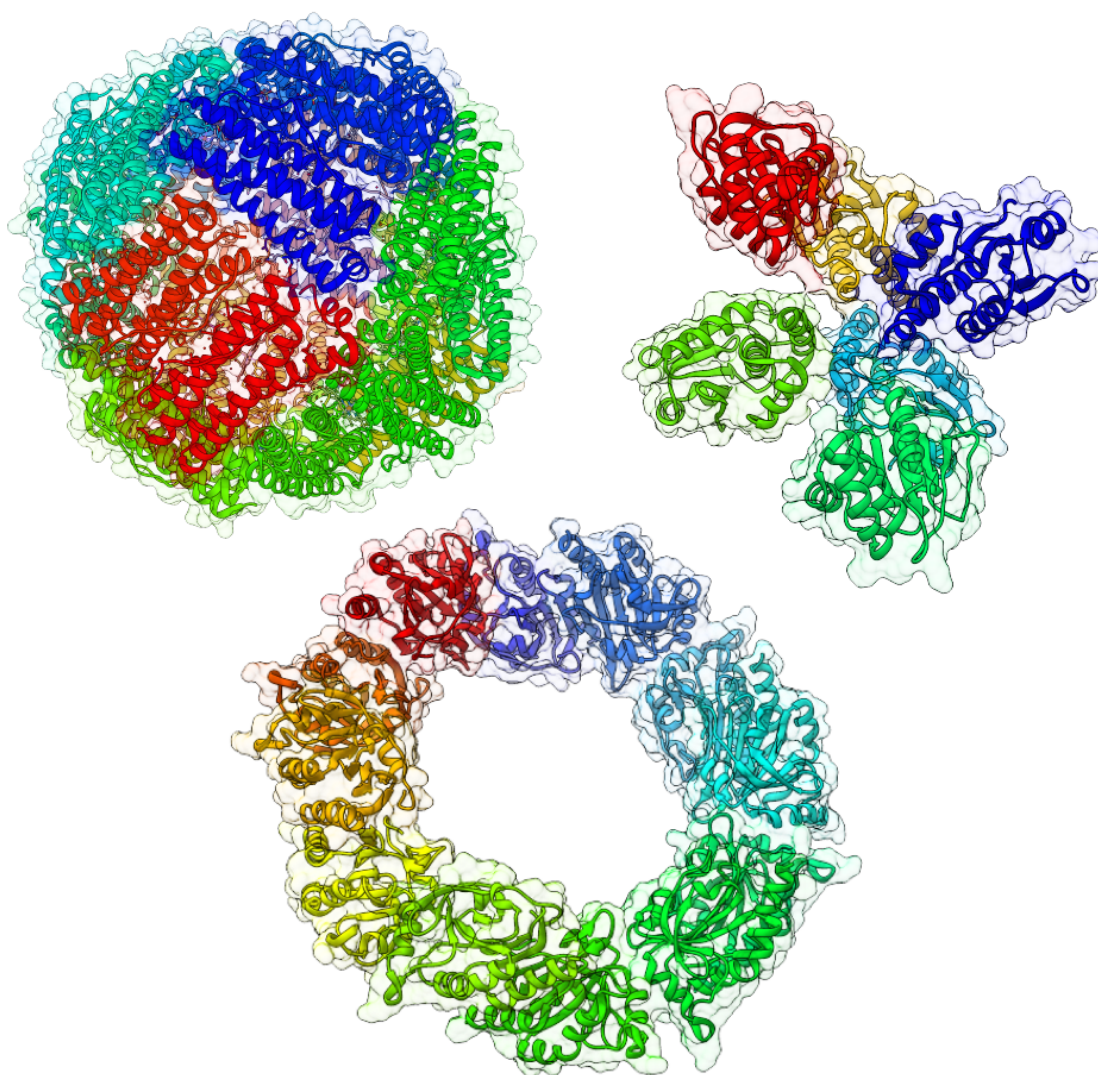
Estudos moleculares aprofundados permitiram determinar o sítio da colchicina na interface  $\alpha\beta$ -tubulina como a região de interação desta classe de compostos. A identificação do sítio de ligação dos derivados indólicos avaliados permitiu o desenvolvimento de estudos de modelagem molecular a fim de elucidar o padrão de interação desta classe e o mecanismo molecular envolvido no processo de inibição da formação de microtúbulos. Estes

estudos computacionais mostraram-se fundamentais para elucidar o mecanismo de ligação descrito e permitiram o planejamento de uma nova série de derivados indólicos. A avaliação desta série resultou na identificação de um inibidor de elevada atividade e perfil farmacodinâmico superior dentre todas as moléculas planejadas, com  $IC_{50}$  de  $5 \mu\text{M}$  para ensaios de polimerização frente ao alvo molecular. Concomitantemente, ensaios de citotoxicidade *in vitro* demonstraram uma interessante seletividade destes compostos por células cancerígenas em comparação a células saudáveis. Estudos de imunofluorescência revelaram a elevada capacidade destes compostos em desorganizar o citoesqueleto celular. Os compostos identificados durante este trabalho apresentam um excelente perfil farmacológico e representam um passo importante para o desenvolvimento de novos candidatos a agentes quimioterápicos contra o câncer.





## 2 Modelagem Baseada em Estrutura e Análise de Acoplamento Direto na Predição de Sistemas Diméricos Macromoleculares





## 2.1 Introdução

### 2.1.1 Oligomerização em Sistemas Biológicos

O elevado sinergismo entre componentes básicos celulares é uma característica essencial para a vida. A interação entre diversas moléculas em um processo complexo e preciso permite às células de um organismo desenvolver ações fundamentais para que o metabolismo ocorra.<sup>157–162</sup> No caso de proteínas, componentes mais abundantes em sistemas biológicos, as interações intermoleculares necessárias para o funcionamento de uma via metabólica foram submetidas à pressão evolutiva ao longo do tempo, fazendo com que suas estruturas tridimensionais fossem moldadas a fim de garantir a manutenção da atividade fisiológica.

Dentre os processos biológicos celulares, uma considerável fração é desenvolvida por complexos proteicos resultantes de interações estáveis entre subunidades de mesma ou diferentes composições (Figura 45). Exemplos importantes incluem enzimas, como a oxidases, caspases, proteína G (Figura 45); e proteínas estruturais como aquaporinas ou proteínas globulares como actina e tubulina, que formam filamentos de actina e microtúbulos, respectivamente (Figura 45).<sup>163–167</sup>

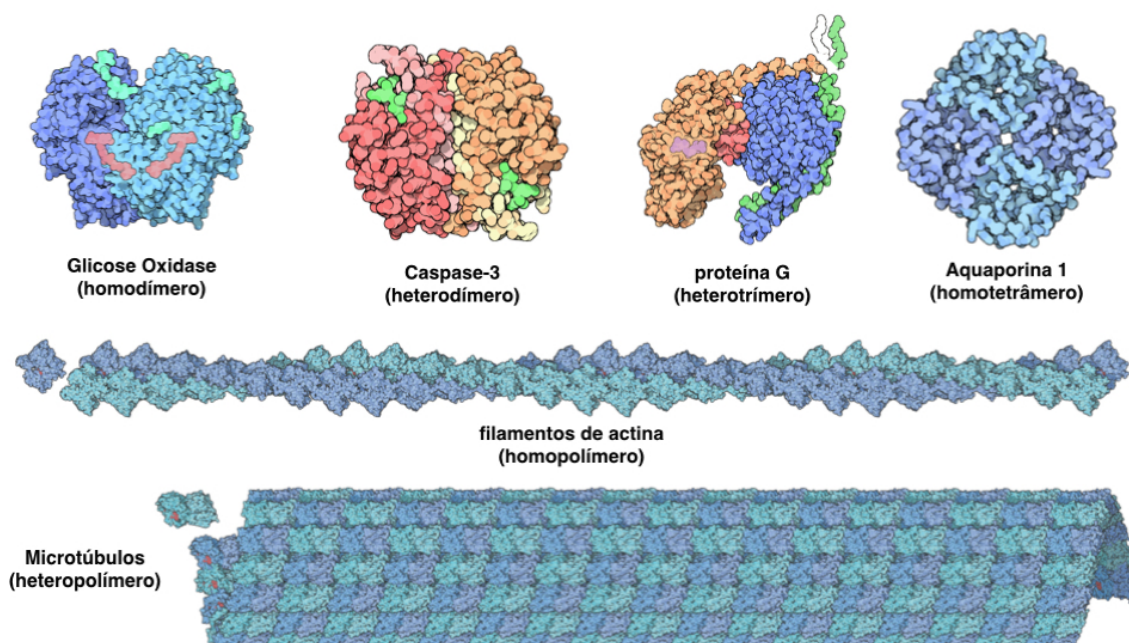


Figura 45 – Exemplos de complexos proteicos formados pela associação estável de diversas proteínas.

Fonte: disponível em [http://www.rcsb.org/pdb/101/motm\\_archive.do](http://www.rcsb.org/pdb/101/motm_archive.do)

Outros exemplos de associação entre proteínas são macrocomplexos capazes de formar anéis que auxiliam no processo de divisão celular (como a proteína FtsZ), proteínas

envolvidas em processos de degradação (protease FtsH AAA) e fatores de transcrição que se ligam ao DNA permitindo modular a atividade de determinados genes.<sup>168–172</sup> Dentre as estruturas oligoméricas formadas pela associação de duas ou mais cadeias proteicas, sistemas formados por proteínas de mesma constituição (homo-oligomeros) constituem um grupo predominante.<sup>173</sup> Ainda dentro deste subgrupo, complexos homodiméricos são os mais abundantes, apresentando em média o dobro de interações com outras proteínas em relação a proteínas incapazes de oligomerizar.<sup>173–176</sup>

### 2.1.2 Predição de Interações entre Resíduos por Análise de Acoplamento Direto

Apesar das interações diméricas serem comuns e fundamentais para o funcionamento celular, a simples análise da sequência de uma proteína não permite identificar nenhum indício se a proteína é ou não capaz de dimerizar. Mesmo quando ensaios experimentais confirmam a capacidade de um sistema em dimerizar, encontrar a interface de dimerização e os mecanismos moleculares envolvidos neste processo é ainda um grande desafio. Interações diméricas devem satisfazer requisitos estruturais tanto monoméricos quanto diméricos e essas restrições são refletidas em acoplamentos entre aminoácidos ao longo da evolução das sequências proteicas de uma determinada família de proteínas. Esse acoplamento surge pelo fato de que, para que uma mutação em um resíduo importante para atividade ou estabilidade estrutural de uma proteína seja viável, ela deve ser compensada por uma segunda mutação em um resíduo em contato, a fim de manter a conformação e propriedades químicas da região da macromolécula (Figura 46 A).<sup>177</sup>

Estes tipos de mutações acumuladas ao longo do processo evolutivo de uma classe de proteínas podem ser identificados pela análise de uma biblioteca de sequências proteicas e a correlação entre essas mutações pode ser quantificada por um método de regressão linear (Figura 46 B). A informação de acoplamento evolutivo tem sido utilizada para estudar a co-evolução de resíduos para prever a estrutura de proteínas monoméricas e sua plasticidade conformacional (Figura 46 C).<sup>177–184</sup> Um exemplo destes métodos é a Análise de Acoplamento Direto (DCA, do inglês *Direct-Coupling Analysis*), desenvolvida no Centro Teórico de Física Biológica CTBP (do inglês, Center for Theoretical Biological Physics), na Rice University, e utilizada neste trabalho (Figura 46).<sup>183</sup>

A co-evolução de resíduos também tem sido estudada no contexto de interações entre proteínas. Uma das primeiras aplicações foi a predição de complexos entre kinases de histidina e reguladores de resposta.<sup>185,186</sup> Outros trabalhos incluem estudos de especificidade em sistemas de dois componentes, predição de modos de interação para novos alvos cancerígenos, estudos da formação do macro-complexos e da função de proteases AAA.<sup>187–191</sup>

Grande parte das aplicações dos métodos de análise de co-evolução para estudar

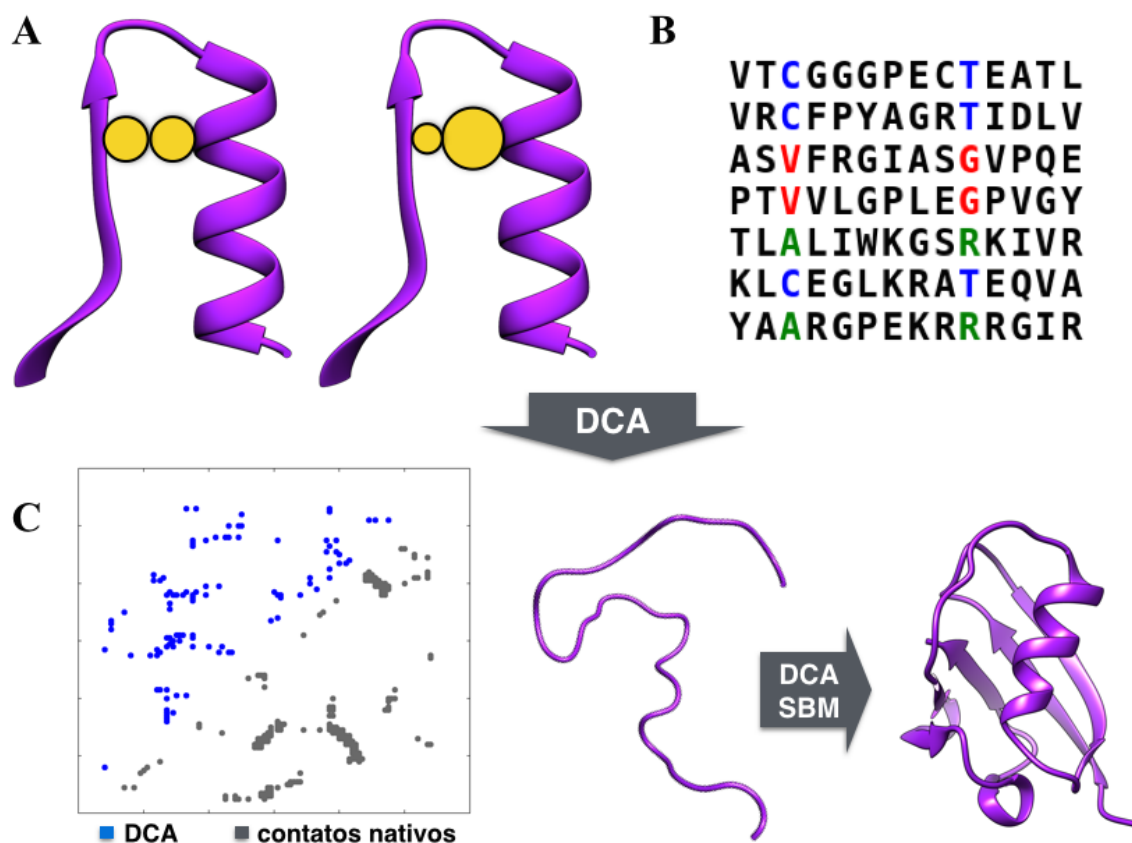


Figura 46 – (A) Processo de acoplamento coevolutivo. No exemplo apresentado, um resíduo pertencente a um par estereoquimicamente complementar (esferas amarelas) é mutado para um resíduo de menor volume. Para que a proteína mantenha suas características estruturais e funcionais, o resíduo complementar deve ser mutado (no caso, para um resíduo mais volumoso) a fim de manter a complementariedade do par interagente. (B) Resíduos que encontram-se fisicamente em contato na estrutura tridimensional de uma proteína devem apresentar uma correlação de mutação. Isso pode ser observado analisando um conjunto de sequências de proteínas da mesma família. (C) As correlações entre resíduos podem ser quantificadas por um método de correlação (como por exemplo o DCA) e estes contatos preditos podem ser utilizados para guiar o enovelamento ou mudanças conformacionais de proteínas.

Fonte: Elaborada pelo autor.

interações heterodiméricas envolve a construção de um banco de dados de pares de heterodímeros conhecidos utilizando dados já conhecidos *a priori*. Por exemplo, no pressuposto de que duas sequências de proteínas que estão codificadas muito próximas no genoma interajam ou que a presença de cópias únicas no organismo indicaria esta correspondência. Homodímeros representam um caso especial em que o sinal de interação é obtido a partir sequências pertencentes a uma única família. O desafio é que pares de resíduos co-evoluídos com elevado grau de acoplamento geralmente refletem contatos que ocorrem, em sua maioria, na estrutura monomérica tridimensional de uma única proteína. Dessa forma, contatos diméricos ou oligoméricos encontram-se misturados a esses contatos monoméricos, tornando-os difícil de identificar.

### 2.1.3 Predição de Complexos Proteicos Utilizando Sinais de Coevolução em Simulações de Dinâmica Molecular com Modelos Baseados em Estrutura

A predição de complexos homodiméricos é uma área bastante ativa atualmente, com estudos sendo desenvolvidos por diferentes abordagens, como por exemplo algoritmos de docagem e simulações de dinâmica molecular.<sup>192–195</sup> Dentre os métodos de simulação mais atuais, uma ferramenta bastante moderna e de grande utilidade na exploração de variações conformacionais de proteínas é o uso de Modelos Baseados em Estrutura (SBM, do inglês *Structure-Based Models*). Os SBMs foram desenvolvidos para estudos de enovelamento e de mudanças conformacionais de proteínas através de simulações de dinâmica molecular.<sup>190, 196, 197</sup> Este método difere dos métodos de dinâmica molecular tradicionais pois os potenciais (também conhecidos como Hamiltonianos) que determinam as forças interatômicas são parametrizados tendo como referência a conformação molecular de uma proteína específica (geralmente determinada por métodos experimentais, como cristalografia de raios-X ou ressonância magnética nuclear). Por outro lado, métodos de dinâmica tradicionais utilizam campos de força generalizados (*ab initio*) adequados para representar grande parte dos sistemas biológicos. SBM são construídos de acordo com o princípio de frustração mínima, no qual as mudanças conformacionais de uma macromolécula estão submetidas a um perfil de energia que possui seu ponto de mínimo correspondente à sua estrutura nativa.

Uma vantagem do método de SBM é que ele permite o ajuste do Hamiltoniano para considerar interações que não estão presentes na estrutura de referência, mas que podem ser observados através de estudos experimentais. Além disso, devido à sua simplicidade, outras modificações do campo de força e a exploração conformacional devido a diferentes interações no sistema são de fácil implementação. Portanto, essas características fazem do SBM um método bastante poderoso para estudar o enovelamento e mecanismos funcionais de proteínas.

Neste trabalho, uma metodologia efetiva e prática foi desenvolvida para extrair contatos diméricos co-evoluídos em meio ao sinal de contatos monoméricos obtidos pela aplicação do método de DCA.<sup>183</sup> Essa metodologia foi aplicada juntamente com SBMs na predição de 11 complexos homodiméricos com diferentes características estruturais e níveis de complexidade.<sup>198</sup> Estes sistemas incluem dímeros simétricos e assimétricos, complexos com múltiplos domínios e proteínas de médio e grande porte (até 444 resíduos).

## 2.2 Motivações e objetivos

A aplicação de métodos experimentais como cristalografia de raios-X, ressonância magnética nuclear (RMN) e microscopia eletrônica em sistemas biológicos permitiram elucidar estruturas tridimensionais de uma grande variedade de proteínas ao longo dos últimos anos. Por exemplo, somente no ano de 2014, 9658 estruturas novas foram depositadas no Protein Data Bank (a principal base de dados de estruturas de proteínas). Apesar de todo este avanço na área de biologia molecular estrutural e do considerável número de estruturas moleculares e mecanismos já solucionados, este conhecimento representa somente uma pequena parcela de uma infinidade de estruturas, funções e interações moleculares ainda a serem descobertos. Grande parte destes sistemas ainda não foram caracterizados devido às limitações de aplicabilidade dos métodos atualmente disponíveis.

Neste trabalho, estudos computacionais foram executados buscando o desenvolvimento de um novo método capaz de identificar as conformações funcionais de sistemas proteicos multiméricos, através da análise de sinais de coevolução e simulações moleculares.

Os objetivos deste trabalho são:

- Avaliar a capacidade de métodos de análise de coevolução utilizando correlação estatística em identificar sinais envolvidos na oligomerização de sistemas proteicos.
- Desenvolver um método eficiente para filtrar sinais oligoméricos dentre os diversos sinais de coevolução presentes na sequência primária de proteínas.
- Desenvolver um método para implementar os sinais de coevolução como potenciais para em um campo de força para simulações de dinâmica molecular.
- Identificar uma estratégia eficiente e generalizada para conduzir a formação de complexos multiméricos.

## 2.3 Metodologia

### 2.3.1 Análise de Correlação Evolutiva de Sequências Proteicas

Para prever os contatos entre monômeros propagados por co-evolução, conjuntos de dados contendo múltiplos alinhamentos de sequências de todas as famílias presentes em cada proteína selecionada foram extraídos da base de dados Pfam.<sup>199</sup> Todos os conjuntos de dados obtidos do Pfam possuíam mais que 3.500 sequências (Tabela 10), satisfazendo um pré-requisito para garantir significância estatística e um considerável nível de precisão na predição do DCA. Uma lista de todas as famílias utilizadas para as predições e as proteínas correspondentes é apresentada na Tabela 10. Para estimar os contatos físicos entre resíduos acoplados por co-evolução, o método *mean field* foi implementado, de acordo com os procedimentos descritos na literatura.<sup>183</sup>

Neste método, regiões individuais de interação são representados por frequências e pares em alinhamentos múltiplos de sequência são definidos como probabilidades de combinação.<sup>177,183</sup> Os pares de interação são ordenados seguindo um índice de informação direta (DI, do inglês *Direct Information*) que corresponde à probabilidade destes resíduos estarem em contato em um espaço tridimensional na proteína.<sup>180</sup>

As posições dos alinhamentos das famílias Pfam ao longo de cada proteína foram determinadas usando o módulo *hmmscan* do programa HMMER, que utiliza modelos ocultos de Markov (do inglês, *Hidden Markov Models*) para executar o alinhamento.<sup>200</sup> Os contatos classificados quanto à probabilidade de ocorrência foram mapeados aos seus resíduos correspondentes nas estruturas cristalográficas utilizando os resultados do alinhamento pelo *hmmscan* e um *script* de mapeamento desenvolvido pelo aluno.

Pelo fato dos contatos preditos entre proteínas nestes estudos serem homo-diméricos, o sinal contendo as interações diméricas estão misturados às interações monoméricas intrínsecas de cada cadeia. Com a finalidade de filtrar os contatos monoméricos preditos pelo DCA e obter somente o sinal de co-evolução correspondente às interações de dimerização, um processo de filtragem compreendido em duas etapas foi desenvolvido.

Na primeira etapa, a área de superfície acessível ao solvente (SASA, do inglês *Solvent Accessible Surface Area*) de cada resíduo da proteína foi calculada através do programa GetArea.<sup>201</sup> De forma geral, um resíduo precisa estar significativamente exposto na superfície da interface entre monômeros para participar em interações que permitem a formação do complexo. Dessa forma, todos os pares de interação preditos pelo DCA que não apresentavam ao menos um dos resíduos com sua superfície mais de 50% exposta ( $SASA > 0,5$ ) foram removidos da lista de contatos. Este valor de 50% corresponde ao mínimo de superfície acessível para classificar o resíduo como exposto, de acordo com o servidor do GetArea.



Tabela 10 – Sistemas selecionados para predição de homodímeros utilizando os métodos de SBM e DCA.

<b>Proteína</b>	<b>Código PDB</b>	<b>Número de resíduos</b>	<b>Família (Pfam)</b>	<b>Sequências</b>
Proteína da Triade de Histidina	3O0M	149	HIT	8213
GAPDH	4LSM	346	Gp_dh_N	14213
Isocitrato Desidrogenase	2IV0	412	Iso_dh	12354
Álcool Desidrogenase	1HDY	374	ADH_zinc_N ADH_N	42002 42970
Aminomutase	2ZSM	434	Aminotran_3	23135
tRNA metiltransferase	1UAL	274	tRNA_m1G_MT	5148
IspG	4G9P	406	GcpE	3640
Cetoacyl Sintase	2IX4	431	ketoacyl-synt	24208 23531
Glicose 6-Fosfato Isomerase	3FF1	446	PGI	7325
RegX3	2OQR	230	Response_reg Trans_reg_C	47512 47512
PhoB	1ZES	125	Response_reg Trans_reg_C	47512 47512

Fonte: Elaborada pelo autor

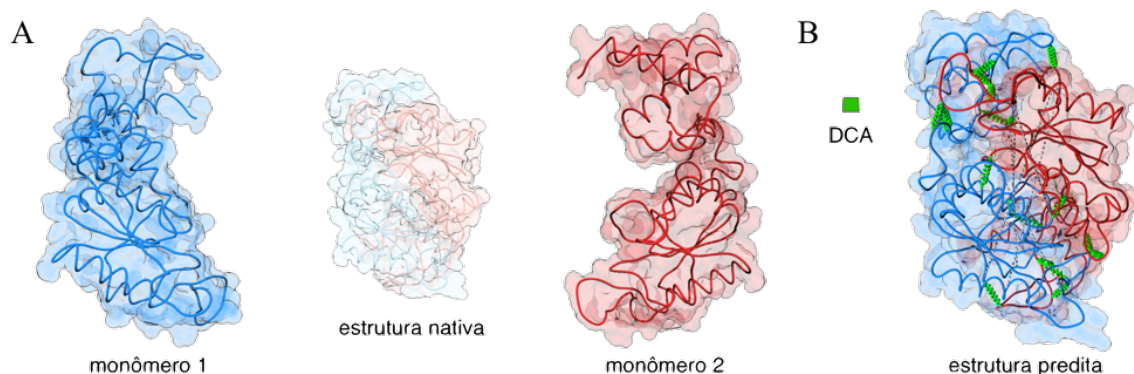


Figura 47 – Processo de predição do complexo proteico tendo como exemplo a proteína tRNA metiltransferase (código PDB: 1UAL). (A) Os monômeros são separados a uma distância de 50 Å. Um dos monômeros é rotacionado para representar uma situação de orientação aleatória entre as macromoléculas. (B) As distâncias de equilíbrio das interações diméricas provenientes do método de DCA são reduzidas até que seja observada a formação de um complexo estável.

Fonte: Elaborada pelo autor.

Em uma segunda etapa, os 100 primeiros pares com maiores valores de DI preditos pela aplicação do DCA e resultantes da filtragem por SASA foram selecionados. Esses pares foram comparados aos contatos presentes em um único monômero da estrutura nativa com uma distância menor ou igual a 8 Å entre átomos  $C\alpha$ . Em seguida, todos os pares em comum aos contatos nativos monoméricos foram removidos e os contatos resultantes desta segunda filtragem foram adicionados como potenciais de interação nas topologias utilizadas para conduzir simulações de dinâmica molecular.

### 2.3.2 Modelagem Estrutural de Sistemas Proteicos

Todas as estruturas moleculares dos homodímeros utilizadas nestes estudos foram obtidas da base de dados PDB (Protein Data Bank).<sup>100</sup> O código de acesso de cada proteína é apresentado na Tabela 10. A resolução das estruturas cristalográficas das proteínas selecionadas apresenta um valor médio de 1,97 Å e uma distribuição variando entre valores de 1,55 e 2,50 Å. Estruturas de homodímeros que apresentavam apenas um monômero na unidade assimétrica foram duplicadas de acordo com a simetria descrita nas suas estruturas por operações de rotação e translação utilizando o programa Pymol 1.7, gerando estruturas dimericas completas.<sup>202</sup> Regiões de loop ausentes em algumas estruturas cristalográficas foram modeladas através do servidor SWISS-MODEL, utilizando a própria estrutura como modelo de referência.<sup>155,203</sup> Todos os átomos que não faziam parte da cadeia proteica foram removidos antes de submeter o sistema a estudos de simulação com SBM.

