



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE CIÊNCIAS AGRÁRIAS**  
**DEPARTAMENTO DE ENGENHARIA DE PESCA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA DE RECURSOS**  
**NATURAIS**

**DISRAELI CAVALCANTE ARAUJO VASCONCELOS**

**MODELAGEM COMPUTACIONAL DAS INTERAÇÕES DA INTEGRINA VLA-4**  
**HUMANA COM LIGANTES COMO FERRAMENTA AO PLANEJAMENTO DE**  
**FÁRMACOS**

**FORTALEZA**

**2023**

DISRAELI CAVALCANTE ARAUJO VASCONCELOS

MODELAGEM COMPUTACIONAL DAS INTERAÇÕES DA INTEGRINA VLA-4  
HUMANA COM LIGANTES COMO FERRAMENTA AO PLANEJAMENTO DE  
FÁRMACOS

Tese apresentada ao Programa de Pós-Graduação em Biotecnologia de Recursos Naturais da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Ciências. Área de concentração: Biotecnologia em Recursos Naturais.

Orientador: Prof. Dr. João Herminio Martins da Silva

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

- V45m Vasconcelos, Disraeli Cavalcante Araujo.  
Modelagem computacional das interações da integrina VLA-4 humana com ligantes como ferramenta ao planejamento de fármacos / Disraeli Cavalcante Araujo Vasconcelos. – 2023.  
113 f. : il. color.
- Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências Agrárias, Programa de Pós-Graduação em Biotecnologia de Recursos Naturais, Fortaleza, 2023.  
Orientação: Prof. Dr. João Herminio Martins da Silva.
1. integrina. 2. peptidomimético. 3. BIO-5192. 4. reposicionamento de fármacos. I. Título.  
CDD 660.6
-

DISRAELI CAVALCANTE ARAUJO VASCONCELOS

MODELAGEM COMPUTACIONAL DAS INTERAÇÕES DA INTEGRINA VLA-4  
HUMANA COM LIGANTES COMO FERRAMENTA AO PLANEJAMENTO DE  
FÁRMACOS

Tese apresentada ao Programa de Pós-Graduação em Biotecnologia de Recursos Naturais da Universidade Federal do Ceará, como requisito parcial à obtenção do título de doutor em Ciências. Área de concentração: Biotecnologia em Recursos Naturais.

Aprovada em: 26/07/2023.

BANCA EXAMINADORA

---

Prof. Dr. João Herminio Martins da Silva (Orientador)  
Fundação Oswaldo Cruz (Fiocruz/CE)

---

Prof. Dr. Bruno Anderson Matias da Rocha  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Bruno Lopes de Sousa  
Universidade Estadual do Ceará (UECE)

---

Prof. Dr. Fernando Berton Zanchi  
Fundação Oswaldo Cruz (Fiocruz/RO)

---

Prof. Dr. José Hélio Costa  
Universidade Federal do Ceará (UFC)

A Deus.

Aos meus pais, Marta e Valdemir.

## AGRADECIMENTOS

Ao Divino, ao meu Anjo da Guarda, pelo amparo espiritual, nas horas leves e nas mais atribuladas da vida.

Aos meus pais e à Lúcia, minha segunda mãe, pelo apoio financeiro e emocional durante todos os meus anos de estudo.

Aos meus avós, Aloisio (*in memoriam*), Izaura (*in memoriam*) e Margarida (*in memoriam*), por terem acreditado, incessantemente, em meu potencial para seguir em frente nos estudos.

Aos meus irmãos de sangue, em especial ao Marcus, a quem me espelho como ser humano e profissional, e à irmã do coração Bárbara, pelas conversas “jogadas fora” que foram importantes para aliviar as tensões do dia a dia.

Aos meus queridos alunos que fizeram parte da minha vida a partir de fevereiro de 2022. Vocês são minha endorfina e loucura diárias, porém sem vocês eu não seria o profissional que me tornei hoje.

Aos meus amigos de UFC e Fiocruz/CE por todo o suporte no trabalho, pelos momentos de diversão e pelas reflexões divagadas acerca da vida, de um modo geral.

Aos meus amigos da EQ, por terem feito me sentir em casa quando lá cheguei, em fevereiro de 2022. Nossos momentos de descontração são exemplos de nossa perfeita simbiose no trabalho em grupo.

Ao meu orientador, Dr. João Martins, pela convivência, pela orientação e pelo aprendizado adquirido durante todos esses anos de trabalho, desde que entrei para a Fiocruz/CE, em 2013.

Ao nosso pós-doc, Dr. Geraldo Sartori, por todos os ensinamentos a mim passados e pela ajuda durante o processo de confecção deste trabalho.

Aos membros da banca, por terem aceitado o convite e pelas valiosas contribuições acerca do trabalho.

À coordenação do Programa de Pós-Graduação em Biotecnologia de Recursos Naturais, por todo o suporte administrativo a mim concedido e pelas céleres resoluções dos problemas que, porventura, apareciam.

À Funcap, CAPES e CNPq pelo apoio financeiro com a manutenção da bolsa durante um período do doutorado. Esse trabalho foi parcialmente fomentado através do Edital Universal 2014 do CNPq, processo 457937/2014-9.

“O que prevemos raramente ocorre; o que menos esperamos geralmente acontece.”  
(BENJAMIN DISRAELI).

## RESUMO

As integrinas são receptores transmembrana que desempenham um papel crítico em muitos processos biológicos de migração celular, os quais podem ser moduladas terapeuticamente usando antagonistas de integrinas, como ligantes peptidomiméticos. Este trabalho teve como objetivo o estudo das interações a nível atômico-molecular de três ligantes (1592, 2363 e 973) inéditos derivados do composto-líder BIO-5192 e que são potenciais inibidores da integrina  $\alpha4\beta1$  (VLA-4), assim como o de fármacos comerciais, inicialmente, utilizados para outra finalidade. Para tanto, foram analisados os modos de ligação dos compostos em VLA-4 por *docking* molecular e, posteriormente, foi selecionado um em específico e submetido às simulações de dinâmica molecular e aos cálculos de energia de livre por MM/PBSA, a fim de se avaliar o ligante de maior afinidade ao receptor. Análises por PCA também foram aplicadas à proteína alvo, para observar seus modos de movimento quando da complexação com os diferentes compostos. Os resultados revelaram que o ligante 1592 apresentou um comportamento estável no sítio de interação de VLA-4, a partir de um RMSD médio próximo a 0,5 nm. As análises de PCA também sugeriram que o receptor proteico apresentou modos de movimento diferentes em sua estrutura quando complexada a 1592, em comparação aos demais compostos. Além disso, através dos cálculos de energia livre de ligação, 1592 apresentou maior afinidade ao alvo em relação às outras moléculas estudadas, através do valor de  $\Delta G_{\text{ligação}}$  final próximo a -143 kJ/mol. Portanto, o ligante 1592 representa um possível candidato a um fármaco inibidor de VLA-4. A partir de métodos de triagem virtual, dinâmica molecular e MM/PBSA aplicados aos fármacos comerciais e, por meio de análises de estabilidade e afinidade ao receptor, o fármaco imatinib representou um potencial candidato a antagonista da integrina VLA-4, com um RMSD médio próximo a 0,2 nm e um  $\Delta G_{\text{ligação}}$  final igual a aproximadamente -140 kJ/mol. Logo, os resultados apresentados lançam novas perspectivas ao estudo de diferentes antagonistas de  $\alpha4\beta1$ , abrindo possibilidades ao reposicionamento de fármacos como ferramenta ao planejamento de pequenas moléculas.

**Palavras-chave:** integrina; peptidomimético; BIO-5192; reposicionamento de fármacos.

## ABSTRACT

Integrins are transmembrane receptors that play a critical role in many biological processes of cell migration, which can be therapeutically modulated using integrin antagonists such as peptidomimetic ligands. This work aimed to study the interactions at the atomic-molecular level of three unprecedented ligands (1592, 2363, and 973) derived from the lead compound BIO-5192 and which are potential inhibitors of  $\alpha 4\beta 1$  integrin (VLA-4), as well as the commercial drugs, initially used for another purpose. For this goal, the binding modes of the compounds were analyzed and, subsequently, a specific one was selected and submitted to molecular dynamics simulations and free energy calculations by MM/PBSA, in order to evaluate the ligand with the highest affinity to the receptor. PCA analyses were also applied to the target protein, to verify its motion modes when complexed with different compounds. The results revealed that the 1592 ligand presented a stable behavior at the VLA-4 interaction site, from an average RMSD close to 0.5 nm. PCA analyses also suggested that the protein receptor had different motion modes in its structure when complexed to 1592, compared to the other compounds. In addition, through the binding free energy calculations, 1592 showed higher affinity to the target in relation to the other studied molecules, through the final  $\Delta G_{\text{binding}}$  value close to -143 kJ/mol. Therefore, the 1592 ligand represents a possible candidate for a VLA-4 inhibitor drug. Based on virtual screening, molecular dynamics and MM/PBSA methods applied to the commercial drugs and, through analysis of stability and affinity to the receptor, the drug imatinib represented a potential candidate for VLA-4 integrin antagonist, with an average RMSD close to 0.2 nm and a final  $\Delta G_{\text{binding}}$  equals to approximately -140 kJ/mol. Therefore, the presented results launch new perspectives for the study of different  $\alpha 4\beta 1$  antagonists, opening possibilities for drug repurposing as a tool for drug design.

**Keywords:** integrin; peptidomimetic; BIO-5192; drug repurposing.

## LISTA DE FIGURAS

Figura 1 – Representação esquemática da estrutura de uma integrina, indicando cada um de seus domínios .....	22
Figura 2 – Representação esquemática do processo de biossinalização em integrinas .....	23
Figura 3 – Representação estrutural modificada do modelo gerado por homologia da integrina $\alpha 4\beta 1$ , mostrando, em vermelho, o íon $Mg^{2+}$ da região MIDAS .....	25
Figura 4 – a) Modelo por homologia de $\alpha 4\beta 1$ obtido pelo programa AlphaFold; b) a cadeia $\alpha$ da integrina onde reside o domínio $\beta$ -propeller; c) cadeia $\beta$ da integrina onde está o íon divalente MIDAS (Kozmon, S. Homology model of <i>alfa4beta1</i> integrins. 2022; unpublished results) .....	27
Figura 5 – Estruturas moleculares de PUPA (acima) e de BIO-5192 (abaixo) .....	33
Figura 6 – Equação da energia potencial calculada nos campos de força, a partir de resolução de termos simples .....	37
Figura 7 – Esquema representativo do processo de <i>ensemble docking</i> .....	38
Figura 8 – Na parte superior, a molécula BIO-5192, a qual foi o composto-líder utilizado neste trabalho para gerar, através da técnica do desenho <i>de novo</i> , os outros ligantes de código 1592, 2363 e 973, que estão na parte inferior da figura.....	39
Figura 9 – Ciclo termodinâmico do MM/PBSA .....	45
Figura 10- Esquema proposto para a validação do protocolo de VS no método de reposicionamento de fármacos na integrina VLA-4 .....	50
Figura 11- Modos de ligação de BIO-5192 em VLA-4 .....	52
Figura 12- Modos de ligação dos ligantes 1592 (azul), 2363 (vermelho) e 973 (preto) em VLA-4, em comparação com o líder BIO-5192 (amarelo) .....	53
Figura 13- Diagrama de interações resultantes do <i>docking</i> molecular dos ligantes 1592, 2363 e 973 em VLA-4 .....	54
Figura 14- a) Curva ROC do protocolo de <i>docking</i> e os parâmetros calculados AUC e Fator de Enriquecimento para os primeiros 1% e 20% da base de dados de	

ligantes de VLA-4. b) Modos de ligação dos compostos mais bem classificados selecionados no banco de dados ChEMBL .....	56
Figura 15- a) Interpolação linear dos dados de RMSD dos átomos pesados em três réplicas de 500 ns, para cada um dos ligantes: 1592 (preto) (Média = 0,4916 nm e DP = 0,1819 nm); 2363 (vermelho) (Média = 0,6826 nm e DP = 0,1332 nm); 973 (azul) (Média = 0,9845 nm e DP = 0,3165 nm) e BIO-5192 (verde) (Média = 0,5416 nm e DP = 0,1364 nm). b) Dados de RMSD dos átomos pesados do ligante 973, para 500 ns de simulação, conforme réplicas indicadas em preto (Réplica 1), vermelho (Réplica 2) e azul (Réplica 3) .....	58
Figura 16- a) Interpolação linear de dados da distância entre o oxigênio carboxilato e a integrina VLA-4 em três réplicas de 500 ns, para cada um dos ligantes: 1592 (preto) (Média = 0,2177 nm e DP = 0,1002 nm); 2363 (vermelho) (Média = 0,1898 nm e DP = 0,0058 nm); 973 (azul) (Média = 1,5739 nm e DP = 0,4118 nm) e BIO-5192 (verde) (Média = 0,1986 nm e DP = 0,0062 nm). b) Distâncias entre o oxigênio carboxilato do ligante 973 e o íon MIDAS Mg <sup>2+</sup> de VLA-4, ao longo de 500 ns de simulação, conforme as réplicas indicadas em preto (Réplica 1), azul (Réplica 2) e vermelho (Réplica 3) .....	60
Figura 17- Interpolação linear de dados do número de ligações de hidrogênio em três réplicas de 500 ns, para cada um dos ligantes: 1592 (preto) (Média = 3,2664 e DP = 1,570); 2363 (vermelho) (Média = 2,6382 e DP = 1,634); 973 (azul) (Média = 1,5256 e DP = 1,225) e BIO-5192 (verde) (Média = 1,6412 e DP = 1,124) .....	61
Figura 18- Interpolação linear de dados da área de superfície acessível ao solvente em três réplicas de 500 ns, para cada um dos ligantes: 1592 (preto) (Média = 9,2094 nm <sup>2</sup> e DP = 0,4821 nm <sup>2</sup> ); 2363 (vermelho) (Média = 9,2973 nm <sup>2</sup> e DP = 0,4417 nm <sup>2</sup> ); 973 (azul) (Média = 9,5556 nm <sup>2</sup> e DP = 0,7847 nm <sup>2</sup> ) e BIO-5192 (verde) (Média = 11,0479 nm <sup>2</sup> e DP = 0,3768 nm <sup>2</sup> ) .....	62
Figura 19- a) RMSF (do inglês, <i>Root Mean Square Fluctuation</i> ) e b) RMSD (do inglês, <i>Root Mean Square Deviation</i> ) do <i>backbone</i> de VLA-4 para cada sistema estudado .....	64
Figura 20- Análise de Componentes Principais (PCA) das 3 réplicas das simulações para	

cada sistema estudado de VLA-4 na forma apo e com os ligantes .....	65
Figura 21- Análise de Componentes Principais (PCA) das 3 réplicas das simulações para cada sistema estudado de VLA-4 na forma apo e com os ligantes 1592 e BIO-5192 – a) Resíduos do sítio de ligação de VLA-4 e b) Todos os resíduos de VLA-4 .....	66
Figura 22- Decomposições energéticas a partir de cálculos por MM/PBSA dos sistemas com VLA-4 complexada aos ligantes a) BIO-5192; b) 1592; c) 2363; d) 973 ..	68
Figura 23- Interpolação linear de dados dos potenciais: a) $\Delta G$ (BIO-5192 – Média = -64,85 kJ/mol e DP = 142,10 kJ/mol; 1592 – Média = -117,02 kJ/mol e DP = 41,14 kJ/mol; 2363 – Média = -70,37 kJ/mol e DP = 155,10 kJ/mol; 973 – Média = -86,94 kJ/mol e DP = 38,15 kJ/mol) ; b) Coulomb (BIO-5192 – Média = -460,25 kJ/mol e DP = 76,95 kJ/mol; 1592 – Média = -306,28 kJ/mol e DP = 87,17 kJ/mol; 2363 – Média = -1170,48 kJ/mol e DP = 419,5 kJ/mol; 973 – Média = -106,99 kJ/mol e DP = 67,15 kJ/mol; c) van der Waals (BIO-5192 – Média = -269,26 kJ/mol e DP = 19,83 kJ/mol; 1592 – Média = -194,87 kJ/mol e DP = 29,59 kJ/mol; 2363 – Média = -134,85 kJ/mol e DP = 24,77 kJ/mol; 973 – Média = -173,77 kJ/mol e DP = 56,73 kJ/mol) para os cálculos de MM/PBSA dos últimos 250 ns de simulação de 3 réplicas dos sistemas formados por VLA-4 e os ligantes estudados neste trabalho.....	70
Figura 24- Diagrama de interações de 5 compostos mais bem posicionados no <i>docking</i> do banco “fda” do ZINC em VLA-4 .....	76
Figura 25- Modos de ligação dos melhores (a) e dos piores (b) compostos com 80% de similaridade a BIO-5192 em VLA-4 .....	81
Figura 26- Diagrama de interações de 3 compostos mais bem posicionados no <i>docking</i> das moléculas com 80% de similaridade a BIO-5192 em VLA-4 .....	82
Figura 27- Curva ROC do protocolo de <i>docking</i> com o banco de moléculas com 80% de similaridade estrutural a BIO-5192 e os parâmetros calculados AUC e Fator de Enriquecimento para 1% e 20% de compostos inativos encontrados .....	85
Figura 28- Modos de ligação de alguns dos piores compostos do banco de moléculas que se ligam a VLA-4 .....	86

Figura 29- Diagrama de interações de 3 compostos mais bem posicionados no <i>docking</i> do banco de moléculas que possuem como alvo VLA-4 .....	87
Figura 30- RMSD (do inglês, <i>Root Mean Square Deviation</i> ) do <i>backbone</i> de VLA-4 para cada sistema estudado, em duas réplicas .....	90
Figura 31- RMSD (do inglês, <i>Root Mean Square Deviation</i> ) dos átomos pesados de a) imatinib; b) saquinavir; c) lumacaftor, em duas réplicas .....	91
Figura 32- Número de ligações de hidrogênio nas duas réplicas entre VLA-4 e os fármacos a) imatinib; b) saquinavir; c) lumacaftor .....	93
Figura 33- Interpolação linear de dados dos potenciais: a) $\Delta G$ (Imatinib – Média = -139,98 kJ/mol e DP = 18,33 kJ/mol; Lumacaftor – Média = -90,37 kJ/mol e DP = 44,57 kJ/mol; Saquinavir – Média = -97,13 kJ/mol e DP = 26,54 kJ/mol); b) Coulomb (Imatinib – Média = -87,79 kJ/mol e DP = 33,73 kJ/mol; Lumacaftor – Média = -38,60 kJ/mol e DP = 37,68 kJ/mol; Saquinavir – Média = -67,53 kJ/mol e DP = 39,03 kJ/mol); c) van der Waals (Imatinib – Média = -238,74 kJ/mol e DP = 14,40 kJ/mol; Lumacaftor – Média = -162,98 kJ/mol e DP = 33,67 kJ/mol; Saquinavir – Média = -206,81 kJ/mol e DP = 16,47 kJ/mol) para os cálculos de MM/PBSA dos últimos 50 ns de simulação de 2 réplicas dos sistemas formados por VLA-4 e os compostos estudados no reposicionamento de fármacos .....	95
Figura 34- Contribuições energéticas de G, Energia de Coulomb, Energia de van der Waals .....	97

## LISTA DE TABELAS

Tabela 1 – Alguns fármacos comerciais que foram submetidos ao processo de reposicionamento.....	49
Tabela 2 – Avaliação dos dez ligantes de VLA-4 melhor classificados de acordo com o Vina score e as atividades medidas pelos valores correspondentes de IC50 (Meia Concentração Inibitória Máxima).....	57
Tabela 3 – Valores finais de $\Delta G_{\text{ligação}}$ para os complexos de VLA-4 e os ligantes estudados, a partir dos cálculos por MM/PBSA de uma réplica de simulação..	69
Tabela 4 – Os dez fármacos com as menores energias calculadas no <i>docking</i> dos compostos do banco “fda” do Zinc em VLA-4, por ordem decrescente do Vina Score.....	75
Tabela 5 – As dez moléculas com as menores energias calculadas no <i>docking</i> dos compostos com 80% de similaridade a BIO-5192 em VLA-4, por ordem decrescente do Vina Score.....	79
Tabela 6 – As dez moléculas com as maiores energias calculadas no <i>docking</i> dos compostos com 80% de similaridade a BIO-5192 em VLA-4, por ordem decrescente do Vina Score.....	80
Tabela 7 – As dez moléculas com as maiores energias calculadas no <i>docking</i> dos compostos que tem como alvo VLA-4, por ordem decrescente do Vina Score.	86
Tabela 8 – Valores finais de $\Delta G_{\text{ligação}}$ para os complexos de VLA-4 e os fármacos estudados, a partir dos cálculos por MM/PBSA de uma réplica de simulação..	97

## LISTA DE ABREVIATURAS E SIGLAS

ADMET	Absorção, Distribuição, Metabolismo, Excreção e Toxicidade
AUC	<i>Area Under the Curve</i>
DII	Doença Inflamatória Intestinal
DM	Dinâmica Molecular
DP	Desvio-Padrão
EF	<i>Enrichment Factor</i>
FBDD	<i>Fragment-based Drug Discovery</i>
FDA	<i>Food and Drug Administration</i>
MEC	Matriz Extracelular
MIDAS	<i>Metal Ion-Dependent Adhesion Site</i>
MM/PBSA	<i>Molecular Mechanics/Poisson Boltzmann Surface Area</i>
PCA	<i>Principal Component Analysis</i>
PDB	<i>Protein Data Bank</i>
PME	<i>Particle Mesh Ewald</i>
RMSD	<i>Root Mean Square Deviation</i>
RMSF	<i>Root Mean Square Fluctuation</i>
ROC	<i>Receiver Operating Characteristics</i>
SyMBS	<i>Synergistic Metal Ion-Binding Site</i>
VLA-4	<i>Very Late Antigen-4</i>
VS	<i>Virtual Screening</i>
ZBB	<i>Zinc Building Blocks</i>

## LISTA DE SÍMBOLOS

Å	Ångstrom
$\alpha$	Alfa
$\beta$	Beta
$\Delta G$	Varição da Energia Livre de Gibbs
%	Porcentagem
°C	Grau Celsius

## SUMÁRIO

1	INTRODUÇÃO.....	19
1.1	Estrutura e função das integrinas.....	21
1.2	A integrina $\alpha 4\beta 1$ como integrante da família $\beta 1$ .....	26
1.2.1	<i>Sequências de reconhecimento para a integrina <math>\alpha 4\beta 1</math></i> .....	28
1.3	O planejamento de fármacos.....	29
1.3.1	<i>Integrinas como alvos terapêuticos</i> .....	31
2	OBJETIVOS.....	34
2.1	Objetivo Geral.....	34
2.2	Objetivos Específicos.....	34
3	METODOLOGIA (PARTE I).....	35
3.1	<i>Docking</i> molecular.....	35
3.2	Dinâmica Molecular (DM).....	36
3.2.1	<i>Ensemble docking e simulações dos novos ligantes derivados de BIO-5192 em VLA-4</i> .....	38
3.2.1.1	<i>Análises de Componentes Principais (PCA)</i> .....	43
3.3	Cálculos de energia livre de ligação.....	44
3.3.1	<i>Cálculos de energia livre de ligação nos complexos <math>\alpha 4\beta 1</math>-ligantes</i> .....	46
4	METODOLOGIA (PARTE II).....	47
4.1	Reposicionamento de fármacos.....	47
4.1.1	<i>Abordagens computacionais no reposicionamento de fármacos</i> .....	50
5	RESULTADOS E DISCUSSÃO – PARTE I.....	52
5.1	Análises do <i>docking</i> molecular de BIO-5192 e seus derivados em VLA-4.....	52
5.2	Análises das simulações de BIO-5192 e seus derivados em VLA-4.....	57
5.3	Análises conformacionais de VLA-4 a partir das simulações de DM.....	63
5.4	Análises de cálculos de energia livre de ligação para os complexos de BIO-5192 e seus derivados em VLA-4.....	67
6	RESULTADOS E DISCUSSÃO – PARTE II.....	75
6.1	Análises do protocolo de <i>Virtual Screening</i> (VS) no reposicionamento de fármacos na integrina $\alpha 4\beta 1$ .....	75
6.2	Análises da validação do <i>Virtual Screening</i> (VS) no reposicionamento de fármacos na integrina $\alpha 4\beta 1$ .....	79

6.3	Análises das simulações de fármacos oriundos do reposicionamento de fármacos na integrina $\alpha 4\beta 1$ .....	89
6.4	Análises de cálculos de energia livre dos fármacos oriundos do reposicionamento de fármacos na integrina $\alpha 4\beta 1$ .....	94
7	CONCLUSÕES.....	100
	REFERÊNCIAS.....	101
	ANEXO A – ARTIGO PUBLICADO NO PERIÓDICO <i>LIFE</i> .....	113
	ANEXO B - ARTIGO PUBLICADO NO PERIÓDICO <i>ACS OMEGA</i> .....	114
	ANEXO C - ARTIGO PUBLICADO NO PERIÓDICO <i>JOURNAL OF CHEMICAL INFORMATION AND MODELING</i> .....	115

## 1 INTRODUÇÃO

A migração celular é um processo fundamental para muitas das funções do sistema linfóide hematopoiético, incluindo o tráfego fisiológico de leucócitos e a gênese e manutenção de diferentes respostas inflamatórias. Dentre as interações moleculares que regem a migração leucocitária, estão aquelas mediadas por selectinas, quimiocinas e componentes da matriz extracelular (MEC). Os receptores primários da superfície celular envolvidos na adesão aos elementos da MEC pertencem à família das integrinas (HYNES, 1987, 1992, 2002).

As integrinas da família  $\beta 1$  são expressas por alguns tipos de células sanguíneas, incluindo linfócitos T, e desempenham um papel crítico em uma ampla gama de processos biológicos e doenças inflamatórias crônicas, carcinogênese e metástase (DARGAHI et al., 2017; OH et al., 2018). Elas são, também, alvos de muitas moléculas terapêuticas contra trombose e inflamação (HYNES, 2002). Desse modo, alterações em genes que codificam para as integrinas são o principal fator para muitas doenças humanas, sejam elas genéticas, autoimunes etc.

As doenças autoimunes são representantes de muitas desordens inflamatórias, a saber: asma, artrite reumatoide, doença de *Crohn*, esclerose múltipla, dentre outras. Elas são caracterizadas pelo ataque dos leucócitos a células sadias do corpo, identificadas erroneamente como estranhas. As integrinas ( $\alpha 4$ ) têm sido utilizadas como alvos terapêuticos importantes em doenças autoimunes, porém ainda não existe um tratamento definitivo (GUBATAN et al., 2021; KHOY et al., 2020).

Dentre as abordagens terapêuticas envolvendo o bloqueio das funções das integrinas, uma delas diz respeito ao uso de ligantes peptidomiméticos, baseado na capacidade da maioria das integrinas em reconhecer porções peptídicas às suas moléculas de ligação. Foi demonstrado que a adição de uma estrutura química à porção N-terminal do tripeptídeo LDV aumentaria consideravelmente sua capacidade inibitória. Essa estrutura, batizada de 4-[N'-(2-metilfenil)ureido]fenilacetil (PUPA), trouxe uma nova gama de compostos, tornando-se um modelo para o desenvolvimento de medicamentos por empresas farmacêuticas (LIN et al., 1999).

Anticorpos monoclonais também têm sido utilizados para tratar doenças associadas à expressão anormal de integrinas, como vedolizumabe (SINGH et al., 2016) e natalizumabe (BLOOMGREN et al., 2012). No entanto, a produção dessas macromoléculas

tem alto custo associado, o que torna viável a escolha de ligantes pequenos (menor peso molecular).

Frequentemente, o conhecimento teórico-computacional sobre o reconhecimento do complexo receptor-ligante agrega valioso entendimento sobre as bases moleculares que regem o reconhecimento para o planejamento de novos fármacos com potencial terapêutico funcional. Por exemplo, novos métodos para construir ligantes através de fragmentos moleculares (*fragment-based de novo design*), combinados com a técnica de dinâmica molecular (DM) e metodologias de *docking* molecular, configuram uma abordagem funcional inicial para planejamento e idealização de novos compostos protótipos (PATEL; BAUMAN; ARNOLD, 2014; TRIPATHI; SINGH; SING, 2016; ZHENG; LEFTHERIS, 2020).

No presente trabalho, aplicamos uma combinação de técnicas computacionais envolvendo *docking* molecular, dinâmica molecular, cálculos de energia livre e Análises de Componentes Principais (PCA, do inglês *Principal Component Analysis*) para estudar interações de compostos desenvolvidos pelo grupo (VASCONCELOS et al., 2022) com a integrina VLA-4 ( $\alpha 4\beta 1$  ou *Very Late Antigen-4*). Além disso, nos valem dessas técnicas para estudar a interação de fármacos comerciais aprovados em VLA-4, em uma abordagem de reposicionamento de fármacos. Ao longo dos anos, numerosos antagonistas de integrina  $\beta 1$  foram relatados (BAIULA et al., 2019a; HEDRICK et al., 2016), no entanto, como essas moléculas interagem com alvos moleculares específicos permanece desconhecido. Assim, nosso objetivo é lançar luz sobre essas questões desafiadoras.

## 1.1 Estrutura e função das integrinas

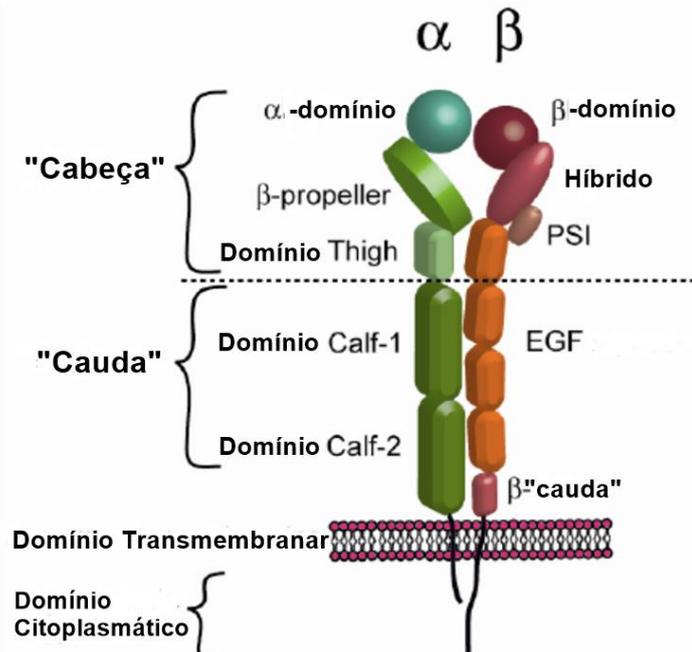
Integrinas são receptores de adesão celular e de adesão à MEC que possuem uma origem evolutiva antiga (JOHNSON et al., 2009). Trata-se de glicoproteínas transmembranares que possuem um papel crucial em processos fisiopatológicos. O controle da migração de alguns tipos celulares sanguíneos é primariamente mediado por componentes da MEC e seus receptores. Desse modo, a expressão diferencial de integrinas representa uma fase fundamental no processo de direcionamento de células do sangue.

As integrinas representam uma das mais abundantes classes de proteínas estudadas. Um dos fatores determinantes que ensejam dificuldades na caracterização da família de integrinas é que muitos dos seus ligantes fazem parte de um vasto grupo de moléculas de adesão à MEC que, além de interagirem com integrinas, reconhecem também outros receptores de MEC, fatores de crescimento, citocinas etc. (BARCZYK; CARRACEDO; GULLBERG, 2010).

Outra característica relevante das integrinas é a de serem âncoras para microrganismos, como vírus (ESFANDIAREI et al., 2006) e bactérias (HOFFMANN; OHLSEN; HAUCK, 2011). Um dos mecanismos para o processo de infecção viral acontece através da interação de proteínas presentes no envelope ou capsídeo do vírus com integrinas nas células do hospedeiro, apesar de que o modo como esse processo ocorre ainda permanece obscuro (HUSSEIN et al., 2015).

Do ponto de vista estrutural, as integrinas compreendem uma superfamília formada por heterodímeros de cadeias  $\alpha\beta$ . Em humanos, são conhecidos pelo menos 18 domínios  $\alpha$  e 8  $\beta$  (SHIMAOKA; TAKAGI; SPRINGER, 2002). As subunidades  $\alpha$  e  $\beta$  da integrina são totalmente distintas, sem homologia detectável entre elas; a identidade de sequência entre as subunidades  $\alpha$  é de cerca de 30% e entre as subunidades  $\beta$  de 45%, indicando que ambas as famílias de genes  $\alpha$  e  $\beta$  evoluíram por duplicação. Em humanos, os genes para as subunidades  $\alpha$  e  $\beta$  estão localizados em vários cromossomos (TAKADA; YE; SIMON, 2007). Portanto, as integrinas são consideradas a mais diversificada família de moléculas de adesão celular (HUMPHRIES, 2000; HYNES, 1992). A Figura 1 retrata, com mais detalhes, a estrutura de uma integrina.

Figura 1 – Representação esquemática da estrutura de uma integrina, indicando cada um de seus domínios. PSI = domínio *Plexin-Semaphorin-Integrin*. EGF = Epidermal Growth Factor



Fonte: Adaptada de (GAHMBERG et al., 2009; NEVO, 2011).

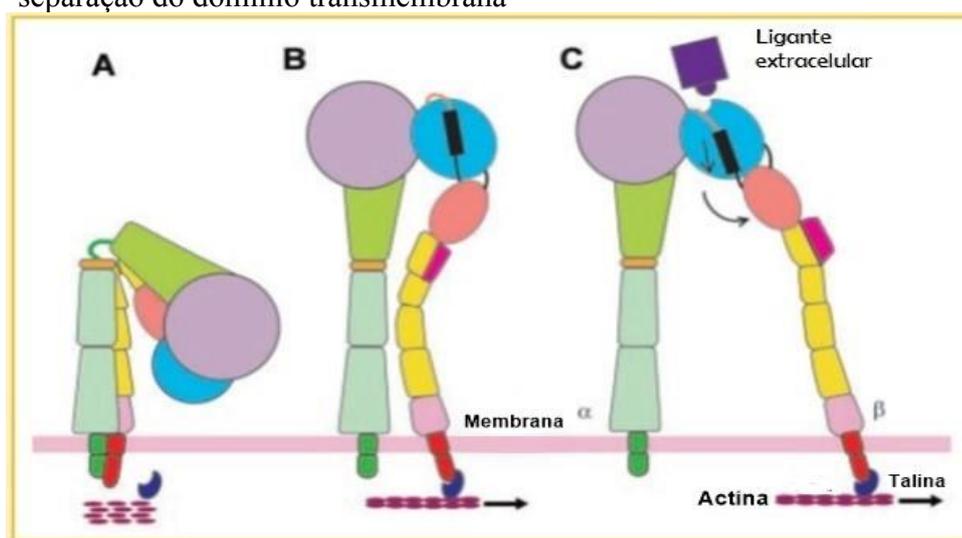
A região do sítio de ligação, conhecida por cabeça da integrina, é formada pela associação do N-terminal das cadeias alfa e beta, porção localizada na região extracelular do receptor. O domínio C-terminal perpassa a membrana plasmática da célula e encontra a porção intracelular, onde ocorre os diversos mecanismos de transdução de sinais, no interior da célula, mediados pelo citoesqueleto de actina (GIANCOTTI; RUOSLAHTI, 1999).

De acordo com a afinidade das integrinas por seus ligantes, ocorrem mecanismos de ativação desses receptores, a partir de mudanças conformacionais. No mecanismo conhecido por *inside-out*, as integrinas podem ser ativadas intracelularmente por sinais de receptores acoplados à proteína G que levam à fosforilação do domínio citoplasmático da subunidade  $\beta$ . A associação das caudas citoplasmáticas  $\alpha$  e  $\beta$  parece ser necessária para manter uma integrina no estado inativo; a associação é interrompida pelo uso de agonistas, como quimiocinas, que são conhecidas por causar ativação da integrina e que sinalizam vias de receptores acoplados à proteína G. A proteína adaptadora do citoesqueleto talina desempenha um papel na regulação da afinidade da integrina. A ligação da região da cabeça da talina à cauda citoplasmática da integrina  $\beta$  causa a dissociação das caudas  $\alpha$  e  $\beta$  e induz uma mudança conformacional na região extracelular que aumenta sua afinidade por seu ligante (SCHWARTZ; GINSBERG, 2002).

Dois modelos foram propostos para esta mudança na afinidade: modelo trava e canivete. Em ambos, a integrina inativa está em uma conformação dobrada, com a cabeça voltada para a membrana. Os eventos de mudança conformacional se correlacionam em segundos com a ativação da integrina, levando a mudanças conformacionais no sítio de ligação na cabeça que aumentam sua afinidade pelo ligante (LUO; CARMAN; SPRINGER, 2007).

