

UNIVERSIDADE FEDERAL DE PELOTAS
Faculdade de Agronomia Eliseu Maciel
Departamento de Fitossanidade
Programa de Pós-Graduação em Fitossanidade

Dissertação



**Triagem virtual de moléculas inseticidas: a busca de potenciais análogos de
agonistas da ecdisona**

Larissa Longaray Machado

Pelotas, 2022

Larissa Longaray Machado

Triagem virtual de moléculas inseticidas: a busca de potenciais análogos de agonistas da ecdisona

Dissertação apresentada ao Programa de Pós-Graduação em Fitossanidade da Faculdade de Agronomia Eliseu Maciel da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Mestre em Fitossanidade (Área do Conhecimento: Entomologia).

Orientador: Prof. Dr. Moisés João Zotti

Pelotas, 2022

M149t Machado, Larissa Longaray

Triagem virtual de moléculas inseticidas : a busca de potenciais análogos de agonistas da ecdisona / Larissa Longaray Machado ; Moisés João Zotti, orientadora. — Pelotas, 2022.

71 f. : il.

Dissertação (Mestrado) — Fitossanidade, Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas, 2022.

1. Bioinformática. 2. Receptor de ecdisona. 3. Ecdise. 4. Descoberta de moléculas inseticida.. I. Zotti, Moisés João, orient. II. Título.

CDD : 632.951

Banca examinadora:

Prof. Dr. Moisés João Zotti (Orientador)
Doutor em Fitossanidade pela Universidade Federal de Pelotas

Prof. Dr. Daniel Bernardi
Doutor em Entomologia pela Universidade de São Paulo

Dr. Eduardo de Paula Costa
Doutor em Ciência da Computação pela KU Leuven (Bélgica)

Dr. Fred Schmitt Kremer
Doutor em Biotecnologia pela Universidade Federal de Pelotas

Dedico à minha família.

Agradecimentos

Agradeço aos meus pais, Rosângela Maria Longaray Machado e Dilomar Braga Machado, pelo apoio aos meus estudos e ajuda constante a conquista dos meus sonhos! Vocês são minha inspiração diária.

À minha avó, Sarita, por ser o melhor abraço em momentos de dificuldade.

À minha irmã Vanessa, por ser sempre a minha grande amiga, defensora e entusiasta, que sempre me aconselhou a me tornar a profissional que sou hoje.

À minha namorada Pamela, por estar do meu lado em momentos difíceis (e mais difíceis), por toda a paciência, por me incentivar a ser sempre melhor e por deixar minha vida mais alegre todos os dias.

E a toda minha família.

Aos meus amigos Tamires e Danrley, por serem as pessoas que sempre me mostraram os melhores caminhos, por falarem palavras de apoio e por serem presentes em toda a caminhada da minha vida acadêmica e profissional.

Aos meus colegas do programa de Pós-Graduação em Fitossanidade, em especial aos meus amigos dos estudos de bioinformática, das madrugadas de taxonomia e das melhores pizzas de Pelotas, Bruno e Juliana. À Indyra, pela parceria e amizade de anos e aos colegas e amigos Ericmar e Daniel. E à minha amiga Eduarda pelo incentivo e ajuda constante durante esses dois anos de crescimento juntas. E a todos meus amigos e colegas presentes nesse momento.

Ao meu orientador Professor Zotti, pela oportunidade e ensinamentos. E a todos os professores pelo conhecimento compartilhado e o esforço para dar aulas de qualidade durante a pandemia.

Ao Programa de Pós-Graduação em Fitossanidade e Universidade Federal de Pelotas pela estrutura e ensino. À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES).

Assim, agradeço a todos que me incentivaram e me ajudaram a persistir e agradeço também aqueles que seguem lutando pela pesquisa e educação do nosso país.

Muito obrigada!

“Feliz aquele que transfere o que sabe e aprende o que ensina”

Cora Carolina

Resumo

MACHADO, Larissa Longaray. **Triagem virtual de moléculas inseticidas: a busca de potenciais análogos de agonistas da ecdisona**. Orientador: Moisés João Zotti. 2022. 69f. Dissertação (Mestrado) – Programa de Pós-Graduação em Fitossanidade. Universidade Federal de Pelotas, Pelotas, Brasil.

Ferramentas de bioinformática se tornaram uma abordagem racional e promissora no campo da agroquímica. Os modelos *in silico* se mostram eficientes na predição de novos alvos. Dentro desse método promissor, destaca-se a técnica de triagem virtual de moléculas e a realização da acoplagem (docking) das moléculas em determinadas proteínas, formando o complexo proteína-ligante, por meio de abordagens computacionais. Para a busca de novas moléculas na área agrícola, o receptor da ecdisona se revela como uma via promissora para novos inseticidas, uma vez que faz parte dos reguladores de crescimento, os quais são considerados mais seguros para os inimigos naturais, insetos benéficos e organismos não-alvo. Essa segurança se dá, em partes, à ausência de receptores de ecdisona em mamíferos. Com base no exposto, a presente pesquisa teve o objetivo de utilizar a tecnologia de triagem virtual para identificar moléculas com alta possibilidade de ligação no sítio ativo do receptor de ecdisteroide (EcR), realizando a triagem de análogos de moléculas referência (tebufenozida, agonista sintético e o hormônio natural de muda 20HE) e o docking molecular da proteína nos ligantes selecionados. Foram analisados também características químicas e estruturais de um milhão, duzentos e setenta e seis mil, quatrocentos e setenta e seis (1.276.476) moléculas triadas, que geraram cento e quarenta milhões, setecentos e quarenta e quatro mil, setecentos e sessenta e oito (140.744.768) conformers. Nos resultados dos análogos de Tebufenozida, os compostos C₂₂H₁₅F₂N₃O₃S (código: ATDoc1), C₂₃H₁₇CLFN₃O₂(código:ATDoc2) e C₁₇H₁₄F₃N₃OS (código: ATDoc3;) mostraram os melhores valores de docking. Para os análogos de 20HE, os compostos C₂₂H₂₈O₅ (código:A20HEDoc1), C₂₁H₃₀O₄ (código:A20HEDoc2) e C₂₂H₂₉FO₄ (Código: A20HEDoc3) obtiveram as melhores pontuações. Por sua vez, para os compostos análogos do agonista sintético BY06830, as moléculas C₂₃H₁₉N₃O₄ (código: AASDoc1), C₂₀H₁₅N₃OS (código: AASDoc1) e C₂₃H₁₇N₃O₄S (código: AASDoc1) resultaram em melhores valores de acoplamentos. Por fim, a molécula C₂₃H₁₉N₃O₄ (código: AASDoc1) demonstrou melhor pontuação de docking.

Palavras-chave: Bioinformática. Receptor de ecdisona. Ecdise. Descoberta de moléculas inseticida.

Abstract

MACHADO, Larissa Longaray. **Virtual screening of insecticidal molecules: the search for potential analogues of ecdysone agonists**. Advisor: Moisés João Zotti. 2002. 69f. Dissertation. Postgraduate Program in Plant Health. Federal University of Pelotas, Pelotas, Brazil.

Bioinformatics tools have become a rational and promising approach in the field of agrochemistry. The in silico models are efficient in the prediction of new targets. Within this promising method, the technique of virtual screening of molecules and coupling (docking) of the molecules in certain proteins, forming the protein-ligand complex, through computational approaches, stands out. In the search for new molecules in the agricultural area, the ecdysone receptor is revealed as a promising route for new insecticides, since it is part of the growth regulators, which are considered safer for natural enemies, beneficial insects and non-target organisms. This safety is due, in part, to the absence of ecdysone receptors in mammals. Based on the above, the present research aimed to use virtual screening technology to identify molecules with a high possibility of binding at the active site of the ecdysteroid receptor (EcR), performing the screening of analogs of reference molecules (tebufenozide, synthetic agonist and the natural molting hormone 20HE) and the molecular docking of the protein in the selected ligands. Chemical and structural characteristics of one million, two hundred and seventy-six thousand, four hundred and seventy-six (1,276,476) sorted molecules were also analyzed, which generated one hundred and forty million, seven hundred and forty-four thousand, seven hundred and sixty-eight (140,744,768) conformers. In the results of Tebufenozide analogues, the compounds C₂₂H₁₅F₂N₃O₃S (code: ATDoc1), C₂₃H₁₇CLFN₃O₂ (code: ATDoc2) and C₁₇H₁₄F₃N₃OS (code: ATDoc3;) showed the best docking values. For the 20HE analogues, compounds C₂₂H₂₈O₅ (code: A20HEDoc1), C₂₁H₃₀O₄ (code: A20HEDoc2) and C₂₂H₂₉FO₄ (Code: A20HEDoc3) had the best scores. In turn, the analogous compounds of the synthetic agonist BY06830, the molecules C₂₃H₁₉N₃O₄ (code: AASDoc1), C₂₀H₁₅N₃OS (code: AASDoc1) and C₂₃H₁₇N₃O₄S (code: AASDoc1) resulted in better coupling values. Finally, the molecule C₂₃H₁₉N₃O₄ (code: AASDoc1) showed better docking scores.

Keywords: Bioinformatics. Ecdysone receptor. Ecdysis. Discovery of insecticide molecules.

Lista de Figuras

Figura 1 - Representação do modelo de chave fechadura em comparação com o docking. Onde (a) a chave equivale a (d) moléculas ligantes, (b) fechadura e o “buraco da fechadura” como (e) proteína e o sítio de ligação respectivamente e por fim a chave que consegue encaixar e abrir a fechadura (c), está representada como a molécula acoplada no receptor (o docking) (f).. **Erro! Indicador não definido.**

Figura 2 - Ilustração das linhagens mantidas em laboratório de entomologia molecular da UFPel para a realização dos testes, (a) SF9 e (b) S2, representando o docking da molécula no receptor de ecdisona em experimentação in vitro. **Erro! Indicador não definido.**

Figura 3 - (a) Etapas do processo de muda (Ecdise), adaptado de SALGADO, V. L. (2013) e (b) imagem do processo de ecdise de cigarra por PHILIP, T (2017). **Erro! Indicador não definido.**

Figura 4 - Molécula Tebufenozida sendo representada em diferentes conformações. **Erro! Indicador não definido.**

Figura 5 - Representação da Triagem Virtual de moléculas, onde as peneiras representam a seleção de moléculas (triagem) e no final a molécula selecionada acoplada com a proteína (docking). **Erro! Indicador não definido.**

Figura 6 - Moléculas referência utilizadas: Tebufenozida (a), BY06830 (b), 20HE (c). Representação em 2D. **Erro! Indicador não definido.**

Figura 7 - Diferentes formas de representar a proteína com a molécula acoplada, sendo da esquerda para direita: (a) em linhas com coloração em de acordo com o elemento – carbono em verde, hidrogênio cinza, nitrogênio azul, oxigênio vermelho, etc.*-, (b) em esferas com coloração por elemento*, (c) em superfície, (d) em fita com coloração por espectro e (e) com fita e colorindo a hélice em vermelho e o ciclo em verde. **Erro! Indicador não definido.**

Figura 8 - Representação da proteína em Surface com a molécula acoplada e em corte transversal (a), bem como o corte transversal com zoom no ligante da bolsa de

ligação mostrando as moléculas análogas acopladas no receptor sendo (b) análogos da Tebufenozida, (c) análogos do 20HE e (d) análogos do ligante de 1r20, com as melhores pontuações de *docking* da esquerda para a direita, dentro de cada grupo.

..... **Erro! Indicador não definido.**

Figura 9 - (a) A estrutura molecular dos aminoácidos envolvidos ligação nas principais moléculas identificadas e (b) estrutura molecular dos aminoácidos envolvidos na ligação nos inseticidas Tebufenozide, Metoxyfenozide, Cromafenozide, Halofenozide, segundo a simulação virtual de docking, gerados pelo software FRED..... **Erro! Indicador não definido.**

Figura 10 - No detalhe, a bolsa de ligação da proteína com Tebufenozide (a) e uma molécula análoga (b) e duas moléculas análogas na bolsa de ligação da proteína (c).

..... **Erro! Indicador não definido.**

Lista de Tabelas

- Tabela 1 - Algumas publicações com ferramentas de bioinformática sendo utilizadas como etapa inicial do Discovery de moléculas. 24
- Tabela 2 - Dados da Triagem Virtual de moléculas semelhantes à Tebufenozida, ao agonista sintético e à 20HE..... **Erro! Indicador não definido.**
- Tabela 3 - Análogos Tebufenozide com melhores pontuações de docking. 47
- Tabela 4 - Análogos 20HE com melhores pontuações de docking..... 48
- Tabela 5 - Análogos do Agonista Sintético BY06830 com melhores pontuações de docking. 49
- Tabela 6 - Inseticidas que agem no receptor de ecdisona os “agonistas de receptores de ecdisteroides” de acordo com IRAC 2022. 50

Lista de Abreviaturas e Siglas

20HE	20-hidroxiecdisona
A20HEDoc1	Análogos do 20HE - pontuação de docking número 1
A20HEDoc2	Análogos do 20HE - pontuação de docking número 2
A20HEDoc3	Análogos do 20HE - pontuação de docking número 3
AASDoc1 número 1	Análogos do Agonista Sintético BY06830 - pontuação de docking número 1
AASDoc2 número 2	Análogos do Agonista Sintético BY06830- pontuação de docking número 2
AASDoc3 número 3	Análogos do Agonista Sintético BY06830- pontuação de docking número 3
ATDoc1	Análogos de Tebufenozida – pontuação de docking número 1
ATDoc2	Análogos de Tebufenozida – pontuação de docking número 2
ATDoc3	Análogos de Tebufenozida – pontuação de docking número 3
ECR	Receptor de ecdisona
FRED	Fast Exhaustive Docking
LogP	Coefficiente de partição octanol-água
OMEGA	Gerador de bases de dados de estrutura multiconformer com alta velocidade.
PM	Peso molecular
ROCS	Rapid Overlay of Chemical Structures

Sumário

1.	Introdução geral	15
2.	Capítulo I - Revisão de Literatura	18
2.1.	Ferramentas de bioinformática: da triagem ao encaixe da molécula no receptor	18
2.1.1.	Bioinformática aplicada à Biologia molecular e estrutural	18
2.1.2.	Relação proteína-ligante e Triagem de novas moléculas	20
2.1.3.	Teoria do encaixe: <i>docking</i> molecular	22
2.2.	Funcionamento do processo natural do organismo do inseto e a ação do inseticida.....	28
2.2.1.	Processo de muda e ecdise.....	28
2.2.2.	Modo de ação dos inseticidas e receptores da ecdisona	31
3.	Capítulo II: Artigo 1 – Triagem virtual de moléculas inseticidas: a busca de potenciais análogos de agonistas da ecdisona	33
3.1.	Introdução.....	34
3.2.	Materiais e métodos.....	35
3.2.1.	Obtenção e preparação das bibliotecas de compostos e do receptor.....	35
3.2.2.	Triagem virtual e acoplamento das moléculas	37
3.2.3.	Modelagem tridimensional	39
3.3.	Resultados	40
3.3.1.	Validação do preparo antecedendo o <i>screening</i> e diminuição significativa do número de compostos a serem testados <i>in vitro</i>	40
3.3.2.	Coeficientes de <i>docking</i> entre a proteína e o ligante que demonstraram melhores acoplamentos	43
3.3.3.	Propriedades físico-químicas (regra dos 5), coeficiente de Tanimoto (ROCS) e ChemGauss4 (FRED)	46
3.4.	Discussão	51

3.4.1. Validação do preparo antecedendo o <i>screening</i> e diminuição significativa do número de compostos a serem testados <i>in vitro</i>	51
3.4.2. Coeficientes de <i>docking</i> entre a proteína e o ligante que demonstraram melhores acoplamentos.....	52
3.4.3. Propriedades físico-químicas (regra dos 5), coeficiente de Tanimoto (ROCS) e ChemGauss4 (FRED)	53
3.5. Conclusão	56
4. Considerações Finais	57
Referências.....	58
Apêndices	63
Anexos.....	67

1. Introdução geral

O Brasil é um dos grandes produtores de alimentos do mundo, e, de acordo com o Centro de Estudos Avançados em Economia Aplicada (Cepea), em 2020, o agronegócio brasileiro obteve uma participação de 26,6% no Produto Interno Bruto (PIB) no Brasil. Dentre os problemas que atingem a cadeia do agronegócio, figuram os danos causados por fatores bióticos, podendo chegar a perdas anuais de US\$ 314 bilhões, considerando apenas Austrália, África do Sul, Brasil, EUA, Índia, e Reino Unido, conforme o relatório da Agência das Nações Unidas para Alimentação e Agricultura (FAO, 2020).

Um importante inseto praga a nível mundial desta ordem é a *Spodoptera frugiperda* (Lepdoptera: Noctuidae), que faz com que, somente no Brasil, tenha se atingido gastos de mais de 600 milhões por ano para tentar controlar as infestações (FAO, [sd]). Esta lagarta, além dos danos causados no Brasil e de estar presente no continente Americano, segundo a Organização das Nações Unidas para a Alimentação e Agricultura, está em propagação mundial crescente desde 2016. No início de 2016, *S. frugiperda* foi detectada pela primeira vez na África Central e Ocidental, e posteriormente, relatada e confirmada em praticamente todo o continente da África Austral e Subsaariana. Em julho de 2018, foi confirmada também na Índia, no Iêmen, em Bangladesh, no Sri Lanka e na Tailândia. Em 2019, constatou-se a *S. frugiperda* em Mianmar, China, Indonésia, Laos, Malásia, Vietnã, Egito, República da Coreia, Japão. Em 2020, observou-se a sua propagação na Austrália, na Mauritânia e em Timor-Leste (FAO, 2020). Em razão dessa expansão, a lagarta-do-cartucho, como popularmente se conhece a *S. frugiperda*, tornou-se uma praga de importância mundial.

Espécies de mariposas podem gerar também grandes impactos em hortaliças, dado que, em fase larval, broqueiam os frutos (Michereff Filho et al., 2012), causando grandes danos na produção. Neste grande grupo de noctídeos, pode-se citar *S. frugiperda*, *S. cosmioides* e *S. eridania*, *Chrysodeixis includens*, *Helicoverpa zea*, *H. armigera*, *Chloridea virescens*, além de outros lepidópteros

como Tuta absoluta, *Neoleucinodes elegantalis* (Michereff Filho et al., 2012; Souza et al., 2016; Czepak et al., 2016; Michereff Filho et al., 2018).

Neste contexto, pode-se destacar o grupo 18 de modos de ação, os “Agonistas de receptores de ecdisteroides”, o qual possui apenas o grupo das diacilhidrazinas como representante, altamente eficiente e seletivo no controle de lagartas, com os seguintes ingredientes ativos: Cromafenozida, Halofenozida, Metoxifenozida e Tebufenozida (IRAC, 2021), todos de crescimento e desenvolvimento.

