

UNIVERSIDADE FEDERAL DE MINAS GERAIS

FACULDADE DE FARMÁCIA

**PROGRAMA DE PÓS-GRADUAÇÃO EM ANÁLISES CLÍNICAS E
TOXICOLÓGICAS**

ALESSANDRA LOURES ROCHA

**ESTUDO *IN SILICO* DE POTENCIAIS ALVOS PROTEICOS PARA MOLÉCULAS
CITOTÓXICAS EM LINHAGENS DE CÉLULAS LEUCÊMICAS HUMANAS**

Belo Horizonte

2020

ALESSANDRA LOURES ROCHA

**ESTUDO *IN SILICO* DE POTENCIAIS ALVOS PROTEICOS PARA MOLÉCULAS
CITOTÓXICAS EM LINHAGENS DE CÉLULAS LEUCÊMICAS HUMANAS**

Dissertação apresentada ao Curso de Pós-Graduação em Análises Clínicas e Toxicológicas da Faculdade de Farmácia da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do grau de Mestra em Análises Clínicas e Toxicológicas.

Orientador: Prof. Dr. Adriano de Paula Sabino.

Coorientador: Prof. Dr. Vinícius Gonçalves Maltarollo

Belo Horizonte

2020

R672e Rocha, Alessandra Loures.
Estudo *in silico* de potenciais alvos proteicos para moléculas citotóxicas em linhagens de células leucêmicas humanas / Alessandra Loures Rocha. – 2020.
82 f. : il.

Orientador: Adriano de Paula Sabino.
Coorientador: Vinícius Gonçalves Maltarollo.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Farmácia, Programa de Pós-Graduação em Análises Clínicas e Toxicológicas.

1. Leucemia – Teses. 2. Triterpenos – Teses. 3. Alcaloides – Teses.
4. Células – Teses. I. Sabino, Adriano de Paula. II. Maltarollo, Vinícius Gonçalves. III. Universidade Federal de Minas Gerais. Faculdade de Farmácia. IV. Título.

CDD: 616.99419



UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM ANÁLISES CLÍNICAS E TOXICOLÓGICAS



FOLHA DE APROVAÇÃO

ESTUDO IN SILICO DE POTENCIAIS ALVOS PROTEICOS PARA MOLÉCULAS CITOTÓXICAS EM LINHAGENS DE CÉLULAS LEUCÊMICAS HUMANAS.

ALESSANDRA LOURES ROCHA

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ANÁLISES CLÍNICAS E TOXICOLÓGICAS, como requisito para obtenção do grau de Mestre em ANÁLISES CLÍNICAS E TOXICOLÓGICAS, área de concentração ANÁLISES CLÍNICAS E TOXICOLÓGICAS.

Aprovada em 30 de novembro de 2020, pela banca constituída pelos membros:

Prof. Adriano de Paula Sabino - Orientador
UFMG

Prof. Vinicius Goncalves Maltarollo
UFMG

Prof. Michell de Oliveira Almeida
USP

Prof. Adolfo Henrique de Moraes Silva
UFMG

Belo Horizonte, 30 de novembro de 2020.

AGRADECIMENTOS

A minha família pelo apoio, incentivo e compreensão nos momentos que precisei ficar ausente;

Ao meu marido pelo amor, carinho e suporte;

Aos meus orientadores Adriano de Paula Sabino e Vinícius Gonçalves Maltarollo, pelo conhecimento, disponibilidade, inspiração e dedicação;

Às pesquisadoras Fernanda Cristina Gontijo Evangelista e Mariana Guerra de Aguiar pelo fornecimento dos dados preliminares.

Aos colegas de trabalho pela cooperação, disponibilidade e suporte no meu dia a dia;

Ao Programa de Pós-Graduação em Análises Clínicas e Toxicológicas da Universidade Federal de Minas Gerais por fornecer a estrutura de trabalho.

RESUMO

As leucemias são neoplasias malignas caracterizadas pela proliferação de células imaturas do sistema hematopoiético. A Leucemia Mieloide Aguda (LMA) é uma neoplasia maligna caracterizada pelo acúmulo de mieloblastos na medula óssea e sangue periférico, sendo uma das leucemias mais prevalentes no adulto. A heterogeneidade genômica da doença dificulta o estabelecimento de protocolos de tratamento. Já a Leucemia Mieloide Crônica (LMC) é uma doença mieloproliferativa rara caracterizada pelo acúmulo de células mieloides e presença do cromossomo Filadélfia (Ph). Os inibidores de tirosina quinase são os principais agentes quimioterápicos disponíveis, porém existem pacientes que não respondem adequadamente ao tratamento. Portanto, a busca por novos compostos com atividade antitumoral é justificada. Neste trabalho foi realizado estudo *in silico* para a busca dos alvos proteicos de nove compostos com atividade antitumoral conhecida frente às linhagens THP-1 e K562. Três desses compostos são triterpenos (1 ao 3) e seis são análogos alcaloides de 3-alquilpiridina (4, 5, 6, 7, 8 e 9). Foi utilizada a metodologia de *target fishing* para a busca dos alvos, baseado no princípio da similaridade em que moléculas semelhantes podem se ligar aos mesmos alvos. O banco de dados ChEMBL foi utilizado na busca de moléculas com atividade citotóxica frente às linhagens THP-1 e K562, para em seguida serem comparadas aos compostos de interesse por meio do cálculo do coeficiente de Tanimoto (Tc), que avalia a similaridade entre moléculas. O Tc foi calculado usando os *fingerprints* MACCS e PubChem. Os compostos com valores de Tc maiores ou igual a 0,75 foram considerados para a pesquisa bibliográfica de alvos já descritos na literatura. Outra estratégia foi o uso de servidores para a predição de alvo molecular. As proteínas 6-fosfofruto-2-quinase/frutose-2,6-bifosfatase 3 (PFKFB3) e nicotinamida fosforibosiltransferase (NAMT) foram indicadas como possíveis alvos para o composto 7 e a fosfatase 2 indutora da fase M (CDC25B) para os compostos 1 e 3. As estruturas cristalográficas desses potenciais alvos foram obtidas por meio do *Protein Data Bank* (PDB). Os compostos 7, 1 e 3 foram submetidos a simulações de acoplamento molecular com o programa GOLD. O estudo *in silico* foi bem-sucedido na busca de alvos moleculares para o composto 7, visto que interações importantes descritas na literatura com o sítio de ligação das proteínas PFKFB3 e NAMPT foram observadas. Por outro lado, os compostos 1 e 3 não se ligaram no sítio de ligação da proteína CDC25B de forma favorável.

Palavras-chave: Leucemia. Triterpenos. Análogos alcaloides de 3-alquilpiridina. Alvo molecular. *Target fishing*. Acoplamento molecular.

ABSTRACT

Leukemias are malignant neoplasms characterized by the proliferation of immature cells of the hematopoietic system. Acute Myeloid Leukemia (AML) is a malignant neoplasm characterized by the accumulation of myeloblasts in the bone marrow and peripheral blood, being one of the most prevalent leukemias in adults. The genomic heterogeneity of the disease makes it difficult to establish treatment protocols. On the other hand, Chronic Myeloid Leukemia (CML) is a rare myeloproliferative disease characterized by the accumulation of myeloid cells and the presence of the Philadelphia (Ph) chromosome. Tyrosine kinase inhibitors are the main chemotherapeutic agents available, but there are patients who do not respond adequately to treatment. Therefore, the search for new compounds with antitumor activity is justified. In this work, an *in silico* study was carried out to search for protein targets for nine compounds with known antitumor activity against the THP-1 and K562 strains. Three compounds are triterpenes (1 to 3) and six are alkaloid analogs of 3-alkylpyridine (4, 5, 6, 7, 8 and 9). The target fishing methodology was used to search for targets, based on the principle of similarity in which similar molecules can bind to the same targets. The ChEMBL database was used to search for molecules with cytotoxic activity against the THP-1 and K562 strains, to then be compared to the compounds of interest by calculating the Tanimoto coefficient (Tc), which assesses the similarity between molecules. The Tc was calculated using the MACCS and PubChem fingerprints. Compounds with Tc values greater than or equal to 0.75 were considered for the bibliographic search of targets already described in the literature. Another strategy was the use of servers for the prediction of molecular targets. The proteins 6-phosphofruct-2-kinase/fructose-2,6-bisphosphatase 3 (PFKFB3) and nicotinamide phosphoribosyltransferase (NAMPT) were indicated as possible targets for compound 7 and phase M inducing phosphatase 2 (CDC25B) for the compounds 1 and 3. The crystallographic structures of these potential targets were obtained using the Protein Data Bank (PDB). Compounds 7, 1 and 3 were submitted to docking simulations with the GOLD program. The *in silico* study was successful in the search for molecular targets for compound 7, since important interactions described in the literature with the binding site of the PFKFB3 and NAMPT proteins were observed. On the other hand, compounds 1 and 3 did not bind at the CDC25B protein binding site favorably.

Keywords: Leukemia. Triterpenes. Alkaloid analogues of 3-alkylpyridine. Molecular target. Target fishing. Molecular docking.

LISTA DE FIGURAS

Figura 1 - A origem das células-tronco leucêmicas.....	16
Figura 2 - Desenvolvimento e progressão da LMC.....	24
Figura 3 - Estrutura dos triterpenos 1, 2 e 3.....	27
Figura 4 - Estrutura dos análogos alcaloides de 3-alkilpiridina 4 a 9.....	28
Figura 5 - Representação hipotética de um <i>fingerprint</i> de subestrutura de 10 <i>bits</i>	31
Figura 6 - Compostos do ChEMBL que possuem estudos de mecanismo de toxicidade publicados e $T_c \geq 0,75$ em relação aos triterpenos pentacíclicos.....	41
Figura 7 - Compostos do ChEMBL que possuem estudos de mecanismo de toxicidade publicados e $T_c \geq 0,75$ em relação aos análogos alcaloides de 3-alkilpiridina.....	42
Figura 8 - Mecanismo de ação do ácido betulínico na indução da apoptose.....	43
Figura 9 - Via de sinalização PI3K/PTEN.....	45
Figura 10 - Mecanismo de proliferação celular na presença de PFKFB3.....	47
Figura 11 - Visão geral da função reguladora das fosfatases CDC25A, B e C na progressão do ciclo celular.....	50
Figura 12 - Representação esquemática da seleção de alvos moleculares.....	51
Figura 13 - Estrutura molecular do ligante GV5.....	52
Figura 14 - Compostos A e B derivados de N-aryl 6-aminoquinoxalina.....	53
Figura 15 - Sobreposição das poses mais bem ranqueadas no <i>redocking</i> (A) e acoplamento (B, C e D) da proteína PFKFB3.....	54
Figura 16 - Estrutura molecular do ligante 20P.....	56
Figura 17 - Estrutura molecular do FK866.....	57
Figura 18 - Sobreposição das poses mais bem ranqueadas no <i>redocking</i> (A) e acoplamento (B, C e D) da proteína NAMPT.....	58
Figura 19 - Estrutura molecular do ligante 3M8.....	60
Figura 20 - Estrutura molecular do inibidor 2-fluoro-4-hidroxibenzonitrila.....	60
Figura 21 - Sobreposição das poses mais bem ranqueadas no <i>redocking</i> (A) e acoplamento (B, C e D) da proteína CDC25B com o composto 1.....	62
Figura 22 - Sobreposição das poses mais bem ranqueadas no <i>redocking</i> (A) e acoplamento (B e C) da proteína CDC25B com o composto 3.....	65

LISTA DE GRÁFICOS

Gráfico 1 - Número de compostos com similaridade > 0,75 em relação aos compostos de interesse utilizando o <i>fingerprint</i> MACCS.....	40
Gráfico 2 - Número de compostos com similaridade > 0,75 em relação aos compostos de interesse utilizando o <i>fingerprint</i> PubChem.....	40

LISTA DE QUADROS

Quadro 1 - Massa molecular, <i>score</i> e <i>score</i> normalizado do ligante GV5, composto 7 e poses mais representativas e mais bem ranqueadas para o <i>redocking</i> e acoplamento com a proteína PFKFB3.....	55
Quadro 2 - Massa molecular, <i>score</i> e <i>score</i> normalizado do ligante 20P, composto 7 e poses mais representativas e mais bem ranqueadas para o <i>redocking</i> e acoplamento com a proteína NAMPT.....	59
Quadro 3 - Massa molecular, <i>score</i> e <i>score</i> normalizado do ligante 3M8, composto 1 e poses mais representativas e mais bem ranqueadas para o <i>redocking</i> e acoplamento com a proteína CDC25B.....	63
Quadro 4- Massa molecular, <i>score</i> e <i>score</i> normalizado do ligante 3M8, composto 3 e poses mais representativas e mais bem ranqueadas para o <i>redocking</i> e acoplamento com a proteína CDC25B.....	64

LISTA DE ABREVIATURAS E SIGLAS

2D	Bidimensional
3D	Tridimensional
μM	Micromolar
Å	Ångström
ABL	Oncogene de Abelson
Akt	Proteína quinase serina/treonina
AG	Algoritmo genético
AMPK	Proteína quinase ativada por AMP
APA	Alquilpiridínico
APAF-1	Fator-1 de ativação de protease apoptótica
ASCII	<i>American Standard Code for Information Interchange</i>
ATCC	<i>American Type Culture Collection</i>
ATP	Adenina trifosfato
Bcl-2	<i>B-cell lymphoma protein 2</i>
Bcl-xl	<i>B-cell lymphoma extra large</i>
BCR	Receptor de células B
CDC25A	Fosfatase 1 indutora da fase M
CDC25B	Fosfatase 2 indutora da fase M
CDC25C	Fosfatase 3 indutora da fase M
CDK	Quinase dependente de ciclina
CI ₅₀	Concentração de inibição do crescimento de 50% das células
CTH	Célula-tronco hematopoética
CTL	Célula-tronco leucêmica
DNA	Ácido desoxirribonucleico
DIT	Duplicação interna em tandem
F1,6-BP	Frutose-1,6-bisfosfato
F2,6-BP	Frutose-2,6-bisfosfato
F6P	Frutose-6-fosfato
FAB	Franco-Americano-Britânico
FIA	Fator indutor de apoptose
FDA	<i>Food and Drug Administration</i>
FISH	Hibridização fluorescente <i>in situ</i>

FN-κB	Fator nuclear kappa B
GI ₅₀	<i>Growth inhibitory dose 50%</i>
GO	Gemtuzumab ozogamicina
IDH	Isocitrato desidrogenase
IS	Índice de seletividade
LLC	Leucemia Linfocítica Crônica
LMA	Leucemia Mieloide Aguda
LMC	Leucemia Mieloide Crônica
LPA	Leucemia Promielocítica Aguda
m ²	Metro quadrado
MC	Monte Carlo
mg	Miligrama
MM	Massa molecular
MOMP	Permeabilização da membrana externa mitocondrial
mTOR	Alvo de mamíferos de rapamicina
MTT	3-(4,5-dimetiltiazol-2-il)-2,5-difeniltetrazol
NADH	Nicotinamida adenina dinucleotídeo
NAD ⁺	Dinucleotídeo de nicotinamida
NAMPT	Nicotinamida fosforibosiltransferase
OMS	Organização Mundial da Saúde
PARP	Poli-ADP-ribose polimerase
PDB	<i>Protein Data Bank</i>
PBMC	Células mononucleares do sangue periférico
PI3K	Fosfatidilinositol 3-quinase
PFK-1	Fosfofrutoquinase-1
PFKFB	6-fosfofruto-2-quinase/frutose-2,6-bifosfatase
PFKFB3	6-fosfofruto-2-quinase/frutose-2,6-bifosfatase 3
PKA	Proteína quinase A
PKB	Proteína quinase B
PKC	Proteína quinase C
Ph	Philadelphia
PTEN	Fosfatase com homologia de tensina
PTPN7	Não receptor tirosina-proteína fosfatase tipo 7
kDa	Quilodalton

RC	Remissão completa
RMN	Ressonância magnética nuclear
RMSD	Desvio quadrático médio
RNA	Ácido ribonucleico
ROS	Espécie reativa de oxigênio
RT-PCR	Reação em cadeia da polimerase por transcrição reversa
SAR	<i>Structure-Activity Relationships</i>
SBDD	<i>Structure-based drug design</i>
SMILES	<i>Simplified Molecular Input Line Entry Specification</i>
STAT3	Transdutor de sinal e ativador da transcrição 3
Tc	Coefficiente de Tanimoto
TCTH	Transplante de células-tronco hematopoiéticas
TKI	Inibidor da tirosina quinase
TLR	Receptor <i>toll-like</i>
U	Unidade
UFMG	Universidade Federal de Minas Gerais

SUMÁRIO

1 INTRODUÇÃO	12
2 REFERENCIAL TEÓRICO	14
2.1 Aspectos gerais do câncer	14
2.2 Leucemias	15
2.3 Leucemia Mieloide Aguda	15
2.4 Leucemia Mieloide Crônica	22
2.5 <i>Maytenus quadrangulata</i>	26
2.6 Alcaloides marinhos	27
2.7 Determinação <i>in silico</i> de alvo molecular (<i>target fishing</i>)	29
2.7.1 Análise de similaridade química	29
2.7.2 Acoplamento Molecular	32
3 JUSTIFICATIVA	34
4 OBJETIVOS	35
4.1 Objetivo Geral	35
4.2 Objetivos específicos	35
5 METODOLOGIA	35
5.1 Identificação de alvos moleculares por similaridade química	35
5.2 Predição do modo de ligação utilizando acoplamento molecular	37
6 RESULTADOS E DISCUSSÃO	39
6.1 Análise de similaridade química	39
6.2 Predição de alvo molecular empregando Servidores	49
6.3 Acoplamento molecular	51
6.3.1 Proteína PFKFB3	52
6.3.2 Proteína NAMPT	56
6.3.3 Proteína CDC25B.....	59
7 CONCLUSÃO E PERSPECTIVAS	66
REFERÊNCIAS	68
ANEXO I – Compostos ChEMBL com Tc \geq 0,75 em relação aos compostos de interesse	76
ANEXO II – Estruturas cristalográficas depositadas no PDB para cada alvo proteico estudado	81

1 INTRODUÇÃO

As leucemias são neoplasias malignas caracterizadas pela proliferação de células imaturas do sistema hematopoiético e são classificadas de acordo com a linhagem acometida, mieloide ou linfoide, e o estado de maturação das células leucêmicas, agudo ou crônico (AN; FAN; XU, 2017). A leucemia mieloide aguda (LMA) consiste em um grupo heterogêneo de neoplasias malignas relacionado com as células hematopoiéticas, representando um dos tipos mais comuns da leucemia em adultos. A LMA é caracterizada por uma transformação clonal das células precursoras hematopoiéticas, mediante processo de múltiplas etapas, por meio da aquisição de rearranjos cromossômicos ou diferentes mutações genéticas. A diminuição da velocidade de autodestruição e parada na diferenciação celular, leva ao acúmulo de mieloblastos na medula óssea e sangue periférico (HASSERJIAN, 2013). Essa heterogeneidade genômica interfere diretamente na classificação da LMA e na definição dos protocolos de tratamento (YU *et al.*, 2020). Já a leucemia mieloide crônica (LMC) é caracterizada pela presença da translocação t (9; 22) (q34; q11.2) que resulta em um cromossomo 22 encurtado chamado cromossomo Filadélfia (Ph), que codifica a oncoproteína de fusão BCR-ABL1. A LMC é uma doença mieloproliferativa caracterizada pelo acúmulo de células mieloides (POPHALI; PATNAIK, 2016). A LMC é rara, com uma incidência anual de 1-2 casos por 100.000 indivíduos e um pico de incidência na sexta a sétima década de vida (KOSCHMIEDER; VETRIE, 2018). O principal tratamento disponível é o uso dos inibidores de tirosina quinase de primeira, segunda e terceira gerações FACHI *et al.*, 2018). No entanto, uma proporção de pacientes não responde adequadamente à terapia ou experimenta uma progressão da doença durante a terapia devido ao desenvolvimento de resistência (POPHALI; PATNAIK, 2016).

No trabalho da pesquisadora Mariana Guerra de Aguiar do Departamento de Química do Instituto de Ciências Exatas da UFMG em parceria com Grupo de Pesquisa em Oncohematologia do Departamento de Análises Clínicas e Toxicológicas da Faculdade de Farmácia da UFMG, foi avaliada a atividade citotóxica *in vitro* dos compostos 1, 2 e 3, sendo estes triterpenos pentacíclicos. A citotoxicidade das amostras foi avaliada contra as linhagens celulares tumorais humanas THP-1 (células de leucemia mieloide aguda, ATCC-TIB-202), K562 (células de leucemia mieloide crônica, ATCC-CRL-3344), MDA-MB-231 (células de carcinoma mamário, ATCC-HTB-26) e TOV-21G (células de carcinoma de ovário, ATCC-HTB-26). A citotoxicidade dos compostos frente à linhagem Wi-26VA4 (células de fibroblastos de pulmão, ATCC- CCL-75) foi usada para estabelecer o índice de seletividade.

Foram considerados neste trabalho apenas os resultados referentes às linhagens leucêmicas THP-1 e K562. Para a linhagem THP-1 os resultados para os compostos 1, 2 e 3 foram comparáveis ao quimioterápico padrão para o modelo em estudo, a citarabina ($p > 0,05$). Para a linhagem K-562, o composto 3 foi comparável ao inibidor de tirosina quinase imatinibe ($p > 0,05$). Em relação à seletividade frente à linhagem THP-1, os compostos 1 (IS = 3,59), 2 (IS = 6,53) e 3 (IS = 4,82) apresentaram seletividade próxima a da citarabina (IS = 4,62). Frente à linhagem K562 o composto 3 apresentou seletividade moderada (IS > 3) em relação ao controle Imatinibe (IS = 6,63).

Além disso, em estudos prévios do nosso grupo de pesquisa foi avaliada a atividade citotóxica *in vitro* dos compostos 4, 5, 6, 7, 8 e 9, moléculas sintéticas análogas ao alcaloide 3-alkilpiridínico, nas linhagens celulares humanas THP-1 (leucemia monocítica aguda ATCC TIB-202) e K562 (leucemia mieloide crônica ATCC CCL-243) obtidas da *American Type Culture Collection* (ATCC) e células mononucleares do sangue periférico (PBMC) (não tumoral), utilizando como compostos citotóxicos padrões os quimioterápicos comumente utilizados nos modelos avaliados, citarabina e imatinibe. Foi observado ainda que os compostos induziram majoritariamente a morte celular por apoptose nas células leucêmicas (K562 e THP-1). Além disso, a progressão do ciclo celular para a fase de síntese (fase S) nas linhagens leucêmicas foi impedida pelos compostos testados. A inibição da progressão do ciclo celular depende principalmente da inibição da expressão das ciclinas, que ativam as CDKs (quinases dependentes de ciclinas) que são controladoras positivas do ciclo celular. Para a linhagem THP-1, os compostos 4, 5, 6 e 7 foram comparáveis e mais seletivos em relação à citarabina, composto já utilizado como quimioterápico ($p < 0,05$). Já na linhagem K562 os compostos mais efetivos e seletivos foram o 8 e 9 ($p < 0,05$), em comparação ao imatinibe.

Nos últimos anos, têm sido amplamente utilizados os métodos de determinação *in silico* de alvos moleculares, também chamados *target fishing*. Em geral, estes métodos se baseiam no princípio da similaridade, indicam que moléculas similares podem se ligar aos mesmos alvos e apresentarem atividades semelhantes. Em combinação com a aplicação dos métodos de acoplamento molecular, estas estratégias são úteis para a predição do modo de ligação entre as substâncias estudadas e o(s) alvo(s) molecular(es) preditos (CERETO-MASSAGUÉ *et al.*, 2015a).

Nesse contexto, o emprego de técnicas computacionais para determinação de alvos moleculares e estudo de mecanismo de ação de compostos citotóxicos se torna uma ferramenta com boa acurácia e baixo custo. Portanto, o presente estudo teve por objetivo

identificar os possíveis alvos moleculares dos compostos 1 a 9 que apresentaram atividade citotóxica frente às linhagens celulares THP-1 e K-562, utilizando métodos *in silico* de similaridade e acoplamento molecular.

2 REFERENCIAL TEÓRICO

2.1 Aspectos gerais do câncer

O câncer é uma doença genética que ocorre quando a informação contida no DNA celular é corrompida, levando a padrões anormais de expressão gênica. O principal mecanismo pelo qual ocorre essa alteração do código genético é através do acúmulo de mutações, embora haja um reconhecimento crescente do papel das mudanças não mutacionais (epigenéticas) no processo. A expressão gênica aberrante leva a uma série de mudanças em processos biológicos fundamentais dentro das células cancerosas (HARRINGTON, 2011).

As principais alterações fisiológicas das células cancerígenas incluem: proliferação autossuficiente, evasão da supressão do crescimento, invasão de tecido e metástase, imortalidade replicativa, indução de angiogênese, resistência à morte celular, desregulação do metabolismo energético celular, instabilidade genômica, promoção tumoral da inflamação e evasão imunológica (HANAHAN; WEINBERG; 2011).

A classificação do câncer baseada no órgão de origem tem sido utilizada com sucesso por muito tempo, sendo divididos em tumores sólidos (por exemplo: câncer de mama, câncer de pulmão, câncer de cólon, câncer colorretal, câncer de próstata) e tumores líquidos (as leucemias). Com a evolução das técnicas moleculares de diagnóstico, um número crescente de alterações genéticas e epigenéticas têm sido consideradas na identificação dos tipos de cânceres. Apesar de cada tumor ser único, a classificação do tumor e o diagnóstico molecular podem beneficiar os indivíduos com base na premissa de que tumores semelhantes surgem, evoluem e se comportam de maneiras semelhantes (OGINO *et al.*, 2012).

O câncer é a segunda causa de morte em todo o mundo e relatórios recentes revelaram que até 2025, um aumento de 19,3 milhões de novos casos por ano é esperado. Cirurgia, radioterapia e quimioterapia são as principais terapias anticâncer convencionais. No entanto, a eficácia limitada dessas terapias e seus graves efeitos colaterais têm incentivado a busca por novos agentes anticâncer baseados em produtos naturais como terapias únicas ou em combinação com outros agentes (TALIB, 2018).