Ao conectar ligantes extracelulares, as integrinas de mamíferos se agrupam na membrana e transduzem sinais para o interior da célula (mecanismo *outside-in*). A interação extracelular do ligante induz alterações conformacionais, incluindo o balanço externo do domínio híbrido, separação dos domínios  $\alpha$  e  $\beta$  e separação dos domínios transmembranares, que levam à interação das caudas citoplasmáticas com a sinalização intracelular de moléculas (ARNAOUT; MAHALINGAM; XIONG, 2005). Estas incluem enzimas (por exemplo, a quinase de adesão focal/cSrc e as pequenas GTPases Ras e Rho) e adaptadores (por exemplo, Cas/Crk e paxilina) que se montam dentro de estruturas de adesão dinâmica, incluindo adesões focais que ligam as células à MEC e aos podossomos (pequenas extensões semelhantes a pés da membrana plasmática) (GINSBERG; PARTRIDGE; SHATTIL, 2005). A figura abaixo destaca o que foi discutido acima

Figura 2 – Representação esquemática do processo de biossinalização em integrinas. O processo, em resumo, ocorre a partir de um estímulo interno (*inside-out*) ou externo (*outside-in*). Então, a “cabeça” da integrina se mantém exposta a uma molécula ligante, fazendo com que haja uma separação do domínio transmembrana



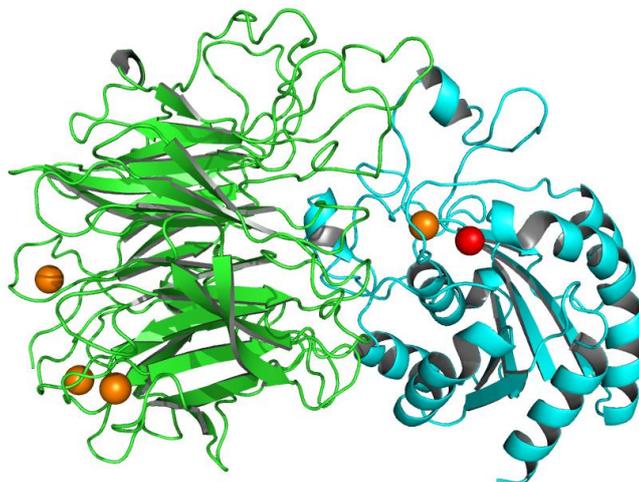
Fonte: Adaptada de Wang; Luo (2010).

Estruturalmente, as integrinas ainda são classificadas de acordo com a presença ou ausência de um domínio extracelular, chamado de domínio inserido (domínio I) (HUMPHRIES, 2000). Esta porção estrutural está presente em aproximadamente metade das integrinas. Naqueles receptores que não são dotados do domínio I, outra região característica chamada de *β-propeller* (SHIMAOKA; TAKAGI; SPRINGER, 2002; TAKAGI; SPRINGER, 2002) se faz fundamental no reconhecimento de ligantes específicos. Trata-se de uma região globular N-terminal presente na subunidade  $\alpha$ . Já nas integrinas que possuem o domínio I, o *β-propeller* funciona de modo a cooperar indiretamente na ligação de moléculas específicas ou não tem função relevante comprovada.

O *β-propeller* contém sete repetições de cerca de 60 aminoácidos cada, que se dobram em uma estrutura em hélice  $\beta$  de sete lâminas semelhante à subunidade  $\beta$  de uma proteína G heterotrimérica. O domínio I contém uma região chamada de Sítio de Adesão Dependente de Íon Metálico (do inglês, *Metal Ion-Dependent Adhesion Site* - MIDAS) e os domínios semelhantes a I contêm um motivo de ligação de metal estruturalmente semelhante. O sítio de ligação RGD (arginina, glicina e ácido aspártico) está localizado na interface entre o domínio *β-propeller* e o domínio I e os resíduos de aminoácidos dos dois domínios interagem diretamente com o peptídeo RGD de um ligante (XIONG et al., 2002).

A presença da região MIDAS (Figura 3) na integrina é uma característica estrutural relevante, já que esta porção abriga um cátion divalente. Logo, a coordenação do íon  $Mg^{2+}$ , na região MIDAS, com um grupamento carboxilato de um ligante representa o principal mecanismo de reconhecimento de integrinas com moléculas específicas (NAGAE et al., 2012). Outros sítios de íons metálicos também foram encontrados semelhantes ao MIDAS; dos dois locais ADMIDAS (adjacentes ao MIDAS), um deles é chamado de Sítio Sinérgico de Ligação de Íons Metálicos (SyMBS).

Figura 3 – Representação estrutural modificada do modelo gerado por homologia da integrina  $\alpha\beta 1$ , mostrando, em vermelho, o íon  $Mg^{2+}$  da região MIDAS. As cadeias  $\alpha$  e  $\beta$  foram coloridas com as cores verde e ciano, respectivamente. Figura produzida no programa PyMOL



Fonte: O autor.

Um grande passo foi dado no sentido da elucidação estrutural e do funcionamento das integrinas, quando foi determinada a estrutura cristalográfica de  $\alpha\beta 3$  (código PDB 1L5G) (XIONG et al., 2001, 2002). O cristal foi obtido em conjunto com o ligante peptidomimético RGD, a partir da resolução de 3,2 Å. Desse modo, mesmo com esse nível de detalhamento estrutural, foi possível estudar como as cadeias laterais dos aminoácidos interagem com o sítio ativo da integrina e, assim, definir diversos parâmetros físico-químicos entre receptor e ligante, tais como: interações intramoleculares, posicionamento de íons, localização de pontes dissulfeto na proteína, entre outros. A estrutura resolvida exibe uma conformação fechada, o que resultou em controvérsias quanto ao processo de ativação da integrina. Ao final, foi elucidado que a forma estendida é a estrutura ativa, mediante a interação com moléculas específicas no sítio de ligação (TAKAGI et al., 2002).

Outro cristal que ajudou no desvendamento de questões funcionais de integrinas é o do ectodomínio da integrina  $\alpha 5\beta 1$  (código PDB 3VI4) (NAGAE et al., 2012). O cristal foi resolvido a 2,9 Å em conjunto com um anticorpo alostérico em contato com o bolsão de interação. A estrutura ainda foi cristalizada na presença e na ausência do peptídeo RGD, o que lançou luz, a partir de simulações de dinâmica molecular, o papel dos íons divalentes nas mudanças conformacionais que estão interligadas entre o sítio de ligação e o resto da molécula.

## 1.2 A integrina $\alpha 4\beta 1$ como integrante da família $\beta 1$

A integrina  $\alpha 4\beta 1$  (VLA-4 ou *Very Late Antigen-4*) faz parte de uma família desses receptores que possuem a subunidade  $\beta 1$ , também chamada de CD29, o qual é um gene humano que codifica para a proteína de mesmo nome e possui um total de 58048 pares de bases (pb) (AKSORN; CHANVORACHOTE, 2019).

Sendo a subunidade  $\beta$  mais comum da família das integrinas, a integrina  $\beta 1$  é atualmente conhecida por se ligar a várias subunidades  $\alpha$  para formar 11 integrinas diferentes. Entre elas, a integrina  $\beta 1$  forma os complexos (VLA)-1 a 6,  $\alpha 7\beta 1$ ,  $\alpha 8\beta 1$  e  $\alpha v\beta 1$  (receptor de vitronectina, VNR) com integrinas  $\alpha 1$  a  $\alpha 8$  e subunidades  $\alpha v$ , respectivamente, constituindo o grupo VLA, que é amplamente distribuído por todo o corpo (JUSTO; JASIULIONIS, 2021). Uma série de estudos revelaram que as integrinas compostas pela subunidade  $\beta 1$  desempenham um papel fundamental na manutenção de propriedades das células tumorais, bem como na promoção de metástase e resistência à quimioterapia/radiação participando da transdução de várias vias de sinalização intracelular (PELLINEN et al., 2018; ZHANG et al., 2019).

As integrinas da família  $\beta 1$  são expressas em alguns tipos celulares, como os linfócitos T. Logo, a expressão nesses leucócitos medeia a ligação a componentes da MEC, como o colágeno e a fibronectina (HEMLER, 1990). A interação entre essa família de integrinas e componentes da MEC promove uma sinalização intracelular, mediada por enzimas tirosinas cinases, a qual resulta em mecanismos de proliferação e migração celulares. Assim, os linfócitos T conseguem modular a expressão de integrinas na membrana, a partir de sinalizações celulares externas, as quais controlam o fenótipo adesivo desses leucócitos (DIAMOND; SPRINGER, 1994).

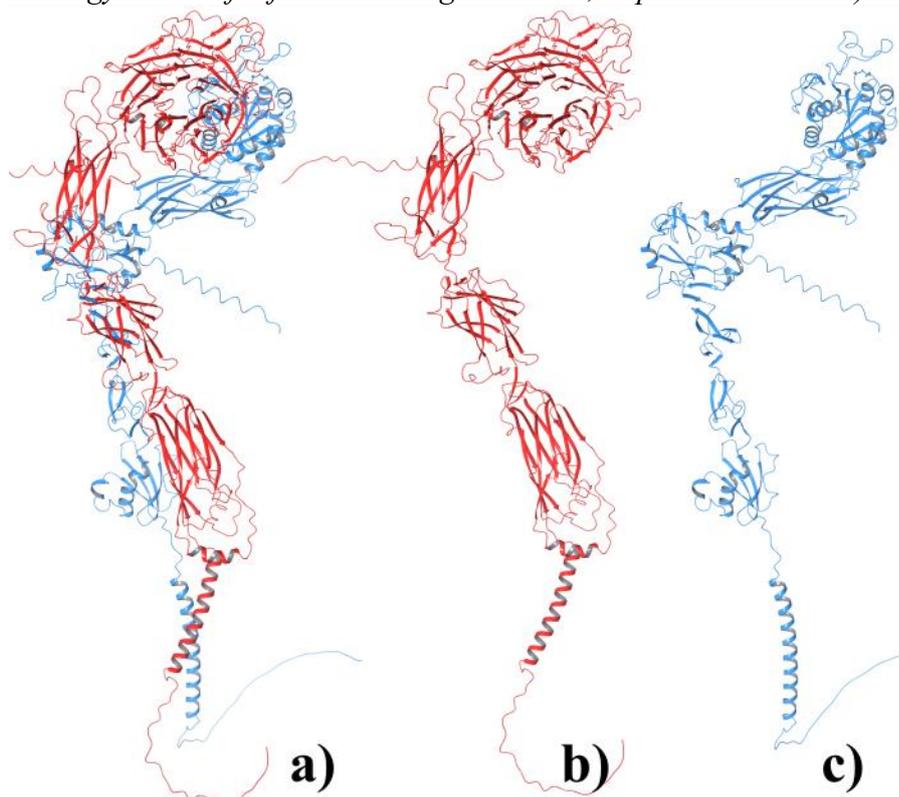
Diante disso, estudos recentes têm apontado o papel das integrinas  $\beta 1$  em mecanismos envolvidos com a migração celular e, portanto, associando esses receptores como alvos terapêuticos a diversos tipos de câncer (FIGUEIREDO et al., 2022; HAAKE et al., 2022; ZHANG et al., 2022a).

O conjunto de ligantes possíveis para interação com todas as integrinas da família  $\beta 1$  é amplo. Algumas interagem com apenas uma proteína na MEC ( $\alpha 5\beta 1$  e  $\alpha 6\beta 1$ ), outras já se ligam a mais de uma molécula na matriz ( $\alpha 1\beta 1$ ) e, por último, existem aquelas que interagem com proteínas na MEC e com receptores de superfície celular ( $\alpha 4\beta 1$ ) (NIU; CHEN, 2011).

A integrina VLA-4 (CD49d/CD29) (Figura 4) é uma das mais bem estudadas, já que é alvo terapêutico para o tratamento de inúmeras patologias relacionadas aos processos de

migração celular. O receptor é expresso em muitas células do sistema imune, como células T e B, NK, eosinófilos e monócitos.

Figura 4 – a) Modelo por homologia de  $\alpha 4\beta 1$  obtido pelo programa AlphaFold; b) a cadeia  $\alpha$  da integrina onde reside o domínio  $\beta$ -propeller; c) cadeia  $\beta$  da integrina onde está o íon divalente MIDAS (Kozmon, S. *Homology model of alfa4beta1 integrins*. 2022; unpublished results)



Fonte: Adaptada de Li et al. (2023).

Esta integrina reconhece a porção CS-1 da fibronectina, a qual é um componente da MEC. Além disso, VLA-4 interage com a molécula de adesão celular vascular-1 (VCAM-1) no endotélio inflamado; esta interação é fundamental na infiltração de células do sistema imune no sistema nervoso central (SNC) e com a tríade peptídica LDV (leucina, ácido aspártico e valina) (HUMPHRIES; BYRON; HUMPHRIES, 2006).

Além disso, o heterodímero desempenha um papel central no recrutamento de leucócitos, em processos inflamatórios e em doenças autoimunes (BERLIN et al., 1993; CHAVAKIS, 2012). Durante o mecanismo de inflamação, essa proteína participa de eventos como transporte, adesão e infiltração de linfócitos.

O bloqueio de  $\alpha 4$  impediu a paralisia em um modelo murino de esclerose múltipla de célula T induzida, e o bloqueio de  $\alpha 4\beta 1$  inibiu especificamente a ligação de células do sistema imune a vasos cerebrais inflamados (YEDNOCK et al., 1992). Esse estudo forneceu a

justificativa para o direcionamento de  $\alpha 4\beta 1$  em doenças do SNC imunomediadas.

A expressão anormal de  $\alpha 4\beta 1$  é frequentemente observada no câncer, levando a vários ensaios clínicos com inibidores direcionados a diversos receptores (HE et al., 2022; ZHANG et al., 2021). No entanto, a expressão e a função desta integrina no câncer permanecem incertas, o que é muito provavelmente devido à dificuldade intrínseca em obter a estrutura tridimensional de receptores de membrana através da cristalografia de raios-X (KELSH-LASHER et al., 2017).

### ***1.2.1 Sequências de reconhecimento para a integrina $\alpha 4\beta 1$***

Várias integrinas reconhecem o segmento da sequência do tri peptídeo RGD (Arg-Gly-Asp) presente em diferentes proteínas que compõem a MEC, enquanto outras reconhecem sequências alternativas, como EILDV (Glu-Ile-Leu-Asp-Val) ou REDV (Arg-Glu-Asp-Val) entre outros. Porém,  $\alpha 4\beta 1$  pertence à família de integrinas leucocitárias (não-RGD) com diferentes sequências de reconhecimento pertencentes a ligantes não-RGD (HATLEY et al., 2019).

A integrina VLA-4 reconhece a sequência peptídica Gln-Ile-Asp-Ser (QIDS) em VCAM-1 e Ile-Leu-Asp-Val (ILDV) na fibronectina (FN), respectivamente. Bloquear a interação de  $\alpha 4\beta 1$  com seus ligantes tem sido usado como estratégia terapêutica para inflamação e autoimunidade. Um dos primeiros estudos publicados na área foi de Liu e colaboradores que usaram a sequência ILDV em FN como um ponto de partida para propor novos antagonistas. Eles sintetizaram uma série de 51 inibidores específicos para a integrina  $\alpha 4\beta 1$  visando o câncer linfóide (LIU et al., 2006).

Nesse estudo, ao verificar uma biblioteca de peptídeos aleatórios com uma linhagem celular de leucemia Jurkat T, os autores identificaram a sequência de aminoácidos D-Pro-Leu-Asp-Ile (pLDI) como o único motivo de proteína que interagiu preferencialmente ao receptor de integrina  $\alpha 4\beta 1$  em células linfóides, células malignas humanas, mas não em linfócitos periféricos normais. Assim, os autores propuseram que a integrina  $\alpha 4\beta 1$  nas células linfóides malignas é constitutivamente ativada e que é um alvo terapêutico para a leucemia linfocítica aguda e a leucemia linfocítica crônica.

### 1.3 O planejamento de fármacos

A pesquisa e o desenvolvimento de medicamentos é um processo de várias etapas que inclui a descoberta de protótipos, testes clínicos e aprovação para a produção. A descoberta de medicamentos é um processo longo, caro e complexo que dura anos e custa milhões de dólares (DIMASI; GRABOWSKI; HANSEN, 2016; SIMOENS; HUYS, 2021). Esse processo consiste na identificação do alvo, descoberta de compostos-líder (*leads*), otimização de *leads* (incluindo síntese molecular e testes de bioatividade) e avaliação pré-clínica (SABE et al., 2021).

Apesar do investimento em grandes quantias e do tempo despendido, a taxa de sucesso dos testes clínicos é inferior a 15% (ZHU et al., 2023). A causa de falha em aproximadamente 50% do processo de planejamento de fármacos é devido a distorções nas propriedades farmacocinéticas (absorção, distribuição, metabolismo, excreção e toxicidade – ADMET) (KIM; KIM; SEONG, 2010). A velocidade e a taxa de sucesso na descoberta de medicamentos aumentaram fortemente com o desenvolvimento de abordagens computacionais (HILL; RICHARDS, [s.d.]).

Atualmente, as abordagens computacionais e de aprendizagem de máquina desempenham um papel cada vez mais vital no planejamento de fármacos. A rápida evolução de métodos e algoritmos encurtou o tempo e custos financeiros para encontrar os candidatos a medicamentos. Na descoberta de fármacos, as contribuições da biologia computacional incluem a caracterização de mecanismos moleculares de interação de ligantes, englobando a identificação de sítios de ligação/ativos e o refinamento das estruturas dos modos de ligação das moléculas-alvo. A maioria dessas abordagens indica que os sítios de ligação/ativos na proteína-alvo devem ser bem determinados (ZHANG et al., 2022b).

A biologia computacional, especialmente a simulação de biomoléculas, é um método poderoso para revelar o mecanismo molecular da proteína-alvo e fornecer novas perspectivas para o desenho de fármacos. O desenho de fármacos auxiliado por computador (do inglês, *Computer-aided drug design* (CADD)) é normalmente usado para descobrir, desenvolver e analisar candidatos a medicamentos e moléculas ativas com propriedades bioquímicas semelhantes (SURABHI; SINGH, 2018) e para acelerar a descoberta de medicamentos reduzindo custos e taxas de erro.

Candidatos a fármacos descobertos por CADD geralmente vêm de bibliotecas de moléculas pequenas. Essas descobertas são feitas usando vários métodos, incluindo *docking*

*molecular*, modelagem farmacófora, VS (Varredura Virtual, do inglês, *Virtual Screening*) e Relação Quantitativa Estrutura-Atividade (QSAR).

Entre essas abordagens, VS é a principal ferramenta aplicada para rastrear novos compostos com propriedades desejadas em grandes bancos de dados químicos. VS é classificado como VS baseado em estrutura (do inglês, *Structure-based Virtual Screening - SBVS*) (DOS SANTOS; FERREIRA; ANDRICOPULO, 2018) e VS baseado em ligante (do inglês, *Ligand-based Virtual Screening - LBVS*) (BANEGAS-LUNA; CERÓN-CARRASCO; PÉREZ-SÁNCHEZ, 2018). VS é usado para acelerar a descoberta de fármacos e encurtar o número de compostos a serem testados no laboratório. Além disso, o VS também desempenha um papel importante no reaproveitamento ou reposicionamento de medicamentos, otimizando os compostos candidatos, o que acelera todo o processo de planejamento de fármacos (SOHRABY; BAGHERI; ARYAPOUR, 2019).

Outra abordagem computacional é a descoberta de fármacos baseada em fragmentos (do inglês, *Fragment-based Drug Discovery - FBDD*) que se tornou um procedimento prático para explorar o espaço químico de uma proteína-alvo, levando a compostos de maior afinidade (PATEL; BAUMAN; ARNOLD, 2014). Particularmente, o desenho *de novo* baseado em fragmentos envolve a criação de novas entidades químicas com base na estrutura do sítio de ligação com o qual elas devem interagir (TRIPATHI; SINGH; SING, 2016). Os projetos de FBDD têm algumas vantagens atraentes, como investimentos mais baixos em pesquisa e desenvolvimento e o uso de bibliotecas menores de moléculas de baixa complexidade. Essas vantagens têm encorajado pesquisadores no desenvolvimento de novos inibidores para diversos alvos de proteínas. Como as integrinas são receptores multidomínios, o uso de FBDD pode ser útil na busca por novos sítios de ligação e por novos compostos inibidores (ZHENG; LEFTHERIS, 2020).

Empregando a mecânica newtoniana, a simulação de dinâmica molecular, que tem sido amplamente empregada no planejamento de fármacos, pode capturar a posição e o movimento de cada átomo em um sistema (DURRANT; MCCAMMON, 2011). Esta abordagem pode revelar os detalhes das interações moleculares e de alterações conformacionais da proteína-alvo, que fornecem informações complementares aos testes experimentais (ZHAO; CAFLISCH, 2015). Além disso, simulações de DM podem fornecer dados de termodinâmica, cinética e perfis de energia livre do complexo receptor-ligante (HUANG; CAFLISCH, 2011). Estas informações podem ser úteis para melhorar a afinidade de ligação do composto-líder.

### 1.3.1 Integrinas como alvos terapêuticos

As integrinas tornaram-se potenciais alvos terapêuticos porque elas estão envolvidas na fisiopatologia de diversas doenças. Antagonistas de integrinas, atualmente no mercado ou em ensaios clínicos, visam a interação nos sítios de ligação, os quais são expressos em células sanguíneas, leucócitos ou plaquetas (LEY et al., 2016).

A cadeia  $\alpha 4$  pode parear com  $\beta 1$  para gerar VLA-4 ou com  $\beta 7$  para formar  $\alpha 4\beta 7$ ; portanto, medicamentos que têm como alvo  $\alpha 4$  podem efetivamente inibir duas integrinas,  $\alpha 4\beta 1$  e  $\alpha 4\beta 7$  (WOODSIDE; VANDERSLICE, 2019).

Natalizumabe, o único antagonista de  $\alpha 4$  aprovado pela FDA (*Food and Drug Administration*), é um anticorpo monoclonal neutralizante (IgG4k) que se liga a integrinas  $\alpha 4$  e bloqueia o reconhecimento de proteínas fisiológicas ligantes, como VCAM-1, que é expresso em células endoteliais inflamadas, macrófagos e outras células, e fibronectina também. Essas propriedades já foram demonstradas *in vitro*, pois o natalizumabe preveniu efetivamente a adesão de células Jurkat humanas, que expressavam  $\alpha 4\beta 1$ , a recombinantes purificados de VCAM-1 (LEY et al., 2016; PETERSEN et al., 2019).

O natalizumabe foi aprovado pela primeira vez em 2004 para o tratamento da esclerose múltipla, após resultados promissores obtidos em 2 ensaios de fase III, que mostraram uma redução na progressão da doença. A droga permaneceu no mercado até que foi removido voluntariamente pela Biogen em 2005 devido ao aparecimento de leucoencefalopatia progressiva multifocal em 2 pacientes com esclerose múltipla. Em 2006, o natalizumabe foi reintroduzido no mercado em um programa de avaliação de risco sob o nome comercial Tysabri (PERNECZKY; SELLNER, 2022).

Vedolizumabe (Entyvio) é um anticorpo monoclonal humanizado direcionado à integrina  $\alpha 4\beta 7$  já no mercado. O medicamento trata a doença inflamatória intestinal (DII), incluindo colite ulcerativa e doença de Crohn. Recentemente, um anticorpo monoclonal humano AMG 181 foi projetado para tratamento de DII pela Amgen contra a integrina  $\alpha 4\beta 7$ , agora em ensaios de Fase II (LI et al., 2023).

Embora o objetivo da maioria dos antagonistas de  $\alpha 4$  seja diminuir a infiltração excessiva de células imunes, essas terapias têm o potencial para serem usadas contra células malignas que se beneficiam da integrina  $\alpha 4$  para o extravasamento das células tumorais. O sucesso em direcionar um ligante para a integrina  $\alpha 4$ , para bloquear sua adesão e função de transdução de sinal em doenças como o câncer, dependerá da capacidade de minimizar a imunossupressão ou afetar indiretamente a função de  $\alpha 4$ . Ao inibir a interação do ligante,

muitos antagonistas interrompem com sucesso as conexões celulares com o ambiente extracelular e as vias de sobrevivência que são necessárias para a progressão do tumor.

Outra abordagem terapêutica na tentativa de bloquear a ação de integrinas diz respeito ao uso de ligantes peptídicos ou peptideomiméticos. Esta iniciativa baseou-se no fato de que boa parte das integrinas tem a capacidade de reconhecer porções peptídicas de suas moléculas ligantes.

A vantagem dessas moléculas em relação àquelas puramente peptídicas reside no fato de que estas são instáveis e de baixa biodisponibilidade, comparadas aos peptideomiméticos. Além disso, existe a contraindicação do uso oral em humanos de moléculas peptídicas, já que sofrem degradação enzimática no estômago.

Várias moléculas pequenas foram desenvolvidas para reconhecer integrinas leucocitárias; detalhes podem ser encontrados em alguns trabalhos (BAIULA et al., 2019a; DOTAN et al., 2020; LEY et al., 2016). Eptifibatide (Integrilin), tirofiban (Aggrastat) e lifitegrast (Xiidra) são três moléculas pequenas antagonistas de integrinas que entraram no mercado. O eptifibatide é um heptapeptídeo RGD mimético e liga-se às plaquetas através da integrina  $\alpha\text{IIb}\beta_3$ ; foi aprovado em 1998. O medicamento previne a formação de trombos em uma artéria coronária. É usado para reduzir o risco de eventos isquêmicos cardíacos agudos. O segundo fármaco, o tirofiban emprega o mesmo mecanismo e tem a mesma indicação do eptifibatide. Foi aprovado em 1998. O terceiro medicamento, o lifitegrast foi aprovado em 2016 e é um inibidor da integrina  $\alpha\text{L}\beta_2$ . Lifitegrast bloqueia a interação da integrina ao seu ligante ICAM-1 e, assim, diminui a inflamação. É usado para tratamento de olhos secos (xerofthalmia), incluindo disfunção da glândula meibomiana. Nenhum desses agentes é administrado por via oral.

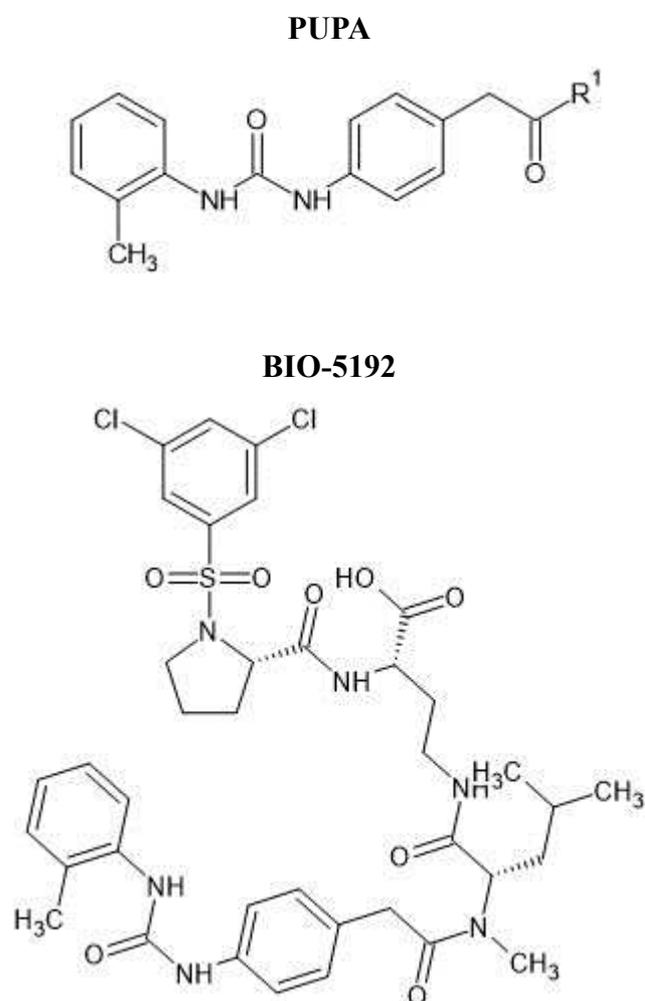
Ao longo do tempo, foi descoberto que a adição de uma estrutura química na porção N-terminal de LDV aumentaria, consideravelmente, a capacidade inibitória dessas moléculas. Esta estrutura é denominada de 4-[N'-(2-metilfenil)ureido]fenilacetil (PUPA) (Figura 5). Desse modo, uma nova gama de compostos foi se tornando alvo de desenvolvimento por parte de empresas farmacêuticas (LIN et al., 1999).

Existem também as moléculas antagonistas de integrinas baseadas na estrutura de N-acilfenilalanina. Um exemplo é o composto valategrast. Muitas dessas moléculas exibem atividade antagonista dupla para VLA-4 e  $\alpha_4\beta_7$  (HIJAZI et al., 2004).

Neste trabalho, os estudos se concentrarão no composto-líder chamado de BIO-5192 (Figura 5). Trata-se de um composto peptideomimético cuja atividade inibitória in vitro contra a integrina VLA-4 já foi demonstrada (LEONE et al., 2003; RAMIREZ et al., 2009).

Nestes trabalhos, foi reportado um IC<sub>50</sub> (Concentração Inibitória da metade) de 0.0018  $\mu$ M (LEONE et al., 2003). Em experimentos com células expressando as integrinas  $\alpha 4\beta 7$ ,  $\alpha 9\beta 1$ ,  $\alpha 2\beta 1$  e  $\alpha I\text{I}\beta 3$ , BIO-5192 mostrou maior seletividade a  $\alpha 4\beta 1$ . A afinidade medida para o complexo  $\alpha 4\beta 1$ /BIO-5192 foi de 250 a 1000 vezes maior que para  $\alpha 4\beta 7$  (LEONE et al., 2003). Além disso, mostrou-se que BIO-5192 proporcionou um aumento de 30 vezes na mobilização de células-tronco hematopoiéticas e de células progenitoras murinas acima dos níveis basais (RAMIREZ et al., 2009).

Figura 5 – Estruturas moleculares de PUPA (acima) e de BIO-5192 (abaixo). R<sup>1</sup> = radicais químicos que podem ser adicionados à estrutura



Fonte: O autor.

## 2 OBJETIVOS

### 2.1 Objetivo Geral

Compreender as interações atômico-moleculares e o comportamento dinâmico de ligantes candidatos a fármacos que são potenciais antagonistas da integrina VLA-4 e de medicamentos disponíveis no mercado candidatos ao reposicionamento de fármacos.

### 2.2 Objetivos Específicos

- Analisar os modos de ligação dos novos compostos derivados de BIO-5192 em VLA-4;
- Compreender o comportamento conformacional dinâmico dos complexos integrina-ligante;
- Estudar modos de movimento de VLA-4 quando complexada aos ligantes;
- Predizer as afinidades dos complexos integrina-ligante;
- Estudar os modos de ligação de fármacos aprovados pelo FDA na integrina VLA-4;
- Analisar as características estruturais dinâmicas dos complexos fármaco-integrina;
- Predizer as afinidades dos complexos fármaco-integrina.

### 3 METODOLOGIA (PARTE I)

#### 3.1 *Docking* molecular

A triagem experimental de grandes bibliotecas de compostos contra alvos moleculares, ou seja, *High-Throughput Screening* (HTS), tem representado o padrão ouro para descobrir *hits* biologicamente ativos. No entanto, os altos custos necessários para estabelecer e manter essas plataformas de triagem muitas vezes dificultam seu uso para a descoberta de medicamentos (SLIWOSKI et al., 2014). Além disso, considerando os recentes desenvolvimentos na tecnologia de computadores e o rápido aumento de dados estruturais, químicos e biológicos disponíveis em um número cada vez maior, é facilmente compreensível como o uso de abordagens *in silico* aumentou significativamente nas últimas décadas (D'AGOSTINO et al., 2013; MACALINO et al., 2015). De fato, as abordagens *in silico* agora permitem a triagem virtual de milhões de compostos em um tempo acessível, reduzindo assim os custos iniciais e melhorando as chances de encontrar os candidatos a medicamentos desejados.

Os métodos baseados em estrutura dependem das informações derivadas do conhecimento da estrutura 3D de um alvo de interesse, e permitem classificar bancos de dados de moléculas de acordo com a estrutura e complementaridade eletrônica de ligantes para um determinado alvo. Neste contexto, o *docking* molecular é entre um dos métodos *in silico* baseados em estrutura mais populares e bem-sucedidos, que ajudam a prever as interações que ocorrem entre ligantes e alvos biológicos. Este processo é geralmente aplicado a partir da previsão da orientação molecular de um ligante dentro de um receptor e, em seguida, estimando sua complementaridade através do uso de uma função de pontuação (KITCHEN et al., 2004).

Dado o potencial oferecido por esse método, esforços crescentes têm sido direcionados para a melhoria dos algoritmos de *docking* e para a superação de suas limitações (AIN et al., 2015; GUEDES; PEREIRA; DARDENNE, 2018). De fato, os principais desafios que caracterizam o *docking* incluem uma amostragem restrita das conformações do ligante e do receptor na predição dos modos de ligação e o uso de funções de pontuação aproximadas, que muitas vezes fornecem resultados que não se correlacionam com as afinidades de ligação experimentais (PANTSAR; POSO, 2018).

No entanto, a aplicação do *docking* no planejamento de fármacos é limitada a alvos biológicos para os quais as estruturas cristalinas são conhecidas. Várias abordagens têm

sido adotadas para superar este último aspecto. Por exemplo, a indisponibilidade de estruturas 3D é muitas vezes contornada pela construção de modelos através da modelagem por homologia. Além disso, esses métodos também podem ser usados em conjunto com a dinâmica molecular para validar e refinar ainda mais os complexos modelados *in silico* (DE VIVO et al., 2016; SALMASO; MORO, 2018).

Metodologias de triagem virtual mais rigorosas também foram desenvolvidas para melhorar as previsões do complexo receptor-ligante baseado em *docking* (RASTELLI; PINZI, 2019). De fato, esses métodos de refinamento e repontuação pós-*docking* são de grande interesse na descoberta de medicamentos porque geralmente fornecem taxas de acerto mais altas no processo de triagem virtual e permitem uma melhor correlação com dados experimentais (GENHEDEN; RYDE, 2015; RASTELLI; PINZI, 2019).

Em particular, o uso da metodologia de *docking* ampliou-se para novos horizontes no planejamento de fármacos, impulsionado pela melhoria dos algoritmos e pelo aumento das informações publicamente acessíveis sobre ligantes e receptores. Por exemplo, graças à velocidade aprimorada e ao poder de previsão, o *docking* também foi incorporado a protocolos de triagem em larga escala para identificar, por exemplo: sítios de ligação em proteínas (HAZARIKA et al., 2018); novos alvos moleculares de ligantes conhecidos (LI et al., 2006); potenciais reações adversas a medicamentos (LABUTE et al., 2014) e ligantes com novos arcabouços químicos ativos contra um determinado alvo ou um conjunto de alvos desejados.

### 3.2 Dinâmica Molecular (DM)

A simulação por DM é um método eficiente para estudar os movimentos físicos de átomos e moléculas. Essas partículas interagem por um período fixo de tempo, mostrando uma visão da evolução dinâmica do sistema.

Geralmente, as trajetórias de átomos e moléculas são determinadas resolvendo numericamente as equações de movimento de Newton para um sistema de partículas em interação. Suas energias potenciais são calculadas usando métodos de mecânica molecular com diferentes campos de força. O campo de força descreve um sistema através da equação de termos simples, que representam as interações entre os átomos (Figura 6).

Figura 6 – Equação da energia potencial calculada nos campos de força, a partir de resolução de termos simples

$$\begin{aligned}
 U(R) = & \sum_{bonds} k_r (r - r_{eq})^2 & \text{bond} & \img alt="Diagram of a bond interaction between two atoms connected by a red spring." data-bbox="755 140 830 165"/> \\
 & + \sum_{angles} k_\theta (\theta - \theta_{eq})^2 & \text{angle} & \img alt="Diagram of an angle interaction between three atoms with a red arc indicating the angle." data-bbox="755 175 830 200"/> \\
 & + \sum_{dihedrals} k_\phi (1 + \cos[n\phi - \gamma]) & \text{dihedral} & \img alt="Diagram of a dihedral interaction between four atoms with a red arrow indicating the rotation around a bond." data-bbox="730 205 855 255"/> \\
 & + \sum_{impropers} k_\omega (\omega - \omega_{eq})^2 & \text{improper} & \img alt="Diagram of an improper interaction between four atoms with a red arrow indicating the out-of-plane displacement." data-bbox="725 265 855 335"/> \\
 & + \sum_{i < j}^{atoms} \epsilon_{ij} \left[ \left( \frac{r_m}{r_{ij}} \right)^{12} - 2 \left( \frac{r_m}{r_{ij}} \right)^6 \right] & \text{van der Waals} & \img alt="Diagram of van der Waals interaction between two atoms with a double-headed arrow." data-bbox="740 340 845 360"/> \\
 & + \sum_{i < j}^{atoms} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} & \text{electrostatic} & \img alt="Diagram of electrostatic interaction between two charged atoms with a double-headed arrow." data-bbox="740 395 845 420"/>
\end{aligned}$$

Fonte: Chang et al. (2016).

Como podemos analisar a partir da Figura 6, a função de energia potencial é dividida em termos de interações ligadas, como os ângulos planares, ângulos torcionais e diedros próprios e impróprios, além dos termos não-ligados, a saber as interações de van der Waals e de Coulomb (ou eletrostáticas).

Os campos de força mais comumente usados em química e biologia são baseados na mecânica molecular e incorporam um tratamento mecânico clássico de interações partícula-partícula que podem reproduzir mudanças estruturais e conformacionais. Em relação ao método experimental, a simulação MD pode não apenas mostrar a variação estrutural em mudanças ambientais, como pH, temperatura e mutações de resíduos, mas também pode exibir o processo dinâmico de desdobraimento e agregação de proteínas ou peptídeos (CHEN; VAN DER KAMP; DAGGETT, 2014).