A Figura 47 apresenta a metodologia empregada tendo como exemplo o dímero da proteína tRNA metiltransferase. Com o objetivo de reproduzir os complexos originais

dos homodímeros selecionados para este estudo, cada estrutura nativa foi separada em dois monômeros por uma distância de 50 Å (Figura 47 A). Em seguida, um dos monômeros foi rotacionado 180 graus em um eixo perpendicular em relação ao eixo que liga ambas as moléculas, com a finalidade de reproduzir uma condição de aleatoriedade na orientação entre os monômeros (Figura 47 A). Diferentes ângulos e orientações entre os monômeros foram testadas gerando resultados muito similares aos obtidos com o procedimento descrito.

A estrutura final correspondendo aos dois monômeros separados foi submetida ao servidor SMOG gerando modelos contendo átomos C $\alpha$  e topologias compatíveis para o método de simulação por SBM. Os contatos entre resíduos, obtidos pela análise de co-evolução, foram utilizados como interações para unir as duas cadeias proteicas em um processo gradual, permitindo aos monômeros se reorientarem ao longo da formação do complexo homodimérico (Figura 47 B). Este processo é executado pela redução da distância de equilíbrio dos contatos de co-evolução (representados em verde, Figura 47 B) até a distância média de interação entre resíduos C $\alpha$ , que é de 8 Å. As simulações de dinâmica molecular foram desenvolvidas utilizando uma versão modificada do GROMACS 4.5.7, com suporte à utilização de potenciais de gaussianas e disponibilizada no site do servidor SMOG.<sup>112, 198</sup>

### 2.3.3 Modelagem Baseada em Estrutura Utilizando Potenciais Gaussianos

SBM são construídos de acordo com o princípio de frustação mínima, no qual as mudanças conformacionais de uma macromolécula estão submetidas a um perfil de energia que possui seu ponto de mínimo correspondente à estrutura nativa.

O Hamiltoniano de um SBM pode ser representado como a seguir:

$$H(r_{ij}) = V_B(r_{ij}) + V_{NB}(r_{ij}) \quad (2.1)$$

onde  $V_B(r_{ij})$  e  $V_{NB}(r_{ij})$  são os potenciais entre as interações covalentes e não covalentes do sistema, respectivamente. O potencial que descreve interações covalentes pode ser descrito mais detalhadamente por potenciais que contabilizam cada grau de liberdade intramolecular:

$$V_B(r_{ij}) = \sum_{\text{ligação}} k_b (r_{ij} - r_{ij}^N)^2 + \sum_{\text{ângulos}} k_a (\theta_{ijk} - \theta_{ijk}^N)^2 + \sum_{\text{diedros}} k_d [1 - \cos(n(\phi_{ijkl} - \phi_{ijkl}^N))] \quad (2.2)$$

A equação 2.2 é analoga à apresentada na Figura 15.  $\mathcal{N}$  corresponde ao estado nativo de referência utilizado para gerar o Hamiltoniano. O primeiro termo em (2.2) corresponde às variações de energia devido às mudanças no comprimento da ligação entre 2 átomos.

Analogamente, o segundo e terceiro termos consideram variações de energia devido a alterações nos ângulos e diedros ( $\theta_{ijk}^N$  e  $\phi_{ijkl}^N$ , respectivamente) em relação às posições de equilíbrio da estrutura de referência. As constantes  $k_a$ ,  $k_b$  e  $k_d$  penalizam desvios dos ângulos e ligações em relação à estrutura nativa, direcionando as mudanças conformacionais do sistema para a estrutura de mínima energia. O potencial que descreve interações não-covalentes em um potencial de SBM pode ser expresso como a seguir:

$$V_{NB}(r_{ij}) = \sum_i \sum_{j>i+3} (1 - C_{ij})V_{ij}^R(r_{ij}) + C_{ij}V_{ij}^C(r_{ij}) \quad (2.3)$$

O tipo de interação pode ser puramente repulsivo para pares que não estão em contato na estrutura nativa considerando  $C_{ij} = 0$ , ou uma função com um mínimo de energia na distância de equilíbrio para resíduos que estão em contato na estrutura nativa considerando  $C_{ij} = 1$ . O termo de repulsão é parametrizado baseando-se no volume de exclusão entre átomos, e pode ser descrito por:

$$V_{ij}^R(r_{ij}) = \epsilon \left(\frac{d}{r_{ij}}\right)^{12} \quad (2.4)$$

onde  $\epsilon$  é uma unidade reduzida de energia ( $\epsilon = kbT$ ) e  $d$  é o raio atômico. Historicamente, o Hamiltoniano utilizado no método de SBM representa as interações atrativas através de um potencial do tipo Lennard-Jonnes (LJ), permitindo estabilizar o sistema com contatos nativos. No caso de um sistema do tipo *coarse-grained*, contendo apenas átomos  $C\alpha$ , este potencial pode ser expresso como a seguir:

$$V_{ij}^C(r_{ij}) = V_{ij}^{LJ}(r_{ij}) = 5\epsilon \left(\frac{r_{ij}^N}{r_{ij}}\right)^{12} - 6\epsilon \left(\frac{r_{ij}^N}{r_{ij}}\right)^{10} \quad (2.5)$$

O formato deste potencial LJ é completamente definido pela distância  $r_{ij}$  entre os pares de átomos do sistema. Um potencial de SBM desenvolvido recentemente, bastante poderoso e utilizado neste trabalho utiliza funções gaussianas para representar interações não covalentes entre átomos, substituindo o potencial LJ.<sup>204</sup> Este potencial permite controlar o formato através da modificação independente de diversos parâmetros. A figura 48 descreve a diferença de parametrização entre os dois tipos de potenciais.

O potencial gaussiano pode ser definido por:

$$V_{ij}^C(r_{ij}) = V_{ij}^G(r_{ij}) = V_{ij}^R(r_{ij}) + A_{ij}G_{ij}(r_{ij}) + V_{ij}^R(r_{ij})G_{ij}(r_{ij}) \quad (2.6)$$

onde  $G_{ij}(r_{ij})$  é dado por:

$$G_{ij}(r_{ij}) = -\epsilon \exp\left(-\frac{(r_{ij} - r_{ij}^N)^2}{2w_{ij}^2}\right) \quad (2.7)$$

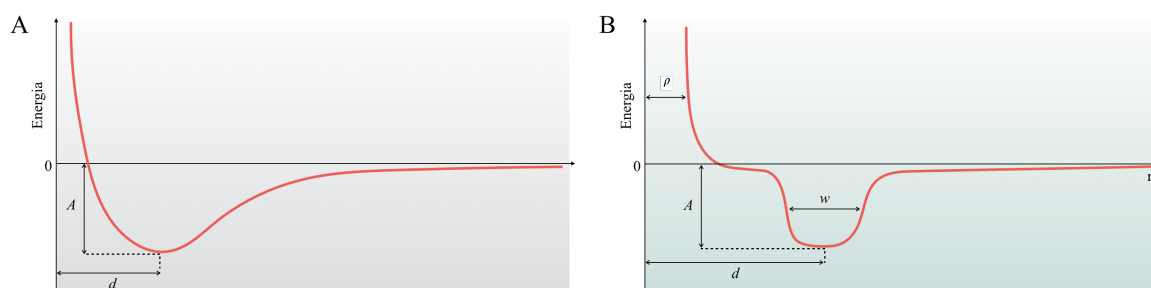


Figura 48 – Parâmetros ajustáveis em potenciais de modelos baseados em estrutura (SBM). (A) Potencial do tipo Lennard-Jones. (B) Potencial do tipo gaussiano. Os parâmetros ajustáveis são definidos como:  $A$  = amplitude,  $d$  = distância de equilíbrio estável (mínima energia),  $w$  = largura do poço de potencial e  $\rho$  = volume de exclusão atômica.

Fonte: Elaborada pelo autor.

O produto  $V_{ij}^R(r_{ij})G_{ij}(r_{ij})$  define o mínimo de energia do potencial na distância nativa. Os parâmetros  $A_{ij}$  e  $w_{ij}$  permitem o controle da amplitude (profundidade do poço de potencial, ou a força da interação) e do decaimento da função gaussiana (largura do poço de potencial) do potencial, respectivamente.<sup>204</sup>

### 2.3.4 Dinâmica Molecular

A fim de reproduzir o processo de dimerização dos sistemas selecionados, as estruturas dos complexos homo-diméricos foram separadas por 50 Å em dois monômeros. O sistema resultante foi submetido ao servidor SMOG gerando topologias específicas para o desenvolvimento de simulações do tipo *coarse-grained* com modelagem baseada em estrutura (SBM). Os contatos de DCA obtidos após o processo de filtragem descrito acima foram inseridos na topologia gerada pelo SMOG como potenciais Gaussianos representando interações entre resíduos que devem estar em contato na condição de mínima energia do sistema. Com a finalidade de evitar alterações substanciais nas conformações de cada monômero durante o processo de formação do complexo, as constantes que descrevem alterações nos diedros de cada cadeia foram multiplicadas por um fator de 100 em relação aos valores originais gerados pelo SMOG.

As simulações foram conduzidas em 7 etapas distintas, em que as distâncias de equilíbrio dos potenciais de DCA ( $r^N$ ) foram modificadas, junto com os parâmetros que descrevem a amplitude e o decaimento do potencial gaussiano ( $A$  e  $w$ , respectivamente, ver item 2.2.3). A distância de equilíbrio dos potenciais de DCA foi sistematicamente reduzida de 50 até 8 Å (distância média de interação entre dois átomos  $C\alpha$ ), a fim de permitir a orientação correta entre os monômeros ao longo do processo de aproximação entre eles. O decaimento do potencial gaussiano foi modificado para aumentar a capacidade de exploração conformacional na etapa em que ambos os monômeros começam a ter contato físico (com uma

distância de equilíbrio a 15 Å). Na equação 7, maiores valores de decaimento ( $w$ ) resultam em mínimos de energia mais largos nos potenciais gaussianos, o que se reflete em um maior número de conformações possíveis para uma determinada energia do sistema. Além disso, a amplitude dos potenciais, que reflete a força de interação entre o par de resíduos, foi elevada de um fator 3 para 5 (em relação as forças de interação monoméricas, de fator 1) nos 3 últimos passos de simulação, a fim de maximizar a estabilidade do complexo final encontrado. Cada etapa de simulação foi desenvolvida até a observação de uma estabilidade do sistema em uma população (*cluster*) conformacional, observada por uma redução na variação do RMSD. As combinações de parâmetros utilizadas em cada estágio de simulação são descritas na Tabela 11.

Tabela 11 – Parâmetros da função gaussiana utilizados ao longo do processo de simulação na formação de complexos homodiméricos.

Etapa de simulação	parâmetros		
	$r^N$ (Å)	$A$	$w$
1	50	3	2
2	30	3	2
3	15	3	4
4	11	3	4
5	8	5	4
6	8	5	2
7	8	5	0,5

Fonte: Elaborada pelo autor

## 2.4 Resultados e Discussão

### 2.4.1 Estudos de Correlação Evolutiva em Sequências Proteicas

Uma metodologia foi desenvolvida para extrair contatos diméricos de co-evolução em meio ao sinal de contatos monoméricos obtidos pela aplicação do método de DCA e para utiliza-los para guiar a formação de complexos proteicos em simulações de dinâmica molecular. Essa metodologia foi aplicada na predição de 11 complexos homodiméricos com diferentes características estruturais e níveis de complexidade (Tabela 10). Estes sistemas incluem dímeros simétricos e assimétricos, complexos com múlti-domínios e proteínas de médio e grande porte (até 444 resíduos), como apresentado na Tabela 10.

A principal idéia de aplicação da metodologia apresentada neste trabalho, no estudo de co-evolução entre resíduos envolvidos em processos de dimerização é a combinação da predição confiável de contatos entre resíduos através do DCA com a disponibilidade de dados das estruturas monoméricas (obtidas, por exemplo, por cristalografia de raios-X ou

por ressonância magnética nuclear). Esta abordagem proporciona um filtro natural para identificar os pares de resíduos que estão altamente acoplados por co-evolução, mas estão localizados na região interna e hidrofóbica da proteína. Estes resíduos são em sua maioria contatos necessários para manter a estrutura do monômero estável e garantir o enovelamento da proteína. Dessa forma, após filtrar todos os pares de interação que apresentavam uma baixa fração de superfície acessível ao solvente (SASA), resíduos localizados nas regiões do contatos monoméricos dentre os contatos preditos por DCA foram excluídos, a fim de obter somente os pares de interação relacionados à formação de complexos. Esse processo foi realizado comparando estes contatos preditos ao mapa de contato monomérico e excluindo os contatos sobrepostos. Os contatos de DCA resultantes do processo de filtragem são apresentados na Figura 49, junto com os contatos nativos e preditos para cada sistema.

Apesar de em alguns casos serem observados alguns contatos presentes tanto nas interações diméricas como monoméricas, a exclusão deles não resultou em uma redução significativa na precisão dos resultados de simulação. Os contatos remanescentes do processo de filtragem foram incorporados em um campo de força com potenciais gaussianos, de acordo com o método de SBM.<sup>204</sup>

Os contatos nativos presentes na estrutura cristalográfica original estão apresentados na parte superior esquerda dos mapas de contato na Figura 49. Os contatos apresentados do lado oposto (inferior esquerdo) correspondem aos contatos obtidos pela estrutura de melhor resolução predita através da simulação de dimerização utilizando sinais de DCA junto ao método de SBM, para cada sistema. Em ambos os mapas de contato, são apresentados dois tipos de interação: monoméricas e diméricas, apresentadas em cores distintas de acordo com a legenda na Figura 49. Pode-se observar uma considerável sobreposição entre contatos preditos pelo DCA (círculos pretos) e contatos diméricos presentes na estrutura nativa, confirmando a capacidade do método em prever contatos diméricos entre cadeias. Além disso, enquanto contatos de DCA próximos a estas regiões diméricas ocorrem em elevada frequência densidade, em outras regiões os contatos de DCA aparecem em uma distribuição esparsa. Estes contatos correspondem a pares falso-positivos que surgem pelo ruído intrínseco devido ao erro na correlação do conjunto de dados. Em alguns casos, como no complexo da proteína isocitrato desidrogenase (Figura 49 F), a proporção de contatos preditos em regiões não diméricas é muito pequena, e praticamente todos os contatos preditos estão localizados dentro das regiões diméricas de interação. Em outros casos, como no da tRNA metiltransferase, apesar da presença de diversos contatos corretamente preditos na região dimérica do mapa de contato (Figura 49 I), uma considerada fração das predições de DCA encontram-se esparsas ao longo do mapa, revelando um substancial número de predições incorretas. Apesar disso, a influência dos contatos incorretos mostrou-se mínima na predição dos sistemas selecionados. Isso ocorre pela a anulação destas interações incorretas devido a bloqueios estereoquímicos que as impedem de ocorrer. Além disso, devido à aleatoriedade dos pares incorretos, as forças de interação são muito pequenas para resultar

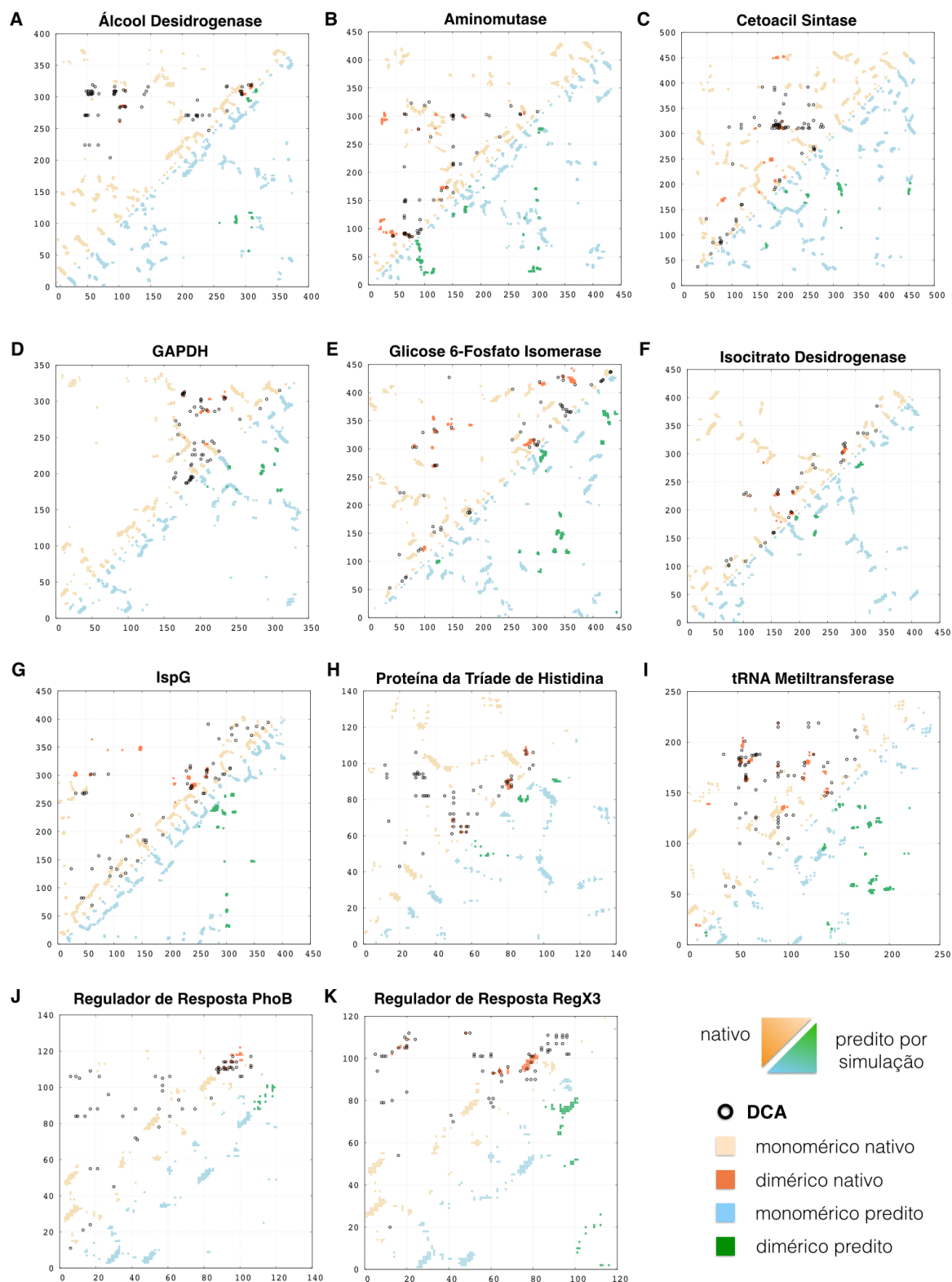


Figura 49 – Pares acoplados por co-evolução preditos por DCA e comparação com os mapas de contato dos complexos homodiméricos nativos e preditos por simulação.

Fonte: Elaborada pelo autor.



em uma mudança substancial na conformação do sistema, em comparação com a alta densidade local de pares diméricos corretamente preditos pelo DCA, que se reforçam de forma a direcionar a formação de interações estáveis.

Além das previsões incorretas do DCA, alguns dos pares de interações que não se sobrepõem exatamente à região correta do mapa de contato podem corresponder a interações elásticas devido à flexibilidade de certas interações entre cadeias laterais de resíduos.<sup>184</sup> Estes contatos podem ser identificados por estudos mais aprofundados das estruturas do tipo *all atom* destes complexos.

A elevada similaridade entre os contatos entre resíduos, observada pela comparação entre os mapas de contato dos complexos nativo e predito na Figura 49 (contatos apresentados em laranja e verde, respectivamente), nos permite confirmar a capacidade do método em reproduzir com exatidão os complexos diméricos através de interações provenientes da aplicação do DCA.

#### 2.4.2 Modelagem de Sistemas Proteicos Aplicando Modelos Baseados em Estrutura e Potenciais Gaussianos

Os contatos submetidos ao processo de filtragem descrito foram aplicados como interações atrativas entre monômeros para direcionar a formação dos complexos homodiméricos. O desvio quadrático médio (RMSD) de cada sistema em relação à estrutura correta de referência, ao longo do processo de simulação, é apresentado na Figura 50. O número de acoplamentos utilizados nas simulações variam entre 30 e 75 para todos os sistemas, de acordo com o número de contatos excluídos pela filtragem de interações monoméricas.

Utilizando o protocolo descrito na Figura 47, complexos diméricos de 11 sistemas distintos foram preditos. As estruturas resultantes das previsões são apresentadas na Figura 51. Os dois monômeros estão coloridos em azul e vermelho, enquanto os potenciais gaussianos provenientes do DCA e utilizados para uni-los são representados em verde.

De forma geral, as estruturas preditas por simulação apresentaram um RMSD médio total de 1,38 Å considerando a conformação de menor RMSD de cada simulação. O RMSD de interface (iRMSD, desvio quadrático médio entre interações inter-cadeia entre átomos a uma distância de no máximo 10 Å) médio total também apresentou um baixo valor de 1,25 Å, confirmando a alta precisão do método.

Além disso, a predição das estruturas dos complexos das proteínas com multi-domínios álcool desidrogenase e cetoacil sintase mostram a capacidade do método em lidar com sistemas em que várias famílias proteicas estão envolvidas. Dessa forma, além de predizer contatos diméricos como descrito na metodologia desenvolvida, também foram estimados contatos entre diferentes domínios, de forma equivalente à predição de interações em hete-

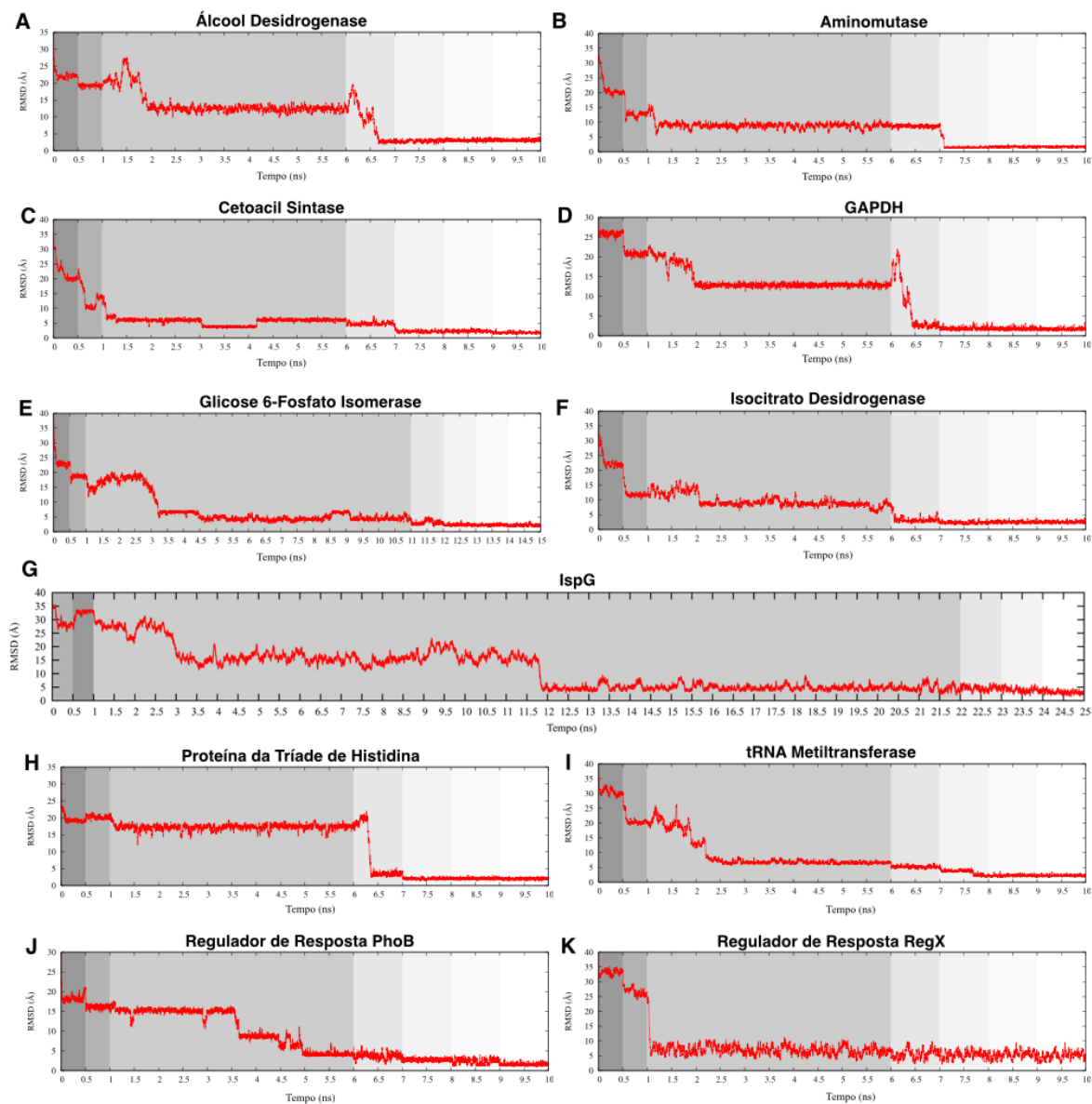


Figura 50 – RMSD das estruturas previstas (em relação aos homodímeros nativos) pelo método de simulação SBM/DCA, ao longo do tempo, em relação à estrutura do homodímero nativo. Diferentes gradientes de cinza representam as diferentes distâncias de equilíbrio e parâmetros das dos potenciais gaussianos utilizados em cada uma das 7 etapas de simulação (descritas na seção 2.2.4). Pode-se observar que todos os sistemas atingem uma conformação bastante estável e de baixo RMSD em todos os casos, indicando a formação de complexos proteicos.

Fonte: Elaborada pelo autor.

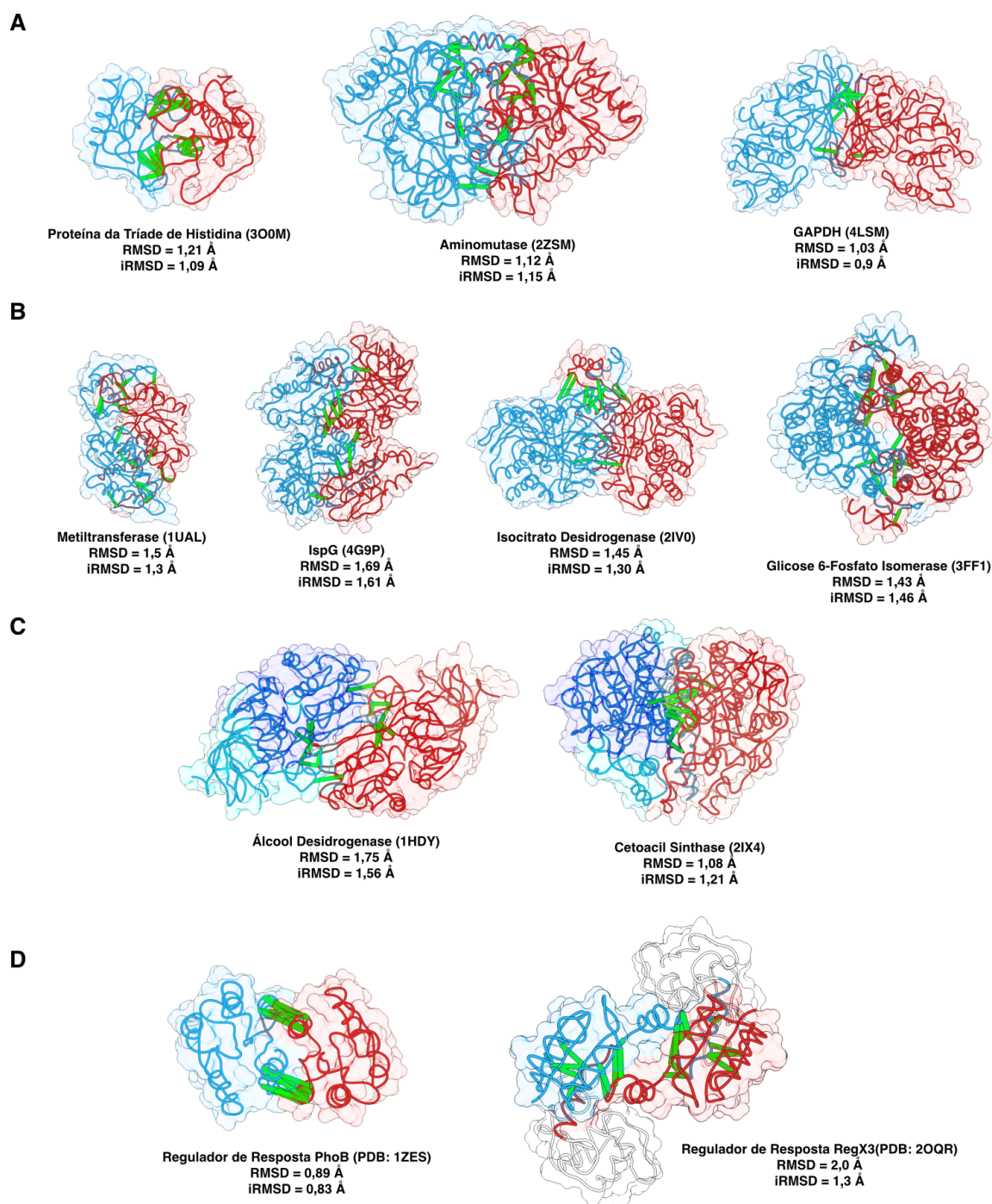


Figura 51 – Estruturas dos complexos homodiméricos preditos pelo método de DCA acoplado ao SBM. Os sistemas escolhidos representam proteínas de diferentes níveis de complexidade: (A) proteínas globulares: proteína da tríade de histidina, aminomutase e GAPDH; (B) proteínas não globulares: tRNA metiltransferase, IspG, isocitrato desidrogenase e glicose 6-fosfato isomerase; (C) proteínas com multidomínios: álcool desidrogenase e cetoacil sintase; e (D) proteínas capazes de dimerizar de diferentes conformações: PhoB e ReX3.