2.2 Leucemias

As leucemias são neoplasias malignas caracterizadas pela proliferação de células imaturas do sistema hematopoiético, resultante da parada da maturação celular ou diferenciação anormal, levando ao acúmulo de células imaturas na medula óssea (PEJOVIC; SCHWARTZ, 2002).

As leucemias são classificadas conforme a linhagem celular acometida, mieloide ou linfóide, e de natureza aguda ou crônica. As leucemias crônicas se caracterizam pela proliferação excessiva de células maduras e são raras em pacientes pediátricos, sendo mais comuns em adultos e idosos. Nas leucemias agudas, por outro lado, as células são imaturas e incapazes de se diferenciarem, comumente ocorrem em pacientes de todas as idades e são potencialmente fatais com evolução rápida se não forem prontamente tratados (AN; FAN; XU, 2017).

2.3 Leucemia Mieloide Aguda

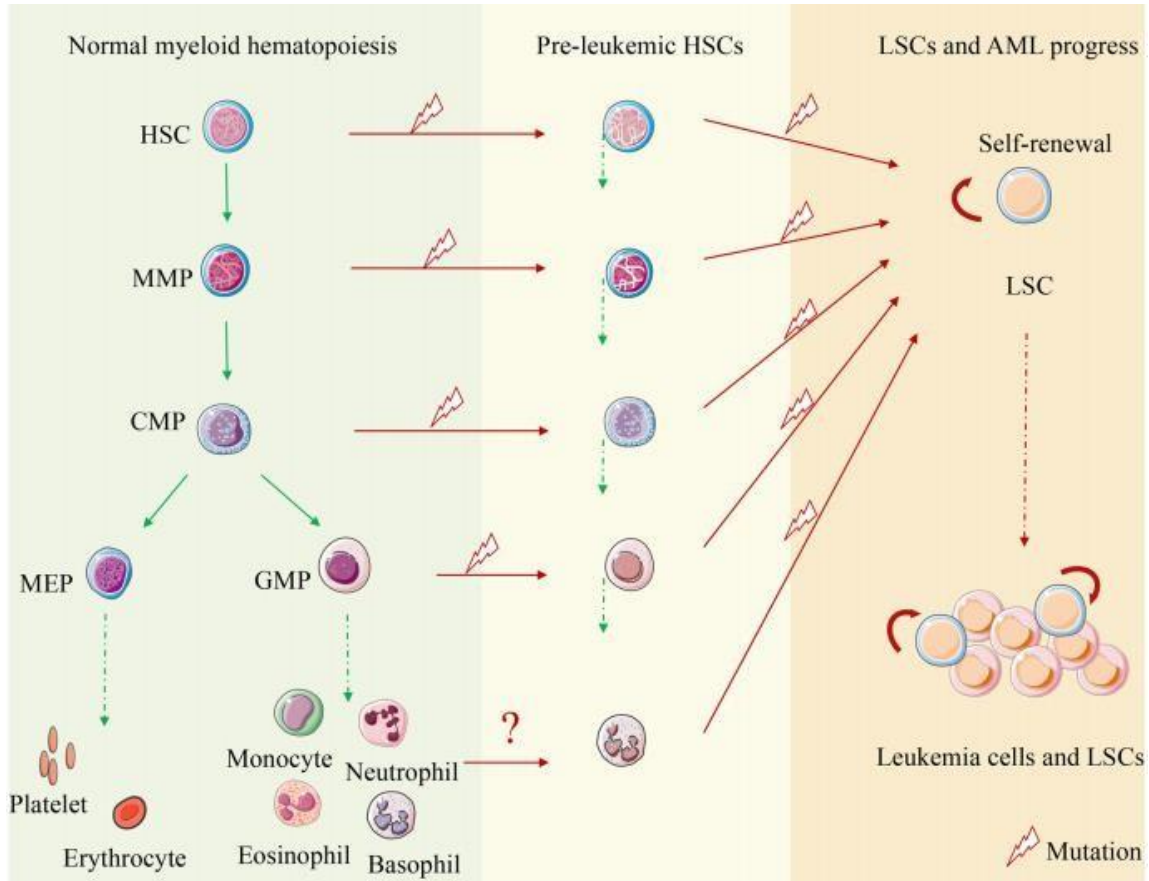
A leucemia mieloide aguda (LMA) é uma doença clonal do tecido hematopoiético, caracterizada pela proliferação anormal das células progenitoras mieloides, resultando em uma geração insuficiente de células sanguíneas maduras normais (SONG *et al.*, 2018). A LMA é uma neoplasia mieloide clonal agressiva com parada da maturação da mielopoiese, levando ao acúmulo de mieloblastos na medula óssea e/ou no sangue (HASSERJIAN, 2013). A LMA tem a maior incidência em comparação com outras leucemias agudas e também a maior taxa de mortalidade entre as doenças do sistema hematológico (LIN; CHEN, 2018).

As células-tronco leucêmicas (CTLs) estão envolvidas na patogênese e são a fonte de manutenção das células leucêmicas. Pacientes com um número maior de CTLs no diagnóstico têm risco significativamente maior de recidiva do que pacientes com menos LSCs. Células-tronco hematopoéticas (CTHs) normais e progenitores podem ser convertidos em CTHs pré-leucêmicas quando ocorrem mutações precoces. CTHs pré-leucêmicas têm melhor vantagem clonal que as CTHs normais e não perdem sua capacidade de diferenciação. Com o acúmulo de mutações, as CTHs pré-leucêmicas são eventualmente transformados em CTLs, que iniciam a LMA (TAN *et al.*, 2020) (**Figura 1**).

A LMA é uma neoplasia caracterizada por anormalidades cromossômicas, mutações genéticas recorrentes, modificações epigenéticas que afetam a estrutura da

cromatina e desregulam micro RNAs. Heterogeneidade genômica, variabilidade individual dos pacientes, e mutações genéticas recorrentes são alguns dos obstáculos entre os muitos fatores que afetam a eficácia do tratamento dos pacientes com LMA (YU *et al.*, 2020).

Figura 1 - A origem das células-tronco leucêmicas.



Fonte: TAN *et al.*, 2020.

Legenda: HSCs, células-tronco hematopoiéticas; LSCs, células-tronco leucêmicas; AML, leucemia mieloide aguda; MMP, progenitor multipotente; mCMP, progenitor mielóide comum; MEP, progenitor megacariócito-eritrócito; GMP, progenitor granulócito-macrófago.

De 1977 até 2001, o sistema criado pelo grupo colaborativo de hematologistas Franco-Americano-Britânico (FAB), que é baseado em características citomorfológicas e citoquímicas e imunofenotípicas das células leucêmicas, foi utilizado para classificar a LMA em oito grupos (M0-M7). Em 2001, uma nova classificação da LMA foi desenvolvida e publicada com contribuição da Organização Mundial da Saúde (OMS), Associação Europeia de Hematopatologia e Sociedade de Hematopatologia (TAMAMYAN *et al.*, 2017).

Com a aplicação de novas técnicas moleculares, como as tecnologias de sequenciamento de nova geração, uma enorme diversidade de mutações genéticas tem sido identificada (YU *et al.*, 2020). Quase todos os pacientes com LMA (97%) carregam pelo

menos uma mutação somática. As diretrizes da OMS utilizam alterações citogenéticas e mutações genéticas recorrentes, descritas pela primeira vez em 2008 e atualizadas em 2016, para a classificação da LMA. Alterações citogenéticas e anormalidades genéticas são úteis para estratificação do risco e também pode ajudar na escolha do tratamento (TALATI; SWEET, 2018).

Em 2016 a OMS lançou uma nova versão da classificação das neoplasias mieloides e leucemias agudas. Para a LMA foram incorporadas informações genéticas à morfologia, imunofenotipagem e apresentação clínica para definir seis principais entidades patológicas: LMA com anormalidades genéticas recorrentes; LMA com características relacionadas à mielodisplasia; LMA relacionada à terapia; LMA não especificada de outra maneira; sarcoma mieloide; proliferação mieloide relacionada à Síndrome de Down. Entre os casos de LMA com anormalidades genéticas recorrentes, 11 subtipos foram delineados de acordo com translocações cromossômicas distintas (KOUCHKOVSKY; ABDUL-HAY, 2016). A OMS continuou a definir a LMA com foco em citogenética e subgrupos genéticos moleculares. Um grande número de anormalidades citogenéticas foi reconhecido na LMA, e a maioria das que não foram formalmente reconhecidas pela classificação são raras (ARBER *et al.*, 2016).

O diagnóstico baseia-se na demonstração de que a medula ou sangue tem mais que 20% de blastos de linhagem mieloide. A linhagem dos blastos é avaliada também por citometria de fluxo, sendo CD33 e CD13 marcadores de superfície tipicamente expressos por blastos mieloides (ESTEY, 2016). O diagnóstico da LMA também pode ser estabelecido na presença de tecido extramedular infiltrado, ou presença de t(8;21), inv(16) ou t(15;17) em um cenário clínico apropriado, independentemente do percentual de blastos (KOUCHKOVSKY; ABDUL-HAY, 2016).

Translocações cromossômicas, como t(8;21) ou t(15;17) na Leucemia Promielocítica Aguda (LPA) resultam na formação das proteínas quiméricas RUNX1-RUNX1T1 e PML-RARA, respectivamente, que alteram o processo normal de maturação das células precursoras mieloides. Além dos grandes rearranjos cromossômicos, alterações moleculares também implicam no desenvolvimento da LMA (KOUCHKOVSKY; ABDUL-HAY, 2016). Mutações em genes que codificam modificadores epigenéticos, como DNMT3A, ASXL1, TET2, IDH1 e IDH2, são adquiridas precocemente e estão presentes no clone fundador. Esses genes mutados são frequentemente encontrados em indivíduos idosos com expansão clonal da hematopoiese. Além disso, essas mutações podem persistir após o tratamento, levar à expansão clonal durante a remissão hematológica, e eventualmente à

recidiva da doença. Por outro lado, mutações envolvendo NPM1 ou moléculas de sinalização como FLT3 e RAS, tipicamente são eventos secundários que ocorrem durante a leucemogênese (BULLINGER; DÖHNER K.; DÖHNER H., 2017).

Mutações que resultam na ativação das vias pró-proliferativas como FLT3 (duplicações internas em tandem - DIT), e mutações no domínio tirosino quinase, K/NRAS, TP53 e c-KIT são encontradas em aproximadamente 28, 12, 8 e 4% dos casos, respectivamente. Estudos de neoplasias malignas sólidas e hematológicas também destacaram o papel do transdutor de sinal e ativador da transcrição 3 (STAT3) no estímulo da sobrevivência e proliferação celular. O aumento da fosforilação de STAT3 devido à elevação na secreção de interleucina 6 ou mutações no receptor de tirosina quinase (por exemplo, duplicações FLT3 ou menos frequente JAK2) é observado em mais de 50% dos casos de LMA e significa pior prognóstico. Mutações que incluem os genes NPM1 e CEBPA, que são encontradas em aproximadamente 27% e 6% dos casos, respectivamente, conferem melhor prognóstico. Mutações em genes relacionados com a metilação do DNA como DNMT3A, TET2, IDH-1 e IDH-2, são encontradas em mais de 40% dos casos de LMA (KOUCHKOVSKY; ABDUL-HAY, 2016).

Embora mais de um fator conduzam à doença, certas anormalidades são mais comuns que outras. Por exemplo, mutações na NPM1 (ocorrendo em aproximadamente 30% de todos os casos de LMA e 50%–60% dos casos de LMA citogeneticamente normais) e FLT3 (ocorrendo em aproximadamente 30% dos pacientes com LMA de novo) estão entre as mutações mais comuns na LMA. Mutações e anormalidade cromossômicas podem co-ocorrer em uma variedade de combinações, que podem alterar seu impacto prognóstico (TALATI; SWEET, 2018).

Existem várias abordagens para o tratamento da LMA, que incluem quimioterapia, terapias alvo específicas, imunoterapia, e transplante de células-tronco hematopoiéticas. As dificuldades do tratamento refletem-se principalmente na resistência aos medicamentos, incluindo a primária, que existe antes do tratamento, e a secundária, que se desenvolve devido ao uso repetido de alguns quimioterápicos de indução. Isso leva aos refratários e/ou a recaída e reduz a sobrevida dos pacientes com LMA (LIN; CHEN, 2018).

A quimioterapia de indução intensiva permaneceu inalterada por mais de quatro décadas e é a espinha dorsal da terapia. Para adultos jovens (com idade <60 anos) e idosos aptos, o regime intensivo de citarabina e antraciclina “7 + 3” é o tratamento padrão, que combina sete dias de infusão contínua de citarabina (100–200 mg/m² diariamente) com três dias de antraciclina, tipicamente daunorrubicina (45-90 mg/m² diariamente) ou idarrubicina

(12 mg/m² diariamente). O objetivo da quimioterapia de indução é alcançar a remissão completa (RC) morfológica. Assim, pacientes com LMA de novo atingem a RC em 65-73% dos casos usando indução padrão com '7 + 3', enquanto apenas 38-62% dos pacientes com mais de 60 anos com LMA atingem RC (MERDINGER *et al.*, 2016). A dose padrão de citarabina consiste em 100–200 mg/m² administrados diariamente como infusão contínua por sete dias. Embora estudos tenham demonstrado maior eficácia em doses mais altas, esse benefício adicional é pequeno considerando o aumento da toxicidade, a terapia de indução com altas doses de citarabina é geralmente reservada para a doença refratária. A combinação de fludarabina, citarabina, G-CSF e idarrubicina (FLAG-IDA), tradicionalmente utilizada para o tratamento da recaída, também se mostrou uma alternativa razoável para os regimes de indução padrão e apresentou resultados similares nas taxas de RC e sobrevida (KOUCHKOVSKY; ABDUL-HAY, 2016).

Uma abordagem ideal para pacientes idosos ainda não foi estabelecida. Indivíduos com idade acima de 65 anos são menos propensos a responder à quimioterapia e são frequentemente mais suscetíveis a toxicidades relacionadas ao tratamento. No entanto, apesar de um significativo pior prognóstico, a terapia de indução melhora a sobrevida em pacientes com idade acima de 65 anos quando comparada aos cuidados de suporte e quimioterapia paliativa (KOUCHKOVSKY; ABDUL-HAY, 2016).

A consolidação ou terapia de pós-indução é dada para prevenir a recaída e erradicar a leucemia residual mínima na medula óssea após a indução como uma ponte para o transplante ou para alcançar a cura. A avaliação da doença residual mínima usando reação em cadeia da polimerase em tempo real ou sequenciamento de próxima geração está sendo cada vez mais usada para ajudar a rastrear a resposta ao tratamento e tem se mostrado superior à morfologia isolada na previsão de recaída iminente (MERDINGER *et al.*, 2016).

Em geral, existem duas estratégias principais para consolidação; quimioterapia (incluindo agentes direcionados) e transplante de células-tronco hematopoiéticas (TCTH). Ambas as estratégias podem ser utilizadas isoladamente ou mais comumente em combinação, dependendo do tipo de leucemia, da adequação do paciente e da disponibilidade de um doador de células-tronco. Uma das decisões de tratamento mais importantes na LMA é estimar o benefício e o risco associados ao TCTH alogênico em primeira remissão para um determinado paciente. (MERDINGER *et al.*, 2016).

Durante a última década, estudos genômicos de neoplasias hematológicas identificaram um espectro de alterações somáticas recorrentes que contribuem para a iniciação e manutenção de LMA, e que conferem sensibilidade a terapias direcionadas molecularmente.

Algumas dessas novas drogas inibidoras direcionadas aumentaram a eficácia do tratamento da LMA (LIN; CHEN, 2018).

Muitos estudos investigaram o papel dos agentes seletivos, particularmente os inibidores de tirosina quinase em pacientes com LMA FLT3-mutado (POLLYEA, 2018).

Outras estratégias que incluem os inibidores IDH-1/IDH-2, CPX-351, gemtuzumab ozogamicina (GO) e o venetoclax foram direcionadas a uma subpopulação específica de pacientes com LMA e já foram aprovadas ou seguem em estudo (BURNETT; STONE, 2020).

Os inibidores de FLT3 podem ser considerados a maior família de agentes alvo-específicos. Estes são divididos em duas categorias principais. Os medicamentos de primeira geração incluem sorafenibe, lestaurtinibe, midostaurina, sunitinibe, e tandutinibe. O sorafenibe, amplamente usado, demonstrou desempenhar um papel importante no tratamento da LMA refratária ou recidiva em pacientes com mutações em FLT3. Os agentes de segunda geração, prometendo ter maior potência e causar menos efeitos colaterais, incluem crenolanibe, quizartinibe e cabozantinibe. No entanto, o grande desafio no uso de inibidores de FLT3 é a resistência. As mutações pontuais que levam à resistência incluem N676, F691 e D835, juntos com o FLT3-DIT (LIN; CHEN, 2018).

Ambas as isoformas mutadas da enzima isocitrato desidrogenase (IDH) contribuem para a leucemogênese através da superprodução do oncometabolito 2-hidroxi-glutarato (ao invés de alfa-cetoglutarato produzido pela enzima selvagem). O enasidenibe foi aprovado pelo FDA em agosto de 2017 para pacientes com LMA recidivada ou refratária com IDH-2 mutado. O ivosidenibe também foi desenvolvido para pacientes com LMA recidivada/refratária (com IDH-1 mutado) e apresentou atividade similar ao enasidenibe. A aprovação pelo FDA ocorreu em julho de 2018. Desde então, o medicamento foi aprovado como agente único em pacientes com IDH-1 mutado que antes não eram tratados por não se adequarem à quimioterapia padrão (BURNETT; STONE, 2020; MARTELLI *et al.* 2020). Entre os pacientes com LMA recidivada/refratária, ambos os agentes produziram taxas de RC da medula óssea de 30% a 40% e sobrevida média de 8 a 9 meses. Ambos os agentes estão sendo investigados em combinação com quimioterapia intensiva ou terapias de hipometilação entre pacientes com as mutações correspondentes (KANTARJIAN; KEATING; FREIREICH, 2018).

O CPX-351 é um lipossomo em nanoescala que possui a relação molar fixa (5:1) de citarabina e daunorrubicina. Foi provado que este lipossomo possui uma combinação ideal, maximizando o sinergismo e evitando antagonismo. Ensaios clínicos mostraram uma vantagem óbvia em pacientes idosos com LMA secundária (LMAs, ou seja, LMA relacionada

à terapia ou LMA com histórico de alteração hematológica prévia). Um estudo randomizado, multicêntrico, aberto, fase III comparou CPX-351 (100 U/m²) com daunorrubicina (60 mg/m²) mais citarabina (100 mg/m²) em 309 pacientes idosos com idade entre 60 e 75 anos com LMA de alto risco. Os resultados indicaram que o grupo CPX-351 alcançou melhor tempo de sobrevida e maiores taxas de RC (LIN; CHEN, 2018).

O gentuzumab ozogamicina (GO) é um conjugado anticorpo-droga anti-CD33 que atua nas células leucêmicas com a caliqueamicina tóxica associada ao anticorpo. CD33 é expresso por mais de 80% das LMAS. O GO foi inicialmente aprovado em 2000 para pacientes idosos (com mais de 60 anos) com LMA refratária/recidiva, mas foi retirado voluntariamente em 2010 por preocupações com a segurança. A dosagem foi revisada em novos estudos sendo associada a menor toxicidade. Ele foi aprovado novamente em setembro de 2017 tanto para as leucemias recém-diagnosticadas quanto para LMA refratária/recidiva com CD33 positivo (KOENIG *et al.*, 2020).

A sinalização da proteína Bcl-2 prolonga a sobrevivência das células cancerígenas. O venetoclax, um inibidor da Bcl-2, demonstrou benefício no tratamento da leucemia linfocítica crônica (LLC). Como agente único, teve atividade modesta na recuperação da LMA, com uma taxa de resposta de 18%. No entanto, com base em estudos pré-clínicos, foram realizados ensaios combinando venetoclax com azacitidina, decitabina e citarabina em baixas doses. Esses ensaios demonstraram altas taxas de RC de 60% a 70% em pacientes idosos recém-diagnosticados com LMA que não eram adequados para quimioterapia intensiva. A mortalidade por indução foi inferior a 5% (KANTARJIAN; KEATING; FREIREICH, 2018).

A LMA é uma doença complexa, com uma paisagem genética diversa. O campo está se expandindo rapidamente, com maior compreensão da fisiopatologia e potenciais novos alvos de drogas (MERDINGER *et al.*, 2016). Os tratamentos medicamentosos para LMA incluem quimioterapia, terapia alvo específica, e imunoterapia. Dentre as diferentes quimioterapias, o regime padrão “7+3” é o de primeira escolha, mas os pesquisadores fizeram progresso significativo neste regime com base na condição dos pacientes (LIN; CHEN, 2018).

A natureza heterogênea da LMA dificulta a escolha do tratamento mais adequado uma vez que os pacientes muitas vezes preenchem os requisitos para uso de mais de uma droga. Dessa forma, mais estudos precisam ser conduzidos para avaliar a segurança e eficácia das associações das drogas já disponíveis.

Apesar das melhorias na terapia direcionada, tornou-se evidente que as terapias de fármaco único têm menor probabilidade de sucesso do que as terapias de multifármacos com

diferentes alvos. Embora o transplante de células-tronco alogênicas tenha sido tradicionalmente considerado a melhor estratégia nesse cenário, os dados disponíveis sugerem que ele pode não ser a estratégia mais eficaz para a erradicação da doença residual mínima. Novos agentes como fármacos com alvos moleculares (inibidores de FLT3 ou IDH) ou agentes baseados em anticorpos monoclonais, incluindo conjugados anticorpo-fármaco e anticorpos biespecíficos podem melhorar as estratégias terapêuticas para erradicar a persistente doença residual mínima após regimes citotóxicos (MERDINGER *et al.*, 2016).

Dessa forma, as opções de tratamento atualmente aplicadas na LMA possuem significativas limitações, sendo necessária a busca de novos alvos terapêuticos e o desenvolvimento de novas estratégias de tratamento (KAYSER; LEVIS, 2018; STEIN; TALLMAN, 2016).

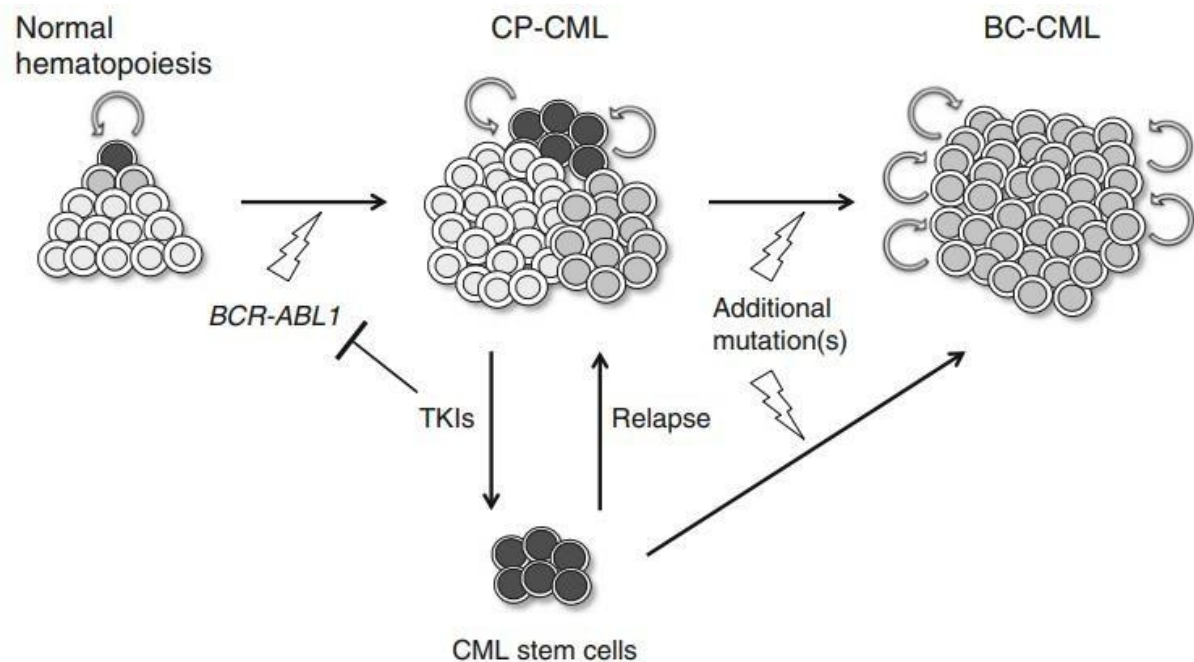
2.4 Leucemia Mieloide Crônica

A leucemia mieloide crônica (LMC) é um distúrbio clonal de células-tronco hematopoiéticas, caracterizado por uma translocação recíproca que envolve a fusão do oncogene de Abelson (ABL) localizado no cromossomo 9q34 com a região do ponto de interrupção (BCR) no cromossomo 22q11.2.1., t (9;22) (q34; q11.2), que resulta em um cromossomo 22 encurtado chamado cromossomo Filadélfia (Ph), que codifica a oncoproteína de fusão BCR-ABL1. Essa proteína tem uma atividade constitutiva de tirosina quinase, que estimula a transformação hematopoiética e a mieloproliferação. A isoforma predominante da BCR-ABL é uma proteína de 210 kDa que está presente em mais de 90% dos pacientes com LMC (POPHALI; PATNAIK, 2016). A proteína resultante da fusão BCR-ABL1 causa uma ativação anômala das vias de transdução de sinal intracelular, levando a um genoma instável, proliferação celular anormal e amplificação dos clones de LMC (LOSCOCCO *et al.*, 2019).