A busca por grupos cada vez mais seletivos de inseticidas é de suma importância, uma vez que os incorporar no Manejo Integrado de Pragas (MIP) poderia minimizar os efeitos deletérios do controle químico em organismos não- alvo (FERREIRA et al., 2006). O impacto dos inseticidas não seletivos na comunidade de artrópodes benéficos no ecossistema dificulta a combinação harmoniosa entre esses inseticidas e o controle biológico; por sua vez, os inseticidas seletivos se mostram bastante compatíveis com o controle biológico e podem ser utilizados no MIP (MACHADO et. al, 2019).

Um dos fatores que demonstra a necessidade do desenvolvimento de novos mecanismos de ação está ligado à resistência de insetos a inseticidas (APRD, [sd]). Já foram documentadas cerca de 800 espécies de insetos e ácaros resistentes a, pelo menos, uma classe de composto (FRAGOSO, 2014). Cada vez mais, há maiores impulsos para descobrir e desenvolver meios e métodos mais seletivos de controle de insetos que sejam compatíveis com os requisitos biológicos, toxicológicos, ambientais e sociais (ALTSTEIN et al., 1993).

A demanda das indústrias e dos grupos de pesquisa por descobertas de moléculas (fármacos ou agrotóxicos) é crescente, e por essa razão, metodologias mais rápidas, eficazes e de baixo custo mostram-se necessárias. Assim, a triagem virtual (virtual screening) tem se destacado como uma importante ferramenta na procura de compostos promissores (VERLI, et al., 2014). O grande avanço dessas técnicas para o planejamento de fármacos se dá devido ao fato de que a triagem virtual representa um processo mais dinâmico e eficiente na busca de substâncias bioativas de interesse (RODRIGUES et, al. 2012).

O receptor de ecdisteroides (EcR) é um importante alvo na pesquisa e desenvolvimento de novas moléculas e se divide em: agonistas da ecdisona, agonistas de esteroides da ecdisona e as dibenzohidrazinas (DBHs) (HU et al,

2017). Desta forma, pode-se afirmar que os inseticidas reguladores de crescimento se assemelham aos hormônios naturais e apresentam ação inibitória seletiva, baixa toxicidade em mamíferos e baixa resistência, sendo seus principais grupos de moléculas os inibidores da síntese de quitina, agonistas do hormônio juvenil e agonistas do hormônio da muda (TUNAZ; UYGUN, 2004).

Nesse contexto, os ecdisteroides apresentam-se como candidatos atraentes para inseticidas, pois possuem um bom potencial de segurança em mamíferos, incluindo seres humanos, uma vez que, embora os vertebrados tenham mecanismos de transcrição gênica similar, os ecdisteroides não perturbam esses sistemas (HARADA et al., 2011). Mostram-se ainda mais adequados por serem os hormônios ecdisteroides os coordenadores dos principais estágios do desenvolvimento do inseto, a muda e metamorfose, ligando-se ao receptor EcR. (BILLAS, et al., 2003).

Atualmente, o valor estimado de descoberta, desenvolvimento e registro de um novo ingrediente ativo para o mercado agrícola ultrapassa 180 milhões de dólares, não incluindo o valor das instalações especializadas (WHITFORD et al., 2009) e o tempo médio, para o mesmo processo, é de 10 anos. Devido ao alto valor para descoberta de moléculas inseticidas, cresce exponencialmente a busca por formas de reduzir custos e tempo. Assim, a triagem virtual de compostos, bem como outras ferramentas de bioinformática, mostra-se satisfatórias para este objetivo.

Assim, com base no proposto, o presente trabalho visou utilizar a tecnologia de triagem virtual para identificar moléculas com alta possibilidade de ligação no sítio ativo do EcR, realizar docking molecular das moléculas testadas e analisar suas características em comparação com as moléculas já consagradas no mercado.

2. Capítulo I - Revisão de Literatura

2.1. Ferramentas de bioinformática: da triagem ao encaixe da molécula no receptor

2.1.1. Bioinformática aplicada à Biologia molecular e estrutural

O início da bioinformática ocorreu antes mesmo do sequenciamento do DNA, há mais de 50 anos, e sofreu grande mudanças no decorrer dos anos (GAUTHIER, J. et al. 2019), com evolução a partir da década de 60, momento em que se utilizaram métodos computacionais na eletroquímica (OAKLEY, M. B.; KIMBALL, G. E. 1949; GAUTHIER, J. et al. 2019) e o primeiro banco de dados de sequência de proteínas. Após, entre 70 de 80, em paralelo com o salto de evolução da biologia molecular, a análise de genomas completos foi possível e, com o advento da internet, os anos 90 e 2000 demonstraram um salto ainda maior com mais ferramentas de bioinformática sendo desenvolvidas e descritas (GAUTHIER, J. et al. 2019).

A bioinformática pode ser definida como um estudo de dois eixos de informação na área da biologia molecular. O primeiro está relacionado com o Dogma Central da Biologia, uma vez que DNAs são transcritos em sequências de RNA, os quais são traduzidos em sequências de proteínas que, por sua vez, se enovelam em estruturas tridimensionais de proteínas funcionais. O segundo é o Método Científico, por meio de criação de hipóteses sobre atividades biológicas, desenho de experimentos, avaliação dos dados, extensão e modificação das hipóteses em decorrência dos dados (GU E BOURNE, 2009). Desta forma, ela é uma aplicação de ferramentas computacionais para coleta, análise, manipulação, apresentação e compartilhamento de dados (NCBI, [sd]). Nos últimos anos, os recursos de bioinformática foram amplamente utilizados na área da biotecnologia, da farmacologia e da medicina, destacando-se como uma das soluções para se trabalhar os problemas mundiais. Desde a descoberta e vigilância do SARS-Cov-2 (HU, T. et al, 2021), foi empregada no rastreamento de novas variantes (HODCROFT, E. B. et al., 2021) e na descoberta de potenciais alvos e caminhos

para o desenvolvimento de medicamentos (AUWUL, M. R. et al., 2021). Adotando conceitos parecidos, essas técnicas podem ser utilizadas para o desenvolvimento e descoberta de moléculas a serem utilizadas em potenciais trabalhos também no ramo agrícola. A biologia computacional trata-se de um distinto conceito da bioinformática, uma vez que é uma área de pesquisa específica com foco principal em problemas em biologia e, estes conceitos, apesar de parecidos, não podem ser confundidos (NCBI, [sd]). Ocorre que, no Brasil e em grande parte do mundo, apesar da distinção mencionada, ambos conceitos são intercambiáveis, utilizados de forma indistinta, e denotam o mesmo sentido, de forma que, neste trabalho, utilizar-se-á o termo bioinformática em seu amplo sentido.

As ferramentas computacionais se tornaram uma abordagem racional e promissora no campo da agroquímica com o rápido crescimento da biologia molecular e estrutural (SCHERKENBECK et. al, 2009; GONG et al., 2013). Essas ferramentas exploram grandes volumes de dados e utilizam bancos de dados que são mantidos por consórcios internacionais, Nucleotide Sequence Database (EMBL-Bank) no Reino Unido; o Banco de Dados de DNA do Japão (DDBJ), e GenBank do Centro Nacional de Informações sobre Biotecnologia (NCBI) (LESK, 2020), com o objetivo que servir como armazenamento, produção e organização destes dados.

A biologia estrutural é um ramo da molecular que explica e demonstra as conformações de macromoléculas, principalmente das proteínas (SILVA et, al. 2020). Dessa forma, este ramo é essencial para o desenvolvimento de “modelos *in silico*” que vem sendo utilizados para o design de fármacos desde os anos 70 e para auxiliar na descoberta de moléculas e suas conformações. Na década de 90, começam a ser desenvolvidos e estimulados os sistemas de triagem de alto desempenho, triagem com receptores específicos, química combinatória utilizando a biologia estrutural, entre outros conceitos que serviram como salto na época (LAMBERTH et al., 2013).

Os modelos *in silico* se mostram eficientes na predição de novos alvos moleculares, possibilitando uma otimização no processo de descoberta, de modo que podem sugerir novas indicações mais vantajosas que as propostas, possibilitar uma possível otimização de compostos-líder e indicá-los para alvos mais específicos, entre outras inúmeras propostas (SILVA et, al. 2020).

Dentro desse cenário promissor, a técnica de triagem virtual é um exemplo de

técnica que possui alta capacidade de analisar compostos quantitativamente e qualitativamente, podendo selecioná-los de acordo com critérios pré-definidos que se mostraram ativos para determinado alvo escolhido, o receptor (VERLI, et al., 2014).

2.1.2. Relação proteína-ligante e Triagem de novas moléculas

O modelo chave-fechadura, que explica o atracamento entre proteína e ligante, foi proposto por Emil Fischer em 1894. Nesse modelo, o receptor equivaleria à “fechadura”, o sítio de ligação ou do receptor corresponderia ao “buraco da fechadura” e o ligante se equipararia à possível “chave”, sendo que a sua interação seria a “abertura” ou “fechamento” desta porta (Figura 1). No entanto, o modelo proposto por Fischer acaba sendo limitado e pode gerar interpretações errôneas em relação ao fato de que a “fechadura” aparentemente é algo rígido e estático, ao passo que, em meio biológico, tanto o ligante quanto a proteína são flexíveis (VERLI, et al., 2014).

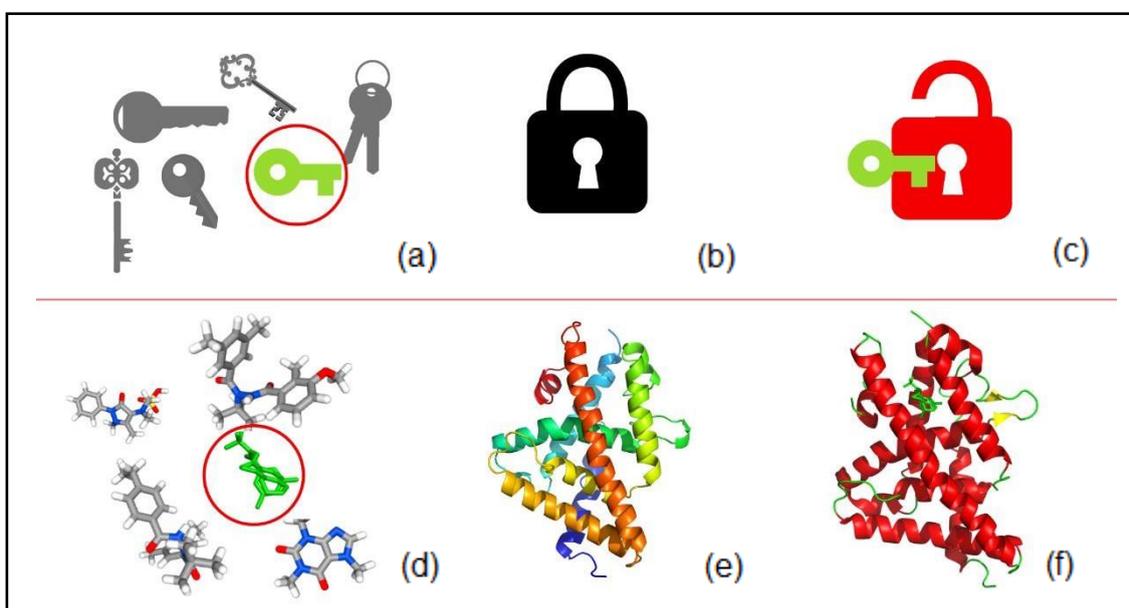


Figura 1 - Representação do modelo de chave fechadura em comparação com o docking. Onde (a) a chave equivale a (d) moléculas ligantes, (b) fechadura e o “buraco da fechadura” como (e) proteína e o sítio de ligação respectivamente e por fim a chave que consegue encaixar e abrir a fechadura (c), está representada como a molécula acoplada no receptor (o docking) (f).

Muitas técnicas estão sendo desenvolvidas nesse campo, inclusive computacionais (*in silico*), e entre elas, a triagem virtual é a experimentação

que consegue selecionar prováveis moléculas para se ligar no receptor alvo. Essa técnica afasta o problema de uma busca ampla no espaço químico, na medida em que seleciona bibliotecas específicas e acessíveis, evitando assim sínteses dispendiosas, com alto custo e tempo, e permite, ainda, a aplicação de filtros que garantem que a biblioteca atenda a um alto padrão de relevância biológica e similaridade com compostos já descobertos (SHOICHET, 2004).

Apesar das técnicas serem essenciais no auxílio para a descoberta de novas moléculas, existem dois principais problemas a serem ultrapassados quando se trata de atracamento molecular. O primeiro problema diz respeito à investigação e à predição das conformações e orientações de uma molécula ligante no seu sítio de complexação, e o segundo problema guarda relação com a predição da afinidade em um complexo receptor-ligante (energia livre de ligação) (VERLI, et al., 2014).

Pode-se utilizar diversos softwares e ferramentas de modelagem molecular com o intuito de otimizar o processo da busca e facilitar a solução dos problemas já citados. O *virtual screening* (ou triagem virtual) é uma abordagem bastante eficiente quando utilizada na descoberta de medicamentos, por exemplo, na qual os prováveis compostos ativos são selecionados com base em propriedades físico-químicas e/ou na conformação molecular favorável para a interação ligante-receptor, antecedendo um bioensaio da atividade de interação (HARADA et. al, 2011).

A triagem pode ser realizada com diferentes metodologias, algumas irão depender de moléculas já conhecidas (*ligand-based*) e outras irão se basear em informações da estrutura alvo (*structure-based*) (COURNIA et al., 2020).

Apesar de bastante interessante, quando nenhum ligante conhecido está presente no receptor, a técnica tomará um tempo maior e exigirá dados mais precisos do receptor e um longo tempo de experimentação *in silico* (computacional). Enquanto o outro, é baseado em um algoritmo mais simples, com menor velocidade e maior precisão no cálculo quando há um ligante conhecido para o receptor (HARADA et al, 2011). Isso ocorre porque, os métodos de triagem baseados em ligante vão utilizar as informações presentes em ligantes ativos conhecidos (HAMZA et. al, 2012). A respeito desse método, explica VERLI (2014):

O método baseado na estrutura é mais utilizado quando a estrutura tridimensional da molécula receptora está disponível com boa qualidade. Nesta metodologia, é realizado um estudo de atracamento molecular de todos os ligantes previamente selecionados, ao invés de apenas uma molécula. É possível, assim como no estudo de atracamento molecular tradicional, considerar a flexibilidade do receptor diretamente pelo programa de atracamento ou utilizar um conjunto de conformações da molécula receptora. Entretanto, o custo computacional aumenta significativamente ao se incluir a flexibilidade do receptor em estudos de triagem virtual. Quando não é possível obter a estrutura tridimensional do receptor, ainda que por técnicas sofisticadas de predição de estruturas de macromoléculas, então o método baseado na estrutura do ligante é empregado. Esta abordagem consiste na análise de similaridade de propriedades estruturais e físico- químicas de compostos ativos e inativos. (VERLI, et al., 2014 p. 206).

Assim, tanto métodos baseados na estrutura quanto no ligante, são ótimas ferramentas que se encarregam na descoberta de novas moléculas e também para compreender as relações entre estruturas e atividade entre eles (SAINI; KUMAR, 2014).

2.1.3. Teoria do encaixe: *docking* molecular

O *docking* molecular prevê a estrutura do complexo proteína-ligante utilizando técnicas de bioinformática (MENG, et al. 2011). Dentre as diferentes formas de realizar este encaixe por abordagens computacionais, pode ser (1) realizando a comparação do ligante e sua semelhança com os compostos triados, aplicando algoritmo e produzindo uma previsão das posições, (2) realizando o exame sistemático das posições possíveis e utilizando o encaixe de rendimento ultra-alto (> 1 milhão de compostos) triando via características de pontuação e (3) o encaixe por ser guiado pelo ligante, ou seja, reduz a busca com base em forma e complementariedade química para as moléculas já conhecidas (OPENEYE SCIENTIFIC, 2022). Todas essas ferramentas resultarão em prováveis compostos que irão encaixar no receptor, porém algumas focam em técnicas que há um receptor (proteína) já conhecido e cristalografado e outras focadas em ligantes (moléculas) já utilizadas no mercado.

Quando ocorre a ausência completa da proteína, utiliza-se metodologias baseadas somente na similaridade da molécula que utilizam coeficientes como o coeficiente de Tanimoto (OPENEYE SCIENTIFIC, 2022), uma vez que este índice é identificado como um dos melhores métricos para cálculos de similaridade (BAJUSZ,

2015).

O desenvolvimento de trabalhos de previsão de poses e triagem de compostos dos últimos anos demonstra excelentes resultados (Tabela 1). Esses trabalhos auxiliam na validação de ferramentas e fluxos de trabalho de encaixe molecular. Abordando o *docking*, similaridade química entre compostos e *desing*, como previsões complementares para ferramentas de triagem virtual.

Tabela 1 - Algumas publicações com ferramentas de bioinformática sendo utilizadas como etapa inicial do Discovery de moléculas.

Título	Ano de publicação	Autores	Assunto Geral
“Alchemical Free Energy Methods Applied to Complexes of the First Bromodomain of BRD4”	2022	Guest, Ellen E., et al.	Previsão precisa e rápida da afinidade de ligação
“Identification, Characterization, and Synthesis of Natural Parasitic Cysteine Protease Inhibitors: Pentacitidins Are More Potent Falcitidin Analogues”	2022	Brinkmann, Stephan, et al	Caracterização e síntese de análogos
“EViS: An Enhanced Virtual Screening Approach Based on Pocket–Ligand Similarity”	2022	Zhang, Wenyi, e Jing Huang	Triagem virtual integrando o encaixe de ligantes, pesquisa de modelos de bolso de proteínas e cálculo de similaridade de formas de ligantes.
“Ulotaront: A TAAR1 Agonist for the Treatment of Schizophrenia”	2022	Heffernan, Michele L. R., et al.	Descoberta de análogos de agonista de receptor para o tratamento de doenças
“Discovery by Virtual Screening of an Inhibitor of CDK5-Mediated PPAR γ Phosphorylation”	2022	O’Mahony, Gavin, et al.	Descoberta por triagem virtual de um inibidor
“A Blueprint for High Affinity SARS-CoV-2 Mpro Inhibitors from Activity-Based Compound Library Screening Guided by Analysis of Protein Dynamics”.	2021	Gossen, Jonas, et al.	Abordagem computacional para rastrear alvos de proteínas com sítios de ligação maleáveis
“Lean-Docking: Exploiting Ligands’ Predicted Docking Scores to Accelerate Molecular Docking”Computationally	2021	Berenger, Francois, et al	Triagem virtual baseada em estrutura

Screened Inhibitors”. “Identification of Phellgridin-Based Compounds as Novel Human CD73 Inhibitors”	2021	Lyu, Sifan, et al.	Identificação de novos inibidores para o desenvolvimento de fármacos
“Computational Graphics Software for Interactive Docking and Visualization of Ligand–Protein Complementarity”	2021	Baskaran, Saravana G., et al	Ancoragem e complementaridade ligante-proteína
“Efficient Exploration of Chemical Space with Docking and Deep Learning”.	2021	Yang, Ying, et al.	Triagem virtual e exploração do espaço químico
“XGraphBoost: Extracting Graph Neural Network-Based Features for a Better Prediction of Molecular Properties”.	2021	Deng, Daiguo, et al.	Rede neural para melhor previsão de propriedades moleculares
“Identification of a Chlorogenic Ester as a Monoamine Oxidase (MAO-B) Inhibitor by Integrating ‘Traditional and Machine Learning’ Virtual Screening and In Vitro as Well as In Vivo	2021	Dhanabalan, Anantha Krishnan, et al	Identificação de um inibidor por triagem virtual e validação in vitro e in vivo
Validation: A Lead against Neurodegenerative Disorders?” “Molecular Simulation of Avβ6 Integrin Inhibitors”.	2020	Guest, Ellen E., et al.	Simulações virtuais e previsão de afinidade de ligação
“Augmenting Hit Identification by Virtual Screening Techniques in Small Molecule Drug Discovery”	2020	Yan, Xin Cindy, et al.	Triagem virtual para a descoberta de drogas

“Combining Docking Pose Rank and Structure with Deep Learning Improves Protein–Ligand Binding Mode Prediction over a Baseline Docking Approach”.	2020	Morrone, Joseph A., et al.	Previsão de ligação proteína-ligante
“Structure-Based Discovery of Novel Ligands for the Orexin 2 Receptor”.	2020	Gunera, Jakub, et al	Descoberta de novas moléculas baseada na estrutura

Os trabalhos de triagem virtual são o primeiro passo, o passo seguinte é a seleção *in vitro* para a descoberta de novas moléculas utilizando diferentes linhagens de células (Figura 2), para além do organismo alvo. A partir disso, obtém-se uma previsão de como será a seletividade nos organismos *in vivo* e se demonstra a ligação do alvo no receptor que foi determinada via parâmetros computacionais.

