



MINISTÉRIO DA EDUCAÇÃO
Universidade Federal de Alfenas . Unifal-MG
Rua Gabriel Monteiro da Silva, 714 – Alfenas/MG – CEP 37130-000
Fone: (35) 3299-1000 . Fax: (35) 3299-1063



ICARO ALVES PINTO

**IDENTIFICAÇÃO *IN SILICO* DE POTENCIAIS INIBIDORES DA VIA WNT
CANÔNICA EM ADENOCARCINOMA MAMÁRIO HUMANO**

Alfenas, 2018

ICARO ALVES PINTO

**IDENTIFICAÇÃO *IN SILICO* DE POTENCIAIS INIBIDORES DA VIA WNT
CANÔNICA EM ADENOCARCINOMA MAMÁRIO HUMANO**

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Ciências Biológicas pela Universidade Federal de Alfenas, Área de concentração: Biologia Molecular.
Orientador: Prof. Dr. Nelson José Freitas da Silveira
Co-orientadora: Prof^a. Dra. Cibele Marli Cação Paiva Gouvêa

Alfenas/MG

2018

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema de Bibliotecas da Universidade Federal de Alfenas
Biblioteca Central – Campus Sede

Pinto, Icaro Alves.

Identificação *in silico* de potenciais inibidores da via Wnt canônica em adenocarcinoma mamário humano / Icaro Alves Pinto. -- Alfenas, MG, 2018.
63 f. : il. –

Orientador: Nelson José Freitas da Silveira.
Dissertação (Mestrado em Ciências Biológicas) – Universidade Federal de Alfenas, 2018.
Bibliografia.

1. Biologia Computacional. 2. Neoplasias da Mama. 3. Via de Sinalização Wnt. 4. Desenho de Drogas. I. Silveira, Nelson José Freitas da. II. Título.

CDD 572.8

Ficha Catalográfica elaborada por Fátima dos Reis Goiatá



ICARO ALVES PINTO

“IDENTIFICAÇÃO *IN SILICO* DE POTENCIAIS INIBIDORES DA VIA WNT CANÔNICA EM
ADENOCARCINOMA MAMÁRIO HUMANO”

A Banca Examinadora, abaixo assinada, aprova a
Dissertação apresentada como parte dos requisitos para
a obtenção do título de Mestre em Ciências Biológicas
pela Universidade Federal de Alfenas.

Área de concentração: Biologia Molecular

Aprovado em: 25 de julho de 2018.

Prof. Dr. Nelson José Freitas da Silveira

Instituição: UNIFAL-MG

Assinatura:

Prof. Dr. Tiago Henrique

Instituição: FAMERP

Assinatura:

Prof. Dr. Marcos José Marques

Instituição: UNIFAL-MG

Assinatura:

"Computers are like Old Testament gods: lots of rules and no mercy."

(CAMPBELL, 1990, p. 201)

RESUMO

O câncer de mama é a principal causa mundial de morte por câncer em mulheres, sendo o triplo-negativo o de pior prognóstico. A prospecção de novas substâncias com potencial antitumoral é importante para a contenção e tratamento da doença. O presente trabalho teve como objetivo a identificação e de um fármaco ou ligante comercial, a partir de análises de bioinformática, que potencialmente se ligue à proteína transmembrana FZD7 de forma a inativar a via de sinalização Wnt em células de adenocarcinoma mamário humano triplo-negativo (TMTN). Como objetivos específicos, pretendeu-se modelar computacionalmente as proteínas FZD7, Wnt3 e Wnt3a, disponibilizá-las em bancos de dados de modelos e realizar análises de *docking* para aferir a energia de ligação entre a FZD7 e os ligantes disponíveis comercialmente selecionados. Para a escolha das proteínas a serem estudadas foi feita uma pesquisa por proteínas superexpressas em TMTN no banco de dados GEO. As proteínas Wnt3 e Wnt3a foram modeladas por homologia, utilizando o software Modeller, e a FZD7 por homologia e por modelagem *ab initio*, com auxílio da ferramenta online I-TASSER. Foram gerados gráficos de Ramachandran e conduziu-se análises de RMSD para avaliar a qualidade estereoquímica dos modelos. Os ligantes foram selecionados com base em sua similaridade ao ácido palmitoleico, e foram tirados do banco de dados ZINC. Análises de *docking* com o software Autodock Vina foram conduzidas para aferir a energia de ligação entre o domínio rico em cisteína (CRD) da FZD7 e os ligantes encontrados. Os melhores modelos gerados para as proteínas Wnt3 e Wnt3a foram depositados no banco de dados público PMDB. Foram encontrados 30 ligantes disponíveis comercialmente em bancos de dados e foram realizadas análises de *docking* entre eles e o CRD da proteína FZD7. Os resultados apontam que os ligantes zinc08221009, zinc13546050, zinc05260769, zinc04529321 e zinc05972969 são candidatos promissores para o desenvolvimento de novos fármacos. A conhecimento dos autores, não há modelos de qualidade das proteínas Wnt3 e Wnt3a disponíveis atualmente em bancos de dados públicos. Os modelos gerados e análises conduzidas pelo presente trabalho serão disponibilizados publicamente e irão certamente beneficiar futuras pesquisas sobre a via de sinalização Wnt.

Palavras-chave: Biologia computacional. Neoplasias da Mama. Via de sinalização Wnt. Desenho de drogas.

ABSTRACT

Breast cancer is the leading worldwide cause of cancer-related deaths in women, being the triple-negative breast cancer type the one presenting the worst prognosis. The prospection and development of new substances with antitumoral potential is of great importance for the treatment of this disease. The objective of this work was to identify a commercial drug or ligand, through bioinformatics analysis, that could potentially bind to the FZD7 transmembrane protein and inactivate the Wnt signaling pathway in triple-negative breast cancer cells (TNBC). We aimed to computationally model the FZD7, Wnt3 and Wnt3a proteins, make them available in protein model databases, and conduct docking analysis to assess the binding free energy between FZD7 and the selected commercially available ligands. To select the proteins to be studied and modeled, a search for overexpressed TNBC proteins was done within the GEO database. The Wnt3 and Wnt3a proteins were modeled by homology modeling, using the Modeller software, and the FZD7 protein was modeled by homology modeling and *ab initio* modeling, using the I-TASSER online tool. Ramachandran plots were generated for the constructed models and RMSD analysis were conducted to assess their stereochemical quality. The ligands were selected based on their similarity to the palmitoleic acid and were gathered from the ZINC database. Docking analysis were conducted, using the Autodock Vina software, to assess the binding energy between the FZD7's cysteine-rich domain (CRD) and the selected ligands. The best generated models for Wnt3 and Wnt3a were deposited in the PMDB. A total of 30 commercially available ligands were found in the ZINC database. The docking results show that the ligands zinc08221009, zinc13546050, zinc05260769, zinc04529321, and zinc05972969 are good candidates for novel drug development. To the best of our knowledge, there are no available quality models of the Wnt3 and Wnt3a proteins in public databases. The created models and conducted analysis by the present work will most certainly help in future researches on the Wnt signaling pathway and its components.

Keywords: Breast cancer. Homology modeling. *Ab initio* modeling. Wnt signaling. Docking.

LISTA DE FIGURAS

Figura 1 - Via de sinalização Wnt.....	19
Figura 2 - Comparação entre as proteínas Wnt1 e Wnt3a de ratos, mostrando sítios de N-glicosilação preditos na literatura e sítios confirmados experimentalmente.....	21
Figura 3 - Estrutura do ácido palmitoleico.....	21
Figura 4 - Número de estruturas experimentais de proteínas disponibilizadas anualmente no PDB.....	23
Figura 5 - Esquema da sequência de passos para a modelagem por homologia.	25
Figura 6 - Representação esquemática de docking molecular.....	29
Figura 7 - Alinhamento de sequências entre o <i>template</i> pdb4f0a e as proteínas Wnt3 (a) e Wnt3a (b).....	37
Figura 8 - Resultados da modelagem por homologia da proteína Wnt3.....	39
Figura 9 - Resultados da modelagem por homologia da proteína Wnt3a.....	40
Figura 10 - Análise de ERRAT do modelo 49 de FZD7 gerado por homologia.....	41
Figura 11 - Resultados da modelagem <i>ab initio</i> da proteína FZD7.....	41
Figura 12 - Estrutura química dos 5 melhores ligantes apresentados na tabela 3.....	45
Figura 13 - Interação do CRD da proteína FZD7 (pdb5t44) com os melhores ligantes resultantes do docking molecular.....	46

LISTA DE TABELAS

Tabela 1 - Estoque atual do PDB. Fonte: https://www.rcsb.org/pdb/statistics/holdings.do (Fev. 2018).....	23
Tabela 2 - Valores de Lipinski dos 30 compostos retirados do banco de dados ZINC.....	42
Tabela 3 - Energia de ligação entre os 30 compostos retirados do banco de dados ZINC e o CRD da proteína FZD7 obtidos pelo AutoDock Vina e X-Score.....	44

LISTA DE ABREVIACOES

ALA	- Alanina
AN	- cidos nucleicos
ARG	- Arginina
ASP	- Asparagina
BL1	- Basal-like 1
BL2	- Basal-like 2
BLAST	- Basic Local Alignment Search Tool
BRCA1	- Gene breast cancer 1
CA	- Carbono alfa
CRD	- Cysteine rich domain
CYS	- Cistena
Dkk1	- Dickkopf-1
DM	- Dinmica Molecular
DNA	- Deoxyribonucleic acid
DVL2/3	- Segment polarity protein Disheveled homolog 2 e 3
DOPE	- Discrete Optimized Protein Energy
EMBOSS	- European Molecular Biology Open Software Suite
ER+	- Estrogen Receptor-Positive
ER-	- Estrogen Receptor-Negative
FASTA	- FAST-All
FZD	- Frizzled
GEO	- Gene Expression Omnibus
GLN	- Glutamina
GLU	- cido glutmico
GSK3 β	- Glycogen synthase kinase 3 beta
HBA	- Hydrogen Bond Acceptors
HBD	- Hydrogen Bond Donors
HER2	- Human Epidermal growth factor Receptor-type 2

HER2+	- Human Epidermal growth factor Receptor-type 2-Positivo
HER2-	- Human Epidermal growth factor Receptor-type 2-Negativo
HIV	- Human Immunodeficiency Virus
HIS	- Histidina
IM	- Imunomodulatório
kcal	- Quilocalorias
LAR	- Luminal Androgen Receptor
LEU	- Leucina
LRP6	- Lipoprotein receptor-related protein 6
LYS	- Lisina
M	- Mesenchymal
MCF-7	- Michigan Cancer Foundation-7
MET	- Metionina
MM-PBSA	- Molecular Mechanics Poisson-Boltzmann Surface Area
MSL	- Mesenchymal-like
NCBI	- National Center for Biotechnology Information
Log-P	- Log calculado P
PDB	- Protein Databank
PDBsum	- Protein Databank Summaries
PDBTM	- Protein Data Bank of Transmembrane Proteins
PHE	- Fenilalanina
PR+	- Progesterone Receptor-Positive
PR-	- Progesterone Receptor-Negative
PRO	- Prolina
PSI-Blast	- Position Specific Iterative Basic Local Alignment Search Tool
PTM	- Post-translational Modification
RMSD	- Root Mean Square Deviation
TASSER	- Threading/Assembly/Refinement
TCF/LEF	- T-cell factor/lymphoid enhancer-binding factor
TMTN	- Tumor Mamário Triplo-Negativo

THR	- Treonina
TRP	- Triptofano
VAL	- Valina
Wnt	- Wingless-related integration
ZINC	- Banco de dados de compostos disponíveis comercialmente

SUMÁRIO

1	INTRODUÇÃO	12
2	JUSTIFICATIVA	14
3	OBJETIVOS	15
3.1	Geral	15
3.2	Específicos	15
4.	REVISÃO DE LITERATURA	16
4.1	Câncer de mama	16
4.2	Via de sinalização Wnt-β-catenina	18
4.2.1	Wnt3a	20
4.3	Modelagem de proteínas	22
4.3.1	Modelagem por homologia	24
4.3.2	Limitações da modelagem por homologia	27
4.3.3	Modelagem <i>ab initio</i>	28
4.4	Docking molecular	29
5.	MATERIAIS E MÉTODOS	32
5.1	Proteína superexpressa em TMTN	32
5.2	Modelagem molecular	32
5.3	Avaliação da qualidade dos modelos	33
5.5	Pesquisa de ligantes	34
5.6	Docking molecular	34
6.	RESULTADOS	36
6.1	Modelagem molecular e aferição de qualidade	36
6.2	Pesquisa de ligantes e triagem virtual	42
6.3	Docking Molecular	44
7.	DISCUSSÃO	51
	REFERÊNCIAS	53
	APÊNDICE A - ARTIGO PUBLICADO COMO COAUTOR	62
	APÊNDICE B - MODELOS DEPOSITADOS NO PMDB	63

1 INTRODUÇÃO

O câncer de mama é a enfermidade de maior ocorrência na população feminina, é o segundo tipo de câncer mais frequente no mundo e representa mais de 20% de novos casos a cada ano. Apesar de ser considerado um câncer de, relativamente, bom prognóstico se diagnosticado precocemente, com taxas de cura que chegam a 90%, no Brasil os índices de mortalidade continuam elevados, principalmente devido ao diagnóstico tardio. Nos países desenvolvidos, a sobrevivência, a partir de 5 anos do tratamento da doença, é de 85%, enquanto nos países em desenvolvimento essa taxa cai para 70% (WORLD HEALTH ORGANIZATION, 2017).

Apesar da mesma origem tecidual, os carcinomas de mama são heterogêneos. Para o tratamento clínico são divididos em três classes principais: ER+ e PR+, que expressam receptores hormonais de estrógeno α e progesterona, respectivamente; HER2+, os que apresentam amplificação do gene HER2 (receptor de fator de crescimento epidérmico humano tipo 2) e o ER-, PR- e HER2-, triplo-negativo. Além destas, existem outras classes e subclasses de câncer de mama de acordo com o perfil de marcadores moleculares. Esses marcadores são comumente usados para definição do tratamento e estabelecimento do prognóstico da doença (PAYNE et al., 2008).

Os tumores de mama triplo-negativos (TMTN) não apresentam ainda alvo terapêutico molecular. Nesses tumores ocorre a superexpressão da via de sinalização Wnt canônica e não canônica, que contribuem para a transformação celular e progressão tumoral. Essa via é ativada por ligantes da proteína transmembrana FZD7, ativando a cascata de sinalização Wnt com consequente acúmulo de β -catenina no núcleo, resultando na ativação de fatores de transcrição que contribuem para o desenvolvimento de neoplasias (DEY et al., 2013).

A bioinformática, apesar de ainda muito nova, vem se mostrando de grande valor para a pesquisa, uma vez que proporciona análises rápidas e de muito mais baixo custo quando comparadas a análises em laboratório. Com os avanços em tecnologias e técnicas de *high throughput*, há muita geração de informação relevante que é armazenada em bancos de dados públicos, como o *Gene Expression Omnibus database* (GEO), *Protein Database* (PDB) e *ZINC database* (BARRETT et al., 2013; BERMAN et al., 2000; IRWIN et al., 2012). Essa informação, por ser massiva, requer técnicas de bioinformática para seu manuseio, interpretação, organização e análise. Assim, o presente trabalho pretende identificar por

análises de bioinformática, potenciais ligantes da proteína FZD7 disponíveis comercialmente. Espera-se que os ligantes escolhidos se liguem à proteína FZD7 e dificulte a ligação das proteínas Wnt3 e Wnt3a, consequentemente inibindo a via de sinalização.

2 JUSTIFICATIVA

O câncer de mama triplo-negativo ainda não apresenta alvo molecular para direcionar o tratamento, assim, faz-se necessária a busca de tal para a criação e desenvolvimento de tratamentos e terapias curativas. Há uma quantidade massiva de ligantes comerciais disponíveis que podem, potencialmente, serem reposicionados para o tratamento de doenças como câncer, alguns tendo ainda a vantagem de serem fármacos e anticorpos já testados em humanos e de terem sua dinâmica molecular conhecida. A proteína FZD7, apesar de ser sugerida por diversos estudos como marcador e alvo molecular para o câncer de mama triplo negativo, ainda não possui sua estrutura tridimensional nem a de seus ligantes disponível em bancos de dados, o que dificulta os estudos sobre a via de sinalização Wnt. Apesar desta via não se restringir à proteína FZD7, optou-se pela inibição limitada à mesma devido ao impacto negativo à homeostase do trato gastrointestinal que a inibição generalizada da via pode causar, como demonstrado por Kuhnert et al. (2004).

A falta de entendimento detalhado sobre as especificidades receptor-ligante de vários dos componentes da via de sinalização Wnt dificulta o desenvolvimento e descoberta de novos agentes terapêuticos. Dessa forma, fica evidenciada a necessidade da geração e disponibilização de tais estruturas para o avanço de estudos relacionados à seleção e criação de possíveis ligantes para inibição desta via, possibilitando o tratamento e cura da doença.

3 OBJETIVOS

Nesta seção do trabalho são apresentados o objetivo geral e específicos do mesmo.

3.1 Geral

Modelagem por homologia das proteínas FZD7, Wnt3 e Wnt3a, e identificação de um fármaco ou ligante disponível comercialmente, a partir de análises de bioinformática, que se ligue à proteína transmembrana FZD7 de forma a inativar a via de sinalização Wnt canônica em células de adenocarcinoma mamário humano triplo-negativo (TMTN).

3.2 Específicos

Como objetivos específicos, o presente trabalho pretendeu:

- a) Modelar computacionalmente as proteínas FZD7, Wnt3 e Wnt3a e disponibilizá-las em um banco de dados de modelos;
- b) Selecionar ligantes disponíveis comercialmente que potencialmente inativem a via de sinalização Wnt;
- c) Realizar análises de *docking* para aferir a energia de ligação entre os ligantes encontrados e a proteína FZD7.