Existem três fases da LMC durante a progressão natural da doença: uma fase crônica inicial, seguida de uma fase acelerada para uma fase blástica ou diretamente seguida pela crise blástica. A fase acelerada e a fase blástica são frequentemente associadas com anormalidades citogenéticas adicionais (QING *et al.*, 2018) (**Figura 2**). Mais de 90% dos casos de LMC são diagnosticados na fase crônica. Cerca de 50% dos pacientes são assintomáticos nessa fase e a doença é diagnosticada em exames de sangue de rotina. Nesse estágio inicial, granulócitos maduros ainda são produzidos, mas existe um aumento de células progenitoras mieloides no sangue periférico (BAVARO *et al.*, 2019). A maioria dos casos de

LMC na fase crônica pode ser diagnosticada por achados no sangue periférico combinado com a detecção de t(9;22) (q34.1;q11.2), ou mais especificamente, BCR-ABL1 por técnicas de genética molecular. No entanto, um aspirado de medula óssea é essencial para garantir material suficiente para um cariótipo completo e avaliação morfológica para confirmar a fase da doença (ARBER *et al.*, 2016). Apenas 5% dos pacientes com LMC são diagnosticados nas fases acelerada ou blástica (QING *et al.*, 2018). Embora a fase acelerada da LMC tenha se tornado menos frequente na era da terapia com inibidores de tirosina quinase, ainda não existe um consenso universal em relação a esta definição. O critério para fase acelerada na classificação revisada pela OMS inclui parâmetros hematológicos, morfológicos e citogenéticos que são complementados por parâmetros adicionais geralmente atribuídos à evolução genética e manifestados por evidências de resistência aos inibidores de tirosina quinase (ARBER *et al.*, 2016). A fase blástica da LMC é um evento clínico grave que pode ser definido quando há 20% ou mais blastos no sangue ou na medula óssea ou quando é observada uma proliferação extramedular de blastos (QING *et al.*, 2018). No entanto, como o início da fase blástica linfoide pode ser bastante repentino, a detecção de qualquer linfoblasto “bona fide” no sangue ou medula, deve suscitar preocupação por uma possível fase blástica linfoide iminente e estudos laboratoriais e genéticos devem ser solicitados para excluir essa possibilidade (ARBER *et al.*, 2016).

Figura 2 – Desenvolvimento e progressão da LMC.



Fonte: ITO, 2013.

Legenda: A leucemia mieloide crônica (LMC) é iniciada pela translocação BCR-ABL1 em células-tronco hematopoéticas, que leva à expansão das células mieloides enquanto permite a diferenciação (CP-CML, LMC em fase crônica). Translocações cromossômicas secundárias, como NUP98-HOXA9 ou AML1-EVI1, ou mutações em p53 ou INK4A desencadeiam a progressão da doença para uma fase mais avançada (BC-CML, crise blástica LMC), com perda progressiva da capacidade de diferenciar e aumento da expansão de células blásticas imaturas. Embora os inibidores da tirosina quinase (TKIs) direcionados à atividade da quinase BCR-ABL1 possam matar as células progenitoras que expressam oncogene, eles são incapazes de eliminar as células-tronco da LMC, levando a um risco aumentado de recidiva da doença após a descontinuação da terapia com TKI.

A LMC é rara, com uma incidência anual de 1-2 casos por 100.000 indivíduos e um pico de incidência na sexta a sétima década de vida. A grande maioria (85-90%) dos pacientes se apresenta na fase crônica e, se não tratada, progride para a fase acelerada e para a crise blástica em cinco anos, assemelhando-se a uma leucemia aguda com mau prognóstico (KOSCHMIEDER; VETRIE, 2018).

O diagnóstico de LMC requer a detecção de BCR-ABL1, geralmente realizada por reação em cadeia da polimerase por transcrição reversa (RT-PCR) e hibridização fluorescente *in situ* (FISH). No entanto, pacientes com LMC geralmente são assintomáticos, a menos que estejam em estágio avançado ou tenham esplenomegalia (OGASAWARA *et al.*, 2019).

A LMC é diagnosticada por meio de contagem completa das células sanguíneas com diferencial, esfregaço de sangue periférico, exame de esplenomegalia e aspiração de medula óssea com biópsia. O diagnóstico definitivo da LMC baseia-se na presença do

cromossomo Ph (t [9; 22]) e / ou a translocação BCR-ABL. O Ph é analisado por análise citogenética das células da medula óssea, que é essencial no diagnóstico, porque isso também irá revelar se existem lesões citogenéticas adicionais, o que implica doença de fase avançada (RADICH; MAURO, 2017).

A farmacoterapia da LMC baseia-se principalmente em inibidores da tirosina quinase (TKIs). O primeiro inibidor BCR-ABL a ser utilizado na prática clínica foi o mesilato de imatinibe, que demonstrou taxas de resposta superiores, tolerabilidade, menor probabilidade de progressão para fase acelerada, e benefício de sobrevida comparado ao interferon- α associado à citarabina (FACHI *et al.*, 2018).

Com o uso de imatinibe, a sobrevida de pacientes com LMC melhorou drasticamente. No entanto, com o tempo, 20 a 30% dos pacientes desenvolvem resistência ao TKI, comumente atribuível a mutações no domínio da quinase BCR-ABL1. Alguns pacientes falham na terapia apesar da inibição da BCR-ABL1, implicando a ativação de mecanismos alternativos de resistência. E 5-10% descontinuem a terapêutica secundária devido à má tolerabilidade. Isso levou ao desenvolvimento de novos TKIs, como nilotinibe, dasatinibe, bosutinibe e ponatinibe (POPHALI; PATNAIK, 2016).

A segunda geração (dasatinibe, nilotinibe e bosutinibe) é considerada mais potente e eficaz do que o imatinibe em doentes resistentes ao imatinibe recentemente diagnosticados. O ponatinibe e o radotinibe (TKI de terceira geração) são eficazes em pacientes refratários ao dasatinibe ou ao nilotinibe e à mutação T315I (FACHI *et al.*, 2018). No início, esses pacientes com T315I foram tratados até que obtivessem uma boa resposta e, em seguida, receberam transplante de células tronco hematopoiéticas (TCTH) alogênico. No entanto, os dados de longo prazo com ponatinibe, particularmente com redução da dose de 45 mg por via oral diariamente (dose aprovada pela FDA, mas associada a efeitos colaterais graves, incluindo hipertensão em $\geq 50\%$, doença vaso-oclusiva grave em 20% a 30% e erupções cutâneas em 20% a 30%) a 30 ou 15 mg por dia, são positivos. Isso dá aos pacientes a opção de continuar terapia oral, incluindo TKIs de segunda ou terceira geração, se for observada uma boa resposta (citogenética) ou passar para o TCTH alogênico devido a questões de custo, toxicidade ou escolha do paciente / médico (KANTARJIAN; KEATING; FREIREICH, 2018).

Apesar do sucesso do imatinibe e de outros TKIs na melhora dos resultados da LMC, uma proporção de pacientes não responde adequadamente à terapia ou experimenta uma progressão da doença durante a terapia. A proporção de pacientes que responde depende do tempo de avaliação e da escolha do TKI. Os TKIs de segunda e terceira geralmente

fornecem uma resposta mais rápida e taxas superiores de resposta molecular profunda. As taxas de progressão, no entanto, parecem relativamente constantes em todo o espectro de TKIs (YURTTAŞ; EŞKAZAN, 2020).

2.5 *Maytenus quadrangulata*

Os recursos naturais têm sido usados há milhares de anos para combater doenças humanas. Durante a última década, o interesse pelos produtos naturais e seus mecanismos de ação foi intensificado e, agora, substâncias que ocorrem naturalmente desempenham um papel crescente na descoberta e desenvolvimento de medicamentos (BEDNARCZYK-CWYNAR; ZAPRUTKO, 2015).

Produtos naturais têm sido a base da quimioterapia para câncer nos últimos 50 anos. Três quartos dos fármacos atualmente disponíveis são produtos naturais ou relacionados a eles. Mais de 60% dos 140 agentes antitumorais aprovados desde 1940 podem ser relacionados a um produto natural. Em 2000, 57% de todos os fármacos em ensaios clínicos para câncer eram produtos naturais ou seus derivados. Quando comparados aos compostos químicos sintéticos, os produtos naturais apresentam um perfil favorável de absorção e metabolismo pelo organismo com baixa toxicidade. Também é demonstrado que os produtos naturais possuem mecanismos diversos e podem atingir várias vias que estão alteradas nas células cancerosas para alcançar maior eficácia terapêutica (SIVEEN *et al.*, 2017).

Apesar da grande diversidade estrutural dos terpenoides, estes compostos possuem a mesma origem biossintética sendo formados a partir da fusão repetida de unidades de isopreno (C_5H_8), e o número de unidades fundidas determina sua classificação (AGUILAR, 2019). Suas estruturas podem ser lineares (principalmente derivados de esqualeno) ou podem ser compostas por um número de ciclos compreendidos entre um e cinco. Os triterpenos cíclicos são de interesse medicinal devido à ampla gama de atividades biológicas que eles podem exercer, como: antitumoral, anti-inflamatória, antioxidante, antiviral, antibacteriana e antifúngica. Considerando essas atividades, existe a interação dos triterpenos bioativos em diferentes alvos, incluindo fator de necrose tumoral, STAT3, receptor *toll-like* (TLR) e PI3K/Akt/mTOR (PERON *et al.*, 2018).

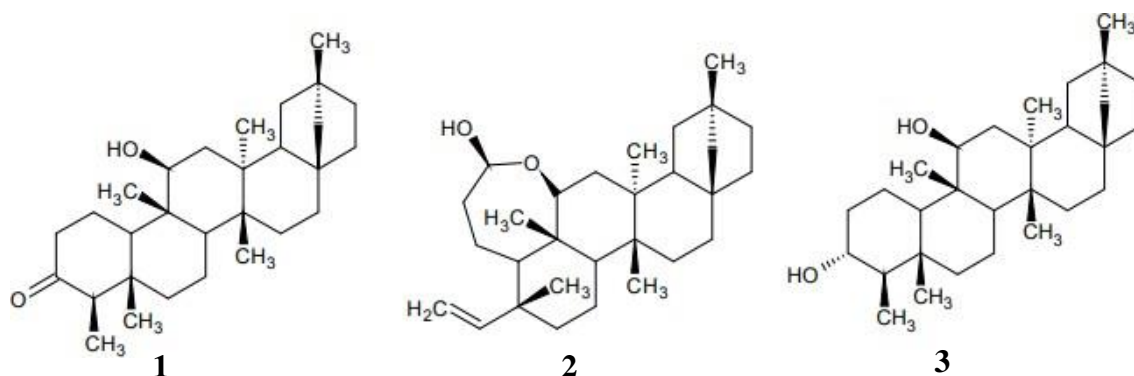
Os triterpenos pentacíclicos não só possuem atividades farmacológicas significativas, mas são isentos de toxicidade proeminente. Portanto, estes triterpenos podem ser considerados como compostos promissores para o desenvolvimento de fármacos. O ácido

ursólico é um exemplo recente de um triterpeno bem conhecido como modelo de composto bioativo a ser usado na descoberta de fármacos (PERON *et al.*, 2018).

A *Maytenus quadrangulata* é uma planta encontrada na caatinga e na mata atlântica brasileira, popularmente conhecida como espinho-de-deus. Espécies do gênero *Maytenus* são utilizadas na medicina popular como agentes antitumorais, analgésicos, antiulcerogênicos, antiasmáticos, anti-inflamatórios, antimicrobianos e para o tratamento de problemas estomacais (AGUILAR, 2019).

Os compostos 1, 2 e 3 (**Figura 3**), que são triterpenos pentacíclicos, foram isolados do extrato hexânico das folhas de *Maytenus quadrangulata*. Os resultados da avaliação da atividade citotóxica desses compostos evidenciaram o potencial dos extratos das folhas de *M. quadrangulata* frente às linhagens leucêmicas THP-1 e K562. Portanto, no contexto de desenvolvimento de novos fármacos e considerando a diversidade de características fenotípicas das células tumorais provenientes das mutações genéticas e alterações epigenéticas presentes nas leucemias se faz necessário o estudo do mecanismo de ação de compostos com potencial antitumoral.

Figura 3 – Estrutura dos triterpenos 1, 2 e 3.



Fonte: Elaborada pela autora.

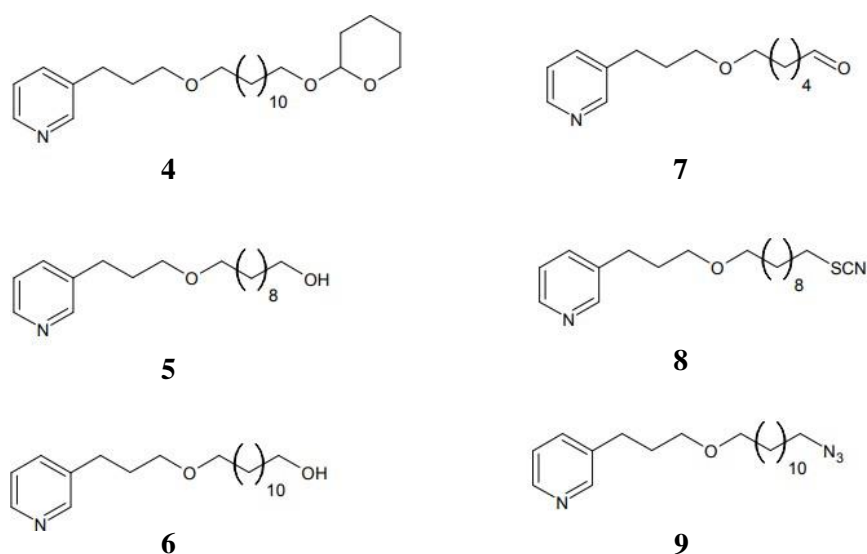
2.6 Alcaloides marinhos

Nos últimos dez anos, muita atenção tem sido dada ao desenvolvimento de fármacos a partir de produtos naturais marinhos, bem como seus análogos sintéticos e semissintéticos (LONG *et al.*, 2018). Os alcaloides marinhos são uma das principais classes de metabólitos secundários isolados de diferentes organismos marinhos tais como cianobactérias, tunicados, ascídios, anêmonas e esponjas (MEHBUB *et al.*, 2014). Esses

compostos possuem diferentes atividades biológicas descritas na literatura tais como anti-inflamatória, antimicrobiana e antimalárica. A incidência de atividade biológica em compostos derivados de fontes marinhas é alta, principalmente no que diz respeito à citotoxicidade, em que extratos de espécies marinhas superam os de origem terrestre (IMPERATORE *et al.*, 2014). Portanto, como os alcaloides marinhos bem como os seus análogos sintéticos destacam-se pela sua marcante citotoxicidade, isso os torna atraentes para o desenvolvimento de novas drogas para o tratamento do câncer.

As espécies de esponjas *Theonella* sp (ordem Lithistida) são uma rica fonte de alcaloides 3-alkilpiridínicos (3-APAs) estruturalmente diversos e biologicamente ativos, como as teoneladinas A-D. Os 3-APAs existem como estruturas abertas, cíclicas ou policíclicas e consistem em um piridínio ou porção tetra-hidropiridina que está conectada a uma cadeia alquil de comprimento variável, geralmente nas posições 1 e / ou 3 (BARBOSA *et al.*, 2018). Vários estudos já confirmaram a atividade citotóxica dos análogos sintéticos de alcaloides marinhos, relatando ações antiangiogênica, anti-proliferativa, inibição da atividade da topoisomerase, polimerização da tubulina, indução de apoptose e citotoxicidade. Os compostos 4 a 9 (**Figura 4**) são análogos sintéticos de alcaloides do tipo 3-APA. Os resultados da avaliação da atividade citotóxica desses compostos evidenciaram seus potenciais frente às linhagens leucêmicas THP-1 e K562, sendo necessários estudos complementares na investigação do mecanismo de ação.

Figura 4 – Estrutura dos análogos alcaloides de 3-alkilpiridina 4 a 9.



Fonte: Elaborada pela autora.

2.7 Determinação *in silico* de alvo molecular (*target fishing*)

Encontrar compostos para um alvo conhecido é a abordagem convencional em programa de química medicinal. No entanto, a descoberta de moléculas bioativas pode ocorrer por triagens biológicas e, em muitos casos, compostos promissores podem ser identificados e seu mecanismo de ação pode e deve ser explorado em sequência. Com a quantidade de dados de atividades biológicas cada vez mais disponíveis, a busca de alvos para estes compostos pode ser auxiliada por métodos computacionais. Por exemplo, a técnica chamada *target fishing* é uma abordagem que visa prever alvos biológicos de moléculas baseado em sua estrutura química (PERAGOVICS *et al.*, 2013).

Estratégias computacionais de *target fishing* são empregadas para identificar os alvos moleculares mais prováveis de uma substância em estudo. Este processo permite a identificação do modo de ação de fármacos conhecidos, reposicionamento de fármacos ou a predição dos efeitos adversos de um composto. A grande quantidade de informações disponíveis sobre a bioatividade de milhares de moléculas permite o desenvolvimento desses tipos de estratégias. Nos últimos anos, temos testemunhado o surgimento de muitos métodos para *target fishing*. Alguns desses métodos baseiam-se no princípio da similaridade molecular, ou seja, moléculas similares podem se ligar aos mesmos alvos e apresentarem bioatividades semelhantes, além do emprego de métodos de acoplamento molecular para a predição do modo de ligação entre as substâncias estudadas e o(s) alvo(s) molecular(es) preditos (CERETO-MASSAGUÉ *et al.*, 2015a).

2.7.1 Análise de similaridade química

O princípio da similaridade química estima que moléculas semelhantes sejam suscetíveis a apresentarem propriedades semelhantes (CERETO-MASSAGUÉ *et al.*, 2015a; ALVES *et al.*, 2018). A análise de moléculas semelhantes é extremamente importante para se estabelecer relações entre estrutura química e atividade biológica ou propriedade de interesse (SAR, do inglês, *Structure-Activity Relationships*) e compreender o comportamento de determinado conjunto de moléculas, além de ser uma ferramenta importante para triagem virtual (ALVES *et al.*, 2018). A estratégia de *target fishing* com base na similaridade é conduzida sugerindo os alvos moleculares para estudo, identificando ligantes que possuem atividade experimental determinada contra determinado alvo e escolhendo o método de similaridade para comparar os ligantes (LAVECCHIA; CERCHIA, 2018). Métodos de

predição de alvo focados em moléculas individuais são essencialmente um exercício de busca por similaridade: de um banco de dados de moléculas bioativas, como o ChEMBL (BENTO *et al.*, 2014) ou o BindingDB (GILSON *et al.*, 2016), estruturas similares a que está sendo estudada são identificadas. As atividades experimentalmente medidas disponíveis no banco de dados são então usadas como uma referência para a molécula em estudo (KOUTSOUKAS *et al.*, 2011). Assim, os alvos de uma molécula podem ser preditos através da identificação de proteínas com ligantes conhecidos que são altamente similares às moléculas de interesse. A vantagem desse método é que ele requer apenas o cálculo de similaridade entre os compostos, o que é considerado computacionalmente barato em relação a outras simulações que consideram a estrutura química tridimensional do alvo molecular (CERETO-MASSAGUÉ *et al.*, 2015a).

A similaridade pode ser calculada com base em representações moleculares 2D e 3D. Os métodos de similaridade 2D dependem de informações deduzidas de gráficos moleculares, geralmente a estrutura em formato SMILES (*Simplified Molecular Input Line Entry Specification*), que é um modo de representação química somente com caracteres do ASCII (*American Standard Code for Information Interchange*). Os métodos baseados em gráfico comparam diretamente as estruturas moleculares umas com as outras e identificam as subestruturas comuns (HENTABLI *et al.*, 2014). Em geral, os descritores moleculares que capturam informações gráficas, como fragmentos ou *fingerprints* moleculares, são muito populares e não exigem grande poder computacional (MAGGIORA *et al.*, 2014).

A informação codificada pelos descritores depende do tipo de representação molecular e o algoritmo definido para seu cálculo. Dentre os descritores conhecidos estão os constitucionais, índices topológicos e geométricos (DANISHUDDIN; KHAN, 2016).

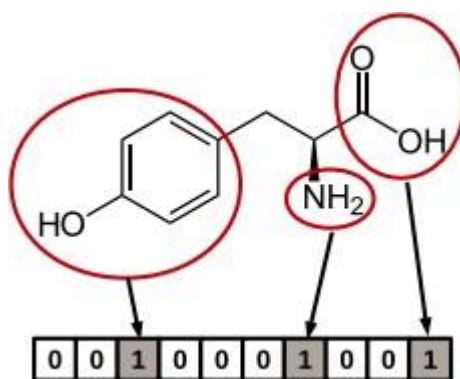
Descritores constitucionais são simples e comumente usados, refletindo a composição molecular de um composto sem qualquer informação sobre sua topologia. São considerados o número de átomos e anéis, contagem de ligações, tipo de átomo e peso molecular. Já os índices topológicos são descritores 2D que consideram o arranjo atômico interno dos compostos. Esses índices codificam informações sobre o tamanho molecular, forma, presença de heteroátomos e ligações múltiplas de forma numérica. Eles têm um significativo papel na identificação de diferentes propriedades físico-químicas, atividades biológicas e propriedades farmacocinéticas (DANISHUDDIN; KHAN, 2016).

Os descritores geométricos são calculados a partir das coordenadas 3D dos átomos na molécula. Esses descritores são ricos em informação e poder de discriminação

entre estruturas químicas semelhantes e conformações de moléculas em comparação com os topológicos (DANISHUDDIN; KHAN, 2016).

Em quimioinformática, *fingerprints* moleculares são as formas mais comuns de representação das estruturas químicas. Representações deste tipo são simplificações da informação química contida em qualquer estrutura química através de vetores binários. Cada posição no vetor indica a ausência (0) ou presença (1) de características pré-determinadas no projeto do *fingerprint* (FERNÁNDEZ-DE GORTARI *et al.*, 2017). Tais representações moleculares podem ser eficientemente comparadas computacionalmente, permitindo assim cálculos de similaridade em larga escala (MAGGIORA *et al.*, 2014).

Figura 5 - Representação hipotética de um *fingerprint* de subestrutura de 10 *bits*.



Fonte: CERETO-MASSAGUÉ *et al.*, 2015b.

Legenda: Três *bits* definidos porque as subestruturas que eles representam estão presentes na molécula (circulado em vermelho).

Neste estudo foram utilizados os *fingerprints* MACCS (DURANT; LELAND; HENRY; NOURSE, 2002) e PubChem (*Pubchem Substructure Fingerprint*, ftp://ftp.ncbi.nlm.nih.gov/pubchem/specifications/pubchem_fingerprints.pdf). O *fingerprint* MACCS, também chamado de chaves estruturais MACCS, é um *fingerprint* baseado em 166 fragmentos estruturais contendo de 1 a 10 átomos não-hidrogênio e é uma das ferramentas mais populares para o cálculo de similaridade. Cada posição de *bit* é atribuída a um fragmento ou chave específica e sua presença ou ausência em um composto é calculada (MAGGIORA *et al.*, 2014). O PubChem *fingerprint* possui 881 chaves estruturais e cobre uma ampla gama de diferentes subestruturas. É o *fingerprint* utilizado pelo site PubChem para pesquisa de similaridade (CERETO-MASSAGUÉ *et al.*, 2015b).

Os *fingerprints* 2D têm sido amplamente utilizados em *target fishing*, pelo fato de serem calculados rapidamente. No entanto, os descritores 3D também podem ser usados,

embora o cálculo seja computacionalmente mais caro. Os métodos de similaridade baseados na forma (ou volume molecular) utilizam comparações das conformações 3D das moléculas, mas outras maneiras podem ser usadas, como a superfície eletrostática. Métodos tridimensionais têm o potencial de detectar similaridades entre moléculas com diferentes estruturas moleculares, mas com volume e disposição de grupos funcionais parecidos. Portanto, a combinação de medidas da similaridade química e de forma, aumenta significativamente a acurácia da predição de alvo (CERETO-MASSAGUÉ *et al.*, 2015a).

Em geral, o coeficiente de similaridade é calculado empregando descritores como *fingerprints* moleculares (para representar a similaridade estrutural entre duas moléculas) ou a própria estrutura tridimensional considerando seu formato e o posicionamento dos grupos funcionais (LIMA *et al.*, 2016).

Dentre as funções de similaridade mais utilizadas, o coeficiente de Tanimoto (também conhecido como similaridade de Jaccard) e as distâncias Euclideana e de Mahalanobis podem ser citados (ALVES *et al.*, 2018). O coeficiente de Tanimoto é geralmente definido pela Equação 1:

$$Tc(A, B) = \frac{c}{(a + b) - c} \quad (1)$$

sendo que a e b são o número de características presentes nos compostos A e B, respectivamente, e c é o número de características compartilhadas por A e B. Portanto, Tc quantifica a fração de características comuns à A e B no número total de características de A e B, e o termo c no denominador corrige a contagem dupla de características. O cálculo do coeficiente de Tanimoto é baseado nas representações chamadas de *fingerprints* moleculares. (MAGGIORA *et al.*, 2014). Quanto mais similares forem dois compostos, mais próximo de 1 será o coeficiente de Tanimoto (CERETO-MASSAGUÉ *et al.*, 2015b).

Neste estudo foram utilizadas estratégias computacionais de *target fishing* por análise de similaridade química 2D e 3D e, após a identificação de moléculas similares com alvos moleculares já conhecidos, foram conduzidos estudos de acoplamento molecular.

2.7.2 Acoplamento Molecular

As estratégias de planejamento de fármaco baseadas na estrutura do alvo molecular (SBDD, do inglês, *structure-based drug design*) se fundamentam nas informações derivadas do conhecimento da estrutura 3D de um alvo de interesse. Desta forma, é possível

estudar moléculas de acordo com a complementaridade estrutural e eletrônica com um determinado alvo. Nesse contexto, o acoplamento molecular é um dos métodos *in silico* baseados na estrutura do alvo mais populares, empregado para prever as interações que ocorrem entre as moléculas e os alvos biológicos (PINZI; RASTELLI, 2019). O processo de acoplamento envolve duas etapas básicas: predição da conformação do ligante bem como sua posição e orientação dentro dos locais de ligação (chamadas de poses) e predição da afinidade de ligação (MENG *et al.*, 2011).

Programas de acoplamento molecular utilizam um algoritmo de busca no qual a conformação do ligante é variada e avaliada recursivamente dentro do sítio de ligação do alvo molecular até que a convergência para a energia mínima seja alcançada (MALTAROLLO, 2013). Cada programa é baseado em um algoritmo de pesquisa específico, tais como o algoritmo incremental, algoritmo genético (AG) e Monte Carlo (MC) (DIAS; AZEVEDO, 2008).