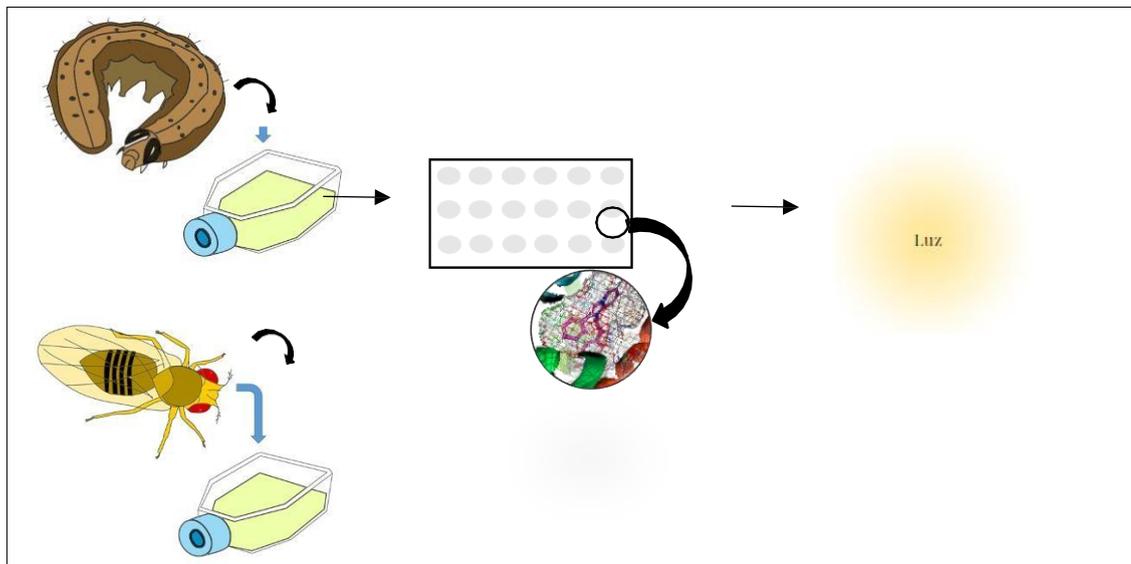


Figura 2 - Ilustração das linhagens mantidas em laboratório de entomologia molecular da UFPel para a realização dos testes, (a) SF9 e (b) S2, representando o docking da molécula no receptor de ecdisona em experimentação *in vitro*.

As linhagens celulares devem ser mantidas em laboratório, e a síntese das moléculas deve ser feita após a triagem como em SMAIL, M. I. et al. (2020) que realizou o trabalho de descoberta, síntese e avaliação biológica dos novos compostos. Para a realização da etapa *in vitro*, é necessário também a verificação da viabilidade das linhagens celulares, além do ensaio propriamente dito e após, devem ser realizadas etapas *in vivo* como ocorreu em DHANABALAN, A. K. et al. (2021) e SERAFINI, M. et al. (2020), trabalho em que a identificação do inibidor foi validado *in vitro* e *in vivo*.

2.2. Funcionamento do processo natural do organismo do inseto e a ação do inseticida

2.2.1. Processo de muda e ecdise

Para o desenvolvimento de inseticidas, é importante ter a compreensão plena dos processos envolvidos. O crescimento e desenvolvimento dos insetos, por exemplo, requer a produção periódica de um novo exoesqueleto cuticular, um processo denominado ecdise (Figura 3). Entre os insetos, existe uma riqueza de variações nos fatores que controlam este processo, e essas variações são evidentes tanto em nível de espécie quanto de estágio. A ecdise é geralmente dividida em fases: (1) Pré-ecdise, em que o inseto se fixa ao substrato, seguindo movimentos que rompem as conexões musculares remanescentes entre o inseto e sua velha cutícula. Nesta fase, o inseto absorve fluido e enche sua nova traqueia com ar. As principais características da fase inicial é que os padrões motores usados têm como objetivo de se preparar para liberar um exoesqueleto (2) Ecdise, correspondente à troca efetiva do exoesqueleto (3) Pós-ecdise: ocorre à expansão e o endurecimento da nova cutícula. Nesta fase, ocorre uma mudança na atividade cardíaca e a deglutição de ar ou água para expandir a cavidade do corpo (TRUMAN, 2005).

Todo esse processo é controlado por uma complexa cascata de hormônios peptídicos de células endócrinas e neuropeptídeos no sistema nervoso (ZITÑAN et al., 2007), o que determina que a sequência de ecdise tenha sido bem sucedida. Assim, as células produzem hormônios desencadeadores de pré-ecdise e ecdise que ativam a sequência de ecdise por meio de ações mediadas por receptores em neurônios específicos no sistema nervoso central (MESCE et al., 2002). Um neuropeptídeo, o hormônio protorácicotrópico, é o sinal que inicia a liberação do precursor esteroide ecdisona das glândulas protorácicas (THUMMEL, 2001).

De modo, que os ecdisteroides são os hormônios esteroides coordenam as principais transições de desenvolvimento iniciais, e tanto aumentos quanto diminuições nos níveis circulantes destes hormônios são importantes para o tempo dos eventos de desenvolvimento (THUMMEL, 2001). Assim, a primeira camada de cutícula é depositada enquanto os níveis de ecdisteroides estão em níveis altos. Já a eliminação da cutícula velha (fase intermuda) ocorre quando as concentrações de ecdisteroides atingem níveis mais baixos (CHARLES, 2010). No final do processo,

ocorre o aumento dos níveis dos hormônios juvenis que são aqueles que previnem a metamorfose da cutícula (WILLIS, 1996). Estes hormônios regulam, direta ou indiretamente, a expressão dos genes das proteínas das cutículas em específicos estágios da metamorfose (CHARLES, 2010).

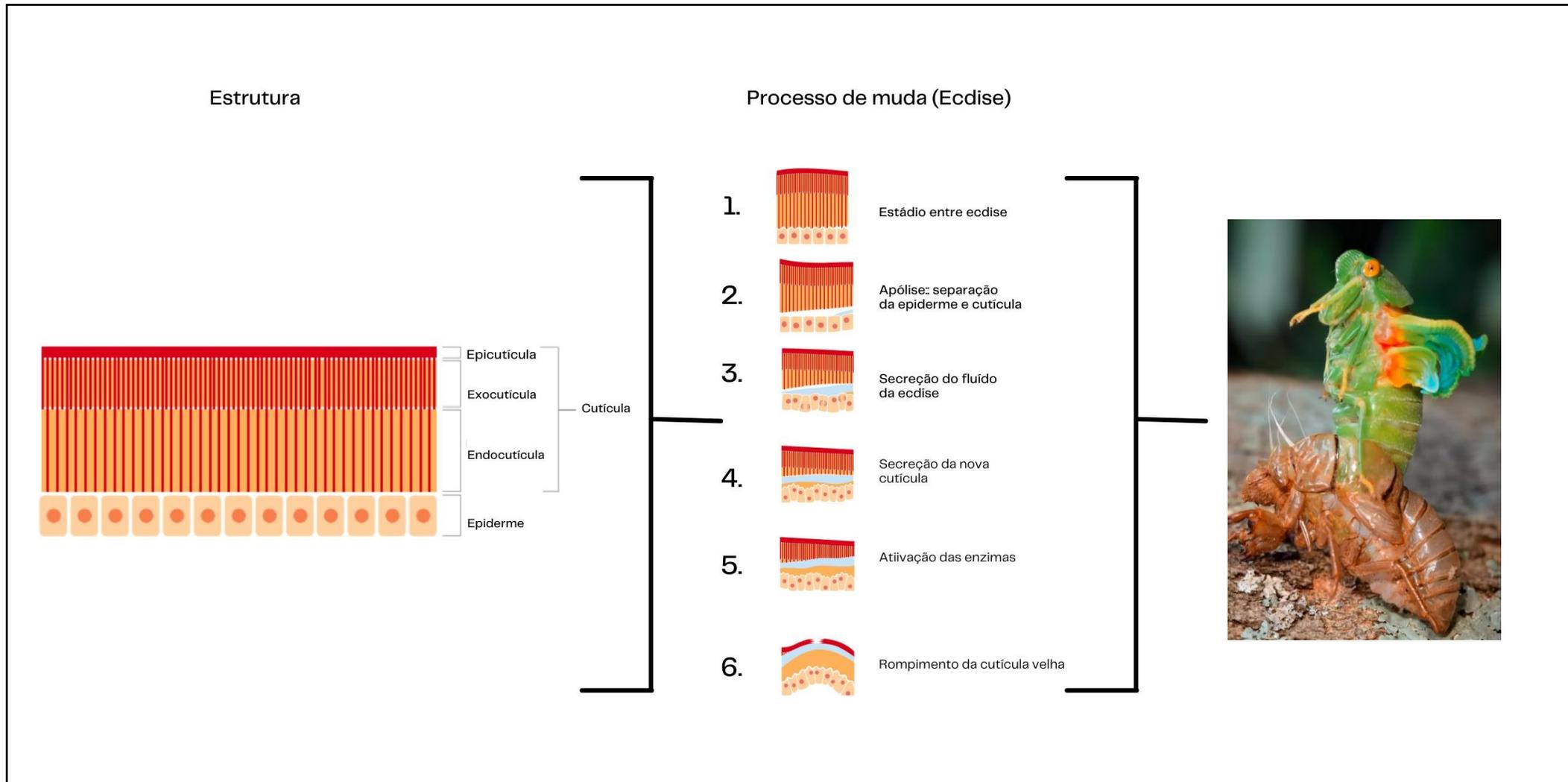


Figura 3 - (a) Etapas do processo de muda (Ecdise), adaptado de SALGADO, V. L. (2013) e (b) imagem do processo de ecdise de cigarra por PHILIP, T (2017).

2.2.2. Modo de ação dos inseticidas e receptores da ecdisona

A descoberta de novas moléculas com formulação química melhor vem sendo pesquisada e desenvolvida cada vez mais, devido a necessidade de encontrar características como maior segurança para o meio ambiente e seletividade, ou seja, é necessário encontrar compostos ativos somente para o inseto praga e não tóxicos para os insetos benéficos, mamíferos e outros organismos não-alvos. Isso é essencial para o futuro da agricultura, pois até os dias atuais parte dos inseticidas ainda são tóxicos e são aqueles com mecanismos descobertos há mais de 50 anos, com difícil degradação ambiental (DELANEY et al., 2006). Assim, descobertas de novas moléculas que se mostrem ambientalmente aceitáveis, inseticidas planejados de forma racional e o uso de novos métodos de controle de pragas são alguns dos grandes desafios atuais enfrentados pelo Manejo Integrado de pragas (MIP) neste século (ZOTTI et al., 2013).

O crescimento do inseto é regulado por eventos orquestrados de ecdisteroides e suas proteínas receptoras. Sendo assim, prevê-se que tanto agonistas quanto antagonistas desse receptor tenham a capacidade de perturbar o crescimento normal do inseto, e com base nessa afirmação, esses seriam bons candidatos a inseticidas (HARADA et al., 2011).

Os reguladores de crescimento são considerados mais seguros para os inimigos naturais, insetos benéficos e organismos não-alvo, quando comparados aos inseticidas neurotóxicos de largo espectro de ação, em função de sua especificidade nos alvos bioquímicos e seu planejamento racional (DARVAS; POLGAR, 1998). O alto nível de segurança se dá em partes à ausência de receptores de ecdisona em mamíferos (BILLAS, et al., 2003). No entanto, em que pese a suposta segurança, estes receptores devem ser testados e avaliados, bem como sua compatibilidade, com testes de seletividade para o registro de novos produtos para sua utilização junto ao Manejo Integrado de pragas (BUENO; FREITAS, 2004).

Os repetidos processos de muda e metamorfose são eventos de dois hormônios não peptídicos periféricos, a muda e os hormônios juvenis. 20-hidroxiecdisona (20E) é o hormônio de muda para a grande maioria dos insetos, sendo a ação destes mediada pelo complexo proteico heterodimérico do receptor de ecdisteróides (EcR) e o ultraspiracle (USP). A ação desse complexo irá ativar a tradução dos genes associados à muda após a ligação com a molécula de ligante a

EcRs (HARADA et. al, 2011), sendo este ligante um campo a ser pesquisado e explorado para o desenvolvimento de novas moléculas inseticidas.

Assim sendo, afora o seu papel natural de controle da expressão de redes de genes no desenvolvimento e reprodução de artrópodes, EcR também podem ser empregados como prováveis alvos para inseticidas mais ecológicos e no controle de transgenes (por exemplo, genes terapêuticos) (BILLAS, et al., 2003). Os primeiros antagonistas não esteroides atuantes no receptor de ecdisteroides são as dibenzohidrazinas (DBHs) (WING; ALLER, 1990). Apesar do fato já citado de que muitos insetos utilizam a 20-Hidroxiecdisonio como o hormônio natural de muda, moléculas ligantes como tebufenozida, methoxifenozida, cloromafenozida e halofenozida (todas dibenzohidrazinas) são inativas para *Bemisia tabaci* (Genn, 1889) (Hemiptera: Aleyrodidae), ou seja, estes seriam inativos para Hemiptera (GRAHAM et. al, 2007) enquanto alguns são extremamente tóxicos para Lepidopteras. Assim, a suscetibilidade para inseticidas não esteroides agonistas do ecdisonio mostra determinada especificidade para espécies, e por esse motivo, os inseticidas com moléculas dibenzohidrazinas podem ou não ser usados para o controle de insetos-praga. Destaca-se, ainda, que as DBHs se mostraram inativas para polinizadores, predadores e parasitóides (RETNAKARAN et al., 2003), característica muito importante ao Manejo Integrado de Pragas (MIP).

3. Capítulo II: Artigo 1 – Triagem virtual de moléculas inseticidas: a busca de potenciais análogos de agonistas da ecdisona

Resumo do Capítulo

A técnica de Triagem Virtual (*Virtual Screening*) mostra ser uma possibilidade eficiente, dinâmica e bastante útil como instrumento para descoberta de novos inseticidas. Isso porque utiliza ferramentas computacionais dentro da biologia molecular e estrutural para descobrir moléculas potenciais que se ligam na estrutura desejada. Somente essa técnica gerou dados precisos e reduziu consideravelmente o tempo necessário para encontrar o ligante. Os objetivos dessa pesquisa consistiram em utilizar a tecnologia de triagem virtual para identificar moléculas com alta possibilidade de ligação no sítio ativo do receptor da ecdisona (EcR). Este receptor nuclear é um importante alvo para a identificação de inseticidas seletivos, na medida que apresenta as características de ação inibitória seletiva, baixa resistência e se trata de um receptor que não é encontrado em mamíferos. A presente pesquisa foi desenvolvida em duas etapas principais, a triagem de análogos de moléculas referência e o *docking molecular* da proteína nos ligantes selecionados. Na primeira, ferramentas computacionais foram programadas para encontrar potenciais ligantes de acordo com características específicas (químicas e estruturais) que os tornam análogos das moléculas referência e, na segunda fase, os ligantes encontrados na triagem virtual foram acoplados na proteína e após, realizadas as modelagens tridimensionais do composto. O *screening* de um milhão, duzentos e setenta e seis mil, quatrocentos e setenta e seis (1.276.476) e cento e quarenta milhões, setecentos e quarenta e quatro mil, setecentos e sessenta e oito (140.744.768) conformers, resultou em seis mil, trezentos e setenta e cinco (6.375) análogos (hits) de Tebufenozida e deste número, quinhentos e cinquenta e quatro (554) docados na proteína e três (3) analisados suas propriedades e comparados a inseticidas. Para os análogos do agonista sintético, foram quatro mil, oitocentos e três análogos (4.803) hits e quinhentos e vinte docados e três (3) analisados. E, por fim, a triagem dos compostos comparados ao 20HE resultaram em vinte (20) hits e dezessete (17) docados, com três (3) compostos analisados. Os resultados mostram que o composto C23H19N3O4 (código: AASDoc1), demonstrou maior destaque somando potencial de similaridade química e estrutural com um inseticida do grupo estudado, porém, sugere-se que futuras avaliações biológicas não devem descartar os demais compostos que apresentaram bons resultados químicos e estruturais

Palavras-chave: Bioinformática. Triagem virtual. Acoplamento molecular. Receptores de nucleares. Descoberta de moléculas inseticidas.

3.1. Introdução

A triagem virtual *in silico* (*virtual screening* ou triagem virtual de alto rendimento -HTVS-) é uma técnica que tem se destacado como um excelente complemento para outras experimentações que se tornam mais caras e demoradas (BERRY, M. et al; 2015). Esta técnica pode ser baseada na estrutura (*structure-based*) ou em moléculas já conhecidas (*ligand-based*) (COURNIA et al., 2020) e a aplicabilidade de alguma dessas metodologias irá depender das informações disponíveis para a realização da triagem.

A disponibilidade de compostos e grandes bibliotecas para a triagem e ensaio virtual cresceu nos últimos anos e, com bibliotecas maiores e mais diversificadas, que superam um bilhão de compostos, ocorre a necessidade de aprimorar cada vez mais os processos de ancoragem molecular, com explorações eficientes do processo com *docking* (YANG, Y. et al; 2021). Este tipo de experimentação pressupõe que se pode descobrir o melhor posicionamento para cada composto na proteína (BASKARAN, S. et al; 2021). E para que esse fato ocorra com sucesso, deve-se atender a diferentes desafios da triagem, como a comparação de ligantes com diferentes estruturas, propriedades químicas e as conformações das moléculas analisadas (PEROLAE., et al, 2005).