4 REVISÃO DE LITERATURA

Nesta seção do trabalho é apresentada a revisão de literatura que embasou o desenvolvimento do mesmo.

4.1 Câncer de mama

O câncer de mama é o tipo de neoplasia maligna mais comum em mulheres (com exceção do câncer de pele não-melanoma), tanto nacional quanto internacionalmente. Em 2012, o câncer de mama foi, mundialmente, responsável por cerca de 1.700.000 novos casos e 522.000 mortes, sendo a maior causa de morte em países subdesenvolvidos e a segunda maior em países desenvolvidos, ficando atrás apenas do câncer de pulmão. Prevê-se que, de 2012 a 2017, haverá 6.200.000 de casos registrados de câncer de mama mundialmente (TORRE et al., 2016). Para o Brasil, foram estimados cerca de 58.000 novos casos de câncer de mama para 2016 (INSTITUTO, 2015; WHO, 2017).

Especificamente para o estado de Minas Gerais, em 2010 as neoplasias foram a segunda maior causa de morte, sendo responsáveis por 15,4% dos óbitos registrados, dos quais 6% (1.123 mortes) foram causados pelo câncer de mama feminino. A idade mediana de diagnóstico neste estado foi de 55 anos, sendo o intervalo entre o diagnóstico e o início do tratamento de 70 dias. Três a cada dez casos são diagnosticados em estágios já avançados da doença (estágios III e IV). Em pacientes diagnosticados no estágio 0 da doença, a sobrevivência de 5 anos é de até 92,7%; 87,8% em tumores em estágio I; entre 74% e 81,4% quando diagnosticado em estágio II; de 41% a 66,7% quando em estágio III e de 14,8% quando em estágio IV. O Instituto Nacional do Câncer José Alencar Gomes da Silva (INCA) estimou 57.960 casos de câncer de mama no ano de 2016 (BRASIL, 2017). Apesar do câncer de mama humano ter bom prognóstico quando diagnosticado em seus estágios iniciais, no Brasil, o número de mortes por esta doença continua aumentando (MINAS GERAIS, 2013; AMIN et al., 2017).

O câncer de mama é uma doença heterogênea, com diversas mutações gênicas, assim, marcadores moleculares específicos podem ser utilizados para auxiliar em seu diagnóstico. Esses marcadores, juntamente com informações como o tamanho atual do tumor, podem ser utilizados para categorizar as pessoas afetadas por essa doença em diferentes grupos de

prognóstico. Dentre esses marcadores destacam-se o receptor de estrógeno α (ER), o receptor de progesterona (PR) e o receptor de fator de crescimento epidermal humano tipo 2 (HER-2), sendo o primeiro o mais amplamente utilizado para diagnóstico, prognóstico e tratamento com terapia hormonal. Pacientes ER+ apresentam prognóstico e sobrevida melhores quando comparados a pacientes ER-, apesar do comportamento biológico dos dois grupos parecerem similares. Isto se dá pelo fato de pacientes ER+ poderem ser tratados por terapia hormonal com antiestrogênicos ou cirurgia endócrina ablativa, diferentemente dos ER- (PAYNE et al., 2008).

Neoplasias mamárias do tipo triplo-negativas (TMTN) são aquelas que não apresentam os marcadores moleculares: ER, PR, e HER-2, não havendo, então, um tipo de tratamento específico e eficiente para as mesmas. São, portanto, o tipo de neoplasias mamárias com o pior prognóstico atualmente. Devido a isso, são necessários avanços na pesquisa para o desenvolvimento de fármacos e tratamentos eficientes para esta doença (KUMAR; AGGARWAL, 2016).

Aproximadamente 15% dos tumores de mama diagnosticados são triplo-negativos, e cerca de 80% destes apresenta mutação no gene BRCA1, o que piora o prognóstico devido à perda de função no reparo de mutações do DNA. Os TMTN podem ainda expressar a proteína p53 mutante e baixos níveis de claudina, proteína responsável pela junção oclusiva em células epiteliais. Afetam preferencialmente mulheres jovens afro-americanas e estão associados a um pior prognóstico, com alta taxa de proliferação, invasão e recorrência da doença, além da maior incidência de metástases hematogênicas para pulmões e cérebro, pois, normalmente, respondem muito pouco aos tratamentos disponíveis. A taxa de recorrência média, em TMTN, é de 3 anos e a probabilidade de morte por metástases em 5 anos é três vezes maior que no caso de tumores ER+ (BERTUCCI; FINETTI; BIRNBAUM, 2012; LEHMANN et al., 2011; MASILI-OKU et al., 2016).

Apesar das características histológicas semelhantes e de não expressarem ER, PR e HER2, os TMTN são um grupo heterogêneo em termos moleculares e podem ainda ser subdivididos em 6 diferentes subtipos de acordo com o perfil de expressão gênica em: *basal-like*, incluindo *basal-like 1* (BL1), *basal-like 2* (BL2) e imunomodulatório (IM); *mesenquimal-like*, incluindo os subtipos mesenquimal (M) e tronco *mesenquimal-like* (MSL) e o subtipo receptor de andrógeno luminal (LAR) (LEHMANN et al., 2011).

No Brasil há poucos estudos com casuísticas avaliando a frequência de TMTN, assim como a expressão de diferentes marcadores nestes tumores e a evolução clínica dos indivíduos acometidos pela doença (CINTRA et al., 2012; CARVALHO et al., 2010; DE BROT et al., 2009; DUTRA et al., 2009; LOPES et al., 2015; MASILI-OKU et al., 2016).

O processo de metástase no câncer de mama contribui enormemente à sua taxa de mortalidade (AHMED; ALAWIN; SYLVESTER, 2016). Tal processo envolve a perda de polaridade celular e adesão, acarretando no ganho de mobilidade e invasão pelas células cancerígenas. Células epiteliais metastáticas podem apresentar um fenótipo de células mesenquimais e a esse processo é dado o nome de transição epitelial-mesenquimal, o que piora o prognóstico da doença (SAVAGNER, 2010).

4.2 Via de sinalização Wnt- β -catenina

A via de sinalização Wnt- β -catenina é uma via dependente de ligantes que regula decisões de transformação celular durante o desenvolvimento embrionário e pode levar ao desenvolvimento e progressão de neoplasias devido ao acúmulo da proteína β -catenina no núcleo da célula quando desregulada, sendo essa proteína um fator amplificador de transcrição. A proteína transmembrana *Frizzled 7* (FZD7) atua como transdutor de sinal para iniciar a cascata de β -catenina, podendo ser ativada pelos ligantes Wnt3, Wnt3a e Wnt5 (*Wingless-related integration*). Sem a ativação desta via, a β -catenina se liga ao complexo de axina e é constantemente degradada. Ao ser ativada, entretanto, a β -catenina não degradada se transloca para o núcleo da célula e ativa genes de fatores de transcrição, o que resulta no desenvolvimento de neoplasias como o TMTN (DEY et al., 2013). O domínio rico em cisteína (“*cysteine-rich domain*” ou CRD) é a parte extracelular da proteína FZD7 que entra em contato direto com os ligantes Wnt para ativação da via, sendo um alvo em potencial para inativação da via.

A sinalização é iniciada pela ligação das proteínas endógenas Wnt3, Wnt3a e Wnt5a/b, superexpressas em TMTN, ao complexo de ativação composto pelas proteínas transmembrana FZD7 e o co-receptor LRP6. Essa ligação recruta e ativa proteínas *dishevelled* (DVL2/3) que por sua vez inibem GSK3 β . A inibição de GSK3 β impede a fosforilação e degradação da β -catenina, aumentando sua disponibilidade e translocação para o núcleo. Uma vez no núcleo, a β -catenina se ligará ao fator de transcrição celular-T/fator intensificador linfóide (TCF/LEF) o que levará à expressão de genes que promovem o desenvolvimento de câncer, como c-Myc e ciclina D1 (FIGURA 1) (AHMED; ALAWIN; SYLVESTER, 2016).

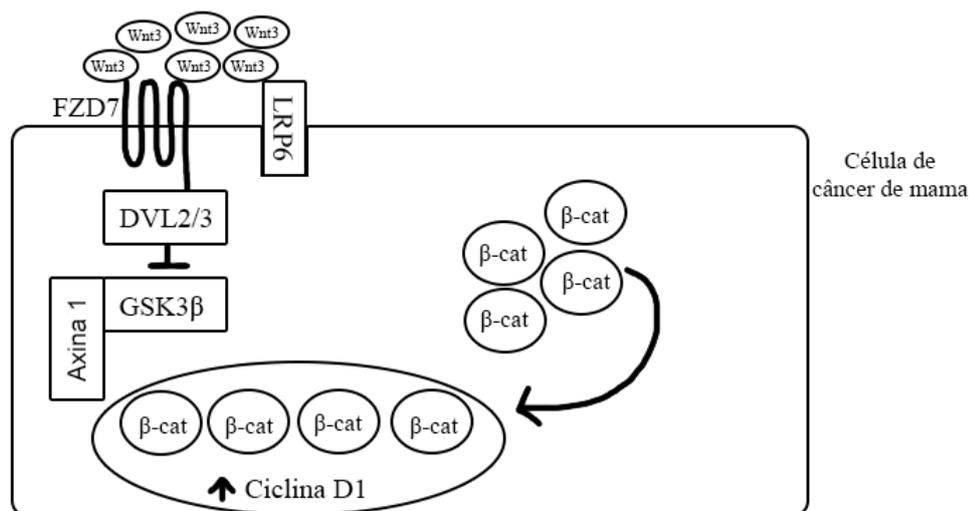


Figura 1 - Via de sinalização Wnt.

Fonte: Adaptado de Ahmed, Alawin e Sylvester (2016).

Como comprovado por Dey et al. (2013), há a hiper-regulação da via Wnt-β-catenina em células TMTN quando comparadas a modelos de subtipos ER+ e HER2+. Foi ainda observado, no mesmo trabalho, que alguns componentes-chave da via, incluindo mRNA de β-catenina, são mais expressos em células TMTN do que em células não triplo-negativas como células MCF-7. Espera-se, então, que ligantes que afetem a ativação da via Wnt-β-catenina *upstream* apresentem efeitos reduzidos em células não triplo-negativas quando comparados aos efeitos em células TMTN. A interação entre a proteína Wnt3 e a proteína transmembrana FZD7 também está ligada à hepatocarcinogênese (KIM et al., 2008).

A relação entre a via de sinalização Wnt canônica e o desenvolvimento e progressão de neoplasias tem sido estudada há mais de 30 anos, partindo da observação feita por Nusse e Varmus (1982) de que a falha na expressão do gene Wnt1 poderia levar ao desenvolvimento de câncer de glândula mamária em camundongos. Desde então, diversos estudos foram realizados com o intuito de inativar a via de sinalização da β-catenina, tanto canônica como não canônica. Nambotin et al. (2012) conseguiram inativar a ligação da proteína FZD7 ao domínio PDZ da proteína DVL, no carcinoma hepatocelular, utilizando pequenos peptídeos de interferência, com consequente diminuição da proliferação celular por ativação da apoptose. Apesar do sucesso em sua abordagem *downstream*, os pesquisadores sugerem que estratégias *upstream* para a inativação da via, como a inibição da ligação da Wnt ao receptor FZD, possam ser mais eficazes pelo seu potencial impacto não somente na via canônica como também na não canônica.

4.2.1 Wnt3a

A família Wnt de proteínas é um grupo de glicoproteínas extracelulares ricas em cisteína que regulam decisões de especificação celular durante o desenvolvimento embrionário e em tecidos adultos (MACDONALD; TAMAI; HE, 2009). O genoma de mamíferos codifica 19 diferentes tipos de Wnt de tamanhos entre 350-400 aminoácidos e com um padrão de 23 ou 24 cisteínas presentes e conservadas (TANAKA; KITAGAWA; KADOWAKI, 2002). Ainda existe pouca informação estrutural sobre as Wnts, e sua sequência primária não se assemelha a nenhum enovelamento de proteína conhecido (JANDA et al., 2012). Presume-se que todas as proteínas Wnt presentes em mamíferos apresentem dois tipos de modificações pós-traducionais (PTMs): a ligação covalente de ácidos graxos em duas posições distintas, o que torna a molécula hidrofóbica (TAKADA et al., 2006; WILLERT et al., 2003); e a N-glicosilação de múltiplas asparaginas (FIGURA 2).

O gene da proteína Wnt3a se encontra no genoma humano no cromossomo 1q42 (KIRIKOSHI; SEKIHARA; KATOH, 2001). A superexpressão desse gene leva à ativação irregular da via de sinalização Wnt canônica, com consequente acúmulo de β -catenina no núcleo da célula e o desenvolvimento e progressão de neoplasias. A proteína Wnt3a também participa dos processos de proliferação e diferenciação celular em vários tipos de células tronco e está ligada à proliferação, migração e invasão de vários tipos celulares (POLLHEIMER et al., 2006; GUJRAL; MACBEATH, 2010; NAGAOKA et al., 2013). O gene Wnt3a é mais expresso na maioria das células de tumores mamários quando comparado a células mamárias normais (BENHAJ; AKCALI; OZTURK, 2006), o que também acontece no carcinoma oral de células escamosas, quando comparado a tecidos normais (ANDRADE, 2011).

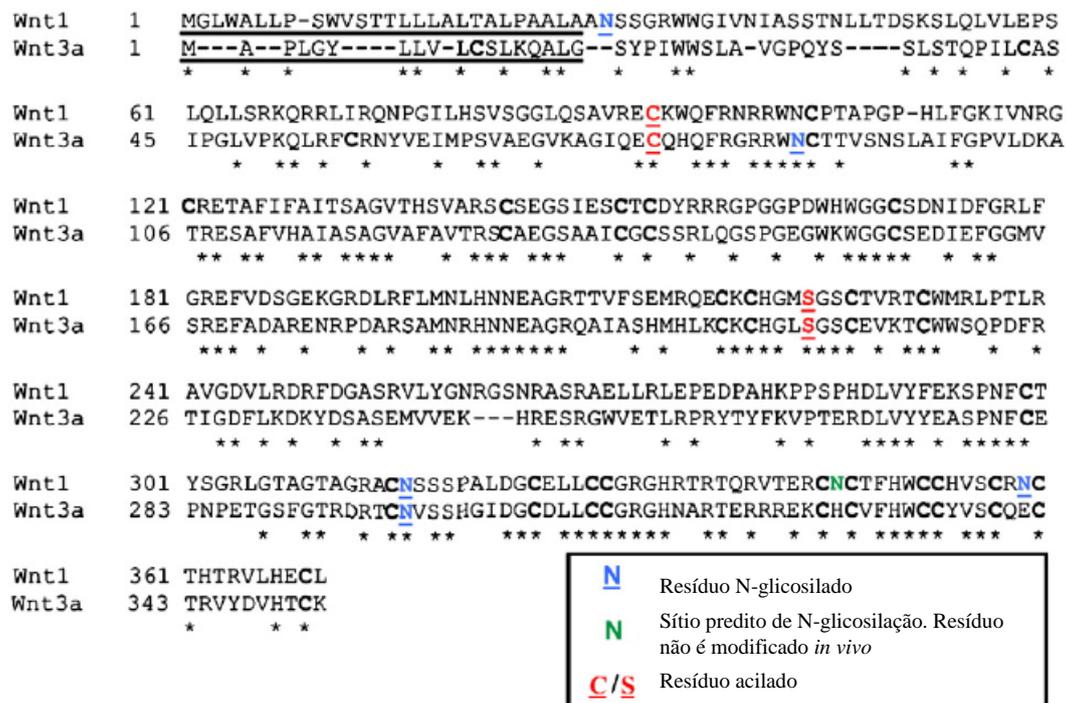


Figura 2 - Comparação entre as proteínas Wnt1 e Wnt3a de ratos, mostrando sítios de N-glicosilação preditos na literatura e sítios confirmados experimentalmente. Notas: Os peptídeos de sinalização putativa estão sublinhados. Aminoácidos conservados entre as duas proteínas estão marcados por asteriscos. Os resíduos de cisteína e serina modificados estão destacados e cisteínas não modificadas se encontram em negrito. Fonte: Adaptado de Doubravska et al, 2011.

A proteína Wnt3a madura presente em ratos apresenta uma modificação por ácido palmítico no resíduo de cisteína 77 (C77). Há também uma modificação pós-traducional na serina 209, onde se liga uma molécula de ácido palmítico (FIGURA 3) (DOUBRAVSKA et al. 2011). De acordo com Doubravska et al. (2011), apenas a modificação presente na serina 209 é crucial para a sinalização com o receptor FZD7 e desencadeamento da cascata da β -catenina.

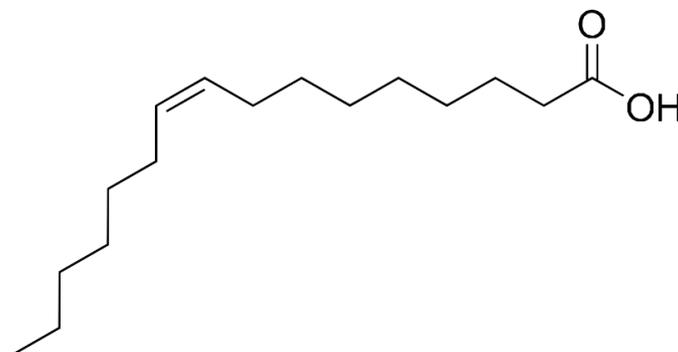


Figura 3 - Estrutura do ácido palmítico. Fonte: Do autor.