Fonte: Elaborada pelo autor.

rodímeros.<sup>186,188,191</sup> Esses resultados demonstram a aplicabilidade do método em sistemas de alta complexidade, com diversas proteínas envolvidas.

A Tabela 12 apresenta um resumo dos melhores valores de RMSD e iRMSD obtidos para todos os sistemas

Como pode-se observar, para todos os sistemas estudados neste trabalho o método desenvolvido apresentou uma excelente capacidade de predição, com valores de RMSD e iRMSD abaixo de 2 Å e 1,6 Å, respectivamente, em todos os casos.

### 2.4.3 Identificação das Diferentes Interfaces de Proteínas Reguladoras de Resposta

Um caso especial estudado neste trabalho é o das proteínas PhoB e RegX3, que atuam como reguladores de resposta. Estas macromoléculas são membros de uma grande família de proteínas, principalmente procariotas, com mais de cem mil sequências catalogadas no servidor Pfam. Eles estão envolvidos em vias de sinalização, onde seu domínio receptor (Pfam PF00072) é fosforilado por uma histidina quinase (Pfam PF00512). Este evento provoca uma alteração conformacional e promove a dimerização da proteína fosforilada, activando a sua função como fator de transcrição. Em seguida, ela se liga ao DNA e propaga uma série de eventos em um processo de cascata metabólica.<sup>205-207</sup> A homodimerização do domínio receptor (REC) é fundamental para possibilitar a conformação do seu estado ativo.<sup>208</sup> A proteína PhoB, que efetua este processo de regulação em *E. coli*, dimeriza em uma interface tipicamente reportada como interface entre cadeias ( $\alpha4$ - $\beta5$ - $\alpha5$ ).

A Figura 49 J mostra o resultado da aplicação da metodologia de SBM e DCA para o complexo formado pelo domínio REC da PhoB após ativação. Uma série de contatos diméricos entre resíduos na região 90-120 (laranja) são detectados por DCA (círculos pretos). O complexo foi previsto com uma precisão de 0,89 Å de RMSD, com respeito à estrutura cristalográfica (PDB: 1ZES), para o melhor caso e com uma média de 1,57 Å para a última etapa de simulação em que o sistema encontram-se em uma conformação estável. Isto sugere a presença de um sinal de co-evolução claro para o estado a formação do complexo ativo. Estudos recentes também sugerem que os dímeros do estado ativo contendo o domínio REC de proteínas reguladoras de resposta podem apresentar conformações alternativas. Uma dessas conformações envolve a troca de domínio (*domain swapping*) das hélices  $\alpha4$  e  $\alpha5$  e folha  $\beta5$ , bem como a formação de contatos diméricos que diferem dos previamente reportados.<sup>209</sup> A metodologia proposta foi aplicada para as estruturas monoméricas da proteína de transdução sensorial de *M. bovis* RegX3 (PDB: 2OQR), que forma um dímero utilizando esta interface ativa alternativa proposta recentemente. O mapa de contato da Figura 49 K mostra que alguns dos contatos monoméricos presentes na Figura 49 J tornam-se diméricos para o dímero da proteína RegX3. Isto pode ser observado com mais detalhes na Figura

Tabela 12 – Valores de RMSD e iRMSD obtidos para as estruturas moleculares preditas para os sistemas de estudo.

<b>Proteína Composto</b>	<b>código PDB Estrutura Molecular<sup>a</sup></b>	<b>RMSD clogP</b>	<b>iRMSD</b>	<b>RMSD médio</b>	<b>iRMSD médio</b>
Proteína da Triade de Histidina	3O0M	1,21	1,09	1,98	1,95
GAPDH	4LSM	1,04	0,90	1,71	1,37
Isocitrato Desidrogenase	2IV0	1,46	1,31	2,57	2,57
Álcool desidrogenase	1HDY	1,75	1,56	3,13	2,81
Aminomutase	2ZSM	1,12	1,15	1,67	1,87
tRNA metiltransferase	IUAL	1,50	1,28	2,27	1,80
IspG	4G9P	1,69	1,61	3,15	2,69
Cetoacil Sintase	2IX4	1,09	1,21	1,82	2,10
Glicose 6-Fosfato Isomerase	3FF1	1,43	1,47	2,22	2,29
RegX3	2OQR	2,01	1,33	5,09	2,32
PhoB	1ZES	0,89	0,84	1,57	1,67

Fonte: Elaborada pelo autor

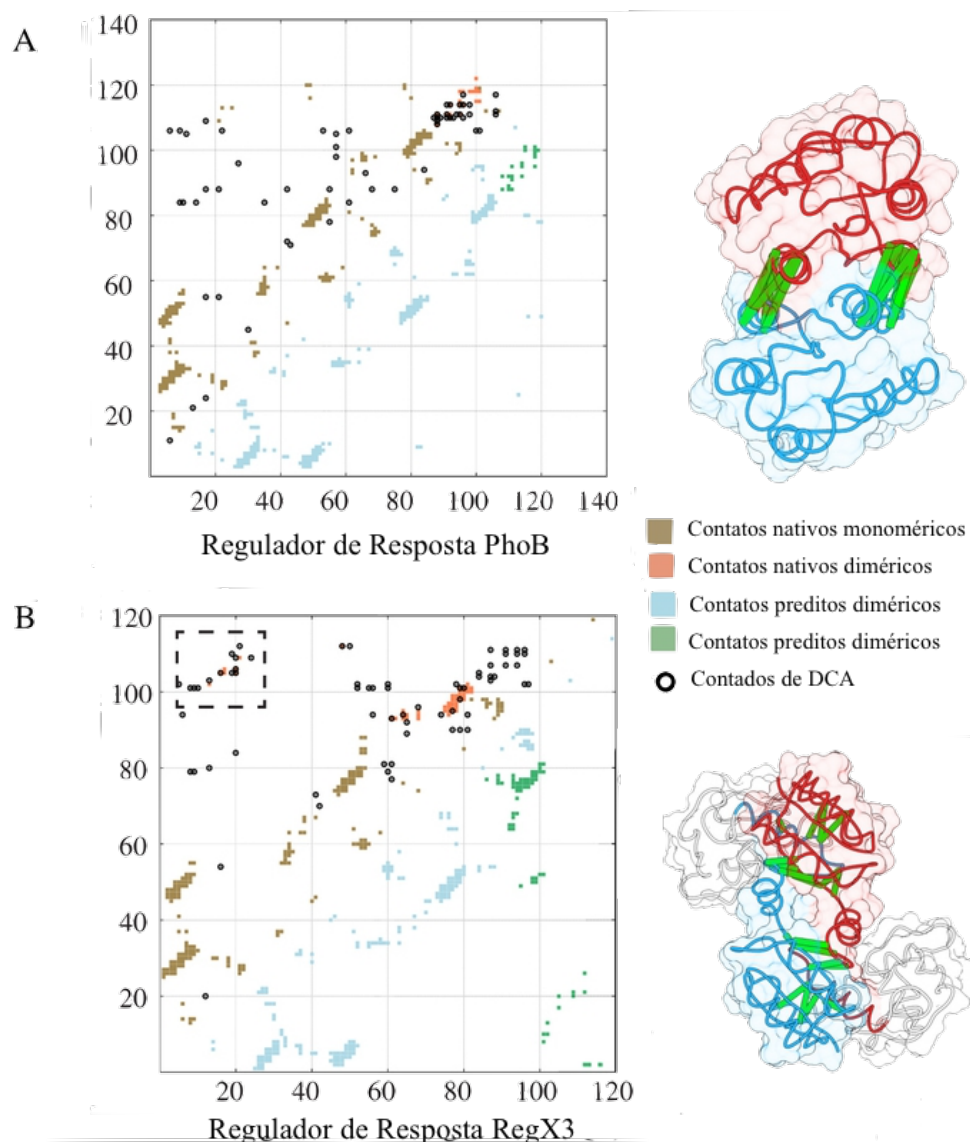


Figura 52 – Comparação entre as estruturas dos complexos homodiméricos preditos para as proteínas reguladoras de resposta (A) PhoB e (B) ReX3. Estas proteínas são capazes de dimerizar de diferentes conformações.

Fonte: Elaborada pelo autor.

52. Além disso, o sinal de acoplamento de uma outra região de contatos que envolve a interação dos resíduos 10-20 com os resíduos 100-110, também é capturado usando análise de co-evolução (Figura 52). Estas duas regiões de contato levaram à formação deste complexo alternativo com uma resolução de 2 Å de RMSD (com iRMSD = 1,33 Å).

A RegX3 é uma proteína de múltiplos domínios, contendo o domínio efetor comum à proteína PhoB. Calculando-se o RMSD apenas para o domínio regulador de resposta (REC), presente em ambas as proteínas, chega-se a um valor de RMSD = 1,13 Å. Isto implica que os sinais co-evolutivos para conformações múltiplas diméricas estão presentes em ambos os casos e podem ser utilizados para prever várias configurações fisiologicamente

relevantes (Figura 52).

## 2.5 Conclusão

O uso da informação de co-evolução no estudo de alterações conformacionais e interações de proteínas é um método promissor e alvo de muitos estudos atualmente. O número de interações utilizados neste método necessário para reconstruir precisamente complexos moleculares é muito menor que métodos baseados em predição de estruturas, tornando-o uma abordagem particularmente efetiva na identificação de interações intermoleculares.

Dentre os tipos de interações entre proteínas no ambiente celular, a formação de homo-oligômeros é bastante comum e essencial a na função de sistemas vitais à célula, como regulação genética e manutenção de estrutura celular. Os resultados apresentados neste trabalho descrevem um método bastante eficaz de predizer interações entre proteínas e formação de complexos proteicos com considerável precisão, para os mais diversos sistemas moleculares.

A predição de homodímeros apresenta o desafio adicional comparado à interação hetero-oligomérica, por ter sinais de interação entre cadeias misturados aos sinais monoméricos de enovelamento. Um método foi desenvolvido para extrair estes diméricos e utilizá-los no estudos de interações entre proteínas. Esta metodologia também mostrou-se capaz de identificar conformações múltiplas de dímeros, como apresentado no caso de dímeros da proteína reguladora de resposta. Estes estudos aqui apresentados abrem possibilidades para explorar sistemas moleculares de alta complexidade e para elucidar o funcionamento de diversos sistemas macromoleculares ainda pouco conhecidos, difíceis de explorar por técnicas de cristalografia de raios- X ou RMN. Além disso, este método pode ser utilizado no desenho racional de drogas voltadas à inibição por interação as interfaces diméricas de um determinado sistema alvo. Finalmente, as interações homo-oligoméricas são relevantes para o estudo específico de doenças geradas por agregação proteica, como Alzheimer e Parkinson.<sup>210-213</sup>





### 3 Construção de uma Base de Dados Virtual de Produtos Naturais





## 3.1 Introdução

### 3.1.1 A importância dos Produtos Naturais no Desenvolvimento de Novos Fármacos

Os produtos naturais têm apresentado um papel fundamental na identificação de novos medicamentos ao longo da história humana.<sup>214-217</sup> Uma análise atual dos fármacos revela que cerca de 64% de todos os fármacos disponíveis tiveram um produto natural envolvido em alguma etapa de seu desenvolvimento.<sup>218</sup> Esta fração inclui produtos naturais não modificados, derivados de produtos naturais e fármacos que tiveram seu planejamento inspirado em um produto natural.

Alguns dos exemplos mais famosos de fármacos de origem natural são o paclitaxel (já apresentado na primeira parte desta tese de doutorado), a morfina e a quinina (Figura 53 A-C, respectivamente). Mesmo atualmente, os produtos naturais continuam sendo utilizados no planejamento de novos fármacos, estando envolvidos em aproximadamente 50% de todos os fármacos baseados em pequenas moléculas desenvolvidas em 2010.<sup>218</sup> A grande diversidade e complexidade estrutural desta fonte de compostos e de seus metabólitos secundários faz com que os produtos naturais ainda representem uma rica fonte de idéias para o desenvolvimento de novos tratamentos terapêuticos.<sup>219</sup> Apesar da importância dos produtos naturais na química medicinal e do Brasil apresentar um vasto e rico bioma de complexidade única, a exploração de compostos isolados de origem natural com potencial atividade biológica é escassa. O maior motivo deste problema é a falta de acesso a informações disponíveis de uma forma organizada em um banco de dados e de interação entre grupos de diferentes especialidades. Portanto, a criação de uma base de dados contendo todas as informações possíveis de um conjunto relevante de produtos naturais isolados do rico bioma brasileiro representa um passo extremamente importante e de grande interesse científico para o desenvolvimento de novos fármacos no Brasil.<sup>220</sup>

### 3.1.2 A Biodiversidade no Brasil e os Compostos Identificados pelo NuBBE

O Brasil possui uma das maiores biodiversidades do mundo, contribuindo com aproximadamente 20% de todas as espécies vivas conhecidas, distribuídas em vários biomas importantes como a Floresta Amazônica e a Mata Atlântica.<sup>221</sup> Existem vários grupos de pesquisa no Brasil focados na exploração racional desta rica biodiversidade. Um deles é o Núcleo de Bioensaios Biossíntese e Ecofisiologia de Produtos Naturais (NuBBE), localizado na Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP) no campus de Araraquara. O NuBBE esteve envolvido com os últimos avanços na química de produtos naturais, incluindo a busca de compostos biologicamente ativos de plantas do Cerrado e Mata Atlân-

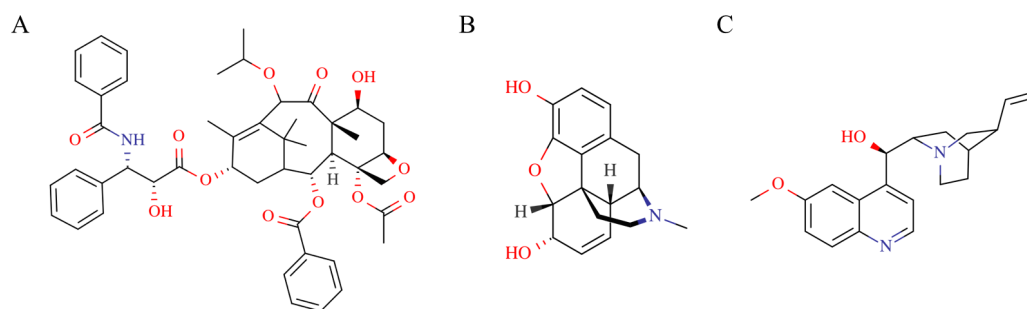


Figura 53 – Estruturas moleculares de alguns fármacos de origem natural. (A) Paclitaxel, utilizado no tratamento de diversos tipos de tumores e isolado do teixo do pacífico. (B) Morfina, um potente anestésico e narcótico proveniente da papoila do ópio. (C) Quinina, um composto com atividades antitérmicas, antimaláricas e analgésicas extraído da árvore cinchona.

Fonte: Elaborada pelo autor.

tica, de organismos marinhos e de fungos.<sup>222</sup> Algumas estruturas, apresentadas na Figura 54, ilustram a diversidade química e estrutural dos compostos isolados pelo NuBBE.<sup>223–228</sup>

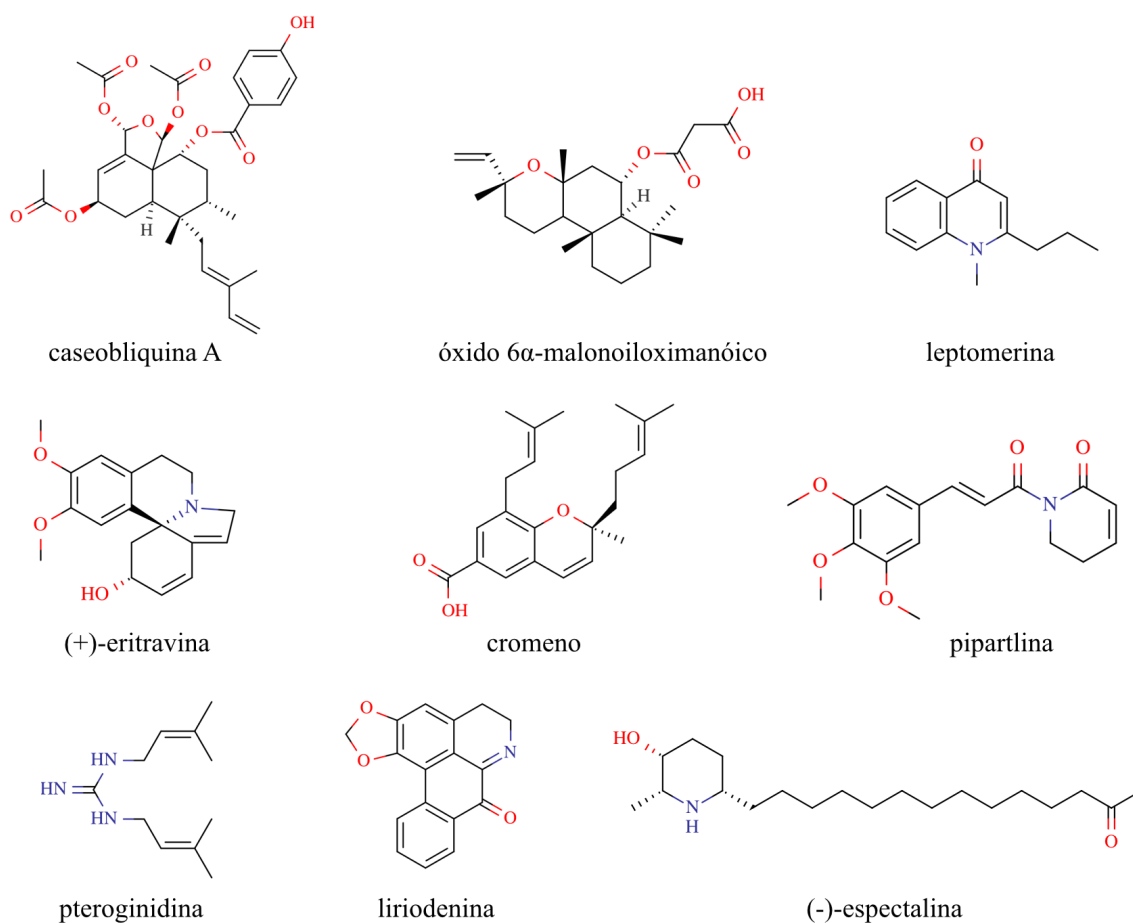


Figura 54 – Estruturas moleculares de alguns dos produtos naturais isolados pelo grupo NuBBE.

Fonte: Elaborada pelo autor.

## 3.2 Motivações e objetivos

O Brasil possui uma das maiores biodiversidades do mundo, contribuindo com aproximadamente 20% de todas as espécies vivas conhecidas, distribuídas em vários biomas importantes como a Floresta Amazônica e a Mata Atlântica. Apesar deste grande privilégio do nosso país, a exploração da biodiversidade brasileira e sua aplicação na criação de produtos de alto valor agregado e de alta necessidade para a população é ainda bastante precária. Dentre os recursos de maior potencial econômico e social estão os produtos naturais presente na flora de nosso país, de alta complexidade química e que poderiam ser utilizados para o desenvolvimento de novos tratamentos terapêuticos e para um crescimento da indústria farmacêutica no Brasil. Grande parte da subutilização desta fonte natural e renovável de recursos se deve à falta de centralização, organização e disponibilidade das informações já obtidas ao longo dos anos através do estudo de nosso bioma.

Neste trabalho, todas as informações de compostos já isolados pelo Núcleo de Bioensaios, Biossíntese e Ecofisiologia de Produtos Naturais (NuBBE), localizado na Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP) foram reunidas e organizadas. Estes dados foram utilizados para construir uma base de dados de livre acesso para auxiliar a comunidade científica e as indústrias farmacêuticas fornecendo compostos naturais de origem brasileira para planejamento e a identificação de novos compostos fármacos.

Os objetivos deste trabalho são:

- Reunir e organizar todas as informações de produtos naturais já isolados em uma base virtual de dados.
- Construir uma biblioteca de produtos naturais através da construção das estruturas tridimensionais e cálculo de propriedades físico-químicas importantes para o perfil farmacocinético.
- Disponibilizar todos os dados obtidos de uma forma clara e eficiente em um servidor gratuito na internet.

## 3.3 Metodologia

### 3.3.1 Construção de uma Base Virtual de Produtos Naturais de Livre Acesso

Como parte dos estudos desenvolvidos ao longo deste projeto de doutorado, a rica coleção de produtos naturais isolados e catalogados pelo grupo NuBBE ao longo de 11 anos

de pesquisa foi utilizada para construir uma base de dados virtual de produtos naturais de acesso livre, intitulada NUBBE<sub>DB</sub> (do inglês, *NuBBE database*). Este trabalho foi realizado através da idealização e colaboração entre os pesquisadores Dr. Adriano D. Andricopulo (coordenador do Laboratório de Química Medicinal e Computacional, LQMC) e Dra. Vanderlan S. Bolzani (coordenadora do NuBBE). Todas as etapas do desenvolvimento da base foram realizadas em colaboração com a aluna de doutorado Marília Valli (Processo FAPESP: 2010/17329-7), orientada pela prof. Dra. Vanderlan S. Bolzani, que desenvolveu o processo de busca e organização dos compostos já isolados no NuBBE.

A base de dados desenvolvida consiste na reunião de todas as informações existentes de 1455 compostos de origem natural isolados pelo NuBBE. Para cada composto estão descritas informações como propriedades físico-químicas, organismo de origem, referências bibliográficas e propriedades farmacológicas já determinadas, além da sua estrutura tridimensional em um formato apropriado para estudos de modelagem molecular (Figura 55). As informações e número de compostos da base de dados serão atualizados constantemente. Este trabalho tem como objetivo principal tornar acessível para a comunidade acadêmica e industrial as informações já obtidas a partir da rica biodiversidade brasileira, permitindo o desenvolvimento de estudos na área de química medicinal e a identificação de novos compostos bioativos em nosso país.

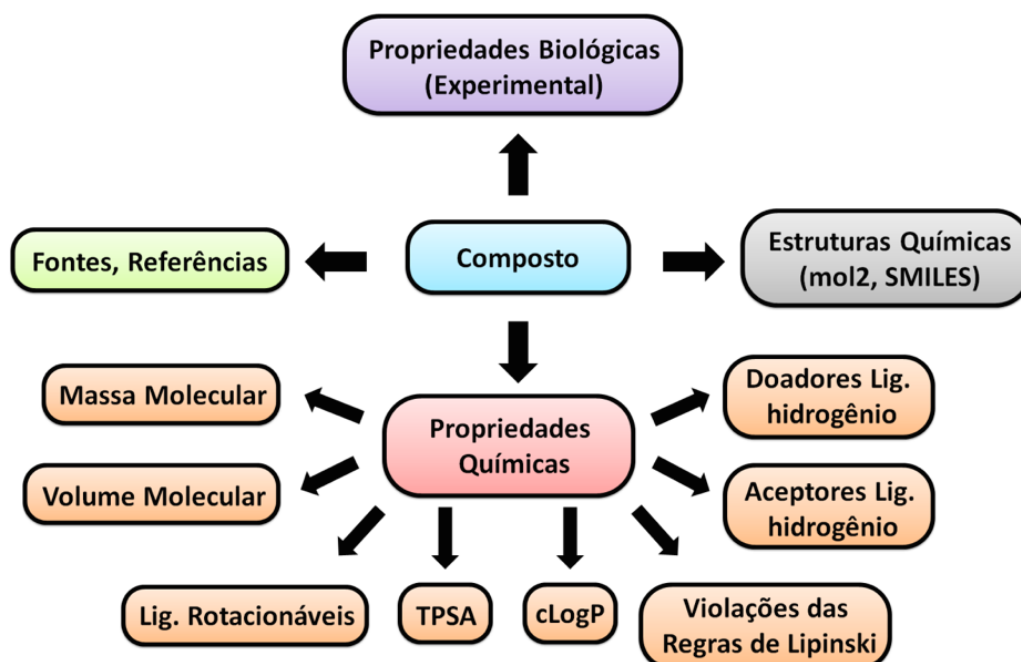


Figura 55 – Descrição da metodologia utilizada na organização e construção da base de dados NuBBE<sub>DB</sub>.

Fonte: Elaborada pelo autor.

### 3.3.2 Organização dos Compostos e Construção das Estruturas Moleculares 3D

As publicações acadêmicas resultado de onze anos de pesquisa do NuBBE foram analisadas individualmente permitindo a obtenção de todas as informações dos compostos já isolados. O NuBBE<sub>DB</sub> dispõe de 1455 compostos, que serão constantemente atualizados conforme a publicação de novos dados. As seguintes informações foram disponibilizadas para cada composto da base: código NuBBE, nome químico, fórmula molecular, classe química, fonte (plantas, fungos, etc.) e a espécie a partir da qual o composto foi isolado. Para os compostos que apresentam propriedades biológicas, farmacológicas ou toxicológicas (dados quantitativos ou qualitativos), estas informações foram disponibilizadas.