Para a realização de uma boa triagem é necessária uma pontuação com critérios bem definidos, com a intenção de ranquear os melhores compostos, fornecendo uma classificação quantitativa confiável e a obtenção correta das coordenadas espaciais, para descrever a colocação correta do ligante (WANG, J. et al, 2001).

O acoplamento molecular (*docking*) foi descrito a primeira vez em 1982 (KUNTZ et al., 1982) e hoje é a ideia central da triagem virtual baseada na estrutura (BERRY. M et al., 2015). As funções de pontuações para o encaixe das moléculas são capazes de filtrar e classificar grandes bibliotecas de compostos, nas quais as melhores classificadas correspondem a um potencial composto (PHATAK et al., 2009).

Ensaio de acoplamento *in silico* possuem limitações e imperfeições nas pontuações, podendo gerar falsos positivos e falsos negativos (LILL, 2011, ALVAREZ, 2004). Tal ocorre, principalmente, quando não se estima a natureza dinâmica das estruturas biológicas, motivo pelo qual é necessário pensar em

métodos que compensem essa situação.

Um bom método de desempenho é via enumeração exaustiva de todas as torções rotativas, o qual gera um grande conjunto de conformações, reproduzindo um ambiente não estático. Este modelo de predição apresenta um bom desempenho na reprodução de conformações (HAWKINS, P. et, al. 2010), demonstrando a importância da preparação dos bancos de dados e diminuindo o erro dos falsos positivos e falsos negativos.

A identificação de compostos ativos por comparação de forma pode ser também uma ferramenta importante na experimentação *in silico*, demonstrando uma forma consistente e algumas vezes superior a abordagens centradas em encaixe (HAWKINS, P. et, al. 2007).

Com base no exposto, buscou-se utilizar a união das ferramentas computacionais de torção rotativa, comparação de forma, encaixe da proteína e ligante e análise química do composto para a realização da descoberta de prováveis novas moléculas que se demostrem análogos de compostos agonistas da ecdisona.

3.2. Materiais e métodos

3.2.1. Obtenção e preparação das bibliotecas de compostos e do receptor

As moléculas foram obtidas por meio de uma seleção aleatória, com o intuito de permitir uma maior diversidade estrutural, por meio de sites especializados para a obtenção destas bibliotecas de compostos, somando o total de trinta e quatro bibliotecas, das seguintes empresas: APExBIO, Asinex, CHEMDIV, Life Chemicals, Selleckchem.

A preparação das moléculas, obtidas por meio das bibliotecas de compostos, foi feita a partir do programa OMEGA da *OpenEye Scientific*. Nesse software, foram realizadas todas as possíveis conformações para as moléculas bioativas (Figura 4), preparo que foi essencial para o trabalho, devido ao fato de que moléculas não estão estáticas na natureza, elas se mantêm em movimento e em ambiente aquoso. Desta forma, o preparo do OMEGA proporcionou essa simulação, com alta velocidade e confiabilidade, permitindo maiores possibilidades de acoplamento entre o receptor e o ligante, uma vez que cada molécula se mostrou

com uma média superior a cem possíveis conformações, antecedendo o processo de triagem virtual, ou *screening* dos compostos.

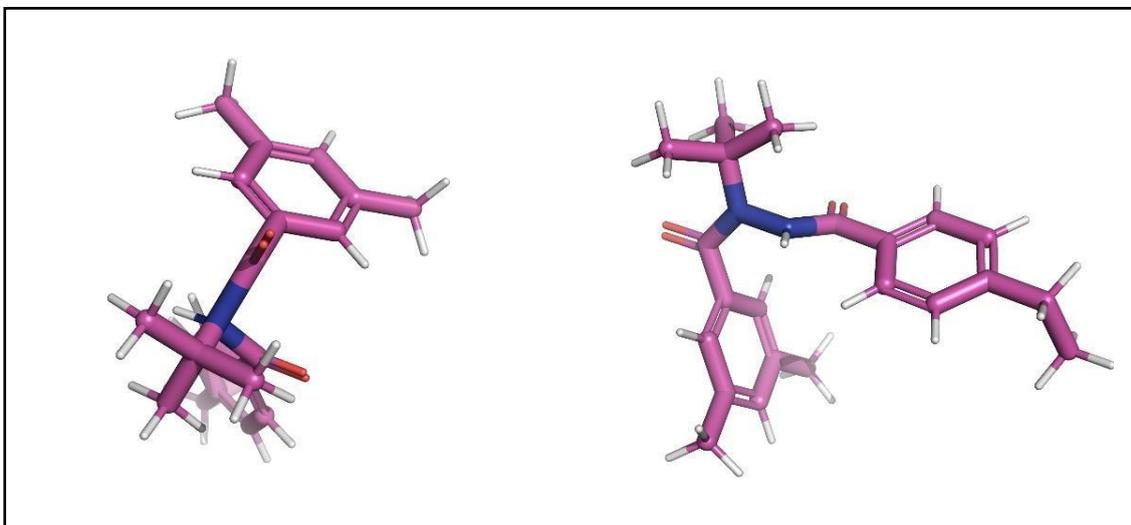


Figura 4 - Molécula Tebufenozida sendo representada em diferentes conformações.

O organismo alvo para a busca de novas moléculas inseticidas é a *Spodoptera frugiperda*, e o receptor nuclear empregado para a realização dos experimentos *in silico*, foi o receptor da Ecdisona (EcR) da ordem Lepidoptera, no organismo *Chloridea virescens* (Lepidoptera: Noctuidae), registrado no *Protein Data Bank* (PDB) com o antigo gênero *Heliothis*. Para a obtenção desta estrutura, a macromolécula foi adquirida no banco PDB, validada pelo método experimental de difração de raios X, por Billas e colaboradores (2003). O conjunto biológico da estrutura de cristal dos domínios de ligação ao ligante do heterodímero EcR/USP ligado ao agonista sintético BY06830, foram registrados como “1R20” no PDB.

O Receptor da Ecdisona foi preparado no momento que antecedeu o acoplamento das moléculas no receptor, ou seja, o *Docking*. Essa preparação foi feita com o programa *MakeReceptor*, também da *OpenEye Scientific*, que é um utilitário gráfico que foi programado para editar o receptor, retirando a cadeia que não se mostrou interessante para *docking* e deixando a bolsa de ligação em evidência.

3.2.2. Triagem virtual e acoplamento das moléculas

Após a preparação das moléculas com o OMEGA, o procedimento foi conduzido pelos pacotes *OpenEye Scientific* com os softwares ROCS para a Triagem Virtual de análogos dos compostos referência selecionados e FRED para o acoplamento das moléculas (Figura 5), ambos aplicados em linha de comando no sistema operacional Linux. A metodologia foi conduzida com o software ROCS de acordo com Rush et al., 2005; Hawkins et al., 2010 e FRED como o desenvolvido por McGann 2003 e 2011.

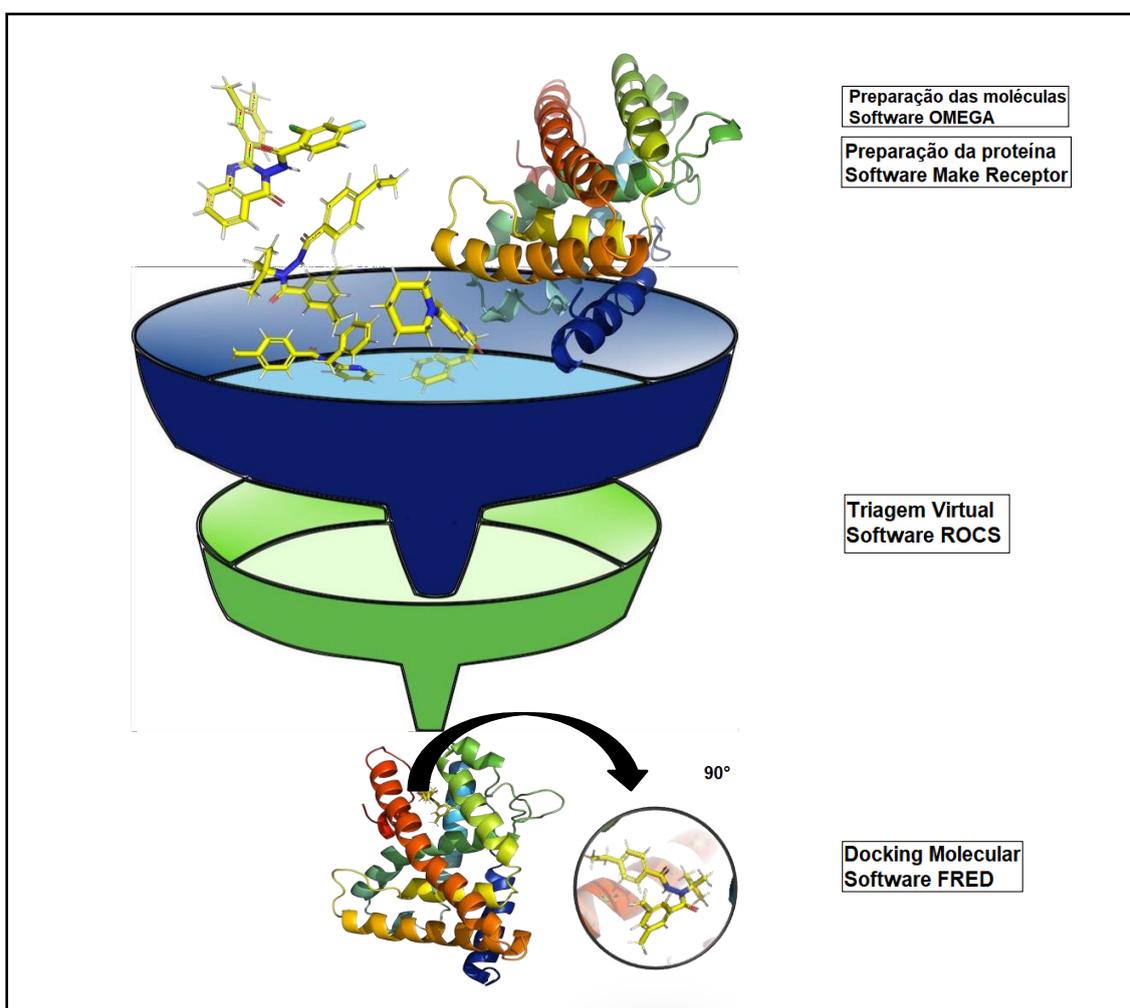


Figura 5 - Representação da Triagem Virtual de moléculas, onde as peneiras representam a seleção de moléculas (triagem) e no final a molécula selecionada acoplada com a proteína (docking).

Para a Triagem Virtual, foram utilizadas três moléculas como referência: a molécula inseticida agonista da ecdisona Tebufenozida, o agonista sintético BY06830 e o hormônio 20-hidroxiecdisona (20HE) (Figura 6). Assim, todas as moléculas e suas determinadas conformações foram analisadas em três vezes, buscando análogos para as três moléculas referência, sendo realizada uma comparação de formas, gerando um resultado baseado no coeficiente de Tanimoto, onde Tanimoto combo será igual à soma de Tanimoto Color (semelhança química entre as moléculas referência com as moléculas analisadas) e Tanimoto Shape (semelhança estrutural entre as moléculas referência e as moléculas analisadas), dentre outros parâmetros.

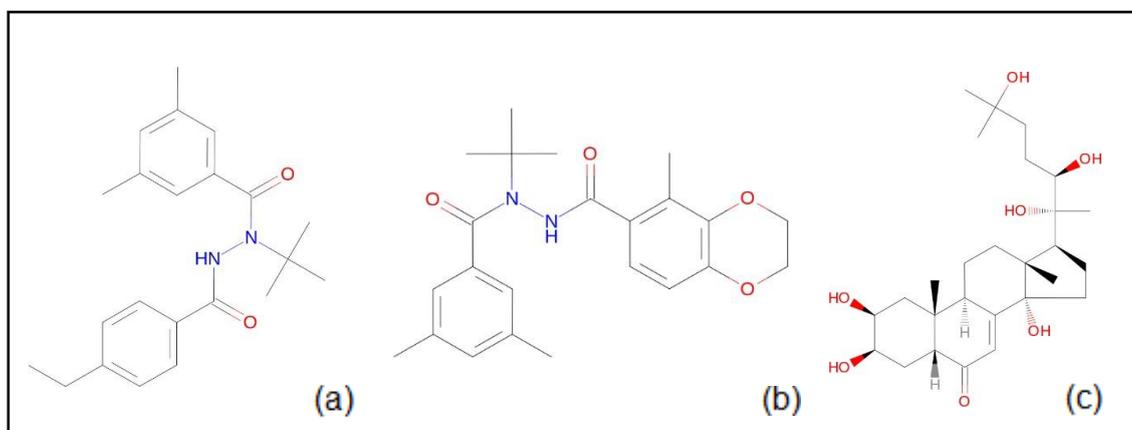


Figura 6 - Moléculas referência utilizadas: Tebufenozida (a), BY06830 (b), 20HE (c). Representação em 2D.

ROCS foi a ferramenta escolhida para a realização da triagem virtual, que serviu na etapa de identificação dos compostos que poderão se mostrar ativos, uma vez que as moléculas referência são aquelas que acoplam no receptor, logo, presume-se que todas aquelas semelhantes, ou seja, seus análogos, terão maiores chances que realizarem o acoplamento também. Esse software realizou sobreposições com base na estrutura 3D das moléculas com todas as conformações possíveis. A consulta foi realizada com a combinação dos parâmetros pré-definidos de geometria e similaridade química (Tanimoto), que aumentaram a probabilidade de acerto.

Após a comparação entre formas, o programa ROCS gerou bancos estruturais que considerou semelhante, ou seja, os chamados *hits*. Utilizando o arquivo dos *hits* gerados, que são todas as moléculas que se mostraram semelhantes de acordo com a soma entre geometria e similaridade química, foi

realizado o próximo passo, consistente no procedimento de *docking* ou acoplamento molecular.

O software FRED foi utilizado no processo de acoplamento molecular e assim foi testado o encaixe entre o receptor e as moléculas que passaram pela triagem. Todas as moléculas que encaixaram no receptor receberam pontuações com base no parâmetro *Chemgauss4*, realizando uma análise de todas as possíveis posições da proteína com os ligantes. Além das pontuações com base no parâmetro também foi gerado uma análise de *docking* por átomo e a comparação entre as moléculas dentro de cada biblioteca, visando uma melhor seleção, além dos resíduos (aminoácidos) interagindo. A referência para este programa será o agonista sintético BY06830, uma vez que é ele que está ligado por cristalografia de raio X na macromolécula que foi adquirida no banco PDB.

Por fim, foi realizado uma nova análise e com os compostos selecionados de acordo com as melhores pontuações de *docking*, sendo avaliados de acordo com 'Lipinski's Rule of Five' (TICE, 2001). As propriedades físico-químicas foram calculadas por *SwissADME*.

3.2.3. Modelagem tridimensional

As representações gráficas e modelagens tridimensionais foram a última etapa da fase de experimentação *in silico*, a qual foi construída utilizando os pacotes Discovery Studio e PyMol, gerados por meio das coordenadas atômicas dos complexos do EcR de *Chloridea virescens* em complexo com ligantes selecionados após o *screening* e *docking* molecular. As modelagens em 3D foram representadas em diferentes formas para uma melhor visualização da proteína (Figura 7) e performance de acoplamento.

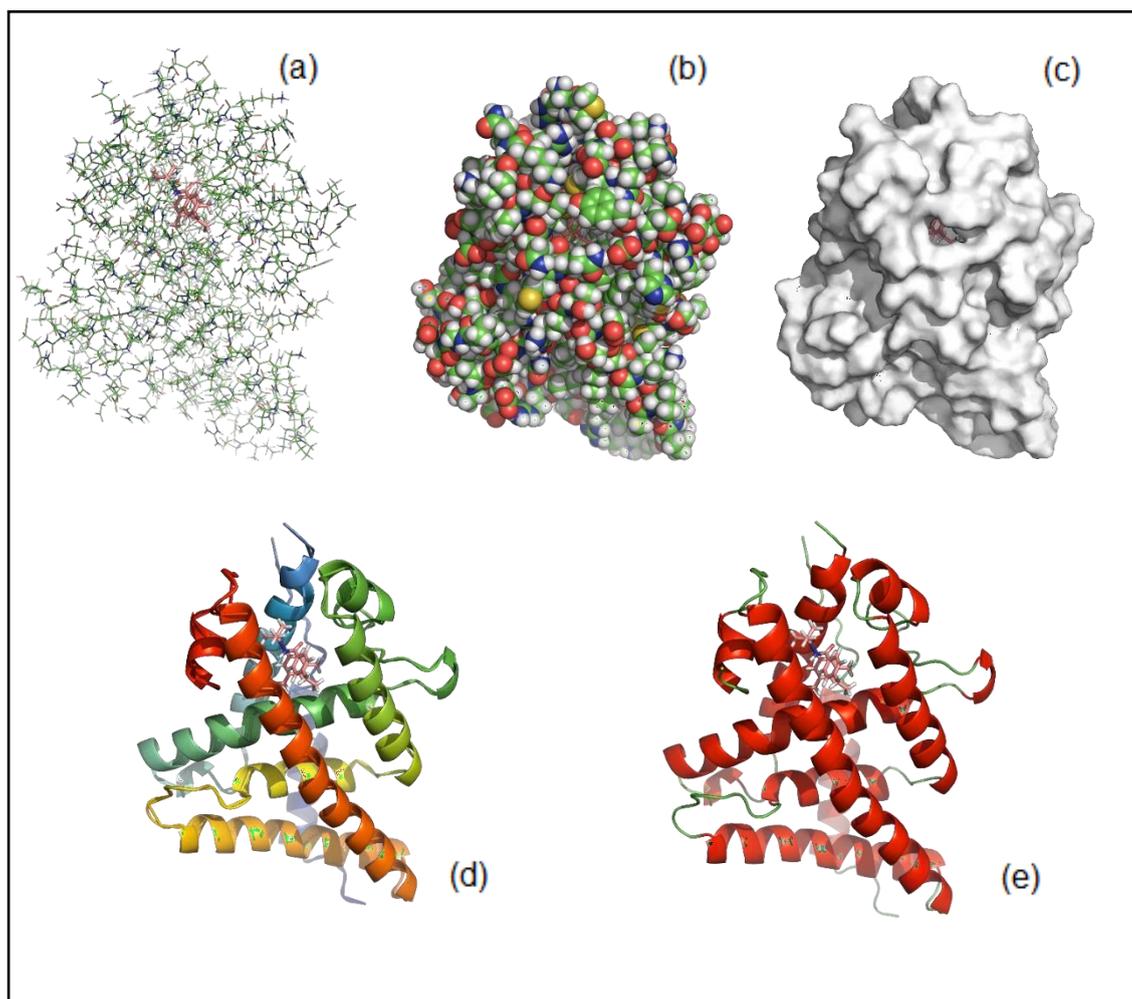


Figura 7 - Diferentes formas de representar a proteína com a molécula acoplada, sendo da esquerda para direita: (a) em linhas com coloração em de acordo com o elemento – carbono emverde, hidrogênio cinza, nitrogênio azul, oxigênio vermelho, etc.*-, (b) em esferas com coloração por elemento*, (c) em superfície, (d) em fita com coloração por espectro e (e) com fita e colorindo a hélice em vermelho e o ciclo em verde.