Já existem estudos que demonstram a inibição da sinalização por Wnt3a por algumas proteínas e substâncias químicas (HE et al., 2015). A proteína plasmática kallistatina se liga ao co-receptor LRP6 da via de sinalização Wnt e antagoniza a via, inibindo angiogênese, inflamação e crescimento tumoral em células mamárias triplo-negativas MDA-MB-231 (ZHANG et al. 2013). O anti-helmíntico niclosamida e o composto natural silibinina inibem a expressão do co-receptor LRP6 em vários tipos celulares, incluindo células MDA-MB-231 de câncer de mama triplo-negativo, impedindo a ativação da via de sinalização Wnt e o acúmulo de β -catenina no núcleo celular (LU et al. 2011; LU et al. 2012). Entretanto, o estudo de Kuhnert et al. (2004) indica que a inibição da via Wnt canônica através de interações com o co-receptor LRP6 podem implicar na perda da homeostasia do trato gastrointestinal, de acordo com suas observações utilizando o inibidor Dickkopf-1 (Dkk1) em ratos adultos.

4.3 Modelagem de proteínas

Informações sobre a estrutura tridimensional de proteínas são de valor inestimável no estudo de suas funções e interações moleculares, tal qual o estudo de vias de sinalização celulares, e estão intimamente ligadas à conformação estrutural das proteínas envolvidas. Estudos envolvendo mecanismos moleculares, em geral, tem seu desenho experimental baseado e facilitado por informações detalhadas sobre o arranjo de átomos e resíduos de aminoácidos de seus alvos. Conhecimento da estrutura tridimensional de proteínas chave no desenvolvimento de doenças podem ser o primeiro passo para o desenvolvimento de tratamentos e terapias direcionados.

Pesquisas envolvendo características estruturais de proteínas comumente utilizam estruturas determinadas por métodos experimentais, como cristalografia de raio-X, microscopia eletrônica de alta resolução e ressonância nuclear magnética. Apesar do grande investimento e rápido desenvolvimento dessas técnicas, elas ainda continuam sendo processos caros, demorados e nem sempre aplicáveis. De acordo com o banco de dados de proteínas PDB (BERMAN et al. 2000), existem atualmente em seu acervo 127.641 estruturas experimentais de proteínas (TABELA 1), entretanto, apesar de seu crescimento exponencial (FIGURA 4),

esse número ainda é muito pequeno quando comparado às mais de 550 mil sequências de proteínas avaliadas por curadores no Swiss-Prot do UniProtKB, uma seção do banco de dados sobre informações funcionais de proteínas (BOUTET et al., 2016). O acervo do PDB se mostra ainda menor quando comparado à totalidade de estruturas de proteínas conhecidas, mas não ainda totalmente avaliadas, presente no UniProtKB que chega a mais de 108 milhões (fevereiro de 2018).

Tabela 1 - Estoque atual do PDB.

		Proteínas	Ácidos nucleicos (AN)	Complexos de proteína/AN	Outros	Total
Método	Raio-X	115258	1909	5879	4	123050
Experimental	NMR	10633	1234	247	8	12122
	ME	1437	30	497	0	1964
	Híbrido	107	3	2	1	113
	Outro	206	4	6	13	229
	Total	127641	3180	6631	26	137478

Fonte: <https://www.rcsb.org/pdb/statistics/holdings.do> (Fev. 2018).

Aumento do acervo de estruturas do PDB

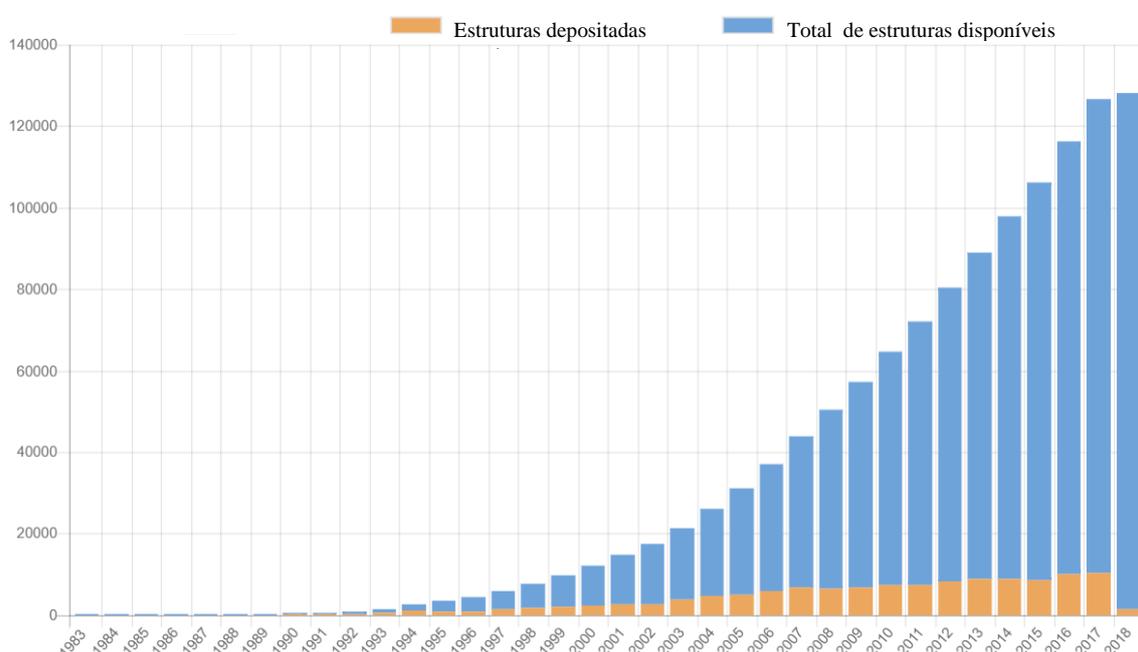


Figura 4 - Número de estruturas experimentais de proteínas disponibilizado anualmente no PDB.

Fonte: <https://www.rcsb.org/stats/growth/protein> (Fev. 2018).

A modelagem computacional se mostra como a melhor opção para cobrir essa diferença entre o número de sequências de proteínas disponível e a quantidade de informação estrutural

das mesmas. Existem duas abordagens principais para a modelagem computacional de proteínas, a modelagem comparativa (ou modelagem por homologia) e a modelagem *ab initio* (ou modelagem *de novo*).

4.3.1 Modelagem por homologia

A modelagem baseada em *templates*, ou modelagem por homologia, é fundamentada na proximidade evolutiva entre a proteína a ser estudada e seus *templates*, que possuem estrutura tridimensional determinada experimentalmente. Uma sequência de passos é seguida para a construção de modelos através desta abordagem, e esta pode ser repetida para melhorar a qualidade dos modelos gerados. Os passos seguidos podem ser simplificados como: 1) identificação de um ou mais *templates* compatíveis com a sequência primária da proteína alvo; 2) alinhamento entre a sequência primária da proteína alvo e a sequência primária do(s) *template(s)* 3) modelagem das regiões estruturalmente conservadas entre a proteína alvo e seu(s) *template(s)*; 4) refinamento do melhor modelo gerado; e 5) avaliação da qualidade do modelo ou modelos finais (FIGURA 5).

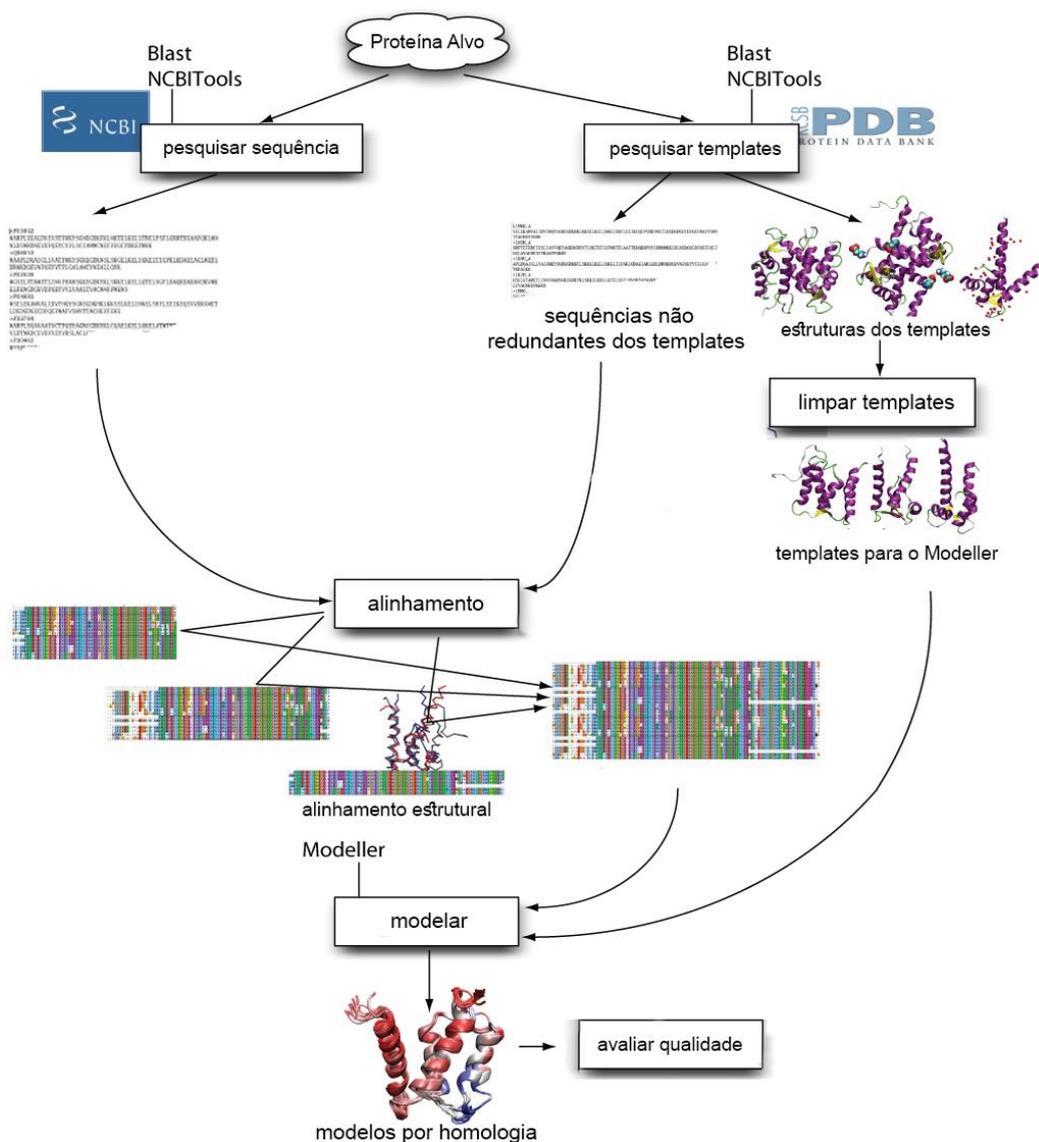


Figura 5 - Esquema da sequência de passos para a modelagem por homologia.
 Fonte: <https://www.rcsb.org>

Para a identificação dos *templates*, são escolhidas preferencialmente proteínas com estrutura tridimensional de alta resolução ($\leq 2\text{\AA}$) definida experimentalmente e que tenham ao menos 40% de identidade de sequência com a proteína alvo. A identidade de sequência é um bom indicador inicial da qualidade do modelo a ser gerado, uma vez que modelos baseados em alinhamentos com baixa identidade de sequência (abaixo da faixa entre 30% e 40%) podem apresentar enovelamento incorreto (MELO; SALI, 2007). O alinhamento de sequências consiste na comparação e análise de correspondências na sequência de aminoácidos entre duas (alinhamento de sequências par a par) ou mais proteínas (alinhamento múltiplo de sequências). Para a detecção da homologia entre as sequências (alvo e *template*) são utilizadas ferramentas

como o *Position Specific Iterative Basic Local Alignment Search Tool* (PSI-Blast), que realiza uma busca de similaridade entre a sequência primária de sua proteína alvo e proteínas com estrutura tridimensional definida experimentalmente em um ou mais bancos de dados de proteínas (ALTSCHUL, 1997).

Com a sequência primária da proteína alvo e um ou mais *templates* em mãos, é então iniciada a modelagem. Existem softwares especializados na construção de modelos de proteína por homologia como o Modeller (SALI; BLUNDELL, 1993) e o SWISS-MODEL (ARNOLD et al., 2006). O software Modeller utiliza a sequência a ser modelada e seus *templates* para calcular uma série de estruturas plausíveis contendo todos os átomos exceto os de hidrogênio. Em seguida, são feitas tentativas de otimização de energia com base na satisfação de restrições espaciais obtidas do alinhamento da sequência alvo e seu(s) *template(s)* e da mecânica molecular de campos de força. O software Modeller deriva distâncias e restrições de ângulos diedros no alinhamento da sequência alvo com a estrutura tridimensional definida do seu *template*. As restrições espaciais na sequência alvo são obtidas da análise estatística das relações entre várias características da estrutura da proteína. As restrições derivadas da estrutura do *template* para a composição do conjunto de restrições total do modelo, violando a própria estereoquímica, compõem a função objetivo. Finalmente, o modelo é obtido pela otimização da função objetivo no espaço cartesiano. Diversos modelos ligeiramente diferentes podem ser calculados variando a estrutura inicial.

A confiabilidade do modelo e sua interpretação estão diretamente ligadas à qualidade do mesmo. A avaliação da qualidade do modelo tem também como objetivo identificar possíveis problemas que podem ter surgido durante sua modelagem. Uma combinação de métodos de avaliação é geralmente utilizada para selecionar o melhor modelo gerado durante a modelagem. Um dos requerimentos mínimos para a boa qualidade de um modelo é que o mesmo tenha uma boa qualidade estereoquímica (MELO; SALI, 2007). O software PROCHECK (LASKOWSKI et al., 1993) é um dos mais utilizados para análises de estereoquímica. Este software avalia características do modelo como comprimento de ligação, ângulos de ligação, ligação peptídica e planaridade de anéis da cadeia lateral, quiralidade, ângulos de torção da cadeia principal e cadeia lateral e choques entre pares de átomos não ligados. Seus resultados podem ser apresentados em forma de um gráfico de Ramachandran, que mostra a porcentagem de resíduos em regiões favoráveis, regiões adicionais favoráveis, regiões generosamente favoráveis e regiões não permitidas. Um modelo de alta qualidade não apresenta resíduos em regiões não permitidas. Existem também softwares, como VERIFY3D (EISENBERG; LÜTHY; BOWIE, 1997) e ERRAT (MACARTHUR; LASKOWSKI;

THORNTON, 1994), que avaliam o ambiente químico de cada resíduo do modelo construído comparado ao ambiente químico esperado de se encontrar em proteínas de estrutura tridimensional definida experimentalmente.

Outro parâmetro utilizado para avaliar a qualidade de modelos gerados por homologia é a análise de seu RMSD (*Root Mean Square Deviation*). O RMSD avalia o valor do desvio médio entre os átomos da estrutura do modelo e os átomos da estrutura do *template*. Quanto maior o valor de RMSD, maior é a diferença estrutural entre as duas proteínas. Um modelo de boa qualidade gerado por homologia é geralmente baseado em uma identidade de sequência entre a sequência alvo e seu *template* de 40-50%. Estes modelos tendem a ter aproximadamente 90% da cadeia principal modelada com um RMSD de 1,5 Å. Modelos de proteínas com identidade de sequência menor que 30% são geralmente de baixa qualidade.

4.3.2 Limitações da modelagem por homologia

A modelagem comparativa, ou modelagem por homologia, se baseia e depende da relação evolutiva entre a proteína alvo e seu(s) *template(s)*, partindo da observação de que sequências evolutivamente relacionadas geralmente possuem estruturas tridimensionais similares (CHOTHIA; LESK, 1986). Isso, entretanto, traz consigo desvantagens e limitações para esse tipo de modelagem. Uma das maiores limitações é a necessidade da disponibilidade de estruturas *templates* definidas experimentalmente que tenham uma boa identidade de sequência (> 30%) com sua proteína alvo. Apesar de existirem estudos que argumentem que a maioria das estruturas conhecidas apresentam um número limitado de conformações tridimensionais (ORENGO; THORNTON, 2005) e que, portanto, para a maioria das proteínas alvo já haveria uma estrutura *template* disponível no PDB (ZHANG; SKOLNICK, 2005), uma boa identidade de sequência continua sendo crucial para prever a função e especificidade da proteína modelada.

Proteínas de membrana participam de diversos processos celulares, incluindo sinalização, comunicação intercelular, translocação de proteínas e transporte de vesículas. Devido a isso, são alvos terapêuticos de muitos fármacos que utilizamos atualmente. Existe, entretanto, uma enorme dificuldade em se definir experimentalmente suas estruturas. Isso se traduz em uma baixa disponibilidade de estruturas tridimensionais determinadas experimentalmente no PDB, contando com apenas 3227 proteínas (fevereiro de 2018) no banco

de dados de proteínas transmembrana (PDBTM), aproximadamente 2,5% do total de proteínas disponíveis (KOZMA; SIMON; TUSNÁDY; 2013). A modelagem de proteínas transmembrana por homologia é, portanto, limitada pela baixa quantidade de estruturas que podem ser utilizadas como *template*.

4.3.3 Modelagem *ab initio*

A modelagem por homologia só é capaz de gerar modelos de proteínas que tenham *templates* com boa identidade de sequência e informação estrutural disponível. Portanto, partes da sequência alvo que não se alinham com a sequência do *template* não são modeladas adequadamente. A modelagem *de novo*, ou *ab initio*, não se limita a utilizar estruturas completas como *template*, mas faz uso de funções de energia que se baseiam em interações físico-químicas entre os átomos. Esta abordagem é considerada a mais difícil para a predição de estrutura de proteínas (FLOUDAS et al., 2006). A modelagem *de novo*, entretanto, se limita à modelagem de pequenas proteínas, ou de fragmentos de proteínas, de até 150 resíduos geralmente. Proteínas grandes são extremamente difíceis de se modelar com esta abordagem (ZHANG; SKOLNICK, 2004a). Esta técnica se baseia na asserção de que, em seu estado nativo, proteínas se encontram em seu mínimo de energia livre global, e então realiza buscas extensivas de espaço conformacional de estruturas terciárias de proteínas que apresentem baixa energia livre.