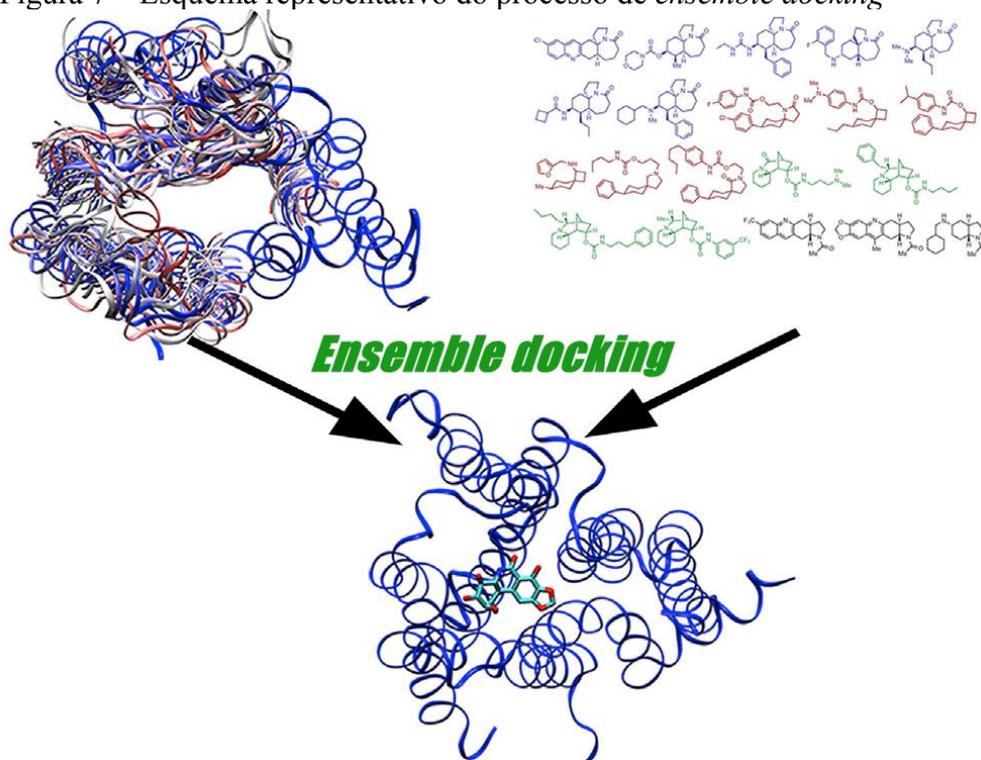
Ainda no campo do planejamento de novos fármacos utilizando sistemas computacionais, as simulações de DM combinadas com cálculos de energia livre de ligação também demonstraram ser uma boa estratégia para melhorar o fator de enriquecimento da triagem virtual pois as simulações auxiliam no aumento da precisão da previsão da capacidade de interação entre os ligantes e o alvo (RASTELLI et al., 2010). Para a otimização de compostos e obtenção de moléculas com alta atividade e propriedades ADMET adequadas, entender o reconhecimento entre os ligantes e o alvo é de grande valia. A simulação por DM

combinada com o cálculo da energia livre de ligação foi amplamente aplicada para estudar a interação entre fármacos e alvos (KING et al., 2021a; PROCACCI, 2021).

### 3.2.1 *Ensemble docking e simulações dos novos ligantes derivados de BIO-5192 em VLA-4*

Na abordagem por *ensemble docking*, a estrutura da proteína-alvo é caracterizada, muitas vezes, usando simulações de dinâmica molecular, e várias conformações assim geradas do receptor são usadas em muitos cálculos de *docking*, em vez de realizá-lo em uma única estrutura (experimental ou modelada) do alvo. Essa técnica visa reproduzir um mecanismo de seleção conformacional, no qual conformações específicas de proteínas são selecionadas pelo(s) ligante(s) para formar um complexo proteína-ligante termodinamicamente favorecido, como mostra a figura abaixo.

Figura 7 – Esquema representativo do processo de *ensemble docking*



Fonte: Adaptada de Amaro et al. (2018).

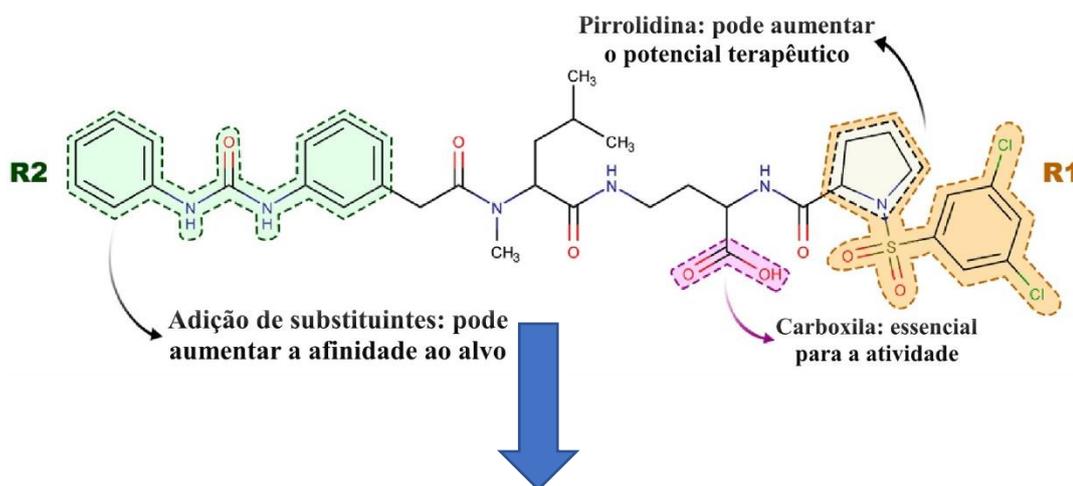
O planejamento de fármacos baseado no *ensemble docking* é aplicado na identificação de novos *hits* e *leads* em uma variedade de vias metabólicas e alvos moleculares. Além dessa abordagem, a técnica baseada em *ensemble* pode ser usada para prever interações de receptores com *hits/leads* fora do sítio no alvo levando a efeitos adversos em ensaios pré-

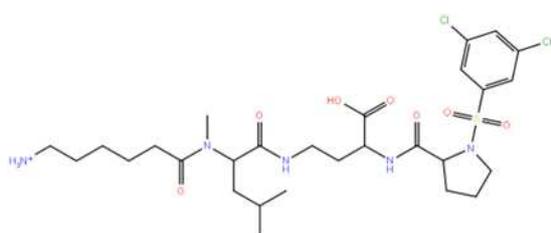
clínicos e clínicos (EVANGELISTA et al., 2016), também para redirecionar medicamentos existentes (ZHAO et al., 2016) e para identificar os alvos de produtos naturais.

O *ensemble docking* é um processo de duas etapas: geração de um conjunto de conformações de proteínas a serem usadas nos cálculos e o encaixe *per se* nas estruturas do receptor selecionadas. Tomadas de forma independente, essas duas tarefas requerem poder computacional significativo, tanto para longas simulações de DM quanto para aplicar em grandes bancos de dados de compostos. Para mitigar o efeito da grande demanda computacional nos cálculos, o que se tem feito é selecionar apenas um subconjunto de conformações de proteínas, por exemplo, agrupando as suas conformações (ou de uma subestrutura, como sítios de ligação) e escolhendo uma estrutura representativa para cada agrupamento. A hipótese subjacente a essa abordagem de *clustering* (agrupamento) é que cada conformação representativa de um *cluster* (grupo) será também representativa de todo o *cluster* em termos de desempenho de *docking* (EVANGELISTA FALCON et al., 2019).

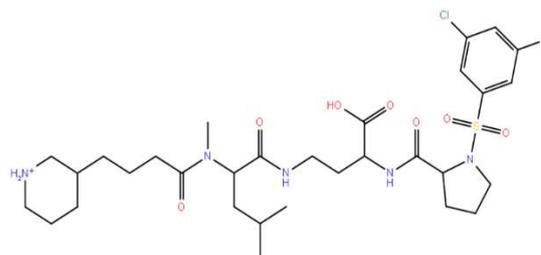
Neste trabalho, foi realizado o método de *ensemble docking* no receptor VLA-4 obtido por modelagem por homologia em trabalho prévio (VASCONCELOS et al., 2022) (apenas o sítio de ligação foi considerado), utilizando como ligantes os compostos de código 1592, 2363 e 973, além do composto-líder BIO-5192, do qual os ligantes derivam. Essas pequenas moléculas são potenciais antagonistas de VLA-4 desenvolvidas pelo grupo em trabalho anterior (VASCONCELOS et al., 2022). A figura abaixo descreve o processo de desenvolvimento dos novos ligantes pela técnica do desenho *de novo*, a partir do *lead* BIO-5192.

Figura 8 – Na parte superior, a molécula BIO-5192, a qual foi o composto-líder utilizado neste trabalho para gerar, através da técnica do desenho *de novo*, os outros ligantes de código 1592, 2363 e 973, que estão na parte inferior da figura





Molécula 1592



Molécula 2363



Molécula 973

Fonte: O autor.

Como podemos observar, os compostos 1592, 2363 e 973 são resultados de variações moleculares a partir da porção R2 de BIO-5192. Tal mudança química favoreceu a interação dos ligantes em VLA-4, além de gerarem estruturas com baixa capacidade de toxicidade, como demonstrado em trabalho prévio (VASCONCELOS et al., 2022).

O método de *ensemble* aconteceu a partir da realização de simulações de DM de VLA-4, a fim de se gerar diversas conformações da proteína. As simulações realizadas neste trabalho foram feitas no pacote Gromacs v. 2018.3 (ABRAHAM et al., 2015; BERENDSEN; VAN DER SPOEL; VAN DRUNEN, 1995).

As simulações foram realizadas em um tempo total de 200 ns de duração. As energias de interação foram tratadas com o campo de força Gromos 53a6 (OOSTENBRINK et al., 2005). O algoritmo de integração de Verlet tratou as interações de Coulomb de e van der Waals com um corte de distância de 10 Å. O método *Particle Mesh Ewald* (PME) (DARDEN; YORK; PEDERSEN, 1998) foi escolhido para tratar as interações eletrostáticas de longo alcance.

Os sistemas foram colocados em uma caixa cúbica e solvatados com água SPC (MARK; NILSSON, 2001). A adição de contra íons neutralizou os complexos. Em seguida, os sistemas foram minimizados seguindo o algoritmo *steepest-descent* e passaram por uma equilíbrio em três etapas. Na primeira, o conjunto NVT (número de mols; volume e temperatura constantes) foi aplicado por 100 ps até que a temperatura atingisse 300 K. Em seguida, usamos duas etapas de simulação no conjunto NPT (número de mols; pressão e

temperatura constantes): uma com restrições de posição por 100 ps e a última aplicando simulação irrestrita por 1 ns.

Ao final das simulações, foi aplicada a função *gmx cluster* do Gromacs para gerar agrupamentos de conformações similares de VLA-4, obtidas ao longo da dinâmica. Utilizamos o *cluster* mais representativo para realizar o *docking* molecular.

Realizamos os ensaios de *docking* usando o programa Autodock Vina (TROTT; OLSON, 2010), implementado no software Pyrx 0.9.2. (DALLAKYAN; OLSON, 2015), uma versão otimizada do Autodock que implementa um novo método de otimização de gradiente. O cálculo desse gradiente efetivamente direciona o algoritmo em relação às menores energias. O Vina é capaz de ser executado em ambientes computacionais com múltiplos núcleos de processamento, aumentando a velocidade e a capacidade de identificação de compostos líder, já que essa ferramenta não considera no cálculo as interações eletrostáticas. O programa Vina consegue ser duas ordens de magnitude mais rápido do que seu antecessor, o AUTODOCK 4 (MORRIS et al., 2009).

Os ligantes BIO-5192, 1592, 2363 e 973 tiveram sua energia minimizada no programa Avogadro (HANWELL et al., 2012). A estereoquímica das moléculas foi respeitada e os estados de protonação dos grupos ionizáveis foram gerados em pH 7,4.

Definimos o centro geométrico da caixa para delimitar a região de encaixe dos ligantes como sendo as coordenadas do íon  $Mg^{2+}$  do motivo MIDAS. As arestas da caixa foram ajustadas em 20 Å, 20 Å e 25 Å para os eixos x, y e z. Definimos um valor do parâmetro exaustividade (do inglês, *exhaustiveness*) do Vina de 150 para este trabalho. Elevando este parâmetro do valor padrão, que é de 8, é permitido ao usuário que se aumente o tempo de busca conformacional linearmente, aumentando a probabilidade de encontrar o mínimo global exponencialmente (informação disponível no Manual do Autodock Vina - <http://vina.scripps.edu/manual.html#exhaustiveness>).

Os melhores resultados foram classificados de acordo com a função de pontuação do Vina. Além disso, classificamos os resultados em relação às distâncias euclidianas entre os átomos de oxigênio do grupo carboxilato do ligante e o íon  $Mg^{2+}$  da região MIDAS da integrina, tomando um raio de corte de 3 Å. Por fim, selecionamos um modo de ligação específico no qual o composto permaneceria em uma forma esticada no sítio da integrina.

Para validarmos o *docking*, estudamos os modos de ligação de 500 compostos que têm alguma atividade em VLA-4 e os ligantes *decoys*, isto é, aqueles que interagem a seu receptor, porém não possuem atividade. O banco de dados de moléculas do ChEMBL (GAULTON et al., 2012) foi usado (ID: ChEMBL1907599). A quantidade total de *decoys*

contidos no banco de dados foi de 28%. Usamos as mesmas configurações do protocolo anterior. Assim, classificamos as moléculas mais bem avaliadas com base na função de pontuação do Vina e checamos a precisão do método na separação de compostos ativos e *decoys* traçando uma Curva Característica de Operação do Receptor (do inglês, *Receiver Operating Characteristic Curve* - ROC) para medir a área sob a curva (do inglês, *Area Under the Curve* - AUC) e o Fator de Enriquecimento (do inglês, *Enrichment Factor* - EF).

Após o protocolo do *ensemble docking*, os complexos integrina-ligantes formados foram submetidos às simulações de DM, a fim de avaliarmos a estabilidade dos sistemas, a forma de interação dos ligantes no sítio do receptor, além de estudarmos modos de movimento da proteína ao longo do tempo para entendermos o que cada composto pode influenciar no comportamento proteico.

O programa Antechamber (WANG et al., 2006) foi usado para preparar os ligantes usando o campo de força Gaff2 (VASSETTI; PAGLIAI; PROCACCI, 2019) e, posteriormente, aplicamos o *script* acpype (SOUSA DA SILVA; VRANKEN, 2012) para gerar os parâmetros e topologias para o pacote Gromacs. O estado de protonação da proteína VLA-4 foi ajustado para pH 7,4 usando a ferramenta H++ (GORDON et al., 2005).

As simulações foram realizadas em triplicata, com cada réplica de 500 ns de duração, totalizando 4,5  $\mu$ s de dados. As energias de interação foram tratadas com o campo de força Amber ff99sb (LINDORFF-LARSEN et al., 2010). O algoritmo de Verlet tratou as interações de Coulomb e de van der Waals com uma distância de corte de 10 Å. O método *Particle Mesh Ewald* (PME) (DARDEN; YORK; PEDERSEN, 1998) foi escolhido para tratar as interações eletrostáticas de longo alcance. Todas as ligações foram limitadas usando o algoritmo LINCS (“LINCS: A linear constraint solver for molecular simulations - Hess - 1997 - Journal of Computational Chemistry - Wiley Online Library”, [s.d.]).

Os sistemas foram colocados em uma caixa cúbica e solvatados com água TIP3P (PRICE; BROOKS, 2004). A adição de três contra íons de sódio neutralizou os complexos. Em seguida, os sistemas foram minimizados seguindo o algoritmo *steepest-descent* e passaram por uma equilibração em três etapas. Na primeira, o conjunto NVT (número de mols; volume e temperatura constantes) foi aplicado por 100 ps até que a temperatura atingisse 300 K. Após isso, usamos duas etapas de simulação no conjunto NPT (número de mols; pressão e temperatura constantes): uma com restrições de posição por 100 ps e a última aplicando simulação irrestrita por 1 ns.

Ao final das simulações, algumas análises foram aplicadas, a partir de ferramentas próprias do pacote Gromacs. Calculamos o RMSD (do inglês, *Root Mean Square*

*Deviation*) da proteína e dos átomos pesados dos ligantes; o RMSF (*do inglês, Root Mean Square Fluctuation*) de VLA-4 ao longo das simulações; as ligações de hidrogênio realizadas entre ligantes e receptor; o SASA (Área de Superfície Acessível ao Solvente) dos complexos ao longo do tempo e as distâncias do grupo COO<sup>-</sup> nas pequenas moléculas e o íon MIDAS da integrina. Esses dados foram submetidos à interpolação por mínimos quadrados. Os resultados foram plotados no *software* GraphPad Prism 8.0.2.

### 3.2.1.1 Análises de Componentes Principais (PCA)

Uma estratégia adotada para o estudo dos modos de movimento da proteína ao longo de simulações por DM e se cada ligante complexado ao receptor interfere nesse processo, é a aplicação dos cálculos de componentes principais.

A Análise de Componentes Principais (PCA) é uma técnica estatística usada para reduzir a dimensionalidade de grandes conjuntos de dados, mantendo o máximo de informações possível. Ela funciona identificando as direções (componentes principais) nos dados que capturam a maior variação e projetando os dados nesses componentes (BRO; SMILDE, 2014; SHLENS, 2014).

A PCA é comumente utilizada em ciência de dados, aprendizado de máquina e outros campos em que dados de alta complexidade precisam ser analisados. Tem aplicações em processamento de imagem, processamento de sinal, finanças e muitas outras áreas (ASOKAN et al., 2020).

A ideia básica da PCA é transformar um conjunto de variáveis correlacionadas em um novo conjunto de variáveis não correlacionadas chamadas de componentes principais. O primeiro componente principal captura a direção na qual os dados variam mais, o segundo componente principal captura a direção que representa a maior variação restante e assim por diante. Cada componente principal é uma combinação linear das variáveis originais e, juntas, formam uma base ortogonal para os dados.

O processo de encontrar os componentes principais envolve o cálculo dos autovalores e autovetores da matriz de covariância dos dados. Os autovetores correspondem aos componentes principais e os autovalores indicam a quantidade de variação capturada por cada componente.

A PCA pode ser usado para visualização de dados, extração de recursos e compactação de dados. Também pode ser combinado com outras técnicas, como agrupamento e classificação, para melhorar o desempenho dos modelos de aprendizado de máquina (HU et

al., 2018).

Um dos benefícios do PCA é que ele pode ajudar a reduzir a complexidade computacional dos algoritmos, reduzindo o número de dimensões nos dados. No entanto, deve ser usado com cautela, pois a interpretação dos componentes principais pode ser difícil e a perda de informações pode afetar a precisão dos resultados.

Neste trabalho, foi utilizado o programa pyPcazip (SHKURTI et al., 2016) para a geração dos autovetores e autovalores. Nos cálculos, usamos o receptor VLA-4 após as simulações de DM. Para diminuir o custo computacional, selecionamos resíduos de aminoácidos que estivessem a uma distância de 8 Å do íon MIDAS em VLA-4, logo fizemos uma análise apenas do sítio de ligação da proteína.

### 3.3 Cálculos de energia livre de ligação

Os cálculos de energia livre de ligação são uma ferramenta essencial para prever a força das interações moleculares entre diferentes biomoléculas, como proteínas, ácidos nucleicos e pequenas moléculas. Esses cálculos usam simulações de dinâmica molecular e métodos de mecânica estatística para estimar as mudanças de energia associadas à ligação entre duas moléculas.

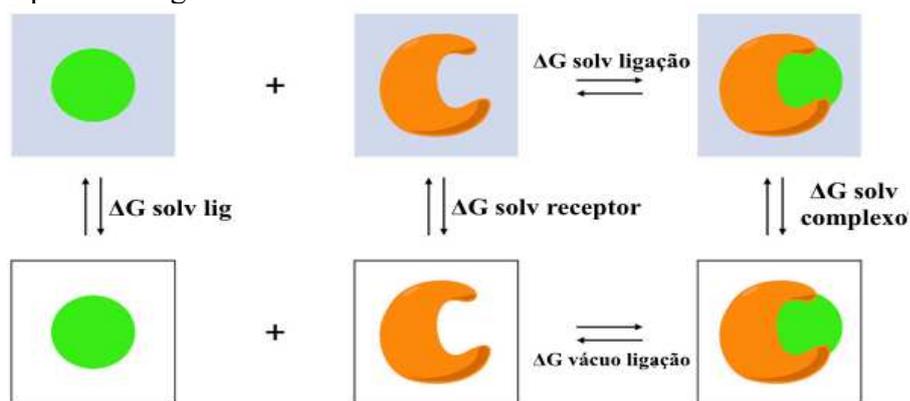
A energia livre de ligação é definida como a diferença entre a energia livre do complexo ligado e a soma das energias livres dos componentes individuais no estado não ligado. O cálculo preciso da energia livre de ligação requer um tratamento abrangente de todas as contribuições de energia relevantes, como eletrostática, forças de van der Waals, ligações de hidrogênio, solvatação e mudanças conformacionais (GILSON; ZHOU, 2007; KOLLMAN et al., 2000; WANG et al., 2006).

Várias abordagens foram desenvolvidas para o cálculo da energia livre de ligação, incluindo integração termodinâmica (TI), perturbação de energia livre (FEP), mecânica molecular/área superficial de Poisson-Boltzmann (MM/PBSA) e energia de interação linear (LIE). Cada método tem suas vantagens e limitações, dependendo do sistema em estudo e dos recursos computacionais disponíveis.

Neste trabalho, aplicamos o método MM/PBSA para os complexos integrina-ligante após as simulações de DM. O método MM/PBSA aplicado à interação de moléculas pequenas é um método que estima a diferença de energia livre de ligação entre o complexo proteína-ligante e os componentes não ligados separados, o complexo, ligante e a proteína sozinha (WANG et al., 2016, 2019) (Figura 9). O MM/PBSA fornece uma abordagem

equilibrada caracterizada por melhorar o rigor e a precisão sobre o *docking* molecular, e com redução da demanda computacional em comparação com outros métodos (SLYNKO et al., 2014; SUN et al., 2014).

Figura 9 – Ciclo termodinâmico do MM/PBSA. A energia livre de ligação em meio aquoso é calculada como a diferença entre a soma da interação no vácuo e a solvatação do complexo com a solvatação do receptor e do ligante individualmente



$$\Delta G_{\text{solv ligação}} = \Delta G_{\text{vácuo ligação}} + \Delta G_{\text{solv complexo}} - (\Delta G_{\text{solv lig}} + \Delta G_{\text{solv receptor}})$$

Fonte: Adaptada de King et al. (2021b).

Duas abordagens principais são empregadas para gerar os dados para previsões de energia de ligação via MM/PBSA com ambas a partir da simulação de dinâmica molecular em solvente explícito: múltiplas trajetórias com os três componentes, complexo, receptor apo e ligante separadamente, ou uma única trajetória com o complexo proteína-ligante que é dividido depois nos três componentes (KOLLMAN et al., 2000; WANG et al., 2016).

A simulação de DM é realizada com solvatação explícita para maximizar a precisão da amostragem conformacional, e os *frames* são pós-processados pela remoção de moléculas de solvente e íons. A trajetória convergente é avaliada com cada *frame* como um ponto de amostra individual para gerar médias de conjunto e valores de incerteza para as quantidades de energia (KING et al., 2021b)

Apesar dos avanços no campo, a previsão precisa da energia livre de ligação continua sendo uma tarefa desafiadora devido à complexidade das interações moleculares subjacentes e às limitações dos métodos de simulação atuais. No entanto, os cálculos de energia livre de ligação tornaram-se uma ferramenta essencial na descoberta de medicamentos e outras aplicações biotecnológicas, pois podem fornecer informações sobre os mecanismos moleculares de ligação e ajudar no planejamento de novos fármacos (MUEGGE; HU, 2023).

Um dos desafios nos cálculos de energia livre é a precisão dos campos de força usados nas simulações. O desenvolvimento de campos de força mais precisos tem sido um dos principais focos de pesquisa no campo da biofísica computacional (CAVASOTTO, 2020).

Outro desafio encontrado na área é o problema de amostragem. Para estimar com precisão a energia livre de ligação, as simulações devem amostrar o espaço conformacional relevante do sistema, o que pode ser uma tarefa computacionalmente intensa. Vários métodos avançados de amostragem, como dinâmica molecular de troca de réplicas e troca de réplicas hamiltonianas, foram desenvolvidos para resolver essa abordagem (MARKTHALER; KRAUS; HANSEN, 2022).

Os métodos de ligação de energia livre têm evoluído rapidamente nos últimos anos, com avanços no poder de computação, melhorias dos algoritmos e o desenvolvimento de novas técnicas. Um avanço recente é o uso de aprendizado de máquina para prever energias livres de ligação. Os modelos podem ser treinados em grandes conjuntos de dados de afinidades de ligação medidas experimentalmente para “aprender” a relação entre características moleculares e energias livres de ligação. Esses cálculos podem então ser usados para prever a afinidade de ligação de novos compostos com alta precisão e velocidade. No entanto, a acuidade desses modelos é limitada pela qualidade e diversidade dos dados de treinamento (BADAOUÏ et al., 2022).

### ***3.3.1 Cálculos de energia livre de ligação nos complexos $\alpha\beta 1$ -ligantes***

Ao final das simulações, as trajetórias de cada réplica dos complexos VLA-4 com os ligantes BIO-5192, 1592, 2363 e 973 foram submetidas a cálculos de energia livre de ligação por MM/PBSA, usando o programa GMXPBSA 2.1 (PAISSONI et al., 2015).

Esse *script* calcula os termos provenientes da mecânica molecular (Coulomb e Lennard-Jones), além das energias de solvatação polar e não polar, a partir dos dados de DM. O programa não reproduz valores absolutos de energia livre de ligação e ao final, faz um *ranking* de afinidades relativas entre complexos de proteínas com diferentes ligantes.

Para os cálculos, foram considerados somente os últimos 250 ns de simulação para cada réplica dos complexos integrina-ligante sem as moléculas de solvente. Os dados da decomposição de energias dos complexos foram tratados e submetidos a análises estatísticas no programa GraphPad Prism 8.0.2. Por fim, o parâmetro  $\Delta\Delta G$  (energia livre de ligação relativa) de cada complexo foi calculado e analisado qual composto possuía maior afinidade a VLA-4.

## 4 METODOLOGIA (PARTE II)

### 4.1 Reposicionamento de fármacos

O reaproveitamento de medicamentos (também chamado de reposicionamento) é uma estratégia para identificar novos usos para medicamentos aprovados ou em investigação que estão fora do escopo da indicação médica original (ASHBURN; THOR, 2004). Essa estratégia oferece várias vantagens sobre o desenvolvimento de um medicamento totalmente novo para uma determinada indicação. Primeiro, e talvez o mais importante, o risco de falha é menor; como o medicamento reaproveitado já foi considerado suficientemente seguro em modelos pré-clínicos e em humanos, se os ensaios em estágio inicial tiverem sido concluídos, é menos provável que falhe, pelo menos do ponto de vista da segurança, em ensaios de eficácia subsequentes. Em segundo lugar, o prazo para o desenvolvimento do medicamento pode ser reduzido, porque a maior parte dos testes pré-clínicos, avaliação de segurança e, em alguns casos, o desenvolvimento da formulação já terá sido concluído. Em terceiro lugar, é necessário menos investimento, embora isso varie muito, dependendo do estágio e do processo de desenvolvimento do candidato a reaproveitamento (BRECKENRIDGE; JACOB, 2019).

Os custos regulatórios e de fase III podem permanecer os mesmos para um medicamento reaproveitado e para um novo medicamento na mesma indicação, mas ainda pode haver economias substanciais nos custos pré-clínicos e nas fases I e II (PUSHPAKOM et al., 2019). Juntas, essas vantagens têm o potencial de resultar em um retorno de investimento menos arriscado e mais rápido no desenvolvimento de medicamentos reaproveitados, com custos médios associados mais baixos, uma vez que as falhas foram contabilizadas (de fato, os custos de trazer um medicamento reaproveitado para o mercado é estimado em US\$ 300 milhões em média, em comparação com cerca de US\$ 2 a 3 bilhões para uma nova entidade química) (NOSENGO, 2016). Finalmente, fármacos reaproveitados podem revelar novos alvos e caminhos que podem ser mais explorados.

Historicamente, o reposicionamento de medicamentos tem sido amplamente fortuito; uma vez que se descobriu que um fármaco tinha um efeito fora da indicação da bula ou um efeito recém-reconhecido, ele era levado para exploração comercial. De fato, os exemplos mais bem-sucedidos de reaproveitamento de medicamentos até agora não envolveram uma abordagem sistemática; o reposicionamento do citrato de sildenafil para disfunção erétil baseou-se na experiência clínica retrospectiva, e o reaproveitamento da

talidomida para o eritema nodoso hansênico (ENH) e mieloma múltiplo foi baseado no acaso (ASHBURN; THOR, 2004).

A sildenafil foi originalmente desenvolvida como um medicamento anti-hipertensivo, mas quando reaproveitado pela Pfizer para o tratamento da disfunção erétil e comercializado como Viagra, detinha uma participação de mercado de 47% no mercado de medicamentos para disfunção erétil em 2012, com vendas mundiais totalizando US\$ 2,05 bilhões (“Pfizer’s Expiring Viagra Patent Adversely Affects Other Drugmakers Too”, [s.d.]).

A talidomida, um sedativo originalmente comercializado em alguns países em 1957, foi retirada do mercado em 4 anos devido à sua ligação com defeitos congênitos esqueléticos graves em crianças nascidas de mães que tomaram o fármaco durante o primeiro trimestre de gravidez. No entanto, descobriu-se por acaso que era eficaz primeiro no tratamento do ENH, em 1964, e décadas depois no mieloma múltiplo (SINGHAL et al., 1999). Desde então, teve um sucesso comercial substancial no mieloma múltiplo e levou ao desenvolvimento e aprovação de derivados ainda mais bem-sucedidos, como a lenalidomida (Revlimid, Celgene), que teve vendas globais em 2017 de US\$ 8,2 bilhões (URQUHART, 2018).

A partir do surgimento do surto pandêmico das novas infecções virais, como a do novo coronavírus (COVID-19), a comunidade científica começou uma corrida contra o tempo para desvendar soluções rápidas e eficientes para a doença emergente. Além disso, o desenvolvimento de uma vacina para qualquer doença, incluindo COVID-19, leva tempo e, mesmo que o processo seja colocado em modo acelerado, levaria de 18 a 20 meses para introduzi-lo como produto pronto para uso. Diante disso, uma das estratégias adotadas como solução dessas questões é o uso do reposicionamento de fármacos.

Um exemplo bem-sucedido de reaproveitamento foi visto recentemente na aprovação do FDA do remdesivir (Veklury; Gilead Sciences) para o tratamento da COVID-19 (<https://www.fda.gov/drugs/drug-safety-and-availability/fdas-approval-veklury-remdesivirtreatment-covid-19-science-safety-and-effectiveness>) abrindo caminho para muitos medicamentos aprovados e em investigação com potencial de reaproveitamento em ensaios clínicos para COVID-19, como dexametasona (P et al., 2021) e tocilizumabe (TONIATI et al., 2020).

A Tabela 1 abaixo mostra outros exemplos de reaproveitamento de medicamentos bem-sucedidos, juntamente com as abordagens de reposicionamento empregadas, a maioria dos quais até agora deriva de um entendimento da farmacologia do medicamento ou de análises retrospectivas do efeito clínico de um medicamento quando prescrito para sua

indicação original.

Tabela 1 – Alguns fármacos comerciais que foram submetidos ao processo de reposicionamento

<b>Fármaco</b>	<b>Indicação Original</b>	<b>Nova indicação</b>	<b>Data da aprovação</b>	<b>Abordagem de reposicionamento utilizada</b>
Zidovudina	Câncer	HIV/AIDS	1987	Varredura <i>in vitro</i> de bancos de compostos
Minoxidil	Hipertensão	Perda capilar	1988	Análise clínica retrospectiva (identificação do crescimento capilar como efeito adverso)
Rituximabe	Câncer	Artrite reumatoide	2006	Análise clínica retrospectiva (remissão de artrite reumatoide coexistente em pacientes com linfoma não-Hodgkin tratados com rituximabe)
Cetoconazol	Infecções fúngicas	Síndrome de Cushing	2014	Análise Farmacológica
Aspirina	Analgesia	Câncer colorretal	2015	Análise clínica e farmacológica retrospectiva

Fonte: Adaptada de Pushpakom et al. (2019).

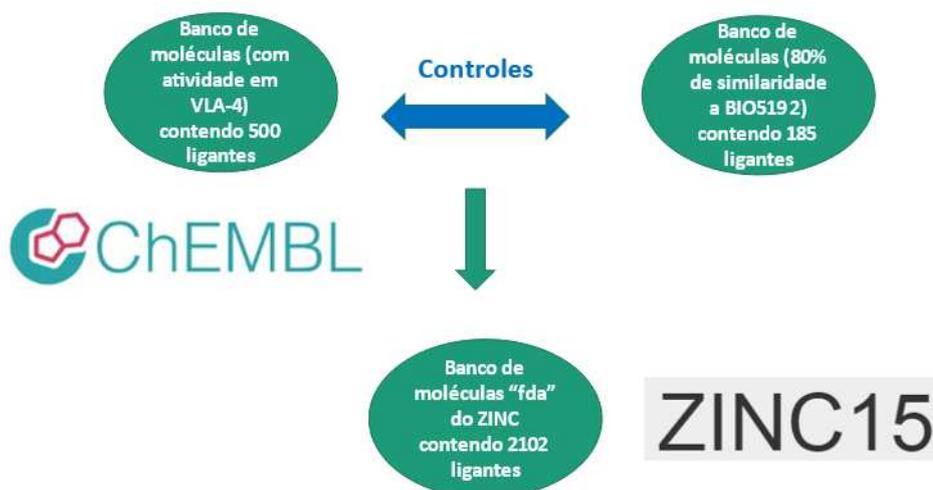
#### 4.1.1 Abordagens computacionais no reposicionamento de fármacos

Abordagens computacionais são amplamente orientadas por grandes volumes de dados; eles envolvem análise sistemática (como expressão gênica, estrutura química, genótipo ou dados proteômicos), que podem levar à formulação de hipóteses de reaproveitamento de medicamentos (HURLE et al., 2013).

Neste trabalho, foi utilizada a abordagem computacional do *docking* molecular, mais precisamente a técnica de varredura virtual ou *Virtual Screening* (VS), para o reposicionamento de medicamentos aprovados pela FDA usando como receptor a integrina VLA-4, a mesma estrutura utilizada na Parte I da Metodologia. Nesse sentido, utilizamos o banco de moléculas chamado “fda” do ZINC (STERLING; IRWIN, 2015) para aplicarmos o VS. Foram aproximadamente 2000 compostos analisados em interação com VLA-4, no programa Vina. Os parâmetros do *docking* aplicados foram os mesmos já implementados nas outras metodologias deste trabalho, com exceção do quesito *exhaustiveness* que, nesse caso, foi ajustado para o valor padrão de 8, já que no VS trabalhamos com maiores quantidades de ligantes. Os resultados foram analisados de acordo com a função de pontuação do Vina e com os modos de ligação gerados.

A fim de validar esse protocolo de VS, decidimos montar um esquema conforme a Figura 10.

Figura 10 – Esquema proposto para a validação do protocolo de VS no método de reposicionamento de fármacos na integrina VLA-4



Esmiuçando melhor a figura 10, fizemos dois VS como controles: um, com um banco de compostos com atividade em VLA-4 e outro, com um banco apresentando moléculas com, no mínimo, 80% de similaridade ao *lead* BIO-5192. Ambos os bancos de ligantes foram extraídos da base de dados ChEMBL. Após isso, foram separados os compostos ativos dos inativos/*decoys*, a partir de um valor de corte de IC50 de 150 nM. Diante disso, foram geradas as curvas ROC e calculados os AUCs e os Fatores de Enriquecimento. Os modos de ligação gerados foram analisados, em conjunto com as interações formadas entre receptor e ligante.

Desse modo, 3 dos melhores fármacos classificados após o VS no Vina foram selecionados para as simulações de DM. Foram considerados os modos de ligação de menores energias. As simulações seguiram o mesmo protocolo de parametrizações da parte I. As exceções foram que, neste caso, utilizamos um sistema de caixa triclínica com a adição de íons neutralizantes a uma concentração fisiológica de 0,15 mol/l. O campo de força aplicado foi o Amber ff14sb (MAIER et al., 2015). Outros diferenciais do protocolo foram o tempo de simulação e o número de réplicas. Neste caso, foram estabelecidas 2 réplicas de 100 ns cada uma. Ao final, foi avaliada a estabilidade dos fármacos no sítio de ligação de VLA-4.