3.3. Resultados

3.3.1. Validação do preparo antecedendo o *screening* e diminuição significativa do número de compostos a serem testados *in vitro*

Na fase de preparação das moléculas, o número analisado de conformações obteve um aumento em decorrência da preparação inicial, porém, sequencialmente, todos os softwares auxiliaram na triagem e seleção do número total, demonstrando uma diminuição do número de compostos e auxiliando na orientação das moléculas a serem selecionadas para futuros testes. Para as trinta e quatro bibliotecas de compostos analisadas com um milhão duzentos e setenta e seis mil quatrocentos e trinta e sete compostos, o programa OMEGA gerou cento e quarenta

milhões, setecentos e quarenta e quatro mil, setecentos e sessenta e oito conformações diferentes, obtendo diferentes resultados no programa ROCS e FRED para as três moléculas ligantes referência: Tebufenozida, agonista sintético BY06830 e 20HE (Tabela 2).

Tabela 2 - Dados da Triagem Virtual de moléculas semelhantes à Tebufenozida, ao agonista sintético e à 20HE.

Moléculas Referência	Fase	Software utilizado	Informações geradas	Resultados
Tebufenozida	Escolha das bibliotecas	-	Bibliotecas – nº total de Moléculas	34 bibliotecas - 1.276.437 moléculas
	Preparação das moléculas	Omega	Conformações	140.744.768
	<i>Screening virtual</i>	ROCS	Hits (análogos)	6.375
	<i>Docking molecular</i>	FRED	Dockados (acoplados) na proteína (ou receptor)	554
Agonista SintéticoBY06830	Escolha das bibliotecas	-	Bibliotecas – nº total de Moléculas	34 bibliotecas - 1.276.437 moléculas
	Preparação das moléculas	Omega	Conformações	140.744.768
	<i>Screening virtual</i>	ROCS	Hits (análogos)	4.803
	<i>Docking molecular</i>	FRED	Dockados (acoplados) na proteína (ou receptor)	520
0HE	Escolha das bibliotecas	-	Bibliotecas – nº total de Moléculas	34 bibliotecas - 1.276.437 moléculas
	Preparação das moléculas	Omega	Conformações	140.744.768
	<i>Screening virtual</i>	ROCS	Hits (análogos)	20
	<i>Docking molecular</i>	FRED	Dockados (acoplados) na proteína (ou receptor)	17

Os dados foram gerados nos softwares Omega, Rocs e Fred. A média de conformações por molécula ficou em torno de 158,5.

3.3.2. Coeficientes de *docking* entre a proteína e o ligante que demonstraram melhores acoplamentos

Foi selecionado os três melhores ligantes para análogos de Tebufenozida (apêndice 1,2 e 3), agonista sintético BY06830 (apêndice 4,5 e 6) e 20HE (apêndice 7,8,9) e após comparados entre si onde o composto C₂₃H₁₉N₃O₄ (código: AASDoc1), se mostrou superior na comparação geral, demonstrando um coeficiente de Chemgauss (FRED) de -16,88. Dos análogos do Tebufenozida os compostos C₂₂H₁₅F₂N₃O₃S (código: TDoc1; SMILE: c1cc(cnc1)CN(c1nc2c(cc(cc2s1)F)F)C(=O)c1ccc2c(c1)OCCO2), C₂₃H₁₇CLFN₃O₂ (código: ATDoc2; SMILE: c1ccc2c(c1)c(=O)n(c(c1ccc(c(c1)C)C)n2)NC(=O)c1ccc(cc1Cl)F) e C₁₇H₁₄F₃N₃O₃S (código: ATDoc3; SMILE: c1cc(CC(=O)Nc2c(c3ccc(cc3)C(F)(F)F)n[nH]c2C)sc1) mostraram os melhores valores de docking. Para os análogos de 20HE, os compostos C₂₂H₂₈O₅ (código: A20HEDoc1; SMILE: C1=C[C@@]2(C(=CC1=O)CC[C@@H]1[C@@H]2C(=O)C[C@@]2([C@H]1C[C@@H]([C@@]2(C(=O)CO)O)C)C), C₂₁H₃₀O₄ (código: A20HEDoc2, SMILE: C1=C2CC[C@@H]3[C@H](CC[C@@]4([C@H]3CC[C@@]4(C(=O)CO)O)C)[C@]2(CCC1=O)CC21H30O4) e C₂₂H₂₉FO₄ (Código: A20HEDoc3; SMILE: C1=C[C@@]2(C(=CC1=O)CC[C@H]1[C@@H]3C[C@H]([C@H](C(=O)CO)[C@]3(C[C@@H]([C@]21F)O)C)C)) obtiveram as melhores pontuações. E para os compostos análogos do agonista sintético BY06830, as moléculas C₂₃H₁₉N₃O₄ (código: AASDoc1; SMILE: c1cc(c(cc1c1c(/N=C/c2ccc3c(c2)OCO3)n2c(cc(cc2)C)n1)OC)O), C₂₀H₁₅N₃O₃S (código: AASDoc1; SMILE: c1ccc(cc1)n1c(=O)c2cccc2nc1SCc1ccncc1) e C₂₃H₁₇N₃O₄S (código: AASDoc1; SMILE: c1cc(cnc1)CN(c1nc2cc3c(cc2s1)OCO3)C(=O)c1ccc(cc1)C(=O)C) mostraram melhores valores de acoplamentos. Estes compostos, que se mostraram as melhores moléculas acopladas na proteína, estão representados na Figura 8 com corte em 3D com a proteína representada em surface e as moléculas em *sticks*.

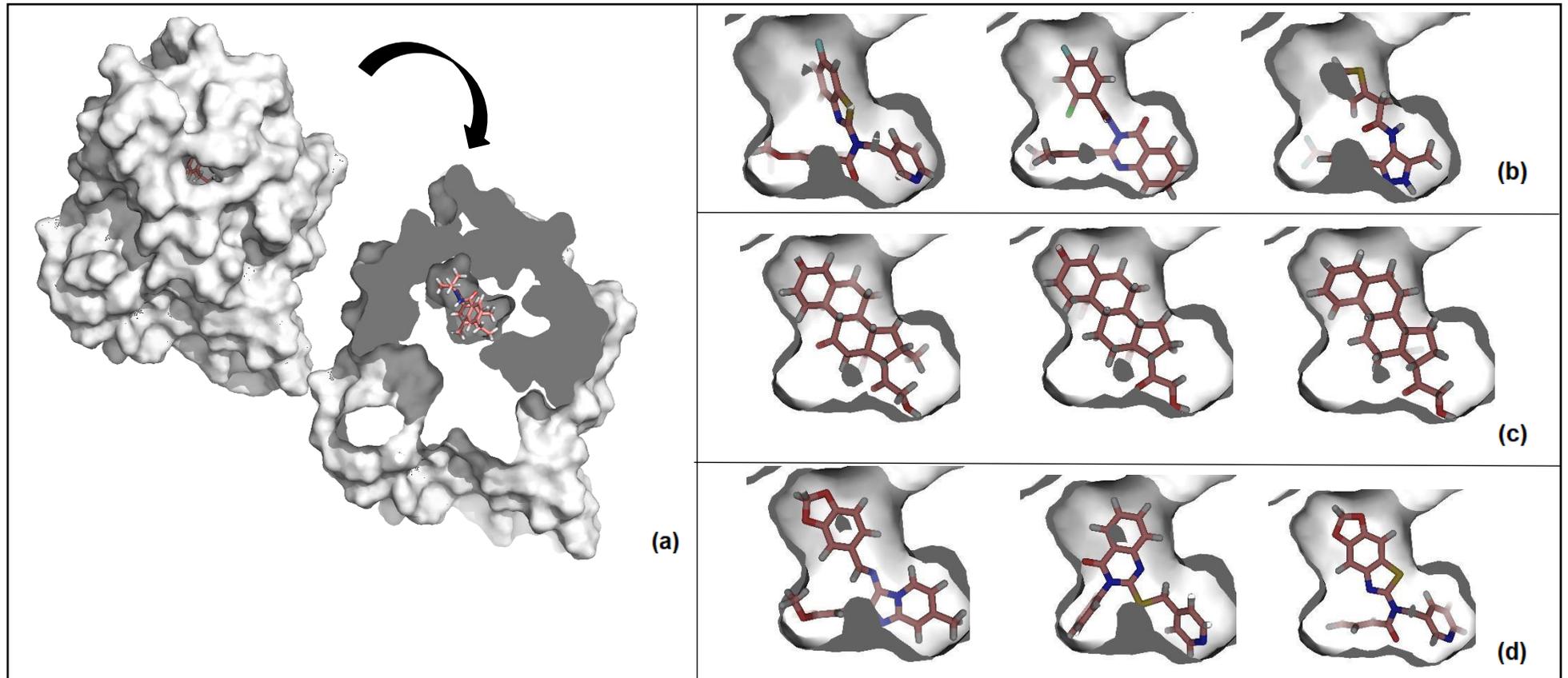


Figura 8 - Representação da proteína em Surface com a molécula acoplada e em corte transversal (a), bem como o corte transversal com zoom no ligante da bolsa de ligação mostrando as moléculas análogas acopladas no receptor sendo (b) análogos da Tebufenozida, (c) análogos do 20HE e (d) análogos do ligante de 1r20, com as melhores pontuações de *docking* da esquerda para a direita, dentro de cada grupo.

Na análise dos resíduos, ou seja, dos aminoácidos presentes na ligação, realizado pelo programa FRED, podemos perceber que, ao comparar os doadores e receptores de hidrogênio presentes nas ligações das moléculas triadas como potenciais, com os inseticidas “receptores de ecdisteroides” liberados no Brasil, obteve-se semelhança nos aminoácidos envolvidos (Figura 9), além de semelhança estrutural e do local de ligação da proteína com a molécula (Figura 10).

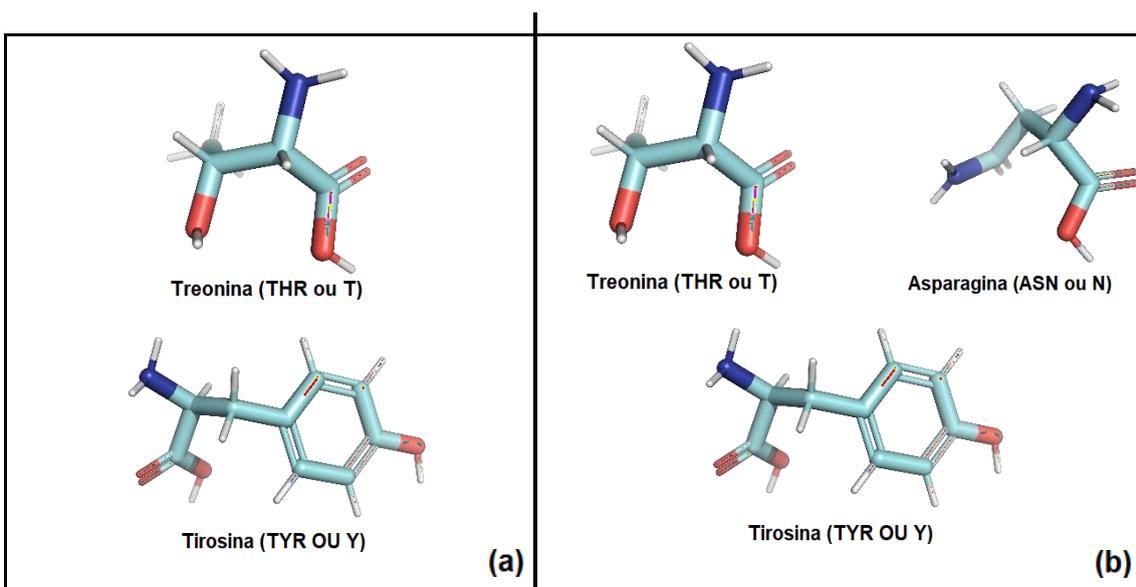


Figura 9 - (a) A estrutura molecular dos aminoácidos envolvidos ligação nas principais moléculas identificadas e (b) estrutura molecular dos aminoácidos envolvidos na ligação nos inseticidas Tebufenozide, Metoxyfenozide, Cromafenozide, Halofenozide, segundo a simulação virtual de docking, gerados pelo software FRED.

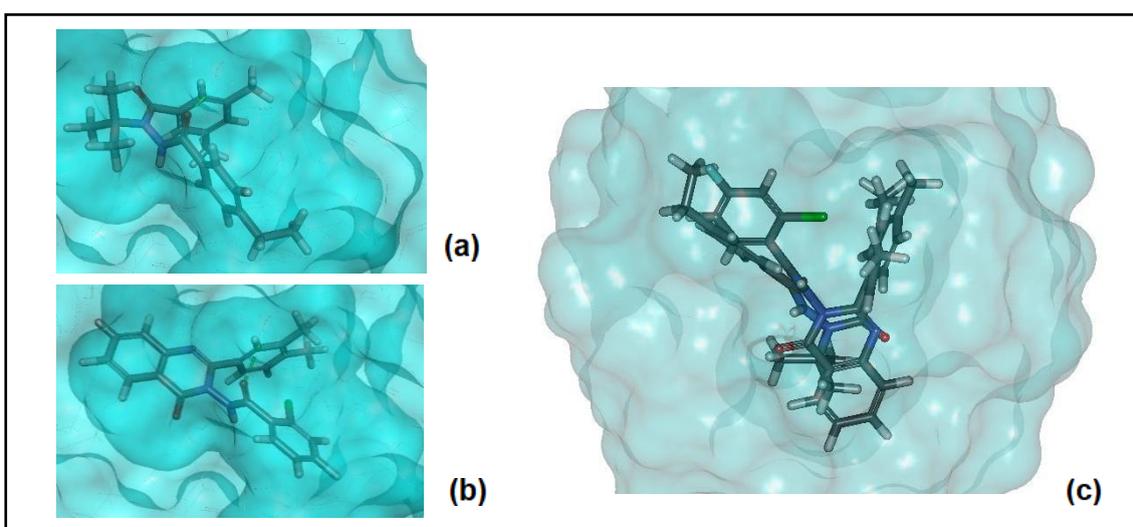
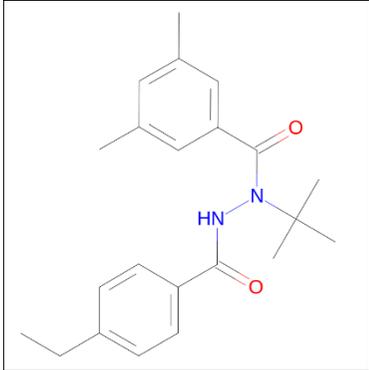
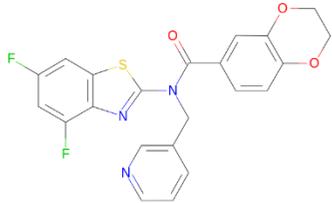
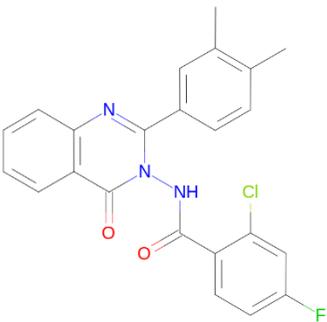
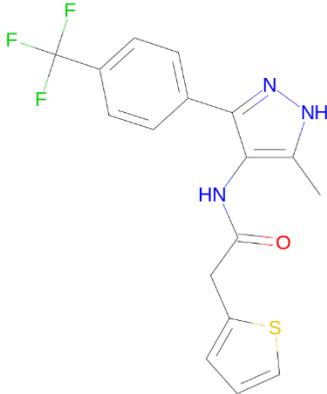


Figura 10 - No detalhe, a bolsa de ligação da proteína com Tebufenozide (a) e uma molécula análoga (b) e duas moléculas análogas na bolsa de ligação da proteína (c).

3.3.3. Propriedades físico-químicas (regra dos 5), coeficiente de Tanimoto (ROCS) e ChemGauss4 (FRED)

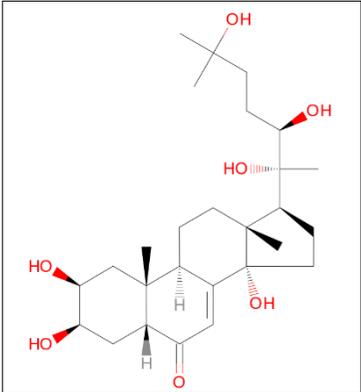
As melhores pontuações de acoplamento foram triadas, de acordo com a regra dos 5, que demonstrou que os melhores compostos triados estavam dentro dos valores de das regras (Tabela 3, 4, 5). Um análogo de Tebufenozida cometeu uma violação, mostrando-se superior ao valor do coeficiente de partição octanol/água médio para moléculas inseticidas e todos os inseticidas liberados no Brasil, que agem no receptor de ecdisona os “agonistas de receptores de ecdisteroides”, estão dentro dos valores da regra dos 5 (Tabela 6).

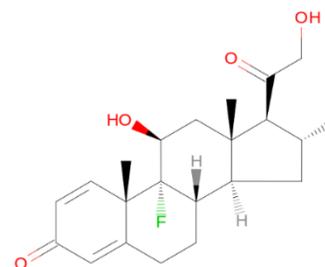
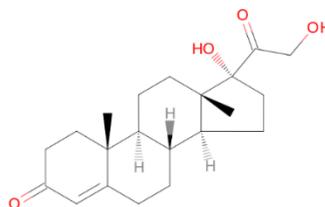
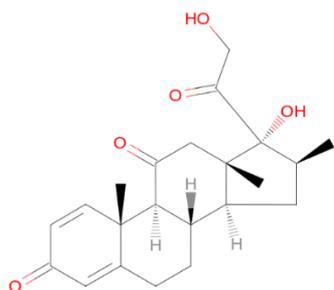
Tabela 2 - Análogos Tebufenozide com melhores pontuações de docking.

Parâmetro	1º molécula (cód: ATDoc1)	2º molécula (cód: ATDoc2)	3º molécula (cód: ATDoc3)	Imagem Modelo: Tebufenozide
Peso molecular	439,43	421,85	365,37	
mLogP	2,58	5,37	2,91	
Nº doadores H	0	1	2	
Nº aceitadores (receptores) H	7	4	5	
Ligações rotativas	5	4	6	
Coefficiente Chemgauss (FRED)	-16,55	-16,37	-16,28	
Coefficiente Tanimoto combo (ROCS)	0,893	1,245	1,074	
Fórmula	C22H15F2N3O3S	C23H17CLFN3O2	C17H14F3N3OS	
Imagem 2D				

As propriedades físico-químicas foram calculadas por SwissADME e sinalizadas em vermelho as violações dos índices médios para inseticida de TICE (2001) e em azul, os índices do coeficiente de Tanimoto menor que 1,00.