Todos os compostos selecionados foram convertidos em uma biblioteca virtual através do desenho das estruturas bidimensionais através do módulo *msketch* do programa MarvinSuite (ChemAxon, Budapeste, Hungria). A estrutura química 3D de cada composto foi gerado utilizando as ferramentas disponíveis no programa OpenBabel para converter as estruturas dos compostos do formato bidimensional para tridimensional, gerando automaticamente as coordenadas de cada molécula.<sup>229</sup> A hibridização de cada átomo das estruturas geradas foram analisadas pela opção *Atom Type* do programa SYBYLX2.0 (Tripos, Saint Louis, MO, EUA). As cargas totais de todas as moléculas da base foram consideradas com neutras. A representação geométrica 3D de cada molécula do banco de dados teve a sua energia conformacional minimizada utilizando o campo de força Tripos e o método de Powell de minimização de energia. As cargas parciais atômicas das estruturas minimizadas foram calculadas através do método de Gasteiger-Huckel, também disponível no SYBYL 8.0.<sup>36,82-84</sup> Durante estas etapas, as moléculas foram consideradas em um solvente implícito aquoso (constante dielétrica de 80,0). Finalmente, a conformação 3D de cada molécula foi adicionada à base de dados como um arquivo no formato *Mol2*, reconhecido pela maioria dos programas de modelagem molecular.<sup>87</sup>

### 3.3.3 Predição de Propriedades Físico-Químicas

As seguintes propriedades físico-químicas dos compostos isolados pelo NuBBE foram calculadas computacionalmente: massa molecular, número de ligações rotacionáveis, coeficiente de partição calculado ( $\log P$ ), número de doadores e aceptores de ligação de hidrogênio, número de violações da "regra de 5" de Lipinski, área de superfície polar topológica (TPSA) e volume molecular.<sup>87, 143, 149, 230</sup> Estas propriedades foram preditas por meio do programa Molinspiration, disponível na internet e de livre acesso.<sup>149</sup> A abordagem baseada fragmento utilizada pelo Molinspiration provou ser confiável, sendo empregada em publicações científicas relevantes de bases de dados de moléculas.<sup>231-233</sup>

### 3.3.4 Montagem do Site de Busca NuBBE

O processo de construção e divulgação da página da base de dados na internet teve o auxílio do técnico especialista em programação Leandro D. Figueira, funcionário da Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP) e membro integrante do grupo NuBBE. A base de dados NuBBE<sub>DB</sub> foi instalada em um sistema operacional Linux utilizando um servidor Apache Tomcat para disponibilização dos dados na Web. Todos os conjuntos de dados foram armazenado em uma base de dados PostgreSQL. Um sistema de busca na base NuBBE foi desenvolvido a fim de permitir a fácil aquisição dos dados de um composto específico, em qualquer navegador comum, pela inserção de características estruturais ou de critérios físico-químicos. A interface de navegação foi implementada utilizando ferramentas comuns da web, como HTML, CSS e JavaScript (AJAX), enquanto que o próprio servidor foi implementado usando as aplicações Java/Servlets e Hibernate. Todos os programas utilizados são gratuitos e reconhecidos pela indústria e pela comunidade acadêmica como robustos e confiáveis. Para permitir ao usuário a busca de compostos pelo desenho de estruturas moleculares, a interface de desenho molecular WebME (fornecida pelo grupo desenvolvedor do Molinspiration) foi implementada em associação com o método de busca de subestruturas fornecido pelo aplicativo CDK (Chemistry Development Kit).<sup>149,234</sup>

## 3.4 Resultados e Discussão

### 3.4.1 Caracterização do Site de Busca NuBBE de Produtos Naturais

A base virtual de produtos naturais NuBBE<sub>DB</sub> encontra-se atualmente disponível na página da web: <http://nubbe.iq.unesp.br/nubbeDB.html>. Analisando a organização dos compostos de forma geral, é possível classifica-los a partir das fontes naturais de obtenção, como apresentado na Figura 56. A NuBBE<sub>DB</sub> é composta atualmente por 80% de compostos isolados de plantas, 6% a partir de fungos ou microorganismos, 7% de compostos sintéticos inspirados em produtos naturais, 5% de compostos semi-sintéticos e 2% de produtos de biotransformação utilizando plantas ou extratos de fungos (Figura 56). Levando em conta esta distribuição, é evidente que o banco de dados NuBBE está principalmente relacionado com produtos naturais, apresentando compostos únicos em estrutura e complexidade química. A diversidade química é consideravelmente grande, com compostos pertencentes às mais diversas classes (terpenos, alcalóides, flavonóides, iridóides, lignanos, etc).

Algumas propriedades, tais como massa molecular, número de doadores e aceptores de ligações de hidrogênio, *clogP*, número de ligações rotacionáveis e área de superfície polar topológica (TPSA) são parâmetros úteis para a predição de propriedades farmacoló-



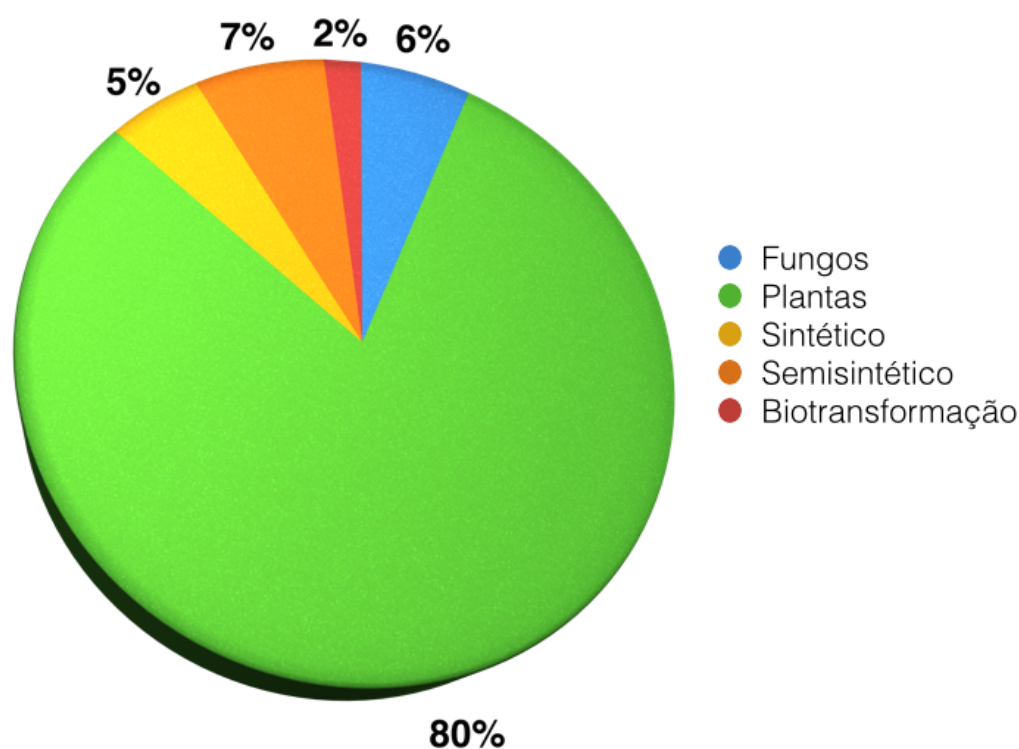


Figura 56 – Distribuição das fontes de origem dos compostos presentes na base NuBBE<sub>DB</sub>.

Fonte: Elaborada pelo autor.

gicas.<sup>98,99,143,235,236</sup> As distribuições destas propriedades estão apresentadas na Figura 57.

A massa e o volume molecular apresentam distribuições muito semelhantes, mostrando alta correlação como esperado (Figura 57 A e B). A massa molecular média do conjunto de compostos é de 386,21. Os valores calculados do coeficiente de partição ( $\log P$ ) seguem uma distribuição gaussiana simétrico (Figura 57 C), abrangendo desde substâncias muito hidrofílicas ( $\log P$  com valores na faixa de -4) até muito hidrofóbicas ( $\log P$  na faixa de 9). O  $\log P$  médio para o conjunto de compostos é de 2,98 (Figura 57 C), portanto, espera-se que a maior parte dos compostos da base de dados seja solúvel em condições padrões de ensaio. A área de superfície polar topológica (TPSA) é um parâmetro útil para prever a capacidade de absorção. Compostos com valores maiores que 140 Å<sup>2</sup> geralmente apresentam um baixo nível de absorção.<sup>237</sup> A distribuição da TPSA do NuBBE<sub>DB</sub> apresenta um pico em valores no intervalo de 50 a 100 Å<sup>2</sup>, sugerindo um bom nível de absorção para boa parte dos compostos (Figura 57 D).

Notavelmente, as porcentagens de compostos com não mais do que 10 aceptores de ligação de hidrogênio (529 compostos) e 5 doadores de ligação de hidrogênio (539 compostos) são ambas de aproximadamente 80% (Figuras 57 E e F), seguindo exatamente a proporção dador/aceptor proposta por Lipinski para compostos com características de fármacos.<sup>98</sup> O número de ligações rotacionáveis é um parâmetro útil para avaliar a flexibili-

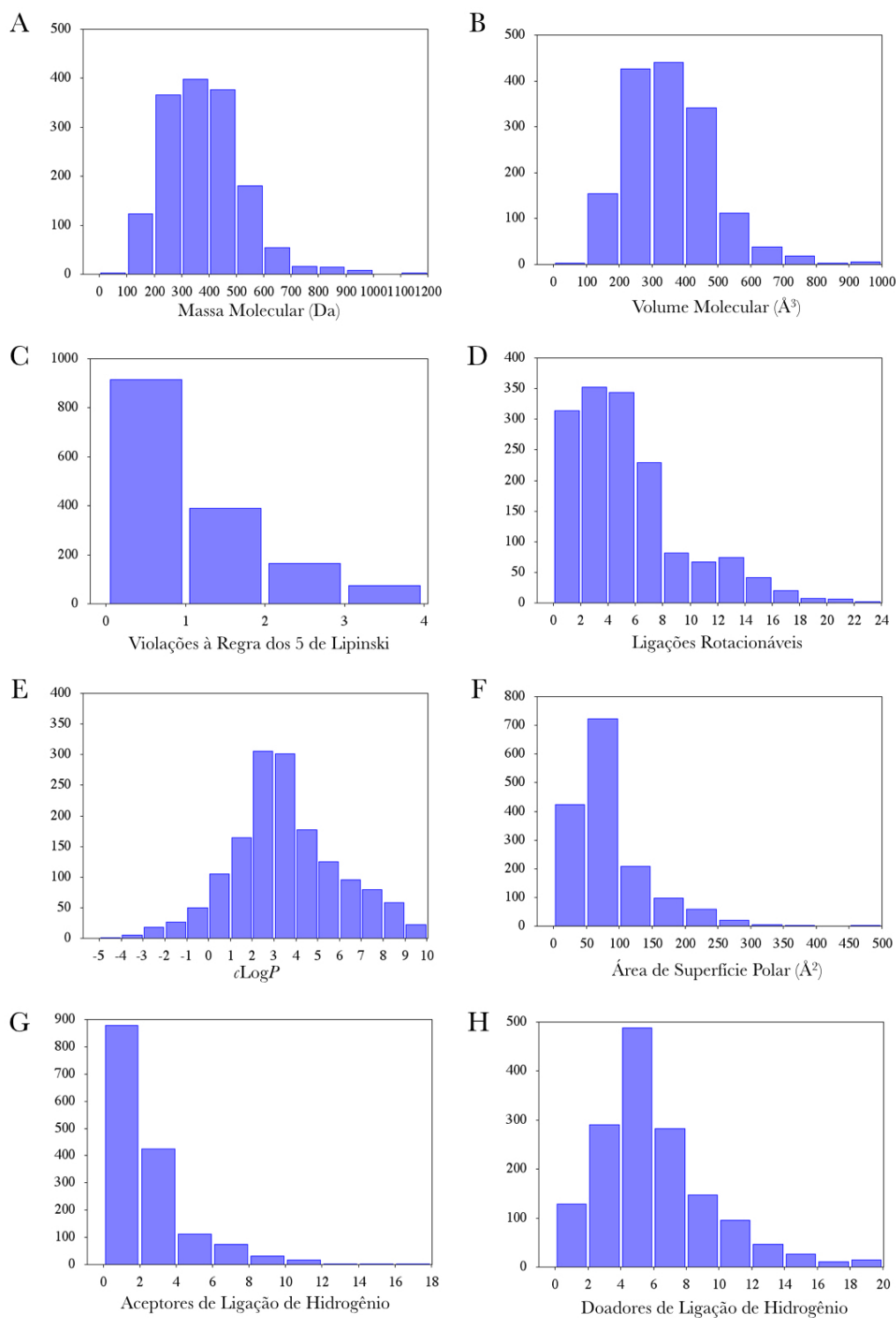


Figura 57 – Distribuições das características físico-químicas da base de dados NuBBE<sub>DB</sub>. (A) Massa molecular, em Daltons. (B) Volume molecular, em Å<sup>3</sup>. (C) Número de violações da “Regra dos 5” de Lipinski. (D) Número de ligações rotacionáveis. (E) Coeficiente de partição calculado, *cLogP*. (F) Área de superfície polar topológica calculada, TPSA, em Å<sup>2</sup>. (G) Número de aceptores de ligação de hidrogênio (H) Número de doadores de ligação de hidrogênio.

Fonte: Elaborada pelo autor.

dade molecular e para prever a biodisponibilidade oral de fármacos. Números maiores que 10 ligações rotacionáveis têm sido associados à diminuição da biodisponibilidade oral.<sup>235</sup> A maioria dos compostos de base de dados apresenta até 6 ligações rotativas (67%, Figura 57 G). A distribuição diminui drasticamente para valores maiores, mas estende-se contabilizando compostos muito flexíveis, com até 24 ligações rotacionáveis. O número médio de ligações rotacionáveis é de 5,8. A aplicação do critério de Lipinski para avaliar a semelhança de compostos com fármacos de sucesso (Regra dos 5 de Lipinski) mostra que um total de 355 compostos, 55% do banco de dados, obedecem os critérios de Lipinski. Além disso, nenhum dos compostos em toda a base viola todos os quatro critérios de Lipinski (Figura 57 D). Levando em conta todas as propriedades consideradas, pode-se notar que o banco de dados apresenta uma diversidade química bastante elevada e um perfil muito interessante para o desenvolvimento de estudos direcionados ao planejamento e à identificação de compostos bioativos.

A interface de busca desenvolvida para a base de dados NuBBE é extremamente intuitiva e de fácil utilização (Figura 57). A busca de compostos pode ser realizada rapidamente através da seleção de propriedades químicas, biológicas, fonte de extração ou referências (Figura 57 A) ou pelo desenho de subestruturas moleculares utilizando a ferramenta de desenho WebME (Figura 57 B). Os resultados de cada busca são apresentados em tabelas (Figura 57 C) a partir das quais é possível ser direcionado para uma página contendo todas as informações detalhadas do composto, além do link para download da sua estrutura tridimensional (Figura 57 D).

**A**

**General Information**  
 Name:   
 Chemical Class:   
 SMILES:   
 Mol. Formula:   
 Additional Info:

**Species**  
 [Choose an option, or refine below...]  
 Family:   
 Genus:   
 Species:   
 Refine

**Chemical Information**  
 ≤ Molecular Weight ≤   
 ≤ cLogP ≤   
 ≤ TPSA ≤   
 ≤ Lipinski Violations ≤   
 ≤ H-bond acceptors ≤   
 ≤ H-bond donors ≤   
 ≤ Rotatable Bonds ≤   
 ≤ Molecular Volume ≤

**Source**  
 Any  
 Synthesis  
 Semisynthesis  
 Biotransformation product  
 Isolated from a plant  
 Isolated from a microorganism  
 Isolated from a marine organism

**Reference**  
 [Choose an option, or refine below...]  
 Title:  Author:  Year:  to   
 Refine

**Biological property**  
 [Any property, or else choose one...]

**B**

**Structure**  
 molinspiration WebME  
 [Add group]  
 [Search compound(s)]

**C**

**Results >>**  
 218 compounds.  
 Download all molecular structure files of the results: [.mol2](#)  
 Download a PDF file containing all information of the results: [NuBBE\\_Compounds\\_Info.pdf](#)  
 Compounds: [1...30](#) [31...60](#) [61...90](#) [91...120](#) [121...150](#) [151...180](#) [181...210](#) [211...218](#)

<a href="#">NuBBE 563</a>	(-)-Butin Flavonoid Antioxidant <a href="#">NuBBE563.mol2</a>	
<a href="#">NuBBE 741</a>	(-)-Epigallocatechin Catechin (flavonoid) Antioxidant <a href="#">NuBBE741.mol2</a>	

**D**

**General Information**  
 NuBBE:   
 Name:   
 Chemical Class:   
 Mol. Formula:   
 SMILES:   
 Additional Info:

**Chemical Information**  
 Molecular Weight:   
 cLogP:   
 TPSA:   
 Lipinski Violations:   
 H-bond acceptors:   
 H-bond donors:   
 Rotatable Bonds:   
 Molecular Volume:

**Biological property**  
 Antioxidant  
 Add Info

Figura 58 – Apresentação da página de acesso da base de dados NuBBE<sub>DB</sub>. (A) Opção de busca de compostos por informações ou propriedades físico-químicas (B) Opção de busca pelo desenho de estruturas químicas (C) Exemplo de resultado de uma busca de compostos na NuBBE<sub>DB</sub>. (D) Exemplo das informações disponibilizadas de um composto da base.

Fonte: Elaborada pelo autor.

## 3.5 Conclusão

A base de dados NuBBE<sub>DB</sub> descrita neste trabalho é uma iniciativa inovadora e pioneira que tem como principal objetivo reunir e disponibilizar o máximo possível de toda a informação obtida através do estudo da biodiversidade brasileira ao longo de vários anos. Ela representa um esforço para incentivar a colaboração entre pesquisadores de áreas distintas, como as de produtos naturais, síntese, quimio-informática e química medicinal. A NuBBE<sub>DB</sub> será de grande utilidade para a comunidade científica, fornecendo informações para estudos de modelagem molecular, metabolômica, derreplicação e principalmente para o planejamento e a identificação de novos compostos bioativos.

## 4 Artigos publicados ou submetidos em periódicos científicos.

### 4.1 Artigos publicados

- SANTOS, R. N.; GUIDO, R. V. C.; OLIVA, G.; DIAS, L. C.; ANDRICOPULO, A. D. Quantitative structure-activity studies on a series of migrastatin analogs as inhibitors of cancer cell metastasis. *Medicinal Chemistry*, v. 7, n. 3, p. 155–164, 2011. doi: 10.2174/157340611795564240.
- VALLI, M.; SANTOS, R. N.; FIGUEIRA, L. D.; VIEIRA JR, G. M.; FUNARI, C. S.; REGASINI, L. O.; LOPES, M. N.; CAVALHEIRO, A. J.; ARAÚJO, A. R.; FURLAN, M. Tubulin ligands identified on screening natural products from nubbe database. *Planta Medica*, v. 78, n. 11, p. 1101-1101, 2012. doi: 10.1055/s-0032-1320486.
- VALLI, M.; DOS SANTOS, R. N.; FIGUEIRA, L. D.; NAKAJIMA, C. H.; CASTRO-GAMBOA, I.; ANDRICOPULO, A. D.; BOLZANI, V. S. Development of a natural products database from the biodiversity of Brazil. *Journal of Natural Products*, v. 76, n. 3, p. 439–444, 2013. doi: 10.1021/np3006875.
- PAULI, I.; DOS SANTOS, R. N.; ROSTIROLLA, D. C.; MARTINELLI, L. K.; DUCATI, R. G.; TIMMERS, L. F.; BASSO, L. A.; SANTOS, D. S.; GUIDO, R. V.; ANDRICOPULO, A. D.; NORBERTO DE SOUZA, O. Discovery of new inhibitors of Mycobacterium tuberculosis InhA enzyme using virtual screening and a 3D-pharmacophore-based approach. *Journal of Chemical Information and Modeling*, v. 53, n. 9, p. 2390–2401, 2013. doi: 10.1021/ci400202t.
- SANTOS, R. N.; ANDRICOPULO, A. D. Physics and its interfaces with medicinal chemistry and drug design. *Brazilian Journal of Physics*, v. 43, n. 4, p. 268–280, 2013. doi: 10.1007/s13538-013-0149-7.
- FERREIRA, L. G.; SANTOS, R. N.; ANDRICOPULO, A. D. Structure-based drug design studies on a series of aldolase inhibitors. *Journal of the Brazilian Chemical Society*, v. 24, n. 2, p. 201–211, 2013. doi: 10.5935/0103-5053.20130026.
- FERREIRA, L. G.; DOS SANTOS, R. N.; OLIVA, G.; ANDRICOPULO, A. D. Molecular docking and structure-based drug design strategies. *Molecules*, v. 20, n. 7, p. 13384–13421, 2015. doi: 10.3390/molecules200713384.

- DOS SANTOS, R. N.; MORCOS, F.; JANA, B.; ANDRICOPULO, A. D.; ONUCHIC, J. N. Dimeric interactions and complex formation using direct coevolutionary couplings. *Nature: Scientific Reports*, v. 5, p. 13652, 2015. doi: 10.1038/srep13652.

## 4.2 Artigos submetidos para publicação.

- SANTOS, R. N.; FERNANDES, D. C.; ALTEI, W. F.; MAGALHÃES, L. G.; SANTOS, M. S.; RODRIGUES, M. T.; COELHO, F.; ANDRICOPULO, A. D. Discovery of a novel class of colchicine binding site agents. Submetido para *eLife* em Setembro de 2015.
- DOS SANTOS, R. N.; MORCOS, F.; JANA, B.; ANDRICOPULO, A. D.; ONUCHIC, J. N. The influence of sequence composition and quality in the performance of direct coupling analysis predictions. Submetido para *Bioinformatics* em Setembro de 2015.

# Referências

- 1 LEVINTHAL, C. *Drugs, behavior, and modern society*. New Jersey: Prentice Hall, 2012.
- 2 PARSONS, G. *History and philosophy of the life sciences*. New York: Springer, 2005.
- 3 REMERS, W. *A history of medicinal chemistry*. Ohio: Clarice Publications, 2011.
- 4 SNEADER, W. *Drug discovery: a history*. New Jersey: John Wiley & Sons, 2005.
- 5 RAVINA, E.; KUBINYI, H. *The evolution of drug discovery: from traditional medicines to modern drugs*. New Jersey: John Wiley & Sons, 2011.
- 6 LI, J.; COREY, E. *Drug discovery: practices, processes and perspectives*. New Jersey: John Wiley & Sons, 2013.
- 7 SEDDON, G. et al. *Drug design for ever, from hype to hope*. New York: Springer, 2012.
- 8 HILL, R.; RANG, H. *Drug discovery and development: technology in transition*. Amsterdam: Elsevier, 2012.
- 9 HAJDUK, P.J.; GREER, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nature Reviews Drug Discovery*, v. 6, n. 3, p. 211 – 219, 2007.
- 10 FREARSON, J. A.; COLLIE, I. T. HTS and hit finding in academia - from chemical genomics to drug discovery. *Drug Discovery Today*, v. 14, n. 23, p. 1150-1158, 2009.
- 11 GUIDO, R. V. C.; OLIVA, G.; ANDRICOPULO, A. D. Virtual screening and its integration with modern drug design technologies. *Current Medicinal Chemistry*, v. 15, n. 1, p. 37 – 46, 2008.
- 12 BUNKER, J. P. The role of medical care in contributing to health improvements within societies. *International Journal of Epidemiology*, v. 30, n. 6, p. 1260 – 1263, 2001.
- 13 RAHMAN, S. Z. et al. Lifestyle drugs: concept and impact on society. *Indian Journal of Pharmaceutical Sciences*, v. 72, n. 4, p. 409 – 413, 2010.
- 14 RILEY, J. *Rising life expectancy: a global history*. Cambridge: University Press, 2001.
- 15 OEPPEN, J.; VAUPEL, J. W. Demography: broken limits to life expectancy. *Science*, v. 296, n. 5570, p. 1029 – 1031, 2002.
- 16 WHO. *Life expectancy at birth, 1990-2013*. Disponível em: [http://www.who.int/gho/mortality\\_burden\\_disease/life\\_tables/situation\\_trends\\_text/en/](http://www.who.int/gho/mortality_burden_disease/life_tables/situation_trends_text/en/). Acesso em: 19 set. 2015.
- 17 BUNKER, J. P.; FRAZIER, H. S.; MOSTELLER, F. Improving health: measuring effects of medical care. *The Milbank Quarterly*, v. 72, n. 2, p. 225 – 258, 1994.
- 18 MULLARD, A. 2014 FDA drug approvals. *Nature Reviews Drug Discovery*, v. 14, n. 2, p. 77 – 81, 2015.

- 19 CIOCIOLA, A. A.; COHEN, L. B.; KULKARNI, P. How drugs are developed and approved by the FDA: current process and future directions. *The American Journal of Gastroenterology*, v. 109, n. 5, p. 620 – 3, 2014.
- 20 KINCH, M. S. et al. An overview of FDA-approved new molecular entities: 1827-2013. *Drug Discovery Today*, v. 9, n. 8., p. 1033 – 1039, 2014.
- 21 RAWLINS, M. D. Cutting the cost of drug development?. *Nature Reviews Drug Discovery*, v. 3, n. 4, p. 360 – 364, 2004.
- 22 KAITIN, K. I. Deconstructing the drug development process: the new face of innovation. *Clinical Pharmacology and Therapeutics*, v. 87, n. 3, p. 356 – 361, 2010.
- 23 KAITIN, K. I.; DIMASI, J. A. Pharmaceutical innovation in the 21st century: new drug approvals in the first decade, 2000-2009. *Clinical Pharmacology and Therapeutics*, v. 89, n. 2, p. 183 – 188, 2011.
- 24 ADAMS, C. P.; VAN BRANTNER, V. Market watch : estimating the cost of new drug development: is it really \$802 million?. *Health Affairs*, v. 25, n. 2, p. 420 – 428, 2006.
- 25 WENLOCK, M. C. et al. A comparison of physicochemical property profiles of development and marketed oral drugs. *Journal of Medicinal Chemistry*, v. 46, n. 7, p. 1250 – 1256, 2003.
- 26 DOAK, B. et al. Oral druggable space beyond the rule of 5: insights from drugs and clinical candidates. *Chemistry & Biology*, v. 21, n. 9, p. 1115 – 1142, 2014.
- 27 FALLER, B. et al. Evolution of the physicochemical properties of marketed drugs: Can history foretell the future? *Drug Discovery Today*, v. 16, n. 21-22, p. 976 – 984, 2011.
- 28 MERZ, K.; RINGE, D.; REYNOLDS, C. *Drug design: structure- and ligand-based approaches*. Cambridge: Cambridge University Press, 2010.
- 29 PAULI, I. et al. Discovery of new inhibitors of mycobacterium tuberculosis InhA enzyme using virtual screening and a 3D-pharmacophore-based approach. *Journal of Chemical Information and Modeling*, v. 53, n. 9, p. 2390 – 2401, 2013.
- 30 TALELE, T. T.; KHEDKAR, S. A.; RIGBY, A. C. Successful applications of computer aided drug discovery: moving drugs from concept to the clinic. *Current Topics in Medicinal Chemistry*, v. 10, n. 1, p. 127 – 141, 2010.
- 31 WLODAWER, A.; VONDRASEK, J. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annual Review of Biophysics and Biomolecular Structure*, v. 27, p. 249 – 284, 1998.
- 32 HOELDER, S.; CLARKE, P. A.; WORKMAN, P. Discovery of small molecule cancer drugs: successes, challenges and opportunities. *Molecular Oncology*, v. 6, n. 2, p. 155 – 176, 2012.
- 33 ANDERSON, A. C. The process of structure-based drug design. *Chemistry and Biology*, v. 10, n. 9, p. 787 – 797, 2003.
- 34 AGUIAR-PULIDO, V. et al. Evolutionary computation and QSAR research. *Current Computer-Aided Drug Design*, v. 9, n. 2, p. 206 – 225, 2013.