O método TASSER (*Threading/ASSEMBly/Refinement*) busca por estruturas no PDB que sejam representativas de fragmentos locais da sua sequência alvo, e então utiliza restrições de estrutura terciária para restringir a busca do espaço conformacional para construção do modelo (ZHANG; SKOLNICK, 2004b). Os modelos são gerados ao conectar as estruturas secundárias dos fragmentos gerados, que passam então por refinamento pelo método de amostragem de Monte Carlo. O método I-TASSER, que implementa o método TASSER de forma iterativa (WU; SKOLNICK; ZHANG, 2007), é um dos métodos mais utilizados para a construção de modelos pela abordagem *ab initio*.

4.4 Docking molecular

Análises de *docking* (ou ancoramento molecular) são simulações que calculam a estrutura do complexo e a energia livre de ligação entre duas proteínas, ou entre uma proteína e um ligante. A simulação computacional de *docking* visa estimar a posição e conformação de um ligante dado um sítio de ligação predito ou pré-determinado de uma proteína. O cálculo da energia livre de ligação de complexos proteína-ligante é considerado o objetivo principal do desenho computacional de fármacos, sendo que menores valores de energia livre significam uma maior afinidade entre as moléculas estudadas. O *docking* é atualmente utilizado como uma ferramenta para otimização de compostos líderes e para estudos de triagem virtual em busca de moléculas biologicamente ativas (DIAS; DE AZEVEDO, 2008).

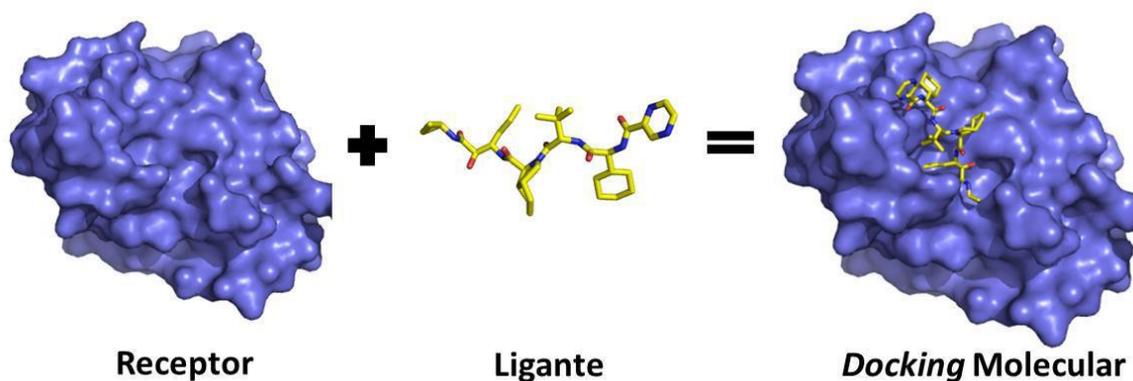


Figura 6 - Representação esquemática de *docking* molecular.
Fonte: Adaptado de Santoyo et al., 2013.

A metodologia de *docking* consiste em 3 passos conectados entre si: predição de poses moleculares, triagem virtual e cálculo da energia livre de ligação (JAIN; NICHOLLS, 2008). Tanto o receptor, quanto o ligante, possuem flexibilidades intrínsecas, implicando no tratamento de milhares de graus de liberdade para ambos pelo algoritmo de *docking* utilizado. Algoritmos de busca são utilizados para explorar o panorama energético e encontrar as melhores poses para o ligante. Em teoria, o mínimo global do panorama energético corresponde à conformação experimental do complexo receptor-ligante, entretanto, devido ao fato de que algoritmos de *docking* trabalham com aproximações, há a chance de o mínimo global não representar, de fato, a pose de ligação nativa. O resultado final de um estudo bem-sucedido de

docking deve ser a predição aproximada da pose nativa do ligante em contato com o sítio de ligação do seu receptor e suas consequentes interações físico-químicas.

Um software de *docking* é composto pela combinação de um algoritmo de *docking* e uma função de *scoring*, que geralmente variam de software para software. A função de *scoring* se baseia em dados experimentais estruturais de complexos e em suas afinidades de ligação e então, com base nestes, os compara aos dados obtidos pelo algoritmo de *docking* para avaliar e ranquear as melhores poses obtidas do ligante.

Quanto à flexibilidade do receptor e ligante, 3 estratégias podem ser utilizadas no *docking*. A primeira considera o receptor como rígido, e apenas os graus de liberdade translacionais e rotacionais do ligante são explorados, ignorando seus graus de liberdade conformacionais (portanto o ligante é considerado como uma molécula rígida sem graus de liberdade internos). Na segunda, o ligante também é considerado rígido e todos os graus de liberdade do ligante (rotacionais, translacionais e conformacionais) são considerados (*docking* semi-flexível). Na terceira estratégia, o receptor é considerado como flexível ou como parcialmente flexível e todos os graus de liberdade do ligante são explorados (*docking* flexível). A demanda computacional aumenta com o número de graus de liberdade explorados, sendo então a flexibilidade, tanto do receptor quanto do ligante, fatores limitantes nas análises de *docking* (GUEDES; MAGALHÃES; DARDENNE, 2014).

Um dos programas mais citados na literatura para análises de *docking* é o AutoDock Vina, que foi o software utilizado neste trabalho. O AutoDock Vina é amplamente utilizado devido à sua facilidade de uso e velocidade, sendo cerca de 100 vezes mais rápido que seu predecessor, o AutoDock 4, também desenvolvido pelo Scripps Research Institute (TROTT; OLSON, 2010). No AutoDock Vina apenas a flexibilidade do ligante é considerada, sendo o receptor considerado como rígido. O algoritmo de busca utilizado por esse software é similar ao do tipo genético, onde uma pontuação é dada a cada membro de um conjunto de poses aleatórias geradas inicialmente para o ligante. Essa pontuação é baseada na energia de interação entre aquela pose e o receptor e também na energia intramolecular que o ligante apresenta na mesma. As poses com menores pontuações são descartadas e as demais sofrem pequenas modificações em seus ângulos e são então ranqueadas novamente. Esse processo se repete até que uma determinada pontuação seja alcançada (GOLDBERG, 1989). A função de *scoring* do AutoDock Vina se baseia na seguinte fórmula (1):

$$c = \sum_{i < j} f_{titj}(r_{ij}) \quad (1)$$

onde o valor c é a soma das contribuições energéticas intramoleculares e intermoleculares e a somatória é feita com todos os pares de átomos que podem se mover em relação uns aos outros. O objetivo final do algoritmo de otimização deste programa é encontrar o mínimo global do valor c e outras poses também de baixo *score* e então ranqueá-las (TROTT; OLSON, 2010). Quanto menor o valor alcançado, maior é a afinidade entre o ligante e o receptor.

5 MATERIAIS E MÉTODOS

Nesta seção do trabalho as técnicas e metodologias aplicadas durante seu desenvolvimento são apresentadas.

5.1 Proteína superexpressa em TMTN

Inicialmente, foi feita uma busca na *GEO Database* (BARRETT et al., 2013), um banco de dados público de expressão gênica, por genes superexpressos em células TMTN. Tal pesquisa resultou na escolha da proteína Frizzled 7 (FZD7) de registro GDS4069 associado ao trabalho de Yang et al. (2011). A proteína FZD7 participa da via de sinalização Wnt canônica em conjunto com Wnt3 e Wnt3a (YANG et al., 2011).

5.2 Modelagem molecular

A sequência primária das proteínas FZD7, Wnt3 e Wnt3a, obtidas do UniProt (UNIPROT, 2007), de códigos O75084, P56703 e P56704 respectivamente, foram utilizadas para pesquisar por suas estruturas experimentais tridimensionais no banco de dados de proteínas PDB (BERMAN et al., 2000). Não há, atualmente, estruturas experimentais disponíveis em bancos de dados referentes à sequência completa das proteínas pesquisadas. A sequência primária das proteínas foi então utilizada na realização de um PSI-BLAST (*Position Specific Iterated – Basic Local Alignment Search Tool*) (NCBI) contra o banco de dados PDB, de forma a encontrar proteínas homólogas a tais sequências. Os alinhamentos globais entre a sequência das proteínas alvo e seus respectivos *templates* foram realizados pela ferramenta online EMBOSS *Needle* (LI et al., 2015).

As proteínas homólogas encontradas pelo PSI-BLAST foram utilizadas como *templates* para a modelagem por homologia utilizando o software Modeller (SALI; BLUNDELL, 1993). Para a modelagem, foram utilizados *scripts* disponibilizados pelo tutorial do programa em seu site oficial. Foram gerados 100 modelos para a proteína Wnt3 e 100 para a proteína Wnt3a

utilizando apenas um *template*, e 100 modelos para Wnt3 e 100 Wnt3a utilizando múltiplos *templates*. Para a proteína FZD7, 100 modelos foram construídos utilizando múltiplos *templates*. Devido à baixa qualidade dos modelos gerados para FZD7, recorreu-se à modelagem *ab initio* com o auxílio da ferramenta online I-TASSER (YANG et al., 2015). O software PyMOL (DELANO, 2015) foi utilizado para auxiliar na edição e visualização dos *templates* e dos modelos gerados, bem como no cálculo de RMSD entre os modelos gerados e seus respectivos *templates*.

5.3 Avaliação da qualidade dos modelos

Os modelos construídos por homologia foram ranqueados com base em seus potenciais estatísticos DOPE (*Discrete Optimized Protein Energy*) (SHEN; SALI, 2006) gerados pelo Modeller, sendo considerado como melhor o modelo que apresentou o valor mais negativo. DOPE é um potencial estatístico dependente de distâncias atômicas, é calculado a partir de uma amostra de estruturas nativas de proteínas e é embasado na teoria probabilística. Análises de ERRAT e VERIFY3D foram conduzidas para avaliar a distribuição dos átomos e o perfil 3D dos modelos, respectivamente. Para análises de ERRAT, uma qualidade geral $\geq 80\%$ é considerada um bom resultado. Em análises de VERIFY3D, um bom resultado é definido com 80% ou mais dos resíduos apresentando um *score* médio 3D-1D $\geq 0,2$.

Os melhores modelos foram avaliados quanto a sua qualidade estereoquímica e tiveram seus gráficos de Ramachandran gerados com o auxílio da ferramenta online PROCHECK (LASKOWSKI et al., 1993) através do banco de dados PDBsum (DE BEER et al., 2014). O gráfico, ou diagrama, de Ramachandran é um gráfico de ângulos diedros ϕ versus ψ dos resíduos de aminoácido de uma proteína, permitindo a visualização de regiões energeticamente permitidas para tais ângulos por resíduo. Esse gráfico considera átomos como esferas rígidas com dimensões correspondentes ao seu raio de van der Waals. Ângulos de torção ϕ e ψ que causem a colisão entre tais esferas correspondem às regiões não permitidas no gráfico (RAMACHANDRAN; RAMAKRISHNAN; SASISEKHARAN, 1963). Foram escolhidos como melhores os modelos que apresentaram a menor porcentagem de resíduos em regiões não permitidas no gráfico de Ramachandran e com o menor valor de RMSD em relação aos seus *templates* calculado pelo programa PyMOL (DELANO, 2015).

5.5 Pesquisa de ligantes

Para a triagem virtual de ligantes foi utilizado o banco de dados ZINC (IRWIN et al., 2012). Foram buscados compostos comercialmente disponíveis que apresentassem pelo menos 99% de similaridade com o ácido palmitoleico, por meio do coeficiente de Tanimoto (WILLETT, 2006). Foi feita ampla pesquisa na literatura de potenciais ligantes análogos à Wnt3a, de forma a impedir ou reduzir a ligação do ligante endógeno Wnt3a à proteína FZD7. Passou-se então à pesquisa por anticorpos, com estrutura definida e disponível, e com afinidade à proteína FZD7. Os melhores candidatos retornados pelo banco de dados ZINC foram então selecionados para as análises de *docking*.

5.6 *Docking* molecular

O *docking* foi realizado entre os ligantes encontrados pela triagem virtual e o domínio rico em cisteína (CRD) da proteína FZD7, representado pela proteína de código pdb5t44 de resolução 1,9 Å (NILE et al., 2017). A metodologia de *docking* utilizada foi a do tipo rígido-flexível com o auxílio do software AutoDockTools (MORRIS et al. 2009), para visualização gráfica do processo e edição dos arquivos da proteína e ligantes, e Autodock Vina (TROTT; OLSON, 2010) para o *docking* em si. A caixa de *docking* foi definida visualmente baseando-se na região de ligação entre a proteína XWnt8 e o CRD da proteína FZD8 do *template* pdb4f0a (JANDA et al., 2012), utilizado na modelagem por homologia da proteína Wnt3a. As dimensões da caixa foram 22 x 24 x 26, com o centro x em -12,207, o centro y em -14,689 e o centro z em 15,32. O ácido palmitoleico (ZINC08221009) foi considerado como ligante controle nas análises de *docking*, por ser a molécula responsável pela ativação da via Wnt (DOUBRAVSKA et al. 2011).

Os arquivos dos ligantes obtidos no banco de dados ZINC no formato mol2 foram transformados para o formato pdbqt, utilizado pelo programa AutoDockTools, com auxílio do programa OpenBabel (O'BOYLE et al. 2011). O programa X-Score (WANG; LAI; WANG, 2002) foi utilizado para validar os resultados da energia de ligação estimada pelo Vina. O

software PyMOL (DELANO, 2015) foi utilizado para adicionar átomos de hidrogênio aos ligantes, após o processo de *docking*, e para transformar os arquivos do formato pdb em mol2 para utilização dos mesmos no programa X-Score.

6 RESULTADOS

Nesta seção do trabalho são apresentados os resultados obtidos pela presente pesquisa.

6.1 Modelagem molecular e aferição de qualidade

Devido à falta de estrutura tridimensional das proteínas Wnt3 e Wnt3a disponível, fez-se por meio de BLAST (NCBI) contra o PDB uma pesquisa de proteínas similares, utilizando-se as sequências FASTA de Wnt3 e Wnt3a, excluindo-se as de alinhamentos redundantes (ALTSCHUL et al., 1990). A pesquisa inicial foi feita apenas para proteínas humanas, mas o resultado da mesma não foi satisfatório. A seguir, foram pesquisadas proteínas similares sem especificação do organismo alvo. Foi encontrado apenas um alinhamento significativo para as proteínas Wnt3 e Wnt3a, o pdb4f0a_B de resolução de 3,25 Å e identidade de sequência de 41% para ambas as sequências alvo (JANDA et al., 2012). A proteína de pdb4krr_A de resolução 2,1 Å (CHU et al., 2013) também foi resultado do PSI-BLAST, mas apresentando apenas 24% de identidade de sequência.

Para Wnt3, os resíduos de 1-59 e o resíduo 355 não tiveram equivalência com a estrutura da proteína alinhada, portanto não foram considerados na construção do modelo. Para Wnt3a, os resíduos de 1-56 e o resíduo 352 não tiveram equivalência com a estrutura da proteína alinhada, portanto não foram considerados na construção do modelo. Nenhum dos resíduos desconsiderados na modelagem participa ativamente da ligação dessas proteínas à FZD7. As sequências primárias das proteínas Wnt3 e Wnt3a foram utilizadas para o alinhamento com a sequência de seu *template* pdb4f0a_B (FIGURA 7) com auxílio da ferramenta online EMBOSS Needle (LI et al., 2015).

Nota: Os aminoácidos conservados entre as sequências estão marcados por barras (|). Os pontos (.) indicam conservação entre resíduos com fracas similaridades em suas propriedades. Os dois pontos (:) representam conservação entre aminoácidos com fortes similaridades em suas propriedades. Destacados pelos retângulos vermelhos estão os resíduos conservados de serina, onde se prediz haver uma modificação pós-traducional conservada, considerados cruciais na sinalização entre essas proteínas e seu receptor FZD7. Destacados pelos retângulos verdes estão os resíduos conservados responsáveis pelo contato direto das proteínas com o receptor FZD7.

Fonte: Do autor.

A proteína FZD7 possui 3 estruturas experimentalmente definidas de frações de sua estrutura: pdb4z33, de resolução de 2,45 Å, representando os resíduos de 569 a 574; pdb5t44, de resolução de 1,99 Å, representando os resíduos de 31 a 168; e pdb5urv, de resolução de 2,20 Å, representando os resíduos de 30 a 168. Foi feita pesquisa por meio de BLAST (NCBI) contra o PDB, utilizando-se a sequência primária da FZD7 em busca de proteínas de sequências similares, excluindo-se os alinhamentos redundantes. Foram encontrados 11 alinhamentos significativos (identidade de sequência > 30%). Devido à redundância de alinhamentos, apenas 2 foram escolhidos como *templates* para a modelagem (pdb4jky_A e 5t44_A), entretanto, os resíduos de 1-29, de 169-229 e de 566-574 não tiveram correspondência com as estruturas dos *templates*.

Dentre os 100 modelos construídos para a proteína Wnt3 com múltiplos *templates*, utilizando pdb4f0a_B e pdb4krr, o modelo de número 91 foi escolhido como melhor por apresentar o menor potencial estatístico DOPE, -26226,7. O modelo foi então submetido a análise de estereoquímica pelo PROCHECK através da ferramenta online PDBsum, que gerou seu gráfico de Ramachandran. Seu gráfico apresenta 90,3% dos resíduos na região mais favorável, 8,1% dos resíduos em regiões adicionalmente permitidas, 0% em regiões generosamente permitidas e 1,6% em regiões não permitidas (FIGURA 8). A análise de RMSD, realizada pelo programa PyMOL, entre o modelo da Wnt3 construído e seus *templates* resultou em 0,341 Å. O fator de qualidade geral gerado pela ferramenta ERRAT foi de 71,58%. A análise do perfil 3D do modelo, realizada pelo software VERIFY3D, resultou em 79,66% dos resíduos com um *score* médio 3D-1D $\geq 0,2$.