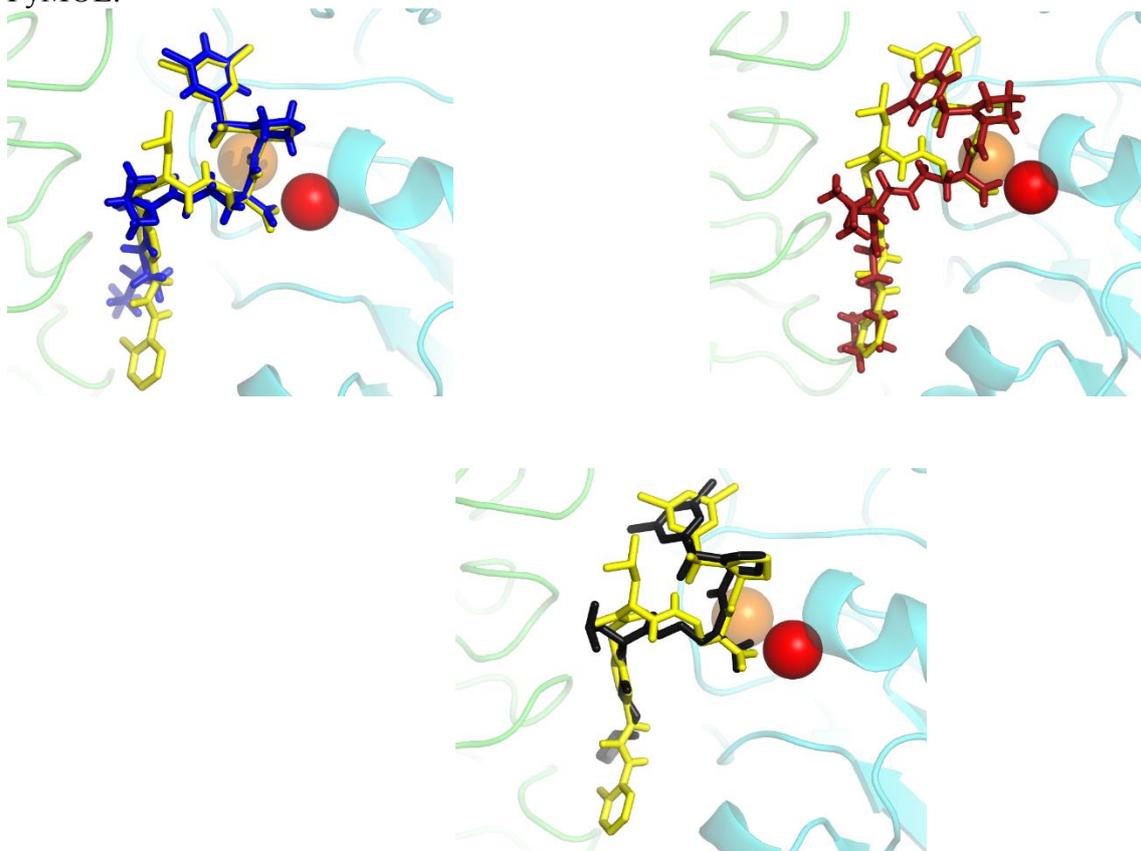
Após os dados de DM terem sido colhidos e avaliados, os sistemas foram submetidos a cálculos de energia livre de ligação pelo método do MM/PBSA. O mesmo programa GMXPBSA 2.1 utilizado na metodologia com os novos compostos derivados de BIO-5192 foi aplicado. Os últimos 50 ns de simulação foram levados em consideração nos cálculos.



VLA-4 devem apresentar um acceptor de ligação de hidrogênio, por exemplo, uma porção carboxilato e grupos hidrofóbicos para acomodar os bolsões formados por aminoácidos apolares (BAIULA et al., 2019b).

Analisando os dados gerados pelo *ensemble docking*, selecionamos os melhores modos de ligação dos novos compostos derivados de BIO-5192 (1592, 2363 e 973) seguindo critérios de sobreposição de grupamentos químicos em comparação com o composto-líder e de interação carboxilato-íon MIDAS (Figura 12). Os valores da função de pontuação do Vina para os modos de ligação selecionados foram: 1592 = -7,9 kcal/mol; 2363 = -8,9 kcal/mol e 973 = -8,4 kcal/mol.

Figura 12 – Modos de ligação dos ligantes 1592 (azul), 2363 (vermelho) e 973 (preto) em VLA-4, em comparação com o líder BIO-5192 (amarelo). As esferas vermelhas representam os íons divalentes presentes na cadeia  $\beta$  da integrina. As cadeias  $\alpha$  e  $\beta$  de VLA-4 estão representadas em verde e azul, respectivamente. As imagens foram geradas no programa PyMOL.



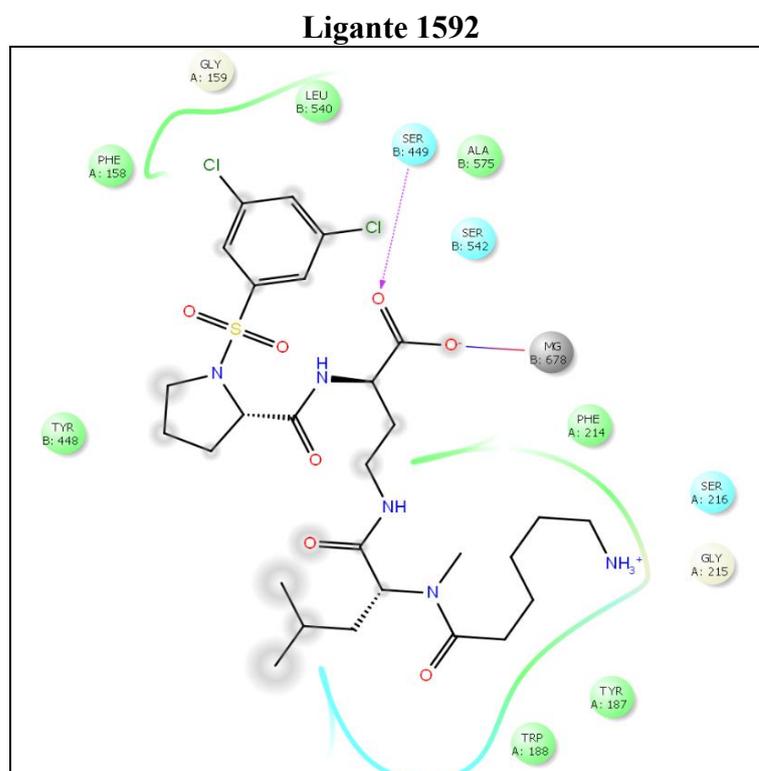
Fonte: O autor.

Analisando esses modos de ligação, podemos identificar que os novos compostos apresentaram boa sobreposição dos seus radicais químicos em comparação a BIO-5192, além

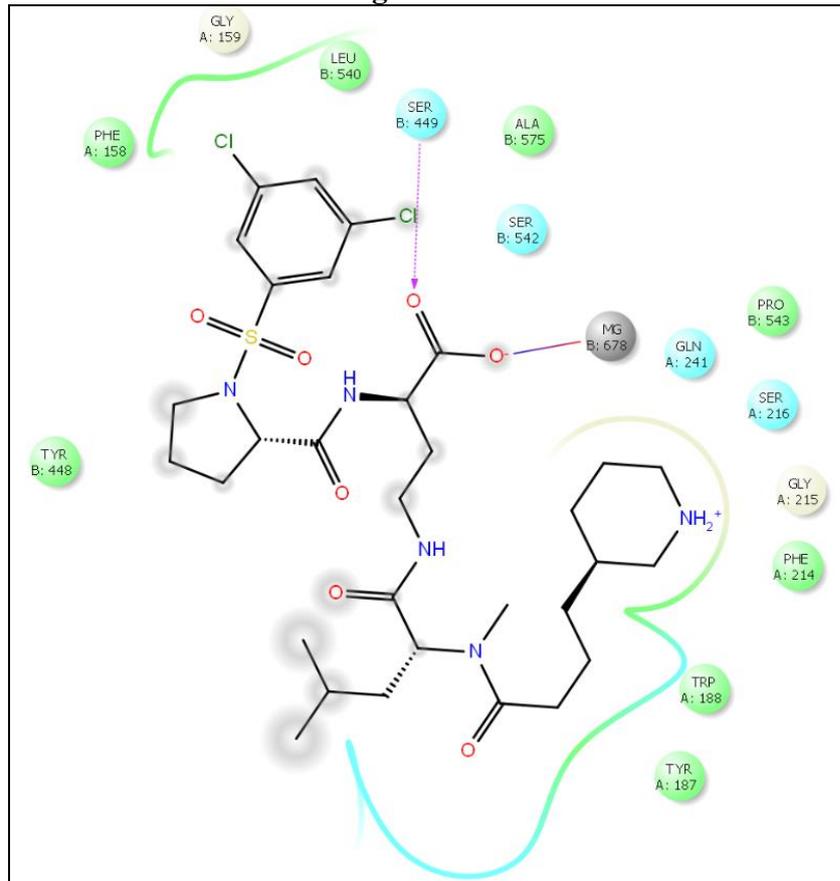
de mostrarem estrita interação do  $\text{COO}^-$  com o íon MIDAS na integrina, mantendo a média de 2,5 Å de distância (Figura 13). O ligante 1592, em azul, apresentou forte correlação dos radicais R1 (anel da prolina - pirrolidina), na parte superior do sítio de interação, que é o mesmo de BIO-5192, e R2 (hexilamina), na parte inferior do bolsão, confirmando a sobreposição do *docking*. Outra constatação é a de que nos modos de ligação preditos, os compostos peptideomiméticos adotam a forma esticada de “S” reverso, abrangendo a interface entre as subunidades  $\alpha 4$  e  $\beta 1$ .

A figura abaixo revela os padrões das interações conservadas entre os derivados de BIO-5192 e VLA-4 (Figura 13). Os resultados sugerem a manutenção do núcleo de interação carboxilato e íon MIDAS, além da ligação de hidrogênio da SER 449 com o oxigênio carboxilato dos ligantes. Na porção R2, onde os novos substituintes químicos foram adicionados, percebemos a presença de aminoácidos apolares, como Tyr187 e Trp188 (esferas verdes – Figura 13) sugerindo as interações hidrofóbicas centradas na parte inferior do sítio de ligação na integrina.

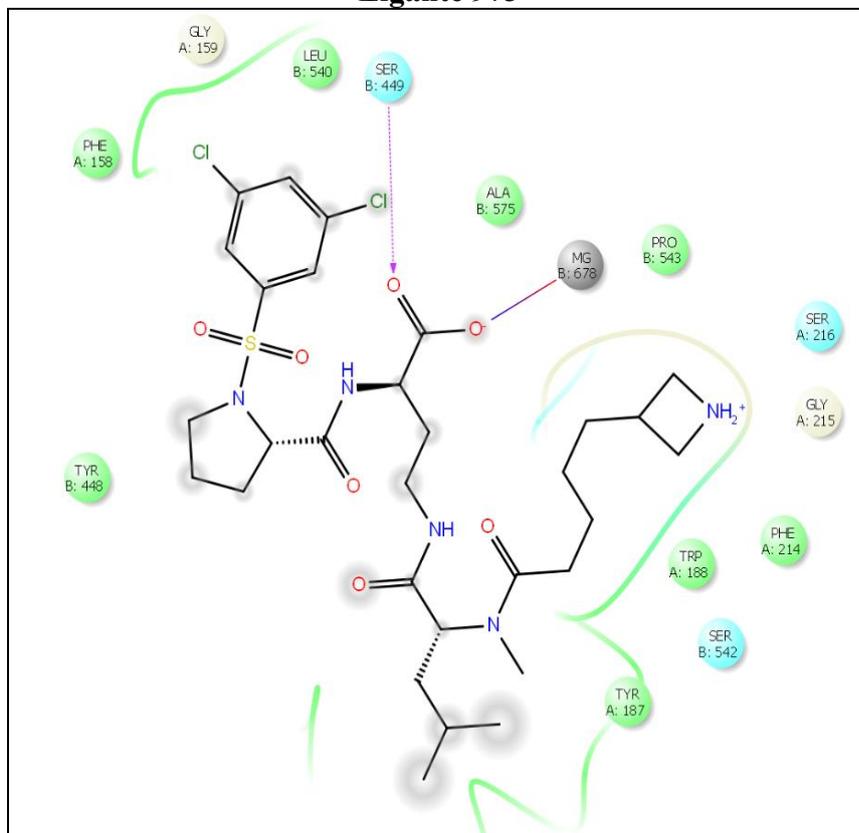
Figura 13 – Diagrama de interações resultantes do *docking* molecular dos ligantes 1592, 2363 e 973 em VLA-4. Figura produzida no Maestro 11.8, com raio de corte de 4 Å



### Ligante 2363



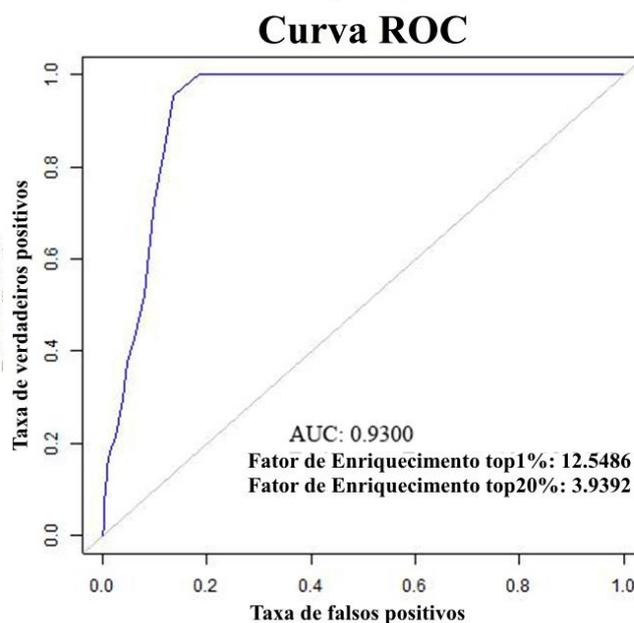
### Ligante 973



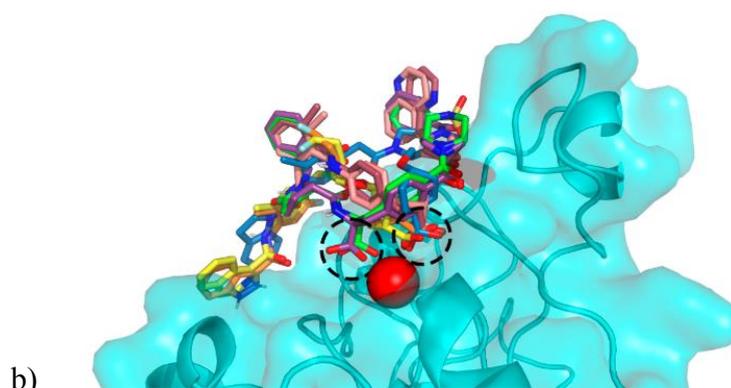
Fonte: O autor.

Os resultados da validação do *docking* sugerem um bom desempenho do método em separar os compostos ativos dos *decoys*, em um banco de ligantes de VLA-4 extraído da base de dados ChEMBL, através do parâmetro  $AUC = 0,93$  (Figura 14a). Isto é, o método é capaz de ranquear as moléculas com os melhores *scores* do Vina nas primeiras posições. Além disso, os valores de  $EF = 12,55$  e  $EF = 3,93$  mostram a capacidade do método de classificar antecipadamente os compostos ativos quando porcentagens de 1% e 20% de moléculas inativas são encontradas, respectivamente, no banco de dados. Podemos observar também que as poses das moléculas melhor classificadas mantiveram a interação da ponte salina entre o oxigênio carboxilato e o íon MIDAS  $Mg^{2+}$ , e todos os ligantes permaneceram em uma conformação esticada dentro do sítio de ligação (Figura 14b).

Figura 14 – a) Curva ROC do protocolo de *docking* e os parâmetros calculados AUC e Fator de Enriquecimento para os primeiros 1% e 20% da base de dados de ligantes de VLA-4. b) Modos de ligação dos compostos mais bem classificados selecionados no banco de dados ChEMBL. Os círculos tracejados indicam o grupo carboxilato das moléculas e a esfera vermelha corresponde ao íon MIDAS  $Mg^{2+}$  na cadeia beta da integrina (superfície ciano). A imagem foi produzida usando o PyMOL



a)



Fonte: O autor.

A tabela abaixo mostra alguns parâmetros calculados dos 10 compostos mais bem ranqueados pelo Vina, utilizados no protocolo de validação do *docking*. Como podemos constatar, entre 10 moléculas nas primeiras posições, apenas 2 *decoys* foram incluídos pelo cálculo do Vina. Logo, pela sobreposição dos modos de ligação (Figura 14b) e pelos dados da Tabela 2, concluímos que o protocolo de docking aplicado neste trabalho está validado.

Tabela 2 – Avaliação dos dez ligantes de VLA-4 melhor classificados de acordo com o Vina *score* e as atividades medidas pelos valores correspondentes de IC50 (Meia Concentração Inibitória Máxima). As informações da atividade foram retiradas do banco de dados ChEMBL

Posição	Vina Score (kcal/mol)	Referência ChEMBL	IC50 (nM)	Atividade
1	-10.2	1940754	4.7	Ativo
2	-10.2	285374	11	Ativo
3	-9.9	2391499	430	<i>Decoy</i>
4	-9.9	1940758	43	Ativo
5	-9.9	2391310	760	<i>Decoy</i>
6	-9.9	280746	3	Ativo
7	-9.9	567180	8.5	Ativo
8	-9.8	572802	32	Ativo
9	-9.8	569302	0.2	Ativo
10	-9.8	566091	2.7	Ativo

Fonte: O autor.

## 5.2 Análises das simulações de BIO-5192 e seus derivados em VLA-4

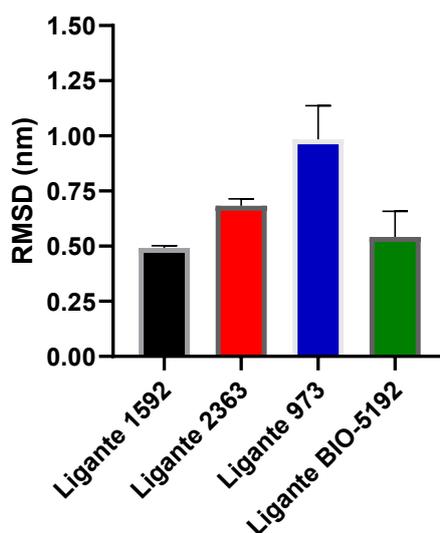
Aplicamos a ferramenta de DM para estudar a estabilidade dos complexos integrina-ligante e se suas interações foram mantidas ao longo do tempo. Escolhemos o composto líder BIO-5192 e três novos ligantes derivados promissores, 1592, 2363 e 973, para analisarmos as simulações.

Em relação ao RMSD dos átomos pesados dos ligantes analisados (Figura 15a),

estimativa que mede o quanto as estruturas das moléculas se desviaram dos modos de ligação iniciais do *docking*, notamos que, para as três réplicas de 500 ns cada, a média do RMSD permaneceu próxima de 0,5 nm para o BIO-5192 e para a molécula 1592, indicando estabilidade para os sistemas. Para este último composto, inclusive, notamos somente que a segunda replicata extrapolou as demais, apresentando valores de RMSD maiores e com mais flutuações (Figura 15b). Entretanto, o ligante 973 apresentou um deslocamento significativo do sítio de interação na integrina ao longo das simulações, com uma média superior a 1 nm contando com o desvio-padrão, indicando que a molécula ocupou outros espaços no bolsão da proteína. Podemos perceber, na Figura 15c, que, nas 3 réplicas, o ligante 973 não apresentou estabilidade estrutural, mostrando um pico no RMSD na segunda réplica, próximo a 400 ns de simulação.

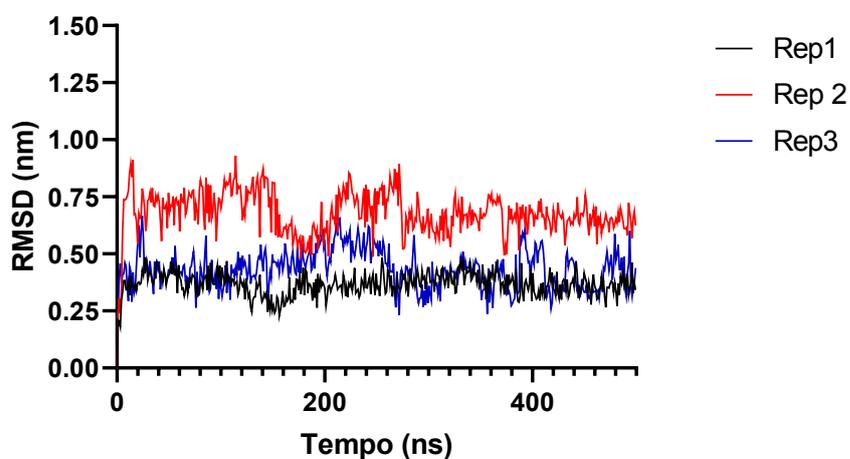
Figura 15 – a) Interpolação linear dos dados de RMSD dos átomos pesados em três réplicas de 500 ns, para cada um dos ligantes: 1592 (preto) (Média = 0,4916 nm e DP = 0,1819 nm); 2363 (vermelho) (Média = 0,6826 nm e DP = 0,1332 nm); 973 (azul) (Média = 0,9845 nm e DP = 0,3165 nm) e BIO-5192 (verde) (Média = 0,5416 nm e DP = 0,1364 nm). b) Dados de RMSD dos átomos pesados do ligante 973, para 500 ns de simulação, conforme réplicas indicadas em preto (Réplica 1), vermelho (Réplica 2) e azul (Réplica 3). Gráficos produzidos no programa GraphPad Prism 8.0.2. DP = Desvio-Padrão

#### Análise do RMSD dos átomos pesados



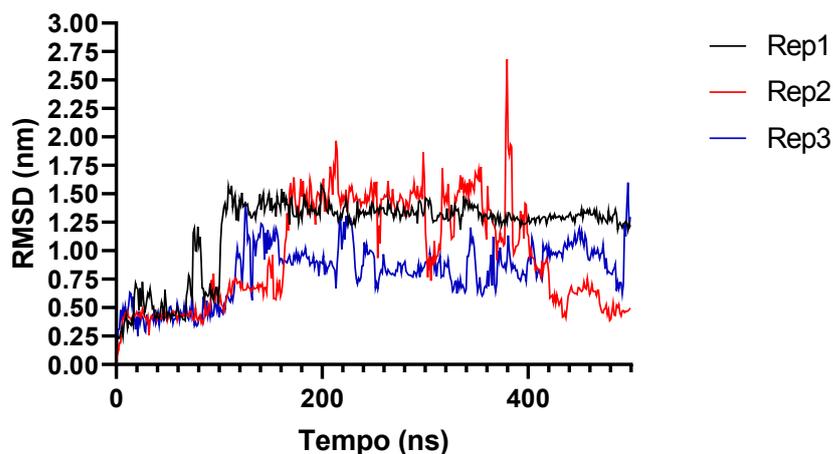
a)

### RMSD dos átomos pesados - Ligante 1592



b)

### RMSD dos átomos pesados - Ligante 973



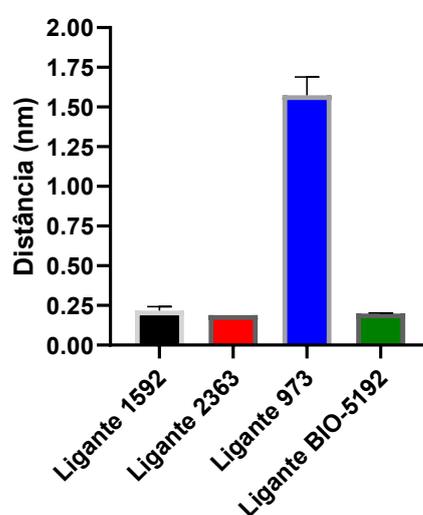
c)

Fonte: O autor.

Ao monitorar a interação entre o íon MIDAS  $Mg^{2+}$  em VLA-4 e o carboxilato de cada ligante (Figura 16a), podemos observar a preservação desse reconhecimento crítico ao longo do tempo para BIO-5192 e seus derivados 1592 e 2363 com uma distância média de aproximadamente 0,25 nm, reforçando a relevância da ponte salina no reconhecimento molecular envolvendo integrinas. No caso do 1592, há uma manutenção da interação formada em todas as réplicas analisadas, durante todo o tempo de simulação (Figura 16b). No entanto, para o ligante 973, a interação foi perdida (Figura 16c), assim como indicado nos valores elevados de RMSD, reforçando a hipótese de que esse composto prefira ocupar outras regiões no sítio de ligação de VLA-4, diminuindo assim sua afinidade com o receptor.

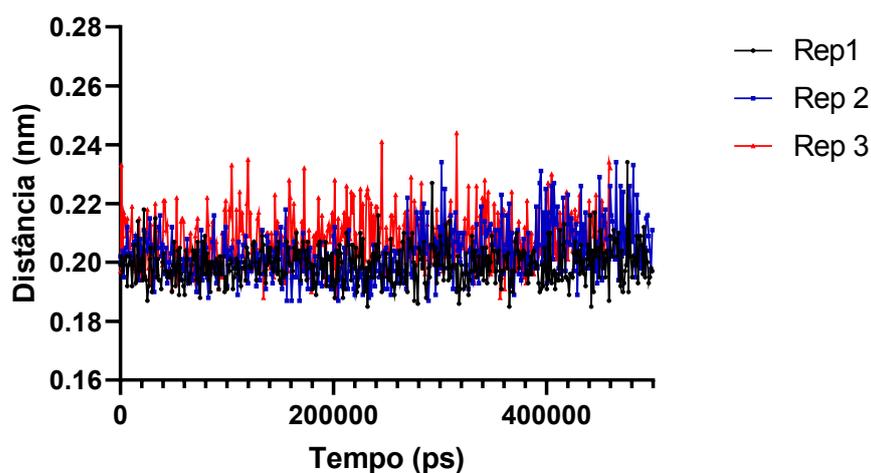
Figura 16 – a) Interpolação linear de dados da distância entre o oxigênio carboxilato e a integrina VLA-4 em três réplicas de 500 ns, para cada um dos ligantes: 1592 (preto) (Média = 0,2177 nm e DP = 0,1002 nm); 2363 (vermelho) (Média = 0,1898 nm e DP = 0,0058 nm); 973 (azul) (Média = 1,5739 nm e DP = 0,4118 nm) e BIO-5192 (verde) (Média = 0,1986 nm e DP = 0,0062 nm). b) Distâncias entre o oxigênio carboxilato do ligante 973 e o íon MIDAS  $Mg^{2+}$  de VLA-4, ao longo de 500 ns de simulação, conforme as réplicas indicadas em preto (Réplica 1), azul (Réplica 2) e vermelho (Réplica 3). Gráficos produzidos no programa GraphPad Prism 8.0.2. DP = Desvio-Padrão

**Distância entre  $COO^-$  e MIDAS  $Mg^{2+}$**



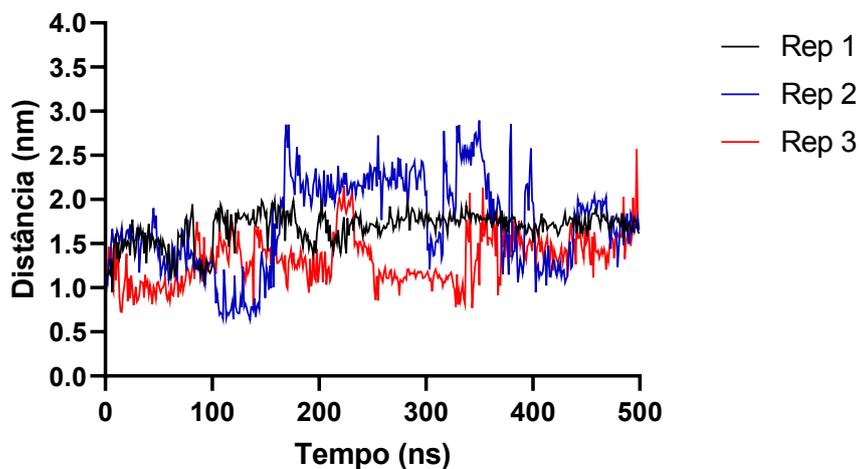
a)

**Distância entre  $COO^-$  e MIDAS  $Mg^{2+}$  - Ligante 1592 e VLA-4**



b)

### Distância entre COO<sup>-</sup> e MIDAS Mg<sup>2+</sup> - Ligante 973 e VLA-4

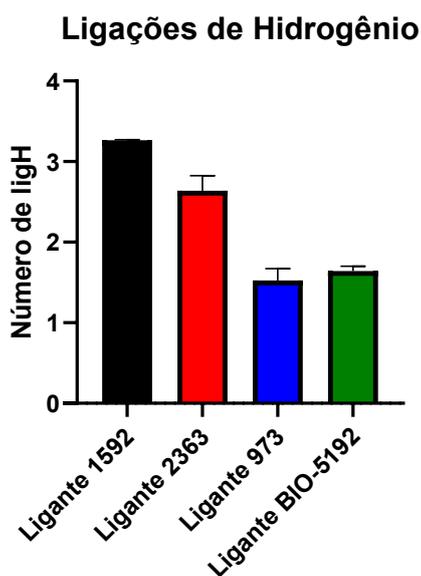


c)

Fonte: O autor.

Ainda no intuito de avaliarmos as interações e o comportamento dos ligantes no sítio de VLA-4, estudamos a quantidade de ligações de hidrogênio realizadas ao longo das simulações para os complexos analisados (Figura 17).

Figura 17 – Interpolação linear de dados do número de ligações de hidrogênio em três réplicas de 500 ns, para cada um dos ligantes: 1592 (preto) (Média = 3,2664 e DP = 1,570); 2363 (vermelho) (Média = 2,6382 e DP = 1,634); 973 (azul) (Média = 1,5256 e DP = 1,225) e BIO-5192 (verde) (Média = 1,6412 e DP = 1,124). Gráficos produzidos no programa GraphPad Prism 8.0.2. DP = Desvio-Padrão



Fonte: O autor.

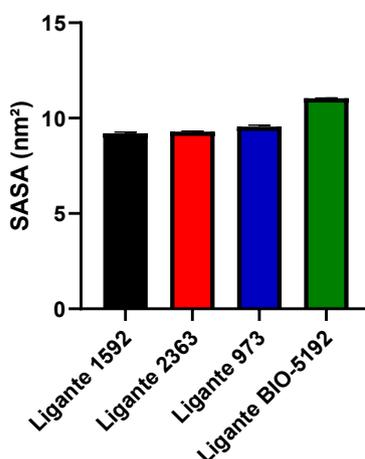
Percebemos que os ligantes 1592 e 2363 despontam como aqueles que fazem um maior número de ligações de hidrogênio com o receptor VLA-4, até mesmo em comparação com o líder BIO-5192. Já a molécula 973 aparece em nível de igualdade com o composto-líder, assumindo o desvio-padrão.

As ligações de hidrogênio são interações químicas importantes que ocorrem entre grupos doadores e aceptores de hidrogênio em moléculas diferentes. Em muitos casos, as ligações de hidrogênio são responsáveis pela especificidade e afinidade de uma interação ligante-proteína, uma vez que a formação dessas ligações pode favorecer uma conformação específica do ligante ou da proteína. Além disso, as ligações de hidrogênio são relativamente fortes em comparação com outras interações não covalentes, como hidrofóbicas ou de van der Waals, o que pode contribuir para a estabilidade do reconhecimento ligante-proteína. Portanto, quanto mais ligações de hidrogênio um composto faz com sua proteína receptora, maior a probabilidade de uma interação forte e específica entre eles.

Ainda com o propósito de avaliarmos o quanto cada ligante interfere na estrutura do receptor, além de analisarmos as interações existentes entre alvo-ligante, um parâmetro frequentemente avaliado é a Área de Superfície Acessível ao Solvente (SASA). Trata-se de uma medida da área de superfície exposta à água de uma molécula de proteína, que pode fornecer informações valiosas sobre a relação proteína-ligante (Figura 18).

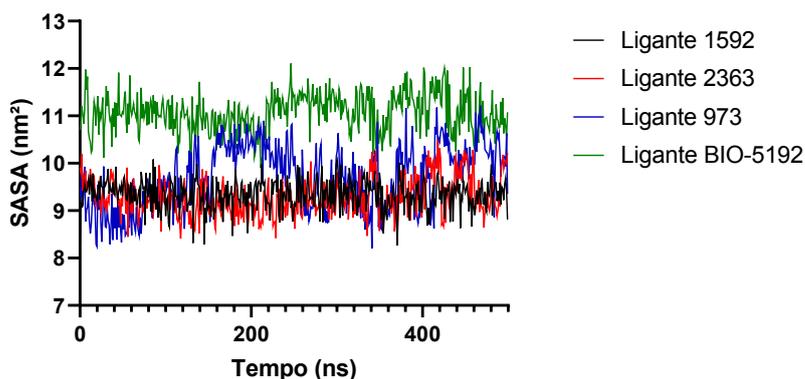
Figura 18 – Interpolação linear de dados da área de superfície acessível ao solvente em três réplicas de 500 ns, para cada um dos ligantes: 1592 (preto) (Média = 9,2094 nm<sup>2</sup> e DP = 0,4821 nm<sup>2</sup>); 2363 (vermelho) (Média = 9,2973 nm<sup>2</sup> e DP = 0,4417 nm<sup>2</sup>); 973 (azul) (Média = 9,5556 nm<sup>2</sup> e DP = 0,7847 nm<sup>2</sup>) e BIO-5192 (verde) (Média = 11,0479 nm<sup>2</sup> e DP = 0,3768 nm<sup>2</sup>). Gráficos produzidos no programa GraphPad Prism 8.0.2. DP = Desvio-Padrão

#### Área de Superfície Acessível ao Solvente (SASA)



a)

### Área de Superfície Acessível ao Solvente (SASA)



b)

Fonte: O autor.

Uma área superficial maior, logo mais exposição de resíduos ao solvente, pode levar a um reconhecimento mais fraco, pois pode reduzir o número de interações favoráveis entre proteína e ligante. Por outro lado, valores de SASA menores podem levar a uma seletividade maior, pois pode acontecer a restrição da entrada de ligantes alternativos (EYAL et al., 2004).

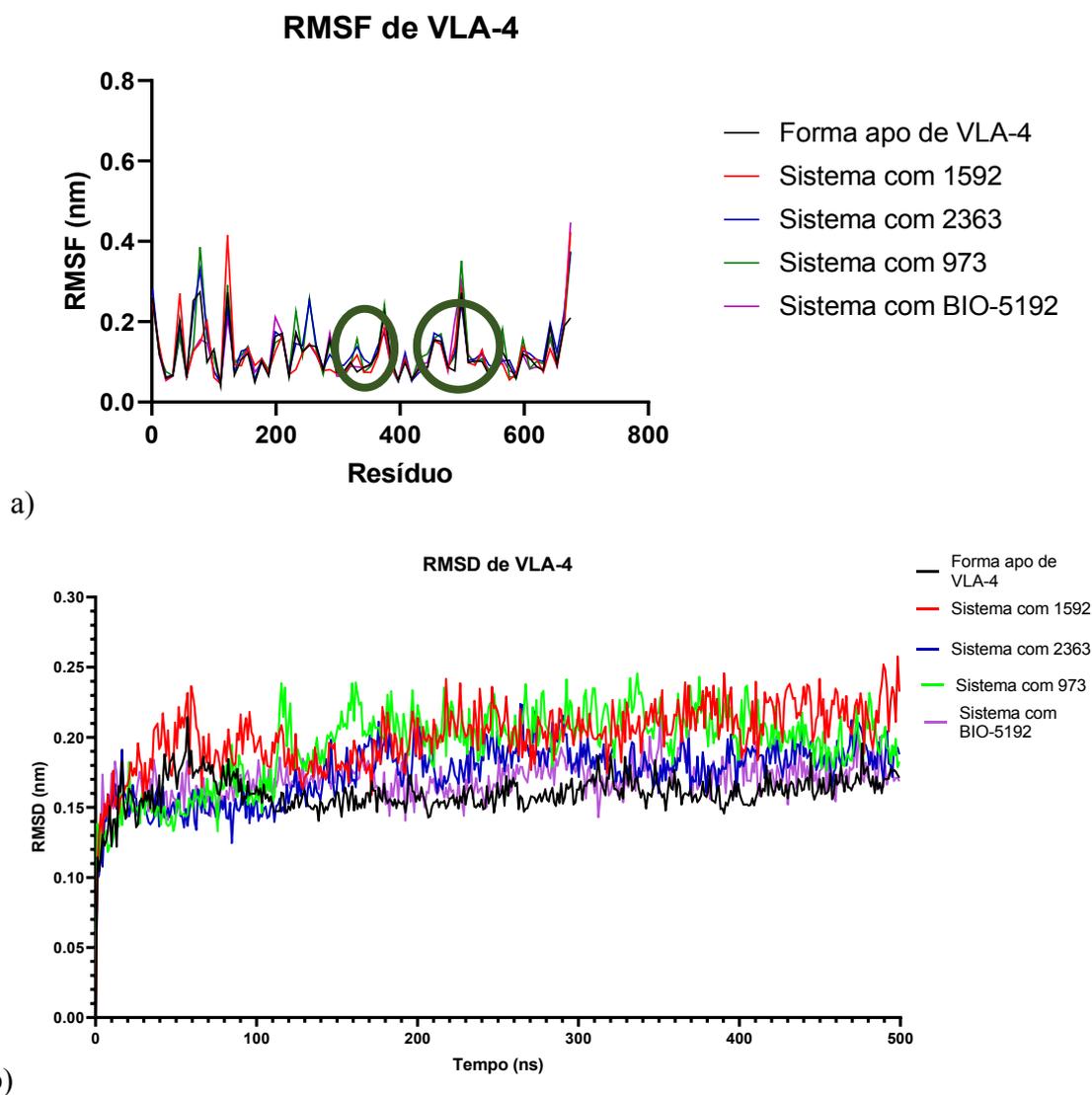
Podemos observar que o complexo com o ligante BIO-5192 apresenta, em média, uma SASA maior ao longo do tempo de simulação. Os outros complexos mostram valores de SASA próximos e menores que os do composto-líder, sem grandes variações, apesar de alguns picos apresentados pelo complexo com a molécula 973 (Figura 18b).