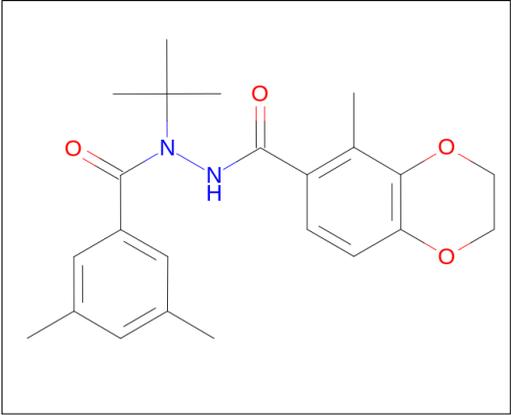
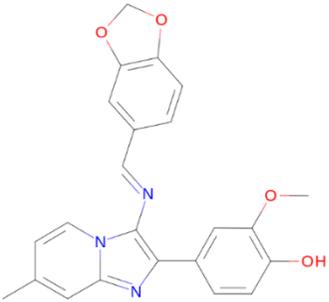
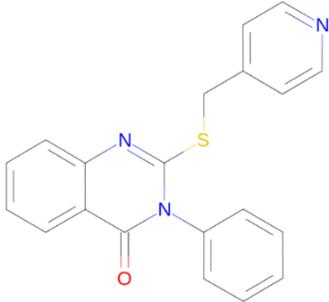
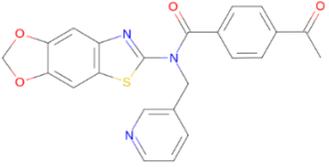
Tabela 3 - Análogos 20HE com melhores pontuações de docking

Parâmetro	1º molécula (cód: A20HEDoc1)	2º molécula (cód: A20HEDoc2)	3º molécula (cód: A20HEDoc3)	Imagem modelo: 20HE
Peso molecular	372,45	346,461	376,462	
mLogP	1,43	2,22	2,35	
Nº doadores H	2	2	2	
Nº aceitadores (receptores) H	5	4	5	
Ligações rotativas	2	2	2	
Coefficiente Chemgauss (FRED)	-13,43	-12,58	-12,08	
Coefficiente Tanimoto (combo)	1,006	1,012	1,023	
Fórmula	C22H28O5	C21H30O4	C22H29FO4	
Imagem 2D				



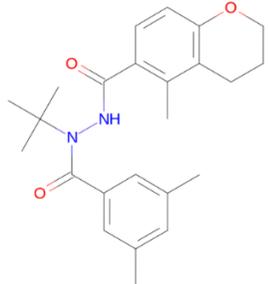
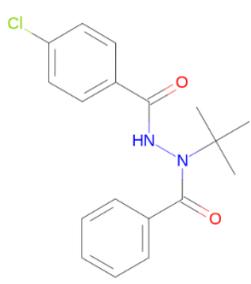
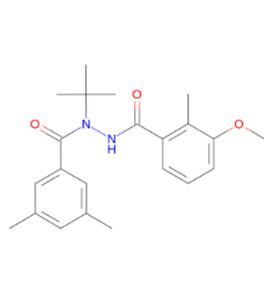
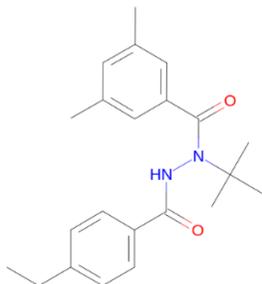
As propriedades físico-químicas foram calculadas por SwissADME e sinalizadas em vermelho as violações dos índices médios para inseticida de TICE (2001) e em azul, os índices do coeficiente de Tanimoto menor que 1,00.

Tabela 4 - Análogos do Agonista Sintético BY06830 com melhores pontuações de docking.

Parâmetro	1º molécula (cód: AASDoc1)	2º molécula (cód: AASDoc2)	3º molécula (cód: AASDoc3)	Imagem modelo: Agonista Sintético BY06830
Peso molecular	401,41	345,42	431,46	
mLogP	2,73	3,13	1,58	
Nº doadores H	1	0	0	
Nº aceitadores (receptores) H	6	3	6	
Ligações rotativas	4	4	6	
Coefficiente Chemgauss (FRED)	-16,88	-16,79	-16,66	
Coefficiente Tanimoto combo (ROCS)	1,034	0,931	1,057	
Fórmula	C23H19N3O4	C20H15N3OS	C23H17N3O4S	
Imagem 2D				

As propriedades físico-químicas foram calculadas por SwissADME e sinalizadas em vermelho as violações dos índices médios para inseticida de TICE (2001) e em azul, os índices do coeficiente de Tanimoto menor que 1,00.

Tabela 5 - Inseticidas que agem no receptor de ecdisona os “agonistas de receptores de ecdisteroides” de acordo com IRAC 2022.

Parâmetro	Cromafenozone	Halofenozide*	Metoxyfenozide	Tebufenozide	Intervalo	Média
Peso molecular	394,51	330,81	368,47	352,47	330,81~394,51	361,56
mLogP	4,05	4,17	4,01	4,57	4,01~4,57	4,2
Nº doadores H	1	1	1	1	1	1
Nº aceitadores (receptores) H	3	2	3	3	2~3	2,75 aprox~3
Ligações rotativas	3	2	3	2	2~3	2,5 aprox~2
Fórmula	C ₂₄ H ₃₀ N ₂ O ₃	C ₁₉ H ₁₉ ClN ₂ O ₂	C ₂₂ H ₂₈ N ₂ O ₃	C ₂₂ H ₂₈ N ₂ O ₂		
Imagem 2D						

As propriedades físico-químicas foram calculadas por SwissADME. A molécula Halofenozide*, apesar de estar presente na tabela do IRAC, não tem nenhum registro de produto no Ministério da Agricultura.

3.4. Discussão

3.4.1. Validação do preparo antecedendo o *screening* e diminuição significativa do número de compostos a serem testados *in vitro*

O software OMEGA da *OpenEye*, como descrito em Hawkins, et al., 2010, cumpriu satisfatoriamente seus objetivos. O programa OMEGA realizou a preparação das bibliotecas para geração de conformadores (GUNERA, et al, 2020) resultando em um banco de dados conformacional utilizado posteriormente para cálculos de similaridade de forma (CHEN, et al. 2021). OMEGA demonstrou capacidade de validar o espaço conformacional das moléculas as quais buscaram-se ser semelhantes (HAWKINS, et al, 2010), além de um bom desempenho ao analisar os critérios de precisão e velocidade (HAWKINS, et al., 2020).

Todas as conformações analisadas e seus respectivos compostos resultaram em uma diminuição significativa para o prosseguimento do *discovery* de compostos, o que é um resultado bastante positivo, uma vez que gerou dados que sugerem que os compostos destacados, após a finalização do *docking*, representam candidatos promissores para a descoberta de moléculas, assim como discutido em BRUS, et. al 2014, estudo em que foi avaliada a estrutura de um novo inibidor seletivo na descoberta de alvos promissores na indústria farmacêutica.

Na busca de compostos derivados ou análogos de compostos já bem dotados de bons resultados, seja na indústria farmacêutica ou agrícola, a triagem virtual se mostra uma ótima ferramenta para diminuir custos e tempo em relação a outros métodos. Assim, o método demonstrou as interações de nível atômico e ofereceu oportunidades para o desenvolvimento dos compostos encontrados, como o encontrado em DIGHE e colaboradores (2016). E, assim como a validação realizada na presente pesquisa, o design dos novos derivados se mostrou uma importante etapa em Bortolami, et. al. 2021. A técnica fornece diversos estudos para o *discovery* de novos compostos (MILES e ROSS, 2021).

3.4.2. Coeficientes de *docking* entre a proteína e o ligante que demonstraram melhores acoplamentos

Os compostos que mostraram os melhores resultados, demonstraram também características parecidas com os compostos referência que foram utilizados na busca de análogos. Esse resultado demonstra a relevância de ter utilizado um programa de similaridade estrutural e química antecedendo o *docking* dos compostos e não somente um método de docagem.

Em comparação de *Shape-matching*, métodos que, assim como o ROCS, utilizam a similaridade como ferramenta de triagem, podem se mostrar superiores em relação a eficácia, que métodos que utilizam o acoplamento do ligante e proteína como triagem. Por essa razão, foram levados em comparação ambos coeficientes na presente pesquisa, uma vez que alguns resultados mostraram que uma abordagem centrada em ligantes baseada em forma é mais consistente e superior que uma centrada em proteínas adotada pelo encaixe (HAWKINS et. al, 2007). Porém, em MCGAUGHEY e colaboradores (2007), foi mostrado resultados bastante positivos no programa FRED, caracterizando-se como um dos melhores programas de encaixe analisados.

Em relação aos resíduos, utilizou-se o mesmo software de *docking* para identificá-los, onde mostrou-se os principais aminoácidos envolvidos nas ligações. As ligações de pontes de hidrogênio são importantes meios de ficar a molécula na proteína, assim como outras ligações. O dado de "Residue" gerado no programa, mostra que os aminoácidos que podem estar envolvidos nessas ligações, segundo a simulação virtual, e a localização espacial deste. Com isso, obteve-se que, os 3 compostos análogos de Tebufenozida analisados mostraram os aminoácidos Treonina e Tirosina; os 3 compostos análogos de 20HE mostraram os aminoácidos Treonina e Tirosina e os 3 compostos análogos do agonista sintético BY06830 mostraram Treonina como resíduo.

Ao analisar as moléculas inseticidas cromafenozida, halofenozida, metoxyfenozida e tebufenozida com o mesmo programa e utilizando os mesmos parâmetros de programação, os aminoácidos foram Treonina e Tirozina em sua grande maioria e apenas a molécula inseticida Metoxifenozida não apresentou Treonina, e sim aspargina e tirosina como resíduos, segundo a análise do software FRED. A natureza dos resíduos ao redor da cavidade pode conferir especificidade ou preferência para diferentes ligantes, podendo esses aminoácidos definir a

cavidade de ligação (SIERRA, et. al. 2021), justificando a importância da compatibilidade de resíduos das moléculas selecionadas com as moléculas inseticidas já registradas no país.

3.4.3. Propriedades físico-químicas (regra dos 5), coeficiente de Tanimoto (ROCS) e ChemGauss4 (FRED)

De todas as moléculas analisadas, a molécula descrita com o código AASDoc1 apresentou o melhor acoplamento se comparado com todas as moléculas aqui avaliadas e as melhores pontuações ranqueadas por Chemgauus dos análogos do hormônio 20HE, apresentaram pontuação maior que -15,81 (referência agonista sintético) definido pelo software FRED, o que demonstra um acoplamento inferior ao definido como parâmetro para a presente pesquisa. Apenas uma violação foi encontrada na molécula ATDoc2, no parâmetro mlogP e duas moléculas demonstraram similaridade relativamente baixa com a molécula modelo, a molécula AASDoc2 em relação ao agonista sintético e a molécula ATDoc1 em relação à tebufenozida.

As propriedades físico-química analisadas foram selecionadas através de parâmetros que demonstraram correlação com biodisponibilidade oral de drogas por Lipinski. 2000 e após revisadas especificamente para inseticidas comerciais em Tice 2001. Os intervalos de parâmetros simples utilizados foram: peso molecular, log P, doadores e receptores de ligações de hidrogênio e ligações rotativas com a média utilizada por Tice. Verificou-se também através do programa SwissADME os valores destes parâmetros para os inseticidas registrados como “agonistas de receptores de ecdisteroides” de acordo com IRAC 2022 e foi realizada a média simples destes parâmetros e o intervalo dos valores.

Coeficiente de ChemGauss (FRED) foi analisado de acordo com a molécula de referência agonista sintético BY06830, uma vez que ele foi considerado a testemunha para este parâmetro, pois a molécula se encontra noPDB encaixada na proteína 1R20, utilizada para a realização dos experimentos *in silico*, via cristalografia de raio X. A cristalografia de raio X nos demonstra exatamente onde a molécula está encaixada na proteína e é uma referência de que o *docking* desta molécula com aconteceu fora das experimentações *in silico*,

por isso, a referência deste coeficiente foi feita a partir de dados de cristalografia de raio X. Esse coeficiente foi de 15,81. E, para o coeficiente de Tanimoto (ROCS) foi utilizado o combo, onde é somado a molécula sobre valores estruturais (1 ponto) e químicos (1 ponto).

Os análogos de Tebufenozida obtiveram valores dentro do intervalo de inseticidas determinado por Tice (2001), exceto o composto nomeado de ATDoc2 (C₂₀H₁₅N₃O₅), que demonstrou um logP superior ao intervalo, constatando assim uma violação da regra. Dentre os análogos de Tebufenozida, o melhor acoplamento (menor coeficiente de ChemGauss, ranqueado pelo software FRED) não foi do composto com maior semelhança estrutural com a molécula referência, o que nos mostra que utilizar o coeficiente de Tanimoto é uma importante orientação, mas somente a análise dele nem sempre nos demonstrará os melhores encaixes entre a molécula e a proteína.

Ao analisar os parâmetros obtidos para os análogos de Tebufenozida em relação aos valores dos inseticidas agonistas de receptores de ecdisteroides, temos alguns valores próximos as médias destes. Obteve destaque a molécula 3 ATDoc3 (C₁₇H₁₄F₃N₃O₅) que apresentou peso molecular bastante próximo ao dos inseticidas agonistas de receptores de ecdisteroides. Esse é um importante dado, uma vez que, no decorrer dos anos, observa-se que o aumento do peso molecular pode ser correlacionado com menor solubilidade e menor penetração pelas membranas (RAO et. al. 2015; LIPINSKI et. al, 2012).

Os análogos do hormônio 20HE demonstraram os valores de *docking* menos interessantes, todos superiores a -15, porém, não demonstraram nenhuma violação da regra de inseticidas de Tice (2001) e, ainda, apresentaram bons valores de peso molecular, bastante próximos a média de inseticidas agonistas de receptores de ecdisteroides e dentro do intervalo analisado nesta pesquisa destas moléculas. Os resultados mostraram ainda uma similaridade estrutural ao hormônio 20HE, com todos os valores do coeficiente de combo Tanimoto superior a 1, resultado de extrema relevância ao considerar que a bolsa de ligação teve um raio definido por meio de estimativa e que, segundo Hawkins et. al, 2007, parâmetros estruturais, como o utilizado, podem apresentar superioridade em relação a parâmetro centrados no encaixe da proteína e ligante.

Os análogos do agonista sintético BY06830, não demonstraram violações a regra, obtiveram resultados bastante inferiores a -15, o que resulta em bons acoplamentos de acordo com o programa FRED. Porém, a molécula AASDoc2 mostrou-se um pouco inferior a 1, no coeficiente de combo Tanimoto, por outro lado, apesar de todas as moléculas analisadas estarem no intervalo dos inseticidas registrados para Tice (2001), a molécula 2 foi a que se posicionou de forma mais próxima para o intervalo inseticidas agonistas de receptores de ecdisteroides.

É essencial salientar ainda que todos os parâmetros auxiliam em uma orientação de uma possível molécula que se tornará um bom inseticida, levando em consideração parâmetros similares a maior parte dos inseticidas registrados, semelhanças químicas e estruturais, compatibilidade de encaixe da proteína e ligante e provável penetração de membrana, porém, existem compostos que irão ir contra os parâmetros citados, mas que *in vivo*, funcionam de forma exemplar. Um modelo, bastante referência no mercado agrícola e de importância econômica é a molécula herbicida glifosato, que se mostra fora dos limites estabelecidos por Tice (2001), esse fato pode ocorrer devido a maneira que o herbicida penetra na folha, entre outros prováveis motivos.

Um parâmetro bastante estudado é o coeficiente de partição octanol água, uma vez que ele auxilia na previsão da toxicidade dos produtos, via bioacumulação ou capacidade de atravessar na barreira hematoencefálica (LIPINSKI et al, 2000; TICE 2001). E em estudos como Webb e colaboradores (2022) este parâmetro é utilizado para auxiliar na determinação da sorção.

A partir das regras de Lipinski para medicamentos, outros diversos estudos analisaram e discutiram os parâmetros físico-químicos (HAO et al. 2011; GANDY et al. 2015; HUANG, et al. 2021; SUKHOVERKOV, et al. 2021),

apesar de haverem falhas nas aplicações, as regras são uma ferramenta bastante útil ao usar abordagem de alto rendimento (SUKHOVERKOV, et al. 2021).

3.5. Conclusão

O *screening* de um milhão, duzentos e setenta e seis mil, quatrocentos e setenta e seis (1.276.476) e cento e quarenta milhões, setecentos e quarenta e quatro mil, setecentos e sessenta e oito (140.744.768) conformers, resultou em seis mil, trezentos e setenta e cinco (6.375) análogos (hits) de Tebufenozida e deste número, quinhentos e cinquenta e quatro (554) docados na proteína e três (3) analisados suas propriedades e comparados a inseticidas. Para os análogos do agonista sintético, foram quatro mil, oitocentos e três análogos (4.803) hits e quinhentos e vinte docados e três (3) analisados. E, por fim, a triagem dos compostos comparados ao 20HE resultaram em vinte (20) hits e dezessete (17) docados, com três (3) compostos analisados. Os resultados mostram que o composto C23H19N3O4 (código: AASDoc1), demonstrou maior destaque somando potencial de similaridade química e estrutural com um inseticida do grupo estudado, porém, sugere-se que futuras avaliações biológicas não devem descartar os demais compostos que apresentaram bons resultados químicos e estruturais

Com base nos fatores analisados, demonstrou-se que a molécula com o melhor acomplamento entre todas as analisadas foi a AASDoc1. Porém outras também demonstraram bons resultados para uma provável penetração de membrana. A similaridade residual entre as moléculas analisadas e os inseticidas comercializados do grupo “receptores de ecdisteroides”, foi um resultado bastante favorável também.

Os dados que geraram semelhanças químicas e estruturais com as moléculas de referência, o encaixe entre a proteína e ligante e provável penetração de membrana, geraram importantes orientações para realizar a testagem *in vitro* e *in vivo* e direções que demonstram os prováveis compostos que se demonstrarão bioativos. Porém as violações analisadas nos compostos, não podem ser excludentes, uma vez que moléculas de agrotóxicos com grande importância econômica, como o herbicida glifosato, mostraram intervalos fora da regra.

Este exemplo, auxilia na justificativa da necessidade de não descartar as moléculas que estavam fora dos parâmetros avaliados como a molécula ATDoc2, pois os parâmetros orientam, mas não descartam os passos de avaliação *in vitro* e *in vivo*.

4. Considerações Finais

Os maiores desafios do projeto giram em torno do entendimento de áreas que, no geral, se distanciam de alunos da área agrícola: bioinformática, linguagem de programação e conceitos moleculares e físico-químicos aplicados. Deste modo, para realização de um projeto deste tipo, o primeiro passo, não constado na dissertação é o estudo do assunto e as primeiras experiências com linha de comando e linguagens de programação.

Através da presente pesquisa, confirmou-se que as técnicas de bioinformática podem auxiliar na descoberta de novas moléculas inseticidas. Dentro da bioinformática, para esse processo, os programas ROCS e FRED revelaram-se ferramentas úteis e eficazes. Do mesmo modo, constatou-se que a utilização de um só parâmetro pode não ser tão satisfatório, quanto a associação deles.