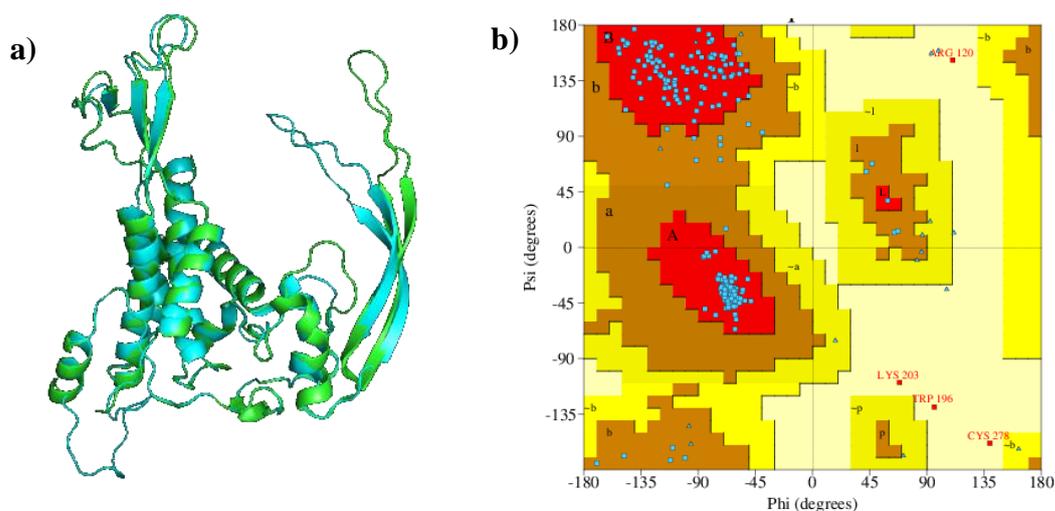


Figura 8 – Resultados da modelagem por homologia da proteína Wnt3.

a) Sobreposição do modelo 91 de Wnt3 (azul) e seu *template* pdb4f0a_B (verde) realizado através do programa PyMOL.

b) Gráfico de Ramachandran do modelo 91 de Wnt3.

Fonte: Do autor.

Com relação aos 100 modelos gerados a partir de múltiplos *templates*, pdf4f0a_B e pdb4krr, para a proteína Wnt3a, foi escolhido o modelo de número 1 por apresentar o menor valor DOPE, -25848,1. O gráfico de Ramachandran mostra 86,8% dos resíduos na região mais favorável, 8,9% em regiões adicionais permitidas, 2,7% em regiões generosamente permitidas e 1,6% em regiões não permitidas. Os resultados da análise de ERRAT apresentou uma qualidade geral de 66,19% para o modelo, enquanto a análise do seu perfil 3D pelo VERIFY3D apresentou 78,31% dos resíduos com um *score* médio 3D-1D $\geq 0,2$. A análise de RMSD, realizada pelo programa PyMOL, entre o modelo da Wnt3a e seus *templates* resultou em 0,430Å (FIGURA 9).

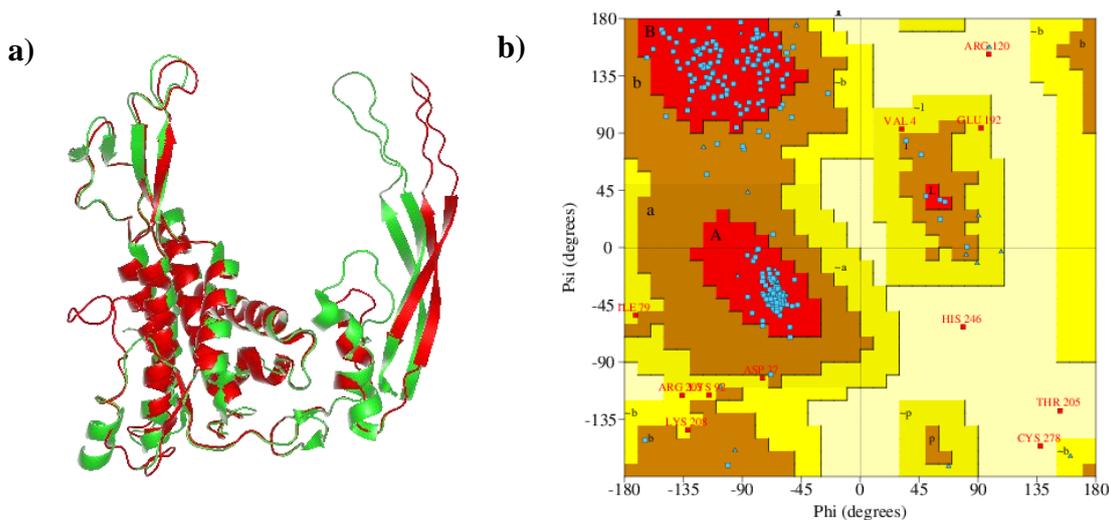


Figura 9 – Resultados da modelagem por homologia da proteína Wnt3a.

a) Sobreposição do modelo 1 de Wnt3a (vermelho) e seu *template* pdb4f0a_B (verde) realizado através do programa PyMOL.

b) Gráfico de Ramachandran do modelo 91 de Wnt3a.

Fonte: Do autor

Para a proteína FZD7, dos 100 modelos construídos a partir dos *templates* pdb4jvkv_a e pdb5t44_a, o modelo de número 49 foi escolhido com um valor de DOPE de -61017,5. Foi gerado um gráfico de Ramachandran através de análise pelo PROCHECK, que resultou em 88,1% dos resíduos na região mais favorável, 9,2% dos resíduos em regiões adicionais permitidas, 1,2% dos resíduos em regiões generosamente permitidas e 1,5% dos resíduos em regiões não permitidas. O modelo foi então submetido a análise de ERRAT que resultou em uma pontuação de 43,2% (FIGURA 10). Em seguida, o perfil 3D do modelo foi avaliado pelo VERIFY3D, que mostrou que 52,44% dos resíduos apresentam um *score* médio 3D-1D $\geq 0,2$. Os resultados dessas análises deixam evidente que o modelo gerado por homologia não apresenta uma boa qualidade. Decidiu-se então realizar uma modelagem *ab initio* para a proteína FZD7 pelo software I-TASSER.

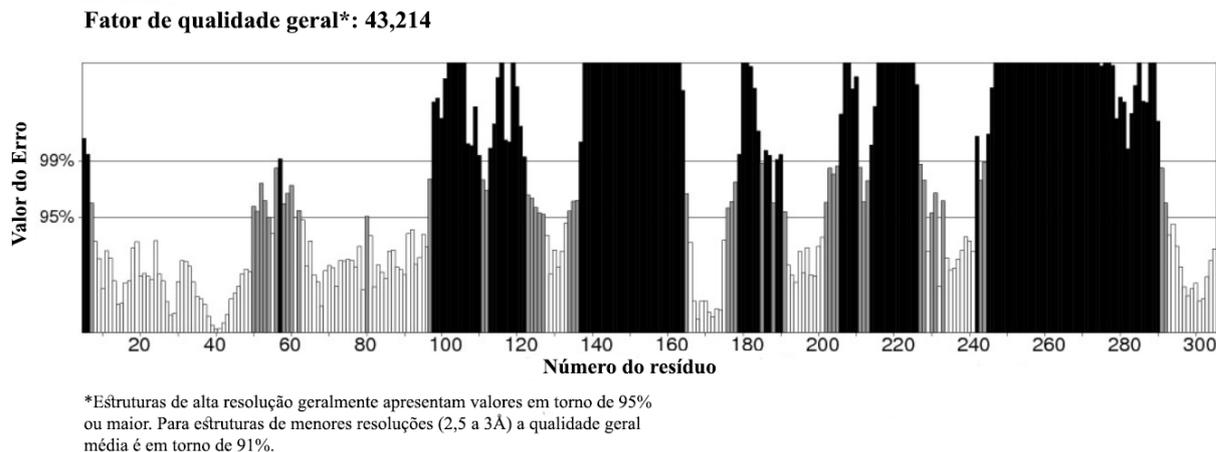


Figura 10 - Análise de ERRAT do modelo 49 de FZD7 gerado por homologia.
Fonte: Do autor.

Os melhores modelos construídos pela abordagem *ab initio* foram ranqueados pelo próprio software a partir de seus valores de C-score. Uma variação de 3 *templates* foi utilizada para a geração dos modelos (pdb517dA, pdb4jkvA, pdb4qinA). O melhor modelo gerado apresentou um valor de RMSD de $9,3 \pm 4,6$ Å e C-score de -0,74. A análise de ERRAT do modelo teve pontuação 83,6%, enquanto sua análise do VERIFY3D resultou em 64,29% dos resíduos apresentando um *score* médio 3D-1D $\geq 0,2$. O gráfico de Ramachandran gerado pelo PROCHECK para este modelo apresentou 75,4% dos resíduos na região mais favorável, 19,4% dos resíduos em regiões adicionalmente permitidas, 2,5% dos resíduos em regiões generosamente permitidas e 2,7% dos resíduos em regiões não permitidas (FIGURA 11).

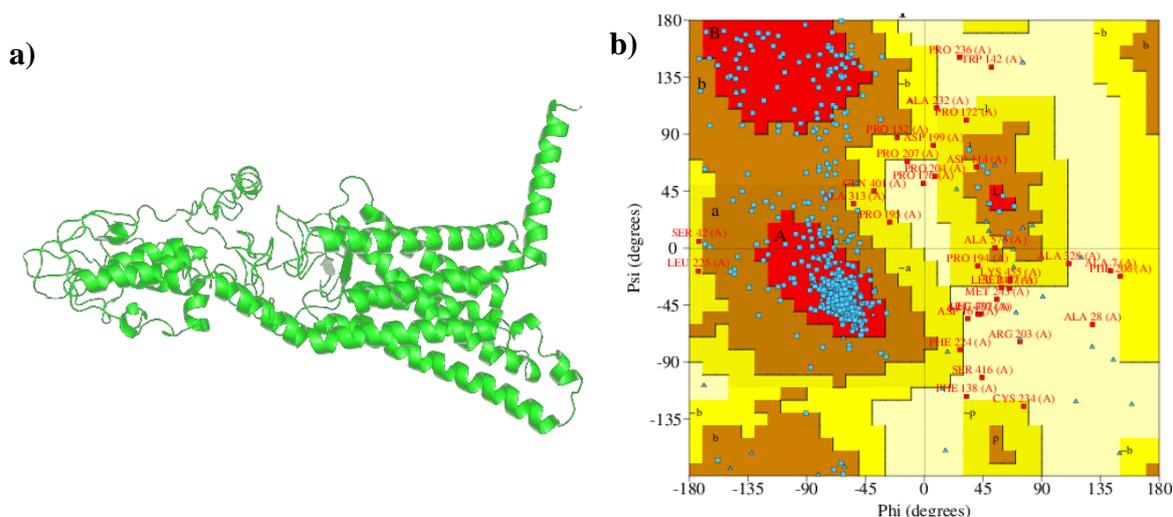


Figura 11 – Resultados da modelagem *ab initio* da proteína FZD7.

a) Modelo de FZD7 gerado pelo I-TASSER.

b) Gráfico de Ramachandran do modelo de FZD7 construído pelo I-TASSER.

Fonte: Do autor.

Para aumentar a confiabilidade das análises de bioinformática, optou-se por utilizar a estrutura tridimensional definida experimentalmente, de código pdb5t44, para as análises de *docking*. Essa estrutura apresenta resolução de 1,99 Å, corresponde ao domínio rico em cisteína (*cysteine rich domain*, CRD) da proteína FZD7 e abrange os resíduos de 31 a 168. O CRD da proteína FZD7 é o sítio onde ocorre a interação da mesma com as proteínas Wnt3 e Wnt3a (JANDA et al., 2012).

6.2 Pesquisa de ligantes e triagem virtual

A busca no banco de dados ZINC retornou 29 compostos, e os arquivos de suas estruturas, juntamente com o arquivo da estrutura do ácido palmitoleico, foram baixados nos formatos mol2 e sdf. Foi utilizado o servidor online FAF-Drugs4 (LAGORCE et al., 2015; ALLAND et al., 2005) para avaliar se os compostos estão em conformidade com a regra dos 5 de Lipinski (TABELA 2). A regra dos cinco de Lipinski é um conjunto de regras que descrevem propriedades moleculares para determinar se um composto químico possui características similares às de fármacos. Para uma molécula ser rejeitada como provável fármaco, ela deve violar duas ou mais das seguintes regras: peso molecular ≤ 500 g/mol; log calculado P (LogP) ≤ 5 ; número de doadores de ligação de hidrogénio (HBD) ≤ 5 ; número de aceptadores de ligação de hidrogénio (HBA) ≤ 10 (LIPINSKI et al., 2001). Nenhum dos ligantes avaliados violou mais de 1 das regras de Lipinski, o que indica que estes possuem propriedades moleculares similares a fármacos.

Tabela 2 - Valores de Lipinski dos 30 compostos retirados do banco de dados ZINC.

Moléculas	Fórmula Química	Peso Molecular (g mol ⁻¹)	logP	HBD	HBA	Violações de Lipinski
ZINC01531062	C ₁₄ H ₂₇ O ₂	226,36	5,50	1	2	1
ZINC02504617	C ₁₅ H ₂₉ O ₂	240,38	6,04	1	2	1
ZINC04529321	C ₁₄ H ₂₇ O ₂	226,36	5,50	1	2	1
ZINC05260607	C ₁₈ H ₃₅ O ₂	282,46	7,64	1	2	1
ZINC05260741	C ₁₈ H ₃₅ O ₂	282,46	7,64	1	2	1
ZINC05260769	C ₁₅ H ₂₉ O ₂	240,38	6,04	1	2	1

ZINC05972969	C ₂₀ H ₃₉ O ₂	308,50	7,79	1	2	1
ZINC06920375	C ₂₂ H ₄₃ O ₂	338,57	9,81	1	2	1
ZINC06920439	C ₂₀ H ₃₉ O ₂	310,51	8,72	1	2	1
ZINC06920440	C ₂₄ H ₄₇ O ₂	366,62	10,89	1	2	1
ZINC08220981	C ₂₂ H ₄₃ O ₂	338,57	9,81	1	2	1
ZINC08221009	C ₁₆ H ₃₁ O ₂	254,41	6,58	1	2	1
ZINC08221028	C ₁₈ H ₃₅ O ₂	282,46	7,64	1	2	1
ZINC08221053	C ₁₈ H ₃₅ O ₂	282,46	7,64	1	2	1
ZINC13507105	C ₂₀ H ₃₉ O ₂	308,50	7,79	1	2	1
ZINC13507107	C ₂₀ H ₃₉ O ₂	308,50	7,79	1	2	1
ZINC13507109	C ₂₀ H ₃₉ O ₂	308,50	7,79	1	2	1
ZINC13546050	C ₁₆ H ₃₁ O ₂	254,41	6,56	1	2	1
ZINC14616324	C ₁₈ H ₃₂ O ₄	312,44	5,90	2	4	1
ZINC14881154	C ₁₆ H ₃₁ O ₂	254,41	6,58	1	2	1
ZINC14881303	C ₁₈ H ₃₅ O ₂	282,46	7,64	1	2	1
ZINC15149683	C ₂₀ H ₃₉ O ₂	310,51	8,72	1	2	1
ZINC15149685	C ₂₀ H ₃₉ O ₂	310,51	8,72	1	2	1
ZINC32787113	C ₁₈ H ₃₅ O ₂	282,46	7,64	1	2	1
ZINC32787341	C ₁₆ H ₃₁ O ₂	254,41	6,56	1	2	1
ZINC32788451	C ₂₂ H ₄₃ O ₂	338,57	9,81	1	2	1
ZINC35465169	C ₁₈ H ₃₂ O ₄	312,44	5,90	2	4	1
ZINC83332374	C ₁₄ H ₂₇ O ₂	226,36	5,50	1	2	1
ZINC85590856	C ₁₉ H ₃₇ O ₂	296,49	8,18	1	2	1
ZINC85663738	C ₂₀ H ₃₉ O ₂	310,51	8,72	1	2	1

Fonte: Do autor.

A pesquisa de literatura por ligantes de ação similar à da proteína Wnt3a resultou em apenas um resultado, o anticorpo anti-CRD da proteína FZD7 vantiactumab (Anti-Fzd ou OMP-18R5). Entretanto, os pesquisadores responsáveis pelo trabalho não possuem a estrutura tridimensional do mesmo disponível (GURNEY et al., 2012). Não existem anticorpos disponíveis em bancos de dados públicos que se liguem ao CRD da proteína FZD7.

6.3 Docking Molecular

Os resultados do *docking* molecular entre os 30 ligantes e a proteína FZD7 se encontram na tabela 3, juntamente com seus valores de X-Score. O software OpenBabel (O'BOYLE et al. 2011), foi utilizado para converter os arquivos dos ligantes do formato mol2, obtidos no banco de dados ZINC, para pdbqt.

Tabela 3 - Energia de ligação entre os 30 compostos retirados do banco de dados ZINC e o CRD da proteína FZD7 obtidos pelo AutoDock Vina e X-Score.