- 35 LILL, M. A. Multi-dimensional QSAR in drug discovery. *Drug Discovery Today*, v. 12, n. 23-24, p. 1013 – 1017, 2007.
- 36 SANTOS, R. N. et al. Quantitative structure-activity studies on a series of migrastatin analogs as inhibitors of cancer cell metastasis. *Medicinal Chemistry*, v. 7, n. 3, p. 155 – 164, 2011.
- 37 LEWIS, R. A.; WOOD, D. Modern 2D QSAR for drug discovery. *Wiley Interdisciplinary Reviews*, v. 4, n. 6, p. 505 – 522, 2014.
- 38 CUMMING, J. G. et al. Chemical predictive modelling to improve compound quality. *Nature Reviews Drug Discovery*, v. 12, n. 12, p. 948 – 62, 2013.
- 39 PELLEGGATTI, M. Preclinical in vivo ADME studies in drug development: a critical review. *Expert Opinion on Drug Metabolism & Toxicology*, v. 8, n. 2, p. 161 – 72, 2012.
- 40 MODA, T. L. et al. PK/DB: database for pharmacokinetic properties and predictive in silico ADME models. *Bioinformatics*, v. 24, n. 19, p. 2270 – 2271, 2008.
- 41 GLEESON, M. P.; HERSEY, A.; HANNONGBUA, S. In-silico ADME models: a general assessment of their utility in drug discovery applications. *Current Topics in Medicinal Chemistry*, v. 11, n. 4, p. 358 – 381, 2011.
- 42 PECORINO, L. *Molecular biology of cancer: mechanisms, targets and therapeutics*. Oxford: Oxford University Press, 2012.
- 43 WEINBERG, R. *The biology of cancer*. New York: Garland Science, 2014.
- 44 GIBBS, W. W. Untangling the roots of cancer. *Scientific American*, v. 289, n. 1, p. 56 – 65, 2003.
- 45 SUN, Y.; MA, L. The emerging molecular machinery and therapeutic targets of metastasis. *Trends in Pharmacological Sciences*, v. 36, n. 6, p. 349 – 359, 2015.
- 46 LEE, J. J.; LOTZE, M. T. Molecular basis of metastasis. *New England Journal of Medicine*, v. 360, n. 16, p. 1678 – 1679, 2009.
- 47 ALBERTS, B. et al. *Molecular biology of the cell*. New York: Gargland Science, 2014.
- 48 FLETCHER, O.; HOULSTON, R. S. Architecture of inherited susceptibility to common cancer. *Nature Reviews Cancer*, v. 10, n. 5, p. 353 – 361, 2010.
- 49 SETON-ROGERS, S. Tumorigenesis: mutant relationships. *Nature Reviews Cancer*, v. 15, n. 3, 2015.
- 50 HENDERSON, B. E. et al. The influence of race and ethnicity on the biology of cancer. *Nature Reviews Cancer*, v. 12, n. 9, p. 648 – 653, 2012.
- 51 KHANDEKAR, M. J.; COHEN, P.; SPIEGELMAN, B. M. Molecular mechanisms of cancer development in obesity. *Nature Reviews Cancer*, v. 11, n. 12, p. 886-895, 2011.
- 52 GRANDO, S. A. Connections of nicotine to cancer. *Nature Reviews Cancer*, v. 14, n. 6, p. 419 – 29, 2014.

- 53 WHO. *Cancer fact sheets*. Disponível em: <http://www.who.int/features/factfiles/cancer/en/>. Acesso em: 19 set. 2015.
- 54 INCA. *Estimativa 2014 - Incidência de Câncer no Brasil*. Disponível em: <http://www.inca.gov.br/estimativa/2014/>. Acesso em: 19 set. 2015.
- 55 SIEGEL, R. L.; MILLER, K. D.; JEMAL, A. Cancer statistics, 2015. *Cancer Journal for Clinicians*, v. 65, n. 1, p. 5 – 29, 2015.
- 56 NERO, T. L. et al. Oncogenic protein interfaces: small molecules, big challenges. *Nature Reviews Cancer*, v. 14, n. 4, p. 248 – 62, 2014.
- 57 GALLUZZI, L. et al. Metabolic targets for cancer therapy. *Nature Reviews Drug Discovery*, v. 12, n. 11, p. 829 – 846, 2013.
- 58 MOEN, M. D. et al. Imatinib: a review of its use in chronic myeloid leukaemia. *Drugs*, v. 67, n. 2, p. 299 – 320, 2007.
- 59 JORDAN, V. C. Tamoxifen (ICI46,474) as a targeted therapy to treat and prevent breast cancer. *British Journal of Pharmacology*, v. 147, p. S269 – S276, 2006. Supplement 1.
- 60 BROGDEN, R. N.; CHRISP, P. Flutamide. A review of its pharmacodynamic and pharmacokinetic properties, and therapeutic use in advanced prostatic cancer. *Drugs & Aging*, v. 1, n. 2, p. 104 – 115, 1991.
- 61 CRAGG, G. M. Paclitaxel (Taxol): a success story with valuable lessons for natural product drug discovery and development. *Medicinal Research Reviews*, v. 18, n. 5, p. 315 – 331, 1998.
- 62 JORDAN, M. A.; WILSON, L. Microtubules as a target for anticancer drugs. *Nature Reviews Cancer*, v. 4, n. 4, p. 253 – 265, 2004.
- 63 ROWINSKY, E. K.; DONEHOWER, R. C. Paclitaxel (Taxol). *New England Journal of Medicine* v. 332, n. 15, p. 1004 – 1014, 1995.
- 64 FOJO, A. *The role of microtubules in cell biology, neurobiology, and oncology*. New York: Humana Press, 2009.
- 65 ZHOU, X. J.; RAHMANI, R. Preclinical and clinical pharmacology of vinca alkaloids. *Drugs*, v. 44, p. 1–16, 1992. Supplement 4.
- 66 VANDECANDELAERE, A.; MARTIN, S. R.; ENGELBORGHS, Y. Response of microtubules to the addition of colchicine and tubulin-colchicine: evaluation of models for the interaction of drugs with microtubules. *Biochemical Journal*, v. 323, n. Pt 1, p. 189 – 196, 1997.
- 67 ZHOU, J.; GIANNAKAKOU, P. Targeting microtubules for cancer chemotherapy. *Current Medicinal Chemistry - Anti-cancer Agents*, v. 5, n. 1, p. 65 – 71, 2005.
- 68 HOLOHAN, C. et al. Cancer drug resistance: an evolving paradigm. *Nature Reviews Cancer*, v. 13, n. 10, p. 714 – 26, 2013.
- 69 DE DONATO, M. et al. Class III alfa-tubulin and the cytoskeletal gateway for drug resistance in ovarian cancer. *Journal of Cellular Physiology*, v. 227, n. 3, p. 1034 – 1041, 2012.

- 70 KAVALLARIS, M. Microtubules and resistance to tubulin-binding agents. *Nature Reviews Cancer*, v. 10, n. 3, p. 194 – 204, 2010.
- 71 ISBRUCKER, R. A. et al. Tubulin polymerizing activity of dictyostatin-1, a polyketide of marine sponge origin. *Biochemical Pharmacology*, v. 66, n. 1, p. 75 – 82, 2003.
- 72 MARZARO, G. et al. Discovery of biaryl aminoquinazolines as novel tubulin polymerization inhibitors. *Journal of Medicinal Chemistry*, v. 57, n. 11, p. 4598 – 4605, 2014.
- 73 MUHLRADT, P. F.; SASSE, F. Epothilone B stabilizes microtubuli of macrophages like Taxol without showing Taxol-like endotoxin activity. *Cancer Research*, v. 57, n. 16, p. 3344 – 3346, 1997.
- 74 SALUM, L. B.; DIAS, L. C.; ANDRICOPULO, A. D. Fragment-based QSAR and molecular modeling studies on a series of discodermolide analogs as microtubule-stabilizing anticancer agents. *QSAR and Combinatorial Science*, v. 28, n. 3, p. 325 – 337, 2009.
- 75 DUCKI, S. et al. Quantitative structure-activity relationship (5D-QSAR) study of combretastatin-like analogues as inhibitors of tubulin assembly. *Journal of Medicinal Chemistry*, v. 48, n. 2, p. 457 – 465, 2005.
- 76 LA REGINA, G. et al. New indole tubulin assembly inhibitors cause stable arrest of mitotic progression, enhanced stimulation of natural killer cell cytotoxic activity and repression of hedgehog-dependent cancer. *Journal of Medicinal Chemistry* v. 58, n. 15, p. 5789 – 5807, 2015.
- 77 BROCKSOM, T. J. et al. First comprehensive bakkane approach: stereoselective and efficient dichloro ketene-based total syntheses of (.+.-)- and (-)-9-acetoxylukinanolide, (.+.-)- and (+)-bakkenolide A, (-)-bakkenolides III, B, C, H, L, V, and X, (.+.-)- and (-)-homogynolide A, (.+.-)-homogynolide B, and (.+.-)-palmosalide C. *ChemInform*, v. 34, n. 19, p. 117 – 119, 2003.
- 78 COELHO, F.; DIAZ, G. Studies on the synthesis of ( $\pm$ )-pathylactone A, a nor-sesquiterpene lactone isolated from marine sources. *Tetrahedron*, v. 58, n. 9, p. 1647 – 1656, 2002.
- 79 OLIVEIRA, D. J.; COELHO, F. Diastereoselective formation of a quaternary center in a pyroglutamate derivative. formal synthesis of Monatin. *Tetrahedron Letters*, v. 42, n. 39, p. 6793 – 6796, 2001.
- 80 GOMES, J. C. et al. Aqueous Morita-Baylis-Hillman reaction of unprotected isatins with cyclic enones. *Organic Letters*, v. 15, n. 22, p. 5838 – 5841, 2013.
- 81 AMARANTE, G. W. et al. Bronsted acid catalyzed morita-baylis-hillman reaction: a new mechanistic view for thioureas revealed by ESI-MS(MS) monitoring and DFT calculations. *Chemistry*, v. 15, n. 45, p. 12460 – 12469, 2009.
- 82 CLARK, M. Validation of the general purpose Tripos 5.2 force field. *Journal of Computational Chemistry*, v. 10, p. 982 – 1012, 1989.
- 83 POWELL, M. J. D. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The Computer Journal*, v. 7, n. 2, p. 155 – 162, 1964.

- 84 GASTEIGER, J.; MARSILI, M. Iterative partial equalization of orbital electro-negativity – a rapid access to atomic charges. *Tetrahedron*, v. 36, n. 22, p. 3219 – 3228, 1980.
- 85 FERREIRA, L. G. et al. Molecular docking and structure-based drug design strategies. *Molecules*, v. 20, n. 7, 2015.
- 86 KITCHEN, D. B. et al. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery*, v. 3, n. 11, p. 935 – 949, 2004.
- 87 LEACH, A.; GILLET, V. *An introduction to chemoinformatics*. Berlin: Springer, 2007.
- 88 OPREA, T. *Chemoinformatics in drug discovery*. New Jersey: John Wiley & Sons, 2005.
- 89 ENYEDY, I. J.; EGAN, W. J. Can we use docking and scoring for hit-to-lead optimization?. *Journal of Computer-Aided Molecular Design*, v. 22, n. 3-4, p. 161 – 168, 2008.
- 90 BATOOL, M. *Protein modelling & molecular docking: modeller, autodock*. Saarbrücken: Lambert Academic Publishing, 2012.
- 91 MITEVA, M. A. et al. Fast structure-based virtual ligand screening combining FRED, DOCK, and surflex. *Journal of Medicinal Chemistry*, v. 48, n. 19, p. 6012 – 6022, 2005.
- 92 JONES, G. et al. Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology*, v. 267, n. 3, p. 727 – 748, 1997.
- 93 KELLENBERGER, E. et al. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins*, v. 57, n. 2, p. 225 – 242, 2004.
- 94 LIEBESCHUETZ, J.; COLE, J.; KORB, O. Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. *Journal of Computer-Aided Molecular Design*, v. 26, n. 6, p. 737 – 748, 2012.
- 95 JAIN, A. N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *Journal of Medicinal Chemistry*, v. 46, n. 4, p. 499 – 511, 2003.
- 96 TROTT, O.; OLSON, A. J. Software news and update AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, v. 31, n. 2, p. 455 – 461, 2010.
- 97 MORRIS, G. M. et al. Software news and updates AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, v. 30, n. 16, p. 2785 – 2791, 2009.
- 98 LIPINSKI, C. A. et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, v. 23, n. 1-3, p. 4 – 17, 2012.
- 99 VEBER, D. F. et al. Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, v. 45, n. 12, p. 2615 – 2623, 2002.
- 100 BERMAN, H. M. et al. The protein data bank. *Nucleic Acids Research*, v. 28, n. 1, p. 235 – 242, 2000.

- 101 SEELIGER, D.; DE GROOT, B. L. Ligand docking and binding site analysis with PyMOL and Autodock Vina. *Journal of Computer-Aided Molecular Design*, v. 24, n. 5, p. 417 – 422, 2010.
- 102 ALOY, P.; RUSSELL, R. B. Structural systems biology: modelling protein interactions. *Nature Reviews Molecular Cell Biology*, v. 7, n. 3, p. 188 – 197, 2006.
- 103 CRAVEUR, P. et al. Protein flexibility in the light of structural alphabets. *Frontiers in Molecular Biosciences*, v. 2, n. 20, p. eCollection, 2015.
- 104 DURRANT, J. D.; MCCAMMON, J. A. Molecular dynamics simulations and drug discovery. *BMC Biology*, v. 9, n. 1, p. 71-80, 2011.
- 105 TEAGUE, S. J. Implications of protein flexibility for drug discovery. *Nature Reviews Drug Discovery*, v. 2, n. 7, p. 527 – 541, 2003.
- 106 RAPAPORT, D. *The art of molecular dynamics simulation*. Cambridge: Cambridge University Press, 2004.
- 107 MCCAMMON, J. A.; GELIN, B. R.; KARPLUS, M. Dynamics of folded proteins. *Nature*, v. 267, n. 5612, p. 585 – 590, 1977.
- 108 CHRISTEN, M. et al. The GROMOS software for biomolecular simulation: GROMOS05. *Journal of Computational Chemistry*, v. 26, n. 16, p. 1719 – 1751, 2005.
- 109 BROOKS, B. R. et al. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, v. 4, n. 2, p. 187 – 217, 1983.
- 110 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, 1995.
- 111 WANG, J. et al. Development and testing of a general Amber force field. *Journal of Computational Chemistry*, v. 25, n. 9, p. 1157 – 1174, 2004.
- 112 PRONK, S. et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, v. 29, n. 7, p. 845 – 854, 2013.
- 113 PHILLIPS, J. C. et al. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, v. 26, n. 15, p. 1781 – 1802, 2005.
- 114 BOWERS, K. et al. Scalable algorithms for molecular dynamics simulations on commodity clusters. Disponível em: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4090217>. Acesso em: 17 set. 2015.
- 115 CASE, D. A. et al. The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, v. 26, n. 16, p. 1668 – 1688, 2005.
- 116 PALL, S. et al. Tackling exascale software challenges in molecular dynamics simulations with GROMACS. *Lecture Notes in Computer Science*, v. 8759, n. 1, p. 3-27, 2015.
- 117 AKYUZ, N. et al. Transport domain unlocking sets the uptake rate of an aspartate transporter. *Nature*, v. 518, n. 7537, p. 68 – 73, 2015.

- 118 MCNULTY, R. et al. Mechanisms of molecular transport through the urea channel of *Helicobacter pylori*. *Nature Communications*, v. 4, n. 2900, p. 1-10, 2013.
- 119 BADIEYAN, S.; BEVAN, D. R.; ZHANG, C. Probing the active site chemistry of beta-glucosidases along the hydrolysis reaction pathway. *Biochemistry*, v. 51, n. 44, p. 8907 – 8918, 2012.
- 120 CAPELLUTO, D. G. S. et al. Biophysical and molecular-dynamics studies of phosphatidic acid binding by the Dvl-2 DEP domain. *Biophysical Journal*, v. 106, n. 5, p. 1101 – 1111, 2014.
- 121 MACKERELL, A. D.; NILSSON, L. Molecular dynamics simulations of nucleic acid-protein complexes. *Current Opinion in Structural Biology*, v. 18, n. 2, p. 194 – 199, 2008.
- 122 BERMAN, H. M. The protein data bank. *Nucleic Acids Research*, v. 28, n. 1, p. 235 – 242, 2000.
- 123 PROTA, A. E. et al. The novel microtubule-destabilizing drug BAL27862 binds to the colchicine site of tubulin with distinct effects on microtubule organization. *Journal of Molecular Biology*, v. 426, n. 8, p. 1848 – 1860, 2014.
- 124 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, 2010.
- 125 LEMKUL, J. A.; BEVAN, D. R. The role of molecular simulations in the development of inhibitors of amyloid beta-peptide aggregation for the treatment of Alzheimer's disease. *ACS Chemical Neurosciences*, v. 3, n. 11, p. 845 – 856, 2012.
- 126 SCHUTTELKOPF, A. W.; VAN AALTEN, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D: biological crystallography*, v. 60, n. 8, p. 1355 – 1363, 2004.
- 127 HINCHLIFFE, A. *Molecular modelling for beginners*. New Jersey: John Wiley & Sons, 2008.
- 128 LODISH, H. *Molecular cell biology*. New York: W. H. Freeman, 2008.
- 129 DESAI, A.; MITCHISON, T. J. Microtubule polymerization dynamics. *Annual Review of Cell and Developmental Biology*, v. 13, n. 1, p. 83 – 117, 1997.
- 130 MITCHISON, T.; KIRSCHNER, M. Dynamic instability of microtubule growth. *Nature*, v. 312, n. 5991, p. 237 – 242, 1984.
- 131 CYTOSKELETON *Tubulin polymerization assay using 99% pure tubulin, fluorescence based*. Disponível em: <http://www.cytoskeleton.com/bk011p/>. Acesso em: 19 set. 2015.
- 132 BONNE, D. et al. 4',6-diamidino-2-phenylindole, a fluorescent probe for tubulin and microtubules. *Journal of Biological Chemistry*, v. 260, n. 5, p. 2819 – 2825, 1985.
- 133 DYRAGER, C. et al. Inhibitors and promoters of tubulin polymerization: synthesis and biological evaluation of chalcones and related dienones as potential anticancer agents. *Bioorganic & Medicinal Chemistry*, v. 19, n. 8, p. 2659 – 2665, 2011.

- 134 COPELAND, R. *Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists*. New Jersey: John Wiley & Sons, 2013.
- 135 RUAN, B. et al. Synthesis, biological evaluation and molecular docking studies of resveratrol derivatives possessing curcumin moiety as potent antitubulin agents. *Bioorganic & Medicinal Chemistry*, v. 20, n. 2, p. 1113 – 1121, 2012.
- 136 GASKIN, F.; CANTOR, C. R.; SHELANSKI, M. L. Turbidimetric studies of the in vitro assembly and disassembly of porcine neurotubules. *Journal of Molecular Biology*, v. 89, n. 4, p. 737 – 755, 1974.
- 137 ZHOU, J. *Microtubule protocols*. New York: Humana Press, 2007.
- 138 MULLIGAN, J. M. et al. Identification of tubulin as the molecular target of proapoptotic pyrrolo-1,5-benzoxazepines. *Molecular Pharmacology*, v. 70, n. 1, p. 60 – 70, 2006.
- 139 JIANG, J. D. et al. 3-(Iodoacetamido)-benzoylurea: a novel cancericidal tubulin ligand that inhibits microtubule polymerization, phosphorylates bcl-2, and induces apoptosis in tumor cells. *Cancer Research*, v. 58, n. 23, p. 5389 – 5395, 1998.
- 140 GURURAJA, T. L. et al. R-253 disrupts microtubule networks in multiple tumor cell lines. *Clinical Cancer Research*, v. 12, n. 12, p. 3831 – 3842, 2006.
- 141 EVANGELIO, J. A. et al. Fluorescent taxoids as probes of the microtubule cytoskeleton. *Cell Motility and the Cytoskeleton*, v. 39, n. 1, p. 73 – 90, 1998.
- 142 CLARK, J. I.; GARLAND, D. Fluorescein colchicine: synthesis, purification, and biological activity. *Journal of Cell Biology*, v. 76, n. 3, p. 619 – 627, 1978.
- 143 ERTL, P.; ROHDE, B.; SELZER, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *Journal of Medicinal Chemistry*, v. 43, n. 20, p. 3714 – 3717, 2000.
- 144 NAKAGAWA-GOTO, K. et al. Antitumor agents. part 236: synthesis of water-soluble colchicine derivatives. *Bioorganic & Medicinal Chemistry Letters*, v. 15, n. 1, p. 235 – 238, 2005.
- 145 BARRON, D. M. et al. A fluorescence-based high-throughput assay for antimicrotubule drugs. *Analytical Biochemistry*, v. 315, n. 1, p. 49 – 56, 2003.
- 146 BEESLEY, R. M.; INGOLD, C. K.; THORPE, J. F. The formation and stability of spiro-compounds. Part I. spiro-compounds from cyclohexane. *Journal of Chemical Society, Transactions*, v. 107, n. 1, p. 1080 – 1106, 1915.
- 147 SWENTON, J. S.; BRADIN, D.; GATES, B. D. Spiro-fused 2,5-cyclohexadienones from thermal 1,3-shifts in quinol vinyl ethers. reactions in nonbenzenoid systems and limitations of the chemistry. *ChemInform*, v. 23, n. 12, p. 1, 1992.
- 148 HAAK, R. A.; BECK, K. R. Synthesis of substituted spiro[4.5]deca-3,6,9-triene-2,8-diones: an expeditious route to the spiro[4.5]decane terpene skeleton. *Tetrahedron Letters*, v. 30, n. 13, p. 1605 – 1608, 1988.
- 149 MOLINSPIRATION *Molinspiration property explorer*. Disponível em: <http://www.molinspiration.com/>. Acesso em: 19 set. 2015.

- 150 SANDER, T. et al. OSIRIS, an entirely in-house developed drug discovery informatics system. *Journal of Chemical Information and Modeling*, v. 49, n. 2, p. 232 – 246, 2009.
- 151 BERLIER, J. E. et al. Quantitative comparison of long-wavelength Alexa Fluor dyes to Cy dyes: fluorescence of the dyes and their bioconjugates. *The Journal of Histochemistry and Cytochemistry*, v. 51, p. 1699 – 1712, 2003.
- 152 CORY, A. H. et al. Use of an aqueous soluble tetrazolium/formazan assay for cell growth assays in culture. *Cancer Communications*, v. 3, p. 207 – 212, 1991.
- 153 FERREIRA, L. G.; SANTOS, R. N.; ANDRICOPULO, A. D. Structure-based drug design studies on a series of aldolase inhibitors. *Journal of the Brazilian Chemical Society*, v. 24, n. 2, p. 201 – 211, 2013.
- 154 KESKIN, O. et al. Relating molecular flexibility to function: a case study of tubulin. *Biophysical Journal*, v. 83, n. 2, p. 663 – 680, 2002.
- 155 BIASINI, M. et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*, v. 42, n. 1, W252-258, 2014.
- 156 DORLEANS, A. et al. Variations in the colchicine-binding domain provide insight into the structural switch of tubulin. *Proceedings of the National Academy of Sciences of the United States of America*, v. 106, p. 13775 – 13779, 2009.
- 157 PICCOLINO, M. Biological machines: from mills to molecules. *Nature Reviews Molecular Cell Biology*, v. 1, n. 2, p. 149 – 152, 2000.
- 158 ALBERTS, B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell*, v. 92, n. 3, p. 291-294, 1998.
- 159 HAKES, L. et al. Protein-protein interaction networks and biology—what’s the connection? *Nature Biotechnology*, v. 26, n. 1, p. 69 – 72, 2008.
- 160 ANDRAS, P.; ANDRAS, C. The origins of life - the 'protein interaction world' hypothesis: protein interactions were the first form of self-reproducing life and nucleic acids evolved later as memory molecules. *Medical Hypotheses*, v. 64, n. 4, p. 678 – 688, 2005.
- 161 BONETTA, L. Protein-protein interactions: interactome under construction. *Nature*, v. 468, n. 7325, p. 851 – 854, 2010.
- 162 EWING, R. M. et al. Large-scale mapping of human protein-protein interactions by mass spectrometry. *Molecular Systems Biology*, v. 3, n. 89, p. 1-17, 2007.
- 163 HOLMES, K. C. et al. Atomic model of the actin filament. *Nature*, v. 347, n. 6288, p. 44 – 49, 1990.
- 164 DOMINGUEZ, R.; HOLMES, K. C. Actin structure and function. *Annual Review of Biophysics*, v. 40, p. 169 – 186, 2011.
- 165 REISLER, E. Actin molecular structure and function. *Current Opinion in Cell Biology*, v. 5, n. 1, p. 41 – 47, 1993.
- 166 CAUDRON, N. et al. Microtubule nucleation from stable tubulin oligomers. *Journal of Biological Chemistry*, v. 277, n. 52, p. 50973 – 50979, 2002.



- 167 COOPER, J. R.; WORDEMAN, L. The diffusive interaction of microtubule binding proteins. *Current Opinion in Cell Biology*, v. 21, n. 1, p. 68 – 73, 2009.
- 168 BERMUDES, D.; HINKLE, G.; MARGULIS, L. Do prokaryotes contain microtubules?. *Microbiological Reviews*, v. 58, n. 3, p. 387 – 400, 1994.
- 169 BIENIOSSEK, C. et al. The molecular architecture of the metalloprotease FtsH. *Proceedings of the National Academy of Sciences of the United States of America*, v. 103, n. 9, p. 3066 – 3071, 2006.
- 170 LANGKLOTZ, S.; BAUMANN, U.; NARBERHAUS, F. Structure and function of the bacterial AAA protease FtsH. *Biochimica et Biophysica Acta*, v. 1823, n. 1, p. 40 – 48, 2012.
- 171 LEE, K. A. Dimeric transcription factor families: it takes two to tango but who decides on partners and the venue? *Journal of Cell Science*, v. 103, n. 1, p. 9 – 14, 1992.
- 172 KLEMM, J. D.; SCHREIBER, S. L.; CRABTREE, G. R. Dimerization as a regulatory mechanism in signal transduction. *Annual Review of Immunology*, v. 16, p. 569 – 592, 1998.
- 173 ALI, M. H.; IMPERIALI, B. Protein oligomerization: how and why. *Bioorganic and Medicinal Chemistry*, v. 13, n. 17, p. 5013 – 5020, 2005.
- 174 MARIANAYAGAM, N. J.; SUNDE, M.; MATTHEWS, J. M. The power of two: protein dimerization in biology. *Trends in Biochemical Sciences*, v. 29, n. 11, p. 618 – 625, 2004.
- 175 MATTHEWS, J. *Protein dimerization and oligomerization in biology*. New York: Springer, 2012.
- 176 ISPOLATOV, I. et al. Binding properties and evolution of homodimers in protein-protein interaction networks. *Nucleic Acids Research*, v. 33, n. 11, p. 3629 – 3635, 2005.
- 177 MORCOS, F. et al. *Protein structure prediction: methods in molecular biology*. New York: Humana Press, 2014.
- 178 EKEBERG, M. et al. Improved contact prediction in proteins: using pseudolikelihoods to infer Potts models. *Physical Review E - statistical*, v. 87, n. 1, p. 1 – 19, 2013.
- 179 KAMISSETTY, H.; OVCHINNIKOV, S.; BAKER, D. Assessing the utility of coevolution-based residue-residue contact predictions in a sequence- and structure-rich era. *Proceedings of the National Academy of Sciences of the United States of America*, v. 110, n. 39, p. 15674 – 15679, 2013.
- 180 SULKOWSKA, J. I. et al. Genomics-aided structure prediction. *Proceedings of the National Academy of Sciences of the United States of America*, v. 109, n. 26, p. 10340 – 10345, 2012.
- 181 MARKS, D. S. et al. Protein 3D structure computed from evolutionary sequence variation. *PLoS ONE*, v. 6, n. 12, p. e28766, 2011.
- 182 HOPF, T. A. et al. Three-dimensional structures of membrane proteins from genomic sequencing. *Cell*, v. 149, n. 7, p. 1607 – 1621, 2012.