Além disso, a SASA também pode ser usada para avaliar a mudança conformacional da proteína após a interação do ligante. Na Figura 18, por exemplo, observamos que os derivados de BIO-5192 causaram mudanças nos valores de SASA em comparação ao líder e, logo, hipotetizamos alterações na estrutura de VLA-4 nesse fenômeno. Desse modo, a partir dos tópicos seguintes estudaremos essas modificações.

### 5.3 Análises conformacionais de VLA-4 a partir das simulações de DM

Analisaremos como os ligantes podem interferir na estrutura de VLA-4, a partir das simulações realizadas. Na figura abaixo, detalhamos o RMSF e o RMSD (*backbone*) da proteína em todos os sistemas estudados. Esses parâmetros nos informam a respeito da flexibilidade de movimento de resíduos individuais na estrutura proteica e do grau de desvio da proteína em comparação com o modelo inicial, antes das simulações, respectivamente.

Figura 19 – a) RMSF (do inglês, *Root Mean Square Fluctuation*) e b) RMSD (do inglês, *Root Mean Square Deviation*) do backbone de VLA-4 para cada sistema estudado. Os círculos indicam a região do sítio de ligação em VLA-4. Gráficos produzidos no programa GraphPad Prism 8.0.2



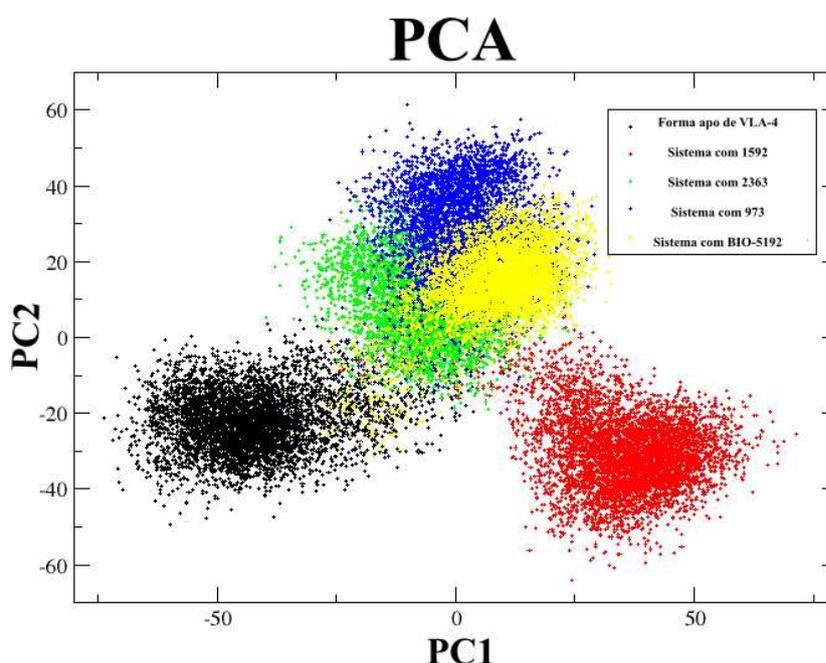
Fonte: O autor.

Neste caso, recortamos uma réplica representativa de cada sistema para as análises. Podemos concluir que as regiões onde há maior flutuação de resíduos são onde há a presença de alças (ou *loops*) na proteína, o que é comum de acontecer, devido à maior flexibilidade de movimento dessas estruturas. Nos resíduos envolvidos no sítio de ligação de VLA-4, notamos uma baixa flutuação, o que denota que essa região crucial do receptor se mantém de forma mais rígida ao longo das simulações (Figura 19a). Podemos notar também que, para todos os sistemas analisados, o RMSD de VLA-4 se mantém em uma certa estabilidade, variando, em média, de 1,5 Å a 2 Å. Isto é, a estrutura proteica simulada não se diferenciou tanto daquela

inicial, mesmo quando da interação do receptor com diferentes ligantes.

E com o intuito de estudarmos os modos de movimento de VLA-4 para cada sistema analisado, implementamos cálculos de Análises de Componentes Principais (PCA) (Figura 20) nas simulações de DM. Para o estudo, foram incluídos apenas os resíduos que estivessem a 8 Å do íon MIDAS da proteína, ou seja, aqueles envolvidos no sítio de ligação. Além disso, foi utilizado somente o *backbone* de VLA-4, isto é, aminoácidos da cadeia principal (C, C $\alpha$ , N, O), a fim de diminuir o custo computacional da operação.

Figura 20 – Análise de Componentes Principais (PCA) das 3 réplicas das simulações para cada sistema estudado de VLA-4 na forma apo e com os ligantes. Gráfico produzido no programa Xmgrace



Fonte: O autor.

Como podemos observar na Figura 20, há a sobreposição de pontos para os sistemas de VLA-4 com os ligantes 2363, 973 e o líder BIO-5192. Já para os outros complexos, temos o aparecimento de um conjunto de pontos em posições distintas do gráfico (ligantes 1592 e forma apo de VLA-4).

Os cálculos de PCA são úteis para identificarmos modos de movimento em uma estrutura proteica, além de avaliarmos a formação de conformações alternativas, elucidando o conhecimento a respeito de como os movimentos proteicos podem afetar sua função (PALMA; PIERDOMINICI-SOTTILE, 2023). Os nossos dados indicam que o complexo de VLA-4 com

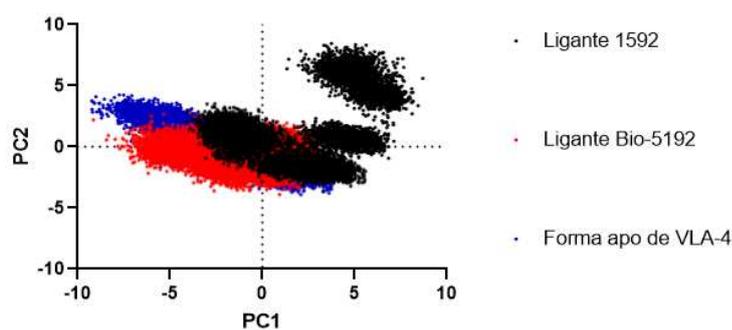
o composto 1592 (pontos pretos) mostrou um comportamento diferente daquele observado aos outros sistemas, incluindo a forma apo do receptor (pontos vermelhos). Isto é, o ligante 1592 afetou de modo diferencial a região do sítio de ligação da proteína, causando alterações únicas em sua estrutura, em relação aos demais complexos estudados.

Tais análises vêm sendo implementadas na comunidade científica no estudo de proteínas e ligantes, como no nosso caso. No estudo de Bruder e colaboradores (BRUDER; POLO; TRIVELLA, 2020), os pesquisadores analisaram a ação de moduladores naturais alostéricos (alloNPs) no mecanismo conformacional de seus alvos biológicos, a partir de dados de PCA.

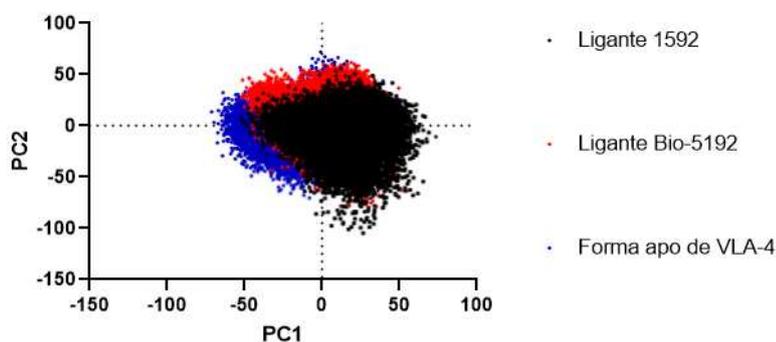
Nesse sentido, como houve uma diferenciação da região de pontos no gráfico do sistema com o ligante 1592, resolvemos analisar, isoladamente, as simulações dos sistemas com 1592, BIO-5192 e a forma apo de VLA-4. Os dados foram coletados a partir de resíduos situados no sítio de interação do receptor (Figura 21a) e contando todos os resíduos da proteína (Figura 21b).

Figura 21 – Análise de Componentes Principais (PCA) das 3 réplicas das simulações para cada sistema estudado de VLA-4 na forma apo e com os ligantes 1592 e BIO-5192 – a) Resíduos do sítio de ligação de VLA-4 e b) Todos os resíduos de VLA-4. Gráfico produzido no programa GraphPad Prism 8.0.2

**PCA - Ligante 1592, ligante BIO-5192 e forma apo de VLA-4 (sítio de ligação): 3 replicatas de 500 ns**



**PCA - Ligante 1592, ligante BIO-5192 e forma apo de VLA-4 (todos os resíduos): 3 replicatas de 500 ns**



b)

Fonte: O autor.

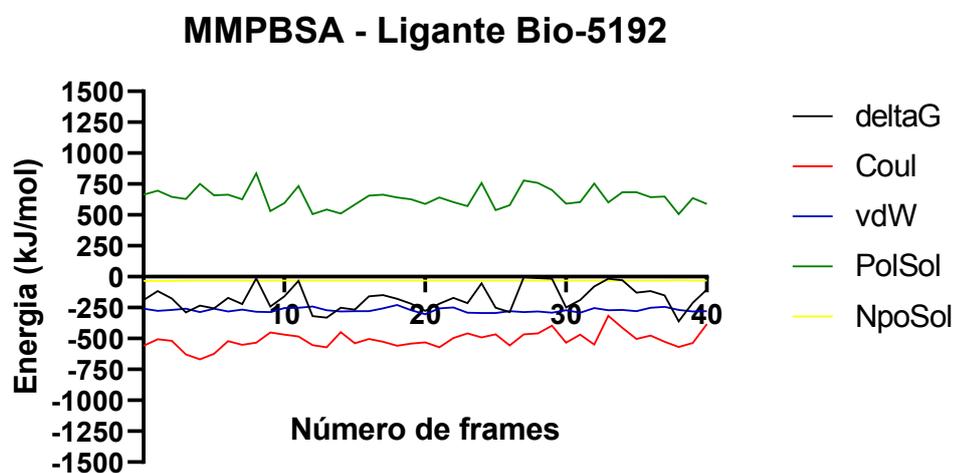
Percebemos, a partir da Figura 21a, que há uma concentração de pontos vermelhos (BIO-5192) e azuis (VLA-4 apo) se sobrepondo no centro do gráfico. Já para o sistema em preto (1592), alguns pontos divergem e se localizam no primeiro quadrante do gráfico. Tais pontos representam, em sua maioria, as simulações da segunda réplica do sistema com o ligante 1592. Enquanto na Figura 21b, há uma concentração de pontos para os 3 sistemas estudados na região central do gráfico, indicando não haver grandes diferenças conformacionais nos sistemas analisados, a partir dos dados coletados de todos os resíduos proteicos.

Portanto, o complexo com o composto 1592 se diferenciou dos demais, isto é, houve o aparecimento de modos de movimento no sítio de ligação de VLA-4 distintos, a partir da interação dessa molécula, que se sobressaíram no gráfico de PCA.

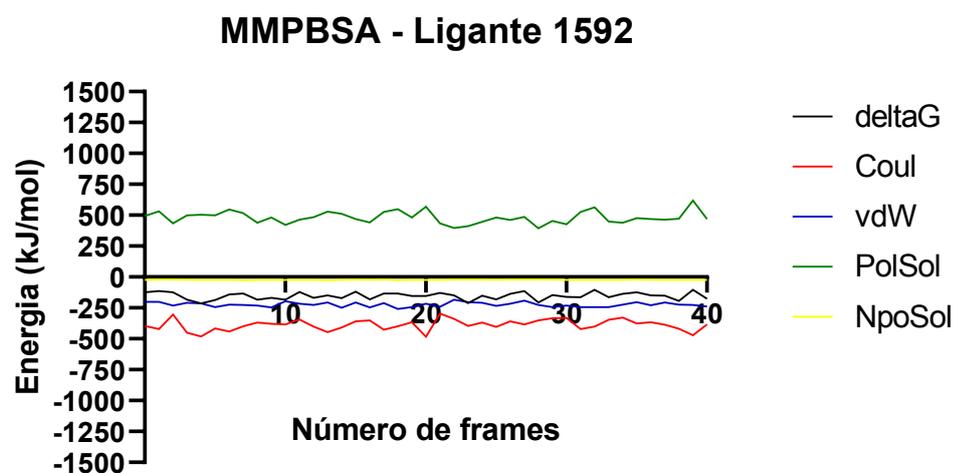
#### **5.4 Análises de cálculos de energia livre de ligação para os complexos de BIO-5192 e seus derivados em VLA-4**

Os dados apresentados a seguir dizem respeito às decomposições energéticas dos cálculos de MM/PBSA dos sistemas estudados de VLA-4 (estrutura da “cabeça”) complexada aos ligantes BIO-5192 e seus derivados. Para tanto, foram considerados apenas os últimos 250 ns de simulação perfazendo, ao todo, 40 *frames* analisados de uma réplica representativa das 3 totais para cada sistema (Figura 22).

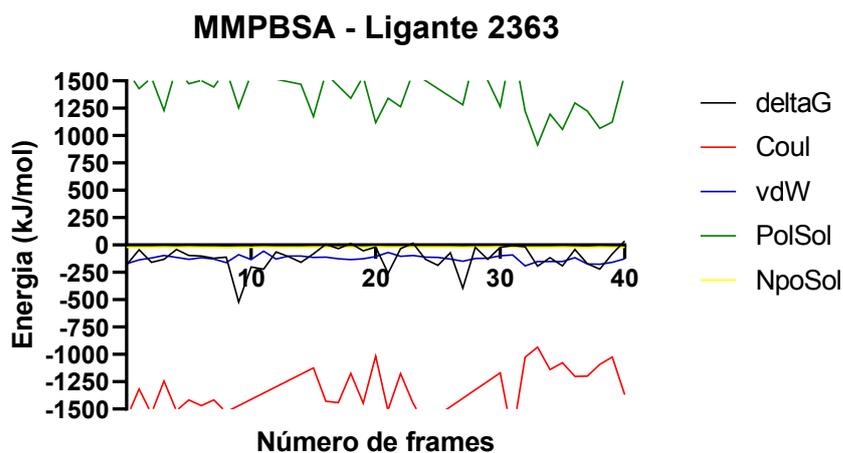
Figura 22 – Decomposições energéticas a partir de cálculos por MM/PBSA dos sistemas com VLA-4 complexada aos ligantes a) BIO-5192; b) 1592; c) 2363; d) 973. Gráficos produzidos no programa GraphPad Prism 8.0.2



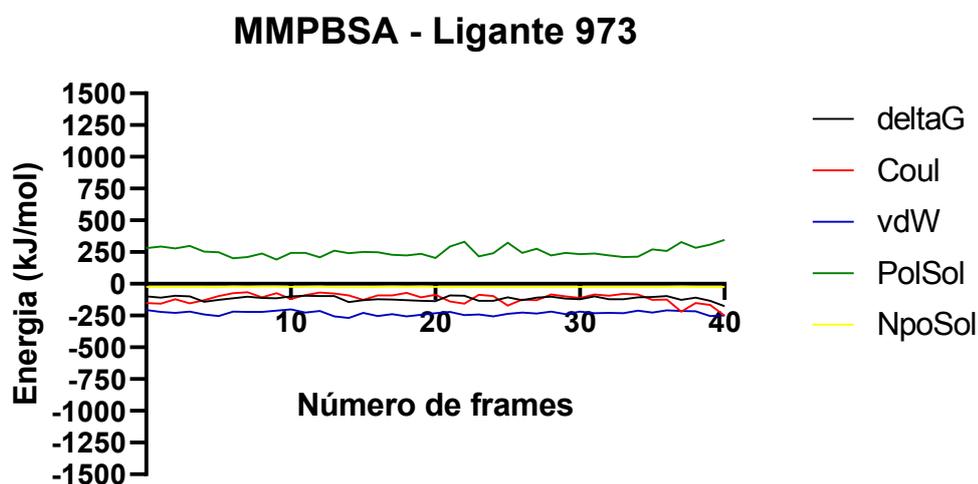
a)



b)



c)



d)

Fonte: O autor.

A tabela abaixo mostra os valores finais de  $\Delta G_{\text{ligação}}$  calculados a partir das contribuições de cada potencial energético, como mostrado na Figura 22.

Tabela 3 – Valores finais de  $\Delta G_{\text{ligação}}$  para os complexos de VLA-4 e os ligantes estudados, a partir dos cálculos por MM/PBSA de uma réplica de simulação

Ligante	$\Delta G_{\text{ligação}}$ final		
BIO-5192	-123.062	+/-	18.907 kJ/mol
1592	-146.849	+/-	5.249 kJ/mol
2363	-80.018	+/-	22.015 kJ/mol
973	-109.854	+/-	3.520 kJ/mol

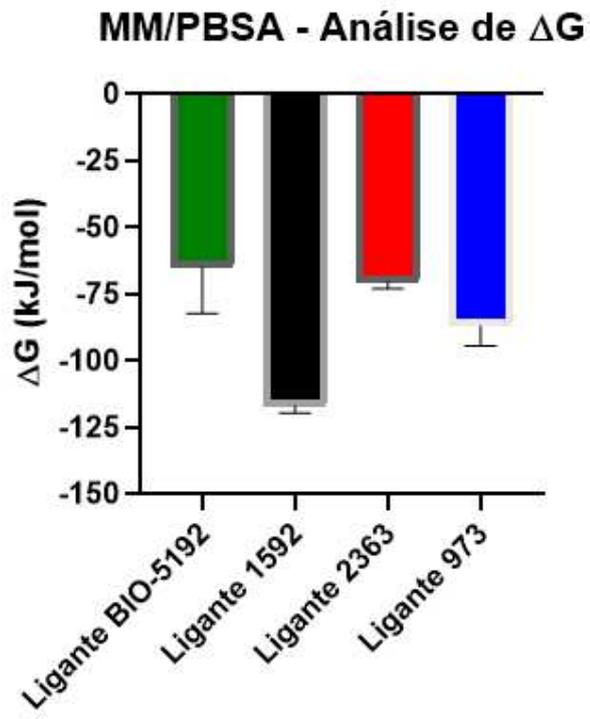
Fonte: O autor.

Analisando os resultados, concluímos que o ligante que apresentou maior afinidade a VLA-4 foi o 1592, com um valor de energia final menor que o composto-líder do qual foi derivado. Além disso, o sistema com 1592 mostrou um resultado baixo de desvio-padrão, ou seja, para os *frames* analisados, os potenciais calculados sofreram poucas flutuações nos seus resultados, demonstrando uma estabilidade do complexo ao longo da simulação.

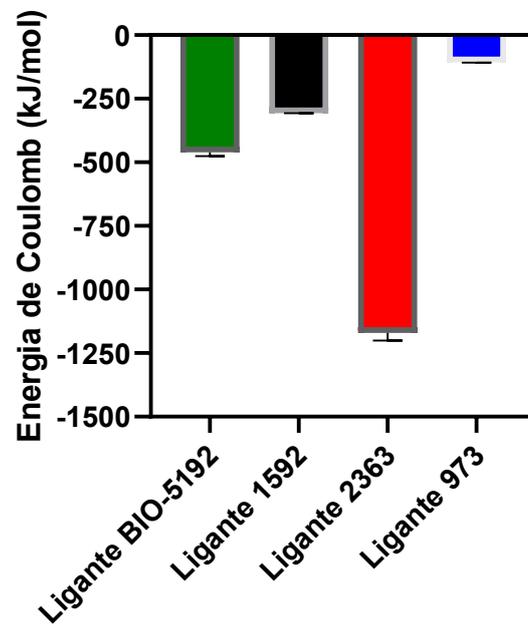
Em contrapartida, os sistemas com os ligantes 2363 e 973 apresentaram os maiores valores de  $\Delta G_{\text{ligação}}$  demonstrando menores afinidades a VLA-4. O complexo com 2363, inclusive, demonstrou bruscas variações dos potenciais energéticos calculados a partir de um valor alto de desvio-padrão. É importante notar que, em alguns *frames*, o potencial  $\Delta G$  foi positivo, o que denota uma baixa afinidade da pequena molécula a seu alvo.

A Figura 23 abaixo mostra a análise estatística de 3 potenciais ( $\Delta G$ , Coulomb e van der Waals) para cada sistema de VLA-4 complexada aos ligantes estudados, nos cálculos de MM/PBSA a partir dos últimos 250 ns de simulação das 3 réplicas de cada sistema.

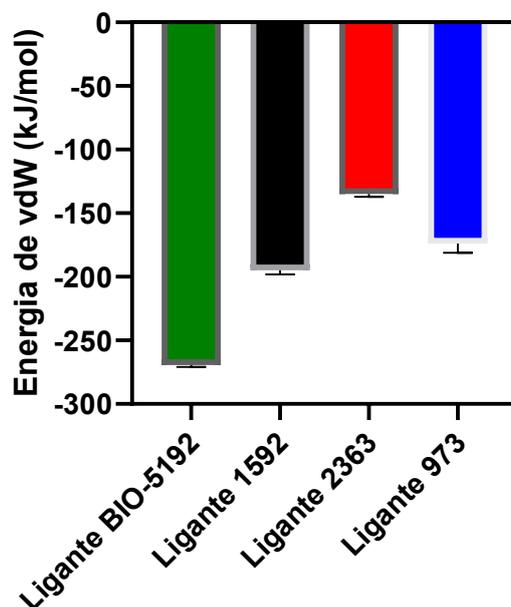
Figura 23 – Interpolação linear de dados dos potenciais: a)  $\Delta G$  (BIO-5192 – Média = -64,85 kJ/mol e DP = 142,10 kJ/mol; 1592 – Média = -117,02 kJ/mol e DP = 41,14 kJ/mol; 2363 – Média = -70,37 kJ/mol e DP = 155,10 kJ/mol; 973 – Média = -86,94 kJ/mol e DP = 38,15 kJ/mol); b) Coulomb (BIO-5192 – Média = -460,25 kJ/mol e DP = 76,95 kJ/mol; 1592 – Média = -306,28 kJ/mol e DP = 87,17 kJ/mol; 2363 – Média = -1170,48 kJ/mol e DP = 419,5 kJ/mol; 973 – Média = -106,99 kJ/mol e DP = 67,15 kJ/mol); c) van der Waals (BIO-5192 – Média = -269,26 kJ/mol e DP = 19,83 kJ/mol; 1592 – Média = -194,87 kJ/mol e DP = 29,59 kJ/mol; 2363 – Média = -134,85 kJ/mol e DP = 24,77 kJ/mol; 973 – Média = -173,77 kJ/mol e DP = 56,73 kJ/mol) para os cálculos de MM/PBSA dos últimos 250 ns de simulação de 3 réplicas dos sistemas formados por VLA-4 e os ligantes estudados neste trabalho. Gráficos produzidos no programa GraphPad Prism 8.0.2. DP = Desvio-padrão



**MM/PBSA - Análise da Energia de Coulomb**



## MM/PBSA - Análise da Energia de van der Waals



c)

Fonte: O autor.

Para o potencial  $\Delta G$ , o ligante 1592 apresentou, em média, um menor valor, isto é, foi o composto que mostrou maior afinidade a VLA-4, em comparação aos demais, até mesmo melhor que seu líder BIO-5192. Por outro lado, analisando a energia de Coulomb, que calcula o potencial eletrostático do sistema, notamos que o ligante 2363 se sobressaiu em relação aos outros. Para todos os ligantes, o potencial eletrostático foi negativo, o que denota uma força atrativa entre receptor e ligante. Logo, tal composto apresentou uma contribuição maior na interação entre cargas elétricas, quando do reconhecimento com VLA-4. Já para o quesito potencial de van der Waals, que nos dá a ideia da soma de interações fracas entre ligante e seu alvo, concluímos que o composto-líder BIO-5192 mostrou uma maior contribuição energética nesse aspecto.

Diante de todos os aspectos que foram discutidos, o ligante 1592 aparece como uma molécula promissora a ser avaliada como potencial inibidor da integrina  $\alpha 4\beta 1$  em futuros ensaios *in vitro* e *in vivo*.

\*\*\*

As pequenas moléculas que se ligam seletivamente à integrina  $\alpha 4\beta 1$  foram projetadas com base nas sequências de reconhecimento mínimo com as proteínas da matriz

extracelular. Em particular, o tripeptídeo LDV (Leu-Asp-Val) que foi reconhecido como a sequência de ligação encontrada na região CS1 (do inglês, *alternatively spliced Connecting Segment*) da fibronectina (KOMORIYA et al., 1991) é homólogo e bastante isostérico ao fragmento IDS (Ile-Asp-Ser), presente no local de ligação do VCAM-1 com VLA-4.

Devido à falta de estruturas cristalográficas de complexos ligante-receptor (JONES et al., 1995), as sugestões sobre as características tridimensionais necessárias que podem garantir uma afinidade ideal foram deduzidas apenas pela combinação de modelos por homologia deduzidos pelo modelo de integrinas  $\beta 2$  (CD11A/CD11B) (YOU et al., 2002) e estudos QSAR em uma pequena biblioteca de ligantes que mimetizam LDV (AMIN et al., 2018; HUTT; SAUBERN; WINKLER, 2011; SINGH et al., 2002). Os estudos de dinâmica molecular e *docking* de receptores e ligantes, aliados a outras técnicas computacionais, realizados por nosso grupo (SILVA et al., 2010; VASCONCELOS et al., 2022) tentam lançar luz sobre aspectos atômico-moleculares do reconhecimento integrina-ligante. Alguns trabalhos completos e detalhados sobre os ligantes sintéticos de  $\alpha 4\beta 1$  já foram relatados (HURYN et al., 2005; JACKSON, 2002), mas falta uma coleção de resultados mais recentes no projeto e síntese desses compostos bioativos.

Neste trabalho, estudamos compostos peptidomiméticos derivados da potente molécula BIO-5192 que é uma quimera entre dipeptídeos de prolina arilsulfonamida e compostos substituídos por PUPA. Algumas modificações químicas já foram reportadas na literatura, a fim de maximizar o desempenho desses compostos e diminuir problemas de ordem farmacocinética e de biodisponibilidade oral (HAGMANN et al., 2001).

Por exemplo, ao substituir o anel da prolina por uma 3-alkil-isoxazolina-5-carboxamida, foram obtidos ligantes com atividade nanomolar e estabilidade em microsomas (SONI et al., 2013). O composto fimategrast, desenvolvido pela companhia GSK, pertencente à mesma classe de compostos, alcançou fase II de ensaios para esclerose múltipla e é atualmente estudado como um facilitador no transplante de células hematopoiéticas “*in utero*”, uma abordagem pioneira para o tratamento de doenças fetais críticas (KIM et al., 2016).

No caso específico deste trabalho, os novos ligantes foram desenvolvidos mantendo o anel da prolina (R1 – Figura 8) e modificando a região da cadeia bifênil (R2 – Figura 8) de BIO-5192. De forma análoga ao que foi estudado, o composto TR14035 (SIRCAR et al., 2002), é uma benzamida substituída por dicloro da estrutura de bifênilalanina. Essa molécula apresentou IC<sub>50</sub> nanomolar, mas atuou como um ligante duplo, sendo ativo tanto nas integrinas  $\alpha 4\beta 1$  quanto nas integrinas  $\alpha 4\beta 7$ . De qualquer forma, representou o principal composto para o desenvolvimento de derivados refinados, como o AJM-300, onde a

cadeia bifenil foi substituída por uma fenil-pirimidindiona. Este composto está atualmente na fase clínica III para colite ulcerativa (SUGIURA et al., 2013).

## 6 RESULTADOS E DISCUSSÃO – PARTE II

### 6.1 Análises do protocolo de *Virtual Screening* (VS) no reposicionamento de fármacos na integrina $\alpha 4\beta 1$

Como mencionado anteriormente (vide METODOLOGIA – PARTE II), realizamos o método de reposicionamento de fármacos, aplicando um protocolo de VS tendo como alvo VLA-4. Foram 2102 compostos aprovados pela FDA extraídos do banco de dados “fda” do ZINC. Os resultados obtidos a partir do procedimento realizado são mostrados abaixo (Tabela 4).

Tabela 4 – Os dez fármacos com as menores energias calculadas no *docking* dos compostos do banco “fda” do Zinc em VLA-4, por ordem decrescente do *Vina Score*

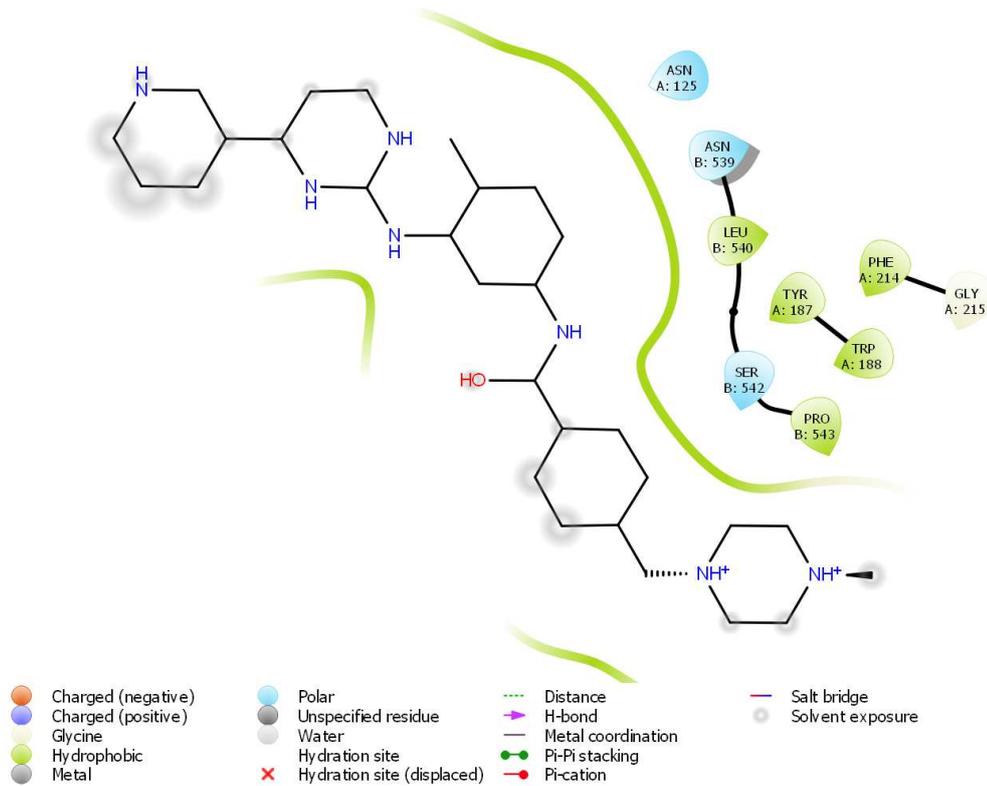
Nome do Fármaco	Nome Comercial	VINA Score (kcal/mol)	Doença Tratada	Administração
Imatinib	Gleevec, Glivec	-9,9	Leucemia Mieloide Crônica (LMC)	Oral
Rifaximina	Xifaxan, Zaxine	-9,9	Doenças do trato gastrointestinal	Oral
Simeprevir	-	-9,9	Hepatite C	Oral
Nilotinib	Tasigna	-9,8	LMC	Oral
Midostaurina	Rydapt	-9,7	Leucemia Mieloide Aguda (LMA)	Oral
Saquinavir	Invirase	-9,7	AIDS/HIV	Oral
Lumacaftor	Orkambi	-9,6	Fibrose cística	Oral
Ergotamina	Cafergot, Ergomar, Migergot	-9,5	Enxaqueca e cefaleia	Sublingual/Oral
Pazopanib	Votrient	-9,5	Câncer renal	Oral
Conivaptan	Vaprisol	-9,5	Hiponatremia	Intravenoso

Fonte: *Drugbank*.

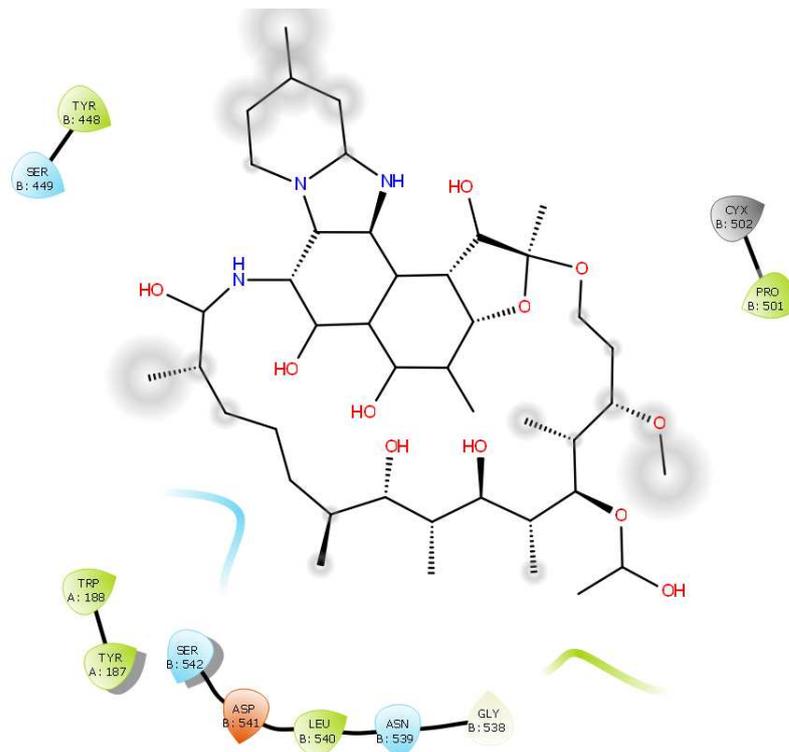
Na tabela acima, percebemos que há uma variedade de fármacos com áreas de atuação distintas, com energias de ligação em VLA-4 calculadas próximas a -10 kcal/mol. Desse modo, analisamos, portanto, as interações que alguns desses compostos fazem quando no sítio do receptor (Figura 24).