Moléculas	Fórmula Química	Vina (kcal/mol)	X-Score (kcal/mol)	logP
ZINC05972969	C ₂₀ H ₃₉ O ₂	-8,1	-9,45	7,79
ZINC04529321	C ₁₄ H ₂₇ O ₂	-8,1	-8,65	5,5
ZINC08221009	C ₁₆ H ₃₁ O ₂	-8	-8,97	6,58
ZINC13546050	C ₁₆ H ₃₁ O ₂	-8	-8,92	6,56
ZINC05260769	C ₁₅ H ₂₉ O ₂	-8	-8,8	6,04
ZINC14881303	C ₁₈ H ₃₅ O ₂	-7,9	-9,18	7,64
ZINC14881154	C ₁₆ H ₃₁ O ₂	-7,9	-8,9	6,58
ZINC02504617	C ₁₅ H ₂₉ O ₂	-7,9	-8,78	6,04
ZINC13507109	C ₂₀ H ₃₉ O ₂	-7,9	-8,69	7,79
ZINC83332374	C ₁₄ H ₂₇ O ₂	-7,9	-8,41	5,5
ZINC15149685	C ₂₀ H ₃₉ O ₂	-7,8	-9,51	8,72
ZINC13507107	C ₂₀ H ₃₉ O ₂	-7,8	-9,33	7,79
ZINC85663738	C ₂₀ H ₃₉ O ₂	-7,8	-9,31	8,72
ZINC05260607	C ₁₈ H ₃₅ O ₂	-7,8	-9,29	7,64
ZINC85590856	C ₁₉ H ₃₇ O ₂	-7,8	-9,21	8,18
ZINC08221028	C ₁₈ H ₃₅ O ₂	-7,8	-9,13	7,64
ZINC01531062	C ₁₄ H ₂₇ O ₂	-7,8	-8,78	5,5
ZINC08221053	C ₁₈ H ₃₅ O ₂	-7,6	-9,05	7,64
ZINC32787113	C ₁₈ H ₃₅ O ₂	-7,6	-9,04	7,64
ZINC06920439	C ₂₀ H ₃₉ O ₂	-7,5	-9,21	8,72
ZINC32788451	C ₂₂ H ₄₃ O ₂	-7,5	-9,04	9,81
ZINC05260741	C ₁₈ H ₃₅ O ₂	-7,5	-8,97	7,64
ZINC08220981	C ₂₂ H ₄₃ O ₂	-7,4	-8,81	9,81
ZINC13507105	C ₂₀ H ₃₉ O ₂	-7,3	-9,63	7,79
ZINC35465169	C ₁₈ H ₃₂ O ₄	-7,3	-8,7	5,9
ZINC14616324	C ₁₈ H ₃₂ O ₄	-7,2	-8,65	5,9
ZINC15149683	C ₂₀ H ₃₉ O ₂	-7,1	-9,14	8,72

ZINC06920440	C ₂₄ H ₄₇ O ₂	-4,9	-6,99	10,89
ZINC06920375	C ₂₂ H ₄₃ O ₂	-4,7	-6,92	9,81
ZINC32787341	C ₁₆ H ₃₁ O ₂	-4,5	-6,85	6,56

Fonte: Do autor.

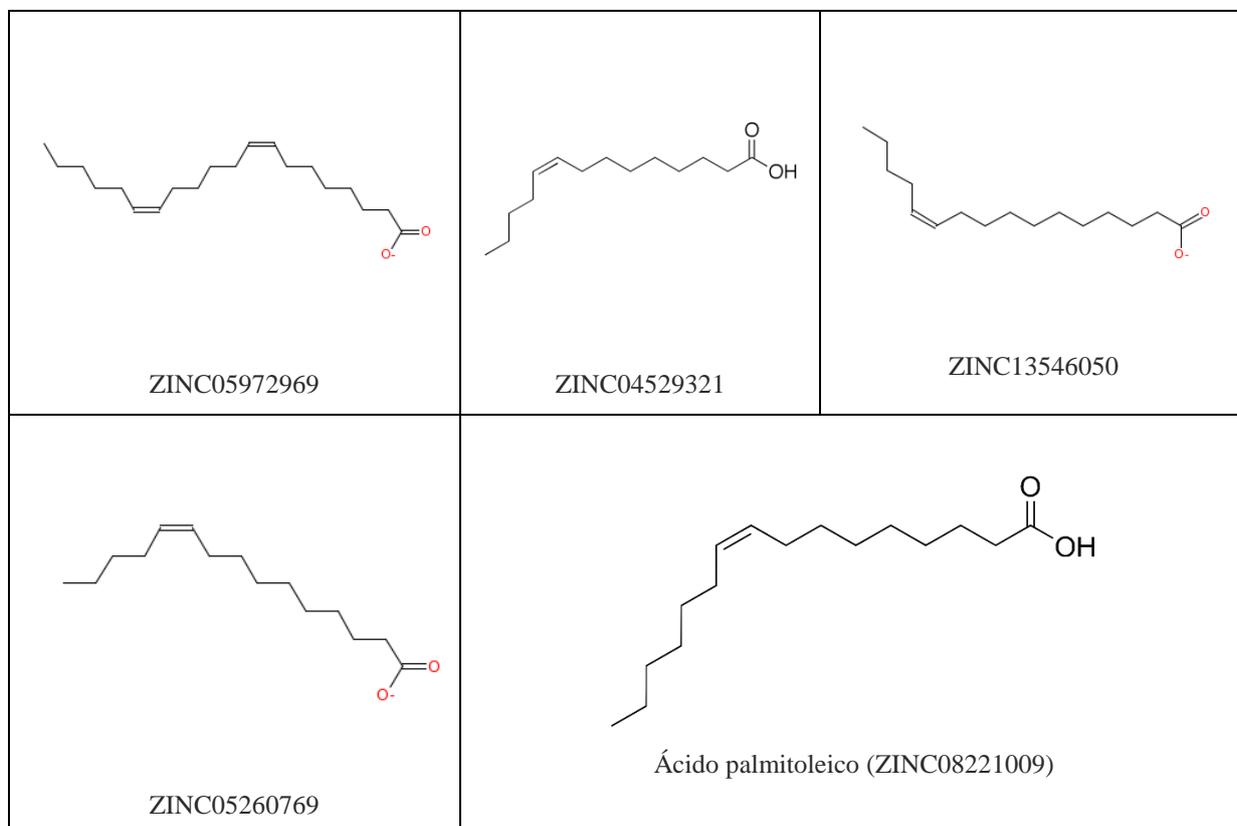


Figura 12 - Estrutura química dos 5 melhores ligantes apresentados na tabela 3.
Fonte: Do autor.

O ácido palmitoleico (zinc08221009), considerado como controle, obteve uma energia de ligação de -8,0 kcal/mol. Dois ligantes, zinc13546050 e zinc05260769, obtiveram a mesma energia de ligação que a molécula controle. Dois outros ligantes, zinc04529321 e zinc05972969, obtiveram energias de ligação de -8,1 kcal/mol, superiores à da molécula controle. O software LigPlot (WALLACE; LASKOWSKI; THORNTON, 1996) foi utilizado para demonstrar visualmente a interação entre os ligantes e a proteína FZD7. Os gráficos das melhores interações são mostrados na figura 13 juntamente com a visualização do docking através do software PyMOL.

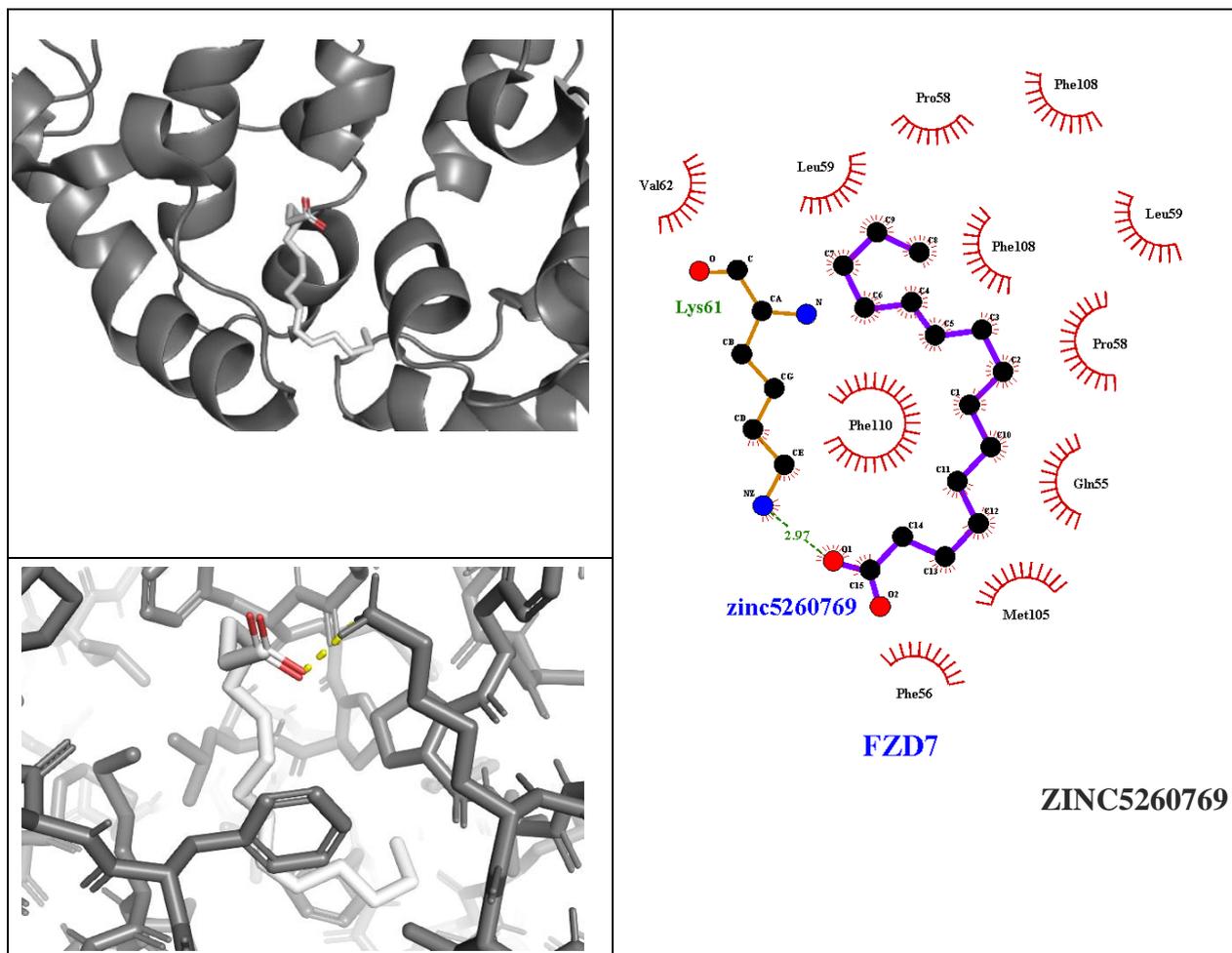


Figura 13 - Interação do CRD da proteína FZD7 (pdb5t44) com os melhores ligantes resultantes do *docking* molecular.

Nota: No gráfico gerado pelo software LigPlot, os semicírculos indicam os contatos de van der Waals entre resíduos da proteína e o ligante. Linhas tracejadas representam ligações de hidrogênio.

Fonte: Do autor.

Figura 13 - Interação do CRD da proteína FZD7 (pdb5t44) com os melhores ligantes resultantes do *docking* molecular. (Continua)

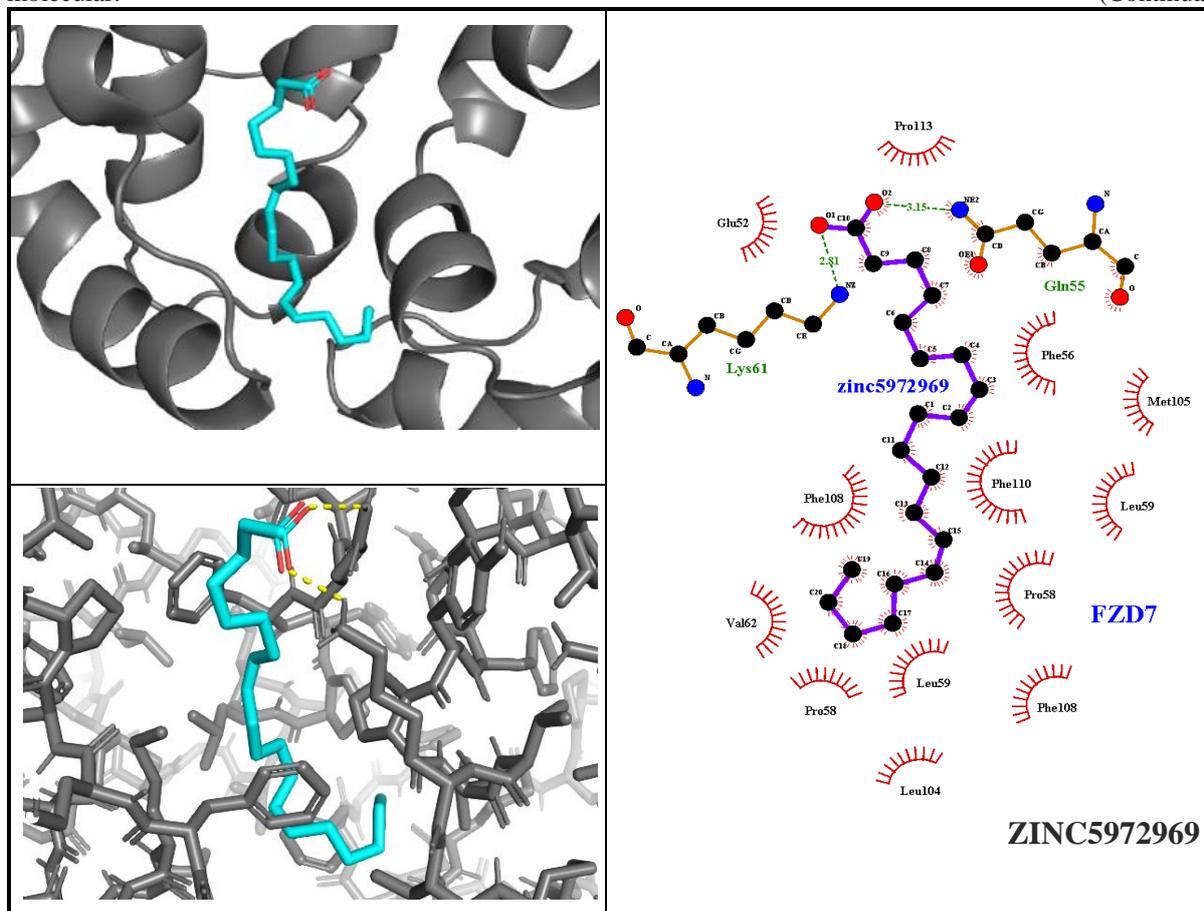


Figura 13 - Interação do CRD da proteína FZD7 (pdb5t44) com os melhores ligantes resultantes do *docking* molecular. (Continua)

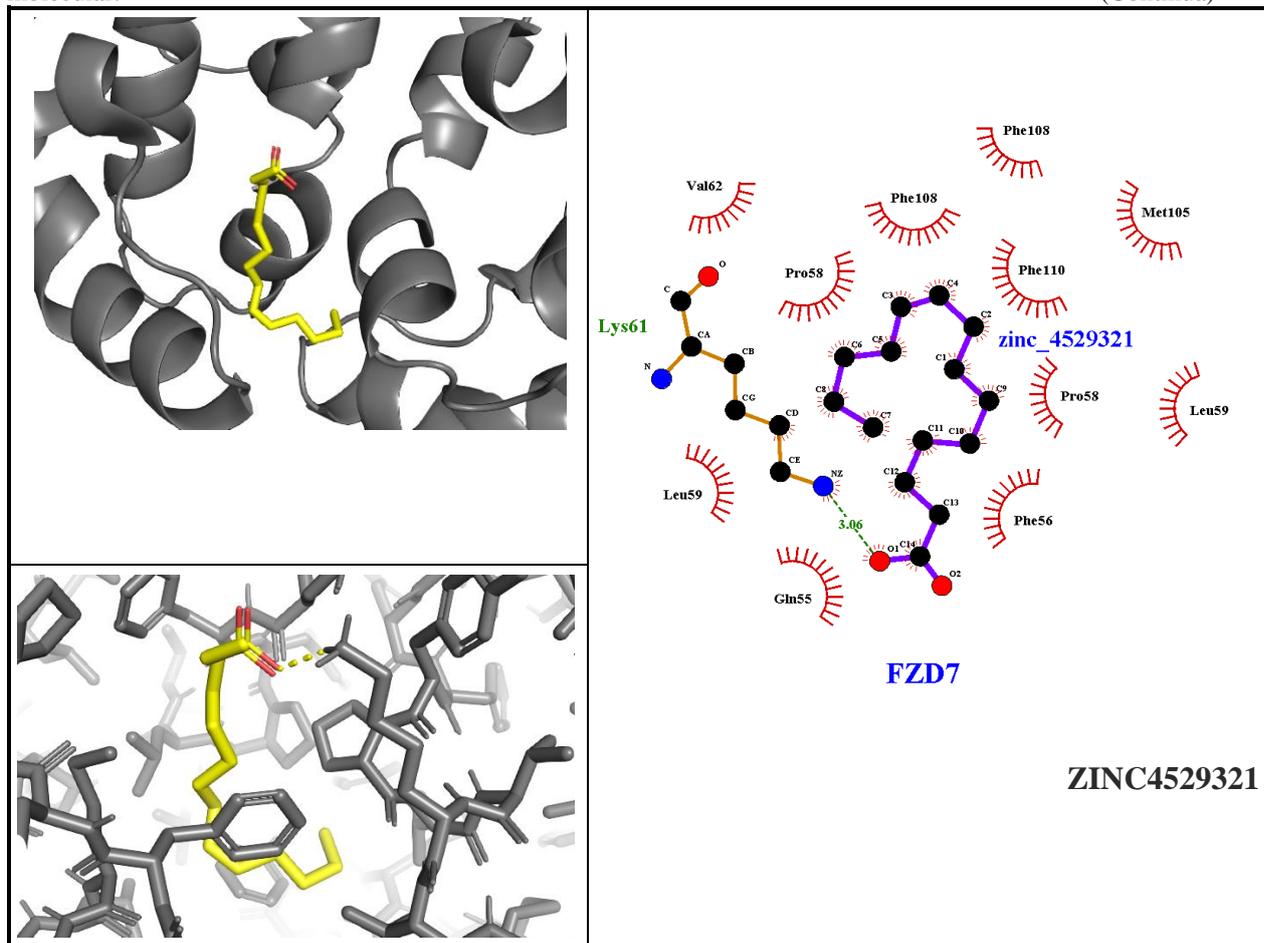


Figura 13 - Interação do CRD da proteína FZD7 (pdb5t44) com os melhores ligantes resultantes do *docking* molecular. (Continua)