- 183 MORCOS, F. et al. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proceedings of the National Academy of Sciences of the United States of America*, v. 108, n. 49, p. E1293 – E1301, 2011.
- 184 MORCOS, F. et al. Coevolutionary signals across protein lineages help capture multiple protein conformations. *Proceedings of the National Academy of Sciences of the United States of America*, v. 110, p. 20533 – 20538, 2013.
- 185 WEIGT, M. et al. Identification of direct residue contacts in protein-protein interaction by message passing. *Proceedings of the National Academy of Sciences of the United States of America*, v. 106, n. 1, p. 67 – 72, 2009.
- 186 SCHUG, A. et al. High-resolution protein complexes from integrating genomic information with molecular simulation. *Proceedings of the National Academy of Sciences of the United States of America*, v. 106, n. 52, p. 22124 – 22129, 2009.
- 187 PROCACCINI, A. et al. Dissecting the specificity of protein-protein interaction in bacterial two-component signaling: orphans and crosstalks. *PLoS ONE*, v. 6, n. 5, p. 19729, 2011.
- 188 CHENG, R. R. et al. Toward rationally redesigning bacterial two-component signaling systems using coevolutionary information. *Proceedings of the National Academy of Sciences of the United States of America*, v. 111, n. 5, p. E563 – 71, 2014.
- 189 TAMIR, S. et al. Integrated strategy reveals the protein interface between cancer targets Bcl-2 and NAF-1. *Proceedings of the National Academy of Sciences of the United States of America*, v. 111, n. 14, p. 5177 – 82, 2014.
- 190 JANA, B.; MORCOS, F.; ONUCHIC, J. N. From structure to function: the convergence of structure based models and co-evolutionary information. *Physical Chemistry Chemical Physics*, v. 16, n. 14, p. 6496 – 6507, 2014.
- 191 OVCHINNIKOV, S.; KAMISSETTY, H.; BAKER, D. Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife*, v. 2014, n. 3, p. 1 – 21, 2014.
- 192 PIERCE, B.; TONG, W.; WENG, Z. M-ZDOCK: a grid-based approach for Cn symmetric multimer docking. *Bioinformatics*, v. 21, n. 8, p. 1472 – 1478, 2005.
- 193 VRIES, S. J. de; DIJK, M. V.; BONVIN, A. M. J. J. The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols*, v. 5, n. 5, p. 883 – 897, 2010.
- 194 MUKHERJEE, S.; ZHANG, Y. Protein-protein complex structure predictions by multimeric threading and template recombination. *Structure*, v. 19, n. 7, p. 955 – 966, 2011.
- 195 ZHENG, W. et al. Predictive energy landscapes for protein-protein association. *Proceedings of the National Academy of Sciences of the United States of America*, v. 109, n. 47, p. 19244 – 19249, 2012.
- 196 ONUCHIC, J. N.; WOLYNES, P. G. Theory of protein folding. *Current Opinion in Structural Biology*, v. 14, n. 1, p. 70 – 75, 2004.

- 197 WHITFORD, P. C.; SANBONMATSU, K. Y.; ONUCHIC, J. N. Biomolecular dynamics: order–disorder transitions and energy landscapes. *Reports on Progress in Physics*, v. 75, n. 7, p.76601, 2012.
- 198 NOEL, J. K. et al. SMOG@ctbp: Simplified deployment of structure-based models in GROMACS. *Nucleic Acids Research*, v. 38, p. W657-W661, 2010. Supplement 2.
- 199 FINN, R. D. et al. Pfam: the protein families database. *Nucleic Acids Research*, v. 42, n. 1, p. D222–D230, 2014.
- 200 FINN, R. D.; CLEMENTS, J.; EDDY, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research*, v. 39, p. W29-W37, 2011.
- 201 FRACZKIEWICZ, R.; BRAUN, W. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *Journal of Computational Chemistry*, v. 19, n. 3, p. 319 – 333, 1998.
- 202 DELANO, W. L. The PyMOL molecular graphics system, version 1.7.4. *Schrödinger LLC*, 2002. Disponível em: <www.pymol.org>. Acesso em: 17 set. 2015.
- 203 BORDOLI, L.; SCHWEDE, T. Automated protein structure modeling with swiss-model workspace and the protein model portal. *Methods in Molecular Biology*, v. 857, n. 1, p. 107 – 136, 2012.
- 204 LAMMERT, H.; SCHUG, A.; ONUCHIC, J. N. Robustness and generalization of structure-based models for protein folding and function. *Proteins*, v. 77, n. 4, p. 881 – 891, 2009.
- 205 LAUB, M. T.; GOULIAN, M. Specificity in two-component signal transduction pathways. *Annual Review of Genetics*, v. 41, n. 1, p. 121 – 145, 2007.
- 206 HOCH, J. A. Two-component and phosphorelay signal transduction. *Current Opinion in Microbiology*, v. 3, n. 2, p. 165 – 170, 2000.
- 207 FABRET, C.; FEHER, V. A.; HOCH, J. A. Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. *Journal of Bacteriology*, v. 181, n. 7, p. 1975 – 1983, 1999.
- 208 BACHHAWAT, P. et al. Mechanism of activation for transcription factor PhoB suggested by different modes of dimerization in the inactive and active states. *Structure*, v. 13, n. 9, p. 1353 – 1363, 2005.
- 209 KING-SCOTT, J. et al. The structure of a full-length response regulator from *Mycobacterium tuberculosis* in a stabilized three-dimensional domain-swapped, activated state. *Journal of Biological Chemistry*, v. 282, n. 52, p. 37717 – 37729, 2007.
- 210 BAUMKOTTER, F. et al. Amyloid precursor protein dimerization and synaptogenic function depend on copper binding to the growth factor-like domain. *Journal of Neuroscience*, v. 34, n. 33, p. 11159 – 11172, 2014.
- 211 BAULAC, S. et al. Dimerization of Parkinson’s disease-causing DJ-1 and formation of high molecular weight complexes in human brain. *Molecular and Cellular Neuroscience*, v. 27, n. 3, p. 236 – 246, 2004.

- 212 TOMPA, P. et al. The role of dimerization in prion replication. *Biophysical Journal*, v. 82, n. 4, p. 1711 – 1718, 2002.
- 213 ZHENG, W.; SCHAFFER, N. P.; WOLYNES, P. G. Free energy landscapes for initiation and branching of protein aggregation. *Proceedings of the National Academy of Sciences of the United States of America*, v. 110, n. 51, p. 20515 – 20, 2013.
- 214 KOEHN, F. E.; CARTER, G. T. The evolving role of natural products in drug discovery. *Nature Reviews Drug Discovery*, v. 4, n. 3, p. 206 – 220, 2005.
- 215 NEWMAN, D. J. Natural products as leads to potential drugs: an old process or the new hope for drug discovery?. *Journal of Medicinal Chemistry*, v. 51, n. 9, p. 2589 – 2599, 2008.
- 216 CRAGG, G. M.; NEWMAN, D. J. Natural products: a continuing source of novel drug leads. *Biochimica et Biophysica Acta*, v. 1830, n. 6, p. 3670 – 3695, 2013.
- 217 WERMUTH, C. *The practice of medicinal chemistry*. London: Academic Press, 2011.
- 218 NEWMAN, D. J.; CRAGG, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *Journal of Natural Products*, v. 75, n. 3, p. 311 – 335, 2012.
- 219 LI, J. W. H.; VEDERAS, J. C. Drug discovery and natural products: end of era or an endless frontier?. *Biomeditsinskaya Khimiya*, v. 57, n. 2, p. 148 – 160, 2011.
- 220 DE LUCA, V. et al. Mining the biodiversity of plants: a revolution in the making. *Science*, v. 336, n. 6089, p. 1658 – 1661, 2012.
- 221 HAZEKAMP, A. et al. *Comprehensive natural products II*. Amsterdam: Elsevier, 2010.
- 222 JOLY, C. A. et al. Ecology, biodiversity conservation research, training, and policy in São Paulo. *Science*, v. 328, n. 5984, p. 1358 – 1359, 2010.
- 223 VIEIRA, G. M. et al. Cytotoxic clerodane diterpenoids from *Casearia obliqua*. *Journal of Natural Products*, v. 72, n. 10, p. 1847 – 1850, 2009.
- 224 VIEIRA-JÚNIOR, G. M. et al. Cytotoxic clerodane diterpenes from *Casearia rupestris*. *Journal of Natural Products*, v. 74, n. 4, p. 776 – 781, 2011.
- 225 VIEGAS, C. et al. Further bioactive piperidine alkaloids from the flowers and green fruits of *Cassia spectabilis*. *Journal of Natural Products*, v. 67, n. 5, p. 908 – 910, 2004.
- 226 VIEGAS, C. et al. New selective acetylcholinesterase inhibitors designed from natural piperidine alkaloids. *Bioorganic and Medicinal Chemistry*, v. 13, n. 13, p. 4184 – 4190, 2005.
- 227 VALLI, M. et al. Anticholinesterasic, nematostatic and anthelmintic activities of pyridinic and pyrazinic compounds. *Current Medicinal Chemistry*, v. 18, n. 22, p. 3423 – 3430, 2011.
- 228 FLAUSINO, O. et al. Anxiolytic effects of erythrinian alkaloids from *Erythrina mulungu*. *Journal of Natural Products*, v. 70, n. 1, p. 48 – 53, 2007.
- 229 O'BOYLE, N. M.; BANCK, M.; JAMES, C. A.; MORLEY, C.; VANDERMEERSCH, T.; HUTCHISON, G. R. Openbabel: an open chemical toolbox. *Journal of Cheminformatics*, v. 3, n. 33, p. 1 – 14, 2011.

- 230 IRWIN, J. J. et al. ZINC: a free tool to discover chemistry for biology. *Journal of Chemoinformatics and Modeling*, v. 52, n. 7, p. 1757 – 1768, 2012.
- 231 AKHOON, B. A. et al. Virtual screening of specific chemical compounds by exploring E.coli NAD<sup>+</sup>-dependent DNA ligase as a target for antibacterial drug discovery. *Journal of Molecular Modeling*, v. 17, n. 2, p. 265 – 273, 2011.
- 232 PIERONI, M. et al. NOC chemistry for tuberculosis—further investigations on the structure–activity relationships of antitubercular isoxazole-3-carboxylic acid ester derivatives. *ChemMedChem* v. 5, n. 10, p. 1667 – 1672, 2010.
- 233 HSIN, K. Y. et al. EDULISS: A small-molecule database with data-mining and pharmacophore searching capabilities. *Nucleic Acids Research*, v. 39, p. D1042-D1048, 2011. Supplement 1.
- 234 STEINBECK, C. et al. The chemistry development kit (CDK): an open-source Java library for chemo- and bioinformatics. *Journal of Chemical Information and Computer Sciences* v. 43, n. 2, p. 493 – 500, 2003.
- 235 LIPINSKI, C. A. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies*, v. 1, n. 4, p. 337 – 341, 2004.
- 236 LIPINSKI, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *Journal of Pharmacological and Toxicological Methods*, v. 44, n. 1, p. 235 – 249, 2000.
- 237 CLARK, D. E.; PICKETT, S. D. Computational methods for the prediction of 'drug-likeness'. *Drug Discovery Today*, v. 5, n. 2, p. 49 – 58, 2000.



## APÊNDICE – Recursos computacionais

Neste apêndice são apresentados os Programas utilizados ou desenvolvidos durante a execução deste trabalho.

Tabela 13 – Ferramentas de Desenvolvimento.

<b>Nome</b>	<b>Versão</b>	<b>Descrição</b>
AWK	4.1.2	Linguagem de programação implementada em GNU (GAWK). <a href="https://www.gnu.org/software/gawk/manual/gawk.html">https://www.gnu.org/software/gawk/manual/gawk.html</a>
Python	2.7	Linguagem de programação interpretada e de alto nível <a href="http://www.python.org/">http://www.python.org/</a>

Fonte: Elaborada pelo autor

Tabela 14 – Ferramentas de Elaboração de Texto.

<b>Nome</b>	<b>Versão</b>	<b>Descrição</b>
Latex	1.40.16	Software gerador de documentos para publicação. <a href="http://www.latex-project.org/">http://www.latex-project.org/</a>
TexShop	3.51	Editor de documentos Latex. <a href="http://www.uoregon.edu/~koch/texshop/">http://www.uoregon.edu/~koch/texshop/</a>

Fonte: Elaborada pelo autor

Tabela 15 – Ambientes Operacionais.

<b>Nome</b>	<b>Versão</b>	<b>Descrição</b>
Mac OS X	10.10	Sistema Operacional UNIX baseado em Darwin. <a href="http://www.apple.com/macosx/">http://www.apple.com/macosx/</a>
Linux Ubuntu	12.04 e 14.10	Sistema Operacional do tipo Unity baseado na distribuição Debian. <a href="http://www.ubuntu.com">http://www.ubuntu.com</a>
Linux Red Hat	9	Sistema Operacional do tipo Gnome 3. <a href="http://www.redhat.com/en/technologies/linux-platforms/enterprise-linux">http://www.redhat.com/en/technologies/linux-platforms/enterprise-linux</a>

Fonte: Elaborada pelo autor

Tabela 16 – Programas de Análise de Sistemas Moleculares.

<b>Nome</b>	<b>Versão</b>	<b>Descrição</b>
Chimera UCSF	1.10	Programa de visualização de estruturas moleculares. <a href="https://www.cgl.ucsf.edu/chimera/">https://www.cgl.ucsf.edu/chimera/</a>
Pymol	1.7r	Programa de visualização de estruturas moleculares. <a href="http://sourceforge.net/projects/pymol/?source=directory">http://sourceforge.net/projects/pymol/?source=directory</a>
VMD	1.9	Programa de visualização de trajetórias de simulações moleculares. <a href="http://www.ks.uiuc.edu/Research/vmd/">http://www.ks.uiuc.edu/Research/vmd/</a>
SYBYL	X2.0	Interface de visualização e edição de estruturas moleculares. <a href="http://www.ks.uiuc.edu/Research/vmd/">http://www.ks.uiuc.edu/Research/vmd/</a>

Fonte: Elaborada pelo autor

Tabela 17 – Programas de Docagem e Dinâmica Molecular.

<b>Nome</b>	<b>Versão</b>	<b>Descrição</b>
AutoDock Vina	1.1.2	Programa de docagem molecular baseado em algoritmo genético com evolução Lamarkiana. <a href="http://vina.scripps.edu">http://vina.scripps.edu</a>
GOLD Suite	5.3	Programa comercial de docagem molecular baseado em algoritmo genético. <a href="http://www.ccdc.cam.ac.uk/Solutions/GoldSuite/Pages/GOLD.aspx">http://www.ccdc.cam.ac.uk/Solutions/GoldSuite/Pages/GOLD.aspx</a>
SurflexDock	1.0	Programa comercial de docagem molecular baseado em construção incremental. <a href="http://www.certara.com/products/molmod/sybyl-x/sbd/">http://www.certara.com/products/molmod/sybyl-x/sbd/</a>
GROMACS	4.5	Programa de dinâmica molecular baseado em mecânica clássica newtoniana. <a href="http://www.gromacs.org">http://www.gromacs.org</a>
SMOG	1.0	Servidor de topologias e parametrizações para simulações de dinâmica molecular utilizando modelos baseados em estrutura (SBMs). <a href="http://smog-server.org">http://smog-server.org</a>

Fonte: Elaborada pelo autor



Tabela 18 – Programas de Desenho Bidimensional e Análise de Propriedades Físico Químicas de Pequenas Moléculas.

<b>Nome</b>	<b>Versão</b>	<b>Descrição</b>
Marvin Sketch	15.7.6	Programa de visualização e desenho bidimensional de estruturas moleculares. <a href="https://www.chemaxon.com/products/marvin/marvinsketch/">https://www.chemaxon.com/products/marvin/marvinsketch/</a>
Molinspiration	1.0	Programa de cálculo de propriedades físico-químicas de pequenas moléculas. <a href="http://www.molinspiration.com">http://www.molinspiration.com</a>
OSIRIS	2	Programa de cálculo de propriedades físico-químicas de pequenas moléculas. <a href="http://www.organic-chemistry.org/prog/peo/">http://www.organic-chemistry.org/prog/peo/</a>

Fonte: Elaborada pelo autor

Tabela 19 – Programas Desenvolvidos durante o Projeto.

<b>Nome</b>	<b>Versão</b>	<b>Descrição</b>
NuBBE Database	1	Base virtual de produtos naturais extraídos do bioma brasileiro. <a href="http://nubbe.iq.unesp.br/portal/nubbedb.html">http://nubbe.iq.unesp.br/portal/nubbedb.html</a>

Fonte: Elaborada pelo autor

Tabela 20 – Programas Utilizados para Elaborar e Analisar os Gráficos Apresentados neste Trabalho.

<b>Nome</b>	<b>Versão</b>	<b>Descrição</b>
gnuplot	4.2	Programa baseado em linha de comandos para elaboração de imagens gráficas. <a href="http://www.gnuplot.info">http://www.gnuplot.info</a>
QtiPlot	0.9.9-rc	Programa interativo para análise e visualização de dados científicos. <a href="http://www.gnuplot.info">http://www.gnuplot.info</a>

Fonte: Elaborada pelo autor



# Anexos



ANEXO A – Estruturas Moleculares dos  
Compostos Avaliados na Identificação de  
Novos Moduladores de Polimerização de  
Microtúbulos.



### aciloínas

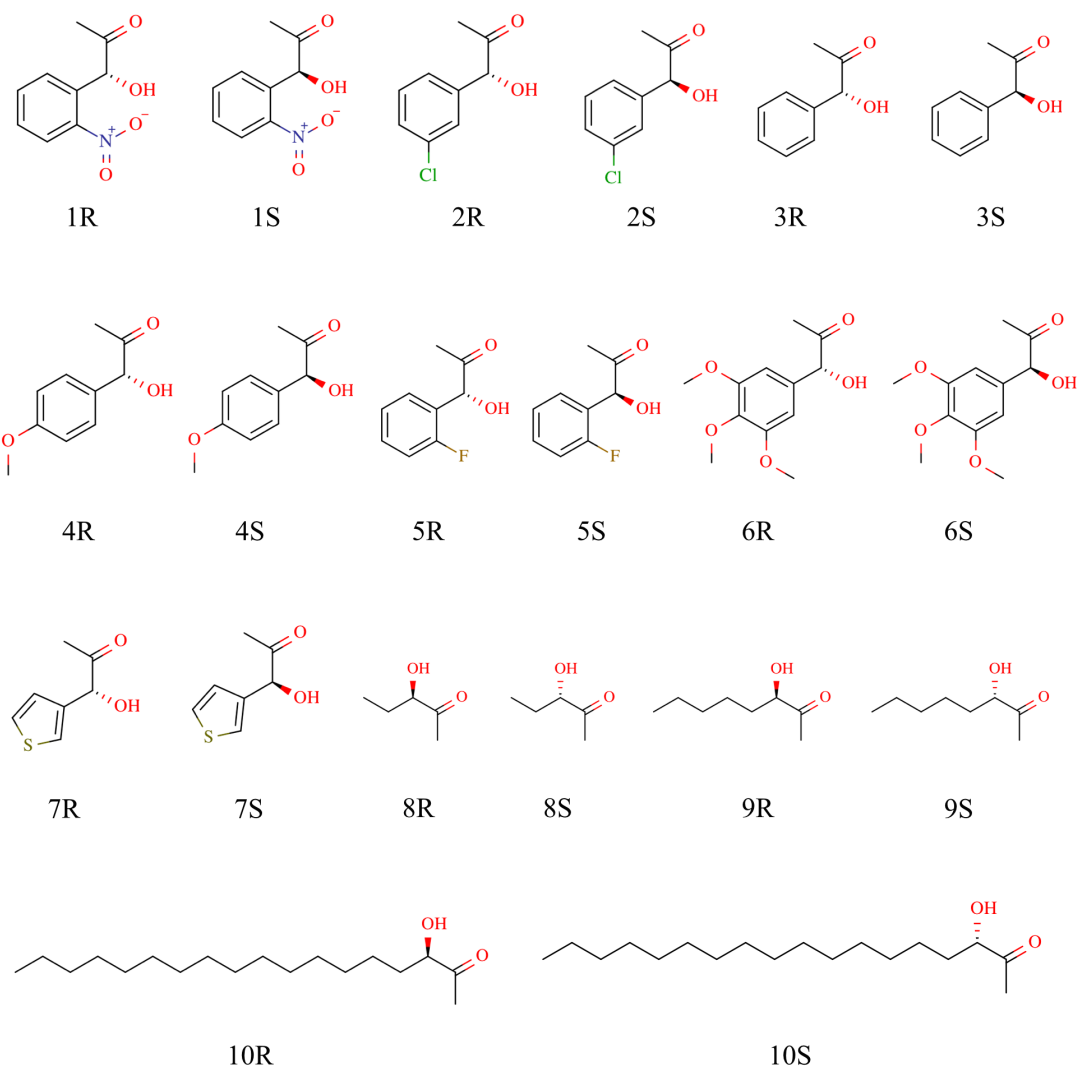


Figura 59 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: aciloínas.

Fonte: Elaborada pelo autor.

## adutos de Baylis-Hilman

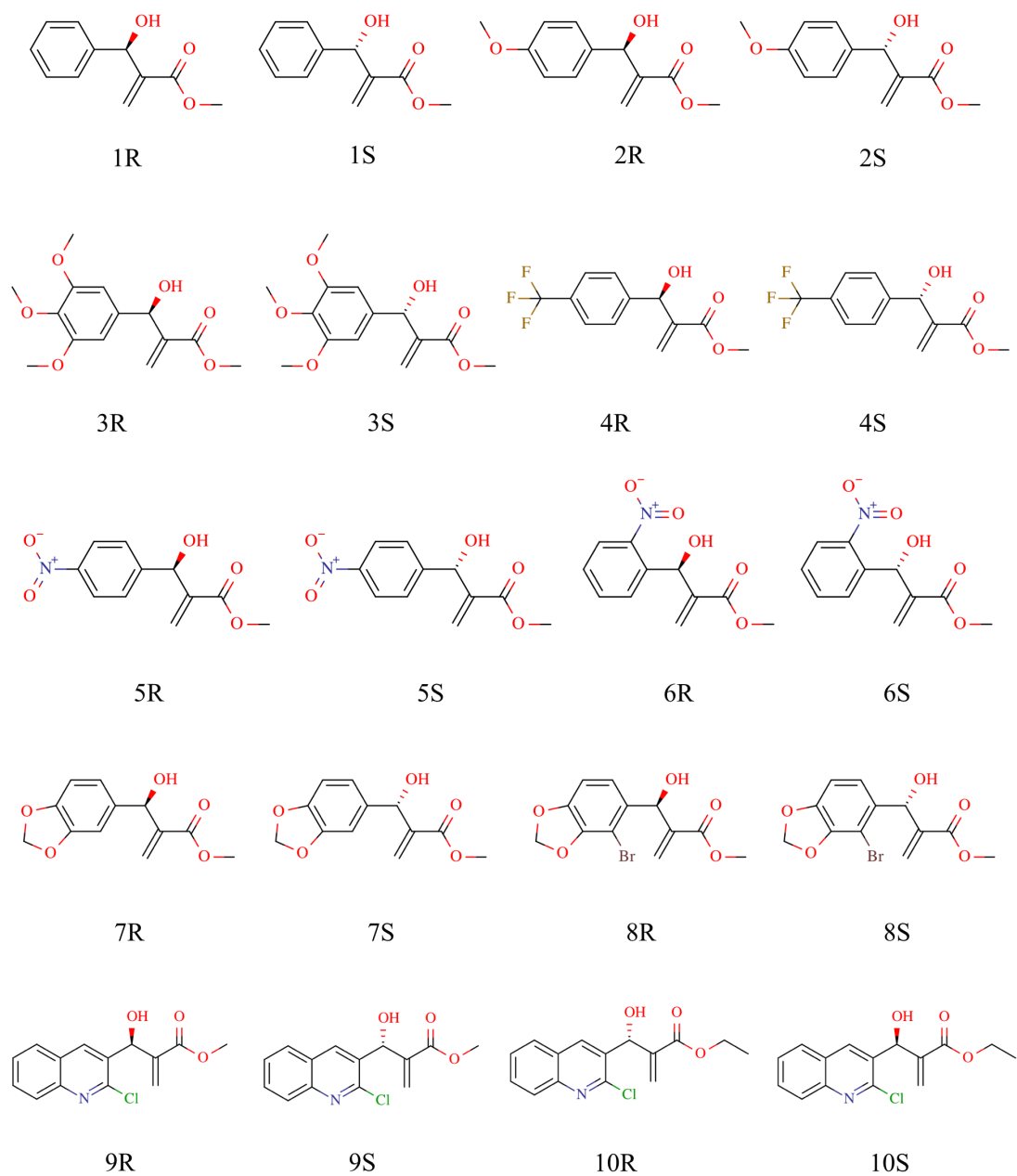


Figura 60 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: adutos de Morita-Baylis-Hilman - parte 1.

Fonte: Elaborada pelo autor.



## adutos de Baylis-Hilman

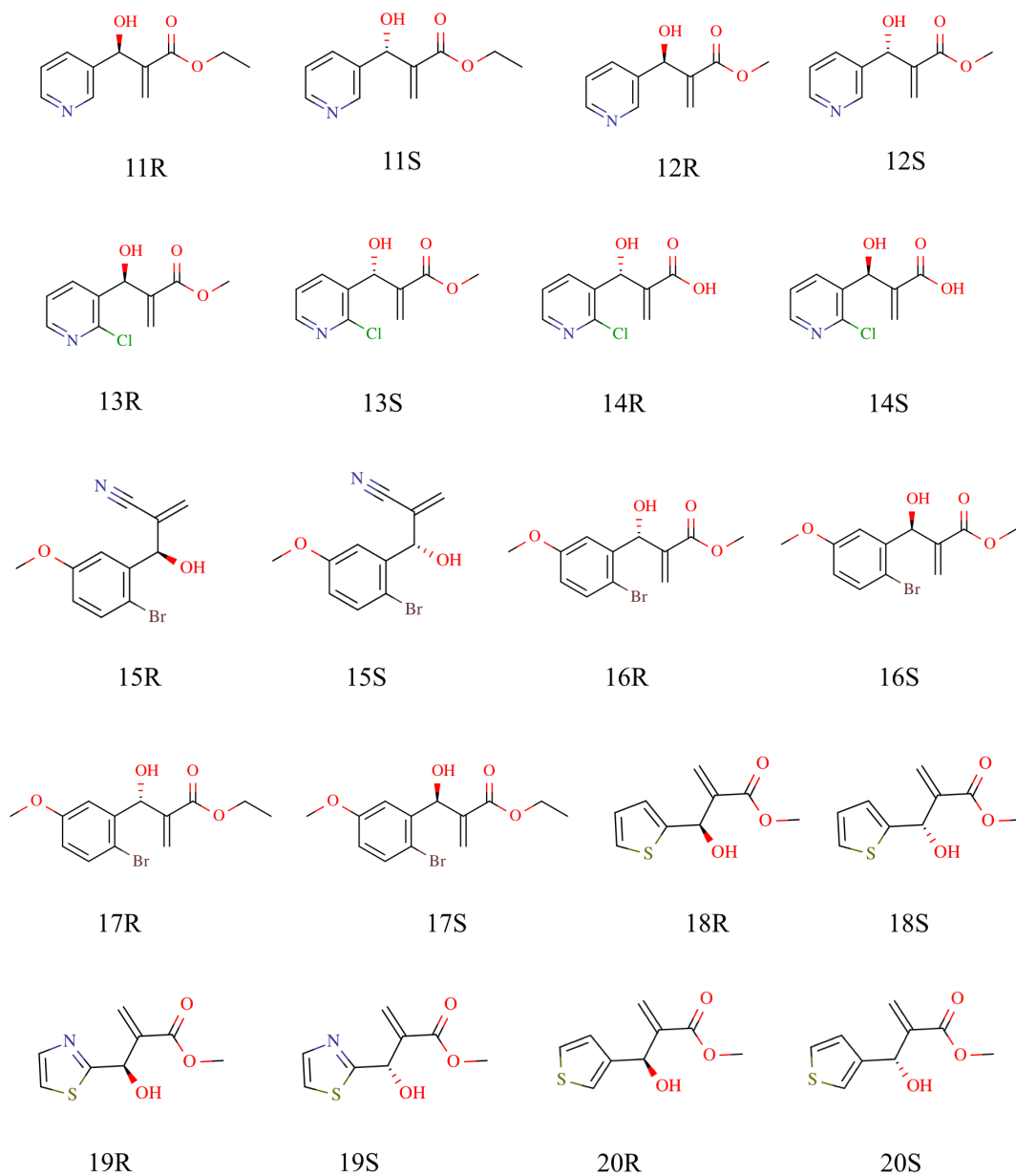


Figura 61 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: adutos de Morita-Baylis-Hilman - parte 2.