O algoritmo incremental divide o ligante em fragmentos que são encaixados separadamente no local do receptor. Essa fragmentação permite que o algoritmo considere a flexibilidade do ligante. Fragmentos rígidos que são encaixados inicialmente funcionam como âncoras que são unidos secundariamente pelas partes flexíveis do ligante. Desta forma, o ligante é gradualmente construído dentro do local de ligação do receptor. O algoritmo incremental é empregado em vários programas de acoplamento como FLEXX (RAREY *et al.*, 1996) e Surfex (JAIN, 2003) (DIAS; AZEVEDO, 2008).

O AG é um tipo de algoritmo de busca estocástica baseado nos conceitos de evolução Darwiniana. Todos os AGs começam com uma população de soluções propostas para um determinado problema (PEGG *et al.*, 2001). Nos programas de acoplamento molecular, as poses geradas são avaliadas em quão bem interagem com o sítio de ligação sendo criadas novas populações, a partir da população inicial, com as poses mais favoráveis, ou seja, de menores energias. O ciclo do algoritmo continua por gerações de avaliação até que soluções suficientemente boas forem encontradas ou até um valor definido pelo usuário.

Os programas de acoplamento utilizam diferentes métodos e parâmetros para o cálculo da energia de interação proteína-ligante, tais como o emprego das funções de pontuação (*score*), a geração de conformações dos ligantes e a geração de diferentes conformações da proteína ou de alguns aminoácidos presentes no sítio de ligação. A partir de uma conformação de um ligante gerado no sítio de ligação de seu alvo, as funções de *score* calculam a energia de interação e, assim, podem ranquear as poses (conformações) geradas com o intuito de reproduzir as conformações de mais baixa energia (ou maior afinidade) ou

identificar possíveis candidatos a ligante de um determinado alvo (MALTAROLLO, 2013). Os principais programas de acoplamento que podem ser citados são AutoDock (FORLI *et al.*, 2016), DOCK (ALLEN *et al.*, 2015), Glide (HALGREN *et al.*, 2004) e GOLD (JONES *et al.*, 1997).

O acoplamento molecular pode ser classificado em rígido, semi-flexível e flexível. No acoplamento rígido, a estrutura do alvo molecular não se altera em termos conformacionais. O método de cálculo é relativamente simples e estuda principalmente o grau de correspondência da conformação. No acoplamento semi-flexível, a conformação de alguns resíduos de aminoácidos pode variar dentro de um intervalo específico, portanto ele é mais utilizado nas interações entre proteínas e pequenas moléculas. Em geral, a estrutura das pequenas moléculas pode ser alterada livremente, enquanto as macromoléculas permanecem rígidas ou retêm alguns dos resíduos de aminoácidos rotativos para garantir a eficiência computacional. No acoplamento flexível, a conformação simulada de ambas as moléculas (ligante e alvo molecular) é livre para alterações, consumindo mais recursos de computação e melhorando a precisão (LIN *et al.*, 2020).

3 JUSTIFICATIVA

A busca por novos antitumorais que possam substituir ou melhorar a quimioterapia atualmente disponível para tratamento da LMA e LMC é uma necessidade na pesquisa, visto que as terapias existentes para tratamento ainda não atendem a todos os grupos de pacientes. Em relação à LMC, mesmo com o desenvolvimento de inibidores de tirosina quinase de segunda e terceira gerações, uma parte dos pacientes não responde à terapia e desenvolvem resistência ao longo do tratamento. Em relação à LMA, a heterogeneidade genômica característica da doença dificulta a definição de protocolos de tratamento.

Considerando o exposto acima e os resultados preliminares de atividade citotóxica dos compostos 1 a 9, estes foram selecionados como candidatos promissores à atividade anti-leucêmica. Nesse contexto, o uso de ferramentas computacionais como o *target fishing* e acoplamento molecular são de grande valor na determinação de alvos moleculares e estudo de mecanismo de ação a um custo reduzido. Em contraste com a triagem virtual, que é usada para pesquisar grandes bibliotecas de compostos para moléculas com maior probabilidade de se ligar a um alvo específico, o objetivo da triagem reversa, também conhecido como *target fishing*, é identificar os alvos mais prováveis de uma molécula de interesse (CERETO-MASSAGUÉ *et al.*, 2015a). A abordagem de acoplamento molecular pode ser usada para

modelar a interação entre uma pequena molécula e uma proteína no nível atômico, o que nos permite caracterizar o comportamento dessas moléculas no local de ligação de proteínas alvo (MENG *et al.*, 2011). Dessa forma, as técnicas citadas de estudo *in silico* foram aplicadas para a elucidação do mecanismo de ação dos compostos de interesse.

4 OBJETIVOS

4.1 Objetivo Geral

Predizer *in silico* o(s) potencial(is) alvo(s) molecular(es) de moléculas citotóxicas em linhagens celulares leucêmicas humanas (THP-1 e K562).

4.2 Objetivos específicos

4.2.1 Procurar na base de dados ChEMBL substâncias com atividade biológica determinada experimentalmente contra as linhagens leucêmicas THP-1 e K562;

4.2.2 Calcular similaridade estrutural 2D entre as moléculas de interesse e as moléculas ativas obtidas na literatura;

4.2.3 Fazer a predição de potenciais alvos proteicos para as substâncias ativas utilizando servidores disponíveis;

4.2.4 Realizar cálculos de acoplamento molecular com as moléculas e os alvos preditos.

5 METODOLOGIA

5.1 Identificação de alvos moleculares por similaridade química

Com o intuito de avaliar potenciais alvos das moléculas previamente selecionadas nos testes *in vitro*, inicialmente, os dados dos compostos com ação citotóxica frente às linhagens celulares THP-1 e K562 foram obtidos no banco de dados ChEMBL v24 (BENTO *et al.*, 2014). Foram selecionados os compostos com concentração de inibição do crescimento de 50% das células (CI₅₀) e GI₅₀ (do inglês, *Growth inhibitory dose 50%*) definidos. Os quatro bancos de dados obtidos foram filtrados para remover entradas sem atividade experimental determinada, moléculas inativas, ou seja, com atividade superior a um dado valor, e misturas

de compostos. Todas as substâncias com CI_{50} e GI_{50} determinados experimentalmente foram consideradas. As estruturas em formato SMILES contidas nas planilhas foram convertidas para o formato de .mol2 e, em seguida, as conformações de menores energias dos compostos testados com atividade citotóxica foram calculadas pelo programa Open Babel 2.4.0.

Para todos os compostos, foram calculados os estados de ionização de acordo com o $pH = 7,4$ usando o programa Open Babel 2.4.0 e, em seguida, as conformações de menores energias foram geradas novamente usando o programa Open Babel 2.4.0 para corrigir possíveis influências de grupos ionizados na conformação.

O programa PaDEL *descriptors* (YAP, 2011) foi utilizado para o cálculo dos *fingerprints* MACCS e PubChem dos compostos de interesse e das moléculas provenientes do banco de dados ChEMBL. O programa EXCEL foi utilizado para o cálculo do coeficiente de Tanimoto baseado nos *fingerprints* gerados pelo PaDEL. As moléculas com Tc maior ou igual a 0,75 foram consideradas para a pesquisa bibliográfica de alvos moleculares ou mecanismos de ação definidos.

Além do mais, foram utilizados os seguintes servidores para predição de alvo molecular:

1) ChemMapper Server (GONG *et al.*, 2013) é um servidor que permite realizar cálculos de similaridade molecular 3D utilizando o algoritmo SHAFTS, um método que combina a sobreposição de moléculas usando a forma molecular e a correspondência de grupos químicos. A escala de similaridade entre os compostos é calculada e dimensionada entre 0 e 2. Quanto mais próximo de 2,0 for a pontuação, maior o potencial de associação farmacológica entre as moléculas. É possível selecionar o valor mínimo de similaridade, sendo 1,2 o padrão, e o banco de dados em que será realizada a pesquisa. Neste estudo foi utilizado 1,5 como valor mínimo (<http://www.lilab-ecust.cn/chemmapper/help.html>).

2) Hitpick (MALO *et al.*, 2006; SCHUFFENHAUER *et al.*, 2003; NIDHI *et al.*, 2006) é um servidor que aplica a combinação de dois métodos baseados em similaridade 2D, 1-*Nearest-Neighbour* (1NN) e um método de aprendizado de máquina baseado em modelos Bayesianos (<http://mips.helmholtz-muenchen.de/hitpick/cgi-bin/index.cgi?content=help.html>).

3) SuperPred (NICKEL *et al.*, 2014) é um servidor que calcula a similaridade 2D, por fragmentos e 3D. A predição de alvo é baseada na distribuição de similaridade entre os ligantes dos alvos conhecidos. Seus modelos foram construídos utilizando um banco de dados contendo cerca de 341.000 compostos, 1.800 alvos e 665.000 interações composto-alvo (<http://prediction.charite.de/>).

4) Swiss Target Prediction (DAINA; MICHIELIN; ZOETE, 2019) é um servidor

que realiza predições calculando moléculas similares bi e tridimensionalmente, em uma coleção com mais de 376.342 compostos, conhecidos por serem experimentalmente ativos em um conjunto de 3.068 alvos macromoleculares (<http://www.swisstargetprediction.ch/>).

5.2 Predição do modo de ligação utilizando acoplamento molecular

Dentro deste conjunto de dados quimicamente similares, compostos com atividade experimental e com alvos moleculares previamente identificados foram considerados os nossos potenciais alvos moleculares e utilizados em estudos de acoplamento molecular. Esses alvos foram recuperados do *Protein Data Bank* (PDB). As estruturas cristalográficas das proteínas resultantes da busca no PDB foram selecionadas baseadas na presença de ligantes, ausência de mutações no sítio de ligação, resolução, R_{free} e elucidadas pelo método experimental de difração por raios-X. Para estas estruturas o critério mais aceito para avaliar os dados experimentais é a resolução do cristal. Estruturas de alta resolução têm resolução abaixo de 1,8 Å, e as acima de 2,7 Å são consideradas de baixa resolução. Os intermediários são classificados como de média resolução. (BAGARIA *et al.*, 2013). Neste trabalho foram consideradas estruturas de alvo molecular com resolução média e alta, com o *cut-off* de 2,0 Å.

No ANEXO II estão descritas todas as proteínas disponíveis no PDB para cada alvo proteico estudado, conforme os critérios de seleção definidos. O desvio quadrático médio (RMSD) foi calculado entre todas as estruturas cristalográficas do PDB e os alvos selecionados para o acoplamento molecular, utilizando o programa PyMOL versão 0.99c (DELANO SCIENTIFIC, 2002).

Além da resolução, o fator R_{free} também foi considerado na escolha da proteína no PDB. O fator R é a medida da qualidade do modelo atômico obtido a partir dos dados cristalográficos, mede o quão bem o padrão de difração simulado corresponde ao padrão de difração observado experimentalmente. Um conjunto totalmente aleatório de átomos dará um fator R de cerca de 0,63, enquanto um ajuste perfeito teria um valor de 0. Os valores típicos são cerca de 0,20. Porém, existe um problema potencial com o uso de valores R para avaliar a qualidade de uma estrutura. O processo de refinamento é frequentemente usado para melhorar o modelo atômico de uma determinada estrutura para fazê-lo se ajustar melhor aos dados experimentais e melhorar o fator R. Antes do refinamento começa, cerca de 10% das observações experimentais são removidas do conjunto de dados. Em seguida, o refinamento é executado usando os 90% restantes. O valor R_{free} é então calculado ao ver quão bem o modelo prevê os 10% que não foram usados no refinamento e minimiza o viés do refinamento. O R_{free}

é um pouco mais alto que o fator R, com um valor adequado abaixo de 0,26 (KLEYWEGT; JONES, 1997).

As proteínas a serem identificadas foram preparadas adicionando os átomos de H, corrigindo estados de protonação dos aminoácidos utilizando o programa GOLD. Entre os resíduos com dupla ocupância foram considerados os que tinham maior probabilidade. A simulação de acoplamento molecular foi realizada em torno do ligante co-cristalizado da proteína usando as configurações padrão do programa GOLD (VERDONK *et al.*, 2003). Os resíduos de aminoácidos foram considerados rígidos e moléculas de água estruturais foram mantidas durante o cálculo.

As águas estruturais foram mantidas durante a análise de acoplamento molecular, sendo consideradas as moléculas de água que interagem por meio de ligações de hidrogênio com o ligante co-cristalizado. Para definição das águas estruturais, foram analisadas até dez estruturas cristalográficas de melhores resoluções disponíveis no PDB para cada alvo molecular, utilizando o programa PyMOL versão 0.99c (DELANO SCIENTIFIC, 2002). As moléculas de água em posições comuns a maioria das proteínas e descritas na literatura foram mantidas para o cálculo de acoplamento molecular.

O protocolo de acoplamento foi validado com simulações de *redocking*, em que o ligante co-cristalizado foi submetido às mesmas condições de acoplamento molecular dos compostos em estudo, utilizando o alvo proteico obtido no PDB (VALLONE *et al.*, 2018; SERAFIM *et al.*, 2019). O acoplamento foi realizado considerando o sítio de ligação como os resíduos que estão a um raio de 10 Å do ligante cristalográfico. No entanto, nas análises das interações do ligante com o alvo molecular foram empregados o *cut-off* de 3,5 Å para as ligações de hidrogênio e de 5,0 Å para as interações hidrofóbicas.

No total, foram geradas 200 poses para cada composto que posteriormente foram agrupadas de acordo com o desvio quadrático médio (RMSD) considerando conformações com desvios de posição atômica menores que o *cut-off* de 1,5 Å para ser incluído no mesmo grupo (TONHOLO *et al.*, 2020). A afinidade pelo alvo molecular foi calculada utilizando a função de pontuação GOLD *score* (JONES *et al.*, 1997) sendo então as poses geradas ranqueadas de acordo com o maior valor calculado. Os valores de *score* foram normalizados pela divisão do valor de *score* de cada pose pela massa molecular (MM) do composto analisado. As funções empíricas de *score* geralmente exibem uma dependência do peso molecular, resultando em uma tendência de moléculas maiores ficarem no topo do ranque de classificação, gerando resultados falsos positivos. Portanto, a normalização pela MM compensa parcialmente essa tendência em pontuar mais as moléculas maiores (CARTA *et al.*,

2007; JACOBSSON; KARLÉN, 2006). Além do mais, os dois grupos mais populosos foram analisados em relação ao modo de ligação dos ligantes com os respectivos alvos utilizando o programa PyMOL versão 0.99c (DELANO SCIENTIFIC, 2002). No *redocking* foi avaliado modo de ligação da pose mais bem ranqueada de cada ligante selecionado no PDB.

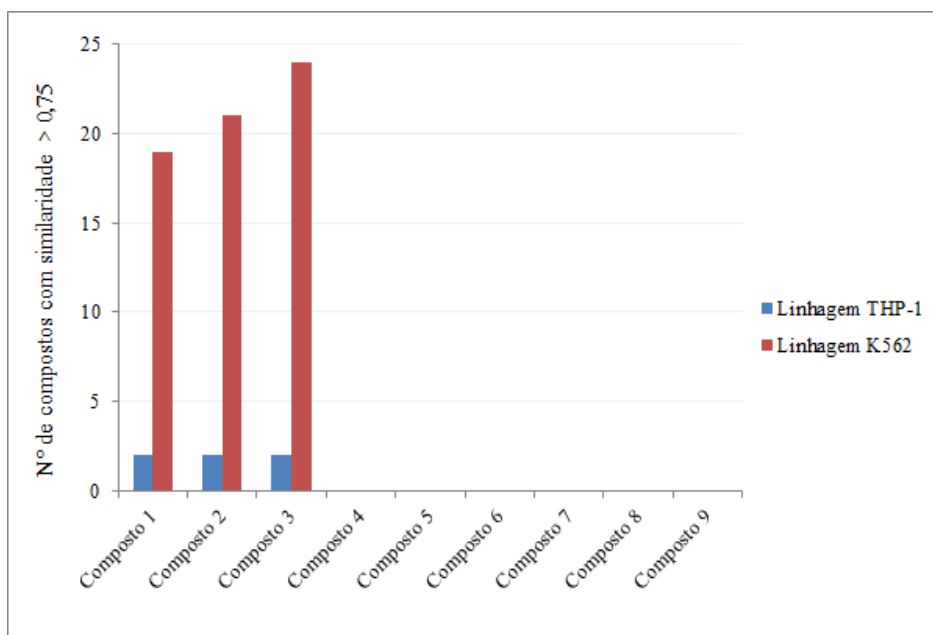
6 RESULTADOS E DISCUSSÃO

6.1 Análise de similaridade química

Utilizando como busca no banco de dados ChEMBL as linhagens celulares THP-1 e K562, foram selecionados os compostos com CI_{50} e GI_{50} definidos. Inicialmente, foram encontrados no ChEMBL 2.868 compostos reportados contra a linhagem THP-1 e 53.017 compostos com atividade contra a linhagem celular K562. Após a remoção de moléculas inativas, entradas sem atividade reportada e misturas, 14.785 moléculas foram selecionadas para os cálculos de similaridade.

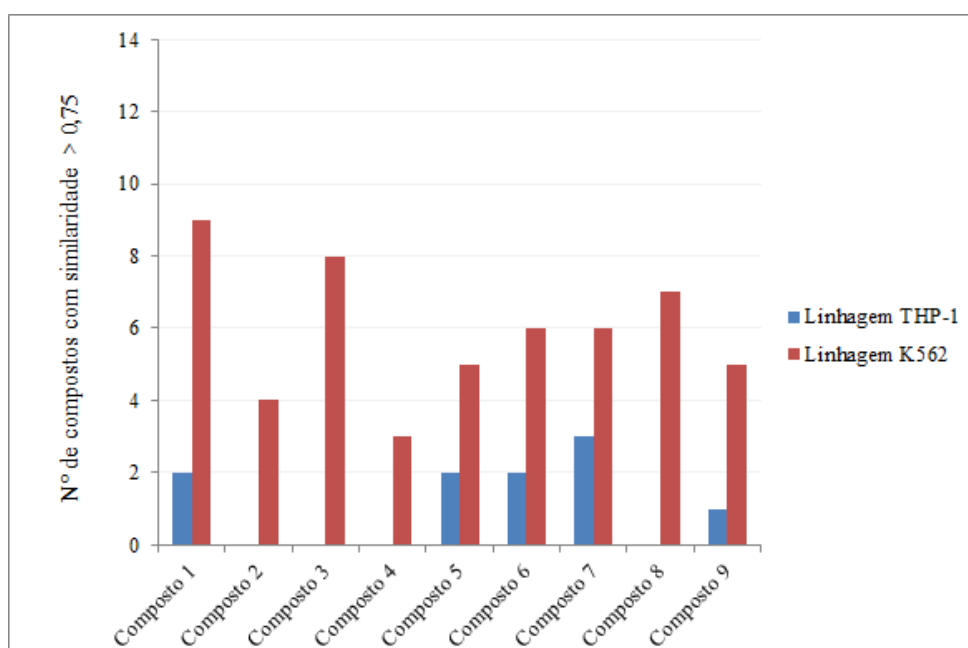
A similaridade estrutural foi calculada entre todas as substâncias de interesse (moléculas 1 a 9) e os compostos provenientes da pesquisa no ChEMBL. O coeficiente de Tanimoto foi calculado usando os *fingerprints* MACCS e PubChem. Os compostos com valores de Tc maiores ou igual a 0,75 estão representados no **ANEXO I**. Os gráficos 1 e 2 representam o número de compostos com similaridade $\geq 0,75$ em relação aos compostos de interesse utilizando os *fingerprints* MACCS e PubChem, respectivamente.

Gráfico 1 - Número de compostos com similaridade $\geq 0,75$ em relação aos compostos de interesse utilizando o *fingerprint* MACCS.



Fonte: Elaborado pela autora.

Gráfico 2 - Número de compostos com similaridade $\geq 0,75$ em relação aos compostos de interesse utilizando o *fingerprint* PubChem.

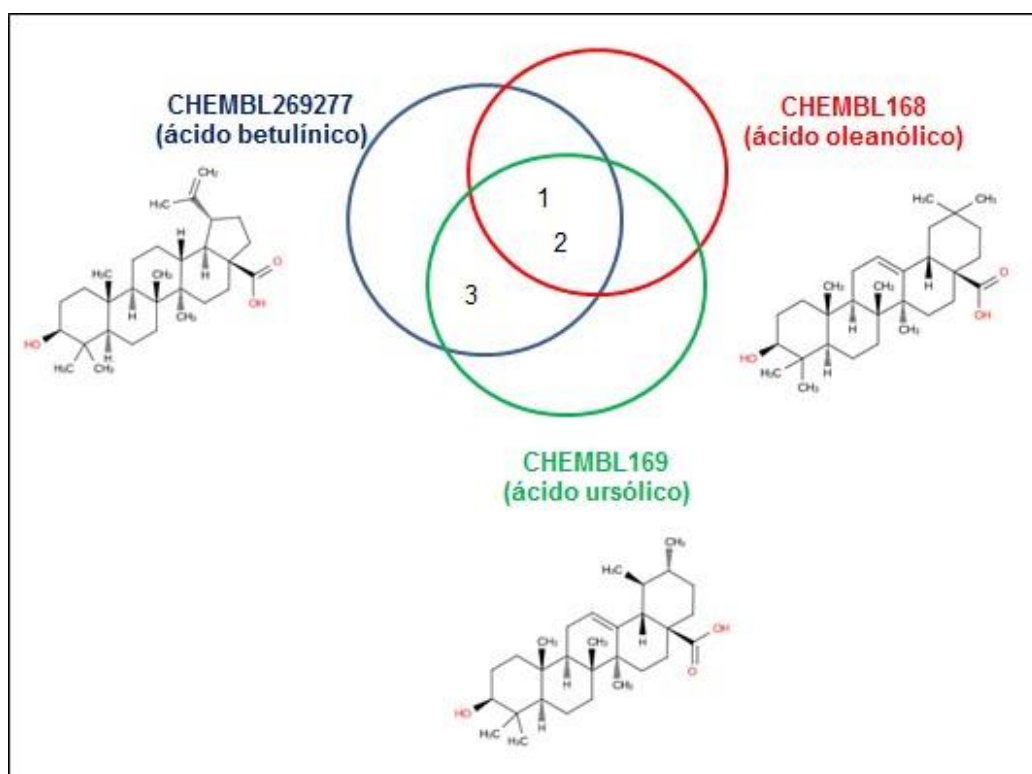


Fonte: Elaborado pela autora.

A partir dos cálculos de similaridade estrutural entre as moléculas de interesse e as moléculas ativas já reportadas na literatura provenientes do ChEMBL, 65 moléculas foram

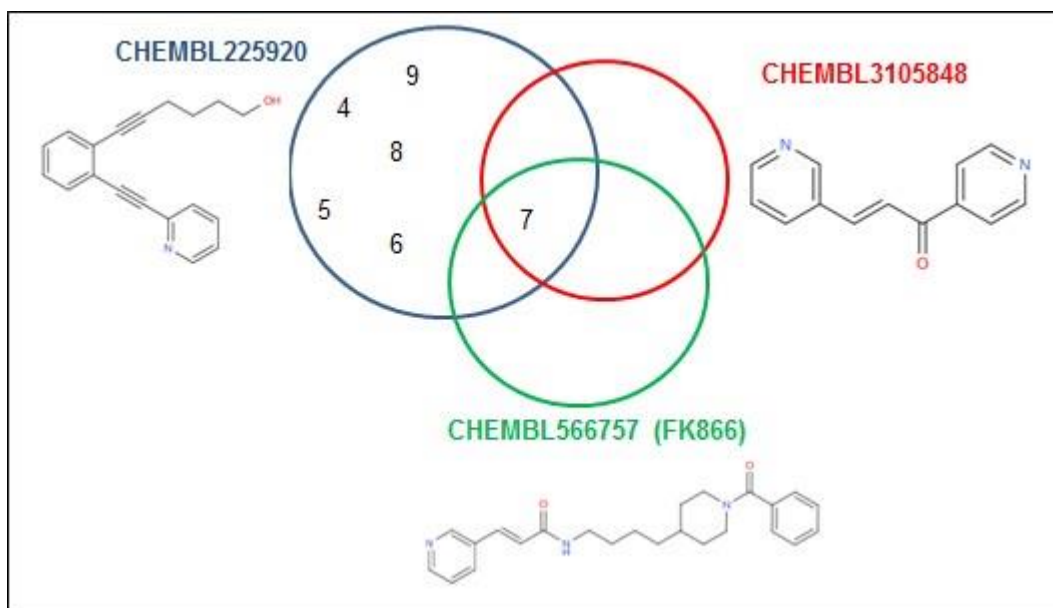
consideradas similares ao nosso banco de dados interno. Entretanto, com o intuito de direcionar a predição de alvo molecular por similaridade química, foi realizada a pesquisa bibliográfica de estudos de mecanismos de ação para todos os compostos com $T_c \geq 0,75$ representados no **ANEXO I**. Nas **figuras 6 e 7** estão representados os compostos que possuem estudos publicados e com quais moléculas de interesse eles possuem $T_c \geq 0,75$.

Figura 6 - Compostos do ChEMBL que possuem estudos de mecanismo de citotoxicidade publicados e $T_c \geq 0,75$ em relação aos triterpenos pentacíclicos.



Fonte: Elaborada pela autora

Figura 7 - Compostos do ChEMBL que possuem estudos de mecanismo de citotoxicidade publicados e $T_c \geq 0,75$ em relação aos análogos alcaloides de 3-alkilpiridina.