Figura 24 – Diagrama de interações de 5 compostos mais bem posicionados no *docking* do banco “fda” do ZINC em VLA-4. Figuras produzidas no programa Maestro 11.8, com um raio de corte de 4 Å

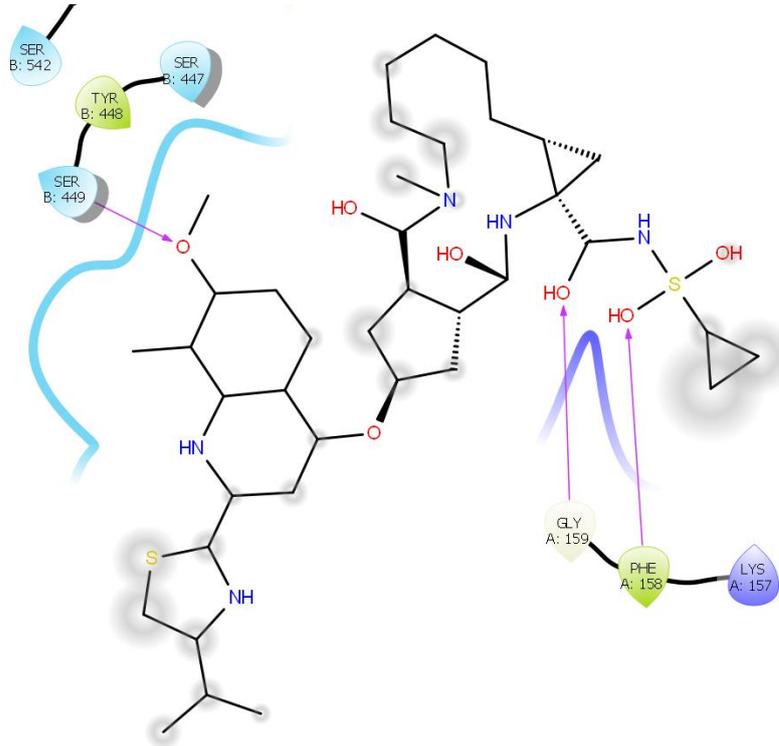
### Imatinib



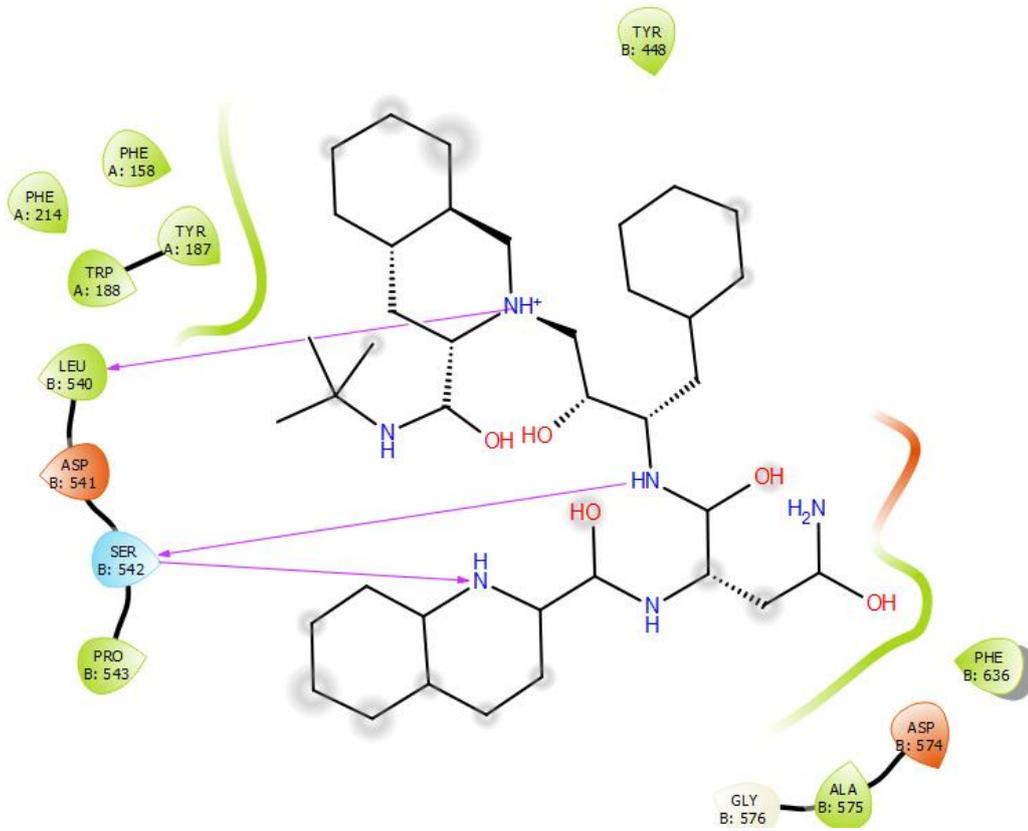
### Rifaximina



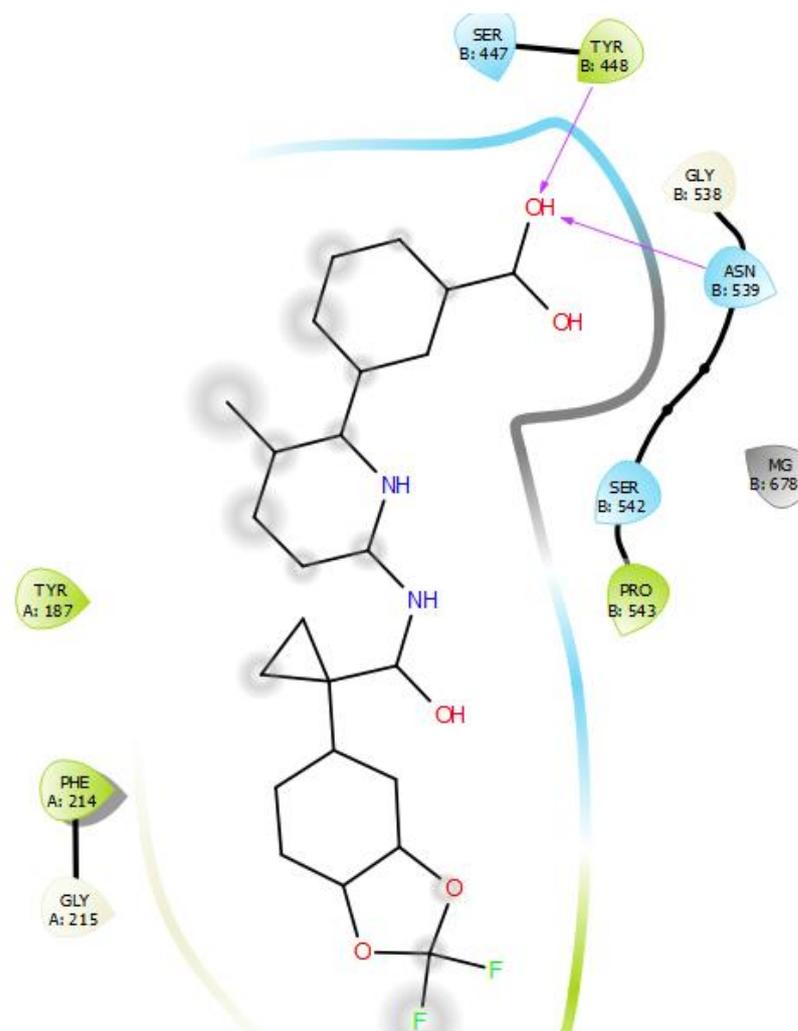
### Simeprevir



### Saquinavir



### Lumacaftor



Fonte: O autor. A legenda da primeira figura (Imatinib) é a mesma para todas as outras desse diagrama.

Concluimos que para os fármacos imatinib e rifaximina, há a presença de aminoácidos apolares (linha verde) interagindo com as moléculas, que são policíclicas. Já para o simeprevir, observamos três interações explícitas. Todas são ligações de hidrogênio. A Ser449 com um oxigênio de um grupo éter; a Gly159 com um oxigênio de um grupamento amida e a Phe159 com oxigênio de um ciclopropano sulfonato. É interessante notar que a estrutura do dióxido sulfúrico presente neste último radical químico é comum, inclusive, para o composto-líder BIO-5192 e seus derivados mostrados neste trabalho. Os fármacos saquinavir e lumacaftor também realizam múltiplas ligações de hidrogênio com o alvo. Este interage com VLA-4 a partir de ligações de hidrogênio onde o carboxilato está envolvido com os resíduos Asn539 e Tyr448. Já aquele está participando de interações de grupos NH com Leu540 e Ser542.

Assim, esses fármacos mostrados contêm radicais químicos compatíveis com moléculas antagonistas da integrina  $\alpha 4\beta 1$ , mostrando um certo padrão atômico-molecular do reconhecimento integrina-ligante.

## 6.2 Análises da validação do *Virtual Screening* (VS) no reposicionamento de fármacos na integrina $\alpha 4\beta 1$

Como mostrado na Figura 10, montamos um esquema de controles para validarmos nosso VS, no processo de reposicionamento de fármacos para a integrina VLA-4. A seguir, são apresentados os resultados do *docking* do banco de 185 moléculas com 80% de similaridade ao composto-líder BIO-5192 em VLA-4 (Tabelas 5 e 6).

Tabela 5 – As dez moléculas com as menores energias calculadas no *docking* dos compostos com 80% de similaridade a BIO-5192 em VLA-4, por ordem decrescente do *Vina Score*

<b>Vina Score (kcal/mol)</b>	<b>Nome ChEMBL</b>	<b>Alvo</b>	<b>IC50 (nM)</b>	<b>Atividade</b>
-10	330687	VLA-4	0,1	Ativo
-9,3	91061	VLA-4	0,15	Ativo
-9,3	328890	VLA-4	0,38	Ativo
-9,3	432215	VLA-4	0,63 e 0,65	Ativo
-9,1	507665	$\alpha 4$	0,09 e 0,4	Ativo
-9,1	91775	VLA-4	0,23	Ativo
-9,1	40295	VLA-4 e $\alpha 4\beta 7$	0,45	Ativo
-9,1	30192	VLA-4	0,46	Ativo
-9	312458	VLA-4	196	<i>Decoy</i>
-9	254336	$\alpha 4$	28	Ativo

Fonte: O autor.

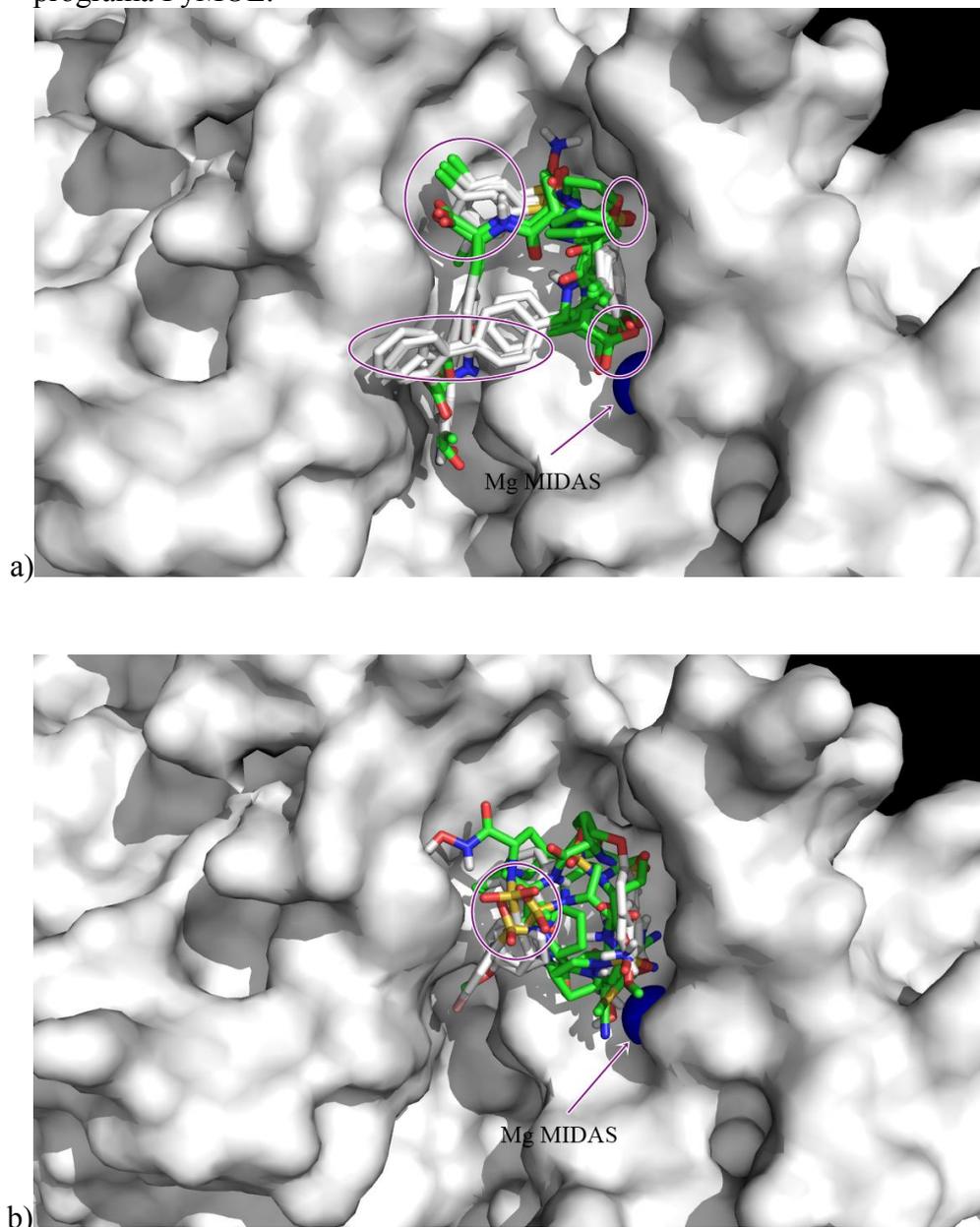
Tabela 6 – As dez moléculas com as maiores energias calculadas no *docking* dos compostos com 80% de similaridade a BIO-5192 em VLA-4, por ordem decrescente do *Vina Score*

<b>VINA Score (kcal/mol)</b>	<b>Nome ChEMBL</b>	<b>Alvo</b>	<b>IC50 (nM)</b>	<b>Atividade</b>
-6,6	357871	Metalopeptidase humana	380	<i>Decoy</i>
-6,6	3742218	Legumaina humana	593	<i>Decoy</i>
-6,6	93912	VLA-4	45	Ativo
-6,5	3741126	Legumaina humana	282	<i>Decoy</i>
-6,5	3261747	Histona desacetilase humana	>100000	<i>Decoy</i>
-6,4	3742323	Legumaina humana	10000	<i>Decoy</i>
-6,3	3485748	GMP-AMP Sintase	-	Inespecífico
-6,3	147629	Colagenase humana	0,8 a 920	Inespecífico
-6,1	2088997	Fator Letal ( <i>Bacillus anthracis</i> )	10000	<i>Decoy</i>
-6,1	3483927	tRNA sintetase ( <i>Leishmania</i> )	-	Inespecífico

Fonte: O autor.

Como podemos notar nas tabelas supracitadas, nosso protocolo de VS foi capaz de ranquear, nas primeiras posições, compostos ativos a VLA-4 em sua maioria, deixando os inativos (*decoys*) nas últimas colocações. Os modos de ligação (primeira *pose* de cada ligante) dessas moléculas no sítio de VLA-4 também foram analisados e são mostrados abaixo (Figura 25).

Figura 25 – Modos de ligação dos melhores (a) e dos piores (b) compostos com 80% de similaridade a BIO-5192 em VLA-4. Os círculos indicam sobreposição de radicais químicos. Figuras produzidas no programa PyMOL.

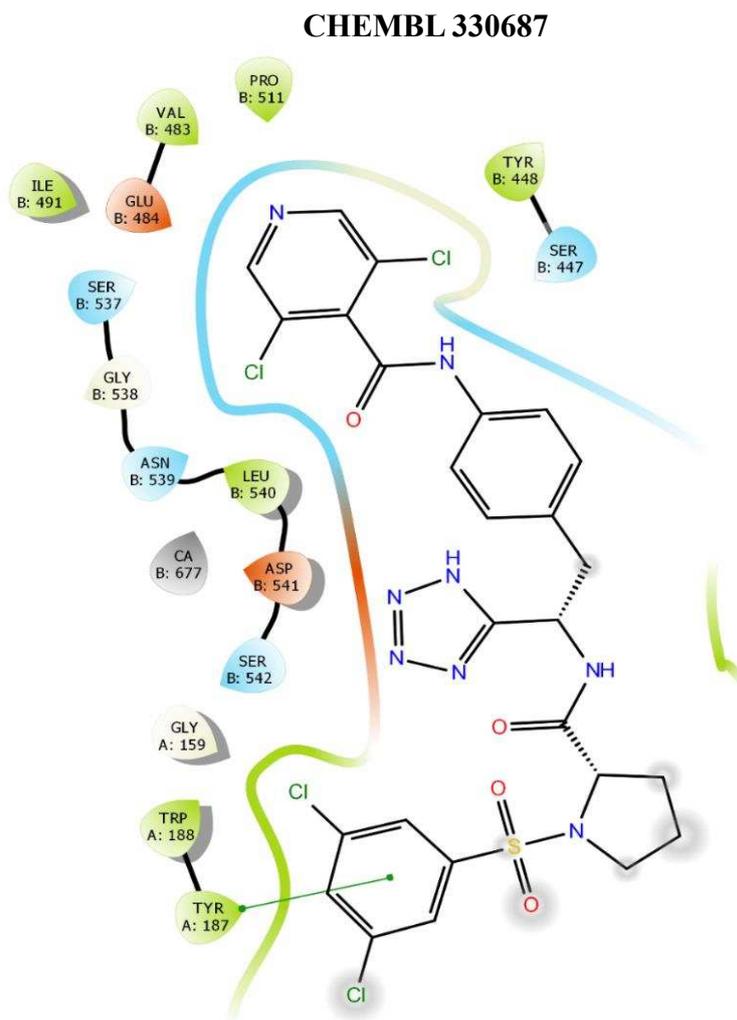


Fonte: O autor.

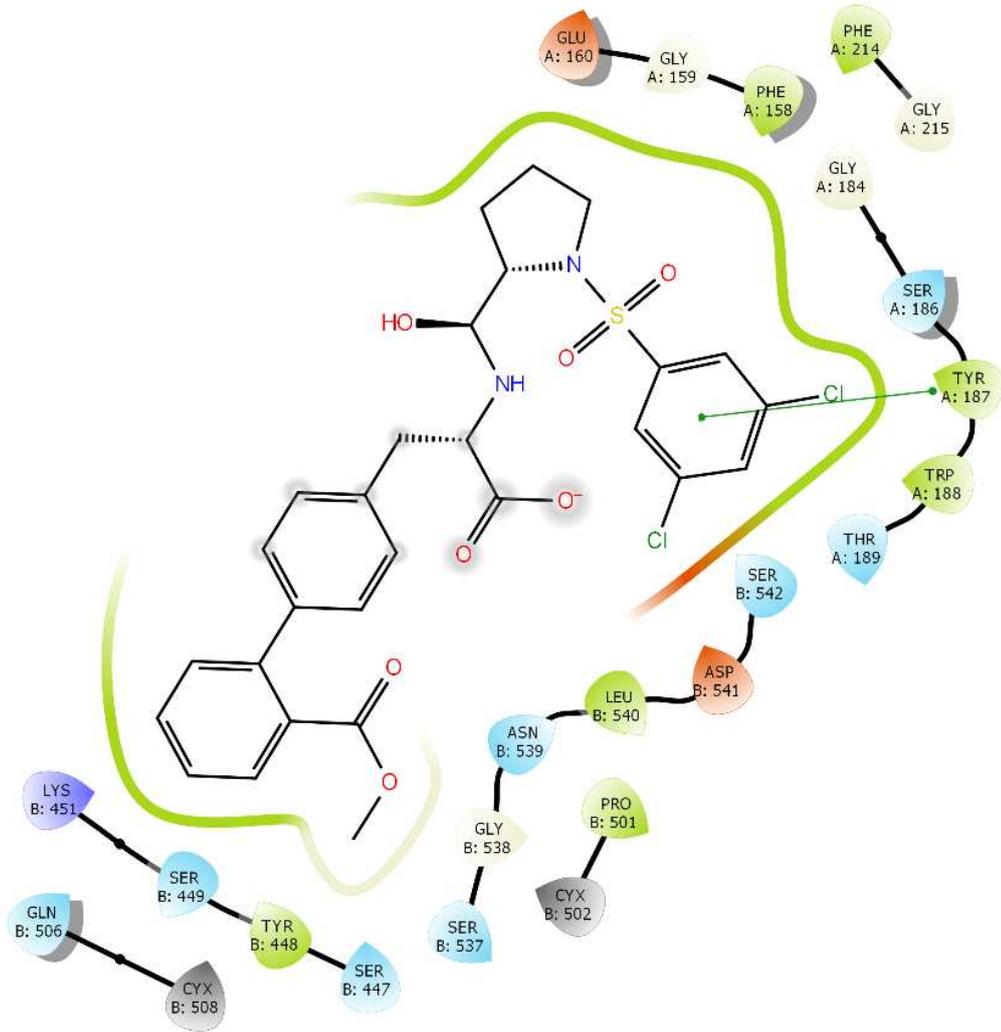
Nossos resultados indicam que, para os compostos mais bem posicionados no VS (Figura 25a), há uma sobreposição de radicais semelhantes para os modos de ligação estudados, inclusive com os grupamentos carboxilatos apontando em direção ao íon  $Mg^{2+}$  MIDAS. Isso não acontece para as moléculas com pontuações piores no *docking* (maiores energias calculadas – Figura 25b), onde não há quase nenhuma sobreposição de radicais, indicando baixo potencial de encaixe no sítio de ligação do receptor.

As interações de alguns dos compostos mais bem ranqueados com VLA-4 são mostradas abaixo (Figura 26).

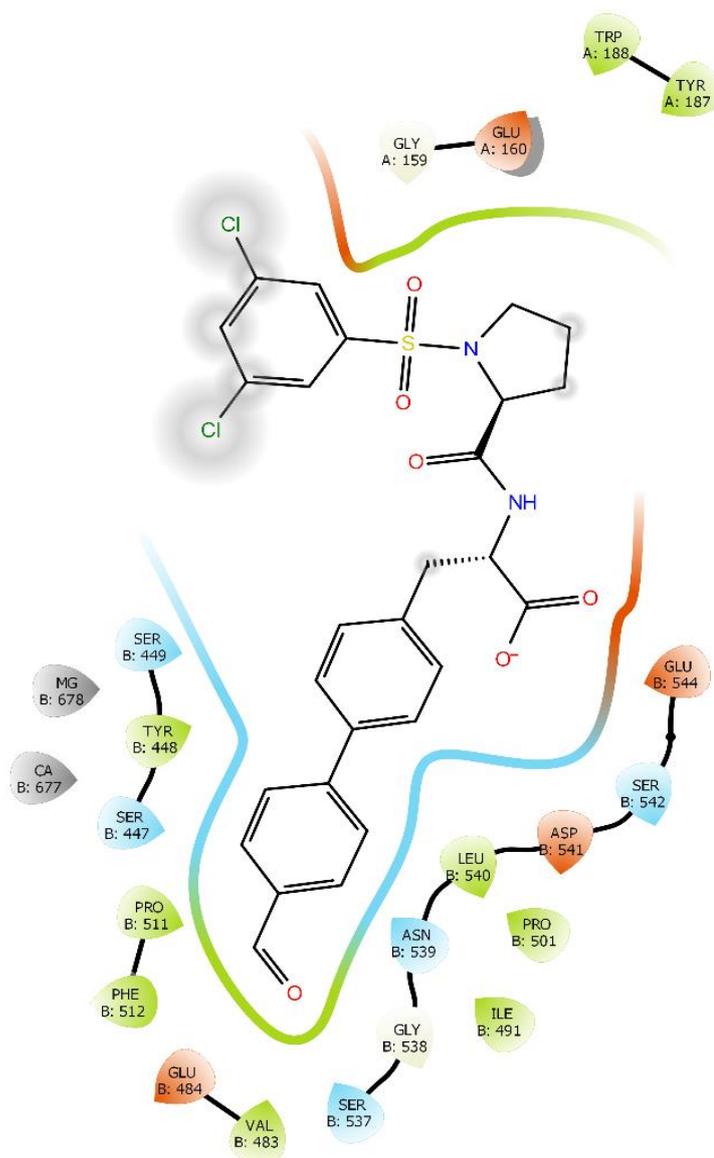
Figura 26 – Diagrama de interações de 3 compostos mais bem posicionados no *docking* das moléculas com 80% de similaridade a BIO-5192 em VLA-4. Figuras produzidas no programa Maestro 11.8, com um raio de corte de 4 Å



## CHEMBL 3742218



## CHEMBL 93912

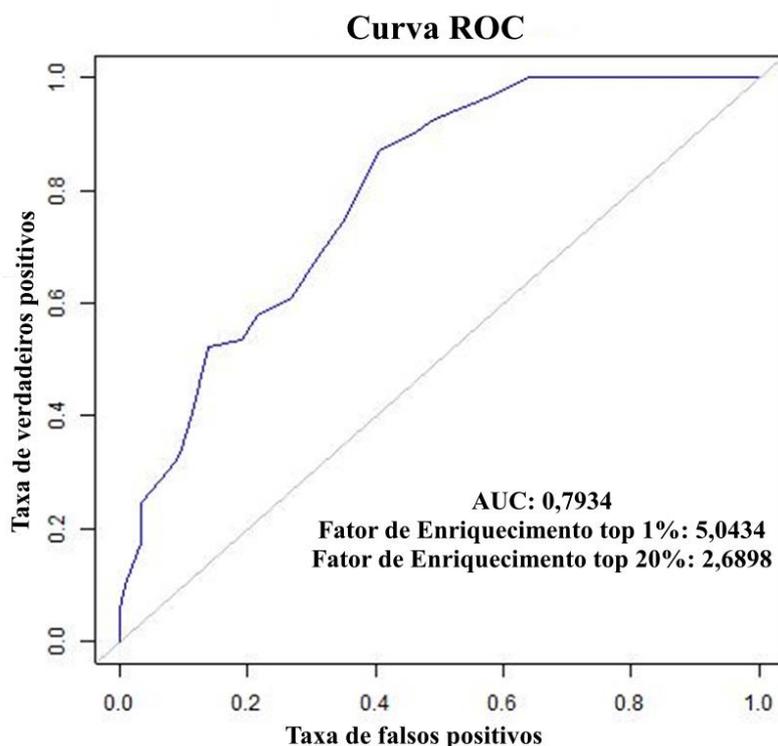


Fonte: O autor.

Analisando as interações, concluímos, para os dois primeiros compostos da figura, a Tyr187, um aminoácido aromático, participando de uma interação  $\pi$ -stacking com o grupamento 1,3-diclorobenzeno. No último composto (CHEMBL 93912), há a presença de aminoácidos polares (linha azul claro) e com carga negativa (linha vermelha) envolvendo o grupo fenilbenzeno, enquanto o 1,3-diclorobenzeno está exposto ao solvente.

Para validarmos todo esse método de VS aplicado ao banco de moléculas com 80% de similaridade estrutural a BIO-5192, plotamos a curva ROC do processo, assim como os cálculos de AUC e Fatores de Enriquecimento (Figura 27).

Figura 27 – Curva ROC do protocolo de *docking* com o banco de moléculas com 80% de similaridade estrutural a BIO-5192 e os parâmetros calculados AUC e Fator de Enriquecimento para 1% e 20% de compostos inativos encontrados



Fonte: O autor.

Como observamos, a curva ROC produziu uma área sob a curva (AUC) de quase 80%, o que significa que nosso protocolo de *docking* está devidamente validado. Isto é, o VS aplicado foi capaz de posicionar os ligantes verdadeiros positivos (ativos) nas primeiras colocações do *docking* (menores energias de ligação calculadas) e que há uma probabilidade de quase 80% de que um composto ativo receba um escore melhor que um composto inativo. Além disso, os Fatores de Enriquecimento preditos mostram que, em torno de 5 e 3 compostos ativos são encontrados pelo protocolo aplicado, quando da presença de 1% e 20%, respectivamente, de moléculas inativas. Tais métricas são amplamente usadas nas estimativas de erros em acertos em protocolos de VS e na seleção e otimização de compostos-líder (MCGANN; NICHOLLS; ENYEDY, 2015).

Quanto ao outro esquema de controle que aplicamos neste trabalho (Figura 10), já analisamos o resultado do VS para o banco de 500 moléculas que interagem com VLA-4, como mostrado na Figura 14 e Tabela 2. Abaixo, são apresentados os dados dos piores ligantes classificados pelo protocolo aplicado (Tabela 7).

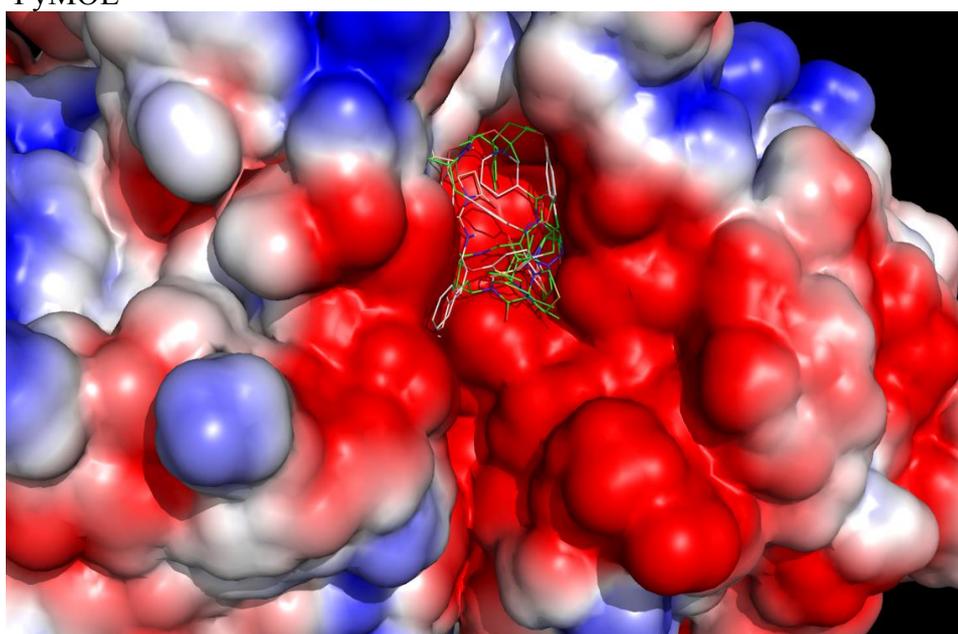
Tabela 7 – As dez moléculas com as maiores energias calculadas no *docking* dos compostos que tem como alvo VLA-4, por ordem decrescente do *Vina Score*

<b>VINA Score (kcal/mol)</b>	<b>Nome CHEMBL</b>	<b>IC50 (nM)</b>	<b>Atividade</b>
-6,1	3289604	-	-
-6,2	163622	3800	<i>Decoy</i>
-6,3	163868	112	<i>Decoy</i>
-6,4	324503	250000	<i>Decoy</i>
-6,4	435647	60	Ativo
-6,5	368331	360	<i>Decoy</i>
-6,5	2408069	12	Ativo
-6,5	378238	500	<i>Decoy</i>
-6,5	331602	50000	<i>Decoy</i>
-6,6	366857	310	<i>Decoy</i>

Fonte: O autor.

Notamos, na tabela acima, a presença, em sua maioria, de ligantes inativos/*decoys*, o que era de se esperar que o protocolo de *docking* fizesse para as últimas colocações. Desse modo, os modos de ligação desses compostos no sítio de VLA-4 são mostrados abaixo (Figura 28).

Figura 28 - Modos de ligação de alguns dos piores compostos do banco de moléculas que se ligam a VLA-4. A superfície eletrostática da proteína é mostrada, indicando regiões avermelhadas como sendo mais eletronegativas e regiões azuladas como sendo mais eletropositivas. A cor branca revela neutralidade de cargas. Figura produzida no programa PyMOL

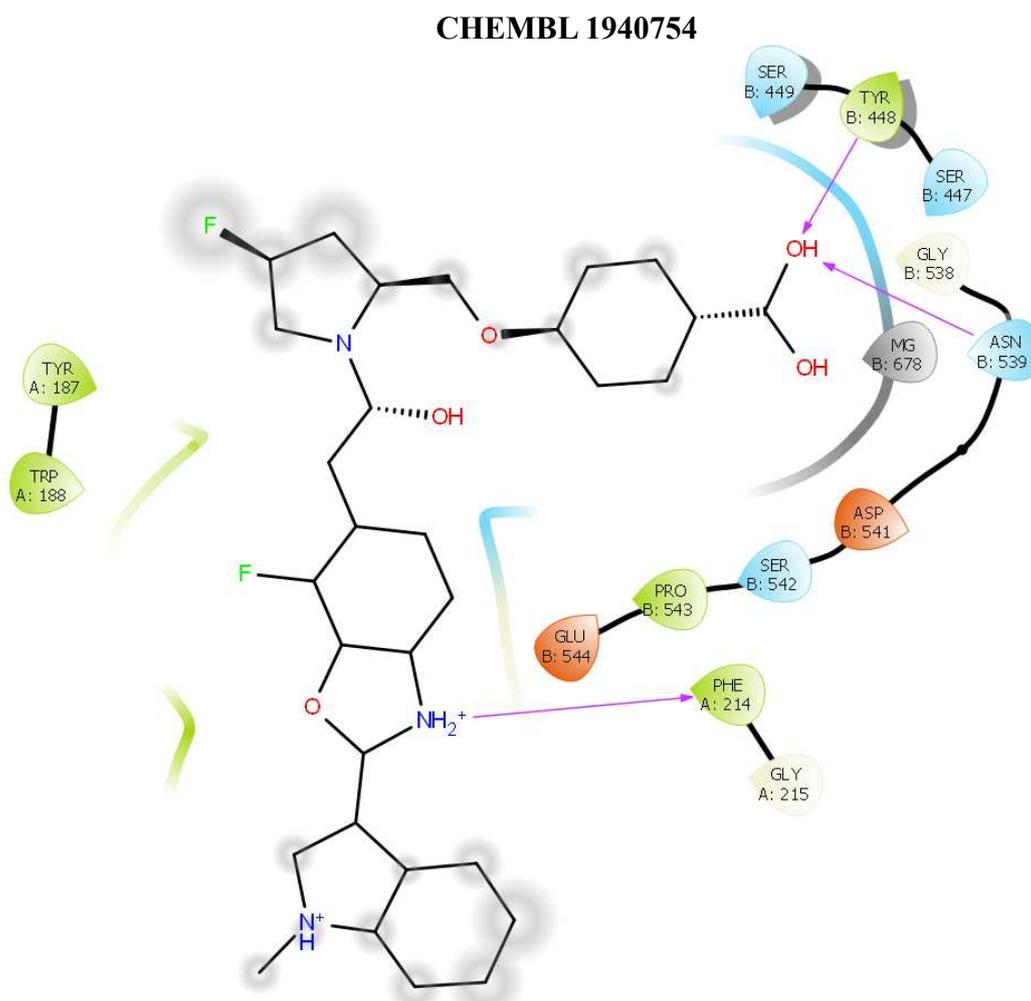


Fonte: O autor.

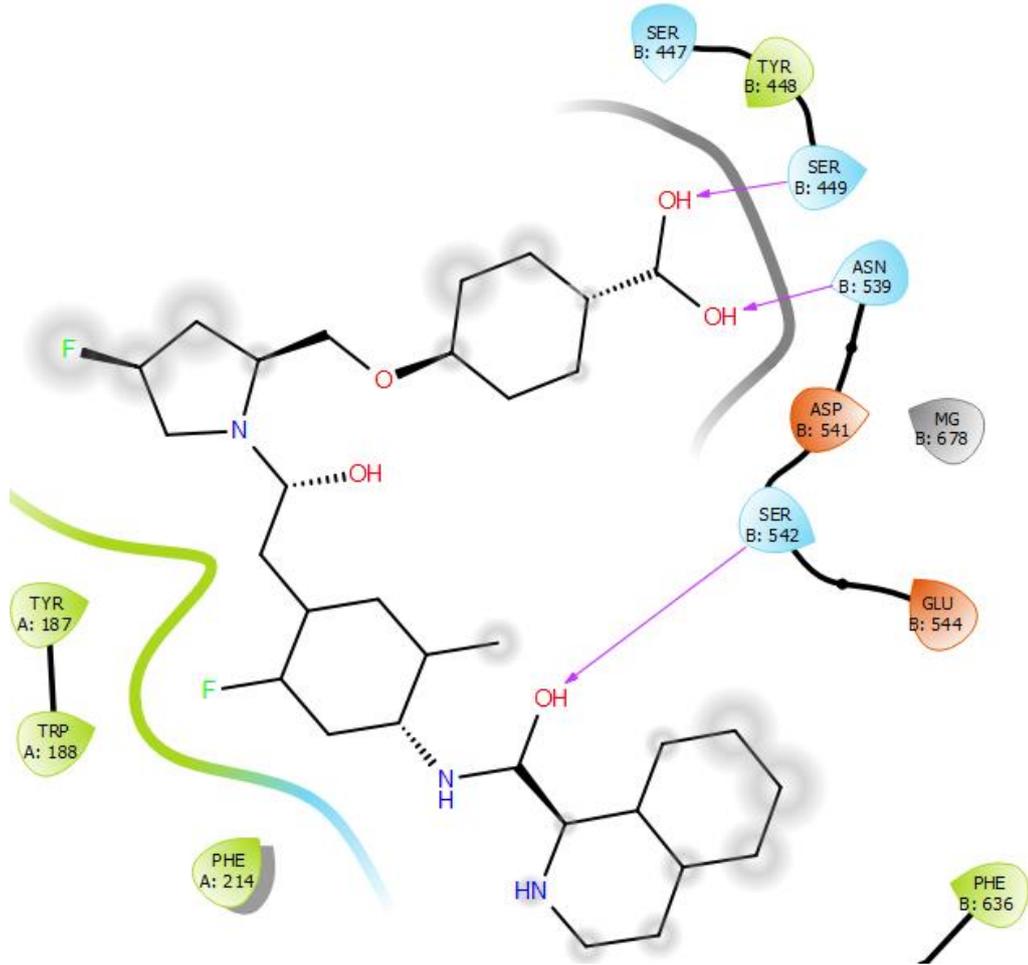
Percebemos os ligantes encaixados no sítio de ligação de VLA-4, o qual é bastante eletronegativo, de uma maneira que não é a usual para moléculas ativas ao receptor. Ou seja, não há a adoção da forma esticada de “S” reverso, abrangendo a interface entre as subunidades  $\alpha 4$  e  $\beta 1$ . Os compostos se aglomeram no sítio de uma maneira caótica e desordenada, sem sobreposições claras de radicais químicos semelhantes.

Por fim, detalhamos as interações conservadas de algumas das melhores moléculas ranqueadas no protocolo de VS dos compostos que têm VLA-4 como alvo (Figura 29).