Existe uma crescente necessidade de prever de forma precisa e racional as prováveis novas moléculas na indústria agroquímica e a utilização de ferramentas de bioinformática são um instrumento em potencial para atingir esse objetivo com um menor preço e de forma mais precisa. Ainda que algumas moléculas estejam fora dos parâmetros químicos e estruturais já pré-definidos as regras são uma ferramenta eficaz ao usar abordagem de alto rendimento, ou seja, realizar a análise de mais de um milhão de moléculas em um curto período de tempo. Desta forma, as técnicas se mostraram bastante satisfatórias para o seu objetivo, porém para ocorrer uma validação ainda mais forte, é preciso a realização de mais testes além do in sílico.

Referências

- AUWUL, M. R. et al. Bioinformatics and machine learning approach identifies potential drug targets and pathways in COVID-19. **Briefings in Bioinformatics**, v. 22, n. 5, p. bbab120, 2 set. 2021.
- BAJUSZ, D.; RÁCZ, A.; HÉBERGER, K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? **Journal of Cheminformatics**, v. 7, n. 1, p. 20, 20 maio 2015.
- BASKARAN, S. G.; SHARP, T. P.; SHARP, K. A. Computational Graphics Software for Interactive Docking and Visualization of Ligand–Protein Complementarity. **Journal of Chemical Information and Modeling**, v. 61, n. 3, p. 1427–1443, 22 mar. 2021.
- BERENGER, F. et al. Lean-Docking: Exploiting Ligands' Predicted Docking Scores to Accelerate Molecular Docking. **Journal of Chemical Information and Modeling**, v. 61, n. 5, p. 2341–2352, 24 maio 2021.
- BILLAS, I. M. L. et al. Structural adaptability in the ligand-binding pocket of the ecdysone hormone receptor. **Nature**, v. 426, n. 6962, p. 91–96, nov. 2003.
- BILLAS, I. M. L. et al. The Structure and Function of Ecdysone Receptors. In: SMAGGHE, G. (Ed.). . *Ecdysone: Structures and Functions*. Dordrecht Springer Netherlands, 2009. p. 335–360.
- BIOVIA Discovery Studio » EDC. EDC**, [s.d.]. Disponível em: <<https://edctecnologia.com.br/solucoes/biovia-discovery-studio/>>. Acesso em: 24 abr. 2022
- BORTOLAMI, M. et al. New Pyrimidine and Pyridine Derivatives as Multitarget Cholinesterase Inhibitors: Design, Synthesis, and In Vitro and In Cellulo Evaluation. **ACS Chemical Neuroscience**, v. 12, n. 21, p. 4090–4112, 3 nov. 2021.
- BRINKMANN, S. et al. Identification, Characterization, and Synthesis of Natural Parasitic Cysteine Protease Inhibitors: Pentacitidins Are More Potent Falcitidin Analogues. **ACS Chemical Biology**, v. 17, n. 3, p. 576–589, 18 mar. 2022.
- BRUS, B. et al. Discovery, Biological Evaluation, and Crystal Structure of a Novel Nanomolar Selective Butyrylcholinesterase Inhibitor. **Journal of Medicinal Chemistry**, v. 57, n. 19, p. 8167–8179, 9 out. 2014.
- CHEN, W. et al. Structure-based virtual screening of highly potent inhibitors of the nematode chitinase CeCht1. **Journal of Enzyme Inhibition and Medicinal Chemistry**, v. 36, n. 1, p. 1198–1204, 1 jan. 2021.
- CHARLES, Jean-Philippe. The regulation of expression of insect cuticle protein genes. **Insect biochemistry and molecular biology**, v. 40, n. 3, p. 205-213, 2010.
- COURNIA, Z. et al. Rigorous Free Energy Simulations in Virtual Screening. **Journal of Chemical Information and Modeling**, v. 60, n. 9, p. 4153–4169, 28 set. 2020.

- DENG, D. et al. XGraphBoost: Extracting Graph Neural Network-Based Features for a Better Prediction of Molecular Properties. **Journal of Chemical Information and Modeling**, v. 61, n. 6, p. 2697–2705, 28 jun. 2021.
- DHANABALAN, A. K. et al. Identification of a Chlorogenic Ester as a Monoamine Oxidase (MAO-B) Inhibitor by Integrating “Traditional and Machine Learning” Virtual Screening and In Vitro as well as In Vivo Validation: A Lead against Neurodegenerative Disorders? **ACS Chemical Neuroscience**, v. 12, n. 19, p. 3690–3707, 6 out. 2021.
- DIGHE, S. N. et al. Discovery and Structure–Activity Relationships of a Highly Selective Butyrylcholinesterase Inhibitor by Structure-Based Virtual Screening. **Journal of Medicinal Chemistry**, v. 59, n. 16, p. 7683–7689, 25 ago. 2016.
- GAUTHIER, J. et al. A brief history of bioinformatics. **Briefings in Bioinformatics**, v. 20, n. 6, p. 1981–1996, 27 nov. 2019.
- GLAAB, E.; MANOHARAN, G. B.; ABANKWA, D. Pharmacophore Model for SARS-CoV-2 3CLpro Small-Molecule Inhibitors and *in Vitro* Experimental Validation of Computationally Screened Inhibitors. **Journal of Chemical Information and Modeling**, v. 61, n. 8, p. 4082–4096, 23 ago. 2021.
- GOSSEN, J. et al. A Blueprint for High Affinity SARS-CoV-2 Mpro Inhibitors from Activity-Based Compound Library Screening Guided by Analysis of Protein Dynamics. **ACS Pharmacology & Translational Science**, v. 4, n. 3, p. 1079–1095, 11 jun. 2021.
- GUNERA, J. et al. Structure-Based Discovery of Novel Ligands for the Orexin 2 Receptor. **Journal of Medicinal Chemistry**, v. 63, n. 19, p. 11045–11053, 8 out. 2020.
- GUEST, E. E. et al. Molecular Simulation of $\alpha\beta6$ Integrin Inhibitors. **Journal of Chemical Information and Modeling**, v. 60, n. 11, p. 5487–5498, 23 nov. 2020.
- GUEST, E. E. et al. Alchemical Free Energy Methods Applied to Complexes of the First Bromodomain of BRD4. **Journal of Chemical Information and Modeling**, v. 62, n. 6, p. 1458–1470, 28 mar. 2022.
- GUNERA, J. et al. Structure-Based Discovery of Novel Ligands for the Orexin 2 Receptor. **Journal of Medicinal Chemistry**, v. 63, n. 19, p. 11045–11053, 8 out. 2020.
- HAWKINS, P. C. D. et al. Conformer Generation with OMEGA: Algorithm and Validation Using High Quality Structures from the Protein Databank and Cambridge Structural Database. **Journal of Chemical Information and Modeling**, v. 50, n. 4, p. 572–584, 26 abr. 2010.
- HAWKINS, P. C. D.; SKILLMAN, A. G.; NICHOLLS, A. Comparison of Shape-Matching and Docking as Virtual Screening Tools. **Journal of Medicinal Chemistry**, v. 50, n. 1, p. 74–82, 1 jan. 2007.
- HAWKINS, P. C. D.; WLODEK, S. Decisions with Confidence: Application to the Conformation Sampling of Molecules in the Solid State. **Journal of Chemical Information and Modeling**, v. 60, n. 7, p. 3518–3533, 27 jul. 2020.

HEFFERNAN, M. L. R. et al. Ulotaront: A TAAR1 Agonist for the Treatment of Schizophrenia. **ACS Medicinal Chemistry Letters**, v. 13, n. 1, p. 92–98, 13 jan. 2022.

HODCROFT, E. B. et al. Want to track pandemic variants faster? Fix the bioinformatics bottleneck. **Nature**, v. 591, n. 7848, p. 30–33, mar. 2021.

HU, T. et al. Bioinformatics resources for SARS-CoV-2 discovery and surveillance. **Briefings in Bioinformatics**, v. 22, n. 2, p. 631–641, 22 mar. 2021.

INETIS. **Home page**. Disponível em: <<https://www.sib.swiss/>>. Acesso em: 24abr. 2022.

ISMAIL, M. I. et al. Design, Synthesis, and Biological Evaluation of Novel 7 *H*-[1,2,4]Triazolo[3,4-*b*][1,3,4]thiadiazine Inhibitors as Antitumor Agents. **ACS Omega**, v. 5, n. 32, p. 20170–20186, 18 ago. 2020.

LIPINSKI, C. A. et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. **Advanced Drug Delivery Reviews**, MOST CITED PAPERS IN THE HISTORY OF ADVANCED DRUG DELIVERY REVIEWS: A TRIBUTE TO THE 25TH ANNIVERSARY OF THE JOURNAL. v. 64, p. 4–17, 1 dez. 2012.

LYU, S. et al. Identification of Phelligrudin-Based Compounds as Novel Human CD73 Inhibitors. **Journal of Chemical Information and Modeling**, v. 61, n. 3, p. 1275–1286, 22 mar. 2021.

MCGANN, M. FRED and HYBRID docking performance on standardized datasets. **Journal of Computer-Aided Molecular Design**, v. 26, n. 8, p. 897–906, 1 ago. 2012.

MCGANN, M. R. et al. Gaussian docking functions. **Biopolymers**, v. 68, n. 1, p.76–90, 2003.

MCGAUGHEY, G. B. et al. Comparison of Topological, Shape, and Docking Methods in Virtual Screening. **Journal of Chemical Information and Modeling**, v. 47, n. 4, p. 1504–1519, 1 jul. 2007.

MENG, X.-Y. et al. Molecular Docking: A powerful approach for structure-based drug discovery. **Current computer-aided drug design**, v. 7, n. 2, p. 146–157, 1 jun. 2011.

MILES, J. A.; ROSS, B. P. Recent Advances in Virtual Screening for Cholinesterase Inhibitors. **ACS Chemical Neuroscience**, v. 12, n. 1, p. 30–41, 6 jan. 2021.

MORRONE, J. A. et al. Combining Docking Pose Rank and Structure with Deep Learning Improves Protein–Ligand Binding Mode Prediction over a Baseline Docking Approach. **Journal of Chemical Information and Modeling**, v. 60, n. 9, p. 4170–4179, 28 set. 2020.

MESCE, Karen A.; FAHRBACH, Susan E. Integration of endocrine signals that regulate insect ecdysis. **Frontiers in neuroendocrinology**, v. 23, n. 2, p. 179–199, 2002.

OAKLEY, M. B.; KIMBALL, G. E. Punched Card Calculation of Resonance Energies. **The Journal of Chemical Physics**, v. 17, n. 8, p. 706–717, 1 ago. 1949.

O'MAHONY, G. et al. Discovery by Virtual Screening of an Inhibitor of CDK5-Mediated PPAR γ Phosphorylation. **ACS Medicinal Chemistry Letters**, v. 13, n.4, p. 681–686, 14 abr. 2022.

OPENEYE SCIENTIFIC. **Cheminformatics Software | Molecular Modeling Software | OpenEye Scientific**. Disponível em: <<https://www.eyesopen.com>>. Acesso em: 23 jan. 2022.

PEROLA; WALTERS, ; CHARIFSON. An Analysis of Critical Factors Affecting Docking and Scoring. In *Virtual Screening in Drug Discovery*; Alvarez, J., Shoichet, B., Eds.; CRC Press, 2005.

PyMOL | pymol.org. Disponível em: <<https://pymol.org/2/>>. Acesso em: 24 abr. 2022.

RAO, H. et al. Physicochemical Profiles of the Marketed Agrochemicals and Clues for Agrochemical Lead Discovery and Screening Library Development. **Molecular Informatics**, v. 34, n. 5, p. 331–338, maio 2015.

RUSH, T. S. et al. A Shape-Based 3-D Scaffold Hopping Method and Its Application to a Bacterial Protein–Protein Interaction. **Journal of Medicinal Chemistry**, v. 48, n. 5, p. 1489–1495, 1 mar. 2005.

SERAFINI, M. et al. Discovery of Highly Potent Benzimidazole Derivatives as Indoleamine 2,3-Dioxygenase-1 (IDO1) Inhibitors: From Structure-Based Virtual Screening to *in Vivo* Pharmacodynamic Activity. **Journal of Medicinal Chemistry**, v. 63, n. 6, p. 3047–3065, 26 mar. 2020.

SIERRA, I. et al. Transcriptomic analysis and molecular docking reveal genes involved in the response of *Aedes aegypti* larvae to an essential oil extracted from Eucalyptus. **PLOS Neglected Tropical Diseases**, v. 15, n. 7, p. e0009587, 16 jul. 2021.

SOFTWARE, O. S. **Cheminformatics Software | Molecular Modeling Software | OpenEye Scientific**. Disponível em: <<https://www.eyesopen.com>>. Acesso em: 24 abr. 2022.

THUMMEL, Carl S. Molecular mechanisms of developmental timing in *C. elegans* and *Drosophila*. **Developmental cell**, v. 1, n. 4, p. 453-465, 2001.

TICE, C. M. Selecting the right compounds for screening: does Lipinski's Rule of 5 for pharmaceuticals apply to agrochemicals? **Pest Management Science**, v. 57, n. 1, p. 3–16, jan. 2001.

TRUMAN, J. W. Hormonal control of insect ecdysis: endocrine cascades for coordinating behavior with physiology. **Vitamins & Hormones**, v. 73, p. 1-30, 2005.

WANG, J.; MORIN, P.; WANG, W.; KOLLMAN, P. A. Use of MM-PBSA in Reproducing the Binding Free Energies to HIV-1 RT of TIBO Derivatives. *J. Am. Chem. Soc.* 2001, 123, 5221–5230.

WHITFORD, F. et al. The Pesticide Marketplace. **Purdue University**, 2009. Disponível em: <<https://edustore.purdue.edu/>>. Acesso em: 22 maio. 2022.

WILLIS, JUDITH H. Metamorphosis of the cuticle, its proteins, and their genes. In: **Metamorphosis**. Academic Press, 1996. p. 253-282.

YAN, X. C. et al. Augmenting Hit Identification by Virtual Screening Techniques in Small Molecule Drug Discovery. **Journal of Chemical Information and Modeling**, v. 60, n. 9, p. 4144–4152, 28 set. 2020.

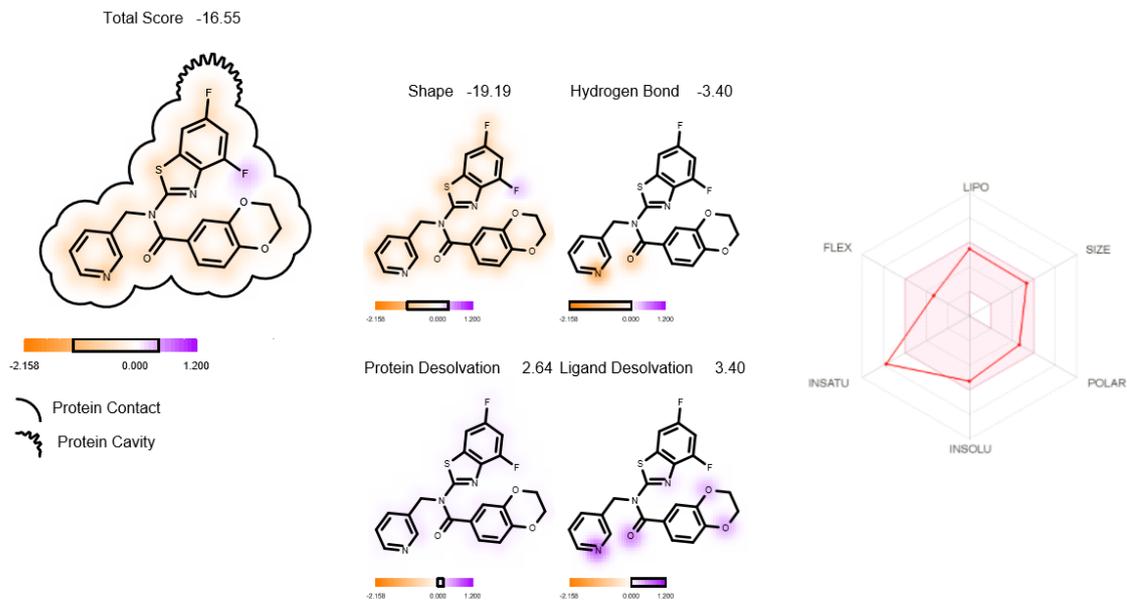
YANG, Y. et al. Efficient Exploration of Chemical Space with Docking and Deep Learning. **Journal of Chemical Theory and Computation**, v. 17, n. 11, p. 7106–7119, 9 nov. 2021.

ZHANG, W.; HUANG, J. EViS: An Enhanced Virtual Screening Approach Based on Pocket–Ligand Similarity. **Journal of Chemical Information and Modeling**, v. 62, n. 3, p. 498–510, 14 fev. 2022.

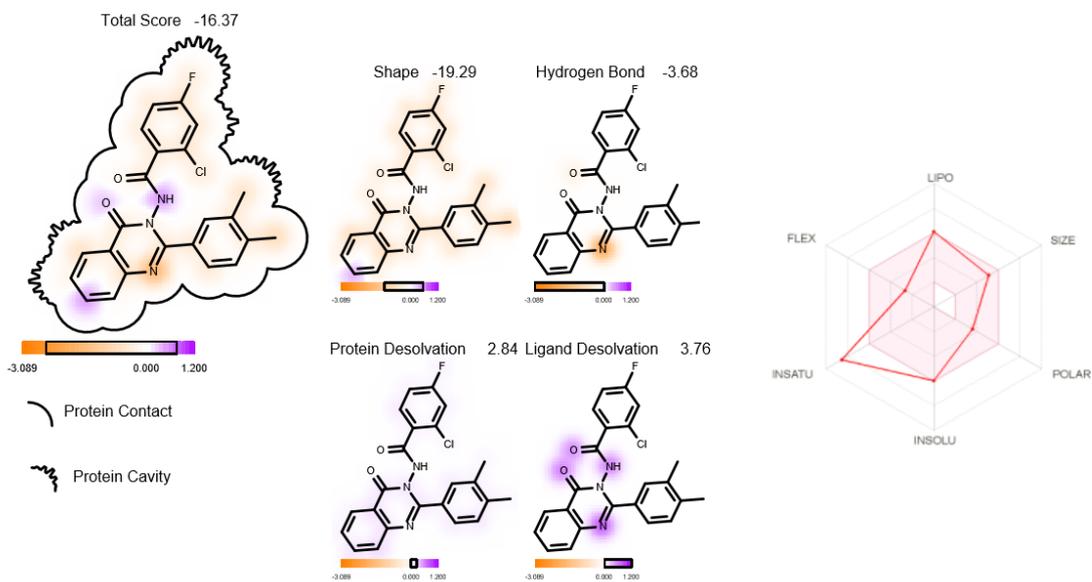
ŽITŇAN, D. et al. Complex steroid–peptide–receptor cascade controls insect ecdysis. **General and comparative endocrinology**, v. 153, n. 1-3, p. 88-96, 2007.