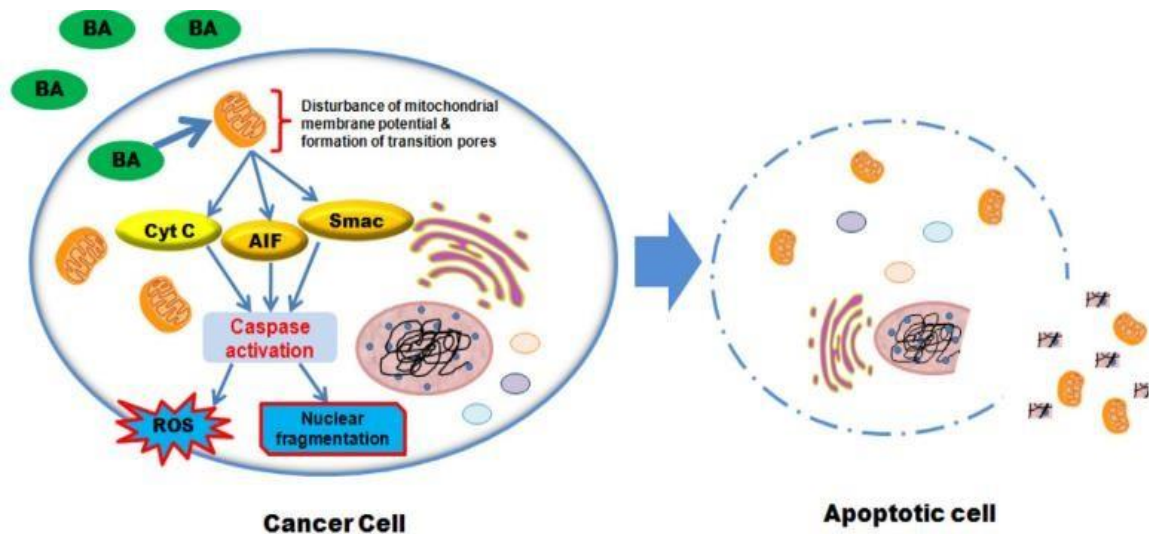
Fonte: Elaborada pela autora.

Dos compostos que possuem mecanismo de ação estudado, três são comuns aos triterpenos pentacíclicos de interesse. A molécula CHEMBL269277 (ácido betulínico) é um triterpeno pentacíclico que possui diferentes atividades biológicas incluindo atividade antitumoral. A atividade antitumoral do ácido betulínico tem sido bem documentada devido a sua capacidade de desencadear a apoptose em diversos tipos de células tumorais. Evidências mostram que o ácido betulínico induz apoptose em células cancerosas através da via intrínseca (mitocondrial) e não pela via extrínseca (KUMAR *et al.*, 2018).

A via intrínseca da apoptose está relacionada à permeabilização da membrana externa mitocondrial (MOMP), que é regulada por vários fatores incluindo proteínas da família Bcl-2 pró e antiapoptóticas, espécies reativas de oxigênio (ROS), proteínas do complexo de poros de transição da permeabilidade mitocondrial, íons cálcio e esfingolipídeos. A MOMP resulta na liberação de proteínas solúveis, incluindo citocromo c, Smac, fator indutor de apoptose (FIA) e endonuclease G. Citocromo c e Smac promovem a ativação de caspases neutralizando as "proteínas inibidoras de apoptose" - mediado por inibição de caspase -3 e -9, enquanto FIA e endonuclease G contribuem para a cromatólise independente de caspases (**Figura 8**). As proteínas pró-apoptóticas da família Bcl-2 (Bax, Bak, Bad) estão envolvidas na MOMP para liberar citocromo c no citosol, que se liga no fator-1 de ativação de protease apoptótica (APAF-1) e pró-caspase-9 para liberar a apoptossoma. A apoptossoma

produz a caspase-9 ativa, que ativa caspase-3, -6 e -7 e leva à fragmentação do DNA e apoptose. Além da ativação de caspases, o ácido betulínico também estimula o aumento na produção de ROS que é considerado um fator de estresse envolvido na iniciação da MOMP (KUMAR *et al.*, 2018).

Figura 8 - Mecanismo de ação do ácido betulínico na indução da apoptose.



Fonte: KUMAR *et al.*, 2018.

Legenda: O ácido betulínico (BA) promove alteração do potencial da membrana mitocondrial e consequente formação de poros, acarretando na liberação no citoplasma de proteínas solúveis como: citocromo c (cyt C), fator indutor de apoptose (FIA) e Smac. Cyt c e Smac promovem a ativação de caspases neutralizando as "proteínas inibidoras de apoptose", enquanto o FIA contribui para a cromatólise independente de caspases. O ácido betulínico também estimula o aumento na produção de ROS, promovendo a apoptose.

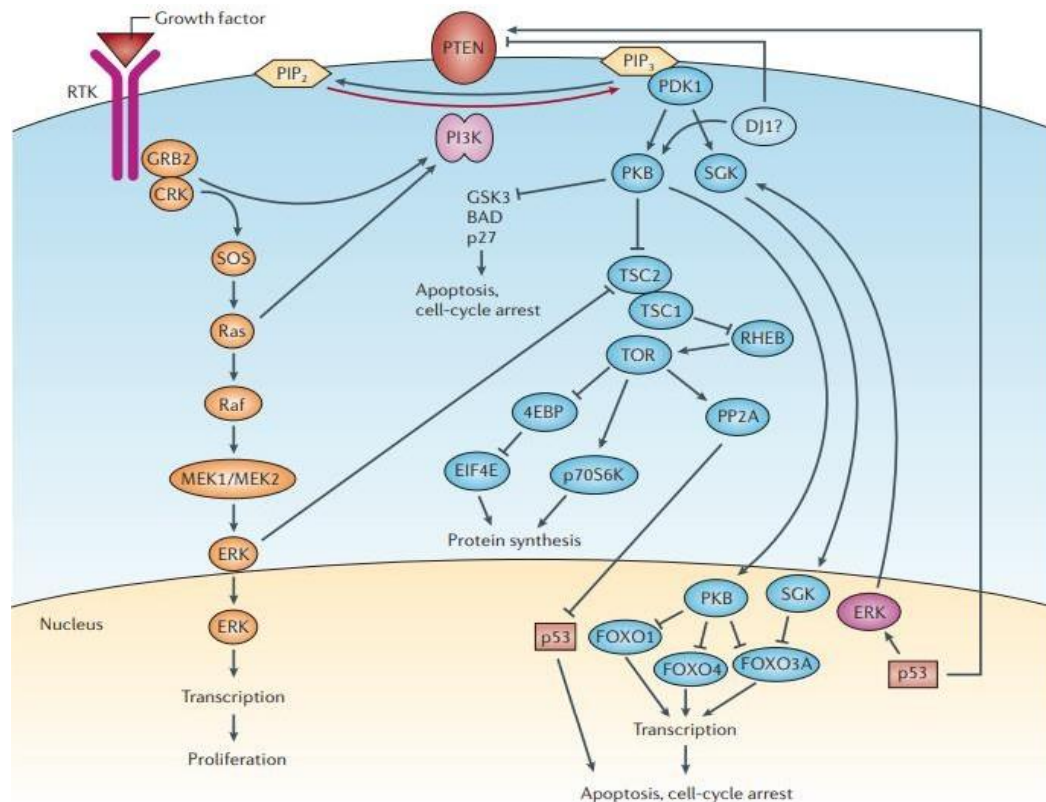
Os efeitos anticancerígenos da molécula CHEMBL168 (ácido oleanólico) têm sido avaliados em diversos tipos de câncer, incluindo a leucemia. O ácido oleanólico induz apoptose por via dependente da mitocôndria, alterando o potencial da membrana mitocondrial, liberando ativadores de caspase, como citocromo c, no citoplasma, levando à fragmentação do DNA nas células cancerígenas humanas. Os efeitos podem ser mediados pela alteração dos níveis de expressão dos genes pró e antiapoptóticos da família Bcl-2, evidenciado pela diminuição da expressão de genes Bcl-2 anti-apoptóticos e aumento da expressão de genes pró-apoptóticos Bax. O ácido oleanólico e seus derivados atuam como inibidores da fosfatidilinositol 3-quinase/proteína quinase B/alvo de mamíferos de rapamicina/ fator nuclear kappa B (PI3K/Akt/mTOR/FN-κB) de maneira dependente da dose. O ácido oleanólico reduz a fosforilação de PI3K e Akt, que são as moléculas que ativam as proteínas mTOR e NF-κB. Porém, não ocorre alteração significativa nos níveis de expressão

total de PI3K ou Akt (ŽIBERNA *et al.*, 2017). NF- κ B é o principal regulador da transcrição genética e sua regulação positiva em linhagens celulares resistentes à quimioterapia está relacionada à falha terapêutica. O desenvolvimento de resistência a múltiplas drogas (MDR) está relacionado à via PI3K/Akt parcialmente devido à ativação de NF- κ B (LIU *et al.*, 2020). Esses efeitos são provavelmente mediados por ROS, pois o tratamento das células com antioxidantes, como a n-acetilcisteína, reverteu a ação inibitória do ácido oleanólico. A ativação da via dependente de proteína quinase ativada por AMP (AMPK) pelo ácido oleanólico também induz alterações metabólicas nas células cancerígenas, como a supressão da lipogênese, síntese de proteínas, e glicólise aeróbia, tendo assim atividade supressora tumoral (ŽIBERNA *et al.*, 2017). Pan *et al.* (2015) demonstrou que derivados do ácido oleanólico inibiram a proliferação celular e induziram apoptose em células da linhagem K562. Os resultados mostraram que o ciclo celular foi interrompido na fase G1 evitando a entrada das células na fase S. A indução da apoptose foi associada à inibição da expressão de Akt1 na via PI3K/Akt. A Akt/PKB, uma serina treonina quinase que é uma molécula crítica na promoção do crescimento celular e vias de sobrevivência, é frequentemente desregulada em muitos cânceres e diabetes.

A molécula CHEMBL169 (ácido ursólico) é um triterpeno pentacíclico natural presente em diversos alimentos como azeitonas, cranberries, maçãs, peras e ameixas. Este composto natural tem recebido atenção crescente nos últimos anos devido ao seu efeito antitumoral, exibindo atividades antiproliferativas e apoptóticas em várias células de carcinoma humano. Assim como para o ácido betulínico e ácido oleanólico, vários mecanismos moleculares estão envolvidos na parada do ciclo celular e apoptose induzida por ácido ursólico, como inibição da replicação do DNA, aumento dos níveis intracelulares de ROS, supressão da ativação do NF- κ B bem como regulação negativa de proteínas anti-apoptóticas (por exemplo, Bcl-2, Bcl-xl e survivina) e ativação da caspase. Além disso, a indução da diferenciação celular, inibição da angiogênese e invasão de células tumorais também foram atribuídas ao ácido ursólico. Recentemente se descobriu que a inativação da via de sinalização PI3K/Akt está envolvida em alguns de seus efeitos antitumorais descritos acima. Uma vez ativado pela PI3K, a Akt se move da membrana celular para o citoplasma e/ou núcleo, onde promove a progressão do ciclo celular e evita a apoptose de várias maneiras, controlando assim as proteínas reguladoras celulares essenciais (por exemplo, ciclina D, c-Myc ou β -catenina) e inibidores da quinase dependente de ciclina (CDK) (por exemplo, p21/WAF1) (VILLAR *et al.*, 2016). No estudo de Wu *et al.* (2012) a caspase-3 e a caspase-9 foram ativadas nas células K562 tratadas com ácido ursólico, além do aumento de

citocromo c no citoplasma, confirmando a indução da apoptose mitocondrial nas células da linhagem K562. Os resultados demonstraram também que o ácido ursólico tem efeitos na inibição do crescimento e indução da apoptose nas células K562 *in vitro* via regulação positiva da fosfatase com homologia de tensina (PTEN) (Figura 9), um regulador negativo da via PI3K/Akt (WU *et al.*, 2012).

Figura 9 - Via de sinalização PI3K/PTEN.



Fonte: CULLY *et al.*, 2006.

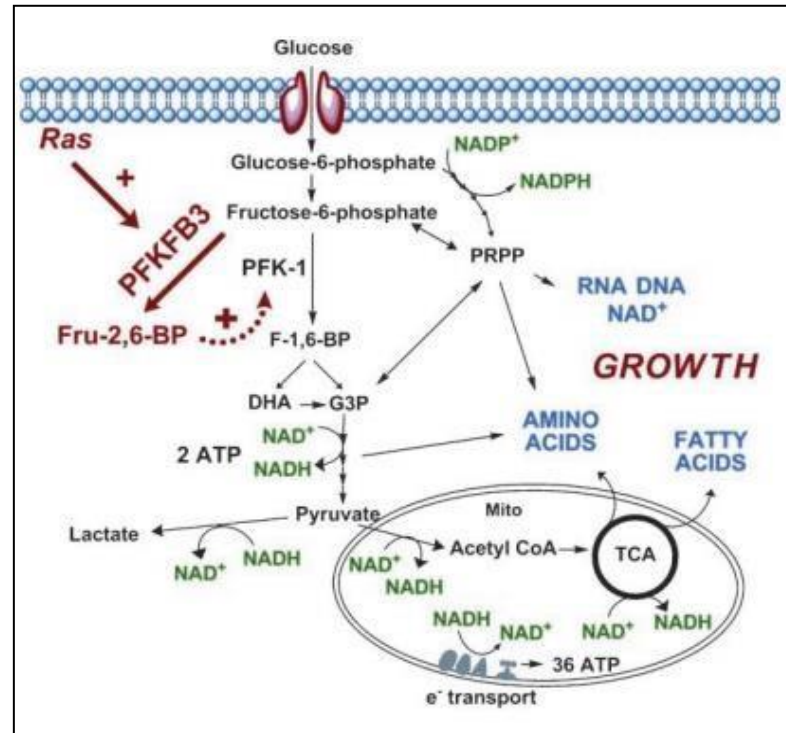
Legenda: A via de sinalização central da fosfatidilinositol 3-quinase (PI3K) (indicada pelos símbolos azuis) começa com a ativação da PI3K (indicada pelo símbolo rosa) pelo receptor tirosina quinase (RTK). A PI3K fosforila e converte o segundo mensageiro lipídico fosfatidilinositol (4,5) bifosfato (PIP₂) em fosfatidilinositol (3,4,5) trifosfato (PIP₃), que recruta e ativa a quinase 1 dependente de fosfatidilinositol (PDK1). A PDK1 por sua vez fosforila e ativa a proteína quinase B (PKB, também conhecida como Akt), que inibe as atividades dos fatores de transcrição *forkhead* (FOXO) (que são mediadores da apoptose e parada do ciclo celular), resultando em proliferação e sobrevivência celular. A proteína supressora de tumor fosfatase com homologia de tensina (PTEN) regula negativamente a sinalização de PI3K desfosforilando PIP₃, convertendo-a de volta em PIP₂. A via de sinalização Ras (símbolos laranja) pode ser acionada por um conjunto de RTKs que são ativados por fatores de crescimento. Ras pode ativar PI3K direta e indiretamente. O status de ativação de p53 também pode afetar o resultado da sinalização de PI3K ao interagir com os fatores de transcrição FOXO regulados por PKB e com quinase 1 extracelular-regulada (ERK1) e ERK2. Outros membros da via de sinalização PI3K incluem SGK (quinase induzida por soro e glicocorticoide), TSC1/TSC2 (esclerose tuberosa 1 e 2), RHEB (homólogo Ras enriquecido no cérebro), TOR (alvo de rapamicina), 4EBP ((iniciação eucariótica proteína de ligação ao fator 4E (EIF4E)), p70S6K (proteína ribossomal, S6 quinase 70kD) e PP2A (proteína fosfatase 2A). Os membros da via de sinalização RTK-Ras incluem GRB2 (proteína ligada ao receptor do fator de crescimento 2), SOS, Ras, Raf, MEK (ERK quinase ativada por mitogênio) e ERK - a ativação desta via leva à proliferação celular.

Em relação aos análogos de alcaloides de 3-alkilpiridina, três moléculas do ChEMBL com mecanismo de ação estudado são comuns às moléculas de interesse. O composto ChEMBL225920 é um análogo de enedinas e os resultados do estudo de Lo *et al.* (2007) sugerem que estes compostos induzem a parada do ciclo celular na fase G2/M e induzem a apoptose por ativação de caspases.

O estudo de Clem *et al.* (2008) sugere que o composto ChEMBL3105848 é um inibidor da PFKFB3, que reduz o fluxo glicolítico resultando na inibição do crescimento celular, é seletivamente citostático para células transformadas e inibe o crescimento tumorigênico *in vivo*. É conhecido que, mesmo na presença de oxigênio, muitas células cancerígenas geram mais ATP a partir da glicose principalmente via glicólise, do que através do processo mais eficiente de fosforilação/oxidação mitocondrial (respiração). Esta mudança do metabolismo energético nas células cancerígenas para esse processo de "glicólise aeróbica" (células normais favorecem a glicólise apenas sob condições anaeróbicas) é conhecido como o Efeito Warburg (BROOKE *et al.*, 2014).

A enzima limitante do processo de conversão de glicose em piruvato, o precursor da produção anaeróbica de ATP, é a fosfofrutoquinase-1 (PFK-1), que converte a frutose-6-fosfato (F6P) em frutose-1,6-bisfosfato (F1,6-BP). A atividade desta enzima é regulada negativamente por ATP, mas essa regulação pode ser alterada pelo ativador alostérico frutose-2,6-bisfosfato (F2,6-BP), que é sintetizado a partir da F6P pela família de enzimas bifuncionais 6-fosfofruto-2-quinase/2,6-bisfosfatase (PFKFB). A alta atividade de PFK-1 causa aumento do fluxo glicolítico, permitindo o aumento da produção de macromoléculas (RNA, DNA, aminoácidos, ácidos graxos etc.) e energia (NADH e ATP) necessárias para a proliferação celular. Dentre as quatro isoformas dessa família, PFKFB3 é a mais importante nesse contexto, com a proporção quinase/fosfatase de 740:1 sendo a mais alta. A PFKFB3 é superexpressa em vários tipos de câncer como de cólon, próstata, pancreático, tireoide e leucemia. A atividade de PFKFB3 também é regulada pela fosforilação de proteínas, que ocorre por várias vias de sinalização como, AMPK, PKA, PKC and PI3K/Akt (BROOKE *et al.*, 2014). Na **figura 10** está representado o mecanismo que promove o crescimento celular na presença das enzimas citadas anteriormente.

Figura 10 - Mecanismo de proliferação celular na presença de PFKFB3.



Fonte: CLEM *et al.*, 2008.

Legenda: PFKFB3 produz F2,6-BP, um potente estimulador da glicólise. O oncogene Ras aumenta a expressão e atividade de PFKFB3, que produz F2,6-BP, um ativador alostérico de PFK-1. Alta atividade de PFK-1 promove aumento no fluxo glicolítico, elevando a produção de macromoléculas (RNA, DNA, aminoácidos, ácidos graxos, etc.) e energia (NADH e ATP) necessárias para a proliferação celular. Devido à glicólise, o ciclo do ácido tricarboxílico e o transporte de elétrons serem associados ao NAD⁺/NADH, a interrupção de qualquer uma dessas vias resultará na diminuição do anabolismo e da produção de energia, levando à supressão do crescimento. F-1,6-BP (fructose-1,6-bisfosfato); DHA (dihidroxiacetona); G3P (gliceraldeído-3-fosfato); PRPP (5-fosfoD-ribosil-1-pirofosfato).

A molécula CHEMBL566757 (FK866) foi reportada por induzir apoptose de células tumorais, reduzindo os níveis intracelulares de dinucleotídeo de nicotinamida (NAD⁺), mediado pela inibição da nicotinamida fosforibosiltransferase (NAMPT). A NAMPT também é conhecida como fator 1 de aumento de colônias células pré-B (PBEF) ou visfatina, e é a enzima limitante para a biossíntese de NAD⁺ a partir da nicotinamida. O NAD⁺ é um substrato da poli-ADP-ribose polimerase (PARP) e de uma subclasse de histonas desacetilases conhecidas como sirtuínas, que estão envolvidas na estabilidade genômica, apoptose, envelhecimento, resistência ao estresse e metabolismo. Como as células cancerígenas têm ativação contínua de PARP por danos ao DNA e instabilidade genômica e elas têm maiores demandas energéticas do que células não transformadas espera-se que as células cancerígenas sejam mais vulneráveis à inibição da síntese de NAD⁺ (YOU *et al.*, 2011 e TAKEUCHI *et al.*, 2014). A superexpressão de NAMPT mantém os níveis adequados de NAD⁺ nas células cancerígenas. De fato, a inibição sozinha dessa enzima parece ter um

impacto importante nos níveis de NAD^+ em vários tipos de células. Assim, as enzimas envolvidas no metabolismo de NAD^+ tornaram-se alvos atraentes na terapia do câncer (YOU *et al.*, 2011). Nesse mesmo contexto, o estudo de Mitchell *et al.* (2019) foi um dos primeiros a descrever a NAMPT como um alvo seletivo nas células de LMA *in vitro* e *in vivo*.

A similaridade entre as moléculas 4, 5, 6, 7, 8 e 9 e o composto CHEMBL225920 sugere que os derivados análogos de alcaloides de 3-alkilpiridina promovem a morte celular por apoptose e parada do ciclo celular.

A similaridade da molécula 7 com os compostos CHEMBL3105848 e CHEMBL566757 sugere que a apoptose e inibição do crescimento celular estão relacionados também à interferência no metabolismo das células cancerígenas. Especificamente a inibição das enzimas PFKFB3 e NAMPT, diminuem o influxo de glicose e disponibilidade de NAD^+ para células, o que leva à morte celular.

Considerando os dados preliminares de atividade citotóxica (CI_{50}) dos compostos 1 ao 9 e os dados da literatura de atividade citotóxica das moléculas do banco de dados ChEMBL similares aos compostos em estudo, foram feitas as relações a seguir.

No estudo de Majeed *et al.* (2013) o CI_{50} do composto CHEMBL269277 frente à linhagem THP-1 foi de 20 μM . Já no estudo da pesquisadora Mariana Guerra de Aguiar, os valores obtidos experimentalmente de CI_{50} frente à linhagem THP-1 para os compostos 1, 2 e 3 foram 14 μM , 13 μM e 10 μM , respectivamente, mostrando atividades equivalentes em baixo μM , considerando-se ensaios independentes. Dados similares foram observados para as demais comparações, conforme descritos a seguir. Atividade frente à linhagem K562, com CI_{50} dos compostos 1, 2 e 3 (16 μM , 51 μM e 11 μM , respectivamente) quando comparados ao estudo de Urban *et al.* (2004).

Em relação à atividade citotóxica da molécula CHEMBL168 (ácido oleanólico), o CI_{50} frente à linhagem K562 obtido no estudo de Kang *et al.* (2012) foi de 58 μM . Para os compostos 1 e 2 que foram similares ao CHEMBL168, o CI_{50} obtido frente à linhagem K562 foi de 16 μM e 51 μM , respectivamente, conforme já descrito anteriormente. Para a molécula CHEMBL169 (ácido ursólico) similar aos compostos 1, 2 e 3, o CI_{50} frente à linhagem THP-1 foi de 9 μM (RASHID *et al.*, 2013).

O composto CHEMBL225920 possui GI_{50} frente à linhagem K562 de 29 μM (LO *et al.*, 2007). Para os compostos similares a essa molécula, 4, 5, 6, 7, 8 e 9, foi obtido experimentalmente o CI_{50} de 49,78 μM , 140,55 μM , 74,09 μM , 194,25 μM , 21,66 μM , 15,37 μM , respectivamente. Dessa forma, apenas os compostos 8 e 9 possuem maior atividade citotóxica frente à linhagem K562 que o composto CHEMBL225920.

As moléculas CHEMBL3105848 e CHEMBL566757 foram similares ao composto 7. O CI_{50} da molécula CHEMBL3105848 frente à linhagem K562 registrado no banco de dados ChEMBL é igual a 3,2 μ M e para a molécula CHEMBL566757, 0,96 nM (ZHANG *et al.*, 2019). O composto CHEMBL566757 (FK866) é um inibidor da NAMPT e no estudo de Zhang *et al.* (2019) o CI_{50} frente à enzima NAMPT também foi determinado, sendo igual a 6,8 nM. O composto 7 não possui atividade frente à linhagem K562 (CI_{50} = 194 μ M), porém sua atividade frente às enzimas NAMPT e PFKFB3 não foi determinada experimentalmente. No entanto, o composto 7 possui atividade frente à linhagem THP-1, com CI_{50} = 37,99 μ M, sendo comparável à citarabina conforme descrito nos dados preliminares, o que indica sua atividade citotóxica.

6.2 Predição de alvo molecular empregando Servidores

Os resultados provenientes do servidor ChemMapper foram obtidos utilizando o limiar de similaridade de 1,5 e o ChEMBL como banco de dados. Os alvos moleculares preditos não relacionados ao câncer foram desconsiderados. Dessa forma, apenas o alvo não receptor tirosina-proteína fosfatase tipo 7 (PTPN7), com similaridade de 1,502, foi predito para o composto 9. Porém, essa proteína não está relacionada diretamente com a patogênese das leucemias em estudo. Portanto, ela não foi considerada para a análise de acoplamento abaixo.

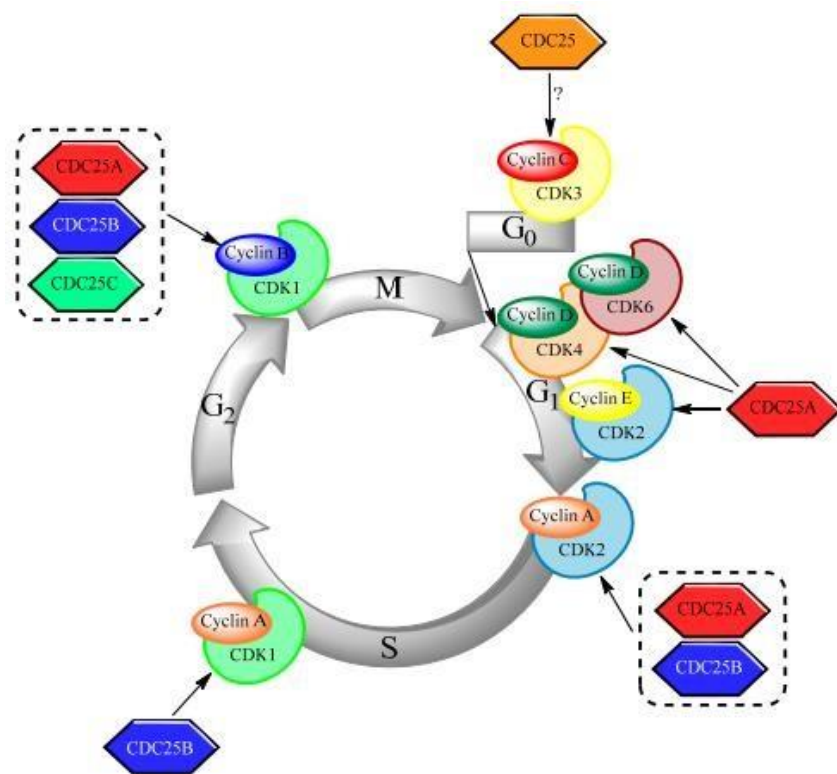
Os servidores Hitpick e SuperPred não realizaram predições de alvos moleculares relacionados à patogênese do câncer, portanto não foram apresentados os resultados dos demais alvos preditos.