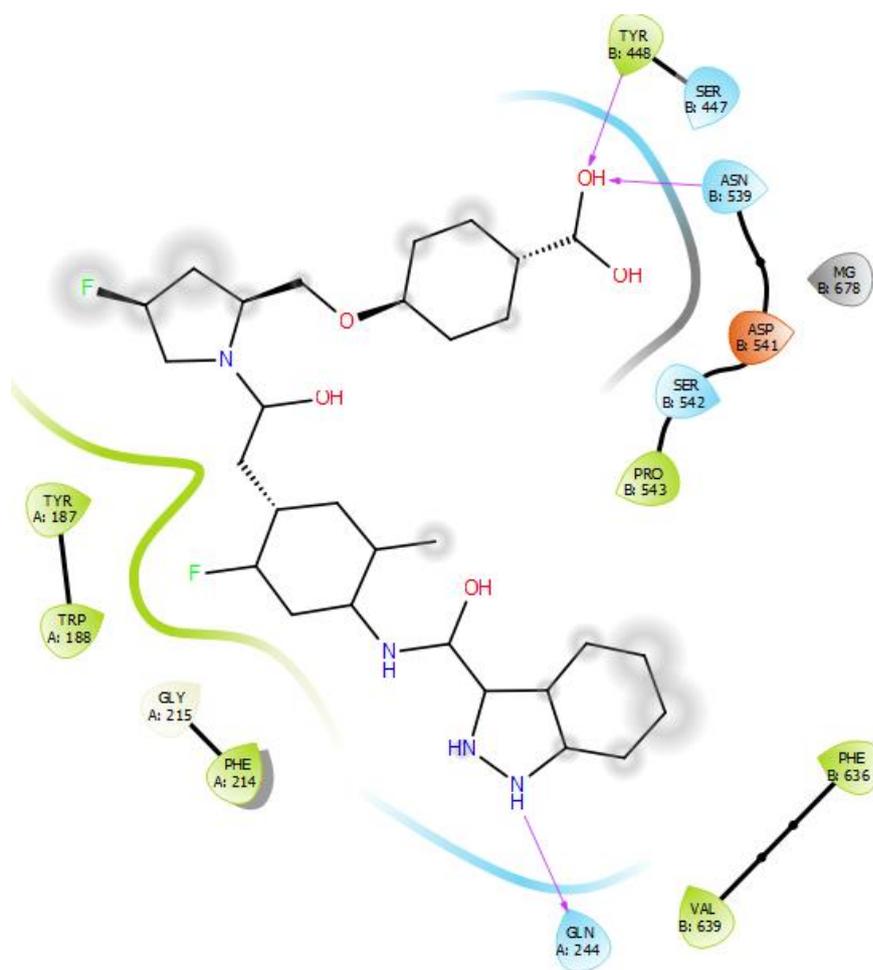
Figura 29 – Diagrama de interações de 3 compostos mais bem posicionados no *docking* do banco de moléculas que possuem como alvo VLA-4. Figuras produzidas no programa Maestro 11.8, com um raio de corte de 4 Å



## CHEMBL 572802



## CHEMBL 567180



Fonte: O autor.

Portanto, tendo como base as imagens acima, alguns aminoácidos são chave em alguns reconhecimentos. A Tyr448 e a Asn539 desempenham interações por ligações de hidrogênio com o oxigênio carboxilato dos ligantes. Como já mencionado anteriormente, esse tipo de reconhecimento é crucial na interação integrina-ligante. Notamos também o grupamento carboxilato  $\text{COO}^-$  apontando em direção ao íon  $\text{Mg}^{2+}$  MIDAS em VLA-4, indicando mais uma característica de compostos que possuem essa integrina como alvo.

### 6.3 Análises das simulações de fármacos oriundos do reposicionamento de fármacos na integrina $\alpha 4\beta 1$

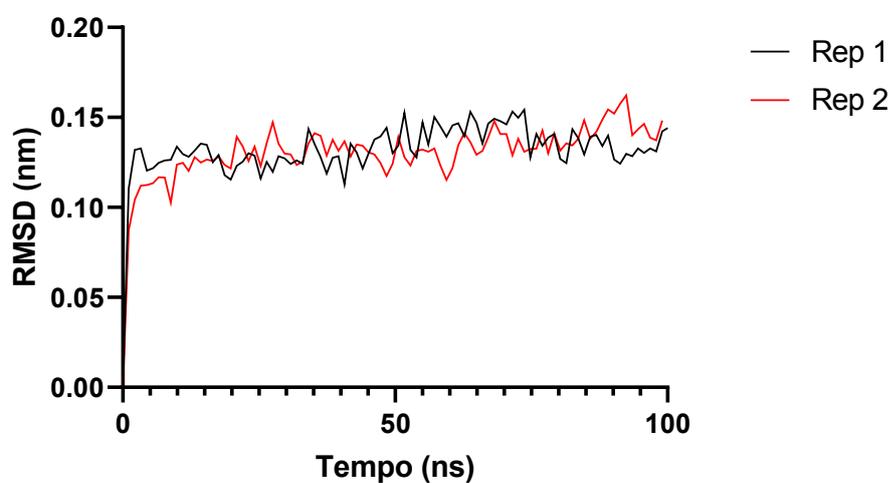
Nesta etapa, selecionamos 3 fármacos que saíram do VS realizado e descrito no tópico 6.1 deste trabalho. Os compostos escolhidos foram imatinib, saquinavir e lumacaftor. A escolha se deu, pois, esses fármacos foram bem ranqueados no protocolo de VS e por

possuírem características químicas que se assemelham a antagonistas da integrina VLA-4. Suas fórmulas estruturais estão descritas na Figura 24.

Primeiramente, analisamos o RMSD da proteína ao longo das simulações para todos os sistemas estudados (Figura 30).

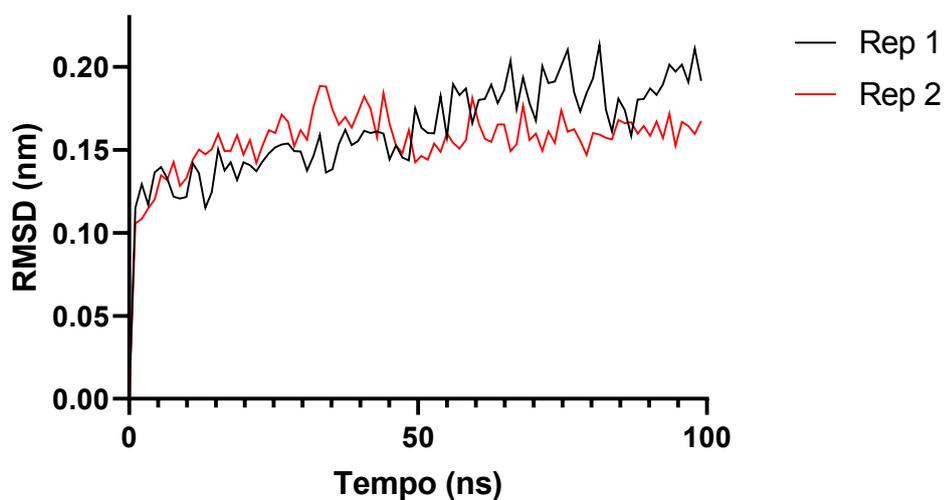
Figura 30 - RMSD (do inglês, *Root Mean Square Deviation*) do *backbone* de VLA-4 para cada sistema estudado, em duas réplicas. Gráficos produzidos no programa GraphPad Prism 8.0.2

### RMSD - VLA-4 (Sistema com imatinib)

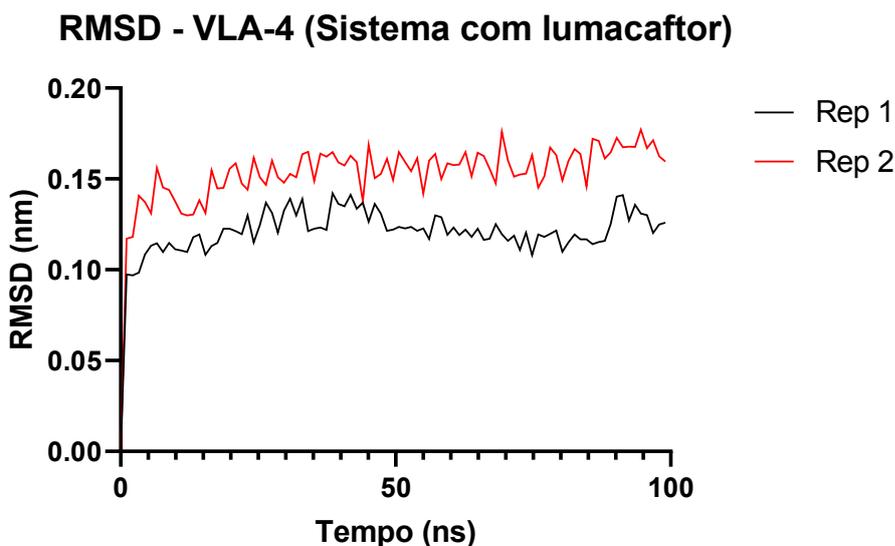


a)

### RMSD - VLA-4 (Sistema com saquinavir)



b)

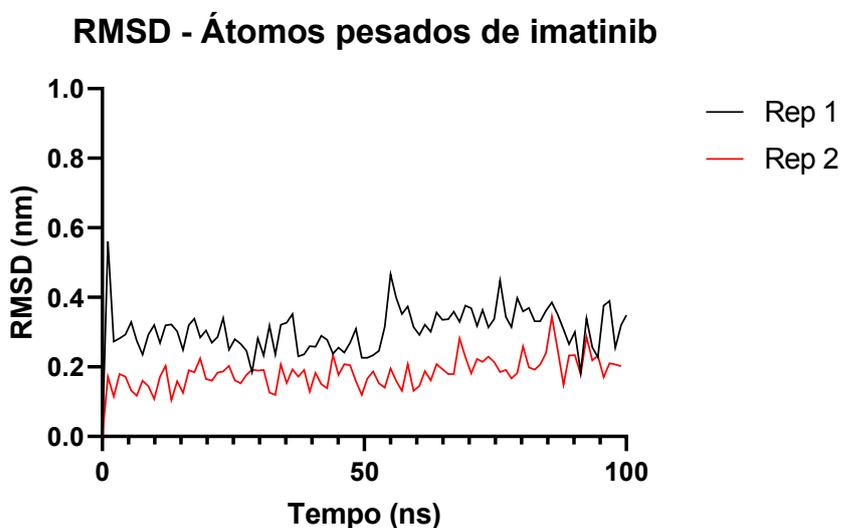


Fonte: O autor.

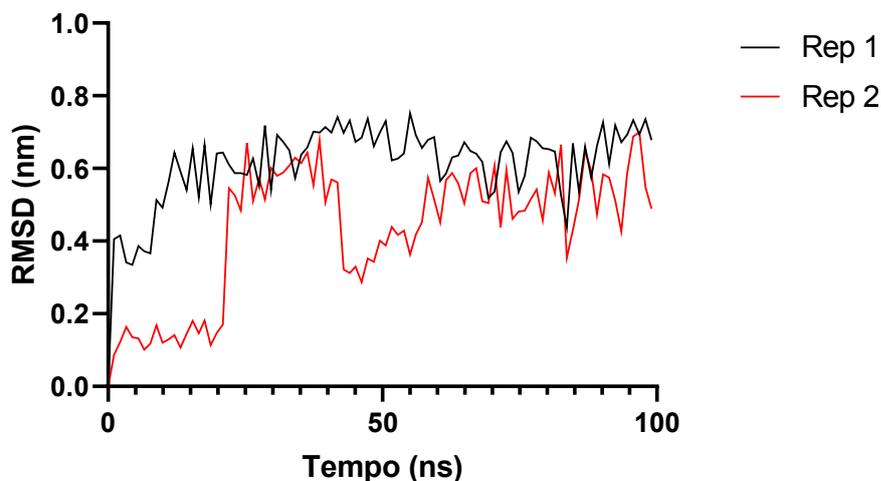
A proteína, de modo geral, se comporta de maneira estável ao longo das simulações realizadas, em torno de 1,5 Å. O sistema com o fármaco saquinavir apresenta maiores flutuações de RMSD, com a réplica 1 chegando próximo a 2 Å, porém ainda sem garantir maiores problemas estruturais na proteína.

Em seguida, estudamos o RMSD dos átomos pesados dos fármacos nas duas réplicas aplicadas (Figura 31).

Figura 31 - RMSD (do inglês, *Root Mean Square Deviation*) dos átomos pesados de a) imatinib; b) saquinavir; c) lumacaftor, em duas réplicas. Gráficos produzidos no programa GraphPad Prism 8.0.2

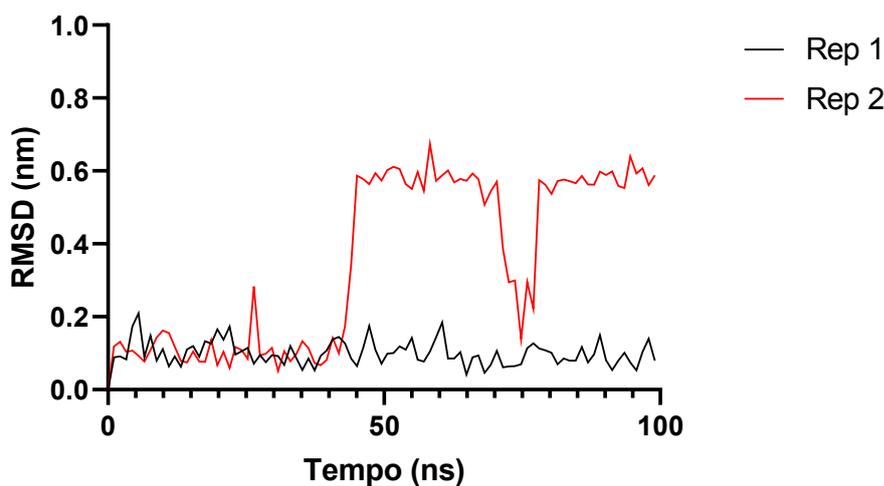


### RMSD - Átomos pesados de saquinavir



b)

### RMSD - Átomos pesados de lumacaftor



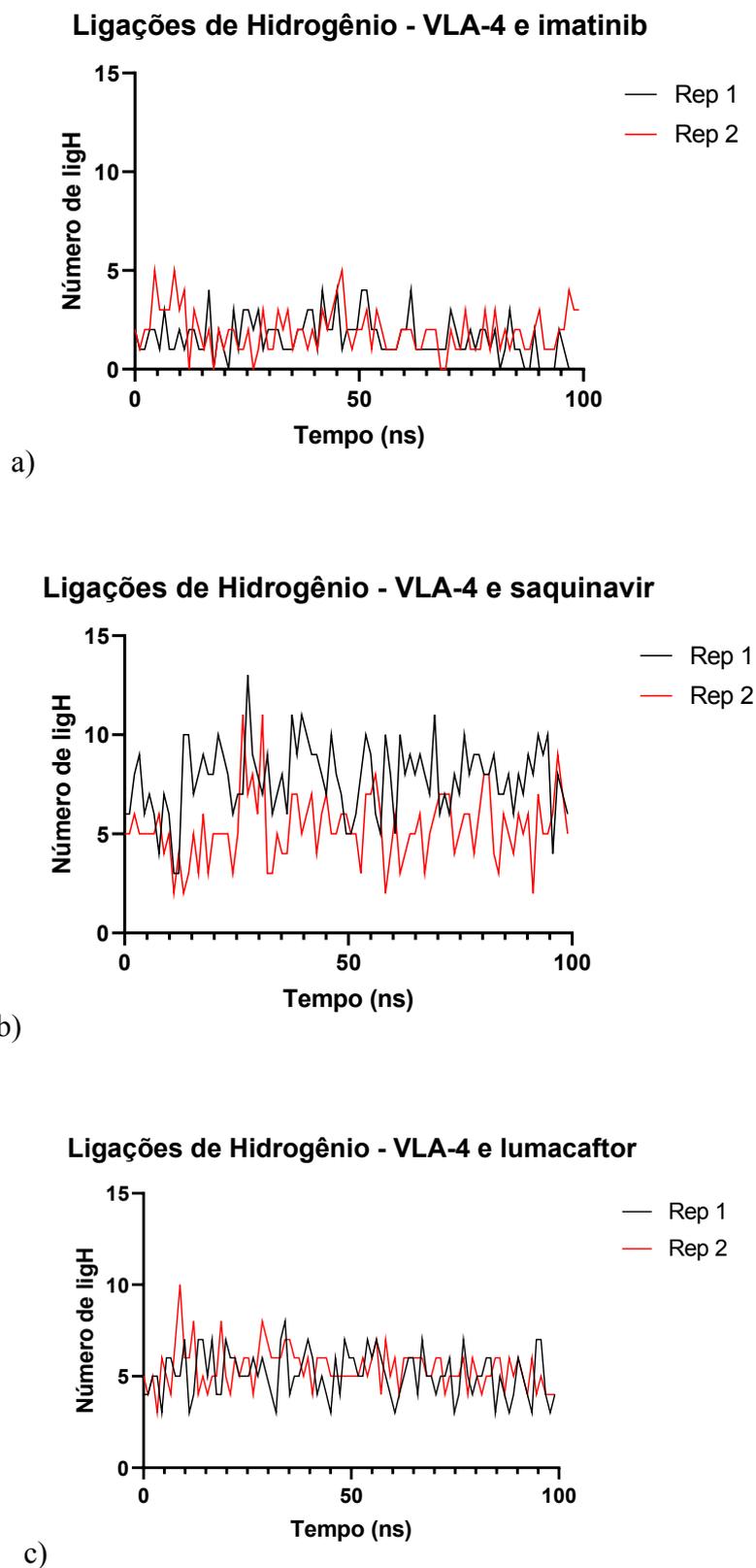
c)

Fonte: O autor.

Portanto, o sistema que se manteve a maior parte do tempo de simulação estável foi com o fármaco imatinib. O composto lumacaftor apresentou, na réplica 1, relativa estabilidade estrutural no sítio de VLA-4, porém, na segunda réplica, houve grandes flutuações nos valores de RMSD, representando instabilidade. O mesmo aconteceu para o fármaco saquinavir.

O número de ligações de hidrogênio entre integrina e fármaco também foi avaliado, para cada sistema estudado, conforme a figura abaixo (Figura 32).

Figura 32 – Número de ligações de hidrogênio nas duas réplicas entre VLA-4 e os fármacos a) imatinib; b) saquinavir; c) lumacaftor. Gráficos produzidos no programa GraphPad Prism 8.0.2



Fonte: O autor.

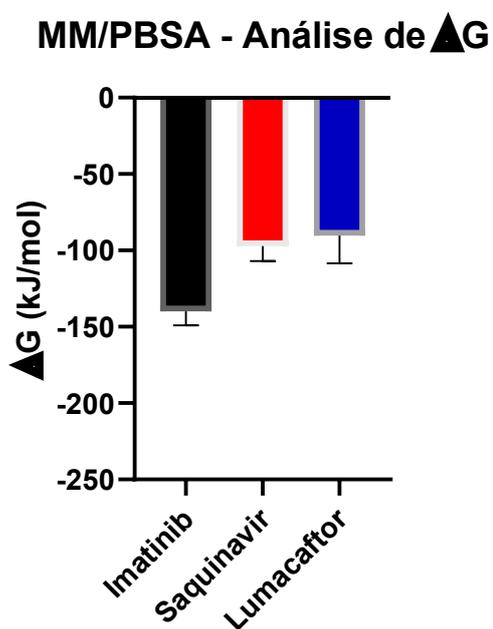
Percebemos uma variação do número de ligações de hidrogênio com o alvo em torno de 5, com o fármaco saquinavir apresentando maiores flutuações nos valores obtidos e o imatinib mostrando uma média menor que os demais, apesar de que, em alguns momentos, na segunda réplica, atingir 5 ligações com o receptor.

#### **6.4 Análises de cálculos de energia livre dos fármacos oriundos do reposicionamento de fármacos na integrina $\alpha 4\beta 1$**

Por último, apresentamos os resultados das análises de MM/PBSA realizadas para os complexos fármaco-integrina. Novamente, utilizamos os mesmos compostos cujas análises de DM foram mostradas no tópico anterior. Para os cálculos de energia livre de ligação, foram considerados apenas os últimos 50 ns de simulação dos complexos, logo, obtivemos 20 frames de simulação para analisar.

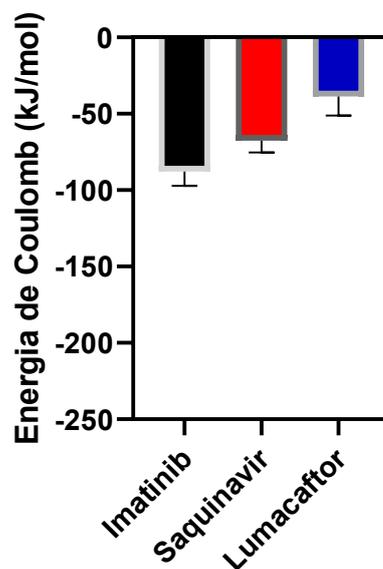
As contribuições energéticas de cada fármaco em complexo com VLA-4 são descritas a seguir (Figura 33).

Figura 33 – Interpolação linear de dados dos potenciais: a)  $\Delta G$  (Imatinib – Média = -139,98 kJ/mol e DP = 18,33 kJ/mol; Lumacaftor – Média = -90,37 kJ/mol e DP = 44,57 kJ/mol; Saquinavir – Média = -97,13 kJ/mol e DP = 26,54 kJ/mol); b) Coulomb (Imatinib – Média = -87,79 kJ/mol e DP = 33,73 kJ/mol; Lumacaftor – Média = -38,60 kJ/mol e DP = 37,68 kJ/mol; Saquinavir – Média = -67,53 kJ/mol e DP = 39,03 kJ/mol); c) van der Waals (Imatinib – Média = -238,74 kJ/mol e DP = 14,40 kJ/mol; Lumacaftor – Média = -162,98 kJ/mol e DP = 33,67 kJ/mol; Saquinavir – Média = -206,81 kJ/mol e DP = 16,47 kJ/mol) para os cálculos de MM/PBSA dos últimos 50 ns de simulação de 2 réplicas dos sistemas formados por VLA-4 e os compostos estudados no reposicionamento de fármacos. Gráficos produzidos no programa GraphPad Prism 8.0.2. DP = Desvio-padrão



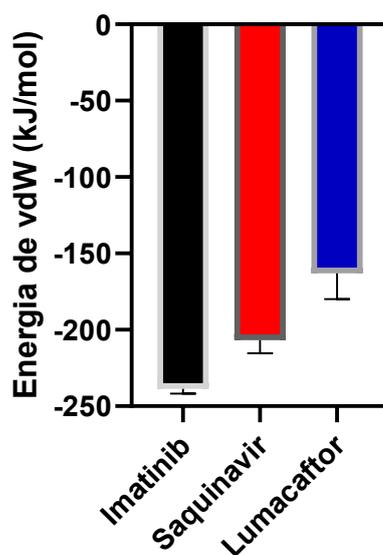
a)

### MM/PBSA - Análise da Energia de Coulomb



b)

### MM/PBSA - Análise da Energia de van der Waals



c)

Fonte: O autor.

O fármaco imatinib apresenta maior afinidade em VLA-4 quando comparado aos demais, por possuir menor média de  $\Delta G$  e menor  $\Delta G_{\text{ligação}}$  final calculado (Tabela 8). Imatinib também mostra maiores contribuições nas energias de Coulomb e de van der Waals, representando maior magnitude em interações eletrostáticas e na soma das interações fracas com o alvo.

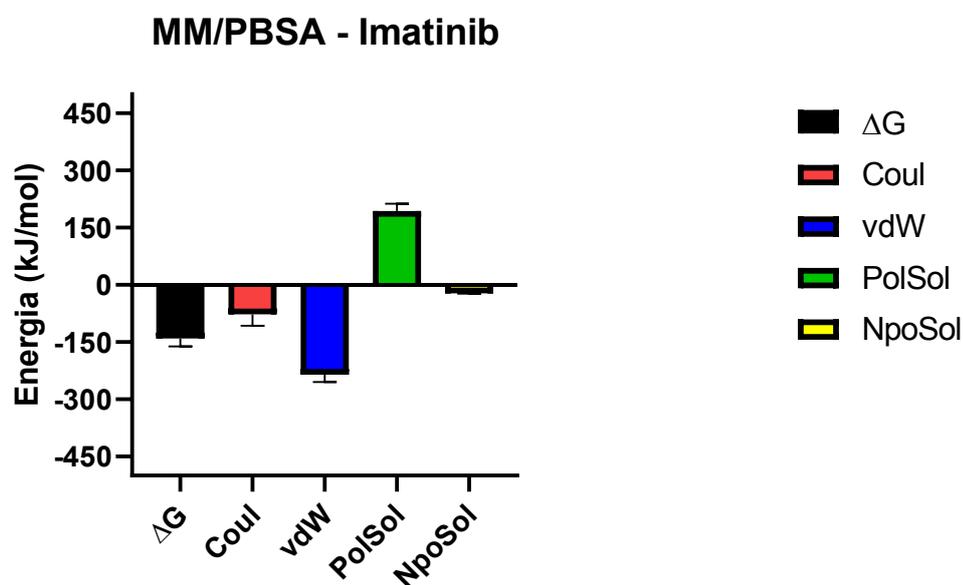
Tabela 8 – Valores finais de  $\Delta G_{\text{ligação}}$  para os complexos de VLA-4 e os fármacos estudados, a partir dos cálculos por MM/PBSA de uma réplica de simulação

Fármaco	$\Delta G_{\text{ligação}}$ final			
Imatinib	-140,291	+/-	4,391	kJ/mol
Saquinavir	-95,309	+/-	7,468	kJ/mol
Lumacaftor	-68,589	+/-	5,608	kJ/mol

Fonte: O autor.

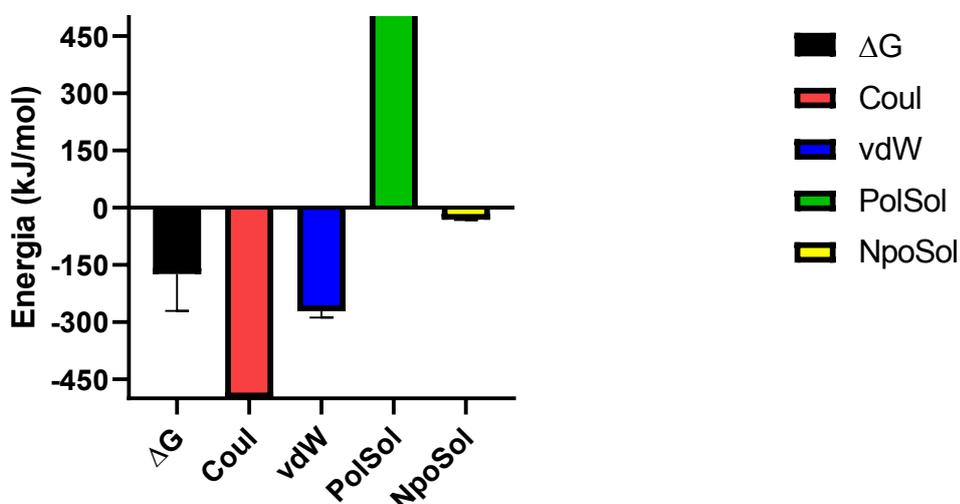
É importante notar, através da Tabela 8, que o imatinib possui um  $\Delta G_{\text{ligação}}$  final bem menor que os demais compostos, mostrando, assim, uma afinidade bem maior a VLA-4, em termos comparativos. E para exemplificar esse dado, detalhamos abaixo o resultado, após análise estatística, das contribuições energéticas nos cálculos de MM/PBSA para o fármaco imatinib em comparação ao composto-líder BIO-5192, levando em consideração, em ambos os casos, todas as réplicas de simulação dos ligantes complexados à integrina VLA-4 (Figura 34).

Figura 34 – Contribuições energéticas de  $\Delta G$ , Energia de Coulomb, Energia de van der Waals, Energia de Solvatação Polar e Energia de Solvatação não polar para os cálculos de MM/PBSA dos complexos de a) imatinib e b) BIO-5192 em VLA-4



a)

### MM/PBSA - BIO-5192



b)

Fonte: O autor.

Se visualizarmos a média de  $\Delta G$  para os dois compostos, podemos concluir que ambos apresentam afinidades próximas a VLA-4. Esse fato corrobora o reconhecimento favorável que o fármaco imatinib possui com a integrina. Se compararmos as interações fracas integrina-ligante (forças de van der Waals), também obtivemos médias parecidas. Os potenciais que destoam são aqueles relacionados à polaridade das moléculas, como energias de Coulomb e de solvatação polar, fazendo com que o líder BIO-5192 mostrasse maior magnitude nesse aspecto. Tal fato era previsível devido à presença de mais grupamentos ionizáveis em BIO-5192, quando comparado ao fármaco imatinib.

\*\*\*

Recentemente, o uso de técnicas de reposicionamento de fármacos para o tratamento do câncer vem sendo reportada. Kandasamy e colaboradores aplicaram um método multialvo na busca por moléculas que podiam combater o câncer de mama, utilizando análises de redes funcionais integradas. Nesse estudo, um dos receptores envolvidos foi a integrina  $\beta 1$  (ITGB1) (KANDASAMY; SEN; GHOSH, 2022).

Outro estudo recente, envolvendo o receptor ILK (do inglês, cinase ligada à integrina) mostrou que o fármaco nilotinib, também analisado nesta tese, apresentou bons resultados no VS de compostos aprovados pelo FDA. Inclusive, o fármaco foi submetido a testes *in vitro* por Ressonância Plasmônica de Superfície (SPR) e de citotoxicidade,

sugerindo-o como potencial inibidor da proteína ILK.

Apesar disso, estudos envolvendo a integrina  $\alpha4\beta1$  em reposicionamento de fármacos aprovados pela FDA são inexistentes. Desse modo, este trabalho tenta lançar luz nesse tema emergente.

## 7 CONCLUSÕES

Este estudo apresenta a descoberta e análise de compostos protótipos e de fármacos comerciais que poderão servir de base para o tratamento de doenças inflamatórias crônicas, nomeadamente: esclerose múltipla, artrite reumatoide, DII, entre outras.

Ao final das simulações de dinâmica molecular, descobrimos que o ligante 1592 apresentou um comportamento estável no sítio de interação VLA-4, indicando pouca mudança quanto à posição do modo de ligação inicial do *docking*, um resultado similar ao do composto líder BIO-5192. As análises de PCA também sugeriram que o receptor VLA-4 apresentou modos de movimento em sua estrutura diferentes quando complexada a 1592, em comparação aos demais compostos. Além disso, pelos cálculos de energia livre de ligação, 1592 apresentou maior afinidade à VLA-4 em relação às outras moléculas estudadas. No geral, nossos dados indicam fortemente que o ligante 1592 representa um possível candidato a um fármaco inibidor de VLA-4.

O mesmo ocorreu para o fármaco imatinib. Estudando os dados oriundos do VS, método que foi devidamente validado neste trabalho, das simulações de DM e do MM/PBSA, sugerimos o fármaco um potencial candidato a antagonista da integrina VLA-4.

Dessa forma, os resultados apresentados lançam novas perspectivas no estudo de diferentes pequenas moléculas potenciais antagonistas da integrina  $\alpha 4\beta 1$  e abre possibilidades ao estudo de sítios alostéricos no receptor proteico, que ainda carece de mais detalhes na comunidade científica.

Logo, propusemos a estratégia do uso de fármacos comerciais como alternativa às dificuldades impostas à síntese química de novas pequenas moléculas, técnica que deveria ser adotada como próxima etapa, a fim da aplicação de experimentos *in vitro* e *in vivo* dos novos derivados de BIO-5192, para atestar sua eficácia em inibir a atividade de  $\alpha 4\beta 1$  em células de mamíferos.

## REFERÊNCIAS

- ABRAHAM, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. **SoftwareX**, v. 1–2, p. 19–25, 1 set. 2015.
- AIN, Q. U. *et al.* Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. **Wiley Interdisciplinary Reviews. Computational Molecular Science**, v. 5, n. 6, p. 405, 1 nov. 2015.
- AKSORN, N.; CHANVORACHOTE, P. Integrin as a molecular target for anti-cancer approaches in lung cancer. **Anticancer Research**, v. 39, n. 2, p. 541–548, 1 fev. 2019.
- AMARO, R. E. *et al.* **Ensemble Docking in Drug Discovery**. **Biophysical Journal** Biophysical Society, , 22 maio 2018.
- AMIN, S. A. *et al.* An integrated QSAR modeling approach to explore the structure-property and selectivity relationships of N-benzoyl-L-biphenylalanines as integrin antagonists. **Molecular diversity**, v. 22, n. 1, p. 129–158, 1 fev. 2018.
- ANDRADE, L. *et al.* Investigation of Unprecedented Sites and Proposition of New Ligands for Programmed Cell Death Protein 1 through Molecular Dynamics with Probes and Virtual Screening. **Journal of chemical information and modeling**, v. 62, n. 5, p. 1236–1248, 14 mar. 2022.
- ARNAOUT, M. A.; MAHALINGAM, B.; XIONG, J. P. Integrin structure, allostery, and bidirectional signaling. **Annual review of cell and developmental biology**, v. 21, p. 381–410, 2005.
- ASHBURN, T. T.; THOR, K. B. Drug repositioning: identifying and developing new uses for existing drugs. **Nature Reviews Drug Discovery** **2004 3:8**, v. 3, n. 8, p. 673–683, 2004.
- ASOKAN, A. *et al.* Image Processing Techniques for Analysis of Satellite Images for Historical Maps Classification—An Overview. **Applied Sciences**, 2020, Vol. 10, Page 4207, v. 10, n. 12, p. 4207, 19 jun. 2020.
- BADAoui, M. *et al.* Combined Free-Energy Calculation and Machine Learning Methods for Understanding Ligand Unbinding Kinetics. **Journal of Chemical Theory and Computation**, v. 18, n. 4, p. 2543–2555, 12 abr. 2022.
- BAIULA, M. *et al.* Novel Ligands Targeting  $\alpha 4\beta 1$  Integrin: Therapeutic Applications and Perspectives. **Frontiers in Chemistry**, 2019a.
- BAIULA, M. *et al.* Novel Ligands Targeting  $\alpha 4\beta 1$  Integrin : Therapeutic Applications and Perspectives. **Frontiers in Chemistry**, v. 7, n. July, p. 1–12, 2019b.
- BANEGAS-LUNA, A. J.; CERÓN-CARRASCO, J. P.; PÉREZ-SÁNCHEZ, H. A review of ligand-based virtual screening web tools and screening algorithms in large molecular databases in the age of big data. **Future Med Chem.**, v. 10, n. 22, p. 2641–2658, 30 nov. 2018.

BARCZYK, M.; CARRACEDO, S.; GULLBERG, D. Integrins. **Cell Tissue Res**, v. 339, n. 1, p. 269–280, 2010.

BERENDSEN, H. J. C.; VAN DER SPOEL, D.; VAN DRUNEN, R. GROMACS: A message-passing parallel molecular dynamics implementation. **Computer Physics Communications**, v. 91, n. 1–3, p. 43–56, 2 set. 1995.

BERLIN, C. *et al.* Alpha 4 beta 7 integrin mediates lymphocyte binding to the mucosal vascular addressin MAdCAM-1. **Cell**, v. 74, n. 1, p. 185–195, 1993.

BLOOMGREN, G. *et al.* Risk of natalizumab-associated progressive multifocal leukoencephalopathy. **N Engl J Med**, v. 366, n. 20, p. 1870–1880, 2012.

BRECKENRIDGE, A.; JACOB, R. Overcoming the legal and regulatory barriers to drug repurposing. **Nature reviews. Drug discovery**, v. 18, n. 1, p. 1–2, 28 dez. 2019.

BRO, R.; SMILDE, A. K. Principal component analysis. **Analytical Methods**, v. 6, n. 9, p. 2812–2831, 10 abr. 2014.

BRUDER, M.; POLO, G.; TRIVELLA, D. B. B. Natural allosteric modulators and their biological targets: molecular signatures and mechanisms. **Natural Product Reports**, v. 37, n. 4, p. 488–514, 29 abr. 2020.

CAVASOTTO, C. N. Binding Free Energy Calculation Using Quantum Mechanics Aimed for Drug Lead Optimization. **Methods in molecular biology (Clifton, N.J.)**, v. 2114, p. 257–268, 2020.

CHANG, C. E. A. *et al.* Investigation of Structural Dynamics of Enzymes and Protonation States of Substrates Using Computational Tools. **Catalysts (Basel, Switzerland)**, v. 6, n. 6, 1 jun. 2016.

CHAVAKIS, T. Leucocyte recruitment in inflammation and novel endogenous negative regulators thereof. **Eur J Clin Invest**, v. 42, n. 6, p. 686–691, 2012.

CHAVES, B. *et al.* Guidelines to Predict Binding Poses of Antibody-Integrin Complexes. **ACS Omega**, v. 5, n. 27, p. 16379–16385, 14 jul. 2020.

CHEN, W.; VAN DER KAMP, M. W.; DAGGETT, V. Structural and Dynamic Properties of the Human Prion Protein. **Biophysical Journal**, v. 106, n. 5, p. 1152, 3 mar. 2014.

D'AGOSTINO, D. *et al.* Cloud infrastructures for in silico drug discovery: Economic and practical aspects. **BioMed Research International**, v. 2013, 2013.

DALLAKYAN, S.; OLSON, A. J. Small-molecule library screening by docking with PyRx. **Methods in Molecular Biology**, 2015.

DARDEN, T.; YORK, D.; PEDERSEN, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. **The Journal of Chemical Physics**, v. 98, n. 12, p. 10089, 31 ago. 1998.