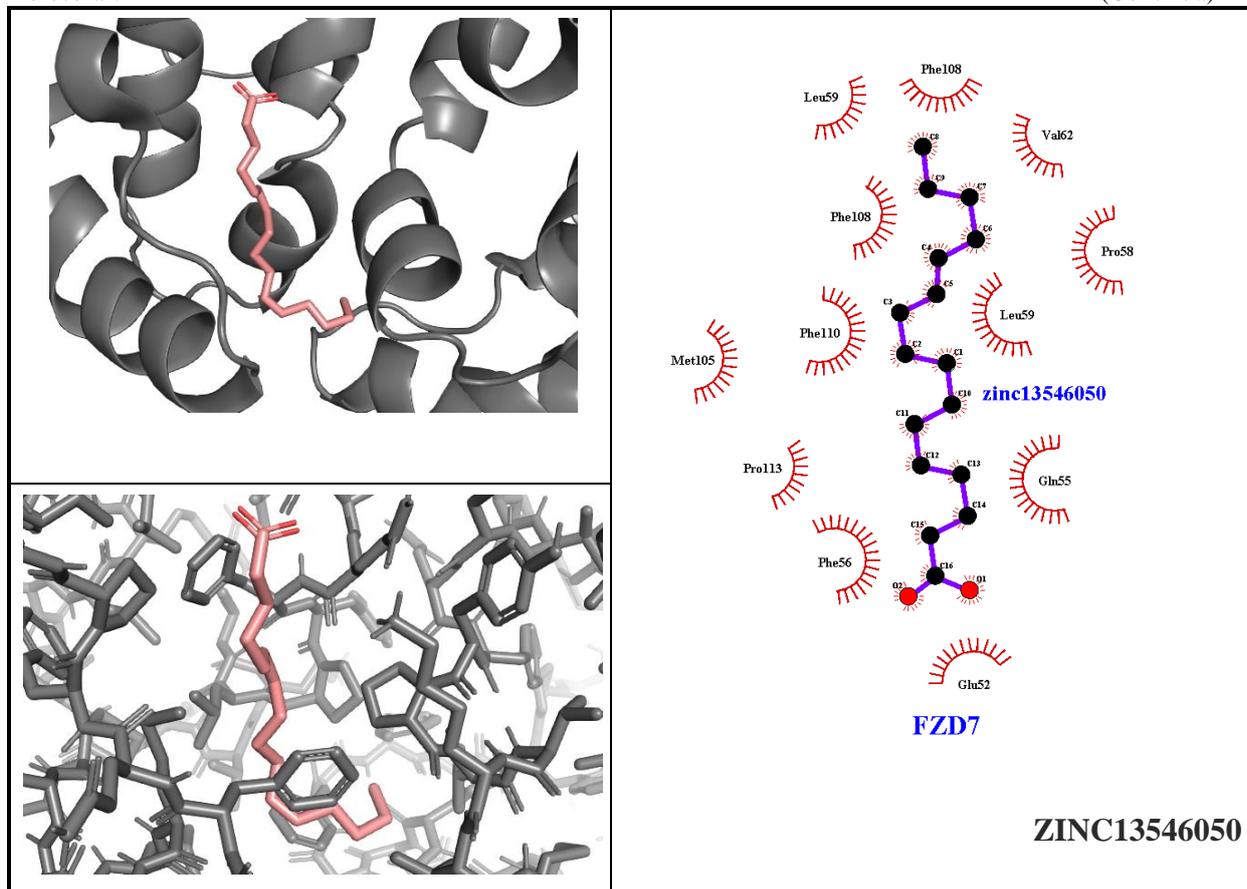
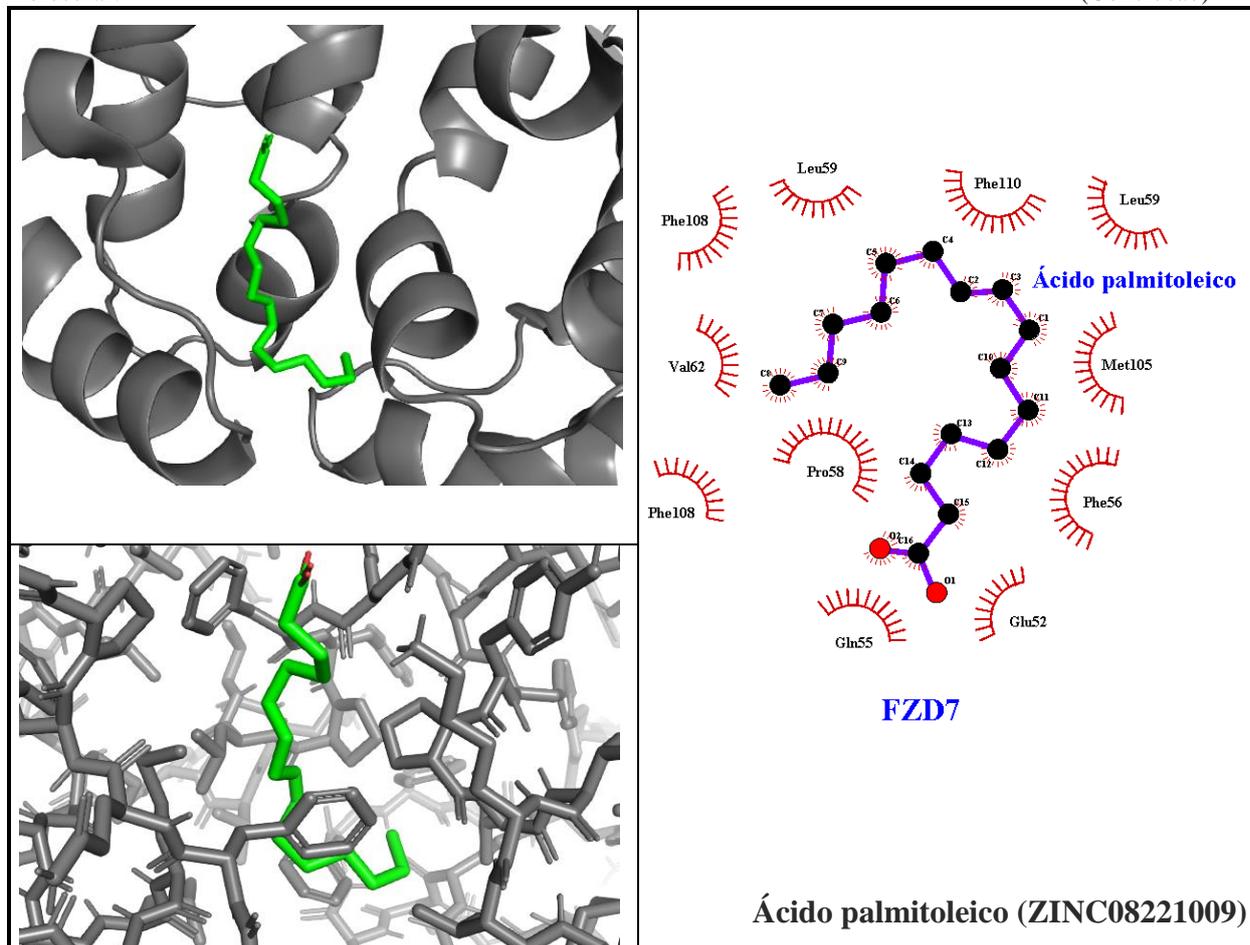


Figura 13 - Interação do CRD da proteína FZD7 (pdb5t44) com os melhores ligantes resultantes do *docking* molecular. (Conclusão)



As representações obtidas através do LigPlot evidenciam que a interação entre o ácido palmitoleico (zinc08221009), considerado como molécula controle no estudo, e o CRD da proteína transmembrana FZD7 é mediada por contatos hidrofóbicos. O mesmo acontece com o ligante zinc13546050. Os ligantes zinc05260769 e zinc04529321 apresentam, além dos contatos hidrofóbicos, ligação de hidrogênio com o resíduo Lys61. O ligante zinc05972969, além dos contatos hidrofóbicos, apresenta uma ligação de hidrogênio com o resíduo Lys61 e uma ligação de hidrogênio com o resíduo Gln55.

7 DISCUSSÃO

O câncer de mama triplo-negativo é o tipo de neoplasia mamária de pior prognóstico atualmente. Isso se dá pela falta de marcadores moleculares em células desse tipo e por ainda não existir um tratamento específico e eficiente para as mesmas. Pelo fato dessas células não apresentarem possíveis receptores-alvo para drogas, a quimioterapia ainda é o tratamento mais recomendado, apesar de seus efeitos adversos. A prospecção de novas drogas mais efetivas, específicas, e menos citotóxicas para o tratamento dessa doença é de extrema importância atualmente. O desenvolvimento de novos tratamentos, entretanto, é dificultado justamente pela falta de marcadores moleculares bem definidos nesse tipo de câncer. Um melhor entendimento dos mecanismos moleculares envolvidos no desenvolvimento e progressão dessa neoplasia é necessário para permitir a criação de novos e melhores fármacos e terapias. Estudos desse tipo podem também levar à descoberta de novos alvos terapêuticos.

A relação entre a via de sinalização Wnt canônica e o desenvolvimento e progressão de neoplasias vem sendo estudada há mais de 30 anos (NUSSE; VARMUS, 1962). A falta de entendimento detalhado sobre as especificidades receptor-ligante, e sobre as estruturas de vários dos componentes da via de sinalização Wnt, dificulta o desenvolvimento e descoberta de novos agentes terapêuticos. Ainda não há disponível um fármaco que inative essa via de sinalização de forma efetiva. Além disso, as diversas mudanças pós-traducionais, que ocorrem até a secreção das proteínas Wnt3 e Wnt3a da célula, dificultam a determinação de sua estrutura experimentalmente, o que aumenta o interesse em sua modelagem *in silico*.

A conhecimento do autor deste trabalho, não existem ainda estudos definindo potenciais moléculas líderes que tenham como alvo o CRD da proteína FZD7. Existem trabalhos que evidenciam moléculas que impedem a ligação da Wnt3a ao co-receptor LRP6 (ZHANG et al. 2013; LU et al. 2011; LU et al. 2012) interrompendo a via de sinalização Wnt. Entretanto, como evidenciado por Kuhnert et al. (2004), a interrupção da via bloqueando interações com o LRP6 implica na perda de homeostasia do trato gastrointestinal em ratos, tornando o CRD da proteína FZD7 um alvo mais atrativo para o desenvolvimento de tratamentos.

Os modelos das proteínas Wnt3, Wnt3a e FZD7 gerados por este trabalho, inéditos ainda a bancos de dados públicos, podem ser usados em análises de *docking* e dinâmica molecular em futuros estudos. Dessa forma, este estudo pode servir de base para a construção de novos modelos e estudos moleculares focando na via de sinalização Wnt como alvo para

novas terapias. Foram depositados no banco de dados de modelos PMDB (*Protein Model DataBase*) os melhores modelos de Wnt3 e Wnt3a construídos nesse trabalho sob os códigos de identificação PM0081550 e PM0081549, respectivamente. O banco de dados PMDB requer que os modelos não possuam uma distância máxima de 3,6 a 4,0 Å entre átomos de carbono alfa (CA) consecutivos e uma distância máxima de 3,6 Å entre átomos CA não consecutivos, e checa isso automaticamente durante a submissão dos modelos (CASTRIGNANÒ et al., 2006). Análises de dinâmica molecular podem ser utilizadas futuramente para otimizar a qualidade dos modelos aqui gerados.

Os resultados do presente trabalho revelam como potenciais candidatos a moléculas líderes os ligantes zinc13546050, zinc05972969, zinc05260769, zinc04529321 e o ácido palmitoleico, tomando como base seus valores de energia de ligação nas análises de *docking* molecular contra o CRD da proteína FZD7. A energia de ligação dos ligantes zinc05972969 e zinc04529321 foi superior à energia de ligação da molécula considerada como controle, o ácido palmitoleico, evidenciando seu potencial como moléculas líderes. Os 4 ligantes e o ácido palmitoleico possuem características similares às de fármacos, como evidenciado pela aplicação da regra dos 5 de Lipinski nos mesmos. Futuros estudos *in vitro* e *in vivo* são necessários para avaliar os efeitos e a efetividade dessas moléculas em células TMTN. Os resultados aqui obtidos devem servir como base e ponto de partida para futuros estudos envolvendo a via de sinalização Wnt, a interação molecular das proteínas Wnt3, Wnt3a e FZD7, e o desenvolvimento de novos fármacos que tenham como alvo tal via.

REFERÊNCIAS

AHMED, R.; ALAWIN, O.; SYLVESTER, P. γ -Tocotrienol reversal of epithelial-to-mesenchymal transition in human breast cancer cells is associated with inhibition of canonical Wnt signaling. **Cell proliferation**, v. 49, n. 4, p. 460-470, 2016.

ALLAND, C. et al. RPBS: a web resource for structural bioinformatics. **Nucleic Acids Research**, v. 33, p. W44-W49, 2005.

ALTSCHUL, S.F.; MADDEN, T.L.; SCHAFFER, A.A. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v. 25, p. 3389-3402, 1997.

ALTSCHUL, S.F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, 215:403-410, 1990.

AMIN, M. B. et al. **AJCC cancer staging manual**, 7 ed. New York: Springer, 2010.

ANDRADE, F.P. et al. Insights from studies with oral cleft genes suggest associations between WNT-pathway genes and risk of oral cancer. **Journal of Dental Research**, v. 90, p. 740-746, 2011.

ARNOLD, K. et al. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. **Bioinformatics**, v. 22, p. 195-201, 2006.

BARRETT, T. et al. NCBI GEO: archive for functional genomics data sets--update. **Nucleic Acids Research**, (Database issue):D991-5, 2013,

BASAK, S.C. Chemobioinformatics: the advancing frontier of computer-aided drug design in the post-genomic era. **Current Computer-Aided Drug Design**, v. 8, p. 1-2, 2012

BENHAJ, K.; AKCALI, K.C.; OZTURK, M. Redundant expression of canonical Wnt ligands in human breast cancer cell lines. **Oncology Reports**, v. 15, p. 701-707, 2006.

BERMAN, H.M. et al. The Protein Data Bank. **Nucleic Acids Research**, v. 28, p. 235-242, 2000.

BERTUCCI, F.; FINETTI, P.; BIRNBAUM, D. Basal breast cancer: a complex and deadly molecular subtype. **Current Molecular Medicine**, v. 12, p. 96-110, 2012.

BRASIL. Ministério da Saúde, Instituto Nacional do Câncer. **Tipos de câncer: Mama**.

Disponível em:

<<http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama>>. Acesso em: 9 de nov. 2017.

BOUTET, E. et al. UniProtKB/Swiss-Prot, the Manually Annotated Section of the UniProt KnowledgeBase: How to Use the Entry View. **Methods in Molecular Biology**, v. 1374, pp. 23-54, 2016.

CAMPBELL, J. **The hero's journey: Joseph Campbell on his life and work**. São Francisco: Harper & Row, 1990.

CARVALHO, L. V. et al. Molecular characterization of breast cancer in young Brazilian women. **Revista da Associação Médica Brasileira**, v. 56, p. 278-287, 2010.

CASTRIGNANÒ, T. et al. The PMDB Protein Model Database. **Nucleic Acids Research**, v. 34, p. D306-D309, 2006.

CHONG, C. R.; XU, J.; LU, J.; BHAT, S.; SULLIVAN, D.J.Jr.; LIU, J.O. Inhibition of angiogenesis by the antifungal drug itraconazole. **ACS Chemical Biology**, v. 2, p. 263-270, 2007.

CHOTHIA, C.; LESK, A.M. The relation between the divergence of sequence and structure in proteins. **EMBO Journal**, v. 5, p. 823-826, 1986.

CHU, M. L. et al. Structural studies of Wnts and identification of an LRP6 binding site. **Structure**, v. 21, n. 7, p. 1235-1242, 2013.

CINTRA, J. R. D. et al. Perfil imuno-histoquímico e variáveis clinicopatológicas no câncer de mama. **Revista da Associação Médica Brasileira**, v. 58, p. 178-187, 2012.

DE BEER, T. A. P. Et al. PDBsum additions. **Nucleic Acids Research**, v. 42, p. D292-D296, 2014.

DE BROT, M. et al. Carcinomas mamários de tipo basal: perfil clínico-patológico e evolutivo **Revista da Associação Médica Brasileira**, v. 55, p. 529-534, 2009.

DELANO, W. L. The PyMOL Molecular Graphics System, Version 1.8. In., 1.8 edn: **Schrodinger**, LLC; 2015.

DEY, N. et al. Differential Activation of Wnt- β -Catenin Pathway in Triple Negative Breast Cancer Increases MMP7 in a PTEN Dependent Manner. **PLoS ONE**, v. 8, n. 10, p. 1–17, 2013.

DIAS, R.; DE AZEVEDO, W. F. Molecular Docking Algorithms. **Current Drug Targets**, v. 9, p. 1040-1047, 2008.

DOUBRAVSKA, L. et al. Fatty acid modification of Wnt1 and Wnt3a at serine is prerequisite for lipidation at cysteine and is essential for Wnt signalling. **Cellular Signalling**, v. 23, p. 837-848, 2011.

DUTRA, M.C. et al. Imunofenótipo e evolução de câncer de mama: comparação entre mulheres muito jovens e mulheres na pós-menopausa. **Revista Brasileira de Ginecologia e Obstetrícia**, v. 31, p. 54-60, 2009.

EISENBERG, D.; LÜTHY, R.; BOWIE, J. U. VERIFY3D: assessment of protein models with three-dimensional profiles. **Methods in Enzymology**, v. 277, p. 396-404, 1997.