O servidor Swiss Target Prediction retornou como alvos para os compostos 1 e 3 as proteínas fosfatase 1 indutora da fase M (CDC25A) e fosfatase 2 indutora da fase M (CDC25B). Para os demais compostos não foram preditos alvos relacionados ao câncer.

As fosfatases CDC25 (CDC25A, B e C) são os reguladores principais da atividade das quinases dependentes de ciclina (CDKs) (**Figura 11**). Essas três enzimas são importantes reguladoras de várias etapas do ciclo celular e elas possivelmente, tem um papel no desenvolvimento de uma variedade de neoplasias humanas. A inibição de CDC25 é considerada, portanto, como uma possível estratégia terapêutica no tratamento do câncer (BRENNER *et al.*, 2014). A CDC25A tem funções importantes durante a replicação e mitose, bem como durante a fase G1 do ciclo celular. Sua superexpressão foi descrita em diferentes categorias de câncer e é associada a um prognóstico adverso (BERTOLI *et al.*, 2015). Existem

várias vias que regulam o gene e a proteína CDC25A, e está bem estabelecido que variações discretas do nível da proteína afetem a estabilidade genômica e a transformação celular induzida por vários oncogenes (SUEUR *et al.*, 2020). A CDC25B regula a transição de fase G₂/M por remover dois grupos fosfato inibitórios do loop de ligação do ATP da CDK2. A CDC25B é superexpressa em vários cânceres, levando à ativação excessiva de CDK2/Ciclina A e progressão aberrante do ciclo celular resultando em pior prognóstico. Estudos genéticos demonstram o papel essencial de CDC25B no crescimento das células tumorais, sendo a CDC25B um alvo terapêutico atrativo para inibição por pequenas moléculas (LUND *et al.*, 2015).

Figura 11 - Visão geral da função reguladora das fosfatases CDC25A, B e C na progressão do ciclo celular.



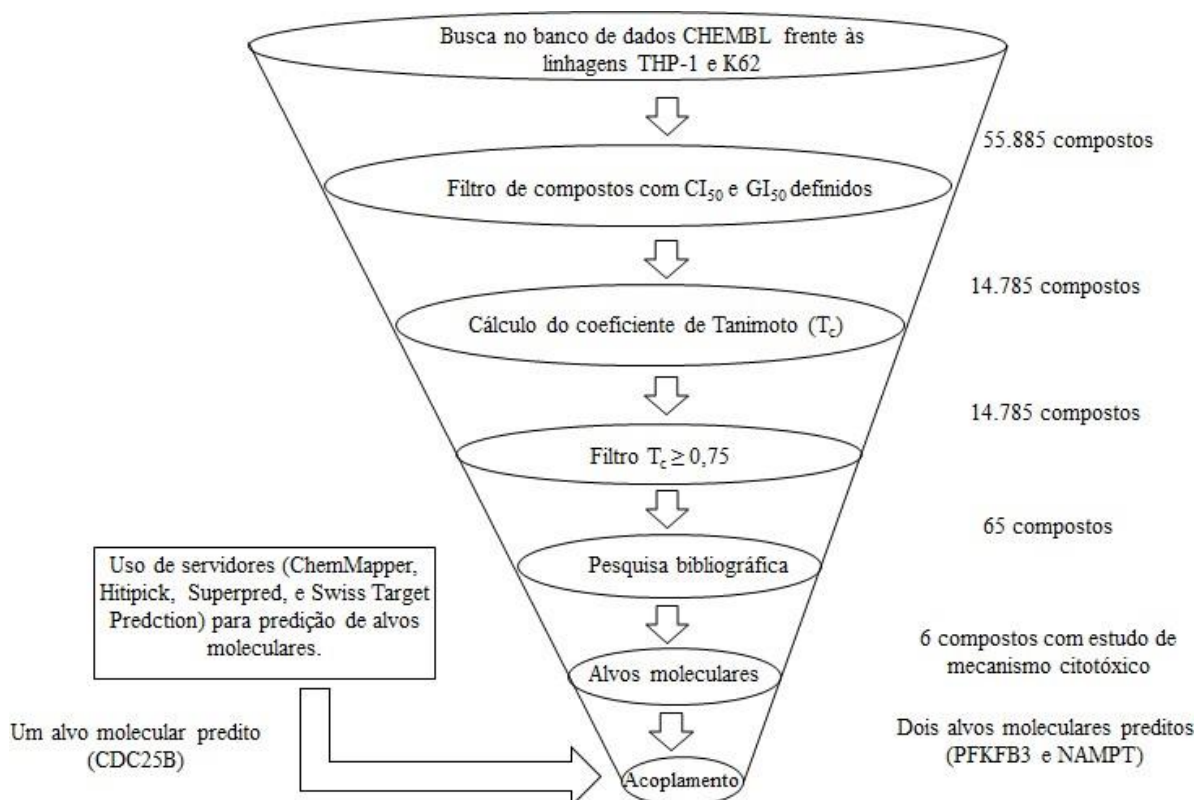
Fonte: BRENNER *et al.*, 2014.

Legenda: Na fase G₀, as células que estavam quiescentes, voltam a entrar no ciclo celular após a ativação da CDK3/ciclina C. A desfosforilação da CDK4/ciclina D, CDK6/ciclina D e CDK2 em complexo com a ciclina E e ciclina A pelas CDC25s levam à transição para a fase de replicação do DNA. Na fase S tardia, a CDC25B ativa a CDK1/ciclina A. Por fim, a desfosforilação de CDK1/ciclina B desencadeia a entrada mitótica, e nesta etapa todas as três isoformas de CDC25 estão envolvidas. No final da mitose, tanto a CDK1/ciclina B quanto as CDC25s são degradados e o ciclo pode ser reiniciado.

6.3 Acoplamento molecular

Após as predições dos potenciais alvos moleculares e a devida seleção (**Figura 12**), baseada nos dados da literatura sobre a relação destas proteínas com os possíveis mecanismos de citotoxicidade, os estudos de acoplamento foram realizados com o intuito de simular o modo de ligação entre os inibidores estudados e os alvos preditos. Nesta etapa, foi considerado então se o modo de ligação dos compostos analisados é condizente com o modo de ligação de ligantes conhecidos dos alvos moleculares preditos.

Figura 12 - Representação esquemática da seleção de alvos moleculares.



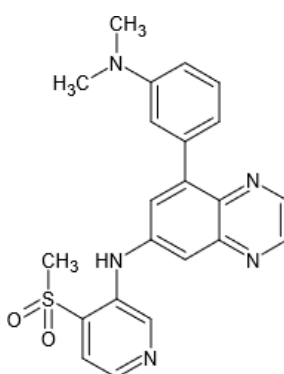
Fonte: Elaborada pela autora.

Considerando a análise de predição de alvos por similaridade estrutural, as proteínas PFKFB3 e NAMPT foram consideradas potenciais alvos para o composto 7. Portanto, o composto 7 foi submetido ao estudo de acoplamento molecular com ambas as proteínas preditas. Como resultado da predição utilizando servidores, o Swiss Target Prediction retornou como alvos para os compostos 1 e 3 as proteínas CDC25A e CDC25B. Porém, no PDB, está disponível apenas a estrutura cristalográfica da CDC25B. Portanto, somente essa proteína foi utilizada para o acoplamento com os compostos 1 e 3.

6.3.1 Proteína PFKFB3

A estrutura tridimensional da proteína PFKFB3 depositada no PDB sob o código de identificação 6HVI, com resolução de 1,96 Å e R_{free} 0,214 foi selecionada. Foram mantidas as moléculas de águas estruturais (702) e (907) para os cálculos de acoplamento molecular, pois fazem ligações de hidrogênio com ligante. Nesta estrutura, a proteína está complexada com um inibidor, denominado GV5 (**Figura 13**), derivado de N-aryl-6-aminoquinoxalina.

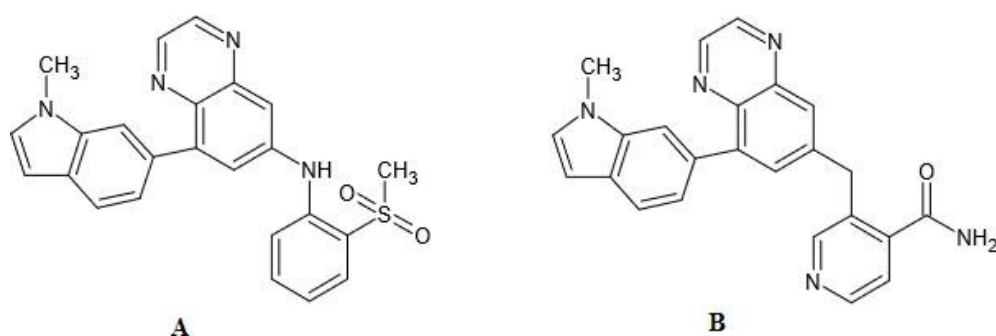
Figura 13 - Estrutura molecular do ligante GV5.



Fonte: Elaborada pela autora.

Segundo Boutard *et al.* (2019), considerando os compostos A e B derivados de N-aryl-6-aminoquinoxalina (**Figura 14**), o núcleo quinoxalina se posiciona no sítio de ligação do ATP da PFKFB3 através de prováveis ligações de hidrogênio com os resíduos de aminoácido Asn163 e Ser152, bem como a potencial interação hidrofóbica com Cys154. Anéis aromáticos bicíclicos não polares na posição 8 da quinoxalina interagem com um bolso lipofílico relativamente plano e o substituinte amino aril orto-substituído na posição 6 interage por empilhamento- π com o resíduo Tyr49 bem como prováveis interações ânion- π ou ligações de hidrogênio com o resíduo Glu166. Os derivados de sulfonamida N-substituídos demonstraram um aumento de atividade no alvo, provavelmente devido ao surgimento de interações adicionais hidrofóbicas com resíduos da região do sítio de ligação, como Val243, Val159, Val217, Pro421 e Ser162.

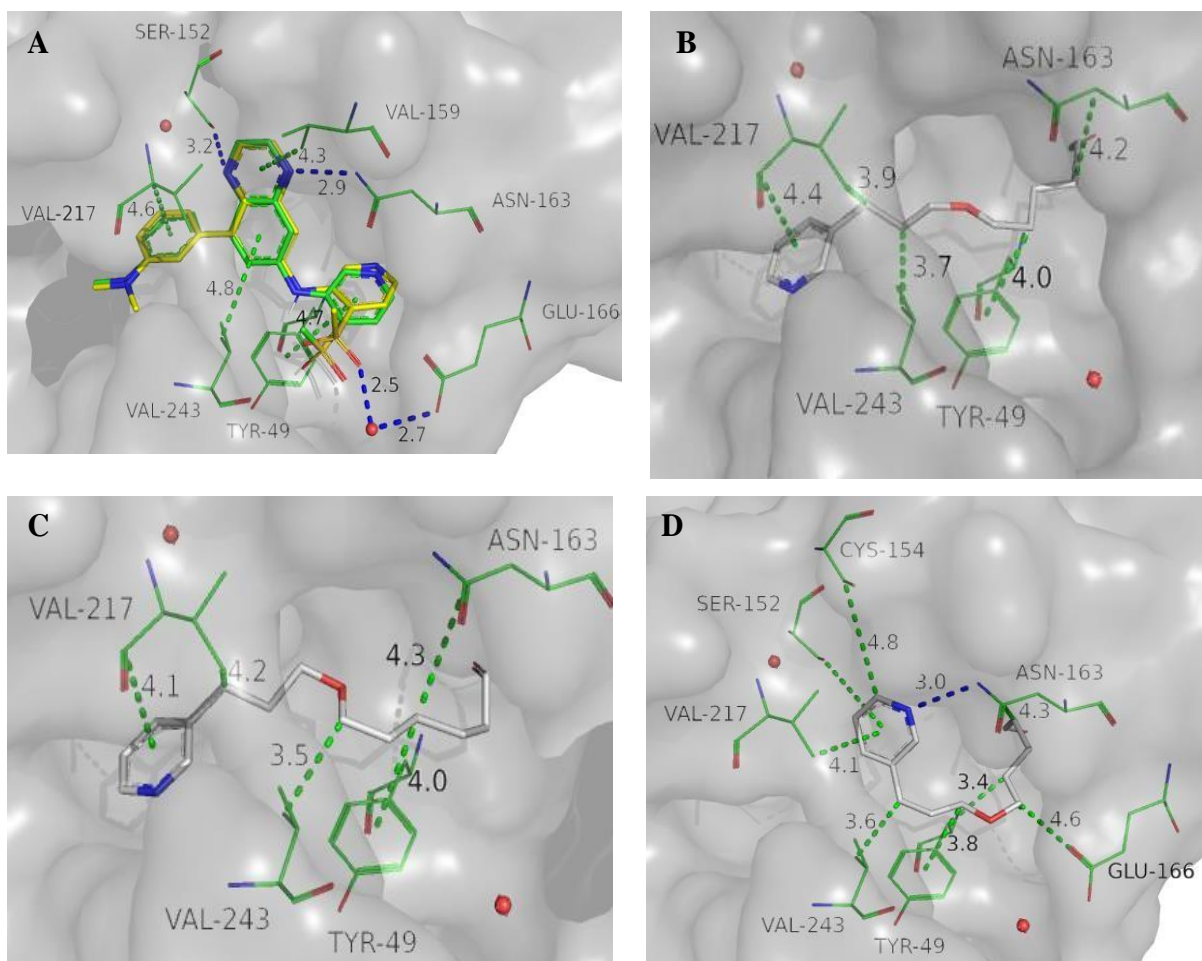
Figura 14 - Compostos A e B derivados de N-aril-6-aminoquinoxalina.



Fonte: Elaborada pela autora.

A pose melhor ranquada do *redocking* do ligante GV5 apresenta a maioria das interações com o sítio ativo descritos na literatura (**Figura 15A**). A ligação de hidrogênio com o resíduo Glu166 é mediada por uma molécula de água (702) e não diretamente como descrito na literatura. Portanto, pode-se dizer que o protocolo de acoplamento adotado para este estudo é adequado para predizer o modo de ligação entre inibidores e a proteína PFKFB3.

Figura 15 - Sobreposição das poses mais bem ranqueadas no *redocking* (A) e acoplamento (B, C, D) da proteína PFKFB3.



Fonte: Elaborada pela autora.

Legenda: A) Sobreposição da pose mais bem ranqueada do ligante GV5 co-cristalizado com a enzima PFKFB3 durante o experimento de *redocking*. A pose do ligante proveniente do *redocking* está representada em cor predominante amarela e em verde está representado o sítio de ligação e o ligante GV5 co-cristalizado. Todas as interações do ligante com os aminoácidos dos sítios de ligação foram mantidas no *redocking*. B) Pose melhor ranqueada do segundo grupo mais populoso do acoplamento do composto 7 com a proteína PFKFB3. C) Pose mais bem ranqueada do grupo mais populoso do acoplamento do composto 7 com a proteína PFKFB3. D) Pose melhor ranqueada do acoplamento do composto 7 com a proteína PFKFB3. O composto 7 está representado em cor predominante cinza.

As moléculas de água estão representadas por esferas vermelhas, as ligações de hidrogênio por linhas tracejadas azuis e as interações hidrofóbicas por linhas tracejadas verdes, com suas respectivas distâncias em Å.

As poses mais representativas do composto 7 agrupadas com 1,48 Å de RMSD foram analisadas. Para isso, a pose de maior *score* dos dois grupos mais populosos foi selecionada: um dos grupos contém 21 poses (10,5%) e o outro, 29 poses (14,5%). As poses mais bem ranqueadas destes grupos foram a pose 70 (**Figura 15B**) e a pose 82 (**Figura 15C**), respectivamente. A pose 160 (**Figura 15D**) foi a melhor ranqueada dentre as 200 e também foi analisada.

Quadro 1 - Massa molecular, *score* e *score* normalizado do ligante GV5, composto 7 e poses mais representativas e mais bem ranqueadas para o *redocking* e acoplamento com a proteína PFKFB3.

Composto	Massa molecular (g/mol)	Pose	Score	Score normalizado
GV5	419	-	104,82	0,25
7	235	70	80,05	0,34
		82	73,77	0,31
		160	82,96	0,35

Fonte: Elaborado pela autora.

As poses melhores ranqueadas do primeiro e segundo grupos mais populosos apresentam interações hidrofóbicas com Asn163 e com Tyr49, que apesar de não serem as mesmas interações descritas na literatura, são interações na mesma região do sítio ativo da proteína PFKFB3. Outras ligações hidrofóbicas que contribuem para o posicionamento semelhante ao GV5 no sítio de ligação são com os aminoácidos Val217 e Val243, que estão entre as que possivelmente contribuem para aumento de atividade (BOUTARD *et al.*, 2019). Porém, as importantes interações com a Glu166 e Cys154 não estão presentes nessas poses. Essas duas poses são muito semelhantes e juntas representam 25% de todas as poses geradas.

Por último a pose de maior *score*, possui interações hidrofóbicas com Asn163 e com Tyr49, além da provável ligação de hidrogênio do anel piridínico com a Asn163, sendo importante para o ligante na ancoragem ao sítio de ligação. Adicionalmente, interage de forma hidrofóbica na região dos resíduos de Glu166, Cys154 e Ser152, que compõem o sítio ativo da proteína PFKFB3.

As três poses analisadas do composto 7 apresentaram prováveis interações importantes no sítio de ligação do ATP da proteína PFKFB3 e valores de *score* normalizado semelhantes e superiores ao da pose do *redocking*, indicando que a composto 7 pode ter afinidade pelo alvo molecular.

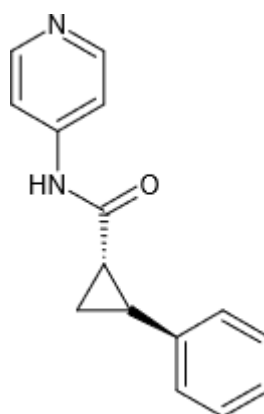
6.3.2 Proteína NAMPT

Em relação à proteína NAMPT, foi selecionada para o acoplamento a estrutura tridimensional depositada no PDB sob o código de identificação 4LVF, com a resolução de 1,50 Å e R_{free} 0,172 complexada ao ligante 20P (**Figura 16**). A estrutura depositada sob o código 6E68, apesar de apresentar resolução igual a 1,50 Å e R_{free} 0,167, não foi selecionada, pois não possui artigo de referência publicado.

Foi mantida a molécula de água estrutural (827) para os cálculos de acoplamento molecular, pois dentre as dez estruturas cristalográficas analisadas em relação à posição das águas estruturais, em quatro delas (estruturas cristalográficas: 6E68, 6TA0, 5KIT e 5UPF) o ligante co-cristalizado possui ligação de hidrogênio com o resíduo de aminoácido Asp219, mediada por uma molécula de água na posição semelhante à água (827).

As estruturas cristalográficas disponíveis para a proteína NAMPT no PDB, representadas no **ANEXO II**, possuem RMSD < 0,20 em relação à 4LVF indicando alta similaridade entre elas.

Figura 16 - Estrutura molecular do ligante 20P.

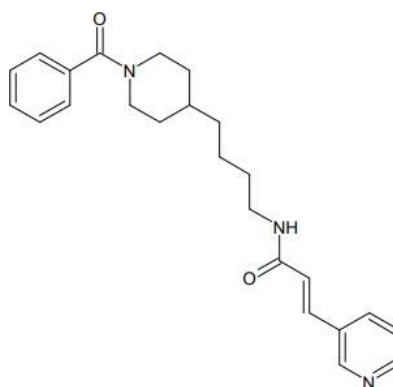


Fonte: Elaborada pela autora.

A NAMPT é uma proteína composta por duas cadeias (A e B) e somente o resíduo de Tyr18 da cadeia B faz parte do sítio de ligação, sendo o restante dos resíduos pertencentes à cadeia A. Segundo Clarck *et al.* (2016), o FK866 (**Figura 17**), potente inibidor da NAMPT, ocupa um sítio de ligação extenso, porém existe apenas uma provável ligação de hidrogênio direta com a proteína (com a Ser275) e duas ligações de hidrogênio mediadas por moléculas de água (242) e (645) com a Asp219. As possíveis interações do tipo π - π entre o

grupo vinil-piridina e a região delimitada pelos resíduos de aminoácido Phe193 e Tyr18B, são importantes para o posicionamento do inibidor no sítio de ligação da proteína. Em contraste, a cauda benzoil-piperidina é mais fracamente ligada, com o anel fenil ocupando uma pequena bolsa delimitada pelas cadeias laterais dos resíduos de Lys189 e Arg349. O estudo de Dragovich *et al.* (2014), demonstrou interações hidrofóbicas favoráveis do ligante com o túnel hidrofóbico da proteína formado por Ile309, Ile351 e His191.

Figura 17 - Estrutura molecular do FK866.

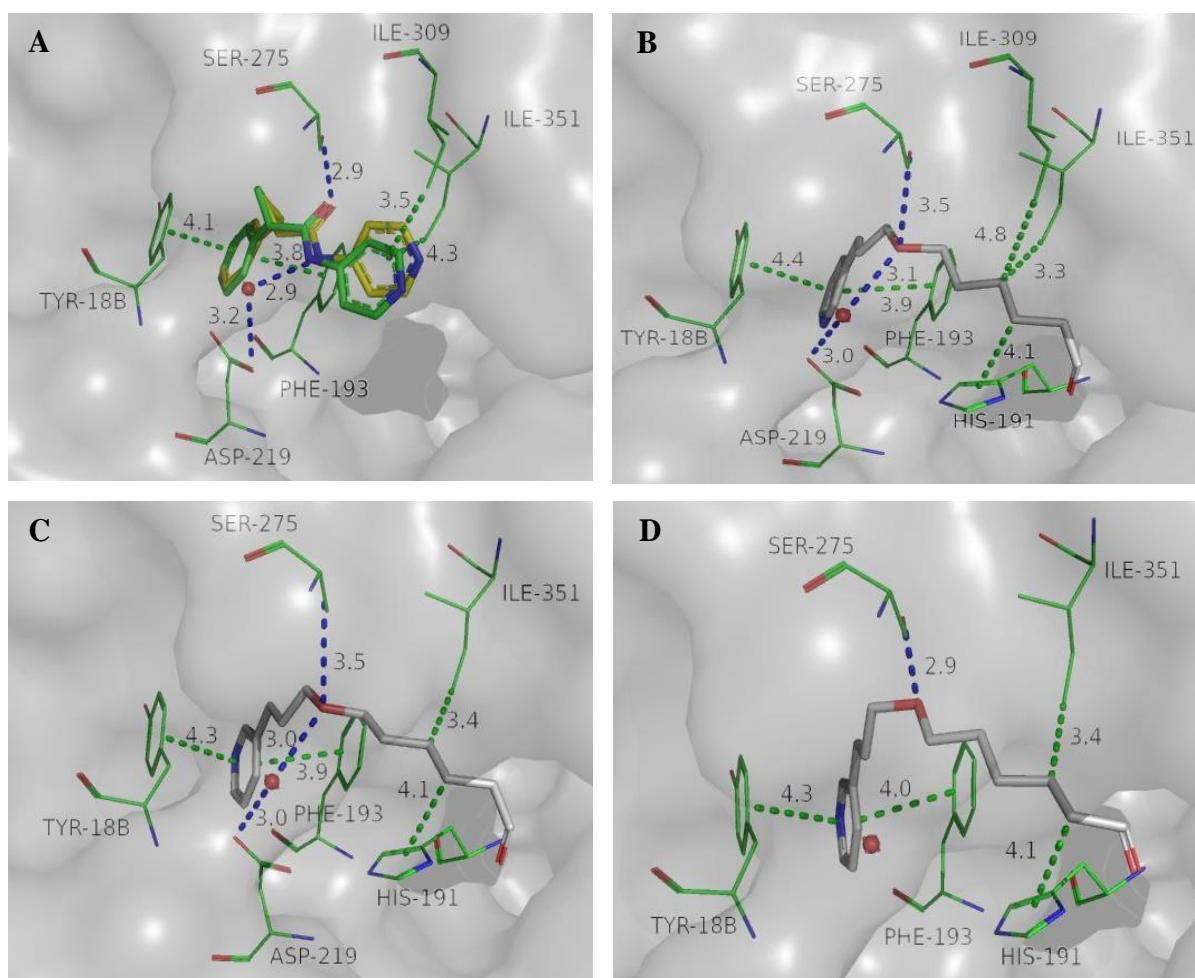


Fonte: Elaborada pela autora.

No estudo de Giannetti *et al.* (2014), o ligante 20P apresentou atividade inibitória frente à NAMPT. A porção piridil-anilina estendeu-se para a região do túnel hidrofóbico, e o substituinte fenil-ciclopropano ocupou o sítio de ligação. A amida possui ligações de hidrogênio com as cadeias laterais dos resíduos Asp219, mediada pela molécula de água (827), e direta com a Ser275.

A pose mais bem ranqueada do *redocking* do ligante 20P apresenta as prováveis ligações de hidrogênio mediada pela molécula de água estrutural (827) com a Asp219 e direta com a Ser275, além da interação do tipo π - π com os resíduos de Phe193 e a Tyr18B, importantes interações descritas na literatura (**Figura 18A**). A interação do ligante com o túnel hidrofóbico é mediado pelas prováveis interações hidrofóbicas com a Ile309 e Ile351. Portanto, pode-se dizer que o protocolo de acoplamento adotado para este estudo é adequado para prever o modo de ligação entre inibidores e a proteína NAMPT.

Figura 18 - Sobreposição das poses mais bem ranqueadas no *redocking* (A) e acoplamento (B, C e D) da proteína NAMPT.



Fonte: Elaborada pela autora.