- DARGAHI, N. *et al.* **Multiple sclerosis: Immunopathology and treatment update.** *Brain Sciences*, 2017.
- DE VIVO, M. *et al.* Role of Molecular Dynamics and Related Methods in Drug Discovery. *Journal of Medicinal Chemistry*, v. 59, n. 9, p. 4035–4061, 12 maio 2016.
- DIAMOND, M. S.; SPRINGER, T. A. The dynamic regulation of integrin adhesiveness. *Curr Biol*, v. 4, n. 6, p. 506–517, 1994.
- DIMASI, J. A.; GRABOWSKI, H. G.; HANSEN, R. W. Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, v. 47, p. 20–33, 1 maio 2016.
- DOS SANTOS, R. N.; FERREIRA, L. G.; ANDRICOPULO, A. D. Practices in molecular docking and structure-based virtual screening. *Methods in Molecular Biology*, v. 1762, p. 31–50, 2018.
- DOTAN, I. *et al.* The role of integrins in the pathogenesis of inflammatory bowel disease: Approved and investigational anti-integrin therapies. *Medicinal Research Reviews*, v. 40, n. 1, p. 245, 1 jan. 2020.
- DURRANT, J. D.; MCCAMMON, J. A. Molecular dynamics simulations and drug discovery. *BMC Biology*, v. 9, n. 1, p. 1–9, 28 out. 2011.
- ESFANDIAREI, M. *et al.* Novel role for integrin-linked kinase in modulation of coxsackievirus B3 replication and virus-induced cardiomyocyte injury. *Circ Res*, v. 99, n. 4, p. 354–361, 2006.
- EVANGELISTA FALCON, W. *et al.* Ensemble Docking in Drug Discovery: How Many Protein Configurations from Molecular Dynamics Simulations are Needed to Reproduce Known Ligand Binding? *Journal of Physical Chemistry B*, v. 123, n. 25, p. 5189–5195, 27 jun. 2019.
- EVANGELISTA, W. *et al.* Ensemble-based docking: From hit discovery to metabolism and toxicity predictions. *Bioorganic & Medicinal Chemistry*, v. 24, n. 20, p. 4928–4935, 15 out. 2016.
- EYAL, E. *et al.* Importance of solvent accessibility and contact surfaces in modeling side-chain conformations in proteins. *Journal of Computational Chemistry*, v. 25, n. 5, p. 712–724, 15 abr. 2004.
- FIGUEIREDO, J. *et al.* Integrin  $\beta 1$  orchestrates the abnormal cell-matrix attachment and invasive behaviour of E-cadherin dysfunctional cells. *Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association*, v. 25, n. 1, p. 124–137, 1 jan. 2022.
- GAHMBERG, C. G. *et al.* Regulation of integrin activity and signalling. *Biochimica et biophysica acta*, v. 1790, n. 6, p. 431–444, jun. 2009.
- GAULTON, A. *et al.* ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic*

**Acids Research**, 2012.

GENHEDEN, S.; RYDE, U. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. **Expert Opinion on Drug Discovery**, v. 10, n. 5, p. 449, 5 maio 2015.

GIANCOTTI, F. G.; RUOSLAHTI, E. Integrin signaling. **Science**, v. 285, n. 5430, p. 1028–1032, 1999.

GILSON, M. K.; ZHOU, H. X. Calculation of protein-ligand binding affinities. **Annual review of biophysics and biomolecular structure**, v. 36, p. 21–42, 2007.

GINSBERG, M. H.; PARTRIDGE, A.; SHATTIL, S. J. Integrin regulation. **Current opinion in cell biology**, v. 17, n. 5, p. 509–516, 2005.

GORDON, J. C. *et al.* H<sup>++</sup>: a server for estimating pK<sub>a</sub>s and adding missing hydrogens to macromolecules. **Nucleic Acids Research**, v. 33, n. Web Server issue, p. W368, jul. 2005.

GUBATAN, J. *et al.* Anti-Integrins for the Treatment of Inflammatory Bowel Disease: Current Evidence and Perspectives. **Clinical and experimental gastroenterology**, v. 14, p. 333–342, 2021.

GUEDES, I. A.; PEREIRA, F. S. S.; DARDENNE, L. E. Empirical scoring functions for structure-based virtual screening: Applications, critical aspects, and challenges. **Frontiers in Pharmacology**, v. 9, n. SEP, p. 1089, 24 set. 2018.

HAAKE, S. M. *et al.* Ligand-independent integrin  $\beta$ 1 signaling supports lung adenocarcinoma development. **JCI insight**, v. 7, n. 15, 8 ago. 2022.

HAGMANN, W. K. *et al.* The discovery of sulfonylated dipeptides as Potent VLA-4 antagonists. **Bioorganic and Medicinal Chemistry Letters**, v. 11, n. 20, p. 2709–2713, 22 out. 2001.

HANWELL, M. D. *et al.* Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. **Journal of Cheminformatics**, v. 4, n. 8, p. 1–17, 13 ago. 2012.

HATLEY, R. J. D. *et al.* The Design of Potent, Selective and Drug-Like RGD  $\alpha$  $\beta$ 1 Small-Molecule Inhibitors Derived from non-RGD  $\alpha$ 4 $\beta$ 1 Antagonists. **ChemMedChem**, v. 14, n. 14, p. 1315–1320, 17 jul. 2019.

HAZARIKA, R. R. *et al.* Large-scale docking predicts that sORF-encoded peptides may function through protein-peptide interactions in *Arabidopsis thaliana*. **PLOS ONE**, v. 13, n. 10, p. e0205179, 1 out. 2018.

HE, Q. *et al.* Dipeptidyl Peptidase-4 Stabilizes Integrin  $\alpha$ 4 $\beta$ 1 Complex to Promote Thyroid Cancer Cell Metastasis by Activating Transforming Growth Factor-Beta Signaling Pathway. **Thyroid : official journal of the American Thyroid Association**, v. 32, n. 11, p. 1411–1422, 1 nov. 2022.

HEDRICK, E. *et al.* NR4A1 Antagonists Inhibit  $\beta$ 1-Integrin-Dependent Breast Cancer Cell Migration. **Molecular and Cellular Biology**, 2016.

- HEMLER, M. E. VLA proteins in the integrin family: structures, functions, and their role on leukocytes. **Annu Rev Immunol**, v. 8, p. 365–400, 1990.
- HIJAZI, Y. *et al.* Pharmacokinetics, safety, and tolerability of R411, a dual  $\alpha 4\beta 1$ - $\alpha 4\beta 7$  integrin antagonist after oral administration at single and multiple once-daily ascending doses in healthy volunteers. **J Clin Pharmacol**, v. 44, n. 12, p. 1368–1378, 2004.
- HILL, R. G.; RICHARDS, D. Drug discovery and development : technology in transition. p. 373, [s.d.].
- HOFFMANN, C.; OHLSEN, K.; HAUCK, C. R. Integrin-mediated uptake of fibronectin-binding bacteria. **Eur J Cell Biol**, v. 90, n. 11, p. 891–896, 2011.
- HU, Y. *et al.* Application of Machine Learning Approaches for the Design and Study of Anticancer Drugs. **Current Drug Targets**, v. 20, n. 5, p. 488–500, 25 ago. 2018.
- HUANG, D.; CAFLISCH, A. The Free Energy Landscape of Small Molecule Unbinding. **PLOS Computational Biology**, v. 7, n. 2, p. e1002002, 2011.
- HUMPHRIES, J. D.; BYRON, A.; HUMPHRIES, M. J. Integrin ligands at a glance. **Journal of cell science**, v. 119, n. Pt 19, p. 3901–3903, 1 out. 2006.
- HUMPHRIES, M. J. Integrin structure. **Biochem Soc Trans**, v. 28, n. 4, p. 311–339, 2000.
- HURLE, M. R. *et al.* Computational drug repositioning: from data to therapeutics. **Clinical pharmacology and therapeutics**, v. 93, n. 4, p. 335–341, abr. 2013.
- HURYN, D. *et al.* The Identification and Optimization of Orally Efficacious, Small Molecule VLA-4 Antagonists. **Current Topics in Medicinal Chemistry**, v. 4, n. 14, p. 1473–1484, 18 mar. 2005.
- HUSSEIN, H. A. *et al.* Beyond RGD: virus interactions with integrins. **Arch Virol**, v. 160, n. 11, p. 2669–2681, 2015.
- HUTT, O. E.; SAUBERN, S.; WINKLER, D. A. Modeling the molecular basis for  $\alpha 4\beta 1$  integrin antagonism. **Bioorganic & Medicinal Chemistry**, v. 19, n. 19, p. 5903–5911, 1 out. 2011.
- HYNES, R. O. Integrins: a family of cell surface receptors. **Cell**, v. 48, n. 4, p. 549–554, 1987.
- HYNES, R. O. Integrins: versatility, modulation, and signaling in cell adhesion. **Cell**, v. 69, n. 1, p. 11–25, 1992.
- HYNES, R. O. Integrins : Bidirectional , Allosteric Signaling Machines In their roles as major adhesion receptors. **Cell**, v. 110, n. 6, p. 673–687, 2002.
- JACKSON, D. Y. Alpha 4 integrin antagonists. **Curr Pharm Des**, v. 8, n. 14, p. 1229–1253, 2002.
- JOHNSON, M. S. *et al.* Integrins during evolution: evolutionary trees and model organisms.

**Biochimica et biophysica acta**, v. 1788, n. 4, p. 779–789, abr. 2009.

JONES, E. Y. *et al.* Crystal structure of an integrin-binding fragment of vascular cell adhesion molecule-1 at 1.8 Å resolution. **Nature**, v. 373, n. 6514, p. 539–544, 9 fev. 1995.

JUSTO, B. L.; JASIULIONIS, M. G. **Characteristics of timp1, cd63, and  $\beta$ 1-integrin and the functional impact of their interaction in cancer.** **International Journal of Molecular Sciences** MDPI, , 1 set. 2021.

KANDASAMY, T.; SEN, P.; GHOSH, S. S. Multi-targeted Drug Repurposing Approach for Breast Cancer via Integrated Functional Network Analysis. **Molecular informatics**, v. 41, n. 8, 1 ago. 2022.

KELSH-LASHER, R. M. *et al.* Integrin  $\alpha$ 4 $\beta$ 1 and TLR4 Cooperate to Induce Fibrotic Gene Expression in Response to Fibronectin's EDA Domain. **Journal of Investigative Dermatology**, v. 137, n. 12, p. 2505–2512, 1 dez. 2017.

KHOY, K. *et al.* Natalizumab in Multiple Sclerosis Treatment: From Biological Effects to Immune Monitoring. **Frontiers in immunology**, v. 11, 24 set. 2020.

KIM, A. G. *et al.* Enhanced in utero allogeneic engraftment in mice after mobilizing fetal HSCs by  $\alpha$ 4 $\beta$ 1/7 inhibition. **Blood**, v. 128, n. 20, p. 2457–2461, 17 nov. 2016.

KIM, K. H.; KIM, N. D.; SEONG, B. L. Pharmacophore-based virtual screening: A review of recent applications. **Expert Opinion on Drug Discovery**, v. 5, n. 3, p. 205–222, mar. 2010.

KING, E. *et al.* Recent Developments in Free Energy Calculations for Drug Discovery. **Frontiers in molecular biosciences**, v. 8, 11 ago. 2021a.

KING, E. *et al.* Recent Developments in Free Energy Calculations for Drug Discovery. **Frontiers in molecular biosciences**, v. 8, 11 ago. 2021b.

KITCHEN, D. B. *et al.* Docking and scoring in virtual screening for drug discovery: Methods and applications. **Nature Reviews Drug Discovery**, v. 3, n. 11, p. 935–949, nov. 2004.

KOLLMAN, P. A. *et al.* Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. **Accounts of chemical research**, v. 33, n. 12, p. 889–897, 2000.

KOMORIYA, A. *et al.* The minimal essential sequence for a major cell type-specific adhesion site (CS1) within the alternatively spliced type III connecting segment domain of fibronectin is leucine-aspartic acid-valine. **Journal of Biological Chemistry**, v. 266, n. 23, p. 15075–15079, 15 ago. 1991.

LABUTE, M. X. *et al.* Adverse Drug Reaction Prediction Using Scores Produced by Large-Scale Drug-Protein Target Docking on High-Performance Computing Machines. **PLOS ONE**, v. 9, n. 9, p. e106298, 5 set. 2014.

LEONE, D. R. *et al.* An assessment of the mechanistic differences between two integrin  $\alpha$ 4  $\beta$ 1 inhibitors, the monoclonal antibody TA-2 and the small molecule BIO5192, in rat

experimental autoimmune encephalomyelitis. **J Pharmacol Exp Ther**, v. 305, n. 3, p. 1150–1162, 2003.

LEY, K. *et al.* Integrin-based therapeutics: biological basis, clinical use and new drugs. **Nature reviews. Drug discovery**, v. 15, n. 3, p. 173–183, 1 mar. 2016.

LI, H. *et al.* TarFisDock: a web server for identifying drug targets with docking approach. **Nucleic Acids Research**, v. 34, n. Web Server issue, p. W219, 7 jul. 2006.

LI, S. *et al.* Molecular Modeling Insights into the Structure and Behavior of Integrins: A Review. **Cells** **2023**, Vol. 12, Page 324, v. 12, n. 2, p. 324, 14 jan. 2023.

LIN, K. *et al.* Selective, tight-binding inhibitors of integrin  $\alpha 4\beta 1$  that inhibit allergic airway responses. **J Med Chem**, v. 42, n. 5, p. 920–934, 1999.

**LINCS: A linear constraint solver for molecular simulations - Hess - 1997 - Journal of Computational Chemistry - Wiley Online Library.** Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/%28SICI%291096-987X%28199709%2918%3A12%3C1463%3A%3AAID-JCC4%3E3.0.CO%3B2-H>>. Acesso em: 16 abr. 2022.

LINDORFF-LARSEN, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. **Proteins**, v. 78, n. 8, p. 1950, jun. 2010.

LIU, R. *et al.* Structure-activity relationship studies of a series of peptidomimetic ligands for  $\alpha 4\beta 1$  integrin on Jurkat T-leukemia cells. **Biopolymers - Peptide Science Section**, v. 84, n. 6, p. 595–604, 2006.

LUO, B. H.; CARMAN, C. V.; SPRINGER, T. A. Structural basis of integrin regulation and signaling. **Annual review of immunology**, v. 25, p. 619–647, 2007.

MACALINO, S. J. Y. *et al.* Role of computer-aided drug design in modern drug discovery. **Archives of Pharmacal Research** **2015** **38:9**, v. 38, n. 9, p. 1686–1701, 25 jul. 2015.

MAIER, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. **Journal of Chemical Theory and Computation**, v. 11, n. 8, p. 3696–3713, 7 jul. 2015.

MARK, P.; NILSSON, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. **Journal of Physical Chemistry A**, v. 105, n. 43, p. 9954–9960, 1 nov. 2001.

MARKTHALER, D.; KRAUS, H.; HANSEN, N. Binding free energies for the SAMPL8 CB8 “Drugs of Abuse” challenge from umbrella sampling combined with Hamiltonian replica exchange. **Journal of computer-aided molecular design**, v. 36, n. 1, 1 jan. 2022.

MCGANN, M.; NICHOLLS, A.; ENYEDY, I. The statistics of virtual screening and lead optimization. **Journal of Computer-Aided Molecular Design**, v. 29, n. 10, p. 923–936, 1 out. 2015.

MORRIS, G. M. *et al.* Software news and updates AutoDock4 and AutoDockTools4:

Automated docking with selective receptor flexibility. **Journal of Computational Chemistry**, v. 30, n. 16, p. 2785–2791, dez. 2009.

MUEGGE, I.; HU, Y. Recent Advances in Alchemical Binding Free Energy Calculations for Drug Discovery. **ACS medicinal chemistry letters**, v. 14, n. 3, 9 mar. 2023.

NAGAE, M. *et al.* Crystal structure of alpha5beta1 integrin ectodomain: atomic details of the fibronectin receptor. **J Cell Biol**, v. 197, n. 1, p. 131–140, 2012.

NEVO, J. Novel Players in the Integrin Signaling Orchestra: TCPTP and MDGI: Dissertation. 2011.

NIU, G.; CHEN, X. Why integrin as a primary target for imaging and therapy. **Theranostics**, v. 1, p. 30–47, 7 jan. 2011.

NOSENGO, N. Can you teach old drugs new tricks? **Nature**, v. 534, n. 7607, p. 314–316, 14 jun. 2016.

OH, J. *et al.* Age-related tumor growth in mice is related to integrin  $\alpha$  4 in CD8<sup>+</sup> T cells. **JCI insight**, 2018.

OOSTENBRINK, C. *et al.* Validation of the 53A6 GROMOS force field. **European biophysics journal : EBJ**, v. 34, n. 4, p. 273–284, jun. 2005.

P, H. *et al.* Dexamethasone in Hospitalized Patients with Covid-19. **The New England journal of medicine**, v. 384, n. 8, p. 693–704, 25 fev. 2021.

PAISSONI, C. *et al.* GMXPBSA 2.1: A GROMACS tool to perform MM/PBSA and computational alanine scanning. **Computer Physics Communications**, v. 186, p. 105–107, 1 jan. 2015.

PALMA, J.; PIERDOMINICI-SOTTILE, G. On the Uses of PCA to Characterise Molecular Dynamics Simulations of Biological Macromolecules: Basics and Tips for an Effective Use. **ChemPhysChem**, v. 24, n. 2, p. e202200491, 17 jan. 2023.

PANTSAR, T.; POSO, A. Binding Affinity via Docking: Fact and Fiction. **Molecules** **2018**, Vol. **23**, Page **1899**, v. 23, n. 8, p. 1899, 30 jul. 2018.

PATEL, D.; BAUMAN, J. D.; ARNOLD, E. Advantages of crystallographic fragment screening: Functional and mechanistic insights from a powerful platform for efficient drug discovery. **Progress in Biophysics and Molecular Biology**, v. 116, n. 2–3, p. 92–100, 2014.

PELLINEN, T. *et al.* ITGB1-dependent upregulation of Caveolin-1 switches TGF $\beta$  signalling from tumour-suppressive to oncogenic in prostate cancer. **Scientific reports**, v. 8, n. 1, 1 dez. 2018.

PERNECZKY, J.; SELLNER, J. Natalizumab extended-interval dosing in multiple sclerosis to mitigate progressive multifocal leukoencephalopathy risk: initial study evidence and real-world experience. **Journal of Central Nervous System Disease**, v. 14, p. 117957352211354, jan. 2022.

- PETERSEN, E. R. *et al.* Expression of melanoma cell adhesion molecule-1 (MCAM-1) in natalizumab-treated multiple sclerosis. **Journal of Neuroimmunology**, v. 337, 15 dez. 2019. **Pfizer's Expiring Viagra Patent Adversely Affects Other Drugmakers Too**. Disponível em: <<https://www.forbes.com/sites/investor/2013/12/20/pfizers-expiring-viagra-patent-adversely-affects-other-drugmakers-too/?sh=5cafe3968d45>>. Acesso em: 25 mar. 2023.
- PRICE, D. J.; BROOKS, C. L. A modified TIP3P water potential for simulation with Ewald summation. **The Journal of Chemical Physics**, v. 121, n. 20, p. 10096, 11 nov. 2004.
- PROCACCI, P. Methodological uncertainties in drug-receptor binding free energy predictions based on classical molecular dynamics. **Current Opinion in Structural Biology**, v. 67, p. 127–134, 1 abr. 2021.
- PUSHPAKOM, S. *et al.* Drug repurposing: progress, challenges and recommendations. **Nature reviews. Drug discovery**, v. 18, n. 1, p. 41–58, 28 dez. 2019.
- RAMIREZ, P. *et al.* BIO5192, a small molecule inhibitor of VLA-4, mobilizes hematopoietic stem and progenitor cells. **Blood**, v. 114, n. 7, p. 1340–1343, 2009.
- RASTELLI, G. *et al.* Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. **Journal of Computational Chemistry**, v. 31, n. 4, p. 797–810, 1 mar. 2010.
- RASTELLI, G.; PINZI, L. Refinement and rescoring of virtual screening results. **Frontiers in Chemistry**, v. 7, p. 498, 11 jul. 2019.
- SABE, V. T. *et al.* Current trends in computer aided drug design and a highlight of drugs discovered via computational techniques: A review. **European Journal of Medicinal Chemistry**, v. 224, p. 113705–113705, 15 jul. 2021.
- SALMASO, V.; MORO, S. Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. **Frontiers in Pharmacology**, v. 9, n. AUG, 22 ago. 2018.
- SCHWARTZ, M. A.; GINSBERG, M. H. Networks and crosstalk: integrin signalling spreads. **Nat Cell Biol**, v. 4, n. 4, p. E65-8, 2002.
- SHIMAOKA, M.; TAKAGI, J.; SPRINGER, T. A. Conformational regulation of integrin structure and function. **Annu Rev Biophys Biomol Struct**, v. 31, p. 485–516, 2002.
- SHKURTI, A. *et al.* pyPcazip: A PCA-based toolkit for compression and analysis of molecular simulation data. **SoftwareX**, v. 5, p. 44–50, 1 jan. 2016.
- SHLENS, J. A Tutorial on Principal Component Analysis. 3 abr. 2014.
- SILVA, J. H. M. DA *et al.* Analysis of  $\alpha 4 \beta 1$  integrin specific antagonists binding modes: structural insights by molecular docking, molecular dynamics and linear interaction energy method for free energy calculations. **Journal of the Brazilian Chemical Society**, v. 21, n. 3, p. 546–555, 2010.
- SIMOENS, S.; HUYS, I. R&D Costs of New Medicines: A Landscape Analysis. **Frontiers in**

**Medicine**, v. 8, 26 out. 2021.

SINGH, H. *et al.* Vedolizumab: A novel anti-integrin drug for treatment of inflammatory bowel disease. **J Nat Sci Biol Med**, v. 7, n. 1, p. 4–9, 2016.

SINGH, J. *et al.* Identification of potent and novel  $\alpha 4\beta 1$  antagonists using in silico screening. **Journal of Medicinal Chemistry**, v. 45, n. 14, p. 2988–2993, 4 jul. 2002.

SINGHAL, S. *et al.* Antitumor activity of thalidomide in refractory multiple myeloma. **The New England journal of medicine**, v. 341, n. 21, p. 1565–1571, 18 nov. 1999.

SIRCAR, I. *et al.* Synthesis and SAR of N-benzoyl-L-biphenylalanine derivatives: Discovery of TR-14035, a dual  $\alpha 4\beta 7/\alpha 4\beta 1$  integrin antagonist. **Bioorganic and Medicinal Chemistry**, v. 10, n. 6, p. 2051–2066, 2002.

SLIWOSKI, G. *et al.* Computational Methods in Drug Discovery. **Pharmacological Reviews**, v. 66, n. 1, p. 334–395, 1 jan. 2014.

SLYNKO, I. *et al.* Virtual screening of PRK1 inhibitors: ensemble docking, rescoring using binding free energy calculation and QSAR model development. **Journal of chemical information and modeling**, v. 54, n. 1, p. 138–150, 27 jan. 2014.

SOHRABY, F.; BAGHERI, M.; ARYAPOUR, H. Performing an In Silico Repurposing of Existing Drugs by Combining Virtual Screening and Molecular Dynamics Simulation. **Methods in Molecular Biology**, v. 1903, p. 23–43, 2019.

SONI, A. *et al.* Synthesis and evaluation of 4,5-dihydro-5-methylisoxazolin-5-carboxamide derivatives as VLA-4 antagonists. **Bioorganic & Medicinal Chemistry Letters**, v. 23, n. 5, p. 1482–1485, 1 mar. 2013.

SOUSA DA SILVA, A. W.; VRANKEN, W. F. ACPYPE - AnteChamber PYthon Parser interfacE. **BMC Research Notes**, v. 5, p. 367, 2012.

STERLING, T.; IRWIN, J. J. ZINC 15--Ligand Discovery for Everyone. **J Chem Inf Model**, v. 55, n. 11, p. 2324–2337, 2015.

SUGIURA, T. *et al.* Oral treatment with a novel small molecule alpha 4 integrin antagonist, AJM300, prevents the development of experimental colitis in mice. **Journal of Crohn's & colitis**, v. 7, n. 11, 2013.

SUN, H. *et al.* Assessing the performance of MM/PBSA and MM/GBSA methods. 5. Improved docking performance using high solute dielectric constant MM/GBSA and MM/PBSA rescoring. **Physical Chemistry Chemical Physics**, v. 16, n. 40, p. 22035–22045, 24 set. 2014.

SURABHI, S.; SINGH, B. COMPUTER AIDED DRUG DESIGN: AN OVERVIEW. **Journal of Drug Delivery and Therapeutics**, v. 8, n. 5, p. 504–509, 18 set. 2018.

TAKADA, Y.; YE, X.; SIMON, S. The integrins. **Genome biology**, v. 8, n. 5, p. 215, 2007.

- TAKAGI, J. *et al.* Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling. **Cell**, v. 110, n. 5, p. 511–599, 2002.
- TAKAGI, J.; SPRINGER, T. A. Integrin activation and structural rearrangement. **Immunological Reviews**, v. 186, n. 1, p. 141–163, 2002.
- TONIATI, P. *et al.* Tocilizumab for the treatment of severe COVID-19 pneumonia with hyperinflammatory syndrome and acute respiratory failure: A single center study of 100 patients in Brescia, Italy. **Autoimmunity reviews**, v. 19, n. 7, 1 jul. 2020.
- TRIPATHI, S. K.; SINGH, P.; SING, S. K. Fragment-based de novo design of cyclin-dependent kinase 2 inhibitors. **Methods in Molecular Biology**, v. 1336, p. 47–58, 2016.
- TROTT, O.; OLSON, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **J Comput Chem**, v. 31, n. 2, p. 455–461, 2010.
- URQUHART, L. Market watch: Top drugs and companies by sales in 2017. **Nature reviews. Drug discovery**, v. 17, n. 4, p. 232, 28 mar. 2018.
- VASCONCELOS, D. *et al.* Development of New Potential Inhibitors of  $\beta$ 1 Integrins through In Silico Methods—Screening and Computational Validation. **Life**, v. 12, n. 7, p. 932, 1 jul. 2022.
- VASSETTI, D.; PAGLIAI, M.; PROCACCI, P. Assessment of GAFF2 and OPLS-AA General Force Fields in Combination with the Water Models TIP3P, SPCE, and OPC3 for the Solvation Free Energy of Druglike Organic Molecules. **Journal of Chemical Theory and Computation**, v. 15, n. 3, p. 1983–1995, 12 mar. 2019.
- WANG, C. *et al.* Calculating protein-ligand binding affinities with MMPBSA: Method and error analysis. **Journal of computational chemistry**, v. 37, n. 27, p. 2436–2446, 15 out. 2016.
- WANG, E. *et al.* End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. **Chemical reviews**, v. 119, n. 16, p. 9478–9508, 28 ago. 2019.
- WANG, J. *et al.* Automatic atom type and bond type perception in molecular mechanical calculations. **Journal of molecular graphics & modelling**, v. 25, n. 2, p. 247–260, out. 2006.
- WANG, W.; LUO, B. H. Structural basis of integrin transmembrane activation. **J Cell Biochem**, v. 109, n. 3, p. 447–452, 2010.
- WOODSIDE, D. G.; VANDERSLICE, P. Inflammation and Regulation by Integrin Cell Adhesion Antagonists. **Translational Inflammation**, p. 43–68, 1 jan. 2019.
- XIONG, J. P. *et al.* Crystal structure of the extracellular segment of integrin alpha Vbeta3. **Science (New York, N.Y.)**, v. 294, n. 5541, p. 339–345, 12 out. 2001.
- XIONG, J. P. *et al.* Crystal structure of the extracellular segment of integrin alpha Vbeta3 in complex with an Arg-Gly-Asp ligand. **Science (New York, N.Y.)**, v. 296, n. 5565, p. 151–155,

5 abr. 2002.

YEDNOCK, T. A. *et al.* Prevention of experimental autoimmune encephalomyelitis by antibodies against alpha 4 beta 1 integrin. **Nature**, v. 356, n. 6364, p. 63–66, 1992.

YOU, T. J. *et al.* A 3D structure model of integrin alpha 4 beta 1 complex: I. Construction of a homology model of beta 1 and ligand binding analysis. **Biophysical journal**, v. 82, n. 1 Pt 1, p. 447–457, 2002.

ZHANG, L. *et al.* MicroRNA-3653 inhibits the growth and metastasis of hepatocellular carcinoma by inhibiting ITGB1. **Oncology reports**, v. 41, n. 3, p. 1669–1677, 1 mar. 2019.

ZHANG, S. *et al.* Macrophage-mediated vascular permeability via VLA4/VCAM1 pathway dictates ascites development in ovarian cancer. **The Journal of clinical investigation**, v. 131, n. 3, 1 fev. 2021.

ZHANG, Y. *et al.* Binding blockade between TLN1 and integrin  $\beta$ 1 represses triple-negative breast cancer. **eLife**, v. 11, 1 mar. 2022a.

ZHANG, Y. *et al.* Application of Computational Biology and Artificial Intelligence in Drug Design. **International Journal of Molecular Sciences**, v. 23, n. 21, 1 nov. 2022b.

ZHAO, H.; CAFLISCH, A. Molecular dynamics in drug design. **European Journal of Medicinal Chemistry**, v. 91, p. 4–14, 16 fev. 2015.

ZHAO, Z. *et al.* Drug repurposing to target Ebola virus replication and virulence using structural systems pharmacology. **BMC Bioinformatics**, v. 17, n. 1, p. 1–12, 18 fev. 2016.

ZHENG, Y.; LEFTHERIS, K. Insights into Protein–Ligand Interactions in Integrin Complexes: Advances in Structure Determinations. **Cite This: J. Med. Chem**, v. 63, p. 5696, 2020.

ZHU, S. *et al.* Intelligent Computing: The Latest Advances, Challenges and Future. **Intelligent Computing**, 3 jan. 2023.

ANEXO A – ARTIGO PUBLICADO NO PERIÓDICO *LIFE*




---

Article

## Development of New Potential Inhibitors of $\beta 1$ Integrins through In Silico Methods—Screening and Computational Validation

Diraelli Vasconcelos <sup>1,2</sup>, Beatriz Chaves <sup>1</sup>, Aline Albuquerque <sup>1</sup>, Luca Andrade <sup>1</sup>, Andrielly Henriques <sup>1</sup>, Geraldo Sartori <sup>1,3</sup>, Wilson Savino <sup>2,3,4</sup>, Ernesto Caffarena <sup>4,5</sup> and João Hermínio Martins-Da-Silva <sup>1,4,6\*</sup>

**1** Laboratório de Biologia Estrutural e Funcional em Biotecnologia, Fundação Oswaldo Cruz Casa, RuaSilo-41773-270, Brazil; diraelli.vasconcelos@focruz.br (D.V.); beatriz.chaves@focruz.br (B.C.); aline.albuquerque@focruz.br (A.A.); lucaandrade@focruz.br (L.A.); andriellyhenriques@focruz.br (A.H.); geraldo.sartori@focruz.br (G.S.)

**2** Laboratório de Pesquisa e Desenvolvimento, Instituto Oswaldo-Cruz, Fundação Oswaldo Cruz, Rio de Janeiro 21045-360, Brazil; wilsavino@focruz.br

**3** Instituto Nacional de Câncer e Neuroimunomodulação, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro 21045-360, Brazil

**4** Rede de Pesquisa em Neuroimunomodulação, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro 21045-360, Brazil

**5** Grupo de Biologia Computacional e Modelagem Molecular, Programa de Computação Científica (PRCC), Fundação Oswaldo Cruz, Rio de Janeiro 21040-222, Brazil; ernesto.caffarena@focruz.br

**\*** Correspondence: joao.martins@focruz.br

**Abstract:** Integrins are transmembrane receptors that play a critical role in many biological processes which can be therapeutically modulated using integrin blockers, such as peptidomimetic ligands. This work aimed to develop new potential  $\beta 1$  integrin antagonists using modeled receptors based on the aligned crystallographic structures and docked with three lead compounds (BIO1211, BIO192, and TCS2114), widely known as  $\alpha 4 \beta 1$  antagonists. Lead-compound complex optimization was performed by keeping intact the carboxylate moiety of the ligand, adding substituents in two other regions of the molecule to increase the affinity with the target. Additionally, pharmacokinetic predictions were performed for the ten best ligands generated, with the lowest docking interaction energy obtained for  $\alpha 4 \beta 1$  and BIO192. Results revealed an essential salt bridge between the BIO192 carboxylate group and the  $Mg^{2+}$  MIDAS ion of the integrin. We then generated more than 200 new BIO192 derivatives, some with a greater predicted affinity to  $\alpha 4 \beta 1$ . Furthermore, the significance of retaining the pyrrolidine core of the ligand and increasing the therapeutic potential of the new compounds is emphasized. Finally, one novel molecule (1392) was identified as a potential drug candidate, with appropriate pharmacokinetic profiles, similar dynamic behavior at the integrin interaction site compared with BIO192, and a higher predicted affinity to VLA-4.

**Keywords:** integrin; VLA-4; peptidomimetic inhibitors; molecular docking; de novo design; molecular dynamics; lead optimization; ADMET; MM/PBSA

---

**1. Introduction**

Cell migration is fundamental for many of the lymphohematopoietic system's functions, including physiological leukocyte traffic and the genesis and maintenance of different inflammatory responses. Among the molecular interactions that govern leukocyte migration, there are those mediated by selectins, chemokines, and extracellular matrix (ECM) components. The primary cell surface receptors involved in adhesion to the ECM elements belong to the integrin family.

Integrins of the  $\beta 1$  family are expressed by some blood cell types, including T lymphocytes, and play a critical role in a wide range of biological processes and chronic

Life 2022, 11, 932. <https://doi.org/10.3390/life11207932>

<https://www.mdpi.com/journal/life>

Fonte: (VASCONCELOS et al., 2022).



## ANEXO C – ARTIGO PUBLICADO NO PERIÓDICO *JOURNAL OF CHEMICAL INFORMATION AND MODELING*

**JCIM** JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

### Investigation of Unprecedented Sites and Proposition of New Ligands for Programmed Cell Death Protein 1 through Molecular Dynamics with Probes and Virtual Screening

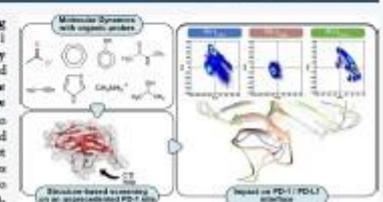
Luca Andrade, Aline Albuquerque, Andrielly Santos-Costa, Disraeli Vasconcelos, Wilson Savino, Geraldo Rodrigues Sartori, and João Herminio Martins Da Silva\*

Cite This: *J. Chem. Inf. Model.* 2022, 62, 1236–1248

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** Cancer immunotherapy has attracted increasing attention over the last few years. Programmed cell death protein 1 (PD-1) promotes self-tolerance and inhibits immune responses by modulating the T-cell function. The interaction between PD-1 and programmed cell death ligand-1 (PD-L1) leads to immune exhaustion, protecting cancer cells from destruction. Here, we computationally designed a novel ligand named 1508 that binds to an unprecedented PD-1 cavity identified by MixMD and defined by amino acid residues Lys78 and Val97. We showed through a set of MD simulations totaling 12.5  $\mu$ s that ligand 1508 establishes frequent cation- $\pi$  and hydrogen bonding interactions with amino acid residues Lys78 and Arg86, respectively, and stabilizes the PD-1 C'D loop in a conformation that does not favor PD-1–PD-L1 complex formation. This study highlights the power of MixMD in exposing new cavities prone to protein–protein complex inhibition and establishes the basis for the design of new molecules that target the PD-1 C'D cavity as an alternative for exploring the modulation of the PD-1–PD-L1 complex in cancer therapy.



#### 1. INTRODUCTION

Cancer is one of the leading public health problems worldwide and is among the top four causes of death in most countries. Global estimates indicate that in 2020 there were 19.3 million cases linked to cancer, with 10.3 million reported deaths. One in 5 people develop cancer during their lifetime, and 1 in 8 men and 1 in 10 women die from the disease.<sup>1</sup> Conceptually, cancer is a group of diseases that originate from genetic modifications associated with dysregulation in cellular proliferation and survival. Genetic and behavioral factors, such as alcohol consumption, exposure to radiation, and obesity, lead to changes in the control of the cell cycle, damaging DNA and altering repair genes, tumor suppressor genes, and oncogenes. These changes result in the generation of processes called hallmarks of cancer, such as cellular hyperproliferation, avoidance of growth-suppressive factors, tumor invasion and metastasis, replicative immortality, induction of angiogenesis, and resistance to apoptosis.<sup>2</sup>

It is well established that the interaction between PD-1, an immune checkpoint receptor present in the T cell, and its ligand, PD-L1, present in cancer cells, results in immune evasion and maintenance of tumorigenesis.<sup>3,4</sup> Specifically, the interaction between PD-1 and PD-L1 promotes cancer cell survival through PD-L1-mediated anti-apoptotic signals, inhibiting the activation of signaling pathways critical for T

cell survival, expansion, and differentiation required for the identification of tumor antigens.<sup>4,7</sup> PD-1 is a transmembrane protein with 288 amino acids composed of an IgV-like N-terminal domain, a transmembrane domain, and a cytoplasmic tail. The cytoplasmic tail containing ITIM and ITSM motifs is responsible for the recruitment of phosphatases and essential for PD-1 inhibition activity.<sup>8</sup> Crystal structures show that PD-1 uses its front  $\beta$ -face (AGFC  $\beta$ -strands) to bind to the  $\beta$ -face of PD-L1 (AGFC'). These predominantly hydrophobic regions demonstrate the importance of this interaction for establishing the functions of these proteins.<sup>9,10</sup>

Current treatments to block the interaction between PD-1 and PD-L1 and restore the immune response involve monoclonal antibodies, such as nivolumab and pembrolizumab, approved for use by the FDA since 2014.<sup>11,12</sup> Despite their relative success, several limitations to using these molecules have been reported over time, such as the high

Received: September 22, 2021  
Published: February 24, 2022

ACS Publications | © 2022 American Chemical Society | 1236

10.1021/acs.jcim.1c01072  
*J. Chem. Inf. Model.* 2022, 62, 1236–1248

Fonte: (ANDRADE et al., 2022).