Fonte: Elaborada pelo autor.

## adutos de Baylis-Hilman

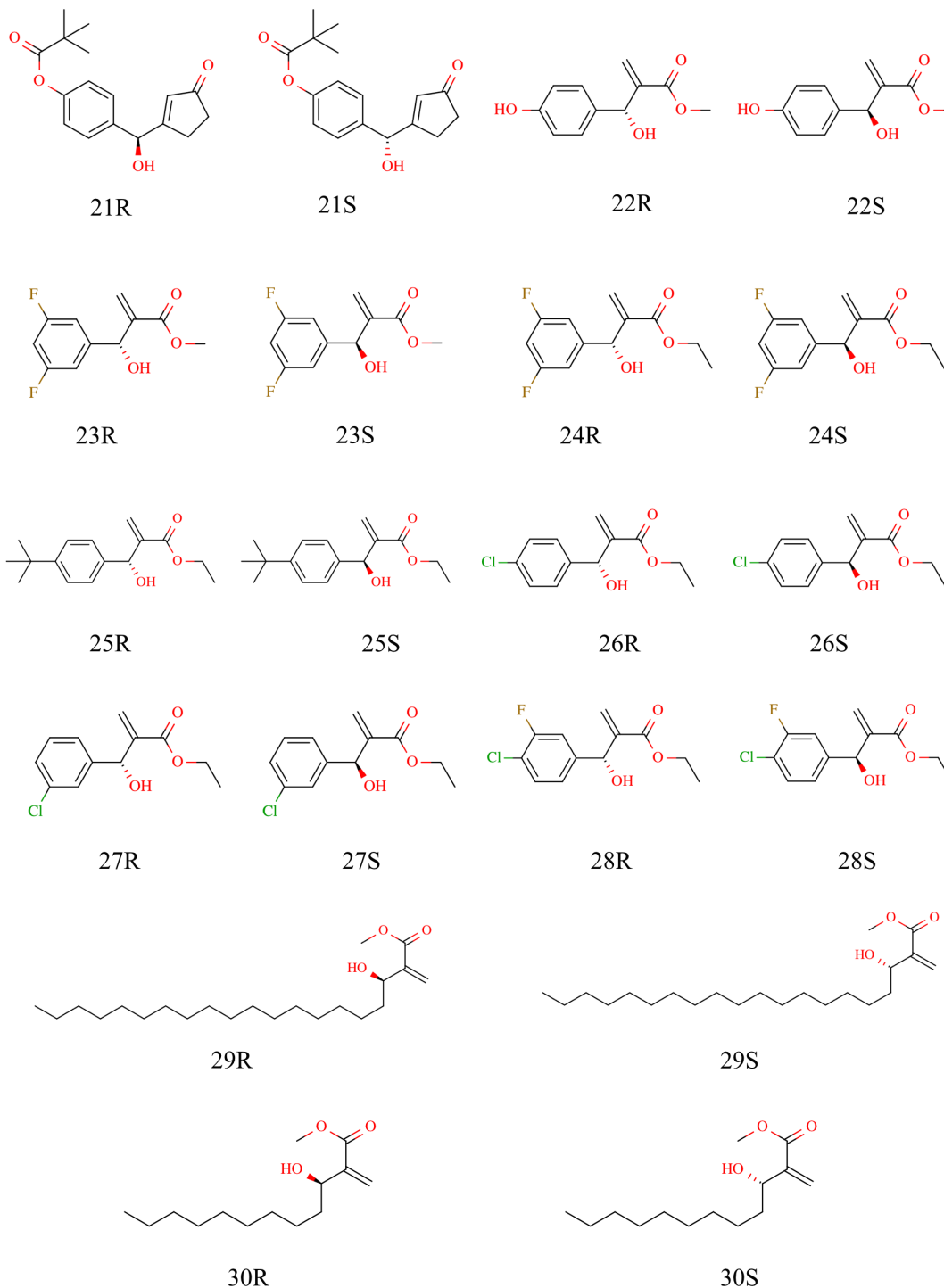


Figura 62 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: adutos de Morita-Baylis-Hilman - parte 3.

Fonte: Elaborada pelo autor.

## adutos de Baylis-Hilman

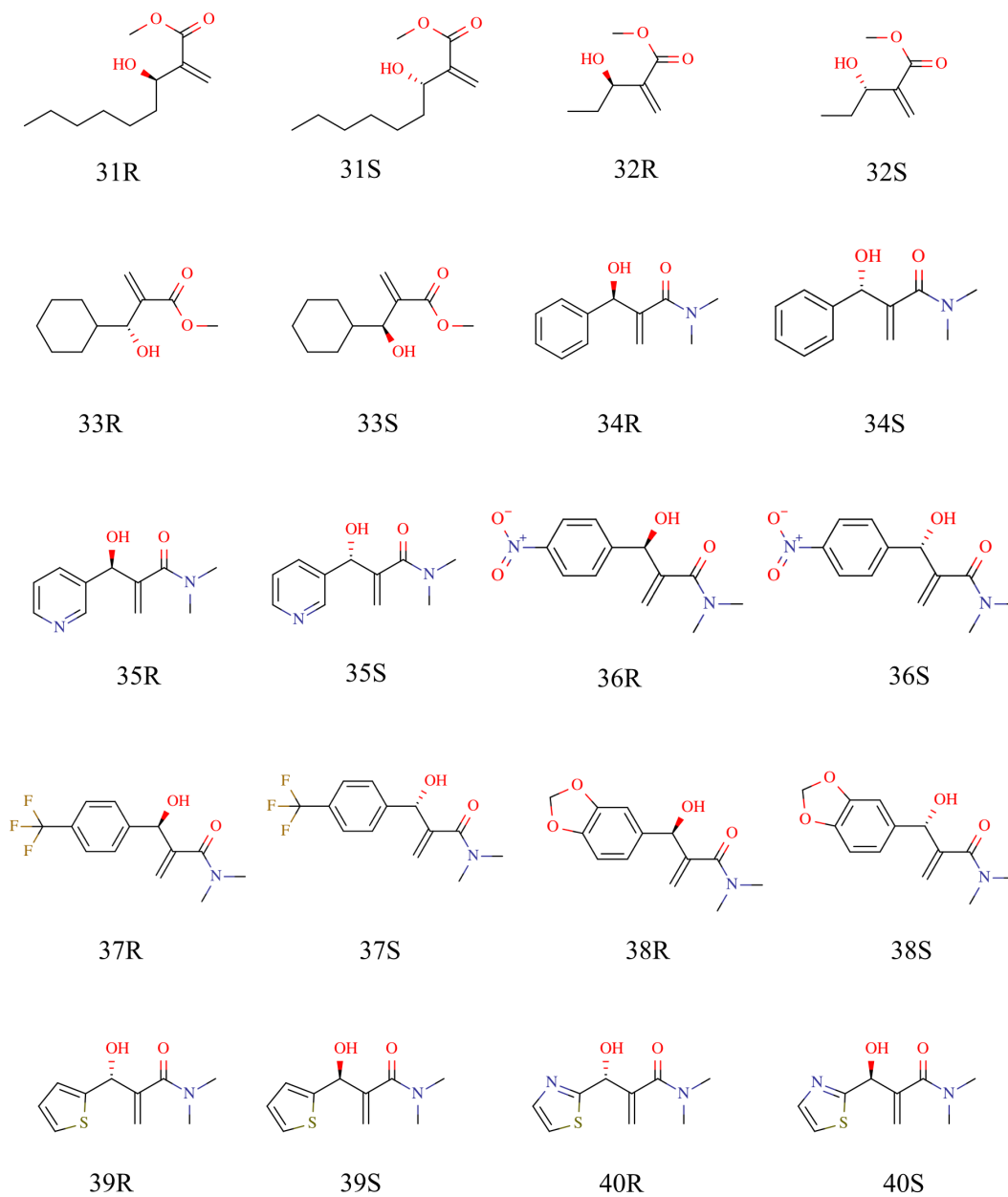


Figura 63 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: adutos de Morita-Baylis-Hilman - parte 4.

Fonte: Elaborada pelo autor.

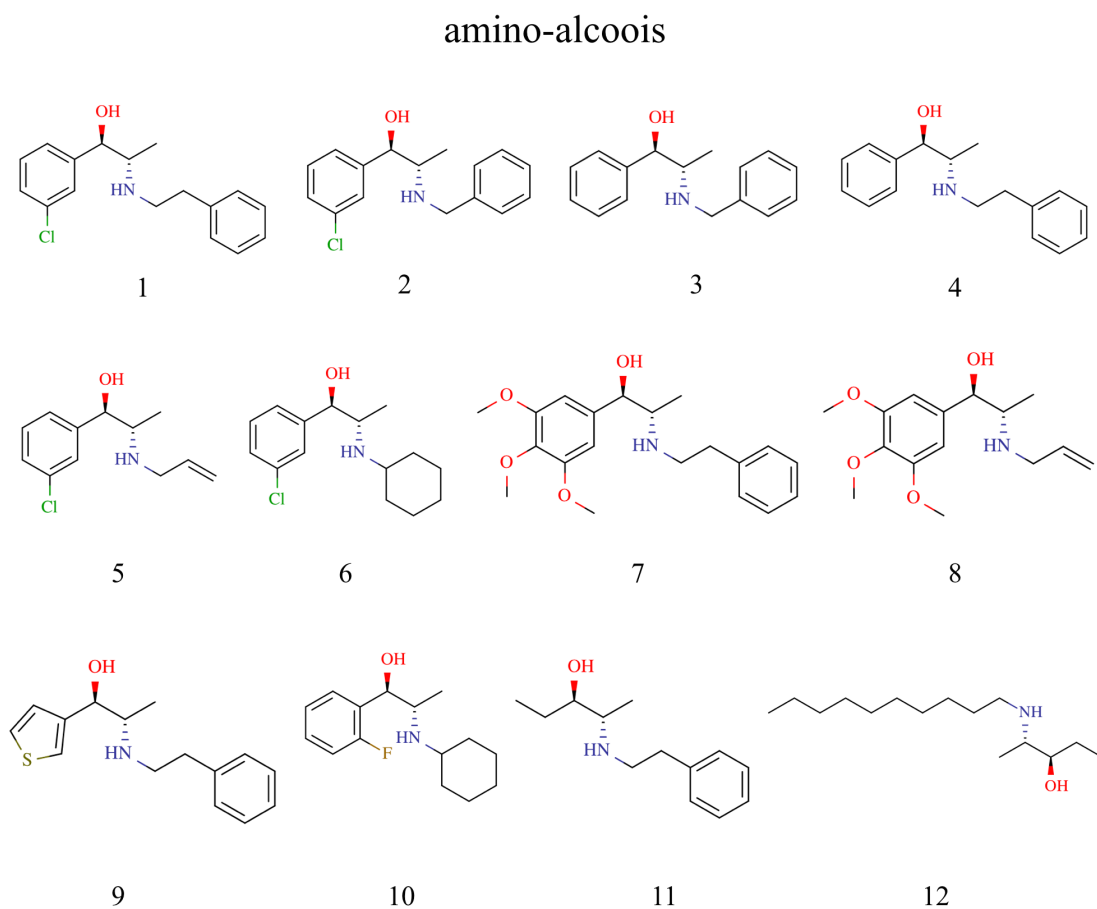


Figura 64 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: amino-alcoois.

Fonte: Elaborada pelo autor.

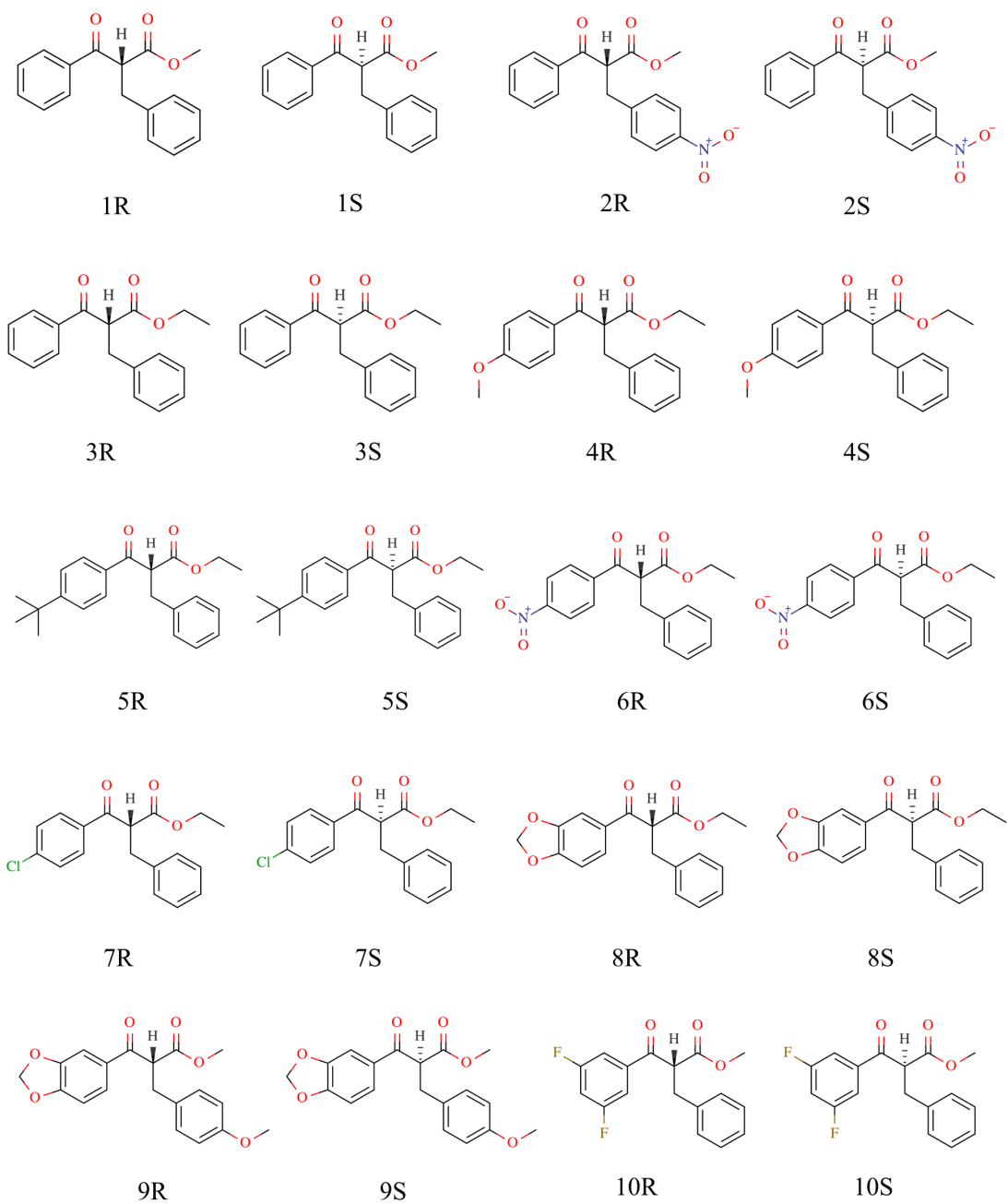
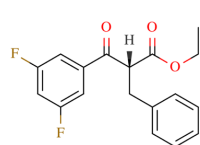
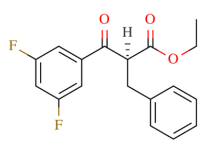
$\beta$ -cetoésteres

Figura 65 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos:  $\beta$ -cetoésteres - parte 1.

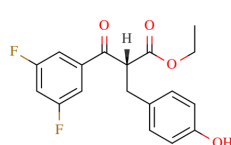
Fonte: Elaborada pelo autor.

$\beta$ -cetoésteres

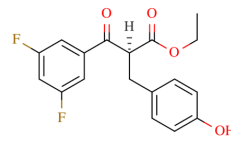
11R



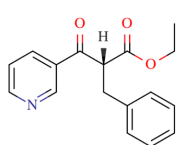
11S



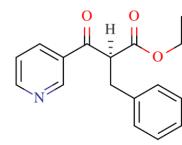
12R



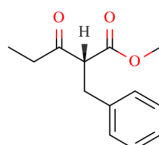
12S



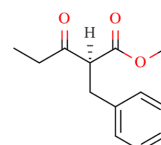
13R



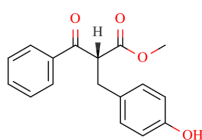
13S



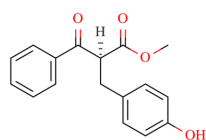
14R



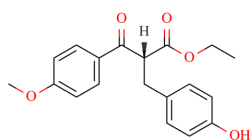
14S



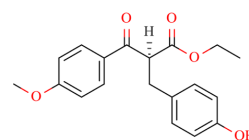
15R



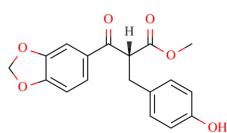
15S



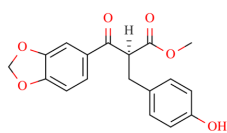
16R



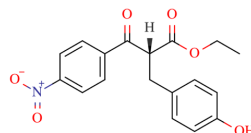
16S



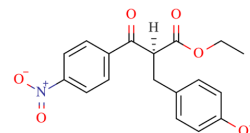
17R



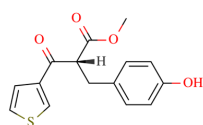
17S



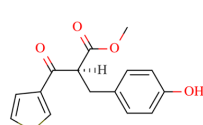
18R



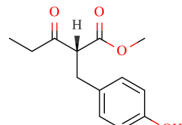
18S



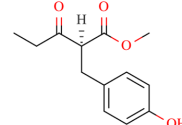
19R



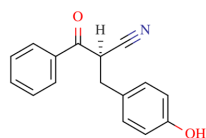
19S



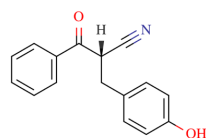
20R



20S



21R



21S

Figura 66 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos:  $\beta$ -cetoésteres - parte 2.

Fonte: Elaborada pelo autor.

## derivados indólicos

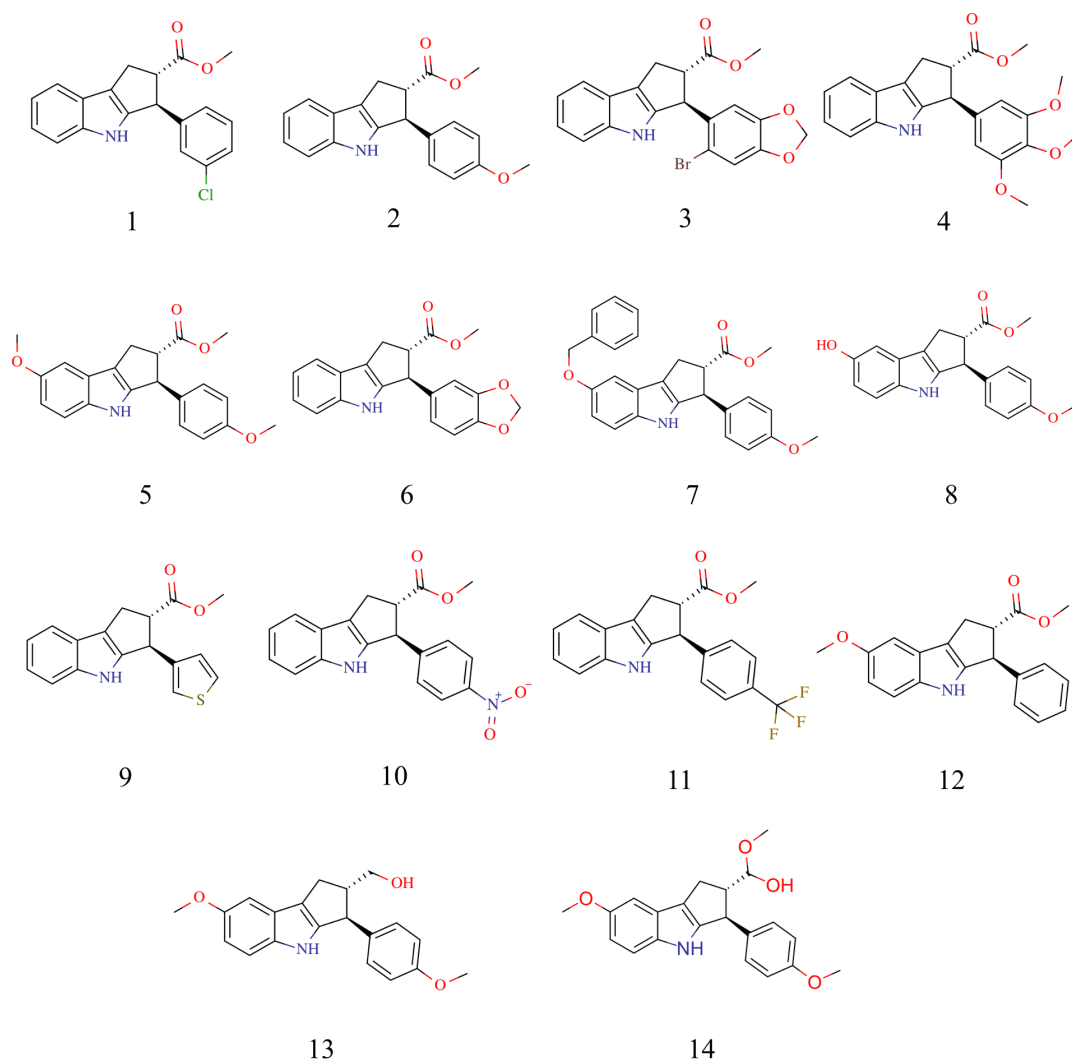


Figura 67 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: derivados ciclopenta- $\beta$ -indólicos.

Fonte: Elaborada pelo autor.

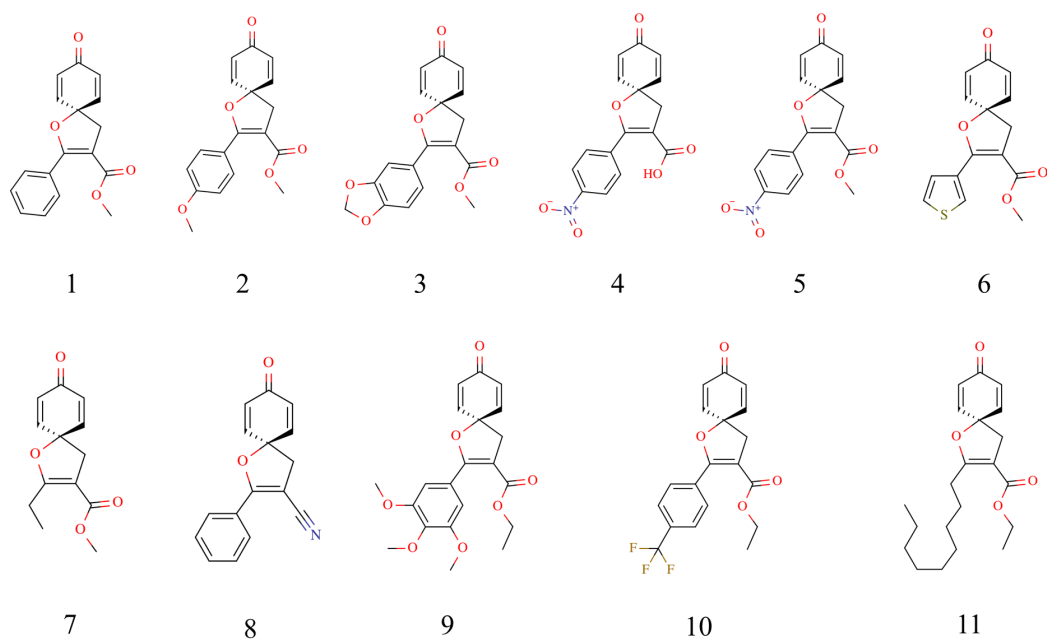
espirocliclopenta- $\beta$ -indóis

Figura 68 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: espirocicloexadienonas.

Fonte: Elaborada pelo autor.



## ftalídeos

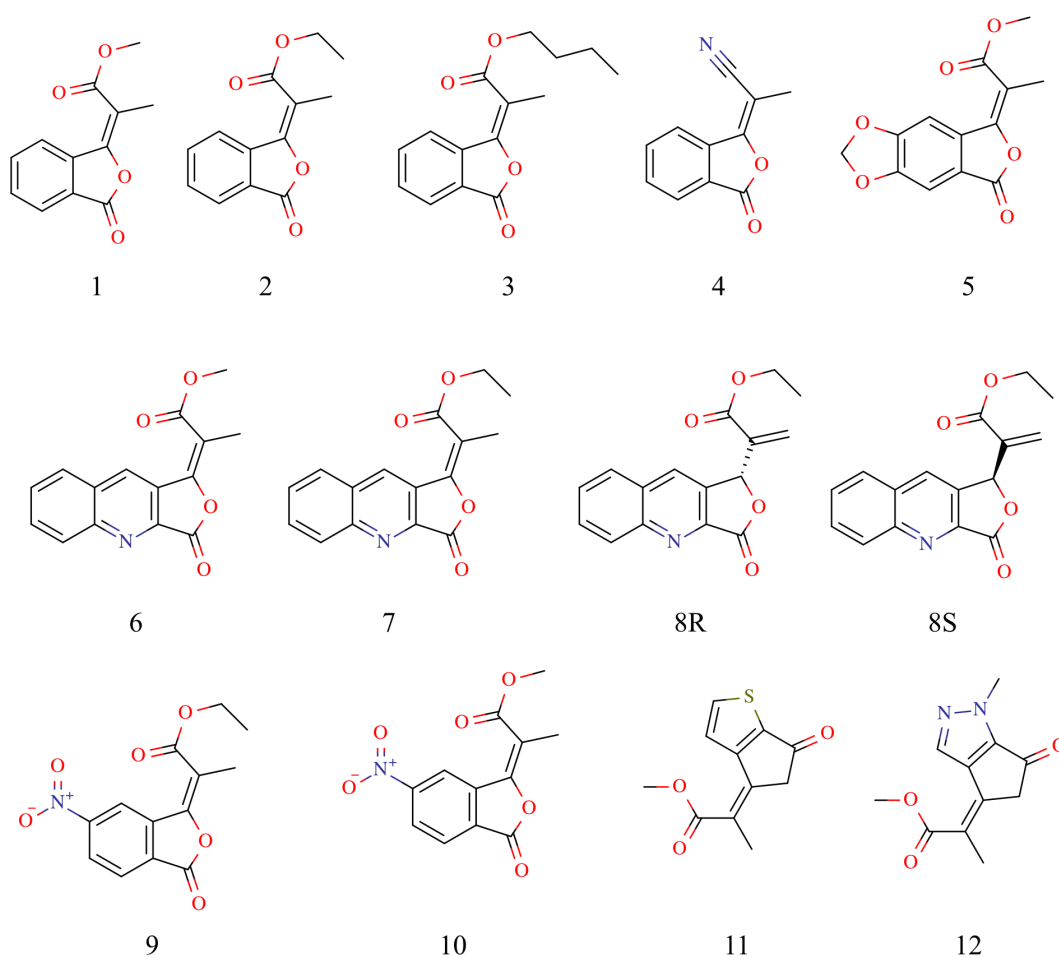


Figura 69 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: ftalídeos.

Fonte: Elaborada pelo autor.