Legenda: A) Sobreposição da pose mais bem ranqueada do ligante 20P co-cristalizado com a enzima NAMPT durante o experimento de *redocking*. A pose do ligante proveniente do *redocking* está representada em cor predominante amarela e em verde está representado o sítio de ligação e o ligante 20P co-cristalizado. Todas as interações do ligante com os aminoácidos dos sítios de ligação foram mantidas no *redocking*. B) Pose melhor ranqueada do grupo mais populoso do acoplamento do composto 7 com a proteína NAMPT. C) Pose mais bem ranqueada do segundo grupo mais representativo do acoplamento do composto 7 com a proteína NAMPT. D) Pose melhor ranqueada do acoplamento do composto 7 com a proteína NAMPT. O composto 7 está representado em cor predominante cinza.

A molécula de água está representada por esfera vermelha, as ligações de hidrogênio por linhas tracejadas azuis e as interações hidrofóbicas por linhas tracejadas verdes, com suas respectivas distâncias em Å.

As poses do composto 7 agrupadas com 1,49 Å de RMSD foram analisadas após o acoplamento. Os três grupos mais populosos foram o quinto com 46 poses (23,0%) e o terceiro com 36 poses (18,0%). As poses mais bem ranqueadas destes grupos foram a pose 186 (**Figura 18B**) e a pose 51 (**Figura 18C**), respectivamente. A pose 156 (**Figura 18D**) foi a melhor ranqueada dentre as 200 e também foi analisada.

Quadro 2 - Massa molecular, *score* e *score* normalizado do ligante 20P, composto 7 e poses mais representativas e mais bem ranqueadas para o *redocking* e acoplamento com a proteína NAMPT.

Composto	Massa molecular (g/mol)	Pose	Score	Score normalizado
20P	238	-	79,35	0,33
7	235	186	81,75	0,35
		51	82,23	0,35
		156	85,30	0,36

Fonte: Elaborado pela autora.

As duas poses mais bem ranqueadas dos grupos mais populosos e a pose melhor ranqueada dentre as 200 poses analisadas possuem a provável ligação de hidrogênio com a Ser275, descrita na literatura. O anel piridínico se posiciona de forma favorável na região do sítio ativo, entre a Phe193 e a Tyr18B, por meio das interações π - π . Já a cadeia apolar do composto 7 se posiciona no túnel hidrofóbico da proteína NAMPT.

A importante ligação de hidrogênio com a Asp219 mediada pela molécula de água estrutural (827) está presente nas duas poses melhores ranqueadas dos grupos mais populosos, porém é perdida na pose mais bem ranqueada.

As três poses analisadas do acoplamento do composto 7 com a proteína NAMPT, apresentaram prováveis interações no sítio ativo da proteína e possuem *score* normalizado próximos e superiores ao do *redocking*, indicando que o composto 7 possa interagir de forma favorável com esse alvo molecular.

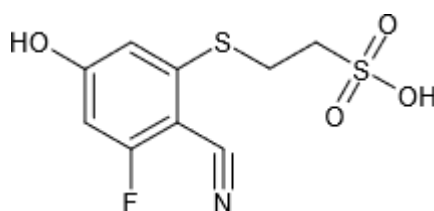
6.3.3 Proteína CDC25B

A estrutura tridimensional depositada no PDB sob o código de identificação 4WH9, com a resolução de 1,50 Å e R_{free} 0,152 foi selecionada para como modelo da proteína CDC25B complexada ao ligante 3M8 (**Figura 19**). Foram mantidas as moléculas de águas estruturais (714), (808), (821) e (831) para os cálculos de acoplamento molecular. Essas

moléculas de água estão em posições semelhantes às águas que fazem ligações de hidrogênio entre o ligante co-cristalizado e a estrutura cristalográfica 4WH7.

A mutação presente nesse alvo proteico é no sítio ativo e não no sítio de ligação do 3M8, portanto foi utilizado para o acoplamento molecular. Além disso, não está disponível no PDB outra estrutura sem mutações.

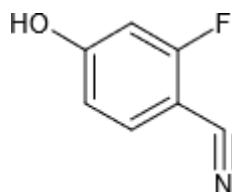
Figura 19 - Estrutura molecular do ligante 3M8



Fonte: Elaborada pela autora.

Segundo Lund *et al.* (2015) o inibidor 2-fluoro-4-hidroxibenzonitrila (**Figura 20**) se liga em uma pequena região bem definida de CDC25B localizada a 15 Å de distância do sítio ativo. Portanto, este é um inibidor alostérico. Esse local de ligação é composto principalmente pelas cadeias laterais da Phe386, Leu398, Cys484, Arg488 e Met505. O anel fenil se posiciona entre as cadeias laterais da Leu398 e Arg488, formando interações hidrofóbicas e cátion- π , respectivamente. Uma característica interessante desse sítio de ligação é a presença de numerosas moléculas de água que formam uma rede de ligações de hidrogênio. Essas águas formam ligações de hidrogênio com a carbonila e amida da Met505 e amidas da Lys399 e Gly380. Com base em estudos de mutagênese, a Arg488 e a Arg492 em CDC25B mostraram ser necessários para recrutamento de substrato CDK2 por meio de interações iônicas com Asp206.

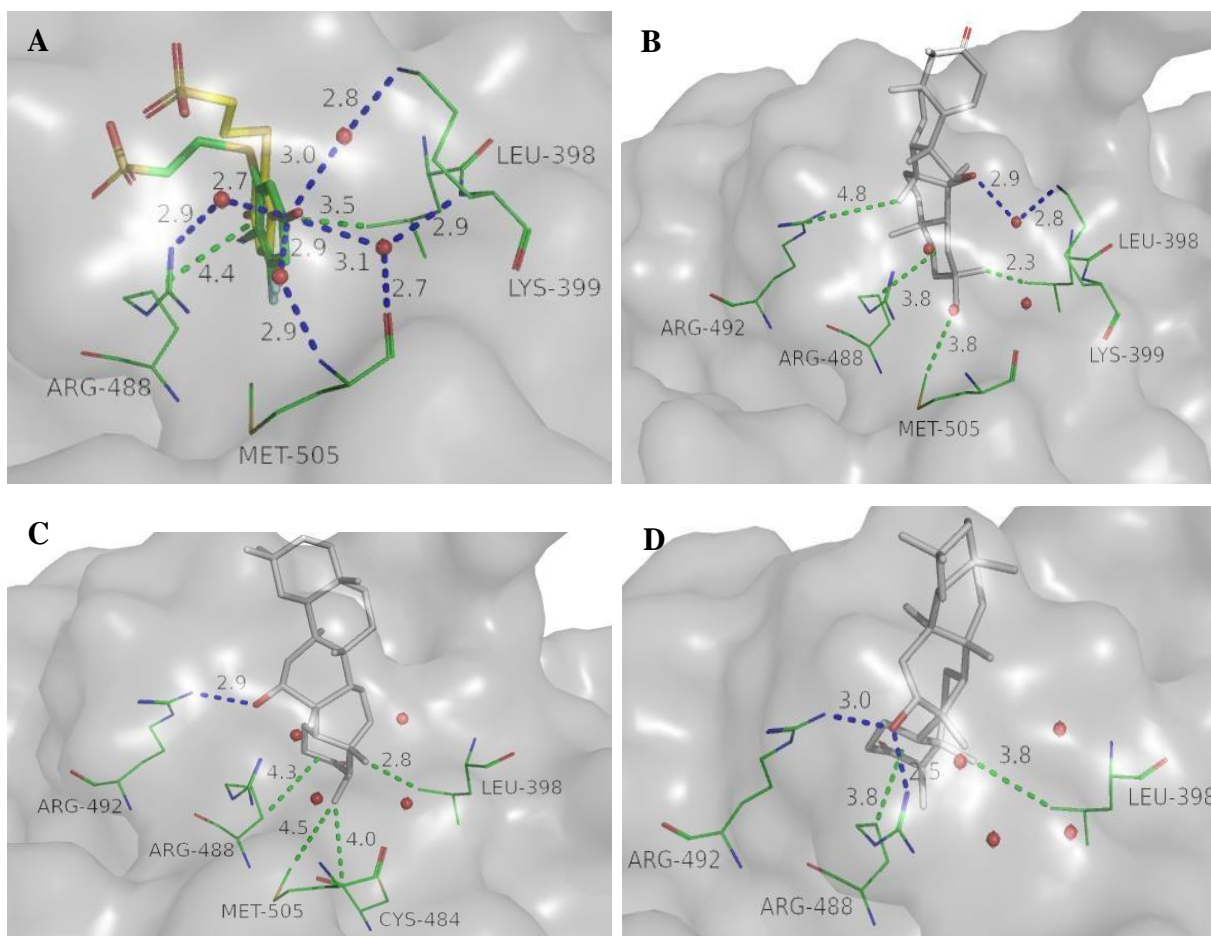
Figura 20 - Estrutura molecular do inibidor 2-fluoro-4-hidroxibenzonitrila.



Fonte: Elaborada pela autora.

Na pose mais bem ranqueada do *redocking* do ligante 3M8 (**Figura 21A**), o anel aromático está posicionado entre as cadeias laterais da Leu398 e Arg488, formando interações hidrofóbicas. Possui prováveis ligações de hidrogênio mediadas pelas moléculas de água (808) e (821) com a Met505 e pelas moléculas de água (714) e (821) com a Lys399, conforme descrito na literatura. Outras interações importantes para a atividade da proteína são a possível ligação de hidrogênio, mediada pela molécula de água (831), e a interação hidrofóbica com a Arg488. Portanto, o protocolo de acoplamento adotado para este estudo é adequado para prever o modo de ligação entre inibidores e a proteína CDC25B.

Figura 21 - Sobreposição das poses mais bem ranqueadas no *redocking* (A) e acoplamento (B, C e D) da proteína CDC25B com o composto 1.



Fonte: Elaborada pela autora.

Legenda: A) Sobreposição da pose mais bem ranqueada do ligante 3M8 co-cristalizado com a enzima CDC25B durante o experimento de *redocking*. A pose do ligante proveniente do *redocking* está representada em cor predominante amarela e em verde está representado o sítio de ligação e o ligante 3M8 co-cristalizado. Importantes interações do ligante com os aminoácidos do sítio de ligação foram mantidas no *redocking*. B) Pose melhor ranqueada do grupo mais populoso do acoplamento do composto 1 com a proteína CDC25B. C) Pose mais bem ranqueada do segundo grupo mais populoso do acoplamento do composto 1 com a proteína CDC25B. D) Pose melhor ranqueada do acoplamento do composto 1 com a proteína CDC25B.

O composto 1 está representado em cor predominante cinza. As moléculas de águas estruturais estão representadas por esferas vermelhas e suas interações em linhas tracejadas, sendo as azuis ligações de hidrogênio e as verdes interações hidrofóbicas, com suas respectivas distâncias em Å.

As poses agrupadas com 1,30 Å de RMSD foram analisadas para o composto 1. Os dois grupos mais populosos foram o segundo com 45 poses (22,5%) e o sexto com 44 poses (22,0%). As poses mais bem ranqueadas destes grupos foram a pose 189 (**Figura 21B**) e a pose 134 (**Figura 21C**), respectivamente. A pose 83 (**Figura 21D**) foi a mais bem ranqueada dentre as 200 e também foi analisada.

Quadro 3 - Massa molecular, *score* e *score* normalizado do ligante 3M8, composto 1 e poses mais representativas e melhor ranqueadas para o *redocking* e acoplamento com a proteína CDC25B.

Composto	Massa molecular (g/mol)	Pose	Score	Score normalizado
3M8	277	-	82,14	0,30
1	426	189	62,61	0,15
		134	57,76	0,14
		83	64,48	0,15

Fonte: Elaborado pela autora.

Na pose mais bem ranqueada do grupo mais representativo, o composto 1 interage na região dos resíduos de aminoácido Arg488, Arg492 e Met505 por meio de prováveis interações hidrofóbicas. Um dos anéis se posiciona no sítio de ligação da proteína CDC25B de forma semelhante ao anel aromático do ligante 3M8 por meio de interações hidrofóbicas. A provável ligação de hidrogênio com a Lys399, mediada pela molécula de água (714), também está presente.

Na pose melhor ranqueada do segundo grupo mais representativo, as interações com a proteína CDC25B são semelhantes à pose anterior, porém a importante interação com a Lys399 é perdida. Em contrapartida, existe a possível ligação de hidrogênio com a Arg492 e a interação hidrofóbica com o resíduo de Cys484. Portanto, a pose mais bem ranqueada do primeiro grupo mais populoso possui mais interações favoráveis com o sítio de ligação que a pose melhor ranqueada para o segundo grupo mais populoso.

Por último a pose de maior *score*, possui prováveis ligações de hidrogênio na região dos resíduos de aminoácido Arg488 e Arg492. Além das interações hidrofóbicas com a Leu398 e a Arg488.

As três poses analisadas do acoplamento do composto 1 com a proteína CDC25B, apresentaram *score* normalizado próximos, porém abaixo ao do *redocking*, indicando não ser favorável a interação do composto 1 com esse alvo molecular ou uma interação com menor afinidade pela proteína.

Considerando o acoplamento para o composto 3, as poses agrupadas com 1,24 Å de RMSD foram analisadas. Os dois grupos mais populosos foram o primeiro com 140 poses (70,0%) e o terceiro com 33 poses (16,5%). As poses mais bem ranqueadas destes grupos foram a pose 66 (**Figura 22B**) e a pose 116 (**Figura 22C**), respectivamente. A pose 66 foi também, a melhor ranqueada dentre as 200.

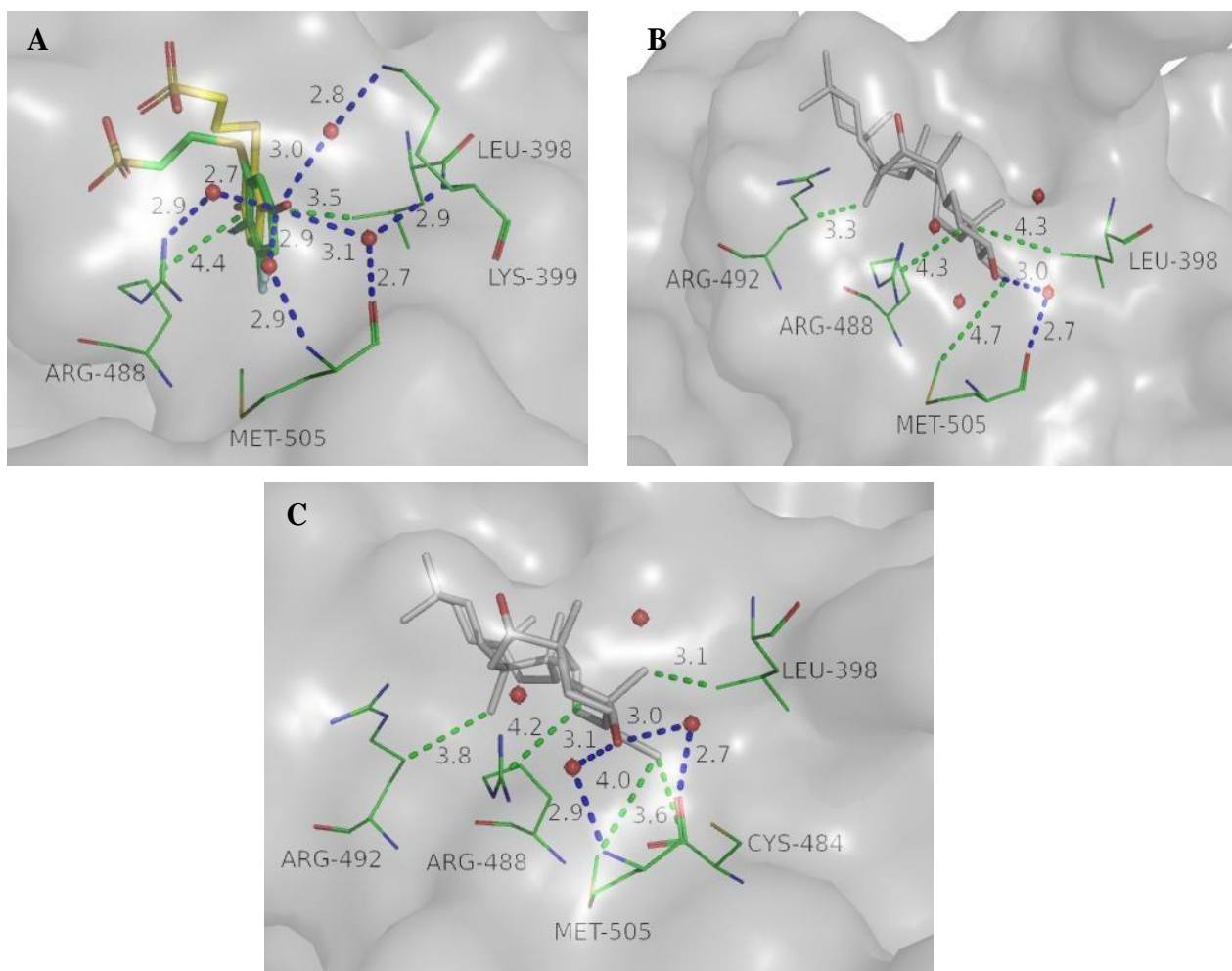
Quadro 4 - Massa molecular, *score* e *score* normalizado do ligante 3M8, composto 3 e poses mais representativas e mais bem ranqueadas para o *redocking* e acoplamento com a proteína CDC25B.

Composto	Massa molecular (g/mol)	Pose	Score	Score normalizado
3M8	277	-	82,14	0,30
3	428	66	61,94	0,14
		116	57,66	0,13

Fonte: Elaborada pela autora.

Os valores de *score* normalizados calculados para as poses geradas na análise do acoplamento do composto 3 com a proteína CDC25B foram muito semelhantes aos das poses geradas para o composto 1, indicando similaridade no posicionamento dessas poses no sítio de ligação da proteína. Portanto, assim como o composto 1, o composto 3 parece não interagir de forma favorável com a CDC25B ou pode interagir com afinidade menor pelo alvo molecular.

Figura 22 - Sobreposição das poses mais bem ranqueadas no *redocking* (A) e acoplamento (B e C) da proteína CDC25B com o composto 3.



Fonte: Elaborada pela autora.

Legenda: A) Sobreposição da pose mais bem ranqueada do ligante 3M8 co-cristalizado com a enzima CDC25B durante o experimento de *redocking*. A pose do ligante proveniente do *redocking* está representada em cor predominante amarela e em verde está representado o sítio de ligação e o ligante 3M8 co-cristalizado.

Importantes interações do ligante com os aminoácidos do sítio de ligação foram mantidas no *redocking*. B) Pose melhor ranqueada do grupo mais populoso do acoplamento do composto 3 com a proteína CDC25B. C) Pose mais bem ranqueada do segundo grupo mais populoso do acoplamento do composto 3 com a proteína CDC25B. O composto 3 está representado em cor predominante cinza. As moléculas de águas estruturais estão representadas por esferas vermelhas e suas interações em linhas tracejadas, sendo as azuis ligações de hidrogênio e as verdes interações hidrofóbicas, com suas respectivas distâncias em Å.

A análise das poses geradas no acoplamento molecular demonstrou que as proteínas PFKFB3 e NAMPT são alvos promissores para o composto 7, pois importantes interações com o sítio ativo já descritas na literatura estão presentes e os valores de *score* normalizado indicaram que o composto 7 se liga de forma favorável a esses alvos moleculares. Porém, o mesmo não acontece para os compostos 1 e 3, já que as poses melhor

ranqueadas geradas no acoplamento indicaram que estes compostos não se ligam de forma favorável no sítio alostérico da CDC25B.

Com o intuito de validar os resultados do estudo *in silico* foi proposto o tratamento da linhagem celular THP-1 com o composto 7 para análise metabolômica, por meio da espectroscopia por ressonância magnética nuclear (RMN). Porém, devido ao acesso restrito aos Laboratórios de pesquisa durante a pandemia não foi possível concluir essas análises.

Metabolômica é o nome aceito para a ciência das ômicas que trata da caracterização do metaboloma, que por sua vez, é definido como todo o conjunto de metabólitos em um determinado sistema biológico, como uma célula, um tecido, um órgão ou um organismo (VIGNOLI *et al.*, 2019). Estudos metabólicos baseados em células *in vitro*, muitas vezes combinados com outras ômicas, têm uso difundido em muitas áreas de pesquisa. Estes incluem: estudos sobre o efeito, ação e toxicologia de drogas, caracterização e compreensão do metabolismo da célula cancerosa, medicina regenerativa, metabolismo imunológico e muitos mais. O objetivo comum de todos eles é compreender a influência e envolvimento do metabolismo sobre/nos efeitos e mecanismos biológicos, e integrar essas informações em mapas metabólicos (KOSTIDIS *et al.*, 2017). Nesse contexto, as culturas de células podem fornecer uma alternativa para a compreensão do metabolismo específico de candidatos a drogas. Para trabalhar com as culturas celulares, dados metabólicos quantitativos detalhados são necessários para ambos os compartimentos, intra e extracelular. Nos últimos anos, várias abordagens metabolômicas foram desenvolvidas para obter esses dados metabólicos quantitativos, incluindo espectrometria de massa e método baseado em NMR (KOSTIDIS *et al.*, 2017).

Considerando que as proteínas PFKFB3 e NAMPT estão envolvidas no metabolismo celular e foram apontadas como prováveis alvos moleculares, a análise metabolômica de células da linhagem THP-1 expostas ao composto 7 seria importante para o esclarecimento do mecanismo de ação de citotoxicidade desse composto, pois indicaria quais vias metabólicas estariam inibidas e quais estimuladas.

7 CONCLUSÃO E PERSPECTIVAS

A metodologia de *target fishing* combinada ao acoplamento molecular é amplamente utilizada para a determinação de alvos moleculares. A disponibilidade de

compostos com atividade citotóxica *in vitro* conhecida foi o que embasou a escolha dessa estratégia para busca de alvos.

As duas ferramentas de *target fishing* utilizadas nesse estudo, pesquisa de moléculas similares aos compostos de interesse a partir do banco de dados ChEMBL e uso de servidores, foram conduzidas com sucesso na busca de potenciais alvos moleculares já descritos na literatura. Os dados obtidos a partir das moléculas obtidas no ChEMBL resultaram como possíveis alvos as proteínas PFKFB3 e NAMPT, ambas envolvidas no metabolismo celular. Já a proteína CDC25B foi predita pelo servidor Swiss Target Prediction.

Após a análise das poses geradas no estudo de acoplamento molecular as proteínas PFKFB3 e NAMPT foram confirmadas como alvos promissores para o composto 7. No entanto, apesar da inibição da progressão do ciclo celular depender principalmente da inibição da expressão das ciclinas, os resultados do acoplamento molecular dos compostos 1 e 3 com a proteína CDC25B, não demonstraram interações favoráveis.

Com intuito de validar o estudo *in silico* realizado, além da análise metabolômica proposta, a avaliação da expressão gênica de PFKFB3 e NAMPT em linhagens de células leucêmicas também é uma ferramenta a ser utilizada no esclarecimento do mecanismo de citotoxicidade dos compostos de interesse.

Outra perspectiva a ser considerada para os triterpenos pentacíclicos (compostos 1 ao 3) seria a análise da permeação/interação com membranas, considerando as ações descritas na literatura para o ácido betulínico, que promove alteração no potencial da membrana das mitocôndrias promovendo a apoptose. Desta forma, não podemos descartar a hipótese do mecanismo de ação dos triterpenos ser mediado por interações com alvos não-proteicos.

As proteínas PFKFB3 e NAMPT estão envolvidas no metabolismo e crescimento celular e possuem expressão aumentada nas leucemias, portanto a busca por inibidores dessas proteínas é de grande importância para a melhoria das quimioterapias disponíveis.

REFERÊNCIAS

- AGUILAR, Mariana Guerra. **Estudo fitoquímico do extrato hexânico e avaliação da atividade biológica de extratos e constituintes das folhas de *Maytenus quadrangulata* (Celastraceae)**. 2019. Dissertação (Mestrado em Química Orgânica) - Departamento de Química do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, Belo Horizonte, 2019.
- ALLEN, *et al.* DOCK 6: Impact of new features and current docking performance. **J Comput Chem**, v. 36, n. 15, p. 1132-1156, 2015.
- ALVES, V. M. *et al.* Quimioinformática: Uma Introdução. **Química Nova**, São Paulo, v. 41, n. 2, p. 202-212, 2018.
- AN, Q; FAN C. H.; XU S. M. Recent perspectives of pediatric leukemia - an update. **Eur Rev Med Pharmacol Sci**, v. 21, n. 4, p. 31-36, 2017.
- ARBER, D. A. *et al.* The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. **Blood**, v. 127, n. 20, p. 2391-2405, 2016.
- BAGARIA, A. *et al.* Estimating structure quality trends in the Protein Data Bank by equivalent resolution. **Computational Biology and Chemistry**, v. 46, p. 8-15, 2013.
- BARBOSA M. C. S. *et al.* Synthesis and evaluation of the mutagenicity of 3-alkylpyridine marine alkaloid analogues with anticancer potential. **Mutat Res Genet Toxicol Environ Mutagen**, v. 825, p. 31-39, 2018.
- BAVARO L. *et al.* Mechanisms of Disease Progression and Resistance to Tyrosine Kinase Inhibitor Therapy in Chronic Myeloid Leukemia: An Update. **Int J Mol Sci**, v. 20, n. 24, 6141, 2020.
- BEDNARCZYK-CWYNAR, B.; ZAPRUTKO L. Recent advances in synthesis and biological activity of triterpenic acylated oximes. **Phytochem Rev**, v. 14, n.2, p. 203–231, 2015.
- BENTO, A. P. *et al.* The ChEMBL bioactivity database: Anupdate. **Nucleic Acids Research**, v. 42, p. 1083-1090, 2014.
- BOUTARD, N. *et al.* Synthesis of amide and sulfonamide substituted N-aryl 6-aminoquinoxalines as PFKFB3 inhibitors with improved physicochemical properties. **Bioorganic & Medicinal Chemistry Letters**, v. 29, p. 646-653, 2019.
- BERTOLI, S. *et al.* CDC25A governs proliferation and differentiation of FLT3-ITD acute myeloid leukemia. **Oncotarget**, v. 6, n. 35, p. 38061-38078, 2015.
- BRENNER, K. A. *et al.* Therapeutic Targeting the Cell Division Cycle 25 (CDC25) Phosphatases in Human Acute Myeloid Leukemia-The Possibility to Target Several Kinases Through Inhibition of the Various CDC25 Isoforms. **Molecules**, v. 19, n. 11, p. 18414-18447, 2014.