FISER, A.; SALI, A. Modeller: generation and refinement of homology-based protein structure models. **Methods in Enzymology**, v. 374, p. 461–91, 2003.

FLOUDAS, C. A. et al. Advances in protein structure prediction and de novo protein design: A review. **Chemical Engineering Science**, v. 61, p. 966-988, 2006.

GOLDBERG, D. E. **Genetic Algorithms in Search, Optimization and Machine Learning**. Massachusetts: Addison-Wesley Longman, 1989.

GUEDES, I. A.; MAGALHÃES, C. S.; DARDENNE, L. E. Receptor–ligand molecular docking. **Biophysical Reviews**, v. 6, p. 75-87, 2014.

GUJRAL, T. S.; MACBEATH, G. A system-wide investigation of the dynamics of Wnt signaling reveals novel phases of transcriptional regulation. **PLoS One**, v. 5, 2010.

GURNEY, A. et al. Wnt pathway inhibition via the targeting of Frizzled receptors results in decreased growth and tumorigenicity of human tumors. **OncoMed Pharmaceuticals**, v. 109, n. 29, p. 11717-11722, 2012.

HE, S. et al. Wnt3a: functions and implications in cancer. **Chinese Journal of Cancer**, v. 34, 2015.

INSTITUTO Nacional de Câncer José Alencar Gomes da Silva. Coordenação de Prevenção e Vigilância. **Estimativa 2016: incidência de câncer no Brasil**. Rio de Janeiro: INCA, 2015. Disponível em: <<http://www.inca.gov.br/estimativa/2016/estimativa-2016-v11.pdf>>. Acesso em: 24 out. 2017.

IRWIN, J. J. et al. ZINC: A Free Tool to Discover Chemistry for Biology. **Journal of Chemical Information and Modeling**, v. 52, n. 7, p. 1757–1768, 2012.

JAIN, A. N.; NICHOLLS, A. Recommendations for evaluation of computational methods. **Journal of Computer-Aided Molecular Design**, v. 22, p. 133–139, 2008.

JANDA, C.Y. et al. Structural basis of Wnt recognition by Frizzled. **Science**, v. 337, p. 59-64, 2012.

KIM, M. et al. Functional interaction between Wnt3 and Frizzled-7 leads to activation of the Wnt/ β -catenin signaling pathway in hepatocellular carcinoma cells. **Journal of Hepatology**, v. 48(5), p. 780–791, 2008.

KIRIKOSHI, H.; SEKIHARA, H.; KATOH, M. Expression of WNT14 and WNT14B mRNAs in human cancer, up-regulation of WNT14 by IFN γ and up-regulation of WNT14B by beta-estradiol. **International Journal of Oncology**, v. 19, p. 1221-1225, 2001.

KOMETIANI, P.; LIU, L.; ASKARI, A. Digitalis-induced signaling by Na⁺/K⁺-ATPase in human breast cancer cells. **Molecular Pharmacology**, v.67, p. 929-36, 2005.

KOZMA, D.; SIMON, I.; TUSNÁDY, G. E. PDBTM: Protein Data Bank of transmembrane proteins after 8 years. **Nucleic Acids Research**, v. 41, p. D524-D529, 2013.

KUHNERT, F. et al. Essential requirement for Wnt signaling in proliferation of adult small intestine and colon revealed by adenoviral expression of Dickkopf-1. **PNAS**, v. 101, n. 1, p. 266-271, 2004.

KUMAR, P.; AGGARWAL, R. An overview of triple-negative breast cancer. **Archives of Gynecology and Obstetrics**, v. 293, p. 247-269, 2016.

KUMARI, R.; KUMAR, R.; LYNN, A. g_mmpbsa - A GROMACS Tool for High-Throughput MM-PBSA Calculations. **Journal of Chemical Information and Modeling**, v. 54, n. 7, p. 1951-1962, 2014.

LAGORCE, D. et al. FAF-Drugs3: a web server for compound property calculation and chemical library design. **Nucleic Acids Research**, v. 43, p. W200-W207, 2015.

LASKOWSKI, R. A. et al. PROCHECK - a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, v. 26, p. 283-291, 1993.

LEHMANN, B. D. et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. **Journal of Clinical Investigation**, v. 121, p. 2750-2767, 2011.

LI, W. et al. The EMBL-EBI bioinformatics web and programmatic tools framework. **Nucleic Acids Research**, v. 43, p. W580-W584, 2015.

LIPINSKI, C. A. et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings¹. **Advanced Drug Delivery Reviews**, v. 46, p. 1-26, 2001.

LOPES, C. M. et al. Clinical, histomorphological, and therapeutic prognostic factors in patients with triple-negative invasive breast cancer. **Jornal Brasileiro de Patologia e Medicina Laboratorial**, v. 51, n. 6, p. 397-406, 2015.

LU, W. et al. Niclosamide suppresses cancer cell growth by inducing Wnt co-receptor LRP6 degradation and inhibiting the Wnt/ β -catenin pathway. **PLoS One**, v. 6, 2011.

LU, W. et al. Silibinin inhibits Wnt/ β -catenin signaling by suppressing Wnt co-receptor LRP6 expression in human prostate and breast cancer cells. **Cell Signal**, v. 24, p. 2291-2296, 2012.

MACARTHUR, M. W.; LASKOWSKI, R. A.; THORNTON, J. M. Knowledge-based validation of protein structure coordinates derived by X-ray crystallography and NMR spectroscopy. **Current Opinion in Structural Biology**, v. 4, p. 731-737, 1994.

MACDONALD, B.T.; TAMAI, K.; HE, X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. **Developmental Cell**, v. 17, p. 9, 2009.

MASILI-OKU, S. M. et al. The apocrine profile of triple-negative breast carcinomas in patients aged 45 years or younger: favorable but rare features. **Revista Brasileira de Ginecologia e Obstetrícia**, v. 38, n. 10, p. 512-517, 2016.

MELO, F.; SALI, A. Fold assessment for comparative protein structure modeling. **Protein Science**, v. 16, p. 2412-2426, 2007.

MINAS GERAIS. Secretaria de Estado da Saúde. Programa de Avaliação e Vigilância do Câncer e seus Fatores de Risco – PAV/MG. **Situação do câncer em Minas Gerais e suas macrorregiões de saúde**: estimativas de incidência e mortalidade para o ano 2013, válidas para 2014: perfil da mortalidade: perfil da assistência na alta. Belo Horizonte, v. 1, 2013.

MORRIS, G. M. et al. Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. **Journal of Computational Chemistry**, v. 16, p. 2785-2791, 2009.

MULLARD, A. 2015 FDA drug approvals. **Nature Reviews Drug Discovery**, v. 15, p. 73-76, 2016.

NAGAOKA, T. et al. Cripto-1 enhances the canonical Wnt/beta-catenin signaling pathway by binding to LRP5 and LRP6 co-receptors. **Cellular Signaling**, v. 25, p. 178–189, 2013.

NAMBOTIN, S. B. et al. Pharmacological inhibition of Frizzled-7 displays anti-tumor properties in hepatocellular carcinoma. **Journal of Hepatology**, v. 54, p. 288-299, 2011.

NILE, A. H. et al. Unsaturated fatty acyl recognition by Frizzled receptors mediates dimerization upon Wnt ligand binding. **Proceedings of the National Academy of Sciences**, v. 114, p. 4147-4152, 2017.

NUSSE, R.; VARMUS, H. E. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. **Cell**, v.31, p.99–109, 1982.

O'BOYLE, N. M. et al. Open Babel: An open chemical toolbox. **Journal of Cheminformatics**, v. 3, p. 33, 2011.

ORENGO, C. A.; THORNTON, J. M. Protein families and their evolution – a structural perspective. **Annual Review of Biochemistry**, v.74, p. 867-900, 2005.

PAYNE, S. J. L. et al. Predictive markers in breast cancer - the present. **Histopathology**, v. 52, p. 82-90, 2008.

POLLHEIMER J., et al. Activation of the canonical wntless/T-cell factor signaling pathway promotes invasive differentiation of human trophoblast. **The American Journal of Pathology**, v. 168, p. 1134–1147, 2006.

RAMACHANDRAN, G. N.; RAMAKRISHNAN, C.; SASISEKHARAN, V. Stereochemistry of polypeptide chain configurations. **Journal of Molecular Biology**, v. 7, p. 95–99, 1963.

REGO, N.; KOES, D. 3Dmol.js: molecular visualization with WebGL. **Bioinformatics**, v. 31, n. 8, p. 1322-1324, 2015.

SALI, A.; BLUNDELL, T. L. Comparative protein modelling by satisfaction of spatial restraints. **Journal of Molecular Biology**, v. 234, p. 779-815, 1993.

SANTOYO, H. A., et al. Protein-Protein and Protein-Ligand Docking. In: OGAWA, T. **Protein Engineering**, InTech, p. 63-81, 2013.

SAVAGNER, P. The epithelial–mesenchymal transition (EMT) phenomenon. **Annals of Oncology**, v. 21, p. 89-92, 2010.

SHEN, M.; SALI, A. Statistical potential for assessment and prediction of protein structures. **Protein Science**, v. 15, p. 2507-2524, 2006.

SHIM, J. S. et al. Effect of nitroxoline on angiogenesis and growth of human bladder cancer. **Journal of the National Cancer Institute**, v. 102, p. 1855-73, 2010.

SHIM, J. S.; LIU, J.O. Recent Advances in Drug Repositioning for the Discovery of New Anticancer Drugs. **International Journal of Biological Sciences**, v. 10, p. 654-663, 2014.

SIEVERS, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**, v. 7, 519. 2011.

SRIRANGAM, A. et al. Effects of HIV Protease inhibitor ritonavir on Akt-regulated cell proliferation in breast cancer. **Clinical Cancer Research**, v. 12, p. 1883-96, 2006.

UNIPROT, C. The Universal Protein Resource (UniProt). **Nucleic Acids Research**, v. 35(Database issue), p. D193-D197, 2007.

TAKADA, R. et al. Monounsaturated fatty acid modification of Wnt protein: its role in Wnt secretion. **Developmental Cell**, v. 11, p.791, 2006.

TANAKA, K.; KITAGAWA, Y; KADOWAKI, T. Drosophila segment polarity gene product porcupine stimulates the posttranslational N-glycosylation of wingless in the endoplasmic reticulum. **The Journal of Biological Chemistry**, v. 277, p. 12816–12823, 2002.

TORRE, L. A. et al. Global cancer incidence and mortality rates and trends – an update. **Cancer Epidemiology, Biomarkers and Prevention**, v. 25, n. 1, p. 16-27, 2016.

TROTT, O.; OLSON, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal of Computational Chemistry**, v. 31, p. 455-461, 2010.

WANG, R.; LAI, L.; WANG, S. Further development and validation of empirical scoring functions for structure-based binding affinity prediction. **Journal of Computer-Aided Molecular Design**, v. 16, p. 11-26, 2002.

WALLACE, A. C.; LASKOWSKI, R. A.; THORNTON, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. **Protein Engineering**, v. 8, p. 127-134, 1996.

WILLERT, K. et al. Wnt proteins are lipid-modified and can act as stem cell growth factors. **Nature**, v. 423, p. 448-452, 2003.

WILLETT, P. Similarity-based virtual screening using 2D fingerprints. **Drug Discovery Today**, v. 11, p. 1046–1053, 2006.

WORLD HEALTH ORGANIZATION. **International agency for research on cancer**. Globocan 2012. Disponível em: <http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx>. Acesso em 10 out 2017.

WU, S.; SKOLNICK, J.; ZHANG, Y. Ab initio modeling of small proteins by iterative TASSER simulations. **BMC Biology**, v. 5:17, 2007.

YANG, L. et al. FZD7 has a critical role in cell proliferation in triple negative breast cancer. **Oncogene**, v.30, p. 4437-46, 2011.

YANG, J. et al. The I-TASSER Suite: Protein structure and function prediction. **Nature Methods**, v. 12, p. 7-8, 2015.

ZHANG, J. et al. Kallistatin antagonizes Wnt/ β -catenin signaling and cancer cell motility via binding to low-density lipoprotein receptor-related protein 6. **Molecular and Cellular Biochemistry**, v. 379, p. 295-301, 2013.

ZHANG, Y.; SKOLNICK, J. Tertiary structure predictions on a comprehensive benchmark of medium to large size proteins. **Biophysical journal**, v. 87, p. 2647-2655, 2004a.

ZHANG, Y.; SKOLNICK, J. Automated structure prediction of weakly homologous proteins on a genomic scale. **Proceedings of the National Academy of Sciences**, v. 101, p. 7594-7599, 2004b.

ZHANG, Y.; SKOLNICK, J. The protein structure prediction problem could be solved using the current PDB library. **Proceedings of the National Academy of Sciences**, v. 102, p. 1029-1034, 2005.

APÊNDICE A - ARTIGO PUBLICADO COMO COAUTOR

Send Orders for Print-Reprints and e-prints to reprints@benthamscience.ae

Letters in Drug Design & Discovery, 2017, 14, 1-12

1

RESEARCH ARTICLE

Computational Approach to the Discovery of Phytochemical Molecules with Therapeutic Potential Targets to the PKCZ protein

Poliany G. Freitas^{1*}, Thiago C. Elias¹, Icaro A. Pinto¹, Luciano T. Costa^{1,2}, Paulo V.S.D. de Carvalho³, Daniel de Q. Omote⁴, Ihosvany Camps¹, Tati Ishikawa⁵, Helen A. Arcuri⁶, Susana Vinga⁷, Arlindo L. Oliveira⁸, Walter F.A. Junior⁹ and Nelson J.F. da Silveira^{1,8}

¹Laboratory of Molecular Modeling and Computer Simulations-MolMod-CS, Institute of Exact Sciences, Federal University of Alfenas, Alfenas, Brazil; ²Laboratory of Molecular Modeling and Computer Simulation-MolMod-CS, Institute of Chemistry, Federal Fluminense University, Niterói, Rio de Janeiro, Brazil; ³Laboratory of Cellular and Molecular Genetics, Federal University of Minas Gerais Brazil and Department of Mathematics and Computer Science, University Southern of Denmark, Odense, Denmark; ⁴Sleep Laboratory, Heart Institute (InCor), Faculty of Medicine, University of São Paulo, São Paulo, Brazil; ⁵Department of Medicine and Food, Faculty of Pharmaceutical Sciences, Federal University of Alfenas, Alfenas, Brazil; ⁶Center of the Study of Social Insects, Department of Biology, Institute of Biosciences of Rio Claro, São Paulo State University, Rio Claro, SP, São Paulo, Brazil; ⁷IDMEC, Higher Technical Institute, University of Lisboa, Lisboa, Portugal; ⁸INESC-ID / Higher Technical Institute, Lisboa, Portugal; ⁹Laboratory of Computational Systems Biology, Faculty of Biosciences, Pontifical Catholic University of Rio Grande do Sul (PUCRS), Porto Alegre, Brazil

Abstract: Background: Head and neck squamous cell carcinoma (HNSCC) is one of the most common malignancies in humans and the average 5-year survival rate is one of the lowest among aggressive cancers. Protein kinase C zeta (PKCZ) is highly expressed in head and neck tumors, and the inhibition of PKCZ reduces MAPK activation in five of seven head and neck tumors cell lines. Considering the world-wide HNSCC problems, there is an urgent need to develop new drugs to treat this disease, that present low toxicity, effective results and that are relatively inexpensive.

Method: A unified approach involving homology modeling, docking and molecular dynamics simulations studies on PKCZ are presented. The *in silico* study on this enzyme was undertaken using 10 compounds from latex of *Euphorbia tirucalli* L. (aveloz).

Results: The binding free energies highlight that the main contribution in energetic terms for the compounds-PKCZ interactions is based on van der Waals. The per-residue decomposition free energy from the PKCZ revealed that the compounds binding were favorably stabilized by residues Glu300, Ileu383 and Asp394. Based on the docking, Xscore and molecular dynamics results, euphol, β -sitosterol and taraxasterol were confirmed as the promising lead compounds.

Conclusion: The present study should therefore play a guiding role in the experimental design and development of euphol, β -sitosterol and taraxasterol as anticancer agents in head and neck tumors. They are potential lead compounds, better than other ligands based on the best values of docking and MM-PBSA energy.

ARTICLE HISTORY

Received: June 13, 2017
Revised: July 31, 2017
Accepted: August 01, 2017

DOI:
10.2174/1570180814666170810115134

Keywords: HNSCC, PKCZ, molecular marker, euphorbia tirucalli, homology modeling, molecular docking, molecular dynamics.

1. INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) develops in the oral cavity and is the sixth leading cancer by

incidence worldwide. This year, globally, 600,000 new cases of head and neck cancer are likely to occur and of these, 40-50% with HNSCC will only survive for 5 years [1]. According to the World Cancer Report of 2014, adding up pharyngeal, oral and laryngeal cancers in both sexes combined, 686,000 new cases were estimated for the year of 2012 with 375,000 estimated deaths for the same year. A decline in incidence can be seen since the 1990s for some countries such as India, China, USA, and Australia. However, conversely, increasing trends in incidence are being shown in

*Address correspondence to this author at the Laboratory of Molecular Modeling and Computer Simulation-MolMod-CS, Institute of Exact Sciences, Federal University of Alfenas – UNIFAL-MG, Gabriel Monteiro da Silva 700, Centro, 37130-001, Alfenas, Brazil; Tel/Fax: +55 35 3701-9602; E-mail: poliany.santos@gmail.com

APÊNDICE B - MODELOS DEPOSITADOS NO PMDB

PDB: [MELHORmulti_Wnt3.B99990091.pdb](#)
AUTHOR: [Icaro Alves Pinto](#)
SUBMITTER: [Alves Pinto Icaro](#)

It has been assigned
the following **PMDB id** : **PM0081550**

PDB: [MELHOR-Wnt3a_multi-1-56.pdb](#)
AUTHOR: [Icaro Alves Pinto](#)
SUBMITTER: [Alves Pinto Icaro](#)

It has been assigned
the following **PMDB id** : **PM0081549**