Apêndices

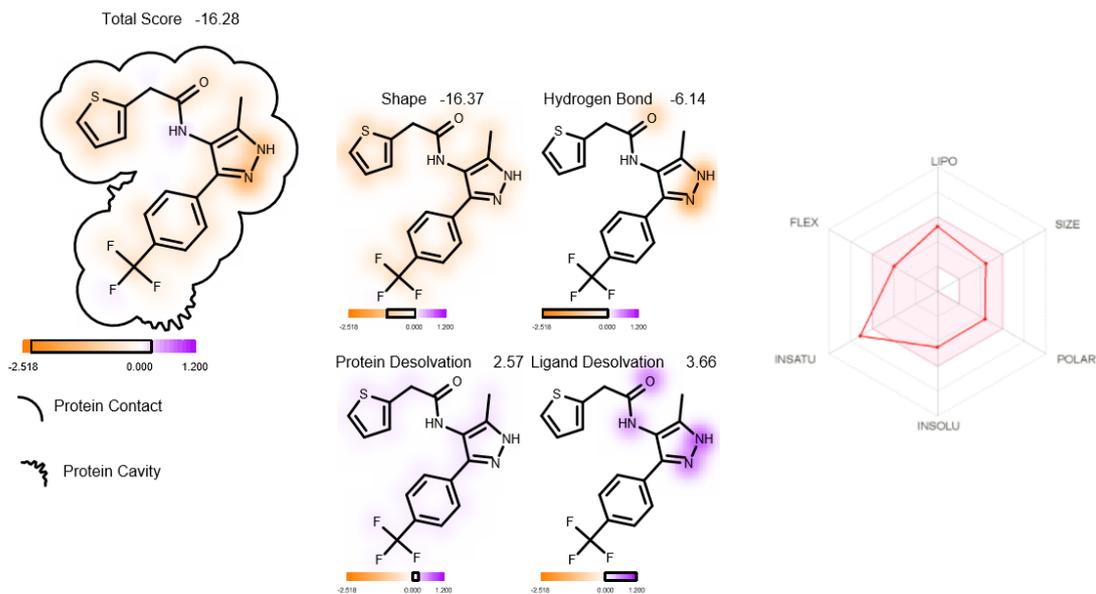
Apêndice A. Análogo Tebufenozida com o melhor docking



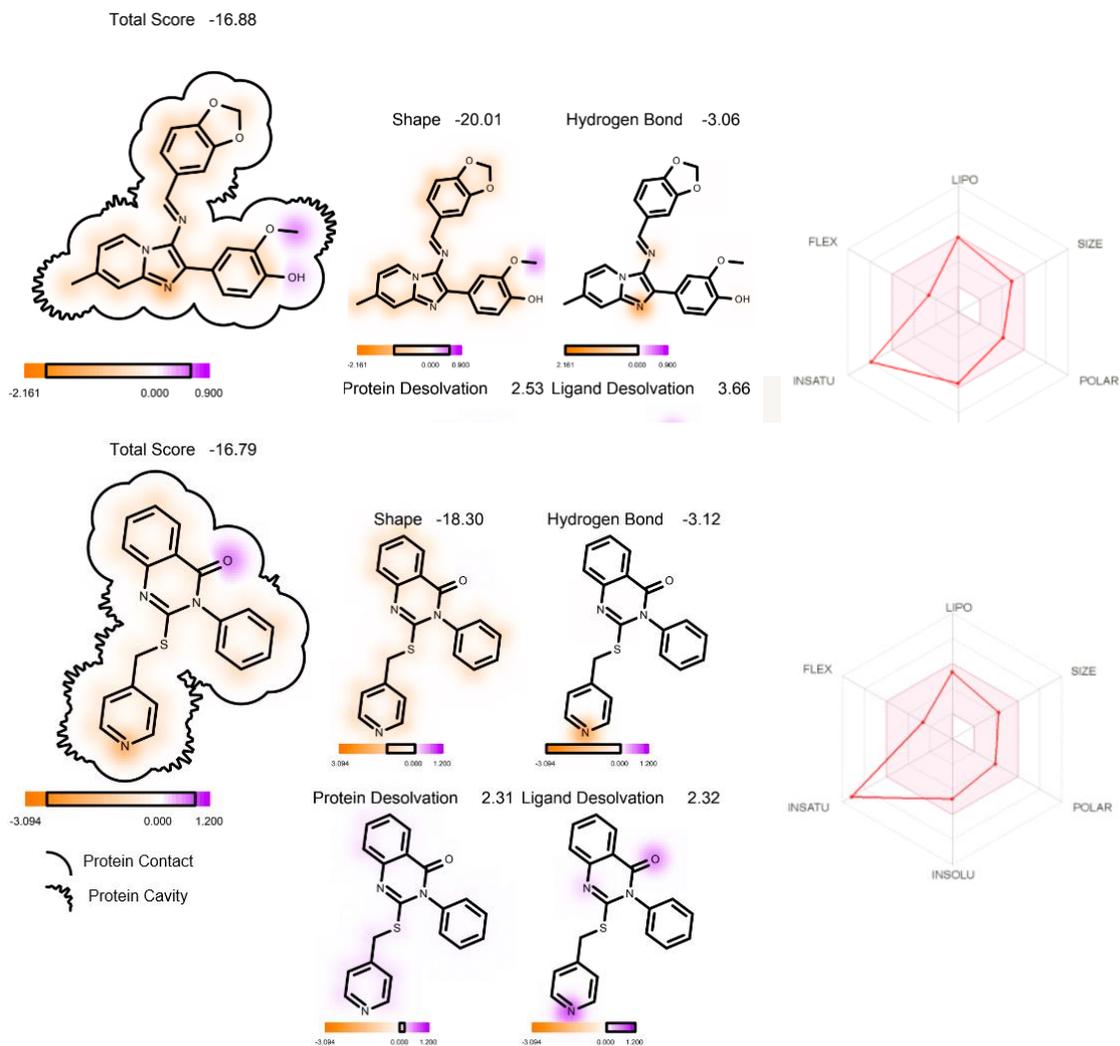
Apêndice B. Análogo Tebufenozida com o segundo melhor docking



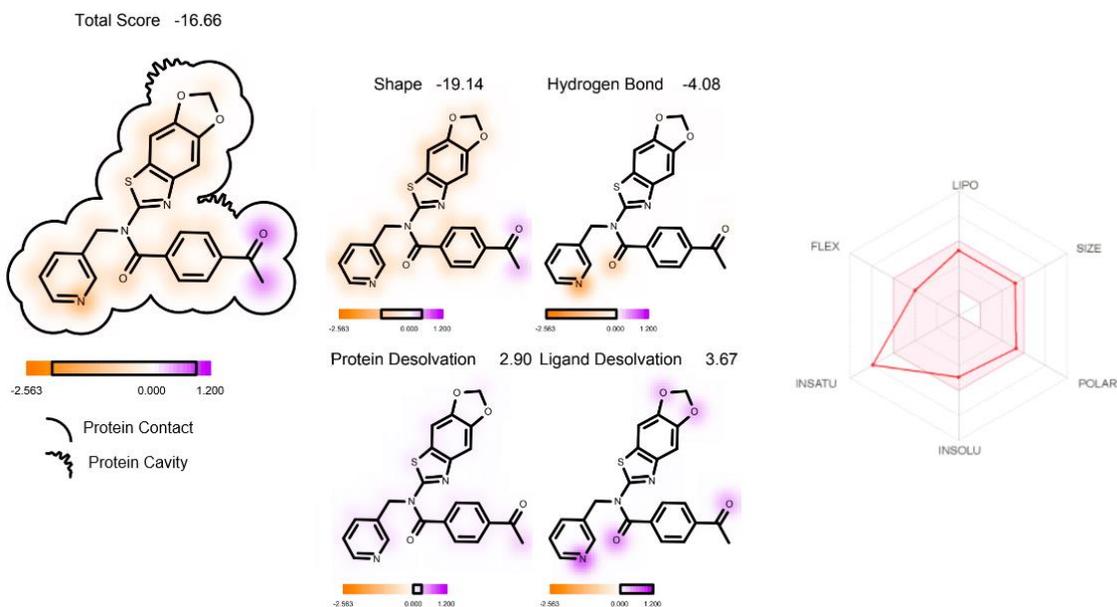
Apêndice C. Análogo Tebufenozida com o terceiro melhor docking



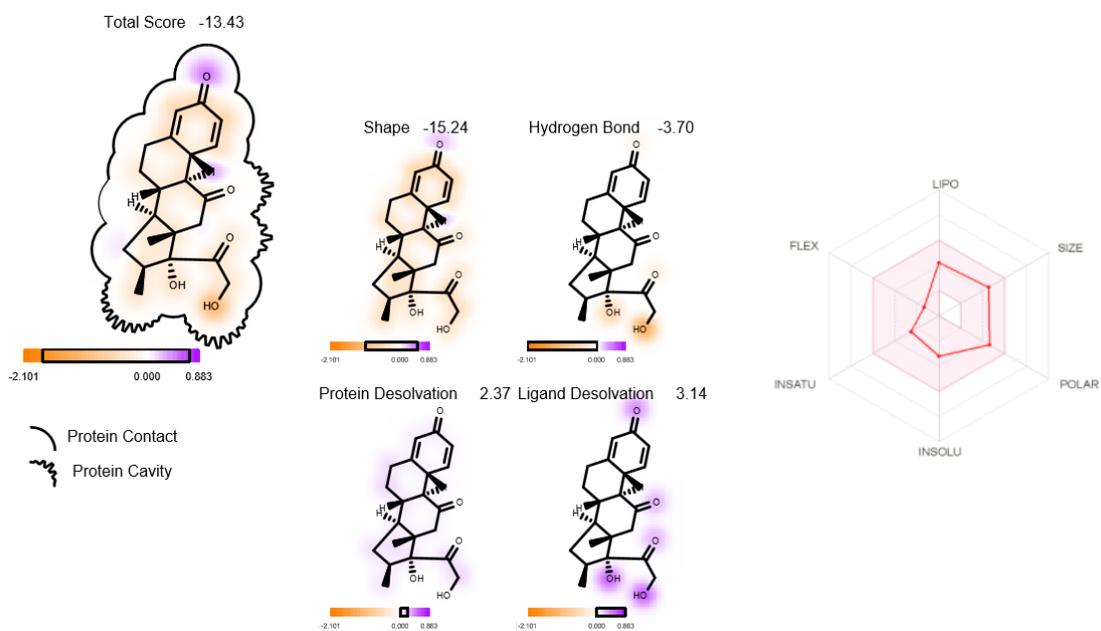
Apêndice D. Análogo agonista sintético com o melhor docking



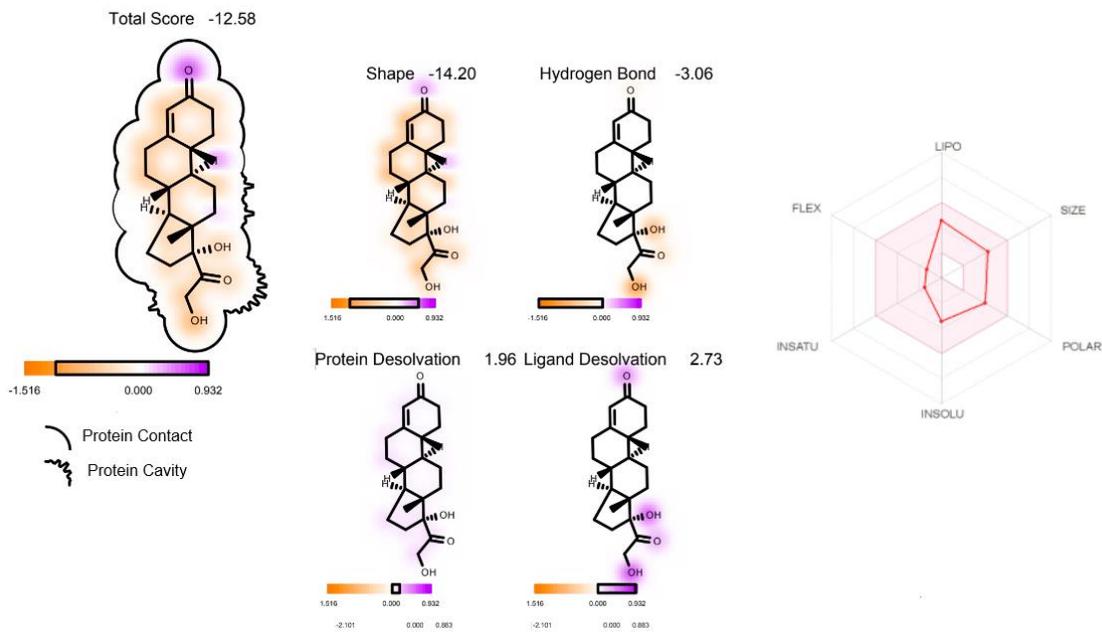
Apêndice F. Análogo do agonista sintético com o terceiro melhor docking



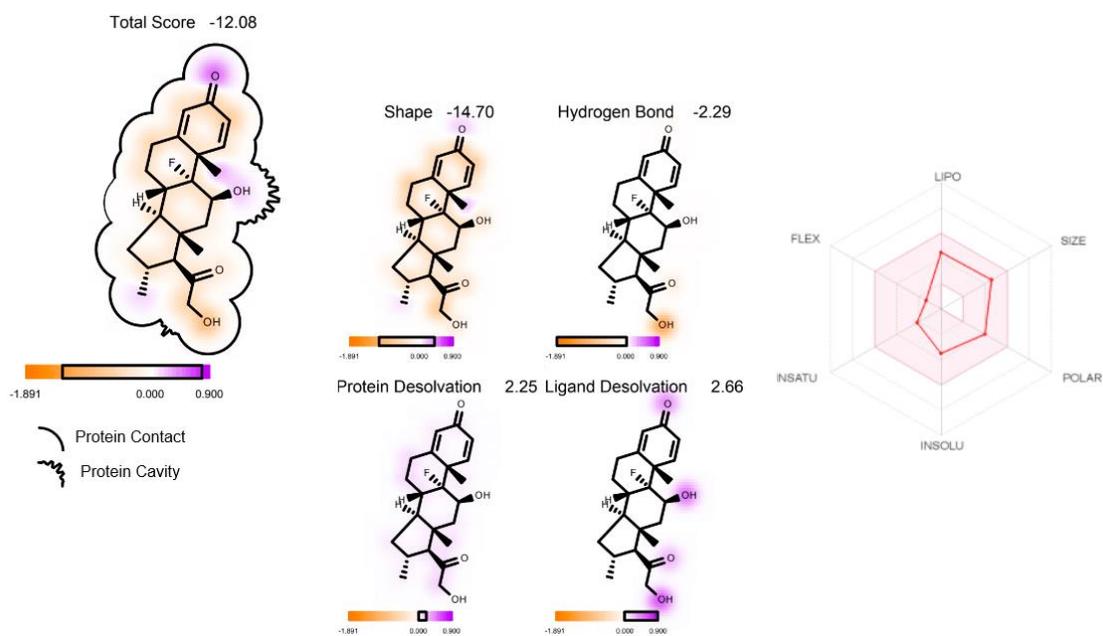
Apêndice G. Análogo do 20HE com o melhor docking



Apêndice H. Análogo do 20HE com o segundo melhor docking



Apêndice I. Análogo do 20HE com o terceiro melhor docking



Anexos

Anexo A – Resultados do software OMEGA

Origem Biblioteca	Nome Biblioteca	Número de moléculas	Média de conformações por molécula	Conformers
Asinex	Asinex BioDesign Library	150.440	154.573	23.223.068
Asinex	Asinex Gold & Platinum Collections	263.557	157,756	3.799.238
Lifechemicals	Screening receptor based_approach	87.883	122,937	1.079.758
Lifechemicals	Nuclear Receptors Similarity	6.107	118,487	723.484
Lifechemicals	Nuclear Receptors Docking	4.969	139,187	691.623
Lifechemicals	Insecticides Library	1.831	127,672	233.768
Lifechemicals	Bioactive compound	8.000	121,106	968.853
Lifechemicals	Natural product (similarity)	913	92,131	84.116
lifechemicals	Natural product (chemoinformatics)	2.029	113,54	230.377
lifechemicals	STING Target	1.602	157,95	253.038
Apexbt	L1022 Discovery probebioactive	3.216	108,676	349.503
Apexbt	Discovery Probebioactive ?	2.909	107,983	314.123
Selleckchem	Bioactive compound library	7.534	102,448	771.850
Chemdiv	3D Diversity	21.703	159,608	3.465.541
Chemdiv	Agro	46.346	117,834	5.461.171
Chemdiv	CDI BMS DivSet_300	300.521	154,758	46.508.039
Chemdiv	Div_150	150.000	154,054	23.009.581
Chemdiv	Div_set_50	50.000	132,861	6.622.901
Chemdiv	Div_set_100	100.000	139,59	13.911.820
Chemdiv	RNA-V	2111	101,705	214.701
lifechemicals	Focused by Machine 1	6865	190,77	1.309.639
lifechemicals	Focused Similarity 2	3014	133,946	403.714
lifechemicals	Focused Rule of four	3368	177,882	598.932
lifechemicals	Focused Similarity 4	14396	133,6741	1.924.240
lifechemicals	Antifungal	3912	103,899	406454
lifechemicals	Screening targeted	9864	111,383	1.098.683
lifechemicals	Epigenetic targeted	7019	147,639	1036380
lifechemicals	Fungicides	1237	132,474	163.871
lifechemicals	Herbicides	7744	140, 466	1.087.769
lifechemicals	Hormone related focused	749	99,24	74.338
lifechemicals	Microbiocides	1912	113,974	217.920
lifechemicals	RNA Focused	4450	108,546	483.030
lifechemicals	Veterinary Focused	236	98,495	23.245

Anexo B – Resultados do software ROCS

Nome Biblioteca	Tebufenozida HITS	Agonista sintético HITS	20HE HITS
Asinex BioDesign Library	500	500	0
Asinex Gold & Platinum Collections	180	446	0
Screening receptor based_approach	399	190	0
Nuclear Receptors Similarity	98	121	0
Nuclear Receptors Docking	0	0	0
Insecticides Library	58	23	0
Bioactive compound	197	150	0
Natural product (similarity)	5	2	0
Natural product (chemoinformatics)	12	12	0
STING Target	19	28	0
L1022 Discovery probebioactive	19	17	2
Discovery Probebioactive ?	17	16	2
Bioactive compound library	85	39	12
3D Diversity	500	0	0
Agro	500	500	0
CDI BMS DivSet_300	500	500	0
Div_150	500	223	0
Div_set_50	500	500	0
Div_set_100	500	500	0
RNA-V	19	7	0
Focused by Machine 1	9	14	0
Focused Similarity 2	20	19	0
Focused Rule of four 3	10	33	0
Focused Similarity 4	299	246	0
Antifungal	126	80	0
Screening targeted	500	297	0
Epigenetic targeted	119	134	0
Fungicides	20	11	0
Herbicides	151	141	0
Hormone related focused	1	0	1
Microbiocides	9	6	0
RNA Focused	500	48	0
Veterinary Focused	3	0	0