BROOKE, G. D. *et al.* Targeting the Warburg Effect in cancer; relationships for 2-arylpriidazinones as inhibitors of the key glycolytic enzyme 6-phosphofructo-2-kinase/2,6-bisphosphatase 3 (PFKFB3). **Bioorganic & Medicinal Chemistry**, v. 22, n. 3, p. 1029-1039, 2014.

BULLINGER L.; DÖHNER K.; DÖHNER H. Genomics of Acute Myeloid Leukemia Diagnosis and Pathways. **Journal of Clinical Oncology**, v. 35, n. 9, p. 934-946, 2017.

BURNETT A.; STONE R. AML: New Drugs but New Challenges. **Clin Lymphoma Myeloma Leuk**, v. 20, n. 6, p. 341-350, 2020.

CARTA, G. *et al.* Unbiasing Scoring Functions: A New Normalization and Rescoring Strategy. **J ChemvInf Model**, v. 47, n. 4, p. 1564-1571, 2007.

CERETO-MASSAGUÉ, A. *et al.* Tools for in silico target fishing. **Methods**, v. 71, p. 98-103, 2015a.

CERETO-MASSAGUÉ, A. *et al.* Molecular fingerprint similarity search in virtual screening. **Methods**, v. 71, p. 58-63, 2015b.

CLARCK, D. E. *et al.* Application of virtual screening to the discovery of novel nicotinamide phosphoribosyltransferase (NAMPT) inhibitors with potential for the treatment of cancer and axonopathies. **Bioorganic & Medicinal Chemistry Letters**, v. 26, n. 12, p. 2920-2926, 2016.

CLEM, B. *et al.* Small-molecule inhibition of 6-phosphofructo-2-kinase activity suppresses glycolytic flux and tumor growth. **Mol Cancer Ther**, v. 7, n. 1, p. 110-120, 2008.

CULLY, M. *et al.* Beyond PTEN mutations: the PI3K pathway as an integrator of multiple inputs during tumorigenesis. **Nat Rev Cancer**, v. 6, n. 3, p. 184-192, 2006.

DAINA A.; MICHIELIN O.; ZOETE V. Swiss Target Prediction: updated data and new features for efficient prediction of protein targets of small molecules. **Nucleic Acids Research**, v. 47, n. 1, p. 357-364, 2019.

DANISHUDDIN; KHAN A. U. Descriptors and their selection methods in QSAR analysis: paradigm for drug design. **Drug Discov Today**, v. 21, n. 8, p. 1291-1302, 2016.

DELANO, W. L. The PyMOL Molecular Graphics System. **Delano Scientific**, Palo Alto, CA, USA, 2002. <http://www.pymol.org>.

DIAS R.; DE AZEVEDO W. F. Jr. Molecular docking algorithms. **Curr Drug Targets**, v. 9, n. 12, p. 1040-1047, 2008.

DRAGOVICH, P. S. *et al.* Fragment-based design of 3-aminopyridine-derived amides as potent inhibitors of human nicotinamide phosphoribosyltransferase (NAMPT). **Bioorganic & Medicinal Chemistry Letters**, v. 24, n. 3, p. 954-962, 2014.

DURANT, J. L.; LELAND B. A.; HENRY D. R.; NOURSE J. G. Reoptimization of MDL keys for use in drug discovery. **J Chem Inf Comput Sci**, v. 42, p. 1273-1280, 2002.

ESTEY, E. Acute myeloid leukemia: 2016 Update on risk-stratification and management. **American Journal of Hematology**, v. 91, n. 8, p. 825-846, 2016.

FACHI, M. M. *et al.* Comparative efficacy and safety of tyrosine kinase inhibitors for chronic myeloid leukaemia: A systematic review and network meta-analysis. **European Journal of Cancer**, v. 104, p. 9-20, 2018.

FERNÁNDEZ-DE GORTARI, E. *et al.* Database fingerprint (DFP): an approach to represent molecular databases. **Journal of Chemoinformatics**, v. 9, 2017.

FORLI, S. *et al.* Computational protein-ligand docking and virtual drug screening with the AutoDock suite. **Nat Protoc**, v. 11, n. 5, p 905-919, 2016.

GIANNETTI, A. M. *et al.* Fragment-based identification of amides derived from trans-2-(pyridin-3-yl)cyclopropanecarboxylic acid as potent inhibitors of human nicotinamide phosphoribosyltransferase (NAMPT). **J Med Chem**, v. 57, n. 3, p. 770-792, 2014.

GILSON, M. K. *et al.* BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology. **Nucleic Acids Research**, v. 44, n. 1, p. 1045-1053, 2016.

GONG, J. *et al.* ChemMapper: a Versatile Web Server for Exploring Pharmacology and Chemical Structure Association Based on Molecular 3D Similarity Method. **Bioinformatics**, v. 29, n. 14, p. 1827-1829, 2013.

HALGREN, T. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. **J Med Chem**, v. 47, n. 7, p. 1750-1759, 2004.

HANAHAHAN, D.; WEINBERG, R. A. Hallmarks of Cancer: The Next Generation, **Cell**, v. 144, n. 5, p 646-674, 2011.

HARRINGTON, K. J. Biology of cancer. **Cancer biology and imaging**, v. 39, n. 12, p. 689-692, 2011.

HASSERJIAN, R. P. Acute myeloid leukemia: advances in diagnosis and classification. **International Journal of Laboratory Hematology**, v. 35, n. 3, p. 358-366, 2013.

HENTABLI, H. *et al.* A New Graph-Based Molecular Descriptor Using the Canonical Representation of the Molecule. **Scientific World Journal**, 2014.

IMPERATORE, C. *et al.* Alkaloids from Marine Invertebrates as Important Leads for Anticancer Drugs Discovery and Development. **Molecules**, v. 19, n. 12, p. 20391–20423, 2014.

ITO T. Stem cell maintenance and disease progression in chronic myeloid leukemia. **Int J Hematol**, v. 98, n. 6, p. 641-647, 2013.

JACOBSSON, M.; KARLÉN, A. Ligand Bias of Scoring Functions in Structure-Based Virtual Screening. **Journal of Chemical Information and Modeling**, v. 46, n. 3, p. 1334-1343, 2006.

- JAIN, A. N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. **J Med Chem**, v. 46, n. 4, p. 499-511, 2003.
- JONES, G. *et al.* Development and Validation of a Genetic Algorithm for Flexible Docking. **Journal of Molecular Biology**, v. 267, n. 3, p. 727-748, 1997.
- KANG, X. *et al.* Synthesis, anti-proliferative and proapoptotic activity of novel oleanolic acid azaheterocyclic derivatives. **Med Chem Commun**, v.3, p. 1245, 2012.
- KANTARJIAN, H. M.; KEATING, M. J.; FREIREICH, E. J. Toward the Potential Cure of Leukemias in the Next Decade. **Cancer**, v. 124, n. 22, p. 4301-4313, 2018.
- KAYSER S, LEVIS M. J. Advances in targeted therapy for acute myeloid leukaemia. **Br J Haematol**, v. 180, n. 4, p-484-500, 2018.
- KLEYWEGT, G. J.; JONES, T. A. Model Building and Refinement Practice, **Methods in Enzymology**, v. 277, p. 208-230, 1997.
- KOENIG, K. *et al.* The Changing Landscape of Treatment in Acute Myeloid Leukemia. **American Society of Clinical Oncology Educational**, v. 40, p. 1-12, 2020.
- KOSCHMIEDER, S.; VETRIE, D. Epigenetic dysregulation in chronic myeloid leukaemia: A myriad of mechanisms and therapeutic options. **Seminars in Cancer Biology**, v. 51, p. 180-197, 2018.
- KOSTIDIS, S. *et al.* Quantitative NMR analysis of intra- and extracellular metabolism of mammalian cells: A tutorial. **Analytica Chimica Acta**, v. 980, p. 1-24, 2017.
- KOUTSOUKAS, A. *et al.* From in silico target prediction to multi-target drug design: Current databases, methods and applications. **Journal of Proteomics**, v. 74, p. 2554-2574, 2011.
- KOUCHKOVSKY, I.; ABDUL-HAY, M. Acute myeloid leukemia: a comprehensive review and 2016 update. **Blood Cancer Journal**, v. 6, n. 7, e441, 2016.
- KUMAR, P. *et al.* Betulinic acid as apoptosis activator: Molecular mechanisms, mathematical modeling and chemical modifications. **Life Sciences**, v. 209, p. 24-33, 2018.
- LAVECCHIA, A.; CERCHIA, C. In silico methods to address polypharmacology: current status, applications and future perspectives. **Drug Discovery Today**, v. 21, n. 2, p. 288-298, 2018.
- LIMA, A. N. *et al.* Use of machine learning approaches for novel drug discovery. **Expert Opinion on Drug Discovery**, v. 11, n. 3, p. 225-239, 2016.
- LIN, M.; CHEN, B. Advances in the drug therapies of acute myeloid leukemia (except acute promyelocytic leukemia). **Drug Design, Development and Therapy**, v. 12, p. 1009-1017, 2018.
- LIN, X. *et al.* A Review on Applications of Computational Methods in Drug Screening and Design. **Molecules**, v. 25, n. 6, 1375, 2020.

- LIU, R. *et al.* PI3K/Akt pathway as a key link modulates the multidrug resistance of cancers. **Cell Death and Disease**, v. 11, n. 9, 797, 2020.
- LO, H. Y. *et al.* Novel acyclic enediynes inhibit Cyclin A and Cdc25C expression and induce apoptosis phenomenon to show potent antitumor proliferation. **Bioorganic & Medicinal Chemistry**, v. 15, p. 4528–4536, 2007.
- LONG, S. *et al.* Antitumor Activity of Quinazolinone Alkaloids Inspired by Marine Natural Products. **Marine Drugs**, v. 16, n. 8, 261, 2018.
- LOSCOCCO, F. *et al.* BCR-ABL Independent Mechanisms of Resistance in Chronic Myeloid Leukemia. **Frontiers in Oncology**, v. 9, 939, 2019.
- LUND, G. *et al.* Inhibition of CDC25B Phosphatase Through Disruption of Protein-Protein Interaction. **ACS Chemical Biology**, v. 10, n. 2, p. 390-394, 2015.
- MAGGIORA, G. *et al.* Molecular Similarity in Medicinal Chemistry. **Journal of Medicinal Chemistry**, v. 57, n. 8, p. 3186-3204, 2014.
- MAJEED, R. *et al.* Synthesis of 3-O-propargylated betulinic acid and its 1,2,3-triazoles as potential apoptotic agentes. **European Journal of Medicinal Chemistry**, v. 63, p. 782-792, 2013.
- MALO, N. *et al.* Statistical practice in high-throughput screening data analysis. **Nature Biotechnology**, v. 24, n. 2, p. 167-175, 2006.
- MALTAROLLO, Vinícius Gonçalves. **Aplicação de estratégias in silico para o desenvolvimento de ligantes com afinidade pelo receptor PPAR δ** . Orientador: Káthia Maria Honório. 2013. 145 f. Tese (Doutorado em Ciência e Tecnologia/Química) - Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André, 2013.
- MARTELLI, M. P. *et al.* Enasidenib and ivosidenib in AML. **Minerva Med**, v. 111, n. 5, p. 411-426, 2020.
- MEHBUB, M. F. *et al.* Marine sponge derived natural products between 2001 and 2010: trends and opportunities for discovery of bioactives. **Marine Drugs**, v. 12, n. 8, p. 4539-4577, 2014.
- MENG, X. Y. *et al.* Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. **Current Computer-Aided Drug Design**, v. 7, n. 2, p. 146-157, 2011.
- MERDINGER, M. *et al.* Novel Prognostic and Therapeutic Mutations in Acute Myeloid Leukemia. **Cancer Genomics and Proteomics**, v. 13, n. 5, p. 317-329, 2016.
- MITCHELL, R. S. *et al.* Selective targeting of NAMPT by KPT-9274 in acute myeloid leukemia. **Blood Advances**, v. 3, n. 3, p. 242-255, 2019.
- NICKEL, J. *et al.* SuperPred: update on drug classification and target prediction. **Nucleic Acids Res**, v. 42 (Web Server issue), W26-31, 2014.

RADICH, J. P.; MAURO M. J. Tyrosine Kinase Inhibitor Treatment for Newly Diagnosed Chronic Myeloid Leukemia. **Hematology Oncology Clinics**, v. 31, n. 4, p. 577-587, 2017.

RAREY, M. *et al.* A fast flexible docking method using an incremental construction algorithm. **J Mol Biol**, v. 261, n.3, p. 470-489, 1996.

RASHID, S. *et al.* Synthesis and biological evaluation of ursolic acid-triazolyl derivatives as potential anti-cancer agentes. **European Journal of Medicinal Chemistry**, v. 66, p. 233-245, 2013.

SERAFIM, M. S. M. *et al.* Antibacterial activity of synthetic 1,3-bis(aryloxy)propan-2amines against Gram-positive bacteria. **Microbiologyopen**, v. 8, n. 11, e814, 2019.

SCHUFFENHAUER, A. *et al.* Similarity metrics for ligands reflecting the similarity of the target proteins. **Journal of Chemical Information and Computer Sciences**, v. 43, n. 2, p. 391-405, 2003.

SIVEEN, S. K. *et al.* Targeting acute myeloid leukemia stem cell signaling by natural products. **Molecular Cancer**, v. 16, 13, 2017.

STEIN E. M., TALLMAN M. S. Emerging therapeutic drugs for AML. **Blood**, v. 127, n. 1, p71-78, 2016.

SUEUR, G. *et al.* STAT5-dependent regulation of CDC25A by miR-16 controls proliferation and differentiation in FLT3-ITD acute myeloid leukemia. **Sci Rep**, v. 10, n. 1, p. 1906, 2020.

SONG, T. Y. *et al.* Improvement of treatment outcome over 2 decades in children with acute myeloid leukemia. **Blood Research**, v. 53, n. 1, p. 25-34, 2018.

TALATI C.; SWEET K. Recently approved therapies in acute myeloid leukemia: A complex treatment landscape. **Leukemia Research**, v. 73, p. 58-66, 2018.

TALIB, W. H. Melatonin and Cancer Hallmark. **Molecules**, v. 23, n. 3, p. 518, 2018.

TAMAMYAN, G. *et al.* Frontline treatment of acute myeloid leukemia in adults. **Critical Reviews in Oncology/Hematology**, v. 110, p. 20-34, 2017.

TAN, Y. *et al.* Targeting acute myeloid leukemia stem cells: Current therapies in development and potential strategies with new dimensions. **Crit Rev Oncol Hematol**, v. 152, 102993, 2020.

TAKEUCHI, M. *et al.* Discovery of a Novel Nicotinamide Phosphoribosyl Transferase (NAMPT) Inhibitor via in Silico Screening. **Biol Pharm Bull**, v. 37, n. 1, p. 31-36, 2014.

TONHOLO, D. R. *et al.* Preclinical toxicity of innovative molecules: In vitro, in vivo and metabolism prediction. **Chem Biol Interact**, v. 315, 108896, 2020.

URBAN M. *et al.* Synthesis of A-Seco Derivatives of Betulinic Acid with Cytotoxic Activity. **J Nat Prod**, v. 67, p. 1100-1105, 2004.

VALLONE, A. *et al.* Antimalarial agentes against both sexual and asexual parasites stages: Structure-activity relationships and biological studies of the Malaria Box compound 1-[5-(4-bromo-2-chlorophenyl)furan-2-yl]-N-[(piperidin-4-yl)methyl]methanamine (MMV019918) and analogues. **European Journal of Medicinal Chemistry**, v. 150, p. 698-718, 2018.

VERDONK, M. L. *et al.* Improved protein–ligand docking using GOLD. **Proteins**, v. 52, n. 4, p. 609-623, 2003.

VILLAR, H. V. *et al.* Down-Regulation of Akt Signalling by Ursolic Acid Induces Intrinsic Apoptosis and Sensitization to Doxorubicin in Soft Tissue Sarcoma. **PLoS One**, v. 11, n. 5, e0155946, 2016.

VIGNOLI, A. *et al.* High-Throughput Metabolomics by 1D NMR. **Angew Chem Int Ed Engl**, v. 58, n. 4, p. 968-994, 2019.

WU, B. *et al.* Ursolic Acid-induced Apoptosis in K562 Cells Involving Upregulation of PTEN Gene Expression and Inactivation of the PI3K/Akt Pathway. **Arch Pharm Res**, v. 35, n. 3, p. 543-548, 2012.

YAP, C. W. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. **J Comput Chem**, v. 32, n. 7, p. 1466-1474, 2011.

YOU, H. *et al.* Design, synthesis and X-ray crystallographic study of NAMPTase inhibitors as anti-cancer agents. **European Journal of Medicinal Chemistry**, v. 46, n. 4, p. 1153-1164, 2011.

YUAN, X.; XU, Y. Recent Trends and Applications of Molecular Modeling in GPCR–Ligand Recognition and Structure-Based Drug Design. **Int J Mol Sci**, v. 19, n. 7, 2105, 2018.

YU, J. *et al.* Clinical implications of recurrent gene mutations in acute myeloid leukemia. **Exp Hematol Oncol**, v. 9, n. 4, 2020.

YURTTAŞ, O. N.; EŞKAZAN A. E. Novel therapeutic approaches in chronic myeloid leukemia. **Leukemia Research**, v. 91, 106337, 2020.

ZHANG, K. *et al.* Discovery of trans-3-(pyridin-3-yl)acrylamide-derived sulfamides as potent nicotinamide phosphoribosyltransferase (NAMPT) inhibitors for the potential treatment of cancer. **Bioorg Med Chem Lett**, v. 29, n. 12, p. 1502-1506, 2019.

ŽIBERNA, L. *et al.* Oleanolic Acid Alters Multiple Cell Signaling Pathways: Implication in Cancer Prevention and Therapy. **Int J Mol Sci**, v. 18, n. 3, 643, 2017.

ANEXO I - Compostos ChEMBL com Tc \geq 0,75 em relação aos compostos de interesse.

Compostos de Interesse	Compostos ChEMBL (Tc)^a	Compostos ChEMBL (Tc)^b	Compostos ChEMBL (Tc)^c	Compostos ChEMBL (Tc)^d	Compostos ChEMBL (Tc)^e	Compostos ChEMBL (Tc)^e
1	CHEMBL169 (0,79) CHEMBL269277 (0,77)	CHEMBL269277 (0,78) CHEMBL169 (0,78)	CHEMBL3426505 (0,83) CHEMBL1651318 (0,81) CHEMBL205518 (0,81) CHEMBL3426487 (0,79) CHEMBL168 (0,79) CHEMBL2088122 (0,79) CHEMBL462795(0,78) CHEMBL269277 (0,77) CHEMBL224814 (0,77) CHEMBL2088123 (0,76) CHEMBL3311279 (0,76) CHEMBL1651317 (0,75) CHEMBL1915760 (0,75) CHEMBL295602 (0,75)	CHEMBL2088122 (0,78) CHEMBL168 (0,78) CHEMBL269277 (0,78) CHEMBL295602 (0,77) CHEMBL194524 (0,76) CHEMBL388775 (0,75)	CHEMBL400000 (0,83) CHEMBL269277 (0,77) CHEMBL398831 (0,77) CHEMBL251244 (0,76) CHEMBL517756 (0,76)	CHEMBL458696 (0,78) CHEMBL269277 (0,78) CHEMBL456957 (0,77)

2	CHEMBL269277 (0,79) CHEMBL169 (0,76)	-	CHEMBL3311279 (0,82) CHEMBL3426505 (0,80) CHEMBL462795 (0,80) CHEMBL269277 (0,79) CHEMBL205518 (0,78) CHEMBL1644101 (0,77) CHEMBL1957345 (0,77) CHEMBL295602 (0,77) CHEMBL3426487 (0,76) CHEMBL3426496 (0,76) CHEMBL431525 (0,76) CHEMBL168 (0,76) CHEMBL375698 (0,76) CHEMBL1165514 (0,76) CHEMBL3311277 (0,76) CHEMBL179443 (0,76) CHEMBL1774188 (0,76) CHEMBL2088124 (0,75)	CHEMBL462795 (0,79) CHEMBL205518 (0,78) CHEMBL3311278 (0,76) CHEMBL3311279 (0,76)	CHEMBL269277 (0,79) CHEMBL517756 (0,78) CHEMBL1978954 (0,76)	-
---	--	---	---	--	--	---

3	CHEMBL169 (0,8) CHEMBL269277 (0,79)	-	CHEMBL205518 (0,89) CHEMBL462795 (0,86) CHEMBL2088123 (0,84) CHEMBL2063136 (0,82) CHEMBL3814944 (0,82) CHEMBL168 (0,82) CHEMBL375698 (0,82) CHEMBL2088122 (0,82) CHEMBL2063137 (0,79) CHEMBL269277 (0,79) CHEMBL224814 (0,79) CHEMBL3311279 (0,78) CHEMBL1915760 (0,78) CHEMBL3426505 (0,76) CHEMBL194309 (0,76) CHEMBL239075 (0,76) CHEMBL239077 (0,76) CHEMBL1651319 (0,76) CHEMBL2088124 (0,76)	CHEMBL462795 (0,80) CHEMBL205518 (0,79) CHEMBL3311278 (0,77) CHEMBL257595 (0,77) CHEMBL2088123 (0,77) CHEMBL3311279 (0,77) CHEMBL375698 (0,75)	CHEMBL1627439 (0,81) CHEMBL269277 (0,79) CHEMBL1978954 (0,76) CHEMBL400000 (0,76) CHEMBL398831 (0,75)	CHEMBL456958 (0,79)
4	-	-	-	-	-	CHEMBL225920 (0,78) CHEMBL389183 (0,75) CHEMBL225697 (0,75)
5	-	CHEMBL494073 (0,78) CHEMBL494674 (0,76)	-	CHEMBL2041096 (0,77) CHEMBL2041094 (0,75)	-	CHEMBL225920 (0,86) CHEMBL389183 (0,83)

						CHEMBL225697 (0,83)
6	-	CHEMBL494073 (0,77) CHEMBL494674 (0,75)	-	CHEMBL2041096 (0,78) CHEMBL2041094 (0,75) CHEMBL2041088 (0,75)	-	CHEMBL225920 (0,85) CHEMBL389183 (0,82) CHEMBL225697 (0,82)
7	-	CHEMBL371474 (0,77) CHEMBL414666 (0,76) CHEMBL363390 (0,76)	-	CHEMBL3105848 (0,77) CHEMBL566757 (0,75)	-	CHEMBL225920 (0,77) CHEMBL479945 (0,75) CHEMBL389183 (0,75) CHEMBL225697 (0,75)
8	-	-	-	CHEMBL2397593 (0,76)	-	CHEMBL72739 (0,80) CHEMBL387746 (0,79) CHEMBL225920 (0,78) CHEMBL389183 (0,76) CHEMBL225697 (0,76) CHEMBL227023 (0,75)
9	-	CHEMBL494073 (0,75)	-	CHEMBL231041 (0,75)	-	CHEMBL225920 (0,79) CHEMBL389183 (0,79) CHEMBL225697 (0,79) CHEMBL1988584 (0,75)

Legenda: a) Tc calculado entre os compostos de interesse e os compostos resultantes da pesquisa ChEMBL para a linhagem THP-1, CI50, *fingerprint* MACCS. b) Tc calculado entre os compostos de interesse e os compostos resultantes da pesquisa ChEMBL para a linhagem THP-1, CI50, *fingerprint* Pubchem. c) Tc calculado entre os compostos de interesse e os compostos resultantes da pesquisa ChEMBL para a linhagem K562, CI50, *fingerprint* MACCS. d) Tc calculado entre os compostos de interesse e os compostos resultantes da pesquisa ChEMBL para a linhagem K562, CI50, *fingerprint* Pubchem. e) Tc calculado entre os compostos de interesse e os compostos resultantes da pesquisa ChEMBL para a linhagem K562, GI50, *fingerprint* MACCS. f) Tc calculado entre os compostos de interesse e os compostos

resultantes da pesquisa ChEMBL para a linhagem K562, GI50, *fingerprint* Pubchem. Nenhuma molécula obtida na busca no ChEMBL com os filtros: linhagem THP-1 e GI50, apresentou $T_c \geq 0,75$.

ANEXO II – Estruturas cristalográficas depositadas no PDB para cada alvo proteico estudado.

Alvo proteico	Código PDB	Resolução (Å)	R_{free}	RMSD sobreposição*
PFKFB3	6HVI	1,96	0,208	-
NAMPT	6E68	1,50	0,167	0,132
	4LVF	1,50	0,172	-
	4LVA	1,55	0,207	0,116
	4O10	1,55	0,179	0,100
	6TA0	1,58	0,177	0,108
	4KFN	1,60	0,183	0,102
	4KFO	1,60	0,181	0,079
	5KIT	1,60	0,180	0,118
	4LWW	1,64	0,179	0,077
	5UPF	1,69	0,241	0,141
	4LTS	1,69	0,190	0,111
	4N9D	1,70	0,183	0,103
	4LVG	1,70	0,171	0,089
	4O1D	1,71	0,208	0,150
	4N9E	1,72	0,189	0,093
	4WQ6	1,72	0,188	0,122
	5U2N	1,73	0,192	0,123
	4M6P	1,75	0,241	0,114
	4LVD	1,75	0,213	0,190
	4O13	1,75	0,227	0,102
4N9C	1,75	0,201	0,126	

NAMPT	3DHF	1,80	0,196	0,194
	4LV9	1,81	0,213	0,148
	4LVB	1,84	0,215	0,172
	4KFP	1,84	0,171	0,081
	5LX5	1,88	0,196	0,173
	3DKL	1,89	0,214	0,184
	5U2M	1,89	0,220	0,131
	4JR5	1,91	0,210	0,126
	5UPE	1,93	0,217	0,142
	5WI1	1,99	0,227	0,175
CDC25B	4WH9	1,50	0,152	-
	4WH7	1,62	0,198	0,099

Legenda: Em destaque estão as proteínas selecionadas para o acoplamento molecular. *RMSD calculado entre as estruturas cristalográficas depositadas no PDB e a estrutura cristalográfica da proteína selecionada para o acoplamento molecular.