## naftiridinas

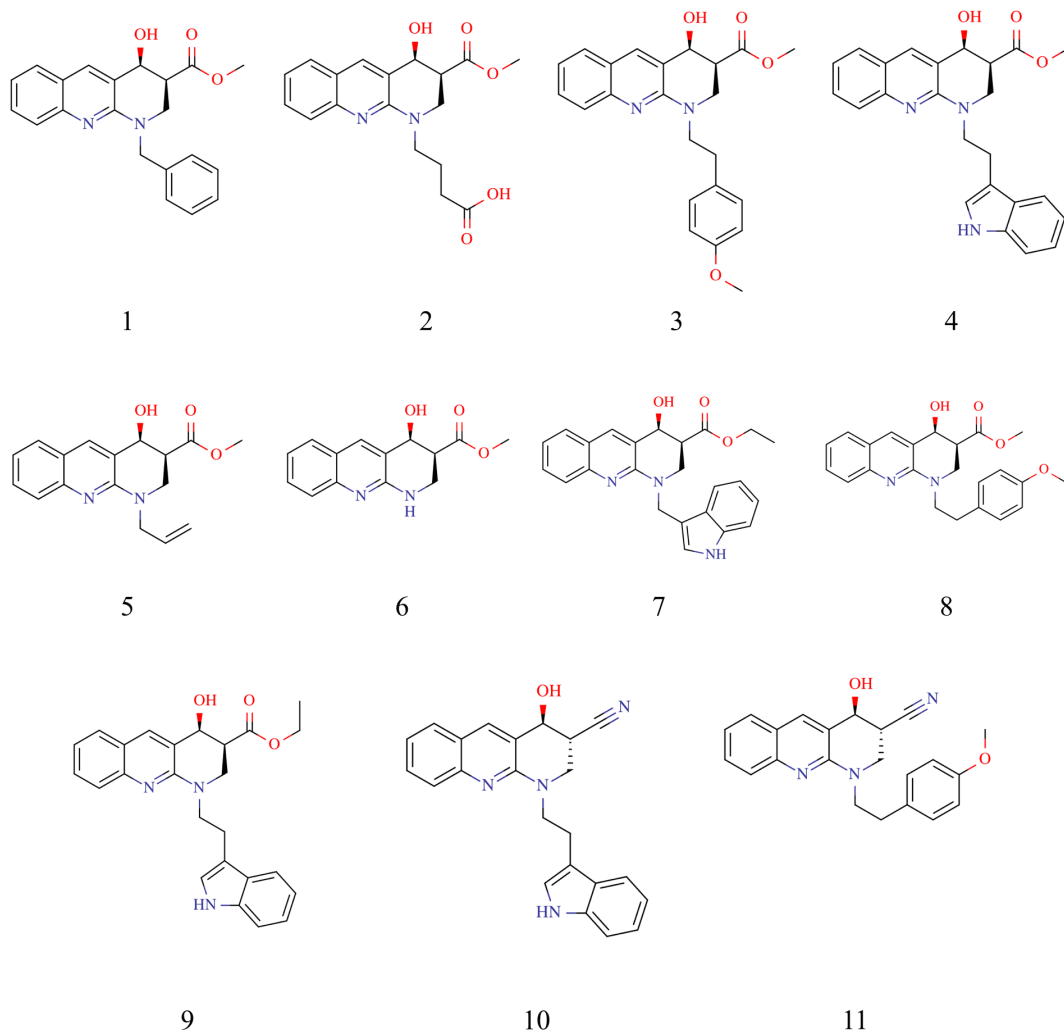


Figura 70 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: naftiridinas.

Fonte: Elaborada pelo autor.

## pirrolizidinas

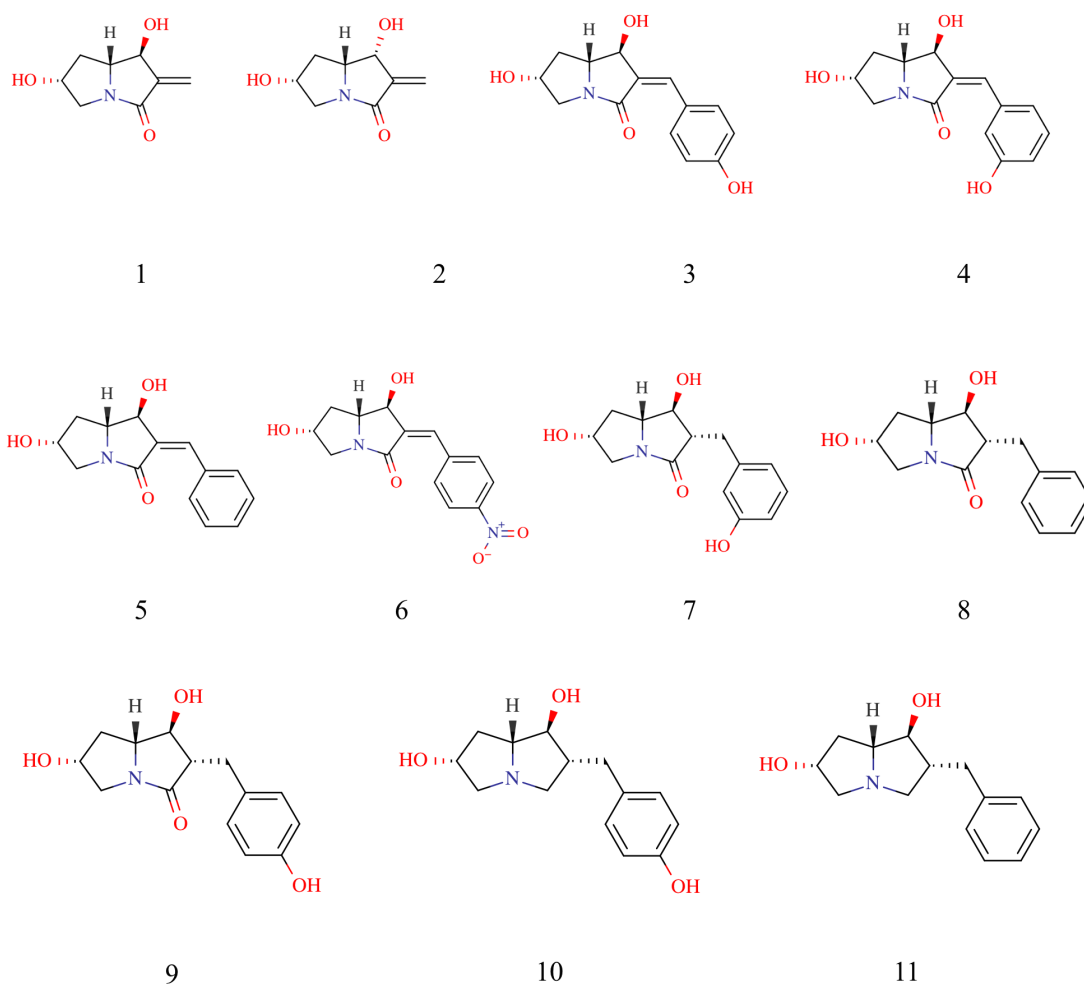
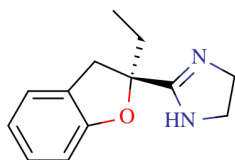
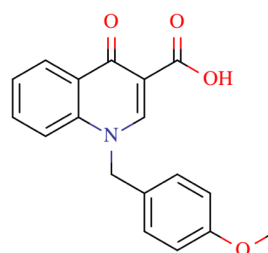


Figura 71 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: pirrolizidinas.

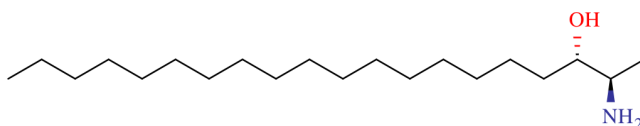
Fonte: Elaborada pelo autor.

derivado de benzofurano  
(Efaroxan)

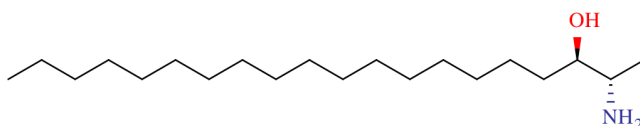
derivado de quinolina



aminoicosanóis



R



S

Figura 72 – Compostos utilizados para estudos de triagem virtual e experimental frente a microtúbulos: derivados de benzofurano, quinolina e aminoicosanóis.

Fonte: Elaborada pelo autor.

ANEXO B – Triagem bioquímica inicial dos compostos selecionados por triagem virtual.



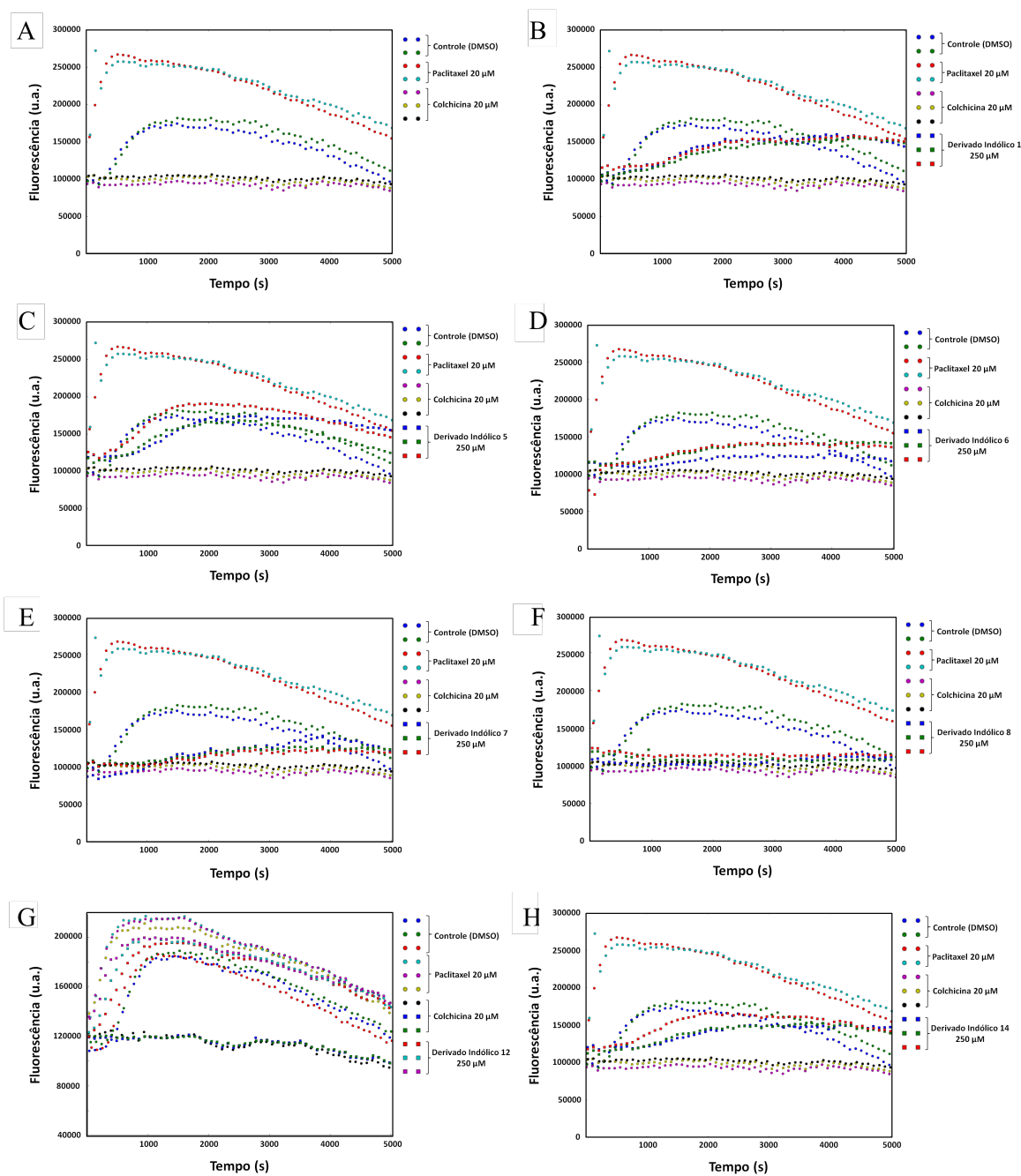


Figura 73 – Triagem bioquímica inicial dos compostos selecionados por triagem virtual frente a microtúbulos: (A) Ensaio controle. (B-H) derivados indólicos.

Fonte: Elaborada pelo autor.

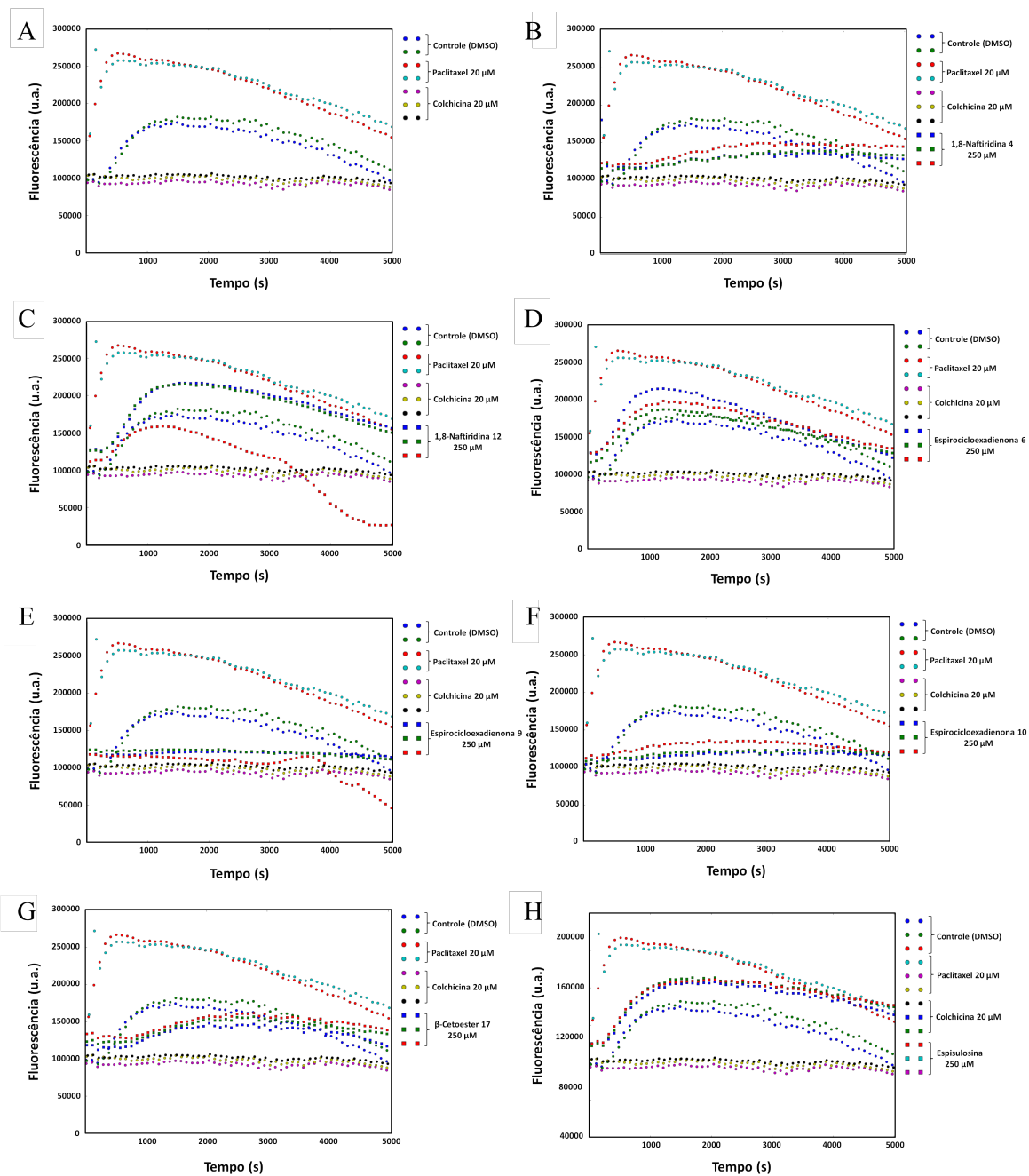


Figura 74 – Triagem bioquímica inicial dos compostos selecionados por triagem virtual frente a microtúbulos: (A) Ensaio Controle. (B-C) Naftiridinas. (D-F) Espirocicloexadienonas. (G)  $\beta$ -cetoéster. (H) Espisulosina.

Fonte: Elaborada pelo autor.



ANEXO C – Desvios quadráticos médios das simulações de dinâmica molecular com ligantes no sítio da colchicina do dímero de tubulina.



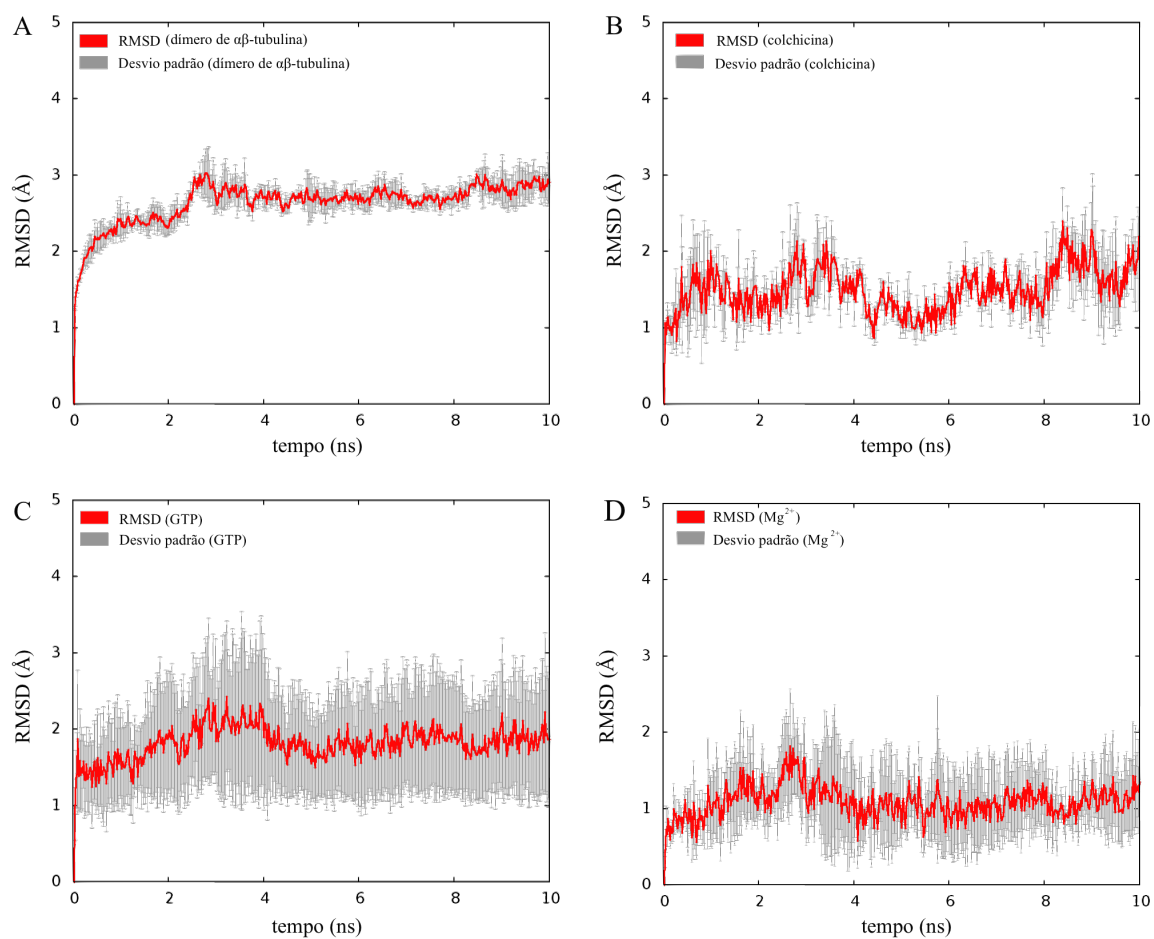


Figura 75 – Curvas de RMSD para os estudos de dinâmica molecular para o sistema contendo colchicina em complexo com o dímero de tubulina. O perfil médio de 3 simulações é apresentado em vermelho e o desvio padrão em cinza. (A) Dímero de tubulina. (B) Colchicina. (C) Co-fator GTP. (D) Co-fator  $Mg^{2+}$ .

Fonte: Elaborada pelo autor.

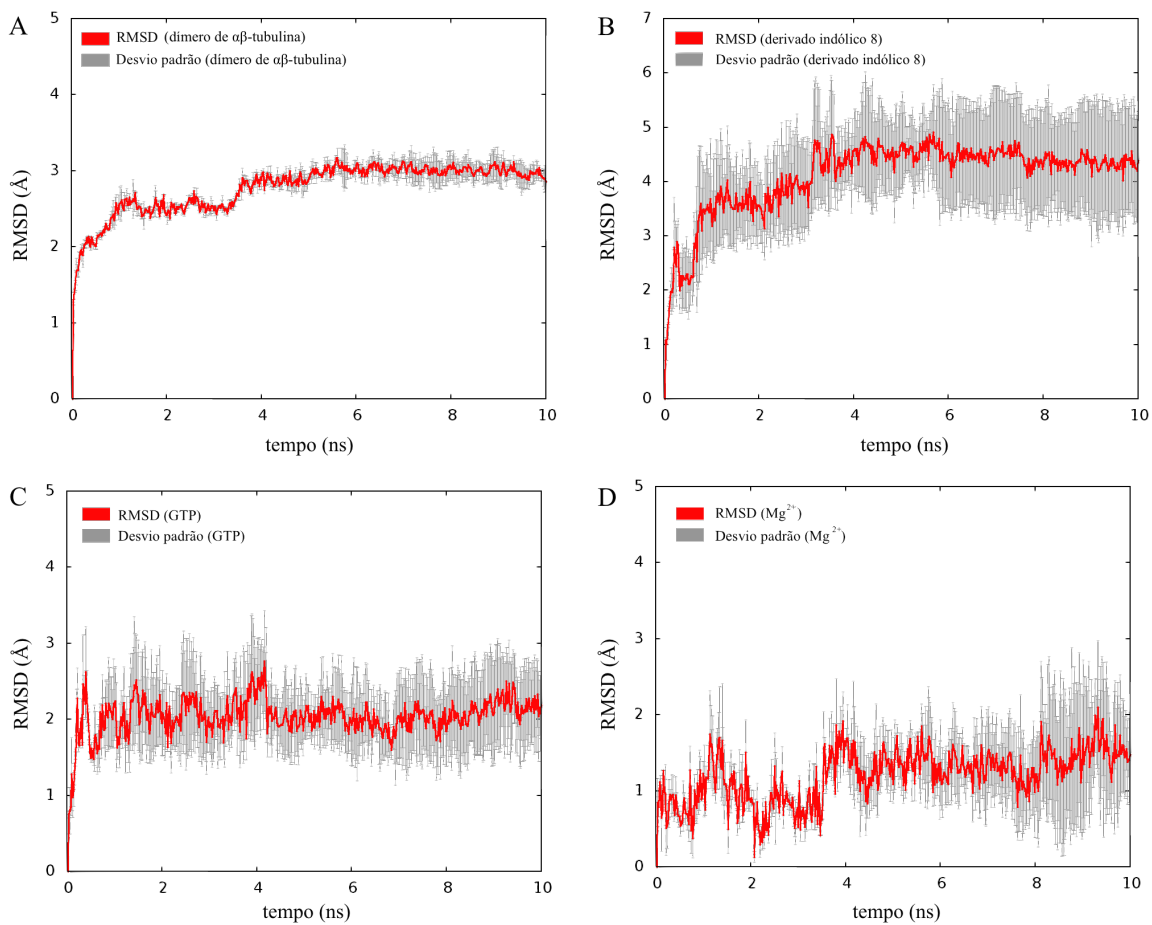


Figura 76 – Curvas de RMSD para os estudos de dinâmica molecular para o sistema contendo o derivado indólico **8** em complexo com o dímero de tubulina. O perfil médio de 3 simulações é apresentado em vermelho e o desvio padrão em cinza. (A) Dímero de tubulina. (B) Derivado indólico **8**. (C) Co-fator GTP. (D) Co-fator Mg<sup>2+</sup>.

Fonte: Elaborada pelo autor.

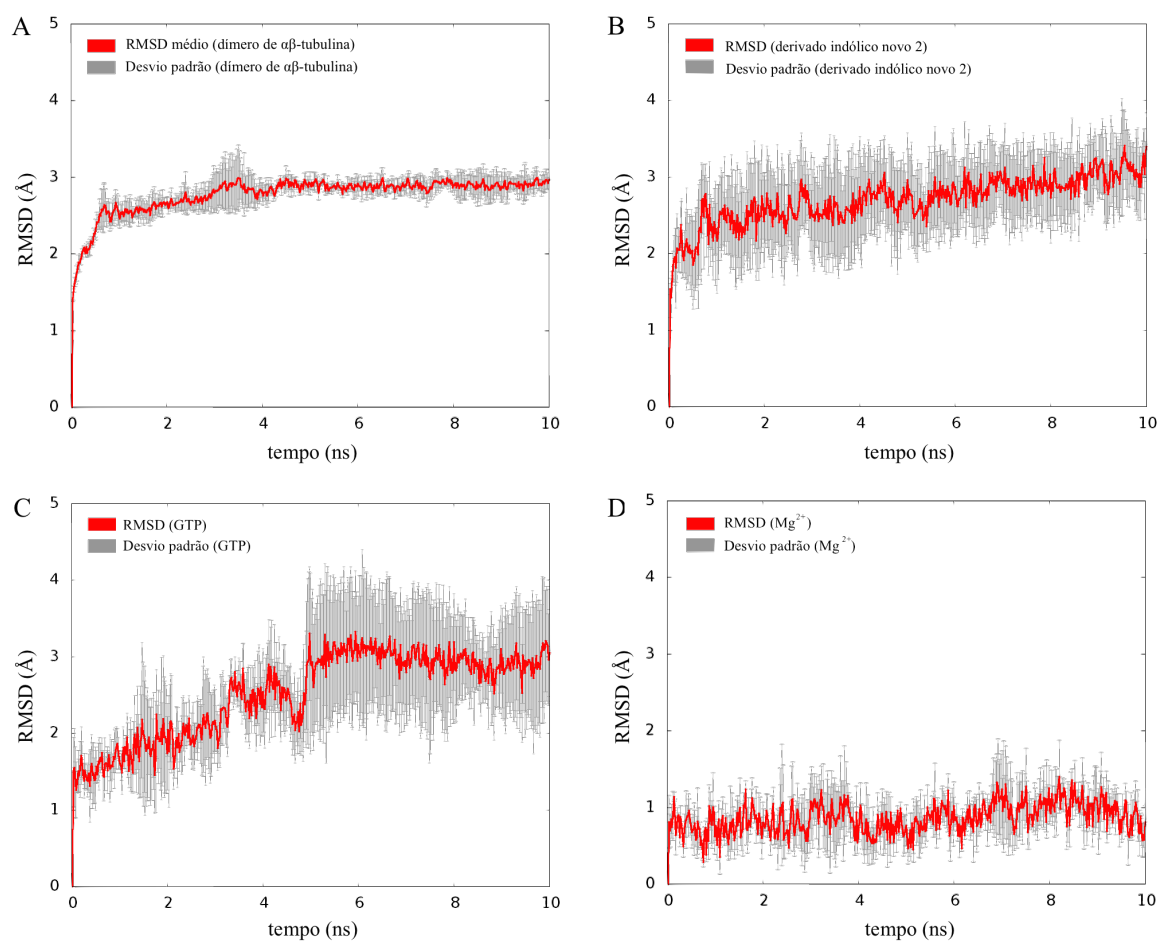


Figura 77 – Curvas de RMSD para os estudos de dinâmica molecular para o sistema contendo o derivado indólico **novo 2** em complexo com o dímero de tubulina. O perfil médio de 3 simulações é apresentado em vermelho e o desvio padrão em cinza. (A) Dímero de tubulina. (B) Derivado indólico **novo 2**. (C) Co-fator GTP. (D) Co-fator  $Mg^{2+}$ .

Fonte: Elaborada pelo autor.